

湛庐 CHEERS

“高手必备的 经济学思维”

系列套装（ 套装12册 ）



版权信息

本套装纸版浙江教育出版社于2021年8月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有•侵权必究

书名：“高手必备的经济学思维”系列套装（套装12册）

作者：斯科特•佩奇；阿图•葛文德；西蒙•斯涅克；罗伯特•弗兰克；威廉•桑代克；塞德希尔•穆来纳森；威廉•庞德斯通；罗伯特•西奥迪尼；阿莫斯•特沃斯基；丹尼尔•卡尼曼；格雷戈•麦吉沃恩；德内拉•梅多斯

电子书定价：1049.99元

总目录

[模型思维](#)

[清单革命](#)

[无限的游戏](#)

[牛奶可乐经济学4](#)

[巴菲特最推崇的8大企业家特质](#)

[无价](#)

[稀缺](#)

[影响力（全新升级版）](#)

[特沃斯基精要](#)

[噪声](#)

[精要主义](#)

[系统之美：决策者的系统思考](#)

CHEERS
湛庐

模型思维

What You
Need to Know
to Make Data
Work for You

超过100万用户的
模型思维课
首创者佩奇教你
如何聪明决策



圣塔菲 Santa Fe Institute
书系 Series



浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

THE MODEL THINKER

[美] 斯科特·佩奇 著
Scott Page

贾拥民 译

版权信息

本书纸版由浙江人民出版社于2019年12月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

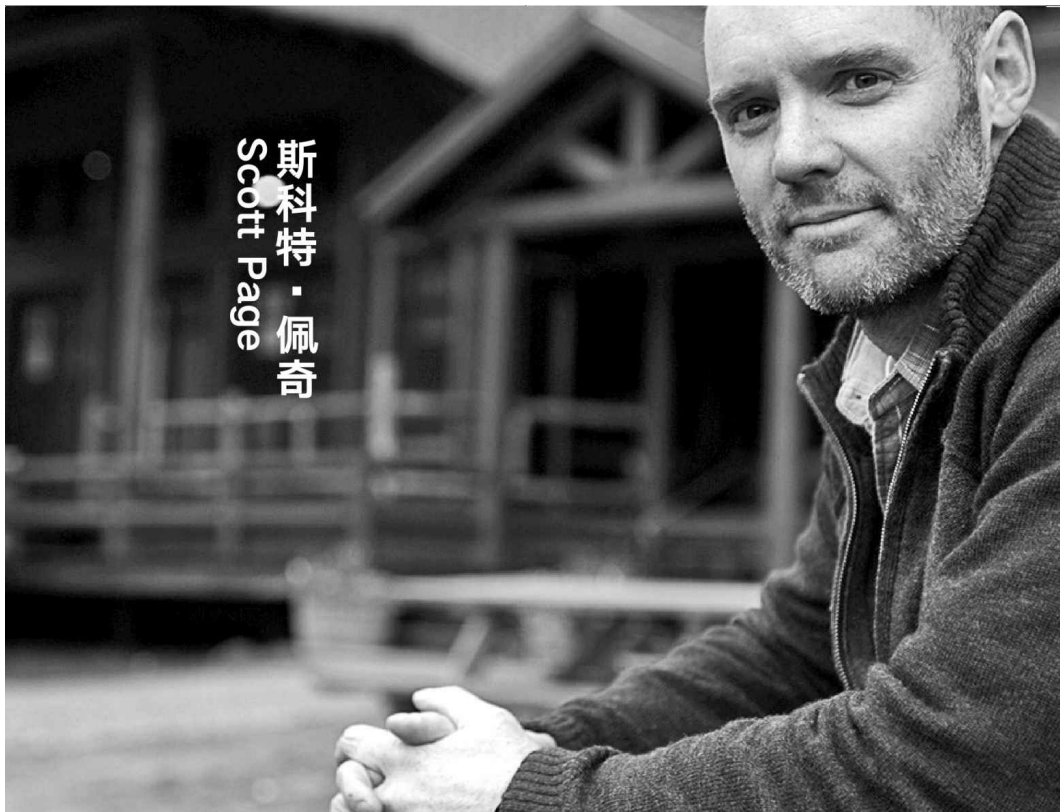
书名：模型思维

著者：斯科特·佩奇

电子书定价：134.99元

THE MODEL THINKER: What You Need to Know to Make Data Work
for You By Scott E. Page.

Copyright © 2018 by Scott E. Page. All rights reserved



广受欢迎的“模型思维课”主讲人

密歇根大学复杂性研究中心“掌门人”

圣塔菲研究所外聘研究员

研究复杂性与多样性的专家



斯科特·佩奇于1985年获得密歇根大学安阿伯分校数学学士学位，1988年获得威斯康星大学麦迪逊分校数学硕士学位，1990年获得西北大学凯洛格商学院管理经济学硕士学位，1993年获得西北大学凯洛格商学院管理经济学和决策科学博士学位。

佩奇以对社会科学的多样性和复杂性的研究和建模而闻名。具体研究方向包括路径依赖、文化、集体智慧、适应和社会生活的计算模型。研究领域涉及多个学科，包括经济学、政治学、计算机科学、管理学、物理学、公共卫生、地理学、城市规划、工程学和历史学。

他曾多次在高中、大学、公司、非营利组织以及政府演讲，介绍他关于多样性和复杂性的研究。也曾经为国际货币基金组织、美国教育部、福特汽车公司、奔驰汽车公司等提供咨询。

佩奇获得了多项奖金，包括2002年的IGERT奖和2001—2006年的生物复杂性项目SLUCE奖，以及2013年的古根海姆奖（Guggenheim Fellowship）。他还曾多次获得加州理工学院、西北大学和密歇根大学颁发的杰出教学奖，这些大学对他多年来在复杂性和多样性方面的教学成果给予了高度认可。佩奇于2011年当选美国艺术与科学学院院士。

广受欢迎的“模型思维课”主讲人

密歇根大学前校长玛丽·苏·科尔曼（Mary Sue Coleman）说：“我们

的教师渴望与全世界分享他们的知识，我们的学生对于可以体验全新的教学方式也同样激动。”斯科特·佩奇对此也积极响应，他在Coursera平台上上线了“模型思维课”，该课程包括一百多个视频和阅读资料。自课程上线以来，有超过5万名学生注册该课程，超过120万人次观看了课程视频，受到了来自世界各地学生的好评。



佩奇教授鼓励对课程好奇的人们注册学习，并亲身参与其中。他说：“这是一个很好的机会，让我们的校友和想要就读密歇根大学的学生们体验一下什么是密歇根大学的教育方式。”

在这门课程里，佩奇讲授了理解和应用模型如何帮助人们做出更好的决策。有证据表明，具备模型思维的人要比没有这种思维的人更优秀，而且能够运用多种模型思考并解决问题的人要比只运用一种模型思考的人更优秀。

佩奇在课程里还着重介绍了几种模型，展示出多样性对创新的重要

性。并具体讲解了拥有多样性视角、启发式的群体是怎么比个人表现更好的。

佩奇的课程引用大量的例子，内容生动有趣。他曾在模型思维课程里讲过一个观点：在离散状态马尔科夫（数学模型）过程中，如果把人生看成努力和努力两个状态，只要状态转移矩阵确定了，长期来看，在每个状态下所停留的时间比例也就都确定了。如果人生的动力源泉是固定的，努力的百分比就是固定的，那么短期内努力或者不努力并不会有什么影响。也就是说，问题的根本不在于你的状态，而在于源动力！所以，在瓶颈期遇到困难实在不想努力的话，多去找一找自己的源动力，想想当初为什么出发。

密歇根大学复杂性研究中心“掌门人”圣塔菲研究所外聘研究员



斯科特·佩奇于2002—2009年担任密歇根大学复杂性研究中心副主任，于2010—2015年担任主任。复杂性研究中心成立于1999年，其创始成员是一个现已成为传奇的研究小组——巴赫小组（BACH Group）。巴赫小组始于20世纪80年代，成员包括来自不同学科的研究人员，他们都对各种复杂的适应性系统感兴趣。



巴赫小组最初的成员包括美国数学家亚瑟·伯克斯（Arthur Burks）、遗传算法之父约翰·霍兰德（John Holland）等人。美国进化生物学家威廉·汉密尔顿（William Hamilton）、认知科学家侯世达、认知科学家梅勒妮·米歇尔（Melanie Mitchell）也是巴赫小组的成员。现在，巴赫小组由政治学家罗伯特·阿克塞尔罗德（Bob Axelrod）、物理学家马克·纽曼（Mark Newman）、数学家和公共政策学家卡尔·西蒙（Carl Simon）和斯科特·佩奇等人组成。其中，马克·纽曼、卡尔·西蒙也是复杂性研究中心的成员。

佩奇于1999年被圣塔菲研究所聘为研究员，开始了在圣塔菲研究所十几年的研究生涯，主要研究方向依然是复杂性和多样性。

作者演讲洽谈，请联系

speech@cheerspublishing.com

更多相关资讯，请关注



湛庐文化微信订阅号

湛庐CHEERS

特别
制作

献给
迈克尔·科恩

不应否认，任何理论的终极目标都是尽可能让不可简化的基本元素变得更加简单且更少，但也不能放弃对任何一个单一经验数据的充分阐释。

阿尔伯特·爱因斯坦



扫码下载“湛庐阅读”APP，
搜索“模型思维”，
获取更多精彩内容。

序言 这本书是怎样写成的

对我来说，成功意味着我在这个世界上的有效性：我能够把我的思想和价值观带给这个世界，我能够以积极的方式改变它。

汤亭亭

这本书源于我与美国复杂系统专家迈克尔·科恩（Michael Cohen）一次偶然的见面。那是在2005年，密歇根大学西厅旁边购物中心的花园。作为一名学者，迈克尔素以慷慨大方而闻名，他在这次见面时说的一番话，彻底改变了我的教学生涯。迈克尔眼中闪烁着光芒，他说：“斯科特，我曾经根据查尔斯·拉夫（Charles Lave）和詹姆斯·马奇（James March）写的一本书，开设过一门名为‘社会科学建模导论’的课程。你应该重开这门课程，它需要你。”

它需要我？我有点困惑。回到办公室后，我马上找出了那门课的课程大纲。看了之后，我发现迈克尔错了，不是这门课程需要我，而是我需要它。我一直在寻找这样一门课程：既能够向学生介绍复杂系统领域的知识，同时又不脱离他们的日常生活和未来的职业规划。通过讲授一门关于模型的课程，我可以向学生展示各种相关的工具和思想，提高他们推理、解释、设计、沟通、行动、预测和探索的能力。

开设这样一门课程的基本动机是，我们必须利用多种多样的模型去应对复杂性。学了一个学期之后，学生们在看待这个世界时，就不会拘泥于某个特定的角度，相反，他们将透过多种不同的视角去观察世界。他们将站在有许多扇窗户的房子中，拥有看向多个方向的能力。我的学生应该能够更好地应对他们所面对的复杂挑战：改善教育，减少贫困，实现可持续增长，在人工智能时代找到有意义的工作，管理好资源，设计出强大而稳健的金融、经济和政治体系。

第二年秋天，我真的重新开设了这门课程。我本来打算将这门课程命名为“32种使你变成天才的模型”，但是密歇根大学的传统不允许这个有“王婆卖瓜”之嫌的名字。于是，我沿用了迈克尔的课程名称“建模导论”。事实早已证明，查尔斯·拉夫和詹姆斯·马奇的书是非常不错的入门读物。不过，在过去的几十年中，建模技术已经取得了巨大的进步。我需要对这门课程进行更新，以便将长尾分布（long-tailed distributions）、网络模型、崎岖景观模型（rugged-landscape model）和随机游走模型（random walk model）等全都包括进来。我需要一本讨论复杂性的书。

于是，我开始了写作，但这并非易事。刚开始的两年，写作进展得非常缓慢。后来，在一个春天，我又一次遇到了迈克尔，这次是在西厅的拱门那里。我自己一直对这门课程有所疑问，尽管那时课程已经吸引20多名学生了。对于本科生来说，这些模型是否过于抽象了？我是不是应该针对不同的问题或政策领域开设不同的课程呢？面对这些问题，迈克尔笑了笑，他指出任何值得做的事情都必定会招致质疑。告别时，迈克尔再一次强调，帮助人们清晰地地进行思考是非常重要且非常有价值的。他叮嘱我不要放弃。

2012年秋季学期，这门课程又发生了根本性变化。密歇根大学副教务长玛莎·波拉克（Martha Pollack）邀请我加入在线开放课程体系，开设一门在线课程，也就是现在所说的慕课（MOOC）。于是，借助于一台联网的电脑，一个29美元的摄像头和一个90美元的麦克风，“模型思维”（Model Thinking）这门在线课程正式诞生了。

在来自密歇根大学、斯坦福大学以及Coursera的无数师友的帮助下，我将上课用的讲义重新改编成了适合在线课程的形式，包括将每个主题划分为若干模块，并删除了所有受版权保护的资料。在这里，我要对这些热心人士表示感谢，尤其是汤姆·希基（Tom Hickey），他总是随叫随到，帮我解决了很多疑难问题。为了保证效果，我一遍又一遍地

重新录制课程。感谢我的狗Bouncer，它是我忠实的听众。

“模型思维”第一次上线时就吸引了6万名学生。到现在，这门课程的注册学生人数已经接近100万了。在线课程如此受欢迎，以至于我暂时放弃了写作本书的计划。我当时认为也许没有必要写一本书。但是很快，我的电子邮件收件箱就被学生们要求得到一本可以作为在线课程补充书籍的请求塞满了。在迈克尔·科恩不幸因癌症去世后，我更加觉得自己必须尽快完成这本书，于是我重新摊开了书稿。第1稿，第2稿.....第9稿.....现在，你们看到的是第30稿。

写一本书需要大量的时间和开阔的空间。美国著名现代诗人华莱士·史蒂文斯（Wallace Stevens）曾经这样写道：“也许，真理取决于在湖边散步的时间。”就我自己写作这本书的经历来说，思考大大得益于数百次横穿怀南斯湖（Winans Lake）的游泳，每到夏天，我和我的家人总是在那里过周末。

在整个写作过程中，我和我生命中的挚爱珍娜·贝德纳（Jenna Bednar）、我们的儿子奥里（Orrie）和库珀（Cooper），以及我们养的3只爱犬Bouncer、Oda和Hildy一起开怀大笑、享受生命、面对挑战、把握机遇。在那里，奥里订正了本书倒数第二稿中的数学错误，珍娜花了两个星期阅读了全部书稿并修正了数百个有问题的地方：字体和语法上有错误的、思路不清晰的、逻辑有缺陷的、举例混乱的.....事实上，与我撰写的所有文字作品一样，对这本书的一个更加准确的描述是：斯科特·佩奇写出了第一稿，最后定稿则由珍娜·贝德纳完成。

在写作本书的过程中，我看着我的孩子们从青春期过渡到了成人阶段。我们一家人也吃掉了太多的石锅拌饭、意大利面和燕麦片巧克力饼干。我们用锯子和修枝剪，修剪了无数棵大大小小的树木，补好了篱笆上的几十个破洞；花费了很多时间和精力，试图让地下室和车库变得整齐一些，但最终却总是无功而返。我们也一年又一年地盼望湖上的冰厚到可以滑冰.....

从草拟大纲到提交定稿，用了整整7年的时间，怀南斯湖冬天结冰、春天又化开，前前后后也有7次了。然而，有的冰冻期比其他冰冻期更加难挨。在写作本书的过程中，我的母亲玛丽莲·坦博·佩奇（**Marilyn Tamboer Page**）不幸因心脏病突发去世，当时她正在散步，那是她每天的例行事务之一。有些深洞永远无法填满，它们会提醒我们珍惜生命所提供的宝贵机会。

现在，这本书终于已经完成了。奥里已经上了大学，库珀明年也要去上大学了。如果您，我亲爱的读者，认为本书中的模型和思想是有用的、有创造性的，而且能够将它们应用于现实世界，并以积极的方式去改变世界，那么我将这些资料组织成书的努力就得到了最大的回报。如果有一天，当我在某个教授或研究生的办公室里（最有可能是在美国中西部的某所大学里）浏览书架时，在拉夫和马奇的书旁边发现了我的这本书，那么我将会觉得非常幸福。

目 录

[序言 这本书是怎样写成的](#)

[Part 1 为什么需要模型思维](#)

[01 做一个多模型思考者](#)

[02 模型的7大用途](#)

[03 多模型思维](#)

[04 对人类行为者建模](#)

[Part 2 模型思维](#)

[05 正态分布](#)

[06 幂律分布](#)

[07 线性模型](#)

[08 非线性模型](#)

[09 与价值和权力有关的模型](#)

[10 网络模型](#)

[11 广播模型、扩散模型和传染模型](#)

[12 熵：对不确定性建模](#)

[13 随机游走](#)

[14 路径依赖模型](#)

[15 局部互动模型](#)

[16 李雅普诺夫函数与均衡](#)

[17 马尔可夫模型](#)

[18 系统动力学模型](#)

[19 基于阈值的模型](#)

[20 空间竞争模型与享受竞争模型](#)

[21 博弈论模型](#)

[22 合作模型](#)

[23 与集体行动有关的问题](#)

[24 与机制设计有关的模型](#)

[25 信号模型](#)

[26 学习模型](#)

[27 多臂老虎机问题](#)

[28 崎岖景观模型](#)

[结语 像芒格一样智慧地思考——多模型思维的实际应用](#)

[译者后记](#)

[注释](#)

[参考文献](#)

The Model Thinker

What You
Need to Know to
Make Data Work
for You

Part 1 为什么需要模型思维

多模型思维要求掌握多个模型，但是我们并不需要懂得非常多的模型，只要知道每个模型都可以有多种应用场景。要成为一个多模型思考者，需要的不仅仅是数学能力，更需要的是创造力。

01 做一个多模型思考者

要想成为一个有智慧的人，你必须拥有多个模型。而且，你必须将你的经验，无论是间接的，还是直接的，都放到构成这些模型的网格上。

查理·芒格（Charlie Munger）

这是一本关于模型的书。我在书中用简洁的语言描述了几十个模型，并解释该如何应用它们。模型是用数学公式和图表展现的形式化结构，它能够帮助我们理解世界。掌握各种模型，可以提高你的推理、解释、设计、沟通、行动、预测和探索的能力。

本书提倡多模型思维方法，应用模型集合理解复杂现象。本书的核心思想是：多模型思维能够通过一系列不同的逻辑框架“生成”智慧。不同的模型可以将不同的力量分别突显出来，它们提供的见解和含义相互重叠并交织在一起。利用多模型框架，我们就能实现对世界丰富且细致入微的理解。本书还包括了一些正式的论证，阐述了如何对现实世界应用多模型框架。

本书非常实用。多模型思维具有十分重要的实用价值。运用这种思维方式，你就能更好地理解复杂现象，就能更好地推理。你将会在职业生涯、社区活动和个人生活中表现出更小的差距，做出更加合理的决策。是的，你甚至还可能会变得更有智慧。

25年前，像本书这样讲解模型的著作主要是供教授们和研究生们研究商业、政策和社会科学所用的，金融分析师、精算师和情报界人士也是潜在的读者。这些人都是应用模型的人，他们也是与大型数据库关系最密切的人，这并不是偶然。不过到了今天，关于模型的书已经拥有了更多的读者：广大的知识工作者们。由于大数据的兴起，他们现在已经

把模型作为日常生活的一部分了。

如今，用模型组织和解释数据的能力，已经成了商业策略家、城市规划师、经济学家、医疗专家、工程师、精算师和环境科学家等专业人士的“核心竞争力”。任何人，只要想分析数据、制订业务发展策略、分配资源、设计产品、起草协议就必须应用模型，哪怕是做出一个简单招聘决策，也要运用模型思维。因此，掌握本书的内容，特别是那些涉及创新、预测、数据处理、学习和市场准入时间选择的模型，对许多人都有非常重要的实际价值。

使用模型来思考能够带给你的，远远不仅仅是工作绩效的提高。它还会使你成为一个更优秀的人，让你拥有更强的思考能力。你将更擅长评估层出不穷的经济事件和政治事件，更能识别出自己与他人推理中的逻辑错误。有了这种思维方式，你将懂得辨识什么时候意识形态取代了理性思考，并对各种各样的政策建议有更丰富、更有层次的洞见，无论是扩建城市绿地的建议，还是强制药物检测的规定。

所有这些好处都来自与多种多样模型的“亲密接触”，幸运的是，我们用不着一下子掌握千百种模型，而只需先掌握几十种就足够了。本书给出的这些模型就为你提供了一个很好的出发点。它们来自多门学科，其中包括许多人耳熟能详的囚徒困境博弈模型，逐底竞争（**Race to the Bottom**）和关于传染病传播的**SIR**模型，等等。所有这些模型都有一个共同的形式：它们都假设一些实体，通常是人或组织，并描述他（它）们是如何相互作用的。

本书所讨论的模型可以分为三类：对世界进行简化的模型、用数学概率来类比的模型以及人工构造的探索性模型。无论哪一种形式，模型都必须是易处理的。模型必须足够简单，以便让我们可以在模型中应用逻辑推理。例如，我们讨论了一种传染病模型，这个模型由易感者、感染者和痊愈者组成，可以给出传染病的发生概率。利用这个模型，我们可以推导出一个传染阈值，也就是一个临界点，超过这个临界点，传染

病就会传播。我们还可以确定，为了阻止传染病传播，需要接种疫苗人数的比例。

尽管单个模型本身可能就已经相当强大了，但是一组模型可以实现更多的功能。在拥有多个模型的情况下，我们能够避免每个模型本身所固有的局限性。多模型方法能够消除每个单个模型的盲点。基于单一模型的政治选择可能忽略了世界的一些重要特征，如收入差距、身份多样性以及与其他系统的相互依赖关系。¹有了多个模型，我们可以达成对多个流程的逻辑推理，可以观察不同因果过程是如何重叠和相互作用的，也拥有了理解经济、政治和社会世界复杂性的可能。而且，我们在这样做的时候并不需要放弃严谨性，因为模型思维能够确保逻辑的一致性。由此，推理将建立在扎实的证据基础之上，因为模型需要用数据检验、改进和精炼。总而言之，当我们的思维得以在多个逻辑上一致、处在通过了经验验证的框架中时，我们更有可能做出明智的选择。

大数据时代的模型

在当今这个大数据时代，像本书这样一本讨论模型的书可能看上去有些不合时宜。现在，数据正以前所未有的维度和粒度急速地涌现出来。过去，消费者的购买数据只能以每月汇总表的形式打印出来，而现在却可以与空间、时间信息及消费者“标签”一起实时传输。学生的学习成绩数据，现在也包括每一份作业、每一篇论文、每一次测验和考试的分数，而不再仅仅是一个期末总成绩了。过去，农场工人也许只能在每月一次的农场会议上提出土壤过于干燥的问题，而现在，他们却能够用拖拉机自动传输以平方米为单位的关于土壤肥力和水分含量的实时数据了。投资公司要跟踪数千只股票的数十种比率和趋势，并使用自然语言处理工具来解析文档。医生则可以随时提取包括相关遗传标记在内的患者记录。

仅仅在25年以前，大多数人获得的知识只能来自书架上的几本书。

也许你工作的地方有一个小型图书馆，或者你家里有全系列的百科全书和几十本参考书。学术界、政府和私营部门的研究者则可以利用大型图书馆的馆藏资料，但是他们也经常不得不亲身前往查阅。就在20世纪末21世纪初，为了获得必要的信息，学者们仍然不得不在卡片目录室、缩微胶片阅览室、图书馆书架以及私人收藏家的“宝库”之间来往穿梭。

现在，这一切都发生了颠覆性的变化。几个世纪以来一直受到纸张束缚的知识内容，今天已经以数据包的形式在“空中”自由流动了。关于此时此地的实时信息也是如此。以前，新闻是刊载在报纸上的，最高以每天一次的频率送到我们手上；而现在，新闻却是以连续的数字流形式流入我们的个人设备。股票价格、体育赛事比分、关于政治经济事件和文化事件的新闻，全都可以实时查询、实时访问。

然而，无论数据给我们留下的印象如何深刻，它都不是灵丹妙药。我们也许可以通过数据了解到已经发生了什么和正在发生什么，但是，由于现代世界是高度复杂的，我们可能很难能理解为什么会发生这种情况。更何况，经验事实本身也可能是误导性的。例如，关于计件工资制的统计数据往往会显示，工人每生产一件产品获得的报酬越高，他们生产的产品就会越少。对此，用一个薪酬取决于工作条件的模型可以很好地解释相关数据。如果工作条件很差，导致很难生产出产品，那么每单位产品的工资可能很高；如果工作条件很好，那么每单位产品的工资就可能会很低。因此，并不是更高的计件工资导致了更低的生产率，而是更加糟糕的工作条件导致了这种结果。²

此外，我们社会中的大多数数据，也就是关于经济、社会和政治现象的数据，都只是时间长河上的瞬间或片断的记录。这种数据是不能告诉我们普遍真理的。我们的经济、社会和政治世界并不是固定不变的。在这个十年内，男孩在标准化考试中的成绩超过了女孩，但是下个十年就有可能变为女孩的成绩好于男孩。人们今天投票的原因，可能与未来几十年投票的原因截然不同。

我们需要模型，不然就无法理解计算机屏幕上不断滑过的数据流。因此，这个时代，可能恰恰因为我们拥有如此多的数据，也可以被称为多模型时代。纵观学术界、政府、商界和非营利部门，你基本上无法找到任何一个不受模型影响的研究领域，甚至可以说根据不存在不需要模型的决策领域。麦肯锡（McKinsey）和德勤（Deloitte）等咨询业巨头要通过构建模型来制订商业策略；贝莱德集团（BlackRock）和摩根大通集团（JPMorgan Chase）等金融业大公司要利用模型来选择投资，州立农业保险公司（State Farm）和美国好事达保险公司（Allstate）等公司的精算师要借助风险校正模型来给保险单定价。谷歌公司的人力资源部门要利用预测分析模型来为超过300万求职者进行评估。各大学和学院的招生人员也要建立模型，以便从成千上万的申请入学者当中选出合格的新生。

美国行政管理和预算局（Office of Management and Budget）通过构建经济模型预测税收政策的影响。华纳兄弟公司通过数据分析模型评估观众对电影的反应。亚马逊公司开发机器学习模型向消费者推荐商品。由美国国家卫生研究院（National Institutes of Health）资助的研究团队建立了人类基因组学的数学模型，用于寻找和评估癌症潜在的治疗方法。盖茨基金会使用流行病学模型设计疫苗接种策略。甚至运动队也都使用模型来预测选秀结果和交易机会，并制订比赛策略。例如，芝加哥小熊队（Chicago Cubs）之所以能够在经历了一个多世纪的失败后赢得世界职业棒球联赛的冠军，就是因为很好地利用了模型去选择球员、设计比赛策略。

对于使用模型的人来说，模型思维的兴起还有一个更简单的解释：模型能够让我们变得更聪明。如果没有模型，人们就会受到各种认知偏差的影响：我们会对近期发生的事件赋予过高的权重、会根据“合理程度”分配概率、会忽略各种基本比率。如果没有模型，我们处理数据的能力就会受到极大的限制。有了模型，我们就能澄清相关假设且更有逻辑地进行思考，还可以利用大数据来拟合、校准、检验因果关系与相关

性。总之，有了模型，我们的思考会更有效。有证明表明，如果让模型与人面对面直接“竞争”，模型将会胜出。³

为什么需要多模型

在本书中，我们主张在给定情况下不仅使用一个模型，而要使用多个模型。多模型方法背后的原理基于这样一个古老的思想，那就是“管中窥豹需多管齐下”。这个思想至少可以追溯至亚里士多德，他强调将许多人的优点集中起来这个做法的价值。呈现视角和观点的多样性，也是美国历史上“名著运动”（great-books movement）背后的一大动力。在这个运动中涌现出来的《伟大的思想：西方世界名著中伟大的思想观念合集》（*The Great Ideas: A Syntopicon of Great Books of the Western World*）一书，就收集了102个重要的可永世流传的思想。

现在，这种方法也在汤亭亭所著的《女勇士》（*The Women Warrior*）一书中得到了回响，她这样写道：“我已经学会了如何让我的思想变得博大；因为宇宙很大，所以给悖论留下了存在的余地。”这种方法也构成了现实的商业和政治世界有实际意义的行动基础。最近的一些论著指出，如果我们想要理解国际关系，就不能只将世界建模为一组具有明确目标的自利国家，也不能只将世界建模为跨国公司和政府间组织之间的联系枢纽，而应该把世界同时建模为这两者。⁴

尽管多模型方法看上去似乎很平常，但请注意，它其实是与我们讲授模型和构建模型的传统方法相悖的。传统的方法，那些在高中时老师教授的方法，依赖一对一的逻辑，也就是说一个问题需要一个模型。比如，老师会告诉我们，在这种情况下，我们应该运用牛顿第一定律；在那种情况下，我们应该运用牛顿第二定律；在第三种情况下，则应该运用牛顿第三定律。又或者，在这里，我们应该使用复制因子方程

（replicator equation）来说明下一期兔子种群的大小。在这种传统的方法中，目标是确定一个适当的模型并正确应用这个模型。而多模型思维

所要挑战的，恰恰正是这种传统方法。多模型方法主张尝试多个模型。如果你在九年级时就使用过多模型思维，你可能会被阻止，但是现在使用多模型思维，你将会取得很大进步。

大部分学术论文也遵循传统的一对一的方法，尽管有时它们是在使用单一的模型去解释复杂的现象。例如，有人声称，在美国2016年选举中投票给特朗普的那些人，都是经济上的失败者。又或者，小学二年级时老师的素质决定了孩子长大成人后能够取得经济成就的大小。⁵不过，近年来，一系列畅销的非虚构作品的诊断，使这种基于单个模型的传统思维方式的弊端呈现在人们面前：教育成功只取决于毅力；资本集中导致不平等；糖消耗导致民众健康状况不佳……这些单个模型中的每一个都可能是正确的，但没有一个是全面的。面对各种复杂的挑战，创造一个包容更广泛教育成就的世界，我们需要的不是单个模型，而是多个模型构成的格栅。

通过学习本书中的模型，你就可以着手构建自己的格栅模型。这些模型来自多个学科，涉及各种现象，例如收入不平等的原因、权力的分配、传染病和流行风尚的传播、社会动乱的前置条件、合作的发展、秩序的涌现，以及城市和互联网的结构等。

模型的假设和结构各不相同。有些模型描述了少量理性的、自私的行为主体之间的互动，有些模型则描述了大量的遵循规则的利他主义者的行为。一些模型描述了均衡过程，还有一些模型讨论路径依赖性和复杂性。这些模型的用途也各不相同。一些模型是用来帮助预测和解释的，一些模型是用来指导行动、推动设计或促进沟通的，还有一些模型则创造了有待我们去探索的虚拟世界。

所有模型都有三个共同特征。第一，它们都要简化，剥离不必要的细节，抽象掉若干现实世界中的因素，或者需要从头重新创造。第二，它们都是形式化的，要给出精确的定义。模型通常要使用数学公式，而不是文字。模型可以将信念表示为世界状态的概率分布，可以将偏好表

示为各备选项之间的排序。通过简化和精确化，模型可以创造易于处理的空间，我们可以在这些空间上进行逻辑推理、提出假说、设计解决方案和拟合数据。模型创建了我们能够以符合逻辑的方式进行思考的结构。正如维特根斯坦在《逻辑哲学论》（*Tractatus Logico-Philosophicus*）一书中所写的：“逻辑本身就能解决问题，我们所要做的，就是观察它是如何做到的。”是的，逻辑有助于解释、预测、沟通 and 设计。但是，逻辑也不是没有代价的，这就导致模型的第三个共同特征是：所有模型都是错误的，正如统计学大师乔治·博克斯（George Box）所指出的那样。⁶所有模型概莫能外，即使是牛顿提出的那些定律和法则，也只是在特定的条件下成立。所有模型都是错误的，还因为它们都是简化的，它们省略掉了细节。通过同时考虑多个模型，我们可以实现多个可能情况的交叉，从而克服单个模型因严格而导致的狭隘性。

只依靠单个模型其实是过于狂妄自大的表现，这种做法会导致灾难性的后果。相信只凭一个方程，就可以解释或预测复杂的现实世界现象，会使真理成为那种很有“魅力”的简洁的数学公式的牺牲品。事实上，我们永远不应指望任何一个模型能够准确预测1万年后的海平面将上升多少，甚至也不应该指望任何一个模型能够准确预测10个月后的失业率。我们需要同时利用多个模型才能理解复杂系统。政治、经济、国际关系或者大脑等复杂系统永远都在变化，时刻都会涌现出介于有序和随机之间的结构和模式。当然，根据定义，复杂现象肯定是很难解释或预测的。⁷

因此在这里，我们面临着一个严重的脱节。一方面，我们需要模型来连贯地思考。另一方面，任何只具有少数几个活动部件的单个模型都无法解释高维度的复杂现象，例如国际贸易政策中的模式、快速消费品行业的发展趋势或大脑内部的适应性反应。即便是牛顿，也无法写出一个能够解释就业水平、选举结果或犯罪率下降趋势的三变量方程。如果我们希望了解传染病的传播机制、教育成效的变化、动植物种类的多样性、人工智能对就业市场的冲击、人类活动对地球气候的影响，或者社

会动乱的可能性，就必须通过多个模型去了解它们：机器学习模型、系统动力学模型、博弈论模型和基于主体的模型等。

智慧层次结构

为了论证多模型思维方式的优点，我们先从诗人和剧作家T. S.艾略特的一个疑问入手：“我们迷失于知识中的智慧到哪里去了？我们迷失于信息中的知识到哪里去了？”在这里，我们还可以加上一句：我们迷失于数据中的信息到哪里去了？

我们可以把艾略特的这个疑问形式化为一个智慧层次结构（wisdom hierarchy），如图1-1所示。在这个智慧层次结构的最底部是数据，也就是原始的、未编码的事件、经历和现象。出生、死亡、市场交易、投票、音乐下载、降水、足球比赛，以及各种各样的（物种）发生事件等。数据既可以是一长串的0和1，也可以是时间戳，或是页面之间的链接等。数据是缺乏意义、组织或结构的。

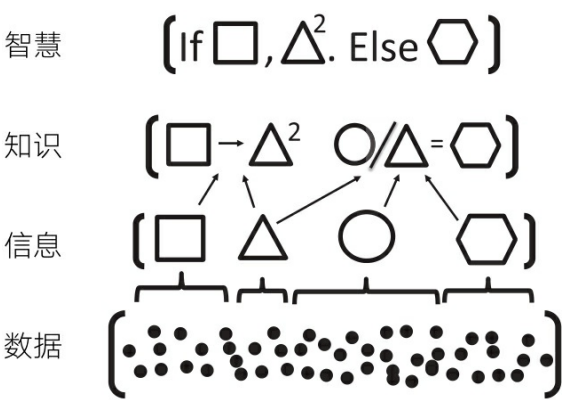


图1-1 智慧层次结构

信息用来给数据命名并将数据归入相应的类别。为了说明数据与信息之间的区别，看看这几个例子：落在你头上的雨是数据，佛蒙特州伯灵顿市和安大略湖的7月份总降水量则是信息；威斯康星州麦迪逊市国会大厦旁边周六市场上的鲜红辣椒和金黄玉米是数据，而农民的总销售

额则是信息。

我们生活在一个信息极大丰富的时代。一个半世纪以前，掌握信息可以带来很高的经济和社会地位。英国小说家简·奥斯汀（Jane Austen）笔下的爱玛就曾问过，弗兰克·丘吉尔（Frank Churchill）是不是“一个拥有着很多信息的年轻人”。^①如果放到今天，她肯定不会在意这个问题。如果穿越到现在，那么弗兰克·丘吉尔会和其他人一样有一部智能手机，问题只在于他有没有能力很好地利用这些信息。正如陀思妥耶夫斯基在《罪与罚》一书中所写的那样：“他们说，我们已经得知了事实。但事实不是一切，至少有一半的分歧就出在人们怎样利用事实上！”

柏拉图将知识定义为合理的真实信念。更现代的定义则认为知识就是对相关关系、因果关系和逻辑关系的理解。知识组织了信息，呈现为模型的形式。市场竞争的经济学模型、网络的社会学模型、地震的地质学模型、生态位形成的生态学模型以及学习的心理学模型都体现了知识。这些模型能够解释和预测。化学键模型解释了为什么金属键会使我们无法将手伸进钢制的门，为什么当我们潜入湖水中时氢键会影响我们的体重。⁸

层次结构的基础就是智慧。智慧就是指识别和应用相关知识的能力。智慧需要多模型思维。有时，智慧体现在懂得如何选出最优模型，就好像将箭从箭袋中抽出来一样。还有时，智慧可以通过求出各种模型的平均结果来实现，这是在进行预测时的一种常见做法。采取行动时，有智慧的人都会应用多个模型，就像医生会让病人做好几种检查来帮助诊断一样。他们使用模型来排除某些行为、选择某些行为。有智慧的个人和团队会有意让模型之间相互“对话”，探索不同模型之间的重叠和差异。

智慧包括选择正确的知识或模型。考虑一下这个物理问题：一个小小的毛绒玩具猎豹从一架飞在6千米高的飞机上掉下来，当它着地时会

造成多大的伤害？学生可能已经掌握了引力模型和自由降落速度模型。这两个模型会给出不同的答案。引力模型的预测是，这个玩具猎豹会撕裂汽车的顶棚。自由降落速度模型的预测则是这个玩具猎豹的最高速度可以达到每小时16千米。⁹在这个问题上，智慧意味着，知道应该如何运用自由降落速度模型。事实上，站在地上一个人，完全可以将这只柔软的毛绒玩具抓在手中。在此，不妨引用进化生物学家J. B. S.霍尔丹（J. B. S. Haldane）的一段话来说明这个问题：“你可以将一只小鼠丢到一口深达千米的矿井，当它坠落到井底时，只要地面是相当柔软的，那么小鼠只会受到轻微的震荡，而且能够自行走开。但如果是大鼠的话就会摔死，人则会粉身碎骨，马更将尸骨无存。”

回到上面这个毛绒玩具的问题上来，要想得到正确的答案需要信息（这个玩具的重量）、知识（自由降落速度模型）和智慧（选择正确的模型）。商界和政界领袖也依靠信息和知识做出明智的选择。例如，2008年10月9日，冰岛的货币冰岛克朗（króna）开始自由落体般的急剧贬值。当时的软件巨头甲骨文公司（Oracle）的财务主管埃里克·鲍尔（Eric Ball）必须做出一个决定。就在几个星期之前，他刚刚处理了国内住房抵押贷款危机带来的冲击。冰岛的情况引发了国际关注，而甲骨文公司持有数十亿美元的海外资产。鲍尔先考虑了关于金融崩溃的网络传染模型，然后他又考虑了讨论供给和需求的经济学模型（在这种模型中，价格变化的幅度与市场冲击的大小相关）。2008年，冰岛的国内生产总值仅为120亿美元，只相当于麦当劳公司6个月的销售收入。事后，鲍尔回忆当时的思考过程：“冰岛的经济规模比美国弗雷斯诺市还要小呢。回去工作吧，不用多管。”¹⁰

要理解这个例子，或者理解多模型思维方法，关键是要认识到鲍尔并没有去探索过多的模型，他找到了一个模型来支持已经决定采取的行动。是的，他没有尝试很多模型后找到一个能证明自己行为合理性的模型。相反，他只评估了两个可能有用的模型，然后选择了一个更好的模型。鲍尔拥有正确的信息（冰岛很小），选择了正确的模型（供需模

型），并做出了一个明智的选择。

接下来，我们重新反思两个历史事件来说明如何让多个模型展开“对话”。这两个历史事件是：2008年的全球金融市场崩溃，它使总财富（或者说至少是人们所认定的总财富）减少了数万亿美元，进而导致了长达4年之久的全球经济衰退；以及1961年的古巴导弹危机，它几乎引发了一场核战争。

对于2008年全球金融市场崩溃的原因，已经出现了多种解释：外国投资过多；投资银行过度杠杆化；抵押贷款审批过程缺乏监督；家庭消费者过分乐观的情绪；金融工具的复杂性；对风险的误解，以及贪婪的银行家明知泡沫存在却铤而走险并期望获得救助；等等。表面证据似乎与这些解释保持了一致：从外国流入了大量资金；贷款发起人发放了“有毒”（低质量）的抵押贷款；投资银行的杠杆率确实非常高；金融工具太过复杂导致大多数人无法理解；不少银行预计政府会出台救助计划；等等。通过模型，我们可以在这些解释之间加以“裁决”，可以分析其内在一致性：它们是否符合逻辑？我们还可以用数据进行校准、对推断进行检验。

经济学家罗闻全运用多模型思维方法，对关于这场危机的20种不同解释进行了评估。他发现，每一种解释都有不足之处。而且，没有理由认为投资者在明知自己的行为会导致全球危机时还会为泡沫作贡献。因此，泡沫的严重程度一定是出乎许多人的意料的。金融公司可能假定其他公司已经做好了尽职调查，而事实上并没有。回想起来，明显“有毒”的抵押贷款组合也找到了买家。如果全球金融市场崩溃成为定局，那么买家就不会存在。虽然杠杆率自2002年以来一直在上升，但却并没有比1998年的时候高出很多。而对于政府必定会救助银行的观点，雷曼兄弟银行的遭遇说明了一切：雷曼兄弟银行于2008年9月15日倒闭，它的资产超过6 000亿美元，这是美国历史上最大的破产案，然而政府并没有介入。

罗闻全认为，每种解释都包含了一个逻辑上的缺憾。从数据本身来说，没有任何一个解释是特别有根据的。正如罗闻全所总结的：“我们应该从一开始就努力对同一组客观事实给出尽可能多的解释，并寄希望于时间。当时机成熟的时候，关于这场危机更细致和更一致的解釋就会浮现出来。”他还说：“唯有通过收集多样化且往往相互矛盾的解釋，我们才能最终实现对危机更完整的理解。”任何单个的模型都是不足的。¹¹

在《决策的本质》（*Essence of Decision*）一书中，美国政治学家格雷厄姆·艾利森（Graham Allison）采用多模型思维方法解释了古巴导弹危机。1961年4月17日，一支由美国中央情报局训练出来的半正规武装队伍在古巴海岸登陆，企图推翻菲德尔·卡斯特罗的政权，加剧了美国与古巴的盟友苏联之间的紧张关系。作为回应，时任苏联总理尼基塔·赫鲁晓夫将短程核导弹运到了古巴。而时任美国总统约翰·肯尼迪则以对古巴的封锁作为回应。最终，苏联做出让步，危机结束了。

艾利森用三个模型解释了这个事件。首先，他运用理性行为者模型（*rational-actor model*）阐明，肯尼迪当时有三种可能的行动：发动核战争、入侵古巴或者进行封锁，最终他选择了封锁。理性行为者模型假设肯尼迪为每种行动绘制了一棵博弈树，并附上苏联可能做出的反应，然后，肯尼迪根据苏联的最优反应来思考自己的行动。例如，如果肯尼迪选择发动核战争，那么苏联就会反击，最终可能会造成数百万人死亡。如果肯尼迪决定封锁古巴，他就会使古巴人挨饿，而苏联则可能选择撤退或发射导弹。考虑到这个选择，苏联应该让步。这个模型揭示了核心策略逻辑（*central strategic logic*），并为肯尼迪大胆选择封锁古巴提供了合理的理由。

然而，尽管如此，像所有模型一样，这个模型也是错误的。它忽略了一些重要的相关细节，使它乍看起来比实际情况更好。这个模型也忽略了苏联已经将导弹运入古巴这个事实。如果苏联是理性的，他们应该会和肯尼迪一样画出博弈树，并认识到他们必须拆除导弹。理性行为者

模型也无法解释为什么苏联没有将导弹藏起来。

其次，艾利森用组织过程模型（organizational process model）解释了这些不一致性。缺乏组织能力是苏联未能隐藏导弹的原因。这个模型也可以解释为什么肯尼迪选择封锁古巴，因为当时美国空军不具备在一次打击中就摧毁导弹的能力。即便只剩下一枚导弹，也会造成数百万美国人的伤亡。艾利森巧妙地结合了这两个模型。来自组织过程模型的洞察力，改变了理性选择模型（rational-choice model）中的结果。

最后，艾利森又使用了政府过程模型（governmental process model）。之前的两个模型都将国家化约为它们的领导者：肯尼迪代表美国行动，赫鲁晓夫代表苏联行动。政府过程模型则认为，肯尼迪不得不与国会抗衡，而赫鲁晓夫则必须维持支持自己的政治基础。因此，赫鲁晓夫在古巴部署导弹是一种力量的宣示。

艾利森这本书分别展示了模型本身以及模型之间对话的威力，每一个模型都能使思路变得更加清晰。理性行为者模型确定了导弹到达古巴后可能采取的行动，并帮助我们看清了这些行动的含义。组织过程模型让我们注意到了是组织而不是个人在实施这些行动。政府过程模型则突出了入侵的政治成本。在通过所有这三个视角评估了这个事件后，我们就有了更全面、更深刻的理解。所有模型都是错的，但是同时运用多个模型确实非常有用。

在这两个例子中，不同的模型解释了不同的因果因素。此外，多模型思维方法也可以专注在不同的尺度上。在一个经常被人提及的故事中，一个孩子声称地球是驮在一头巨大的大象背上的。一位科学家问这个孩子，那么大象又是站在什么东西上呢。孩子回答道：“一只巨大的乌龟的背上。”然后，科学家继续问，孩子继续答。不难预料接下来会发生的事情，孩子的回答是：“你不要再问啦！乌龟驮乌龟，一直驮下去！”¹²

如果我们这个世界真的就是通过乌龟驮乌龟这样维持着的，或者

说，如果这个世界是自相似（self-similar）的，那么最顶层的模型将适用每个层面。但是经济、政治世界和社会都不可能是这样的乌龟队列，大脑也不可能。在亚微米水平上，大脑由构成突触的分子组成，突触组成了神经元，神经元在神经网络中结合。不同的神经网络相互重叠，具体模式可以通过脑成像技术来加以研究。这些神经网络存在的层级低于功能性系统（如小脑）。既然大脑在每个层级都有所不同，我们就需要多个模型，而且这些模型也各不相同。表征神经网络稳健性的模型与用于解释脑细胞功能的分子生物学模型几乎没有任何相似之处，而后者又与用于解释认知偏差的心理学模型有所不同。

多模型思维的成功取决于一定程度的可分离性。在分析2008年金融危机的成因时，我们需要依赖外国人购买资产模型、资产组合模型、金融杠杆模型等多个模型。艾利森在根据博弈论模型进行推导时，不需要考虑组织过程模型。与此类似，在研究人体时，医生会将骨骼系统、肌肉系统、大脑系统和神经系统分开。也就是说，多模型思维并不要求这些不同的模型将系统分割为互不相关的部分。面对一个复杂的系统，用柏拉图的话来说，我们不能“将整个世界雕刻在关节上”。但是，我们可以部分地将主要的因果关系分离出来，然后探讨它们是如何交织在一起的。在这个过程中，我们将发现经济、政治和社会系统产生的数据会表现出一致性。这样一来，社会数据就不会再像家里养的猫一样吐出令人费解的毛球序列了。

做一个多模型思考者

现在总结一下。我们生活在一个充斥着信息和数据的时代。同时，这些数据得以产生的技术条件还极大地缩短了时间和空间上的距离。它们让经济、政治和社会行动者变得更加敏捷，能够在一瞬间就对经济和政治事件做出反应。它们还增加了连通性，因而也增加了复杂性。我们面临着一个由技术引发的悖论：在我们对世界的了解变得更多、更深入

的同时，这个世界也变得更加复杂了。考虑到这种复杂性，任何单个模型都更有可能遭到失败。当然，我们不应该抛弃模型，恰恰相反，我们应该将逻辑一致性置于比直觉更优先的位置；我们不能满足于双重模型、三重模型甚至四重模型，我们要成为多模型思考者。

要成为一个多模型思考者，必须学习掌握多种模型，我们可以从中获得实用的知识，需要理解对模型的形式化描述，并知道如何应用它们。当然，我们也不一定非要成为专家不可。因此，这本书在可阅读性和论证深度之间做了一些权衡，它既可以作为学习资源也可以作为学习指导，书中对各个模型的正式描述都放在独立的专栏中。我还保证不会出现一行接一行都是方程式的情况，如果那样的话，即便是最专注的读者可能也无法忍受。不过，本书还是包括了少数几处包含方程式的论述，但它们都是容易理解的，也是应该被掌握的。构建模型是一门艺术，只能通过不断实践才能熟练掌握，这不是一项以观赏为目的的活动，需要刻意地练习。在建模中，数学和逻辑扮演着专家教练的角色，它们会纠正我们的缺漏。

本书其余各章安排如下。第2章和第3章讨论了多模型思维方法，第4章讨论了对人类建模的挑战。接下来的20几章，每章分别讨论一个模型或一类模型。由于一次只讲解一个模型，所以可以非常方便地将模型的假设、含义和应用厘清。这种章节结构也意味着，我们既可以阅读纸质书，也可以阅读电子书，而且可以直接去阅读与自己感兴趣的模型相关的章节每一章，我们都会应用多模型思维方法去解决各种各样的问题。本书最后给出了两个深度分析：一是针对类药物流行的现象，另外一个则涉及收入不平等问题。

02 模型的7大用途

了解现实就意味着构建转换系统，这些转换系统或多或少都必须与现实相对应。

让·皮亚杰（Jean Piaget）

在本章中，我们定义了模型的类型。人们通常认为，模型就是对世界的简化。是的，模型可以是对世界的简化，但是模型也可以采用类比的形式，或者，模型本身可能就是为探索思想和总结观点而构建的虚拟世界。在本章中，我们还描述了模型的7大用途。在学校里，我们应用模型来解释数据。在现实世界中，我们应用模型来预测、设计和采取行动，也可以使用模型来探索新思想和新的可能性，还可以利用模型来交流思想、增进理解。

模型的价值还体现在，它们能够把特定结果所需要的条件清晰地揭示出来。我们所知道的大多数结论都只是在某些情况下成立。例如，三角形最长边的平方等于另两边平方之和这个结论，只有当最长边是直角的对边时才成立。模型还可以揭示直觉结论可能成立的条件。我们可以分析传染病在什么情况下会传播、市场在什么条件下能正常运行、投票在什么环境下能够得到好的结果、群体在什么条件下能够给出准确预测.....。这些都不是确定的事件。

本章分为两部分。在第一部分，我们描述了构建模型的3种方法。在第二部分，我们介绍了模型的7大用途：推理（reason）、解释（explain）、设计（design）、沟通（communicate）、行动（act）、预测（predict）和探索（explore）。这些用途的首字母，构成了一个缩略词“REDCAPE”。这个缩略词的字面含义为“红色披风”，提醒我们：多模型思维可以赋予我们强大的力量。¹

构建模型的3种方法

要构建一个模型，我们可以在如下所述的3种方法中选择一种。

构建模型的第一种方法是具身法（**embodiment approach**）。用这种方法构建的模型包括重要部分，同时对于不必要的维度和属性，要么剥离，要么将它们整合在一起考虑。生态沼泽模型、关于立法机构和交通系统的模型都是用这种方法构建的，气候模型和大脑模型也是如此。

构建模型的第二种方法是类比法（**analogy approach**），可以对现实进行类比与抽象。我们可以将犯罪行为传播类比为传染病传播，将政治立场的选择类比为在一个左-右连续线段上的选择。球形牛是类比方法的一个最直观的例子：为了估计一头牛身上牛皮的面积，我们会假设那头牛的形状是球形的。之所以要这样做，是因为微积分教科书所附积分表中的公式，会出现 $\tan(x)$ 和 $\cos(x)$ ，但是不会出现类似 $\text{cow}(x)$ 这样的东西。²

相比而言，具身法更强调现实主义，而类比法则致力于刻画过程、系统或现象的本质。当一位物理学家假设不存在摩擦，同时又以其他方式做出符合现实的假设时，他所采用的就是具身法。当一位经济学家将相互竞争的公司视为不同的物种并在此基础上定义产品利基时，就是在做类比，用一个模型来表示不同的系统。但是，在具身法与类比法之间并没有一条明确的界限。例如，关于学习的心理学模型，在给不同的备选项分配权重时，往往会合并考虑多巴胺反应与其他因素，这种模型还会用我们在不同备选项之间进行权衡的方案做类比。

构建模型的第三种方法是另类现实法（**alternative reality approach**），也就是有意不去表征、不去刻画现实。这类模型可以作为分析和计算的“演练场”，我们可以利用这类模型探索各种各样的可能性。这种方法使我们能够发现适用于物理世界和社会世界之外的一般结论。这类模型有助于我们更好地理解现实世界中各种约束条件的含义，

比如如果能够通过空气安全有效地传输能量，那么将会怎样？这类模型还允许我们进行现实世界中不可能的（思想）实验：如果我们能够加快大脑的进化，那么将会怎样？本书包含了不少这种类型的模型，其中一个“生命游戏”（Game of Life），它是一个很大的棋盘，棋盘上的每一个方块要么是活的（黑色），要么是死的（白色），并根据某个特定规则在生死之间切换。虽然这个模型与现实世界并不一致，但是它能够帮助我们加深对自组织、复杂性现象的认识，甚至是许多关于生命本身的洞见。

无论是表征更复杂的现实世界、创造一个类比，还是建立一个用来探索思想的虚拟世界，任何一个模型都必须是易于处理且便于交流的。我们能够用形式化的语言对模型编码，比如数学符号或计算机代码。在描述模型时，我们不能在不给出正式描述的情况下直接抛出诸如信念或偏好之类的东西。信念通常可以表示为一系列事件或先验的概率分布。而偏好则可以用多种方式来表示，比如用对一组备选项的排序或者一个数学函数来表示。

易于处理则是指适合分析的性质。在以往，分析依赖于数学运算或逻辑推理，因此建模者必须能够证明论证中的每一个步骤。这个约束条件导致了一种崇尚极致简约模型的“审美倾向”。神学家、哲学家奥卡姆的威廉（William of Ockham）提出了流传至今的“奥卡姆剃刀”原则：如无必要，勿增实体（Plurality must never be posited without necessity）。爱因斯坦则把“奥卡姆剃刀”原则进一步阐释为：事情应该力求尽可能简单，但是不可过于简单化。不过到了今天，当遇到用解析方法难以处理的问题时，我们还可以求助计算方法，可以构建由许多不断变化的组件的精细模型，而无须考虑解析上是否易于处理。科学家在构建全球气候模型、大脑模型、森林火灾模型和交通模型时，就采用了这种方法。当然，他们仍然不会忘记“奥卡姆剃刀”原则，只不过已经认识到“尽可能简单”还会要求很多不断变化的组成部分。

模型的7大用途

模型有几十种用途，不过在这里，我们只专注讨论其中的7种用途：推理、解释、设计、沟通、行动、预测和探索。

模型的7大用途（**REDCAPE**）

推理：识别条件并推断逻辑含义。

解释：为经验现象提供（可检验的）解释。

设计：选择制度、政策和规则的特征。

沟通：将知识与理解联系起来。

行动：指导政策选择和战略行动。

预测：对未来和未知现象进行数值和分类预测。

探索：分析探索可能性和假说。

REDCAPE：推理

在构建模型时，我们要先确定最重要的行为人（行动者）、实体以及相关特征。然后，描述这些组成部分如何互动和聚合，我们能够推导出一些东西，并说明原因何在。这样一来，也就提高了我们的推理能力。虽然，能够推导出的东西取决于我们的假设，但是我们通过模型发现的绝不仅仅是重言式（tautology）⁽²⁾。因为我们很少能仅凭检验推断出假设的全部影响，我们需要形式逻辑。逻辑还可以揭示不可能性和可能性。利用模型进行推理，我们可以得到精确的，甚至是令人出乎意料的关系。我们可以发现自身直觉的制约性。

阿罗定理（Arrow's Theorem）就是一个可以说明逻辑如何揭示不可能性的极佳例子。这个模型解决了个人偏好是否集结为集体偏好的问

题。在这个模型中，偏好表示为各备选项之间的排序。以对餐馆进行排名为例，假设有5家意大利餐馆，分别用字母A到E表示，这个模型允许120种排序中的任何一种。阿罗要求集体排序是单调的（如果每个人都将A排在B之前，那么集体排序也是如此）、独立于无关的备选项（在其他备选项的排名发生了变化的情况下，如果任何人对A和B的相对排名都没有发生变化，那么A和B在集体排名中的顺序也不会改变），且是非独裁的（没有任何一个人能够决定集体排序）。然后阿罗证明，如果允许任何偏好都存在，那么就不存在集体排序。³

逻辑也可以揭示悖论。利用模型，我们可以证明，每个亚种群中的女性人口比例大于男性，但是在整个种群中却是男性人口的比例更高，这种现象被称为“辛普森悖论”（Simpson's paradox）。在现实世界中，这种情况已经发生过了：1973年，加州大学伯克利分校的绝大多数院系都录取了更多的女生，但是从总体上看，它却录取了更多的男生。模型还表明，两个没有胜算的赌局，当交替轮流进行时，是有可能带来正的预期回报的，这就是人们熟知的“帕隆多悖论”（Parrondo's paradox）。通过模型，我们可以证明，在向网络中添加节点的同时，是可以减少连接所有节点所需边的总边长的。⁴

需要注意的是，我们不能把上面这些模型的例子简单地视为数学上的新奇事物。事实上，每一个模型都有很大的实际应用价值：提高女性在人口中比例的努力可能会适得其反；将没有机会赢利的投资适当地组合起来可能会带来收益；电线、管道网、以太网线路或道路网的总长度可以通过增加更多的节点来减少等。

逻辑也可以揭示数学关系。根据欧几里得定理，三角形可以由任意两个角和一条边，或任意两条边和一个角唯一确定。根据对消费者和公司行为的标准假设，当市场上有大量的相互竞争的企业时，价格等于边际成本。但是，这里也会出现一些出乎意料的结果，其中一个所谓的“友谊悖论”（friendship paradox），它说的是，在任何一个由友人组

成网络中，平均而言，一个人的朋友拥有的朋友要比这个人更多。

“友谊悖论”之所以会出现，是因为非常受欢迎的那些人有更多的朋友。图2-1显示的是扎卡里（Zachary）的空手道网络。在图中，黑色的圆圈所代表的人有6个朋友，这些朋友用灰色圆圈表示，他的朋友们平均每个人有9个朋友。

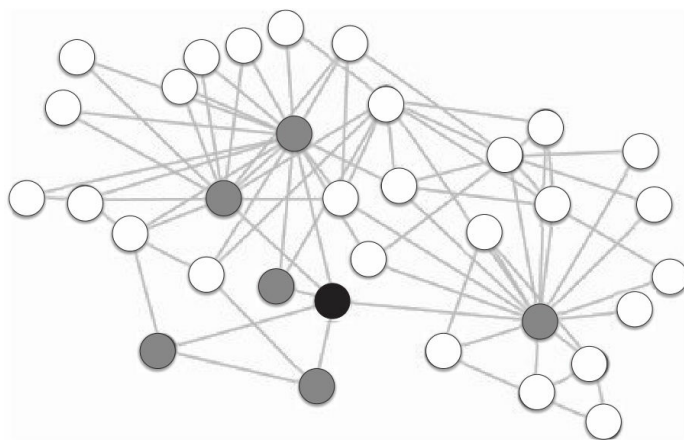


图2-1 友谊悖论：朋友拥有的朋友比自己多

在整个网络中，34人中有29人拥有比他们自己更受欢迎的朋友。⁵稍后在下文中，我们还将了解到，只要加入更多的假设，那么大多数人的朋友平均来说会比自己更加好看、更加善良、更加富有、更加聪明。

最重要的是，逻辑还揭示了真理的条件性。政客可能会声称降低所得税会通过促进经济增长，从而增加政府收入。但是，根据政府收入等于收入水平乘以税率的基本模型，我们很容易就可以证明，只有当收入的百分比增幅超过了减税的百分比时，政府收入才会增加。⁶因此，收入税减少10%的政策，只有在它能够导致收入增幅超过了10%时，才能带来政府收入的增加。政客的逻辑只适用于某些特定条件，而模型就能将这些条件识别出来。

当我们将模型中推导出来的主张与叙述性主张进行比较时，这种“条件性”的威力将会变得更加明显，即便后者有经验证据支持时也是如此。我们先来考虑一下这个管理名言：重要的事情先做（first thing

first)。它说的是，在面对多项任务时，你应该首先完成最重要的那项任务。这个原则有时也被称为“大石头优先”原则，意思是当你要将一些大小不一的石头装入一只桶中时，你应该先装入大石头，如果你先放入小石头，那么大石头就放不下了。

“大石头优先”原则，是专家从观察中总结出来的，在许多时候确实不失为一个很不错的原则，但是它也不是无条件的。基于模型的方法将会先对任务提出具体的假设，然后推导出最优规则。例如，在“装箱问题”（bin packing problem）中，必须将一系列不同大小或不同重量的物体装入容量有限的箱子中，目标是保证所用的箱子尽可能少。

不妨想象这样一个场景：你准备搬家，要把家中的所有东西打好包，放入若干个50厘米×50厘米的箱子里。把你的所有东西按大小排好序，然后将每一件东西放入第一个有足够空间的箱子，这种方法称为“首次适应算法”（first fit algorithm），事实证明相当有效。这就是说，“大石头优先”原则的效果非常不错。

但是，假设我们要考虑一个更加复杂的任务：在国际空间站上，为若干研究项目分配空间。每个项目都对有效载荷重量、空间大小和动力有一定要求，对宇航员的时间和认知能力也有自己的要求。而且，每个项目都有做出科学贡献的潜在能力。在这个问题中，即便我们想出了一个衡量这种“大石头”（重要性）的方法，对上面这些属性求加权平均值的权重，但在给定的相互依赖性的维度下，“大石头优先”原则也已经被证明是一个相当糟糕的原则。更复杂的算法以及可能的市场机制则会更好地发挥作用。²因此，在某些条件下，“大石头优先”原则可能是一个很好的原则。但是，在另外一些条件下，“大石头优先”原则就不行了。通过利用模型，我们可以划出一条界线：什么时候应该采用、什么时候不能采用。

形式主义的批评者声称，说到底，模型只不过是对于我们已经知道的东西进行了重新包装而已，只不过是將旧酒倒入闪闪发亮的“数学”新瓶

中而已。这些批评者可能会说，难道我们不知道“三个臭皮匠，赛过诸葛亮”吗？难道我们不知道“三思而不行，终将无所得”吗？我们不需要模型就能知道这些道理。他们还认为，我们可以通过阅读《荷马史诗》中奥德赛将自己绑在桅杆上的故事，懂得承诺的价值。

但这些批评者没有认识到，从模型中得出的推论总是采用条件判断形式：如果条件A成立，那么可以得出结果B。例如，如果你要装箱，而大小是唯一的约束条件，那么就先装好最大的东西。我们从经典文献和伟大思想家的名言中吸取的教训却通常不包括任何条件。如果我们试图依据这种“原则”来生活或管理他人，就肯定会迷失在众多意思相反的谚语海洋当中，既然有“三个臭皮匠，赛过诸葛亮”，也会有“厨子多了烧坏汤”（表2-1）！

表2-1 相互对立的谚语

谚语	相反的谚语
三个臭皮匠，赛过诸葛亮 (Two heads are better than one)	厨子多了烧坏汤 (Too many cooks spoil the broth)
三思而不行，终将无所得 (He who hesitates is lost)	一针及时顶九针 A stitch in time saves nine
破釜沉舟 (Tie yourself to the mast)	留条后路 (Keep your options open)
“完美”是“优秀”之敌 (The perfect is the enemy of the good)	要做就要尽善尽美，不然不如不做 (Do it well or not at all)
事实胜于雄辩 (Actions speak louder than words)	笔尖强过干戈 (The pen is mightier than the sword)

而在模型中，我们可以在给定的假设下证明定理。相反的谚语经常共存，但是相反的定理却不会出现。两个定理，如果对何为最优行动有不同看法，必定会做出不同的预测；或者，给出了不同解释的定理必定有不同的假设。

REDCAPE：解释

模型为经验现象提供了清晰的逻辑解释。经济学模型解释的是价格变动和市场份额等现象；物理学模型可以解释坠落物体的轨迹和轨迹形状的变化；生物学模型可以解释物种的分布；流行病学模型解释了传染病传播的速度和模式；地球物理学模型能够解释地震的大小和分布。

模型可以解释点值（point values）和点值的变化。例如，某个模型可以解释五花肉期货的当前价格以及过去6个月来价格上涨的原因。另一个模型可以解释为什么美国总统会任命持温和立场的最高法院法官，以及为什么美国总统候选人会向左翼或右翼靠拢。模型还可以解释形状：关于思想、技术和传染病传播的模型，都会产生S形的采用曲线（或传染曲线）。

我们在物理学中学到过不少模型，例如玻意耳定律，这个定律告诉我们，氧气的压力乘以体积等于一个常数（ $pV=k$ ），这个定律非常完美地解释了许多现象。⁸如果知道了体积，就可以估计出常数 k ，然后就可以解释压力 p ，或者预测作为 V 和 k 的函数的压力 p 。这个模型的准确性可以归因于如下事实：气体由大量存在的简单成分组成，而且遵循一个固定不变规则，即任何两个氧分子在相同情况下必定遵循相同的物理定律。氧分子的数量如此之多，以至于统计上的平均值可以抹去任何随机性。

但是，大多数社会现象都不具备这三种性质：社会行动者是异质性的、互动是在小群体内展开的、行为人也不遵守固定的规则。此外，人还会思考。更加重要的是，人会对社会上的风吹草动做出反应，而这就意味着行为变化可能是无法相互抵消的。因此，社会现象要比物理现象更加难以预测。⁹

最有效的模型既能解释简单的现象，也能解决令人费解的问题。教科书中关于市场的经典模型能够解释为什么对于像鞋子或薯片这样正常

商品需求的意外增加，会在短期内提高它们的价格，这是一个非常直观的结果。这些模型还可以解释，为什么从长期来看，需求增加对价格的影响会小于生产商品的边际成本的影响。需求的增加甚至有可能导致价格下降，这种现象在规模收益增加的情况下确实会出现。这无疑是一个更令人惊讶的结果。这些模型还可以解释一些悖论，例如水和钻石悖论：钻石只具有很小的实用价值，但是价格却很高；水虽然是人类生存的必需品，但价格却很低。

有人说，模型可以解释任何东西。这种说法没有错，模型确实可以。然而，基于模型的解释必须包括正式的假设和明确的因果链条，而且这些假设和因果链条都要面对数据。例如，有个模型说，用低被捕概率可以解释犯罪率的居高不下，这样的模型就是可检验的。

REDCAPE：设计

模型还可以通过提供框架来帮助设计，因为只有在适当的框架内我们才可以考虑不同选择的含义。工程师使用模型设计供应链；计算机科学家使用模型设计Web协议；社会科学家使用模型设计制度。

1993年7月，一群经济学家在位于加利福尼亚州帕萨迪纳市的加州理工学院开会，设计一种拍卖方法，拍卖对象是手机所用的电子频谱。在那之前，美国政府一直将频谱的使用权分配给大型公司使用。1993年通过的《统一综合预算协调法案》（*Consolidated Omnibus Budget Reconciliation Act*）则允许政府拍卖频谱以筹集资金。

从一座信号塔发射的无线电信号只能覆盖一定的地理区域。因此，政府可以出售各个特定地区的许可证，例如，俄克拉何马州西部、加利福尼亚州北部、马萨诸塞州、得克萨斯州东部等。这就提出了一个设计问题。一家公司所拥有的任何一张给定的许可证的价值，取决于该公司得到的其他许可证。例如，加利福尼亚州南部许可证对于拥有加利福尼

亚州北部许可证的公司来说更有价值。经济学家将价值的这种相互依赖性称为外部性。这里的外部性有两个主要来源：建设成本和广告市场。持有相邻地区的许可证意味着更低的建设成本和利用重叠的媒体市场的潜力。

这种外部性对同时举行的拍卖提出了挑战。一家试图赢得一组许可证的公司可能会在其中某一张许可证的拍卖中输给另一个竞标人，并因此而失去所有外部性，也就是可以带来的收益。那样的话，这家公司就可能会希望退出其他许可证的拍卖。然而另一方面，连续拍卖也有一个缺点。竞标人在前面的许可证拍卖中会出低价，以对冲在后面的拍卖中竞买失败可能导致的损失。

成功的拍卖制度设计必须符合这样一些要求：不会受策略性操纵的影响、能够产生有效率的结果，同时又容易被拍卖参与者所理解。为此，参加加州理工学院会议的那些经济学家，利用博弈论模型分析了策略性竞标人可能会利用的各种特征，采用计算机模拟比较了各种设计方案的效率，还通过统计模型选择了真人实验的参数。最终，他们设计出了一种多轮拍卖方法，做到了允许参与者退出竞标并禁止早期竞标人掩盖真实意图。事实证明，这是成功的。过去的30年以来，美国联邦通信委员会已经使用这种拍卖方法筹集了将近600亿美元资金。¹⁰

REDCAPE：交流

由于创造了一种共同的表示方法，模型能够有效地改进交流。模型要求对相关特征及其关系给出正式的定义，这使我们能够精确地进行交流。例如，模型 $F=ma$ ，涉及3个可测量的量——力、质量和加速度，并将它们之间的关系用方程式的形式表示出来。每一项都可以表示为可测量的单位，因而可以很方便地就这个模型进行交流，而不必担心会有什么误解。相比之下，“更大、更快的东西会产生更大的力”这种说法的准确度却要低得多。因为这需要翻译，而翻译会令很多人“迷失方

向”。“更大”指的是重量还是体积？“更快”指的是速度还是加速度？“力”指的是能量还是力？“更大”和“更快”的结合又怎么产生“力”呢？对这种说法的定义，也有不同的方向：可以将“力”写为重量与速度之和（ $P=W+V$ ）、重量与速度之积（ $P=W V$ ），又或者写为重量与加速度之和（ $P=W+A$ ）.....

当我们根据可复制性的要求，给像“政治意识形态”这样的抽象概念下了一个定义之后，这些概念也就具有了与质量和加速度等物理概念等量齐观的某些特征。我们可以通过一个模型给出这样的论断，根据他们的投票记录，某个政客比另一个政客更“自由”（“保守”）。然后我们可以准确无误地用这种论断与他人交流。“自由”是有明确定义的，而且是可度量的。其他人可以使用相同的方法去对其他政客进行比较。当然，投票记录可能不是衡量“自由”与“保守”的唯一标准。这时我们可以构建出第二个模型，根据演讲的文本分析来分配意识形态立场。有了这种模型，也可以将我们所说的更加“自由”的意思准确无误地传达给其他人。

很多人都低估了交流对人类社会进步的影响。一个无法交流的思想，就像一棵淹没在森林中的树，没有人会注意到它。启蒙时代显著的经济增长在很大程度上取决于知识的可传播性（知识通常表现为模型形式）。事实上，有充分证据表明，在那个时代，思想的可传播性对经济增长的贡献，比教育水平还要大。其中一个有力的证据是，在18世纪的法国，各城市的经济增长与狄德罗（Diderot）的《百科全书》

（*Encyclopédie*）的订阅数量之间的相关性，远远高于与识字率之间的相关性。¹¹

REDCAPE：行动

弗朗西斯·培根曾经这样写道：“人生的伟大目标，不在于知，而在于行。”良好的行动需要良好的模型。政府、企业 and 非营利组织都要使用模型来指导行动。无论是提高价格（降低价格）、开设新的分支机

构、兼并其他公司、提供全民医疗保健，还是资助某个课外计划，决策者都要依赖模型。在最重要的行动中，决策者要使用多个复杂的模型，模型与数据紧密相关。

2008年，作为《问题资产救助计划》的一部分，美国联邦储备银行提供了1 820亿美元的金融救助款，以拯救跨国保险公司美国国际集团（AIG）。根据美国财政部的报告，政府之所以决定拯救美国国际集团，是“因为它在金融危机期间如果破产，就会对我们的金融体系和经济产生破坏性影响”。¹²救助的目的不是为了拯救美国国际集团本身，而是为了支持整个金融体系。每天都有企业破产，但是政府通常不会介入。¹³

根据《问题资产求助计划》做出的每一项具体决策都是以特定模型为基础的。图2-2显示了国际货币基金组织给出的一个网络模型。在这里，节点（圆圈）代表金融机构，边代表这些金融机构的持有资产价值之间的相互关系。连接的颜色和宽度代表相关性的强度，更深和更粗的线条意味着更大的相关性。¹⁴

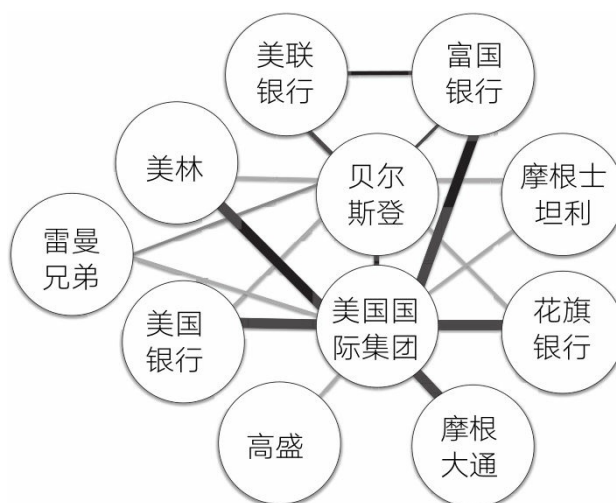


图2-2 金融机构的网络模型

从图2-2可见，美国国际集团在这个金融机构网络中占据了中心位置，因为它向其他公司出售保险。如果这些公司的资产价值蒙受了损

失，美国国际集团根据承诺要向它们支付赔偿金。这就是说，如果资产价格下跌，那么美国国际集团就欠了这些公司钱。这个网络的隐含义是，如果美国国际集团破产了，那么与它相关的公司也会破产，从而很可能会引发一连串的破产。通过稳定美国国际集团，美国政府可以为网络中的其他公司的市场价值提供支持。¹⁵

图2-2也有助于解释为什么政府会让雷曼兄弟公司倒闭：因为雷曼兄弟公司并没有在网络中占据中心地位。我们不能让历史重演，所以我们无法确知美国联邦储备委员会当时是不是采取了正确的行动。但是我们确实知道，雷曼兄弟公司的破产没有导致金融体系崩溃，而且政府向美国国际集团的贷款还为它带来了230亿美元的利润。

指导行动的模型通常依赖于数据，但并不是全部模型都依赖于数据。大多数政策模型都需要使用数学公式，但也并非总是如此。过去，决策者也曾经建立过物理模型。在20世纪中期，菲利普斯（Phillips）为英国经济构建了一个水力模型，它一度被用于考虑政策选择。另外，关于旧金山湾的物理模型，也对终结将旧金山湾改造为淡水湖的计划起到了很大作用。¹⁶密西西比河流域模型水道试验站（Mississippi River Basin Model Waterways Experiment Station）建造的流域缩微模型，位于密西西比州克林顿市附近，占地近80万平方米，按1:100的比例完整复制了整个流域。这个模型可以检验建造新水坝和水库对流域上游和下游地区的效应，因为放出来的水会遵循物理结构中的物理定律。这样的物理模型中，嵌入数据的物理实体和物理定律自然会“完成”逻辑推理。

到目前为止，我们举的例子都是关于组织如何使用模型来采取行动的，个人当然也可以这么做。在日常生活中，当我们准备采取某个重要行动时，也应该使用模型。例如，在决定购买房屋、更换工作、回到大学攻读更高的学位，或者在决定是购买还是租赁汽车时，都可以使用模型来指导决策。用到的模型可能只是一些定性的模型而不一定有相应的数据支持，但是多模型思维会“迫使”我们向自己提出一些重要的问题。

REDCAPE: 预测

模型长期以来被用来预测。天气预报员、专家、顾问和许多国家中央银行行长，都在使用模型进行预测。警察机构和情报部门也使用模型预测犯罪行为，流行病学家则使用模型预测下个季节哪种流感病毒将最为流行。现在，随着数据可得性的提高和精细度的改进，利用模型进行预测的做法变得更加常见了。例如，Twitter上的跟帖和谷歌上的搜索关键词，都已经被用于预测消费趋势和潜在的社会活动了。

模型既可以用来预测特定的个别事件，也可以用来预测一般趋势。2009年6月1日，法国航空公司的AF 477航班在从里约热内卢飞往巴黎的途中，在大西洋上空坠毁。在接下来的几天里，救援人员发现了一些漂浮的碎片，但是无法找到尸体。到7月份，飞机上的水下信标中的电池耗尽了电力，搜索不得不中止。一年后，伍兹霍尔海洋研究所（Woods Hole Oceanographic Institution）率领的搜寻队使用美国海军的侧扫声呐船和水下自动航行器进行了第二次搜索，也没有得到什么结果。于是，法国国家统计分析局（French Bureau d'Enquêtes et d'Analyses）不得不求助于模型。他们将概率模型应用于大海洋流，并识别出了一个坠毁的飞机最有可能沉没的矩形区域（面积并不很大）。根据模型给出的这个预测，搜索队在一个星期之内就找到了飞机残骸。¹⁷

过去，解释和预测往往是齐头并进的。解释电压模式的电气工程模型也可以预测电压大小，解释政客过去投票行为的空间模型也可以预测他们在未来的投票。运用原本用于解释的模型进行预测的最著名的一个例子是，法国数学家、天文学家奥本·勒维耶（Urbain Le Verrier）运用解释行星运动的牛顿定律，预测还存在另一颗行星，进而以此来解释天王星运行轨道的异常。勒维耶证明，那些轨道与太阳系外围地区存在另一颗大行星时的轨道一致。1846年9月18日，勒维耶将预测发给了柏林天文台。5天后，天文学家就在勒维耶预测的那个位置上发现了海王星。

不过话说回来，预测毕竟是与解释不同的。有的模型可以用来预测，但是却不一定能解释什么。深度学习算法可以预测产品的销售情况、明天的天气变化、价格演变趋势和身体健康状况，但是它们几乎没有提供什么解释。这些模型类似于“嗅弹犬”。尽管这些狗可以利用它们灵敏的嗅觉系统确定一个包裹是不是包含着爆炸物，但是我们确实不应该要求它们解释为什么知道那里有炸弹，也不能去问它们工作原理是什么、怎样才能拆除炸弹。

此外还要注意到，有些模型有很强的解释力，但是在预测上却没有什么价值。板块构造论模型虽然可以解释地震是怎样发生的，但是却不能预测地震何时发生；动力系统模型虽然可以解释飓风是怎样形成的，但是却无法准确预测飓风什么时候袭来，也不能准确预测飓风的移动路径；生态模型虽然可以解释物种的形成的模式，但是却无法预测出现的新物种类型到底是什么。

REDCAPE：探索

最后，我们还会用模型来探索直觉。这种探索可能与政策相关：如果让所有城市公交车都免费，会怎么样？如果让学生自主选择作业来证实他们的课程成绩，会怎么样？如果在草坪上标出能量消耗数量，又会怎么样？我们可以提出很多假说，而且所有这些假说都可以用模型进行探索。我们还可以利用模型来探索某些在现实世界中不会出现的情况。如果法国生物学家拉马克（Lamarck）的观点是正确的、如果后天获得的性状真的可以遗传给我们的后代，那么那些把牙齿矫正好了的父母的孩子们就再也不需要牙套了吗？在这样的世界还会发生什么？提出这样的问题并探索它们的含义可以帮助我们揭示进化过程的局限性。暂且将现实世界的约束丢到一边，可以极大地激发我们的创造力。也正是出于这个原因，批判性设计运动的许多倡导者都利用科幻小说来促进思考并提出了不少新的思想。¹⁸

探索有时还涉及对共同假设进行跨领域比较。例如，为了理解网络效应，建模者可能会从一系列程式化的网络结构入手，然后追问网络结构是不是会影响以及如何影响合作、传染病传播或社会动乱。又或者，建模者可能会将一系列学习模型应用于决策、双人博弈和多人博弈；但是他们这样做的目的不是为了解释、预测、行动或设计，而只是为了探索和学习。

当我们在实践中应用一个模型时，也能以多种方式使用它。同一个模型既可以用来解释、预测，也可以用来指导行动。例如，2003年8月14日，俄亥俄州托莱多市附近，树木倒塌压断了电线，造成了局部电力中断，但是，由于监控软件出了故障，没有及时发出警报，让技术人员去对电力进行重新分配，最后导致一天之内，美国东北部和加拿大有超过5 000多万人遭受停电之苦。同一年，意大利和瑞士之间的电线受风暴袭击，导致6 000万欧洲民众无法用电。为此，工程师和科学家求助于将电网表示为网络的模型。这些模型不但有助于解释故障是如何发生的，而且有助于预测未来可能出现故障的区域。它们还能够识别出为了增强电网的稳定性，应该在哪些地方增加新的线路、新的变压器和新的电源，从而起到指导行动的作用。

将一个模型用于多种用途，正是本书中将会反复出现的一个主题。正如接下来将会看到的，一对多是运用多个模型来理解各种复杂现象这一中心主题的必要补充。

03 多模型思维

没有什么比现实主义更不真实了.....细节令人困惑。只有通过选择、通过消除、通过强调，我们才能获得事物的真正意义。

乔治娅·奥·吉弗（Georgia O' Keeffe）

本章将通过科学的方法来引入多模型思维。我们先从孔多塞陪审团定理（Condorcet jury theorem）和多样性预测定理（diversity prediction theorem）入手讨论，这两个定理为证明多模型思维在帮助人们行动、预测和解释方面的价值提供了可量化的论据。需要指出的是，这两个定理可能夸大了许多模型的情况。为了说明原因，我们又引入了分类模型（categorization model），它将世界划分为一个个箱子。使用分类模型的目的是表明构建多模型可能会比预想的更难。然后，我们利用这类模型讨论了模型粒度（model granularity），也就是模型应该有多具体，并帮助我们决定是采用一个大模型还是多个小模型。选择取决于用途：在预测时，我们经常需要大模型；而在解释时，小模型则更好一些。

我们得到的结论解决了一个长期以来挥之不去的忧虑：多模型思维可能需要学习非常多的模型。是的，虽然我们 must 学习掌握一些模型，但是并不需要学习像有些人想象的那么多。我们不需要掌握100个模型，甚至连50个也不需要，因为模型具有一对多的性质。我们可以通过重新分配名称、标识符，或者修改假设来将任何一个模型应用于多种情况。模型的这个性质很好地平衡了多模型思维的需求。事实上，在新的领域应用模型对创造力、开放性和怀疑精神的要求也非常高。我们必须认识到，并非每个模型都适合每项任务。如果一个模型无法解释、预测或帮助我们推理，那就必须将它放到一边，考虑其他模型。

这种一对多的技能，与许多人所认为的要成为一名优秀建模者所必

需的数学和分析才能是不同的。一对多的过程对创造力的要求很高，它实际上相当于在问这样一个问题：对于随机游走，我能够想到多少种用途？作为这种创造力的一个例子，在本章的最后，我们将几何学中的面积公式和体积公式作为模型，解释了超级油轮的大小、评估了身体质量指数、预测了新陈代谢的比例.....并解释为什么我们很少看到女性CEO。

孔多塞陪审团定理和多样性预测定理

现在来看看正式模型，它们有助于理解多模型思维的好处。在这些模型的情境下，我们描述了两个定理：孔多塞陪审团定理和多样性预测定理。

孔多塞陪审团定理是从一个解释多数规则长处的模型中推导出来的。在这个模型中，陪审员要做出要么有罪、要么无罪的二元决策。每个陪审员正确决策的时候比错误的时候多。为了将这个定理应用于模型集合而不是一组陪审员，我们将每个陪审员的决策解释为模型的一个类别。这种分类可以是行动（买入或卖出），也可以是预测（美国民主党胜出还是共和党胜出）。孔多塞陪审团定理告诉我们，通过构建多个模型并使用多数规则，将比只使用其中一个模型更加准确。这个模型依赖于世界状态（state of the world）的概念，它是对所有相关信息的完整描述。对于一个陪审团来说，世界状态包括了审判时呈现的所有证据。对于那些衡量某个慈善项目的社会捐献的模型来说，世界状态则可能与项目的团队、组织结构、运营计划以及项目所要解决的问题的特征或状况相对应。

孔多塞陪审团定理

总数为奇数的一组人（模型）将未知的世界状态分为真或假。每个人（模型）正确分类的概率为 $p > 1/2$ ，并且任何一个人（模

型）分类正确的概率在统计上都独立于任何其他人（模型）分类的正确性。

孔多塞陪审团定理：多数投票正确的概率比任何人（模型）都更高；当人数（模型数）变得足够大时，多数投票的准确率将接近100%。

那么，如何将这个定理的原理应用于多模型方法呢？生态学家理查德·莱文斯（Richard Levins）对此给出了详细的阐述：“因此，我们尝试用几个不同的模型来处理同一个问题，这些模型的简化方法各不相同，但都有一个共同的生物学假设。如果这些模型（尽管它们有不同的假设）都导致相似的结果，那我们就得到了一个强有力的定理，它基本上不受模型细节的影响。因此，我们的真理就是若干独立的谎言的交集。”¹需要注意的是，在这里，莱文斯渴望达成一致的分类。当许多模型都给出了相同的分类时，我们会信心大增。

多样性预测定理则适用于给出数值预测或估值的模型，它量化了模型的准确性和多样性对所有模型平均准确性的贡献。²

多样性预测定理

多模型误差=平均模型误差-模型预测的多样性，即：

$$(\bar{M} - V)^2 = \sum_{i=1}^N \frac{(M_i - V)^2}{N} - \sum_{i=1}^N \frac{(M_i - \bar{M})^2}{N}$$

在这里， M_i 表示模型*i*的预测， $\bar{M}$ 等于模型的平均值， V 等于真值。

多样性预测定理描述了一个数学恒等式。我们用不着费心检验，因为它总是成立。下面举一个例子来说明这一点。假设我们用两个模型来

预测某一部电影会获得多少项奥斯卡奖。一个模型预测它将获得两项奥斯卡奖，另一个模型则预测它将获得8项。这两个模型预测的平均值，也就是多模型预测的结果等于5。如果最后这部电影获得了4项奥斯卡奖，那么第一个模型的误差等于4（ 2^2 ），第二个模型的误差等于16（ 4^2 ），而多模型误差则等于1，模型预测的多样性等于9（因为每个模型的预测与平均预测均相差3）。这样一来，多样性预测定理就可以表达为：1（多模型误差）=10（平均模型误差）-9（模型预测的多样性）。

这个定理的原理在于，相反类型的误差（正负）会相互抵消。如果一个模型的预测值太高，同时另一个模型的预测值太低，那么这些模型就会表现出预测多样性。两个模型的误差相互抵消，模型的平均值将比任何一个模型更加准确。即便两个模型的预测值都太高，这些预测值的平均误差仍然不会比两个高预测值的平均误差更糟。

但是，多样性预测定理并不意味着任何不同模型的集合的预测必定是准确的。如果所有模型都有一个共同的偏差，那么它们的平均值也会包含那个偏差。不过，这个定理确实意味着，任何多样性的模型（或人）的集合将比其普通成员的预测更加准确，这种现象就是通常所说的“群体的智慧”（wisdom of crowds）。这是一个数学事实，它解释了计算机科学中集成方法（ensemble method）成功的原因，这种方法对多个分类加以平均，也解释了使用多个模型和框架进行思考的人比使用单个模型的人预测的准确性更高的事实。任何一种看待世界的单一方式都会遗漏掉某些细节，使我们更容易产生盲点。单模型思考者不太可能准确预测到重大事件，例如2008年的金融危机。³

这两个定理为我们利用多个模型提供了令人信服的理由，至少在进行预测的情况下。然而，这个理由在一定意义上可能显得过强。孔多塞陪审团定理意味着，如果有足够多的模型，我们几乎永远不会犯错。多样性预测定理则意味着，如果能够构建一组多样的中等准确性的预测模

型，我们就可以将多模型误差减少为接近于零。但是，正如接下来将会看到的，我们构建多个多样性模型的能力是有限的。

分类模型

为了说明为什么这两个定理可能会“夸大其词”，现在来讨论一下分类模型。这类模型为孔多塞陪审团定理提供了微观基础。分类模型将世界状态划分为不相交的。最早的分类模型可以追溯到古希腊时代。在《范畴篇》（*The Categories*）一书中，亚里士多德描述了对世界进行分类的10个范畴，包括了实体（substance）、数量（quantity）、地点（location）和状态（positioning）等，每个范畴都会创建不同的类别。

当我们使用一个普通名词时，“裤子”是一个类别，“狗”、“勺子”、“壁炉”和“暑假”也是如此。我们就是在使用类别去指导行动。我们按种族，比如意大利人、法国人、土耳其人或韩国人，来对餐馆进行分类，以便决定在哪里吃午餐；按照市盈率对股票进行分类，并根据市盈率高低买卖股票。当人们声称亚利桑那州的人口之所以增长是因为该州气候宜人时是在用分类方法进行解释。我们还使用类别进行预测，例如预计身为退伍军人的候选人在选举中会有更大的获胜机会。

我们还可以在智慧层次结构中解释分类模型的作用。对象构成了数据，将对象分为不同类别就能创造出信息，而将估值分配给各个类别则需要知识。为了评价孔多塞陪审团定理，我们依赖一个二元分类模型，它将对象或状态分为两个类别，一类标记为“有罪”，另一类标记为“无罪”。关键的思想是，相关属性的数量限制了不同类别的数量，因此也就限制了有用模型的数量。

分类模型

存在一组世界的对象或状态，每个对象或状态都由一组属性定

义，每个属性都有一个值。根据对象的属性，分类模型 M 将对象或状态划分为一个有限的类别集 $\{S_1, S_2, \dots, S_n\}$ ，然后给每个类别赋值 $\{M_1, M_2, \dots, M_n\}$ 。

假设有100份学生贷款，其中有一半是按期还款的，另一半是违约的。我们知道每一笔贷款的两个信息：第一，贷款金额是否超过了5万美元；第二，贷款者主修的工科还是文科。这是两个属性。通过这两个属性，我们可以区分出4种类型的贷款：主修工科学生的大额贷款、主修工科学生的小额贷款、主修文科学生的大额贷款以及主修文科学生的小额贷款。

二元分类模型将上面这4种类型中的每一种都分为按期还款与违约。一种模型可能将小额贷款归为按期还款，将大额贷款则归为违约。另一种模型则可能将主修工科学生的贷款归为按期还款，将主修文科学生的贷款归为违约。我们有理由认为这两种模型中的任何一个都可能在超过一半的情况下是正确的，而且这两种模型大体上相互独立。

但是，当我们尝试构建更多的模型时就会出现这个问题。要将4个类别映射为两个结果，最多只有16个模型。上面这两个模型是其中的两个，它们将所有贷款分为按期还款或违约。剩下的14个模型中的每一个都有一个完全相反的模型，只要某个模型的分类是正确的，那么与之相反的那个模型的分类就是错误的。因此，在14个可能的模型中，最多只有7个可能在超过一半的情况下是正确的。而且，如果任何一个模型碰巧在一半的情况下是正确的，那么与它相反的模型也必定如此。

数据的维数限定了可以创建的模型数量，最多可以有7个模型。我们无法创建出11个独立的模型，更不用说77个了。而且，即使我们有更高维度的数据，比如，假设我们知道贷款者的年龄、平均成绩、收入、婚姻状况和住址，那么依赖这些属性的分类一定能产生准确的预测。每个属性子集都必须与贷款是否已经偿还相关，同时还必须与其他属性无

关。这两者都是很强的假设。例如，如果收入、婚姻状况和住址是相互相关的，那么交换这些属性的模型也将是相互相关的。⁴在严格的概率模型中，独立性是合理的：不同的模型会产生独立的错误。运用分类模型的原理分析孔多塞陪审团定理的逻辑时，我们看到了构建多个独立模型的困难。

在试图构建一组多样性的、准确的模型时，也可能会遇到类似的困难。假设我们想要构建一个分类模型来预测500个中型城市的失业率。一个准确的模型必须将这些城市划分为多个类别，以便让同一个类别中的城市具有相似的失业率，而且该模型必须能够准确地预测该类别的失业率。对于两个进行多样性预测的模型来说，它们必须对城市进行不同的分类或给出不同的预测，或两者兼而有之。这两个标准虽然并不冲突，但却很难同时满足。如果一个分类依赖于平均教育水平，而另一个分类依赖于平均收入，那么它们分类的结果可能是类似的。如果确实是这样，这两个模型可能都将是准确的，但却不是多样性的。根据每个城市名称的第一个字母创建26个类别，可以构造多样性的分类，但却很可能无法成为一个准确的模型。最重要的是，在实践中，“许多”实际上可能更接近5，而不是50。

预测的实证研究结果与这种推论一致。虽然增加模型可以提高准确性（根据多样性预测定理，必定会是这样），但是在已经拥有了一定数量的模型之后再继续增加模型，每个模型的边际贡献就会下降。例如，谷歌公司在实践中发现，仅用一位面试官评估求职者（而不是随机挑选），会使录用一名高于平均水平雇员的概率从50%提高到74%，加入第二位面试官可以把这个概率提高到81%，再加入第三位面试官则只能把这个概率进一步提高到84%，加入第四位面试官也只能提高到86%……使用20位面试官也只能将这个概率提高到90%多一点。这些证据表明，增加面试官人数的作用是有限的。

类似的结果也出现在经济学家对失业率、经济增长率和通货膨胀率

进行的成千上万次的预测中。在这种情况下，我们应该把每位经济学家视为一个模型。加入一位经济学家会使预测的准确性提高大约8%，加入两位可以提高12%，加入3位可以提高15%，加入10位经济学家则能够将准确率提高大约19%。顺便说一句，假设你知道谁是最好的经济学家，那么最好的经济学家的预测只比平均水平高出大约9%。因此，3位随机选择出来的经济学家的表现就已经优于那位最好的经济学家了。⁵相信多位经济学家的平均预测、而不依赖历史上表现最好的经济学家的另一个原因是世界一直在变化。在今天的预测中表现优异的经济学家，明天就可能会泯然众人。同样的逻辑也可以解释为什么美国联邦储备系统要依赖一系列经济模型，而从来不会只依赖某一个经济模型。

这里的教益非常明确：如果能构建出多个多样性的、准确的模型，我们就可以做出准确的预测和估值，并选择正确的行动。这些定理验证了多模型思维逻辑的可靠性。但是，构建出满足这些假设的许多模型，却不是这些定理所能做到的，也不是它们所应该做到的。在实践中，我们可能会发现我们可以构建出3个或5个很不错的模型。如果是这样，那就太好了。我们刚刚讲过，加入1个模型后可以改进8%，加入3个模型后改进幅度可以达到15%。请不要忘记，第二个和第三个模型不一定比第一个模型更好，它们也许会更糟。但是，即使它们的准确性稍差，但只要分类（字面意义）有所不同，就应该把它们加入进来。

适当的模型粒度

许多模型都能在理论上和实践中起到作用，但这并不意味着它们就一定代表正确的方法。有时，我们最好构建一个单一的大型模型。现在，我们就来分析什么情况下应该使用什么策略，同时考虑粒度问题，也就是我们应该在怎样的精细程度上划分数据。

关于应该只用一个大型模型，还是使用多个小型模型的问题，我们先回顾一下模型的7大用途：推理、解释、设计、沟通、行动、预测和

探索。其中有4种用途——推理、解释、沟通和探索都要求我们进行简化。通过简化，我们可以应用逻辑来解释现象、交流思想，并探索各种各样的可能性。

回想一下孔多塞陪审团定理。在这个定理中，我们可以分析内在逻辑，解释为什么使用多模型方法更有可能产生正确的结果，也更有利于传播我们的发现。如果我们构建了一个以人格类型分类的陪审员模型，并将证据描述为语词的载体，我们就会迷失在细节的丛林中。阿根廷著名作家豪尔赫·路易斯·博尔赫斯（Jorges Luis Borges）在一篇科学论文中阐明了这一点。他描述了一批总想制作更精细地图的制图师：“制图师协会决定制作一幅国家地图，它的大小与国家大小相同，而且一对一地将土地上的每一点都标记在地图上。但是，他们的后代不像他们的祖先这样喜欢研究制图，并认为这种巨大的地图毫无用处。”

模型的另外3种用途——预测、设计和行动，却可以因高保真模型而受益。因此，如果有大数据，那么就应该利用它。根据经验，我们拥有的数据越多，模型就越精细。这一点可以通过用来梳理思维的分类模型来说明。假设我们想构建一个模型来解释数据集中的变化。为了给问题提供一个背景，不妨再假设我们从很多杂货店获取了大量数据，详细列出了数百万家庭每个月的食物支出。这些家庭的消费金额不同，我们用变差（variation）来衡量这种变化，也就是每个家庭的支出与所有家庭的平均支出之间的差的平方和。如果每个月的平均支出是500美元，而某个特定家庭每个月的支出为520美元，那么这个家庭对总变差（total variation）的“贡献”就是400（ 20^2 ）。统计学家把一个模型中能够解释的变差比例称为该模型的 R^2 。

如果数据的总变差为10亿，而模型解释了其中的8亿，那么这个模型的 R^2 是0.8。解释的变差比例对应于模型在平均估计上的改进程度。如果某个模型估计某家庭每个月的支出为600美元，而且这个家庭的实际支出确实为每个月600美元，那么这个模型就解释了该家庭对总变差的

全部贡献。如果家庭支出为800美元，但是模型的预测是700美元，那么对总变差的贡献就从原来的9万 $[(800-500)^2]$ ，变成了1万 $[(800-700)^2]$ 。从而模型解释了8/9的变差。

R^2 ：解释变差的百分比

$$R^2 = \frac{\sum_{x \in X} (V(x) - \bar{V})^2 - \sum_{x \in X} (M(x) - V(x))^2}{\sum_{x \in X} (V(x) - \bar{V})^2}$$

其中， $V(x)$ 等于 X 中的 x 的值， $\bar{V}$ 等于平均值， $M(x)$ 等于模型的估值。

在这种情况下，分类模型将家庭划分为不同类别，并估计了每个类别的值。更精细的模型会创建更多的类别，而要创建这些类别就需要更多的家庭属性。如果加入了更多的类别，可以解释的变差比例就会更大。如果我们像博尔赫斯所说的那些制图师一样思考，将每个家庭都分为一类，我们就可以解释所有的变差。但是，这种解释，就像比例为1:1的地图一样，没有多大用处。

创造过多的类别会导致对数据的过度拟合，而过度拟合会破坏对未来事件的预测。假设我们想利用上个月的食物采购数据来预测本月的数据，而家庭每月的支出是会有变化的。如果一个模型将每个家庭都分为一类，那么就可以预测家庭的支出与上个月相同。由于存在月度波动，这个模型并不是一个好的预测器。通过将某个家庭与其他类似的家庭归入同一个类别中，我们可以通过对类似家庭在食品上的平均支出来构建一个更准确的预测器。

为此，我们假设每个家庭的月支出是从某个分布中抽取出来的（我们将在第5章详细讨论各种分布），再假设分布的均值和方差已知。创建分类模型的目的是根据属性构建类别，使同一类别中的家庭具有类似

的均值。如果能做到这一点，那么某个家庭在第一个月内的消费就能够告诉我们其他家庭在第二个月的支出大概是多少。当然，没有任何一种分类方法是完美的。在每个类别中，家庭的均值可能会略有不同，我们称这种情况称为分类误差（categorization error）。

构建的类别越大，分类误差就越大，因为类别越大，我们就越可能将具有不同均值的家庭集中到同一个类别中。但是，更大的类别依赖更多的数据，又可以使我们对每个类别均值的估计更加准确（参见第5章中讨论的平方根规则）。因估计均值错误而出现的误差称为估值误差（valuation error）。估值误差随类别数量的增加而减少。如果不同家庭的月支出不同，那么包含一个家庭的类别（甚至包含10个家庭的类别也一样）将无法准确估计均值，但包含1 000个家庭的类别则能够准确地估计均值。

现在，我们已经得到了关键的直觉：增加类别的数量能够通过将具有不同均值的家庭归入同一个类别减少分类误差。统计学家将这种情况称为模型偏差（model bias）。但是同时，构建更多类别则会增加对每个类别均值估计的误差，统计学家将这种情况称为均值方差的增加。因此，我们在决定要构建许多个类别时就面临着一个权衡。对于这种权衡，我们将它总结为模型误差分解定理（model error decomposition theorem），统计学家则将这个结果称为偏差-方差权衡（bias-variance trade-off）。

模型误差分解定理

偏差-方差权衡

模型误差=分类误差+估值误差

$$\sum_{x \in X} (M(x) - V(x))^2 = \sum_{i=1}^n \sum_{x \in S_i} (V(x) - V_i)^2 + \sum_{i=1}^n (M_i - V_i)^2$$

其中， $M(x)$ 和 M_i 分别表示数据点 x 和类别 S_i 和 $V(x)$ 的模型值， V_i 表示它们的实际值。⁶

一对多

学习模型需要时间精力以及广泛的兴趣和知识。为了减少学习成本，我们可以采用一对多的方法。我们提倡掌握适量的、比较灵活的模型，并学会创造性地应用它们。例如，我们可以使用流行病学模型来解释玉米良种的扩散、Facebook 的风行、犯罪行为的传播和流行明星的“吸粉”。我们将信号传递模型应用于对广告、婚姻、孔雀羽毛和保险费的分析。我们利用进化适应的崎岖景观模型解释为什么人类不需要鲸鱼那样的喷气孔。当然，我们不能随便拿起一个模型就将它应用到任何情境之中。但是，大多数模型都是灵活的。而且，即使失败了我们也会有所获益，因为尝试创造性地使用模型能够暴露它们的局限，这是一件很有趣的事情。

一对多方法是一个相对较新的方法。过去，特定的模型只属于特定的学科。经济学家有供求模型、垄断竞争模型和经济增长模型；政治学家有选举竞争模型；生态学家有关于物种形成和复制的模型；物理学家有描述运动规律的模型，等等。所有这些模型都是针对特定目的而构建的。那个时候，科学家们不会将物理模型应用到经济学领域，也不会用经济学模型去研究大脑，就像普通人不会用缝纫机来修理泄漏的水管一样。

但是今天，将模型从各自所属的学科孤岛中“释放”出来，并将它们以一对多的方法应用到其他领域中去的做法已经取得了显著成功。经济学家保罗·萨缪尔森（Paul Samuelson）重新诠释了物理学中的模型，以解释市场如何实现均衡。经济学家安东尼·唐斯（Anthony Downs）利用经济学中描述海滩上冰激凌商店之间的竞争的模型，解释了相互竞争的政治候选人在意识形态空间上的定位。社会学家应用粒子相互作用的模

型，分析不同国家的贫困陷阱、犯罪率的变化，甚至经济增长。经济学家则已经开始采用基于经济原理的自我控制模型来理解大脑的功能。⁷

一对多：更高的幂

要想创造性地应用模型，需要不断实践。为了说明“一对多”这种方法的巨大潜力，在这里以一个大家熟悉的数学公式 X^N ，也就是求一个变量的 N 次方为例，并将它作为模型应用。当幂等于2时，这个公式给出的是正方形的面积；当幂等于3时，它给出的是立方体的体积。当幂变为更高的值时，这个公式则刻画了几何膨胀或几何衰减。

超级油轮：第一个应用是考虑一艘长方体状的超级油轮，其长度是深度和宽度的8倍，表示为 S 。如图3-1所示，超级油轮的表面积为 $34S^2$ ，体积则为 $8S^3$ 。建造一艘超级油轮的成本主要取决于它的表面积，因为这决定了所需钢材的数量。而超级油轮能够产生的收入数量则取决于它的体积。先计算一下体积与表面积之比，为 $8S^3/34S^2 \approx S/4$ ，这表明，随着尺寸的增加，盈利能力呈线性增长。



图3-1 长方体状的超级油轮：表面积= $34S^2$ ，体积= $8S^3$

航运业巨头斯塔夫罗斯·尼阿科斯（Stavros Niarchos）掌握了这个比例关系，他建造了第一艘现代超级油轮，并在第二次世界大战后的重建期间赚了数十亿美元。第二次世界大战期间使用的T2油轮长152米、深7米多、宽15米多。而现代超级油轮，例如诺克·耐维斯号（Knock Nevis），则长450多米、深20多米、宽50多米。要想象诺克·耐维斯号这样超级油轮的大小，不妨想象一下，将芝加哥的西尔斯大厦放倒，并让

它漂浮在密歇根湖的水面上会是什么样子。诺克·耐维斯号大体上相当于将T2油轮放大了3倍多。然而，与T2油轮相比，诺克·耐维斯号的表面积是T2邮轮的10倍，体积则是T2邮轮的30倍。有人也许会问，那么为什么超级油轮不造得更大一些呢。答案很简单：超级邮轮必须通过苏伊士运河。事实上，诺克·耐维斯号每一次通过苏伊士运河时，都是“挤”过去的，它的两侧都只能剩下一点儿缝隙。⁸

身体质量指数：医学界通常用身体质量指数（BMI）来定义身体质量的不同类别。身体质量指数最早出现在英国，计算方法是一个人的体重与身高的平方比。⁹因此，保持身高不变，身体质量指数会随体重呈线性增长。如果一个人比身高相同的另一个人重20%，那第一个人的身体质量指数就会高出20%。

为了应用模型，我们先将人假设为近似一个完美的立方体，由脂肪、肌肉和骨骼的某种混合物构成。 M 表示1立方米立方体的重量。那么“人体立方体”的重量就等于它的体积乘以每立方米的重量，即 $H^3 \times M$ ，立方体的“身体质量指数”就等于 $H \times M$ 。到这里，这个模型还有两个缺陷：身体质量指数随身高呈线性增长；而且考虑到肌肉比脂肪更重，更健美的人会有更高的 M ，因此会有更高的身体质量指数。身高本应与肥胖无关，而肌肉发达本应是肥胖的对立面。即便我们使这个模型变得更加“真实”，这些缺陷仍然存在。

如果使用参数 d 和 w 来表示一个人的“深度”（前胸到后背的厚度）和“宽度”，并与高度成比例，那么身体质量指数可以写成：

$$BMI = \frac{H \times (dH) \times (wH) \times M}{H^2} = dwHM$$
。这样一来，许多NBA以及其他球类运动明星的身体质量指数将会把他们归入超重类别（ $BMI > 25$ ），甚至许多世界顶尖男子十项全能运动员也不能幸免于难。¹⁰由于即便是身材适中、身体健康的人也可能有很高的身体质量指数，我们不应该对如下结果感到惊讶：对涉及样本总数高达数百万人的近百项研究进行的一个荟萃分析表明，体重稍稍超标的人寿命更长。¹¹

代谢率：现在，应用模型来预测动物大小与代谢率之间的反比关系。每个生物体都要进行新陈代谢，也就是重复进行的一系列化学反应，分解有机物质并将之转化为能量。以卡路里计量的生物体代谢率等于维持生命所需的能量。如果我们构建小鼠和大象的立方米模型，那么从图3-2可知，小立方体的表面积与体积的比值要大得多。

我们可以把小鼠和大象建模为：身体由1立方英寸体积大小的细胞组成，每个细胞都进行新陈代谢，同时这些代谢反应产生的热量必须通过动物的体表皮肤散发掉。小鼠的表面积为14平方英寸，体积为3立方英寸，表面与体积之比约为5:1。¹²因而，对小鼠来说，每立方英寸的细胞，就有5平方英寸的体表皮肤来散热。相比之下，大象的每个发热细胞则仅有1/15平方英寸的体表皮肤来散热。这就是说，小鼠散热的速度是大象的75倍。

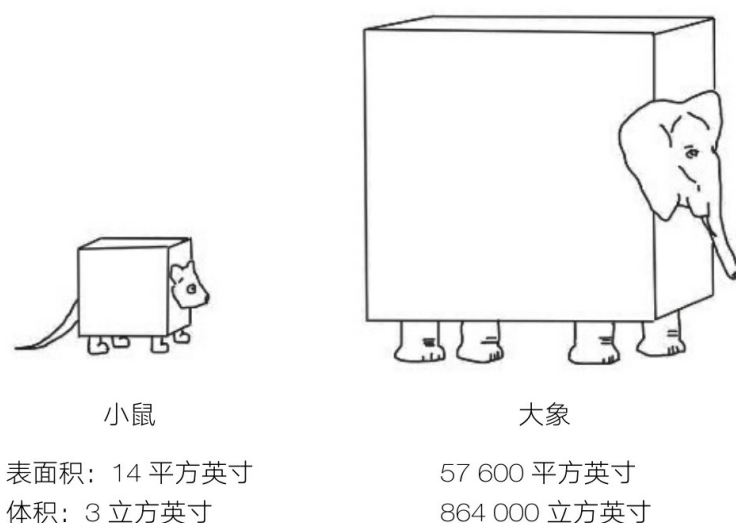


图3-2 膨胀的大象

因此，对于这两种动物来说，要想保持相同的体内温度，大象的新陈代谢就必须更慢。事实也确实如此。如果一头大象的新陈代谢速度与小鼠一样，那么这头大象每天将会需要吃下6 800千克的食物。那样的话，大象的细胞所产生的大量热量无法完全通过它的体表皮肤散发出去。最后，大象将会热到冒烟，然后爆炸。在现实世界中，大象之所以

没有爆炸，原因就在于它们的代谢率为小鼠的1/20。这个模型不能预测新陈代谢随体形大小而变化的速度，但是准确地预测了方向。更精细的模型还可以解释比例定律。¹³

女性CEO：最后，我们进一步增大公式中的指数，并以此来解释为什么只有较少的女性能够成为CEO。根据统计，2016年，只有不到5%的财富500强企业是由女性CEO掌管的。一个人要成为一名CEO，必须经历多次升职。我们可以将这些升职机会建模为概率事件，即一个人有一定概率可以升职。然后进一步假设，要成为CEO，必须做到每一个升职机会都不会错过。

我们假设，要成为一名CEO，至少要升职15次，这大体上相当于每两年升职一次、在30年内成为CEO。大量证据表明，升职时会出现有利于男性的“温和”的偏差。我们可以将这种偏差建模为男性升职的概率更高一些。¹⁴具体地说就是将这种偏差描述为男性的升职概率略高于女性的升职概率。如果将这两个概率分别设定为50%和40%，那么男性最终成为CEO的可能性几乎是女性的30倍！¹⁵这个模型揭示了“温和”的偏差会累积成为非常巨大的差异。10%的升职概率差异，最终变成了成为CEO可能性的30倍的差距。

这个模型也可以为如下现象提供一个新的解释：为什么女性大学校长的比例（大约25%）要比女性CEO的比例高得多？与财富500强企业相比，学院和大学的管理层级较少。一名教授只需升职3次，就可以成为大学校长：系主任、院长，然后就是校长。既然只有3个层级，那么偏差累积的程度就不会太过严重。因此，女性大学校长的比例更高，并不意味着教育机构比企业更加平等。

多模型思维

在本章的一开始，我们通过孔多塞陪审团定理和多样性预测定理为

多对一的方法奠定了逻辑基础。然后，我们使用分类模型说明了模型多样性的局限性，也阐述了多个模型是怎样改进我们在预测、行动和设计等方面的能力的，同时也指出，要想构建多个不同的模型并不容易。如果可以的话，也就能达到接近完美的预测准确度了，但是我们很清楚这是不可能的。无论如何，我们的目标是尽可能多地构建有用的、多样性的模型。

在接下来的各章中，我们将会描述一系列核心模型。这些模型突出了世界的不同部分，它们对因果关系做出了不同的假设。通过它们的多样性，这些模型创造了多模型思维的可能性。通过强调更复杂整体的不同部分，每个模型都可以发挥自己的作用，还可以成为更强大的模型集合的一部分。

如前所述，多模型思维确实要求我们掌握多个模型，但是我们并不需要懂得非常大量的模型，只需要知道每个模型都可以应用到多个领域，但这并不容易。成功的一对多思维取决于创造性地调整假设和构建新的类比，以便将为某个特定目的而开发的模型应用到新的领域。因此，要成为一个多模型思考者，需要的不仅仅是数学能力，更需要的是创造力。这一点我们已经看得很清楚了。

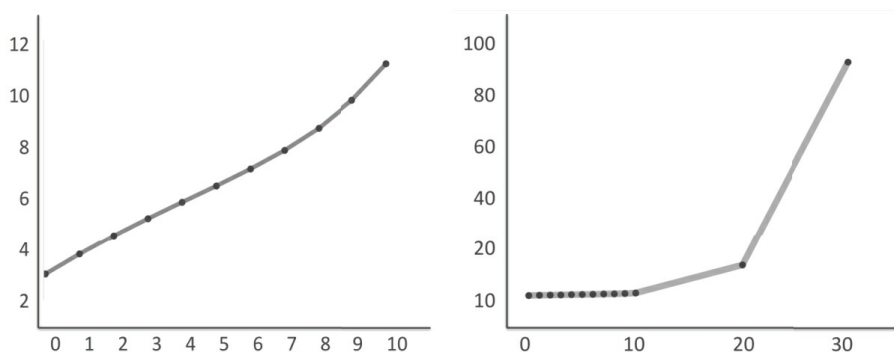
装袋法与多模型

通常，我们会用模型与现有数据集中的样本拟合，然后用其余数据来检验这个模型。而在其他一些时候，我们会用模型去拟合现有数据集，然后用该模型去预测未来的数据。然而，这种构建模型的过程会产生一种张力：模型中包含的参数越多，就越能够很好地拟合数据，同时也越有可能过度拟合。好的拟合不一定意味着好的模型。

物理学家弗里曼·戴森（Freeman Dyson）曾经谈到物理学家恩利克·费米（Enrico Fermi）对他的一项研究的评论。那项研究的模

型拟合度极高。“无奈之下，我问费米是不是对我们计算出来的数值与他测量出来的数值之间的高度一致性没有什么印象。他反过来问我：‘你是用多少个任意参数进行计算的？’我回忆了一下我们的截止程序，然后告诉他‘4个’。他说：‘我记得我的朋友约翰·冯·诺伊曼曾经说过，有4个参数，就可以拟合一头大象；有5个参数，就可以让大象摆动它的大鼻子了。’然后，对话就结束了。”¹⁶

用于“摆动大象鼻子”的估计量通常包括了更高阶的项：平方、立方，甚至四次方。高阶项的存在会带来大误差的风险，因为高阶项有很强的放大效应。10只是5的两倍，但是 10^4 却是 5^4 的16倍。下图显示了过度拟合的一个例子。



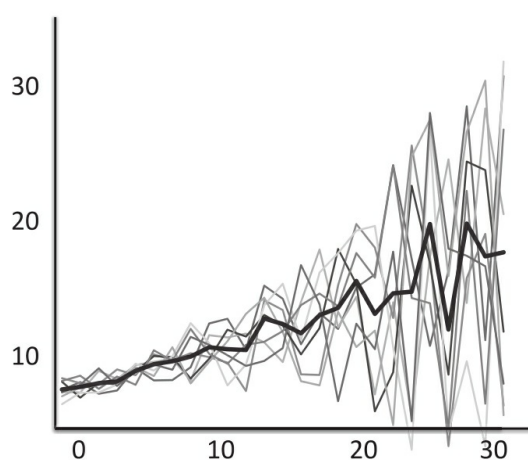
过度拟合与样本外误差

左图显示了一家生产工业用3D打印机的企业销售数据（假设），销售数据是该公司的销售团队每月（平均）上门推销次数的函数。左图显示的是一个非线性最优拟合，包括非线性项的5次方。右图则表明，如果销售团队上门推销的次数达到了30，那么该模型预测3D打印机的销售量将达到100台。如果一个客户最多只购买一台3D打印机，那这个预测就不可能是正确的。因此，由于存在过度拟合，这个模型出现了巨大的样本外误差。

为了避免过度拟合，可以避免使用高阶项。不过，一种更巧妙的解决方法是，可以采取自举聚合法（bootstrap aggregating）或装

袋法（bagging）来构建模型。为了引导数据集，我们从原始数据中随机抽取若干数据点，创建多个规模相同的数据集。抽取这些数据点时，采取的是抽出后放回的方法，也就是说，在抽取了一个数据点之后，我们又将它放回到“袋子”中，下一次仍然可能会抽到它。这种技术产生了一组规模相同的数据集，每个数据集都包含某些数据点的多个副本而不包含其他数据点的副本。

然后，我们将（非线性）模型拟合到每个数据集上，以便生成多个模型。¹⁷这样一来，就可以把所有数据集都绘制在同一组数轴上，从而得到一幅如下所示的“意大利面图”（spaghetti graph），图中颜色最深的那条线表示不同模型的平均值。



意大利面图

装袋法能够刻画鲁棒的非线性效应，因为它们在数据的多个随机样本中都清晰可见，同时又能避免在任何单个数据集中去拟合特殊模式。通过随机样本构建多样性，然后对多个模型求平均值，装袋法很好地应用了多样性预测定理背后的逻辑。它构建了多样性的模型，如前所述，这些模型的平均值将比模型本身更加准确。

04 对人类行为者建模

时至今日，我们仍然无法说出哪一种人类行为理论是设定很成功并在各种环境下都通过了检验的。

埃莉诺·奥斯特罗姆（Elinor Ostrom）

这一章讨论本书的一个核心问题：应该怎样对人进行建模？在接下来将要给出的众多模型中，人都将成为分析的基本单元。我们将构建关于人们投票、合作、参与时尚活动、投资退休账户，以及毒品上瘾的模型。在每一个模型中，我们都必须对“人”做出假设：他们的目标是什么？他们是只关心自己的利己主义者还是利他主义者？他们可能采取的行动是什么？他们如何选择自己要采取的行动？或者说他们是否拥有选择权？

我们可以为每个模型构建任意的特殊假设，但这种做法会引起混乱并错过真正的机会。如果这样做，最终将只剩下一组特殊的构造，而且每一次要构建新模型时，都需要对人们的行为方式进行新的思考。由此产生的异质性会限制我们思考和组合模型的能力，我们将不可能成为有效率的多模型思考者。

我们遵循的方法强调一致性和多样性。或者将人建模为基于规则的行为者（rule-based actor），或者将人建模为理性行为者（rational actor）。在基于规则的行为者集合中，我们考虑那些基于简单固定规则行事的人以及基于适应性规则行事的人。基于适应性规则行事的人能够根据信息、过去的成功或者通过观察他人的行为而改变自己的行为。正如我们在下文中将会详细讨论的那样，在这些不同的情况之间并不存在明确的界线：适应性规则有时可以解释为一个固定规则，理性行为有时也会采取简单规则的形式。

我们怎样对人建模，归根到底取决于问题的背景和想要实现的目标。我们是在预测还是在解释？是在评估政策行动吗？是在尝试设计一种制度吗？或者，是在探索？在低风险的环境中，例如要构建一个预测人们会购买什么颜色的外套或者他们会不会在看完演出后起立鼓掌的模型时，我们通常会假设人们采用固定规则。而在构建关于人们决定是不是要合作创业或信任他人的模型时，我们假设人们会学习和适应。而在高风险的环境中，我们将假设知晓相关信息的、经验丰富的人会做出最佳选择。

在更详细地描述我们的方法之前，先来澄清一些常见的误解。许多人都是在经济学入门课程中第一次接触到描述社会现象的正式模型的，而且那些经济学模型通常依赖一个基本的理性行为者模型。在这个基本模型中，每个人都是自利的，并且有能力实现优化。这个模型通常还假设每个人都有相同的偏好和收入水平，经济学家在这些模型中求解均衡，并在此基础上评估各种冲击对市场或政策变化的影响。这些模型虽然基于不准确的假设，但却很有用。这些模型方便了经济学家之间的交流，也更有利于学生理解。

基于这种经验，许多人推断，构建正式的模型需要一种狭隘的、不切实际的“人性观”，也就是说，必须假设所有人都是自私的，而且从来不会犯错。但事实并非如此。事实上，即便是经济学家也不会这样认为。在经济学的前沿领域，早就出现了包含不完全信息和异质性行为主体的模型。在这些模型中，行为者会根据他们所了解到的东西做出调整，他们有时（尽管不总是）会关心他人的收益。当然，人们会在何种程度上表现出涉他偏好（other-regarding preferences）则取决于具体情况。例如，当向慈善机构捐款或从事志愿工作时，一个人可能会显得比在购房时更加关心他人。

尽管如此，令人遗憾的是：经济学模型总是假设自私的、不切实际的理性行为者。我们必须放弃这种观点。打个比方，如果你只是在沙滩

边的海水里走了几步，那么你可能会推断海水是浅的。但是，当你游到更远的地方时，你就会开始感受到海水的深度。在这里，就让我们从近岸的浅水地带开始尝试。有时，我们会冒点儿险，进一步说明模型如何能够容纳关心他人的、有限理性的行为者。

无论做出什么假设，我们都无法摆脱假设的影响。我们被绑在逻辑一致性的“桅杆”上，不能随便制造影响。如果假设消费者的选择具有强大的社会影响力，那模型就会产生若干占据很大市场份额的产品。如果假设人们通过网络获得信息，那么填补结构漏洞的那些人将会拥有权力。

在本章的其余部分，我们先概述了在对人建模时会遇到的一些挑战：人是多样性的、易受社会影响的、容易出错的、有目的的、有适应能力且拥有自己主体性的。我们不能在一个模型中包含所有这些特征而不会产生复杂的混乱，因此我们必须做出选择。如果异质性无关紧要，那我们也许可以假设完全同质的行为主体。如果问题很简单或人们很精明，那也许可以假设人们不会犯错误。

接下来描述理性行为者模型，并讨论其理论基础以及运用这种模型的理由（尽管它在描述层面上是不准确的）。我们的结论是，理性行为者模型是起到“黄金标准”的作用，还是“稻草人”的作用，抑或是介于这两者之间，都取决于模型的目的。理性行为者模型在预测人类行为方面的作用，不如作为沟通、评估行动和设计政策的工具那么成功。

然后，我们阐明了如何在标准的理性行为者模型中加入心理偏差和利他偏好。是否让模型包含心理偏差或（和）关心他人的偏好，仍然取决于我们正在研究的内容。某些人类心理偏差，例如损失厌恶和现世主义偏差（presentist bias）。这些假设对于退休储蓄或社会骚乱的模型可能很重要，但对于驾驶行为或疾病传播的模型可能不那么重要。

然后，我们将描述基于规则的行为。这类模型的优点是既灵活（我们可以把任何行为记下来，作为一个规则），又易处理。我们所要做

的，是将这种行为用计算机程序编好码，也就是一个基于主体的模型，然后观察接下来会发生什么。这种自由当然也伴随着责任。由于我们可以选择任何一种行为规则，所以必须小心不要做出特殊的假设。在某些情况下，当给定目标函数时，可以证明所用的行为规则是一种最优行为，尽管情况并非总是如此。

最后，在本章的结束部分，我们又回过头去重新讨论理性行为作为基准行为的价值。即使人们没有优化，他们也会适应不断变化的环境和新的知识。这个观察结果带来了各种各样的难题。如果我们根据人们有心理偏差的假设，或者他们会做出不符合自身利益的行为的假设设计制度或政策，我们就不得不承担人们会改变行为的风险。你也许可以愚弄人们一次，但很难愚弄他们两次、三次。尽管不一定能得出理性是唯一合理的假设这种结论，但是逻辑确实支持将理性作为相关的基准。逻辑还支持考虑作为理性下限的简单行为规则。而且，在对任何给定情况进行建模时，我们可以应用任意数量的适应性规则和心理规则，作为探索这些极端之间巨大空间的方式。

对人建模的挑战

对人建模是一个很大的挑战，虽然模型要求低维表征，但人却是天生无法简单地加以表征的。人是多样性的、易受社会影响的、容易出错的、有目的的、有适应能力且拥有自己主体性的，也就是说，我们有行动的能力。

相比之下，诸如碳原子和台球之类的物理对象是没有上述这6个属性的。碳原子不具备多样性，尽管它们可以在化合物中占据不同的位置，例如在丙烷中。碳原子从不违反物理定律，也不会主导有目的的生命。它们不会根据过去的经验改变自己的行为，没有主体性，也不会发起行动或转行。因此，社会科学家会时不时地讽刺：如果电子可以思考，那么物理学就会面临非常大的困难。如果电子也拥有构建模型的能力

力，那么物理学无疑会变得更加困难。

我们可以从多样性所带来的问题开始讨论。人们的偏好、行动能力不同，形成的社交网络、利他主义倾向以及分配给不同行动的认知资源（注意力）也有所不同。如果每个人都一样，那么建模工作就会轻松得多。有时我们会根据统计原理假设行为的多样性可以相互抵消。例如，我们可以构建一个模型，预测慈善捐赠额是收入水平的函数。对于给定的收入水平和税率，有些人可能比我们所假设的（偏好）更利他，而另一些人则可能比假设的更利己。如果偏离模型的偏差达到平均值（在第5章中，我们将给出一些能够解释为什么会是这样的分布模型），那么这个模型的预测就可能是准确的。当然，除非不同人的行动是相互独立的，否则不会出现多样性可以抵消掉的结果。在行为受到社会影响的时候，极端行为会产生溢出效应（spillovers）。当政治活动家鼓动选民时，就会发生这种情况。在下文中模拟社会骚乱时，我们会讨论多样性的这种影响。

人们所犯的错误是否能够相互抵消，取决于具体情境。认知依恋（cognitive attachment）缺失导致的误差就可能是随机且独立的，认知偏差导致的误差可能是系统性的、相关的。人们对最近发生的事件往往更加重视，并且更容易回忆起故事性的情节而不是统计数字。这类共同偏差不会被消除。

还有一个挑战与人们所渴望得到的东西有关。构建与人相关的模型时，一个主要的挑战是如何准确评估他们的目标和目的。有些人渴望财富和名声，有些人则希望为自己所在的社区乃至全世界变得更好做出贡献。在理性行为者模型中，我们直接以函数的形式表示一个人的收益。在基于规则的模型中，目的可能更加隐而不露。这是一种行为规则，人们愿意生活在一个“融合”的社区中，但如果与自己同一种族的人在社区中的比例低于10%，人们就选择离开。这样的规则显然包括人们对自己渴望得到的东西的信念。

对人建模的最后一个挑战来自人的主体性：我们有采取行动的能力，改变行为的能力以及学习的能力。也就是说，在某些情况下，人类可能是一种“习惯生物”：行动可能会超出我们的控制范围。也很少有人会主动选择沉迷于阿片类药物或贫穷。但是，归根到底是人们采取的行动产生了这些结果。

通常，当人们采取的行动产生了不好的结果时，他们会修正自己的行为。我们可以通过在模型中加入学习来捕捉这一点。人们采取的学习方式因环境而异。为了搞清楚自己需要学习多少个小时才能在考试中取得好成绩，或者自己需要每个星期锻炼多少次才能保持好身材时，人们可以根据个人经历或通过内省来学习。而在了解要到哪家杂货店购买食品，或者要不要为某个慈善项目捐款时，人们可以通过观察他人来学习。在第26章中，我们证明，在非策略性行为的环境中，学习机制一般能发挥作用，人们能够学习最好的行动。我们还将证明，在策略性的博弈环境中，则“世事难料”。而且，无论是个人学习还是社会学习，都不一定会带来好的结果。

人的这6个特征的每一个都是潜在的模型特征。如果建模时决定只包含一个特征，那么我们还必须决定在多大程度上来体现它。例如，我们要如何使演员变得多样性？需要包括多少社会影响力？人们会向他人学习吗？要如何定义目标？能拥有多大的主体性？我们所拥有的主体性（或活力）可能比自己所认为的要少。美国社会心理学家乔纳森·海特（Jonathan Haidt）⁽³⁾用骑手和大象的比喻描述了我们缺乏主体性的状态。他这样写道：“当我对自己的弱点感到惊讶时，我所想到的自我形象是一个骑在大象背上的骑手。我手中紧握着缰绳，以为自己只要动一动缰绳，就可以指挥大象，告诉它是该转弯、停步或前行。我是可以指挥它的，但只有当大象没有它自己的欲望时，我才能这么做。一旦大象自己真的想做什么事情，我根本不能左右它。”¹我们有时候能够驾驭大象，有时候却不能。没有任何一种单一对人进行建模的方法适用于所有环境，我们不得不用多种方法对人进行建模。

理性行为者模型

理性行为者模型假设人们在给定收益或效用函数的情况下做出最优选择。这里所说的行为既可以是决策，其收益只取决于行为者个人的行为；也可以发生在博弈中，其收益取决于其他人的行为。在同时进行选择或信息不完全的博弈中，理性行为者模型还需要设定关于其他人将会做什么的信念。

理性行为者模型

行为者个体的偏好由在一组可能的行为上定义的数学形式的效用函数或收益函数（payoff function）来表示。行为个体选择函数值最大化的行动。在博弈中，这种选择可能需要相信其他博弈参与者的行为。

我们构建了一个原始的理性行为者模型，用于描述一个人如何决定将多大比例的个人收入分配给住房支出。这个模型将行为者的效用描述为住房和所有其他消费的函数，后者包括食品、服装和娱乐。这个模型假设了住房的价格和所有其他商品的价格。当然，这个模型并不完全符合现实世界，它认为所有住房都是一样的，并将所有其他商品都归为一个名为消费品的类别，并认为它们完全等价。在这里可以暂且将这些不准确的情况放在一边不予考虑，因为这个模型的目的是解释住房支出占收入的比例。

消费的理性行为者模型

假设：行为者个体的效用来自总消费 C 和住房支出 H ，其效用函数可以写成如下形式：

$$U(C, H) = C^{\frac{2}{3}} H^{\frac{1}{3}}$$

结果：效用最大化的行为者个体会将自己收入的 $\frac{1}{3}$ 用于住房。²

在这个模型中，个体用于住房的收入比例不取决于住房价格与收入水平。这两个结果在数据上都是合理近似值。³除了位于收入分布极端位置上的那些人之外，大多数人都将自己收入的 $\frac{1}{3}$ 用于住房。这个发现具有重要的政策意义：如果房价下跌了10%，人们将会多购买10%的住房。这个结果也为假设同质性行为主体提供了理由。既然人们将收入的某个固定比例用于住房支出，那么住房总支出将只取决于平均收入。

使用效用函数能使模型成为可分析、可检验且易于处理的。我们可以用数据估计函数，可以推导出最优行为，还可以通过更改参数值来提出各种各样“如果.....将会怎样”的问题。假设了一个效用函数，也就意味着假设了偏好一致性，尽管那在现实世界中可能并不存在。任何偏好，要想用效用函数来表示，就必须满足某些公理。证明效用函数存在的定理需要假设一个备选方案集并确定偏好排序。想象一下，我们可以列出一个人可能购买的所有可能商品的组合，偏好排序就是这些商品组合从最受欢迎到最不受欢迎的排序。一个人可能更喜欢加牛奶的咖啡，而不怎么喜欢加柠檬的茶，如果真的是这样，那么他就将咖啡、牛奶的组合排在茶、柠檬的组合前面。

当且仅当在偏好排序中，A排在了B的前面时，效用函数赋予A的值高于赋予B的值，我们就说这个效用函数表示了偏好。偏好要与效用函数一致，就必须满足完备性、传递性、独立性和连续性。完备性要求对所有备选方案定义偏好排序。传递性排除了偏好循环，也就是说，如果有人偏好A甚于B，偏好B甚于C，那么他必定偏好A甚于C。换句话说，如果一个人在苹果与香蕉之间更喜欢苹果，在香蕉与奶酪之间更喜欢香蕉，那么他在苹果与奶酪之间必定更喜欢苹果。这个条件排除了不一致的偏好。

独立性要求人们分别评估彩票的结果。彩票是备选方案的一种概率分布，而不是备选方案本身，比如，一个可能的彩票形式是A的概率为60%，是B的概率为40%。如果，某个人对A的排序高于B，而且对于任何结果中包含了B的彩票，这个人都偏好用A代替B，那么他的偏好就满足独立性。独立性排除了过于强烈的风险规避。一个厌恶风险的人可能会在“去新奥尔良玩一趟”与“到迪斯尼世界玩一下”之间更偏好“去新奥尔良玩一趟”；同时在“肯定可以去迪斯尼世界”与“一半的机会去新奥尔良、一半的机会去迪斯尼世界”之间更偏好“肯定可以去迪斯尼世界”。

连续性要求，如果一个人偏好A甚于B、偏好B甚于C，那么必定存在这样一个彩票，以概率 p 得到A、以概率 $(1-p)$ 得到C，他对B的偏好与对这个彩票的偏好完全一样。连续性条件排除了对某些确定结果的强烈偏好。⁴

除了人们进行优化这一可疑主张外，人们违背独立性和传递性的假设导致许多人质疑理性行为者模型的广泛使用，特别是经济学家。但是，作为建模者，我们有充分的理由采用理性行为者模型。

第一，人们往往会表现得“似乎”在最优化。他们可能会应用产生近似最优行为的规则。当人们打桌球、玩飞盘或开车时，他们当然不会写下一堆数学方程式。为了计算出接住飞盘的准确时机所需要的数学方程式的高深程度，可能会让所有人震惊。然而，人们确实能够接住飞盘。顺便说一句，狗也接得住。因此，从人和狗接飞盘的行为来看，“似乎”两者都解决了一个困难的最优化问题。

同样的逻辑还可以扩展应用到更高维的问题上。对威斯康星州麦迪逊市大都会巴士公司（Metropolitan Bus Company）运维主管哈罗德·泽克（Harold Zurcher）的工作进行分析发现，他在是否要更换公交车引擎以及什么时候更换方面做出的决策，近乎最优。⁵虽然泽克没有用过任何数学方程式，但是他依靠启发式方法取得了成功。这些启发式来自经验，通过利用它们，泽克就能表现得“似乎”一个（几乎）完美的理性行

为者了。

第二，即便人们确实会犯错，但在重复的情况下，人们的学习能力也会推动人们接近最优行为。

第三，在“赌注”（利害关系）很大的情况下，人们更应该投入足够的时间和精力来做出接近最优的选择。人们可能会为了喝一杯咖啡或买一节电池而多付30%，但他们不会在购买汽车或房子时多付30%。学习与更大的“赌注”会使人们表现得更理性，这个观点有充分的经验和实验证据支持。⁶

第四，理性行为者模型简化了分析。大多数效用函数都只有一个唯一的最优行为。一个人可以有上千种次优行为，说人们没有实现最优，就打开了一个拥有巨大可能性的盒子。如果假设人们会通过选择来维护自己的身份或捍卫文化规范，那么我们就可能无法得到一个清晰的答案。理性选择也许是不现实的，但现实主义却是以混乱为代价的。即使知道某个答案是错误的，它也可能比完全没有答案更有用，因为它至少允许我们将模型转化为数据，并讨论某些变量的变化会带来什么影响。⁷

第五，理性行为者假设保证了内部一致性。如果模型假设了次优行为且模型在公共域中，它就可以用来学习。人们可以改变自己的行为，可能不会最优化，但除了最优之外的任何假设都会受到批评，也就是不一致。我们在本章的末尾还会回到这一点。

第六，有人认为这是最重要的一个原因，也就是理性可以作为基准。⁸在设计政策、做出预测或选择行动时，我们应该考虑如果人有理性偏好并且进行最优化时会发生什么。这种做法可能帮助我们找出思维中存在的缺陷。还应该接受这样一种可能性，也就是这种做法会使我们得出这样的结论：理性行为者模型不适用，我们应该选择其他模型。有鉴于此，我们可能会再增加第七个原因：多模型思维。如果人们应用多模型方法，犯错的可能性就会大大减少。

选择理性行为者模型的理由

“似乎”：基于智能规则做出的行为可能与最优或近似最优行为无法区分。

学习：在重复的情况下，人们应该能够接近最优行为。

大的“赌注”：在重大决策中，人们会收集信息并认真思考。

唯一性：最优行为通常是唯一的，从而使模型成为可检验的。

一致性：最优行为创建一致的模型。如果人们学会了利用这样的模型，就不会改变自己的行为。

基准：最优行为提供了一个基准，作为人们认知能力的上限。

损失厌恶和双曲贴现

理性行为者模型受到心理学家、经济学家和神经科学家的挑战。他们指出，这种模型与人类的行为方式不符。来自实验室和自然实验的经验证据表明，人在决策时会受到各种各样的偏差（包括现状偏差）的影响。我们在进行概率计算时会忽略基本比率，对确定的事情赋予的权重过高，也会表现出损失厌恶。

随着越来越多的研究者开始将行为、信念与大脑内的神经过程联系起来，硬连线偏差的证据变得非常引人注目。例如，神经经济学使用脑成像技术来研究与经济相关的行为，如对风险的态度、信心水平和对信息的反应等。⁹著名心理学家丹尼尔·卡尼曼（Daniel Kahneman）指出，到目前为止，我们已经掌握了大量支持区分两种思维方式的证据：快速、直观的基于规则的思考（快思考）和深思熟虑（慢思考）。快思考更容易受到上述各种偏差的影响。¹⁰从长远来看，我们可以从大脑的结构中推断出一些行为模型，但是一定要记住，大脑具有巨大的可塑性。能够通过慢思考来克服各种偏差。

此外，对于仅在少数研究中可见的任何发现，我们都应该保持谨慎。许多心理学研究的结果尚未完全得到证实。2015年一项研究表明，在主要心理学期刊上发表的100个研究结果中，有一半都无法复制。¹¹更何况可复制性本身也并不意味着普遍性。而且，许多研究的被试池也没有足够的经济和文化上的多样性。¹²如果利用更加多样性的被试池，我们应该会看到更少的行为规律，从而提供更大的理由避免对行为进行概括。

在尝试构建更符合现实的模型时，我们必须牢记易处理性这个原则。更符合现实的模型可能需要更复杂的数学。¹³这些困难或担忧当然并不意味着我们必须放弃那些心理现实行为模型，但它们确实意味着我们应该谨慎行事，并将更多的注意力放到那些已经得到很好证明的行为规律上。

下面就来讨论两种已经多次复制成功的偏差：损失厌恶和双曲贴现（hyperbolic discounting）。损失厌恶是指面对收益时，人们表现为风险厌恶，面对损失时，人们却表现为风险偏好。卡尼曼和行为科学家阿莫斯·特沃斯基（Amos Tversky）提出了一个关于这种行为的一般理论，也就是前景理论（prospect theory）。¹⁴损失厌恶初看上去似乎并不是非理性的，但是它意味着对于一个相同的情景，在呈现为潜在损失与潜在收益时，人们会选择不同的行为。¹⁵

例如，人们更偏好肯定能赢得400美元，而不怎么喜欢有机会赢得1000美元的彩票。然而，他们却更愿意选择有可能损失1000美元的彩票，而不愿意选择肯定会损失600美元。同样的不一致性也延伸到非货币领域。医生在收益情境时的选择是风险厌恶的，而当备选方案以损失的形式呈现给他们时，他们则愿意冒更大的风险。

前景理论：示例

收益框架：有两个备选方案。

备选方案A：肯定可以赢得400美元。

备选方案B：如果硬币正面朝上，可以赢得1 000美元；如果背面朝上，什么也得不到。

损失框架：先给你1 000美元，然后给你两个备选方案。

备选方案 $\hat{A}$ ：肯定会损失600美元。

备选方案 $\hat{B}$ ：如果硬币正面朝上，不会损失什么；如果背面朝上，你将损失1 000美元。

在这里，A和 $\hat{A}$ 是等价的，B和 $\hat{B}$ 也是等价的。根据前景理论，会有更多的人选择A和 $\hat{B}$ 。

双曲贴现意味着，人们对近期的贴现更强。标准经济模型假设的是指数贴现（**exponential discounting**），也就是说，人们对未来会以恒定的贴现率贴现。对于一个年贴现率为10%的人来说，明年的1 000美元，相当于今天的900美元；而且他在未来的每一年，都会以10%的贴现率对下一年贴现。但是，大量证据表明，大多数人都不会以固定的贴现率去贴现未来。相反，他们会受即时性偏差的影响：他们对近期的贴现率远远高于更远的未来。¹⁶例如，如果你问人们，在从今天起20年后得到9 500美元与从今天起20年多一天后得到10 000美元之间，更愿意选择哪一个，几乎每个人都会再等一天以便多得到500美元。但是，如果你问同样一批人，在今天就可以得到9 500美元与明天才能得到10 000美元之间，更愿意选择哪一个，那么多数人都选择现在就得到9 500美元。这就是即时性偏差的一个例子。¹⁷

这种偏差会导致时间不一致的行为。20年后，大多数人更愿意再等一天，以得到10 000美元。这种偏差在逻辑上并不一致。双曲贴现可以解释人们为什么会欠下巨额信用卡债务、吃不健康的食品、做出无保护措施性行为，也可以解释许多人不能为退休进行储蓄的原因。

总之，根据对模型用途的设想，我们可以选择假设损失厌恶和双曲贴现，只要这些假设似乎更能匹配大多数人的行为。但我们也可以不这样做，主要原因是，它们可能使模型更加复杂，而不能改变我们所发现的东西的性质；或者，如果假设双曲贴现，模型可能产生不符合实际的行为。

基于规则的模型

现在讨论基于规则的模型。¹⁸基于最优化的模型假设人们最大化的效用函数或收益函数，而基于规则的模型则假设特定的行为。基于规则的模型可能会假设，在拍卖中，一个人的出价总是比拍卖物品的真实价值低10%；或者，如果一个人的朋友一直可以获得更高回报的话，那么这个人会“复制”这位朋友的行为。

许多人将基于最优化的模型等同于数学（模型），而将基于规则的模型等同于计算（模型）。但是基于最优化的模型和基于规则的模型之间的区别并不像人们想象的那样清晰。不妨回想一下前面给出的住房支出模型。最优行为是以一个简单规则的形式呈现的：将1/3的收入用于住房。这两种方法的关键区别在于它们的基本假设。在基于最优化的模型中，对偏好或收益的假设是最基本的；而在基于规则的模型中，对行为的假设才是最基本的。

行为规则既可以是固定的，也可以是适应性的。固定规则意味着始终适用相同的算法。正如理性选择模型可以作为人类认知能力的上限，固定规则模型则可以作为人类认知能力的下限。在市场中，一个常见的固定规则是零智能规则（zero intelligence rule），也就是接受任何能够带来更高收益的报价。这个规则意味着永远不会采取愚蠢的（即减少效用的）行动。假设我们想要衡量单边市场机制的效率，在这种市场中，卖方对某种商品发布报价，买方要么接受、要么放弃。遵循零智能规则的卖方会随机选择一个高于该商品价值的价格，买方则会以低于该商品

价值的价格购买。当我们在计算机模型中对这些行为进行编码时发现，在该市场中，零智能交易者可以得到接近完全有效的结果。因此，交易市场即便在买卖双方不理性的情况下，也可以良好地运行。¹⁹

而适应性规则可以在一系列行为之间切换，演变出新的行为或者复制其他行为。之所以要采取这些行动，是为了提高收益。因此，与固定规则不同，适应性规则需要效用函数或收益函数。这种方法的支持者认为，在任何情况下，只要人们倾向于采取简单而有效的规则，就应该采用基于适应性规则的模型，也就是说，既然人们以这种方式行事，那么就得起按这种方式来建模。²⁰虽然基于规则的模型没有对理性做出明确的假设，但适应性规则确实表现出了生态理性（ecological rationality）——更好的规则会占据主导地位。²¹

为了解释基于适应性规则的模型的工作原理，我们在这里不妨以“爱尔法鲁”（El Farol）²²自组织协调模型为例。²²爱尔法鲁是美国新墨西哥州圣塔菲的一家夜间营业的酒吧，每个星期二的晚上都会举办很吸引人的舞会。每个星期，都有100名潜在舞者要决定是去爱尔法鲁酒吧跳舞还是留在家里。所有这100个人都喜欢跳舞，但是如果酒吧过于拥挤，他们也就不想去了。这个模型假设了一个明确的偏好结构：一个人留在家里的收益为0；如果只有小于或等于60个人参加，那么收益为1；如果有超过60个人参加，收益为-1。

如果我们构建一个固定规则模型，那么任何结果都可能出现。假设为每个人分配这样一个规则：第一个星期，去酒吧，如果发现到场的人超过了60人，那么下一个星期就不去；再下一星期，去。那么，在爱尔法鲁酒吧，第一个星期将会涌进100个人，第二个星期却一个人都不会来，然后第三个星期又会有100个人来……与此不同，爱尔法鲁模型通过赋予每个人一组规则来创建适应性规则。每条规则都告诉个体是不是应该去爱尔法鲁酒吧。规则有几种形式。有些是固定的规则，例如，每隔一星期去一次。其他规则是根据最近几个星期前往爱尔法鲁酒吧的人

数变化趋势来制订的。例如，其中一条规则可能预测这个星期去爱尔法鲁酒吧的人数将与上个星期相同。如果上个星期到场的人数少于60人，那么这个规则就会告诉你这个星期应该去。

基于适应性规则的模型将会给每个规则分配一个分数，这个分数等于该规则给出正确建议的星期所占的百分比。然后每个人都可以采取规则集合中分数最高的那个规则。最好（分数最高）的规则将在几个星期内发生变化。对这类模型的模拟发现，如果每个人都拥有大量的规则集合，那么每个星期二都会有大约60人到场，这就是说，在没有任何中央计划者的情况下实现了协调。或者换句话说，这个适应性规则系统通过自我组织实现了几乎完全有效的结果。

爱尔法鲁模型：适应性规则

有100个人，每个人每个星期都要独立地决定是否前往爱尔法鲁酒吧。如果决定前往，且只有60个人或更少的人到场，那么这个人的收益为1，否则收益为-1，决定不前往爱尔法鲁酒吧的人收益为0。

每个人都有一套规则来决定是否参加。这些规则可以是固定的，也可以依最近一段时间以来的参加人数而定。每个星期，每个人都要按照遵循他的规则集合中曾产生过最高收益的规则行事。

我们可以在（图4-1）微观-宏观循环的框架内解释适应性规则模型（例如，爱尔法鲁模型）中的行为。在微观层面，一组个人（用 a_i 表示）根据规则采取行动，这些规则创建了宏观层面的现象（用 Macro_1 和 Macro_2 表示），如图4-1中向上的箭头所示。在爱尔法鲁酒吧问题中，宏观现象是过去的博弈参与者人数的序列，向下的箭头表示这些宏观现象是如何反馈到个人的行为中的。在爱尔法鲁模型中，每个人可能在应用不同的规则。如果人们所用的规则连续4个星期都导致爱尔法鲁酒吧

人满为患，那么规则就会告诉人们较少参加将会带来更高的收益。当一些人转而采用这些规则后，前往爱尔法鲁酒吧的人数就会减少。微观层面的规则产生宏观层面的现象（过高的“出勤”率），后者又反馈回微观层面的规则。

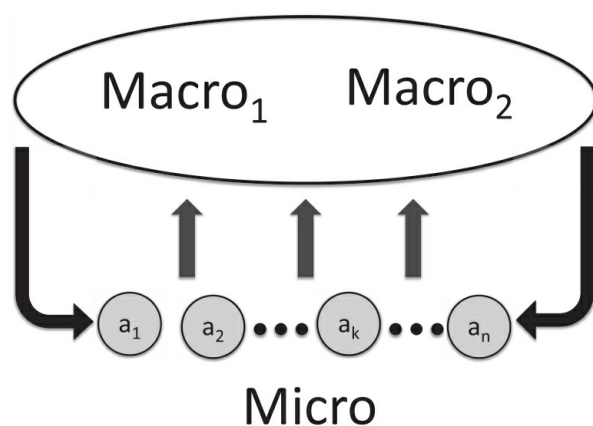


图4-1 微观-宏观循环

模型产生了什么样的结果

微观-宏观循环揭示了一个关键问题：模型中的行为主体应该有多聪明？人们有能力推断出他们行为的所有后果吗？这个循环还暗示了本书中要讨论的一个更大的问题：模型会产生什么样的结果？它是否会达成均衡？还是会产生随机性、创建一个循环，抑或导致一系列复杂的结果？

我们先从行为主体应该有多聪明这个问题开始讨论。现在不妨假设，我们认为个体都只拥有适度的认知能力，因此我们构建了一个零智能行为主体模型。个体的行为集结到一起会产生宏观层面的总体现象。如果宏观层面出现了有效或近乎有效的结果——由买方和卖方组成的单边市场就是如此，那么我们的假设就可能是合理的。既然如此易于遵循的固定规则会产生很好的结果，那么人们将几乎没有动力去开发更精致复杂的规则。

但是，当我们的模型只能产生效率低下甚至糟糕的宏观结果时，就会产生张力，爱尔法鲁模型就是这种情况。在爱尔法鲁模型中，一个共同的固定规则可能会导致完全无效率的周期性结果：这个星期爱尔法鲁酒吧是过度拥挤的，而下个星期却空无一人。面对这种效率低下的结果，我们可能更倾向认为人们会适应。他们可能会试错，可能会考虑整个形势，然后制订新的行动规则。如果将这个逻辑推向极致并假设思考成本很低，我们就会发现自己实际上已经在倡导理性行为者模型了。任何表现不佳的人都可以做得更好。虽然这是事实，但是人们也得有能力制订更好的行动规则才行。

这也就引出了一个大问题：模型产生了什么样的结果？我们有四种选择：均衡、周期、随机性或复杂性。结果的类别将决定我们对于人们应该学会实现均衡的论点的重视程度。如果模型会在宏观层面上产生随机性，那么个人可能无法学到任何东西。好在我们的模型没有这个问题。同样的逻辑也适用于产生复杂模式的模型。在这些情况下，我们可以假设人们能够继续适应新规则，但不能假设他们可以选择最优规则。相反，宏观现象的复杂性会使最优反应显得难以置信。人们更有可能像在爱尔法鲁模型中一样，通过一系列简单的规则来应对复杂性。

产生周期或均衡的模型则可以创造一个稳定的环境，因此我们有理由期望人们可以学习，没有人会持续采取次优的行动。假设现在有这样一个交通模型：每个人都利用一个固定规则选择一条通勤路线。在这个模型中，交通系统处于某个均衡状态。假设在这个均衡中，有一个名叫莱恩的人，每天早上都要花费75分钟从卡拉巴萨斯前往洛杉矶市中心。在给定这个均衡的情况下，如果莱恩从托潘加峡谷（Topanga Canyon）抄近路走，那么她这段行程将只需要45分钟。考虑到每天能够节省30分钟的价值以及住在洛杉矶的人谈论交通的频率，莱恩确实很可能会找到这条更短的路线。事实上，她有很多种方法：可以利用地图软件的线路推荐，或者问一下自己的邻居，也可以自己多探索几次。

因此，如果模型产生了均衡（或简单的周期），并且均衡与优化行为不一致，那就意味着我们的模型存在逻辑缺陷。如果人们可以采取更好的行动，他们应该可以弄清楚，他们应该学习。需要注意的是，为了达到均衡，我们并不需要假设最优化行为。人们可以通过遵循简单的规则来产生均衡；在这种均衡中，任何人都无法通过改变自己的行为来让自己受益。在这种均衡状态下，人们看上去“似乎”正在最优化，事实上确实是。同样，这个逻辑不一定适用复杂或随机的结果。如果洛杉矶的交通模式产生一系列复杂的交通拥堵和通行减速，那么我们没有理由相信莱恩每天都能选择最优路线。事实上，她几乎肯定不能。

如果可以采取任何行动的适应性规则产生了均衡，那么这种均衡必定与致力于最优化的行为主体的行为一致。如果同样这些适应性规则产生了复杂性，那么行为主体的行为就不一定是最优的。我们还可以将这个观点表述为如下形式：最优行为可能是一种不切实际的假设，特别是在复杂情况下。另一方面，如果一个系统产生了稳定的结果，而且某个人可以采取更好的行动，那么这个人很可能会找到这种更好的选择。

同样的逻辑还延伸适用于政策干预。假设我们现在要利用数据来估计人们的行为规则，比如，一个人因为很轻微的健康问题而在午餐时间出现在医院急诊室的可能性。如果假设一个固定规则，那么我们可能会扩建医疗设施，以保证求医的人们不必过多等待。如果人们一直持续遵循这个固定规则，就会达到新的一个均衡：中午的等待时间将会变得很短。然而，在等待时间变得更短了之后，本来不会因为扭伤了脚踝或小小的感冒就去急诊的人，也可能会决定去。这种均衡依赖于人们对次优行为的选择，例如，即使不必等待也不去急诊室。如果人们会学习，我们就不能依靠过去的数据来预测政策变化之后的结果。这个见解被称为“卢卡斯批判”（Lucas critique），是坎贝尔定律（Campbell's law）的一种变体，它指出人们对任何措施或标准的反应都会使其效率降低。²³

卢卡斯批判

政策或环境的变化可能引起受影响者的行为反应。因此，使用过去的行为数据估计的模型将不准确。模型必须考虑到人们对政策和环境变化做出反应这一事实。

这一点我们应该已经清楚了：对于如何对人进行建模这个问题，并不存在一个固定不变的最优答案。如何理性地制订规则及如何根据具体情况制订。我们需要的是在每种情况下尽可能做出最好的判断。考虑各种各样的不确定性，我们应该往构建更多模型的方向试错，而不是更少的模型。

即便我们倾向认为理性选择模型不切实际，也必须认识到它们的易处理性，它们所拥有的揭示激励的正确方向的能力，以及它们作为基准的价值。简单的基于规则的行为，比如零智能，也是不现实的。虽然这些假设是“错”的，但是它们仍然可以使用。它们都很容易分析，可以用来揭示给定环境下智能的重要性。

毫无疑问，人类行为发生在零智能与完全理性这两种极端情况之间，因此构建行为个体利用适应性规则的模型是有意义的。这些规则应该考虑到人们在同一个领域内的认知依恋和认知能力各不相同这一事实。因此，我们应该期待行为多样性会涌现出来，也可以期待群体内部的某种一致性。这些也都可以包含在模型当中。²⁴

总而言之，考虑到对人建模所涉及的复杂性，我们有充分的理由去利用多种不同的模型。我们可能无法准确地预测人们会做什么，但是也许能够确定一系列可能性。如果可以的话，就应该多构建一些模型，我们已经从构建模型中获益，因为我们知道会发生什么。

最后，我们呼吁大家保持谦卑和同理心。在构建关于人的模型时，建模者必须非常谦卑。由于面临着多样性、社会的影响、认知错误、目的性和适应性等多种挑战，我们的模型不可避免地会出现这样那样的问题，这也正是需要采用多模型方法的原因。严谨的行为模型能够很好地

拟合某些情况，并使我们能够专注于环境的其他方面。当我们拥有更多更好的数据时，更丰富的行为模型将会更合适。我们必须保持适度的期望。人是多样性的、易受社会影响的、容易出错的、有目的的、有适应能力且拥有主体性的。怎么能认为单一的人类行为模型不会出错呢？一定会。我们的目标是构建许多模型，作为一个整体，它们将是有用的。

The Model Thinker

What You
Need to Know to
Make Data Work
for You

Part 2 模型思维

要成为一个多模型思考者，必须首先学习掌握多个模型。我们需要理解对模型的形式化描述，并知道如何应用它们。构建模型是一门艺术，只能通过不断实践才能熟练掌握。在建模中，数学和逻辑扮演着专家教练的角色，它们会纠正我们的缺漏。



05 正态分布

我不敢说自己比其他65个人都更聪明——但是我当然要比那65个人的平均水平更高。

理查德·费曼（Richard Feynman）

分布构成任何建模者核心知识库的一部分。从本章开始，我们将利用各种分布来构建和分析路径依赖、随机游走、马尔可夫模型，以及各种搜索模型和学习模型。如果想要度量权力、收入和财富的不平等，并进行统计检验，也需要关于分布的知识。在本书中，我们花了篇幅不大的两章专门讨论分布。本章先讨论正态分布（normal distribution），下一章讨论幂律分布（长尾分布）。我们都是从建模者而不是从统计学家的角度来讨论的。作为建模者，我们对两个主要问题感兴趣：为什么要这样看待分布？为什么分布很重要？

要解决第一个问题，就需要重新认识分布。分布以数学的方式刻画变量的变差（在某个类型内部的差异）和多样性（不同类型之间的差异），将变量表示为在数值上或类别上定义的概率分布。正态分布的形状是我们熟悉的钟形曲线形状。大多数物种的高度和重量都满足正态分布，它们围绕着均值对称分布，而且不会包含特别大或特别小的事件，例如，我们从来没有遇到过1米长的蚂蚁，也没有看到过1千克重的麋鹿。我们可以通过中心极限定理（Central Limit Theorem）来解释正态分布的普遍性。中心极限定理告诉我们，只要把随机变量加总或求其平均值，就可以期望获得正态分布。许多经验现象，特别是像销售数据或投票总数这样的总量数据，都可以写成随机事件总和的形式。

当然，并不是所有事件的规模（大小）都是正态分布的。地震、战争死亡人数和图书销量都呈长尾分布，这种分布主要由很小的事件组

成，也包括极少数非常巨大的大型事件。加利福尼亚州每年都发生超过10 000次地震，但是除非你一直盯着茉莉花的花瓣看它们是否在颤动，否则你不会注意到这些地震。然而，偶然也会出现大的地震：地面裂开、高速公路塌陷，整个城市都在颤抖。

了解系统是否由于多种原因产生正态分布或长尾分布是非常重要的。例如，我们可能想了解电网是否会受到大规模停电的冲击，或者市场体系是否会产生少数亿万富翁和数十亿穷人。有了相关的分布知识，就可以预测洪水超过堤坝的可能性、达美航空238航班准时抵达盐湖城机场的可能性，以及交通枢纽成本超过预算金额两倍的可能性。分布知识对设计也很重要。正态分布意味着不会有太大的偏差，因此飞机设计师不需要为身高5米的人预留腿部空间。对分布的理解也有利于指导行动。正如我们在下文中将会了解的那样，防止骚乱在更大程度上取决于能不能在极端情况下安抚人群，而不在于平时能不能减少不满情绪的平均水平。

在本章中，我们按结构—逻辑—功能的顺序来展开论述。我们先定义了何为正态分布，并描述它们是怎样产生的，然后回答它们为什么这么重要。我们将应用分布知识，解释为什么好的东西总是以小样本的形式出现，检验哪些效应是有显著性的，解释六西格玛（Six Sigma）过程管理为什么有效。然后回到逻辑问题，追问如果我们将随机变量相乘而不是相加会发生什么，结果是获得对数正态分布（lognormal distribution）。对数正态分布可以包括更大的事件，且均值不对称。由此，我们可以推导出，多重效应会导致更大的不平等，这个深刻的结论对提高工资的政策如何影响收入分配有重要的意义。

结构：正态分布

分布为事件或价值分配概率。每日降雨量、考试分数或身高的分布为每一个可能的结果值分配一个概率。各种统计量将分布中包含的信息

压缩为单个数值，例如均值，分布的平均值。德国黑森林中树木的平均高度可能达到24米，开胸手术后的住院时间平均为5天。社会科学家经常通过均值来比较各个国家的经济和社会条件。2017年，美国的人均国内生产总值为57 000美元，远超法国的42 000美元，但是法国人的平均预期寿命则比美国人高出3年。

均值之外的第二个重要统计量是方差，可以衡量一个分布的离散程度，也就是数据与均值之间距离的平方的平均值。¹如果分布中的每个点具有相同的值，那么方差等于零。如果一半数据的值为4，一半的值为10，那么平均来说，每个点与均值的距离为3、方差等于9。分布的标准差是另一个常用的统计量，等于方差的平方根。

可能的分布集合是无限的。我们可以在纸上任意画出一条线并将它解释为概率分布。幸运的是，我们经常遇到的分布一般都属于有限的几种类型。最常见的分布就是正态分布，也就是钟形曲线，如图5-1所示。

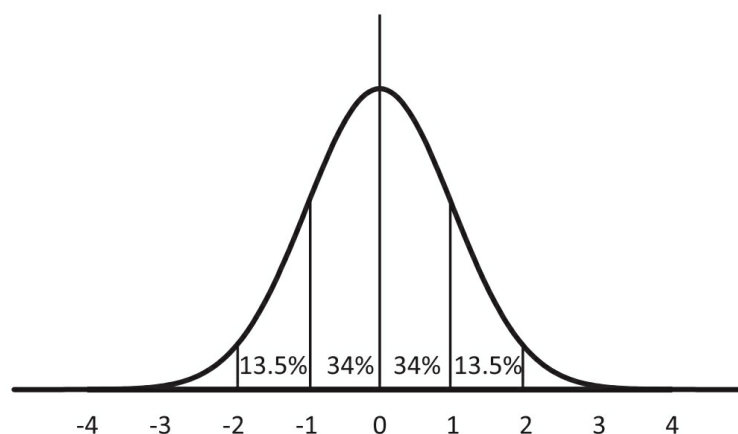


图5-1 正态分布及其标准差

正态分布的均值是对称的。如果一个正态分布的均值等于零，那么抽取到大于3的概率等于抽取到小于-3的概率。正态分布的特征在于其均值和标准差（或者等价地，其方差）。也就是说，所有正态分布的图

形看上去都是相似的，大约68%的结果在均值的一个标准差内，大约95%的结果在两个标准差内，并且超过99%的结果在三个标准差内。正态分布允许任何大小的结果或事件，不过“大”事件是非常罕见的，与均值距离超过五个标准差的事件发生的概率为200万分之一。

我们可以利用正态分布的规律给各种范围的结果分配概率。如果位于美国威斯康星州密尔沃基市房子的平均面积是2 000平方英尺（1平方英尺 $\approx$ 0.09平方米）、标准差为500平方英尺，那么那里68%的房子面积介于1 500平方英尺到2 500平方英尺之间，95%的房子面积介于1 000平方英尺到3 000平方英尺之间。如果2019年的福特福克斯汽车平均每加仑（1加仑 $\approx$ 3.79升）汽油可以行驶40英里（1英里 $\approx$ 1.6千米），且标准差为每加仑1英里，那么超过99%的福特福克斯汽车每加仑汽油可以行驶37英里至43英里。尽管消费者希望自己的汽车越省油越好，但是一般来说不可能每加仑汽油行驶80英里。

逻辑：中心极限定理

非常多的现象都表现为正态分布：动物和植物的体型大小，学生在考试中的成绩，便利店每天的销售额，海胆的寿命，等等。中心极限定理表明为什么对随机变量求和或取均值会产生正态分布。

中心极限定理

只要各随机变量是相互独立的，每个随机变量的方差都是有限的，且没有任何一小部分随机变量贡献了大部分变差，那 $N \geq 20$ 个随机变量的和就近似一个正态分布。²

中心极限定理一个非常重要的特征是，随机变量本身不一定是正态分布的。它们可以有任何分布，只要每一个随机变量都具有有限的方差，并且它们中的任何一小部分随机变量都不贡献大部分方差。假设，

在一个500人的小城镇中，人们的购买行为数据显示，每个人平均每个星期花费100美元。在这些人中，可能有些人这个星期只花50美元、下个星期则花150美元，另一部分人可能每3个星期花费300美元。而其他则可能每个星期的花费在20至180美元之间。只要每个人的支出都只有有限的变差并且没有任何一小部分人贡献了大部分变差，那么分布的总和必定是一个正态分布，其均值为50 000美元。每个星期的总支出也将是对称的：可能高于55 000美元，也可能低于45 000美元。根据同样的逻辑，人们购买的香蕉、牛奶以及炸玉米饼的数量也都是正态分布的。

我们还可以应用中心极限定理来解释人类身高的正态分布。一个人的身高取决于基因、环境以及两者之间的相互作用。基因的贡献率可能高达80%，因此不妨假设身高只取决于基因。研究表明，至少180个基因有助于人体长高。³例如，一个基因可能有助于长出较长的颈部或头部，另一个基因可能有助于长出更长的胫骨。虽然基因之间存在相互作用，但我们可以假设在“长高”这件事情上，每个基因都是相互独立的。如果身高等于180个基因贡献的总和，那么身高将呈现正态分布。相同的逻辑可以证明，狼的体重和大熊猫的拇指长度也是如此。

功能：应用分布知识

我们对正态分布的第一个应用将揭示：为什么罕见结果在规模小的群体中更常见，为什么最好的学校往往规模较小，为什么癌症发病率最高的郡县人口较少。回想一下，在一个正态分布中，95%的结果位于两个标准偏差内，99%的结果位于三个标准偏差内，根据中心极限定理，一组独立随机变量的均值将是正态分布的（当然方差要满足前述要求）。由此可见，我们可以非常确信：考试分数的总体平均值也将是正态分布的。然而，随机变量平均值的标准差并不等于变量标准差的平均值，而且总和的标准差也不等于标准差的总和。相反，这些关系取决于

总体大小的平方根。

平方根法则（**The square root rules**）

N 个相互独立的随机变量，都具有标准差 σ ，对这些随机变量的值的标准差 σ_μ 和对这些随机变量总和的标准差 σ_Σ ，分别由以下公式给出：⁴

$$\sigma_\mu = \frac{\sigma}{\sqrt{N}} \qquad \sigma_\Sigma = \sigma \sqrt{N}$$

均值的标准差公式表明，大的总体的标准差要比小的总体的标准差低得多。由此可以推断，在小的群体中应该会观察到更多的好事和更多的坏事。事实上我们确实观察到了：最安全的居住地是小城镇，但最不安全的地方也是小城镇；肥胖率和癌症发病率最高的那些郡县的人口较少。这些事实都可以通过标准差的差异来解释。

如果不考虑样本量，直接根据离群值（异常值）推断因果关系可能会导致相当糟糕的政策行为。出自这个原因，美国统计学家霍华德·魏纳（Howard Wainer）将均值标准差公式称为“世界上最危险的方程式”。例如，在20世纪90年代，盖茨基金会和其他一些非营利机构以“最好的学校都是小学校”为依据，倡导将大学校分拆为小学校。⁵为了揭示这种推理的逻辑缺陷，试想一下，现在有两所学校，一所是只有100名学生的小学校，另一所是有1 600名学生的大学校，并假设这两所学校学生的成绩均来自相同的分布，平均分为100，标准差为80。在小学校中，平均值的标准差等于8，即学生成绩的标准差80除以学生人数的平方根10。而在大学校中，平均值的标准差则等于2。

如果以平均分为标准，把那些平均成绩在110以上的学校称为“优秀”，把平均成绩在120以上的学校称为“非常优秀”，那么将只有小学校才有可能达到这个标准。对于小学校而言，平均成绩为110时，只比总体均值高出了1.25个标准差，这类事件发生的概率大约为10%。而平均

成绩为120时，则比总体均值高出了2.5个标准差，这类事件大约150所学校发生一次。对大学校进行相同的计算时，我们却会发现“优秀”阈值意味着比均值高5个标准差，而“非常优秀”阈值则比均值高10个标准差！实际上这类事件永远不会发生。因此，最好的那些学校普遍规模较小这个“事实”并不能证明小学校的表现更好。即便学校规模本身完全没有影响，“最好的学校都很小”这种事情也会发生，因为平方根法则会起作用。

检验显著性

我们还可以利用正态分布的规律来检验各种平均值的显著性差异。如果经验均值与假设均值之间的偏差了超过两个标准差，那么社会科学家就会拒绝这两种均值相同的假设。⁶现在提出这样一个假设，即巴尔的摩的通勤时间与洛杉矶的通勤时间相同。假设数据表明，巴尔的摩的通勤时间平均为33分钟，而洛杉矶为34分钟。如果这两个数据集的均值标准差都是1分钟，那么我们就不能拒绝巴尔的摩和洛杉矶两地通勤时间相同的假设。虽然二者的均值不同，但只存在1个标准差。如果洛杉矶的平均通勤时间为37分钟，那么我们会拒绝这个假设，因为均值之间相差4个标准偏差。

但是，物理学家可能不会拒绝这样的假设，至少当数据来自物理实验时不会。物理学家采用更严格的标准，因为他们拥有更大的数据集（原子的数量远远超过了人的数量），数据也更“干净”。物理学家在2012年证明希格斯玻色子（Higgs boson）存在时所依据的证据，在700万次试验中随机出现不到一次。

美国食品药品监督管理局（FDA）所使用的药物批准程序也包含了显著性检验。如果一家制药公司声称自己研发的某种新药可以减轻湿疹的严重程度，那么这家公司就必须进行两项随机对照试验。为了构建一项随机对照试验，该公司组织了两个相同的湿疹患者群体。一组接受这

种药物治疗，另一组则只使用安慰剂。试验结束后，比较平均严重程度和平均副作用发生率。然后，该公司还要进行统计检验。如果药物显著地缓解了湿疹症状（以标准差衡量）且没有显著地导致副作用，则可以批准该药物。美国食品药品监督管理局并没有使用严格的双标准差规则。治疗某种致命疾病且同时只会导致轻微副作用的药物比能够缓解真菌导致的灰指甲症状但同时却会导致骨癌发病率高于预期的药物的统计标准更低。美国食品药品监督管理局还关注统计检验的效力，也就是测试能够证明药物有效的概率。

六西格玛方法

这里要讨论的正态分布规律的最后应用是六西格玛方法，我们将说明正态分布是如何通过六西格玛方法为质量控制提供有效信息的。六西格玛方法是摩托罗拉公司于20世纪80年代中期提出的，目的是减少误差，该方法根据正态分布对产品属性进行建模。试想这个例子：一家企业专业生产制造门把手所用的螺栓。它生产的螺栓必须天衣无缝地与其他制造商生产的旋钮组装在一起。规格要求是螺栓直径为14毫米，但是任何直径介于13毫米与15毫米之间的螺栓也可以接受。如果螺栓的直径呈正态分布，均值为14毫米，标准差为0.5毫米，那么任何超过两个标准差的螺栓都是不合格的。两个标准差事件发生的概率为5%，这个概率对于一家制造企业来说太高了。

六西格玛方法涉及缩减标准差的大小从而降低生产出不合格产品的可能性。各企业可以通过加强质量控制来降低误差率。2008年2月26日，星巴克超过7 000家门店停止营业3小时，目的是重新培训员工。与此类似，航空公司和医院所用的检查清单也有助于减少变差。⁷六西格玛方法降低了标准差，这样即使出现了6个标准差的误差，也可以避免出现故障。在生产螺栓这个例子中，就要求必须把螺栓直径的标准差减少至1/6毫米。而6个标准差的含义是，误差率仅为十亿分之二。实际使

用的阈值假设1.5个标准差的出现是不可避免的。因此，一个六西格玛事件实际上对应于一个四个半西格玛事件，这时允许的误差率大约为三百万分之一。

在六西格玛方法中应用中心极限定理（即隐含的加性误差模型）是如此微妙，因而几乎没有什么人注意到。螺栓制造企业不可能精确地测量每个螺栓的直径，它可能会抽样几百个，并根据这样一个样本来估计均值和标准差。然后通过假设直径的变差源于多种随机效应的总和，例如机器振动、金属质量变化以及压力机温度和速度的波动，就可以利用中心极限定理推断出正态分布。这样一来，这家螺栓制造企业就可以得出一个基准标准差，然后花大力气去降低它。

对数正态分布：乘法冲击

中心极限定理要求我们对随机变量求和或求平均值，以获得正态分布。如果随机变量是不可相加而是以某种方式相互作用的，或者如果它们不是相互独立的，那么产生的分布就不一定是正态分布。事实上，一般情况下都不会是。例如，独立随机变量之间的乘积就不是正态分布，而是对数正态分布。⁸对数正态分布缺乏对称性，因为大于1的数字乘积的增长速度比它们的和的增长速度快，比如， $4+4+4+4=16$ ，但 $4 \times 4 \times 4 \times 4=256$ ；而小于1的数字的乘积则比它们的和小，比如，

$\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$ ，但 $\frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} = \frac{1}{256}$ 。如果将20个不均匀地分布在0到10之间的随机变量相乘，那么多次相乘后所得到的乘积将会包括一些很接近于零的结果与一些相当大的结果，从而生成如图5-2所示的对数正态分布。

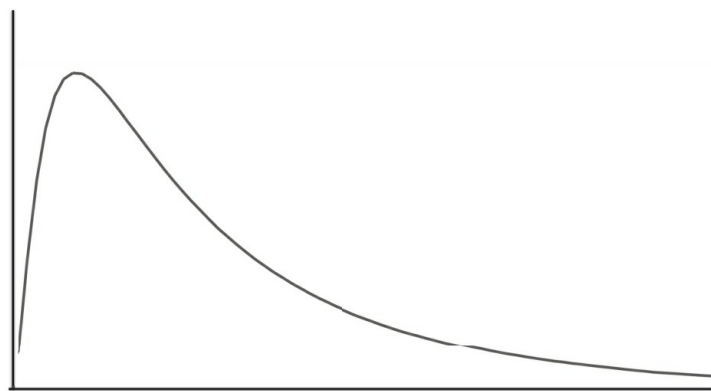


图5-2 一个对数正态分布

一个对数正态分布的尾部长度取决于随机变量相乘的方差。如果它们的方差很小，尾巴就会很短，如果方差很大，尾巴就可能会很长。如前所述，将一组很大的数相乘会产生一个非常大的数字。在各种各样的情况下都会出现对数正态分布，包括英国农场的大小，地球上的矿物质的浓度，从受到感染到症状出现的时间，等等。⁹大多数国家的收入分布也近似于对数正态分布，尽管在最顶端，许多点会偏离对数正态分布，因为高收入的人“太多”了。

一个简单的模型可以解释为什么收入分布更接近于对数正态分布而不是正态分布。这个模型将与工资增长有关的政策与这些政策所隐含的分布联系起来。大多数企业和机构都按某种百分比来分配加薪，表现高于平均水平的人能够得到更高百分比的加薪，表现低于平均水平的人则只能得到更低百分比的加薪。与这种加薪方法相反，企业和机构也可以按绝对金额来分配加薪，例如普通员工可以获得1 000美元的加薪，表现更好的人可以获得更多，而表现更差的人则只能获得更少。

百分比加薪方法与绝对金额加薪方法两者之间的区别乍一看似乎只是语义上的区别，但其实不然。¹⁰如果每一年的绩效都是相互独立且随机的，那么根据员工绩效按百分比加薪，就会产生一个对数正态分布。即使后来的表现相同，未来几年的收入差距也会加剧。假设一名员工因过去几年表现良好，收入水平达到了80 000美元，而另一名员工则只达

到了60 000美元。在这种情况下，当这两名员工的表现同样出色并都可以获得5%的加薪时，前者能够获得4 000美元的加薪，后者却只能得到3 000美元的加薪。这就是说，尽管绩效完全相同，不平等也会导致更大的不平等。如果企业按绝对数额分配加薪，那么两名绩效相同的员工将获得相同的加薪，由此产生的收入分布将接近正态分布。

小结

在本章中，我们讨论了正态分布的结构、逻辑和功能。我们看到，正态分布可以用均值和标准差来表示。中心极限定理说明，当我们将有限方差的独立随机变量相加或求平均值时，正态分布是如何产生的。还给出了随机变量的均值与总和的标准差公式，阐述了这些性质会带来的后果。我们现在已经知道，小的群体更有可能呈现异常事件，如果缺乏对这类事件的洞察力，就会做出不正确的推断并采取不明智的行动。我们还了解到，假设随机变量服从正态分布，科学家就可以对统计检验的显著性和效力做出判断。本章还分析了在过程管理中如何利用正态假设来预测失败发生的可能性。

并不是每个量都可以写成独立随机变量的总和或平均值，因此并非所有事件都满足正态分布。有一些量是独立随机变量之间的乘积，因此它们是对数正态分布的。对数正态分布只取正值，有更长的尾巴，意味着更大的事件和更多非常小的事件。当高方差的随机变量相乘时，尾部会变得更长。长尾分布的可预测性较差，而正态分布则意味着很强的规律性。作为一个预测规则，我们当然更倾向于规律性，而不是发生很大事件的可能性。因此，如果了解了生成各种各样分布的逻辑，我们将会获益匪浅。我们可能更希望随机冲击相加，而不是相乘，以减少发生很大事件的可能性。

06 幂律分布

每个基本定律都有例外，但是你仍然需要定律，否则你所拥有的只是毫无意义的观察。那不是科学，只是做笔记。

杰弗里·韦斯特（Geoffrey West）

在本章中，我们将讨论幂律分布。幂律分布就是通常所称的长尾分布或重尾分布。在把这种分布绘制在图上时，会产生对应大事件的沿水平轴运行的长尾。例如城市人口分布、物种灭绝、万维网上的链接数量以及企业规模等，所有这些分布都有很长的尾巴，视频下载量、书籍销量、学术论文引用数量、战争中的伤亡人数、洪水和地震的分布也是如此。换句话说，在这些分布中，都包括了非常大的事件：东京有3 300万居民，J. K.罗琳的“哈利·波特”系列畅销书卖出了5亿本，1927年密西西比河的大洪水将面积相当于西弗吉尼亚州的地区淹在了9米深的水下.....¹

要想对幂律分布与正态分布之间的巨大差异有一个直观的了解，不妨想象一下人类身高的幂律分布。如果人类身高与城市人口的幂律分布类似，而且假设所有美国人的平均身高为175厘米，那么美国人当中将会有一个人比帝国大厦还高，有超过1万人比长颈鹿还高，同时身高小于18厘米的人也将超过1.8亿人。²

产生幂律分布要求非独立性，通常以正反馈的形式出现。³图书销售、森林火灾的发生和城市人口都不同于光顾杂货店的次数，这些并不是独立的。当某个人买了一本《哈利·波特》后，其他人也可能跟着买；当一棵树着火时，火势会蔓延到邻近的树木；当一个城市的人口增加时，这个城市的基础设施会随之改善，工作机会也会随之增加，从而对其他人更具吸引力。社会学家罗伯特·默顿（Robert Merton）把这种

已经拥有更多的人未来也能够得到更多的现象称为马太效应（Matthew effect），正如《圣经》中所说：“凡有的，还要加给他，叫他有余；凡没有的，连他所有的，也要夺去。”（马太福音25:29）

既然在各种领域中都能发现幂律分布，那么如果有某个机制可以解释所有这些幂律分布就太好了，可惜的是，这种机制并不存在。如果幂律分布的每一个实例都有一个独特的解释，那将更好，可惜的是，这也不是真的。相反，我们只拥有一系列能够生成幂律分布的不同模型，每个模型都能解释不同的现象。

在本章中，我们将重点放在两个幂律分布模型上。第一个模型是优先连接模型（preferential attachment model），它能够解释城市规模、图书销量和网络链接等；第二个模型是自组织临界模型（self-organized criticality model），它能够解释交通拥堵、战争伤亡，以及地震、火灾和雪崩的大小等。在第12章中讨论熵时，我们还会研究第三个幂律分布模型，在那个模型中，幂律会在给定均值的条件下最大化不确定性。在第13章中，我们将证明随机游走模型中的返回次数也满足幂律分布。还有其他一些模型则表明幂律会从最优编码、随机停止规则和组合分布中产生。⁴本章还将讨论幂律分布的结构、逻辑和功能。在讨论中，我们重新评估了特别大的事件的影响，并描述在预防和规划这些事件上的能力局限。

幂律分布的结构

在幂律分布中，事件发生的概率与事件大小的某个负指数成比例。例如，我们熟悉的函数 $\frac{1}{x}$ 就描述了一种幂律。在这个幂律分布中，一个事件的概率与其大小成反比：事件越大，发生的可能性越小。因此，在幂律分布中，小事件的数量要比大事件要多得多。

幂律分布

一个定义在区间 $[x_{min}, \infty)$ 上幂律分布⁵可以写成如下形式：

$$P(x) = Cx^{-a}$$

其中，指数 $a > 1$ 决定了尾部的长度，同时常数项 $C = (a - 1)x_{min}^{a-1}$ 确保总概率的分布。

幂律中指数的大小决定了大事件的可能性和大小。当指数等于2时，事件的概率与其大小的平方成比例。大小为100的事件，发生的概率与 $\frac{1}{100^2}$ （或一万分之一）成比例。当指数增加到3时，该事件的概率与 $\frac{1}{100^3}$ 成比例。对于2或更小的指数，幂律分布缺乏一个可明确定义的均值。例如，从指数为1.5的幂律分布中抽取出来的数据均值永远不会收敛。换句话说，它会无限地增加。

图6-1显示了网页链接数量分布的近似图。

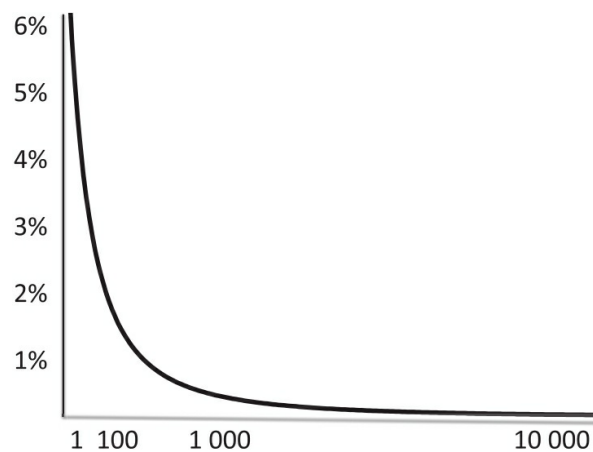


图6-1 网页链接的近似幂律分布

大事件的可能性将幂律分布与正态分布区分开来，因为在正态分布中，我们实际上从未见过大事件，而在幂律分布中，大事件虽然也很少

见，但是它们发生的频率足以引起注意和准备。即使是百万分之一的事件也必须加以考虑。例如，地震大小的分布接近于指数大约为2的幂律。如果发生了震级大于里氏9.0级的地震，不但建筑物会被夷为平地，整个地形地貌都会变得面目全非。这是一个发生的可能性只有百万分之一的大事件，在一个世纪的时间中，这种规模的地震发生的概率为3.5%。⁶

为了更清楚地分析概率为百万分之一的大事件在正态分布与长尾分布之间的差异，现在来看一看由于恐怖袭击所造成的死亡人数的分布，它遵循幂律分布，且指数为2。⁷在长尾分布中，概率为百万分之一的恐怖袭击事件是一个差不多有800人死亡的事件。如果由于恐怖袭击造成的死亡人数满足一个均值为20、标准差为5的正态分布，那么概率为百万分之一的事件将只会导致不到50人死亡。

幂律分布有明确的定义，不是每一个长尾分布都是幂律分布。要想快速地检验某个分布是不是幂律分布，可以用双对数坐标系把该分布画出来：双对数坐标系可以将事件大小及其概率转换为相应的对数值，并将幂律分布转换为直线（图6-2）。⁸

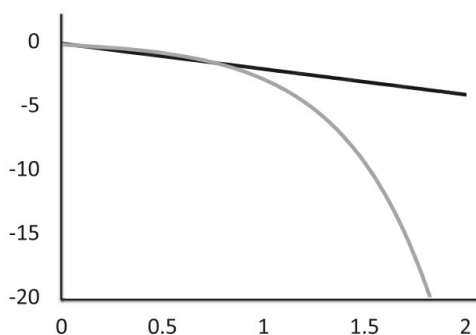


图6-2 双对数坐标中的幂律分布（黑色）与对数正态分布（灰色）

换句话说，在双对数坐标系中，自始至终都呈直线的图形就是幂律分布的证据，而一开始是直线然后逐渐下降的图形则与对数正态分布（或指数分布）相对应。对数正态分布图形向下弯曲的速率取决于产生

分布的变量的变化。⁹当我们增大对数正态分布的方差时，对数正态分布的尾部增大，从而使在双对数坐标系中的图形更接近线性。¹⁰

齐普夫分布（Zipf distribution）是幂律分布的一个特例，即指数等于2的幂律分布。指数等于2的幂律分布的一个重要特征是，事件的等级排列序号乘以其概率等于常数，这个规律被称为齐普夫定律（Zipf's Law）。单词符合齐普夫定律，最常见的英语单词the出现的频率为7%，第二最常见的英语单词of出现的频率为3.5%。请注意，of的等级排列序号2乘以频率3.5%，恰恰等于7%。¹¹

齐普夫定律

对于指数为2的幂律分布（ $a=2$ ），事件的等级排列序号乘以它的大小等于常数，即：

$$\text{事件等级} \times \text{事件大小} = \text{常数}$$

包括美国在内的许多国家的城市人口分布大体上符合齐普夫定律。从美国2016年的城市人口数据可以看出，每个城市的人口排名乘以它的人口总数的值接近800万（表6-1）。

表6-1 城市人口分布

排名	城市	2016 年人口	排名 × 人口
1	纽约，纽约州	8 600 000	8 600 000
2	洛杉矶，加利福尼亚州	4 000 000	8 000 000
3	芝加哥，伊利诺伊州	2 700 000	8 100 000
4	休斯敦，得克萨斯州	2 300 000	9 200 000
5	菲尼克斯，亚利桑那州	1 600 000	8 000 000

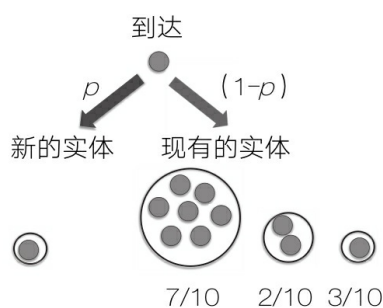
幂律分布的逻辑

现在，我们着手讨论若干产生幂律分布的模型。如果没有适当的模型，幂律分布就只是一种无法解释的模式。

我们要讨论的第一个模型是优先连接模型。模型假设实体以相对于其比例的速度增长。优先连接模型刻画了罗伯特·默顿所说的马太效应：更多导致更多。这个模型考虑了通过新移民到来而实现增长的人口。新到达的人，要么加入现有的某个实体，要么自己创建新的实体。如果是前者，那么加入现有某个实体的概率与该实体的大小成正比。

优先连接模型

一连串物体（人）一个接一个地到达。第一个到达者创建一个实体。后续每次有人到达时都应用以下规则：在概率 p （较小）的情况下，新到达者创建一个新的实体；在概率 $(1-p)$ 的情况下，新到达者加入现有的某个实体。加入某个特定实体的概率等于该实体的大小除以到目前为止所有到达者的数量。



不妨想象一下大学新生进入大学校园时的情景。第一个来到学校的学生创建了一个新的社团，第二个到达的学生以较小的概率创建了自己的社团，更有可能的是，他会加入第一个学生创建的社团。前10个到达的学生可能会创建3个社团：一个有7个成员，一个有两个成员，一个有一个成员。第11个到达的学生只会以极小的概率创建第4个社团，如果不创建新的社团，她就加入现有的社团。如果这样做，那么她有70%的可能性加入已有7个学生的社团，有20%的可能性加入已有两个学生的

社团，只有10%的可能性加入只有一个学生的社团。

优先连接模型有助于解释为什么网络链接、城市规模、企业规模、图书销量和学术引用数量的分布都是幂律分布。在这些情况下，一个行动（比如一个人购买了一本书）会增加其他人也这样做的可能性。如果从某家企业购买商品的概率与它在当前市场的份额成正比，同时如果新企业进入市场的概率较低，那么优先连接模型预测企业规模的分布将是幂律分布。同样的逻辑也适用于图书销量、音乐下载量和城市发展。

我们要讨论的第二个模型是自组织临界模型，它通过在系统中建立相互依赖关系的过程产生幂律分布，直到系统达到临界状态为止。自组织临界模型有很多种。其中，沙堆模型（sand pile model）假设有人将沙粒从距桌面几十厘米的地方洒落到桌子上。随着沙粒不断增多，一个沙堆开始形成。最终，沙子的堆积会达到临界状态，此后每加一次沙子都可能导致“沙崩”。在这种临界状态下，多加入的沙子通常要么没有影响，要么最多只会导致一些沙子下滑。这些属于幂律分布中的数量众多的小事件。但有时，只要再加入一粒沙子就会导致大规模的“沙崩”，这就是大事件。

森林火灾模型（forest fire model）也是自组织临界模型的一种。假设树木可以在一个二维网格上生长，这些树木也可能会随机地被闪电击中。当树木的密度较低时，由闪电引发的任何火灾的规模都很小，最多只会蔓延到几个格点。当树木密度变得足够高时，再被闪电击中就会导致森林大火。

森林火灾模型

“森林”最初只是一个空的 $N \times N$ 网格。每个周期在网格上随机选择一个格点。如果该格点为空，那么就以概率 g 在那里种上一棵树。如果该格点上已经有树，那么闪电会以概率 $(1-g)$ 击中该格点。如果该格点有一棵树，那么树会着火，火势会蔓延到所有连接

到该格点的有树的格点。

这里需要注意的是，在森林火灾模型中，被闪电击中的概率等于1减去种树的概率。这种结构使我们能够改变种树与闪电击中树的相对速度。这是一种简化，减少了模型中参数的数量。在对各种各样的种树速度进行试验之后，我们发现，当种树的速度接近1时，树木的密度会增加到一个临界状态：在这个相对茂密的森林中，被闪电击中有可能摧毁很大一片森林。在这种临界状态下，森林中斑块大小的分布，以及火灾大小的分布，都满足幂律分布。此外，森林还会自然而然地趋向这种密度水平。如果密度较低，密度会增加（因为火灾很小）。如果密度超过了阈值，那么任何火灾都会毁掉整个森林。因此，树木密度自组织地达到了一个临界状态。¹²

在沙堆模型和森林火灾模型中，宏观层面的变量，也就是沙堆的高度或森林的密度，都具有一个临界值。当有像沙崩或火灾这样的大事件发生时，宏观层面的变量值会减小。这两个模型的一些变体可以解释太阳耀斑、地震和交通拥堵的分布。不过，当事件发生时，不断增加的宏观层面的变量会减少，这虽然是必要的，但对于自组织临界性来说是不够的。均衡系统也具有这种特征。水通过溪流，流入和流出湖泊，但是由于水流很平稳，所以湖水的水位是逐渐变化的。通过自组织达到临界状态的关键假设是压力平稳地增加（就像水流入湖中一样），同时压力在爆发时迅速减少，这包括可能发生的大事件。

长尾分布的含义

在这里，我们讨论长尾分布的三个含义，即它们对公平、灾难和波动性的影响。根据定义，与正态分布相比，长尾分布意味着少数几个大“赢家”（大崩溃、大地震、大火灾和严重的交通拥堵）和很多的“输家”；而正态分布则是关于均值对称的。长尾分布也可能增加波动性，

因为更大实体中的随机波动会产生更大的影响。

公平

如果某一个人写的书更好、创作的歌曲更有吸引力、发表的论文学术水平更高，那么他应该比其他人获得更大的名声和更多的金钱。但是，如果另一个人只是因为表现得稍微好一点，或者完全靠碰巧走运就比其他人赚到了多得多的钱、获得了大得多的名声，那就有失公平了。就像我们在优先连接模型中看到的，因为马太效应，正反馈创造了少数大赢家。在市场中，要发生正反馈，人们必须知道别人买了什么商品，而且人们必须有能力购买商品。就手机上的应用程序而言，根本不存在可能会减慢正反馈的生产限制，但是卡车就会面临这种约束。福特公司不可能无限增加F-150卡车的产量，但是财捷集团（Intuit）却可以无限量地销售TurboTax应用程序，只要有人愿意下载。

实证研究表明，社会效应会创造更大的赢家。在音乐实验室的实验研究中，研究者让大学生挑选和下载歌曲。在第一个实验组中，被试不知道其他人下载了哪些歌曲，下载量的分布具有较短的尾部，没有出现下载量超过200次的歌曲，且下载量少于30次的歌曲也只有一首。在第二个实验组中，被试知道其他人下载了哪些歌曲，下载量的分布具有较长的尾巴，有一首歌的下载量超过300次。而且，超过一半歌曲的下载量都不到30次。

尾巴变长了，社会影响增加了不平等。如果社会影响只会导致人们下载更好的歌曲，那么这种不平等也不会造成什么问题。但事实上，这两个实验组的下载量之间的相关性并不强。我们可以将第一个实验组中每一首歌的下载次数解释为歌曲质量的一个表征，那么这项研究表明，社会影响并没有导致人们去下载更好的歌曲。大赢家的出现不是随机的，但它们其实并不一定是最好的。¹³

当然，我们必须非常小心：不能从一项研究中就得出太强的推论。然而，我们确实可以推断，卖出了5 000万册书的畅销书作家、学术论文得到了20万次引用的科学家当然是值得赞许的，但是这种极端的成功本身就表明中心极限定理是不成立的。人们不会独立地购买书籍或引用论文。惊人的成功可能意味着正反馈，也许还有一点运气。在本书的最后一章讨论收入不平等的原因时，我们还会回到这些思想上来。¹⁴

灾难

长尾分布还包括灾难性事件：地震、火灾、金融崩溃和交通拥堵。尽管模型无法预测地震，但确实可以深入解释为什么地震的分布会满足幂律。这些相关的知识告诉我们各种强度的地震发生的可能性。我们至少知道会发生什么，尽管不知道什么时候会发生。¹⁵

而且，森林火灾模型已经可以指导行动了。人们可以通过选择性地在森林中采伐一些树木来降低树木的密度，以防止大火灾的发生，也可以制造防火带。有人会说，在模型告诉我们应该这样做之前，我们早就懂得采伐树木或建造防火带了。这当然是事实，但重要的是，森林火灾模型能够让我们意识到临界密度的存在。临界密度可能因森林而异，可能取决于树木的类型、盛行风速和地形。这个模型有效地解释了为什么森林会出现自组织临界状态。

我们还可以使用这个模型来做一个很好的类比。请回想一下，第1章中讨论了席卷整个体系的金融机构的破产，我们可以将森林火灾模型应用到那种情况下：把银行和其他金融机构想象为网格上的树，网格上的邻接则表示存在未偿还的贷款。一个银行破产相当于一棵树着火，而火势有可能会蔓延到邻近的银行。

当银行的“密度”变得越来越高时，这种看似浅显的森林火灾模型就预示着大规模的银行破产随时可能发生。不过，在深入探析这个类

比时，我们可以发现它存在四个方面的缺点。第一，金融机构的网络并未嵌入物理空间，各家银行的连接数也不相同，有的银行可能拥有几十项金融债务，而有些银行则可能只有一两项金融债务。第二，森林中的树木不能主动采取行动来减少火势蔓延的可能性，但是银行却可以，它们可以提高自己的储备水平。

第三，一家银行拥有的连接越多，其破产会产生连锁反应的可能性就越低，因为它的损失已经分散到了更多的银行身上。例如，如果一家银行只从另一家银行借款，那么如果它在借来的1亿美元的贷款上出现了违约，第二家银行可能会破产。但是，如果第一家银行是从其他25家银行分别借款的，那么任何一家银行都不至于受到重创。在这种情况下，银行体系可以很好地消化这个违约事件而不会崩溃。¹⁶

第四，从一家银行的破产到另一家银行的破产，这种蔓延会不会出现还取决于银行的投资组合。如果两家“相连”的银行拥有相似的投资组合，那么当一家银行破产时，另一家银行也可能早就脆弱不堪了，这时银行破产蔓延的可能性就很大。如果整个网络中的所有银行都拥有相同的投资组合，那么最糟糕的情况就很可能出现。在这种情况下，当一家银行破产时，就可能会出现普遍的银行破产。¹⁷但是，如果每家银行分别持有不同的投资组合，那么一家银行表现不佳并不意味着其他银行也表现不佳。在这种情况下，银行破产就可能不会蔓延。因此，一个模型要想真正有用，就必须考虑到各种不同的投资组合。如果没有这些信息，那么即便知道哪些银行对其他银行负有未偿还债务也不足以预测或防止银行破产，而且银行之间的高互连性的净效应也是不明确的。

波动性

最后，我们讨论最微妙的一个含义。如果组成幂律分布的实体规模出现了波动，那么幂律的指数就可以作为衡量系统层面波动性的一个代表。由此可以推断，企业规模的分布应该会影响市场波动性。例如，我

们可以将某个国家的国内生产总值视为数千家企业的总产量。如果各家企业的生产水平相互独立且变差有限，那么根据中心极限定理，这个国家的国内生产总值分布将服从正态分布。也就是说，企业生产水平的差异越大，总体波动性就越大。如果企业规模的长尾分布导致生产水平上更大的变差，那么这种长尾分布也必定与更大的总体波动性相关。

对美国波动性模式的实证研究表明，波动性在20世纪70年代和80年代有所上升，然后在接下来的20年间又下降了，有人将后面这20年称为“大稳健”（Great Moderation）。¹⁸但是，从2000年前后开始，波动性再次上升。研究显示，可以通过企业规模分布的变化来解释这种波动性演变的模式。¹⁹随着企业规模分布的尾部变得越来越长（越来越短），最大的企业对波动性的影响越来越大（越来越小）。换句话说，总体波动性会随企业规模分布的尾部变长（变短）而增加（减少）。1995年，当总体波动性较低时，沃尔玛的营业收入为900亿美元，相当于美国国内生产总值的1.2%。到了2016年，沃尔玛的营业收入增加到了4 800亿美元，占国内生产总值中的百分比提高到了2.6%，沃尔玛在美国国内生产总值中所占的份额增加了一倍多。2016年，沃尔玛收入的增加或减少可能导致的总体波动性因而也增加了一倍多。

没有人能够反驳这个观点的逻辑。因此，相关的问题是，一个经过校准的模型到底能不能生成与实际波动水平相对应的效应。校准拟合结果表明，确实非常接近。企业规模的分布很好地对应于大稳健时期的历史证据。虽然这种相关性并不能证明是企业规模分布的变化（而不是政府对经济的有效管理或更好的库存控制）导致了这种变化，但是它确实足以“阻止”我们拒绝这个模型。²⁰这些证据还为我们在未来评估各种波动时将这种模型加入我们的工具箱提供了很好的理由。

设想长尾分布的世界

在长尾分布中，大事件发生的概率必须加以考虑。在本书讨论的多

个模型中，长尾分布是由于反馈和相互依赖性而产生的。我们应该高度注意这个结果。随着世界中相互联系性的提高和反馈的增加，我们应该会观察到更多的长尾分布，同时现在关注的这些长尾分布的尾部也可能会进一步拉长。这就是说，不平等可能会增加，灾难可能变得更大，波动性也会变得更加剧烈。这些都是不可取的。

到目前为止，我们都是在宏观层面上讨论这些事件的可能性的。它们也同样可能发生在更小的尺度上。波士顿的中央隧道工程（Big Dig）是一条穿过市中心的长达5 000多米的隧道，它是一个中等规模的灾难的典型例子。这个项目花费了140亿美元（相当于最初预算的3倍多），并成了美国有史以来最昂贵的公路项目。根据模型思维的方法，我们不会把这个项目简单地视为一个单独的项目，而是作为很多子项目的总和：挖掘深沟、浇筑混凝土隧道、设计排水系统、建造墙壁和“顶盖”。项目的总成本等于各个子项目成本的总和。

如果每个子项目的成本都是相加的，那么这个项目的成本分布将是正态分布。²¹然而，各个子项目的成本是相互关联的。原本计划用来将顶盖黏合到位的那种环氧树脂强度不够时，就不得不用成本更高、强度更大的另一种环氧树脂来代替，从而增加了项目的成本。而且，第一种环氧树脂的失效还产生了移除和更换折叠顶盖的额外成本。这些工作反过来又需要重做项目的其他几个部分。于是总体成本增加了一倍以上，因为每个项目必须撤销然后重做。这种相互依赖性最终导致了一个大型且昂贵的事件。

大事件发生的可能性使计划变得非常困难。像地震这样自然灾害的分布符合幂律分布。因此，大多数事件都是很小的事件，但是有些事件一旦发生就会很大。如果灾难性事件遵循指数为2的幂律分布，那么政府就必须时刻保留大量的储备金或者至少做好应对的准备，必须未雨绸缪。如果政府为了这个目的而在应急基金账户中保持了巨额盈余，那么如果没有大事件发生，政府可能会阻止自己花掉这笔钱或减税。

搜索与机会

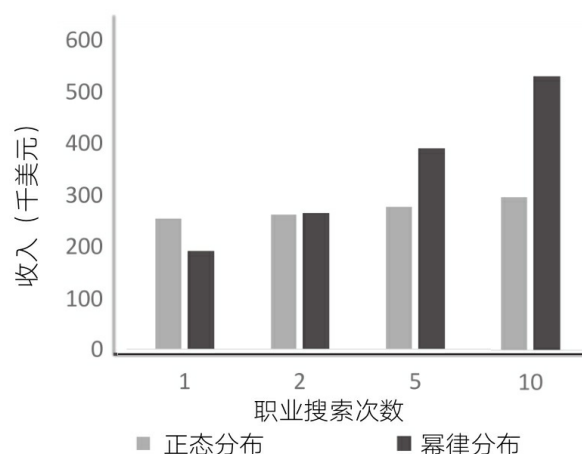
我们可以在某些搜索模型中应用关于分布的知识来解释为什么一个人获得机会的数量可能与他的成功经历密切相关。在这里，事实上是将一类模型（分布模型）嵌入了另一类模型（搜索模型）。我们在搜索的时候，无论是搜索新鞋、工作职位还是度假胜地，其实是不知道所选择的价值的，直到去真的尝试它。不过，我们可能会对所选择的价值的分布有所了解，例如它的均值、标准差，以及这种分布是正态分布还是长尾分布等。

在这里，我们将职业选择建模为一个搜索过程。给定某个行业，一个人尝试某条职业道路。我们将这种行为建模为从一个分布中抽取某个事件。假设，这个人可以坚持这个职业选择或再试一次，再试一次对应于从分布中的另一次抽取。例如，考虑一个有才华的年轻科学家的职业选择。她可以选择去医学院深造，也可以选择去研究量子计算。医学院提供了一条更安全的道路，选择研究量子计算则可能成为一名创业企业家并承担更多风险。为了解释这些差异，我们将医生的工资分布表示为均值25万美元，标准差25 000美元的正态分布，并把量子计算企业家的工资分布表示为指数为3、期望工资为20万美元的幂律分布。²²

再假设，在每个行业内，这位科学家也可以尝试多种职业。也就是说，她可以搜索。医生可以从肿瘤科转入放射科，一个企业家破产后也可以重整旗鼓继续尝试创业。但是每一次转换职业都要付出一定的成本：对于一个医生来说，这意味着要接受更多的培训；对于一个从事量子计算的企业家来说，这意味着需要付出更多的时间去从事没有报酬的工作。

另外再假设，这位科学家认为这两个职业同样有意思，并且会根据薪资水平来做出选择。我们的模型证明，哪种选择更好取决于

有多少次尝试新职业的机会。如果她必须坚持自己的第一个职业选择，那么成为一名医生就可以获得更高的期望工资；如果她有足够的资源持续尝试，努力成为一名企业家，那么最终她将从长尾中获得高薪。假设在每个职业中分别进行1次、2次、5次和10次职业搜索，下图显示在20次测试中得到的平均最高工资。如果这位科学家有机会在量子计算初创企业中尝试10次，那么她的薪资将会是她选择进入医学院深造并尝试10个职业后收入的两倍。



如果拥有的财富和家庭的支持与一个人不得不尝试新职业的机会数量相关，那么模型的预测是，更富有的人将选择风险较高的职业。²³专利证据与我们的模型是一致的。一个人成功申请专利的可能性与他的数学技能相关，数学能力排名前1%的人更容易获得专利。而且在这数学能力排名前1%的人当中，收入排名前10%的家庭的人更有可能拥有专利。²⁴至少有两个模型可以解释这种差异：一个模型假设更贫穷且有才华的学生没有上大学的机会，他们可能正在从事日常工作，从未有机会在进入医学院深造还是去研究量子计算之间做出选择，另一个模型则假设更贫穷的学生会选择更安全的职业。

机会的增加可以创造风险激励，这个逻辑可以应用到很多领域。风险资本家经常冒险，因为他们有机会进行多项投资。只要投

中了一个独角兽（市值10亿美元以上的公司），不仅可以补偿多次失败的投资，还可以带来很大的利润。研究药物的实验室也愿意承担风险，花费数十亿美元用于药物开发。甚至在决定午餐吃什么时，我们也可以应用同样的逻辑。长途旅行并在某个不熟悉的小镇短暂停留时，我们一般更喜欢选择连锁餐厅用餐；但是，如果真的搬到那个小镇去居住，我们就会尝试多家餐厅。

07 线性模型

是的，我承认我在说谎。但为什么你非要强迫我给出一个线性解释呢！线性解释几乎总是谎言。

埃莱娜·费兰特（Elena Ferrante）

模型通常假定变量之间存在某种特定的函数关系。这种关系可以是线性的，也可以是非线性的，或者可以包括阈值效应。在这些模型中，线性模型是最简单且应用最广泛的。本章的重点就是线性模型。教育对收入的影响、因锻炼而增加的期望寿命，以及收入对选民投票率的影响，都可以用线性模型来解释。

在本章的开头部分，先回顾一下单变量线性函数。然后讲解了如何通过回归将数据与线性函数拟合，并揭示各种效应的符号、大小和显著性。我们还讨论了为什么误差、噪声和异质性意味着数据不会全部落在回归线上。接着，我们扩展了线性模型以容纳更多的变量，并讨论了如何拟合多元线性模型。为了建立多变量模型的直觉，我们将成功建模为技能和运气的线性函数。本章的结尾部分总结了如何依靠数据和回归指导行动、减少错误，但是这样做也可能会导致边际行为，进而导致保守的行为。确实，“唯大系数论”思维可能会扼杀创新。为了确定更多的创新项，我们可能需要考虑构建其他更具推测性的模型。

线性模型

在线性关系中，由于第二个变量的变化而导致的第一个变量的变化量不依赖于第二个变量的值。假设树木的高度与树木的年龄呈线性关系，那么树木每年生长的高度相同。假设房子的价值随它的面积（平方

米)线性增加,那么房子面积扩大200平方米所带来的房子价值的增量,等于房子面积扩大100平方米所带来的房子价值增量的两倍,400平方米的扩大使房子的价值增加了4倍。

线性模型

在线性模型中,自变量 x 的变化,会导致因变量 y 的线性变化,用如下方程表示:

$$y=mx+b$$

其中, m 等于直线的斜率, b 等于截距,即当自变量等于0时的因变量值。

线性回归模型的目标是找到能够最小化到各数据点的直线。线性回归可以解释犯罪、洗衣机销量,甚至可以解释葡萄酒价格的变化。¹假设我们找到了一组年龄介于20岁到60岁之间的成人的年龄数据以及他们每个星期走路的距离,可以发现如下回归方程:

$$\text{第}i\text{个人步行的英里数}=-0.1\times\text{年龄}_i+12+\varepsilon_i$$

这个回归方程不仅告诉我们这种效应的符号(距离随年龄的增长而减少),还告诉我们这种效应影响的大小(年龄每增加1岁,距离减少1/10英里)。在这个例子中,截距并不重要,因为它位于数据范围之外,也就是说,数据原本就不包括年龄接近于0的人。根据这个方程,我们可以预测,一个40岁的人每个星期步行8英里,而50岁的人则每个星期步行7英里。但是,用于产生回归的数据不会全部都落在回归线上。

图7-1显示了用于生成回归线的假想数据。其中灰色圆圈代表的人名叫博比,他40岁了,每个星期步行11英里,比模型估计的要多走3英里。为了使数据与模型一致,我们在方程中给每个数据点增加了一个误

差项，用 ϵ 表示，等于模型估计值与因变量实际值之间的差异。博比的误差项等于+3英里。

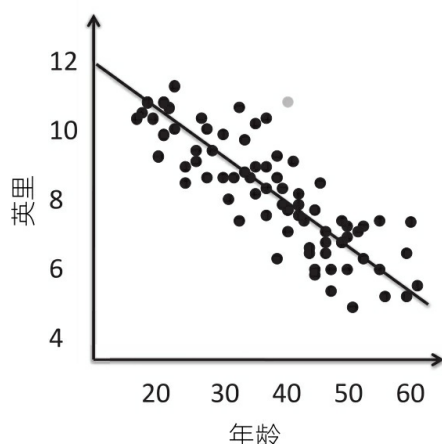


图7-1 散点图和回归线

在社会环境和生物环境中，我们不能期待会有完美的线性拟合。结果通常取决于多个变量，但是根据定义，单变量回归只能包含一个变量。由于这些被省略的变量，预测值可能会偏离实际值。博比走的路可能要比预期更多，因为作为植物学教授，他要带他的学生到树林里采集标本。这个模型没有把职业作为一个变量，但是职业有助于解释为什么图7-1中的数据点没有落在回归线上。 ϵ 项也可能由测量误差引起。如果人们忘记随身携带智能手机，或者将智能手机借给他人，那么利用智能手机收集的健身数据就会包含误差。此外，环境噪音也可能导致误差，比如人们可能会因为坐在颠簸的汽车上而获得额外的步行里程。²

回归线越靠近数据，模型解释的数据越多， R^2 就越大（得到解释的百分比越大）。如果数据全部都恰好位于回归线上， R^2 就等于100%。

符号、显著性和大小

线性回归可以告诉我们关于自变量系数的如下内容：

符号：自变量与因变量之间的正相关或负相关。

显著性（ p 值）：系数上非零符号的概率。

大小：对自变量系数的最佳估计。

在单变量回归中，回归线与数据拟合得越好，我们对系数的符号和大小就越有信心。统计学家使用 p 值来表示系数的显著性， p 值等于基于回归的系数不为零的概率。 p 值为5%意味着数据由一个系数等于零的过程生成的概率为1/20。显著性的标准阈值是5%（通常用*表示）和1%（通常用**表示）。但是，显著性并不是我们唯一关心的东西。一个系数可能是显著的，但是却很小。如果真的是这样，就可以对相关关系很有信心，但是变量的影响其实不大。又或者，也可能系数虽然不显著但却很大，这通常发生在有噪声数据或数据带有许多遗漏变量的情况下。

为了阐明如何利用回归来指导行动，不妨想象一下这样一家销售香料的公司。该公司供应超过100种香料。客户会购买包含6种、12种或24种香料的包装。客户下单后，员工负责包装和运输。将每8小时的班次的订单数量作为员工工作年限的函数进行回归，结果如下：

$$\text{完成的订单数} = 200 + 20^{**} \times \text{工作年数}$$

在上面的方程中，工作年数前面的系数20的显著性水平为1%。我们可以确信它是正的。如果这种关系是因果关系，那么这个模型就可以用来预测每个员工每个班次可以完成的订单数量（作为工作年数的函数），还可以使用这个模型来预测某个在职员工明年可以完成的订单数量。在这里，有一个模型的实例，既可以给出预测，也可以指导行动。

相关关系vs. 因果关系

回归所揭示的是变量之间的相关关系，而不是因果关系。³如果先构建了某个模型，然后用回归检验模型的结果是否得到数据的支持，但那也不能证明因果关系。但是，在我们能够用回归发现显著的相关性之

前，有一种方法远比回归方法好，这种方法就是通常所称的“数据挖掘”（data mining）。但是，数据挖掘存在识别与其他因果变量相关的某个变量的风险。例如，数据挖掘可能会揭示维生素D的水平与身体总体健康程度之间存在显著的正相关关系。人们多晒太阳有利于吸收维生素D，因此这种效应可以归于生活方式更积极的那些人在户外度过的时间更长，从而健康状况更好。或者回归可能会发现，某个大学的学术表现与参加马术队的学生人数存在显著相关。但是，马术队与学术水平之间可能并不存在直接的因果关系，但它们与平均家庭收入和学校资助水平相关。

数据挖掘还可能导致虚假的相关关系，即两个变量只是偶然相关。我们可能会发现，名字较长的公司可以获得更高的利润，或者居住在比萨店附近的人更容易患流感。事实上，使用5%的显著性水平阈值，每检验20个变量就会发现有一个是显著的。因此，如果尝试足够的变量，肯定会发现某些显著但虚假的相关性。

我们可以通过创建训练集（training set）和检验集（testing set）来避免报告虚假相关。在训练集上发现的相关性，如果也存在于检验集上，就更可能是真实的。但即便是这样，我们仍然无法保证那就是因果关系。为了证明因果关系，还需要进行一个实验来操纵自变量并观察因变量是否会随之发生变化，或者也可以想办法找到可以证明这类因果关系的自然实验。

多元线性模型

大多数现象都有不止一个因果变量和相关变量。一个人的幸福可以归因于身体健康、婚姻美满、子女、宗教信仰和财富等。一栋房子的价值取决于室内面积、庭院大小、浴室数量、卧室数量、建筑类型以及当地学校的质量等。在解释房子价值的时候，可以把所有这些变量都包含在回归中。但是必须记住，随着添加更多的变量，也就需要更多的数

据，不然无法得到显著的系数。

实力-运气方程

在讨论多元回归之前，先引入迈克尔·莫布森（Michael Mauboussin）的实力-运气方程，以便对多元方程有一个直观的认识。⁴这个方程说的是，任何成功，无论是日常工作中的成功、体育运动上的成功，还是游戏时的成功，都可以视为实力-运气的一个加权线性函数。

实力-运气方程

$$\text{成功} = a \times \text{实力} + (1-a) \times \text{运气}$$

其中， a 位于区间 $[0, 1]$ 上，是技能的相对权重。

如果给实力和运气分配适当的权重（也许通过利用现有数据进行回归，可以得到这样的权重），我们就能够运用这个模型来预测结果。例如，假设一家休闲汽车销售公司的经理发现，用销售数量来衡量的成功有很大的运气成分，那么他就会期待回归均值：本月取得了很好业绩的销售人员下个月可能会回到平均水平。然后，这个经理就可以利用这个模型来指导行动了。比如，他不会为了争取一个连续两个月都取得非常不错业绩的销售人员而付出比竞争对手高很多的薪资。相反，如果回归表明运气对成功几乎没有任何作用，那么连续两个月的业绩就可以作为未来业绩表现的一个很好的预测器。在这种情况下，经理就应该为这个最佳销售人员提供有竞争力的报酬。⁵

同样的逻辑也适用于CEO薪酬的决定。在那些“运气决定了成功”的行业中，董事会不应该向CEO发放高额奖金。石油公司的利润取决于原油的市场价格，那是一个公司无法控制的变量。因此，一家石油公司的

董事会不应该因为某一年公司业绩不错就给CEO发放巨额奖金。相反，广告公司则不然：如果广告公司业绩表现良好，那么给CEO发放巨额奖金就是一件明智的事情。简而言之，要奖励实力，而不要为运气去买单。事实上，那些很成功的公司都不会为运气付出太多。

即便是最简单的模型，例如上面这个实力-运气方程，也能帮助我们得出深刻的见解。进一步思考这个方程可以发现，即便是在那些成功几乎完全取决于实力的环境中，例如跑步、骑自行车、游泳、下棋或网球比赛，如果不同的参赛者之间实力差异很小，那么运气就会在很大程度上决定谁输谁赢。我们可以预期，在竞争最激烈的比赛中，比如奥运会，进入决赛的选手之间的实力差异很小，因而运气就非常重要了。莫布森把这种情况称为“实力悖论”（paradox of skill）。

历史上最伟大的运动员之一迈克尔·菲尔普斯（Micheal Phelps）可以说同时位于这个悖论的两端。在2008年奥运会的一场决赛中，菲尔普斯在100米蝶泳快结束时仍然落后于米洛拉德·卡维奇（Milorad Cavic）。然而幸运女神眷顾了他，菲尔普斯率先触到了池壁。然而，在2012年奥运会的一场决赛中，菲尔普斯一直领先于查德·勒·克洛斯（Chad le Clos），但是幸运女神这次没有眷顾他，勒·克洛斯率先触到了池壁。菲尔普斯拥有令人难以置信的实力，但是上一次胜利和这一次失败，却都是运气的产物。

多元线性回归

多元线性回归模型拟合了具有多变量的线性方程，当然同样要最小化到数据的总距离。这些方程包括每个自变量的系数。下面的方程反映了这样一个假设的回归输出：学生在数学考试中的成绩，是学生学习的小时数（HRS）、学生家庭社会经济状况（SES）和上“快班”课程的数量（AC）的函数。

$$\text{数学成绩} = 21.1 + 9.2^{**} \times \text{HRS} + 0.8 \times \text{SES} + 6.9^{*} \times \text{AC}$$

根据回归分析的结果，学生每多学习一个小时，数学成绩会提高9.2分。这个系数有两个*号，因此它在1%的水平上显著，这意味着很强的相关性，尽管不是因果关系。这个方程也表明，每参加一个“快班”课程，数学成绩能够提高近7分，这个系数也是显著的，但仅仅在5%的水平上显著。家庭社会经济状况这个变量的取值为从1（低）到5（高），系数也为正，但是与零没有显著差异，因此我们可以认为它可能没有什么因果关系。

有了这样一个（或任何形式的）回归方程，我们就可以预测结果。这个模型预测，如果花7个小时学习，并同时参加一个“快班”课程，数学成绩就能够达到90分左右。这个模型还可以用来指导行动，但必须保持谨慎，因为我们无法推断因果关系。数据表明，花时间学习和参加“快班”课程的学生成绩更好。但是，花时间学习和参加“快班”课程这两个因素也可能没有什么用，因为也许存在选择性偏差（selection bias），那些花更多时间学习、参加“快班”课程的学生，数学成绩可能本来就更好。

即便回归不能说明是什么原因导致数据呈现出来的特定模式，但是至少可以排除其他解释。以美国种族之间的巨大财富差距为例：2016年，白人家庭的平均财富（约11万美元）是非洲裔美国人家庭和拉美裔美国人家庭的10倍。各种各样的原因都可以用来解释这种差距，包括制度因素、收入差距、储蓄行为差异或结婚率差距等。回归可以为其中一些解释提供支持并排除其他解释。例如，回归分析表明，非洲裔美国人的婚姻状况与家庭财富之间没有显著关系，因此婚姻状况不能成为这种财富差异的原因。此外，收入差距虽然相当大，但是也不足以解释这种财富差距。⁶

大系数与新现实

如前所述，线性回归模型在科学研究、政策分析和战略决策中都发挥着重要作用，部分原因是因为线性回归模型容易估计和解释。而且，随着数据可得性的不断改善，线性回归模型得到了更广泛的应用。“要信只信上帝，要认只认数据”（In God we trust. Everyone else must bring data.）这句话在商界和政界都可以经常听到。对数据的这种依赖（通常意味着线性回归模型），可能会导致我们过于倾向边际行动（marginal action），远离重要的新思想。企业、政府或基金会，都致力于收集数据，拟合线性回归模型，试图找到有最高统计显著性系数的变量，这种努力几乎肯定会导致调整该变量并获得边际收益的行为。

在采取行动的时候，最好选择具有较大系数的变量，而不要选择具有较小系数的变量。与此同时，“大系数至上”这个思路建立在“保守主义”的基础上，它会使我们将注意力集中到较小的改进上，而无法再关注全新的政策。“大系数至上”的另一个问题是大系数的大小对应于给定现有数据的边际效应。正如我们在下一章中会阐述的，通常这种效应将会随着变量值的增大而减少。如果确实是这样，那么当我们试图利用它时，大系数就会变小。

大系数与新现实

线性回归揭示了自变量与我们感兴趣的（因）变量之间的相关程度。如果这种相关是因果关系，那么具有大系数变量的变化就会产生很大的影响。基于大系数的政策在保证能够带来改进的同时，排除了涉及更多根本性变化的新现实。

“大系数至上”思维方式的替代者是“新现实思维”。如果说，大系数思维可以拓宽道路、建造高利用率的车道以减少交通拥堵，那么新现实思维就相当于建造了铁路和公共汽车系统。大系数思维为低收入家庭的学生购买计算机提供补贴，新现实思维则直接为每个人都提供了计算机。大系数思维改变了飞机上座位的宽度，新现实思维则创造了一个使

用可互换吊舱的飞机机舱。大系数思维已经相当不错了，因为基于证据的行为是明智的，但我们也必须同时关注重要的新思想。当我们遇到重要的新思想时，可以用模型去探究它们是否可行。对青少年交通事故的回归也许会告诉我们，年龄的系数是最大的，这意味着提高驾驶年龄的政策也许能起到一定作用。还可以采取更多的政策，例如禁止夜间驾驶的宵禁、通过智能手机自动监控青少年驾驶员或限制青少年驾驶汽车的乘客数量等。这些新现实政策带来的效果可能比大系数带来的要好。

小结

总而言之，线性模型需要假定效应大小不变。线性回归为我们对数据进行第一轮加工提供了一个强大的工具，有了它，我们能够识别出变量的符号、大小和显著性。如果我们希望了解咖啡、酒精或苏打水对健康的影响，就可以进行回归分析。我们可能会发现，喝咖啡会降低心血管疾病的风险，适量饮酒也有同样的效果。这也就是说，在现有数据范围之外推断线性效应时必须非常小心。我们绝对不能推断，喝30杯咖啡、6瓶葡萄酒会是个好主意。我们不应该用线性模型对过于久远的未来进行预测。从1880年到1960年，加利福尼亚州的人口增长率为45%，如果进行线性预测，那么我们将会预测2018年加利福尼亚州的人口会达到1亿人，但这超出了实际人口水平的两倍。

请记住，线性模型只是一个开始，大多数有趣的现象都不是线性的。因此，回归模型通常会包括非线性项，例如年龄的平方、年龄的平方根，甚至包括年龄的对数。为了解释非线性，我们还可以将线性模型首尾相连，这些连接起来的线性模型可能近似于曲线，就像我们可以使用直边的砖块来砌出弯曲的路径一样。虽然线性可能是一个强大而不切实际的假设，但是它至少提供了一个很好的起点。在给定了数据的情况下，可以使用线性模型来检验我们的直觉判断。然后，我们可以构建更精细的模型，其中变量的影响会随着它的增加（收益递减）或变得更强

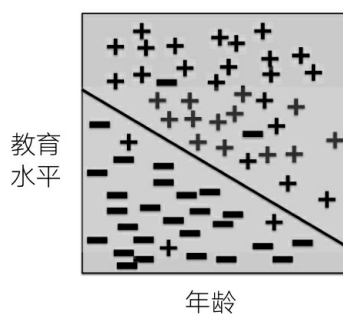
大（正回报）而减弱。这些非线性模型正是下一章要研究的重点。

对数据的二元分类

在当今这个大数据时代，组织普遍使用根据模型建立的算法对数据进行分类。政党可能想要了解哪些人投了谁的票，航空公司可能想知道常客的特点，某项活动的组织者可能想要了解哪些人会参加这项活动。在所有这些情况下，它们所使用的方法都将相关的人分成了两组：一组是“正”的（+），也就是购买了、贡献了、注册了的人；另一组是“负”的（-）。

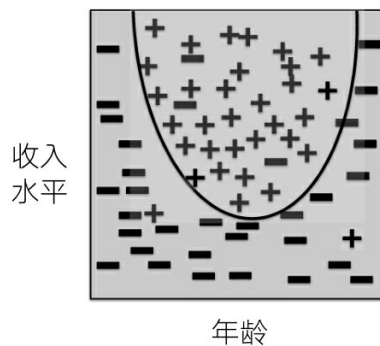
分类模型应用算法根据人们的年龄、收入、教育水平或在互联网上花费的时间等性质，将人划分为不同的类别。不同的算法意味着不同的属性与结果之间关系的基础模型。应用多种算法，也就是使用许多模型，能够产生更好的分类。

线性分类：在图a中，“正”（+）代表参加投票的人，“负”（-）代表没有参加投票的人。在此基础上，可以用一个反映人们年龄与教育水平的线性函数来对某个人是否会参加投票进行分类。数据表明，受过更多教育的人更有可能去投票，同时年龄大的人也更有可能去投票。在这个例子中，图中的直线近乎完美地实现了分类。⁷



图a 用线性模型对投票行为进行分类

非线性分类：在图b中，“正”（+）代表航空公司的常客（每年飞行超过1万英里的旅客），“负”（-）代表航空公司的所有其他旅客。中年人和收入更高的人更有可能乘坐飞机旅行。要对这些数据进行分类，需要先利用某个深度学习算法（如神经网络算法）找到一个非线性模型。神经网络模型包含多个变量，因此它们几乎可以拟合任何曲线。



图b 用非线性模型对航空公司的常客进行分类

决策树森林：在图c中，“正”（+）表示参加科幻大会的人，基于他们的年龄和每个星期花在互联网上的小时数。在这里，我们使用了三棵决策树对数据进行分类。决策树根据各种属性不同的条件组合进行分类。图中显示的三棵决策树分别为：

决策树1：如果年龄 <30 岁

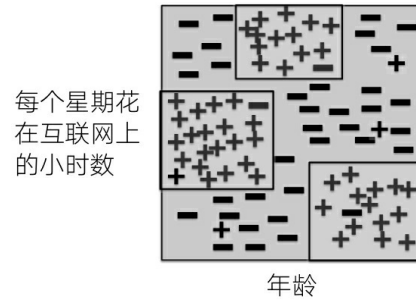
且每个星期花在互联网上的小时数介于 $[15, 25]$ 岁之间

决策树2：如果年龄介于 $[20, 45]$ 岁之间

且每个星期花在互联网上的小时数 >30

决策树3：如果年龄 >40 岁

且每个星期花在互联网上的小时数 <20



图c 用决策树森林对会议参与者进行分类

树木的集合称为森林。机器学习算法会在一个训练集上随机构造出树，然后将那些在检验集上准确分类的树保存下来。

08 非线性模型

讨论非线性科学，就类似于讨论无大象的动物学。

约翰·冯·诺伊曼

在本章中，我们介绍非线性模型和非线性函数。非线性函数可以向下或向上弯曲，可以形成S形，还可以扭结、跳跃和波动。在下文中，我们将会讨论到所有这些可能性。现在先从依赖于凸性和凹性的模型开始讨论。我们阐明了增长和正反馈是如何产生凸性的，收益递减和负反馈又是如何产生凹性的。在绝大多数学科中，都包含了这两类模型。

关于生产的经济学模型假设交货期和库存成本会随着企业规模的增大而减少，从而使每单位产品的销售利润成了企业规模的一个凸函数，这也就解释了为什么沃尔玛能够获得如此高的利润。¹关于消费的经济学模型则假设效用（或价值）是凹的，也就是说，第5块比萨带给我们的享受比第1块比萨小。在一个生态系统中，当一个新物种入侵并无须面对任何天敌时，其“人口”会以恒定的速率增长，这就产生了一个凸函数。但是随着“人口”的增长，它们的食物就会减少。因此，作为种群规模函数的适合度（fitness）是凹的。

本章由三个部分组成。第一部分讨论凸函数，包括了人口增长和衰退模型。第二部分讨论凹函数。在这一部分中，我们将会看到凹性意味着风险规避和对多样性的偏好。在第三部分中，我们研究了一系列经济学中的增长模型，它们结合了凹函数和线性函数。

凸函数

凸函数的斜率是递增的：函数值随度量值的增加而增加。例如，在

一个人群中，可能结成的“对”的数量是这个群体人数的凸函数。一组3人，可以结成3个不同的“对”；一组4人，可以结成6个不同的“对”；一组5人，则可以结成10个不同的“对”。群体规模每增大一些，都会导致“对”的数量有更大的增加。与此类似，每一次，当厨师增加了一种新的香料时，他可以使用的香料组合数量就会增加很多。

我们要讨论的第一个凸函数模型是指数增长模型（exponential growth model），它描述的是一个变量的数量（通常是指人口或资源）与它的初始值、增长率和周期数之间的函数关系。

指数增长模型

时间 t 的资源值 V_t ，其初始值为 V_0 ，且以速率 R 增长，可以写成如下方程：

$$V_t = V_0 (1+R)^t$$

这个单方程模型在金融、经济、人口、生态以及技术等领域中都发挥着核心作用。当我们把它应用于金融问题时，这里的变量就是货币。利用这个方程，我们可以计算出，年利率为5%时1 000美元债券在一年后的价值会增加50美元，而到第20年将增加100多美元。为了得出清晰的推论，我们假设增长率固定不变。根据这个假设，可以利用指数增长方程推导出72法则（Rule of 72）。

72法则

如果一个变量在每个周期内以 R （增长率小于15%）的百分比增长，那么下面提供了一个很好的近似：

$$\text{翻倍所需的周期数} \approx \frac{72}{R}$$

72法则量化了最高增长率的累积效应。1966年，津巴布韦的人均国内生产总值为2 000美元，是博茨瓦纳的两倍。但是在接下来的36年里，津巴布韦几乎没有增长，而博茨瓦纳的年平均增长率则达到了6%，这意味着博茨瓦纳的人均国内生产总值每12年就会翻一番。于是在36年中，博茨瓦纳的人均国内生产总值翻了三番（增加了8倍）。因此，到了2004年，博茨瓦纳的人均国内生产总值达到了8 000美元，相当于津巴布韦的4倍。

同样是这个公式，还揭示了为什么房地产泡沫必定会结束而技术进步则不会。2002年，美国的房价上涨了10%。这个增长率意味着每7年翻一番。如果这种趋势一直持续35年，那么美国的房价将会翻五番，即增长32倍。这也就是说，一栋在2002年价格为20万美元的房屋在2037年将上涨到640万美元。当然，价格不可能一直这么涨下去，泡沫必定会破灭。与此不同，摩尔定律（Moore's law）则指出，可以安装在一块集成电路上的晶体管数量每两年会增加一倍。摩尔定律之所以持续存在，是因为用于研发的投入带来了近乎恒定不变的进步速度。

人口学家则用指数增长模型研究人口问题。如果每年增长6%，那么人口在12年内就会翻一番，在36年内会翻三番，在100年内则会翻八番，即增长256倍。早在1798年，英国政治经济学家、人口学家托马斯·马尔萨斯（Thomas Malthus）就观察到了人口数量呈指数增长的现象，并在给出的一个模型中指出，如果经济体生产粮食的能力是呈线性增长的，就会出现粮食危机。短期的变化如下：人口增长模式为1、2、4、8、16、32……；而粮食生产的增长模式则为：1、2、3、4、5……。马尔萨斯预测，灾难很快会发生。幸运的是，出生率不久之后就下降了，工业革命的到来也极大地提高了生产率。如果这两件事情都没有发生，那么马尔萨斯的预测应该是正确的。关键是，马尔萨斯忽视了创新的潜力。本章下面将要给出的模型重点就是创新，它颠覆了马尔萨斯担心的趋势。

指数增长模型也可以用于研究物种的增长，当然不仅仅适用于兔子。当你受到细菌感染时，那些肉眼不可见的细菌会以极高的速度繁殖。人类鼻窦中的细菌每分钟都在以4%的速度增加。应用72法则，我们可以计算它们每20分钟就会翻一番。一天之内，每个初始细菌都会繁衍出超过10亿的后代。²当然，由于鼻窦的物理空间有限，它们不可能一直繁殖，没有空间时增长就会停止。食物的限制、天敌的存在、生存空间的缺乏，都会减缓增长。有些物种，例如生活在美国郊区的鹿，或者被毒梟巴勃罗·埃斯科巴（Pablo Escobar）带入哥伦比亚的河马，虽然繁殖速度远远不如细菌，但是由于受到的限制很少，它们的种群迅速增大。³

具有正斜率的凸函数会以递增的值增加，具有负斜率的凸函数就会变得不那么陡峭，也就是说，最初具有较大负斜率的凸函数将逐渐走平。半衰期模型（half-life model）中的方程就是如此，这个模型可以用来刻画分解、折旧和遗忘。

在半衰期模型中，每 H 周期，数量就会衰减一半。因此，我们把 H 称为该过程的半衰期。对于某些物理过程，半衰期是恒定的。所有有机物都包含两种形式的碳：不稳定的同位素碳-14，以及稳定的同位素碳-12。在活的有机物中，这些同位素是以固定比例存在的。当有机体死亡后，体内的碳-14开始分解，其半衰期为5 734年，碳-12的数量则不会改变。美国物理化学家威拉德·利比（Willard Libby）意识到，通过测量碳-14与碳-12的比例，就可以估计化石或人工制品的“年龄”，这种技术被称为放射性碳年代测定法（radiocarbon dating）。现在，古生物学家已经将放射性碳年代测定法应用于测定恐龙、猛犸象和史前鱼类遗骸的年代了。考古学家则用这种方法来判断古生物的真伪。利用这种方法，考古学家估计，在意大利阿尔卑斯山发现的冰人“奥茨”（Ötzi）的遗骸有5 000年的历史。而于1357年首次出现在公众眼前的“都灵裹尸布”则被认定为是14世纪的物品，而不是某些人所声称的用于耶稣基督葬礼的那一块。

半衰期模型

如果每 H 周期，剩余数量的一半会衰减，那么在 t 周期后，剩余的比例为：

$$\text{剩余比例} \approx \left(\frac{1}{2}\right)^{\frac{t}{H}}$$

半衰期模型的一个新应用是在心理学中。早期的心理学研究表明，人们几乎以接近固定不变的速度忘记信息。人们记忆的半衰期取决于事件的显著性。⁴2016年，电影《聚焦》（*Spotlight*）获得了奥斯卡最佳影片奖。假设，人们对奥斯卡获奖记忆的半衰期为两年，那么到了2018年，有1/4的人会记住这一事实；但是到了2026年，将只有1/1024的人还会记得这件事情。但对任何特定事件的回忆因人而异，对《聚焦》的导演汤姆·麦卡锡（Tom McCarthy）来说，他可能永远不会忘记他是哪一年获得奥斯卡奖的。

凹函数

凹函数与凸函数相反。凹函数的斜率是递减的。具有正斜率的凹函数会呈现收益递减的特点：当我们拥有的东西越来越多的时候，每个额外东西所能带来的价值会越来越少。几乎所有商品的效用或价值都呈递减趋势。闲暇越多、金钱越多、冰激凌越多，甚至与爱人共度的时光越多，对我们的价值就越小。一个直观的证据源于如下事实：包括巧克力在内，对任何事物的消费越多，我们就会越不觉得享受，同时愿意为它付出的代价也就越少。⁵

收益递减可以解释很多现象，包括为什么异地恋往往能够带来很大的幸福感。如果你每月只能与你的伴侣相聚几个小时，那么每多一分钟都是一个莫大的惊喜。而在一个月不间断的相处后，幸福曲线的斜率就

会变平，从而额外增加的相聚时间就变得不那么重要了。⁶同样的逻辑也可以解释为什么房地产开发商喜欢邀请人们在周末免费去他们的海滨公寓。在短暂的周末，你无法在海滩上享受足够长的时间，你会很想把房子买下。相反，如果让你在海滩上连续待上十天半个月，你可能就会觉得无聊。

当我们假设了凹性时，也就隐含地假设了对多样性和风险规避的偏好向。要证明前者，只需要给出一个有多个参数的凹函数就可以。如果人们的幸福曲线是凹性的，而且闲暇和金钱都在增加，那么人们就会更偏好休闲和金钱的组合，而不怎么喜欢只有金钱、没有闲暇或只有闲暇、没有金钱。而风险规避则意味着更偏好确定的有把握的事情而不怎么喜欢彩票，也就是不确定的事情。例如，一个厌恶风险的人会更喜欢得到100美元，而不是只有一半机会得到200美元，另一半机会什么也得不到。一个厌恶风险的人也更喜欢双层冰激凌甜筒，而不怎么喜欢要么没有冰激凌、要么可以得到四层冰激凌。

图8-1说明了为什么凹性就意味着风险规避。这幅图描绘了3种结果的幸福价值（幸福感）：高结果（ H ）、低结果（ L ），以及前两种结果的平均值（ M ）。给定形状向下的曲线，平均结果的幸福感会超过低结果和高结果的平均幸福感。凸函数情况下则相反。凸性意味着风险爱好：我们更喜欢的是极端值，而不是平均值。可以购买的股票数量是其价格的一个凸函数，因此，股票买家更喜欢价格波动。如果价格不断上涨和下跌，买家最终能够获得的股票比价格保持不变时获得的更多。⁷

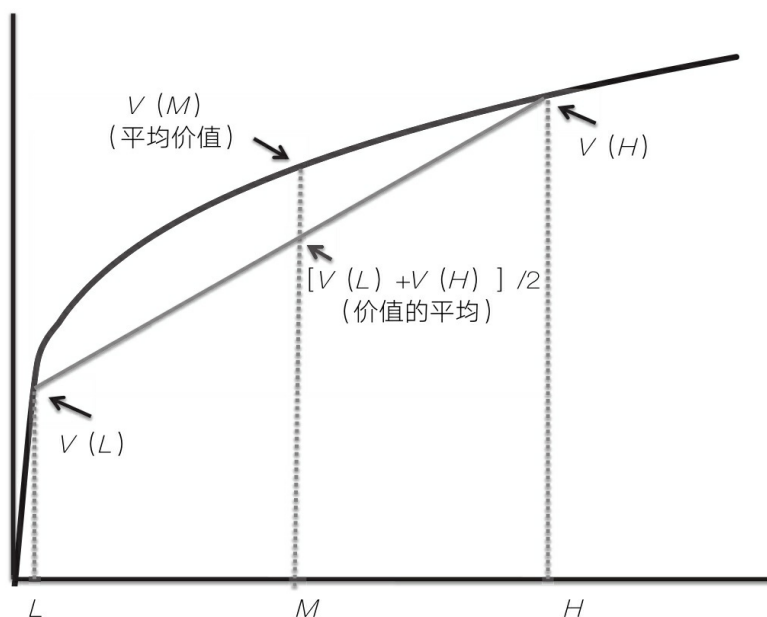


图8-1 风险规避：平均价值>价值的平均

经济增长模型

接下来，我们将构建一系列经济增长模型。这些模型不仅揭示了增长的原因，还可以解释和预测各国的增长模式，还可以指导例如提高储蓄率等的行为。为了给对增长模型的研究奠定基础，我们先引入一个标准的经济生产模型，其中产出取决于劳动和实物资本。经验证据和逻辑都支持产出是劳动力和资本凹函数的假设。保持固定资本，随着投入的劳动力的增加，劳动力的价值应该变得越来越低。同样，在保持工人数量不变的情况下，添加更多的机器或计算机会增加更少的价值。逻辑推理还表明，产出应该是线性的，工人数量和资本总额翻番应该能使产出翻番。这就是说，一家拥有60名工人和一栋厂房的扫帚生产企业，在新建了一栋同样大小的厂房并多雇用了60名工人后，它的产出肯定应该翻番。柯布-道格拉斯模型（Cobb-Douglass model）是经济学中使用最广泛的模型之一，它同时包括了这两种性质。产出是劳动力和资本的凹函数，而且从规模上看是线性的。这个模型既可以应用于单个企业，也可以应用于行业或整个经济生产。⁸

柯布-道格拉斯模型

给定 L 个工人和 K 个单位资本，总产出如下所示：

$$\text{产出} = \text{常数} \times L^a K^{(1-a)}$$

其中 a 是介于0到1之间的实数，表示劳动力的相对重要性。

接下来，我们利用柯布-道格拉斯模型来构建经济增长模型。简化起见，我们假设经济体中有10 000名工人，并暂且不考虑工资和价格，这使我们能够专注于分析机器数量的变化是如何影响总产出的。然后，将资本投资与增长联系起来。为了使模型尽可能简单，再假设只生产一种商品——椰子。椰子含有丰富的椰汁和椰奶，可以作为食物。然而，椰子长在高高的树上，所以工人需要使用某种机器才能把它们摘下来。接下来，再做出一个非常不现实的假设，即机器本身也是用椰子制成的。这样可以简化模型，同时也保持了当前消费与未来投资之间的关键权衡。作为柯布-道格拉斯模型的一个特例，我们将生产函数写为工人数量的平方根乘以机器数量的平方根，即：

$$\text{产出} = \sqrt{\text{工人数量}} \sqrt{\text{机器数量}} = 100 \sqrt{\text{机器数量}}$$

如果经济体中只有一台机器，那么产量等于100吨。如果人们消费掉了所有100吨椰子，就不能投资制造新机器了，从而明年的产量将保持不变，也就是经济没有增长。如果他们投资1吨椰子制造了第2台机器，产量将增加到141吨，增长率为41%。如果他们制造了第3台机器，那么产量将增加到173吨。⁹通过不断地投资，经济增长率却逐渐下降，因此产出是一个凹函数。

我们已经大体知道投资是如何推动增长的。现在可以构建一个包含投资规则的更加精细的模型。假设投资等于储蓄率乘以产出，并假设机器按某个不变的折旧率折旧。例如，到了年底，不能再用的机器数量等于机器总数的某个固定比例。然后就可以得出，下一年的机器数量等于

上一年的机器数量加上新投资的机器数量，再减少因折旧而减少的机器数量。于是，这个完整的简单增长模型由4个方程组成。

简单增长模型

$$\text{产出函数: } O(t) = 100 \sqrt{M(t)}$$

$$\text{投资规则: } I(t) = s \times O(t)$$

$$\text{消费-投资方程: } O(t) = C(t) + I(t)$$

$$\text{投资-折旧方程: } M(t+1) = M(t) + I(t) - d \times M(t)$$

其中， $O(t)$ = 产出， $M(t)$ = 机器， $I(t)$ = 投资， $C(t)$ = 消费， s = 储蓄率， d = 折旧率。

假设这个经济体中有100台机器，储蓄率为20%，折旧率为10%，产量等于1 000吨椰子，消费量等于800吨椰子，新投资200台机器。再假设，因折旧而损失的机器为10台，也就是在新的一年开始时将有290台机器。通过类似的计算可知，在第2年，产出将增长为1 702吨，而第3年的产出则将为近2 500吨。¹⁰由此可见，在这前3年，产出以递增的速度在增长。但是这种凸性只会在前几年出现，原因是机器的初始数量很少，因而折旧几乎完全不会产生任何影响。

随着时间的推移，机器数量的增加和折旧开始变得十分重要。从长远来看，产出的增长将完全停止（图8-2）。只要分析一下模型，就可以找到原因。投资是线性的，因为增加的新机器的数量是随产出呈线性增加的，同时产出则是机器数量的凹函数。因此，随着经济的增长，投资与机器数量的关系也是凹性的。然而，折旧与机器数量之间却是线性关系。最终线性折旧会赶上产出的凹性增长。

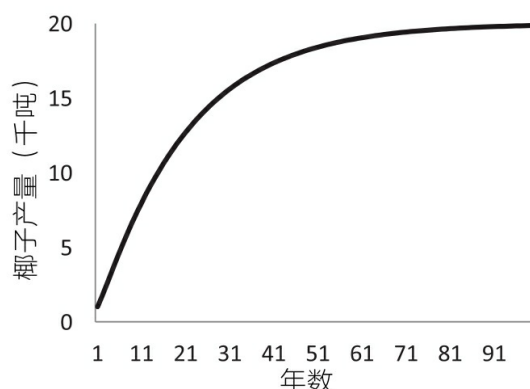


图8-2 简单增长模型中前100年的产出

在经济的长期均衡中，投资的新机器数量等于折旧损失的机器数量。在这个简单增长模型中，当经济体拥有40 000台机器并生产20 000吨椰子时，这种长期均衡就会出现。在这一点上，经济体在新机器上投入了20%的产出或4 000吨椰子，恰恰等于因折旧而损失的机器数量，也就是40 000台机器当中的10%。因此，因折旧而损失的机器数量等于通过投资和停止增长所创造的新机器数量。¹¹

索洛*增长模型

现在构建一个更一般的模型，它是索洛增长模型（Solow Growth Model）的简化，因此我们在索洛后加了一个星号。我们用实物资本取代机器，并将劳动力视为一个变量。此外，还添加了一个技术参数，它可以线性地增加产出。创新会使这个参数增大。与简单增长模型一样，当投资等于折旧时，长期均衡就会出现。不过，在这里，人们认为均衡时的产出水平取决于劳动力数量和技术参数，以及储蓄率和折旧率。¹²

索洛*增长模型

经济体中的总产出由以下方程给出：

$$\text{产出} = A\sqrt{L}\sqrt{K}$$

其中， L 表示劳动量， K 表示实物资本量， A 表示技术水平。长期均衡产出 O^* 由下面的方程给出：[13](#)

$$O^* = A(t)^2 L \frac{s}{d}$$

长期均衡产出随劳动力数量的增加、技术的进步、储蓄率的提高而增加，同时随折旧率的上升而下降。这些结果都不足为奇。更多的工人、更先进的技术和更多的储蓄理应可以增加产出，更快的折旧理应减少产出。但是，产量随着劳动力数量的增加和储蓄率的提高而线性增长这个事实却着实令人惊讶。劳动只能带来递减的收益，因此如果不考虑模型，我们可能会预测长期产出与劳动力数量之间是凹性关系。但是，随着劳动力数量的增加，产出也会增加，投资也会增加，进而带来更高的产出。从长远来看，投资的正反馈恰好抵消了收益递减。均衡产出与折旧率之间的关系是凸性的，折旧率降低20%可以使产出增加25%。

长期均衡产出也会随技术改进的平方而增加。因此，创新增加的产出要比线性增长更快。我们可以使用这个模型来了解其原因。如果从一个长期均衡的经济体开始，并将技术参数提高50%，那么产出将会增加50%，投资也将增加50%。然后投资超过了折旧，经济继续增长，投资将继续超过折旧，直到经济再增长50%为止，在这一点上，因折旧而造成的资本损失抵消了投资。这个计算过程揭示了创新乘数（**innovation multiplier**）的存在，创新有两个效应。首先，创新直接增加产出；其次，创新间接导致更多的资本投资，从而导致产出再次增加。因此，创新是持续增长的关键。[14](#)

需要注意的是，产出的这些增加不是瞬间发生的。当技术出现了一个突破时，技术参数的变化是相当缓慢的。直接效应的影响需要随着时间的推移显现。旧的实物资本必须被新技术的新实物资本所取代。当计算机技术进步刚刚发生时，一般企业的计算机不会马上变得更快，只有当技术发生了变化并且企业购买了新计算机之后，它们才会变得更快。

实物资本投资增加导致的二阶增长则会发生在更长的时间范围内。技术与技术对增长的影响之间的滞后，可能意味着创新在出现后的几十年时间内都会导致增长。火车是在19世纪早期发明的，但是镀金时代（Gilded Age）并没有马上开始，直到19世纪的后半期才到来，这是一个长达50多年的“时滞”。另一个例子是，在阿帕网（ARPANET）出现整整30年后，互联网才开始步入繁盛期。¹⁵

国家缘何成功与失败

我们还可以将增长模型应用于一系列重大政策问题，例如，落后国家是否可以赶上发达国家？为什么有些国家取得了成功而有些国家却失败？政府在促进增长方面能够发挥什么作用？这些研究揭示了增长模型的价值和局限性。为了便于说明，不妨从低国内生产总值的国家实现快速增长的能力开始讨论。模型证明，资本积累可以实现快速增长，技术投资也可以。一个实物资本较少的落后国家，有可能通过新的资本投入进入技术前沿，从而实现难以置信的高速增长。¹⁶

创新对长期增长来说是必不可少的，这种必要性也意味着一次性进口新技术有很大的局限性，而持续增长需要创新。

这些模型也表明，攫取和腐败，也就是政府将经济体的产出挪用于政府开支，会减少储蓄，进而削弱增长。对经济增长率的跨国比较研究的结果支持以下这些观点：减少攫取和腐败以及促进创新，都能推进经济增长。实现这些目标，需要一个强大但有限的中央政府来促进多元化。强大的中央政府能够保护产权、贯彻法治。多元主义能够阻止精英的俘虏，精英往往更喜欢现状，可能不会接受创新，因为创新往往可能具有很大的破坏性。

举一个破坏性创新的例子：克雷格列表网（Craigslist）。用户可以自行在这个网站上发布待售和求助广告。在21世纪初，克雷格列表网导

致美国平面媒体行业失去数十万个工作岗位，但其实在那个时候，克雷格列表网本身只雇用了几十名员工而已。虽然许多人失去了工作，但是克雷格列表网通过增加技术参数使整体经济更有效率。而在一个多元化程度较低的社会中，平面媒体行业可能会游说政府禁止克雷格列表这样的网站。很显然，这样做将会减缓经济增长。

中国的经济优势

线性模型+72法则：从1960年到1970年，日本的国内生产总值以每年10%的速度增长。根据线性模型的预测，连年10%的增长，会使日本经济每7年翻一番（运用72法则）。1970年，日本人均国内生产总值约为2 000美元（以当前美元计）。如果这种增长趋势持续下去，那么到2012年，日本的人均国内生产总值将会翻六番，也就是说，人均国内生产总值将达到128 000美元。

增长模型：增长模型对日本经济增长的解释为实物资本投资的结果。这个模型还预测日本经济增长在一段时间内将会是凹性的。具体来说，这个增长模型的预测是，当日本的国内生产总值接近美国和欧洲时，日本的经济增长率会降低到1%~2%的跨国平均值。¹⁷证据支持这个预测。从1970年到1990年，日本国内生产总值的年增长率大约为4%。但是从1990年到2017年，它的增长率仅为1%或更低。

中国的增长：从1990年到2010年，中国国内生产总值的增长率接近10%。2016年，中国的人均国内生产总值达到了8 000美元左右；正如增长模型预测的那样，增长速度已经放缓了。从2013年至2017年，增长率接近6%。同样，在中国，10%的增长率不可能一直持续下去，这与72法则相悖。如果中国经济在整个21世纪一直保持10%的增长水平，那么到这个世纪结束时，中国的人均国内生产总值将会超过1亿美元。

这毕竟是一个非线性的世界

之所以要构建非线性模型，是因为我们感兴趣的现象很少是线性的。在本章中，我们看到收益递减和收益递增是许多经济、物理、生物和社会现象的共同特征。我们还看到，在模型中包含曲率是有重要含义的。也许最重要的是，我们看到了函数形式能够影响我们的思维，用函数拟合数据有助于做出精确的表述。科学家可以使用碳-14数据来计算人工制品的年龄，经济学家还可以估计经济小幅增长的长期影响。

本章的一个核心结论是，一旦包括了非线性，直觉就变得不够用了。直觉可以告诉我们影响的方向：储蓄的增加、劳动力的增加和技术创新可以加快增长。模型还揭示了这些影响的形状和形式。正如我们所料，储蓄具有线性效应。从长远来看，劳动力的增加也是如此，即便模型假设短期收益递减。创新的增加还会产生乘数效应，我们对这种效应取其平方。第一个增长是创新的直接影响，产出的第二次增长则来自资本的增加。

在模型的帮助下，这些见解会变得很清晰。如果没有模型，我们通常可以推断出上升和下降的内容，但缺乏对功能关系形式的理解。我们会倾向于以线性的方式来思考，从而得出日本的经济将很快成为世界霸主的结论。利用模型，我们可以更好地思考非线性效应。也就是说，本章中介绍的凹函数型和凸函数，在非线性模型的巨大海洋中，只不过是沧海一粟。如果我们希望提高在复杂世界中推理、解释和行动的能力，就需要更深入地研究非线性现象。

09 与价值和权力有关的模型

你的价值不在于你知道了什么，而在于你能够分享什么。

罗睿兰（Ginni Rometty）

在本章中，我们讨论对个体行为者的价值和权力进行量化分析的模型。有些情况很容易处理。当一个群体的总产出等于每个成员个人贡献的总和时，每个人的价值就等于自己的贡献。但是，当集体产出不能分解为单独的组成部分时，例如当一组计算机程序员编写软件程序时，或者当一群创业企业家提出了新技术的某种创造性用途时，要分清每个人的贡献就会很困难。在美国，将权力授予政党时也会出现类似的问题：一方政党控制的议会数量与权力相关，但是这种相关并不是完美的相关。

在本章中，我们定义了度量价值和权力的两个标准。第一个标准是“最后上车者价值”（last-on-the-bus value，简称LOTB），它等于一位行动者在团队已经形成的情况下加入团队时的边际贡献。第二个标准是夏普利值（Shapley value），它等于行动者遍历所有可能的加入团队的序列，加入团队时的边际贡献平均值。例如，在一个由三个人组成的团队中，要求出一位行动者的夏普利值，先要求出他以第一、第二、第三位加入者的身份加入时的边际贡献，再计算平均值。我们是在合作博弈模型的框架下定义这些度量标准的。合作博弈模型由一组博弈参与者和一个价值函数组成。这个价值函数为每个可能的博弈参与者子集分配一个集体收益。

本章由四个部分组成。在第一部分中，我们定义了合作博弈、“最后上车者价值”和夏普利值，并给出了一些实例。在第二部分中，我们讲述了夏普利值的公理基础，并证明它是唯一能满足四个条件的度量标

准。其中有两个条件分别是：对永远不能为团体增加价值的博弈参与者必须赋予零值，所有博弈参与者的价值总和必定等于博弈的总价值。在第三部分中，我们将夏普利值概念应用于执行某个创造性任务的团队。在这里，创造性的团队指每个成员都有新想法的团队。我们将阐明在这种情况下，夏普利值是怎样产生直观的价值衡量标准的。在第四部分中，我们考虑如何将夏普利值方法应用于投票博弈这个特殊情况。我们利用这个概念区分了投票权与投票百分比，结果发现两者之间并不总是一致的。某个拥有20%席位的政党，在这一次投票中可能完全没有权力，但是在另一次投票中却可能得到1/3的总权力。

合作博弈

合作博弈由一组博弈参与者和一个价值函数组成。这个价值函数为博弈参与者的每个可能的子集（通常称为联盟）分配一个值。合作博弈的目标是刻画集体工作和联合项目。在合作博弈模型中，假设人们都会参与，以便我们可以专注于讨论如何为他们的参与分配价值。

合作博弈

合作博弈由 N 个博弈参与者和一个价值函数组成。这个价值函数为任何子集 $S \subseteq N$ 分配一个值 $V(S)$ 赋值。这些子集称为联盟。没有博弈参与者组成的联盟的价值等于零，即 $V(\emptyset)=0$ 。所有 N 个博弈参与者的价值 $V(N)$ 等于博弈的总价值。

在合作博弈中，一个博弈参与者的“最后上车者价值”等于当他是最后一个加入团队的人时，他所能增加的价值。“最后上车者价值”刻画了边际博弈参与者的价值。如果雇用4个人来搬运一张桌子，假设搬运这张桌子产生的价值为10，并且要4个人一起动手才搬得动，那么每个人的“最后上车者价值”均为10。如果只需要三个人就可以搬动这张桌子，

那么每个人的“最后上车者价值”均为零。这里需要注意的是，“最后上车者价值”不一定是博弈的总价值相加。特别是，如果价值函数表现出了规模收益递减的性质，那么“最后上车者价值”的总和将小于博弈的总价值；如果增加的价值表现出了规模收益递增的性质，那么“最后上车者价值”的总和将超过博弈的总价值。

一个博弈参与者的夏普利值，等于他在所有可能加入的联盟的次序下对联盟边际贡献的平均值。换句话说，我们要在想象中按顺序将博弈参与者加入联盟中并计算每个博弈参与者为每个序列增加的价值。例如，考虑一家同时在西班牙和法国运营的小公司，它至少需要一位会讲法语的人和一位会讲西班牙语的人开展日常业务。假设该公司有三名员工：一名会讲西班牙语的人、一名会讲法语的人和一名既会讲法语又会讲西班牙语的双语人士。

现在假设，这个合作博弈为任何一位能讲法语和西班牙语的人分配了1 200美元的价值。如果该公司能够运营，这个金额就等于公司每日的收入。如果任何两名员工来上班了，那么第三名员工就不是必需的。因此，在这个例子中，每个博弈参与者的“最后上车者价值”为零。

为了计算只会讲法语的那个人的夏普利值，我们要考虑这三个人来上班的所有6种可能的次序。在这6种次序中，只有在一种情况下，也就是只会讲西班牙语的人第一个到，然后这个只会讲法语的人第二个到时，这个只会讲法语的人才增加了价值。因此，这个只会讲法语的人的夏普利值就等于 $1/6$ 乘以1 200美元，即200美元。与此类似，只会讲西班牙语的那个人只有当他第二个到且只会讲法语的那个人第一个到时，才能增加价值，因此他的夏普利值也等于200美元。而在其他四个次序中，既会讲法语又会讲西班牙语的人第一个到或者第二个到都能增加价值，因此，他的夏普利值等于800美元。所有这三个人的夏普利值总和等于1 200美元，也就是这个博弈的总价值。

夏普利值

给定合作博弈 $\{N, V\}$ ，夏普利值的定义如下：

N 个博弈参与者加入联盟的次序有 $N!$ 个，让 O 代表这所有 $N!$ 个次序。对于 O 中的每一个次序，将博弈参与者 i 增加的价值定义为当博弈参与者 i 加入时价值函数发生的变化。博弈参与者 i 的夏普利值等于他在 O 中所有次序上增加价值的平均值。

在了解了上述基本概念的基础上，现在可以构建一个更加复杂的例子了。想象一下，在赛艇比赛中，一个团队通常由四名桨手和一名舵手（舵手通常个子较小，控制划桨节奏和方向）组成。现在想组建一支赛艇队，就需要找到六名赛艇运动员，也就是合作博弈中的博弈参与者：五名高大强壮的桨手和一名舵手。在参加比赛的时候，四名桨手和一名舵手上场的团队价值为10；或者由五名桨手上场，但由于重量过重，整个团队表现不佳，这样的团队价值为2。

为了计算出夏普利值，假设这些博弈参与者以各种可能的顺序加入。如果舵手以第一、第二、第三或第四位的次序加入，那么他不会增加任何价值；当他以第五位的次序加入时，他增加的价值为10，这种情况出现的概率为 $1/6$ ；如果他以第六位的次序加入，那么他将取代一位桨手，所增加的价值为8。将所有这些情况平均，可以发现舵手的夏普利值等于3。

而对于任何一个桨手来说，当且仅当他以第五位的次序加入时，才能增加价值，这种情况发生的概率为 $1/6$ 。如果舵手没有加入，那么以第五位加入的桨手所增加的价值为2。如果舵手已经加入，那么以第五位加入的桨手所增加的价值为10。由于舵手最后一位加入的机会是 $1/5$ ，同时舵手在前四位加入的机会是 $4/5$ ，因此可以求出每一个桨手的夏普利值为 $7/5$ 。¹从直观上就可以看出，舵手的价值应该超过单个桨手的价值，同时考虑到桨手可以在没有舵手的情况下参加比赛（尽管成绩

会“很差”），因此舵手的价值应该比所有桨手的总价值低。有无数种方法都可以在满足上面这两个约束的前提下分配价值，夏普利值只是给出了其中一个特定的分配方案：给舵手分配3，给所有桨手分配7。

夏普利值的公理基础

我们现在讲述夏普利值唯一满足的公理，这也就解释了为什么要优先考虑夏普利值而不是其他。第一，我们是通过对所有可能次序中博弈参与者的边际贡献来计算夏普利值的，因此任何永远不能增加价值的博弈参与者的夏普利值都为零。第二，对任何两个相同的博弈参与者，即对每个联盟贡献相同的任两个博弈参与者，也必须分配给他们相同的夏普利值。第三，由于所有次序的价值总和等于博弈的总价值，所以夏普利值的总和也必定等于与博弈的总价值。这里需要注意的是，“最后上车者价值”虽然满足前两个性质，但是却不满足最后一个公理。

在这三个公理的基础上，还可以增加第四个公理——可加性。这个性质要求，如果合作博弈的价值函数可以分解为两个价值函数，并把每个分解出来的价值函数分配给一个不同的合作博弈，那么复合博弈中一个参与者的价值应该等于他在两个分博弈中的价值总和。很容易看出，夏普利值也满足这个性质。不过，这四个公理唯一刻画了夏普利值这一点其实并不太明显。

证明一种度量唯一满足一组公理，也就为这种度量奠定了坚实的逻辑基础。如果没有公理基础，或许也可以找到某种直观的度量，但是我们可以认为它是武断的、似是而非的。上述公理告诉我们，如果选择任何其他度量，就不得不至少放弃其中一个公理。当然，这并不意味着夏普利值是唯一合理的标准。罗依德·夏普利（Lloyd Sharpley）这位伟大的经济学家和数学家，可能是先写出了这个标准，然后才构造了这些只有它唯一满足的公理的。当然，谁先谁后其实并不重要。即便这些公理是用后向方法构造出来的，只要我们接受，就应该采用这种方法。衡量

标准的适当性取决于公理的合理性。就这些公理而言，前三个公理是无可争议的。第四个公理（可加性）虽然看上去比较复杂，但也是合理的，如果没有这个公理，博弈参与者就会有强烈的动机去合并或分割联盟。

夏普利值的公理基础

夏普利值唯一满足以下公理：

零性：如果博弈参与者为任何联盟增加的价值都等于零，那么该博弈参与者的价值等于零。

公平性 / 对称性：如果两个博弈参与者对任何联盟都具有相同的增加价值，那么这两个博弈参与者具有相同的价值。

完全分配性：博弈参与者价值的总和等于博弈的总价值 $V(N)$ 。

可加性：给定两个定义在相同博弈参与者集合之上的博弈，它们的价值函数分别为 V 和 $\hat{V}$ ，那么在博弈 $(V + \hat{V})$ 中，一个博弈参与者的价值等于该博弈参与者在 V 和 $\hat{V}$ 的价值的总和。

夏普利值的应用

现在，我们将夏普利值应用在基于替代用途测试（alternative uses test）的合作博弈中。在测试中，每个人都必须为一种常见的物品想出一些新的用途，比如砖块。这种测试的目的是根据人们想出来的用途或用途类别来衡量一个人的创造力。我们在计算夏普利值的过程中，发现了一个直观的评分规则。

想象一下有三个人参加了某替代用途测试，分别是阿伦、贝蒂和卡洛斯。测试要求他们想出区块链的替代用途，这是一种分布式记账技

术。如图9-1所示，阿伦和卡洛斯分别提出了6个想法，每个人的创造力得分均为6；贝蒂则提出了7个想法，因而得到7分。他们这三个人组成的团队的总创造力得分为9，因为总共有9个不同的想法（不同人提出的想法，有些是重合的）。

为了计算夏普利值，可以写下这个团队能够形成的所有6种可能的排序，而且只有当某个人为团队提供了独特的想法时才“给分”，然后再对所有6种情况求平均值。或者，在计算夏普利值的过程中，我们可能已经注意到了，某个人因某个想法而“得分”的概率等于1除以所有提出了这个想法的人的数量。任何一个提出了一个他人没有的独特想法的人都可以获得满分。图9-1用粗体字来表示这类想法，例如阿伦提出的用区块链进行艺术交易的想法。

如果两个人提出了同一个想法，那么每个人都有1/2的机会首先加入该团队。同样，如果所有三个人都想到了同一个想法，那么每个人都有1/3的机会首先加入。这就是说，在想到同一个想法的人之间平等地分配得分能够产生夏普利值。因此，它是分配满足4个公理的值的唯一方法。这些值表明，虽然阿伦不是提出最多想法的那个人，但是他却增加了最多的价值。²

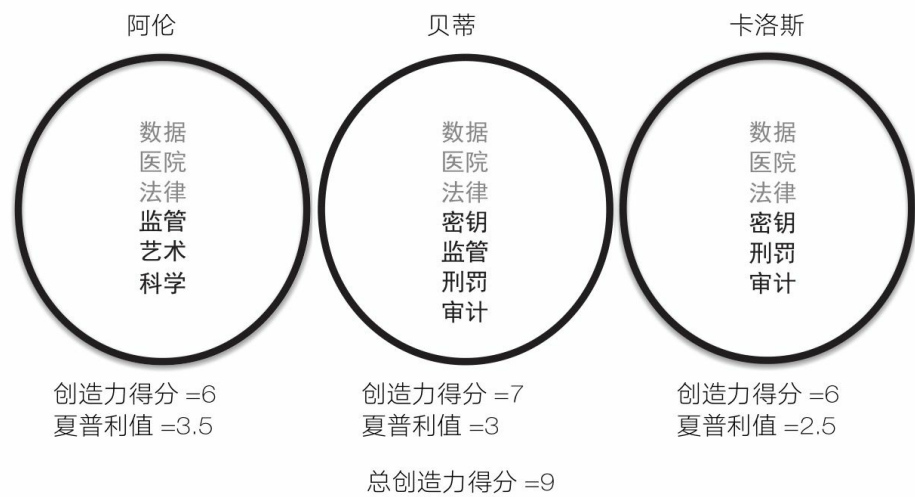


图9-1 在替代用途测试中应用夏普利值

夏普利-舒比克权力指数

接下来，我们将夏普利值应用于一类投票博弈。在这种投票博弈中，每个博弈参与者（代表某个政党或官员）控制着固定数量的席位或投票权，而且要采取行动，就必须获得多数席位或支持票。在投票博弈中，夏普利值通常被称为夏普利-舒比克权力指数（Shapley-Shubik index of power）。³通过对这个指数的计算，我们发现一个博弈参与者（政党）控制席位（投票权）的百分比与其权力之间并不存在直接的转换。

为了计算权力指数，考虑各个政党加入联盟所有可能的次序。如果某个政党加入了一个联盟并获得绝对多数，那么这个政党所增加的价值等于1。在这种情况下，我们就称这个政党是“关键的”。否则，这个政党不会增加任何价值。

假设议会中共有101个席位，分别由4个政党掌握：A党控制了40个席位、B党控制了39个席位、C党和D党则各控制了11个席位。在这个例子中，如果A党首先或最后加入，那么A党就不会成为“关键的”政党。但是，如果A党在第二位或第三位加入，就肯定会成为“关键的”政党。因此，A党的权力指数为 $1/2$ 。如果B党在第一位或最后一位加入，那么它就不能增加任何价值；如果B党在第二位加入，那么当且仅当A党已经在第一位加入时，B党才可能成为“关键的”政党。如果B党在第三位加入，那么它要想成为“关键的”政党，唯一的机会是A党在最后一位加入。这两个事件组合发生的概率分别为 $1/12$ 。因此，B党的权力指数等于 $1/6$ 。C党和D党也可以在两个与B党类似的事件组合中成为“关键的”政党。如果C党或D党在第一位加入，那么不可能成为“关键的”政党。如果A党在第一位加入，那么只要C党或D党在第二位加入，就能成为“关键的”政党。如果A党在最后一位加入，那么只要C党或D党在第三位加入，也能成为“关键的”政党。因此，C党和D党的权力指数也分别为 $1/6$ （图9-2）。

政党	席位数量	权力指数
A	40	1/2
B	39	1/6
C	11	1/6
D	11	1/6

图9-2 席位与权力之间的脱节

这个例子表明，一个政党控制的席位百分比与它实际拥有的权力之间可能存在着脱节。A党和B党控制的席位数量几乎相同，但是A党的权力却是B党的三倍；B党控制的席位虽然比C党或D党多得多，但是所拥有的权力却不比它们大。与这个例子相似的席位分配在现实世界的议会制度中经常出现。因此，只拥有少量席位的政党往往可以掌握很大的权力。例如，在以色列议会中共有120个席位。2014年，利库德集团领导的联盟共有43个席位，反对派联盟则拥有59个席位（仅略低于多数席位），正统派联盟拥有18个席位，但所有这三方都拥有相同的夏普利-舒比克权力指数。当然，相同的权力指数并不意味着正统派联盟在现实世界中确实拥有完全相同的权力，不要忘记，所有模型都是错的。但是它确实表明，正统派联盟的影响力超过了他们的席位数所占的百分比。

20世纪60年代中期，纽约拿骚县（Nassau County）监事会出现过惊人的席位与权力脱节的情况。该监事会由6名成员组成，每个成员控制的选票与该成员所代表的地区的人口成比例（图9-3）。投票事项要多数通过，需要得到115张票中的58张或以上。在三个最大的地区中，任何两个地区合作都可以稳占多数。因而另外三个地区的投票永远不可能是决定性的，这三个地区没有权力。例如，虽然北亨普斯特德（North Hempstead）地区拥有21票，超过了总票数的18%，但是并不能影响投票结果。

地区	票数	权力指数
亨普斯特德 1 区 (Hempstead 1)	31	1/3
亨普斯特德 2 区 (Hempstead 2)	31	1/3
奥伊斯特贝 (Oyster Bay)	28	1/3
北亨普斯特德	21	0
长滩 (Long Beach)	2	0
格伦科夫 (Glen Cove)	2	0

图9-3 拿骚县席位与权力脱节

理论上说，夏普利-舒比克权力指数适用于任何席位或投票权分配不均等的情况，比如欧盟或美国的选举团，但这并不意味着它在所有情况下都是适当的方法。就美国选举团制度而言，50个州可以有50! (3×10^{64}) 的不同次序。

当然，考虑到选民偏好的区域相关性，并非所有联盟都是可能的。密西西比州几乎不太可能与纽约州组成联盟。为了提供更有效的权力衡量标准，我们需要将某些联盟置于更优先于其他联盟的位置，或将某些联盟排除出去。在后面的章节中，我们描述了允许排除某些联盟的迈尔森值 (Myerson Value)。

小结

个体的夏普利值与为联盟增加的平均贡献相对应。它是衡量增加价值的一种标准。在投票博弈中，也可以将夏普利值解释为权力的一种度量。不过，夏普利值可能并不一定总是最好的衡量标准。假设威胁是可信的，那么在一个群体已经形成的情况下，个人的“最后上车者价值”可能是衡量权力的一个更好标准，因为它能够度量每个人通过威胁离开可以攫取多大利益。

在这些情况下，联盟会希望减少“最后上车者价值”。通过扩大联盟

规模，可以创建出一个具有很高的总价值、同时“最后上车者价值”又足够低的联盟。不断加入新成员，会使现有成员变成“可以放弃的”，从而使“最后上车者价值”趋向于零。我们在实践中确实可以观察到这一点。例如，雇主会通过雇用多余的工人来削弱工人的权力，制造业企业会向多个相互竞争的供应商采购中间产品，政府会与多个承包商签订合同，等等。

同样的直觉也可以用于解释美国立法机构中出现的联盟。国会游说者和政党领导人希望通过法案（价值的一种结果），同时又试图限制个别众议员和参议员的权力。⁴如果游说者努力争取到了通过法案所必需的最低数量的众议员和参议员的支持，那么每一个众议员和参议员都会拥有很大的“最后上车者价值”。任何一个人都可以通过改变自己的投票来推翻那个法案。在这种情况下，游说者可以通过收买绝大多数众议员和参议员来降低他们的“最后上车者价值”。同样的逻辑也意味着，只拥有微弱多数的政党可能是非常难以驾驭的，因为每一个成员都拥有很大的“最后上车者价值”。而在某个政党拥有了绝大多数席位（投票权）的时候，没有任何众议员或参议员能够拥有太大的权力。

将视野放大到现代互联网世界，我们发现应用“最后上车者价值”和夏普利值的概念来思考权力问题非常有用。无论是个人、组织、企业，还是政府，抑或是恐怖组织的权力，都部分取决于偏离合作制度可以造成的损害的程度，也就是“最后上车者价值”。一个技术高超的计算机黑客，由于拥有摧毁大量财富的力量，因而拥有巨大的权力。即便黑客完全不能给社会创造价值，这个结论依然成立。

在考虑跨国企业或其他跨国组织的价值时，夏普利值可能是一个更好的衡量标准。在这些情况下，退出本身就是一个不可行的选择。能源公司必须参与能源生产博弈、能源分配博弈、房地产博弈、环境博弈、就业博弈等。这样的公司的总增加值等于各个领域的增加值之和。

通过合作博弈论的视角来思考权力和价值，可以得出很多深刻有力

的洞见。合作博弈还指出了我们下一步应该关注的地方。在政界和商界，并不是所有联盟都是合理的。不过目前的模型假设它们合理。更丰富的模型需要考虑到世界的连通性。咨询公司和金融公司要从科技公司购买软件，科技公司和咨询公司通过金融公司进行投资和借贷，金融公司和科技公司要聘请咨询师。在这些网络中，每个参与者都能增加价值并发挥影响力。要计算出这种环境下的权力，我们需要网络模型。

10 网络模型

网络理论是科学的一个完整分支。但是就过去的二三十年来说，它相对较新。我们还没有机会把所有这些理论从大学中拿出来，然后问自己：“我们应该建立什么样的网络？应该将网络用于什么样的目的？”

安妮-玛丽·斯劳特（Anne-Marie Slaughter）

本章将介绍网络模型。对网络进行全面研究，需要写好多本书。因此，我们在这里只专注于一个更加温和的目标：只希望了解有关网络的基础知识，能够给网络的各个组成部分命名，并讨论它们对于建模的重要性。我们得出的答案是，网络几乎总是很重要。我们构建的任何模型，无论是市场模型、传染病传播模型，还是信息传播模型，都可以通过将参与者嵌入网络中而变得更加丰富。¹

网络无处不在。人们经常会谈起贸易网络、恐怖主义网络，以及志愿者网络。不同物种会组织成食物链，那是一种网络形式。企业会建立供应链，那也是一种网络。如前所述，将金融系统视为一个支付承诺的网络会很有效。网络对于理解社会关系一直都很重要。在人类历史的大部分时间里，社交网络受到地理限制，难以扩展。由于技术的进步，许多社会互动和经济交易现在都是通过虚拟网络进行的，并且可以使用模型进行分析。

本章内容仍然遵循前面用来讨论分布时所用的模式，即结构—逻辑—功能。我们首先用一系列统计量来表征网络结构，包括：度、路径长度、聚类系数和社区结构等。然后我们讨论了一些常见的网络类别：随机网络、中心辐射网络、地理网络、小世界网络和幂律网络。之后探索网络形成的逻辑，我们构建了一些微观层面的流程，生成所能观察到的

网络结构。最后讨论功能，也就是网络结构为什么是一个重要的问题。

本章主要关注网络结构的五个重要含义。我们首先从友谊悖论入手，在对它进行了全面分析之后，描述六度分隔理论和弱关系属性的强度。最后讨论了网络在节点或出现故障时的鲁棒性，解决了网络上的信息集结问题。本章最后还讨论了网络会如何影响模型结果。

网络的结构

网络由节点以及连接节点的边（edge）组成。由边连接起来的节点互为邻居。如果沿着边，可以从任何一个节点到达任何其他节点，就将这样的网络称为连接的网络。网络可以用图形来表示，也可以用边的列表表示，或者也可以用由0和1组成的矩阵表示，其中第A行、第B列的一个数字表示节点A和节点B之间的边。虽然人们更喜欢用图形来表示网络，但是其实用列表和矩阵来表示网络，才更适用于计算网络统计数据。

网络中的边可以是定向的，也就是说，可以从一个节点指向另一个节点。在信息网络中，一条有向边表示一个人从另一个人获取信息。在生态系统网络中，从红尾鹰到灰松鼠的一条有向边表示红尾鹰吃松鼠。边也可以是非定向的；连接两个朋友的边就是非定向的。在非定向网络中，一个节点的度（degree）等于连接到它的边的数量。

网络以一组网络统计数据为特征。对于每个统计量，我们可以计算网络平均值和所有节点的分布。例如，友谊网络的平均程度告诉我们平均每个人有多少个朋友。度分布（degree distribution）告诉我们某些节点是否比其他节点连接得更多。社交网络的分布比万维网、互联网和引文网络更加平等，后面这几类网络都有很长的尾巴。

路径长度，指两个节点之间的最小距离，与度成反比。当增加边时，就缩短了节点之间的平均距离。在航空公司的航线网络中，路径长

度对应于人们从航线网络中的某个城市到另一个城市所需的航班数量。如果要在两家航空公司之间做出选择，在其他所有条件（即价格）都相同的情况下，旅客会更喜欢平均路径长度更低的那家航空公司。平均路径长度也与信息丢失相关。经过多人中转传递的信息比直接在两个人之间传递的信息更容易遭到扭曲。最短路径上的节点在网络中起着关键作用。如果信息是通过最短路径传递的，那么就必定会经过最短路径上的节点。节点的介数得分（**betweenness score**）等于通过该节点的最小路径的百分比。在社交网络中，介数得分高的人掌握更多信息并且拥有更多权力。

最后一个统计量是聚类系数，它等于节点的邻居节点对当中，同时彼此也互为邻居节点对所占的比例。例如，一个人有10个朋友，这些朋友可以组成45个对。如果在这45个对当中，有15个对本身也是朋友，那么这个人的聚类系数就等于 $1/3$ 。如果所有这45对都是朋友，那么这个人的聚类系数就等于1，这也是所有可能值当中最大的一个。整个网络的聚类系数等于各个节点聚类系数的平均值。

网络统计量

度：节点的邻居数（即边数）。

路径长度：从一个节点到另一个节点必须遍历的最小边数。

介数：经过某个节点连接两个其他节点的最短路径数量。

聚类系数：一个节点的邻居对当中，同样也由一条边连接的邻居对所占的百分比。

图10-1显示了一个辐射网络和一个地理网络，它们各具有13个节点。在这个辐射网络中，中心节点的度为12，所有其他节点的度均为1，因此平均度小于2。这种度分布是“不平等”的。中心节点与其他每个节点的距离均为1。所有其他节点与中心节点的距离为1，与中心节点之

外的任何一个节点的距离为2。因此，这个辐射网络的平均路径长度也小于2。中心节点位于任何两个其他节点之间的最小路径上，于是介数得分为1。任何一个分支节点都不位于连接其他节点的任何最小路径上，因此它们的介数得分均为0。最后，在这个辐射网络中，连接到某个节点的任何节点都不彼此连接。因此，网络的聚类系数为0。

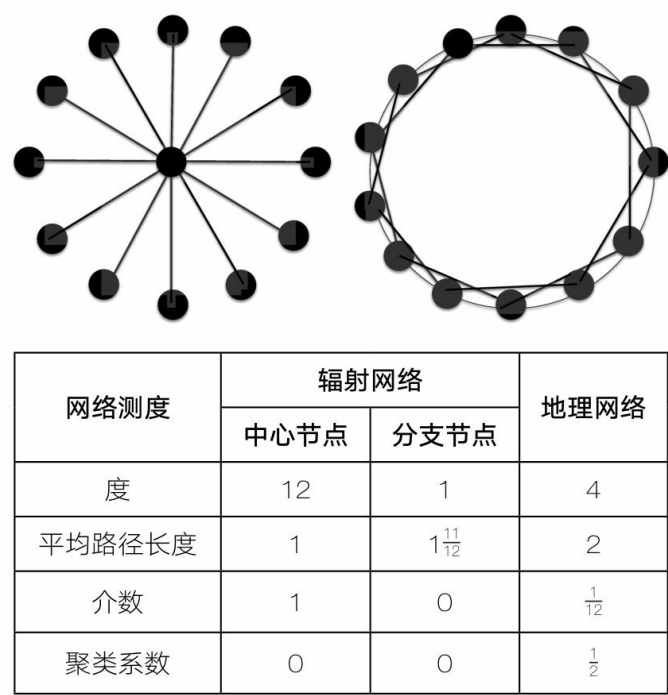


图10-1 一个中心辐射型网络和一个地理网络

在这个地理网络中，每个节点都连接到位于它右侧和左侧的两个节点，因此平均度等于4。每个节点到4个节点的距离为1、2、3，因此平均距离恰好等于2。从图10-1可见，这个地理网络的度和距离分布都是简并性（degenerate）的，因为每个节点都具有相同的度和相同的平均距离。可以看出，每个节点的介数都等于1/12。²每个节点都有4个邻居，可以构成6个对。在这6个对中，恰好有3对是相互连接的：直接靠着该节点的左右两个节点分别连接到再外一点的节点，并相互连接。因此，聚类系数等于1/2。

刻画网络的聚集程度的另一种方法是将节点划分为不同的社区

（community）。在年轻人的友谊网络中，这里所说的社区可能对应于对艺术、体育或科学感兴趣的青少年。或者，社区也可以通过种族和性别来定义。政治联盟网络可能可以分为地区性的或意识形态上的盟友。可以用来确定社区的方法有很多，其中一种方法是依次移除具有最高介数的边，因为介数高的边更有可能将不同的聚类连接起来。还有一种方法是将社区的数量视为给定的，并在特定的目标函数下寻找最佳划分方法，例如最小化社区之间边的数量或最大化社区内部边的比例。³

我们可以使用社区检测算法（community detection algorithm）去看网络数据的问题。有研究表明，人们可能会生活在“网络泡沫”（online bubbles）当中。这也就是说，我们所属的社区，可能是由只从类似来源获取新闻的人组成的。如果真的是这样，那么无疑会对社会凝聚力产生重要的影响。在互联网出现之前，情况可能也是这样的，但是很难用数据证明。现在，数据科学家可以在网络上抓取海量数据，从而将人们的新闻来源准确地识别出来，实际上我们确实一定程度上生活在“泡沫”当中。现在，模型可以给出社区的正式定义，数据可以揭示这些社区的力量。结合我们的判断力，就可以根据数据做出明智的推断。

常见的网络结构

在分析网络时，我们遇到了网络过于多样性的问题（图10-2）。少数几个网络统计量无法确定具体的网络结构：人们可以构建出数十亿个具有10个节点且平均度为2的不同网络。还可以通过检验它的统计指标是否与某个常见的网络结构有显著差异来表征网络。例如，当一位法官引用了另一位法官的判决时，一个研究者可能会收集司法引用的数据并将其以网络的形式表示。这种网络从图形上看似乎有一个很特殊的结构和集群。我们可以将这个网络的统计数据与具有相同数量的节点和边的随机网络进行比较，以检验这个网络是不是随机的。在第一种常见的网

络随机网络（random network）中，聚类系数等于一条随机的边的概率，因为一个节点的两个邻居并不比任何其他随机选择的节点更可能包含一条边。

随机网络的蒙特·卡罗方法^⑤

为了检验一个具有 N 个节点和 E 条边的网络是不是随机网络，可以创建大量具有 N 个节点和 E 条边的随机网络，并计算出度、路径长度、聚类系数和介数的分布。然后，执行标准的统计检验，以确定接受还是拒绝那个网络的统计数据可能抽取自该模拟分布的假设。⁴

理论模型通常假设某种特定的网络结构。有的研究者偏好假设随机网络，而有的研究者则偏好假设规则的地理网络，例如这样的网络：节点排列成圆形并且每个节点在每个方向上都连接到最近的节点。这也是第二种常见的网络，有一种地理网络将节点排列在棋盘上，并让每个节点与自己东、南、西、北的邻居相连。大多数常见的地理网络都具有较低的度，即节点仅连接到本地邻居，并且具有相对较大的平均路径长度。在地理网络上，介数和聚类系数不会有变化。

第三种常见的网络是幂律网络，这种网络的度分布是幂律的。少数节点有许多连接，同时大多数节点的连接则非常少。第四种常见的网络是小世界网络，它结合了地理网络和随机网络的特征。⁵要想构建一个小世界网络，可以从一个地理网络开始，然后进行“重新布线”，方法是随机地选择一条边并把这条边所连接的其中一个节点替换为一个随机的节点。如果“重新布线”的概率等于零，所拥有的就是一个地理网络；如果“重新布线”的概率等于1，那么就有了一个随机网络；而当概率介于这两者之间时，就会得到一个小世界网络，以小集群区别于通过随机链接连接到其他集群的地理网络。社交网络看起来类似于小世界，每个人都有一群朋友，以及若干随机的朋友。

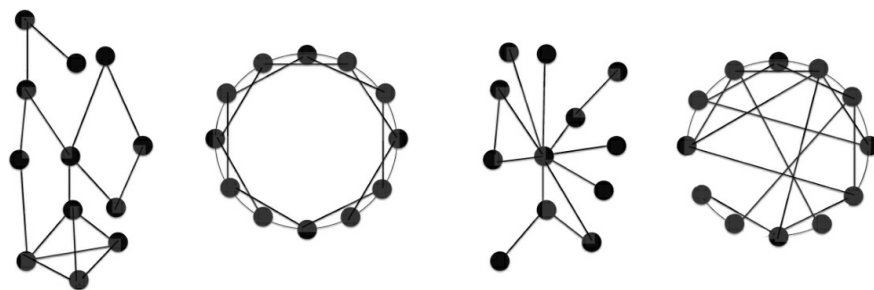


图10-2 随机网络、地理网络、幂律网络和小世界网络

网络形成的逻辑

现在简要讨论一下几个描述网络形成的模型，这些模型给出了解释网络结构的逻辑。我们遇到的大多数网络结构都是从个体行为者做出的关于建立连接的选择中涌现出来的。友谊网络、万维网和电网都是如此。这些网络不是计划的结果。不过，也有一些网络，例如供应链网络，确实是计划的产物。我们希望按计划构造的网络对节点的故障具有鲁棒性。当然，自发涌现的网络结构都具有鲁棒性这个事实是一个谜。

我们已经讨论了如何创建随机网络和小世界网络。只需要随机创建一组节点，然后绘制连接随机节点对的边，就可以创建随机网络。通过构建一个规则的地理网络（常用的方法是在一个圆周上排列节点并在每个方向上连接 k 个邻居），然后随机“重新布线”一部分边，就可以创建一个小世界网络。

电网形成的模型依赖于经济和工程原理。电网必须为家庭、企业和政府提供电力。无论生产者是否营利性企业还是公共事业公司，都没有动力去创造聚集程度很高的网络，因为那种网络效率低下。但是，聚类的缺乏降低了网络的鲁棒性。经济和工程方面的考虑也排除了远距离跳跃——跨越网络连接的可能性。电力公司不会建立从芝加哥到达拉斯的直接连接。然而，人和企业却可以，一个芝加哥人可能会与达拉斯的某个人建立起友谊，新加坡的一家公司可能会与底特律的一家公司进行交易。正如后面将看到的，这些远距离跳跃有助于提高网络的鲁棒性。

要创建一个具有长尾分布的网络，可以利用优先连接模型的一个变体。先随机创建一些节点，然后画出从新节点到现有节点的边。如果我们令连接到节点的概率与节点的度成正比，就可以产生幂律的度分布。在这个模型中，越早“到达”的节点的度越大。但是这个模型有一个缺点，那就是，它不允许节点质量有任何差异。更高质量的节点本应具有更高的度。不过，质量和度的网络形成模型（quality and degree network formation model）纠正了这种缺憾，并且产成了长尾分布。

质量和度的网络形成模型

创建 d 个互不连接的节点。在每周期 t 中创建一个质量为 Q_t 的、从分布 F 中抽取出来的新节点。根据其他 d 个节点的度将这个新节点连接到那些节点上。用 D_{it} 表示在时间 t 时节点 i 的度，那么给定 N 个节点时选择节点 i 的概率等于：

$$\frac{D_{it} + Q_{it}}{\sum_{j=1}^N (D_{jt} + Q_{jt})}$$

如果新节点质量的均值和方差都足够低，那么这个模型就类似于标准的优先连接模型。如果质量分布有一条长尾，那么质量很高的新节点的度可以增长到非常大的程度。⁶

网络的功能

前文中，已经提到过友谊悖论；事实上，在任何网络上，平均来说，人们确实不可能比他们的朋友拥有更多的朋友。我们可以利用辐射网络来说明出现这种情况的原因及其背后的逻辑。在辐射网络中，12个人中的每人都只有1个朋友，只有1个人有12个朋友。有1个朋友的那12个人都连接到了中心节点，中心节点有12个朋友。这个特征，也就是度更高的人与更多的人连接在一起的事实，驱动了结果。在中心网络上，

平均来说，所有人只有不到两个朋友。然而，平均而言，每个人的朋友都有超过11个朋友。

友谊悖论适用于任何网络：电子邮件网络、学术引文网络、银行网络和国际贸易网络等。平均而言，一篇学术论文引用的参考文献被引用的次数比这篇文章本身更多；与一个国家的贸易伙伴进行贸易的国家数量，要比与这个国家进行贸易的国家更多；食物网络中与单一物种相连接的多个物种的连接比该物种自身更多。在具有更加分散的度分布的网络上，朋友的数量与朋友的朋友的数量之间的差异会变得更加明显。例如，根据Facebook的数据对友谊网络进行的一项研究结果表明，一个人平均大约有200个朋友，而他们的朋友平均来说有超过600个朋友。⁷

友谊悖论

如果网络中任何两个节点的度不同，那么平均而言，节点的度会低于其相邻节点。换句话说，平均而言，人们的朋友比他们自己更受欢迎。⁸

友谊悖论的逻辑可以扩展到任何与朋友数量相关的性质。如果活跃、快乐、聪明、富有和友善的人平均而言会拥有更多的朋友，那么一个人的朋友平均来说会更活跃、更快乐、更聪明、更富有、更友善。⁹想象一下这样一个网络，在不快乐的人中，90%的人有4个朋友、10%的人有10个朋友。而快乐的人的情况则相反：在快乐的人中，10%的人有4个朋友，90%的人有10个朋友。人们的朋友将不成比例地由拥有10个朋友的人组成。由于那些人中的大多数都是快乐的，因此绝大多数人的朋友都会比自己更快乐。

现在，我们来看一下六度分隔理论，也就是地球上的任何两个人都可以通过6个或更少的朋友联系在一起。虽然友谊悖论适用于任何网络，但是六度分隔却只适用于某些类型的网络。这个术语源于美国社会

心理学家斯坦利·米尔格兰姆（Stanley Milgram）在20世纪60年代进行的一项实验。米尔格兰姆向内布拉斯加州奥马哈市和堪萨斯州威奇托市的296人寄出了一些包裹，那些包裹最终需要转寄给在马萨诸塞州波士顿市的一个人。收到包裹的人必须遵守相同的规则：所有参与者只能通过邮政系统将包裹寄给他们认识且他们认为更有可能认识那个波士顿人的人，并附上同样的指示。每个参与实验的人都要在一份记录路径的名册上签名，并邮寄明信片给研究者，以便研究者可以跟踪链条上的断点。最终，有64个包裹抵达了波士顿。这些抵达波士顿的包裹所经历的平均路径长度略小于6，因此就有了“六度分隔”这种说法。

50年后，研究者们又组织了一场更大规模的“六度分隔”实验。这一次是利用电子邮件进行的，实验组织者在全球范围内设定了18个最终收件人，邀请了20 000多个人参加了这个实验。电子邮件链的中间路径长度在5到7之间，具体取决于最初的发送者与最终目标之间的地理距离。由于实验中运用的路径长度不等于实验参与者之间的最小路径长度。因此，这些证据表明，大多数人之间用不着6度就可以连接起来了。¹⁰

在这里，我们构建了一个简化版的小世界网络，以便直观地理解六度分隔理论。这个小世界网络假设每个人都有一个由若干个圈内好友构成的小群体，这些人彼此认识，而且每个人都拥有不属于这些圈内的朋友，我们把这些圈子外的朋友称为“随机朋友”（random friends）。¹¹图10-3表明，某人（用黑色圆圈表示）有5个圈内好友和两个随机朋友。它还显示了这个节点的朋友（用浅灰色圆圈表示）的部分“朋友圈”。

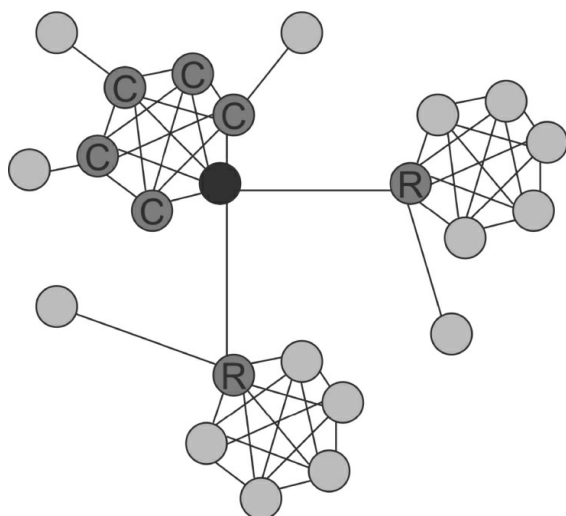


图10-3 一个节点的圈内好友（C）和随机朋友（R）

这些随机朋友也可以认为是一种弱关系，他们可以将你连接到其他群体的人。我们的弱关系，也就是网络中的随机朋友，由于连接了具有不同兴趣和信息的社区，从而发挥了重要的信息作用。因此，社会学家很强调弱关系的力量。¹²

在这种网络结构中，我们可以计算出二度邻居，也就是朋友的朋友的数量，方法是将随机朋友的所有朋友人数相加，但是不把圈内好友计算在内，因为他们本身就是节点群体的成员。与此类似，我们可以计算出朋友的朋友的朋友的数量，方法是把所有“圈内好友”的随机朋友的朋友加进来，但是不把随机朋友的“圈内好友”的“圈内好友”计算在内，因为他们在计算二度邻居时已经加进来了。为了产生六度分隔，我们将相同的逻辑应用于一个具有100个“圈内好友”和20个随机朋友的网络。

六度分隔

假设每个节点有100个“圈内好友”（C），他们彼此都是朋友；以及20个随机朋友（R），他们没有与节点共同的朋友。

一度： $C+R=120$

二度： $CR+RC+RR=2\ 000+2\ 000+400=4\ 400$

三度： $CRC+CRR+RCR+RRC+RRR=328\ 000$

四度¹³： 17 360 000

五度： >10亿

六度： >200亿

由于假设随机朋友的朋友之间没有重叠，这个模型隐含地假定人口是无限的。但是在现实世界中，随着度数的增大，真实的社交网络会出现朋友之间的重叠。在包括了重叠和其他真实世界特征（例如朋友数量的异质性）的网络中，实际值将会与上面计算出来的值不同。不过，每个度的邻居数量的相对大小将保持相似：一个人的三度邻居（朋友的朋友的朋友）会比二度邻居（朋友的朋友）多得多。

三度朋友可能是相当重要的，在我们这个例子中，三度朋友的人数超过了25万。与一个人的“圈内好友”不同，一个人的三度朋友往往会住在不同的城市，就读于不同的学校，拥有不同的信息，他们会更加多样性。他们也足够接近，可以建立起信任关系：朋友的朋友的朋友可能是你的室友的母亲和同事，或者是你妹妹的男朋友的姨妈。三度朋友的数量很重要，他们的多样性以及相对接近性使他们成了你的重要资产，他们可以提供新的信息和工作机会。这些人最有可能帮助你找到工作，促使你搬到新的城市，或者成为生活中、商业上的伙伴。

网络结构的鲁棒性

现在来讨论网络结构的鲁棒性，看看当网络的一个节点或一条边发生故障时会怎样。网络最重要的性质是，它在受到冲击时是不是仍然能保持连接。我们可以使用模型来计算网络保持连接的概率——作为移除节点数量的函数。还可以考察当移除某些节点时平均路径长度会发生什么变化。例如，在航空公司的航线网络中，这种分析可以告诉我们，如

果某个机场由于天气原因或电力故障而关闭了，将会需要多少额外的航班。

在这里，我们考虑网络中最大连接部分，即巨型组件的大小怎样随着节点的随机失败而变化的问题。图10-4显示了一个大型随机网络和一个大型小世界网络中巨型组件的大小。在这个随机网络中，巨型组件的大小一开始会呈线性下降。在边概率等于1除以节点数这个临界点上，巨型组件的大小会下降到原始网络的任意小的比例。但是，小世界网络却不会出现这种突然的变化。因为在小世界网络中，大多数连接都存在于地理聚类中。每个聚类都可以顶住多个节点发生故障的冲击。这个性质与随机连接相结合，可以有效地防止整个网络崩溃。

从图10-4可以看出：缺乏局部聚类的稀疏网络更容易出现故障。我们可以将这个结论应用于对电网的分析。电网缺乏远距离跳跃连接和紧密聚类，而保证小世界网络鲁棒性的就是这些。在电网中，节点或连接如果发生了故障，是不能通过聚类中的其他连接或与仍然有效节点的远距离连接来克服的。因此，局部故障就级联放大地传播到整个网络。¹⁴ 与此形成鲜明对照的是互联网。具有长尾度分布的互联网对随机节点故障具有很强的鲁棒性。互联网的度分布意味着，绝大多数节点的连接很少，因此即便它们发生了故障，网络也能保持连接。

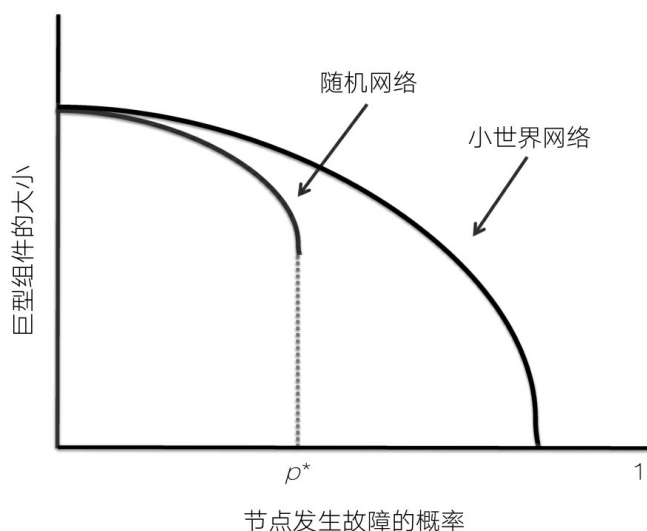


图10-4 巨型组件的大小作为节点故障的函数

到目前为止，我们一直假设节点是随机发生故障的。我们也可以考虑移除战略性节点。有些具有长尾分布的网络，比如说互联网，现在也变得比以前脆弱了。战略性地移除度最高的节点会破坏整个网络。可以从辐射网络看出其逻辑。当随机移除节点时，辐射网络仍然可以保持连接，除非直接移除中心节点，但那是一个低概率事件。然而，战略性地移除节点，即直接毁灭中心节点，却可以一步切断网络连接。

对于某些网络，例如恐怖主义网络和毒品贸易网络是我们想要切断的。如果这些网络是像电网那样的稀疏网络或者具有长尾度分布，那么就可以通过移除战略节点来断开这些网络。对于恐怖主义网络来说，这意味着必须逮捕拥有链接最多的恐怖组织成员。如果这些网络类似于小世界网络，那么它们就会具有鲁棒性，甚至在战略性的节点被移除之后仍然能屹立不倒。尝试切断这种网络任何“区段”的努力都将失败，因为随机重新连接的存在会将这些“区段”重新连接到网络上。

小结

当我们尝试构建关于由人构成的网络模型时，目的通常是用它们来刻画社会影响，也就是在社会网络中，一个人的成功、行为、信息或信念，会影响他们的朋友的成功、行为、信息或信念。一个人的行为既可能是依赖于情境的，也可能是由内在因素决定的；个人对共同事业的价值或贡献也是如此。一个人的价值或贡献可能是源于他本身的某种性质，例如聪明才智、努力水平或好运气，但是，一个人的成功或许也可以归功于其朋友和同事的网络。这其实是一个非常古老的问题：成功到底取决于你所知道的东西，还是取决于你所认识的人？

试想一下，一群科学家在一个实验室里一起进行研究工作。他们分享想法和知识，互相提建议。因此，某个科学家发表的学术论文、申请

的专利或取得科学突破的数量不仅取决于他自己知道些什么，而且也依赖于他与其他科学家的互动，他会受到他所认识的人的影响。我们要把环境特征和内在性质一起考虑，然后再来确定某位科学家的成功应该在何种程度上归因于自己的努力程度。

成功主要取决于人才的信念，有些投资公司雇用明星基金经理来操盘，但是往往无法取得预期的结果。经验证据表明，顶级投资者也依赖于向他们提供特定类型信息的同事网络。¹⁵放在更大的文献背景内，这个发现其实并不令人意外。许多文献（其中有一些是基于模型的）都表明，一个人在一个组织中的位置会影响他的成功。

当然，成功肯定与能力相关。一个能够让投资者获利数百万的商业创意当然是一个好主意，发表数百篇论文并获得无数奖项的科学家肯定具有很高的科研能力。但是另一方面，我们也不能否认，恰恰是那些在网络中占据了最核心位置的人做出了最大的贡献。我们可以使用介数和其他中心性测度来衡量一个人的位置。在网络中占据了介数很高位置的那些人，填补了著名社会学家罗纳德·伯特（Ronald Burt）所说的社区之间的“结构洞”（structural holes），我们可以使用特定的算法识别出这些结构洞。¹⁶由于能够从多个社区获取信息和思想，这些填补了结构洞的人拥有很大的权力和影响力。当然，要想去填补结构洞，你得有相当高的才华和能力才行。看到一个洞就跳下去并不算填补结构洞。要填补结构洞，你必须让社区的每一个人信任和理解你，你必须熟悉每个社区的知识库。

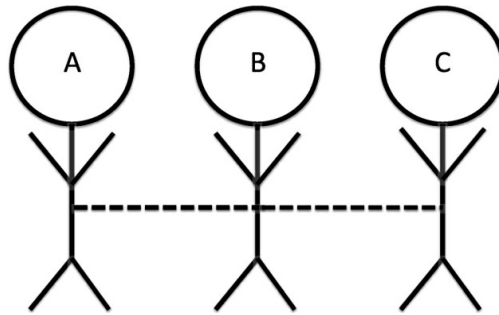
运用同样的原理，我们还可以对企业的价值进行评估并讨论国家的权力分配。我们可以将企业的价值视为内在价值，着重从资产负债的角度分析。还可以考察该企业运营的情境，例如它在供应链中的位置。与此类似，一个国家的权力取决于它的资源和联盟。无论是对于企业还是对于国家，它们的内在性质都是与连通性相关的。那些在网络中占据很高位置的人也拥有某些举足轻重的特性。

这里的分析以及大多数文献都将节点视为分析的单位。事实上，边同样也很重要。而且，从一个更广泛的角度来看，网络本身可能就是一个合适的分析单位。例如，如果教师网络能够让思想和信息在课堂之间顺畅流动，那么就可以改善教育成果，因此一个连通性很好的管理者可以有效地协调课程改革。与此类似，一个二年级的老师知道很多关于刚上三年级的学生的信息，而这些信息对三年级的老师非常有帮助。数学老师知道学生还有什么概念没有掌握好，这类信息可以帮助科学老师更好地设计课程。因此，优秀的学校都拥有强大的教师网络。这只是一个例子；在许多其他公共领域和私人领域，网络都可以改进我们的思考。¹⁷

迈尔森值和结构洞

填补结构洞的人能够将网络中的不同社区联系起来，从而具有更大的影响力。网络的各种统计数据，例如介数，都与能不能占据结构洞相关。对于一个人在网络中的影响力的另一个衡量指标是迈尔森值，它依赖于夏普利值的原理。为了计算迈尔森值，我们在网络上构建一个合作博弈。在这个合作博弈中，只允许存在包含了连接组件的联盟。

考虑排成一行的三个人。假设他们的位置代表他们的政治（意识形态）立场。如图所示，**B**位于中间。如果我们限定只允许直接相邻的人组成联盟，那么位于最左的**A**将不能与位于最右边的**C**连接，除非**B**也加入联盟。为了计算出每个博弈参与者的迈尔森值，我们先为所有可行的联盟分配增加值；再对每个可能的联盟计算夏普利值，将每个联盟视为一个单独的博弈。最后，将每个联盟博弈的夏普利值相加，就可以得到迈尔森值。



可能的次序：**ABC**，**BAC**，**BCA**，**CBA**排除掉的次序：**ACB**，**CAB**

举例来说，假设两个博弈参与者的任何联盟产出的价值为10，并且所有三个博弈参与者在一起时的产出的价值为14，那么就可以求出如下迈尔森值：博弈参与者A为3；博弈参与者B为8；博弈参与者C为3。¹⁸

中心性测度（例如介数）只以网络为基础，而迈尔森值则取决于价值函数。同时采用这两种测度，可以将一个人的权力对他在网络中位置的依赖性，与对他所发挥作用的依赖性分开。在这个例子中，三个博弈参与者的迈尔森值（3，8，3）与他们的介数得分（0，1，0）是完全相关的。但是情况并非总是如此，特别是对于更加复杂的网络和价值函数。

11 广播模型、扩散模型和传染模型

就像传染病的传染导致更多的传染病一样，信任的“传染”也可以促成更大的信任。

玛丽安·穆尔（Marianne Moore）

在本章中，我们运用广播模型、扩散模型和传染模型分析信息、技术、行为、信念和传染病在人群中的传播。这些模型在通信科学、市场营销学和流行病学的作用中发挥着核心作用。所有这三类模型都将整个人口划分为两个群组：知道或拥有某种东西的人与不知道或不拥有某种东西的人。随着时间的推移，个人会在这两个群组之间移动。有人会从业感者变为感染者，或者从不了解新产品新思想的人变成知情达意者。

串上某种传染病、购买了某种产品或掌握了某个信息的人的数量随时间推移而演变的曲线，也就是采用曲线的经验图形往往是凹的或S形的。人们如何获悉信息或怎样患上传染病——无论是通过广播传播还是扩散传播，决定了这种图形的形状。本章的主要内容就在于，将思想和传染病传播的微观过程与这些采用曲线的形状联系起来。为此，本章首先分析了广播模型。这种模型适用于人们从某个单一来源知悉思想或罹患传染病的情形。广播模型生成的图是r形的。然后，我们讨论扩散模型。在扩散模型中，传播始于接触，就像传染病在人与人之间传播时那样。扩散模型会产生S形曲线。

许多产品、应用程序、思想和信息都是通过广播传播和口碑传播的。我们可以同时允许广播和扩散来对这些情况建模，由此而得到的模型被称为巴斯模型（Bass Model），它在营销学中起着核心作用。巴斯模型会生成r形曲线还是S形曲线，则取决于广播过程和扩散过程之间的相对优势。

本章中讨论的最后一个模型来自流行病学的关于传染病传播的SIR模型（易感者、感染者和痊愈者模型）。这个模型包括了痊愈率，它能够刻画抵御传染病的免疫系统、突显性逐渐弱化的信息，以及被流行浪潮甩下的时尚行为等。SIR模型会产生一个临界点，在临界点上，产品或传染病性质的微小变化，就意味着失败与成功之间巨大分野。病毒毒力的轻微减少，就可以使大规模传染变为轻微的发病；歌曲传唱概率的小幅变动，就可能把一只热门的新乐队送上天堂或打入地狱，区别之大，恰如披头士乐队与一支在利物浦某个地下酒吧卖唱的乐队之间的区别。

广播模型

本章中介绍的所有模型都要假设存在一个相关人群，用 N_{POP} 表示。相关人群包括那些可能患上传染病、了解信息或采取行动的人。相关人群所指的并不是一个城市或国家的全部人口。如果我们要为连续主动脉缝合法的扩散建模，那相关人群就是指心脏外科医生，而不是居住在费城的所有人。

在任何时候，总会有些人患上了某种传染病、了解特定信息或采取了一定行动。我们将这些人称为感染者或知情者（用 I_t 表示），相关人群中除了感染者或知情者之外的其余成员则是易感者（用 S_t 表示）。这些易感者可能会感染传染病、了解信息或采取行动。¹相关人群的总人数等于感染者或知情者人数加上易感者人数的总和： $N_{POP}=I_t+S_t$ 。

广播模型

$$I_{t+1}=I_t+P_{broad}\times S_t$$

其中， P_{broad} 表示广播概率， I_t 和 S_t 分别等于时间 t 上的感染者（知情者）和易感者的数字

初始状态为 $I_0=0$ ，且 $S_0=N_{POP}$ 。

广播模型刻画了思想、谣言、信息或技术通过电视、广播、互联网等媒体进行的传播。大多数时事新闻都是通过广播形式传播的。这个模型的目标是描述一个信息源传播信息的过程，可以是政府、企业或报纸。它也适用于通过供水系统传播污染的情况。但是，这个模型不适用于在人与人之间传播的传染病或思想。由于广播模型更适合描述思想和信息的传播（而不是传染病的传播），所以我们在这里说知情者的人数，而不说感染者的人数。

在给定时间段内，知情者人数等于前一期的知情者人数加上易感者听到信息的概率乘以易感者人数。按照惯例，初始人口全部由易感者组成。要计算未来某期的知情者人数，只需要将知情者人数和易感者人数代入上述方程即可。由此得到的将是一个r形采用曲线。

想象一下：某个拥有100万居民的城市的市长宣布了一项新的税收政策。在他宣布之前，没有人知道这项政策。假设某人在任何一天听到这个新闻的概率等于30%（即， $P_{broad}=0.3$ ），那么第一天会有30万人听到这个新闻。在第二天，剩下的70万人中有30%的人，即21万人会听到这个新闻。在每一个时期，知情者的人数都会增加，并且以一个递减的速度增加，如图11-1所示。

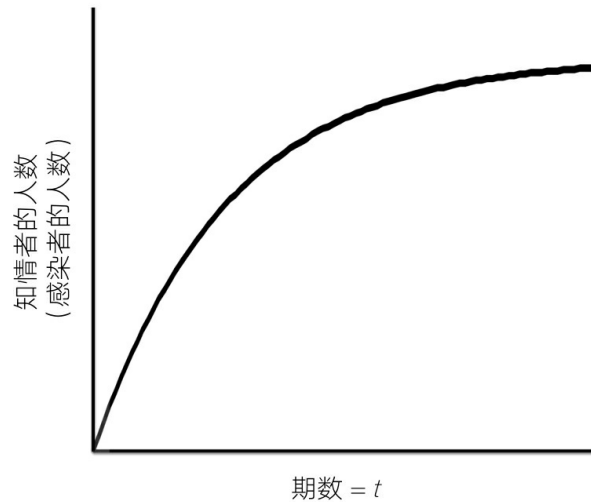


图11-1 广播模型产生的r形采用曲线

在广播模型中，相关人群中的每一个人最终都会知悉信息。如果有适当的数据，就可以估计出相关人群的规模。假设一家企业为练习太极拳的人推出了新设计的运动鞋，并在第一个星期就收到了20 000双鞋的订单。如果在第二个星期收到了16 000双鞋的订单，那么我们可以大致估计出他们最终的总销售量，也就是相关人群的规模为100 000。

用广播模型拟合销售数据

$$\text{第1期: } I_1 = 20\,000 = P_{\text{broad}} \times N_{\text{POP}};$$

$$\text{第2期: } I_2 = 36\,000 = 20\,000 + P_{\text{broad}} \times (N_{\text{POP}} - 20\,000)$$

$$\text{于是, 总销售量为}^2: P_{\text{broad}} = 0.2, P_{\text{broad}} = 100\,000$$

当然，对于根据仅有的两个数据点估计出来的任何结果，我们都不应该抱以太大的信心。这个模型无疑遗漏了许多现实世界的特征。人们既可能通过传媒获悉相关消息，也可能通过口耳相传听到消息，而且有些人可能会购买不止一双鞋子，或者可能存在针对潜在消费者的广告，等等。如果把这些因素都包括进去，估计出来的结果肯定会有所不同。尽管必须牢记这个注意事项，但是这个模型确实提供了一个粗略的估

计。这个企业不应该期望能够卖出200万双鞋，但是应该有信心可以卖出不止100 000双鞋。随着更多数据的出现，估计结果是可以得到改进的。如果第三个星期的销售额是13 000双（这等于模型预测的数量），那么这个企业对当初的预测可以寄予更大的信心。

扩散模型

大多数传染病，以及关于产品、思想和技术突破的信息，都是通过口口相传而传播开来的，扩散模型刻画了这些过程。扩散模型假设，当一个人采用了某种技术或患上了某种传染病时，这个人有可能将之传递或传染给与他接触的人。在传染传染病的情况下，个人的选择不会在其中发挥任何作用。一个人患上某种传染病的概率取决于诸如遗传、病毒（细菌），甚至环境温度等因素。在炎热潮湿的季节，疟疾的传播速度要比在寒冷干燥的季节快得多。

技术的传播则与采用者的选择有关，因此更有用的技术被采用的概率更高。但是在这里，我们并没有在模型中明确将这种情况选择考虑在内。这样一来，苹果智能手表的新潮性就发挥了与流感病毒同样的作用。

在这里，我们更看重的是信息的传播，因此我们将人们分为知情者或不知情的新人。如果新人与知情者相遇且信息在他们之间传播，那么新人就会变成知情者。这种事件的发生，因环境而异。生活在城市中的人，相遇的概率可能比生活在农村的人更高，同时也有更高的接触概率。非常吸引人眼球的新闻也比一般的新闻被分享的概率更高，例如，关于外星人降临登陆的新闻被分享的概率比关于M&M公司的椒盐卷饼重新上市的新闻更容易被分享。因此，我们可以将扩散概率（diffusion probability）定义为接触概率（contact probability）和分享概率（sharing probability）的乘积。我们可以根据扩散概率来构建模型，但是在估计或应用模型时，必须独立地跟踪接触概率和分享概率。

扩散模型假定随机混合（random mixing）。随机混合的含义是，相关群体中任何两个人接触的可能性都相同。对于这个假设，我们应该保持警惕。就描述幼儿园内传染病传播的扩散模型而言，这可能是一个准确的假设，因为幼儿园里儿童之间的相互接触是高频率的。但是，如果将它应用于城市人口则是有问题的。在城市中，人们并不是随机混合的。人们在一定的社区中生活，在一定的场所内工作，他们属于工作团队、家庭和社会团体，他们的互动主要发生在这些群体中。但是同时也不要忘记，一个假设要成为有用模型的一部分，其实不一定非得十分准确不可。因此，我们将继续使用这个假设，同时保持开放的心态，在需要改变的时候随时改变这个假设。

扩散模型

$$I_{t+1} = I_t + P_{diffuse} \times \frac{I_t}{N_{POP}} \times S_t$$

其中， $P_{diffuse} = P_{spread} \times P_{contact}$ 。

在这个模型中，与在传播模型中一样，从长期来看，相关人群中的每个人都会掌握信息。不同的是，扩散模型的采用曲线是S形的。最初，几乎没有人知情， I_0 很小。因此，能够与知情者接触的易感者人数也必定很小。随着知情者人数的增加，知情者与不知情者之间接触的机会增加，这又使知情者的人数更快地增多。当相关人群中几乎每个人都成了知情者时，新知情的人数会减少，从而形成了S形的顶部。技术的采用曲线通常也具有这种形状。例如，杂交种子的采用曲线虽然因州而异（艾奥瓦州采用杂交种子的速度比亚拉巴马州更快），但是所有州的采用曲线都是S形的。³

在广播模型中，根据数据估算相关人群规模是一件相当简单的事

情。采用者的初始数量与相关人群规模密切相关。与此相反，利用扩散模型的数据估计相关群体的规模可能会非常困难。产品销售量的增加，可能是由于一个很小的相关人群内部的高扩散概率，也可能是由于一个很大的相关人群中的低扩散概率。

图11-2显示了两个假想的智能手机应用程序的相关数据。在第一天，每个应用程序都有100人购买。在接下来的5天中，应用程序1拥有更高的总销量和更快的销量增长。如果没有模型，我们很可能会预测应用程序1拥有更大的市场。但是，用模型拟合这两组数据的结果表明，事实与我们猜想的恰恰相反。

上市日	智能手机应用程序 1	智能手机应用程序 2
1	100	100
2	136	130
3	183	169
4	242	220
5	316	286
	...	
365	1 000	1 000 000

图11-2 智能手机应用程序的两条采用曲线

应用程序1拟合的扩散概率为40%，相关人群规模为1 000人；而应用程序2的扩散概率为30%，相关人群规模为100万人。⁴事实上，只要再过几天，我们会观察到应用程序2的相关人群更大。但是，如果没有模型，如果不能根据前5天的数据来进行分析，我们就可能会对总销售额给出不正确的推断。

在使用扩散模型来指导行动的时候，我们必须将扩散概率分解为分享概率和接触概率的乘积。为了提高应用程序的销售速度，开发人员既可以设法提高人们相互接触的概率，也可以设法加大他们分享关于应用程序信息的概率。要想改变第一个概率是很困难的。为了增大第二个概

率，开发人员可以为带来了新注册用户的老用户提供一些激励，事实上，许多开发人员都是这样做的，比如游戏开发者可能会给带来了新注册玩家的老玩家奖励游戏积分。虽然这样做能够增加扩散速度，但是并不会影响总销量，至少根据这个模型来看不会有影响。如上所述，总销量等于相关人群的规模，而与分享概率高低无关，提高销售速度不会带来长期的影响。

大多数消费品和信息都是通过广播和扩散传播的。而巴斯模型则将这两个过程组合在一起了。⁵巴斯模型中的差分方程等于广播模型和扩散模型中的差分方程之和。在巴斯模型中，扩散概率越大，采用曲线的S形就越显著。电视、收音机、汽车、电子计算机、电话机和手机的采用曲线形状都是r形和S形的组合。

巴斯模型

$$I_{t+1} = P_{broad} \times S_t + P_{diffuse} \times \frac{I_t}{N_{POP}} \times S_t$$

其中， P_{broad} =广播概率， $P_{diffuse}$ =扩散概率。

SIR模型

到目前为止，在我们已经讨论过的模型中，一旦有人采用了一项技术，则永远不会放弃它。对于电力、洗碗机和电视等技术来说，确实如此：一旦采用之后，一般永远不会不采用。但这并不适用于所有通过扩散传播的事物，例如我们患上了某种传染病之后不久就会恢复健康，或者当我们采用了某种流行款式或参加了某项潮流运动之后（例如，某种时装或舞蹈），是可以放弃的。遵循惯例，我们将放弃所采用的某种事物的人称为痊愈者。由此产生的模型，即SIR模型（易感者、感染者、痊愈者），在流行病学中占据了中心位置。

由于这个模型起源于流行病研究领域，同时也因为考虑传染病的痊愈更为自然，因此我们以传染病的传播为例来描述SIR模型。为了避免过于复杂的数学计算，我们假设治愈传染病的人会重新进入易感人群，也就是说治愈传染病并不会产生未来对传染病的免疫力。

SIR模型

$$I_{t+1} = I_t + P_{\text{contact}} \times P_{\text{spread}} \times \frac{I_t}{N_{\text{POP}}} \times S_t - P_{\text{recover}} I_t$$

其中， P_{recover} ， P_{spread} ，和 P_{contact} 分别等于传染病的传播概率、接触概率和痊愈概率。

流行病学家对接触概率和传播概率会进行单独跟踪，我们也会这样做。接触概率取决于传染病如何从一个人传播到另一个人。艾滋病通过性接触传播；白喉通过唾液传播；流感病毒通过空气传播。因此，流感的接触概率高于白喉，白喉的接触概率又高于艾滋病。而且，在发生接触后，各种传染病的传播概率也会有所不同。白喉比SARS更容易传染给另一个人。

SIR模型会产生一个临界点，就是所谓的基本再生数 R_0 ，也就是接触概率乘以扩散概率与痊愈概率之比。某种传染病，如果 R_0 大于1，那么这种传染病就可以传遍整个人群，而 R_0 小于1的传染病则趋于消失。在这个模型中，信息（或者，在这个例子中是传染病）并不一定会传播到整个相关人群。能不能做到这一点取决于 R_0 的值。因此，像疾病控制中心这样的政府机构必须依据对 R_0 的估计来指导政策制定。⁶

基本再生数 R_0

$$R_0 = \frac{P_{\text{spread}} \times P_{\text{contact}}}{P_{\text{recover}}}$$

如表11-1所示，麻疹可以通过空气传播，因而它的再生数高于艾滋病，艾滋病只能通过性接触和共用针头传播。对 R_0 的估计假设人们不会为了应对传染病而改变行为。然而，当学校里虱子肆虐时，家长的反应可能是让孩子待在家中，以降低接触概率，还可能会剃光孩子的头发，减少接触发生时传播的可能性。这两种行为变化都会降低虱子传播的 R_0 。

表11-1 各传染病的基本再生数 R_0

	麻疹	脊髓灰质炎	艾滋病	流感
R_0	15	6	4	3

在没有疫苗的情况下，检疫是一个选择，但是成本很高。⁷如果存在疫苗，那么疫苗接种可以预防传染病传播。即便做不到每个人都接种疫苗，也可以预防传染病传播。必须接种疫苗的人的比例，即疫苗接种阈值（vaccination threshold），可以通过公式 $V \geq \frac{R_0 - 1}{R_0}$ 求出。我们可以从上述模型中推导出这个公式。⁸

疫苗接种阈值随 R_0 的增加而提高。例如，脊髓灰质炎的 R_0 为6，因此为了防止脊髓灰质炎的传播，疫苗必须覆盖5/6的人群。而麻疹的 R_0 为15，为了阻止麻疹的传播，疫苗必须覆盖14/15的人口。疫苗接种阈值的数学推导也为决策者提供了指引，如果接种疫苗的人数太少，这种传染病就会传播开来，因此政府接种疫苗的次数会超过模型估计的阈值。对于麻疹和脊髓灰质炎等 R_0 非常高的传染病，政府将努力保证所有人都接种疫苗。

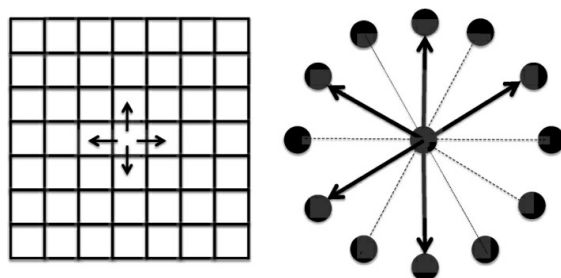
有些人担心疫苗有副作用，选择不参加疫苗接种计划。如果这些人只占人口的一小部分，那么其他人接种疫苗也可以防止这些人感染这种传染病，流行病学家将这种现象称为群体免疫力。选择不接种疫苗的人事实上是搭了其他接种疫苗的人的便车，对于搭便车的现象，我们将在

本书后面的章节中详细研究。⁹

R_0 、超级传播者，以及度的平方

假设随机混合 R_0 的推导如下：在每一步，人群中的个体随机相遇。如前所述，随机混合假设可能与通过空气传播的传染病或通过接触传播的传染病有关，但对于通过性行为传播的传染病则不太合理。

如果将SIR模型嵌入到网络中，就会观察到度分布对传染病传播的重要性。在这里，我们比较一下矩形网格网络（棋盘格）与中心辐射型网络。在矩形网格网络中，每个节点都连接到东、南、西、北的节点；而在中心辐射型网络中，则由一个中心节点连接到所有其他节点。



假设传染病会随机发生在某个节点上。我们在网络中设定 $p_{\text{contact}}=1$ ，以保证每个人都会与他所连接的每个人接触。在下一个时期，传染病可能以一个与传染病毒力相对应的给定概率，独立地扩散给每个邻居。

首先考虑矩形网格网络。在每个时期，传染病都可以扩散到东、南、西、北4个节点中的任何一个。我们预计，如果传染病传播的概率超过了1/4，传染病就会蔓延。展望未来的一个时期就会看到，如果一个新节点患上了传染病，那么这个节点有3个可能患上传染病的邻居。如果原始节点的北部和东部的两个邻居患上了传

染病，那么传染病可能传播的节点就会达到6个。因此，这种网络似乎对传染病传播的潜力发挥没有太大影响。

接下来，考虑中心辐射型网络。第一个患上传染病的节点可能是中心节点，也可能是外围节点。如果中心节点患上了某种传染病，那么它可以将传染病传播到任何一个其他节点。我们预计这种传染病会扩散，即便传播的概率很低也是如此。如果是一个外围节点患上了传染病，那么唯一可能被传染的节点就是中心节点。正如在前面讨论过的，如果中心节点患上了传染病，那么即使传播的可能性很小，传染病也会蔓延。

对于中心辐射型网络， R_0 携带的信息量很有限，因为如果中心节点患上了传染病，传染病就会传播开来。流行病学家们将位置在度很高的中心节点上的人称为“超级传播者”（superspreaders）。超级传播者加速了艾滋病和SARS的早期传播。¹⁰超级传播者不一定是社交明星或“人脉”特别广的人，可能从事某种特定的行业职业，比如收费站的收费员、银行柜员、牙科医生，这类职业使他会与属于不同社交网络的人接触。生活在19世纪与20世纪之交的“伤寒玛丽”（Typhoid Mary）只是纽约的一名上门服务的厨师。她从这一家再到另一家，将伤寒感染给每一个接触者。当她被确认为传染源之后，就被强制隔离了。

为了推导出高度数节点的影响，我们首先要注意到一个事实：高度数节点不但能够更快地传播传染病，而且会更快地患上传染病。如果一个人朋友的数量是另一人的三倍，那么他患上传染病的可能性也是后者的三倍，同时传播这种传染病的可能性也是后者的三倍。因此，他对传染病传播的总贡献将是另一个人的九倍。因此，节点对传染病（或思想）传播的贡献与节点的度的平方相关。如果节点A的度数是节点B的K倍，那么节点A传播传染病的可能性是节点B的K倍，同时传染病传播到节点A的概率也是节点B的K

倍。因此，节点A对传染病传播的总效应将是节点B的 K^2 倍。这种现象被称为度的平方。

一对多

尽管SIR模型原本是用来分析传染病传播的，但是我们也可以将它应用于所有先通过扩散传播，然后趋于消失的社会现象，例如书的销售、歌曲的流行、舞步的风行，“热词”的传播、食谱和健身方法的流传等。在这些情形下，我们也可以估计接触概率、传播概率和“痊愈”概率，以及基本再生数 R_0 。这个模型意味着，这些概率只要发生了微小的变化，就可以使 R_0 移动到高于零的水平，从而造成成功与失败之间的天壤之别。

确实，成功可能取决于非常微小的差异，正如美国作家约翰·厄普代克（John Updike）在描述棒球明星特德·威廉姆斯（Ted Williams）最后一次击球时所说的：“一件事情做得很好与搞砸了之间，只有极其细微的差异。”¹¹假设你构思了一个新的笑话；只要让这个笑话更有趣一点点，就可能会把 R_0 推到高于1的水平，从而使这个笑话广泛传播开来。同样的逻辑也适用于想法的“黏性”。如果一个想法能够在人们的思维中再坚持一小段时间，那么他们摆脱它的“痊愈”率就会降低，从而提高了 R_0 。

当然，并不是所有情况都会位于阈值上。披头士乐队拥有巨大的才华，他们的 R_0 肯定超过了1，尽管这只是一个猜想。对于现在的流行歌星，我们可以使用互联网下载量来估计他们的 R_0 。流行歌星贾斯汀·比伯（Justin Bieber）的 R_0 估计为24，这就是说，他的传染“毒力”比麻疹更强。¹²

在SIR模型中，我们推导出了两个关键阈值，即 R_0 和疫苗接种阈值。这两个阈值都是属于敏感依赖于环境的临界点，环境（情境）中的

微小变化都会对结果产生很大的影响。这种临界点不同于直接临界点（direct tipping point）。在直接临界点，特定时刻的微小行动会永久性地改变系统的路径。直接临界点出现不稳定的点是，例如当球停在山顶上时。在任一方向上稍微推一下，都将会使球从山的这一侧或另一侧滚下去，这个小小推动是一个直接的倾覆。¹³

而在依赖于环境的临界点上，参数的变化会改变系统的行为方式。在直接临界点上，未来的结果轨迹急转直下。折弯，例如由扩散模型产生的S形采用曲线中的第一个弯曲，不满足这两种临界点的定义。采用曲线中的折弯对应于斜率增长率最大的点。在那一点上，扩散一发不可收拾，但是并没有发生倾覆。

图11-3显示了Google+发布后前两个星期的用户数。¹⁴从图中可以看出，在发布6天后，出现了一个折弯。在那一点上，扩散的过程正在顺利展开。从两个星期内就获得了超过1 600万用户这个结果来看，我们不能说Google+很早就陷入了困境，更不能说它在第6天就出现了直接临界点。将倾覆与急剧上升（下降）混淆起来，导致临界点这个术语被过度滥用了。新闻媒体和互联论坛上所说那些临界点，几乎有很少符合正式定义的。

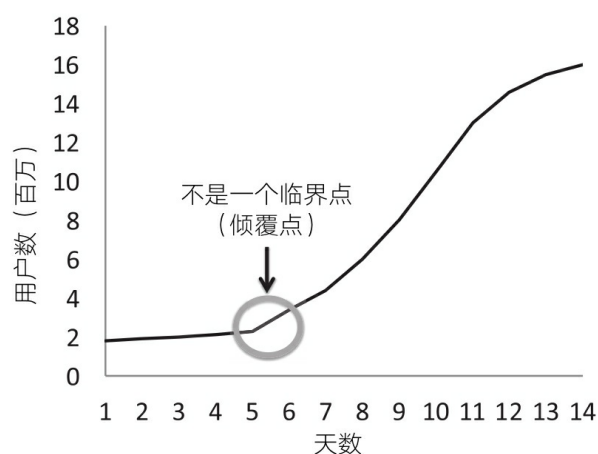


图11-3 Google+用户数量上的一个折弯点（不是一个倾覆点）

我们不妨将肥胖症视为一种流行病来考虑。尽管人们不会像患上感

冒那样感染肥胖症，但是他们可能会受到某种社会影响而做出一些容易导致肥胖的行为。¹⁵要想扭转肥胖这种流行病，我们必须降低它的 R_0 。而要降低 R_0 ，则可以通过降低接触概率或者提高分享概率和痊愈概率来实现。当然，在很多方面，用SIR模型来研究肥胖症的传播、学校辍学率或犯罪率，并不比经济学模型或社会学模型更好。它只是一个不同的模型，因此会给出不同的解释和预测，它也可能指向不同的行动或政策。它扩大了我们的模型集合，帮助我们更好地理解世界，但它不是解决问题的灵丹妙药。

在将广播模型、扩散模型和传染模型应用于社会现象时，我们可能会发现某些假设是成立的，而其他一些假设则不能成立。例如，在某种传染病的传播中，每一次接触导致该传染病传播的概率是独立的。但是在社交领域，由于采用本身也是一种选择，因此传染有可能会因更多的接触（曝光）而变得更有可能。流感不是我们选择的，我们只是得了流感。但是我们会选择买紧身牛仔裤，随着越来越多的人穿上了紧身牛仔裤，我们所有人都更可能穿紧身牛仔裤。类似的逻辑也适用于分析社交运动的参与率、新技术的采用率，甚至分析文身的人有多少。在这些情况下，我们可能必须对基本模型进行修正，以允许每次接触的采用概率会随着接触次数的增多而增大。¹⁶信仰或信任行为的“传染”也是如此。这种修正，在扩大模型的应用范围时通常是必不可少的。

12 熵：对不确定性建模

信息是不确定性的解。

克劳德·香农（Claude Shannon）

在本章中，我们讨论熵。熵是对不确定性的一个正式测度。利用熵，我们可以证明不确定性、信息内容与惊喜之间的等价性。低熵对应于低不确定性，同时揭示的信息很少。如果某个结果发生在低熵系统中，例如太阳从东方升起，我们并不会感到惊讶。而在高熵系统中，比如在抽奖时抽中了某个数字，结果是不确定的，并且实现的结果能够揭示信息。在这个过程中，我们经历了惊喜。

利用熵，可以比较不同的现象。我们可以判断新西兰的选举结果是不是比联合国对谴责某个国家的方案的投票结果更不确定，还可以将股票价格的不确定性与体育赛事结果的不确定性进行比较，也可以利用熵的概念来区分四类结果：均衡、周期性、复杂性和随机性。我们可以将看似随机的复杂模式和真正的随机性区分开来，并且可以分辨出哪些现象看起来像是有一定模式的，但事实上是随机的。

我们还可以使用熵来表征分布。在没有控制或调节力量的情况下，一些群体可能会向最大熵漂移。给定特定的约束条件，例如不变的均值或方差，就可以解出最大熵分布。最大熵分布的结果还可以用来证明某些分布比其他分布更优，从而能够对我们在建模时的选择起到指导作用。

本章分为五个部分。在第一部分中，我们讨论对信息熵（information entropy）的直观认识，然后给出信息熵的正式定义。在第二部分中，我们描述了关于熵的公理基础。在第三部分中，我们讨论如何使用熵来区分均衡、周期性、随机性和复杂性。在第四部分中，我们

研究了会在给定约束条件下产生最大熵的系统。最后，我们探讨了这样一个问题：为什么在有的时候，我们更喜欢复杂性而不是均衡。

信息熵

熵是用来度量与结果的概率分布相关的不确定性的。因此，它也可以衡量意外。熵与方差不同，方差度量一个数值集合或数值分布的离散程度。不确定性与离散程度有关，但是两者并不是一回事。在具有高不确定性的分布中，许多结果的概率都是有意义的，这些结果并不一定有数值，具有高离散度的分布则只是具有一些极端的数值。

通过比较具有最大熵的分布与具有最大方差的分布，可以将这种区别鲜明地呈现出来。给定取值范围为从1到8的整数的若干结果，能够使最大化熵的分布对每个结果赋予相同的权重。¹而能够使方案最大化的分布则是以1/2的概率取值1、以1/2的概率取值8（图12-1）。

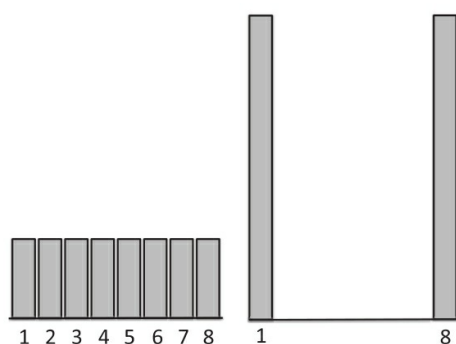


图12-1 最大熵与最大方差

熵是在概率分布上定义的。因此它可以应用于非数值数据分布，例如森林中鸟儿的种类或不同口味果酱的市场份额。熵在数学上等于概率与它们的对数之和的相反数。这个数学公式听起来似乎很复杂，但是事实并非如此。

我们先从信息熵这种特殊情况开始讨论。对于信息熵，可以把它理

解为根据随机抛硬币的结果来衡量不确定性的一种方法。假设每个家庭都只有两个孩子，男孩和女孩的可能性相同。某个家庭的孩子们的性别列表（按出生顺序排列）相当于抛两次硬币。因此，结果分布的信息熵为2，因为它对应于两个随机事件。其信息内容也等于2，因为我们只要提出两个“是或否”问题要求他们回答，就可以掌握结果。

与此类似，在有三个孩子的家庭中，性别列表相当于抛3次硬币。要了解这样的家庭的孩子的性别，也只需要提出三个“是或否”问题。同样的逻辑适用于任何数量的儿童。在一般情况下，要了解 N 个孩子的性别，只需要提出 N 个“是或否”问题。

这里需要注意的是，这 N 个问题区分出了 2^N 种可能的出生顺序。这种数学关系是理解熵测度的关键所在： N 个二元随机事件会产生 2^N 个可能的结果序列，并且，与之等价，我们可以通过提出 N 个“是或否”问题知悉结果序列。这也意味着，信息熵将不确定性水平（和信息内容） N 分配给了 2^N 个结果上的一个等可能分布。

于是，挑战变成了如何用数学公式刻画这种关系。每个结果序列的概率均为 $\frac{1}{2^N}$ 。要将这个数值转换为 N ，需要一个相当复杂的数学公式

$$N = -\log_2\left(\frac{1}{2^N}\right)$$
。我们可以将这个公式推广到任意概率的情形下。如果某个结果序列出现的概率为 p ，就分配一个不确定性 $\log_2(p)$ ，它近似于识别该序列所需要提出的“是或否”问题的数量。为了计算出一个分布的信息熵，我们只需求得所有结果（或者像在前面那个例子中那样的结果序列）需要提出的问题的期望数量的平均值。

信息熵

给定一个概率分布 $(p_1, p_2, \dots, p_N)$ ，信息熵， H_2 等于：

$$H_2(p_1, p_2, \dots, p_N) = - \sum_{i=1}^N p_i \log_2(p_i)$$

注：上面的下标2表示使用的是以2为底的对数。

乍一看，这个数学公式带来的混淆似乎比它所能澄清的还要多。通过举例说明，应该能够使这个公式更加直观。想象一下这种情况：第一胎是女孩的家庭不再生任何孩子，而第一胎是男孩的家庭则还要再生两个孩子。从而，所有家庭中将有一半家庭只有一个女孩。而另一半家庭则等可能地分别属于如下四个结果之一：三个男孩；两个男孩与一个女孩；一个男孩与两个女孩；一个男孩一个女孩与一个男孩。这四种结果中的每一种出现的概率均为1/8。

信息熵等于我们想了解一个家庭的子女排列状况时必须提出的“是或否”问题的期望数量。我们首先会问，第一个孩子是不是女孩，回答“是”的概率为1/2。如果是这个答案，那么就不需要继续问下去了。因此有一半的时间，我们只需问一个问题。我们可以把这写成

$-\frac{1}{2} \log_2\left(\frac{1}{2}\right)$ 。如果答案是否定的，那么我们还必须再提出两个问题，于是总共要问三个问题。这四种情况中的每一种都以1/8的概率出现，因此每种情况对信息熵的贡献为1/8×3，我们对每种情况可以写出

$-\frac{1}{8} \log_2\left(\frac{1}{8}\right)$ 。从而，信息熵等于2，即上述五项的总和。³虽然这里使用的对数和负号可能会让有些人觉得困扰，但是直观含义仍然是非常清楚的：信息熵就对应着“是或否”问题的期望数量。如果我们不得不提出很多问题，那么分布就是不确定的。而知道了结果，也就揭示了信息。

熵的公理基础

公理基础：熵

$$H_a(p_1, p_2, \dots, p_N) = - \sum_{i=1}^N p_i \log_a(p_i), \text{ 其中, } a > 0。$$

这种熵测度是唯一满足以下四个公理的测度：

对称性：对于任何转置概率，都有连续函数 $H(\sigma(\vec{p})) = H(\vec{p})$ 。

最大化：对于所有 N ， $H(\vec{p})$ 在 $p_i = \frac{1}{N}$ 处最大化。

零性： $H(1, 0, 0, \dots, 0) = 0$ 。

可分解性：如果 $\vec{P} = (p_{11}, p_{12}, \dots, p_{nm})$

$$H(\vec{P}) = H(P_1, P_2, \dots, P_N) + \sum_{i=1}^N H(Q_{P_i})$$

其中， $P_i = \sum_{j=1}^m p_{ij}$ 且 $Q_{P_i} = (\frac{p_{i1}}{P_i}, \frac{p_{i2}}{P_i}, \dots, \frac{p_{im}}{P_i})$ 。

为了得到熵的一般表达式，我们采用公理化的方法。数学家克劳德·香农对他给出的这种测度施加了四个条件。前三个条件很容易理解，它必定是连续的和对称的，而且在所有结果以相同的概率发生时最大化，同时在某些结果上等于零。第四个条件可分解性则要求在具有 m 个子类别的 n 个类别上定义的概率分布的熵，等于各类别上的分布的熵与每个子类别的熵的总和。香农证明，有一类熵测度是唯一满足这些公理的测度。虽然这里的分布的乘积是一个不那么直观的自然假设。例如，在结果是两个独立事件的乘积的情况下，这意味着联合事件的信息内容是每个事件单独发生时的信息内容的总和。

正如夏普利值的公理基础一样，这些公理对存在性的贡献大于它们本身的合理性。聪明的数学家总是可以构造出能唯一定义一个函数的公理。香农的前两个公理很难质疑。有的人可能吹毛求疵地指责，将已知分布的不确定性设置为零过于任意了，但这只是一个适当的基准，另一种可能性是将已知分布的不确定性指定为1。⁴可分解性虽然解释起来不

是很容易，但是也很难去挑战它。两个组合随机事件的不确定性理应等于每个事件的不确定性之和。总的来说，这些公理不仅仅是可辩护的，事实上，它们是难以辩驳的。

利用熵区分结果类别

我们现在阐明，如何利用熵测度来对经验数据进行分类，并在计算机科学家、数学家斯蒂芬·沃尔弗拉姆（Stephen Wolfram）给出的四大类别的框架下建模：均衡、周期性、随机性和复杂性。⁵在沃尔弗拉姆的这个分类中，放在桌子上的铅笔处于均衡状态，绕太阳运转的行星处于循环当中，抛硬币的结果序列是随机的，纽约证券交易所的股票价格也是近似随机的（我们在下一章中将会说明原因）。最后，一个人大脑中的神经元发放则是复杂的：它们既不会随意发放，也不会以某个固定的模式发放。图12-2以图形方式呈现了这四个类别。

平衡结果没有不确定性，因此其熵等于零。周期性过程具有不随时间变化的低熵。当然，完全随机过程具有最大的熵。复杂性具有中等程度的熵，因为复杂性位于有序性和随机性之间。虽然熵在两种极端情况下能够为我们给出明确的答案——均衡和随机性；但是这并不适用于周期性和复杂性的结果。在这些情况下，通常还必须善用我们的判断力。

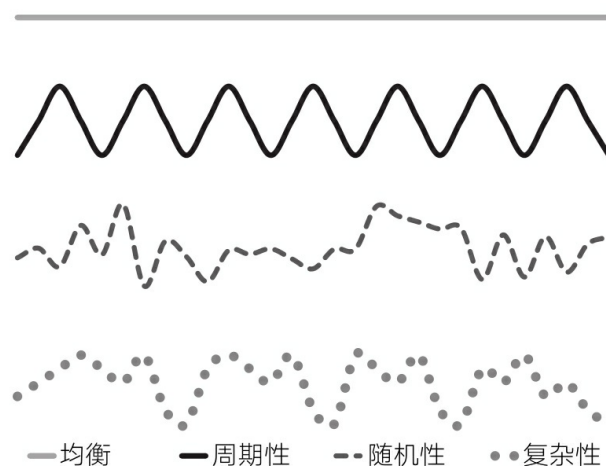


图12-2 沃尔弗拉姆的四种类别

为了对时间序列数据进行分类，我们需要先计算出不同长度的子序列中的信息熵。假设，有个人会把他每天戴的帽子的类型一一记录下来。假设他只在两种帽子之间进行选择，一种是贝雷帽，记为B，另一种是浅顶软呢帽，记为F。这样过了一年，他对帽子的选择生成了一个有365个事件的时间序列。我们先计算长度为1的子序列的熵，也就是说，先计算戴每种类型帽子的概率的熵。假设他喜欢这两种类型的帽子的程度相同，那么长度为1的子序列的熵等于1。因此，我们可以先把均衡排除掉，因为他会改变他的选择，但是其他三种类别中的任何一种都是可能的。

为了确定类别，我们接下来计算长度为2到6的子序列的熵。如果所有都具有最大的熵，那么我们可以将简单的周期性排除掉。假设当我们考虑更长的序列时，熵会缓慢增加，直到达到最大值8为止。换句话说，无论子序列有多长，熵都不会超过8。熵为8相当于256个结果的等可能分布，这不可能是一个简单的循环。熵为8更可能代表具有特定结构和模式的复杂过程序列。我们不能确定地说，这个时间序列是复杂的。一种可能的情况是，这个人试图做到随机化，但是却失败了。

最大熵和分布假设

在很多情况下，我们建模时都必须把不确定性包括进来；因而作为建模者，必须对有关的分布做出假设。这里的原则是，我们要尽量避免做出任意特殊假设（ad hoc assumption）。也许，我们对产生分布的过程已经有了一些了解。如果是这样，通常可以运用逻辑—结构—功能方法，推导出该过程产生的统计结构。

例如，假设我们想要对一个房地产拍卖中的所有拍卖对象的总价值的分布做出一个假设。总价值等于各个项目的价值总和。因此，我们可

可以根据中心极限定理假设这是一个正态分布。对于一栋房子的可能价值，我们也可以假设一个正态分布，因为房屋的价值取决于它的多个性质：卧室的数量、浴室的数量和占地大小等。

对于艺术珍品或稀有手稿的可能价值，正态分布却可能没有意义。在这些情况下，我们对决定它们价值的过程几乎一无所知。一种方法是假设一个具有最大不确定性的分布，即最大熵分布。

最大熵分布的形状取决于各种约束条件。正如我们已经看到的，如果假设了一个最小值和一个最大值，那么均匀分布会使熵最大化。教科书和学术期刊中的许多社会科学模型都假设均匀分布，我们可能会质疑这个假设，因为均匀分布在现实世界中确实很少出现。然而，无差别原则（principle of indifference）可以证明假设均匀分布的合理性。如果只知道范围或可能集，那么就应当予以无差别的对待。

在某些情况下，我们可能知道分布的均值，也知道所有值都必定是正数。给定这些约束条件，最大熵分布必定具有长尾，因为我们要将分布置于更多的值上，从而必须使少数高值结果与许多低值结果保持平衡。不难证明，熵最大化分布是一个指数分布。因此，如果我们正在构建一个模型，需要假设网站点击量或市场份额的分布形式，那么在没有可用数据的情况下，指数分布是一种自然的假设。

如果我们确定了均值和方差（并且允许出现负值），那么最大熵分布则是正态分布。这里的逻辑与前一种情况类似。为了创造更多的不确定性，我们创造了一些极端值，在这里，可以平衡正值和负值，而不用改变均值。但是，这样做会增大方差，因此我们必须在均值附近添加更多值，从而创造出钟形曲线。

我们可以在逻辑—结构—功能框架内解释这些最大熵分布。如果我们认为在给定的社会、生物或物理环境中，某个微观层面的过程能够最大化熵，那么我们应该期待上面这些分布中的某一个会出现。或者也可以假设一个微观过程，并能够证明熵在增加。如果是这样，上述分布中

的某一个也会涌现出来。

最大熵分布

均匀分布：给定范围 $[a, b]$ ，使熵最大化。

指数分布：给定均值 μ ，使熵最大化。

正态分布：给定均值 μ 和方差 σ^2 ，使熵最大化。

我们也可以将这些结果解释为探索性的。我们可能会遇到一些指数分布或正态分布的数据。虽然没有“义务”去追问某种潜在的行为是否会在一定约束条件下使熵增加，但这样做确实可以帮助我们获得一些新的洞见。在本书前面的章节中，我们利用中心极限定理解释了物种的高度、重量和长度为什么会服从正态分布。在这里，我们再给出一个不同的、基于模型的解释：如果一种突变能够最大化熵（以便探索最好的生态位），并且假设平均规模和总离散度是固定的，那么规模的分布将会是正态的。关键不在于这种最大熵方法是不是提供了一个更好的解释，而在于给定约束下最大化熵必定会导致正态分布。因此，当我们看到正态分布时，它可能是最大化熵的结果。

熵的实证含义和规范含义

前面我们已经讨论了，熵如何衡量不确定性、信息和惊喜，如何与测量离散度的方差不同，以及如何有助于我们对不同类别的结果进行分类和比较。在本书第13章和第14章中研究随机游走和路径依赖时，还会利用熵来识别随机性并测量路径依赖的程度。事实上，我们可以将熵测度用于任何实际应用，可以用它来衡量对金融市场的干预是增加了还是减少了不确定性，可以检验选举、体育赛事或博彩中的结果到底是不是随机的。

在这些应用中，熵都是作为一个实证的衡量标准来使用的。它告诉我们世界是什么样的，而不是世界应该是什么样。一个系统中的熵的本质，不能简单地说好，也不能简单地说不好。我们想要多少熵，取决于具体情况。在制定税法时，我们可能需要一种均衡行为模型，并不希望有随机性。在规划城市时，我们可能会希望看到复杂性，均衡或者周期性都会显得过于平淡。我们希望一个城市充满生机活力，为偶然的相遇和互动提供无限机会。在这种情况下，更多的熵会更好，但是又不能太多。我们不喜欢随机性，随机性会使计划变得非常困难，并可能导致我们的认知能力崩溃。最理想的情况是，世界会产生适度的复杂性，以保证我们生活在一个有趣的时代。

建筑师克里斯托弗·亚历山大（Christopher Alexander）证明，诸如强中心、厚边界和非独立这类的几何属性，能够生成复杂的生活建筑、社区和城市。⁶亚历山大渴望城市和生活空间中的复杂性。中央银行的规划者可能不太喜欢复杂性，在金融市场中，他们可能更喜欢可预测的均衡结果。不过幸运的是，使用模型，我们既可以探索复杂性，也可以讨论均衡的可能性。

13 随机游走

醉鬼能找到回家的路，但是一只醉酒的小鸟可能永远回不了家。

角谷静夫（Shizuo Kakutani）

在本章中，我们讨论两个来自概率论和统计学的经典模型：伯努利瓮模型（Bernoulli urn model）和随机游走模型。¹这两个模型都描述了随机过程，即使看上去它们似乎在生成某种复杂的结构。如果不收集数据，随机性是很难辨别的。我们经常想当然地以为能够在选举结果、股票价格和体育赛事得分中总结出一定的模式，但这只是一厢情愿。借用学者、风险分析师纳西姆·塔勒布（Nassim Taleb）的一句俏皮话来说，我们都被随机性所惑，不过是一些“随机漫步的傻瓜”！²

伯努利瓮模型描述了产生离散结果的随机过程，例如抛硬币或掷骰子。这个模型在几个世纪以前出现时，是为了解释赢得赌注的概率，现在已经在概率论中占据中心位置。随机游走模型就是建立在伯努利瓮模型的基础上的，保持了正面和反面的总数。这个模型可以刻画液体和气体中粒子的运动，动物在物理空间中的活动，以及从出生到童年人体身高的增长，等等。³

本章首先简要介绍伯努利瓮模型，并对条纹长度（length of streaks）进行了分析。然后描述随机游走模型，我们将会了解到，一维和二维随机游走会无限次地回到起点，而三维随机游走则可能完全不需要回到起点。我们还会了解到，对于一维随机游走，回到零点之间的时间间隔分布遵循幂律分布。对于这个发现，有人可能认为它除了满足人们的好奇心之外没有什么用，但事实上，它可以解释物种和企业的生命周期。我们还将使用随机游走模型评估有效市场假设，并用它来确定网

络规模。

伯努利瓮模型

伯努利瓮模型由一个装了灰球和白球的瓮组成。从瓮中抽取的球代表随机事件的结果。每次抽取都与之前和之后的抽取无关，因此我们可以应用大数定律：从长远来看，抽出每种颜色的球的比例将会收敛到这个球在瓮中的比例。当然，这并不意味着从一个装了7个白球和3个灰球的瓮中抽取1 000次，将会恰好抽出700个白球，它的意思是抽取出来的白球比例会收敛到70%。⁴

伯努利瓮模型

每一次，从一个装了 G 个灰球和 W 个白球的瓮中随机抽取一个球，结果等于抽取出来的球的颜色。在下次抽取之前，球要先放回瓮中。令 $P = \frac{G}{(G+W)}$ 表示灰球的比例。在抽取 N 次的情况下，可以计算出抽取出来的灰球的期望数量 N_G ，及其标准差 σ_{N_G} ：

$$N_G = N \times P, \text{ 以及, } \sigma_{N_G} = \sqrt{N \times P \times (1 - P)}$$

伯努利瓮模型的结果产生了可预测长度的条纹。在灰球和白球数量相等的瓮中，抽取出白球的概率等于1/2，连续抽取出两个白球的概率等于1/2乘以1/2，以此类推。一般情况下，如果瓮中白球的比例为 P ，那么连续抽取 N 个白球的概率等于 P^N 。通过计算概率，我们可以评估某种条纹是不是有可能出现（尽管很令人吃惊），或是几乎完全不可能（因而基本上可以肯定“有诈”）。当一名篮球运动员连续9次投中了三分球时，只是有热手效应吗？或者，我们是否应该期待有一个这种长度的随机序列？数学计算表明，一个很优秀的三分投手在长达10年的职业生涯中，也几乎完全不可能连续9次投中三分球。⁵

我们可以进行类似的计算以确定投资者是幸运、能力出众还是在欺诈。自1965年至2014年，由沃伦·巴菲特（Warren Buffett）经营的集团伯克希尔哈撒韦公司（Berkshire Hathaway），在50年中有42年的表现优于市场。1964年伯克希尔哈撒韦公司的1美元在2016年的价值已经超过了1万美元，而投资标准普尔500指数的1美元价值大约为23美元。如果伯克希尔哈撒韦公司有50%的机会击败市场，那么它在50年来的表现应该超过市场的25倍，标准差为3.5年 ($3.5 \approx \sqrt{50 \times \frac{1}{2} \times \frac{1}{2}}$)。而事实上，伯克希尔哈撒韦公司击败市场的实际年数大约高于均值四个标准偏差，这是一个概率仅有百万分之一的事件，因此，我们可以排除这完全是运气的可能。由于伯克希尔哈撒韦公司定期公布它的投资，所以也可以排除欺诈的可能。与此相反，前纳斯达克主席、美国历史上最大的诈骗案制造者伯纳德·麦道夫（Bernard Madoff）从来不透露他的投资情况，如果客户要求投资透明度的话，麦道夫是不可能连续几十年取得“成功”，连续几十年得到正回报的。⁶

随机游走模型

接下来讨论简单随机游走模型，它建立在伯努利瓮模型的基础上，并将过去结果的和保持下来。我们将初始值，也就是模型的初始状态设置为零。如果我们抽取出一个白球，就在总数上加1；如果抽取出一个灰球，就从总数中减1。模型在任何时候的状态都等于先前结果的总和，也就是抽取出来的白球总数减去抽取出来的灰球总数的值。

简单随机游走模型

$$V_{t+1} = V_t + R(-1, 1)$$

其中， V_t 表示时间 t 上的随机游走值， $V_0=0$ ， $R(-1, 1)$ 是一个可能等于-1或1的随机变量。在任何时间段内，这个随机游走的期

望值都等于零，且标准差为 $\sqrt{t}$ ，其中的 t 等于周期数。⁷

图13-1给出了一个简单随机游走。这幅图看上去似乎有一个模式：先是一个长期下降的趋势，然后是一个上升趋势；在上升过程越过零线时出现了一个适度的崩溃。但这个模式只是偶然发生的。

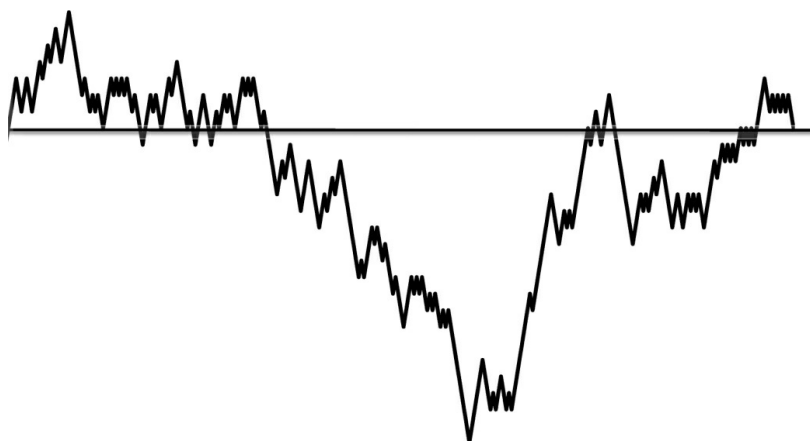


图13-1 一个300周期的简单随机游走

简单随机游走既是周期性的（会无限次地返回零点），又是无界性的（会超过任何正的或负的阈值）。如果等待足够长的时间，随机游走会高于正的1万、低于负的100万，也会无限次地穿过零线。此外，返回零点所需的步数分布满足幂律。⁸在大多数时候，返回零只需几步。所有游走中，有一半是两步返回的，然而有些游走需要很长时间才能返回。鉴于随机游走的无界性，这必定是真的。一个超过100万步阈值的游走，需要超过200万步才能到达那里并返回零点。

幂律分布结果还有一个意想不到的应用领域。如果我们将企业的销售水平或员工规模建模为随机游走，那么企业的生命周期就会成为一个幂律分布。更准确地说，当销售强劲时，企业会新招聘一名员工；当销售不佳时，会解雇一名员工；当不再拥有任何员工时，企业也就“寿终正寝”了。这样一来，返回次数的分布就等于企业生命周期的分布，而且是一个幂律分布。再者，就其第一近似而言，企业的生命周期是一个

幂律。⁹我们可以应用相同的逻辑来预测生物分类单元（界，门，纲，目，科，属和种）的寿命。如果某个分类单元的成员数量遵循随机游走，例如，如果某个属中的物种数量随机地上下变化，那么，这个分类单元的大小就应该满足幂律。这方面的数据支持了这个模型的预测。¹⁰

对于随机游走模型，还可以做这样一个类比：将随机游走视为冰川沿着地面的移动。根据模型的预测，冰川湖泊的大小分布将满足幂律。每一次，当冰川落到了陆地表面以下又返回顶部时，就会形成一个直径等于返回时间的湖泊。在这里，相关数据再一次与模型基本对应。¹¹

这个基本随机游走模型可以通过多种方式加以修正。我们可以创建一个正态随机游走（normal random walk）。在正态随机游走中，每一周期的值的变化都服从正态分布。正态随机游走不会完全回到零点，但它会无限次地穿过零点。

我们还可以令某一种结果比另一种结果更有可能发生，从而创建一个有偏差的随机游走。我们可以利用这种有偏差的随机游走模型来预测在博彩中获胜的概率。轮盘赌中，在红色结果上下注时赢的概率等于9/12。¹²我们可以将赌轮盘赌的总收益或总损失建模为这样一个随机游走：增加1的概率为9/19（大约47.4%），而减少1概率则为10/19。那么在下注100次之后，预期损失为5美元，标准差为10美元。这也就是说，我们可以在95%的置信水平上，认为损失不超过25美元、收益不超过15美元。在下注1万次之后，预期损失等于526美元，标准差为100美元。因此，在95%的置信水平上，我们的损失介于325美元与725美元之间。¹³同样，在下注1万次之后，我们还能赢是一个相当于超过均值5个标准偏差的事件，也就是说我们赢的可能性不到百万分之一。因此，要想在轮盘赌中赢，应该做的事情是下一个大赌注而不是下很多个小赌注。

一些体育比赛，例如篮球比赛，可以建模为两个有偏差的随机游

走。在球场上，每支球队在每次攻守中都有可能得分。这个概率可以根据一支球队的进攻能力和对方球队的防守能力来估计。我们将球队在球场上的“行程”模拟为一个随机事件。每支球队的得分对应一个随机游走值，得分较高的球队更有可能获胜。来自NBA的数据分析表明，实际比赛结果与这个模型匹配得相当好。只有当一支球队获得了巨大的领先优势时，得分才会偏离随机性，在那种情况下领先优势继续扩大的可能性低于领先优势缩小的可能性。这种现象可以解释为领先的球队失去了继续得分的动力，同时落后的球队则必须至少让分数看上去不那么“丢脸”。¹⁴

我们似乎会认为篮球比赛的结果肯定不是随机的。聪明、健壮且灵活的篮球运动员，拥有很多巧妙的进攻手段，并能在关键时刻实现扭转乾坤的得分。这当然也是事实，但是球员们的努力效果可能会被抵消。额外的进攻得分可能会因为额外的防守努力而被抵消。一个重要的抢断后的快速上篮，可能会被冲刺了大半个球场的对方球员破坏。这个模型还提出了一个策略：更强的那支球队应该加快比赛节奏，以创造更多的进攻回合。占有优势的球队应该更频繁地玩“轮盘赌”，因为随机“漂移”对他们有利。

简单随机游走模型只在一个维度上进行。我们还可以对高维随机游走建模。二维随机游走从平面中的原点（0，0）开始，然后在每个周期中随机走向东、南、西、北。二维随机游走类似于在一张纸上绘制出来的一条弯弯曲曲的线，同时也满足递归性（recurrence）和无界性，有点儿类似于在你的起居室中随机搜索一只丢失的耳环时的路线。这种递归性使随机觅食成了蚂蚁的一个觅食策略。¹⁵如果二维随机游走不是递归性的，那么蚂蚁就需要更复杂的内部地图或更强的信息踪迹才能找到它们的巢穴。

但是在有三个维度的情况下，随机游走将不再满足递归性。在一个房间里到处飞的苍蝇和在空气中弹跳的分子都只会有限次地返回到它们

的起点。¹⁶（正因为如此，才会在本章开头引用角谷静夫的那段话。）

随机游走的无递归性为模型如何阐明我们的思考提供了一个很好的例子。直觉告诉我们，当添加维度时，返回起点的次数应该会减少，而逻辑则表明，这里会出现一个突然的变化。在一维和二维的情况下，随机游走会无限次地返回起点。而在三维的情况下，它将“永恒在外游荡”。要得到这种结果必须利用数学，只靠直觉是不够的。

使用随机游走估计网络规模

我们可以利用低维随机游走的递归性来估计某个网络的规模。方法很简单，随机选择一个节点，然后沿着网络的边开始随机游走，并跟踪它回到初始节点的频率。返回到初始节点所需的平均时间与网络的规模相关。例如，为了估计一个社交网络的大小，可以要求某人指定一个朋友，然后让那个朋友再说出一个朋友的名字，一直继续这个过程，看需要多久才会返回到同一个人。

图13-2显示了两个网络。左边的网络有3个节点，它们组成了一个三角形。右边的网络有6个节点，组成了两个三角形。在左边的网络上，我们不妨从A开始随机游走。假设它先移动到B，然后再移动到C，最后再返回到A。这也就是说，随机游走只需3步就可以返回它的起点。而在右边的网络上，从D开始的随机游走可能需要7步（F—G—H—F—E—F—D）才能回到起点。如果将这样的实验重复多次，那么左边网络的平均返回时间显然会比右边网络要短。虽然对这些小型网络来说，要衡量它们的规模并不一定需要这种方法，但对于大型网络（如万维网或大型电子邮件网络）来说确实非常有用。

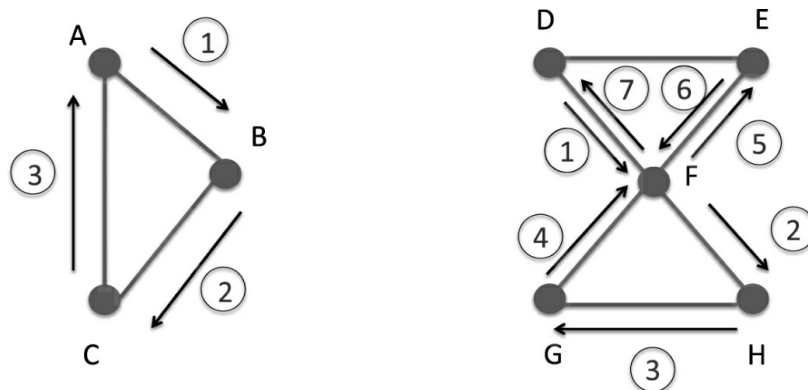


图13-2 网络上的随机游走

随机游走与有效市场

事实已经证明，股票价格接近正态随机游走，带有正漂移，以获得市场收益。许多个股的价格也接近随机。图13-3显示了Facebook在2012年5月18日首次公开发行后一年中的每日股票价格数据。Facebook公开发行时的价格为每股42美元。截至2012年6月1日，股票价格已经下跌到了28.89美元。一年后，价格进一步下降至24.63美元。图13-3还显示了另一个已经校准为具有类似变差的随机游走。

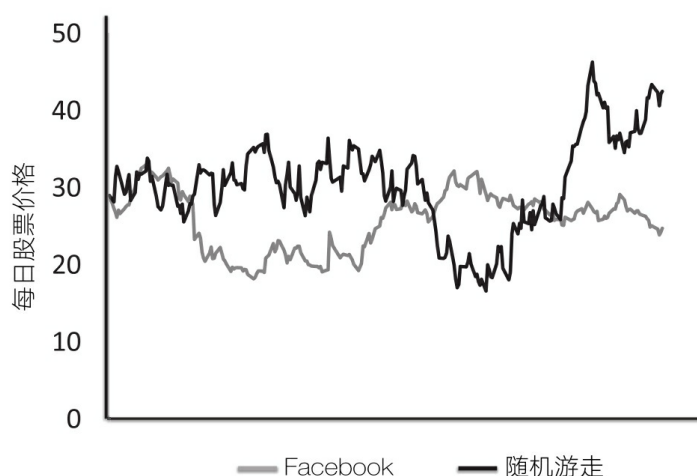


图13-3 2012年6月—2013年6月，Facebook每日股票价格vs. 一个随机游走

我们可以对Facebook的股价序列进行统计检验，以确定它是不是真

的满足正态随机游走的假设。首先，价格应该以相同的概率上下波动，在这个序列所涵盖的249个交易日内，Facebook的股票价格在127天内是下跌的，占总交易日数的51%。其次，在随机游走中，增加的概率应该与前一周期的增加无关，Facebook的股票价格连续两天在同一方向上发生变化的时间只占总时间的54%。最后，持续出现在同一方向上的最长波动应该是8天，在这一年时间里，Facebook的股票价格曾连续10天上涨。因此，总的来说，我们不能否认Facebook的股票价格与正态随机游走一致的假设。

同样的分析也适用于所有股票的日交易价格。为了做到这一点，我们必须先去除股票价格中所包含的平均上涨趋势。研究表明，从20世纪50年代到80年代，每日股票价格略有正相关关系。在进行了去趋势处理之后，一天上涨之后再出现上涨的概率略超过50%。20世纪80年代之后，由于投资者开始变得更加精明，一天上涨之后再出现上涨的概率下降到了50%，从而与随机游走完全一致。

股票价格可能遵循随机游走的原因是，聪明的投资者能够识别出并消除这种模式。例如，在20世纪90年代，分析师注意到，股票价格往往会在每年年初出现上涨，这种现象被称为“一月效应”（January effect）。聪明的投资者可以在12月以低价购买股票，并在来年1月卖出以获取利润。这个策略看起来好得让人难以置信，而事实是，如果投资者在12月购买股票，他们就会抬高价格，从而抵消“一月效应”。事实上，我们不应该对“一月效应”的消失感到惊奇。

经济学家将市场价格的可识别持久模式类比为人行道上的百元钞票。如果有人看到人行道上有张一百元的钞票，就会把它捡起来，然而只要这样做了，钞票就会消失。同样的逻辑适用于股票价格模式：如果它们存在，它们就会消失。因此，充满了聪明的投资者的市场几乎必定不会包含什么可预测的价格模式。既然价格不会呈现出任何模式，那就只能是随机游走了（需要注意的是，必须先去除一般的上行趋势）。

经济学家保罗·萨缪尔森（Paul Samuelson）构建了一个能够生成随机游走的模型。他的模型并不要求投资者知晓未来所有期间的股票价值，而只要求他们知晓股票价值的分布。正如萨缪尔森所说：“人们不能过于迷信现有的定理，它不能说明实际的竞争性市场运作良好。”¹⁷但不是每个经济学家都能认同他的观点。

一些经济学家将这种随机游走思想进行了扩展，提出了有效市场假说（efficient market hypothesis）。这个假说指出，在任何时候，股票的价格都反映了所有的相关信息，未来的价格必定遵循随机游走。有效市场假说依赖于一个自相矛盾的逻辑。¹⁸因为要确定准确的价格需要付出时间和精力，财务分析师必须收集数据并构建模型。如果价格真的是随机游走的，所有这类活动都将无法得到预期的回报。然而，如果真的没有任何人花费时间和精力去估计价格，那么价格就会变得不准确，也就意味着人行道上会铺满百元钞票。

简而言之，正如格罗斯曼和斯蒂格利茨悖论（Grossman and Stiglitz paradox）所强调的，如果投资者相信有效市场假说，他们就会停止分析，从而导致市场效率低下；而如果投资者认为市场效率低下，他们就会应用模型进行分析，从而提高市场效率。

事实上，股票市场上的价格变动与随机游走确实相当接近，尽管利用复杂的统计技术确实能够揭示某些短期模式。¹⁹这也就是说，虽然人行道上可能没有铺满百元钞票，但是在草地上确实能够找到一些四叶草，只要足够努力。

有些批评有效市场假说的人还指出，许多投资者持续战胜市场的时间明显不能用偶然性来解释。²⁰此外，股票价格之所以随机波动，也可能是由于一些其他原因，例如复杂的交易规则的总体影响。日常价格的波动性超过了流入市场的信息量，而且，在现实世界中似乎并没有发生什么重要的事情时，股票市场也会出现大幅飙升或跳水，这就表明市场上存在泡沫。给某个人带来很大不利的某个事件，对另一个人来说也许

不过是“尽管有这些问题，但是.....”。是的，波动性确实很高，但是很少的信息就可能会产生很大的影响。即便市场真的出现了大幅飙升或跳水，市场仍然可能是遵循长尾随机游走的。在长尾随机游走中，股票的日常波动源于长尾分布。

虽然，股票价格始终准确的说法似乎令人难以置信，但从长远来看，价格确实不会与真实价值相差太远。我们可以应用72法则来证明这一点。如果经济每年增长3%，那么在半个世纪中，经济总量将增长4倍。如果回到1967年，当时美国的国内生产总值相当于今天的4.2万亿美元（按2009年美元计算），而到了2017年，美国的国内生产总值增长到了将近17万亿美元（按2009年美元计算），增长了4倍，这正是我们所预期的：每年增长3%，半个世纪就可以增长4倍。在同一时期，标准普尔500指数股票的实际价值也增加了大约4倍。如果股票市场每年上涨12%（以实际美元价值计算），那么股票价格就会增加256倍，这应该是不可能的。²¹

从长远来看，有效市场假说或类似的假说是合理的。但是从短期来看，押注价格修正却可能存在不小的风险。在这方面，长期资本管理公司（LTCM）的成败经历很有启发性。这是一家对冲基金，其董事会成员中包括了两位诺贝尔经济学奖得主。在1996年和1997年，长期资本管理公司公布的年回报率都超过了40%，原因是它发现了市场上的效率低下问题，并预测市场会做出修正。1998年，他们（正确地）注意到了俄罗斯债券价格与美国国债价格之间的不一致性，于是他们下了一个很大的赌注。然而，俄罗斯的违约（自1917年以来的首次违约）在短期内进一步增大了这种不一致性。长期资本管理公司一下子亏损了46亿美元，几乎导致整个金融市场崩溃。在长期资本管理公司得到救助后不久，债券价格确实恢复了一致性，但是这个修正来得太慢了。长期资本管理公司给我们的教训是深刻的，也是显而易见的，那就是不要过分相信一个模型。

小结

在本章中，我们讨论了伯努利瓮模型和随机游走模型，然后将这些模型应用到了很多领域。我们看到，这些模型能够从连续得分现象中解析出随机性，还可以用来制订博彩策略、评估股票价格变化的时间序列，以及理解篮球比赛的结果。我们还懂得了如何应用随机游走返回时间的幂律分布来增进对企业生命周期和生物分类单元的理解。

从这些应用中，我们看到随机游走模型为评估时间序列提供了一个很有用的框架。我们不能被一两年的成功所愚弄，因为那并不意味着持续的卓越。在《从优秀到卓越》（*Good to Great*）这本有史以来最畅销的商业书籍之一中，吉姆·柯林斯（Jim Collins）描述了那些能够持续取得成功的公司的特点，例如拥有谦逊的领导者、选择合适的人进入团队、保持严格的纪律。柯林斯以6次铁人三项世界冠军戴夫·斯科特（Dave Scott）的习惯为例，将之称为“冲洗你的奶酪”。戴维·斯科特会清洗奶酪以减少身体的脂肪含量。柯林斯在这本书中特意列出了11家坚持了他所说的那些原则的“伟大公司”。但是，在他的书出版后的10年中，只有一家公司实现了强劲增长。另外10家公司中，一家被其他企业收购了，一家由政府接管了，另外8家则只带来了零回报。

伟大的企业确实会拥有一些共同特征，但这个事实并不意味着这些特征就必定有助于成功。也许，很多表现糟糕的公司也拥有这些特征。挑选一些看上去很好的公司出来，列出它们的特征，这并不是模型思维。模型思维的要求是，推导出能够导致成功的那些特征，例如才华横溢的工人，然后再根据数据来检验相关结论。如果可能的话，最好寻找一些自然实验，也就是相关特征随机变化的实例。

其他模型，例如我们在第28章中将会介绍的舞动景观模型和崎岖景观模型，更是对吉姆·柯林斯的全部理论提出了根本性的质疑。如果经济是复杂的，那么今天证明成功的特征在未来并不一定同样有效。

按“大石头优先”原则，当前的“伟大”在10年后甚至有可能连“不错”都算不上。在得出一般性的结论之前，必须应用多个模型，以避免“犯大错”的风险。也应该注意避免被某些“模式”所惑，看上去似乎是一个趋势，其实可能是随机。

14 路径依赖模型

人不能两次踏进同一条河流，因为无论是这个人，还是这条河，都已经不同了。

赫拉克利特（Heraclitus）

本章将讨论路径依赖模型。在任何领域，人们的行为都建立在他人行为的基础上，无论是国际事务、艺术、音乐、体育、商业、宗教、技术还是政治，我们都应该会看到某种程度上的路径依赖。大学生选修的课程，导致他未来会倾向于选择某些职业而不会选择其他职业；对某个候选人的认可，可能会使他开启政治生涯；友谊可能会导致其他社会关系；我们穿的衣服、读的书、看的电影，以及任何需要我们付出时间的活动，都会表现出一定程度的路径依赖性。

路径依赖也存在于更大的尺度上。普通法裁决会确立并强化先例，从而影响未来的裁决。¹早期的制度形式会影响后来的制度选择。美国政府制定的由私营公司提供健康保险的政策，导致了一个宏大的私营医疗保险行业以及一系列保健机构，还导致了公立医院和私立医院并存的结果。²特定的制度还会引发特定的行为模式，如自私或合作的倾向；而这种倾向反过来又会影响未来制度的效力。³

在本章中，我们构建了一系列动态瓮模型（dynamic urn model），该模型将生成表现出路径依赖性的结果序列。这种模型扩展了伯努利瓮模型，允许瓮内球的分布随过去的结果而变化。利用这些模型构建我们的思维，给出路径依赖的正式定义，并将路径依赖的结果与均衡区分开来。这些正式的定义还可以将路径依赖与临界点区分开来，而临界点是结果中更加突然的变化。

本章由四个部分组成。前两个部分分别讨论了波利亚过程（Polya

process) 和均衡过程 (balancing process)。波利亚过程假设正反馈并产生路径依赖的结果和均衡。关于路径依赖的许多典型例子，包括“QWERTY”打字机的发展，都是基于正反馈的，这种情况也被称为收益递增。均衡过程则假设负反馈并产生路径依赖的结果，但不会产生路径依赖的均衡。第三部分定义了一个基于熵的路径依赖度量，最后一个部分则讨论了模型的进一步应用。

波利亚过程

波利亚过程利用伯努利瓮模型的扩展来刻画正反馈效应。在波利亚过程中，我们会往瓮中加入与抽取出来的球相匹配的球，这个过程会产生结果路径依赖 (outcome path dependence)。结果路径依赖是指每一周期的结果都取决于先前的结果。说结果的长期分布，也就是均衡路径依赖 (equilibrium path dependence) 也取决于结果，也是对的。⁴这两种类型的路径依赖之间的区别，是本章要讨论的一个核心内容。一个均衡路径依赖的过程，必定是结果路径依赖的。如果长期的结果取决于路径，那么这个路径“沿途”的结果也必定如此。不过，一个过程是结果路径依赖的，并不意味着它一定是均衡路径依赖的。现在发生的事情可能取决于过去，但是长期均衡却可能在一开始就确定下来了。

波利亚过程

一只瓮里面装着一个白球和一个灰球。每一周期，都随机抽取出一个球并将这个球与和它颜色相同的另一个球一起放回到瓮中。抽取出来的球的颜色表示结果。

波利亚过程可以用来刻画多种多样的社会和经济现象。一个人选择学习打网球，还是打壁球，可能取决于其他人的选择。如果更多的朋友选择学习打网球，那么这个人就更有可能也选择学习打网球，因为这会

增加他找到伙伴打比赛的机会。与此类似，一个人决定购买什么类型的软件、学习哪种语言或购买哪款智能手机，也可能取决于他的朋友以前做出的选择。类似的逻辑同样适用于企业对技术标准的选择，它们可能会根据其他企业的选择来做出选择。

这个模型通过改变球的分布来刻画这些社会影响。如果灰球代表选择学习打网球的人，白球代表选择学习打壁球的人，那么，当更多的人选择学习打网球时，瓮里面就包含了更多的灰球，从而导致后来的人更加有可能选择学习打网球。对更多人所选择的结果的这种不断增长的牵引力创造了路径依赖。

我们可以从波利亚过程中推导出两个令人意想不到的性质。首先，具有相同数量的白色结果的任何序列都会以相同的概率发生。其次，白球和灰球的每个分布都以相同的概率发生。第二个性质意味着极端的路径依赖。任何事情都可能发生，一切皆有可能。在1 000个周期之后，瓮中包含了40%的白球的概率等于包含了2%的白球的概率。

要理解为什么会这样，不妨考虑前三个周期内所有可能的结果序列。第一个周期的结果是灰色的概率为 $1/2$ 。如果这个结果发生了，那就要加入一个灰球，从而使第二个周期的结果为灰色的概率增加到 $2/3$ 。如果第二个周期的结果也是灰色，那就要加入第三个灰球，从而使第三个周期的结果为灰色的概率增加到 $3/4$ 。由此可以得出，三个灰球（或三个白球）的总概率等于 $1/2$ 乘以 $2/3$ 乘以 $3/4$ ，其乘积等于 $1/4$ 。

前三个结果由两个白球和一个灰球组成的结果序列如图14-1所示。在最上面一行中，结果的顺序是先灰色、再白色、然后又是白色。这个序列的概率是 $1/12$ ，其他两个序列的概率也是一样。因此，获得这三个序列之一的概率等于 $1/4$ 。根据对称性，选择两个灰球和一个白球的概率也必定等于 $1/4$ 。因此，每个可能的结果集——三个白色、三个灰色、两个白色以及一个灰色、两个灰色、一个白色，出现的概率都是 $1/4$ 。而且，产生两个白色和一个灰色的每一个序列都以相同的概率出

现。对任意周期数，都可以得到类似的结果。⁵

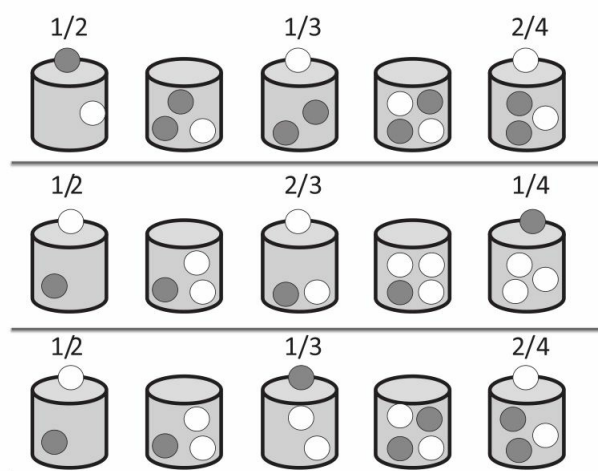


图14-1 由两个白球和一个灰球组成的各种结果

如果我们将波利亚过程加以扩展，也就是进一步加入其他颜色的球，这两个性质仍然成立。各种颜色的任何比例都会出现而且概率相同。这些结果为消费品生产商带来了一个难题：消费者对产品某些性质的长期偏好可能是随机的。无法预测结果的知识仍然可以用来指导行动。例如，福特公司不会希望出现这样的情况：生产出了4万辆黄色皮卡车，后来却发现在一个路径依赖的过程中，红色成了消费者最喜欢的颜色。不受欢迎的颜色的产品会大量积压在库存中，这种可能性指向两种可能的行动。生产企业可以考虑建立这样的供应链：把颜色的决定放在最后一个环节。例如，服装公司可以先不给毛衣染色，直到看清楚流行的颜色是什么为止。或者，生产企业也可以不给消费者选择的机会。亨利·福特只提供一种颜色的T型车，即黑色T型车。苹果公司在推出第一款iPhone时也做了同样的事情：你可以买一部黑色的iPhone，或者以同样的价格买另一部黑色的iPhone。

均衡过程

第二个模型是均衡过程，它的假设与波利亚过程恰恰相反。在抽取

出某种颜色的球后，要加入一个相反颜色的球。如果在前两个周期都抽取出了白球，那么瓮中将包含三个灰球和一个白球，从而导致下一周期抽取灰球的概率增大为 $3/4$ 。这个过程也会产生路径依赖的结果，因为任何一个周期结果的可能性取决于过去的结果的历史。但是，它不会产生依赖于路径的均衡。从长远来看，瓮收敛为每种颜色的球的比例都相同。⁶

均衡过程

一个瓮包含一个白球和一个灰球。每一周期都随机抽取出一个球，并将与抽取出来的球颜色相反的球与抽取出来的那个球一起放回瓮中。球的颜色表示结果。

均衡过程可以用来刻画有趋向平等分配压力的决策或行动序列。有两个孩子的父母可能会尝试给每个孩子分配相同的时间。与一个孩子共度了一个下午之后，父母会产生与另一个孩子共度更多时光的愿望。

平衡过程甚至可以用来对努力实现公平的组织行为建模。国际奥委会希望世界各地都能举办奥运会。2013年，国际奥委会宣布东京成为2020年夏季奥运会和残奥会的主办城市，与东京竞争的两个欧洲城市伊斯坦布尔和马德里败北。四年后，国际奥委会决定把2024年和2028年奥运会主办权分别授予欧洲城市巴黎和北美城市洛杉矶。东京之所以能够赢得2020年奥运会的主办权，部分原因当然是因为它很有竞争力，但是还有一部分原因则是因为自1964年以来夏季奥运会从未在日本举行过，在地域上保证公平这种考量似乎发挥了作用。欧洲、亚洲和大洋洲以及美洲在第二次世界大战后举办奥运会的次数大体上相同：欧洲8次、美洲6次、亚洲和大洋洲7次。

路径依赖还是临界点

路径依赖是对结果的逐渐影响，临界点则意味着结果的突然变化。微软公司的发展为路径依赖提供了一个很好的例子。微软公司于1975年成立后，为BASIC计算机语言开发了解释器。1979年，微软与国际商用机器公司（IBM）达成协议，为IBM的个人电脑提供操作系统。这笔交易使微软走上了将一个只拥有40名员工的公司转变为世界上最有价值公司之一的道路。

IBM的合同促进了微软的发展，但是并不能保证微软长期的成功。当时，个人电脑的市场还很小，互联网并不存在，也没有复杂的文字处理程序、商业软件或视频游戏。此外，个人计算机的成功还部分取决于微软开发的DOS操作系统。在个人计算机市场发展起来后，其他公司也开发出了很多与DOS兼容的软件，从而提供了更多的正反馈。DOS的成功、个人计算机市场的发展以及与DOS兼容的软件开发，都可以被视为从瓮中抽取出来的始终是同一种颜色的球。每个结果都使下一个结果更有可能。计算机时代的到来也许是不可避免的，但是微软的核心作用和个人计算机的发展则只代表了许多种可能的发展路径之一。

我们可以将微软公司成长的路径依赖与1914年6月28日弗朗茨·斐迪南大公（Archduke Franz Ferdinand）遇刺事件对比一下，许多人都认为，这个暗杀事件是导致第一次世界大战爆发的临界点。在暗杀发生6年前，奥匈帝国吞并了波斯尼亚和黑塞哥维那。很多塞尔维亚人都对这种情况心怀不满，于是加夫里洛·普林斯（Gavrilo Princip）枪杀了弗朗茨·斐迪南大公夫妇。奥匈帝国将这个事件归咎于塞尔维亚政府，这是一个几乎不可避免的反应。奥匈帝国准备向塞尔维亚宣战时，决定向德国靠拢，寻求德国皇帝威廉二世的支持。紧张局势迅速升级。塞尔维亚随后与俄罗斯结盟，而俄罗斯又与法国和英国结盟。到了1914年8月2日，德国就向法国宣战了。8月3日，比利时拒绝让德国借道进入法国，世界大战全面爆发。这个过于简化的事件版本表明，考虑到世界各国的结盟状况，斐迪南大公的遇刺是世界开始向战争倾斜的临界点。

我们可以利用可能结果的概率变化来度量路径依赖和临界点。⁷对于波利亚过程，初始概率在瓮中的所有分布上都是均匀的，这是一个最大熵分布。随着事件的展开，分布逐渐变窄，标志着路径依赖的形成：当结果出现后，可能发生的事情也会变化。这种熵的减少是渐进的。而对于临界点，概率分布是突然改变的，熵可能会迅速下降。图14-2显示了这两个过程的差异，每个过程都有两个可能的结果。事件发生后，也就是微软获得了IBM的合同或斐迪南大公遇刺后，每次的概率就变了，后续事件也会改变概率。具有临界点的过程出现大幅度的转折，而路径依赖的过程则变化缓慢。

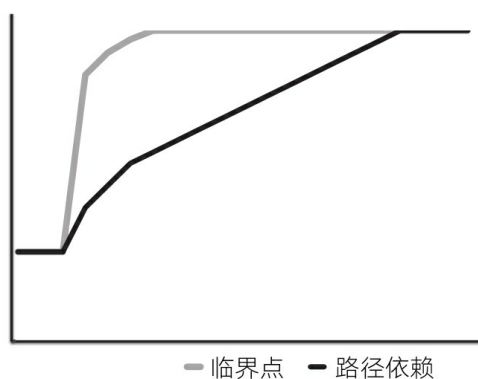


图14-2 临界点与路径依赖

路径依赖模型的应用

在现实世界中，路径依赖可能不会像波利亚过程那样极端。然而，我们还是可以从模型中推断出当行为具有很大的社会性成分时，几乎任何事情都有可能发生。在一个大学里，大多数学生在冬季都可能穿黑色大衣，而在另一个大学里，学生们则可能穿孔雀蓝大衣。模型思维告诉我们，行为差异既可能是社会影响的结果，也可能是不同的内在偏好所致。在任何时候，只要人们在一组固定的备选项中进行选择，而且他们的选择依赖于其他人先前做出的选择时，就会出现这种情况。这样的例子包括在选举中投谁的票、要看哪一部电影，以及购买哪一种技术等。

模型还可以扩展到允许社会影响因所选方案的不同而变化。香草冰激凌可能会有一定程度的反馈。而具有异国情调的绿茶冰激凌则可能得到更具多样性的反馈：一个朋友可能不喜欢它并且劝你不要尝试，另一个朋友可能很喜欢它并鼓励你一定要试一试。这可以证明，反馈的变化越少，选择这种结果的可能性就越大。⁸模型也可以改变，使人们对社会影响的敏感性不同，也就是说，人们对加入瓮中的球赋予不同的权重。

在模型的任何变体中，我们都可以测量（或估计）路径依赖的程度并将之与其他模型的变体进行比较。如果我们在构建一个关于新产品的引入模型时所做的假设表明结果取决于路径的早期部分，那么尽早进入、进行干预或补贴可能是一个不错的策略。这个模型为企业尽早将产品推向市场或提供大幅度的折扣以便尽可能多地吸引早期用户提供了逻辑支持。当然，其他假设则可能表明，拥有更好的产品可能比更早进入市场还重要，也就是说，更好的策略是关注质量。利用各种模型，我们可以识别出与特定情况相关的不同特征，比如个人偏好与社会影响的相对重要性及反馈的变化以及质量的相对差异，并利用这方面的知识去制订策略和指导数据收集。

最后，波利亚过程表明，正反馈是怎样生成结果路径依赖和均衡路径依赖。路径依赖现象会出现在各种各样的情况下。只要一个行动会与未来行动“相遇”或与未来行动相互作用，就会出现某种程度的路径依赖（如果不是均衡路径依赖，也会是结果路径依赖）。在做出关于大型公共项目的决策时也是如此。⁹建造公园或高速公路的决定会对未来的规划决策构成约束。路径依赖程度通常取决于项目的规模。例如，纽约中央公园对纽约市的发展就产生了深远的影响。虽然波利亚过程揭示了交互导致路径依赖的这个核心思想，但是要利用这个洞见来指导行动，还需要更加真实的模型。

风险价值与波动性

我们可以将时间序列数据中的标准差解释为波动性。对股票、房地产和私人股权的投资都会表现出波动性。风险价值（value at risk, VaR）衡量的是在某个特定时间段内损失特定金额的概率。一项风险价值为一年5%的总额为1万美元的投资，在一年结束时损失超过1万美元的可能性为5%。¹⁰银行使用风险价值法来确定为了避免破产手头必须持有的资产数量。例如，为了保证一项风险价值为两周40%的1万美元投资的安全，投资者可能被要求必须持有10万美元的现金。

如果投资遵循简单随机游走，且每一周期投资规模的增大量或减少量为M，那么它的风险价值为N期2.5%的 $2M\sqrt{N}$ 。¹¹因此，每天随机上涨或下跌1 000美元投资的风险价值为9天2.5%的6 000美元，或者，一年2.5%的38 000美元。考虑到风险价值在步骤大小上是线性增长的，但它会像周期数的平方根一样增长，我们可以使用风险价值来解释为什么美国联邦存款保险公司（Federal Deposit Insurance Corporation）只要求银行保留相当于资产总额的2%隔夜现金，而银行则要求消费者支付房子总价的20%。这种差异是由于贷款期限不同所致。隔夜贷款的期限是一天，住房贷款可能持续数十年。3 065天的平方根约为60。

这里，我们假设了一个正态随机游走，计算风险价值的分析师通常考虑过去的经验收益分布。如果经验分布有一个较长的尾部，也就是说，如果它包含更多的大事件，那么风险价值会随着大事件的发生而增加。

虽然风险价值起源于金融领域，但是它所包含的思想却是可以广泛应用的。一个由非营利组织运营并由志愿者提供服务的施粥处通常需要25名志愿者，该组织可能想了解缺乏足够志愿者的可能性。如果志愿者的数量是一个每个星期增加或减少一个人的简单随机游走，那么就可以利用上面的风险价值公式，并且将参数设定为

M=1和N=52，从而可以确定，风险价值为一年2.5%的15（在一年、2.5%的概率下，风险价值为15），这意味着该非营利组织有2.5%的概率会缺少一名志愿者。

15 局部互动模型

每一代人都嘲笑旧时尚，同时又虔诚地追随新时尚。

亨利·戴维·梭罗（Henry David Thoreau）

在本章中，我们要研究两个局部互动模型（local interaction model），即局部多数模型（local majority model）和生命游戏。这两个模型都是建立在一个由单元格组成的棋盘上，棋盘由处于两种状态之一的元胞组成。表面上看，这两个模型似乎没有太大的不同，其实不然。在局部多数模型中，元胞通过与它的大多数邻居的状态相匹配来更新。而在生命游戏中，元胞的更新规则要更加复杂，它依赖于多个阈值。这两个模型的结果也不同。局部多数模型总是收敛到均衡，生命游戏则取决于其初始条件，可能会产生任何类型的结果：均衡、周期性、复杂性或随机性。

局部多数模型可以用于解释和预测社会系统和物理系统中的实际结果。它可以通过符合个体或者刻画像自旋玻璃这样的磁性粒子与自己的“邻居”对齐的物理系统来表示离散选择。相比之下，生命游戏则完全是探索性的。开发生命游戏的目的就在于探索简单的规则如何集结并产生各种复杂的现象。在生命游戏中，周期性模式、复杂序列和随机性都可能从相互作用中涌现出来。生命游戏突显了整体与部分的不同。这里，我们可以给出一个粗略的类比：在人类的大脑中，诸如情绪、认知和意识等涌现现象也会从更简单的部分涌现出来。

首先分析局部多数模型。我们将阐明，一个标准的协调博弈如何为模型所假设的行为规则提供微观基础。因此，我们既可以将模型中的参与者解释为遵循规则的行为者，也可以解释为应用某个最优反应策略的理性行为者。然后，我们分析了生命游戏，并说明它如何从简单的规则

中产生复杂性。本章末尾的讨论强调了利用多个局部互动模型探索的价值。

局部多数模型

局部多数模型假设元胞是排列在棋盘上的。¹每个元胞处于两种状态中的一种：开或关。初始时，我们随机地给元胞分配状态，此后，元胞的状态取决于它“邻居”的状态。邻居可以通过多种方式加以定义。我们将元胞C的邻居定义为位于它东、南、西、北的4个元胞以及4个对角上的相邻元胞，因此它的领域大小为8。

局部多数模型

二维方格上的每个单元都处于两种状态之一：开或关。每个单元有8个邻居。²在每个周期中，随机选择一个元胞。³当且仅当其中它的5个或更多邻居处于另一个状态时，这个元胞才会改变自己的状态。

1	2	3
4	C	5
6	7	8

局部多数模型中的局部互动包括正反馈：元胞要与其他元胞的状态相匹配。图15-1显示了局部多数模型的典型均衡配置。

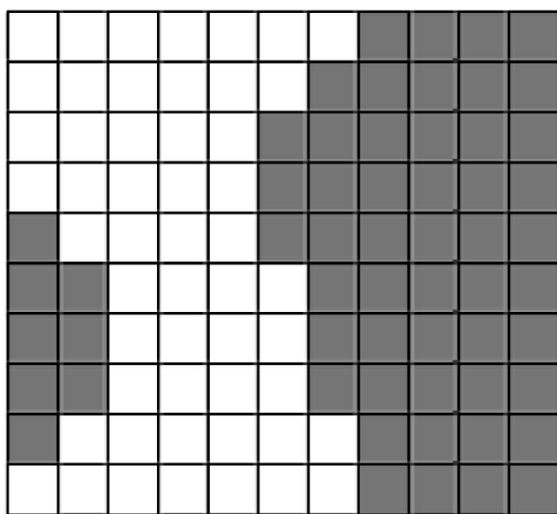


图15-1 局部多数模型中的典型均衡配置

在均衡状态下，每个元胞的状态都与它的大多数邻居的状态相匹配。均衡配置类似于荷尔斯坦奶牛身上的黑白斑块。虽然均衡配置取决于各元胞的初始配置，但是该模型对初始条件没有表现出极高的敏感性。改变一个元胞的状态最多只会导致最终配置出现很小的变化。最终的模式还取决于元胞被激活的顺序。因此，这个模型也表现出了路径依赖。它的均衡数量非常巨大。模型产生的两个均衡之间的相似度，不会比同一块地里的两头荷尔斯坦奶牛身上的黑白斑块更高。

局部多数模型最初是为了用来刻画一类物理系统而开发的。在这类物理系统中，每个元胞的状态代表一个原子的自旋态，我们不妨将每个元胞想象为一个具有负电荷或正电荷的磁体。每个磁体都位于一个局部磁场中，物理力量会驱使这个磁体与它的邻居的自旋相匹配。同样的模型也可以用来表征玻璃体和水晶体。

在这里，我们利用这个模型来刻画人与人之间的局部协调或一致性。假设每一个元胞都代表一个人的行动。这里的行动可以是任何行为习惯，例如握手或鞠躬，打断或举手。任何一个人都想选择一个与邻居相匹配的行动。棋盘代表社交网络，在很多情况下，棋盘确实是代表社交网络的适当模型。例如，当房主决定维护自家的院子、植树造林或创

建生态景观的时候，又或者礼堂内的人决定是否为表演者起立鼓掌时。⁴虽然棋盘最多只能作为一个粗略的近似，但是我们利用它可以得出核心直觉结论。

如果在计算机上对这个模型进行模拟，就会发现它总是能实现斑块状的均衡配置。在第16章中，我们将给出这种均衡出现的原因。在局部多数模型的物理解释中，斑块状均衡模式对应于受挫状态（frustrated state）。许多元胞在开的状态下有一些邻居，在关的状态下也有一些邻居。

如果我们通过社会视角解释模型，就可以把这种受挫状态理解为次优均衡（suboptimal equilibrium）。如果开对应于以握手的方式表示欢迎、关对应于以鞠躬的方式表示欢迎，那么位于“斑块”边界上的人在与他们的邻居互动时可能会碰到一些尴尬的情况：当其他人准备握手时他们却准备鞠躬，或者当其他人准备鞠躬时他们却准备握手。如果每个人都选择了相同的动作，也就是说，如果他们解决了协调博弈问题，那么从整体上看，所有人都会更加开心。

由于这种互动只作用于局部，因此出现了次优均衡，即受挫状态。相反，如果元胞是根据全局多数原则来匹配的，那么很快所有元胞都会处于相同的状态。这种观点意味着，创建共同行为可能需要影响更加广泛的网络。如果人们只在局部与自己的邻居协调，他们就会创造出各种各样的行为。因此矛盾的是，恰恰是协调导致了多样性。

纯粹协调博弈

在纯粹协调博弈（pure coordination games）中，每个博弈参与者都要在两个行动A或B中选择一个。如果两个博弈参与者都选择了相同的行动，那么每个博弈参与者的收益为1。如果两个人各自选择了不同的行动，那么每个人的收益都为零。

行动	A	B
A	1, 1	0, 0
B	0, 0	1, 1

一个纯粹协调博弈有两个有效的均衡：两个博弈参与者选择A或者两个博弈参与者选择B。它还有一个无效的均衡，也就是每个博弈参与者都在A和B之间进行随机选择。我们可以从纯粹协调博弈的角度重新解释局部多数模型：在这个博弈中，每个元胞都是一个博弈参与者，它必须选择一个共同行动来对抗它的8个邻居。如果博弈参与者只有在随机激活时才能改变行动，那么某个博弈参与者就可以通过选择与它的大多数邻居所选择的行动相匹配的行动来增加自己的收益。这种策略被称为短视最优响应（myopic best response），因为它没有考虑到邻居可能的未来行为。

一个博弈参与者有5个邻居选择了B，那么这个博弈参与者可以通过从A切换到B来在短期内增加自己的收益，但是，如果这个博弈参与者和邻居是被其他选择A的博弈参与者的“海洋”所包围的一个孤岛，那么这个博弈参与者保持A不变反而可能获得更高的期望收益。最关键的一点是，局部多数模型中的行为规则虽然只是一个假设的规则，但却可以植根于博弈论模型中。

协调的悖论（paradox of coordination）将不同群体之间的差异解释为一种特异性的分歧。对于某些行为，比如你是将酱油或番茄酱存放在橱柜中还是冷藏在冰箱里，或者当别人走进你家里的时候，是会脱鞋进门还是直接穿着鞋进来，而与其他人协调一致是明智的选择。由此而产生的地区多样性使我们的生活更加丰富多彩。以咖啡为例，意大利的小杯力士烈特（ristretto）、法国的中杯浓缩，以及华沙的大杯咖啡加奶油（kawa ze smietanka）极大地增加了人们在欧洲旅行的乐趣。

但是，其他很多差异却可能是低效的。各种各样的电插头可能会令

人抓狂——这里是两个插脚的，到那里却是三个插脚的。随着全球一体化程度的提高，技术协调失败的代价可能是非常高昂的。例如，瑞典政府决定将靠左行驶改为靠右行驶，以便与欧洲大陆其他地区保持一致。这件事非同小可，转为靠右行驶的那一天在瑞典被称为“H日”（Dagen H）。1967年9月3日凌晨4点45分，瑞典行驶在路上的所有车辆以及许多特意在凌晨时分上路参加这个活动的瑞典人都突然停下来，然后，在接下来的15分钟内，所有的汽车都从左侧行驶到右侧。凌晨5点，所有汽车开始在道路的另一侧再次启动，开始靠右行驶。尽管达成协调的动机相当强烈，但是许多时候人们却无法做到。例如，虽然英国已经通过隧道连接到了欧洲大陆，但是英国的车辆至今仍然靠道路的“错误”一侧行驶，而且英国的某些（尽管不是全部）前殖民地也是如此。

协调的悖论

如果人们是在局部进行协调的，那么从全局的角度来看，整体配置将会是斑块状的、多样性的。

应用这个模型时，必须记住，许多帮助我们实现协调的文化习俗，例如人们如何哀悼死者的逝去、如何庆祝孩子的出生，并不是简单的“奇风异俗”，而是文化的重要组成部分，是一系列内在一致的行为规则、惯例的组合；它们定义了一个人到底是谁，给了人们生活的意义和归属感。⁵

与任何模型一样，对于局部互动模型，我们也可以对各种参数进行模拟实验，看看改变它们会如何影响结果。在局部互动模型中，均衡时形成的斑块大小增加得比邻域的大小更快。如果我们使邻域，也就是影响网格上一个格子的格子数量增加到原来的两倍，那么斑块会变得比原来的两倍还大。因此，这个模型表明，当技术和城市化使人与人之间的联系更加紧密之后，协调的力量可以产生更大的同质性行为 and 信念。

模拟实验还表明，如果我们将整个配置变成一个狭长的矩形，那么模型一般来说会产生水平和垂直的条纹，如图15-2所示。⁶斑马状的条纹是一个均衡，因为每个处于开（关）状态的元胞，都有5个处于开（关）状态的邻居。在正方形上，这种类型的模式也是一个均衡，但是很少会真的出现。尽管这种令人困惑的发现有可能导致我们误入歧途，陷入没有什么实证意义或理论价值的问题，但是它们也有可能引导我们得出一些深刻的洞见，从而带来更深层次的意外发现。

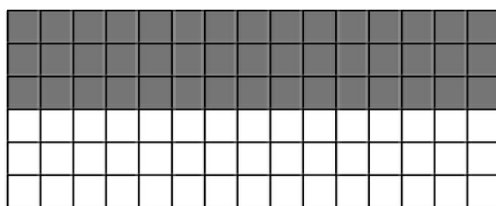


图15-2 局部多数模型中稳定的线条

在这个例子中，正方形网格会产生荷尔斯坦奶牛式斑块状模式，而狭长的矩形网格则会产生斑马状的模式这个结果，几乎就是在直接要求我们回答：是不是可以用这里的模型去解释不同动物皮毛上斑点和条纹的特有模式？大量科学研究文献告诉我们，确实可以。⁷

生命游戏

接下来我们讨论下一个模型：生命游戏。这个模型也假设位于棋盘上的元胞处于两种状态中的某一种。生命游戏与局部多数模型的关键区别在于，在这里，元胞的更新规则有两个阈值，并且所有的元胞都同步更新自己的状态。因此，在生命游戏中，我们可以说初始配置如何如何，时间1的配置如何如何，时间2的配置如何如何，等等。还可以把同步更新视为一个“进行曲”动态机制（更新！更新！更新！）。⁸

生命游戏

方格上的每个元胞都或者是活的（开的）或者是死的（关的）。每个元胞的邻居由网格上的8个相邻元胞组成。元胞根据如下两个规则同步更新自己的状态：

活的规则：对于一个死元胞，当恰好有三个活的邻居时，这个死元胞就会变活。

死的规则：对一个活元胞，当活的邻居小于两个时或当有三个以上的活邻居死去时，这个活元胞就会死去。

我们从排在同一条水平线上的三个活的元胞开始讨论，如图15-3所示。在下一个时期，对每个元胞应用上述死活规则之后，我们会得到排在同一条垂直线上的三个元胞。中间的活元胞有两个活着的邻居，所以它还活着。两端的两个活元胞分别只有一个活的邻居，所以它们都死了。最后，中间的活元胞的上方和下方的元胞都变活了，因为这两个元胞分别各有三个活着的邻居。根据对称性，等再下一个时期更新后，又会回到三个元胞组成水平线的情形。如果继续迭代运用上述规则，模型就会在水平线与垂直线之间不断交替，这也就是说，它将会不断“闪烁”。

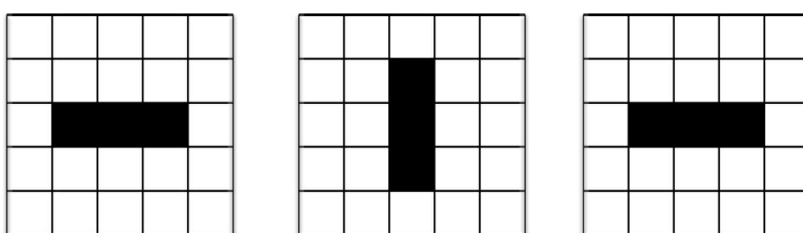


图15-3 生命游戏中的“闪光灯”

闪光灯是由元胞之间的互动产生的，而不能直接从假设中推导出来。复杂系统学者将这种宏观现象称为涌现。闪光灯是生命游戏产生的涌现结构中最常见的、最不能给人留下深刻印象的一种。图15-4显示了其他三种简单涌现结构：方块、滑翔机和右五连（R-pentomino）。这

是一个均衡配置：每个活元胞都有三个活的邻居，每个死元胞最多有两个活的邻居，因此，既没有活元胞死亡，也没有死元胞复活。

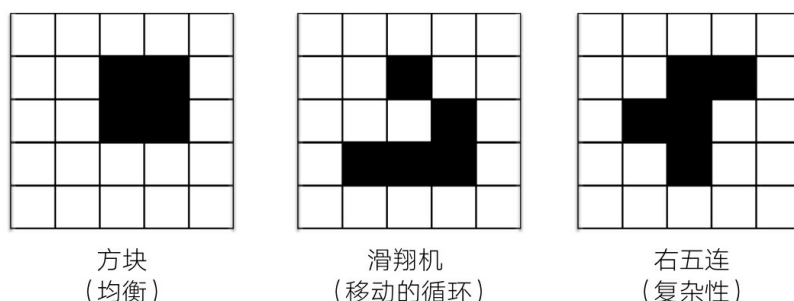


图15-4 生命游戏中的三种涌现结构

图15-4中间的配置会产生一个大小为4的循环，该循环沿着对角线向下和向右滑动，成了滑翔机。与之相关的另一个更精致的配置被称为“滑翔机枪”，它能产生无穷无尽的滑翔机。第三种配置右五连能产生一系列复杂的模式。如果我们在大型网格上重复运行模型超过1 000次，这个右五连会生成许多滑翔机、闪光灯以及其他一些较小的稳定配置。生命游戏也可以产生随机性。⁹因此，生命游戏可以根据初始状态产生任何类别的结果。

这种能力甚至引发了一些哲学问题。生命游戏由排列在网格上的、只有两种状态的元胞组成，并运用简单的规则进行更新。它可以生成精美的图案，通过适当的编码，还可以变成一台通用计算机。我们可以把初始模式视为输入，根据规则产生的结果则可以解释为计算。因此，我们可以在这个模型和人类的大脑之间进行一个粗略的类比，因为人类的大脑也是由依赖于基于阈值规则的、在空间上相互连通的简单组成部分构成的，尽管那些规则要复杂得多。

当然，这并不是说我们在生命游戏中观察到模式就可以解释意识了。现在还没有任何一本书名为《生命游戏：意识的解释》，尽管丹尼尔·丹尼特（Daniel C. Dennett）^⑥确实写过一本以《意识的解释》（*Consciousness Explained*）为名的书，而且他也认为像生命游戏这样的

简单模型可以提供关于意识如何演化的洞见。丹尼特这个思想得到了不少人的共鸣，物理学家斯蒂芬·霍金（Stephen Hawking）就这样写道：“我们完全可以想象，像生命游戏这样的东西，只有少数几个基本定律，就可以产生高度复杂的特征，甚至可能产生智能。”¹⁰

小结

在本章中，我们研究了两个关于排列在网格上的元胞之间的互动模型。第一个模型是局部多数模型，它总是可以达到某个均衡（尽管可能的均衡有许多种），我们可以将这个模型与多种物理过程和社会过程进行类比。第二个模型是生命游戏，它能够产生从均衡到随机的任何一种类型的结果。这个模型与现实世界之间不存在任何明确的联系。生命游戏给出了一个很好的例子，它说明，构建替代现实的模型是怎样帮助我们产生洞察力的，也就是从微观的规则中涌现出动态的宏观层面结构，这可以极大地加深我们对世界的理解。正如生命游戏所呈现的那样，整体可以执行远远超出其各个组成部分的功能。例如，如果我们把两个 3×3 的方格的角连接起来，画出一个斜8字图形，那么生命游戏就会产生一个长度为8的循环模式。它循环通过一系列模式，然后在恰好8个步骤内回到那个8字图形上。这个模式，从图形上看像8字，它的行为也“好像”知道它要数到8。这确实是非常惊人的。

为什么生命游戏会产生复杂性，而局部多数模型则不可避免地走向均衡？为了理解个中原因，我们还需要更多的分析工具和框架。在第16章中，我们介绍了李雅普诺夫函数（Lyapunov function），它运用差分方程对世界状态进行分类。细致地构建出一个李雅普诺夫函数之后，我们就可以解释为什么局部多数模型必定会达到平衡，而生命游戏则不一定。

最后要强调的是，正是在我们对模型进行探索的过程中，才使模型（以及现实世界）是否会产生均衡、周期性、复杂性或随机性这个问题

突显了出来。确实，模型既能够回答问题，也能够提出问题。它们关上了一些门，同时又打开了更多的门。

16 李雅普诺夫函数与均衡

数学之美，只会呈现给那些更有耐心的追求者。

玛丽安·米尔札哈尼（Maryam Mirzakhani）

在本章中，我们学习李雅普诺夫函数，它给出了模型能够实现均衡的条件。李雅普诺夫函数是在按时间索引的配置系统上定义的实值函数。在每个时间步骤中，李雅普诺夫函数都要给配置分配一个值。如果配置发生了变化，也就是说，如果模型不处于均衡状态，那么李雅普诺夫函数的值就会减小一个固定量。李雅普诺夫函数也具有最小值，这意味着到了某个时间点上，它的值最终必定会停止递减。当发生这种情况时，模型就达到均衡了。例如，我们可以使用李雅普诺夫函数来证明，为什么局部多数模型一定会收敛。

本章的核心结论是，如果能够为模型构建一个李雅普诺夫函数，那么模型必定会达到均衡。在这种情况下，我们不可能得出周期性、随机性或复杂性。更重要的是我们甚至可以确定模型收敛到均衡所需的时间，这一点在为局部多数模型构造李雅普诺夫函数中可以看得很清楚。

本章的内容分为6个部分。我们首先定义了李雅普诺夫函数，然后将李雅普诺夫函数应用于对逐底竞争博弈（Race to the Bottom Game）的分析。然后，我们分别为局部多数模型和自组织活动模型构建了李雅普诺夫函数。接着，我们证明了为什么可以为某些交易市场构建李雅普诺夫函数，却不能为另外一些交易市场构建李雅普诺夫函数，同时也分析了为什么生命游戏没有对应的李雅普诺夫函数。接着，我们讨论了一个看似令人烦恼的数学问题：在不能找到李雅普诺夫函数的情况下达到均衡。本章最后讨论了均衡是否可取的问题。

李雅普诺夫函数

一个离散动态系统（discrete dynamical system）由可能的配置空间，以及一个转移规则（transition rule）组成。我们可以把配置空间视为世界的多维状态，例如生命游戏中的活元胞和死元胞的初始集合，而转移规则则是将时间 t 时的配置映射到时间 $t+1$ 时的配置上。

一个李雅普诺夫函数是从配置到实数的一个映射，它满足两个假设：第一，如果转移函数不处于均衡状态，则李雅普诺夫函数的值就会减少某个固定的数量（对此，稍后会给出更多的解释）；第二，李雅普诺夫函数具有最小值。如果这两个假设成立，那么该动态系统必定达到均衡。

李雅普诺夫函数

给定一个离散时间动态系统，它的转移规则由 $x_{t+1}=G(x_t)$ 组成。对于实值函数 $F(x_t)$ ，如果对于所有的 x_t ，都有 $F(x_t) \geq M$ ，而且存在一个 $A>0$ ，那么下式成立，这个实值函数 $F(x_t)$ 是李雅普诺夫函数：

$$F(x_{t+1}) \leq F(x_t) - A, \text{ 如果 } G(x_t) \neq x_t$$

如果对于 G ， F 是一个李雅普诺夫函数，那么从任何 x_0 开始，必定存在一个 t^* ，使 $G(x_{t^*}) = x_{t^*}$ ，即该系统会在有限时间内达到均衡。

首先为逐底竞争博弈构建一个李雅普诺夫函数。这个函数刻画了该博弈的博弈参与者的策略环境，以便每一个博弈者都更愿意提供恰好低于平均水平的水平。

逐底竞争博弈

有 N 个博弈参与者，每个博弈参与者在每个时期都要提出一个支持水平，其取值范围为 $\{0, 1, \dots, 100\}$ 。提出了最接近平均支持水平 $2/3$ 的博弈参与者可以获得那个期间的奖励。

这个博弈可以用来解释美国各州政府削减社会项目支出的行为，例如减少对贫困人口的援助。没有任何一个州的州长或州立法机关希望公众认为自己冷酷无情，但是，又没有人愿意提供慷慨的援助贫困人口的计划，因为那会把邻州的穷人吸引过来。于是每个州都会提供一些资金，但只愿意提供低于平均水平的资金。在通过环境保护法规或制定税率时，相互竞争的国家也会有类似的动机。各国宁愿采取只对环境施加较少限制的政策，并且为了吸引企业投资，税率低于平均水平。

逐底竞争博弈是否能达到均衡，取决于博弈参与者的行为规则。如果博弈参与者随机选择支持水平，那么结果就将会是随机的。然而，考虑到这个博弈的收益结构，随机选择支持水平并没有意义。在这里，我们假设如下行为规则，这种规则与实验研究的结果一致。¹在第一个时期，我们假设每个博弈参与者会随机选择一个低于50的支持水平。在以后各期中，每个博弈参与者选择的水平至少等于1且低于上一期平均水平的 $2/3$ 。如果上一期的平均水平已经小于零，那么每个博弈参与者都会选择零。

很容易证明，来自任何博弈参与者的最大支持水平满足李雅普诺夫函数的条件，即最大支持水平的最小值为零。而且，在每个时期内，如果支持水平采用整数值，那么最大支持水平至少会下降1。因此，到了某个时间点上，每个博弈参与者都会提议零支持水平。也就是说，博弈参与者在逐底竞争中都碰到了“底”。在这个例子中，模型产生了一个不好的结果。为了防止出现这种逐底竞争，必须改变博弈结构。例如，为了增加对贫困人口的扶持，法律可以规定由联邦政府来提供资金，或者对各州规定一个最低“扶贫”支出水平。²

另外，假设我们允许博弈参与者在0至100的区间内选择任何实数值，而不是整数值。再假设在每一轮中，博弈参与者选择的支持水平等于上一轮平均水平的2/3，那么平均支持水平将会随着时间的推移而逐渐降低，但是永远不会达到零的均衡。正如色诺悖论（Xeno's paradox）一样，这个过程会越来越接近于零，但是却永远不会达到零。因此，为了确保均衡的实现，还必须假设一个最小减量（A）。

局部多数模型

现在，用李普雅诺夫函数分析一下局部多数模型。我们将局部多数模型中的李雅普诺夫函数定义为总体中的总不一致性（total disagreement），即与相反状态的元胞相邻的所有元胞数量的总和。为了证明这个模型必定会达到均衡，我们必须先证明，如果一个元胞改变了自己的状态，那么总不一致性至少会下降一个固定的数量。

这个证明的数学过程并不太复杂。首先，如果一个元胞改变了自己的状态，那么相对于它的邻居而言它必定是少数。由此我们知道，那时它至少有5个邻居处于相反状态，最多有3个邻居处于相同状态。因此，当这个元胞切换状态后，与这个元胞不一致的元胞数量至少减少了两个（图16-1）。为了计算总不一致性的变化，还必须将与这个元胞相邻的元胞对总不一致性变化的贡献考虑进去。由此，那5个或更多的现在状态一致的元胞的不一致性减少了（每个元胞减少1），同时之前状态一致的那3个或更少的元胞现在则具有更高的不一致性了（每个元胞增加1）。因此，所有相邻元胞的总不一致性至少下降了4。

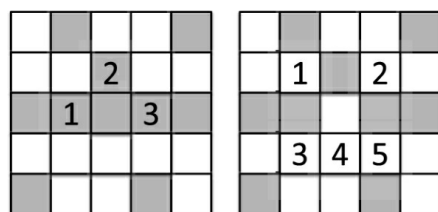


图16-1 局部多数模型中总不一致性的减少

这样一来，我们就证明了，即使某些元胞可能带来更多的不一致性，总不一致性也满足李雅普诺夫函数的条件。因此，局部多数模型必定会收敛到均衡，不仅仅是有些时候或大部分时候，而是所有时间都是如此。我们还可以确定收敛的速度：无论什么时候一个元胞改变自身的状态，总不一致性至少会下降4。由此可知，一个总一致性为100的配置，必定会在25个时期内达到均衡。换句话说，一个总一致性为D的配置，必定会在D/4个时期内达到均衡。正如我们在第15章中已经指出过的，实现的均衡几乎总是一个低效率的斑块状模式，它包括了不少受挫的元胞。

自组织活动模型

接下来，讨论李雅普诺夫函数的下一个应用。这一次，我们要证明自组织活动模型中也存在均衡。这个模型中由一群人和每个人可以选择去做的活动的集合组成。这个模型的关键假设是，每个人都更偏好不那么拥挤的活动，例如，更少人参加同一项活动意味着在健身房不用等待、在面包店和咖啡店不用排队。

这个模型的灵感来自著名经济学家托马斯·谢林（Thomas Schelling）的著作《微观动机与宏观行为》（*Micromotives and Macrobehavior*）。谢林在这本书中描述了城市中出现的各种令人叹为观止的自组织现象：在几乎完全不存在中央计划者的情况下，交通模式、行人流量、停留在公园和餐馆的人数，以及商店库存，到底是怎样自动达到各自的适当水平的？街头那家小店，是怎么每天都能进到4瓶来自密歇根州锡达维尔原产的纯枫糖浆的？巷尾那家面包房，又是怎么做到每天都在关门前大约20分钟卖完当天制作的黑麦面包的？尽管城市内有如此多样性的行动者——游客、店主、居民和送货人等，而且他们对整个城市的信息所知有限，但这种秩序也会涌现出来。

自组织活动模型

一个城市里，有 A 种活动可以参加，每一天都由 L 个时间段组成。在人口规模为 M 的城市中，每个人都要选定一个日程安排。在这里，日程安排是指这个人在 L 个时间段内分配 L 种活动（从一个更大的 K 种可能性的集合中）。一个人要面对的拥挤水平则设定为等于同时选择同一种活动的其他人的数量。

为了证明这个模型是收敛的，我们要证明总拥挤度（total congestion），也就是整个人群拥挤水平的总和满足李雅普诺夫函数的条件。当一个人降低了他的拥挤水平时，就会降低自己对总拥挤度的贡献，并且会使他不再遇到的每个人的拥挤水平减少1，同时使他新遇到的每一个人的拥挤水平增加1。既然他降低了自己的拥挤水平，第一组人的数量会比第二组人更多。例如，假设一个人原本在早上8点去一个拥挤的健身房，在下午4点去一个拥挤的咖啡馆，如果她改为在上午8点去咖啡馆，并在下午4点去健身房，结果发现在上午8点那个时段咖啡馆几乎是全空的，健身房在下午4点时只是有一点点拥挤，那么她降低了自己以及之前遇到的所有人的拥挤水平。她确实使她现在遇到的少数人提高了拥挤水平，但是总拥挤度却下降了（而且至少减少了1）。既然总拥挤度不可能低于零，系统必定会达到均衡。

虽然，一般来说，我们无法保证系统能够找到一个有效率的均衡，但是这种自组织活动模型几乎总能收敛到总拥挤度近乎最小的配置。而在效率低下的配置中，更多人选择在同一个时间段参加某一项活动（比如去健身房），而不是去参加另一项活动（比如去咖啡馆）。如果在这两项活动中拥挤程度的差异很大，那么一个人就可以通过切换自己去健身房和咖啡馆的时间来降低拥挤水平。如果咖啡馆和健身房在另一个时间段有相同数量的人前往，这种切换就可以减少总拥挤度。³

这个模型可以解释当今世界的很多秩序，它可以让我们更加深刻地

理解城市如何在没有中央计划者的情况下通过自组织实现近乎完全有效的配置。它还告诉我们，为什么许多游乐园，比如迪斯尼，却做不到这一点。迪斯尼每天都有新的入园参观者，他们没有时间去尝试新的路线。如果没有“中央计划者”的帮助，迪斯尼将会出现某些景点大排长龙，同时其他景点门可罗雀的情况。因此，为了减少这种低效率状况的出现，迪斯尼对于某些特定的景点，只允许游人提前预约在特定时间内进入，同时让员工引导人们参观不那么拥挤的其他景点。

纯交换经济

我们也可以使用李雅普诺夫函数来探索纯交换经济什么时候可以达到均衡、什么情况下可能无法实现均衡。纯交换经济由一个消费者集合组成，每个消费者都有自己的商品禀赋和偏好。对此，我们可以设想，一群人在市场或集市上与他人交易一些东西，比如茄子、奶酪或地毯。每笔交易都需要双方付出一定的时间和精力。为了让双方有动力完成交易，每一方都必须有所获益，得到超过此交易成本的某个金额。

不过在这里，我们要做的不是直接构建一个总是按固定数量减少并具有最小值的李雅普诺夫函数。恰恰相反，我们要先证明，总幸福感（total happiness）总是按固定数量增加并具有最大值。根据假设，任何时候，只要双方完成交易，他们的幸福水平至少会提升与交易成本相当的数量。此外，由于每个人都拥有固定的商品禀赋，因此总幸福感有一个最大值。李雅普诺夫函数的假设得到了满足，系统可以达到均衡。但是，在这种均衡下，配置不一定是有效的。当然，既然不是有效的，那么市场上的一些人应该能够识别出一些让他们自己更幸福的交易。

在构建这一论点时，我们假设只有参与交易的人才有可能获得幸福（或不幸福）。但是在其他交易形式中，情况可能并非如此。试想一下，假设伊拉克与巴基斯坦达成协议，用自己的石油去交换巴基斯坦的核武器，那么这两个国家的领导人可能会因此而觉得更加幸福，但是从

全球的角度来衡量，总幸福感肯定会下降，世界上其他国家可能会对伊拉克建立核武器库深感不安。

其他国家的人所感受到的这种影响被称为负外部性（negative externality）。当市场上的交易包含了负外部性时，交易就不一定能提高总幸福感了。在前面给出的那个纯交换经济的例子中，当人们交易水果、蔬菜、地毯和工具时，几乎完全不存在外部性。外部性的存在意味着我们不能直接说系统是不是会达到均衡。上面说的那种石油换核武器的交易可能会引发其他交易。例如伊拉克不断增加核武器储备，很可能导致沙特阿拉伯要求自己的盟国提供更多的军事支持，而这反过来又可能导致该地区的其他国家采取行动。在这类交易过程中，全球总幸福感或全球总安全性可能会随着每一个行动而波动不休，甚至不能确定这类交易会会不会在某一点上停止。

李雅普诺夫函数在交易环境中是否存在，取决于负外部性的大小。我以前教过的一个本科生告诉我的一个例子，可以很好地说明这一点。她所服务的公司要搬进新的办公场所，其中有一个很大的房间，配有开放式办公桌，供分析师使用。她的经理提议给所有分析师随机分配办公桌，随后让他们自由进行“办公桌交易”。经理认为这可以导致一个很好的结果，因为他认为自由交易能够带来效率的提高。

这个学生却意识到，即便任何两个相互交易的人都变得更幸福了，他们的前邻居和新邻居也不一定会有同样的感受。如果一个人决定搬到房间的另一侧，那么他当前的邻居，特别是当这个邻居是特意选择在他旁边的时候，可能会感觉受到了伤害。而且，这位前邻居也许不喜欢新邻居，因为新邻居可能会在讲电话时大喊大叫。因此，这位前邻居也可能选择搬工位。这种搬来搬去可能会持续很长时间，每个人的士气都可能因此而受到打击。这个计划看起来相当危险。组织需要自己的员工彼此信任并相互尊重，但是在通过办公桌交易远离对方的两个人身上，是很难维持这种关系的。当经理理解了这个模型后，就放弃了他的计

划。⁴

然而，故事并没有就此结束。这位经理还购买了各种风格和颜色的办公椅，计划随机分配椅子，让人们进行交易。在这种情况下，我的学生再一次使用模型思维告诉他，进行椅子交易不会产生负外部性，而且可以给员工带来不少乐趣。椅子交易是一个纯交换市场。这两种情况，对我们如何运用模型指导行动很有启发。纯交换市场适用于椅子，但是却不适用于办公桌。

没有李雅普诺夫函数的模型

有的时候，我们尝试为模型构造李雅普诺夫函数，但却无法取得成功，尽管如此，我们仍然可以积累知识。通常，我们至少能够了解模型为什么不会产生均衡。在生命游戏中，某些配置会产生均衡，而其他配置则不能产生均衡。当某个配置确实产生了均衡时，就可以写出一个特定于该配置的李雅普诺夫函数。例如，在生命游戏中，采用对角线形式的任何初始配置，每个周期长度都会减少2，因为位于对角线两端的那两个活元胞会死去，并且没有任何新的元胞会变活。这个配置以所有元胞都死光而结束。对于这样的配置，活元胞的数量就是一个李雅普诺夫函数。如果从另一个配置开始，例如，能够产生复杂配置序列的“右五连”，就无法构造出李雅普诺夫函数来，因为系统不会达到平衡。

我们无法构建李雅普诺夫函数，并不意味着模型或系统肯定不能达到均衡。一些系统在所有已知情况下都能够达到均衡，但是没有人能够构造出李雅普诺夫函数。一个著名的例子是所谓的“取一半或乘三加一法则”（HOTPO），它也被称为考拉兹猜想（Collatz conjecture），看似简单，其实不然。“取一半或乘三加一法则”以整数开始，如果是奇数，就乘以3并加上1；如果是偶数，就除以2；当得到的结果等于1时，这个过程就停止。假设从数字5开始，由于这是一个奇数，所以将它乘以3并加1，得到16，然后，将16除以2得到8，8再除以2得到4，然后4再除以2

得到2，最后得到1，达到了均衡。对于任何一个不到 2^{64} 的数字，“取一半或乘三加一法则”都会止步于均衡。

然而，没有人能够证明“取一半或乘三加一法则”是不是最后总能达到均衡。坊间传闻，匈牙利数学家保罗·厄多斯（Paul Erdos）曾说：“对于这样的问题，数学还不够发达。”⁵数学家们无法确定“取一半或乘三加一法则”是否总能达到均衡，这说明了一个更普遍的真理：模型所能提供的，只是证明结果的可能性。我们无法保证肯定可以推导出它们，很多时候，我们提出了一个模型之后，最终却发现要证明结果非常困难（如果不是不可能的话）。

小结

在本章中，我们已经看到，李雅普诺夫函数不仅可以帮助我们证明一个系统或模型能不能达到均衡，还可以告诉我们达到均衡的速度有多快。即便构建李雅普诺夫函数的努力遭到了失败，这种尝试也是有意义的。它们可以提供一些关于复杂性成因的线索。具有外部性的纯交换经济以及所举的交易办公桌的例子都属于这种情况。在这种情况下，我们不能构建一个总是减少或总是增加的全局变量，因此，无法保证这些过程能够达到均衡。

回顾一下模型的7大用途——推理、解释、设计、沟通、行动、预测和探索，我们就会发现李雅普诺夫函数在每个用途上都有所帮助。如上所述，利用李雅普诺夫函数，我们可以推断出系统走向均衡的原因，还可以解释系统收敛到均衡的速度、设计信息系统（例如迪斯尼世界所采取的分时段游览的预约系统）、采取行动（例如对办公桌进行交易是不可取的）、沟通系统如何达到均衡的途径、预测系统达到平衡的时间以及进行探索。我们可以尝试提出假设和构建模型来解释一些令人惊讶的现象，例如自组织的城市。

17 马尔可夫模型

历史是一部循环诗，由时间写在人的记忆上。

珀西·比希·雪莱（Percy Bysshe Shelley）

马尔可夫模型用来刻画以一定概率在一组有限的状态之间不断转换的系统。政治体系可能在民主制度与独裁制度之间转变，市场可能在不稳定与稳定之间变化，一个人也可能在快乐、沉思、焦虑和悲伤之间转换。在马尔可夫模型中，状态之间转移发生的概率是固定的。一个国家在某一年从专制转向民主的概率可能是5%，一个人在一小时内从焦虑过渡到倦怠的概率可能是20%。此外，如果系统可以通过一系列过渡从任何一个状态转换为任何其他状态，并且不存在简单的循环，那么马尔可夫模型就可以达到唯一的统计均衡（statistical equilibrium）。

在统计均衡中，单个实体可以继续在各种状态之间移动，但是各种状态之间的概率分布仍然是固定的。例如，一个关于意识形态的马尔可夫模型，统计均衡允许人们在持自由主义立场、保守主义立场和中立立场之间转换，但是秉持每种意识形态的人口比例将保持不变。在应用于单个实体时，统计均衡意味着实体处于每种状态的长期概率不会改变。例如，当一个人处于统计均衡状态时，他在60%的时间内感到高兴，而在40%的时间里感到悲伤。这个人的精神状态可能每小时都会发生变化，但是他在这些精神状态之间的长期分布却不会发生变化。

这种独特的统计均衡还意味着，结果的长期分布不可能取决于初始状态或事件的路径。换句话说，初始条件是无紧要的，历史也是无紧要的，会改变状态的干预措施也不重要。随着时间的推移，满足这些假设的过程就会“不可抗拒地”走向那一个独特的统计均衡，然后保持不变。在这里，模型再一次揭示了条件逻辑：如果世界符合马尔可夫模型

的假设，那么从长远来看历史并不重要。当然，马尔可夫模型并没有说历史永远是不重要的。首先，马尔可夫模型允许结果是路径依赖的，接下来会发生的事情将取决于当前的状态。其次，马尔可夫模型还允许对历史记录进行长期建模，但对于这种情况，模型的某个假设必定会被违背。

马尔可夫模型可以应用于很多领域。我们可以用马尔可夫模型来解释动态现象，例如民主转型、战争升级和药物干预，还可以用于对网页、学术期刊和体育运动队进行排名，甚至可以用来辨别书籍和文章的作者身份。本章将介绍其中一些应用。我们将从两个简要的例子开始讨论，接着描述证明统计均衡存在性的一般定理。然后，我们转而讨论马尔可夫模型的应用。在本章末尾，会根据我们所掌握的关于马尔可夫模型的知识，重新阐述历史是如何及何时重要的问题。

马尔可夫模型的两个应用

马尔可夫模型由一组状态与这些状态之间的转移概率构成。在第一个例子中，我们将某个人在某一天的精神状态描述为“充实”或“无聊”。这两种精神状态就是模型的两种状态。转移概率表征在状态之间变动的概率。我们可以假设，当精神上“充实”时，一个人有90%的机会继续停留在该状态上，同时有10%的机会变得“无聊”；而当“无聊”时，这个人有70%的机会继续觉得“无聊”，同时有30%的概率变成一个精神上的“充实”的人。

假设上面这些转移概率适用于100名修读生物学课程的学生。在这些学生中，第一天有一半学生觉得这门课程令他们很“充实”，另一半学生则觉得很“无聊”，如图17-1所示。应用上述转移概率，可以预计第二天，会有5名原本觉得“充实”的学生（10%）会变得“无聊”，同时会有15名原本觉得“无聊”的学生（30%）会变得“充实”。这样就会有60名觉得“充实”的学生和40名觉得“无聊”的学生。到第三天，这60名原本觉

得“充实”的学生中应该会有6名感到“无聊”，同时应该会有12名原本觉得“无聊”的学生变成觉得“充实”，从而导致66名学生觉得“充实”、34名学生觉得“无聊”。继续运用这个转移规则，这个马尔可夫过程会收敛到75名学生觉得“充实”、25名学生觉得“无聊”的统计均衡。在这个统计均衡中，学生们会继续在这两种状态之间转移，但是处于每一种状态的学生总数将保持不变。

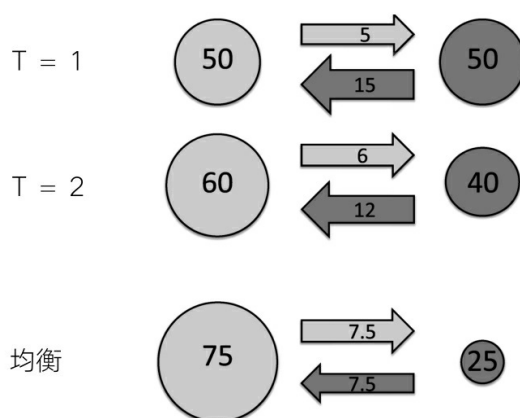


图17-1 一个马尔可夫过程

与此不同，假设这个过程开始时，初始状态是所有100名学生都觉得“充实”，那么到了第二天，将只有90名学生仍然觉得“充实”。到第三天，觉得充实的学生会下降到84名。继续迭代这个过程可以发现，从长远来看，最终仍然会收敛到75名学生觉得“充实”、25名学生觉得“无聊”的统计均衡。这个模型得出了同样的统计均衡。

接下来讨论第二个例子。在这个例子中，我们将不同国家分为三个类别：自由的、部分自由的，以及不自由的。图17-2显示了当今世界上每个类别的国家的百分比，这个分类结果是根据美国自由之家

（Freedom House）截止于2010年的数据（样本期间为35年）得出的。这幅图表示，民主化呈现出了明显的加速趋势。在过去的35年间，自由国家的比例上升了20%。如果这种线性趋势一直持续下去，那么到2040年，所有国家中将会有2/3以上是自由的，而到2080年，将会有8/9的国

家是自由的。

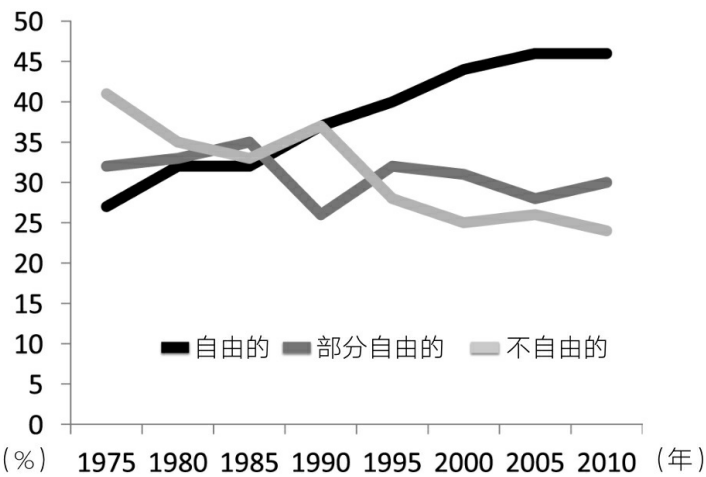


图17-2 对全世界国家的分类：自由的、部分自由的和不自由的国家的百分比

马尔可夫模型会导致不同的预测。为了进行预测，我们将每一个周期的长度设定为5年，并根据过去的数据粗略地校准了转移概率（表17-1）。¹

表17-1 民主化的马尔可夫模型转移概率

		下一周期的状态		
		自由的	部分自由的	不自由的
当前状态	自由的	95%	5%	0%
	部分自由的	10%	80%	10%
	不自由的	5%	15%	80%

如果使用1975年时属于每个类别的国家的百分比来初始化模型，那么如我们所料，校准后的模型几乎完美地匹配了2010年的实际分布：48%的国家是自由的，31%的国家是部分自由的，21%的国家是不自由的（在2010年，这三个类别的国家的实际百分比分别为46%、30%和24%）。如果继续运行这个模型，那么它预测2080年时的情况将会是这

样：62.5%的国家是自由的，25%的国家是部分自由的，12.5%的国家是不自由的。

这个马尔可夫模型的预测似乎并不乐观，原因在于这样一个事实：线性投影假设没有考虑到自由国家既可以转变为部分自由的国家，也可以转变为不自由的国家。随着越来越多的国家变成了自由的国家，从自由的国家转变为不自由的国家的数量也在增加。在现实世界中，出现这种情况的原因是多方面的。首先，实现民主要求国家财政权力机构和行政机构有很高的执行能力。借用政治学家托马斯·弗洛雷斯（Thomas Flores）和伊尔凡·努尔丁（Irfan Nooruddin）的话来说，在某些国家或地区，民主可能不容易扎下根来。²在那些地方，我们有理由预期，会出现从自由国家转变为部分自由的国家的情况，马尔可夫模型也刻画了这种情况。

佩龙-弗罗宾尼斯定理

上面的这两个例子都会收敛到一个唯一的统计均衡，这并非偶然。任何一个马尔可夫模型，只要状态集是有限的、不同状态之间的转移概率是固定的、在一系列转移后能够从任何一个状态变换为任何其他状态，而且状态之间不存在固定的循环，就必定会收敛到唯一的统计均衡。

这个定理意味着，如果满足这四个假设，那么改变初始状态、历史和干预措施，都不能改变长期中的均衡。如果各个国家根据某个固定的概率在独裁统治和民主制度之间变换，那么对其中的某些国家强加民主制度，或对它们进行干预、鼓励民主化，从长期的角度来看不会产生什么影响。如果主流的政治意识形态的变动满足上述四个假设，那么历史也不能影响意识形态的长期分布。如果一个人的精神状态可以用马尔可夫模型来表征，那么鼓励和支持的话语都不会产生长期影响。

佩龙-弗罗宾尼斯定理 (**Perron-Frobenius Theorem**)

一个马尔可夫模型必定收敛于一个唯一的统计均衡，只要它满足如下四个条件：

状态集有限： $S=\{1, 2, \dots, K\}$ 。

固定转换规则：状态之间的转移概率是固定的，即在每个周期中，从状态A转换为状态B的概率总是等于 $P(A, B)$ 。

遍历性（状态可达性）：系统可以通过一系列转换从任何状态到达任何其他状态。

非循环性：系统不会通过一系列状态产生确定的循环。

需要强调的是，从佩龙-弗罗宾尼斯定理中得出的结论不应该说明历史是不重要的，而应该是：如果历史确实是重要的，那么必定会违背模型的其中一个假设。有两个假设，即状态集有限和非循环性，几乎总是成立的。遍历性似乎有可能会被违背，就像当盟国发动战争并且不能转变为联盟时，会出现非遍历性。尽管确实有一些这样的例子，但是遍历性通常也能成立。

状态之间的转移概率是固定的这个限制是最有可能被违背的假设。因此，这个模型表明，如果历史确实是重要的，那么必定存在某种潜在的结构因素改变了转移概率（或者改变了状态集）。考虑一下帮助贫困家庭摆脱困境的扶贫政策。事实证明，导致社会不平等的那些因素通常不会受到政策干预的影响。³同时，马尔可夫模型又表明，改变家庭状态的政策干预措施，例如旨在帮助成绩落后的学生的特殊帮扶计划，或者食物募捐活动，只能在短期内带来改善，不会改变长期均衡。相比之下，提供资源和培训，以提高人们保住工作的能力，进而减少从就业变为失业概率的干预政策则有可能会改变长期结果。无论如何，马尔可夫模型至少为我们提供了一些术语，使我们能够理解状态与转移概率之间

的区别。它也告诉我们一个基本道理——与其改变当前状态，还不如改变结构因素，而后者更有价值。

销量-耐久性悖论

销量-耐久性悖论（sales-durability paradox）说的是，产品或创意的流行程度与其说取决于它们的相对销量，不如说取决于它们的耐久性。只需要将拥有某种类型商品的人的比例设定为状态，就可以用马尔可夫模型来解释这个悖论。在这里我们考虑两种不同的地板，一种是瓷砖（耐用品），另一种是油毡（销量更大的商品）。当销量更大的商品（在这里这个例子中是油毡）不那么占主流时，就会产生这种悖论。

在我们的模型中，假设油毡的销量是瓷砖的3倍。为了刻画耐久性差异，假设每年有 $1/10$ 的人必须更换他们的油毡地板，而需要更换瓷砖的人则只有 $1/60$ 。在由此而得到的马尔可夫模型的均衡中，有 $2/3$ 的地板都是用瓷砖铺就的。⁴

销量-耐久性悖论背后的逻辑，也可以用来解释市场份额与品牌忠诚度（某人改用其他品牌的产品的可能性）之间的正相关关系。在马尔可夫模型中，更低的品牌忠诚度在均衡状态时必然意味着更低的市场份额，因为忠诚度所起的作用就像耐久性一样。这种经验规律有时也被称为“祸不单行法则”（double jeopardy law）。如果一个企业的产品的品牌忠诚度较低，那么其销售量往往也较低。⁵

一对多的马尔可夫模型

马尔可夫模型可以应用于各种各样的环境。我们可以用马尔可夫模型来对四种核酸之间的遗传漂变进行建模分析：腺嘌呤（A）、胞嘧啶

(C)、胸腺嘧啶 (T) 和鸟嘌呤 (G)。如果每种核酸都有很小且相等的概率成为其他三种类型之一的核酸，就可以写出一个刻画遗传漂变的转移矩阵。

我们也可以用马尔可夫模型对身体健康演变的轨迹进行建模，方法是将不同的健康类别（如优秀、中等和糟糕等）设定为不同状态。这样的模型可以用来评估药物治疗、行为改变和手术等健康干预措施如何转移概率和均衡分布。那些能够产生更好均衡的健康干预措施是值得追求的。⁶

马尔可夫模型还可以用于识别国际危机的不同模式，并能够用于区分会导致战争的过渡与会带来和平的过渡。⁷不过，在这个领域的应用要求我们估计两种不同的模型：一种模型中，危机导致了战争，另一种模型中，战争爆发前实现了和解。如果这两个模型中的转移概率有显著差异，那就可以对现有的各种模式进行比较，例如轰炸、劫持人质、不交换囚犯以及逐步升级的强硬姿态等，然后看哪个过程对数据的拟合更优。

这种通过马尔可夫模型将不同模式区分出来的方法，还可以用来辨别书籍或文章的作者。只要事先已经确定了某个作者的某些已知著作，就可以估计出一个词之后出现另一个词的概率。例如，在本书中，单词 *the* 跟在 *for* 后面的次数，相当于单词 *example* 跟在 *for* 后面次数的四倍。我们可以将这类信息表示为一个大转移概率矩阵。本书的矩阵看上去肯定与其他人的书的矩阵有所不同。如果为 Melville（梅尔维尔）和 Morrison（莫里森）分别构建单词转移矩阵，那么我们肯定会看到它们在单词对之间的转换是不同的。⁸

运用这种方法，我们可以通过构建模型来确定《联邦党人文集》（*Federalist Papers*）所收录各篇文章的作者。《联邦党人文集》共收录了85篇文章，分别由亚历山大·汉密尔顿（Alexander Hamilton）、约翰·杰伊（John Jay）和詹姆斯·麦迪逊（James Madison）在1787年和1788年

写成，目的是说服当时的纽约州民众支持美国宪法。每篇文章的署名，都用了统一的笔名“Publius”。尽管大多数文章的真正作者都已经确定了，但仍有少数几篇文章一直存在争议。

一项利用马尔可夫模型来辨别作者的研究认为，所有有争议的文章都是詹姆斯·麦迪逊写的。⁹即使汉密尔顿或杰伊写了这些文章，他们也是在模仿麦迪逊的风格。对哲学家阿琳·萨克森豪斯（Arlene Saxonhouse）发表的4篇论文和未发表的12篇短篇论文的分析结果表明，其中至少有3篇论文都有很高的概率可以归入霍布斯名下。¹⁰当然，在这两个例子中，模型都不一定能给出准确的答案。但是，这些模型确实生成了知识。我们要依靠自己的智慧做出判断：对这个模型与其他模型或直觉结论，该如何进行权衡。

我们要讨论的最后一个应用，是谷歌公司如何运用马尔可夫模型构造谷歌最初的网页排名（PageRank）算法。网页排名是万维网的搜索方式。¹¹万维网由链接连接起来的网站组成，为了估计出每个站点的相对重要性，可以计算一个站点连入和接出的链接数量。在图17-3所示的站点网络中，站点B、C和E各有两个链接，A有一个链接，D没有链接。这种方法提供了对重要性的粗略估计，但是它有很大的缺陷。站点B、C和E都有两个链接，但是站点E看上去似乎比站点B更加重要，这是由它在网络中的位置决定的。

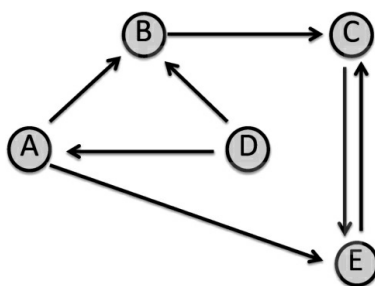


图17-3 万维网上站点之间的链接

网页排名算法将每个站点都视为马尔可夫模型中的一个状态。如果两个站点共享一个链接，那么就在这两个站点之间分配一个正的转移概

率。我们暂且为任何链接分配相等的概率，也就是说，假设在A上的搜索者有同样的可能性移动到B或E上。如果搜索者来到了E上，那么将永远交替出现在C和E之间。或者，如果搜索者选择了B，那么他还是会去C，然后再一次开始在C和E之间交替出现。实际上，从任何站点开始都会导致交替出现在C和E之间这个结果。我们发现C和E似乎是更加重要的站点。然而不幸的是，这个模型不满足佩龙-弗罗宾尼斯定理的两个假设。该系统无法从任何站点到达任何其他站点：无法从C到达D，转移概率在C和E之间创建了一个循环。

为了解决这两个问题，谷歌公司的算法加入了一点：从任何站点都能够以一个很小的随机概率移动到任何其他站点，如图17-4所示。现在，这个模型就满足佩龙-弗罗宾尼斯定理的所有假设了，而且存在唯一的均衡。于是，所有站点都可以根据它们在那个均衡中的概率进行排序。一个从A开始的搜索者，最有可能在几次搜索后以到达C或E结束。一旦到达C或E之后，他将会在这两个站点之间反复来回，直到尝试前往一个随机站点为止。如果他到了A或者D，那么回到C的路径很可能会经过B或者E。因此，B的排名应该高于A或D。图17-5所示的唯一统计均衡表明，这个结论是正确的。A、B和D都很少被访问，其中，B的访问量最大。

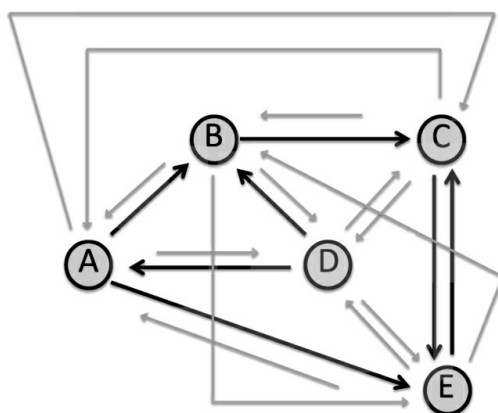


图17-4 在站点之间添加随机移动

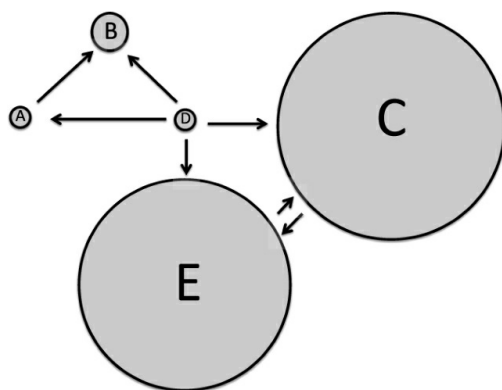


图17-5 网页排名模型的统计均衡

网页排名可以看作随机游走与马尔可夫模型的组合。如果将网页排名视为一种算法，就会发现可以用它来生成任何网络的排名。我们可以让节点代表棒球队或足球队，再用转移概率表示一支球队击败另一支球队的时间百分比。¹²如果球队之间只打一场比赛，那么可以根据胜率来分配转移概率。由此而得出的排名虽然不是最终的，但却是对专家主观评估意见的有益补充。我们还可以利用食物链数据，通过网页排名算法来计算物种之间的相对重要性。¹³

小结

马尔可夫模型描述了以固定的转移概率在不同状态之间转换的动态系统。如果再假设这个过程能够在任何两个状态之间转移，并且这个过程不会产生循环，那么马尔可夫模型就可以得到唯一的统计均衡。在均衡中，人或实体在各个状态之间移动，但是各个状态的概率分布不会发生改变。由此可见，当一个过程接近均衡时，概率的变化就会减弱。用曲线图表示，就表现为曲线的斜率走平。回想一下，我们在讨论线性模型时对美国加利福尼亚州人口增长的讨论。加利福尼亚州的人口增长已经放缓，因为随着人口的增长，离开加利福尼亚的人数也在增加。即便离开的人数所占的比例没有发生变化，这个结论也是成立的。

在应用马尔可夫模型解释现象或预测趋势时，建模者对状态的选择

至关重要。状态的选择决定了这些状态之间的转移概率。一个简单的关于药物成瘾行为的马尔可夫模型，可能只需要呈现两种状态：或者是药物成瘾者，或者是正常的人。而更精细的模型则可以根据使用频率来区分药物成瘾者。无论对状态的选择如何，如果上面的四个假设都成立（并且在这种情况下，关键检验将变为转移概率是不是能够保持固定不变），那么系统将会存在一个唯一的统计均衡。系统状态的任何一次性变化都最多只能产生一些暂时性的影响，减少均衡中的药物依赖必须改变其转变概率。

按照同样的逻辑，我们可以推断，那些试图通过为期只有一两天的活动来激发学生学习兴趣的做法，可能不会产生什么有意义的影响。与此类似，进入社区“送温暖”、来到公园“捡垃圾”的志愿者也可能无法带来什么长期收益。任何一次性的资金涌入，无论其规模大小，影响都会消失，除非它改变了转移概率。2010年，马克·扎克伯格（Mark Zuckerberg）向新泽西州纽瓦克市的公立学校捐赠了1亿美元，并吸引了不少跟风捐赠者。这种一次性捐赠，尽管摊到每个学生头上达到了大约每人6 000美元，但对考试成绩却几乎没有产生任何可衡量的影响。¹⁴

马尔可夫模型是通过区分以下两类政策来指导行动的：一类政策能够改变转移概率，而改变转移概率可以产生长期影响；另一类政策只能改变状态，并且只能产生短期影响。如果转移概率无法改变，那么我们必须定期重置状态才能改变结果。沉溺于辛劳工作可能会产生导致好强、自私和压抑的心理状态转移概率，而每天锻炼、冥想或参加宗教活动则可能帮助人们以一个感恩的、富有同情心的、放松的心理状态迎接每一天。周末休息也有类似的功能，已婚夫妇不时过一过约会之夜也有很好的效果。这两者的共同作用是，能够暂时使一个人的状态远离均衡。

当然，并不是每个动态系统都满足马尔可夫模型的假设。在不满足马尔可夫模型假设的情况下，历史、干预政策和事件都可能会产生长期

影响。例如，在波利亚过程中，结果改变了长期均衡。对系统的重大干预或冲击可能会改变转移概率甚至是整个状态集。蒸汽机、电力、电报或互联网等重大技术变革，改变了经济的可能状态集。重新界定权力架构或制定新政策的政治和社会运动，也会改变状态集。因此，我们也许更应该将历史视为一个马尔可夫模型序列，而不是视为一个向不可避免的均衡方向发展的过程。

马尔可夫决策模型

马尔可夫决策模型（Markov decision model）是对马尔可夫模型的一种修正，方法是将行动包括进来，行动会带来回报，而回报则以状态为条件，还会影响状态之间的转移概率。考虑到行动对转移概率的影响，最优行动并不一定是能够最大化即时回报的那个行动。

例如，要在上网与学习这两个行动之间做出选择。上网总能带来相同的回报。而当学生选择学习时，则有两种可能，既可能觉得充实，也可能觉得无聊。如果觉得充实，学习就可以获得高回报，如果觉得无聊，学习就只能获得低回报。

为了加入行动对转移概率的影响，假设一个觉得学习无聊的学生转为在上网时，仍然会处于无聊状态；而一个觉得学习充实的学生转为在上网时，有一半的时间会变得无聊。假设一个学习的学生有75%的机会在下一个时期处于觉得充实的精神状态，而无论他当前的状态如何。于是：

行动：上网（U），学习（S）

状态：觉得无聊（B），觉得充实（E）

奖励结构

	无聊 (B)	充实 (E)
上网 (U)	6	6
学习 (S)	4	8

转移映射

	无聊 (B)	充实 (E)
上网, 无聊 (U, B)	1	0
上网, 充实 (U, E)	1/2	1/2
学习, 无聊 (S, B)	1/4	3/4
学习, 充实 (S, E)	1/4	3/4

马尔可夫决策模型的解决方案由每个状态下采取的行动构成。之前讨论过的短视最优反应行为，在每个状态下都选择能够最大化奖励的行为。在现在这个例子中，这种选择对应于无聊时上网、精神充实时学习。

但是，这种短视的解决方案会导致学生陷入无聊状态。一旦发生了这种情况，他们就会选择上网，并在所有剩余时间内一直保持无聊状态。因此，他们的长期平均回报等于6。而总是选择学习的解决方案则会在他们75%的时间里处于充实状态，只在25%的时间里处于无聊状态，从而得到的长期平均回报为7。这个解决方案产生了更高的平均回报，因为他们更多地处于充实的精神状态。

正如这个例子所表明的，将一个决策问题表达为一个马尔可夫决策模型，可以告诉我们更好的行动是什么。通过考虑行动对状态的影响，我们会做出更明智的选择。晚睡与早起和锻炼相比，会产

生一个更高的直接回报，购买昂贵的咖啡比自己动手制作咖啡产生更高的回报。然而，从长远来看，我们可能会更乐于坚持锻炼和节省咖啡钱。那么，我们需要一个模型吗？不一定。相反，我们也许只需要时时记起《圣经·箴言》21:17就可以了：“爱宴乐的，必致穷乏；好酒爱膏油的，必不富足。”这可能是对的；但是我们同时可能记得《圣经·传道书》8:15所说的：“我颂赞喜乐，因为世人在天日之下再好不过的，就是吃喝欢乐。”是的，我们总能找到一对相反的谚语。通过将我们的选择嵌入马尔可夫决策模型中，可以使用逻辑来确定在给定的情境下，哪些常识性的建议真的有用。

18 系统动力学模型

我们尚未充分理解支配系统行为的原理。

杰伊·赖特·福雷斯特（Jay Wright Forrester）

在本章中，我们将介绍一系列系统动力学模型。¹这些模型可以分析那些有反馈和相互依赖性的系统，可以用于对生态和经济、供应链和生产过程建模。系统动力学模型提高了我们通过包含正反馈和负反馈的逻辑链进行思考的能力。系统动力学模型通常要包括源（source）、汇（sink）、存量（stock）、流量、速率和常数等组成部分。源产生对系统的输入，汇吸收输出，存量跟踪变量的水平，流量刻画各存量水平之间的反馈，速率和常数用于流量，流量可以是随时间变化的，也可以是固定不变的。

系统动力学模型可以同时包括正反馈和负反馈。当变量或属性的增加导致同一变量或属性的更大增加时，就会出现正反馈，第6章中讨论的马太效应就是如此，成功会带来更大的成功，畅销会带来更大的畅销。就学术论文和专利而言，引用会导致更多的引用。

负反馈会抑制趋势。我们必须小心，不要从“负”这个词的字面含义推出任何规范性的含义。负反馈往往能够带来理想的属性，防止泡沫和崩盘。当我们吃东西吃到饱时，大脑会收到停止进食的信号；当企业的利润超过了正常的经济回报时，竞争者就会进入，使在位者的利润减少并防止他们剥削消费者；当一个物种过度繁殖时，它的成员会为食物展开竞争，从而使增长率下降。在这些情况下，负反馈都有助于系统层面的鲁棒性。

使用系统动力学模型，我们通常可以确定复杂性的原因。当系统既包括正反馈又包括负反馈时，就会产生复杂性。在生命游戏中，活元胞

会导致新元胞变活，但是过度拥挤又会导致元胞死亡。

这些模型可以将流量和存量水平表示为数学函数。我们可以校准这些定量模型，以便解释过去的存量值、预测未来值，同时估计干预措施的影响。然后，我们就可以利用这些模型解释、预测和指导行动。这些模型也可以是定性的，可以用加号或减号标记每个箭头以表明内在逻辑。²

本章内容由五部分组成。在第一部分中，为了引入系统动力学的术语体系，我们先构建了一个定性的面包店模型。在第二部分中，我们构建了一个基于洛特卡-沃尔泰拉方程（Lotka-Volterra equation）的捕食者-猎物模型。我们这个版本假设了存在狐狸和野兔两个相互作用的种群，并同时嵌入了负反馈和正反馈。在第三部分中，我们阐述了如何利用系统动力学模型来预测恶性循环。在第四部分中，我们描述了WORLD3模型，它是一个关于全球经济的巨型模型。最后，在本章的结论部分，我们讨论了系统动力学模型为什么经常会产生违反直觉的结果，这些结果证明了人类直觉推理的局限性和模型作为逻辑辅助工具的价值。

系统动力学模型的各个组成部分

任何一个系统动力学模型都由源、汇、存量和流量组成。源产生存量；存量是某个变量的数量或水平；流量描述了存量水平的变化；汇能够捕获来自存量的流量输出；汇和源是不包含在模型中的过程的“占位符”；存量水平会根据源和流量随时间推移而变化。例如，在一个关于游乐园的系统动力学模型中，公园中的人（一个存量）的人数会随着人们的到来（一个源）而增加，而增加的速度可能又取决于其他参数，例如天气状况、广告投入和门票价格等。

系统动力学模型使用如图18-1所示的表征系统。源和汇用云表示，

存量用方框表示，流量用箭头表示并附以加号或减号标识，可变流量用倒三角形表示，而不变流量则用被流量箭头对穿的圆圈表示。正箭头表示正反馈（更多会带来更多），负箭头表示从一个变量到另一个变量的负反馈。

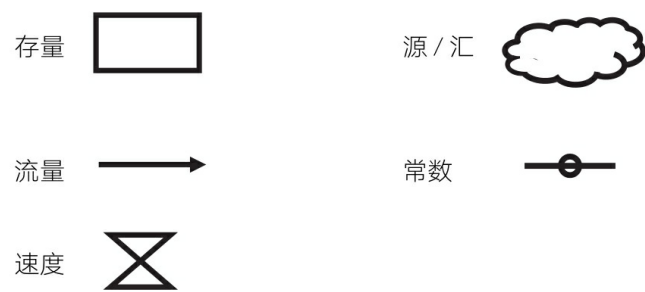


图18-1 系统动力学模型的组成部分

为了让读者尽快熟悉这套符号，在这里先给出一个由面包师、面包和顾客组成的简单的面包店系统动力学模型：面包师制作面包，顾客购买面包。如果面包师生产面包的速度超过了顾客购买面包的速度，面包的库存量就会增加，面包店将会堆满面包。或者，如果卖出面包的速度超过了面包师生产面包的速度，面包店将永远处于卖光的状态。为了使模型更加真实，我们可以假设让面包师根据面包的存量来调整面包生产的速度，如图18-2所示。在这幅图中，包括了一个从面包库存量到面包师生产面包的速度的流量（一个箭头）。我们在这个箭头上放了一个负号，以表示随着面包库存量的增加，生产面包的速度会下降。如果适当地调整速度，模型将产生一个均衡，使面包生产速度收敛于顾客购买面包的速度。

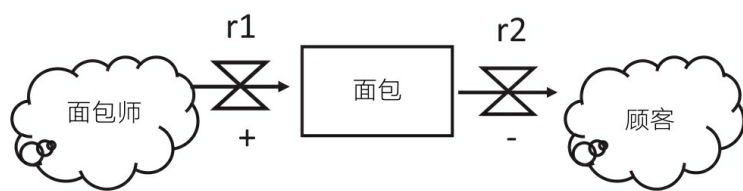


图18-2 一个关于面包店的系统动力学模型

为了使模型更加真实，还可以加入第二个存量排队长度，它等于在面包店外等待的人数。然后再加入第二个源，即潜在顾客，它会增加排队的人数。排队长度较短或适当，可能会吸引更多顾客，而过长的排队长度则会让客户流失，导致他们调头而去。为了刻画排队长度对来自源到达速度的影响的这种可变性，我们在这个箭头上标记了（+/-）的符号，还在从排除存量线条到顾客购买面包速度的箭头上加了一个加号，表示排队的人越多，客户购买面包时做决策的速度越快，如图18-3所示。

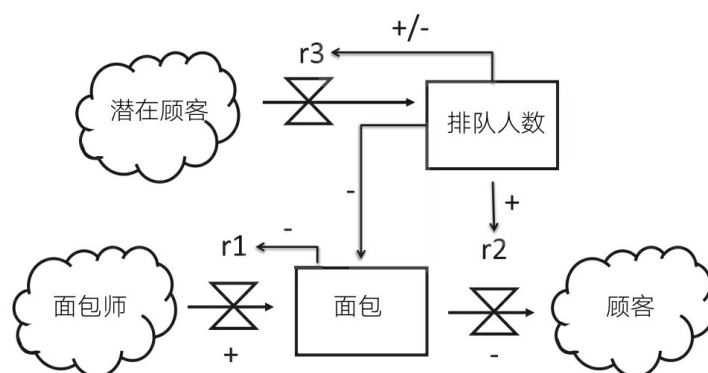


图18-3 一个更加精细的关于面包店的系统动力学模型

这个模型可以根据数据进行校准。我们可以估计出人们根据排队长度加入排队队伍的速度。然后，面包师可以根据面包存量和排队长度确定烘焙的最佳速度。这个速率将提供一个起点，从中可以获得更好的速率。即使没有校准，编写模型的行为本身也增加了价值。面包师很清楚排队的长度对整体销售量的重要性。

捕食者-猎物模型

我们现在介绍捕食者-猎物模型，这是一个用来刻画野兔数量（猎物）与狐狸数量（捕食者）之间关系的生态模型。这个模型包括两个正反馈：野兔生下野兔、狐狸生下狐狸。它还包括一个负反馈：狐狸吃野

兔。该模型假设野兔的存量水平很高，狐狸会产生更多的后代。图18-4定性地表示了这些假设，但没有量化关系。从图中可以看出，随着狐狸数量的增加，野兔数量的减少，从而又导致狐狸数量减少。而随着狐狸数量的下降，野兔数量应该增加，进而导致更多的狐狸。逻辑表明了循环的可能性，也可能是均衡，但我们无法确定。

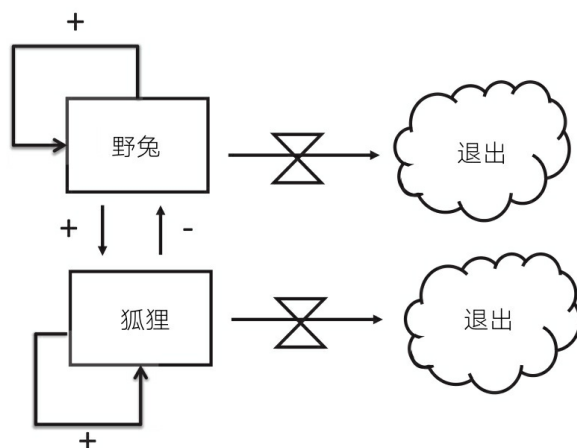


图18-4 捕食者-猎物模型的系统动力与模型

为了更加深入地了解到底发生了什么，我们需要构建出这个模型的定量版本。假设流量是线性的，取决于存量水平。再假设在没有狐狸的情况下，野兔的数量以固定的速度增长，而在没有野兔的情况下，由于缺乏食物，狐狸的数量以固定的速度减少。再假设，在这个模型中，野兔和狐狸相遇的概率与狐狸数量乘以野兔数量的积成正比。为了捕捉发生这种相互作用时吃野兔的狐狸的行为，假设狐狸的数量以某个恒定的速度乘以上述乘积的速度生长，同时野兔的数量则以某个恒定的速度乘以上述乘积的速度减少。由此而得到的方程就是通常所称的洛特卡-沃尔泰拉方程。

洛特卡-沃尔泰拉方程

假设一个生态系统由 H 只野兔和 F 只狐狸构成。野兔的数量以 g 的速度增长，狐狸的数量则以 d 的速度减少。当野兔和狐狸相遇

时，野兔以 a 的速度死亡，狐狸以 b 的速度增长。根据这些假设，可以给出如下微分方程组：³

$$\dot{H} = gH - aFH \quad \dot{F} = bFH - dF$$

从这个微分方程组可以推导出两个均衡，一个是灭绝均衡，即 $F=H=0$ ；一个是内点均衡，即 $F = \frac{g}{a}$ 和 $H = \frac{d}{b}$ 给出。

这两个微分方程分别描述了野兔和狐狸的数量随时间而发生的变化。当方程等于零时，野兔和狐狸的数量不再改变，系统处于均衡状态。第一个均衡，即灭绝均衡（extinction equilibrium），是说野兔和狐狸都不复存在。因此，这个模型预测，在一定条件下，捕食者-猎物这种关系会导致两个物种同时灭绝。当然，不是所有情况下都会发生这种事情，不然的话，地球上也就不会有任何物种存活了。

第二个均衡是内点均衡（interior equilibrium）。在内点均衡中，狐狸和野兔的数量均为正。在这种均衡状态下，狐狸的数量随野兔的增长速度而增加，但是如果狐狸和野兔之间的每次相遇都会导致野兔种群规模以更快的速度减少的话，狐狸的数量就会减少。这两个结果都很直观：如果野兔繁殖更快，那么这个生态系统可以养活更多的狐狸；如果每只狐狸都需要更多的野兔才能活下去，那么这个生态系统就只能养活更少的狐狸。两个结果都符合我们的直觉，我们想要的也正是这样的结果：模型应该能产生符合直觉的结论。

但模型也应该能够产生不那么直观的结论。幸运的是，这个模型也正是如此。它表明，狐狸的均衡数量完全不依赖于狐狸的死亡率。如果狐狸以更快的速度死亡，野兔的均衡数量就会增加，剩下的狐狸的食物就会更加丰富，而这就意味着狐狸的增长速度更快。狐狸的快速增长速度恰好抵消了狐狸更高的死亡率。

类似的逻辑同样适用于野兔种群。野兔的均衡数量不依赖于野兔的增长速度，也不依赖于野兔被狐狸吃掉的速度，相反，野兔的数量取决

于狐狸的死亡速度和狐狸将野兔吃掉“转化为”更多狐狸的速度。对于这些结果，直觉“失败”了，因为我们无法直观地想清楚这里面的反馈机制。野兔增长速度提高的直接影响是会有更多的野兔，间接影响是会有更多的狐狸，而这又意味着更少的野兔，这两种效应会相互抵消。能够得出像这样的非直观结论，正是系统动力学模型的标志特征。直觉在这里“失败”了，因为我们只锁定了直接影响而未能思考整个逻辑链。即便增加（或减少）速度或流量的直接影响是增加（或减少）存量，系统以正反馈和负反馈的形式产生的影响也意味着，其他存量的值也会发生改变，因此速度或流量变化的净效应可能会减少、会被抵消，甚至可能会被逆转。

通过数学演算，我们可以证明洛特卡-沃尔泰拉方程存在这两个均衡，但是并不知道这两个均衡当中哪一个会实现。确实，如果模型在一开始就处于某个均衡状态，那么它将保持在那个均衡状态。但在运行模型之前，我们并不知道这个方程组是否将会产生均衡、周期性、随机性，抑或复杂性。我们所知道的只是，均衡的确存在。

对方程的模拟会产生滞后循环（lagged cycles）。首先，其中一个物种变得“人口众多”，然后，它的数量开始减少，另一个物种的数量开始增加。实证研究表明，这种循环确实很常见。图18-5显示了位于苏必利尔湖中的一个纵长72千米的名为罗亚尔岛（Isle Royale）的岛屿上，狼（捕食者）的数量与驼鹿（猎物）的数量50年来一直处在此消彼长的过程中。不难注意到，捕食者和猎物物种规模的波动，都呈现出了滞后循环的特征。由于模型省略了地理因素、其他物种的影响、天气的变化、物种内部的异质性等，因此其模式并不像我们预期的那么规则。

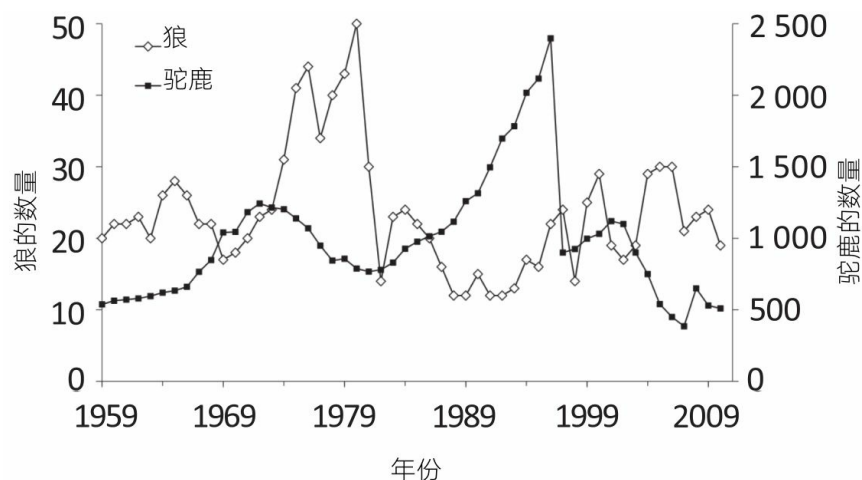


图18-5 罗亚尔岛上的狼和驼鹿

对洛特卡-沃尔泰拉方程的上述分析进一步支持了我们早先的观察结论，也就是不应该将均衡的存在与均衡的实现混为一谈。在图中所示的这种情况下，系统产生的是循环而不是均衡。不过，动态循环是围绕着均衡发生的。因此，均衡可以告诉我们，狐狸和野兔的数量平均来说是多少。由此可见，我们早先得到的违反直觉的结果，也就是提高狐狸（或野兔）的增长速度对狐狸（或野兔）的均衡规模没有影响，在总体层面上仍然是成立的。

系统动力学模型的应用

系统动力学模型既可以包括正反馈回路，也可以包括负反馈回路。正反馈回路会导致良性循环，例如，当两个国家之间的信任度提高，会导致贸易上升、军事对抗减少，从而进一步增进信任。但是，正反馈回路也可能导致恶性循环，例如，某个地区就业机会的减少，可能会削弱人们学习掌握技能的动力，而这反过来又可能导致企业由于缺乏合格的劳动力而迁离这个地区，进而进一步减少工人去努力获得技能的动力。

系统动力学模型可以帮助我们预测恶性循环。2008年，许多国家的经济都面临着严重的金融压力。当资产价格下跌时，过度杠杆化的银行

离破产只有一步之遥，投资者和储户开始担心他们投资的安全。有些国家，如美国，只在一定限度内对银行存款进行保险；而在其他一些国家，例如在澳大利亚，则根本不存在存款保险。

为了防止恐慌，澳大利亚决定引入存款保险制度。逻辑上看似乎是合理的：对存款保险可以防止银行挤兑。但是，这种想法其实只考虑到了系统的一部分。在引入存款保险制度的过程中，这种制度本身的一个缺陷会导致悲剧性的后果。如果把相关的系统动力学模型写出来，这一缺陷就会变得很明显。在关于金融系统的系统动力学模型中，每个银行（都是一个存量）都有一定的资产规模。储户将钱存入银行并获得一定回报，银行的借款人用这笔钱进行投资。存款保险制度保证存款人放在银行的资金的安全。

此外，人们还会将资金投入股票市场和货币市场基金。每种类型的投资都是存量。一旦我们开始在各种存量之间绘制箭头，即存量之间的流量，存款保险这种制度的缺陷就会变得很明显。存款保险制度的直接影响能够提高银行的安全性，从而使银行更具吸引力（图18-6中的箭头1）。但是，它同时还会使其他类型的投资变得不那么有吸引力。想象一下，假设在那个动荡的金融危机时期，你自己是一个投资者，在银行和货币市场上都投入了一定资金。现在，你的银行存款已经投保了，但是你的货币市场基金则没有。因此，谨慎的行动是增加银行存款（箭头2）并退出货币市场（箭头3）。

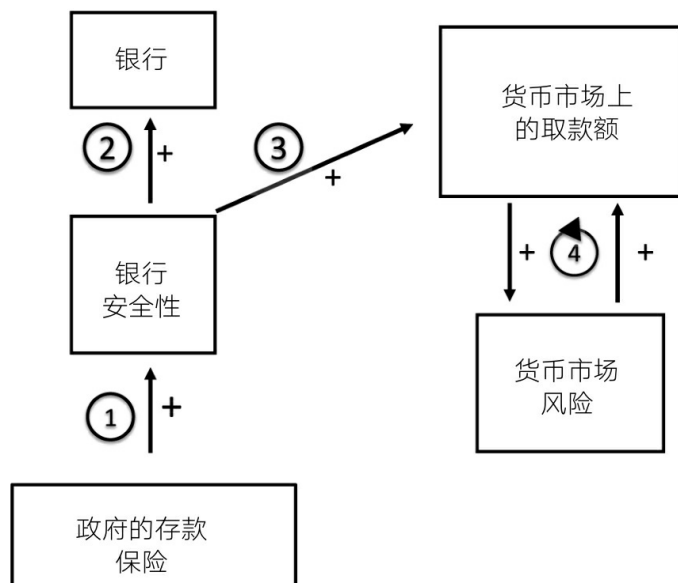


图18-6 金融系统的系统动力学模型

随之而来的是一个恶性循环：货币市场投资的减少使它的风险增加，而风险增加又导致更多的资金从货币市场撤出，从而形成了一个正反馈回路（图18-6中的循环路径4）。撤出资金提高了风险，而风险的上升又会导致更多的提款，从而导致更大的风险。这个政策似乎肯定会造成货币市场的崩溃，现实也确实如此。在决定为银行存款提供保险的短短四天之后，政府就不得不冻结了货币市场上的所有账户，不然整个行业都会彻底崩溃。显然，这个决定带来了灾难性的后果。数以百万计的退休人员，都需要从这些账户中提取资金去购买食品、交纳房租，支付其他基本必需品。⁴

虽然从“事后诸葛亮”的角度来看，这种恶性循环似乎是显而易见的，但是我们其实无法保证，即使澳大利亚的决策者当时已经构建了一个系统动力学模型，他们明白采取这项政策会带来的恶劣后果。重要的是，写出这样一个模型，至少有可能会帮助他们看到，银行存在保险制度会在金融体系内产生更加广泛的影响。如果是这样，他们应该是有可能注意到这个恶性循环的。这个例子还说明了数据的局限性。来自其他国家的数据表明，保险存款制度有助于稳定金融体系。但是，在这些

国家，存款保险制度并不是在金融危机期间创立的，因此这种经验数据对澳大利亚当时的政策选择没有意义。

WORLD3模型

接下来，考虑一个更复杂的涵盖了全球经济的系统动力学模型，这个模型通常被称为WORLD3模型。WORLD3模型最早出现于20世纪70年代，当时这个模型预测世界经济行将崩溃，除非各国政府改变自己的经济增长政策和环境政策。⁵ WORLD3模型包括了在一个共同的框架内以不同速度增长的多个彼此互动的过程，因而利用它，决策者能够看到各种各样的相互依赖关系。⁶但是，主流经济学家往往不看重WORLD3模型，因为在他们看来，这个模型一方面过于复杂，另一方面又没有考虑到经济行为主体的理性反应。

这个模型假设，人口和经济产出每年以某个不变的百分比增长，但是经济产出会造成污染。随着时间的推移，土地的生产力下降，人口规模将超过经济提供商品的能力限度，那时世界经济就会崩溃。这个预测不禁让我们联想到托马斯·马尔萨斯在差不多两个世纪前提出的那个可怕的警告。

这个模型包含了大约150个变量、300个方程和500个参数，生育率、经济增长率和土地使用率等无不包括在内。要校准这个模型，必须根据数据估算出这些参数的增长率。WORLD3模型还包括了变量之间的相互作用，这意味着多个参数的变化通常会产生非线性效应。因此，检验模型的鲁棒性需要不断尝试同时更改两个变量、三个变量的多种组合。500个参数可以产生超过12 000个参数对和超过2 000万个参数三元组。对任何人来说，这无疑都太多了。

这个模型的预测是，到2100年，全世界人口将会减少到40亿。不过，约翰·米勒（John Miller）发现，只要调整两个参数——分配给消费

的工业产出比例和女性的生育周期长度，就能够使模型预测的世界人口几乎增加一倍，即达到74亿。正反馈带来了巨大的增长。更长的生育周期意味着更多的孩子，他们又需要更多的食物。增大食物在产出中的份额可以让更多儿童生存，生存下来的女性有更长的生育周期，于是生下来的孩子更多。最终结果将是人口大规模增长。⁷

参数的小小变化，就会导致人口增加一倍，这个结果会带来很多麻烦。不过，结果依赖于参数这个事实本身并不是一个弱点。相反，构建这个模型的目的就是为了指导行动、帮助识别有效的政策。例如，这个模型证明，降低生育率能够降低人口增长速度（事实证明确实如此）。再者，由于这个模型已经经过了校准，因此它还可以给出人口增长速度到底可以减少多少的估计值。因此，我们可以将这个模型包括在一个模型集合中，以便得到更加准确的预测。

随着时间的推移，这个模型最初的预测变得不那么准确了，部分原因是随着人口的增长，人口增长率已经放缓，它们不再符合模型的假设，这正是经济学家早就预料到的一种适应性反应。⁸虽然WORLD3模型的倡导者接受了这种批评，但是他们也指出，这个模型的许多预测，包括与经济增长和世界总人口有关的预测都是非常准确的。至于生育率的降低，这个模型的倡导者认为，如果WORLD3模型的失效是它自己的原因所导致的——因为这个模型使人们认识到了人口过剩问题的严重性和环境的重要性，他们对这种错误的态度是乐见其成的。

小结

在构建系统动力学模型时，我们选择了关键的组成部分（存量），描述了这些部分之间的关系（流量）；然后我们用系统动力学模型进行模拟实验，以发现它的含义。这些模型与马尔可夫模型的不同之处在于，在这里速度是可以调整适应的（它起到了转移概率的作用）。因此，这种模型不一定会达到均衡。我们必须运行模型才能看清楚会发生

什么。此外，因为不必求解结果，所以也不用担心假设的易处理性。

系统动力学模型可以包含许多变量，而且可以包括这些变量之间任何类型的反馈。我们当然也可以构建不存在这些东西的模型，但是一旦绘制出了定义存量的方框，建模者就几乎无法在它们之间绘制出一些箭头来。建模者觉得自己有义务追问：“还有哪些其他变量也可能会受到影响？这些变量的变化又如何反馈回当前的模型中？”这些问题会带来更加精细的模型。

当然，这种灵活性也可能是需要付出代价的：创建的存量和流量越多，模型就会变得越难以理解。构建一个有用的系统动力学模型的“艺术”体现在，既要包括足够多的细节来揭示我们的直觉哪里行不通，但是也不能包括过多的细节，以至于会创造一个像现实世界一样混乱的泥淖。最有用的系统动力学模型都位于那个边界上。这些模型可以揭示出意想不到的效应，并有助于更好地采取政策行动。正如在本章前面的内容中看到的，即使是意图极好的政策，例如澳大利亚政府推出的存款保险制度，也可能产生不良后果。

系统动力学模型还阐明了负反馈回路会怎样限制干预政策的效果。强制要求汽车增加安全功能的法律，例如安装防抱死制动系统或安全气囊，可能会导致人们更加鲁莽地开车，进而导致更多人受到伤害。道路加宽可能会导致更多人搬到郊区，从而进一步加剧拥堵。减少香烟中的尼古丁含量则可能会导致吸烟者消费更多的香烟。为性传播疾病开发更好的治疗方法可能会使人们更容易做出不安全的性行为。这样的例子不胜枚举。⁹从事后回顾的角度来看，这些负反馈中有许多似乎都是显而易见的，但是要想提前做出预测，可能就不是一件容易的事情了。写出一个定性的系统动力学模型能够使这些反馈回路变得清晰，从而帮助我们成为更好的思考者。

系统动力学模型会鼓励我们把反馈包含进模型中，这个事实恰恰是这种方法的优势所在。1696年，英国国王威廉三世推出了一项宅基地

税，规定每栋房子的基本税率为两先令，另外还根据窗户的数量收取额外的税收：窗户数量在10扇至20扇的房子要额外缴纳4先令的税，窗户数量超过了20扇的房子则要额外缴纳8先令的税。国王之所以要对窗户征税，是因为它们是可观察的（可以客观地加以测量），并与住房价值相关。如果国王依赖对房子不动产价值的评估，就会出现很多偏袒和贿赂现象。这种对窗口征税的方法是一个好主意。因此在接下来的一个世纪里，这个方法传播到了法国、西班牙和苏格兰。直到1926年，法国才最终废除了窗口税。

作为模型思考者，我们有理由预期有目的会适应的人们会对税收政策做出反应。当然，他们确实以各种各样的方法做出了反应。有些人把房屋上原有的窗户用砖头砌上了。聪明的建筑师很快就提出新的住房结构设计。在这种税收期间建造的许多中产阶级住宅，二楼卧室都没有窗户。在爱丁堡，有一大排房子的卧室完全没有窗户。¹⁰于是税收收入随之下降了。在这里，坎贝尔定律再一次发挥了作用：政客们制定了一个制度，但是人们找到了“绕过它”的方法。更精细的系统动力学模型应该还会考虑到窗口减少带来的影响。它将包括从被称为“窗口”这种存量到其他性质的箭头，例如民众的健康——民众的健康会因缺乏新鲜空气和日照而受到损害。

系统动力学模型的巨大价值部分在于，它们能够帮助我们深入思考自己行动的影响。我们通常都能够考虑到政策的直接影响：对窗口征税能够增加收入；安装防抱死制动器能够挽救生命。但是我们并不一定随时都会考虑到间接影响，也就是各种正反馈和负反馈。这些模型恰恰能帮助我们更清晰、更深入地进行思考。

19 基于阈值的模型

种族、民族以及其他阶层的融合标志着显著的社会不平等，这是民主社会的至关重要的理想。

伊丽莎白·安德森（Elizabeth Anderson）

在本章中，我们将介绍基于阈值的模型。当外部变量超过或低于特定的阈值时，人们的行为所发生的变化，就是基于阈值的行为（变化）。例如，当一个人决定在价格低于100美元时购买一件外套，或者当某项社交活动的参与者人数达到1 000人就加入进去时，就会出现基于阈值的行为。基于阈值的行为很直观，也不难分析，但往往会产生违反直觉的结果，例如当宽容行为反而导致隔离时。基于阈值的模型也可以产生临界点。例如，当一个人加入社交活动的决定取决于已经参与了该项活动的人数时，随着越来越多的人参与该活动，参与者的总人数也超过了其他人的阈值，从而导致了更多的人加入。

本章中讨论的模型也可以归类为基于主体的模型。在计算机领域，基于主体的模型指可以单独为每个主体建模的计算机程序。基于主体的模型比系统动力学模型具有更细的粒度。系统动力学模型用单个存量变量表示整个群体，而基于主体的模型则做到了在空间中定位主体，并且可以包括行为多样性。由此增加的自由度当然可以带来一些优势，但是我们必须记住，过多的细节会与要建模的原因相冲突。例如，在构建一个用来讨论人们如何选择参加某项社交活动的模型时，就不应该对人们大脑中的每个神经元进行建模。最优粒度水平取决于模型的目的。

本章从格兰诺维特（Granovetter）的骚乱模型（riot model）开始讨论，接着引入这个模型的双重骚乱版本，用来对初创企业的成长过程进行建模。然后介绍两个隔离模型：第一个讨论参加聚会的一群人如何在

不同房间之间移动，第二个讨论城市规模上的隔离。接下来，讨论乒乓球模型（ping pong model），它包括了负反馈，可以产生循环或均衡。在本章末尾部分的讨论中，我们将重新回到生命游戏，分析双重阈值规则如何导致了正反馈和负反馈的结合。正反馈会产生相关行为，而那正是负反馈所要抑制的。本章还会进一步讨论模型粒度问题。

格兰诺维特的骚乱模型

在基于阈值的模型中，个体根据某个总量变量是否超过阈值而决定采取两种行动中的哪一种。如果变量的值超过阈值，个体就采取一个行动，否则，就会采取另一个行动。我们要讨论的第一个阈值模型刻画的是骚乱和社会活动。在这个模型中，每个人都可以选择参加骚乱或者不参加，如何选择取决于参与骚乱的人数。模型没有预设任何规范立场，这里所说的骚乱或社会运动，既可以指对暴君的正义反抗，也可以指足球流氓毁坏财物的非法行为，模型对这两类情况都适用。

骚乱模型为每个人分配一个阈值。当参加骚乱的人数超过那个阈值时，这个人就会参加骚乱。¹一开始，只有那些阈值为零的人才会参加骚乱。考虑到本书的研究目的，这里所说的“骚乱”是一般的社会活动而不是暴乱，因此在这种情况下，参加骚乱也无非是指参与聚会。假设第一天，有200个阈值为零的人发动了一场社会活动。第二天，这200人继续参加，于是参与阈值低于200的人也加入了他们的行列。假设第二天新加入的人有500人，那么第三天，阈值低于700的人也会加入。而这可能会涉及好几千人。

骚乱模型

假设有 N 个人（用 i 来索引），每一个人都有一个“骚乱阈值”（riot threshold），不妨记为 $T(i) \in \{0, 1, \dots, N\}$ 。在一开始，骚乱总产值为零，即 $T(i) = 0$ 的那些人都会参加骚乱。令

$R(t)$ 等于在时间 t 参加骚乱的人的数量。在时间 t 上, 如果 $T(i) < R(t-1)$, 那么第 i 个人就会参加骚乱。

骚乱模型的三种情况

对骚乱模型的分析表明, 阈值的多样性至少与平均阈值一样重要。原因可以通过比较如下三种情况来说明(它们都涉及1 000人)。在第一种情况下, 每个人的阈值都为10, 因此不会发生社会运动。在第二种情况下, 有5个人的阈值为零, 10个人的阈值为1, 其他人的阈值均为20。那么在这种情况下, 一开始会有5个人发起社会活动, 第二天会有10个人加入, 但此后就再没有其他加入者了。

在第三种情况下, 这1 000人中每个人都有一个独特的阈值, 范围从0到999。为了方便起见, 我们可以根据这些人的阈值将他们编号为0至999号, 第 i 个人的阈值为 i 。第一天, 第0号人发起一场社会活动, 第二天第1号加入, 第三天第2号加入.....如此类推, 每天都有一个“新”人加入, 直到所有1 000人都参加这场社会活动为止。

不难看出, 第一种情况的平均阈值最低, 但是不会发生社会活动, 因为没有人的阈值为零。在第二种情况下, 虽然有一些人的阈值为零, 但是仍然不足以发动一场广泛参与的社会活动。只有在最后一种情况下, 社会活动才会成功。

这个模型表明, 阈值的总体分布是非常重要的, 而不仅仅只有阈值的均值才重要。因此它也说明, 要预测哪些社会活动将会成功有很大的困难。这个模型还可以指导行动, 它可以告诉那些希望发动社会活动的人们: 要取得成功, 除了组织一个核心革命团体之外, 还需要“发动群众”, 也就是“创建”一个愿意加入他们的群体。

骚乱模型有各种各样的变体, 可以用来分析人们什么时候会起立鼓掌、人们的观点会怎样变化(例如, 会不会接受同性婚姻)、时尚潮流

的改变（例如，愿不愿意佩戴蝶形领结），以及市场动态演变（例如，是否投身于股市泡沫或房地产泡沫）。在所有这些情况下，人们的行为都可以通过基于阈值的规则来近似，而且该阈值因人而异。在所有这些情况下，发生大事件的可能性——无论是大规模社会活动，还是厚边框眼镜的大范围流行，都可能在很大程度上取决于阈值的分布而不是阈值的均值。

市场创造和双重骚乱

骚乱模型还可以扩展到用来分析互联网初创企业。这些初创企业能够创造出新的买家和卖家市场。要想创造一个新的市场，初创企业必须创造一群买家和一群卖家。例如，致力于将狗主人与狗保姆匹配起来的网站，就需要让狗保姆和狗主人到网站上注册，对于提供包裹配送、运输或清洁服务的网站来说也是这样。这些网站要想取得成功，必须先“创建”出这样两个群体，而且必须保证这两个群体以大致相同的速度增长，否则，卖家或买家将无法找到匹配的另一方。换句话说，初创企业必须同时制造出两场“骚乱”。

成功的初创公司爱彼迎（Airbnb）提供了一个关于双重骚乱的例子。爱彼迎很好地将试图出租房子的人与寻找短期住处的人匹配了起来。为此，爱彼迎需要创建两个人群：求租房子的人和出租房子的人。只有当爱彼迎的网站有足够多的可供出租的房子时，想租房子的人才会访问他们的网站。因此，需要让愿意出租房子的人到爱彼迎的网站上注册，并将他们的房子挂牌出租。爱彼迎的前两次发布都失败了，因为将房子在网站上挂牌出租需要房东付出不少时间，需要上传不少图片并填写其他一些信息。在爱彼迎拥有大量租房者之前，没有房东愿意将自己的房源放到他们的网站上挂牌出租。

因此，一方面，爱彼迎需要有足够多的挂牌房源在租房者当中造成一场“骚乱”，也就是让足够多的租房者来到他们的网站求租房子；另一

方面，他们也需要足够的租房者，以便在那些想要出租房子的人当中制造一场“骚乱”。爱彼迎能否成功，取决于这两个人群的阈值。相比之下，难度更大的是怎样让房东将房源放到他们的网站上挂牌出租，因为这需要房东付出更多的努力。爱彼迎克服这个难题的方法是：挨家挨户地上门，帮助房东将他们拥有的房子在爱彼迎的网站挂牌出租。一旦完成了这个任务，租房“骚乱”就开始了，随后又出现了挂牌“骚乱”。²爱彼迎之所以能够成功地创造出这个市场，是因为它的创始人能够培育出足够数量的初始租户，并以此触发双重骚乱。他们先创造出了一条狗尾巴，然后用狗尾巴拉出了一只狗。

谢林派对模型与谢林隔离模型

接下来要讨论的两个模型都是经济学家托马斯·谢林提出的，谢林用它们来分析种族隔离问题。人们在很多个尺度上按种族和民族隔离，先按不同的国家隔离，在国家内部又按地区隔离。在美国，城市内部还会出现社区层面的种族隔离。对于这些，有人说这是种族不宽容的证据。但是，这个结论与越来越多的跨种族、跨民族通婚的证据相矛盾。同一群人，既然决定不与不同种族的人住在同一个社会，甚至不愿意与不同种族的人在一张桌子上吃午餐，但是他们却选择与不同种族的人结婚，怎么会这样呢？

如果跨种族婚姻中的那些人，和与不同种族的人共坐一张桌子上的那些人，分别属于不同的社会阶层，那么我们或许可以解释这些事实，但是事实并非如此。跨种族婚姻在所有收入水平上都有发生，同时在美国最顶尖的大学和学院里也可以发现“种族隔离”的午餐桌。幸运的是，谢林的模型可以同时解释这两组事实，它们说明了宽容的行为也会导致种族隔离。

谢林派对模型

第一个是谢林派对模型（Schelling's party model）。在一定意义上，可以认为这个模型是随机游走和骚乱模型的“混搭”。谢林派对模型描述了一个在一幢有两个房间的房子举行的派对。这个派对邀请的客人可以分为两个类型，这里的类型可以是男性和女性、黑人和白人、西班牙人和澳大利亚人、哥特人和印第安人……但这都不会影响结果。关键的假设是，每个人都能够区分其他人所属的类型。

谢林派对模型

N 个人参加一个派对，每个人都有一个可观察的类型A或B。每个人随机选择两个房间中的一个。在每个时刻，每个人都有 p 的概率走到另一个房间去。第 i 个人的宽容阈值为 T_i 。对于这个人，如果他所属的类型的人在当前房间内所有人当中所占的比例低于这个阈值，他就离开这个房间。

为了理解隔离是如何在这种存在“宽容”的情况下产生的，假设前来参加派对的是20名澳大利亚人（A）和20名巴西人（B）。每个人都是宽容的，只要同一个房间里有25%的人与自己属于相同的种族，他们就会留在这个房间里。再假设一开始，有一个房间里待了12名澳大利亚人和9名巴西人，另一个房间则有8名澳大利亚人和11名巴西人。这样没有人觉得自己必须走到另一个房间去，但是两个房间之间会有随机移动，这种随机移动会改变每个房间内的“种族结构”。在图19-1中，一个房间里待了11名澳大利亚人和4名巴西人。这个配置处于一个临界点上：只要任何一名巴西人离开，巴西人所占的“人口比例”就会降低到25%以下，从而导致其余3名巴西人也选择离开。如果发生了这种情况，就永远也不会有巴西人再进入那个房间。

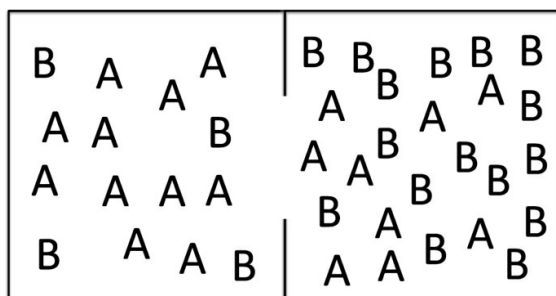


图19-1 一个很容易导致隔离的配置

回忆一下随机游走，我们说过随机游走能够越过任何阈值。一个房间里的澳大利亚人的数量是一维随机游走。因此，如果派对持续时间足够长，那么隔离就是不可避免的。当然，即便是最好的聚会也不一定能持续得“足够久”，所以并不是所有派对都会以隔离而告终。我们也知道，派对规模越大，隔离结果越不可能出现，因为当参加派对的人更多时，随机游走就必须超过一个更高的阈值。例如，在一个有1 000人参加的派对中，不太可能出现某个房间里某个特定类型的人所占比例低于25%的情况。因此，我们预测，在小型派对中出现隔离的可能性更大。

而且，当参加派对的人有不同的宽容阈值时，我们也应该会观察到更多的隔离现象。要理解个中缘由，不妨假设参加派对的是10名巴西人和10名澳大利亚人，并为每个人分配5%至45%的宽容阈值，使每种类型的人的平均宽容阈值为25%，如图19-2所示。

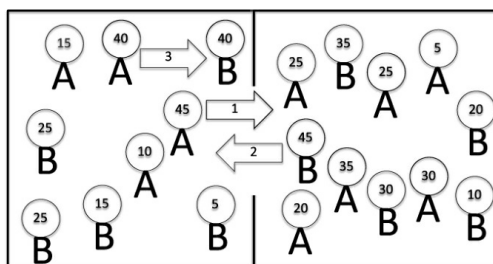


图19-2 异质性阈值导致的“重新安置”

在图19-2中，左边的房间里有5名巴西人和4名澳大利亚人，因此澳大利亚人的比例低于45%，这导致最不宽容的那名澳大利亚人离开这个房间，进入右边的房间（用箭头1表示）。这名澳大利亚人进入了右边

那个房间后，就降低了巴西人在那个房间里的比例，导致最不宽容的巴西人换房间（用箭头2表示）。这两次移动使左边房间里澳大利亚人所占的百分比降低到了低于40%的水平，从而导致第二名最不宽容的澳大利亚人也离开左边的房间，进入右边的房间（用箭头3表示）。就这样，一个级联式反应出现了。然而，如图19-3所示，结果不一定是完全隔离，因为最宽容的那些人在任何一个房间里都会觉得舒服自在。在这个模型中，多样化阈值产生了两种不同的影响：它们使隔离更容易发生，同时又使完全隔离的可能性降低，因为非常宽容的人在任何一个房间都不会觉得有问题。

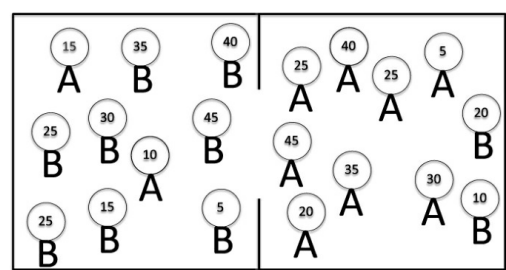


图19-3 异质性阈值导致的“隔离”

这个模型可以用来解释不同职业之间的性别比例差异，例如为什么会有更多的女性担任护士，更多的男性担任推销员。这种差异也许是由于偏好不同所致，但是只要对这个谢林派对模型稍稍加以扩展，就会发现，如果人们更愿意与相同性别的人一起工作，那么这种不同职业之间的性别比例差异就可能会出现。对于这一点，我们还可以用旋转门模型（revolving-door model）来给出更加正式的解释。旋转门模型有两个基于经验基础的假设：退出某个职业的女性会选择女性更多的新职业；女性退出某个职业的速度略高于男性。³例如在生命科学领域中，女性退出生物医学研究的速度比男性更快一些，并且会在雇用了更多女性的医疗保健等职业中谋职，这种行动就会提高这两个职业的性别隔离程度。

谢林隔离模型

谢林隔离模型（Shelling's Segregation Model）将行为主体放在地理空间的不同位置上，就像是放在棋盘的不同格子上（图19-4）。其他方面则都与谢林派对模型相同。这个模型同样假设有两种类型的人，对他们的行为假设也与谢林派对模型相同。

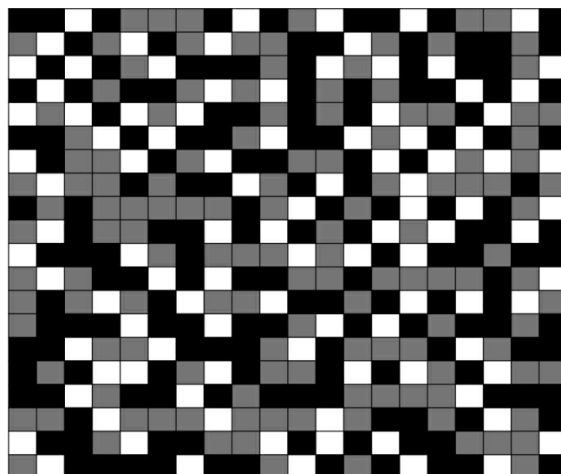


图19-4 谢林隔离模型的初始配置

谢林隔离模型

有 N 个人，每个人都属于类型A或类型B，随机排列在一个 $M \times M$ 的棋盘上。棋盘上还有部分空间未被占用。第 i 个人的宽容阈值为 T_i 。对于这个人，如果他所在方格的8个相邻方格中同一类型的人所占的百分比低于他的阈值，就会重新定位到一个随机的新方格上。

如果个体的平均阈值接近50%，这个模型会产生地域性的隔离，如图19-5所示。出现这种隔离是因为个体只考虑自己所在的“社区”，这种社区最多只能有8个居住者。几乎任何随机的初始配置都包括了被相反类型的其他人包围的一些人。如果个体进入了具有更多相同类型的个体区域，就可能引起另一种类型的人的重新安置。随着重新安置的不断积累，最终发生了隔离。毫无疑问，阈值多样性会加剧这些影响，原因如

前所述，这里不再重复。

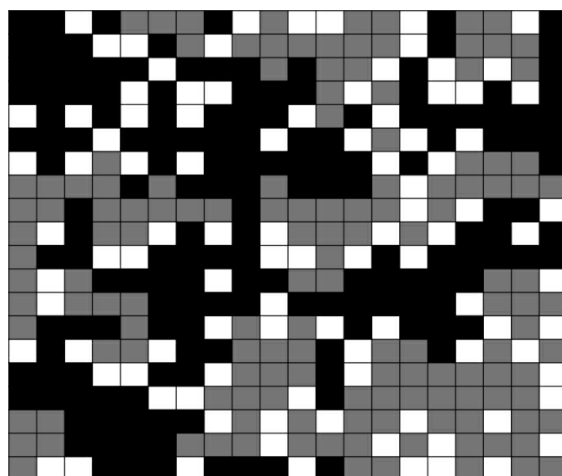


图19-5 谢林隔离模型中“重新安置”后的配置

即使是宽容的人，也会产生隔离的居住模式，这就是托马斯·谢林的开创性名著《微观动机与宏观行为》一书最根本的洞见。

乒乓球模型

现在讨论乒乓球模型，它假设基于阈值的行为会产生负反馈。负反馈可以起到稳定系统的作用，还可以产生循环。乒乓球模型假设了有限数量的实体，它们可以是人，也可以是机械的、生物的或化学的装置。在每个时期，每个实体采取一个正（+1）的行动或一个负（-1）的行动。在第一个时期，每个主体随机选择一种行动，初始状态等于行动的总和。系统所有的后续状态，都等于前一个时期的状态加上所有动作的平均值和一个随机项。每个实体都有一个阈值（从特定的分布中抽取），如果状态绝对值超过了阈值，就选择能够减少状态绝对值的操作。简而言之，如果状态超过了实体的阈值（无论是负的，还是正的），这个实体就尽其所能减少状态的大小。

乒乓球模型

规模为 N 的种群中的每一个实体随机采取一个初始行动，或者为正（+1），或者为负（-1）。系统的初始状态 S_0 设置为零。系统的所有未来状态 S_t 则等于平均行动再加上一个随机变量，即：

$$S_t = \frac{\sum_{i=1}^N A_i(t)}{N} + S_{t-1} + \varepsilon_t$$

每个实体 i 都有一个响应阈值（response threshold） $T_i > 0$ ，它从 $[0, \text{RANGE}]$ 中均匀抽取出来。如果状态的绝对值大小 $|S_t|$ 小于它的阈值，该实体就采取与以前相同的行动，否则，就采取减少状态绝对值大小的行动。即：

如果 $|S_t| \leq T_i$ ，那么 $A_i(t+1) = A_i(t)$ ，否则 $A_i(t+1) = -\text{sign}S_i(t)$

其中， ε_t 是从 $\{-1, +1\}$ 中随机抽取出来的。

乒乓球模型有多种应用。人们可能会将时间和资源分配给多个慈善项目。如果某个慈善项目得到了太多的关注或金钱，人们就会开始向其他项目捐钱以保证平衡。例如，一个国家需要维护两个联合国教科文组织世界遗产，人们会在两个遗产上都投入一些时间和资源，但是如果某个遗产受到的破坏更严重，人们就可能会重新分配资源。

正如“乒乓”这个名称所预示的那样，乒乓球模型可以生成围绕着均衡的循环行为。在一个时期内，太多的人选择了这个行为，而在接下来的那个时期，却有太多的人选择了另一个行为。当所有实体的阈值都为零时，在这一个时期内所有实体都会选择这个行为（+1），而在下一个时期内所有实体都会选择另一个行为（-1）。

为了探索阈值多样性如何影响人们的乒乓球式行为模式，还是会找到一个均衡，我们考虑如下三种情况（每一种情况都有100个人参与）。在第一种情况下，假设阈值均匀分布在0到10之间。如果第一个

时期中的状态等于6，那么它将超过大约60个人的阈值。这60人中大约有30人将采取行动1，并将改变行动。这些行动总和现在将超过50，因此系统的新状态（前两个时期的平均值）将超过20。这个值超过了所有阈值，因而产生了图19-6所示的乒乓球效应。

在第二种情况下，如果我们增大阈值的多样性，使它们在0到100之间均匀分布，那么乒乓球效应几乎消失。如果再次假设状态在第一个时期等于6，那么平均来说只有6个人的阈值将得到满足。如果这6个人中有3个改变行动，那么状态将趋向于零。这种对偏差的抑制可以从图19-6中看得很清楚（它对应于0到100之间的阈值）。

在第三种情况下，考虑一个适中的范围，让阈值均匀分布在0到60之间，我们将会观察到更温和的周期，如图19-6所示。因此，在具有负反馈的系统中，阈值多样性会导致稳定性，但是在具有正反馈的模型中则具有恰好相反的效果。

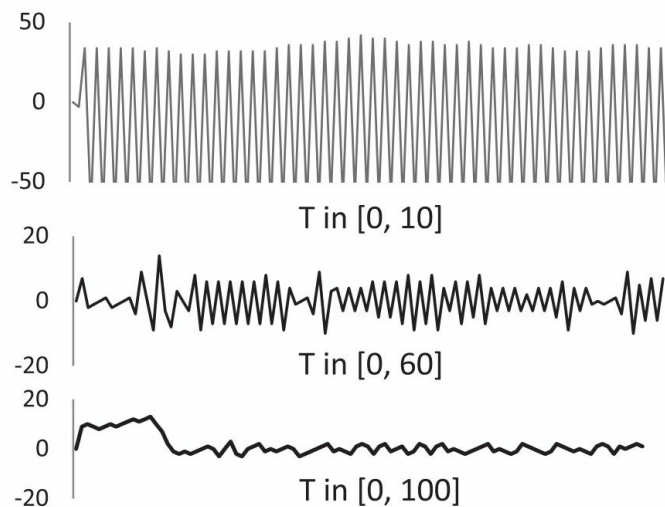


图19-6 三种阈值范围下乒乓球模型的时间序列

模型粒度

反馈的基本逻辑很简单：正反馈强化行动，负反馈抑制行动。只有

正反馈的系统会爆炸或崩溃，只有负反馈的系统将稳定或循环，同时具有正反馈和负反馈的系统则可能会产生复杂性。

在系统动力学模型中，反馈作用于存量变量（例如，野兔的数量）和速度（例如，增加面包店的购买率）。基于主体的模型（例如本章中介绍的阈值模型）会产生作为个体行动的结果反馈。这类更精细的模型可能包含多个阈值，阈值多样性会加剧正反馈的影响并缓和负反馈的影响。分布尾部的多样性使骚扰更加容易发生，宽容阈值的多样性则促进了隔离。具有负反馈系统中响应的多样性阻止了同质性反馈可能导致的大幅波动。在企业之间的经济竞争模型中，生产成本的异质性也可以起到类似的作用。随着价格的上涨或下跌，由于各自成本的差异，企业做出的反应会各不相同。⁴

这两种模型之间的差异引发了模型粒度的问题。是创建一个以“野兔”为名的变量（或方框）并描述野兔的数量如何随着其他变量而增加或减少？还是对每一只野兔单独建模？细分变量可以提高描述的准确性，但是模型不是以描述准确性为标准来评分的。请一定要记住统计学大师乔治·博克斯的格言：所有模型都是错的。至于博尔赫斯的地图，虽然与现实世界一样大，却全无用处。许多建模者，包括爱因斯坦，都认为，我们应该寻求适当的粒度水平。例如，如果要构建一个模型来解释人类的手臂能够施加的力，就不需要考虑DNA。

而在研究社会系统的时候，则可能并不存在理想粒度水平。我们可能需要在多个粒度水平上分别进行探索。通过构建多个模型，每个模型都有不同的粒度水平，然后进行跨模型对话。假设我们试图了解瑞典与芬兰之间的贸易模式，可以这样入手：把两个国家的贸易总量作为变量，并确定宽泛的宏观模式。然后，将每个国家的进出口总额按行业分解，并进一步分解为各个行业内的企业。行业级数据将使我们能够更好地解释过去的模式，并对未来的模式做出更加精确的预测。对各个行业内的企业做进一步深入了解，包括成本结构和增长轨迹等信息，可以得

到更好的结果，但是我们需要大量的信息才有可能构建出一个包含这么不断变动着的组成部分的、真正有用的模型。我们甚至还可以对这些企业的领导力进行建模。最后一个层次的拆解式研究，很可能不会带来多少有用的结果，但可能会使一些致力于扩张海外的企业领袖名扬四海。

总而言之，更精细的粒度不一定更好，有的模型可能过于注重细节。我们能够理解更细粒度的模型，也可以从相对粗糙的模型中获益。通过比较不同模型的预测、解释之间的差异，我们可以看清楚假设如何影响结果，也可以看清楚假设的条件性。许多模型不仅在各自包含的变量方面存在差异，而且在何种粒度上分解变量也理所当然地有所不同。

算法骚乱

骚乱模型和乒乓球模型都可以用来解释股市崩盘及随后的价格反弹。我们在此考虑两个案例，它们很能说明一些问题。

第一个案例发生在1987年10月19日星期一，那一天，道琼斯工业平均指数下跌了超过22%。第二天，这场股价崩溃在全球范围引发了连锁反应。这次崩溃的原因直到今天仍然是许多研究的主题。当时，美国经济扩张已经进入了第四个年头。那一年的前8个月，道琼斯工业平均指数的涨幅超过了40%。也许就是因为股票市值的这种上涨，使许多人认为股票价格被高估了。在崩溃前的那个星期天，美国时任财政部部长詹姆斯·贝克（James Baker）威胁说，如果德国再不降低关税，美国就要采取削弱美元政策。这种说法在当天似乎没有引起什么反响。但是第二天，股票市场就崩盘了。14个月后，股票市场又恢复到了崩盘之前的价位。

为了应用本章给出的这些模型，我们可以用一种金融资产来代表整个市场。假设，持有该资产的每个人都有一个“崩溃阈值”（crash threshold）。如果资产价格在某一天的下跌幅度超过了崩溃阈值，投资者就会出售资产，将资金从市场中撤走。这个规则

可以捕获趋势交易者或噪声交易者的行为。在此基础上，我们可以构建一个骚乱模型。如果某些投资者在10月19日一早醒来就决定出售大量资产，他们的行为就会导致股票价格下跌。如果股票价格的下跌幅度超出了其他投资者的崩溃阈值，那么其他投资者也会出售，从而导致股票价格呈螺旋式下降。于是在这里就有了一个经典的正反馈循环：抛售导致价格下降，价格下降导致更多抛售。

现在加入乒乓球模型。如果价格下跌得太多，人们会应用第二条规则，换句话说，人们还有一个“便宜货阈值”（bargain threshold），如果价格低于这个阈值，人们就会买入这种金融资产。在这里，投资者是基于价值而非趋势行事。当价格急剧下跌时，便宜货阈值会产生负反馈。买家急于以便宜的价格买入，导致价格停止下跌。

当然，现实世界中的市场要比这个模型（卖家跌破阈值就卖，买家等着机会入货）复杂得多。股票市场包含了许多类型的交易者，包括大型机构、养老基金、各国政府、投机者、投资组合保险公司，以及散户投资者。⁵正因为存在这种多样性，在价格下跌时几乎永远有人愿意买进，从而产生了稳定市场所需的负反馈。

对崩溃的早期分析强调了（计算机）程序交易盛行的影响，这些其实是用计算机程序编码并实施的基于阈值的规则。如果股票市场指数低于某个预先设定的值，就卖出所有股票。诸如此类的规则都是在无人监督的情况下自动执行的。大多数分析师现在认为，程序交易是促成1987年股市崩盘的一个原因，但并不是主要原因。

对1987年股市崩盘更细致的分析表明，大量投资组合保险公司作为必须保证投资者投资组合回报率的交易者，产生了强烈的正反馈，而且这些正反馈并未受到负反馈的抵消。在股票市场大幅下跌的情况下，这些投资组合保险公司会通过抛售股票来止损。随着股票的进一步下跌，这些投资组合保险公司抛售的股票越来越多。实

际上，这些投资组合保险公司的行为就好像他们是具有不同阈值的个人组成的群体。有一家投资组合保险公司抛售了超过10亿美元的股票，而10月19日当天全天的成交量只有200亿美元。

第二次案例是一个发生在2010年5月6日下午的闪电崩盘。在三分钟内，道琼斯工业平均指数就下跌了5%。这个闪电崩盘是算法交易的结果。由于现代金融市场的高度复杂性和交易的极高速度，没有人确切知道在闪电崩盘期间究竟发生了什么。但是我们确实知道，有一个庞大的共同基金发出了一个超级大卖单，将超过40亿美元的股票期货倾销到了市场上，而市场上则充斥着各种各样试图利用有利交易的高频交易算法。这种高频算法能够感知价格趋势并开始以极快的速度执行交易，对此，我们可以想象一个有极高速度的骚乱模型。

这就产生了一个“有毒”的市场，交易者担心大型机构投资者知道他们不知道的东西，因此他们退出了市场。⁶鉴于市场行为异常，许多算法停止了交易；但是其他算法则继续销售，然后就发生了闪电崩盘，所有这一切都在几分钟内完成。20分钟后，便宜货阈值规则开始发挥作用，正如乒乓球模型所预测的那样，价格回升到了（接近于）原先的水平上。

20 空间竞争模型与享受竞争模型

我们的理论是，如果你需要用户来告诉你，你正在销售的是什么，那么你其实不知道你正在销售的是什么，而且这不可能给客户留下好的体验。

玛丽莎·梅耶尔（Marissa Mayer）

在本章中，我们将介绍关于个人选择的模型，人们选择时所针对的备选项由其属性来代表。开发这些模型主要是为了刻画消费者的选择行为。买房子的人要考虑它的面积、卧室和浴室的数量以及建筑质量。这些模型也适用于招生的考官、选聘董事长的董事会，以及参加投票的选民，他们都要在申请人或候选人之间做出选择。招生考官要考虑申请人的SAT成绩、平均成绩和课外活动。选民要根据候选人在教育、基础设施、犯罪和税收方面的立场进行评估。除了帮助我们理解个人选择之外，这些模型还可以解释为什么会这样做出选择，例如，为什么我们有这么多种早餐谷物可供选择。

讨论这些模型时，我们会将某些属性描述为空间属性，而将其他属性描述为享受属性。空间属性，例如夹克的颜色或一片面包的厚度，没有最优值。每个人都喜欢特定“数量”的这类属性：一个购买猪排的消费者有自己喜欢的辣度，一个业余滑雪运动员在滑下斜坡时有自己喜欢的下降角度。产品的属性越接近理想水平，消费者对产品的评价就越高，而且理想水平因人而异：一个人可能比另一个人更喜欢辣一些的猪排。

与此不同，在享受属性上，更多（或者在某些情况下更少）总是意味着更好。人们喜欢智能手机的待机时间更长一些，房间的面积更大一些，皮鞋的鞋底更耐磨一些，自己的汽车更省油一些。不过，在现实世界中，大多数选择都是“混合型”的：人们既会考虑空间属性，也会考虑

享受属性。

在本章中，我们假设人们总是会选择他们估价最高的备选项（备选方案）。之所以要这样做是出于第4章中已经提到过的一些原因：在对人类行为进行建模的时候，理性行为提供了一个基准，基于理性假设构建的模型不仅容易处理，能够给出符合唯一性要求的预测，并且这种预测在事情重复发生且利害关系很大时与经验相符。

空间（属性）竞争模型（**spatial competition model**）和享受（属性）竞争模型（**hedonic competition model**）在经济学和政治学中早就得到了广泛应用，部分原因在于它们便于用数据检验。¹在本章中，我们可以体会到这种适用性。我们先从空间竞争模型入手分析。接着将这些模型应用于政治领域，并阐明如何使用这种模型来分析现状效应、议程权力和拥有否决权的人的影响力。然后，我们介绍享受竞争模型及其混合模型，以揭示对价格竞争的若干结论。在这个过程中，我们还将说明，如何在模型中应用数据，根据候选人和法官在相关法案和法律案件上的投票历史来推断出他们的立场。我们还介绍了如何推断出清洁空气或更少的通勤时间等没有明确价格属性的隐含价格。²

空间竞争模型

空间竞争模型假设备选项可以用一组属性来定义，消费者则可以用一系列理想点来定义。最简单的空间竞争模型考虑的是只有单一属性的产品。在由数学家、统计学者哈罗德·霍特林（**Harold Hotelling**）提出的最早的空间竞争模型中，这个属性就是地理位置。³

霍特林的模型假设一些消费者分散地分布在海滩上（如图20-1中的圆圈所示），海滩上有两个卖冰激凌的小贩（分别用A和B表示）。每个消费者都会从离自己更近的那个小贩手中购买一个冰激凌。图中的分割点位于这两个小贩之间的中点位置，它决定了哪些消费者会从哪个小

贩手中购买冰激凌。分割点左侧的那6个消费者会从A小贩那里购买，而分割点右侧的那7个消费者则会从B小贩手中购买。

考虑到消费者更喜欢离自己更近的商品，我们可以对距离给出一种更加抽象的解释。例如，我们可以想象这两个冰激凌小贩位于同一个地点，但是分别出售乳脂含量不同的冰激凌。因此，同样的数字可以表示B小贩出售冰激凌的乳脂比A小贩出售的冰激凌更加浓稠。根据这种解释，消费者的位置就不再是他们在海滩上的物理位置了，而是他们所喜欢的乳脂含量。

我们还可以继续应用一对多的模型思维，用这个模型去分析政治竞争。唐斯模型（Downsian model）将霍特林模型中的地理空间重新定义为从左到右的意识形态连续系统。于是，我们可以对图20-1重新解释如下：A小贩代表自由派的政治候选人，B小贩代表保守派的政治候选人，圆圈则代表选民的意识形态理想点。然后进一步假设，选民更喜欢政治立场与自己更接近的候选人。

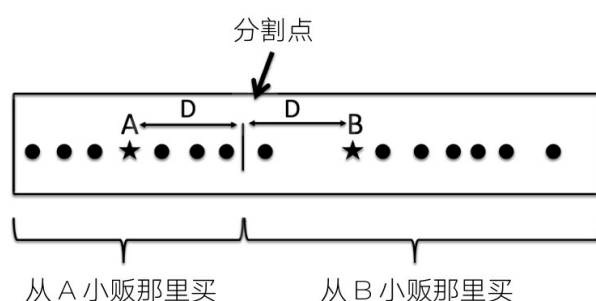


图20-1 海滩卖冰激凌的地理模型

从企业的地理位置和产品属性，转换为政治意识形态，意味着从诸如位置和乳脂含量等物理属性，转换为更抽象的思想观念。虽然我们对于物理属性已经有了明确的测量方法，但是要分配意识形态立场，则需要找到一种新的方法将候选人的行为转化为数字。如果候选人有投票记录，我们只要收集到候选人以往的选票数据，就可以对他们分配意识形态立场。

在这个过程中，我们应该忽视所有缺乏意识形态成分的选票，例如，当所有有权投票的人都一致同意设立全国牛奶日时，这样的投票数据就无须考虑。而对于所有其他投票，则可以依靠专家意见来确定自由派立场和保守派立场。候选人在意识形态区间上的空间位置可以设定为投票给保守派的时间百分比。⁴永远采取保守派立场的候选人将被放置在最右边，一半时间投保守派的票、一半时间投自由派的票的候选人则位于最中间的位置上。

构建了这样一个模型之后，我们就可以通过实证检验来判定美国各政党的意识形态立场的差异是不是变得更大了。如图20-2所示，有一项研究表明，美国民主、共和两党的平均理想点出现了显著且不断加剧的极化趋势。虽然我们不能就此得出美国政党极化程度增加的结论，但是它确实提供了支持这种结论的证据。分析还表明，这种极化主要应归因于共和党的向右倾斜——变得更加保守。

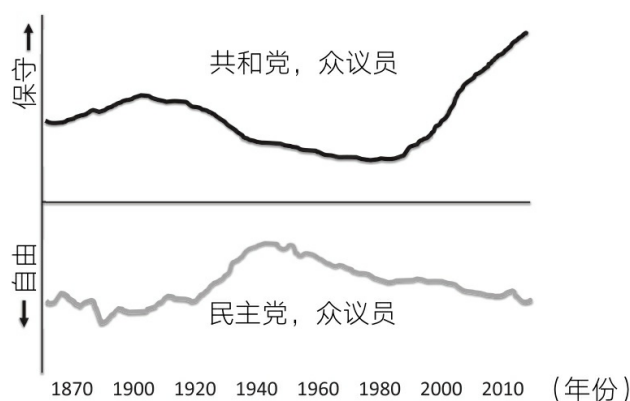


图20-2 美国国会众议院意识形态极化的加剧

增加属性的数量

一般的空间竞争模型可以包括任意数量的属性。比如，沙发可以通过多个物理维度来描述：长度、宽度和深度，结构类型，以及室内装饰类型。消费者从产品中获得的价值（或效用）取决于产品在各个维度上与理想点的距离。我们可以将这种价值函数写为常数项减去备选项与消

费者理想点之间的距离。⁵

空间竞争模型

备选项包括 N 个空间属性： $\vec{a} = (a_1, a_2, \dots, a_N)$ 。

做出选择的个体的理想点： $\vec{x} = (x_1, x_2, \dots, x_N)$ 。

个体选择了备选项后得到的收益（效用）等于：

$\pi(\vec{x}, \vec{a}) = C - (x_1 - a_1)^2 - (x_2 - a_2)^2 - \dots - (x_N - a_N)^2$ ，其中， $C > 0$ 是一个常数。

示例： $\vec{x} = (3, 4, 6)$ ， $\vec{a} = (2, 1, 8)$ ， $C = 20$

$$\pi(\vec{x}, \vec{a}) = 20 - (3 - 2)^2 - (4 - 1)^2 - (6 - 8)^2 = 6$$

在这种一般的空间竞争模型中，对于两种巧克力棒，可以用可可含量的百分比以及含糖量来表示，如图20-3所示。分割线是连接这两个产品的直线的垂直平分线。理想点位于分割线左侧的消费者更喜欢A，而理想点位于分割线右侧的消费者则更喜欢B。⁶

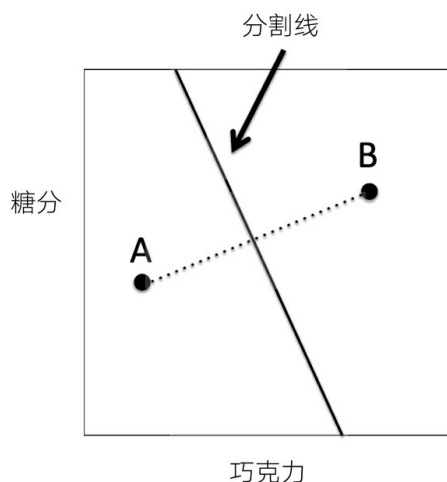


图20-3 两种巧克力棒的产品属性与分割线

这个模型还可以容纳任意数量的产品。要加入第三种巧克力棒，我

们只需在属性空间中增加另一个点即可。为了确定消费者会选择购买哪种巧克力棒，可以再绘制一条分割线，如图20-4所示。这两条分割线可以将理想点的空间划分为三个区域，它们被称为沃罗诺伊邻域（Voronoi neighborhood），领域根据它们与产品的距离划分理想点空间。

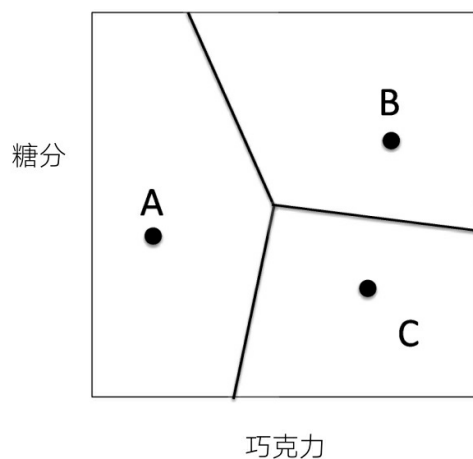


图20-4 有三种产品的空间模型（沃罗诺伊邻域）

唐斯空间竞争模型

接下来，我们应用空间竞争模型来分析政治候选人的意识形态立场“定位”。为了便于讨论，假设候选人都是追求投票最大化，也就是说，他们的首要目标是赢得选举。我们可以从一个简单的例子开始考虑，思考候选人的动机。图20-5显示有13名选民和两名候选人的两种不同情况。回想一下，前面我们讲过，选民更喜欢立场与自己更接近的候选人。在图20-5的上图中，L表示的自由派候选人获得了5票，R表示的保守派候选人获得了8票。在图20-5的下图中，自由派候选人向中间位置移动后，获得了7票，并赢得了选举。

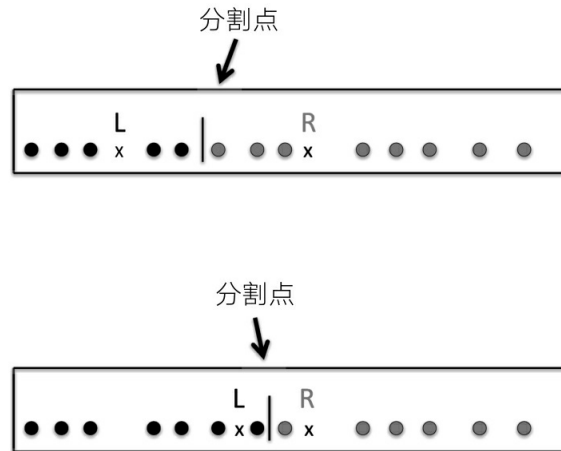


图20-5 向中间位置移动的自由派候选人赢得了选举

自由派候选人有动机向中间位置移动。同样的逻辑，保守派候选人也有这样的动机。也就是说，保守派候选人可以向左侧移动，同时保持在比自由派候选人更靠右的位置上，从而赢得选举。继续按这个逻辑推理下去，自由派候选人也可以选择更接近中间选民理想点的立场。如果继续应用这个逻辑，我们会看到两个候选人应该会趋同到同一点上。这个结果就是通常所称的中位投票者定理（median voter theorem），或称“中间选民定理”。

中间选民定理又可以分别解释强的形势和弱的形势。在强的形势下，中间选民定理意味着各候选人都来到了中间选民的理想点上，采取了相同的立场，但是这显然不符合经验事实。在弱的形势下，中间选民定理意味着候选人有动机采取较温和的立场，经验证据表明这种情况确实会发生。在选举过程中，候选人会趋向中间立场。当然，这种变化不会演变为一场“夺命狂奔”。意识形态信念坚定或受益于位于意识形态光谱极端位置的核心支持者的候选人，对重新定位会非常谨慎。

这个模型将每个候选人和每个选民都归结为意识形态光谱上的一个简单的点，这是一个强有力的假设。捷克作家和政治家瓦茨拉夫·哈维尔（Václav Havel）对这种基于单维意识形态的模型预测非常不以为然，他认为：

对于永恒徘徊的人类来说，它（意识形态）提供了一个立即可以安身的家园，所有人必须做的就是马上接受它。突然之间，一切都变得清晰了，生活呈现出了新的意义。所有的奥秘，所有悬而未解的问题，所有的焦虑和孤独感，全都消失了。当然，人们必定要为这个“廉租房”付出沉重的代价。这代价就是放弃自己的理性、良心和责任，因为这种意识形态最重要的一个方面就是向更高权威拱手献出自己的理性和良心。⁷

哈维尔没有说错。我们不应该为了意识形态而放弃理性。模型为我们提供了更好的推理工具，这种特殊的模型有助于我们理解为什么政治家会做出那种行为。利用数据，我们可以确定每个政治家在意识形态的左右区间上的位置的可信度。一个总是采取温和立场的政治家，可以充满信心地站在这个区间中间。

顺便说一句。哈维尔否认人可以简化为一点，即可以用数据进行检验的程度。但是，即便他的批评成立，即便不能根据投票记录来确定一个候选人的意识形态立场，我们也并不一定要放弃这个模型。我们可以通过赋予他一个区间（而不是一个点）来代表哈维尔所说的意识形态的不确定性。或者，可以为候选人构建一个意识形态时间序列，看他是否保持前后一致。一项对最高法院大法官意识形态立场的研究表明，对有些法官来说，担任法官的时间长了，立场会变得更加自由。⁸

我们还可以增加模型的维度。一个二维模型可以区分社会政策和财政政策。我们可以用这样的模型刻画在社会政策上采取自由主义立场，同时在财政政策上采取保守主义立场的政客。对于美国国会来说，一个维度的模型可以解释大约83%的选票变化，加入第二个维度后，解释力只能增加4%。⁹

除了能够更准确地预测行为之外，增加维度也可能会改变我们的理论结论。我们可以从一个二维模型入手来探讨这个问题。根据对一维模

型的分析，我们知道如果第一个候选人在两个议题的立场都不位于中位数上，那么另一个候选人可以通过在第一个议题上采取与第一个候选人完全一样的立场（从而使该议题变得无关紧要），并在第二个议题上采取中位数立场来赢得选举。同样，如果第一个候选人在第一个问题上取中位数立场但是在第二个问题上却没有，那么另一个候选人可以在两个议题上都取中位数立场来赢得选举。由此可以得出这样一个结论，唯一无法被打败的位置是二维中位数。但是，图20-6却表明，图中用圆圈表示的二维中位数立场是可以被击败的。如果用方形表示的另一个候选人在第一个问题上取偏左的立场，在第二个问题上取偏右的立场，那么就会产生一条分割线，在这条分割线上，他以三票的优势胜出。

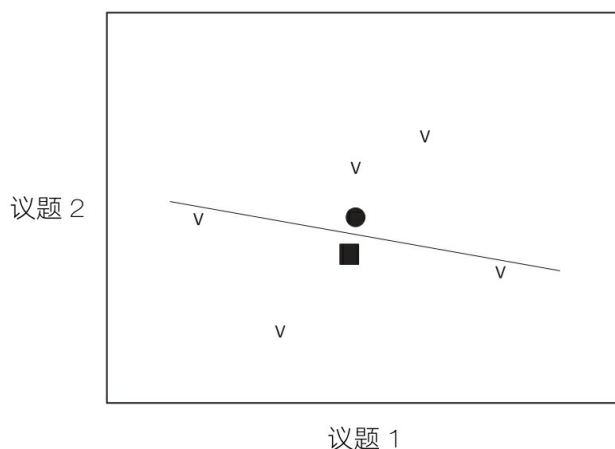


图20-6 二维情况下，中位数候选人会输给挑战者

从这个例子中，我们还可以直观地看出，当且仅当选民理想点的排列，在从中位数出发的每个方向上都不到一半时，二维中位数才可能是不可击败的，这种情况称为径向对称（radial symmetry）。如果选民的理想点在圆形或正方形上均匀分布时，就满足径向对称，这无疑是一个非常强大的假设。任何位置都是有可能被击败的这个结果，被称为普洛特无赢家结果（Plott's no-winner result），它在二维或多维的情况下都成立。[10](#)

结果的差异非常明显。一维模型意味着候选人要位于中位数立场，

高维模型则意味着他们不应该这么做。我们应该相信哪种模型呢？正确的做法是不应该完全相信任何一种模型，而是从这两种模型中都汲取一些洞见。一维模型揭示了追求选票最大的候选人有趋向温和政策的强大动机，高维模型则证明了这些动机的局限性。没有哪个位置可以保证获胜，所以不应该预期会出现均衡。我们应该期待的是复杂性，观察到各路政客时而联盟，时而拆台，纵横捭阖，翻手为云，无休止地为选票竞争。¹¹

我们也可以应用唐斯模型解释美国国会及其各委员会等立法机关和总统制度通过的各种法案的意识形态维度。这里关键是将各个法案映射到与委员会成员相同的单一意识形态维度上。在这里，我们主要考虑三个策略性因素：当前政策的影响（现状效应）、控制议程的权力，以及加入拥有否决权的立法成员的影响。

我们通过一个例子来说明问题。一个由三个人组成的委员会，三位委员的理想点分别为40、60和80，政策提案由理想点为40的委员提出，必须得到多数人同意才能通过。图20-7表明了现状对最终政策的影响。如果现状是80，那么这个提案人就需要中间投票人，也就是理想点为60的那个委员认为该项提案不比现状差。在这种情况下，中间投票人将接受提案人提出的理想点为40的提案，因为它与现状一样好。¹²如果我们将现状移至70，那么这个中间投票人将会拒绝理想点为40的提案，提案者必须提供理想点为50的政策。最后，如果现状位于中间选民的理想点60上，提案者就没有任何权力。因此，我们可以得出以下推论：当现状具有极端值时，提案者拥有最大的权力。

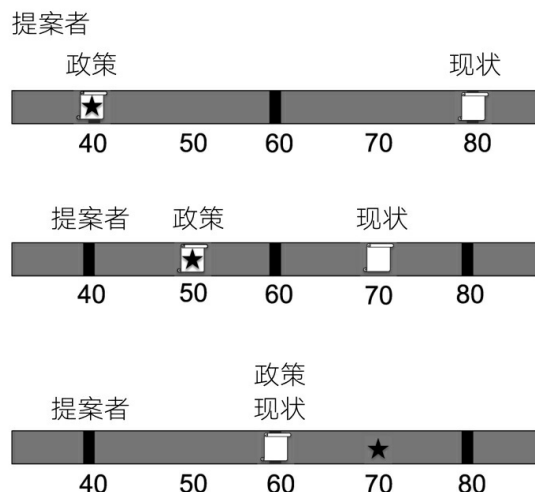


图20-7 现状对最终政策的影响

这个结果适用于需要人们投票做出决策且意见可以映射到一个维度的任何情况下。假设，一个非营利组织的负责人向董事会提出了一个增加用于资助穷人租用经济适用住房的支出建议，如果当前的支出水平与董事会里面的中位数成员的意愿相符，这个负责人的议程权力就非常有限。如果现行政策与董事会成员的理想点不一致，那么他就可以拥有相当可观的权力。

为了说明提出方案的人所拥有的权力，考虑现状为70的情况（如图20-8所示）。在图20-8中，最上面的图显示的是前面说的那种情况，其中提议者的理想点为40，并提出了一个理想点为50的政策。中间的图显示了中位数投票者能够提议自己的理想点60，并且能够获得理想点位于40的委员会成员的支持票的情况。最下面的图则显示了提议者理想点为80的情况。此时，他不能提出他自己和中间选民都喜欢（甚于现状）的政策，因此只能接受现状。

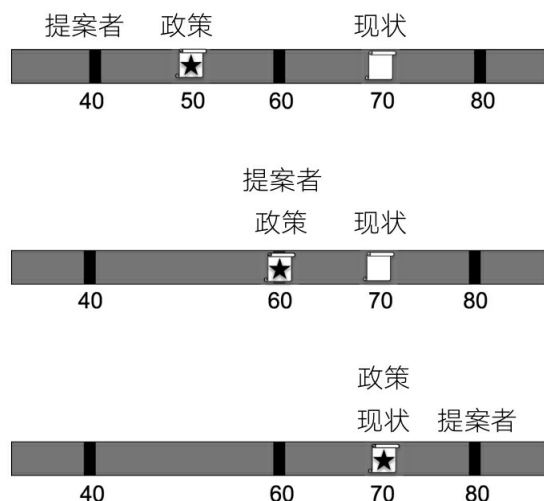


图20-8 提案者对最终政策的影响

这个假想的例子说明了提案者权力的限度。法案可能会朝着拥有权力的人的理想点的方向靠拢，但是正如在上面已经看到的，对于这种权力会不会过大的担忧，只要考虑一下现状的代表性就可以得到缓解。¹³

最后，我们还可以使用这个模型来讨论多层级政府和更多人拥有否决权的情况。在这里，我们将委员会成员解释为众议院、参议院和总统，并假设要想通过一项法案，必须得到全部这三个投票者的批准。在这种情况下，每一个投票者都拥有否决权。再次参见图20-8中现状为70的情形。如果所有这三个投票者都必须批准才能做出任何改变，那么任何提案都无法击败现状。任何位于70左侧的政策都会被位于80处的投票者否决。另一方面，任何位于80右侧的政策都会被其他两个投票者否决。¹⁴如果所有这三个投票者都可以否决立法，那么没有任何立法能够通过，除非现状处于40到80之间，也就是说，如果当前现状位于任何人的理想点的左侧或右侧。这个模型揭示了拥有否决权的投票者的数量和意识形态多样性与出现政治僵局的可能性之间的紧密联系。这是一个具有普遍意义的深刻洞见。拥有多样性否决权的组织很可能无法采取任何行动。

享受竞争模型

在享受竞争模型中，各备选项（通常是各种产品）也是用属性来表示的。但是，在这种模型中，属性包括了质量、效率或价格，而且更多或者更少总是更受欢迎的。为了刻画异质性，享受竞争模型允许个体给不同维度赋予不同的权重。

利用线性回归方法，我们可以运用享受竞争模型推断商品属性的隐含价值。具体程序很简单。这种模型假设收益是商品属性和个人赋予权重的线性函数。如果我们获得了数千幢房子的售价以及关于这些房子属性的数据（面积、卧室和浴室的数量，当地学校的质量，院子的大小和建筑质量等），就可以通过产生回归估计出购买房屋的人赋予每个属性的平均权重（以美元计）。这种回归方法通常称为享受回归。

享受竞争模型

备选项包括 N 个效价属性（valence attribute）：
 $\vec{a} = (v_1, v_2, \dots, v_N)$ 。

个体用各个属性上的权重来表示： $\vec{x} = (w_1, w_2, \dots, w_N)$ 。

该个体得自该备选项的收益（效用）为：

$$\pi(\vec{x}, \vec{a}) = w_1 \times v_1 + w_2 \times v_2 + \dots + w_N \times v_N$$

示例： $\vec{v} = (3, 1, 2)$, $\vec{w} = (4, 2, 5)$

$$\pi(\vec{v}, \vec{w}) = 4 \times 3 + 2 \times 1 + 5 \times 2 = 24$$

其中有些属性，如一个游泳池或新厨房，则是有市场价格的。作为对模型的一种检验方法，我们可以将估计出来的价格与市场成本比较一下。如果回归结果表明，游泳池使房子的价格增加了15万美元，同时建造游泳池的成本却只有15 000美元，我们就知道这个模型遗漏了某些属性。如果回归结果表明增加的价值仅为8 000美元，那么这可能意味着无法收回加建游泳池的全部成本。

其他属性，例如从房屋到市中心的通勤时间，是没有市场价格的。在这种情况下，回归所能得到的是该属性的隐含价格。需要注意的是，隐含价格也可能包括大量信息。表20-1显示了六栋假想房屋的假想价格数据。

表20-1 六栋假想房屋的假想价格

房子坐落地	房子面积 (平方英尺)	卧室数量 (个)	通勤时间 (分钟)	价格 (美元)
松树街 204 号	2 000	3	30	\$200 000
枫树街 312 号	3 000	5	60	\$380 000
格迪斯道 211 号	2 500	4	10	\$310 000
马丁路 342 号	1 500	2	20	\$150 000
克拉克巷 125 号	2 000	4	30	\$220 000
布朗苑 918 号	4 000	6	50	\$360 000

如果假设房子的所有其他属性都完全相同，那么可以运行享受回归。假设得到的回归方程如下：

$$\text{价格 (美元)} = 100 (\text{平方英尺数}) + 20\,000 (\text{卧室个数}) - 2\,000 (\text{通勤时间})$$

这就是说，这个回归方程估计，每增加1平方英尺的价值为100美元，每间卧室的价格为20 000美元；另外，通勤时间每增加1分钟影响房子价值2 000美元。假设一个人买下房子后，要住上20年，在20年间，上班的总天数为4 000天至5 000天。如果取较低的那个数字，那么假设用于通勤的时间每天都能省1分钟，就能省下4 000分钟（或60多个小时）的通勤时间。因此，2 000美元这个估计值相当于每小时30美元左右。换句话说，人们为了住在更接近上班地点的房子里而付出的代价，就“好像”是他们为了免受1小时交通堵塞之苦而付出了30美元一

样。[15](#)

产品竞争的混合模型

空间竞争模型和享受竞争模型在如何表示对各种属性的偏好这一点上有所不同。在空间竞争模型中，每个人对每种属性都有一个更偏好的水平，而且某个备选项在哪些维度上与他的理想点接近，对他来说价值就越大。而在享受竞争模型中，每个属性越多（或越少），人们就越喜欢。

消费品和服务、理想的生活伴侣、公共政策、宗教，以及求职者，全都既包括空间属性，也包括享受属性。每个人也许都有各自喜欢的薯条脆度，但是所有人都希望每份食品的价格更低。脆度是空间属性，价格则是享受属性。不同的雇主在寻找潜在的雇员时，喜欢的人格特征一般会有所不同。有些企业更喜欢外向的人，也有一些企业可能更喜欢内向的人，但是所有企业都更喜欢诚实和正直的人。个性类型是空间属性，而诚实正直则是享受属性。

因此，我们可以创建一个混合模型，让备选项同时包含空间属性和享受属性。这种混合模型可以用来分析市场进入、产品差异化竞争，以及价格竞争的激烈程度。在此，不妨再分析一下前面讨论过的巧克力棒的例子。在确定新产品的属性之前，打算进入巧克力棒市场的企业可能会先将这三种现有产品放在属性空间中分析，然后给消费者发放问卷，以了解他们理想点的分布状况。这个拟进入者可以估计出自己想推出的产品的沃罗诺伊邻域，如果那个邻域的消费者数量很少，就不应期待产品会热销。

任何考虑进入某个市场的企业家都可以采用这种方法。一个靴子设计师可以绘制出现在所有绝缘靴的设计图（那很可能有数十种），然后他可能会发现没有一种设计是用闪亮漆皮的。设计一个用于制作待办事

项列表的智能手机应用程序的开发者，可以列出现有类似应用程序的功能，估计市场总需求并预测可能的销售额。

我们可以将空间竞争模型中的降价行为视为分割线的移动。回顾一下前面的图20-3，图中显示了两种巧克力棒的情形。分割线对应于那些在A巧克力棒与B巧克力棒之间无差异的消费者的理想点。如果生产B巧克力棒的企业进行了降价促销，并且如果消费者更愿意付出更少的钱来得到同样多的糖果，那么这种降价会把分割线推向A并使B的市场份额增大。我们不需要模型就知道B的降价会增大其市场份额，但是我们确实需要这个模型来估计这种影响的大小。关键在于要将拥挤的市场（crowded market）与稀疏的市场（sparse market）区分开来。前者指在一个低维的属性空间挤进了大量的产品，后者则指市场中只有很少的竞争对手。在拥挤的市场中，每种产品都只有一个小小的沃罗诺伊邻域，而在稀疏的市场中，沃罗诺伊邻域是很大的。

价格变化在这两种不同类型的市场中会产生不同的影响。图20-9显示了B巧克力降低10%（假设价格从每根2美元下降为1.8美元）的影响。左图显示的是一个稀疏的市场中的情形。降低B巧克力棒的价格会改变A与B之间的分割线，并将B的市场份额从50%提高到54%（B的市场份额增加了8%）。价格下降了10%，但是销量只增长了8%，综合起来使销售收入减少了3%。因此，在这种情形下，降低价格不是一个好主意。右图显示的是一个拥挤的市场中的情形，市场中有7种不同的巧克力棒。在这里，价格下降对B的绝对市场份额影响也不算太大，只是从15%增加到了20%，但这5%的增幅却代表着B的市场份额有相当大的增长比例（达到了33%）。总体影响是使销售收入增加了20%。¹⁶

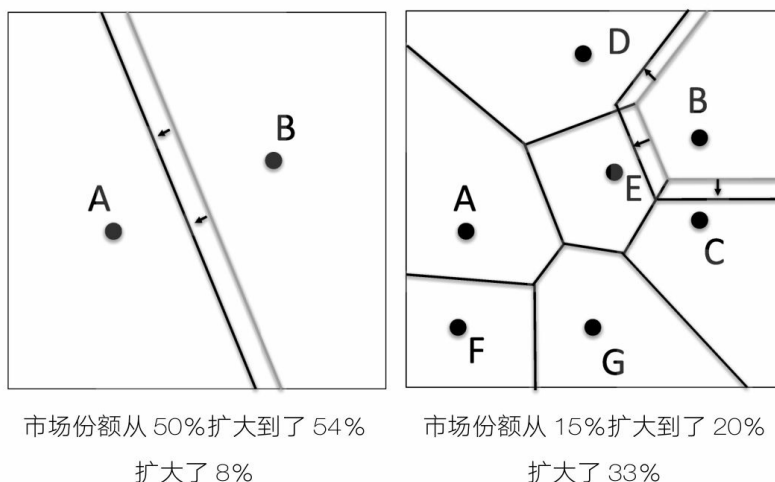


图20-9 稀疏的市场与拥挤的市场中的价格竞争

因此，这个模型的预测是，拥挤的市场中的价格竞争比稀疏的市场中的价格竞争影响更大；同时，大宗商品^[4]市场上的价格竞争则可以称得上是“极端的竞争”，其影响最大。这个模型还预测，高端时尚商品的价格竞争将很少出现，因为产品的高维性能够创造出一个稀疏的市场，设计师可以维持大幅度的价格加成。

属性数量与价格竞争激烈程度之间的这种关系表明，增加新属性不失为一个好策略。这将使市场更加稀疏，从而降低价格竞争激烈度，并带来更高的利润。然而，即便这种推断本身是正确的，要实现这个策略也并不容易。无论如何，人们必定会给新的属性较高的估值。事实上，每一个成功的新属性的出现（例如无绳立体声扬声器），背后都隐藏着多次失败尝试的历史（例如比克公司命运多舛的一次性产品）。

小结

空间竞争模型、享受竞争模型，以及两者的混合模型提供了一个很好的框架，在这个框架内，我们可以表征不同的产品、政治候选人，甚至是求职者。这些模型可以测量意识形态“位置”、隐含价格属性，并评估潜在的市场进入者。它们提供了很多深刻的洞见，有助于我们理解市

市场竞争如何产生差异化的激励、政治竞争如何产生建立联盟的激励，以及只具有较少属性的产品的价格竞争为什么会更加激烈。

这些模型都做出了相当强的、从经验上看相当可疑的假设。例如，在前面讨论的模型中，我们假设人们的偏好是不会改变的，他们不会屈服于社会压力。如果真的是这样的话，为什么企业和政客会花费大量金钱和资源去改变消费者和选民的偏好呢？对于这种批评，我们可以再次援引博克斯的格言来辩解：所有的模型都是错的。

还有一种更精细的回应方法可以对基本偏好与工具偏好加以区分。前者指人们的欲望（想得到什么），后者指人们对能够产生根本性结果的属性的偏好。学生的基本偏好可能会在成为一个受欢迎的人、身体健康和取得学术成就之间维持平衡。为此，一个学生可能会通过提早起床、喝果汁、尽快完成自己的家庭作业（以保证晚上有时间参加社交活动），来实现这些根本性的目标。而这个学生对水果奶昔的选择则是工具性的，水果奶昔有助于实现他对健康的基本偏好。如果哪一天，他看到一篇科学论文说水果奶昔的糖分很高，不利于身体健康，他就可能会改喝饮用水。如果是这样，他的工具偏好就会发生改变，但是他的基本偏好并没有变化。在这里，我们再一次看到，模型本身并不是目的，只是提供了一个框架来构建我们的思维。

多个价值模型

在市场经济中，我们可以通过一个愿意支付的价格来衡量一件商品对这个人的价值。一个人可能会认为一个五香熏牛肉三明治值7美元、戈雅（Goya）^⑧的一幅画值300万美元、佛罗里达州奥卡拉市的一块土地值75 000美元。但是，许多经济模型都忽略了这些估值的来源。美国经济学家乔治·施蒂格勒（George Stigler）引用了大文豪陀思妥耶夫斯基（Dostoyevsky）的《卡拉马佐夫兄弟》（*The Brothers Karamazov*）一书中那个著名的情感主义者米蒂亚

（Mitya）说的一句话。他说：“De gustibus non est disputandum。”（口味各异，难言好坏）我们这里所讨论的模型确实比较少讨论口味本身，而更多讨论口味如何转化为人们赋予商品的货币价值。

享受模型：这个模型根据商品的内在属性解释商品的价值。估值的差异取决于对商品属性的不同潜在偏好。这些属性可能是商品的物理组成部分。例如，一个五香熏牛肉三明治中包括了黑麦面包、五香熏牛肉、瑞士奶酪、芥末、泡菜和洋葱。然后，我们可以将它的价值写成这些组成部分的价值的加权线性组合。更精细的享受模型还可以包括交互项。例如，如果与烤黑麦面包一起售卖，那么五香熏牛肉可能会更有价值。

协调模型：这种模型将价格解释为社会建构的。戈雅画作的价值取决于其他人相信它真的有那么高的价值。在最初，人们对戈雅画作的价值有自己的看法。然后，他们与社会网络中的其他人互动并更新自己的看法。两个人都可以将他们的价值设定为等于两人估价的平均值，一个人可以改变他的估价以匹配另一个人给出的价值，或者每个人都可以根据其他人的估价来调整自己的估价。给定这三种假设中的任何一种，估价都可以实现局部收敛。这样，相互之间有联系的人将会具有类似的估值。分配给商品的最终价值将取决于价值的初始分布、社会网络以及配对发生的顺序。

预测模型：这类模型将价格解释为对未来价值的预测。根据这种模型，佛罗里达州奥卡拉市的土地价值、比特币或股票价值，都取决于人们将来愿意为它们付出多少钱。这些估值取决于预测模型，而预测模型又取决于属性和类别。我们可能会将奥卡拉市归类为气候温暖的、税率很低的非沿海城市。人们的估值的变化源于不同的预测模型。投资者会使用多种预测模型。这些模型可能依赖于属性，或者，就像对比特币进行估价一样，也会对协调做出假设。

这三种模型为商品的价值提供了三种不同的解释。没有任何一个模型在所有情况下都是最好的。在特定情况下，每个模型都可能成为最好的那个。模型就像箭袋中的箭。五香熏牛肉三明治的价值很可能来自其内在性质，戈雅画作的价值则可能在很大程度上是社会建构的——只要人们认为他的画作有价值，那么它们就有价值。对佛罗里达州土地价格的估计则可能取决于对未来房地产价值的预测。

21 博弈论模型

演绎推理是从最抽象到最不抽象的推理。它从一套公理开始，运用逻辑定律和数学规律来操纵，形成对世界的预测。

雷切尔·克罗松（Rachel Croson）

本章以及随后各章中讨论的许多模型，包括合作模型、信号模型、与机制设计有关的模型和与集体行动有关的模型都要涉及博弈。我们不会在这里非常深入地探讨博弈本身，因为实际上整本书都是讨论这个主题的。本章的目标是提供一个适当的入门介绍，为此，我们给出了三种主要类型的博弈的一些例子：标准式博弈，博弈参与者在一组离散的行动（通常为两种）中做出选择；序贯博弈（*sequental game*），博弈参与者按顺序选择行动；连续行动博弈，博弈参与者可以选择任意尺度或效果的行动。我们通过这些例子介绍了博弈的主要概念，有助于理解后面章节中给出的模型。当然，它们本身也有讨论的价值。

在本章的其余部分，我们首先讨论 2×2 标准式零和博弈。在零和博弈中，两个博弈参与者中的每一个都要在两个行动中做出选择，无论某个博弈参与者选择什么行动，一个博弈参与者得到的收益，都会被另一个博弈参与者遭受的损失所抵消。我们利用零和博弈的例子来定义博弈论的基本术语，区分策略和行动，并引入迭代消除被占优策略的概念。然后，研究市场进入博弈（*market entry game*），市场进入博弈是序贯博弈的一种。我们将在重复市场进入博弈的框架下讨论所谓的连锁店悖论（*chain store paradox*）。接着，考虑一个“努力博弈”，这是一种连续行动博弈。在这个博弈中，个体选择努力水平以赢得固定金额的奖励，付出努力越大，博弈参与者赢得奖励的机会越高。本章最后简要讨论了博弈论模型的一般价值。

标准式零和博弈

在本节中，我们分析两个双人标准式零和博弈（two-player normal-form zero-sum games）。在这种博弈中，每个博弈参与者选择一个行动，并根据博弈参与者自己的行动和另一个博弈参与者的行动获得一定收益。此外，博弈参与者双方的收益总和为零。

在图21-1所示的第一个硬币配对博弈中，每个博弈参与者都要在两个行动中做出选择：猜硬币正面朝上还是背面朝上。行博弈参与者希望自己的选择与另一个博弈参与者的选择匹配，而列博弈参与者则不希望匹配。收益如矩阵所示：

	正面朝上	背面朝上
正面朝上	1, -1	-1, 1
背面朝上	-1, 1	1, -1

图21-1 硬币配对博弈

博弈的策略是如何进行博弈的规则，它可以是对单个动作的选择、在不同行动之间的随机化，或者，正如在下一节中将会看到的，也可以是一个行动序列。博弈的纳什均衡（Nash equilibrium）是指这样一种策略，它们能够使每个博弈参与者的策略在给定其他博弈参与者策略的情况下是最优的。在硬币配对博弈中，存在一个唯一的均衡策略，那就是，两个博弈参与者都以相同的概率在两个行动之间进行随机化。为了证明随机化是一种均衡，只需要证明，如果某个博弈参与者随机化，那么另一个博弈参与者选择任何行动都不可能比随机化更好。

要证明这一点很简单。如果行博弈参与者（在图21-1中，行博弈参

与者的行动以黑体显示) 以 $1/2$ 的概率选择正面朝上、 $1/2$ 的概率选择背面朝上, 无论他的选择到底是什么, 列博弈参与者的收益都为零。正因为如此, 随机化是列博弈参与者的最优策略。根据对称性, 随机化也是行博弈参与者的最佳选择。

随机化策略的最优性, 对策略互动环境中的行为有很大的意义。体育运动也是零和博弈: 一方获胜, 另一方就要落败。在点球大战中, 一名前锋希望在瞄准球门左侧与球门右侧之间进行随机选择; 在网球比赛中, 发球方要随机将球发到内角或外角; 在足球比赛中, 进攻方希望在跑动与传球之间随机选择。而且, 在所有这些比赛中, 另一方也会随机化他们的反应。任何非随机性都可能会被对手利用, 扑克等纸牌游戏也是如此。一个优秀的扑克玩家会随机地虚张声势。如果他一直虚张声势, 对手就会了解这种策略, 他就会落败。当然, 对手的最优策略也是随机地虚张声势, 这样就同样有可能赢或输。

现在讨论最小化风险博弈 (minimize risk game)。在图21-2所示的这个博弈中, 每一个博弈参与者都可以选择采取冒险的行动或安全的行动, 这是一个非对称的零和博弈。博弈参与者的收益不仅取决于自己的行动, 还取决于哪一个博弈参与者采取了哪一个行动。在这个博弈中, 行博弈参与者有一个占优策略, 即采取安全的行动。无论列博弈参与者选择哪一个种动作, 对于行博弈参与者来说, 选择安全的行动总是更好的。但是对于列博弈参与者来说, 情况却并非如此。如果行博弈参与者选择冒风险, 那么列博弈参与者也应该选择冒风险; 如果行博弈参与者选择了安全的行动, 那么列博弈参与者也应该选择安全的行动。

	选择冒险	选择安全
选择冒险	-10, 10	0, 0
选择安全	10, -10	0, 0

图21-2 最小化风险博弈

通过考虑对行博弈参与者激励的情况下，列博弈参与者可以推断出行博弈参与者总是会选择安全的行动，那么他也会选择安全的行动。这种一方为另一方排除最优策略被称为迭代消除被占优策略。因此，两个博弈参与者都选择安全的行为是这个博弈的纳什均衡。

序贯博弈

在序贯博弈中，博弈参与者按照某个特定的顺序采取行动。由此，可以用一棵博弈树（game tree）来表示一个序贯博弈。博弈树由节点和边组成，每个节点对应于博弈参与者必须采取行动的时刻，该节点的每条边分别表示可以采取的某个行动。在博弈树最末尾的分支上，我们写下相应行动路径的收益。图21-3所示的博弈树显示了市场进入博弈。

在市场进入博弈中，有两个博弈参与者：拟进入者和现有企业。如果拟进入者选择不进入市场（博弈树的左侧分支），那么它的收益为零，现有企业的收益为5。如果拟进入者决定进入市场，那么现有企业必须做出选择：是接受新进入者，同时自己的收益从5下降为2，还是发动与新进入者的商战，但这会导致自己的收益变为零，同时令新进入者的收益为负。之所以假设这种情况下新进入者的收益为负，因为它必须为进入市场付出一定的成本。

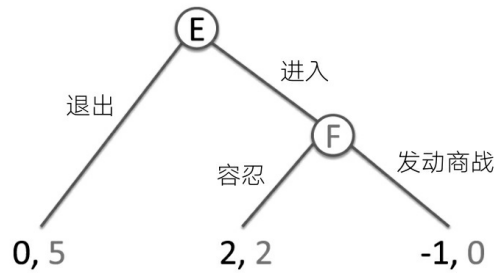


图21-3 市场进入博弈

在序贯博弈中，策略对应于每个节点处的行动选择。假设现有企业在发现有新企业进入时决定发动商战。那么，如果拟进入者知道这一点，就不会选择进入，因为这种情况下进入会产生负收益。这个行动序列——拟进入者选择不进入、现有企业在拟进入者进入时就会发动商战，是一个纳什均衡。然而，这并不是唯一的纳什均衡，也不是最有可能出现的结果。拟进入者选择进入市场，现有企业决定接受（不发动商战），这是第二个均衡。

那么，应该如何在这两个均衡之间做出选择呢？我们可以利用细化准则。在序贯博弈中，一种常见的细化准则是选择子博弈完美均衡（subgame perfect equilibrium）。可以运用逆向归纳法（backward induction）来求解子博弈均衡：从最末端的节点开始，并在每个节点处选择最优行动。然后沿着博弈树逆向倒推，假设每个博弈参与者会在给定另一个博弈参与者在后续节点上的行动时选择最优行动。例如，在市场进入博弈中，我们从现有企业的末端节点开始推导。它有一个最优行动，即接受对方进入。然后移动到博弈树上的节点，不难发现拟进入者的最优策略是进入。

这个博弈在重复进行时会变得更加有趣。试想一下，现有企业也可能存在于许多个市场中。也许它是一家连锁企业，在几十个城市都有门店。再假设存在一系列的拟进入者。那么，这个企业将陆续地进行一系列市场进入博弈。

如果现有企业从最后一个市场开始使用逆向归纳法进行推理，那么

它将接受最后一个市场中的进入者。根据同样的逻辑继续推导，现有企业将接受倒数第二个市场中的进入者，以此类推，它将接受所有市场中的进入者。在序贯博弈唯一的子博弈完美均衡中，所有潜在的进入者都选择进入，现有公司接受所有。

虽然每个市场的进入和接受都是唯一的子博弈完美均衡，但实际上，这可能不会发生。假设我们是现有公司董事会的成员，我们面对的是第一个进入者曾经学过博弈论并已经进入市场。我们可能想要竞争，试图阻止其进入其他市场。如果竞争是可信的，那这将是一个明智的策略，也就是说，如果能够建立一个愿意竞争的声誉。我们希望创造的结果不同于子博弈完全均衡。

博弈理论家将这种情况称为连锁店悖论。这是一个例子，博弈论认为的最优行为可能不是一个老练的行为者在利害关系很大时所选择的行为。这个例子并没有反驳博弈论或破坏理性选择假设，而是揭示了为什么我们总是必须挑战假设。

连续行动博弈

我们现在研究另一种博弈。在这种博弈中，博弈参与者可以在连续的可能行动集中进行选择。在连续行动博弈中，行动对应于努力水平。通过选择更大的努力，博弈参与者能够增大自己赢得奖励的概率。这个博弈还允许考虑任意大数量的博弈参与者。

努力博弈

N 个博弈参与者中的每一个人都要选择以货币形式表达努力水平，以赢得价值为 M 的奖励。一个博弈参与者赢得奖励的概率等于他的努力水平除以所有博弈参与者的总努力水平。如果令 E_i 表示博弈参与者 i 的努力水平，那么他的获胜概率由以下方程式表示：¹

$$P(i \text{ wins}) = \frac{E_i}{(E_1 + E_2 + \dots + E_N)}$$

均衡努力水平为： $E_i = \frac{M}{N} - \frac{M}{N^2}$

均衡努力水平的表达式揭示了很多重要的含义，正如我们所预料的那样，个人的努力水平会随着奖金的增多而增大。同样，在均衡状态下，总努力水平将会小于奖金的价值。在假设博弈参与者会进行最优化的情况下也会得到这些结果。博弈参与者应该付出一定努力以赢得奖励，但是不应该付出不合理的努力水平。

通过增加博弈参与者的数量，可以看到其对个人和总体努力的影响。根据模型，即便每个人的努力水平都下降了，总努力水平也会增加。这个结果说明，那种吸引了大量“参赛者”的研究课题竞标、建筑设计竞赛和征文比赛，反而可能会产生水平不那么优异的赢家（与“参赛者”较小的竞赛相比），因为在这种“参赛者”众多的竞赛中，个体参与者付出的努力水平会较低。

小结

本章一开始，我们先讨论了零和博弈，这类博弈不包括互利的行动组合。任何对一个博弈参与者有益的行动都必定会损害另一个博弈参与者。在零和博弈中，任何行为对一方的“益”，在数量上都等于对另一方的“损”。从一个人那里拿钱给另一个人，就是一种零和行动。许多个人行动和政策选择至少在一个方面是零和的，我们每天只有这么多的时间可用，这么多钱可花，这么多资源可分配。也就是说，在这个维度上的零和行动，在另一个维度上可能不是零和的。例如，重新安排预算，在货币的数量这个维度上是零和的，但是就人的幸福感或满足程度这个维度而言，却可能是正和的或负和的。

我们应该始终探究提议的政策变化是否会导致零和博弈。许多人都

对家长的择校行为颇有微词（这里所说的择校问题，主要是父母将孩子送进什么样的学校学习的能力），因为它加剧了竞争。但是迫于竞争的压力，学校的教学质量将会得到提高，至少从逻辑上看是这样。

然而，只有在产能过剩的情况下，学校才有动力提高质量，否则，择校会在学生中造成零和博弈。假设一个城市有1万名学生和10所学校，每个学校可以接受1 000名学生。如果所有学生都以相同的方式对学校进行排名，那么最好的学校的名额将只能通过抽签来分配，中签的学生将去更好的学校，未中的学生将去更差的学校。学生们玩的是一场零和博弈。如果新的学校开放或现有学校改善，就将不再是零和博弈，每个人都可以赢。

市场进入博弈与零和博弈都提供了有益的洞见。市场模型揭示了学校改进质量的动力。零和博弈则表明有些学生将受益、有些学生将受损。每个人所承受损害或获得的收益的大小，则取决于具体情况：学生和家长对学校的质量的了解程度如何，学校还有存在多少剩余名额，学校是否真的知道如何提高教学质量，会创办新的学校吗？

这两个博弈都没有给我们一个正确的答案，但是都产生了有用的见解。择校会带来竞争。它还产生了一个具有零和博弈特征的大规模排序问题。竞争的积极方面是否会超过消极的排序成本取决于环境。我们必须把模型排列在一系列事实的基础上，才能做出正确的政策选择。

识别问题

关于人们行为的数据经常揭示出人类行为的“聚类倾向”。优秀的学生更有可能与其他优秀的学生成为朋友，而不怎么可能与学习困难的学生成为朋友。有过犯罪前科的人比从未犯过罪的人更有可能与犯罪的人交往。在社交网络中，各种各样的“社会善”和“社会恶”——吸烟嗜好、健美的人、肥胖者，甚至幸福也都会“聚类”。人们还会根据信仰或意识形态聚在一起。

有两个模型可以解释这种聚类：同伴效应模型（peer effect model）和分类模型（sorting model）。同伴效应模型用博弈论来解释聚类现象，即一个人与他的朋友一起进行协调博弈。而在分类模型中，人们会“迁移”到与他们相似的其他人附近。一群优秀的学生之所以聚在一起，可能是因他们要协调完成某个共同行动（同伴效应），或者也可能是因为优秀的学生就喜欢找优秀的学生一起玩（分类效应）。如果只有数据快照（snapshot of data），那么这两者是无法区分的。

数据：学生既有可能获得高分H，也有可能获得中等分数M，两者是等可能的（概率相同）。假设每一个学生都会加入某个人数为4的“小团体”，则有如下分布： $p(\{H, H, H, H\}) = P(\{M, M, M, M\}) = 1/16$ ， $p(\{H, H, H, M\}) = P(\{M, M, M, H\}) = 3/16$ ， $p(\{H, H, M, M\}) = 0$ 。

同伴效应模型：学生最初先会形成一些人数为4的随机小组： $p(\{H, H, H, H\}) = P(\{M, M, M, M\}) = 1/16$ ， $p(\{H, H, H, M\}) = P(\{M, M, M, H\}) = 1/4$ ，并且 $p(\{H, H, M, M\}) = 6/16$ 。属于仅包含一种类型的小组的人数保持不变。与小组内所有其他人类型都相反的人会切换类型，因此， $\{H, H, H, M\}$ 这样的小组会变为 $\{H, H, H, H\}$ 。在每种类型的人数相同数量的小组中，会有一个成员切换类型，也就是说， $\{H, H, M, M\}$ 这样的小组有同样的概率会成为 $\{H, H, H, M\}$ 或 $\{M, M, M, H\}$ 。

分类模型：学生最初先会形成一些人数为4的随机小组。在具有两种类型的任何一个小组中，与至少两个其他人的类型相反的人，会与相反类型的某个人交换小组。也就是说， $\{H, H, H, M\}$ 会变为 $\{H, H, H, H\}$ ，而且 $\{M, M, M, H\}$ 会变为 $\{M, M, M, M\}$ ；并且，任何 $\{H, H, M, M\}$ 这种形式的小组有同样的概率成为 $\{H, H, H, M\}$ 或 $\{M, M, M, H\}$ 。

两个模型都与数据一致，从而导致了识别问题。只有数据快照，我们无法确定吸烟、看漫画书或喜欢滑板到底因为同伴效应导致的，还是因为分类效应导致的。在某些情况下，我们可以推断出使用哪种模型更好。例如，美国中西部地区的人们喜欢说“pop”（泡泡）、沿海地区的人们喜欢说“soda”（苏打），对于这种倾向，我们可以有把握地认为这是由同伴效应驱动的，因为很少从外地移居波士顿的人，会将可口可乐称为“苏打水”。但是，对于某些利害关系更大的行为，例如学业成绩、滥用药物、肥胖和幸福，就需要用更多的时间序列数据来识别哪种模型才适用了。利用时间序列数据，我们就可以分辨出人们到底是在改变自己的行为（同伴效应），还是在更换他们的朋友（分类效应）。在许多情况下，这两个因素都有。²

22 合作模型

从来没有人因施舍而变得贫穷。

安妮·弗兰克（Anne Frank）

经常会有人要求科学家列出他们心目中最重要的问题，各个领域的专家给出的最重要的问题往往各不相同，例如，宇宙是如何形成的？意识是怎样出现的？我们能否找到根治癌症的方法？等等。在这些问题当中，有一个是社会科学家和生物科学家都认为是重要的问题，那就是，合作是怎么产生的？¹合作要求合作者采取不符合自身利益的行动，而这就意味着我们不会经常观察到合作现象。但是在现实世界中，我们却看到合作出现在无数领域中，而且达到了非常大的规模。合作在细胞层面上就存在：细胞通过黏附作用实现了合作，一个细胞会产生细胞外物质，供其他细胞黏附之用。我们观察到，蚂蚁、蜜蜂、人类、人类组织之间，甚至国家之间都存在着广泛合作，不同国家会在制定条约和国际法方面进行合作。

在本章中，我们运用模型来讨论合作如何产生、如何维持以及怎样做才能创造更多的合作。当然，本章给出的这些模型无法完美地解释世界上广泛存在的各种合作，比如，为什么乌鸦会告诉其他乌鸦所发现的腐肉地点，为什么裸鼯鼠会共同防御以它们为猎物的天敌，为什么攀爬藤蔓种植在“亲属”旁边时根系较不发达，为什么白蚁和蜜蜂会建造出精致复杂的巢穴，为什么蚂蚁能用身体和附肢搭起运送食物的桥梁，等等。但是，这些模型确实能够告诉我们很多重要的结论。²

尽管我们看到，物种内部和物种之间广泛存在着合作的例子，但是我们同时还应该看到合作失败的情况。合作的程度取决于具体环境。联盟既能吸引人们加入，也会失去原有成员；英国是欧盟的创始成员国之

一，但是后来却要退出欧盟。一个积极为学校筹款活动服务的人，却可能在超市排队时插队或者偷税漏税。一头狮子，既可能加入围猎水牛的团体，也可能自己偷偷地去猎杀野猪。并不是每个物种都合作。黑胡桃树的根须将一种名为胡桃醌（juglone）的物质释放到土壤中，以抑制附近植物的生长。

细胞、树根、乌鸦、人、企业和国家等合作型实体行为特征的多样性，要求采取多模型方法。也许，我们最好将细胞和植物建模为遵循固定规则的，将乌鸦、蚂蚁和狮子建模为运用更多依赖于环境或过去结果规则的，将人、企业和国家建模为有前瞻能力并会进行成本收益计算的。

本章的第一个要点是：合作可以通过多种机制、在多种环境下涌现出来并维持下去。我们讨论了四种促进合作的机制：重复、声誉、局部聚类 and 群体选择。这些机制都能在没有外部干预或管理的情况下进行合作。它们可以适用于合作的鼯鼠、蜜蜂和人类，人类还有更正式的促进和保持合作的方式。在本章最后的讨论中，我们描述了其他制度性的解决方案，包括付钱让人们进行合作，如果不这样做就惩罚，或者制定法律强制人们的合作行为。

本章的第二个要点是，这些机制中任何一个的“效力”都取决于那些正在合作的实体所拥有的“行为曲目”（behavioral repertoires）。一些机制，特别是通过重复实现合作这个机制，几乎适用于任何行为。声誉和规范这两个机制则需要前瞻性行为和信息共享，对于那些更“老练”的行为者来说更加有效。

局部聚类对合作的效果取决于具体模型。由演化力量选择支持或反对的行为者之间的合作，最常出现在稀疏的网络上，而通过规范进行合作则需要密集的网络。群体选择的有效性则取决于行动者的前瞻能力和适应速度的细微差别。行为者所拥有的更强的前瞻能力，能够增强群体选择的力量，而允许行为者更快地适应则可能会阻碍群体选择发挥作

用。为了探索这些问题并分析清楚行为假设与合作结果之间的相互关系，要利用我们熟悉的囚徒困境博弈模型和合作行动模型。合作行为模型允许我们刻画有利于多个参与者的行动，并对网络上的合作行为进行建模。

本章安排如下。我们首先描述了囚徒困境博弈，同时说明理性行为者之间的合作是怎样得以维持的。接着，我们阐述了重复行动是如何促进基于规则的行动者之间的合作的，以及为什么不断发展合作比单纯维持合作更加困难。然后，我们考虑不太复杂的生物学意义上的行为者之间的合作，并阐明了亲缘选择和局部聚类如何促进了合作。最后两部分讨论了群体选择以及我们如何综合这些模型实现更高程度上的合作的问题。

囚徒困境博弈

囚徒困境博弈的名称源于如下故事。有两个人，被控共同犯下了某种罪行。有关当局只掌握了间接证据，因此给他们每个人都提供了认罪减刑的机会。两人因此面临着两难选择：如果两人都不认罪，那么每个人都会（根据现有证据）受到轻微的惩罚；如果只有一人认罪，那么认罪的这个人不会受到惩罚，而另一个人则会受到很严厉的惩罚；如果两人都认罪，那么两人都会受到较严厉的惩罚，但是不会像只有一个人认罪时那么严厉。

图22-1将这个故事表述为一个双人博弈。每个博弈参与者可以选择合作或背叛。图中的灰色数字表示列博弈参与者的收益，黑色数字表示行博弈参与者的收益。对每个博弈参与者来说，背叛都是占优策略，无论其他博弈参与者采取什么行动，背叛都能带来更高的收益。但是，如果两个博弈参与者都背叛，每个博弈参与者的收益都会低于双方合作时的收益。因此，追求自身利益的行为导致了集体利益的恶化。

	合作	背叛
合作	3, 3	1, 4
背叛	4, 1	2, 2

图22-1 囚徒困境博弈的例子

囚徒困境博弈紧紧抓住了现实世界中许多情况下的核心激励，它可以用来建模一些国家之间的军备竞赛：背叛对应于将资源用于开发武器，合作对应于发展经济。还可以用来建模竞选活动中的广告战：背叛对应于投放负面广告，合作对应于投放正面广告。它甚至可以用来解释为什么雄孔雀会有如此之长的尾巴，每只孔雀都有很强的动力使自己看上去比其他孔雀更强壮、更健美。

很多囚徒困境博弈都是在事后才认识到的。许多新技术的最早一批采用者，例如，最早使用ATM机的银行，发现自己的利润因此大为增加。但是，当其他银行也跟进时，利润就因竞争加剧而下降了。事后证明，可以把选择使用ATM机类视为一种“背叛”。³

如图22-2所示，一般形式的囚徒困境博弈假设如果两个参与者都选择背叛，那么基线收益为零。这样一来，我们就可以用三个变量来表示这个博弈：来自合作行为的奖励 R ，背叛的诱惑 T ，以及被损害的一方的收益 S （参见图22-2中的收益矩阵）。收益矩阵下面的不等式确保了选择背叛是一个占优策略，而选择合作则能够产生有效率的结果。

	合作	背叛
合作	R, R	-S, T
背叛	T, -S	0, 0

$$T > R > 0 > -S$$

$$2R > T - S$$

图22-2 一般形式的囚徒困境博弈

通过重复和声誉机制实现合作

我们首先阐明，博弈的重复进行和声誉的建立为什么能够维持理性行为者之间的合作。能够维持合作这个事实，并不能保证合作真的能够实现，它只是说明，如果合作“不知怎么”出现时，理性的博弈参与者能够维持它。为了证明重复行动能够维持合作，我们构建了一个重复博弈模型。在这个重复博弈中，每次博弈结束之后，都以概率 P 再一次进行。从理论上说，这种博弈可以永远持续下去。

博弈参与者在重复博弈中，要根据以往的博弈历史选择行动。在这里，我们考虑一个被称为冷酷触发（grim Trigger）的重复博弈策略。具体来说，这个策略是，在第一次博弈中选择合作，并且，只要另一个博弈参与者不背叛，那么就在未来的所有博弈中一直选择合作；但是，一旦另一个博弈参与者背叛了，那么就永远选择背叛。冷酷触发策略是“永不饶恕”的。如果两个博弈参与者都采用冷酷触发策略，那么双方将会永远合作。

要想证明冷酷触发策略能够在重复博弈中维持合作，我们只需要证明，如果一个博弈参与者选择了冷酷触发策略，那么另一个博弈参与者也可以通过采用冷酷触发策略获得最高的收益。由于第二个博弈参与者

的任何一个偏离合作的行为都会导致第一个博弈参与者无休止的背叛，所以第二个博弈参与者只需要对一直合作的预期收益，与一次背叛再加上两个博弈参与者此后都一直背叛的预期收益加以比较就行了。⁴而冷酷触发策略能否带来更高的收益，则取决于诱惑的大小、合作回报的多少以及重复博弈的概率。

重复博弈维持合作

在重复囚徒困境博弈中，如果继续进行下一次博弈的概率 P 超过了诱惑收益 T 减去奖励收益 R 的差与诱惑支付的比，那么采用冷酷触发策略就能维持合作，即：⁵

$$P > \frac{(T-R)}{T}$$

这个结果告诉我们，如果诱惑收益超过了奖励收益的三倍，即 $T > 3R$ ，那么博弈必须以超过 $2/3$ 的概率重复。这个不等式还告诉我们，如果合作的奖励增大了、博弈重复进行的可能性增加了，或者背叛的诱惑减少了，合作就会变得更加容易维持。这些含义中的每一个，都意味着一条直观的促进合作的途径：增加合作奖励，让重复进行博弈的可能性更大，以及减少背叛的诱惑。虽然这些都只是非常简单的推论，但是在写出这个模型之前，它们可能并不是一目了然的。

通过思考合作的必要条件，我们还可以推断出一些不那么直接的结论。上面这个不等式还意味着，如果博弈参与者认为博弈重复进行的概率在未来会下降到低于阈值，那么理性的博弈参与者将会在这种概率变化发生之前就停止合作，而不会等到变化发生时再停止合作。⁶

而且，重复博弈能不能维持理性的博弈参与者之间的合作，还取决于这个模型的一个特别假设：博弈会以一定概率不断重复进行下去。相反，如果转而假设博弈只会重复一定次数，比如只会重复进行3次，那么理性的博弈参与者将不会选择合作。

我们可以利用逆向归纳法来证明这一点。假设博弈只重复进行3次并且第一个博弈参与者将采用冷酷触发策略。再假设 $T=3$ 、 $R=2$ 、 $S=1$ 。给定这样的收益矩阵，如果第二个博弈参与者在所有三轮博弈中都选择合作，那么他的总收益为6。接下来需要确定是不是所有其他策略都不能带来更高的收益。如果第二个博弈参与者在第一轮博弈中就背叛，那么他所能得到的收益仅为3，因为在他背叛之后，第一个博弈参与者在后面两轮博弈中都选择背叛。如果第二个博弈参与者在第二轮博弈时再背叛，那么他可以得到的收益为5。但是这两个策略都是不理性的。如果在第三轮博弈中才背叛，那么他可以得到的收益为7（前两轮博弈中各得到2，最后一轮博弈中得到3）。因此，理性的博弈参与者会在最后一轮博弈背叛。

但是，第一个博弈参与者（他宣布自己会采取冷酷触发策略）应该会意识到第二个博弈参与者在第三轮博弈中背叛，因而他也会在第三轮背叛。第二个博弈参与者也会意识到两个博弈参与者都会在第三轮博弈中背叛，因此他在第二轮博弈中就会背叛。根据同样的逻辑，第一个博弈参与者也会在第二轮博弈中就背叛，这种推理也适用于第一轮博弈。事实上，只要博弈只重复有限次数（无论多少次），上述结果就都适用。在最后一轮博弈中，理性的博弈参与者会背叛。结果，两个博弈参与者都有动力在倒数第二轮博弈中背叛，以此类推，他们会在所有轮次的博弈中都背叛。从而，唯一符合理性的策略就是永远背叛。

目前为止，我们一直是在一个孤立的环境中考虑两人博弈的，并没有考虑一个人的背叛行为可能会影响其他人对待背叛者的态度。这样做，其实是将这种两人博弈圈进了一个封闭世界中。我们可以扩展这个模型，让博弈在一个社区的成员之间进行，并让这些人有机会监督其他参与博弈的人的行为。

现在假设，每一天，先让这些人随机两两配对，然后进行囚徒困境博弈。在这种情况下，我们假设这个社区的所有成员相信博弈将会永远

重复进行下去，因此未来继续博弈的概率等于1。在这些假设条件下，每一个人都不太可能在第二天仍然与前一天博弈过的那个人博弈，所以他们背叛的动机会更加强大。但是，由于我们假设同一社区的人可能会认出谁是背叛者，所以那些背叛过的人很可能会留下不好的名声，根据假设，这个社区中的任何人将来都不会与声誉不好的人合作。

如果我们用 P_D 表示一个人背叛且被发现、从而留下了一个背叛者的坏名声并在未来的所有博弈中受到惩罚的概率，那么，通过声誉机制维持合作的条件，也就与重复博弈条件下维持合作的条件一样了。这个条件可以写为 $P > \frac{(T-R)}{T}$ ，用一个背叛被发现的概率取代博弈重复进行下去的概率 P 。

在声誉模型中，合作是通过社区来实施的。背叛并被发现的人，在未来将会遭到所有博弈参与者的背叛。在这里，个体还是会计算背叛的收益和成本。他们还必须相信其他人会坚持惩罚到底，而这就意味着其他所有人都会背叛那些背叛的人。要做到这一点，所有个体必须要么彼此认识，要么有某种方法来识别或标记过去的背叛者。因此，在其他条件相同的情况下，规模较小的社区的成员应该能够更好地通过这种机制来实施合作。在美国北方的小城镇，人们在冬季会将车停在商店停车场里。他们不用担心汽车会不会被人偷走（“背叛”），因为他们认识城里的每个人。任何偷车的人，即便是恶作剧，都会导致自己的声誉下降。

实物标签可以使声誉变成一个公共信息，从而有利于能够维持合作。在纳撒尼尔·霍桑（Nathaniel Hawthorne）的小说《红字》（*The Scarlet Letter*）中，海丝特·白兰（Hester Prynne）被迫穿着一件上面有一个猩红色的“A”的衣服，以表示她犯下了通奸的过错。在有些社会中，小偷被定罪后，手会被砍掉，这无疑是一个代价极其高昂的标签。

背叛者会被打上标记，这种情况甚至也发生在了人类之外的其他物种中。在大海中，医生鱼——裂唇鱼（*labroides dimidiatus*），既可以选

择清除作为它的“邻居”的其他鱼类身上的寄生虫（合作），也可以吃某种更美味的其他食物（背叛）。如果裂唇鱼合作，那么它的“邻居”就可以少受寄生虫的困扰。其他鱼也可观察到寄生虫有没有减少。于是，裂唇鱼“邻居”的清洁程度就成了一个标签，一个代表声誉的具体形象。⁷

连通性与声誉

通过声誉机制维持合作有一个条件：个体必须能够知悉自己邻居的行为偏离合作的可能性。为了评估关于这种行为偏差的信息传播出去的可能性，我们可以应用在向传染病模型中加入网络时学到的三个结果。首先，网络的度越大，关于偏离合作行为的信息传播出去的可能性就越大。其次，度分布的变化，特别是超级传播者的存在，也会增加信息传播出去的可能性。再次，如果一个人所背叛的那个受害者，与这个人的其他邻居没有任何联系，那么这个人的邻居就不会知道这个人背叛了他人。因此，要保证声誉扩散，网络必须具有很高的聚类系数，而聚类系数又是社会资本的一个衡量指标。

规则行为者之间的合作

现在，我们放松关于理性的假设，转而假设博弈参与者只会遵循诸如冷酷触发策略之类的规则行事。我们将利用这个更一般的模型探析合作是否可能出现以及如何出现。在这个模型中，假设一群人重复进行囚徒困境博弈，并假设博弈将以如上所述的概率进行下去。我们将证明，如果博弈继续重复进行下去的概率足够高，那么理性的博弈参与者会在这种情况下合作。

与前面那个模型不同，在这里我们假设博弈参与者直接应用特定的行为规则。有些博弈参与者可能会采用冷酷触发策略，有些博弈参与者

可能始终合作，而另一些博弈参与者则可能始终背叛。这些策略的某些变体甚至可以在人类之外的其他物种身上看到。雄鸣鸟（Warbler）可能会采取“爱自己的敌人”策略，它们不会大声唱歌，也不会以牺牲邻居为代价来扩大领地。我们可以将这种行为视为一种合作行为。⁸

为了便于解释，我们假设每个个体都与其他人一起博弈。在完成了自己所有的博弈之后，每个人都要公布自己的“成绩”，也就是在博弈中的平均收益。之所以要使用每一场博弈的平均收益而不是所有博弈的总收益来考量，是因为给定博弈以一定概率继续进行，有些博弈参与者（出于偶然）参加的博弈可能会比其他博弈参与者多一些。在这个模型设定中，策略的效能取决于策略的分布。因此，胜出的策略也可能取决于初始分布。如果合作策略在一开始的时候表现是最好的，那么种群中的合作者数量就可能会增加。

在这个例子中，我们随机地向每个博弈参与者分配如下五种遵循行为规则的策略中的一种：始终合作（C）、始终背叛（D）、冷酷触发（GRIM）、针锋相对（TFT）、欺负好人（TROLL）。冷酷触发策略是一开始选择合作，后面也一直继续合作，直到另一个博弈参与者背叛为止，然后就一直背叛。始终合作和始终背叛这两种策略与名字的含义一样：盲目地选择合作或背叛，无论其他博弈参与者的行为如何。针锋相对（或一报还一报）是指在第一次合作，然后每一次都复制另一个博弈参与者在前一次中的行为，两个人都使用针锋相对策略的博弈参与者将永远合作。欺负好人策略则剥削始终合作的博弈参与者，更具体地说，这种策略是，在前两次选择背叛，如果另一个博弈参与者在这两次都没有背叛过，那么就选择永远背叛；而如果另一个博弈参与者在前两次已经背叛过了，那么就先转而在接下来的两次选择合作，然后一直采用冷酷触发策略。

我们首先根据如图22-1所示的囚徒困境博弈的收益矩阵，将每种行为规则策略面对其他每一种策略时的收益计算出来。我们先计算始终背

叛这个策略在面对各种策略时的收益。如果面对的是始终合作这个策略，那么它在每一轮博弈中都可以得到的收益为4。与此对应，在这些博弈中，始终合作所能得到的平均收益则为1。如果始终背叛策略“对阵”的是针锋相对策略或冷酷触发策略，那么它在第一轮博弈中获得的收益为4，之后每一轮博弈中都获得2的收益。

如果我们假设博弈会重复多次，那么所有轮次博弈的平均收益将只能略超过2，我们将它记为2⁺。而如果始终背叛策略与欺负好人策略相遇，前两轮博弈双方都背叛，然后欺负好人策略在第三轮和第四轮博弈中合作，但是此后一直背叛。因而始终背叛策略还是可以获得2⁺的平均收益；同时欺负好人策略的平均收益则略低于2，我们将它记为2⁻。

与此类似，我们可以计算出每一对策略的预期收益。⁹表22-1显示了每一个策略对所有其他策略的收益。

表22-1 行策略对列策略的平均收益

博弈参与者的策略	对方的策略				
	始终合作	始终背叛	针锋相对	冷酷触发	欺负好人
始终合作	3	1	3	3	1
始终背叛	4	2	2 ⁺	2 ⁺	2 ⁺
针锋相对	3	2	3	3	3 ⁻
冷酷触发	3	2	3	3	2 ⁺
欺负好人	4	2	3 ⁻	2	3 ⁻

表22-1显示了相互合作、相互背叛和利用其他策略中的缺陷策略的各种情况。仔细研究这张表可以发现，这五种策略中其实有四种是在与自己合作的，因此我们可以将这些策略视为潜在的合作策略。只有针锋相对这一种策略是所有这四种潜在的合作策略都能合作的策略。因此，如果这四种策略的任何一个组合在人口中占了大部分，那么针锋相对策

略就能够表现得非常好，尽管不一定总是最好的。¹⁰

研究者以人为被试，进行了数千次实验，结果充分揭示了人们所选择的策略的巨大异质性。我们将使用表22-1中的收益来考虑给定不同分布时的结果。基于这些策略的不同组合的收益多样性，最优策略将取决于人口的构成。在主要由总是选择合作策略的人组成的人群中，始终背叛这个策略的表现最佳。如果个人选择采用这个最优策略，或者自然选择的作用发挥得非常快，那么人们可能永远无法合作。如果学习或选择以适中的速度发生，那么博弈参与者也会逐渐远离始终合作策略。然而，一旦人口中只包含了很少的采用始终合作策略的人，那么始终背叛策略的表现将不如冷酷触发策略、欺负好人策略和针锋相对策略。这时，这三种策略中的某一种策略将会在人群中扎下根来。无论是在以人类为被试的实验中，还是在计算机仿真实验中，都会发现这种模式的广泛存在：一开始背叛策略的表现很好，但是不久之后，合作也能扎下根来。我们可以把这些情况下发生的这种事情，称为合作的涌现和合作的演化。

我们不难想象出这五种策略或任何其他策略集上的某种分布，再计算出该分布的平均收益，然后思考通过学习或自然选择接下来可能会发生什么事情。在本书后面的章节中，我们构建了一些关于学习和（自然）选择的正式模型。在这里，我们只是非正式地提出这样的观点，因为我们的目的只是指出合作是否能够出现、取决于种群中的初始战略分布以及人们如何学习或发展新策略。

合作出现或发展的一个必要条件是，合作带来的收益超过了背叛者能够获得的收益。否则，选择和学习都会导致整个种群趋向背叛。为了简化分析，不妨想象如下这个由采取冷酷触发策略、始终合作策略和针锋相对等合作策略的人，以及采取始终背叛策略的人所组成的种群。然后，我们可以计算出，要想让合作策略平均来说表现得更好，必须具备什么条件。这个计算表明，不断发展的合作比简单地维持合作更加困

难，而且合作是无法自我引导的，少数合作者无法促成合作的出现。¹¹

合作的维持、合作的出现和合作的不断发展，以及合作的自我引导之间的区别值得再三审视。如果当所有参与者都合作时，合作的表现是最好的，那么合作是自我维持的。合作的维持所对应的情况是，通过冷酷触发策略实现的合作是重复博弈的纳什均衡。如果在种群中配对时，合作策略的平均表现优于那些不合作的策略，合作就能够出现或发展起来。

正如刚才已经指出过的，合作出现的条件要比维持合作的条件更难满足。事实上，数学推理告诉我们，以自我引导的方式让合作出现几乎不可能。如果合作者的比例接近于零，合作者的收益就会低于背叛者。这样说并不意味着合作的自我引导永远不会发生，而只是在这个模型中不会发生。为了实现合作，我们需要一部分人从一开始就是合作的。这种情况有可能发生在那些会反思博弈结构的人身上，但是似乎不太可能发生在蜜蜂和树根“身上”。要想理解这种自我引导怎样才能发生，我们需要一些更加精细的模型，以允许局部学习、进化和群体选择。下面就来讨论这些模型。

合作行动模型

为了研究合作怎样才能实现自我引导，在这里引入一个合作行动模型（cooperative action model）。在这个模型中，人们可以采取合作行动，也可以不采取合作行动。¹²合作行动要求个人承担一定成本，会给他人带来收益。在这个模型中，聚类 and 群体选择都可以产生合作。

合作行动模型与重复囚徒困境博弈之间存在着一些差异。首先，在合作行动模型中，个人并不是两两配对重复进行博弈并在博弈中使用策略、获得收益的，相反，个人要么是合作者，要么不是合作者。其次，合作行动模型不假设理性行为者，也不假设个体会采用更复杂的规则。

再次，这个模型中的个体属于一个交互网络。他们的合作行动只会影响与他们有联系的人，也就是他们的邻居。最后，因为个体有固定的类型，所以他们会对所有邻居都采取相同的行动。例如，一个有五个邻居的合作类型的个体，要付出五次合作的成本，并且为另外五个人带来收益。

合作行动模型

一个种群由 N 个人组成，他们或者是合作者，或者是背叛者，连接于一个网络中。在每一次互动中，合作者都要承担合作成本 C ，而其他人则可以获得合作收益 B 。背叛则不会产生任何成本和收益。合作优势比率 B/C 刻画了合作的潜在收益。

在这个模型中，网络发挥了关键的作用。网络的存在，使合作得以出现，甚至可以实现合作的自我引导。一个主要在内部成员之间进行互动的合作者团体或合作者群组会有很好的表现，能够使合作在种群中扩展开来。在生态系统中，后代通常位于父母附近。如果合作者的后代更有可能成为合作者，那么合作的自我引导将会变得更加容易。

为了证明聚类可以导致合作的自我引导，我们从一个只有一部分已经被“充满”的网络开始。这个网络上的每个节点都是一个人可以“居住的住处”。在生物学背景下，这种“住处”就是生物的可行栖息地。然后，我们让合作者或背叛者“住进”网络的一部分。例如，可以先绘制出一个平均度数为10的随机网络，然后在每个节点上掷骰子，如果掷骰子掷出了“6”，就在那个节点上放一个人进去。如果没有掷出“6”，就将这个节点留空。如果我们已经决定要在一个节点上放一个人，那就再掷一次骰子。如果掷出了“5”，就在那个节点上放一个合作者，否则，就放一个背叛者。这个过程结束后，网络中 $1/6$ 的节点将会被人占有，而且在这些被占有的节点中，只有 $1/6$ 是被合作者占有的。

鉴于这个网络结构，每个人的邻居数量将会有所不同，有些人没有邻居，有些人会有四五个邻居。为了在这个网络上实现合作的增长或消亡，我们通过迭代地填充与被占用节点相邻节点的方法来填充网络的其余部分。假设填充空节点的人的类型将与这个节点的邻居中表现最好的那种类型相同（合作者或叛逃者）。图22-3给出了一个线性网络的两个片段，合作者用黑色线条表示，背叛者用灰色线条表示，黑色虚线则表示空节点。每个片段都在中心处包含了一个空节点，它有两个邻居、一个背叛者和一个合作者。在这图22-3中，合作创造的收益为2，而发生的成本则为1。

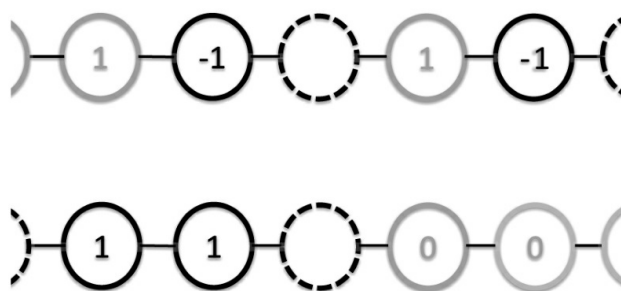


图22-3 两个线性网络中的一个空节点的邻居的收益

在图22-3的上图中，空节点右边的背叛者有一个合作者的邻居，因此可以获得1的收益。空节点左边的合作者有一个背叛者的邻居，因此可以获得-1的收益。根据规则，由于空节点的邻居中，背叛者获得的收益更高，所以这个空节点将由一个背叛者占据。在图22-3的下图中，空节点的背叛者的邻居的邻居也是背叛者，同时，空节点的合作者的邻居则连接到了另一个合作者。在这种情况下，空节点的邻居中以合作者的表现更好，因此空节点将成为合作者。

这个例子表明，一个单独的合作者不能产生一个额外的合作者，但是两个相邻的合作者可以。这就是说，一个小小的合作聚类就可以将合作扩展到空单元上。因此，合作区域可以从少数几个合作者中产生。

我们可以根据相邻的合作者和背叛者的比例以及合作优势比例，写

出决定空单元是会成为合作者还是会成为背叛者的更一般的条件。因此，在度数更低的网络中，合作的自我引导更加容易实现。这个发现与我们在分析声誉机制如何维持合作时所得到的结果相反，在那种情况下，更多的连接网络会增加背叛行为破坏某人声誉的可能性，因此更多的连接有助于合作的维持。这也是多模型思维能够产生依赖于特定条件的知识的又一个很好的例子。连通性高的网络能够产生更大的合作还是更少的合作，这个问题没有单一的答案。如果合作有赖于运用声誉机制的老练的行为者维持，那么连接更多的网络将更有利于合作。如果合作是在不成熟的行为者（如树木或蚂蚁）中自我引导或演化的，那么连接较少的网络应该更能促进合作。

聚类自我引导合作

如果一个空节点的邻居包括了一个合作者（其度数为 D 且有 K 个作为合作者的邻居），同时这个空节点的所有非合作者邻居都没有合作者的邻居，那么这个空节点会成为一个合作者，当且仅当合作优势比例高于与合作者数量之间的比例时，即：[13](#)

$$\frac{B}{C} \geq \frac{D}{K}$$

群体选择

我们要讨论的最后一个自我引导、发展和维持合作的机制是群体选择。这个机制依赖于群体之间的竞争或选择。[14](#)为了构建群体选择模型，我们将种群进一步划分为若干个群体。在每个群体内，个人的行动满足某种形式的合作行动模型——每个人或者选择合作或者背叛。与以前一样，我们可以认为每个人都有各自的表现。我们还为每个群体分配一个表现，它等于该群体成员的平均表现。群体选择模型假设选择是在

群体与群体之间进行的，表现最佳的群体的复制体（副本）将替换表现较低的群体。这种选择有利于合作者组成的群体，它们的表现将会更好。

然而，合作者组成的群体在群体选择时会占优势，这个直觉结论无法回避这样一问题：在任何一个群体内部，背叛者都比合作者有优势。作为例子，不妨考虑两个规模均为10人的群体：第一个群体包含了两名合作者和8名背叛者，第二个群体包含了两名背叛者和8名合作者。如上所述，假设收益等于2且成本等于1。在第一个群体中，每个背叛者的绩效等于4，因为他可以从每名合作者那里获得2的好处；每名合作者的成本为9，并且只获得2的收益，因此其绩效等于-7。第一个群体所有成员的平均绩效等于1.8。在第二个群体中，每个背叛者从8名合作者中的每一个中获得2名，因此其绩效等于16，每名合作者的绩效等于5，因为他从其他7名合作者那里得到14，但是支付9的成本。第二个群体的平均绩效等于7.2。

这些计算结果揭示了一个矛盾：在每个群体内部，背叛者对合作者有优势，但是表现更好的群体却必须包含更多的合作者。这里的张力是非常明显的：个体选择有利于背叛，但是群体选择却有利于合作。这种张力在各种各样的生态、社会、政治和经济环境下都会出现。例如，让自己的根系与其他树木合作的树木，个体生存条件可能会变得更加糟糕，但是这种合作有助于形成一个更加强大的生态系统，并使之更快速、更有效地扩散到更大的土地上。在一个社区内，合作的个人可以获得的收益少于背叛者，但是合作的社区规模将会扩大。支持自己所属政党的政治家可能比那些只专注于个人支持率的政客更加不容易再次当选，但是凝聚力更强的政党将更有可能发展壮大。在某家企业就职的人如果只专注于学习掌握与本企业有关的知识和技能，对自己可能不利，但是他所属的企业则可能胜过其他企业。

合作行动模型能够帮助我们识别和量化这种张力。为了确定群体选

择能否引导、发展或维持合作，还需要往模型中加入更多的细节。为此，特劳森（Traulson）和诺瓦克（Nowak）构建了一个精致的模型。在他们的模型中，种群的人口会增长，而且新出现的成员会复制表现最好的成员的类型。这个模型内置了个体选择和群体选择。选择发生在个体层面，同时表现更好的人更有可能来自合作的群体。当一个群体变得足够大时，它会一分为二，创建出一个新的群体。为了防止种群人口过多，新群体的形成会随机地导致现有的某个群体消失。这最后一个特征引入了一种较弱形式的群体选择。¹⁵

这些模型证明，群体选择能够增进合作，条件是合作行动的利益相对较大，同时最大群体的规模相对于群体的数量来说比较小。群体选择的效力部分取决于最大群体的规模与群体数量之间的比，这个结果揭示了竞争的必要性。有更多的群体，意味着全部由合作者组成的群体更有可能出现，它也隐含地假设了更多的竞争。最意想不到的一个结果是，最大群体的规模越小，导致的合作更多。较小的最大群体规模可以防止合作者组成的群体被背叛者所支配，也就是说，这限制了个人选择的影响。回想一下在前面举的有8名合作者和两名背叛者的群体的例子中，背叛者的表现更好。如果允许群体的规模扩大为80，那么在群体“分裂”出新的群体之前，它就会包含更大比例的背叛者。如果这个群体在有12名成员时就拆分为两个群体，那么在最坏的情况下，这个群体在拆分时也会包括2/3的合作者。

群体选择拥有增进合作的潜力，这个结论还可以应用到组织内部。大多数组织主要根据个人绩效来分配薪酬。将员工分成若干相互竞争的团队，并根据团队绩效分配奖金和机会，能够诱导合作行为的出现。如果资源流向团队，个人就有动力在这些团队中努力工作，即相互合作。¹⁶如果合作带来的好处很大，并且团队规模相对于团队数量来说很小，那么这种激励措施应该能够增进团队内部的合作。

但在评估群体选择的潜力时，我们必须仔细考虑个体行动者的复杂

情况。树木的适应速度非常缓慢，因此不必要求群体选择很快发挥作用。但是，人类的适应是非常迅速的，因此如果个人背叛的动机很高，那么相应地，群体选择就必须以很高的速度进行。而且，人们也可能会认识到群体选择效果，他们可能会考虑群体之间的竞争，并理解创建一个强大的群体需要符合个体的自身利益。这种认识会使合作更有可能实现。所有这些都表明，我们应该很小心，不要对特定模型中证明的能够促进合作的特定条件过于自信。恰恰相反，我们应该坚持多模型思维，善用判断力，追问定性结论是否仍然成立。

小结

合作如何出现、发展和维持，是来自众多学科领域成千上万的学者一直致力研究的一个难题。模型对这种研究很有帮助，其中最突出的模型是囚徒困境博弈模型。如果我们采用重复博弈框架，并假设行为者是理性的，这个合作难题会暂时消失。模型表明，可以利用进行惩罚的威胁来维持合作。惩罚可以通过重复博弈机制直接实施，也可以通过声誉机制间接实施。这些机制或许可以解释，在利害关系很大且博弈参与者是成熟老练的个体时，合作是怎样产生的；但是它们无法解释为什么蚂蚁、蜜蜂、树木和裸鼯鼠也会合作，而且合作程度是如此之高。当我们考虑了遵循规则的博弈参与者之间的合作时，我们发现要让合作不断发展并不那么容易。理性行为者可以在遵循规则的行为者无法发展合作的环境中维持合作。

我们还发现，像针锋相对这样的简单规则虽然不是最优的，但是可以实现相互合作而且不会被“剥削”。后续研究还表明，如果博弈中会出现随机误差，那么针锋相对策略的表现就不是很好。如果允许出现误差，那么每个使用针锋相对策略的博弈参与者都会在另一个博弈参与者出现误差之后产生一个背叛行动和合作行动组成的循环。如果两名博弈参与者都因误差意外采取了背叛行动，那么针锋相对策略将导致双方一

直背叛，直到另一个误差出现为止。

在现实世界的囚徒困境博弈中，这种误差确实会发生。1983年9月1日，韩国航空公司007航班从阿拉斯加安克雷奇起飞，前往韩国首尔，途中偏离航线进入了苏联领空。一架苏联SU-15战斗机击落了这架民航客机，机上269人全部不幸遇难。美国人认为这是苏联人的“背叛”行径，而苏联人则认为这架飞机在执行间谍任务，认为这是美国人的“背叛”行径。

其他的策略，例如“赢则坚持，输则改之”（Win Stay, Lose Shift）策略，在这样的情况下可以做得更好。在“赢则坚持，输则改之”策略下，相互合作时的收益和诱惑收益都编码为“赢”，另外两种收益则编码“失”。“赢则坚持，输则改之”策略从合作开始，此后，如果赢了，就坚持上一轮所做的一切；如果输了，就转变为另一种行动。只要考察一些例子，你就可以观察到“赢则坚持，输则改之”策略会回归到合作行为上。¹⁷

我们还描述了另外两种机制。聚类使合作能够实现自我引导。这种机制依赖于合作者之间的互动并通过选择来发展合作。群体选择发挥作用的原理也类似。合作者组成的群体表现得更好并取代了背叛者组成的群体。在正式的模型中，我们发现通过聚类和群体选择实现合作的条件要比通过重复博弈或声誉机制维持的合作更加严格。

我们还了解到，这些机制的成功与否，取决于我们如何对个人进行建模。我们不应期望这些机制对人类、蚂蚁和树木发挥作用的方式完全相同。更精明老练的行为者可能因为他们拥有前瞻能力而能够更好地维持合作，但是，当周围都是合作者时，他们也更有可能发现背叛的好处。

我们的大多数讨论都假设合作是有益的。但是，有些组织也可能通过合作来剥削他人。企业组成卡特尔（cartel）¹⁹，人为地压低价格；各产油国结成联盟，限制石油产量，以扩张自己的利益，而不管人类的利

益是否受到了损害。癌细胞会合作抵抗我们的免疫系统。¹⁸因此，当我们研究合作的时候，还应该记住，合作不一定是为了共同利益，例如野生水牛并不会受益于狮子之间的合作。

23 与集体行动有关的问题

自从大约5万年前智人发展出了现代创造力、高效和狩猎技能以来，如何可持续地管理环境资源一直是一个非常困难的问题。

贾雷德·戴蒙德（Jared Diamond）

在本章中，我们讨论集体行动问题，也就是那些自身利益与集体利益出现了不一致的情况。集体行动问题在各种情况下都会出现。在机场，乘客可能通过尽可能地挤到行李传送带旁边而获益，因为可以更快地拿到自己的行李，但是如果所有乘客都站到离行李传送带几步以外的地方，那么每一个人都会感觉更好。在美国，人们几乎没有任何动力去成为一个了解有关信息的选民，因为仅凭自己的一票改变选举结果的可能性非常低，尽管如果选民了解有关议题的话，国家的整体绩效将会更好。我们可以把集体行动问题视为一个多人囚徒困境博弈：每个人都有动机去背叛，但是从集体的角度来说，每个人都能够通过合作使自己的境遇得到改善。

学者们经常在历史案例的背景下研究集体行动，例如苏格兰公地的管理或纽芬兰和缅因州沿海龙虾栖息地的演变。¹其中最著名的案例是贾雷德·戴蒙德讲述的复活节岛上的波利尼西亚人的衰落。²复活节岛位于智利以西3 000多千米的南太平洋，方圆1 000多千米内没有其他可以供人居住的岛屿。这个特殊的地理位置决定了，复活节岛上的居民必须自己管理自己。据估计，到17世纪初，复活节岛上的人口已经超过了1.5万人。从16世纪开始，复活节岛上的居民积累了足够多的资源，调集了富余劳动力去建造一种名为莫埃（maoi）的巨型石头人像，它们往往重达80吨以上。

但是，复活节岛上的居民在忙于建造莫埃的同时，却没能共同合作

管理好他们的森林。到了1722年，当欧洲探险家第一次登上复活节岛时，岛上的食物已经相当匮乏了，人口也减少到了2 000人左右。岛上已经基本看不到超过3米高的树木了，许多鸟类和动物都已经灭绝。用戴蒙德的话来说，岛上的文明已经崩溃了。当欧洲人带来的病毒害死了岛上几乎所有剩余的原住民时，这个崩溃过程就最终完结了。

根据戴蒙德的解释，复活节岛上文明的崩溃，以及中美洲的玛雅人、美国西南部的阿纳萨齐人（Anasazi）和格陵兰岛的温兰人（Vinlander）的没落和崩溃，都是由于自然资源的过度开发和气候变化共同造成的，而资源的耗竭，又是因为制度和文化的失败所致。温兰人不但在极地边缘地带放养动物，而且还从原本就非常脆弱的草地上揭起草皮来建造房屋。结果在很短的时期内，他们的土地就因过度使用而变得极度贫瘠了，于是温兰人开始挨饿。与复活节岛上的居民一样，温兰人也未能管理好公共资源。他们砍掉了太多的树木、毁坏了太多的草皮，文明崩溃了。

虽然这些例子都很引人注目，也促人深思，但是由于它们都发生在“遥远”的历史时期，所以许多人可能据此认为，集体行动问题只是在久远的过去才是重要的问题。这是一种相当不幸的框架效应。事实上，由于世界已经变得更加复杂，各个部分之间的相互联系也更加紧密了，集体行动问题其实比以往更加重要。在人类社会的所有方面几乎都面临着集体行动问题。公共教育、身体和精神的保障、基础设施、公共安全、司法系统和国防支出，核心都是集体行动问题；管理全球渔业、应对气候变化，特别是减少对大气层的碳排放量，与集体行动问题就更加息息相关了。

此外，由于几乎所有工作都变得更加依赖于团队合作，也必然会产生集体行动问题。员工有很强的动机去“搭便车”——自己偷懒卸责，让别人去努力工作，他们也会对共享工作空间提出过度需求，以确保自己的团队有足够的工作空间。

本章的结构如下。首先定义一个一般的集体行动问题，然后具体分析三类重要的集体行动问题。我们的讨论是从公共物品的供应问题入手的，在这类问题中，个体要为道路、学校和社会服务提供资金，或者为清理公园或水源而付出时间和精力。然后，我们研究了拥挤问题，即个人对诸如道路系统、海滩或公园之类的公共资源的使用必须受到限制。最后，我们分析了可再生资源的开采利用问题——个人需要消费这些可以再生的资源，例如鱼，龙虾和树木。拥挤问题每天都会发生。如果太多的汽车塞满了城市的街道，那么该市或许可以通过向进入城市的汽车征收拥堵费来解决这个问题，因此以往的过度使用不会产生长期影响。然而，森林被过度采伐了、渔场被过度捕捞了，却至少需要数十年时间才能恢复元气。这就是说，我们不得不承担过去合作失败的恶果。

个体激励与集体目标之间不一致的性质各不相同，因此问题的解决方案也不同。我们可以通过征税来解决公共产品供应问题，在某些情况下，也可以通过分类来解决这个问题。拥挤问题则可以通过收费或实施使用限制来解决，解决可再生资源问题则需要监督、制裁，还需要建立健全解决冲突的机制。

说到底，我们在这里给出的解决方案只是一些最基本的见解，它们必须因地制宜地应用于各种具体情况。任何真实情况都包括了模型无法完全覆盖的复杂层面。巴厘岛的水神庙能够解决水资源分配问题。水资源的分配是一个序列拥挤问题：上游的人可以先利用水资源，然后是中游的人，最后才是下游的人。国际捕鱼权制度通过限制进入的方法解决了可再生资源的公地悲剧问题，不然的话，挪威限制本国渔民近海捕捞的努力，可能因瑞典、俄罗斯和丹麦等国渔民在近海的过度捕捞而无效。³在现实世界中，任何一个解决方案的有效性，都部分依赖于我们在第22章中讨论过的关于如何在囚徒困境博弈中实现合作的机制：重复行动、声誉、网络结构和群体选择等。群体选择是间接发挥作用的：成功解决了这些问题的社区和国家将繁荣发展，它们的成功会被其他人所复制。

集体行动问题

在集体行动问题中，每个人都可以在做贡献与免费搭便车之间进行选择。搭便车符合个人利益最大化动机，因为这能为个人带来更高的收益。然而，当每个人都做出贡献时，整个群体能够获得更大的收益。

集体行动问题

在集体行动问题中， N 个人中的每个人都要选择是搭便车（ f ）还是为集体行动做贡献（ c ）。个人的收益取决于自己的行动和合作者的总数。个人可以通过搭便车获得更高的收益，即，收益 $(f, C) > \text{收益}(c, C+1)$ ，但是当每个人都做出贡献时，所有人的收益总和实现最大化。

我们可以把集体行动问题视为一个多人囚徒困境博弈。因此，我们可以参考第22章中提出的解决方案，来解决如何形成合作和维持合作的问题。然而，那些方法是不完整的，原因如下：第一，集体行动问题涉及群体和社区，而不仅仅涉及配对博弈的个体；第二，许多集体行动问题都有特定的形式，因此其解决方案必须量身定制。

集体行动问题一：公共物品供应问题

我们要讨论的第一个具体的集体行动问题与公共物品的供应有关。公共物品满足非竞争性（non-rivalry），也就是一个人对公共物品的使用不会影响到其他人的使用，以及非排他性（non-excludability），也就是无法禁止个人使用公共物品。公共物品包括清洁的空气、国防、龙卷风到来之前的警报信号和知识。美国宪法列举了政府有建立司法体系、确保国内安全、防御外国侵略的责任，这些也属于公共产品。

私人物品，比如自行车、燕麦饼干和量角器，既不是非竞争性的，也不是非排他性的。但知识既是非竞争性的，也是非排他性的。比较一下燕麦饼干和三角函数知识，就可以将这种差异突显出来。老师可能会说：“梅丽莎，我很抱歉，卡拉已经把最后一块燕麦饼干吃掉啦。”但他永远不可能说：“梅丽莎，我很抱歉，卡拉已经用过了毕达哥拉斯定理，现在它永远消失啦。”

公共物品的非排他性和非竞争性导致了集体行动问题。这个问题的出现，不是因为人们不想做出贡献，而是在于人们低估了贡献的价值。每个人贡献出来的每一美元，都可以增加每个人的效用。我们在这里将给出一个正式的模型：每个人都要将自己的收入在一种公共物品和一种代表性的私人物品之间进行分配，可以把私人物品想象成可以花在任何其他东西上的钱。我们可以将这个模型扩展到为包括多种公共物品和多种私人物品的情形，但那只会使分析复杂化，并没有特别大的意义。

公共物品供应问题

有 N 个人，每人要将自己的收入 I ($I > N$) 配置到一种公共物品 (PUBLIC) 和一种私人物品 (PRIVATE) 上，每个单位的成本为1美元。每个人都有以下形式的效用函数：

$$\text{效用}(\text{PUBLIC}, \text{PRIVATE}) = 2\sqrt{\text{PUBLIC}} + \text{PRIVATE}$$

社会最优配置：PUBLIC= N （如果 $N=100$ ，那么每个人捐献100美元）。

均衡配置：PUBLIC= $1/N$ （如果 $N=100$ ，每个人捐献0.01美元）。⁴

在这里，我们把效用模型描述为公共物品的凹函数和私人物品的线性函数。对于这样的假设，我们必须说明其理由。凹性对应于收益递减：一个人对某种东西估价随着对它消费的增多而减少。效用对公共物

品数量的凹性，意味着公共物品的边际收益递减。这是一个标准假设。人们因高速公路新增的第三条车道而获得的好处，大于因新增第四条车道而获得的好处。污染严重的空气变得清洁对人们的好处要比将清洁的空气中最后一点微尘彻底清理干净更大。

之所以假设私人物品的效用是线性的，是因为它其实代表着所有私人物品的组合。虽然实用函数对任何一种私人物品可能都是凹性的，无论是巧克力、电视机还是牛仔夹克，但是它对于所有商品来说则更可能接近于线性。这个假设还有一个额外的优势，那就是，会使模型更加容易分析。

先求解社会最优配置问题。我们把社会最优配置定义为能够最大化整个种群效用总和的配制，也就是大多数人的最大幸福。⁵社会最优配置要求，每个人都得为种群中的每个人捐献一美元用于提供公共物品。这里需要注意的是，每个人对公共物品的贡献随着种群规模的增大而增加。这个结果不依赖于我们选择的效用函数的形式，因为种群规模越大，能够因非竞争性的公共物品而受益的人越多。这也就是说，享受清洁空气或国防安全的人越多，应该提供的公共物品就越多。

均衡贡献等于1除以种群总人数。随着人口的增多，人们有更大的动机去搭其他人贡献的便车。为了说明个中缘由，我们可以看看让人口规模增大1（加入一个人）时会发生什么。这个“新”人从公共物品中获得的效用，与之前的其他人一样。如果其他人的贡献保持不变，那么这个“新”人为公共物品捐献的动力就会比其他人在以往贡献时更弱。因此，他的贡献会比他们更少。另一方面，当他真的有所贡献时，他将使公共物品的数量有所增加，从而又增强了其他人贡献更少的动机。

因此，这个模型表明，随着人口规模的扩大，公共物品供给（不足）问题也会加剧。公共物品的最优水平提高了，但是人们贡献的动力却下降了。从模型中推导出来的最优贡献水平和均衡贡献水平确实依赖于对效用函数形式的假设，但是公共物品供给不足的现象确实是非常普

遍的。

这种分析假设是人都是自私自利的，这也是经济模型中的常见假设。不过，来自问卷调查、实验，以及日常观察的证据表明，人们还有涉他偏好（other-regarding preferences），会考虑他人的利益。人们希望他人也能够与自己一样，上好的学校、利用完善的道路网络。我们可以在模型中加入一个利他主义参数来模拟这种行为，该参数取零值时就对应于经济学家所假设的自利理性行为者，取值1则对应于每个人为他人考虑的程度与为自己考虑的程度完全一样。纯粹的利他主义者，也就是对每个人都一样关心的人，做出的贡献达到了社会最优水平。只要达不到纯粹的利他主义，就会导致供应不足。

计算表明，在规模很大的种群中，人们的贡献水平与最优配置水平之间的比例，近似地等于利他主义参数的平方。虽然公共物品供给不足的程度取决于效用函数的形式，但是这个例子说明了利他主义的局限性。关心他人的程度相当于关心自己的程度一半的那些人，贡献水平只相当于最优水平的1/4。关心他人的程度相当于关心自己的程度一半的那些人，贡献水平则只占最优水平的1/9。

利他主义者提供的公共物品

N 人有利他主义偏好，其对总效用的权重为 α ：

$$(1 - \alpha) \times \text{Utility}_j(\text{PUBLIC}, \text{PRIVATE}) + \alpha \times \sum_{i=1}^N \text{Utility}_i(\text{PUBLIC}, \text{PRIVATE})$$

均衡纯粹的利他主义者（ $\alpha=1$ ）：PUBLIC= N

均衡一般解： $\text{PUBLIC} = \frac{[(1-\alpha) + \alpha N]^2}{N}$

实例： $\alpha=1/2$ ：PUBLIC $\approx N/4$

考虑到我们并不生活在一个由纯粹的利他主义者组成的世界中，我们必须寻找其他机制来解决公共物品供给不足的问题，例如税收。政府征税以支付国防、公共道路、教育和刑事司法体系的支出，并提供其他公共物品。不过，确定税额需要包括个人收入和偏好异质性的更精细的模型。人们可以就税额和税率进行投票。空间投票模型预测税率等于中间选民所偏好的公共物品的水平。但是，如果人们有异质性的收入和偏好，这个水平就可能不是社会最优的。

许多公共物品，如学校、道路网络和资源回收系统，可以认为是属于“本地”的。本地社区可以将本社区之外的人排除在这些公共物品的使用者之外；但是在社区内部，这些公共物品仍然是非竞争性的和非排他性的。对于这种本地公共产品，人们可以根据自己的偏好将它们归入不同的社区。这就是所谓的蒂布特模型（Tiebout model），它为公共物品的供应问题提供了一个可能的解决方案。想要更好的学校、公园、游泳池和公共安全的人，可以投票支持征收更高的税收，以支付提供这些公共物品的费用。而那些不想要这些本地公共产品的人，则可以住进另一个社区，并只需要支付较低的税。

当然，蒂布特模型也不是万能的，它也不可避免地带来了一些成本，其中之一就是社会凝聚力会因此下降。此外，如果高收入人群以这种方式将自己与社会其他群体隔离开来，就会导致较贫困社区的公共物品供应进一步下降，并减少可以传递信息和知识的网络互动。⁷

集体行动问题二：拥塞模型

第二个集体行动问题是拥塞模型，它刻画了道路、海滩、供排水系统等公共物品对个人的价值随着用户数量的增加而减少的现象。任何一个受过交通堵塞之苦的人，都明白这里所说的含义。道路通畅比严重堵塞能够带给我们更多的乐趣和效用。有人估计，美国每年的交通延误成本已经达到了1 000亿美元以上，尤其是洛杉矶和华盛顿特区，人们每

年堵在路上的时间超过了60小时。

拥塞模型假设，资源的总量是固定不变的。每一天，每个人都可以选择使用这种资源或不使用。使用资源的个人收益随着其他用户数量的增加而线性地减少。⁸模型中的斜率，即拥塞参数，描述了这种拥塞效应的严重程度。

拥塞模型

在 N 人中， M 人选择使用一种资源。其效用可以写成如下形式：

$$\text{效用}(M) = B - \theta \times M$$

B 表示最大收益， θ 是拥塞参数。其余 $(N-M)$ 人不使用这种资源，并且效用为零。⁹

$$\text{社会最优解： } M = \frac{B}{2\theta}, \text{ 且效用}(\frac{B}{2\theta}) = \frac{B}{2\theta}$$

$$\text{纳什均衡解： } M = \frac{B}{\theta}, \text{ 且效用}(\frac{B}{\theta}) = 0$$

在社会最优解中，使用这种资源的人的数量等于最大可能收益除以拥塞参数的两倍。这些发现与我们的直觉相符，使用一种资源的人的数量应该随着最大可能收益的提高而增加、随着拥塞效应的加剧而减少。在纳什均衡解中，使用这种资源的人的数量恰好是社会最优解下使用这种资源的人的数量的两倍。这时，拥塞变得如此严重，以至于没有人能够获得任何收益。这个结果依赖于不使用这种资源的效用为零的假设。这个发现有一个违反直觉的含义：建一个美丽的公园可能并不能为社区居民带来太大的效用。在均衡状态下，公园将会变得非常拥挤，以至于在公园休憩并不会比待在家里更好。

当一个模型产生的结果与常识相反时，就需要对结果细加思量。人

们肯定更希望拥有一个公园，因此这个模型必定是错的。首先，它之所以出错，是因为我们假设所有人都有相同的偏好。如果不同的人对公园的“享受”程度各不相同，那么在有些人无法因公园而获益的情况下，另一些人仍然可能会获得正效用。其次，这个模型假设公园总是会拥挤不堪，但那不可能是事实。再次，人们的替代选择可能是去海滩，而不是待在家里。因此，新公园的出现，可能会使海滩变得不那么拥挤。最后，人们能够从各种体验中获得效用。如果一个城市里有滑板公园、宠物公园和水上公园，那么人们可以在一定时间内从多种不同的体验中获益。

尽管存在这样一些问题，但是这个简单的拥塞模型的主要结果仍然具有一定的说服力。在繁忙时期，拥塞确实会上升到公园所产生的好处不会比任何其他活动更多的程度。拥塞是不可避免的，尽管不会像只有一个公园时那么严重。更何况，即便建成多个公园也不能保证人们能够在多个公园之间实现最优配置。在下面的例子中，有太多的人都挤到了一个更大的那个公园中。

除了创建更多的公园，社区还可以尝试其他解决方案，例如采取配给制或轮换进入制、进行抽签、收取入场费和扩大容量等。配给制的核心是给每个人或每个家庭分配一定数量的资源。这种解决方案适用于像水这样可分割的资源，但不太适用于公共道路。轮换是按时间分割资源的使用权。例如，为了减少空气污染，城市可以限定在某些日子里只允许车牌号为偶数或奇数的汽车上路行驶。但是很多资源，如受欢迎的公立学校，是不能配给或轮换的，在这种情况下，则可以通过抽签来配置资源。

存在多个拥塞性公共物品的情况

M 个人去公园1， $(N-M)$ 个人去公园2。相对而言，公园2更大更好，为了体现这一点，假设如下形式的效用函数：[10](#)

公园1：效用 $(M) = N - M$

公园2：效用 $(N - M) = 3N - 3 \times (N - M)$

社会最优解： $M = N/2$ ，创造出的总效用为 N^2 ；

纳什均衡解： $M = N/4$ ，创造出的总效为 $\frac{9}{16} N^2$

对于道路系统，收取通行费是一种普遍的解决方案。伦敦市会收取车辆进入中心城区的费用，世界各地的收费公路也是如此。收费其实是将资源配置给那些愿意为使用资源付出更多钱的那些人，但是这些人也许不是能够从资源中得到最高效用的人。新加坡同时使用了收费和限制进入的方法。每一年，新加坡都会拍卖固定数量的机动车许可证，许可证的有效期为10年，一份许可证的售价往往超过了一辆普通汽车的价格。为了减少高峰期的交通堵塞，新加坡和伦敦一样，还向进入中央商务区的车辆收费。在同等规模的城市中，新加坡的交通状况算得上相当不错，而且新加坡政府还通过上述方法筹集了大量资金，可以用它们进一步发展公共交通。

扩大道路通行能力，能不能结束拥堵？并不一定。当一个城市通过增加高速公路的车道来缓解交通问题时，会使高速公路附近的房子升值，从而形成一种正反馈回路：房子的数量会随之增加，从而使交通量进一步上升，进而需要更宽的道路。这种正反馈回路，与第18章所描述的系统动力学模型中的正反馈回路类似。

集体行动问题三：可再生资源开采模型

第三个集体行动问题是可再生资源开采模型。在这种模型中，个人所利用的资源是能够再生的。这种模型适用于森林、河流、草原和渔业资源。在这些情况下，未来可用资源的数量取决于现在使用的数量。如果使用得太多，资源就可能无法足够快地再生。资源可能无法快速再生

这个事实使可再生资源利用问题比公共物品问题或拥塞问题更加脆弱。如果一个城市在某一年内资金不足，无法提供足够的公共照明，那么它还可以在下一年想办法加以改进，并且不会因为这种错误产生长期不良影响。但是，如果人们过度捕捞或过度砍伐森林，就会付出持久的代价，因为要生产鱼就必须先有鱼，但并不需要用路灯来制作路灯。此外，可再生资源可能是必需品：食物、饮用水和保暖用的燃料。人们需要开采可再生资源才能生存下去。

可再生资源开采模型

令 $R(t)$ 表示第 t 期开始时的可再生资源数量，再令 $C(t)$ 表示第 t 期内耗用的资源的总量， g 表示资源的增长率。那么，第 $t+1$ 期的资源数量由以下差分方程给出：[11](#)

$$R(t+1) = (1+g) [R(t) - C(t)]$$

$$\text{均衡消费水平： } C^* = \frac{g}{(1+g)} R$$

可再生资源开采模型表明，资源的耗用水平有一个临界点。任何高于均衡开采率的资源开采率都会导致崩溃，这一点在上面的正式模型中可以看得很清楚。我们可以将资源总量视为一个“饼”，开采资源相当于从这个饼中咬掉了一口。由于资源是可再生的，它会产生与剩余资源数量成比例的一定资源。如果资源开采水平较低，资源将会增加；如果开采水平过高，就无法通过再生来弥补了。在这两者之间，存在着一个均衡：开采的资源与再生的资源恰恰相等。

如果开采水平超过了均衡水平，那么模型预测，资源总量将加速下降，最终导致非常突然的崩溃。缓慢的下降之后是急剧的下降，这是对那些管理难以准确测量的资源（例如鱼类资源）的人发出的一个严重警告。年度渔获量提供了关于鱼类资源总量的一丝线索，但是它们并不准确。因此，我们不应该感到惊讶：正如贾雷德·戴蒙德在他的关于社会

崩溃的书中所描述的那样，北大西洋的鳕鱼捕捞业在现代也出现了崩溃，就像温兰人在历史上所遭遇过的一样。鳕鱼捕捞业在北大西洋已经有500多年的历史了。最早来到加拿大海岸的英国探险家们讲述了一些关于鳕鱼的“神话”：把一个篮子抛进海，就可以提上一篮子鳕鱼上来；浅滩上聚集了大量鳕鱼，连小船都划不过去！然而到了1992年，加拿大政府不得不严令暂停鳕鱼捕捞。¹²

可再生资源开采模型假设，它们会以不变的速度增加。有了这种假设，我们就可以求出均衡开采水平。但是在现实世界中，可再生资源的增长率每年都在变化。对于牧场来说，增长率取决于温度和降水量；对于鱼类种群来说，增长率取决于可用食物的数量，而食物的数量又取决于气温变化或气候变化。

在另外两个模型中，这种变化不会产生长期后果。在有些年份，我们可能有比较充裕的公共物品，拥塞情况不会太严重；而在有些年份，公共物品供给不足，拥塞情况就会非常严重。这些变化确实会影响效用，但是影响不会比天气等不可避免的变化更长。但是在可再生资源开采的问题上，情况就完全不同了。如果行为不变，细微的变化就可能导致崩溃与丰裕之间的天壤之别。在图23-1中，我们假设平均再生率为25%、资源的数量为100个单位。给定这些假设，均衡开采水平等于每年20个单位。这幅图给出了介于20%至30%之间的随机抽取的多个增长率下的长期结果。这个模型还假定，资源的最高水平为150个单位。

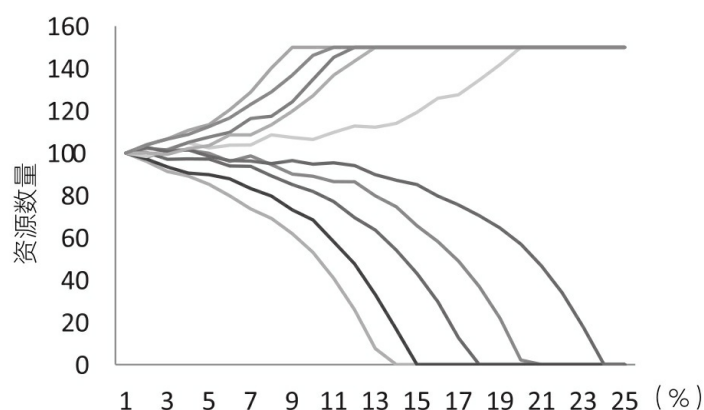


图23-1 资源增长率可变的10种可能路径

在大约一半的路径中，资源水平崩溃了。在另一半路径中，资源的水平增大到最大可能水平。这些变化是不能相互抵消的，恰恰相反，增长率下降或提高的影响会随着时间的推移而累积。¹³从这个模拟实验的结果中，我们可以看出，最优开采政策要求，当某一年出现了资源增长乏力的情况后，下一年必须减少开采，以防止崩溃。

由于可再生资源增长率的变化要求开采率跟随资源增长情况而变化，我们知道管理可再生资源必须做到随时调整开采水平。用来实现这种调整的方法或机制取决于资源的特征。俗话说得好，“世上没有灵丹妙药”。¹⁴不存在适用于任何情况下的解决方案。到底怎么解决这类问题，取决于资源和管理当地本身的特征。

例如，鱼与牛就不同。管理着一块共享公共草地的多个牛群的社群，可以监督每个牛群的放牧者的行为和资源的水平，即草的数量。过度放牧问题可以通过分配给每个牛群一定时间或一小块草地，或通过轮换放牧等方法来解决。而且，配给或轮换的具体形式，都可以根据牧草的情况随时加以调整。但是，对于一个捕鱼的社群来说，管理资源就需要更精细的机制来精确地监控个人的行为，因为海中鱼的数量是不可能准确计算的，只能根据捕获量来估算。因此渔业资源开采问题比养牛的不确定性程度更高。管理共同的水生资源，需要更严格的保护和更严密的监测。

集体行动问题的发生

在集体行动问题中，自利行为导致的结果与个人的目标是不一致的。如上所述，这类问题在各种各样的情况下都会发生：在为使用非竞争性和非排他性的公共物品付费时，在决定什么时候将车子开上高速公路时，等等。甚至在我们决定如何在高速公路上驾驶汽车时，也会出现

这类问题。在繁忙的高速公路上，如果发生了事故，有些驾驶员会自顾自地打电话、设路障，而全然不考虑他们这种行为对后面所有车辆造成的困扰。

这类问题也会发生在各个层面上。它们会出现在家庭内部：打扫房屋、做晚餐、购物，以及为外出度假存钱，所有这些都可能使个人动机与集体福利不一致。它们还会发生在社区、地区和国家等层面上，只要涉及公共产品的供给以及有限资源的使用和管理就会发生。它们还会发生在全球层面上，例如碳排放问题。大多数国家都希望全球排放水平更低，但是自己又想生产更多能源，也就是排放更多的碳，这时个体的理性行动与共同利益不一致。

集体行动问题在自然界也会发生。在森林里，树木会争夺光线和水分。如果某种树木演化出更高的树冠或更深的根系，那么它将增大自身的生存机会，但是会对其他的树种造成伤害。树木不可能“通过立法”去防止树冠长得过高或根系伸展太深，它们无法实现“社会”最优解决方案。¹⁵

当涉及的个人或参与者所属的群体越小、越同质化、信息越透明（采取行动更容易、系统状态可容易监测）时，集体行动问题往往更容易解决。家庭通常能顺利解决集体行动问题，但是国际组织却发现全球合作极其困难。要减少碳排放，需要大量不同的行动者之间的能力合作，而且它们所使用的监测机制是不精确的。解决这样的问题需要协调和执行机制。

历史告诉我们，过度捕捞或过度放牧，会导致可再生资源的崩溃。我们可以将同样的逻辑应用于今天面临的集体行动问题。美国著名政治经济学家埃莉诺·奥斯特罗姆（Elinor Ostrom）花了数十年时间研究现实世界中解决集体行动问题的各种机制。她发现，除了严密监督各种偏离行为之外，能够有效解决集体行动问题的那些社群有如下共同特点：能够就某些明确的界限达成一致、同意明确界定的规则、授权实施渐进式

制裁、拥有解决纠纷的机制。[16](#)

24 与机制设计有关的模型

制度的目的在于改变人类行为。为了保证制度不会随着时间流逝而归于无效，制度必须适应制度所要规范的环境或社会的变化。

珍娜·贝德纳

在本章中，我们讨论如何运用模型去设计政治制度和经济制度。任何一个制度都要包括两个因素：一是人们用来交流信息的渠道，二是人们用来根据所揭示的信息做出决策、重新配置资源或安排生产的程序。在市场中，个人和企业通过价格进行沟通以执行交易并做出生产决策。在层级机构中，人们通过书面语言进行交流以组织工作并推进计划。在民主国家，人们通过投票进行沟通。投票规则决定政策。良好的制度可以促进沟通和行动，从而产生理想的结果，效率低下的制度则相反。

在本章中，我们给出了一个通常称为机制设计（mechanism design）的对制度进行建模的框架。这个框架强调了真实制度的如下四个方面：信息，指参与者知道些什么及应该向他们揭示什么；激励，即采取特定行动的利益和成本；集结，个人行为如何转化为集体结果；计算量，这是对参与者认知能力的要求。

机制设计思想起源于分析资源配置的一般问题，即中央计划体制和市场机制究竟哪一个才能实现资源的最优配置。在早期的模型中，建模者会先给出若干行为规则，例如在市场中充当价格接受者，或者如实投票；然后再研究这些行为的含义，也就是看这些微观行为是怎样集结为集体结果的。后来，这种方法被放弃了，取而代之的是假设人们采取最优化行为的方法。这种假设使这些问题很适合运用博弈论的工具来分析。这就是机制设计的思路。在构建了博弈模型后，机制设计专家求解出纳什均衡，然后在共同的行为假设的基础上进行制度比较。

事实证明，这个分析框架是很有用的。它可以用于搜索现有规则和程序中的缺陷、解释制度成功或失败的原因，可以用来预测结果，还可以用来设计各种各样的制度，包括第2章中描述的频谱拍卖，以及在线交易体系、政府投票系统，甚至还可以用来设计为航天飞机上的研究项目分配空间的程序。¹

本章讨论的内容包括六个部分。在第一部分中，我们首先使用芒特-赖特尔图（Mount-Reiter diagram）描述了机制设计的框架。在第二部分中，我们研究了三个人在两个备选项之间进行选择的问题。在第三部分中，我们分析了三种拍卖机制，发现所有这些机制都产生相同的结果。在第四部分中，我们证明上述结果不是一个巧合，并描述了一个基本结果，即收入等价定理（revenue equivalence theorem）。收入等价定理的含义是，只要满足某些假设条件，任何拍卖机制都会产生相同的结果。在第五部分中，我们对多数投票平均分担机制（majority-vote equal sharing mechanism）与枢轴机制（pivot mechanism）进行了比较，它们都是决定是否进行某个公共项目建设的方法。最后，我们沿着在对纳什均衡进行批评时引入的思路，进一步扩展了对机制设计问题的分析。

芒特-赖特尔图

一种机制由六个部分组成：一个环境（世界的相关特征）、一个结果集、一个行动集（也称为消息空间），一个行为规则（人们根据这个规则来做出行动）、将行动映射到结果的结果函数，以及将环境映射到一组希望得到的结果的社会选择对应（social choice correspondence）。社会选择对应通常包括能够最大化参与者效用的结果或帕累托有效（pareto efficiency）配置的集合，当且仅当不存在每个人都喜欢的其他结果，某个结果是帕累托有效的。帕累托效率是下限标准。

帕累托有效

在一个结果集中，对于某个结果，如果存在每个人都喜欢的另一个备选方案，就说这个结果是帕累托占优的。相对应，所有其他结果都是帕累托有效的。²

芒特-赖特尔图以图形方式描述了一个机制的上述基本组成部分（见图24-1）。需要注意的是，芒特-赖特尔图将我们想要的东西和已有的东西并列在一起。图的顶部是社会选择对应，它描述了我们希望得到的规范的结果。在图的底部，则列出了我们在现实世界中能做的事情——人们应用他们的行为规则来发送消息或采取行动。结果函数将这些行动映射到结果中去。在理想情况下，下部这个更加复杂的路径会产生与上部路径相同的结果，即人们期望得到的结果。

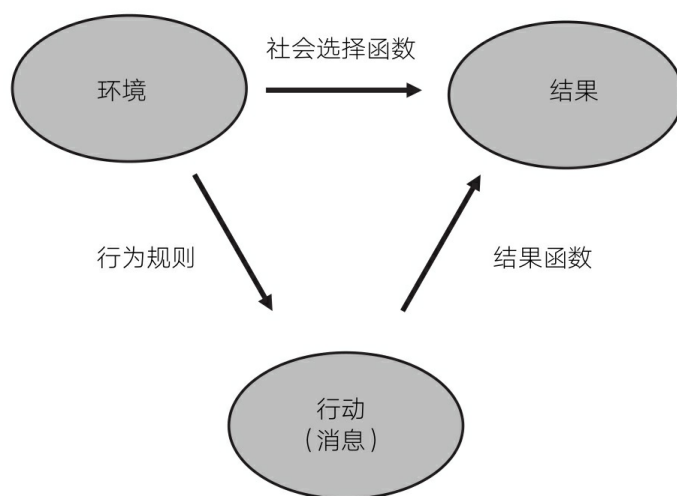


图24-1 芒特-赖特尔图

当然，并不是所有机制都能取得成功。例如，如果环境是由对某种公共物品有偏好的人组成的，那么社会选择对应会将他们的偏好映射到这种公共物品的最优水平上。然而，正如我们在第23章中已经看到的那样，自愿捐献机制，也就是人们按照自己的意愿支付公共物品费用的机制会导致每个人只提供 $1/N$ 单位公共物品而不是最优的 N 单位公共物品。当一种机制产生的结果与目标不一致时，就说这种机制未能实现社会选

择对应。

我们希望机制满足哪些性质？答案必须“因地制宜”。我们在这里只描述其中五个性质。第一，我们会希望这种机制的均衡结果与社会选择对应一致（帕累托有效）。第二，在理想情况下，参与者将会采用占优策略，即他们的最优行动不依赖于他人的行动。如果是这样，就说有效的结果是占优策略可实施的。第三，我们不想强迫人们参与这种机制（自愿参与）。第四，如果这种机制涉及资源的转移或支付，我们不希望增加额外成本或破坏资源（预算平衡）。在本章的后面，当我们分析决定某个公共项目的机制时，会发现这个性质往往很难满足。第五，在许多情况下，我们希望参与者讲真话。我们希望人们发送的消息能够揭示他们的真实信息或真实类型。博弈理论家把这个性质称为激励相容（incentive compatibility）。在大多数有趣的情况下，通常没有任何一个机制可以满足所有这些要求。因此，机制设计理论的一个重要贡献就在于证明什么是可能的、什么是不可能的。

多数规则和拥王者机制

我们要考虑的第一种情况是，人们如何通过投票决定是否采取联合行动或通过某项法案。假设，乌玛、维拉和威尔三个人想要一起去看电影，在出门前必须决定到底是看动作片、剧情片，还是喜剧片。类似的问题也会出现在军事决策中，比如三位高级军官，要决定到底是主动出击敌军、守卫阵地，还是割让土地。在这两个决策问题中，环境都由三个人组成，其偏好定义在三个备选项上。我们使用排序来表示他们偏好。排序动作片 > 喜剧片 > 剧情片，对应于最喜欢的是动作片，其次是喜剧片，然后是剧情片。我们假设以下的偏好排序：

乌玛：动作片 > 喜剧片 > 剧情片

维拉：喜剧片 > 剧情片 > 动作片

威尔：喜剧片 > 剧情片 > 动作片

在这个例子中，我们将社会选择对应设定为帕累托有效的选择集合。按照上面给出的偏好，喜剧片和动作片是帕累托有效的，而剧情片则被喜剧片占优。

我们首先评估作为一种机制的多数规则。在出现平局的情况下，我们假设选择是随机的。如果人们如实投票，那么喜剧片将会获得两票。然而，假设维拉和威尔都认为另外两个人将会在剧情片与动作片之间出现分歧，而且每个人都会投票给剧情片。再假设投票是连续的，维拉率先投票并选择了剧情片。威尔第二个投票，并同样选择了剧情片。这里，乌玛投什么票其实已经无关紧要了，但是为了避免冲突，她也选择了剧情片。这三张选票构成了一个纳什均衡，也就是说任何一个人都没有动力改变自己的投票。在这种情况下，多数规则并不总是能实现帕累托有效结果。

接下来考虑一下拥王者机制（kingmaker mechanism）。³在这个机制中，先在群体中随机地选择一个人当“拥王者”。然后，由“国王”选择一个“国王”来决定这个群体的选择。如果威尔成了“拥王者”，那么他必须在乌玛和维拉之间选出一个“国王”来。他选择的任何人都将成为“国王”并决定大家一起去看什么电影。

如果是一个理性行为者，那个被选为“拥王者”的人会选择自己最喜欢的电影，由此得到的结果将是帕累托有效的。因此，拥王者机制能够实现帕累托有效的结果。这个机制还有一个额外的优点，那就是，如果任何两个人最喜欢的电影是相同的，那么这个机制肯定会选择这个结果。逻辑如下。假设威尔是“拥王者”。如果乌玛和维拉最喜欢的电影类型相同，那么无论威尔选择谁当“国王”，都会选出这种电影来。另一方面，如果是威尔和乌玛最喜欢的电影类型相同，那么威尔应该选择乌玛当“国王”。

三种拍卖机制

现在我们已经对机制有了最基本的印象了。接下来，我们转向对拍卖的研究。由于像eBay这样的在线拍卖市场的普遍存在，大多数人对拍卖已经有一定的了解。当然，拍卖也广泛用于其他环境，包括政府采购、二手车市场和大多数网络广告。我们在这里将注意力限定在卖家只有一个而竞买人有很多个的情形。拍卖标的可以是房屋、汽车、足球比赛门票或艺术品。再假设每个竞买人赋予拍卖的价值都是唯一的，目的是排除出现平局的可能性。帕累托有效结果是指拍卖标的转移到了对它（们）估价最高的竞买人手中这种结果。任何其他结果，都是帕累托占优的结果。接下来，我们比较三种类型的拍卖：出价递增拍卖（ascending-bid）、第一价格拍卖和第二价格拍卖。

在出价递增拍卖中，拍卖师喊出一个价格，任何愿意支付这个价格的竞买人都会举手。然后拍卖师不断提高价格，直到只剩下一个竞买人为止。拍卖成交后，最后这个竞买人要支付的是出价第二高的竞买人的价格。在出价递增拍卖中，理性的竞买人会一直参与拍卖，直到价格达到他心目中的价值为止，而他会在那一点上退出拍卖。一方面，在价格达到竞买人的价值之前退出，就有可能无法以该价值的价格买下拍卖标的；另一方面，在价格超过竞买人的价值之后，仍然继续参与拍卖则会带来风险，这意味着竞买人或许可以买下该标的，但是支付的价格却超过了价值，从而导致净损失。

如果所有竞买人都理性行事，那么具有最高价值的竞买人赢得拍卖标的并支付相当于价值第二高的竞买人的价格。这一点可以通过下面的例子来说明。假设有三个竞买人，拍卖标的对他们的价值分别为30美元、60美元和80美元。当拍卖师喊出的价格超过了30美元时，第一个竞买人就会退出。当价格达到了60美元时，第二个竞买人没有理由继续提高出价。因此，第三个竞买人在这个拍卖中胜出，并付出60美元，或者也可能再多一点点。为了简化分析，我们假设第三个竞买人要付出60美

元。⁴

在第二价格拍卖中，每个竞买人都以密封的方式提交出价，其他所有竞买人都不知道这个竞买人的出价金额。拍卖标的由出价最高的竞买人获得，但是竞买人只需支付相当于第二高出价的金额。第二价格拍卖这种机制设计能够保证说实话是最优行动。想象如下这个拍卖的例子。一个对拍卖标的估价80美元的竞买人要决定在第二价格拍卖中如何出价。不妨假设其他竞买人已经提交了出价，那么这个竞买人必须考虑三种可能的情况：其他竞买人的最高出价可能低于80美元、可能等于80美元，也可能超过80美元。不管在哪一种情况下，竞买人如实出价都是最优行动。

通过一个数值实例，上述逻辑会变得更加清晰。我们先假设拍卖标的对这个竞买人的价值为80美元，再考虑其他竞买人提交的如下四种最高出价：70美元（明显低于其价值）、80美元（等于其价值）、85美元（略高于其价值），或90美元（明显高于其价值）。然后进一步假设这个竞买人考虑了四种出价方案，从65美元到95美元不等，如表24-1所示。

表24-1 当标的对竞买人的价格为80美元时，作为出价函数的净收益

其他竞买人的 最高出价	对拍卖标的估价为 80 美元的竞买人的出价			
	65 美元	80 美元	85 美元	95 美元
70 美元（明显低于）	0	+10	-10	+10
80 美元（等于）	0	0	0	0
82 美元（略高于）	0	0	-2	-2
90 美元（明显高于）	0	0	0	-10

从表24-1中可以看出，出价80美元总能带来与任何其他出价一样高

的收益。这就是说，按拍卖标的对自己的真正价值如实出价，始终是最优行动（占优策略）。同样的逻辑也适用于所有竞买人，因此所有人都应该按他们的真实价值出价，所以这种机制是激励相容的。因此，在第二价格拍卖中，具有最高价值的竞买人会在拍卖中胜出，并且支付的金额等于价值第二高的竞买人的价值。

在第一价格拍卖中，每个参与拍卖的竞买人都提交一个出价，最高出价者胜出，并且付出的金额就等于该出价。与第二价格拍卖一样，所有参与拍卖的竞买人都同时提交出价，因此没有人知道其他人的出价。参与拍卖的某个竞买人在第一价格拍卖中的最优出价策略，则取决于该竞买人对于其他竞买人的价值以及可能的出价的信念。我们一般假设，竞买人无法知悉其他竞买人的准确价值，但是他们确实拥有关于这些价值分布的正确信念。我们假设竞买人的价值均匀分布在0美元至100美元之间，并且所有竞买人都知道这种分布。竞买人也知道拍卖标的对他们自己的价值。

通过简单的微积分计算，我们可以证明，如果竞买人的出价是均匀分布的，并且如果所有竞买人都以最优方式出价，那么在只有两个竞买人的情况下，每人的出价都应该是自己真实价值的一半；在有 N 个竞买人的情况下，每个竞买人的出价应该是自己价值的 $N-1/N$ 。这也就是说，与其他19个人一起参加拍卖的竞买人应该按自己真实价值的95%出价。根据这个出价规则，具有最高价值的竞买人肯定能够买下拍卖标的。我们还可以证明，这个竞买人付出的金额等于价值第二高的竞买人的预期价值。因此，第一价格拍卖也能够产生帕累托有效的结果，且拍卖成交价格与对拍卖标估值第二高的竞买人的预期价值相对应。⁵

在写下这个模型之前，我们许多人就已经知道，参与拍卖的竞买人越多，任何一个竞买人的出价就应该越高。但是，不利用模型进行推导，我们就不能得出均衡出价规则。这里第一价格拍卖模型为我们提供了一个精确的公式。这个公式告诉我们，出价会随拍卖标的对竞买人价

值的增加而提高，而这就意味着具有最高价值的竞买人将赢得拍卖，就像在其他两种拍卖中一样。

收入等价定理

在三种拍卖机制的任何一种中，都是具有最高价值的竞买人在拍卖中获胜。因此，所有这三种拍卖机制都产生了有效的结果。此外，中标的竞买人付出的预期金额等于第二高出价者的价值。换句话说，所有这三种拍卖都导致相同的预期收入并将拍卖标的分配给了同一个竞买人。这一点非常引人注目。而且，更加值得注意的是，我们还可以证明，在竞买人采取最优行动、且最高出价者在拍卖中胜出，同时价值为零的竞买人不会得到任何收益的任何拍卖中，胜出者和预期收入都是相同的。换句话说，所有拍卖机制都会产生相同的预期结果，这个结论就是通常所称的收入等价定理。⁶

收入等价定理

在任何拍卖中，如果竞买人从已知的共同分布中抽取独立私人价值，那么只要满足如下条件，拍卖就必定会给卖方带来同样的收入、给买方产生同样的预期收益：每个竞买人的出价都是为了最大化自己的预期收益、最高出价的竞买人总能赢得拍卖标的，同时价值为零的竞买人的预期收益为零。

收入等价定理还意味着全支付拍卖（all-pay auction），在这种拍卖中，每一个竞买人，即便是未成功的竞买人，都要按自己的出价付款，这也会产生与第二价格拍卖相同的结果。⁷即便像第三价格拍卖这样“几近疯狂”的拍卖设计（最高出价者胜出并支付相当于第三高出价的金额），也会产生相同的胜出者和相同的收入。

当然，收入等价定理并不意味着拍卖规则是无关紧要的。在拍卖

中，竞买人可能不一定使用最优策略。或者，在第一次价格拍卖中，他们可能会对其他竞买人的价值分布有不同的信念。只要出现了任何一种情况——竞买人不进行最优化或存在不同的信念，那么各种拍卖的收入就可能会有所不同。经验证据和实验确实表明，理论与现实之间存在着一些差异。

但是，正如我们前面在讨论什么情况下可以预期行为者理性行事时所指出的，利害关系越大，竞买人越精明老练，他们理性行事的可能性就越高。在消费品的在线拍卖中，我们会预期一些人可能会遵循经验法则或者受认知偏差影响做出不理智的行动（例如每次以10美元增价竞拍）。但是，在价值千百万美元的石油租赁拍卖中，竞买人肯定会事先做好准备，了解全部可以掌握的信息，并学会必要的技能。

此外，拍卖形式本身也可能会影响竞买人的数量。在木材拍卖中，第一价格拍卖所能吸引的竞买人数量要比出价递增拍卖更多，因为如果大竞买人（大公司）提交的出价较低，小竞买人也有机会胜出。但是，在出价递增拍卖中，小竞买人是不会有机会胜出的，因为大竞买人可以观察到小竞买人的出价并给出更高的出价。⁸

不同的拍卖形式对参与者认知能力的要求也不同。在某些拍卖形式中，最优行为是很容易学习的。例如，在出价递增拍卖中，竞买人应该一直参与拍卖，直到价格达到自己的价值为止。其他未遵循最优策略的竞买人可能会导致采取最优策略的竞买人获得更高或更低的预期收益，但是他们还是不会改变最优策略：只要价格低于自己的价值，竞买人就应该继续参与拍卖。与此类似，在第二价格拍卖中，竞买人应该始终遵循同样的策略，按自己的真实价值出价。但是，要真正理解按自己的真实价值出价是最优策略，需要多个逻辑步骤。

请回想一下占优策略的定义：无论其他人采取什么策略，占优策略都是最优的。出价递增拍卖和第二价格拍卖中都存在占优策略。但是，第一价格拍卖则不存在占优策略。在第一价格拍卖中，一个竞买人出价

策略的变化可以改变另一个竞买人的最优策略。如果一个竞买人的出价总是为0或50，那么另一个竞买人就应该总是出价1或51，后者没有理由出价60或70。考虑到那个竞买人的出价行为，出价60肯定能够赢得拍卖标的，但是51的出价也已经足够了。

而且，即便某个拍卖中存在占优策略，也并不是所有的占优策略都同样容易推断。在出价递增拍卖中，只要价格低于竞买人的价值就一直继续参加拍卖，这个最优策略并不是一目了然的，而是至少需要通过一步推理才能得知的：如果价格低于你的价值，请以这个价格买下。在第二价格拍卖中，竞买人必须考虑过多种可能发生的情况才能看清楚按自己的真实价值出价是最优策略。当然，如果一个人参加过好几次第二价格拍卖，那么他应该能够学会在未来的拍卖中采取这种策略。

关于各种形式的拍卖，要考虑的最后一个问题是，拍卖是不是会鼓励非最优行为。在第一价格拍卖和第二价格拍卖中，竞买人要在不知道其他竞买人出价的情况下提交自己的出价。而在出价递增拍卖中，竞买人可以直接观察到价格一路上行并知道哪些对手仍然在参与拍卖。这可能会导致竞买人赋予“在拍卖中胜出”这个结果某些附加价值，或者甚至赋予“出价压过别人一头”这种行为某些附加价值。主持慈善拍卖的拍卖师试图通过诉诸情感来调动竞买人提高出价，例如展示儿童在用拍卖所得购买的新游乐设备上快乐嬉戏的视频。

不同的策略能否成功，还取决于竞买人的精明程度。我们无法想象，在木材拍卖会中，竞买人会被人引诱报出严重高于自己预测估值的出价。但是，在慈善拍卖会上出价过高却不算非常罕见。有人说，也许竞买人在投标过程中会改变自己对拍卖标的的估值，但这只是一个猜想，我们只需要承认它确实可能会发生。在第一价格拍卖和第二价格拍卖中，竞买人只有一个出价机会，因而无法在拍卖期间通过感情因素来影响他们的决定。

最后，在第一价格拍卖和出价递增拍卖中，价格等于最高的出价。

而第二价格拍卖中，价格等于第二高的出价。这就给人留下了这样的印象：在第二价格拍卖中，卖方本来应该可以获得更高的价格，并在一定程度上解释了为什么政府不愿意使用第二价格拍卖。想象这样一个例子，在美国政府组织的油田开采权的拍卖中，政府收到了三个投标，第一个是600万美元，第二个是800万美元，第三个是1 200万美元，如果采用第二价格拍卖，那么第二天报纸上的新闻标题就可能是：“政府得到1 200万美元的出价，但是却以800万美元的价格售出了油田。”那无疑是一场公关灾难。事实上，任何了解拍卖理论的人都知道，如果政府进行的是第一价格拍卖或出价递增拍卖，那么最高出价是不可能达到1 200万美元的，而只能达到800万美元。

正如我们在本书中一直反复强调的，正式的模型可以揭示特定结果成立所需的条件。等价定理并不意味着所有拍卖机制都会产生相同的结果。它只说明，在满足如下条件时，所有拍卖都是等价的：竞买人采取最优出价策略、最高出价者赢得拍卖标的，同时对拍卖标的估价为零的竞买人的预期收益为零。卖方可以通过放宽这三个假设中的任何一个来筹集更多资金。卖方很难让人们违背自身利益行事，而且也可能无法从认为拍卖标的完全没有价值的人那里榨取到钱财。

因此唯一剩下的可能性是，不将拍卖标的出售给出价最高者。而要做到这一点，最直接的一个方法根本不出售拍卖标的。如果卖方知道竞买人的价值分布，那么就可以设定一个保留价格或底价，即最低出价。在某些情况下，这可以增加卖方的预期收入。例如，假设卖方确定，某个拍卖标的对三个竞买人的价值分别为5美元、10美元和60美元。在上述三种拍卖形式中的任何一种中，胜出者的出价都是60美元，而且都只需支付10美元。在这种情况下，卖方可以规定，保留价格为60美元，然后再进行一个第一价格拍卖，从而获得更高的收入。

公共项目决策问题

接下来，我们比较一下用来决定启动建设一个公共项目的两种机制，如新学校、新高速公路或新体育场。我们假设，新项目会给每个人带来一定价值，同时也要每个人付出一定成本。

公共项目决策问题

$V_1, V_2, \dots, V_N$ 表示 N 个人赋予一个公共项目的货币价值，并假设该公共项目的成本为 C 。那么当且仅当 $C < V_1 + V_2 + \dots + V_N$ 时，这个项目才会启动。

我们首先考虑多数投票平均分担机制（majority-vote equal sharing mechanism）。在这个机制中，个人投票决定是否启动某个公共项目。如果多数人投了赞成票，那么项目启动，而且成本由所有人平均分摊。

多数投票平均分担机制

个人投票表示赞成或反对启动某个公共项目。如果多数人投票支持该项目，那么该项目启动，并且每个人都承担 C/N 的成本。如下面的例子所表明的，这种机制可能会违背效率条件和自愿参与原则。

从空间投票模型中可以看出，项目是否启动取决于中间选民的偏好。在我们现在讨论的情况下，中间选民就是指公共项目对他的价值位于中位数的人。根据定义，这个机制满足预算平衡条件和激励相容要求。但是，这个机制不一定能够满足效率条件和自愿参与原则。

假设有三个人的价值分别是0美元、120美元和150美元，而公共项目的成本是300美元。有效的结果是不应启动该项目，因为300美元的费用超过了个人价值的总和。然而，考虑到成本将平分，每个人将投票决定是否以每人100美元的成本进行该项目。因此，这三个人中有两个将

投票支持这个项目，而且这个项目将会启动，这是一个低效的结果。此外，价值为0美元的个人获得-100美元的回报，因此这个例子也违反了自愿参与原则。

接下来讨论公共项目决定的第二种机制，枢轴机制。在这种机制下，每个人都提交自己对公共项目的估值，如果估值总和超过了项目成本，就启动该公共项目，否则就不启动。同时，对某个人征税的金额等于项目成本减去所有其他个人估值的总和。如果其他个人的估值已经超过项目成本，这个人就不用支付任何费用。

枢轴机制

个人 i 对一个成本为 C 的项目提交自己的估值 $\hat{V}_i$ 。如果所有个人的值的总和超过了成本，那么就启动这个公共项目，即：

$$\hat{V}_{\Sigma} = \hat{V}_1 + \hat{V}_2 + \cdots + \hat{V}_N \geq C$$

如果 $C - (\hat{V}_{\Sigma} - \hat{V}_i) < 0$ ，那么个人 i 不用交税；如若不然，个人 i 就要缴纳数额为 $C - (\hat{V}_{\Sigma} - \hat{V}_i)$ 的税收。这个机制是激励相容的（ $\hat{V}_i = V_i$ ）、有效率的，而且个人行为也是符合理性。它还实现了占优策略的有效结果。但是，正如下面的例子所表明的，这个机制可能会违背预算平衡条件。

实例：（ V_1, V_2, V_3 ）=（60, 120, 150）， $C=300$ 。

这个公共项目本应启动，因为 $300 < 60 + 120 + 150$ 。个体1要缴纳的税收为30，即总成本减去其他人估值之和的结果（ $300 - 270$ ）；个体2要缴纳的税款为90；个体3要缴纳的税款为120。由此得到的总税收为240，低于项目的成本。

这个机制满足激励兼容条件，原理与第二价格拍卖类似。现再举例说明。假设一个公共项目的成本为300美元，而且某个人对这个公共项

目的价值估计为80美元。有三种情况需要考虑。如果其他人的估值的总和低于220美元，那么这个人没有动机提交超过80美元的估价，因为到时必须由他来支付该金额。如果在另一个极端上，其他人的估值总和超过了300美元，那么这个人什么都不用付出，他可以给出任何估价。但是，如果其他人的估值总和介于220美元到300美元之间，且这个人提交了80美元的估价，那么他要承担的成本将等于300美元减去那个总和的差，并且这个项目将启动（这是一个有效率的结果）。他将不会提交70美元的估值，因为其他人的估值总和可能是225美元，那么他提交的这个低估值将使这个公共项目无法启动。而如果他提交的估值为80美元，那么他所要承担的成本仅为75美元。

由于枢轴机制满足激励兼容性，因此它也满足有效性。只有在估值总和超过了成本的情况下，才能实施公共项目。需要注意的是，因为报告一个人的真实价值是一个占优策略，所以有效结果也是占优策略可实施的结果。此外，由于每个人最多支付项目对自己的价值，这个机制也满足自愿参与原则。但是，这个机制不一定会得到预算平衡的结果，事实上，这个机制只是在极少数情况下才能做到这一点。

对于决定公共项目的决策问题，任何机制都无法满足我们可能想要达到的所有标准。事实上，当用模型证明了这一点之后，就可以省下很多时间，不会再无谓地去尝试一些不可能实现的事情。正如工程师不会浪费时间去建造永动机一样，机制设计专家也不会去尝试为公共项目决策问题寻找一个满足激励相容、个体理性、有效率且预算平衡的机制。事实上根本不存在这样的机制。

枢轴机制已经相当不错了，但是它不能满足预算平衡条件。而且，这个缺陷无法通过提高人们为项目缴纳的税额来解决，因为那样做会使这个机制不再是激励兼容的或个体理性的。那样的话，个体会有动机去撒谎，有些人可能会被要求为项目贡献出超过其价值的东西。一种可能的解决方法是通过其他途径来增加税收，以便为项目提供资金池。当

然，那种途径本身也会产生激励问题，但那不是直接的。

更好的解决方法是同时拥有其他更多的资金来源。例如，一所大学，如果有规模很大的校级基金，同时组成大学的各个学院又各自都有独立的基金，就可以用这种机制来决定是否建立一个新的学生会。大学每个学院的院长都有动机真实地揭示学生会对自己学院的价值，同时大学校长则可以用校级基金来弥补可能资金的不足。由拥有预算权限的分支机构组成的企业也可以考虑这样做。例如，当这样的企业在决定要不要切换到一个基于云的系统时，就可以采用枢轴机制来决策，而且更高的管理层可以解决任何可能出现的缺陷。

小结

作为一个框架，机制设计理论使我们能够依据各种标准对不同的机制进行比较。机制设计能不能产生有效率的结果？人们会说实话吗？人们会自愿参加吗？某个机制是会产生盈余、还是会导致损失？利用机制设计框架，我们还可以推导出可能的结果。当然，一般来说，我们无法在同一个机制下满足所有想要达到的标准。在进行机制设计的时候，建模者摇身一变，成了工程师。我们使用模型来尝试构建可行的解决方案。

随着技术的进步和变化，机制也会发生变化。以谷歌等互联网搜索网站所使用的广告拍卖算法为例。最初，谷歌是按固定价格收费的，每千次点击收取多少费用。随着信息技术的发展，谷歌能够同时进行数百万次拍卖，在这种情况下，固定收费就不再是一种最优机制了。通过引入拍卖方法，谷歌不仅增加了收入，并且更有效地分配了广告空间。谷歌现在使用的是一种广义的第二价格拍卖。每个竞买人都会提交自己对每次点击的出价，目的是通过特定的关键词进行推广，例如，治疗一种因接触石棉而导致的癌症的药物，最高出价者得到第一个广告位，第二高出价者得到第二个广告位，第三高出价者得到第三个广告位……以此

类推。这些竞买人付出的价格则通过第二价格拍卖来确定。

假设前四个最高出价分别是每次点击10美元、7美元、6美元和3美元。那么，第三高出价者获得第三个广告位，同时支付的价格等于第四高出价者的出价，即3美元；第二高出价者支付的价格则等于第三高出价者的出价，即6美元；最高出价者则要付出7美元的价格。⁹在了解到了广告商的估值后，谷歌还可以设定保留价格（底价）并收取更高的费用。但是，如果竞买人也知悉了谷歌的这些计划，这个结果就不一定是有效的。认为自己有可能成为高出价者的竞买人不希望谷歌知道他的估值。同时设定底价也会损害谷歌的声誉，而且，底价也会被视为一种非合作行为，因为谷歌无法证明自己有权对网页上的位置拥有保留价值。除非卖出去了，否则关键字搜索上的顶级广告位对谷歌本身来说几乎没有任何价值。而对于销售古董家具或二手车的人来说，情况却并非如此。这些物品对卖家是有价值的，因此设定一个底价是合理的。而且，谷歌是一个重视声誉的企业，设定底价可能会激怒广告客户。

总而言之，机制设计框架可以帮助我们设计制度，也能够指导我们在不同的制度之间进行选择。有了这个框架，就可以推断出什么是可行的、什么是不可能实现的。也许，我们很难设计出这样一种机制：既能够产生有效率的结果，又能够引导人说真话，同时还满足预算平衡条件。如果确实是这样，我们就不应该浪费时间和精力去设计这样的机制，而应该将更大的时间精力投入到如何实现（例如）效率和平衡预算之间的权衡上。

我们还可以利用机制设计框架来探索一些更宏大的问题，例如，我们在什么情况下应该利用市场、在什么条件下应该投票、在什么时候应该依靠等级体系，在什么环境中应该转而采用自愿的集体行动来分配资源或采取行动。¹⁰市场、民主、等级制度和集体行动这四个机制中的每一个，都只在某些环境下运行良好，而在另外一些环境下则表现不佳。例如，我们不会用投票来决定人们购买什么商品，也不想让市场去决定

谁来当美国总统。

在社会和组织内部，我们都观察到了这些制度形式。以大学为例。大学要面对一个教职市场，同时又要通过民主程序来雇用教师；要通过等级体系分配课程作业，并运用集体原则制定战略、计划。非营利机构、以获取利润为目标的企业，以及政府机构，也都需要混合运用各种不同的制度形式。利用机制设计工具箱，我们可以比较每种制度的运行方式，然后再把各种任务更好地适配各种制度。

25 信号模型

诚实的人不会隐瞒自己的所作所为。

艾米莉·勃朗特（Emily Brontë）

在本章中，我们研究信号模型（signaling model）。这类模型确定了人们发送“昂贵”的信号以揭示信息或类型的条件。一个人可以通过购买昂贵的艺术品表明自己的财力，通过攀登很高的山峰来展现自己的体力，或者通过在社交媒体上发声支持受难者来表达自己的同情心。利用发送信号来揭示自己的身份一直都是人性的一个部分。

早在19世纪，经济学家托尔斯坦·凡勃仑（Thorstein Veblen）就提出了“炫耀性消费”的概念，大大增进了我们对信号的理解。凡勃仑观察到，人们经常选择通过炫耀性消费来表明他们的社会地位，而不仅仅会购买那些能够带来直接享受或效用的商品。如果活到今天，当凡勃仑看到现代人的炫耀性消费行为，他肯定将会心一笑：例如迈巴赫敞篷车，每辆售价将近150万美元；10年陈酿的克丽丝特尔酒（Cristal），每瓶售价超过1 500美元；徕卡相机，每台售价数万美元……

炫耀性消费由来已久，部分原因在于人类很在意别人对自己的看法。这种消费行为之所以经久不衰，还因为消费可以起到信号的作用。¹我们不能完全看清某个人，所以我们依赖于他们穿的衣服、开的汽车、喝的酒来推断他们的“隐藏属性”。如果我们看到一个人开着昂贵的汽车，那么大体上可以推断出他拥有一定财富。一个人向慈善组织大笔捐款，表明他是一个慷慨大方的人，因为没有自私的人会做出这样的行为。一个人在社交媒体上宣布自己获得了理论生物学博士学位，那是传达关于他的智力水平和所从事职业的信号。几乎所有行动都在一定程度上传递了某种信号。当政客们投票决定是对某个国家宣战、还是实施

制裁时，他们就发出了关于自己意识形态立场的信号。某些有长期目标的政客（例如打算日后竞选总统）可能会试图通过投票发出最有利于自己未来政治前途的信号，因而不一定会给最有利的政策投赞成票。

在本章中，我们首先研究离散信号模型。在这种模型中，个体可以选择发送信号或不发送信号；同时，不同的个体发送信号的成本也不相同。要让信号发挥作用，它们就必须是昂贵的（有成本的）或可验证的。这将是本章的一个重点内容。例如，一个雇主打算从新入职员工中选一些人夏天到西班牙巴塞罗那出一趟“美差”。所有申请的员工都在简历中声称自己会西班牙语。但是，简单地说自己会说西班牙语只是一个没有成本的信号。为此，雇主可以启动一个“西班牙语争章活动”，要获得一枚徽章，需要用西班牙语完成一个小时的演示。对于真的熟练掌握了西班牙语的员工，发出这个信号，也就是用西班牙语完成演示的成本较低。但是对于那些不懂西班牙语的员工来说，准备长达一个小时的西班牙语演示的成本却高得令人望而却步。用信号模型的术语来说，这个徽章就将熟练西班牙语的人与不懂西班牙语的人区分开来了。

接下来，我们简单地介绍一个连续的信号模型。在这个模型中，信号的大小是可变的。一个夏令营的皮划艇队一般只能有一个领划的皮划艇运动员，人们希望在这个位置上的的是一个耐力非常好的人，那么，怎样才能挑选出这样一个人来呢？营地主管可能会要求两个提出申请的选手连续划皮划艇10个小时，以便将皮划艇划到尽可能远的距离上。那两个皮划艇运动员中耐力更强的那个选手可以划到另一个实力较弱的选手无法企及的距离上，从而让自己脱颖而出。

离散信号模型和连续信号模型都能为我们提供信号什么时候分离、什么时候不能分离的条件。因此，它们给我们提供了比文字描述更加准确深入的见解。文字描述能够告诉我们，人、动物、政客和政府发出了什么信号和为什么要发出信号，但是却不能向我们明确表征信号发出的时间和信号的强弱。这些模型还可以非常清晰地解释，为什么学生会如

此努力地试图证明自己对学院的价值。在本章的结论中，我们讨论了信号模型的理论贡献及其政策含义，还将讨论生态学、人类学下的信号模型，以及商业活动中的信号模型应用。

离散信号模型

我们从离散信号模型开始讨论。在这种模型中，人们要决定是否采取某种行动。你可以买一块昂贵的手表来证明你拥有大量财富，可以通过主修物理学来证明你智力超群，可以通过横渡英吉利海峡来证明自己身体健康。但是你不能半途而废：要么发送信号，要么不发送。这个模型假设，存在两种类型的人，强者和弱者。这两种类型在现实世界中，可以对应于身体健康有资格进入海军陆战队的年轻人和身体孱弱的人，也可以对应于会两种以上外语的员工和只会本国语言的员工，等等。

发送信号的成本取决于个体的类型。这里所说的信号，可能是为有可能成为海军陆战队员的申请人提供的为期一个月的魔鬼训练计划，也可能是求职者用西班牙语完成的长达一个小时的演示。强健的准海军陆战队员会发现完成训练计划的成本更低。在模型中，我们假设发送信号的每个人可以平等地分享总收益。对于这个假设，可以从两个角度加以解释。在某些情况下，某种资源可能会在发送信号的所有人之间分配。例如，向学校捐赠了1 000美元的每个人（捐赠是慷慨的信号）的名字都会被刻在一面墙上。在另外一些情况下，例如对于准海军陆战队员和求职者，则可以从发送信号的人的集合中随机挑选一些出来作为“中奖者”。

这个模型支持三种不同的结果：混同（pooling），所有人都发送相同的信号；分离（separating），每种类型的人各自发送一个独特的信号；部分混同，其中一些类型区分开来了，其他类型则没有区分开来。

离散信号模型

一个规模为 N 的种群由 S 个强者类型的个体和 W 个弱者类型的个体组成；这两种类型的个体发送信号的成本分别为 c 和 C ，且 $c < C$ 。种群中所有发送信号的成员平均分配 B 的收益（ $B > 0$ ）。这个模型有三种可能的结果：

混同（ $C < \frac{B}{N}$ ）：两种类型的个体都发送信号。

分离（ $c < \frac{B}{S}$ ，且 $\frac{B}{S+1} < C$ ）：只有“强者”类型的个体发送信号。

部分混同（ $c < \frac{B}{N} < C < \frac{B}{S+1}$ ）：“强者”类型的所有个体和“弱者”类型的部分个体发送信号。

在这个模型中，我们假设个体在给定其他个体行动的情况下做出最优选择。也就是说，我们将它视为一个博弈，并求解它的均衡。在混同均衡中，每个人都发出信号，如果收益很高并且弱者类型的人发送信号的成本较低，就会存在这种均衡，确切的条件是收益除以人数的商必须超过弱者类型的人的成本。例如，假设一位捐助者捐赠了100万美元设立了一个奖学金，用于奖励某所高中所有100名毕业生。假设有50名学生属于强者类型，只需每周学习两小时就能够从高中毕业，而另外50名学生则属于弱者类型，每周必须花费10小时学习才能完成高中学业。对于强者类型的学生，我们可以将学习成本估计为2 000美元，而对于弱者类型的学生，学习成本则为5 000美元。如果所有100名学生都顺利毕业，那么每人都可以获得10 000美元的奖学金。因此这两种类型的学生都有很强的动机去学习。

但是，如果假设我们将奖学金总额减少到了20万美元。现在，如果所有100名学生都顺利毕业，那么每人都只能获得2 000美元的奖学金。这样，学习就不再符合弱者类型学生的自身利益了。而对于强者类型的学生来说，现在每人可以得到4 000美元奖学金了，因此学习仍然是有意义的。但是，这个数额仍然不足以诱导弱者类型的学生毕业，哪怕只有一个都不可能。在这种情况下，奖学金的设置导致了分离均衡。

最后，假设奖学金的总额为40万美元。再一次，如果所有100名学生都毕业，那么弱者类型的每个学生的所能得到的奖学金为4 000美元，低于5 000美元。因此，他们不会全都选择学习。但是，如果弱者类型的学生都不学习，那么强者类型的学生每人将会获得8 000美元的奖学金，这个数额对弱者类型的学生来说也很有吸引力。因此在均衡中，最终将会有30名弱者类型的学生与所有50名强者类型的优秀学生一起毕业。结果是总共有80名学生毕业，每人得到5 000美元的奖学金，这也正是弱者类型的学生的学习成本。我们将这种结果称为部分混同，因为有部分弱者类型的学生与强者类型的学生混同在了一起。

部分混同均衡比其他两个均衡更加复杂，因为它需要弱者类型的学生实现彼此之间的某种协调。我们可以假设存在某个过程，弱者类型的学生会与其他人沟通，告诉别人自己计划采取能够确保毕业的行动。或者也可以假设弱者类型的学生的努力恰恰达到了这样一个水平：他们能不能毕业完全是随机的，并且该努力水平会导致30名弱者类型的学生毕业的期望。第二种情况似乎不那么合理。一般而言，我们应该将部分混同均衡解释为一个基准，即如果人们试图最优化，会发生什么。是否能达到部分汇集均衡，可能取决于具体情况，尤其取决于人们是否可以交流各自预期中的行动。

连续信号模型

在离散信号模型的部分混同均衡中，强者类型的人在有些时候可能会觉得沮丧。如果他们能够发出足够强烈的信号，就可以完全与弱者类型的人区分开来，并且获得更高的收益。为了在模型中包含这种可能性，我们可以改变假设并允许强者类型的人自行选择它们要发送的信号大小。这只需要对模型稍做修改即可。为此，我们将离散信号的发送成本重新解释为连续信号的每单位发送成本。此外，我们假设，对于任何固定数量的信号，强者类型的人每单位成本更低一些。

为了在这个新模型中实现分离均衡，强者类型的人必须愿意选择一个对于弱者类型的人来说成本极高的信号，当然条件是在考虑了收益和成本之后，这个信号仍然是值得发送出去的。通过模型推导，我们发现至少有一些强者类型是可以分离的，但不一定是全部。

令人惊讶的是，随着强者类型的群体规模的增大，信号的量级反而会变小。这种情况之所以会发生，是因为强者类型的人发送信号的好处减少了。成为规模更大的群体的一部分，能够得到的好处反而会更少。完全分离这个条件意味着，当强者类型的人数很少，或者强者类型的人发送信号成本要比弱者类型的人低很多时，分离均衡更有可能实现。

连续信号模型

一个规模为 N 的种群由 S 个强者类型的个体和 W 个弱者类型的个体组成，两种类型的个体发送信号的单位成本分别为 c 和 C ($C > c$)。发送最大信号的所有个体分享利益 B 。任何大小为 $M \geq \frac{B}{SC}$ 的信号都能够将强者类型分离开来。如果 $CW \geq cN$ ，那么所有强者类型都会分离开来。如若不然，就存在部分混同均衡，其中一部分弱者类型的个体也会发送信号。²

这个模型可以解释，为什么昂贵的手表和珠宝能够作为财富的信号。一个人的房子或汽车也标志着其拥有的财富，但是人们无法随时随身携带房屋和汽车。衣服也可以发出财富信号，但是却可能无法创造分离。只要花上几百美元，任何一个人可以穿得像一个拥有大量财富的人。但由于成本很高，手表和珠宝作为财富信号会更加有效。一个穷人或中产阶级人士买不起售价一万美元的手表。戴这样的手表，足以证明自己拥有可观的财富。通过发送这样的信号能够获得的好处可能是，人们会更尊重他，假设人们认为财富在某种程度上与一个人的重要性相关的话（尽管有人可能会质疑这种推断）。

信号的用途和价值

信号能够把隐藏的属性突显出来。我们的行动标志着我们的健康、财富、智慧和慷慨。我们的一些行动产生了作为副产品的信号。一个纯粹是因为对长跑有兴趣而参加马拉松比赛的人，可能会传递出身体健康和做事专注投入的信号，尽管这可能这不是他的本意。信号模型为解释几乎任何行动提供了另一种视角。但是，一个人选择参加某种活动、学习掌握某种技能、购买某个商品，到底是完全出于个人兴趣，还是在发出某种信号呢？我们也许无法分辨。

例如，信号模型为大学文凭的价值提供了另一种解释。关于收入的大量数据表明，大学毕业生的工资显著高于没有接受过大学教育的人。我们可以推断，较高的工资源于大学期间获得的技能和知识。同时，相关数据还表明，数学和科学专业的大学毕业生工资更高。由此可以推断，在这些专业中学到的技能具有更大的经济价值。然而，如果认真观察一下人们所实际从事的工作，我们可能会发现很少有人在工作中需要使用微积分。而且，求职者在接受面试时，几乎从来没有人被问到过余弦函数的导数怎么求、玻意耳定律怎么解释。有鉴于此，我们可以推断大学学位，特别是科学和数学学位，代表了一个人获取知识能力的信号。毕业生获得的较高薪酬完全取决于学位的信号价值，而不是毕业生在大学期间所学到的知识。³

可以考虑一下成为一名医生所必须发送出去的信号。医学专业的学生必须通过物理、有机化学和微积分等课程的考试。但是，医生看病是否使用微积分？医生在为你诊疗耳朵和鼻子的时候，会先在他的记事本上写出一个微积分方程吗？当然不是这样。在很大程度上，微积分知识与医生执业可能完全无关，但是它可能是医生掌握知识体系能力的良好信号。如果真的是这样，那么即便与从事的职业几乎没有任何直接相关性，通过微积分考试也会成为医生的一个有用的信号。

只要有可能，任何人在构建信号的时候，都更愿意在生成信号的同时也能掌握有用的技能。例如，事实证明，要成为一名成功的医生，记忆能力是很重要的。为了传递能够证明记忆能力的信号，面试者可能会要求申请人背出每个国家的首都和货币。成功地通过这项考核，能够证明申请人确实有很强的记忆能力，但是所记忆的这些内容对成为一名好医生并无意义。当你觉得自己的肠胃非常不舒服，匆忙赶到急诊室时，你并不会在乎给你看病的医生是否知道布拉迪斯拉发是斯洛伐克的首都，你只是希望那个医生对消化系统的各个部分都了如指掌。出于这个原因，医疗委员会要求医生通过解剖学考试。通过解剖学考试能够证明一个人的记忆能力，而且记住身体的各个部位也确实是有用的。因此，通过解剖学考试是一个功能性信号（functional signal）。

小结

信号模型的应用范围非常广泛。如前所述，雄孔雀的美丽尾羽是它“身体健康”的信号。众所周知，雄孔雀装饰性极强的扇形尾羽几乎没有任何功能性价值，事实上，这种夸张的尾羽不但无用，而且可能还会给它们带来糟糕的结果。雄孔雀如果选择发展更强壮的爪子，有用性要高得多。但是，强壮的爪子很难让雌孔雀在很远的地方就注意到，这一点比尾羽差得太多了，因此尾羽在演化过程中胜出了。⁴雄性果蝇的彩色尾部也具有与雄孔雀的尾羽类似的功能，蚱蜢和鸟类的鸣叫声也是如此。啁啾需要付出可观的能量，只有吃饱了的蚱蜢才可以花时间啁啾而不用去忙着寻找食物。因此，啁啾声可以起到信号的作用。

在人类社会，不同的文化会通过不同的行动来表明健康状况。人类学家区分出了三种形式的昂贵信号：无条件的慷慨（unconditional generosity），浪费性的维生方式（wasteful subsistence behavior）、精美的传统手工艺制品。⁵“夸富宴”是居住在太平洋西北地区的印第安土著居民举行的一种仪式，这可能是发送这种慷慨信号的最为突出的一个例

子。为了庆祝一个事件，比如一个成员出生或去世时，酋长会送出大量的财富，甚至直接毁坏财富，并对其他酋长提出挑战，要求他们也做出同样的行为。其他酋长如果做不到，就会失去声望。将自己的财物赠予他人，还可以说是有利于社会，但是将它们毁坏无疑是极大的浪费。

事实上，当人们（通常是男性）在预期收益比采集种子或浆果更低时，仍然坚持远行狩猎时，就已经采取了浪费性的维生方式了。男人这样做是因为他们能够希望获得额外的尊重。狩猎成功，猎人就发出了说明他力量和勇气的信号，这在其他环境中也可能很有用。作为一名成功的浆果采集者，能够发送自己拥有良好的视力和耐心的信号，这些当然也是有用的个人特点，这一点毫无疑问，但是在很多方面的预测性能不如狩猎技能好。对生活在澳大利亚北部一个群岛上的梅里亚姆人

（Meriam）的一项研究表明，平均来说，作为海龟猎手的男性居民，在50岁时存活的后代人数是其他同龄男性居民的两倍多。⁶

复杂精美的传统工艺品制作需要付出非常多的时间和资源。当然，这种活动也可能生产出有用的物品，如地毯。但是，大多数传统工艺品都是没有太大实用价值的礼仪性物品。一些人类学家将这类传统工艺品的制作解释为信号的发送。创作生产这些物品的意义，不依赖于它们能够实现的功能，因这它们具有重要的文化意义。

很多广告也可以解释为昂贵的信号。例如，购买昂贵的超级碗总决赛的商业广告位，可以说是在发送关于自己产品的“合法性”信号。因为这意味着企业相信消费者会非常喜欢自己的产品，从而可以赚回足以覆盖广告成本的利润。想象一下，假设现在有两家企业分别推出了一款新咖啡机。第一家企业知道自己开发了一个“伟大”的产品。而第二家企业则知道，尽管自己的工程师付出了最大的努力，但是这个产品仍然可能会故障频出导致消费者大量投诉。第二家企业预计将会出现20%的退货率。

每一年，都会有数百万人购买咖啡机。如果不做广告，那么这两家

企业可能会平分市场。假设生产出了更好产品的企业决定投入200万美元来宣传自己产品的质量。这家企业预计，在广告攻势下，早期购买者都会购买自己的产品，并且从长远来看，这又会导致更大的销量。这家企业的决策者的脑袋中，可能有一个波利亚瓮模型。相比之下，生产质量较差产品的另一家企业则不会花钱做这样的广告，因为它预计自己的产品应该不会非常畅销。花大钱来表明产品的质量，这种行为有时被称为“烧钱”（burning money）。就像雄孔雀用尾羽吸引到了潜在的配偶一样，企业通过“烧钱”吸引了消费者。

在所有这些情况下，发送信号都要付出成本。那些发出信号的人会发现，信号的成本，会导致他们更大的财富、能力以及慷慨个性被他人识别所能带来的好处而有所减少。此外，发送信号所耗费的时间和精力也可以被认为是一种机会成本：如果把这些资源用在其他用途上，可能会产生更大的社会盈余。例如，一个年轻人可能会花费数小时去决定穿什么衣服，以便表明自己的“社会意识”；或者，一个高中生可能会将大量时间和精力投入到某种“非生产性的”活动中去，因为他相信这样能够提高他被精英大学录取的机会。

为了减少发送信号的社会成本，我们应努力使信号尽可能有效地发挥作用。例如，最好是让年轻人通过参加团体性的运动来证明自己身体健康和勇敢，通过这种运动，他们能够学会体育精神和尊重集体利益，而不要让他们冒着生命危险从飞驰的摩托车上跳下以证明自己的勇敢。最好是要求医生记住人体解剖图谱，而不是考他们记得多少句《魔戒》的精灵语。

尽可能地多尝试吧。浪费性的信号肯定会继续存在。我们的挑战是利用模型，特别是机制设计工具，来构建制度和协议，以保证发送出去的信号确实携带了充分的信息。

26 学习模型

一个人可以养成的最重要的习惯就是对继续学习的渴望。

约翰·杜威（John Dewey）

本章研究个体学习模型和社会学习模型，我们会在两种情况下应用它们。第一种情况，如何学会在一个备选方案集合上做出最优选择。在这种情况下，个体学习和社会学习将会汇聚到最优选择上，而学习规则的不同只能影响收敛速度。第二种情况，如何在博弈中应用学习规则来采取适当的行动。在博弈中，某个行动的收益取决于本人和其他博弈参与者的行动。在这两种情况下，学习规则都更有利于规避风险的均衡结果而非有效率的均衡结果。我们还发现，个体学习并不一定会产生与社会学习相同的结果，而且任何一种学习都不可能所有环境下都比另一种学习表现得更好。

这些发现为我们的主张——采用多模型方法来表征行为，提供了有力的支持。学习模型介于理性选择模型与基于规则的模型之间。理性选择模型假设人们会审慎考虑所处的环境和要完成的博弈，然后采取最优行动；基于规则的模型则直接根据规则来指定行动。学习模型假设人们会遵循规则，但是，正是这些规则使行为能够发生改变。在某些情况下，行为会趋向最优行为。在这些情况下，学习模型可以用来证明假设人们会采取最优行动的合理性。但是，学习模型也不一定会收敛到均衡，它们也可能生成循环或复杂的动态。而且，如果学习模型确实收敛了，它们可能会有比其他模型更多的均衡可以选择。

本章的内容安排如下。我们首先描述强化学习模型，并将这种模型应用于如何选择最优备选方案的问题。强化学习模型通过更高的奖励来强化行动。随着时间的推移，学习者会学会只采取最优行动。这是一个

基准模型，非常适合研究学习模型。它与实验数据也拟合得相当好，而且不仅仅适用于人类。海蛞蝓、鸽子和老鼠，都会强化成功的行动。相比之下，强化学习模型也许更适用于海蛞蝓，它只有不到2万个神经元，而不那么适用于拥有超过850亿个神经元的人类。如此巨大的脑容量使人类能够在学习时考虑反事实，而这种现象是强化学习模型无法考虑的。

然后，我们介绍社会学习模型。在社会学习模型中，个体能够从自己的选择和他人的选择中学习。个体会复制最流行的或表现高于平均水平的行动或策略。社会学习假设行为者能够观察或沟通。有些物种是通过所谓的共识主动性（*stigmergy*）来实现社会学习的：成功的行动会留下其他个体可以追随的痕迹或残留物。例如，当山羊在群山间走动时，会留下被踩踏的草，从而强化了通往水或食物的路径。

接着，我们将这两种类型的学习模型应用于博弈分析。如前所述，博弈给出了一个更加复杂的学习环境。同样的行动，可能会在这一个时期内带来高收益，在下一个时期内却产生低回报。正如人们通常可以预料到的那样，我们发现社会学习模型和个体学习模型都不一定会收敛到有效的均衡，而且它们也可能产生不同的结果。最后，我们讨论了一些更加复杂的学习规则。¹

个体学习模型：强化学习模型

在强化学习中，个体要根据各个行动的不同权重来选择行动。权重较大的行动比权重较小的行动更经常被选中。分配给某个行动的权重取决于这个个体在过去采取这个行动时所获得的奖励（收益）。这种高回报收益的强化可以导致个体选择更好的行动。在这里，我们要探讨的问题是，强化学习是不是会收敛为只选择具有最高奖励（收益）的那个备选方案。

乍一看，只选择最有价值的那个备选方案似乎是一个非常容易完成的微不足道的任务。如果奖励是完全以数值形式来表示的，例如金钱的数额或时间的长短，那么我们有理由相信人们会选择最好的那个备选方案。在第4章中，我们就是用这种思路来说明一个在洛杉矶工作的人在选择通勤路线时会选择最短的路线。

但是，如果奖励没有采用数值形式（通常情况下都是如此），人们就必须依赖自己的记忆。我们在一家韩国餐厅吃过一次午餐，发现那里的泡菜很美味，所以我们更有可能再次光顾那家餐厅。星期一，我们在跑步前一小时吃了燕麦饼干，结果发现我们连续跑上10千米都不觉得累。如果星期三，我们又在跑步前吃了燕麦饼干并且步履如飞，我们就会加大这个行动（跑步前吃点燕麦饼干）的权重，因为我们已经知道燕麦饼干可以改善跑步成绩。

除了人类之外，其他物种也会这样做。早期研究学习的心理学家爱德华·桑代克（Edward Thorndike）设计了一个经典实验。在这个实验中，一只通过拉动杠杆逃离了箱子的猫得到了奖励。在回到箱子中之后，这只猫在几秒钟内就再次拉动了杠杆。桑代克得到的数据表明，猫会持续进行尝试。他发现猫（以及人）在奖励增大时学习得更快。他将这个规律称为效果律（law of effect）。²桑代克的这个发现是有神经解剖学基础的。重复一个行动会构建出一个神经通路，而这个神经通路在未来会引发相同的行为。桑代克还发现，更出人意料的奖励，也就是远远超出过去奖励水平或预期奖励水平的奖励，会使人们学习得更快，他把这个规律称为惊奇律（law of surprise）。³

在强化学习模型中，分配给一个所选备选方案的权重，是根据该备选方案在何种程度上超过了预期，即“渴望水平”（aspiration level）来进行调整的。这样的模型，既考虑了效果律，也就是会采取那些能够更经常地产生更高回报的行动，也考虑了惊奇律，也就是对某个备选方案赋予的权重取决于它所带来的奖励超过了渴望水平的程度。⁴

强化学习模型

假设一个由 N 个备选方案组成的集合 $\{A, B, C, D, \dots, N\}$ 、与各备选方案对应的奖励的集合 $\{\pi(A), \pi(B), \pi(C), \pi(D), \dots, \pi(N)\}$ ，以及一个严格为正的权重的集合 $\{w(A), w(B), w(C), w(D), \dots, w(N)\}$ 。那么，选择备选方案 K 的概率如下：

$$P(K) = \frac{w(K)}{w(A) + w(B) + w(C) + w(D) + \dots + w(N)}$$

在选中了备选方案 K 之后， $w(K)$ 会增大 $\gamma \times P(K) \times (\pi(K) - A)$ ，其中 $\gamma > 0$ 等于调整速率（rate of adjustment）， $A < \max_K \pi(K)$ 等于渴望水平。⁵

这里需要注意的是，渴望水平必须设定为低于至少一个备选方案的奖励水平。否则，被选中的任何一个备选方案在未来再次被选中的可能性会很低，而且所有备选方案的权重都会收敛到零。不难证明，如果渴望水平低于至少一个备选方案的奖励水平，那么最终几乎所有权重都会被赋予在最优备选方案上。之所以会发生这种情况，是因为每选择一次最优备选方案，权重的增加幅度都会最大，从而给这个备选方案创造了更强的强化。

即便我们将渴望水平设定为低于任何一个备选方案的奖励水平，这种情况也必定会发生。在将渴望水平设定为低于任何一个备选方案的奖励水平时，每种备选方案被选中时权重都会有所增加；因此，这个模型可以用来刻画习惯形成：之所以更频繁地做某件事情，只是因为我们在过去已经做过这件事情。而且，即便将渴望水平设定得很低，会带来最高水平奖励的那个备选方案也会以最快的速度增加权重，因此从长远来看，最优备选方案将会胜出。但是，要收敛到最优备选方案上，所需的时间可能会很长。另外，当我们增加了更多的备选方案时，收敛时间也

会变长。

为了避免这些问题，我们可以构造内在愿望（endogenous aspiration）。为此，我们修改上面的模型，将渴望水平设置为平均奖励，从而让它随时间推移而不断调整。想象一下，假设父母试图确定自己的孩子到底是更喜欢苹果薄饼还是香蕉薄饼。设定选择苹果薄饼的奖励为20，选择香蕉薄饼的奖励为10，并将两个备选方案的初始权重都设置为50，将调整速率设定为1，并将渴望水平设定为5。假设父母在第一天准备的是香蕉薄饼，这样香蕉薄饼的权重将增加到55。假设父母在第二天也准备了香蕉薄饼，那么10的奖励等于新的渴望水平，香蕉薄饼的权重不会改变。

假设父母在第三天准备了苹果薄饼。这会带来20的奖励，超出了渴望水平。这会使苹果薄饼的权重增加到60，从而使苹果薄饼变成了更可能被选中的备选方案。更高的奖励也提高了平均收益，因而也使渴望水平上升到了10以上。因此，如果父母再一次准备香蕉薄饼，香蕉薄饼的权重就会减少，因为香蕉薄饼的奖励水平已经低于新的渴望水平了。也就是说，强化学习将收敛为只会选择苹果薄饼。

我们很容易就可以证明，强化学习将趋向于以概率1选择最优备选方案。这个结论意味着，与所有其他备选方案的权重相比，最优备选方案的权重将会变得任意大。

强化学习的效果

在学会选择最优备选方案模型的框架中，当渴望水平被设定为等于平均获得的奖励时，强化学习（最终）几乎总是会选择最优备选方案。

社会学习模型：复制者动态

强化学习假设个体是孤立采取行动的。但是，人们也会通过观察他人来学习。社会学习模型假设个体能够观察到他人的行动和奖励，这可以加快学习速度。现在学界研究得最充分的社会学习模型是复制者动态（replicator dynamics），它假设采取某个行动的概率取决于该行动的奖励和它的受欢迎程度。我们可以将前者称为奖励效应（reward effect），把后者称为从众效应（conformity effect）。⁶在大多数情况下，复制者动态模型都要假定一个无限种群。在这个假设的基础上，我们可以将所采取的行动描述为各种备选方案之间的概率分布。在标准的复制者动态模型中，时间是不连续的，所以我们可以通过概率分布的变化来刻画学习。

复制者动态

假设一个由 N 个备选方案组成的集合 $\{A, B, C, D, \dots, N\}$ 、与各备选方案对应的奖励的集合 $\{\pi(A), \pi(B), \pi(C), \pi(D), \dots, \pi(N)\}$ 。在时间 t ，一个种群的行动可以用这 N 个备选方案上的概率分布来描述： $(P_t(A), P_t(B), \dots, P_t(N))$ 。且这个概率分布随如下复制者动态方程而变化：

$$P_{t+1}(K) = P_t(K) \times \left(\frac{\pi(K)}{\bar{\pi}_t} \right)$$

其中， $\bar{\pi}_t$ 等于第 t 期中的平均奖励。

考虑这样一个例子，父母可以选择准备苹果薄饼、香蕉薄饼和巧克力薄饼。假设所有的孩子都有同样的偏好，再假设这三种薄饼分别能产生20、10和5的奖励。如果最初有10%的父母制作苹果薄饼、70%的父母制作香蕉薄饼、20%的父母制作巧克力薄饼，那么平均奖励等于10。应用复制者动态方程，在第2期中选择三个备选方案中的每一个的概率如表26-1所示：

表26-1 备选方案概率

备选方案	π	P_1	$\pi/\bar{\pi}_t$	P_2
苹果薄饼	20	0.1	20/10	0.2
香蕉薄饼	10	0.7	10/10	0.7
巧克力薄饼	5	0.2	5/10	0.1

复制者动态方程告诉我们，在接下来一段时间里，制作苹果薄饼的父母的比例将会增大到原来的两倍。这是因为苹果薄饼的奖励等于平均奖励的两倍。而制作巧克力薄饼的父母将会减少，因为巧克力薄饼的奖励只相当平均水平的一半。最后，制作香蕉薄饼的父母的比例则不会改变，因为香蕉薄饼的奖励恰恰等于平均奖励。结合所有这些变化，我们发现平均奖励增大到了11.5。

如前所述，复制者动态同时包括了从众效应（更受欢迎的备选方案更有可能被复制）以及奖励效应。从长期来看，奖励效应占主导地位，因为高奖励的备选方案总是会与奖励水平成比例增长。在复制者动态中，平均奖励发挥的作用，与强化学习中当渴望水平随着平均奖励水平而调整时、渴望水平所发挥的作用类似。两者之间唯一的区别是，在复制者动态中，我们要计算整个种群的平均奖励，而在强化学习中，渴望水平等于个体的平均奖励。只要种群是一个相当大的样本，这种区别就是很重要的。因此，复制者动态产生的路径依赖要远小于强化学习。

在构建复制者动态时，我们需要假设每个备选方案都已经存在于初始种群中。由于最高奖励的备选方案总是具有高于平均奖励水平的奖励，而且它的比例在每个时期都会增加，因而复制者动态（最终）会收敛到整个种群都选择最优备选方案的结果。²因此，在这种学会选择最优的环境中，个体学习和社会学习都会收敛到拥有最高奖励水平的那个备选方案上。但是在博弈中则不一定会这样。

复制者动态能够学会最优行动

在学会从一个有限的备选方案集中选择最好的备选方案的过程，无限种群复制者动态几乎总是收敛到整个种群都选择最优备选方案。

博弈中的学习

现在，我们将两种学习模型应用于博弈。⁸请先回想一下，在博弈中，博弈参与者的收益不仅取决于他自己的行动，同时也取决于其他博弈参与者的行动。某个特定行动的收益，例如在囚徒困境中的合作，可能会在一个时期内很高而在下一个时期却很低，这取决于另一个博弈参与者的行动。

我们先从油老虎车博弈（Guzzler Game）开始讨论。这是一个双人博弈，每个博弈参与者都必须选择是驾驶一辆经济型汽车还是一辆高油耗但很坚固的车。选择高油耗的车总能带来2的收益。当一个博弈参与者选择经济型汽车，另一个博弈参与者也选择经济型汽车时，双方都可以得到3的收益，因为两个司机都会有更好的视野，汽车耗费燃油也更少，而且都不必担心被巨大的耗油量压垮。但是，如果另一位博弈参与者选择了油老虎车，那么驾驶经济型汽车的博弈参与者必须非常注意那个人的行为。为了刻画这种影响，我们假设在这种情况下，驾驶经济型汽车的博弈参与者的收益会降低为零。图26-1中给出了这个收益矩阵。

	开油老虎车	开经济型汽车
开油老虎车	2, 2	2, 0
开经济型汽车	0, 2	3, 3

图26-1 油老虎车博弈

这个油老虎博弈有两个纯策略均衡：两个博弈参与者同时选择经济型汽车，或者两个博弈参与者同时选择油老虎车。⁹双方都选择经济型汽车的这个均衡会带来更高的收益，这是这个博弈中的有效均衡。

我们先假设，这两个博弈参与者都会进行强化学习。图26-2给出了4个数值实验的结果，其参数为：每个行动集的初始权重都等于5、渴望水平为零、学习速度（ γ ）为1/3。在这所有4个数值实验中，两个博弈参与者都学会了选择油老虎车，即低效率的纯策略均衡。为什么会这样？为了分析这种情况发生的原因，只需要看一看收益矩阵即可。选择油老虎车的博弈参与者总能得到2的收益；而选择经济型汽车的博弈参与者则有时会得到3的收益，有时则什么也得不到（收益为零）。根据假设，两个行动在初始人口中出现的概率是相等的。因此，选择经济型汽车的平均收益仅为1.5，而选择油老虎车的平均收益则为2。于是就会有更多的博弈参与者选择油老虎车，而这又使选择经济型汽车的收益进一步下降。

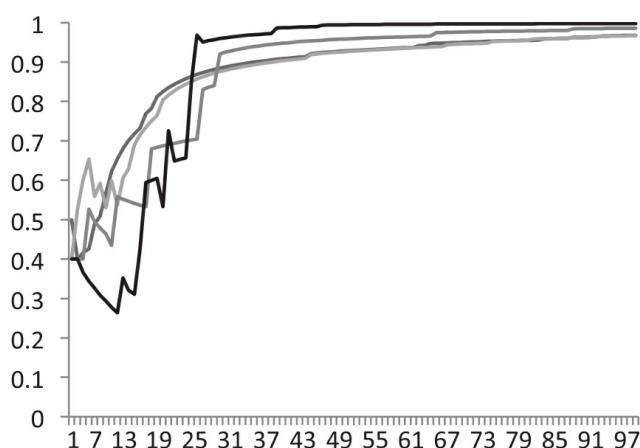


图26-2 油老虎博弈中的强化学习：选择油老虎车的概率

接下来，我们将复制者动态应用到这个博弈中来。我们假设初始人口由相同比例的选择油老虎车和经济型汽车的人组成。然后进一步假设每个博弈参与者与每个其他博弈参与者博弈。选择油老虎车的人会获得

更高的收益，因为最初选择每种行动的人的数量相等，所以在第二期会有更多的人选择油老虎车。¹⁰如果再次应用复制者动态方程，那么选择油老虎车的博弈参与者的数量将会进一步上升。持续不断地应用复制者动态方程，最终将导致所有的人都选择油老虎车。图26-3显示的是对一个有100名博弈参与者的油老虎车博弈，运用离散复制者动态进行4个数值实验的结果。

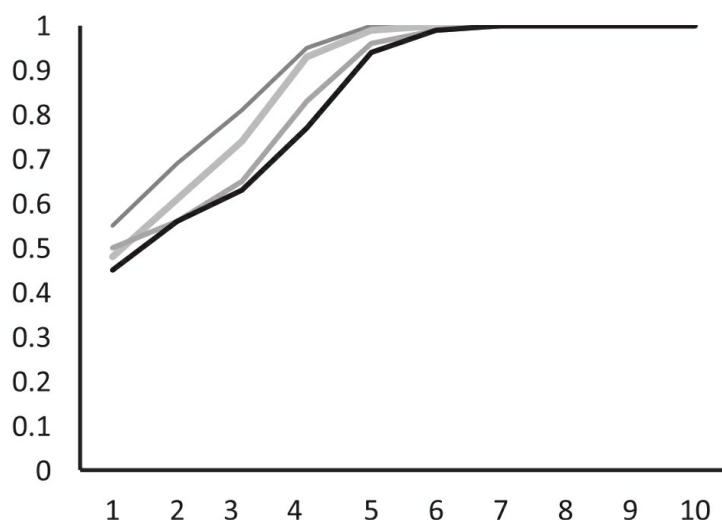


图26-3 油老虎博弈中的复制者动态（100名博弈参与者）：选择油老虎车的概率

由于假设了有限种群，所以要引入少量随机性，这样采用每个行动的比例可能不完全等于复制者动态方程所给出的比例。在这4个数值实验中，所有博弈参与者全都只在过了7期之后就都选择了油老虎车。收敛之所以如此迅速地发生，原因在于从众效应和奖励效应促使人们在第一期后尽快选择油老虎车。例如，当90%的人选择了油老虎车时，选择经济型汽车的人的收益将低于选择油老虎车的人的收益的1/6。从众效应极大地放大了奖励效应，使社会学习比个体学习更快。在个体学习中，平均来说要花费100多期才能达到99%的人都选择油老虎车的情况。

在这个博弈中，两个学习规则都收敛到了选择油老虎车上，这是因

为当两个行动的可能性相同时，选择油老虎车有更高的收益。这种行动被称为风险主导。两种学习规则都更“青睐”风险主导均衡而不是有效均衡。接下来，我们还要再构造一个博弈模型，在那里，这两个学习规则将收敛于不同的均衡。

慷慨 / 妒忌博弈

我们要讨论的下一个博弈，慷慨 / 妒忌博弈（the generous/spiteful game），建立在一个备受关注的行为倾向的基础上：我们更加关心的是绝对收益，还是相对收益？假设一个人在以下两个奖金分配方案中选择前者，那么他就是更关心绝对收益：（1）所有同事都能获得15 000美元的奖金，而他自己只能得到10 000美元；（2）所有同事都只能得到5 000美元的奖金，而他自己却能够得到8 000美元。相反，宁愿得到更少奖金（在上面两个奖金分配方案中选择后者）的人则更关心相对收益。这种更关心相对收益的偏好，在“居心险恶的人与魔法灯”的寓言故事中得到了很好的体现。

居心险恶的人与魔法灯

在一次考古探险中，一个居心险恶的人发现了一盏青铜灯。他擦了一下灯，结果召唤出了一个精灵。精灵说：“我会赐予你一个愿望，因为我是一个仁慈的精灵。我可以给你想要的任何东西！不但如此，对于你认识的每一个人，我都将给予他们给你的两倍。”这个居心险恶的男人仔细想了一会，然后抓起了一根棍子，递给那个精灵，说：“好吧。现在请你抠出我的一只眼睛吧。”

这个居心险恶的人所采取的这个行动，在带给他一个很低的绝对收益的同时，又给了他一个高的相对收益。¹¹在外交事务中也会出现类似的紧张局势。新自由主义者认为，每个国家都希望最大化自己的绝对利

益，这可以用军事力量、经济繁荣和国内稳定来衡量。但是另外一些人，他们通常被称为新现实主义者，却认为各国更重视的是相对利益，即一个国家宁可只能获得较低的绝对利益，但是一定要比自己的敌人更强。在冷战高潮期间，新现实主义者肯尼思·华尔兹（Kenneth Waltz）这样写道：“各国首先关心的不是最大化自己的力量，而是要维持他们在国际体系中的地位。”¹²

我们可以将绝对收益与相对收益之间的这种潜在冲突嵌入到一个 N 人博弈中。在这个博弈中，存在两类行为，一种行为是“慷慨大度”的，它在增加行为者自己绝对收益的同时，也会增加其他人的收益；另一种行为则是“妒忌狭隘”的，它只会增加行为者自己的收益。这个博弈不同于集体行动博弈。在集体行动博弈中，慷慨大度是要付出成本的。¹³在这个模型中，慷慨的行动是占优策略：无论其他博弈参与者采取的行动是什么，选择慷慨行动的博弈参与者都能获得更高的收益。然而，选择妒忌行动的博弈参与者得到的收益却比选择慷慨行动的博弈参与者更高。

这些陈述，乍一看似乎是自相矛盾的，其实不然。通过选择慷慨的行动，这个博弈参与者能够将自己的绝对收益提高3，而且同时也会将所有其他博弈参与者的收益提高2。而选择妒忌行动的博弈参与者则只能使自己的收益提高2，而且不能提高其他博弈参与者的收益。因此，每个博弈参与者都能通过选择慷慨的行动来提高自己的收益。相反，当一个博弈参与者选择妒忌的行动时，他反而会降低自己本来可以得到的收益，而且（这是关键假设）会使其他人的收益下降得更多。

慷慨 / 妒忌博弈

在这个博弈中，有 N 个博弈参与者，每一个博弈参与者都要选择：是慷慨（G），还是妒忌（S）。

$$\text{收益} (G, N_G) = 1 + 2 \times N_G$$

$$\text{收益}(S, N_G) = 2 + 2 \times N_G$$

如果我们在慷慨 / 妒忌博弈中应用强化学习规则，那么博弈参与者们都会学会做一个慷慨的人。为什么会这样？要理解这一点，不妨假设博弈参与者处于几乎已经收敛到均衡的状态了， N_G 个博弈参与者都选择了慷慨的行动。这时，如果有一个妒忌的博弈参与者，那么他可以获得 $2 + 2 \times N_G$ 的收益。这将成为他的渴望水平。如果他选择的是G（哪怕发生概率很小），那么他的收益为 $1 + 2 \times (N_G + 1) = 3 + 2 \times N_G$ ，这比他的渴望水平还要高。因此，他将更有可能变成慷慨的。不断应用这种逻辑，我们就会发现所有的博弈参与者都将学会慷慨。

但是，如果我们在慷慨 / 妒忌博弈中应用复制者动态，人们就会学会做一个妒忌的人。只要仔细观察一下复制者动态方程，就可以得出这个结论。在每一期，选择妒忌行动的那些博弈参与者所获得的收益都要高于选择慷慨行动的博弈参与者。因此，在每一期，选择妒忌行动的博弈参与者的比例都会上升。

这些结果突显了个体学习与社会学习之间的关键差异。个体学习会引导人们选择更好的行动，因此人们会学会采取占优行动（如果占优行动存在的话）。而社会学习则引导人们选择相对于其他行为来说表现更好的行动。在大多数情况下，这些行动通常也会产生更高的收益。但是，在慷慨 / 妒忌博弈中却并不是这样。在这种模型中，妒忌行动的平均收益更高，但是占优的却是慷慨行动。特别需要注意的是，我们的分析得出了一个悖论性质的结果：如果人们进行的是个体学习，那么他们就能够学会做一个慷慨的人——比通过社会学习能够学会的更加慷慨。之所以会出现这种情况，是因为在社会学习中，博弈参与者会复制表现相对较好的博弈参与者的行动。

现在考虑一下之前的观点：我们可以将复制者动态视为一种适应性规则，或者视为发生在若干固定规则之间的选择。如果假设了后者，那

么我们的模型就意味着选择可能有利于妒忌这种类型。（自然）选择不一定会导致合作。这个结果与我们在研究重复囚徒困境博弈时发现的结果背道而驰——在那里，重复导致了合作。我们在那里考虑的是重复博弈，并允许更复杂的策略。

将不同的学习模型组合起来

我们已经看到了个体学习模型和社会学习模型都能在一组固定的备选方案中找到最优解决方案，但当把它们应用于博弈时，也可能产生不同的结果。缺乏协议也可以是一种力量。不妨想象一个由所有可能的博弈组成的巨大集合。再想象一个由所有学习模型组成的巨大集合。我们可以将第二个集合中的每个学习模型应用于第一个集合中的每个博弈，并评估它们的表现。然后我们可以将这个包括了所有博弈的集合划分为两个集合：学习规则产生了有效结果的博弈的集合，以及不能得到有效结果的博弈的集合。我们还可以考察实验数据并评估每个学习规则在作为对实际行为的预测器时表现如何。毫无疑问，这种研究将会揭示一些我们未曾预料过的东西。每个学习规则都会在某些博弈中带来有效的结果，但是对其他一些博弈则不能。同时，每个学习规则本身在准确描述了行动的不同情况下也会有所不同。因此，我们提倡多模型思维。

在本章中，我们介绍了两个标准的学习模型。每一个模型都只包括了少数几个会变化的组成部分。我们的目标是对这些令人兴奋的文獻给出一个适中的介绍。只要往这两个标准模型中加入更多细节，就能够更好地拟合实验数据和经验数据。请读者回想一下，在强化学习模型中，个体会根据一个已采用过的行动或备选方案的奖励（或收益）是否超过了渴望水平来加大或减少该行动或备选方案的权重。个体不会增加从未采取过的行动的权重：在强化学习模型中，我们不会因为假想采取了某个行动会带来很高的收益，而提高采取这种行动的概率。

在所有情况下，这种假设都没有意义。假设一个员工决定在休假时

不带手机。当他去度假时，他的老板打电话要他解决一个重要问题，这个员工没有接到这个电话，并因此错过了一个升职机会。在强化学习模型中，员工不会给“度假时带手机”这个行动赋予更大的权重。有鉴于此，人们对标准模型进行了修正，提出了罗斯-伊雷夫学习模型（Roth-Erev learning model），让未被选择过的备选方案也可以根据其假想的收益来获得权重。在这个例子中，这个员工将会给“度假时带手机”赋予更大的权重。

这个修正导致了基于信念的学习规则。未被选择过的备选方案权重的增加量可以通过一个实验参数来确定。实验参数越高，人们对其他人行为的影响的考虑就越多，对那些行为赋予的权重也增加得越多。经济学家埃尔文·罗斯（Alvin Roth）和伊多·伊雷夫（Ido Erev）还考虑到，其他博弈参与者也在进行学习，他们的策略也可能在发生变化，因此还对过去进行了贴现处理。¹⁴

这些额外的假设具有其直观意义，并且都得到了经验证据的支持，但是它们并不适合于所有情况。如果回到前面举过的制作薄饼的例子，那么第一个假设意味着在父母制作好了香蕉薄饼之后，还要赋予制作苹果薄饼的备选方案额外的权重，而且该权重要与苹果薄饼的收益成比例。只有当父母知道苹果薄饼的收益时，这样的假设才是有意义的。但是，只有当人们能够观察到或凭直觉感知未被选择行动的收益时，才会出现这样的情况。

第二个修正模型来自行为经济学家科林·凯莫勒（Colin Camerer）和何（Ho）。他们构建了一个通用的函数形式，把强化学习和基于信念的学习都作为特殊情况包括了进去。这个函数的关键是一个可以用数据拟合的、确定每种类型学习规则相对强度的参数。¹⁵将多个模型组合在一起，正是我们学习掌握许多模型的一个重要动机。也就是说，由于参数的增加，组合模型必定能够导致更好的拟合。即便考虑到了参数增多这个因数，凯莫勒和何的模型也能给出更好的预测和更深刻的解释。

对学习建模带来了一些挑战。在一个模型设置中运行良好的学习规则可能完全无法适用其他情况。此外，人们学习的东西可能取决于他们最初的信念，因此两个人可能在同一个环境中以不同方式学习，同一个人也可能在不同的环境中以不同的方式学习。即便我们真的构建出了一个准确的学习模型，也会遇到可利用性原则（exploitability principle）带来的难题：如果一个模型解释了人们如何学习，那么其他模型就可以应用这个模型来预测相关知识，并在某些情况下利用该知识。这样一来，人们就可能会学会如何不会被利用，从而使我们原来的学习模型不再准确。在本书前面的章节中，当我们讨论卢卡斯批判和对有效市场假设的分析时，我们就已经遇到过这种现象了。我们不一定能得出结论说那是因为人们会了解到他们在最优化，然而，学习毕竟倾向于淘汰不良行为、从而有利于更好的行为。

文化能否压倒战略

我们现在将传染模型和学习模型结合起来，以便剖析组织理论中由来已久的一个理论观点：文化压倒战略。¹⁶简而言之，这个观点声称，改变行为的战略激励终将归于失败。理论组织家强调，文化——即现有的既定规则和信念的力量实在太强大了。经济学家的观点则相反：推动行为的，只能是激励。

为了将这些相对立的谚语式诊断转变成条件逻辑判断，我们首先必须应用网络传染模型的一个变体。在这个模型中，经理，或者也可能是CEO，宣布了一个新战略，并给出了推动变革所能带来的好处的多项证据。这位经理或CEO甚至可能会对组织的核心原则加以重新界定，以便反映这种新行为的要求。然后，组织中的其他个体决定是否采取这种行为，这取决于经理或CEO对其战略的说服力有多大。一开始，只有一部分人执行这个计划。当他们在工作网络中与他人互动时，就会热情洋溢地传播新战略。当然新战略也会面

临挑战，会有一种反向的力量拉动人们不去采用新战略。有三个特征决定了新策略能否顺利展开：接触率（ P_{contact} ）、扩散率（ P_{spread} ）和放弃率（ P_{recover} ），它们很自然地映射到了基本再生数中的参数 R_0 上，即：

$$R_0 = \frac{P_{\text{spread}} \times P_{\text{contact}}}{P_{\text{recover}}}$$

如果再加入存在超级传播者的可能性，就可以得出这样的结论：只要如下三个条件中任何一个条件成立，文化就会压倒战略，否则，战略就能压倒文化。这三个条件分别是：如果人们不相信新战略，如果人们很快就放弃了新战略，如果新战略的拥护者相互之间的连通性不够好。

我们的第二个模型是，将复制者动态方程应用于这个用来表征员工之间互动的文化战略博弈。我们可以将员工的不同选择用博弈论的语言分别表示为文化行动（做他们目前所做的事情）和创新的战略行动。我们还假设，经理或CEO已经确定了收益结构，如果两个博弈参与者都选择创新的话，他们都能获得更高的收益；但是，如果只有一个博弈参与者选择创新，那么他的收益将会减少。

	文化	创新
文化	200, 200	220, 180
创新	180, 220	300, 300

文化 / 战略博弈

这个博弈有两个严格的纯策略纳什均衡：一个是两个博弈参与

者都创新（战略胜过了文化），另一个是两个博弈参与者都不创新（文化胜过了战略）。乍一看，经理或CEO似乎已经给出了足够大的激励，能够保证员工会选择创新的行动。但是通过分析，我们发现，经理或CEO必须动员起足够多的初始支持者才能使创新成为现实。如果一开始就支持新战略的人的比例没有超过20%，那么文化就会胜过战略。如果要增加创新战略的收益，那初始支持者的比例可能会更低，但仍然会产生有效的结果。¹⁷

这两个模型表明，字面上相反的两个谚语“文化压倒战略”和“人们会对激励做出反应”都是正确的。根据第一种模型，具有很高人格魅力的CEO可以制订能够胜过文化的新战略。根据第二种模型，文化能胜过“弱激励”，但是不能胜过“强激励”。

27 多臂老虎机问题

有一件事我确实特别擅长，那就是将网球击过网，打在界内。
在这件事情上，我是最棒的。

塞雷娜·威廉姆斯（Serena Williams）

在本章中，我们在如何学会选择最优备选方案的学习模型中加入不确定性，从而生成了一类被称为“多臂老虎机问题”（multi-armed bandit problems）的模型。在一个多臂老虎机问题中，不同备选方案的奖励源于一个分布，而不是固定的金额。多臂老虎机问题模型适用于各种各样的现实环境。在收益不确定的行动之间进行的任何选择，无论是药物试验，还是对树立广告牌位置的选择、技术路线的选择，抑或是要不要允许在教室中使用笔记本电脑的决定，都可以建模为多臂老虎机问题。当然，如何选择一个可以出人头地的职业，也可以用多臂老虎机问题模型来建模。¹

在面对一个多臂老虎机问题时，人们必须对各个备选方案多加尝试，以便通过这种学习过程来了解收益的分布。多臂老虎机问题的这个特征，也导致我们必须在探索（寻找最佳备选方案）和利用（选择迄今为止表现最佳的备选方案）之间善加权衡。在探索与利用权衡中找到最优平衡点，需要非常精妙复杂的规则和行为。²

本章的主体内容分为两个部分，最后对模型的应用价值进行了讨论。在本章的第一部分，我们描述了一类特殊的伯努利多臂老虎机问题，其中每个备选方案都是一个伯诺利瓮（瓮中灰球和白色球的比例是未知的）。我们描述并比较多种启发式求解方法，然后说明这些解是如何有助于改进药物疗效的比较检验、广告计划和教学策略的。在第二部分中，我们描述了一个更一般的模型，其中收益分布可以采取任何形

式，并且决策者对其类型有一个先验分布。我们还阐明了如何求解确定最优选择的吉廷斯指数（Gittins index）。

伯努利多臂老虎机问题

我们从一类特殊的多臂老虎机问题开始讨论。在这类多臂老虎机问题中，每个备选方案都能以固定的概率产生成功的结果。因此，这类多臂老虎机问题相当于在一系列伯努利瓮之间进行选择，且每个瓮都包含着不同比例的灰球和白球。因此，我们将这类多臂老虎机问题称为伯努利多臂老虎机问题，也经常被称为频率问题，因为决策者对分布一无所知。不过，当决策者对各个备选方案进行了多次实验（探索）之后，他会对这些分布有所了解。

伯努利多臂老虎机问题

一个备选方案集 $\{A, B, C, D, \dots, N\}$ 中的每一个备选方案都能够产生一个成功的结果，但是各自的概率 $\{P_A, P_B, P_C, P_D, \dots, P_N\}$ 都是未知的。在每一个时期，决策者选择一个备选方案 K ，并以概率 P_K 得到一个成功的结果。

假设一家烟囱清洁公司获得了一批最近购买了房子的人的电话号码，然后打算向他们推销烟囱清洁服务。这家公司测试了三种推销策略。第一种策略是“笃定预约式”（“你好，我打电话来是为安排你家每年一度的烟囱清洁。”）；第二种策略是“关心提问式”（“你好，你知道烟囱堵塞是火灾的最大风险因素吗？”）；第三种策略是“人性感动式”（“你好，我的名字是希尔迪，我已经和我父亲一起打理这家烟囱清扫公司整整14个年头了。”）。

每一种推销策略都有可能成功，但是成功概率在事前是未知的。假设该公司首先尝试的推销策略是“笃定预约式”，但是失败了。然后又尝

试了“关心提问式”，结果成功地获得了一个客户；而且，这种策略紧接着又成功了一次。但是，在接下来的3次尝试中，这种策略都失败了。于是，该公司尝试了第三种策略。第三种策略的第一次尝试是成功的，但是接下来却连续失败了4次。这样，在总共进行了11次尝试之后，第二种推销策略的成功率是最高的，但是第一种策略只尝试了一次。于是，决策者面临着在利用（选择最有效的备选方案）或探索（回过头去继续尝试其他两种推销策略以获得更多信息）之间的权衡。医院在不同的外科手术方案之间的选择，制药公司对药物检验的不同方案的权衡，也都会碰到同样的问题。每一种“协议”都有未知的成功概率。

为了进一步深入理解这种探索-利用权衡，我们比较了两种启发式。第一种启发式是取样并择优启发式（sample-then-greedy），即先对每个备选方案都尝试固定的次数 M ，然后选择具有最高平均收益的备选方案。而在确定尝试次数 M 大小的时候，我们可以参考伯努利瓮模型和平方根规则。平均比例的标准差有一个上界 $1/2\sqrt{M}$ 。如果每种备选方案都进行了100次的测试，那么平均比例的标准差将等于5%。如果应用两个标准差规则来识别显著差异，当两个比例相差大约10%时，我们就能够自信地将它们区分开来了。例如，如果一个备选方案在70%的时间内都取得了成功的结果，而另一种方法则只在50%的时间内取得了成功，那么就有95%以上的置信水平相信我们能够选中正确的备选方案。

第二种启发式称为自适应探测率启发式（adaptive exploration rate heuristic）。它的程序是，第一阶段，先让每种备选方案各完成10次试验。第二阶段，再进行总共20次试验，但是试验次数根据各备选方案在第一阶段的成功率按比例分配。例如，如果第一阶段的10次试验中，有一个备选方案成功了6次而另一个只成功了2次，那么第一个备选方案将获得接下来的20次试验中的3/4。到第三个阶段，可以根据成功率的平方确定本阶段要进行的20次试验的分配比例。如果那两个备选方案的成功率仍然与前面一样，那么更好的备选方案在第三阶段的20次试验中将

分配到 $\frac{(0.6)^2}{[(0.6)^2 + (0.2)^2]}$ 或90%。

以此类推。对于每一个阶段的20次试验，计算分配比例时所用的成功概率的指数可以以某种速率递增。随着时间的推移，通过提高利用率，第二种算法相比第一种算法有所改进。如果一个备选方案的成功概率比另一个备选方案高得多，比如80%对10%，那么第二种算法就不会在第二个备选方案上“浪费”100次试验。另一方面，如果两个备选方案的成功概率非常接近，就要继续进行试验。³

对于第一种启发式，取样并择优启发式，如果在使用时过分执着，那么不仅效率低下，而且可能是不道德的。当美国著名外科医生罗伯特·巴特莱特（Robert Bartlett）在测试人工肺时，发现它的成功率远远超过了其他备选方案。既然人工肺的表现已经如此优异了，那么继续测试其他备选方案就会导致不必要的死亡。于是巴特莱特停止测试其他备选方案，让每个患者都使用上了人工肺。事实上，可以证明这是一个最优规则：如果某个备选方案总能取得成功，那么就继续选择这个备选方案。增加实验可能没有任何价值，因为没有其他备选方案能够表现得更好。

贝叶斯多臂老虎机问题

在贝叶斯多臂老虎机问题中，决策者对各备选方案的收益分布有先验信念。考虑到这些先验信念，我们可以对探索与利用之间的上述权衡进行定量分析，并（至少在理论上可以）在每个时期都做出最优决策。然而，即便是对于最简单的多臂老虎机问题，要确定最优行动也需要进行大量的计算。在真实世界的实际应用中，要精确计算出结果是不可行的。因此决策者通常都会利用近似方法。

贝叶斯多臂老虎机问题

给定备选方案集 $\{A, B, C, D, \dots, N\}$ ，以及对应的收益分布 $\{f(A), f(B), f(C), f(D), \dots, f(N)\}$ 。决策者对每个分布都有先验信念。在每一期，决策者选择一个备选方案，并获得收益，并根据收益计算出新的信念。

要确定最优行动，需要经过如下四个步骤。首先，要计算出每个备选方案的即时期望收益。其次，对于每个备选方案，都要更新关于收益分布的信念。再次，在得到的关于收益分布的新信念的基础上，根据我们所掌握的信息确定所有后续时期的最优行动。最后，我们将下一期行动的期望收益与未来的最优行动的期望收益相加。最后得到的这个结果就是通常所称的吉廷斯指数。在每一个时期，最优行动的吉廷斯指数都是最大的。

这里需要注意的是，计算指数的过程同时也量化了探索的价值。而且，如果我们尝试某个备选方案，那么吉廷斯指数也不会等于期望收益。相反，吉廷斯指数等于假设根据所掌握的知识采取最优行动时，所有未来收益的总和。但是，计算吉廷斯指数非常困难。下面举一个相对简单的例子。假设备选方案有两个，一个是肯定能带来500美元的安全的备选方案，另一个是有10%概率可以带来1 000美元的有风险的备选方案，在其余90%的时间里，有风险的这个备选方案不会带来任何收益。

为了计算出这个有风险的备选方案的吉廷斯指数，我们首先要问清楚会发生什么：它要么总是收益1 000美元，要么总是没有任何收益。然后再思考每个结果将会怎样影响我们的信念。如果我们知道这个有风险的备选方案会让我们收益1 000美元，那么我们总是会选择它。如果我们知道这个有风险的备选方案没有任何收益，那么我们在未来将总是选择安全的那个备选方案。

由此可见，有风险的这个“臂”的吉廷斯指数对应于每个时期获得1 000美元奖励的概率为10%，以及除了第一期之外的每个时期获得500美

元的概率为90%。平均来说，要对备选方案进行多次选择的情况下，这相当于每一期大约550美元。因此，这个有风险的备选方案才是更好的选择。⁴

吉廷斯指数

为了说明如何计算吉廷斯指数，考虑下面这个只有两个备选方案的例子。备选方案A产生的收益抽取自{0, 80}，且0和80出现的概率相等。备选方案B在{0, 60, 120}当中产生一定的收益，而且这三个收益的概率也是相等的。我们假设，决策者试图最大化10个时期内的总奖励。

备选方案A：收益等于零的概率为1/2，在出现了这种结果之后，在剩下的全部9期内都会选择备选方案B（备选方案B的期望收益60），这样就得到了540的期望收益（9乘以60）。收益等于80的概率也为1/2，即便出现了这个结果，在第二期的最优选择仍然是选择备选方案B。于是有1/3的概率，备选方案B产生了120的收益，因此总收益等于1 160（80加上9乘以120）。同样，有1/3的概率，备选方案B产生了60的收益，在这种情况下，备选方案A是剩下的所有8期的最优选择，这样产生的总收益等于780（60加上9乘以80）。最后，还有1/3的概率，备选方案B产生了零的收益，在这种情况下，备选方案A是剩下的所有8期的最优选择，这样产生的总收益等于720（9乘以80）。

把上面这些可能性全都考虑进去，可以得出，在第一期，备选方案A的吉廷斯指数如下：

$$\text{Gittins index}_1(A) = \frac{1}{2} \times 540 + \frac{1}{2} \left(\frac{1}{3} \times 1160 + \frac{1}{3} \times 780 + \frac{1}{3} \times 720 \right) = \frac{2140}{3}$$

备选方案B：有1/3的概率，收益等于120。如果发生了这种情

况，那么所有未来时期的最优选择也仍然是备选方案B。因而10个时期的总收益将等于1 200。如果收益等于零（概率为1/3），那么所有未来时期的最优选择都将是备选方案A（备选方案A的期望收益为40），因而，期望总收益将等于360（9乘以40）。如果收益等于60，那么决策者在所有未来时期都应该选择替代方案B，总回报为600；但是，如果在第二个时期选择了备选方案A，那么有一半时间备选方案A总是产生80的收益，此时总回报为780（60加上9乘以80）；另一半时间它产生零收益，并且所有后续时期的最优选择都将是备选方案B（会产生60的收益），于是得到的总收益为540（9乘60）。由此可知，在第二期中选择备选方案A才是最优选择，这种选择产生的期望收益等于 $660 \times (1/2 \times 780 + 1/2 \times 540)$ 。

把所有可能性都考虑进去，不难推出，备选方案B在第一期中的吉廷斯指数如下：

$$\text{Gittins index}_1(B) = \frac{1}{3} \times 1200 + \frac{1}{3} \times 660 + \frac{1}{3} \times 360 = \frac{2220}{3}$$

根据这些计算结果，备选方案B是第一期的最优选择。最优长期选择则取决于第一期内学习的结果。如果备选方案B产生了120的结果，那么我们将永远坚持选择备选方案B。

上述的分析表明，在采取行动时，我们更关心的是备选方案能够成为最优选择的概率而不是它的期望收益。此外，如果某个备选方案会产生非常高的收益，我们应该更有可能在将来选择它。相反，如果它只能产生平均收益，那么即便收益水平高于另一种备选方案的期望收益，我们一直坚持这个备选方案的可能性也不会太高。在尝试的早期阶段尤其如此，因为我们希望找到更高收益的备选方案。这些结果在我们讨论的诸多应用中都是成立的。如果采取行动没有风险，也不需要付出高额成本，那么这个模型告诉我们，即使高收益行动的概率很低，我们也要努力去探索它们。

小结

本书一再强调的一个核心要点是，通过学习模型，我们可以做出更好的决策。在对人们的多臂老虎机问题中“应该”做些什么与人们“实际”做了些什么进行一番比较之后，我们可以对这一点有更深刻的理解。绝大多数人在遇到多臂老虎机问题时，都不会去试图计算一个吉廷斯指数。之所以没能这样做，部分原因是他们没有将必要的数据保存下来。例如，直到最近，医生才开始记录各种手术方法的效果，比如不同类型的人工关节的功效，或者不同类型的心脏支架的优缺点，等等。没有这些数据，医生就无法确定哪一种手术方法会带来最高的期望收益。

其实是所有人，都需要足够的数据，这样才能把模型教给我们的东西应用起来。因此，如果你真的想了解晚餐前散步还是晚餐后散步对你的睡眠更有利，你就需要跟踪记录你的睡眠状况，并运用一些相当复杂的启发式去了解哪种散步模式效果更好。初看上去，这可能会显得小题大做，而且要付出的时间精力相当可观，的确如此，但是现在已经好很多了。新技术的不断涌现，使我们能够非常方便地收集有关睡眠模式、脉搏率、体重，甚至情绪好坏的大量数据。

我们每个人都要做出很多与自己的身体健康息息相关的决策，比如什么时候去锻炼，但是绝大多数人并没有去收集必要的数据并计算出吉廷斯指数。但这其实非常重要，关键在于我们是否能够做到这一点，而且如果做到了，我们的睡眠模式、身体健康状态就会得到改善。心理学家塞思·罗伯茨（Seth Roberts）探索了整整12年，结果发现自己每天至少站立8小时才可以改善自己的睡眠状态（尽管他还是睡得更少）。他还发现，迎着早晨的阳光站立，可以减轻他上呼吸道的感染症状。⁵当然，我们一般人可能很难具备他这种用自己身体来做实验的奉献精神。但是，由于不保存数据，也不对相关结果进行比较，我们可能会更容易放任自己不吃早餐或暴饮暴食——尽管我们有更好的选择，比如吃西柚。

在利害关系很大的商业决策、政策制定和医疗决策中，数据更容易收集，应用多臂老虎机问题模型也早就成了一种常见的做法。企业、决策者和非营利组织，都会先对各种备选方案进行探索，然后利用那些表现最好的备选方案。而且在实践中，备选方案往往不会保持固定不变。例如，鼓励参加农业补贴计划的政府邮件可能每一年都会改变，比如将上一年的强健男子的照片换为性感美女的照片，等等。⁶这种类型的连续实验可以通过将在下一章中讨论的模型来刻画，那就是：崎岖景观模型。

用不同模型分析美国总统选举

我们可以应用至少三种模型来分析美国总统选举：空间竞争模型、分类模型和多臂老虎机问题的模型。

空间竞争模型：民主、共和两党的总统候选人在意识形态空间中互相竞争以吸引选民。我们有理由预计，各候选人倾向于温和的中间立场、选情会比较胶着，不同政党候选人获胜的顺序是随机的。除了少数例外情况外，总统选举不会以某一方压倒多数获胜结束。为了检验美国总统大选获胜者的顺序是不是随机的，我们构建了从1868年到2016年38次总统大选的获胜政党的时间序列。该序列如下（字母R、D分别表示共和党、民主党）：

RRRRDRDRRRRRDDRRRRDDDDDDRRDDRRDRRRDDRRDDR

然后我们可以计算出不同长度的子序列（块）的熵。长度为1的子序列的熵为0.98。长度为4的子序列的熵为3.61。统计检验表明，我们不能否认这个序列是随机的。作为比较，在长度为38的随机序列中，长度为1的子序列的熵为1.0，长度为4的子序列的熵为3.58。

分类模型：如果我们将每个州视为一个类别，同时假设不同州

之间存在着异质性，那么空间竞争模型意味着一旦候选人选定了初始位置，某些州就不再具有竞争力了。这个模型的预测是，在少数几个立场温和的州，选举竞争将特别激烈。2012年，奥巴马和罗姆尼都在10个州花掉了自己电视广告预算的96%以上。他们每个人都将广告预算的一半多用于3个温和的州：佛罗里达州、弗吉尼亚州和俄亥俄州。2016年，希拉里·克林顿和特朗普也将一半以上的电视广告预算花在了3个温和的州：佛罗里达州、俄亥俄州和北卡罗来纳州。⁷

多臂老虎机问题的模型（回溯性投票）：选民将更有可能将选票再一次投给有良好执政业绩的那个政党。给绩效好的政党投票，相当于拉一个会带来高收益的杠杆。经济繁荣通常会使竞选连任者受益。有证据表明，当经济表现良好时，选民更有可能投票给执政党的候选人。而且，在执政党内部，现任候选人的影响也大于非现任候选人。⁸

28 崎岖景观模型

当你费心去寻找时，就会发现令人惊奇的事情。

传为萨卡加维亚（Sacagawea）所说

在本章中，我们研究崎岖景观模型。与空间竞争模型和享受竞争模型一样，崎岖景观模型也将一个实体定义为属性的集合。每个属性的集合都映射到一个价值上。崎岖景观模型的目标是修改属性，以构造出一个具有最高价值的实体。这类模型起源于生态学中对演化的研究。现在，崎岖景观模型已经广泛用于探索各种问题的求解方法、研究企业之间的竞争和创新，以及其他领域，这也是我们在本书中要研究的重点。在本章中，我们将应用崎岖景观模型揭示，属性影响的相互依赖性如何使创新变得困难、导致所找到的解决方案呈现出路径依赖性、并且还造成了解决方案本身的多样化。同时，我们也会阐明，许多更困难的问题是怎样通过更加多样化的问题求解方法而得以解决的。

本章由三个部分组成，然后，我们讨论了如何扩展模型以刻画竞争。在第一部分中，我们先描述了一个适合度景观模型（fitness landscape model），然后阐明了怎样将它重新解释为一个关于问题求解和创新的模型。在第二部分中，我们讨论了一维模型中崎岖度的含义。在第三部分中，我们提出了崎岖景观的NK模型，它将一维模型扩展到了任意数量的二元维数。

适合度景观模型

适合度景观模型假设物种拥有能够促进其适合度的特征或性状，我们可以不那么严格地将之定义为繁殖潜力，同时种群中不同成员所拥有

的特定性状的数量或程度可能不同。如果用横轴表示性状，用纵轴表示物种的适合度，就可以绘出一张适合度景观的图，其中高海拔点对应高适合度。

举例来说，为了绘制出一张对应于土狼尾巴长度性状的适合度景观的图，我们应该令土狼的所有其他性状都保持不变，而只改变尾巴的长度，并测量尾巴长度的变化对适合度的影响。这就是说，要绘制出这张图，我们必须先了解土狼尾巴为什么有助于提高它的适合度。

假设一只土狼的尾巴有助于土狼在跳跃时保持平衡，而且土狼可以将它作为表示幸福、恐惧或即将发动攻击的信号。我们从横轴的最左侧开始，在那里，尾巴长度为零，这种情况下它不能执行任何一种功能，因此它的适合度为零。随着尾巴长度的增加，维持平衡和传递信号的功能也随之提高。因此，适合度先是随尾巴长度的增加而上升的。

但是，到了某一点上，比如当尾巴长到18英寸时，可能就是有助于土狼保持平衡的最理想长度。如果尾巴变得更长，土狼运动的敏捷度将会下降。不过，更长的尾巴可能还会继续提高它传递信号的价值，因此，长度为20英寸的尾巴可能会产生最大的整体适合度。一旦尾巴的长度超过了20英寸，适合度就会开始下降。结果如图28-1所示，它具有一个单峰。

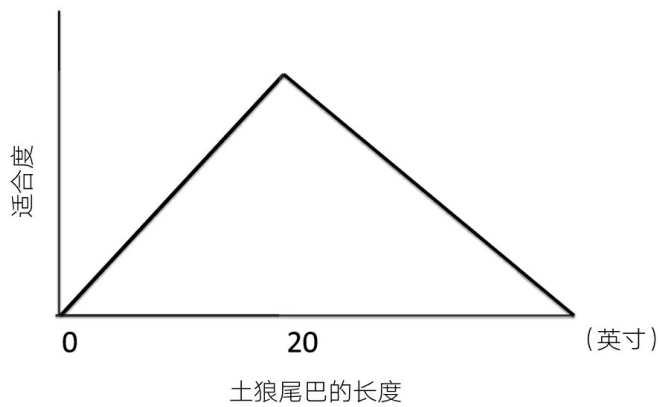


图28-1 富士山景观

这种单峰景观被称为富士山景观。在现实世界中，这样的景观是经常出现的。有富士山景观的问题通常是比较容易解决的问题。我们可以指望演化或学习在遇到这类问题时找到这个山峰。不妨想象一下，假设有一个由尾巴长度各不相同的土狼组成的种群，那么自然选择的压力最终会导致土狼的尾巴长到大约20英寸。拥有这个长度的尾巴的土狼，能够同时将维持身体平衡和传递信号这两个任务完成得最好。因此，这样的土狼的适合度最高，能够留下最多的后代，从而导致更多的土狼拥有20英寸长的尾巴。如果我们认为这是一个优化问题，那么任何一个爬山算法都能找到这个山峰。

我们也可以应用一对多的思维方法，将这个问题重新解释为一个产品设计问题，比如设计一把煤铲。假设我们已经决定了煤铲的长度和形状，只剩下煤铲的大小有待决定。于是，煤铲的面积就是要反映在横轴上的特征。而在纵轴上，我们给出的是一个工人在给定煤铲大小的前提下，一个小时内能够铲煤的数量。

与之前的例子一样，我们仍然从最左侧开始，那一点对应于面积为零的煤铲。面积为零的煤铲其实只是一根棍子。当然，用一根棍子去铲煤是完全无效率的，其价值为零。我们逐渐增大煤铲的面积，先是一茶匙大，接着是一汤匙大，然后是像玩具铲那么大……煤铲就变得越来越有效。在这个过程中，煤铲的适合度图形是向上倾斜的。然后，到了某个点上，当煤铲的面积变得太大之后，要用它去铲煤就变成了一件苦差事。一个小时内可以铲煤的数量就会随煤铲的面积进一步增大而减少。最后，当煤铲面积变得足够大时，将没有任何人能够抬起这把煤铲，因而它的适合度又一次变为零。我们又看到了富士山景观。我们有理由期待肯定能够找到山峰，也就是要设计的煤铲的理想面积。

事实上，将铲子的效率作为铲子面积的函数，以此来确定最优铲子大小的思想，正是著名经济学家、管理学家弗雷德里克·泰勒（Frederick Taylor）提出的。在19世纪90年代，泰勒和其他一些人开创了科学管理

的新纪元。在泰勒生活的那个时代，制造业决策，比如流水装配线的移动速度有多快，焊接强度要多高，工人的休息时间是多少，等等，都被建模为崎岖景观问题。进入20世纪后，许多伟大的实业家，例如福特汽车公司创始人亨利·福特（Henry Ford）、“石油大王”约翰·D. 洛克菲勒（John D. Rockefeller）和“钢铁大王”安德鲁·卡内基（Andrew Carnegie）都为这个现在用泰勒主义一词来概括的运动做出了卓越的贡献。

从制造个性化的、“只此一家别无分店”产品的工匠生产，转为大规模制造，是一个重大的变革。在大规模生产中，制造流程要分解为多个部分，每个部分都要经过优化，然后变成常规操作。这样一来，效率得到了大幅提高。但是在许多人看来，这个过程也是劳动的非人化过程。这种分歧就是一个提示：我们需要多个模型。任何单一模型都是对世界的过分简化，只能突出其中的某些维度。科学管理模型侧重于流程效率。这种偏向导致了批评。以产出效率为准则做出的决策，会导致其他目标遭到忽视，例如工人的快乐和福祉。

从表面上看，景观模型似乎只是一个相当浅显的想法：将适合度、效率或价值作为特征或性状的函数绘制在图上，然后爬上山顶，找到那个特征或性状的最优值。而且，把解决问题想象为“爬山”，似乎也不过是一个简单的比喻。这当然都是有效的批评。但是，如果构建了正式的景观模型，我们将能够得到一些非凡的结论。

崎岖景观

当我们同时考虑多个属性并且允许一个属性的贡献与其他属性的贡献相互作用时，就会得到一个崎岖景观，也就是具有多个山峰的景观。考虑一个设计沙发的问题，我们必须决定坐垫的厚度和扶手的宽度。我们用沙发在市场上的预期销售额来代表设计的价值，而沙发的销售额与设计的美感相关。如果沙发有厚厚的垫子，那么较宽阔的扶手可能会使

沙发更具美感。如果沙发的垫子很薄，那么扶手窄一点会更好。作为扶手宽度和坐垫厚度的函数，预期销售的二维图将具有两个山峰。一个山峰对应于窄扶手、薄垫子的沙发设计；另一个山峰则对应于宽扶手、厚垫子的沙发设计。

变量之间的相互依赖效应，使得景观出现了崎岖的特点。这种崎岖性有好几个重要含义。首先，在崎岖景观中寻找到最高点时所用的不同方法，可能会以找到不同的山峰而告终。如果从不同的起点出发，也可能会找到不同的山峰。因此，崎岖性导致了对初始条件的敏感性和路径依赖的可能性。而这些都意味着，景观的崎岖性有助于结果的多样性。崎岖性也意味着出现次优结果的可能性，在崎岖景观中，次优结果表现为局部高峰。

图28-2显示了一个有5个山峰的崎岖景观。在这些山峰中，有4个是局部高峰，它们只是比与它们相邻的点的值高一些，只有一个是全局高峰，即具有最高值的点。要理解搜索是怎样止步于依赖初始搜索点的局部高峰的，可以想象从一个点开始往山峰爬的过程。这种过程被称为梯度启发式（gradient heuristic）或爬山算法（hill-climbing algorithm）。在崎岖景观中，梯度启发式技术会“卡”在局部高峰上。

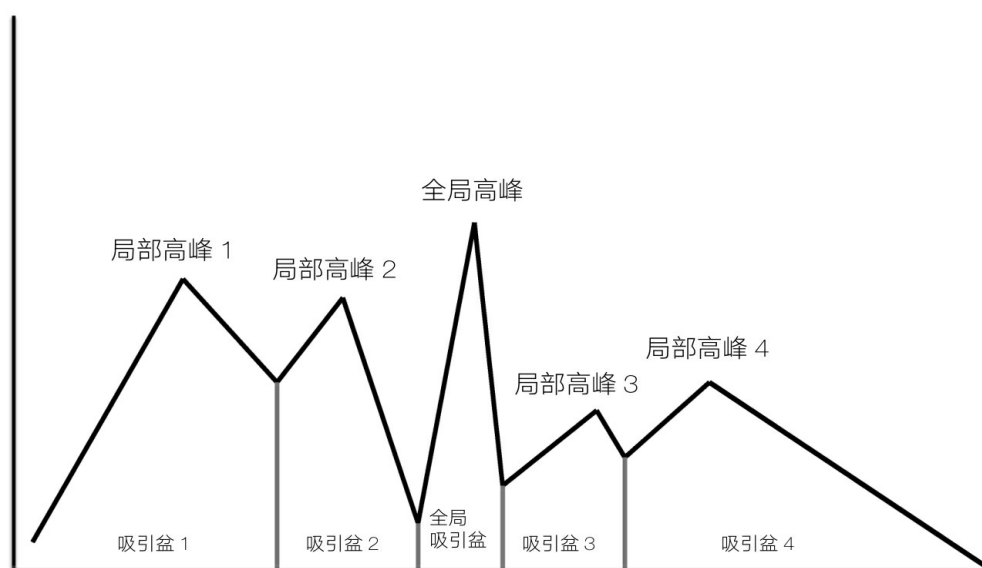


图28-2 一个有5个山峰的崎岖景观

如果起点位于最左侧，那么梯度启发式将定位于局部高峰1上，但它不是最优的。如果梯度启发式从图28-2中标识为“吸引盆2”的区域开始，那么它将定位于局部高峰2上。其他每个山峰，包括全局高峰，都有这样一个区域：如果梯度启发式从那个区域开始，就会找到那个局部高峰。这些区域被称为吸引盆（basin of attraction），如图28-2所示。从图中可以看出，全局高峰吸引盆的面积是最小的。如果我们随机选择一个起点并应用梯度启发式，那么全局高峰恰恰是最不容易被找到的那一个山峰。

吸引盆取决于启发式。如果我们使用了不同的启发式，就可能得到不同的吸引盆。例如，我们也可以不用梯度启发式，转而使用一个名为“一直向右走”的启发式。这个启发式一直向右侧搜索，直到找到一个局部高峰为止。对于这个例子，这两种启发式具有相同的局部高峰但却会产生不同的吸引盆，只要比较一下图28-3和图28-2，就可以看出这一点。

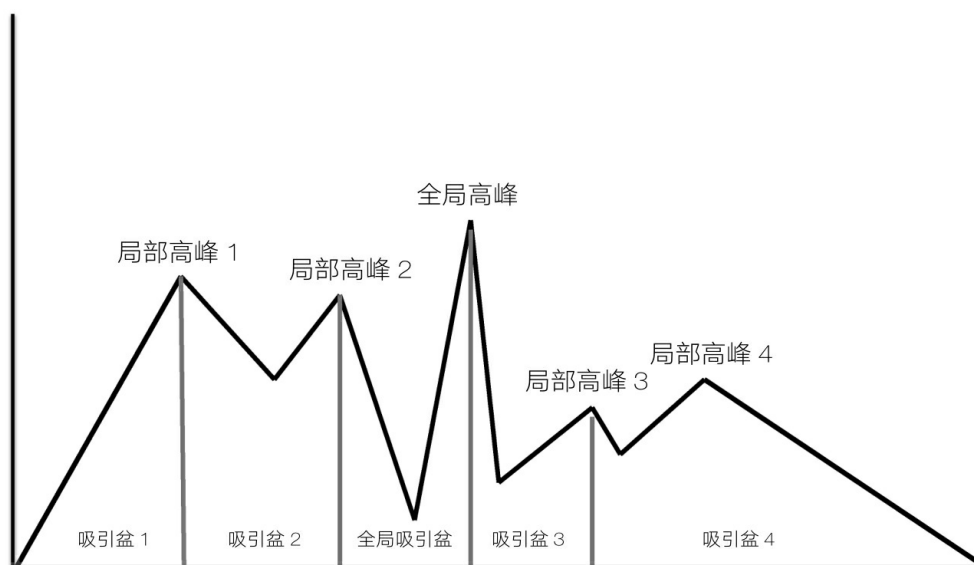


图28-3 由“一直向右走”启发式产生的吸引盆

要在崎岖景观中找到最优或接近最佳的高峰，需要多样性和复杂性的方法。多样性的价值是不言而喻的，如果不同的启发式能够找到不同的峰值，那么对一个问题应用多个不同的启发式就能够产生多个不同的局部峰值，然后就可以从这些局部峰值中选择一个最优的。¹或者换一个思路，如果从不同的起点应用同一个启发式，那么也可以得到相同的结果：先找到若干个不同的局部最优点，然后选择其中最优的。

另外还应该注意到，景观的崎岖度（以山峰的数量来衡量）与问题的难度相关。当然，有的问题很难解决，但却不一定有崎岖的景观。在一大块玉米地里找一枚金币这个问题可以用一个平坦的景观来表示，只是在硬币的位置有一个单峰。它的景观不是崎岖的，但是要找到硬币确实很难。

NK模型

现在描述NK模型。利用这个模型，我们可以对属性的相互作用与崎岖性之间的联系加以形式化。²这个模型将对象，或我们这里所称的备选方案，表示为一个长度为 N 的二进制字符串，这就是NK模型中“ N ”的含义。至于“ K ”，则指与该字符串的每一位交互以确定这一位的值的其他位的数量。如果 K 等于零，那么价值函数就是线性的。如果 K 等于 $N-1$ ，则所有的位都相互交互，每个字符串的值都是随机的。因此，我们可以考虑增大 K ，将景观的崎岖度调整到富士山景观与随机景观之间的某个适当位置。

NK模型

一个对象由 N 位二进制符号组成， $s \in \{0, 1\}^N$ 。

这个对象的价值表示为 $V(s) = V_{k1}(s_1, \{s_{1k}\}) + V_{k2}(s_2, \{s_{2k}\}) + \dots + V_{kN}(s_N, \{s_{Nk}\})$ ，其中， $\{s_{ik}\}$ 等于一个从原字符串中除

了第 i 位之外随机选择出来的有 k 位的子字符串的集合，而且 $V_{k1}(s_1, \{s_{1k}\})$ 是从区间 $[0, 1]$ 中抽取出来的一个随机数。

$K=0$ ：得到的是一个关于位的线性函数。

$K=N-1$ ：任何位的变化都使每个位产生新的随机贡献。

NK模型作为一个框架，为探索想法和提出问题创造了一个非常好的空间。我们想要问，局部最优值的数量是怎样取决于交互项的数量。我们还可以问，全局最优值的大小又是如何取决于交互项的数量。不过眼下，还不是回答这两个问题的适当时机，因为还没有定义好怎样去搜索可能性空间，这就是我们正在使用的启发式，局部最优集及其吸引盆的集合取决于搜索时所用的启发式。

在下文中，我们主要依赖单翻转算法（single flip algorithm）来进行搜索。这个算法按顺序选择每一个属性并切换该属性的状态。如果改变那个属性会产生一个更高的价值，就保留这个切换。否则，那个属性就退回到它原来的状态。之所以选择这个算法，有两方面的动机。首先，我们可以把它解释为描述基因突变的一个粗略模型，让好的基因变异逐步接管整个种群，并让坏的基因变异消失。其次，它也是在我们这个空间中表征爬山算法最自然的一个形式。

作为例子，我们先来求解 $N=20$ 且 $K=0$ 时的NK模型。当 $K=0$ 时，每个属性对总价值的贡献独立于其他属性，这时单翻转算法可以识别出每个属性更好的状态和全局最优值。因此， $K=0$ 意味着不存在相互作用，这种情况对应于富士山景观。每个状态的值均匀分布在区间 $[0, 1]$ 中。不难证明，从区间 $[0, 1]$ 上的均匀分布中随机抽取出来的两个值中较高的那一个期望值为 $2/3$ 。只要我们对这20个属性的贡献求平均值，就不难推出全局最优值的期望值也是 $2/3$ 。

而在另一个极端，当 $N=K-1$ 时，每一个属性都与其他每个属性相互作用。在这种情况下，切换任何一个属性的状态，其他每个属性的贡献

都会改变。那将是从区间 $[0, 1]$ 中均匀抽取出来的一个新随机数。同时对象的值则将是这20个新随机数的总和（每个属性一个）。这就意味着，属性的每一次翻转，都会导致整个对象产生一个新值，而且它与之前的值不相关。因此，景观将会非常崎岖，每一点都可能隆起、每一点都可能下陷。

在上面这两个结果的基础上，我们可以推导出局部峰值的期望数量。如果我们从任何一个备选方案开始，单翻转算法将会对该备选方案与 N 个备选方案中的每一个进行比较。例如，假设我们从所有位都取零值的备选方案开始，那么单翻转算法将评估 N 个备选方案，每一个备选方案都恰好有一位的取值为1。

初始备选方案：00000000000000000000

属性1切换后的备选方案：10000000000000000000

属性2切换后的备选方案：01000000000000000000

.....

属性20切换后的备选方案：00000000000000000001

要成为一个局部高峰，其值必须高于这 N 个备选值中的每一个。原来的备选方案拥有最高值的概率等于 $1/N$ 。因此，局部峰值的数量近似等于可能的替代方案数量 2^N 除以 N 的商。很容易计算，对于 $N=20$ ，大约有5万个局部峰值。由于具有如此之多的局部最优，只靠单翻转算法很难准确定位全局峰值。

因此，在这里重要的不是局部最优值的数量，而是它们的值有多大。因此还需要将这些最优值的期望均值与全局最优值的期望值进行比较。通过这种比较，我们可以确定单翻转算法的性能如何。而在计算这些值的时候，我们可以利用中心极限定理。在这个例子中，很容易证明

局部最优值的期望值大约等于0.6，而全局最优值的期望值则略大于0.75。³将这些值与 $K=0$ 时的全局最优值比较一下，就可以发现崎岖景观的局部峰值低于富士山景观的峰值，但是崎岖景观的全局峰值则比富士山景观的峰值更高。

由此自然而然地引出了这样一个问题：在这两个极端之间，也就是当我们将交互作用的属性数量 K 从零增加到 $N-1$ 的过程中，发生了什么事情？答案是，我们可以观察两种效果。交互作用的属性数量的增加，在产生了更高的全局峰值的同时，也产生了更多的（更低的）局部峰值。假设我们在搜索时使用了单翻转算法，那么对这个模型的计算表明，对于比较小的 K 值，互动的好处（得到了更高的全局峰值）超过了互动的坏处（出现了更多的局部峰值）。因此，在开始阶段，局部峰值的期望值会随着 K 的增加而提高。同时，局部峰值数量的增加则意味着它们的平均值将减小。因此，如果决定使用单翻转算法，人们一般更喜欢用相对较小的 K 值，比如3或4。但是，我们为什么要把自己局限在这种只切换单个属性的简单启发式上呢？经由变异而完成的进化也许适用这个启发式，但是我们却不必受它所限。我们可以切换两个属性甚至三个属性的状态。更复杂的算法将会减少局部最优值的数量。

崎岖性与舞动的景观

NK模型的一个重要含义是，我们需要适度的相互依赖性，因为这种互动能够产生更高的峰值。多模型思维则要求我们跳出模型的特定假设，并仔细思考是什么原理驱动了这些结果。我们不难发现，背后的逻辑由两部分组成。第一个组成部分源于组合学：两个元素的组合数会随对数的平方而增加，而三个元素的组合数则会随三元组数量的立方而增加。因此，这种相互依赖效应有可能会创造出更多有益的互动。

第二个组成部分则源于我们只需要保持更好的组合这个事实。想象一下，假设我们准备利用4种食材来制作一份食物。有4种食材意味着：

如果选用其中的两种，那么有6种可能的组合。假设我们所用的是以下4种食材：泡菜，香蕉，鸡肉，焦糖。由此产生的6种配对是：香蕉和泡菜、泡菜和鸡肉、焦糖和泡菜、香蕉和鸡肉、焦糖和香蕉，以及焦糖和鸡肉。你可能会觉得只有一种搭配对你有吸引力，那么你就会选择那种搭配。假设我们喜欢的是焦糖和香蕉，就会忽略其余配对。⁴

类似的逻辑也适用于进化系统。能够产生正面相互作用的表型组合，例如坚硬的外壳与强健的短腿，会在种群中保存下来。适者生存法则与产生负面相互作用的组合相反。因此，我们现在看不到有如下这种组合的动物：跑得很慢的脚+味道鲜美的皮肉+外表鲜艳的外表。它们也许曾经存在过，但是早就被捕获并被吃光了。

我们在搜索模型中也会遇到类似的情况。当拥有非常多的可能性时，我们更喜欢变化。同样的逻辑在这里也是适用的：组合（两个元素、三个元素）会产生丰富的可能性。而且我们还希望，这些可能性的价值有很大的变化范围。然后，我们更有可能发现其中一个具有非常高的价值。由于相互作用效应会增加变异，因此总体上说，它们是有利的，但只在一定程度上有利。正如我们在上面已经阐述过的，太多的变化会使景观随机化。在理想情况下，我们会有适度的互动。有的学者认为，如果相互作用的数量和大小可以演化或适应，那么系统应该会自然而然地演变为具有高峰值的崎岖景观。如果真的是那样，那就表明系统倾向于向复杂性而不是均衡或随机性演化。⁵当然，什么时候能够达到这个结果、是不是真能达到这个结果，本身也是一个可以通过模型来探索的有趣问题。

最后要强调的一点是，我们一直将景观视为固定的。但是在生态和社会系统中，物种或企业要面对的景观还取决于他人的行为和属性。任何一个物种的适应，或任何一个企业策略的改变，都会改变和重组它们的竞争对手的适合度景观。

现在，我们可以将空间竞争模型和享受竞争模型重新解释为舞动的

景观上的运动模型。这种运动可能会导致均衡，每个博弈参与者都站在局部或全局山峰上。或者，在舞动的景观上的竞争，也可能导致复杂的行动模式和结果。只要粗略地观察一下生态系统、政治领域和经济社会，就会明白后一种情况更容易出现。

我们之所以会观察到如此多的复杂性，一个很重要的原因可能是，我们这个世界在很大程度上是由自适应的、有目的的行为者组成的，它们有能力操纵舞动的景观。为了理解这种复杂性，我们需要多模型思维。

我们可以对知识授予专利权吗

我们今天的幸福源于数百年来知识积累。知识体现在所有方面：物理定律、内燃机、复式记账法、传染性细菌致病论、X射线和HTML等。知识通常是一种公共物品，永远是非竞争性的，不过知识既可能是、也可能不是排他性的。要想排他，必须有方法验证，当知识已经呈现为特定形式的人工制品时，要验证是比较容易的。例如，要想验证某个人是不是使用了某种算法或技术来解决了某个问题，那往往是不可能的；但是，要想验证某人是不是在软件程序中嵌入了某种算法，就是可以做到的。

当知识的排他性可以保证时，我们就会面临一个选择。我们可以像对待道路和国防那样去对待知识，并通过向民众征税来生产知识。政府可以向那些思考者支付报酬或者直接补贴他们，或者通过支持大学、研究机构来间接地加以支持。政府还允许人们获得专利权。专利制度之所以能够鼓励知识生产，关键就在于它为专利所有人创造了一定期限的使用知识的独占权、并允许他们向其他使用专利的人收费。在美国和欧洲，专利权的期限为自申请提交之日起20年。⁶专利倡导者认为，如果任何人都可以免费使用别人发明的东西，那么私人（个人或机构）就不会有什么动力去开发更好的捕

鼠器、计算机算法或音响系统了。他们强调，专利制度能够克服知识生产中固有的激励问题。

但是，经济学家米歇尔·博尔德林（Michele Boldrin）和戴维·莱文利用多模型思维，提出了一个有力的反对专利制度的理由。² 在他们给出的允许思想（创意）组合的模型中，引入专利权会限制不同思想的组合，从而阻碍创新。如果一家公司获得了触摸屏技术专利，那么就很可能会减少其他企业设计采用这种技术的新产品的动力。如果没有专利保护，就会有更多的产品采用这种技术。也就是说，创新将会增加。

专利制度的支持者则反过来指出，就算专利制度真的会阻碍创新（那将很糟糕），但是如果没有专利保护，那么投资的减少幅度将会大得多。博尔德林和莱文基于我们在本书中讨论过的扩散模型反驳了这种说法。利用新知识设计生产的有用产品会迅速通过消费者传播开来。收音机、电视和谷歌搜索引擎都是如此。这会创造出一种先发优势，创新者仍然会受益，尽管获利程度与专利保护下有所不同。

博尔德林和莱文还对某项发明应该在多大程度上归功于发明者提出了疑问。如果重大突破都是某个孤独的天才在密室中做出的，而且如果没有动力，大多数新思想都不会出现，那么专利制度就是有理由的。但是，崎岖景观模型表明，大多数困难的问题都有很多种可行的解决方案。新发明，特别是那些结合了现有思想和技术的发明，例如汽车、电话和在线拍卖，也许是本来就会发生的“自然事件”，而不是某个天才人物行为的结果。如果各种想法和创意都能够在思考者的群体中自由流动，那么很多人可能早就实现了这些创新。从历史上看，许多重大发明（发现）都有一种引人注目的同时性，例如，微积分是由艾萨克·牛顿和戈特弗里德·莱布尼茨发明的、电话是由亚历山大·格雷厄姆·贝尔和伊莱莎·格雷发明的，以及

进化的自然选择理论是由查尔斯·达尔文和阿尔弗雷德·拉塞尔·华莱士发现的。

总而言之，多模型思维能够呈现专利制度的优点和缺点。这些模型提供的更深入、更周详的结论支持一种更加灵活的专利制度。也许，对于其中一些想法，那些许多人都能够发现的想法，以及可以与许多其他想法重新组合的想法，我们应该采取与今天的专利制度不同的专利制度，例如授予更短的保护期限、更宽松的使用条件。甚至，有些想法根本不应该被授予专利。

结语 像芒格一样智慧地思考——多模型思维的实际应用

一切都是复杂的；如果不是这样，那么生活、诗歌以及所有一切，都只会成为烦恼和负担。

华莱士·史蒂文斯

这是全书的最后一部分内容。在这一部分中，我们用多模型思维分析两个重要的政策问题：阿片类药物滥用和经济不平等。我们将阐明，如果同时运用多个模型，不但可以更好地分析这些问题，而且可以帮助我们理解为什么它们如此难以解决。我们还将会看到，特别是在阿片类药物滥用这个问题上，专家们如何利用多个模型在危机真的发生之前就预测到它。但是在这里，我们并不想说，利用模型就可以避免灾难的发生，那过于夸大其词了。我们这里对阿片类药物滥用问题的分析其实并不深刻，事实上，我们只是试图就如何利用多模型思维来考虑政策制定和政策实施问题给出一个粗糙的模板。我们没有收集过数据，也没有校准模型。相反，只是定性地应用模型得到一些见解。

然而另一方面，我们对收入不平等的分析则包括了更多细节，而且与学术文献紧密地结合在一起。它代表了多模型思维的另一个极端，对各个模型都深入地进行探讨。无论是对阿片类药物滥用问题，还是对收入不平等问题，利用多个模型来进行思考都会使我们变得更有知识、更聪明。但从定义上讲，复杂的系统是很难预测和理解的。我们肯定会犯错误，但是可以从这些错误中吸取教训，变得更加明智。

多模型思维与阿片类药物滥用

在美国，阿片类药物滥用情况有多严重？只要举出一个数字就够了：2016年，医生开出了超过2亿多张阿片类药物处方，这相当于差不多每个人一个处方。在那一年，美国有超过10万人死于与阿片类药物有关的服药过量。而有阿片类药物滥用问题的人则超过了1 000万人，其中有200万多人已经被归类为阿片类药物使用障碍患者。

医生之所以开出了如此之多的阿片类药物处方，主要原因当然是它们有疗效，阿片类药物可以减轻疼痛。数以千万计的美国人都需要止痛，从而对这类药物产生了巨大的需求。但是没有人料到，这类药物会被滥用到这个程度。为了解释阿片类药物的滥用，我们采用了多模型思维方法。关于这个危机产生的原因，有四个模型都给出了一些重要的直觉性结果。

第一个模型是多臂老虎机问题，它解释了为什么阿片类药物会被批准使用。在申请药物上市时，制药公司要进行临床试验，以证明药物有显著的疗效并没有有害的副作用。我们可以将药物临床试验建模为一个多臂老虎机问题，其中一只手臂对应新处方药，另一只手臂则对应安慰剂或现有药物。

阿片类药物批准模型——多臂老虎机问题

为了证明阿片类药物的疗效，制药公司要进行药物与安慰剂的比对试验。在临床试验中，患者随机分为两组，一组服用阿片类药物，另一组服用安慰剂。我们可以把阿片类药物建模为一个“双臂老虎机”问题模型中的一只手臂、而把安慰剂建模为另一只手臂。在试验结束时，每个试验都归类为“成功”或“失败”。临床试验发现，接受阿片类药物治疗的患者的疼痛（在统计学的意义上）显著减轻了。对接受过髋关节置换术、牙科手术和癌症治疗的患者进行的药物试验都表明，阿片类药物的效果显著优于安慰剂。

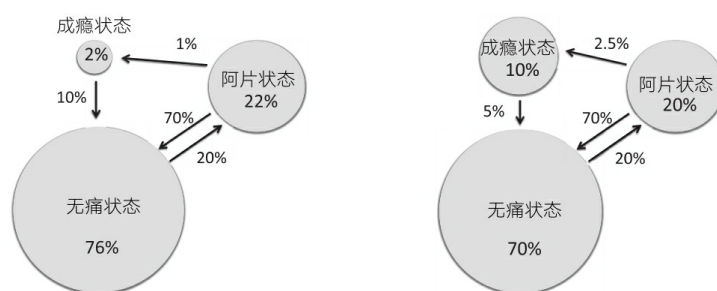
对于任何一种药物，成瘾的可能性都是一个非常值得关注的问题。只有当临床试验的结果可以证明，药物成瘾的患者比例极小（不到1%）时，这种药物才有可能获得批准。然而，这种试验并没有考虑过医生会给患者开出“大处方”的可能性，在某些情况下，医生一次开出的药就足够服用一个月。一个人服用阿片类药物的时间越长，成瘾的可能性就越大。数据表明，服药期较长的患者的成瘾率将会超过2.5%。下面的马尔可夫模型表明，成瘾率从1%提高到2.5%，就可以使阿片类药物成瘾者的均衡人数增加5倍。

这些转移概率仅用数据进行了不严格的校准。因为我们在这里用这个模型只是为了得出一个直观结论：相对较小的成瘾率是如何导致了大量成瘾者的。对这个模型的数值实验表明，只要我们稍稍降低成瘾者戒瘾成功的概率并提高从无痛苦状态转变为阿片状态的概率，那么阿片类药物成瘾者的比例会急剧增加。例如，如果我们在第二个模型中将从成瘾状态转变到无痛苦的转移概率降低为1%，那么成瘾者的比例就会增加到35%。这种模型思维的含义是非常明显的，在现实世界中也开始得到了落实：有的医疗保健服务提供者，现在已经对医生可以开出阿片类药物的数量进行了限制。

成瘾模型——马尔可夫模型

为了计算出成瘾的概率，我们创建了一个三态马尔可夫模型。这三个状态分别表示不受疼痛折磨的人（无痛状态）、使用阿片类药物的人（阿片状态）和成瘾者（成瘾状态）。我们要估计这三种状态之间的转移概率（用下图中的箭头表示。左侧的模型假设使用阿片类药物的人当中，会有1%的人成瘾，10%的成瘾者会恢复为无痛状态，并假设20%的处于无痛状态的人会成为阿片类药物使用者。在均衡状态下，只有2.2%的人是成瘾者。在考虑了“大处方”的情况下，左边的模型假设2.5%使用阿片类药物的人会成瘾，同时只有5%的成瘾者会恢复为无痛状态，并假设20%处于无痛状态的人会

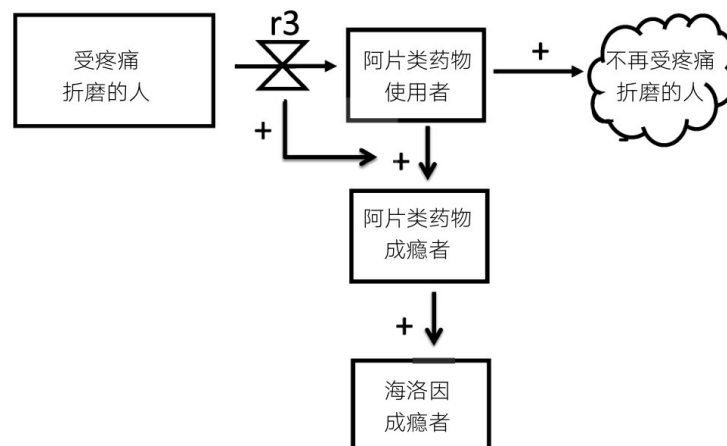
成为阿片类药物使用者。现在，在均衡状态下，会有10%的人是成瘾者。¹



我们的第三个模型是一个系统动力学模型。与马尔可夫模型类似，在这个模型中，我们也假设了有三种类型的人（三个状态）：受疼痛折磨的人、使用阿片类药物的人和不受疼痛折磨的人。然而，这个系统动力学模型并不能直接写出这些状态之间的转移概率，而是想象存在一个“流”：由受疼痛折磨的人，到阿片类药物的使用者，再到不受疼痛折磨的人。更精细的系统动力学模型还可以包括其他一些源（其他药物的提供者），并允许阿片类药物使用者和海洛因使用者之间互动。此外，更精细的模型可能包括其他类型的情况。患有焦虑和抑郁的人更容易成瘾这一事实也可能被纳入这个模型。²

海洛因成瘾之路——系统动力学模型

模型描述了受疼痛折磨的人群产生阿片类药物使用者和海洛因成瘾者的过程。阿片类药物使用者会变为无痛苦状态，也会变为成瘾状态。而成瘾者则可以进一步发展成为海洛因使用者。人们使用海洛因的一个原因是他们不能再服用阿片类药物了。因此，随着阿片类药物的流量的加大，海洛因使用者的数量也在增加。



可以用来分析成瘾问题的最后一个模型是社会网络模型，我们没有在这里给出它的正式形式。这个模型依靠社会网络解释为什么人均阿片类药物使用量的地图会出现类似于农村地区那样的聚类。根据我们对平方根规则的分析，人口规模越小，变差越大。回想一下前面举过的例子：表现最好的学校和表现最差的学校都是小学校。农村地区阿片类药物的使用更大，也可能是因为医生会给农村病人开出服药时间更久的大处方，因为他们距离药店更远。除了这些解释之外，阿片类药物的使用也出现了聚类现象，而不再像是随机发生的。

如果阿片类药物是通过药物贩卖的渠道扩散的，那么就可能会出现这种聚类。与销售二手旧家具不一样（在销售二手家具时，人们会刊登广告），阿片类药物是通过个人关系销售出去的。因此，这样一个社会网络模型将会从家庭和朋友的社交网络开始。模型可能会假设，人们只向亲密的朋友出售阿片类药物。如果真的是这样，那么阿片类药物滥用者将会出现局部聚类现象。³

用多模型法分析经济不平等问题

接下来我们讨论最后一个多模型应用。在这里，我们将深入研究经济不平等的各种原因。之所以要讨论这个问题，有三个原因。首先，不

平等是我们这个时代最重要的政策问题之一。收入和财富与人类社会的繁荣和未来息息相关。高收入人群拥有更好的健康状况、更长的预期寿命、更高的生活满意度和幸福感。处于收入分布底部的人无论是谋杀率、离婚率，还是精神性疾病患病率都更高，他们普遍更加焦虑。⁴不过，我们必须非常小心，不要将相关关系与因果关系混淆起来。这种相关关系的很大一部分可以通过更健康、更快乐的人赚的钱更多这个事实来解释。但是，几乎所有研究都表明收入与生活幸福之间存在联系。没有人喜欢自己穷困潦倒。其次，我们有各种各样的不平等模型，这些模型分别由经济学家、社会学家、政治学家，甚至物理学家和生物学家提出。再次，我们拥有丰富的关于收入和财富的国内数据和跨国数据。我们不仅拥有当前的数据，还拥有可以回溯数百年的时间序列数据。

我们要先总结一下有关收入分配的若干经验规律。首先，在任何时代、任何国家，收入分布都有一个很长的尾部——许多低收入者和一小部分高收入者。在历史上，收入分布曾经被校准为对数正态分布或帕累托分布。最近，颗粒度更细的数据表明，收入分布的尾部长于对数正态分布，但是又不完全符合幂律分布。财富分布也同样是偏斜的。

其次，在大多数发达国家，近几十年来收入和财富不平等状况（无论用什么指标来衡量）一直处于不断恶化的趋势中。目前，美国的收入和财富不平等程度已经接近了镀金时代。由于整个分布内部的变化很难辨别，因此按照惯例，我们描述了归属于分布的上尾部的收入份额的变化情况。图29-1显示了顶层0.1%所占的收入份额随时间流逝而演变的情况。该图表明最顶层千分之一的家庭所占的收入份额在20世纪50年代后稳步下降，直到20世纪80年前后一直稳定在不到4%的水平。但是现在，这个数字已经上升到了10%左右。

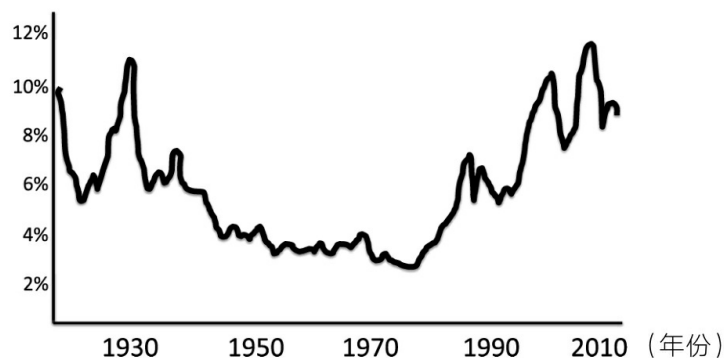


图29-1 最顶层0.1%的人的收入份额，1916—2010年。

资料来源：Piketty, 2011。

再次，在全球范围内，生活在极端贫困中的人的数量也急剧下降了。我们应该看到，这些看似对立的趋势之间其实不存在逻辑矛盾。贫穷国家收入的快速增长，显著地减少了收入的跨国差异，但是并没有抵消国内不平等程度的加剧。群体选择模型也产生了类似的效果。利他主义社区数量的增长，压倒了每个社区内部自私者上升的趋势。

不平等有很多原因，而且这些原因往往是相互交织的。经济力量、社会趋势、政治因素，以及发展历史，都会导致不平等。因此，正如经济学家史蒂文·杜鲁夫（Steven Durlauf）所指出的那样，我们不应该试图只用一个方程来解释所有不平等的水平或趋势，也不应该把所有政策建立在一个模型的基础上。⁵我们必须想得更加细致周到一些。使财富和收入集中在顶层1%或最顶层0.1%的过程，可能与将底层20%人困在贫穷陷阱中的因素无关。要深入了解收入分化的原因，需要采取多模型方法。

我们首先描述解释收入分配变化的模型。收入有以下几个来源：工资和薪金、营业收入、资本收入和资本收益。收入的这些组成部分股票的相对大小因收入水平而异。低收入人群的资本收益或资本收入很少。收入最高的那些人则从每个来源都可以获得很可观的收入。

我们的第一个模型扩展了柯布-道格拉斯（Cobb-Douglass）生产函

数模型。在这个模型中，劳动包括两种类型：受过教育的人提供的劳动和未受过教育的人提供的劳动。支付给某种类型的劳动的工资取决于该类型的相对供给和技术。⁶这个模型可以解释近期基于供求关系不平等程度的上升。在20世纪50年代，制造业的增长增加了对未受过教育的工人的需求。与此同时，大学入学人数的大幅上升（这部分是因为美国《退伍军人安置法案》的实施），增加了受过教育的工人的供给。到了20世纪80年代，由于人们上大学的动力下降，减缓了大学毕业生数量的增长，再加上受教育程度较低的移民流入，增加了低技能工人的供给。与此同时，技术变革——自动化制造的兴起和向更加数字化经济的转型，增加了受过教育的工人的相对价值。他们的工资上涨反映了价值的这种变化。

按教育程度划分的平均收入时间序列数据与这个模型的拟合相当好。出于这个原因，许多经济学家依靠这个模型来提出政策建议。根据这个模型，他们主张增加所有人受教育的机会，因为这能抑制受过教育的工人工资的上涨趋势并减少不平等。这个模型很好地解释了总体趋势，但是它无法解释每个收入阶层内部的变差的扩大。

技术和人力资本模型——增长模型

产出取决于实物资本（K）、受过教育的人的劳动（S）和未受过教育的人的劳动（U），具体生产函数如下：

$$\text{产出} = AK^{\alpha}S^{\beta}U^{\gamma}$$

参数A、 α 、 β 和 γ 刻画了技术和三种投入要素相对价值。高技能工人和低技能工人的相对市场工资是：⁷

$$\text{Wage}_S = \left[\frac{U}{S} \right] \times \left[\frac{\beta}{\gamma} \right] \times \text{Wage}_U$$

（ Wage_S 是高技能工人的工资， Wage_U 是低技能工人的工资）

不平等的原因：有利于受过教育的工人的技术变革会使 β 增大、使 γ 减少。这一点，再加上低技能工人供给的增加，会增加不平等。

不过，正反馈模型却可以。这个模型着重关注分布的尾部上的人群，特别是企业家群体。2011年，企业家在美国400名最富有的人中所占的比例达到了70%。⁸这个模型假设技术，特别是互联网和智能手机使人们之间的联系更加紧密，并使人们在更大程度上受他人的选择的影响。⁹一个想购买无线立体声扬声器的人，可以在线阅读评论，并从多个可选项中选出“最好的”。在过去，这个人也许只能在当地的立体声商店买，并只能选择唯一可选的某个型号。一个扭伤了膝盖的人，现在可以在网上搜索并了解他最喜欢的运动医生的身份。这种连通性会产生正反馈并导致更大的不平等。为了给这种受社会网络影响的经济选择建模，我们修改一下优先连接模型，以便将正反馈与人才联系起来，从而为受社会影响的经济选择建模。

虽然正反馈模型与时间序列数据的拟合程度不如前面那个强调技术导致增长的模型高，但是我们可以通过运行数值实验来了解反馈是如何导致不平等的。回想一下在第6章中描述过的音乐下载实验，也就是让随机分成两个组的大学生在两个不同的情境中下载音乐。在第一个情境中，被试们无法看到其他人在下载什么音乐。这种情境刻画了互联网出现前的世界。在第二个情境中，被试们可以看到每首歌的下载数字。在不能观察到“社交”信息的情境下，没有一首歌的下载次数超过了200次，只有一首歌的下载次数少于30次。然而，当人们可以看到下载次数时，有一首歌的下载次数超过了300次，同时一半以上的歌曲的下载次数少于30次。信息和社会影响放大了马太效应。富人变得更加富裕了，而穷人则变得相对更穷了。

对才华的正反馈——优先连接模型

存在 N 个生产者。开始时，每个生产者的销量均为零。第一个消费者随机选中了一个零销量的生产者，购买了产品，使该生产商的销量为正。随后，每个后续消费者都以概率 p 从销量为零的生产者那里购买、以概率 $(1-p)$ 从具有销量为正的生产者那里购买。当从具有销量为正的生产者那里购买时，消费者选择生产者的概率与该生产者的当前销量成正比。

不平等的原因：更多的联系增加了社会影响，创造了正反馈。

显然，我们可以将同样的逻辑应用于对经济问题的分析。¹⁰社交网络的正反馈效应导致不平等的可能性部分取决于人们所购买东西的性质。没有重量的商品，比如说可下载的电影、音乐和网络应用程序，以及某些技术，都很容易传播。点击一下图标是不能复制拖拉机、汽车和洗衣机的。因此，新的智能手机应用程序的销量可以几乎不需要付出任何资本支出就能够扩大，但是汽车却不能，即便是最畅销的汽车也不能。一个例子是，2015年5月，沃尔沃宣布，将在南卡罗来纳州生产S60轿车。这家新公司将于2015年9月破土动工，而第一批汽车将在2018年末才能下线。

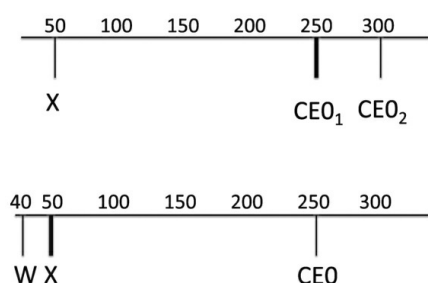
下一个模型源于空间投票模型。我们用它来解释公司高管（如CEO）薪酬的上升，它不是由社会网络因素决定的。2012年，财富500强企业CEO的平均收入超过了1 000万美元，大约相当于当年工人平均工资的300倍。相比之下，在1966年，CEO的工资仅为工人平均工资的25倍左右。其他国家CEO的收入则要少得多。在日本，CEO的收入大约是普通员工的10倍。在加拿大和整个欧洲，CEO的工资大约是普通工人的20倍。

在大多数公司内，CEO的薪酬是由一个由董事会成员组成的薪酬委员会决定的，通常包括了工资、奖金和股票期权。这就是说，决定CEO薪酬的人往往是其他CEO。他们有很强的动机提高其他CEO的薪酬，进

而提高自己的薪酬。我们可以使用一个空间模型来表示薪酬委员会的偏好。根据空间投票模型，工资将被设定为中间选民的偏好。CEO的薪酬跨国差异可以通过董事会和薪酬委员会的组成来解释。在德国，董事会包括工人成员，他们更倾向于减少CEO的工资。

CEO的薪酬——空间投票模型

CEO薪酬由薪酬委员会投票决定。在美国，薪酬委员会通常由现任和前任CEO（他们当然更喜欢高薪）以及薪酬专家（X）组成。而在其他国家，薪酬委员会的组成人员中还有工人（W），因而导致中间选民更偏好比较低的薪酬水平。



不平等的原因：CEO通过相互“俘获”来确定自己的薪酬。任何一位CEO的薪酬的增加，都会使所有CEO更加偏好更高的薪酬。

这个模型根据什么样的薪酬才是适当的来解释CEO薪酬的上升。在一点上，我们也可以回过头去考虑分析价值的多模型方法。薪酬委员会成员的理想价值可能是基于数据的，也可能在社会影响下形成的，或者也可能是基于对CEO未来价值的预测。

下一个收入不平等模型来自著名经济学家托马斯·皮凯蒂（Thomas Piketty）的畅销书《21世纪资本论》（*Capital in the Twenty-First Century*）。与其说这是一个正式的模型，还不如说它是一个观察结果：资本回报率总是会超过经济增长率。只要这个关系成立，那么高收入者从资本回报中获得的收入部分就会随着时间的推移而不断增加。如

果在增长模型的基础上构建一个更精致的模型，应该不难证明资本回报率始终会超过整个经济体的增长率。从长远来看，经济增长率可能不到2%或3%，但是资本回报率则可能会高出一倍以上。

由此可以推出，在一个由赚取工资的工人和从租金中获得收入的资本家组成的经济体中，资本家的收入份额将会增加。更正式地说，资本增加的速度将取决于三个比率：消费率、税率和资本回报率。消费取决于资本存量水平。一个没有什么资本的人需要把自己收入的很大一部分消费掉，而拥有大量资本的人的消费只占收入的很小一部分。如下面的专栏所示，如果我们将消费率表示为一个常数除以资本水平，那么消费量将不会依赖于资本水平。较富裕的人将以较低的速度消费，而这会使他们的净资本更有可能增加。

资本租金模型（皮凯蒂）——72法则

经济由工人和资本家组成。工人的工资增长率为 g ，即经济增长率。资本家在时间 t 有财富 W_t 。资本的回报率为 r （税后净额），且资本家的消费为一个不变的常数 A 。资本家的收入将比工人的增长更快，当且仅当：

$$r - \frac{A}{W_t} > g$$

不平等的原因：在市场经济中，资本回报率必定超过经济总增长率（即 $r > g$ ）。拥有大量财富的资本家只将自己资本收入的一小部分用于消费，因此他们在总收入中的份额将随着时间的推移而不断提高。

为了说明资本回报率与经济增长率之间的差异是怎样导致不平等的，我们可以应用72法则。如果在最初的时候，工人的收入与资本家的收入相等，工资的年增长率为2%，而资本的年增长率为6%，那么在36

年后，工资将增加一倍，但是源于资本的收入将增加8倍。而在72年内，资本家的收入将会达到工人收入的60倍。

皮凯蒂就是运用这样一个模型来解释收入和财富不平等的长期趋势的。用法国和英国过去三个世纪以来的数据对这个模型进行校准，结果非常好。这个模型还揭示了过去一个世纪在美国和欧洲各国的不平等的演变模式：两次世界大战摧毁了欧洲的资本存量，使那里的收入和资本分配较为平衡。这个模型能够很好地适合数据的一个原因是它略去了两个相互抵消的效应。由于将企业家排除在外，这个模型低估了不平等。而通过假设资本家的后代会明智地投资（尽管并非所有资本家的后代都能做到这一点），这个模型又夸大了不平等。因此，它在创造了一个新的富人阶层的同时，抹去了一个原来的富人阶层。这个模型，由于同时出现了这两个疏漏，因而比只有一个疏漏的模型更加准确。

这个模型的含义是，只要资本在增加，资本家就会从经济蛋糕中获得越来越大的份额。只要继续应用72法则，任何人都能看到资本家的收入将使工人的收入相形见绌。对于这种原因导致的收入不平等，似乎有一个非常简单的解决方案：向富人征收财富税。作为征税的替代方案，有的人可能会期待战争和革命，也就是以暴力的方式重新分配财富；或者，等待能够产生很多新资本家的技术突破的出现。

接下来要考虑的两个模型都优先考虑社会力量，而且两者都有很强的经验证据支持。第一个模型解释了选择性婚配，也就是所谓的门当户对的婚姻选择所导致的不平等恶化。家庭的收入取决于夫妻双方的收入。如果一个低收入者与一个高收入者结婚，那么这种婚姻将有助于平衡收入分配。如果一个高收入者与另一个高收入者结婚，那么收入差距将会增大。虽然大多数人结婚时，都无法确知未来的人生伴侣的终身收入；但是，人们确实可以了解潜在结婚对象的受教育程度和身体健康状况，并且能够收到他们是否拥有雄心壮志的信号。有证据表明，男性（女性）受教育程度越高、收入越高（技术和人力资本模型），他们越

有可能选择受教育程度较高的人生伴侣。

不平等的加剧是由以下因素造成的。第一，获得大学学位的女性越来越多。第二，相对收入会随着接受教育的程度的增加而提高。第三，受过良好教育的男性（女性）更喜欢受过良好教育的人生伴侣。因此，由两个受教育程度很高的人组成的家庭，更有可能拥有两个高收入者，从而导致家庭之间的收入不平等状况恶化。这里的逻辑似乎很严密。唯一的问题是这种效应的影响空间有多大。¹¹

社会学家通常会将人们按教育水平分为五类：辍学、高中毕业、上过大学、有大学学位，以及研究生。然后，他们计算出每个教育水平的平均收入，并拟合关于每一对教育水平之间的婚姻数量的数据，从而粗略估计出选择性婚配的影响。

选择性婚配模型——分类模型与分类

每个人都有自己的教育水平： $\{1, 2, 3, 4, 5\}$ 。（其中1=辍学，2=高中毕业，3=上过大学，4=大学学位，5=研究生）

令 $P(m, j)$ 和 $P(w, j)$ 分别表示男人和女人具有教育水平 j 的概率。收入 (g, ℓ) 表示性别为 g 、收入水平为 ℓ 的人的（估计）收入。一对夫妇组成的家庭的收入，包括了一个受教育程度为 ℓ_M 的男子和一名受教育程度为 ℓ_w 的妇女的收入。其家庭收入估计如下：¹²

$$\text{收入}(M, \ell_M) + \text{收入}(W, \ell_w)$$

不平等的原因：受过良好教育的女性人数的增加、高教育水平的工人工资的增加，以及选择性婚配（人们喜欢与收入水平相同的异性结婚的倾向）导致了家庭之间收入不平等的增加。

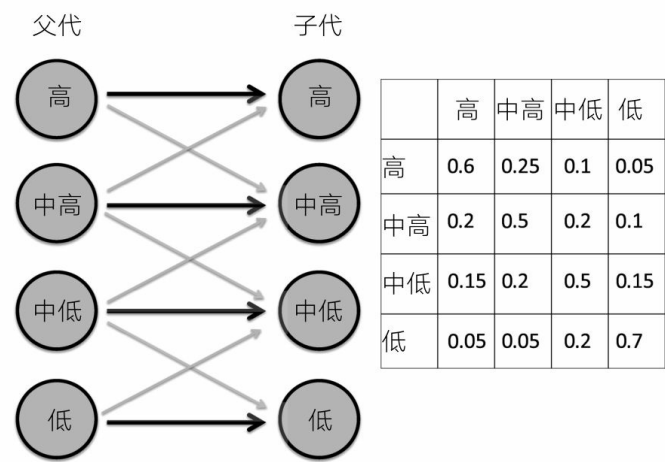
如果婚姻是随机的而不是非得门当户对不可的，收入不平等的程度

就会轻得多。一项研究表明，如果婚姻是随机的，那么以基尼系数衡量的收入不平等程度将会减少25%。[13](#)

下一个模型则使用马尔可夫模型分析不同收入类别之间的变动。这个模型按收入水平将人们（或家庭）分成四类：高、中高、中低、低。每个类别包括了分布的1/4。选定一个时间段，可能是1年、10年，也可能是一代人的时间，然后估计收入类别之间的转移概率，以刻画收入流动性。

代际收入（财富）动态变化——马尔可夫模型

将所有人口划分为4个人数相等的收入（或财富）类别。我们可以估计一个类别中的个体（或家庭）在一代的时间内流动到另一个类别中的转移概率（如下图所示）。更平等的转移概率对应更大的社会流动性。



收入水平（收入等级）之间的转移概率

不平等的原因：社交技能、隐性知识、对风险和教育的态度以及遗产，减少了收入阶层之间的流动性。

如果代与代之间不存在黏性，那么高收入父母的孩子的收入属于4

个收入阶层中的任何一个的概率都相同，即所有的转移概率都等于1/4。在完全不存在流动性的最极端情况下，转移概率将仅包括沿对角线的那些¹。经验结果估计，表明现实介于这两个极端之间。

我们可以拿100个随机挑选出来的高收入或低收入水平家庭来进行仿真实验，计算出每一个后代收入的概率分布。如果使用上面的专栏中给出的概率，高收入者的子女有60%的机会成为高收入者，只有5%的机会成为低收入者，那么高收入者的孙子女成为高收入的可能性不足43%，成为低收入的可能性却会超过10%。¹⁴

这个收入动态模型也可以作为评估收入流动原因的基准模型。我们可能会使用一个线性模型来估计作为父母的财富、收入和能力水平的函数的子女收入（假设我们拥有相关的数据）。皮凯蒂的模型意味着父母的财富与子女的收入之间是正相关的。基于能力的模型则意味着父母的能力与子女的收入是正相关的，因为父母的能力和后代的能力之间存在某种相关性。

需要注意的是，要确定父母收入系数需要每个子女和每个父母的收入数据。只是过去几十年以来，我们才开始拥有比较完整的个人收入数据。幸运的是，在《太阳照常升起》（*The Sun Also Rises*）这本著名小说中，经济史学家格雷戈里·克拉克（Gregory Clark）发现了一种解决数据匮乏问题的新方法，这种方法利用了人的姓氏。克拉克计算了1888年所有名叫撒切尔的人的平均收入，并将之与1917年所有名叫撒切尔的人的平均收入进行了比较。30年的时间基本相当于一个人一辈子的工作时间。结果，克拉克发现姓氏的平均收入存在显著的相关性，这是表明收入缺乏流动性的一个证据。

有了这种模型，我们还可以识别代际转移中的种族差异。非洲裔美国人虽然也有进入收入分布顶端的，但是表现出了更低的财富持久性；而另一方面，他们在低端则表现出更大的持久性。这也就是说，富裕的非洲裔美国人不太可能拥有富裕的子女，而贫穷的非洲裔美国人却更有

可能拥有贫穷的子女。¹⁵

我们要讨论的最后一个关于收入不平等的模型基于邻域效应，它是以杜鲁夫的持续不平等模型（persistent inequality model）为基础的，利用了人们按收入类别分离居住的经验规律性。这也就是说，高收入的人倾向于与其他高收入的人生活在同一个社区中，而低收入的人则与其他低收入的人住在一起。按收入类别分离居住会产生经济上、社会上和心理上的外部性，从而导致流动性下降。在这个模型中，个人收入取决于能力、教育支出和溢出效应。

教育属性还包含教育方面的公共支出。从经验上看，公共教育支出又与平均收入水平相关：高收入地区在公共教育方面的支出高于低收入地区，从而又可以为高收入社区的儿童带来更好的教育和更高的（未来）收入。

溢出可以解释为获取适当工具所需知识的社会传播。在这里，我们可以将杜鲁夫的模型与居住在高收入社区的人们如何获得关于何为适当工具的知识过程联系起来。还可以将模型与社会网络模型和“弱关系有大力量”现象联系起来：生活在高收入社区的人们能够间接地与更多掌握了很有经济价值的信息的人联系起来。

我们还可以将溢出效应解释为一种社会传播行为，例如学习或工作所花费的小时数。如果收入包含了随机成分，那么低收入社区的人就会观察到，花费在自我完善、自我提高上的时间不会带来多少回报（这种观察结论是正确的）。与此相关，溢出还可能包括心理属性，比如对生活积极的或消极的态度、对社会的安全感，以及对自己的信念。

持续不平等（杜鲁夫）模型——谢林隔离模型+局部多数模型

所有个人分别属于不同收入类别，并按不同收入类别分离居住。个人将自己的部分收入用于教育，从而产生正面的溢出效应，这种溢出效应随社区收入水平的上升而增强。生活在社区C中的孩

子的未来收入取决于自己天生的能力、教育支出和溢出效应。教育支出和溢出效应的贡献取决于这个社区的收入水平， I_C 。

个人收入 $C=F$ （能力、教育支出（ I_C ）、溢出效应（ I_C ））

不平等的原因：在低收入社区长大的儿童获得的教育机会较少、受益于经济溢出效应的可能性也更低。

在完整的模型中，杜鲁夫求解出了教育支出的均衡水平，并推导出了持续不平等产生的条件。这种持续的不平等源于他所称的“贫困陷阱”。生活在低收入社区的个人缺乏提高收入所必需的教育资源，当地也不存在可以令他受益的溢出水平，因此无论他们的能力水平如何，都很难走出贫困陷阱。杜鲁夫的模型有助于解释收入水平上的巨大种族差距，例如，非洲裔美国人生活在贫困社区的比例过高。杜鲁夫的模型解释了为什么他们更可能陷入低收入的人生轨迹无法自拔。

上面给出的这些模型突出了收入不平等的各种不同原因。对于收入分布的顶层，经验证据最支持的是那些以技术变革为基础的模型。¹⁶ 20多年来，美国国税局追踪了收入最高的400个美国人的收入情况。最顶层这些人的收入主要来自新技术、大众零售和金融行业。这三个行业的共同特点是，它们都可以快速扩展。这种特别高的增长率可能源于像搜索引擎和社交网站这样的赢者通吃市场。但是，以技术变革为基础的模型不能解释收入分布底层的情况。关于收入流动性，它们也无法告诉我们什么信息。此外，它们也不能解释为什么美国CEO的薪酬远远超过了其他国家。

为了解释这些现象，我们需要引入其他模型，比如收入流动性模型、杜鲁夫的持续收入不平等模型和空间投票模型。只有在考虑了所有这些模型之后，我们对收入（财富）不平等才会有更加深刻且多方面的理解。我们看到，在不平等的产生和维持中，有许多不同的过程都起到了作用。我们还可以观察到，这些不同的过程之间存在重叠和交叉。当

我们对不平等的复杂性和自我强化的因果关系有了更深刻的理解之后，我们就会对任何声称能够“快速”解决不平等问题的简单方案持怀疑态度。我们认为，减少不平等将需要在很多个方面共同努力。

让智慧入世

在本章中，我们讨论了如何将多个模型结合成一个整体来使用。利用这种方法，我们可以解释阿片类药物滥用和收入不平等的多种原因，并揭示任何一个解释框架的局限性。如果我们是制定政策的专家，就可以用这些模型中的某一个或某一些模型去拟合数据、衡量政策效果，还可以组织自然实验来指导政策选择。

我们还可以针对任何一种社会挑战来进行类似的分析，例如，扭转肥胖趋势，改善教学成绩，缓解气候变暖，管理水资源，甚至改善国际关系。

在每种情况下，即使添加一个新模型也可能产生巨大的后果。以预测金融崩溃为例。美国联邦储备委员会依靠传统的经济模型，使用通货膨胀、失业和存货等国民核算数据。但是这些数据存在滞后性，它们每周、每季度或每年发布一次。这些数据也来自调查，即整个经济的样本。

复杂性学者多恩·法默（J. Dooyne Farmer）主张基于从网络上获取的实时数据创建第二类模型。这些新模型将依赖于更细粒度的实时数据，因此与传统的模型不同。法默认为，这样的模型可能比现有的模型要好得多，他可能是对的。然而，在预测和预防金融灾难方面，这些新模型并不需要更准确。考虑到新模型将使用不同的数据并依赖于不同的假设，它们将做出不同的预测。从多样性预测定理中我们知道，只要新模型的精度不差很多，当与现有模型相结合时，这些新模型将提高预测的准确性。用法默的话来说，政策制定者将会更有集体意识。

在做商业决策时，高管可能会从事类似的工作。高管可以应用多个模型来决定产品属性、产品发布时间，设计薪酬计划，构建供应链并预测销售。因为这些操作都发生在一个复杂的系统中，所以任何一个模型都是错误的。多个模型将会带来更好的行动。

总而言之，当面临选择、预测或设计方面的挑战时，我们应该采取多模型方法。基于多模型思维的“谋定而后动”肯定要比仅仅基于冲动和直觉就行动更好。当然，这也就意味着，我们无法保证成功。即便有很多模型，我们也可能无法确定最相关的逻辑链。与要解决的问题有关的领域可能非常复杂，甚至利用许多个模型也可能仅能解释变差的一小部分。

在应用模型来辅助设计时，我们可能会发现自己无法构建出有用的模型抽象。在这种情况下，模型的简单性恰恰可能会成为它们失败的原因。面对复杂性，我们可能会发现模型无助于我们交流思想、做出准确预测和选择最优行动。我们的探索也可能是几乎没有价值的。这本书中讨论的模型的七大用途并不一定能提供顺利登顶的云梯。但是，即便是在这些情况下，我们也能受益。我们可以揭示出相互依赖性，能够理解为什么复杂的过程往往很难理解。

即便是在模型的帮助下，我们的推理能力也会受到限制，所以我们必须保持谦卑，也必须保持好奇心。我们必须继续构建新模型并改进现有模型。如果某个模型遗漏了世界的某些关键特征，例如社会影响、正反馈或认知偏差，那么我们就应该构建能够包含这些假设的其他模型。当采取这些方法后，就可以分辨出哪些属性在什么时候是重要的、重要程度有多高。所有模型都是错的，这当然是事实。但是，我们不会因这个事实而泄气；恰恰相反，它会成为我们通过多个模型来追求智慧的动力。

同样重要的是，我们还应该在这种努力中追求乐趣。虽然本书一直在强调一些务实的目标，比如成为更好的思考者、在工作中取得更好的

成绩、做一个有知识有智慧的世界公民。但是，它同时也隐含着另一个重要的目标，那就是，揭示建模的乐趣。建模的实践可以成为一个非常美丽的邂逅。我们做出假设、制订规则，然后根据规则、运用逻辑。正是通过这种合乎逻辑的努力，我们才能提高自己、使自己变得聪明。我们要将这种智慧带入世界，并积极地利用它去改变世界。

译者后记

当今时代，大数据至热。但是许多人其实只是被大数据裹胁了。有了数据还远远不够，我们还要懂得如何让数据告诉我们信息，懂得如何从信息中总结出知识。说到底，我们需要形成自己直面世界的思维体系。斯科特·佩奇这本书为我们指出了实现这个目标的方向和路径：多模型思维。

佩奇是个跨学科天才，他在书中给出的模型，涉及各个学科领域。有些模型其实相当复杂，但是佩奇总能说得既简明又透彻。在《多样性红利》一书中，佩奇强调了群体智慧可以改善我们的决策；在这一本书中，他进一步落实了如何实现多样性红利，那就是，利用多个模型，让模型组成一个群体来帮助我们。这本书是一个完整的工具包。我相信，只要你愿意尝试成为模型思考者，就肯定能从本书中收获很多东西。

当然，也要防止另一个极端。“只要不断地拷打数据，数据总会说话”，类似的说法已经屡见不鲜。我们理当保持警惕。不要忘记佩奇写这本书的思想前提和知识基础。他是复杂性科学的大师级人物，已经到了驭繁如简的境界。模型就是帮助我们像大师那样去思考的工具。但是，我们切不可把我们要面对的世界简单化。不要以为在土里插个温度计，就可以测量整个地球的温度了。恰恰相反，模型思维的前提就是承认复杂性。作为本书核心的多模型范式，实际上就可以说是复杂性的另一种表述。

译书从来不易。我首先要感谢的是我的太太傅瑞蓉。本书与我以前的每一本译著一样，她的贡献至少要占到一半。感谢儿子贾岚晴带给我的动力和快乐。这本书献给他们。

感谢钟睽睽，有情怀、重情义、善创新、爱读书的他，既是好老板，也是好老师，让我在工作之余能够腾出时间来完成此书。

我还要感谢汪丁丁教授、叶航教授和罗卫东教授的教诲。我对复杂性科学的兴趣，源于他们。同时还要感谢何永勤、虞伟华、余仲望、鲍玮玮、傅晓燕、傅锐飞、陈叶烽、童乙伦、罗俊、邓昊力、陈姝、黄达强、李燕、李欢、丁玫、何志星等好友的帮助。

这本书，是我翻译的第二本佩奇的著作。佩奇的书，简洁而丰富，纯粹而复杂，因而总是有一种神奇的吸引力，我翻译起来也觉得特别得心应手。这是非常难得的体验。感谢湛庐文化和简学给我这个机会。

书中错漏之处在所难免，敬请专家和读者批评指正。

贾拥民

于杭州嵩谷阁

注释

01 做一个多模型思考者

1. 例如：凯西·欧奈尔（Cathy O'Neil）在2016年阐述了基于数据的简单模型如何忽略人口的某些部分以及本书第4章提到的自适应。
2. 请参见帕阿斯克和希勒分析森林工业的论文（Paarsch & Shearer, 1999）。在他们的模型中，木材加工厂主和林场主要向工人支付按树计件工资。关于植树造林的原始数据显示，计件工资越高，工人种植的树木越少；向某个种植树木的人支付的报酬越多，这个人种植的树木就越少。直接得出的推论与标准经济学原理背道而驰——经济学理论告诉我们，如果每种植一棵树都可以得到更多的报酬，那么种植树木的工人就应该更加努力地工作才是。

不过，如果我们构建一个包含了小时工资率的市场模型，假设每小时的体力劳动报酬为20美元，那么数据和我们的直觉就都可以证明是正确的。从这个假设可以得出这样一个等式：

$$20\text{美元} = \text{每小时种植的树木数量} \times \text{种一棵树的美元报酬}$$

这个模型意味着，如果树木很难种植，那么工资将会很高。如果一个人可以在一小时内种植10棵树，那么种每棵树的工资就等于2美元。如果树很容易种植，那么工资就会很低。如果一个人可以在一小时内种植20棵树，那么种每棵树的工资将为1美元。因此，这个模型预测，计件工资率与种植的树木数量之间呈负相关关系。这样一来，数据中看似矛盾的东西通过这个模型的镜头来看，在逻辑上是解释得通的。这个模型通过种植条件的变化来解释相关性。事实上，这个模型预测的不仅仅是这种相关性，它还给出了一种特定的

数学关系：计件工资率乘以种植的树木数量应该等于一个常数。我们可以对这个结果进行检验。

3. 关于模型胜过了人的证明，请参见：Dawes, 1979; Tetlock, 2005; Silver, 2012; Cohen, 2013。关于认知偏差，请参见：Kahneman, 2011。
4. Slaughter, 2017 & Ramo, 2016.
5. 研究表明，最具影响力的科学发现和专利是不成比例地从多个学科中汲取灵感的。有一项研究在分析3 500万篇论文后指出，从长远来看，跨学科论文的影响更大（Van Noorden, 2015）。思想的组合不一定会表现为模型的组合，但在许多情况下确实是如此；请参见：Jones、Uzzi&Wuchty, 2008; Wuchty、Jones&Uzzi, 2007。如果我们将专利解释为创新的证据，那么有两个独立的研究思路都可以将思想的多样性与成功联系起来。Shi、Adamic、Tseng、Clarkson（2009）的研究表明，跨类别的专利被引用得更多。Youn、Strumsky、Bettencourt、Lobo（2015）则证明，大多数专利都包括了多个子类别。跨学科研究也稳步发展到了这样一个临界点：平均而言，社会科学家引用其他学科的论文，多于他们所属学科的论文。
6. Box & Draper, 1987.
7. Page, 2010a.
8. 我这样说并不意味着可以将知识与模型等同起来。我的意思是，模型可以表示知识并提供一个清晰的形式来交流对有关知识的理解。“知识”一词的含义非常广泛，它还包括了嵌入在身体中的、隐性的技能，例如怎么打网球、说法语或讨价还价。我在本书中使用的是更狭义的知识定义。关于更宽泛的知识概念，请参见：Adler, 1970。

9. 只要注意到，从高空往下掉落的跳伞者在终点时速度可以达到320千米 / 小时，就可以得出这个近似结果。终端速度与质量的倒数成比例。假设跳伞者的质量比毛绒玩具猎豹大400倍。400的平方根等于20，因此，毛绒玩具猎豹的终端速度大体上等于320千米 / 小时除以20，即大约16千米 / 小时。
10. 他没有说错。弗雷斯诺市比冰岛还要大30%。埃里克·鲍尔和约瑟夫·李普曼（Joseph LiPuma）（2012）讨论过怎样把学术研究得出的教训应用到商业世界中。
11. Lo, 2012.
12. 这个故事的不同版本在美国哲学家与心理学家威廉·詹姆斯（William James）、斯蒂芬·霍金和法学家安东宁·斯卡利亚（Antonin Scalia）的著作中都可以找到。

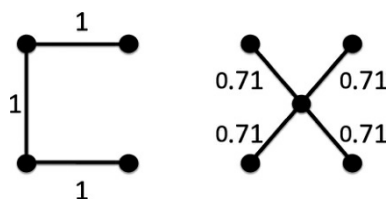
02 模型的7大用途

1. 关于对构建模型原因的更细致的分类剖析，请参见：Epstein, 2008。
查尔斯·拉夫和詹姆斯·马奇（1975）描述了模型的3个用途：解释经验现象、预测其他现象和新的现象，以及建造和设计系统。他们也都提倡利用模型进行探索。
2. 请参见：Harte, 1988。这种方法也借鉴了一篇阐述的关于模型在社会科学中的各种用途的观点的论文（Johnson, 2014）。这两种方法通常也称为伽利略和极简主义的理想化。请参见：Weisberg, 2007。关于类比的作用更进一步的讨论，请参见：Pollack, 2014；Hofstadter&Sander, 2013。侯世达在《表象与本质》一书中将类比为思考之源和思维之火。另外也请参见：Schelling, 1978、1987；他也详细地说明了模型的各种类别。丹尼尔·利特尔（Daniel Little）的博客“理解社会”（understanding society）也提供了相当不

错的社会科学本体论的入门知识。

3. 请参见：Arrow, 1963。如果我们对可能的个人排序施加限制，那么进行集体排序就是有可能的。比如，如果令每个人都保持相同的排序，那么集体的排序也会如此。但是一般而言，我们无法将个人排序映射为一致的集体排序。
4. 与我同一时代的杰出人士一定已经注意到了，我从《嚎叫》（*Howl*）中借用了“这是真的发生过的事情”的手法。请参见：Bickel、Hammel&O'Connell, 1975。

下图是一个添加一个节点会减少总边长的诸多可能例子之一。左侧的网络有4个点（分别作为正方形的一个角）。右侧的网络则多了位于中心的第5个节点。如果将正方形的边的长度假设为1，那么左侧图中各边的总长度为3，右侧图中各边的总长度则等于 4×0.71 ，小于3。



当向录取率更高的院系提出入学申请的男性多于女性时，辛普森悖论就会出现。例如一所拥有医学院和兽医学院的大学。假设有900名男生申请入读医学院，其中480名（53%）被录取；同时有300名女生申请医学院，180名（60%）被录取。并假设有100名男生申请兽医学院，20名（20%）被录取；同时有300名女性申请兽医学院，90人（30%）被录取。在每一个学院，女生的录取率都比较高，但是就整个学校总体而言，男生的录取率达到了50%（1000人申请，500人被录取）；女生的录取却只有45%（600人提出申请，270人被录取）。

作为“帕隆多悖论”的一个例子，假设有这样两个赌局：在第一个赌

局中，一个人总是会输1美元；在第二个赌局中，如果轮数不能被3整除，那么输2美元，如果能够被3整除，就赢3美元。每一次投注都会形成一个期望损失，但是如果你只在赢的轮次上参加第二个赌局，而在其他轮次上则参加第一个赌局，那么你每三轮就肯定能够赢得1美元。

5. Kooti、Hodas&Lerman, 2014.

6. 假设每个人都可以赚得相同的收入，并以固定税率 t 缴纳税收。令 c 表示百分比减量、 r 表示收入的增量。当前的政府收入等于 $I \times t$ 。减税政策实施后，政府的收入等于 $I(1+r) \times t(1-c)$ 。只有当 $I \times t < I(1+r)t(1-c)$ 时，政府收入才会增加。重新排列各项，我们有 $r > c(1+r)$ 。

7. 基于市场的机制如何能为多维有效载荷问题提供一个更好的解，请参见：Ledyard、Porter&Wessen, 2000。

8. 物理学家尤金·维格纳在他的论文（Eugene Wigner, 1960）中使用过的一个形容词“好得不合情理”。他说，物理科学中使用数学模型的有效性“好得不合情理”。

9. Ziliak&McCloskey, 2008.他们讨论了社会科学模型解释变差的能力。

10. Porter&Smith, 2007.

11. Squicciarini&Voigtländer, 2015.Mokyr, 2002.

12. 请参见美国财政部网站。

13. 例如，在20世纪90年代中期，在美国俄亥俄州哥伦布市新开业的餐馆中，大约有60%后来都关门了。没有任何一家得到过政府的救助，当然政府也不应该救助它们。健康的市场经济本身必定包含了破产。请参见：Parsa et al., 2005。

14. 摘自国际货币基金组织的《2009年全球金融稳定报告》。连接的强

度基于投资组合价值的相关性。但是这种相关性又是基于极端事件的，也就是这些机构表现得特别好或特别糟糕的日子。这个指标可以刻画一家金融机构的破产扩散到另一家的可能性。事实上，业绩表现的相关性可能源于投资组合的相似性以及一家银行对另一家银行资产的持有。

[15.](#) Geithner, 2014.

[16.](#) 要了解旧金山湾的物理模型及其在政策中的用途，请参见：Weisberg, 2012。

[17.](#) Stone et al., 2014.

[18.](#) Dunne, 1999&Raby, 2001.

03 多模型思维

[1.](#) Levins, 1966.

[2.](#) Page, 2007、2017.

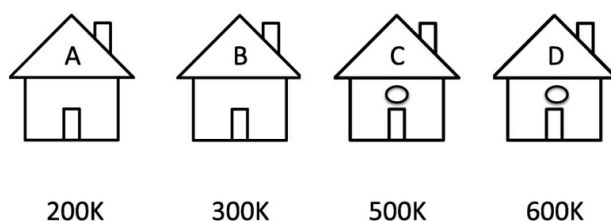
[3.](#) 关于群体的智慧，请参见：Surowiecki, 2006；关于狐狸是如何胜过刺猬的，请参见：Tetlock, 2005；关于计算机科学的集成方法，请参见：Patel et al., 2011。

[4.](#) 卢红和佩奇（2009）证明，独立的模型需要一个唯一的分类集。这也就是说，在二元分类模型中只存在一种创建一组独立预测的方法。

[5.](#) 关于多样性预测定理的详细阐述，请参见我以前出版的几本书：《多样性红利》《多样性与复杂性》等。关于经济预测所用的数据，请参见：Mannes、Soll&Larrick, 2014。

[6.](#) 考虑下图中用A、B、C和D表示的四幢平房及其市场价值。根据这些平房是否包含了一个录音室由“门”上方的圆圈符号表示），创建两

个类别。平房A和B不包含录音室，因此它们属于同一个类别；而平房C和D则包含了一个工作室，因此它们属于第二类。



四幢平房及其市场价值

首先计算出这些平房价格的总变差，即每幢房子的价值与均值之差的平方和。这四幢平房的均值等于400 000美元，因此总变差等于100 000：总变差= $(200-400)^2 + (300-400)^2 + (500-400)^2 + (600-400)^2$ 。

为了计算分类误差，假设我们知道每个类别内的真实均值：第一个类别为250 000美元，第二个类别为550 000美元（平房已转变为录音工作室）。但是这种分类将不同价值的房子混同在了一起。剩余的变差等于分类误差：分类误差A和B= $(200-250)^2 + (300-250)^2 = 5 000$ 、分类误差C和D= $(500-550)^2 + (600-550)^2 = 5 000$ 。总分类误差等于这两个数字的总和，即10 000。

为了计算出估值误差，假设模型预测平房A和B的价值均为300 000美元、同时预测平房C和D的价值均为600 000美元。估值误差等于每个类别的预测值与真实均值之间的差的平方和。估值误差A和B= $(300-250)^2 + (300-250)^2 = 5 000$ ，估值误差C和D= $(600-550)^2 + (600-550)^2 = 5 000$ 。于是总估值误差等于10 000。

总模型误差等于预测值与实际值之间的差的平方和：模型误差= $(200-300)^2 + (300-300)^2 + (500-600)^2 + (600-600)^2 = 20 000$ 。因此不难注意到，模型误差等于分类误差和估值误差之和。

7. 读者如果想了解二维社会互动的旋转玻璃模型，请参见：

Brock&Durlauf, 2001。爱德华·格莱泽（Edward Glaeser）等人则给出了一个关于犯罪的一维模型。另外，德鲁·弗登伯格（Drew Fudenberg）和戴维·莱文（David Levine）也构建了一个关于大脑的经济模型。

8. 斯塔夫罗斯·尼阿科斯并不是第一个试图利用货船超大空间的人。1858年，因设计建造大西部铁路的传奇英国工程师伊桑巴德·金德姆·布鲁内尔（Isambard Kingdom Brunel）建成了一艘近220米长的大船，并命名为“大东方号”（SS Great Eastern）。事实证明，作为一艘货轮，大东方号是失败的。由于缺乏适当的水动力模型，导致大东方号整体设计不佳。这艘船只有在最慢的速度下才能够航行。幸运的是，它最终还是派上了用场——被用于铺设跨越大西洋的海底电缆。关于多模型思维在设计船舶时的作用，请参见：West, 2017。

9. 以磅和英寸为单位时，身体质量指数 = $\frac{703 \times \text{体重}}{\text{身高的平方}}$ 。

10. 勒布朗·詹姆斯（LeBron James）身高6英尺8英寸，体重大约250磅，身体质量指数为27.5。凯文·杜兰特（Devin Durant）身高6英尺9英寸，体重235磅，身体质量指数为25.2。2012年和2016年奥运会十项全能金牌得主阿什顿·伊顿（Aston Eaton）身高6英尺1英寸，体重185磅，身体质量指数为24.4。另一位著名运动员，2008年奥运会十项全能金牌得主布莱恩·克莱（Brian Clay）的身体质量指数则为25.8。

11. Flegal et al., 2012.

12. 假设这只小鼠3英寸长、1英寸高、1英寸宽，同时假设这只大象10英尺高、10英尺长、5英尺宽。大象的表面积为400平方英尺，也就是57 600平方英寸。大象的体积为500立方英尺，相当于864 000立方英寸。

13. 杰弗里·韦斯特（Geoffrey West, 2017）和他的同事们已经构建了一个更精细和更准确的模型，它预测新陈代谢的规模与质量的四分之三次幂成比例。而且在80多年之前，科学家就已经在数据中找到了与之相近的定律，即克雷伯定律（Kleiber's law）。
14. 发送的简历内容完全相同，但是申请人的名字改变了的实验表明，女性获得的报酬低于男性，得到的评价也低于男性；请参见：Moss-Racusina et al., 2012。
15. 一个男性员工成为CEO的概率等于他连续获得15次晋升的概率。概率比，也就是男性成为CEO的概率与女性成为CEO的概率之间的比值。按照我们的假设，男性每次升职的概率为50%，女性每次升职的概率为40%，于是概率为 $(1.25)^{15}=28.4$ 。
16. Dyson, 2004.
17. Breiman, 1996.

04 对人类行为者建模

1. Haidt, 2006; Simler&Hanson, 2018.
2. 要证明这个结果，需要利用一些微积分知识。假设一个人的收入为 M ，一单位食品的价格为1美元，一个单位的住房服务的价格等于 P_H 。那么我们可以将他的预算约束写成 $M=F+P_H \times H$ 。而这当然就意味着 $F=M-P_H \times H$ 。可以将其效用函数写成

$$U(H)=(M-P_H \times H)^{\frac{2}{3}} H^{\frac{1}{3}}$$

为了求这个函数的最大值，我们对 H 求导数并令它等于零：即

$$-\frac{2}{3}P_H\left(\frac{H}{M-P_H\times H}\right)^{\frac{1}{3}}+\frac{1}{3}\left(\frac{M-P_H\times H}{H}\right)^{\frac{2}{3}}=0$$

把第一项移到等式的右侧，交叉相乘，可以得到： $2P_H\times H=M-P_H\times H$ 。在预算约束的条件下，用 $2P_H\times H$ 替换 $M-P_H\times H$ ，可以得到 $M=2P_H\times H+P_H\times H$ ，或者 $M=3P_H\times H$ 。因此，他将收入的1/3用于住房。

3. 在美国，这是一个不错的近似值。
4. 这个定理的正式表达如下： $X=\{A, B, C, \dots, N\}$ 表示一个结果的有限集，并设定一个抽奖，令它成为这个结果集上的一个概率分布，即 $L=(pA, pB, \dots, pN)$ 。如果对彩票的偏好($>$)满足完备性、传递性、独立性和连续性，那么这个偏好可以表示为一个连续的有实数值的效用函数，可以给每个彩票赋予一个效用。完备性是指，任何两个彩票， L 和 M ，都可以进行比较；传递性是指，如果 $L>M$ ，并且 $M>N$ ，那么 $L>N$ ；独立性是指，如果 $L>M$ ，那么给定任意一个彩票 N 和任意一个概率 $p>0$ ，都有 $pL+(1-p)N>pM+(1-p)N$ ；连续性：如果 $L>M$ 且 $M>N$ ，那么存在概率 p 使得 $pL+(1-p)N\sim M$ 。这个定理的正式证明有点繁杂，下面只给出证明的简要思路。假设存在最优结果 B 和最差结果 W ，并且假设 B 的效用等于1、 W 的效用等于零。给定任何其他结果 A ，根据连续性公理，存在概率 p ，使得人们在肯定能够得到 A ，与以 p 的概率得到 B 并以 $(1-p)$ 的概率得到 W ，两者之间无差异。利用上面的符号，我们把它写成 $A\sim pB+(1-p)W$ 。然后，我们再将 A 的效用指定为 p 。按此思路推导（这需要一点点数学知识），可以证明，一个人越喜欢某个结果（或彩票）， p 就越大。这样一来，我们也就魔术般地将排序转变成了数字。对于这个定理的完整的证明，请参见：Von Neumann&Morgenstern, 1953。

5. Rust, 1987.

6. Camerer, 2003.
7. Harstad&Selten, 2013.
8. 关于以理性选择模型为基准的理由，请参见：Myerson, 1999。
9. 对于这个领域的一个早期的综述，请参见：Camerer、Loewenstein&Prelec, 2005。
10. Kahneman, 2011.
11. 开放科学协作组织的原始论文（2015），导致了更多的复制结果的研究，它们得出相似的百分比。
12. 要想了解关于被试多样性的重要性，请参见：Medin、Bennis&Chandler, 2010。
13. Berg&Gigerenzer, 2010.他们充分表达了这种批评意见，认为它破坏了基于数学的心理模型。
14. Kahneman&Tversky, 1979.
15. 例如，治疗方案A——肯定能够拯救40%的患者，治疗方案B——有50%的机会拯救所有人。损失框架：治疗方案A'——肯定会有60%的患者死亡，治疗方案B'——有50%的机会没有任何人死亡，有50%的机会所有人都死亡。根据前景理论，大多数医生在收益框架下都会选择治疗方案A，而在损失框架下都会选择B'。
16. 关于双曲贴现的含义的早期研究，请参见：Thaler, 1981; Laibson, 1997。
17. 双曲贴现的公式可以写成如下更一般的形式：

$$H(r, t) = \frac{1}{(1 + rt)^{\frac{\beta}{\alpha}}} \circ$$

18. Gigerenzer&Selten, 2002.

[19.](#) Gode&Sunder, 1993.

[20.](#) Gigerenzer&Selten, 2002.

[21.](#) 在弗农·史密斯（Vernon Smith）的诺贝尔奖获奖演说中，他指出：“生态理性，要使用推理，也就是用理性的重建来考察个人的行为，但根据的是他们的经验和常识。人们遵循规则但不一定能清楚地将它们表达出来。但是，它们是可以被发现的。”请参见：Smith, 2002。

[22.](#) Arthur, 1994.

[23.](#) Lucas, 1976; Campbell, 1976.

[24.](#) 关于模型如何阐明可能发生的事情的进一步讨论，请参见：de Marchi, 2005。关于基于方程和规则的模型（行为者在类似的博弈中做出类似的行为），请参见：Gilboa&Schmiedler, 1995; Bednar&Page, 2007、2018。

05 正态分布

[1.](#) 给定一个数据集 $\{x_i, \dots, x_N\}$ ，方差等于数据与均值之间距离的平方的平均值，可以写为：

$$\text{VAR} = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$$

标准差则等于数据与均值之间距离的平方的平均值的平方根，即：

$$\text{s.d.} = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$$

[2.](#) 这里有好几个条件，而且只要满足其中任何一个条件就足够了。林德伯格条件（Lindeberg condition）是比较常见的一种，它要求，随着

变量的数量的无限增加，来自任何一个变量的总变差的比例将收敛为零。

3. Lango et al., 2010.

4. 对于一般情况，假设独立性，我们可以得到如下表达式：

$$\bar{\mu} = \frac{\sum_{i=1}^N \mu_i}{N}$$

$$\bar{\sigma} = \sqrt{\frac{\sum_{i=1}^N \sigma_i^2}{N^2}}$$

用 σ 代替 σ_i ，就可以得到 $\bar{\sigma} = \sqrt{\frac{N\sigma^2}{N^2}} = \frac{\sigma}{\sqrt{N}}$ 。

5. Wainer, 2009.

6. 两个标准差的阈值（5%显著性）是一个惯例。尽管经常受到批评，但它仍然是社会科学家最常使用的。当然，6%显著性水平的大相关系数，可能比4.9%显著性水平的小相关系数更值得注意。进一步的讨论，请参见：Ziliak&McCloskey, 2008。

7. Gawande, 2009.

8. 随机变量的乘积分布之所以称为对数正态分布，是因为分布的对数将会是正态分布的。个中的原理，我们可以简述如下。首先，我们写出一些数字的乘积， $y = x_1 \times x_2 \times x_3 \times \dots \times x_n$ ，并将它写为如下以10为底的幂的乘积的形式：

$$10^{\log_{10}(y)} = 10^{\log_{10}(x_1)} \times 10^{\log_{10}(x_2)} \times 10^{\log_{10}(x_3)} \dots 10^{\log_{10}(x_n)} = 10^{\log_{10}(x_1) + \log_{10}(x_2) + \log_{10}(x_3) + \dots + \log_{10}(x_n)}$$

然后，我们对两边同时取以10为底的对数，得到以下的等式：

$$\log_{10}(y) = \log_{10}(x_1) + \log_{10}(x_2) + \log_{10}(x_3) + \dots + \log_{10}(x_n)$$

因此，变量 y 的对数可以写为随机变量的对数之和。这些随机变量

的对数也是随机变量，只要它们的方差满足中心极限定理的条件，它们的和 $\log_{10}(y)$ 就是正态分布的。

9. Limpert、Stahel&Abbt, 2001.

10. 这个思想最早可以追溯到罗伯特·吉布拉特（Robert Gibrat）1931年的研究。

06 幂律分布

1. 关于这场大洪水的影响和文化意义，请参见：Parrish, 2017。

2. 在这个例子中，指数假设为2。这个例子借用自：Clauset、Young&Gleditsch, 2007。

3. 关于本章所述模型的技术性细节以及其他关于幂律的例子，请参见：Newman, 2005。

4. Newman, 2005; Piantadosi, 2014.

5. 常数 C 使所有结果的总概率等于1。根据这个定义，幂律分布满足标度不变性（scale invariance），如果改变了用来衡量结果的单位，那么分布的形状仍然不会发生变化。

6. 在计算这些概率的时候，可以先求出事件在一年内未发生的概率。例如，如果事件的概率为 $1/2$ ，那么这个事件在一年内不发生的概率等于 $(0.999)^{365}=0.69$ 。因此，该事件发生的概率等于31%。一个概率为一百万分之一的事件在一个世纪内不发生的概率也可以用类似的方法计算。

7. Cederman, 2003; Clauset、Young&Gleditsch, 2007; Roberts&Turcotte, 1998. 恐怖袭击导致 x 人死亡的概率大体上可以写为一个常数项： 0.06 除以 x 的平方。对于 x 只取整数值的离散分布，幂律分布则可以写为 $p(x) = 0.608x^{-2}$ 。之所以选择系数 0.608 ，是为了使概率总和

等于1。因为 $\sum_{n=1}^{\infty} \frac{1}{n^2} = 1.644934$ ；0.608与1.644934的乘积等于1。

8. 对于幂律，我们对两边同时取对数，并将 $y=Cx^{-a}$ 变换为 $\log(y) = \log(C) - a\log(x)$ ，这是一个 $\log(y)$ 关于 $\log(x)$ 的线性方程。当我们绘制出 $\log(y)$ 随 $\log(x)$ 而变的图形时，将会得到一条直线。对于指数分布， $y=C \times A^{-x}$ 。如果取两边的对数，就可以得到 $\log(y) = \log(C) - x\log(A)$ ，这意味着 $\log(y)$ 对于 x 是线性的。换句话说， $\log(y)$ 将随 $\log(x)$ 而快速减少，从而形成凹的图形。

9. 如果我们取对数正态分布的对数，就会得到以下形式的方程： $\log(y) = C - b \times \log(x) - \frac{\log(x)^2}{\sigma^2}$ ，其中， σ 是这个对数正态分布的标准偏差的自然对数，作为该分布方差的一个代理变量。对于很大的 σ ， $\log(x)^2$ 的贡献将很小，直到 $\log(x)$ 的值变得足够大，可以引起图形出现下降为止。

10. 关于对数正态分布和幂律分布的正式区别，请参见：Broido&Clauset, 2018。研究认为，许多曾经被认为具有幂律分布特点的网络，实际上并不是幂律分布的。

11. 要了解齐普夫的单词频率定律和其他可选模型，请参见：Piantadosi, 2014。如果事件的大小分布满足幂律，那么排名也符合这个定律。在开区间 $[1, \infty)$ 上，指数为 a 的幂律分布具有形式 $p_a(x) = ax^{-a}$ 。假设有100个事件。令 S_R 表示第 R 大的事件的预期大小。那么某个事件大于 S_R 的概率必定等于 $\frac{R}{100}$ 。例如，如果 $R=3$ ，那么一个事件大于 S_3 的概率必定等于3%。因此：

$$\int_{S_R}^{\infty} ax^{-a} = \frac{R}{100}$$

从中可以解得 $S_R^{-(a-1)} = \frac{R}{100}$ ，而它又可以改写为：

$$S_R = \left(\frac{R}{100} \right)^{\frac{-1}{(a-1)}}。$$

在 $a=2$ 的特殊情况下，上述表达式就变成了 $S_R = \left(\frac{R}{100} \right)^{-1}$ 。

12. Bak, 1996. 这个模型的适用范围很广，至今仍然没有看到它的边界。学者们已经利用这个模型来解释经济波动、战争中的死伤、恐怖主义行为、进化中的间断均衡以及交通拥塞等理论和现实问题了。请参见：Paczuski&Nagel, 1996; Sneppen et al., 1995。
13. Salganik、Dodd&Watts, 2006; Ormerod, 2012. 幂律分布还意味着，构成概率分布的很大一部分是非常多的小事件。这些小事件合到一起，也足以产生与大事件同等规模的经济价值。请参见：Anderson, 2008b。互联网的出现，使零售商能够出售大量的书籍、电影和音乐，尽管有些东西只能吸引少数人。一个只卖一本畅销书、卖出了500万的零售商，获得的利润可能与另一个销售1万种书、每种书只卖出500本的零售商差不多。
14. 有一个模型说明了这种情况是怎样出现的，请参见：Denrell&Liu, 2012。
15. 地质学家使用里氏标准来表示地震震级。震级等于地震强度的对数。因此，里氏6级地震的强度是里氏5级地震的10倍。地震的强烈可以用齐普夫定律来预测（但不是地震的时间），请参见：Merriam&Davis, 2009。
16. Eliot、Golub&Jackson, 2014. 他们给出了增大连通性如何使失败的可能性降低的明确模型。
17. Mayt、Levin&Sugihara, 2008.

- [18.](#) Stock&Watson, 2003.
- [19.](#) Carvalho&Gabaix, 2003.
- [20.](#) Claridat、Galí&Gertler, 2000.
- [21.](#) 我要感谢塞思·劳埃德（Seth Lloyd），是他提醒我注意到了这个例子。
- [22.](#) 我们让工资分布等于随机变量 x 的100 000倍，其中， $p(x)=2x^{-3}$ 从1到 ∞ 。变量 x 的均值等于2，因此工资分布的均值为200 000美元。
- [23.](#) Weitzman, 1979.他给出的模型在更一般的情况下证明了这个结果。
- [24.](#) Bell et al., 2018.

07 线性模型

- [1.](#) 有价值的葡萄酒，如波尔多葡萄酒，都要由专家加以排名。这些葡萄酒同时也在市场上定价。价格和排名都可以作为质量的代理变量。奥利·阿申费尔特（Orley Ashenfelter）根据冬季降雨量、收获季降雨量和9月份的平均温度，拟合出了关于波尔多葡萄酒的质量的（对数）线性模型。请参见：Ashenfelter, 2010。线性模型将因变量的对数表示为自变量的对数的线性和，例如：

$$\log(y) = b_0 + b_1 \log(x_1) + b_2 \log(x_2)$$

这个表达式意味着，因变量也可以写成自变量的乘积。只需对上式的两边同时取以 e 为底的指数，就可以做到这一点。于是，我们得到：

$$y = e^{b_0} (x_1)^{b_1} (x_2)^{b_2}$$

通过取对数，乘式就可以变加式，然后就可以应用线性回归的工具了。以年份葡萄酒的价格为因变量，阿申费尔特的模型的拟合度指

标 (R^2 ，即，得到了解释变差的百分比) 达到了83%左右。有证据表明，它比葡萄酒专家的判断更加准确地预测了葡萄酒的价格。阿申费尔特的模型甚至可以预测专家的估值的变化。罗伯特·帕克 (Robert Parker)，一位著名的葡萄酒评酒专家，最初将Pomerol和St. Emilion于1975年出品的葡萄酒评为95分（满分为100分）。但是，在阿申费尔特的模型的预测中，对这些酒的质量排名则较低。后来，到了1983年，罗伯特·帕克将他对这些酒的排名降低到了低于平均水平，一如阿申费尔特的模型的预测。请参见：Storchmann, 2011。

2. Xie, 2007.

3. Ryall&Bramson, 2013.

4. Mauboussin, 2012.他讨论了如何利用这个方程式指导合理的管理决策。

5. Bertrand&Mullainathan, 2001.

6. Shapiro、Meschede&Osoro, 2013.他们并没有把相关性和因果关系混为一谈。如果两个变量不相关，就不应该期待它们之间存在因果关系。

7. 为了找到最优的那条线，我们使用的是支持向量机 (SVM)，这是一种类似于回归的方法。两者之间关键区别在于，支持向量机将所有数据分成正的组或负的组，然后找到最大化到每组中最近点的直线。如果不存在这样的直线，就会对偏离行为进行处罚。回归则考虑到所有点的距离。

08 非线性模型

1. Arthur, 1994.

2. 30次倍增为 2^{30} ，超过10亿。

3. Karlsson, 2016.

4. Ebbinghaus, 1885.

5. 关注大脑反应的研究者发现，即便是享用巧克力，消费了一定数量后，人们也会开始觉得厌恶，请参见：Small et al., 2001。

6. 与本书中的许多例子一样，这个例子我也是从拉夫和马奇出版于1975年的那本书中借用来的。

7. 假设你每年投资3 000美元。如果股票的价格维持在每股15美元的水平不变，那么你每年可以买200股。如果价格在20美元到10美元之间变动，那么在价格最高的年份你只能购买150股，而在价格较低的年份你可以购买300股。平均而言，你可以购买225股，这比你在价格不变的情况下购买的数量更多。

8. 关键假设是，在柯布-道格拉斯模型中的指数 a 和 $(1-a)$ 的总和为1。这个假设意味着如果我们将工人数量和资本数量加倍，总产出也会增加一倍，即：

$$\text{产出} = \text{常数} \times (2 \times \text{工人})^a (2 \times \text{资本})^{(1-a)}$$

展开各项并重新排列，就可以得到输出加倍的方程式：

$$\text{产出} = \text{常数} \times (2 \times \text{工人})^a (2 \times \text{资本})^{(1-a)}$$

9. 在第二年，每天产出等于 $100\sqrt{2} = 141$ 。在第三年，产出则 $100\sqrt{3} = 173$ 。增长率等于每年的产出增长百分比。

10. 计算过程如下。第2年，机器：290；产出：1 702 ($=100\sqrt{290}$)。投资 $= (0.2) \times 1\ 702 = 340$ ，因此消费 $= 1\ 362$ 。折旧 $= (0.1) \times 290 = 29$ 。第3年，机器：601，产出：2 453 ($=100\sqrt{601}$)。投资 $= (0.2) \times 2\ 453 = 491$ ，因此消费 $= 1\ 962$ 。折旧 $= (0.1) \times 601 = 60$ 。

11. 长期均衡可以通过求解 M^* 求出，即，使得 $0.2 \times 100 \sqrt{M^*} = 0.1 M^*$ 。 $M^* = 40\,000$ 。
12. 在完整的索洛增长模型中，会用一个参数 a 替换掉这里的平方根函数，像在柯布-道格拉斯模型中一样，还包括劳动市场。
13. 为了求出均衡，我们设定投资等于折旧： $s \times A \sqrt{L} \sqrt{K^*} = d \times K^*$ 。因此，均衡机器数量 K^* 要满足等式 $\sqrt{K^*} = A \frac{s}{d} \sqrt{L}$ 。将之代入产出函数，可以求得产出等于 $A^2 L \frac{s}{d}$ 。
14. 经济学家罗伯特·戈登（Robert Gordon, 2016）则采取了一种悲观的看法，认为即将出现的新技术将使得 A 出现大幅增长。更高级的增长模型将技术视为其他变量的函数。例如，保罗·罗默（Paul Romer, 1986）的模型假设增长源于商品种类的增加：随着经济的增长，产品种类也在增长。经济学家马丁·魏茨曼（Martin Weitzman, 1998）则对思想的产生和组合给出了明确的模型。
15. 关于技术变革中的滞后，请参见：Arthur, 2011。
16. 例如，以前从来没有通过固定电话的乡村地区，可以直接建造无线基部并提供手机服务。亚历山大·格申克伦（Alexander Gerschenkron, 1952）将这种现象称为“后发优势”。
17. 皮凯蒂（2014）揭示，从1700年到2012年，世界GDP的年平均增长率仅为1.6%，而且这种增长中的一半是来自人口增长。将72法则应用于这0.8%的年增长率，我们可以发现，在300年的时间内，平均生活水平提高了大约10倍。

09 与价值和权力有关的模型

1. 这个值是这样计算出来的：将桨手能够增加价值的概率 $1/6$ 乘以桨手能够增加的价值期望值（后者等于 $4/5 \times 10$ 再加 $1/5 \times 2$ ）。同时注意

到，夏普利值的总和为10，也就是这个博弈的总价值。

2. 正式的计算过程如下：在阿伦的6个想法中，只有两个是他独有的，一个是贝蒂也提出来的，3个是每个人都提出来的。如果他是最先进入的那个人（这个事件发生的概率为 $1/3$ ），那么增加的价值为6。如果他是第二个进入的，那么能够增加两个独有的想法，这两个想法都是其他人所不知道的；而且，在贝蒂和卡洛斯进入之前，他还有 $1/2$ 的机会再增加一个想法。所以，他总共增加了2.5个想法。如果他是第三个进入的，那么他可以增加两个他独有的想法。因此，他的夏普利值等于3.5。贝蒂提出了4个与另外一个人共同的想法、3个与所有人共同的想法。因此，她的夏普利值等于3。最后，卡洛斯提出了3个与另一个人共同的想法以及3个与所有人共同的想法，所以他的夏普利值等于2.5。需要特别注意的是，夏普利值的总和为9，即所有想法的总数。
3. 衡量投票权力的另一种方法是，班茨哈夫-彭罗斯指数（Banzhaf-Penrose index）。它的计算方法是，根据所有可能的获胜联盟确定所有可能成为“关键投票者”的政党的数量，而赋予一个政党的值就等于它发挥关键作用的次数的值除以这个数字。请参见：Banzhaf, 1965。
4. Groseclose&Snyder, 1996.

10 网络模型

1. 对网络模型的全面讨论，请参见：Newman, 2010；想要了解网络模型在社会科学领域的应用，请参见：Jackson, 2008；Tassier, 2013。
2. 给定任何一个节点，它的最小路径长度可以为1到4个节点，也可以为2到4个节点，还可以为3到4个节点，由此，从每个节点开始的最小路径共有12个节点。平均而言，节点的最小路径是访问另一个节

点。由此可以得出，每个节点的平均介数等于1/12。根据对称性，所有节点必定都具有相同的介数。

3. 关于社区探寻算法的更多细节，请参见：Newman, 2010。由于可能性的数量不同，用这些算法构造出来的分区可能有所不同。具有100个节点的网络可以用超过1.9亿个不同的方法进行分区。删除边的次序的随机性，相同的算法通常也会产生不同的分区。通过应用多个算法并多次应用每个算法，就可以增强推理的鲁棒性。
4. 根据中心极限定理，我们知道度是正态分布的，并且其均值为 $\frac{2E}{N}$ ，因为每个边连接两个节点。
5. Watts&Strogatz, 1998.
6. 关于网络形成的正式分析，请参见：Newman, 2010。
7. Ugander et al., 2011。
8. 给定一个具有 N 个人的网络，令 d_i 表示节点 i 的邻居的数量，即度。平均度 $\bar{d}$ 可以表示为如下形式：

$$\bar{d} = \frac{\sum_{i=1}^N d_i}{N}$$

平均度 $\bar{d}$ 等于这个节点的期望邻居数。而在计算邻居的邻居的平均数量时，度为 d_i 的节点会被“数上” d_i 次，即，每个邻居一次。因此，节点的邻居的邻居的总数 N_2 可以表示如下：

$$N_2 = \sum_{i=1}^N d_i^2$$

因此，为了求得一个节点的邻居的平均度，我们必须用这个数除以邻居的总数，得出的结果为 $N \bar{d}$ 。因此，只要证明下式就足够了：

$$\frac{\sum_{i=1}^N d_i^2}{Nd} \geq \bar{d}$$

而它可以重写为

$$\frac{\sum_{i=1}^N d_i^2}{N} - \bar{d}^2 \geq 0$$

左边的项等于度分布的方差。如果任意两个节点具有不同的度数，那么度分布就具有正的方差，因此，节点的邻居的平均度，超过节点的平均度。

[9.](#) Eom&Jo, 2014.

[10.](#) Dodds、Muhamad&Watts, 2003.

[11.](#) Newman, 2010; Jackson, 2008.

[12.](#) Granovetter, 1973.

[13.](#) 四度朋友的数量是通过对如下8组节点求和来计算的：CRCR=4 000 000，CRRC=4 000 000，RCRC=4 000 000，CRRR=800 000，RCRR=800 000，RRCR=800 000，RRRC=800 000，RRRR=160 000。

[14.](#) Albert、Albert&Nakarado, 2004.

[15.](#) Groysberg, 2012. 条纹模型可以解释它们为何无法成功回归均值。

[16.](#) 关于填补结构洞的人的价值，请参见：Burt, 1995。

[17.](#) 关于对教师网络的影响，请参见：Frank et al., 2018。

[18.](#) 具有非零价值的联盟是{A, B}、{B, C}和{A, B, C}。在前两个联盟中，每个联盟的价值都等于10。第三个联盟的附加价值等于-6。第三个联盟的独立价值等于14，但是这个联盟是由两个联盟组成

的，每个联盟的价值都是10。因此，联盟的价值等于14减去20：-6=（14-10-10）。然后，我们可以为每个联盟中的每个博弈参与者分配如下夏普利值：联盟{1, 2}：博弈参与者1为5，博弈参与者2为5；联盟{2, 3}：博弈参与者2为5，博弈参与者3为5；联盟{1, 2, 3}：博弈参与者1为-2，博弈参与者2为-2，博弈参与者3为-2。对这些值求和，就可以得到迈尔森值。

11 广播模型、扩散模型和传染模型

1. 这些模型都假设了离散的时间（步长），例如，天或周，并使用差分方程来描述“明天”被感染（或变得知情）的人的数量与“今天”被感染（或变得知情）的人的数量之间关系。连续时间模型则需要运用微分方程和微积分。但是，如果改用连续时间模型，那么结果从定性的角度来看不会有所不同。
2. 将第一个方程式插入第二个方程式，可以得出： $36\ 000 = 20\ 000 + 20\ 000 - P_{\text{broad}} \times 20\ 000$ ，进一步化简为 $4\ 000 = P_{\text{broad}} \times 20\ 000$ ，所以我们有： $P_{\text{broad}} = 0.2$ ，以及 $N_{\text{POP}} = 100\ 000$ 。
3. Griliches, 1988.
4. 假设，每个应用程序的初始销售量（ I ）等于100。一开始，设定 $P_{\text{diffuse}} = 0.4$ ，同时 $\text{POP} = 1\ 000\ 000$ 。第三期的新销售量等于 $0.4 \times \frac{100 \times 900}{1\ 000} = 36$ 。与此类似，我们可以计算出未来各期的销售数据。对于第二组数据，令 $P_{\text{diffuse}} = 0.3$ ， $\text{POP} = 100\ 000$ 。第二期的新销售量 $0.3 \times \frac{100 \times 999\ 900}{1\ 000\ 000} = 30$ 。用同样的方法，可以求得后续各期的销售量。
5. 巴斯把率先采用某种技术或购买某种产品的人称为“创新者”，而把跟着这样做的人称为“模仿者”（Bass, 1969）。

6. 对 R_0 的正式推导始于这样一个观察结果：当感染者的人数很小的时候，易感者的数量大体上相当于相关人群的数量。因此，为了减少变量的数量，我们可以用相关人群的数量代替易感者的数量。然后，我们可以将感染者人数的变化写成最初感染者人数的线性函数的形式（见正文中的专栏）。对 R_0 的正式推导过程如下。当出现了一种新的传染病时，它最初只会感染少数人。我们用 I_0 表示这个很小的感染者的数量。将之代入SIR模型，第一期中感染者的数量等于：

$$I_1 = I_0 + P_{\text{contact}} \times P_{\text{spread}} \times \frac{I_0}{N_{\text{POP}}} \times S_0 - P_{\text{recover}} I_1$$

以 N_{POP} 为 S_0 的近似，上面这个式子就变

为 $I_1 = I_0 + P_{\text{contact}} \times P_{\text{spread}} \times I_1 - P_{\text{recover}} I_1$ 。由此我们可以推出，感染者的数量将上升，当且仅当 $P_{\text{contact}} \times P_{\text{spread}} > P_{\text{recover}}$ ，而这又等价于下式：

$$\frac{P_{\text{contact}} \times P_{\text{spread}}}{P_{\text{recover}}} \geq 1$$

7. 强制隔离会使接触的概率降低到接近零的水平，从而有效降低基本再生数，但是成本很高。在20世纪初，结核病（ $R_0 \approx 3$ ）在美国每年都会导致超过10万人死亡，各州都不得不加征财产税，以建立强制性的疗养院以收容隔离病人，因为手术治疗如摘除肺叶和整个肺，甚至往肺部空洞处塞入乒乓球，但最终全都被证明是无效的。请参见：Dubos, 1987。
8. 为了求疫苗接种阈值，我们需要对SIR模型进行一些修正，以便容纳接种了疫苗的人群。一种传染病要得到传播，感染者必须与未接种疫苗的人（概率为 $(1-V)$ ）接触（接触概率为 P_{contact} ），并且传染病必须是会扩散的（ P_{spread} ）。第一期的差分方程可以写为如下形式：

$$I_1 = I_0 + P_{\text{contact}} \times P_{\text{spread}} \times \frac{I_0}{N_{\text{POP}}} \times (1 - V) \times S_0 - P_{\text{recover}} I_t$$

运用近似值 $S_0 = N_{\text{POP}}$ （就像在推导 R_0 时一样），上式就可以改写为：

$$I_1 = I_0 + P_{\text{contact}} \times P_{\text{spread}} \times I_0 \times (1 - V) - P_{\text{recover}} I_t$$

当且仅当 $P_{\text{contact}} \times P_{\text{spread}} \times (1 - V) > P_{\text{recover}}$ 时，感染者的人数才会增加。这个不等式可以重写为 $R_0 (1 - V) \leq 1$ 。展开并重新排列各项，我们有 $R_0 - 1 < V \times R_0$ 。只要两边同时除以 R_0 就可以得到我们想要的结果。

9. 关于从社会科学角度对SIR模型和群体免疫力的更详尽的分析，请参见：Tassier, 2013。
10. Stein, 2011.
11. Updike, 1960.
12. Tweedle&Smith, 2012.
13. Lamberson&Page, 2012b.
14. 数据源自网络。
15. Christakis&Fowler, 2009.
16. 对此，达蒙·森托拉（Damon Centola）和迈克尔·梅西（Michael Macy）（2007）称之为复杂的传染病（complex contagion）。

12 熵：对不确定性建模

1. Smaldino, 2013.
2. 数字 x 的以2为底的对数必定等于使2产生 x 的幂，因此 $\log_2(4) = 2$ ，且 $\log_2(2^N) = N$ 。在一般情况下， $\log_a(x)$ 等于使 a 产生 x 的幂。因

此，如果 $a^y=x$ ，则 $\log_a(x) = \log_a(a^y) = y$ 。

3. 我们可以将信息熵写成如下的长形式：

$$H = -\frac{1}{2} \times \log_2\left(\frac{1}{2}\right) - \frac{1}{8} \times \log_2\left(\frac{1}{8}\right) - \frac{1}{8} \times \log_2\left(\frac{1}{8}\right) - \frac{1}{8} \times \log_2\left(\frac{1}{8}\right) - \frac{1}{8} \times \log_2\left(\frac{1}{8}\right)$$

这个式子可以化简为：

$$\frac{1}{2}(1) + \frac{1}{8}(3) + \frac{1}{8}(3) + \frac{1}{8}(3) + \frac{1}{8}(3) = \frac{1}{2}(1+3) = 2$$

4. 多样性指数，即，概率的平方和的倒数，可以满足前两个公理以及乘法公理。以多样性指数来衡量，已知结果的多样性等于1。请参见：Page, 2007、2010a。

5. Wolfram, 2001; Page, 2010a.

6. 亚历山大总共列出了15个这样的属性。他将自己的想法，再加上很多精英的照片集结成册，自行出版了4本精美的书：《秩序的本质，第1册：生命现象》（*The Nature of Order, Book 1: The Phenomenon of Life*）（2002）、《秩序的本质，第2册：创造生命的过程》（*The Nature of Order, Book 2: The Process of Creating Life*）（2002）、《秩序的本质，第3册：生活世界的愿景》（*The Nature of Order, Book 3: A Vision of a Living World*）（2005），以及《秩序的本质，第4册：流光溢彩的大地》（*The Nature of Order, Book 4: The Luminous Ground*）（2004）。这4本书中，第2册与我们这里讨论的主题的关系最密切。

13 随机游走

1. 欲了解更多关于随机游走的知识，请参见：Mlodinow, 2009。

2. Taleb, 2001.

3. Turchin, 1998; Suki&Frey, 2017.

4. 需要注意的是，大数定律说明平均比例是收敛的，而中心极限定理则告诉我们白球比例的分布是正态的。

5. 一位三分球命中率为46%的篮球运动员连续投中9个三分球的概率大约为 $1/1\,000$ (0.46^9)。如果这个篮球运动员一直继续投三分，那么在10年的NBA生涯中（总共参加大约800场比赛），一次都做不到三分球9投9中的概率大约为47% (0.999^{800})。

30多年来，统计学家一直在争辩，篮球运动员和其他职业运动员是不是真的会出现“热手效应”，或者换句话说，任何一次投篮或罚球命中的概率是否与前一次的成功相关。请参见：Chance, 2009。他分析了乔·迪马吉奥（Joe DiMaggio）的56场比赛连胜纪录。在考虑“热手效应”的证据时，我们必须考虑到行为。如果一名篮球运动员认为自己的手感很好，那么可能会尝试更加困难的投篮。此外，如果防守方觉得对方哪个运动员手感太好，他们就会收紧对他的防守。这些行为反应在模型中可以通过加大难度来体现。特沃斯基等人没有发现可以证明“热手效应”的证据。乔舒亚·米勒（Joshua Miller）和亚当·圣胡尔霍（Adam Sanjurjo）（2015）则在以往的研究中的条件概率计算中发现了一个推理错误，进而证明以前那些声称没有发现“热手效应”的研究实际上支持了“热手效应”假说。以前研究的这种误差源于采样技术中的缺陷。这些研究收集了多名（例如）篮球运动员投篮命中和不中的序列；然后再计算随机选择的多次投篮后某次投篮命中的概率。这种采样程序会引入一种微妙的统计偏差。

为了说明这种偏差，不妨考虑如下这种情况：多名篮球运动员每人只投篮4次，而且每次投篮都有相同的命中概率，这样就有16种可能的投篮命中和未命中的序列。我们用B代表投篮命中，用M代表投篮不中。在这16个序列中，有6个是连续两次命中后再接着投篮

一次产生的，它们是：BBBB，BBBM，MBBB，BBMB，BBMM和MBBM。这些都是连续两次投篮命中后再投第三次篮的情形。如果抽到了BBBB，那么无论选择哪一个由两个B组成的序列，那么接下来再投篮命中的概率都等于100%。如果选择了MBBB，那么B跟着BB出现的概率也等于100%。但是，如果选择了BBBM，那么BB之后出现M的概率等于50%。最后，如果选择的是MBBM、BBMB或BBMM，那么M必定会在BB之后出现。求这6种情况的平均值：

$$p(B|BB) = \frac{1}{6} \left[1 + 1 + \frac{1}{2} + 0 + 0 + 0 \right] = \frac{5}{12}$$

之所以会产生偏差，是因为在序列BBBB中，有两个BB可以选择，但是在其他序列中，则只有一个（例如，BBMB）。前述采样程序使得BBBB的这两个子序列中的每一个被选中的概率都仅为BBMB中的那一个子序列的一半。这种偏差的含义是，如果不存在“热手效应”，那么这种采样程序就已经证明投篮命中后再投篮时更有可能不中。当然事实并非如此，而这就意味着投篮命中之后再投篮实际上更有可能命中。

6. 麦道夫几十年来一直宣称自己的客户每月的回报率为1.5%。他声称，无论市场大环境（“大盘”）如何变化，他的投资每个月都在升值。在经济衰退期间维持正收益，要比表现优于市场更加困难。当市场下跌时，一个人即使战胜了市场也仍然可能会出现负收益。整体市场下跌了80多个月，但是麦道夫却一直声称自己实现了正收益。如果我们愿意冒风险做出一个“英勇”的假设：即使在大盘下跌的时间超过3/4的情况下，麦道夫也能以某种神奇的方法取得正收益，那么他连续80期成功的概率大约为十亿分之一（ 0.75^{80} ）。
7. 只要了解到，一个随机游走的值是相同的、独立的随机变量的总和，我们就可以求解标准差。每个随机变量的均值都为零，而且取值为

+1或-1。因此，每个都具有1的标准差。设置 $\sigma=1$ 并且对所得到的总和应用西格玛平方根公式，就可以得出结果。

8. Newman, 2005.

9. Levinthal, 1991; Axtell, 2001.

10. Newman, 2005; Sneppen et al., 1995.

11. 湖泊是按面积计算的。我们的模型产生幂律直径。面积等于直径平方的常数倍，因此面积也将是幂律分布的。具体数据请参见：Downing et al., 2006。

12. 在平衡的轮盘赌轮上，所有间隔都应有相同的可能性。如果桌子有任何倾斜，那么球在上坡时更可能从外缘掉落。关于多恩·法默、诺曼·帕卡德（Norman Packard）和他们的朋友通过“可穿戴计算机”，利用这个现象并击败轮盘赌的故事，请参见网络。

13. N 次投注后，这随机游走的期望值等于 $N \times (\frac{18}{38} - \frac{20}{38}) = \frac{-N}{19}$ 。由于获胜的概率大约为 $\frac{1}{2}$ ，我们可以将这个值的标准偏差近似为 $\sqrt{N}$ 。确切的值则等于 $2\sqrt{\frac{18}{38} \times \frac{20}{38} N} = 0.9986 \sqrt{N}$ 。

14. L.皮尔（L. Peel）和A.克劳塞特（A. Clauset）（2015）将每个博弈模型化为一个单个序列，并发现得分序列表现出了反持续性：上次得过分的队伍不太可能接下来马上得分。鉴于球队交替发动进攻，这个结果应该是预料之中的。

15. Baxter, 2009.

16. 这个结果的证明需要先计算在 N 个步骤内返回起点的概率，然后把所有可能的 N 的结果都加总起来。

17. Samuelson, 1965.

[18.](#) Grossman&Stiglitz, 1980.

[19.](#) Lo&MacKinlay, 2007.席勒（2005）证明，市盈率相对较低的股票的
表现优于市场。

[20.](#) Mauboussin, 2012.

[21.](#) 从1967年到2017年的这个窗口期恰好正逢标准普尔500指数上涨周
期。在大多数窗口期，股票价格增长速度高于经济整体的增长速
度。

14 路径依赖模型

[1.](#) Hathaway, 2001.

[2.](#) Pierson, 2004.

[3.](#) Bednar&Page, 2007、2018.

[4.](#) Page, 2006.

[5.](#) 这里是我在2006年给出的一个证明概要。显然，只要证明在前 N 期抽
取出 K 个白球的概率等于 $1/(N+1)$ 即可。存在 $(N+1)$ 种可能性，
因为 K 最小可以取零值，最大则可以等于 N 。 N 个周期中抽取出给定
 K 个白球序列的概率可以写为 N 个分数的乘积。这些分数的分母分
别是数字2到 $N+1$ 。分子分别为数字1到 K （白球）和1到 $(N-K)$
（灰球）。分子的乘积等于 $K!$ 乘以 $(N-K)!$ 而分母的乘积则等
于 $(N+1)!$ 。通过这个计算，就可以给出 K 个白球的特定序列的
概率。在 N 个期内排列 K 个球的可能方式的数量等于

$\binom{N}{K} = \frac{N!}{K!(N-K)!}$ 。因此，正好抽取出 K 个白球的总概率等于

$$\frac{K!(N-K)!}{(N+1)!} \cdot \frac{N!}{K!(N-K)!} = \frac{1}{(N+1)}$$

6. 我们用反证法来证明，假设结果不成立。从长期看，60%的结果是白色的，因此，瓮中60%的球将是灰色的。但是，这又意味着60%的结果将是灰色的。于是，出现矛盾，原命题得证。
7. 要想了解更多依赖用熵来度量不确定性的模型，请参见：Lamberson&Page, 2012b。
8. Lamberson&Pag, 2012b.
9. Page, 1997.
10. 风险价值也可以用年内任何时候损失超过10 000美元的概率来计算。
11. 这些计算依据的是以下事实：长度为 N 的随机游走的值的标准差等于 $\sqrt{N}$ 。2.5%对应于两个标准差。

15 局部互动模型

1. 在物理学中，局部多数模型通常称为伊辛模型（Ising model）。局部多数模型是投票者模型的一个变体。投票者模型假设了随机选择各种规模的邻居。请参见：Castellano、Fortunato&Loreto, 2009。
2. 对于沿4条边的元胞，我们将顶部的边连接到底部的边，创建一个圆柱体，然后将左侧的边连接到右侧的边，创建一个圆环。
3. 在局部多数模型的另一个版本中，元胞可以被同时激活，或者根据更新的动机来激活，其中处于相反状态的最大多数邻居元胞最先移动。如果元胞同时更新，局部多数模型可以生成循环。
4. 关于起立鼓掌的模型，请参见：Miller&Page, 2004。
5. 关于包括了跨域行动一致性的文化模型，请参见：Bednar et al., 2010。

6. 在这里，也是先将顶部连接到底部创建一个圆柱体，然后连接两端以形成一个“瘦小”的圆环。
7. 有一类称为扩散反应的模型也会在较窄的形状上产生条纹，并在较宽的形状上产生斑点。使用这些模型，科学家可以预测哪些动物会有条纹，哪些动物会有斑点，哪些动物会是纯色。答案取决于在发育阶段的哺乳动物胚胎的大小，即在模式形成时的大小，而不是成年动物的大小。否则，大象就会有斑点。请参见：Murray, 1988。
8. 我要感谢伯纳多·休伯曼（Bernardo Huberman）提供给我这个比喻。
9. 要证明这一点，需要编写一个计算机程序。这个程序能够生成随机数字序列或随机模式，并证明生命游戏模仿了计算机程序。正式的证明需要说明生命游戏在元胞自动机的集合中是通用的。请参见：Berlekamp、Conway&Guy, 1982。
10. Dennett, 1991; Hawking&Mlodinow, 2011.

16 李雅普诺夫函数与均衡

1. Nagel, 1995.
2. 我要感谢珍妮·贝德纳，她提供了关于逐底竞争的几个例子，除此之外，她还提供了本书的很多例子。
3. Page, 2001.
4. 外部性（即便是负外部性）的存在，不能排除构建李雅普诺夫函数的需要。局部多数模型和路径选择模型都包含了负外部性。在局部多数模型中，当一个元胞改变了自己状态时，它会对处于相反状态的邻居施加负外部性。然而，它与现在匹配的邻居产生的正外部性更大。
5. Guy, 1983.如果你够大胆，那么就从27开始试一试。

17 马尔可夫模型

1. 正式的研究论文需要利用统计技术更准确地估计转移概率并计算出误差范围。它还要检验转移概率在那个时间段内是不是保持不变的。如果转移概率取决于人均收入，那么就不会。请参见：Przeworski et al., 2000。

2. Flores&Nooruddin, 2016.

3. Tilly, 1998.

4. 采用瓷砖的人的均衡百分比等于从油毡转移到瓷砖的概率除以从油毡转移到瓷砖的概率与从瓷砖转移到油毡的概率的总和。这就等于

$$\frac{\frac{1}{4} \cdot \frac{1}{10}}{\frac{1}{4} \cdot \frac{1}{10} + \frac{3}{4} \cdot \frac{1}{60}} = \frac{\frac{1}{40}}{\frac{3}{80}} = \frac{2}{3}。我们还可以通过如下更一般的模$$

型。令D表示拥有昂贵的耐用品的人的百分比，C表示拥有更便宜的商品的人的百分比。再令BUY (C) > $\frac{1}{2}$ 表示有人购买更便宜的商品的概率。令REPLACE (C) 和REPLACE (D) 表示更换两种产品的概率。如果以下不等式成立：

$$\text{REPLACE (C)} \times 1 - \text{BUY (C)} > \text{REPLACE (D)} \times \text{BUY (C)}$$

那么会有更多的人买更便宜的商品，但是同时更多的人拥有耐用品。根据假设，命题的第一部分显然成立。为了证明第二部分，我们要先解出转移概率。某个人从D移动到C的概率是P (D, C)

=REPLACE (D) ×BUY (C)。与此类似，某个人从C移动到D的概率是P (C, D) =REPLACE (C) ×1-BUY (C)。在均衡状态下，D-D×P (D, C) +CP (C, D) =D。设定C= (1-D)，我们得到

$$D = \frac{P(C,D)}{P(C,D) + P(D,C)}。如果P (C, D) > P (D, C)，那么上式大于$$

0.5。这个不等式与REPLACE (C) ×1-BUY (C) >REPLACE (D)

×BUY (C) 是等价的。

5. McPhee, 1963; Ehrenberg, 1969.他们给出了一些证明双重危险的经验证据。为了证明这个结果，我们只需要证明消费者在转换品牌时会以同样的概率购买任何一种产品。

6. Briggs&Sculpher, 1998.

7. Schrod, 1998.

8. Khmelev&Tweedie, 2001.

9. Khmelev&Tweedie, 2001.

10. Reynolds&Saxonhouse, 1995.

11. 请参见网络。

12. 巧妙构建此类模型的关键在于，定义有用的状态并分配准确的转移概率。请参见：Langville&Meyer, 2012。

13. 在食物链中，一个物种与它所消费的物种相连。请参见：Allesina&Pascual, 2009。

14. Russakoff, 2015.

18 系统动力学模型

1. Sterman, 2000.

2. 关于定性系统动力学模型的价值分析，请参见：Wellman, 1990。

3. 这个模型假设所有野兔都是因被狐狸吃掉而死亡的。为野兔的其他死亡方式增加变量，只会使模型复杂化，但不会改变其结果，因为这样做其实只降低了现在这个模型中野兔增长率的取值范围。这里的符号 $\dot{H}$ 表示每单位时间内 H 的变化率，换一种写法是 $\frac{\partial H}{\partial t}$ 。当野兔

和狐狸数量的变化率都等于零时，即 $\dot{H} = \dot{F} = 0$ 时，均衡就实现了。为了求解均衡，取方程 $gH - aFH = 0$ ，然后两边同时除以 H ，可以得到 $g - aF = 0$ 。解出 F 就可以得到结果。接下来，再取方程 $bFH - dF = 0$ ，然后两边同时除以 F ，可以得到 $bH - d = 0$ 。求解 H 就可以得到结果。

4. 我要感谢多伦多大学的迈克尔·赖亚尔（Michael Ryall），他提供了这个例子。
5. 对于这个模型的相关结果的总结，请参见：Meadows et al., 1972。
6. 这套模型就是通常人们所称的“罗马俱乐部”模型。戴维·洛克菲勒（David Rockefeller）于1968年创立的罗马俱乐部，资助了围绕这些模型的研究，并推广了这些模型的发现。
7. 通过对变量在更小范围内进行操纵，约翰·米勒还可以将模型的预测推向近300亿。请参见：Miller, 1998。
8. Hecht, 2008; MacKenzie, 2012.
9. Sterman, 2006.
10. Glantz, 2008.

19 基于阈值的模型

1. Granovetter, 1978.
2. 爱彼迎创始人采取的方法是，在2008年美国总统大选期间，开发出了奥巴马奥氏（ObamaO's）和船长麦凯恩氏（Cap'n McCain's）谷物早餐，然后挨家挨户上门推销。
3. 关于旋转门模型，请参见：Jacobs, 1989。实证研究表明，在几乎不需要接受传统“正规”教育的职业中（例如调酒行业和园艺行业），当

某个工作场中女性员工的比例占到了15%时，男性就会选择离开（或选择不进入），请参见：Pan, 2015。

4. Syverson, 2007.

5. Gammill&Marsh, 1988.

6. Easley et al., 2012.

20 空间竞争模型与享受竞争模型

1. Clark、Golder&Golder, 2008.

2. 关于对法律立场的分析，请参见：Martin&Quinn, 2002。

3. 霍特林（1929）研究了地理区位决策。凯尔文·兰开斯特（Kelvin Lancaster）（1966）扩展了霍特林的模型，用来研究经济选择之间的享受竞争；唐斯（1957）则将这个模型应用于政治学研究。

4. 提名分数模型提供了一种更复杂的方法来分配意识形态立场，但是所依据的基本思想是一样的，请参见：Poole&Rosenthal, 1985。

5. 选择常数项（表达式中的 C ）时，要让所有收益均为正。为了实现这一点，可以将它设定为任何理想点和可选点之间的最大可能距离。

6. 在以这种方式构建分割线时，我们假设消费者赋予这两个属性的权重是相同的，也可以将权重不同的情况包括进来。如果人们认为含糖量比可可含量更重要，那么我们会以逆时针的角度倾斜这条分割线。而在极端情况下，如果（例如）人们只重视含糖量一个属性，那么分割线将是水平的，并且会在垂直轴上均匀地划分A和B。空间模型提供了一个清晰的例子，说明我们应该怎样将直觉——在这里直觉是我们更喜欢且更接近我们的理想的东西，转换为正式的模型。一旦将备选方案（巧克力棒）和消费者的理想点在空间上画出来，并根据它们与理想点之间的距离来定义偏好排序（将备选方案

从最优到最差排列好），我们事实上就已经根据这些备选方案给出了一个效用函数了。一种产品的效用等于它与理想点之间的距离的倒数。

7. Havel, 1978.

8. Martin&Quinn, 2002.

9. McCarty, 2011.这些百分比可能会随着对党派忠诚度的变化或者提交投票的法案类型的变化而改变。

10. 从形式上说，如果我们假设选民的人数为奇数并且在二维中位数上存在一个选民理想点，那么这个条件就要求通过二维中位数的任何一条线都会将剩下的选民理想点划分为两个大小相等的集合。请参见：Plott, 1967。

11. 理查德·麦凯尔维（Richard McKelvey, 1979）证明，连续进行一系列投票，会导致具有两个维度或多个维度的任何一个政策都有可能被选中，从而导致“混乱结果”的发生。不过，麦凯尔维谨慎地声明，他这个结论只针对给定偏好下可能产生的结果序列，而不是对系列投票会发生什么事情的预测。也请参见：Kollman、Miller&Page, 1997；他们给出了多维空间模型的一个计算版本，在那个版本中，候选人在各种行为假设下会移动到中心位置。

12. 提议者可能不得不提议41。

13. McCarty&Meirowitz, 2014.他们对这个模型以及应用于政治学领域的其他博弈论模型进行了更加深入的分析。

14. 关于拥有否决权的投票者的影响的进一步的分析，请参见：Tsebelis, 2002。

15. 有一项关于洛杉矶房价的研究估计，通勤成本大约为每小时28美元；请参见：Bajari&Kahn, 2008。

16. 这里的计算过程如下。初始收入等于价格乘以数量，我们可以将它写成 $p \times q$ 。在稀疏的市场中，价格下跌后，收入下降10%，销售额增加8%，因此收入等于 $0.9p \times 1.08q = 0.972p \times q$ 。而对于拥挤的市场，价格下跌后的收入等于 $0.9p \times 1.33q = 1.197p \times q$ 。

21 博弈论模型

1. 正式模型以及相关证明如下。令 E_i 表示博弈参与者 i 的努力水平。如果胜出，那么博弈参与者 i 获得的收益等于 $M - E_i$ ，否则为 $-E_i$ 。假设博弈参与者胜出的概率等于他在总努力水平中所占的比例，即：

$$P(i \text{ wins}) = \frac{E_i}{\sum_{j=1}^N E_j}$$

为了求解这个模型唯一的（对称）纳什均衡，我们考虑这个问题：假设所有其他博弈参与者都选择相同的努力水平 E^* ，博弈参与者 i 的努力水平是什么。博弈参与者 i 的收益等于：

$$\frac{E_i \times M}{E_i + (N-1)E^*} - E_i$$

这个函数的一阶导数等于：

$$\frac{(N-1)E^* \times M}{E_i^2 + 2(N-1)E^*E_i + (N-1)^2(E^*)^2} - 1$$

为了求出最大值，我们令一阶导数等于零，于是可以得出 $(N-1)E^* \times M = E_i^2 + 2(N-1)E^*E_i + (N-1)^2(E^*)^2$ 。在对称均衡中， $E_i = E^*$ 。代入上式就可以得出结果。而为了证明一阶导数确实给出了最大值，我们只需检验收益函数的二阶导数是不是为负即可。

2. 欲了解如何梳理网络效应、以及为什么使用快照数据很难得出关于网

络效应的结论，请参见：Shalizi&Thomas, 2011。关于聚类行为和属性的实例，请参见：Christakis&Fowler, 2009。

22 合作模型

1. 请参见2005年出版的《科学》杂志创刊125周年纪念特刊。
2. Martin et al., 2008; Biernaskie, 2011。
3. Zaretsky, 1998.在银行的例子中，如果银行能够赚取大量的地理租金，ATM会降低它们利润吗？地理租金指，银行可以利用客户无法承担长途跋涉到另外一个城镇的银行办理业务的成本而获得的额外利润。我要感谢西蒙·威尔基（Simon Wilkie）以及其他一些朋友，他们与我一起讨论了这个例子。
4. 从技术上讲，我们这样说其实是在证明冷酷触发策略是概率性重复囚徒困境博弈中的均衡策略。其他策略，例如针锋相对策略，在与冷酷触发策略配对时，也可以成为均衡策略。
5. 当另一方博弈参与者采取冷酷触发策略时，选择背叛的博弈参与者在第一个时期内可以获得收益 T ，但是在未来的任何一次博弈中都只能获得最多为零的收益。如果博弈参与者在每个时间段使用冷酷触发策略，那么在博弈进行的第一期都能获得 R 的收益。博弈进行两期的概率为 P ，进行三期的概率为 P^2 ，进行 N 期的概率为 P^{N-1} 。因此，期望收益等于：

$$\sum_{t=1}^{\infty} P^{t-1} \times R = \frac{R}{1-P}$$

对 $(1+P+P^2+P^3+\dots) = \frac{1}{1-P}$ 的证明如下。假设结果为真，然后对等式的两边同时乘以 $(1-P)$ ，于是右侧等于1，左侧等于

$(1+P+P^2+P^3+\dots) - (P+P^2+P^3+\dots)$ ，它也等于1。

6. 我们只需要把保持合作时的那个计算过程重做一遍，并使 P 下降到 $\hat{P}$ 即可。如果 $\hat{P}$ 不支持合作，那么，继续以概率为 P 的博弈进行几期后，再以 $\hat{P}$ 的继续概率进行博弈，也不能支持合作。
7. 诺瓦克和卡尔·西格蒙德（Karl Sigmund）（1998）称之为图像评分。另外也请参见：Bshary&Grutter, 2006。
8. 如果假设它们能够增大鸣鸟的领地，那么这两种形式的攻击都可以映射为背叛。请参见：Godard, 1993。不过，将某些行动映射为具有收益的博弈形式，需要广泛的实地观察。
9. 其中有四对值得重点阐述。当针锋相对策略与冷酷触发策略配对时，两者会永远合作下去，而且每个博弈参与者的平均收益均为3。当欺负好人策略与欺负好人策略配对时，前两轮双方都会背叛，然后永远合作，其平均收益略小于2。当冷酷触发策略与欺负好人策略配对时，冷酷触发策略在第一轮合作、欺负好人策略在第一轮就背叛；在第二轮，双方都背叛；在第三轮和第四轮中，冷酷触发策略背叛和欺负好人策略合作；此后，两者都一直背叛。冷酷触发策略的收益序列可以写成1, 2, 4, 4, 然后一直是2，因此平均收益为 2^+ 。欺负好人策略的收益序列可以写成4, 3, 1, 1, 然后一直是2，平均收益正好为2。当针锋相对策略与欺负好人策略配对时，针锋相对策略在第一轮合作，而欺负好人策略在第一轮就背叛。第二轮双方都背叛。在第三轮中，欺负好人策略合作，但是针锋相对策略继续背叛。在第四轮中，欺负好人策略第二次合作，针锋相对策略恢复合作。此后，双方都永远合作。双方的收益均为一个1、一个4、一个2，以及一直为3，因此双方的平均收益均为略小于3。
10. 事实上，始终合作策略和冷酷触发策略都是被针锋相对策略占优的。无论针对什么策略，针锋相对策略至少与始终合作策略或冷酷

触发策略一样好（或更好）。针锋相对策略之所以占优永远合作策略，是因为针锋相对策略不会被始终背叛策略和欺负好人策略所“剥削”。针锋相对策略之所以占优冷酷触发策略，是因为它能够与欺负好人策略合作，因为针锋相对策略能够原谅欺负好人策略的背叛，但是冷酷触发策略却不能。

在一个后来成为经典的著名实验中，罗伯特·阿克塞尔罗德要求各个领域的学者参加一项“锦标赛”——提交用于有一定概率继续进行下去的重复囚徒困境博弈中的不同策略。在学者们提交的14个策略中，针锋相对策略表现最佳。然后他又组织了一场，要求大家提交新的策略。这一次有62个人提交了各种策略。针锋相对策略再一次胜出。阿克塞尔罗德将针锋相对策略的成功归结为该策略的如下性质：它会合作，又会惩罚，同时它也是宽容的。冷酷触发策略不宽容，因此冷酷触发策略无法重启与欺负好人策略的合作。请参见：Axelrod, 1984。表22-1没有给出针对所有策略的平均收益，因为那将意味着假设每个策略具有相同的可能性。一个种群可能大多数人都采取针锋相对策略；另一个种群可能包括了很大比例的采取欺负好人策略的人；而第三个种群则可能有很多人采取始终合作策略和始终背叛策略。

11. 假设，诱惑收益等于奖励收益的4倍，即， $T=4S$ ，而且种群中5%的人合作。这样一来， P 必定超过 $\frac{20T-R}{20T}$ 。要维持合作，只需要

P 超过 $\frac{(T-R)}{T}$ 即可。在 $T=4$ 且 $R=3$ 的情况下，合作的出现要求

$$P \geq \frac{77}{80}, \text{ 而维持合作则要求 } P \geq \frac{1}{4}.$$

接下来给出一般性的证明。假设在种群中，有 θ 比例的人采取针锋相对策略（或，冷酷触发策略），比例 $(1-\theta)$ 的人采取始终背叛策略。再假设每个人都要与整个种群中的所有其他人博弈。与针锋相对策略配对的针锋相对策略（或冷酷触发策略）在每一期都能获得 R 的收益，因而其期

望收益为 $\frac{R}{(1-P)}$ 。与始终背叛策略配对的针锋相对策略只能获得 $-S$ 的回报。而始终背叛策略与始终背叛策略相配对时，双方都只能获得零收益。而且，始终背叛策略与针锋相对策略配对时，始终背叛策略能够获得 T 的收益。因此，针锋相对策略的平均收益等于 $\theta \times \frac{R}{(1-P)} - (1-\theta) \times S$ ，同时始终背叛策略的平均收益则为 $\theta \times T$ 。这样一来，当且仅当 $\frac{R}{1-P} - S \times \frac{(1-\theta)}{\theta} \geq T$ 时，针锋相对策略胜过始终背叛策略。因此，当且仅当下式成立时，针锋相对策略获得更高的收益：

$$P \geq \frac{T - R + S \times \frac{(1-\theta)}{\theta}}{T + S \times \frac{(1-\theta)}{\theta}}$$

如果 θ 很小，那么 $\frac{(1-\theta)}{\theta}$ 将变得非常大，上述条件很难成立。关于合作的出现比合作的维持更加困难的进一步分析，请参见：Boyd, 2006。

12. 这个模型，以及对于它的分析，是我们从诺瓦克的论文中借用的（Nowak, 2006）。诺瓦克证明，重复博弈、声誉和亲属选择也能支持合作。
13. 边缘上的节点会复制表现最好的邻居的行动。根据假设，所有相邻的背叛者获得的收益都等于零。而合作的邻居则可以获得 $K \times B - D \times C$ 的收益——当且仅当 $\frac{B}{C} \geq \frac{D}{K}$ 时，这一收益大于零。
14. 关于群体选择理论的基本概念，请参见：Wilson, 1975。
15. 对于特劳森和诺瓦克的模型，在这里要简单地介绍一下。将一个由 N 个人组成的种群划分为 M 个不同大小的群体。在每个群体中，都应

用合作行动模型并为每一个人分配一个表现值。令选中个体*i*的概率等于*i*的表现除以所有*N*个人的表现之和的结果。然后，将这个人的一个克隆体添加到同一个群体中。

如果加入这个克隆体后，该群体的大小超过了阈值 $\bar{S}$ ，那么就以概率 $(1-q)$ 随机地从这个群体中的某个个体移除，同时以概率 q 将该群体分成两个群体，每个成员随机地进入其中一个群体中。为了保持群体的总数不变，再随机选择一个现有群体将之移除。当*M*很大且群体分裂很少发生（ q 很小）时，当且仅当 $\frac{B}{C} \geq 1 + \frac{\bar{S}}{M}$ 成立，合作者的数量就会增加。关于这个模型的更加完整的说明，请参见：Nowak, 2006。

16. 当敏捷管理的倡导者米歇尔·佩鲁索（Michelle Peluso）成了IBM的首席营销官时，她组建了一些相互竞争团队，每一个团队的表现都是其他团队可观察的。然后对表现最好的团队加以奖励（Dan, 2018）。这种类型的敏捷管理实践借鉴了敏捷编程的思想。敏捷编程采用同时推进代码编写、程序测试和与用户互动的方法，取代了标准的按顺序依次推进的瀑布式方法。
17. 还有一种策略是慷慨的针锋相对，即，一开始选择合作，且只在部分时间惩罚背叛行为。在会出现误差的博弈中，这种策略优于针锋相对策略和“赢则坚持，输则改之”策略。请参见：Rand et al., 2009; Wu&Axelrod, 1995。

18. Axelrod&Pienta, 2006.

23 与集体行动有关的问题

1. Hardin, 1968.
2. Diamond, 2005.

3. Ostrom, 2005; Ostrom、Janssen&Anderies, 2007.

4. 为了得出社会最优解，假设每个人在公共物品上花费 X 。整个种群的总效用等于：

$$N [2\sqrt{NX} + \text{收入} - X]$$

对 X 求导，并令得到的导数等于零，可以得到：

$$N \left(\frac{\sqrt{N}}{\sqrt{X}} - 1 \right) = 0$$

得到 $X=N$ 。

为了求解对称纳什均衡，我们假设其他每个人都为公共物品贡献相同的数额，并将这个数额记为 A 。令 Y 表示这个人贡献的数额。于是这个人的效用等于：

$$2\sqrt{(N-1)A+Y} + \text{收入} - Y$$

对 Y 求导，并令得到的导数等于零，我们可以得到 $\frac{1}{\sqrt{(N-1)A+Y}} = 1$ 。重新排列各项并对两边同时取平方，得到 $Y + (N-1)A = 1$ 。在对称均衡中，每个人贡献相同的数额（ $Y=A$ ），因此有 $Y = \frac{1}{N}$ 。

5. 哲学家会认为这是功利主义。约翰·罗尔斯（John Rawls, 1971）提出了另一种选择：最大最小原则（maxmin principle）。根据这个原则，理想的社会结果就能最大化最不富裕的人的效用。罗尔斯争辩说，我们应该站在“无知之幕”背后去考虑配置问题。“无知之幕”指我们对自己社会中的地位毫不知情。有了“无知之幕”，我们不知道我们是贫是富、有无名声，也不知道自己拥有多大的能力、会不会受到环境的阻碍。罗尔斯要求对最不富裕的人赋予最大的权重，而

功利主义则平等地给每个人赋予权重。

6. 用*i*来给这*N*个人编号。我们可以把第*j*个人的效用写成如下形式：

$$\text{Utility}_i(\text{PUBLIC} + \text{PRIVATE}) = \sqrt{\text{PUBLIC}} + \text{PRIVATE}_i$$

为了解出对称纳什均衡，我们假设每一个其他人都为公共物品贡献了数量*A*。

再令*Y*表示个体*j*所贡献的数额。用*I*表示所有人的共同的收入水平。这样一来，个体*j*的效用等于：

$$(1 - \alpha) \left(2\sqrt{(N-1)A + Y} + I - Y \right) + \alpha N \left(2\sqrt{(N-1)A + Y} \right) + \alpha (N-1)(I - A) + (I - Y)$$

对*Y*求导，并令得到的导数等于零，我们可以得到：

$$\frac{1 - \alpha}{\sqrt{(N-1)A + Y}} + \frac{\alpha N}{\sqrt{(N-1)A + Y}} = 1$$

重新排列各项，我们有： $(1 - \alpha) + \alpha N = \sqrt{Y + (N-1)A}$ 。在对称平衡中，*Y*=*A*。从而，遵循 $(1 - \alpha) + \alpha N = \sqrt{NY}$ 。等号两边求平方，可以得到 $[(1 - \alpha) + \alpha N]^2 = NY$ ，这就意味着，

$$Y = \frac{[(1 - \alpha) + \alpha N]^2}{N}。$$

7. Cornes & Sandler, 1996.

8. 更现实的模型还要假设非线性拥塞成本（也许是*S*形的）。这种假设能够刻画像道路这样的资源的如下特点：前几位其他用户对个人的利益不会有任何影响，但是到了某个时点上，资源会因过于拥挤而归于无用。

9. 使用资源的*M*个人的总效用等于 $(B - \theta \times M)$ 。对*M*求导，并令得到的导数等于零，我们可以得到 $(B - 2M\theta) = 0$ ，求解可得 $M = \frac{B}{2\theta}$ 。为

了求解纳什均衡，我们将从不使用资源的收益设定为零值。于是，人们会使用资源，直到收益等于外部可选项，即： $M = \frac{B}{\theta}$ 。

10. 需要注意的是：我们将最大收益 B 设定为等于种群规模 N ，目的是减少变量的数量。为了求解社会最优结果和纳什均衡，我们首先注意到，总效用等于 $(N-M) \times M + 3N - (N-M) \times (N-M)$ ，从而得到 $4(N-M)M$ 。对 M 求导，得 $4N - 8M = 0$ 。求解可得 $M = N/2$ 。于是总效用等于 $4(\frac{N}{2})^2 = N^2$ 。为了求解均衡，我们要找到一个 M ，使两个公园的边际效用相等。这种情况在 $(N-M) = 3N - 3(N-M)$ 时发生。这个式子可以重写为 $N = 4M$ 。将 M 和 $N-M$ 的值代入效用函数，就可以计算出总效用。

11. 为了求解均衡消费，只要设定 $R^* = (1-g)(R^* - C^*)$ 并解出 R^* 即可。

12. Kurlansky, 1998.

13. 为了说明为什么变化没有抵消，可以考虑一个两期模型。第一年的增长率为20%，因此只有96个单位的资源（ $80 \times 1.2 = 96$ ）。第二年的增长率为30%，这导致了 $98.8 = (96 - 20) \times 1.3$ 个单位的资源。如果我们将这两年的增长率互换一下，那么在第一年之后，就会有104个单位的资源，而且在第二年之后会存在 $100.8 = (104 - 20) \times 1.2$ 个单位的资源。

14. Ostrom、Janssen&Anderies, 2007.

15. Craine&Dybzinski, 2013.

16. Ostrom, 2010; Ostrom, 2004.

24 与机制设计有关的模型

1. Ledyard、Porter&Rangel, 1997.

2. 以帕累托有效为例。考虑3个人的以下4种收益情况：

$$\{ (3, 3, 4), (9, 0, 0), (0, 8, 1), (2, 2, 3) \}$$

除了(2, 2, 2)之外，都是帕累托有效的。但是配置(2, 2, 3)被(3, 3, 4)占优。

3. Hurwicz&Schmeidler, 1978.

4. 第三个竞买人可能出价略高于60美元，为了简化分析，假设正好为60美元。

5. 我们在这里给出的证明假设了这些值是均匀分布的，但是这个结果在其他类型的分布下仍然成立。假设其他 $(N-1)$ 个竞买人的出价都是真实价值乘以 $\frac{N-1}{N}$ 。如果其他竞买人的价值乘以 $\frac{N-1}{N}$ 小于 b ，那么出价 b 就将高于另一个竞买人的出价。发生这种情况的概率等于 $b \frac{N}{N-1}$ 。由此可知，大于所有其他 $(N-1)$ 个出价的概率等于这个值的 $(N-1)$ 次幂。因此，如果竞买人的真实价值为 V ，则出价 b 的预期价值等于该价值减去出价的差 $(V-b)$ 乘以 b 为最高出价的概率。这样一来，就可以将预期价值写为：

$$\text{预期价值} = (V-b) \left(b \frac{N}{N-1}\right)^{N-1}$$

为了使上式的值最大化，我们对 b 求导，并令求得的导数等于零，于是可以得到如下条件：

$$V(N-1) \left(\frac{N}{N-1}\right)^{N-1} b^{N-2} - N \left(\frac{N}{N-1}\right)^{N-1} b^{N-1} = 0$$

简化后可以得到这个表达式： $V(N-1) - Nb=0$ ，重写为

$b = \frac{V(N-1)}{N}$ 。为了证明出价最高的竞买人付出的价格等于出价第二高的竞买人的出价的预期价值，我们要先注意到，如果从区间 $[0, 1]$ 上的均匀分布中抽取 N 个随机的值，那么最高出价的预期价值等于 $1 \times \frac{N}{N+1}$ 。第二高的出价的预期价值则等于 $1 \times \frac{N-1}{N+1}$ 。因此，出价最高的竞买人的预期价值等于：

$$\left(\frac{N-1}{N}\right) \times 1 \times \left(\frac{N}{N+1}\right) = 1 \times \left(\frac{N-1}{N+1}\right)$$

这就等于第二高的竞买人的预期价值。

6. 罗杰·迈尔森（Roger Myerson）是我的博士生导师，他因这个贡献获得了诺贝尔经济学奖。
7. 在全支付拍卖中，当竞买人的价值位于区间 $[0, 1]$ 上时，其最优策略可以表示为：拥有价值 V 的竞买人出价 $V^N \frac{(N-1)}{N}$ 。因此，如果有三个竞买人，那么价值为 $\frac{1}{2}$ 的竞买人将出价 $\frac{1}{8} \times \frac{2}{3} = \frac{1}{12}$ 。
8. 相关的实验证据，请参见Lucking-Reiley, 1999。相关的eBay拍卖实验证据，请参见Morgan&Hossain, 2006。相关的木材拍卖证据，请参见Athey、Levin&Seira, 2011。
9. Ostrovsky、Edelman&Schwarz, 2007.
10. Page, 2012.

25 信号模型

1. 有关信号如何驱动行为和选择的更多例子，请参见：Simler&Hanson , 2018。

2. 一个弱者类型的个体发送信号成本等于 MC 。假设所有 S 个强者类型的个体都发出了信号，那么收益等于 $\frac{B}{S+1}$ 。因此，如果

$MC \geq \frac{B}{S+1}$ ，那么弱者类型个体就不会发送信号。如果他们源于分离的收益 $\frac{B}{S} - cM$ ，超过了所有 N 个个体分担信号时的 $\frac{B}{N}$ ，强者类型更偏好发送信号。从先前的计算可知，使弱者类型的个体不发送信号的最小信号等于 $M = \frac{B}{(S+1)C}$ 。因此，如果我们

$\hat{M} = \frac{B}{SC} M$ ，那么弱者类型将不会发出信号。而要让强者类型更偏好发送信号 $\hat{M}$ ，我们必须保证 $\frac{B}{S} - c \frac{B}{SC} \geq \frac{B}{N}$ 。上式两边同时除以 B 并乘以 C ，就可以得到 $\frac{C}{S} - \frac{c}{S} \geq \frac{C}{N}$ 。这个式子可以简化为 $(C-c)N \geq CS$ ，可以改写为 $C(N-S) \geq cN$ 。

3. 经济学家迈克尔·斯宾塞（Michael Spence）提出了这种可能性。斯宾塞因提出了一个假设教育没有功能性价值的就业市场模型（Spence, 1973）获得了诺贝尔经济学奖。
4. 尾羽过大就会影响孔雀的适合度。请参见：Zahavi, 1975。
5. Bird&Smith, 2005.
6. Smith&Bird, 2003.

26 学习模型

1. 关于学习的心理学研究，涵盖的背景要比我们在这里所讨论的更加广泛。一个人可以通过学习，了解一个事实，例如阿肯色州的首府在哪里。一个人也可以通过学习，获得某种隐性知识，例如怎样烤面

包、修理发动机或编写计算机程序。一个人还可以通过学习，掌握某种知识体系，例如有机化学。

2. Thorndike, 1911.

3. Rescorla&Wagner, 1972.

4. 我描述的这个模型是以下述论著中模型为基础的：Rescorla&Wagner, 1972; Herrnstein, 1970; Bush&Mosteller, 1955; Cyert&March, 1963; Bendor、Diermeie&Ting, 2003; Epstein, 2014。

5. 对参数 γ 的选择，必须保证备选方案的权重为正。如果 γ 超过了最高可能渴望水平与任何备选方案的最小奖励之间差异的倒数，就可以做到这一点。

6. Bendor&Swistak, 1997.

7. 如果我们构建出一个有限种群的复制者动态模型，随后的种群是随机选择的，那么就可能无法重新复制出最优备选方案。如果真的是这样，那么复制者动态模型将找不到最优选择，因为它无法将备选方案重新引入到种群中。

8. Fudenberg&Levine, 1998; Camerer, 2003.

9. 这个博弈还有一个混合策略均衡，其中2/3选择经济型汽车，1/3选择一个油老虎车。但是，根据我们的学习规则，这种均衡是不稳定的，因此忽略它。

10. 更正式的表示是， $P(\text{经济型汽车}, 1) = 0.5$ ， $P(\text{油老虎车}, 1) = 0.5$ ，收益(经济型汽车, 1) = 1.5，收益(油老虎车, 1) = 2，平均收益 = 1.75。应用复制者方程，我们可以得出 $P(\text{经济型汽车}, 2) = 0.5 \times \frac{1.5}{1.75} = 0.43$ ，以及 $P(\text{油老虎车}, 2) = 0.5 \times \frac{2}{1.75} = 0.57$ 。

11. Frank, 1985.

[12.](#) Waltz, 1979; Powell, 1991.

[13.](#) Vriend, 2000.这篇论文分析了类似的收益结构，并将它解释为生产相同产品并同时选择数量的企业之间的竞争。经济学家将这种模型称为古诺竞争模型（Cournot competition model）。

[14.](#) 罗斯-伊雷夫学习模型是通过以下公式更新第 t 期中的备选方案 k 的权重 $W(k, t)$ 的： $W(k, t+1) = (1-r) \times W(k, t) + \Delta(k, t, e)$ 。参数 r 是一个新近度参数（recency parameter）。如果行动 k 被选中，那么 $\Delta(k, t, e) = (1-e) \times \text{收益}(k, t)$ ；如果行动 k 没有被选中，那么 $\Delta(k, t, e) = e \times \text{收益}(k, t)$ 。参数 e （实验参数）决定了未被选中的备选方案的权重。

[15.](#) Camerer&Ho, 1999.

[16.](#) 这个分析紧密追随了贝德纳和佩奇（2007，2018）的思路，并借鉴了格雷夫（2006）的模型。贝德纳和佩奇强调了初始行动对于均衡的涌现的重要性，而格雷夫则更侧重于信念的作用。也请参见：Gilboa&Schmiedler, 1995.他们提出了基于案例的决策理论。另外，关于身份在经济选择中的作用，请参见：Akerlof&Kranton, 2010.

[17.](#) 为了证明复制者动态的结果，请注意当且仅当以下情况成立时，按文化规范采取行动才会有更高的收益：

$$(1-B) \times 200 + B \times 220 > (1-B) \times 180 + 300B$$

重新排列各项，得 $20(1-B) > 80B$ 。结果得证。

27 多臂老虎机问题

[1.](#) 关于这类问题与经济现象的相关性，请参见：Bergemann&Valimaki, 2008。

[2.](#) Hills et al., 2015.

3. 欲了解多臂老虎机问题的更多技术细节，以及对各种启发式的详细比较，请参见：Scott, 2010。
4. 吉廷斯和琼斯（1972）率先描述了这些最优规则。可以将吉廷斯指数重新表述为一个贝尔曼方程（Bellman equation）。贝尔曼方程适用于任何通过一系列选择完成、且每个选择都会产生收益的问题。贝尔曼方程依赖于价值函数的构造。该价值函数的值等于收益序列的总和（其中，未来的收益要根据当前利率进行贴现）。
5. Roberts, 2004.
6. 请参见杰克·鲍尔斯（Jake Bowers）等人（2017）对美国农业部农场服务机构小额贷款项目的实验分析。
7. 数据源于2012年的《华盛顿邮报》，以及Dann, 2016。
8. 如果他们是在位领导人，那么阿尔·戈尔（Al Gore）、乔治·H. W. 布什（George H. W. Bush）和希拉里·克林顿将会因经济繁荣而获得更大声誉。关于较早期的政治选情与经济景气的关系的分析，请参见：Markus, 1988。关于更晚近的证据，请参见：Fair, 2012。关于候选人与政党效应大小的相关证据，请参见：Campbell、Dettrey&Yin, 2010。

28 崎岖景观模型

1. 关于多样性的价值的更加深入的阐述，请参见《多样性红利》一书。
2. 对NK模型更加全面的介绍，请参见：Kauffman, 1993。
3. 当 $N=20$ ， $K=19$ 时，我们可以推导出局部峰值和全局峰值的期望值。每个属性的贡献均匀分布在区间 $[0, 1]$ 上，其均值为 $\frac{1}{2}$ 、方差为 $\frac{1}{12}$ 。这个结果只要通过观察均匀分布就可以证明。每个备选方案

的价值等于20个属性的贡献的平均值。因此，根据中心极限定理，这些值服从均值为 $\frac{1}{2}$ 、方差为 $\frac{1}{12N}$ 的正态分布。从而，在 $N=20$ 的情况下，每一个标准偏差的大小为 $0.0645 = \sqrt{\frac{1}{240}}$ 。然后我们可以估计出局部峰值的平均值：0.609。我们可以认为，从这个分布中随机抽取21次当中最好的那一次就是一个局部峰值。因此，它的期望值等于近似地从正态分布中抽取出来的、使 $\frac{21}{22}$ 的备选方案都有较低的值的一个值。它将会比平均值略低两个标准偏差。使用正态分布计算器，可以算出期望平均值等于0.609。为了估计出全局峰值的期望值0.759，我们注意到，全局峰值具有所有 2^{20} 个备选方案的最高值。每个备选方案可以被视为是从该分布中抽出的一个值，因此其期望值等于近似地从正态分布中抽取出来的、使 $\frac{2^{20}}{2^{20}+1}$ 的备选方案都有较低的值的一个值。利用正态分布计算器，可以算出该期望平均值等于0.759。全局峰值的值，应该比100万次抽取的最高值还更高。

4. Wright, 2001.研究认为这些整合重组有助于人类以及人类社会的出现，也有利于这个时代的技术和科学进步。
5. Kaufman, 1993; Miller&Page, 2007.他们把复杂性描述为“有趣的中间状态”。
6. 在美国，如果申请专利后获批的时间超过了3年，那么专利保护期为从专利发布之日起计算17年。
7. Boldrin&Levine, 2010.

结语 像芒格一样智慧地思考——多模型思维的实际应用

1. 根据正文所述的概率，确切的统计均衡是：无痛状态为70.7%，阿片状态为19.5%，上瘾状态为9.8%。而对于第一种情况，相应的百分比则分别为76.3%、21.5%和2.2%。
2. Wakeland、Nielsen&Geissert, 2015.
3. 我要感谢艾比·雅各布斯（Abbie Jacobs），她帮助我得到了这个结果。她有所需要的模型和数据。
4. Wilkinson&Pickett, 2009.
5. 这是杜尔劳夫在芝加哥大学贝克尔·弗里德曼研究所（Becker Friedman Institute）主办的一次会议上的评论。会议的主题是“了解不平等及其应对措施”，于2015年11月6日举行。
6. Goldin&Katz, 2008; Acemoglu&Autor, 2011; Murphy&Topel, 2016.
7. Mas-Colell、Whinston&Green, 1995.
8. Kaplan&Rauh, 2013a; Jones&Kim, 2018.他们的模型将能力作为企业家创意的可扩展性的代理变量。另请参见：Frank, 1996；该文是关于不同职业中的不平等是如何产生的早期研究。另请参见：Xie、Killewald&Near, 2016。
9. Ormerod, 2012.
10. Ormerod, 2012.该文详细描述了日益增强的连通性如何导致了不平等。
11. Cancian&Reed, 1999; Schwartz&Mare, 2005.
12. Greenwood et al., 2014.
13. 估计的结果是0.34而不是0.43。请参见：Greenwood et al., 2014。基尼系数可以衡量收入分布与平均分布之间的距离。令 $S(P)$ 表示收入分布中最底部 $P\%$ 的人口所占的总份额（例如，如果位于收入最

底部30%的人口获得了总收入的2%，那么 $S(30)=2$ ，计算出来的基尼系数为：

$$\text{基尼系数} = \frac{2}{99} \times \sum_{P=1}^{100} \left[\frac{P}{100} - S(P) \right]$$

如果收入是均匀分布的，即 $S(P) = \frac{P}{100}$ ，那么基尼系数=0。如果所有的收入都归于最顶部的1%，那么对于所有 $P < 100$ ，都有 $S(P) = 0$ ，且 $S(100) = 1$ ，因此基尼系数=1。

14. 计算可以按如下过程进行。儿童属于四个类别的概率为（0.6, 0.25, 0.1, 0.05）。这也就是说，高收入者的下一代中60%的人有高收入，中高收入者的下一代中20%的人有高收入，中低收入者的下一代中15%的人有高收入，而低收入者的下一代中只有5%的人有高收入。高收入者的孙子女中有高收入的人的百分比则等于 $0.6 \times 0.6 + 0.25 \times 0.2 + 0.1 \times 0.15 + 0.05 \times 0.05 = 0.4275$ ，低收入者的孙子女中有高收入的人的百分比则等于 $0.6 \times 0.05 + 0.25 \times 0.1 + 0.1 \times 0.15 + 0.05 \times 0.7 = 0.105$ 。

15. Pfeffer&Killewald, 2017.

16. Kaplan&Rao, 2013b.

参考文献

- Acemoglu, Daron, and David Autor. 2011. “Skills, Tasks and Technologies: Implications for Employment and Earnings.” In Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, 4: 1043–1171. Amsterdam: Elsevier-North Holland.
- Acemoglu, Daron, and James Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Cambridge, MA: Harvard University Press.
- Adler, Mortimer Jerome. 1970. *The Time of Our Lives: The Ethics of Common Sense*. New York: Holt, Rinehart and Winston.
- Akerlof, G., and R. Kranton. 2010. *Identity Economics*. Princeton, NJ: Princeton University Press.
- Albert, Rika, Istvan Albert, and Gary L. Nakarado. 2004. “Structural Vulnerability of the North American Power Grid.” *Physical Review E* 69: 025103.
- Allesina, Stefano, and Mercedes Pascual. 2009. “Googling Food Webs: Can an Eigenvector Measure Species' Importance for Coextinctions?” *PLOS: Computational Biology* 9, no. 4.
- Allison, Graham. 1971. *Essence of Decision: Explaining the Cuban Missile Crisis*. New York: Little, Brown.
- Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty, and Emmanuel Saez. 2013. “The World Top Incomes Database.” <https://www.inet.ox.ac.uk/projects/view/149>.

- Anderson, Chris. 2008a. "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete." *Wired* 16, no. 7.
- Anderson, Chris. 2008b. *The Long Tail: Why the Future of Business Is Selling Less of More*. New York: Hachette.
- Anderson, Phillip. 1972. "More Is Different." *Science* 177, no. 4047: 393–396.
- Arrow, Kenneth. 1963. *Social Choice and Individual Values*. New Haven, CT: Yale University Press.
- Arthur, W. B. 1994. "Inductive Reasoning and Bounded Rationality (The El Farol Problem)." *American Economic Review Papers and Proceedings* 84: 406–411.
- Arthur, W. B. 2011. *The Nature of Technology: What It Is and How It Evolves*. New York: Free Press.
- Ashenfelter, Orley. 2010. "Predicting the Quality and Prices of Bordeaux Wine." *Journal of Wine Economics* 5, no. 1: 40–52.
- Athey, Susan, Jonathan Levin, and Enrique Seira. 2011. "Comparing Open and Sealed Bid Auctions: Evidence from Timber Auctions." *Quarterly Journal of Economics* 126, no. 1: 207–257.
- Austin, David. 2008. "Percolation: Slipping Through the Cracks." American Mathematical Society. www.ams.org/publicoutreach/feature-column/fcarc-percolation.
- Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.
- Axelrod, David, Robert Axelrod, and Kenneth J. Pienta. 2006. "Evolution of Cooperation Among Tumor Cells." *Proceedings of the National*

- Academy of Sciences* 103, no. 36: 13474–13479.
- Axtell, Robert L. 2001. “Zipf Distribution of U. S. Firm Sizes.” *Science* 293: 1818–1820.
- Bajari, Patrick, and Matthew E. Kahn. 2008. “Estimating Hedonic Models of Consumer Demand with an Application to Urban Sprawl.” In *Hedonic Methods in Housing Markets*, 129–155. New York: Springer.
- Bak, Per. 1996. *How Nature Works: The Science of Self-Organized Criticality*. New York: Springer.
- Baldwin, Carliss Y., and Kim B. Clark. 2000. *Design Rules. Vol. 1, The Power of Modularity*. Cambridge, MA: MIT Press.
- Ball, Eric, and Joseph LiPuma. 2012. *Unlocking the Ivory Tower: How Management Research Can Transform Your Business*. Palo Alto, CA: Kauffman Fellow Press.
- Banzhaf, John F. 1965. “Weighted Voting Doesn't Work: A Mathematical Analysis.” *Rutgers Law Review* 19, no. 2: 317–343.
- Barber, Gerald M. 1997. “Sequencing Highway Network Improvements: A Case Study of South Sulawesi.” *Economic Geography* 53, no. 1: 55–69.
- Bass, Frank. 1969. “A New Product Growth Model for Consumer Durables.” *Management Science* 15, no. 5: 215–227.
- Baxter, G. William. 2009. “The Dynamics of Foraging Ants.” Paper presented at the annual meeting of the American Physical Society, March 16–20, abstract H40.00011.
- Bednar, Jenna. 2007. “Credit Assignment and Federal Encroachment.” *Supreme Court Economic Review* 15: 285–308.
- Bednar, Jenna. 2008. *The Robust Federation: Principle of Design*.

Cambridge: Cambridge University Press.

- Bednar, Jenna, Aaron Bramson, Andrea Jones-Rooy, and Scott E. Page. 2010. "Emergent Cultural Signatures and Persistent Diversity: A Model of Conformity and Consistency." *Rationality and Society* 22, no. 4: 407–444.
- Bednar, Jenna, and Scott E. Page. 2007. "Can Game(s) Theory Explain Culture? The Emergence of Cultural Behavior Within Multiple Games." *Rationality and Society* 19, no. 1: 65–97.
- Bednar, Jenna, and Scott E. Page. 2018. "When Order Affects Performance: Culture, Behavioral Spillovers and Institutional Path Dependence." *American Political Science Review* 112, no. 1: 82–98.
- Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen. 2018 "Who Becomes an Inventor in America? The Importance of Exposure to Innovation: Executive Summary." www.equality-of-opportunity.org.
- Bendor, Jonathan, Daniel Diermeier, and Michael Ting. 2003. "A Behavioral Model of Turnout." *American Political Science Review* 97, no. 2: 261–280.
- Bendor, Jonathan, and Piotr Swistak. 1997. "The Evolutionary Stability of Cooperation." *American Political Science Review* 91: 290–307.
- Bendor, Jonathan, and Scott E. Page. 2018. "A Model of Team Problem Solving." Unpublished manuscript.
- Berg, Nathan, and Gerd Gigerenzer. 2010. "As-If Behavioral Economics: Neoclassical Economics in Disguise?" *History of Economic Ideas* 18, no. 1: 133–166.

- Bergemann, Dirk, and Juuso Välimäki. 2008. "Bandit Problems." In *The New Palgrave Dictionary of Economics*, 2nd ed., ed. Steven N. Durlauf and Lawrence E. Blume. London: Palgrave Macmillan.
- Berlekamp, Elwyn R., John H. Conway, and Richard K. Guy. 1982. "What Is Life?" In *Winning Ways for Your Mathematical Plays*. Vol. 2, *Games in Particular*. London: Academic Press.
- Bertrand, Marianne, and Sendhil Mullainathan. 2001. "Are CEOs Rewarded for Luck? The Ones Without Principles Are." *Quarterly Journal of Economics* 116: 901–932.
- Bickel, P. J., E. A. Hammel, and J. W. O'Connell. 1974. "Sex Bias in Graduate Admissions: Data from Berkeley." *Science* 187 (4175): 398–404.
- Biernaskie, Jay, M. 2011. "Evidence for Competition and Cooperation Among Climbing Plants." *Proceedings of the Royal Society B* 278: 1989–1996.
- Bird, Rebecca, and Eric Smith. 2004. "Signaling Theory, Strategic Interaction, and Symbolic Capital." *Current Anthropology* 46, no. 2: 222–248.
- Boldrin, Michele, and David Levine. 2010. *Against Intellectual Monopoly*. Cambridge: Cambridge University Press.
- Borges, Jorge Luis. 1974. *A Universal History of Infamy*. Trans. Norman Thomas de Giovanni. London: Penguin.
- Bowers, Jake, Nathaniel Higgins, Dean Karlan, Sarah Tulman, and Jonathan Zinman. 2017. "Challenges to Replication and Iteration in Field Experiments: Evidence from Two Direct Mail Shots." *American Economic Review Papers & Proceedings* 107, no. 5: 1–3.

- Bowles, Samuel, and Herbert Gintis. 2002. "The Inheritance of Inequality." *Journal of Economic Perspectives* 16, no. 3: 3–30.
- Box, George E. P., and Norman Draper. 1987. *Empirical Model-Building and Response Surfaces*. New York: Wiley.
- Boyd, Robert. 2006. "Reciprocity: You Have to Think Different." *Journal of Evolutionary Biology* 19: 1380–1382.
- Breiman, Leo. 1996. "Bagging Predictors." *Machine Learning* 24, no. 2: 123–140.
- Briggs, Andrew, and Mark Sculpher. 1998. "An Introduction to Markov Modeling for Economic Evaluation." *Pharmaco Economics* 13, no. 4: 397–409.
- Brock, William, and Steven Durlauf. 2001. "Discrete Choice with Social Interactions." *Review of Economic Studies* 68: 235–260.
- Broido, A. D., and A. Clauset. 2018. "Scale-Free Networks Are Rare." Working paper.
- Bshary, R., and A. S. Grutter. 2006. "Image Scoring and Cooperation in a Cleaner Fish Mutualism." *Nature* 441, no. 7096: 975–978.
- Burt, Ronald. 1995. *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.
- Bush, Robert, and Frederick Mosteller. 1954. *Stochastic Models for Learning*. New York: John Wiley and Sons.
- Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.
- Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. 1997. "Labor Supply of New York City Cabdrivers: One Day at a

- Time.” *Quarterly Journal of Economics* 112, no. 2: 407–441.
- Camerer, Colin, and Tek Ho. 1999. “Experience-Weighted Attraction Learning in Normal Form Games.” *Econometrica* 67, no. 4: 827–874.
- Camerer, Colin, George Loewenstein, and Drazen Prelec. 2005. “Neuroeconomics: How Neuroscience Can Inform Economics.” *Journal of Economic Literature* 43: 9–64.
- Campbell, Donald T. 1976. “Assessing the Impact of Planned Social Change.” Public Affairs Center, Dartmouth College.
- Campbell, James E., Bryan J. Dettrey, and Hongxing Yin. 2010. “The Theory of Conditional Retrospective Voting: Does the Presidential Record Matter Less in Open-Seat Elections?” *Journal of Politics* 72, no. 4: 1083–1095.
- Cancian, Maria, and Deborah Reed. 1999. “The Impact of Wives' Earnings on Income Inequality: Issues and Estimates.” *Demography* 36, no. 2: 173–184.
- Carvalho, Vasco, and Xavier Gabaix. 2013 “The Great Diversification and Its Undoing,” *American Economic Review* 103, no. 5: 1697–1727.
- Castellano, Claudio, Santo Fortunato, and Vittorio Loreto. 2009. “Statistical Physics of Social Dynamics.” *Review of Modern Physics* 81: 591–646.
- Cederman, Lars Erik. 2003. “Modeling the Size of Wars: From Billiard Balls to Sandpiles.” *American Political Science Review* 97: 135–150.
- Centola, Damon, and Michael Macy. 2007. “Complex Contagions and the Weakness of Long Ties.” *American Journal of Sociology* 113: 702–734.
- Chance, Donald. 2009. “What Are the Odds? Another Look at DiMaggio's Streak.” *Chance* 22, no. 2: 33–42.

- Christakis, N. A., and J. Fowler. 2009. *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives*. New York: Little, Brown.
- Churchland, Patricia, and Terry J. Sejnowski. 1992. *The Computational Brain*. Cambridge, MA: MIT Press.
- Chwe, Michael. 2013. *Jane Austen: Game Theorist*. Princeton, NJ: Princeton University Press.
- Clarida, Richard, Jordi Galí, and Mark Gertler. 2000. “Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory.” *Quarterly Journal of Economics* 115, no. 1: 147–180.
- Clark, Gregory. 2014. *The Son Also Rises: Surnames and the History of Social Mobility*. Princeton, NJ: Princeton University Press.
- Clark, William, Matt Golder, and Sona Nadenicheck Golder. 2008. *Principles of Comparative Politics*. Washington, DC: Congressional Quarterly Press.
- Clauset, Aaron, M. Young, and K. S. Gleditsch. 2007. “On the Frequency of Severe Terrorist Attacks.” *Journal of Conflict Resolution* 51, no. 1: 58–88.
- Cohen, Tyler. 2013. *Average Is Over: Powering America Beyond the Age of the Great Stagnation*. New York: Dutton.
- Cooke, Nancy J., and Margaret L. Hilton, eds. 2014. *Enhancing the Effectiveness of Team Science*. Washington, DC: National Academies Press.
- Cornes, Richard, and Todd Sandler. 1996. *The Theory of Externalities, Public Goods, and Club Goods*. 2nd ed. Cambridge: Cambridge University

Press.

Craine, Joseph, and Ray Dybzinski. 2013. "Mechanisms of Plant Competition for Nutrients, Water and Light." *Functional Ecology* 27: 833–840.

Cyert, Richard M., and James G. March. 1963. *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice-Hall.

Dan, Avi. 2018. "How Michelle Peluso Is Redefining Marketing at IBM." *Forbes*, January 18.

Dann, Carrie. 2016. "Pro-Clinton Battleground Ad Spending Outstrips Trump Team by 2." NBC News, November 4.

Dawes, Robyn. 1979. "The Robust Beauty of Improper Linear Models in Decision Making." *American Psychologist* 34: 571–582.

de Marchi, Scott. 2005. *Computational and Mathematical Modeling in the Social Sciences*. Cambridge: Cambridge University Press.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal. 2009. "Optimal Versus Naive Diversification: How Inefficient Is the $\frac{1}{N}$ Portfolio Strategy?" *Review of Financial Studies* 22, no. 5: 1915–1953.

Dennett, Daniel C. 1991. *Consciousness Explained*. Boston: Back Bay Books.

Dennett, Daniel C. 1994. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon & Schuster.

Denrell, Jerker, and Chengwei Liu. 2012. "Top Performers Are Not the Most Impressive When Extreme Performance Indicates Unreliability." *Proceedings of the National Academy of Sciences* 109, no. 24: 9331–9336.

- Diamond, Jared. 2005. *Collapse: How Societies Choose to Fail or Succeed*. New York: Viking Penguin.
- Dodds, Peter, Robby Muhamad, and Duncan Watts. 2003. "An Experimental Study of Search in Global Social Networks." *Science* 301: 827–829.
- Downing, John A., et al. 2006. "The Global Abundance and Size Distribution of Lakes, Ponds, and Impoundments." *Limnology and Oceanography* 51, no. 5: 2388–2397.
- Dragulescu, Adrian, and Victor M. Yakovenko. 2001. "Exponential and Power-Law Probability Distributions of Wealth and Income in the United Kingdom and the United States." *Physica A* 299: 213–221.
- Drucker, Peter. 1969. *The Age of Discontinuity: Guidelines to Our Changing Society*. New York: Harper and Row.
- Dubos, Jean. 1987. *The White Plague: Tuberculosis, Man and Society*. New Brunswick, NJ: Rutgers University Press.
- Dunne, Anthony. 1999. *Hertzian Tales: Electronic Products, Aesthetic Experience and Critical Design*. London: Royal College of Art.
- Dyson, Freeman. 2004. "A Meeting with Enrico Fermi." *Nature* 427: 297.
- Easley, David, and Jon Kleinberg. 2010. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge: Cambridge University Press.
- Easley, David, Marcos Lopez de Prado, and Maureen O'Hara. 2012. "Flow Toxicity and Liquidity in a High Frequency World." *Review of Financial Studies* 24, no. 5: 1457–1493.
- Easterly, William, and Stanley Fischer. 1995. "The Soviet Economic Decline." *World Bank Economic Review* 9, no. 3: 341–371.

- Ebbinghaus, Herman. 1885. *Memory: A Contribution to Experimental Psychology*. Online in *Classics in the History of Psychology*. <http://psychclassics.yorku.ca/Ebbinghaus/index.htm>.
- Ehrenberg, Andrew. 1969. "Towards an Integrated Theory of Consumer Behaviour." *Journal of the Market Research Society* 11, no. 4: 305–337.
- Einstein, Albert. 1934. "On the Method of Theoretical Physics." *Philosophy of Science* 1, no. 2: 163–169.
- Eliot, Matt, Ben Golub, and Matthew Jackson. 2014. "Financial Networks and Contagion." *American Economic Review* 104, no. 10: 3115–3153.
- Eom, Young-Ho, and Hang-Hyun Jo. 2014. "Generalized Friendship Paradox in Complex Networks: The Case of Scientific Collaboration." *Scientific Reports* 4: 4603.
- Epstein, Josh. 2006. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton, NJ: Princeton University Press.
- Epstein, Joshua. 2008. "Why Model?" *Journal of Artificial Societies and Social Simulation* 11, no. 4: 12.
- Epstein, Joshua. 2014. *Agent Zero: Toward Neurocognitive Foundations for Generative Social Science*. Princeton, NJ: Princeton University Press.
- Ericsson, K. A. 1996. "The Acquisition of Expert Performance: An Introduction to Some of the Issues." In *The Road to Excellence: The Acquisition of Expert Performance in the Arts and Sciences, Sports, and Games*, ed. K. A. Ericsson, 1–50. Mahwah, NJ: Erlbaum.
- Fair, Raymond. 2012. *Predicting Presidential Elections and Other Things*. 2nd ed. Stanford, CA: Stanford University Press.
- Farmer, J. Doyne 2018. "Collective Awareness: A Conversation with J.

Doyne Farmer.” *The Edge*. https://www.edge.org/conversation/j_doyne_farmer-collective-awareness.

Feld, Scott L. 1991. “Why Your Friends Have More Friends than You Do.” *American Journal of Sociology* 96, no. 6: 1464–1477.

Flegal, Katherine M., Brian K. Kit, Heather Orpana, and Barry I. Graubard. 2012. “Association of All-Cause Mortality with Overweight and Obesity Using Standard Body Mass Index Categories: A Systematic Review and Meta-analysis.” *Journal of the American Medical Association* 309, no. 1: 71–82.

Flores, Thomas, and Irfan Nooruddin. 2016. *Elections in Hard Times: Building Stronger Democracies in the 21st Century*. Cambridge: Cambridge University Press.

Florida, Richard. 2005. *Cities and the Creative Class*. New York: Routledge.

Foster, Dean, and H. Peyton Young. 2001. “On the Impossibility of Predicting the Behavior of Rational Agents.” *Proceedings of the National Academy of Sciences* 98, no. 22: 12848–12853.

Frank, Kenneth, et al. 2018. “Teacher Networks and Educational Opportunity.” In *Handbook on the Sociology of Education*, ed. Barbara Schneider and Guan Saw. New York: Oxford University Press.

Frank, Robert. 1984. *Choosing the Right Pond*. Oxford: Oxford University Press.

Frank, Robert. 1996. *The Winner-Take-All Society: Why the Few at the Top Get So Much More than the Rest of Us*. New York: Penguin.

Freeman, Richard, and Wei Huang. 2015. “Collaborating with People Like Me: Ethnic Coauthorship Within the U. S.” *Journal of Labor Economics*

33 no. S1: S289-S318.

Fudenberg, Drew, and David Levine. 1998. *Theory of Learning in Games*. Cambridge, MA: MIT Press.

Fudenberg, Drew, and David Levine. 2006. "A Dual-Self Model of Impulse Control." *American Economic Review* 96: 1449–1476.

Gammill, James F. , Jr., and Terry A. Marsh. 1988. "Trading Activity and Price Behavior in the Stock and Stock Index Futures Markets in October 1987." *Journal of Economic Perspectives* 2, no. 3: 25–44.

Gawande, Atul. 2009. *The Checklist Manifesto: How to Get Things Right*. New York: Henry Holt.

Geithner, Timothy. 2014. *Stress Test: Reflections on Financial Crises*. New York: Crown.

Gerschenkron, Alexander. 1952. "Economic Backwardness in Historical Perspective." In *The Progress of Underdeveloped Areas*, ed. B. F. Hoselitz. Chicago: University of Chicago Press.

Gertner, Jon. 2012. *The Idea Factory: Bell Labs and the Great Age of American Innovation*. New York: Penguin.

Gibrat, Robert. 1931. *Les inégalités économique*. Paris: Sirely.

Gigerenzer, Gerd, and Reinhard Selten. 2002. *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press.

Gigerenzer, Gerd, and Peter Todd. 2000. *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.

Gilboa, Itzhak, and David Schmeidler. 1994. "Case-Based Decision Theory." *Quarterly Journal of Economics* 110: 605–639.

- Gilovich, Thomas, Amos Tversky, and R. Vallone. 1984. "The Hot Hand in Basketball: On the Misperception of Random Sequences." *Cognitive Psychology* 17, no. 3: 295–314.
- Glaeser, Edward, Bruce Sacerdote, and Jose Scheinkman. 1996. "Crime and Social Interactions." *Quarterly Journal of Economics* 111, no. 2: 507–548.
- Glantz, Andrew. 2008. "A Tax on Light and Air: Impact of the Window Duty on Tax Administration and Architecture, 1696–1851." *Penn History Review* 15, no. 2: 18–40.
- Glasserman, Paul, and H. Peyton Young. 2014. "Contagion in Financial Networks." Office of Financial Research Working Paper.
- Godard, Renee. 1993. "Tit for Tat Among Neighboring Hooded Warblers." *Behavioral Ecology and Sociobiology* 33, no. 1: 45–50.
- Gode, Dhananjay K., and Shyam Sunder. 1993. "Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality." *Journal of Political Economy* 101, no. 1: 119–137.
- Goldin, Claudia, and Lawrence F. Katz. 2008. *The Race Between Education and Technology*. Cambridge, MA: Harvard University Press.
- Gordon, Robert J. 2016. *The Rise and Fall of American Growth: The U. S. Standard of Living Since the Civil War*. Princeton, NJ: Princeton University Press.
- Granovetter, Mark. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78, no. 6: 1360–1380.
- Granovetter, Mark. 1978. "Threshold Models of Collective Behavior."

- American Journal of Sociology* 83, no. 6: 1360–1443.
- Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos. 2014. “Marry Your Like: Assortative Mating and Income Inequality.” *American Economic Review: Papers & Proceedings* 104, no 5: 348-353.
- Greif, Avner. 2006. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge: Cambridge University Press.
- Griliches, Zvi. 1957, 1988. “Hybrid Corn: An Exploration of the Economics of Technological Change.” In *Technology, Education and Productivity: Early Papers with Notes to Subsequent Literature*. New York: Basil Blackwell.
- Groseclose, Tim, and James Snyder. 1996. “Buying Supermajorities.” *American Political Science Review* 90: 303–315.
- Grossman, S., and J. Stiglitz. 1980. “On the Impossibility of Informationally Efficient Markets.” *American Economic Review* 70, no. 3: 393–408.
- Groysberg, Boris. 2012. *Chasing Stars: The Myth of Talent and the Portability of Performance*. Princeton, NJ: Princeton University Press.
- Guy, Richard. 1983. “Don't Try to Solve These Problems.” *American Mathematical Monthly* 90: 35–41.
- Haidt, Jonathan. 2006. *The Happiness Hypothesis: Finding Modern Truth in Ancient Wisdom*. Basic Books. New York: NY.
- Haldene, Andrew. 2012. “The Dog and the Frisbee.” Speech given at the Federal Reserve Bank of Kansas City's 36th Economic Policy Symposium, Jackson Hole, WY.
- Haldene, Andrew. 2014. “The Dappled World.” Speech given at the University of Michigan Law School, Ann Arbor, October 23.

- Haldane, John B. S. 1928. "On Being the Right Size." Online version available at <http://irl.cs.ucla.edu/papers/right-size.html>.
- Hardin, Garret. 1968. "The Tragedy of the Commons." *Science* 162, no. 3859: 1243–1248.
- Harrell, Frank E. 2001. *Regression Modeling Strategies with Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York: Springer.
- Harstad, Ronald M., and Reinhard Selten. 2013. "Bounded Rationality Models: Tasks to Become Intellectually Competitive." *Journal of Economic Literature* 51, no. 2: 496–511.
- Harte, John. 1988. *Consider a Spherical Cow*. Mill Valley, CA: University Science Books.
- Hathaway, Oona. 2001. "Path Dependence in the Law: The Course and Pattern of Change in a Common Law Legal System." *Iowa Law Review* 86.
- Havel, Václav. 1985. *The Power of the Powerless: Citizens Against the State in Central-Eastern Europe*. Ed. John Keane. Armonk, NY: M. E. Sharpe.
- Hawking, Stephen, and Leonard Mlodinow. 2011. *The Grand Design*. New York: Bantam.
- Hecht, Jeff. 2008. "Prophecy of Economic Collapse 'Coming True.'" *New Scientist*. November 17.
- Herrnstein, Richard J. 1970. "On the Law of Effect." *Journal of the Experimental Analysis of Behavior* 13: 243–266.
- Hills, Thomas, Peter M. Todd, David Lazer, A. David Redish, Iain D. Couzin, and the Cognitive Search Research Group. 2015. "Exploration

Versus Exploitation in Space, Mind, and Society.” *Trends in Cognitive Science* 19, no. 1: 46–54.

Hofstadter, Douglas, and Emmanuel Sander. 2013. *Surfaces and Essences: Analogy as the Fuel and Fire of Thinking*. New York: Basic Books.

Holland, John. 1975. *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press.

Hong, Lu, and Scott E. Page. 2009. “Interpreted and Generated Signals.” *Journal of Economic Theory* 144: 2174–2196.

Hotelling, Harold. 1929. “Stability in Competition.” *Economic Journal* 39, no. 153: 41–57.

Huffaker, Carl Burton. 1958. “Experimental Studies on Predation: Dispersion Factors and Predator-Prey Oscillations.” *Hilgardia* 27, no. 14: 343–383.

Hurwicz, Leo, and David Schmeidler. 1978. “Outcome Functions Which Guarantee the Existence and Pareto Optimality of Nash Equilibria.” *Econometrica* 46: 144–174.

Inman, Mason. 2011. “Sending Out an SOS.” *Nature Climate Change* 1: 180–183.

International Monetary Fund. 2009. *Global Financial Stability Report*.

Jackson, Matthew. 2008. *Social and Economic Networks*. Princeton, NJ: Princeton University Press.

Jackson, Matthew and Asher Wolinsky. 1996. “A Strategic Model of Social and Economic Networks.” *Journal of Economic Theory* 71: 44–74.

Jacob, Francois. 1977. “Evolution and Tinkering.” *Science* 196: 1161–1166.

Jacobs, Jane. 1989. *Revolving Doors: Sex Segregation and Women's Careers*.

Stanford, CA: Stanford University Press.

Johnson, James. 2014. "Models Among the Political Theorists." *American Journal of Political Science* 58, no. 33: 547–560.

Johnson-Laird, Philip. 2009. *How We Reason*. New York: Oxford University Press.

Jones, Benjamin F., Brian Uzzi, and Stefan Wuchty. 2008. "Multi-University Research Teams: Shifting Impact, Geography and Social Stratification in Science." *Science* 322: 1259–1262.

Jones, Charles, and Jihee Kim. 2018 "A Schumpeterian Model of Top Income Inequality." *Journal of Political Economy*. Forthcoming.

Kahneman, Daniel. 2011. *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decisions Under Risk." *Econometrica* 47, no. 2: 263–291.

Kalyvas, Stathis. 1999. "The Decay and Breakdown of Communist One - Party Systems." *Annual Review of Political Science* 2: 323–343.

Kamin, Leon J. 1969. "Predictability, Surprise, Attention and Conditioning." In *Punishment and Aversive Behavior*, ed. B. A. Campbell and R. M. Church, 279–296. New York: Appleton-Century-Crofts.

Kaplan, Steven, and Joshua D. Rauh. 2013a. "Family, Education, and Sources of Wealth Among the Richest Americans, 1982–2012." *American Economic Review Papers and Proceedings* 103, no. 3: 158–162.

Kaplan, Steven, and Joshua D. Rauh. 2013b. "It's the Market: The Broad - Based Rise in the Return to Top Talent." *Journal of Economic*

Perspectives 27, no. 3: 35–56.

Karlsson, Bengt. 2016. “The Forest of Our Lives: In and Out of Political Ecology.” *Conservation and Society* 14, no. 4: 380–390.

Kauffman, Stuart. 1993. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford: Oxford University Press.

Kennedy, John F. 1956. *Profiles in Courage*. New York: Harper & Brothers.

Khmelev, Dmitri, and F. J. Tweedie. 2001. “Using Markov Chains for Identification of Writers.” *Literary and Linguistic Computing* 16, no. 4: 299–307.

Kleinberg, Jon, and M. Raghu. 2015. “Team Performance with Test Scores.” Working paper, Cornell University School of Information.

Knox, Grahame. n.d. “Lost at Sea.” *Insight*, http://insight.typepad.co.uk/lost_at_sea.pdf.

Kollman, Ken, J. Miller, and S. Page. 1992. “Adaptive Parties in Spatial Elections.” *American Political Science Review* 86: 929–937.

Kooti, Farshad, Nathan O. Hodas, and Kristina Lerman. 2014. “Network Weirdness: Exploring the Origins of Network Paradoxes.” Paper presented at the International Conference on Weblogs and Social Media (ICWSM), March.

Kurlansky, Mark. 1998. *Cod: A Biography of the Fish That Changed the World*. New York: Penguin.

Kydland, Finn E., and Edward C. Prescott. 1977. “Rules Rather than Discretion: The Inconsistency of Optimal Plans.” *Journal of Political Economy* 85, no. 3: 473–491.

Lai, T. L., and Herbert Robbins. 1985. “Asymptotically Efficient Adaptive

- Allocation Rules.” *Advances in Applied Mathematics* 6, no. 1: 4–22.
- Laibson, David. 1997. “Golden Eggs and Hyperbolic Discounting.” *Quarterly Journal of Economics* 112, no. 2: 443–477.
- Lamberson, P. J., and Scott E. Page. 2012a. “The Effect of Feedback Variability on Success in Markets with Positive Feedbacks.” *Economics Letters* 114: 259–261.
- Lamberson, P. J., and Scott E. Page. 2012b. “Tipping Points.” *Quarterly Journal of Political Science* 7, no. 2: 175–208.
- Lancaster, Kelvin J. 1966. “A New Approach to Consumer Theory.” *Journal of Political Economy* 74: 132–157.
- Landemore, Helene. 2013. *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton, NJ: Princeton University Press.
- Lango, Allen H., et al. 2010. “Hundreds of Variants Clustered in Genomic Loci and Biological Pathways Affect Human Height.” *Nature* 467, no. 7317: 832–838.
- Langville, Amy N., and Carl D. Meyer. 2012. *Who's #1?: The Science of Rating and Ranking*. Princeton, NJ: Princeton University Press.
- Lave, Charles, and James G. March. 1975. *An Introduction to Models in the Social Sciences*. Lanham, MD: University Press of America.
- Ledyard, John, David Porter, and Antonio Rangle. 1997. “Experiments Testing Multiobject Allocation Mechanisms.” *Journal of Economics and Management Strategy* 6, no. 3: 639–675.
- Ledyard, John, David Porter, and Randii Wessen. 2000. “A Market-Based Mechanism for Allocating Space Shuttle Secondary Payload Priority.”

Experimental Economics 2, no. 3: 173–195.

Levins, Richard. 1966. “The Strategy of Model Building in Population Biology.” *American Scientist* 54: 421–431.

Levinthal, Daniel A. 1997. “Adaptation on Rugged Landscapes.” *Management Science* 43: 934–950.

Levinthal, Daniel. 1991. “RandomWalks and Organizational Mortality.” *Administrative Science Quarterly* 36, no. 3: 397–420.

Levitt, Steven, and Stephen Dubner. 2009. *SuperFreakonomics: Global Cooling, Patriotic Prostitutes, and Why Suicide Bombers Should Buy Life Insurance*. New York: William Morrow.

Lewis, Michael. 2014. *Flash Boys: A Wall Street Revolt*. New York: W. W. Norton.

Limpert, Eckhard, Werner A. Stahel, and Markus Abbt. 2001. “Log-normal Distributions Across the Sciences: Keys and Clues.” *BioScience* 51, no. 5: 341–352.

Little, Daniel. 1998. *Microfoundations, Method, and Causation: On the Philosophy of the Social Sciences*. Piscataway, NJ: Transaction Publishers.

Lo, Andrew W., and A. Craig MacKinlay. 2007. *A Non-Random Walk Down Wall Street*. Princeton, NJ: Princeton University Press.

Lo, Andrew W. 2012. “Reading About the Financial Crisis: A Twenty-One-Book Review.” *Journal of Economic Literature* 50, no. 1: 151–178.

Lucas, Robert. 1976. “Econometric Policy Evaluation: A Critique.” In *The Phillips Curve and Labor Markets*, ed. K. Brunner and A. Meltzer, 19–46. Carnegie-Rochester Conference Series on Public Policy 1. New

York: Elsevier.

- Lucking-Reiley, David. 1999. "Using Field Experiments to Test Equivalence Between Auction Formats: Magic on the Internet." *American Economic Review* 89, no. 5: 1063–1080.
- MacKenzie, Debora. 2012. "Boom and Doom: Revisiting Prophecies of Collapse." *New Scientist*, January.
- Mannes, Albert E., Jack B. Soll, and Richard P. Larrick. 2014. "The Wisdom of Select Crowds." *Journal of Personality and Social Psychology* 107: 276–299.
- Markowitz, Harold M. 1952. "Portfolio Selection." *Journal of Finance* 7, no. 1: 77–91.
- Markus, Greg B. 1988. "The Impact of Personal and National Economic Conditions on the Presidential Vote: A Pooled Cross-Sectional Analysis." *American Journal of Political Science* 32: 137–154.
- Martin, Andrew D., and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999." *Political Analysis* 10: 134–153.
- Martin, Francis, et al. 2008. "The Genome of *Laccaria bicolor* Provides Insights into Mycorrhizal Symbiosis." *Nature* 452: 88–92.
- Martinez Peria, Maria Soledad, Giovanni Majnoni, Matthew T. Jones, and Winfrid Blaschke. 2001. "Stress Testing of Financial Systems: An Overview of Issues, Methodologies, and FSAP Experiences." IMF Working Paper no. 01/88.
- Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green. 1994. *Microeconomic Theory*. New York: Oxford University Press.

- Mauboussin, Michael. 2012. *The Success Equation: Untangling Skill and Luck in Business*. Cambridge, MA: Harvard University Press.
- May, Robert M., Simon A. Levin, and George Sugihara. 2008. "Ecology for Bankers." *Nature* 451: 893–895.
- McCarty, Nolan. 2011. "Measuring Legislative Preferences." In *Oxford Handbook of Congress*, ed. Eric Schickler and Frances Lee. New York: Oxford University Press.
- McCarty, Nolan, and Adam Meirowitz. 2014. *Political Game Theory: An Introduction*. Cambridge: Cambridge University Press.
- McKelvey, Richard. 1979. "General Conditions for Global Intransitivities in Formal Voting Models." *Econometrica* 47: 1085–1112.
- McPhee, William N. 1963. *Formal Theories of Mass Behaviour*. New York: Free Press of Glencoe.
- Meadows, D., G. Meadows, J. Randers, and W. W. Behrens III. 1972. *The Limits to Growth*. New York: Universe Books.
- Medin, Douglas, Will Bennis, and Michel Chandler. 2010. "The Home-Field Disadvantage." *Perspectives on Psychological Science* 5, no. 6: 708–713.
- Merriam, Daniel F., and John C. Davis. 2009. "Using Zipf's Law to Predict Future Earthquakes in Kansas." *Transactions of the Kansas Academy of Science* 112, nos. 1&2: 127–129.
- Merton, Robert C. 1969. "Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case." *Review of Economics and Statistics* 51, no. 3: 247–257.
- Merton, Robert K. 1963. "Resistance to the Systematic Study of Multiple Discoveries in Science." *European Journal of Sociology* 4, no. 2: 237–

- Milgrom, Paul, and John Roberts. 1986. "Pricing and Advertising Signals of Product Quality." *Journal of Political Economy* 94, no. 4: 796–821.
- Miller, John H. 1998. "Active Nonlinear Tests (ANTs) of Complex Simulation Models." *Management Science* 44, no. 6: 820–830.
- Miller, John H. 2015. *A Crude Look at the Whole*. New York: Basic Books.
- Miller, John H., and Scott E. Page. 2004. "The Standing Ovation Problem." *Complexity* 9, no. 5: 8–16.
- Miller, John H., and Scott E. Page. 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton, NJ: Princeton University Press.
- Miller, Joshua B., and Adam Sanjurjo. 2015. "Surprised by the Gambler's and Hot Hand Fallacies: A Truth in the Law of Small Numbers." IGIER Working Paper no. 552.
- Mitchell, Melanie. 1996. *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press.
- Mitchell, Melanie. 2009. *Complexity: A Guided Tour*. Oxford: Oxford University Press.
- Mlodinow, Leonard. 2009. *The Drunkard's Walk: How Randomness Rules Our Lives*. New York: Penguin.
- Mokyr, Joel. 2002. *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Princeton, NJ: Princeton University Press.
- Morgan, John, and Tanjim Hossain. 2006. ". . . Plus Shipping and Handling: Revenue (Non)Equivalence in Field Experiments on eBay." *Advances in Economic Analysis & Policy* 6, no. 2: 3.

Moss-Racusin, Corinne, John F. Dovidio, Victoria L. Brescoll, Mark J. Graham, and Jo Handelsman. 2012. "Science Faculty's Subtle Gender Biases Favor Male Students." *Proceedings of the National Academy of Sciences*. 1647–1649.

Munger, Charles. 1994. "A Lesson on Elementary, Worldly Wisdom as It Relates to Investment Management & Business." University of Southern California Business School.

Murphy, Kevin M., and Robert H. Topel. 2016. "Human Capital Investment, Inequality and Growth." *Journal of Labor Economics* 34: 99–127.

Murray, J. D. 1988. "Mammalian Coat Patterns: How the Leopard Gets Its Spots." *Scientific American* 256: 80–87.

Myerson, Roger B. 1999. "On the Value of Game Theory in Social Science." *Rationality and Society* 4: 62–73.

Myerson, Roger B. 1999. "Nash Equilibrium and the History of Economic Theory." *Journal of Economic Literature* 37, no. 3: 1067–1082.

Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study." *American Economic Review* 85, no. 5: 1313–1326.

Newman, Mark E. 2005. "Power Laws, Pareto Distributions and Zipf's Law." *Contemporary Physics* 46: 323–351.

Newman, Mark E. 2010. *Networks: An Introduction*. Oxford: Oxford University Press.

Nowak, Martin. 2006. "Five Rules for the Evolution of Cooperation." *Science* 314, no.5805: 1560–1563.

Nowak, Martin A., and Karl Sigmund. 1998. "Evolution of Indirect

Reciprocity by Image Scoring.” *Nature* 393: 573–577.

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.

O’Neil, Cathy 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, NY: Crown.

Open Science Collaboration. 2015. “Estimating the Reproducibility of Psychological Science.” *Science* 349: 6251.

Organization for Economic Co-operation and Development. 1996. *The Knowledge Based Economy*. Paris: OECD.

Ormerod, Paul. 2012. *Positive Linking: How Networks Can Revolutionise the World*. London: Faber and Faber.

Ostrom, Elinor. 2004. *Understanding Institutional Diversity*. Princeton, NJ: Princeton University Press.

Ostrom, Elinor. 2010. “Beyond Markets and States: Polycentric Governance of Complex Economic Systems.” *Transnational Corporations Review* 2, no. 2: 1–12.

Ostrom, Elinor, Marco A. Janssen, and John M. Anderies. 2007. “Going Beyond Panaceas.” *Proceedings of the National Academy of Sciences* 104: 15176–15178.

Ostrovsky, Michael, Benjamin Edelman, and Michael Schwarz. 2007. “Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords.” *American Economic Review* 97, no. 1: 242–259.

Paarsch, Harry J., and Bruce S. Shearer. 1999. “The Response of Worker Effort to Piece Rates: Evidence from the British Columbia Tree-Planting

- Industry.” *Journal of Human Resources* 34, no. 4: 643–667.
- Packer, Craig, and Anne E. Pusey. 1997. “Divided We Fall: Cooperation Among Lions.” *Scientific American*, May, 52–59.
- Paczuski, Maya, and Kai Nagel. 1996. “Self-Organized Criticality and $1/f$ Noise in Traffic.” arXiv: cond-mat/9602011.
- Page, Scott E. 1997. “An Appending Efficient Algorithm for Allocating Public Projects with Complementarities,” *Journal of Public Economics* 64, no 3: 291–322.
- Page, Scott E. 2001. “Self Organization and Coordination.” *Computational Economics* 18: 25–48.
- Page, Scott E. 2006. “Essay: Path Dependence.” *Quarterly Journal of Political Science* 1: 87–115.
- Page, Scott E. 2007. *The Difference: How the Power of Diversity Creates Better Groups, Teams, Schools, and Societies*. Princeton, NJ: Princeton University Press.
- Page, Scott E. 2010a. *Diversity and Complexity*. Princeton, NJ: Princeton University Press.
- Page, Scott E. 2010b. “Building a Science of Economics for the Real World.” Presentation to the House Committee on Science and Technology Subcommittee on Investigations and Oversight, July 20.
- Page, Scott E. 2012. “A Complexity Perspective on Institutional Design.” *Politics, Philosophy and Economics* 11: 5–25.
- Page, Scott E. 2017. *The Diversity Bonus*. Princeton, NJ: Princeton University Press.
- Pan, Jessica. 2014. “Gender Segregation in Occupations: The Role of Tipping

and Social Interactions.” *Journal of Labor Economics* 33, no. 2: 365–408.

Parrish, Susan Scott. 2017. *The Flood Year 1927: A Cultural History*. Princeton, NJ: Princeton University Press.

Parsa, H. G., John T. Self, David Njite, and Tiffany King. 2005. “Why Restaurants Fail.” *Cornell Hospitality Quarterly* 46, no. 3: 304–322.

Patel, Kayur, Steven Drucker, James Fogarty, Ashish Kapoor, and Desney Tan. 2011. “Using Multiple Models to Understand Data.” *Proceedings of the International Joint Conference on Artificial Intelligence*, 1723–1728.

Peel, L., and A. Clauset. 2014. “Predicting Sports Scoring Dynamics with Restoration and Anti-Persistence.” *Proceedings of the International Conference on Data Mining*. Philadelphia: SIAM.

Pfeffer, Fabian T., and Alexandra Killewald. 2017. “Generations of Advantage: Multigenerational Correlations in Family Wealth.” *Social Forces*, 1–31.

Piantadosi, Steven. 2014. “Zipf’s Word Frequency Law in Natural Language: A Critical Review and Future Directions.” *Psychonomic Bulletin & Review* 21, no. 5: 1112–1130.

Pierson, Paul. 2004. *Politics in Time: History, Institutions, and Social Analysis*. Princeton, NJ: Princeton University Press.

Piketty, Thomas. 2014. *Capital in the 21st Century*. Trans. Arthur Goldhammer. Cambridge, MA: Belknap Press.

Pollack, John. 2014. *Shortcut: How Analogies Reveal Connections, Spark Innovation, and Sell Our Greatest Ideas*. New York: Gotham.

Poole, Keith T., and Howard Rosenthal. 1984. “A Spatial Model for

- Legislative Roll Call Analysis.” *American Journal of Political Science* 29, no. 2: 357–384.
- Porter, David, and Vernon Smith. 2007. “FCC Spectrum Auction Design: A 12-Year Experiment.” *Journal of Law, Economics, and Policy* 3, no. 1: 63–80.
- Powell, Robert. 1991. “Absolute and Relative Gains in International Relations Theory.” *American Political Science Review* 85, no. 4: 1303–1320.
- Przeworski, Adam, Jose Antonio Cheibub, Michael E. Alvarez, and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Material Well-Being in the World, 1950–1990*. Cambridge: Cambridge University Press.
- Raby, Fiona. 2001. *Design Noir: The Secret Life of Electronic Objects*. Basel: Birkhauser.
- Ramo, Joshua Cooper. 2016. *The Seventh Sense: Power, Fortune, and Success, in the Age of Networks*. New York: Little, Brown and Company.
- Rand, David G., Hisashi Ohtsukia, and Martin A. Nowak. 2009. “Direct Reciprocity with Costly Punishment: Generous Tit-for-Tat Prevails.” *Journal of Theoretical Biology* 256, no. 1: 45–57.
- Rapoport, Anatol. 1978. “Reality-Simulation: A Feedback Loop.” *Sociocybernetics*, 123–141.
- Rauch, Jeffrey. 2012. *Hyperbolic Partial Differential Equations and Geometric Optics*. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society.

- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Rescorla, Robert, and Allan Wagner. 1972. "A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement." In *Classical Conditioning II*, ed. A. H. Black and W. F. Prokasy, 64–99. New York: Appleton-Century-Crofts.
- Reynolds, Noel B., and Arlene Saxonhouse. 1994. *Three Discourses*. Chicago: University of Chicago Press.
- Roberts, D. C., and D. L. Turcotte. 1998. "Fractality and Self-Organized Criticality of Wars." *Fractals* 6: 351–357.
- Roberts, Seth. 2004. "Self-Experimentation as a Source of New Ideas: Ten Examples About Sleep, Mood, Health, and Weight." *Behavioral and Brain Sciences* 27, no. 2: 227–262
- Romer, Paul. 1986. "Increasing Returns and Long-Run Growth." *Journal of Political Economy* 94: 1002–1037.
- Rosen, Sherwin. 1981. "The Economics of Superstars." *American Economic Review* 71: 845–858.
- Roth, Alvin, and Ido Erev. 1995. "Learning in Extensive Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term." *Games and Economics Behavior* 8: 164–212.
- Russakoff, Dale. 2015. *The Prize: Who's in Charge of America's Schools?* Boston: Houghton Mifflin Harcourt.
- Rust, Jon. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica* 55, no. 5: 999–1033.
- Ryall, Michael D., and Aaron Bramson *Inference and Intervention: Causal*

Models for Business Analysis. New York: Routledge.

Salganik, Matthew, Peter Dodds, and Duncan J. Watts. 2006. "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market." *Science* 311: 854–856.

Samuelson, Paul. 1964. "Proof That Properly Anticipated Prices Fluctuate Randomly." *Industrial Management Review* 6: 41–49.

Schiller, Robert. 2004. *Irrational Exuberance*. 2nd ed. Princeton, NJ: Princeton University Press.

Schrodt, Philip. 1998. "Pattern Recognition of International Crises Using Hidden Markov Models." In *Non-Linear Models and Methods in Political Science*, ed. Diana Richards. Ann Arbor: University of Michigan Press.

Schwartz, Christine R., and Robert D. Mare. 2004. "Trends in Educational Assortative Marriage from 1940 to 2003." *Demography* 42, no. 4: 621–646.

Schelling, Thomas. 1978. *Micromotives and Macrobehavior*. New York: W. W. Norton.

Scott, Steven L. 2010. "A Modern Bayesian Look at the Multi-Armed Bandit." *Applied Stochastic Models in Business and Industry* 26: 639–658.

Shalizi, Cosma, and Andrew C. Thomas. 2011. "Homophily and Contagion Are Generically Confounded in Observational Social Network Studies." *Sociological Methods and Research* 40: 211–239.

Shapiro, Thomas, Tatjana Meschede, and Sam Osoro. 2013. "The Roots of the Widening Racial Wealth Gap: Explaining the Black-White Economic

Divide.” Research and Policy Brief, Institute on Assets and Social Policy, Brandeis University, Waltham, MA.

Shi, Xiaolin, Lada A. Adamic, Belle L. Tseng, and Gavin S. Clarkson. 2009. “The Impact of Boundary Spanning Scholarly Publications and Patents.” *PLoS ONE* 4, no. 8: e6547.

Silver, Nate. 2012. *The Signal and the Noise: Why So Many Predictions Fail — but Some Don't*. New York: Penguin.

Simler, Kevin, and Robin Hanson. 2018. *The Elephant in the Brain: Hidden Motives in Everyday Life*. Oxford: Oxford University Press.

Simmons, Matthew, Lada Adamic, and Eytan Adar. 2011. “Memes Online: Extracted, Subtracted, Injected, and Recollected.” Paper presented at the International Conference on Web and Social Media.

Slaughter, Ann Marie. 2017. *The Chessboard and the Web: Strategies of Connection in a Networked World*. New Haven, CT: Yale University Press.

Smaldino, Paul. 2013. “Measures of Individual Uncertainty for Ecological Models: Variance and Entropy.” *Ecological Modelling* 254: 50–53.

Small, Dana M., Robert J. Zatorre, Alain Dagher, Alan C. Evans, and Marilyn Jones-Gotman. 2001. “Changes in Brain Activity Related to Eating Chocolate: From Pleasure to Aversion.” *Brain* 124, no. 9: 1720–1733.

Smith, Eric, Rebecca Bliege Bird, and D. Bird. 2003. “The Benefits of Costly Signaling: Meriam Turtle Hunters.” *Behavioral Ecology* 14: 116–126.

Smith, Vernon. 2002. “Constructivist and Ecological Rationality.” Nobel Prize lecture.

- Sneppen, Kim, Per Bak, Henrik Flyvbjerg, and Mogens Jensen. 1994. "Evolution as a Self-Organized Critical Phenomenon." *Proceedings of the National Academy of Sciences* 92: 5209–5213.
- Solow, Robert M. 1956. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70, no. 1: 65–94.
- Spence, A. Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics* 87, no.3: 355–374.
- Squicciarini, Mara, and Nico Voigtländer. 2015. "Human Capital and Industrialization: Evidence from the Age of Enlightenment." *Quarterly Journal of Economics* 30, no. 4: 1825–1883.
- Starfield, Anthony, Karl Smith, and Andrew Bleloch. 1994. *How to Model It: Problem Solving for the Computer Age*. Minneapolis, MN: Burgess International.
- Stein, Richard A. 2011. "Superspreaders in Infectious Diseases." *International Journal of Infectious Diseases* 15, no. 8: e510–e513.
- Sterman, John D. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. New York: McGraw-Hill.
- Sterman, John. 2006. "Learning from Evidence in a ComplexWorld." *American Journal of Public Health* 96, no. 3: 505–515.
- Stiglitz, Joseph. 2013. *The Price of Inequality: How Today's Divided Society Endangers Our Future*. New York: W. W. Norton.
- Stock, James H., and Mark W. Watson. 2003. "Has the Business Cycle Changed and Why?" In *National Bureau of Economic Research Macroeconomics Annual 2002*, vol.17, ed. Mark Gertler and Kenneth Rogoff, 159–218. Cambridge, MA: MIT Press.

- Stone, Lawrence D., Colleen M. Keller, Thomas M. Kratzke, and Johan P. Strumpfer. 2014. "Search for the Wreckage of Air France Flight AF 447." *Statistical Science* 29, no. 1: 69–80.
- Storchmann, Karl. 2011. "Wine Economics: Emergence, Developments, Topics." *Agrekon* 50, no. 3: 1–28.
- Suki, Bela, and Urs Frey. 2017. "A Time Varying Biased Random Walk Model of Growth: Application to Height from Birth to Childhood." *Journal of Critical Care* 38: 362–370.
- Surowiecki, James. 2006. *The Wisdom of Crowds*. New York: Anchor Press.
- Syverson, Chad. 2007. "Prices, Spatial Competition, and Heterogeneous Producers: An Empirical Test." *Journal of Industrial Economics* 55, no. 2: 197–222.
- Taleb, Nassim. 2001. *Fooled by Randomness*. New York: Random House.
- Taleb, Nassim. 2007. *The Black Swan: The Impact of the Highly Improbable*. New York: Random House.
- Taleb, Nassim. 2012. *Antifragile: Things That Gain from Disorder*. New York: Random House.
- Tassier, Troy. 2013. *The Economics of Epidemiology*. Amsterdam: Springer.
- Tetlock, Phillip. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton, NJ: Princeton University Press.
- Thaler, R. H. 1981. "Some Empirical Evidence on Dynamic Inconsistency." *Economic Letters* 8, no. 3: 201–207.
- Thompson, Derek. 2014. "How You, I, and Everyone Got the Top 1 Percent All Wrong: Unveiling the Real Story Behind the Richest of the Rich." *Atlantic*, March 30.

- Thorndike, Edward L. 1911. *Animal Intelligence*. New York: Macmillan.
- Tilly, Charles. 1998. *Durable Inequality*. Berkeley: University of California Press.
- Tsebelis, George. 2002. *Veto Players: How Political Institutions Work*. Princeton, NJ: Princeton University Press.
- Turchin, Peter. 1998. *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants*. Sunderland, MA: Sinauer Associates.
- Tweedle, Valerie, and Robert J. Smith. 2012. “A Mathematical Model of Bieber Fever: The Most Infectious Disease of Our Time?” In *Understanding the Dynamics of Emerging and Re-Emerging Infectious Diseases Using Mathematical Models*, ed. Steady Mushayabasa and Claver P. Bhunu. Cham, Switzerland: Springer.
- Ugander, Johan, Brian Karrer, Lars Backstrom, and Cameron Marlow. 2011. “The Anatomy of the Facebook Social Graph.” arXiv: 1111.4503.
- Updike, John. 1960. “Hub Fans Bid Adieu.” *New Yorker*, October 22.
- US Bureau of Labor Statistics. 2013. *Consumer Expenditures in 2011*. Report 1042, April. Washington, DC: BLS.
- Uzzi, Brian, Satyam Mukherjee, Michael Stringer, and Ben Jones. 2013. “Atypical Combinations and Scientific Impact.” *Science* 342: 468–471.
- Van Noorden, Richard. 2015. “Interdisciplinary Research by the Numbers.” *Nature*, September 16.
- von Neumann, John, and Morgenstern, Oskar. 1953. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.
- Vriend, Nicolaas J. 2000. “An Illustration of the Essential Difference

Between Individual and Social Learning, and Its Consequences for Computational Analyses.” *Journal of Economic Dynamics and Control* 24: 1–19.

Wainer, Howard. 2009. *Picturing the UncertainWorld*. Princeton, NJ: Princeton University Press.

Wakeland, W., A. Nielsen, and P. Geissert. 2015. “Dynamic Model of Nonmedical Opioid Use Trajectories and Potential Policy Interventions.” *American Journal of Drug and Alcohol Abuse* 41, no. 6: 508–518.

Waltz, Kenneth. 1979. *Theory of International Politics*. New York: McGraw-Hill.

Washington Post. 2012. “Mad Money: TV Ads in the 2012 Presidential Campaign.” <http://www.washingtonpost.com/wp-srv/special/politics/track-presidential-campaignads-2012>.

Watts, Duncan. 2011. *Everything Is Obvious Once You Know the Answer*. New York: Crown Business.

Watts, Duncan, and Steven Strogatz. 1998. “Collective Dynamics of ‘Small-World’ Networks.” *Nature* 393, no. 6684: 440–442.

Weisberg, Michael. 2007. “Three Kinds of Idealization.” *Journal of Philosophy* 104, no. 12: 639–659.

Weisberg, Michael. 2012. *Simulation and Similarity: Using Models to Understand the World*. Oxford: Oxford University Press.

Weisberg, Michael, and Muldoon, Ryan. 2009. “Epistemic Landscapes and the Division of Cognitive Labor.” *Philosophy of Science* 76, no. 2: 225–252.

- Weitzman, Martin L. 1979. "Optimal Search for the Best Alternative." *Econometrica* 77: 641–654.
- Weitzman, Martin L. 1998. "Recombinant Growth." *Quarterly Journal of Economics* 2: 331–361.
- Wellman, Michael. 1990. "Fundamental Concepts of Qualitative Probabilistic Networks." *Artificial Intelligence* 44: 257–303.
- Wellman, Michael. 2013. "Head to Head: Does US High-Frequency Trading Need Stricter Regulatory Oversight? (YES)." *International Financial Law Review*, September.
- West, Geoffrey. 2017. *Scale: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies*. New York: Penguin.
- Whittle, Peter. 1979. "Discussion of Dr Gittins' Paper." *Journal of the Royal Statistical Society, Series B* 41, no. 2: 148–177.
- Whitty, Robin W. 2017. "Some Comments on Multiple Discovery in Mathematics." *Journal of Humanistic Mathematics* 7, no. 1: 172–188.
- Wigner, Eugene. 1960. "The Unreasonable Effectiveness of Mathematics in the Natural Sciences." *Communications in Pure and Applied Mathematics* 13, no. 1.
- Wilkinson, Richard, and Kate Pickett. 2009. *The Spirit Level: Why Greater Equality Makes Societies Stronger*. London: Bloomsbury.
- Wilson, David Sloan. 1975. "A Theory of Group Selection." *Proceedings of the National Academy of Sciences* 72, no. 1: 143–146.
- Wolfram, Stephen. 2001. *A New Kind of Science*. Champaign, IL: Wolfram Media.

- Wright, Robert. 2001. *Nonzero: The Logic of Human Destiny*. New York: Vintage.
- Wu, Jianzhong, and Robert Axelrod. 1995. "How to Cope with Noise in the Iterated Prisoner's Dilemma." *Journal of Conflict Resolution* 39, no. 1: 183–189.
- Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi. 2007. "The Increasing Dominance of Teams in the Production of Knowledge." *Science* 316, no. 5827: 1036–1039.
- Xie, Yu. 2007. "Otis Dudley Duncan's Legacy: The Demographic Approach to Quantitative Reasoning in Social Science." *Research in Social Stratification and Mobility* 25: 141–156.
- Xie, Yu, Alexandra Killewald, and Christopher Near. 2016. "Between - and Within-Occupation Inequality: The Case of High Status Professions." *Annals of the American Academy of Political and Social Science* 663, no. 1: 53–79.
- Youn, Hyejin, Deborah Strumsky, Luis Bettencourt, and José Lobo. 2015. "Inventions as a Combinatorial Process: Evidence From US Patents." *Journal of the Royal Society Interfaces* 12: 0272.
- Zagorsky, Jay. 2007. "Do You Have to Be Smart to Be Rich? The Impact of IQ on Wealth, Income and Financial Distress." *Intelligence* 35: 489–501.
- Zahavi, Amotz. 1974. "Mate Selection: A Selection for a Handicap." *Journal of Theoretical Biology* 53, no. 1: 205–214.
- Zak, Paul, and Stephen Knack. 2001. "Trust and Growth." *Economic Journal* 111, no.470: 295–321.
- Zaretsky, Adam. 1998. "Have Computers Made Us More Productive? A

Puzzle.” *Regional Economist*, Federal Reserve Bank of St. Louis.

Ziliak, Stephen T., and Deirdre N. McCloskey. 2008. *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives*. Ann Arbor: University of Michigan Press.

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

湛庐阅读App：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本

怎么读是读者面临的最大阅读障碍

“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”App。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。

从单一到复合，从知道到精通，从理解到创造，湛庐希望建立一个“与最聪明的人共同进化”的社区，成为人类先进思想交汇的聚集地，与你共同迎接未来。

与此同时，我们希望能够重新定义你的学习场景，让你随时随地收获有内容、有价值的思想，通过阅读实现终身学习。这是我们的使命和价值。

湛庐阅读App玩转指南

湛庐阅读App 结构图:



三步玩转湛庐阅读App:



App获取方式:

安卓用户前往各大应用市场、苹果用户前往 App Store 直接下载“湛庐阅读”App，与最聪明的人共同进化！

使用App扫一扫功能， 遇见书里书外更大的世界！

大咖优质课、
献声朗读全本一键了解，
为你读书、讲书、拆书！

你想知道的彩蛋
和本书更多知识、资讯，
尽在延伸阅读！

快速了解本书内容，
湛庐千册图书一键购买！



延伸阅读

《多样性红利》

- ◎ 一个人是否聪明不是由智商决定的，而取决于认知工具的多样性！
- ◎ 《多样性红利》创造性提出多样性视角、启发式、解释和预测模型四个认知工具箱框架。
- ◎ 广受欢迎的“模型思维课”主讲人斯科特·佩奇经典名作！



《表象与本质》

- ◎ 闻名世界的认知科学家侯世达凭借独特的智慧与天赋，联合法国心理学家桑德尔，终于向世人展示了这部极具开创性的著作，一解人类认知之谜。
- ◎ 人类大脑中的每个概念都源于多年来不知不觉中形成的一长串类比，这些类比赋予每个概念生命，我们在一生中不断充实这些概念。大脑无时无刻都在作类比。类比，就是思考之源和思维之火。



《复杂经济学》

- ◎ 作为“复杂经济学”的创始人，布莱恩·阿瑟在本书中汇集了多年对复杂经济学的研究。其核心思想可以归结为：经济不一定处于均衡状态，演绎推理将被归纳推理所取代。
- ◎ 这是一本见证“复杂经济学”成长的著作，你将看到“复杂性思维”在经济学领域是如何发展起来并形成一门崭新的学科。
- ◎ 详细讨论了“爱尔法鲁酒吧问题”和“圣塔菲人工股票市场”两个重要项目。



《财富的起源》

- ◎ 物联网时代的《国富论》！牛津大学新经济思想掌门人埃里克·拜因霍克重磅力作！用复杂性重新定义经济、商业和财富，帮你理解

复杂时代的复杂经济！

- ◎ 《财富的起源》回溯了经济学的发展脉络，用动态、主体、互动、涌现和进化5大视角审视当下经济，阐释商业设计、物理技术、社会技术促进财富增长的进化力量，提出应用复杂经济学的4大应用领域：战略、组织、金融与公共政策，助你利用这部物联网时代的《国富论》决胜商业社会！



-
- (1) 二人均为简·奥斯汀的小说《爱玛》中的人物。——编者注
- (2) 重言式，又称永真式。逻辑学名词。如果一个公式，对于它的任一解释，其真值都为真，就称为重言式。经济上称为套套逻辑。——编者注
- (3) 乔纳森·海特的著作《象与骑象人》详细讲述了这个比喻。其中文简体字版已由湛庐文化策划，浙江人民出版社出版。——编者注
- (4) “爱尔法鲁酒吧问题”最早由经济学家布莱恩·阿瑟提出，他的著作《复杂经济学》《技术的本质》中文简体字版已由湛庐文化策划，浙江人民出版社出版。——编者注
- (5) 蒙特·卡罗方法，又称模拟统计方法，是一种以概率统计理论为指导的一类重要的数值计算方法，是使用随机数来解决很多计算问题的方法。——编者注
- (6) 丹尼尔·丹尼特是著名的哲学家，他的另一本著作《直觉泵和其他思考工具》中文简体字版已由湛庐文化策划，浙江教育出版社出版。——编者注
- (7) 大宗商品指无法区分的产品，如原油、猪肚和2号红小麦等。——编者注
- (8) 戈雅，全名弗朗西斯科·何塞·德·戈雅·卢西恩特斯（Francisco José de Goya Lucientes），西班牙浪漫主义画派画家，代表作《裸体的马哈》。——编者注

- (9) 卡特尔，由一系列生产类似产品的独立企业所构成的组织，目的是提高该类产品价格并控制其产量。——编者注

CHEERS
湛庐

全球思想家正在读的
20本书之一

THE CHECKLIST MANIFESTO: HOW TO GET THINGS RIGHT



清单革命

如何持续、正确、安全地
把事情做好

ATUL GAWANDE
BESTSELLING AUTHOR
OF BETTER AND COMPLICATIONS

[美]阿图·葛文德 著
王佳艺 译

 浙江教育出版社
ZHEJIANG EDUCATION PUBLISHING HOUSE

版权信息

书名：清单革命

作者：〔美〕阿图·葛文德（Atul Gawande）

译者：王佳艺

出版社：浙江教育出版社·湛庐

出版时间：2022年09月

ISBN：9787572243004

字数：173千字

版权所有·侵权必究

各方赞誉

巴曙松 北京大学汇丰金融研究院执行院长 中国银行业协会首席经济学家

中国传统文化中凝聚着许多从容的智慧，例如，“花繁柳密处，拨得开，才是手段；风狂雨急时，立得定，方见脚跟”。在生活节奏日益加快、社会分工日趋多元的现代社会，如何才能做到这样从容？本书提供了一个简洁易行的清单工具，让从容的智慧在现代社会中找到了一个载体。

廖新波 广东省卫生健康委员会前副主任

没有医生不想把病人治好，但没有谁能保证一定治好所有病人。世界上没有能确保成功挽救生命的清单，但是医生可以通过流程清单减少错误的出现。医学的清单是用生命书写的，也是用生命完善的！记住，生命不能重来，珍惜你手上的清单吧！

王杉 北京大学人民医院前院长

一张小小的清单，就让一家医院的中心静脉置管感染事件发生率从11%下降到了0，在15个月内避免了43起感染事件和8起死亡事故，同时还让医院员工的工作满意度上升了19%，手术室护士的离职率从23%下降到了7%。本书不仅仅给了医院管理者一个强大的思维工具，更带来了一场深入人心的观念革命。

于莺 北京协和医院急诊科前主治医师

有了清单，临床医学就有了统一的操作标准，诊疗方案不会因为医生的个体化差异而明显不同。在某种程度上，这是质量控制标准。本书不只是一本管理学书籍，更是一本清单使用说明书。人们在阅读后，可以给自己列个清单。自查，自省，才能少走弯路，避免不必要的损失。

修金来 《中国医院院长》杂志前副社长兼主编

还是一名医学生时，听老师说，医院是人间修罗场。而工作以后才

发现，这一说法虽然有些夸张，却基本是事实。要挽救患者的生命，医生仅有医术和责任心是不够的，还必须拥有应对复杂局面的有效方法。幸运的是，有效方法不一定复杂，它可以简单到只是一张工作清单！

薄三郎 上海长海医院麻醉科副主任、副教授

清单是帮助你正确地工作的必要工具。乍看之下，清单并不算高科技，但使用它绝非多此一举。清单的本质在于沟通，它的实施离不开划定角色，遵照流程。使用清单，就是检视那些微小且容易忽视的问题。这些问题看似无关全局，却往往会给人致命一击。

秦朔 新媒体品牌“秦朔朋友圈”发起人

生活在信息、知识、技术的海洋中，你常常会有无知感、无力感、沮丧和挫败感。面对这种外部复杂性，你需要具备一种庖丁解牛、化繁为简、一次就把事情做对的思维和方法。《清单革命》是你的必读书！

董雷 中国东方航空公司机长 中国民航管理干部学院客座教授

我是用飞行员的眼光来阅读这本书的。掩卷之余，“飞行清单”已不仅仅是飞行程序的一部分。本书作者让我明白了清单的真正含义：如何让复杂的知识真正发挥作用，有效避免“无能之错”。读完本书，相信你也会有一个清晰的答案。

马尔科姆·格拉德威尔 畅销书《引爆点》作者

本书涉及的主题几乎与现代社会的所有方面都息息相关。我们该如何应对越来越复杂的工作？在书中，作者赋予了清单伟大的责任。他认为，现代社会要求我们重新审视专业分工：专家也会犯错，他们也需要帮助；但是否接纳清单，还要看他们是否承认这一事实。

史蒂芬·列维特 畅销书《魔鬼经济学》作者

我一口气读完了这本高水准的著作。书中不仅有很多有趣的故事，还改变了我对世界的看法。这是很长一段时间以来我读过的最佳作品。

《纽约时报》

很少有人能够像阿图·葛文德这样，把手术中突然出错的场景描写

得如此活灵活现，也很少有人能够像他那样，把专家正在为降低这种风险所做的努力描写得如此清晰。

《华盛顿邮报》

即使是抱着怀疑态度的读者，也会觉得本书列出的论据非常具有说服力。作者用生动的描述和有力的证据，呼吁医生采用清单这一看似稀松平常的工具来提高工作效率。

《美国医学会杂志》

本书文字优美，故事动人，论据和论点都很有说服力。在过去几年里，作者亲自参与了世界手术安全清单的研发和实施工作，并投入了大量精力。在书中，他与大家分享了这段不寻常的经历，并全力呼吁医护人员使用简单却被证明很有效的清单来改变工作方式。他的呼吁应该得到关注和回应。

《纽约时报书评》

航空业和建筑业的实例证明了清单的有效性，本书极具吸引力。

《商业周刊》

本书涉及的领域远不止医疗行业。读过本书后，你或许会想尝试制作清单，而且你会从中获益良多。

《出版人周刊》

本书生动地展现了这样一个有趣的情况：不起眼的清单能让你做得更好。

《西雅图时报》

清单看似很傻，但非常有用，本书用大量事例证明了这一点。作者在书中给出的理由极为充分，如果有人看了书还不认可清单的作用，我们就要向他提出质疑了。

《卫报》

本书通俗易读，却意义深远。作者用生动的事例和大家讨论了一个严肃的话题。

彭博新闻社

本书十分引人入胜。作者用有力的事实告诉我们：使用清单，大有裨益。

测一测 你知道如何用清单解决复杂问题吗？

· 随着医疗技术的不断发展，手术失误可以完全被避免，这是对的吗？（ ）

A. 对

B. 错

· 清单是大而全的操作手册吗？（ ）

A. 是

B. 否

· 在以下哪件事情上，清单的适用性最差？（ ）

A. 烘焙蛋糕

B. 将火箭送上月球

C. 飞行前安全检查

D. 培养子女

扫描下面二维码查看本书更多测试题

扫码鉴别正版图书
获取您的专属福利



扫码获取全部测试题及答案，
一起了解如何用清单解决复杂问题

推荐序

一场捍卫安全与正确的清单革命

马尔科姆·格拉德威尔
畅销书《引爆点》作者

阿图·葛文德对现代医学面临的问题和挑战有着深入的思考。《清单革命》一书也正是从医学角度切入，讲述了作者的诸多从医经历。但我们很快就发现，本书的主题，即我们该如何应对越来越复杂的工作，几乎与现代社会的所有方面都息息相关。如此震撼又发人深省的好书，我很久没有读到过了。

阿图·葛文德认为，人类的错误主要分为两类：一类是“无知之错”，一类是“无能之错”。“无知之错”是我们因为没有掌握正确知识而犯下的错误，“无能之错”是我们掌握了相关知识，却没有正确使用这些知识而犯下的错误。在书中，他为我们列举了许多医学方面的案例，从中我们可以知道，手术的例行程序异常复杂，医护人员不可避免会犯下这样或那样的错误。在压力重重的环境中，即使是再优秀的医生也难免漏掉其中一个步骤、少问一个关键问题，以致在手术过程中出现失误。阿图·葛文德还拜访了飞行员和建造摩天大楼的建筑人员，并从他们那里找到了应对复杂问题的方法：制作清单。即使是超级专家也需要一张清单把一些关键步骤写在小卡片上以帮助其完成工作。在书中，阿图·葛文德还描述了他的研究团队如何采用这一理念，研发出手术安全清单。目前，该清单已在世界各地投入使用，并取得了非凡的成果。

虽然我的评论篇幅较短，但读者千万不要错误地认为这本书的视角带有局限性，得出的结论也稀松平常，事实并非如此。阿图·葛文德擅长写作，而且很会讲故事，在书中，他为清单赋予了伟大的责任。他认为，现代社会要求我们重新审视专业分工：专家也会犯错，他们也需要帮助；但是否接纳清单，还要看专家是否承认这一事实。

引言

“无知之错”与“无能之错” 可以原谅的与不被原谅的

·为什么在美国本该在90分钟内完成的心脏急救检查，完成率不到50%？为什么会有高达68%的死刑判决存在错误？我们到底能掌控多少事情？又有哪些事情根本不在我们的可控范围内？

·或许，我们犯错，是因为没有掌握相关的知识；或许，我们犯错，不是因为没有掌握相关的知识，而是没有正确使用这些知识。人类的错误分为两大类型：“无知之错”，可以原谅；“无能之错”，不被原谅。

为什么会有一件事谁都没有做

我的大学同学约翰是旧金山的一名普外科医生，只要有机会凑在一起，我们会交流自己在行医过程中碰到的惊险故事，每个外科医生都有很多这类经历。一天，约翰就给我讲述了一个这样的故事。

在万圣节的晚上，他的医院接收了一个被刺伤的病人，这名男子因在化装舞会上和别人发生争执而受伤。

起初，病人情况稳定，呼吸正常，也没有表现出疼痛难忍的样子。他似乎只是喝多了，嘴里嘟嘟囔囔地不知道在说些什么。创伤组医护人员迅速用剪刀将他的衣服剪开，然后对他的身体进行仔细检查。这名男子略显肥胖，大概有90公斤重，赘肉主要集中在肚子上，而伤口也在这个部位。伤口长5厘米，就像张开的鱼嘴，把病人的腹腔大网膜也翻了出来。约翰他们判断，只需把这名男子推进手术室，进行仔细检查，确保他的内脏没有受到损伤，然后将那个小伤口缝合就行了。

“没什么大事。”约翰当时说。

如果病人伤势很严重，场面就会截然不同：创伤组会冲进手术室，病人的担架床会被飞快地推进去，护士们则会迅速准备好各种手术器械，而麻醉医生也不得不匆匆就位，他们没有时间仔细查阅病人的病历。但是，当时病人的情况并不严重。创伤组觉得有充足的时间，不用火急火燎。所以，他们让病人躺在创伤诊疗室的担架床上，等待手术室准备就绪。

但情况急转直下，一个护士发现那个病人不说话了。他心跳过速，眼睛上翻，而且在护士推他的时候一点反应也没有。这位护士立刻发出了急救警报，创伤组成员蜂拥而至。那时候，病人的血压都快没了。医生和护士立刻为他输氧，并迅速为其补液，但病人的血压还是没有上升。

于是，我刚刚提到的那个假想场面不幸变成了现实：创伤组冲进了手术室，病人的担架床被飞快地推了进去，护士们迅速准备好各种手术器械，麻醉医生不等仔细查阅病人的病历就匆匆就位，一名住院医师将一整瓶消毒液倒在病人的肚子上。约翰抄起一把大个儿的手术刀，干净利落地在病人的肚子上划出了一条上至肋骨、下至耻骨的长口子。

“电刀。”

约翰将电刀头沿着切口的皮下脂肪向下移动，将脂肪分开，然后再将腹肌的筋膜鞘分开。就在他打开病人腹腔的一刹那，大量鲜血从腹腔内喷涌而出。

“糟了！”

到处都是血。这可不是一般的刺伤，那把刀子扎进去足足有30多厘米深，一直扎进了脊柱左侧的主动脉，就是那根将血液从心脏送出的大动脉。

“谁会做出这么疯狂的事情来？”约翰说。另一个外科医生马上用拳头压在血管破裂处的上方，可怕的大出血终于得到了一定的控制，危急的局势渐渐稳定下来。约翰的同事说，自从越南战争结束后，他就再也没有见过这么严重的创伤。

结果还真是和这个医生说的类似。约翰后来才知道，在那天的化装舞会上，行凶者扮成了一名士兵，他的枪还装上了刺刀。

这位病人在死亡线上挣扎了几天，最终还是挺了过来。直到现在，约翰只要一提起这件事还是会懊悔地连连摇头。

当病人被送到急诊室的时候，医护人员几乎做了他们应该做的一切：他们对病人从头到脚进行检查，仔细跟踪测量病人的血压、心率和呼吸频率，检查病人的意识是否清醒，为病人输液，打电话让血库准备好血袋，而且给病人插上了导尿管以确保其尿液排尽。该做的他们几乎都做了，但就是有一件事情谁都没有做，那就是询问送病人过来的急救人员，到底是什么器械造成了创伤。

“你绝不会想到有人会在旧金山被刺刀扎伤。”约翰只好这么解释。

他们差点害死那个病人

约翰还给我讲述了另一个惊心动魄的故事。一位病人因为胃癌而需要进行肿瘤切除手术。在手术过程中，病人的心跳突然停止了。约翰看了看心脏监视器，然后对麻醉医生说：“你看，病人的心跳是不是停止了？”监视器屏幕上的心跳波形变成了一条平平的直线，就像仪器的电极根本没有接到病人的身体上一样。

麻醉医生说：“肯定是电极脱落了。”他一点都没有觉得病人的心跳真的会停止，因为这个40多岁的病人非常健康，医生只是在不经意间才发现了他胃部的肿瘤。

病人在看其他病，可能是在看咳嗽的时候提到自己有一点胃灼热的症状，实际上那种感觉连胃灼热都算不上。他只是觉得在吃东西的时候，有时候食物会卡在食道里，这让他感觉有点胃灼热。医生随即让他做了一次钡餐造影检查。他站在X光机前面，将乳白色的钡剂喝下。医生在屏幕上看到，在病人胃部的顶端有一团老鼠般大小的组织时不时地压迫胃的入口，就像一个瓶塞时开时闭一样。由于发现得早，癌细胞并没有扩散。要想根治此病只有靠手术，医生需要将病人的整个胃部切除，手术大约需要进行4个小时。

一开始，手术进行得很顺利，病人的胃部成功切除，什么问题都没有。但就在医生准备重建病人消化道的时候，心脏监视器上的波形变成了直线。过了5秒钟，手术团队才意识到，监视器的电极并没有脱落，因为麻醉医生根本无法摸到病人的脉搏，他的心脏真的停止跳动了。

约翰把盖在病人身上的手术单扯掉，马上对其进行胸外心脏按压，病人的肠子随着约翰的动作一起一伏。一位护士立刻发出了蓝色警报。

说到这里，约翰停了下来，他问我：“如果你是我的话会怎么处理？”

我使劲儿思考可能会出现的问题。在实施大手术的过程中，病人的心脏停止了跳动，这很可能是由于大出血造成的。所以，我会加大输血量，并寻找出血点。

麻醉医生也是这么说的。但病人的腹腔完全敞开着，并没有任何大出血的迹象。

“麻醉医生得知这一情况后根本无法相信，”约翰说，“他一直不停地叫道，‘一定有大出血！一定有大出血！’但真的没有。”

缺氧也会造成心跳停止。我告诉约翰，我会将输氧量开到最大，然后检查病人的气道，还会将病人的血样送到实验室进行化验，以排除其他异常情况。

约翰说，他们也想到了，但病人的气道没有任何问题。而血样的化验结果至少要等20分钟才能出来，那时候病人可能早就命归黄泉了。

气胸会不会是造成病人心跳停止的原因呢？这种可能也被排除了。医生用听诊器进行了仔细检查，结果两侧肺部的气流运动情况良好。

那么，一定是肺栓塞引起的，我继续猜测原因，一定有血凝块随同血流进入了病人的心脏，并阻止了血液循环。

这种情况非常罕见，但动大手术的癌症患者的确会面临这样的风险。如果情况真是这样，那么我们没有太多办法挽救病人的生命。医生可以给病人注射大剂量的肾上腺素，或实施电击除颤，但这些举措可能没有太大用处。

约翰说，他的团队得出了相同的结论。手术团队对病人进行胸外按压已经整整15分钟了，但显示器上依然是一条意味着死亡的直线。大伙儿都快绝望了。前来帮忙的医护人员中有一位高级麻醉医生，他在病人被麻醉的时候曾经来过手术室。他走后，一切似乎都很正常，没有出什么状况，但他觉得一定是有人出了纰漏。

他问当值的麻醉医生，在病人心跳停止前的15分钟里是否进行过什么与往常不同的操作。

当值麻醉医生说没有，但他马上又改了口。手术的第一次例行化验结果显示，病人除了血钾水平较低以外一切正常，于是麻醉医生为病人注射了一剂氯化钾进行修正。

我怎么就没想起来血钾水平呢？血钾水平异常是造成心脏停止跳动的典型原因之一，任何一本教科书里都会提到。我真不敢相信自己竟然忽视了这一点。血钾水平过低会造成心脏停止跳动，医生要为病人注射一定剂量的氯化钾以使其血钾水平恢复正常，但血钾水平过高同样会让心脏停止跳动。美国各州就是用这种方法来执行死刑的。

那位高级麻醉医生随即要求查看氯化钾溶液的包装袋。有人从垃圾桶里把包装袋翻了出来，当值麻醉医生果然使用了错误浓度的氯化钾溶液，其浓度是病人所需浓度的100倍。也就是说，他给病人注射了致命剂量的氯化钾。

如此折腾了大半天以后，医护人员不知道还能不能挽救病人的生命。他们已经耽搁了太长时间。但从那一刻起，他们所做的一切都是正确的。手术团队给病人注射了胰岛素和葡萄糖，以此来降低病人过高的血钾水平。由于这些举措需要15分钟才能奏效，病人等不了那么久，所以他们又给病人静脉滴注氯化钙注射液，并让病人吸入一种叫作沙丁胺醇的药物。病人的血钾水平迅速下降，他的心脏再一次开始跳动了。

手术团队终于缓了一口气，但他们不知道自己是否还能完成这台手术。他们差点害死这个病人，而且对此后知后觉。不过，他们还是坚持把手术做完了。约翰走出手术室，将早先发生的一切告知了病人家属。他和病人都非常幸运，因为这个病人最终完全康复了，就好像这一切根本没有发生过一样。

人类错误的两大类型

作为外科医生，我们总是互相交流一些意外的经历，有时候也会讲述一些让自己非常懊恼的故事，因为一些故事中不幸的事情原本是完全可以避免的。我们和同行分享的既有在手术室里获得的巨大成功，也有在那里遭遇过的惨痛失败。每个外科医生都有失手的时候，这是我们工作的一部分。我们喜欢认为一切都在自己的掌控之中，但约翰的故事让我开始反思一个问题：我们到底能掌控多少事情？有哪些事情根本不在我们的可控范围之内？

哲学家塞缪尔·戈罗维兹（Samuel Gorovitz）和阿拉斯代尔·麦金泰尔（Alasdair MacIntyre）曾于20世纪70年代写过一篇关于人类谬误本质的短文。在一次手术培训中，我读到了这篇文章，从那以后，文章中提到的问题就一直萦绕在我心头。两位哲学家想要回答的问题是：我们为何会在实践过程中遭遇失败？他们将导致我们犯错的一类原因称为“必然的谬误”，也就是说，我们所做的事情完全超出了我们的能力范围。人类并非全知全能，即便拥有先进科技，我们的能力也是有限的。关于世界和宇宙，其中有很多领域是我们还未理解也无法掌控的。现在如此，将来也是如此。

但是，的确有不少领域是人类多多少少可以控制的。我们能够建造摩天大楼，能够预测暴风雪，能够挽救心脏病发作或被刺伤的人。戈罗维兹和麦金泰尔指出，人类的错误可以分为两大类型。

第一类错误是“无知之错”，我们犯错是因为没有掌握相关知识。目前我们还只是部分理解了世界的运行规律。有些设计新颖的摩天大楼我们还不知道该怎么建造，有些自然灾害我们还无法预测，有些棘手的疾病我们还不知道该如何预防和救治。

第二类错误是“无能之错”，我们犯错并非因为没有掌握相关知识，而是因为没有正确使用这些知识。所以，一些摩天大楼因为错误的设计或施工而倒塌了，一些突如其来的暴风雪原本是可以预测出来

的，还有一些致命的刺伤被医护人员忽视了，因为他们忘记询问急救人员病人到底是被什么器械刺伤的。

清单宣言 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

我们在实践过程中遭遇失败的原因是：我们所做的事情完全超出了自身的能力范围。人类并非全知全能，即便拥有先进科技，我们的能力也是有限的。

如果把约翰讲述的故事看作**21**世纪初医学问题的一个缩影，我们会惊奇地发现，原来倾向于犯“无知之错”的我们现在越来越倾向于犯“无能之错”了。在人类历史的绝大部分时间里，我们的生活主要被“无知之错”所主宰，给人类带来巨大痛苦的疾病最能说明上述观点。就大多数疾病而言，我们以前并不知道病因是什么，也不知道该如何治疗。但仅仅就在过去的几十年时间里，科学的发展为我们积累了大量知识，使得我们现在较少应对“无知之错”的挑战，更多是投入大量精力来应对“无能之错”的挑战。

以心脏病为例。在20世纪50年代，我们对于如何预防和治疗心脏病几乎还是一无所知。比如，我们并不知道高血压的危害。即便了解了这一点，我们也不知道该如何降低血压。直到60年代，第一种治疗高血压的安全药物才被开发出来，并被证明是有效的。但那时候，我们依然不知道胆固醇、遗传因素、吸烟及糖尿病对心脏造成的危害。

不仅如此，在那个时代，如果有人不幸心脏病发作了，我们也不知道该怎么办。我们会给病人注射吗啡止疼，或许还会让他们吸氧。病人必须绝对静卧休养几周，我们甚至不允许他们起身上厕所，生怕这些动作会给他们受损的心脏造成更大压力。病人家属所能做的只有向上帝祈祷，希望病人能够挺过去。即便他们能够如愿，病人回家后也基本上做不了什么事情，只能悲惨地度过余生。

人类的错误可以分为两大类。第一类是“无知之错”，我们犯错是因为没有掌握相关知识。第二类是“无能之错”，我们犯错并非因为没有掌握相关知识，而是因为没有正确使用这些知识。

但在今天，我们至少有十几种办法，能够有效降低人们心脏病发作的概率。比如控制血压，让高血压患者服用药物以降低胆固醇水平并消除炎症，建议人们控制血糖水平，鼓励人们经常锻炼，帮助人们戒烟等。如果我们发现有人出现了心脏病早期症状，那么还会督促他接受进一步检查。而对于那些心脏病发作的人，我们现在也有很多可供选择的治疗手段，这些手段不仅能够挽救病人的生命，还会降低其心脏受损的程度。我们可以让病人服用药物以疏通其受阻的冠状动脉；我们还有能够撑开这些血管的支架；我们还能做开胸手术，利用搭桥技术绕过受阻血管建立供血通道。在某些情况下，病人需要做的仅仅是卧床休息、吸氧并服用药物。过不了几天，病人就能出院回家过正常生活了。

不被原谅的“无能之错”

现在，我们面临的错误更多的是“无能之错”，也就是无法持续、正确地运用我们所掌握的知识。在众多选项中选择真正适合病人的心脏病治疗方法非常困难，即使对经验丰富的专家来说，也不是一件容易的事情。不仅如此，每种治疗方法都存在很多隐患，可能会引发众多并发症。

现在，我们面临的错误更多的是“无能之错”，也就是无法持续、正确地运用我们所掌握的知识。

研究显示，如果心脏病病人要接受心血管球囊扩张治疗，那么手术必须在发病的病人到达医院后的90分钟内开始。不然，病人生存的概率就会大幅下降。从医院的操作实践角度来说，这意味着，对于每一位急诊就医说自己胸痛的病人，医生必须在90分钟内完成所有检查和化验，做出正确的诊断，制订出相应的治疗方案，再将方案告知病人并获得许可，确认病人没有过敏或其他需要考虑的健康问题，然后通知手术团队，让心导管手术室做好准备，最后将病人送入手术室开始手术。

那么在一家普通的医院里，所有这些步骤能够在90分钟内完成的概率是多少呢？2006年，这一概率还不到50%。这不是什么特例，此类问题在医院里司空见惯。研究发现，至少有**30%**的中风病人、**45%**的哮喘病人以及**60%**的肺炎病人没有得到及时的妥善治疗。即使你知道该怎么做，正确实施治疗的各个步骤还是非常困难的。

我花了不少时间想搞明白：现代医学面临的困境和压力都是由哪些原因造成的？结果发现，主要原因并不是缺少资金或政府干预，也不是法律诉讼或保险公司制造的麻烦，而是现代科学的复杂性以及我们在运用复杂知识时所面临的紧张和压力。这类问题不只发生在美国，它们遍及世界各地。无论在欧洲还是在亚洲，无论在发达国家还是在发展中国家，你都能看到这类问题的踪影。而且让我感到惊讶的是，这类问题甚至不仅仅存在于医学领域。

在人类所有的实践领域中，知识及其复杂性与日俱增。同样，正确运用所掌握的知识也越来越难。比如，政府提出的应对飓风等自然灾害的措施经常出错。而在2005—2008年间，针对律师失误的诉讼案件数量增长了36%，有的律师搞错了开庭时间，有的把卷宗弄丢了，有的则用错了法律条款。当然，漏洞百出的还有各种软件、情报和银行交易。实际上，任何一个需要我们掌控大量复杂知识的领域都容易出现这类问题。

“无知之错”可以原谅，
“无能之错”不可以原谅。如果还没有找到解决某类问题的最佳方法，那么只要尽力了，无论结果如何，我们都能接受；但是，如果明明知道该怎么做，却没有做到，那么这类错误就会让我们暴跳如雷。

“无知之错”可以原谅，“无能之错”不可以原谅。如果还没有找到解决某类问题的最佳方法，那么只要尽力了，无论结果如何，我们都能接受；但是，如果明明知道该怎么做，却没有做到，那么这类错误就会让我们暴跳如雷。我们怎么可以允许半数心脏病患者得不到及时治疗呢？我们怎么能够允许68%的死刑判决过程中出现错误？两位哲学家将此类错误冠以“无能”的恶名，不是没有道理的。被此类错误祸害的当事人会使用更为激烈的词语，比如认为责任人“渎职”，甚至是“丧尽天良”。

不过，对于这些责任人来说，无论是医生、律师，还是政府官员，他们可能会觉得这类评价有失公允。因为他们觉得自己面临的问题非常复杂，需要处理的信息越来越多，每天要不断学习新知识并做出艰难的选择。即使这些从业人员已经付出了很大努力，但由问题复杂性造成的失败远比因为没有责任心而导致的失败多得多。所以，在大多数技术含量很高的专业领域，对于失败的正确处理方法不是惩罚，而是鼓励从业人员积累更多经验和接受更多培训。

即使我们已经付出了很大努力，但由问题复杂性造成的失败远比因为没有责任心而导致的失败多得多。

经验是非常重要的，这一点毋庸置疑。要治疗遭受创伤的病人，外

科医生只看教科书是不够的。即使你熟知贯通伤相关的所有知识，即使你了解各种诊断和治疗方法，即便你懂得快速行动的重要性，也不足以让你成功地治愈病人。一名优秀的外科医生还必须熟悉诊疗的真实环境，熟知实施各种步骤的次序和正确时机。不断练习、不断积累经验会让你熟能生巧，让你获得成功。如果导致失败的原因仅仅是缺乏某些技能，那么只需接受更多的培训并练习就能解决问题。

但这并不是约翰所碰到的问题的症结。他是我所知道的最为训练有素的外科医生之一，他在一线操刀已经有十多年了。这种情况很普遍。无论是在医学领域，还是在其他领域，个人能力往往不是最难克服的障碍。因为在大多数专业领域，人们接受训练的时间往往比在普通领域中要长，培训的强度也是前所未有的。对专业人士来说，无论是医生、教授，还是律师或工程师，在单独执业之前都要进行高强度的基础知识学习，每周花在学习上的时间长达60小时甚至是80小时。我们会利用一切可能的机会来完善自己的技能，并且我们用来掌握和磨炼技能的方法已经非常高效了，但我们还是会经常犯错，无论个人能力多么强都不足以改变这种状况。

我们需要一场真正的变革

这就是我们在21世纪面临的情形：我们已经积累了数量惊人的知识。掌握这些知识的是人类社会里最训练有素、技艺最高超、工作最努力的一群人。而且，他们的确已经通过运用这些知识获得了非凡的成就。但是，要恰当使用这些复杂的知识并不是一件容易的事。在各个领域，从医疗到金融，从商业到行政，可以避免的错误和失败比比皆是，它们让从业人员受挫，让他们士气低落。而造成这一困境的原因越来越明显：我们需要掌握的知识的数量和复杂程度，已经超出了个人正常的能力范围。知识的确拯救了我们，但也让我们不堪重负。

我们需要掌握的知识的数量和复杂程度，已经超出了个人正常的能力范围。知识的确拯救了我们，但也让我们不堪重负。

我们需要开展一场伟大的变革来防止错误与失败。我们将立足于已有的经验，充分利用我们所掌握的知识，利用这一变革弥补人类不可避免的缺陷和不足。这一变革并不十分艰难，甚至简单至极，特别是对那些花了多年时间来培养和磨炼高超技艺的专业人士来说，投身这一变革简直吹糠见米。

这个变革就是：清单革命！

第一部分

清单革命是一场观念变革



THE
CHECKLIST
MANIFESTO
HOW TO GET
THINGS RIGHT

第1章 人人都会犯错 观念的转变

·我们的身体可能以13 000多种不同的方式出问题。在重症监护室中，每位病人平均24小时要接受178项护理操作，而每一项操作都有风险。

·大量复杂的知识会让我们不堪重负。请承认，我们每个人都会犯错；请承认，无论我们进行多么细致的专业分工和培训，还是会忽略一些关键的步骤，还是无法避免一些错误。

为什么一个进了鬼门关的小女孩能被拉回人间

医学学术期刊《胸外科年鉴》（*Annals of Thoracic Surgery*）上曾刊登过一篇论文，作者用干瘪晦涩的学术语言描述了一场发生在阿尔卑斯山一个奥地利小村庄里的噩梦。一对夫妻带着自己3岁大的女儿去屋后的林子里散步，结果一不留神孩子滑进了一个只结了一层薄冰的池塘。他们纵身跳入池塘试图将女儿拉上来，但孩子很快就沉入了水底。过了好一会儿，他们才把孩子救上岸。这对夫妻随即拨打了急救电话，急救人员立刻通过电话指导他们对孩子实施心肺复苏。

8分钟后，急救人员赶到了，但他们发现孩子已经没有生命迹象，她的血压和脉搏都测不到，呼吸也停止了。孩子的体温只有19摄氏度，瞳孔已经放大，对光刺激没有任何反应，这说明她的大脑已经停止了工作。

但急救人员并没有放弃，他们依然给女孩实施心肺复苏。一架直升机将孩子火速送往附近的医院。一路上，急救人员不停地按压女孩的胸腔，等直升机到达医院后，他们直接将女孩推进了手术室，并把她抬到医院的轮床上。一个外科小组随即赶到，以最快的速度为孩子接上人工心肺机。

人工心肺机的个头和一张办公桌相当，外科医生必须将孩子右侧腹股沟的皮肤切开，将一根硅胶导管插入股动脉，让血液流入机器，并把另一根导管插入股静脉，使氧合后的血液被送回体内。一位体外循环灌注师打开人工心肺机的血泵，调整氧含量、温度和流量等参数。女孩的血液开始经体外循环，心肺机的管子也随之变成了鲜红色。直到一切就绪后，急救人员才停止按压女孩的胸腔。

把女孩送到医院的时间和为她接入人工心肺机的时间加起来总共是1.5小时。不过，在“2小时关口”就要到来之际，女孩的体温上升了6摄氏度，她的心脏重新开始跳动，这是她身上第一个恢复功能的内脏器官。

6小时后，女孩的核心体温已经达到正常的37摄氏度。医生试图用机械式呼吸机替换人工心肺机，但是池塘里的水和杂物对孩子的肺造成了严重损伤，输入的氧气无法经由肺部进入血液，所以医生只能为她接上体外膜肺氧合（extracorporeal membrane oxygenation, ECMO）设备。为此，医生必须用一把锯子打开女孩的胸腔，并且将便携式体外膜肺氧合设备的导管直接插入女孩的主动脉和跳动的心脏中。

待体外膜肺氧合设备启动后，医生将人工心肺机的导管从女孩的身

体里移除，然后对血管进行了修复，并将腹股沟上的切口缝合。外科小组随后将女孩送到了重症监护室，她的胸腔依然打开着，上面覆盖着无菌塑料薄膜。在接下来的一整天时间里，重症监护团队一直用纤维支气管镜吸除女孩肺里的积水和杂物。一天后，她的肺恢复良好，可以直接使用机械式呼吸机了。于是，医生又把她送回手术室，将插在她身体中的体外膜肺氧合设备的导管拔掉，修复血管，并将胸腔闭合。

在接下来的两天时间里，女孩的肝、肾和肠等器官都恢复了功能，但脑还是没有反应。CT扫描显示，女孩是整个脑部都有肿胀的迹象，这是弥漫性损伤的特征，但脑部没有任何区域死亡。所以，医生决定采取进一步行动，在女孩的颅骨上钻一个小孔，放入一个探头以监控其脑压，并通过引流脑脊液和使用药物等手段不停地对其脑压进行调整。在随后的一周多时间里，女孩一直处于昏迷状态。但最终，她渐渐地苏醒了过来。

先是她的瞳孔对光线有反应了。然后，她能够自主呼吸了。终于有一天，她张开了双眼，从昏迷中醒了过来。溺水两周后，这个女孩出院回家。虽然右腿和左臂变得不太灵活，说起话来还有些模糊，但只要经过大量的康复治疗，她就可以在5岁前完全恢复。生理检查和神经检测结果都显示，她和其他女孩没什么两样，完全是一个健康的正常人。

这个故事之所以令人惊奇，并不只是因为医生能够把一个进了鬼门关2小时的小女孩拉回人间，更因为他们能够在混乱的医院里有条不紊地成功实施那么多复杂的治疗步骤。

人们常常会在电视里看到这样的场景：溺水者被救上岸后，急救人员用力按压其胸部，并对其实施人工呼吸。随着溺水者一阵咳嗽，肺部的积水被咳出，心跳恢复正常，神智也变得清晰了。但在现实中，急救可没有这么简单。

为了挽救这个小女孩的生命，数十位医护人员要正确实施数千个治疗步骤，比如在插入血泵导管的时候，不能把气泡注入病人的体内，要时刻保证各种导管、女孩敞开的胸腔以及她与外界接触的脑脊髓液不被细菌感染，他们还要使用一堆难伺候的设备，让它们正常运转。上面提到的每一步都很艰难，而要将这些步骤按照正确的顺序一个不落地做好更是难上加难。在整个过程中，医护人员没有任何犯错的余地。

实际上，很多因溺水而心脏停止跳动的孩子没有被挽救回来，不是因为他们的心跳停止太久了，而是因为其他各种原因，比如机器坏了，手术团队动作不够快，或者有人没洗手而导致患者发生了感染。虽然这些案例没有写进《胸外科年鉴》，但它们非常普遍，只不过人们没有意识到这一点罢了。

医学，从非常复杂到更加复杂

我觉得我们都被青霉素忽悠了。亚历山大·弗莱明（Alexander Fleming）在1928年的这项重大发现，为我们展现了一个颇具诱惑力的前景：只要服用一粒小药丸，或者打一针，就能治愈很多疾病。毕竟，青霉素治愈了许多原本无法医治的感染性疾病。那么，我们为什么不能用相似的简单方法治愈各种不同的癌症呢？为什么不能用同样简单的方法来治疗烧伤、心血管疾病和中风呢？

清单宣言 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

医学已经变成一门需要掌控极端复杂性的艺术，成了测试人类是否能够驾驭这种复杂性的一种考验。

现代医学并没有如我们所希望的那样先进。虽然在过去的100年间医学在突飞猛进地发展，但是我们却发现，大多数疾病都有各自的特点，它们的共同之处并没有我们想象的那么多，而且许多疾病仍然难以治愈。即使曾经用青霉素创造过无数奇迹的感染科医生，也面临非常严峻的情况：并非所有细菌菌株都对抗生素敏感，而那些敏感的也会很快发展出耐药性。今天，我们需要使用非常个性化的治疗手段，有时候甚至需要使用多种手段来治疗感染。需要考虑的因素有很多，比如特定菌株的药敏性特征、病人的病情及受影响的器官系统。如今的治疗过程变得越来越复杂，离青霉素为我们展现的前景越来越远。医学已经变成一门需要掌控极端复杂性的艺术，成了测试人类是否能够驾驭这种复杂性的一种考验。

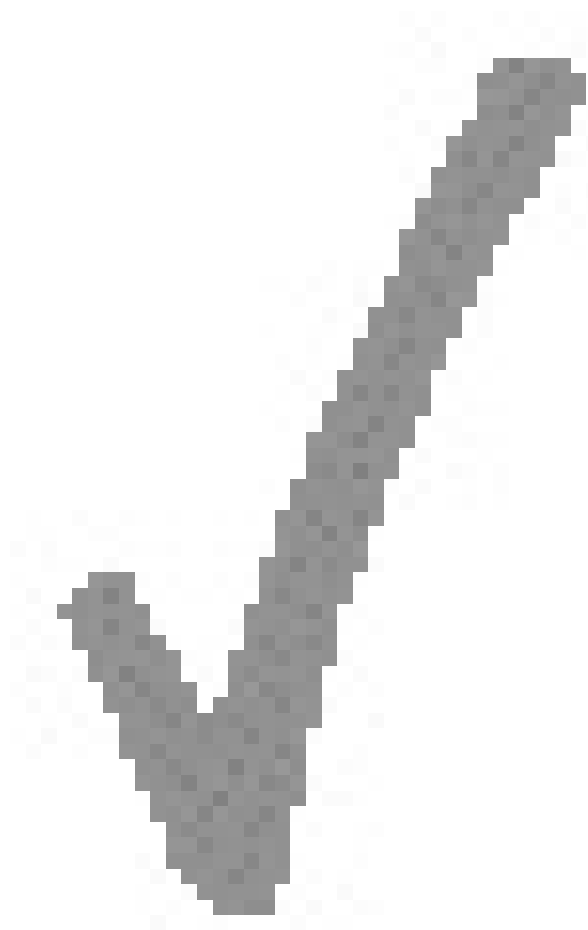
在世界卫生组织发布的《国际疾病分类》（第九版）中，我们可以找到13 000多种不同的疾病、综合征和损伤类型。也就是说，我们的身体可能以13 000多种不同的方式出问题，而科学几乎给上述每一种情况都提供了解决问题的方法。就算我们无法治愈疾病，也能尽量减少疾病给患者带来的损伤和痛苦。不过，每种疾病的治疗方法都是不同的，而且基本上都不简单。现在，医生们手边就有6 000多种药物和4 000多种治疗手段可供选择，每一种都有不同的要求、风险和注意事项，这让医生们很难不出错。

我们医院在波士顿的肯莫尔广场（Kenmore Square）有一家附属社区诊所。“诊所”这个名称可能会让人觉得这家医疗机构很小，其实不

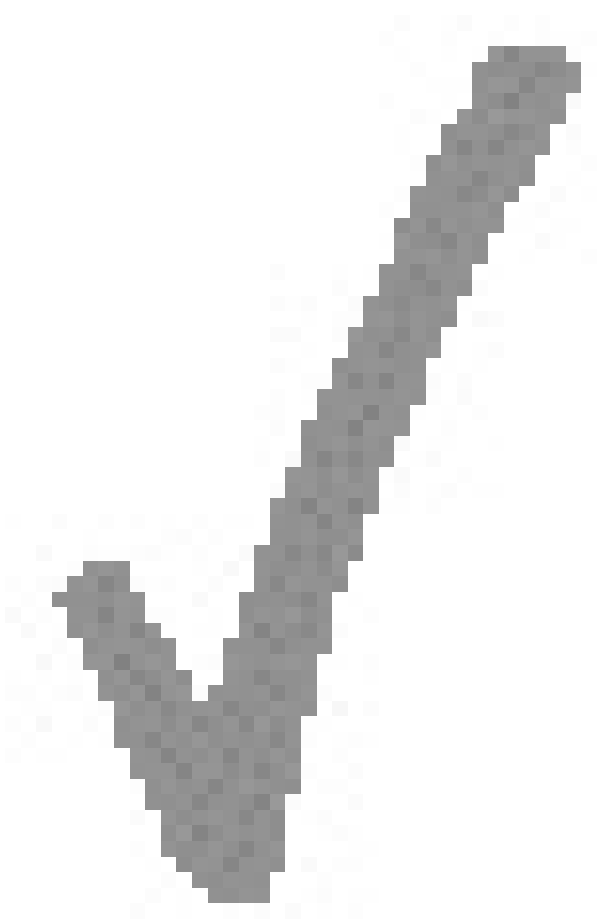
然。这家诊所始建于1969年，现在的名称为“哈佛先锋”（Harvard Vanguard）。该诊所为各个年龄段的人们提供各类门诊服务，但就算只是提供门诊服务也不轻松。

为了跟上医学飞速发展的步伐，这家诊所建立了20多个科室，雇用了600多名医生和其他1 000多名专业人士。他们涉及的专业多达59种，其中有很多专业在诊所建成的时候还没诞生。如果你乘电梯到五楼，从电梯口一路走到普外科，会依次经过多个科室，如普内科、内分泌科、遗传科、手外科、化验实验室、肾科、眼科、整形外科、放射科和泌尿科。这还只是位于一条走廊两边的科室。

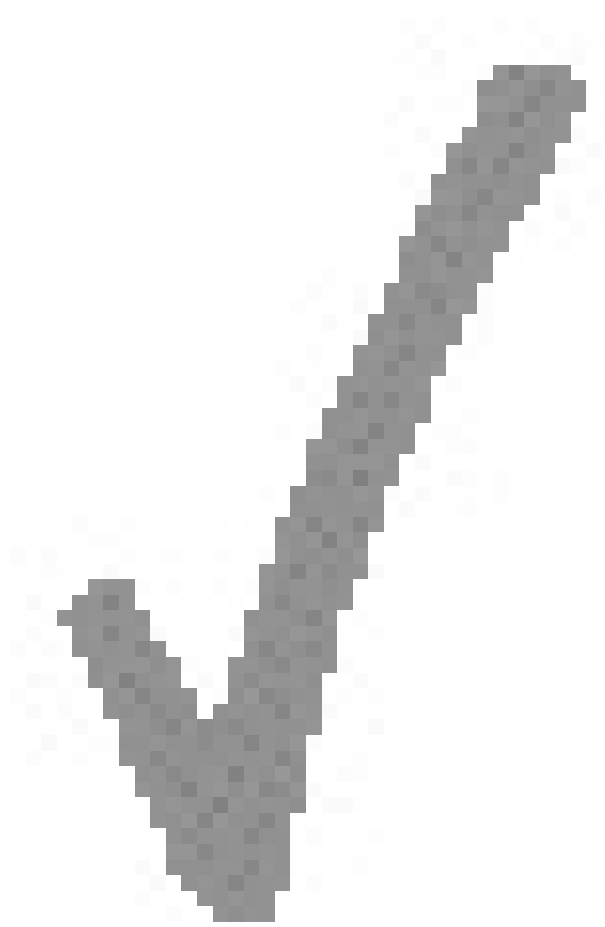
为了应对疾病复杂性这个问题，不同专业的医务人员进行了分工。尽管如此，我们的工作依然非常繁重。就拿我自己某一天的工作来说吧。



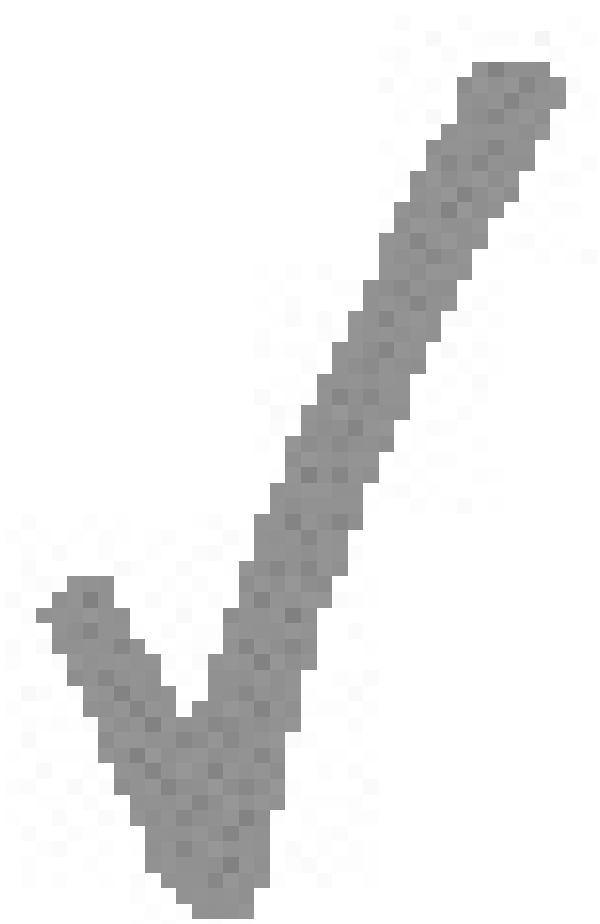
急诊室让我去查看一名25岁的女病人。病人右下侧腹部越来越痛，还伴有发烧、呕吐等症状。我怀疑她得了阑尾炎，但她是个孕妇，所以不能让她接受CT扫描，这么做可能会威胁胎儿的安全。



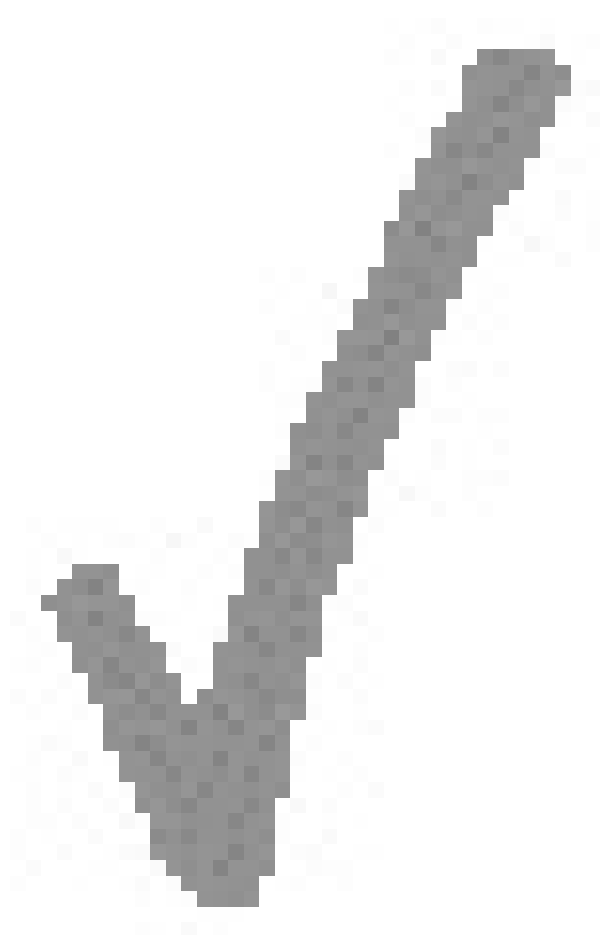
随后，一名妇科肿瘤医生让我去手术室看一看。一位女病人的卵巢上长了一个肿块，医生在切除这个肿块的时候发现，这可能是胰腺癌转移形成的。他想让我检查一下病人的胰腺，看看是否要一并切除。



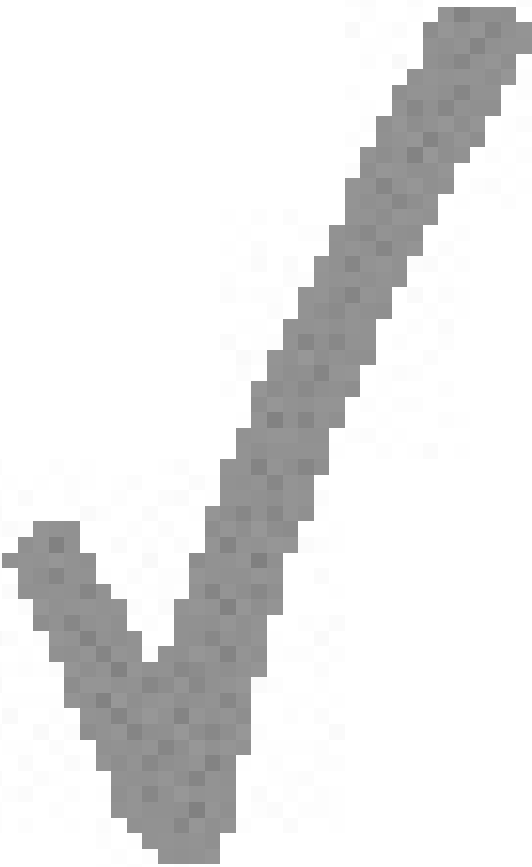
邻近医院的一名医生打电话给我，说要让一名重症病人转到我们医院。这个病人长了一个巨大的肿瘤，这个肿瘤已经阻塞了她的肾脏和肠子，并造成了大出血，情况非常严重，已经超出了他们可以掌控的范围。



后来，我们医院的内科又打电话给我，让我过去看一位61岁的病人。他的肺气肿非常严重，所以医生无法对他进行髋关节置换手术，因为担心他的肺容量不够大，手术风险性高。但是他现在得了很严重的结肠感染，也就是急性憩室炎，医生用了3天抗生素都无济于事。所以，手术是他能活下去的唯一希望。



可能是觉得我的工作还不够乱，另一位同事让我去帮忙看一位52岁的患者，他患有糖尿病、冠心病、高血压、慢性肾衰竭，而且严重肥胖。他得过中风，现在又得了绞窄性腹股沟疝。



还有一名内科医生让我一起给一位直肠瘘病人做诊断，她身体状况很好，但因为这个病可能要进行一次手术。

正如各位看到的那样，我要在一天的时间里处理那么多不同而且复杂的病例。上述6个病人所患的疾病完全不同，此外，我还要进行26次不同的诊断。也许会有人觉得其他医生可能没那么忙，但实际上，我的每个同事每天都要面临许多复杂而棘手的问题。

根据哈佛先锋诊所病历管理部门的统计数据，一年内，仅仅在门诊诊疗过程中，每个医生平均就要诊断250种疾病，而患有这些疾病的病人人们还有其他900多种不同的健康问题需要医生仔细考虑。每个医生平均要开300种药物和100多种不同的化验单，还要使用40种不同的治疗方法，如给病人接种疫苗、接骨等。

在所有医生中，普内科医生面临的问题最繁杂。他们每人平均每年要诊断371种疾病，要考虑其他1 010种健康问题，要开出627种药，还要使用36种治疗方法。只要想一想他们需要掌握多少种不同门类的知

识，就会令人头脑发胀。而这还只是门诊的工作量，医疗管理系统并没有记录医生为住院病人提供服务时需要完成的各种工作。

但就算是只考虑门诊，统计数据也并不完整。因为在使用系统的时候，医生常常会使用“其他”这个选项。如果病人很多，而你的进度已经落下2小时，候诊的病人变得越来越不耐烦，这时候你往往不会在浩瀚的数据库中慢慢寻找相应的诊断代码，而是会选择“其他”这个万能选项草草了事。但话说回来，即便你有时间，也未必能够从系统中找到你想要记录的病症。

大多数美国医院使用的电子病历系统跟不上医学发展的速度，没有把新近发现的或刚刚和其他疾病区分开来的疾病包括在内。我曾经碰到过一个患有肾上腺神经节细胞瘤的病人，这是一种罕见的肾上腺疾病。我还碰到过李-佛美尼综合征（Li-Fraumeni syndrome）患者，这是一种可怕的遗传病，会让病人周身器官都发生癌变。但是，在医疗管理系统的下拉菜单里找不到上述两种疾病的名称，于是，我只能选择“其他”选项了。科学家几乎每周都会有新的医学发现，会找到新的癌症亚型，会制订新的诊断标准，还会发明新的治疗手段。医疗的复杂程度与日俱增，连计算机系统都跟不上了。

但是医疗的复杂性并不只是源于知识的飞速积累，实践问题也是重要的原因。各位已经看到了医务人员面临的问题多么复杂棘手。下面让我们再来看看重症监护室（ICU）。

178项操作，每一项都有风险

“重症监护”这个词听起来不太容易让人理解，这个领域的专业人员更喜欢用“重症护理”一词，不过大家可能还是会觉得不太清楚。但如果我们用“生命支持”这个非医学名称的话，一切就明明白白了。如今，人们可能会遭受许多可怕的病痛，如挤压、灼伤、炸伤、主动脉破裂、结肠破裂、严重心脏病发作、重度感染等。但不可思议的是，即使遭受了如此严重的病痛，很多人还是能活下来。要知道，上面提到的任何一种状况，在过去都能轻而易举地夺走人们的生命。而现在，重症病人存活下来已经变得司空见惯，这在很大程度上要归功于重症监护，因为相关医疗技术让我们能够用人造的方法替代人体失效的功能。当然，这不是轻而易举就能办到的，我们需要有大量科学知识和技术手段作为支撑。

如果病人的肺部出问题，我们需要机械式呼吸机，或许还需要将病人的气管切

开。如果病人的心脏出了问题，我们需要主动脉内气囊泵。如果病人肾脏机能失效，我们则需要血液透析机。如果病人昏迷不醒，无法进食，我们可以通过手术将硅胶导管插入病人的胃部或小肠，这样我们就能将由特殊配方制成的流质食物直接通过导管灌入病人的消化道；要是这位不幸的病人连肠胃都被破坏了，那么我们会将氨基酸、脂肪酸和葡萄糖直接注射进病人的血液里。

在美国，每天有将近**9**万人住进重症监护室，每年有**500**多万人需要接受重症监护。几乎每个人在一生中都有机会光顾这个没人想去的地方。现代医疗在很大程度上要依赖重症监护室里的生命支持系统，如早产儿、重伤患者、中风患者、心脏病患者，以及接受脑部、心脏、肺部等部位大手术的病人。在医院的各种治疗活动中，重症监护占据的比重越来越大。而在**60**多年前，重症监护室基本上不存在。但如果你到现在的医院去看一看，在每天接受诊治的**700**位病人中，差不多会有**155**位接受重症监护。这些病人平均会在重症监护室里待上**4**天时间，存活率大约是**86%**。所以，进入重症监护室，接入机械式呼吸机和各种电线、导管并不意味着被判了死刑。不过，待在那里的几天会是你人生中情形最为危急的几天。

以色列科学家曾对重症监护病人在**24**小时内接受的各项护理进行了调查。研究发现，每位病人平均每天要接受**178**项护理操作，如服药和吸除肺部积液等，而且每项操作都有风险。令人惊讶的是，医护人员操作的错误率只有**1%**。即便如此，这也意味着每位病人平均每天要承受两次左右的失误操作。只有不断降低医护人员操作的错误率，提高成功率，重症监护才能成功挽救更多病人，但这很难做到。

对于一个昏迷的病人来说，仅仅在床上躺上几天就可能会出问题。

他的肌肉会萎缩，骨质会变得疏松，身上会长出褥疮，血管则开始阻塞。医护人员必须帮助病人活动他那软弱无力的四肢以避免发生痉挛。此外，医护人员还必须给病人皮下注射降低血液黏稠度的药物，每过几小时就帮他们翻个身，在给他们擦身和换床单的时候还不能碰掉导管或电线。他们每天要给病人刷两次牙，以防止口腔内繁殖的细菌引发肺炎。如果病人还要使用呼吸机、血液透析机，或者有开放性伤口的话，重症监护的难度就更大了。

威胁无处不在

不妨听我讲述一个故事，大家对重症监护的困难就会有直观的认识了。

有个病人名叫安东尼·德菲利波（Anthony DeFilippo），他已经**48**岁了，是一位来自马萨诸塞州埃弗里特的豪华轿车司机。他因为患有疝气

和胆结石而在一家社区医院接受手术。不幸的是，在手术期间，他出现了大出血的危急情况。虽然值班外科医生最终止住了大出血，但他的肝脏受到了严重损伤。在术后的几天里，他的病情不断恶化，社区医院已经没有能力对他进行医治了。于是，我同意他转院，并接手负责他的治疗，希望能稳定住他的病情。某个周日凌晨的1:30，他被送到了我们的重症监护室。他那乱蓬蓬的黑发紧紧贴在被汗水打湿的额头上，他的身体在颤抖，心率高达每分钟114次。因为高烧、休克和供氧不足，他神志不清、语无伦次。

“我要出去！”他大声叫道，“我要出去！”德菲利波不断用手撕扯他的衣服、氧气面罩和覆盖在腹部伤口上的敷料。

“德菲利波，没事的，”一位护士对他说，“我们会帮助你的。你在医院里。”

德菲利波是个大个子，他一把推开护士，翻身就要下床。我们立刻加大了输氧量，把他的手腕绑在床沿上，并试图和他讲道理。他最终精疲力竭，任由我们抽血和注射抗生素。

化验报告显示，他的肝功能衰竭，白细胞数量异常的高，这说明发生了感染。没过多久，我们发现他的肾功能也衰竭了，因为他的尿袋里一滴尿液也没有。在接下来的几个小时里，他的血压不断下降，呼吸变得越来越困难，而且也不再那么吵吵嚷嚷，因为他已经失去了意识。他的每一个器官，包括他的大脑都开始衰竭。

我给他的姐姐打电话，并将德菲利波的病情告诉了她。她焦急地说：“您一定要想办法救救他。”

我们的确尽了全力。我们给他注射了麻醉剂，一位男住院医师将一根呼吸管插入了他的气管，另一位女住院医师将各种设备连接上他的身体。她将一根5厘米长的细细针头插入德菲利波右手腕的桡动脉，并且用丝线将与针头相连的导管缝在他的皮肤上。随后，她将一根30厘米长的中心静脉置管插入他左侧颈部的颈静脉。当导管缝好后，X射线成像结果显示，导管的顶端正好位于它应该在的位置，也就是心脏入口处的腔静脉中。随后，她将另一根稍粗一点的透析导管穿过德菲利波的右胸上侧，插入他的锁骨下静脉。

我们将呼吸导管的另一头接上呼吸机，并将呼吸频率调整到每分钟14次。我们不断调整呼吸机的压力和流速，就像工程师在操控面板前不断调整参数一样，直到德菲利波血液中氧气和二氧化碳的含量达到我们

预先设定的水平为止。

我们通过插入桡动脉的导管来监测他的动脉血压，并不断调整药物使其达到理想水平；还通过插入颈静脉的导管测量静脉压来调整静脉输液的速度，将插入锁骨下静脉导管的另一头插入血液透析机。这个人工肾脏每隔几分钟就能把他全身的血液过滤一遍，并将透析后的血液送回病人身体。我们还可以通过各种细微的调整来控制他的血钾水平以及血液中碳酸氢盐和氯化钠的含量。德菲利波就像我们手中可以操控的一台机器一样。

他当然不是一台简单的机器。我们面临的困难就好比仅仅依靠几个简单的仪表和操控装置，就要把一辆从山上飞奔而下的18轮汽车安全开到山脚下。光是维持德菲利波的血压就需要几万毫升静脉注射液和一架子的药物。呼吸机的功能已经发挥到了极限。德菲利波的体温已经上升到40摄氏度。在与他病情相仿的重症病人中，只有不到5%的人能够走出重症监护室。只要我们稍有闪失，这一点点的生存概率就会荡然无存。

在随后的10天里，情况渐渐好转起来。德菲利波的主要病因是手术过程造成的肝损伤。由于他肝内的胆总管被切开了，所以胆汁不断漏出来。而胆汁具有腐蚀性，它能消化食物中的脂肪，也能把人体内部各种脏器都消化掉。德菲利波的身体太虚弱了，经不起修复手术的折腾。所以，一旦他的病情稳定下来，我们就请放射科医生帮忙采取临时性措施。他们在X射线的帮助下，将引流支架穿过腹壁，并将其放入切开的胆总管，这样就能把胆汁引出来。由于病人体内漏出的胆汁太多，他们不得不放置3个引流支架，一个在胆总管里，两个在胆总管周围。胆汁被排出后，德菲利波的烧退了。他的输氧量和补液量都下降了，血压也恢复到了正常水平。他渐渐开始恢复。但到第11天，就在我们准备把机械式呼吸机撤掉的时候，德菲利波的体温骤然升高，血压骤降，血氧水平再次大幅下降。他的皮肤上都是汗水，身体却在不断地打冷战。

我们不知道到底发生了什么。虽然我们怀疑发生了感染，但X射线检查和CT扫描都找不出感染源在什么地方。尽管我们给德菲利波使用了4种抗生素，他还是高烧不退，最终导致其心脏发生了纤维性颤动。蓝色警报启动了。一大群医护人员闻讯赶来，对其实施电击除颤。这一举措奏效了，他的心跳又恢复到了正常的频率。又过了两天时间，我们才找出问题的原因。我们怀疑接入德菲利波体内的某根导管被感染了，所以他换上了新导管，并将撤下的导管送到实验室进行培养。48小时后，结果出来了，所有导管都发生了感染。很可能是一根导管在接入时被细菌污染了，随后细菌通过德菲利波的血液蔓延到其他导管。就这

样，所有导管都成了毒源，不断地将大量细菌引入德菲利波体内，使他高烧不退，病情恶化。

这就是重症监护室里的现实，我们能拯救病人的生命，也同样能威胁他们的生命。中心静脉置管感染非常普遍，我们也将其视为一种常见的并发症。

在全美各家医院的重症监护室里，医护人员每年要为病人插入500万根静脉置管。美国**2008**年全国性的统计数据显示，静脉置管插入**10**天后，就会有**4%**的病人发生感染。在美国，每年有**8**万人发生此类感染，死亡率为**5%~28%**。感染威胁性的大小取决于病人病情的严重程度，那些挺过去的病人平均要在重症监护室里多待一周时间。但这里的威胁还远不止这些。在美国，插入导尿管**10**天的重症监护病人有**4%**会发生膀胱感染，而接入呼吸机**10**天的病人有**6%**会发生细菌性肺炎，死亡率为**40%~45%**。总而言之，有将近一半的重症监护病人会发生严重并发症。此类情况一旦发生，他们的存活率将大幅降低。

又过了一周时间，德菲利波才从感染中恢复过来，我们这才把呼吸机撤掉。但他还要耐心等上两个月才能出院。德菲利波的身体在出院后依然非常虚弱，这让他失去了工作和家庭，他也不得不搬到姐姐那里生活。我们为他插入的胆汁引流管依然在他的腹部上挂着。当德菲利波恢复元气后，我们会对他实施手术，帮他重建胆总管。但不管怎样，他顽强地活了下来，而大多数病情和他一样严重的病人最终都没能挺过来。

超级专家也会犯错

这就是现代医疗面临的主要困惑：为了挽救垂死的病人，我们要掌握正确的知识，并确保每天对病人实施的**178**项治疗措施都不出错。请注意，在治疗过程中你会受到各种各样的干扰，比如一些监视器的警报会莫名其妙地响起，临床的病人可能会吵吵闹闹，还会有护士把头探过来问你是否可以帮把手，把病人的胸部打开。这些都会让原本已经非常复杂的任务变得更加复杂，即便是进行分工都未必能应付得过来。那么，我们到底该怎么办呢？

现代医疗对于上述问题的解决之道是专业分工，让每个人的专业领域变得越来越窄。在我讲述德菲利波的故事的时候，你可能会觉得好像是在无时无刻地照料他，是我在进行各类繁杂的操作。不，这些其实是由重症监护医师做的。作为一名普外科医生，我希望自己能够应付大多数临床状况。但是重症监护的繁杂和细致让我不得不把工作交给专科护理专家。欧美的大多数主要城市都开设了有关特殊护理的培训课程。

在如今的美国医院里，有一半重症监护室要依靠这些专科护理专家。

专业分工是现代医学的金科玉律。在20世纪初，人们只需要在高中毕业后读一年医学专业就可以行医了。而到了20世纪末，如果你想当医生，无论是在儿科、外科，还是在神经科单独执业，都必须完成4年的医学本科学习，然后还要接受3~7年的住院医师培训。不过，这些准备似乎已经不足以应对日益复杂的医疗了。如今，大多数年轻医生在结束住院医师培训后还要接受1~3年的专业进修培训，如有关腹腔镜手术、小儿代谢疾病诊治、乳腺癌放射学或重症监护等的培训。今天的年轻医生往往已经不再年轻。一般来说，他们要到34~35岁才能开始独立执业。

我们生活在一个由超级专家主导的时代。医生们在各自的狭小领域内不断磨炼自己的技艺，直到自己在这个领域比其他人干得更好。与一般专家相比，超级专家有两大优势：他们知道更多重要的细节，而且学会了如何掌控特定工作的复杂性。但无论是在医疗行业，还是在其他领域中，一些工作的复杂性远远超出个人可以掌控的范围，即使是最能干的超级专家也难免会犯错。

可能没有哪个领域的分工细化程度能够比得上外科手术了。我们可以把手术室看作操作极为频繁的重症监护室。在手术过程中，麻醉医生只负责消除病人的疼痛和稳定他们的体征。但麻醉医生可分为小儿麻醉医生、心脏麻醉医生、产科麻醉医生、神经手术麻醉医生等。相似地，“手术室护士”这一称谓已经名存实亡，因为他们的专业也要进一步细分。

清单宣言 | THE CHECKLIST MANIFESTO HOW TO GET THINGS RIGHT

无论是在医疗行业，还是在其他领域中，一些工作的复杂性远远超出个人可以掌控的范围，即使是最能干的超级专家也难免会犯错。

外科医生当然也要顺应潮流，其专业细分程度到了近乎荒唐的地步。我们经常开玩笑说，以后会有左耳外科专业和右耳外科专业。虽然我们是在开玩笑，但谁也保不准以后外科医生不会有看左耳的和看右耳的之分。

我是一名普外科医生，但是现在除了在非常偏远的农村地区，你根本找不到什么外科手术都能做的医生。于是，我决定将我的专业聚焦在肿瘤外科手术上，但这个目

标还是太大。所以，虽然我已经尽全力掌握各种普外科手术技术，特别是与急救相关的技术，但我还是不得不把目标集中在内分泌腺肿瘤切除手术上。

近几十年，专业分工细化程度的不断提升为外科手术技术的飞速发展做出了巨大贡献。以前，即使是小手术，死亡率也都高达两位数，而且病人往往恢复得很慢，常常会留下残疾。但在今天，手术后病人当天出院已经是家常便饭了。

不过，如今的手术数量也在迅猛增长。一个美国人在一生中平均要接受**7**次手术，而美国的外科医生们一年要做**5 000**万台手术。所以，手术伤害的绝对数量依然居高不下。在美国，每年有**25**万人没能走下手术台，不仅如此，许多研究都显示，至少有一半致死因素和严重并发症都是可以避免的。我们并不无知，但无论我们进行多么细致的专业分工，无论接受数量多么巨大的培训，还是会忽略一些关键的步骤，还是无法避免一些错误。

成功和失败并存的现代医疗向我们发出了严峻的挑战：如果专业分工都不足以解决问题，那该怎么办呢？如果超级专家都会失败，我们还有什么办法吗？渐渐地，我们看到了问题的答案，但它来自一个意想不到的领域，一个和医疗完全没有关系的领域。

清单宣言 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

我们并不无知，但无论我们进行多么细致的专业分工，无论接受数量多么巨大的培训，还是会忽略一些关键的步骤，还是无法避免一些错误。

第2章 “抓取关键点”比“大而全”更重要 系统要素的重塑

·一张小小的清单，让约翰斯·霍普金斯医院的中心静脉置管感染事件发生率从11%下降到了0，而且在15个月内避免了43起感染事件和8起死亡事故。

·清单从来都不是大而全的操作手册，而是人们通过理性选择关键内容后创造的思维工具。懂得在操作过程中抓取关键点，不仅仅确保了基准绩效，更是达成高绩效的保证。

坠落的“飞行堡垒”

1935年10月30日，在俄亥俄州代顿的莱特机场（Wright Air Field），一场非同寻常的比赛正在进行。比赛的主办方是美国陆军航空队，也就是美国空军的前身，而参赛“选手”则是波音公司和马丁与道格拉斯公司（Martin and Douglas）分别研制的下一代远程轰炸机。比赛的结果本不该有什么悬念，因为在第一阶段评估中，波音公司研制的299型铝合金机身轰炸机遥遥领先，把马丁与道格拉斯公司研制的飞机远远地甩在后面。波音公司的轰炸机的载弹量是军方招标要求的5倍，而飞行速度几乎是早先轰炸机的两倍。一位看过299型轰炸机试飞的西雅图新闻记者将其称为“飞行堡垒”，后来人们就一直这么称呼它。美国军事历史学家菲利普·迈林格（Phillip Meilinger）说这场比赛只是走走过场而已，军方准备向波音公司订购至少65架新型轰炸机。

一群高级军官和公司高管目送299型轰炸机滑向跑道。这架飞机的机身外表非常光滑，外形非常抢眼，它的翼展有31米，机翼下悬挂了4台发动机，而以前的轰炸机往往只配备2台发动机。只见试验机呼啸着冲向跑道的尽头，略一抬头便腾空而起，以大仰角迅速爬升至近100米的高度。但突然之间，飞机就像一个醉汉一样倒向一侧，随即失速坠地，发生巨大爆炸。5人机组中有2人不幸遇难，其中就包括试飞员普洛耶尔·希尔少校（Major Ployer Hill）。

调查结果显示，这起事故并不是机械故障引起的，而是因为人为失误造成的。这架飞机比以往的飞机复杂许多，飞行员要控制4台发动机，而且每台发动机的燃油混合比都有所不同。此外，飞行员还要操控起落架、襟翼、电动配平调整片（让飞机以不同速度飞行时保持稳定的装置）和恒速液压变距螺旋桨等。由于忙于各种操作，希尔少校忘记了一项简单却很重要的工作。研发人员为飞机设计了一套全新的控制面锁定机制，但希尔在起飞前忘记对升降舵和方向舵实施解锁了。当地的报纸认为波音公司的新飞机“太过复杂，以致单人无法操控”。最后，军方不得不选择马丁与道格拉斯公司设计的较小的轰炸机。波音公司几乎因此而破产。

不过，军方购买了几架波音299型轰炸机作为测试机。一些试飞专家依然坚信这一型号的飞机是可以操控的，所以他们聚到一起出谋划策。

这些试飞专家决定不做的事情和他们决定要做的事情一样有趣。他们并没有要求驾驶299型轰炸机的飞行员接受更长时间的培训，因为希尔少校的经验和技术已经是超一流的了，他是美国陆军航空兵的首席试

飞行员。这些专家想出了一个非常巧妙的办法，那就是编制一份飞行员检查清单。这张清单的诞生，从侧面说明航空业取得了多么惊人的发展。虽然驾驶早期的飞机会让人感到紧张，但那对于受过专业训练的飞行员来说不算难。在那个时候，使用清单开飞机就好比看着使用手册把汽车倒出车库一样荒唐可笑。但是，新飞机的操控复杂程度大幅提升，任何一个飞行员，无论他的记忆力有多好，飞行技术有多么高超，都不能保证自己可以在飞行的各个阶段准确无误地完成所有操作。

在编制清单的时候，试飞专家尽量做到简明扼要，他们把起飞、巡航、着陆和滑行各阶段的重要步骤浓缩在一张索引卡片上。卡片上列出的事项飞行员都知道该怎么操作，他们会根据清单的提示检查刹车是否松开，飞行仪表是否准确设定，机舱门窗是否完全关闭，还有升降舵等控制面是否已经解锁。这些对飞行员来说真的都是一些简单得不能再简单的操作了，你可能会因此而质疑清单的有效性，那么让我们看看清单的力量吧！

清单的力量 | THE CHECKLIST MANIFESTO HOW TO GET THINGS RIGHT

这一看似愚蠢的飞行清单投入使用以后，299型轰炸机的无事故安全飞行里程达到了290万千米。最终，美国军方总共订购了13 000多架这个型号的飞机，并将其命名为B-17。由于飞行员驯服了这头巨兽，盟军在第二次世界大战中取得了决定性的空中优势，并利用它们对纳粹德国实施了毁灭性的轰炸。

抓取“关键”要素与“必不可少”的基本要素

和当年的试飞专家一样，现在很多领域的专家，如软件设计师、财务经理、消防员、警官、律师，还有医生等，也面临着越来越复杂的工作，单凭记忆他们很难万无一失地完成自己的工作。也就是说，现在很多领域的工作任务已经成了不能由单人操作的“飞行堡垒”。

但是，人们还是很难相信像清单这么简单的东西真的会有用。人们承认自己在工作中会大意，会犯一些低级错误，甚至是毁灭性的低级错误。但是他们坚持认为自己的工作内容太复杂了，根本无法将其概

括到一张卡片上。

比如，现在疾病的种类有很多，远远超过了飞机的种类。一项对宾夕法尼亚州41 000名创伤病人进行的调查研究显示：光是对于不同的创伤，医生就需做出1 224种不同的诊断。结合不同病人的具体情况，这些诊断的组合有32 261种之多，这就好比有32 261种不同类型的飞机要降落。为每种飞机在各种条件下的降落过程编写正确的程序是不可能做到的。所以，医生们对于一张小小的清单到底能起到多大的作用从心底里表示怀疑。

即便如此，医生们也并非完全不采纳与清单相似的做法。比如，每家医院都会跟踪记录病人的4项重要体征，分别是体温、心率、血压和呼吸频率。这些数据能够让医护人员对病人的健康状况有一个基本了解。我们很清楚，这4项指标都具有重要的医学意义。比如，某位病人的3项指标都非常正常，看起来没有什么问题，你都认为他可以出院了，但第4项指标显示这名病人正在发烧，或者血压偏低，或者心跳过速，此时你忽视任何一个信号都可能会危及病人的生命。

在20世纪初，水银体温计开始广泛使用，而后俄国医生尼科拉·科罗托科夫（Nicolai Korotkoff）发明了通过使用充气袖套和听诊器来测量血压的方法。从那时起，医生就掌握了测量病人重要体征的方法。虽然与单独测量某项体征相比，测量4项体征数据更有利于医生准确掌握病人的病情，但是当时的医生并没有做到定期记录。

在复杂的环境中，人们往往要应对两大困难。第一种困难是人类记忆和注意力的谬误。在重压之下，人们特别容易忽视一些单调的例行事项。比如当一个病人呕吐不止，而他的家人不停地向医生询问病情时，医生很容易忘记为病人测量脉搏。对于被工程师称为“或全或无”的过程来说，记忆错误和注意力不集中是非常危险的。无论是去食品店买做蛋糕的原料，还是为飞机起飞做准备，或是在医院里诊断病人的病情，错过任何一条重要信息都会让你付出的努力彻底泡汤。

但是，第二种困难一样不可小觑，那就是人们会麻痹大意，会故意跳过一些明明记得的步骤。毕竟，在很多复杂的过程中，某些步骤并不总是那么重要。飞机的升降舵一般来说总是处于解锁状态，在大多数情况下，检查升降舵纯粹是在浪费时间。但与及时跟踪病人的4项体征相类似，或许这种做法就能在50个病人中发现一次险情。人们会一直说“以前从来就没出过这类问题”，直到真的发生了严重后果为止。

清单会提醒操作者不要忘记一些必要的步骤，并让他们明白该干什么。这不仅是一种自我检查方法，而且还是一种保障安全和效率的纪律。

清单似乎能帮助我们防范此类错误。它们会提醒操作者不要忘记一些必要的步骤，并让他们明白该干什么。这不仅是一种自我检查方法，而且是一种保障安全和效率的纪律。清单促使医生定期记录病人体征，这在病人的治疗过程中发挥了重要作用。当然，这并不是医生的功劳。



直到20世纪60年代，对病人的4项重要体征（体温、心率、血压和呼吸频率）进行例行记录才在西方医院里成为操作规范。而想出这个主意的是护士。他们自己编制了用来进行记录的表格，这无异于创造出一种清单。护士每天要为病人提供很多服务，如帮他们服药，为其伤口进行包扎换药，解决各种问题。而体征记录表能确保他们每6小时，或在必要的时刻测量病人的体温、心率、血压和呼吸频率，并对其整体健康状况进行评估。

在大多数医院里，护士还有一个记录项目，那就是病人的疼痛感，病人要在一个10级量表上对自己的疼痛感打分。护士们还进行了其他几项革新，他们设计了给药时间表和简要护理计划表。虽然没有人将这些表格称为清单，但它们实际上就是清单。

虽然护士们非常欢迎这些革新，但是医生们对相似举措却并不热衷。他们认为：表格和清单是护士把弄的玩意儿，都是些无聊的东西，接受更多专业训练的医生根本不需要这些东西。

让清单变成理性的选择

不过2001年，约翰斯·霍普金斯医院一位名叫彼得·普罗诺弗斯特（Peter Pronovost）的重症监护专家决定试验一种为医生准备的清单。他并没有把重症监护团队在一天中所要进行的所有操作都写进这张清单，而只是在数百种项目中挑选了一种进行试验，那就是与防止中心静脉置管感染相关的操作。正是这种感染差点夺去德菲利波的生命。



普罗诺弗斯特把防止插入中心静脉置管引发感染的步骤写在一张纸上，这些步骤分别是：（1）用消毒皂洗手消毒；（2）用氯己定消毒液对病人的皮肤进行消毒；（3）给病人的整个身体盖上无菌手术单；（4）戴上医用帽、医用口罩、无菌手套并穿上手术服；（5）待导管插入后在插入点贴上消毒纱布。

实施这些步骤并非什么难事，医学院这些年来一直就是这么教学生的。所以，给这些简单步骤编写清单听起来似乎没有必要，甚至还有点傻。不过，普罗诺弗斯特还是要求自己重症监护室里的护士对医生插入中心静脉置管的操作观察一个月时间，并记录他们实施上述步骤的情况。调查结果显示，1/3以上的操作不够规范，医生们总是至少跳过了一个步骤。

在随后的一个月里，普罗诺弗斯特和他的团队说服约翰斯·霍普金斯医院的管理层，授权护士在发现医生跳过清单上所列步骤的时候叫停操作。此外，护士每天还要询问医生是否有留置管需要移除以减少导管接入时间，避免不必要的感染。这些举措具有革命性的意义。虽然护士总是有办法提醒医生做正确的事情，他们会委婉地提醒医生，如“医生，你是不是忘记戴口罩了”；有时也会采取强力手段，一个护士用身体挡在医生面前，因为她觉得医生在病人身上盖的无菌手术单不够多，但很多护士还是不清楚自己该不该和医生对峙，也不清楚一些问题是否值得说出来，比如导管是插入病人

胸部的，那么有必要在他的腿上覆盖无菌手术单吗？幸好新规定说得很明确：如果医生不按照清单上列出的步骤操作，护士就能进行干预，这是管理层赋予护士的权利。

清单的力量 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

普罗诺弗斯特和他的同事在随后的一年中一直对清单使用情况进行跟踪。试验结果令人惊奇：在所有插入中心静脉置管10天的病人中，出现感染的病人比例从11%下降到了0。医院在15个月内只发生了两起置管感染事件。统计显示，在约翰斯·霍普金斯医院，清单的实施共防止了43起感染事件和8起死亡事故。

普罗诺弗斯特开始动员更多同事参加试验，他们还测试了其他几种清单。

清单的力量 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

有一种清单旨在督促护士至少每隔4小时观察一下病人的疼痛状况，并及时为病人提供止痛药物，这让病人长期忍受疼痛却得不到及时治疗的情况发生率从41%下降到3%。

一种为机械式呼吸机相关操作准备的清单，会提醒医生给病人服用抗胃酸药物，以防止病人患上胃溃疡；它还会提醒医护人员确保病床床头至少向上抬高30度，以防止病人的口腔分泌物进入气管。结果，不当处置率从70%下降到4%，病人的肺炎发生率下降了1/4。与上一年同期相比，重症监护室的死亡人数减少了21位。

研究人员发现，仅仅是要求重症监护室医护人员为日常操作编写清单，就能显著提高他们的绩效水平，并能让病人在重症监护室里的停留时间减少一半。

普罗诺弗斯特发现，除了上述效果之外，医院里的清单也可以在其他项目中发挥作用。它们帮助医护人员记住关键步骤，并且清晰地列出了操作过程中必不可少的基本步骤。普罗诺弗斯特惊奇地发现，即便是一些经验丰富的医务工作者也会忘记某些预防措施的重要性。在引入呼吸机检查清单之前对重症监护室医护人员所做的一项调查显示，虽然大量科学证据显示，给接入呼吸机的病人服用抗胃酸药物是非常重要的，但还是有一半数量的医护人员并没有意识到这一点。因此，普罗诺弗斯特认为，清单为基准绩效建立了更高的标准。

不过，这个法子看起来非常原始，非常不起眼。普罗诺弗斯特总是被同事们称为“聪明绝顶”“灵感泉涌”的“天才”，他在约翰斯·霍普金斯大学的公共卫生学院获得了医学博士和哲学博士学位，他所接受的专业训练包括急诊治疗、麻醉学和重症监护。但清单并不是什么新发明，难道真的要接受那么多高等教育才能想到运用这个简单的方法吗？或许，还真的需要。

虽然普罗诺弗斯特的初步试验取得了巨大成效，但重视这一举措的人寥寥无几。他到美国各地进行推广，任何一个有助于该清单推广的人他都不会放过，其中包括医生、护士、保险商，还有医院管理层。他平均每个月要跑7个城市，但鲜有采纳这一做法的人。

人们拒绝使用清单的理由有很多。一些医生觉得这一建议冒犯了他们的尊严，其他人则对普罗诺弗斯特的试验结果提出了合理的质疑。因为，试验结果只说明清单在一家医院获得了成功，这家医院就是他所在的约翰斯·霍普金斯医院。要知道，这家医院的重症监护室有充足的资金和人员保障，而且普罗诺弗斯特一直在监督清单的正确实施。但在现实世界中，按清单操作依然能取得这么好的效果吗？重症监护室的医护人员往往人手不足，他们时间紧，压力大，已经有很多烦人的表格要填写了，而现在你又要让他们填写一张清单，这实在是件非常困难的事情。

但在2003年，密歇根医疗与医院协会（Michigan Health and Hospital Association）与普罗诺弗斯特取得了联系，他们希望能够在密歇根州各家医院的重症监护室里试验他发明的中心静脉置管检查清单。这是一项规模浩大的试验，普罗诺弗斯特终于找到了验证清单是否能够在更多医院奏效的机会。

密歇根州在全州范围内的医院试验清单几年之后，我来到了底特律

的西奈-格雷斯医院（Sinai-Grace Hospital）。在那里，我切身体会到了普罗诺弗斯特所面临的严峻挑战。

这是一家典型的城镇医院，位于底特律西区八英里路（Eight Mile Road）的南侧。医院由一些红色砖房构成，周围有许多废弃的房屋、支票兑换店和假发店。当我访问那里的时候，这家医院有800名医生、700名护士和2 000名其他医务工作者，而他们的服务对象是全美收入最低的人群。底特律有超过25万居民没有医疗保险，有30万人接受密歇根州的医疗援助。这意味着医院面临着长期的财务问题。不过，底特律资金最紧张的医院并不是西奈-格雷斯医院，而是底特律收容医院（Detroit Receiving Hospital）。在这家医院收治的病人中，有1/5以上的人无力支付医药费。在2000—2003年间，西奈-格雷斯医院等9家底特律医院被迫裁掉了1/3的员工，州政府费尽周折凑了5 000万美元才避免这些医院破产。

西奈-格雷斯医院有5个成人重症监护室和1个婴儿重症监护室。重症监护室主任哈桑·麦基（Hassan Makki）向我介绍了2004年普罗诺弗斯特和密歇根医疗与医院协会刚刚开始推广清单时医院的情况。他说：“员工士气低落，很多护士离职了，即便留下来的护士也不确定自己以后会不会走。”很多医生也在考虑要不要跳槽。与此同时，大家的工作压力却有增无减，因为新的规定严格限制了住院医生连续工作的时间。而就在这样一个时刻，普罗诺弗斯特却给医护人员增加工作负担，要他们抽时间填写日常检查清单。

汤姆·皮斯科罗斯基（Tom Piskorowski）是一位重症监护室医生，他当时听到消息后的第一反应是：“把文案工作忘掉吧，我们真正应该关心的是病人。”

我早上7点钟跟随医疗团队来到一个重症监护室，那里有11个病人。其中4人被枪械击伤：1人胸部中弹，1人的肠子、肾脏和肝脏被打穿，还有2人颈部中弹，因此四肢瘫痪。有5个病人脑出血：3个病人的年龄超过79岁，他们下楼的时候因不慎跌倒而引起脑出血，一个中年男子的颅骨和左侧颞叶被歹徒用钝器击伤，还有一个工人从8米高的梯子上摔下来，头部着地，自颈部以下全身瘫痪。另外，一个癌症患者刚刚做完肺叶部分切除手术，一个病人刚刚做完颅内动脉瘤切除手术。

值班的医护人员希望能够挨个顺利地各项操作，但总有事情不停地打断他们：一个病情似乎已经稳定下来的病人又开始大出血；一个刚刚拔掉呼吸导管的病人呼吸困难，医护人员不得不为他重新接上呼吸机。此类紧急状况总是不断发生，很难想象他们还能抽出时间来考虑填写什么清单。

但我发现他们真的做到了。

在大多数情况下，负责这项工作的是护士。每天早上，一个手里拿着写字板的高级护士会查房，她会检查每个接入呼吸机的病人其病床床头抬起的角度是否正确，确认每个病人是否服用了正确的药物、接受了正确的检查。每当医生要为病人插入中心静脉置管的时候，护士会根据清单进行检查，在填写完清单后将其放入病人的记录档案内。我还查阅了医院的档案，结果发现在2006—2008年，他们一直认真地实施这一举措。

普罗诺弗斯特是个聪明的人。在清单推广初期，他并没有要求医院管理层立刻使用中心静脉置管清单，而是请他们收集医院中心静脉置管感染率的相关数据。他们发现，在2004年初，密歇根州各个医院重症监护室的感染率高于美国平均水平，有些医院的感染率高得出奇。西奈—格雷斯医院的感染率高于美国75%的医院。与此同时，密歇根州的蓝十字蓝盾协会（Blue Cross Blue Shield）同意为每家参与普罗诺弗斯特推广计划的医院颁发一小笔奖金。突然之间，试验清单变成了理性的选择。

于是，“基石行动”（Keystone Initiative）开始了。每家参与行动的医院要委任一名项目经理专门负责清单的推广，这位负责人每两个月还要和普罗诺弗斯特碰一次头，共同商讨如何解决项目开展过程中碰到的问题。普罗诺弗斯特还坚持要求每家医院指派一名高层领导分管这项工作，并要求其每月至少下基层一次，倾听员工的意见和建议，协助项目经理解决问题。

医院高管有些不情愿。办公室和会议室是他们的主战场，他们关心的是医院的战略问题和财务问题，而不是冒险到病区和大家打交道。虽然一些医院的高层在下基层的时候碰了钉子，但事实说明，他们的参与对于项目的推广至关重要。

在行动开始的第一个月里，管理层发现能够有效减少导管感染的氯己定消毒皂配备不足，只有不到1/3的重症监护室配备了这种消毒皂。此类问题只有医院管理层才能解决。接下来在数周时间里，密歇根州的每一个重症监护室都配备了氯己定消毒皂。医护人员还向管理层反映，虽然清单要求医生在为病人插入中心静脉置管时要给病人全身覆盖无菌手术单，但全尺寸的手术单不那么容易获得，管理层随即增加了此型手术单的订购量。他们还说服中心静脉置管的最大生产商之一箭牌国际公司（Arrow International）设计生产全新的导管套件，这种新套件将无菌手术单和氯己定消毒皂也包括在内。

2006年12月，《新英格兰医学杂志》（*New England Journal of Medicine*）刊登了一篇里程碑式的文章，公布了“基石行动”所获得的丰硕成果。

“基石行动”开展3个月后，密歇根州重症监护室由插入中心静脉置管而引起的感染事件发生率下降了66%。大多数重症监护室，其中包括西奈-格雷斯医院的重症监护室病房的季度感染率下降为零。密歇根州的平均感染率显著下降，在全美名列前茅。

项目实施18个月后，参与行动的医院总共节省了1.75亿美元，挽救了1 500多人的生命。

这项行动取得的辉煌成就已经延续了好几年，这都要归功于一张看似多此一举的清单。

抓住关键，就抓住了一线生机

不过，人们很容易认为这一成功只是个例。预防中心静脉置管感染的举措可能非常特殊，毕竟这一清单并不能防止其他各种因插入导管而引发的并发症。如果针头插入过深会引发气胸，如果针头扎破血管又会引发出血。普罗诺弗斯特的清单只能预防感染。没错，在这个例子中，医生的确会忘记一些最基本的步骤，如洗手、戴无菌手套并穿上手术服等。对于此类失误，清单的确非常有效，但医生进行的其他各种操作是否还具有这种特点就很难说了。

我开始思考这个问题。

在得知普罗诺弗斯特取得的非凡成就之后，我咨询了马库斯·塔尔曼（Markus Thalmann），他就是撰写那篇溺水女孩奇迹生还论文的心外科医生。让我感到好奇的是，这一系列复杂操作不是发生在一家先进的大型医学研究中心里，而是在一家普通的社区医院里。



这家医院位于奥地利阿尔卑斯山区一个名叫克拉根福（Klagenfurt）的小镇，离女孩溺水的地点很近。我让塔尔曼告诉我如何才能确保正确地实施那么多复杂的操作步骤。

塔尔曼说，那个女孩被送来的时候，他已经在这家医院工作6年了。这不是第一个被送来时就已经体温过低而且窒息的病人，他们医院每年都会接收3~5名这样的病人，大多数人几乎已经进了鬼门关，其中有一些是不幸溺水，还有一些则是服药自杀，晃悠悠倒在阿尔卑斯山的冰天雪地之中。在很长一段时间里，他们无论怎么努力都无法救活病人。大多数病人在被发现时心跳和呼吸已经停止很长时间了。尽管如此，塔尔曼还是觉得一些病人仍然有一线生机，但他和同事还是没能抓住。

他认真研究了所有病历，发现准备工作是最大的困难。要想挽救病人的生命，就必须在很短的时间里让许多医护人员就位，并且准备好一堆设备。医疗团队中必须有创伤外科医生、心脏麻醉医生、体外循环灌注师、手术和重症监护护士以及重症监护医师，但总是有人或设备不能及时就位。

他试图使用常规方法来解决问题，他对每个人大喊大叫，试图让他们加快行动速度，但还是没有病人能够生还。于是，他和一些同事决定尝试一些新方法。最终，他们编写出了一份清单。

在整个救援过程中，急救队和医院接线员的权力级别最低，但塔尔曼却把清单交给了他们，并为他们详细介绍各个环节。就拿挽救落水女孩的这个例子来说，急救队按照清单通知医院准备好人工心肺机和复温操作。只要有可能，急救队会在到达医院之前发出通知，医院准备时间非常紧张，但准备是否充分对病人能否生还至关重要。随后，医院的接线员会列出需要通知的医护人员名单，并通知他们尽快做好各种准备。

清单投入使用后，塔尔曼及其同事第一次成功挽救的这类危重病人，就是那个3岁的小女孩。此后不久，塔尔曼去了维也纳的一家医院供职，而他原来的同事又成功挽救了至少两个危重病人的生命。其中一个病人试图自杀，被发现时已经冻僵了。而另一个病人是个16岁的女孩，她的母亲不慎把车开下悬崖，坠入冰冷的河水之中。母亲当场身亡，而女孩来不及爬出车外就随着汽车沉入水底。当急救人员到达时，她的呼吸和心跳已经停止很久了。

但当女孩被救出后，一切救援步骤都有序进行。当急救人员开始为女孩进行心肺复苏时，医院已经得到了通知。几分钟后，女孩被送到医院。外科团队把她直接送到了手术室，并迅速为其接上了人工心肺机。就这样，一步接一步依次进行。由于抢救及时，女孩有了生还的希望。

当女孩的体温逐渐回升时，她的心脏恢复了跳动。在重症监护室里，接入呼吸机、补液和静脉药物注射帮助她渐渐恢复。第二天，医生就把各种导管拔掉了。再过一天，女孩就能坐起来，可以出院了。

第3章 团队犯错的概率比单个人要小 智慧的差别

·美国每年发生的严重建筑事故只有20起，这意味着建筑行业每年的严重事故发生率不到0.002%。面对复杂的摩天建筑，他们是如何做到的？

·团队的力量是巨大的。它可使我们不再依赖单枪匹马，不再听命于唯我独尊的大师，而是依靠团队的智慧。一个人免不了会犯错，一群人犯错的可能性会变得小一些。

当简单的个人“强制函数”不再奏效时

清单诞生以来，已经为4代人所使用。它的使用范围比我们想象的似乎要广得多，哪怕是经验丰富的人也能使用它来避免失误。清单为我们提供了一种认知防护网，能够抓住每个人生来就有的认知缺陷，如记忆不完整或注意力不集中。正是由于清单的强大功能，它常常会给我们带来惊喜。

但没有什么东西是万能的，清单也不例外。所以在使用之前，我们必须搞清楚清单在哪些情况下能帮助我们，而在哪些情况下没有什么用处。

加拿大约克大学的布伦达·齐默曼（Brenda Zimmerman）和多伦多大学的肖洛姆·格鲁伯曼（Sholom Glouberman）是两位专门研究复杂性科学的教授。他们提出了一种理论，将世界上的问题分为三类：简单问题、复杂问题和极端复杂的问题。

简单问题是那些具有明确解决方法的问题。例如如何烘焙蛋糕。对于这类问题，你可能需要学习一些基本技巧，一旦掌握了这些技巧，成功的可能性就会非常大。

复杂问题类似于如何把火箭发射到月球上的问题。虽然有时候你可以将其分解为一系列简单问题，但是无法找到直接的解决方案。一般来说，这类问题需要掌握不同专业技能的人组成团队、通力合作才能成功解决。时机和协调是成功的关键因素。

极端复杂的问题类似于培养子女这类问题。一旦成功地将火箭发射到了月球上，那么在发射其他火箭的时候就能重复和完善这一过程。毕竟，火箭就是火箭，它们之间有很大的相似性。但培养子女就不同了，因为每个孩子都是独一无二的。虽然成功培养一个孩子能让你积累经验，但这并不意味着你在培养下一个孩子的时候也能成功。专业技术固然重要，但不是获得成功的充分条件。培养不同孩子的方法可能截然不同。这体现出极端复杂问题的一大特性：结果的不确定性非常大。虽然我们知道有可能成功培养孩子，但是这一过程极其复杂。

清单为我们提供了一种认知防护网，能够抓住每个人生来就有的认知缺陷，如记忆不完整或注意力不集中。

虽然很多问题看似非常不同，但它们本质上非常相似。1935年，试飞专家要想办法避免飞机坠毁；2003年，医护人员要想办法避免病人因中心静脉置管而感染；而在不久之前，医生要想办法挽救一个3岁溺水女孩的生命。影响这些问题的因素虽然非常不同，但它们的核心都是一个简单的问题，那就是人们的注意力是否集中。

对第一个问题来说，飞行员需要集中注意力给飞行控制面解锁；对第二个问题来说，医护人员需要集中注意力使中心静脉置管保持无菌状态；而对第三个问题来说，外科小组需要集中注意力准备为女孩接上人工心肺机。

这些问题都是可以解决的，只要我们使用被工程师称作“强制函数”的方法，即用相对简单而直接的方法来迫使关键群体必要行为的发生，如使用清单来塑造行为。

从事不同工作的人都可能会被大量简单问题所包围。医生可能会在为病人插入中心静脉置管时忘记戴口罩，而在病人出现心跳停止时忘记这可能是由血钾水平过高引起的。律师在为逃税的犯罪嫌疑人辩护时可能会忘记重要的辩护策略，或者记不清递交各种法律文书的截止日期。当然，警察在工作中同样会犯低级错误，比如在让目击证人指证犯罪嫌疑人时忘记告诉证人犯罪嫌疑人可能不在队列之中，或者让知道犯罪嫌疑人是谁的人员出现在指认现场无意间暗示证人。而清单能够防止此类低级错误的出现。

但是，人们所做的大多数重要工作没有那么简单。要知道，插入中心静脉置管只是重症监护室医护人员每天所要进行的178项操作中的一种，他们的工作极端复杂。我们真的能够为每种工作都编制出清单，并在工作中按部就班地使用吗？这种想法是不是太不切实际了？我们很难找到一个简单明了的方法来照料重症监护室的病人。我们需要不同专业的专家在不同的情况下通力合作，以完成不同的任务，而这些复杂的任务很难用简单的“强制函数”加以控制。

而且人和火箭可不一样，人与人的差异非常大。没有两个肺炎患者是完全相同的。即使他们被相同的细菌感染，即使他们都咳嗽、气短、缺氧，即使医生让他们服用了相同的抗生素，两个人的恢复情况也相差甚远。医生必须处置难以预料的突发事件，很难在事前就准备好处理清单，因为这些事件是不可预测的。医疗过程包含了处理所有出现的三大类问题，其中既有简单问题，也有复杂问题和极端复杂问题。医生常常要随机应变，打破现有程序的条条框框，从病人的具体病情出发，为其着想。

对于上面提到的问题，我已经思考了很久。要想当一个好医生，必须弄明白什么时候应该相信自己的判断，什么时候应该严格遵守既定规程，这对于处理各种艰巨的工作也是非常重要的。你当然要把简单的事情做好，不要犯低级错误，但也要为随机应变和主观判断留出足够的空间。清单对于简单问题的价值是不言而喻的，但是对于复杂的和极端复杂的问题，它们是否同样有效呢？

我碰巧找到了这个问题的答案。

抛弃大师单枪匹马的工作体系

那是2007年1月的一个早上，天气十分晴朗，我开车去医院上班。停好车后，我沿着一条便道走向医院的大门。这时，我注意到旁边有一幢正在建造的大楼，那是我们的新医疗中心。大楼的外立面当时还没有建好，只能看到一个由大型钢梁撑起来的骨架。但是整个建筑已经有11层楼高了，占据了整个街区。我驻足抬头仰望，只见在4层楼的钢梁上有一个建筑工人正在焊接一个焊点。这时，我忍不住心想：这个工人和他的工友怎么知道自己在建造大楼的时候没有犯错呢？他们怎么确保大楼不会倒塌呢？

清单宣言 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

要把简单的事情做好，不要犯低级错误，也要为随机应变和主观判断留出足够的空间。

这幢大楼不是很大。建好以后，它能为我们医院提供150个单人病房和16个超级先进的手术室。除此以外，这幢大楼真的没有什么特别之处。我可以和你打赌，在美国2007年新建的大楼中，比它规模大的有几十座。

即便如此，这也不是一项轻松的工程。后来，医院的不动产经理告诉我，这幢大楼的总面积达到3.3万平方米，地上11层，地下还有3层，总造价高达3.6亿美元，需要消耗3 885吨钢铁、1万米钢筋混凝土管，共配备19台通风设备、16部电梯、1个冷却塔和1个备用发电机。在建造过程中，建筑工人要挖出7.6万立方米土石，铺设19.5千米长的铜管、75千米长的排水管和153千米长的电线，这些管子和电线的长度加起来足以从这里到达缅因州。

我暗自祈祷，它可千万不能倒下来。

要知道，这么高的大楼就算是碰到了地震也不能倒。我很好奇，建筑工人凭什么确定一切都万无一失呢？我意识到这个问题有两个层面：首先，他们怎么知道自己掌握了相关的知识？其次，他们怎么知道自己正确运用了这些知识？要解决上面提到的两个问题可不容易。在设计大楼的时候，专家必须考虑各种因素，如当地土壤的成分、单个结构的理想高度、可用材料的强度及建筑的几何外形等。把计划变成现实同样是困难重重，专家们必须确保各种工人和机械依次正确完成各个步骤，此外还要随机应变，处理那些意想不到的困难和变化。但是，这些建设者显然获得了巨大的成功，他们已经在全世界建造了几百万座大楼。虽然近几十年来建筑工程变得越来越复杂，但是困难并不能阻止他们获得更多成功。而且，和医生、教师等行业的从业人员一样，一线建筑工人也有细致的分工，他们有的专门负责打桩，有的则专门负责给各分区布线。

为了解决我心中的疑问，我决定拜访乔·萨尔维亚（Joe Salvia）。萨尔维亚是负责我们新大楼建设的结构工程师，我想让他告诉我建筑行业是怎么完成如此复杂的工程的。事实证明我找对人了。

自20世纪60年代以来，波士顿大多数主要的医院建筑都是由他创办的麦克纳马拉 / 萨尔维亚（McNamara/Salvia）公司负责设计工程结构的，波士顿的很多宾馆、办公楼和公寓楼也是他们负责建造的。他们还重建了波士顿红袜队（Boston Red Sox）的主场芬威球场（Fenway Park），该球场设有36 000个座位。这一项目重建的还包括球场具有标志性的、被称为“绿怪兽”（Green Monster）的左野墙，这堵墙高达11.3米，让球手很难在左野打出本垒打。萨尔维亚公司最在行的就是对大型而复杂的摩天大楼进行工程结构设计。

该公司设计的最高的大楼是迈阿密一幢高达80层的摩天大楼。在罗得岛的普罗维登斯（Providence），他们负责一家购物中心的结构设计，这个建筑的用钢量创造了东海岸之最（总共使用了2.4万吨钢）。他们还参与了一个大型项目，这个项目将建造的可能是全球规模最大的体育娱乐建筑群，即位于新泽西州东卢瑟福的梅兰多兹胜境（Meadowlands Xanadu）。这一建筑群包括纽约巨人队（New York

Giants）和纽约喷气机队（New York Jets）的体育场，还包括一座设有3000个座位的音乐厅、全美最大的多放映厅影院，以及美国第一家室内滑雪场冰雪公园（Snow Park）。在过去几十年的大多数时间里，这家公司的工程师每年要完成50~60个项目，平均每周就要完成一幢新楼的结构设计和建造，而其中没有一幢出现过任何可能要倒塌的迹象。

在萨尔维亚位于波士顿市中心的办公室里，我问他如何确保他们的设计和施工是准确无误的。萨尔维亚61岁，头发几乎都掉光了，说话带有一口浓重的波士顿口音，他性格随和，不会让人感到紧张，一点也不像我印象中的工程师。萨尔维亚告诉我，他设计的第一个项目是一家小型购物广场的屋顶。

那时候他刚刚大学毕业，是一个来自美国坎布里奇的23岁小伙子。他出生和成长的地方可不是干净整洁的城市公寓。他的父亲是个维修工，母亲在一家肉类加工厂当工人。不过萨尔维亚的学习成绩非常好，他成了家族里第一个上大学的人。他去了塔夫茨大学，就读医学专业，但一堂有机化学课改变了他的人生。

萨尔维亚告诉我：“有一次老师说，‘你们要记住这些公式。’我回应说，‘既然可以在书里找到这些公式，那么为什么一定要记住它们呢？’老师说，‘你想当医生吗？如果想的话就必须记住这些，医科学生要记住一切。’在我看来，这实在是很荒唐。而且我也不擅长记忆，于是我转系了。”

但是他非常善于解决复杂问题。他试图向我解释如何心算二次方程的解，但我脑子里盘旋的却是以前从来没有听别人用波士顿口音说过“二次方程”这几个字。他说：“我还非常喜欢创造。”所以，他从医学院来到了工程学院。工程学既是一门科学，又是一门务实的学问，他非常热爱这个专业。他学习了静力学和动力学，还学习了钢铁、混凝土和土壤的化学物理特性。

刚刚拿到学士学位的萨尔维亚加盟了一家名为萨姆纳·沙恩（Sumner Shane）的建筑工程公司。这家公司的业务主要是建造购物中心，他们刚刚在得克萨斯接到了一单生意，萨尔维亚负责屋顶系统的设计。他发现教科书曾教给他许多关于如何建造牢固屋顶的知识，而且详细的建筑施工法规也为他提供了许多信息。

萨尔维亚说：“我在大学里就学过如何使用钢结构，也就是用钢柱和钢梁来进行设计。”当地的建筑法规列出了很多具体规定，如钢的强度、土壤的组成、建筑的积雪承压能力、抗风压能力和抗地震能力等。

他要做的就是根据这些要求来设计新购物中心的具体细节，如建筑的大小、层数以及仓库和装卸平台的位置。萨尔维亚边说边在纸上描绘他的设计，他先画了一个长方形，然后画出了墙壁、出入口和过道，购物中心的样子渐渐浮现在纸上。

“你要画出可以承受屋顶重量的支撑点网格，”萨尔维亚边说边在可以安放立柱的地方打上小叉，“剩下的工作就是解方程了，要解出 x 。”要根据屋顶的面积和厚度算出它的重量，如果打算每隔10米设置一根立柱，那么就要算出每根立柱的直径和承重强度。你必须仔细检验自己的计算以确保工程质量符合所有要求。

这些知识都是在学校里学的。但是，萨尔维亚也发现，还有很多东西是他在学校里学不到的。

他说：“你可以做出几何学上的最佳设计方案，但在实践中，这种方案或许并不可行。”比如，你要考虑成本。但他当时设计的时候并不知道这一点。所以，他设计的屋顶大小和用料增加了项目的成本。当然，设计师还要考虑美学因素，没有客户会希望在楼面的中央矗立着一根立柱。有时候，一些特殊的视线也不允许被阻挡。

萨尔维亚说：“如果工程师说了算的话，那么所有建筑都会是长方形的盒子。”但是每个建筑大小不一，形状也不相同，都是独一无二的，是复杂的，你在教科书里找不到标准化的方程来解决以前从未遇到过的问题。就拿萨尔维亚创立自己的公司后所负责的一个项目来说吧。

这个项目是由建筑师菲利普·约翰逊（Philip Johnson）设计的波士顿地标性建筑——高46层的钢结构塔楼国际广场（International Place）。这个建筑非同寻常，它的外形是圆柱体和长方体的结合，以前不曾有人将这样的形状用于摩天大楼。萨尔维亚解释说，从结构工程学的角度来看，圆柱体是有问题的。建筑在风中或在地震中要抵抗得住自身弯曲的倾向，而方形的坚固程度比圆形高出60%。如果用圆柱体的话，难度就会大幅度上升。为了实现约翰逊的美学设计，萨尔维亚和他的团队必须对建造技术和过程进行创新。

萨尔维亚第一次设计的那个屋顶要简单许多，但对当时的他来说，那已经是困难重重了。除了要考虑成本和美学以外，他还要考虑其他系统的设计要求。参与设计工作的还有管路工程师、电气工程师、机械工程师，他们希望铺设管路、电线和设置空调通风系统的地方，可能正是萨尔维亚想要放置立柱的地方。

萨尔维亚说：“建筑就像身体一样。”它有皮肤，有骨架，还有血管

系统（管路）、呼吸系统（通风）和神经系统（电线）。所以，如今的建筑项目总共要涉及16个行当。萨尔维亚拿出一份施工计划，这份计划是为他们正在建造的一幢约120米高的摩天大楼编制的。他把计划翻到目录部分指给我看。

每个行当都分管一部分工作，如运输系统（电梯和自动扶梯）、机械系统（供热、通风、管路、空调、防火）、砖石结构、混凝土结构、金属结构、电气系统、门窗、温度和湿度控制系统（包括防水和隔热）的安装，粗木工和精木工，现场管理（挖掘、废水和雨水收集、走道保洁）。其他更加细致的工作还有地毯铺设、内部装饰、景观美化和灭鼠等。

这些工作缺一不可，但是每种工作又要符合大局。此外，各种工作必须互相协调，有序而精确地展开。从表面上来看，建筑工程的复杂性非常高。为了应对这么高的复杂性，整个行业不得不进行革新。

萨尔维亚告诉我，自中世纪以后，人们想建造楼房，可以去找建造大师。建造大师负责建筑物的外形设计、工程结构设计，并监督整个建造过程，事无巨细，一切都在他的掌控之中。建造大师们的代表作有巴黎圣母院、圣彼得大教堂和美国国会大厦。但到20世纪中叶，这个行当彻底终结，建造大师们从历史的舞台上消失了。这是因为，建筑过程每个阶段的复杂性和多样性都超出了个人能力的极限。

建筑设计和工程设计首先分道扬镳了。然后，每个行当进一步细分，直到一边是分工明确的设计师，而另一边是具有不同专业技能的工程师。当然，施工过程也进行了细分，比如塔吊承包商和精木工就不熟悉对方的工作。换言之，建筑行业 and 医疗行业一样，也经历了不断分工的演进过程。

实际上，医疗体系还沿用着“建造大师时代”建立的工作体系。在这一体系中，大师级医生或单枪匹马，或带领一群小跟班主导整个诊治过程。调查显示，**2007**年有**1/3**的病人在接受治疗的过程中，有**10**个以上的专科医生参与其中，参与的其他医务人员可能就更多了，其中包括护士、药剂师助理，还有义工等。已经有大量证据显示，我们适应现实的速度很慢，很多诊治措施是重复的、有瑕疵的，或者是没有经过统一协调的。

无论在设计时有多么困难，
都没有犯错的余地。因为错误
意味着很多人要送命，意味着
公司要损失大笔金钱。

萨尔维亚解释说，在建筑行业，这种错误是不能容忍的。虽然在设计第一个屋顶系统的时候觉得非常困难，但他很快意识到，自己没有犯错的余地。因为错误意味着很多人要送命，意味着公司要损失大笔金钱。早在20世纪初，建筑设计师、工程师和建筑工人就开始正视建造大师体系不再有效这一现实。所以，他们抛弃了这个体系，并想出新办法来确保工程不会出任何问题。

组织一个井然有序的团队

为了让我对如今的建筑行业有个直观的了解，萨尔维亚带我去他们附近的一个工地参观。他们的这项工程是名为“俄罗斯码头”（Russia Wharf）的32层办公楼公寓，建筑总面积达6.5万平方米，占地8 000平方米。

整个建筑群非常壮观。俄罗斯码头原来是用来停靠商船的，这些商船来往于圣彼得堡和波士顿之间，并为波士顿原来的造船业运送铁、粗麻和帆布等原材料。波士顿倾茶事件就发生在这附近，而新建的玻璃幕墙钢结构建筑将矗立在俄罗斯码头的原址之上。大楼内的天井有10层楼高，而且新建筑还将保留原始建筑有着110年历史的新古典立面，使其成为新楼的一部分。

当我们来到工地后，萨尔维亚看了看我的布克兄弟（Brooks Brothers）蓝色休闲西装和黑色平底便鞋，不禁低声轻笑：“在工地，穿合适的鞋子是非常重要的。”



旧建筑的内部结构很早就拆除干净了，新建筑的钢铁骨架也已经建造了一半，有14层楼那么高。一台塔吊高高在上，比新建筑的现有高度还要高4层。如果你从那里往下看，我们一定像蚂蚁一样小。我跟着萨尔维亚穿过一些混凝土搅拌车，跨过一些小泥坑，最终进入了约翰·莫里亚蒂建筑公司（John Moriarty and Associates）的一楼现场办公室。这家公司是这个项目的主承包商，那里和我在电影里看到的脏兮兮、破破烂烂的工地办公室完全不一样，也看不到叼着香烟、冲下属大吼大叫发号施令的工头。我眼前的办公室由6个房间组成，里面的工作人员有很多都穿着工作靴、牛仔裤和安全反光背心。他们有的端坐在办公桌前，盯着眼前的计算机屏幕，有的则聚在会议桌旁，就投影银幕上的报告或方案进行讨论。

他们给了我一顶蓝色的安全帽，然后让我签了一份保单。萨尔维亚把芬·奥沙利文（Finn O’Sullivan）介绍给我。这个身高一米九的爱尔兰人是这里的项目经理，他面带微笑，说起话带着一口浓重的爱尔兰口音，声调抑扬顿挫。他们现在不再使用“工头”这个称谓了。奥沙利文告诉我，他每天都要管理200~500名工人，这些工人来自6个不同的分包商。让我感到惊讶的是，他所掌握的知识的广度和复杂程度丝毫不亚于任何一个艰涩复杂的医学领域的知识。他努力向我解释自己和同事们如何确保所有工人的操作正确无误，确保整个施工过程没有瑕疵。要知道，他们需要考虑各个方面的因素，而且奥沙利文不可能熟知大多数具体工作的细节。虽然他讲得很认真，但我还是丈二和尚摸不着头脑。当他把我带进主会议室的时候，出人意料的一幕浮现在我眼前，在大型白色椭圆形会议桌周围的墙上，贴着很多砧板大小的清单。

贴在我右手边墙上的是施工日程安排表。我凑近仔细看了看，上面把每天需要完成的各项任务都标注了出来，而且所有工作的顺序和完成时间也一目了然。比如，第15层楼的混凝土浇筑工作在当月13日完成，而钢结构在当月14日运输到工地。

这份日程安排表要用好几张纸才能打印完整，而且需要彩打，因为在表上，不同的颜色表示不同的含义。比如，用红色标注的是重要步骤，在工人开始其他工作之前必须完成这一步。每当一项工作完成的时候，监工会向奥沙利文报告，而他会在自己的项目管理程序中给这项工作打上一个钩。奥沙利文每周都会将下一阶段的日程安排打印出来。当然，如果工程进展顺利，他打印日程安排的频率会加快，施工日程安排表实际上就是一张长长的清单。

由于每个建筑都有自己的特点，所以他们每次都要制订新的清单。在制订清单的过程中，每个行当都会派自己的代表前来参加，而萨尔维亚的公司会派出代表以确保建筑的工程结构没有问题，并符合其他各方的要求。清单完成后，会被送到各分包商和其他独立专家的手中再一次接受检查。这么做是为了确保不忽略重要环节。

就这样，详尽的清单陆续出炉，用来指导施工建设的具体开展，而且这个方法可以确保数以百计乃至千计的专业人士能够在正确的时间、正确的地点以正确的方式贡献自己的宝贵知识。

俄罗斯码头项目的施工日程安排是一个多层次的施工计划。当我到工地实地参观的时候，真的能看到施工过程的不同层次。带我参观的是萨尔维亚公司为这个项目委派的首席结构工程师伯尼·鲁伊拉德（Bernie Rouillard）。我事先声明，我不喜欢去太高的地方。但我还是戴上安全帽，跟着鲁伊拉德走过写有“闲人免入”的标示牌，跨过一堆废弃的生锈钢筋，沿着一排木板走道进入一台橙色的升降机前。待大门关闭后，升降机开始爬升，发出“咔嚓咔嚓”的声音。我们最后到达了第14层的楼面。走出升降机后我发现，厚厚的混凝土楼板上空空如也，整个楼面没有外墙，只有一些3米多高的钢柱围在四周，楼层的中央是呈长方形的巨型混凝土内芯。不用走到楼层的边上，就能轻轻松松地把波士顿的繁华景象尽收眼底。

“你可以从这里一览整个工程的全貌。”鲁伊拉德边说边示意我跟着他往楼层边缘走。我小心翼翼地挪到离边缘还有一米的地方停了下来。鲁伊拉德好心地把旁边的各项工程指给我看，而我实在没心思仔细端详，因为大风正“呼呼”地从我身边吹过，而距离地面十几层楼的高度也让我觉得头晕目眩。当我们把身体转向楼面中央的时候，我感觉好多了。鲁伊拉德让我看铺设在天花板上的裸露钢结构桁架，这些结构是用来支撑上一层楼板的。

鲁伊拉德说，他们随后会在这些桁架上喷耐火材料。

“你们要对钢结构进行防火处理？”我问道。

他回答“是的”，因为钢铁在大火中会塑化、变软，会像面条一样发生弯曲，这就是世贸中心倒塌的原因。他带我沿着楼梯下到第13层，我看到这里的桁架已经喷上了耐火材料。这种材料含有石膏成分，它们让天花板看起来灰灰的、毛茸茸的。

我们又往下走了几层。我看到建筑的“皮肤”，也就是玻璃幕墙和钢结构外立面已经安在这几层的混凝土楼板上。我们越往下走，建筑的层次就越多。分包商的一个团队已经在下面这些楼层的“皮肤”内砌起了分隔不同房间的墙壁，而排管工人则在铺设输水管和排水管。随后安装的是通风管路。当我们下到底层的时候，很多工序已经完工了，大理石内饰做好了，电线、水管都铺设好了，就连楼梯的扶手栏杆都已经安装完毕。这么繁杂的工程却进行得如此有序，这让我对萨尔维亚这样的建筑专家感到由衷敬佩。

团队沟通的智慧

还在较高楼层的时候，我发现了一个问题。即使是我这个门外汉也能察觉出有些事情不太对劲儿。那几天阴雨不断，而在那几层敞开的楼面上，有大量雨水积在靠近混凝土内芯的楼板上，就好像楼面向内倾斜一样。我向鲁伊拉德提出了这个问题。

“是的，业主也发现了这个问题，对此他们不太高兴。”鲁伊拉德说。他猜想问题的成因是这样的：混凝土内芯的重量非常大，再加上当地的土壤相对比较松软，所以内核的下沉速度可能比预想的要快。而建筑外层的钢结构还没有担负许多载荷，毕竟还有18层楼没有建好。鲁伊拉德认为，正是建筑内外下沉速度的差异导致楼板向内侧倾斜。但他觉得，只要大楼完工了，外层结构承担足够的负荷，这个问题就会迎刃而解。

让我最感兴趣的倒不是鲁伊拉德的这个解释，因为我无从判断他的猜测是否正确。真正让我感兴趣的是在施工过程中出现了意料之外的情况：上层楼板倾斜了。他们应当先清理掉这些积水，但施工计划肯定会因此而拖延。这会涉及方方面面，这么一件小事有可能让整个项目乱套。不仅如此，相关专家还必须研究清楚这个问题是否意味着建造过程

存在严重缺陷。我很好奇他们将会如何处理这个情况。不确定性总是无处不在的，他们凭什么判断这只是正常的沉降呢？又凭什么认为外层钢结构负重后真的会和混凝土内芯持平呢？正如鲁伊拉德所说的那样，“天有不测风云”。这的确是一个非常复杂而棘手的问题。

等回到底楼的现场办公室之后，我问奥沙利文他们将如何处置这一情况。我相信他们这些建造摩天大楼的人一定碰到过许多类似的问题，许多清单无法事先预计的问题。如果医生在诊治病人的过程中碰到类似问题的话，虽然病人的情况千差万别，但他们会完全凭借某个专家的判断来应对挑战。也就是说，单个医生拥有决定权，能够全权负责病人的诊治。而在俄罗斯码头项目这个例子中，鲁伊拉德就是这样的专家。如果我们把工地看作医院，那么他的个人判断就能决定事态的发展。但是，奥沙利文指出，这种方法存在缺陷。



和诊治病人一样，建造大楼也需要各个领域的专家（来自16个不同的行当）。既然全知全能的建造大师已经不存在了，那么让单个专家拥有决定权会引发灾难。不同的专家可能会做出相互冲突的决定，而一些重要问题又可能被大家忽视，这样造出来的建筑早晚会的倒塌。我实在无法想象，如果在医院里犯下这样的大错，后果会有多么严重。

“那你们会怎么做呢？”我问道。

奥沙利文指了指我的左边，在正对着施工日程安排表的墙上贴着另一张巨幅图表，样子看起来和施工计划没有什么不同，只不过这张图表的名字是“建议日程安排表”。这也是一张清单，不过，这张清单详细规定的并不是施工任务，而是沟通任务。工地项目经理应对突发问题的方法就是，确保专家们彼此充分沟通交流，建议日程安排表上会写明在X日进行有关Y问题的讨论。专家们可以做出个人判断，但他们要明白自己是团队的一员，也要充分考虑其他人的观点。经过充分讨论之后，他们会得出最终的行动方案。虽然没

有人能够事先预料所有问题，但他们至少可以事先预估可能会在什么时候和什么地方碰到问题。建议日程安排表规定了哪些人在什么时候就什么问题进行讨论，在下一步工作开始之前分享他们的专业知识或提出建议。

清单宣言 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

如果让合适的专家聚在一起，并且让他们作为一个团队进行充分讨论，那么严重的问题是可以被发现的。

比如，这张建议日程安排表规定，承包商、安装专家和升降机工程师在月底前要对负责1~10楼运输工作的升降机进行评估。升降机是由专业厂家制造的，在出厂前经过了严格检测，而且是由专家安装的。但是，建筑项目负责人并没有假设一切都会顺利进行。相反，他们认为任何环节都可能出现意外。这就是极端复杂问题的特性。不过，他们还认为，如果让合适的专家聚在一起，并且让他们作为一个团队进行充分讨论，那么严重的问题是可以被发现的。

所以，建议日程安排表的作用就是让专家们互相交流。承包商、安装专家和升降机工程师必须在月底之前进行讨论；还要在当月25日之前和耐火材料专家交流；而在两周前，结构工程师、咨询顾问和业主已经就楼板向内侧倾斜进而引起积水的问题进行了讨论。

我看到这个任务旁边已经打了钩，说明这项工作已经完成了。我问鲁伊拉德讨论进行得怎么样。

一切都很顺利。大家一起对问题进行了评估，业主和承包商现在也认为外部钢结构最后会加速沉降，和混凝土内芯持平。排水清理工作已经安排好了，施工计划进行了一定的调整，参与讨论的每个人都签了字。

极端复杂的情况总是充满了不确定性，人们总是不能确定一切是否能够顺利进行。面对未知，建筑专家们相信沟通的力量，而不相信某个人的智慧，即使他是经验丰富的工程师。他们依赖的是集体的智

慧，要确保各方面的专家都对问题进行了仔细的评估，并一起讨论得出解决方案。

清单宣言 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

一个人免不了会犯错误，
但许多人犯错的可能性或许会
变得小一些。

一个人免不了会犯错误，但许多人犯错的可能性或许会变得小一些。

在现场办公室的另一个房间里，一个发型时髦、身穿黄色反光背心的年轻男子坐在两个大型液晶显示屏前。这人名叫瑞安·沃尔什（Ryan Walsh），他的工作是将各行当专家提交上来的施工方案输入电脑，并将它们整合成大楼的三维图。在图中，大楼每一层的结构都非常详细。他在屏幕上调出了顶楼的三维图。那时候，他已经将9个专业的施工方案整合在一起，其中包括结构工程、电梯安装、管路铺设等。他点击鼠标，带领我们在这一楼层参观，就好像我们真的在过道里行走一样。你可以在屏幕上看到所有东西，比如墙、门、安全阀等。更重要的是，你能在屏幕上看到问题，比如，有个地方的高度不足以让一般的成年人通过。沃尔什为我演示了一个名叫“冲突侦探”（Clash Detective）的软件，这个软件能将不同专业施工方案互相冲突的地方，以及施工方案和建筑法规冲突的地方都找出来。

沃尔什说：“如果一根结构梁跨越的地方应该安装灯具，那么‘冲突侦探’软件就会在屏幕上用不同的颜色来标注这根横梁。有时能发现几百处冲突。有一次，我甚至发现了2 000多处冲突。”但仅仅发现冲突是远远不够的，问题必须得到解决。为此，必须确保专家们对问题进行讨论。所以，软件会将相关问题纳入建议日程安排表，并给相关专家发去电子邮件。

他们还有一个名叫“项目中心”（Project Center）的软件。只要有人，哪怕是一线的建筑工人发现了问题，都可以通过这个软件发邮件通知相关专家，并通过该软件对讨论进程实施跟踪，最后还要确保问题经过了充分讨论，并得以解决。当我们回到萨尔维亚的办公室时，鲁伊拉德给我看了一封这样的电子邮件。

这封邮件是鲁伊拉德在上周收到的，铆接4米钢梁的一位工人发现4个接口中有两个不吻合。他不知道这是不是大问题，于是用数码相机拍了照，通知了鲁伊拉德。鲁伊拉德回信说这个情况不正常。他们共同商讨出解决方案，将钢梁焊接到位。电子邮件还自动抄送了要对解决方案签字的专家。每个人有3天时间对方案进行评估，并做出答复。在这一过程中，每个参与讨论的人都要确认自己与他人进行了充分沟通，因为即使是这么一个小小的修复任务，也会对多方面的工作产生影响。

萨尔维亚曾经告诉我说，在过去的几十年里，建筑领域的最大进步就是对进程跟踪和沟通过程进行了不断完善。直到这个时候，我才明白他的话是什么意思。

团队沟通不畅的代价

在面对巨大挑战的时候，建筑行业愿意把自己的命运交付给集体决策的决心让人感到惊讶。萨尔维亚的合伙人是罗伯特·麦克纳马拉（Robert McNamara），他是曼哈顿花旗大厦的一名结构工程师。

花旗大厦非常有特色，它那标志性的屋顶由一个斜坡构成，整个建筑高达279米，给人留下最深印象的是它底部的4根巨柱。这些柱子有9层楼高，它们把整个建筑腾空托起。更奇特的是，它们并不是位于建筑的4个角上，而是位于四条边的中央。首席结构工程师威廉·勒梅萨里尔（William LeMessurier）为这些柱子设计了隐形的倒V形巨型支架对其进行加固。建筑的外形非常抢眼，整个摩天大楼就好像漂浮在第53街的上空一样。但风洞测试显示，由于这座大楼比周围的建筑物高出许多，所以强风和湍流会对其施加强大的作用力。要使建筑承受这么大的力量都快超出结构工程师的能力范围了，似乎只有飞机设计师才能应付得了。对于大楼可以接受的摇摆程度，大家心里实在没底。

那么他们该怎么办呢？他们既没有放弃这个独特的设计，也没有降低建筑的高度。麦克纳马拉提出了一个新奇的方案，这个方案叫作“调谐质量阻尼器”。麦克纳马拉认为，他们可以在大楼第59层的楼顶用巨型弹簧悬挂一个重达400吨的混凝土块。当强风把大楼推向一侧的时候，这个混凝土块会向反方向摆动，并以此抵消强风的作用力。

这是一个非常聪明而简洁的解决方案。工程师们用一个较小尺寸的模型进行了风洞测试，结果令人非常满意。但是，对这么复杂的项目来说，错误和不确定性总是存在的。为了尽量减小犯错的可能，专家们在开工前的最后时刻进行充分沟通。大厦的业主与建筑设计师、城市规划部门官员、结构工程师等专家进行了讨论。他们对整个设计方案和相关数据进行了仔细评估，确认所有可以预想的问题已经得到妥善解决。最后，他们依次在施工计划上签字，建设工作由此正式拉开帷幕。

他们知道不能依赖单枪匹马，而是要依靠集体的智慧。他们用一套清单来保证不遗漏任何简单的问题，不跳过任何简单的步骤；用另一套清单来保证所有专家都对困难的和意料之外的问题进行充分讨论，并共同商讨出解决方案。

在人口稠密的市中心地区建造结构如此复杂的建筑实在让人有点不放心。要知道，单是在大厦里工作的就有几千人，还有几万人在附近工作和生活。这么做要冒很大的风险，似乎是个不明智之举。但我们相信专家们有能力应对复杂的问题。他们知道不能依赖单枪匹马，而是要依靠集体的智慧。他们用一套清单来保证不遗漏任何简单的问题，不跳过任何简单的步骤；用另一套清单来保证所有专家都对困难的和意料之外的问题进行充分讨论，并共同商讨出解决方案。

奥沙利文告诉我：“在这个行业，造成严重错误的最大原因往往就是沟通不畅。”

就拿花旗大厦来说吧，在计算建筑牢固程度的时候，工程师假设底部基座上的巨型支架是焊接起来的，但是焊接工作非常耗费人力，成本较高。承建这一项目的伯利恒钢铁公司（Bethlehem Steel）提议使用螺栓对支架进行铆接，但这么做会让建筑变得不那么牢固。他们通过计算，认为铆接的强度完全可靠。但是，《纽约客》后来披露，他们的计算结果不知是何原因，没有经过勒梅萨里尔的审阅，这个检查点被跳过了。然而，当时即便勒梅萨里尔审查了计算结果，也未必能够发现问题。

在1978年，也就是大厦投入使用一年之后，普林斯顿大学的一名工程专业学生发现建筑强度不够，勒梅萨里尔这才发现他的设计被更改了。更糟糕的是，这意味着建筑存在致命缺陷：它将无法抵抗时速在112千米以上的强风。而纽约的气象资料显示，这样的强风每55年至少会发生一次。如果这一情况真的发生，那么支架的连接点会断裂，大厦会从第30层开始塌陷，进而倒塌。但他们发现问题的时候大厦已经完全投入使用了。勒梅萨里尔把这个坏消息告诉了大厦业主和市政官员。那年夏天，当飓风艾

拉（Ella）向纽约逼近的时候，一个紧急施工队在大家下班后秘密地将5厘米厚的钢板焊接在200个重要的螺栓上面。大厦终于保住了，直到今天依然稳稳地矗立在那里。

虽然建筑业的清单制度并不能发现所有问题，但是它的成功让人啧啧称奇。

清单的力量 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

美国有将近500万座商住楼、1亿座低层住宅和800万座高层住宅，每年要新建大约7万座商住楼和100万座住宅楼。这些建筑部分或全部倒塌的情况非常少见，特别是摩天大楼。俄亥俄州立大学于2003年进行的一项研究显示，美国每年发生的严重建筑事故只有20起，这意味着建筑行业每年的严重事故发生率不到0.002%。

正如萨尔维亚对我说的那样，虽然如今的建筑比历史上任何一个时期的建筑都要复杂，设计师和工程师要考虑抗震、能效等诸多问题，但是施工的时间却比几十年前缩短了1/3。

清单真的非常有效！

第二部分

清单革命的行事原则



THE
CHECKLIST
MANIFESTO
HOW TO GET
THINGS RIGHT

第4章 权力下放 清单奏效的关键

·350万美元、2 498箱救援物资，为什么沃尔玛能抢先一天把水和食物送到灾民手中？生死时速面前，卡特里娜让政府救援队汗颜。

·每个人都在等待救世主，但中央极度集权的解决方法只会让人等得望眼欲穿。将决策权分散到外围，而不是聚集在中心，让每个人担负起自己的责任，才是让清单奏效的关键。

中央极度集权行不通了

建筑业为了应对极端复杂的问题会使用特殊清单。清单所体现出的一些理念非常引人注目，那就是把权力分给更多人。在面对风险的时候，大多数当权者喜欢把权力、决策权等攥在自己手里。一般来说，高层会用清单列出下属应该完成的工作，并督促他们及时正确地完成。我所见到的第一张施工清单，那张贴在奥沙利文办公室右侧墙上的施工日程安排表，就属于此类清单。那张清单写明了每一个关键步骤的具体要求，写明了人们在碰到简单或平常问题时的例行处理程序。它起到的作用就是“强制函数”。

但是，在奥沙利文办公室的另一面墙上还有一张清单。当人们遇到非常规的、可能引发危险的复杂问题时，比如在建高楼的楼板突然出现意料之外的内倾状况时，这张清单就能派上用场了。它所体现出的理念与前者截然不同，因为它将决策权分散到外围，而不是聚集在中心。这种理念给了人们足够的空间，让他们能够根据自己所掌握的经验 and 专业技术适应环境，应对挑战。高层需要做的并不是直接进行决策，而是督促大家积极参与讨论，让他们担负起自己的那一份责任。这就是此类清单奏效的关键所在。

奥沙利文告诉我，如今的施工监理行业也将这种方法奉为标准。他说，现在的监理不会自己重新计算建筑的抗风强度，也不会评估钢梁的连接处是应该铆接还是应该焊接。无论一个人有多么聪明，知识有多么渊博，他都不可能独立完成俄罗斯码头项目或花旗大厦这样复杂的建筑监理工作。所以，虽然他们会尽全力监督施工质量，但现在要做的是确保施工人员及时、正确地完成施工。监理将自己的权力和责任分散给了相关各方。

奥沙利文说：“这么做是有道理的。如果什么事都要监理亲力亲为，那么就算是只建两间小屋，其安全系数也不见得比大型项目的更高。”我觉得，一些当权者已经意识到，如果他们该放权时不放权，就会遭遇失败。我们只要看看卡特里娜飓风席卷新奥尔良之后发生的一切，就能明白这个道理。

高层需要做的并不是直接进行决策，而是督促大家积极参与讨论，让他们担负起自己的那一份责任。这就是清单奏效的关键所在。

2005年8月29日早上6点，卡特里娜飓风在新奥尔良市的普拉克明县（Plaquemines Parish）登陆。起初的灾害预报大幅低估了灾害的严重程度。由于电力被切断，各种通信中断，灾害预警消息无法发送出去。到了下午，市区周围的防洪堤决口了，市区大部分地方被洪水淹没。有关灾情的报道在电视上不断播报，但是美国联邦紧急事务管理局（Federal Emergency Management Agency, FEMA）的负责人迈克尔·布朗（Michael Brown）却对此视而不见，并在新闻发布会上宣称，局面尽在掌控之中。

FEMA的信息渠道有很多，但是该机构在新奥尔良常驻的代表只有一人。这位代表在飓风登陆的当天下午费尽周折搭上海岸警卫队的直升机查看了市区的灾情，并发出了紧急情况报告。但大多数通信都被切断了，他只能通过电子邮件发出这份报告。这位代表在报告中说，新奥尔良市内洪水肆虐，他亲眼看到有尸体漂浮在水面上，还有成百上千人站在楼顶等待救援。但政府高层官员并不使用电子邮件。事后的一次国会听证会披露，直到飓风登陆的第二天，政府高官才接到相关信息的通报。

但那个时候，新奥尔良市区80%的地方已经被洪水淹没。有两万多难民滞留在超级圆顶体育场，还有两万多难民聚集在莫里尔会议中心（Morial Convention Center）。有5 000多人被转移到了10号州际公路的立交桥上，其中一些人是救援队员转运过来的，但大多数人是自己逃难出来的。他们来不及搬出自己的财产，许多人只是随身带了几件衣服。市内的各家医院因为电力被切断而处于瘫痪状态。人们因为得不到维持生存的食物和水而开始抢劫，暴力事件的上升成了一个严重问题。

很多当地政府官员和自发组织者竭尽全力联络联邦政府，希望他们能够知道灾区最需要什么，但一直无法和对方取得联系。当他们最终通过电话找到某个官员的时候，却被告知要耐心等待，因为信息需要逐级上报，传统的指令和控制系统很快已被大量信息和指令淹没。有太多决策要做，但是决策者得到的信息不够及时，不够具体，他们不知道应该向什么地方提供什么样的帮助。

尽管如此，政府仍拒绝放弃传统指令模式。灾情正在不断恶化，而各级政府却在争论决策权的归属问题。联邦政府不愿意把权力下放给州政府，州政府不愿意把权力下放给当地政府，更没有人愿意把权力下放给企业、非政府组织或个人。

无政府状态和奥威尔式官僚造成了可怕的后果。当局拒绝满载饮用水和食物的卡车进入灾区，因为他们的救援计划并不包括这些内容。公车征用令迟迟不发，美国交通运输部直到两天后才收到正式的征用令公文。数万人已经被围困，需要及时疏散。但与此同时，新奥尔良却有两百多辆公共汽车被闲置在附近的高地上。

清单宣言 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

在面对极端复杂的问题时，高层应该尽可能把权力下放给一线人员，而不是将大权集中在自己手中。极端复杂的问题本来就是出乎人们意料的，对于此类问题，传统的中央极度集权处理范式是行不通的。

问题并不是这些高官缺乏同情心，而是他们没有认识到，在面对极端复杂的问题时，他们应该尽可能把权力下放给一线救援人员和当地官员，而不是将大权集中在自己手中。每个人都在等待救世主，但中央极度集权的解决方法只会让灾民们等得望眼欲穿。事后，时任国土安全部部长的迈克尔·切尔托夫（Michael Chertoff）被要求对政府迟缓的救援行动做出解释。他说：“卡特里娜飓风是一次超级灾难，是‘完美风暴’，超出了决策者甚至是任何一个人的预料。”但这一解释说不通，极端复杂的问题本来就是出乎人们意料的，这是它们的基本属性之一。上述事件说明，对于此类问题，传统的中央极度集权处理范式是行不通的。

每个人都是清单的主宰者和参与者

让人意想不到的是，在所有组织中，把救灾问题的复杂性理解得最透彻的竟然是沃尔玛。哈佛大学肯尼迪政府学院的教授对沃尔玛在救灾过程中的表现进行了研究。研究报告向我们展现了一个灵活应变、充分下放权力的零售巨擘。得知新奥尔良遭受严重风灾的消息后，沃尔玛首席执行官李·斯科特（Lee Scott）只是发布了一条非常简短的指令：“本公司将对风灾做出相应级别的响应。”有人记得他在公司高层会议上是怎么说的：“在座各位将要做出超出自己级别的决定，请务必根据所掌

握的信息及时做出最佳的选择。记住，最重要的就是做正确的事情。”

就这样，简简单单的指令从高层传递到一线门店经理那里。这条指令真是给他们松了绑，让他们放开手脚便宜行事。



沃尔玛共有126家门店因为被淹或停电而歇业，被困员工及家属超过两万人，公司的救援行动最早把帮助他们作为工作重点。有一半以上受损的门店在行动开始后的48小时内重新开业。但是，当管理人员从员工口中得知灾情非常严重的时候，救援重点立刻从重新开业变成了紧急救援。他们心里想的是：“天哪，我们能为那些处于水深火热中的人们做些什么？”

沃尔玛的门店经理自行决定为当地居民分发物资，其中包括纸尿布、饮用水、婴儿配方奶粉和冰块等。当联邦紧急事务管理局的官员还在思考应该如何征用各种补给品的时候，沃尔玛的门店经理已经启用了手动记账系统来记录提供给先遣救援队的各种物资，其中包括食品、睡袋、卫生纸，还有像短柄小斧、绳索和靴子这样的救援设备。10米高的暴风潮席卷了一家门店，待洪水退去后，这家门店的经理助理开了一辆铲车进入店里，把所有能够抢救出来的物资搬到停车场上进行分发。当得知一家医院药物短缺的时候，她再次冲到店里的药品柜台，把可以用的药物都取了出来。后来，她所做的一切得到了公司高层的称赞。

沃尔玛的高层把工作重点放在设定目标、监控进度以及保持和一线员工及政府机构的联络上。也就是说，在应对这一极端复杂的危机时，他们没有发布具体的指令。由于情况出乎意料，而且瞬息万变，所以，他们努力确保人们能够及时有效地沟通。沃尔玛的救援团队甚至包括来自红十字会的成员（美国联邦政府拒绝参加沃尔玛的救援团队）。公司还为员工开设了24小时呼叫中心，这个中心一开始只有8名接线员，但由于话务量不断攀升，接线员的数量很快就增加到80位。

在尽可能彼此协调、救助他人这一共同目标的指引下，沃尔玛的员工自创了许多了不起的办法。他们在市里设立了3个移动药店，并计划免费为有紧急需要的灾民提供药品，即使没有处方也没有关系；他们还设立了服务点，为持有工资支票等各类支票的人们兑换现金；他们甚至开设了临时诊所，为灾民接种预防灾后疫情的疫苗。最了不起的是，仅仅在飓风登陆两天之后，沃尔玛的物流团队已经想办法让一辆辆满载着食品、饮用水和救援设备的卡车绕过重重障碍，到达迫切需要援助的灾区。他们在政府救援力量到达灾区的前一天，就把水和食物送到灾民甚至是美国国民警卫队的手里。在整个救援过程中，沃尔玛总共运送了**2 498**个集装箱救援物资，并为灾民和指挥中心捐赠了价值达**350**万美元的物资。

后来，杰斐逊县（Jefferson Parish）县长亚伦·布劳萨德（Aaron Broussard）在接受电视采访时说：“如果美国政府也能像沃尔玛那样做出及时响应，这场灾难就不会造成这么巨大的损失了。”

我们需要在矛盾中寻求平衡

我们应该从以上故事里吸取什么样的教训呢？对于这个问题，人们可能并没有想清楚。有些人认为，在处置这类极端复杂的情况时，私营部门比公共部门更加高效。但事实并不像他们所说的那样。沃尔玛的确是一个成功案例，但我们也能找出一大堆没有做出恰当响应的大公司——当地的公用事业企业无法修复通信线路和电网，而石油公司根本没有为这么大的灾难储备足够的原油，他们的炼油能力也跟不上。当然，政府部门也并非一事无成。比如，在灾难发生后的几天里，新奥尔良当地的警察和消防队员虽然缺乏各种救援设备，但是他们组织很多体格健壮的志愿者，用平底船从水中、房顶上和阁楼里救出了62 000多人，这是非常了不起的。

我们从中真正应该吸取的教训是，由于极端复杂问题往往不可预测，这类问题的解决超出了个人能力的范围。所以，所有指令都由核心层、最高层来决定的做法是注定要失败的。人们需要行动和适应的余地。但是，即便有再多的人手，如果他们彼此孤立，取得成功也只能是异想天开，因为这种情况无异于无政府状态。为了获得成功，我们要把一些看似相互矛盾的元素结合起来，在矛盾中寻求平衡，并需要不断跟踪各项工作的进度。

由于极端复杂问题往往不可预测，这类问题的解决超出了个人能力的范围。所以，所有指令都由核心层、最高层来决定的做法是注定要失败的。

建筑行业从业人员对于应对极端复杂问题的理解就很深入。更了不起的是，他们将这一理解转变为简单易行的清单，把对复杂性的管理纳入平时的例行工作之中。

这需要他们在一对对矛盾之中寻找合适的平衡点，如自我和纪律、个人发挥和标准程序、单兵作战和集体协作。为了用清单达到上述平衡，他们设计出两种截然不同甚至几乎对立的清单。他们使用一整套清单来确保那些看似简单但十分重要的步骤不被人们忽略，而用另一套清单来确保人们充分沟通，互相协调，承担责任，并赋予他们权力，让他们用所知的最佳方法来解决瞬息万变和出人意料的问题。

从卡特里娜飓风和建筑行业的通行做法，我们可以知道：在极端复杂的情况下，清单不只有助于成功，而且是人们获得成功所需的必要条件。当然，主观判断也是不可或缺的，但是我们需要使用清单来协助主观判断，甚至用清单来提高主观判断的质量。

清单，安全与高质量服务的关键

茅塞顿开的我开始意识到清单无处不在。你可以在职业足球教练那里找到它们，也可以在炫目的舞台上发现它们的身影。我曾在广播中听到这样一个有趣的故事。



故事的主人公是著名摇滚乐手大卫·李·罗斯（David Lee Roth），他是范·海伦（Van Halen）乐队的主唱之一。每次签订巡演合同的时候，罗斯都会坚持在合同中包含这样一个条款：后台化妆间里必须摆放一碗M&M's巧克力豆，而且里面不能有棕色巧克力豆。这一点如果主办方做不到的话，演唱会将被取消，而且主办方还要对乐队进行全额赔偿。至少有那么一次，范·海伦乐队因为上述原因霸道地取消了科罗拉多的一场演唱会，原因是罗斯在化妆间里找到了棕色巧克力豆。有人或许会认为大明星总是喜欢摆谱，提出不近人情的苛刻要求。其实不然，这是罗斯用来保障演唱会安全的一块试金石。

罗斯在其自传《来自热浪的疯狂》（*Crazy from the Heat*）一书中写道：“范·海伦是第一支将演唱会开到偏远城市的乐队。我们的设备足足装了9辆18轮卡车，而一般的演唱会只需要3辆卡车就行了。工作人员一不留神就会犯技术错误，比如横梁因为无法负重而倒塌，地板也会因为不堪重负而塌陷；有的舞台的门不够大，舞台置景无法通过。演出的合同附文读起来就像是黄页，因为设备实在是太多了，调试安装工作需要大量人手。”为此，他们设计了一个小测试，也就是合同附文的第126条那个关于巧克力豆的条款。罗斯写道：“如果在后台放置巧克力的碗里发现了棕色巧克力豆，我们就会对各项装配工作进行逐一检查。我保证会发现技术错误，会碰到各种各样的问题。”这些可不是鸡毛蒜皮的小事，一些错误会威胁到人们的生命安全。就拿那次被取消的科罗拉多演唱会来说吧，乐队发现当地主办方没有仔细阅读有关舞台重量的要求。如果演出如期进行的话，舞台完全有可能在演出中坍塌。

“大卫·李·罗斯也用清单！”我对着收音机大喊。

我还找到乔迪·亚当斯（Jody Adams）来检验我的理论，看看使用清单是否必要。亚当斯是里阿尔托（Rialto）意大利餐厅的主厨和老板，这家餐厅是我最喜欢的波士顿餐厅之一。在20世纪90年代，亚当斯被《美食与美酒》（*Food and Wine*）杂志评为美国十佳新秀厨师之一，并于1997年获得詹姆斯·比尔德基金最佳厨师奖（James Beard Foundation Best Chef），这一奖项相当于美食界的奥斯卡。里阿尔托经常登上全美最佳餐厅的榜单，《绅士》（*Esquire*）杂志也曾将其列入榜单。亚当斯的专长是制作意大利风味的美食。当然，她对菜谱做了不少革新。

亚当斯自学成才，她在布朗大学（Brown University）学的是人类学，也从未上过厨师学校。但她对做菜很感兴趣，于是就去了一家餐厅打工。她从切洋葱学起，一步一步，直到创造出自己的厨艺风格。

里阿尔托餐厅的菜非常精美，色香味俱全，而且多年来一直保持着非常高的水平。我很想知道亚当斯是怎么做到的。我很清楚像汉堡王和塔可钟（Taco Bell）这类快餐连锁店是如何在全球范围内进行标准化运营的。这些店里的每一种食品都是根据既定规程，用流水线生产出来的。但在大饭店里，厨师要对菜谱不断进行改进和雕琢，每一道菜都是不同的。而且，他们要日复一日、年复一年地保持高水平。要知道，他们晚上最多要接待300位顾客。我为这一完美的表现找到了一种理论，但这一理论真的说得通吗？于是，我接受了亚当斯的邀请，来到她的餐厅进行实地考察。



那是一个周五的晚上，我整晚都坐在凳子上看厨师们忙忙碌碌。餐厅的厨房很狭长，厨师们的喊叫声和食物煎炸烧烤发出的嘶嘶声混杂在一起，好不热闹。那晚，亚当斯和她的员工要在5小时内接待150多名顾客，为他们提供几十种美味菜肴。

亚当斯手下的厨师技艺都非常高超，他们中有一半人接受过厨师学校的专业培训，大多数人都有10年以上的工作经历。这些厨师各有分工，有的只做面点，有的负责烘焙，还有专门的烧烤厨师、煎炸厨师、甜点厨师、副主厨和酒侍。多年来，他们的技艺日臻完美。他们所做的大部分工序我实在是搞不清楚。即便我是个外科医生，他们也不会让我动他们的刀。

在人们的印象里，烹饪靠的是技巧和创意。如今，厨师们的大胆发挥是厨艺比拼节目大受欢迎的重要原因，但我在里阿尔托看到的却是严明的纪律，而不是光鲜亮丽的外表和天马行空的发挥。厨房的正常运转需要的是纪律，而清单在其中起到了关键作用。

第一张清单就是菜谱，这是最基本的清单，涉及每道菜的制作方法。菜谱打印出来后要塑封，然后放置在每个操作台上。亚当斯将这些菜谱奉为“圣经”，并要求下属严格执行。她说：“遵从菜谱对于长期维持高质量服务是至关重要的。”

甜点操作台的旁边挂着一块告示板，上面贴着被亚当斯称为“厨房批注”的条子，这其实是她发给员工的一些关于菜品质量的电子邮件。最近一封是前一天夜里00:50发出的，上面写道：“油炸馅饼中要多放些香料和大蒜，做得更加松软些；穗丝要用玉米做；奶油玉米浓汤要放在椭圆形的盘子里，不要放在方形的盘子里；蘑菇汤要多放点葱、蒜和马沙拉白葡萄酒。记住：按照菜谱操作！”

清单宣言 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

菜谱是最基本的清单。遵从菜谱对于长期维持高质量服务是至关重要的。

厨师们并不总是喜欢按照菜谱进行操作。对于一个做了几百次奶油玉米浓汤的厨师来说，他会觉得整个制作过程自己早就烂熟于心了，不需要再按部就班了。但亚当斯却认为，这时候菜肴质量水平就会开始下滑。

当然，这些菜谱并非一成不变。我所见到的菜谱在页边上都写了一些改动意见，其中有许多都是亚当斯的下属提出的，他们有时候甚至会重新做一张菜谱。

那时候亚当斯新开发出一道菜：整只龙虾佐干邑白兰地和鱼汤，配上小圆蛤和西班牙香肠。这道菜被收入了茱莉亚·蔡尔德（Julia Child）的著名菜谱中。不过，在把新菜写上菜单之前，亚当斯总是会让厨师们试做几次以充分暴露问题。亚当斯的菜谱要求厨师把龙虾一分为二，然后用橄榄油小火慢煎，但菜肴的质量总是不能精确控制，龙虾要么没煎透，要么煎过火了。此外，这道菜的调味酱要根据顾客的要求定制，但这道工序需要的时间太长了，而顾客一般只能等8~10分钟。

所以，亚当斯和两位厨师对制作工序进行了改进。他们决定事先做好调味酱，并将龙虾煮成半熟。一次又一次的尝试让这道菜变得越来越完美，菜谱也重新修改了一遍。

此外，他们还为每位顾客准备了一张清单。每当顾客点餐的时候，厨房就会打印出一张纸条，上面写明了顾客点的菜肴、顾客的桌号、座位号、顾客特地关照的注意事项或数据库中记录的相关事宜，如顾客是否对特定食物过敏、牛排要几分熟、当天是否是顾客的生日、是否有贵宾到场、是否需要亚当斯出去打个招呼等。作为厨房运营现场指挥的副主厨会把每张条子都大声朗读出来。

“蘑菇汤现做。莫泽雷勒干酪现做。龙虾等一等。不加面筋的十分熟牛排等一等。”

“现做”意味着这是第一道菜，而“等一等”则意味着这是第二道菜。而且这位顾客对谷蛋白过敏，所以牛排里不能加面筋。负责操作的各位厨师在听到顾客的要求以后要进行重复，以确保他们接收到的指令准确无误。

“蘑菇汤现做。莫泽雷勒干酪现做。”一位厨师重复道。

“龙虾等一等。”海鲜厨师喊道。

“不加面筋的十分熟牛排等一等。”烧烤厨师喊道。

和建筑行业一样，在厨师这个行当里，并不是每件事情都能被简单地写进菜谱。所以，亚当斯也开发了沟通清单，以确保大家在面对意料之外的问题时能够像一个完整的团队一样沟通协作。每天下午5:00，也就是离餐厅开门迎客还有一个半小时的时候，当值厨师会聚在一起召开一个碰头会。他们要简短地讨论一下事先没有料到的问题，并以此来应对这些复杂问题的不确定性。在我访问的那天晚上，他们在开会的时候讨论了当天顾客的预定菜单，对菜单进行了两处修改，想办法为一名请病假的员工补缺，还为一个突发情况制订了应对方案。这个突发情况是这样的：当天会有20多个女孩来餐厅开生日派对，但她们在路上堵车了，预计到达的时间正好是餐厅的营业高峰，厨师们可能会忙不过来。在开会的时候，每个人都能发言，而且大家会共同商讨行动计划。

当然，这些举措还是不能确保不出任何问题。汤可能装盆太早了，等到上菜的时候已经凉掉了，鹌鹑的调味酱料可能不够了，鲈鱼可能烤得太干了。所以，亚当斯还设置了最后一道检查：每道菜在被端出厨房之前都必须经过她或副主厨检查。他们要确保菜肴看起来没有问题，要核对顾客的订单，要闻一闻，或许还要用干净的勺子舀一勺尝一口。

我数了数，至少有5%的菜被退了回去。副主厨对煎炸厨师说：“这盘鱿鱼要炸得再熟一点，炸得再金黄一点。”

待考察完毕后，我终于有机会坐下来品尝一下美食了。我点了炸橄榄、烤蛤蜊、夏日玉米粥、当地农场的绿色沙拉，当然还有美味的龙虾。这些菜的味道实在是鲜美至极。离开餐厅的时候，我的胃里塞满了食物，但我的大脑还在飞速运转。亚当斯的厨艺更像一门艺术，而不是一门科学。但即使在这个高度专业化、分工细致并且需要高超手艺的行业里，人们还是离不开清单。我收集到的各项证据都指向相同的结论，我找不到清单没有用武之地的行当或职业。对于我的工作来说，这一论点或许同样适用。

第5章 简单至上 清单要素的选择机制

·据估计，每年全球至少有700万人在术后残疾，至少有100万人没有走下手术台。如果我们把手术不同阶段的清单合并成一张清单来执行，是不是可以减少术后残疾或术中死亡的病人数量？

·从来没有全面的高效，从来没有一张清单能涵盖所有情况，冗长而含糊不清的清单是无法得到高效和安全执行的。清单要素的遴选，必须坚守简单、可测、高效三大原则。

一切为了安全和正确

2006年下半年的时候，我接到一个来自日内瓦的电话，打电话的是一位操着英国口音的女士。她自称是世界卫生组织的工作人员，她说给我打电话是想问我能否帮助他们组织一批专家解决一个小问题。他们收集的数据显示，全世界的外科手术数量正在不断增加，而其中有很大一部分没有达到安全标准，会给病患造成生命危险。所以，他们想设立一个全球性的项目，以尽量减少手术中的死亡和伤害事故。

我当时的回答是这样的：“你们想怎么做呢？”

“我们会召开一次会议。”她说。

我问他们会为这个项目投入多少经费。

“哦，钱真的不多。”她说。

我拒绝了她。没有钱我可做不到，我很忙。

但是对方很清楚谈判的艺术。她在电话里说了这么一句话：“哦，对不起，我以为您是外科手术安全方面的专家，是我搞错了。”

于是，我同意为他们组织召开一次会议。

世界卫生组织从193个会员国那里收集了大量数据和报告，为他们工作的一大好处就是能够查阅和使用这些数据。在对现有数据进行整理之后，我的研究团队发现，世界卫生组织官员得出的结论是正确的。

在全球范围内，外科手术的数量正在飞速增长。到2004年，全球每年要实施2.3亿台大型手术。也就是说，每25人中就有一人要在一年内接受一次大型手术。而这一数量还将继续增长。大多数人没有意识到，每年全球外科手术的数量已经超过了全球新生儿的数量。虽然大多数时候手术的进行是非常顺利的，但这并不意味着没有问题。

据估计，医院外科手术术后并发症的发病率在**3%~17%**。虽然手术的切口变得越来越小，病人恢复得越来越快，但手术的风险却依然很大。据估计，每年全球至少有**700**万人在术后残疾，而至少有**100**万人没有走下手术台。手术造成的伤害，与疟疾、结核病等传染性疾病造成的伤害不相上下。

世界卫生组织是一个致力于解决大规模公共卫生问题的组织，而手术安全问题却非常具体，有着较高的技术因素。在分析完数据以后，我才明白，他们为什么会突然对这一问题感兴趣。近几十年来，全球各地的经济状况显著改善，人们的寿命越来越长，所以对大型手术的需求也就不断增加。需要接受手术治疗的疾病有很多，如癌症、骨折、其他创伤性损伤、分娩并发症、严重出生缺陷、肾结石、胆结石和疝气等。虽

然还有20多亿处于贫困地区的人们无法接受手术治疗，但全世界手术数量飞速增长是一个不争的事实。所以，在世界各地，手术的质量和安全性就成了一个大问题。

但我们又能做些什么呢？在全球范围内减少外科手术造成的伤害，和开展根治脊髓灰质炎的运动可不能相提并论。我曾经跟随世界卫生组织的医生到世界各地看他们如何为当地的儿童接种疫苗。即使是这类看似简单的工作，在全球范围内开展起来也会变得困难重重。而外科手术更加复杂，即使是在一家医院里减少外科手术造成的损伤也不是一件轻松的事。而要想办法让全球各地的医院提高手术安全性，简直就是不可能完成的任务。要知道，外科手术有2 500多种，从脑组织活检到脚趾切除，从心脏起搏器植入到脾脏切除，再从阑尾切除到肾脏移植，你根本不知道应该从何处着手。我或许能够帮助世界卫生组织提高某种手术，如中心静脉置管插入的安全性，但这一小小的努力，对于全球手术安全性的提高又能有多少作用呢？

不是“无所不包”，而是简单、可测与高效

2007年1月，我们在日内瓦的世界卫生组织总部召开了为期两天的讨论会。参加会议的有来自世界各地的外科医生、麻醉医生、护士、安全专家，甚至还有病人。他们之中既有来自欧洲、加拿大和美国顶级医院的医生们，也有来自红十字国际委员会的外科医生。红十字国际委员会负责向世界各个遭受天灾人祸的地方派遣医疗队伍，救治当地患病和受伤的难民。参加会议的还有一位来自津巴布韦的父亲，他的女儿在接受手术治疗时不幸因为缺氧而窒息。在与会人士介绍各自的经历和发现之后，我对成功更加不抱希望了，我们怎么可能解决那么多地方的不同问题！

有一位40多岁的医生来自加纳西部，那里的可可种植业和掘金业很发达，让当地百姓摆脱了贫困。这位医生为我们讲述了当地医院的状况，他说没有外科医生愿意待在那里。加纳面临着大量的人才流失，许多高级人才都到海外寻求发展。这位医生介绍说，他们当地的医院只有3位医生，而且都是没有接受过外科手术专业训练的全科医生。但是，如果有病情十分严重的病人，如因为难产失血过多而奄奄一息的产妇、因为患阑尾炎而高烧不退的病人或因为交通事故而造成气胸的病人前来就诊，那么这些医生别无选择，只能硬着头皮为他们实施手术。

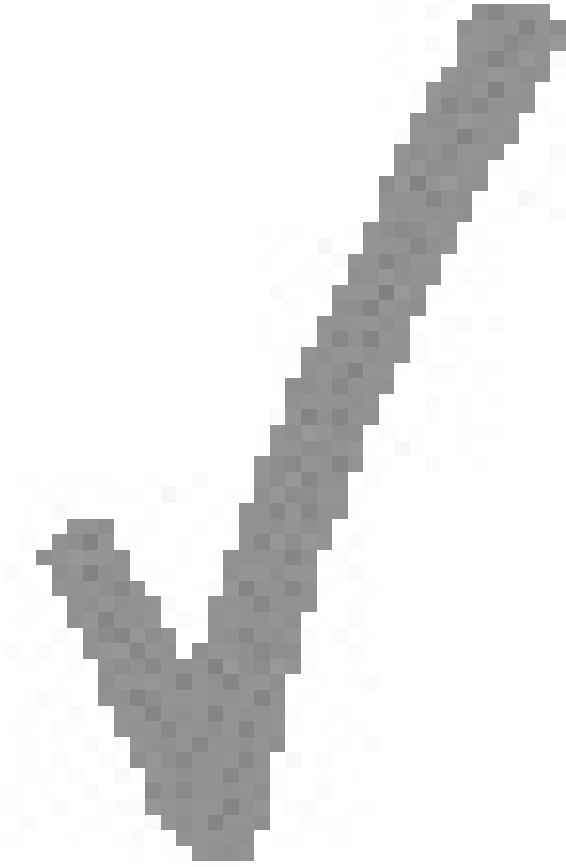
他说：“你要知道，什么样的活儿我都要干，既要当儿科医生，又要当妇产科医

生，还要当外科医生。”他的手上有一些教科书，还有一本介绍手术实施基本技巧的手册。他的一个助手并没有接受过专业训练，靠自学掌握了实施麻醉的技巧。他们医院的条件很差，没有高级医疗设备，手术常常会出问题。但他觉得尽力而为总比什么都不做要好。

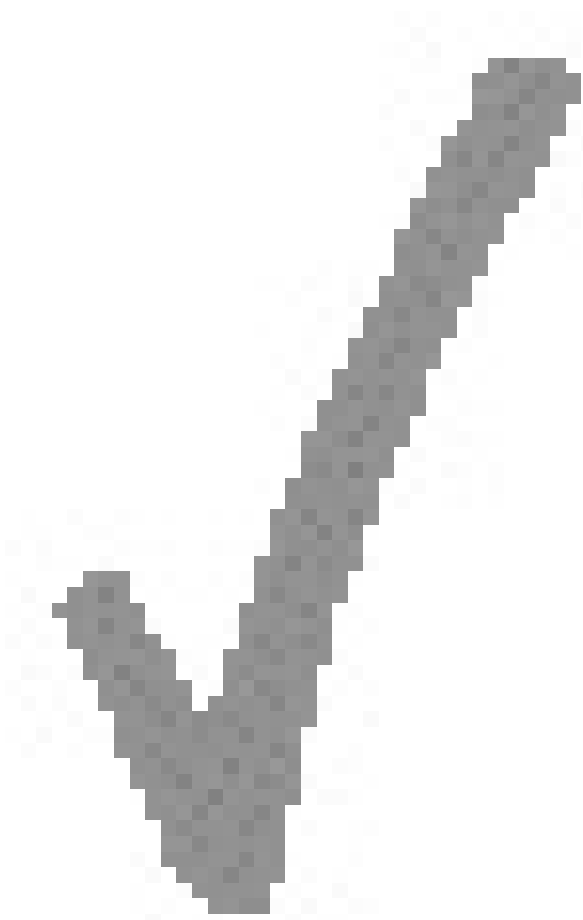
一位来自俄罗斯的生物工程学家也在会上发了言。他职业生涯的很长一段时间，都花在了监督世界各地医院医疗设备的使用情况上。他发现，无论是在贫穷地区，还是在富裕地区，医疗设备的使用和维护都存在问题。他碰到过手术设备因维护不当而引发火灾或导致病人触电身亡的情况，碰到过医护人员因为没有接受恰当的培训而错误使用先进科技产品的情况，甚至还碰到过在病人生命垂危时急救设备却被锁在柜子里的情况。

蒙古最大一家医院的外科主任医师告诉我们，那里缺少止痛药物。而来自亚洲、非洲和中东地区的医生也反映了相同的问题。新西兰的一位学者则提到了贫困国家因为麻醉不当而导致手术死亡率居高不下的问题。虽然在非洲的一些地区，普通手术的死亡率小于1/5 000，但是非洲其他地区的手术死亡率要高出十多倍。在多哥进行的一项调查显示，每**150**个接受手术治疗的病人中，就会有一人因麻醉不当而丧生。

一位来自印度的麻醉医生紧接着这个话题说开了，她认为之所以会出现这些问题，是因为外科医生不尊重麻醉医生。她说，在印度，外科医生会对麻醉医生大喊大叫，而且对麻醉医生提出的安全警告不屑一顾。看到这样的场景后，很多医学生决定自己以后不当麻醉医生。而麻醉往往是手术过程中风险最大的一部分，训练有素的高水平麻醉医生却并不多见。来自印度的麻醉医生得出那样的结论不只是因为她的个人经历，她还去过很多国家，在那些地方也看到了相似的情况。不过，来自爱尔兰的一位护士认为，麻醉医生的待遇还不是最差的，在手术室里护士更不受尊重，医生往往不把护士当成团队的一分子，不把他们放在眼里。他们提出的警告也往往被忽略。



一方面，与会人士一致认为，手术对于挽救世界各地人们的生命是非常重要的，应该尽可能多地增加人们接受手术治疗的机会。即使是在最糟糕的医疗条件下，手术也常常能够挽救生命。而且在世界很多地方，术后严重并发症的发病率不算太高，大概在5%~15%之间。



但另一方面，即使术后并发症的发病率不高，但这一比例也是让人难以接受的，毕竟每个百分点意味着有数百万人因为手术而残疾或死亡。仅仅是在美国进行的研究就让我们看到，至少有一半的术后并发症是可以避免的。不过导致并发症的原因实在太多了，我们需要采取行动，但究竟应该从何做起呢？

有人提议开设更多的培训项目。但话一说出口，他们就觉得这个方案行不通。如果这样的问题在每个国家甚至每家医院都有，那我们就没有办法开展范围如此之广的培训来改变现状。我们既没有钱，又没有精力来办成这件事情。

我们讨论了激励计划。美国也曾试验一些与绩效挂钩的激励计划。这些计划要求，医生如果一直按照标准程序进行诊疗，就能获得奖励。但他们如果不这么做，就会受到惩罚。虽然这个办法有一点效果，但效果并不明显。美国最大的一项此类计划，只是将术后并发症发病率降低了2%~4%。此外，这一举措的效果也不容易测量。医生往往是自己报告结果，所以数据的准确性无法保障。不仅如此，结果的好坏还常常与

病人初始病情的严重程度有关。一些医生所进行的手术，其术后并发症发病率显著下降，但这可能是由于他后来治疗的病人病情没有先前那些病人那么严重。所以，这个办法将耗费大量财力却收效甚微。我们实在无法想象在全球范围内实施这一举措的可行性。

我们能够想到的最简便的方法，就是以世界卫生组织的名义制定和发布一套手术安全标准。专家委员会经常这么做。这些指导方针往往无所不包，从预防手术感染的具体措施到促进培训和协作的具体方法，什么都有。这将成为我们的手术安全“日内瓦公约”和防止手术室混乱状况的“赫尔辛基协定”。但是，你只需在世界卫生组织总部地下室那个昏暗的过道里走一走，便会开始质疑这个计划。

我在地下抄近路往来于各幢建筑物之间的时候，发现地下室里放着许多托架，上面摆放的是一本本厚重的两百多页的指导手册，堆放得很高。这些手册是由一些专家小组精心编制的，里面的内容无所不包。有的是关于如何预防疟疾的，有的是关于如何治疗艾滋病的，还有的是关于如何控制流感的。这些手册都塑封好堆放在那里，上面积了厚厚的一层灰尘。毫无疑问，专家们为这些标准付出了很多心血，他们一定进行了许多思考，动了很多脑筋。其中的一些标准，的确有可能让相关的诊疗实践水平获得提升。但在大多数情况下，它们只是被分发到了世界各地。而在曼谷、布拉柴维尔、波士顿和布里斯班的医院里，医生们依然我行我素，对这些标准不理不睬。

我曾经向一位世界卫生组织官员提出如下问题：世界卫生组织是否为成功实施全球性的公共卫生项目制订过指导手册？这位官员看了我一眼，她的眼神让我觉得自己像个无知的孩童，我的行为就如同一个孩子在狗嘴里寻找发出狗叫声的东西一般：这个想法很可爱，但十分愚蠢。

不过我还是四处寻找启示。我向世界卫生组织各部门的官员打听可以借鉴的公共卫生干预的成功案例。他们告诉我一些范例，一个著名的成功案例是天花疫苗全球接种运动。这项运动于1979年开始实施，并让全世界的人们从此免受天花的祸害。另一个著名的成功案例发生在1854年的伦敦，约翰·斯诺（John Snow）成功发现了一场霍乱疫情的暴发源头。

那年夏天，霍乱在伦敦的一个居民区暴发，在疫情开始后的3天时间里就有200多人死亡。人们惊恐万分，大部分居民纷纷逃离家园。但是疫情并没有结束，而是进一步恶化。第二周，又有500多人不幸病死。当时人们普遍认为，像霍乱这样的疾病是由瘴气造成的，但斯诺对这一理论提出了质疑。他在地图上——标出所有病人的住址，发现这些记号散布在一个水源周围，那就是索霍地区（Soho）博德大街（Broad Street）的一口水井。他询问了死者的家属，调查了死者的生活习惯，并对可能的致病原因进行了详细的统计分析。最终，斯诺得出结论：是受污染的水源导致了霍乱的暴发。人们后来发现水井的附近有一个泄漏的化粪池。斯诺劝说当地政府移除了这口井的水泵把手。这样，人们就无法再饮用这口井的水了。此后，疫情渐渐平息下来。正

是斯诺医生创立了流行病学调查的基本方法，直到今天，这些方法依然被广泛使用。

我发现这些成功案例有一些共同之处：它们所采取的干预措施非常简单，如接种疫苗和移除水泵把手，这些举措的效果能被精确地测量，而且非常显著。用财务分析师的话来说，它们的投资回报率非常高。而在阿基米德看来，它们就如同杠杆一样。

切中要害的“关键点”在哪里

简单、可测、高效这三种特征让我想起了一项公共卫生研究，这是我最喜欢的研究之一。此项公共卫生计划是由美国疾病控制中心和巴基斯坦一个名叫希望（HOPE）的慈善组织共同实施的。这个项目旨在降低卡拉奇贫民窟儿童过高的夭折率。卡拉奇四周的贫民窟居住着400多万人，那里可能是世界上人口密度最大、居住条件最差的地方了。多年的贫穷和食物短缺导致这个地区30%~40%的儿童营养不良。街道污水横流，几乎所有水源都被污染了。每10个孩子中会有1个活不到5岁，主要死因在于腹泻和急性呼吸道感染。

这个问题不是一时半会儿可以解决的，而且导致这一问题的原因有很多，除了供水系统和排污系统不完善以外，当地居民不识字也是一个重要原因，他们学习基本卫生知识非常困难。由于当地政治不稳定，腐败问题、官僚问题严重，所以很少有人愿意在那里投资兴业，当地居民也就很难找到工作，很难挣到钱来改善自己的生活条件。全球农产品价格持续走低，这让很多农民无法维持生计，成千上万的人涌向都市寻找工作，这更加剧了贫民窟的拥挤程度。在这种情况下，除非社会和政府发生自上而下的重大变革，我们很难想出办法来改善当地儿童的卫生状况。

终于，一个年轻的公共卫生工作者想出了一个主意，此人便是在内布拉斯加州奥马哈长大的斯蒂芬·卢比（Stephen Luby）。他的父亲在克雷顿大学（Creighton University）妇产科任教，他自己曾经在得克萨斯大学西南医学中心学习。不知是出于什么原因，他总是对公共卫生问题非常感兴趣。后来，他在疾控中心找到了一份工作，负责对南卡罗来纳州暴发的疾病进行流行病学调查。当美国疾控中心在巴基斯坦设立办公室并招募员工的时候，他毫不犹豫地报了名。后来，他和身为教师的妻子一起来到了卡拉奇。卢比于20世纪90年代后期开始发布有关当地卫生医疗状况的调查报告。

我曾经问卢比他是怎么想出这个主意的。他说：“如果当地也有像奥马哈那样的供水系统和污水系统，那么我们就能够解决问题。但是，你可能要等上几十年时间，这些重大的市政工程才会完工。”所以，他开始寻求技术含量低的解决方法。他想出来的方法非常不起眼。在他同事的眼里，这个想法甚至非常可笑，那就是使用香皂。



卢比打听到一个消息：日用消费品巨头宝洁公司刚刚研发出一种新型的抗菌香皂，他们迫切地想要证明这种新产品的功效。虽然卢比的同事对他是否能够成功没有信心，但他还是说服了宝洁公司提供研究资金，并且提供成箱含有和不含三氯卡班消毒成分的香皂。卢比在卡拉奇的贫民窟里随机挑选了25个居民区，来自“希望”慈善机构的志愿人员每周都会在这些地方走街串巷，发放香皂，其中一些是含有新型消毒成分的香皂，另一些则是一般的香皂。他们鼓励人们在6种情况下使用香皂：每天洗澡时、每次大便之后、擦拭婴儿的时候、做饭之前、吃饭之前，以及给他人喂食之前。随后，志愿人员在测试地区收集当地儿童疾病发病率的相关信息，再收集11个控制组地区的相关信息进行参照对比。而志愿人员在这11个地区并不分发香皂。

《柳叶刀》（*Lancet*）杂志于2005年发表了一篇里程碑式的论文，卢比和他的研究团队在这篇论文里报告了他们的研究成果：在一年时间里，住在试验居民区的家庭平均每周能够得到3.3块香皂。在这段时间里，无论试验区的居民使用何种香皂，当地儿童的腹泻发病率都比控制组儿童低**52%**，肺炎发病率下降了**48%**，而脓疱疹这种皮肤细菌感染的发病率下降了**35%**。他们取得的结果让人感到非常震惊。要知道，当地的穷人不识字，居住地十分拥挤，而且不管他们使用何种香皂，他们的饮用水和清洗用水都是受到污染的。

具有讽刺意味的是，宝洁公司认为研究结果非常令人失望。因为添加了消毒成分的香皂并不比普通香皂更加有效，普通香皂就能

达到非常理想的效果。或许正因为贫民窟各方面的条件都非常糟糕，使用普通香皂的效果更加让人啧啧称奇。

卢比告诉我说，这个项目的成功秘诀就在于，香皂不仅仅是香皂，它还改变了人们的行为。因为志愿者并不只是分发香皂，他们还通过分发宣传册或口述向人们说明应该在哪些情况下使用香皂。这一点是至关重要的。如果你仔细查看数据就会发现一个惊人的事实：无论是在试验组居民区还是在控制组居民区，在研究刚开始的时候，每个家庭平均每周的香皂使用量并不是零，而是每周两块，也就是说，他们早就开始使用香皂了。

那么，这项试验到底改变了什么呢？卢比告诉我，有两件事情发生了变化。首先，人们不再因为经济拮据而少买香皂。住在那里的人们并不觉得香皂很贵，大多数家庭都会购买香皂。但是光购买还不够，我们希望他们能够经常使用香皂。这项试验让当地的贫困居民不再为香皂而过多考虑经济问题。其次，这个项目让人们学会了更加正确地使用香皂。这一点也是非常重要的。

在巴基斯坦、孟加拉国等南亚地区，卢比和他的团队对人们的洗手行为进行了研究。他们发现，几乎每个人都会在大便后洗手。他说：“南亚的人民很爱干净。”即使洗手的地方离得很远，在80%的情况下，人们还是会跑过去洗手。这一数据会让大多数在机场如厕的人感到汗颜。但研究人员发现，当地居民洗手往往不够充分，他们要么洗得太快了，要么只在大便后洗手，或者是用灰或泥土，而不是用清水和香皂进行清洗。

卢比的这项试验的确改变了人们的行为。志愿者在当地教人们如何正确洗手，教他们要把手充分打湿，均匀地涂抹香皂，用较长时间搓手，最后还要把香皂泡沫冲洗干净，而且不要用衣服来擦洗干净的手。志愿者还教当地居民，在他们原本不习惯洗手的一些时候也要洗手。

卢比解释说：“当地的人们本来并不习惯在做饭之前，或在给婴儿喂食之前洗手。”香皂本身也是一个关键因素，卢比说：“宝洁公司提供的香皂非常好闻，而且比当地人平时购买的香皂更容易起泡，人们很喜欢用它来洗手。跨国公司真的很注重顾客的消费体验。而公共卫生部门的官员往往不注重这一点。”还有，人们喜欢从志愿者那里获得免费的香皂，他们将其视为礼物，而且志愿者们不会对他们指手画脚。当地居民收到的不只有礼物，还有能够改善他们生活质量并大幅减少疾病发

生的基本理念。

促进有效的执行与有效的分权

在回顾这次试验的时候我欣喜地意识到，这不只是一项关于香皂的研究，还是一次关于清单的研究。我想知道的是，在解决手术安全问题的时候，清单是否也能像香皂那样简单、廉价、有效，而且能广为传播。手术安全问题涉及很多方面，虽然要开发出简单有效、能够在全球范围内解决这一问题的清单还需要很长时间，我也不确定是否真的能够开发出这样的清单，但是当这一建议在会上被提出的时候，其他与会人士却做出了积极的响应。

有人提到了哥伦布儿童医院（Columbus Children's Hospital）的成功做法。他们开发出了旨在减少手术感染的清单。感染是儿童手术中最常见的并发症之一，防止感染最有效的方法除了医护人员正确掌握消毒技巧，还有在病人皮肤被切开前的**60**分钟内为病人注射预防性抗生素。

注射的时机是非常关键的，待医生下刀后再注射就晚了，但如果注射得太早，等到手术开始的时候，药力就会减退。研究显示，如果抗生素注射的时机恰到好处，那么仅仅是这一项举措就能将感染的风险降低一半。就算是在切开病人皮肤30秒之前匆匆为其注射抗生素，人体的循环系统也有足够时间将药物输送到即将被切开的组织中。

但是，这个关键步骤常常被忽略。2005年，哥伦布儿童医院对各项手术记录进行了检查。他们发现，有1/3以上患儿在接受阑尾切除术时，医生没有在正确的窗口期为其注射抗生素，有的注射得太早，有的注射得太晚，有的则根本没有注射。

这看起来是非常愚蠢的错误，给病人注射抗生素能难到哪里去呢？即便是医学院的学生都不会认为医院会在这种地方犯错，但事实确实如此，在病人被推入手术室的时候，医护人员往往手忙脚乱，最可能被他们忽略的就是这种不起眼的步骤。

给病人注射抗生素应该是麻醉医生的职责。而在那个时候，他们最关心的是如何让病人安全而平静地入睡。对于一个赤身裸体、躺在冰冷的手术台上的8岁儿童来说，这可不是一件简单的事情。因为手术室里满是陌生人，这些人将要切开他的身体，这会让他感到十分害怕。手术室里的意外状况和干扰因素还有很多，比如，设备会发生故障（“那个

红色警示灯应该这么一闪一闪吗？”），再比如病人突然哮喘发作，或者急救医生传呼主刀医生，让他马上赶到急救室。当这些情况发生的时候，像注射抗生素这样的步骤就很容易被遗忘。



哥伦布儿童医院的外科主任不仅是一位小儿心脏手术专家，而且曾经是一名飞行员，他决定借鉴航空业的做法来解决这一问题。他专门为外科医生设计了一份“准许起飞”清单，并将其贴在手术室的白板上。清单上列出的内容非常简单，在手术开始之前，护士要督促手术团队执行检查程序，并口头确认病人的身份和手术具体部位，这些是医护人员在手术开始前本来就要做的事情。不过，这张清单还包括了一个检查项目，那就是确认已经给病人注射了抗生素，或者手术团队已经确认病人没有必要使用抗生素。

这张清单上就是这些内容。但是，要想让手术团队停下来执行检查程序，并将其变成一种习惯，却不是一件容易的事情，仅仅是设置这些检查项目并不会起到多大的作用。所以，这位外科主任将外科医生、麻醉医生和护士召集起来，向他们详细说明这么做的原因。

他还做了一件看似奇怪的事情。他设计了一个金属小帐篷，上面刻着“准许起飞”这几个字，并且将其放置在手术器械套件中。这个帐篷有15厘米长，正好可以把手术刀盖在下面。护士必须在手术开始前将这个帐篷放置在手术刀的上面。这样，手术团队就会记得执行检查程序。还有一件事情也非常重要，在护士同意手术开始，并将帐篷拿开之前，外科医生不能下刀。这是一个微妙的变化，一张小小的清单起到了分权的作用。

哥伦比亚儿童医院外科主任对清单的效果进行了评估。3个月以后，医护人员为89%接受阑尾切除术的患儿在手术前及时注射了抗生素。10个月之后，这一比例达到了100%。使用清单成了一种习惯，而且大家对其中的一点达成了共识，那就是在必要的步骤完成之前，任何一个手术团队成员都可以叫停手术。

没有一张清单能应对所有的意外

虽然被这个故事深深吸引，但我还是心存疑虑。不错，清单的使用的确让这家医院在手术实施过程中的某个方面得到了显著改善，我很愿意相信他们的手术感染率因此而显著下降。但是，如果我们想让术后并发症的整体发病率也显著下降，那么就必须找到一种综合性的方法，以避免医护人员在手术过程中犯各种各样的错误。

最后，多伦多大学外科医学系主任理查德·雷兹尼克（Richard Reznick）发了言。他说，他的医院开发了一种范围更广、包括21个检查项目的手术清单，并且完成了这一清单的可行性测试。这张清单旨在帮助外科手术团队避免手术过程中可能发生的各类错误。医护人员要对各个检查项目逐一进行口头确认，如是否给病人注射了预防性抗生素，血库是否准备好了病人可能需要的血袋，重要的扫描和化验报告是否在手术室里，以及特殊的医疗设备是否已经就位等。

这张清单还规定，手术团队在手术开始之前必须进行“团队简报”。他们要花一些时间就一些重要事项进行交流，如手术的预计时间有多长，病人的失血量可能是多少，病人是否有其他疾病或需要注意的事项。

雷兹尼克从未听说过建造大师退出历史舞台的故事，但是他本能地想到了建筑行业的解决方法，即通过任务检查和沟通检查来应对越来越复杂的问题。对此，其他一些专家也表示了赞同。

一位来自约翰斯·霍普金斯大学的胰腺外科医生马丁·马卡里（Martin Makary）给我们看了一张含有18个检查项目的清单。在过去的5个月中，他们医院有11名医生参与

了这张清单的测试。还有一些专家来自加利福尼亚州的凯泽（Kaiser）医疗集团，他们开发的清单有30个项目，比多伦多大学和约翰斯·霍普金斯大学的清单诞生得更早。但是，上面提到的这些清单都遵循相同的设计理念。

世界各地手术死亡的主要原因不外乎4种，分别是感染、出血、不彻底的消毒和意外。对于前3种致死原因，科学和经验已经为我们提供了一些简单而有效的预防措施。我们自认为这些举措一直实施得很好，但事实并非如此。不过，这只是一些简单的错误，经典的清单对于它们非常有用。所以，上述清单都包含了针对这3种致死原因的检查项目。

但第4种致死原因，也就是意外，与前3种截然不同。这类错误源于手术本身的复杂性。毕竟，我们要给病人开膛破肚，去除他们体内的病灶。不过，这些清单的开发者似乎都清醒地意识到一点：没有一张清单能够应对所有意外情况。所以他们一致认为，当此类情况发生时，最有效的解决方法是让医护人员停下来，进行充分的交流，让他们像团队一样，共同对每个病人面临的独特潜在威胁做出判断。

清单宣言 | THE CHECKLIST MANIFESTO HOW TO GET THINGS RIGHT

团队在开始工作之前，必须进行“团队简报”，就一些重要事项进行交流。

也许这一切看起来都是明摆着的事情，但清单的使用意味着传统手术实施方式发生了重大变革。人们以前把手术看作个人表演，外科医生就像钢琴演奏家一样，是演奏大师。所以手术室在英语里不仅叫作“operating room”，还可以叫作“operating theater”（剧院）。后面这种叫法不无道理，因为手术室就是外科医生的舞台。外科医生在灯光下走上舞台，病人静静地躺在手术台上，辅助人员各就各位，大师的精彩演出就要开始。

现在，外科手术变得越来越复杂，外科医生认为自己也随着手术的发展而不断进步。我们自认为手术室里的医护人员已经能够像一个真正的团队一样互相协作，但令人尴尬的是，研究发现，其他手术团队成员常常不清楚病人可能会面临什么风险，不清楚自己应该为哪些可能出现的问题做好准备。他们有时甚至不清楚为何要进行某一台手术。在一项调查中，研究者对300位刚刚进行完手术的医护人员进行了访问。结果

每8个被访者中就有1人报告说，直到主刀医生下刀时，他们才知道手术的具体部位在哪里。

布莱恩·塞克斯顿（Brian Sexton）是约翰斯·霍普金斯大学心理学的领军人物。他所做的一系列研究揭示了这样一个严酷的现实，那就是现在的手术团队距离成为真正的团队还有很长的路要走。在一项研究中，他对来自5个国家的1 000位手术室医护人员进行了调查，这些国家分别是美国、德国、以色列、意大利和瑞士。研究发现，虽然有64%的被访外科医生认为所在团队的合作质量很高，但只有39%的麻醉医生和28%的护士也持一致观点，麻醉科住院实习医生也这么想的比例更小，只有10%。塞克斯顿还发现，每4个外科医生中就有1人认为，资历较浅的团队成員没有资格质疑资历较深的外科医生所做出的决定。

清单宣言 | THE CHECKLIST MANIFESTO HOW TO GET THINGS RIGHT

团队进行有效合作的阻力来自“事不关己，高高挂起”的消极态度。细致的分工让团队成员只关心自己手上的事情，而对其他成员碰到的问题不闻不问。

团队进行有效合作的最大障碍其实不是偶尔暴跳如雷、挥舞着手术刀、让人胆战心惊的外科医生。不错，有些外科医生的确做得比较过火。但实际上，团队进行有效合作的更大阻力来自“事不关己，高高挂起”的消极态度。细致的分工让团队成员只关心自己手上的事情，而对其他成员碰到的问题不闻不问。殊不知，无论是在手术即将开始的时候，还是满载乘客的飞机滑向跑道的时候，或者是数百米高的摩天大楼拔地而起的时候，这种态度都是非常危险的。而这个问题在医疗行业尤为突出，我在自己的手术室里就碰到过。

几年前，当我还是一名实习医生的时候，在一次手术中，一位资深的外科医生突然对我身边的另一名住院实习医生大发雷霆，因为这位实习医生对他的手术计划提出了质疑，所以他命令这个没大没小的家伙离开手术台，站到墙角去好好面壁思过。但这位实习医生拒不从命，所以那个脾气暴躁的前辈就把他赶出了手术室，还以犯上为理由，劝说院方给予那个实习医生停职处分。

对于某些工作来说，团队合作可能的确是非常困难的。在极端复杂的情况下，我们不得不依靠分工和各自掌握的专业技能。就拿手术室来说吧，你能在那里找到主刀医生、助理医生、助理护士、巡回护士和麻醉医生。他们在各自的专业领域里都是大师，每个人都接受了多年训练，这些训练就是要让他们成为专家。但是，我们不应该割裂地看待各

项任务，认为只要把自己分内的工作做好就行了，而应该为了更好地实现团队的目标贡献自己的力量。这就要求我们想方设法把看似独立的任务有机地衔接起来，并不断学习用团队协作的方式来应对各种可能发生的问题。

我曾经认为和谐的团队合作是可遇而不可求的。我偶尔会碰到这样难得的经历，每个人都开足马力，通力合作，积极面对各种困难，整个团队就像一个人一样。

清单宣言 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

我们不应该割裂地看待各项任务，认为只要把自己分内的工作做好就行了，而应该为了更好地实现团队的目标贡献自己的力量。

THE
CHECKLIST MANIFESTO
How to Get Things Right
✓ **清单革命在行动·医疗行业**



有一次，我为一个80岁的老人实施急救手术。在此之前一周，他刚刚做完心脏手术，而且恢复得很好。但在前一天晚上，他的腹部突然开始剧烈疼痛。到了早上，疼痛感进一步加剧了。我到普外科病房探视这位病人，只见他躺在病床上，身体因为疼痛而蜷作一团。他的心跳超过每分钟100次，而且心律不齐，血压也已经开始下降。每当我触碰到他腹部的时候，他都疼得好像要从床上蹦起来一样。

他很清楚自己出了大问题，但他似乎并没有感到害怕。

“我们该怎么办？”他咬着牙问我。

我告诉他，很可能有一个血凝块阻断了他的肠道动脉供血。这种情况就和中风一样，只不过被切断供血的不是脑部，而是肠道。

没有了新鲜血液，他的肠道会坏死，还会破裂。如果不动手术的话，命就保不住了。但就算是动了手术，也未必能挺过去。这类手术的存活率只有50%左右。即使他能走下手术台，等待他的也可能是很多严重的并发症。他可能需要用呼吸机呼吸，或用插管进食。这个病人已经动过一次大手术，身体非常虚弱，而且年事已高。我问他是否愿意再动一次手术。

他给出了肯定的回答，但希望我能先和他的妻儿取得联系。我打电话通知了他们，他们也同意动手术。于是，我打电话给调度台说明了情况，让相关人员马上为我准备好手术室并召集手术团队，找谁都可以，只要把人凑齐了就行。

在一个小时之内，病人就被送进了手术室。当医护人员在那里集结并准备开始工作的时候，我觉得一个真正的团队正在形成。巡回护士杰·鲍尔斯病人做了自我介绍，并向病人说明在场每一个人的职责。手术助理护士史蒂夫已经穿上了手术衣，戴上了无菌手套，站在无菌手术器械旁做好了准备。高级麻醉医生熊郅正在和他指导的住院实习医生索尔进行交流，他们一边准备药物和器械，一边检查手术计划是否有问题。实习外科医生乔亚基姆手拿导尿管站在我的身旁，待病人入睡后，他会立刻将导尿管插入病人的膀胱。

时钟嘀嗒嘀嗒地走着，我们耽搁得越久，病人坏死的肠子就会越多，病情会变得越重，而他的生存希望也会随之变得越渺茫。在场的每一位医护人员都清醒地意识到了这一点，这不是一件容易做到的事。真正做到换位思考，急病人之所急，想病人之所想，并没有说的那么简单。但我的手术团队的确做到了，他们行动敏捷，有条不紊，步调一致。眼下的这台手术非常困难，但没有什么困难能够阻挡我们。

这个病人个头很大，脖子很短，肺容量很小。所以，在麻醉后给他插入呼吸导管并不容易。但郅已经提醒过大家可能会发生此类问题，大家都做好了预案，他和索尔将可能需要的器械也准备妥当了。乔亚基姆和我用手术刀切开了病人的腹部，我们发现他右侧的结肠已经变黑坏死了，但并没有破裂，剩下3/4的结肠和所有小肠看起来似乎没有什么问题。这实在是个好消息，因为受损范围不大。但在我们切除病人右侧结肠的时候却发现，剩下的结肠情况也不妙。

正常的结肠看起来应该是健康的粉色，但我们发现病人的结肠上散布着硬币大小的紫色斑块，看来阻塞右侧结肠主动脉的血凝块也阻塞了给左侧结肠供血的动脉分支。所以我们不得不把病人1米多长的结肠全部切除，并实施肠造口术，把他的肠子接入一个用来盛放排泄物的袋子。史蒂夫事先已经想到了这种可能，他让杰拿来我们需要的牵开器。乔亚基姆轻推了我一下，提醒我把病人腹部的切口开得更大一些。我所进行的每一个操作都得到了他的全力协助。我们将病人阻塞的动脉血管一段一段地钳住、切除和结扎。病人的每一处创面都在不停地流血，因为坏疽制造的毒素破坏了血凝功能。但是郅和索尔早就做好了准备，不停地为病人输液输血。实际上，自手术开始以来，病人的血压已经回升了许多。当我提到病人需要接受重症监护的时候，郅说他已经做好了安排，并把情况通知了重症监护医生。

由于我们通力合作，而不是各自为政，这位病人活了下来。我们只花了两个多小时就完成了这台手术。病人的重要体征都十分稳定，几天后就可以出院了。病人家属把功劳归于我一个人，但我心里很清楚，少了我的团队我什么也干不成。整台手术就像一曲交响乐一样充满了和谐之美。

那一次，我们的确进行了有效的团队合作。但我还是认为这一切是可遇而不可求的，是天意让我们这些人那天下午凑到了一起，并发生了奇妙的化学反应。虽然我和郅经常一起做手术，但我和杰以及史蒂夫已经数月没有合作了，乔亚基姆也已经很久没有和我同处一个手术室了，而我和索尔只合作过一次。在此之前，我们这6个人从未作为一个完整的团队实施过一台手术。在任何一家大医院里，这种情况并不少见。我所在的医院有42个手术室，相关的医护人员有1 000多人，你总是能够在团队中找到新面孔。虽然团队合作对于手术的成功非常重要，但能否有效做到的确是很难预测的。不知是何原因，当我们这6个人凑到一起的时候，一切就这么自然而然地发生了。这实在是一件妙不可言的事情。

激活“团队合作”的检查项目

虽然我觉得只有足够幸运才能碰到可以进行有效合作的团队，但上文提到的那些清单的设计者却创造出一种有意为之的可能。他们坚持认为：医护人员应该在手术前就手术的具体情况进行交流，至少要在手

术开始前花一分钟时间进行简短的讨论，他们认为这是促进团队合作的一个好方法。橄榄球运动员之所以在每次开球之前聚在一起商讨战术，也是出于相同的原因。这些清单上还列出了一个有助于促进团队合作的检查项目，那就是医护人员在手术开始之前应该熟知彼此的姓名，这让我感到非常好奇。

对于这项内容，约翰斯·霍普金斯大学研发的清单写得最为明确：新的团队在手术开始之前，每一个团队成员必须介绍自己的姓名和职务，比如，“我是主刀医生阿图·葛文德”，“我是巡回护士杰·鲍尔斯”，“我是麻醉医生熊郅”。

开始的时候，我觉得这个做法有点矫情，并对其实际效果深表怀疑。但事实证明，这个步骤是经过精心设计的，各种心理学研究都支持这样一个不言而喻的结论：团队成员不知道彼此姓名时，往往不能很好地合作。而约翰斯·霍普金斯大学的心理学家塞克斯顿所做的研究说明，这一结论同样适用于手术室。

清单宣言 | THE CHECKLIST MANIFESTO HOW TO GET THINGS RIGHT

清单上要有一个有助于促进团队合作的检查项目，比如熟知彼此的姓名，因为团队成员不知道彼此姓名时，往往不能很好地合作。

在一项研究中，他和他的团队硬着头皮对刚刚进行完手术且疲惫不堪的医护人员进行调查。每个被访者都被问及两个问题：在刚刚进行完的这台手术中，你对团队成员之间的沟通情况如何评分？其他团队成员的名字叫什么？研究发现，在半数手术中，团队成员不清楚彼此的姓名。但是，当团队成员知道彼此姓名的时候，他们对于沟通状况的评分显著提高了。

约翰斯·霍普金斯大学等大学的学者还发现，当护士有机会在手术开始之前介绍自己的名字，并提醒大家注意手术中可能遇到的风险时，他们会更加主动地在手术中注意各类问题，并提出解决方案。研究者将这种情况称为“激活现象”。一开始就让人有机会发言，能够提高他们主动参与和表达意见的积极性，能够增强他们的责任感。

这些研究的调查范围有限，得出的结论还不具有普适性，但是这些初步的调查结果让我看到了希望。在此之前，要想提高外科医生的能力

并大幅度减少手术对病人造成的伤害，只能依靠加强外科医生的专业训练和增加他们的实践经验。但在一些城市里，手术团队却对这些不同凡响的清单进行了测试，而且都取得了令人欣慰的结果。

清单的力量 | THE CHECKLIST MANIFESTO | HOW TO GET THINGS RIGHT

约翰斯·霍普金斯大学的研究人员测量了他们开发的清单对促进团队合作所起到的作用。共有11个外科医生同意参加试验，其中包括7名普外科医生、2名整形外科医生和2名神经外科医生。3个月后，在他们带领的手术团队中，认为团队合作状况十分理想的人数比例从68%跃升到92%。

加利福尼亚州凯泽医疗集团的研究人员也对他们开发的清单进行了为期6个月的测试，总共涉及3 500台手术。在此期间，手术团队成员对于团队合作氛围的评分均值由“好”变成了“非常好”，员工的工作满意度上升了19%，手术室护士的离职率从23%下降到了7%。

清单还让医护人员避免了许多错误。比如在一次执行检查程序的过程中，手术团队发现，一瓶氯化钾被换成了抗生素。这一错误如果没有及时纠正，病人可能会因此而丧命。还有一次，手术调度室把手术的类型写错了。手术团队得到通知要为病人实施胸廓切开术，这种手术的切口非常大。但实际上，病人需要接受的只是胸腔镜检查，这类手术的切口不到一厘米长。

而在多伦多，研究人员在手术室里实地观察了手术清单的使用效果。虽然他们只观察了18台手术，但在其中的10台手术中，清单让手术团队发现了严重的问题，比如没有在术前为病人注射抗生素，预防意外情况的血袋没有准备好。清单不止一次发现了这些问题。此外，还有一些与病人独特病情相关的具体问题竟然也被发现了。我真的没想到清单能够起到这样的作用。



多伦多研究人员的报告中提到了一个使用脊髓麻醉的腹部手术。在进行这类手术时，哪怕病人只是觉得有一点点疼痛，他也应该及时向医护人员报告这一情况，因为这可能意味着麻醉失效了，麻醉医生需要补充麻醉剂。但这个病人有严重的神经疾病，无法用语言进行沟通，只能用手比画进行表达。在手术中，我们一般会固定住病人的手和手臂，以防止病人不经意地用手抓手术单，或触碰外科医生和手术部位。但在这台手术中，如果病人的手和手臂也被固定住的话就会出大问题。在手术开始之前，手术团队并没有意识到这个问题。外科医生走进手术室，穿上手术衣，带上无菌手套，走到手术台前。因为需要执行检查清单，他没有直接开始手术，而是停下来和其他成员商讨当天的手术计划，报告里记录了他们的一段对话。

“在麻醉方面有什么问题需要注意吗？”外科医生问道。

“有一个问题，病人有发音障碍。”麻醉医生答道。

外科医生想了想说：“这可能会让我们无法了解他的神经功能状况。”

麻醉医生对此表示同意，并说道：“我和病人商量好了用手势进行沟通。”

“这样的话，我们就不能固定他的手臂。”外科医生说。麻醉医生点了点头。手术团队随即想出了一个办法，这个办法既能让病人的手臂自由活动，但又不会让他抓到手术单。

“我还担心手术室里的人太多了，”麻醉医生继续说道，“噪声和人员走动可能会干扰我们与病人的沟通。”

“我们可以要求在场的人保持安静。”外科医生说。问题就这样

解决了。

清单冗长和含糊不清的代价

这些研究还不足以说明清单能够实现世界卫生组织制定的宏伟目标，那就是以较低成本大幅度降低术后并发症的发病率，而且项目的效果还要能够被准确测量。但是，在这次日内瓦会议结束的时候，我们一致认为，值得进行一次大规模的手术安全检查清单测试。

一个工作小组负责将经过测试的不同清单合并成一张清单。这张清单有3个检查点，每到一个检查点，手术团队就必须停下来完成相关的检查程序，然后才能继续。这3个检查点分别是实施麻醉前、开刀前以及手术结束后患者离开手术室之前。工作小组根据实施时间节点的不同，将有关过敏、抗生素、麻醉设备等的检查项目分别归于不同的检查点之下。他们把所有能够想到的可以提高手术安全性的其他检查项目也都加了进去。他们还在检查表中设置了沟通检查项目，手术室里的每个人都要在手术开始之前知道其他成员的姓名和职责，而且每个人都可以对手术计划和可能发生的问题发表意见。

我们决定先在全球各地的一些医院进行小规模测试性研究，世界卫生组织承诺为这项研究提供资金。我感到非常兴奋，对试验的结果非常乐观。回到波士顿以后，我迫不及待地想要亲自测试一下这张清单了。



我把清单打印出来，并带到了手术室，还把自己在日内瓦获得的启示告诉了护士和麻醉医生。

我说：“我们能不能试试这张了不起的清单？”清单的内容十分详细，从设备检查到抗生素注射再到术前讨论，无所不包。手术团队的其他成员用怀疑的眼光看了看我，但他们并没有提出反对意见。“行啊，不管你说什么，我们照做就是了。”这不是我第一次想出稀奇古怪、滑稽可笑的主意了，他们对此已经习以为常。

我把清单交给了巡回护士迪，并让她在恰当的时候进行第一部分检查。15分钟后，我们将要对病人实施全身麻醉，于是我说：“等一下，清单呢？”

迪说：“我已经完成啦。”她把清单递给我看，所有检查项目都打上了钩。

“不不不，每个检查项目都需要经过口头确认，这是一张团队清单。”我说。

“哪里有写？”迪问道。我又看了看清单，她是对的，上面并没有写清楚这一点。

“好吧，还是请你把它念出来。”

迪耸了耸肩，然后开始逐一念出检查项目。但有些项目模棱两可，比如，护士是只需要确认每个团队成员都了解病人的过敏症呢，还是需要说出具体的过敏症名称。这些困惑让在场的每个人都觉得有些不耐烦，手术台上的病人也开始变得躁动不安。

“一切都好吧？”她问道。

我告诉她没有问题，我们只是在执行安全检查程序，并请她放心。

实际上，我自己也被这张冗长且含糊不清的清单给惹恼了。我觉得这张清单已经开始分散我们的注意力，而不是帮助我们更好地关心手术台上的病人了。

那台手术还没结束，我们就停止使用这张清单了。就连在我们自己的手术室里它都不能发挥作用，那就更别指望它能在全球范围内提高手术的安全性了。

第6章 人为根本 清单外的应急反应机制

·曾经发生在哈德逊河上的飞机迫降奇迹挽救了155个人的生命。而事实上，在危急时刻，飞行员能够反应的时间只有60秒。

·清单的力量是有限的。在最危急的情况下，解决问题的毕竟是人而不是清单，是人的主观能动性在关键时刻决定成败。

人们需要应对各种复杂的突发状况

在经历了第一次痛苦的尝试后，我做了一件早就该做的事情，那就是去图书馆查阅有关如何制作飞行清单的资料。虽然建筑行业的清单看起来很不错，但建筑工程往往持续数月之久，而外科手术却是分秒必争。时间问题看来是把清单运用于手术的一大限制。相类似的是，飞行员也会面临紧急状况，时间对他们来说同样是一个巨大的挑战，但飞行清单似乎能够帮助他们应对各种复杂的突发状况。

我曾读过一篇由丹尼尔·布尔曼（Daniel Boorman）撰写的文章，他在波音公司工作，于是，我打电话给他，从而发现他是一名资深飞行员。在过去的20年里，他一直为波音客机编制飞行清单，并参与新机型驾驶舱控制设备的设计。他参与的项目包括波音747-400型飞机及其之后的所有机型。布尔曼是设计波音787梦想飞机驾驶舱的技术骨干，他参与设计的内容包括驾驶控制、显示系统以及清单系统。

他是波音公司“飞行理念”的缔造者之一。每当你乘坐波音飞机出行的时候，都有一整套理念掌控着飞行员驾驶飞机的方式。比如，在飞行各阶段必须完成的动作或检查项目有哪些，哪些是必须亲自完成的，哪些可以交给计算机去完成，在碰到意外情况的时候应该如何应对。必须有人将波音公司的“飞行理念”转化为具体实践，而没有几个人会比布尔曼更胜任这一工作。他的父辈在70多年前就驾驶过配备了第一代飞行清单的B-17轰炸机。而他多年来研究了数千起坠机事故以及相当严重的飞行事故，发现了旨在防止飞行员犯错的飞行科学方法。

我正好要去西雅图，于是就询问布尔曼我是否能够去拜访他，他慷慨地接受了我的请求。到了西雅图后，我驾驶着一辆租来的汽车，沿着城郊宽阔平坦的道路驶向波音公司的总部。一排排低矮的长方形建筑出现在我的眼前，波音公司总部看起来就像研究所一样，周围有一所州立大学，所以波音公司的这些建筑和周围的环境倒是能够融为一体，而在建筑群的后面还有停机坪、跑道和机库。

布尔曼走到门口来接我。他51岁，穿着休闲裤和开领牛津衬衫，十足的飞行员派头。这让他看起来更像一个工程学教授，不太像公司雇员。他带我来到了3-800楼的门前，这是一幢外表平平但功能齐全的办公楼。大楼的门口立着一个展示箱，里面陈列着发黄的照片，照片上有几个穿着银色飞行服的飞行员。展示箱上落满了灰尘，我猜这个箱子可能已经有几十年没人碰过了。试飞部所在的工作区域光线非常明亮，员工的办公隔间是灰褐色的。我们来到位于楼层正中一间没有窗户的会议室，只见一摞摞为各航空公司制作的飞行清单手册靠墙堆放着，相关的航空公司包括全美航空公

司、达美航空公司和联合航空公司等。

布尔曼把一本手册递给我看。手册是用螺旋金属圈装订的，将近200页，还贴着许多黄色的标签。早期的飞行清单只是一张卡片，上面记录了飞机在滑行、起飞和降落阶段飞行员必须注意的事项，和现在的手册简直无法相比。让我感到好奇的是，这本厚厚的手册究竟该如何使用？打开手册后我才发现，里面远不止一张清单，而是有许许多多不同类别的清单。每一张清单都非常简洁，一般只有几行字，用大号的、容易识别的字体印刷。每张清单都是针对特定情况制作的，所以这本厚厚的手册涵盖了飞行各个阶段可能出现的各种状况。

翻开手册后你首先会看到“正常”清单，上面列出了飞行员日常操作所需完成的主要动作及需要注意的重要事项。飞行员在每个操作阶段，如启动发动机前、飞机被推出前、开始滑行前等，都必须按照清单逐一完成各项检查。不过，正常清单总共也就只有3页长，接下来的都是“非正常”清单。设计人员列出了他们能想到的所有紧急情况，如驾驶舱冒烟、各种警示灯亮起、无线电失灵、副机长无法操作、发动机失灵等，并为应对每一种情况设计了正确的操作流程。手册涉及的很多紧急情况飞行员可能一辈子都碰不到，但万一碰到的话，他们就能依靠清单来化解危机。

布尔曼为我详细介绍了一张清单，这张清单讲述了如何处置飞行中前货舱门警示灯亮起这一情况。在飞行途中，如果飞机的前货舱门没有关闭或没有锁好是极度危险的，这种情况一旦发生，前货舱门警示灯就会亮起。布尔曼向我讲述了1989年发生的一起飞行事故，他对此进行了仔细研究。



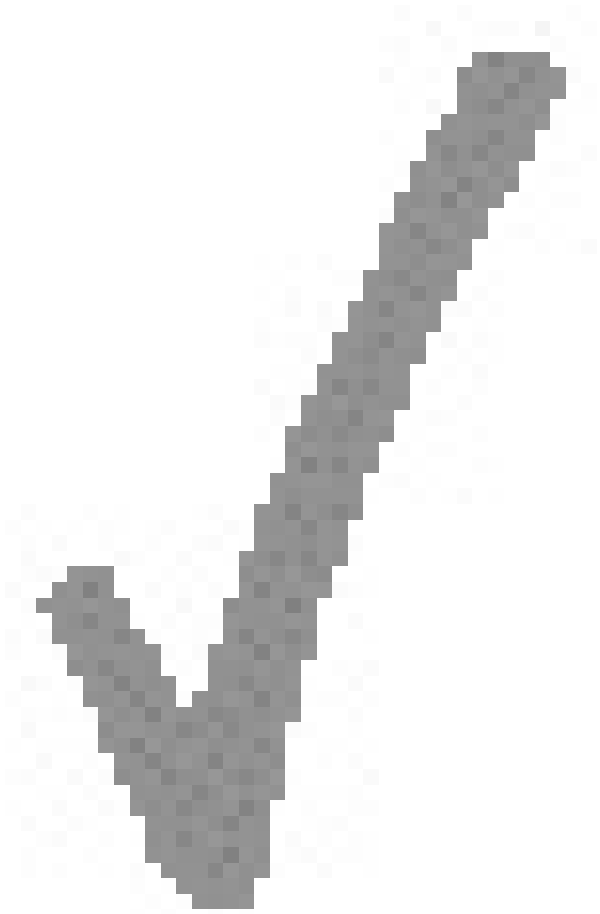
1989年，美国联合航空公司（后简称“美联航”）的一架波音747客机从火奴鲁鲁起飞，飞往新西兰奥克兰市，机上有18名机组人员和337名乘客。在飞行途中，前货舱门的栓销因为线路短路而开启。当时，飞机刚刚爬升到6 700米高度，飞机机舱是加压的，

这样就能维持正常的空气氧气含量。在这样的高度，飞机货舱门的栓销脱开会造成严重威胁：如果舱内的空气顺着货舱门的门缝外泄，那么机舱内外巨大的气压差就会导致外爆。这就好比在打开可乐罐之前用力摇晃瓶罐，在拉开瓶盖的时候，可乐就会喷涌而出。在这起事故中，货舱门瞬间就被炸飞了，上甲板的几个窗户连同商务舱的5排座椅也随即飞了出去，有9位乘客掉进了大海，相邻座位的乘客因为系上了安全带而幸免于难，正在走道里为乘客服务的一位空乘人员也差点被吸出舱外，还好一位身手敏捷的乘客死死拽住她的脚踝，她离机身上被炸开的洞口只有几厘米远。

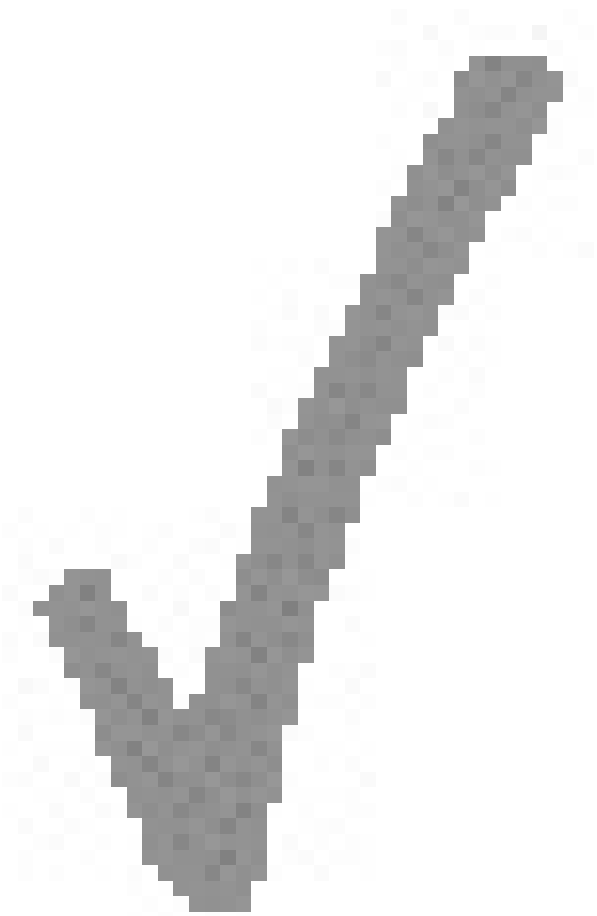
一切发生得太快，飞行员根本来不及阻止灾难的发生。从栓销脱开到货舱门被炸掉，再到乘客被吸出舱外，只过了短短1.5秒时间。后来，波音公司重新设计了货舱门的供电系统，而且增加了额外的栓销，因为这种装置不可能做到万无一失。这样，就算有一个销子脱开了，前货舱门警示灯亮起，那么飞行员也还有时间做出响应。额外的栓销能够为他们争取到防止货舱门被吹走的机会窗口。这时，飞行员就能按照清单进行恰当的处置。

布尔曼解释道，当有栓销脱开的时候，飞行员不应该浪费时间修补舱门，也不应该认为其他的栓销不会脱开，而是应该想办法尽快消除机舱内外的气压差。舱压下降得越多，舱门脱落的可能性就越小。

有一个办法能够很快降低舱内的气压：飞行员可以按下超控开关，打开通气阀门，这样机舱内外的气压差就能够在30秒内消除。不过，这种做法是有问题的。



首先，对乘客来说，座舱突然失压会让人感觉非常难受，特别是耳膜会非常疼痛。婴儿更是无法承受这种剧烈的气压变化，因为他们的咽鼓管尚未发育完全。



其次，在6 000～10 000米的高度降低舱压，就好像把乘客扔到珠穆朗玛峰上一样。由于空气非常稀薄，他们的身体和脑部的供氧会有困难。

美联航的这次事故让我们了解到飞机的货舱门被炸掉是一件多么可怕的事情。在机舱迅速失压后，乘客和机组所要面临的主要威胁就是缺氧。没有被吸出舱外的人虽然能够远离那个直径达3～5米的大洞，但是，舱内的温度会骤降到冰点，空气的氧气含量也会急剧下降，机组人员会觉得昏昏欲睡，意识模糊。感应装置会自动放下氧气面罩，但飞机的紧急供氧系统只能维持10分钟时间。而且，这套系统还可能彻底失灵，这架客机就碰到了这一状况。

舱音记录器让我们了解到了整个事件的细节：

机长：发生什么事情了？

副机长：我不知道。

机长随即通知地面控制塔台飞机出了问题。两秒后，座舱内外就没有气压差了，舱内的空气变得和舱外一样稀薄。

副机长：戴夫，把你的氧气面罩戴上。

机长：好。

副机长：火奴鲁鲁管控中心，这是美联航1号重型，你是否让我们左转？

无线电：美联航1号重型，指令确认。

副机长：现在就转向。

机长：我的氧气面罩里没有氧气。

随机工程师：你现在想让我干什么？

（一段无法辨识的声音）

副机长：你还好吗？

机长：还行。

副机长：你有氧气吗？我们没有任何氧气。

随机工程师：我也没有氧气。

后来进行的事调查发现，爆炸把供氧管路破坏了。机组勉勉强强把飞机降到氧气充足的高度，这真是不幸中的万幸。随后，他们把飞机开回了火奴鲁鲁机场，18名机组成员和余下的328名被吓坏的乘客得以生还。

飞行员从这起事故中吸取的教训非常复杂。如果你在1万米高空飞行，此时前货舱门警示灯突然亮起，那么尽快消除机舱内外的气压差以避免舱门外爆脱落的确是个非常好的主意。但是，通过按动紧急泄压按钮来达到这一目的不是一个明智的选择，因为机舱里的每个人都会因此而缺氧。布尔曼说，最好的方法是，在可控的情况下，尽快把飞机降到2 500米左右。在那个高度，可以安全地打开通气阀，消除机舱内外的气压差。因为这一高度的空气已经不太稀薄，人们不会呼吸困难（科罗拉多阿斯潘镇的海拔高度就有2 500米）。上述操作能安全消除因舱门栓销脱开而发生外爆脱落的风险。

解决问题的毕竟是人，而不是清单

上面提到的这些步骤全都列在了前货舱门检查清单上。布尔曼强调说，这些清单是为机组人员在紧急情况下使用而精心设计的。波音公司每年要发布100多份飞行清单，这些清单要么是全新的，要么是经过修改的新版本，它们都是由编制人员一丝不苟制作出来的。布尔曼的运营小组简直就是一个清单制作工厂，多年来，他们积累了不少如何让清单变得实用有效的经验。

布尔曼向我解释说，清单有好坏之分。糟糕的清单中的检查项目模糊不清、不精确，而且表述冗长、不便使用。总之，就是非常不实用。

一些编制人员不熟悉这些清单的使用环境，他们只是坐在办公桌前闭门造车，并且把清单的使用者当成智力障碍者，把所有操作步骤都列在了上面。这样的清单只会让人们的思维变得迟钝，而不是变得清晰。

清单宣言 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

清单有好坏之分。糟糕的清单中的检查项目模糊不清、不精确，而且表述冗长、不便使用。优秀的清单中的检查项目往往非常精确、高效、切中要害，即便在最危急的情况下也便于使用。

优秀的清单中的检查项目往往非常精确、高效、切中要害，即使是最危急的情况下也便于使用。它们不会把所有操作步骤都列出来，清单不会开飞机，它们应该只提醒人们那些最关键的步骤，那些即使是资深的专业人士也可能会遗忘的步骤。总之，优秀的清单非常实用。

布尔曼强调说，清单的力量是有限的。它们能够帮助专家记忆如何操作复杂的程序和设备，它们能够帮助人们搞清楚哪些事情是最重要的，并且促使人们进行团队合作，但解决问题的毕竟是人，而不是清单。

我可以想象，当前货舱门警示灯亮起的时候，飞行员的第一本能反应可能并不是查阅清单手册。毕竟在许多时候，闪烁的警示灯最终被证明只是虚惊一场。飞行可能没有出任何问题，没有噪声，没有爆炸，没有奇怪的“砰砰”声，只有警示灯在那里一闪一闪。地勤人员已经在起飞前检查过舱门，而且没有发现问题。我们都知道，飞行事故的发生概率非常低，大约每50万次飞行才会出现一次状况。所以，飞行员的本能反应很可能是进行仪表故障检查，看看线路是不是出了问题，然后才会想到前货舱门可能真的没有关闭或关紧。

清单的力量是有限的。它们能够帮助专家记忆如何操作复杂的程序和设备，它们能够帮助人们搞清楚哪些事情是最重要的，并且促使人们进行团队合作，但解决问题的毕竟是人，而不是清单。

但飞行员之所以会向清单寻求帮助，主要是出于两大原因。首先，他们接受的训练要求他们这么做。他们刚进入飞行学校就会学到，人类的记忆和判断是有缺陷的。为了保障乘客的生命安全，他们必须认清真实情况。其次，清单被证明是有效的。但无论飞行员受过的教育如何反复强调要相信程序，而不是轻易相信直觉，也并不意味着他们会盲从。飞行清单也不可能是完美的，其中有一些检查项目让人感到困惑、模棱两可或存在缺陷，但它们还是赢得了飞行员的信任。在命悬一线的时候，飞行员愿意依靠清单渡过难关，这的确让我感到非常惊讶。

美联航的这起事故让我们看到，飞行员对清单是多么的信任。



当时飞行员的处境非常危险，机身残片到处乱飞，机舱的噪声震耳欲聋，他们当时一定会心跳加速，却要冷静地处理各种状况。除了供氧系统被破坏以外，飞机还出现了很多问题：飞机残片飞进了右侧机翼的3号发动机里，使其停止运行；残片还击中了4号发动机并使其起火；机翼外侧的襟翼已经损坏，无法使用；驾驶舱位于机身的前部，虽然机组成员想掌控局面，但他们依然搞不清发生了什么事情。

他们猜测有人在机舱内引爆了炸弹，但并不了解机身的损坏情况到底有多严重，而且也不知道是否还会发生爆炸。即使搞不清楚状况，他们还是需要关闭损坏的发动机，将紧急情况通知空管中心，把飞机降到安全高度，测试飞机的可控性，并判断仪表板上的哪些警报是可以忽略的，哪些是需要立即处理的，最后还要决定是在海面迫降，还是飞回火奴鲁鲁。如果你想测试机组成员究竟是相信他们的直觉，还是相信程序，那么观察他们在危机中的表现是最好的办法。

他们会做何选择呢？答案是他们抓起了清单手册。

机长：你是不是想让我看一下清单？

随机工程师：是的，我已经把它拿出来了。你准备好了告诉我一声。

机长：我准备好了。

要进行的操作有很多，他们必须明智地选择先做哪些，后做哪些。机组按照清单上列出的程序降低了飞行高度，安全地关闭了两台受损发动机，对飞机的降落能力进行了测试，放掉了多余的油料以减轻配重，并成功地把飞机降落在火奴鲁鲁机场的跑道上。

对于飞行员来说，清单被证明是值得信赖的。这多亏了布尔曼这样的专家，是他们编制出优秀而实用的清单。显然，编制我们的手术安全清单还有很长的路要走。

编制清单必须注意的6大要点

布尔曼告诉我说，在制作清单的时候，有几点需要考虑。

第一，清单应含有清晰的检查点，令使用者在这些节点根据清单列出的项目执行检查程序。当然，在某些情况发生的时候也需要进行检查，如警示灯亮起或发动机停车。

第二，编制人员需要在“操作-确认”和“边读边做”这两种清单类型中做出选择。在使用前一种清单的时候，团队成员先根据记忆和经验完成各自的操作，然后再一同确认是否都做好了。而在使用“边读边做”清单的时候，使用者一边念出检查项，一边进行检查，这种清单更像菜谱。所以，在编制新的清单时，必须根据具体情况选择合适的类型。

第三，清单千万不能太长。有一种说法认为，检查项目的数量应该在5~9项之间，因为人类工作记忆的容量也就这么大，但布尔曼认为不必恪守这一法则。他说：“我们要具体情况具体分析。在某些情况

下，你只有20秒的判断时间，但在其他情况下，你可能有几分钟。”

但是，如果某个检查点的停留时间超过了60~90秒，使用者就会觉得不耐烦，他们会偷工减料，跳过一些步骤。所以要尽量让清单做到简明扼要。因此，我们应该把注意力放在那些一旦跳过就可能会造成严重威胁、但又常常被人们忽视的步骤上。它们被称为“杀手项目”。航空专家会对各个步骤的重要性进行评估，并统计每个步骤被忽略的频率。我对此十分羡慕。

第四，清单的用语要做到精练、准确，语言为使用者所熟悉的专业用语。

第五，清单的版式也很重要。描述检查项目的文字长度最好不要超过一页，不要排列得杂乱无章，也不要随便使用各种颜色，大小写字母要结合起来使用以便阅读。布尔曼的介绍甚至详细到推荐我们使用像黑体这样的无衬线字体。

在某种程度上，我们在编制清单的时候，多少已经考虑到了布尔曼提到的这些要点。当然，我们还需要对清单进行一些改进，让上面的检查项目变得更精练、更准确一些。我觉得我们完全有能力轻松搞定它。但布尔曼向我指出，还有一点是非常重要的。

清单宣言 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

无论编制人员在编制清单的过程中多么用心、多么仔细，清单必须在现实中接受检验，因为现实往往比想象更为复杂。

第六，无论编制人员在编制清单的过程中多么用心、多么仔细，清单必须在现实中接受检验，因为现实往往比想象更为复杂。布尔曼说，第一稿往往很难通过。编制人员需要对失败原因进行仔细研究，对清单进行改进，并不断测试，直到使清单适用于各种现实环境中。

我对布尔曼说，在手术中进行试验可不容易。但他对此进行了反驳，因为航空业也容不得一点过错。他们不能在飞机飞行途中有意让舱门的栓销脱钩，然后观察机组人员如何处理这一紧急情况，这就是他们

为何要使用飞行模拟器的原因。布尔曼说可以带我去参观一下。

我努力掩饰住自己的情绪，不让布尔曼看出我像个孩子一样因为有机会到驾驶舱看一看而异常兴奋。我让语调尽量显得非常平静，说道：“好吧。”

清单宣言 | THE CHECKLIST MANIFESTO HOW TO GET THINGS RIGHT

清单编制6大要点：1.清单应含有清晰的检查点；2.清单类型应合理；3.清单内容简明扼要，不宜太长；4.清单用语精练、准确；5.清单版式整洁，切忌杂乱无章；6.清单必须在现实中接受检验。

是人决定了关键时刻该做什么

我们走了一小段路就来到了毗连的一幢建筑内。穿过一扇普通的铁门，我看见了一个外形怪异、像盒子一样的太空舱。这个太空舱由三根巨型液压支柱支撑。我们来到了一个像平台一样的东西上面。那个太空舱慢慢降到和平台持平的高度，3根巨型液压支柱消失在地板之下。布尔曼带我走了进去，这时我才知道，里面原来是波音777-200ER型客机的模拟驾驶舱。他让我在左侧的机长座位上坐好，而他自己坐在右侧的副机长座位上。布尔曼教我如何系好飞行员的安全带。模拟驾驶舱的风挡玻璃实际上是由三块等离子屏幕构成的。按下开关后，屏幕随即亮了起来。

“你想从哪个机场起飞？”布尔曼问道，“全世界的机场几乎都可以在我们的数据库里找到。”

我选择了西雅图的塔科马国际机场。前一天，我乘坐的航班就是在那里降落的。突然之间，机场停机坪的画面就出现在了屏幕上，这让我惊喜不已。我们的飞机停在登机廊桥旁，运送乘客行李的小拖车在我们面前开来开去，我还看见远处有飞机在滑行道上滑行。

布尔曼开始带领我执行检查程序。在我左手边的舱壁上有一个凹

槽，里面放着清单手册以便随时取用。但这本清单只是备份，飞行员一般会使用多功能显示屏上显示的电子清单。布尔曼调出了一张清单，为我演示如何操作。

“氧气。”他边说边用手指了指氧压表。

“测试完毕，100%。”我应该这样回答。

“飞行仪表。”布尔曼继续说道，随后指了指仪表板上航向和飞行高度的读数。

在起飞前的座舱准备阶段，我们只需完成4个检查项目。在启动发动机之前，我们需要完成6个检查项目，而且要完成滑行和起飞简报，机长和副机长会就滑行、起飞计划和可能出现的问题进行简短的讨论。布尔曼将详细的起飞计划告诉了我。

我尽力理解布尔曼说了些什么。他说，我们会从16左跑道正常起飞，起飞的速度大约是100多节每小时。起飞后，我们会根据标准离场程序向东南方向飞行，然后爬升到6 000米高度。我能搞懂的只有这些了。他好像还说了一些有关无线电设置的重要事项，然后开始介绍各种紧急情况的处置预案。如果飞机达到起飞决断速度之前两台发动机都停止运行了，那么我们将采取紧急制动。如果还有一台发动机能够正常工作，我们将继续爬升。不然，我们将要在附近选择合适的着陆地点。我装作很有经验的样子点了点头。

“你还有什么注意事项需要提出来吗？”他问道。

“没有了。”我说。

布尔曼启动了引擎。虽然模拟舱里没有真正的引擎，但我还是能够听到发动机加快转速的轰鸣声，我们必须提高嗓音才能让对方听清。

在飞机滑行之前，我们又停下来进行了5项检查：确认防结冰装置已经开启；确认自动刹车设置完毕；确认飞行控制面检查完毕；确认地面设备已经清除；确认没有警示灯亮起。

上面这些清单执行起来很迅速，每张大约只需要30秒就能完成。进行滑行和起飞简报的时间大约是1分钟。布尔曼告诉我，为了让清单变得尽量简洁，编制人员会在模拟舱中花大量时间对飞行员的使用情况进行观察。他们对每张清单的执行过程进行计时，然后不断加以改进和删减，直到清单上只留下最重要的步骤。

当一切准备就绪后，布尔曼让模拟的地勤人员将飞机推出。令人不敢相信的是，这次由我驾机升空。布尔曼让我用力踩脚舵松开刹车，当飞机开始滑动的时候我感到座舱抖了抖。我用左侧的一个小舵柄来控制前轮的方向，如果将其向上推，飞机就会向右转，反之，飞机就会向左转。我通过调整中央操作台上的油门杆来控制飞机在地面滑行的速度。一开始，飞机就像醉鬼一样在滑行道上扭来扭去。但当我们就要进入跑道的时候，我已经可以轻松操控飞机的行驶方向了。我将油门杆拉到最小推力的位置，双脚用力蹬脚舵踩刹车，等待塔台发出起飞指令。布尔曼随即调出了起飞前检查清单。

“襟翼。”他说。

“设置完毕。”我说。

我越来越喜欢驾驶飞机的感觉了。这时，塔台通知我们准许起飞。我松开了刹车，布尔曼告诉我应该把油门杆推到什么位置。飞机开始在跑道上加速，一开始速度还比较慢，但不一会儿，飞机就像火箭一样疾速向前冲去。我不断蹬脚舵修正航向，尽力让机头对准跑道中心线。当飞机达到起飞速度的时候，布尔曼给了我一个手势，我随即向后拉操纵杆（我曾经把这个东西当成了方向盘）。飞机腾空而起的感觉非常真实，我觉得自己真的飞到了空中，但是我并不知道模拟舱到底是怎么办到的。

只见朵朵白云飘在我们四周，脚下的城市变得越来越小。我们慢慢爬升到6 000米的高度。突然，前货舱门警示灯亮了起来。我把这次演练的根本目的忘得一干二净。多功能显示屏上立即出现了清单的前几项内容，但我还是将放在左侧舱壁凹槽里的手册取了出来，因为我想看看整张清单是什么样子。

我注意到，这是一张只有7行内容的“边读边做”清单，上面的说明告诉我们，前货舱门没有关闭或没有锁紧，我们的首要目标是减小舱门脱落的风险。虽然这只是一次模拟，但我还是感到心跳加速。

清单指导我们如何降低舱压，上面是这么写的：“拉出着陆高度选择旋钮，并调整到2 400米。”布尔曼随即向我指了指顶部仪表板上一个用来控制座舱气压的旋钮，我按照说明进行了操作。

随后，清单指示我们将飞机降低到安全飞行高度或2 400米，两者哪个更高就选择哪个。我推了推操纵杆，让机头下沉。布尔曼把高度仪指给我看。几分钟后，我们到达了2 400米高度，并改成平飞状态。然后，我按照说明，手动打开泄气阀门，并按住

按钮30秒，将剩余的机舱内外气压差全部消除。最终，飞机没有爆炸，我们安全了。
我很想和布尔曼击掌庆祝一下，我想对他说，有了清单，飞行也没有那么困难啊！

不过，有很多步骤清单并没有详细说明。比如通过无线电通知塔台发生了紧急情况，通知空乘人员做好相关准备，寻找附近最安全的降落机场，还有让人检查货舱门等。这些步骤我没有做，它们都由布尔曼代劳了。他告诉我，在编制清单的时候，他们有意省略了这些步骤。虽然这些步骤也很重要，但经验告诉他们，专业飞行员从来没有忽略过这些必要的步骤。所以，他们不需要把这些步骤放进清单。实际上，布尔曼认为，它们也不应该被放进去。

清单宣言 | THE CHECKLIST MANIFESTO HOW TO GET THINGS RIGHT

清单不是无所不包的操作手册，而是帮助使用者发挥专业技能的有效工具，要做到简洁、精准、实用。

对于清单应该在复杂的工作中扮演怎样的角色这一问题，人们可能会产生误解。无论是对建造摩天大楼的工程师来说，还是对应对险情的飞行员来说，清单都不是无所不包的操作手册，而应该成为帮助使用者发挥专业技能的有效工具，要做到简洁、精准、实用。也正是因为达到了这些要求，清单才挽救了成千上万人的生命。

让事故的教训转变为实用的清单

让我再为大家讲述一个有关飞行清单的故事，这是在2008年发生的一起飞行事故。

2008年1月17日，英国航空公司38号航班迫近伦敦，准备降落。这架飞机是从北京起飞的，已经飞了将近11个小时，机上载有152人。飞机正在进行最后阶段的下降，马上就要降落在伦敦希思罗机场。那时刚过正午，空中的云体很薄，云量也不多，可视距离超过9 000米，风也不大。虽然当时是冬天，但气温有10摄氏度，不算太冷。机上的人们怎么也不会想到，已经飞过了千山万水、就在马上要到达目的地的时候会大祸临头。

飞机下降到220米高度，飞过一片居民区，距离机场还有3 000米远。此时，飞行员需要稍稍增加推力以减小飞机下降的速度。但突然之间，引擎推力瞬时减小，先是右侧引擎，然后是左侧引擎。此时，副机长正在操控飞机，但无论他怎样用力推油门杆，发动机就是没有任何响应。隆隆的引擎轰鸣声消失了，留下的只有死一样的寂静。

副机长放下襟翼以增加升力，希望飞机能够滑翔到跑道上。与此同时，他不断蹬脚舵，保持飞机的航向，但飞机的飞行速度下降得太快了，整架飞机就像一块重达160吨的大铁砣一样重重地向地面砸去。英国航空事故调查局（Air Accidents Investigation Branch）后来对这起事故进行了调查。调查人员发现，飞机当时的下降速度是7米每秒。飞机在距离跑道400米的地方接地。据估计，当时飞机的飞行速度是200千米每小时。

这架飞机差一点就会撞进附近的居民区，驾车行驶在机场周围道路上的人们看着飞机直扑而下，他们心想这下自己必死无疑了。更巧的是，这起事故差点引发政治危机，因为当时英国首相戈登·布朗正坐车赶往希思罗机场，他将要启程对中国进行上任后的第一次正式访问。一位随行助手在接受伦敦《每日镜报》（*Daily Mirror*）访问时，这样描述当时的情形：“飞机距离我们的头顶只有几米远，擦着灯柱‘嗖’地飞了过去。”不幸中的万幸是，这起事故没有造成任何人员死亡。

飞机刚刚掠过机场周边的道路就重重地撞击在跑道前端的草地上，发出巨大的响声。飞机前轮在撞击时立刻折断，右侧的主起落架脱离了机身，有两个轮子飞了出去，并撞穿了29~30排这一段机身。左侧的主起落架则向上穿透了机翼，1 400升航空燃油随即倾泻而出。有目击者看到机身和地面擦出了火花。但幸运的是，燃油没有被点燃，飞机没有发生爆炸。虽然整架飞机被强大的撞击力彻底毁坏，但乘客们大多毫发无伤。飞机在地面滑行了300多米才停了下来。这次事故导致十几位乘客需要住院治疗，所幸他们中伤情最严重的仅是腿部骨折。

事故发生后仅仅过了1小时，调查人员就赶到现场展开调查。他们分别在1个月和4个月后发布了阶段性调查结果，但导致这起事故的原因仍是一团疑云。调查人员将发动机、燃油系统、飞行数据记录仪拆了下来，逐一进行检查，但是都没有找到任何引擎缺陷。飞行数据显示，发动机的燃油流量不知因为什么而减少了。所以，调查人员用光纤内窥镜对油路系统进行了检查，却没有发现任何缺陷或阻塞，对燃油阀门及燃油流量控制电路的测试也没有找出问题来，在油箱内也没有找到任何可能阻塞油路的残片。

因此，调查的重点开始转移到燃油本身。测试显示，油箱里装的是

一般的A-1航空燃油。由于飞机的航线经过北极圈，调查人员想搞清楚燃油会不会在飞行途中结了冰，从而引起这次飞行事故，然后又在他们能够发现任何痕迹之前解了冻。

飞行数据显示，当飞机飞临中国和蒙古边境的时候，飞机周围的环境气温是零下65摄氏度；而当飞机飞临乌拉尔山和斯堪的纳维亚半岛的时候，飞机周围的气温下降到了零下76摄氏度。对于需要穿越北极圈的航线来说，这么低的气温并不算异常。虽然A-1航空燃油的凝固点是零下47摄氏度，但航空专家已经解决了这一问题，需要穿越北极圈的飞机都经过了特殊设计以抵御极度严寒，而且飞行时飞行员会一直监测燃油温度。

跨越极地的商业航线于2001年2月开辟。至事故发生时，已有数千架飞机安全地完成了飞行任务。实际上，在英航那架飞机的飞行途中，飞行记录仪记录到的最低燃油温度是零下34摄氏度，远远高于燃油的凝固点。不仅如此，飞机是在天气晴朗的伦敦上空，而不是在天气恶劣的乌拉尔山上空失去发动机动力的。

但是调查人员还是怀疑飞行路线是事故的诱因之一。他们提出了一套精妙的理论。

航空燃油一般会含有少量水汽，每4升燃油一般含有不超过两滴的水。当飞机穿越极地的时候，燃油里的少量水分会结成小冰晶，悬浮在燃油中。人们从来没有认为这是个大问题，但他们忽视了这样一种可能：在进行长途极地飞行的时候，燃油的流速很慢，所以小冰晶有时间沉淀下来，并聚积在油箱的某个地方。当飞机即将降落的时候，燃油的流量突然加大，这可能会导致聚积起来的冰晶阻塞油路。

可是，调查人员找不到任何有力的证据来证明这一理论。这就好像有人躺在床上因窒息而死，侦探推测说，室内的氧气分子突然全部聚集到了房间的另一侧，进而导致这个人在熟睡中窒息而亡。这一解释在理论上的确说得通，但很少有人真的相信这样的事情会发生。不过调查人员并不死心，他们在低温情况下直接向油路注水，小冰晶果然形成了，而且阻塞了油路。

事故发生8个月后，调查人员只能找到这样的解释。每个人都焦虑不安，他们很希望自己能够在相似的事故再次发生之前做些什么。所以，调查人员根据上述未经证实的理论，为飞行员制订了解决这一潜在问题的具体操作步骤。

当发动机失去动力的时候，飞行员的本能反应是推油门杆以增加发动机的推力。但是，如果油箱内有冰晶聚积，那么增加燃油流量只会向油路中增加更多冰晶。所以，调查人员认为，此时飞行员应该反向操作，将油门杆拉回最小推力位置，并保持一小段时间。这样燃油流量就会减小，油管的热交换装置也就有时间融化冰晶，只要

几秒钟时间，发动机就能恢复正常工作状态。虽然调查人员提出的理论并没有得到证实，但这是他们能够想到的最佳解决方法了。

所以，在2008年9月，美国联邦航空管理局发出了具体警告，详细说明了在跨越极地飞行过程中，飞行员应该如何防止冰晶在油箱中聚积的操作规程，还说明了当发动机失去动力后，飞行员应该如何恢复动力的操作程序；同时要求世界各地的飞行员要在30天内知晓相关信息，并且熟练掌握相关操作。最值得一提的是，飞行员真的这么做了，这也正是我要讲这个故事的原因。通过它，我们能够得到很多启示，而不只是因为上述解决方案涉及清单。但在总结经验之前，还是让我们看看大多数行业在碰到重大事故时都是怎么做的吧。人们很少会对事故进行仔细调查，医生和教师不会这么做，律师和金融顾问也不会这么做，除非媒体将丑闻公之于众。某些问题可以影响成千上万人，但由于它们往往每次只涉及少数人，所以人们不太会努力寻找导致事故或错误的原因。

不过，我们有时候也会对失败原因进行调查，并想出改进的具体措施。但这些发现一般会出现在研讨会上，发表在专业期刊上，或被写入教科书。我们最常做的就是发布几厘米厚的指导意见，或制定新的标准，这并不意味着会使人们的行为立刻发生改变。人们行为的改变往往需要花费数年时间。

一项医学研究对9种重大治疗技术的发明所产生的影响进行了调查。这些技术的发明意义非凡，如肺炎球菌疫苗不仅能够保护儿童，而且能够保护成人避免患上呼吸道感染这种常见的致死疾病。但研究报告说，在美国，医生平均要花费17年时间，才会在半数治疗过程中采纳新的治疗技术。

但像布尔曼这样的专家指出，上述现象的发生并不是因为人们懒惰或情绪抵触，而是因为新知识没有被系统地转变为简单、实用的操作方法。如果航空专家也只是在有了关于航空安全的重要发现时公布厚厚的报告，就像各种医学专家每年发布70多万篇论文一样，这些信息就很难被有效利用了。

但是他们并没有止步于此。当航空事故调查员发布了厚厚的调查报告之后，布尔曼和他的团队会埋头工作，努力将这些信息转变为实用的操作步骤。他们会起草一份标准的极地飞行清单，并不断筛选和精简检查项目，不断推敲如何设置检查点，如飞行员应该怎样确认发动机失去动力是因为油箱内聚积了冰晶，而不是其他原因。最后，编制小组和飞行员一起在模拟舱中对清单进行测试，及时改正发现的问题，并再次进行测试。

布尔曼团队花了两周时间不断完善和测试清单。每一家使用波音777型飞机的航空公司随后都会收到这张清单。一些公司会直接将其投入使用，但很多公司会根据自己的具体情况对其进行进一步修改，这就像学校和医院往往有自己独特的操作规程一样，航空公司也是如此，而且波音公司和航空管理当局也鼓励他们这么做。这也是为什么当航空公司进行合并的时候，飞行员往往会为使用哪一家公司的清单而吵得不可开交。事故调查报告发布1个月之后，成熟的清单就被送到了飞行员手中，或被存进了飞机的计算机里，随后便开始正式投入使用。

我们怎么知道他们真的使用了这些更新了的清单呢？2008年11月26日，相似的事故差点再次发生。



达美航空公司一架波音777客机从上海飞往亚特兰大，机上载有247人。飞机飞临蒙大拿大瀑布的时候，飞行高度是12 000米。这时候，2号引擎突然失去动力。后来调查显示，冰晶阻塞了供油系统，英国航空公司38号航班事故调查人员先前提出的理论终于得到了证实（后来，波音公司对系统进行了改进，以防此类事件再次发生）。如果这架飞机当时在飞跃蒙大拿山的时候失去了所有动力，后果将不堪设想。

与上次不同的是，飞行员这次知道应该如何正确处置这一情况，他们根据清单上列出的步骤进行操作。由于飞行员处置得当，发动机恢复了正常工作，机上的247人得救了。整个过程进行得非常顺利，甚至乘客都没有发现飞机出现了异常情况。

我真希望这一幕有一天也能出现在外科手术台上。

第7章 持续改善 保持清单的自我进化能力

·在8家试点医院中，持续改善的清单让近4 000名病人术后严重并发症的发病率下降了36%，术后死亡率下降了47%。

·就算是关于最简单的操作步骤的清单也需要不断改进。简洁和有效永远是矛盾体，只有持续改善，才能确保清单始终安全、正确和有效。

每个清单都需要适时调整

回到波士顿以后，我让我的研究团队对尚不成熟的手术清单进行改进，让它变得更加实用。我们试图吸取航空业制订清单的经验，让它变得更加清晰、简洁。我们主要采用了“操作-确认”这种类型的清单，而没有使用“边读边做”这一类型。这样一来，手术团队就能一方面灵活地完成各种操作，另一方面能在关键节点停下来，确认是否遗漏了重要步骤。手术清单得到了进一步的完善。

最后，我们在模拟手术室里进行了测试。那里实际上是一个会议室，位于我所在的公共卫生学院的大楼里。一个助手躺在会议桌上，他就是我们的“病人”。其他人各扮演一个角色，这些角色包括外科医生、助理外科医生、手术助理护士、巡回护士，还有麻醉医生。不过，在我们就要开始进行测试的时候，问题冒了出来。

该由谁来负责让手术团队停下来执行检查程序呢？我们在编制清单的时候并没有说清楚这一点，但这不是一个小问题。引起手术室里每一个人的注意需要一定的魄力。一般来说，只有外科医生才有这样的权威和能力。我建议说，或许应该让外科医生负责这项工作。但有人对此表示反对。在航空业，负责启动检查程序的人是“不把杆的飞行员”。这么做不是没有道理的，“把杆飞行员”可能会因为忙于各种飞行操作而不便执行清单检查程序。不仅如此，分权还传递了这样一个信息，那就是飞行安全不是机长一个人的事情，每个人都要负起责任来，每个人都有对飞行操作提出质疑的权利。我的同事认为，如果我们想显著提高手术质量，那么也应该遵循这一原则：让更多的人分担责任并拥有质疑的权利。所以，我们决定让巡回护士负责启动检查程序。

护士是否需要在清单上做记号呢？没有这个必要。我们之所以执行清单，并不是想留下什么记录，而是希望手术团队能够停下来进行充分的交流，评估可能发生的问题，共同商讨各种处置措施，尽可能地提高手术安全性。

清单上列出的每一项都需要进行调整。我们仔细测量了执行每个版本的清单所需的时间，并希望每个检查点，也就是在患者接受麻醉前、皮肤被切开前和离开手术室前的检查时间都不超过60秒，但我们还没达到预设标准。如果我们希望清单能够在手术室这样的高压环境下被接受，就必须做到使清单内容简明扼要。我们必须删除一些检查项目，那

些无关病人生死的项目。

清单宣言 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

分权传递了这样一个信息，安全不是机长一个人的事情，每个人都要负起责任来，每个人都有对飞行操作提出质疑的权利。

放弃比保留更艰难

事实证明，在改进过程中，要删除一些项目的确是一件难度很大的事。简洁和有效之间存在着矛盾。如果删减的项目过多，那么关乎手术安全性的许多步骤就无法检查。但如果保留的项目太多，清单又会变得十分冗长，不便使用。不仅如此，不同专家对于哪些步骤比较重要意见不一。2007年春天，曾在日内瓦世界卫生组织总部参加会议的各方代表又聚首伦敦，共同商讨这些问题。结果，到底留下哪些检查项目成了大家分歧最大的问题。

比如，欧美多国的研究发现，时间较长的手术容易让病人形成深静脉血栓，在病人腿部形成的血凝块一旦进入肺部就会产生致命后果。如果为病人注射小剂量的血液稀释剂，如肝素，或给他们穿上压力袜，那么病人形成深静脉血栓的风险就会显著降低。但来自中国和印度的研究者却认为这么做没有必要，因为他们的病人不像西方病人那样非常容易形成血凝块，在手术中因为深静脉血栓而死亡的情况极其罕见。此外，对于中低收入国家的医院来说，血液稀释剂和压力袜价格都不便宜，而且如果医生过量注射血液稀释剂会引发危险后果，缺乏经验的医生很难把握好分寸。所以这个检查项目被删除了。

在改进过程中，要删除一些项目的确是一件难度很大的事。简洁和有效之间存在着矛盾。如果删减的项目过多，那么关乎手术安全性的许多步骤就无法检查。但如果保留的项目太多，清单又会变得十分冗长，不便使用。

我们还讨论了手术室火灾这个棘手问题。在手术中，医生往往要使用像电刀这样的高压电气设备，在使用过程中这些设备偶尔会产生电弧。而手术室里有一种可怕的助燃剂，那就是高浓度氧气。所以，世界各地的医疗机构都面对着发生手术室火灾的风险。这种火灾非常可怕，因为纯氧几乎可以把所有东西瞬间变成可燃物，如覆盖在病人身上的无菌手术单，甚至是插入他们喉咙的通气管。但手术室火灾是完全可以避免的，我们可以采取各种有效的预防措施，如事先检查氧气有没有泄漏，在不影响手术的前提下尽量降低氧气的浓度，尽量不使用含酒精的消毒剂，避免氧气聚集到手术部位。只要做到上面提到的几点，手术室火灾就不会发生。此外，如果手术团队的每一个成员在手术开始前都搞清楚氧气阀、警铃和灭火器的位置，就算是火灾真的发生了，也不会造成严重的后果。正因如此，与感染、出血和麻醉不当这些手术杀手相比，手术室火灾的发生率很低。

美国每年要进行几千万台手术，其中只有大约100台手术发生了火灾，真正危及病人生命的火灾更是罕见。但造成病人手术部位感染的手术却多达30万台，每年有8 000多人因此而丧生。与预防感染相比，我们在预防火灾方面已经做得很好了。

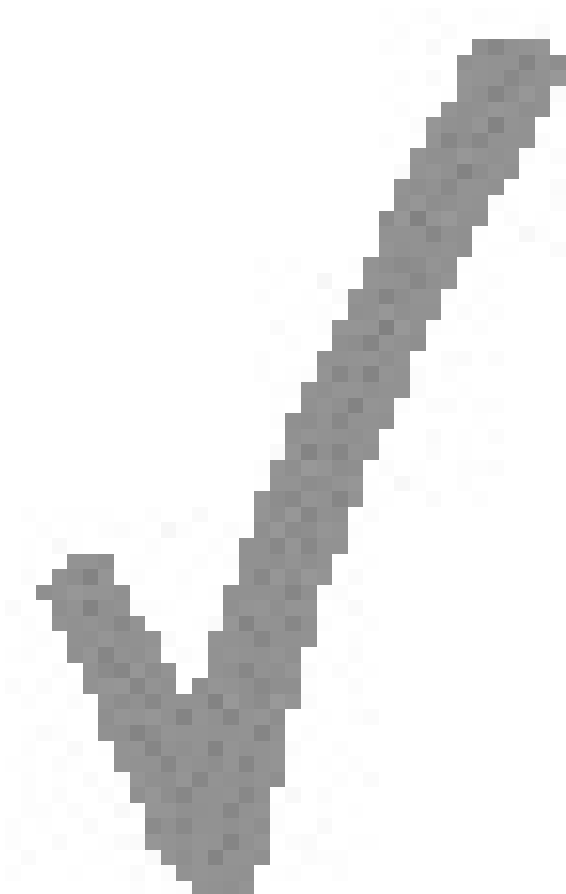
如果把消除火灾隐患的检查项目也包含在清单里，那么清单就会变得非常冗长，所以我们决定放弃它们。

我们的决策过程并不像进行科学实验那样系统。虽然搞错手术对象或手术部位的情况非常罕见，但是预防此类情况发生的措施很容易执行，而且也被一些国家接受了，其中就包括美国。此外，这类医疗事故很容易引发关注。所以，这些检查项目留在了清单上。

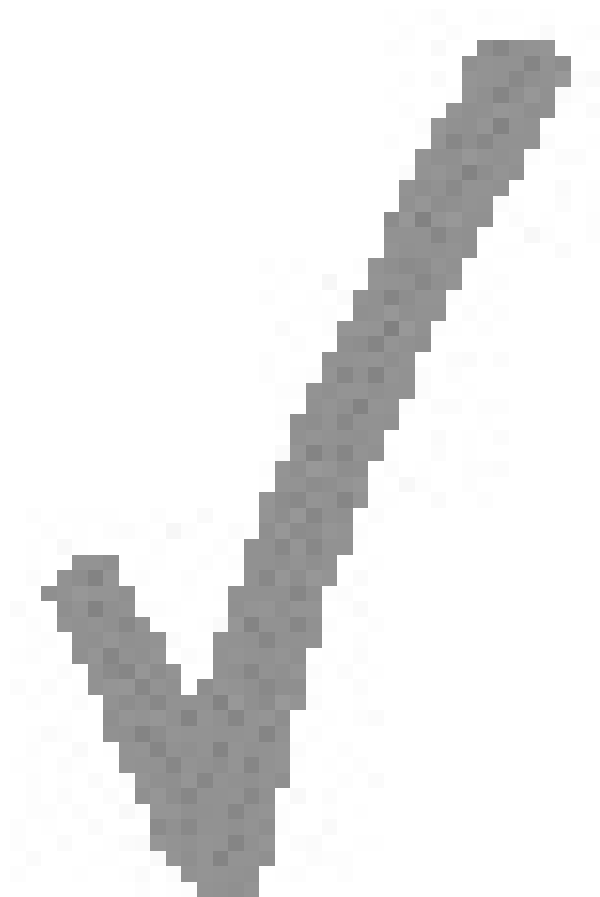
我们意识到沟通不畅是造成手术失败的一个重要原因。所以在手术开始之前，我们要求手术团队成员进行正式的自我介绍，并简短地讨论手术的关键部分和可能发生的问题。虽然这一方法的有效性尚未被证实，但由于团队合作对于手术成功极为重要，所以我们愿意把这一条留在清单里，以试一试它们的效果。

伦敦会议结束后，我们依次进行了一些小规模测试。一个伦敦的团队开始进行测试，并提出改进意见。随后，香港的团队又开始进行新一轮测试。随着测试次数的增加，清单日臻完美。皇天不负有心人，我们终于能够拿出一份可以在全球范围内推广使用的清单了。

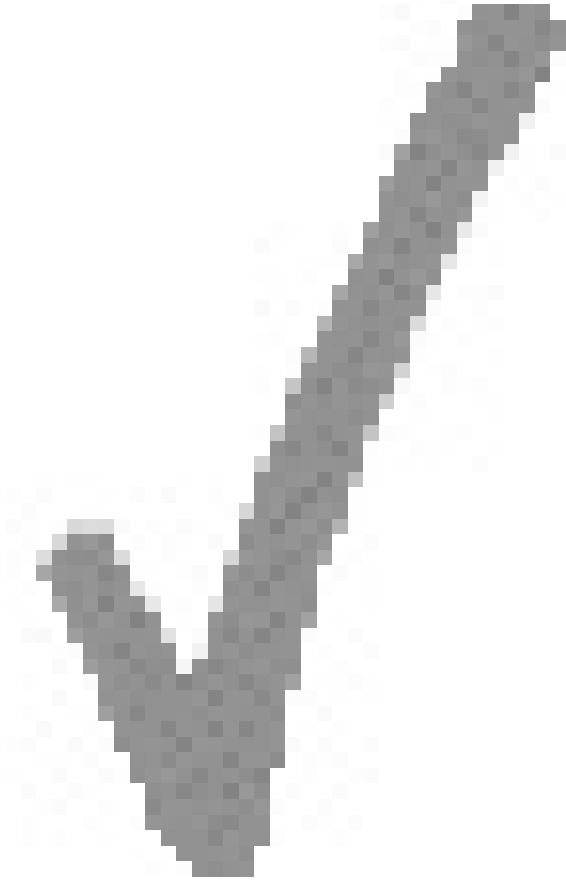
最终版本的世界卫生组织手术安全清单（Safe Surgery Checklist）由19个检查项目构成。其中，在实施麻醉前有7个检查项目：



患者本人或家属是否已经确认了患者的身份，并同意进行手术；

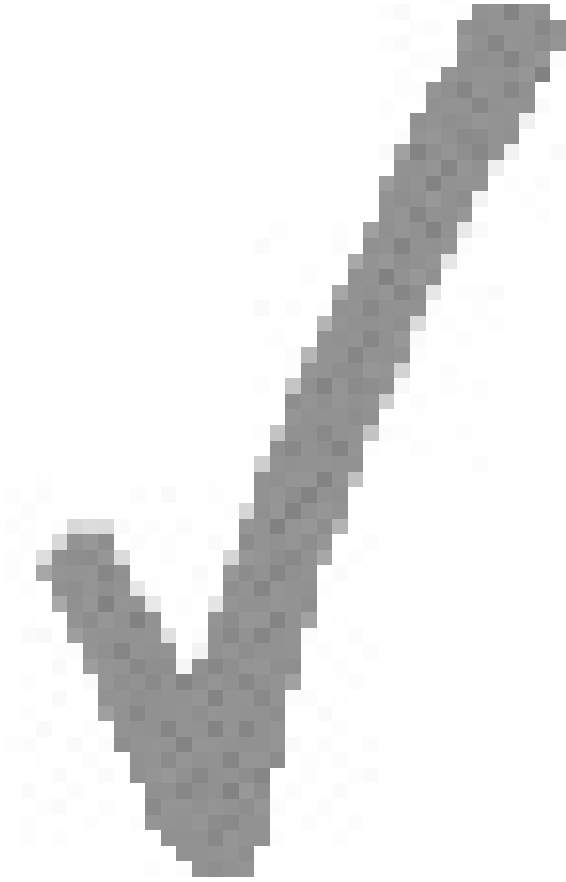


是否已经标记手术部位；

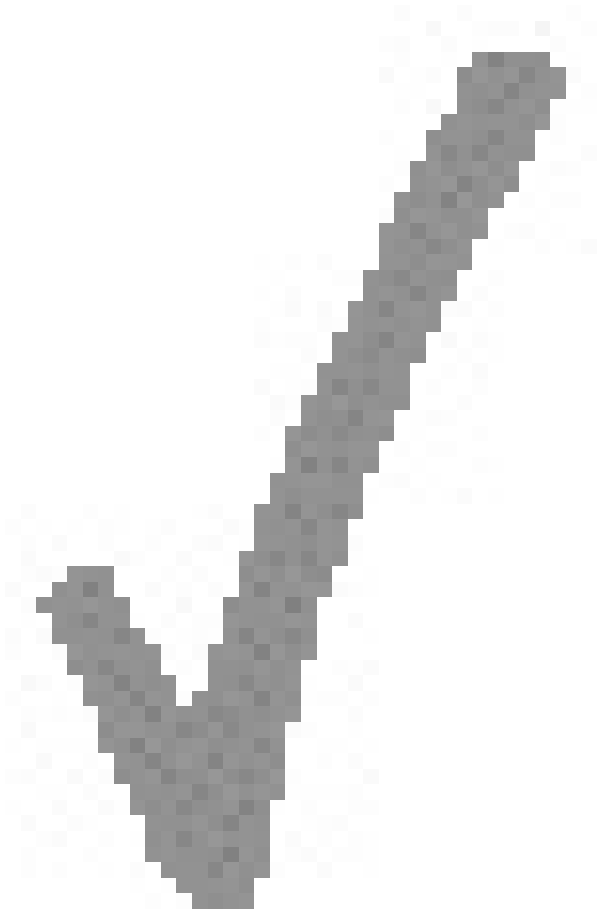


监测，相关仪器运转是否正常；

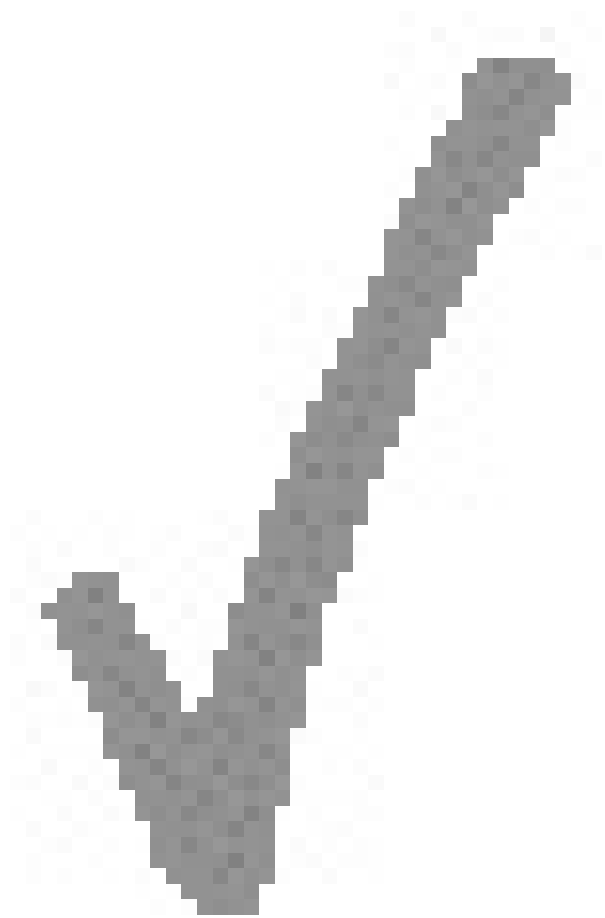
是否对患者进行了血氧饱和度



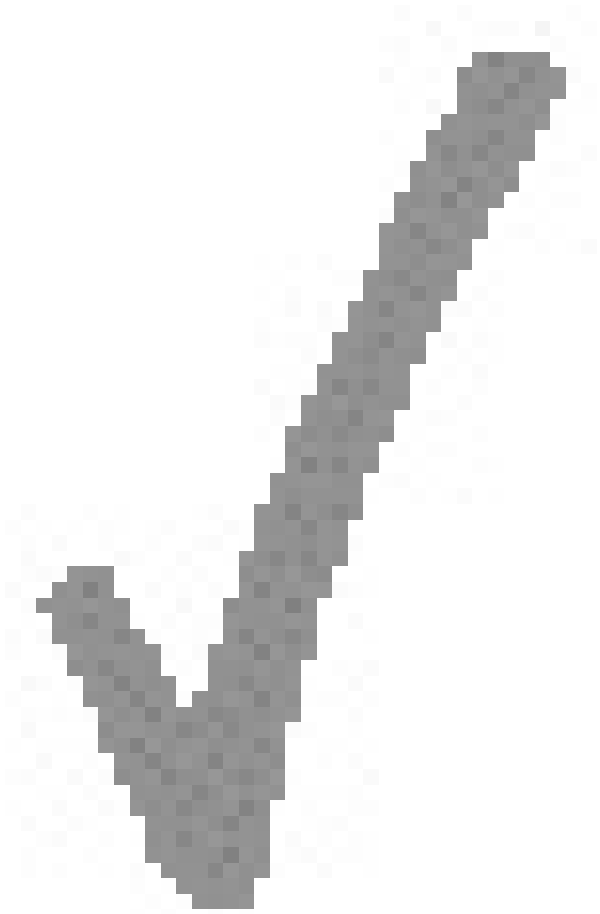
患者是否有既往过敏史；



患者是否存在气道困难和误吸
的风险（这是实施全身麻醉最危险的一个地方），所需设备和辅助
人员是否已经就位；

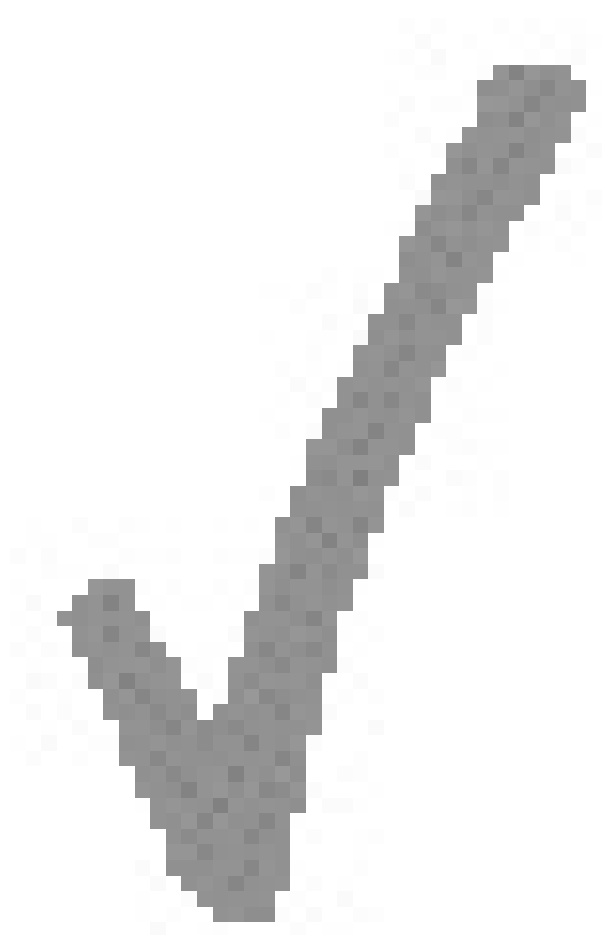


患者是否存在失血量大于500毫升的风险，患儿的判断标准为7毫升每千克；



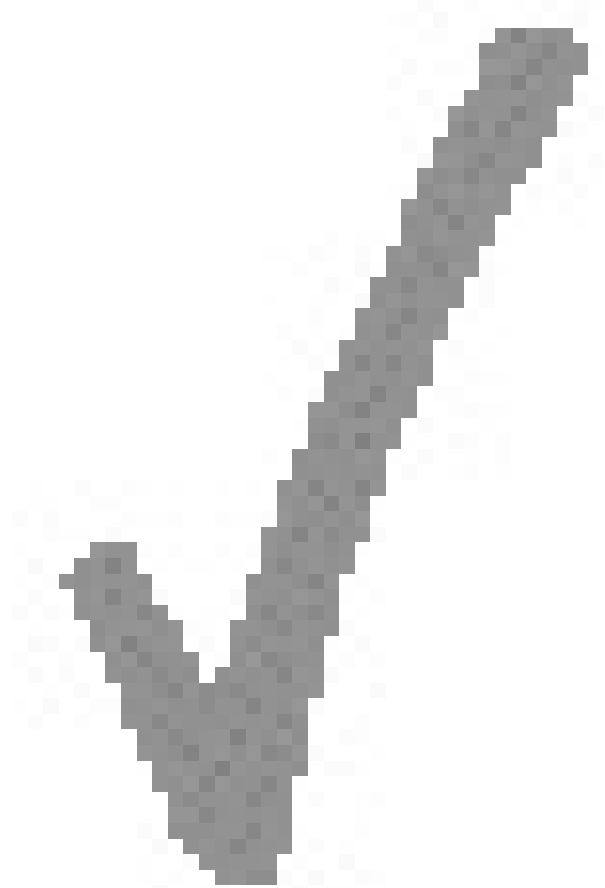
必需的 中心静脉置管、血袋和注射液是否已准备就绪。

在切开患者的皮肤前，手术团队还要进行7项检查：



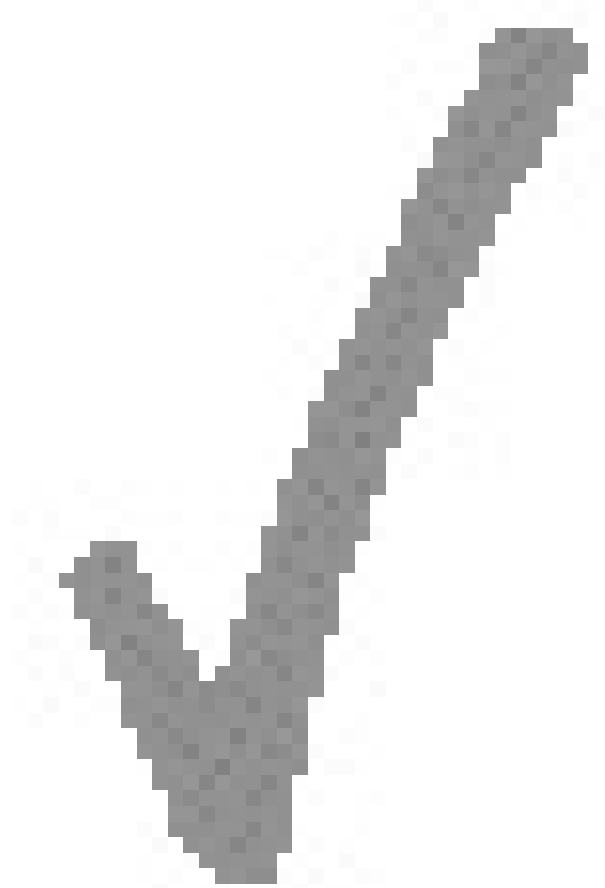
名和职责进行了自我介绍;

团队所有成员是否对各自的姓



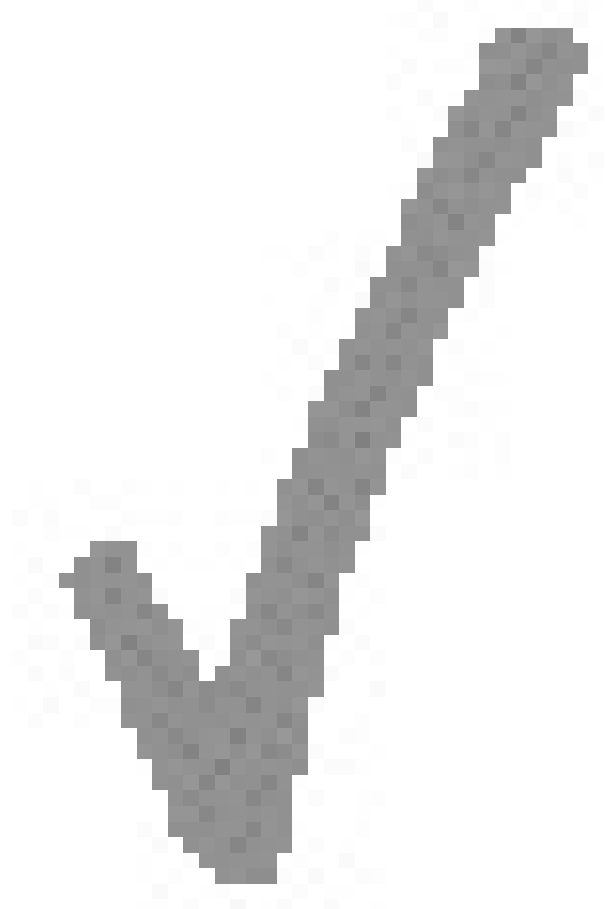
称和手术部位；

是否确认了患者姓名、手术名



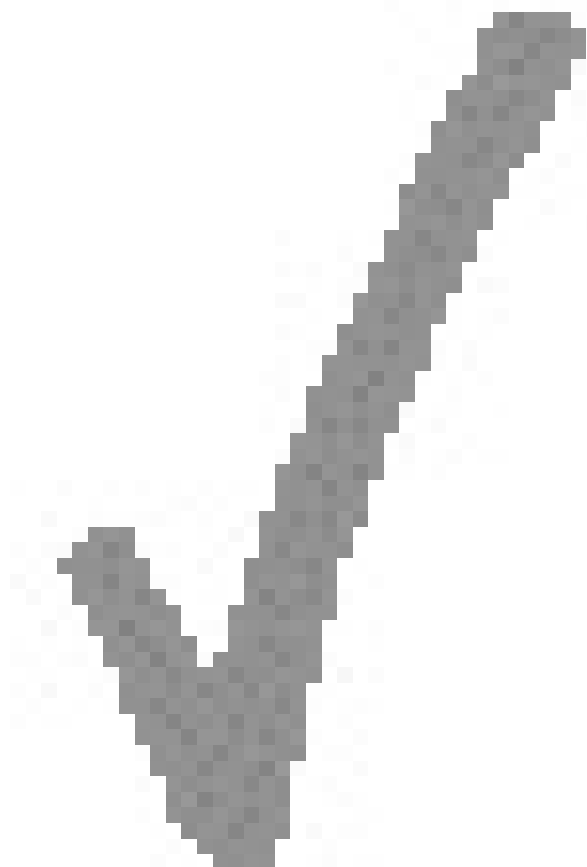
注射了预防性抗生素；

是否在手术前60分钟内给患者

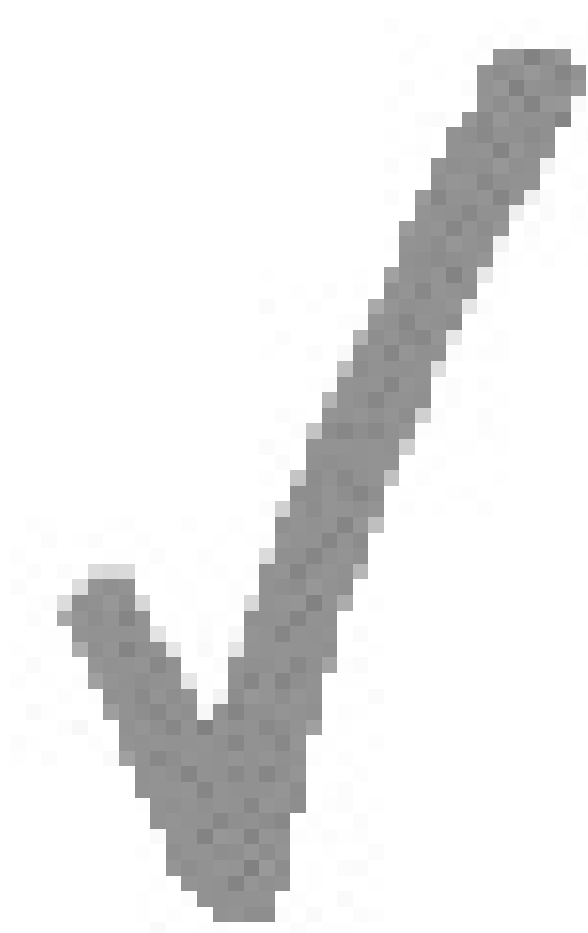


影像资料；

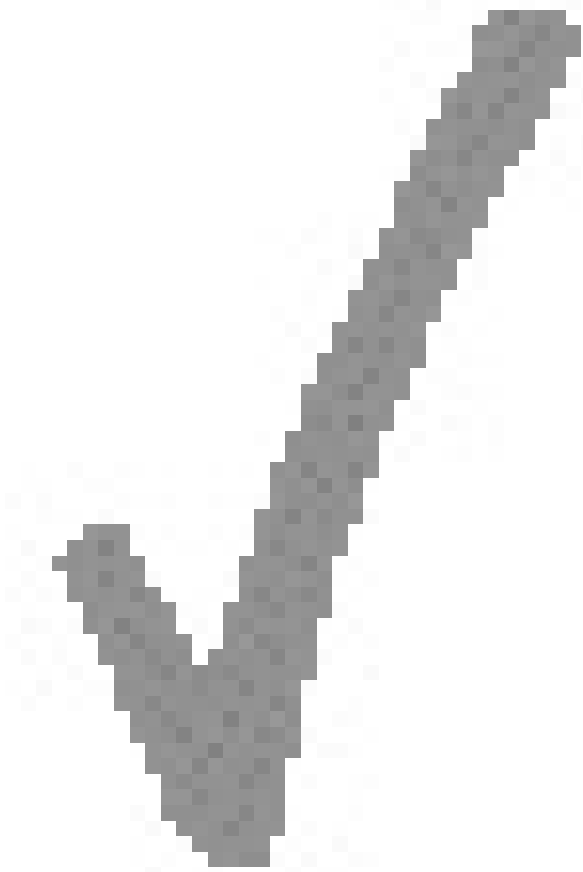
是否已展示手术所需的扫描和



是否已经讨论过有关手术的关键信息，主刀医生是否介绍了手术的关键步骤是什么、手术需要进行多长时间、预计病人的失血量是多少，相应剂量的血袋是否已准备就绪；

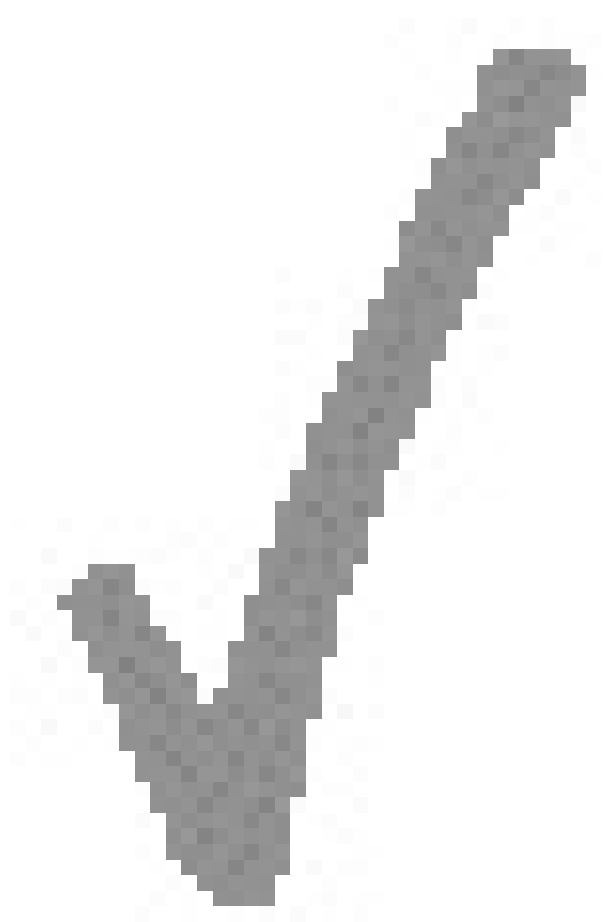


麻醉医生是否评估了麻醉计划，是否还有其他需要特别注意的事项要提出；



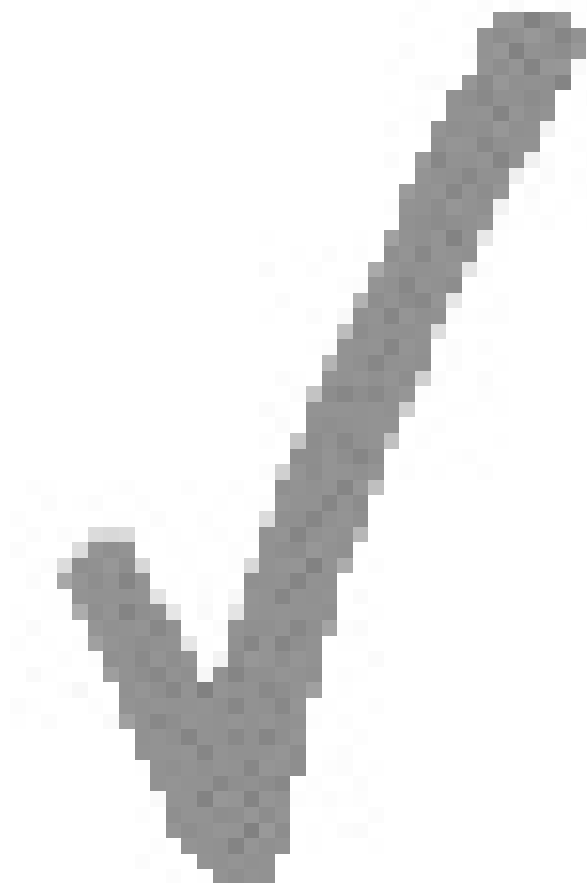
护士是否确认设备已经就位、消毒已经完成，是否还有其他注意事项需要提出。

在手术结束后患者离开手术室前，还要进行最后5项检查，巡回护士必须对下列事项进行口头确认：



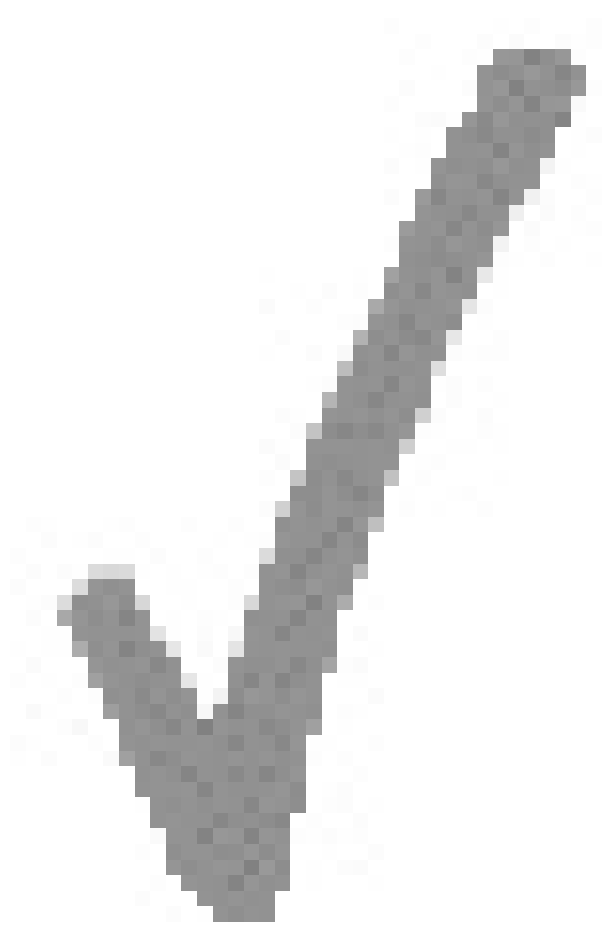
的手术是否一致；

记录里的手术名称和刚刚完成



点完毕；

手术器械、敷料和针头是否清

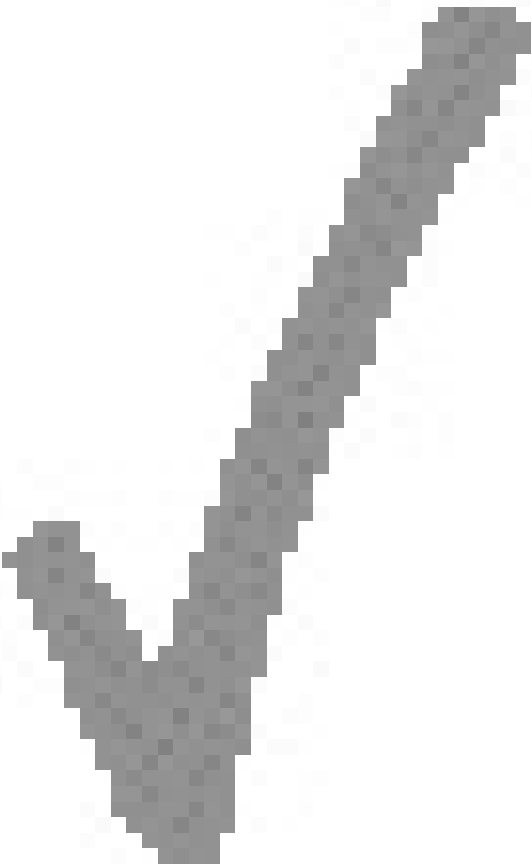


是否已经标记；

需要进行病理检验的组织标本

题;

设备是否存在需要解决的问



主刀医生、麻醉医生和护士是否说明患者术后康复治疗方面的注意事项，是否确定没有重要信息遗漏，是否已就重要问题进行了沟通。

当然，手术的整个过程远不止这19个步骤，但我们希望能够像建筑专家那样化繁为简，用为数不多的检查项目确保一些重要的步骤没有被遗漏，如确认是否已经给患者注射抗生素、确认患者有无过敏史、患者的身份是否得到确认等。我们还用一些沟通检查项目来确保手术团队成员切实进行团队合作，确保他们对手术中可能发生的问题进行了讨论，并为此做好了准备。至少我们希望这张精心设计的清单能够发挥这些作用，但我们能否如愿以偿呢？

设置清晰的测试标准

为了找到问题的答案，我们决定在全球8家医院对手术安全检查清单进行实地测试。这一大规模研究所得出的结果有充分的说服力，而且对人数不多的研究团队和数目不大的研究经费来说，这样的研究还是可以完成的。有很多医院申请参加这个项目，所以我们设置了一些选拔标

准。

首先，医院的负责人必须会说英语。虽然我们可以将英文的清单翻译成不同语言的版本，但如果8家医院的负责人说不同的语言，我们没有能力维持复杂的日常沟通。

其次，医院所在地的安全必须没有问题。我们收到了一家伊拉克医院的热情申请。这本是一件好事，但由于伊拉克战乱不断，我们考虑再三，还是没有接受这一申请。

再次，参加测试的医院必须具有广泛的多样性，既要有来自富裕国家的医院，也要有来自低收入国家或中等收入国家的医院。这一决定让世界卫生组织的官员觉得很错愕。他们解释说，世界卫生组织的首要任务是帮助中低收入国家，在高收入国家收集数据的高额成本，可以用来帮助中低收入国家做更多事情。但我去过很多医院，其中既有印度的乡村诊所，又有哈佛大学先进的医疗中心。那些地方无论是富裕还是贫穷，当地医院的手术中那些可以避免的错误都会反复发生。所以我认为清单可以在任何地方发挥作用。况且，如果试验在高收入国家取得成功，那么我们就更容易劝说中低收入国家的医院采纳这一措施。所以我们做出决定，只要高收入国家的医院愿意承担所有或大部分研究经费，它们就有资格入选。

最后，参选的医院还必须允许我们的研究人员在清单投入使用之前和之后客观地对其手术并发症发病率、术后死亡率和失误率进行调查。这对医院来说可不是一件小事，大多数医院，甚至是高收入国家的医院，对自己在这些方面的现状其实并不了解，客观的调查可能会让他们感到难堪。虽然这些条件比较苛刻，但我们还是如愿在全球挑选出了8家试点医院。

8家医院中有4家来自高收入国家，它们也是全世界顶级的医疗机构，其中包括位于西雅图的华盛顿大学医疗中心（University of Washington Medical Center）、加拿大的多伦多综合医院（Toronto General Hospital）、伦敦的圣玛丽医院（St. Mary's Hospital）和新西兰最大的奥克兰市立医院（Auckland City Hospital）。另外4家医院来自中低收入国家，但他们的接诊量也非常巨大，这4家医院分别是马尼拉的菲律宾综合医院（Philippine General Hospital，这家医院的规模是上述高收入国家医院的两倍）、约旦阿曼的哈姆扎王子医院（Prince Hamza Hospital，这家医院是约旦政府为了向数量急剧增加的难民提供医疗服务而新建的）、新德里的圣斯蒂芬医院（St. Stephen's Hospital，这是一家城镇慈善医院）和坦桑尼亚伊法卡拉的圣弗朗西斯区医院（St. Francis Designated District Hospital，这家医院要为附近的100万农村人口提供服务）。

这些医院之间的差异非常大，把它们放在一起进行研究看似有些不可靠。在高收入国家，人均医疗费用达到数千美元。但在印度、菲律宾和东非，人均医疗费用连10美元都达不到。就拿华盛顿大学医疗中心来说吧，它每年的预算超过10亿美元，是坦桑尼亚全年医疗费用总和的两倍。这8家医院实施的手术也有天壤之别。

高收入国家的那些医院往往装备精良，人才充沛，从机器人前列腺切除术到肝脏移植，什么样的高难度手术他们都能做。对那些普通的日间手术更是不在话下，如疝气修补术，乳房活组织检查，还有为慢性中耳炎患儿排除耳内脓液所做的鼓膜置管术。

对缺少资源和人手的低收入国家医院来说，他们优先考虑的是急救手术，如为难产的产妇实施剖宫产，或是挽救事故重伤患者的生命。即使是进行相同的手术，如阑尾切除术、乳房切除术、在患者断裂的股骨中打入钢钉等，这些手术在不同医院除了名称相同外，其实质内容存在非常大的差异。在资金不足的医院里，医护人员往往缺乏训练，他们的设备非常简陋，而且患者被送来的时候往往病情已经十分严重，比如阑尾可能已经破裂，乳房肿瘤可能已经长得很大，股骨不仅断裂，而且可能已经开始感染。

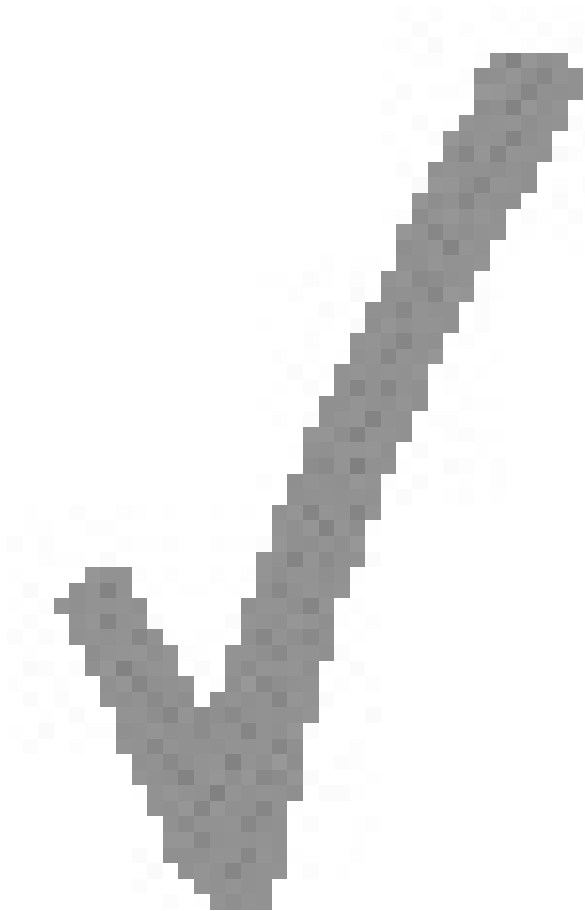
但我们还是对这8家医院进行了研究。毕竟我们的目标不是对它们进行一一比较，而是测试清单能否提高手术的安全性。我们在这些医院的所在地招募了研究人员，并就如何收集有关术后死亡率和并发症发病率的准确信息对他们进行了培训。我们制定的评价标准是比较保守的，研究人员只记录比较严重的术后并发症，如肺炎、心脏病发作、再次手术或出血量大于4个单位的大出血和严重的伤口感染等，而且只记录在医院实地发现的并发症，不能引用其他报告资料。

在清单投入使用之前，我们对每家医院的4个手术室进行了为期3个月的跟踪观察，这就像是对世界各地医院的手术质量进行活体检查一样。调查结果并不令人满意。

调查涉及的成人患者近**4 000**人，其中有**400**多人出现了术后并发症，有**56**人不幸死亡。将近半数患者并发症与感染有关，另外**1/4**患者的并发症与手术的技术问题有关。手术团队不得不把病人再次推进手术室进行止血，或修复遗留问题。在各个医院中，术后并发症的整体发病率在**6%~21%**之间。需要指出的是，在被调查手术室里进行的手术往往难度较大，较为简单的手术情况会好一些。但调查数据证实了我

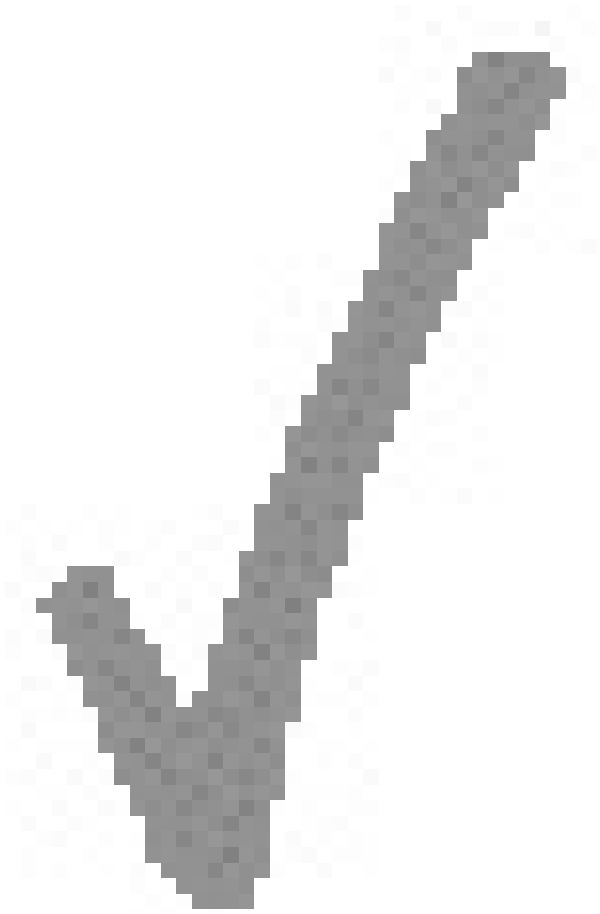
们的观点，无论医院在什么地方，手术都是充满风险和危险的。

我们还发现每家医院都有巨大的改进空间，这一点并没有出乎我们的意料。比如，没有一家医院规定手术团队必须在手术前对病人的失血量进行评估，并做好相应的准备，也没有一家医院规定医护人员在手术前必须就手术的具体情况进行简短的交流。我们对各家医院医护人员执行6个安全步骤的情况进行了跟踪调查，这些步骤分别是：

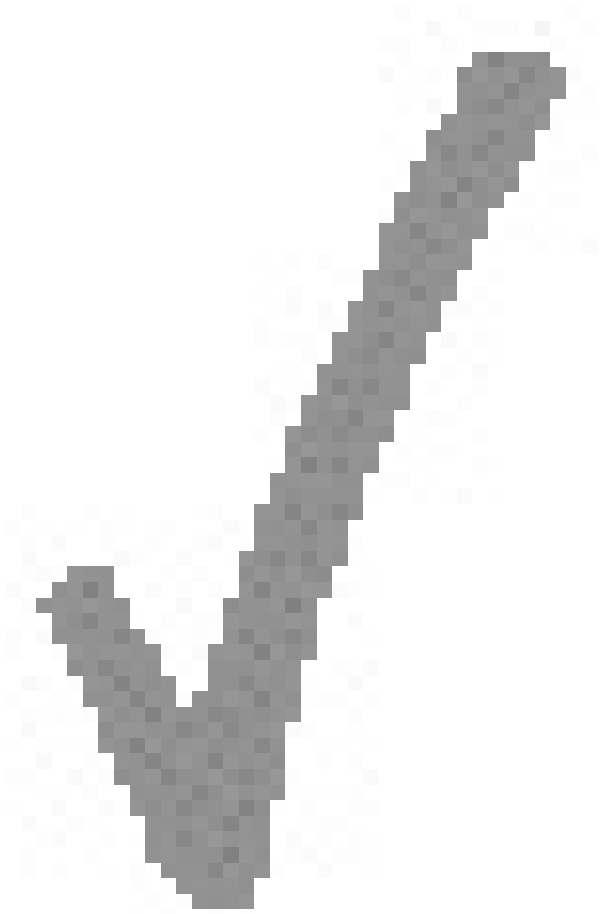


素；

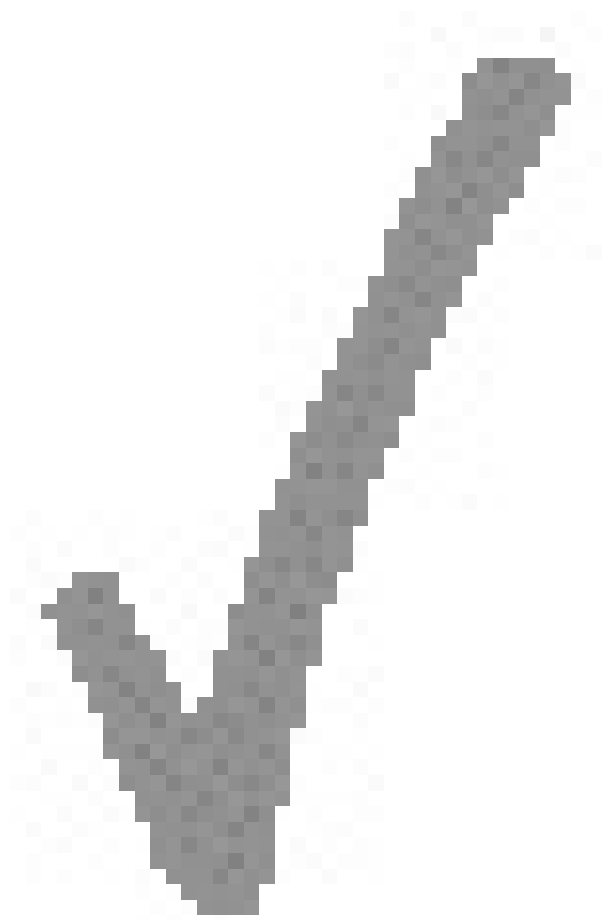
及时给患者注射预防性抗生



对患者进行血氧饱和度监测；

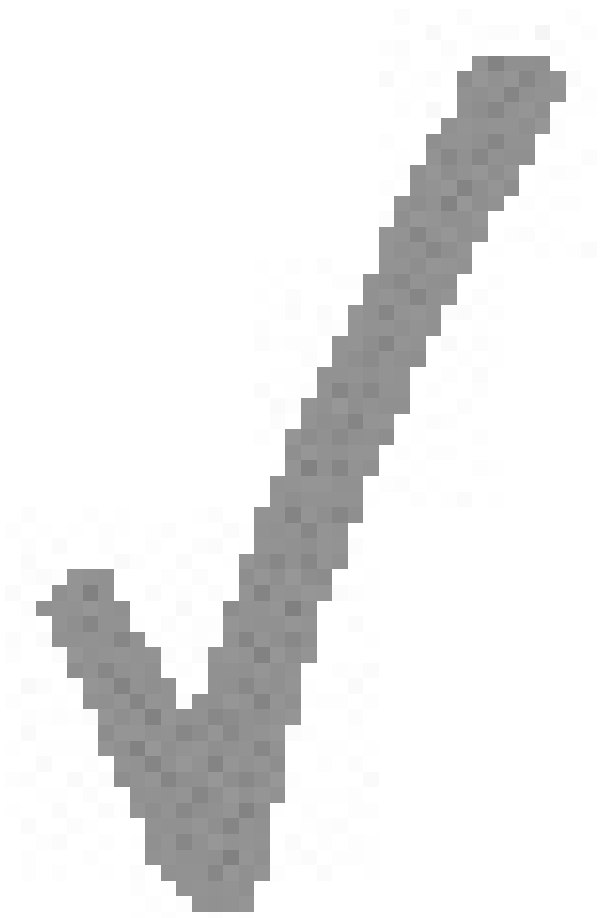


确认患者是否存在气道困难；



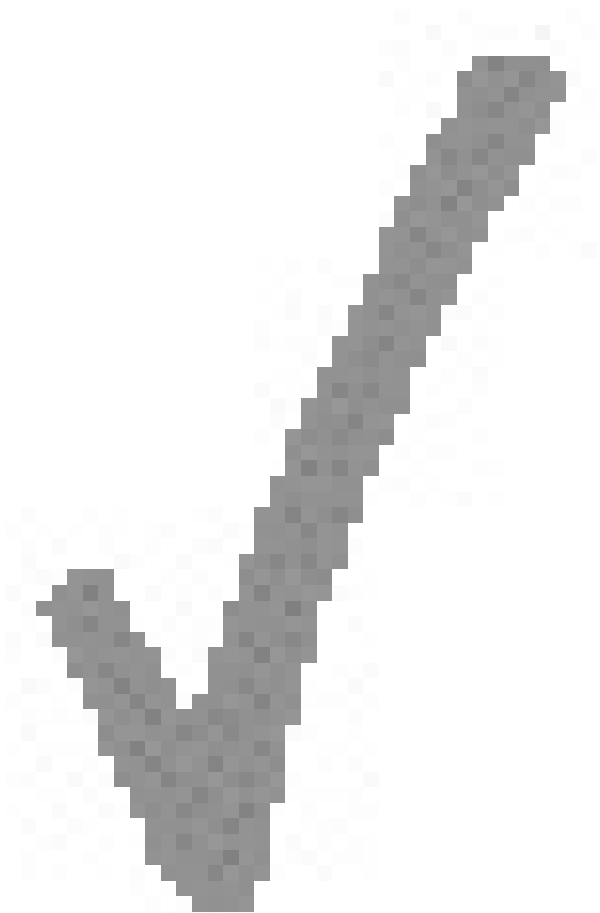
称；

口头确认患者的身份和手术名



中心静脉置管；

为失血量较大的患者安全插入



敷料和针头。

在手术结束后清点手术器械、

这些都是非常基本的步骤，就像飞行员在起飞前要对升降舵进行解锁一样。但我们却发现，各医院在此方面需要改进的地方比比皆是。即使在表现最佳的医院里，医护人员没有完成其中至少一个步骤的比例达到**6%**。也就是说，每**16**次手术就会出现一次失误。但**8**家医院的整体平均失误率更加可怕，无论当地的经济状况如何，医护人员遗漏至少一个步骤的比例占到手术总数的2/3。全球的手术安全状况是多么不容乐观啊！



从2008年春天开始，所有试点医院开始使用包含19个检查项目、执行时间长达2分钟的清单。我们当然不会天真地认为，只要把这些清单放进手术室里就能改变一切，医院的负责人需要系统地将这一做法纳入医院的日常运营之中。接受培训的不仅有外科医生，还有麻醉医生、护士和其他相关人员。我们将各家医院不足之处的详细数据交给了他们，这样医护人员就能明白需要解决哪些问题。我们为他们提供了一些PPT演示文件和YouTube视频。其中一段视频的标题是“如何正确使用手术安全清单”，另一段视频可能更加有趣些，标题是“如何错误使用手术安全清单”。这段视频想要告诫他们，在使用清单的时候，一不小心就会犯错。

我们还要求医院负责人首先选定一个手术室进行推广，手术团队最好是由外科主任医师、资深的麻醉医生和护士组成，他们一定会发现清单的意义。

每家医院要根据自己的具体情况和当地的专业用语习惯，对检查项目的顺序和措辞进行调整。一些医院使用经过翻译的清单，还有一些则表示他们要增加额外的检查项目。对某些医院来说，清单的使用将带来系统性的改变，如在手术室里存放更多的抗生素。我们希望先期进行试验的手术团队用他们的经验和耐心对清单进行必要的修改，而不是不经努力就彻底否定清单。

清单的有效使用还需要手术室里的文化发生变革，因为医护人员的责权和预期会发生改变，医院管理层必须意识到这一点。我们大胆猜测，如果在项目一开始的时候，管理层就让大家清楚地意识到他们大力支持清单的推广，那么医护人员就会更加容易接受这一新举措。

推广，就是更深入的跟踪研究

我和我的团队奔赴世界各地跟踪试验进展。我从未见过这么复杂多样的手术环境，不同医院之间反差之强烈，所碰到的问题种类之繁多，远远超出了我的想象。

坦桑尼亚的那家试点医院距离第一大城市达累斯萨拉姆有300多公里远，连接两地的只有一条单车道泥土路。每当雨季到来的时候洪水肆虐，这条生命线就会被切断，有时医院会数周无法得到药品和麻醉剂等补给。在那里等待接受手术的病人有数千人之多，但外科医生只有5个。麻醉医生的人数更少，只有4个，而且没有一个麻醉医生持有医学

学位。

手术用血大多是由病患家属提供的。当血量不够的时候，医护人员会毫不犹豫地挽起袖子。为了节省麻醉剂，他们通常为病人实施脊髓麻醉。此前我根本无法想象某些手术竟然能够在脊髓麻醉的情况下实施。他们会在消毒后多次使用手术手套，一直用到手套破了洞为止。他们还自制纱布，每天下午茶歇时，护士和麻醉医生会围坐在一张老旧的木桌旁，根据第二天手术的需要，将一卷卷白棉布剪成大小合适的纱布。

与坦桑尼亚的医院或印度的农村诊所相比，新德里那家城镇慈善医院的条件要好很多。它的供给更加充沛，医护人员的素质也更高，但它要为这座城市的1300万人口提供医疗服务，这实在是让人难以想象。医院共有7名接受过完整培训的麻醉医生，但他们每年要进行两万台手术。新西兰的试点医院虽然每年也要实施相似数量的外科手术，但那里有92名麻醉医生。印度的医护人员还要面临许多问题，如设备短缺、停电。他们每天的工作时间长达14小时，等待接受手术的病人排着长长的队伍。即使如此，我也很少听到他们怨天尤人。而在美国，你经常可以听到医护人员牢骚满腹。

这些医院并不只是贫富悬殊，它们在其他方面也非常不同。实际上，每家医院都是非常独特的。就拿伦敦的圣玛丽医院来说吧，这家医院已经有150多年历史了，占据了伦敦帕丁顿的一个街区，医院的大楼是由红砖和白色岩石建成的。弗莱明于1928年在此发现了青霉素。如今，在达西·德纳姆勋爵（Lord Darzi of Denham）的领导下，这家医院的外科团队依然是全球微创手术和手术模拟领域的先锋。

圣玛丽医院非常现代，设施先进，曾经吸引了很多社会名流，威廉王子和哈里王子就在此出生，英国前首相戴维·卡梅伦（David Cameron）严重残疾的儿子也在此接受治疗。但它并非只为权贵提供服务，而是作为英国国家卫生服务体系中的一家公立医院，一视同仁地为英国公民提供医疗服务。

我把圣玛丽医院的16个手术室逛了个遍，它们和我们医院的手术室看起来非常相似，设备齐全，技术先进，但两个医院手术的具体实施程序却存在巨大差异。在那里，病人是在麻醉之后才被推进手术室的，这意味着清单的第一部分必须修改。此外，那里的麻醉医生和巡回护士也不戴口罩。在我看来，这简直是对医疗事业的亵渎。但我不得不承认，远离病人手术部位的医护人员是否有必要戴口罩还有待证明。那里的医护人员使用的术语我几乎都不熟悉，虽然我们说的都是英语，但我常常搞不清楚他们在说些什么。

约旦试点医院的手术室环境也让我觉得既熟悉又陌生。约旦是一个

发展中国家，那里医院中的手术设备比较陈旧，而使用频率却很高，手术室里没有任何华而不实的东西。不过，这样的条件我完全可以接受，也可以适应。那里的手术质量看起来也很不错。访问期间，我碰到了一个伊拉克籍的外科医生。他在巴格达学医，毕业以后就一直在那里做医生。2003年美国发动伊拉克战争，他不得不放弃自己的财产和工作，与家人一起逃到了约旦。萨达姆·侯赛因在执政的最后几年把伊拉克的整个医疗系统都掏空了，而在此之前，巴格达是整个中东地区医疗水平最高的地方之一。现在的情况大不相同了，那位医生告诉我说，约旦似乎已经取代了当年巴格达的地位，他觉得自己能够来到这里很幸运。从对话中我了解到，每年有20多万名外国人到约旦求医，给这个国家带来了10亿美元的收入。

在约旦，男女地位非常不平等。让我感到吃惊的是，在手术室里，这种不平等多多少少被打破了。

刚到那里的时候，我曾经坐在一家餐馆外面观察从我身边经过的人。我发现，男性和女性几乎被完全隔开。大多数妇女用头巾把头发盖住或包起来。负责接待我的是一位20多岁的外科住院实习医生。有一次，我们一起出去看了一场电影。他告诉我说，自己已经和一名女研究生交往了两年时间。我问他花了多长时间才看到对方的头发。

“我从没见过。”他说。

“别糊弄我了，真的没见过？”

“真的没有。”他的女朋友偶尔会露出一小缕头发，所以他知道女友的头发是深棕色的。即便是对这些接受过高等教育、以比较西方的方式来恋爱的年轻人来说，情况也没有好到哪里去。

在手术室里，所有外科医生都是男性，大多数护士都是女性，麻醉医生的性别比例则是对半开。这个国家的社会等级如此森严，我很怀疑清单有效使用所需的团队合作是否真的能够在这里实现。手术室里的女性依然戴着头巾，她们中的大多数会回避男性的目光。但我渐渐发现，团队成员之间的关系并不像看起来的那样等级森严。只要有必要，医护人员会毫不犹豫地抛弃这些条条框框。

在我观摩的一次手术中，外科医生下意识地用手调整了一下无影灯，弄脏了自己的手套，而他自己并没有注意到这个动作，但一旁的护士却看得真真切切。

“你必须更换手套。”护士用阿拉伯语说（有人把她的话翻译给我听）。

“没什么问题。”医生说。

“不行，”护士坚持说，“别做傻事。”随后，她帮助医生更换了手套。

虽然这8家医院的差异十分巨大，但它们毕竟是医院，参与试验的

毕竟都是外科手术团队，所以很容易在彼此之间找到共通之处。你总会发现有病人躺在手术台上，他们的内心充满了希望，但也充满了恐惧。病人把自己的身体托付给你，相信你能挽救他的生命。无论在哪家医院你都会发现，手术室里的医护人员各司其职，竭尽全力，为的就是不辜负病人给予他们的信任。

当然，清单的推广并不是一帆风顺的。在应用过程中，我们会时不时地碰到一些小问题。比如在马尼拉，每4台手术才能分配到一个护士，因为训练有素的护士都被美国和加拿大的医院挖走了。前来补缺的医学院学生还非常稚嫩，哪里敢负责启动检查程序。所以我们必须劝说麻醉医生来承担这项职责。而在英国，当地的医护人员不知道应该如何根据他们独有的麻醉程序来改动清单。

学习总是有个过程的。无论清单看似多么简单，如果你以前没有养成使用它的习惯，那么现在要迅速养成这种习惯就很难。有时候，手术团队会忘记执行其中的某个部分。特别是手术结束后，病人即将被推出手术室时进行的那部分检查。另外一些时候，医护人员觉得一丝不苟地执行清单太困难了。这倒不是因为检查程序非常复杂，而是因为觉得有些话碍于情面不太好说出口。

护士在手术开始前发现病人没有注射预防性抗生素，依照程序，整个团队必须停下来，待病人注射完以后才能继续手术。但这么做可能会让爱面子的其他成员感到难堪，所以希望维护团队表面和谐的人往往难以启齿。手术室里的每一个人都有自己的行事风格，特别是外科医生，他们有的沉默不语，有的喜怒无常，有的则滔滔不绝，很少有人知道应该如何调整自己的风格以便使团队进行充分交流。

至于每个人是否应该在手术开始前对自己的姓名和职责进行自我介绍这个问题，人们的看法差异很大。从新德里到西雅图，护士们对这一环节都非常赞成，但一些外科医生却觉得这个环节非常烦人。不过，大多数医生还是接受了这一做法。

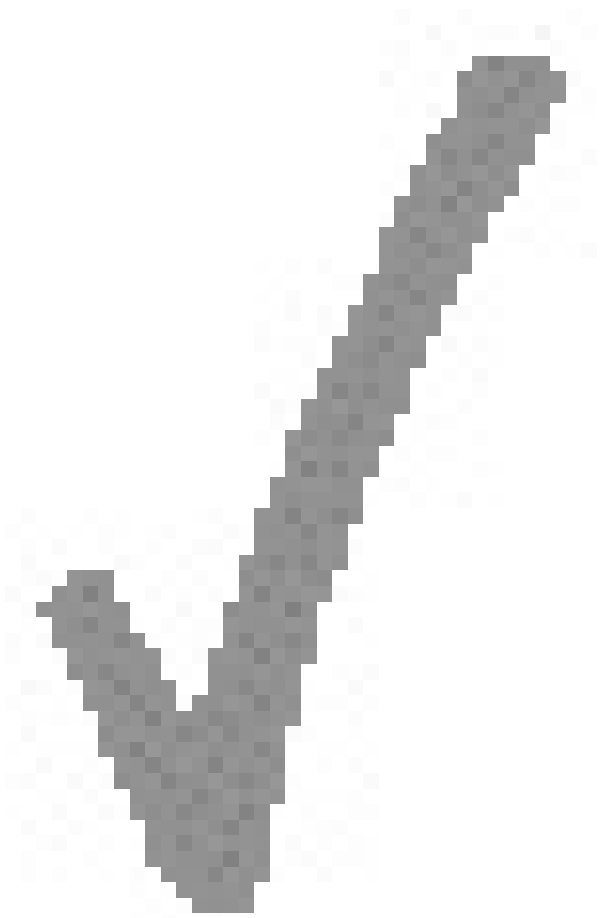
请注意，我说的是大多数，而不是所有人。在世界各地的试点医院，我们都曾被赶出手术室。有的医生曾对我们说：“执行这张清单实在是浪费时间。”在一些地方，管理层会斥责不听话的医生，并强迫他们使用清单。但我们并不鼓励这种做法，强迫只会引起更大的反弹，这种抵触情绪会传染给其他团队成员。我们希望管理层在介绍这张清单的时候，只是将其描述为一种简单的工具，一种可能会提高手术安全性的工具。毕竟，这张清单的效果还有待证实。持有反对意见的医生也可能是对的，或许这又是一次善意而徒劳的尝试。

虽然我们碰到了一些抵制，但试点工作很快就在各家医院全面展开了。在我们进行跟踪研究的每个手术室里，每台手术的进行都少不了这张清单，我们还对病人术后的恢复情况进行了调查。当一切步入正轨后，我回到波士顿等待试验结果。

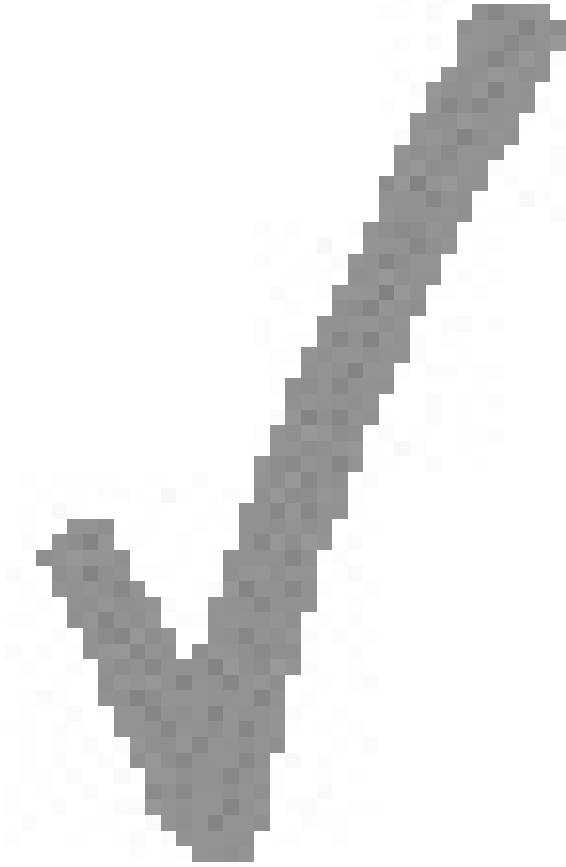
试验结果的有效性

对于试验究竟能取得怎样的结果，我心里实在没底。在清单投入使用之后，我们只进行了为期3个月的跟踪调查。如果在这么短的时间里发生了改变，应该可以归因于清单的使用，而不太可能是由医疗技术改善等长期趋势变化造成的。但我担心的是，在这么短的时间里，改变是否真的会发生。手术团队还在不断磨合，也许他们需要更多时间进行学习。而且说心里话，他们得到的帮助实在是非常有限，我们既没有为他们增添设备，又没有为他们增加人手或其他医疗资源。贫穷的医院依然很穷，我很担心在切实提高他们的医疗条件之前，可能真的很难发生什么改变。毕竟，我们只给了他们一张包含了19个检查项目的清单，他们接受的培训也只是学习如何使用它。虽然竭尽全力让这张清单变得简明扼要，但我们或许做得有些过火了，去掉了太多不该去掉的部分。也许，我们根本不该听从航空专家的建议？

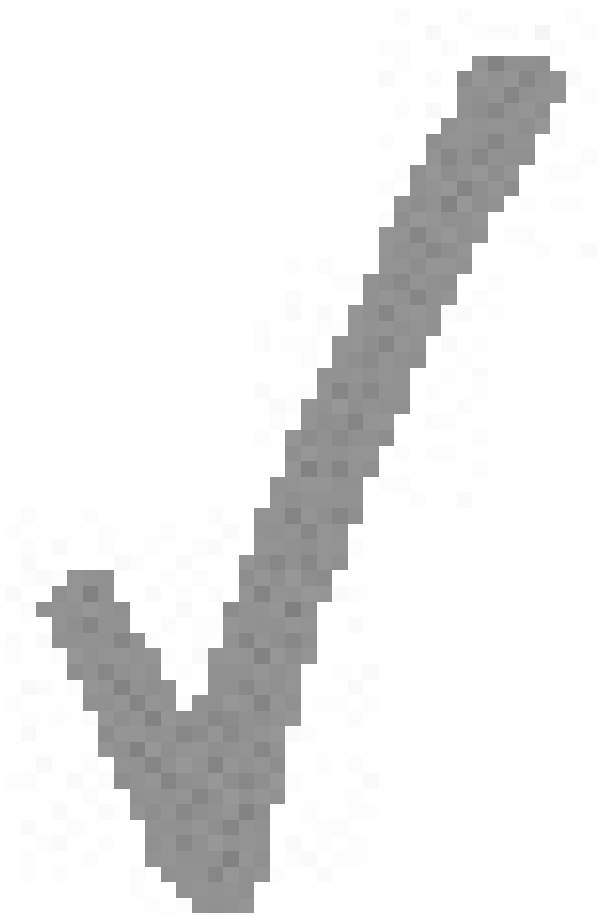
但渐渐地，一些振奋人心的消息从四面八方传来。



在伦敦圣玛丽医院进行的一台膝关节置换手术中，手术团队在执行检查程序的时候意识到，他们准备的人造膝关节尺寸大小不对，而且在医院里找不到正确尺寸的备件。幸好医生还没有下刀，否则悲剧就无法挽回了。这台手术的主刀医生曾经猛烈抨击清单。但在这件事情发生后，他立刻180度大转弯，成了清单的支持者。



在印度的试点医院，手术清单的使用暴露出了当地手术流程的固有缺陷。以前，病人是在手术室外的等候区接受抗生素注射，但手术往往会延误，医护人员在使用了手术清单后意识到，早先注射的抗生素很可能在手术开始前数小时就已经失效了。所以，医生根据清单对手术流程进行了改进，现在他们会等到病人被送进手术室后才为其注射抗生素。



在西雅图，华盛顿大学医疗中心的一个朋友告诉我，养成使用清单的习惯非常容易。但我更加关心的是，清单是否帮助他们发现了潜在问题。她告诉我，这一举措的确暴露了很多问题，如抗生素没有及时注射，设备没有就位，其他注意事项被忽略等。不仅如此，她还认为，手术清单能够帮助整个团队更好地应对手术实施过程中可能发生的各种问题，如发生病人大出血或碰到手术技术难题。她说：“我们的团队合作越来越好。”

这些故事让我的内心充满了希望。

2008年10月，试验结果出炉了。我有两个研究助理，他们都是普外科住院医师。其中一个名叫亚历克斯·海恩斯（Alex Haynes），他全脱产协助我进行调查。海恩斯花了一年多时间，在8家医院组织整项研究，并且把所有数据都整理好。另一个助理名叫汤姆·韦泽（Tom Weiser），他在这张清单的开发上已经投入了两年时间，并负责对数据进行复查。为了确保分析结果准确无误，退休外科专家威廉·贝里（William Berry）对所有试验程序和数据进行了最后一遍检查。终于，

在某一天的下午，他们来到了我的办公室。海恩斯说：“你该看看这个。”他把一沓统计表放在我面前，并向我一介绍分析结果。

清单的力量 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

在手术安全清单投入使用之后，8家试点医院的病人术后严重并发症的发病率下降了36%，术后死亡率下降了47%，而且所有结果都具有统计学意义。

病人术后感染的发生率几乎下降了一半，因大出血或手术技术问题而需要再次接受手术治疗的病人数量减少了1/4。从整体上来看，在研究涉及的近4 000名病人中，原本可能有435人发生严重并发症，但实际的发病人数只有277人。清单的使用让150多人免受伤害，更让死亡人数减少了27人。

你可能会觉得我当时一定高兴得从椅子上跳了起来，然后冲出房间边跑边叫：“我们成功啦！我们成功啦！”但我并没有这么做。实际上，我一下子变得非常紧张。我一头扎进这一堆统计数据中，寻找能够推翻这些分析结果的错误和问题。

这样的结果会不会只是一种巧合，与清单的使用无关呢？我努力猜测其他能够解释这一结果的原因，如手术团队在使用清单后实施的高风险手术和急救手术比较少，所以统计数据显示良好。于是，海恩斯又对数据进行了一遍分析。结果显示，我提出的这些假设并不成立。事实上，在清单投入使用后，急救手术的数量甚至比以前更多了一些，而且各类手术的比例，如产科、整形外科、胸部和腹部手术的比例并没有发生改变。

那么，这会不会是霍桑效应造成的呢？也就是说，因为有人在手术室里进行观察，所以手术团队变得更加谨慎，工作也变得更加卖力了，而我们却误认为是清单的功劳。毕竟，在研究涉及的手术中，研究人员对其中的20%进行了实地观察。但研究团队反驳了我的观点，因为项目一启动，研究人员就进入了手术室开始进行实地观察。但在清单使用前，各项指标并没有显著改善。而且，我们还记录了研究人员对哪些手术进行了实地观察。统计分析结果显示，无论研究人员是否在场，各

项指标都因为清单的使用而显著提高了。

好吧，我又提出了一种假设，清单会不会只在某些地方，如中低收入国家的医院里发挥作用？但调查数据显示，虽然4家高收入国家的医院的初始并发症发病率的确较低，但是清单的使用还是让这些医院的病人术后严重并发症发病率下降了1/3，效果是非常明显的。

研究团队向我逐一介绍了各家医院的试验结果。

清单的力量 | THE CHECKLIST MANIFESTO HOW TO GET THINGS RIGHT

在这次试验中，无论在哪一家医院，清单的使用都显著降低了病人术后并发症的发病率。在8家医院中，有7家医院的这一指标下降了10%以上。

试验结果是真实有效的。

沟通质量的改善是关键

2009年1月，《新英格兰医学杂志》迅速发表了我们的试验成果。实际上，在我们将分析结果告诉试点医院的时候，各相关方就已经得知消息。华盛顿州的很多医院在听到西雅图医院的试验结果之后，纷纷主动尝试使用清单。很快，他们就与保险公司、波音公司和政府结成了同盟，在全州范围内全面推广清单的使用，并对相关数据进行记录。在英国，圣玛丽医院的外科主任达西勋爵恰巧被提名为卫生部部长。当他和英国驻世界卫生组织最高代表利亚姆·唐纳森爵士（Sir Liam Donaldson，唐纳森爵士一开始就非常支持手术清单的推广）看到研究结果的时候，他们也在全英国范围内发起了推广这一举措的运动。

然而，外科医生的反应比较复杂。虽然执行检查程序并不像他们原先担心的那样占用太多时间（实际上，在一些医院里，这一举措还节省了一些时间），但一些医生仍提出异议，认为我们的研究并没有说明清单到底是如何发挥巨大作用的。

事实证明，在8家试点医院里，我们的确观察到医护人员在手术安

全方面有了长足的进步，他们更加注意为病人及时注射预防性抗生素，在手术前严格确认病人的身份和手术名称，在手术中更加注意监测病人的血氧饱和度。但是，这些改进无法解释为何一些不相关的并发症，如大出血的发生率也显著下降了。我们分析，手术团队成员之间沟通质量的改善是关键所在。在清单投入使用后，我们对刚刚走出手术室的医护人员进行了随机调查。结果发现，医护人员对团队合作质量的打分与病人术后的恢复状况存在显著的相关性。团队合作质量改善得越多，术后并发症的发病率下降得就越多。

也许，医护人员的反馈意见最能说明问题。有250多名医护人员，其中包括外科医生、麻醉医生、护士等，在清单投入使用3个月后匿名填写了调查问卷。起初，大部分人对这一举措是否能够奏效表示怀疑。但使用了清单后，80%的被访者认为清单便于使用，耗时很少，并能提高手术的安全性；有78%的被访者在自己的工作中凭借清单发现过隐患。

但还是有人表示怀疑。毕竟，还有20%的被访者认为清单不便使用，耗时过多，而且并没有提高手术的安全性。

于是，我们还让被访者回答了一个问题：“如果你自己要接受手术，那么你是否希望手术团队使用清单？”

结果，有93%的人给出了肯定的回答。

第三部分

让使用清单成为一种习惯



第8章 清单，让世界更简单

·一张清单，竟然让投资家建立的投资组合市值增长了160%；一张清单，得到全世界2 000多家医院的积极推广且使医院医疗事故发生率显著降低。

·每个人都会犯错，别让相同的错误一再发生，别让我们为那些错误付出沉痛的代价。清单不是写在纸上的，而是印在心上的。清单正在一步步改变我们的生活，改变这个复杂的世界。

超越复杂性

有一个机会就放在我们面前。这个机会不仅仅属于医疗行业，而属于每个行业。即使是最资深的专家，也能通过它寻找错误和失败的原因，并制订相应的检查项目来显著提高自己的业绩。但我们真的会这么做吗？我们是否为此做好了准备？这些问题的答案尚不清晰。

就拿手术安全清单来说吧，这份清单能够显著降低术后并发症的发病率。如果有人发现了某种效果一样好的新药，那么他们就会请名人在电视上做广告，大肆宣传；推销员则会忙着联络医生，希望他们能够给病人使用这种药物；政府则会斥资对其进行研究；丰厚的利润会吸引更多制药公司对这种药物进行改进，并生产出效果更好的新药。如果清单是一种医疗设备，那么外科医生会在展示会上争相尝试，然后缠着管理层，吵着嚷着让他们引进这种新设备。而这些管理层，也不得不为了提高医院的医疗质量而答应外科医生的请求。

手术机器人就是这么火起来的，每台价值170万美元的遥控机器人的诞生，就是为了让医生在进行腹腔镜手术的时候动作更加灵活，并尽量减少术后并发症的发生。外科医生对其垂涎三尺。其实，此类机器人的使用大幅度提高了外科手术的成本，而它们的效果并不比标准腹腔镜手术好多少。不过，世界各地的医院还是花费了数十亿美元来采购这些设备。

与之相比，清单受到的待遇又如何呢？人们并未对其置之不理。自从试验结果公布以来，有十几个国家在全国范围内对其进行了推广，其中包括澳大利亚、巴西、加拿大、哥斯达黎加、厄瓜多尔、法国、爱尔兰、约旦、新西兰、菲律宾、西班牙和英国。一些国家还对清单的使用效果进行了进一步跟踪，这对确保清单的成功推广是非常重要的。

清单的力量 | THE CHECKLIST MANIFESTO HOW TO GET THINGS RIGHT

在美国，20个州的医院协会也承诺要对清单进行推广。到2009年底，有10%的美国医院已经或准备使用清单。在世界其他地方，也有2 000多家医院已经这么做了。

这些消息的确非常鼓舞人心，但要让医生全面支持这一举措还有很长的路要走。清单的使用往往是出于外力的影响，如指手画脚的政府卫生官员，而外科医生对他们多少有点抵触情绪；再比如伸长耳朵，到处

打探消息的医院安全管理员，在外科医生的心目中，他们和小时候学校里的校规巡视员一样令人反感。有时候，负责推广的是外科主任，这意味着外科医生尽管不敢当面抵触，但他们会认为上司有越俎代庖之嫌。他们会想：这是我的病人、我的手术室，我怎么做手术是我自己的事情，我对手术全权负责，他们凭什么告诉我应该怎么做。

既然外科医生已经开始使用清单，即便他们有些不情愿，那又有什么关系呢？结果是最重要的，难道不是吗？

未必如此。我们的目标并不是让医生们在检查项目旁边乖乖打钩，而是要培养医疗团队注重合作和纪律的文化。如果我们清醒地意识到这是一个机会的话，那么这张耗时2分钟的清单仅仅是一个开始。我们手上的这张清单只能帮助外科医生们发现手术中一些共有的问题，但实际上我们能够使用清单方法做更多事情。

我们可以为髋关节置换手术、胰腺手术、主动脉瘤修复术制作专门清单；再比如，我们可以对每一种大型手术的操作过程进行研究，找出可以避免的常见问题，最后使用清单来消除这些隐患。我们甚至可以像航空业那样，针对一些特殊情况制订紧急清单。不知道大家是否还记得约翰告诉我的那个故事，医生在手术过程中可能会忘记血钾水平过高也会导致心脏停止跳动。如果有相应的紧急清单，这类问题就不会再发生。

清单宣言 | THE CHECKLIST MANIFESTO HOW TO GET THINGS RIGHT

我们的目标并不是让医生在检查项旁边乖乖打钩，而是要培养医疗团队注重合作和纪律的文化。

即便在手术室外，医生也要面对很多挑战。这些挑战也很危险，非常容易诱发错误，其中包括对心脏病、中风、药物过量使用、肺炎、肾衰竭、癫痫等病症的治疗；还有一些情况看似没有那么复杂危急，如对病人的头痛、胸部疼痛、肺部结节和乳房肿块进行诊断，实际上这些情况没有看起来的那么简单，它们都存在风险和不确定性。所以，即使在日常诊治过程中，清单也是值得使用的。优秀的清单能够像听诊器一样成为医生和护士的好助手。关键问题是，医疗行业的特定文化是否能够帮助我们抓住这个机遇。对此，我们还不能给出清晰的答案。

改变我们的核心价值观

在《太空先锋》（*The Right Stuff*）这本书里，汤姆·沃尔夫（Tom Wolfe）不但讲述了美国第一代宇航员的故事，还记录了一个辉煌时代的终结。

在20世纪50年代，极端危险的试飞工作主要靠像查克·耶格尔（Chuck Yeager）这样的王牌试飞员来完成。他们驾驶着推力巨大、难以驾驭的复杂飞行器冲向蓝天，将生死置之度外。据统计，有1/4的试飞员最终以身殉职。试飞员必须英勇、睿智，不仅要集中注意力，而且要随机应变，这些就是沃尔夫所说的“不可或缺的那些东西”。

但是，随着控制飞行风险的知识不断积累，随着飞行清单和模拟器变得越来越普及、越来越精确，试飞的危险系数正在不断下降。谨慎和一丝不苟成了更加重要的决定因素，试飞员从此不再像摇滚明星那样叱咤风云。

相似的过程也发生在医疗行业。对于最复杂、最危险的那些工作，如手术、急救和重症监护等，我们已经找到了更好的方法来完成它们。但是这些方法的使用需要我们大刀阔斧地改变传统文化，特别是改变我们的核心价值观：在高风险的复杂情况下，专家的胆大心细才是最重要的，这些也属于沃尔夫所说的“不可或缺的那些东西”。清单和标准的手术程序与这些东西迥然不同，这也是很多人对清单的使用产生抵触情绪的原因。

当然，如果缺少了勇气、睿智和随机应变的能力，医生光靠清单是不能治病救人的。医学是一门超级复杂的学问，如果少了建立在扎实专业基础之上的勇气，普通医生无法成为优秀的医生。但与此同时，一名优秀的医生也应该学会如何与他人进行有效的团队合作。

别让“可卡因头脑”作祟

还有很多行业看到了清单这个工具，但对于它的抵制也同样无处不在。金融业就是一个很好的例子。最近，我和莫尼斯·帕伯莱（Monish Pabrai）进行了交谈，他是帕伯莱投资基金（Pabrai Investment Funds）的合伙人之一，这家基金管理公司的总部位于加利福尼亚州欧文市。我最近碰到了3位将清单从航空业和医疗行业引入自己工作中的投资家，帕伯莱就是其中一位。

他们3人都掌管着巨额资金。帕伯莱管理着价值5亿美元的投资组合；盖伊·施皮尔（Guy Spier）是蓝宝石资本管理公司（Aquamarine Capital Management）的老总，该公司总部位于瑞士的苏黎世，旗下基金的规模达到了7 000万美元；第三位投资家不希望我披露有关他和他旗下基金的具体信息，我能告诉大家的是，他所管理的基金规模

在全世界名列前茅，价值高达数十亿美元。这3位习惯遵从价值投资理念，都喜欢买入价值被严重低估的优质公司股票，他们关注的焦点不是市场的波动趋势。这些投资家不会根据某些计算机算法来进行投资决策，他们希望自己独具慧眼，能够先于他人买入未来的“可口可乐公司”。

帕伯莱向我介绍了他是如何进行价值投资的。在过去的15年间，他每过3~6个月就要进行一次新的投资。为了找到一家真正有投资价值的公司，他往往要深入调查10家公司。他的消息来自各种渠道，如路牌广告、有关巴西房地产的报刊文章，或者是他随便拿起的一本矿业杂志。帕伯莱的阅读面非常广，视野非常宽阔，他会睁大眼睛去发现那些埋在土里但露出一丝光芒的宝石，也就是那些现在不为人知、但将来会飞黄腾达的企业。

帕伯莱能发现成百上千的投资机会，但是只需稍加分析，他就会放弃大多数所谓的机会。大约每隔一周，他就会因为发现了一笔可能的好买卖而心跳加速，他实在很难相信没有人发现这么好的机会。于是，帕伯莱开始在心里美滋滋地想，如果这次一切顺利的话，他能够赚到上千万美元，不，说不定是几亿美元。

帕伯莱说：“此时，你的内心会被欲望占据。”施皮尔将这种状态称为“可卡因头脑”。神经学家发现，赚大钱的预期能够刺激头脑中的原始奖励中枢，其效果和吸食可卡因是一样的。帕伯莱说，对一个专业投资人来说，越是在这个时候就越要保持头脑清醒，对意向公司进行系统而全面的调查。在进行分析的时候，他们会尽力排除非理性情绪的影响，做到既不过于乐观，又不过于悲观。他们会仔细分析公司的财务报表，了解公司的债务和风险，调查公司管理团队历年来的表现。他们还会对被调查公司的竞争对手进行分析，考虑整个市场的前景，对投资机会和风险进行全面评估。

价值投资这块圣地的守护神是沃伦·巴菲特，他是史上最成功的投资者之一。尽管2008年的经济危机给他造成了不小的损失，但在全球富豪榜上，他依然是数一数二的。帕伯莱对巴菲特以及伯克希尔-哈撒韦公司做出的所有投资决策，无论是成功的还是失败的都进行了仔细分析。他把自己能够找到的所有关于巴菲特的书籍都读了一遍。

帕伯莱甚至在慈善义拍中花65万美元拍下了与巴菲特共进午餐的机会。在和我谈论巴菲特的时候，帕伯莱直呼其名。我想，在花了65万美元与巴菲特共进午餐后，他直呼对方的名字也是顺理成章的事情。帕伯莱告诉我说，巴菲特的脑子里有一张清单，他用这张清单对潜在的投资机会进行评估。所以在旗下的基金成立之后，帕伯莱就一直坚持这一

原则。他严守纪律，每次在对公司进行研究的时候，他一定会要求自己不慌不忙。整个分析过程可能要花费数周时间，但坚持使用清单让他表现优异。可他并不是神，所以他也会犯错，甚至犯下灾难性的大错。

帕伯莱所说的错误并不是一般意义上的错误或错失良机。投资是一个高风险的行业，有些错误是不可避免的。而帕伯莱所说的是那一类原本可以避免的错误，是由于计算失误或分析不到位造成的“无能之错”。比如，帕伯莱在总结经验教训的时候发现，他总是在评估被调查公司财务杠杆风险⁽¹⁾的时候屡屡犯错。他能得到所有相关信息，但是他没有进行仔细分析。

帕伯莱相信，他之所以会在同一个地方屡屡跌倒，主要是因为无法让“可卡因头脑”冷静下来。

帕伯莱45岁，曾经是一名工程师。他的故乡在印度，而印度是一个人口大国，教育资源非常有限，他能够从异常激烈的竞争中脱颖而出实属不易。后来，他被南卡罗来纳州的克莱姆森大学（Clemson University）录取，并在那里学习工程学。大学毕业后，他在芝加哥和加利福尼亚的高科技公司谋求发展。在进入投资行业之前，他自己成功地创办了一家信息技术公司。

我之所以说这些，是想强调帕伯莱不是一个会轻易头脑发热的人，他知道要抵御追求满足当下欲望的诱惑。但他说，对于令人兴奋的投资机会，无论他如何努力保持冷静客观，头脑总是会诱骗他，让他能够轻易看到可以支持自己初始判断的证据，而对那些负面因素却视而不见。这就是“可卡因头脑”干扰理性思维的特征。

他说：“如果你抵挡不住诱惑，便将开始投机取巧。”

但在熊市里，情况恰恰相反。帕伯莱说，他的头脑会进入一种恐慌模式。因为看到周围有很多人输了个精光，所以他会不由自主地高估潜在的危险。

他还发现，自己在应对复杂性的时候也会犯错。如果他想做出高质量的投资决策，那么就需要对公司各个方面的情况进行多方位的调查。帕伯莱发现，即使“可卡因头脑”没有作祟，他还是会常常考虑不周，因为他头脑里的清单还不够完善。他说：“我不是巴菲特，我的智商达不到天才的水平。”所以，他需要一种即使智力一般的人也能使用的方法。因此，他设计了一张纸质清单。

显然，巴菲特也应该做一张纸质清单。帕伯莱注意到，即便是巴菲特也会经常重复某一类错误。他说：“这让我知道巴菲特并没有使用真正的清单。”

他将自己见到过的所有错误都列了出来，其中包括巴菲特、其他投资人和他自己犯过的错误，清单上很快就累积了几十种错误。为了防止这些错误再次发生，他为每种错误都设置了相应的检查项目，所以他的清单上已经有了70项检查内容，其中有一项就来自伯克希尔-哈撒韦公司所犯的错误。

该公司于21世纪初收购了科特家具租赁公司（Cort Furniture），这是一家总部位于弗吉尼亚的家具租赁公司。在过去的10年里，科特公司的业务量和利润增长迅猛。巴菲特的长期投资伙伴查理·芒格认为，科特公司抓住了美国经济转型期的一个重要机遇。由于商业环境变得越来越动荡，如今的企业需要掌握迅速扩张和迅速收缩的能力。所以，他们更愿意租用而不是购买办公楼，这同样适用于办公家具。芒格认为，科特公司会因为抓住这个机遇而大赚一笔。这家公司其他方面的情况也很不错，他们的财务状况很稳健，管理团队也很优秀，所以芒格下决心购买了这家公司。

然而这是一个巨大的错误。芒格忽略了一个事实，那就是该公司前3年收入的飞速增长，全都要归功于20世纪90年代末网络公司的兴起，科特公司为上千家初创企业提供了办公家具租赁服务。当网络经济泡沫破灭后，这些公司在一夜之间消失得无影无踪，科特公司的收入也因此而急剧下滑。

帕伯莱说：“芒格和巴菲特对网络经济泡沫一清二楚，但他们却忽视了科特公司的收入主要来源于网络公司这一事实。”芒格后来认为他在这次交易中犯了“宏观经济错误”。

“科特公司的盈利能力曾经非常出众。但在这次危机之后的较长时间里，他们几乎丧失了盈利能力。”在向股东承认错误的时候芒格如是说。

所以，帕伯莱将下面这个检查项目添加到了他的清单里：在对一家公司进行分析的时候，必须确认这家公司的收入是否因为经济周期的大幅波动而被高估或低估。

效率，清单带来的额外优势



和帕伯莱一样，那位匿名的投资家，暂且称他库克，也制作清单，但他做得更加细致。他会将投资过程各个阶段可能出现的问题

分别罗列出来。他把整个投资过程分为4个阶段，即研究阶段、决策阶段、执行阶段，以及交易完成后的问题监测阶段。他会根据各阶段的特点来设计清单以避免出现错误。这些清单都设置了清晰的检查点，每当到了这些关键节点，他和投资团队就会进行相应的检查。

比如，他有一张第三日清单。当对某家公司的分析进入第三天的时候，他们会使用这张清单进行检查。在这个检查点到来之前，他们应该已经完成了对这家公司过去10年间财务报表的分析。库克为报表上的每个项目和不同报表的关联项都制订了详细的检查项目。

库克说：“在一张财务报表上作假并不难，但要用各类财务报表把谎话给编圆了就没那么容易了。”

比如，有一个检查项目要求研究人员仔细分析现金流量表的脚注，而另一项则要求他们仔细阅读重大管理风险陈述，还有一项要求研究人员检查公司的现金流和成本是否与公司报告的收入增长相吻合。

他说：“这些都是最基本的分析步骤。但是你瞧，人们就是会经常忘记它们，这非常不可思议。”就拿安然公司的惊天丑闻来说吧，人们完全可以从公司公开的财务报表中发现他们作假。

库克告诉我这样一个故事，曾有一家公司看似非常有前途，这让投资者的“可卡因大脑”又开始“嗡嗡”作响。但事实上，公司高层管理人员在向潜在投资者吹嘘公司前景的同时，已经把自己持有的股份全都悄悄变卖了。这家公司即将遭受巨大亏损，而投资者却全然不觉。在库克的第三日清单中有这么一项：研究人员必须仔细阅读公司披露的重要股权变更信息，于是，他发现了这个不可告人的秘密。库克说，每**50**次检查可能会有**49**次一无所获，但就是这一次发现却能让你避免重大损失。

清单并不会列出具体的操作步骤，但它能帮助库克在投资过程中尽力保持冷静而睿智的头脑，确保他能在必要的时候得到所需的重要信息，从而系统地进行决策，并和每一个应该沟通的人进行充分交流。库克相信，在清单的帮助下，他和他的伙伴能够将人类的决策能力发挥到极致，所以他完全有信心击败市场。

清单能够帮助人们在投资过程中尽力保持冷静而睿智的头脑，确保在必要的时候得到所需的重要信息，从而系统地进行决策，并和每一个应该沟通的人进行充分交流。

我问库克他怎么知道清单一定奏效。

他借用外科手术的例子（他看过手术安全清单）向我解释说：在外科医生确认自己洗干净了手，而团队成员彼此进行了充分交流后，在没有增长专业技能的情况下，他们提高了自己的绩效水平；在我们这个行当里，使用清单也能起到相同的作用。

库克不会将使用清单的具体结果告诉我，他的基金不会公开披露投资收益。但他说，他已经看见了明显的效果。他是从2008年开始使用这套清单的，至少他的基金平安度过了那一年的金融风暴。其他人对他的基金给出了更高的评价，认为他是业内的佼佼者。他取得的成功有多少能直接归功于清单的使用呢？库克无法给出明确的答案。毕竟，自清单投入使用以来只过去了两年。但库克可以肯定的是，在变化越来越快的商业环境里，清单让他们具有了额外的优势，那就是效率。

对此我感到非常意外。当库克刚刚引入清单的时候，他认为这会增加决策所需的时间和工作量，投资团队的决策速度会慢下来。对此，他已经做好了心理准备，并愿意付出相应的代价。毕竟，少犯错能够让他们避免巨大损失，这笔生意划得来。而事实证明，他们进行前期调研的时间虽然增多了，但从整体上来看，他们却能在较短的时间里对更多的投资机会进行评估，而整体效率提高了不止一星半点。

在变化越来越快的商业环境里，清单让人们具有了额外的优势，那就是效率。

库克介绍说，在使用清单之前，他们往往要花费数周时间、开数次会议才能决定到底是放弃一家候选公司，还是对其进行更加深入的调查。这一过程并没有截止日期，常常会引发危险。当人们对一家公司仔细研究了一个月以后，他们往往会决定对其进行投资。但在使用了清单后，他和他的团队发现，他们完全可以在执行第三日清单的时候就做出这一决定。他说：“虽然决策过程变得更加仔细，但速度却更快了。这就像棒球比赛，我们一旦挥杆把球击出去，就可以全神贯注向前飞奔了。”

帕伯莱和施皮尔也发现了相同的现象。施皮尔以前还雇用了一个投资分析师，而现在不需要这个帮手了。帕伯莱使用清单已经有一年时间长了，他管理的基金市值已经增长了一倍以上。当然，这不能完全归功于清单的使用。不过他发现，清单确实能够帮助他大幅缩短投资决策的时间，而且让投资分析变得更加完整和系统。

2008年下半年，股票市场开始暴跌，投资人纷纷恐慌性地抛售手中的股票，物美价廉的股票在市场上比比皆是。在一个季度里，帕伯莱对100多家公司进行了调查，并将其中10家公司的股票添加到自己的证券组合中。他说如果缺少了清单的帮助，自己根本无法完成那么大量的分析工作，而且对分析的结果也不会那么有信心。一年之后，他旗下投资组合的市值增长了**160%**。帕伯莱完全没有犯错。

这些投资家的经历之所以让我感到好奇，不仅因为使用清单在投资领域同样有效，还因为如此有效的方法在这个行业里同样乏人问津。在金融市场，每个人都在绞尽脑汁赚更多的钱。如果有人表现不错，那么人们就会像饥肠辘辘的鬣狗一样扑上去，试图找出此人获得成功的秘诀。各种投资诀窍层出不穷，如投资网络公司，购买抵押贷款证券等，但其中绝大多数都很快被市场这个巨大的黑洞无情地吞噬了，而使用清单这种方法却生存了下来。我问库克有多少人对他使用的清单感兴趣，他回答说：“没有人。”事实确实如此，人们对他购买了哪些公司的股票及购买这些股票的原因非常感兴趣，但当他刚刚说出“清单”这

两个字的时候，人们却一哄而散。即便是在自己的公司里，他也发现很难推销这一理念。

库克说：“我四处碰壁，我的团队花了几个月的时间才发现清单的价值。”直到今天，库克的部分合伙人依然不使用这种方法，特别是当库克不在的时候，他们就会把清单晾在一边。

他说：“让我感到十分惊讶的是，其他投资人连试一试都不愿意。有些人也曾问过我具体的方法，但没有人真的把它当回事。”

运用清单系统地完成评估工作

人们对清单的抵触可能是不可避免的。几年前，克莱蒙特研究生大学（Claremont Graduate University）一位名叫杰夫·斯玛特（Geoff Smart）的心理学家进行了一项足以说明问题的研究。

他对51位风险投资家进行了调查，这些人要对尚未盈利的初创企业进行高风险投资，每次出手都有几百万美元。他们与帕伯莱、库克和施皮尔这样的基金经理非常不同，后者只投资成熟的企业，这些企业有多年的公开财务报表可供分析；而风险投资人要在一堆两眼放光、头发油腻、乳臭未干的创业者身上下注，除了一些可能会赚钱的想法和一些笨拙的设计原型以外，他们没有任何参考依据。但谷歌和苹果就是这么诞生的，风险投资家相信自己能够找出下一匹黑马并拥有它。

对风险投资人来说，最难做的决定就是是否要把大把钞票投给某个创业者。斯玛特对他们的决策过程进行了仔细研究。你可能会认为风险投资家会根据创业者的创业理念的好坏来进行决策。但实际上，找到好的创业理念并不是一件特别困难的事情，真正困难的是找到能够把好想法变成现实的创业者。要把好的理念变成可行的行动计划，创业者需要投入大量的时间和精力建设团队，应对各种压力和挫折，解决技术和人事方面的问题，还需要有钢铁般的意志和非同一般的专注力并坚持数年。这样的人非常罕见，简直是凤毛麟角。

斯玛特发现，风险投资家挑选合适人选的方法有很多种。实际上，不同方法反映出的思维方式也不相同。

有一类投资者被称为“艺术评论家”，他们只要看一眼，就能知道眼前的创业者够不够格，就像艺术评论家品鉴一幅油画是不是真品一样。他们靠的是直觉和多年积累下来的经验。

还有一类投资者被称为“海绵”投资家，他们会花更多时间收集有关候选人的信息。他们像海绵一样，在面试、实地考察以及和举荐人的交流过程中收集有用的信息，然后再根据自己的感觉进行决策。正如一位“海绵”投资家告诉斯玛特的那样，他们会随机挑选某些方面收集信息，这样就不会被大量的分析工作累垮。

还有“检察官”投资家，他们会像审犯人一样对创业者进行盘问，会用刁钻的问题来考验对方，让对方解决一些专业难题，或让他们回答如何处置一些假设的情况。而“求婚者”投资家在乎的是如何与对方联姻，而不是如何评估对方。“终结者”投资家则认为选择合适人选的努力是徒劳的，这类投资家不会浪费时间对创业者进行评估，而是会从他们手里买下很有前景的商业计划，炒掉无能的创业者，然后雇人代替他们。

还有一类投资家被斯玛特称为“机长”投资家，他们用清单系统地完成评估工作。他们会对自己和他人犯过的错误进行研究，并从中吸取教训。他们还会编制正式的清单以防止此类错误再次发生。“机长”投资家强迫自己严守纪律，即便是直觉告诉他们某个候选人非常出色，他们还是会严格执行各项检查程序。

斯玛特对各类风险投资家的业绩进行了跟踪调查。到底谁能更胜一筹呢？我想你能轻松猜到这个问题的答案，那就是“机长”投资家。研究发现，“机长”投资家后来因为创业者无能而将其炒掉，或承认自己评估错误的概率是**40%**，比其他类型的投资者低**10%**。

“机长”投资家的投资回报也很有说服力。

清单的力量 | THE CHECKLIST MANIFESTO HOW TO GET THINGS RIGHT |

调查显示，“机长”投资家的投资回报中位数是80%，而其他类型投资家则不到35%。这并不意味着其他投资家的投资能力比“机长”投资家逊色，经验的确很重要，但清单的使用能够给投资者带来更大的成功。

最有趣的发现是，大多数风险投资家属于“艺术评论家”和“海绵”投资家，他们都是不进行系统分析的直觉型决策者。在调查对象

中，“机长”投资家的人数只占1/8。你或许会认为，这是由于大多数人不知道清单而造成的。但事实上，即使人们知道了也不会做出什么改变。斯玛特在很多年前就发表了自己的研究成果，后来他还撰写了一本名为《聘谁》（*Who*）的书对此进行详细说明。现在大家应该已经知道采用这一方法的好处了，但当被问及“机长”投资家所占的比例是否有所提高时，斯玛特却说：“没有，他们所占的比例依然很小。”

建构注重合作和纪律的文化

人们不喜欢清单，因为它们需要人们投入大量精力，而且并不有趣。但我并不认为人们这么想仅仅是出于懒惰。我觉得，大家之所以对能够帮助人们救命或挣钱的清单视而不见，有着思想上的深层原因。我们有时候会不由自主地觉得使用清单会让人感到尴尬。在我们内心深处，那些站在风口浪尖、从容不迫的大英雄是不会使用清单的。他们英勇果敢，随机应变，从不循规蹈矩。

也许，我们需要给英雄重新下个定义了。



2009年1月14日，世界卫生组织发布了手术安全清单。碰巧的是，就在第二天，全世界见证了航空史上的一次伟大奇迹。那一天，全美航空公司1549号航班从纽约拉瓜迪亚机场（La Guardia Airport）起飞，机上载有155人。不幸的是，飞机在曼哈顿上空撞上了一大群加拿大鹅，导致飞机两台发动机全部停止运行。但不可思议的是，飞行员成功地将飞机迫降在冰冷的哈德逊河上，而且没有一人在事故中死亡。因此，媒体将这一事件称为“哈德逊河上的奇迹”。美国交通运输安全委员会（National Transportation Safety Board）的一位官员认为这是“航空史上最成功的一次水面迫降”。57岁的机长切斯利·萨伦伯格三世（Chesley B. Sullenberger III）[\(2\)](#)曾是一名空军飞行员，飞行时间累计两万多小时。事件发生后，他成了举世瞩目的英雄。

在报道这一事件的时候，《纽约邮报》（*New York Post*）使用了这样的大字标题：“萨伦伯格机长是一位低调的飞行英雄。”ABC新闻则称他为“哈德逊河上的英雄”；德国报纸毫不吝啬地称他为“纽约英雄”；西班牙媒体也称他为“纽约的英雄”；法国媒体使用的则是“新诞生的美国英雄”；小布什亲自打电话向他表示感谢；而奥巴马则邀请他和他的家人参加5天后举行的就职仪式。萨伦伯格的家住在加利福尼亚的丹维尔，全球各地的摄影记者纷至沓来，想要抓拍他妻子和孩子的照片，他们家的草坪都被踩烂了。在当地，人们还为萨伦伯格举行了盛大的游行，而书商则与他签下了价值300万美元的出版合约。

当公众通过各种渠道得知有关这一事件的更多细节之后，他们发现萨伦伯格并不是单枪匹马的孤胆英雄，还有其他很多因素共同创造了这样一个奇迹，如飞行员使用的标准程序和清单、帮助飞行员进行完美滑翔的电传操纵系统、与萨伦伯格共同沉着应战的副机长，还有在飞机迫降后帮助所有乘客迅速撤离的空乘人员。正如萨伦伯格在首次接受新闻采访时所说的：“我想更正一下大家的看法，奇迹的创造不是我一个人的功劳，这是所有机组成员共同努力的结果。”他认为，在这起事件中，他个人的飞行技能可能很重要，但不容忽视的是，团队合作和坚决执行标准程序也同样非常重要。

不过，公众最后还是将他的这番话理解为自己的谦虚。一个月以后，虽然纽约市为5位机组成员而不是为萨伦伯格一个人颁发了荣誉钥匙，虽然他们共同接受了各大新闻网的“独家”采访，而且在当年的橄榄球总决赛开赛之前，他们来到了坦帕湾（Tampa Bay）的雷蒙德·詹姆斯体育场（Raymond James Stadium）接受全场7万多名观众的起立喝彩，但是媒体已经为这一系列报道定下了基调，他们不想谈论团队合作和标准程序，而只是关心一个故事，那就是萨伦伯格如何用当年在军校里学到的滑翔机驾驶技巧，创造出航空史上的水面迫降奇迹。萨伦伯格说：“驾驶滑翔机是在很多年前的事情了，而且那些滑翔机和现在的喷气客机非常不同，我觉得这些经验没有太大作用。”

正如哈德逊河迫降事件，人们似乎就是不能接受这样一个现实：如今，创造奇迹光靠单枪匹马是不行的，而要靠各方因素的有效协同。

萨伦伯格驾驶的是一架由空客生产的A320喷气客机。这种飞机有两台发动机，两侧机翼上各吊挂了一台。事故发生当天，天气非常寒

冷，但非常晴朗。1549号航班于当天下午3:25起飞，目的地是北卡罗来纳州的夏洛特。当时，飞机由副机长杰弗里·斯基尔斯（Jeffrey Skiles）驾驶，萨伦伯格则担当“不把杆飞行员”的职责。值得一提的是，两人在此次飞行之前从来没有合作过，但他们都是经验丰富的老飞行员，斯基尔斯的飞龄与萨伦伯格相当。他以前的主飞机型是波音737，并一直担任机长。后来因为公司精简人员，他才不得不改飞空客A320，并降格为副机长。你可能会认为两位飞行员的经验都很丰富是一件好事，但事实未必如此。

不妨让我们来设想一下这样的情景，两位经验非常丰富但素未谋面的律师为被告人在法庭上进行辩护，或者在NBA总决赛即将打响之际，两个互不相识的顶级教练同时执教一支队伍。这时候，事情很可能会被搞砸。

不过，在使飞机离开登机口和启动发动机之前，两位飞行员严格执行了标准的操作流程。其他大多数行业的专业人士往往不会这么做。他们共同执行了检查程序，确认每一位机组成员都进行了自我介绍。他们还进行了飞行简报，对飞行计划、可能发生的问题，以及相应的应急预案进行了讨论。正是因为没有跳过这几分钟的标准程序，他们不但检查了飞机的安全状况，而且顺利地组成了一个团队。这个团队将有效分工合作，共同面对将要发生的一切。

在常人看来，像萨伦伯格和斯基尔斯这样的资深专家可以轻松跳过这些无聊的标准程序。所有机组成员的飞龄加起来已经有150年，在如此漫长的职业生涯里，他们一次又一次地执行清单，不断在模拟舱里进行练习，而且每年还要学习更新过的检查项目。这些例行检查在大多数时候是没有什么意义的，他们从未出过飞行事故，完全有理由相信自己能够安全地结束职业生涯。他们碰到飞行问题的概率极低，远远小于医生、投资者或律师等专业人士遇到问题的概率，但他们仍然一丝不苟地执行清单。

他们并不一定需要这么做。仅仅是在30多年前，一些飞行员在执行清单的时候依然马马虎虎。他们会说自己从来就没有发生过问题，或者说：“让我们起飞吧，一切都会顺利的！”他们还会说：“我才是这架飞机的机长，这些清单是在浪费我的宝贵时间。”就拿1977年的特内里费（Tenerife）特大空难来说吧，那是有史以来死亡人数最多的一次空难。

两架波音747大型客机在加纳利群岛一条大雾弥漫的跑道上高速相撞，导致583人死亡。肇事飞机属于荷兰皇家航空公司（KLM），该机机长错误地将塔台发出的等待

起飞指令理解为准许起飞指令。随机工程师觉得塔台的指令不清，向机长发出了警告，但机长对此置若罔闻。当飞机在跑道上开始加速的时候，泛美航空公司的另一架747飞机尚未撤离跑道。

“泛美的飞机是不是还在跑道上？”随机工程师向机长提问说。

“已经撤离了。”机长固执己见，并继续推油门杆。

机长错了。虽然随机工程师的判断是正确的，但是机组并没有为这一判断做好准备，他们没有成为一个真正的团队。所以随机工程师从未想过自己可以打断机长，让他停下来搞清楚状况，就更别提想到自己有责任这么做了，而一意孤行的机长将他们全都送上了黄泉路。

人们之所以不喜欢执行清单，是害怕自己会变得死板。对于以前的飞行员来说，他们可不想让自己变成机器人，只知道埋头于清单，而对风挡外的现实世界不管不顾。但精心设计的清单恰恰不会让你变得死板，它们会帮助你节省有限的脑力，不让你的头脑被繁杂的检查项目所占据，将你解放出来去处理更加重要的问题。如：我们应该在哪里降落？升降舵是否已经解锁？病人是否已经注射了预防性抗生素？基金经理是否已经卖出了所有股份？所有相关人员是否已经到场？

清单宣言 | THE CHECKLIST MANIFESTO HOW TO GET THINGS RIGHT

精心设计的清单会帮助你节省有限的脑力，不让你的头脑被繁杂的检查项目所占据，将你解放出来去处理更加困难的问题。

不妨让我们看看下面这张非常简短的清单都包括哪些内容。



这是一张专门为塞斯纳（Cessna）单引擎飞机可能遇到的发动机在空中停止运行的情况（与全美航班遇到的情况相同）而设计的清单。这种小飞机由单人驾驶，不像全美航空的A320那样由双人

驾驶。这张清单精简到只有6个步骤，如不要忘记重新启动发动机，确认燃油阀位于开启状态，打开备用油泵等。但最吸引人的是这张清单上的第一步，这个步骤说得非常直白，那就是驾驶飞机！飞行员有时候会惊慌失措，一味地埋头重启发动机，或一个劲儿地思考出了什么问题，却把最基本的任务抛在了脑后。可见这样的清单一点也不死板，它们会帮助人们尽可能减少出错的概率。

清单创造的生命奇迹

起飞一分半钟以后，全美航空1549号航班爬升到1 000米高度。就在这时，他们撞上了一群加拿大鹅。撞击发生得非常突然，萨伦伯格的第一反应是俯身躲避。飞鸟撞击风挡和引擎而发出的声音非常大，舱音记录器也将它们记录了下来。正如后来的新闻报道所说的那样，飞鸟撞击飞机而引起的事发生过成百上千起，但造成严重后果的非常罕见，遭受两只以上飞鸟撞击的飞行事故就更加罕见了。

工程师在设计飞机引擎的时候专门考虑过这一问题。不幸被吸入引擎的飞鸟多多少少会被汽化，但加拿大鹅的体型很大，在鸟类中名列前茅，体重往往会超过5千克，任何一种飞机引擎都没办法完全汽化它们。如果不能将飞鸟汽化，喷气发动机会自动停止运行，这样就不会发生爆炸，发动机残片也不会飞出去刺穿机翼和机身，或伤害到旅客。而这架飞机的两台引擎竟然吸入了超过3只加拿大鹅，这种情况少之又少。所以，在撞击发生后，两台引擎立刻停止了工作，飞机随即失去所有动力。

当这一切发生后，萨伦伯格做出了两个重要决定。首先，他从副机长斯基尔斯那里接管了飞机的操控权。其次，他决定在哈德逊河上迫降。在当时的情况下，似乎没有其他选择，萨伦伯格不假思索地做出了决定，因为他们发现飞机的速度不够快，飞不回拉瓜迪亚机场，也飞不到空管人员建议的新泽西州泰特伯乐机场。

至于谁来驾驶飞机这个问题，两人的飞行经验都非常丰富，但萨伦伯格驾驶A320的经验比斯基尔斯多很多。而且飞机左侧有很多需要避开的地标性建筑，如曼哈顿的摩天大楼、乔治·华盛顿大桥等，而机长的座位位于驾驶舱左侧，所以机长驾机比较容易看清这些地标，并躲开它们。此外，斯基尔斯刚刚完成A320紧急情况处置培训，对于马上就要用到的清单也只是刚刚熟悉。所以，由萨伦伯格驾机是明智的选择。

“我驾驶飞机。”萨伦伯格一边用标准用语说，一边把手放到了操纵杆上。

“你驾驶飞机。”斯基尔斯答道。他们并没有争论接下来该做些什

么，甚至都没有进行讨论。实际上他们也不需要这么做，因为在执行起飞前的准备程序时他们已经组成了团队，交流了紧急情况的处置预案。萨伦伯格将负责寻找距离最近、安全性最高的着陆场所，而斯基尔斯则负责执行应对发动机停止运行情况的清单程序，尝试再次启动发动机。除了近地警告系统不断发出的“拉起”警报声以外，机舱里听不到其他声响。两个飞行员都专注于自己的任务，并时不时地观察同伴的动作以便使得整个操作协调一致。

在整个事件中，两个飞行员都扮演着非常重要的角色。我们总觉得副机长是摆设，以为他们只是因为闲着无聊才找些事情做。但现在的飞机非常复杂，安全飞行既离不开机长的努力，也离不开副机长的协助。就好像手术的实施既需要外科医生，也需要麻醉医生一样。

在执行飞行任务的过程中，两个飞行员中的一个负责操控飞机，而另一个则主要负责操控各种飞行仪表和执行清单，他们往往会在返程时交换彼此的职责。而当飞机碰到紧急情况的时候，我们很难判断谁的工作更艰巨，更具挑战性。这架飞机的滑翔时间只有3分半钟，在这么短的时间里，斯基尔斯要完成重启发动机的所有程序，如果发动机重启失败，他还要为飞机进行水面迫降做好准备。但仅仅是重启一台发动机的时间就很可能超过3分半钟，他必须迅速做出选择。

飞机快速向地面冲去，斯基尔斯果断做出判断，他认为求生的最大希望是迅速启动一台发动机，所以他决定把绝大部分精力集中在执行应对发动机停止运行的清单程序上，而且要以最快的速度执行完毕。他们并不清楚发动机的损坏情况到底有多严重，但只要能够恢复部分动力，飞机就能重返机场。最后，斯基尔斯竟然完成了重启两台发动机所需的所有操作。调查人员事后模拟了当时的情况，但他们无法在那么短的时间里完成那么多操作。

更不可思议的是，斯基尔斯并没有放弃水面迫降程序。虽然他没有时间完成清单上列出的所有步骤，但他至少让飞机成功发出了遇难信号，并尽全力让飞机的各种装置和操控面符合水面迫降要求。

“襟翼是否放下？”萨伦伯格问道。

“襟翼已经放下。”斯基尔斯答道。

萨伦伯格竭尽全力控制飞机向水面滑去。但即便是在这个时候，萨伦伯格也不是凭借一己之力完成高难度动作的。正如也曾是飞行员的记者威廉·朗格维舍（William Langewiesche）所说的那样，飞机的电传操纵系统能够让飞行员轻松驾机并做出完美的滑翔动作，这种系统并不需

要飞行员掌握特别高超的飞行技艺，并且能够防止飞机发生机头偏转或机身摇晃的情况。

在飞机转向的时候，电传操纵系统能够自动调整飞机的尾舵以避免机头发生侧移。在飞机的电子仪表主显示器上会显示出一个绿点，萨伦伯格只需将飞机的矢量标记对准这个绿点就能让飞机完美滑翔。而且这套系统还能让飞机保持获得升力的最佳攻角，防止飞机抬头过高，这样飞机就不会因为失速而坠落。在电传操纵系统的帮助下，萨伦伯格可以把注意力集中到更加重要的问题上，如尽量在渡口旁寻找着陆点以便乘客获得救援，还有在飞机接触水面的时候尽量让机翼保持水平。

与此同时，希拉·戴尔（Sheila Dail）、唐娜·登特（Donna Dent）和多琳·韦尔什（Doreen Welsh）也在座舱内执行着她们的清单。

她们要求乘客低头俯身，用双手抓住脚踝以缓冲着陆时的巨大撞击。待飞机迫降后，空乘人员指导乘客穿上救生衣。她们要确保在飞机停下后迅速打开舱门，确保乘客不会因为拿行李而浪费宝贵的逃生时间，确保他们不会因为救生衣提前充气而被卡在飞机里。冰冷的河水从机身裂缝中渗进了飞机的后舱，而在后舱值守的韦尔什蹚着齐胸深的刺骨冰水坚持完成各项操作。虽然在4个紧急出口中只有2个能够顺利开启，但在空乘人员的共同努力下，所有乘客在3分钟内全部撤离随时会下沉的飞机，没有比标准要求多用1秒钟时间。

当疏散工作正在进行的时候，萨伦伯格来到座舱，查看乘客撤离的情况，并检查飞机迫降后的状况。而斯基尔斯则留在驾驶舱里执行疏散清单，防止起火等情况发生。直到所有程序都执行完毕后，他才撤离。大量渡轮和船只从四面八方赶来，哪怕飞机上有更多的乘客，也不会有人因为船上空间不够而继续泡在冰冷的河水中。由于飞机的油箱只装了一半油，里面的空气让整架飞机稳稳地漂浮在水面上。于是，萨伦伯格有时间进行最后一项检查，他沿着过道走遍了整个座舱以确保没有一位乘客被落下。最后，他的身影终于出现在舱门口。

这起事件的整个处理过程出奇地顺畅。萨伦伯格后来对媒体说：“当我再次见到副机长斯基尔斯的时候，我们不约而同地说了一句，‘情况看起来并没有我们想象的那么糟糕。’”

到底谁是真正的英雄呢？我们实在是找不出答案。这的确是一个奇迹。首先，他们的运气非常不错。由于事故发生在白天，所以飞行员能够目视寻找安全的着陆地点。其次，着陆地点附近有大量船只，这样乘客就不会因为浸泡在冰冷的河水里而体温过低。事故发生时，飞机具有一定的高度，可以避开乔治·华盛顿大桥。最后，飞机是在河上顺着波

浪运动的方向降落的，所以，迫降造成的冲击力不会像飞机逆流降落或在海上迫降那么大。

即便如此，机上的155人还是有可能命丧黄泉。拯救他们生命的是比高超飞行技艺更不寻常、更难得的一种品质：全美航空**1549**号航班的全体机组人员能够在千钧一发的时刻坚决执行重要的清单，并保持沉着冷静。他们很清楚何时应该随机应变，何时应该严守纪律。他们知道如何应对复杂而危险的情况，并为此通过有效的团队合作预先做好了准备。

这才是他们真正与众不同的地方，是他们能够在这个时代成为英雄的关键所在。这些品质是非常稀罕的，我们必须清醒地意识到，其他行业也同样需要这些宝贵的品质。

纪律，职业精神的内涵

所有行业都有相对应的职业精神。职业精神反映出从业人员应该具有的品质和必须承担的责任。这些准则有时是成文的，有时是不成文的，但至少要有4个共同的组成要素。

第一条是无私。无论我们是医生、律师、教师、政府官员，还是士兵或飞行员，作为一个有职业精神的人，我们应该把委托人的利益放在自己的利益之上。

第二条是训练有素。我们应该追求卓越，掌握更多的知识和专业技能。

第三条是值得信赖。我们要对个人行为和肩负的担子负责。

第四条是遵守纪律。这是飞行员为他们的职业精神特别添加的一条。他们坚决执行审慎的程序，并与其他机组成员通力合作。大多数行业的职业精神往往少了这一条，包括医疗行业在内。医生总是把自主当作最重要的职业素养，这与严守纪律是相互矛盾的。在如今这个复杂的世界里，成功往往与大型企业、专业团队和高风险技术这些元素，以及超出个人认知范围的复杂知识密切相关。我行我素不应该成为我们追求的风格，这种风格反映的更多的是自我保护，而不是卓越不凡。个别行业的职业精神偶尔会要求人们达成“共治”，此时人们在工作中仅仅做到彼此尊重是远远不够的，还需要遵守纪律。

我行我素不应该成为我们追求的风格，这种风格反映更多的是自我保护，而不是卓越不凡。

遵守纪律是很困难的，这要比做到值得信赖、训练有素甚至是无私都要困难。我们天生就有弱点，而且善变，很多人就连坚持不吃零食都做不到。我们天生不喜欢遵守纪律，而喜欢寻求新鲜和刺激的东西，不喜欢关注细枝末节。为了遵守纪律，我们必须付出一定的努力。

这可能就是航空业要求将遵守纪律变成一种规范的原因。20世纪30年代，一小批空军飞行员发明了起飞前清单。这一发明的巨大力量让许多机制应运而生。在美国，我们有交通运输安全委员会专门对航空事故原因展开独立调查，并就如何避免同类事故提出建议。我们还制定了全国性的法规以确保这些建议变成实用的清单，并让各家航空公司的飞行员认真执行。

防范错误，我们别无选择

我们不能让清单变成僵化的教条，应不断地对其进行改进，就算是最简单的清单也不例外。飞机制造厂商在每一张清单上都印刷了发布日期，这说明清单的内容会时常发生变动。最后，我们必须认识到，清单只是一种提供帮助的工具，如果它起不到这个作用，就没有存在的必要；但如果它真能起到这个作用，我们就应该敞开心扉接受它。

我们不能让清单变成僵化的教条，应不断地对其进行改进，就算是最简单的清单也不例外。

人们很乐意向计算机寻求帮助，这种高科技产品向我们展现了通过使用自动化工具来防止人类犯错的美好前景。计算机的确能帮我们做很多事情，如计算、处理、存储和传输。毫无疑问，计算机所代表的科技能够增强我们的能力。但还有很多事情是科技难以做到的，如应对不可预测的事情，建造高耸入云的大厦，或实施能够挽救生命的手术。而且

在很多方面，科技让这些活动变得越来越复杂。人类赖以生存的很多系统已经非常复杂了，而科技在某种程度上增加了这些系统的复杂性，让我们更容易犯错。

现代生活的一大特点是对于各种系统的依赖。我们要在系统中将不同的人、不同的技术整合起来。由此引发的最大问题就是如何让系统有效运作。比如在医疗过程中，如果我希望病人能够得到最好的治疗，那么光靠我一个人的努力是不够的，手术团队的所有成员和各种手术设备必须有效协同运作。从系统的角度来看，医疗过程和汽车的生产过程是一样的，这是唐纳德·伯威克（Donald Berwick）提出的观点。他是波士顿医疗卫生改善协会（Institute for Healthcare Improvement）主席，对如何构建理想的医疗系统做了很多深入的思考。在他看来，无论是生产汽车，还是治病救人，在工作系统中只有优秀的组成部分是远远不够的。

清单宣言 | THE CHECKLIST MANIFESTO HOW TO GET THINGS RIGHT

清单是一种提供帮助的工具，如果它起不到这个作用，就没有存在的必要；但如果它真能起到这个作用，我们就应该敞开心扉接受它。

在医疗系统中，我们总是热衷于拥有完美的组成部分，想得到最好的药物、最好的设备和最优秀的专家。但我们却很少思考如何让这些部分互相匹配，进而组成最有效的系统。伯威克认为这种理念是非常错误的，他说：“所有对系统概念略知一二的人都会很快意识到，仅仅通过最优化系统的各个部分并不能创造出最优的系统。”他为我们举了一个有关用最优秀的部件组装汽车的例子：如果把法拉利的引擎、保时捷的刹车、宝马的悬挂和沃尔沃的车身组装在一起，你能得到什么呢？他说：“我们得到的不可能是最好的汽车，只能是一堆昂贵的垃圾。”这听起来似乎非常好笑，但在医疗行业，我们恰恰就是这么做的。

美国国家卫生研究院每年花费300亿美元开展各项医学研究，取得的成果非常丰硕。但我们却没有致力于分析如何将这些伟大的研究成果应用于日常实践的国家卫生系统革新研究院，没有像交通运输安全委员会那样迅速对医疗事故原因展开调查的组织，没有像波音公司那样编制相应医疗过程清单的机构，更没有每月跟踪改进措施实施效果的组织。

相同的情况也出现在其他很多行业。教师、律师、政府官员和金融从业者也不会对经常犯的错误进行仔细研究，不会为防止这些错误再次发生而寻找解决方案，更不要提对解决方案不断进行改进了。

清单宣言 | THE CHECKLIST MANIFESTO
HOW TO GET THINGS RIGHT

请承认人类的不完美，每个人都会犯错。在极端复杂的世界里，我们必须改变观念，因为我们别无选择。

但我们应该这么做，这是我最想强调的一点。每个人都会犯错，会忽略细节，会一时想不起学过的知识。但在大多数情况下，人们想到的只是努力，再努力。一方面，努力提高自己的技艺以减少犯错的可能性；另一方面，在问题发生后努力减少损失。但几十年前发明清单的飞行员可没有这么想。当崭新的299型轰炸机展现在他们面前时，他们想到的是没有人能够保证自己在驾驶这么复杂的飞机时不犯一点错误。他们也可以让飞行员接受更多培训，或者把飞行事故简单地归咎于飞行员技艺不精。但他们没有这么做，而是承认人类的不完美，并且发明了简单、实用且有效的清单。

我们也可以学习他们的做法。实际上，在极端复杂的世界里我们必须这么做，因为我们别无选择。只要稍加留意我们就会发现，人们会一遍又一遍地犯相同的错误，即便是那些英勇果敢、才华横溢的人也不例外。我们已经知道了人类犯错的必然性，但仍为此付出了沉重的代价。所以说，该尝试一下新办法了。

让清单革命融入你的观念

2007年春，当我们的手术安全清单渐渐成形的时候，我开始在自己的手术室里使用这张清单。之所以这么做，并不是因为当时迫切需要它，而是因为我想确认它的确有用。此外，我们即将在全球8家医院展开测试，我也不希望别人觉得我只动员别人，自己不用，所以，我想最好亲自试用一下这张清单。你可能会问我是否打心底里觉得“这一发明真的可以让我自己的手术变得更安全，让我少犯很多愚蠢的严重错误”。假如你扒光我的衣服，让我躺在手术台上，并且威胁我，如果我

不说实话，你就不实施麻醉的情况下活生生地切除我的阑尾，那么我会告诉你：“别，千万别！我可不想犯什么大错，清单的效果最好不要体现在我身上！”

但让我感到难堪的是，自从这张清单投入使用以来，我的团队没有一周能逃过它的法眼。就拿上周来说吧，仅仅在5次手术中，我的团队就被抓住了3次。

第一次是我发现有一个病人在手术前没有注射预防性抗生素，这是最常见的一个错误。麻醉团队忙于各种事务，所以分了心，光是寻找适合注射的静脉血管就花了很多时间。还有一个监视器也突然“罢工”了，给他们制造了不少麻烦。这时，护士让大家停下来，一起执行针对病人皮肤切开前的检查清单。

“病人是否在60分钟内注射了抗生素？”我一边看着墙上贴着的清单一边说。

“嗯，还没有，我马上就给他注射。”住院实习麻醉医生回答道。整个手术团队等了1分钟时间，然后助理护士才把手术刀递给我。

第二次是一个病人特地提醒我们不要给她注射抗生素。她说抗生素会让她感到肠胃不适，而且会引发酵母菌感染。她很清楚注射预防性抗生素的好处，但由于她的这台手术引发伤口细菌感染的可能性非常小，不到1%，所以她愿意冒这个风险。但注射抗生素是一种例行操作（当我们没有分心时），以致我们差一点就为她注射了。

第一次是在麻醉前，她自己发现了问题，并提醒了我们。第二次则是在麻醉后，我们在执行清单的时候发现了这个问题。当时手术就要开始了，在交流注意事项时，护士提醒大家不要给病人注射抗生素。麻醉医生感到非常惊讶，因为她正准备给病人注射，此前病人和我们交谈的时候她并不在场。

第三次则是在为一个60多岁的老妇实施甲状腺切除术的时候。她的甲状腺可能会发生病变，所以我要将她的半个甲状腺切除。她患有许多疾病，需要服用和注射一堆药物。她以前还有吸烟史，烟龄很长，几年前刚把烟戒掉，不过戒得还算彻底。这位老妇能毫不费力地爬两层楼梯，身体状况看起来还不错。我在用听诊器为她进行检查的时候没有发现她的肺部有杂音，病历记录也没有显示她有肺病。但她在手术前告诉麻醉医生说，自己以前曾做过两次手术，术后发生了呼吸困难的情况，每次都需要在家吸氧好几周才能恢复，其中有一次她甚至要求接受重症监护。

这是一个很严重的问题。麻醉医生了解这一情况，但我不了解。在手术前执行清

单的时候，麻醉医生问我为什么不让病人留院观察，他说病人有呼吸问题。

“什么呼吸问题？”我问道。这时我才从麻醉医生的口中了解到全部情况。于是，我们为这位病人安排了留院观察。而且，我们还决定让她在术中和术后使用吸入器以防止呼吸问题。结果，这位病人不需要额外吸氧就平安出院了。

这个世界需要一场清单革命

无论一台手术多么平常，每个病人都是独特的。在清单的帮助下，我们发现了很多与手术安全相关的潜在问题，其中包括病人药物过敏、设备故障、用药不当、活检标本标记错误等。比如“不，这个是右侧的，这个才是左侧的”。这些都要求我们制订的清单更加明确细致，手术的准备工作的更加充分。如果没有清单，我们不知道会出现多少可能给病人造成巨大伤害的差错。有人可能会说，我们平时的警觉程度和注意力集中程度足以让我们发现严重问题，没有被我们发现的问题不会给病人造成太大伤害。

但有一次，清单的确挽救了我收治的一个病人的生命。



这位病人是哈格曼先生，他53岁，是两个孩子的父亲，还是当地一家公司的老总。他的右侧肾上腺长了一个罕见的嗜铬细胞瘤，所以我要通过手术将其右侧的整个肾上腺摘除。这种罕见的肿瘤会让肾上腺素分泌过量，而且很难摘除。虽然此病非常罕见，但我除了负责普外科手术，还对内分泌腺手术具有浓厚的兴趣，并掌握了不少手术技术，积累了不少实践经验。我已经成功实施了近40台肾上腺肿瘤手术，而且无一例发生并发症。所以，哈格曼先生前来就医的时候，我对为他成功实施手术非常有信心。不过，我向他解释说，发生严重并发症的风险总是存在的。这台手术的危险主要在于把肾上腺从下腔静脉剥离这一操作上。下腔静脉是让血液回流心脏的主静脉，在进行这一操作时，如果下腔静脉破裂，就会造成大出血，危及病人的生命。但我向他保证，发生这种情况的可能性微乎其微。

不过一旦进了手术室，谈论概率的大小是没有意义的，危险情况要么发生，要么不发生。不幸的是，这次我碰到了大麻烦。

我用腹腔镜实施手术，插入哈格曼体内的光纤摄像头将他体内的影像传到屏幕上。我一边看着屏幕上的图像，一边控制随着腹腔镜进入他体内的手术工具切除肿瘤。手术进行得非常顺利，我掀起他的肝脏，有一团软软的棕黄色腺体出现在我眼前，这团东西看起来就像是熟鸡蛋黄，我开始动手将其剥离下腔静脉。虽然这一操作要求医生非常细心，但其难度并不是特别高。就在我快要将腺体整体剥离的时候，一件从未发生过的事情不幸发生了：我把病人的下腔静脉弄破了！

这是一场巨大的灾难！我造成的伤害好比在哈格曼先生的心脏上直接扎了个洞，大量鲜血从裂口处喷涌而出。在短短60秒钟内，他体内所有的血液差不多全都流到了腹腔里。他的心脏停止了跳动，我以最快的速度打开他的胸腔和腹腔，并用手握住他的心脏，每隔3秒挤压一次，以维持他的脑部供血。我身旁的住院医师则用力压住病人的下腔静脉以减缓出血的速度。但我明显感到，哈格曼先生的心脏变得越来越空。我心想这下完了，哈格曼先生没法活着走出手术室了，我害死了他。

但不幸中的万幸是，我们在手术前执行了清单程序。在估计病人的失血量时，我说失血量应该不会太大，在这类手术中，我从来没有让病人失血超过100毫升。我很自信，很期待成功完成这台手术。但我又说，由于肿瘤紧紧贴着下腔静脉，所以至少在理论上大出血是有可能发生的。听到这番话后，护士马上通知血库准备好4个单位浓缩红细胞以防万一。想不到这一次它们真的派上了用场。正是因为清单让血库做好了充分准备，仅仅这一个检查项目就挽救了这个病人的生命。

不仅如此，执行清单这一过程也起到了很大作用。手术室里除了我以外，还有麻醉医生、麻醉护士、外科住院医师、手术助理护士、巡回护士和医学院学生。此前，我只和他们中的两个人有过合作，而且只和外科住院医师比较熟识。但在手术开始前，我们都进行了自我介绍：“我是主刀医生阿图·葛文德；我是外科住院医师里奇·布拉福德（Rich Brafford）；我是护士苏·特沃迪（Sue Twardy）……”你可以感觉到，一个真正的团队正在形成。我们对病人腕带上的姓名进行了核实，确认了哪一侧的肾上腺将被摘除；麻醉医生和护士确认了没有需要

提醒的注意事项；大家共同确认病人已经注射了预防性抗生素，并给他盖上了加热毯，让他穿上了防止静脉血栓形成的充气靴。进入手术室的时候我们还是陌生人，但在病人的皮肤被切开之前，我们已经成了齐心协力、共同应对挑战的团队。

所以，当我失手酿成大祸的时候，每个团队成员都保持了镇定。巡回护士发出警报让其他医护人员赶到手术室，她还立刻从血库里取出准备好的血袋，麻醉医生则开始不停地给病人输血补液。大家有条不紊地开展各项工作，有的负责搬来我需要的设备，有的则呼叫能够帮上忙的血管外科医生，有的协助麻醉医生输血，有的则将进展通知血库以便做好进一步备血准备。整个团队为我和病人赢得了宝贵的时间。

最终，我们给病人输了30个单位的血，他的失血量是他原来体内所含血量的3倍。我一边盯着监视器上显示的血压读数，一边用手挤压他的心脏，他的血液循环终于没有因为失血过多而停止。闻讯赶来的血管外科医生和我花了很长时间才修补好病人下腔静脉上的裂口，我感到病人的心脏恢复了自主跳动。最后，我们将他的伤口缝好，哈格曼先生挺了过来。

我不想骗大家说哈格曼安然无恙地离开了手术室。持续的低血压破坏了他的一条视神经，他有一只眼睛几乎完全失明。在术后的几天里，他一直靠呼吸机维持呼吸。在此后的数月里，他都无法工作。这次失误让我遭受了沉重打击，虽然我向他道了歉，并重新回到了工作岗位，但好长时间我都没缓过这口气。此后，每当实施肾上腺切除术的时候，我都会想起这次经历。从另一个角度看，这对我而言也许是件好事，它迫使我不断尝试改进技术，希望能够想出办法更好地保护下腔静脉，防止此类情况再次发生。

正是因为这次遭遇，我对清单感激不尽。如果那天我们没有执行清单程序，事情又会变成什么样子呢？这个问题的答案我连想都不敢想！我不愿意想象自己走出手术室将噩耗告诉病人家属的场景。

不久前，我和哈格曼先生通了电话。他成功地卖掉了自己的公司，并且正在努力让另一家公司扭亏为盈。他每周抽时间跑步锻炼3次，甚至能够自己开车了。

他说：“我必须小心我的盲点，但我能应付。”

我在他的脸上看不到一丝痛苦和愤怒，这是常人难以做到的。他一直说：“我能活下来就已经很幸运了。”我问他是否允许我把他的故事写进这本书。

他说：“没问题，如果你真的那么做，我会很高兴。”

现在，已经有无数的事实向我们证明：

这个世界需要一场清单革命！

译者后记

我是一个不折不扣的飞行迷。虽然买不起私人飞机，也买不起波音公司的专业模拟器，但至少买得起高配置电脑和《微软飞行模拟》（Microsoft Flight Simulator）这套全球发行量最大的家用飞行模拟软件。你可别小瞧了它，美国的很多飞行学校就用它进行教学，很多专业飞行员在家里的时候也会用它来熟悉和复习操作程序。

我在模拟飞行世界里算得上是一名经验丰富的老飞行员了。所以，当接到这本书的翻译任务时，我的内心既兴奋又纳闷。兴奋的是终于能够翻译一本与飞行有关的畅销书了，纳闷的是一个外科医生怎么会和飞行清单扯上关系呢？

通读全书我才知道，原来作者阿图·葛文德想借用这种简单而实用的方法来应付超级复杂的问题和情况。有趣的是，书中提到的299型轰炸机试飞员因为没有给操控面解锁而造成飞机坠毁的情况，我在模拟程序里也发生过，而且不止一次。我的“不良飞行记录”还包括因操作不当导致发动机起火，在降落的时候忘记放下起落架，意外地进行了“完美的”机腹迫降。幸好我驾驶的不是真飞机！

之所以会发生这一系列愚蠢的错误，真的是因为现在的飞机太复杂了，而且我痛恨使用清单。要知道，它们真的有几百页厚。一些公司，如PMDG和Level-D利用这一软件制作非常专业的飞机插件，如MD-11、波音737、波音767等。每一型模拟飞机的系统介绍和操作手册都有600～700页，光是各类清单就有200～300页，看得我头都大了。说实话，我没钱也没工夫把它们都打印出来。即便是打印出来的那几张“常用清单”，在“飞行”的时候我也懒得看，因为我觉得自己很熟悉各种操作，不会出事。但事实证明，离开了它们，我已经在模拟飞行世界里闯下很多大祸了。

我知道清单很有用。如果我真的是飞行员，一定不会抛弃它，但我不知道它竟然还能在那么多行业里发挥作用。我觉得推动清单使用的无非是两种力量。

第一种是外在压力。比如，虽然航空事故的总体发生率和伤亡率很低，但单起事故造成的人员和财物损失很大，而且事故不论大小，哪怕是单引擎小飞机坠毁，全球各家媒体都会竞相报道，产生的负面影响极大，容不得飞行员犯错。

第二种就是内在驱动力。即作者提到的职业精神。他曾因为自己不慎弄破病人的下腔静脉造成大出血而向病人诚恳道歉。说实话，这起事故在手术的正常风险范围内，换个医德不高的医生，说不定还会向病人和家属邀功，吹嘘自己挽救了病人的生命。如果作者少了这份职业道德，我很难想象他会费尽心思、不辞辛劳地搞什么新发明。他已经是一名非常优秀的外科医生了。所以，我对他的这份真诚和责任心感到由衷的钦佩。总之，这是一本充满了智慧和让人感动的好书，不容错过。

最后，我还要感谢在本书翻译过程中给予我大力支持的朋友。陈晋、王军和王佩幸不仅承担了部分章节的翻译工作，还让我以非常愉悦的心情进行翻译。此外，尹洪波和贺皖琍也在资料查阅和收集过程中给了我很大帮助，在此一并表示衷心的感谢！

王佳艺

未来, 属于终身学习者

我这辈子遇到的聪明人(来自各行各业的聪明人)没有不每天阅读的——没有, 一个都没有。巴菲特读书之多, 我读书之多, 可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式; 指数型技术在迅速颠覆着现有的商业世界; 人工智能已经开始抢占人类的工作岗位……

未来, 到底需要什么样的人才?

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才, 而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力, 获得异于众人的思考和整合能力。未来, 将属于终身学习者! 而阅读必定和终身学习形影不离。

很多人读书, 追求的是干货, 寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代, 答案不会简单地出现在书里, 因为生活根本就没有标准确切的答案, 你也不能期望过去的经验能解决未来的问题。

而真正的阅读, 应该在书中与智者同行思考, 借他们的视角看到世界的多元性, 提出比答案更重要的好问题, 在不确定的时代中领先起跑。

湛庐阅读 App : 与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上, 把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本

怎么读是读者面临的最大的阅读障碍

“读书破万卷”不仅仅在“万”, 更重要的是在“破”!

现在, 我们构建了全新的“湛庐阅读”App。它将成为你“破万卷”的新居所。在这里:

- 不用考虑读什么, 你可以便捷找到纸书、电子书、有声书和各种声音产品;
- 你可以学会怎么读, 你将发现集泛读、通读、精读于一体的阅读解决方案;
- 你会与作者、译者、专家、推荐人和阅读教练相遇, 他们是优质思想的发源地;
- 你会与优秀的读者和终身学习者为伍, 他们对阅读和学习有着持久的热情和源源不绝的内驱力。



下载湛庐阅读 App,
坚持亲自阅读,
有声书、电子书、阅读服务,
一站获得。

CHEERS

本书阅读资料包

给你便捷、高效、全面的阅读体验

本书参考资料

湛庐独家策划

- ✓ 参考文献
为了环保、节约纸张，部分图书的参考文献以电子版方式提供
- ✓ 主题书单
编辑精心推荐的延伸阅读书单，助你开启主题式阅读
- ✓ 图片资料
提供部分图片的高清彩色原版大图，方便保存和分享

相关阅读服务

终身学习者必备

- ✓ 电子书
便捷、高效，方便检索，易于携带，随时更新
- ✓ 有声书
保护视力，随时随地，有温度、有情感地听本书
- ✓ 精读班
2~4周，最懂这本书的人带你读完、读懂、读透这本好书
- ✓ 课 程
课程权威专家给你开书单，带你快速浏览一个领域的知识概貌
- ✓ 讲 书
30分钟，大咖给你讲本书，让你挑书不费劲



湛庐编辑为你独家呈现
助你更好获得书里和书外的思想和智慧，请扫码查收！

（阅读资料包的内容因书而异，最终以湛庐阅读App页面为准）

(1)财务杠杆风险是指公司运用负债融资及优先股筹资后，公司投资者和普通股股东所需负担的额外风险。——编者注

(2)英雄机长萨伦伯格亲笔讲述这次传奇迫降的著作《最高职责》简体中文版，已由湛庐策划，并由北京联合出版公司于2016年出版。——编者注

CHEERS
湛庐

THE
INFINITE
GAME
无限
的
游戏
THE
INFINITE
GAME
THE
INFINITE
GAME
THE
INFINITE
GAME
THE
INFINITE
GAME
THE

[英]西蒙·斯涅克
Simon Sinek 著
石雨晴 译

天津出版传媒集团
天津科学技术出版社

版权信息

本书纸版由天津科学技术出版社于2020年8月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

书名：无限的游戏

著者：西蒙·斯涅克

电子书定价：62.99元

The Infinite Game by Simon Sinek

Copyright © by 2019 by SinekPartners, LLC

All rights reserved including the right of reproduction in whole or in part in any form.

This edition published by arrangement with the Portfolio, an imprint of Penguin Publishing Group, a division of Penguin Random House LLC

亲爱的祖母：

您的一生，活得宛如人生永无终点。

希望我们都能学会这般无限地活着。

爱你的，西蒙

Dear Grandma,

Because you lived as if there was no finish line.

May we all learn to live such an infinite life.

Love, Simon

岔路口立着一块路牌，
一个方向名为“胜利”，
另一个方向名为“满足”。
如果必须选择一个方向，
我们要选择哪一个？

如果选择胜利，
目标就是赢！
当我们冲向终点时，
将会体验到竞争的刺激。
众人聚首，为我们欢呼！
然后一切结束，
曲终人散。
（或许我们希望再来一次）

如果选择满足，
那么旅途漫漫。
有时必须谨慎地踏出每一步，
有时可以驻足欣赏沿途风光，
我们永不停步，
我们一往无前。
众人聚首，与我们同行。

当我们生命终结时，
在满足之路上的同行者，
没有我们，也将继续前行，
还将激励更多的人与之同行。

赞誉

张瑞敏

海尔集团董事局主席、首席执行官

产品会被场景替代，行业将被生态覆盖。在无价、无界、无序的时代，让我们与斯涅克一起融入以“人的价值最大化”为宗旨的无限游戏。

彭剑锋

中国人民大学教授、博士生导师

在数字化与智能化时代，对于企业而言，最大的敌人不是竞争对手，而是企业自己，是成功的思维习惯，是面对环境变化思维的局限。《无限的游戏》这本书，将为中国企业家突破思维局限，走出成功陷阱，实现创新成长提供无限思维。

肖知兴

领教工坊学术委员会主席

西蒙·斯涅克的“黄金圈思维”揭示了人们对意义、目的与价值观的渴求。《无限的游戏》更进一步地强调了组织的精神性和终极性。一个能提供清晰的使命感的组织，自然能够激发组织成员内心最深层次的力量。未来的领导者，更需要这样的无限思维。

徐中

领导力学者，智学明德国际领导力中心创始人

西蒙·斯涅克看问题总是格局宏大、目光长远、刨根问底、一针见血，处处展现“第一性原理”思维，不时让人醍醐灌顶。坦率来说，无限游戏并不是一个新概念，对未来领导者来说却是最重要的概念之一。未来卓越的领导者必须走出“短期的、竞争的、输赢的”有限思维，迈向“长远的、竞合的、共赢的”无限思维，为了组织不忘初心的基业长青，为了人类生生不息的世界大同，着眼一起做大蛋糕，而不是只想着如何瓜分眼前的蛋糕！

秦朔

秦朔朋友圈创始人

所有人都深感世界正变得越来越不确定。而所有“不确定”其实都是对有限思维、有限人格的挑战。当你具有超越性的无限思维，并不断激发出新的自驱力，就能从容地驾驭“不确定”。《无限的游戏》提示人们，一项事业的崇高才是无限思维的源泉。

张建国

人瑞人才科技集团公司董事长、行政总裁

每个人都有他自己的局限性，但自己很难感受到，因为他本身就在局限之内，思想的局限性尤其如此。但我们现在所处的世界变化万千，在商业社会中，打败你的往往不是你习惯认为的竞争对手，而是新的行业外的玩家。《无限的游戏》这本书，引导企业经营者从有限思维上升到无限思维，也许能为企业寻找到根本性的发展动力。这对我也是一种新的启发。相信这本书对很多实际的企业经营者都有突破性的思想贡献。

推荐序

教育者亟待觉悟“无限”

清华经管领导力研究中心主任 杨斌

承担教书育人重任的教育者，在与学生的互动中传递着许多关键的价值观因素。校长也好，教师也罢，越是被学生爱戴、喜欢、信任，就越是具有价值观影响力，正所谓“亲其师，信其道；尊其师，奉其教；敬其师，效其行”。年纪小的学生，被海绵型的学习观主导，尚不懂得应该以淘金式的方式去批判性、建构性地学习，这一阶段所播撒下的价值观种子，更是有着持久的威力。

也正是从这个角度，应该格外重视教育者的思维模式，这是一种潜藏起来的价值观。许多时候，教育者自身都没有意识到这种思维模式对自己的影响，更不要说通过语默作止产生了对学习者的迁移。这里的教育者，当然也包括家庭教育中的父母亲友，这部分的影响融于日常生活中，并潜移默化到下一代的骨髓中。

尽管叫思维，但也有其道德性，只是表现形式上容易感觉是方法论层面的东西。它会影响到是非对错的认识，人生该怎么度过，人与人、组织与组织之间到底是什么关系。

比如，到底是看重团队还是个体？不是看你怎么说、口头上提倡什么，而是看你真正怎么做、评价什么。独善其身的文化基因是被教育过程中的很多指挥棒所诱发、强化、固化并影响终身的。从小到大，对于团队的奖励总是太少，放眼去都是对个体成绩的认可。科研论文的合作者文化不彰，得从中小学阶段就对团队协作鼓励严重不足中找原因。要允许、包容学生中有一定程度的搭便车情形，而不要单单为了对个体准

确地衡量、有效地比较而舍弃孩子在团队中体会不分彼此的价值、体验融为一体的力量。

再比如，由心理学家卡罗尔·德韦克（Carol S. Dweck）界定而引发广泛讨论的成长型思维（growth mindset）模式的精髓是认为“人的能力、智力等是变化的，可以拓展的”，看起来这该是教育行当自身价值的证明，但成长型思维面向的却是成年人、终身学习、人力资源开发这个群体。试想一下，孩子们是如何经由固定型思维的教育者而培养为成长型思维的成年人的，也许就更该把注意力放在教育者的成长型思维——不是当教育理论来学，而是对照改造自我。不仅关注教育者如何看待学生的潜力，更要看待已是成年人的教育者自身的发展可能，毕竟学生和教育者之间最基础的是社会性学习，也就是“有样学样”。

同样，自“无限的游戏”这个概念体系而来的“无限思维”，是另一种重要的思维模式。詹姆斯·卡斯（James P. Carse）教授1986年第一次从哲学上区分了世界上的这两种游戏：有限游戏以取胜为目的，而无限游戏以延续为目的。无限游戏在我们的生活中大量存在，没有终点，也没有赢家，期待着传奇，改写着规则和剧本，升华着人类的进程。但充斥在我们周围的，是以有限游戏的思维，在无限游戏中追求着“只不过是瞬间”的赢，却消耗着继续参与游戏所必需的意志与资源，带来信任、合作与创新的衰减。

商业领域中，有限思维看到的是阶段性的财务数据、财富目标，无限思维关注的却是“崇高的事业”，我把它称为“正因志业”。这让世界进步且美好，让人更接近且展现了人的本质，虽不能至，然心向往之。

有限思维享受我赢而对手失利，无限思维却享受棋逢对手而带动彼此的活力；有限思维以控制来带来下属的亦步亦趋，无限思维以信任解放团队的主人翁感、主动积极性和创造力。

有限思维惧怕颠覆希望在稳定性中得到边界内的利益，无限思维拥抱变革在超越中进入突破边界而恍然的新天地。有限思维的焦点是获得

稀缺的权力（power），无限思维则聚焦于开放而非封闭的力量（strength）。有限思维与可能性为敌，消灭可能性带来的纷乱；无限思维与可能性为伍，激发可能性而释放红利。

有限思维在意的是领导人的自我安全与权力维系，希望代表着自己过去成绩的头衔能延续到未来；无限思维的领导者无我而富有勇气，听从内心深处正因志业的召唤，以朴素并具有无尽可能的名字开放了所有能让彼此关系变得深刻而互惠的可能性。有限思维中，你的未来和我的未来有着零和的剧本结局，无限思维却让你的未来和我的未来成为我们的未来，我们彼此为层出不穷的惊奇做好了准备。

还可以一直这样对比下去。团队思维、成长思维，一定意义上，都是无限思维以某种角度的呈现。简而言之，无限思维的益处，显而易见，无处不在。但与此相对应的，我们现行的教育体系的许多方面，却沉浸在有限游戏的设置中，不知不觉地、“照常经营地”养成着下一代人的有限思维。

在一个学校中，我们最经常看到的，就是以赢家的头衔来结束这场游戏。成为赢家，就要为自己而战，打败其他所有玩家。制订的每一个计划、采取的每一个行动都是为了赢，在（有限）游戏中的人坚信这是唯一正确的选择。其实，引导他们这么做的，是他们自己的越来越根深蒂固且自我持续强化的思维模式。考试的目的是为了让学生了解自身的进步与下一步的方向，还是为了在一群学生中分出高低胜败，甚至赋予优胜者特定的头衔以进入更高级的（但也仍然是有限的）游戏？讨论问题是为了寻求更多的视角与宽阔、深入的理解，还是为了比赛时谁先击中标准答案并最终让所有人的意见都收敛到标准答案？已经司空见惯的奖学金安排，究竟是奖励过去的表现，还是鼓励未来的探索？是奖励与他人比较的赢家，还是鼓励跳出舒适区的新我？是奖励按照标准剧本完美无瑕的演出，还是鼓励书写新传奇的不完美的探索？

校长们、老师们一旦觉得“我们已经知道”“一切尽在掌握”“只需按

方培养”，有限游戏的大幕就开启了，而敢于承认“我们还不知道”“一切皆有可能”“成长千方各样”，无限思维才在敬畏和拥抱不确定性中登场。某种意义上，精神、创造、改变、解放意义上的“年轻”并不是按年龄来区分的，而应该是“无限思维”的代名词，这样的话，我们绝不能让校园中充满着年龄虽小却已不年轻的有限人。

在人的发展和人类进步的这个无限游戏中，有限思维的模式，让人们过度执着于有限的目标，在某段时间内，根据某个游戏规则（衡量评价标准），他们也许能够达成“第一”或“最好”的目标。但这一点儿也不意味着，这将有利于他们在一个无限游戏中不被淘汰，因为规则在不断变化，目标在持续升级，边界被跨越，范式被颠覆。人的创新，是典型的无限思维所带来的无限游戏，就像丹尼尔·平克（Daniel Pink）在《驱动力》（*Drive*）[\[4\]](#)中所揭示的，有限思维中的赢家面对开放、创新和合作的“竞赛”（无限游戏）时，训练有素所带来的优势都成了束缚羁绊的劣势。唯有丢掉有限带来的枷锁，才能拥抱无限孕育着的传奇。

美国篮球传奇人物约翰·伍登教练，其影响远超篮球或者体育界，是一位“无限思维”的信仰者、实践者与布道者。作为主教练的他拿过10个全美大学联赛冠军，却不把篮球或者体育运动当作“有限游戏”来玩。“在我的执教生涯里，我很少或几乎没有说过‘赢’这个字，也很少论及‘打败’对手，以及激励某个队伍争夺第一名。”他曾经这样说，“我对成功的定义是这样的：竭尽全力，达己最佳而感到自足，由此得至内心平静，谓之成功。”这是“伍登教导力”的核心理念，让队员们从记分牌的焦虑中解脱出来、沉静下来、努力起来。

有限与无限的游戏所带来的哲学思考，有限与无限的思维所产生的实践分野，对所有人来说，都是重要的启蒙，而对于教育者来说，则是迫在眉睫的警示。因为教育者（教育管理者）很多时候是游戏模式的设计者、选择者，是有限思维的助长者、熏陶者（而不自觉），影响着下一代人、下一代组织、未来的社会和人类的明天。

西蒙·斯涅克、詹姆斯·卡斯提供的这副看世界的透镜，有着大道至简的真、上善若水的善、生生不息的美。对一个人而言，生有涯，对于所投身的正因志业而言，则是无涯。把有限的生命投入到无限的正因志业中，无我而大我，这便是觉悟。

教育者，须先觉悟“无限”，才孕育新世界。



扫码下载“湛庐阅读”App
搜索“无限的游戏”
获取更多精彩内容

序

我为什么要写这本书

出人意料的是，这本书居然有存在的必要。纵观人类历史，无限思维的益处比比皆是，一目了然。伟大社会的兴起、科学与医学的进步，以及对太空的探索，都源自一大群为了共同的事业而团结在一起的人，对他们来说，纵使前景晦暗不明，仍要携手共进。这就像一艘冲向星空的火箭坠毁了，我们必须找出问题所在，然后一遍又一遍地尝试。就算成功了，我们也还是会继续前进。驱使我们这样做的，并不是年终奖励，而是我们感觉自己正在为比我们自身更伟大的事业做贡献，这项事业的价值存续远比我们自己的一生要长。

尽管有诸多益处，但以长远、无限的眼光行事并不容易，需要真正的努力。生而为人，面对令人不适的问题，我们天生倾向于寻找立竿见影的解决方案；为了实现自己的雄心壮志，我们倾向于优先考虑速赢之法。我们往往从成功与失败、胜者与败者的角度看待这个世界。这种默认的输赢模式有时在短期内是有效的，不过，若要作为公司和组织的运营策略，从更长远来看，这样做会产生严重的后果。

这种默认的思维模式所产生的结果我们都很熟悉：为达到“武断”的预期目标而进行的年度大裁员；充斥着残酷竞争的工作环境；将股东的需要置于员工及客户的需要之上；不诚实、不道德的商业行为；在团队内，奖励只看绩效高低，而对高绩效者可能给其他成员造成的损害视而不见；奖励流向那些在乎自己远超在乎下属的领导者。所有这些都会降低员工的忠诚度与其工作的积极性，并增加人们的不安全感与焦虑感，而这正是我们中太多人当下的处境。这种纯交易性的商业做法在工业革命之后得到了迅猛的发展，在当今的数字时代，依然在加速发展。我们

对商业和资本主义的全部理解似乎都受制于短视的有限思维。

尽管我们中的许多人为这种现状感到遗憾，但不幸的是，市场想要保持现状的愿望似乎比想要改变现状的动力更大。当我们说出“人不能唯利是图”这样的话时，阻力往往就会随之而来。当前商业体系的许多掌控者和领导者，都认为我们太天真，不了解商业运行的现实，有太多人因此退却。我们屈从于现实，每天在对工作的担忧中醒来，无法安心工作，在生活中挣扎着寻找满足感。这种工作与生活失衡的问题已经严重影响了整个社会。对此，我很好奇，我们有没有另一种可行的选择？

其实完全有可能，现实或许，只是或许，并非真如愤世嫉俗者所一直喋喋不休的那样。或许是现在的商业体系不正确，或不够好。我们习以为常的或许只是少数人所偏爱与推进的，并不能够代表大多数人。若事实当真如此，我们便有机会塑造另一种不同的现实。

我们完全有能力创造出一个这样的世界：大多数人都会热情中醒来，能安心地在外工作，能带着工作的满足感踏上回家的路。我所提倡的这种改变并非易事，却是有可能实现的。有了优秀的领导者、伟大的领导者，这一愿景就能实现。所谓伟大的领导者，就是那些能够跳出“短期”与“长期”框架的人，他们知道所要思考的并不是下一季度，也不是下一届选举，而应该关乎下一代人。他们所设定的目标应该是让自己的组织获得比自己的生命更长久的成功，当他们真的这样去做时，对我们、对企业，甚至对股东来说，都会带来非同寻常的益处。

我写作本书的目的，不是改变那些捍卫现状的人，而是召集那些已经准备好挑战现状、用新现实取代旧现实的人。这一新现实更有助于满足我们内心的需求——获得安全感，为比自身更伟大的事业做贡献，以及为自己和家人创造更好的生活条件。这样的现实才符合我们作为个体、企业、社会和一个物种的最大利益。

如果我们相信有这样的世界存在——能让自己每天都感觉振奋、安心与满足，如果我们相信领导者就该是可以实现这一愿景的人，那么，

寻找并支持这样的领导者就成为我们共同的责任，而这样的领导者也将致力于用更有可能实现这一愿景的方式来领导我们。为此，我们所需采取的行动之一就是，了解成为无限游戏的领导者到底意味着什么。

2019年2月4日

于英国伦敦

前言 赢

某一年的新年，A国对B国军队及其盟军发动了突袭。在接下来的24小时内，有8.5万多A国人共计袭击了遍布A国的125个目标。B国军队毫无准备，被打了个措手不及。突袭开始时，许多指挥官甚至都不在岗——他们正在附近的城市庆祝新年。

新年对A国人来说十分重要，因此，A国有一个传统，就是新年当天绝不开战。然而，A国领导人认为这是一次击溃B国军队并有望迅速结束这场战争的机会，因此决定要打破传统，发动突袭。

出人意料的是，B国军队竟击退了A国所有的袭击，击溃且重创了发起猛攻的A国。在突袭开始的大概一周后，大部分主要战役结束了，B国军队阵亡人数不足1000人。与之形成鲜明对比的是，A国士兵阵亡超过3.5万！虽然B国赢了大多数战役，最终却输掉了整场战争。

所有这些加在一起就引出了一个问题：几乎每场战役的赢家都是你，那你究竟是怎么输掉整场战争的？

目 录

[推荐序 教育者亟待觉悟“无限”](#)

[序 我为什么要写这本书](#)

[前言 赢](#)

[第1章 有限游戏与无限游戏](#)

[要素1 追求崇高的事业](#)

[第2章 崇高的事业](#)

[第3章 此事业非彼事业](#)

[第4章 事业的守护者](#)

[第5章 企业责任](#)

[第6章 意志和资源](#)

[要素2 建立彼此信任的团队](#)

[第7章 彼此信任的团队](#)

[第8章 道德的堕落](#)

[要素3 尊重值得较量的对手](#)

[第9章 值得较量的对手](#)

[要素4 拥有攸关存亡的变通性](#)

[第10章 攸关存亡的变通性](#)

[要素5 表现领导的勇气](#)

[第11章 领导的勇气](#)

[后记 用无限思维玩无限游戏](#)

[致谢](#)

[注释](#)

第1章

有限游戏与无限游戏

只要有至少两个玩家存在，游戏即可成立。但游戏分两种：有限游戏与无限游戏。

有限游戏的玩家都是已知的。他们有固定的规则，有一致认同的目标，一旦目标达成，游戏就结束了。足球就是这样的有限游戏，球员都身着球衣，很容易辨认；球赛有一套既定的规则，由在场的裁判负责监督执行，所有球员都同意并遵守这些规则，一旦犯规，也愿意接受处罚；所有人也都同意，在规定时间内得分高的一方为获胜方，然后游戏结束。在有限游戏中，总会有开端、中点和结束。

相比之下，无限游戏则既有已知玩家，也有未知玩家，且没有确切的或一致认同的规则。尽管可能存在一些惯例或法律，对玩家的言行有一定的约束力，但这些也只圈定了一个非常广阔的范围，在该范围内，玩家想怎么玩都可以，打破惯例，有时也是可以的。每个玩家的游戏方式完全由自己决定，可以随时因任何原因而改变。

无限游戏没有时限，也没有终点。既然没有明确的游戏终结点，自然就不存在所谓的“赢”。在无限游戏中，玩家的首要目标就是一直玩下去，也就是让游戏一直进行下去。

我对这两种游戏的理解源自哲学家詹姆斯·卡斯教授于1986年出版的著作《有限与无限的游戏：一个哲学家眼中的竞技世界》（*Finite and Infinite Games: A Vision of Life as Play and Possibility*）。正是卡斯的这本书让我第一次摆脱了输赢、平局和僵局的传统思维。用卡斯有限与无限游戏的视角来观察我们的世界，看得越多，就越能发现我们周遭存在着

的无限游戏——既没有终点也没有赢家的游戏。比如，在婚姻或友谊中就没有所谓的赢家；学校可能是有限的，但在教育中并不存在所谓的赢家；在找工作或升职的竞争中，我们可以打败其他人，但从来没有人能够得到“职业生涯第一”的桂冠；国家之间可以在世界范围内竞争，比拼影响力或经济实力，但就全球政治而言，并不存在所谓的赢家。无论我们一生取得了多少成就，在我们死后，都没有人能被冠以“人生赢家”的美誉，商业赢家当然也是不存在的。所有这些过程，而非事件。

然而，现在许多领导者仿佛全然不知自己在玩的是什么游戏。他们不停谈论着“赢”，沉迷于“赢得竞争”，向世界宣称自己是“最棒的”，说自己的愿景就是“成为第一”。遗憾的是，这些目标在没有终点的游戏中都是不可能实现的。

若用有限的思维去领导无限游戏，各种问题就会接踵而至，最常见的是信任、合作与创新的减少；相反，若用无限的思维去领导无限游戏，则能让我们朝着更好的方向前进。具备无限思维的团体所拥有的信任、合作与创新水平都要远远高于有限思维的团体，自然也能享受到随之而来的各种益处。若我们所有人都势必要成为各种无限游戏的玩家，那么学会识别游戏，学会无限思维的领导方式，就是关乎我们切身利益的大事了。与之同等重要的还有，学会识别有限思维，在其造成真正损害之前及时调整。

商业的无限游戏

商业游戏完全符合无限游戏的定义。我们不一定知道除了自己还有哪些玩家以及新的玩家会在何时加入。除法律（法律也可能因国而异）外，无限游戏中并不存在一套固定的、所有人都认同的规则，所有玩家都能自行制定策略与战术。商业游戏也没有既定的开端、中点和结束。尽管许多人同意在一定的时间范围内评估自己相较于其他玩家的表现，比如财年，但这些时间范围所标志的只是游戏过程中的节点，没有一个

时间能标志游戏的结束。商业游戏是没有终点的。

尽管所有企业都在玩一场没有输赢的游戏，但仍有太多的商业领袖以赢为目的。他们总是号称自己是“最棒的”或“第一名”，这样的说法我们已经司空见惯，以至于很少，甚至不曾停下来思考其中的一些说法有多荒谬。每当看到一家企业声称自己是第一名或最好的公司时，我总喜欢研究一下为了支撑这一说法，他们是如何精心挑选衡量标准的。例如，多年来，英国航空公司（British Airways）一直在广告中声称自己是“世界上最受欢迎的航空公司”。¹理查德·布兰森（Richard Branson）创办的维珍航空公司（Virgin Atlantic Airways）向英国广告标准局（Advertising Standards Authority，简称ASA）投诉称，乘客调查显示，英国航空公司的说法并不成立。然而，英国广告标准局基于“英国航空公司的国际客运量是所有航空公司中最大的”这一理由，批准其继续使用这一说法。因此，在英国航空公司的广告语中，“最受欢迎”的意思只是其业务量很大，并非一定是乘客的首选。

对一家自称第一的公司来说，其依据可能是所服务的顾客数量；而对另一家自称第一的公司来说，其依据可能是收入、股票表现、员工数量或在全球各地开设的分部数量。这些公司为了支撑自己的说法，甚至会自行选择不同的时间段，有时是一个季度或8个月，有时是3年、5年或12年。但同行是否认同这样的时间段呢？在有限游戏中，判定输赢的衡量标准只有一个，且所有玩家都认同，比如进球数、速度或力量。但在无限游戏中，衡量标准有很多，我们永远无法判定谁是赢家。

在有限游戏中，时间一到，游戏就结束了，直到另一时间游戏再开始。而在无限游戏中，情况正好相反，不断继续的是游戏本身，时间会耗尽的是玩家。无限游戏没有胜负之分，当玩家耗尽了参与游戏所需的意志与资源后，就会自行退出。在商业活动中，我们称之为破产，有时也为合并或收购。这意味着，要想在商业这场无限游戏中取得成功，我们必须停止思考赢，转而思考如何建立一个足够强大、足够健康的组

织，让它能够经久不衰，不被游戏淘汰。有趣的是，这一长远思维往往也能帮助企业在短期内变得更加强大。

两个不同的玩家

多年前，我分别在微软和苹果举办的教育峰会上做演讲，前后相隔几个月。在微软的峰会上，大多数演讲者都花了大量时间谈论如何打败苹果。而在苹果的峰会上，几乎所有演讲者都在用几乎全部时间谈论苹果公司该如何帮助教师教学以及学生学习。似乎，一方只沉迷于打败竞争对手，另一方则一心只想推进教育事业。

微软峰会的演讲结束后，他们送了我一件礼物——新款音乐播放器Zune，这在当时是很厉害的一款产品。这是微软对苹果音乐播放器iPod的回应，iPod是当时MP3播放器市场的主导者。微软不甘示弱，推出了音乐播放器Zune，试图争夺这一主要竞争对手的市场份额。当时是2006年，时任微软首席执行官的史蒂夫·鲍尔默（Steve Ballmer）对微软能够最终“击败”苹果充满信心，尽管他也知道这并非易事。²如果产品质量是唯一的决胜因素，那么鲍尔默的乐观就是正确的。不得不说，微软送给我的这款产品非常出色，它设计巧妙，用户界面简单、直观、好操作，我非常喜欢。但出于诚实，我得承认，我将Zune转送给了朋友，因为我的iPod可以兼容微软的Windows系统，而Zune却无法兼容苹果的iTunes。

在苹果峰会的演讲结束后，我与苹果高管约翰·库奇（John D. Couch）²同乘一辆出租车返回酒店，准确地说，他是苹果的第54号员工，这意味着他在苹果创立之初就加入了，苹果的文化与信念已经完全深入他心。身边有这么一位“跑不了”的听众，我实在忍不住，想挑点儿事。我对他说：“你知道……我之前在微软演讲，他们送了我新款的音乐播放器Zune，不得不承认，它比你们的iPod好太多了。”库奇看着我，笑了笑，回答说：“我对此毫不怀疑。”他只说了这么一句。

对于微软有一款更好产品的事实，这位苹果高管泰然自若。这或许只是出于一个市场主导者的勇气，或许只是装作不在意（但表现得非常自然），又或许是出于别的什么原因。尽管我当时并不知道他的确切心思，但他的回应是一位具有无限思维的领导者会有的很自然的反应。

无限思维的益处

一个组织可能会自行挑选一个时间段以及一组衡量标准，用以评估自己有多成功，但在无限游戏中，要评估一个组织的真正价值，却不能以此为依据。一个组织的真正价值取决于他人想要为该组织做贡献，令其能够一直成功下去的意愿，而这个“一直”不仅指他们自己任职期间，也包括他们卸任之后的长远未来。有限思维的领导者努力从员工、客户和股东那里获取，从而满足其自选的衡量标准；无限思维的领导者则努力激励员工、客户和股东，让他们能主动且持续地为该组织的发展做出贡献。无限思维的玩家希望让自己的组织变得比他们加入时更好。乐高能发明出经受住时间考验的玩具，并不是因为幸运，而是因为几乎所有的乐高人都希望为公司的发展出一分力，希望公司能够存续得比他们的生命更长久。他们的动机不是要在某一个季度争先，而是要“继续创造出新颖的游戏体验，每年都能走到更多的孩子身边”。³

在詹姆斯·卡斯看来，有限思维的玩家参与游戏的目的是以赢家的身份终结这场游戏。若想成为赢家，就势必会有输家。有限思维的领导者为自己而战，想要打败其他所有玩家。他们所制订的每一个计划，所采取的每一个行动都以赢为目的。他们坚信这是他们必须做的，然而事实上，他们本不必如此，也并没有规定要求他们必须如此，引导他们这样做的是他们自己的思维方式。

卡斯认为，无限思维的玩家参与游戏的目的就是一直玩下去。在商业领域，这意味着组织本身能够存在得比其领导者更长久。无限思维的玩家会按照对游戏有益的方式去玩。在商业领域，这意味着不能只看到

自己的净收益。有限思维的玩家生产的是他们认为可以卖得掉的产品，无限思维的玩家生产的则是人们想要购买的产品。前者主要关心的是销售这些产品对公司有何益处，后者主要关心的是这些产品对购买它们的消费者有何益处。

有限思维的玩家所做的往往是，遵循有助于实现他们个人目标的标准，但较少考虑可能因此而产生的连锁反应。“什么对我最有利”就是有限的思维方式，而“什么对我们最有利”才是无限的思维方式。为参与无限游戏而创立的公司并不会只考虑自己，在做决定时，还会考虑这些决定对员工、社会、经济、国家乃至整个世界的影响，所做的是对整个游戏都有利的。

柯达公司创始人乔治·伊士曼（George Eastman）致力于实现自己的愿景：让摄影对每个人来说都简单易学。他意识到，推进自己的愿景与实现员工和他们所在社群的福祉密切相关。1912年，柯达成为第一家会基于公司业绩为员工分红的企业。⁴数年后，柯达发行了股票期权。柯达还为员工提供一系列优厚的福利，例如提供带薪病假（这在当时是一个新理念），并为进修当地大学课程的员工提供学费补贴。这些举措现已为众多公司采用，换言之，这不仅对柯达这一家公司有利，对整个商业游戏都有利。除了提供数以万计的工作岗位，伊士曼还建了一家医院，办了一所音乐学校，并慷慨资助了一些高等学府，包括罗切斯特大学和罗切斯特机械学院（后更名为罗切斯特理工学院）。

卡斯告诉我们，有限思维的玩家始终抱着一个终极目的在参与游戏，因此并不喜欢惊喜，也害怕任何形式的混乱。他们不喜欢无法预测或无法控制的事情，因为这些事情有可能会打乱原本的计划，增加输的概率。无限思维的玩家则期待惊喜，甚至沉醉其中，他们时刻准备着迎接这些惊喜所带来的改变，享受游戏所赋予的自由，乐于接受游戏中的一切可能性。他们所寻找的并不是用于应对已发生事件的方法，而是创新的方法。无限的视角可以解放我们的思维，让我们能够专注于更宏大

的愿景，而非过度关注其他公司正在做些什么。以新技术为例，无限思维的玩家更擅长预测新技术在未来的应用，而非忙于应对新技术将对现有商业模式构成的挑战。

说到这里，就可以理解约翰·库奇为何会对微软精心设计的音乐播放器如此淡然处之了。他明白，在商业这场无限游戏中，有时是苹果的产品更好，有时又是其他公司的产品更好。苹果的目标并不是设法超越微软，而是设法超越自己，一直在展望未来，思考在推出了iPod之后还可以研发些什么。苹果的这种无限思维不仅仅是帮助公司跳出既定的框架，更是超越这种框架去思考问题。微软的音乐播放器推出大约一年后，苹果就发布了第一代苹果手机。苹果手机重新定义了整个智能手机类别，音乐播放器几乎一瞬间就过时了。尽管有些人认为苹果具有预测消费者偏好和看透未来的能力，但事实并非如此。真正帮苹果开辟出创新之路的是其所具备的无限思维，这恰恰是有限思维的领导者所无法做到的。

有限思维的企业可能会想出一些“创新”的方法来提高利润，但这些决策通常不会带来利润之外的益处，比如造福于该组织本身，及其员工、客户和所在社群。这些决策也不一定有利于企业的长远发展。原因很简单，制订这些决策的主要目的往往是让那些不具备无限思维的决策者获益，但也只能短期获益。相反，无限思维的领导者不会要求员工专注于有限的目标，他们会以实现人人皆可受益的无限愿景为目标，号召员工与自己一起努力，寻找方法。有限的目标只是朝这个愿景迈进的阶段性标志。当每个人都专注于实现无限愿景时，不仅能推动创新，还能推动企业各项数据的增长。事实也证明，无限思维的领导者所领导的公司往往能获得创纪录的利润。更重要的是，无论是在经济稳定时期，还是不稳定时期，无限思维的领导方式都能为企业带来更多的灵感、创新、合作、品牌忠诚度和利润，能帮助企业在顺境中更加繁荣，也能帮助企业坚韧地走出逆境。

若要创办一家有适应力的企业，企业结构就应该是有利于企业永久存续的，这与创办一家稳定的企业不同。稳定，顾名思义，就是保持不变。理论上说，一个稳定的组织能够经受住风暴的考验，并能像之前一样从风暴中安全走出。实际上，若称一家公司稳定，往往是在与另一家风险更高、业绩更好的公司对比。“增长缓慢但稳定”，这就是人们的想法。然而，一家以稳定为目的创办的企业无法理解无限游戏的本质，因为它可能并没有准备好应对不可预知的情况，新的技术、新的竞争对手、市场的变化或全球性的大事件，这些都可能在瞬间打乱他们的既定战略。无限思维的领导者想要的不仅仅是一家能够安然挺过外界变化的公司，更是一家可以随外界变化而变化的公司。他们想要建立一家能够拥抱惊喜、适应变化的公司。有适应力的公司，在经历剧变的前后可能会呈现截然不同的样子，而且往往对这种变化充满感激。

以生产瑞士军刀闻名世界的维氏瑞士军刀（Victorinox）的业务就曾受到“9·11”事件的严重冲击。“9·11”事件发生前，该公司的产品深受许多企业采购部门的青睐，也是人们钟爱的退休、生日和毕业礼物。但在“9·11”事件之后，瑞士军刀被禁止出现在人们的手提行李中。面对这样的情况，大多数企业会摆出一种防御性的姿态，将目光锁定在自身传统模式所遭受的打击与企业将会蒙受的损失上。而维氏却选择主动出击，将这一突发事件视为机遇而非威胁，这是无限思维的玩家的典型举动。维氏的领导者并没有采取极端的成本削减与裁员措施，而是通过创新的方法保住了员工的工作（他们没有裁掉任何一个人），增加了对新产品开发的投资，并鼓励员工主动思考如何利用已积攒到的品牌优势打入新的市场。

维氏在顺境时就建立了现金储备，因为他们知道未来或许会出现更艰难的时期。⁵正如其首席执行官卡尔·埃尔森纳（Carl Elsener）所说：“回顾一下世界经济的发展史，就会发现经济的发展总是如此。总是如此，未来亦将如此！经济永远不可能只上升，也永远不可能只下降。它会起起伏伏，伏伏起起……我们不能以季度为单位去思考，要考

虑永存。”正是得益于这种无限思维，维氏既做好了思想上的准备，也做好了财力上的准备，当危机事件发生时，对其他企业来说可能是致命的，但维氏却能从容应对。与“9·11”事件发生之前相比，维氏大不一样了，甚至更加强大了。过去，刀具占了维氏总销售额的95%，仅瑞士军刀就占了80%。如今，瑞士军刀只占其总收入的35%，旅行工具、手表和香水的销售则使维氏的收入较“9·11”事件之前翻了近一番。维氏不是一个稳定型的企业，而是一个适应型的企业。

无限思维的益处显而易见，而且是多方面的。那么，若将有限思维用到商业这一无限游戏中，又会发生什么呢？

在无限游戏中运用有限思维会增加我们陷入困局的概率，会令我们迅速消耗掉继续参与游戏所必需的意志和资源。

商业上的困局

微软在发布音乐播放器Zune时，并没有想借助该产品来进一步发展宏大的愿景，考虑的不是未来存在着怎样的可能性，而只是将之当作一场关乎市场份额和金钱的竞争。在该竞争中，微软的表现并不是太好。史蒂夫·鲍尔默对Zune能够“击败”iPod的预测真是大错特错。Zune刚上市时的市场份额是9%，然后逐渐下降，到2010年时已经跌至1%，2011年就停产了。⁶相比之下，同一时期，iPod的市场份额一直保持在70%左右。

一些人认为Zune失败的原因在于，微软在广告上的投入不足，但这个理论站不住脚。美国内衣品牌Spanx、辣椒酱品牌Sriracha和运动相机品牌GoPro就是依靠口口相传来提高品牌知名度的。⁷没有采用传统广告，这些公司一样从默默无闻中崭露头角，蓬勃发展。还有一些人则认为，Zune的失败是因为微软进入MP3播放器市场的时间太晚，这个理论也没有比前一个好多少。苹果推出iPod的时间，距离MP3播放器成为一

个广为人知的产品类别也有整整5年。当时，索尼等品牌都已经在深入开发该技术，且产品销售情况很好了。然而，自2001年问世后的4年内，iPod就成功占据了美国数字音乐播放器市场的最大份额，而且这个数字还在不断增长，完全没有回落。⁸

一如微软的音乐播放器这样可能十分出色的产品，它们的问题并不出在设计、营销或产品发布的时机上。若想在商业的无限游戏中生存并发展，需要的远不止这些。历史上，失败的伟大产品比比皆是。我们必须考虑的一大要素就是企业的领导方式。在有限思维的领导者眼中，与对手进行比较和赢是企业的头等大事，因此在制定企业战略、产品战略、激励机制和招聘决策时，他们都是以帮助实现这些有限目标为导向的。随着有限思维模式几乎深深渗透到企业的方方面面，该企业的视野就会变得狭隘。最终，在该企业中，几乎人人都会将过多的注意力放在紧急而非重要的事情上。管理者会本能地根据已知因素来做出应对，而非主动探索或推进未知的可能性。有时，领导者可能会过分执着于竞争对手在做什么，并错误地认为对方每动一步，自己都必须有所应对，如此一来，他们就会忽视掉大量更有利于其组织发展的选择。这就像是想通过防守策略来赢得比赛。在有限思维的诱惑下，微软陷入了一场没完没了的“打地鼠”游戏。

微软的领导者并不理解他们所参与的是一场无限游戏，也不理解苹果在其中所运用的无限思维。尽管史蒂夫·鲍尔默有时会提及“愿景”或“长期”这样的词，但与其他爱用这类无限语言的有限思维的领导者一样，他的讲话几乎总是基于一些有限的主题，比如地位、股票表现、市场份额和金钱。因为思维方式的错误，微软所追逐的目标——赢，根本就不可能实现。

然而，微软似乎并没有从与iPod的竞争上吸取教训。2007年苹果手机问世时，鲍尔默的反应再次凸显了他的有限视野。某次采访中，当被问到苹果手机时，他嘲讽道：“苹果手机绝不可能拿到多大的市场份

额，绝不可能……他们或许能赚到很多钱。但你们好好想想，看看全球售出的13亿部手机，我更乐意做的是，让其中60%、70%或80%的手机都安装上我们的软件，而不是仅仅拥有其中2%或3%的市场份额，而苹果也就能拿到这么多的市场份额。”⁹受限于有限思维，鲍尔默更关注苹果手机相对其他企业可能的业绩表现，而不是它可能对整个市场带来的改变，更不用说是完全重新定义手机在我们生活中所扮演的角色了。然而，一定会气疯鲍尔默的是，苹果手机上市仅仅5年，其销量就超过了微软所有产品的总销量。¹⁰

2013年，在以微软首席执行官身份出席的最后一场新闻发布会上，鲍尔默以有限思维的方式总结了自己的职业生涯。他选了一些自己在任期内完成的指标来定义成功。“在过去的5年中，苹果赚的钱可能比我们多，”他说，“但我敢断言，在过去的13年中，我们赚到的钱超过了地球上的几乎任何人。坦率地说，这就是我的一大骄傲所在。”¹¹鲍尔默想说的似乎是，在他领导的这13年，微软“赢”了。试想一下，若是在这场新闻发布会上鲍尔默做的不仅仅是回顾企业的资产负债表，还能够分享一下微软为了实现比尔·盖茨最初提出的无限愿景——“予力全球每一人、每一组织，成就不凡”做了什么，那这会是一场多么不一样的发布会啊。

有限思维的领导者会用公司的业绩来证明自己职业生涯的价值，而无限思维的领导者会用自己的职业生涯来提升公司的长期价值。在这些价值中，仅有一部分可以用金钱来衡量。这场游戏并不会因为鲍尔默的退休而结束。没有他，微软仍将继续玩下去。在这场无限游戏中，鲍尔默究竟对公司的财务状况做了多大贡献其实并不重要，真正重要的是他是否帮公司做好了充足的文化准备，让公司能够在下一个13年、33年或300年中继续存续并繁荣下去。若按照这个标准评估，鲍尔默输了。

在商业的无限游戏中，领导者若是坚持有限思维的模式，或者过度执着于有限的目标，那么在某个时间段内，根据某个衡量标准，他们或

许能够达成“第一”的目标。但这并不意味着他们所做之事一定有利于企业在长久游戏中不被淘汰。事实上，更常见的情况是，这样的决策不仅有害于公司的内部运作，而且若不加以干涉，还会加速公司的灭亡。

有限思维的领导者往往过分关注短期结果，因此，他们所运用的战略或战术往往也着眼于业绩数字的提高，比如，减少研发投入、大幅削减成本（例如经常性裁员，选择更便宜、更劣质的产品原料，在制造或质量控制上偷工减料）、通过收购其他公司来实现增长以及股票回购。这些决策反过来会动摇公司的文化。人们开始意识到没有什么东西是安全的，也没有什么人是安全的。由此，一些人的行为仿佛本能地切换到了自我保护模式。他们可能会私藏信息、掩盖错误，并在行事时更谨慎、更注意规避风险。为了保护自己，他们不信任任何人。另一些人则会更加依赖“唯适者才能生存”的心态。他们的策略会变得过于富有攻击性，自我变得不受遏制。他们学会通过获得高层青睐往上爬的方法，有时，这甚至意味着会故意阻碍自己的同事。为了保护自己，他们不信任任何人。无论是自我保护还是自我推销，这些行为加在一起，都会导致整个公司范围内合作度的普遍下降，进而导致公司内难以出现任何真正有新意或有创意的想法。

出于对有限游戏的执着，微软非常痴迷季度数据，反而丧失了灵感、想象力和创新能力，对此，许多在微软成立之初就加入的老员工深感痛惜。¹²公司的内部团队开始相互竞争而非相互支持，彼此间的信任与合作受到了影响。正如《名利场》（*Vanity Fair*）的报道所言，这家公司曾“拥有富有远见卓识且才华横溢的年轻领导者，是一台精益竞争机器，可如今，却变成了一个臃肿、官僚主义盛行的企业，其内部文化无意中激励了那些扼杀创新思维的管理者，这些管理者害怕新的思维会威胁公司现有的秩序”。¹³换言之，有限思维将微软的企业文化搞得一团糟。

对超大型企业来说，无限思维的领导者所积累下来的意志和资源是

足够有限思维的领导者“挥霍”很长一段时间的。因此，在鲍尔默的领导下，微软仍然是游戏中的主导者，尤其在商业市场上。这在很大程度上要归功于比尔·盖茨所打下的基础，盖茨拥有着更无限的思维。然而，若鲍尔默继续留任，或者由另一个有限思维的领导者接任，那么，那些坚持良性竞争的人的意志，以及继续游戏所需的资源，终将耗尽。对一家公司来说，仅仅规模庞大、财力雄厚，并不意味着就已强大到能够一直存续下去。

微软的经历并非特例，商业史上不乏类似的警示故事。比如通用汽车，他们曾执着于市场份额超过利润，若非有政府的紧急财政援助，早就破产淘汰了。至于西尔斯百货（Sears）、电路城（Circuit City）、雷曼兄弟（Lehman Brothers）、美国东方航空（Eastern Airlines）和百视达（Blockbuster Video）就没有这么幸运了。这些企业给商业史又添了几个例子，让我们看到那些曾经强大的老牌企业，是如何在沉迷于有限思维的领导者的领导下走上毁灭之路的。

可悲的是，在过去的三四十年里，有限思维的领导方式已经成为现代商业的标准模式，既受到华尔街的欢迎，也成为众多商学院的授课内容。与此同时，企业的寿命似乎越来越短了。麦肯锡的一项研究显示，自20世纪50年代以来，标准普尔500指数涵盖的公司平均寿命已经减少了40多年，从平均61年降至如今的不足18年。¹⁴耶鲁大学教授、麦肯锡荣休董事理查德·福斯特（Richard Foster）称，变化“正在以前所未有的速度加快”。¹⁵我承认，上述数据是多因素作用的结果，但我们也必须考虑到，如今有太多的领导者在企业创办之初就没有为其奠定长久存续的基础。这是很讽刺的，因为即使是最以目标为导向的有限思维的领导者也不得不承认，一个组织只有存续、繁荣的时间越长，其实现所有目标的可能性才越大。

过多的有限思维的领导方式会影响到的不仅仅是企业。随着越来越多的有限思维者成为各个生活领域中的权威者，公共政策也会被迫向着

进一步巩固有限思维的方向转变。然后，用不了多久，整个经济运行都会受限于有限思维，经济运行所遵循的规则也将与我们所处的这个游戏不符。这种处境是不堪一击的，数据也印证了这一点。1929年，股市大崩盘导致经济大萧条，美国政府出台了《格拉斯-斯蒂格尔法案》

（*Glass-Steagall Act*）。¹⁶自从《格拉斯-斯蒂格尔法案》通过并实施后，股市大崩盘的情况就再也没有发生过，一直到二十世纪八九十年代，该法案名存实亡，三次经济灾难又随之而来：1987年的黑色星期一、2000年的互联网泡沫破裂，以及2008年的金融危机。

若在无限游戏中运用有限思维，我们就会不断做出有碍实现自身抱负的决定。这就像以“享受生活”为名大吃甜点，最终令自己患上了糖尿病一样。若是有过多有限思维的玩家参与游戏，确实可能会埋下“股市大崩盘”的隐患，但这是一种极端的结果。更有可能出现的情况是，组织内部的信任、合作和创新普遍减少，而这将大大增加该组织在快速变化的商业世界中存续和繁荣的难度。若我们相信信任、合作和创新对自身组织的长期前景至关重要，那么就只剩下一个选择——学习如何使用无限思维。

无限思维的五个要素

在选择领导方式时，我们必须考虑以下三点：

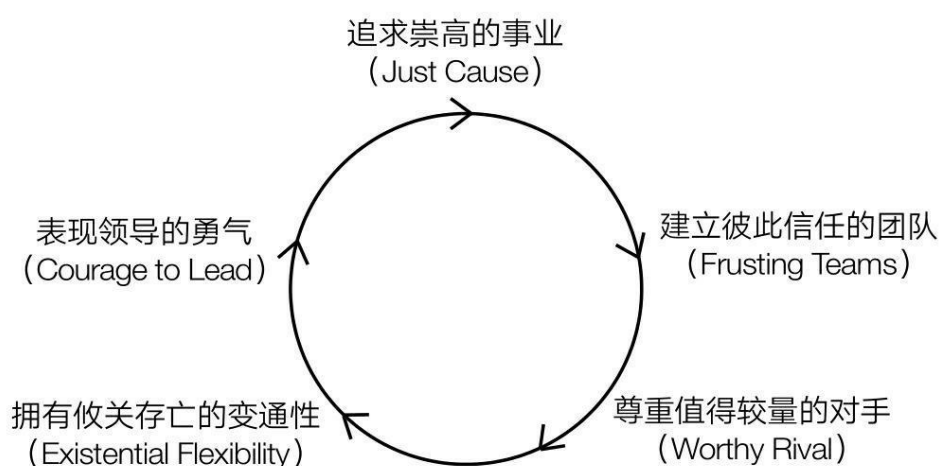
1. 某一特定游戏是有限的还是无限的不是由我们决定的。
2. 我们可以选择是否要参与该游戏。
3. 如果决定参与，我们可以选择到底是运用有限思维还是无限思维。

若参与有限游戏，则显然要遵守对的规则才能增加胜算。比如足球

赛，若按篮球赛的规则来准备显然是毫无用处的。若要领导团队参与一场无限游戏，道理也是一样的。只有按照这场游戏的规则来玩，才能增加自己生存并繁荣的可能性。

用无限思维来领导团队的这一选择与其说像备战足球赛，不如说像你决心要练出好身材。这不是只做一件事就足够的，我们不能指望去健身房一次性锻炼9个小时就能立刻拥有健美的身材。不过，若是能坚持每天都去健身房练上20分钟，就一定能练出好身材，坚持比强度更重要。但这里的问题是，没人知道我们到底什么时候才能看到成果。事实上，出成果的时间点也是因人而异的。但有一点毫无疑问，就是我们都知知道这样做是百分之百有用的。当然，我们可能会设定一些有限的健身目标，但若要尽可能地保持健康，所选择的生活方式比我们是否有按时完成上述健身目标更加重要。所有的健康计划都有一些共性，比如要求我们多吃蔬菜、经常锻炼和保证充足的睡眠。运用无限思维也是同样的道理。

若想运用无限思维来领导团队，就必须做到下图中的五点。



无限游戏的五个要素

如果我们选中了某个健康计划，可以选择只听从其中的部分而非全

部建议，比如，积极锻炼，但不吃蔬菜。这样做，我们或许能享受该计划所带来的部分益处，但若想获得全部的益处，就必须全部照做。同理，无限思维的上述要点，你只做到一部分确实也能获益，但若想让一个组织具备在无限游戏中长久、健康发展下去的能力，就必须无一遗漏地做到所有。

坚持无限的思维方式是很难的，非常难。偏离也在意料之中，毕竟人无完人。我们可能会受到各种因素的影响，比如贪婪、恐惧、野心、无知、外部压力、利益冲突、自我等。雪上加霜的是，有限游戏是充满诱惑力的，它们可能很有趣、很令人兴奋，有时甚至会令人上瘾。就像打赌，每赢一次（相当于每实现一个目标）都会刺激我们的身体释放多巴胺，这会驱使我们再做一次同样的事，想要设法再赢一次。我们的心志必须足够坚强，才能克制住这股冲动。

我们不能奢望自己或每位领导者都拥有完全无限的思维方式，或者奢望所有无限思维的领导者都能始终维持这种思维。正如专注于固定的有限目标比专注于未来的无限愿景更容易，用有限思维领导公司也会更容易，尤其是在经济困难或衰退时期。我在本章中引用的每一个例子，包括正面的例子，这些公司都曾在自身发展的某个阶段出现过有限思维的领导者，他们抛弃了作为该公司创立之本的无限根基，转而执着于追求有限的目标。事实上，所有这些公司都差一点毁灭于有限思维的模式。只有那些被无限思维的领导者救下的幸运儿才得以存续，并变得更加强大，变得比以前更能激励自己的员工，更能生产有吸引力的产品。

无论我们选择的是哪种游戏方式，有一点都至关重要：我们要诚实面对以及诚实告知他人我们的选择，因为我们的选择会产生涟漪，影响他人。只有当我们周围的同事、客户、投资者清楚知道了我们的选择，他们才能相应地调整自己的期望和行为。只有知道了我们的思维方式，他们才能判断自己可能受到什么样的短期和长期影响。他们有权知道我们将采取什么样的行事方式，只有这样，作为求职者、消费者或投资者

的他们才有可能做出明智的选择。我们若能做到无限思维的领导者所必需的五大要素，就能让他们相信，我们会专注于前进的目标，会致力于在前进中相互支持。也能让他们相信，我们所创办的这个组织，是想要长久地存续与繁荣下去的，因此我们会努力地抵御短期诱惑，采取合乎道德的行动。

对于我们这些选择拥抱无限思维的人来说，无限思维将引领我们踏上这样的人生旅途：每天醒来都干劲满满，在外工作倍感安心，一天结束心满意足。当属于我们的游戏时间结束时，回顾这一生，回顾职业生涯，我们可以问心无愧地说：“我这一生活得很值。”更重要的是，在畅想未来时，我们眼中所浮现的，会是许许多多曾受我们激励的人，即使没有我们同行，他们也将继续这一旅途。

THE INFINITE GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE

要素 1

追求崇高的事业

第2章 崇高的事业

从1941年9月到1944年1月，纳粹对列宁格勒^③进行了近900天的围困，列宁格勒有超过100万市民死亡，其中包括40万儿童，许多人都是被活活饿死的。与此同时，不为人知的是，有数以万计的种子与成吨的土豆、大米、坚果和谷物藏在市中心。

大约是距离围城开始的25年前，一位名叫尼古拉·瓦维洛夫（Nikolai Vavilov）的年轻植物学家开始收集种子。在他小时候，遭遇过多次大饥荒，导致数百万人丧生，因此他决定要毕生致力于消除饥饿，防止类似的生态灾难再次出现。最终，这一理想主义的目标成为瓦维洛夫一生追求的事业。他周游世界，收集各种各样的粮食作物，并研究其中一些作物比另一些更具适应性的原因。没过多久，他就收集到了超过6000种农作物的种子。同时，他还研究遗传学，尝试培育新的作物品种，使它们更好地抵抗虫害和疾病、更快地生长、更耐受恶劣环境，实现更高的粮食产量。随着工作的推进，瓦维洛夫明确了要建立一家种子银行的愿景。就像我们会备份重要数据，以防计算机崩溃导致数据丢失一样，瓦维洛夫想为全世界的所有粮食都备份一份种子，以防有任何物种因自然灾害或人为灾难而灭绝或无法生长。

1920年，业已享有盛誉（且拥有了更丰富种子收藏）的瓦维洛夫结束了高校教师的生涯，成为列宁格勒应用植物学研究所的负责人。在政府的资助下，瓦维洛夫组建了一支科学家团队，共同推进他的事业。瓦维洛夫在刚加入该部门时曾写道：“我希望将该部门建成一个不可或缺的机构，尽可能让它对每一个人都有用处。我想将全世界的植物种子都收集起来，将这个部门打造成所有农作物及其他植物的宝库。”一如任

何具有无限思维与远见卓识的人一样，他总结道：“我并不知道最终结果会如何……但无论如何，我都想努力一试。”¹

1940年，瓦维洛夫遭人诬陷被捕。1943年，这位拥有远见卓识且毕生致力于终结饥饿的植物学家因营养不良在狱中去世，年仅55岁。

瓦维洛夫去世时，列宁格勒正陷于被围困的水深火热之中。就在这一战区的中心地带，圣以撒广场（St. Isaac's Square）的一幢毫不起眼的建筑里，藏着瓦维洛夫团队所有的研究记录，当然，还有他们的无价之宝——数以万计的各种农作物种子。除了炮击这样显而易见的风险，这些种子还面临着市内老鼠成灾的威胁。更糟糕的是，瓦维洛夫的收藏还引来了纳粹的关注。

被围困期间，尽管面临重重威胁，尽管处境与其他列宁格勒市民一样恶劣，瓦维洛夫所带领的科学家团队仍在继续工作。他们会在隆冬时节冒险外出，在靠近前线的田地里秘密播种土豆。尽管他们能够将一些劳动成果偷偷带出城去，但剩下的都被他们藏了起来，小心看管着。这些科学家无私地投身于实现瓦维洛夫的愿景，做好了不惜一切代价保住种子银行的准备，哪怕牺牲自己的生命。尽管身边有着数以万计的种子，成吨的土豆、大米、坚果、谷物和其他农作物，但这些科学家都不肯吃，最终，有9名科学家死于饥饿。

曾有人引用瓦维洛夫的话来谈论自己所追求的事业：“我们将会走进柴堆，我们将会烈火焚身，但我们不会放弃自己的信念。”²那些共同投身于此的人，不单单是被瓦维洛夫的话所鼓舞，他们在践行瓦维洛夫的话。瓦迪姆·列赫诺维奇（Vadim Lekhnovich）就是幸存者之一，他曾帮忙种植过土豆，并在枪林弹雨中守护过它们，后来曾有人问他，明明储藏了海量的食物，为什么不吃。“当时已经虚弱到步子都挪不动了。每天早上想要起身，想要动一动手、动一脚，都是难如登天。”他说，“但想要克制自己不去吃那些储藏的粮食却一点也不难。因为其中包含的可是你毕生追求的事业，是你的同伴们毕生追求的事业。”³

这些科学家之所以能在围困中仍然继续推进瓦维洛夫的工作，是因为他们觉得自己在参与一项远比他们个人更伟大的事业。这一崇高的事业被瓦维洛夫称为“全人类的使命”，因为它，他们的工作和生命都有了目的和意义，这个目的和意义超越了任何个人，超越了任何人在被围困期间所面临的痛苦和挣扎。这些科学家的目光更长远，他们所设想的未来是可以用自己的工作拯救全人类，他们参与这场游戏的目的是让全人类尽可能长久地生存下去。

什么是崇高的事业

霍华德是少年棒球联盟的球员，他所在的球队是全联盟最差的球队之一。不过，每次输比赛，教练都会在赛后告诉球员们：“输赢并不重要，重要的是我们在比赛中是怎么打的。”这时，霍华德就会举手问教练：“那我们为什么要记分呢？”

在参与有限游戏时，我们的目标是赢。即使我们只期待自己能打得不错，只是为了享受比赛，但也绝对不会以输为目标。参与无限游戏的动机则截然不同，并不是为了赢，而是为了能继续玩下去，是为了推进比我们自身或所在的组织更伟大的事情。任何想要在无限游戏中成为领导者的人都必须拥有一项愿意为之努力的、非常明确的崇高的事业。

崇高的事业是对尚不存在的未来状态的具体设想，这个未来状态极具吸引力，以至于人们愿意为实现它而做出牺牲，就像瓦维洛夫团队的科学家们一样，甚至愿意献出自己的生命。但并非所有崇高的事业都需要做到这一步。我们要做的可能是拒绝一份薪水更高的工作，继续留在致力于推进我们所相信的崇高的事业的组织。这可能意味着熬夜工作，可能意味着经常出差。尽管可能不喜欢这些牺牲，但因为有了崇高的事业，我们会觉得这些牺牲是值得的。

赢只能让你获得短暂的喜悦，虽能大幅增强我们的自信心，但这种

效果转瞬即逝。无论一年前的你实现了某个目标，无论获得了晋升，还是赢得了锦标赛的冠军，当时那种美妙的成就感都是不可能一直延续的。这些感觉会过去，若要再次体会那种感觉，我们就必须努力再赢一次。然而，若有一项崇高的事业，一个比任何一场特定胜利都更重要的理由，我们的日子就会过得更有意义，过得更有满足感。这样的感觉才能周复一周、月复一月、年复一年地持续下去。若就职于一个以有限目标为动力的企业，我们或许会有喜欢这份工作的时候，但可能永远无法爱上它。如果就职于一个为崇高的事业而奋斗的企业，我们或许会有不喜欢这份工作的时候，但会一直热爱它。就像自己的孩子，我们或许觉得他们时而讨人喜欢，时而又讨人厌烦，但对他们的爱是不会中断的。

崇高的事业不同于“为什么”。“为什么”来自过去，是关于起源的故事，是关于我们是谁的声明，是价值观与信念的总和。^④而崇高的事业则关乎未来，决定我们前进的方向，描绘了我们想要生活其中并愿意投身建设的那个世界。每个人都有属于自己的“为什么”，而且，只要愿意挖掘，每个人都能清楚知道自己的“为什么”，但并非每个人都必须拥有一份属于自己的崇高的事业，我们可以选择加入他人的事业。我们既可以自己发起一场运动，也可以选择加入一场运动，然后将其变成自己的。与只有一个的“为什么”不同，我们可以努力推进不止一项崇高的事业。我们的“为什么”是固定不变的。但崇高的事业关乎的是迄今尚未实现的未来，因此我们并不知道它最终会呈现何种样貌。我们可以按照自己喜欢的方式，坚持不懈地建设，还可以在推进的过程中不断改进。

“为什么”就像一栋房子的地基，是起点，无论我们在它之上修建什么，都能从中获取力量与坚持的动力。崇高的事业则是我们对这栋房子的构想，是我们希望它最终能够呈现的样子。或许，我们需要穷尽一生，也或许穷尽一生都未必能将它建成。但无论如何，我们的努力都是在为它的最终成形添砖加瓦。随着构想一点点变为现实，会有更多的人受到激励，加入我们，继续推进这项事业，直到永远。

我的“为什么”是激励人们做能激励他们自己的事情，这样一来，我们才能一起将这个世界变得更加美好。这个“为什么”对我来说是独一无二的。再说说我为之努力的崇高的事业，我想要打造一个这样的世界：大多数人都会在想要工作的热情中醒来，能安心地在外工作，能带着工作的满足感踏上回家的路。我正在尽一切可能寻找愿意加入我，与我一同推进这项事业的人。

让我们的工作与生活充满意义的正是我们为之奋斗的崇高的事业。崇高的事业能让我们保持专注，不为有限的奖励与个人的胜利所诱惑。在为崇高的事业奋斗的路上，也会有必须参与的有限游戏，但一切有限游戏都离不开崇高的事业这个大背景。崇高的事业是让我们想要坚持游戏的动力。无论是在科学研究、国家建设还是商业领域，领导者若想吸引我们加入，并为他们的无限追求而奋斗，就必须用清楚、明确、肯定的语言描绘出所构想的那个理想未来。

当美国的开国元勋们宣布美国要脱离英国独立时，他们知道，要发起这样一场颠覆性的运动，离不开一个正当的理由。他们在《独立宣言》中写道：“我们认为这些真理是不言而喻的，人人生而平等，造物主赋予每个人若干不可让与的权利，其中包括生命权、自由权和追求幸福的权利。”他们提出的愿景不是建立一个由边界界定的国家这么简单，这个理想的未来国家是用人人自由且平等的原则勾勒出来的。1776年7月4日，56人签署了这份宣言，同意“以我们的生命、财富和神圣的荣耀互相共同保证”。这充分说明了这份事业对他们的重要性。为了实现“建立新国家”这一无限的愿景与理想，他们愿意付出自己有限的生命和利益。他们的牺牲也能激励后来者，让后来者继续用自己的鲜血、汗水和泪水推进这一事业。

当我们坚信自己为之献身的事业会有后来者继承并发扬时，我们就知道它是正确的。美国的开国元勋们必定就是如此，尼古拉·瓦维洛夫也是如此。瓦维洛夫的愿景是，让这个世界上所有人不再挨饿，从而

让人类尽可能长久地生存下去，这一愿景一直被继承并发扬到今天。目前，全球共有近2000个种子银行，分布在100多个国家，那里的工作人员正在继续瓦维洛夫于一个世纪前开始的工作。位于挪威的斯瓦尔巴全球种子库是全球最大的种子银行之一。斯瓦尔巴种子库位于北极地区，选取的储存场所因地利原因，天然温度就已符合种子的储藏需求，现存有来自近6000种植物的超过10亿颗种子。它的存在是为了确保即便在最坏的情况下，我们仍然能有维持自己物种生存的必要食物来源。全球作物多样性信托基金（The Crop Trust）是一个与联合国合作建立的组织，旨在支持全球各种种子银行的工作。该基金的总干事玛丽·哈加（Marie Haga）认为瓦维洛夫可被视为该事业的开创者。⁴“一个世纪以前，瓦维洛夫首次迈上收集种子的旅途，”她说，“现在，新一代作物多样性的支持者们继续献身于此，前往世界各地，不仅为了保护种质，也为了保护瓦维洛夫的遗产。”

现在有许多企业都把自己的目标、愿景或使命写到办公室的墙上，领导者这么做的目的是想要激励员工。然而，这些目标、愿景或使命大都不具备成为崇高的事业的资格。有的虽不具备激励人心的作用，但无伤大雅，最糟的情况是，它们甚至会令我们误入歧途，让我们一直被困在有限游戏的圈子里。即使是一些再善意不过的声明，若是用有限、笼统、以自我为中心或模棱两可的语言来表述，也是无法在无限游戏中发挥任何作用的。比如“你不想做的事我们来做，这样你就能专注于自己爱做的事情了”。这说的可能是真心话，但太啰唆了，尤其是在一个企业的环境中。另外，这也不是一个可以凝聚人心的口号。再比如，有一种表述笼统的常见愿景是：尽可能提供最高价值、最高质量的产品，等等。此类声明对那些想在无限游戏中领导我们的人来说是毫无帮助的。这样的愿景无法让人产生参与感，是以自我为中心的，考虑的只是公司的利益，只关注了企业内部，而不是产品或服务的长远发展。

VIZIO是一家电视机、扬声器制造商，总部位于美国加利福尼亚州，其官网上写着，他们存在的目的是“通过不断创新，以更优惠的价

格为消费者提供更高性能、更智能的产品”。⁵我相信他们全都说到做到了。但这句话真的能激励员工吗？换作是你，读到这段文字时，是否会产生要去这里求职的冲动？若我们之中真的有人会因此而起鸡皮疙瘩，或是因此感受到内心的召唤，想要加入他们，那也只可能是极少数人。在这样的话语中，我们既看不到有可供我们投身其中的事业，也完全看不出该企业所全力支持的到底是什么，而这二者在无限游戏中都是必不可少的。

再次强调，崇高的事业是一种特定的愿景，描述的是尚未实现的一种未来状态。若想用崇高的事业指引我们的工作方向，激励我们做出牺牲，并让这一崇高的事业存续得比我们的生命更长久，就必须符合下面的五个标准。你若不确定自己的目标、使命或愿景是否属于崇高的事业，或对用崇高的事业来领导团队很感兴趣，那么不妨用下面的标准来做一个简单的测试。

崇高的事业必须是：

- 支持某事的——积极的、乐观的；
- 有凝聚力的——欢迎所有愿意为之做出贡献的人；
- 服务导向的——为了他人的主要利益；
- 有适应力的——能够经受住政治、技术和文化的变迁；
- 理想主义的——宏伟、大胆但最终无法实现。

支持某事的——积极的、乐观的

崇高的事业是我们所支持和相信的，而不是反对的。领导者想要召集人们共同反对某事非常容易，甚至可以怂恿人们陷入疯狂。毕竟人在生气或害怕时，可能会出现很激烈的情绪。相较而言，若要支持某事，就需要被激励。对某事的支持能够点燃我们的激情，使我们充满希望与乐观。反对就是要诋毁、妖魔化或拒绝某事。支持是要请所有人共同来

推进一项事业。反对会令我们执着于业已存在的事物，从而诱导我们做出反应。支持会令我们专注于尚未实现的未来，激发我们的想象力。

想象一下，我们不说与贫困做斗争，而说捍卫每个人养家糊口的权利。前者创造了一个共同的敌人，是我们所反对的。它的目标是“可以赢的”，是一场有限游戏，让我们相信自己是能够一劳永逸地战胜贫困的。后者则给了我们不断前进的理由。这两个角度的影响不仅仅是语义上的，还会影响我们看待问题与愿景的方式，进而影响我们对自己该如何参与其中的判断。前者提供了一个等待我们解决的问题，后者则是提供了一个关乎可能性、尊严和赋权的愿景。它不是激励我们去“减少”贫困，而是激励我们“增加”能够养活自己与家人的人数。支持与反对之间的区别虽不易察觉，但影响重大。《独立宣言》的起草者似乎依靠直觉理解了这一区别。

短期来看，带领美国走向独立的开国元勋们确实是在反抗英国，《独立宣言》中有超过60%的篇幅是在具体控诉对英国的不满。不过，他们为之奋斗的事业才是其持久激励的真正来源，这一事业在《独立宣言》中也有着最为重要的地位，是其开篇最先言明的内容，也为《独立宣言》的剩余部分奠定了基调，指明了前进的方向。这一理想与每个人都息息相关，也很容易被美国民众记住。除了学者与狂热的历史爱好者，几乎没有哪个美国人能迅速背出那超过60%篇幅的控诉，但大多数美国人都能毫不费力地背出“人人生而平等”这句话，也往往能迅速说出“生命、自由、追求幸福”这三大基本原则。这些话在美国民众的文化心理上留下了不可磨灭的印记，频频被美国的爱国者、政治家引用，用来提醒美国人想要努力成为怎样的人，以及作为国家建立之基的理想是什么。这些话告诉美国民众，他们所支持的到底是什么。

有凝聚力的——欢迎所有愿意为之做出贡献的人

人类想要有参与感，渴望有归属感。我们享受作为团体一分子的感受。

觉，正因如此，我们会去教堂，会参加集会，或是在看球赛时穿上我们最喜爱的球队的队服。崇高的事业就像是一张邀请函，邀请他人与我们共同推进比自身更伟大的事业。若这则崇高的事业宣言能够在我们的脑海中勾勒出一一种积极且具体的未来愿景，就能触动我们的内心，让我们想要主动加入，与他人共同奋斗。

一份精心撰写的事业宣言能为我们勾勒出关于未来的崭新愿景，并激励我们贡献自己的想法、时间、经验，以及任何可能有助于实现它的东西。运动就是这样形成的，它始于少数人，这些人理想化的未来愿景吸引了其他相信他们的人。那些早期参与者并不是为了得到什么，而是为了给予。他们想要帮忙，希望在建设这个新未来的道路上贡献自己的一分力量。最初吸引他们的这项事业最终也成为他们自己的事业。

有些组织只是承诺要“改变世界”或“产生影响”，但其中几乎没有任何关乎具体目标的线索。这些承诺本身是好的，但太宽泛了，无法成为可供我们使用的有效过滤器。再提醒一下，崇高的事业应是对尚未实现的未来状态的特定愿景，这一未来状态是吸引人的，能够让人们愿意为推动它的实现而做出牺牲。我们称之为“愿景”，因为它必须是我们能“看到”的“景”。崇高的事业若想具备号召力，所勾勒的愿景就必须是具体、可感知的，能让我们知道自己可以产生什么样的影响，或知道那个更美好的世界究竟是什么样子。只有当脑海中能够浮现出组织或领导者想要实现的那个世界时，我们才能知道自己更愿意为什么贡献精力与我们自己。只有清楚知道了该事业的目标是什么，才有可能点燃自己的激情。

许多招聘方常常把这样一个标准挂在嘴边：“我们只聘用有激情的人”。而他们怎么知道某个应聘者是不是只对面试充满激情，而对他们的事业毫无兴趣呢？事实上，每个人都有激情，但不是所有人都会对同一件事充满激情。无限思维的领导者会积极寻找与他们有共同激情的员工、客户和投资者。对员工而言，这就是我们所说的“招聘要先考虑文

化，技能总是可以后面再教的”。至于客户和投资者，共同的激情也将成为他们热爱并忠于该组织的根源。

以沙拉快餐公司Sweetgreen为例，他们追求着比销售沙拉更伟大的事业，也欢迎同道中人加入。他们是这样描述自己的使命的：将人们与真正的食物连接起来，推动建立更健康的社区。Sweetgreen对“真正的食物”的定义是来自当地农场的食材，同时还能支持当地农场的发展。正因如此，Sweetgreen门店的菜单会因地制宜。或许有许多人只是单纯因为喜欢他们家的沙拉而在那里消费，但也有一些人是因为热爱本地食材，想要支持当地农场的发展，才受其吸引，成为最忠实的支持者。哪怕购买这些商品需要承受某种形式的牺牲，比如不够方便或价格更高，这些消费者也愿意购买。对消费者而言，以某种形式或方式支持这家公司是在践行自己的价值观和信仰，也是在为建造自己心目中那个更美好的世界做贡献，他们觉得自己也是这项事业的参与者。

服务导向的——为了他人的主要利益

崇高的事业必须至少涉及两方——贡献者与受益者、给予者与接受者。贡献者会为了推进该事业而给出一些东西，比如他们的想法、努力或金钱，接受这些贡献的人会从中获益。若想让自己的事业通过服务导向测试，那么该组织的主要受益人就不能是贡献者。

假设我的老板要给我提供一些职业建议，那么这些建议的主要作用就必须有利于我的职业发展而非他的职业发展。假设我是一名投资者，那我投资某家企业的主要目的就必须是有利于该企业朝着他们崇高的事业迈进。假设我是一名领导者，那我付出时间和汗水，以及决策的主要目的必须有利于我所领导的那些人。假设我是一名一线员工，那我辛勤劳动的主要目的必须是造福于购买我们公司产品与服务的那些消费者。若只存在一方，即自己是工作的唯一受益者，那我们所做的就不是崇高的事业，而只是形象工程罢了。

当Sweetgreen谈论自身贡献的受益者时，他们谈论的是社区和人，而不是这些贡献能为自己的公司做些什么。《独立宣言》的起草者们很清楚，他们的努力与美国独立战争的主要受益者都是“我们的人民”，而非“我们这些领导者”。如果这场战争的主要受益者是他们自己，那么美国最终很可能会沦为一个独裁国家或寡头国家。掌握了这个分析角度，再看到声称自己工作的主要受益者是股东而非顾客的企业时，我们便能立刻预料到它会有怎样的未来。

这里的关键词是“主要”。服务导向并不意味着做慈善。做慈善时，所有贡献的受益者，或至少绝大部分贡献的受益者必须是慈善的接受方。贡献者所能得到的好处就是因做贡献而得到的精神满足。但是在商业上，我们当然可以考虑如何通过工作来让自己获益或通过工作改善自己的生活状况，当然可以期待并要求获得与我们的努力和成果相匹配的薪水和认可，也可以让投资者获益，只是不能以牺牲公司、员工或消费者为代价。企业不应该为了让股东获利，而将节省成本的主意打到任何主要受益者的身上，比如逼迫顾客购买不合格的产品，或让员工失去他们的工作。毕竟归根结底，股东也只是众多贡献者中的一个。只有当一项事业的主要受益者是该组织外的另一方时，该事业才可能是正当的。

这就是“仆人式领导”（servant leadership）的含义。它意味着贡献所产生的主要益处是向下流动的。在缺乏服务导向（或将服务导向视为次要而非主要原则）的企业中，利益往往是向上流动的。投资者投资的主要目的就是比其他人更先赚到钱；领导者做决策的目的就是让自己比员工更先获益；销售员会无视顾客的需求，为赚到更多销售分红而不择手段。这就是当下普遍存在的利益流形式。在现代文化中，以保护自身及上位者利益为优先的人随处可见，他们将自己本应服务的对象放在了次要的位置上。

崇高的事业应使以服务为导向的要求符合无限游戏的玩法。无限游戏的玩家希望这场游戏一直继续下去，让后来者也能参与。作为领导

者，若想创办一家有能力参与无限游戏的组织，那么在做决策时，就绝不能以自己的收入作为唯一目标。他们的努力方向应该是让该组织具备参与无限游戏的能力。作为投资者，所投金钱的主要受益者也绝不能是自己，而是所信任的这个组织，希望帮助他们推进崇高的事业。无限思维的投资者所希望的是，为推进比自身更伟大的事业做贡献，这项事业若成功了，也能带来丰厚的回报。有限思维的投资者则更像赌徒，下注的唯一原因就是为了赚钱。我们不要混淆这两种行为。

服务导向之所以在无限游戏中如此重要，是因为它能培养一批忠诚的员工和客户（以及投资者），这些人将愿意与该组织同甘共苦、不离不弃。任何组织都能从这一强大的忠诚基石中获取长久发展的力量，这是单靠金钱无法做到的。最忠诚的员工能够感受到领导的由衷关心，因为事实也确实如此。作为回报，他们会为了公司的利益贡献自己最好的创意，自觉且负责任地工作，努力解决问题。最忠实的客户能感受到，这家公司是真的在意他们的所想、所需和所求，因为事实也确实如此。作为回报，这些忠实的客户愿意改变习惯或花更多的钱去购买这家公司而非其他公司的产品，并且会说服自己的朋友也这样做。而拥有最佳领导者的企业也能感受到，投资者是真心想要帮助他们尽可能地发展壮大，帮助推进他们所追求的崇高的事业，因为事实也确实如此。这样做的结果就是，各方都能获益。

有适应力的——能够经受住政治、技术和文化的变迁

任何希望成为具有无限思维的领导者的人都最好牢记《独立宣言》的例子。美国的开国元勋们承诺要实现平等，保障美国人民不可被剥夺的人权，这一承诺是经久不衰的。在美国独立后的200多年里，尽管这个国家的领导者、人民和文化都有过改变，但这一崇高的事业的重要性与激励作用仍然一如既往，从未减退。这是一项有着无限生命的崇高的事业。

在商业的无限游戏中，崇高的事业必须是比我们所提供的产品与服务更伟大的东西。这些产品和服务本身并不是崇高的事业，而是为推进崇高的事业服务的。如果用自己的产品来定义自己的事业，那整个企业的存在就会依附在这些产品上。任何新技术都有可能一夕之间将这些产品淘汰，我们的事业与整个企业也会随之一并淘汰。

以美国的一些铁路公司为例，这些公司曾是美国最大的企业之一。后来，随着汽车技术的进步以及高速公路网络的发展，人们有了比搭乘火车更快捷，有时也更便宜的选择。如果这些公司能将自身存在的必要性定义为人与物的运输，而非推动铁路的发展，那这些公司现在很可能已是一些大型车企或航空公司的所有者了。一些出版商也因为将自己所在的行业定义为图书出版而非思想传播，错过了利用新技术推进自身事业发展的机会。若非如此，这些出版商可能就是数字电子阅读器的发明者了。如果唱片公司能将自己定义为音乐的分享者，而非唱片、磁带和光盘的销售者，也许就能在如今这个数字流媒体的世界里活得更轻松一些。这些唱片公司若能用比自己所售产品更伟大的事业来定义自己，很可能就已是iTunes等服务的发明者了。只是很可惜，它们没有这样做，也正在为此付出代价。

市场会有起伏，人会有来去，技术会有发展，产品和服务会随着消费者取向与市场需求的变化而变化。我们需要一些持久性的东西来凝聚我们。这些东西要能经受住变化与危机的考验。为了让我们玩好无限游戏，我们的事业必须是经得起考验的、有适应力的、永恒的。

理想主义的——宏伟、大胆但最终无法实现

《独立宣言》签署者在申明“人人生而平等”且被“赋予……若干不可剥夺的权利”时，主要指的是白人、盎格鲁撒克逊人和新教徒。然而，人们几乎在看到它的同时，就对该理想有了更广泛也更包容的理解，并开始为之奋斗。在美国独立战争期间，乔治·华盛顿禁止在军队

中组织反天主教的活动，他也会定期参加天主教活动，以行动告诉他的士兵，他希望人们怎么做。在将近100年后，美国通过南北战争结束了奴隶制，随后不久，美国宪法的《第十四条修正案》赋予了非裔美国人和曾是奴隶的公民们公民的身份与平等的权利。1920年，美国的妇女选举权运动为女性赢得了投票权，这是为推进美国的崇高的事业所迈出的又一步。1964年《民权法案》和1965年《投票权法案》又为此迈出了两步，旨在保护非裔美国人及其他少数族裔美国人不受歧视。2015年，美国又迈出了重要的一步，美国最高法院就“奥伯格费尔诉霍奇斯案”（*Obergefell v. Hodges*）做出判决，将《第十四条修正案》的权利保障范围扩大到了同性婚姻。

如果美国的开国元勋们定下的目标是“赢得独立”，那这个目标一旦实现，他们可能会围坐在一起，举杯畅饮，互相吹捧他们打赢这一战有多么伟大。但事实并非如此。战争胜利后，他们就着手起草宪法（这部宪法直到独立战争正式结束的7年后才完全通过，获批生效），让他们所应遵循的原则成为经久的法条，以保护和推进他们对未来宏伟、大胆、理想主义的愿景。美国人自白纸黑字写下这个愿景以来，就一直在努力保护和推进它，也将继续保护和推进下去。美国的崇高的事业尚未完全实现，实际上也永远不可能完全实现，但美国民众会一直为之努力奋斗。这才是最重要的。

奴隶制的废除、妇女获得选举权、《民权法案》的出台都是美国为推进其事业而做出的一些重要努力。这些运动本身也是无限的，还远远算不上完成，但这标志着美国正在大步迈向《独立宣言》所勾勒的那个理想世界。庆祝胜利固然重要，但人们不能沉溺于此，止步不前。这场无限游戏还在继续，人们还有大量的工作要做。这些胜利只是迈向理想未来路途中的一个里程碑，给了人们瞥见那个理想未来的机会，从而激励他们继续前进。

崇高的事业的理想前进轨迹就是如此，无论取得了多少成就，我们

总觉得自己还有很长的路要走。如果把崇高的事业想象成冰山的话，那么我们所看到的，就只是冰山一角。在一个组织中，关于未来愿景中未知的部分，往往是创始人和早期贡献者的心里最清楚，但其他人看不到的东西。将崇高的事业阐述得越清楚，就越有可能吸引和邀请到创新者与早期参与者，也就是愿意冒险做最先推进这项事业的人，哪怕这项事业还几乎只存在于他们的想象之中。每成功一次，这座冰山都会多露出一一点，让更多人能够看到，这个愿景也才能在他人的眼中越来越清晰。当这个愿景一点点变为现实，怀疑者就会变成相信者，这种愿景完全实现的可能性才会激励更多的人，让他们甘愿献出自己的时间、精力、想法和才华，帮忙进一步推进这一事业。不过，无论展现在我们眼前的冰山有多巨大，领导者都有责任提醒我们，这座冰山的绝大部分依然沉在水下，等着我们去探索。因此，无论获得过多少次成功，我们为之奋斗的崇高的事业一直都在前方。

当你找到自己的事业时，将它记录下来

美国的开国元勋们都是些有着英雄色彩的传奇人物。他们活着就是为了推进自己的崇高的事业。商界那些能够激励人心的领导者也是如此。但是，若这些富有号召力的崇高的事业守护者停下来，比如退休或者去世了，会发生什么呢？常令我感到惊讶的是，许多富有远见卓识的领导者都认为，他们并不需要用文字将自己的事业记录下来。他们以为，自己能看清的愿景，自己组织中的其他人自然也都能看得清。但事实并非如此。

若没有用恰当的语言将崇高的事业记录下来，那么该事业随时间一点点模糊甚至完全消失的风险就会加剧。若没有崇高的事业的引领，企业就会变成一艘没有指南针的船，往往会偏离航线。船上的人会将关注点从远处的地平线转移到面前的仪表盘上。若没有崇高的事业的引领，有限思维就会开始潜入。领导者会为当前的航行速度很快，或航行距离

很远而欢呼喝彩，但意识不到自己的航程缺乏方向或目标。

记录下来的崇高的事业可以一代又一代地流传下去，创始人的直觉却不能。就像《独立宣言》一样，崇高的事业的书面声明将大大增加这一事业长久流传下去的可能性，不知多少年后，美国人或许已经不知道美国的开国元勋们都有谁了，但这一流传下来的崇高的事业依旧能够引领并激励他们前行。这也是口头合同与书面合同的区别，两者都有法律效力，都可强制执行，但写下来的合同可以避免任何因记忆模糊而对交易条款产生的混淆或分歧，这一点对那些合同确立时并不在场的人来说尤为重要。

写下来的事业就像一个指南针。有指南针在手，才能保证每一位继任的领导者都目视远方，即使创始人不在了，也能轻松地航行在技术、政治和文化规范所形成的浪潮之上，而不会迷失方向。

第3章

此事业非彼事业

让我们快速地说明一下什么叫此事业非彼事业。

似乎有越来越多的企业意识到拥有一个核心目标的重要性，这是好事。但问题是，许多人所说的只是听起来像是崇高的事业而已。他们甚至可能仔细琢磨过语言，所提出的也满足崇高的事业必备的部分标准。然而，除非符合全部五个标准，否则所提出的就不是崇高的事业。

提不出真正崇高的事业主要有以下几个原因。有时，富有远见的事业驱动型领导者也会不小心给出假的崇高的事业，因为他们找不到合适的语言来描述对未来的构想。有时，领导者其实一点对未来的愿景都没有，但想让别人相信他们是事业驱动型的。常见的“冒名顶替者”有登月计划、“成为最好的”和“以增长为目标”。有时，企业会将自己的企业社会责任（CSR）项目误以为是崇高的事业，这或许能在有限游戏中发挥一定作用，但在无限游戏中是绝对无法引导一家企业生存和发展的。

我们需要具备识别这些陷阱的能力。首先是警示，提醒我们，这些“假”的事业是无法让组织适应无限游戏的，只会让它继续使用有限思维。其次是让我们能自行判断自己是否拥有真正的崇高的事业，若没有，就要及时回头，重新构思。我们甚至可以在一开始就避免提出一个非崇高的事业。事业是“假”的不代表企业就一定是不好的，只是意味着可能会走更多弯路。无论是作为投资者、员工还是客户，具备识别事业的能力都有避害的作用。如果我们怀疑某个企业从事的不是崇高的事业，就可以去寻找另一个真正有崇高的事业的企业。

真正崇高的事业能让听到的人产生强烈共鸣，它必须与那些拥护它

的人息息相关，越能让人们感觉这项事业与自己有关，越能激发出想要帮忙推进这项事业的热情。如果你只是想用崇高的事业来提升品牌形象、吸引热情员工或实现一些短期目标，比如一次消费、一次投票或对企业的一份支持，那么，它的影响力也将是短暂的。一个组织所给出的崇高的事业究竟是空话，还是真正的信仰所在？其实只要我们加入该组织，或开始与该组织中的人打交道，就能很快辨别出来。

登月计划并非崇高的事业

“我们选择登月，”美国前总统约翰·肯尼迪坚定地说，“我们选择在这个10年完成登月……并非因为简单，而是因为难，因为这个目标能够衡量出我们最大的能力与最强的技术，因为这一挑战是我们愿意接受的，是我们不愿意推迟的，是我们想要战胜的。”¹就在肯尼迪发起这项挑战的短短8年后，宇航员尼尔·阿姆斯特朗（Neil Armstrong）迈出了“个人的一小步，人类的一大步”。

用所谓的登月计划来支撑自己观点的领导者往往只是想激励人们去做一些看似不可能完成的任务。登月计划符合崇高的事业的大多数标准，因此通常是有效的。就肯尼迪号召大家参与的登月计划而言，它是积极且具体的、有凝聚力的、以服务为导向的，也是值得为之牺牲的。然而，它却不是无限的。无论这项挑战有多艰难，看起来有多不可能实现，它都是一个可以实现的有限目标，而并不是一个理想的未来状态。人们很容易将《从优秀到卓越》（*Good to Great*）和《基业长青》（*Built to Last*）的作者吉姆·柯林斯（Jim Collins）所说的一个“宏伟、惊险、大胆的目标”（Big, Hairy, Audacious Goal，简称BHAG）误认为是崇高的事业，因为它们确实有着惊人的激励作用，往往也需要很多年才能实现。²但在完成登月计划之后，这场游戏还能继续吗？若只是再选一个宏伟、惊险、大胆的目标来替代它，并不符合无限游戏的玩法，只是开启了另一场有限的追逐而已。

在通用电气的员工大会上，一些员工常会表达出对公司过于关注短期目标的担忧。前首席执行官杰克·韦尔奇（Jack Welch）喜欢这样回答：“长期就是一系列的短期而已。”当员工表达出这样的担忧时，他们真正想问的可能是：“我们所做的这一切究竟是为了什么？”我们所有的努力除了满足业绩要求、获得物质奖励，还有什么贡献？韦尔奇的回答表明，在他看来，没有比这些更伟大的事业要追求了。目标很简单，就是执行、执行、再执行。对韦尔奇来说，能把每一个有限目标都完成就足够了。但是，商业是一场无限的游戏，这意味着一系列的短期目标会不断叠加下去，永远不会有结束。

一个又一个地完成目标确实很有趣，但若一直如此，时间一长，每次成功的快感都会递减。我时常遇见似乎“被有限目标耗得精疲力竭”的高管。他们的工作能力很强，而且每完成一个目标都能获得丰厚的报酬，因此他们会不断重复这个模式。但从某个时刻开始，他们就再也感受不到为比自己更伟大的事业做贡献的美妙感觉了，只觉得自己像一只跑滚轮的仓鼠，重复着毫无成就感却又激烈的无意义竞争。这种有限的胜利无论积累了多少，都无法带领我们走向无限。

崇高的事业必须回答的一个问题是：登月计划能有助于推进什么无限且长久的愿景呢？崇高的事业是我们追求所有其他目标的大背景，这些有限的目标有大有小，但它们的完成必须有助于推进作为其大背景的那个崇高的事业。我们若是过度关注一个有限目标，那么无论该目标有多激励人心，都容易使我们做出只有利于实现该目标，但可能会阻碍实现无限愿景的决策。

肯尼迪提出登月计划的背后，是美国开国元勋们清楚阐明的一个更宏大的无限愿景——我们的进步不是为了少数人的利益，而是为了众人的利益。肯尼迪在提出登月计划之前，也提到了这一有限目标背后的无限愿景：“面对这片未知的海洋，我们扬帆起航，因为前方有待习得的新知识和待赢得的新权利，我们必须得到它们，并将它们用于推动全人

类的进步。”他提出的许多目标都是建立在这一信念之上的，将人类送上月球并平安地返回家园也不例外。

尽管登月计划确实有激励人心的作用，但这一作用是存在有效期的。登月计划只是在无限游戏框架内提出的一个大胆的、能激励人心的目标，但并不是无限游戏本身。

成为最好的并非崇高的事业

“我们的愿景是要在每一个市场领域都成为主导者，以惊艳的设计、卓越的质量，以及最佳的价值广受客户喜爱。”³这是一个听来十分典型的企业愿景声明或者使命宣言。佳明公司（Garmin）是一家全球定位系统设备的制造商，产品可服务于从跑步者到飞行员的很多群体。尽管此类声明有众多不同的版本，但基本公式是一模一样的：我们是最好的，每个人都想要我们的产品，因为我们的产品是最好的……它们有极大的价值。

愿景声明或使命宣言能够发挥指南针一样的作用，能为我们指引方向。但因为没有任何撰写标准，上述声明已经变得太过普遍。它们太过宽泛、笼统，对想要运用无限思维的公司来说，几乎毫无价值。“成为最好的”这类声明都是以自我为中心的，将企业视作愿景中的首要主体（因此也是主要受益者），而完全无助于提升企业在客户心中的重要性。事实上，往往要到最后，这些声明中才会出现客户或该企业所能提供的价值。若将这份以自我为中心的声明置于首位，领导者就会将自己的努力集中在企业内部，而非关注那些可能购买其产品的人。人们就算购买或喜欢这些产品也不意味着就相信该企业的事业，他们甚至可能连该企业所追求的事业是什么都不知道。

有限思维的领导者常常将拥有一款成功的产品和拥有一家强大的企业混为一谈。这有点像是洛杉矶湖人队的老板因为勒布朗·詹姆斯

（LeBron James）在球迷心中很重要，就认为自己的球队在球迷心中也很重要一样。拥有一个伟大的球员、一款受欢迎的产品或一款大火的应用程序并不意味着就具备了参与无限游戏的能力。愿景声明若以产品为中心的话，一旦更好的产品出现、市场环境发生改变、新技术诞生后，也就毫无用处可言了。而且过时的愿景声明往往会误导企业固守旧的商业模式，对本可以抓住的机遇视而不见。这似乎就是佳明公司曾经的遭遇。

佳明在2007年或许是“最好的”。就车载与船只仪表板内安装的全球定位系统设备而言，该公司确实是全球领军者。然而，随着智能手机的功能越来越强大、可靠，我们不再需要单独的全球定位系统设备了，佳明也因此而受到了重创，目前的市值已不足其2007年的1/3。对佳明来说，要将自己的损失简单归咎于智能手机的崛起和普及真是再容易不过了（他们也确实是这样做的）。他们没有意识到问题出在自己的愿景声明上，这一愿景令他们将关注点放在了产品上，从而错失了智能手机所提供的机遇。如果他们能始终以为顾客创造价值为重、为先，那么当机遇尚在时，他们或许已经抓住并研发出了人们会优先选用的手机导航应用程序了。曾经，他们的品牌确实强大到足以做到这一点。然而，事实正好相反，他们依旧执着于自己旧有的商业模式，即销售安装在仪表板内的全球定位系统硬件。崇高的事业应该是商业模式的指挥者，而不是被商业模式所指挥。

若愿景声明或使命宣言以产品为基础，那么它们对该企业的文化就会产生负面影响。在产品至上的企业中，除工程师或产品设计师以外的员工都会感觉自己是公司的“二等公民”（有时他们受到的对待也确实如此）。这种情况在技术企业或工程企业中相当普遍。企业若要更充分地调动员工积极性，就应避免让会计、支持、客服等岗位上的任何一个人感觉自己的存在就只是为了满足工程师或产品开发团队的需求。这些岗位上的人也希望感觉到自己在团队中的价值，希望能够与大家一起携手推进比产品或他们自己更伟大的事业。

“成为最好的”完全不可能成为崇高的事业，即使我们现在是最好的（基于我们自己选择的衡量标准和时间段），这个地位也只是暂时的。即使成为最好的，游戏也不会结束。正因为游戏仍将继续，我们常常会因为太在意、太想保住这一排名而开始一味防守。如果你只是想为组建一支团队而要准备一场振奋人心的演讲，那么“我们是最好的”会是不错的演讲素材，但若是要将它作为整个企业的创立之基，就太不堪一击了。无限思维的领导者明白“最好”不是一种永恒的状态。因此，他们会努力做到“更好”。“更好”意味着这将是一段持续改进的旅程，也能让我们觉得，它是在邀请我们为它的进步贡献才华与精力。在无限游戏中，“更好”优于“最好”。

以增长为目标并非崇高的事业

想象一下，某天早上，你走出家门，看到邻居正在往车上搬行李。你问他：“你要去哪儿？”他说：“去度假。”你好奇地追问：“真棒。你要去哪儿？”他再次回答：“告诉你了呀，去度假。”你说：“我知道呀，但你是要去哪儿呢？”结果邻居被你问恼了：“都跟你说了，去！度假！”⁴

当你意识到自己的问法不对，得不到想要的答案时，就换了一种策略。“那你是打算去哪儿度假呢？”这下，邻居立刻就懂了，说出了他的度假计划：“我打算去加拿大自驾游，每天开400千米。”

若你的问题是：“贵公司的事业是什么？存在的理由是什么？”往往就会得到“增长”这个答案。这就像是问“你要去哪儿”，邻居只会告诉你“去度假”这个答案一样。在这些以增长为目标的企业中，领导者可以脱口而出他们的增长战略与目标，但这就像是在解释你度假时打算走哪条高速公路，以及每天走多少公里一样，从中听不出你最初出发的原因，以及你希望抵达的目的地。增长的战略和目标无法提供更大的背景或目的。

金钱是推动事业发展的燃料，但并非事业本身，增长是为了让我们拥有更多推动事业发展的燃料。正如我们买车的目的不是为了买更多的汽油，企业必须具备比赚钱能力更大的价值。一家企业就像一辆汽车，对于选择了它的人来说，只有当它能送我们去的目的地是没有它我们就无法抵达的地方，它才具有更大的价值。所设想的目的地就是崇高的事业。

值得一提的是，企业所提出的目标有太多都很想当然或不切实际，尤其在初创企业这个圈子里。对10亿美元估值的追求并不是一家公司能够健康、持久存续的标志，这只是风险投资行业催生的一个标准（因为估值是该行业赚钱的方式）。⁵企业要想不被淘汰，长期留在这场游戏中，依靠的其实是强大的文化和赚钱支撑自己生存的能力（盈利能力）。

此外，在成熟市场上，持续追求过度增长是会带来问题的。所谓成熟市场就是指，产品、技术或业务不再是新的或特别的，而是已被广为接受且随处可见的。西尔斯百货和通用电气所处的就是成熟市场，对这类企业来说，如果秉持的是“不惜一切代价实现增长”的心态，就无法给市场提供具有吸引力的选项。⁶许多企业会开始采取防御姿态，把钱交给股东以讨好他们，或过度使用股票回购这一方式人为地抬高股价。对于有限思维的成熟企业来说，以收购或合并来实现增长往往会成为其继续保持高增长率的唯一方式。这或许能短期内提升股价，但正如《哈佛商业评论》等许多媒体曾报道过的：“70%到90%的收购都是失败的。”⁷

若将增长视为所追求的事业，为增长而增长，就像为增肥而吃东西一样。这会迫使高管们考虑有助于保持增长的战略，但甚少或压根不考虑增长的目的。就像为增肥而吃东西的人一样，这样的企业最终也会遭遇“健康”问题。以增长为事业往往会催生不健康的文化。在这种文化中短视主义和自我中心将占据主导地位，信任和合作将随之减少。增长是结果，而非原因，这并不是所追求的事业，是一种产出，而非企业存在

的理由。若我们以崇高的事业为目标，便会愿意为推进它而不惜牺牲自己的利益。若我们将金钱或增长视为所追求的事业，则有可能为保护自己的利益而不惜牺牲他人，甚至是这项事业本身。此外，没有什么是可以永远增长下去的。所有的气球和气泡最终都会破裂，金融也不例外。

企业社会责任并非崇高的事业

有一家企业，他们会将自己为社区所做的一切好事都宣传出去，比如会分享一些奖学金受益者的故事。他们希望让自己的客户和员工知道，自己是具有人文关怀的。这本来是件很好的事，但前提是，他们自己的6万名员工不用工作在一种“头重脚轻”“狗咬狗”的有害文化中。

企业社会责任项目并非崇高的事业。企业并不会因为赞助马拉松、捐款慈善机构或让员工带薪休假做志愿者而变成事业驱动型的，也不会因为把产品送给买不起的人而变成事业驱动型的。

企业社会责任项目几乎可算作是为慈善事业捐款的商业用语。尽管有此类项目是伟大且值得赞扬的，但除非本就是慈善机构，否则这也只是企业做的许多事情之一。企业社会责任项目必须纳入一个更广泛的战略中，该战略要以推进崇高的事业为目标，涵盖企业所做的方方面面。企业赚钱的方式以及捐赠的方式都必须有助于推进崇高的事业。“公益事业相关的工作”并不是一个企业的兼职，而是其存在的核心。服务不是为了装点门面，而是一块试金石。再多的企业社会责任项目也不足以抵消或平衡因过度关注有限目标而对企业文化其余部分的消耗。

一个有限思维的领导者，即使充满善意，往往也认为要“先赚钱，再来做好事”。然而，无限思维对服务的看法有些许不同：“做好事来赚钱”（先后顺序很重要）。我会终己一生善待他人，用心服务社区，与此同时，我仍然能够建立一个财力雄厚的企业。这与其说是在平衡二者，不如说是一种生活方式。这些个人和企业是在努力成为服务者，服

务为其工作的人，服务他们所在的社区。这种给予是贯穿其一生，直至其生命终结的，因此更像几十年如一日的坚持，而非只为平衡过去。这种差异是由领导者的思维方式决定的。

第4章

事业的守护者

1962年，山姆·沃尔顿（Sam Walton）创办了沃尔玛，他当时的想法很简单：通过“随时随地”提供“最优惠的价格”来为美国的普通劳动者服务。临终时，沃尔顿曾这样描述他的愿景：“如果我们携手努力，就能降低每个人的生活成本……我们将让这个世界有机会看到花更少的钱拥有更好的生活是什么样子。”¹在沃尔顿的领导下，打造沃尔玛的一切决策都是以该事业为目标制订的，从超市选址到超市规模，均不例外。人们爱上了沃尔玛，无论是沃尔玛的员工，还是店内的购物者，大家都希望沃尔玛的超市能开到自己的社区。沃尔玛的生意越做越大，成长于美国经济大萧条时期的沃尔顿也成了美国最富有的人之一。

遗憾的是，在后续的发展过程中，这个崇高的事业变得模糊了。2009年，迈克·杜克（Mike Duke）接任沃尔玛首席执行官，自那时起，该事业显然就不再是推动该沃尔玛发展的那股驱动力了。沃尔顿最初的愿景几乎已经沦为写在办公室墙上的营销口号与空洞文字。这家企业为利润、增长和业已拥有的主导地位所迷，不惜牺牲最初成就了它的那项事业。

在沃尔玛，迈克·杜克是响当当的效率专家。李·斯科特（Lee Scott）在宣布由杜克接替自己，担任公司首席执行官时，结结巴巴地说道：“我曾认为，我认为董事会也曾认为，公司本可以被管理得更好。”他接着解释道：“迈克不仅是位优秀的领导者，也是位真正优秀的管理者……我认为，在商界，你不能忘记这个事实：你要做的不只有领导，还有管理。”董事会的目的如果是纠正管理问题或提高业绩，那么选杜克这样的人来执掌公司可能是一个完美的选择——仅就短期而言。

但若是担心山姆·沃尔顿的崇高的事业被模糊掉，希望有人能让公司重回正轨，那么像杜克这样的人恐怕就是糟糕的人选了。

杜克在继任这一位置时说的一番话充分暴露了他将以何种思维领导沃尔玛。“沃尔玛在当前经济中的地位非常稳固，市场份额和利润都在不断增长，在顾客心中也比以往任何时候都更重要。”他在宣布就任的新闻发布会上说：“我们的战略是合理的，我们的管理团队能力卓越。我有信心，我们能继续为股东创造价值，为超过200万的同事提供更多机会，并帮助全球1.8亿顾客节省开支，过上更好的生活。”²

注意到他话中的先后顺序了吗？杜克首先想到的是提高市场份额和利润。虽然他提及了企业在顾客心中的重要性，但直到发言快结束时，他才提到了为顾客创造价值。这是人类天生的奇怪癖好。话语中信息点的先后顺序往往会暴露出一个人心中真实的优先排序及其战略的侧重点。山姆·沃尔顿是从人的利益出发的，杜克是从股东的利益出发的。

在杜克的领导下，沃尔玛的股价确实上升了一段时间。然而，看重数字胜于人的做法是要付出代价的。这个一度深受人们喜爱的品牌也卷入了多起有关员工和顾客的丑闻。2011年，沃尔玛成为被告，遭遇了史上最大规模的就业歧视集体诉讼案之一。原告是该公司的女性员工，她们控诉公司故意压低她们的薪酬，并故意不让她们升职。2012年，该公司的员工又以罢工和抗议的形式，要求得到尊严和尊重，以及一份可以维持生计的薪水。曾几何时，居民集会的目的是让沃尔玛门店进驻自己的社区，如今却正好相反。沃尔玛在丹佛市和纽约市的扩张计划就是因大规模的抗议活动而搁浅的。沃尔玛还曾因贿赂外国官员而遭到美国国会调查。³公司员工的士气一落千丈，人们对沃尔玛的喜爱也大都变成了蔑视。

在沃尔玛身上发生的事是上市公司常常会有有的遭遇，事业驱动型企业也不例外。⁴公司常常会迫于华尔街的压力而将有限思维的高管放到最高的领导岗位上，但我们需要的其实是一个富有远见的无限思维的领

导者。正如我们前面讨论过的，史蒂夫·鲍尔默就是一个典型反例。

1983年，约翰·斯卡利（John Sculley）接替史蒂夫·乔布斯出任苹果首席执行官。斯卡利没有继续推进乔布斯所提出的事业，而是更关注与IBM的正面竞争，他对企业文化的破坏重创了苹果的创新能力和文化。2000年，罗伯特·纳德利（Robert Nardelli）在竞聘通用电气首席执行官失败后，接掌了家得宝公司（Home Depot），他对削减成本执着不懈的追求几乎摧毁了家得宝的创新文化。2004年，戴尔的首席运营官凯文·罗林斯（Kevin Rollins）接替迈克尔·戴尔（Michael Dell）成为首席执行官，罗林斯专注于增长，主持了该公司有史以来最大规模的裁员，见证了客户投诉的增长和美国证券交易委员会（SEC）对其审计问题的调查。这些人都是资深高管，然而，他们的有限思维令他们难以胜任被托付的那个位置。事实上，斯卡利对苹果、罗林斯对戴尔的伤害之大，迫使两家公司不得不为了消除他们制造的混乱，分别召回了比他们更具无限思维的前任——史蒂夫·乔布斯和迈克尔·戴尔。在选择由谁担任首席执行官时，问题不是这个候选高管的工作技能是否熟练，而是他们是否具备匹配该职位的正确思维方式。

首席愿景官

每位“首”字头主管的职责都体现在他们的头衔中：首席财务官、首席营销官、首席技术官、首席运营官。他们的职责是什么、需要监督什么，都写在他们的头衔里。头衔的作用之一就是确保我们把对的人放在对的位置上。基本没有人会考虑让讨厌数字、无法理解资产负债表的人担任首席财务官。如果你不懂技术，家里的电视机上还连着过时的磁带录像机，那么短时间内你应该是不会出现在任何首席技术官的候选名单上的。那么，问题来了，首席执行官到底是做什么的？

如今，太多企业领导者都在至少应该考虑一下无限游戏时，仍然固守有限游戏的玩法，而组织中缺乏界定首席执行官角色与职责的清晰标

准正是造成这一现象的原因之一。很多时候，领导者之所以没能做好胜任该职位的准备，是被头衔误导了。“执行”一词并不能告诉我们首席执行官的职责到底是什么。

用词很重要，会给事物指明方向、赋予意义。选错了词，意图就会改变，事情的发展就不一定会如你所愿。以马丁·路德·金为例，他发表的演讲是“我有一个梦想”而非“我有一个计划”。他需要有计划，这一点毫无疑问。我们也知道他开过多次会议，探讨如何制订这个计划。但作为民权运动的“首席执行官”，他的职责并不是制订这项计划，而是对这个梦想负责，确保那些负责执行计划的人能够为实现该梦想而努力。

2018年，洛里·罗宾逊（Lori Robinson）上将从空军退役，她是当时美国军方历史上级别最高的女军官。她对组织中最高职位者的职责是这么解释的：眼光不能局限于该组织，要看得更高、更远。每次接到新命令，她都会这样描述自己的职责：“我会看得更高、更远，也需要你们向内深入。”如果最高管理者需要专注于“更高、更远”的东西，我们就需要修改他们的头衔，从而更好地体现他们的主要职责。

在无限游戏中，若领导者知道自己即将担任的角色是“首席愿景官”（Chief Vision Officer，简称CVO），就更能有针对性地做好履职所需的准备。他们犹如坐在长矛最锋利的尖端，对他们来说，首席愿景才是他们的主要职责。他们是愿景的持有者、传播者和守护者。他们的工作就是确保所有人都能清楚理解崇高的事业，并确保公司内所有其他“首”字头的主管都能将工夫花在推进该事业上。这并不是说，无限思维的领导者就完全不用关心企业的有限利益。只是，作为崇高的事业的守护者，他们有责任做出判断，什么时候值得为推进无限愿景而承担一些短期的有限成本，需要考虑的远不止净收益。作为最重要的无限玩家，首席愿景官必须看得更高、更远。

最高职位的继任者

如今，有太多企业是按照单一的等级制度来组织的。首席执行官最大，首席财务官或首席运营官往往被视为二把手。在绝大部分企业中，首席财务官或首席运营官往往会将自己视为“最高职位”的继任者。曾在杰克·韦尔奇手下工作了17年的迈克尔·丁金斯（Michael Dinkins）解释说：

我认为，之所以有许多首席财务官会被提拔为首席执行官，原因之一在于，首席财务官是企业中为数不多的能够看到企业全貌的人之一。企业里发生的一切.....他们了解企业内部的所有流程以及这些流程发生的时间段.....他们了解人力部门如何招聘.....了解制造厂如何引进新设备.....了解行业的质量控制体系.....他们了解整个企业，这就是首席财务官的一大优势。⁵

如果我们寻找的是有限思维的战术型领导者，那么丁金斯的说法就很有道理。不过，这种说法并不适用于寻找首席愿景官。首席愿景官不是负责运营的，也不是负责财务的。首席愿景官的关注点应该更高、更远，首席财务官和首席运营官要做的才是深入关注企业内部。前者必须有无限的眼界，后者则应着眼于商业计划。前者构想的是非常遥远且抽象的未来，后者看到的则是短期内所应采取的切实步骤。

这就是一流企业往往是由多人同时管理的原因之一。愿景守护者（首席愿景官）与具体操作者（首席财务官或首席运营官）的组合，是一种技能互补的搭档关系。如果我们能调整一下公司内的正式等级制度，令其有助于培养与职责相称的思维方式，就更有可能打造出这样的搭档关系。这意味着，我们不能再把首席执行官视为第一，把首席财务官或首席运营官视为第二，而是要把他们视为追求共同事业的重要搭档。任何一方都不知道如何更优秀地完成对方的工作，这正是他们彼此需要的原因。别忘了，史蒂夫·鲍尔默、约翰·斯卡利和凯文·罗林斯也曾有过十分卓越的表现，而那时，他们身边都有着一位更具无限思维的搭

档。

虽然首席愿景官更常出现在聚光灯下，得到的赞美往往也更多（至少在公开场合是如此），但双方都必须清醒地认识到，他们是搭档，需要相互信任。首席愿景官知道光靠自己是无法推动愿景实现的，他们的身边需要有像迈克尔·丁金斯所描述的那种人。首席运营官或首席财务官知道，若将自己的专业技能用来协助推进崇高的事业，这些技能就能发挥更大的作用，拥有更大的意义，那是比他们自己、比公司都更伟大的事业。当这样的搭档关系奏效时，首席愿景官与首席运营官或首席财务官都会将更多的时间花在相互帮助上，而不是相互竞争谁更受关注。

令许多首席财务官或首席运营官感觉不安的一个真相是，他们其实已经走到了自己技能水平的巅峰。他们已经是企业中最资深、最熟练的财务或运营人员了，这很棒。没有他们，首席愿景官是无法推进愿景实现的。但这并不意味着他们有能力走到最前面去，成为这一愿景的引领者。在得到这一“最高职位”后，许多人都会继续沿用自己熟知且擅长的思维方式——思考让公司发展多大规模，思考让公司获得何种形式的利润。他们并不会拥抱自己的新职责，即考虑未来会出现何种变化，考虑公司该如何继续推进崇高的事业。

这就像是一名销售员被提拔成了销售经理。他可能很擅长销售，但他现在的职责不再是销售了，而是管理那些做销售的人。他如果没能成功换挡，没能及时调整思维模式，没能学会履行新职责所需的技能，问题就会接踵而来。无论是首席财务官、首席运营官，还是其他高管，若能学会适应自己的新角色、新职责，并拥有无限思维，那么无疑也能成为成功的首席愿景官。不过，若没能做到这些，他们很可能就会沿用上一份工作的那些技能，这会增加他们将公司引入一条有限道路的可能性。

无论迈克·杜克是否有成为沃尔玛首席愿景官的资格，他没能适应自己的新角色都是不争的事实，他没能将山姆·沃尔顿的愿景守护到下

一个世纪。相比之下，杜克的继任者道格·麦克米伦（Doug McMillon）可能正是沃尔玛所需的那种首席愿景官。2013年，道格·麦克米伦宣布就职，他在一份新闻稿中说：“获得领导沃尔玛的机会于我是莫大的殊荣。我们公司在为全球顾客创造价值方面有着悠久的历史，无论顾客的需求如何增长、变化，我们都将随时随地为他们服务。我们的管理团队才华横溢、经验丰富，我们的战略也令我信心十足，我们一定会拥有一个光明的未来。我们会信守对顾客的承诺，从而为股东创造价值，为员工创造机会，拉动业务的增长。”⁶道格·麦克米伦的这番话给出了他心中的优先级，与5年前迈克·杜克上任时的优先级完全相反。麦克米伦把山姆·沃尔顿的愿景放在首位。这真是令人兴奋啊，让我们期待着，在他的领导下，重返这场无限游戏的沃尔玛将会有怎样的表现。

第5章 企业责任

如今的商业世界瞬息万变，而所有的变化似乎都有代价。企业被商业游戏淘汰的速度越来越快。20世纪50年代，一家公司的平均寿命是60年多一点，如今却已不到20年。瑞信（Credit Suisse）2017年的一项研究显示，颠覆性技术是导致企业寿命急剧缩短的原因。然而，颠覆性技术的出现并不是现在才有的，信用卡、微波炉、气泡膜、尼龙搭扣、晶体管收音机、电视机、计算机硬盘、太阳能电池、光纤、塑料和微芯片等都是20世纪50年代的发明。“颠覆”可能并不是挑战存在的成因，只是一个隐蔽的真正成因的表象。技术不是失败的理由，从本质上讲，问题的主因并不在技术，而在于领导者。他们未能预见在不断变化的环境中，自己的企业会有怎样的未来。失败是目光短浅的结果，目光短浅是有限思维的领导者的一个内在特征。事实上，这种目光短浅的现象在过去50年的兴起可以追溯到一个人的思想上。

诺贝尔经济学奖得主米尔顿·弗里德曼（Milton Friedman）被认为是当今伟大的资本主义形态理论家之一。他于1970年发表的一篇文章为股东至上理论奠定了基础，而股东至上理论正是如今大量有限思维商业做法的核心所在。弗里德曼认为：“在自由企业、私有财产这一制度下，企业高管只是企业所有者的一名雇员而已，对雇主负有直接责任。该责任要求他们遵照雇主的意愿来经营企业，而雇主所想要的一般都是，在遵守社会基本规则、不违法、不违背道德习惯的同时，赚到尽可能多的钱。”¹事实上，弗里德曼坚称“企业有且只有一个社会责任，那就是在不违背游戏规则的前提下，利用自己的资源，从事旨在提高利润的活动”。²换言之，弗里德曼认为，企业唯一的目的是赚钱，赚到的钱属

于股东。这些思想已牢牢扎根于如今的时代潮流之中，并被广为接受。因而，企业“所有者”走到了这一利益食物链的顶端，而他们的企业只有一个存在的目的——创造财富。我们往往认为，从过去到现在，商业游戏有且只有这一种玩法。只可惜，事实并非如此，过去不是，现在也不是。

弗里德曼对商业的看法似乎很单一。任何曾经领导过企业、为企业工作过或从企业购买过产品的人都知道，企业是动态的、复杂的。这意味着，在过去的40多年中，我们打造企业所依据的定义对企业很可能是不利的，而且会破坏其号称拥护的资本体系。

弗里德曼之前的资本主义

对企业责任的定义，可以追溯到亚当·斯密更具无限思维的定义。亚当·斯密是18世纪的苏格兰哲学家与经济学家，是公认的经济学和现代资本主义之父。他在《国富论》中写道：“消费是一切生产的唯一终点和唯一目的，生产者的利益只有在成为促进消费者利益的要素时，才应当加以关注。”³他进一步解释道：“这一点完全是不言自明的，试图证明它才是荒谬的。”简言之，企业的利益应该永远排在消费者的利益之后。讽刺的是，我正努力用一整本书来证明的事情，在亚当斯密看来“不言自明”，试图证明它才是荒谬的。

不过，亚当·斯密并没有忽视我们的有限偏好。他承认：“重商制度几乎总是在为了生产者的利益而牺牲消费者的利益，该制度似乎认为一切工业和商业的终点与最终目的都是生产而非消费。”简言之，他承认自利是人的本性。他将这一自利倾向称为“看不见的手”。这只看不见的手是一个普遍真理（出于自私的动机，我们都会想要建立强大的企业），因此，最终还是会有利于消费者的。他解释道：“我们不能将自己的晚餐寄望于屠夫、酿酒师或面包师的善意，而是要寄望于他们对自己利益的关心。”屠夫是为了自己的利益才想要提供最优质的肉，他并

不关心酿酒师和面包师的状况。酿酒师想要酿造出最好的啤酒时，也并不关心市场上正在销售什么样的肉和面包。想要做出最美味面包的面包师也不会考虑我们会在自己的三明治中放些什么。亚当·斯密相信，最终的结果是，消费者会得到最好的一切，至少在该体系平衡时确是如此。不过，他从未考虑到会有这样一天，该体系会因外部投资者与分析师群体的自私自利而完全失衡，他没有预料到一群自私自利的局外人能对面包师施加如此庞大的压力，会为了让投资者的收益最大化而削减成本，使用更廉价的原料。

如果你对历史或对18世纪的哲学家不感兴趣，那我们就来看看资本主义在股东至上理论成为主流后又发生了哪些变化。股东至上理论成为主流仅仅是20世纪最后几十年才发生的事。在该理论出现之前，美国的商业运营方式与现在截然不同。美国康奈尔大学法学教授琳恩·斯托特（Lynn Stout）曾经说过：“在20世纪中叶以前，美国上市公司都在证明自己是世界上最有效、最有影响力，也是最赚钱的企业组织形式之一。”那个时代的公司不仅仅能让最富有的人参与投资，也能让普通的美国人参与投资，并享受到不错的收益。最重要的是，“高管和董事都将自己视为大型上市企业的管家或受托人，并认为这些企业不仅应该服务于股东，也应该服务于债券持有者、供应商、员工和社会”。

直到1970年弗里德曼的那篇文章发表，高管和董事才开始认为自己应对企业的“所有者”，即股东负责，而不是服务于某个更伟大的目标。这一观念在20世纪八九十年代越来越盛行，上市公司与银行内部的激励机制也越来越过度地关注短期收益，这些收益也越来越向着少数人集中。也是在此期间，为了实现主观制定的目标，每年一轮的大规模裁员首次成为广为接受的战略之一。而在20世纪80年代以前，这种做法根本就不存在。当时，在一家公司工作一辈子是很平常的事。公司会照顾员工，员工也会照顾公司，信任、自豪和忠诚都是双向的。而现在，我们往往会在远不到退休的年纪时就主动或被动地离开公司。

资本主义的滥用

今天这种有限思维的资本主义形态与曾经无限思维的资本主义形态几乎毫无相似之处，而后者曾启发了美国的开国元勋（托马斯·杰斐逊拥有全套三卷的《国富论》）。今天的资本主义与200多年前亚当·斯密所构想的资本主义截然不同，只是名义上的资本主义。它与福特、柯达、西尔斯百货等公司在19世纪末、20世纪初所践行的资本主义毫无相似之处，然而，在此之后，这些公司也都陷入了有限思维的陷阱，迷失了方向。如今许多商界领袖更像是在滥用资本主义，或者说“资本主义的滥用”（capitalism abuse）。“滥用”一词的定义是对某物的不当使用，偏离了其初衷。如果说资本主义的初衷是让消费者受益，让企业领导者服务于比他们自身更伟大的事业，那么今天就偏离了这一初衷。

我认为，企业的目的不只有赚钱，还有追求崇高的事业。有些人可能会说我的观点是幼稚且反资本主义的。首先，我要郑重提醒大家小心说这些话的人。我认为，那些对弗里德曼的企业观，以及由他所催生的许多现行企业的做法的狂热捍卫者，正是其中最大的受益者。但企业从来就不是只为赚钱而存在的。正如亨利·福特（Henry Ford）所言：“一个只会赚钱的企业是一个糟糕的企业。”⁴企业的存在是为了推动某些东西的发展——技术、生活质量，以及任何可能以某种方式、形式或形态改善生活的东西。人们若愿意付钱购买一家企业的东西，说明人们能够从这些东西中感知到或获取到某种价值。这意味着，一家企业提供的价值越高，能赚到的钱就越多，就有越多“燃料”支撑其进一步发展。资本主义不仅关乎繁荣（以特点、利益、金钱来衡量），也关乎进步（以生活质量、科技进步，以及人类彼此和平共处、携手合作的能力来衡量）。

自20世纪70年代末以来的不断滥用，给我们留下了一个破败不堪的资本主义。这是已经退化的资本主义，它服务于为一己之私滥用该制度的少数人，只为了让这些人获得更大的利益，这对发挥资本主义哲学的真

正优势毫无帮助（遍及全球的反资本主义运动与保护主义运动就是明证）。事实上，整个股东至上理论，以及弗里德曼对企业责任的定义，都是由投资者自己提出来的，只是投资者的工具，目的是激励高管把保护投资者的有限利益放在首位。

比如企业开始将高管的薪酬与公司的短期股价表现而非长期健康挂钩，这一做法主要就是由弗里德曼的观点所催生的，那些支持弗里德曼观点的人从中获益颇丰。美国经济政策研究所（Economic Policy Institute）的一份报告称，1978年时，首席执行官的平均收入大约是工人工资的30倍。到2016年，他们的平均收入已上涨了超过800%，是工人平均收入的271倍。如今，他们的收入平均增长了近950%，而美国工人的收入仅增长了11%。同一份报告显示，首席执行官平均收入的增长速度比股票市场的增长速度快了70%！

你并不需要具备工商管理学硕士学位也能明白这是为什么。正如琳恩·斯托特在其著作《股东价值迷思》（*The Shareholder Value Myth*）中的解释：“如果对首席执行官来说，他们80%的薪酬都取决于公司股价在未来一年中的表现，那么他们就会竭尽全力确保股价上升，哪怕这样做会损害员工、客户、社会、环境，甚至该公司本身的长远利益。”若将薪酬与股价直接挂钩，就会鼓励首席执行官采取关闭工厂、压低工资、极端削减成本、大裁员等策略，这些策略可能会在短期内提升股价，但往往会损害企业在无限游戏中长期生存与繁荣的能力。为了抬高股价，还有另一种被上市公司高管用滥了的做法，就是股票回购，该做法在大多数情况下是合法的。高管会基于供求规律回购自己公司的股票，从而暂时提高市场对其股票的需求，进而暂时抬高股价（让他们的业绩表现暂时看上去是不错的）。

尽管许多用以短期抬高股价的做法似乎都有违道德，但我们若再重新审视一下弗里德曼对企业责任的定义，就会发现他为这类做法敞开了大门，甚至予以鼓励。别忘了，他对企业履职的唯一指导就是，一切行

为都必须合法且符合“道德习惯”的。作为旁观者，我被“道德习惯”这一古怪用词所震撼。为什么不直接说“道德准则”呢？“道德习惯”是否意味着，某件事，只要我们做的次数足够多，就会正常化，彻底变成道德的了？如果有非常多的企业都会经常性地大裁员，用牺牲员工生计的方式满足随意预设的目标，那么是不是意味着这一策略就不再是不道德的了？是否如果人人都这么做，那这件事就一定正确的？

事实上，法律与“道德习惯”的产生往往不是因为对可能出现滥用行为的预测，而是在滥用行为出现后的应对。换言之，法律与“道德习惯”总是落后的。基于对弗里德曼定义的一般解读，企业几乎被要求要利用这一时间差来实现利润最大化，直到未来某一天，有法律和道德告诉他们不能这样做。基于弗里德曼的观点，这么做是企业的责任！

Facebook、Twitter、谷歌等科技公司显然更愿意在违反了道德习惯后再来请求原谅，而非在领导公司时就从根本上考虑到该如何保护它们最重要的资产之一：我们的私人数据。根据弗里德曼的标准来评判，这些公司只是在做他们应该做的事。

如果我们在创办企业时所遵照的企业定义就有缺陷，那么，我们未来所提拔的人和所组建的领导团队，很可能都是最符合弗里德曼所信奉的规则，即有限游戏的规则的。然而，这样的领导团队可能也是最不会好好遵守道德要求的，即做不到不利用体系漏洞来牟取私利。建立团队时若带着错误的目标，那么所建成的团队也更有可能做出一些不好的决策，损害他们本应领导和保护的企业、人和社会的长远利益。正如1757年法国国王路易十五所说：“我死之后，哪管洪水滔天。”换言之，在我死后才到来的灾难都是你的问题，而不是我的问题。这似乎已是如今大量有限思维的领导者的共识了。

迫使你运用有限思维的压力

绝大多数上市公司高管都知道一个公开的大秘密，即股东至上理论和华尔街施加给他们的压力其实都是对企业不利的。但愚蠢的是，尽管他们知道这个秘密，私下里也会抱怨，也会有疑虑，但仍会继续捍卫这个理论，仍会继续屈从于这些压力。

我不会将宝贵的笔墨浪费在长篇大论“当高管们屈从于这些压力时，会给国家和全球经济造成什么样的长期影响”。我只需要提醒大家留意几点就够了：2008年那场人为造成的大萧条，有太多人都因工作而承受着与日俱增的压力和不安，以及太多领导者在乎自己胜过我们。这真是一种莫大的讽刺，有限思维捍卫者的所作所为恰恰将他们期望从中获利的那家公司置于了险境，仿佛他们已经确定了能得到最多樱桃的策略就是砍倒那棵樱桃树。

在很大程度上，由于最初为防止银行产生1929年大萧条那样的影响和投机倾向而引入的监管有所放松，投资银行重掌巨大的权力与影响力。⁵结果显而易见，华尔街能迫使企业去做他们不应该做的事，也能阻止企业去做他们应该做的事。

企业家也不能幸免于压力的影响。他们所承受的巨大压力往往是要让企业表现出持续、高速的增长。为了实现这一目标，或者当增长放缓时，他们就会向风投公司或私募股权投资公司来集资。这在理论上听起来不错。不过，私募股权投资公司的商业模式存在一个缺陷，可能会对任何渴望继续游戏的公司造成严重破坏。私募股权投资公司和风投公司若想赚钱，就只有出售。而出售往往发生在他们投入第一笔钱的三到五年后。私募股权投资公司或风险投资家只要愿意，就可以说出任何符合无限游戏规则的、以事业为中心的漂亮话来。或许在他们决定出售的那一天到来前，他们都是相信自己所说的这些话的。只是到了那一天，许多人对崇高的事业和所有其他利益相关者的关心就会突然锐减。投资者会以要实现有限目标的名义给企业施加压力，这些压力可能会，也往往会对该企业的长远发展造成毁灭性的伤害。有很多目标导向型的高管直

到出售发生之前都认为他们的投资者是不同的。

没有什么是可以一直增长下去的，也没有任何规则告诉我们，高速增长是企业长久存续的必备战略。快速增长在有限思维的领导者眼中是目标，但在无限思维的领导者眼中只是一个可调节的变量。有时，降低增长速度也是一种重要战略，这种战略有助于确保企业的长期安全，有助于确保该企业能够妥善抵御高速增长带来的额外压力。比如，一家快速发展的零售企业可能会选择放慢门店扩张速度，以便将更多资源用于培训与培养员工和门店经理。能让一家公司成功的并不是开新店，而是让已经开了的店能够运营良好。真正对公司有益的做法是，现在就把事情做好，而不是等着解决未来会因高速增长而产生的问题。优秀的领导者要有能看清增长计划的长远眼光，愿意在尚未完全准备好或时机不对时谨慎行事，哪怕这意味着要放缓增速。

在20世纪50到70年代之间，多个机构都认为“预测”这一概念至关重要。他们请来“未来学家”团队，研究技术、政治和文化的发展趋势，以便做出预测，提前做好准备。佳明公司若采用了这一做法，或许就能主动适应移动电话技术的进步，而不是被迫做出应对。就连美国政府也这么做过。1972年，美国国会为了研究拟立法律的长期影响，专门设立了技术评估办公室。世界未来学会（World Future Society）主席爱德华·科尼什（Edward Cornish）说：“他们开始意识到，法律一经出台，通常要在实施了20年或50年后才会重新审议，因此，他们希望确保现在制定的法律不会对接下来的这几十年产生负面影响。”然而，到20世纪80年代，这门学科失宠了，一些政府官员认为试图“预测未来”是在浪费钱。技术评估办公室于1995年正式关闭。尽管今天的商界仍然存在未来学家，但企业让他们预测的趋势通常都是用来营销的，而非用来评估当前选择对未来的影响。

短期收益才是市场最容易看到的，因此，集中于有限目标的领导者往往不愿牺牲短期收益，哪怕是有利于未来发展的正确选择。这种心态

会给公司内的其他人带来压力，迫使他们也集中于短期目标，而这些目标往往会导致所购买的服务或产品质量下降。这就与亚当·斯密所说的背道而驰了。如果投资界遵从亚当·斯密的哲学，就应该不遗余力地帮助企业生产出最优质的产品、提供最好的服务、建立最强大的企业。这才是对消费者和国家财富都有益的做法。股东若真是他们所投公司的所有者，那他们确实会这样做。但他们的做法压根就不像主人，而更像租客。⁶

试想一下我们驾驶自己的车与驾驶租来的车会有何不同，这样你应该就能立刻明白股东为什么更关心能不能抵达目的地，却几乎毫不关心用于前往目的地的工具了。无论哪一天，只要打开美国的消费者新闻与商业频道（CNBC），我们都能看到集中于交易策略与近期市场走势的讨论。这些节目的主题是交易，而非拥有。他们教给观众的是如何购买与抛售房屋，而非如何找到一个可以安居乐业的家。如果短线投资者对待所投资的企业就像对待租来的车而非自己的车一样，那么，这些企业的领导者又为什么必须像对待企业所有者那样来对待他们呢？事实上，上市公司不同于私人企业，并不需要遵循传统的所有权定义。如果我们的目标是建立一家能够历经数代仍然存续的企业，就必须停止自动将股东视为所有者的做法，高管必须彻底消除“自己只为股东工作”的观念。无论是关注短线还是长线的股东，他们的正确做法都应是将自己视为贡献者。

员工贡献的是时间和精力，投资者贡献的是资本（金钱）。这两种形式的贡献对企业的成功来说都至关重要且必不可少，因此双方都应该凭自身的贡献得到公平的回报。从逻辑上讲，为了让一家公司变得规模更大、实力更强或更精通自己的业务，高管们必须如亚当·斯密所言，确保投资者的钱或员工的辛勤劳动首先惠及该公司的消费者。若能做到这一点，该公司想要提升销量、提高价格、拥有更忠实的客户群，以及为公司、投资者等赚到更多的钱都会更容易。益处可能还不止于此，或许我还漏掉了一些什么。此外，高管需要再次将自己视为一个伟大机构

的管家，这些机构的存在目的就是服务于所有的利益相关方。如此一来，得到服务的将不只是少数人的所想、所需和所求，而是与该企业成功息息相关的所有人的所想、所需和所求。

其实，人人都希望自己的工作和生活是有意义的，这是人之所以为人的本质。我们都希望参与比自身更伟大的事业。我不得不认为，正是因为这一点，许多公司才会在将高管与股东视为真正的主要服务对象时，仍然声称自己主要服务于员工和客户。许多人其实都能感觉出现代的资本主义形态与我们的价值观是不相符的，哪怕他们找不到恰当的语言来描述。若我们真的都接受了弗里德曼对企业的定义，那么企业就会视利润的最大化为唯一的愿景和使命，我们也都不会因此而有所不满，但企业并没有这么做。如果企业的真正目的只有赚钱，那么就不会有那么多公司假装自己在追求更伟大的事业或目标。声称企业的存在是为了实现更伟大的目标与真正打造一个这样的企业，完全是两码事。在无限游戏中，这些策略中只有一个是真正有价值的。

变革之鼓已然敲响

2018年，贝莱德集团（BlackRock, Inc.）的创始人、董事长兼首席执行官拉里·芬克（Larry Fink）写了一封给首席执行官们的公开信——《使命感》（A Sense of Purpose），此举在金融业引起了轰动。在这封信中，他强烈建议领导者在创办企业时，设置一些更理想化的目标，不要只着眼于短期的财务收益。他解释道：“若没有使命感，无论是上市公司，还是私人企业，都不可能充分发挥自身的潜力，迟早会失去关键利益相关者的经营授权。他们将屈从于收益分配的短期压力，并渐渐地减少对员工发展、创新和公司基本建设的投入，而这些恰恰又是公司长期增长所必需的要害。”贝莱德集团是全球最大的资产管理公司，管理着超过6万亿美元的资产。尽管像这样要求企业拥抱使命感的呼声并不罕见，但当公开提倡这一观念的人是像拉里·芬克这样在金融界有着至

高地位的人时，关于这一理念的讨论就会从文章中、书中和饮水机旁，真正走入“宫殿”的围墙之内。

股市只有在按其初衷，让普通人也能参与分享国家财富时，才能以最佳的状态运转。不过，美国人对今天的资本主义形态，以及被视为有限游戏工具的股市已经彻底失望。美国人投资股市的比例已经降到了20年来的最低点。流失最多的是中产阶级。人们介意的并非是少数创业者赚了很多的钱，而是该体系本身的失衡以及他们对该体系的不再信任。领导者应该要注意到这一点。

具有讽刺意味的是，所有与上市市场合作过或为其工作过的人都知道，当这个系统过度失衡时，总会迎来修正。这种修正往往是突然且剧烈的。当前的资本主义制度已经严重失衡，内部人士最好能自己进行必要的修正，若做不到，就会增加被迫修正的可能性。因为若拒绝从内部开始改变，人们试图将其完全摧毁的可能性就会上升。无论人们反对的是政府的无能、腐败，还是反对经济模式的失衡，往往都会引发民粹主义起义。别忘了，英国若能放宽对殖民地的经济限制，给殖民地人民更大的参政权，并允许他们分到更多他们帮忙创造的财富，那么美国的独立战争很有可能根本不会发生。道理就是这样，哪里有不平衡，哪里就有动荡。

颠覆一个体系是很严重的事，风险重重，发生得突然也剧烈，而且几乎都会遭遇到反对的抵抗。美洲殖民地的居民们也是在多年呼吁、乞求改变未果后才选择了反抗。意识形态层面的原因只会让他们对革命产生一些兴趣，最终真正将他们推上革命之路的，还是因权力与财富严重失衡而给他们的生活与经济状况加诸的痛苦与限制。面对这些痛苦与限制，他们才开始设想拥有另一种未来。

古罗马时，领导者拒绝向为了保卫罗马而受苦的盟友提供公民身份。在美洲殖民地，尽管殖民地居民的辛勤劳动推动了英国经济的发展，但英国仍然拒绝赋予他们应有的权利。无论是在古罗马，还是在美

洲殖民地，财富与权力的创造都离不开普通民众的支持。在我们这个时代，企业与领导者赚到了钱，却让自己的员工来承担大多数成本。每当企业未能实现随意定下的目标时，员工都要担心自己会不会失业，一旦失业，他们就无法再维持生计，无法养家糊口。员工能感觉到企业和领导者不关心自己，不将他们视为人来对待（请注意：提供免费食物与豪华办公室并不能让员工感觉被关心）。人们希望得到公平的对待，希望分享由他们帮忙创造的财富，从而抵消他们为发展公司而承担的成本。这不是我在要求，而是他们在要求！

与任何对现状的挑战一样，要求抛弃弗里德曼企业观的呼声既可以来自员工，也可以来自领导者，既可以来自外部，也可以来自内部。请留心我们周遭的红色危险信号。在世界许多地方，民粹主义的声音正越来越大。无论在商界还是政界，所有的掌权者都有能力实现变革。不过，别搞错了，无论他们行不行动，变革都已在来的路上。因为这就是无限游戏的原理。有限的体系终有一天会耗尽所有的意志与资源，这一天总会来的。尽管某些人可能积攒了大量的金钱或权力，但这也只是暂时的，这个体系是无法自承其重的，终会分崩离析。如果说历史与几乎每一次的股市崩盘都在设法告诉我们一些什么，那就是，失衡才是最棘手的问题。

变革之鼓已然敲响。越来越多的人开始质疑建立在弗里德曼资本主义理论之上的一些公认原则。企业对弗里德曼企业责任定义的奉行也令越来越多的人感到了不安。自觉资本主义（Conscious Capitalism）、共益企业（B Corp）、共益团队（B Team）等组织都在积极推广利益相关者模式（Stakeholder Model）、三重底线（Triple Bottom Line）等理念，以挑战弗里德曼的理论。在二十世纪八九十年代，美国经济高速发展，涌现了一批商业英雄，比如杰克·韦尔奇，不过，他们的光彩与魅力正在逐渐消失。不言自明的是，我们需要重新定义企业责任，让其与“商业是一场无限游戏”的理念更相符。这个定义必须理解金钱是结果而非目的，这个定义必须让员工和领导者感觉到，他们工作的价值远远

超过了他们为自己、为公司或为股东赚到的钱。

弗里德曼认为企业只有一种责任，就是利润，这种企业观将有限思维体现得淋漓尽致。我们需要用一种超越利润的定义来取代弗里德曼的定义，该定义还必须考虑到让企业得以运转的活力等各个方面。为了增强无限思维对国家、经济以及所有参与该游戏的公司的积极作用，企业责任的定义必须包含以下几点。

1. 推进事业：给人们归属感，让他们感觉自己的生活和工作有超越劳动本身的价值。
2. 保护相关的人：企业的运营方式必须有助于保护员工、消费者，以及我们的生活环境和工作环境。
3. 创造利润：资金是企业维持独立发展所需的“燃料”，能够维持独立发展的企业才有可能继续推进前两个优先事项。

简言之：

企业的责任就是，利用自己的意志和资源来推进一项比自身更伟大的事业，保护与该企业运营相关的人和环境，并且创造更多资源，以确保企业能尽可能长久地继续做到这些事情。企业可以随心所欲地开展业务，前提是要能对自身行动的后果负责。

推进事业、保护相关的人和创造利润这三大支柱在无限游戏中似乎是必不可少的。美国的开国元勋们激励美国人民团结起来，共同争取生命权、自由权和追求幸福的权利。这些关乎人身安全、事业或意识形态的权利都是不可剥夺的，都是美国人民的一部分，都是美国人民可以为自己把握的机会，也正是它们激励了美国这个国家，让美国踏上了它的无限征程。

一个国家必须保护自己的人民，让人民免于生活在恐惧之中。要做到这一点，必须拥有能够抵御外来威胁的武装部队，必须建立司法体系，必须确保国内安宁。同样，公司也必须保护自己的员工，在公司内部打造一种文化，让身处其中的员工能有心理上的安全感，并感觉到雇主是将他们当作人在关心。我们希望看到，公司对我们成长的投资并不逊于它对自身增长的投资。企业不应该仅仅因为没能达到某个主观制定的目标就搞年度大裁员，让员工带着对失业的恐惧来工作。企业若能认真考虑自身产品的生产过程与生产原料会对其产地与销售地社区造成何种影响，便也能为墙外之人提供安全和保护。

对国家而言，归属感以及愿意牺牲自我来推进的意识形态往往被称为某种主义。在商业中，则被称为崇高的事业。无论是在生活的地方，还是工作的地方，我们都应该感觉到自己是在为某样比自身更伟大的东西而奋斗。

国家之间，利益很重要。经济繁荣是一个国家拥有偿债能力的保证。经济强大了，才能有充足的资源熬过困难时期，以及在繁荣时期发展壮大。对企业来说也是一样的道理。无论在一个国家里还是在一家企业中，每个人都希望拥有努力工作赚钱的机会，这样才能维持生计、养家糊口。

一个具备无限思维的国家，它的目标也会是其人民的目标。这样的国家在奋进的同时，会为人民服务，也会让人民参与它的奋斗。这就是一个人与祖国之间的情感纽带，这就是这个人爱国情感的来源。换到商业背景下就是，企业的目标不能仅仅迎合股东的目标，还必须与员工的目标一致。如果我们希望自己的工作不仅能惠及我们自身，而且能惠及我们的同事、客户、社区和整个世界，那么我们就应该就职于与自己拥有一致价值观和目标的企业。如果企业不愿意这么做，我们可以要求他们这么做。无论是谁，只要为实现公司目标付出过血汗与泪水，就有权因为自身的贡献而得到重视，就有权分享自己的劳动成果。

弗里德曼认为，我们努力工作的成果应首先惠及精英统治阶级（所有者），无限思维的领导者则会承诺，只要有共同的目标，所有贡献者就都将共享三大支柱中所言明的益处。在工作中，我们有权拥有心理上的被保护感，有权获得与付出等值的回报，有权贡献于比我们自身更伟大的事业。这些都是我们不可剥夺的权利。与其他任何无限追求一样，商业也只有在得到人们的赋权，且代表人们行使该权力时，才能拥有更强大的力量。无论何时，颠覆一旦发生，就不会很快结束，这一点是不变的，但领导者的反应方式可以改变。弗里德曼对企业责任的有限定义侧重于资源的最大化，而修正后的定义将人们的意志也纳入了考量的范围。

第6章

意志和资源

拉斯维加斯的四季酒店是一家非常优秀的酒店。它的优秀并不在于高档的床，因为任何酒店都能买到高档的床。四季酒店的优秀源自那里的员工。当你走在酒店大堂时，会有工作人员热情地跟你打招呼，这一举动就足以让你感到他们的真心，感到他们不是因为公司有规定才会这么做的。人类是有很高理解力的社会性动物，我们分辨得出其中的差别。

我就曾在该酒店大堂的咖啡厅里有过这样的经历。我当时在拉斯维加斯出差，一天下午，我去买咖啡。那天的咖啡师是个名叫诺亚的年轻人，既风趣又迷人。就因为他，我那天买咖啡的体验比平时更加愉快。我们闲聊了一会儿，我终于忍不住问他：“你喜欢你的工作吗？”诺亚毫不犹豫地回答道：“我爱我的工作！”

对我这个行业的人来说，这样的回答意义重大。他没有说“我喜欢我的工作”，而是说“我爱我的工作”。二者之间的区别是很大的。“喜欢”是理性的，我们可以喜欢与自己一同工作的人，喜欢这项挑战，喜欢这份工作。但“爱”不同，爱是感性的，是很难量化的。这就像问一个人“你爱你的另一半吗”，他们回答“我很喜欢我的另一半”。这个答案显然意味不同，因为爱的标准更高。因此，当听到诺亚说“我爱我的工作”时，我一下就来了精神。从这个回答中，我发现，诺亚与四季酒店之间形成了情感纽带，这个纽带比他赚到的钱和工作本身都更为重要。

我立刻追问道：“请跟我讲讲，你究竟是因为四季酒店的哪一点而让你说出‘我爱我的工作’。”诺亚再次不假思索地答道：“一天里任何时候，经理们只要从我身边经过，就会问我过得好吗，有没有什么需要

的，有没有什么他们能帮得上忙的。不只是我的直属经理，任何经理都是这样。”他继续说道：“我还有另一份工作。”他进一步解释道，在他的另一份工作中，经理们每次从他身边走过时，都会试图抓住他的错处。诺亚叹气道：“我会埋头干活，尽量不引人注目。我所想的只有赶紧熬完这一天，拿到薪水。只有在四季酒店，”诺亚说，“我才觉得我可以做自己。”

四季酒店的工作让诺亚发挥出了最佳的水平，这正是每个领导者对自己员工的期望。正因如此，才会有那么多的领导者，甚至包括一些最善意的领导者，常常问：“我该怎样做才能让我的员工发挥最大的价值？”但这个问题问错了。因为它问的不是该如何帮员工成长得更强大，而是该如何更多地榨取他们的价值。人不是待拧干的湿毛巾，员工也不是可以让我们将最后一点儿价值都榨个一干二净的物品。用这个问题得到的答案，或许能在短时间内提高产出，但从长远来看，这是以牺牲员工和文化为代价的。这样的方式永远无法让诺亚感受到他在四季酒店所感受到的那份爱与投入。更好的问法是：“我该如何创造出能让员工主动发挥出最佳水平的环境？”

很多时候，当业绩未达预期时，我们的第一反应就是责备做事的人。但在诺亚所做的两份工作中，工作者都是他，唯一不同的是他工作时所处的领导环境。另一份工作更看重他的产出，而非他能在工作中感受到多少支持。若我是在那里遇到诺亚的，那他所带给我的体验可能就会截然不同了，我很可能就不会在书中提到他，也不会称赞他所工作的酒店了。真正能够产生更大影响的，不是做事的人，而是领导者。

四季酒店的经理们知道自己的工作就是为诺亚创造出良好的工作环境，让他能够在做自己的同时，越来越好。如果我们在培训领导者时，能够教他们如何将员工置于比结果更优先的位置，那么他们就会努力为员工创造出良好的工作环境。这才是“领导”这份工作的真正含义。对一名经理来说，抽点时间去大堂走走，问问员工们过得怎么样，并且真心

在意员工给出的答案，显然是件不需要什么成本的事。正是因为四季酒店的领导层将员工的意志置于员工所能产出的资源之前，员工才会想要为这份工作奉献一切，四季酒店的顾客也才能通过这些员工感受到这一点。

意志优先于资源

在任何游戏中，几乎都有两种“货币”是必需的——意志和资源。资源是有形的，容易衡量。我们在谈论资源时，说的通常都是金钱。根据组织偏好或当下标准的不同，这些资源可以有多种计算方式——收入、利润、税息折旧及摊销前利润、每股收益、现金流、风险资本、私募股权、股价等。资源的来源一般在企业外部，比如客户或投资者。资源代表的是，能帮助该组织健康发展的所有财务指标的总和。

意志则正好相反，它是无形的，难以衡量。我们在谈论意志时，说的通常都是人们对工作的感受。意志包含了士气、动机、灵感、奉献、参与的欲望、主动付出的欲望等。意志的来源一般在企业内部，比如领导层的素质，以及崇高的事业的清晰度和影响力。意志代表的是，能帮助该组织健康发展的所有人力因素的总和。

无论是有限思维的领导者还是无限思维的领导者，他们都知道资源至关重要，也都认同意志至关重要。在我所见过的首席执行官中，还没有哪个认为员工是不重要的。但问题是，意志和资源永远无法同等优先。两者无法兼得的情况并不少见，这时候，领导者就必须选出自己更愿意牺牲哪一个。问题是他们会选哪一个呢？每个领导者都有自己的偏好。

我们大都有过这样的经历，在会上听某位领导介绍公司最看重的事情，而这些事情的优先顺序往往是：增长、客户、员工。尽管领导者坚称他们关心自己的员工，但员工在该名单上出现的顺序很重要。按照上

面这种顺序，至少有两件事在领导者眼中是比员工更重要的，其一就是资源。领导者的优先事项名单暴露了他们的偏好，这种偏好将会影响他们的决策。

有限思维的领导者往往偏好成绩。因此，他们常常会选择能在短期内出成绩的决策，哪怕这样做会“令人遗憾地”损害员工的利益。在经济困难时期，这类领导者首先想到的会是裁员、极端削减成本等措施，而不会探索虽可能不会产生同样立竿见影的效果，却会对长远发展更有利的替代方案。对偏好资源的领导者来说，他们会更倾向于计算，比如下周裁员10%能给公司节省多少成本，他们往往不会选择那些需要更长时间才能在资产负债表上体现出成效的方案。

无限思维的领导者则与之相反，他们会努力看向更远处，不被当下的财务压力所局限，并尽可能多地让员工与企业共享利润。在经济困难时期，这类领导者不太可能将员工看作一项待削减的开支，他们会更愿意探索节约资金的其他方法，哪怕需要更长时间才能看到成效。无限思维的领导者可能会选择通过休假而非裁员来帮忙管理资源，比如，要求每位员工休两周或三周的无薪假期。员工的薪水可能会减少，但所有人都能保住工作。共克时艰其实会令一个团队更加团结，这正是自然灾害后人们会更加团结的原因所在。若将成本过分地加诸某一部分人身上，企业文化就可能被撕裂。对于眼光长远、不只着眼于当下困局的无限思维的领导者来说，他们更愿意守护员工的意志，哪怕这意味着要花一个季度、一年或更长时间才能节省下足够平衡资产负债表的资金。这些领导者明白，员工的意志正是激励员工主动付出、主动解决问题、主动发挥想象力以及主动开展团队合作的动力，这些都是一个组织要长久存在并繁荣发展所必不可少的。将强大的意志置于资源之上所能产生的价值不可被低估。

那些偏好资源的领导者每次听到我这样的人说员工应该优先于资源时，都会被吓了一跳，他们以为自己听到的是，我认为金钱不重要。他们

还以为自己听到了，我认为他们并不关心自己的员工。但这不是一道非此即彼的选择题。这一偏好其实不用这么极端。昔客堡（Shake Shack）的创始人丹尼·迈耶（Danny Meyer）就分享过他的偏好，他说自己的生意是由49%的技术与51%的感性组成（这就是该餐厅老板对意志和资源的看法）。¹哪怕对意志的偏好只比对资源的偏好多了那么一点点，也有更大的可能创造出更强大的企业文化，以确保意志和资源的供给都是充足的，从而保证企业能够长久地走下去。

意志的成本

博柏利（Burberry）前首席执行官兼苹果前零售高级副总裁安吉拉·阿伦茨（Angela Ahrendts）表示，有太多领导者“把人视为成本”。尤其在人员流动率非常高的零售行业，人们普遍的逻辑是，“我凭什么要在那些迟早要走的人身上投资呢？”对商业运作方式的这一理解是片面且有限的。太多有限思维的领导者只看到了不投资于人所能省下来的钱，却忽视了这种做法本身所带来的额外成本。雇用新员工来填补空缺也是要花钱的，失去有经验的员工，以及等待新员工接受培训并适应新文化，都会影响企业的生产率。再加上高流动率会造成的士气低落，人们不禁好奇，省下来的钱是否真的值得用这一切去换取。阿伦茨也很好奇，因此她算了算账，算出来的结果令她感到吃惊，苹果公司在照顾员工方面的实际增量成本其实是零。

在苹果公司，工作于零售店的全职员工与工作在办公楼的全职员工享受着完全一样的福利，包括全覆盖的医疗保险和牙科保险，如果他们想在工作之余上上课的话，还能享受2500美元的教育报销额度。苹果公司是首批确保新员工工资不低于每小时15美元的公司之一。苹果公司的全职零售员工还享有与其他所有办公室员工完全一样的公司股票认购权。由此产生的所有额外成本都被更少的招聘与培训成本抵消了，而大多数滥用裁员手段的企业都将在未来的某一天被迫支付重新招人填补空

缺的费用（首席执行官在报告裁员节省了多少资金时通常不会将该成本计算在内）。为了填补人员不断流失所造成的空缺，许多大型零售企业都不得不养一支庞大的招聘队伍。苹果则不然，他们用于招聘零售员工的团队人数很少。当然，有些人会说，就员工人均为公司赚到的钱而言，苹果要比大多数零售企业赚的多得多，自然付得起更高的工资。不过，开市客（Costco）付给收银员的平均工资也有15.09美元，还帮他们加入了养老金计划以及购买了健康保险。²开市客发现，此举不仅降低了人员流动率，还提高了生产率，而这足以抵消此举所带来的额外成本。此外，当员工能感受到公司的照顾时，顾客往往也能享受到更好的服务，而这很可能转化为更高的平均销量。

如果实际成本为零，那么我们对待员工的不同就只是思维方式的不同。正因为换了一种思维方式，苹果和开市客的平均人员留存率达到了90%左右，而其他零售企业的平均留存率只有20%到30%。³，⁴有限思维的组织将人视为一种需要管理的成本，无限思维的组织则更愿意将员工视为有血有肉的人，并认为他们的价值是无法像机器一样来计算的。对人的投资不仅仅是为他们提供高薪和好的工作环境，而且意味着要像对待人一样来对待他们。要明白，他们与其他所有人一样，都有抱负、有恐惧、有想法、有观点，最重要的是，也想要感觉到自己的重要性。这对许多有限思维的领导者来说或许更像一场冒险，多花了一大笔钱，却只能“期望”这钱花得有用。相比之下，工资和福利的减少显然更容易计算。不过，这个风险或许是值得一冒的。当企业让员工感到自己很重要时，员工就会更加团结，而这种团结是金钱买不到的。

当意志强大时

在2008年的经济大萧条期间，人们都因经济困难而勒紧了裤腰带，许多人选择暂停购买非必需品，比如为家里和办公室添置的储存和收纳用品。当时，容器商店集团（The Container Store Group）是全美国唯一

一家专卖家用和办公室用储存与收纳整理产品的零售商，他们明显感知到了这一消费变化，因为销售额下降了13%。这让一家不习惯收入突然下滑的公司陷入了麻烦。自1978年第一批门店开业以来，该公司一直维持着20%的年均复合增长率。领导层进行了员工谈话，得出的结论是，他们必须削减开支，而且削减幅度不能低于销售额的下滑量。雪上加霜的是，没人知道这场大萧条会持续多久，也没人知道他们的销售额还会下降多长时间。

容器商店集团一直是一家员工至上的公司，并以此为傲。因此，当大萧条来袭时，他们拒绝用解雇员工的方法来应急，但他们必须做点什么。在提出要无限期暂停加薪和暂停提供养老金配比福利时，领导层并不确定员工会作何反应。他们希望员工能够理解并认同他们的观点，也就是与其让少数人承受更多的损失，不如让所有人来共担损失。

员工的实际反应出乎了领导层的预期，令他们既惊讶又开心。员工们做了一些根本没人要求他们做的事情，他们不仅接受暂停加薪，还主动承担了帮公司寻找更多省钱方法的责任。尽管没人要求，他们却会主动降低出差时的住宿档次。一些人直接放弃了住酒店，选择住到朋友家里或亲戚家里。还有一些人选择自己支付出差时的餐费与出租车费，不再向公司申请报销。任何能帮公司省钱的方法，他们都用了。员工们甚至还会主动联系供应商，询问对方有没有可以帮他们公司省钱的方法。令人惊奇的是，供应商们也很乐意帮忙。这几乎是前所未闻的！供应商显然没有用降价来帮助客户渡过难关的义务。但是，容器商店集团与供应商之间有着牢固的关系，令供应商想要出手相助。

该公司的联合创始人、前首席执行官基普·廷德尔（Kip Tindell）说：“自上而下的举措连这一半的效果都达不到。”他是对的。公司领导层可以要求员工降低酒店档次，可以向员工施压，要求员工竭力说服供应商降价，也可以宣布不再报销出差费用。如果他们这样做了，钱确实能省下来，但也有引发大规模反抗的风险。众所周知，降档的要求会挑

起员工对公司及其领导层虽安静但沸腾的怒火。但在容器商店集团这个案例中，贡献源自员工们自己的欲望，因此才会有如此不同的结果。空气中似乎弥漫着员工对公司强烈的感情，大家都在积极努力地为公司寻找脱困之法。最重要的是，每个人都觉得他们是团结一致的。

有限思维的领导者往往认为，意志来自外部的激励——薪酬、奖金、津贴或员工之间的竞争。如果光靠这些就足以激励一个人，那就好办了。钱可以买到很多东西，我们确实可以用金钱来激励员工，可以付钱让他们努力工作。但金钱买不到真正的意志。⁵同样是为企业付出所有，员工是为了获得外部奖励，还是发自内心地去做，这代表着这个企业是充斥着唯利是图者还是满怀热情者。唯利是图者只会在我们持续支付高额报酬的情况下努力工作。他们对这家企业或这个团队几乎毫无忠诚可言，没有真正的归属感，也没人真正觉得自己是在为比自己更伟大的事业做贡献。唯利是图者不太可能会出于爱和贡献而牺牲自己的利益。满怀热情者正好与之相反，他们希望成为企业的一部分。尽管他们可能会因为自己的工作而变得富有，但富有并不是他们工作的目的。之所以在此工作，是因为他们相信这家企业所追求的是崇高的事业。

廷德尔说：“我们的员工将这一事业置于了自己之前。”崇高的事业固然重要，但只有它还不足以激发员工的意志。即使在大萧条期间，廷德尔仍然能看到长期投资所能带来的回报。在廷德尔的记忆中，2008年经济大萧条期间所发生的那些事展示了什么是“自发的爱与贡献”。他可能觉得那些都是自发的，但事实并非如此。坚强的意志不是一蹴而就的，也不会凭空而来。为员工提供良好的工作环境，为一线员工提供高于其他零售工作的薪资，以及培训领导者，让领导者懂得该将员工的个人成长置于企业的财务增长之前，这些都是容器商店集团多年如一日的坚持。作为回报，员工才会同样多年如一日地关心顾客、公司和供应商；当公司陷入困境时，员工和供应商才会想要做一些能帮到这家公司的事情。别人如何对待我们，取决于我们如何对待他们，正所谓“投我以桃，报之以李”。

在无限游戏中，偏好意志的企业最终能有更出色的表现，原因之一与可由我们掌控的某个因素有关。我们可以掌控自己的花钱方式或金钱管理方式，相较之下，我们对自己赚钱方式的掌控就要逊色许多。政治、经济周期、市场波动、其他玩家的行为、客户偏好、技术进步、天气等不可抗力都会严重影响我们积攒资源的能力。领导者对上述因素的掌控是非常有限的，但他们几乎完全掌握着意志的来源。意志源于企业文化。

与终究有限的资源不同，意志是可以无限供应的。正因如此，相较于以资源为优先来运营的企业，以意志为优先来运营的企业最终会拥有更强大的适应力。当经济不景气（这事常常发生）时，偏好意志与偏好资源的公司相比，前者的员工更有可能团结起来保护彼此、保护公司、保护资源，以及保护他们的领导者。没有人要求他们这么做，这都是他们自己的选择。当员工的意志特别强大时，就会带来这样的结果。廷德尔曾说：“我们建立了家人般的感情——热爱且忠诚于彼此、客户、供应商和社会。我们建立该企业的目的正是让与之相关的所有人都能越过越好。”

THE INFINITE GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE

要素 2

建立彼此信任的团队

第7章

彼此信任的团队

“这有什么用？”乔治问道，“这与油田毫无关系啊。”这也是房间内其他人的普遍共识。他们马上就要前往壳牌公司的尔萨钻井平台（URSA）工作了，根本没有多余的时间参加这次工作坊。尔萨是壳牌公司有史以来建造的最大海上深水钻井平台。

尔萨钻井平台有48层楼高，可在水下深度超过900米的海床上作业，深度超过了1997年世界上的其他所有平台，建造成本为14.5亿美元。巨大的规模与成本也带来了各种各样的新挑战与新危险，壳牌公司希望不要出现任何差错，因此选择由里克·福克斯（Rick Fox）担任该项目的负责人。

福克斯是硬汉中的硬汉，努力且自信。他不能容忍软弱，觉得自己完全有权这样做。这是世界上最危险的工作之一。一个人只要走错一步，朝错的方向瞥一眼，就可能会立刻丧命，被沉重的活动部件活生生撕成两半。他深知这一点。安全是福克斯最关心的事，另外还要确保尔萨钻井平台以最大采油能力运转，在可控范围内尽可能多地从海底抽取石油。

克莱尔·尼埃（Claire Nuer）住在远离壳牌公司新奥尔良总部的北加州，经营着一家领导力咨询公司，在听说了壳牌公司尔萨钻井平台这个项目后，就一直在寻找可分享她的理念的机会。她打电话给福克斯自荐。在询问福克斯面临着哪些挑战时，福克斯大部分时间都在谈论技术上的挑战。在让福克斯解释完运营钻井平台所有的复杂之处后，尼埃给出了一个相当出人意料的提议。如果福克斯真的想让他的工作团队在面对所有新挑战时能同时做到安全和成功，那么他的团队就需要学会表达

自己的感受。

这种提议听起来“很新潮”。一听就觉得，这不可能在任何以业绩为导向的严肃组织中找到一席之地。若在其他任何时候，听到这里，福克斯可能就已经挂断电话了，毕竟在他的观念中，表达自己的感受无益于暴露自己的软弱。但尼埃很幸运，不知出于何种原因，福克斯认真听了她接下来说的话，或许是因为福克斯正苦于与儿子之间的紧张关系吧。他甚至接受了尼埃的邀请，与儿子一起飞到加州，参加了她的一次工作坊。在那里，这对父子拥有了一个可以敞开心扉的安全空间，说出了对彼此的感受。这次工作坊对改善这对父子的关系产生了深远且积极的影响，这让福克斯产生了让其他人也来体验一下的想法。福克斯聘请了这位很特别的姑娘，让她飞到美国的另一端，用她的理论在那群粗野、满手老茧的路易斯安那州员工身上试验一下。福克斯知道，当他要求员工去做这样一件事时，迎接他的会是怀疑与冷嘲热讽。但福克斯真心关心自己的员工，无论要承受何种羞辱与嘲笑，与员工们所能得到的好处相比都是不值一提的。于是，试验开始了。

几乎每一天，尔萨团队的成员们都需要围坐在一起，花几个小时聊聊他们的童年，聊聊他们的感情，分享各种快乐的与不那么快乐的回忆。¹有一次，一名成员在告诉队友自己的儿子得了绝症时，终于忍不住流下了眼泪。成员们不仅要谈论自己，还要倾听他人。另一名成员回忆称，他在他人的鼓励下问出了一个问题：“如果你们能改变我身上的某一点，那会是什么？”“你不听别人说话，”他们告诉他，“你自己说得太多了。”对此，他说：“请再多告诉我一些。”

福克斯的团队比以往任何时候都更深地了解了彼此。不只是作为同事的了解，也是作为人的了解。他们抛去了伪装，敞开了心扉，展露了真实的自己。在这个过程中，事实也逐渐清晰，大多数人所表现出的硬汉形象不过是他们的保护色。在坚硬的外壳下，他们与其他所有人一样，也会感到疑虑、恐惧和不安。过去，他们只是把这些都藏了起来。

在尼埃的指导下，福克斯用一年的时间为尔萨项目打造了一个成员能够安心与彼此相处的团队。

纯工作关系的团队与彼此信任的团队是不同的。在纯工作关系的团队中，成员间主要是事务性的关系，这一关系的基础是完成工作的共同愿望。这并不妨碍我们喜欢与自己一起工作的人，也不妨碍我们享受自己的工作。但这些并不足以打造一个彼此信任的团队。信任是一种感受。我们不可能因为领导者要求我们快乐就快乐，要求我们感觉被激励就被激励，同理，信任也不是有了领导者的要求，就可以做到的，无论是信任领导者，还是信任彼此。要培养信任感，首先要能安心地展露真实的自我。我们必须能够在自己脆弱的时候依然充满安全感。是的，脆弱。光是读到这个词就会让一些人坐立不安。

在一个彼此信任的团队中工作，我们能安心地说出自己的脆弱，可以安心地举手承认自己犯下的错误，坦诚自己的不足，对自己的行为负责并寻求帮助。寻求帮助是表达脆弱的一种方式。在一个彼此信任的团队里，我们敢于寻求帮助，因为我们深信领导或同事会在身边支持我们。“在小事上一点点袒露彼此的脆弱，才会产生日积月累的信任。”休斯敦大学研究教授布琳·布朗（Brené Brown）在她的著作《敢于领导》（*Dare to Lead*）中写道：“信任和脆弱相生相伴，背叛了其中任何一个就是将两个一并摧毁。”

如果不是在一个彼此信任的团队中，我们就不敢在工作中表现出任何脆弱，我们常常要为了掩盖脆弱而撒谎、隐藏和伪装。我们隐瞒错误，表现得好像很清楚自己在做什么（即使我们并不知道），永远不会承认自己需要帮助，因为我们害怕丢脸，害怕被报复，害怕发现自己的名字出现在下一轮的裁员候选名单中。没有彼此信任的团队，组织中的所有裂缝都会被掩盖或忽视。如果这种情况持续一段时间，就会令裂缝进一步扩散，直至崩塌损毁。因此，任何组织想顺利运转，都离不开彼此信任的团队。在石油钻井平台上，这样的团队是能救命的。

哈佛商学院教授罗宾·埃利（Robin Ely）在《哈佛商业评论》发表过一篇关于尔萨钻井平台的文章，埃利指出：“安全感的表现之一就是能够承认错误并乐于学习，能够说出‘我需要帮助’‘我自己抬不动这个东西’‘我不确定该怎么读这个仪表’。”尔萨团队发现，成员彼此之间的心理安全感越强，交流就越高效。团队中的许多人都是第一次在工作中感觉到这种安全感，让他们可以放心地提出自己的担忧。此举成效非凡。尔萨钻井平台的安全记录跻身业内最佳。在壳牌公司内部，尼埃的这些信任构建技巧逐渐得到了广泛应用，令壳牌公司整体的安全事故减少了84%。^{2, 3}

我认为团队成员必须学会向彼此袒露脆弱，必须关心彼此，并将这份关心表现出来，但每当我提出这些建议时，往往会遭到反对。不过，如果这是连里克·福克斯这样的硬汉都能在石油钻井平台上做到的事，那么任何行业的任何领导者都应该做得到才对。我们信任他人的能力与行业无关，这是生而为人就具备的能力。有时，我们只需要做一件事，就是将这些概念翻译成适合我们工作文化的语言。

在商业世界中，阻力往往来自另一个地方。企业领导者告诉我，商业应该是专业的，不应该带入个人情绪。领导者的职责是拉动业绩，而非让员工开心。然而，事实是，个人情绪的存在是无可避免的。如果你曾在工作中感觉到沮丧、兴奋、愤怒、振奋、困惑、嫉妒、自信或不安，那么恭喜你，你是人。作为人，我们不可能仅仅因为是在工作就彻底关闭自我情绪的阀门。

我们不能将“能够安心表达自己的感受”与“缺乏情绪的专业性”混为一谈。我们也绝对不能因对团队中某个成员的不满而发怒或脱队。我们是成年人，必须懂尊重、有礼貌、思虑周全。但这并不意味着我们可以或应该设法关闭自己的情绪阀门。用有限的思维看待领导力会否认员工感受与业绩之间的关联。里克·福克斯这样的领导者懂得，彼此信任的团队是以员工感受为核心建立的。事实也证明，彼此信任的团队才是最

健康、最高效的团队。

就行业历史均值来说，石油钻井平台的运行时间（平台正常运转与操作的时间）为95%，而尔萨钻井平台的运行时间为99%。他们的产量比行业基准高出了43%，比自己定的生产目标多了1400万桶。同时，尔萨钻井平台的环保目标的完成度也远远高于预设目标。换言之，要打造高业绩的团队，信任比业绩更重要。

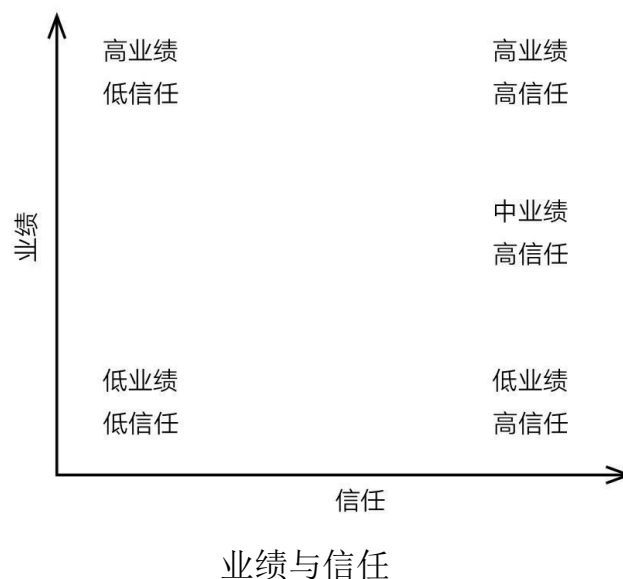
业绩与信任

美国海军海豹突击队因《勇者行动》（*Acts of Valor*）和《菲利普船长》（*Captain Phillips*）等电影为公众所熟知。这支海军特种部队确实是全球表现最佳的团队之一。不过，可能令你惊讶的是，他们挑选的队员却不一定都是表现最佳的。在挑选适合海豹突击队的队员时，要做的事情之一就是评估候选人在两条坐标轴上的表现：业绩与信任。

业绩考察的是技术水平是否过关。这个人的工作能力有多强，够勇敢吗，能在压力下保持冷静吗？而信任考察的是品格，是否谦逊，是否有个人责任感，能给予队友多少支持，是否能对其他队友产生积极的影响。换成一名海豹突击队队员的话来说就是：“我或许敢把自己的性命托付给你，但我敢不敢将自己的钱或妻子托付给你呢？”也就是说，即使我信任你的技术水平，也不意味着我信任你这个人。你也许能在战斗中保护我的安全，但我对你的信任还不足以让我在你面前袒露自己的脆弱。这就是身体安全与心理安全的区别。

看看这张业绩与信任的图，显然没有团队会欢迎处于左下角的人——低信任、低业绩者。但也很显然，所有团队都会想要处于右上角的人——高信任、高业绩者。海豹突击队发现，处于左上角的低信任、高业绩者会成为团队的毒瘤。这类队员有自恋的特征，会毫不犹豫地指责他人，总是把自己放在第一位，会“说别人的坏话”，也可能对队友产生

负面影响，尤其是对团队里的新成员或低级别成员。比起低信任、高业绩者，海豹突击队更愿意选择高信任、中业绩者，有时甚至宁愿选择高信任、低业绩者（这是相对而言的）。海豹突击队是全球表现最佳的团队之一，连他们都认为信任优先于业绩，我们还有什么理由认为在商业中业绩是最重要的呢？



如果企业文化中充斥着实现季度或年度目标的巨大压力，那么许多领导者都会以高业绩为重，几乎不考虑自己团队中的其他人是否信任自己。这些价值观会反映在他们的雇用、提拔与解雇的人选上。在经济飞速发展的20世纪80与90年代，时任通用电气首席执行官的杰克·韦尔奇就是一个极端的例子。韦尔奇太关注赢与第一了，他甚至给自己的一本书起名为《赢》（*Winning*）。他几乎只关注业绩，而这对信任的培养是不利的。与海豹突击队一样，韦尔奇也按照两条坐标轴对高管们进行了排名。不过，不同于海豹突击队的是，他的两条坐标轴分别是业绩和潜力，本质上说就是业绩和未来可能实现的业绩。基于这些衡量标准，考核年度内“赢”的最多的人就会被定为升职对象，业绩不佳者就会被解雇。在打造高业绩文化的过程中，员工的产出成为韦尔奇最看重的因素。

在韦尔奇所营造的这类环境中，受益且获得赞美的往往是高业绩者，包括那些低信任的高业绩者。但问题是，对团队不利的成员往往更关心自己的业绩与职业发展轨迹，而非自己可以为整个团队的提升做些什么。尽管他们可能会在短期内拥有好的表现，但他们取得成果的方式往往助长了一种不利于他人发展的环境。实际上，在执迷于业绩的企业文化中，领导者往往会为了进一步拉动业绩而鼓励内部竞争，从而加剧上述这些倾向。

在韦尔奇等有限思维的领导者看来，让员工相互竞争似乎是个不错的主意。但这个“不错”只是暂时的。员工之间迟早会出现破坏信任的行为，比如私藏信息而非共享信息，窃取他人的功劳而非赞美他人，操控更年轻的团队成员，以及为逃避个人责任而把其他人推出来当替罪羊。某些人甚至会为了自己的发展而故意伤害同事。不难预料，这个组织迟早会为之付出代价，也许最终会被这场游戏彻底淘汰。韦尔奇所打造的通用电气几乎是注定会失败的。事实上，若不是美国政府在2008年股市大崩盘后为其提供了3000亿美元的紧急财政援助，通用电气可能早已不复存在。时间总能很好地揭示真理。

毫不令人意外的是，就连出发点很好且重视信任的领导者也时常落入雇用、提拔高业绩者的陷阱，没有考虑这些人是否值得信任，以及是否具备信任他人的能力。业绩很容易用产出来量化。在商界，我们确实有各种各样的个人业绩衡量指标，但几乎没有任何指标可以有效评估一个人的可信度。有趣的是，在任何团队中，低信任的高绩效者其实都极易识别，去问问团队成员就知道了。

如果我们问，在团队里大家最信任的是谁，是谁总在危急关头支持他们，他们提到的很可能是同一个人。这个人可能是个人业绩最高者，也可能不是，但就能帮忙提升团队业绩这一点来看，这样的人不仅是很棒的队友，也很可能是天生的领导者。这类成员往往具有很高的情商，能够为自己影响团队活力的行为负起责任。他们既想要个人的成长，也

想要帮助身边的人一起成长。如果我们倾向于用业绩而非信任来评估一个人，那么在考虑升职人选时，就很可能忽略一个值得信任的人对团队的价值。

面对别人的评价，低信任的高业绩者鲜少会认同，甚至鲜少会愿意倾听。他们认为自己就是值得信任的，真正不值得信任的是其他所有人。对于自己的表现，他们会找借口推脱而非承担相应的责任。尽管他们能感觉到团队中的其他人好像做什么都不带自己（他们很可能说服自己相信，这都是因为那些人在嫉妒自己），却没能意识到团队中的所有紧张关系都有一个共同要素，那就是这些人的存在。即使得知了团队其他成员对他们的感觉，许多低信任的高业绩者仍然会选择加倍努力提升业绩，而非努力修复失去的信任。毕竟，在以业绩为导向的企业指标下，曾经推动过他们职业发展，让他们拥有过职业安全感的都是业绩。他们有什么理由要改变自己的策略呢？

优秀的领导者不会无缘无故地偏爱那些高信任的低业绩者，也不会立即抛弃那些低信任的高业绩者。如果有人业绩不佳，或者他们的行为方式对团队活力产生了负面影响，那么领导者首先需要问的问题是：“他们是可以通过培训改变的吗？”作为领导者，我们的目标是确保员工拥有工作所需的技能——技术技能、人际交往技能或领导技能，帮助他们发挥出真正的最佳状态，成为团队的宝贵财富。这意味着我们必须帮助低信任者习得人际交往技能，让他们变得更值得他人信任，也更能信任他人；也必须帮助低业绩者习得技术技能，让他们能够提高自己的业绩。只有当事实证明这名成员无法改变，即抗拒别人的反馈且对自己的工作表现不负责任时，才应该认真考虑将他们从团队中剔除。走到这一阶段，若领导者仍然决定留用他们，那么该领导就要为由此而产生的后果负责。

在团队内，成员们天然就会排斥自己不信任的人，或者会与此人保持一定的距离。由此，领导者应该更容易看清，若要提升团队的整体业

绩，自己应该怎么做。那么，一定是这样吗？低信任的到底是这名成员，还是剩余的整个团队？

建立信任的安全圈

他遭到了好几项指控。调查人员正在对其中的一些指控进行调查，包括他是否在外出巡逻时偷偷去健身房睡觉，他是否给自己的私家车窗户非法涂色，以及他是否试图通过出示警徽，要求别人帮他消除在另一个辖区的罚单。杰克·科伊尔（Jake Coyle）警官觉得是其他人一直咬着他不放，就想抓他的把柄。他们就好像一直在用显微镜盯着他一样。他不信任自己的领导，不信任自己的同事，领导与同事也不信任他。

其他警官经常刁难他，要让他清楚地知道他不是小团体中的一员。大家取笑他，捉弄他。比如，会把垃圾扔进他的车里，或者用扫雪机把他的车堵在里面。在其他警官看来，这些都只是玩笑，但其实这些事情对科伊尔警官的影响很严重。大家对他的刁难，让他在警局内无法产生丝毫的信任感和心理上的安全感。问题已经严重到了让他痛恨来上班，他每天只想着赶紧值完班回家。这种情绪越攒越多，他开始考虑换个地方重新开始，也已经在考虑调到另一个警察局去，直到出现了一个转折点。

杰克·考利（Jack Cauley）成了美国岩堡市警察局的新局长，他发现这里的警队与他刚刚离任的那支，以及全美各地不计其数的其他警队（以及如今太多的企业文化）很相似。在这里，许多人都觉得自己被低估、被忽视。他们觉得这里看重业绩数字胜于一切，因此自己不得不拼命做出业绩。一名警官在描述考利上任前的该警察局时说：“我们等于被告知我们是可替代的，没有我们，还有成百上千的人在等着做这份工作。”另一位警官说：“新人有什么想法也不敢提。”在这里，警官们会因开出的罚单不够多而受到惩罚。

有的警察局会将罚单数和逮捕数当作业绩的唯一考核指标，考利局长对他们的这套做法再熟悉不过了。1986年，在堪萨斯州的欧弗兰帕克市，考利正式加入警队，开始了自己的职业生涯。作为一名雄心勃勃的年轻警员，他的晋升之道就是超额完成上级给他下达的指标。他们如果要求他开X张罚单，他就会开出2X倍。经过这么多年，他逐渐意识到，这种对业绩的过分关注是以牺牲警官们与执法文化为代价的。因此，在获得担任岩堡市警察局局长的这个机会时，他立刻就接下了。这是他的机会，他要证明，将警局原本建立在罚单数量、盲目服从或工作不安全感之上的文化改成以信任为基石的文化后，究竟会对警队产生什么样的影响。

考利上任后采取的首批举措之一就是，与警察局里的每位成员都进行一对一谈话，倾听他们的声音，包括所有宣誓就职的警官和其他工作人员。谈话期间，许多人告诉他，他们希望能在停车场周围建一圈围栏，为此已经申请了好多年。停车场环绕着岩堡市警察局的总部大楼，是一圈开放、暴露的沥青地。警官和工作人员抱怨说，夜里下班时，外面又黑又安静，他们都害怕走到车里去，不知道是否会有人藏在暗处，等着偷袭他们。这么多年来，管理层的回复一直都是让他们自己想办法克服。他们被告知有比建围栏更急需用钱的事情，那些事情才跟他们的警察工作更相关，比如添置新枪或新车。

考利逐渐看清，在这个警察局，大家并不觉得领导在背后支持他们。作为新局长，他必须先建立一个“安全圈”（circle of safety）。没有安全圈，他所需采取的其他举措都不会奏效。考利迅速下令建好了停车场的围栏。这一简单的举动等于通知了所有人：一切都将不同了。这只是考利的一系列举措之一，这些举措看似很小，但却向他的下属们发出了响亮的信号——我听到了你们的声音，我是关心你们的。安全圈是信任存在的必要条件，它描述的是一种能让人产生心理安全感的环境，让人们能够安心地向同事袒露脆弱，安心地承认错误，安心地指出领导对他们训练方式的不足，安心地分享自己的恐惧和焦虑，当然，也能安心

地开口求助，并相信对方一定会支持自己，而非利用这些来攻击自己。

也是在最初的一次谈话中，考利与“问题警官”杰克·科伊尔坐到了一起。此时，这位新局长已经知道了内部调查的结果，那些针对科伊尔的严重指控是不成立的，不过，他也确实存在少量违规行为，比如给自己的私家车窗户非法涂色。这些违规行为都不严重，但加在一起足以解雇这名年轻的警官。考利局长本可以看着科伊尔警官说“你是个低信任的低业绩者”，然后将他扫地出门。但考利局长怀疑，真正有害的可能是这里的文化，而不是这名警官本身。如果自己的目的是改变这种文化，那么再给这名警官一次机会似乎再合理不过了。在许多有限思维的领导者看来，这位局长的决定太过冒险，为什么要留住一个已确认是低业绩且不值得信任的人呢？最终，考利没有解雇科伊尔，只罚他停薪停职三天。科伊尔记得考利当时是这么对他说的：“这是你扭转局面的机会，”科伊尔警官笑着继续道，“他等于是告诉我，‘我相信你……’这份工作几乎是我当时唯一拥有的东西了。我已经失去了其他的一切……因此，我的想法是，‘都这样了，那就让我们试试吧！’”

这次谈话后，科伊尔警官也用行动表明，他知道自己该做什么。如果他的上司想要营造一种彼此信任的文化，就必须以一种值得他人信任的方式行事。真正彼此信任的关系需要由双方来共担风险。就像约会或交朋友一样，尽管有一个人必须率先承担信任对方的风险，但若想让这段关系有任何成功的可能，那么另一个人也必须在某个时刻用信任来回报对方。若要在一个组织中建立安全圈，领导者就有责任做那个率先承担风险的人。不过，接下来的成败取决于员工愿不愿意冒险踏进这个安全圈。领导者不能强迫任何人踏进这个圈子。即使在强大的、彼此信任的团队中，仍会有一些人抗拒进入，尤其是在那些长久以来都坚信业绩优先于信任的团队中。这并不意味着这些成员就是对团队不利的，他们只是需要更多的时间。真正的信任是需要花时间培养的，有些人只是需要比其他入更长的时间而已。

建立信任的过程是有风险的。我们会先冒一点风险，当感觉安全时，就会愿意冒更大的风险。有时也会有失误，那我们就再试一次，直到我们终于觉得可以完全坦然地做自己。信任离不开持续、主动的培养。给科伊尔警官第二次机会，让他努力在更健康的文化中自行扭转局面，这对考利局长来说只是个开始。他还会亲自持续地参与科伊尔的成长，时而教导他，时而确认他的状况，并密切关注他对自己工作的看法。他还会确保科伊尔警官的直属上司也像他这么做。考利局长还会让科伊尔对自己的行为负责，并为他提供一个可以安心表达自己感受的空间，不用担心受到羞辱、奚落或惩罚。反过来，科伊尔也必须利用考利正在构建的这个安全空间来分享他的感受，并在必要时开口求助。考利也希望科伊尔的行事方式能够与该组织的价值观一致。考利所做的一切奏效了。

如今，岩堡市警察局的文化已经彻底改变。这里，信任在自由地流动。杰克·科伊尔也已经成为该警局最受尊敬、最受信任的警官之一，新成员的培训也交给了他来负责。至于一直在寻求真相的考利局长，他仍在坚持与下属们的一对一恳谈。

真相不该成为伤人的武器

保护自己是人类的天性。我们会避开危险，寻找能让我们觉得安全的地方。最安全的地方就是在信任的人身边，知道对方会保护我们。而最令人焦虑的地方就是会令人感觉孤立无援的地方，甚至连队友都会伤害我们，我们必须自我保护。当危险存在时，无论是真相还是幻觉，我们都会出于恐惧而行动，而非因自信而行动。想象一下，若我们一直深陷在恐惧的情绪之中，恐惧无法升职，恐惧可能陷入困境，恐惧可能被嘲笑，恐惧可能无法融入团队，恐惧领导可能不看好自己，恐惧自己的名字可能出现在新一轮的裁员候选名单上，那我们会作何反应。

恐惧具有非常强大的驱动力，可以迫使我们做出完全不利于自身或

组织的行为，可以迫使我们做出最有限的选择，而这些选择可能会造成无限的伤害。面对恐惧，我们会隐瞒真相。在任何情况下，这种做法都是非常糟糕的，若该组织原本的处境就已经很糟，那这种做法无异于雪上加霜。2006年，艾伦·穆拉利（Alan Mulally）成为福特公司的新任首席执行官，这正是他当时所面对的局面。

福特公司当时深陷困境，聘请穆拉利也是希望他能拯救公司。正如考利局长在岩堡市警察局里所做的，穆拉利在福特做的第一件事就是，从公司员工那里尽可能详细地了解公司的一切现状。然而，事实证明，这项任务比他预想的要困难得多。

为了充分了解该公司的状况，穆拉利引入了每周商业计划审查。他要求所有高管参会，陈述自己对公司战略计划的执行情况，并采用简单的颜色标记法——绿色、黄色和红色。穆拉利知道公司目前问题严重，然而，一周又一周过去了，每位高管的项目执行情况标记的全是绿色，这令他很吃惊。终于，他沮丧地直言道：“我们今年将会有数十亿美元的亏损，难道公司就没有哪里是有问题的吗？”没人回答。

高管们的沉默是有原因的，他们非常恐惧。穆拉利的前任首席执行官经常会痛骂、羞辱或解雇那些说了逆耳之言的人。人们习得的是能为自己带来奖励的行为，为了保护自己不被首席执行官责罚，这些高管已经条件反射地选择了隐瞒有问题的地方，或隐瞒未能达到的财务目标。穆拉利说他想要的是诚实和责任感，但这话完全没用，在让高管们产生安全感之前，他是得不到他想要的诚实与责任感的。尽管愤世嫉俗者认为，不应该将私人情绪带入工作之中，但这一屋子都是大公司的高管，事实证明就连他们也会因为恐惧而不愿向首席执行官说出真相。但穆拉利很坚持。

在随后的每次会议上，他都会重复同样的问题，终于有一天，有一个人把自己的一个项目执行情况标成了红色，这人就是福特美洲地区运营部主管马克·菲尔兹（Mark Fields）。他以为这个决定会令自己丢掉工

作，但事实上没有，他也没有被当众羞辱。相反，穆拉利则鼓起了掌，说道：“马克，非常清晰！谁能帮马克出主意？”在下一次的会议上，马克仍然是唯一使用了红色标记的高管。事实上，看到菲尔兹还在，其他高管都很吃惊。一周又一周，穆拉利都会重复他的问题：“我们目前仍然在大量亏损，有没有哪里出了问题？”高管们的标记中终于慢慢开始出现了黄色和红色，他们终于可以开诚布公地探讨自己所面临的一切问题了。在这个过程中，穆拉利学到了一些有助于增进团队信任感的小技巧。比如，为了不让高管们感觉受到了羞辱，他会在言语中将他们所遇到的问题与他们本人分开。“你是遇到了一个问题，”他会告诉他们，“但出问题的并不是你。”⁴

随着会议上的标记颜色越来越丰富，穆拉利终于看清了公司内部现状，这也意味着，他可以主动为下属们提供所需的支持了。安全圈一旦建立，彼此信任的团队就会随之形成，高管们也就可以，用穆拉利的话来说：“作为团队齐心协力，把红色变成黄色，把黄色变成绿色”。⁵他知道，只要他们能做到这一点，就能拯救这家公司。

任何东西都不可能永远以百分之百的状态运转，人也是一样。如果在面对挑战时，我们无法坦诚相待、彼此依靠、相互帮助，就不可能走得长远。但对领导者来说，仅仅创造一个可以安心讲真话的环境是不够的。我们必须将我们所鼓励的行为树立为榜样，用奖励激励人们采取有助于增进信任的行为，为他们提供获得职业成功所需的支持，并且在赋予他们自由的同时，让他们为自己的行为负起责任。我们的价值观与行为方式共同决定了整个企业的文化，文化=价值观+行为。⁶

文化=价值观+行为

要建立以信任为基石的文化需要付出大量的努力。第一步就是创造一个能让人们感到安心和自在的空间。我们必须改变自己的思维模式，认识到必须先有衡量信任与业绩的标准，才能评估某人在团队中的价

值。这或许就是考利局长在改造岩堡市警察局时所采取的最有力的举措之一。原本的文化迫使人们只顾满足数字指标，后来的动力则能激励大家彼此照顾，并更好地为社会服务。不过，考利局长也知道，要做到这一点，就必须改变他对下属认可和奖励的方式。

如今，岩堡市警察局的评估主要考察的是，他们解决了什么问题，以及他们对同事、居民的生活产生了什么影响。传统的衡量指标仍在沿用，但不再是评估的重点。除了书面的评估报告，考利有时还会在点名时，给那些最能体现该警局价值观的警官颁发证书，以表示对他们的认可。

意料之中的是，与传统衡量指标相比，考利局长更提倡和认可对团队成员与社区的关心、个人的主动性及解决问题的能力，因此他看到了更多的关心、更多的主动性与更多被解决的问题。我们习得了能为我们带来奖励的行为。岩堡市警察局的警官们解决的问题越多、表现出的主动性越强，警队与社会之间的信任感也就越强。考利局长称之为“逐个执法”（one-by-one policing），因为所获得的益处都是由他们一次一步、一次解决一个问题这样积累而来的。这一制度提倡的是持续而非强度。

领导者的行为若能让员工产生心理上的安全感，员工就会信任他们。而这种行为是指，赋予员工自主权，让他们运用自己的专业技能自行决定该如何完成本职工作。给员工自由，但要求他们为自己的行为负责。考利局长解释道，在旧的制度中，警官们被告知的是“去做一、二、三、四，然后如此循环往复”，而在新制度中，若发现了问题或机会，警官们会说“如果.....不是很酷吗”，考利局长也会真的放手让他们去实践自己的想法。

这就是“逐个执法”理念的核心。优秀的领导者与彼此信任的团队能让成员们以最佳的状态完成工作。最终，新的文化形成，鼓励的是解决问题而非掩盖错误。其中的区别在于，同样是一个事故多发的十字路

口，你是选择开很多很多的罚单，还是首先研究该如何减少事故的发生。该文化也会阻止过分狂热的执法行为，因为这种行为也源自有所偏颇且以指标为重的评估与认可体系。

例如，警局的自行车分队在得知镇上有一条未使用的自行车道后，发现了机会。他们主动发布了要为孩子们举办活动的消息，孩子们只要带上自行车即可参与。在这里，孩子们可以学习骑自行车，可以在自行车道上骑行，还可以与警官们一起享用免费的甜甜圈——他们称这个活动为“泥土、骑行和甜甜圈”。警官们带着自己的自行车，以及当地一家商店捐赠的甜甜圈，摆了一张桌子，等着孩子们的到来。第一次活动时，他们本以为不会有什么孩子出现，结果却来了40多个，而在之后的每个月，参与者的数量都一直维持在这个水平之上。“泥土、骑行和甜甜圈”成为他们融入社区的一大机会。对大多数人来说，只有出了事，或想摆脱麻烦时，才会找警察，才会与警察交谈。过去，警官们可能会去当地学校做演讲，但这些活动往往是一次性的，他们希望能有更多了解孩子，以及让孩子了解他们的机会。在“泥土、骑行和甜甜圈”这一活动中，警官们不做演示，不做要求，只是与孩子们一起骑自行车而已。

请记住，无限的游戏需要无限的策略。犯罪是一场无限的游戏，因此，考利局长下辖警官们所采取的方法远比“进攻并征服”的思维方式更适合这个游戏。这场游戏的目标不是要获得全胜，而是要保持自身意志的强大、资源的充足，同时还要努力挫败其他玩家的意志，尽量耗尽他们的资源。警方永远无法“击败”犯罪，却可以令罪犯更难犯罪。在考利局长领导下的岩堡市警察局，警官们正在研究新的策略，希望这些策略可以更轻松、更安全且成本更低地重复采用，直到永远。

“警察的主要工作并不是打击犯罪，而是提升人们的生活质量，”考利局长解释道，“那警察自己的生活质量呢？”如果他们痛恨自己的工作，光是“去上班”这个动作就要耗费他们的大量精力，那么势必会打击他们的信心，削弱他们的判断力。“如果遇到一个坏脾气的警察，你可

能就完蛋了。”一名警官解释道，“如果他今天本来就过得很糟，你再给他雪上加霜，或者给他增加工作量，那你可能就要受到顶格的处罚了。”正如尔萨钻井平台那个例子告诉我们的，如果你从事的是一份会有生命危险的工作，那么为员工创造一个可以安心敞开心扉的空间就不是简单的锦上添花，而是生死攸关的大问题。

如果一名警官每天去上班时都觉得干劲十足，工作中能与同事彼此信任，还有一个可以安心表达自身感受的地方，那么公众能从与该警官的互动中获益的可能性也会大大增加。一家无法让员工热爱的公司是永远无法获得客户喜爱的，同理，若警官们无法信任彼此，无法信任自己的领导者，那么民众也永远不会信任他们。

岩堡市警察局的做法是，通过改变内部文化的侧重点来应对外部挑战，而此举已经为该警察局75位宣誓就职的警官带去了显著的变化。美国约有12500个警察局，其中超过95%的警局警官人数不足100人，对那些可能苦于内部缺乏信任或与社区间缺乏信任的警局，“逐个执法”的成功案例不失为很好的示范。

当然，考利局长也承认，要改善警局的状况仍需开展大量的工作，旧的思维方式也尚未完全消除。但是他们在前进，他们的文化也比过去健康了很多很多。听警官们说，他们在社区执勤时，只为说一声“谢谢”而挥手叫住他们的居民多了很多，在咖啡店里请他们喝咖啡的人也多了很多。犯罪率得到了控制，社区居民也更愿意出手帮忙。“在社区居民眼中，我们是帮他们解决问题的人，”考利局长说，“而非只是执法者。”

无论哪个行业，领导者若给员工施加过大的数字指标压力，采用失衡的激励机制，就有可能创造出这样一种环境：优先考虑短期的业绩和资源，导致业绩、信任、心理安全感和员工意志的长期下降。这在警界是真理，在商界也是真理。如果从事客服工作的员工压力过大，他们就更有可能会给客户带去糟糕的服务体验。情绪会影响工作方式，这是早就

众所周知的事。如果工作的环境会让人们产生恐惧——害怕陷入麻烦、害怕被羞辱或害怕被解雇，人们就会觉得自己需要用谎言、隐瞒和弄虚作假来掩饰自己的焦虑、错误或在培训上与他人之间的差距，这样的环境会从根本上阻碍人们建立信任。在警界，该影响可能造成的结果会比糟糕的客户服务严重得多。

在脆弱的文化中，人们依靠规则获取安全感。这就是官僚主义者出现的原因。官僚主义者相信，只要严格遵守规则，他们就能保住工作。在此过程中，他们会做出有损组织内外信任的行为。而在强大的文化中，人们依靠与他人间的关系来获取安全感。良好的人际关系是团队高效运转的基础。所有高业绩团队的产生都是从信任开始的。

不过，在今天这场无限游戏中，我们需要的远不止强大的、彼此信任的、高业绩的团队，我们需要建立一种制度，确保这种信任与这种业绩表现能够一直持续下去。如果说，创造出有助于培养信任的环境是领导者的职责，那么我们是否也应该培养一批知道如何做到这一点的领导者呢？

培养领导者

美国海军陆战队的候补领导者都要在位于美国弗吉尼亚州的候补军官学校（Officer Candidate School）参加为期10周的培训和选拔。在众多考验中，有一项“领导力反应课程”（Leadership Reaction Course）。领导力反应课程由20个迷你障碍课程组成，更准确地说，是问题解决课程。这些海军陆战队士兵要四人一组完成各项挑战，比如在规定时间内，只用三种不同尺寸的木板，将所有人员与军备物资运过水障碍区（“池塘”的军事术语）。海军陆战队会利用这门课程来评估候补军官的领导能力。他们会考察候选人服从领导者或应对逆境的能力，也会考察他们快速理解当前局势、确定优先顺序并分配任务的能力。令人惊奇的是，所有这些对未来领导者的能力评估中，并没有“成功跨越障碍”这一

项。就连评估表最下面都没有一个确认障碍挑战是否完成的选项。换言之，海军陆战队的评估重点是士兵们的投入、行为，而非结果。他们这样做是有充分理由的。他们知道，优秀的领导者偶尔也会有任务失败的时候，而糟糕的领导者偶尔也能享受到任务成功的喜悦。成功的能力并不能使一个人胜任领导职位。能够展现出领导才能，才能让一个人成为卓有成效的领导者。正如海军陆战队通过多年试错而学到的那样，诚实、正直、勇敢、坚毅、复原力、判断力和决断力等品质和能力才更有可能带来信任与合作，而这种信任与合作会在未来更长的时间里，提升团队的成功概率，减少他们失败的次数。偏好意志多过资源，偏好信任多过业绩，这样的思维方式能够增加团队持续发挥出更高水平的可能性。

对任何组织来说，培养新领导者的能力都至关重要。你可以把一个组织想象成一株植物。无论这株植物多么强壮，长得多高，只要无法产生新的种子，无法培养新的领导者，就完全不存在持续繁荣的能力。对所有领导者来说，他们的首要任务之一就是培养新的领导者。而这些新人应该学到的是，如何让该组织适应和参与无限游戏。若现任领导者更关注如何让自己的植物长得尽可能大，那么，这个组织就会像杂草一样，不择手段地生长，完全不考虑会对所在的花园产生何种影响（甚至不顾它自身的长期前景）。

我知道有许多人虽然坐到了一个组织的最高位，但不能算是领导者。他们或许有正式的职位，我们或许会因受他们管辖而服从他们的命令，但这并不意味着我们一定会信任他们或追随他们。还有一些人可能没有正式的职位或权力，但他们愿意为照顾下属而承担风险。他们能够创造一个安全的空间，让我们做真实的自己，并安心说出自己的想法。这些人才会得到我们的信任，我们愿意追随他们去任何地方，愿意为他们付出更多的努力，做这一切的理由不是“不得不”，而是“我愿意”。

海军陆战队感兴趣的并不是领导者能否穿越水障碍区，或能否完成

其他任何障碍挑战。他们感兴趣的是让领导者习得这样一种能力：创造一个良好的环境，让所有成员都能彼此信任，从而携手合作，克服障碍。海军陆战队的士兵都知道，以信任为基石的领导氛围才是能让他们成功多于失败的保证。

我将在本书中不止一次地提到这样一句话：领导者的负责对象不是结果，而是那些对结果负责的人。在一个组织中，提高业绩的最佳方法就是创造一个良好的环境，让信息可以自由流动，错误可以高亮展示，帮助可以有来有往。简言之，就是能让成员在团队中感到安全。这是领导者的职责。

里克·福克斯就是这么做的。他为自己的团队成员创造了一个良好的环境，让他们可以安心地彼此袒露脆弱，由此，他建立了一个高业绩的团队。海豹突击队也是这么做的。他们认为个人的可信赖程度比他们创造业绩的能力更加重要，由此，他们建立了许多高业绩的团队。艾伦·穆拉利也曾这样做。他先为下属创造了一个可以安心说出公司真实现状的空间，随后才使福特重新成为一个高业绩的公司。这也是杰克·考利当下正在做的事，且已经让警局有了明显的改善。当领导者愿意将信任的重要性置于业绩之上时，业绩往往也会随之提升。而当领导者过度强调业绩重于一切时，该组织的文化势必遭受重创。

第8章

道德的堕落

难以想象会发生这样的事。无论从何种角度来看，这都是极其不道德的。难以想象这样一群人居然能做出这样的事，我相信，他们都自认为是善良且诚实的人，但他们的行事方式无论用什么标准来衡量，都是大错特错的。

从2011年年中开始，大约到2016年年中这段时间，美国富国银行（Wells Fargo Bank）的员工开立了超过350万个虚假银行账户。正如2016年《纽约时报》的报道：“一些客户发现了这个骗局，因为他们或是被收取了意料之外的费用，或是收到了他们没有申请过的信用卡或借记卡，或是因他们不知道的账户而被催债。但大多数虚假账户都没有被发现，那些员工通常会在开户后不久就销户。”¹

最终，富国银行有5300名员工因参与这些欺诈而被解雇。²时任首席执行官约翰·斯坦普夫（John Stumpf）告诉美国国会，这些行为“完全违背了我们的核心原则、道德准则和企业文化”。该公司在发给媒体的一份声明中，给出了与斯坦普夫一致的说法：“我们团队中的绝大多数成员每天都在代表客户做着正确的事……如果真的发生了这样的事，我们深感痛心，这绝对违背了富国银行的原则。”换言之，富国银行的高管想让我们相信，公司只有少数几颗坏苹果而已。然而，这并不是一小群人的孤立行为，而是上千人多年犯罪所造成的后果！富国银行的企业文化很可能已在这一严重的道德堕落事件中遭到了破坏。

道德的堕落指的是，人们为牟取私利而采取了不道德的行为，但误以为自己的做法并没有违背自己所在文化中的道德准则，而且这样的行为往往是以牺牲他人的利益为代价的。道德的堕落往往始于看似无伤大

雅的越轨行为，如果不加以制止，就会愈演愈烈，不断恶化。

尽管任何地方都有可能出现道德失范行为，但有限思维的组织尤其容易出现道德的堕落。正如前几章讨论过的，过度强调季度或年度财务业绩的文化会给人们施加巨大的压力，迫使他们为达到业绩目标而不惜走捷径、违反规则或做出其他可能不道德的决定。

不幸的是，有些人虽然方法有不道德之嫌，但只因完成了目标，依然能得到奖励，这就将一个组织的优先顺序暴露无遗。实际上，这些组织的奖惩制度恰恰在激励这些行为。这些组织的奖惩往往只看结果，并不考虑用以完成目标的行为是对是错，因此奖金和晋升机会都流向了那些完成了目标的人，而那些诚实正直却未能完成目标的人往往无法获得认可和晋升。这向该组织中的每个人都传递了一个信息：漂亮的业绩数字比在工作中遵守道德准则更为重要。那些本不愿效仿同事不道德做法的人，一旦发现那才是获得奖金、升职，甚至是保住工作的唯一途径，便也会屈从于压力。他们会丧失客观的判断力，将自己的道德越轨行为合理化。合理化的借口有很多，比如“我得养家糊口啊”“这不就是管理层想要的吗”“我别无选择”“这就是行业标准”，这些都是我们为了减轻自己的罪恶感或责任感说给自己或别人听的。

理性思考的能力既是上天对我们的恩赐，也是上天对我们的诅咒。我们一直在设法理解周遭的世界，理解复杂的方程式，也具备自省的能力。正是有了理性思考和分析的能力，我们才能解决难题并发明出更先进的技术。不过在得知自己的行为违反了一些根深蒂固的道德准则时，我们也可以利用自己的分析思维能力来正当化自己的行为，或为自己的行为辩解，从而减轻可能因此而产生的罪恶感。这就像从一个有钱的朋友那里偷了东西，却告诉自己“他们根本注意不到的。再说了，他们也不缺这点钱，再买一个就是了”。我们可以用任何借口来合理化这一行为，却改变不了偷东西的事实。当这种合理化在一个组织内司空见惯时，雪球就会越滚越大，直到整个组织都充斥着不道德的行为，极端情

况下，还会导致腐败。

压力、需求与激励的文化

1973年，普林斯顿大学心理学教授约翰·达利（John M. Darley）和丹尼尔·巴特森（C. Daniel Batson）做了一项实验，目的是进一步了解情境变量会对道德准则产生何种影响，尤其是压力会如何影响我们帮助有困难者的意愿。他们让一群学生到校园各处讲故事，故事讲述的是有一个人被殴打后躺在路边，路过的人都没有停下来帮助他，最后只有一个人停下来救了这个伤者。

为了重现这一幕，教授们雇了一位演员，弯身蜷在小巷里，就像被抢劫或被打伤了那样，学生们在穿越校园时必定要从他身边经过。学生被分成不同的小组进行实验，每次实验，教授们都会给学生施加一定的压力，观察压力对行为的影响。这一压力的作用是迫使他们快速穿过校园，第一组学生的压力非常大。“你们都迟到了！”陪同前往的实验人员催促学生，“助理正等着呢，你们几分钟前就该到了。我们最好快点。助理应该在等着了，你们赶紧啊，最好一分钟内就能赶到。”第二组则是中等压力：“助理已经准备好了，就等你们了，请直接过去。”第三组的压力很小：“助理还要再准备几分钟，不过，你们最好现在就过去。如果你们到了，他们还没准备好，也应该不用等太久的。”

当压力很小时，有63%的学生停下来救助伤者；当压力中等时，有45%会停下来；而当压力很大时，就只有10%的学生停下来帮助那个明显处于困难中的人。有些人甚至会直接从他身上跨过去。结论一目了然。这些学生都是善良的人，关心公共事业。然而，在压力（在这个案例中是时间的压力）之下，他们做正确之事的意愿会让步于听从教授们要求的意愿。而在富国银行的案例中，那些销售部员工都工作在极端的高压之下。

尽管对那些无论外界压力大小，都能够顺利完成业绩目标的人来说，公司确实提供了很多正向的激励措施，但对那些无法完成目标的人来说，恐惧会逐渐在他们的心中滋生。一些员工回忆称，他们被迫每天销售8到20种不同的产品，若完不成目标，就会被经理痛骂。³一名员工还记得经理曾对她说：“你要是完不成你的方案，就是没有团队精神。你要是连累了整个团队，就会被解雇，而且会被永久记录在案。”这名员工告诉她的主管们，她觉得在道德允许的范围内，自己根本不可能完成他们设定的目标，她还曾多次拨打银行的道德监管热线反映这个问题。若有证据表明一个大型组织内存在道德失范行为，她的应对方法才应该是我们希望或期待在员工身上看到的。但富国银行的最终决定不是回应她的担忧，而是解雇了她。富国银行的期望是，员工永远不会说出公司的指标不可能完成这种话，他们只希望员工能想方设法完成指标，哪怕不择手段。正如富国银行另一名员工所承认的：“用不道德的方法提升销量是一种常态。他们就是这么教我们的，我们只是照做而已。”⁴

道德的堕落不是一次性的事件，并不是开关一开，就会突然出现的。它更像是一个被感染的伤口，拖得越久，溃烂越严重。对该丑闻的调查发现，在爆出该丑闻的10年之前，富国银行就进行过一次内部审计，当时就已发现该组织内部确实存在不良的工作氛围和有违道德的做法。⁵当初的审查结论是，对失业的恐惧会“鼓励舞弊行为”。尽管公司的首席审计官、人力资源代表及其他高管都收到了这一审查结果，但领导层并没有采取任何措施来纠正它。不仅如此，截至2010年，也就是虚假账户骗局开始的一年之前，该公司就已累计收到了700份检举报告，称公司内存在疑似有违道德的销售策略（但董事会报告称，他们对此一无所知）。⁶约翰·斯坦普夫早在2013年就发现公司存在系统性问题。更惊人的是，2017年的一份董事会报告称，斯坦普夫早在2002年就知道有些人的做法有问题，这个时间比丑闻爆发早了近15年！⁷这份报告还指责富国银行社区银行前负责人卡丽·托尔斯泰德（Carrie Tolstedt）不仅知道不当销售行为的存在，还“强化了这一高压的销售文化”。该报告

称，她甚至“公然阻碍外部的干预与监督”，并联合其他高层一起“抵制和阻碍审查”。⁸人们只能猜测，要么她也承受着类似的高压，只是不敢说出来，要么她所在部门的业绩成果让她获得了丰厚的奖励。

富国银行在公开声明中称，这起丑闻仅限于零售业务部，公司的大多数人“都在做正确的事”，但有大量证据表明，道德的堕落已经在这个公司中蔓延开来，影响深远。例如，在与虚假账户丑闻有所重叠的时间段内，该银行出现过故意谎报所售贷款质量的问题。因这一问题，富国银行在2018年被罚款20.9亿美元。⁹该银行的汽车业务部还曾卖给一些客户他们并没有签字申购的汽车保险，并因此赔偿客户总计8000万美元。¹⁰批发业务部也曾因其他的道德失范行为而遭到调查。¹¹该部门当时的负责人是蒂姆·斯隆（Tim Sloan），他后来接替约翰·斯坦普夫成为富国银行的首席执行官。

最终，富国银行因开设了数百万虚假账户而被问责，并被处罚1.85亿美元。然而，撇开一时的尴尬与对股价的短期影响不谈，他们受到的处罚其实是很轻微的。客观地说，1.85亿美元还不到富国银行当年220亿美元总利润的1%，也只占其近950亿美元总收入的0.2%。¹²这相当于一个年收入7.5万美元的人被罚了150美元。这压根不算什么惩罚。

富国银行纵容了这种会驱使员工欺诈（欺诈属于犯罪）的文化存在，却没有任何领导者因此而被追究刑事责任，甚至连一份起诉书都没有收到。约翰·斯坦普夫确实因此丢了工作，损失了4100万美元尚未行权的股票，但解雇他也只是迫于巨大的舆论压力。不仅如此，他离职时还获得了总值超过1.34亿美元的养老金和股票。¹³由此可知，对于那些负责监督企业文化的领导者来说，他们不仅不会因该文化中道德的堕落而受到惩罚，还能从中获取切实的利益，这会激励领导者保持现状。在我看来，高管们能因“业绩文化”而获得褒扬，却不用对一个被道德堕落所腐蚀的文化负责，会是一件非常可怕的事。

好人做坏事

有一些人会因接触花生、蜜蜂或贝类出现可能危及生命的严重过敏，这些患者都深知，一针肾上腺素可以挽救他们的生命。“肾上腺素笔”（EpiPen）这一品牌占据了90%的市场份额，你使用的很可能就是这款产品。“肾上腺素笔”是一个肾上腺素自动注射器品牌，该种注射器能够阻止过敏性休克反应。对任何可能出现极端过敏反应的人来说，该产品都是必需品，又因其保质期只有12个月，每年都需更换，所以销售它是非常赚钱的生意。

2007年，一家名为迈兰（Mylan）的公司买下了“肾上腺素笔”这一品牌。当时，该品牌已是市场的主导者，加之没有非注册商标的替代产品，便没有什么能阻止迈兰以每年22%的平均涨幅提高该产品的价格。2014年，董事会在看到涨价对公司股价的影响后，决定提高对员工的要求。¹⁴他们为业绩最佳的员工提供了一个一次性的机会：他们若能在未来5年内让公司的每股收益翻一番，就能分享总计可能达数亿美元的奖金。仅排名前五的高管就能分得总计近1亿美元的奖金。这一激励政策会产生什么样的效果毫无疑问，政策实施的第二年，该公司就将“肾上腺素笔”的涨价幅度从22%提升到了32%，这是该产品自2009年以来的第15次涨价。¹⁵ 2016年，迈兰公司宣布“肾上腺素笔”的价格将创下600美元的历史新高，这意味着该产品的价格在短短6年内就上涨了500%。¹⁶ 要不是遭到了民众的大规模抗议和美国众议院监督委员会的国会问询，该公司可能还会继续涨价。

后来，当被问及是否对所发生的事情感到抱歉时，该公司的首席执行官希瑟·布雷施（Heather Bresch）回答说：“我不会为符合现有体系的运作方式感到抱歉。”¹⁷顺便一提，负责是指我们对自己的行为负责，而不是将自己的行为归咎于体系。

道德在迈兰堕落得如此彻底，布雷施似乎都没有意识到公司做了什

么错事。更令人难以置信的是，布雷施竟然声称，“肾上腺素笔”丑闻是一件好事，因为它唤起了大众对医疗保健体系内权力滥用的关注，并成为变革的催化剂。当然，迈兰的企业文化若能认为道德准则重于收益，若能认为其首要的负责对象是自己的崇高的事业而非它自己或它的股东，那么该公司就能更早地利用其在市场上的强大影响力成为变革的捍卫者，也能少惹很多麻烦。

顺便一提，在“肾上腺素笔”定价丑闻爆发的两年后，迈兰又被罚了，它与美国司法部达成了4.65亿美元的和解协议，原因是将注册商标产品“肾上腺素笔”错误地归类为非注册商标产品，向政府收取了过高的费用。正如美国联邦代理检察官威廉·魏因雷布（William D. Weinreb）的解释：“迈兰公司为从贫困者医疗补助保险（Medicaid）项目中牟利，将其注册商标药物‘肾上腺素笔’不当归类.....纳税人有权期望迈兰等通过国家税收资助项目获得收入的企业能够严格遵守规则。”¹⁸

但我们并不能将好人做坏事完全归咎于有缺陷的激励机制。若是如此，那么做出这些事情的人，按道理应该会内疚不已、彻夜难寐。然而，一切证据都表明，他们似乎完全不为自己的选择而感到不安。研究道德堕落现象的社会科学家称，那些背弃信任的人并不是邪恶之徒，但他们确实在自我欺骗。

自我欺骗

人类很聪明，有各种各样的方法可以合理化自己的行为，可以自我欺骗。明明是在任何理性的人眼中都不公平、不正当的行为，我们却可以说服自己相信这就是公平且正当的。美国圣母大学商业伦理学教授安·特恩布伦塞尔（Ann Tenbrunsel）和美国西北大学凯洛格管理学院管理与组织系荣誉教授戴维·梅西克（David Messick）对将自我欺骗作为组织内道德堕落的一种机制进行了研究。他们在研究中发现，有几种极其简单又常见的方法，可以让我们在采取不道德行为时，不认为是不道德

的，这些方法无论对个人还是对团体都适用。¹⁹

自我欺骗方法之一源自我们的用词，确切地说，是委婉语的使用。我们的决定或行为可能会产生令我们感到反感或难以忍受的影响，委婉语则可以让我们免于受到这些影响。美国人认为酷刑是不人道的，是违背美国价值观的，政治家们知道这点，因此为避免内疚，“9·11”事件发生后，他们选择用“强化审讯”（enhanced interrogation）的方法来保护家园。

在商界，我们也是这么做的。选词时故意弱化或模糊自身行为所带来的影响，这在职场上是很常见的。我们会说要管理“外部性”，而不会直白地谈论“我们的制造工艺对工厂员工及环境造成的伤害”。与“我们找到了可以让消费者对产品上瘾，从而提升销量的方法”相比，自然是“通过游戏化来增强用户体验”更容易让人接受。人类变成了“数据点”，“数据挖掘”一词则更直白地说出了我们正在追踪人们的每一次点击、每一次旅行和每一个个人习惯。我们会说“减少了人数”，在线票务经纪人会说向我们收取的是“服务费”而不是“附加费”。

用词选择可以帮我们远离任何责任感，也可以督促我们采取更合乎道德的行事方式。试想一下，如果我们实事求是地称呼组织内的一切将会如何，或许我们就会花更多时间去寻找更有创意且真正更合乎道德的方法来实现我们的目标。在这个过程中，我们的文化也能得到正向的强化。

另一种会导致道德堕落的自我欺骗是，我们会将自己从因果链中抽离出来，或像迈兰首席执行官所做的那样，把自己的越轨行为归咎于“体系”。有时，我们甚至可以把自己从因果链中摘得一干二净，将产品影响消费者的所有责任全部转嫁给消费者。尽管“货物出售概不退换”是一个合法的法律概念，但企业往往会用它来撇清自己与所做决策之间的关系。这一思维是，消费者如果不喜欢这个产品，大可不必购买。高管们在被质疑应为其产品所产生的负面影响负责时，常常会以此

作为回应。消费者的选择当然是一个因素，但这无法也不能将一个组织从因果链中完全抽离。比如，就吸烟者健康受损这个结果而言，吸烟者自身当然有责任，但烟草公司也脱不了干系。

企业履行了法律责任并不等于不用承担道德责任。比如，许多企业认为，一旦我们勾选了选项，接受了他们提出的条款和条件，他们就可以对接下来发生的一切免责。在法律上可能确实如此，但在道德上绝对不是。例如，现在越来越多人认为，Instagram、Snapchat、Facebook以及不计其数的手游公司都使用了成瘾性的技术，而这些公司不能仅仅因为目前还没有法律禁止他们研发此类技术，就否认他们在这些技术的研发中所扮演的角色。尤其对于他们故意增加了一些能诱导我们延长浏览时间的功能，比如无限下拉、点赞按钮和内容自动播放。这些公司的解释几乎千篇一律，他们之所以新增这些功能或收集我们的个人数据，只是为了“增强用户体验”。我们或许确实能从这些决策中得到一些好处，但同时也得付出代价。提醒我们要权衡其中的利弊，或者要判断它们是否有违我们的价值观，这些就是道德准则的全部意义！天下没有免费的午餐。

在《华盛顿邮报》2019年刊登的一篇评论文章中，Facebook创始人兼首席执行官马克·扎克伯格（Mark Zuckerberg）就一些针对该公司的批评给出了回应，他的回应是要求政府出台更多法律。“我认为，政府和监管机构需要扮演更积极的角色，”他说道，“我们可以通过更新互联网规则来保留其中最好的部分。”他的意思仿佛是，基于米尔顿·弗里德曼对企业责任的定义，只有在法律和“道德习惯”要求他们遵守道德准则时，他们才需要这么做。可悲的是，在某些行业，比如科技行业和社交媒体行业，我们可能已经走到了必须通过立法来维护道德的地步。我们沦落至此的开端究竟是什么呢？

特恩布伦塞尔和梅西克认为，众所周知的滑坡效应也是促使人们自我欺骗、导致道德堕落的因素之一。²⁰我们对道德越轨行为的每一次容

忍，都是在为更多且更严重的道德越轨行为铺路。就这样，我们的文化中对可接受行为的标准也会一点一点地改变。“如果其他所有人都在这么做，那这个行为肯定没问题。”

当领导者过分关注有限游戏时，就会因有利可图而注意不到或刻意忽略滑坡效应。在一个无限思维的组织中，任何旨在提高利润的不道德想法都会被视为“拿着一根十米长杆也不敢碰的坏主意”。在一个痴迷有限游戏、深受堕落道德腐蚀的组织中，同样的想法则会被视为“太棒了，真不敢相信我们居然早没想到！”再加上奖励机制的不平衡，注重绩效而忽视信任，道德的滑坡就会开始加速，这就仿佛是一个浑身涂满了婴儿油的人遇到了涂有特氟龙涂层的滑梯，直到最后道德完全堕落。

就像温水煮青蛙的谚语所言，迈兰公司逐步提高“肾上腺素笔”价格的行为无疑也是为了减轻价格突然暴涨对消费者的冲击（或增加消费者的接受度）。而这也揭示了职场上的道德堕落。通过逐步（甚至是在短时间内）抬高价格，就能实现业绩指标的飙升。随着业绩数字的攀升，许多人或许已经在盘算该怎么花之后能分得的奖金了。迈兰的高管们完全专注于他们个人可以获得的巨大好处，因而不会对自己的决策产生任何道德上的负疚感，他们甚至为了更快地完成目标而加快了提价的速度。

迈兰和富国银行都是道德堕落的极端例子。这些极端例子能帮助我们了解道德堕落在职场中的发生过程。但不要被欺骗，也不要掉以轻心。没有骗局或丑闻发生并不意味着就没有问题。事实上，如果我们仔细观察，就会在许许多多的企业中发现道德堕落的迹象。比如，为帮公司减税而使用的会计诀窍。还有，为某款产品提供返现，但故意给顾客设置太过烦琐的步骤，以至于大多数顾客都会因为嫌麻烦而自动放弃返现，而这都在该公司的意料之中。有的食品和饮料公司会夸大产品的健康益处，设法隐藏一些不健康的配料，或者修改包装上的分量大小，让该产品的含糖量或卡路里含量显得没有它的实际含量那么多。我们越纵

容这些企业决策，这类行为就会越“正常化”，或逐渐成为“行业标准”。

请记住，道德的堕落与自我欺骗息息相关。任何人，无论自身的道德感如何，都可能会屈从于它。一些领导者用不道德的方式经营企业，并因此收获了巨额回报，在被我们指出并批判后，他们仍不认为自己做错了什么。如果自己都不认为自己做的事有什么错，又怎么会有动机去改变呢？在迈兰和富国银行这样的案例中，问题都是由公共丑闻暴露出来的。但聚光灯能暴露问题，却无法解决问题。此类危机可以帮助企业看清一些肮脏的真相，但大多数企业都不会遇到此类危机。若不对道德的堕落加以控制，那么很可能终有一天会有什么东西走向毁灭。最终，不仅是我们的公司，还有我们的员工、客户和投资者，都将为此付出代价，那个代价将远远超过我们现在为纠正自身错误而需承担的成本。

蒂姆·斯隆在接任富国银行首席执行官时就已承认，管理层“实在是太晚才认识到了这些问题的影响范围之广与严重程度之深”，并发誓此类情况“绝不会再出现”。²¹许下这样的承诺很容易，但要履行就没有那么容易了。道德的堕落是极难逆转的。即便领导者试图改变堕落的文化，但若仍在用有限的思维做决策，那也不可能扭转它了。有限思维的领导者若要着手改变被道德堕落所腐蚀的文化，他们会怎么做？答对了。他们会用有限的解决方法。（友情提示：这样的方法并不管用。）

若制度取代了领导

我曾在一家大型广告公司工作。在公司工作一年后，领导层决定实施考勤表制度。我们公司与律师事务所不同，在那里，律师可能根据自己的实际工作时间向客户收费，所以考勤确实是公司跟踪、记录员工工作时间的一种方式……但在我们公司，其实根本没人知道考勤表有什么用，我们只是奉命行事。

有好几个月，我没填考勤表，但也设法躲过了处罚（如果他们跟踪

了解我的工作情况，自然知道我对分配到的客户是百分之百尽力服务的，也就没必要再报告什么考勤了）。当然，我最终还是因没有上交考勤表而惹了麻烦。此后，我会在每月月底一次性填好所有的考勤表，上班时间一律填上午九点半，下班时间一律填下午五点半。但多数时候，我其实都得更早、走得更晚，但谁在乎呢？我还记得，当我把考勤表拿给老板签字时，他会在检查完后挖苦道：“你可真是个始终如一的员工啊，不是吗？”然后才给我签字。

关于推行该制度的原因，我只能认为是财务上出了问题。或许某个客户被多收了钱，因此要求公司证明，承诺管理他们账目的人真的是他们账目的实际管理人……或者诸如此类。为了纠正财务上的问题，才会在全公司范围内增加了一道新程序。这种解决方案就是伦纳德·王（Leonard Wong）博士所说的“懒惰的领导”（lazy leadership）。

一旦发现存在问题、错误、不道德的决策或业绩增长缓慢的情况，懒惰的领导会努力用程序来解决问题，而不是为员工提供支持。毕竟程序是客观、可靠的，信任程序比信任人更容易，至少我们目前是这么认为的。伦纳德·王指出，事实上，“程序告诉我们的总是我们想听到的”。²²程序会给我们开绿灯，但或许不会告诉我们真相。当领导者用程序取代判断时，就会为道德的堕落创造条件，哪怕是在那些会让人们坚持更高道德与伦理标准的文化中。

比如，军人认为自己恪守着比普通大众更为严格的诚实标准与正直标准，公众也这么认为。然而，伦纳德·王与他的研究搭档斯蒂芬·杰拉斯（Stephen Gerras）研究了美国某些军队团体，并在其合著的论文《对自己撒谎：军队中的不诚实行为》（*Lying to Ourselves: Dishonesty in the Army Profession*）中提到，给美国士兵规定的程序、手续或要求过多，会导致系统性的道德堕落。²³领导层对美国士兵提出的一些要求不是不合理，而是压根不可能做到，比如要求他们完成的训练日数量比日历上的日子还多。

与在企业中一样，美国军队里完成任务的压力也是自上而下的，但这里也存在着自下而上的巨大压力。美国军官们为了脱颖而出，就想表现得好像自己无所不能、无所不精。无法完成要求可能会有损指挥官的形象，招来指挥官的训斥，并影响晋升。提交一份伪造的合规报告有助于保持系统的平稳运行，并令他们的职业生涯一帆风顺。说真话有时会招来比撒谎更严重的惩罚，这种环境会迫使美国士兵觉得，若要满足上级对自己的要求，就不得不撒谎或欺骗。这是个自相矛盾的窘境。

这样做所带来的结果是，美国士兵对用“巧”方法来满足要求的做法习以为常，并不觉得这样做会有违他们比常人更高的道德标准。

许多美国士兵并不认为这样的行为是欺骗或撒谎，他们觉得不过就是“勾些选项”“完成些很官僚的程序”，或者只是在做“领导想让他们做的事情”。有些人根本不认为自己的行为是不道德的，他们觉得这些要求太过琐碎、微不足道，根本不在任何正直或诚实的标准约束之内，就像我对考勤表的看法一样。这就像告诉某人，我不得不取消我们的计划了，因为我有点“家事”要处理，然而这个“家事”根本不存在。我们只是既想退出计划，又不想伤害他人的感情罢了。尽管我们说了谎，也不过是“无害的”善意谎言，我们仍然相信自己是诚实的人。

然而，当一种文化到处都充斥着这些看似微不足道的越轨行为时，就说明道德堕落了。请记住，道德堕落指的是，在从事不道德行为时，仍然相信我们的行事方式符合自己的道德或伦理准则。与在企业中一样，士兵的不道德行为若产生了更为严重的后果，引起了公愤，也是很可能要受罚的（接下来，美国军队当然会为了防止此类情况再次出现而给其余士兵增加更多的线上训练任务）。

这其实是一个巨大的讽刺。当我们尝试用有限思维的解决方案来解决有限思维所造成的道德堕落问题时，只会加剧道德的堕落。当我们尝试用程序和制度来解决文化问题时，得到的往往是更多的谎言和欺骗。小谎变大谎，这样的行为也就成为正常的了。

懒惰的领导不是坏领导者或坏人的委婉语。就像选择不锻炼的人并不是坏人一样。懒惰的领导做出的决策往往是充满善意的。就军队那个案例而言，领导层可能真的认为他们对士兵提出的所有额外要求都是对他们有帮助的，同理适用于任何有类似情况的大型组织。但这些额外要求往往不会加诸高层领导自己身上，因此他们可能意识不到自己的“解决方案”会造成什么问题。不过，即便他们意识到了或者也受害于这些虚伪、功能障碍或过度的官僚主义，他们也有可能成为这场闹剧的同谋，就像我所在广告公司的那个老板一样。当这种情况出现时，这些领导者很可能会一并推动这些行为的合理化，同时也自我欺骗。然后，这个坡就会越来越滑。

如果像美国军队这样如此看重正直这一品质的地方都能出现道德的堕落，那任何地方都不可能完全幸免了。事实也确实如此。道德的堕落在企业和机构中极其普遍，这一点再怎么强调都不为过。但增加制度并不是解决这一问题的良药。程序对供应链管理来说至关重要，清晰的流程有助于提高生产效率。然而，道德的堕落是人的问题。尽管听起来似乎不合常理，但要解决人的问题，我们需要的也是人。

治疗和预防道德堕落的最佳良药就是无限思维。领导者如果能为下属提供一项崇高的事业，并给他们一个可以与彼此信任的团队合作推进该事业的机会，就能形成良好的文化，在这种文化中，人们既会努力实现短期目标，也会同时考虑到自己为实现目标所做的决策是否合乎伦理道德，是否会造成更广泛的影响。这么做不是因为有规定，不是因为检查清单上有这些内容，也不是因为上了公司关于“做事要遵守道德准则”的在线课程。他们这样做是因为理当如此。我们不做有违道德的事，因为我们不希望自己的行为会阻碍崇高的事业的推进。当我们感觉自己身处一个彼此信任的团队中时，我们不想让队友失望。我们觉得自己要对团队负责，要对组织的声誉负责，而不仅仅是对自己和个人的野心负责。当我们觉得团队关心我们时，就会希望代表团队做正确的事，让领导者感到自豪。我们的标准自然就会随之提高。

作为社会性动物，我们会对自己所处的环境做出反应。把好人置于道德堕落的环境中，好人也会更容易做出道德失范行为。同理，把一个人，即使是一个可能过去有劣迹的人，置于更强大、价值观更端正的文化中，此人也会遵守标准，遵守该环境中的规范。正如我之前所说，从定义上讲，领导者应予以负责的不是结果，而是对结果负责的那些人。这是一份需要持续保持警惕的工作，因为一旦小事恶化，最终也会带来毁灭性的结果。

无限思维的领导者认为，要创造出更能抵御道德堕落的文化需要耐心和努力，需要对崇高的事业投入，需要重意志多过重资源，需要具备培养彼此信任的团队的能力。这些方面的投入要体现出成效可能需要花上不止一个季度或一年的时间（取决于企业规模）。一旦道德标准确立（或重新确立），我们就必须时刻警醒地去守护它们。如果出现了会催化道德堕落的自欺行为，就必须通过坚持完全的诚实与坚持持续的自我评估来确保自己不会做出有违道德的行为。道德失范时有发生，它是人类的一部分。但道德的堕落不是，道德的堕落是领导层的失败，是企业文化中的一个可控因素。这意味着，文化中道德的强大也要归功于领导者对文化的塑造。

若遵守道德成为标准

2011年11月25日，户外服装公司巴塔哥尼亚（Patagonia）在《纽约时报》上刊登了一整版的广告，标题是“不要买这件夹克”。一些愤世嫉俗者认为，这个标题不过是一个许多人消费不起的高档品牌耍的宣传噱头。然而，如果你仔细看看这则广告，就能发现一些有助于了解该公司文化的线索，了解他们这则广告的灵感到底从何而来。

在广告的正文中，巴塔哥尼亚做了一件令大多数其他公司都觉得匪夷所思的事。他们坦率承认了自家产品在生产过程中对环境造成的负面影响，在这个案例中，就是指他们的畅销产品R2抓绒夹克。这则广告写

道：

生产这件夹克需要消耗135升水，这足以满足45个人的日常所需（按每天3杯水计算）。从最初的原材料（60%的再生聚酯纤维）到放入我们仓库的成品，这一过程产生了近9千克的二氧化碳，是成品重量的24倍。这件夹克在运往仓库的途中还会产生相当于其重量2/3的废物。²⁴

“未竟之事还有很多，我们还有很多事情要做，”该广告最后总结道，“不要买你不需要的东西。在买任何东西之前都请三思。加入我们……设想创造一个这样的世界，在那里，我们只使用天然可再生的资源。”

“我们这么做是出于内疚。”巴塔哥尼亚创始人伊冯·乔伊纳德（Yvon Chouinard）说：“我们都知道自己必须减少消耗。”²⁵其他公司可能会用委婉语在自己与自己行为所造成的影响之间划清界限，或者试图模糊这些影响，但巴塔哥尼亚承认自己在这一因果链中负有全责，也没有为自己的行为找任何可能导致滑坡出现的例外或借口。他们对自己和公众都十分诚实，无论自己的行为给世界带来的影响是好是坏，都毫不遮掩。他们知道，自己若想在无限游戏中幸存下去并进一步发展壮大，就必须做到这般的诚实。他们没有把自己描绘成行业体系的受害者，他们认为自己就是这个体系的一部分，他们正在尽己所能地改变这个体系。

任何丑闻或道德堕落事件的事后调查几乎总会揭示出领导层的失败。如果拥有与迈兰、富国银行类似的文化，企业内的道德就几乎注定会堕落。如果这些企业的领导者听信了米尔顿·弗里德曼的话，认为自己的职责就是提升业绩，那么这些企业内的激励机制也会强化这一信念。最终，他们会认为短期财务目标比任何事业（如果他们有的话）都重要。领导者如果认为资源比意志更重要，那么他们就会调整企业文

化，令其符合这一优先顺序。

在巴塔哥尼亚，领导者是根据崇高的事业来确定优先顺序的，这一优先顺序会指导人们做出相应的行为——这是所有无限思维企业的共同点。这一做法关心的不仅仅是他们今年能赚到多少钱。该公司的人力资源与共享服务副总裁迪恩·卡特（Dean Carter）表示：“我们计划在未来的一百年中继续从事该行业，因此会考虑长期结果。”²⁶巴塔哥尼亚的领导者是在用无限思维运营公司，因此他们的目的就不是赢或击败市场上的哪个竞争对手。巴塔哥尼亚背后的驱动力源自他们对未来的一种愿景，该愿景希望，终有一天，他们可以在生产优质产品的同时将负面影响降至最低，以及“用商业来激发灵感，得出应对环境危机的解决方案，然后通过商业来实施这些解决方案”。²⁷

巴塔哥尼亚绝不是一个完美的公司，他们也会犯错，公司的员工也会有道德失范的时候。但他们敢于承认，也明白自己对崇高的事业的追求是一个不断自我改进的过程。在太多的企业中，“不断改进”这个词往往意味着改进程序和提高效率。在巴塔哥尼亚这样的无限思维企业中，意志与资源都很受重视，他们的不断改进覆盖了组织的方方面面，包括他们的文化，以及该文化所遵循的标准。这有助他们维持高道德标准的文化。巴塔哥尼亚的目标不是成为最好的，而是成为更好的。

就算有部分人对这个标题无动于衷，但这并不是什么一次性的花招。这符合巴塔哥尼亚一贯的行事作风，该公司一直在努力对自己的行为负责，也一直在改进。正如官网所说：

巴塔哥尼亚是一家不断成长的企业，我们希望能够长久经营下去。对我们是否真诚的考验是，我们能否做到销售的每一样产品都是有用的、尽可能多功能的、经久耐用的、美观但又不盲目追求时尚的。我们还没有完全通过这一考验。²⁸

这一声明接着还承认，该公司并非所有产品都符合这些标准，但马

上就介绍了他们的“化纤再生项目”（Common Threads Initiative），希望该项目能有助于他们向自己的目标前进。在该项目中，他们承诺会生产经久耐用的优质服装，这样就不必经常换新（从而减少浪费）；承诺免费维修产品，这样人们就不会把它们直接扔掉（从而减少浪费）；与易贝（eBay）建立伙伴关系，方便人们再利用、购买和销售二手产品（从而减少浪费）；当一个产品的寿命确实彻底终结时，巴塔哥尼亚会从中将它带走，回收再利用（从而减少浪费）。

有些公司为了提高业绩而想方设法地寻找漏洞，但巴塔哥尼亚不同，他们会想尽办法填补漏洞，从而强化自己的价值观与信念。例如，在过去的10年中，该公司一直在与非政府组织维泰（Verité）合作，在其一级供应链中、产品生产工厂中寻找和纠正劳动力滥用的问题。2011年，该公司通过独立的内部审计发现，尽管他们一直在努力打造对社会负责的供应链，但在其供应链中仍存在大量违规行为，包括在二级供应链，即负责将原材料加工成布料等生产材料的工厂中，发现了多起人口贩卖与剥削的事件。值得注意的是，巴塔哥尼亚就连二级供应商的情况也一直在关注，更不用说一直在设法改善了。

“公平劳工协会（Fair Labor Association，简称FLA）会对企业的海外工厂进行现场审计，帮助企业改进他们的企业责任项目。”吉利恩·怀特（Gillian White）发表在《大西洋月刊》（*The Atlantic*）上的一篇文章写道，“就连这样的组织都只要求相关品牌审计、监督和报告一级供应商的情况。”²⁹根除强迫劳动是一项艰难且复杂的工作，需要投入大量的时间和资源。大多数企业只会在出现丑闻或法律问题时，才会被迫去解决。巴塔哥尼亚则不同，他们是自愿承诺并投资的。他们也知道这个问题或许永远无法彻底解决，但仍会不断努力尝试。这就是持续改进和道德行动的全部意义所在。这恰恰符合了有效的、崇高的事业的标准，我们的理想或许是永远无法实现的，但我们会不断努力，至死方休。这让我们在公司的工作有了目的和意义，并会激励我们以端正的品行继续这场战斗。

有限思维的企业可能会担心这种做法成本过高，或会导致利润下降、顾客流失或声誉损毁（如今，几乎没有企业会主动承认自己做错了什么）。巴塔哥尼亚并不担心这些，他们并不害怕把自己暴露在公众眼前，也不害怕承担巨大风险。当然，他们也坦诚自己拥有一大优势，巴塔哥尼亚是一家私人企业。“上市公司确实承担着每个季度都要让利润有所增长的巨大压力，毕竟他们的投资者只对企业最终能给自己在财务上带来的既得利益感兴趣，”³⁰迪恩·卡特提醒我们，“因此，作为私有企业确实更有利，因为我们的既得利益无疑就是要产生更大的影响。”

尽管巴塔哥尼亚是一家获得了认证的共益企业，也就是实行“利益相关者资本主义”（stakeholder capitalism）的企业，但巴塔哥尼亚并不是一家慈善机构，依然是一个逐利组织，希望今年能比去年赚得更多。但他们也深知，赚钱并不是自己存在的理由。与其他所有的无限思维的企业一样，他们将金钱看作继续追求崇高的事业所需的燃料。要获得共益企业的认证，企业必须先明确自己最根深蒂固的社会与环境价值观，然后要遵守这些价值观，履行自己对员工、顾客、供应商和社会的责任，以及对投资者财务健康的责任。³¹巴塔哥尼亚知道，企业越成功，就越有能力坚持这个标准，能对世界产生的积极影响也就越大。他们知道，从长远来看，若能坚持以崇高的事业为引导，时刻提防道德的堕落，他们就能吸引并领导那些与他们有着相同愿景与价值观的人，这些终将成就他们的繁荣。

道德决策不应该以最有利于短期利益为出发点，而应该以“要做正确的事”为出发点。以牺牲道德为代价的短视性思维会慢慢削弱一家公司，“做正确的事”则会让这家公司越来越强大。巴塔哥尼亚一直在竭尽全力做正确的事，并把人类与地球置于利润之前，这种模式已经为他们赢得了员工和顾客的高度忠诚。有了这些，再加上他们已经在市场上拥有的良好意志与信誉，他们便成为同类公司中最成功、最具创新性和最赚钱的公司之一。在过去的10年中，该公司的收入增长了3倍，利润增长了两倍。用巴塔哥尼亚首席执行官罗斯·马卡里奥（Rose Marcario）

的话来说，“为地球做点好事能够创造出新的市场，能够让我们赚到更多的钱”。³²这里请注意她的话语中所体现的优先顺序。

“就我们当前已有的认知而言，”巴塔哥尼亚环境事务副总裁里克·里奇韦（Rick Ridgeway）说，“我们一直坚持生产的服装，都能让选择了这件产品的人过上更舒适的生活。只要还有大量的人没有过上这种生活……我们应该就会继续增长。”不过，里奇韦承认，“我们的增长也存在一个临界点，到那时，增长制造的问题可能会超过它所能解决的问题。”真到那一天，巴塔哥尼亚将会作何应对，我们也只能拭目以待。不过，他们一直都在考虑和谈论（公开谈论也不少）这个问题，而这一做法本身就是其具有道德优势的又一标志。

得益于无限思维的领导，巴塔哥尼亚不仅成为一家更能抵御道德堕落的公司，而且还亲身示范了何为合乎道德的做事方式，为商界树立了榜样。“我们认为，如果能让商界看到我们的成功，”首席运营官道格·弗里曼（Doug Freeman）表示，“我们就能成为很好的示范，让人们看到经商还可以有别的方式。”³³巴塔哥尼亚的行事方式既有益于他们自己，也有益于这个游戏……而且这种方式是管用的。如今，其他公司也在纷纷效仿他们。

THE INFINITE GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE

要素 3

尊重值得较量的
对手

第9章

值得较量的对手

光是听到他的名字，我就会不舒服。若听到有人赞美他，一股嫉妒之情就会涌上我的心头。我知道他是个好人，很友好。我非常尊重他的工作，他每次在专业场合见到我都对我十分友好。我们做着同样的工作——写书、演讲，与世界分享我们的观点。虽然与我们做着类似工作的人还有很多，但不知为何，我就是特别在意他，我想要超越他。我会定期查看网上排名，看看我的书卖得如何，与他比又如何。我对其他人都不会如此，只有他。如果我的排名比他高，我就会笑得洋洋自得，觉得自己高他一等；如果他的排名高一些，我就会面带怒容，心烦不已。他是我主要的竞争对手，我想赢过他。

但后来，发生了一件事。

我们受邀同台。尽管之前我们曾在同一场活动中演讲过，但这是我们第一次真正的同台。这一次，我们要同时出现在台上，肩并肩坐着接受联合采访。主持人认为让我们介绍彼此会很“有趣”。我要先介绍他。

我看了看他，又看了看观众，最后重新看向他，说道：“你让我极其没有安全感，因为你所有的长处都是我的短处。我做起来很艰难的事，你却可以做得很好。”观众笑了。他看向我，回答道：“不，安全感是相互的。”他接着说出了我的一些长处，同时也是他希望自己能有所提高的地方。

在那一刻，我突然明白了自己为什么那么想赢他。我对他的看法与他无关，而与我自己有关。每当他的名字出现时，总会让我想起自己的短处。相比将自己的精力投入于自我提升——克服弱点或增强优势，把

注意力放在打败他上要更容易一些。竞争就是这样，不是吗？它会驱使我们争胜。问题是，所有衡量谁领先、谁落后的标准都是主观的，是我自己设置了比较的标准。此外，这场竞争是没有终点的，我一直想赢的其实是一场赢不了的竞赛。我犯了一个典型的有限思维错误。事实上，尽管我们做着类似的工作，但他并不是我的竞争者，而是我值得较量的对手。

任何看过或玩过游戏和体育比赛的人都熟悉有限游戏的概念：一位玩家或一方击败另一方，赢得冠军或奖品。其实，这是一种根深蒂固的思维方式，无论面对什么性质的游戏，只要场上有其他玩家，我们会自然而然地拿出“我们”要与“他们”对抗的态度。但在无限游戏中，我们必须停止这种思维方式，不能再把其他玩家视为要打败的竞争者，而要把他们视为可以促使我们变得更好的值得较量的对手。

值得较量的对手是这场游戏中另一个值得比较的玩家，可能是我们行业内的玩家，也可能是行业外的参与者，可能是我们的死敌，也可能是我们一时的合作者或同事。只要我们是无限思维在参与游戏，那么他们用的是无限思维还是有限思维并不重要。不管他们是谁，不管在哪里遇到他们，重要的都是他们在某些方面（或许多方面）做得与我们一样好，或者比我们更好。与我们相比，他们可能是制造了更好的产品，拥有更高的忠诚度，是更好的领导者，或者行动时有更清晰的目标。我们不需要赞美他们的一切，不需要赞同他们，甚至不需要喜欢他们，而只需要承认他们身上有我们可以学到的长处和能力。

我们必须为自己选出值得较量的对手，明智的做法是要有策略地去挑选。仅仅为了优越感而挑选一直逊色于我们的玩家是毫无价值的，这对我们自身的成长几乎没有或毫无价值。他们不一定要是最大的玩家，也不一定要是当前的参与者。将他们选为值得较量的对手，仅仅因为他们身上有某些东西能够暴露出我们的弱点，促使我们不断改进。如果我们想要强大到可以不被这场游戏所淘汰，他们的存在对我们而言就是至

关重要的。

从20世纪70年代中期到80年代，克里斯·埃弗特·劳埃德（Chris Evert Lloyd）和玛蒂娜·纳芙拉蒂洛娃（Martina Navratilova）都是女子网球界的领军人物。¹尽管在场上她们是迫切想要战胜对方的竞争者，但真正帮助她们成为更优秀网球选手的，是对彼此的尊重。“我感谢她作为我的对手为我所做的一切，让我表现得更加出色，”劳埃德在提到纳芙拉蒂洛娃时深情言道，“我想她也会感激我为她所做的一切。”例如，正是纳芙拉蒂洛娃的存在，迫使劳埃德改变她的打球方式，她不能再过于依赖底线拿分了，必须学着成为一个更有进攻性的选手。这正是值得较量的对手能为我们做到的。他们会以一种几乎没有其他人能做到的方式推动我们前进，即便是我们的教练也做不到。在劳埃德和纳芙拉蒂洛娃的故事中，这一推动作用不仅提升了她们的表现，也提升了网球这项运动的水准。

这种微妙的思维转变会对我们的决策方式，以及我们对资源的优先排序产生深远影响。传统竞争会迫使我们拿出争胜的态度，值得较量的对手则会激励我们拿出改进自己的态度。前者让我们关注结果，后者让我们关注过程。这种简单的视角转变会即刻改变我们看待自己企业的方式。对过程与持续改进的关注才能帮助我们发现新的技能，增强我们的适应性。过度关注赢不仅会逐渐耗尽我们的精力，还会真真切切地扼杀创新。²

之所以说需要调整自己的视角，将自己领域中的强者视为值得较量的对手，还有一个原因，就是这种思维方式能让我们保持诚实。这就像一个一心只想赢的跑者，他会忘了规则，甚至忘了道德，忘了他开始跑步的初衷。他可能会将时间、精力花在影响比自己快的跑者上，比如绊倒对方。或者，他可能会服用能提高成绩的药物，偷偷地给自己增加优势。这两种策略无疑都会增加他赢得比赛的概率，但也会减小他在这些比赛之外获得成功的概率。而且，这些策略终会有不再管用的一天，到

那时，他仍将是那个跑得不如对手快的跑者。若我们能将其他玩家视为值得较量的对手，就不会再有必须不惜一切赢得比赛的压力，如此一来，我们自然也不那么需要做出不道德或非法的行为了。当坚持我们所践行的价值观比得分更重要时，我们自然会变得更诚实（那些在“做正确的事”与“做有助于自己领先的事”之间选择了前者的组织或政治家就是很好的例子）。

说回那位对我来说值得较量的对手。在我将亚当·格兰特（Adam Grant）视为竞争者时，我没有从中获得任何助益，这种做法只是满足了我的有限思维。它让我更关心自己在主观标准中的排名，而不是对自己事业的推进。我花费了太多的时间、精力担心他在做什么，而没有集中精力思考该如何提升自己的水平。

从那天起，我学会了转变自己的思维方式，我不再比较我与格兰特（或任何人）的图书排名，也不再将自己的不安全感化作与他的对抗，而是选择与他合作推进我们共同的事业。我们成了亲密的朋友（他还好心地帮我校对了这本书，帮我将它变得更好）。如今，再听到他的名字，或再看见他的出色表现时，我会由衷地感到开心。我希望他的思想得到传播。其实，每位阅读本书的读者，也都应该去读一读《沃顿商学院最受欢迎的思维课》（*Give and Take*）和《离经叛道》（*Originals*）。一个有趣的事实是，在无限游戏中，我们双方都可能成功，事实证明，人们其实可以不止买一本书。无限思维拥抱的是充裕，有限思维的基石却是稀缺。在无限游戏中，为“成为最好的”所做的努力都是徒劳的，因为在同一时间可以存在多位表现出色的玩家。

值得较量的对手有助于提升业务水平

2006年，艾伦·穆拉利离开了飞机制造商波音，临危受命，接任了福特公司的首席执行官一职。这是一段新旅程的开始，这段旅程最终成就了汽车史上最伟大的一场翻身仗。在正式宣布就职的新闻发布会上，

穆拉利回答了一些问题。一位记者问他开的是什么车。“一辆雷克萨斯，”穆拉利说，“这是世界上最好的车。”福特的这位新掌门刚刚承认了他开的这辆丰田车比任何福特车都要好！在一些人看来，这是亵渎。但在穆拉利看来，这是一个诚实的评价，穆拉利是个偏好真相的人，即便这个真相会令人感到不舒服。

在穆拉利接手前的15年里，福特损失了25%的市场份额。³彼时，福特正走向破产。穆拉利确实需要能带来转机的策略，但他首先想做的是，尽他所能地了解这家公司。他想了解福特除资产负债表之外的健康状况。而他了解到的事情之一就是，消费者为什么对这个品牌不再抱有幻想。⁴福特汽车（至少在美国）的名声是乏味、不可靠且是个“油老虎”，这或许就是人们不再像过去那样购买福特汽车的原因吧。

从历史上看，包括福特在内的底特律汽车企业都有一个通病，就是太过执着于市场份额，并将其视为竞争与比较的主要指标。⁵但穆拉利知道，世界上有一些最赚钱的汽车公司恰恰也是市场份额最小的公司之一。他很快发现，只是提升市场份额并无益于福特的长远发展，毕竟市场份额这个东西，靠促销活动和成本削减就可以做到（这正是穆拉利刚来福特时，福特提交给他，希望以此来扭转局面的方案），但该战略的效果只能维持短短几年。“我们不会追求市场份额，”他说，“不会在没有需求的地方投放汽车，然后通过打折来拉动销量，这样只会让问题更糟。”⁶如果福特不想被这个游戏所淘汰，就必须改变自己的游戏方式。这意味着，他们必须重新学习如何制造出人们真正想要的汽车。

穆拉利加入福特后做的第一件事情就是，每晚驾驶一辆不同型号的福特汽车回家。在试遍公司的每一款汽车后，他要求换开一辆丰田凯美瑞回家。但问题是，福特连一辆能让他开的凯美瑞都没有。在福特公司，购买其他制造商的汽车是一种常规操作，但工程师会将那些汽车完全拆解，用以研究它们的制造方式，至于真正可供驾驶的——一辆都没有。现在想一想，一家苦于销售汽车的大型车企，高管们竟然几乎完全

不知道别人家的车开起来是什么样的感觉。如果购车者是通过试驾来做选择的，那么福特的高管们难道不应该知道购车者都在试些什么车吗？穆拉利让福特大批采购了其他制造商的汽车，并要求高管试驾。

穆拉利在说雷克萨斯是世界上最好的汽车时，并不是想要让福特人难过，而只是给他们提供了一个值得较量的对手。他相信，要救福特，就必须坦率地评价自己的产品，同时尊重业内的其他玩家，向他们学习。正如穆拉利所说，丰田这家企业制造的是人们想要的产品，但消耗的资源与时间却比世界上的任何车企都要少。丰田为福特提供了一个标杆，福特可以以此来推动自己汽车质量的提高、制造工艺的改进。他们若是能够排除万难，做到这点，利润就会随之而来。在穆拉利看来，研究其他汽车制造商并不仅仅是为了模仿他们或在销量上超过他们，而是为了向他们学习。“我从未试图击败通用汽车或克莱斯勒，”穆拉利说，“我们始终专注于自己的事业，并用标杆分析法替代竞争，为我们找出在运营中可以持续改进的地方。”持续改进生产程序可以制造更好的产品，也能更有效地推进亨利·福特最初确定的崇高的事业：为每个人都提供安全、高效的交通工具，让公路为全人类敞开。亨利·福特确定的这一事业也起到了过滤器的作用，帮穆拉利筛选确定了其他的一些决策，比如低价甩卖捷豹、路虎、沃尔沃等品牌。福特公司当初购买这些品牌是希望自己能在尽可能多的不同类汽车市场中拥有竞争力，但穆拉利认为这分散了福特公司对自己创立初衷的关注。

2008年的股市大崩盘给美国汽车工业造成了毁灭性的打击。如果没有政府的紧急财政援助，通用汽车和克莱斯勒都会破产。但福特自行经受住了这次经济衰退的考验，而这多亏了穆拉利在2006年为帮助福特自我改造而申请的近240亿美元贷款，还有福特在运营和产品方面的稳步改进。紧急财政援助计划实施前，国会举行了听证会，穆拉利也在出席证人之列，鉴于福特的情况，他完全可以坚持要求政府不向通用汽车或克莱斯勒提供资金，但他没有。他若将其他玩家视为竞争者，当然会乐意看着他们破产，让福特成为美国唯一幸存的大型汽车制造商。这当然

是赢了吧？

不过，穆拉利认为其他制造商都是值得较量的对手，因此，他非常支持这项援助计划。他知道这些公司的存在只会成为福特的助力，让福特成为更好的公司。他也知道，福特的对手还是另一个更大生态系统的组成部分。如果他们破产了，许多供应商也会破产。这也可能会摧毁福特。因此，穆拉利还整合了各项方案，对许多汽车供应商伸出了援手，想帮他们渡过这段经济衰退期。遗憾的是，虽然深陷困境，通用汽车与克莱斯勒的领导者们仍然在用有限思维经营公司，他们拒绝了福特的提议，不愿意为了行业的共同利益而合作。本田、丰田、日产则不同，他们与福特合作，一同帮助大型供应商渡过了难关，而他们本身的业务其实也离不开这些供应商。无限思维的玩家明白，对自己生存最有利的选择就是让这场游戏得以继续，而这其实也是无限思维的领导者的最终目标。

值得较量的对手有助于看清从事一份工作的原因

到20世纪80年代初，计算机革命已经发展得如火如荼。苹果公司是这场革命中的领军者之一，对苹果来说，对手的真正价值几乎无关于激励他们改进产品，而是发挥了更大的作用。值得较量的对手让他们对自己的事业有了更清楚的认识，也让他们的员工更加团结。对手的存在提醒了苹果公司内外的每一个人，让他们始终记得自己所代表的是什么，也就是他们最初进入该行业的原因。“他们是海军，我们是海盗。”

20世纪70年代，IBM占据了大型计算机市场的最大份额。大型计算机是一种会填满整个房间的巨型机器，能够为企业提供庞大的运算能力。但IBM拒绝研发我们过去所说的“微型计算机”，他们认为这些计算机运算能力不足，无法满足企业的需求。他们认为，个人计算机在办公室里是不会有立足之地的。

但到1981年，一切都变了。作为个人计算机的拓荒者，康懋达（Commodore）、坦迪（Tandy）和苹果都在企业客户中取得了非常高的销量，看到这些，IBM态度大变。由于现金充裕，IBM能够投入巨额资金研发自己的个人计算机。他们开出了高得离谱的薪水，到其他同行业公司挖走了一些最优秀、最有前途的工程师，苹果也在被他们挖角的企业之列。短短12个月，他们自己的“个人计算机”就问世了。

在IBM加入之前，个人计算机市场上份额最大的是苹果。这意味着，当IBM进入该市场后，苹果的损失最大。听到这一消息，有限思维的玩家很可能会感到恐慌，但像苹果这样无限思维的玩家却恰恰相反。1981年8月，就是IBM的个人计算机首次上市的那个月，苹果在《华尔街日报》上刊登了一整版的广告，标题是“欢迎IBM，真诚地”广告内容将苹果对这一新玩家的看法表达得一清二楚，苹果将其视为值得较量的对手，而非竞争者。⁷

“这是自35年前计算机革命开始以来最激动人心、最重要的市场，欢迎你的加入。”苹果这则广告的开头写道，“将真正的运算力交到个人手中，人们的工作、思考、学习、交流与休闲方式已经在随之改进。在下一个10年中，个人计算机的数量将继续呈指数级、跳跃式的增长。只有通过巨大的努力才能将这项技术推广到全世界，我们期待负责任的对对手加入。我们理解你们的付出。因为我们在做的事情是，通过提高个人生产力来增加社会资本。”苹果在这份致新对手的信上这样写道：“欢迎共担这一使命。”苹果正在努力推进一项崇高的事业，IBM的加入也将帮到它。⁸

IBM接受了这一挑战，而且，得益于在商业领域的主导地位，能够利用早已与大企业建立的合作关系，向对方销售新的个人计算机。当时，IBM被亲切地称为“蓝色巨人”，对于所有负责为公司购买个人计算机的采购经理来说，蓝色巨人都是安全且优先的选择。当时流传着一句话：“从来没有人因为采购IBM的产品而被解雇。”IBM为了进一步拓展

业务，允许其他计算机制造商在自己的产品中“克隆”或使用他们的操作系统。苹果拒绝效仿这一做法。任何想用苹果操作系统的人，都必须购买苹果的产品。由于无法克隆苹果的操作系统，为大众市场自行研发新操作系统的成本又非常高昂，大多数其他计算机制造商都选择向IBM购买操作系统的授权，生产与其操作系统兼容的产品。这一做法也令IBM的个人计算机成为该行业的标准，适用于商界内外。

IBM“帮助”苹果公司把个人计算机变成了每张办公桌上的必需品，以及每个家庭的基本家用电器。但他们为苹果做的还远不止这些。苹果还利用与IBM的对比，更清晰、更有说服力地说出了他们的故事，以及他们代表的是什么。崇高的事业存在于我们的想象之中，但公司和产品都是真实存在的。对任何有明确事业感的个人或企业来说，他们本身就可以成为自己无形愿景的有形象征。对我们来说，追随一个真实存在的企业或领导者，会比追随一个抽象的想法更容易。我们若能在讲述自身崇高的事业的时候，举出另一种选择的有形代表，便更有可能令人信服。

约翰·库奇在苹果创立之初就加入了这家公司。“他们是海军，可以预见，他们的产品是销往企业的，”他是这样描述IBM的，“我们想成为海盗，让每个人都拥有创造力。”IBM和苹果代表了两种意识形态，在寻找各自的追随者。IBM代表着企业、稳定和一致，苹果代表着个性、创造力和差异化的思维方式。通过向公众展示对比，苹果从领导个人计算机的革命，过渡到了领导一群志趣相投的革命者。

我们衡量计算机质量的标准指标有价格、运行速度、内存等，基于这些指标，两家公司生产的个人计算机基本不相上下。事实上，克隆IBM的产品往往还会便宜很多。竞争者所比较的几乎总是产品的特性和优势，苹果却选择在更高层次上与IBM交锋。竞争者争夺客户，对手寻找追随者。在苹果的追随者眼中，IBM代表过去，而苹果代表未来。而在IBM的信徒眼中，苹果不过是给挥洒创意者提供的玩具，IBM才是在

为认真做着重要工作的人服务。这个层面的比较比产品和特性的比较更宏观。这成了一场信仰的游戏。

苹果对IBM进入个人计算机市场的反应与常理完全相反。一个行业如果迎来一个如此强大的新玩家，老玩家们往往会十分紧张。他们往往会忽视自己的愿景，开始专注于与新玩家之间在产品等标准指标上的竞争。这意味着，即便他们之前用的不是有限思维，一旦选择了将新来者视为竞争者而非值得较量的对手，过不了多久，他们就会被这一选择拖入有限思维的泥潭。这就是发生在加拿大手机制造商黑莓（BlackBerry）身上的故事。

曾经，IBM狂风暴雨般地杀入了苹果的市场，大约25年后，苹果也对黑莓做了类似的事。不同的是，苹果选择将IBM视为一个值得较量的对手，认为对方可以帮助自己更好地阐明自己代表着什么，黑莓却选择将苹果视为一个需要打败的竞争者。最终，黑莓为这个有限思维的选择付出了沉重的代价。

在苹果手机出现前，黑莓拥有全球第二大手机操作系统。得益于高性能、高耐用且非常可靠的产品，黑莓手机成为政府和许多公司的必备品。他们占领了企业市场。即使在苹果2007年推出苹果手机之后，黑莓的发展势头仍然强劲，并在2009年创下手机销量的新高，占手机总销量的20%。然而，随着苹果手机越来越受欢迎，黑莓陷入了恐慌。黑莓的领导者本可以如几十年前苹果对IBM所做的那样，选择让双方的理念形成鲜明对比，凸显自己对世界的愿景，即以满足企业和政府对安全性、可靠性的需求为核心的愿景，但是他们没有。面对苹果手机越来越火的事实，黑莓选择努力模仿它。首先，他们开始为现有设备提供应用程序和游戏，这极大降低了自身产品的反应速度。接着，他们放弃了自己标志性的全键盘设计，转而采用触摸屏。然而这些产品没能像苹果手机一样好用，在耐用性上也比以往那些型号的产品差了许多。

遗憾的是，这种情节经常上演。记住，妨碍对手往往是有限思维的

一种征兆。有限思维的领导者往往会错失良机，没能利用行业中的颠覆性事件来更清晰地凸显自己所追求的事业。他们会加倍努力地投入到有限游戏之中，简单粗暴地模仿其他玩家，希望对其他玩家管用的手段，对他们也能管用。然而，事实证明，对苹果有用的，对黑莓并不管用。他们放弃了领导一项事业的机会，选择成为一种产品的追随者。他们一心只想打败苹果，眼中再也没有自己的愿景。他们忘了当初进入该行业的初衷。很快，黑莓的业绩就开始持续地急剧下滑。到2013年，该公司的市场份额已不足1%，短短4年里下跌了近99%。如今，黑莓在这场它曾经主导过的游戏中，已经沦为了一个无足轻重的玩家，没有哪家公司会再将它视为值得较量的对手。

多年来，IBM一直被苹果视为值得较量的对手。最终，随着计算机的普及与市场的改变，IBM退出了个人计算机的游戏。不过，对手的失利并不等于苹果的胜利。苹果很快又找到了另一个代表着安全、稳定和企业思维的对手——微软。不过，与IBM一样，微软也慢慢模糊了自己的事业，不再能像过去那样，与苹果在意识形态上形成鲜明的对比。对现在的苹果来说，新的值得较量的对手又是谁呢？

事业盲症

我有一个朋友，她极其专注于自己的事业，仿佛忘记了世界上除了她自己的观点还存在其他观点。遗憾的是，这位朋友给所有与她持不同观点的人都贴上了错误、愚蠢或道德败坏的标签。我的这位朋友患了事业盲症。

事业盲症是指，我们过度投入自己的事业，或是过度专注于寻找其他玩家事业中的“错误”，因而无法看到他们的长处或自己的短处。仅仅因为不认同对方的观点、不喜欢对方或发现了对方道德上的重大劣迹，就错误地认为他们不配被拿来与我们比较。我们看不到在同一件事情上，他们确实有比我们更有效或更好的地方，而这些地方恰恰是我们可

以学习之处。

事业盲症会削弱谦逊、助长傲慢，从而阻碍创新，降低我们参与这场长期游戏所需的变通性。如果无法做到诚实，或持续改进，就可能会不断重蹈覆辙，或一直将许多事情做得一团糟。另外，组织的任何弱点都有可能沦为其他玩家手中的把柄，而傲慢会增加发生这种情况的概率。要继续游戏，我们就必须拥有意志和资源，但上述这些会导致意志和资源大量流失。每当我试图告诉我的朋友，那些他们认为卑鄙的玩家也有某些长处，她应该尊重他们时，她就会嘲笑我，还会因为我居然夸奖她的竞争者而将我视为背叛者。

我们可能真的很难将某个玩家认作自己值得较量的对手，尤其当对方有令我们讨厌的地方时，但找到值得较量的对手才是帮助我们成为更好玩家的最佳方式。约翰·道格拉斯（John Douglas）是罪犯侧写的先驱，退休前是美国联邦调查局（FBI）行为分析小组的主管，他曾解释道：“我对这些人审问得越多，就越觉得成功的罪犯都是出色的侧写师。”⁹比如，道格拉斯发现，尽管我们都认为连环杀手不可理喻，但抓住连环杀手的最佳方法就是承认FBI擅长之事，他们也非常擅长……这意味着FBI必须做得更好。有了值得较量的对手，即善于躲避FBI调查的罪犯，FBI就不得不持续不断地提升自己的技能。

认可对方是自己值得较量的对手并不意味着认可对方的事业是道德的或合乎伦理的，而只意味着对方有某些长处，能让我们看到自己该改进的地方。他们的游戏方式会对我们构成挑战或压力，或者会激励到我们，无论哪种，最终都会促使我们自我改进。要选谁成为值得我们较量的对手完全取决于我们自己。而且在无限游戏中，最有利的做法恰恰是不固守自己的选择。

不要误以为没有了值得较量的对手就等于赢了这场游戏

柏林墙倒塌后不久，美国犯了一个愚蠢的错误，这可能是美国在20世纪最错误的外交政策之一。美国宣告自己“赢”了冷战，但其实并没有。本书看到现在，我们都知道这一准则：在无限游戏中，根本不存在“赢”这件事。这是商界的真理，也是全球政治的真理。美国并没有打赢冷战，是苏联自己在资源耗尽后主动退出了游戏。

冷战符合无限游戏的所有标准。有限战争会有约定好的游戏规则、容易区分的阵营，以及对战争何时结束的明确定义（比如夺下了土地，或其他一些容易衡量的有限目标）。冷战与有限战争形成了鲜明的对比——没有基本规则，也没有明确的战争目标，各方无从得知战争何时结束。尽管美国等西方国家声称自己“赢”了冷战，但除了没有爆发双方都不希望看到的全面核战争，几乎没有人能准确想到或预测到冷战的胜利会是什么样子的。没有标志冷战结束的条约，双方只是在持续对抗，一直想方设法地改进自己的游戏方式，但又完全不知道这场战争会朝着什么方向发展。因此，1989年柏林墙的倒塌其实并不在人们的预料之中。

在商界，时代会变，玩家也会变，一家大公司的破产，并不意味着游戏的结束，也不意味着任何一家公司的胜利。玩家都知道，在这个行业中，还会有其他公司崛起，会有新的公司加入。在所有玩家中，最能驱使我们前进的，就是我们最重要的值得较量的对手，他退出游戏，并不意味着替补席上就一定有蓄势待发的新玩家，能够立刻出场。无论是新玩家还是老玩家，要想拥有足够取代上一个对手的实力，可能都还需要数年的发展。无限游戏的高级玩家都明白这一点，因此，在失去自己首要的对手后，他们仍会努力保持谦逊。他们知道新玩家的出现只是时间问题，因此会谨慎提防着傲慢与有限思维的滋生，防止自己被它们主宰。在无限游戏中，耐心是优势。

在没有发现值得较量的对手时，强大的玩家会误以为自己可以掌控游戏的方向，或者可以掌控其他玩家。但这些都是不可能的。无限游戏就像股票市场，公司可以上市，可以退市，但没人能掌控这个市场。

如果是资金雄厚、实力强劲且极其成功的玩家，他们就算忽视了自身的弱点，短期内可能也不会遇到什么麻烦，但长期来看并不是这样。例如，在产品、营销和资产负债表方面都表现出色的快速增长型企业，往往会忽视对领导层的培训或是对企业文化的积极塑造。这些做法总有一天会让他们自食苦果。高朋（Groupon）就是一个例子，商业媒体对该企业产品创新和增长率的称赞，令领导者忽视了自己的员工。一旦增速放缓且其他公司推出了与他们不相上下的产品，这种对员工的忽视就会变成致命弱点。优步是另一个例子。优步或许是共享出行技术的先驱，但对该公司来说，忽视文化比任何产品的失败都更加致命。2017年，达拉·科斯洛萨西（Dara Khosrowshahi）接替特拉维斯·卡兰尼克（Travis Kalanick）成为优步的新首席执行官，此举目的明确，就是要修复该公司的文化。

拥有值得较量的对手可为我们提供巨大的优势，它让我们有了单一的关注重点，能更有针对性地制定策略、分配资源，能让内部的不同部门一致对外。

在商界，新玩家的出现势必会改变游戏规则。作为电影租赁行业唯一的超级巨头，百视达在像奈飞（Netflix）这样的小公司和像互联网这样的新兴技术出现时，没能意识到这其实是在要求他们重新审视自己的整个商业模式。当亚马逊出现时，大型出版商没有自问，面对这个新的数字时代，应该如何更新、升级自己的模式，反而是翻倍地投入于旧有的模式。出租车公司也没有自问“我们需要做些什么来与时俱进”，没有学习如何适应新的现状，如何提供更好的出租车服务，而是选择通过起诉共享出行公司来保护自己的商业模式。邮寄纸质产品目录的做法令西尔斯百货在过去的数十年里逐渐成长为一家超大型的公司，积攒了大量的财富，习惯于此的他们没能快速适应沃尔玛等大卖场与电子商务的崛起。作为行业巨头的Myspace认为自己无可匹敌，甚至没有察觉到Facebook的到来。曾经送我们到达今天这个位置的方法，已经无法将我们送得更远，找到值得较量的对手是帮助我们及时自我改进并适应新环

境的最佳方法，以免当我们醒悟时一切都来不及了。

少了值得较量的对手，我们有可能会失去谦逊与敏捷的反应。若是没有值得较量的对手，就算是实力强劲且拥有强烈事业感的无限思维玩家，也很可能会慢慢变成一个追逐胜利的有限思维的玩家。若没有值得较量的对手，曾经将他人利益与事业利益放在首位的组织就很可能变得以追逐个人利益为主。当这种情况发生时，傲慢逐渐滋生，组织的弱点就会很快暴露于人前，玩家会变得过度僵化，失去留在该游戏中所必需的变通性。

THE INFINITE GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE

要素 4

拥有攸关存亡的
变通性

第10章

攸关存亡的变通性

有些人认为他疯了。他开始将资产折换成现金，变卖房产，并以自己的人寿保险为抵押去贷款，甚至还将自己的名字授权给他人。公司经营得这么好，他为什么要在此刻离开，去做那么不同、那么冒险的事情呢？1952年，华特·迪士尼这样做了，而他所做的是关乎攸关存亡的变通。

华特·迪士尼习惯了冒险和创新。作为动画这一新兴领域的年轻艺术家，他在不断地创新。他是最早让真人演员与卡通人物互动的短片制作人之一。1928年，他制作了首部声画同步的动画片——《汽船威利》（*Steamboat Willie*）。不过，华特·迪士尼并不满足于仅仅制作让观众发笑的娱乐短片，他开始制作可以以假乱真的动画电影，一部可以激发人类全部情感的电影。1937年，他推出了有史以来第一部长篇动画电影——《白雪公主和七个小矮人》，这对全世界来说都是过去从未见过的新事物。其作品的这种演变并不是为了实验而实验的结果，也不是为了发财或出名。他所走的每一步，都是在推进他自己的崇高的事业：邀请观众放下生活中所有的压力和焦虑，进入他所创作的那个更加恬静美丽的世界。

华特·迪士尼的崇高的事业源自他还是孩子时播下的种子。在他4岁时，他的父亲埃利亚斯（Elias）带着全家离开芝加哥，搬到了密苏里州马瑟琳乡下的一个农场。小华特总在户外玩耍，常常看到四处漫游的动物，在那里，他有着来自大家庭的关爱，还有一个所有人都相互支持的社区。他的哥哥罗伊（Roy）后来回忆道，那里“就是城里孩子的天堂”。¹但这种恬静美好的童年并没有持续太久。搬到马瑟琳仅仅5年

后，埃利亚斯想做农夫的努力就失败了，全家人不得不再次搬家。

在堪萨斯城定居后，埃利亚斯买下了一条送报路线，小华特被叫去送报，好帮忙家里维持生计，但家里还是入不敷出。随着财务状况越发恶化，埃利亚斯身上的压力越来越大，越来越大的还有他的脾气。“后来发展到，”华特·迪士尼回忆道，“跟父亲说实话也会招来一顿揍。”幸运的是，在农场生活时，华特发现了画画的乐趣，这一爱好成了他的避风港，让他得以从现实生活的艰辛中完全逃离出来。在后来的人生中，他利用自己的绘画能力与想象力，给其他人也提供了一个逃离现实的机会，将他们带到了另一个世界，让他们也能体会到，他童年记忆中曾在马瑟琳那个小镇拥有的快乐。

超越无限

事实证明，华特·迪士尼能够将人们带到另一个世界的能力是非常赚钱的。《白雪公主和七个小矮人》的上映，不仅赢得了评论界与大众的赞誉，而且仅上映第一年就赚到了超过800万美元的票房。凭借这部电影带来的金钱和成功，他在伯班克市建立了自己的工作室，该公司前员工唐·拉斯克（Don Lusk）认为那里的企业文化可以用“就是天堂”来形容。为了偿还建造这间工作室所累积的债务，并推动公司的发展，当时作为公司首席执行官的罗伊·迪士尼想让客户上市，但华特·迪士尼反对这一想法，他担心股东会干涉公司的业务。不过，他最终还是屈服于压力，让客户上市了。

随着公司的发展，出现了大量新的挑战。首先，华特·迪士尼制作公司的文化开始分化出更多不同的阶层。²例如，过去人人都能享有的津贴，现在只提供给更资深的员工。随着工资差距的不断加大，内部的异议也越来越大，华特·迪士尼第一次遭遇了来自工会充满敌意的抗争。由于工作室内乌托邦文化的土崩瓦解，再加上逼他拍摄更节约成本的真人电影的压力，以及公司制度对他创造力的限制，华特·迪士尼对

未来感到了前所未有的挫败和沮丧。公司的思维变得越来越有限，越来越缺乏远见，这令华特·迪士尼确信，这家企业再也无法帮助推进他的事业了。尽管遭遇了挫折，但他无限的视野一如既往。这就是华特·迪士尼决定做出攸关存亡的变通的原因。他辞职了。在《白雪公主和七个小矮人》首次上映的15年后，华特·迪士尼选择离开，去做新的尝试。

1952年，他卖掉了房产、其他资产以及在华特·迪士尼制作公司的股份，还抵押了人寿保险来贷款，然后拿着所有的钱成立了一家新公司，并用自己名字的首字母为它命名——WED⁽⁵⁾。新公司成立后，他便开始了新的项目，他相信这个项目会比以往的任何项目都更有助于推进他的事业——建造一个真实存在的世界，让人们得以暂时逃离日常的现实生活。他要建造世界上最欢乐的地方，他要建造迪士尼乐园。

当时的游乐园往往很脏、很危险，多是一些游乐设施的随机组合，华特·迪士尼想要建造一个安全、干净的乐园，有连贯的故事线贯穿其中。在这里，你看不到丝毫的辛苦或麻烦，也看不到任何的黑暗或肮脏潜伏在阴影中。在这里，人们可以完全沉浸在完美的幻想世界中。他说：“我认为，我对迪士尼乐园的最大期望，就是它能成为的一个欢乐的地方，让成年人与孩子可以一同体验人生和冒险的奇妙，并从中得到快乐。在这里，你会离开今天，进入昨天和明天的世界。”

过去，观众只能看电影，而现在，在迪士尼乐园，人们可以走入电影。而且，不同于有限的电影，这座乐园是可以永远进化下去的。华特·迪士尼以真正的无限思维解释道：“迪士尼乐园永远不会完工。在这里，我们可以不断发展，不断加入新的东西。电影则不同，一旦拍摄结束，完成后续处理，我们的工作就完成了。即便有哪里可以改进，也不能再修改了。我一直想做一些有生命的、会不断成长的事情。这在迪士尼乐园实现了。”

和许多企业家一样，华特·迪士尼每次创业都会赌上自己的一切。不过，着手建造迪士尼乐园可能是他最大的一次冒险，因为他本可以不

这样做的。与他的第一次创业相比，这一次他可能失去的东西多了太多。这就是这位无限思维、富有远见的领导者所处的困境。不过，只要他发现公司所走的道路不再能推进他的事业，他便愿意赌上一切，重新开始。他不是看到了更大的商机才离开，也不是丢下了一个濒临失败的企业；而是发现了可以更好地推进自己事业的方法，然后迫不及待地抓住了它。

变通性与愿景

攸关存亡的变通性指的是一种能力，为了更有效地推进崇高的事业而对商业模式或战略过程进行极端颠覆的能力。这是无限思维的玩家对不可预测性的理解，这份理解让他们拥有了做出此类改变的能力。面对新事物或颠覆性的事物，有限思维的玩家会感到恐惧，无限思维的玩家则会沉迷其中。对有明确事业感的无限思维领导者来说，如果在展望未来时，发现自己现在所走的道路会极大限制自己推进崇高的事业的能力，他们就会变通。或者，当他们发现，有一种新技术比目前使用的技术更能帮忙推进他们的事业时，也会变通。若不具备这种无限的远见，就算改变了战略，哪怕是极端的改变，往往也是被动的或机会主义的。攸关存亡的变通性总是代表主动出击。不要把它与许多公司为了在新技术或新的消费者习惯下存活而采取的防御性手段混为一谈。例如，有许多报纸和杂志彻底改变了自己的商业模式，走向了数字化，但这并不是因为他们找到了推进自身事业的更好方式，只是他们为了在这个不断变化的世界生存下去而不得不做出的改变。这种改变确实是生存的必需，但鲜少能激励到组织内的人员，或重新点燃他们的激情。不过，攸关存亡的变通可以。

更能推动许多初创企业前进的往往不是他们所拥有的资源，而是企业家对愿景的激情。对于一个已经在享受成功的企业来说，攸关存亡的变通能够重新激发他们对新事物的激情。当华特·迪士尼创立WED，从

头再来时，他从原公司带走了一批人，这些人想要追随他踏上新的冒险，一如第一次冒险时那样。他们愿意分担风险，愿意投入时间，愿意为成功实现这一新构想付出一切。他们发现华特·迪士尼的热情是有感染力的，他们非常兴奋能再次参与一场自己做梦都没想过的冒险。这一变通也重燃了华特·迪士尼自己的激情。提到自己的新公司时，他说：“我太爱这里了！”

攸关存亡的变通不是发生在公司创立之时，而是发生在公司完全建成且已在正常运作之时。在所有有限思维的旁观者看来，这种变通确实是攸关存亡的，因为领导者是在用当前盈利之路的确定性交换一条新路的不确定性，风险重重，这可能会导致企业的衰落，甚至覆灭。在有限思维的玩家看来，这一步的风险太大，不值得一试。但在无限思维玩家眼中，继续现在的道路才是更大的风险。他们拥抱不确定性。他们认为，无法变通将大大限制他们推进自己事业的能力，他们担心坚持现在的道路可能终会将组织带向覆灭。

同样，无限思维的玩家变通的动机就是想要推进自己的事业，哪怕这样做会颠覆现有的商业模式。显然，有限思维的玩家之所以不愿变通就是为了保护当前的商业模式，哪怕这样做会阻碍自己事业的推进。如果说公司就是领导者用来推进事业的工具，那么为了让公司在无限游戏中尽可能长久地存续下去，重大的战略转变是必不可少的。

攸关存亡的变通性比正常经营企业所需的日常变通性更大。我们也不能将发光物综合征（shiny-object syndrome）与攸关存亡的变通性混为一谈。世界各地还存在着一类沮丧的员工，他们的领导者充满善意，有时也富有远见，但他们就像看到发光物就想抓的猫一样，遇到什么好点子都想抓住，都会说：“就是这个！要推进愿景我们就必须这样做！”当攸关存亡的变通发生时，所有相信该事业的人都很清楚为什么必须这样做。尽管他们可能不喜欢这种变化可能带来的剧变和短期压力，但他们都认同这样做是值得的，也都希望这样做。相比之下，发光

物综合征带给人们的往往是困惑和疲惫，而非激励。

当富有远见的领导者做出攸关存亡的变通时，在外界看来，他们仿佛拥有了预知未来的能力，而其实并没有。但他们确实对一个尚不存在的未来状态有着清晰且不变的构想，这就是他们的崇高的事业，他们会持续仔细地观察，寻找能有助于朝这个构想迈进的想法、机会或技术。艾伦·穆拉利在福特召开每周商业计划审查会议，也是为了研究其他公司正在做些什么，而这并不局限于福特的传统竞争对手。他解释说：“该制度的目的在于，时刻关注当下的一切发展并从中学习。”有限思维的领导者也会寻找机会，但他们的目光往往聚焦在自己的行业内，聚焦在公司的资产负债表上，或聚焦在近在眼前的未来。追求崇高的事业的无限思维领导者会把目光投向行业之外，看向更远处的未来——一个必须有想象力才能看到的地方。20世纪80年代初，当史蒂夫·乔布斯为苹果做出攸关存亡的变通时，一定就看到了那样的远方。

苹果具有非常清晰的事业感，该事业的种子早在苹果创立的很久之前就种下了。公司创始人史蒂夫·乔布斯和史蒂夫·沃兹尼亚克（Steve Wozniak）对权威集团充满了不信任。他们所热衷的想法是，赋予个人对抗权威的力量。在20世纪70年代的计算机革命期间，这两位年轻的创业者将个人计算机视为个人用来挑战现状的完美工具。他们想象着有一天，个人计算机能让个人拥有反抗企业，甚至与之竞争的力量。

在推出了苹果电脑一代和苹果电脑二代后，苹果已经成为一家非常成功的公司。1979年12月，正是苹果忙于研发新产品迭代的时期，乔布斯带着公司的几位高管参观了位于加利福尼亚州帕洛阿尔托（Palo Alto）的施乐创新中心——施乐帕洛阿尔托研究中心（Xerox PARC）。在参观过程中，苹果高管看到了施乐公司研发的新技术之一“图形用户界面”（graphical user interface）。有了图形用户界面，人们不用学习磁盘操作系统（DOS）命令等计算机语言也能使用计算机。有了图形用户界面，用户有史以来第一次可以用“鼠标”移动屏幕上的“光标”，“点

击”位于“桌面”上的可视“图标”和“文件夹”。若愿景是赋权给个人，那么这一创新就能让计算机走入更多人的生活。

苹果高管离开施乐帕洛阿尔托研究中心后，乔布斯分享了他的想法，苹果必须改变路线，必须投资于图形用户界面。其中一名高管试图发出理性的声音，大声说道：“我们不能这么做。”他提醒乔布斯，苹果已经在一个截然不同的方向上投入了数百万美元的资金和无以计数的工时。放弃这项工作，完全从零开始开发新产品，会给公司带来巨大的压力。这位高管继续道：“史蒂夫，如果投资于此，就等于亲手炸毁自己的公司。”而乔布斯的回答是：“我们自己炸，总好过让别人来炸。”

有限思维的领导者很难就这么放弃既定的战略路线，尤其在已经投入了大量的时间或资金后，他们也很难放弃似乎就快要到手的绩效奖金。尽管会给公司增加成本和压力，但在乔布斯看来，这一攸关存亡的变通是苹果唯一的选择。指导他做出这一选择的是崇高的事业，而非选择的成本。苹果的员工也赞同他的选择。那些热爱在苹果工作的人也热爱让乔布斯推着他们去做那些其他任何人都从未做过的事。就这样，在短短4年内，麦金塔计算机（Macintosh）问世了。它的计算机操作系统彻底颠覆了个人计算机领域。这是第一次，个人计算机的操作真正简单到了几乎任何人都能够使用。微软被迫追随苹果的脚步，在麦金塔计算机推出近4年后，发布了Windows 2.0，这是Windows系统的首个版本，外观和使用感都与今天许多人使用的版本很像。这款软件的目的就是让微软的个人计算机使用起来像麦金塔计算机一样。

攸关存亡的变通性

“用起来就像铅笔一样方便，你只需按下按钮，其余的我们来做。”柯达公司推出了有史以来第一款面向普通大众的相机，而这则广告几乎就是对柯达掌门人乔治·伊士曼愿景的总结。那是19世纪末，摄影几乎只属于专业人士和认真的业余爱好者，普通人没有办法自行拍摄

家人照片或度假照片。摄影设备又大又重，处理底片所需的化学物质带有剧毒。摄影是件既复杂又昂贵的事情，但伊士曼沉迷于为大众简化摄影的想法。尽管他早前已经改进了底片涂覆方式，但真正的突破还是在于，他发明了用一种叫赛璐珞（celluloid）的硝酸纤维素塑料完全取代厚重底片的方法。赛璐珞最初是用来制作假牙床的，但我们更熟悉的还是它的现代应用方式——胶卷。

在让摄影走入大众后，柯达成为全球最大的公司之一，伊士曼也成为世界最富有的人之一。1932年，伊士曼去世，但柯达公司仍在继续推进他的事业，仍在寻找更方便大众捕捉自己记忆的方式。1935年，柯达向大众推出了首款取得巨大商业成功的彩色胶卷。这也为彩色电影和彩色家庭自制影片的出现铺平了道路。柯达还发明了带圆形托盘的幻灯机，能让人们更轻松、更方便地分享他们的假期照片、婚礼照片，以及其他任何想要拉着亲朋好友一起坐下观看的照片。到20世纪60年代早期，柯达发明了胶卷暗盒，使摄影进一步简化、便利化。在此之前，胶卷是穿在相机卷轴上的，一些人觉得这不好操作或者害怕不够牢固。而现在，他们只需将胶卷暗盒“咔嗒”一声装到相机背后，就可以拿着相机去“赛跑”了。到1975年，柯达研发部门又有了一项极其了不起的发明——第一台数码相机。但这时却出现了一个问题。

数字化显然是该公司为推进其崇高的事业所迈出的下一步，但问题是，数码摄影技术的发明直接威胁到了公司现有的商业模式。柯达的盈利来源涵盖了摄影的方方面面。他们生产相机、胶卷、方块闪光灯（处理胶卷的设备）、冲洗胶卷所需的各种化学物质以及显影纸。众人皆知，这一新的数码摄影技术会令柯达当下的业务过时淘汰。若柯达当时的掌舵人是乔治·伊士曼，或任何一位无限思维的领导者，那这根本就不是个问题。他们会将这一新技术视为推进事业的更好方式，会想办法重新配置自己公司的业务。遗憾的是，当时的管理层没有理会公司一直以来的事业。有限思维成为主导。他们的决策不再是为了推进事业，而是为了管理成本和让短期财务利益最大化。

因为缺乏远见，当柯达高管们第一次看到数码技术时的第一反应是，人们是永远不会想在屏幕上看照片的。这些高管告诉工程师，人们喜欢让照片出现在纸上，相纸的使用不需要任何改进。被认为是数码相机发明者的年轻工程师史蒂文·萨森（Steven Sasson）并不认同，他拼命想让高管们想象一下摄影技术在未来20年或30年中可能的发展。令他沮丧的是，他的领导者对推进这项事业毫无兴趣，自然也对任何会改变现状的决定不感兴趣，尤其是在当时现状良好，能让他们每个人都赚得盆满钵满的情况下。他们不愿惹恼华尔街，不愿因“炸毁”公司而暂时陷入低谷，因此他们拒绝为了推进崇高的事业而让柯达重生为一家数码公司。

他们背弃了伊士曼的愿景，没有做出必需的攸关存亡的变通。他们决定压制新技术的发展，尽可能地拖延这一根本不可避免的事情。“当你跟一堆生意人聊公司在未来18年到20年里的发展时，他们根本提不起兴趣，因为到那时，他们都已不在这家公司了……每售出一台数码相机，都会减少售出一台胶卷相机，我们知道自己在胶卷上赚了多少钱。”萨森继续道，“当然，问题是，你的胶卷很快就会卖不出去了。这就是我的立场。”柯达的高管们没有选择领导这场数码革命，而是选择闭上眼睛、捂住耳朵，试图说服自己不会有事的。但这种情况并没有持续太久，也不可能持续太久，因为有限的策略永远用不长久。

无论如何，这只数码“精灵”都已经被从瓶子里放了出来，柯达预测，其他公司得花上10年左右才能真正将数码摄影这一技术变成有影响力的东西。他们是对的，在数码相机诞生的大概10年后，尼康公司才推出了一款单反相机，让用户可以给相机机身外接一个数码处理器（该处理器是柯达造的，因为他们拥有专利）。但真正推出首款全数码相机的是另一家比他们小得多的日本公司——富士，该相机的上市时间是1988年，距离柯达推出首款适合普通大众使用的胶卷相机过去了整整100年。后来，尼康与富士联手，继续创新和完善这项技术。又过去了大约10年，日本电子公司夏普推出了首款移动电话。再10年后，到20世纪初

期，数码相机和内置相机的移动电话成为主流。

柯达确实拥有许多与数码技术相关的原始专利。他们通过这些专利赚到了数十亿美元。这会给人一种错误的印象，即他们这家公司经营得很好。有限思维的领导者误以为拥有漂亮的资产负债表就等于拥有了强大的公司。但其实并不是，至少在无限游戏的背景下并不是。当2007年柯达的专利到期时，公司的资金就枯竭了，5年后柯达申请了破产保护。

破产往往是企业自杀的结果。回顾一下那些将一度非常成功的公司带上破产之路的决策，你会发现，其背后沉迷于有限游戏的领导者之多，真是令人不安。他们抛弃了企业为之奋斗的事业，拼命抓着旧的商业模式不放，这些模式确实曾给他们带去成功，却经不起时间的考验。企业总是用“市场状况”“新技术”或任何其他股票原因来解释自己的覆灭，但真正的原因大都不是这些。真正的问题在于领导者的无能，他们没能在必要时做出攸关存亡的变通。一旦抛弃了事业，也就一并抛弃了变通的能力。每个组织都会有必须做出变通的时刻。这不一定会发生在某一位特定领导者的任期内，但任何领导者都有责任让自己的企业具备攸关存亡的变通性，一旦需要变通，马上就能够实现。这意味着他们必须坚持把崇高的事业当作指路明灯，必须维持良好的企业文化，能包容和培养大量成员彼此信任的团队文化。

2012年1月19日，柯达在《纽约时报》上刊登了破产声明，声明一开篇就完美地总结了这一点：“柯达，131岁的胶卷先驱者，多年来在日益数字化的世界中举步维艰，已于周四早些时候申请了破产保护。”柯达首席财务官安托瓦内特·麦考维（Antoinette McCorvey）的一份声明让我们看到了柯达领导层到底在玩什么样的有限游戏。“自2008年以来，”该声明写道，“柯达尽了最大努力，但重组的成本和引发经济衰退的各种力量一直在给公司的流动资金状况雪上加霜。”一个曾经伟大的无限思维企业，领导者却抛弃了推进伊士曼的崇高的事业的道德责任，

自以为有责任去实现一些有限的野心。³他们将自己企业的未来交由市场力量而非对愿景的激情来决定。最终，整个公司、公司的所有员工以及公司的股东都不得不为此付出代价。

如今，柯达变成了过去那个自己的影子。柯达在发明数码相机时，拥有约12万名员工，而现在只有6000人左右。该公司仍然在生产胶卷，以及一切与胶卷处理相关的产品，但讽刺的是，如今他们的整个业务主要服务于一个群体——专业摄影师。于他们而言，这是最后一颗钉子。柯达已经完全抛弃了作为自己创立之基的事业。

没有崇高的事业的引导，柯达高管们也失去了为公司预见长远成功的远见和勇气。他们最多只能做到被周遭的世界戳一下才跳一下。乔治·伊士曼确实是大众摄影的发明者，柯达人也曾经是该行业几乎各个领域的先驱者。这家曾经伟大的公司覆灭在了他们自己发明的这个富有远见的技术之下，而真正的元凶只有一个，就是他们自己的有限思维。

THE INFINITE GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE
INFINITE
GAME

THE

要素 5

表现领导的勇气

第11章

领导的勇气

在公司总部的大堂里，挂着一个巨型标志，写着公司的愿景：让人们生活得更健康。公司的高管们相信这句话，认为自己的公司有着比赚钱更伟大的目标，他们希望借助公司推进更伟大的事业。他们会定期与医疗保健公司、医院和内科医生开会，讨论如何合作，从而更好地为患者服务。然而，在许多次会议快结束时，总有人矛头直指房间中的大象说：“但你们自己的商店里不卖香烟吗？”¹

2014年2月，CVS Caremark宣布将在其全部超过2800家门店中停售一切与烟草相关的产品。这一决定将使公司每年损失20亿美元的收入。他们是主动做出这一决定的，没有任何来自竞争的压力，没有公众的强烈呼声，没有任何丑闻，也没有任何迫使他们做出这一决定的网络运动。

该消息一出，就得到了广大民众的大力支持。但华尔街的专家们却都不太高兴。美国消费者新闻与商业频道的财经评论员吉姆·克莱默（Jim Cramer）说：“它在奥兹国⁶或许能赚到钱，但华尔街不是奥兹国。华尔街不会说‘你们知道吗？我打算买CVS的股票，因为他们是好市民’。我正在设法计算每股收益，CVS的每股收益就是变糟了。”²

外界的其他评论员也都持同一态度，认为这一决定对CVS的竞争者来说是一大助力。美国伊利诺伊州的一名销售与市场营销顾问指出，该决定等于放弃了每间药店每周700包香烟的销售量，而该销售量会流向其他零售商。该顾问补充道：“零售商知道，争取到这批成年的烟草消费者会同时带动店内其他商品的销量，因为消费者来买烟时很可能还会顺带买些别的东西。”³通过有限游戏与无限游戏的视角仔细研究上述针

对CVS此举的回应，我不禁认为，它们都源自极其有限的思维。如果商业游戏是一场有限游戏且未来很易预测，那么专家们的观点就百分之百正确。然而，事实证明，这场游戏是无限的，未来也难以预测。

事实上，CVS每家门店减少的每周700包香烟的销量并没有流向其他地方，香烟的总销量实实在在地下降了。CVS委托第三方对该决策的影响进行了独立研究，该研究发现，在CVS拥有15%或更高市场份额的各州，所有零售商的香烟总销量都下降了1%。⁴在这些州，吸烟者平均每人少买了5包香烟，也就是说，在该决定正式实施的8个月内，香烟销量总计减少了9500万包。另外，就在CVS停售香烟后的这段时间里，尼古丁贴片的销量增长了4%，这表明CVS的决定确实有助于吸烟者戒烟。至于收入减少的问题，过去，其他一些目标驱动型的公司曾拒绝与CVS合作，后来他们也注意到了CVS的决定。一些维生素和养生片剂等公司的产品，终于也在CVS旗下的药店上架了。⁵此举让CVS有了更丰富的优质品牌可供消费者挑选，也开辟了新的收入来源。当一家公开宣称以“让人们生活得更健康”为事业的公司真正做出了履行承诺的勇敢决定时，这个决定不仅让美国人生活得更健康了，而且对其药店的整体销量产生了积极的影响。

当然，影响CVS股价表现的因素还有很多，比如，CVS在做出上述决定后不久就更名为CVS Health。但在无限游戏中，财务健康就像锻炼一样，是无法用每天的步数来衡量的。这是一个稳定的持续积累的过程，假以时日，就能产生丰硕的成果。吉姆·克莱默指出，华尔街不会因为一家公司是“好市民”就去购买它的股票，但消费者和员工会。更加忠实的顾客与更加忠诚的员工往往会为该公司带来更大的成功。一家公司越成功，股东的收益往往也就越多。

其实，正如克莱默与其他分析师预测的那样，CVS的股价确实在该决定宣布后的第二天下跌了1%，从每股66.11美元跌至65.44美元，下跌后的第二天才回升回来。⁶不过，在该决定宣布的一年半以后，也就是

该方案正式实施的8个月后，该公司的股价创下了历史新高，达到每股113.65美元，是该决定宣布前的两倍。吉姆·克莱默格外担心的，被视为上市公司财务指标中的“黄金标准”的每股收益又表现如何呢？在该决定宣布前的2013年12月，CVS的每股收益为1.04美元。该决定公布后，跌至0.95美元；到下一季度，又回升至1.06美元；然后，在接下来的3年中上涨了70%，达到平均1.77美元。

在一个被有限思维主导的世界中，运用无限思维来做决定完全有可能令领导者丢掉工作。如今，人人都承受着压倒性的压力，这股压力在逼迫我们维持有限的思维。对大多数人来说，几乎所有的职业发展机会都是与我们在有限游戏中的表现相关联的。除此之外，还有很多因素，比如来自分析师群体的持续不断的警示，来自私募股权或风险投资公司的压力，与股价表现而非公司表现挂钩的高管薪酬（出人意料的是，股价表现与公司表现并不总是一致的），我们的自负及许多人自己给自己施加的压力（他们误以为自我价值感要靠自己在有限游戏中的表现来体现），以及我们对跳脱有限思维来行动的希望的完全破灭。我们很容易屈服于来自周围有限思维的玩家的压力，这也是一个更权宜的选择。正因如此，运用无限思维来做决定是一件需要勇气的事。

领导的勇气是一种意愿——愿意为了未知的未来而冒险，而这些风险是真实存在的。修改月度、季度或年度方案相对容易，但着眼于长远未来来做决定则会困难得多。这些决定确实可能让我们在短期内付出代价，这个代价可能是金钱，也可能是我们的工作。要想按照高于法律的标准，即按照道德的标准来经营公司，离不开领导的勇气。当我们被迫做一些违反道德准则的事情时，要有领导的勇气，我们才能大声反对，让那些给我们施压的人明白他们在制造何种状况；要有领导的勇气，我们才敢向他们提供或可解决这一问题的帮助；要有领导的勇气，我们才敢做出有违当前商业标准的决定；要有领导的勇气，我们才敢无视来自不投入或不相信崇高的事业的外部压力。

在无限游戏中，勇气不仅仅关乎我们所采取的行动。有限思维的领导者也可能敢于冒险。在提到无限思维时，领导的勇气就是指愿意彻底改变我们对这个世界运行方式的看法。这个勇气敢于反对米尔顿·弗里德曼所定义的企业，并拥抱另一种定义。当我们拥有了将有限思维变为无限思维的勇气时，所做出的许多决定在外界的传统视角看来，可能是过于大胆的，就像CVS停售香烟的那个决定一样。不过，我敢说，在那些用无限思维观察世界的人眼中，这样的决定才是显而易见且理所当然的。

那么，我们如何才能获得改变思维方式的勇气呢？

方法1：我们可以等待一段改变人生的经历，它会撼动我们的内心，挑战我们看待世界的方式。

方法2：我们可以去寻找一项能够激励自己的崇高的事业，与志同道合者为伍，或与能与我们彼此信任的人为伍，找到一个值得较量的对手，推动我们不断进步，提醒自己要更专注于所推进的事业，而非一味固守当前的道路或策略。

方法1确实是许多伟大领导者变成无限思维者的原因。无论是悲剧、机遇还是神谕，它一定能够促使你用与过去截然不同的方式来看待这个世界。这股推动你的力量有时会来得特别突然。不过，这种方法有点像赌博，我不建议我们只是日复一日无所事事地等待它的降临。

方法2给予了我们更多的控制权。我们所需的只是一点信念、一点自律以及去践行的意愿。对许多人来说，这种转变会带来很深刻的感受。但除了这种感觉，这种思维的转变确实会影响我们的决定和行动。对那些仍然在用有限思维看世界的人来说，我们的行动可能是理想主义的、天真的或愚蠢的。而在那些与我们有共同信念的人眼中，我们的行动是勇敢的。对游戏中的无限思维的玩家来说，这些勇敢的选择也是唯一可行的选择。

目标的力量

“她告诉我，我不能让我们的航空公司倒闭，她是个单亲妈妈，公司倒闭会重创她的生活。”他回忆道。

2001年9月1日，道格·帕克（Doug Parker）被任命为美国西部航空公司的新任首席执行官。10天后，“9·11”事件发生，重创了许多企业，航空业首当其冲。在接下来的两年中，美国航空客运量下降到了第二次世界大战以来的最低水平。美国联合航空和全美航空等公司纷纷申请破产保护。而对美国西部航空这种区域性的小型航空公司来说，他们并没有大航空公司那样的收入缓冲能力，看上去，公司就要彻底倒闭了。

帕克是第一批向当时新成立的航空运输稳定委员会申请政府贷款的人之一，该机构在“9·11”事件后向航空业提供了100亿美元（约合703亿元）的贷款，但进行得并不顺利。

在搭乘美国西部航空公司航班回家的路上，帕克感到十分沮丧。“情况看上去很不妙，”他回忆道，“我才刚当上美国西部航空的首席执行官，就马上要成为有史以来在任时间最短且最失败的首席执行官了。”为了暂时放下今天的烦恼，他决定去厨房和乘务员聊一聊。就在那里，他遇到了玛丽。玛丽是一名出色的乘务员，工作对她来说意味着一切。她所工作的航空公司无力熬过行业剧变完全不是她的错。“她能够避免严重个人危机的唯一希望，”帕克回忆道，“就是自己公司的领导能够想办法让公司继续运营下去。”

在遇到玛丽之前，避免公司倒闭对帕克来说只是件生意上的事，计算出维持公司运营所需的资金缺口只关乎对资源的管理。但在遇到玛丽之后，这变成了一项个人使命，也关乎意志了。帕克解释道：“若能投身于比自己更伟大的目标，我们或许就能完成一些若只为了自己就很可能无法完成的事。”在新发现的激情激励下，这位新任首席执行官和他的团队竭尽全力去争取，最终成功争取到了政府的贷款，这在他当初乘

飞机回家的路上还看似是一件完全不可能的事。为了进一步增强公司实力，获得更具竞争力的航线网络，帕克于2005年带领美国西部航空与全美航空合并，又在2013年与美国航空公司合并。“到那一刻，我的使命完成了，”帕克自豪地说，“美国航空是全球最大的航空公司。我们的团队终于安全了。”

但帕克总觉得缺了点什么。他说：“到2016年初，我发现我自己的工作目标产生了怀疑。我们已经实现了一个比自己更伟大的目标，现在我却不太满足于自己的工作。我只是在为了钱而工作吗？还是为了威望？对于这两个问题，我肯定都是不愿意给出肯定答案的。”帕克开始问自己，是否应该离开这家公司，去做一些别的事情，“可以更好地满足想为比自己更伟大的事业而工作的欲望”的事情。这种情况在极其成功的人群中很常见。他们在结束自己的职业生涯后，会继续付出，或是建立基金会，或是将自己的财富捐给慈善机构，他们是在努力满足自己想要回报社会的欲望，想要做一些更博爱的“事情”。但目标并不是只有在结束成功的职业生涯后才能找到。

帕克为玛丽及其同事们造福的动机，尽管具有令人难以置信的激励作用，却犹如一次登月计划。它是有终点的，一旦完成，就得重新寻找目标。帕克已品尝过为比自己更伟大的目标所驱使是何滋味，这点燃了他内心的激情，这份激情让他率领公司取得了前所未有的成功，不是为了他自己的荣耀，而是为了别人的利益。而他还想要体会那种感觉。

帕克听了制造企业巴里-韦米勒（Barry-Wehmiller）首席执行官鲍勃·查普曼（Bob Chapman）的演讲。查普曼相信最好的领导者和最好的企业都会以人为先，而非以数字为先，他也是该理念的倡导者。正是秉持着“人先于利润”的理念，查普曼所带领的公司一直在以超乎人们预期的表现增长，这为他赢得了很多演讲邀请，听众中既有相信者，也有怀疑者。正是查普曼的一场演讲令帕克顿然领悟，他找到了自己的“登月计划”，但没有找到这一“登月计划”的大背景。在帕克的职业生涯中，

给予下属工作的安全感和更高的薪水可能是一个重要的里程碑，但这并不是一项可以激励他一生的事业。“我们需要创造一个能让他们感受到关怀的环境！在这样的环境中，他们出色的工作会得到认可与赞赏，他们的领导关心在意他们；一天工作结束时，他们能够心满意足地踏上回家路。这就是我一直在寻找的比我自己更伟大的新使命。”找到无限追求的帕克激动不已地说道。

当全球最大航空公司的首席执行官有了改变自己领导方式的勇气后，当他的思维从有限转变为无限后，会发生什么呢？

与许多以数字为先而非以人为先的公司一样，美国航空公司在过去的很长时间内，都存在与员工间的信任问题。该公司的前领导团队以“帮助公司进行破产保护”为由，与工会谈判，希望工会做出重大让步，但与此同时，他们却向7位高管保证，只要这些高管能在公司多留几年，就能获得两倍于薪水的奖金。他们还从政府的拨款中拿出了4100万美元来保障45位高管的养老金，但没有任何保护普通员工的说明。

这件事最终导致时任首席执行官唐纳德·卡蒂（Donald Carty）辞职。⁸在他的离职声明中，他表达了对继任者的希望，希望他们能够努力建立一种“鼓励协作、合作和信任的新文化”。不过，尽管他的继任者杰拉德·阿佩（Gerard Arpey）和汤姆·霍顿（Tom Horton）都公开保证会这么做，却都没能做到。破坏信任的行为与可能的道德堕落仍然存在。除非新的领导团队愿意做出一些艰难的选择和牺牲，向员工证明他们确实是值得信任的，否则一切都不会改变。帕克明白，光有承诺改变的豪言壮语是没用的。他知道，他和他的领导团队必须有勇气去证明一切真的会有所改变。而这正是他们接下来所做的。

他们的第一个重大行动发生在2015年，他们重新协商了给飞行员和乘务员的合同，该合同将使他们的飞行员和乘务员成为业内薪酬最高的人群之一。然而，一年后，达美航空（Delta）和美联航都签署了新的飞行员合同与乘务员合同，令乘务员的收入高出美国航空5%，飞行员高

出8%。当时，该公司内的信任问题尚未解决，因此许多人误以为领导层早已得知这一消息，故意抢先一步和他们签订合同，用更低薪的合同锁住了他们未来的5年。

“说出你信任别人这句话只是口头的，”帕克说，“若要证明这种信任真的存在，我们就必须言行一致。”换作其他许多管理团队，遇到这一问题时，可能只会耸耸肩，承诺在下一次谈合同时解决这个问题。他们可能会说：“这难道不就是签订合同的目的吗？”然而，信任不是通过压力或强迫建立的，信任是通过与自己价值观一致的行动建立的，尤其是在信任最匮乏的时候。只有在我们做对的事情时，尤其是在我们主动去做时，才会产生信任。帕克与美国航空的总裁罗伯特·伊索姆（Robert Isom）共同发表了一份联合声明称，看到自己的员工在未来三四年中的收入都落后于行业平均水平，“新生的美国航空”觉得这“不正确，也与我们的承诺不符”。⁹

高管们决定要给所有乘务员和所有飞行员分别加薪5%和8%，而且不要任何回报。该决定将使公司在未来3年内多承担超过9亿美元的成本。

2017年4月27日，美国航空宣布了这一决定，华尔街的反应可想而知——反对。花旗集团的航空业分析师凯文·克里塞（Kevin Crissey）在给客户的信息中写道：“这真令人沮丧。劳工将再次优先获得报酬，股东只能得到残羹剩饭。”¹⁰摩根大通的分析师团队也写了一封信，表达了同样的看法：“我们对美国航空向其劳工团体转移近10亿美元的财富而感到不安。我们能理解美国航空想要与劳工这一利益相关者‘建立信任基础’的愿望，但我们认为最新的这份协议太过分了。在我们看来，面对更高的工资标准，解决的办法并不是去追赶它。有时，一个人做出承诺的时机完全看运气。”¹¹我相信他们想通过“看运气”一词表达的是“可能不公平，但运气确实站我们这边”。幸运的是，美国航空的领导层有勇气坚持自己“想要增强团队信任”的这一决定，并没有考虑克里塞和摩根

大通团队所建议的年度奖金制度。

遗憾的是，主宰市场的都是像克里塞、摩根大通分析师这样的有限思维者。据美国航空预计，他们的股价可能会下跌5%之多。而在消息公布的第二天，其股价实际下跌了9%。不过，好消息是，这往往也只有短期的影响。不到两周，美国航空的股价就完全恢复到了初始水平，到该年年底还上涨了超过20%。即便如此，许多的华尔街人士仍然主张，美国航空如果没给员工加薪，获得的利润还会更高。这再一次证明了在资源与意志之间，这些人偏爱前者。有限思维者无法理解的一点是：对人的投资终会令公司、客户受益，令这些投资收到回报（他们可能也没能意识到，起初的股价下跌也只是他们引导的结果）。

一家大型上市公司的首席执行官告诉我，华尔街分析师面向的目标群体往往是短线投资者，因此在写东西时更偏好能引起他们兴趣的内容——有限目标。在被问到如何看待那些短线分析师的唠叨时，帕克承认很难对它们完全无视。他认为我们“必须努力应对，因为我们可能会被迅速地卷入其中”。而好消息是，帕克、他的团队和董事会都在努力抵御外界的噪音，努力专注于公司的长期发展。“我们必须关心自己的团队，这样他们才会关心我们的客户，”帕克说，“这就是我们为股东创造价值的方式。”

美国航空仍处于这一新旅程的发展初期。但与过去相比，他们现在所宣扬的更多的是一种长远的理念，自然，这也吸引到了更多长线投资者的关注。这类投资者不太关心短期波动。特德·韦斯切勒（Ted Weschler）就是其中之一。韦斯切勒就职于沃伦·巴菲特的伯克希尔-哈撒韦公司（Berkshire Hathaway Corporation），该公司是出了名的长线投资商，鲜少出售自己的投资标的，韦斯切勒正是经营该公司的4名投资经理之一。事实证明，像伯克希尔-哈撒韦这样的长线股东都有自己的分析师，他们通常不会被24小时一循环的财经新闻所摆布。

被誉为“奥马哈先知”（Oracle of Omaha）的巴菲特是有史以来最成

功的投资者之一，也是世界最富有的人之一，在全球各个金融圈里都备受尊敬，他曾写道，航空公司是一个人可能做出的最糟糕的投资之一。正如他在2007年伯克希尔-哈撒韦致股东的信中所解释的：“最糟糕的一类企业是，发展迅速，但需要大量资本才能维持增长，然后又几乎或完全赚不到钱，比如航空公司。自莱特兄弟时代开始，我们就很难在这类公司身上找到一种持久的竞争优势。如果当初在基蒂霍克（Kitty Hawk）试飞时，有一名富有远见的资本家在场，他可能会帮后来者们一个大忙，直接把奥维尔·莱特射下来。”¹²值得一提的是，现在，伯克希尔-哈撒韦已经是美国航空的单一最大股东。当道格·帕克告诉他们，他打算给公司的乘务员和飞行员追加涨薪时，韦斯切勒给予了支持。可笑的是，所有那些抱怨帕克领导观的有限思维者，一旦认为投资美国航空可以赚钱，可能也会毫不犹豫地蜂拥而至。

坚持有限思维无须勇气

CVS决定要用自己崇高的事业来引导公司的发展，他们成为第一个冒险在自己店内停售香烟的公司。有了示范，其他人本应更容易效仿。然而，截至我写作本书时，它的两个最大竞争者仍然在货架上摆放着待出售的香烟。我先假定他们的做法与其事业是相符的。这两家公司都是经营药店的，他们选择坚持原有道路的原因很可能是他们所追求的事业与CVS不同。他们的决定或许是与他们所宣称的目标一致的。为了确定，我专门调查了一下。

沃尔格林药房属于沃博联（Walgreens Boots Alliance），在沃博联官网中，他们的宗旨是“帮助世界各地的人们生活得更健康、更快乐”。后面还写道，“沃博联致力于构建一个更健康、更快乐的世界。这一宗旨体现在以下核心价值观中”。¹³而其价值观的第一条是“信任：以尊重、正直和坦诚为原则，坚持做正确的事”。在被问到是否打算效仿CVS时，沃尔格林发表了一份声明，提到他们“致力于帮助想要戒烟的

顾客，因此决定主动在部分门店内缩小烟草产品的摆放空间，并将它们放到更不显眼处”。¹⁴沃尔格林“真勇敢”。

沃博联执行主席詹姆斯·斯金纳（James Skinner）也回应了同一问题，他说：“我们在定期研究这个问题，具体怎么做始终要看之后的研究结果。”¹⁵这难道不正好是缺乏勇气或信念的表现吗？如果所做出的决定与公司实际宣称的宗旨一致，那么斯奎那到底有什么可担心的呢？

美国疾病控制与预防中心的数据显示，吸烟是美国人的头号但可预防的杀手。¹⁶每年死于吸烟相关疾病的人数比死于艾滋病、违禁药、酗酒、车祸和枪支相关事故的总人数还要多！由此带来的经济成本也高得离谱。所有吸烟相关疾病每年总计要花掉美国纳税人超过3000亿美元。¹⁷美国国家航空航天局的航天飞机项目历时30多年，也才总计花掉了1960亿美元，其中还包括了建造6架航天飞机（其中5架已经飞向太空）的费用。美国每年与吸烟相关的医疗保健花费几乎是太空旅行费用的50倍！

如果一家石油公司必须为石油泄漏或管道裂缝所带来的损失负责，一家汽车企业必须为汽车设计缺陷所造成的伤害负责，那么烟草公司和出售烟草产品的商店难道不应该对这每年3000亿美元的成本负责吗？致力于帮助人们保持健康的药店却在销售香烟这样高成瘾性且致癌的产品，由此给消费者健康带去的伤害，他们当然是负有一些责任的，不是吗？

要预防吸烟相关疾病可能造成的死亡和资金损失，有且只有一个最佳方法，那就是帮助吸烟者戒烟。大多数吸烟者偶尔也会有主动想要戒烟的时候。近70%的吸烟者都表示自己想要戒烟，这其中有许多人都是在药店买烟的。¹⁸戒烟并不容易，显然有许多人都戒得很艰难。这就是在香烟旁边提供戒烟计划并不会有多大帮助的原因所在。这有点像在减肥产品旁卖甜甜圈。消费者面前有两种选择，一个是他们渴望且容易冲动购买的商品，另一个是必须依靠他们的自制力与辛苦付出才能完成的

目标。任何真心想要帮忙的人都会将那个容易勾起他们冲动的物品撤走，努力让那个更艰难的选项变得更容易完成一点，即使这会令帮忙的人付出一定的代价。这就是真正的领导的勇气！

领导者说出来的事业或目标，必须是他们真正相信的，这一点至关重要。拥有事业或目标声明的全部意义就在于，那必须是他们真正相信的。他们真正相信自己的企业拥有比赚钱更伟大的目标。一项事业要前进，就必须有人去推进。如果他们不打算付诸行动，那把事业写在墙上或挂在网站上又有何意义呢？

有越来越多的人说他们想要为目标驱动型的组织工作，尤其是千禧一代和Z世代（Gen Zers）^[7]。不过，若没有愿意挑战职场公认准则的、立场坚定的无限思维的领导者，事业声明就只会是一场自我感觉良好的营销，企业可能只是用它来讨好组织内外的人，并不一定真的相信它或真的会去践行它。或许，企业领导者们也像普林斯顿神学院的那些学生一样承受着压力，只是他们的压力是必须做出业绩。企业领导者若真的对无限思维毫无兴趣，或是做不到就算弄不明白也至少能对其保持开放的态度，那么至少可以勇敢地说出自己的真实意图，删掉网站上、营销中那些空洞的事业或目标声明，坦诚自己的短期目标其实也是在以正直为原则构建信任，一如沃尔格林在其价值观中所解释的那样，但这也是件需要勇气的事。

在CVS宣布了自己的决定后，同为美国三大连锁药店之一的来德爱也被问到了是否会效仿CVS。毕竟这一做法也符合来德爱所宣称的公司目标。在其官网“我们的故事”栏目中，第一句话就是：“在来德爱，我们的个人兴趣就是你的健康和福祉。正因如此，我们为你，我们重要的顾客，提供的所有产品和服务都是为了让你过上更健康、更快乐的生活。”¹⁹然而，当被问及他们是否打算效仿CVS，让自己的药店停售香烟时，该公司发表了一份米尔顿·弗里德曼风格的声明：“来德爱提供种类多样的产品，包括烟草产品，它们的买卖都是符合美国联邦法、州法和

地方法规的。”²⁰

让我们稍微想一想。一个公司在回答一个道德问题（或为一个不道德的决定辩护）时，用的解释竟然是他们的做法是合法的，这种辩护是无益于带来或重建信任的。

当企业及其领导者表现出了勇气与正直，证明了自己的诚实与坚定时，往往就会赢得客户和员工的意志与信任。就在CVS宣布将在自己的所有药店中停售香烟的第二天，玛丽雅利塞·萨恩斯（Maryalyce Saenz）办公桌上的电话响了，是她妈妈打来的，妈妈几乎是哭着告诉萨恩斯的，她对女儿能为CVS这样的公司工作感到无比自豪。多年来，萨恩斯父亲吸烟的习惯一直是家庭矛盾的根源。“这真是一个非常勇敢的决定，”萨恩斯解释道，“那天来上班的时候我真的感觉很自豪。然后我仔细想了想整件事，就在那天，我坐回位子上，想着，‘这绝对就是最适合我的位置’。”²¹可以肯定地说，一家公司的守法行为并不能给自己的员工和顾客带来同样温暖且难以言喻的感受。

即使不受周围有限玩家的欢迎，仍敢于将企业目标看作比赚钱更具英雄气概的事情是很难的。真正的领导的勇气会令公司及其领导层按照比法律高得多的标准来要求自己。只有当组织按照比法律更高的道德标准运营时，我们才能说他们是正直的。顺便一提，这才是正直的真正定义——坚守尤其符合道德或艺术价值观的准则：不腐败。的确，追求崇高的事业是一条正直的道路。这意味着言行必须一致，也意味着，外界有时会传来要求企业服务于那些未必相信其事业的人的利益的声音，而领导层必须选择无视这些呼声。

正直并不仅仅意味着“做正确的事”，还意味着要在公众大声抗议或丑闻爆发之前就采取行动。若领导者发现了某些不道德的事，却只在引发了强烈抗议后才采取行动，那就不是正直，而是控制损失。“他们在等着公众舆论告诉他们该怎么做，”哈佛商学院教授罗莎贝斯·莫斯·坎特（Rosabeth Moss Kanter）在谈到今天的首席执行官如何做决策时

说，“首席执行官是缺乏勇气的。”²²

分歧与十字路口

人是复杂的且不完美的。世界上并不存在完全无限思维的领导者，当然也不存在完全无限思维的组织。事实上，就连最专注于应用无限思维的企业也可能会走上一条有限思维的道路。当这种情况发生时就需要有领导的勇气，勇敢承认偏离了自身事业的事实，若要将组织带回正轨，同样需要领导的勇气。

很遗憾，在一个组织取得了巨大成功之后，上述这种情况是很常见的。但无限思维的玩家明白，无论自己在传统意义上有多成功，这些成就都只是冰山一角，而有限思维的玩家往往会为了守住自己当下的地位而转向防守。一个企业若想在到达顶峰后，继续留在这场无限游戏中，就必须有领导的勇气。要认识到，无论取得了多大的成功，自己的事业都是无限的。遗憾的是，转向有限思维的诱惑太大。

迪士尼公司也曾一度偏离了自己的无限事业，转而追求更有限的目标，比如主导全球市场、提高股东价值、让那些选择投资自己的人有钱可赚。1993年，迪士尼公司收购了米拉麦克斯影业（Miramax Films），随后就制作了一些适合家庭观看的电影，比如昆汀·塔伦蒂诺（Quentin Tarantino）执导的犯罪电影《低俗小说》（*Pulp Fiction*）、丹尼·鲍尔（Danny Boyle）的黑色喜剧《猜火车》（*Trainspotting*）等，还创办了唱片品牌好莱坞唱片（Hollywood Records）。

每次更换首席执行官，新任领导者的面前都会出现一个十字路口。他们将如何领导这个企业？迈克·杜克和史蒂夫·鲍尔默在各自执掌沃尔玛与微软期间，都选择将公司带上了一条有限思维的道路。这两家公司如果一直坚持这样的道路，可能早已被这场游戏彻底淘汰了。道格·麦克米伦与萨提亚·纳德拉（Satya Nadella）在分别接掌了沃尔玛与微软之

后，做了相同的选择，竭尽全力将各自的公司带回到无限思维的正轨上来。尽管前方仍然挑战重重，但这二人似乎都真心地致力于领导一项事业，而不仅仅是经营一家公司。

公司上市或领导层变更等重大事件也可能会迫使组织更换道路。不过，就算没有大事发生，一个组织也可能从无限走向有限。这样的偏离或分歧其实很常见。人们时时刻刻都在偏离自己的道路。比如，我们常常会违背一些有益健康的常规做法，或其他一些健康潮流。公司是由人管理的，所以可想而知，偏离是难以避免的。组织偏离正轨的原因往往有一个共性，就是领导者对自身有限追求的兴趣超过了对无限游戏的兴趣，这会将整个组织都拖下水。

若领导者相信了关于他们自己的神话——公司的成功全部源自领导者个人的天才，而非员工在他们所领导的事业激励下发挥出的集体的天才，那该组织就走到了命运的十字路口。这些领导者往往会沉迷于为自己赢得更多的名声、财富和荣耀，以至于可能会牺牲公司的利益和公司的事业。管理层开始与员工脱节，信任崩塌。业绩势必会受此影响，而当业绩下滑时，这些领导者会先指责他人，而非先弄清楚公司变更道路的最初原因是什么。为了“解决”问题，他们对员工的信赖变成了对程序的信赖。公司变得更加僵化，一线员工往往会失去决策权。船长本应待在甲板上，把握航行的大方向，若他弄错了自己的角色，进入船舱摆弄引擎，试图加快船只的航行速度，那一定不是什么好事。

Facebook曾经是个无限思维的玩家，现在似乎也正沿着一条有限思维的道路在前进。Facebook成立于2004年，创立之初就有着非常清楚的事业追求：“赋予人们构建社群的力量，让世界更紧密地联系在一起。”然而，现在的它深陷丑闻，所做之事完全违背了初衷。这难道就是马克·扎克伯格所说的“赋予人们力量”吗？Facebook之所以偏离了曾经激励人心的无限道路，是不是迫于来自华尔街的巨大压力，才会让领导者认为自己必须为那些有限的期望负责？是不是因为在一个由广告销售

驱动的商业模式中投入得太多，才没能及时做出重塑企业所需的攸关存亡的变通？领导者是不是失去了与企业愿景之间的联系？是不是失去了与那些要继续游戏就必须首先服务的对象之间的联系？是不是因为傲慢？不管是哪些因素导致Facebook走上了有限的道路，有一个事实是肯定的：他们的行动思维比过去更有限了。大而富并不意味着这家公司就不会失败。在这场永无止境的游戏里，金钱确实能推迟一些不可避免之事的发生，但也能给领导者提供让一切重回正轨的机会，唯一的问题是他们能不能把握住机会。只要具备一点点领导的勇气，他们就能在为时太晚之前，及时赢回曾助他们获得过成功的人的信任。

正如微软、沃尔玛和迪士尼等公司的例子所示，在偏离正轨后，企业还可以撑上一段时间。在此期间，他们将面临着重回无限道路的挑战。尽管面对偏离正轨所带来的成本，有些企业确实能比其他企业撑得更久一些，但钱终究有花完的一天，并不是每个组织都能在偏离无限道路后撑上那么长时间的。无论企业的规模大小，最能防止他们偏离无限道路的都是无限思维的领导者的那些必备要素，这些要素也是我一直试图在本书中阐释清楚的。无限游戏中玩的不是成就清单，而是思维方式。

如何获得领导的勇气

在我所有失败的人际关系中，唯一的共同因素就是我自己。在有限思维的领导者面临的所有挣扎与挫折中，唯一的共同因素就是他们的有限思维。承认这一点是需要勇气的，敞开心扉去接受另一种世界观就更需要勇气了，尤其是在我们知道自己的许多选择都会带来糟糕的结果时。要真正运用无限思维来重塑组织文化似乎更需要难以想象的勇气。事实确实如此。因为承认出问题的部分原因在自己可能会给我们带来尴尬，甚至羞耻，但决定参与解决该问题也可以赋予我们力量与鼓舞。

几乎没有人，或者只有很少的人，能够独自完成从有限思维到无限

思维的转变。我们必须找到与我们有共同责任感的人，找到同样认为是时候做出改变的人，找到同样想要通过合作来完成改变的人。我在这里所举的每个例子都在告诉我们，那些艰难的决定并不是由某位伟大的女性或伟大的男性做出的。做出这些决定的是伟大的合作伙伴和伟大的团队，是团结在一起、深信彼此且拥有共同事业的伟大的人们。没有哪位世界知名的空中飞人表演艺术家会在没有保护网的情况下尝试全新的死亡挑战，同理，要获得领导的勇气，我们离不开他人的帮助，那些与我们有着相同信仰的人就是我们的保护网。

勇敢的领导者之所以强大，因为他们知道自己没有掌握所有的答案，也没有掌握百分之百的控制权。不过，他们有彼此，以及引导他们前行的崇高的事业。只有软弱的领导者才会选择权宜之路。他们自认为掌握了所有的答案，或试图控制一切变量。面对完成预设目标的压力，与探索可能无人验证过的其他选择相比，用年末裁员来迅速压缩成本的做法，显然并不需要多少勇气。只有当一个组织的领导者拿出了领导的勇气，内部员工才会逐渐表现出同样的勇气。就像孩子会模仿父母一样，员工也会效仿领导者。领导者若把自己看得比团队更重要，那么在他所培养的文化中，员工也一定会将自己的发展置于公司的健康之前。唯有领导的勇气才能催生领导者的勇气。

后记

用无限思维玩无限游戏

我们的生命是有限的，但生活是无限的。我们是生活这一无限游戏中的有限玩家。我们来来去去，生生死死，无论有没有我们，生活都将继续。这场游戏中还有其他玩家，有些是我们的对手。我们能享受到胜利的喜悦，也会遭遇失败的痛苦，但无论输赢，第二天，我们仍会继续游戏。无论赚了多少钱，积攒了多大的权力，升职过多少次，我们都永远不会得到“生活赢家”的头衔。

在其他的游戏中，我们都有两种选择。尽管我们不能选择游戏的规则，但可以选择玩不玩，以及怎么玩。生活的游戏却有点儿不同。在这个游戏中，我们只有一种选择。我们一出生就会自动成为玩家，唯一可以选择的就是我们想要在这场游戏中运用有限思维还是无限思维。

如果选择用有限思维来生活，我们的首要目标就会是比别人更富有或升职更快；如果用无限思维来生活，我们的目标就会是推进一项比自身更伟大的事业。我们将那些与我们有着相同愿景的人视为事业伙伴，努力与他们建立彼此信任的关系，希望能够共同推进这项对众人都有益的事业。我们对自己所享有的成功心怀感恩，在自我进步的同时也会努力帮助周围的人进步。用无限思维来生活就是要过服务他人的生活。

记住，在生活中，我们同时参与着多个无限游戏，我们的职业生涯只是其中之一。除此之外，在养育子女、友谊、学习和创造力等方面也不存在所谓的赢家。但我们可以选择用什么样的思维方式来处理这些事情。在养育子女时，若选择有限的思维方式，我们就会不择手段地为孩子提供最好的一切，并确保他们样样都出类拔萃。这看上去似乎很合理，毕竟这些都“将帮助我们的孩子在生活中出类拔萃”。只是，当有限

思维变成我们的首要战略时，道德就可能堕落，我们就可能会因为迫于压力，过分执着于孩子在某个制度中的等级高低，而忘记关心他们作为人是否在学习或成长。

临床心理学家温迪·莫格尔（Wendy Mogel）分享了一个极端的例子。在一次会议上，她正在发言，一位父亲突然举起了手，告诉她，他曾因为儿子的阿普加评分（apgar score）与儿科医生吵了起来，最终吵赢了。阿普加评分是对新生儿健康状况的检测，要在其出生后的1~5分钟内进行。莫格尔解释道：“如果他们的皮肤是青紫的，摸上去是松软的，就得1分；如果是粉色的，摸上去是饱满的，就得5分。”我们稍微想一想就会发现，这位家长似乎更关心“赢”，更想要让自己刚出生的孩子拿个高分，而非真正关心孩子的健康。想想在未来的18年中，这位家长可能会为了让孩子一直得到最高分和一直就读最好的学校而做到何种程度？以及又会对孩子是否有真正学到东西和孩子方方面面的健康忽视到何种程度？

无限思维的子女养育方式与之相反，我们会帮助自己的孩子发现他们的才华，引导他们找到激情所在，并鼓励他们沿着那条路向前走。这意味着，要教孩子懂得服务的价值，教会他们如何交朋友，以及如何与他人打成一片；这意味着，要让孩子懂得教育是持久的，不会因为从学校毕业而结束，教育会贯穿他们的一生，但不一定总有课程或分数的引导；这意味着，要让孩子懂得如何运用无限思维去生活。在这场养育子女的无限游戏中，没有什么贡献是比让孩子能在我们离开后继续成长，继续为他人服务更伟大的了。

用无限思维去生活意味着，在做决定时要考虑到这些决定可能产生的间接影响与再间接的影响，在考虑给谁投票时要换一个不同的视角。一旦做了决定，就要对其后续的影响负责。

与其他所有的无限游戏一样，在生活这场游戏中，我们的目标不是赢，而是让这场游戏继续下去。我们要过一种服务他人的生活。

没人会想在自己的墓碑上刻下银行账户的最终余额。我们希望人们记住的是我们对他人的付出，是尽心尽责的母亲、慈爱的父亲或忠诚的朋友的形象。服务他人才是有利于这场游戏的做法。

在生活的这场无限游戏中，我们只能做一种选择。你会选择什么？

* * *

如果本书有激励到你，请把它分享给另一个你想要激励的人。

致谢

想法会不断进化。它们不像灯，按下开关就会突然亮起。它们也不是随机生成的，我们的想法往往与别人问我们的问题或我们正在努力解决的难题有关。如果真有一个瞬间顿悟的时刻，那也一定会出现在我们大量阅读、大量观看、大量倾听、大量交谈之后，这一切都会推动、激励、指引我们走向想法可能成形的方向。这本书的诞生正是如此。

写作本书的种子是在许多年前种下的。当时，我的好友布莱恩·柯林斯（Brian Collins）送了我一本詹姆斯·卡斯的《有限与无限的游戏》。卡斯，感谢你写了这么一本薄薄的但神奇的书。这个想法令我着迷，它开始影响我看待世界的方式。随后，我买了几十本送给朋友，送的都是一些我认为会理解这一不同观点的人。其中之一就是兰德公司（RAND Corp）的安迪·霍恩（Andy Hohen）。我和霍恩长谈了很多次，讨论如何用无限游戏这一新视角去观察分析全球政治和军事战略。戴维·谢德（David Shedd）是位智慧的思想者，在政府部门工作了很久，他提出了一些棘手的问题，帮我进一步完善了我的想法。我很幸运地接到了美国空军准将布莱恩·霍尔特（Blane Holt，现已退役）和高级行政官员迈克·瑞安（Mike Ryan）的邀请，参加了美国欧洲司令部（EuCom）在德国举办的会议，并在那里分享了我的观点：如何利用无限思维更好地理解美国在世界中所扮演的角色。

后来，在一次企业家大会上，我在赛斯·戈丁（Seth Godin）演讲的激励下，决定放弃自己既定的剧本，做一些新的尝试。那是我第一次将无限思维应用到商业中。显然，我们需要的远不止一个观察世界的新视角——在当今世界中，所有人，或者至少大多数人，都在参与着这样或那样的无限游戏，我们需要理解“领导”在这样的一个世界中究竟意味着

什么。

这个想法不断丰满，终于到了需要检验它的时候。我找到了一些早期参与者，他们愿意冒险让我站到现场观众面前分享我尚未完全成形的想法。在加利福尼亚州的棕榈泉，鲍勃·巴顿（Bob Patton）允许我在其公司的战略发展论坛上分享我的想法；TED演讲为我提供了观众；谷歌也给了我与其员工深入探讨的机会；威廉-莫里斯奋进娱乐公司

（William Morris Endeavor）鼓励我用无限思维的真正含义去挑战他们的领导层。虽然缓慢，但我确实在这个过程中更确定了适用于无限游戏的领导方式。有太多太多的人与我进行过思维的碰撞，给过我在真实观众面前检验自己想法的机会，真心感谢你们。

当我终于把这个想法告诉我的出版商阿德里安·扎克赫姆（Adrian Zackheim）时，他的反应一如既往，笑着对我说：“我会出版它。”阿德里安，衷心感谢你，又一次为我关于这个世界运行方式的疯狂想法下了赌注。接着，真正的写作工作开始了。

写书是研究与写作的结合，一旦开始，就需要进行更多的对话和辩论，不断地修改或重写。这个过程充斥着各种情绪，所有情绪。在我经历所有这些情绪的煎熬时，让·哈勒姆（Jenn Hallam）一直在身边支持着我。从一开始就陪在我身边的老友，是你推着我不断完善我的想法，是你帮着我让文字表达得更加清晰。如果没有你，我不可能写完这本书。谢谢你，难以言喻的感谢，谢谢你。

当我深陷在写作的无底洞中时，我的团队会帮我料理其他的一切。萨拉·托博罗斯基（Sara Toborowsky）、金·哈里森（Kim Harrison）、洛里·杰克逊（Lori Jackson）、梅利莎·威廉姆斯（Melissa Williams）、莫莉·斯特朗（Molly Strong）、莫妮克·赫尔斯特罗姆（Monique Helstrom）、莱拉·苏西（Laila Soussi）以及团队里的其他成员，感谢这么多个月以来你们对我的耐心与悉心的照顾，感谢你们帮我处理其他一切需要处理的事情。

特别感谢汤姆·斯塔格斯（Tom Staggs），你给了我非常非常多宝贵的时间，帮我完善我的想法，完善这本书。同样，我也非常重视你的建议，非常珍视与你的友谊。我还要感谢美国海军陆战队中将乔治·弗林（George Flynn，已退役），在我写作本书的整个过程中，你一直在我身边支持我，从与我一起修改大纲，到与我一起阅读终稿。还有汤姆·加德纳（Tom Gardner）和莫特利富尔投资网站（Motley Fool）的所有同事，感谢你们与我分享了大量的知识。

还有亚当·格兰特，我值得较量的对手，我亲爱的朋友。你是如此出色的写作者，激励着我变得更好。感谢我的事业伙伴鲍勃·查普曼。我们手中的火炬正一天比一天更亮。

感谢斯特夫（STRIVE）摩洛哥活动的全体成员。我是在与你们一起沙漠骑行时受到的启发，那是我第一次想要探讨什么样的人生才是无限的人生（这可能与我那天一大早骑车爬山的感受有一定的关系吧）。

感谢与我分享了自己想法与故事的人，谢谢你们让这本书变得生动：安吉拉·阿伦茨、克里斯汀·贝茨（Christine Betts）、杰克·考利局长、杰克·科伊尔警官以及我在岩堡市警察局遇到的所有杰出的人、萨沙·科恩（Sasha Cohen）、约翰·库奇、美国海军舰长里奇·迪维尼（Rich Diviny，已退役）、卡尔·埃尔森纳、杰夫·伊梅尔特（Jeff Immelt）、柯蒂斯·马丁（Curtis Martin）、史蒂夫·米切尔（Steve Mitchell）、艾伦·穆拉利、道格·帕克、乔·罗德（Joe Rohde）、美国陆军少校威廉姆·斯温森（William Swenson）以及劳琳·萨金特（Lauryn Sargent）和斯科特·汤普森（Scott Thompson），谢谢你们。特别感谢基普·廷德尔，不仅仅因为你分享的故事，还因为你对我的信任和鼓励。

感谢对我敞开心扉，然后挑战并推动我前进的人——萨拉·布莱克利（Sara Blakely）、琳达·博夫（Linda Boff）、美国空军上将凯文·奇尔顿（Kevin Chilton，已退役）、美国空军上校迈克·德劳利（Mike Drowley）、伊利斯·艾贝尔瓦恩（Elise Eberwine）、艾尔·圭多（Al

Guido)、布莱恩·格雷泽 (Brian Grazer)、戴维·科特金 (David Kotkin)、美国海军陆战队舰长莫林·克雷布斯 (Maureen Krebs)、贾米尔·马哈茂德 (Jamil Mahoud)、美国海军直升机海上打击队的C.K.摩根 (C.K. Morgan, 你不知道的是, 正是你的感谢信让我重写了整个大纲)、埃西·诺思 (Essie North)、美国空军少将大卫·罗宾逊 (David Robinson, 已退役)、美国空军上将洛里·罗宾逊 (已退役)、黛西·罗宾顿 (Daisy Robinton)、克雷格·拉塞尔 (Craig Russell)、珍·沃尔德曼 (Jen Waldman)、凯文·沃伦 (Kevin Warren)、迈克·沃斯 (Mike Wirth)。由衷地感谢你们。

感谢美国空军、陆军、海岸警卫队、海军以及海军陆战队的各级领导, 是你们考验了我的魄力。

而最重要, 也最要感谢的就是你们, 我的读者。感谢那些加入了 我, 与我一起推进崇高的事业的人。我很荣幸能为你们服务, 我们将携手共建这样一个世界: 绝大多数人都会在想要工作 的热情中醒来, 能安心地在外工作, 能带着工作的满足感踏上回家的路。

让我们一起加油!

注释

第1章

1. Janet Guyon, “British Airways Takes a Flier,” September 27, 1999.
2. Daniel Eran Dilger, “Microsoft Abandons Zune Media Players in Defeat by Apple’s iPod,” March 14, 2011.
3. Jonathan Ringen, “How Lego Became the Apple of Toys,” *Fast Company*, January 8, 2015.
4. Rick Wartzman, *The End of Loyalty: The Rise and Fall of Good Jobs in America* (New York: PublicAffairs, 2017), 20–21.
5. Epoch Times staff, “Staying True to Values: Interview with Carl Elsener Jr., Victorinox CEO,” *Epoch Times*, August 8, 2016.
6. Tim Beyers, “Too Zune for Hype,” *Motley Fool*, November 20, 2006. Dan Frommer, “Apple iPod Still Obliterating Microsoft Zune,” *Business Insider*, July 12, 2010.
7. Meg Prater, “9 Brands that Survive Without a Traditional Marketing Budget,” *HubSpot*, July 17, 2017.
8. Rachel Rosmarin, “Apple’s Profit Soars on iPod Sales,” *Forbes*, July 19, 2006.
9. Jay Yarrow, “Here’s What Steve Ballmer Thought about the iPhone Five Years Ago,” *Business Insider*, June 29, 2012.
10. Kurt Eichenwald, “Microsoft’s Lost Decade,” *Vanity Fair*, August 2012.

- [11.](#) Mary Jo Foley, “For Steve Ballmer, a Lasting Touch on Microsoft,” *Fortune*, December 10, 2013.
- [12.](#) Matt Weinberger, “How Microsoft CEO Satya Nadella Did What Steve Ballmer and Bill Gates Couldn’t,” *Business Insider*, January 30, 2016.
- [13.](#) Kurt Eichenwald, “Microsoft’s Lost Decade,” *Vanity Fair*, July 24, 2012.
- [14.](#) Stéphane Garelli, “Why You Will Probably Live Longer Than Most Big Companies,” IMD, December 2016.
- [15.](#) Kim Gittleson, “Can a Company Live Forever?,” BBC News, January 19, 2012.
- [16.](#) Simon Sinek, *Leaders Eat Last* (New York: Portfolio/Penguin, 2017) (Glass-Steagall and Stock Market Crashes).

第2章

- [1.](#) Carolyn Fry, *Seeds: A Natural History* (Chicago: University of Chicago Press, 2016), 30–31.
- [2.](#) Jules Janick, “Nikolai Ivanovich Vavilov: Plant Geographer, Geneticist, Martyr of Science,” *HortScience* 50, no. 6 (June 1, 2015): 772–76.
- [3.](#) Gary Paul Nabhan, *Where Our Food Comes From: Retracing Nikolai Vavilov’s Quest to End Famine* (Washington, D.C.: Shearwater, 2009), 10.
- [4.](#) Michael Major, “The Vavilov Collection Connection,” Crop Trust, March 19, 2018.
- [5.](#) “Irvine California Jobs,” Vizio, careers.

第3章

1. President John F. Kennedy, Moon speech, Rice University, Houston, Texas, September 12, 1962, NASA.
2. Jim Collins, *Good to Great: Why Some Companies Make the Leap... and Others Don't* (New York: HarperCollins, 2001). Jim Collins and Jerry I. Porras, *Built to Last: Successful Habits of Visionary Companies* (New York: HarperCollins, 2004).
3. Garmin, “About Us,” “Our Vision,”.
4. 这个例子源自销售教练杰克·戴利（Jack Daly）。
5. Eric Paley, “Venture Capital Is a Hell of a Drug,” *Tech Crunch*, September 16, 2016.
6. Robert J. Samuelson, “Capitalism’s Tough Love: The Real Lessons from the Fall of Sears and GE,” *The Washington Post*, January 13, 2019.
7. Roger L. Martin, “M& A: The One Thing You Need to Get Right,” *Harvard Business Review*, June 2016.

第4章

1. Barbara Farfan, “Overview of Walmart’s History and Mission Statement,” *The Balance Small Business*, July 25, 2018.
2. “Mike Duke Elected New Chief Executive Officer of Wal-Mart Stores, Inc.,” Walmart, November 21, 2008.
3. Josh Eidelson, “The Great Walmart Walkout,” *The Nation*, December 19, 2012.
4. Simon Sinek, “Why Too Many Successions Don’t Succeed,” *Huffington Post*, December 27, 2008.

5. Michael Dinkins, “What Jack Welch Taught This CFO about Leadership,” Spend Culture Stories Podcast, 2018.
6. “Doug McMillon Elected New Chief Executive Officer of Wal-Mart Stores, Inc.,” Walmart, November 25, 2013.

第5章

1. Milton Friedman, “A Friedman Doctrine—The Social Responsibility of Business Is to Increase Its Profits.” *The New York Times Magazine*, September 13, 1970.
2. Friedman, “A Friedman Doctrine.”
3. Adam Smith, *The Wealth of Nations*, Part Two (New York: Collier, 1902), 442.
4. Bryce G. Hoffman, *American Icon: Alan Mulally and the Fight to Save Ford Motor Company* (New York: Crown Publishing, 2012), 398.
5. Caroline Fohlin, “A Brief History of Investment Banking from Medieval Times to the Present,” in *The Oxford Handbook of Banking and Financial History*, ed. Youssef Cassis et al. (Oxford: Oxford University Press, 2014).
6. 这个比喻是汤姆·斯塔格斯提供的参考。

第6章

1. Diane Cardwell, “Spreading His Gospel of Warm and Fuzzy,” *The New York Times*, April 23, 2010.
2. Andrés Cardenal, “Higher Wages Could Pay Off for Wal-Mart Employees, Customers, and Investors,” *Motley Fool*, January 20.

- [3.](#) Zac Hall, “Retail Chief Angela Ahrendts Talks ‘Today at Apple’ and More in Video Interview,” 9to5Mac, May 17, 2017.
- [4.](#) Don Reisinger, “Here’s How Apple’s Retail Chief Keeps Employees Happy,” *Fortune*, January 28, 2016.
- [5.](#) Jonathan Fader, PhD, “Should We Give Our Kids Participation Trophies?,” *Psychology Today*, November 7, 2014.

第7章

- [1.](#) Angus Chen, “Invisibilia: How Learning to Be Vulnerable Can Make Life Safer,” NPR, June 17, 2016.
- [2.](#) Chen, “Invisibilia.”
- [3.](#) Robin J. Ely and Debra Meyerson, “Unmasking Manly Men,” *Harvard Business Review*, July-August 2008.
- [4.](#) Bryce G. Hoffman, *American Icon: Alan Mulally and the Fight to Save Ford* (New York: Crown Publishing, 2012), 110–125.
- [5.](#) Hoffman, *American Icon*, 121.
- [6.](#) 该等式的提出者为Lt. Gen. George Flynn, USMC (ret).

第8章

- [1.](#) Michael Corkery, “Wells Fargo Fined \$185 Million for Fraudulently Opening Accounts,” *The New York Times*, September 8, 2016.
- [2.](#) Chris Arnold, “Former Wells Fargo Employees Describe Toxic Sales Culture, Even at HQ,” NPR, October 4, 2016.
- [3.](#) Arnold, “Former Wells Fargo Employees Describe Toxic Sales Culture.”

- [4.](#) “Wells Fargo Workers Created Fake Accounts,” video, CNN Business, April 10, 2017.
- [5.](#) Matt Egan, “Wells Fargo Claws Back \$75 Million from Former CEO and Top Exec,” CNN Business, April 10, 2017. Independent Directors of the Board of Wells Fargo & Company, “Sales Practices Investigation Report,” April 10, 2017.
- [6.](#) Matt Egan, “Feds Knew of 700 Wells Fargo Whistleblower Cases in 2010,” CNN Business, April 19, 2017.
- [7.](#) Independent Directors of the Board of Wells Fargo & Company, “Sales Practices Investigation Report,” April 10, 2017, 55.
- [8.](#) Independent Directors of the Board of Wells Fargo & Company, “Sales Practices Investigation Report,” April 10, 2017, 13, 8, 46.
- [9.](#) Julia Horowitz, “Wells Fargo to Pay \$2.09 Billion Fine in Mortgage Settlement,” CNN Business, August 1, 2018.
- [10.](#) Emily Glazer, “Wells Fargo to Refund \$80 Million to Auto-Loan Customers for Improper Insurance Practices,” *The Wall Street Journal*, July 8, 2017.
- [11.](#) “U.S. Probing Wells Fargo’s Wholesale Banking Unit: WSJ,” Reuters, September 6, 2018.
- [12.](#) Wells Fargo & Company Annual Report 2016, 37, 88.3B.
- [13.](#) Matt Krantz, “Wells Fargo CEO Stumpf Retires with \$134M,” *USA Today*, October 13, 2016.
- [14.](#) Mark Maremont, “EpiPen Maker Mylan Tied Executive Pay to Aggressive Profit Targets,” *The Wall Street Journal*, September 1, 2016; Aimee Picchi, “Mylan Boosted EpiPen’s Price Amid Bonus Target for

Execs,” CBS News, September 1, 2016; Gretchen Morgenson, “EpiPen Price Rises Could Mean More Riches for Mylan Executives,” *The New York Times*, September 1, 2016.

- [15.](#) Morgenson, “EpiPen Price Rises Could Mean More Riches for Mylan Executives.”
- [16.](#) Catherine Ho, “CEO at Center of EpiPen Price Hike Controversy Is Sen. Joe Manchin’s Daughter,” *The Washington Post*, August 24, 2016; Matt Egan, “How EpiPen Came to Symbolize Corporate Greed,” CNN Business, August 29, 2016.
- [17.](#) Danielle Wiener-Bronner, “Mylan CEO: You Can’t Build a Company in a Quarter,” CNN Business, June 4, 2018.
- [18.](#) “Mylan Agrees to Pay \$465 Million to Resolve False Claims Act Liability for Underpaying EpiPen Rebates,” U.S. Department of Justice, Office of Public.
- [19.](#) Ann E. Tenbrunsel and David M. Messick, “Ethical Fading: The Role of Self-Deception in Unethical Behavior,” *Social Justice Research* 17, no. 2 (June 2004): 223–36.
- [20.](#) Tenbrunsel and Messick, “Ethical Fading,” 228–29.
- [21.](#) Matt Egan, “Elizabeth Warren to Wells Fargo CEO: ‘You Should Be Fired,’ CNN Business, October 3, 2017; “Wells Fargo Statement Regarding Board Investigation into the Community Bank’s Retail Sales Practices,” *Business Wire*, April 10, 2017.
- [22.](#) “Dr. Leonard Wong Discusses a Culture of Dishonesty in the Army,” STEM-Talk, Episode 29, Florida Institute for Human & Machine Cognition, January 17, 2017.

- [23.](#) Leonard Wong and Stephen J. Gerras, “Lying to Ourselves: Dishonesty in the Army Profession,” U.S. Army War College Strategic Studies Institute (Carlisle Barracks, PA: U.S. Army War College Press, 2015),.
- [24.](#) Tim Nudd, “Ad of the Day: Patagonia,” *Adweek*, November 28, 2011.
- [25.](#) Monte Burke, “The Greenest Companies in Fly Fishing,” FlyFisherman.com, February 1, 2016.
- [26.](#) Katya Margolin, “Could Patagonia’s Alternative Leadership Model Unleash the Best in Your People?,” *Virgin*, October 7, 2016.
- [27.](#) Patagonia, “Sustainability Mission/Vision,”.
- [28.](#) Patagonia, “Don’t Buy This Jacket, Black Friday and the *New York Times*,” November 25, 2011.
- [29.](#) Gillian B. White, “All Your Clothes Are Made with Exploited Labor,” *The Atlantic*, June 3, 2015.
- [30.](#) Margolin, “Could Patagonia’s Alternative Leadership Model Unleash the Best in Your People?”
- [31.](#) “Three Guides for Going B—and Why It Matters,” Patagonia.com, August 27, 2018.
- [32.](#) Jeff Beer, “How Patagonia Grows Every Time It Amplifies Its Social Mission,” *Fast Company*, February 21, 2018.
- [33.](#) “Clothing Company Tells Customers to Buy Less,” *PBS NewsHour*, August 21, 2015.

第9章

- [1.](#) Gwen Knapp, “Evert vs. Navratilova—What a Rivalry Should Be,” *San*

Francisco Chronicle, June 19, 2005.

2. Jonathan Fader, PhD, “Should We Give Our Kids Participation Trophies?,” *Psychology Today*, November 7, 2014.
3. Bryce G. Hoffman, *American Icon: Alan Mulally and the Fight to Save Ford* (New York: Crown Publishing, 2012), 109. Sarah Miller Caldicott, “Why Ford’s Alan Mulally Is an Innovation CEO for the Record Books,” *Forbes*, June 25, 2014.
4. Hoffman, *American Icon*, 127.
5. Hoffman, *American Icon*, 97–98.
6. Hoffman, *American Icon*, 139.
7. Bill Murphy Jr., “37 Years Ago, Steve Jobs Ran Apple’s Most Amazing Ad. Here’s the Story (It’s Almost Been Forgotten),” *Inc.com*, August 23, 2018.
8. Murphy Jr., “37 Years Ago, Steve Jobs Ran Apple’s Most Amazing Ad.”
9. John Douglas, *Mindhunter: Inside the FBI’s Elite Serial Crime Unit* (New York: Pocket Books, 1996), 56.

第10章

1. Neal Gabler, *Walt Disney: The Triumph of the American Imagination* (New York: Vintage, 2007), 10.
2. Gabler, *Walt Disney*, 492.
3. Tendayi Viki, “On the Fifth Anniversary of Kodak’s Bankruptcy, How Can Large Companies Sustain Innovation?,” *Forbes*, January 19, 2017.

第11章

1. Larry Merlo, “The Good and the Growth in Quitting,” TED Talk, Wake Forest University, YouTube video, 15:24, April 2015.
2. Jeff Morganteen, “Cramer: CVS’ Tobacco Move Won’t Fly on Wall Street,” CNBC, February 5, 2015.
3. CVS’ Tobacco Exit Draws Reaction, Applause,” *Convenience Store News*, February 6, 2014.
4. “We Quit Tobacco, Here’s What Happened Next,” Thought Leadership, CVS Health Research Institute press release, September 1, 2015.
5. Brian Berk, “CVS Pharmacy Unveils the ‘Next Evolution of the Customer Experience,” *Drug Store News*, April 19, 2017.
6. Andrew Meola, “Rite Aid (RAD) and Walgreen (WAG) Rise on CVS Caremark(CVS) Tobacco Announcement,” *TheStreet*, February 5, 2014.
7. CVS Health Corporation Revenue & Earnings Per Share (EPS), Nasdaq,
8. Edward Wong, “Under Fire for Perks, Chief Quits American Airlines,” *The New York Times*, April 25, 2003.
9. “A Letter to American Employees from Doug Parker and Robert Isom on Team Member Pay,” American Airlines Newsroom, April 26, 2017.
10. “The Case Against ‘Maximizing Shareholder Value,’ “NPR, May 6, 2017.
11. John Biers, “American Airlines Defends Pay Increase As Shares Tumble,” Yahoo Finance, April 27, 2017.
12. Quoted in Adam Levine-Weinberg, “7 Ways Warren Buffett Blasted the

Airline Industry—Before Investing Billions There,” *Motley Fool*, March 5, 2017.

- [13.](#) Walgreens Boots Alliance, “About us,” “Vision, Purpose and Values,”.
- [14.](#) Ronnie Cohen, “When CVS Stopped Selling Cigarettes, Some Customers Quit Smoking,” Reuters, Health News, March 20, 2017.
- [15.](#) Lisa Schencker, “Why Is Walgreens Still Selling Cigarettes? Shareholders Want to Know,” *Chicago Tribune*, January 26, 2017.
- [16.](#) “Health Effects of Cigarette Smoking,” Centers for Disease Control and Prevention.
- [17.](#) “Economic Trends in Tobacco,” Centers for Disease Control and Prevention.
- [18.](#) Centers for Disease Control Prevention (CDC), “Quitting Smoking Among Adults—United States, 2001–2010,” *MMWR. Morbidity and Mortality Weekly Report* 60, no. 44 (November 11, 2011): 1513–19.
- [19.](#) Rite Aid, “About Us,” “Our Story,”.
- [20.](#) Paul Edward Parker, “Rite Aid Responds to CVS Decision to Stop Selling Tobacco,” *Providence Journal*, February 6, 2014.
- [21.](#) CVS Health, CVS Purpose Short, YouTube, October 9, 2017.
- [22.](#) Steve Lohr and Landon Thomas Jr., “The Case Some Executives Made for Sticking with Trump,” *The New York Times*, August 17, 2017.

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

湛庐阅读App：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本
怎么读是读者面临的最大阅读障碍
“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”App。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者为伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。

从单一到复合，从知道到精通，从理解到创造，湛庐希望建立一个“与最聪明的人共同进化”的社区，成为人类先进思想交汇的聚集地，与你共同迎接未来。

与此同时，我们希望能够重新定义你的学习场景，让你随时随地收获有内容、有价值的思想，通过阅读实现终身学习。这是我们的使命和价值。

湛庐阅读App玩转指南

湛庐阅读App 结构图：



三步玩转湛庐阅读App：



使用App扫一扫功能， 遇见书里书外更大的世界！

大咖优质课、
献声朗读全本一键了解，
为你读书、讲书、拆书！

你想知道的彩蛋
和本书更多知识、资讯，
尽在延伸阅读！

快速了解本书内容，
湛庐千册图书一键购买！



延伸阅读

《指数型组织》

- ◎ 《指数型组织》是一本指数级时代企业行动手册。
- ◎ 作者奇点大学创始执行理事萨利姆·伊斯梅尔归纳了指数型组织的11个强大属性，并提出了建立指数型组织的12个关键步骤。



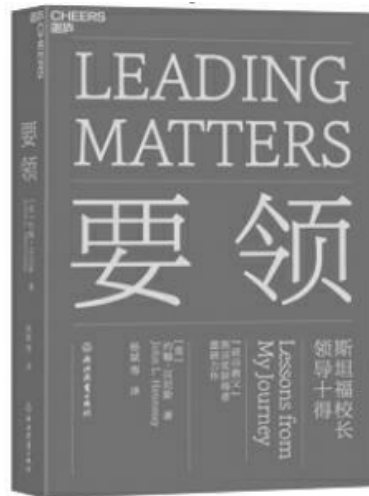
使用“湛庐阅读”App，
“扫一扫”获取本书更多精彩内容。

ISBN 978-7-213-06921-5



《要领》

- ◎ 《要领》是汉尼斯校长给“要成为领导者”之新生代的领导力讲义，总结自己“满脸污泥、汗水和血迹”地在竞技场中拼搏的习得，凝结成十条归真求是的“领导要义”。
- ◎ 斯坦福大学原校长、谷歌母公司Alphabet董事会主席、图灵奖得主、“硅谷教父”约翰·汉尼斯重磅力作，清华经管领导力研究中心主任杨斌教授领衔翻译。



使用“湛庐阅读”App，
“扫一扫”获取本书更多精彩内容。

ISBN 978-7-5536-8313-3



《奈飞文化手册》

- ◎ 一本对奈飞文化进行深入解读的力作。
- ◎ 这本书适合各类型的企业也同样适合管理者与普通员工。管理者发现自己的企业文化中存在的问题，找到改进的方向，员工则了解优秀的企业应该拥有怎样的文化，思考自己应该选择怎样的企业。



使用“湛庐阅读”App，
“扫一扫”获取本书更多精彩内容。

ISBN 978-7-5536-7805-4



《伟大企业的四个关键原则》

- ◎ 全食超市创始人约翰·麦基引领商业觉醒的力作，揭示建立伟大企业的4个关键原则。
- ◎ 世界500强美敦力公司前CEO比尔·乔后倾情作序推荐星巴克创始人、西南航空创始人联袂推荐！



-
- (1) 丹尼尔·平克的著作《驱动力》中文简体字版已由湛庐文化策划、由浙江人民出版社于2012年出版。——编者注
 - (2) 约翰·库奇时任苹果教育副总裁，他的著作《学习的升级》中文简体字版已由湛庐文化策划、浙江人民出版社于2019年出版。——编者注
 - (3) 列宁格勒是圣彼得堡的旧称。——译者注
 - (4) 想要更深刻地理解“为什么”，并找到“为什么”，可以阅读作者的另一本书《如何启动黄金圈思维》，该书中文简体字版已由湛庐文化策划、浙江人民出版社于2019年出版。——编者注
 - (5) 华特·迪士尼的全名为华特·埃利亚斯·迪士尼（Walter Elias Disney）。——译者注
 - (6) 奥兹国是童话《绿野仙踪》中的魔幻世界。——编者注
 - (7) Z世代主要指95后和00后。——译者注

CHEERS
湛庐

第四届“国家图书馆文津图书奖”
获奖作品升级版

牛奶可乐 经济学

[美]
罗伯特·弗兰克 著
Robert H. Frank

王堃 俸绪娴 康溪馨 译

为什么父母不能
自由选择孩子的入学时间？
为什么法律不能禁止
你的邻居晚上弹钢琴？
航班超额售票时
谁该多等一会儿？

Liberty, Competition, and the Common Good
4

带你走入
新奇有趣的
经济学课堂

为什么鹿角
不是进化得越大越好？
为什么实行员工自治的
企业没有繁荣起来？
为什么穿更贵的
西服去面试更有可能成功？

THE DARWIN ECONOMY

浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

版权信息

本书纸版由浙江人民出版社于2020年3月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

书名：牛奶可乐经济学4（合集）

著者：罗伯特·弗兰克

电子书定价：62.99元

THE DARWIN ECONOMY

Copyright © 2011, Robert H. Frank

All rights reserved.

序言

达尔文带来的经济学启示

作为经济学专业领域的学科，行为经济学在近几十年来的发展已呈脱缰之势。目前这个领域的学者主要致力于经济学和心理学的交叉研究，他们的大部分注意力集中于人们的判断和决策中存在的系统性认知偏差。正如已故的斯坦福大学心理学教授阿莫斯·特沃斯基（Amos Tversky）曾经常说的：“我的同行们都在研究人工智能，而我呢，却在研究人们与生俱来的愚蠢。”

在20世纪80年代初，我教授了一门早期的行为经济学本科生课程。由于鲜有学生听说过这一新兴的学科，我面临的首要挑战便是想办法起一个能吸引学生注册的课程名称。最终，我决定称这门课为“对理性选择的背离”。那时候当然还没有什么标准的课程大纲。在经过一番思考之后，我决定将授课内容确定为两大主题：带有悔意地背离理性选择和毫无悔意地背离理性选择。

在第一个主题之下，我列举了关于人们易出现的许多系统性认知偏差的研究。例如，尽管标准理性选择模型认为，人们会忽略沉没成本，也就是在进行决策时已然无法收回的成本，然而这些成本却常常会对决策造成显著的影响。请设想，在你就要动身前往80公里外的体育场观看体育比赛或演唱会时，一场暴风雪意外降临。如果你买的票是不能退款的，那么票价理应不会影响你决定是否仍要开车前往。然而，相比于碰巧得到免费票的人，花费100美元购票的球迷或歌迷显然更有可能冒险开车过去，这些人很可能会因为自己的认知偏差而感到后悔。对基于这些偏差所做的决策，人们通常在意识到以后就会感到后悔。

在第二个主题下，我列举的是关于人们偏离标准理性选择模型的预

期但并不感到后悔的行为的研究。这里，用于佐证的案例是人们如何在最后通牒博弈中对单方提议做出反应。在最后通牒博弈中，实验人员给一名被试（被称为提议者）一些钱，比如100美元，然后让他提出他与另外一名被试（被称为响应者）如何分配这笔钱。如果响应者接受，两人就会分别得到相应的金额。例如，如果提议者提出“我60你40”，而响应者接受了，那么提议者得到60美元，响应者得到40美元。但是，如果响应者拒绝提议者的分配方案，那么这100美元就要归还给实验人员，两名被试什么都得不到。

标准理性选择模型认为，提议者会单方面提出自己得到99美元，而给响应者1美元的分配方案，因为他知道相比于什么都得不到，响应者出于对自己利益的考虑会接受这一方案。但是这样的提议很少出现，即使出现了，多数也都会被响应者拒绝。而且，拒绝单方提议的被试很少会因为这样做而感到后悔。

起初，行为经济学的研究大都集中于对理性选择带有悔意的背离，也就是由认知偏差导致的背离。曾在康奈尔大学与我共事的理查德·塞勒（Richard Thaler）曾与卡斯·桑斯坦（Cass Sunstein）合著了《助推》（Nudge）一书。这是一本2008年出版的畅销书，它总结了多种导致人们误入歧途的认知偏差，并指出政策制定者应如何重构环境以促成更好的决策。我强烈支持这本书中的几乎所有提议。

然而，我从一开始就认为，对理性选择没有悔意的背离会造成更大的损失。这是因为，一旦意识到认知偏差，人们通常有意愿并且有能力单方面进行纠正。相反，人们通常没有办法也没有动力纠正自己并不感到后悔的行为，即使这些行为会耗费大量的社会成本。

想想看，理性选择模型的标准假设是，决定任何物品满意度的首要因素是它的绝对品质。面试装显然不符合这一假设。如果你与其他几名资质差不多的申请者都想得到一份投资银行的工作，那么在面试时外表看起来不错对你来说将非常有帮助。但是“看起来不错”本来就是相对

的，意味着相比于其他申请者看起来更好。如果他们在面试时穿的是价值500美元的正装，你穿的是2 000美元的而不是200美元的正装，那么你将更有可能博得有利的第一印象，并且更有可能赢得复试的机会。

在一些领域中，达成一些重要目标所需要的能力显然取决于相对付出，这时，亚当·斯密的“看不见的手”就失效了。尽管斯密的许多现代的忠实拥趸不加批判地积极鼓吹一些见解，但不受约束的市场常常不能成功地引导自利个体的行为，让其实现集体利益。恰恰相反，正如先驱查尔斯·达尔文清楚看到的，对个体行为的激励常常会导致浪费性的军备竞赛。

例如，达尔文明白了雌孔雀之所以会被雄孔雀显眼的开屏所吸引，或许是因为这样的展示是一种可靠的信号，体现了一种稳健的、可以遗传给后代的免疫系统，因为体内有寄生虫的雄孔雀会由于代谢问题而无法展示出修长又靓丽的尾羽。但达尔文同样意识到，炫耀自己尾巴的雄孔雀更易被天敌捕食，这样的展示对于整个物种来说是一种浪费。如果所有雄孔雀都只开一半屏，它们依然能吸引到原先的伴侣，同时也更不易被捕食。然而对于展示尾巴这件事，作为个体的雄孔雀没有理由后悔，因为如果不尽力展示，它获得配偶的概率就会大大降低。与此类似的是，如果所有求职者都花2 000美元买正装，相比于大家都买300美元正装的情况而言，一个人获得工作的机会并没有增加多少。但是这些求职者同样不会后悔购买更贵的正装。

这些都是集体行动中存在的问题，特别类似于军备竞赛的情形，与认知偏差无关。在正装上花费过多的原因与军费过度支出并无不同。在这些情况下，没有任何一个个人或国家可以通过减少支出而获益。

与此相反的是，如果个体是出于自己的认知偏差而蒙受损失，那么他们既有办法也有动力减少损失。他们可以通过一些方式搜寻更多的信息，例如聘请并咨询专业人士。他们也可以通过签订合约的方式来限制自己的行为，以免发生这些偏差。

集体行动的问题不仅很难通过个体解决，而且造成的损失远比认知偏差造成的要大。但是正如我接下来将要解释的，好消息是，许多由集体行动的问题导致的重大损失，都可以通过简单且不那么激进的税收政策来消除。在我解释的过程中，我将努力论证我的推测：100年后的经济学家更可能将达尔文视作他们理论思想的奠基人，而不是亚当·斯密。

这么多年来，我读过或在现场听过许多作者的采访。令我感到惊讶的是，当被问及对青年作者的建议时，他们经常会回答“写写你知道的就行”。在一个早晨，我坐下来准备开始写作本书第9章（成功与运气）的内容，我心想：对于“成功与运气”我确实了解那么一些！

实际上，自从开始这本书的写作，我所经历的所有情绪中，最强烈的便是怀疑自己究竟能不能写成它。当然了，生活中充满了偶然事件，但是我真的太过幸运。

其中最突出的一段经历是这样的：我多年的好朋友兼同事汤姆·基洛维奇（Tom Gilovich）每个星期六早晨都会抽空跟我在纽约州伊萨卡（Ithaca）附近的室内网球场打一个半小时的球。大概在2006年的一个早上，在我们打第二盘前换场的时候，我突然感到一阵恶心。据说，我当时晕倒在球场上，失去了意识，也没有脉搏。

几天过后，我还躺在医院的病床上，主治医生告诉我，我所经历的是心搏骤停，他解释说，这几乎是致命的，而且绝大多数挺过来的人都会因此落下严重的残疾。

汤姆后来向我详细描述当时的情景。我昏倒后，他立刻喊了人帮忙，并叫了救护车。尽管他之前并没有接受过心肺复苏训练，但他曾经在电影和电视节目里看到过。他迅速将我翻身，开始使劲按压我的胸膛，虽然没有效果，但是他一直坚持这样做，他说，直到过了很久之后，我终于轻微地咳嗽了一下。

那个网球场位置偏僻，距离城镇有几公里远，不过附近碰巧在几分钟前刚发生过车祸。由于调度失误，有两辆救护车被派去了车祸现场，而不是只有一辆。就在第二辆救护车快要赶到的时候，司机接到指令，改变方向，开往网球场，在我晕倒后不久便到了，医护人员和担架都是现成的。

尽管我有几分钟缺氧，但重要的是，我在4天后突破万难出院了，而且没有留下任何后遗症。两星期后，我回到球场跟汤姆打球。恢复运动对我来说并不可怕，因为几天前我成功地通过了压力测试，再说，我也没有晕倒在球场上的记忆。但是我知道，这件事对汤姆来说并不轻松。

汤姆的妻子卡伦·基洛维奇（Karen Gilovich）是一名救生员，这与之前的巧合相比甚至不那么有戏剧性了。怀着深深的感激和欣喜，我将此书献给他们。

这本书的问世也要归功于我能够幸运地在康奈尔大学谋得职位。在我自1972年开始执教后不久，我得知自己正是当年被我所在的院系聘用的所有7名教授中的第7名。这个系其他几年的聘用人数甚至不超过4人。一位同事后来告诉我，当他支持聘用我为第7名教授时，脾气火爆的系主任因为更青睐另一个候选人而恼怒不已，向他丢了一支粉笔。那时候，我仅有的另一个工作机会来自中西部地区的一所不那么出名的大学，那所学校我随时可以去。

令我感到幸运的是，我不仅能够获得康奈尔大学的工作机会，而且还保住了这个职位。在我工作的第4年，我唯一发表过的文章还是由我与研究生院的同学合著的。就在那一年，后来成了美国联邦储备委员会理事的经济学家内德·格拉姆利克（Ned Gramlich）辞去了他在布鲁金斯学会的职务，来康奈尔大学经济学系访问了两个学期。我们很快成了好朋友，尽管我没有什麼拿得出手的成果，他却似乎认为我有潜力。当他问我可否为他正在编辑的一本文集写一篇文章时，我很痛快地答应了。

我迫切地想让他满意，就非常努力地写了那篇文章。

就在文章快要完成的时候，内德垂头丧气地向我道歉，说出版社取消了文集的出版。失望之下，我把文章投了出去，结果6个星期以后，《经济计量学》（*Econometrica*）接收了这篇文章。如今，该期刊已经成了经济学领域的重要期刊之一。对于经济学家来说，与有文章被收入一本文集相比，在这样的期刊上发表一篇文章在职业发展上具有完全不可比拟的优势。

我在1977年，也就是工作的第5年里的成果就多得多了，投了5篇文章，这在量上原本也于事无补，但我投出的每一篇都很快被《美国经济评论》（*American Economic Review*）、《政治经济学杂志》（*Journal of Political Economy*）等前沿经济学期刊接收，还免去了修订。自此之后的几十年里，类似的情况再也没有在我其他文章上发生过。我纯粹只是幸运罢了。

身为一名学者，有机会与优秀的学生和出色的同事共事是难得的幸事。能够在像康奈尔这样的大学度过我的职业生涯，极大地影响了我在专业领域所能学到的和做到的事。在其他领域，我当然也可以快乐地生活，却不一定能写出这本书。

我还想要感谢我的妻子艾伦·麦克莱斯特（Ellen McCollister），感谢她在我进入“写作模式”时表现出极大的耐心和支持。目前为止，她对此已经经历甚多了，但从未表现出厌倦。许多经济学家整天都在证明数学原理，而我喜欢写与人们的实际经历有关的事情，因为这给了我跟艾伦讨论并学习她的丰富洞见的机会。

还有很多人对我的写作提供了莫大的帮助。我要特别感谢：布鲁斯·布坎南（Bruce Buchanan）、加里·伯克（Gary Burke）、菲利普·库克（Philip Cook）、泰勒·考恩（Tyler Cowen）、李·芬内尔（Lee Fennell）、特德·菲舍尔（Ted Fischer）、克里斯·弗兰克（Chris Frank）、赫伯特·甘斯（Herbert Gans）、史丽娜盖什·盖文内尼

(Srinagesh Gaverneni)、汤姆·基洛维奇、马克·格罗杰 (Marc Groeger)、玛丽亚·瓜达卢佩 (Maria Guadalupe)、亨利·汉斯曼 (Henry Hansmann)、奥里·费茨 (Ori Heffetz)、莫里兹·霍伊默 (Moritz Heumer)、鲍勃·霍基特 (Bob Hockett)、格雷厄姆·科斯利克 (Graham Kerslick)、马克·克莱曼 (Mark Kleiman)、吉姆·勒基特 (Jim Luckett)、戴维·莱恩斯 (David Lyons)、迈克尔·马丁 (Michael F. Martin)、雷克斯·米克森 (Rex Mixon)、森德希尔·穆来纳森 (Sendhil Mullainathan)、汤姆·内格尔 (Tom Nagel)、马修·纳格勒 (Matthew Nagler)、迈克尔·奥黑尔 (Michael O'Hare)、萨姆·比齐加迪 (Sam Pizzigati)、凯特·鲁宾斯坦 (Kate Rubenstein)、蒂姆·斯坎伦 (Tim Scanlon)、汤姆·谢林 (Tom Schelling)、埃里克·舍恩伯格 (Eric Schoenberg)、菲利普·西曼 (Philip Seeman)、拉里·塞德曼 (Larry Seidman)、彼得·辛格 (Peter Singer)、杰夫·索默 (Jeff Sommer)、蒂蒙·斯比勒提尼 (Timon Spiluttini)、唐凯 (Kai Tang)、史蒂夫·特莱斯 (Steve Teles)、菲德尔·特沃尔德 (Fidel Tewolde)、迈克尔·瓦尔德曼 (Michael Waldman)、戴维·斯隆·威尔逊 (David Sloan Wilson)、萨斯基亚·维特拉克 (Saskia Wittlake) 和安德鲁·怀利 (Andrew Wylie)，谢谢他们富有洞见的建议。当然，他们不必对本书的任何错误负责。

最后，我非常感激来自普林斯顿大学出版社的彼得·多尔蒂 (Peter Dougherty) 和塞思·迪奇克 (Seth Ditchik)，感谢他们在本书创作计划的早期表现出的热情和他们提供的睿智的建议，这些都有助于本书的成形。我最初选择的书名是“自由主义者的福利状况”。要说这本书能吸引到读者，我要感谢彼得和塞思 (迈克尔·马丁也推了一把) 劝我放弃了这个题目。

我原本想用的另一个书名是“达尔文之楔”，这最终成了本书第2章的标题。我喜欢它所表达出的个体利益和群体利益之间的分歧，这是我命题的主要基础，达尔文也清楚地知道这种分歧的重要性。我同时还

想，正因为它不为人所熟知，所以才有一种朦胧的吸引力吧。某一天晚餐时，我征求几个朋友的意见。在朋友们发话之前，我的妻子说她首先想到的是“达尔文的高跟鞋”，引起了哄堂大笑。第二天早晨，我就发信息给彼得·多尔蒂，提议使用他最初建议的英文书名——The Darwin Economy。



扫描二维码，下载“湛庐阅读”App，
搜索“牛奶可乐经济学4”，获取本书参考文献！

目录

序言 达尔文带来的经济学启示

01 迷恋往昔：现代经济为什么会增速放缓、遭遇危机

对片面的解释保持警惕

市场也会失灵

没有约束的自由不是自由

棘手的间接伤害

只有抛弃错误认知才能解决问题

02 达尔文之楔：竞争无法弥合个体与群体的利益分歧

“看不见的手”是一个非凡洞见

竞争的缺陷

成为更好的，而不是最好的

幸福感大部分取决于相对收入水平

传统经济学没有考虑情境因素

03 尚未被利用的机会：市场失灵传统解释的缺陷

为什么实行员工自治的企业没有繁荣起来

如何验证经济理论的正确性

对“看不见的手”的质疑

更高的收入意味着更高的消费能力

04 干预是为了更好的竞争：你无法通过不吃饭饿死身体里的寄生虫

如何避免公共决策中的浪费

事与愿违的支出削减

我们需要什么样的公共服务与产品

个人造成的浪费远超群体

[收入差距与支出瀑布](#)
[如何解决资金短缺的问题](#)

[05 控制“地位消费”：得不偿失的军备竞赛](#)

[无人获益的军备竞赛](#)
[经济领域的“军备竞赛”](#)
[人们的价值判断会受到情境的影响](#)
[如何让管制发挥有利的作用](#)
[累进消费税：限制个人浪费的重要途径](#)
[我们为什么必须纳税](#)

[06 深度理解竞争：完全竞争中的加害者与受害者](#)

[科斯的成功之处](#)
[协商路上的绊脚石](#)
[成本-收益分析是不道德的吗](#)
[跨种族婚姻中的夫妻是否应被禁止在公共场合牵手](#)
[我们应该捍卫何种权利](#)

[07 效率至上：金钱在调节利益冲突者之间关系中的作用](#)

[收入差异对成本-收益分析的影响](#)
[电台节目应该是纯谈话还是纯音乐](#)
[稀有的钟应当归谁所有](#)
[航班超额售票时谁该多等一会儿](#)
[自由主义者对成本-收益分析的敌视](#)
[为何收入转移往往是实行高效率政策所必需的](#)
[通过政策进行收入转移有助于推行高效率的政策](#)
[采用高效率政策面临的阻碍](#)

[08 “这就是你们的钱”：人们的收入并不只是付出努力的结果](#)

[所有权的神话](#)
[要用税收体系为社会工程买单](#)

[高效率地供应公共产品](#)
[累进税率](#)
[劳动力市场中高排名的隐性市场](#)
[高排名值多少钱](#)
[工作场所之外的社会比较](#)
[高社会排名给他人施加的成本](#)
[机会在召唤](#)

[09 成功与运气：成功对于运气的依赖比多数人想象的要大](#)
[成功与美德](#)
[比尔·盖茨是个幸运的人吗](#)
[与运动有关的幸运](#)
[具备勤勉天赋和能力本身已是一种极佳的运气](#)
[赢家通吃市场是如何将小的差异放大的](#)
[信念至关重要](#)

[10 伟大的权衡：为何降低高收入者的税率会抑制经济增长](#)
[所得税与涓滴理论](#)
[给富人减税并不会带来更多工作机会](#)
[为何降低高收入者的税率通常会抑制经济增长](#)
[公地悲剧](#)
[愚蠢的口号竞争](#)

[11 对自由竞争的规制：对有害行为征税](#)
[庇古税](#)
[消除污染并不是最优的结果](#)
[为何征税手段能够实现最低的减排成本](#)
[征税的目标](#)
[气候变化与二氧化碳税](#)
[拥堵费](#)
[按照汽车重量征税](#)

[烟草税](#)

[酒精税](#)

[对造成非直接伤害的行为征税](#)

[自行车与摩托车头盔法规](#)

[对有害行为征税是一种滑坡谬误吗](#)

[12 重新思考自由主义者的反对主张](#)

[科斯带来的启示](#)

[合租室友抽烟，你还要和他一起住吗](#)

[解决问题的有效性是最大自主权的前提](#)

[妥协的真正原因](#)

[约翰·高尔特会同意强制储蓄计划吗](#)

[安全法则](#)

[什么是权利](#)

[我们需要新一代的自由主义者](#)

[后记](#)

[译者后记](#)

[注释与参考文献](#)

THE DARWIN ECONOMY

THE DARWIN ECONOMY

01

迷恋往昔：
现代经济为什么会增速放缓、遭遇危机

人们时常过分迷恋过往。然而有时候，从生活中的一些重要方面来说，现状确实不胜往昔。比方说，在第二次世界大战结束后的30年里，美国人的收入都以每年近乎3%的速率迅速增长。那时，中产阶级充满经济活力，道路和桥梁都维护良好，每年还会大肆兴建新的基础设施。约翰·肯尼迪总统强烈号召：“不要问国家能为你们做什么，要问你们能为国家做什么。”这令美国人欢欣鼓舞。那时，人们坚信，美国能够在登月竞赛中拔得头筹。那时的人们是乐观的。

然而好景不长。在政府干预的几十年里，美国经济增速极其缓慢，只有那些收入处于顶端的人才赚得盆钵满盈。美国大型公司高管的薪水在此期间增长了10倍，而剔除通货膨胀的影响后，员工工资实际上却下降了，中产阶级大多负债累累。

在当时的美国，建设高速铁路系统、智能电网等紧缺基础设施的提案被一再搁置，已有的基础设施也因失修而日益破败。穷人和富人同样受颠簸的公路和危险的桥梁的困扰。供水和排水系统也频频出现故障。数不胜数的学校被废弃，许多美国人生活在因缺少维护而随时可能倒塌的水坝下。苏联削减了用于隔离未加妥善处理的核原料的开支。

更令人不安的是，即便是在这些紧迫的问题面前，美国政府似乎不能有所作为，几乎陷于彻底瘫痪的境地。之所以会如此，通常是因为他们刻意忽视了支配人类行为的基本事实和逻辑。

对此，一个有力的佐证便是美国没能解决由2008年的经济危机引发的失业“顽疾”。正如约翰·梅纳德·凯恩斯（John Maynard Keynes）在大萧条时期做出的解释：深陷泥潭的经济鲜能自行恢复。¹他指出，消费者是不会带头刺激经济的，因为他们自己都还承受着债务的压力和失业的恐惧（如果还没失业的话）。投资者也不会带头刺激经济，因为大多

数企业的生产已经过剩。最后，凯恩斯推断，政府才是唯一既有能力又有动力刺激支出，从而提振就业的角色。

人们普遍失业的情形，就像飞机每天起飞时座位还都是空的，没有任何价值。毫无疑问，必须采取一切办法避免这种浪费。然而经济刺激计划的批评者立即谴责政府支出本身就是一种浪费，他们无视了那么多迫切需要关注的好项目。例如，根据内华达州交通局的数据，如果立即修复80号州际公路上一段16公里的破损路面，将会花费600万美元；但如果推迟至两年以后才进行修复，那时气候和交通使用会对路面造成更加严重的损害，进行同样修复的花费将达到3 000万美元。²

在经济严重衰退的时期，人员和设备都被空置。同时，伴随着全球市场的萧条，生产所需原料的价格异常低迷，为这些项目融资的利率也接近历史最低水平。这些本都是亟待解决的问题，而且完全不需要额外的经济刺激。然而，无知极大地限制了当时的政治对话，政府因此无法有所作为。

反对经济刺激的人以担忧财政赤字为由支持政府不作为，但赤字是长期问题。没有人觉得基础设施的维护可以被永久拖延下去，这件事情越早做就越节省，长期来看反而会减少赤字。赤字当然是必须应对的问题，但要在经济完全恢复之后再应对。

同样是带头打着担心财政赤字过高的旗号而反对额外经济刺激的这些人，却投票支持削减美国国家税务局的预算。可信的证据表明，美国国家税务局的预算每减少1美元，就会减少10美元税收，导致赤字净增加9美元！这样的政策竟然会被众议院批准，意味着美国政府正渐渐成为一个无知的政府。在这样一个国家里，由无知导致的政治瘫痪，使人们应对哪怕是最为紧迫的问题都会受到阻碍。

同样是这些领导者，他们通过投票削减了对养育低龄儿童的低收入妇女的营养支持，削减数额超过了10亿美元，而且还减少了0.7亿美元的美国国家净水循环基金。这些计划原本不仅能够帮助有需要的人，同

时也能杜绝日后代价高昂的问题。削减这些支出，只会增加未来的赤字，而不是减少赤字。

小布什总统针对最富有的美国人的减税政策计划于2010年取消，这些人在反对取消减税政策时竟也忘记了他们对财政赤字的担忧。许多富裕的人已经拥有了一辈子无法享尽的财富，延长减税政策的期限几乎无法产生刺激经济的效果。因取消减税而增加的政府收入，可以用来支撑更加有效的刺激措施，若将其补贴给各州，可能会帮助缓解大量教师、警察和消防人员的失业问题。然而，参议院少数党领袖米奇·麦康奈尔（Mitch McConnell）在美国有线电视新闻网的一档访谈中不无嘲讽地表示：“在经济衰退期间增税可不是个好主意。”³

还有一个不那么紧迫但可能无比重要的问题：一直以来，针对全球变暖的消极科学预测不断增加，而美国政府对此的态度却是漠视，甚至是敌视。质疑气候变化的人通常合理化他们不作为的理由，那就是这些求变的呼声缺乏严谨的科学理论的支撑。但是，最杰出的科学家们也很快承认，他们的预测存在固有的不确定性。气温上升的幅度比预期要小当然是有可能的，但同样也可能更大，如果是后一种情况的话，极有可能是一场灾难。根据给出的气温可能增加的范围及其增加的概率，“不确定性”就是政府理应采取行动的最有力依据。

例如，2009年由美国麻省理工学院开发的知名的综合全球系统模型（Integrated Global Systems Model）做出预测，全球平均地表温度有10%的可能会在2095年上升6.67摄氏度以上。⁴这种幅度的温度增加最终将意味着地球上生命的终结。简单的政策措施就可轻松化解这一威胁，如对二氧化碳排放课以重税。如果分阶段逐步施行这一措施，人们就可以慢慢适应，而不用承受痛苦的牺牲。

任何理性的政策举措都可以迅速解决这一问题，但是掌控能源政策的议会领袖们甚至固执地否认问题的存在。富有经验的国会观察员表示，有意义的气候法案几乎不可能立即在美国参议院获得通过。在一个

无知政权的领导之下，这样的立法在政治上显然是绝无可能的。

对片面的解释保持警惕

我们有必要对单一的解释保持警惕，然而我们忽视了这样一个事实的重要性：在最近30年中，有一种强大的文化共识在美国民众的头脑中已逐渐根深蒂固，那就是，“政府才是万恶之源”。美国自由主义者向来对政府滥用权力保持警觉，他们正是以上共识的核心宣扬者之一。尽管自由党历来人数较少，他们对公共话语却有着极大的影响，并且这一影响还在与日俱增。

自由党的影响力很大程度上源于他们为了散布“政府才是问题所在”这一信息所抛掷的巨额资金。比如，在《纽约客》杂志上刊登的一篇被广泛引用的文章里，简·迈尔（Jane Mayer）发现，身为自由党的亿万富翁、科氏工业集团掌门人查尔斯·科赫（Charles Koch）和大卫·科赫（David Koch）向极右翼智库捐赠超过一亿美元，这些智库成员是茶党（Tea Party）和其他一些团体的组织者，这些团体的使命便是宣扬“政府才是问题所在”的信息。⁵

鲁珀特·默多克（Rupert Murdoch）的福克斯新闻台尽管宣称公正无私，其实也在不遗余力地宣传同样的信息。在此之前，该频道接受了支持右翼智库的亿万富翁理查德·梅隆·斯凯夫（Richard Mellon Scaife）的巨额赞助。斯凯夫是《匹兹堡先驱评论报》（*Pittsburgh Tribune-Review*）的所有者，而且是梅隆家族财富的继承人。更早些时候，约翰·奥林基金会（John M. Olin Foundation）把将近4 000万美元分发给保守派智库、媒体，以及高校中的经济和法律研究项目，用于散播“政府才是问题所在，自由市场才是解决之道”这一观念。

总体看来，这些投入收效显著，助长了刚开始兴起但无处不在的愤怒，使政府几乎不能有所作为。当然，自由党说对了一点，那就是令人

困扰的政府浪费由来已久。诚然，人们感激自由党对于个人自由被侵害和公共资金被滥用的警惕。但是，仅仅因为政府有缺陷就意味着多数美国公民想要彻底废止一切公共政策吗？毕竟，市场也并不完美，何况有许多重要任务只有政府才适合执行。国防便是一个鲜明的例子，除此之外还有公共基础设施的建设和维护，明确和落实产权亦属政府职责。

每一个运作良好的社会都有政府在经济和社会生活中发挥重要作用。在那些公民最为拥戴政府的国家中，公共物品和公共服务往往是最好的，腐败水平是最低的，人均收入也是最高的。相反，在那些政府最为弱势的国家，如索马里、海地或者苏丹，市场通常运作不良，人均收入极低，犯罪和暴力事件频发，公民们也认为自己的政府腐败无能。如果必须选择，多数美国人宁可住在新西兰也不愿住在海地。政府水平和影响力并非造成以上局面的唯一原因，却是重要的原因。

许多活动必须借助集体之力进行，这便意味着政府必须通过征税来为这些活动付费。自由党和其他反政府激进分子常常谴责强制征税是偷窃行为，但是没有一个政府可以仅依靠志愿捐赠就能正常运作。没有强制税收，就不会有政府。没有政府，就没有军队，没有军队的国家最终会被某些有军队的国家入侵。待尘埃落定之时，你就要被强制向那些国家的政府缴税。

如果我们只能生存在强制征税的政府之下，那么最好的选择就是创立一个能够运用我们的资金产生最大价值的政府。我们必须认真思考该如何设置和监督政府机构，深入探讨什么样的公共服务是自己想要的，以及如何为这些服务付费。然而目前为止，这些事情一件也没有进行。

一个资源丰富的国家，有着教育水平和生产力较高的劳动人口，在这样的国家中，一切不该如此。值得庆幸的是，要打破当前的僵局其实并非难事。这是因为造成当前这种局面的原因，并不是不可调和的价值观念的差异，而仅仅是对竞争机制的误解。

市场也会失灵

毫无疑问，亚当·斯密的“看不见的手”是名副其实的开创性洞见。生产者迫切地改良产品设计并进行降低成本的创新活动，仅仅是为了从竞争对手那里夺取市场份额和利润。短期看来，他们能够如愿以偿。但是竞争企业会立即效仿创新行为，由此导致的竞争将迅速使价格降至更低的水平。斯密认为，消费者会成为所有这些变动的最终受益人。

斯密的许多当代追随者认为他提出了更为鲜明的主张，那就是市场总会借由个体的自利实现整个社会的利益最大化。然而，斯密自己的阐述要谨慎得多。他写道：“追求利润的企业主只是谋取自己的利益，但无论怎样，他都会被一只‘看不见的手’引导着促成原本不在他意图之内的结果。对社会来说，不参与并不总是坏事。”⁶

斯密从未确信由“看不见的手”促成的结果在所有情形下都是好的。他的怀疑是显而易见的，例如，他写道：“哪怕只是为了娱乐消遣，从事相同行业的人也很少聚在一起，他们交谈的结果往往是针对公众的共谋，或者抬高价格的办法。”⁷对斯密来说，神奇的是这些自利行为通常会导致对社会有益的结果。⁸

如斯密一样，现代的进步批评家倾向于将市场失灵归咎于抑制市场竞争的阴谋。但是在斯密所处的时代，市场竞争要比现在更容易受到抑制。“看不见的手”面临的最大挑战在于竞争过程本身的逻辑。

达尔文是最早清楚地发觉问题症结所在的人之一。他的核心洞见之一是，自然选择对某些行为和特征的偏好主要依赖于这些行为和特征对生物个体的影响，而不是对群体的影响。⁹他承认，个体和群体的利益在有些时候是一致的，在这种情况下，我们能得到类似“看不见的手”造成的结果。举例来说，一只鹰的基因突变带来的更敏锐的视力可以增加这只鹰自身的利益，同时，这一突变在鹰群中的扩散使鹰这一物种更加成功。

然而，在另外一些情形下，基因突变虽然对个体有所助益，但对群体来说却是极为不利的。在内部斗争激烈的种群中，基因突变的这种影响是必然的。雄性动物体形大小的演变便是佐证。多数脊椎动物是一夫多妻制的，意味着一个雄性个体可以与多个雌性个体交配。对这些雄性个体来说，关键是获得交配资格，因为一些雄性个体获得多个配偶就意味着另一些雄性个体得不到配偶，无法将自己的基因传递下去，用达尔文的话来说，它们会成为最终的失败者。因此，雄性会通过激烈的争斗来抢夺配偶也就不奇怪了。在这些争斗中，体形大小是很重要的，因此便有了进化的“军备竞争”，从而产生块头越来越大的雄性个体。

象海豹的演化是个极端的例子，但是颇具启发性。¹⁰雄性象海豹的重量通常有将近2.7吨，是雌性象海豹重量的5倍多，相当于一辆林肯领航员越野车的重量。在交配季节，一对对成年雄性象海豹进行持续1小时的激烈斗争，直到一方落败，伤痕累累、精疲力竭地爬走。获胜者会拥有对将近100只雌性象海豹近乎排他的交配权。但是，尽管块头比对手更大使雄性象海豹个体更容易在争夺配偶的斗争中获胜，身躯过于庞大却会明显阻碍雄性象海豹群体的生存，因为这使它们更易受到鲨鱼和其他捕食者的攻击。

如果雄性象海豹有机会通过投票决定每个个体的体重是否减半，它们有充足的理由投赞成票。因为，在斗争中起到关键作用的是相对大小，而不是绝对大小，所以，对于特定的一对一的较量而言，这一改变并不会影响最终的结果，却会降低每个雄性象海豹个体被鲨鱼捕食的风险。然而，并没有切实可行的方法能让雄性象海豹进行这样的投票，每个雄性象海豹个体也不会单方面让体形变小，毕竟，体重比同类差太多的个体是永远无法与雌性进行交配的。

人类互相争夺奖励时也普遍存在类似的冲突。托马斯·谢林（Thomas Schelling）举了一个著名的例子¹¹，借以说明这些冲突的实质。谢林发现，冰球运动员可以自由选择是否在冰场上是否要戴头盔，但他

们总是不戴头盔，而如果可以对此进行投票选择，他们反而会支持制定规则要求他们戴头盔。这使谢林感到疑惑：如果头盔这么重要，他们直接戴上不就得了？为什么还需要制定规则呢？

谢林为了解决他的疑惑而进行了观察。他发现不戴头盔能带来一些竞争优势，可能是不戴头盔可以使运动员看或听得更清楚，也可能是不戴头盔使他们可以恐吓对手。获得竞争优势的吸引力胜过了不那么实际的利益，也就是降低受伤的概率，因此运动员们欣然接受了随之而来的额外风险。问题在于，当每个运动员都不戴头盔时，并没有人取得相对优势，所以才需要外在的规则。

正如谢林所阐明的，冰球运动员遇到的问题与信息不完全、缺乏自控力或者认知能力不足无关，这些缺陷往往成为政府干涉的依据。¹²显然，问题也不是由剥削或其他不完全竞争造成的。相反，这正是再平常不过的集体行动的困境。运动员们支持戴头盔的规定，因为这是唯一可以在比赛中确保他们的安全的途径。像在更衣室放一块牌子，提示大家戴头盔可以降低重伤风险这种简单的“助推”是无法解决问题的。他们需要的是强制性的规定。

自由主义者对此表示不满，认为这一规定剥夺了个体自由选择的权利。这又该如何应对呢？这种反对就类似于反对军备控制协议，因为这剥夺了协议签署方选择军备投入的权利。当然是这样，但是这就是协议的关键所在啊！处于集体行动的困境中的各方，往往会意识到：只有约束他们各自的行动，才能实现想要的结果。

约翰·斯图尔特·穆勒（John Stuart Mill）曾在他所著《论自由》（*On Liberty*）中提出，只有在没有其他干涉性更低的方法来避免伤害他人的情况下，才可以限制个体的行动自由¹³。戴头盔的规则似乎是满足这一条件的。不戴头盔的运动员会给对手造成威胁，降低对方赢得比赛的可能性，这对他们来说是至关重要的。如果戴头盔的规则反而带来更大的伤害，就不会有人支持了。不过这是一个简单的实际问题，并不

涉及深刻的哲学原则。

回报取决于相对优势的情况会催生集体行动的困境，从而导致市场失灵。例如，在达尔文学说中的军备竞赛里，造成个体与集体利益冲突的隔阂，同样可能解释为什么“看不见的手”在工作场所中无法自行实现最大程度的安全。传统的“看不见的手”理论的出发点基于这样的观察：其他条件保持不变的情况下，更危险的工作会带来更高的收入，一是因为雇主节省了安装额外的安全设施的花费，因而能够支付更高的工资；二是因为人们一般偏好安全，倾向于选择更安全的工作，除非做更危险的工作确实能赚得更多。根据“看不见的手”的一般原理，员工之所以会为了更高的工资收入而选择接受更低的安全保障，是因为额外的收入足以补偿安全性的降低。但是这一论述的前提是，对额外收入的价值衡量仅仅基于其所增加的绝对消费量。然而，如果一个员工可以获得更高的收入，他还可以获得其他收益。他当然可以增加绝对消费量，但与其他人相比，他的相对消费量也更多。

例如，多数家长都想把孩子送进最好的学校。一些人可能为了支付更好的学区房的月供，接受薪水更高但更危险的工作。然而，其他员工也是这么想的，学校质量是一个内在的相对概念。因此，如果其他员工也通过牺牲安全来换取更高的工资，最后的结果只会是哄抬了好学校附近的房价。最终，每个人的工作安全性都会降低，然而没有人可以实现自己最初甘愿牺牲安全所想要达到的目的。同样，在军备竞赛中，当各方都大肆置办武器时，没有任何一方会比以往更加安全。

在这种情况下，员工可能更希望世界是一种人人都更加安全的状态，即便这意味着工资水平更低。但是，每个员工只能掌控自己的工作选择，而无法左右他人。如果一个员工选择了一份更加安全的工作，而其他人员并没有这样选，那么这个员工将被迫把孩子送进差一点点的学校。如果想得到他们所希望的结果，这些员工就必须联合行动。此时，轻微的助推仍旧不会有效果。仅仅知道自己可以选择或不采取某种行为，也不

能消除促使个体这样做的动机。

没有约束的自由不是自由

我最初了解穆勒的思想是在高中时代，他在文章中写道，限制个体自由的唯一合理性在于避免对他人的伤害，对此我到现在都十分认同。尽管穆勒不是自由主义者，但自由主义者却对他表示赞同，常常引用他的这一观点。¹⁴然而，那些对自由的合理约束，要远比自由主义者和其他反政府激进分子预想的更多。

针对美国经济重回正轨所需要采取的政策，最激烈的反对意见来自自由主义者和其他右派人士，因此我们应该仔细研究这些人的论点。不同于多数的左派批评人士，我在一定程度上同意自由主义者对世界的最重要的基本假设：市场是竞争性的，人是理性的，政府在约束每个公民行动的自由之前必须充分证实约束的合理性。尽管每个假设都有待商榷，但即便人们全盘接受这些假设，它们之间存在的内在矛盾也是显而易见的。

这个理论框架的致命缺陷在于一个本身毫无争议的论述，那就是在生活的许多重要领域中，表现的优劣取决于相对标准。比如说，一名职业网球运动员的收入水平，并不取决于他在绝对意义上的表现，而是取决于他在职业比赛中相对于其他选手的表现。回报取决于排名这一事实推翻了个体与集体利益一致的预设，而自由主义者所说的完全不受约束的市场体系的基础也随之瓦解。

并非只有反政府激进分子误解了市场交换的逻辑。许多进步思想家长久以来持有的观念也与这一逻辑存在冲突。尽管这些进步思想家指出的经济和政治体系的缺陷是真实存在的，但是他们弄错了造成这些缺陷的原因，因此也常常搞错了应对这些缺陷的最佳途径。

例如，许多左派批评家把市场失灵归咎于不充分的竞争，但是这一

问题本身就是市场竞争的一个基本属性。如今市场的竞争前所未有地激烈，却无益于缩小市场失灵的范围，反而起到了推波助澜的作用。

棘手的间接伤害

令我和自由主义的朋友很容易产生分歧的问题是，什么会对他人构成危害。我们都认为政府对盗窃他人财产或向他人施以暴力等行为的限制是合理的，但真正棘手的是一些间接形式的伤害。

比如，尽管使用合成类固醇的短跑运动员甚至没有与离他最近的对手产生肢体接触，却使对手付出了巨大代价。对手可以选择不使用合成类固醇，并因此输掉比赛，从而使自己付出大量时间和努力却一无所获；这位对手也可以选择使用合成类固醇，恢复自己的竞争优势，由此招致严重的长期健康风险。无论做出哪种选择，最初使用合成类固醇的运动员都会给他造成相比于人身袭击或丢失自行车更重大的伤害。

然而，许多自以为是的自由主义者依然坚持认为运动员有权选择是否使用提高成绩的药物。但是，为什么这项权利要凌驾于其他人避免因此而受到伤害的权利之上呢？仅仅因为伤害是间接的，这种伤害就可以打折扣了吗？

穆勒的伤害原则如果要保持一致性，就必须考虑间接形式的伤害。我对什么会给他人造成伤害的定义可能对有些人来说太过宽泛，但是，即便是从自由主义者的立场也很难驳斥这一定义。就算自由主义者完全可以自由地同他人一起建立想要的任何类型的社会组织，他们也会选择同等对待直接伤害和间接伤害的社会。这有时会让人疑惑，因为相比于直接伤害，间接伤害通常更难以衡量。但是，有时直接伤害也是很难衡量的，而即便是在这样的情况下，一般也不会有人质疑将直接伤害纳入考量。

最根本的问题在于，如果有人对什么会对他人造成伤害进行了合理

的定义，那么现代化工业国家的监管机构就将与穆勒的伤害原则保持一致，实际上这也正是穆勒原则要求的。

只有抛弃错误认知才能解决问题

错误认知制约了政治讨论，这对解决重要问题形成了阻碍。如果我可以摒弃这些认知，许多问题就没有看上去那么棘手了。

例如，迅速增加的财政赤字并没有看起来那么难以克服。仅凭削减支出并不能减少赤字。美国在生育高峰期出生的一代人即将退休，选民们也不支持减少在社会保障与医疗保险上的投入，人们必须同时增加额外收入。好消息是，这并不需要任何人做出牺牲，但前提是美国国会愿意从根本上重新设计税收制度。尽管茶党及类似群体公开反对一切形式的征税，但其实许多税收会让美国变得更富裕，而不会变得更贫穷。国家的发展更多依赖于这些税收。

对一项活动征税不仅能增加收入，同时也会抑制这项活动，税收抑制经济发展的观点也基于此而产生。对于一些有用的活动来说，通常如此，这些活动是当前税收收入的首要来源。例如，工资税会抑制创造就业，所得税（也是一种储蓄税）会抑制投资。

当征税对象变成那些会对他人造成伤害的活动时，情况正好相反。驶入拥挤的高速公路会延缓大家的行进，给他人造成上千美元的损失，但高速公路相比于其他道路节省的时间微乎其微。购买重型汽车会使他人面临风险，其实轻一些的车型照样可以满足人们的需求。

对有害活动征税可以达到一石二鸟的效果：在筹集急需资金的同时，还可以抑制这些弊大于利的活动。

反政府分子明确地将这类税收指责为“社会工程”，试图“控制我们的行为，操纵我们的选择，以及改变我们的生活方式”¹⁵。例如，政府为了减轻对外国石油的依赖而征收汽油税，总会招致这种指责。

这样的抱怨是很空洞的，因为几乎所有的法律、法规都是社会工程。禁止杀人和偷窃的法律旨在控制人们的行为，操纵人们的选择，并改变人们的生活，因而就是社会工程。同样还有噪声条例、速度限制，乃至停车标识和交通信号灯。社会工程是不可避免的，因为如果没有社会工程，狭隘地追求个人利益的行为会给他人造成无法接受的伤害。只有坚定的无政府主义者才会支持一个没有社会工程的世界。

如果彻底禁止有害行为是可以接受的，那么为什么基于同样目的的税收不可以呢？事实上，在抑制这些有害行为的方法中，相比于法律和法规，税收的成本和强制性都更低。那是因为，税收对危害的抑制作用主要依赖于那些更容易改变行为的人们。

例如，如果政府对污染征税，那些能以最低的成本减少污染排放的企业会抢先采取减污的举措，从而避免被征税。同样，如果政府按重量对汽车征税，那些更容易换成轻型汽车的人就会换一辆，其他人则会觉得缴纳税收更划算一些。

会对他人造成不当损害的行为还有很多。酗酒会增加其他人遭遇车祸的概率，吸烟会使他人患相关的疾病，排放二氧化碳会加剧气候变化带来的损害。

对有害活动征收的每1美元税都会相应地减少对有益活动的征税。由此获得的资金不仅可以削减财政赤字，而且可以降低工资税。削减这部分税收可以刺激就业，使低收入家庭担负得起对有害活动新征的税。

当然，无论何时，都应尽量减少政府支出的浪费。国防支出和石油公司的补贴都没有受到预算削减的影响，同样还有臭名昭著的乙醇补贴项目。这类项目需要更仔细的审查。

我要再次强调，未经恰当考量的支出削减通常是弊大于利的。久久不对高速公路进行维护，实际上会增加未来的财政赤字，因为随着维护的不断延迟，成本会急速累加。

对危害性活动征税是削减支出、筹集必要的资金的最佳途径。只有认为人人都有权利对他人造成伤害的人，才会认为这侵犯了所有人的权益，因而反对征税。并且，由于这些税收会使国家的“经济蛋糕”变大，我们更不应以无法承受其代价为由而反对。

只有在经济恢复到充分就业的状态之后，才应分段征收这些新税。但是即使在20世纪50年代联邦税收最低的时期，如果当时的美国领导者仍坚持认为一切税收都是罪恶的，政府就不太有意愿征收新税。

这样转变税收政策可以增加人们的可支配资源。用不着任何牺牲，人们就可以创造绰绰有余的收益，从而消除政府债务，整修长期被忽视的公共基础设施。

这是一个大胆的观点。但是在本书中你将会看到，这一观点直接来源于人们已经熟知的逻辑和事实。简单来说，好消息是，每个能够清楚地了解并聪明地应对有害行为的社会，存在大量可供驱遣的闲置资金。

THE DARWIN ECONOMY

THE DARWIN ECONOMY

02

达尔文之楔：
竞争无法弥合个体与群体的利益分歧



我生于1945年，像我这个年纪的人如果要预测100年后将会发生的事情的话，到那时，就算预测失败，作古之人也用不着担心会被朋友取笑了。如此，我便斗胆做出如下预测：一个世纪以后，如果问及一众经济学专家谁才是经济学之父，多数人的回答将是达尔文。

如果在今天来提出这个问题，显然我的经济学同僚中会有超过99%的人推举亚当·斯密。对我个人而言，我对达尔文的推崇丝毫不影响我对斯密的赞赏。恰恰相反，翻开这位18世纪英国道德哲学家的杰作——《国富论》中的任意一页，我仍然会惊讶于他广博、深刻的洞见。

达尔文自然非等闲之辈，然而鲜有生物学和经济学领域之外的学者将他的名字与经济学的观点联系在一起。精通达尔文进化论的人会发现，他实际上深受托马斯·马尔萨斯（Thomas Malthus）和大卫·李嘉图（David Ricardo）的著作的影响。马尔萨斯曾是斯密的学生，而李嘉图则深受《国富论》的影响。因此，即便我的预言成真，斯密的拥趸照样可以视斯密为经济学的“祖父”。

我所做预测的依据是达尔文和斯密有关竞争过程的观点之间微妙但至关重要的区别。如今，斯密以其“看不见的手”这一理论而闻名，据其现代的信徒们所述，该理论认为客观的市场力量可以引导贪婪的个人行为，实现整体利益的最大化。就像我在第1章中提到的，这样的界定是过于简化的，但是它捕捉到了斯密理解竞争过程的重要维度。不管怎样，对于反政府干预者来说，“看不见的手”理论对于不受管制的市场运行结果的乐观描绘，便是他们世界观的基石。这些人认为管制是不必要的，因为他们相信不受管制的市场可以良好地自行运转。

达尔文关于竞争过程的观点则截然不同。他的观察使他相信，个体动物的利益与其族群更广泛的利益存在深刻的冲突。我预测，迟早有一

天，“看不见的手”将只会成为达尔文更普遍的理论中的一个特例。许多自由主义者笃信的理念，尽管可以在斯密的理论框架下说得通，却并不兼容于达尔文的体系。

“看不见的手”是一个非凡洞见

尽管如此，“看不见的手”理论仍是革命性的天才之见，现在看来更是如此。厂商为什么会乐此不疲地根据消费者的喜好设计新产品？又为什么会为削减成本而改进生产流程？在斯密及他之前的许多人看来，这背后的动机显然只是赚更多钱。而其他没有清楚看到的，则是这些行为激起的竞争者的反应，以及由此引发的动态机制会如何导致完全有异于初衷的结果。

如果一个生产者找到更省钱的生产方式，他便可以稍稍降低价格，继而抢占竞争对手的市场份额。在短期内，他获得的利润会如愿飙升，但是市场份额的丧失会强烈刺激竞争厂商模仿其创新行为。一旦这种行为蔓延至整个行业，由此产生的竞争会驱使价格降至仅比最新的更低的生产成本高出一点点。这一变动过程的最终受益人便是消费者，因为他们能以越来越低的价格享受越来越好的产品。

《国富论》出版于1776年，距今不足250年的历史，斯密的“看不见的手”是多么非凡的阐述！激发了斯密的深刻思考的这一规律，数以千计的杰出头脑也曾在更早时候观察到过，然而竟没有一人意识到其重要性，亚里士多德没有，哥白尼没有，牛顿也没有。

斯密很清楚，不加监管的市场并不总会产生好的结果。他关注的市场失灵多半涉及占有支配地位的市场领导者的卑劣行径。对此，他写道：

扩张市场和缩减竞争总会符合那些以利润为生的厂商的利益。

这只代表人群中的某个阶级，而这些人利益从来都不与公众利益完全一致，他们的利益通常来自欺骗乃至压迫公众，在许多情形下两者兼有。¹

在斯密看来，市场失灵是由于缺乏有效竞争：厂商可能会在产品质量上欺瞒消费者；或者试图降低价格，以便将竞争者驱逐出市场，一旦成功后就重新抬价。这种混乱在斯密所处的年代屡见不鲜，现在尽管已经不那么常见了，却依然存在。

左派的社会批评家长期以来致力于通过非竞争行为来理解市场为何失灵。例如，美国经济学家约翰·肯尼斯·加尔布雷斯（John Kenneth Galbraith）曾强调，亚当·斯密的信徒们所设想的“传统秩序”与斯密所认为的能够更准确地描绘现代市场的“修正秩序”之间存在差异。²在传统秩序下，进入市场的消费者具有完备的偏好，而厂商则以尽量低的代价满足消费者的需求。但是在加尔布雷斯的修正秩序下，具有影响力的厂商首先会考虑提供最容易生产和获利的产品，并雇用美国广告业的营销商们来说服消费者购买这些产品。

许多经济学家对这一主张持质疑态度，并举出一些企业失败的典型案例，比如在加尔布雷斯所处时代的福特埃德塞尔型汽车。³福特汽车公司在1957年9月大张旗鼓地推出这一车型，该车型以亨利·福特的儿子埃德塞尔·福特（Edsel B. Ford）命名，其高额的宣传预算中还包括一档广受欢迎的电视特别节目《埃德塞尔秀》（*Edsel Show*）。然而这从未真正点燃消费者的购买热情，这款车在1960年便停产了。

就在几年前，微软公司花费将近10亿美元开发了一款针对青少年的智能手机Kin。这款手机在2010年4月正式进驻商店，却在仅仅45天之后就因为滞销而被匆匆撤掉。

尽管存在这些失败案例，广告宣传对消费者偏好的影响仍是毋庸置疑的。但是在那些质疑市场的人看来，宣传的魔力是一柄双刃剑。驱

动“看不见的手”的力量是人的贪婪，如果生产者能够以虚高的价格出售次等的产品，他们绝不会放过这样的机会。如果竞争厂商使消费者以为买到的是更好但更便宜的产品，厂商们就可以赚很多钱。现在广告宣传的目的正在于此。如今的市场竞争显然远不够充分，但是相比于亚当·斯密的年代已经非常接近于无摩擦并且信息完全的理想状态了。

竞争的缺陷

达尔文关于竞争的理念并非来自厂商之间的互动，而是来自动植物种群中个体之间的相互作用，他意识到这两个范畴之间有着深刻的相似性。达尔文的分析揭示了竞争机制的缺陷，但这并非由于竞争不充分，而恰恰是源自竞争过程本身的逻辑。他的理论的核心前提是，自然选择只对能促进繁衍的个体性状变异有利。如果一种性状使个体可以更好地生存和繁衍，这种性状就会被选择，否则，就会逐渐消失。达尔文认识到，在许多情况下，对个体有利的变异同样也会有利于种群的利益。但是，他也同时看到，许多增进个体利益的特性却会对种群造成损害。

瞪羚的奔跑速度属于前一种情形。成年瞪羚的长途奔跑速度可保持在每小时48公里，而其短时爆发速度可达到每小时96公里。瞪羚的速度是如何变快的？站在进化的角度，跑得更快似乎是更好的，但事实并非如此，不然所有的种群都应该跑得更快才对。绦虫的行动是很慢的，毕竟在它们的生存环境中，速度快慢并不那么重要。瞪羚之所以跑得快，是因为在其进化环境中，跑得快是生存的决定性因素。在瞪羚的天敌中，非洲猎豹的速度是非常快的，并且在两者生存的地域中很难找到藏身之处。因而对于现代瞪羚的祖先来说，跑得慢意味着更容易被捕食。

促使瞪羚跑得更快的选择压力来自被天敌捕食的威胁，因此更快的奔跑速度对瞪羚个体和整个种群都是有利的。⁴在一定程度上，跑得更快对于所有个体和种群都是有利的。在速度这一特性上，达尔文的自然选择理论与斯密“看不见的手”理论中关于节省成本的创新行为和设计更

有吸引力产品的行为的论述是一致的。

然而，还有其他一些特性，在增进个体繁殖适度的同时却会给整个种群造成损害。这种冲突往往体现在能够增加种群内部个体一对一竞争优势的特性上。

一个例子是雄性麋鹿大到异乎寻常的鹿角。这些作为武器的鹿角并非用来对抗天敌，而是用来与其他雄性麋鹿争夺雌性麋鹿。在这样的争斗中，起到关键作用的是鹿角的相对大小。导致鹿角变大的突变使雄性个体更容易击败它的竞争对手，获胜的雄性麋鹿取得了与雌性麋鹿的交配权，其后代便也带有这种突变特性，因此，这种突变就会很快扩散开来。更多的特性随着一代代的繁衍而逐渐积累，形成一场“军备竞赛”。这一过程似乎已经稳定了下来，目前北美雄性麋鹿最大鹿角的长度超过1.2米，重量超过18千克。

尽管在这一进程中，每种基因突变都增进了个体的繁殖适度，其累积效应却使麋鹿群体的情况变得更糟。例如，庞大的鹿角使雄性麋鹿在繁密树林中行动受限，因此更容易被狼捕杀。鹿角相对较小的雄性个体会更容易从捕食者那里逃脱，但是由于它在与其他雄鹿的斗争中具有先天的劣势，它也很难将小型鹿角的基因传递给下一代。

简而言之，雄性麋鹿面临集体行动的困境。鹿角较大的个体更容易赢得战斗，使对手更容易输掉战斗。因此，更大的鹿角给个体带来的收益远远超过群体。如果每只雄性麋鹿的鹿角都变得更小一些，那么整个群体将会获益。

之所以会存在个体和群体之间的冲突，是因为繁殖适度是一个相对概念。能够在自然选择中胜出的，是具有相对优势的特性。并非任何有用的基因变异都可以在种群中得到扩散，它还必须在与其他变异的竞争中胜出。通常，如果进化中的一个特性有助于个体战胜种群内的其他个体，那么它对整个种群来说就是一个缺陷。在生存竞争的每一个阶段，相对优势使个体之间互相淘汰，最终，当这场竞赛平复下来时，整个种

群的发展却会面临阻滞。

我计划以自由主义者的基本主张为起点进行探究，这些“热爱自由”的人们，在这样的主张下，会选择什么样的政府。对此，达尔文的核心见解可以提供许多关键思路。自由主义者的选择显然有赖于他们对自由竞争市场的期待，而他们的期待一直受到斯密“看不见的手”理论及其对完全竞争的假设的引导。传统的市场怀疑论者担心的竞争障碍在达尔文关注的自然环境中几乎完全不存在，然而达尔文对于竞争过程的见解本身就为市场怀疑论提供了有力支撑。更深入地研究这种见解如何动摇“看不见的手”的解释，是很有意义的。

自由主义者对不受管制的市场的信念建立在几个前提之上，其中最重要的两个是：消费者拥有完全的信息，以及市场是竞争性的。我们必须逐字逐句地仔细研究，才会发现达尔文的理论框架对这两个前提构成的挑战。当然，面对数百万的商品和服务，消费者不可能了解所有可行的选项。但是消费者对于可能重要的选项，可以说是相当了解的，或者他们可以选择去获取这些信息。同样，没有一个市场可以满足完全竞争的严格条件：市场出入完全自由，有大量的生产同质、标准化产品的企业，每个企业只占据市场份额的一小部分。然而，多数市场可以被视作是竞争性的，因为如果出现更好的机会，一些企业主最终会做出改变，并吸引消费者的注意。

简单来说，达尔文所带来的挑战无关乎传统市场怀疑者所主张的市场竞争不充分。如果自由主义者要授权政府来管理企业，原因也不会是竞争不充分。

但是，亚当·斯密的“看不见的手”理论同样需要一些额外的假设，而已有的事实会使任何理性个体质疑这些假设。第一个假设是，面对众多选择，人们会理性地思考相应的成本和收益。第二个假设是，就物质资源对人们幸福感的影响而言，重要的是绝对收益，而非相对收益。有力的证据表明，这两个假设都失效了，从而削弱了“看不见的手”理论的

效力。你们会在本书中看到，第二个假设的失效所造成的问题尤其严重。

至于第一个假设，大量研究表明，人们并不像传统理论中所说的那样在意成本和收益⁵。比如说，对于立即会造成影响的成本和收益，哪怕再小，人们通常都会十分在意；但对于不确定会产生或者延迟很久才产生的成本和收益，即便再高昂，人们也会不为所动。通常，人们对收益和损失的态度也存在差异。这种不对称性本身并不意味着非理性，但是这种差异通常会导致对决策者来说不好的结果。

理性假设失效了，于是人们授权给政府，由政府采取各种措施重整决策环境，以使人们做出更好的选择。但是我在此不会阐释政府会如何应对这些假设的失效，这不是本书的主旨。塞勒和桑斯坦在他们合著的《助推》中早已对此进行了全面的论述。我要关注的是，如果一个自由主义者相信，在经济决策中，人们更关注的是他们的相对处境，那么当这位自由主义者拥有完全的决定权时，他会选择什么样的政府？我们将会看到，这些讨论并没有偏离理性思考。

成为更好的，而不是最好的

正如达尔文清楚看到的，我们的生命在很大程度上被置于一条相对地位的评价曲线之上。一种有益的基因突变不仅应有助于个体繁殖大量后代，而且必须使该个体比不具有这种基因突变的竞争对手能产生更多后代。因此，繁殖适度是一种典型的相对概念。能够生存和繁荣的，不必是宇宙中最强壮的、跑得最快的或者最聪明的个体，它可能很脆弱、行动迟缓、头脑愚笨，但真正关键的是，它能够在种群内争夺资源的斗争中胜出。

为了做到这一点，它的神经系统必须接收有关周围环境的信息，预测不同的行为在多大程度上能实现相应的目标。但同时，神经系统还必

须执行另一项重要的任务：对这些目标进行排序。哪些目标是最为重要的？在紧急情况下，又该放弃哪些目标？

只有了解了人类的动机结构，我们才能对市场的运作机制有所了解。人类的大脑经历了数百万年的进化，人类的动机便蕴藏其中。每一代人类大脑最直接的目的是指导个体做出最有利于将基因蓝图传递给下一代的行为。在解释人和其他动物的行为动机的理论框架中，达尔文提供了唯一具有科学性的框架。

将大脑视作一个进化完成的器官并不意味着否定文化对人类行为的重要影响。很显然，人类的演化及人类文化的演化是同时发生的，相互之间存在显著影响。⁶许多人认为，进化而来的大脑是人类动机系统的硬件，而文化则是软件。⁷正如这一观点所强调的，人类动机异常复杂，并且是多维度的。

经济学家们通过建立模型来解释市场的运作方式时，不得不进行一些简化，但是有些简化过于极端了。例如，多数经济学家假设人是纯粹自私的，但其实存在极有说服力的事实表明人们拥有超越狭隘的个人利益的动机。⁸他们还认为人们从消费中获取的满足感仅仅取决于消费的绝对量，然而，同样有极具说服力的事实表明相对消费量也很重要。⁹

由于繁殖优势总是首先取决于资源的相对占有情况，大脑必然会关注相对地位。包括早期的人类社会在内，多数脊椎动物群体是一夫多妻制的，这意味着雄性个体可以占有多个配偶。在这样的群体中，地位高的雄性个体拥有多个配偶。在这种“抢椅子”游戏的逻辑下，地位低的雄性个体总是什么都得不到。

饥荒是人类进化过程中时常面对的生存威胁，但即使在饥荒最严重的时期，也总会有些食物的。谁可以获得这些食物几乎总是取决于相对收入。那么现在看来，在人类种群中，最穷的个体最有可能饿死。

在这种背景下，请你设想两种基因变异：一种变异下的大脑极为关

心相对地位，另一种则完全相反。一般来说，一个人越在乎某物，就会动用越多的脑力和体力来获取它。因此，更关心相对优势的个体更可能尽一切努力获取和维持较高的地位。这会使这些个体更有可能在饥荒中存活下来，并成功繁衍，从而增加他们的基因在下一代中出现的概率。

当前的环境自然与我们祖先生活的时期截然不同，但是相对地位依然重要，这通常是站在工具理性⁽¹⁾的角度来说的。例如，你要去参加工作面试，打算穿得得体一点，但是衡量看上去是否得体的标准几乎都是相对的。有些面试官可能并没有意识到求职者们的穿着有什么不同。然而，如果你穿的是价值500美元的套装，而其他求职者的衣服只有200美元，相比于他们穿2 000美元的套装的情况而言，你将更有可能得到下一轮面试的机会。正像前面讨论过的，想要把孩子送到好学校的家长，必须出价高于其他家长才能获得好学校附近的住房，相对收入几乎是决定他们能否这样做的唯一因素。这里再次体现了“抢椅子”游戏的逻辑：不管人们赚多少钱，只有一半的孩子可以进入排名在前一半的好学校。

幸福感大部分取决于相对收入水平

有经验证实，许多消费经历带来的心理反应极大程度上取决于所处的环境。例如，有些人通过飙车获得乐趣。但是多快的车速可以带来这种乐趣呢？如果不考虑相关情境，这个问题是完全没法回答的。在20世纪20年代，速度能达到每小时96公里的汽车已经很快了。现在，随着时间的推移，速度的标准发生了相当大的变化。

我的第一辆跑车是1955年款福特雷鸟。这辆车的引擎是V型8缸引擎，体积为4.79立方分米，功率将近150千瓦，可以在11.5秒内使车的速度由0增至每小时96公里。我永远不会忘记那时候这有多快。2010年前后，我开的是2001年款马自达米亚达系列，可以在6.5秒内提速至每小时96公里。接近于性能极限的是2011年款保时捷911Turbo系列，可在3秒内达到同样的速度。这个速度真的很快，但是不会永远如此。

对相对地位的重视是人类大脑进化的一部分，这一假设不仅可以得到现实经验的验证，对于所处环境中的相对地位的特定的神经生理学反应也可以作为证据。比如，相对地位似乎与血清素浓度之间存在相互影响，血清素可以调节情绪和行为。在一定限度内，血清素浓度的升高可以提升幸福感。像百忧解这样广泛用于治疗抑郁症或其他情绪失调的药物，便能有效提升脑内血清素的浓度。

在雄性体内，性激素睾酮的浓度与相对地位之间也有类似的关系。血浆睾酮水平会随着相对地位的下降而下降；而随着相对地位的上升，血浆睾酮的水平也会上升。例如，一个赢了网球比赛的运动员，赛后其体内的血浆睾酮水平必定会上升，输了的对手的血浆睾酮水平则会下降。与血清素相比，睾酮浓度的升高似乎可以增进机体行为，进而有助于提升或保持相对地位。¹⁰

对于决定主观幸福感的因素的研究，为相对优势的重要性提供了更多证据。早期的研究者发现，尽管一个国家的平均幸福水平在面对显著的经济增长时也会在一段时间内非常平稳，但国家的公民在某一时刻的幸福水平仍极大地取决于收入水平。¹¹并且，近期还有研究表明，一个人的幸福感与其周围人的平均收入呈现稳定的负相关关系。¹²

总而言之，想要阐明市场的运作机制，就需要以“环境会对主观评价产生显著影响”这一假设作为出发点。然而荒谬的是，亚当·斯密的“看不见的手”的理论却假定环境不构成任何影响。拥护这一假设的并不只有自由主义者和右派狂热分子。100多年来，美国政府所做的经济预测和政策分析正是以该模型为依据。

本书接下来的三章将致力于探索情境的重要性如何颠覆人们对市场运作机制的认知。为什么即使是在消费者拥有充分信息、企业主和厂商处于完全竞争的条件下，“看不见的手”仍然会失效？你将会看到，将简单的情境因素纳入考虑，会成为解答这一问题的关键。对市场失灵的错误解释导致许多左派人士出于错误的理由而主张管制市场。他们声称，

有必要通过管制来保护人们免遭拥有市场支配力的企业主和厂商的剥削。但是，管制市场的真正原因是避免人们自己过度竞争所造成的后果。

传统经济学没有考虑情境因素

我们将会看到，造成浪费性竞争的，只不过是主观评价在许多领域强烈依赖于具体情境这一事实。情境会影响主观评价，这是没有争议的。有哪个明智的人在一群见多识广的人面前还会坚称情境无关紧要？比如说，他能说服人们接受在相同情境下，一辆车跑得快不快跟其他车的车速没关系吗？这是绝对不可能做到的。但这恰恰引出了另一个问题：我们最初是如何逐渐接受完全不考虑情境因素的理论的？思考这个问题将有助于我们接下来探究情境对经济行为的影响。

最近几年里，我向诸多经济学家请教过这一问题。其中有一位表示，只要可以证实，相比于传统模型，情境因素可以提高对数据的拟合，那么情境的重要性就可以得到充分的认可。然而事实并非如此。例如，在前面我曾指出，即使存在有力的证据表明建立在相对收入基础上的理论对数据拟合得更好，经济学家也不愿放弃基于绝对收入的储蓄理论。¹³

另一位经济学家则推测，许多经济学同行担心的是，如果真的将情境纳入考虑，可能意味着缺失严谨性。毕竟，让许多经济学家引以为傲的是能够将他们的理论表达为数学形式。这类阵营里的一些经济学家可能会担心，关注相对地位会使他们被误认为是社会学家。然而就像过去一些研究所证实的，将相对地位纳入考量并不会对规范化建模构成障碍。¹⁴只要经济学家们乐意，这些模型尽可以更加复杂。

还有一位经济学家觉得，经济学家们之所以不愿意考虑情境因素，根本上是因为这样做会削弱久负盛名的“看不见的手”理论的效力。这种

解释至少是适用于一些经济学家的，但是仍不充分。毕竟，在经济学家做出的政策建议中，已经涉及了许多其他形式的市场失灵。例如，即便是市场最狂热的追捧者也会毫不迟疑地承认，当控制污染的成本很高时，政府干预是富有成效的。

我能想到的最后一个可能性在我看来是最说得通的。在30多年里，我写的书始终是有关相对地位的，最常见的来自自由主义者和其他右派人士的回应是谴责我煽动阶级斗争。他们反驳相对地位的原因与反对施虐狂（政府）的原因如出一辙。但是，将相对地位纳入讨论远远不同于加强政府的施虐倾向。社会理应关注如何消除嫉妒等负面情绪，这是所有人都认同的。总有人拥有更多更好的东西，过于关注这一事实必定会带来不愉快。教育孩子不要嫉妒别人运气好是值得尝试的方案。不过即使这样的教育是成功的，也不能消除由关注相对地位导致的挥霍性后果。这些浪费更多来自这样的事实：生活中的许多重要奖励取决于相对消费，而远不仅仅是出于嫉妒。无论如何，就像开超速罚款单并不等于认可超速驾驶一样，解决集体行动困境的税收方案并不等于认可嫉妒情绪。

THE DARWIN ECONOMY

THE DARWIN ECONOMY

03

尚未被利用的机会：
市场失灵传统解释的缺陷



亚当·斯密对“看不见的手”效力的担心主要在于拥有支配力量的市场主体限制竞争。对市场持怀疑态度的自由派人士很快接受了这些担忧，并继续将竞争受限视为市场失灵的最重要原因。但是达尔文清楚地看到，完全竞争也并不总能引导个体行为增进公共利益。他意识到，个人与集体的利益常常存在很大冲突，在这种时候，个人利益通常会胜出。

达尔文的阐述有一个重要特征，就是认为即使个体利用一切可能的机会来获取潜在收益，市场失灵依然存在。请回忆一下托马斯·谢林所举例子中的冰球运动员，他们在比赛时明明可以戴头盔却不戴，但对于制定要求必须戴头盔的规则，他们却大多投票支持。在这个例子中，个体拥有充足的信息，也足够理性，但是达尔文所设想的市场失灵依然会发生。尽管更小的鹿角对于整个雄鹿群体来说是更有利的，却不利于雄鹿个体。

相比之下，不完全竞争只有在个体总是不能利用可行的选择来使自己获利时，才会导致市场失灵。借用经济学家熟悉的比喻：一定还有尚未被发现的机会。我将证明，这就是市场失灵传统解释的缺陷。

为什么实行员工自治的企业没有繁荣起来

我早些年在康奈尔大学同一群经济学家进行讨论时，第一次意识到这种传统解释的缺陷。这些经济学家里有人开展了一个针对员工自治企业的调查项目。他们坚定地致力于实现社会正义的理想，非常怀疑传统的层级分明的资本家企业。他们认为，这种层级分明的企业，正是污染、异化及很多其他严重的社会问题的源头。他们确信，消除这些弊端

的方法无疑是把权力交给员工，以此取代传统的工业企业组织形式。

基于人们重视个人自主性的事实，上述想法未必完全不可行。¹研究表明，幸福感更高的员工工作效率也更高。²如果一位员工并不认同公司的使命，又整天被一位专制的老板呼来喝去，这显然会打击他的士气。员工自治的经济体系的倡导者很肯定地认为，一旦员工掌握主导权，情况就会得到显著改善。有三位资深研究人员在这一新兴领域出现伊始就这样表述：

在员工参与度更高的工作环境中，不仅员工的工作效率会更高，而且他们对工作的满意度也会更高。这样一来，用不着向员工施加压力或者将他们置于更危险的境地，每小时的工作产出就可以显著增加至少15%。浪费的减少并非来自工作提速，而是来自员工更强的责任感。这样，资本主义企业因照章工作、怠工、偷懒和员工直接的抵抗情绪而白白浪费的工人劳动就可以得到利用。³

这真是非凡的见解。如果任何一家企业可以通过以上方法提高生产力，同时又提高员工的幸福感，那么这家企业便可以获得巨大的竞争优势。一家典型的企业的工资支出约占总成本的70%，生产力提高15%便可以节约超过10%的总成本。企业可以使价格降低差不多的程度，同时还能继续保持盈利，这足以抢走竞争对手绝大多数的客户。同时，由于企业的员工幸福感提升了，最优秀的人才也任由企业挑选，从而进一步增加了企业的竞争优势。简而言之，这意味着，仍存在大量没有被利用的获利机会。拥有这些优势的企业应如烈火燎原一般席卷市场，将巨额利润尽收囊中。

实际上，员工自治的企业在美国和其他工业国家有着悠久的历史。小型家庭作坊或许是最常见的例子，有许多其他合作式企业也属此类。20世纪60年代后期，我还在加州大学伯克利分校读研究生，那时我们很多人会去利奥波德氏（Leopold's）买东西，那是一家由员工管理的音像

店。此外，许多城市都有合作经营管理的食品零售商店。

然而，这类组织在经济体系中所占分量极其有限，也确实没有事实表明这类组织会在市场上盛行。这一现象引发了一个基本问题：如果员工自治的企业这么好，为什么人们看到的这类企业却寥寥无几呢？

我曾向康奈尔大学员工自治系统研究小组的几个同事提出这一问题，那个小组里聚集了一帮可能会认真对待这个问题的聪明人。在一番讨论过后，他们给出了经过深思熟虑的解释：员工自治企业的数量之所以没有迅速增长，是因为金融市场拒绝公平地为这些企业提供资金。他们解释说，银行很长时间以来都与传统的资本主义企业打交道。它们对这类企业了如指掌，并且自认为足够了解这些企业带来的风险，因而能够很好地对贷款申请做出评估。为了支持这一解释，我的一个同事提供了一份银行要求当地的员工自治的企业填写的贷款申请表格。这份表格的长度足有普通申请表格的两倍，并且会问许多尖锐的问题。

我对此的看法是，这些不便并不足以解释为什么员工自治企业没有盛行开来。如果相比于传统企业，这些企业可以享有超过10%的成本优势，他们完全负担得起因填写冗长的贷款申请表格而额外花费的时间。事实上，即便是整个银行系统完全将它们拒之门外，它们也有能力发展壮大。许多企业都是从小规模起步的。一个有着成本优势的企业，如果得不到来自外界的贷款支持，它在一定时期内的发展可能相对缓慢，但是10%的成本优势本身就是可实现扩张的巨大的内部资本来源。

更重要的是，资本主义银行系统不太可能一直不为信用良好的贷款企业提供资金。所有的贷款都必然蕴含一定的风险，但是任何一家比竞争对手有10%成本优势的企业的风险是极低的。如果银行家们担心他们不善于应对员工自治的企业，那么他们的保守做法不应该是完全拒绝贷款，而是应该提高贷款利率。

还有人提出，银行不愿意把钱借给员工自治的企业，是因为它们认为一旦这种所有制结构的企业扩散开来，会威胁资本主义体系的长期利

益。⁴但是如果以为银行家会出于这个原因而错过一次有利可图的机会，那就彻底误解了资本主义的实质。人是贪婪的，资本主义始于这一假设。在《国富论》里一段最广为引用的段落中，亚当·斯密写道：“屠夫、酿酒商或面包师并不是出于仁慈才为我们提供晚餐的，而是出于他们的私利。”⁵银行家不是利他主义者，并不肩负着通过牺牲自己来增进同类利益的使命，他们只是努力想要赚一笔的资本家。所以说，哪怕贷款给员工自治的企业会削弱资本主义体系这一不太合理的假设真的成立，也依然不足以令一个理性的银行家拒绝放贷。

这一问题类似于常见的过度捕捞的悲剧，对此，我会在第10章进行更加详细的阐述。这就是达尔文论述的个体与群体诱因之间的隔阂导致的又一种形式的市场失灵。渔民不会因愚蠢而耗尽渔场资源，因为任何一个单独的个体进入捕鱼行业并不会对最终结果造成显著影响。同样，在银行业，如果足够多的银行将钱借给员工自治的企业，使这些企业繁荣发展，由此可能会导致资本主义体系瓦解，但是即便某个银行家抽身而退，结果也不会发生什么变化。在这种情况下，即使坚定地致力于维持当前的企业形式，一个贪婪的银行家也没有任何理性的动机错过这样一个贷款获利的机会。个人的抑制行为会导致他错失一些利润，并不会挽救资本主义的消亡。

而且，为什么银行家要忠于这样一个目标呢？在多数客户都是员工自治企业的世界里，银行家们照样可以活得很好。在那样一个世界里，成本真的可以降低10%，最终银行本身也会变成员工自治。早期与这类企业有过接触的银行确实可以占据强有力的竞争优势。

然而，正像前面提到的，并没有有力的证据表明员工自治的企业拥有任何竞争优势。在结束研究生学业几年以后，我去了一趟伯克利，发现利奥波德氏，也就是那家由员工管理的音像店，已经不在我上学时候的位置了。我询问朋友这家店怎么样了，朋友告诉我它早在几年前就已经倒闭了。但是它的主要竞争对手，也就是附近的一家折扣音像连锁店

依然在经营。

这又一次说明，虽然员工自治的企业已经存在了很长时间，但是它们从来没有在经济领域占多少比重。无论如何，没有迹象表明这些企业一定会席卷整个市场。唯一能对此做出解释的是，相比于传统企业，在一些能够决定企业成败的关键方面，员工自治的企业可能具有一些优势，但是肯定会在另一些方面表现不佳。例如，利奥波德氏音像店的员工可能从参与日常管理决策中获得一种责任感，从而提高了他们的工作效率。但是久而久之，或许他们开始对频繁参与会议感到疲倦，又或者，问题在于负责管理的员工提供的是在他们看来对人们有益的而不是人们需要的产品。

如何验证经济理论的正确性

我与员工自治企业的倡导者们的交流是一次带有意识形态价值的实际探讨，我的经济学同行们知道市场制度存在严重的缺陷，也同样知道资本主义剥削是导致这些问题的原因。但正因为认识到这一点，他们才不愿或不能剖析他们所持理论的更深层意义。

在过去这些年里，我鼓励学生们在思考市场失灵时尽可能地完全摆脱他们的意识倾向。如果他们对于为什么市场以某些特定的方式失灵持有某种假设，那么对该假设最重要的检验标准则是：这是否意味着尚有人们未加利用的机会。员工自治企业的倡导者们的阐释显然没有通过检验。如果这些企业具有他们的倡导者们所宣称的那些优势，那么企业家们就可以花费上亿美元收购传统企业并将之重组为员工自治的企业。他们只需要买下现有传统企业的股票，将决策权移交给员工，接着将重组企业的股票卖给新的员工股东。由于重组后的企业相比于原先的会更有效率，这些股票的卖出价格会远高于它们的买入成本。然而，先前也提到过，并没有出现以这种方式重组企业的热潮。

简而言之，长期存在大量未加利用的机会的理论并不好。在任何一个稳定地维系了很长一段时期的市场环境下，极有可能不存在大量未被利用的机会。然而，谴责市场结果的左派人士给出的解释也没能通过“没有未加利用的机会”的检验。例如，这些左派批评家常常宣称我们必须对工作场所的安全进行监管，否则员工就会被手握资本的掌权者剥削。起初，这一主张听起来似乎挺有道理。毕竟，企业主常常钱多得花不完，而他们的员工却每天冒着生命危险勉强维持生计。这确实看起来像是剥削。但是我将会解释，其实这也意味着仍存在大量未被利用的机会。

假设由于老板的台锯没有安装刀片防护套，一位木工面临受伤的风险，这位老板应该安装防护套吗？必须对安装这一装置的收益和成本进行分析后，才能做出明智的回答。如果这位木工掌管该企业，并可以自行做出决定，分析结果就是显而易见的了。防护套的成本是很好测算的，为了方便讨论，我们假设每个星期需要花费50美元。这一装置带来的收益就是这位木工为了获得防护套的保护而愿意放弃的最大代价，这自然取决于操作没有防护套的台锯的风险有多高。比方说，如果他愿意承受的代价是100美元，显然是应该安装防护套的。但是，如果说防护套对他的价值只有30美元，那么他的选择就是不安装。

即使这位木工受雇于一家公司，决策逻辑也不会有什么不同，他仍然有必要对比防护套的成本和收益。即使由雇主承担安装费用，成本也不会有变化。而在木工眼中，防护套的价值仍然在于其所能提供的保护。我们假设防护套的收益超过了成本，因为这种情形更有意思。如果对木工来说，每周用于安装和维修防护套的费用只有50美元，而它对工人的价值却有100美元，那么雇主毫无疑问会安装。不这么做意味着会失去原本可以获得的50美元经济剩余。

“看不见的手”理论的质疑者坚持认为，资本主义雇主的贪婪会阻止他安装防护套。但是，这一指责忽略了亚当·斯密理论的本质。如果安

装防护套的收益超过了成本，雇主却不这么做，就会将获利机会拱手让给愿意安装的竞争对手。记得吗？防护套的花费只有每周50美元，而在木工眼里的价值却有每周100美元。我们可以假定，现在这个雇主以当前的工资水平雇用这个木工并没有什么损失。因此，如果一个竞争对手提供安装了防护套的台锯，而工资只比目前的少75美元，那么这个木工会接受这份工作。换工作后，这个木工可以获得每周25美元的额外经济收益，也就是防护套对他的价值100美元减去工资减少的75美元。并且他的新雇主每周也会多得25美元（假定新雇主的成本与当前雇主的相似）。无论何时，如果一个雇主拒绝安装收益超过成本的安全设备，就会存在被遗漏了的获利机会。

许多市场体系的批评者反驳说，使用成本-收益分析来决定是否应该安装安全装置是不道德的，但是这种反驳经不起一丝推敲。防护装置的花费可以用来购买其他对人们有价值的东西。没有任何不好的结果发生的世界是不可能存在的。在决定采取哪一种安全措施时，我们必须权衡收益和成本。

如果你不同意，那么我来问两个问题：你今天检查刹车了吗？如果是，那么你明天还要再检查一遍吗？理智的人对两个问题的回答不可能都是肯定的。如果你刚确认了刹车状态良好，那么它在明天就出故障的概率是极小的。检查刹车会耗费时间和金钱，所以理性的人只会隔一段时间才检查一次。在美国，有汽车检查计划的州通常明确规定一年检查一次。多数人不会频繁地检查刹车，因为这么做不仅花费高而且不能带来显著的好处。如果说成本-收益分析框架适用于考虑检查刹车的频率，那么为什么不能用它来考虑台锯上是否应该安装防护套呢？

其他质疑者的回应是“看不见的手”或许在竞争充分的劳动力市场中是有说服力的，但是他们接着强调，劳动力市场实际上并不是竞争性的。毫无疑问，劳动力市场的流动性确实不如其他类型的市场。例如，一旦一个人与一名雇主签订了合同，相比于其他雇主，他逐渐积累起来

的许多技能会对目前的雇主更有价值，他也就更加不可能获得来自别处的更好的工作机会。从某个时候开始，他实际上就被“困住”了。⁶

然而，即使这位员工被迫要永远为同一名雇主工作，雇主不安装通过成本收益检验的安全设施也是不明智的。如果由雇主支付设施的费用，并通过削减这位员工的工资来抵补这部分花费，员工和雇主的情况都会得到改善。质疑市场的人可能会觉得被“困住”的员工得到的报酬太低了，不能承受他在工资足够高时愿意付出的代价。然而，在这种情况下，对于一个能够实行管制的国家来说，可能会有好得多的选择，只需要要求企业提升工资即可。接下来，员工就可以自行决定是否要承担额外的安全设施的代价。

无论如何，不能过分强调劳动力市场的非竞争性。劳动力流动是不充分的，但是相比于以往，现在人们可以更频繁地换工作了。而且，即使企业知道大多数员工不会离开，有一些也确实还是会离开的，还有另外一些最终也会退休或者去世。因此，企业主们必须维持对新员工的吸引力，这意味着他们必须提高企业的声望。信息时代是藏不住秘密的，一家剥削员工的企业最终会面临严重的用工困难。

总而言之，强调锯刀“太危险”相当于强调提高锯刀安全性的成本将低于相应的收益。如果劳动力市场是竞争性的，员工们也足够重视锯刀的安全性从而愿意承担额外的费用，那么企业主就会有安装防护套的强烈动机，以避免竞争对手挖走他的员工。如果这些都没有发生，那么就可以认为员工事实上并不看重锯刀的安全性，因而不愿意支付额外的费用。他们会乐意无偿获得安全的锯刀，而并不愿意为此付出代价。

关于“看不见的手”更详细的论述可以进一步解释为什么市场上针对工作场所的安全性问题会出现五花八门的选择。一些企业的工资高，工作安全性低；一些企业的工资低，工作安全性高；还有一些企业，两者的水平都是中等。面对这些选择，员工根据自己的偏好在这些公司中寻觅。当然了，偏好在许多维度上都是不同的，但是出于目前的需要，我

们仅考虑两个维度：面对风险的态度和对金钱能买到的东西的渴望。

对于特别谨慎同时又刚好不在乎钱的人来说，最好的选择是一份风险极低的工作。这样的工作可以满足他们对安全的需求，又因为钱在他们眼里不是那么重要，工资低一点也是可以接受的。处在另一个极端的是那些最不关心风险，同时又迫切需要更多收入的人。对这些人来说，最危险的工作会是最好的选择。这样的工作能提供给他们的薪水是最高的，同时他们也可以容忍这份工作额外的风险。对在这两个维度上的偏好都不那么极端的人来说，最好的选择是工资和安全性都处于中等水平的工作。

那么，根据传统的“看不见的手”理论，人们愿意支付多大的代价，就能相应从工作中获得多大程度的安全保障。这是理所当然的，因为用来获得更多安全保障的资源也可以被用来获得其他对自己有价值的东西。如果一位员工没有得到他想要的更多的安全保障，这只意味着他并不愿意付出相应的代价。

对“看不见的手”的质疑

通常，质疑市场机制的人会搬出一些雇主的行为与“看不见的手”理论中理想的市场结果明显不一致的例子，作为有力的反击。沃尔玛，这家美国最大的零售商，就经常被他们用来作为举例对象。比如说，沃尔玛公司有好几次把维修工人锁在店里一整夜，也没有管理人员在紧急情况下放他们出去。沃尔玛为这一举动的辩护是这样可以防止盗窃，但是这也导致至少在某些时候员工无法得到及时的医疗救助。这家公司还因其他有违劳动法的行为受到指控，包括为了不支付加班工资而修改员工的出勤记录。就算这不是剥削，也相差无几。⁷

当然，沃尔玛也为社会带来很多好处，尤其是通过不懈地削减成本而提供的各种各样价格低廉的产品。然而，在这家公司的整个发展历程

中，它频频违反各项法律法规，左派和反对政府管制的人士认为这是对市场“看不见的手”的肆意扰乱。⁸这些法律法规所禁止的行为确实有必要加以制约。然而，就目前已有的事实看来，市场比以往任何时候都更有竞争性。

如果市场真的是竞争性的，那么还有必要制定这些规则吗？是什么在阻止“看不见的手”发挥效力？

更高的收入意味着更高的消费能力

我给出的解释很简单⁹，它源自达尔文的核心洞见：个体利益常常与群体利益相冲突。就像我在第1章中提到的，按照标准的“看不见的手”的理论，一个工人愿意接受安全性更低的工作以换取更高的工资，这意味着额外的收入足以补偿安全性的降低。但是“看不见的手”理论的假设是，收入的价值体现在其所能带来的额外的绝对消费。然而，相比于安全防护只是让工人在少数情况下避免因粗心导致手被刀片伤到，更高的收入可带来另外一个确定且能立刻得到的好处，那就是：消费能力相比于其他人有所提升。

这一点在父母希望将孩子送到尽可能好的学校的例子中体现得最为明显。就像第1章中提到过的，一名员工愿意接受一份更加危险但收入更高的工作，很可能是因为这样他就可以支付更好学区的房子的月付。但是这对其他员工是一样的。而且，学校质量本身是一个相对概念，如果其他人也牺牲工作安全性来换取更高的工资，那么相对来说，没有人会达到自己的目的，他们只会把好学区的房价抬高。

由此，即使是在所有工人完全了解他们会面临的风险且劳动力市场竞争充分的条件下，也是需要对安全问题进行管制的。工人们很有可能更偏好另外一种状态：每个人都从事更安全的工作，即使要付出工资更低的代价。但是，每个工人只能决定自己的工作选择，而不能左右他

人。如果一个工人接受了安全性更高的工作，而其他人没有这样做，那么这个工人就会被迫把孩子送去差一些的学校。要得到他们想要的结果，工人们就必须采取一致的行动。

但是仅仅知道个体的行为会相互抵消并不能消除个体行动的动机。相比于不加管制的私人劳动力市场，世界上所有的国家或组织都会选择采取相似的政策来提高工作场所的安全性。例如，在美国，职业安全与健康管理局就明确规定了不同的行业必须遵守的详细的安全规程。这些国家或组织还要求企业给工人购买工伤补偿险，该保险的费率会随着伤害索赔项目的数量大幅增加。这些手段并不完善，但是毫无疑问，工作场所的安全性将因此提高。

质疑大政府⁽²⁾的人通常会谴责这些政策对个人自由构成了不正当侵犯。例如，塞勒和桑斯坦在描述安全监管政策时写道：“他们纯粹就是在下禁令，这肯定会伤害到一些人。这些法规不允许人们用（政府赋予的）拥有安全工作环境的权利换取更高的报酬，哪怕聪明多识的人想要这么做。”¹⁰这意味着，至少对于信息充分的工人来说，亚当·斯密的“看不见的手”即使在没有管制的情况下也可以提供最优的安全性和工资组合。然而这一理念在人们强烈关注相对状况的时候是站不住脚的。

就像达尔文清楚认识到的那样，生活的许多重要方面被置于一条相对地位的评价曲线之上。是相对收入，而非绝对收入，决定了谁可以在好学区或者景色绝佳的地方买到房子。当相对收入变得重要时，“看不见的手”就失灵了。我们再也没有理由认为个体动机会促使资源实现其最大价值。

这里，读者需要再次注意，我对安全管制所做出的解释与托马斯·谢林对佩戴冰球头盔的规定所做出的解释有惊人的相似之处。¹¹我在第1章曾提过，谢林最初观察到，可以自行选择滑冰装备时，运动员总是不佩戴头盔，但对于制定佩戴头盔的规则却表示强烈赞同。谢林认为导致出现这种矛盾的原因是个体与群体利益之间的冲突。当一个运动员卸下

头盔，他就获得了相对优势，这在他眼里足以与重伤概率的些微上升相抵。其他运动员也面对相同的动机，不可避免的结果就是所有的运动员都不戴头盔。没有哪个队伍的运动员可以获得他们追求的竞争优势，他们却都面临着更高的重伤风险，因此他们才会支持要求佩戴头盔的规定。

仅仅知道导致问题的动机结构并不能解决问题，需要一种强有力的方式来改变个人动机。正是明白了这一点，冰球联盟才制定了佩戴头盔的规定。对于坚持不戴头盔的运动员来说，唯一可行的办法是结成自己的联盟。这与在工作场所制定安全规则是一样的，那些不想受规则管制的人可以搬去另外一个没有安全管制的国家，但是我们将在第12章看到，这样的国家仅有几个，他们也很可能并不想居住其中。

还需要注意的是，我给出的安全管制的理由并不意味着在实行管制之前仍存在未被利用的机会。就拿谢林所举的头盔的例子来说，在没有管制的情况下，市场结果体现的是每个个体的最优选择。如果一个竞争企业提供了更安全但工资更低的工作，也不会有人接受，因为这意味着他们不得不搬到更差的学区。

我做出的解释同样不需要以人们的信息有限、市场流动不充分，或任何人非理性为前提，也不一定要存在一个有力的市场参与者将自己的意志强加给其他参与者。这一解释唯一需要的前提是个体与群体之间存在利益冲突这一达尔文观察到的基本现象。当个体在许多方面追求重要结果的能力显著依赖于相对收入时，就可以满足这一条件。

对相对收入的重视导致的个体和群体动机的差异，有助于我们理解市场失灵的其他方式。在自由主义者和“看不见的手”的支持者看来，市场力量不仅可以引导人们获得最优的工资和安全性组合，也可以获得工资和其他工作特性的最优组合。例如，我们可以考虑专业分工的问题。亚当·斯密强调过，工业革命开始后的生产力爆炸式增长就是更高程度的劳动力专业化分工的结果。然而斯密还意识到，随着工作的专业化程

度越来越强，它对人们在心理上的影响也越来越大。马克思曾强调，资本主义下的专业分工太过了，他认为专业化分工“将劳动从人身上剥离，使人退化为机器的附属物，摧毁了工作的全部吸引力并使之成为令人厌烦的苦工”。¹²

自由主义者和其他支持“看不见的手”的人士对此做出反驳，他们指出，马克思的观点意味着将一些机会闲置。这与他们在关于安全性的案例中给出的理由如出一辙：如果工人们一直认为工作场所缺乏多样性，那这必然意味着在这些工人看来，日常工作的多样性比专业化程度下降所导致的产量损失更重要。但是如果工人们在乎的是相对收入，这一主张就不成立了。如果每个工人的工作多样性都增加一些，同时减少一些收入，那么整个工人群体就会更幸福，因为这一转变不会影响人们实现那些取决于相对收入水平的目标。但是没有哪个个体可以单独换一份更具有多样性的工作而不会损害他实现这些目标的能力。

与此类似，个体与群体利益之间的冲突也可以用来解释为什么人们最终会在不那么理想的环境中工作。劳动力市场为人们提供了获得更高收入的机会，他们可以接受自主性更低的或者有违道德的工作。¹³额外收入的诱惑是人们接受这类工作的重要原因，然而，如果所有工人的绝对收入一起增加，相对购买能力几乎不会受到影响。

对相对收入的重视导致个体利益与集体利益存在分歧，这一事实也质疑了传统经济学中显示性偏好的理论。这一理论认为，相比于人们说的话，我们更能够通过观察人们的行为来了解人们的偏好。¹⁴根据这一理论，如果一个人可以每周削减100美元的工资来换一份更具多样性的工作，而他却没有这么做，那么对他来说，额外的多样性肯定不值100美元。从个人价值评估的角度来说，这是对的。但是个人评价和群体评价并不一定是一致的。如果重要的是相对收入，那些拒绝为增加的多样性支付100美元的人可能会发现，如果所有人都支付100美元获得额外的多样性，增加的多样性的价值将远远不止100美元。

谁也无法否认，在某种程度上，实现人生中的许多重要目标的能力取决于相对购买能力。由此导致的一个直接结果是，当一个人的收入增加时，他实现这些目标的能力提高了，其他人实现这些目标的能力却下降了。或者用经济学家的话来说，同样的活动在带给一个人更多收入的同时，会给其他人带来负的外部效应。

许多自由主义者肯定乐于忽略这一问题，或者坚持认为他们拥有随意给其他人造成直接伤害的权利。但是我将在第6章证明，间接伤害的问题会让正直的自由主义者面临选择的困境。如果一个人想要行使直接给他人造成伤害的权利，那么他就必须为此提供有说服力的理由，而事实证明，这是非常困难的。

我希望这样假设是合理的：对于一个剥夺了球员选择是否戴头盔的权利的规定，一个正直的自由主义者应该不会反对。这个规定当然会剥夺球员的权利，但这也正是球员们会支持这项规定的原因啊！他们知道，如果可以在打冰球时不戴头盔，那么他们一定不会戴。但是他们并不想不戴头盔。

很明显，政府比冰球联盟更有约束力。仅凭这一事实就决定了施行政府禁令需要更加严格的标准。但这只是程度上的差异，本质上并没有差别。我们将在接下来的章节中看到，令人欣慰的是，人们常常不用对行为加以限制就可以减轻对他人的伤害。

THE DARWIN ECONOMY

THE DARWIN ECONOMY

04

干预是为了更好的竞争：
你无法通过不吃饭饿死身体里的寄生虫

经过2005年三次独立的专项拨款，美国国会共计投入3.2亿美元，用于建设一座连接阿拉斯加州凯奇坎（Ketchikan）市区及格拉维纳岛（Gravina Island）上的机场的桥梁。那座桥被戏称为“哪儿都去不了的桥”，这个工程很快就成了政府财政浪费的著名案例。¹

建这座桥从一开始就不是个好主意。凯奇坎的人口在那时只有不到9 000人，格拉维纳岛只有50人。市区和这个岛屿之间的交通靠的是轮渡，每趟的费用是6美元，每趟船的间隔在15分钟到30分钟不等，依具体时间而定。显然，有了桥会更方便，但是远不至于因此而启动这样耗资巨大的工程。

既然这座桥这么失败，为什么一开始还要建呢？反对政府管制的改革者翻来覆去地强调其中的原因。提议开展这项工程的政客希望讨好能从中直接获益的当地居民，以赢得他们的选票，却使远在别处、毫无戒心的纳税者被迫买单，这些纳税者永远不会注意到，更不会抱怨最终由此导致的税收的增加。其他州的立法者之所以支持这一提案，是出于理性的考虑，他们预期当自己有提案时，也可以得到对等的帮助。

让人欣慰的是，这项工程最终在美国民众激烈的反对之下被迫取消。然而，在每笔国会预算中，仍有许多其他提案得以通过，因为这些提案太微不足道，无法进入大众视野。

有一些无意义的活动最终木已成舟。例如，20世纪80年代广为人知的军事采购项目，现在被称为“政府疏忽计划”，据说私人国防用品承包商的一个拔钉锤就向政府要价435美元，一个马桶坐垫要640美元，一个咖啡壶则要7 600美元。²这些事实着实令人不安。如果已经有不够尽责的或者腐败的政府官员被诱使为日常用品支付如此高昂的价格，那么涉及复杂的、难以弄清楚的设备的合同中又可能存在多少浪费呢？

如何避免公共决策中的浪费

反对政府管制的改革者一定清楚地意识到了什么。政府确实存在浪费行为，但问题的关键在于对此该怎么办。许多自由主义者相信，最好的策略是“饿死野兽”⁽³⁾。或者，正如反税收团体——美国税制改革组织的主席格罗弗·诺奎斯特（Grover Norquist）所做的生动描述：“我并不希望废除政府。我只是希望缩小它的规模，小到能塞进浴室或者浴缸吧。”³

“饿死野兽”策略倡导者们的观点很简单。既然上缴到华盛顿、萨克拉门托或奥尔巴尼⁽⁴⁾的钱必然会被浪费，解决方法就是尽量少向这些地方缴钱。加利福尼亚州是“饿死野兽”策略倡导者们的沃土，因为这个州独特的宪法条例允许选民直接决定立法提案。

据说，要想知道美国发展的走向，就应该了解加利福尼亚州。加利福尼亚州率先认真处理汽车尾气排放产生的空气污染问题，率先普及节能电器，也率先为女性和少数族裔争取更多权利，率先应对二手烟污染问题，还催生了在过去30年里引领公众主流意识的反税运动。

1978年6月，13号提案赢得了超过65%的加利福尼亚人的支持，投票率接近历史最高水平。13号提案的官方名称为《人民限制财产税动议》（*People's Initiative to Limit Property Taxation*），俗称“贾维斯-甘恩修正案”，以发起这次公民投票的两位先锋人物霍华德·贾维斯（Howard Jarvis）和保罗·甘恩（Paul Gann）命名。贾维斯是奥兰治县（Orange County）的一名商人，长期反对征税；甘恩是萨克拉门托的保守派积极分子。修正案的主要条款是将加利福尼亚州的财产税降至财产估价的1%，相应地，在随后的每一年里也都不能超过2%。

关于13号提案的一些细则会对加利福尼亚州造成什么影响，人们仍存在争论，但是没有人会怀疑这一提案抑制了长期以来的税收增长趋势。与联邦政府不同的是，州政府通常是不允许持续存在赤字的，毫无

疑问，13号提案也抑制了原本可能产生的政府支出。

由于政府支出总会存在浪费，支持13号提案的人因此声称该提案减少了政府浪费，由此不必担心自相矛盾。但是要想向中立者证明13号提案会让加利福尼亚州变得更好，就困难得多了。立法者提出的所有政府计划总会得到一些选民的支持。这些计划中，有些发挥出了价值，还有一些则没什么用处。资金的亏空迫使政府削减财政预算，这时，哪些计划会被砍掉取决于其支持者力量的大小。然而，就像阿拉斯加州的“哪儿都去不了的桥”所揭示的，仅仅获得一部分人的支持并不意味着该计划会对更大范围的公众有益。由此得到的必然结论是，13号提案也导致其他有价值的计划被取消。

最终结果是什么呢？彼得·施拉格（Peter Schrag）在他1998年的书《失乐园：加利福尼亚的经历，美国的未来》（*Paradise Lost: California's Experience, American's Future*）中试图解答这一问题。⁴施拉格在《萨克拉门托蜜蜂报》（*Sacramento Bee*）当了19年社论版主编，对于在实施13号提案和其他意在限制政府职能的公民票选提案之后的20年里，加利福尼亚州的社会和经济发展历史，他经过细致研究，给出了尽可能中立的解释。

由此涌现的图景与在第二次世界大战之后的一代人心中加利福尼亚州一直具有的“模范且具吸引力”的形象截然不同。在《失乐园：加利福尼亚的经历，美国的未来》出版后的几年里，加利福尼亚州政府的财政状况继续急剧地恶化，相比于美国其他州，其整体发展状况也显著恶化。例如，仅2009年，财政收入不足使加利福尼亚州被迫削减了200亿美元的预算。但即使在13号提案实施的最初20年里，加利福尼亚州也发生了极大的变化。施拉格写道：

在30年前，加利福尼亚州是全美对学校资助最多的州之一，如今排名却落到后25%，无论是学生的身体健康状况、学校获得的公

共拨款，还是考试成绩，都已不及纽约州、康涅狄格州或新泽西州，反而更接近密西西比州的水平。加利福尼亚州一度有名的高速公路系统，如今在美国已经成为最破败的公路网之一。许多公共图书馆缩短了营业时间，甚至有些干脆倒闭了。这个州曾经最为丰厚的社会福利，也被一再削减。曾经一直免学费的院校系统，虽仍然是世界上最庞大的公立教育体系之一，如今却如同所有其他州立大学系统一样，因资金和学费问题苦苦挣扎，甚至更为严重。⁵

13号提案的支持者则称还有其他重要因素导致了加利福尼亚州长期的相对衰退，以此进行反驳。但无法改变的事实是，长期的收入短缺一直是加利福尼亚州的核心问题。

反对政府管制的积极分子坚持认为应对收入短缺的最佳手段是削减浪费性的财政支出。除了浪费性项目的直接受益者以外，会有谁反对这一点呢？困难的是，如何才能在削减浪费支出的同时，避免带来更加严重的附加损失。经验表明，“饿死野兽”的策略并不能解决这一问题。

拥护“饿死野兽”策略就好比让肠道寄生虫病患者不要再吃东西，并解释说因为病人吃的食物是寄生虫的命脉，只要停止进食，这些寄生虫最后就会死掉。好吧，确实如此，但是病人自己可能会先死掉，或者在这个过程中身体受到严重损害。正因如此，杀灭寄生虫的有效策略都是非常有针对性的，试图对寄生虫造成直接伤害，同时尽量减小由此对宿主带来的伤害。

把宿主-寄生虫的类比进一步扩展，会给我们带来更多启示。注意，任何复杂生物体内都是有寄生虫的。减少寄生虫对生物体当然有好处，正是因此，自然选择才总是更青睐拥有有效免疫系统的生物体。但是，自然选择也总是更青睐最有生命力的寄生虫。与寄生虫的对抗会带来好处，也要付出代价。权衡这种对抗的要领在其他领域也同样适用，这个要领是：要先使用代价最小的手段，来消灭最危险的寄生虫。但是

最终会到达一个临界点，这时下一个手段的成本等于由剩下的最危险的寄生虫所带来的损害。超过了这一点，再消灭寄生虫实际上会让生物体的情况变糟。⁶

政府财政浪费的问题中存在着同样的逻辑。减少浪费的最好办法当然是先使用成本最低的手段来消除直接造成浪费的最重要的因素。当然，为此必须先搞清楚为什么会存在浪费。

通常的答案是，政客们会因为重要出资人的要求而支持浪费性的项目。⁷因此，政府财政浪费的反对者可以关注的一个方向是，通过法律规定来减少立法者对大型选举资助的依赖（小额捐助则没有这么严重，因为捐助的个人不能对立法者施加重要的影响）。施行更加严格的竞选财务法的成本相对较低，并且可以抑制一些大的财政浪费。但是美国最高法院最近几年却不怎么愿意支持更严格的竞选财务法。相反，从最高法院2010年针对联合公民诉联邦选举委员会案做出的有争议的裁决中，似乎可以看出他们甚至想要取消由来已久的企业竞选捐款限制。⁸

如果法院不改变想法，政府财政浪费的反对者会继续采取其他应对策略。例如，我们从“哪儿都去不了的桥”案例中得出的经验是，当越来越多的选民了解到浪费的存在时，浪费性的政策在政治上就很难再继续下去了。信息革命大大降低了向选民传递信息的成本，这让我们可以取得一些进步。但与此同时，信息革命还带来了信息爆炸。因此，吸引选民关注某一个特定的浪费性项目仍一如既往地困难。

简而言之，美国民众与政府财政浪费的抗争会持续进行下去。以后，新的技术和更好的制度或许可以带来显著的进步，却无法彻底消除浪费。

事与愿违的支出削减

21世纪初，美国面临的最紧迫的问题之一是，尽管有许多至关重要

的事务待处理，政府却没有资金。例如，在2010年7月，《华尔街日报》报道称，全美各地的预算赤字已经迫使许多州把多处的沥青路降级为砾石路：

平整的公路曾是美国的历史成就，如今在美国的乡下，因为各个郡县苦于财政赤字，以及州政府和联邦政府收入的缩减，这些公路已经破败不堪，取而代之的是砾石路或者其他粗糙路面。随着财政赤字的扩大，许多地方用于修路的州政府拨款也减少了。近年来，在密歇根州的83个郡中，就有38个将一些沥青路改为了砾石路。2009年，南达科他州将至少160公里的沥青路降级成了砾石路。亚拉巴马州和宾夕法尼亚州也开始将沥青路降级成更便宜的砾石路，也被称为“穷人的路”。俄亥俄州的一些郡县则只是坐视公路逐渐被侵蚀和损毁，一点点变成了砾石路。⁹

问题是，这些举动不仅不能节省资金，最终还会带来浪费。坑坑洼洼的不平整路面，每年会给每一辆车造成100多美元的损失，更不用说许多无谓的伤亡了。¹⁰路面保养每拖后两三年，成本就会加倍。因此，现在花1美元保养路面，就可以不用在三年后再多花1美元。预防路面不平整造成的损害和伤亡，可以节省下来一笔资金，即使忽略这一笔资金，这依然是一笔年回报率超过18%的投资。

联邦政府的借款利率则会因此而低得多。比如说，美国10年期国债的利率从2001年起就一直在5%以下。出售国债并用获得的收入得到超过18%的收益可以立即改善美国的财政状况。任何一家私营企业可以借由这样的投资机会一跃而起。而且，这一投资机会并不少。美国有大约50%的主要道路一直未修整，意味着这些路面“破裂、坍塌、老化，需要重新铺砌”¹¹。

然而，反税收、反政府管制的言论仍主导着美国的政治话语，使这类投资机会很难进入讨论范围，即使在一个急需更多公共投资机会的经

济环境下也是如此。当然，我在这里所指的就是由2008年金融危机导致的经济衰退。

在2007年末以前的总支出水平之下，几乎人人都有工作。但是紧接着，房产泡沫开始破裂。基于虚高房价的住房贷款所带来的膨胀消费，也大幅下降。公司开始裁员，从而进一步降低了消费。消费的下降相应降低了投资，因为大多数企业的产出已经超过了人们的消费意愿。并且随着其他国家的经济陷入低潮，美国的出口量也下降了。到了2008年末，相比于历史上的“大萧条”初期，产出量和就业率下降得更为剧烈。

在20世纪30年代，世界经济陷入严重的萧条时，许多经济学家认为最好的应对政策是平衡政府财政收支。并且，由于收入的下降导致税收收入急剧减少，这意味着应减少政府支出或提高税率。赫伯特·胡佛（Herbert Hoover）总统针对“大萧条”所采取的措施正是依据这一盛行的经济学正统观念。因此，虽然这些措施没有改善萧条反而实际上使之加剧，但是以此指责他似乎不太公平。

富兰克林·罗斯福总统似乎凭直觉认识到盛行的正统观念是错的。于是他撇开正统观念，大力增加政府支出，用于公共建设工程，直接增加工作机会。但是直到1936年，在凯恩斯出版了《就业、利息和货币通论》（*The General Theory of Employment, Interest, and Money*）之后，经济学家们才得以通过一个清晰的理论框架来弄清楚为什么罗斯福的直觉是对的。¹²凯恩斯被公认为20世纪最伟大的经济学家，他解释了一个陷入深度衰退的经济体为何无法自行迅速恢复，由此赢得了这一声誉。

凯恩斯认为，消费者不会率先站出来解决问题，因为即使是那些还没有失业的人也会担心丢掉工作。企业也不会继续投资，因为它们的产出已经过剩了。最终凯恩斯得出的结论是，唯有政府才既有动机又有能力在经济低迷时期促进消费。

尽管现在大多数经济学家认同由凯恩斯提出的应对经济衰退的方法，但仍有一批有影响力的人持有不同意见，并促使自由主义者和反政

府管制的积极分子们反对一切经济刺激措施。我曾多次要求这些人解释为什么增加政府支出不会加快经济的恢复。多数情况下，他们要么不愿意回答，要么不能够做出清楚的回应，因此我真的非常感谢李·奥海宁（Lee Ohanian）为此做出的尝试。他给出的关键理由是：伴随着财政支出的增加，针对收入和支出的赋税也相应提高，从而抑制了经济活动。¹³反对经济刺激的人们所持论点可以归结为：如果政府现在借钱，增加财政支出，消费者们会意识到这将在未来带来更高的赋税，从而使他们减少当下的支出，由此带来的损失将等额抵消政府的支出，因此不会产生净刺激。

这样的论断使心理学家们将经济学家看作“智商高却什么都不知道”的人。或许有人会为了避免未来不确定的赋税而减少当前的支出，但是这样的人很少。行为经济学家们在几十年来不断表明，这并不是绝大多数人的行为方式。几乎没有消费者了解政府债务的规模，更不知道这将如何影响未来的税收。

标准的经济学模型预计人们会在工作时存够钱，以防在退休后生活水平下降。然而，事实表明，大多数人的积蓄远远不够。¹⁴即便未来可能会以宠物食品度日，也不足以令人们储蓄更多。然而，反对政府管制的积极分子却希望让人们相信，未来可能面临不确定的税务会使人们增加储蓄，以抵消增加的政府财政支出。

这是一个荒谬的论断，但是在一个由反征税、反政府管制的言论所主导的环境中，只要有一些人声称经济刺激行不通，就足以杜绝其他言论。那些人称，通过借钱来刺激经济会增加政府债务，由此导致的代价则由后代承受。由于大多数人并不想让自己的后代承受这样的代价，争论便就此止息。

事实上，审慎的投资完全不会给子孙们带来损失。恰恰相反，如果政府在经济衰退期间以4%的利息借钱，投资于能够产生18%的收益的项目，最终不仅可以使原本失业的人重新获得工作，还可以为子孙增添财

富。

在充分就业的经济环境中，相比于借钱，用税收来做这样的投资自然更好。但是反征税的言论显然排除了这种选择，即便是对能从政府投资中直接获益的人来说，也是如此。因此，《华尔街日报》记者劳伦·埃特（Lauren Etter）指出，北卡罗来纳州的居民虽然对道路质量恶化满腹牢骚，但似乎并不打算采取这种显而易见的补救方法：

7月，斯塔茨曼县（Stutsman County）的居民拒绝了一项通过增加财产税和销售税来为修路提供更多资金的举措。在斯皮里特伍德（Spiritwood）的运动员酒吧，鲍勃·鲍曼（Bob Baumann）呷着一瓶银子弹啤酒，说：“我宁愿我的孩子在砾石路上开车，也不愿意让他们背负这么高的税。”¹⁵

砾石路上有一种屡见不鲜的情况：一辆车子迸起的石子砸到后面汽车的挡风玻璃，由此导致的替换费用能达到1 000美元。这样的事情在沥青路上就极少发生，这只是修路的诸多好处之一。显然，拒绝维修路面并不是真正的节约。然而，一旦论及征税，人们就会因担心未来会征税而拒绝维修路面。

我们需要什么样的公共服务与产品

在此复述一下我的观点，我并不是说政府不存在浪费。浪费确实存在，但许多反对政府管制的激进分子笃信政府的存在弊大于利。确实，世界上的许多政府曾带来巨大的危害。即使是现在，许多国家的普通民众遭到偷窃后也不敢报警。在这些国家，腐败是稀松平常的事。

比如，针对腐败现象，总部位于柏林的非营利组织透明国际（Transparency International）定期通过调查对各国政府质量进行评估。依据该组织对“腐败”的定义——滥用政府权力谋取私人利益，该组织公

布了全球清廉指数。在调查中，受访者需要回答自己认为其所在国家的“行政人员腐败的程度”。如索马里等一些国家在“透明国际”的全球清廉指数排名中一直垫底。这类全球清廉指数持续偏低的国家，包括阿富汗、海地、汤加等，都属于世界上最贫穷的国家，这并非偶然。

然而，无论典型的政府看起来多么糟糕，令人惊奇的是，任何一个国家都有自己的政府。如果没有政府，一个国家的领土早就被拥有政府和军队的其他国家侵占了。如果政府的存在是不可避免的，那么我们就面对挑战，建立一个尽可能好的政府。

尽管一些改革者反对政府管制，但确实有一些政府相对来说更加清廉，为满足其民众对公共物品和服务的需求，也做了一些说得过去的工作。2007年“透明国际”名单中政府腐败程度并列最低的是丹麦、芬兰和新西兰。新加坡、瑞典、冰岛、荷兰、瑞士、加拿大和挪威也在前10名之列。同样不让人意外的是，这其中的绝大多数国家都属于世界上较富有的国家。

其中的因果关系是双向的。诚实、高效的政府有助于开展能够提高人均收入的活动，更加富裕的民众通常也更有能力和意愿支持更加高效的政府。但是人均收入和全球清廉指数之间的关系并不是绝对的。例如，美国的人均收入比2007年全球清廉指数排在前10位的任何一个国家都更高，但其在名单上却只排到第20位，这主要是因为美国的民众认为在美国的竞选财务制度下，国会存在腐败。

对于有着诚实、高效的政府的国家而言，人们不会认为把支持良性政府作为一项有价值的投资是荒谬的。然而对于美国的反政府人士来说，却并不是这样，他们中的许多人对政府提供的服务有着不加掩饰的轻蔑。正像罗纳德·里根常说的：“政府对经济的态度可以归结为几点：如果经济发展，就对它征税；如果经济过热，就进行管制；如果经济停滞，就出资推助。”诚实、高效的政府是以能带来职业自豪感的职业公务员制度为基础的。在一个轻蔑政府的风气盛行的环境中，招揽到有才

干且甘于奉献的公务人员显然更加困难。

然而，就算反政府人士自己来承担管理国家的职责，他们也未必能成为自己想任命的那种公务人员。不难看出，那些认为政府无用的人，会认为任用不那么在乎这些政务工作的亲信不会带来任何危害。但是正如卡特里娜飓风来袭时的经验所揭示的，任命没有能力胜任的朋友负责他们不那么关心的政府部门的工作，结果通常不好。

有理智的人都不会选择废除政府。如果必定要有政府，那么，我们便应好好考虑如何改善政府。我们想要什么样的公共产品和服务？为此我们如何更好地筹资？我们又如何找到愿意信赖的公务人员？

然而，美国在21世纪初最紧要、最现实的问题，就是融资需求超过了税收收入。生育高峰时期出生的一代必定会在接下来的20年里导致退休人数激增，这一财政缺口也会随之急剧扩大。人们自然应该不断探索新的途径，通过减少财政浪费来尽量弥合缺口。但是如果预算的削减导致沥青路变成了砾石路，那么一定是削减过头了。

反对政府管制的改革者有一个观点是对的，那就是通过消除浪费来节约资源要好过放弃有价值的事，但他们想当然地认为在市场经济中造成最大浪费的就是政府。他们常说，人们在花自己的钱时，远比国会的那帮政客要谨慎得多。

或许是这样。然而，再仔细想想，并非所有著名的政府财政浪费案例都像看上去那么夸张。例如，在本章开头提到的价值7 600美元的咖啡壶是由洛克希德·马丁公司生产的，那时候，我的一个大学同学，名叫罗纳德·杜博斯（Ronald Dubose），就在那里工作。洛克希德·马丁公司还生产了著名的价值640美元的马桶坐垫。杜博斯自视为自由主义者，他倾向于认为政府支出通常是极其浪费的。但是在一封电子邮件中，谈及这些咖啡壶和马桶坐垫的项目时，他却这样说：

那些臭名昭著的咖啡壶生产出来的时候，我就在洛克希德·马

丁公司。事实上，我当时就负责检测咖啡壶。我来描述一下情况。

首先，飞机的动力来源有两种：28伏的直流电和115伏、400赫兹的交流电。这款咖啡壶的设计需要能够在400赫兹交流电下使用，完全不同于我们平常在厨房里用的咖啡壶。这是因为使用400赫兹交流电的发动机和设备比使用60赫兹交流电的体积更小、重量更轻。由于燃料必须支持飞机在40年左右的时间中飞往各地，飞机的每一千克重量在其寿命期限内带来的消耗都是十分巨大的。因此，使用400赫兹交流电是一个明智、经济的选择。其次，C-5运输机的产量只有120架，所以开发费用在总成本中的占比微乎其微。最后，飞机上的任何设备都必须经过测试，以确保能够在军事条件下运作，并且不会起火或发生其他意外。这些测试通常包括MIL-STD-810环境测试，这其中有冲击、震动、高 / 低温测试等。

飞机需要经过硬着陆和几个小时的剧烈震动。仅仅进行这些测试就要花很多钱，要通过测试还需要其他花费。所以抱怨咖啡壶厂商其实一直是一种误解。考虑到初始花费和那么低的产量，我反倒惊讶这款咖啡壶竟然这么便宜。飞行员们本可以带保温瓶，这是大多数军用飞机必备的，但是C-5运输机是军用运输飞机，我确定研发飞机的人乐于进行这样的“投资”。

著名的马桶坐垫并不由我负责，因为这种坐垫用在C-130运输机上。然而，这并不是人们通常在哈德瓦公司（Ace Hardware）看到的坐垫，而是商用飞机上用的模塑玻璃纤维圈，覆盖着马桶，真正的坐垫固定在上面。无论如何，考虑到低量生产需要定制的模具和手工接合工艺，目前的费用已经非常合理了。

即使最臭名昭著的政府财政浪费案例是误解，然而，在其他案例中，肯定还有政府的采购人员没有争取最合适的价格。关于人性的一切事实表明，相比于花别人的钱，一个人在花自己的钱时更可能尽力搜寻

较低的价格。并且，私人领域的竞争非常激烈，这有助于防止价格高出成本太多。

个人造成的浪费远超群体

尽管大多数人在买东西的时候都很谨慎，但实际上私人领域的浪费是显著超过公共领域的，这不仅仅是因为私人领域的范围要大得多。私人领域的许多浪费并不是因为人们在一些物品或服务上花费太多，而是因为人们的一些重要支出取决于相似处境下其他人的支出水平。这又可以证实达尔文的观点，那就是个体和群体的利益通常存在着明显的冲突。

试想，你要为孩子举办一场成年派对，你希望这个特殊的场合看起来“像样”一点。但是怎么才能做得“像样”，取决于你所处的环境。2005年，纽约一位首席执行官花1 000万美元在洛克菲勒中心的彩虹厅为女儿举办了一场成年派对，他邀请了女儿的150多位朋友参加。客人们不仅在派对上欣赏到了一众知名歌星和乐队的表演，聚会结束时，每位客人回家前还拿到一袋派对礼物，里面装着价值300美元的iPod播放器，这是当时的“必需品”。¹⁶

尽管这位首席执行官在当时因为挥霍奢侈而饱受批评，但是可以想见，他身为一个父亲的动机是好的。他可能只是想通过举办派对来向世界宣告自己对女儿的爱，使之成为女儿难以忘怀的独特回忆。但是“独特”只是一个相对概念。由于他所处的金融圈内的其他人也会花重金为孩子举办特别的派对，因此在他所处的环境下，“独特”的标准是非常高的。这一标准，与30多年前相比，自然已高出许多，原因很简单，现在有钱人的家庭要比那时富裕得多。

中等收入的消费者和富裕的企业高管们所处的社交圈在很大程度上没有交集，但是对于中等收入的家庭来说，“独特”的标准也在逐渐提

高。例如，我们可以考虑玛丽·埃伦·菲洛（Mary Ellen Fillo）对康涅狄格州这类家庭近年来发展趋势的描述：

苏珊·里尔登（Susan Reardon）是利奇菲尔德县一个快满16岁的女孩儿格蕾丝的母亲，她计划花12 000美元为女儿举办一个生日派对，她说：“我读到一篇关于青少年派对的文章，标题是《难道你不想拥有一个跟我一样酷的派对吗》，这足以说明一切。”派对计划包括为客人们设计特别的邀请函，在后院搭帐篷，准备铺上亚麻布的圆桌、一个舞池、一支乐队，供应冰激凌圣代，还要播放庆祝的音乐。¹⁷

所有的家庭，无论收入高低，为亲人庆祝特别日子的花费都在逐年增加。2009年，美国一场婚礼的平均花费是28 082美元。¹⁸而在1980年，相应的花费按2009年的美元汇率换算，则为11 213美元。¹⁹这些额外支出的累积效应，在很大程度上只是进一步抬高了“独特”的标准。这些活动的花费最终都远超以往，却没人觉得它们比以往更“独特”。

收入差距与支出瀑布

自由主义者所笃信的亚当·斯密的“看不见的手”理论，是建立在消费行为从根本上来说独立于环境的假设上的。

然而，环境是具有决定性的，这种情况下，导致人们花钱的因素通常会造成极大的浪费。例如，最近几十年来，随着美国企业高管薪酬的爆炸式增长，他们的豪宅越建越大，早就超过了限度，继续扩大也不会再增加效用。多数的高管想要或者需要建更大的房子，仅仅是因为“大”的标准改变了。

这一转变背后的原因是收入增长模式发生了根本的变化。第二次世界大战之后的30年里，美国所有家庭的收入，无论高低，都以不到3%

的速率增长。但是就在1970年前后，这一模式发生了变化。从此以后，家庭收入增长率显著下降，但不包括位于收入分布顶端的家庭，这些家庭收入正以几乎有史以来最高的速率增长。例如，在1980年，美国最大的几家企业的首席执行官的收入大约是一名普通员工收入的40倍，如今却超过了400倍。无论如何分析数据，结论都是相同的。人们的收入水平越高，其收入的增长就越快。²⁰

这一转变催生了一种我称之为“支出瀑布”（expenditure cascades）的现象。²¹收入最高的人仅仅是因为收入更高就要建更大的豪宅。中产阶级对此却并不介意，恰恰相反，许多中产阶级似乎乐于欣赏展示富人和名人生活的图片和电视节目。但是，富人建造的更大的豪宅改变了其他接近富人阶层的人的住房标准，因为他们出入的许多社交圈是相同的。可能在这个圈子里早已约定俗成，要举办36人的晚宴，而不是24人的；或者要在家中举办婚礼，而不是在酒店或俱乐部举办婚礼。所以这些接近富人阶层的人的住房也越建越大，从而改变了其他仅次于他们的人的相对参考框架，由此向收入越来越低的人依次传递。到了2007年，美国新建的独户住房面积的中值超过了214平方米，与1970年相比增长了50%。²²

如今在美国，中等收入的人并不会因为收入多了就建大房子，或者花更多的钱庆祝特殊场合。实际上在过去的30年里，中等收入家庭的收入增长并不显著。他们之所以建更大的房子、花更多的钱，是因为其他跟他们一样的人也在建更大的房子、花更多钱，这是由正在发生变化的收入增长模式导致的。类似的“支出瀑布”不仅发生在住房和庆祝特殊场合的派对上，在其他许多领域也是如此。

当然，并非所有额外支出都是毫无价值的，例如，一颗更大、更纯净的钻石所折射出的光，就连荒岛上的独居者也会觉得赏心悦目。但是这些额外支出中的大部分都被浪费了。一旦住房面积超过一定的限度，对于更大的空间的需求就不再来自其所带来的内在效用，而是几乎只取

决于社会因素。住房面积增大还会带来许多额外的麻烦。人们至少要雇用和管理更多的人来看护房屋，这就增加了心生不满的雇员四处造谣的风险。还有很多诸如此类的麻烦。如果最大的那些豪宅都更小一些，拥有这些豪宅的人其实会更幸福一些。

此外，还要考虑到一个事实，那就是消费者总支出比政府支出的两倍还要多，由此，对于私人领域的浪费超过公共领域这一论断，任何理性的人都不会觉得不可置信了。

如何解决资金短缺的问题

无论何时，重要的问题都不是私人领域的浪费是否超过公共领域，而是我们应该寻找更多的资源来解决最为紧迫的资金短缺问题。消除政府财政浪费的努力一直没有停止，但是经验表明，这并不会带来突如其来的重大突破。饿死政府这头“野兽”的结果似乎弊大于利。收入增长方式的转变改变了人们必要的支出，这是另一头“野兽”。为了强调环境对重要支出决策的影响，我称其为“地位消费野兽”，它比以往任何时候的政府都更加贪婪。

人们目前才刚刚开始寻找消除私人领域浪费的方法。我将会在下一章讲一个好消息，那就是消除私人领域的浪费相较于公共领域而言会更加容易。事实表明，通过使人们的动机发生简单的、不那么剧烈的改变，就可以饿死“地位消费”这头野兽。

THE DARWIN ECONOMY

THE DARWIN ECONOMY

05

控制“地位消费”：
得不偿失的军备竞赛

根据达尔文的观点，个人动机与群体动机之间经常存在激烈的冲突，这在军备竞赛中体现得最为明显。但是，一直以来，尽管我们对军备竞赛这一比喻早已习以为常，却并没有完全理解它。

无人获益的军备竞赛

简单来说，军备竞赛就是这样：一个国家通过扩充武器装备可以在与另一个国家的竞争中胜出，这会促使竞争对手也扩充武器，以维持平衡。接着，前一个国家进一步扩充武器，继而导致竞争对手再一次回应，如此往复。最终，双方只是在武器装备上花费了大量资金，没有任何一方的安全性得以提高。

假定两个竞争对手的军备水平起初相差无几，那么想必大家都能认可的是，军备竞赛这一过程从根本上来说是一种浪费。如果两个国家都减少在军备上的花费，增加在学校、住房、医院、道路和其他非军事物品上的支出，对双方来说都要好得多。

对此，基本的解决方法是签订军事管制协议，双方承诺削减军备开支。缺乏信任可能是达成协议的最大障碍，而成功签订的协议几乎都会赋予双方互相检查的自由。正如罗纳德·里根所说：“信任，但要查证。”

到目前为止，这些说法都很合理。但是仍有一个尚未解答的重要问题：准确地说，在什么情况下可以认定发生了军备竞赛？如果有人说有什么东西过量了，言外之意，肯定是有其他东西存在短缺。在军备竞赛的例子中，过量的是武器，不足的则是日常消费品。但是为什么由此导致的失衡不是反过来的？为什么各个国家互相之间的角逐，不是通过减

少在军备上的支出，来修建更多更好的医院和道路呢？

这个问题的答案很简单，但颇有启发性。尽管一个国家在军备和公共设施这两个领域里的发展都取决于其所处的环境，但环境对军备的影响更甚。在两个国家的案例中，一个国家的军备是否充足几乎完全取决于另一个国家的军备情况。如果一个国家的军备更充足，另一个国家的主权就会面临危险，前者就是安全的。

环境会影响消费品，但是程度远不如对军事装备的影响。比如说，如果消费者了解到敌对国家的公民有更好的电视和更豪华的汽车，他们或许会心生不快，但是这些不快所造成的损失远远比不上失去政治主权。

之所以会发生军备竞赛，正是因为存在这样的不对等。现在我想做一个假设，那就是环境对消费品的影响比对军事装备更大。一个国家可能会通过制造更多的炸弹来超越敌对国家，但是由此带来的直接后果是减少了在面包机上的开销。如果面包机落后于对手国家的代价比在武器装备上落后更大，那么这个国家从一开始就不会倾向于制造过多的炸弹了。相反，两国之间将会进行一场受相对水平影响的、双方致力于制造更多面包机的“军备竞赛”。每一个国家在面包机上的花费会越来越多，以努力超过竞争对手，并在这一过程中不断抽走用在军事装备上的资源。这种情形之所以没有发生，完全只是因为环境对军事装备的影响超过了对消费品的影响。

关于军备竞赛还有其他几个需要注意的要点。其中之一便是军备竞赛并不是出自非理性行为。在许多国家看来，把维持与敌对国家同等的军备力量放在优先位置，是完全合理的。能否维持主权取决于一个国家自我保卫的能力，军备不足则会削弱这一能力。

军备竞赛也并非源自竞争的缺失。相反，一直以来，军备竞赛反而使各国因抢夺重要资源而引发最直接的竞争。竞争越激烈就越容易引发军备竞赛。

军事管制协议也没有什么稀奇。自由主义者如果因为这限制了签署方制造炸弹的自由而反对军事管制协议的话，就太愚蠢了。签署方的意图正是限制自己制造炸弹的自由！各国乐于签署这样的协议，是因为他们知道，如果他们保有自主权，最终在炸弹上的花费就会过度。

到目前为止，有关军备竞赛特点的叙述完全没有争议，但这并不意味着军事管制协议不会有争议。比如说，如果没有切实可行的办法防范对方的欺骗行为，那么一个国家的领导者拒绝签订协议或许才是理性的。

如果领导者们相信，军备竞赛最终会对他们有利，那么拒绝签订军事管制协议同样可能是理性的。例如，据说里根曾在20世纪80年代支持建设战略导弹防御系统，部分原因在于他相信这会迫使苏联也采取同样的行动，从而加速这一美国首要竞争对手的经济崩溃。

总而言之，势均力敌的国家之间开展军备竞赛是一种浪费，共同签订限制军备支出的协议对各方都有利，这几乎是全世界的共识。如果哪个国家不愿签订协议，通常是出于现实层面的担心，比如无法切实有效地查证或者资源分布不均衡，而并非站在抽象的哲学角度——担心丧失独立行动的自由。

经济领域的“军备竞赛”

在军事竞争领域是这样，在经济竞争领域也是如此。例如，在体育运动中，无论是个体还是团队的成绩都在很大程度上取决于相对表现。如果一名短跑运动员服用激素，他拿到奖牌的概率就会提高。如果一名出色的选手加入一个团队，也会提高整个队伍获胜的概率。这些举动也有着明显的弊端。激素会使服用者面临严重的健康风险，而聘请明星运动员的花费则非常昂贵。但是，不论是从个人角度，还是团队的角度看，都是利大于弊的。问题在于，如果所有的短跑运动员都服用激素，

所有的队伍都花费高昂的代价争抢明星运动员，那么将不会有任何人、任何一方获得竞争优势。

与军事竞争领域一样，运动领域的解决方法也是签订“地位军备”控制协议。有些国家的管理机构会进行随机的药物测试，并严厉处罚违规的人。其他社团联盟则对薪水设置上限以限制工资总支出，赛车协会也会对发动机排量施加限制，诸如此类。所有这些限制都是为了解决类似军备竞赛的问题。除了自由主义者之外，并没有多少人觉得这些举措严重侵犯了个体的权利。

阻碍协议签订的现实因素，也同样阻碍了协议的实施。例如，已故的纽约洋基队前经理乔治·施泰因布伦纳（George M. Steinbrenner）曾成功阻止其他棒球队老板在美国职业棒球大联盟中获取更大的利益分成，并设定更为严格的收入限制。施泰因布伦纳可以从当地有线电视的收入中分享可观的收益，他知道自己可以借此给明星职业运动员更高的报价。他还知道，尽管阻止收入设限和收益分成会增加包括他的队伍在内的所有队伍的工资支出，但是拥有一支持续获胜的队伍远不止可以弥补这一劣势。

施泰因布伦纳反对设定工资限制和收益分成，这增加了其他队伍老板的成本，最终这些成本将由球迷承担。按照经济学家的说法，这些政策造成了“地位外部性”（positional externality）。

以上论述是人们熟悉且没有争议的。令人感到陌生的，则是相似的外部性可以使日常消费模式产生相似的浪费。在第4章，我曾论述过迫使人们在住房和特殊场合的庆祝上过度消费的因素。这一论断中暗含着一个推论，那就是他们在别的事物上的花费必然过少。那么这些事物又是什么呢？人们又为什么在这些事物上花费过低呢？决定军备竞赛的逻辑为这些问题提供了系统的解释。

人们的价值判断会受到情境的影响

通过以下一些简单的思想实验，我们很容易从直觉上理解在不同情形下环境的重要性。在每一个实验中，你需要在两个相似的情形中做出选择，两种情境除了文中指出的不同外，其他方面都是一样的。而且实验假定两种情境会一直持续下去。

第一个思想实验探究的是环境对住房估值的重要性：

以下两种情形你会选择哪一个？

情形A：你居住地区的住宅面积为557平方米，其他人居住地的住宅面积则有743平方米；

情形B：你居住地区的住宅面积为372平方米，其他人居住地的住宅面积只有279平方米。

即便是按照美国如今的标准来看，4种住房面积都算是大的。不过，不要妄想你在情形A中可以通过努力工作搬到房子更大的住宅区，选择A意味着永远住在相对更小的房子里。

根据亚当·斯密的“看不见的手”理论所蕴含的标准经济模型，要做出这一选择是不需要多想的。在那些模型中，环境完全不重要。你对住房的评价将仅取决于其绝对面积，由于在情形A中的住宅面积更大，所以你自然会把情形A当作更好的选择。然而在人们实际做出选择时，却并不是这么不假思索。通常，他们会花费一些时间仔细考虑一番，最终多数人选择了情形B。

现在考虑第二个思想实验：

以下两种情形你会选择哪一个？

情形A：在一年里，你死在工作岗位上的概率是万分之二，其他人则只有万分之一。

情形B：在一年里，你死在工作岗位上的概率是万分之四，其他人则为万分之八。

与第一个思想实验类似，你同样需要在绝对优势和相对优势之间做出选择。如果你选择情形A，那么你死在工作岗位上的概率只有情形B下的一半。但是你的工作在情形A下是更危险的，而在情形B中却是相对更安全的。

我曾对不计其数的人进行过这一思想实验，询问对象包括大学新生和退休人员等各种类型的人。几乎所有人都选择了情形A。与第一个思想实验不同的是，相比于相对优势，参与第二个思想实验的人一致偏好绝对优势。

我们能从这些选择模式中知道些什么呢？在第一个思想实验中选择情形B的人给出的理由大多是，尽管情形B中房子面积的绝对数值比情形A中更小，但是他们觉得自己在情形B下的满意度会更高。房子的绝对大小当然也很重要，但是一旦超过一定限度，环境对大多数人的影响就会更大。

参加第二个思想实验的人同样给出了他们的理由，他们当然不希望像情形A那样，自己的工作危险性是其他人的两倍，但是相比于情形B，他们宁愿接受这一现实，因为他们在情形B下工作的危险性是情形A下的两倍。

借用已故经济学家弗雷德·赫希（Fred Hirsch）的说法，两个思想实验所呈现的选择模式表明住房是一种“地位物品”，而工作场合的安全性则是“非地位物品”。¹地位物品的价值对环境极度敏感。现实表明几乎所有的价值判断都会受到环境的影响，但地位物品相对来说受到的影响更大，非地位物品受环境的影响则相对较小。

正如我们从军备竞赛的例子中所看到的，当环境在一些领域的影响更大时，就会导致浪费性的扭曲结果。各个国家之所以制造过量的炸

弹，是因为它们的武器储备是否充足取决于它们之间的对比情况。环境对消费品也有影响，但远远比不上其对武器装备的影响。

相同的逻辑意味着，人们对地位物品的消费会发生扭曲的过剩。在两个思想实验中，扭曲所涉及物品是住房和工作安全性，对此，我们来补充一些细节。两种物品都是人们想要的，也都是需要付出代价的。我们曾在第3章讨论过，在标准的“看不见的手”理论看来，在竞争市场中，人们会选择安全性和住房条件的最优组合。极度厌恶风险的人可以选择更安全的工作，只不过也要接受更低的工资（因为给台锯安装防护套是需要花钱的），同时也意味着不得不减少住房上的花费。不那么在乎风险的人则可以选择更危险但工资更高的工作，从而提高住房上的花费。

自由主义者和热衷于“看不见的手”理论的人辩解称，如果政府对工作场所的安全性进行管制，就会迫使人们增加对安全性的支出，比他们原本愿意支出的要多。例如，试想某个人正考虑是否接受更危险的工作，以赚取更高的收入，购买更贵的房子。依据标准的“看不见的手”理论，如果增加住房面积带来的价值超过了安全性降低造成的代价，这个人就会接受这样的工作。但是，只有增加的住房面积的价值超过可以恰好补偿降低的安全性这一平衡点时，他才会接受风险的提高。“看不见的手”理论的支持者认为，如果强迫他承担增加在安全保障上的支出，那么买到的安全的价值也会比不上他付出的代价。

然而，如果相比于安全，住房的价值更易受环境影响，那么上述观点就不再成立了。因为如果环境会产生影响的话，那么就会有两个全新的因素进入成本-收益的权衡中。从事更危险的工作不仅可以使人住上更大的房子，而且可以住上与别人相比更大的房子；另外，无论是从绝对水平上讲还是相对水平上讲，这同样也意味着安全性降低。

之所以要考虑环境，是因为决策中需要考虑的两个新的因素是极不对等的。只要可以让你的孩子进入更好的学校，相对更贵的房子就非常

重要。相比之下，工作安全性的降低只是次要的。因此，在传统理论解释中加入这两个相对概念，会使最终的决定倾向于接受危险性更高的工作。但是这会产生问题，因为人们追求的这些相对优势是会相互抵消的，无论有多少人希望通过接受更危险的工作来取得相对优势，也只能有一半的人可以最终获得这种优势。

这一简单的事实是对传统的“看不见的手”理论的一记重击。即使所有的决策者拥有充分的信息并且完全理性，即使所有的劳动力和商品市场都是充分竞争的，也并不意味着市场中“看不见的手”会导致对整个社会最有益的结果。不断有事实表明，如果住房是地位物品，安全是非地位物品，不加管制的市场中就会出现过大的房子和过于危险的工作。

如果人们投票选出的立法者，实施的正是限制工作场所安全的制度，那么包括自由主义者在内的一些人就没有理由抱怨这些制度剥夺了人们自主选择承担多大风险的权利。如果虚幻的相对收益是驱使个体选择额外风险的首要因素，那么这一选择或许最好由集体做出。所以，抱怨这些规定限制了个体自由就如同抱怨戴头盔的规定限制了冰球运动员的个体自由。确实如此，但这也正是关键所在。

当然，市场结果不够完美并不意味着政府干预就必然可以导致更好的结果。政府也有其不完善之处。重要的是，管制最终是否奏效只能取决于管理政策是否有效。口头上喊“个体自由至上”是不能解决问题的。

军备竞赛带来的教训是，人们消费了过多的地位物品，而非地位物品却不足。哪些物品是地位物品，哪些又是非地位物品呢？达尔文的理论框架有助于我们思考这一问题。衡量成功的首要标准，就是个体把基因传递给下一代的能力。因此，某一类物品的增加越能提高个体繁衍的成功率，就越具有地位物品的属性。

例如，雄性麋鹿的鹿角之所以成为优势，是因为鹿角可以让雄性麋鹿赢得与雌性麋鹿的交配权，而这几乎完全取决于鹿角的相对大小。鹿角更大的代价是在茂密的树林里行动不便，会使雄性麋鹿更容易被狼或

其他捕食动物捕杀。但是，纯粹从达尔文的观点来看的话，如果行动灵活需要以鹿角相对较小为代价，灵活性是没有意义的，因为这种雄性麋鹿的基因不能传递下去。因此，对雄性麋鹿来说，鹿角是一种地位物品，而灵活性则是一种非地位物品。正是因此，雄性麋鹿群体过度追求大鹿角而忽视了灵活性。

上述达尔文的观点也为衡量其他物品提供了参考。例如，花时间赚更多钱是否比把相同的时间用于休闲更具有地位性。就像前面提到的，在人类神经系统的进化过程中，饥荒频频发生，尽管食物总是充足的，但也只有那些收入相对更高的人才能吃饱。相比于有更强烈的意愿赚更多钱的人，愿意花更多时间休闲的人更容易挨饿。因此，在达尔文的理论框架下，收入比休闲更具有地位性。

这与以下思想实验中体现的行为模式是一致的，这个思想实验类似于我们先前讨论过的两个思想实验：

以下两种情形你会选择哪一个？

情形A：你每年有4周假期，其他人则有6周；

情形B：你每年有2周假期，其他人只有1周。

面对这些选择，多数被试选择了情形A，表示相比于相对优势，他们更偏好绝对优势。接下来的讨论表明，多数被试不希望自己的假期比别人短，但是相比于放弃一半的假期，他们宁愿接受这一现实。

相同的逻辑也适用于其他非物质性消费，比如，避免噪声和污染，应该不如因此减少的收入更具有地位性。²同样，申诉程序、工作丰富度和舒适的条件等工作场所中的便利则不如因此减少的工资更具有地位性。

相比之下，达尔文认为，在抚养后代上的投入是高度地位性的。如

汽车、衣服和珠宝等容易观察到的消费物品，应该比保险等观察不到的物品更具有地位性。

储蓄也不容易被观察到，因此也可以认为是不具有地位性的。然而，与此相反，如果今天的储蓄减少了，意味着未来可以消费的地位性物品也会减少。但是人们普遍倾向于低估未来的成本和收益，反过来又大大削减了成本和收益。今天储蓄减少当然会造成以后可消费的地位物品减少，但是今天的减少是立刻可以感受到的，而未来的减少只存在于想象中。更重要的是，很多生活早期的支出，比如趁孩子还小的时候买学区房，本质上比更晚一些的支出更具有地位性。年轻一些的父母可能会自我安慰：“现在就把孩子送进我们能力范围内最好的学校，退休的事等以后再操心吧。”总而言之，储蓄是非地位性的。³

公共产品出于其所具有的基本属性也是非地位性的。私人产品对不同的人来说数量和质量都会不一样，不同的是，公共产品的数量和质量对所有人来说都是一样的。因此，人们无法通过公共产品获得相对优势。

以上推论与众多研究的实证结果在很大程度上是一致的⁴，与各国的管理、税收和公共支出模式也是一致的。

如何让管制发挥有利的作用

管制可以产生实际影响。反对政府管制的积极分子坚持认为管制是不好的，自然也有很多监管不善的例子。但是我们不能忽略相反的情形，也就是很多管制政策也会带来一些好处。从这个观点出发，研究议员们会通过民主选举选择什么样的管制政策，我们就可以知道如何让管制发挥好的作用。

例如，美国的大多数地区对儿童入学都有这样的规定：如果在秋季学期的某个日期前满5岁，当年就要上幼儿园。这项规定的意图何在

呢？如果父母可以自由选择孩子的入学时间，那么他们都会觉得延迟一年入学会更有利。这样，孩子入学时的年龄更大、更聪明、更强壮，在情绪上比同学更成熟，在学校的表现会更出色，更容易赢得运动比赛，也更有可能考入优异的大学。但是一旦有一些父母通过让孩子延迟入学来积累优势，其他父母也会迫于压力这么做。最终，所有的儿童都晚一年入学，但是能进入好大学的人数并没有因此增加。基于这一情形，也就不难理解人们为什么会希望他们选举出的议员强制规定入学时间了。

多数国家同样也会对工作场所的安全性进行管制。采取的措施基本上要么是促使人们增加储蓄，要么是提供补贴，因为许多人在退休时没有足够的储蓄。许多国家制订计划，致力于促进不同地区之间的教育均衡。多数国家会想办法缩短工作时间，其中有许多设定了工作假期的最低要求。

这些管制恰好印证了我的猜想：私人消费模式在某些领域更容易受到环境的影响，价值判断会由此发生扭曲。而且我在别处也曾论证过，除此之外的其他解释多数是自相矛盾的⁵，它们多数也存在机会没有被充分利用的问题，这一点我在第3章曾经讨论过。如果市场像看起来那样具有竞争性，而且所有领域的价值评估对环境同样敏感，那么多数的管制政策就基本没有必要了，因为“看不见的手”会解决一切问题。

累进消费税：限制个人浪费的重要途径

还好实施管制政策并不是抑制过度消费的唯一途径。如果问题在于人们对地位物品花费过多，而对非地位物品花费不足，那么要纠正失衡，最温和的办法就是改变相对定价。如果信息是充分的，政府也是完善的，我们只需要根据环境对每种物品影响的程度设定不同的税率。最具有地位性的物品税率最高，比较具有地位性的物品税率其次，以此类推。

尽管一些研究者已经开始对某类物品影响的差异进行估算⁶，但是现有的知识太支离破碎，这样的研究根本没法进行。哪怕我们再了解这些数字，为每一种物品设定不同的税率在政治上也要花费很大的代价。立法者们便会被说客包围，这些说客会标榜自己生产的产品如何不具有地位性，因而应被免除税负。

在我以前写的书中，我曾主张采取一种更简单、更有效的方法，也就是取消现行的累进所得税，转而征收高得多的累进消费税⁷。这种方法是基于奢侈品比必需品的地位性更强这一事实而得出的。区分哪些是奢侈品显然存在一些困难，但是，考虑到奢侈品具有依赖环境的内在属性，我们至少可以确定，消费最高的那些人最可能把最后一美元花在奢侈品上。因此，急剧累进的消费税本身就是一种奢侈品税，不必考虑如何界定哪些物品是奢侈品。

实施累进消费税政策的方式简单明了。⁸同现在一样，纳税人要向税务机构申报自己的收入。他们还需像现在的个人退休金账户和其他免税退休账户所要求的那样申报一年里的储蓄。最终，人们需要对他们的“应税消费”缴税，应税消费也就是收入减去年度储蓄，再减去标准免税额。⁹边际税率会随应税消费的增加而上升。如果税收是收入中性的，顶层的边际税率会显著高于当前针对收入的边际税率，以弥补因减免对储蓄征税而减少的税收。但是，如果美国政府想要修复破败的基础设施，处理苏联时期遗留下来的核武器，在生育高峰期出生的一代人退休的时候恢复收支平衡，那么就需要更多的税收，而这需要进一步增长最高段的边际税率。

通过提升边际税率来增加税收收入的提议，总会引起人们的顾虑，认为提高税率可能会抑制储蓄和投资。实际上恰恰相反，累进消费税会促进人们增加消费和储蓄，尽管针对较高消费的边际税率是非常高的。

如果这种税收政策可以直接促进高消费人群的储蓄，那么它也会间接影响其他人的消费。毕竟，每个人的消费都构成了其他人消费的参照

系。考虑到环境的重要性，累进消费税的间接影响肯定比它的直接影响还要大。

我在第4章讨论过的支出瀑布促使美国许多中等收入家庭过度消费。如果收入最高的人增加储蓄并减少在住房上的消费，这会改变收入仅次于他们的人在住房支出上的参照系。所以，这些人也会减少住房消费。以此类推。高收入的人也会减少在庆祝特殊场合上的花费、在礼物上的花费和在珠宝上的花费。这些变化同样带来类似的间接影响。

所有事实表明，这其实是件好事。在2005年至2006年间，美国家庭总储蓄率是负数。实际上，这是美国人在“大萧条”之后的几年里首次出现入不敷出的情况。2008年的金融危机击垮了全球的经济，而低储蓄率则加剧了金融危机。自由派和保守派在数十年里一致认为，如果美国人都减少消费，同时增加投资和储蓄，那么所有人的境况都会变好。要想增加储蓄，人们必须共同行动。累进消费税可以说是实现这一目标的理想工具。

累进消费税可在增加税收的同时，抑制人们给他人造成损害的行为。花1 000万美元为女儿举办成年聚会的首席执行官可能本不想伤害别人，然而他的行为却造成了伤害。他的行为通过改变界定特殊场合的参照系，使他人面对艰难的选择：要么增加支出，要么保持支出水平但显得他们并不重视这一场合。如今，许多中等收入家庭的十几岁的孩子都希望父母雇小丑或者魔术师在自己的生日派对上表演。然而现实表明，派对支出增加后，孩子们并没有变得更快乐，这时候如果还坚持认为参照系的改变不会对其他人造成损害，可就有点儿奇怪了。

你在下一章中可以看到，人们通过避免遭受他人行为的伤害而获得的收益，与限制其他人的行动自由所需的代价，两者孰轻孰重，我们必须做出权衡。累进消费税的作用正在于此。例如，设想一个每年消费1 000万美元的家庭，正在考虑增加200万美元的住房花费。如果最高边际税率是100%，那么这项支出计划将花费400万美元。如果该家庭实施这

项计划，其支付的税可以减少政府200万美元的财政赤字。

该家庭也可以缩减支出，只在住房上增加100万美元的花费。此时需要支付的税为100万美元，同时还增加了200万美元的储蓄。政府赤字会因此下降100万美元，额外的储蓄也会刺激投资、促进经济增长。无论哪一种选择，都不会对富裕家庭造成实质损害，因为如果所有富裕家庭都采用最初的支出计划，由此导致的结果仅仅是重新确定房屋的合理定价。

重要的是，通过极高的消费获得的效用几乎完全是地位性的。能够促进高收入者增加储蓄和减少消费的税费，对于这些人几乎没有直接的负面影响。恰恰相反，正如我在第4章提到过的，如果对所有人来说最贵宅邸的价值涨幅变小，很可能让高收入者更开心。管理昂贵的住宅会很麻烦。

自由主义者最应积极支持税收政策，而不是支持用规范性管制来抑制伤害别人的行为。征税并不阻止人们做他们想做的事，只是增加了人们做这些事的成本。采用征税方法来抑制地位物品的外部效应的优势就类似于通过排污税来抑制环境污染的优势。

在环保领域，对一些企业来说，减少污染的成本是所有成本中最高的，但是即使要交排污税，继续排放污水仍是有利的。与此类似的是，对于一些家庭来说，降低消费非常困难，因而即使税率升高，这些家庭仍可能通过额外的支出来维持原有的消费水平。但无论是在环保领域还是在消费领域，我们想要阻止的伤害更大程度上取决于整体行为，而不是某个个体或者某个企业的行为。正如征收排污税可以使多数企业减少污染，从而减轻污染带来的损害，累进消费税也可以通过促进人们进行储蓄而减轻地位物品的外部效应带来的花费。

尽管当前对地位性的认知还不充分，最新的证据为理解急剧累进的消费税如何影响社会福利带来了一些启发。与现在的所得税相比，累进税制可以减少高端消费而增加公共消费。如果富裕的消费者总体上来说

是理性的，那么削减私人消费应该从对他们来说最不急需的物品开始。尽管政策存在缺陷，政府至少也会将一部分税收用于提供选民支持的公共服务。那么，现实问题在于，削减最不急需的私人高端消费品支出所造成的损失，是否高于增加对人们最重要的公共服务所带来的收益。诚然，如果一个社会的私人消费极低，而公共物品消费又极高，那么实施累进消费税肯定会减少社会福利。那么对于私人消费很高且公共服务水平很低的社会来说呢？近几年来美国的情形便是如此。

在缺乏详尽事实证据的情况下，可以保守推测，面对急剧累进的消费税，高端消费者率先削减的支出恰好就是他们在收入增长时增加得最多的支出。正如我在第4章讨论过的，近年来增长最显著的高端消费是住房和特殊场合的庆祝活动。这些消费都是具有高度地位性的。

在高边际消费税率之下，最富有的家庭一定会减少为孩子举办派对的花费。如果他们这么做了，那么定义他们所处的圈子中特殊场合的标准就会相应发生变化。谁会坚持认为这些变化会给孩子们带来重大损失呢？

急剧累进的消费税增加了财政收入，这对公共服务的影响会是怎么样的呢？即使增加的收入中会有一部分被浪费，其余多数也会用在有益的事情上，比如，可以修理35号州际公路上那座在2007年8月1日坍塌的危桥，这座桥位于明尼阿波利斯市中心，坠入了密西西比河，导致13人身亡，145人受伤；或者可以检查那些目前未经检查就可入港的货物集装箱。

虽然证据有限，但是我们总可以推测特定的税收政策的变化可能对福利造成的影响。想必多数经济学家都会认同的是，住房变小（因而更便宜）给富裕家庭带来的损失，要低于安全性升高（如使用安全性更高的公路）带来的收益。如此，安全性提升带来的好处对非富裕家庭来说纯粹只是意外之财。

传统观点认为，税收政策使人们面临公平与效率之间的艰难抉择，

上述结论为反驳这一观点提供了依据，我将会在第10章进行进一步论述。如果正像我观察到的那样，环境塑造了人们的支出模式，那么无论是从公平的角度讲，还是从当前偏低的经济效率来看，提高针对高收入者的边际税率都是正当的。

我们为什么必须纳税

我从好多年前就开始主张实施累进消费税政策。当我对我的自由主义者朋友们说起的时候，他们通常觉得这是一个有用的办法，但是紧接着他们又会觉得在政治上不可行。主宰美国政治进程的强势的反税收、反政府管制力量使人们很容易受到这种消极态度的影响。但是，不进行税收改革的话，迟早会面临问题。继续借款和缩减必要的公共服务支出可以维持一阵子，却不能永远维持下去。等需要做出权衡的时候，累进消费税就成了可行的选择。

实际上，美国参议院早在1995年就在佐治亚州民主党参议员萨姆·纳恩（Sam Nunn）和新墨西哥州共和党参议员皮特·多梅尼西（Pete Domenici）的支持下提出了这一税种。这一提案因没有争取到预算而没能排在议程首位，但是自此之后的几年里，也不断有其他人宣扬类似的提案。我在1997年发表的文章¹⁰中曾提倡采用累进消费税，之后不久，我收到一封来自米尔顿·弗里德曼（Milton Friedman）的来信，他直到2006年去世时都是美国小政府保守主义理念的拥护者。我发表文章时，美国的财政几乎有巨额的盈余，因此弗里德曼在信中写道，他并不同意我提出的政府应该筹集更多资金的观点。但是他也补充道，如果政府确实需要额外的资金，累进消费税是最好的办法。他附加了一份他在1943年的《美国经济评论》上所发表文章的复印件，在这篇文章中，他声称这种税是为战争筹资的最佳途径。¹¹

保守党中的许多其他人提倡固定税或增值税，这些税本质上来说属于全国性的销售税。但是，由于富人的储蓄占收入的份额相比于穷人更

大，这些税是极具累进性的。对一切消费征收相同比率的税实际上会增加这个国家最富有的人的可支配收入，从而使“支出瀑布”现象更加严重。引发那些“支出瀑布”现象的就是富人增加消费的行为。若要将资源导向更有生产力的用途，对高消费者征收较高的边际消费税就是必要的。累进消费税可以做到这一点，增值税却不能。

经济学家劳伦斯·塞德曼（Laurence Seidman）提出，要实行累进消费税，第一步最好是在维持当前所得税的同时把累进消费税当作一种补充性的附加税。¹²在他的提议下，调整后总收入不足100万美元的家庭可以完全免缴附加税。而那些收入更高的家庭则需汇报他们每年的储蓄，以年收入减去储蓄的差额作为年消费支出。因而，累进的附加税只是对消费中超出50万美元的部分征收的。¹³最终，通过逐渐降低征收附加税的门槛，累进消费税会彻底取代所得税。

这种方法可以暂时刺激支出增长，给疲软的经济带来切实的效用，这使塞德曼的提议更有吸引力。关键一步是，要马上批准附加税，但是延迟至经济状况接近充分就业时再实施。在税收政策生效之前，富裕家庭忙于扩建住宅和大肆举办派对，仅仅宣布税收政策会掀起额外的高端消费浪潮，这肯定不是刺激额外支出的最佳方法。但是，相比于放任总支出低迷，远远无法达到充分就业状态，这种方法显然好一些。

长期来看，累进消费税会逐渐改变最终的支出结构，使之由消费转移至投资，从而逐步提高生产力。如果经济下行，免税或减税会成为有力的刺激手段，因为人们只有理性增加消费才会从中获益。与此相反，临时削减所得税只能起到微弱的刺激效果，因为心怀恐惧的消费者通常会把钱存起来。

自由主义者自然可以像指责其他税收一样指责累进消费税：“这是小偷的行为！这是社会工程！弗兰克好像以为华盛顿的那帮政客比我们还懂怎么好好花国家的钱！这是阶级斗争！”

成熟的人能够理解为什么必须征税。目前，美国对储蓄征税，这会

减少生产性投资。美国还对工资收税，这会减少就业。为什么不对过量的事物收税呢？

THE DARWIN ECONOMY

THE DARWIN ECONOMY

06

深度理解竞争：
完全竞争中的加害者与受害者



几乎在任何一个欧洲国家，个人行为都比在美国要受到更多管制。如今在美国，个人行为也受到比一个世纪前更多的管制。这两组比较结果的根本原因都在于人口密度的差异。欧洲的人口密度一直远大于美国，美国当今的人口密度也比过去要大。人口密度越大，人与人之间发生冲突也就越频繁。

社会并不仅仅是人口密度在增大，人与人之间的相互影响也在加深，而相互影响加深的部分原因恰恰在于人口密度增大。这种趋势还会持续下去，因此人们对于管制的需求会越来越大。一个社会越是能够以明智的方式应对这种需求，就越可能繁荣昌盛。

但是，除非政府能理解使规则成为必要的市场失灵现象，否则就不可能采取最佳补救措施来对市场失灵加以纠正。自由派人士一直认为，市场失灵的根本原因在于自然或人为的阻碍因素使市场无法达到完全竞争状态。然而人口密度增大并不会导致产品市场上出现垄断，也不会让劳动力市场的竞争程度降低，所以最终也并不会降低消费者的理性程度或是损害他们的知情权。因此，管制与人口密度之间的正相关性并不能为传统的自由主义观点提供依据。

然而，这种联系却完全符合达尔文的观点：个人利益与集体利益经常存在严重的冲突。当个人利益与集体利益背道而驰时，人口密度越大，人们能够加害他人的机会显然也就越多。

达尔文本身并未就如何缩小个人利益与集体利益之间的鸿沟发表太多意见。毕竟，大多数动物只有极其有限的集体协作能力。人类却显然是动物中的一个例外。即使是最早期的狩猎采集者，也广泛应用社会规范来限制社会成员的自利行为。¹然而，那个时代的人类却不太了解如何以最有效的方式去限制那些加害他人的行为，可谓仍然处于蒙昧之

中。

每一个加害他人的人类行为都强调了增加管制的必要性，这并不是一个过分夸张的说法。但显然禁止一切有害他人的行为是不可行的。人们应当这样理解穆勒提出的伤害原则：政府限制个人自由的唯一正当理由是防止过度加害他人。一直以来，我们都允许有害他人的行为发生。例如，不友善的评价会伤害他人，但我们仍然有限度地允许人们做出这类评价，因为剥夺言论自由带来的危害会更加严重。与此类似的是，我们也允许一家公司通过降低价格损害商业竞争对手的利益，尽管这样做甚至有可能导致对方破产，但若禁止采用这种竞争手段会给消费者带来更大的坏处。伤害原则必须这样理解才有道理：某种对于个人行为的限制是否正当，应当比较进行限制所需的成本与该行为不受限制时其他人受到的伤害大小，依据这两者孰轻孰重来决定。

科斯的成功之处

罗纳德·科斯（Ronald Coase）于1991年获得诺贝尔经济学奖，主要由于他极大地帮助人们改变了对上述精细权衡的看法。科斯在英格兰出生并接受教育，在职业生涯后半段，任教于芝加哥大学法学院。芝加哥大学里的自由市场坚定支持者们十分尊敬他，将他视为研究加害他人行为的学界权威。

激进的自由主义者不愿承认这些加害他人的行为理应受到规制者的监管。但那些更为认真的自由主义者一直都承认不放任人们肆意实施加害行为是必要的，芝加哥大学杰出的师资团队里有不少人都在其列。²他们视科斯为其领袖，是因为相信科斯的理论框架为缩小管制涉及的范围提供了最令人信服的理由。在某种意义上，他们的看法没有错。但在另一种意义上，我认为他们彻底地误解了科斯。

科斯确实指出，在某些有特定条件限制的情况下，私人主体可以自

行解决许多原本被人认为必须由管制者仔细干预的问题。³然而，科斯虽然非常赞赏许多市场解决方案的精巧与优雅，但他并不是自由市场的狂热支持者。科斯的许多拥护者总是会忽视，科斯本人明确承认许多实际的阻碍时常让私人主体之间无法协商达成高效率的合同。他认可政府干预在这种情况下常常能取得良好效果。

科斯的观点最早见于1959年发表的一篇题为《联邦通信委员会》（The Federal Communications Commission）⁴的论文中的一个简单例子。就和今天一样，价值数百万美元的广播权在当时也由联邦通信委员会的管制者免费授予私人企业家们。科斯认为，若能把广播权卖给出价最高者，由于出价最高者同时也是会将广播权用在最具价值之处的一方，这会使广播权的行使效率更高。

最初催生广播监管的一个原因就在于，行使广播权时常会损害他人的利益。在一个波频上的信号广播总是会干扰到相近波频上的广播。科斯认为，若广播者之间互相协商，他们自己便可以圆满解决这一问题。

科斯给出了一个与广播干扰问题类似的例子，这个例子如今已经广为人知：一位医生对相邻的工厂提起诉讼，因为工厂的噪声妨碍他医治病人。科斯将这一案例总结如下：

一个糖果生产厂已经在某个地方运作多年。有位医生搬到工厂附近，糖果生产厂的机器运转最初并未给医生造成任何困扰。近8年之后，这位医生在他院子尽头靠近工厂的地方建了一间问诊室，然后便发现工厂机器发出的噪声与震动会干扰自己工作。医生于是向法院成功申请到了一项不允许工厂使用机器的禁令。在这件事里，法院的工作实际上是要决定，究竟是医生有权强迫糖果生产厂安装新机器、搬去新厂址，从而给糖果生产厂增加额外的成本，还是糖果生产厂有权要求医生在他院子中的其他位置工作或者搬到别处去，从而使医生承担额外的成本。⁵

在科斯之前，人们对那些有害他人的活动进行政策讨论时，经常采用区分加害者与受害者的分析思路。制造噪声的工厂被视为加害者，邻近处的医生则是遭到噪声干扰的受害者。科斯则指出，噪声、排烟等外部性问题，是纯粹的相互作用现象。工厂的噪声对医生不利，这一点没错；但以医生受到损害为由禁止工厂发出噪声，同样妨害了工厂主的利益。

干扰医生显然并不是工厂主的本来意图，医生也并非存心要阻碍工厂运作。两者间距离较近，各自从事活动的性质又恰好彼此冲突，所以才引发一个亟待解决的共同问题。科斯认为，若能找到成本最低的解决方法，工厂主与医生都会从中获益。因此，除非互相协商是不现实的，否则不论政府是不是要让工厂为噪声损害承担责任，两个当事人总是可以高效地解决问题。

例如，假设医生在不采取任何行动的情况下会因为噪声而蒙受20 000美元的损失，但他若花费10 000美元搬去别处就可以彻底免除噪声干扰。再假设工厂主花5 000美元给自己的机器安装隔音设备就可以消除噪声。由于隔音设备的成本更低，比起让医生搬离，消除噪声是个更好的选择。科斯指出，若医生和工厂主能以低至足以忽略不计的成本彼此协商，他们就应当能够达成安装隔音设备的协议，政府是否判定工厂主为噪声损害负责并不会影响这一结果。

在工厂主确实为噪声损害负责的情况下，我们不难理解为何政府的判定不影响结果。如果政府认定工厂主要为噪声损害负责，这就意味着再继续产生噪声的话，他就必须向医生支付20 000美元的损害赔偿金。对此，工厂主的最佳应对方法一定是安装仅需要5 000美元的隔音设备，以此避免20 000美元的赔偿金。

至于工厂主不需要为噪声损害负责的情况，科斯的结论让大多数读者出乎意料。他指出，如果政府认为工厂主不必为噪声损害负责，医生的最佳选择便是提议由自己出钱、让工厂主安装隔音设备。他的次佳选

择才是花费10 000美元搬离噪声影响之处。

一方面，医生付给工厂主的钱不得小于安装隔音设备所需的5 000美元，否则工厂主一定会拒绝其提议，因为反正他并不需要为噪声损害提供赔偿。另一方面，工厂主要求的钱款金额也不能超过10 000美元，否则医生将会放弃这个提议直接选择花10 000美元搬走。

简而言之，科斯认为双方当事人都有充分的理由采取效率最高的方式解决问题，不论根据法律工厂是否要为噪声损害负责。科斯明确指出，这一结论必须建立在当事双方相互协商是现实可行的这一假设之上。因此，他这样写道：“法律明确了当事人的权利之后，他们自己仍有可能通过协商来改变法律规定为他们设想的安排，只要这样做带来的收益值得他们花费彼此协商所需的成本。”⁶

《联邦通信委员会》这篇论文在1959年发表后，几乎没有引发什么议论。许多人读完之后都觉得科斯肯定弄错了，这其中就包括著名经济学家米尔顿·弗里德曼、乔治·施蒂格勒（George Stigler），以及芝加哥大学的其他知名的自由市场支持者。他们认为，即便在一个能以零成本协商达成协议的假想世界里，生产决策显然也会受到法律是否要求人们为自己给他人造成的损失承担赔偿责任这一因素影响。科斯承认医生与工厂主各自的财富水平确实会受到归责原则的影响，但他坚持认为，双方决定问题解决的方式并不会受其影响。

芝加哥大学的那群经济学家从来不会在学术争论中善罢甘休，他们邀请当时还在弗吉尼亚大学任教的科斯来芝加哥，一起讨论科斯的理论。20名经济学家和科斯一齐聚在《法律与经济学期刊》（*Journal of Law and Economics*）编辑阿伦·迪雷克托（Aaron Director）家中共进晚餐，科斯的论文正是在这一期刊上发表的。多年以后，施蒂格勒这样回忆那晚的谈话：

同平常一样，大多数时候都是米尔顿·弗里德曼在发言，他也

是进行了最多思考的人。经过两小时的争论，刚开始全部反对科斯观点的20个人，最后全都赞成他的意见。那是多么令人振奋的一次聚会啊！事后我十分后悔，我们没有先见之明，没将那晚的对话录下来。

他们说服科斯再写一篇更为详细的论文以阐明观点，这就是一年之后（即1960年）发表在《法律与经济学期刊》上的《社会成本问题》（The Problem of Social Cost）⁷。这篇论文很快就成了有史以来最广受引用的经济学论文，至今仍是如此。⁸然而，尽管这篇论文受到了极高的关注，学者们仍未彻底理解其重要意义。

这篇论文一经发表就很快引来了风暴般激烈的批评。左派学者感到不满，因为他们将科斯的观点理解为，除了产权能得到清晰界定与执行外，对污染或其他加害他人的活动进行政府管制并无必要。另一些人反对科斯的观点，因为他的理论似乎完全摒弃了之前区分加害者与受害者的道德框架。他们认为：医生并没有加害于任何人，为什么他必须为隔音设备付钱？但这两种批评都经不起仔细推敲。

协商路上的绊脚石

先来考虑关于科斯彻底排斥政府监管的指责。任何读过他更早的那篇论文的人都会知道，科斯其实深信“交易成本低至相互协商现实可行”的假设前提在实际生活中不可能得以满足，而他“不论污染者是否要承担法律赔偿责任，都不影响高效率的解决方案的执行”的结论明显是建立在这一假设前提之上的。发表于1937年的《企业的本质》（The Nature of the Firm）是科斯的第一篇重要文章，而促使他写这篇文章的因素，正是他观察到，实际的困难时常使人们无法彼此沟通，达成高效率的协议。⁹这篇论文的内容主要基于他5年前还是本科生时作为研究员前往美国参访大型企业期间进行的田野调查。这篇1937年发表的论文主

要想解答的问题是：到底为什么会出现企业？他想知道，为何独立的承包商之间不直接进行商业交易，而要通过企业这种形式。就如同许多伟大的思想一样，他给出的答案在事后看来也显而易见得令人懊恼：若没有企业的存在，人们就需要进行一系列复杂而高成本的交易来替代企业的作用。

打比方说，你想买一辆汽车，就不得不先和许多矿工协商，签订合同，委托他们帮你挖掘铁矿；再和其他人签订别的合同，让他们把你的铁矿变成钢铁，然后继续签订更多合同将钢铁锻造成你想要的形状，以此类推。即使买家在每一个环节都与效率最高的供应商合作，所需合同的数量与复杂程度仍然会使这辆汽车的最终价格高得惊人。科斯认为，组建一个组织，雇员在其中只需要遵循上级命令，这样整个过程就能得到极大的简化。

自1937年《企业的本质》一文发表以来，经济学的一个新领域开始萌芽并蓬勃发展。这一领域被称作“交易成本经济学”，它将组织的形式与行为解释为人们为了降低交易成本而导致的直接或间接结果。¹⁰

从科斯以往的学术成果来看，他无疑非常清楚现实阻碍时常使私人主体之间的协商成本高得惊人。他想要表达的观点，并不是说政府对于管控加害他人的活动毫无用处。相反，他在1960年发表的《社会成本问题》一文的深层含义恰恰在于，在协商解决并不现实的情况下（实际情况基本一直如此），法律对责任分配做出不同的界定，解决问题的方式也会相应不同。因此，法律责任的界定在科斯的理论中其实至关重要。但其重要性的体现方式不受早先很多学者的待见。当时的主流意见认为，加害者一定要为自己造成的损失承担责任。科斯从根本上动摇了这种逻辑。他认为政府应该让能以最低成本应对外部性问题的一方来承担应对的责任。

回到之前举的医生和工厂的例子，如果医生和工厂主无法自主协商，政府就应当让工厂对噪声损害承担赔偿责任，因为根据这个例子中

的假设条件，工厂减小噪声要比医生搬去不受干扰之地成本更为低廉。但如果反过来是医生搬走的成本更低，政府就不应该要求工厂对噪声损害负责。科斯认为，当政府不要求工厂承担损害赔偿时，医生便会选择自己搬迁，而这正是相应情况之下效率最高的解决方式。

成本-收益分析是不道德的吗

还有一种反对科斯的理由，是认为科斯的分析完全忽略了对加害他人的活动进行管控背后的道德框架。许多批评科斯的学者未能清楚认识到，科斯的分析恰恰揭示出传统道德框架有一个根本错误。如前文所述，这一框架将工厂主视为加害者，而医生则是受害者，没有给出任何原因便预先假设医生有权利在行医时不受任何噪声干扰，而且认为政府有义务以强制措施保护这一权利。

科斯在分析中质问，为何在这个情境下医生就一定要得到更多优待呢？比方说，如果把之前假定条件中设计的成本数额互换一下，医生花费5 000美元就可以搬迁到其他地方，解决噪声问题，而工厂主安装隔音设备需要10 000美元。在这种情况下，成本最低的解决途径便是让医生搬迁。根据科斯的分析，如果当事人彼此协商并不可行，政府就不应该让工厂主为噪声损害负责，这样一来，医生便会有动力通过自行搬走来解决噪声问题。

在上述情况下，让工厂主为噪声损害负责真的是更好的选择吗？如果确实这样做了，在协商解决可行的时候仍能有最佳解决方案，因为工厂主有理由向医生支付一笔钱，让医生自己搬迁，以此避免安装隔音设备的10 000美元开销。但这种协商在现实生活中基本是无法做到的。如果一定要让工厂主为噪声负责，他最好的选择便是花费10 000美元安装隔音设备，否则他就只能为噪声损害支付20 000美元的赔偿金给医生。没错，噪声问题仍然得到了解决，但花费的成本却是让医生自行搬迁的足足两倍。

人们可能会从直觉上认为工厂主是上述案例中的加害者，然而科斯的分析指明这种直觉并无逻辑基础。如果医生比工厂主更早来到此处，他或许能以“将来也可以一直不受噪声干扰地行医”为理由而行动。若能够证明，以这样的理由行动有利于进行高效率投资活动，这或许确实是一个好理由，让后搬来的噪声制造者为噪声损害负责。¹¹然而，在科斯的例子中，工厂主才是先在此地营业的一方。

还有一些人则担心，解决这一问题的成本过高，医生无力负担。然而，总归有人要承担解决问题的成本。例子中并没有任何信息表明，工厂主就一定比医生更有钱，更有能力负担这种成本。或许这是位依靠轮椅才能出行的拉丁裔工厂主，勉强维持着工厂运营，他雇用了50名勤劳、有信仰、不酗酒的优秀公民，而这些雇员每个人都要以一己之力养活妻子、两个孩子再加上一只猫和一条狗。又或许这位医生是个单身的花花公子，他作为皮肤科医生凭为悠闲的阔太太们改善肤质就能赚取惊人的高额利润，每周有三个下午都在打高尔夫球。

既然不存在任何这类与我们的结论相抵触的因素，我们就不得不考虑传统的“加害者-受害者”道德框架是否可能是错误的，或者是建立在有瑕疵的假设之上的。也就是说，我们必须考虑到，在这类情况下界定权利的最合理方式，可能如科斯所说，正是尽可能模拟在协商解决可行的情况下，人们会自主做出的那种选择。这种选择一定是让适应外部性问题所需成本最低的一方负责解决外部性问题。有时这种选择确实会让被人们普遍认为是加害者的那一方来承担责任，但并非总是如此。

科斯迅速成了支持自由市场的保守派们在所有加害行为相关议题上的智识偶像，这一点实在有些讽刺。他们之所以如此推崇科斯，主要还是认为科斯的理论框架能使更多议题被纳入无须管制干预即可解决的范畴中。这种认识并不完全是错的。然而，我已经说过，科斯从来都不是一个自由市场的狂热支持者。他的理论框架完全植根于现实考虑。

自由主义者时常提到那些高于现实考虑的基本人权与自由，然而保

护任何权利都势必会有成本与收益。¹²科斯采用的分析思路归根结底就是在强调，“哪些特定权利应当受保护”仍然是典型的实际问题。

科斯可能会得出这样的结论：在特定情况下，医生应当有权利在行医时不受邻居的噪声干扰。但即便如此，他这样认为的理由也并不是在任何情况下人们都有享受零噪声环境的神圣权利，而是让制造噪声的人消除噪声比让受影响的人自行消除影响的成本更低。

由于大多数道德问题都涉及加害他人的行为，科斯的理论框架同这些问题也有一定联系。如果他被道德哲学学院聘请担任教职的话，他就会被归类为一个结果论者，结果论者认为一个行为正确与否的关键在于是否做出了能够产生最高总体效率的选择。¹³与他们相对立的则是道义论者，道义论者虽然承认结果并非完全无关紧要，但坚信存在一些特定的基本道德原则，无论结果如何人们都必须遵守。

除了要遵守禁止盗窃、袭击和其他直接伤害他人的恶行的法律以外，人们有权免受任何政府管制的干扰，这在自由主义者眼中是完全不需要理由、不言自明的。基于这种观点，大部分自由主义者都无疑属于道义论者的阵营，而许多拥护科斯观点的人则是彻底的结果论者。如果真的仅凭一种如此宽泛的权利就一定能导致最好的整体效益，应该有一个合理解释才对。我们发现，要给出这样一个解释极其困难。

道义论者还面临其他问题，例如无法解释清楚他们援引的基本道德原则到底从何而来。但他们在与结果论者的辩论中时常占据上风，因为他们只需要设计一些具体情景，让其中能够带来最佳整体效益的选择对一般旁观者而言难以接受即可。

道义论者一直以来最喜欢的一个例子就是：有个植物学家来到一个丛林部落中，发现10个无辜的人即将被射杀。¹⁴他被告知，如果他射杀其中一个人，另外9个人就可以被释放。这个植物学家该怎么做呢？结果论者似乎会认为射杀一个无辜的人是正确选择，因为这样就可以救活9条生命。然而大多数普通的、有感情的人都不愿意接受这一结论。道

义论者坚称这类例子显示结果论者的道德逻辑无法成立。

结果论者与道义论者之间的争论已经持续了千年之久，我在这里所说的任何话都不可能解决两方之间的争议。由于我要支持的政策是由科斯的结果论式理论框架衍生而来的，所以在进一步阐述之前我首先想强调的是，结果论者与道义论者的理论框架其实并不像看起来那样完全不可调和。

跨种族婚姻中的夫妻是否应被禁止在公共场合牵手

下面就举一个具体例子，帮助大家更清楚地看到，一些因素使两种道德逻辑之间的不可调和性被人为夸大了。这个例子正是道义论者在证明结果论者的观点属于道德沦丧时经常使用的一种典型。20世纪60年代中期，佐治亚州亚特兰大出现了一个问题：跨种族夫妻是否应当被禁止在公共场合牵手？如今，大多数美国年轻人一定很难相信，曾经还有许多州的法律都禁止不同种族间通婚。在1967年一个里程碑式的公民权利案件中，美国联邦最高法院的大法官们以9对0的完全一致的投票结果宣告这些法律全部违宪。¹⁵但在20世纪60年代中期的亚特兰大，跨种族婚姻仍然是非常罕见的，白人构成了人口的主要部分，而这些白人中有很大一部分人看到跨种族夫妻在公共场合牵手就会感到十分不悦。在这样的环境里，若赋予跨种族夫妻在公共场合牵手的权利，真正因此获益的人非常少，却会有许多人为此受到冒犯。显然，按照最佳整体效果来选择的话，就应当禁止跨种族夫妻在公共场合牵手，然而大多数人都会认为这种禁令有违道德。道义论者因此得出结论，我们绝不能接受以结果论的逻辑来考虑道德问题。

上述问题很容易被改编为一个科斯式的数字计算例子。假设亚特兰大有100对跨种族夫妻，每对都愿意每周支付100美元以换取在公众场合牵手的权利，所以他们加起来一共能够每周支付10 000美元；而亚特兰大共有100万白人愿意每周支付1美元以避免看到任何跨种族夫妻牵手，

那么赋予跨种族夫妻当众牵手权利的成本就是100万美元。如果两方能够协商解决问题，那么只需要每个白人一周贡献0.1美元，加起来就有10万美元之多，这笔钱可以用来每周支付给每一对跨种族夫妻1 000美元，让他们不在公共场合牵手。相比获得牵手权利，跨种族夫妻每周能够多赚900美元（1 000美元的收入减去他们无法当众牵手而损失的100美元），而受到冒犯的白人每人每周多获利0.9美元（每人每周1美元用以阻止牵手的开销减去他们每人分摊到的0.1美元花费）。

科斯的理论似乎表明，若涉事双方能够协商解决，他们将会同意执行禁止跨种族夫妻牵手的法令，并如上所述由白人向这些跨种族夫妻付钱。然而在现实生活中，人们显然并不能进行这种协商。科斯认为，在无法由当事人自行协商解决的情况下，法律应当让那些能以更低成本适应外部性问题的人来承担适应外部性问题的责任。既然跨种族夫妻不能当众牵手的成本共计每周10 000美元，这个数字小于看到他们牵手就感到不悦的白人每周共计100万美元的损失，依据科斯的逻辑，似乎正确选择就是禁止跨种族婚姻者牵手。

这个例子确实能解释为何结果论的道德分析至今让许多人难以接受。但在因此而否定科斯的理论之前，同我一样不赞同禁止跨种族夫妻牵手的人们应当首先考虑一下，科斯的理论有没有可能在上述例子中被错误地运用。

根据前文的描述，这种分析思路完全忽视了一个事实：在一段时间里，不同的人会以完全不同的方式去适应各种实际存在的或想象出的损害。在20世纪60年代，亚特兰大的白人居民愿意为避免看到跨种族夫妻牵手而支付的总额，可能确实大于那一小部分跨种族夫妻愿意为获取牵手权利付出的成本。然而，在那几年里，跨种族亲密关系的发生越来越频繁，人们对跨种族婚姻的态度也在迅速转变，这些变化在当时也都是完全可以预测到的。

那个时候我正在亚特兰大读本科，我的许多同学都坚定地反对任何

形式的跨种族关系。然而，我的孩子从未在美国南方生活过，他们就说自己对于两个不同种族的人牵手完全没有任何情绪反应，毕竟这在他们长大的环境中实在太常见了。任何一个在美国南方以外的地方生活过的人都可以预料到，南方白人那种看到不同种族的情侣牵手就仿佛受到精神伤害的感觉，一定会随着时间流逝，随着他们逐渐适应环境的变化而消减。

与此恰恰相反，也有其他一些类型的损害，不仅不能随着时代发展而逐渐消失，反而愈发严重了起来。背景噪声造成的损害就属于这一种。人们不仅无法慢慢适应喧嚣的环境，还会随着噪声的持续而产生愈发严重的生理负担。¹⁶那些被禁止和伴侣牵手的人，受到的相应伤害也会随着时间发展而愈发严重。一个人不仅不能习惯无法享有其他人总是可以享有的权利，他遭受的情感伤害还会日益加剧。

加害他人的行为天然就是一种相互作用。当跨种族夫妻牵手，有一些人会受到伤害。当跨种族夫妻不被允许牵手，也有一些人一定会受到伤害。科斯的理论表明，对于这种情况的最佳应对方式是尽可能减小总的伤害。要正确运用他的理论，不仅需要评估当事各方当下所受伤害或自称所受伤害的程度，还必须考虑到他们日后慢慢避免或适应这些伤害的能力。简而言之，正确运用科斯的理论，本来就绝不会得出需要禁止跨种族夫妻当众牵手的结论。

我们应该捍卫何种权利

做出特定行为的权利并不是凭空而来的。要捍卫任何权利都会使看重这些权利的人获利，同时也产生成本。不仅公权力机构保护和推行该权利需要成本，那些被这一权利限制了自身行为的人也要承担成本。

即便是道义论者也承认，成本与收益在道德决策中并非无关紧要。一个社会在决定哪些权利应当受到保护时当然要考虑道德问题。我并非

宣称科斯理论就是进行这类社会决策的唯一正确途径。但既然人们能够同意，比较面前各个选项的成本与收益至少在一定程度上是必要的，那我希望大家也一样可以同意，科斯的理论总能帮助我们更清晰地进行这种比较。

许多自由主义者与保守派确实推崇科斯的理论已久。然而，稍加观察便会发现，这种推崇让他们很难拒绝应用穆勒的伤害原则。在他们反对某项自己不喜欢的管制时，若伤害原则能充当他们反对的依据，他们自然很乐意支持伤害原则。但若是依据伤害原则应当支持他们不喜欢的某项管制时，他们往往立刻就否定伤害原则，坚称这一原则阻碍他们以自己希望的方式行使某些他们生来就拥有的权利。既然提出了这种理由，他们就不得不面对以下质问：他们宣称自己本来就享有的权利，究竟从何而来呢？

自由主义者们所推崇的科斯理论表明，社会在界定并强制实行权利的时候，必须模拟在自行公开协商可行的情况下人们会达成的协议结果。根据这一思路，承受某一加害行为结果的人将会愿意牺牲实际财富以避免受到该行为的伤害。这样理解的话，伤害并不局限于肉体上的暴力或是盗窃财产等。所谓伤害，还要包括我在前几章里讨论过的各种形式的间接伤害，就比如一个短跑运动员在对手服用兴奋剂提高成绩时会受到的伤害。

若自由主义者坚持对伤害原则保持墙头草一样的态度，他们势必会遇到难题。为了在穆勒的伤害原则应当适用时拒绝适用它，自由主义者就必须宣称自己有权利给他人施加实质性的、可测量的伤害，哪怕这样做对他们自己并无任何好处。然而这种权利在科斯的理论框架里完全站不住脚，他们因此不得不同时否认了科斯的正确性。想一想，一个诚实的自由主义者怎么可能坚持这样的立场呢？如果你们当真非常重视自由，又怎么会否定一个意在尽可能模拟完全自由意志者自主协商结果的理论框架呢？

为了进一步讨论，我将采用一个不肯放弃科斯理论的自由主义者的视角，试着继续阐释科斯的理论将会如何引导我们看待权利。到目前为止，我们原则上可以同意的一点是，在决定哪些权利应该得到保护时，需要考虑成本与收益的问题。然而，如何计算相关的成本与收益又是另一个话题了。如何比较利益相冲突的各方之间孰轻孰重，基本是各个社会理论家之间争议最大的问题。在下一章，我将阐明科斯的理论框架对回答这一问题也提供了非常有益的指导。

THE DARWIN ECONOMY

THE DARWIN ECONOMY

07

效率至上：
金钱在调节利益冲突者之间关系中的作用

货币的发展是导致劳动分工细化的重要一步，而劳动分工细化的过程在过去的几个世纪里使人均收入增长了数千倍。在决定两种相互冲突的利益之间孰重孰轻时，货币作为一种衡量单位是不可或缺的。

因此，在科斯举的那个工厂机器噪声干扰邻近医生行医的例子中，要确定解决问题的最佳方案，只需要简单地比较一下工厂主为自己的机器安装隔音设备所需的开销与医生搬去安静地方的开销即可。进行这一比较与两人本身的收入高低完全无关，虽然他们各自的收入高低确实会因法律判定哪一方对噪声损害负责而受到影响。

收入差异对成本-收益分析的影响

有的时候，受到有害行为影响者的收入高低，直接决定了成本最低的问题解决方式是什么。例如，假设萨拉喜欢在晚上拉小提琴，而她的邻居萨姆会被琴声打扰。法律应当限制萨拉拉小提琴吗？为了使讨论简短一些，假设立法者只考虑以下两个选项：要么禁止在工作日晚上10点以后演奏小提琴，要么禁止在工作日晚上11点以后演奏。为了确定哪个选项能以更低的成本解决问题，立法者们需要知道两方面的信息：其一，萨拉能够一直演奏到晚上11点而不是提前一小时结束，对她而言具有多少价值；其二，让萨拉在晚上10点就停止演奏而不是一小时后再停止，对于萨姆而言具有多少价值。

萨拉显然更希望能够一直演奏到11点为止，然而她的愿望到底有多强烈呢？对大多数经济学家而言，要回答这个问题至少在理论上是十分简单的，就算实际生活中可能并非一直如此。在经济学家心中，她的愿望强烈程度就由她愿意支付多少钱以换取一直演奏到晚上11点的权利来

衡量。对于萨姆有多需要晚上10点就开始享受安静的权利，经济学家也会用类似方法计算，也就是看他最多愿意支付多少钱以获取这项权利。当然，经济学家心里明白，若是简单询问当事双方他们愿意付多少钱的话，他们一旦知道自己说出的金额会决定“琴声禁令”的每日起始时间，就一定都会夸大自己的答案。所以在进行实际操作时，经济学家并不会直接问他们，而是从间接行为证据中推断出他们各自愿意支付的数额。例如，可以参考其他条件都相同而噪声禁止时间不同的街区之间房价的差异。

对于通过比较涉事各方愿意为实现心愿支付的金额来解决噪声问题的提案，实际计算是其中最不具争议的部分。当事人自身的消费能力会严重影响他们愿意为获取权利而支付的数额，这才是一个麻烦得多的问题。人们愿意为多出一小时安静时间或是多出一小时演奏乐器时间支付的数额，当然取决于许多因素，包括他们对音乐的态度如何、公寓的墙厚不厚，以及工作时间安排。然而与此同时，支付意愿很大程度上也取决于收入水平。其他因素相同的情况下，一个人的收入越高，自然就越乐意付更多钱来换一个得偿所愿。

于是，依照科斯的理论解决噪声问题，结果便会倾向于选择更符合富人意愿的解决方式。例如，假设萨拉非常希望能够练习小提琴到晚上11点，然而她贫乏的收入水平最多允许她每天支付8美元来换取这一权利。另一方面，萨姆其实并不会被萨拉晚间的小提琴声打扰得多么严重，但他极其富有，愿意每天付15美元换取晚上10点后的安静。在科斯的理论框架下，解决问题的高效率方案便是发布禁令，禁止萨拉晚上10点以后演奏小提琴。然而，我曾在一堂课上对90名研究生进行过调查，问他们认为在这个案例中应该如何处理问题，80多名学生都认为噪声宵禁应该设置在晚上11点，而不是晚上10点。

科斯的《社会成本问题》最初在1960年发表时，许多左派评论家给出负面评价的理由都在于，他们认为以支付意愿为基准进行成本和收益

的计算是有悖道德的，这对于公民不论收入高低人人平等的原则是一种令人不齿的背叛。这一观点乍看很有说服力，其认识却是错误的。其实，正是这种观点导致我们总是无法成功推行那些能够使全体公民不论贫富都受益的政策。

大多数人其实基本都能接受，支付意愿决定了一个人能在私人市场中获得什么。打比方说，如果许多人正参与竞拍一幅画，愿意付最多钱的人就是画的得主。利用成本-收益分析来指导公共决策时，支付意愿同样是一个决定性因素。然而，尽管大多数人都能接受以支付意愿作为私人领域的资源分配基准，要在公共领域做同样的事却争议重重，人们经常否决那些在决定主要公共政策时采用传统成本-收益分析的提议。

电台节目应该是纯谈话还是纯音乐

一个简单的例子就可以让我们大致了解一下相关问题。假设有一个社区必须决定是否要将其公共广播的形式从全部由音乐节目组成改为全部由谈话节目组成。为了便于讨论，我们假设制作这两种形式的广播成本相同，除了以下三位之外其他投票者都对这件事无所谓：一位富人希望广播变为谈话节目，另外两位穷人分别同等程度地倾向于反对这一改变。由于他们之间的收入差异，富人愿意付1 000美元促成改变，两个穷人分别只愿意付100美元来阻止改变发生。那么是否应该改变广播节目的形式呢？

根据标准的成本-收益分析，改变节目形式的收益将是富人乐意为此支付的1 000美元，而成本则是两名穷人为了阻止改变发生愿意付出的金额之和，也就是200美元。因此，改变节目形式的净收益是800美元。成本-收益分析的规则就是支持一切收益大于成本的选择，据此，毫无疑问应当改变广播节目形式。

然而，评论家们反对称，以成本-收益分析的方式解决这类问题违

反了之前提到的重要民主原则：法律应对每个公民的利益一视同仁。在刚才讨论的案例中，三名当事人每人对各自支持的节目形式都有同等程度的倾向。所以评论家们质问道，为什么那一个富有者的喜好就要胜过两个穷人的意愿呢？

这个问题明显在措辞上很有力量。在这类问题中，使用成本-收益分析来进行决策确实存在一些问题。出于这些担忧，许多政府都不愿意使用成本-收益分析。例如，克林顿执政期间发布的行政命令要求各联邦机构首脑根据财富分配的情况来修正通过成本-收益分析得出的结论。¹如果我们用支付能力来修正通过支付意愿得出的结果，在广播的例子中进行成本-收益分析也就只是简单的人数统计罢了，将广播由音乐形式转变为谈话形式的提议会因为2对1的局面被否定。

这就是正确的结果吗？问题不在于实施成本-收益分析是否在某些方面存在缺陷，而在于可以替代成本-收益分析的决策方式是否能得出更好的结果。

在这个例子中，舍弃成本-收益分析的一个显著结果就是，拒绝改变节目形式会导致800美元的净损失，这是由富人的1 000美元减去两个穷人共愿意支付的200美元所得。这一损失完全可以轻松避免。比方说，我们可以为改变节目形式附加一个条件，要求富人多支付500美元税金并为两个穷人各减税250美元。这样，相比前一种情况，三个人都能获得更多利益。

此时，富人或许会抱怨自己多交了税。然而从他的角度来看，原本的结局分明就更糟。从改变广播节目中获得的1 000美元价值足够补偿他多支付的税金。同样，尽管两个穷人失去他们喜欢的音乐广播节目后会感到不满，但真要按照原本的情况来安排，他们得到的只会更少。根据他们自己的说法，250美元的减税额足以补偿他们从节目改变中损失的价值了。若没有进行这次收入转移，所有人都是输家。

自由主义者及许多其他群体反对称，政府没有权力强制执行这种收

入转移。他们说：“这本来就是你们的钱，政府没有权力抢走你的钱拿去送给穷人！”虽然他们的这种主张经常被人从公平的角度来反驳，但是，这种主张从效率的角度来看明显更站不住脚。

在民主社会中，贫穷选民的利益总会以各种方式得到关注。正如在前面的例子中，如果排除收入转移的方式的话，就需要以其他成本更高的方式来照顾低收入选民的利益。相反，若是允许进行收入转移，就有可能制定出使贫富双方都能获得更多经济利益的政策。然而，除非政府能有比现在更多的以支付意愿为基准的自由，否则就不可能知道该制定什么政策并加以实施。要获得这种自由，人们就必须更大程度地接受收入转移。

我们已经知道，以支付意愿为基准来决定由谁获得什么东西，在私人领域远比在公共领域受人认可。然而，即便是在私人领域里，参照支付意愿做决定有时也会使一些重要的价值受损。在这里，最核心的问题同样不在于支付意愿是不是能够产生最佳结果的分配基准，而在于其他可行方案能否产生比参照支付意愿进行决策时更好的结果。接下来的这个例子能够让我们清楚地认识到，支付意愿虽然有时会在某些方面导致令人不悦的结果，但它确实是一种非常值得推荐的分配基准。

稀有的钟应当归谁所有

假设有一家古董店将一座1905年的斯蒂克利落地大摆钟展示在店门口的橱窗里。有两位潜在的购买者想要买下这座钟：一位是苏珊，她是教小学四年级学生的教师，她酷爱20世纪早期的落地大摆钟；另一位是马尔科姆，他是个富有的人身伤害案件律师，对20世纪早期或其他时期的钟并没有特别大的兴趣。苏珊经常参加古董研讨会，也一直梦想着拥有一座价格在她财力范围内的斯蒂克利钟。马尔科姆只是恰好路过了这家古董店，他看到橱窗中的钟后觉得放在办公室中供访客等待的房间里应该看起来还不错。苏珊是一个年收入2.8万美元的单身母亲，她愿意

出5 000美元买下这座钟。对她来说这确实是一大笔钱了，这个出价足以显示出她有多么渴望获得这座钟。而马尔科姆每年能赚95万美元，他不怎么特别在意这座钟，却依然愿意出1万美元买下它。在市场体系中，马尔科姆更高的支付意愿无疑会让他获得这座钟。

许多人情感上都会难以接受让马尔科姆买下钟是最佳结果，甚至不少自由企业制度的拥护者也是如此。毕竟，这座钟一定会给苏珊带来长久的、深刻的满足感，但它对于马尔科姆而言就没太大意义。面对这些疑虑，很多人或许还是会接受市场机制运行的结果，因为他们认为不存在一种实际可行的方法，能让我们利用该例子里那些反映购买愿望强烈程度的非正式信息。比如说，实际操作中很难依据人们宣称的购买愿望的强烈程度来分配商品与服务。

然而，即便在这个例子中，其实也存在着一种不那么模棱两可的理由，可以支持我们将支付意愿当作分配基准。我们看到，苏珊比马尔科姆更在乎这座钟，但是，如果我们相信两个买家提出的报价是真心的话，让苏珊得到这座钟对她来说不是好事。原因就在于，马尔科姆挪出1万美元来买钟也无所谓，但苏珊与马尔科姆不同，在几乎所有方面她都更迫切需要手里这笔用来买钟的钱。马尔科姆比她富有太多，他在任何地方花的每一笔钱都不如同样一笔钱对苏珊而言来得举足轻重。例如，苏珊用这笔钱或许可以搬去一个更加安全的街区，或是更好的学区，这对苏珊的影响比对马尔科姆的影响大多了。并不是说马尔科姆不在乎住处的安全性或孩子的学校，而是因为他已经居住在有好学校的安全街区里了。如果让苏珊得到这座钟，根据她心中这座钟的价值数额，她最佳的选择应当是将钟卖给马尔科姆。²毕竟，他付给苏珊的数额会比钟本身的价值还要多出5 000美元。苏珊说她愿意出5 000美元买钟，确实表明这座钟对她来说十分重要，但这并不意味着这座钟就比她能同样用5 000美元买到的其他东西对她来说更加重要。

假如马尔科姆最终以8 000美元的价格买下了钟，苏珊也许会因为

自己并没有足够的财力买钟而感到极其失望。但根据她自己对钟的估价，花8 000美元买钟却无法拥有那些能用8 000美元换取的其他东西，只会让她更加失望。如果我们相信苏珊和马尔科姆各自陈述的价格属实，那么让马尔科姆买下钟对他们两方来说一定都是更好的结果。

并不是说，这就是人们能想象到的最好情况。比如，人们若觉得购买力的分布能更均衡一些的话世界会更美好，也并不是什么离谱的想法。如果能让教师一年挣7.5万美元而人身伤害案件的律师挣20万美元，苏珊可能就会比马尔科姆出价更高，毕竟马尔科姆完全不如她那么重视这座钟。但在现有的收入分布情况之下，能够产生的最好结果就是让马尔科姆买下这座钟。

上述例子表明，依照支付意愿来分配稀有产品时，“经济蛋糕”是最大的，尽管这一分配基准在某些方面不尽如人意。如果我们以其他基准来分配，一定总还能够找出一个再分配的方案，使各方都获得更多利益，就像在钟和广播的例子中那样。这背后的原理就是，当“经济蛋糕”变大以后，每个人都会分到更大一块。因此，穷人和富人都有理由让“经济蛋糕”尽可能大，而做到这一点就需要以支付意愿为基准进行分配决策。

航班超额售票时谁该多等一会儿

考虑一下超额订座的商业航班该如何分配座位的问题，就能更清楚地理解上述简单结论背后的逻辑。对商业航空公司来说，以超出某一航班载客容量的数量发售机票是惯例操作，这种操作可以减少因为旅客取消行程而导致的空座现象。虽然超额售票确实降低了成本，但随之而来的问题也十分明显：在一小部分航班上，实际前来乘坐的旅客数会比座位数量更多。

例如，假设有260个人来搭乘一架从纽约飞往洛杉矶的航班，然而

飞机上只有250个座位。这260人中就包括约翰和埃里克，约翰是第二个抵达登机口的乘客，而埃里克是倒数第二个。约翰是一名办公室保安，他急着赶去洛杉矶看望他病重的母亲。如果错过了这趟航班，他必须再等10小时才有下一趟航班可以乘坐，他愿意花350美元来避免错过这趟航班。埃里克是最晚到达的乘客之一，这并非因为他疏忽大意，而是因为他前来纽约时搭乘的那趟航班晚点了。他是微软的副总裁，正准备去夏威夷度假。虽然他可以乘坐1小时后在西雅图转机的一趟航班去夏威夷，但他愿意付1 000美元以避免等待。这两个人之中，谁应该获得本次航班的座位呢？

在1979年以前，这个问题的解决方式就是依据先来后到。约翰可以乘上飞机，而埃里克只能继续等待。不同于前面那个落地钟的例子，大多数旁观者似乎都认为这是个好结果，至少约翰希望准点到达洛杉矶的理由看起来比埃里克要紧急多了。

问题就在于，先来后到的标准并不是每次都会产生让人从直觉上感到满意的结果，而且即便在眼前的例子中也未必就产生了最好结果。也有可能埃里克是第2个到达的旅客，约翰才是第255个。考虑到许多随机因素会影响转机者的前一趟航班是否准点，我们没有理由假定先到达登机口的旅客就比后到的旅客有更重要的理由获得座位。

先来后到的标准并不是根据支付意愿来分配座位的，因此并不能使“经济蛋糕”最大化。这种分配方式使一些愿意为准时抵达目的地支付一大笔钱的人被迫等待，而另一些并不乐意花多少钱来避免等待的人却乘上了飞机。出于这一理由，美国民用航空委员会（在1984年之前负责管理商业航空公司的联邦机构）提出以支付意愿替代先来后到作为分配座位的基准。

美国民用航空委员会具体的提案是，让航空公司为超额售卖的航班的旅客提供现金赔偿、免费机票或其他补偿，吸引一些主动放弃本次航班座位的志愿者。所有预订了机票的旅客都有权利继续留在航班上，航

班公司因此必须不断提高补偿数额，直到有足够的志愿者站出来放弃座位。

联邦机构提出这类规定之后，社会公众与其他当事主体有60天时间可以发表意见。美国民用航空委员会刚公布议案不久，由拉尔夫·纳德（Ralph Nader）在华盛顿特区建立的组织航空消费者行动计划（Aviation Consumer Action Project，以下简称ACAP）就提交了一份措辞激昂的文件。ACAP似乎意在保护航班乘客不受航空公司压榨。在他们提交的文件中，ACAP反对称，若民用航空委员会采纳该项议案，等待下一班可乘坐航班的负担就会过多地压在穷人身上。ACAP担忧，就如其他领域中的情况一样，富人将会有权花钱避免生活中的各种麻烦，穷人却被迫承受它们。

ACAP的反对意见最开始在许多人心引起了强烈共鸣。然而，若仔细地思考，大多数人都会感到他们的说法十分愚蠢。如果不接受这种拍卖式的程序，就只能维持现状，而先来后到的模式会给穷人带来更不利的结果。例如，想要回家探望病母的保安员约翰若是第255个到达登机口的乘客而非第2个，依照民用航空委员会的拍卖式程序，只要他想，就有权一直待在飞机上不走；或者他也可以接受航空公司提供的补偿，放弃自己在这班飞机上的座位。然而在先来后到的规则下，他就一定会被迫错过这班飞机，继续等待。

在拍卖式的规则下，航空公司需要10位自愿放弃座位的乘客。假设约翰拒绝了他们200美元的最初开价，而有其他3个人接受了；航空公司接着将价格升到300美元，又有一名志愿者站了出来，但约翰还是留在原地；当价格到400美元时，约翰和另外5位乘客（微软副总埃里克并不在其中）都同意放弃座位，这架航班就可以起飞了。想象一下，当约翰就要接受400美元的赔偿时，ACAP的官员突然赶到，拿着一纸禁止航空公司举行这种拍卖活动的法庭禁令，叫停了拍卖。约翰很有理由会认为：为什么ACAP要通过剥夺他主动放弃座位的权利来“保护他的利

益”呢？他不是一定要接受赔偿放弃座位，这只是一个选项而已。他决定接受是因为在他的心里400美元的报价超过了保留原有座位的价值。若非如此，他就会继续拒绝航空公司的开价，留在飞机上。

在很多年里，我先后同几百个人讨论过上述例子，至今没有任何一个人觉得美国民用航空委员会应当接受ACAP的反对意见，并撤销其改革议案。拍卖之所以效率高，是因为它的机制就是以支付意愿为基准进行分配。以支付意愿为基准，确实意味着有时那些有要务在身的穷人不得不等待下一班飞机，而无所事事的富人却得以顺利地继续旅程。但那些主动放弃座位的穷人会得到一笔赔偿，这笔钱可以用来满足他们认为更加要紧的其他需求。约翰可能会晚一些见到母亲，但他多获得了400美元，可以把钱花在其他有价值的用途上。

对有些人来说，决定是否要主动放弃座位是一个艰难的选择。但美国民用航空委员会认为，让人们自行做出选择，总比遵循ACAP的建议、根本不给他们选择的机会要好。如前文所述，最重要的不是支付意愿能否产生完美结果，而在于是否存在一个替代方案能产生较之更佳的结果。

否定ACAP的反对意见并不意味着否定贫穷者面临的困境。前面这个例子也并不表明在任何分配问题中支付意愿都必须充当终极的判断标准。这里，我主要关注的是普通产品与服务的分配问题。谁该获得那只古董钟？谁应该等下一班飞机？诸如此类。

自愿的拍卖只是美国民用航空委员会可能提出的多种依照支付意愿分配座位的方式之一。例如，可以不由航空公司开价提供赔偿来吸引人自愿放弃座位，而是由航空公司把座位分给那些为了保留座位出价最高的乘客。与原来的提案无异，这种方法同样是依据支付意愿来分配座位的。然而，若民用航空委员会要求乘客为留在机上的权利竞价，这项提案肯定永远都不会获得人们的同意。

联邦机构中绝不只有民用航空委员会对这类分配问题十分敏感。就

算美国对于竞选的规定较为宽松，低收入选民在政治进程中仍然相当有发言权。这些人和为其发声的人会反对任何无视他们利益的政策。至于他们是否真的具有阻挠特定政策实施的政治力量，当然就要看具体情况了。

除了ACAP，民用航空委员会的自愿拍卖提案几乎没有遭到穷人或是其代言人的明显反对。一个重要原因是，这一提案赋予低收入乘客一项获得实际经济价值的权利，也就是让他们有权继续保有自己的座位，除非自愿为了获得赔偿而放弃座位。若这一提案没有预设乘客有权继续保有座位，它绝对会被人否决。如果仍然遵循先来后到的老路子，比起目前的自愿拍卖体系，只会对所有乘客都更为不利，不论贫富。

ACAP错误地认为若采用拍卖方式的话，被留下来等待之后的航班的一定是穷人。贫穷但需要准点到达的乘客，往往并不会放弃座位以换取赔偿，而更富有却并不赶时间的乘客往往会站出来接受赔偿。当然，低收入乘客确实比高收入乘客更可能选择放弃座位。但是，若他们这样做了，理由一定是他们能够获得的现金赔偿或免费机票的价值比起不让出座位能节约的时间的价值更高。

自由主义者对成本-收益分析的敌视

正如上述例子里清晰阐释的那样，对于有限的资源应该如何分配的问题，了解人们心中认为自己面对的选项具有多少金钱价值是极其有帮助的。不论在某一时刻人们的收入分布存在多么大的不均衡性，一项公共政策若是依照某种与支付意愿无关的标准来进行资源分配，总还能找到一个替代方案，使结果对所有人来说都变得更好。

由于对整个市场机制怀有不信任感，以及市场机制导致不公平现象十分严重且不断加剧，人们产生了担忧，许多自由主义评论家十分抗拒在公共决策中以支付意愿作为衡量标准。在美国民用航空委员会这个例

子中，依据先来后到本身是一种十分粗糙的分配机制，所以反对依照支付意愿分配会对低收入人士产生极大的不利影响，虽然那些反对人士声称自己是在为低收入者发声。若能精细地设计细节，任何一种无法反映人们愿意为眼前选项支付多少钱的公共决策，基本都能通过改造而变得对所有当事人都更有利。那些不愿意接受这种改造的人不妨想象一下，他们自己就像拉尔夫·内德，在那个希望能接受赔偿、改乘下一班飞机的低收入旅客面前，挥舞着一纸法院禁令加以阻拦。

为何收入转移往往是实行高效率政策所必需的

由于保守派们坚决反对收入转移，他们一直是高效率公共政策得以采纳的更大障碍。这是因为，任何一个民主国家若不进行收入转移，都势必要以成本更高昂的方式来满足低收入群体的利益。

早先那个选择公共广播节目形式的例子就为我们阐释了基本原理。回想一下例子的假设条件，从纯音乐节目转变为纯谈话节目可以为那位富人带来1 000美元的收益，但会给两位穷人分别造成100美元的损失。两位贫穷选民在人数上占优，他们能够通过投票阻止节目形式的改变。然而，改变节目形式能产生大于成本的收益，所以阻止改变对他们而言是不利的。若两位穷人能够允许节目改变形式，并因此为自己换取一份收入转移的话，所有人都会得益。如果这种收入转移不被允许发生，那么改变节目形式的提议就会被他们投票否决，相比能够实现节目改变的情况，穷人与富人得到的好处都变少了。

这样的例子还有很多。比如说，1979年的能源危机促使美国总统吉米·卡特提出要对每加仑（约3.79升）汽油征收0.5美元的税。这一提案意在通过价格机制减少人们的汽油消费，从而减小美国对进口石油的依赖性。提案刚公布，立刻就有反对的声音提出这会给贫穷家庭增加无法承受的负担。不过卡特本就预见到会有这样的反对意见，所以他同时还提出要减少工资税，由此造成的财政收入减少，就用汽油税收入来填补

差额。由于工资税是一种累退税，对这一税种实施减免可以缓解贫穷家庭因汽油税而面临的财务困难。

最终卡特的提案还是被否决了，一部分原因在于批评者错误地认为减少工资税会使人们失去节俭的动力。（这些评论家们没有认识到的是，即便工资税或许会让人们能够比以前购买更多汽油，升高了的税后价格仍会让他们很有理由不去这样做。）但卡特法案未获通过的另一个重要原因在于，它涉及收入转移。那时距离罗纳德·里根当选总统还有一年多，但国会对于收入再分配的反对声音已经在不断增多。

未能通过卡特提案的最终结果就是，不论穷富，所有人的经济利益都受到了损害。舍弃了卡特通过价格机制减少汽油消费的提案，政府便采用价格管制的方式来保护穷人免受高价之苦。结果所有人都因此负担了高昂成本：加油站前大排长龙，宝贵的旅行计划被迫延后，等等。若当初政府采用了卡特的提案，对富裕的公民来说，支付更高的汽油费用以减少排队等候时间并不是多大的负担。贫穷的公民也不必长时间排队，工资税的减少缓解了汽油价格提升带给他们的压力。

不论举多少个例子，背后的规律都是相同的。由于不能或者不愿通过税收制度来实行再分配，我们就只能用其他成本更高的方式来照顾穷人的利益。穷人住不起好房子怎么办？最有效率的解决方式应该是进一步提高所得税抵免额。所谓所得税抵免是罗纳德·里根推行的几个收入转移制度之一，然而近年来美国国会中的保守派们对所得税抵免的反对越来越强烈了。如果无法直接进行收入转移，替代方案便往往是一些会付出更高成本的方式，例如实行租金管制。

勤劳的农民们入不敷出怎么办？最高效率的解决方式是直接对他们进行收入转移。然而，若收入转移不被允许，政府就只能采用一种成本高昂得多的替代方法，比如对农产品进行价格支持。穷人付不起电话费、燃气费和电费怎么办？一样，最高效率的解决方法还是收入转移，但政府采取的方法是设置一个“生存必需”用度内的低费率，这不仅导致

人们使用这些资源或服务时不够节约，还让其他消费者支付了更高的价格。

不愿进行收入转移还导致了一个后果，就是污染条例豁免了老旧汽车。虽然洛杉矶市内已经使用15年或以上的汽车只占车辆总数的一小部分，但洛杉矶盆地的空气污染中超过一半都要归因于这些老旧汽车。驾驶这些车辆的大多数都是穷人，立法者担忧若不豁免这些车辆，会给他们造成无法承受的额外开销，因此污染条例就未将老旧车辆纳入管制范围。

老旧车辆排放污染的问题，可以通过向穷人转移足够的收入和帮助他们购买新一些的二手车来解决。这一策略能够以大约每千克33美元的成本减少碳氢化合物和一氧化二氮的排放。³然而，由于污染法案豁免了这些最老旧的车辆，加利福尼亚州立法机关只好转而推进电动车辆的发展，这一方案每减少1千克污染气体排放就需要大约2 000美元的成本。在这个例子中，政府不愿进行收入转移，这使不论穷富的所有人都承受了更高的代价。

经济学家们估计，如果能施行一种简单的拥堵费制度，就可以节约几十亿的高速公路建设费用，并且帮助避免交通拥堵。⁴只需对现有科技（如美国东北部目前许多公路与桥梁都在使用的一卡通系统）稍加改造并加以利用，就可以让每部车辆安装一个小型电子收款器，从而根据道路拥堵程度收取道路使用费。因此一个驾驶员若在中午高峰时间开过曼哈顿，10美元的额外通行费就会自动通过他车上的收款器来收取。

这一制度面对着一一种显而易见的反对理由，那就是人们担心那些不得不在交通高峰期出门的贫穷机动车驾驶者会承受过重的负担。但是没关系，通过拥堵费制度获取的财政收入以及因为这一制度而节约的其他成本，完全允许政府进行足够的收入转移来帮助穷人承受拥堵费负担。如果政府不允许进行这种必要的收入转移，那么拥堵费制度就永远都不可能被接纳。政府本可以省下很多高速公路建设费用和交通拥堵的

成本，但还是继续负担着这些成本。

这些绝对不是什么罕见的个例。相反，这是一种普遍模式：政府不愿为穷人进行收入转移，于是只能用其他方式解决他们的问题。即便拒绝通过税收系统进行收入再分配，穷人的利益在民主社会必须得到注意这一点还是不会改变。拒绝收入转移只不过迫使政府以成本更高的方式顾及他们的利益罢了。

通过政策进行收入转移有助于推行高效率的政策

虽然未能处理好分配问题时常让政府无法施行高效率的政策，但至少有些时候，政策制定者直接进行收入转移，为高效率政策的推行奠定了基础。有启发性的一例是，纽约州公共服务委员会提出对每一通来自私人电话的查号服务通话都收费0.1美元（在那时，电话用户进行此类通话无须缴费）。促使委员会提出这一方案的原因是，赋予人们免费查号的权利让电话公司不得不聘用数百名接线员帮助人们查号，然而打来电话查号的用户大多数其实都能很方便地自行查到电话号码。

这一提案无可避免地遭到了猛烈抗议。社会学教授以及其他在听证会上做证的人表示，必要的社区通信模式会因为拨号者无法免费获得查号服务而备受影响。眼看议案无法被通过的形势越发明显，委员会主席阿尔弗雷德·卡恩（Alfred Kahn）提出进行如下简单修改：每次查号服务0.1美元的收费仍然不变，但将原先的查号接线员调去其他更有用的岗位工作，由此能够节约一些钱，利用这笔钱可以为每个使用查号服务的用户减去其月账单中的0.3美元。因此，每个月使用查号服务不足3次的用户，电话费将会比原先更低。做出这一修改后，委员会的提案基本上没有遭到任何反对便获得通过了。

这个例子中涉及的数额如此微小，却显示出再分配问题在公共政策领域的影响之大。即便查号服务仍然免费，几乎也没有哪个家庭一个月

能够拨打超过10通查号电话。不足1美元一个月的额外开销对于最贫穷的家庭也实在称不上会对生活水平造成什么严重威胁。这一例子还表明，若能够将现金转移或其他功能相同的手段作为政策工具加以利用，政府就可以成功说服那些本来反对某些高效率政策的人。

农业领域为我们提供了另一个通过直接的收入转移实现高效率政策的实例。背景问题是，由于无法跟上企业化农业带来的迅猛的生产率增长，美国的家庭农场主们长期承受着经济压力。由于大规模扩大农业生产，现代农业已经使粮食价格降到很低，家庭农场主因此入不敷出。由于担忧这些农场主面临的困境，美国国会反复加大对农产品的价格支持的力度。农作物的价格支持政策诱导农民们花费了数十亿美元雇佣劳动力，购买机械、种子、肥料与杀虫剂，并为种植农作物进行其他投入，然而消费者并不愿意以高昂的保护价格购买这些农产品。政府只好收购剩余农作物，然后眼睁睁看着它们在经过长时间、高成本的仓储之后白白腐坏浪费。美国国会简直是以最为效率低下的途径来帮助家庭农场主。

然而，美国农业部至少还开展了另一项工作，以高效得多的方式来解决农民贫困问题。那就是农业部的土地休耕计划，根据这一计划，农民们若空出一些现有土地不进行耕种的话，就可以从政府处获得现金奖励。该计划为挣扎在贫困中的家庭农场主们提供了额外收入，却没有让他们生产任何无人购买的多余农作物。

土地休耕计划比价格支持的效率要高得多，却仍然引来了争议。我们确实很容易明白为何批评者会反对该计划，毕竟农民会因为什么都不种植而得到钱。然而，如果另一个选择是付给他们大笔钱款、让他们种植注定卖不出去而只能浪费的农作物，这一计划的优越性也就不言自明。无论如何，这个例子说明，就算政府无法进行收入转移，并不一定无法将一个效率低下的政策转变为不那么浪费的新政策。然而，随之产生的问题是，政府为何无法更多地进行这种转变呢？

采用高效率政策面临的阻碍

既然政府有时可以协商出高效率的解决方式，为何其他时候就要拒绝成本-收益分析的逻辑呢？一个困难之处在于，由于几乎任何需要公共决策的问题都至少要涉及一些分配问题，对每一个问题逐一实行补偿安排或许是不现实的。也许只有在那些能够很简单地完成补偿安排的情况下，高效率的解决方式才有可能得以实行，例如超额售卖的航班采用自愿拍卖模式，或者为使用查号服务的用户每月抵扣电话费。然而在许多问题中，显然并没有这类可以简单实现的补偿安排。

而且，即便在补偿安排可以轻松实现的时候，或许仍然会难以找到高效率的解决方式。比如说，租金管制效益很低，那为什么大城市的政府不给人们发放代金券，促使他们放弃自己那受到租金管制的租房合同呢？租客可以搬去那些面积更小、按市场价格出租的公寓，然后通过代金券的形式来获取相当于现金转移的补贴。

根据同样的道理，政府为什么不通过拍卖形式来决定监狱或其他不受欢迎的公共设施应该建在哪里呢？每个社区都可以提交一份封好口的投标书，在里面写明能让他们接受将公共设施建在自己这里的最小金额补偿。投标价格最低的社区将是建设地所在，社区可以得到一笔现金补偿，金额就相当于第二低的投标价格。⁵这笔补偿金可以通过向其他社区征收与其投标价格成比例的税款来获取，这样各个社区就很有理由诚实出价，而不往高虚报。

前面的讨论表明，似乎至少存在两方面的阻碍使政府无法通过使用成本-收益分析来高效地解决公共领域问题。其一在于，以激进的自由主义者为主的许多人从原则上就反对政府进行再分配。其二在于，几乎每天都有无数问题亟待解决，要逐一设计补偿方案是不现实的。

这两个阻碍都非常有说服力，然而任何一个都经不起仔细推敲。先看第一个阻碍，有些人对大众宣称：一切收入转移都是非法的，这一点

是原则问题，因为“那本来就是你们的钱”。这里所谓的原则就是，每个人都通过市场中公平、自愿的交换获取财富，因此有权依照自己心愿任意支配这些财富。然而，这一原则若按字面意思理解并不具有说服力，因为按照这一逻辑讲，政府就完全没有权力向公民收取任何税款。然而正如第1章里已经指出的那样，这样的政府将不具备任何能力组建军队，它的公民也迟早会被迫向外国侵略者的政府交税。

“那本来就是你们的钱”这种原则之所以不成立，还有一个重要的现实原因。这种原则的支持者们，大概是相信让人们尽可能高比例地直接控制自己创造的资源对他们而言更有利，所以才会支持这一原则。然而，如我已经论证过的那样，这一原则似乎并不值得这些人去支持，因为实际上依照这一原则行事会使人们最后掌握的资源总量变少。如前文所述，民主国家必须在以下两者中做出选择：要么直接将资源转移给低收入人群，要么通过其他成本更高的方式回应低收入人群的利益诉求。

那么，拉低效率的第二个阻碍又如何呢？政府每天面对无数公共决策问题，要逐一实施补偿措施应该是不可能的。对这一问题的答案是，根本不需要进行这种累赘操作就可以进行成本-收益分析。我并不是说，这类补偿是不必要的，而是说并不需要在每一个具体的公共决策问题上逐一进行相应操作。成本-收益分析的批评者们有一个看法是正确的，他们认为以支付意愿作为分配基准一定会导致许多公共政策都向高收入群体倾斜，因为高收入者能够支付更多钱，原因就这么简单。

然而，与其放弃成本-收益分析，政府不如做一个更好的选择：可以毫不歉疚地使用支付意愿作为分配基准，同时利用福利与税收制度对低收入家庭因此而承受的伤害进行补偿。如前文所述，这种补偿不需要而且也不可能针对每个具体问题分别进行设计。以支付意愿为基准进行成本-收益分析，然后通过一套能反映出这种分析方式为低收入者带来了多少损害的税收制度，让低收入者从税收制度中获得补偿就可以了。除了为成本-收益分析而进行的补偿，还有其他因素让福利与税务减免

成为必要，关于这些其他因素，我们将在下一章继续讨论。

THE DARWIN ECONOMY

THE DARWIN ECONOMY

08

“这就是你们的钱”：
人们的收入并不只是付出努力的结果

17世纪的英国哲学家约翰·洛克在《政府论（下篇）》中写道：“每个人都对他自己的人身享有一种财产权，而且这一权利为他所独享，除他之外的任何人都不能拥有。我们可以说，他以自己的身体做出的劳动，以及他用自己的双手进行的工作，都正当地归他所有。”

那种至今仍坚持视所有课税为盗窃行为的极端自由主义者，因为这番话而将洛克视为自己的智识偶像。尽管一些更为认真的自由主义者都无奈地承认至少有必要存在一定的税收，但不可否认的是，洛克的理论对自由主义者群体之外的人们也有着巨大的影响力。既然你通过自己的天赋和努力使一样东西变得更有价值，被添加的价值仿佛就应该是你的。政府又有何权利来主张要分一杯羹呢？

自罗纳德·里根在演讲中加以力推之后，人们对自己的劳动成果享有天然权利这一理念，在过去的数十年间愈发深入人心。20世纪80年代末，美国众议院筹款委员会前主席比尔·阿彻（Bill Archer）为支持一个对高收入家庭减税的提案而援引里根的说辞，他写道：“要把政府手里过量的税金归还到当初将这些钱交到首都华盛顿来的家庭和劳动者手里，这是原则性问题。”小布什曾提出一项共计1.3万亿美元的减税法案，该法案的受益者主要是全美国最富有的那些家庭，而他为这一法案做出的辩护着实令人难忘：“这就是你们的钱，是你们之前付的。”如果一个美国政治家今天要提议任意一项税收，他的反对者都会立刻严厉地指控他，称他“认为华盛顿的官僚比人民自身更懂如何明智地支配他们的钱”。

所有权的神话

然而，若要仔细审视“保留所有税前收入是人们的天然权利”这一观点，便会很快发现它并不成立。在工业化的现代民主国家中，人们的高收入并不仅仅是他们付出努力的结果。很大程度上，这些收入都得益于政府在基础设施、教育、负责个人私有财产权的界定与实施的制度体系等多方面的大量公共投资。人们总是很容易忽略，这种公共投资一直都在为社会的繁荣发挥核心作用。

数十年前，我曾有机会亲身了解，生活在一个未能获益于这种公共投资的环境之中是怎样一番体验。那时我刚从大学毕业，成为美国和平队志愿者，两年时间里在尼泊尔的一个小村庄担任数学与科学教师。不论在当时还是现在，尼泊尔都是世界上最为贫困的国家之一。

就同当地的其他和平队志愿者一样，我也为自己聘请了一位厨师。他的名字叫博克哈曼·拉伊（Birkhaman Rai），来自邻国不丹的一个偏远的喜马拉雅山村，几年前离开家乡来到这里。他一天学都没有上过，却是我见过的最足智多谋、多才多艺的人。

作为厨师，拉伊手艺娴熟，在集市上与人买卖时他又宛如精明的谈判家，而且他还懂得如何宰羊或修补茅草屋顶。他能刷墙，也会修理坏掉的闹钟。他既是个能干的锡匠，也是个熟练的木匠。我最喜欢的一双鞋穿破之后，他将鞋修复得不比任何专业修鞋匠逊色。他很快就成了村庄里修缮工作方面的权威。

由于拉伊不会读写，我离开尼泊尔之后便再没能与他联系。但如今的尼泊尔几乎就同我住在那里时一样基础建设落后，当时我每年从有限的志愿者生活津贴中付给拉伊的几百美元，很可能至今仍是他一生中的最高收入。

如果拉伊出生在美国，他或许早已成为富人。至少，以他的能力和决心，他理应能赚到比他在尼泊尔的收入高出几百倍的钱。

在美国，每年当4月15日的纳税申报期限临近时，富有的自由主义

者们都要愤怒地站出来指责政府在夺走理应属于他们的钱。他们最好能反思一下，不论他们有多聪慧勤奋，如果他们在尼泊尔或索马里这类国家长大的话，从一开始他们就绝无可能赚到可以被夺去的巨额财富。让他们有机会获取财富的基础设施，都是用税金建成的。他们获得的财富之中，很大一部分都是以其他人所纳的税款进行公共投资而取得的收益，然而那些纳税人却未能从这些收益中分享回报。

利亚姆·墨菲（Liam Murphy）、托马斯·内格尔（Thomas Nagel）等哲学家经过研究指出，“这就是你们的钱”的说法鼓吹了一种“关于所有权的神话”，使人们偏离了本应该思考的那些有关整个法律与社会制度体系的重要问题。¹因此，他们这样写道：

所谓“我们的钱”之中有多大比例可以允许政府通过税收手段获取的问题，从逻辑上是讲不通的，因为其实是包括税收体系在内的法律体系决定了“我们的钱”到底有多少。真正的道德问题在于，我们对管理财产权利的法律体系应当以何种目标进行何种设计。什么样的市场能够最好地鼓励投资、提高生产率？什么样的产品应当通过集体公共决策来提供，而什么样的产品又由私人选择提供？政府应当为每位公民都提供最小限度的经济保障吗？在多大程度上机会平等应当被公开支持？社会与经济上的严重不平等是否应该受到道德谴责？如果是，我们能够做些什么来正当地对其加以纠正？²

如果不讨论什么人应该缴纳何种税款，当然就无法回答上述问题。然而人们选举出的官员们，只要公开讨论任何一种税收，往往就会被右派人士用一句“这就是你们的钱”打断。无法进行这类讨论使所有人的财富都受到了损害。

要用税收体系为社会工程买单

在过去的—个多世纪里，人们激烈地辩论着以下两种方法究竟哪个更有利于经济繁荣：是集中计划管理，还是将大多数经济决策都下放给去中心化的私人市场进行。这一辩论已经有了结果。不论去中心化的市场体系有什么样的缺陷，苏联集体制的经验使大多数中立旁观者都已确信，一个成功的经济系统必须主要由市场体系及其变体来构成。

当然，实际上即便是受管制最少的市场经济往往也需要庞大的公共部门。除了维系一套能够保障私有企业与消费者关系运转的法律体系，公共部门还要为人们提供道路、教育、治安、消防以及国防等公共产品。税收体系必须为这些公共产品买单。

对某一行为征税不仅会带来财政收入，同时也会减少这一行为的发生。对存款、创造就业机会等有益行为征税，会使“经济蛋糕”缩小；相反，若对污染、交通拥堵等有害他人的行为征税，“经济蛋糕”便会相应增大。在第11章里，我还将详细阐释如何从对有益行为征税转变为对有害他人的行为征税。

这种转变能做大“经济蛋糕”，让每个人都有机会分到更大一块，这相对而言理应没什么争议。然而，即便是这种转变，也会遭到来自极端自由主义者的反对。“社会工程！”他们齐声尖叫道。他们口中的社会工程是指任何意在“控制人们的行为，左右人们的选择，改变人们的生活方式”的尝试。举例来说，以减少对进口石油依赖为目标的汽油税就总是遭到他们的这般指责。

然而，正如我在第1章中所说的，他们的反对理由非常无力，因为基本上任何法律与规范都构成一种社会工程。禁止杀人与盗窃的法律控制了人们的行为，左右了人们的选择，改变了人们的生活方式，没错，这就是一种社会工程。噪声禁令、交通限速也一样，甚至停车标志和红绿灯也是如此。要消灭社会工程，就只有采取无政府主义了。比起法律与习惯性规范，以税收来抑制有害行为的成本要低得多，且手段也更柔和。税收主要对那些最容易改变自己行为的人产生影响，通过他们的自

我改变来抑制有害行为，关于这一点我还会在第11章里详细加以阐释。

在这里我们要讨论一个税收政策的话题中更具争议的维度，那就是税收政策对税后收入分配产生的效果，不论是有意为之的效果还是无心栽柳的效果。什么样的收入分配是公正的？如果政府有责任实现公正的收入分配，那么其究竟要承担何种责任？几乎没有其他问题比这些更受哲学家、经济学家、法律学者等人士关注。³然而，这些问题远远超出了此处讨论的范畴。我关注的是“任何情况下人为地进行收入再分配都是不正当的”这一经常由自由主义者和保守主义者提出的观点。但事实上，不论一个人对前面那些更为宽泛的问题持何种意见，都很难为这种观点辩护。

例如，正如我们在前一章里看到的那样，拒绝将富人的收入转移给穷人时常导致民主政府被迫以其他效率更低下的方式来满足低收入选民的利益诉求，这种效率低下对贫富双方都有弊无利。宣称你有权利保有自己100%的税前收入是一回事，但如果行使这种权利会让你与其他所有人都更贫穷，一定要这样做就说不通了。但自由主义者反对带有再分配效果的税收的理由，还存在其他更为关键的问题。

高效率地供应公共产品

一个重要的现实考虑是，若不愿实施收入再分配，政府就无法向高收入的公民供应他们重视且乐于购买的公共产品。就同私人产品一样，人们对公共产品的支付意愿也随着收入增长而高涨。富有阶层比贫困阶层更倾向于重视公共产品的价值，这并非因为财富使人改变品位，只是因为他们有更多的钱可以花费而已。

一个只由富人构成的社会将比只由穷人构成的社会选择为社会成员提供更多、更好的公共产品。但当各种收入水平的人在同一个社会里生活时，公共产品的本质又要求其供应给每个人的数量与质量是相同的。

一个社会必须做出决定，不同收入水平的人应当被征收多少税款以购买这些公共产品。一些简单的例子可以帮助我们解释为何采取累进税制常常是有益于各种收入水平人群的可行做法。

假设两位公民，兰德与保罗，分别拥有位于湖畔的两座相邻的夏季度假别墅。由于近期湖中出现大量入侵物种斑马贻贝，他们两人都不得不每周向自家的水管中加氯，以防水管被这些小型软体动物堵塞。某个厂商向他们介绍了一种新型滤水设备，能为他们省去每周加氯的麻烦。一套设备总价需要1 000美元，可供两家同时使用。两个房主都有同样强烈的购买意愿，然而收入为保罗两倍之多的兰德最多愿意为这台设备支付900美元，保罗则至多愿意出300美元。

他们都不会单独购买这台滤水设备，因为设备对他们任意一人的价值都要低于售价。然而两个人一起购买、共享使用的话就是划算的，因为设备对他们的价值之和共计1 200美元。这样一来，他们两人的经济剩余合计是200美元，这是通过二人共计获得的滤水设备的价值减去购买设备的成本而得的。而且当“经济蛋糕”做大后，任一当事人都一定有机会分享到比从前更大的一块。

既然分享滤水设备对兰德和保罗而言都是有效率的，我们不妨假设他们很快就达成一致意见，准备一起购买。然而，联合购买、共享使用的提案总是说起来容易做起来难。正如科斯所指出的那样，仅仅是商量联合购买事宜便会产生成本。只有两个人时，这种商谈的成本或许不会过于高昂，但如果成百上千人加入类似的讨论中，沟通的成本便很容易高到令交易完全无法进行。

当人数众多时，所谓的“搭便车”问题也会相应出现。如果一项工程成功还是失败不取决于任意一人在其中做出的贡献，那么每个人都有动机省下自己的力气，或者说“搭便车”，然后寄希望于别人会做出贡献。到最后，即便只有一小部分人实际参与贡献，要就人们如何公平地负担成本达成一致意见仍然非常困难。例如，兰德和保罗或许都不愿意透露

他们心中实际愿意为滤水设备支付的金额，其原因就类似于你不想告诉卖家自己最多能接受以多少钱买下他的某块东方地毯。

这些实际问题可能会促使我们授权政府代为购买公共产品。然而即便如此，我们仍然需要就付款方式达成政治协议。

我们假设，兰德和保罗请政府作为他们的代理人购买滤水设备，而政府的税收政策必须遵循“非歧视性”规则，任何公民为公共产品支付的费用都不得高于其邻居所支付的数额。

若只有经过大多数人的同意才能供应公共产品，一个受到上述规则制约的政府就不可能为兰德和保罗提供滤水设备。由于必须实行人头税（一个向每人征收相同数额税款的税种），兰德和保罗每个人都必须为滤水设备支付500美元。然而对保罗来说，滤水设备至多值300美元，他一定会反对这个提议，于是“大多数人都同意”的条件就无法满足。所以，若只允许实行人头税，民主政府一定无法为他们提供滤水设备。

只要人们的收入差异显著，纳税人内部对同一公共产品的估价往往也就会差异显著，而上述例子所体现的问题在这种估价差异显著的情况下一定会存在。对所有人征收相同税款的规则，会导致政府无法为人们提供很多有价值的公共产品。

现在再假设政府可以对人们的收入实行累进税率。如果收入为保罗两倍之多的兰德是个自由主义者，他可能会痛心地抱怨自己要缴纳的税款比保罗高出许多是不公平的。然而，转念一想他会发现，若他提出反对意见，就会因为缺乏“大多数人的同意”而得不到滤水设备的供应。还是让政府收他750美元、收保罗250美元，然后为他们购买滤水设备，这样来得更好些。在这种税制下，他只需要付出750美元就能够享用对他而言价值900美元的滤水设备。选择接受累进税率，他就能为此获得150美元的经济剩余。而保罗则以250美元的钱换来了价值300美元的滤水设备，他也获得了50美元的经济剩余。

要求每位公民为共同消费的产品支付完全相同的数额或他收入中的相同比例是十分荒唐的，我们不妨想象一下，一对夫妻若要按这种规则来生活会发生什么，这样我们能更清楚地了解这一点。例如说，朱莉每年能挣200万美元，而她的丈夫布鲁斯的年收入只有2万美元。鉴于他们的收入差异，比起布鲁斯，朱莉会想要花更多钱在住房、旅游、娱乐以及其他夫妻一同消费的项目上。但若夫妻二人必须平摊这些消费项目的开销，由于布鲁斯收入有限，他们就只能住在狭小的房子里，把他们的孩子送到教学水平不合格的学校就读，在度假、娱乐、享用美食等方面百般节省。这样我们就不难理解，朱莉为何会乐意为她与布鲁斯共享的消费项目分担超过50%的费用。

这个例子为我们说明了累进税率的合理性，而这种合理性完全是基于对现实问题的考虑。作为一种手段的累进税率能够同时帮助穷人与富人更充分地达成各自的目的。这本来就与公正、平衡等道德问题完全无关。但是，对自由主义者而言，他们支持累进税率的重要原因往往就在于这些道德问题，而自由主义者与保守派对累进税率提出批评也正是拒绝承认这些道德上的考虑。然而道德问题与实用问题常常不如人们以为的那样泾渭分明。

累进税率

科斯的见解又一次为我们思考旧有问题提供了富有成效的新方法。等级制之中的排序就如同他在作品里集中讨论的外部性问题一样，是人和人之间相互作用的。换言之，若有人在某一集体中排在了较高等级，就一定同时也有人要处于低位。人们在任意集体中对更高排序的追求，导致了一种我称之为“地位外部性”的结果。如果这些希望身居高位的人要成功达到目的，就必须有其他人居于低位并承受居于低位的成本。上述结论，为我们提示了一种站在自由主义者的视角之下支持累进税率的理论。

为了概略地为大家解释这种理论，我们不妨从一个假设条件非常偏向自由主义者立场的思想实验开始讲起。想象一下，你和其他999个人经历了一场摧毁全部现有社会架构的大洪水，刚刚从诺亚方舟上走下来。你的任务是和你选择的其他一些人组成新的社会。如果其他人希望构建的社会令你不满，你没有必要加入他们；但你也不可以强迫他人加入你的社会，除非他们认同你的计划。简而言之，在这个思想实验里，所有新建社会都必须完全由相互达成一致意见的成员组成。

在一开始，每个组建的社会都可以按比例分得一部分洪水尚未吞噬的土地与其他物产。例如，你和其他99个人决定共同组建一个社会，由于你们社会的人口占世界人口的10%，最初分得的就是世界上现存土地和其他物产的10%。每个社会最初分得的财富中并非每一项都令人称心，而人们形成社会后的任务之一就是要就社会内部各成员之间的财富分配方式达成协议。

假设人的天赋与性格都是完全可知的。你和其他一些人联合组成一个社会后，你就可以依靠自己的努力与能力去自由赚取相应的财富，但你同时也必须遵守你所在的社会决定采纳的任何税收法律与其他规范。

稍加思考便可发现，那些对你来说至关重要的结果都取决于同意与你一同组成社会的人是谁。如果你的社会中大部分人都有较高的生产率，那么你的社会就能比那些由生产率低下者构成的社会购买更多、更好的公共产品。但如果同你居于一个社会之中的大多数人都远比你生产率高，对你而言也会有负面影响。

举例来说，假设你和大多数其他人都想要周围风景优美的房子。然而，并不是所有房子周围都有优美的风景。很不幸但非常可能发生的情况是，你的社会采用了类似市场机制的模式，在社会成员内部分配土地和其他物产。生产率最低的社会成员将无法住上带有景观的房子，这一点基本确凿无疑。

也许房子有无景观对你来说并不重要，但你很可能的确会在意能否

将你的孩子送入一个好学校读书，使孩子将来能在好大学的招生活动或是理想职业的招聘活动之中成为一个有力的竞争者。然而，好学校也是一个相对而言的概念。所谓好学校，就是相比同一范围内其他学校而言更受青睐的学校。

在很多国家，最好的学校基本都位于最富裕的街区。在有些地方，这是因为学校预算的一大部分都来源于财产税。但即便是在每所学校学生人均分得预算额相同的情况下，教学质量最好的依然是那些处于最富裕街区的学校。一个原因在于，学校质量的重要决定因素是学生质量，而最为成功的家长从孩子入学前到就读后都一直为孩子提供优渥程度远超常人的学习条件。正是因此，供这些学生就读的学校，即便预算同其他学校相同，也有更好的教学氛围。

如果孩子的教育对你来说十分重要，那你不论在哪个社会之中都需要购买一套好学区内的房子。我们再次假设，大部分社会中，私人市场都是最主要的分配机制，那么你若加入了一个大多数人的生产率都远高于你的社会，你能够将孩子送进一个相对较好的学校的机会就非常有限。

于是，你在做出加入社会的选择不考虑你相对于其他社会成员在生产率排行榜里居于何种位置。最好的结果当然是成为一个高度繁荣的社会中生产率最高者之一，这样你将同时享受绝对与相对意义上的财富。但对大多数人而言，这都是不可行的。在任何社会中，其他人的平均生产率上升必然导致你在其中的排名下降；相反，当其他人的平均生产率下降时，你的生产率名次就会上升。

因此，生活在一个平均生产率更高的社会带给你的好处，会被你在这样一个社会中收入排名较低带来的坏处所抵消。其他条件一致的情况下，大多数人当然会希望占据更高的社会地位。然而从数学上来说，不可能一个社会中的每个人都是社会地位较高者。于是问题变成，为什么还有人愿意加入一个他会成为其中地位低下者的社会呢？

如果人们决意避免社会等级较低的情况，他们可以与生产率同自己刚好持平的人一起组建社会。最有能力的那群人可以形成一个人人都最有能力的社会，能力其次的可以形成一个人人都能力其次的社会，以此类推。在这些社会中，没有人会成为社会地位更低的人，但这种解决方式也会让人错过一些互利交易的可能性。例如，一个社会若全部由最具天赋的人构成，谁又会去清扫大街呢？而在一个所有人都最缺乏能力的社会里，脑外科手术又要由谁来做呢？

抛开这些实际困难不论，纯粹的层级同质化社会也会使人们失去本能够从交易中取得的潜在好处，而这种好处只要当一些人比另一些人更在乎社会排名时就会出现。如果高排名对一些人而言的价值超过了低排名对其他一些人的成本，能力各不相同的人们可以组建出一个让人人都比其在层级同质化的社会中受益更多的混合型社会。

例如，假设兰德可以保留他所有的劳动成果，那么他一星期就能创造并赚取10 000美元的价值，而保罗的相应数据仅为5 000美元。他们二人都愿意至多支付各自收入的三成以换取社会中较高的地位，因为这一地位可以帮助他们购买带有景观的房屋或是将孩子送进好学校；同时，只要能获得相当于各自收入三成的赔偿，他们也都愿意为此安于一个较低的社会地位。换句话说，我们假设兰德在以下三种情形下都是同样满意的：每周赚得7 000美元并居于社会高位；在同质化社会中每周赚取10 000美元；每周赚取13 000美元并充当社会中最不具有生产率的成员。在最后一种情形中，他不会有机会拥有一套附带景观的房子，但多出来的钱可以让他购买其他有吸引力的东西。至于保罗，他在一个社会中身处低位但每周挣6 500美元，将会和他在一个所有人都挣5 000美元的社会里一样开心。

在这些条件之下，不论是兰德还是保罗，与对方合作都会比他们各自同能力与自己完全持平的人组成社会要好。如果他两人结盟，每周的总收入（将二人各自的产出相加）将会是15 000美元，且显然兰德会是

社会中生产率较高、地位更高的成员。假设兰德分得其中8 000美元，而剩余的7 000美元属于保罗，这样一来，相比他们各自和相同能力者组建社会并获取与其产出相对应的财富，兰德的经济剩余为每周1 000美元，而保罗的经济剩余为每周500美元。

但如果兰德毫不介意自己是否在社会中占据高位呢？那他最好的选择是加入一群比他更有生产力且在意社会等级的人。他在集体中作为低位成员的存在将会成为一种财富，因为正是他处于低位才让其他人能够身居高位，他将会为此得到相应的补偿。

由此，我要主张的观点是，即便在一个为了照顾自由主义者的感受，让他们远离任何可能受到的强制而特意构建的情境之下，自愿形成的社会依然会采用累进税率。在这个思想实验中，累进税率的累进程度将由一些人愿意为社会高位付多少钱以及另一些人愿意为多少钱同意居于社会低位来决定。这些实际操作的问题，我稍后再继续分析。现在我们需要注意的是，在思想实验中没有一个自由主义者有理由抱怨这样的税收体系。重点在于，如果我们禁止使用累进税率，导致的结果便是每个社会都只能由能力层级一致的人组成，这对所有人来说都不是最理想的安排。

上述思想实验支持累进税率的依据，并未直接引入关于公平、公正的考量。每个社会的税收结构之形成，都是自由、独立的个体之间以自利为目的进行交易导致的直接结果。

然而，如果我们能够同意，在社会等级中更高的地位是有价值的，对上述思想实验的解释就完全不与传统讨论中的公平、公正问题相矛盾。前面已经提到，一些人能在社会中占据高位的唯一途径便是让另一些人处于低位。在任何一个由能力不同者组成却又未采取累进税率的社会里，占据社会高位的人便是在免费享用一项宝贵财富。这一财富之所以存在，正是因为其他人负担了处于社会低位的成本，使他们占据社会高位成为可能。就通常意义上的公平、公正而言，有些人不付出任何代

价就利用他人承担的成本来获益，这显然可以说是失公平或公正的。

劳动力市场中高排名的隐性市场

这个思想实验或许看起来有些天马行空，但一种与之相近似的情况实际上经常发生在人们身边高度竞争的劳动力市场中。在美国或是大多数其他民主国家，任何人不得被强迫为某一公司工作。我们也大可推断，当其他条件相同时，多数人更愿意处在一个比其他同事更高而非更低的排名上。从这两个假设可以得出一个清晰的预测：一个由能力不同者组成的工作团队绝不可能保持稳定，除非在团队内部，报酬能由生产率更高的成员转移给生产率更低的成员。这其中隐含的推理逻辑与刚才讨论的思想实验完全相同。

为了验证这一预测，我们可以比较一下工作团队内部的报酬模式与传统劳动力市场理论下预期的报酬模式，后者推定地位排名是无关紧要的。根据这类理论，劳动者所获报酬是他们劳动产出的市场价值。在现实市场中，报酬确实随着劳动者自身生产力的提高而提高，但并非产出提高了多少，报酬就会随之提高同等数额。例如，一个房屋构架团队中最能干的木匠，产出或许比他最不能干的同事多出一倍，但他的工资几乎不会比后者高出三成以上。

因此，我们在任何工作团队内部看到的工资分布，一般都要比个人生产率的分布更为紧密。若想亲自观察这种规律，你不妨想想在自己所在的组织中那些从事相似工作的共事者们（打个比方，你若是一个大型律师事务所的初级律师，那就想想与你同组的其他初级律师）。第一种情况是，你们失去了两个生产率最高的同事；第二种情况是，三个生产率最低的同事离职了。哪一种情况对你们团队造成的损失更大呢？大多数人一定毫不犹豫就会认为是失去两个生产率最高的同事的情况。

假如真是如此，依照传统劳动力市场理论，他们二人的工资之和应

当比那垫底的三人的工资之和更高。然而现实中往往相反：在一个团队完成类似任务的人中，任意三人的工资之和都会比任意两人的工资之和明显更多。简而言之，典型的规律是，一个团队中生产率较高的成员所获报酬比他们创造的价值要少，而最不具有生产率的成员获得了高于其创造价值的报酬。

我的看法是，这一规律反映出工作团队内部有一个高位排名的隐形市场在运作。这一看法若正确，就可以简单地回答一个原本十分恼人的问题：为什么在一个工作团队中最富生产率却又看似未能获得足够报酬的员工，并没有很快被与该团队相竞争的雇佣者以更高的薪酬挖走？虽然他们当下的薪资比传统劳动力市场理论所预计的要低，但他们其实还得到了其他一些对自己而言很有价值的回报。他们留在原来的团队里，正说明了他们所处的高地位完全能够抵消报酬上的不足。同理，团队中那些生产率低下的同事虽然感到排名居于最底端是恼人的，但仍然选择留在这里工作，就说明他们从高薪酬中获得了充分的补偿。

于是，私人劳动力市场对每个工作团队内部薪酬分配的影响，从效果上来说就相当于一种隐形的累进税率。在美国，劳动合同都是自愿签订的，自由主义者若要反对说这些收入上的转移违反了任何人的权利，就非常荒谬了。

高排名值多少钱

高排名的劳动者无形中转给低排名同事的收入，到底有多少呢？或者我换一种方式提问，生产率最高的劳动者为他的雇主创造的价值，到底比他的薪酬要高出多少呢？生产常常是一种复杂的团队活动，个体的生产率差异一般都极难估量。但在有些情况下，我们仍然可以对劳动者为雇主创造利润的个体差异估计出合理范围。

下面以我自己工作的大学为例，阐释如何在特定情境下完成这种估

计。与我同任教职的同事们有两个主要任务，其一是教学，其二是学术研究。每人在每个领域中的贡献都是复杂且多维度的。即便如此，我们也有可能粗略估计出个人总体贡献的差异至少大到何种程度。

例如，虽然学术研究的总体生产率极难计量，其中的一个方面却很容易通过金钱进行量化，那就是看研究者能够从除大学以外的来源获取多少拨款和其他财政支持。一个教员从政府机构或私人基金获取一笔研究经费后，钱通常会被分配给直接的与间接的两类开销。预算中分给直接开销的钱，用来购买实验器材、物资，并负担其他由开展研究直接导致的花费。而预算中分给间接开销的钱，就要用来帮助支付图书馆、公共事业、信息网络、行政与其他支持性职员，以及清扫积雪等方面的花费。

眼前问题的重点在于，需要由预算中的间接开销部分来负担的花费是基本固定不变的。一名教授获得了一笔经费，并不会导致校长的薪水、图书馆的花销、清扫积雪的费用或是其他日常管理费用相应提高。这也就意味着，如果大学新聘请的一位教授从他带来的研究经费中每年贡献出100万美元用以分摊间接开销，大学的年预算就会相应增加100万美元。换言之，相比让这位教授留在他原本那所学校，雇用他以后，这所大学的预算剩余将是每年100万美元。

高质量教员的劳动力市场竞争极其激烈，为了吸引到那些持续贡献大量资金给间接开销的教员，各所大学都在努力相互争抢。依照传统的竞争性劳动力市场理论，生产率的内部排名无关紧要，劳动者对雇主贡献的每一美元都会相应地反映到薪水数额中。这一模型的预测结果是，一位能为大学每年稳妥地带来100万美元用以分摊间接开销的教员，将会比另一位其他条件都相同但无法分摊任何间接开销的教员，薪酬高出100万美元。例如，假设一所大学试图聘请其他大学里一位很擅长获取研究经费的明星教员，且为他开出的薪水相比他现有报酬的涨幅要小于100万美元，那么他目前所在大学一定会为了留下他而将他的薪水涨到

和对方开价持平的数字。

然而，我们发现教员的实际薪水同他通过这种计算方式得出的生产率之间其实完全没有这么紧密相关。例如，在康奈尔大学的某个科学学部，教员对日常管理花费的长期平均贡献每增加1美元，他的薪水只会随之增加9美分。⁴

假如能分摊最多间接开销的教授在其他方面比其同事要生产率低，那么他们所获薪酬相对于他们对间接开销的贡献而言看似明显不合理，或许只是个错觉。也就是说，那些获得最多研究经费的教员，或许是因为他们教学能力不足，又或许是他们在其他方面对大学的贡献不够有价值，才没能得到与其经费获取能力相称的薪水。这是一种说得通的解释，毕竟教授的时间精力十分有限，在教学与研究两者之间很难不加取舍、完美两全。

然而证据显示，获取经费最多的教授往往也在其他重要方面比其同事生产率更高。例如，经费的发放一般要考察研究者的学术声誉，所以那些获得最多经费的人正是那些对大学名望贡献最多的人，这种贡献与其所获经费的多少无关。证据还显示，平均而言，作为研究者生产率更高的人，往往作为教师生产率也略高一些。⁵所以，个体之间对间接开销的分摊数额之差异，基本可以准确地反映整体生产率的差异。然而薪水的差异相比间接开销分摊数额的差异，实在是微不足道。

这一规律非常普遍。在每一个可以由数据支撑相关比较的行业中，任何团队中生产率最高的劳动者，薪水都比他们贡献的价值要小得多，而生产率最为低下的劳动者则恰恰相反。⁶其他一些因素当然可能会影响我们观察到的这种薪水与贡献价值之间的出入。但即便将这些因素都剔除出去，仍然会出现显著的工资压缩现象。⁷在假定人们认为工作团队内部的高排名很有价值的情况下，这正是我们会预测到的结果。

由于较高排名只有在其他人同意处于相对的较低排名的情况下才能存在，高排名显然有其高昂的隐性价格。从效果上看，每个雇佣者都在

进行一种隐性的收入再分配，向更具生产率的劳动者收税，并将他们的收入转移给生产率低下的人。

那些特别重视高排名的人，更倾向于接受作为一个平均生产率相对较低的团队中最具生产率者所必须承担的隐性工资罚金。相反，那些基本毫不在乎排名的人就会甘心在平均生产力较高的团队中居于底端，以此换取隐性工资奖励。至于那些对排名在乎程度中等的人，最好的选择则是在一个生产率适中的团队里居于中等排名，这样他们所获薪酬会大致等于他们产出的价值。

实际上，针对高排名的隐性市场运作方式非常类似针对安全性、自主性等职业特征的隐性市场。那些特别在意安全的人，会选择最为安全的职业并承受工资上的惩罚，以此避免另行采取安全措施的成本。那些将自主性看得很重的人，也会选择给他们最大自由度的工作，并承担工资上的相应惩罚。而格外看重高排名的人，会去选择那些可以让他们居于高排名的工作，并且愿意接受一般来说数额甚高的工资惩罚。

就如同针对其他可取职业特征的隐性市场一样，高排名的隐性市场的运作也在很大程度上是公司与劳动者在主观上意识不到的。受影响的当事各方还未进行任何艰难的、代价高昂的磋商谈判，税收和转移就已经发生。公司只不过公布出他们的薪资安排，而人们只不过选择最适合自己的职位。存在于工作团队内部的高排名隐性市场是难得的一例，在这个例子中，交易成本没有阻止人们自行找出解决外部性问题的有效方式。

工作场所之外的社会比较

在同事中的排名并不是唯一重要的社会等级。人们也会在意自己的收入与工作单位之外的人相比如何。例如，有研究发现，有姐妹做家庭主妇的女性，在其姐妹的丈夫比自己丈夫收入更高的情况下，她出门找

工作的概率会高出16%至25%。⁸与邻居的收入相比如何很重要。⁹与朋友的收入相比如何也很重要。确实，他人的消费会成为塑造我们消费标准的参照物，所以社会共同体中任何一个人的收入多少都不是无关问题。

工作团队内部较高排名的隐性市场，对于在工作场合之外解决社会地位的外部性问题毫无帮助。恰恰相反，由于这一市场让那些在工作团队内部占据高排名的人牺牲了很大一部分他们本应获得的报酬，实际还会导致这些人在工作场合之外面临更加严峻的外部性问题。

不过，对任何人来说，要在一个社会共同体中同陌生人协商处理社会地位的外部性问题，都比在私人工作团队内部解决类似问题艰难得多。正如前文所述，解决工作团队内部的地位外部性问题基本不需要经过任何复杂的协商过程。公司只需要公布一份相对于生产率差异分布更为紧密的薪资计划，然后让劳动者选择最适合自己的职位。相较之下，要在工作场合之外和其他人商量类似的问题就会代价高昂且极为困难。

任何问题只要涉及会祸及他人的活动，科斯的理论都是支持自由市场的保守派们在相关辩论中时常引为己用的权威理论。但科斯的中心观点并不是认为政府不应介入以削减有害他人的活动的危害，而是当过高的交易成本使人们自主协商以解决这些问题变得不现实时，政府应当制定法律、界定财产权利，从而使实际结果接近人们能自主协商时会达成的一致意见。

我认为将富人的收入转移给穷人的税收系统正是可以运用上述理论的一例。¹⁰这就像几乎在每个私人劳动合同中都隐含着的那种转移，它体现了社会层级不同带给人的成本与收益。这种税收系统可以促使形成一个更稳定和多元的社会。

高社会排名给他人施加的成本

我为前文这一思想实验设计的细节，意在建构出一个即便在最极端的自由主义者看来也不可能侵犯到任何个人权利的环境。打个比方，一个人若是想骑摩托时不戴头盔，他绝不会被迫加入一个要求他必须戴上头盔才能骑行的社会。而且，也没有人能强迫他加入一个会对他不戴头盔骑摩托的行为征税的社会。

我们的思想实验得出了惊人的结论：即便在这样一个环境中，自由主义者也没有理由拒绝加入一个向富人征税、将其收入转移给穷人的社会。如果他们足够在意高社会排名，他们就会更愿意被征收税款而不是被一个让他们占据高排名的社会驱逐出去。如果他们不在意社会排名高低，他们会发现加入一个自己处于低排名的社会就能获得更高的税后收入。

许多富有的自由主义者感觉自己理所当然就有一种权利去享有在社会等级中的高排名，对这一点我们并不惊讶。但如果交易成本足够低、社会可以自由组建与解体的话，他们就不可能企图免费获得这一权利。他们将不得不为这一权利而与人进行交易。从工作团队内部的高排名交易这一类似过程中，我们可以知道高排名的价格绝不便宜。

一个有钱人也许刚从屠夫的冷藏库里拿走了三块牛排，却想只付其中一块的钱，他没有权利这样做。一个有钱人也许还想不花一分钱就在社会等级中占据一个高排名的位置，他同样没有权利这样做。

机会在召唤

试图纠正市场失灵的努力从一开始就一直是人类社会的中心任务，虽然人们一般难以注意到这一点。不幸的是，政治光谱左右两端都对这种努力抱有错误认识，导致许多努力都未能成功。

左派所推崇的规制性救济手段，建立在“市场失灵是由经济寡头的剥削行为导致的”这一观点之上。右派则会反对称市场仍然有充分竞

争，有亚当·斯密的“看不见的手”进行调节就足够了，不需要政府的规制。虽然市场确实是高度竞争的，但保守派们对“看不见的手”的笃信却过于夸大了。

正如达尔文清楚认识到的那样，自然中不受限制的竞争时常无法促进获取共同利益，但这并不是因为垄断剥削。这只是个人利益与集体利益之间时常发生严重分歧导致的结果而已。人类以外的动物对此基本束手无策，早期人类虽然有更为优越的沟通技能，也并不比其他动物要好多少。但自从工业革命以来，人类社会取得了重大进步。爆发式材料革新的发生，几乎完全依靠拉近了个人利益与集体利益之间距离的全新社会制度。

然而，还有无数宝贵机会我们仍未充分利用。“这就是你们的钱”的话术在近年的政治辩论中大为流行，导致人们甚至无法讨论税收和那些让所有人受益的财政支出政策。在这一章里我已经阐明，对分配问题的误解导致人们未能成功地抓住许多宝贵机会。

坏消息是，反对者们在分配问题的辩论中以一种救世主心态对各自的立场固执己见。然而我们看到，以其他方式展开这场辩论会带来更好的前景。要支持更平等的资源分配方式，不仅可以通过抽象道德理念来辩论，还可以通过这种分配方式带来的互利效果来阐释。正如我下一章主张的那样，如果我们能舍弃那些并无依据的关于成功与价值之关系的成见，人类社会还能取得更大的进步。不过好消息是，由于当前的安排实在是极其缺乏效率，我们将会有很大的进步空间。

THE DARWIN ECONOMY

THE DARWIN ECONOMY

09

成功与运气：
成功对于运气的依赖比多数人想象的要大

人们经常会谈论感性大脑和理性大脑，就好像我们的身体里住了两个不同的人，一个由理性主导，另一个由情绪主导。但神经科学家认为，大脑的情绪回路和认知回路的相互依赖性极强。¹

在某些特殊情况下，大脑甚至不需要经过明显的认知过程，就能产生强烈的情绪反应。例如，一旦看到与蛇相似的物体，在画面传达到大脑的认知回路之前，人就已经产生了恐惧。²但是，更普遍地来说，人们的情绪反应很大程度上取决于在认知时如何对事件进行构建和解读。

打个比方，要想判断某个人的行为是否属于会引起他人愤怒的冒犯之举，我们必须首先了解该行为发生的环境，以及该环境所适用的规则。在熟悉的环境中，我们的情绪反应通常是非常精准的。但异常的情绪反应并不罕见，尤其是在不熟悉的文化环境中。

有一年休假，我和家人在巴黎的时候就发生过一次类似的事情。海登，我最小的儿子，当时正在城里的一所法国学校上五年级。有一天下午，他回到家的时候怒气高涨。他说他收到了一则学校的纪律警告，但事情本与他毫无干系。这个处分是由操场管理员提出的，理由是有一位学生对这个管理员喊脏话。

管理员认不出具体是哪位学生喊的，于是就指控了所有在附近玩耍的学生，其中包括海登。海登坚称自己没有对管理员说过一个字，他希望我们要求学校召开听证会，以便澄清真相。我咨询了几位朋友，他们告诉我说学校是不可能进行调查的。我还了解到，除非在同一学年里，学生已经累计收到过3次警告，否则这次警告不会造成任何影响。

我向海登解释，法国学校的公平体系似乎跟我们原来在美国时有差异。法国学校并没有投入大量的时间和精力，来查证每一次可能的违纪行为的具体情况，而是采取了另一种办法来保证与事实基本相符的结

果。我解释说，即便进行充分的调查，也有可能得出不正确的结论；虽然如果不进行调查，误判的可能性显然更高，但如果一个学生在一个学年内累计收到4次警告，那么他至少有一次是真的违反了规定，这种可能性还是极高的。令我震惊的是，我儿子在了解到还可以从另外一个角度考虑问题时，他的怒气很快就消了。

人们对税收的认知也能对自身的情绪反应产生同样重大的影响。如果我们认为纳税就等同于陌生人没收本该属于我们的财产，那么我们就几乎不可能不愤怒。但其实通过不同角度来看待税收才是更合理的。比方说，我在前面的章节也曾讨论过，如果没有税收支撑大量的公共投资，现代工业国家也不可能实现很高的平均收入。意识到这点，就能瓦解掉让我们贸然愤怒的认知。

然而，即便认同税收合理性的人，也通常保留着一种固执的观点，那就是他们个人所承受的赋税太重。在税前收入最高、与他人纳税差距极大的人身上，这种观点自然最为集中。尽管平民主义者抱怨富人能找到无数的纳税漏洞，但在2007年，美国收入前10%的人贡献了71%以上的联邦税收，收入前1%的人对税收的贡献则超过了40%。³

成功与美德

在1988年民主党全国大会上，得克萨斯州农业委员吉姆·海托华（Jim Hightower）形容老布什是“出生于三垒，就认为自己能打出三垒安打”。⁴鉴于布什先生的家庭背景在他事业生涯中所起到的显著作用，他应该意识得到幸运女神一直垂青于他。通常来说，自力更生的人其实更符合海托华的描述。自力更生的人明白自己有多努力，他们清楚地记得自己做出的每一次牺牲。相比之下，那些顶尖富豪的继承人更有可能不会意识到好运气对其成功所起的关键性作用。

假如说人们倾向于忽略好运气对于自己成功所起的作用，他们也往

往会忽略霉运对于别人失败的重要性。有时候，有些人申请破产是因为不努力工作，花销没有节制。但在美国，由于重大疾病而失业并失去医疗保险，最终导致破产的情况更为常见。这些影响因素却从来没有在关于贫困的社会评论中被提及。例如，一位评论员曾在由田纳西大学法学教授格伦·雷诺兹（Glenn Reynolds）主办的保守派博客“博学网”（Instapundit）发表以下感想——雷诺兹后来将这一评论选为“今日最佳点评”发表在了博客上。

造成“财富或收入鸿沟”的原因是：聪明的人一直做聪明且赚钱的事，而愚笨的人一直做愚笨的、让自己无法达成目标的事。

受过良好教育、不沾毒品、专心工作、拥有储蓄习惯、明智投资的人，比那些退学、滥用药物、透支买豪华汽车和房子而最终失去它们的人，表现好得多。

我们之间并不存在收入鸿沟，存在的是智商鸿沟。⁵

努力工作、一心一意的人当然更有可能获得成功。然而，经济学家正逐渐意识到，成功对于机遇的依赖性比大多数人想象的要强得多。经济失败在很大程度上也是由于机遇的变幻。

本书的核心主题之一就是：作为个人，我们经常会面临各式各样的诱惑，这些诱惑致使我们损害公共利益；为了抵制这些诱惑，相对于直接管制而言，征税通常更为有效且干扰性更低。然而，在美国提倡新的税种，往往会被认为是政治上不可能实现的事情。这大多是因为，主流的认知将征税描述成对个人自由的侵害。国家迫切需要彻底改革税收政策，然而，只有支持改革的新认知框架能够流行开来，这才有可能实现。

即便进行乐观估计，重新构建框架的速度也不会很快。但是，美国现有的税收结构会造成极大浪费，所以这种做法仍值得去推进。而要想

迈出有效的第一步，就要更加密切地审视成功与运气之间的关系。

与多数父母教导孩子的观点正好相反，天赋与努力工作既不是经济成功的必要条件，也不是充分条件。绝大部分成功人士当然既有极高天赋，又努力过人。然而，也有一部分获得了巨大成就的人两样都不沾。真人秀主角、假唱男子组合的成员、用客户的养老储蓄购买次贷证券并在市场崩溃前成功逃离的基金经理不就是这样的人吗？

其实，更常见的是，天赋异禀并且极其努力的人只能获得中等收入。他们中有数以千计的人都很走运，不过最大的赢家是那些既有天赋、野心、努力，又拥有极佳运气的人。

比尔·盖茨是个幸运的人吗

微软的联合创始人比尔·盖茨曾是多年的世界首富，至今仍居富人榜前三名。他于1955年出生在西雅图的一个富裕家庭，后来进入私立学校学习，13岁的时候就开始接触电脑编程。他在西雅图上的高中有一个电脑俱乐部，这在当时是极为罕见的，并且他能进入华盛顿大学的电脑实验室，他非常喜欢那里。

这些背景，连同他的天分、野心和勤勉，使他与自己的高中朋友、哈佛同学保罗·艾伦一同创建了一家软件创业公司。远在我们听说微软的大名之前，它就已经是一家生意兴隆的企业，但它的创始人团队在1976年为公司取名的时候，肯定没有想到，迎接它的将是如此辉煌的前程。

为开发一款用于IBM即将推出的个人电脑的操作系统，IBM向微软进行咨询，这是故事的开端。伦纳德·蒙洛迪诺（Leonard Mlodinow）在1990年出版的书中回忆道，最开始时，盖茨回复IBM说微软无法胜任这个项目。⁶他建议IBM去联系加里·基尔代尔（Gary Kildall），他的数码研究公司当时已经开发出了一款名为CP/M的个人电脑操作系统。

但基尔代尔的妻子，也就是数码研究公司的管理人，拒绝签订IBM提出的保密协议，于是谈判以失败告终。IBM的杰克·萨姆斯（Jack Sams）又回去找盖茨，讨论是否还有其他可行的办法。萨姆斯和盖茨都清楚，还有另外一款操作系统或许可以派上用场，那就是QDOS，6个星期前由来自西雅图电脑产品公司的蒂姆·帕特森（Tim Paterson）开发出来的一款“快速而卑鄙的操作系统”（quick and dirty operating system），这款系统借用了基尔代尔CP/M系统的操作手册作为基础。帕特森对QDOS做了足够多的明显改动，以便能够合理地拥有法律上的所有权。

根据蒙洛迪诺的叙述，后来盖茨向萨姆斯提出了一个事后被证明价值数十亿美元的问题：“你希望得到.....（QDOS），还是希望我拥有它？”萨姆斯似乎没有领悟到这个问题所蕴含的经济意义，据说他是这样回答的：“当然，由你拥有。”

此后，盖茨提议微软以大约5万美元收购QDOS，然后做了进一步改动，并将之重命名为MS-DOS，作为微软磁盘操作系统。但最大的幸运之处在于，IBM对个人电脑的市场前景持明显的悲观态度，从而允许微软保留操作系统的所有权，只对每一套系统的发行收取少许的许可费。正是这一步，让微软开始平步青云。

如果盖茨的高中没有电脑俱乐部，如果基尔代尔的妻子愿意签署IBM的保密协议，如果帕特森更加慎重地对待与微软的协议，如果以上事情中的任何一件没有发生，那么比尔·盖茨几乎不可能取得如此巨大的成功。

值得称道的是，盖茨一家似乎充分意识到了自己的幸运。盖茨后来被问到有多少青年人在上大学之前能够拥有与他相同的背景，他说道：“假如说能有50人，我就会觉得很惊讶。我在很小的年纪就接触到软件开发，当时其他任何人都不可能拥有这样的条件，所有的一切都归功于一系列不可思议的幸运事件。”⁷近年来，盖茨及其夫人向“比尔和

梅琳达·盖茨基金会”的捐款超过300亿美元，该基金会为减轻全世界人民的疾苦做出了大量贡献。盖茨的父亲，老比尔·盖茨，长期以来都在猛烈抨击有关取消房产税的提议。

与运动有关的幸运

促成比尔·盖茨成功的偶然事件在事后看来都极为重要。但其他许多成功故事中的偶然事件，哪怕我们留心关注，其重要性却不显眼。让我们来看看有哪些事情能够最终决定职业运动员的成功。马尔科姆·格拉德威尔（Malcolm Gladwell）指出，在最顶尖的冰球职业联盟中，大约40%的球员出生于1月、2月或3月，而只有10%的球员出生于10月、11月或12月。⁸他认为，之所以存在这种巨大的差异，是因为几乎各地的青年冰球联盟都将1月1日定为入队选拔的生日截止日期。

在年初出生的球员因此会一直成为队里年纪最大的成员。平均来看，他们的个头、强壮程度、速度和经验都比出生在后面月份的队员略胜一筹。由于他们在每个阶段都具有优势，因而更容易被选拔进入精英巡回队和全明星球队，也更有可能参与配备一流设施和一流教练的项目，或者获得运动奖学金，等等。

具备勤勉天赋和能力本身已是一种极佳的运气

格拉德威尔并不是说，只要在1月出生，就能保证毫不起眼的冰球运动员在日后有很大概率进入美国冰球联盟。他的观点在于，生日时间这样简单的偶然事件，就能对成功的概率产生影响。

尽管，勤勉工作和高超技能不能保证一定取得成功，但成功人士还是孜孜不倦地将自身的成功归功于这两个因素。这是因为，大多数成功人士身上都具备这两个因素，若非如此，他们成功的可能性将大大降

低，这样的表述也是发自内心的。但是，为什么某些人天生就具备比别人更为出色的技能呢？为什么某些人的勤勉程度远远超乎他人呢？通过近距离的观察，我们可以清楚地看到，能够具备这两个因素，本身就含有很大的运气成分。

环境和基因对于个人的影响究竟各占几分，这一直是个争论不休的问题。无论结果如何，事实上这些因素共同在发挥影响。有些人天生聪敏过人，那是因为他的基因里就携带了“聪明”的因子，或者因为他是在一个鼓励、刺激他发展智力的环境中成长起来的，又或者基因和环境的影响同时存在。我几乎可以肯定地说，是最后一种情况。

有些人能够，也更愿意勤勉工作的原因也是如此。这种性格可能部分源自基因，部分源自成长的环境。无论基因和环境各自的影响孰轻孰重，可以肯定的是，具备这两项品质的人在生活中拥有明显的优势。

如何才能出生在拥有出色技能和勤勉的良好基因，并且拥有一个能够促使他们获得这两项品质的成长环境呢？答案很简单，那就是靠运气。即便拥有这些品质不能保证一定成功，但他们成功的可能性远远超过那些缺乏天分和优越环境的人。

成功与运气的相关程度比大多数人想象的要高，但这并不是在否认努力工作能够增加成功的可能性，也不是否认努力工作本身需要付出巨大的努力。有这么一个故事，有人在一个鸡尾酒派对上遇到了一位著名的小提琴家，他说：“如果能够像您演奏得那样动听，我愿意付出一切代价。”演奏家问道：“那您愿意每天练习8个小时吗？”

心理学家K.安德斯·埃里克森（K. Anders Ericsson）在职业技能方面颇有研究，他曾估计，要想真正掌握一门技能，专注练习的时间至少要达到一万小时。⁹这未必是一种辛苦和劳累的负担。甲壳虫乐队在英国一夜成名，而在此之前，他们在德国汉堡市的廉价酒吧里磨炼了不止一万小时。许多传记在描述他们的经历时，都指出他们非常享受在汉堡的那段时光。比尔·盖茨在青年时代，就在编程上投入了不止一万小时，

那大多都是专心致志的愉快时光。

然而，更常见的是，在为技能奠定基础的一万小时里，人们都太过于愉快而无所事事了。正如埃里克森及其合著者所言，要想能够真正达到“有效练习时间”，其实要求很高：

你需要一项特定的练习，也就是刻意练习，才能锻炼技能。大多数人在练习时，只专注于他们已经掌握的技能。刻意练习是不同的。它是针对某些你不能做好，甚至完全不了解的事情，去做精心思虑的、具体的、持续的努力。诸多领域的研究表明，只有在你尚未掌握的事情上下功夫，你才能成为你所希望成为的专家。¹⁰

简而言之，成为一名真正的专家并不容易，这需要付出巨大的耐心和决心。要想树立这种决心，把努力工作的意愿视作由偶然事件决定的是没用的。抱有那种想法的人，通常更容易袖手旁观，看命运是如何为他做好安排的。那些想着“好人都努力工作，而我是个好人”的人也许更容易树立专注于艰巨任务的决心。

夸大道德品质在成功中所起的作用，这或许是一种有益的心理暗示。但是这种心理暗示并不能改变一个事实，那就是具备勤勉天赋和能力本身已是一种极佳的运气。

赢家通吃市场是如何将小的差异放大的

天赋和努力程度的差异能够解释为何人与人的收入之间存在巨大差异。我们在看待收入鸿沟时，还须考虑一个重要的维度，那就是劳动力市场如何将这种差异转换为税前收入的差异。

教科书里的竞争性劳动力市场模型认为，劳动者的收入取决于他们为雇主所增加的利润。有一种直觉式的观点很常见，那就是薪资随着能

力的不同而变化，这与竞争性劳动力市场模型的说法十分相符。比方说，在一家砖厂，每小时生产101块砖的员工的工资应该比每小时生产100块砖的员工高1%。但正如我和经济学家菲利普·库克（Philip Cook）在1995年出版的书中所写的那样，在许多劳动力市场中，薪资和能力的关系并不符合这个简单模型。¹¹

其中一个最显而易见的例子就是公司高管的市场。在20世纪80年代，美国最大规模公司的首席执行官们的收入是普通工人的42倍，到2001年，差距高达500多倍¹²，然而并没有证据表明首席执行官们的聪明程度和努力工作程度比从前高。造成这类群体收入爆发性增长的原因，与其他大多数高端劳动力市场群体的情况类似，因此对此进行仔细研究很有启发意义。亚当·斯密和达尔文曾提出，为何自由的竞争通常会导致许多令人无法接受的结果，深入观察这些劳动力市场有助于我们思考这个问题。

许多左派评论家认为，高管薪资令人惊愕的增长水平能够证明竞争性市场力量已然崩溃。他们指出，行业巨头们合谋起来驱逐竞争对手，所以能够从俘获的客户手中敲诈出前所未有的高薪资。并且，高管们把自己的同伙安排在董事会里，同伙则回报以高得离谱的薪资和奖金。

可以肯定的是，这样滥发薪资的行为确实存在。但是现在的情况比以往有所好转。沟通便利性的提升和交通成本的降低，使这种情况不再那么严重。高管在雇用管理层成员时确实存在信息不对称，但他们所获得的信息比过去要多，并且声誉成了更加有效的绩效预报器。同样，机构持股人的警觉性越来越强，敌意收购的威胁越来越大，这都为高管薪资滥发现象增加了额外的限制。

尽管有了上述进步，企业的管理依然存在问题。某些绩效平庸的高管还可以拿到很高的薪水，但绩效不合格的高管被解雇的速度要比以往快得多了。我和菲利普·库克写道，在所有劳动力市场中，顶层薪资飞速增长的原因有二：一是技术力量能将绩效上的一点小进步急剧放大，

二是公司对绩效最优的高管的争夺也日趋激烈。¹³

绩效决定薪资是我们所谓的“赢家通吃”市场的第一个前提。而第二个前提则是，少数绩效顶尖的人往往能够获得最多的报酬，天赋和努力程度上相差不大的人之间却依然存在巨大的收入鸿沟。这两个方面都在舍温·罗森（Sherwin Rosen）关于古典音乐市场的描述中得到了体现：

如今古典音乐市场空前地大，然而，任何一种乐器的全职独奏家的数量仅为数百人，声乐、小提琴和钢琴之外的独奏家数量更少。在这一小群人中，一流演奏家的数量屈指可数，他们的收入却非常高。第二梯队演奏家的收入与他们的收入存在着天壤之别，哪怕大多数消费者通过“盲听”，很难鉴别出其中的细微差别。¹⁴

逼真录制和回放技术的发展，使天赋最佳的音乐家们拥有了巨大的影响力。由于我们所欣赏的大部分音乐都是预先录制的，我们能随时随处听到世界第一的女高音。从世界第一女高音的录音母带复制激光唱片与从其他歌手的录音母带复制激光唱片相比，成本毫无差别，数百万的乐迷自然愿意为每张多花几分钱购买顶级音乐家的唱片，而不是去购买与之稍有差距的音乐家的唱片。结果就是，一流的女高音能签下七位数的唱片合约，而其他稍稍逊色的音乐家只能勉强过活。

这个逻辑在大型机构领导者身上也同样适用。2007年，聘用戴维·斯克尔顿（David J. Skorton）为康纳尔大学第12届校长的董事会清楚地知道，校长最主要的职责就是完成学校40亿美元的募捐任务，当时募捐才刚刚启动。聘任委员会挑选了几位他们认为可以实现这个目标的人选，他们最终判断，没有人能像斯克尔顿一样出色地完成这个任务。

在了解斯克尔顿过去几年在该职位的表现后，我很轻易地就明白了选择他的理由。斯克尔顿极富幽默感、热情和魅力，在心脏病研究和爵士乐方面具有突出成就，校友们都很爱他。如果他在演讲中所描述的关

于学校未来的愿景能够打动校友，比排位第二的候选人多获得3%的捐款，募款金额就能增加1亿多美元。

我不知道斯克尔顿的薪水是多少。几年前，就有报道称，有几所私立大学的校长年收入已经超过100万美元，许多社会评论家对此表示震惊和愤慨。像斯克尔顿这样的领导者凤毛麟角，正因为他们是如此珍贵，真正令人震惊的反而是他们只能获得这个水平的收入。

在许多私人公司，薪酬差异更是大得惊人。假设一个年收入100亿美元的公司将首席执行官的人选最终确定为两人。假设其中一人每年有数个决定优于另外一人，让公司的年收入增加3%可谓是轻而易举，也就是能够增加3亿美元。如果首席执行官这个职位的竞争人选在决策的质量上有优劣之分，那么经由竞争脱颖而出的最优人选的薪水将远远高于次优人选，尽管他们的天赋差异微乎其微。

近几十年来，管理层决策的影响力急剧增强，公司越大越是如此。造成这种现象的其中一个最重要的原因就是信息革命，再加上沟通和交通成本的降低、制造业技术的发展以及其他因素，许多本地和区域市场变成了美国和国际市场。比方说，在俄亥俄州北部表现最出色的轮胎生产厂商，过去至少能在区域市场占有一席之地，而日渐精明的消费者现在只会从世界排名靠前的几家生产厂商中进行选择。公司的表现一贯是由一小部分最优秀的员工决定的，但由于如今市场的边界发生了变化，存活下来的公司的领导者所发挥的作用要比之前的同职位的人大得多。

在竞争性市场中，更大的影响力意味着更高的薪资。2008年，经济学家泽维尔·加克贝斯（Xavier Gabaix）和奥古斯丁·朗迪耶（Augustin Landier）曾在报纸上撰文写道，高管的工资应该根据公司的市值来定。¹⁵他们发现，在他们研究的大型公司的案例中，首席执行官的报酬在1980年到2003年间增长了6倍，大致与公司市值的涨幅相当。

放松管制既提供了新的市场机会，又带来了新的竞争威胁，进一步提升了美国航空业、货车运输业、银行业、经纪业及其他行业的高管的

价值。除此之外，衍生证券和金融资本的新注入渠道使外部收购的威胁日益增强。这些变化使优秀绩效能够带来更大的潜在收益，而劣等绩效所带来的损失也会恶化，因此，在关键职位雇用最优秀的人才也就变得尤为重要。基于上述原因，顶级高管的经济价值正在不断增长。

然而，决策影响力的增强并不是美国企业高管薪酬增长的唯一原因。毕竟，美国大规模公司首席执行官们的决策影响力向来很大，但百万年薪不过是最近20年才出现的。再者，全球化不仅使美国企业高管的决策影响力增强了，在德国和日本也是如此，近年来这两个国家企业高管的薪资也在上涨，但相对于美国来说涨得还算“节制”。所以，并不是说最优秀的首席执行官能为公司带来数百万的利润，他就一定能获得百万级别的薪资。

在赢家通吃的市场，最优秀的企业高管能获得高额的报酬，他们除了必须创造出非凡的价值外，还得参与有效的竞争。然而，在许多市场领域，种种正式和非正式的规则限制了此类竞争。

比方说，大多数职业运动联盟会设定限制条约，阻止球队老板从其他球队以高价挖走优秀球员。1976年，棒球运动的保留条款被成功废除，随后其他类似的条款也相继废除。如今，在大部分职业运动队中，运动员至少赢得了自由选择经纪公司的有限权利。获得这些权利之后，运动员的薪酬都会大幅增加。¹⁶

与职业运动队老板不同的是，并没有正式的条约规定公司老板不可以竞价挖走优秀人才。但非正式规则似乎发挥着同样的作用。在非正式规则中，从内部提拔人选作为高管曾经一度是通行的惯例。大多数情况下，内部候选人通常只有寥寥几位。

高管的薪资是由公司和内部候选人进行双边谈判决定的。候选人没有别处可去，公司也没有别的人选可以考虑。在这种情况下，公司能够以不到如今薪资水平的1/10留住优秀人才。

企业界的反掠夺规则早已不是什么秘密。其中最著名的一个例子就是IBM聘请路易斯·郭士纳，郭士纳擅长为企业重整旗鼓，他为雷诺兹-纳贝斯克（RJR Nabisco）公司创造了破纪录的收益，但在计算机行业却毫无经验。在早些年，跨行业聘请人才几乎是一件不可思议的事情。然而，IBM的赌注赢得了极大的回报。20世纪90年代，郭士纳带领这位举步维艰的“计算机巨人”实现了惊天逆袭。¹⁷

全新的高管人才“现货”市场极大地影响了高管的薪水，近几十年来，自由代理对职业运动员的影响也是如此。在我和菲利普·库克的一项研究中，通过对美国最大的制造和服务企业中近800名首席执行官进行调查，我们发现，在1970年到1992年期间，从外部聘请高管的比例提高了50%。¹⁸这个趋势有效地打破了原先不成文的将高管绑定在原公司的保留条款。

尽管超过半数的首席执行官仍从内部任命¹⁹，但游戏规则已然完全改变。在美国，表现最优秀的高管越来越倾向于寻求外部机会。为留住最有价值的资深高管，企业必须开出足够高的薪水才能防止他们跳槽。棒球界取消了保留条款，这成为近年来最优秀球员薪资爆发的先决条件。而流动便利性越来越高也为最优秀高管提供了有利条件。

总而言之，飞涨的高管薪资是竞争瓦解的标志这一说法经不起推敲。但有一个显而易见的例外，那就是在金融服务业，金融机构提供大量竞选资金，收买对自己有利的监管条例，这严重阻止了企业进行有效竞争。²⁰但事实上，近年来大部分飞速增长的薪资得益于市场力量的增强。

当然，赢家通吃并不是什么新鲜事。技术使地球上最优秀人才的能力和活动范围以空前的速度扩展。500多年前，印刷机帮助少数几位小说家取代了数百万的乡村说书人。后来，电报机的发明使一小部分联合专栏作家取代了大量的当地记者。个人计算机的普及则让一小部分软件开发者替代了数以千计的地方税务会计师。伴随着这些变化的兴起，竞

竞争性劳动力市场将个体在天赋和努力程度上的差异，放大为前所未有的收入差距。天赋和努力程度的差异在某种程度上源自我们无法掌控的偶然因素，最终，如今运气与成功之间的关联，比人类历史上的任何一个时间点都要强。

开明的自由主义者或许不再认为一切税收皆为窃取，转而支持另外一个不再那么极端的观点：人们有权利像对待税前收入那样来对待税后收入，毫厘必争。假若拥有比他人高1%的天赋，或多付出1%的努力，就能多获得1%收益，在这样一个世界里，这种观点似乎讲得通。然而，在我们所生活的世界里，如此微小的差异，常常会导致数百倍或数千倍的收入差距，因此这种说法是站不住脚的。

信念至关重要

意识决定行动。受“这是你们自己的钱”这种认知框架的启发，小布什政府极大地降低了最高收入人群的所得税。为减少由此带来的联邦预算赤字，官员们竭尽所能地削减开销。被削减开支的项目未必是产生价值最少的项目，它们更有可能是最不引人注目，或者最不容易引起选民攻击的项目。

其中一个被小布什政府终止的项目，是美国能源部的一项关于隔离苏联泄露的核材料的工程。²¹这些核材料存储在质量很差的设备里，并且由经常喝得醉醺醺的、不能按时领到工资的士兵看管。恐怖分子急切地试图获取这些核材料。基于当前的情况，我们并不难想象他们有一天能够得手。

从任何一个收入水平群体的角度来说，削减该项目的开支都不是一个有利的决定。但在当前的决策框架下，诸如此类的决策非常常见。

要想隔离苏联那些保管不当的核材料，需要花钱。如果政府希望促成此事，就得付出代价，也就是说必须征税。我将在后续的章节论证，

基于种种原因（但并非为了社会公平），从高收入人群身上征收更多这类政府需要的费用，不仅符合富人的利益，对穷人也是有益的。但只有改变关于税收的思维模式，这才有可能发生。人们的税前收入存在巨大差异的原因也与这种思维模式有关。

主流的思维模式认为，对于用天赋和努力在劳动力市场所换取的全部报酬，人们天然地拥有全部使用的权利。由于收入在某种程度上受到税收所支持的公共投资的影响，这种观点根本站不住脚。越来越多的证据证明，天赋异禀、努力过人的劳动者在劳动力市场所取得的成功，很大程度上取决于偶然因素，这就进一步否认了上述观点。由于心理作用，人们普遍将成功归因于自身的美德，但事实上，是否拥有天赋以及努力工作的品质，很大程度上取决于无法控制的因素。

在2009—2010年的斯坦利杯冠军赛季中，马利安·霍萨（Marian Hossa）在芝加哥黑鹰队担任右前锋，当年他获得了800万美元的收入。霍萨出生于1979年1月12日，毫无疑问，他为自己冰球事业上的成功付出了极大的努力。我没有见过他，也不知道他如何看待自己当前的缴税负担。但我们可以轻易猜想，如果他知道自己的生日与冰球事业上的成功存在关联，或许他会对现在的缴税负担少一点反感吧。

THE DARWIN ECONOMY

THE DARWIN ECONOMY

10

伟大的权衡：
为何降低高收入者的税率会抑制经济增长



《伊索寓言》里有一个关于一只会产金蛋的鹅的故事：

有个男人和他的妻子运气很好，他们拥有一只每天都会产金蛋的鹅。尽管已经足够幸运了，但不久，他们又开始盘算怎么样才能一夜暴富。他们想，这只鹅肯定里里外外都是金子，要是杀掉它，就能马上得到一大笔金子了。可当他们剖开鹅的身体时，却发现这只鹅的体内与普通的鹅没有任何区别。

这则故事长期以来都是活跃自由主义者的最爱，他们以此来提醒那些倾向于进一步实行累进税政策的人，这样做会使所有人都变得贫困。自称自由主义者的美联储前主席格林斯潘曾发表过类似的观点：“任何税收都会拖累经济增长，只不过程度不同而已。”¹

然而，并不只有自由主义者才有这种想法。这个观点的另一个版本，就是涓滴理论^⑤，由于政界各方人士都经常提到，这个理论似乎已经成了板上钉钉的事实。

当然，“所有税收都会拖累经济增长”这种说法不是事实。我在前面已经讲过，如果不征税，政府就不能组织和维持文明社会，保护国家不受侵犯，也无法享受经济快速增长所带来的种种好处。

我在此不展开论述，但如果不进行征税，我们就不可能讨论税收是如何影响经济增长的。我将在第11章中重点论述如何对伤害他人的行为进行征税。你将会了解到，此类征税不仅不会拖累经济增长，反而能起到促进作用。

所得税与涓滴理论

那么，所得税呢？自由主义者和保守派人士都确信，在超过某种限度后，过高的所得税绝对会成为经济增长的累赘。比方说，如果税率高达100%，由政府分配这些收益，使大家的税后收入都一样，那么就没人愿意再去工作。结果，也就没有收益可以再去分配了。

这种说法几乎没有什么吸引力，因为它并没有说明假若真的提高最高收入人群的税率（现在距离100%还有十万八千里），将会出现什么情况。如果涓滴理论要想引起人们的注意，它必须提出诸如此类的主张：在当今工业国家的实际税率基础上，进一步加重最高收入人群的赋税将会阻碍经济增长。

这种说法正确吗？从表面来看似乎有些道理，这是因为人们向来认为，人的行为是由动机引导的。由于加重最高收入人群的赋税会减少努力和承担风险所带来的回报，我们似乎就有理由相信，这将导致人们缩短工作时间、降低他们所承担的风险。

然而，所有的经济学教材还提出，税收收入的减少还将导致另外的反作用。那就是，由于收入降低使人感觉匮乏，它反倒会刺激人们工作更长时间、承担更多风险，来消除这种匮乏感。比方说，一位高收入者所希望维持的生活水平需要每天花费2 000美元。如果他现在的税后收入是每小时250美元，那么他每天需要工作8小时。但如果税率提高，使他的税后收入变为每小时200美元，他每天不得不多工作两小时，否则就得变卖他的法拉利跑车。

其他人对于税率提高的反应可能有所不同。由于更高的边际税率降低了额外休假的机会成本（就放弃税后收入而言），有些人可能会选择减少工作时间。至于到底会引发哪些反作用，经济学理论没有给出任何答案，完全没有。

假若经济学理论不能证明涓滴理论的合理性，那么数据能不能予以证明呢？同样，数据并没有为涓滴理论提供更多的支持。有一种观点认为：如果税后收入的降低会导致人们缩短工作时间，那么反之亦然，也

就是更高的税后收入会使人们增加工作时间。自1900年以来，剔除通货膨胀因素，美国的平均时薪已经增加了不止5倍。那么，依照涓滴理论，现在美国人的工作时长应该明显增加了。然而，如今美国人每周的工作时间只相当于1900年的一半。

涓滴理论还预测，在实际税后工资更低的国家，每周的工作时长应该更短。然而，数据所证明的事实却完全相反。日本首席执行官们的收入还不及美国首席执行官们的1/10，并且所面临的边际税率也越来越高，但事实上，他们的工作时间比美国首席执行官更长。

涓滴理论十分强调人的动机，导致许多人据此预测，收入差距的扩大应与经济增长率成正比。因为收入差距的扩大会给人们更多的压力，去追赶收入超过他们的人。在第4章中，我们曾讨论过，收入差距确实会对消费模式产生影响。可是，研究人员在对单个国家一段时间内的数据进行分析后发现，经济增长率与收入差距成负相关。在第二次世界大战后的30年里，收入差距处于历史较低水平，然而大多数国家的经济增长水平却极高。相反，到1973年，经济增长率只有之前的一半左右，那时大部分国家的收入和财富差距都在稳步扩大。

从跨国数据来看，经济增长与收入差距也成负相关。通过分析世界银行以及经济合作与发展组织针对65个工业国家采集的数据，经济学家阿尔贝托·阿尔塞纳（Alberto Alesina）与达尼·罗德里克（Dani Rodrik）发现，经济增长率与收入前5%和20%人群在国家总体收入中所占的比例成反比。²现实状况一而再，再而三地与涓滴理论的预测正好相反。

给富人减税并不会带来更多工作机会

2001年，小布什提出了涓滴理论的另一个演化版本；2008年，美国国会议员约翰·麦凯恩（John McCain）再次提出了相似的观点，他们都针对最高收入群体的减税提议进行辩护。这个演化理论并没有说明减

免税收将会如何影响人们努力工作的程度。他们都提出，由于许多富人都拥有小型企业，这些企业每年提供的新工作岗位超过全美新工作岗位总数的一半，减免税收将会立刻刺激产生更多的工作机会。

乍看之下，这种观点似乎有些道理，事实上他们在提出这种观点的时候也几乎无人反对。然而，如果仔细想想，从经济学的角度这根本讲不通。这与人人皆知的关于商业雇佣的经济学逻辑不吻合。

其实这个观点暗含了一个先决条件，那就是如果企业主有经济能力雇用更多员工，他们一定会那么做。但你是否具有经济能力去做某事，与你是否应该去做并无关系。比方说，你能买得起你不喜欢的歌手的唱片，并不意味着你就应该买它。与此类似，企业主能够雇得起更多的员工并不能说明什么问题。问题的关键在于，雇用更多的员工是否能够增加额外的利润。

在每一本入门级经济学教科书中，雇用员工的基本原则都说得非常直白：如果新员工的产出至少能够覆盖自己的薪酬，那么这名员工就应该被雇用，否则就不应该。在满足这个原则的情况下，从经济学的角度来讲，无论企业主多么窘困，雇用新员工都有道理。相反，如果不满足这个准则，哪怕企业主是世界首富，雇用都没有道理可言。企业主个人的税后收入与是否雇用员工无关。

减税提议的拥护者们有时会提出，在新员工开始生产之前，企业主预先需要一定量的资金来覆盖雇用员工和培训的成本。减税可以帮助他们获得这部分资金。话说得没错，但这不符合基本的雇佣原则。

如果企业主利用减税的钱来支付雇佣最初的成本，其实就是自己借钱给自己，以期获得未来的回报。事实上，他们就扮演了为自己服务的银行家的角色。这类借款和外部借款的评测方法是一致的。

如果雇用新员工的最终收益不仅能够支付他们的薪资，还能还上借款和利息，那么银行借款就能通过成本-收益测试。内部借款也必须满

足同样的标准，即公司的最终收益不仅能够支付他们的薪资，还能还上借款以及企业主没有把这笔钱存进银行所失去的利息。在做雇佣决策时，内部借款的隐性成本跟外部借款的显性成本同样重要。

低税率能够刺激小型企业主雇用更多员工这个观点，与入门级经济学教材所指出的基本原理是完全对立的。小型企业被誉为美国新增工作机会的主要引擎，但这并不能否认一个事实，那就是企业主的个人所得税税率与他们雇用员工的决策无关。

如果张三额外雇用一名水管工，每周可以增加2 000美元收入，但只需支付1 500美元薪水，那么张三的水管服务公司就有充分的动机雇用他。诚然，减免税收能使张三口袋里的钱更多。但哪怕没有减免税收，张三也能雇用新的水管工。但反过来，如果新的水管工每月带来的收益低于1 500美元，无论减税力度多大，张三都不能雇用他。

总的来说，无论是经济学理论还是实证经验，都无法证明涓滴理论的演化版本的合理性。我们没有令人信服的理由来证明，向最高收入群体征收更高的税将阻碍经济发展。而正如之前所说，这种说法常常被提及，却很少被质疑。

为何降低高收入者的税率通常会抑制经济增长

更麻烦的是，我们有足够的理由相信，对最高收入群体实行低税率或许会使经济无法充分利用可用资源。比方说，对最高收入群体实行低税率不能鼓励他们更加努力地工作，却会导致他们选择生产效率更低的工作。

在2008年金融危机到来之前，美国最优秀、最聪明的毕业生中希望成为华尔街基金经理的比例一年比一年大。例如，在普林斯顿大学2007届的毕业生中，45%的就业人员选择了金融服务业。统计发现，从许多方面来看，这些工作其实魅力平平。³这些工作的工作时间很长，并且

大部分压力极大。还有人抱怨这些工作对智力的要求不高。但大家对这些工作有一个共识，那就是这些工作的待遇极好。最优秀、最聪明的人都在金融服务领域求职，数量前所未有的，这是因为这个行业的成功人士都可以赚得盆满钵满。

据机构投资者（Institutional Investor）出品的《阿尔法》杂志（*Alpha*）报道，对冲基金经理詹姆斯·西蒙斯（James Simons）在2006年赚了17亿美元，另外还有两名经理的收入也超过了10亿美元。当年，排名前25的对冲基金经理的收入合计超过140亿美元。一年后，对冲基金经理约翰·保尔森（John Paulson）获得了约为40亿美元的收入。

这些基金经理同样享受着极好的税收优惠待遇，没有人可以一本正经地解释其中的原因。比方说，即便是“附股权益”，也就是指他们在投资组合收益上所获得的20%佣金，看起来也像是普通所得，这部分收入是按照15%而非35%（普通所得税的最高标准）的税率来征税的。2007年，单是这一条款就为保尔森节省了约800万美元的税。

美国国会周期性地提出，应以普通所得的税收标准来对附股权益征税。不出所料，金融行业的游说者总是坚持声称，这样的做法无异于杀掉产金蛋的鹅。诚然，他们的客户作为中介所促成的交易具有重大价值。然而，这项立法提议并不会妨碍任何一桩有价值的交易。如果按照普通所得的标准对附股权益进行征税，在促成同样交易的情况下，对冲基金经理的100万美元的税后收入将变成76.5万美元。真有人认为这会让他放弃这桩交易吗？

经济分析表明，向对冲基金经理征收更高的税，实际上可以缓解基金经理竞争人数过多的浪费现象，从而刺激其他领域生产的发展。这个市场是赢家通吃的典型代表：在竞争中，少数几位赢家是从大量的初始竞争者中选拔出来的。这样的市场通常会因为两个原因而吸引过多的竞争者。

第一个原因在于信息偏差。要想做出理智的决策，在决定是否加入

竞争之前，你至少要弄清自己获胜的概率大约有多大。但人们对于自身相关技能的评估往往过于乐观。比方说，超过90%的员工认为自己的生产力高于其他同事的平均生产力。大约有同样比例的司机认为，自己的水平高于其他司机的平均水平。这种偏差似乎不会随教育水平的提高而消除。九成以上的大学教授坚信自己的工作比其他同事的更富有成果。⁴

过度自信通常是人们扭曲职业选择的原因之一，因为它会增强人们的动机。除此之外还有第二个原因，那就是最大的赢家总是辉煌耀眼。NBA明星每周在电视上出现几次，就能挣到8位数的薪水，数以千计的失败者却无人关注。同样，获得10位数收入的对冲基金经理比那些未能进入最后一关的竞争者要耀眼得多。有一个简单的逻辑，那就是如果人们过高地估计了自己获胜的可能性，许多人会放弃在传统市场薪水很高的职位，而选择进入赢家通吃的市场进行竞争。

公地悲剧

赢家通吃市场的潜在竞争者们也会遇到一个被称为“公地悲剧”的问题，这是一个最初为了解释海洋中的过度捕鱼问题而发明的激励机制。⁵北大西洋的鳕鱼曾经产量丰富，但由于过度捕捞，数量减少了95%。导致鳕鱼数量减少的原因与导致过多竞争者进入赢家通吃市场的原因十分类似。

公地悲剧为达尔文提出的“个人利益与集体利益之间通常存在着严重冲突”这一观点提供了生动的例证。我们可以通过一个简单的数值案例来解释这种冲突的本质。假如有100个人要在两个职位中做出选择，一是在一家与世隔绝的渔场捕捞鳕鱼，二是去一家工厂工作，两个职位的年薪都是5万美元。为方便起见，假设除薪酬外，每个人都认为这两个工作具有同等的吸引力。如果在渔场每年至少能挣到5万美元，有人就愿意选择渔场；如若不然，这个人就会选择工厂。

假设最开始的时候，渔场里的鳕鱼很多，第一位入职的捕鱼者每年能挣到10万美元。由于这是工厂年薪的两倍，其他人很快就会被吸引到捕鱼业。但由于渔场里的鳕鱼供应数量有限，平均捕鱼量会随着捕鱼者数量的增加而稳步下降。据经济学理论预测，当捕鱼业的年收入降至5万美元，即与工厂所支付的年薪相同时，人们就会停止进入捕鱼业。我们假设捕鱼业人数增加到40人时就会发生这种情况。

如果每个人的年收入都为5万美元，那么40人一年的总收入则为200万美元，剩下的60人选择了工厂，他们一年的总收入是300万美元。总的来说，100位工人一年能赚500万美元。但如果这100人都选择了工厂，他们会赚一样多的钱！渔场本来是极富潜力的珍贵资源，最终变得完全没有经济价值。问题在于，个人的动机导致太多人涌入捕鱼业，从而使捕鱼业潜在的经济剩余所剩无几。毫无疑问，结果原本可以比这更好：如果只有一人选择捕鱼业，100人的总收入则会增加5万美元，这是因为仅有的一位捕鱼者可以赚10万美元，而其他99位在工厂工作的人每人可以挣到5万美元。

公地悲剧通常是由简单的外部因素引发的。每位潜在的捕鱼者只关注自己预期的捕鱼量所能带来的收益。他们没有理由去考虑，自己进入捕鱼业将会导致原有捕鱼者的收获数量变少。市场趋于平衡时，最后一位进入者捕鱼所能获得的5万美元刚好能够补偿他放弃工厂工作而损失的5万美元。但由于他的进入减少了其他已从业捕鱼者的捕鱼量，他的进入带来的系统影响实际上是负面的。进入捕鱼业的个人动机过大，过度捕鱼将成为意料之中的结果。

跃跃欲试的投资组合经理面临的动机与之类似。正如海里的鱼儿数量是有限的一样，在特定的时间点要达成的交易数量也是有限的。在超过一个平衡点后，基金经理人数的增加与投资管理带来的总佣金金额的增加不再成比例。如果有人在对冲基金行业幸运地获得一个职位，他所能获得的收益，将很大程度上被竞争对手未能获得这个职位而导致的损

失抵消。因此，在这个案例中，个人的动机将会导致人满为患的浪费现象。⁶

然而，不只有金融领域会出现吸引过多竞争者的情况。这是一种几乎所有赢家通吃市场都具备的特征。并且，其实几乎所有能够提供社会最高收入的市场都是赢家通吃市场。通常每个行业都有几个重量级的职位，比方说获得格莱美奖的歌手、全明星棒球游击手、畅销书作家、500强企业的首席执行官、主流网站的新闻主播、获得奥斯卡奖的演员、流行广播脱口秀主持人、一线原告律师等。由于这些职位的收入非常高，竞争不可避免地也就非常激烈。通常会有数以百计，甚至数以千计聪明能干而又野心勃勃的候选人加入竞争之列。

我们再回到之前的话题，在赢家通吃市场，潜在的竞争者往往会高估自己获胜的概率，同时还会忽略一个事实，那就是他们加入竞争后，原有竞争者的胜算又会降低几分。由于这两个原因，此类劳动力市场会吸引过多的竞争者。

在公地悲剧的捕鱼业案例中，如果就业人员从捕鱼业向工厂转移，总体的收入将会上升。同样，如果天赋最低的竞争者能够放弃华尔街或其他赢家通吃市场的竞争，而选择更加传统的职业路径，仍有足够多有天赋的竞争者参与超级明星职位的竞争。因此，如果有一半参与对冲基金和私募股权基金职位竞争的人明天就离开金融行业，这些岗位仍然不缺极其优秀的候选人。这样就会产生更多更优秀的工程师、医学研究员、教师和家庭医生，由此带来的收益将超过赢家通吃市场竞争者变少所带来的损失。如果赢家通吃市场的税后收入降低，为这些职位而竞争的人数就会变少。因此，可以通过提高最高收入群体的税率来实现我们想要的就业转移。

在谈及取消对冲基金和私募基金经理的税收优惠时，一位来自哥伦比亚商学院的金融学教授提出了反对：“私募股权是经济中非常重要的一个部分。”他还补充说提高税率会阻碍这个领域的发展。⁷其他人则认

为此类提议是嫉妒阶级的福利。

然而，以上结论都遗漏了一个最重要的要点。没有人会否认，能干的人可以让资本发挥出最大的作用，这对于社会来说是非常重要的。但能够带来收益的交易的数量不会无限扩张。在达到一定程度后，向这些岗位输送更多优秀毕业生的机会成本将会很高，但增加的经济效益却微乎其微。不出意料，几乎所有涌入金融服务业的哈佛、普林斯顿和耶鲁大学的毕业生都极其聪明、勤奋。如果他们在其他行业有所追求，有的人可能已经发现了治疗绝症的有效疗法，也有人可能已经开发出更加高效的太阳能板。但是，他们却为复杂的金融衍生品卖力，将国家推入了自大萧条以来最严重的经济衰退期。

总而言之，让赢家通吃市场的税后收入稍微看起来不再那么惊人，可以提高其他职业路径的吸引力，而在这些领域，更多的人才能够产出实际的成果。由此产生的税收能够为许多确实需要付诸行动的事情买单。在一个经济体中，如果赢家通吃市场的重要性越来越强，关于公平与效率权衡的传统观点将会被颠覆。在当前环境下，向最高收入群体征收更高的税，其实可以立即达成上述一系列目标。

即便如此，我仍认为向最高收入群体征收更高的税是个糟糕的主意。我曾在第5章解释过，对家庭的总消费开支施行更强有力的累进税，比当前的税收系统有效得多。我反对所得税的理由与自由主义者和其他反税战士的理由有很大差别，他们的口号使我们无法就“如何改革功能严重不良的征税系统”进行理智的讨论。

讨论须集中在几个基本的、超越具体税收细节的问题上。比如说，如何设计产权的归属？私人物品与公共物品之间的最佳平衡点在哪儿？什么样的机制能够最好地促进平衡？我们应该如何为公共设施买单？我们对社会最贫困成员应该承担什么责任（如果有的话）？什么样的体系能够最好地促进环境的可持续性？等等。

愚蠢的口号竞争

进行一场关于税收政策的开明对话能使政界左右两派都受益。它能为各种收入水平的群众带来好处。然而，反税和反政府的口号却到处阻碍对话的发生。几乎不出意外，这些口号明显违背了人类行为的理论和实证经验。然而，人们却继续喊着这些口号而不觉得尴尬。事实上，其中的许多口号甚至被经验丰富、表面中立的评论家充满敬意地重复着。只有当这些口号能够激起广泛的、应得的蔑视时，真正的改革才有可能发生。因此，在奋力向前之前，现在就是一个审视这些口号不当之处的恰当时机。

如果要选出一个最愚蠢的反税口号，我会投票给“一切税收皆为窃取”。这个口号表面上是说，爱管闲事的政府官员不应该拥有没收经济资源的权力，那是人们凭借自己的天赋和努力创造出来的。但在一个缺乏完善的、强制的产权系统，以及高度发达的专业化市场所需的公共基础设施的国家，是没有什么经济价值可以没收的。除非政府有权强制征税，否则这一切都将不存在。没有一个有见识的公民会真正选择在政府缺乏这种权力的社会中生活，他们的担心不仅限于此，这样的社会很快会被拥有强制征税制度的邻国武力征服。

位居第二的愚蠢口号则是这样的：“向部分人征收更高的税是不公平的。”这个口号有几个失败之处。所有的社会都会因为拥有道路、桥梁、警察、消防、国防以及其他公共设施和服务而受益。但没有一个社会可以在不增加税收的情况下，有能力承担这些设施和服务的支出。同私人物品的情况一样，更高的收入通常会催生出对于更多、更好的公共设施的需求。但如果社会不对部分人征收相对更高的税，社会所能提供的公共设施在数量和质量上只能满足最贫困人群的需求。我再次重申，一个有见识的人不会愿意生活在这样的社会。

一个认为累进税不公平的口号则基于一种奇怪的公平理论。在自由

主义者的理想世界里，人们可以按照自己的意愿，自由地与他人组建社会，并且只受全体一致认可的规则约束。在这样一个世界里，如果有人希望组建一个自己位于高层级的社会，他们就得说服他人接受低层级的位置。并且，如果高层级位置的隐形市场同竞争性劳动力市场一样，自愿组成的社会中的高层级位置将具有很高的市场价值。由于此类位置的价值很高，它的存在还基于一个条件，那就是为了高层级位置的存在，必须设立低层级位置，其他人必须承担相关的成本。因此，把累进税说成一种公平的税收反倒符合逻辑。然而，反税战士们却坚持认为，人们有权利自由地获得高层级位置而无须付出代价，他们给出的理由很简单：交易成本的存在，使得组建一个自愿性社会在现实中行不通。如果按照惯常的理解，这种观点实际上是在藐视公平。

还有一个很有影响力的反税口号是：“市场上的成功依赖于天赋和努力，而不是运气。”这句口号会鼓励人们为了成功而更努力奋斗，从这个角度来说，它确实是有益的。但它也有弊病。数百万勤奋而又有天分的人从未获得过成功。对现代劳动力市场的功能了解越深，我们就能越清楚地看到，其实是偶然事件在起决定性作用。甚至个人的天赋和努力工作的能力都在很大程度上取决于基因和环境因素，因此理智的人都不愿意将成功归结于美德。这些口号故意忽视运气所起的作用，这会增强人们的一种信念，那就是他们拥有一种道德权利，可以百分之百地拥有自己的税前收入。这种信念毫无疑问是有害的。

然而，在所有反税口号中，没有任何一个口号能比本章议题所带来的伤害更大：“向富人征税无异于杀死产金蛋的鹅。”尽管没有经济证据或实证的支持，我们仍然可以说这句话明显不如“一切征税皆为窃取”荒谬。但由于种种原因，相信前一句话的人更广泛，因而给公众带来的恶性影响也就更大。比方说，它是小布什针对最富有家庭实行减税政策的基础，致使布什在任期间，国家的债务翻了一倍。在第4章中我们也看到，此类减税会使消费模式朝着不利的方向发展。

在我人生中，每一任当选的美国总统在竞选时都承诺减少政府的浪费性开支。其中有几位总统朝着这个目标取得了真正的进展。然而，在他们中的每一位在任期间，政府的总体开支持续增加。当然，我们应该坚持更大程度地减少浪费。但我们所面临的未来的大量开支是不可任意改变的，因此我们的期望不太现实。

比方说，随着在生育高峰期出生的那一代人进入退休年纪，在我们可预见的未来，美国民众所面临的社会保险和工资税之间的鸿沟也越来越大。医疗保险计划的收入已经严重不足，并且其严重程度的加剧会比社会保险更快，这是因为在医疗方面的支出不可避免地会增加。经验证明，这些福利计划在政治上是神圣不可侵犯的。

另外还有一些额外的公共开支也非常引人注目。比如说，美国的交通基础设施已经被忽视了数十年。到某个时刻，政府真的不得不修补道路和桥梁。美国还没有高铁系统，但许多发达程度不如美国的国家都已经开始建设高铁。⁸随着能源价格的不断飞涨，智能能源网对于公共投资来说越来越具有吸引力。⁹许多地区甚至仍缺乏基本的公共交通系统。政府还需要隔离苏联泄露的核材料。他们应该撤回之前针对科学研究资金削减，因为这是竞争优势的重要来源之一。种种需求不胜枚举。

美国政府或许可以为处于边缘领域的重大福利计划找到借口，从而削减各种经费，节省一些钱。但只有傻瓜才会假装相信，哪怕没有重大的额外收入，我们在未来需要做的事情都能做到。尽管我认为，提高最高边际收益的所得税税率是个糟糕的主意，但我们必须准备好一些向社会上最富有人群征税的办法。正如威利·萨顿（Willie Sutton）在被问道为何抢劫银行时回答的那样，因为钱在那儿。如果政府不向富人征税，人们就不可能获得所需的收入。但有一个好消息，那就是我们将会在后面的章节看到，政府可以向最高收入群体征收某些税而不损害他们的利益。

THE DARWIN ECONOMY

THE DARWIN ECONOMY

11

对自由竞争的规制：
对有害行为征税

对有害行为进行征税，不仅能够产生收益，而且还能遏制这种行为。这一简单的观察结果简直令人欢欣鼓舞，因为美国急切地需要更多税收，并且美国的经济正遭受一系列有害行为的损害。对这些行为征税可以达到一石二鸟的效果，一来可以帮助政府实现预算平衡，二来可以遏制那些弊大于利的行为。

庇古税

在讨论具体案例之前，我最好先解释清楚“对于有害行为，税收比直接监管更有效”背后的逻辑。这个观点是由英国经济学家庇古首先提出来的，因此对有害行为征税被称为庇古税。¹设想有两家公司，一家是清洁型，一家是污染型，整个硅谷地区仅有它们两家排放二氧化硫污染气体。它们中的任何一家都拥有三种生产技术。使用不同技术的区别仅仅在于生产成本不同和二氧化硫的排放量不同。生产成本与相应的二氧化硫排放量请见表11-1。请观察表格中不同变量的模型，它们揭示了对付有害行为的不同模式之间一些重要的普遍属性。

表11-1 不同生产模式对应的生产成本及二氧化硫排放量

	最清洁模式 (0 吨 SO ₂ / 天)	中等清洁模式 (3 吨 SO ₂ / 天)	污染最重模式 (6 吨 SO ₂ / 天)
清洁型公司生产成本	\$100/ 天	\$40/ 天	\$10/ 天
污染型公司生产成本	\$4 500/ 天	\$1 500/ 天	\$500/ 天

首先请注意，从左往右看，你会发现，最清洁的模式成本都是最高的。原因很简单：过滤器的质量越好，成本越高。可以看到，成本的增加可以消除一部分污染，消除得越多，成本越高。这是因为，理性的经

济活动者通常会首先选择最划算的方式，其次才是第二划算的。比方说，清洁型公司如果每天要减少3吨的排放量，成本仅为30美元（也就是污染最重模式与中等清洁模式的成本之差），但如果要继续减少3吨，成本则需要每天60美元（最清洁模式与中等清洁模式成本之差）。

成本变量还揭示了一个重要的模式，不同企业消除污染的成本是不一样的。请注意，第一次减少3吨二氧化硫的排放，污染型公司需要每天花费1 000美元，再减少剩余的3吨排放量，则需要每天花费3 000美元，而清洁型公司的花费则少得多。

我们可以合理假设，受二氧化硫所害的人群不太可能直接去找污染者协商。在经济学的标准假设中，企业追求自身利益的最大化，也就意味着他们将会在法律允许范围内选择最低廉的生产方式。如果缺乏监管，那么这两家公司都会选择最低廉模式，也就是污染最重的模式，从而导致二氧化硫的日排放量达到12吨，也就是每家6吨。

消除污染并不是最优的结果

对于该地区来说，二氧化硫排放量达到何种水平是最理想的呢？许多非经济学专业人士会脱口而出：零。然而，在弄清二氧化硫排放的成本前，我们无法给出正确的答案。毕竟，减少排放也需要成本，如果二氧化硫带来的危害足够低，最好的结果就是容忍它的存在。但是，为了使讨论更富趣味性，我们不妨假设，每吨二氧化硫造成的损失是每天40美元不变。如果企业采取污染最重的模式，总损失将达到每天480美元。

清洁型公司为减少二氧化硫排放量所付出的成本远远低于污染型公司。减少总排放量最低成本的做法是，首先让清洁型公司从污染最重模式转为中等清洁模式。该举措增加的生产成本仅为每天30美元，如果排放二氧化硫，造成的损失将会是120美元，所以很明显，这么做是值得

的。要想进一步减少排放，最划算的第二步做法是让清洁型公司从中等清洁模式转为最清洁模式，这样可以每天再减少3吨排放量，而增加的成本为每天60美元。由于再次避免了120美元的损失，这一步也是值得的。至此，清洁型公司已经完成了它所能做的全部任务。

如果我们还希望进一步减少排放所带来的损害，接下来就应该让污染型公司从污染最重模式转为中等清洁模式。然而，这一举措需要每天花费1 000美元，远远超出了排放将导致的120美元损失。在这个例子中，污染型公司消除有害事物的成本太高了，因此它不可能为减少有害气体的排放做贡献。我们只能接受每天6吨二氧化硫的排放量，别无他法。

许多人很难接受，存在一定程度的污染排放量才是最优的结果。但事实上，排放量减少到一定程度后，减排成本反而会比排放本身所带来的损失要高。如果有人坚持认为，在任何环境中，任何污染物的最优排放量是零，那么你可以问问他，为何不立即用吸尘器吸一遍他的卧室。打扫之后每过一分钟，积累的灰尘都在增加。如果他坚信灰尘最好是零存在，那么他应该持续不断地吸尘，或者雇一个人帮他这么做。

这样做很傻。吸尘会占用宝贵的时间，但其实一丁点灰尘不会带来什么伤害。这个逻辑同样也适用于其他环境的其他污染物，也就是零排放量不一定是最好的结果。最优的减排方案，是首先采用成本最低的减排方法，然后进一步减排，直到所降低的损失不足以抵消所花费的成本。

这个简单的案例说明，政府指挥与控制式的监管不够有效。美国过去对于污染物的监管方法是，要求所有排污方在一个基准年里，按照同样的比例降低排放量。比方说，如果监管者希望减少一半的二氧化硫排放量，他们就得要求清洁型公司和污染型公司都从污染最重模式调整为中等清洁模式。这样就能实现既定的减排目标，但总共增加的经营成本高达每天1 030美元：其中污染型公司增加1 000美元，清洁型公司增加

30美元。而清洁型公司从污染最重模式转变为最清洁模式，减排效果一样，但只需花费90美元，比现在的总成本低940美元。

要求所有排污方按照同等比例减排是一种浪费行为，因为这完全没有考虑到有些排污方的减排成本要比其他排污方低得多。监管者只需指导清洁型公司转变为最清洁模式，便可有效实现目标。总的来说，监管者对于不同企业能够采取哪些不同的减排手段缺乏足够细致的了解，期望他们能够细致地管理到这种程度也是不现实的。

为何征税手段能够实现最低的减排成本

监管方可以确保以成本最低廉的方式实现减排目标。在上述案例中，假设监管方对每吨二氧化硫的排放征收40美元的税。清洁型公司的反应将是污染最重模式调整为最清洁模式，因为这样每天只花90美元就能避免240美元的征税。而污染型公司，将继续采用污染最重模式，因为转为最清洁模式所节省的税收，相对于它所花费的生产成本来说，简直是九牛一毛。

为此，清洁型公司将竭力减排，这样做的效益很高，因为它的减排成本远远低于污染型公司。但由于清洁型公司不再排放二氧化硫，它就不用缴纳排污税。利用税收手段调节减排，不仅有效，而且十分公平。

相对于不对排污采取任何措施来说，税收手段将提升全社会的整体经济福利。征税唯一产生的成本是清洁型公司转变为最清洁模式所需的每天额外90美元的花费。由于降低污染量可以避免每天240美元的损失，全社会的净利润为每天150美元。污染型公司每天缴税240美元，从公司角度来说是一种成本，但对于整个社会来说却不是，因为有了这份税收，就可以从其他渠道减少同等额度的税收。

要想弄清税收手段的效力体现在何处，关键要意识到二氧化硫排放所带来的危害取决于它们在大气中的总浓度，而不在于排放者是谁。社

会的利益在于降低减排的总成本，而非规定指定的排放者减少指定的排放量。税收手段通过激励排污企业以最低的成本采取最大限度的减排措施，从而实现了成本最低化。

这个假设案例表现出的乐观一面在现实中得到了验证：1990年，美国国会将二氧化硫的排放许可证纳入《清洁空气法案》（*Clean Air Act*）的修正案中。二氧化硫是导致酸雨形成的重要原因，20世纪七十年代，在美国东北部，二氧化硫问题导致的损失越来越大。新法令规定，企业需要按吨申请排放许可。随着许可证颁发量的减少，总的二氧化硫排放量也在逐渐降低。

由于企业可以通过芝加哥交易所的拍卖活动对许可证进行买卖，许可证的功能就相当于对二氧化硫的排放进行征税。要想排放额外的数量，许可证的拥有者需要放弃售卖许可证所能带来的收益，因此，使用许可证的机会成本从本质上来说就是一种潜在的征税。

这个项目获得了极大成功，人们提前完成减排任务，并且所花费的成本低于监管与控制的成本。²过去媒体经常发表的有关酸雨问题的文章，现在也完全不见了踪影。

征税的目标

在环保运动的初始阶段，经济学家们就积极倡议采用许可证或税收手段来减少污染，然而，左派环保团体和部分评论家对此持反对态度。据报道，一位环保主义者曾经问道：“接下来，洛杉矶警察局是不是准备从威斯康星州购买公民权利的交易证呢？”³诸如“经济学家希望有钱的企业可以随心所欲地污染环境”等批评比比皆是。

这种言论其实违背了企业行为背后的简单目的。企业不污染环境是因为它们能够从中获益，企业污染环境是因为减轻污染需要付出成本。尽管将经济学家的提议变成法律需要数十年时间，环保团体现在已经成

了税收手段最积极的拥护群体之一。例如，环保组织塞拉俱乐部劝说俱乐部的支持者们购买二氧化硫的排放许可证，然后将许可证撕碎，从而使排放量更低于监管目标。

自由主义者或其他自由市场保守派反对征收二氧化硫排放税的理由是什么呢？目前，我们可以先假设运用经济学家科斯提出的框架是正确的思路（见第6章）。部分自由主义者可能不会同意，但正如我在后面章节解释的那样，他们很难驳倒这个观点。许多自由市场的热衷者非常乐于接受科斯的框架，因为他们相信，这有利于最大限度地减少政府的监管足迹。这种想法是基于科斯的论断：如果受有害活动影响的各方之间进行谈判是切实可行的，那么他们就会有强烈的动机自行达成有效的解决方案。

那样很好。但科斯充分意识到，由于存在交易成本，人们相互之间通常无法进行磋商。此时，有效的解决方案可能来自政府的干预。科斯认为，如果无法进行私下磋商，政府应该实行制度安排，以便指导各当事方在条件允许的情况下自行磋商。

由于二氧化硫的排污方有数百家，数百万人受此影响，私下协商解决方案绝对是不现实的。为此，科斯的框架提出，法律应该将减排的重担分给那些完成成本最低的参与方。

消费者对于此类问题无法提出具有吸引力的建议。相反，排污者则可以有很多方法，例如，在烟囱上安装净化器，或者使用低硫煤炭。科斯的框架认为，对于酸雨的最好补救措施，是用一种能够激励生产者减少二氧化硫排放的方法来确定产权。征收二氧化硫的排放税，或者实行许可证管理，这是能够解决问题的最有效的办法。这些办法对企业的干扰也最小。

这种税收不是窃取。的确，它会导致排放二氧化硫的厂家提高产品价格，但提高的价格反映的是排污方对他人造成的伤害的经济价值。任何会产生恶劣副作用的产品的价格，都反映出伤害所对应的经济价值，

正如价格能够反映出劳动力和原材料的成本一样。税收能够阻止二氧化硫造成经济损失，这部分的经济价值高于增加的产品成本，这就使整体的经济产出价值更大，而不是相反。公民不是征税的受害人，而是受益人。

支持二氧化硫排放税的立法者不会坦白，位于华盛顿的官僚并不比你更清楚应该如何明智地花你的钱。相反，他们明白一个道理：如果宝贵资源是免费的，例如我们所呼吸的空气，那么人们使用它的效率就会很低。实际上，这种征税使你能够购买更多你所珍视的东西。

如果自由主义者反对征税的花言巧语阻碍了这种税收的实施，结果就会像拉尔夫·纳德发起的航空消费者行动计划一样，在发生超员预定的情况时，航空公司不再为让出座位的志愿者提供补偿。我曾在第7章讨论过，使用价格系统来调配稀缺资源能够促使“经济蛋糕”做大，无论这种资源是超员预售的飞机上的座席，还是对二氧化硫容纳有限的空气。

如果为了保证二氧化硫排放税能够成功通过立法，需要进行支付转移，自由主义者同样没有理由反对。使用价格系统来调配稀缺资源能够促使“经济蛋糕”做大，但这并不能保证每个人分到的那块“蛋糕”就会自然而然地比之前大。例如，污染型公司服务的大多数客户可能收入很低，公司缴纳二氧化硫排放税会使产品价格上涨而令他们难以承受。或许可以通过提供补给食品等方式，加强他们的社会保障体系，从而使得他们不再反对上述征税。重点在于，政策变化使“经济蛋糕”变大之后，从中获益的人就有可能对那些遭受损失的人进行补偿。我在第7章曾提出，由于缺乏这样的补偿机制，许多政策无法实施，无论是对富人还是对穷人来说，这都是一种损失。

气候变化与二氧化碳税

二氧化硫排放税的原理也基本适用于对二氧化碳排放进行征税的案例。二氧化碳在大气中的浓度越来越高，它被认为是导致全球变暖的罪魁祸首。二氧化碳排放税的反对者强调，关于气候变化预报的不确定性极高，因此他们反对采取行动。然而，不确定性是一把双刃剑。气候研究专家本身也承认，根据他们提出的模型所得出的预计结果不确定性极高。但这也说明，尽管实际后果可能比预测的中等情况乐观，但也意味着有可能更糟糕。

2009年哥本哈根全球气候大会的组织者提出，希望到21世纪末，将全球变暖的温度控制在3.6华氏度以内。但即便上升的度数极少，也会导致致命的损失，并且根据最受认可的气候变化模型预测，如果听之任之，平均温度的上升值小于目标值的可能性几乎不存在。

根据麻省理工学院开发的地球综合系统模型的最新预测，到21世纪末，如不进行有效干预，上升温度的中位数将为9华氏度。⁴该模型还预计，上升温度超过12华氏度的可能性为10%。如果达到这种程度，永久冻土将会融化，并释放大量甲烷进入大气。而甲烷造成温室效应的力度是二氧化碳的50多倍。因此，根据麻省理工学院模型的预测结果，全球变暖而导致地球上大部分生物灭绝的可能性为10%。

确实，气候模型的预测具有极大的不确定性。事情可能不像预计的那么糟糕，但也有可能更糟。我们应该采取行动吗？在回答这个问题之前，我们必须问一问，采取行动需要付出多大的代价？而答案令人震惊：几乎为零。

联合国政府间气候变化专门委员会预计，到2030年，须对每吨碳排放征收80美元的税，才能在2100年前一直维持气候稳定。⁵如此高额的税费将导致每加仑（约3.79升）汽油的价格上涨70美分。然而，在预估出这些数值之前，他们还不知道麻省理工学院的悲观模型。为稳妥起见，我们假设税费为每吨300美元。此时，货物价格的上涨与碳排放量成正比，拿汽油来说，每加仑汽油的价格将上涨2.6美元左右。

油价突然之间如此大幅度上涨会带来许多痛苦，在2008年美国人曾经历过。但如果逐步进行，它带来的伤害就会小得多。比方说，面对不断上涨的燃油价格，制造商就会争先恐后开发出更高效率的汽车。现在，欧洲民众比美国民众为每加仑汽油多付4美元。但正是因此，世界上许多燃油效率最高的汽车都出自欧洲的汽车制造商之手。实际上，欧洲民众在汽油上的花费低于美国民众，但享受到的乐趣一点儿也不比他们少。

如果一个家庭卖掉福特野马（约6.33公里 / 升）而购入福特福克斯（约13.5公里 / 升），在行驶同样里程的情况下，消耗的油量反倒少了。征税制度可以慢慢引入，要给人们一些时间来适应。比方说，人们可以搬到离工作地点更近的住处，选择拼车，选择距离更近的度假地点，等等。税收可以用于支持低收入家庭，从而减轻汽油涨价给他们带来的负担。政府还可以用一部分税收来减轻负债、重建损坏的基础设施，或者减轻其他赋税。

2009年，美国众议院曾通过一项能源法案，法案里包含了关于碳排放限额与交易综合系统的内容，其功能相当于征收碳排放税。尽管许多共和党立法者长期以来倡议总量管制与排放交易立法，希望以此作为减轻环境外部因素问题的框架，但自由主义者和共和党中的茶党却坚定地反对类似的政策，并谴责此类政策为社会工程。

当然，这就是社会工程！征税的原因在于，当前二氧化碳的排放实际上是免费的，而人们排放得太多，几乎对每个人都造成了威胁。然而，如今没有任何一个共和党国会候选人对二氧化碳的总量管制与排放交易表示支持。共和党在参议院获得了足够多的选票，甚至能够通过阻止支持性的选票进入参议院，从而阻止能源立法。

且不说这样做的赌注过高，他们用以论证自己立场的逻辑也十分滑稽。来自俄克拉何马州的共和党参议员詹姆斯·英霍夫（James Inhofe）曾声称：“全球变暖是人为因素导致的，这一说法其实是错的，这么说

根本没有任何科学基础。”这个声明是荒谬而错误的。只有一小部分研究气候的科学家会假装同意这种说法，实际上这部分科学家大多获得了能源行业的资金支持。大部分气候科学家相信，不受限制的温室气体排放将导致地球平均表面温度持续上升。唯一不确定的是，到底会上升多少。灾难性气候变化的风险可以通过一项税收政策轻易地消除。然而，反对税收的狂热者却时刻准备着不让这样的结果出现。活跃的自由主义者人数不多，但他们的主张所造成的破坏程度却很严重。

拥堵费

对其他伤害他人的活动进行征税，上述逻辑也完全适用。假如你进入一条拥挤的道路，阻碍了他人的前进，那你对他人就造成了伤害。自由主义者和其他自由市场保守主义者没有理由反对征收拥堵税，从2003年2月起，工作日进入伦敦市中心的车辆每天都要缴纳14美元。这项费用所产生的直接结果就是，交通量下降了1/3，部分公共汽车的行程时间减少了一半，二氧化碳的排放量减少了20%，并且城市雾霾的主要成分，也就是微粒和氮氧化物的排放量也得到了显著抑制。⁶

正如之前讨论二氧化碳排放税时论述的那样，缴纳这项费用并继续开车进入伦敦市中心的人所付出的成本是最高的，其他人则会将行程调整到非高峰时间段，或者选择乘坐公共交通。拥堵程度的降低和排放物的减少所带来的总价值，远远高于人们做出改变所带来的损失。此外，征收这些费用后，政府从其他方面征税的需求就会减少。如果说我们必须征税，那么对有害行为进行征税比起对有益行为进行征税则更有利。

2007年，纽约市市长迈克尔·布隆伯格（Michael Bloomberg）提出征收类似的拥堵费，对工作日驶入曼哈顿南部86号大街的汽车收费。早6点到晚6点期间驶入的小汽车须缴纳每天8美元的费用，商用卡车缴纳21美元。相对于伦敦的收费计划，尽管市长的提议能够产生净收益，但

批评者马上提出这对穷人不公平。例如，纽约市专栏作家刘易斯·菲德勒（Lewis A. Fidler）说道：“这将城市分裂为富人的城市和穷人的城市，并且费用单其实是在宣称能够支付得起的人可以进来，支付不起的人滚蛋。”⁷

尽管市长的提议获得了纽约市议会的预批准，反对者还是说服了纽约州首府奥尔巴尼的立法者对提议进行阻挠。其实，批评者的顾虑很容易消除。首先，由于过桥费和停车费很高，穷人很少会开车到曼哈顿上班。大部分人原本选择的就公共交通。但是，低收入者偶尔也需要开车进城，比方说带孩子或老人看医生，或者办其他杂事。如果布隆伯格市长受到右派的阻挠少一点，他或许愿意完善他的提议，比方说，对在城中工作的低收入者实行每年10次免征政策。大部分的免征资格可能最终会被拿到克雷格列表网^⑥出售，售卖所得的收益很有可能会让人们放弃反对的立场。

然而，这样的补充方案会导致活跃自由主义者和其他自由市场保守主义者的抗议。“政府无权对我征税，然后还用我的钱补贴穷人！”此类口号的效果十分惊人。但是，它们确实阻碍了那些可以为每个人带来更好生活政策的实施。

按照汽车重量征税

穆勒的伤害原则还提出，按照汽车重量进行征税并立法。多年来，美国人购买SUV（运动型多功能车）的数量稳步上升，因为很多人认为，这种汽车比小轿车更安全。事实却并非这么简单。假设在其他条件相同的情况下，如果两辆汽车的车头发生碰撞，重量更大的汽车所受伤害较小。比方说，一辆3 266千克的福特远足与一辆1 134千克的本田思域车头相撞，你肯定希望自己乘坐的是福特。但由于SUV的重量更大、重心更高，大型SUV通常更难操控，从而更难避免发生碰撞，并且在进行避险操作时更容易翻车。总的来说，普通SUV的安全性比不上普通轿

车。⁸

然而，就汽车对他人造成的风险而言，毫无疑问，SUV是绝对的赢家。车辆越重，越容易导致其他车辆的乘客死亡或严重受伤。

正如二氧化硫的案例那样，人们也不可能对这种特定的外部效应进行私下协商而获得解决方案。如果政府的挑战在于界定产权，以便获得最接近于人们有条件进行自行协商时所能达成的结果，最简单的方案莫过于按照汽车重量进行征税。此类征税能够让车主在做出购车决定时考虑到为他人带来的风险。而在当前的情况下，他们没有任何动力来考虑这些风险。如此征税还有一个优势，那就是可以降低其他有益活动的税率。我不知道还会有谁有理由坚持说，按照汽车重量进行征税会剥夺汽车车主应有的权利。

烟草税

烟草税也有助于减少伤害，并且有着同样诱人的优势。毋庸置疑，吸烟对于吸烟者本身是有害的，也有确凿证据证明，二手烟也会对他人造成重大伤害。事实上，大部分吸烟者都宁愿自己没有养成吸烟的习惯。一个人是否会养成吸烟的习惯，其实可以通过一个可靠的指标来预测，那就是他朋友吸烟的比例。烟草公司的游说者试图以此为证据，来证明征收更高的烟草税并不会降低吸烟率。

一位烟瘾严重的成年吸烟者确实很可能不会因为烟草税提高而戒烟。但是，很多人是在青少年阶段养成吸烟习惯的。由于青少年的可支配收入十分有限，烟草税对于他们是否吸烟的影响就大得多了。

更重要的是，同伴效应会对双方都起作用。如果烟草税阻止了青少年吸烟，他们同龄群体吸烟的比例就会降低。这会进一步防止其他人吸烟，这种效果可以不断传递。已有研究发现，吸烟者在烟草税提高的时候更快乐。⁹大部分吸烟者都有戒烟的意愿，他们似乎明白，在吸烟者

更少、香烟价格更高的环境下，他们的目标更容易实现。提高烟草税有利于创造这样的环境。

吸烟者对他人的伤害，不仅仅在于将他人暴露在二手烟的环境中，还在于使他人更容易染上吸烟的习惯。假若没有烟草税，世界上吸烟的人数就会增加。如果你的孩子在这样一个世界长大，他们变成吸烟者的可能性会更大，患肺癌的可能性也就更大。他们更有可能患上肺气肿以及心脏方面的疾病。他们死于房屋火灾的概率也会更高。父母们为什么会希望自己的孩子在这样的世界里成长，而不愿意让烟草税减少这些可能性，并且让有益的活动得到税收减免呢？

酒精税

关于酒精税的争论更为复杂。与吸烟者不同的是，大部分酒精饮料的消费者没有明显的后悔的迹象。大部分饮酒者也不太可能对不饮酒的人造成伤害。在美国，由于饮酒造成的伤害，大部分是由占据了40%酒精消费量的5%饮酒者引起的。¹⁰这部分人伤害他人的方式是各式各样的。他们会忽视、虐待家人，他们会因为酒后驾驶而伤害或害死别人，他们会对别人进行肢体攻击，等等。

反对提高酒精税的人往往会提出这样的理由：长期酗酒的人不会为此改变太多。但已有证据却给出了相反的结论。诚然，许多饮酒者会继续大量饮酒，并且他们的饮酒程度给他们的事业和婚姻带来了极大的伤害。所以，他们不会仅仅因为提高酒精税之后，酒的价格提高而减少饮酒量。但正如经济学家长期以来所强调的那样，选择的核心在于物品稀缺与否。

如果一样消费品在总体消费中所占的比例很小，那么大家通常不会在意它价格的提高。打个比方，如果食盐的价格翻倍，大部分人的消费量不会产生变化。但长期酗酒的人在饮酒上的花费占据了他们收入的很

大一部分，这不仅是因为他们喝得多，也因为酗酒会导致他们收入减少。所以，他们不能承受酒的价格进一步提高。

要想研究酒精税究竟会对酗酒者造成何种影响，我们面临着一个挑战，那就是销售数据不能显示是谁购买的以及谁喝的酒。经济学家菲利普·库克发明了一个聪明的办法来规避这个问题。¹¹观察显示，大部分与酒后驾驶相关的车祸是由司机长期酗酒引起的，并且，酗酒者占肝硬化死亡人数的绝大部分。基于这些结论，他获得了一个重要发现：如果在某个州提高酒精税，此后与酒后驾驶相关的车祸数量以及肝硬化死亡人数都会显著下降；但在没有提高酒精税的州，就不会出现这种情况。尽管大部分患有肝病的酗酒者知道最有效的治疗方法是少喝酒，但这个知识本身不足以让他们改变自己的行为。但在面对价格刺激时，他们中的许多人很快就能做出改变。

还有人反对酒精税的理由是，为了抑制一小部分不负责任的酗酒者的行为而惩罚大部分人不公平的。然而，事实上，占据了大部分酒消费量的酗酒者所承担的酒精税也会占大部分。酒精税带来的收益可以使政府降低中度和轻度饮酒者的所得税税率，并补偿减少的这部分税收。

对造成非直接伤害的行为征税

本章讨论的伤害行为全都属于个别人对其他人造成伤害，并且他们自己无法消除这种伤害的案例。例如，作为一名驾驶者，你可以尽量小心，但你还是有可能被冲过停车标志的SUV撞到。目前讨论的都是会对他人造成直接的、身体伤害的案例。因此，正如第1章讨论的那样，按照穆勒的伤害原则对这些行为进行极其严格的管控，也是合理合法的。

第5章中曾提出，在许多消费行为导致的伤害案例中，受害者也无法通过自己的努力避免遭到伤害。但那些案例通常不会直接造成身体上

的伤害。相对于对造成直接身体伤害的行为进行征税，许多自由主义者更难以接受由于改变社会参考体系而导致的问题。但在下一节中，我将论述是否造成身体上的伤害并不是决定因素。真正的问题在于，要想避免由他人的行为所带来的伤害，其成本是多么高。例如，如果别人不惜借更多钱来买好学区的房子，你就没有好的解决办法。你可以自己也去借钱，承受随之产生的财务风险；你也可以不去借钱，但不得不把孩子送到不好的学校。

自行车与摩托车头盔法规

还有一个更加复杂的综合案例，那就是骑自行车和摩托车戴不戴头盔的问题。在我儿子14岁那年，他从自行车上摔下来，并且摔得很重，头部和肩膀重重着地。幸运的是，他只是轻微脑震荡，并且断了一根锁骨。但急救室的医生告诉我，如果他没有戴头盔，这次的伤害可能致命，因为头盔的右半部分已经被摔得粉碎。

他之所以戴了头盔，是因为纽约州的法律规定，16岁以下骑自行车的人必须戴头盔。部分家长反对这项规定，声称为孩子制定安全规则是他们自己的责任，而非州政府的责任。

但我知道，即使没有发生这次事故，我和其他一些父母也会感激这一规定。许多男孩将冒险行为当作一种荣誉，如果没有这样的规定，戴头盔会被其他同龄人嘲笑。在一个理想的世界里，父母能够成功帮助孩子不去理会这种压力。但我们并非生活在理想的世界。我们清楚，如果不强制规定，无论父母怎么努力，戴头盔的孩子都会变少。

纽约州关于戴头盔的规定能通过穆勒的测试吗？也就是说，它是否能够通过限制一部分人的行为，成功避免这些人对其他人造成伤害？毫无疑问，这项规定避免了大量的伤害。若非如此，许多孩子，也许包括我儿子，早就已经丧命。他们失去自己的生命，这也是他们父母以及其他

关心他们的人的损失。即便不对这些损失进行详细的估算，我们也可以肯定，没有人会怀疑这种损失是多么重大。那么，问题就在于，因戴头盔的规定而避免的伤害，相对于被强制要求戴头盔所带来的伤害，孰轻孰重。

当然，许多支持这项规定的人之前都有戴头盔的习惯。即便如此，他们中的许多人依然会暗自感谢这项规定，因为他们发现如果大家戴头盔，自己也不会觉得尴尬了。基于同样的原因，之前没有戴头盔习惯的人也很乐意被要求这么做。对于这两类人来说，这项规定并没有造成任何伤害，这就跟戴头盔的规则基本上不会对冰球运动员造成伤害是一样的。但不可避免的是，仍有一小部分人会觉得这项规定对于他们来说是一种冒犯，但是他们受到的伤害绝对轻于这些规定所能避免的伤害。

在第7章我们曾讨论过，在严格应用穆勒的伤害原则之前，需要对当事双方所受伤害的经济价值进行可靠的评估。纽约州立法机构在不经估算的情况下就通过了关于戴头盔的立法，或许是因为大部分的立法者都认为，这项规定所能避免的伤害将远远大于它所带来的伤害。

反对头盔规定的自由主义者可能提出以下论点：规定所导致的冒犯和反感，比它所能预防的伤害更严重。然而，几乎没有自由主义者会采取这个论点。他们只是在说，州政府没有权利指导人们如何生活，所以不应该通过这项立法。除非他们完全不打算应用穆勒的伤害法则，否则这样做是行不通的。如果自由主义者有权利阻碍大部分人所期待的青少年戴头盔规定的实施，那也就是说他们拥有将伤害强加于家长和其他人身上的权利。请问他们凭什么拥有这样的权利呢？

部分自由主义者或许会反对，认为举这个例子对他们来说不公平，州政府自然有权利实施青少年必须戴头盔的规定，因为毕竟青少年没有足够的判断能力和经验，无法在面对高风险行为时做出慎重的决定。但证据表明，青少年成长为成年人后，并不会发生任何奇迹。缺乏深谋远虑、天真的乐观，以及在面对社会压力时的脆弱，这些问题或许会随着

年龄的增长而逐渐好转，但这些问题在大部分成年人身上依然非常明显。同孩子一样，成年人也有许多爱他们的人，如果受到严重或致命伤害，那些人也会无法承受。

有一年，我在巴黎休假时，曾与一位无法忍受交通拥堵而每天骑车45分钟上班的同事有过一次对话。尽管在巴黎，自行车事故非常常见，这位聪明、成熟的女士也从不戴头盔。我告诉她我儿子经历的事故，并劝说她考虑戴头盔。在她提出反对意见时，我故意揶揄她是由于害怕看起来不时尚才拒绝的。她对此提出强烈抗议，我也完全相信，她并不在意时尚与否。然而，数周之后，她羞怯地向我坦白，其实她曾在周末尝试过戴头盔，但不能想象自己戴头盔出现在大家面前是怎样一种情形。这不足为奇，因为在巴黎的一整年里，我从未见过任何一位骑自行车的女性戴头盔。特立独行确实会让人感觉尴尬。

我也无法想象法国颁布一项强制戴头盔的法律（过去我也曾无法相信他们会通过一项法律，禁止在酒吧和餐馆吸烟）。能够接受穆勒伤害原则的自由主义者不会断然反对这些法规，哪怕是在成年人身上实施的法规。因为社会环境会在很大程度上影响人的决策，许多人发现，要避免他人行为所带来的影响，成本非常高。无论是否强制戴头盔，都会有人会受到伤害。如果发生这种情况，我们的共同利益在于最大限度地减少总体伤害。

在自由主义者的强烈反对声中，许多州都通过了法规，规定骑摩托车的成年人也必须戴头盔。我那位曾经骑自行车时受伤的儿子现在已经成年了。他不骑摩托车，但如果要骑，我会很感激纽约也在这几个强制戴头盔的地区之列。但人们常常迷失在关于是否应该强制规定戴头盔的讨论里，而忽略了比起直接规定，其实还有其他策略，可以让被这项规定冒犯的人承受的成本更低。

对于成年人是否应该戴头盔这个问题，部分有戴头盔习惯的人也十分敏感。在一个只有少数人戴的环境里，例如在巴黎时我的同事所处的

环境，少数戴头盔的人会被认为是另类。一个显而易见的事实是，极少数人戴头盔的状况将会持续，哪怕大部分人都希望自己和他人戴头盔。我们还知道，不戴头盔可能会造成严重伤害，不仅仅是伤害当事人，而且也会伤害那些关心他们的人。因此，要求戴头盔所能减少的总体伤害，使在任何环境下都不愿意戴头盔的人所承受的伤害，不再那么重要。

然而，自由主义者强烈反对这项规定。他们称任何愿意戴头盔的人都有自由去做的权利。他们没有否认与众不同所导致的尴尬会阻碍一些人戴头盔，但那其实是在说，那是那些人自己的问题。或许，那些人应该更认真地思考什么是真正重要的，并且克服他人的看法。如果采取另外的办法，即强迫不愿意戴头盔的人必须戴，那么这部分人将无从选择。但我仍要说，如果在一个环境里，大部分人平时都不戴头盔，并且只有当他人开始戴头盔，自己才会戴，但他们仍希望包括自己在内的大部分人都去戴头盔，那么这个环境是有问题的。

幸运的是，强制要求戴头盔并不是解决这个问题的唯一途径。我们可以采取一种弹性的戴头盔要求，那就是允许任何人都能以适当的价格购买豁免权。比方说，每年缴纳300美元，可以从机动车辆管理局购买许可证，然后将它贴在牌照旁，告诉交警这辆车的车主获得了不戴头盔的豁免权。

这个办法的逻辑跟征收二氧化硫排放税的逻辑是一样的。对于某些企业来说，降低排放量所付出的成本远高于其他企业；类似地，对于某些人来说，必须戴头盔的成本也会比其他人高。正如二氧化硫所带来的伤害取决于总排放量，阻碍人们戴头盔的社会压力取决于不戴头盔的人所占的比例。在上述案例中，有效的解决方案都是让最容易采纳必要改变措施的人做出改变。征收二氧化硫排放税，会引导减排成本低于应缴税款的企业采用清洁技术进行减排，而对不戴头盔的人征税，能够鼓励并非强烈反对戴头盔的人做出改变。

如果自由主义者仍对征税感到不舒服，他们可以这么考虑：这不仅减少了伤害，还能增加税收收入，从而为有益的公共服务买单。这样可以降低对有益活动征税的税率。对不戴头盔骑车的人征税，意味着政府可以减少征收工资税，工资税不利于产生更多的工作机会；或者可以减少征收所得税，所得税不利于人们进行储蓄。

对有害行为征税是一种滑坡谬误吗

我将以近年来一项最受争议的征税提议案例作为总结，那就是针对含糖苏打饮料的征税提议。含糖苏打饮料被认为是引发肥胖症流行的主要因素。在自由主义者和其他人的反对声中，华盛顿特区通过了这项征税的立法。在想办法筹集奥巴马政府提出的医疗保险改革法案所需费用的过程中，参议院财政委员会在经过考虑后，拒绝了一项类似的征税立法。纽约州前任州长戴维·佩特森（David Paterson）为了尽快解决州政府日益扩大的财政赤字，需要额外的收入，从而提出了这项没有成功的征税提议。

批评含糖苏打饮料税的人认为，这是保姆式州政府对人们的终极侵扰。对于他们来说，这跟那些会伤害他人的案例是有差别的。由于过量饮用含糖苏打饮料而肥胖的人，在饮用高果糖玉米糖浆时，也更有可能患上糖尿病和心脏病。他们更有可能被足部截肢，也更有可能在成年前死亡。

然而，这些昂贵的成本都是饮用者给自己增加的成本，不影响别人。当然，过量摄入玉米糖浆所导致的医疗并发症通常需要由公共支出所支持的治疗。但如果苏打饮料饮用者拥有正常的寿命，这些支出在他们临终前也会发生。正如许多经济学家所指出的，在成年前死亡其实为政府节省了数十亿美元的社会保险开支。

简而言之，征收含糖苏打饮料税更像是一种家长式的管制，而不是

对危害他人的行为进行征税。然而，毋庸置疑，含糖苏打饮料税将以一种取悦许多纳税人的方式来改变他们的行为。正如我在第2章中讨论过的那样，证据表明，人们总是过于看重眼前的成本和利益，过于轻视具有不确定性的或者延后发生的结果。许多人购买含糖苏打饮料的理由是这些饮料很好喝，而不是希望自己未来变胖。数年之后，如果税收让他们减少了购买这些饮料的次数，很少人会为之前的行为感到后悔。

从行为学的角度来说，未来的自己对于现在的自己来说，无异于一个陌生人。假设喝含糖苏打饮料引发的即时后果，是导致一位完全陌生的人由于糖尿病循环障碍而被足部截肢。如果这种随机的联系一定会发生，几乎没有人会喝太多的含糖苏打饮料。但由于受害者是存在于遥远未来的自己，并且伤害程度不确定，为避免伤害而限制自己有害行为的动力就会大大减弱。

为反对含糖苏打饮料税，经济学家格里高利·曼昆（Gregory Mankiw）担心批准这些征税政策就会导致滑坡谬误^[7]。他写道：“征收苏打饮料税会鼓励人们保证营养，更好地照顾未来的自己。于是，也可以对糖果、冰激凌和油炸食品征税。接下来，可以资助西兰花、健身房税费和牙线。向愚蠢的电视秀征税以及为严肃文学提供补贴也就离我们不远了。”¹²

这种担忧有其合理性，但我们不得不开始“滑坡”。在进一步经历这个“滑坡”之前，我们无法消除这种担忧。饮用含有大量高果糖玉米糖浆的苏打饮料显然会造成重大损害。但只要我们还对储蓄、创造工作机会以及其他有益的行为进行征税，通过对有害行为征税来取代对有益行为的征税就迫在眉睫。

THE DARWIN ECONOMY

THE DARWIN ECONOMY

12

重新思考自由主义者的反对主张

每个人对于美好生活的构想都不尽相同。几乎人人都重视个人的自主权，自由主义者尤甚。然而，哪怕是最坚定的自由主义者，每天都会放弃许多珍贵的选择权，以便获得其他权益。举个例子来说，如果你是工薪一族，就必须配合他人的日程安排，并且去做许多你并不喜欢做的事情。即便自主经营事业，你的时间也不可能完全属于自己。但是你能坦然接受这样的事实，因为你的放弃使你获得了其他重要的选择权。

简而言之，拥有完全的自主权是一种不可能存在的极端情况。自由主义者所面临的挑战，跟其他人一样，在于如何获得最大价值的选择权组合。

征税将直接削弱你的自主权，无论是何种税，因为你不可能再用这部分钱去做你想做的事情。然而，这并不意味着，纳税就会损害你选择权的总体价值。这取决于许多因素，而其中最重要的因素是，纳税所带来的利益和你直接用税款所能买到的东西相比孰轻孰重。尽管“那本来就是你自己的钱”，但另一方面，公共服务能为你提供很高的价值。正如我在前面章节所述，种种征税还有额外的好处：遏制那些弊大于利的行为。

本章旨在探索一个主题，那就是理性的自由主义者应当自愿承受何种由国家所赋予的限制。我所说的“理性的自由主义者”，是指高度重视个人自主权，并且为了实现自己的目标而对所有选择权进行权衡的人。至于目标为何，我在此不进行假设。自由主义者的目标，有可能是尽量消费，或者尽量积累财富，也有可能是帮助别人，或者成为佛教徒，又或者从政。那么，问题就在于，如果可以任意选择，理性的自由主义者愿意选择何种形态的社会？

“理性的”这个定语十分重要。有些人往往会去抵制一些能够实现自

己目标的制度，理由是这些制度违背了某些虚无缥缈的原则。我在本章的论述并不是为了说服这类人。但我认为，我的论述能够说服理性的自由主义者，诸如现代福利体系之类的社会制度，将超乎想象地，更有可能帮助他们实现目标。

科斯带来的启示

在第6章，我们详细讨论了经济学家罗纳德·科斯所提出的框架，这个框架为我们即将面临的问题提供了行之有效的思路。我们来简单回顾一下，科斯首先提出，伤害他人的行为在本质上会起到相互作用。比方说，汤姆吸烟会伤害山姆，但事实上，伤害山姆并不是汤姆有意为之。汤姆只不过是希望通过成本最低的方式来实现自己的目标。如果我们阻止汤姆吸烟，我们就会伤害汤姆。科斯的洞见的革命性在于，汤姆和山姆在尽量降低双方所受的伤害时，获得了共同的利益。

在第6章中我曾讨论过，通过对科斯的工作进行仔细研究，我们可以确定，现实障碍通常会阻碍双方私下协商，从而无法为外部性找到解决方案。按照他的分析方法，第一步应该分析，如果不存在障碍，事件将会如何发展。为此，他抛出了一个问题：如果不存在协商成本，不存在起草和执行合约的成本，不存在阻碍人们经过讨论得出解决方案的障碍，事情将会如何发展？

科斯认为，在此情况下，人们会选择最有效的解决方案。比方说，如果汤姆戒烟的成本低于山姆所受到伤害的经济价值，无论法律是否有规定，汤姆都会选择戒烟。如果法律没有明确汤姆应为伤害负责，他就可以毫无责任地制造污染。但山姆不会听之任之，为了维护自己的利益，他得付钱给汤姆，以便让他停止吸烟。

如若不然，还有一种更经济的方式，那就是山姆搬家到上风的位置，从而避免受到伤害。如果法律认为汤姆吸烟所引发的伤害是合法

的，山姆可以原位不动，然后用钱来消除自己所受到的伤害。那样的话，汤姆向山姆付钱，让山姆搬家，这是成本更低的方式。

重要的是，如果双方不能就最有效的解决方案达成一致，那么事情就会恶化。科斯的框架描述了一种鲜明的对比，那就是在效率和自主性之间存在着尚未被人意识到其重要性的联系。由于这种联系十分重要，为方便讨论，我们可以将它明确如下：

除非能够有效解决将为他人带来伤害的问题，否则个人的自主性永远要妥协。

合租室友抽烟，你还要和他一起住吗

我们可以用一个简单的例子来加以说明，并且在举例的过程中，我们可以看到有效解决问题的关键在哪里。假设汤姆和山姆正在考虑是否合租一个每月房租为3 000美元的、带有两间卧室的公寓。或者，他们可以各自花2 000美元，各租一个单间公寓。先不论房租，除了汤姆喜欢吸烟而山姆不喜欢闻到烟味这个问题，他们对其他方面没有任何意见。如果他们分开住，吸烟也不再是什么问题。那么问题在于，合租所节省的房租是否足以补偿吸烟所带来的伤害。

在回答这个问题之前，我们必须了解汤姆吸烟的意愿有多强烈，以及山姆对于在无烟环境居住的意愿有多强烈。我在第7章中曾说过，衡量意愿强烈程度最有效的指标在于他们愿意为此支付多少钱。假设汤姆为解除不允许在家吸烟的限制，愿意每个月支付800美元，而山姆觉得在无烟环境居住的价值为每个月1 600美元。为方便起见，我仅从票面价值进行估算，不考虑汤姆或许会因为其他某种激励因素而放弃吸烟。

他们所面临的是一个典型的“成本-收益”问题。两人合租两居室的公寓要花费3 000美元的房租，比两人各租一间单间公寓的总价低1 000

美元。因此，两人合租的收益是每月1 000美元。那么成本是多少呢？对汤姆来说，能够在家吸烟的价值是每月800美元，但他对汤姆造成的损失是每月1 600美元，那么最有效的解决方案就是汤姆放弃在家吸烟。他做出这个改变的成本是每月800美元，也就是他为了吸烟愿意支付的价钱。在这种情况下，两人合租的成本是每月800美元。

如果他们一起合租，并且达成协议汤姆不能在家吸烟，由于合租的收益是每月1 000每月，而成本仅为800美元，那么他们每月就可以分享200美元的额外收益。如果他们愿意平分这200美元，他们在分摊房租后，每个月能比单独租房节省100美元。

除了吸烟问题外，山姆不介意是合租还是单独租房。如果合租，他需要分摊1 900美元的房租（这比他自己租房，也就是在不受吸烟问题困扰的情况下，少花100美元），那么他每个月可以节省100美元。对于汤姆来说，如果他承担的房租是1 100美元，那么合租能够为他节省100美元。这比他单独租一间单间公寓少花900美元，但其中的800美元要作为放弃在家吸烟的补偿。

我在第7章中曾讨论过，定价通常取决于各方的收入。比方说，山姆愿意为无烟环境付出的价钱高于汤姆愿意为继续吸烟支付的价钱，因为山姆的收入比汤姆高。左派批评家常常因为定价取决于人们的支付能力而反对成本-收益测试。他们质疑道，为何富人就可以占优势？我在第7章也曾讨论过，这是因为，如果根据支付意愿来衡量成本和收益，参与各方都能获得他们想要的结果。使用其他定价方法则有可能妨碍他们推动事情往前发展。

举例来说，假设在汤姆和山姆的案例中，定价不是根据两人的支付意愿来决定的。如果双方必须平摊房租，首先汤姆永远不会同意合租。每月支付1 500美元房租，但是不能吸烟，那他肯定愿意花2 000美元自己住。

如果按照之前描述的协议进行合租，他们俩都能获得净收益。如果

要说这个协议侵犯了某方的权利，这是不成立的。毕竟，每个人都有独自居住的权利，如果他们愿意的话，完全可以不达成协议。这样一来，汤姆就没有理由抱怨他失去了在家吸烟的权利。达成协议后他确实不能再在家吸烟，但是他获得的补偿高于他的要求，也就是每个月可以节省900美元房租。山姆同样没有理由抱怨他付的房租份额更高。汤姆须放弃他很看重的习惯，来达成双方可以接受的合租协议，如果山姆所需支付的房租份额没有很大差距，那汤姆可以不同意合租。

相对于他们所面临的其他选择，上述协议（妥协以及其他方面）使两人的自主权达到最大。自主权是指一个人可以做自己想做的事情的权利。合租协议没有要求山姆放弃他所看重的事情，却让他的资产增加了额外的100美元，从而使得他能做的事情更多了。在汤姆看来，同意放弃在家吸烟是一种极大的牺牲。但他获得了他更看重的东西作为回报，那就是每个月拥有额外900美元可以支配。他能做的事情少了重要一项，但他获得了比他所放弃的事情更有价值的选项。因此，他同样享有了更大的自主权。

通过这个案例，我希望强调的是，为伤害他人的行为找到有效的解决方案符合当事各方的利益。无论问题具体是什么都可以这么做：可以是吸烟问题，可以是噪声问题，也可以是为了进入最好的学校而比拼学区房的问题。无论当事人是谁也都适用：可以是共和党人士或者民主党人士，天主教徒或基督教徒，从高中辍学的人或大学毕业生，理性的自由主义者或茶党成员。

上述案例强调了科斯的“外部性是相互的”的观点的重要性。汤姆希望能够在家吸烟，但他并不是一个作恶者。山姆希望得到一个无烟的居住环境，他也不是一个受害者。他们只不过是共同面临一个事关双方利益的问题，须用成本最低的方式来解决它。一旦找到有效的解决方案，他们就能拥有新的选择权，可以选择一些比之前所拥有的更喜欢的东西。也就是说，每个人都能拥有更大的自主权。解决问题的有效性是最

大自主权的前提。

解决问题的有效性是最大自主权的前提

在此基础之上，让我们再来探讨几个理性自由主义者或许愿意接受的具体案例。聪明的读者或许已经注意到，合租的案例与第8章中的组建社会的案例十分相似。我们再做一次之前做过的练习，想象一场洪灾摧毁了社会上的一切事物，你和其他999人登上方舟得以获救。你的任务是组建社会，重建社会文明。

本次练习的目的依然是为那些强烈关心个人自主性的人实现他们利益的最大化。最重要的是，进入一个新社会必须是严格自愿的。没有人可以强迫你进入你不想进入的社会，并且只有在获得每一位社会成员的认可后，社会规则才能生效。

假设你是一位理性的自由主义者，而其他999人代表着从各种能力和品性的人中随机抽取的人员。就对自主权的渴望程度而言，你跟其他一小部分与你类似的人处于整个数值分布的极值尾部。从性格来说，你是一位极端的自由主义者，但你并不是一位空想家。如果你会选择加入一个对你的自由行为有所限制的社会，原因只有一个，那就是接受这些限制符合你的利益。

同之前的设定一样，每个社会都拥有在洪灾中存留下来的一部分土地和其他一些财产。假设一个100人的社会占世界总人口数的10%，他们也相应地拥有10%的土地和其他财产。如何分配这些财产成为一个需要成员们自己解决的问题。他们可以让每个人分到同等份额，也可以先拿出一小部分进行分配，然后允许成员们用已经分到的钱购买其余部分。或者，他们还可以组建一个财产永久共有的社会。所有的方案都可供选择。

假设你可以完全了解他人的才能和品行，那么你是知道自己选择加

入的社会是什么样的。一旦你和其他人达成协议，组建一个社会，在你们制定的规则下，你可以凭借自己的努力和才能赚钱。

我在第8章中曾讨论过，能否取得重要成果取决于你所选择加入的人群的特点。如果他们之中大部分人生产力很高，你所在的社会就能购买更多更好的公共设施。但这也是缺点，比方说会对你竞标房产产生不利影响。假如只有一部分房子能看到美景，而大部分人都想要那样的房子，如果你属于生产力最低的那部分人，那么你得到这样房子的可能性很低。

在第8章中我曾提出，诸如此类的考虑是促使人们支持累进税的原因之一。如果一位理性的自由主义者希望社会中其他人生产力不如自己，使得他能在竞标房产时占据优势，其他人的反应则是要求通过税务机制获得补偿。那么，自由主义者就得判断，这样的社会对他来说是否具有吸引力。

鉴于一些集体主义经济社会的惨淡历史，参与者或许都会同意将大部分的商品和服务移交到自由市场。但在第9章我们可以看到，竞争性市场的收入前景具有天然的不确定性。在天赋或能力方面的微小差异，以及其他看似不重要的随机事件，往往会导致差距极大的市场回报。在仔细考虑这一事实后，哪怕是理性的自由主义者，或许也会希望通过某种社会保障的形式，来补充市场收入。一旦成员们就社会的机构设置和管理规则达成统一意见，他们就能按照自己的意愿自由支配收入。

顺便提一下，至少从某个重要角度来说，这个练习有别于道德哲学家约翰·罗尔斯（John Rawls）提出的一个著名的思想实验。¹罗尔斯让读者想象自己处于无知之幕后边，人们无法认清自己的才能和品性。他认为，在无知之幕的掩盖下，我们可以预测分配规则将会是公平的，因为人们不知道哪些规则对他们是有利的。罗尔斯认为，在这种背景下制定出来的规则，如果只是为了帮助社会中最穷困的人群增加收入，那么收入不平等的情况将会越来越严重。尽管也有人提出，位于无知之幕

后面的大部分人制定出来的制度会比这更不公平，但人们普遍认为，如果需要采取更激烈的措施来限制不公平，罗尔斯的思想实验可以提供强有力的理论基础。

或许也有人会说，我所提出的练习是不公平的，因为人们事先已经知道了自己的才能情况，他们有能力排斥那么运气不佳、天生才能堪忧的人。这种反对意见非常有力。与罗尔斯实验不同的是，我的练习目标不在于呈现一个公平社会的面貌，而在于指出理性的自由主义者会选择加入什么样的社会。我从未宣称社会应该被塑造成理性的自由主义者所愿意选择的社会。我所主张的是那种社会会与现代福利国家高度相似。

由安·兰德（Ayn Rand）创作的小说《阿特拉斯耸耸肩》（*Atlas Shrugged*）里的主人公约翰·高尔特（John Galt）又会选择加入什么样的社会呢？兰德曾说，她写这部小说是为了“展示这个世界是多么迫切地需要发起者，而社会对待发起者的方式又是多么的恶劣”。²但我所提出的实验是将自由主义发起者置于有利的地位。在那里，高尔特没有理由担心遭受恶劣的待遇。尽管这个实验做出了有利于高尔特的设计，但他被赋予的能力在本质上是防御性的。他可以拒绝别人强加给他的令人不悦的条款，但他也不能强加条款给别人。

妥协的真正原因

在最开始的时候，理性的自由主义者将会放弃“一切税收皆为窃取”的主张。除了国防领域，还有许多重要物品是无法由私营部门生产出来的，或者说生产质量不尽人意。比方说，在某些情况下，私营收费公路是经济可行的，但让私营部门在密集交错的城市网格设立收费亭是不现实的。政府通过税收来建设和维护道路的效率更高。理性的自由主义者还会接受一个事实，那就是征税必须是强制性的。自愿性的税收远远不足以支付人们所需的公共设施和服务。

在我提出的练习中，理性的自由主义者能够做出妥协，最主要的原因在于人多所固有的优势。正如本书第8章所说，许多公共设施拥有一个共有特性，那就是它们在面向所有民众时必须是等质等量的。越多人来分摊成本，每个人的付出也就越少。即便他人青睐的公共设施和服务不同于自己的喜好，理性的自由主义者也许也会选择加入他们。这样做有一个很明显的缺点，那就是他不得不通过交税来为某些他不喜欢的公共设施买单。但也会有更多人来跟他一起分摊他所喜欢的公共设施的成本，总的来说他的赋税压力会减轻。

高尔特及其志同道合的伙伴或许梦想着拥有一个完全由他们这样的发起者所构建的社会，因为这样的话他们就无须妥协。然而，在经过短暂的思考后，他们肯定会改变主意。在那样一个社会，谁来为他们提供洗衣服务呢？他们的房子着火时谁来救火？谁来收垃圾？如果将很多时间花费在这样的事情上，他们其实是在严重牺牲自己的自主性。

亚当·斯密最卓越的洞见之一是：如果人们能够分解任务并实行专业化分工，生产力将获得爆发式的增长。比方说，他曾举过一个广为引用的例子，18世纪苏格兰大头针厂在实行专业化分工后，生产力增长了数百倍：

一个人抽铁线，一个人拉直，一个人切截，一个人削尖线的一端，一个人磨另一端，以便装上圆头。要做圆头，就需要有两三种不同的操作。我见过一家这样的雇有10个工人的小工厂，如果工人们竭尽全力地工作，他们一天大约能生产12磅（约5.44千克）大头针。如果是中等大小的大头针，1磅（约0.45千克）最多有4 000枚。那么这10个人，每天最多能生产48 000枚。按照每个人占1/10计算，每个人每天最多能生产4 800枚。但如果他们不是各自负责一个部分，假如他们没有受过类似的训练，那么他们每个人一天的生产量肯定不足20枚，甚至有可能1枚都生产不出来。³

亚当·斯密还意识到，只有在人口密度足够的情况下，这样的专业化分工才能完全发挥出优势。因此，他写道：

由于交换的力量为劳动分工带来了机会，因此，分工的程度必将取决于交换力量的强弱，或者换种方式来说，取决于市场化的程度。在苏格兰高地这样荒芜的地方，散落着一些孤零零的房屋和小村庄，每位农民还必须兼任自己家的屠夫、烘焙师和酿酒师。在此情境下，我们很难在30公里以内的地方找到干铁匠、木匠或泥瓦匠等行当的人。住户之间彼此散落，很难在15公里内的地方找到另一户人家，他们必须学会做许多小事情，而在人口更密集的地方，这些事可以由工人的助手来做。小乡村里的木匠不仅仅是木匠，还得是细工木匠、家具木匠，甚至是木材雕刻师，同时还是车匠、制犁匠、二轮货运马车匠和四轮货运马车匠。⁴

假如让思想实验中1 000人里的一小部分自由主义者组建一个新的社会。即便实验人数大大增加，如果约翰·高尔特不愿意加入跟他意见相左的群体，那么他也只能加入一个非常小的社会。除非我们愿意放弃亚当·斯密的主张，即分工和专业化是社会繁荣的基础，那么我们只能得出这样的结论：一个完全由坚定的自由主义者组建的社会也是非常贫穷的，至少相对于其他人数明显更多的社会来说是如此。

约翰·高尔特会同意强制储蓄计划吗

在牢记这个事实的基础上，让我们假设高尔特需要决定是否加入这样一个群体：这个群体跟他所有意见都一致，但有一个问题除外，那就是是否要实行强制储蓄项目。这个群体希望将所有税种的税率提高10%，然后将这些额外的税收存入老百姓的账户里，退休后才能使用。他们支持这个项目，不是因为他们自认目光短浅或不理性，也不是因为

他们想为自由主义者的生活制造困难。他们希望通过这个项目来减轻在学区择房上的竞争。

我再次声明，好学校是一个相对的概念。它只不过是在相同环境中相对更好。大多数父母都希望孩子能进好学校，但无论何时，只有50%的学生能被录取进入排名前50%的学校，无论父母在好学区花了多少钱买房都无济于事。

结果，年轻的父母们就会陷入一个痛苦的两难境地。他们要么在年轻的时候存下足够多的钱，以便在退休后维持舒适的生活水准；要么用大部分储蓄到更好的学区买房。正如之前讨论的那样，问题在于假如他们花光了所有的储蓄，他们也只不过是成功抬高了进入好学校的“门票”价格。这个群体通过将他们现有的部分收入存进储蓄账户里，直到退休才能使用，由此来缓解这个问题。

我们假设最开始的时候，高尔特和他的自由主义者同伴们觉得被这个项目的提议冒犯了。尽管科斯说过外部性会起相互作用，他们会问：“政府究竟有何权利强制我们存钱？”然而，其他人不会被反对意见所动摇，他们宣布无论自由主义者是否加入，他们都愿意执行这个计划。我相信，高尔特和他的朋友们找不到另外一个人数量更多、跟他们意见相同的群体。那么，他们要么选择自己组建一个小型社会，保留进行储蓄的自主决策权；要么加入一个必须每月储蓄的更大型的社会。

如果他们加入了更大型的社会，他们就能够与更多人分摊公共设施的成本。此外，强制储蓄税与其他税种不同，因为这部分钱存在他们自己名下的账户里。到退休的时候，他们就能拿回所有的钱以及累积的复利。由于劳动分工和专业化程度更高，他们的税前收入能明显提高。无论如何，在更大型的社会里，他们的税后收入也会更高。如果加入大型社会后，在除去储蓄税和其他税费后，他们每周的收入还更高，但他们依然不这么做的话，怎么可以自称是“理性的”呢？

为了保护他们的自主权？那样的话，自主权的概念就会变得很怪

异。如果他们在更大型社会能够获得更多的可支配收入（且不论他们在退休时还能取出强制储蓄的那部分钱），也没有受到更多的限制，那么他们在更大型社会获得的自主权会多得多。

安全法则

如果大型群体希望工作安全得到监管，同样的逻辑也适用。高尔特和他的朋友们或许会说，这样的管制将会剥夺他们决定是否追求更高收入而承担额外风险的权利。是的，但并不是因为大型群体自以为是、好管闲事才希望实行这样的管制，而是因为个体与群体在风险偏好上存在根本的达尔文式冲突。在第3章中我曾写过，风险更高的工作回报更高，这是因为安全措施是需要成本的。因此，如果有人选择了风险更高的工作，他们能够用更多的钱在更好的学区买房，对于个人来说，这完全讲得通。但如果所有人都这么做，他们所期望的结果就会化为泡影，人们花光自己的储蓄，结果只会抬高好学区的房价。

高尔特和他的同伴或许仍存在质疑。然而，事实上希望实行监管的人数多过不希望监管的人数。由于社会越大，能够实现的劳动分工和专业化程度就越高，能够分摊的公共设施成本就越多，所以我们不难想象，大型社会里安全系数更高的工作的税后工资会比小型社会里风险不受监管的工作的工资更高。

高尔特及其同伴仍然拒绝加入，他们声称自主权的重要性高于一切。这样一来，他们就暴露了自己的不理性。视自主权高于一切的人却青睐于自主权更少的选择，这样的人不可能是理性的。在大型社会，低风险的工作让他们因为严重事故而失去自主权的可能性更低一些。更高的税后收入能够从无数方面增加他们选择的可能性。他们可以更自由地选择去何处度假，去哪里用餐，也可以选择缩短每周的工作时间，或者提前退休，等等。

在哲学和经济学领域，对于“理性”一词的定义不尽相同。但所有的定义都规定，理性者将会以最有效的方式实现自己的目标。如果高尔特及其同伴由于不同意大多数人在储蓄或风险方面的立场，而选择了更小的自主权，我们可以说，他们既不理性，也不像他们自己声称的那般重视自主权。

什么是权利

“人民拥有一系列权利”是《无政府、国家和乌托邦》（*Anarchy, State, and Utopia*）开篇的第一句话，这本著作是哲学家罗伯特·诺奇克（Robert Nozick）反对自由主义者立场的标志性作品。⁵许多自由主义者会因为以下理由对我的实验提出异议：只有多数人对少数人拥有仗势欺人的能力。诚然，多数者通常会拥有这种能力。为防止他们滥用这种能力，所有成功的文明社会都会制定宪法，用以保障许多具体的个人权利。比方说，没有一个国家会赋予他人杀害你、窃取你的汽车、阻止你选举的权利。许多社会都赋予广泛的公共演讲自由，哪怕演讲者是在极力反对多数者。但是，实现多数者的意愿，并不总是，或者说并不通常，会侵犯别人的权利。

在任何国家，监管储蓄状况都不是违宪的，也没有任何国家会声称对安全进行监管是违反宪法规定的。这些观察结果里或许藏着某些真相。如果自由主义者坚持认为有人可以不受监管，他们必须解释这种权利从何而来。

无论如何解决这些具有争议性的问题，总有人会成为弱势群体。比方说，在关于储蓄的案例里，拒绝接受监管意味着许多人在退休时储蓄资金不足。但如果进行监管，有些人不得不存下比自己计划的更多的钱。在安全监管的案例中，拒绝接受监管意味着有些人被迫承担更多的风险。但如果进行监管，有些人不得不更多地为安全买单。如果自由主义者坚持对于后一类人的关注要优先于前一类人，他们必须对此做出解

释。

决定是否加入一个有人与你在某些问题上意见相左的社会，与决定是否跟吸烟的人一起合租，是明显不同的。但就目前而言，这两种决定在两个重要的方面是相同的。首先，经济的重要性所占的比重相同。无论是共享公寓还是共享社会，与他人共享都意味着必须在某些方面做出妥协，但最终节省的成本可以让人觉得妥协是值得的。

其次，在两个案例中，对意愿进行估价为找到最有效的限制伤害的办法提供了通用标准。我在第7章曾说过，估价通常取决于各方的收入。在考虑他们收入的基础上，应该通过估价进行权重分配。正如之前说的那样，使用其他估价原则或许会阻碍当事各方继续进行交易。

有些自由主义者可能会反对说，科斯的框架强调成本-收益分析，这不一定适用于与人权相关的所有事务。这也许是对的。但我们并没有在谈论与人权相关的所有事务。我们只不过是谈论日常公共政策的决策，这些政策通常只会对每个人一生的总财富产生极小的影响。

人生漫漫，通常在人的一生中，每个人都会受到数万项公共政策的影响。只选择那些收益大于成本的措施，并不能保证每个人都能从每项政策中获得净收益。每个变化都会产生赢家和输家。但如果以成本-收益分析作为决策规则，长期来看实际上每个人都会成为赢家。这就好比进行1 000次赢面更大的赌博。假如抛硬币出现正面你能赢4美元，出现背面输3美元。你可能连续输好几次。但如果硬币本身没问题，抛几千次之后，最终你绝对能成为赢家。

使用成本-收益分析作为社会决策的规则，会使某些人受到不公正的待遇吗？由于成本和收益是由支付意愿来衡量的，更高收入群体的意愿能够占更大比重，这是因为高收入者愿意为自己想要的东西支付更高的价钱。如果他们想要的与他人想要的存在系统性的不同，在许多领域都是如此，这就会成为一个问题。

正如我在第7章中所讨论的，假设低收入者有能力拒绝成本-收益分析，比方说在实验中可以拒绝加入采用这种方法的社会，但其实他们有更好的选择，那就是接受这种方法，从而获得更多可支配的收入。并且，在更大、更多元化的社会里，劳动分工和专业化的程度也更高，所以，做出这个让步对他人也是有利的。但一旦达成了协议，支付意愿将成为衡量不同选择的相关标准。

最坏的情况是，如果社会规则不能保证尽可能地把总的“经济蛋糕”做大，他们就白白浪费了扩大每一位公民自主权的机会。如果“经济蛋糕”做得更大，每个人都有机会分到更大的一块，也就意味着他们可以选择做更多的事情。

这样的话，自由主义者就有充足的理由来建立一个更大的同盟，大于他们将异见者排除在外时能形成的小同盟。这样做的优势在于人均收入将会大大高于之前仅由自由主义者组成的小同盟。此时，他们不得不遵守一些规则，从而放弃一些选择权，或者以更大代价获取选择权，这在之前讨论的关于储蓄和安全监管的案例中可以看出。但他们也会有许多新的选择。如果这些新的选择比他们牺牲的选择更有价值，正如在储蓄和安全监管案例中那样，几乎可以肯定，只有不理性的自由主义者才会拒绝加入。

人们当然不能随心所欲地建造新的社会，无论这个新社会是多么适合他们。实际上，基本上所有人都出生在一个既定的社会里，要想建造一个新的社会几乎是不可能的。因此，生活就像科斯框架呈现的第二种形式，即在生活中，交易成本的存在，使得人们不可能为外在因素所造成的问题协商出有效的私下解决方案。科斯建议道，社会应该对产权进行界定，采取一个可以推动当事各方找到他们所认可的解决方案的机制。

科斯框架着重强调了效率的重要性。比方说，如果一个群体希望监管工作场地的安全，但另外一个群体不希望如此，通过比较双方为自己

立场愿意付出的代价，就可以找到解决问题的方案。

安全监管对于收入水平较低的工人来说更为重要。像约翰·高尔特那样的自由主义者，他们的工作环境通常比需要安全监管那部分人的工作环境更为安全。因此，理性的自由主义者几乎不愿意付出代价来阻止安全监管。由于希望实行安全监管的工人在人数上多过持反对意见的自由主义者，那么，成本-收益分析将会支持实行安全监管。

自由主义者或许还会抗议，这侵犯了他们自主决定为安全付出多少成本的权利。但自由主义者却是在否认他人拥有决定自己最高承受多少风险的权利。为何第一种权利比第二种权利更重要？自由主义者须对此做出解释。

如果理性的自由主义者最终选择加入一个支持实施安全监管、交易成本为零的大型社会，他们有什么理由声称安全监管剥夺了他们的重要权利呢？在我们所生活的世界里，高额的交易成本意味着起码有一方无法获得他们想要的东西。自由主义者该如何解释，为什么大型社会的意愿就要打折扣？一个自称将自由看作高于一切的群体，会如何解释，在人人都享有完全自由选择权的环境中，有些人却得不到他们想要的东西？自由主义者是否拥有比科斯框架听起来更加中听的权利法则？那是什么样的法则呢？是宗教教条吗？请问是哪一条呢？

我们需要新一代的自由主义者

总而言之，代表传统自由主义者立场的理性自由主义者的观点很难站得住脚。我已经预设了传统自由主义者所需的所有条件，也就是市场是完全的竞争市场，消费者非常理性，以及政府不会限制除伤害他人之外的行为。在这一系列条件中，我只增加了一个重要的条件，一个不存在争论的观察结果：生活中许多重要方面是取决于相对标准的。也就是这个观察结果导致了传统自由主义者立场的崩塌。

报酬与相对业绩之间的关系，是亚当·斯密和达尔文对于竞争性过程的想法存在分歧的具体体现，更确切地说，是亚当·斯密的现代门徒与达尔文在观点上的分歧。前者相信，竞争的力量会引导自利的个体为全体社会创造最大的价值。达尔文的观点则相反，他认为竞争模式会导致个人追求自身的利益。在第2章中我们可以看到，这样的行为有时候会使全体社会受益，但并不是一直如此。尤其是当个体的薪酬取决于相对地位时，个体和集体的利益通常就会产生分歧。

如果是职位竞争而非雇主的压榨导致工人承担更多风险，自由主义者就没有理由说安全监管剥夺了他们自主选择的权利。聪明的自由主义者不会因为军备控制协议规定签署国不能随意制造武器，而去反对这些协议。这种限制恰恰是协议的意义所在。签署协议的人也知道，如果可以自由选择，他们将会无休止地制造炸弹。安全监管就好比军备控制协议，所以它自然会限制个人的选择权。但达成协议的各方的目标就是要实现这种限制！

在回报取决于相对业绩的领域，我们不可避免地会看到为了地位而开展的“军备竞赛”。并且，几乎不出所料，竞争者会通过“军备控制协议”来规定一些限制；因为如果不这么做，就会产生互相抵消的开支。赛车协会会限制发动机排放量，运动联盟会限制参赛人数，学区会限制幼儿园入园日期，等等。如果想从根本上反对这些限制，从而解除他们对于个人自由的限制，办法只有一个：完全不参与这些事务。

部分自由主义者坚持认为，相对地位对于大多数人来说并不重要。但即便是口才最差、知识最浅薄的辩论者也能毫不费力地驳倒这个观点。自由主义者最终必须承认相对地位是重要的，但他们中的许多人或许依然想说事实不应该是这样。他们会反对说，以地位因素决定公共政策，无异于将羡慕、嫉妒这样的负面情绪合法化。他们会说，我们不会考虑虐待狂和受虐狂在公共政策制定上的偏好，我们还应该忽视源自嫉妒和羡慕的偏好。

然而，这个论点同样站不住脚，因为关于地位的因素与此类情绪无关。根据环境做出判断是人类神经系统不可避免的属性。一辆车的速度只有比同一环境中的其他大多数车快，才能称之为快。如果一处房屋明显比同一环境中的其他大部分房子小，不管它的实际面积有多大，它看起来都是小的。正如我们在第5章看到的那样，当环境对于评测结果的影响力高于其他因素时，消费模式就会扭曲。人们在房屋、珠宝、汽车等受环境影响的物品上花费太多，在受环境影响较小的事物上的花费就会过少，比如安全、储蓄和休闲。在一个完全没有嫉妒或羡慕的世界里，此类消费扭曲也会发生。

自由主义者或许还会继续坚持声称累进税侵犯了他们的权利。但正如我们在第8章看到的那样，任何其他形式的征税都会剥夺社会上最富有成员选择他们最喜欢的公共和私人设施的权利。自由主义者或许还会坚持认为，向最富有的社会成员征收相对更高的税，会使“经济蛋糕”变小。正如我们在第10章看到的那样，几乎没有证据可以支持这种论点。相反，我们在第11章可以看到，通过抑制利大于弊的行为，一个设计完善的征税系统实际上可以使“经济蛋糕”变大。如果我们缩减一小部分在婚礼和成年礼上的花费，仪式的主角们也会觉得这些仪式很特别。如果我们减少数吨二氧化碳的排放量，我们也能降低灾难性全球气候变化发生的可能性。

正如穆勒所说的那样，政府可以通过立法限制个人的行为，从而阻止对他人的伤害。但最后发现，严厉的管制几乎从来都不是最有效的。污染税的经验已经证明，增加有害行为的成本通常比直接禁止的作用更好。社会的利益在于减少有害行为的总量，而不是减少具体某个人的有害行为。尽可能少用监管而采用征税手段，可以将总成本控制在最低程度，因为这样可以将减少伤害的任务集中分配给少数几位完成成本最低的参与方。

几十年来，反政府和反税的口号妨碍了政府更好地运用这些简单的

道理。正如我们在第4章看到的那样，有些国家拥有励志的、诚实的政府，它们能够不遗余力地提供公共设施和服务。每一位理性的人都希望生活在这样政府的羽翼之下。但好政府不是凭空而来的，它必须经过细心的培育。“饿死这些野兽”这样的说辞断然不会帮我们锻造出我们所梦想拥有的社会。

我们需要优秀的政府，因为个人和社会的目标通常是冲突对立的。当冲突出现时，期待“无形的手”来解决问题的想法未免会显得太天真。有一个好消息，那就是明智的征税政策通常可以指导我们在不干扰他人的情况下获得好的结果。只需通过对伤害他人的行为进行征税，我们就能轻易地获得维护道路和桥梁的费用，而不用痛苦地牺牲私人消费。

尽管传统自由主义者的立场经不起严格的审查，我也从未幻想可以通过理论或证据，能使自由主义信仰坚定的追随者改变想法。但在写这本书的时候，我希望它至少能说服部分自由主义者重新思考他们对于集体行动不假思索的反对。在公共辩论中，最热切关注个人自主权的人其实占据着非常重要的地位。几乎没有人希望看到一个对任何一项恼人行为都要进行监管的政府。真正有效的政府只会专注于处理那些无法由自己解决的、真正带来伤害的行为，会以干扰性最低的方式来阻止这些行为。

拥有一个这样的政府，就相当于形成了一个成功的自由主义福利国家。为实现这个目标，美国需要新一代的自由主义者，他们愿意接受针对自己行为的立法限制，同时继续对政府进行的不必要的干扰进行激烈的斗争。

每一个时代都有末日预言者。之前几个时代的预言者大多都是宗教狂热者，而当今时代的预言者则是最优秀的科学家。他们告诉我们，如果我们不采取强有力的行动，快速抑制全球变暖，我们很有可能会化为灰烬。能够实现这个目标的政策工具不仅简单，而且易于理解，那就是实行坚定的碳排放税政策。这样一来，我们所需承担的成本就可以承

受。但似乎我们还没有采取任何行动。

阻止我们采取行动的是反征税、反政府的狂热分子，驱动他们的那些理论一旦经受审视便不攻自破。当前，他们在争议中占了上风，但他们的理由并不可信。无论成败与否，我们必须同他们进行斗争。读完全篇，你还能无所事事地袖手旁观吗？

愤世嫉俗的朋友提醒道，我们不过是天真地以为自己能够做得更好。他们说，我不过是在同假想敌作战。或许吧。但正如塞万提斯在《堂吉珂德》中留下的警示所说的那样：“过于清醒或许就是疯狂，而最疯狂的莫过于接受现实，而不去想生活本应怎样。”⁶

后记

在2011年4月，临近本书精装本交稿期限之际，美国经济依然低迷。包括我在内的诸多经济学家坚持认为联邦政府在2009年实施的刺激计划对于解决当时的困境来说简直是杯水车薪，分三年实施总额仅为7870亿美元的刺激方案怎能应对每年2万亿美元的需求缺口？即便如此，经济还是呈现出了缓慢复苏的迹象。

因此在描述美国政治体制对经济衰退响应的不足时，我所使用的时态基于这样的假定：进一步的经济恶化在本书出版时早已发生。例如，我在书中提到，如果政府能像所有人认为应该做的那样，加速推进基础设施的维修工作，经济就会复苏得更快，我这样写道：“在经济衰退的低谷，生产设备空置，工人们无所事事。伴随着世界范围内市场的普遍萧条，生产材料的价格异常低廉。项目融资的利率也会接近历史低点。”

当然了，直到2011年9月精装本发行之时，这场衰退还远没有结束。本书的英文平装本在2012年9月问世，经济恶化仍在“固执”地继续。日渐破败的基础设施在此前没有得到修护，而本可以做这些修护工作的工人依旧无所事事，利率依旧位于历史低点，世界市场上的材料价格也依旧低廉。

经济复苏停滞的部分原因在于，联邦政府最初实施的刺激计划由于州政府和地方政府支出的大幅缩减而被削弱，这些缩减仍在继续。尽管财政紧缩将英国经济推入二度衰退，紧缩政策的拥护者依然反对进一步的经济刺激。

在美国和欧洲，这些紧缩政策的拥护者坚持认为，一个国家不能无限地透支财富，也就是说，就像一个家庭或者一个企业一样，政府最终必须保持收支平衡。如果我们在此讨论的是长期经济发展，这么说是有

道理的。但是这对于短期发展而言却是大错特错。

紧缩政策支持者的理论缺陷显而易见。试想，一个已有负债的家庭必须决定是否借5 000美元在阁楼里添加隔热材料。假定这个方案平均每月会使家庭水电费用减少100美元，并且每月的还贷金额为50美元。实施这一方案在短期必然会增加负债，但是，这一方案无论在短期还是长期都会使这个家庭的经济境况得到改善。即便需要还贷，他们还是能够获得50美元的额外收益。并且，一旦还清借款，他们每月获得的额外收益就成了100美元。对于这样的方案还有什么好争辩的呢？

同样的逻辑完全适用于经济衰退期间的基础设施投资。没错，这些投资会增加政府债务。紧缩政策的拥护者宣称这样的计划会“压榨我们的后代”。然而他们也承认，这些投资不能无限推迟，而且推迟得越久投资成本也越高。因此，正是因为政府现在不实施这些计划，才让我们的后代平白背负更多的债务。

2012年的经济衰退则会从其他方面增加未来的负担。例如，刚刚毕业的大学生被迫在继“大萧条”以后最为艰难的劳动力市场中谋生，如此滞后的发展会减少他们中的很多人一辈子的收入。由于企业投资没有达到正常水平（如果产出已然超出消费者需求，何必还要再建新的工厂呢？），一个国家的资本存量会始终处于较低水平，这意味着生产力和工资的增长也较为缓慢。普遍的失业和滞后的工资在过去曾导致离婚率升高，营养不良的儿童人数增多。无论在哪种情况下，未来的税收都会减少，继而进一步增加美国政府的财政赤字。

为了给后代营造更光明的经济前景，最重要的便是尽快结束这场经济衰退。为了实现这一目标，最可靠的办法之一就是即刻实施大规模的基础设施整修计划。然而，这在强硬的反政府论断引导政治舆论的美国连想都不要想，更遑论实施。

经验表明，相比于我在本书第1章中的描述，2012年的美国政治体制所陷入的困境更加严重。因此，最为紧要的是识别这一现象背后的影

响因素。大量证据证明，其中一个因素便是过去40年来收入差距的急剧扩大。因为在这期间，美国社会几乎所有的财富都积聚在这个国家的最高收入者手中，如今这个群体拥有了越来越大的影响力。伴随着最高法院的公民联合决议，赞助竞选的富人为了自己的目的而影响公共政策的行为所受的限制，较以往已经大为降低。

表面看来，这些富人所想要的，并且已经成功实现的政策，可以进一步增加他们的收入。针对他们的税率已经大幅下降，对他们企业的管制也更为宽松。这些政策的变化反过来进一步扩大了他们的政治影响力，使他们可以收买更多对自己有利的政策，等等。

选民仍在被大肆灌输这样的观点：税收和管制会对经济造成破坏。基于各种证据，这些观点都是自相矛盾的。尽管存在这样的观点，调查显示，克林顿当政时期的多数选民认为，针对富人的税收应该在现有税率范围内调高。但是这些都没有在政治上得到实现，因为任何一个政党，如若做出这一提议，需获得白宫、众议院多数和参议院超过六成赞成票的支持才能得以实现。如今，面对最富有的竞选赞助者的坚决反对，这种程度的选举变动显然无法实现。该问题造成的影响使这一问题日益难以解决。简而言之，我们似乎陷入了一个无法逃离的怪圈。

然而，出路可能还是有的。关键是要更多地关注这样一个事实：富人收买的政策实际上并没有很好地服务于他们自身的利益。诚然，他们设法获取了国家收入的更大份额，但是有力的证据表明，总体而言，受他们所支持的政策的影响，这些富人的境况变差了。这些政策带来的私人收益比他们先前预想的要低得多，而其带来的间接成本则高得多。

根本问题在于，超过一定程度后，私人支出的整体增加几乎不会带来什么福利的改善。例如，富人用减免的税收建造更大的宅邸，但是当豪宅的面积从原先的3 000平方米增加10%，其首要影响不过是提高了他们所在社交圈中住房面积的标准而已。

相反，增加对各种形式的公共物品的支出，能够为每个人带来确定

的收益，这其中也包括富人。例如，相比于布满坑洼的道路，在维护良好的路面上行驶更安全，也更轻松。

当然，为了填补道路的坑洼而提高税收意味着富人在汽车上的花费会减少。但是请你仔细考虑接下来的这个简单的思想实验：假设有这样一个社会情境，所有富人都在维护良好的道路上驾驶价值15万美元的梅德赛斯轿车。在另一个其他方面完全相同的社会情境中，每个富人驾驶的是价值20万美元的宾利轿车，行驶的路面则布满了一两米深的坑洼。在哪种情境下，富人的整体驾驶体验更好呢？

实验中的第二个社会情境与2012年的美国大体类似，但是，如果你像大多数人一样，认为第一个社会情境下的富人满足程度更高，那么随之而来的一个颇具挑战性的问题便是：如果在平坦的路面上驾驶不那么贵的汽车更好，为什么富人们会极力反对为了修护基础设施而提高税收呢？一个合理的解释是，他们不过是对增税的影响存在认知错觉。

税收的增加显然会减少你手中的可支配收入。因此，面对增税的可能性，人们倾向于认为他们未来满足自己需求的能力会下降。然而，一旦收入增加超过一个合理的阈值，人们需要的许多物品就成了经济学家口中的地位物品。这些物品可能本身就是稀缺的，如豪华的滨海别墅；或者其价值极度依赖于情境，如宝石或性能稳定的跑车。由于地位物品供不应求，只有出价最高的人才能买到。人们倾向于忽视这一简单的事实，从而曲解了增税的影响。

之所以会存在这样的认知错觉，是因为人们在生活中经历的多数财务危机通常来源于仅对他们自己造成影响的事件。例如，他们可能遭遇健康或者工作危机，婚姻可能会破裂，珠宝可能被偷窃，洪水可能会毁坏房屋。每种情况造成的影响都降低了他们对地位物品的购买能力。

绝大多数的财务危机都是由千奇百怪的原因导致的，增税带来的收入下降也会造成类似的影响。但是增税是不一样的。它会影响到所有参与竞购地位物品的人。并且，由于增税会使所有人的可支配收入减少，

因此对竞购的结果基本没有什么影响。同一幅油画和同一艘游艇最终还是落入与增税前同样的人手中。

反对税收和管制可以带来赞助竞选的富人们更想要的结果。但是如果这些人的目标是买一套视野更佳的住房，那他们就要失望了。这样的房子总共就那么多，最后还是会被与减税前相同的人买走，因为减税所提高的是所有人的竞拍能力。

总而言之，当反对税收的富人为竞选提供赞助时，他们买到的只不过是更深的坑洼和更污浊的空气，这些是减税和管制松弛的必然后果。

降低最高税率还会对富人造成许多其他损害。比如，由此导致的赤字使联邦政府削减对基础科学研究的资金支持，即使美国获批专利占全球的比重会因此继续下滑，他们也不得不这么做。削减这类资金会危及美国经济长期繁荣的基础。

赤字还会降低一个国家的安全性。譬如，小布什政府将能源部门的项目资金削减了8%，该项目的目的是提高苏联遗留下来的防护松懈的核库存的安全性。美国前参议员萨姆·纳恩意图通过私人筹款努力弥补减税带来的损失，结果并不成功。尽管我们有理由担心恐怖分子可能试图在某个美国城市引爆核炸弹，多数集装箱仍然未经检查就进入了港口。

自利是人类强大的动力。只要富人们相信更低的税率对他们有利，他们就会继续资助站在他们这一边的候选人。这会导致收入变得更加不平衡，恶性循环仍将继续。

道德说教者经常大力劝说富人对穷人更宽容一些，想象这会如何轻而易举地改变这些富人的生活。这固然很好。但是更可取的办法则是使人们注意到这样的事实：从纯粹现实的角度来看，如果不是他们自己要坚持改变政策，那些最富有的家庭经济境况原本会更好。

译者后记

我最初是在北大汪丁丁老师的行为经济学课堂上与本书编辑结缘，我们自然是有着对行为经济学的共同兴趣和关注。未承想本书编辑恰好在湛庐文化工作，而彼时负责出版的策划老师正组织本书的翻译，我便经由本书编辑引荐结识了本书策划，并迅速明确合作意向。我一直对英文翻译抱有兴趣，此番得到机会翻译喜欢的领域的著作，我感到欣喜万分。

可以说，接手本书的翻译是机缘巧合，但也顺理成章。这正如我对本书的理解一般：经济学的理性传统与达尔文进化思想相碰撞，催生出了奇妙而合乎情理的反应。在我看来，行为经济学便是在传统经济学条分缕析的枝丫的缝隙之中布入绵密交织的内容物，似是将经济学理性进行了“柔化”处理，使经济学体系更为完整和自洽，这也是行为经济学在我心中的意义所在。

然而，本书在翻译到第5章时，由于我本职工作较为忙碌，恐不能如期交付全书译稿，遂在中途邀请了两位北大校友，俸绪娴和康溪馨，来协助翻译。绪娴具有出色的翻译工作背景，溪馨虽读法律，却也十分热衷翻译。于是我们三人便通力合作，协同翻译，进展十分顺利。最后由我完成第1章至第5章的译稿，由溪馨完成第6章至第8章的译稿，由绪娴完成第9章至第12章的译稿。

我很感激能通过翻译此书结识诸位益友，期待未来有机会再次分享阅读和翻译的乐趣，也希望我们的努力可以为行为经济学的普及贡献绵薄之力。

此外，受个人学识和经验所限，不免在领会作者意图、遣词造句等方面有所疏漏，还请读者朋友不吝指正。

王堃

01 迷恋往昔：现代经济为什么会增速放缓、遭遇危机

1. John Maynard Keynes, *The General Theory of Employment, Interest, and Money*, Basingstoke, Hampshire, U.K.: Palgrave Macmillan, 2007 (originally published in 1936).
2. Nevada Department of Transportation, FAQs.
3. *State of the Union with Candy Crowley*, interview with Mitch McConnell, CNN, July 18, 2010.
4. A. P. Sokolov, P. H. Stone, C. E. Forest, R. Prinn, M. C. Sarofim, M. Webster, S. Paltsev, C. A. Schlosser, D. Kicklighter, S. Dutkiewicz, J. Reilly, C. Wang, B. Felzer, J. Melillo, and H. D. Jacoby, “Probabilistic Forecast for 21st Century Climate Based on Uncertainties in Emissions (without Policy) and Climate Parameters,” MIT Joint Program on the Science and Policy of Global Change, Report 169, January 2009.
5. Jane Mayer, “Covert Operations: The Billionaire Brothers Who Are Waging a War against Obama,” *New Yorker*, August 30, 2010, pp. 44–55.
6. Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, State College, PA: Penn State University, 2005, p. 364 (originally published in 1776).
7. 同上, p. 111。

8. 想要全面了解崇尚自由的狂热分子如何歪曲亚当·斯密的立场，请参阅阿玛蒂亚·森（Amartya Sen）在纪念亚当·斯密的《道德情操论》出版250周年时写的一篇序言： *The Theory of Moral Sentiments*, New York: Penguin, 2009。
9. Charles Darwin, *The Origin of Species*, 6th London Edition, The Literature Project, 2000–2010. 现代生物学家强调有时选择发生在群体层面，具体参见David Sloan Wilson, *The Natural Selection of Populations and Communities*. Reading, MA: Benjamin Cummings, 1980; David Sloan Wilson, “The New Fable of the Bees,” in *Advances in Austrian Economics*, Vol. 9, Greenwich, CN: JAI Press, 2004, pp. 201–220; David Sloan Wilson and Edward O. Wilson, “Rethinking the Theoretical Foundation of Sociobiology,” *Quarterly Review of Biology* 82, 2007:327–348; 以及A. J. Field, “Why Multilevel Selection Matters,” *Journal of Bioeconomics* 10, 2008: 203–238。就目前而言，重要的一点是，在个人层面具有优势而受到青睐的特质往往对更大的群体不利。
10. Burney J. Le Boeuf, “Male-Male Competition and Reproductive Success in Elephant Seals,” *American Zoologist* 14(1), 1974: 163–176.
11. Thomas C. Schelling, *Micromotives and Macrobehavior*, New York: W. W. Norton, 1978.
12. Richard Thaler and Cass Sunstein, *Nudge*, New Haven, CT: Yale University Press, 2007.
13. John Stuart Mill, *On Liberty*, State College, PA: Penn State University, 1998 (originally published in 1859).
14. 也有一些人对这一原则提出了强有力的反对意见，具体参见Joel Feinberg, *Harm to Others*, New York: Oxford University Press, 1987。

15. The statement of Congressman Dan Burton (R, IN).

02 ∴ 达尔文之楔：竞争无法弥合个体与群体的利益分歧

1. Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, State College, PA: Penn State University, 2005, p. 213 (originally published in 1776).
2. 加尔布雷斯在两本引起广泛讨论的书中提出了这些观点： *The Affluent Society*, Boston: Houghton Mifflin, 1956和*The New Industrial State*, Boston: Houghton Mifflin, 1967。
3. 关于埃德塞尔型汽车的详细历史，具体参见Thomas Bonsall, *Disaster in Dearborn: The Story of the Edsel*, Stanford, CA: General Books, 2002。
4. 这一主张本身是值得肯定的，一定程度上说来，是羚羊之间的相对速度决定了哪些羚羊会被捕获和吃掉。关于这个主张有一个古老的笑话：两位露营者在露营时遇到了一只愤怒的熊，其中一个人醒来看到他的朋友正着急忙慌地穿跑鞋，便不解地问道：“为什么要这么麻烦？难道你不知道你没办法跑过那只熊吗？”“我不必跑过它，”朋友回答，“我只需要跑过你。”
5. Richard H. Thaler and Cass R. Sunstein, *Nudge*, New Haven, CT: Yale University Press, 2007.
6. Peter Richerson and Robert Boyd, *Not by Genes Alone: How Culture Transformed Human Evolution*, Chicago: University of Chicago Press, 2004.
7. Richard Rorty, “The Brain as Hardware, Culture as Software,” *Inquiry*

47(3), 2004: 219–235.

8. 我在1988年写的那本书的主题就是，在激烈的竞争环境中，这种动机是如何演变的，具体参见*Passions within Reason: The Strategic Role of the Emotions*, New York: W. W. Norton。
9. 有关此证据的详细调查，请参阅我1985年出版的书：*Choosing the Right Pond: Human Behavior and the Quest for Status*, New York: Oxford University Press。
10. 关于局部等级、血清素和睾酮之间关系的研究综述，参见Robert H. Frank, *Luxury Fever*, New York: Free Press, 1999, chapter 9。
11. Richard Easterlin, “Will Raising the Incomes of All Increase the Happiness of All?” *Journal of Economic Behavior and Organization* 27, 1995: 35–47. 贝齐·史蒂文森（Betsy Stevenson）和贾斯廷·沃尔弗斯（Justin Wolfers）的研究已经表明，人民的幸福感也与国家内部和一个国家任何时候的平均收入呈正相关：“Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox,” *Brookings Papers on Economic Activity*, Spring 2008: 1–87。
12. Erzo Luttmer, “Neighbors as Negatives: Relative Earnings and Well-Being,” *Quarterly Journal of Economics* 120(3), August 2005: 963–1002。
13. 请参见我1985年出版的书*Choosing the Right Pond*的第5章。
14. Ori Heffetz, “Conspicuous Consumption and the Visibility of Consumption Expenditures,” Princeton University Department of Economics mimeo, 2004; Luis Rayo and Gary Becker, “Evolutionary Efficiency and Happiness,” University of Chicago Department of Economics mimeo, 2004; and Laurie Simon Bagwell and Douglas Bernheim, “Veblen Effects in a Theory of Conspicuous Consumption,”

American Economic Review 86, June 1996: 349–373.

03 : 尚未被利用的机会：市场失灵传统解释的缺陷

1. Greg Hundley, “Why and When Are the Self-Employed More Satisfied with Their Work?” *Industrial Relations* 40(2), April 2001: 293–316; Matthias Benz and Bruno S. Frey, “Being Independent Raises Happiness at Work,” *Swedish Economic Policy Review* 11, 2004: 95–134.
2. 详情参见Andrew J. Oswald, Eugenio Proto, and Daniel Sgroi, “Happiness and Productivity,” Department of Economics Working Paper, University of Warwick, March 2009, 这扩展了早期发现更快乐的工人往往更有成效的研究: T. A. Wright and B. A. Staw, “Affect and Favorable Work Outcomes: Two Longitudinal Tests of the Happy Productive Worker Thesis,” *Journal of Organizational Behavior* 20, 1998: 1–23。
3. Sam Bowles, David Gordon, and David Weisskopf, *Beyond the Wasteland*, Garden City, NY: Anchor Doubleday, 1983, pp. 167, 168.
4. Herbert Gintis, “The Nature of the Labor Exchange and the Theory of Capitalist Production,” *Review of Radical Political Economics* 8, 1976:36–54.
5. Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, State College, PA: Penn State University, 2005, p. 19 (originally published in 1776).
6. Benjamin Klein, Robert G. Crawford, and Armen A. Alchian, “Vertical Integration, Appropriable Rents, and the Competitive Contracting Process,” *Journal of Law and Economics* 21, October 1978: 297–326.
7. 有关沃尔玛公司劳动实践的详细讨论, 请参见Steven Greenhouse, *The*

Big Squeeze, New York: Knopf, 2008; and in Nelson Lichtenstein, *The Retail Revolution*, New York: Metropolitan, 2009。

8. 关于这一点，请参见Bethany Moreton, *To Serve God and Wal-Mart*, Cambridge, MA: Harvard University Press, 2009。
9. 关于这个解释的更详细的描述，请参阅我1985年出版的一本书的第8章：*Choosing the Right Pond: Human Behavior and the Quest for Status*, New York: Oxford University Press。
10. Richard Thaler and Cass Sunstein, *Nudge*, New Haven, CT: Yale University Press, 2007, p. 251.
11. Thomas C. Schelling, *Micromotives and Macrobehavior*, New York: W. W. Norton, 1978.
12. Karl Marx, *Capital*, New York: Modern Library, 1936, pp. 708–709.
13. 具体请参见我在2004年出版的一本书中的一篇文章：*What Price the Moral High Ground?* Princeton, NJ: Princeton University Press。
14. Paul Samuelson, “A Note on the Pure Theory of Consumers’ Behaviour,”*Economica* 5, 1938: 61–71.

04：干预是为了更好的竞争：你无法通过不吃饭饿死身体里的寄生虫

1. Ronald Utt, “The Bridge to Nowhere: A National Embarrassment,”Heritage Foundation, October 20, 2005.
2. Project on Government Oversight, <http://www.pogo.org/pogo-files/reports/national-security/defense-waste-fraud/ns-wds-19990901.html>.

3. “Conservative Advocate,” *Morning Edition*, National Public Radio, May 25, 2001.
4. Peter Schrag, *Paradise Lost: California's Experience, America's Future*, New York: New Press, 1998.
5. 同上, p. 8。
6. B. C. Sheldon and S. Verhulst, “Ecological Immunology: Costly Parasite Defences and Tradeoffs in Evolutionary Ecology,” *Trends in Ecology and Evolution* 11, 1996: 317–321.
7. 这个问题在近几十年来变得尤为严重, 有力的论证请参见Robert Reich, *Supercapitalism*, New York: Vintage, 2007。
8. Adam Liptak, “Justices, 5-4, Reject Corporate Spending Limit,” *New York Times*, January 21, 2010.
9. Lauren Etter, “Economic Crisis Forces Local Governments to Let Asphalt Roads Return to Gravel,” *WSJ.com*, July 17, 2010.
10. “Oakland Road Commissioners Divert Money for ‘Pothole Emergency,’” *PR Newswire*, April 19, 1997.
11. Bruce Van Voorst, “Why America Has So Many Potholes,” *Time*, May 4, 1992, pp. 64–65.
12. John Maynard Keynes, *The General Theory of Employment, Interest, and Money*, Basingstoke, Hampshire, U.K.: Palgrave Macmillan, 2007 (originally published in 1936).
13. Lee Ohanian, “How Stimulating Is Stimulus?” *Forbes.com*, June 17, 2009.
14. Peter A. Diamond, “A Framework for Social Security Analysis,” *Journal of Public Economics* 8(3), December 1977: 275–298.

15. Etter, op. cit.
16. Selim Algar and Andy Geller, “Daddy's \$10M ‘Bad’ Mitzvah,” *New York Post*, October 26, 2007.
17. Mary Ellen Fillo, “MTV's ‘Sweet 16’ show drives fad for lavish birthday parties,” *Hartford Courant*, July 31, 2007.
18. Bethany Kandel, “With This Dress, I Thee Wed and Wed and Wed,” *New York Times*, January 31, 2009.
19. Sarah Wilkins, “For Richer, For Poorer,” *Mother Jones*, January–February 2005.
20. For a more detailed survey of these distributional shifts, see my 2007 book, *Falling Behind: How Rising Inequality Harms the Middle Class*, Berkeley, CA: University of California Press.
21. Robert H. Frank, Adam Seth Levine, and Oege Dijk, “Expenditure Cascades,” Social Science Research Network Working Paper, October 2010.
22. Kelly Evans, “Size of New Homes Starts Shrinking as Builders Battle Housing Slump,” *Wall Street Journal*, September 12, 2007, p. A1.

05：控制“地位消费”：得不偿失的军备竞赛

1. Fred Hirsch, *Social Limits to Growth*, Cambridge, MA: Harvard University Press, 1976.
2. *Falling Behind: How Rising Inequality Harms the Middle Class*, Berkeley, CA: University of California Press, 2007.
3. 然而，一些人已经确定了在某些条件下，地位问题会导致储蓄率升

高，具体请参见H. L. Cole, G. J. Mailath, and A. Postlewaite, “Social Norms, Savings Behavior, and Growth,” *Journal of Political Economy* 100(6), 1992: 1092–1125; and Giacomo Corneo and Olivier Jeanne, “Social Organization, Status, and Savings Behavior,” *Journal of Public Economics* 70(1), 1998: 37–51。

4. *Falling Behind*, op. cit., chapter 7.

5. 具体请参见这本书的第8章: *Choosing the Right Pond: Human Behavior and the Quest for Status*, New York: Oxford University Press, 1985。

6. Sara J. Solnick and David Hemenway, “Is More Always Better? A Survey on Positional Concerns,” *Journal of Economic Behavior and Organization* 37, 1998: 373–383; Sara J. Solnick and David Hemenway, “Are Positional Concerns Stronger in Some Domains Than in Others?” *American Economic Review, Papers and Proceedings* 95, 2005: 147–151; F. Alpizar, F. Carlsson, and O. Johansson-Stenman, “How Much Do We Care about Absolute versus Relative Income and Consumption?” *Journal of Economic Behavior and Organization* 56, 2005: 405–421; Charles Kerwin, Erik Hurst, and Nick Roussanov, “Conspicuous Consumption and Race,” University of Chicago, mimeo, 2007.

7. Robert H. Frank, *Luxury Fever*, New York: Free Press, 1999; and Robert H. Frank, *Falling Behind*. 其他作者也讨论了针对地位外部性的税收补救措施，具体请参见Michael Boskin and Eytan Sheshinski, “Optimal Redistributive Taxation When Individual Welfare Depends on Relative Income,” *Quarterly Journal of Economics* 92(4), 1978: 589–601; Yew Kwang Ng, “Diamonds Are a Government's Best Friend: Burden-Free Taxes on Goods Valued for Their Values,” *American Economic Review* 77(1), 1987: 186–191; Norman Ireland, “On

Limiting the Market for Status Signals,” *Journal of Public Economics* 53, 1994: 91–110; Richard Layard, *Happiness: Lessons from a New Science*, London: Penguin, 2005。

8. 有关细节的讨论请参见Laurence Seidman, *The USA Tax: A Progressive Consumption Tax*, Cambridge, MA: MIT Press, 1997。
9. 贷款偿还将被计入储蓄总额，从而减少人们潜在的纳税义务。新的借款将从储蓄总额中扣除，这反过来又会导致潜在的税收增加。对于业主来说，每年的住房消费量将被算作他们的房屋的隐含租金价值。
10. Robert H. Frank, “The Frame of Reference as a Public Good,” *Economic Journal* 107, November 1997: 1832–1847.
11. Milton Friedman, “The Spendings Tax as a Wartime Fiscal Measure,” *American Economic Review* 33(1), Part 1, March 1943: 50–62.
12. Laurence S. Seidman, *Pouring Liberal Wine into Conservative Bottles: Strategy and Policies*, Lanham, MD: University Press of America, 2006.
13. 比如说，一个美国家庭年消费总额为200万美元，减去储蓄的30万美元，实为170万美元，其中超出50万美元的120万美元被征收附加税。

06：深度理解竞争：完全竞争中的加害者与受害者

1. H. Kaplan and K. Hill, “Food Sharing among Ache Foragers: Tests of Explanatory Hypotheses,” *Current Anthropology* 26, 1985: 223–246.
2. 即便是在大多数议题上都毫无保留支持自由主义者意见的理查德·艾普斯坦（Richard Epstein）也承认，自由主义者所持的一些原则可

以根据实际需要做出让步，见Richard A. Epstein, “Nuisance Law: Corrective Justice and Its Utilitarian Constraints,” *Journal of Legal Studies* 8(1), 1979: 49–102。

3. A. C. Pigou, in *The Economics of Welfare*, 4th ed., London: Macmillan, 1932.
4. Ronald H. Coase, “The Federal Communications Commission,” *Journal of Law and Economics* 2, October 1959: 1–40.
5. 同上， p. 26。
6. 同上， p. 27。
7. Ronald H. Coase, “The Problem of Social Cost,” *Journal of Law and Economics* 3, October 1960: 1–44.
8. A. M. Diamond, “Most Cited Economic Papers and Current Research Fronts,” *Journal of Citation Studies* 50, Part 2, 1989: 10–15.
9. Ronald H. Coase, “The Nature of the Firm,” *Economica*, New Series, 4(16), November 1937: 386–405.
10. Armen Alchian and Harold Demsetz, “Production, Information Costs, and Economic Organization,” *American Economic Review* 62, December 1972: 777–795; Richard M. Cyert and James G. March, *A Behavioral Theory of the Firm*, Englewood Cliffs, NJ: Prentice-Hall, 1963; Michael Jensen and William Meckling, “Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure,” *Journal of Financial Economics* 3(4), 1976: 305–360; Sanford Grossman and Oliver Hart, “Takeover Bids, the Free-Rider Problem, and the Theory of the Corporation,” *Bell Journal of Economics* 11, Spring 1980: 42–64; 以及 Oliver Williamson, *Markets and Hierarchies: Analysis and Antitrust Implications*, New York: Free Press, 1975.

11. 反对科斯的自由主义者往往假定先来者有权在后到者出现后仍然继续不受限制地进行他们从事的活动，见Hans-Hermann Hoppe, “The Ethics and Economics of Private Property,”chapter 2 in Enrico Colombatto, ed., *The Elgar Companion to the Economics of Property Rights*, Cheltenham, U.K.: Edward Elgar, 2004。但赋予先到者这种权利若确实能够产生高效率，支持上述意见与科斯理论就并不矛盾。
12. Steven Holmes and Cass R. Sunstein, *The Cost of Rights: Why Liberty Depends on Taxes*, New York: W. W. Norton, 1999.
13. 与我同在康奈尔大学的博学同事罗伯特·霍基特（Robert Hockett）提示我，科斯在来到美国之前曾是英国经济学家尼古拉斯·卡尔多（Nicholas Kaldor）的同事，卡尔多是个极端的结果论者，也是著名的成本-收益分析的拥护者。
14. 案例来自Bernard Williams. See J.J.C. Smart and Bernard Williams, *Utilitarianism: For and Against*, Cambridge: Cambridge University Press, 1973。
15. *Loving v. Virginia*, 388 U.S. 1 (1967).
16. 这一观点的论据从该书第6章中概括而来： *Luxury Fever*, New York: Free Press, 1999。

07：效率至上：金钱在调节利益冲突者之间关系中的作用

1. Clinton Executive Order 12866, 3 CFR 638 (1993).
2. 为了方便，在这里我假设一个人买某样东西时愿意支付的数额就等于他愿意卖出这样东西的价格。在有些情况下，购买时愿意付的钱要比卖出时愿意接受的钱低很多。然而在这类情况下，我们只需要稍

稍改变条件，假定马尔科姆的两种心理价位都比苏珊的要高就可以了，这一更改并不会影响我们的推理。

3. S. J. Pokharel, G. A. Bishop, and D. H. Stedman, “On-Road Remote Sensing of Automobile Emissions in the Los Angeles Area: Year 2 Final Report Prepared for CRC,” March 2001; “CARB Throws a Hail-Mary Regulation,” Patrick Bedard, *Car and Driver*, November 2001, p. 134.
4. “Exploring the Role of Pricing as a Congestion Management Tool,” Searching for Solutions: A Policy Discussion Series, No. 1, Federal Highway Administration and Federal Transit Administration, Washington, D.C., July 23, 1991.
5. Michael O'Hare, Lawrence Bacow, and Debra Sanderson, *Facility Siting and Public Opposition*, New York: Van Nostrand Reinhold, 1983; and Howard Kunreuther and Paul Kleindorfer, “A Sealed-Bid Auction Mechanism for Siting Noxious Facilities,” *American Economic Review* 76(2), May 1986: 295–299.

08： “这就是你们的钱”： 人们的收入并不只是付出努力的结果

1. Liam Murphy and Thomas Nagel, *The Myth of Ownership*, New York: Oxford University Press, 2001.
2. Liam Murphy and Thomas Nagel, “Tax Travesties,” *Boston Globe*, January 26, 2003.
3. 有关继续开拓令人激动的新天地的相关文献的精彩摘要请参阅 Robert Hockett, “Taking Distribution Seriously,” Cornell University Law School working paper, 2011。

4. Robert H. Frank, "Are Workers Paid Their Marginal Products?" *American Economic Review* 74(4), September 1984: 549–571.
5. Mohammad Qamar uz Zaman, "Review of the Academic Evidence on the Relationship between Teaching and Research in Higher Education," U.K. Department for Education and Skills Research Report RR506, 2004.
6. Robert H. Frank, "Are Workers Paid Their Marginal Products?"; Jeffrey Pfeffer and A. Davis Blake, "Determinants of Salary Dispersion in Organizations," *Industrial Relations* 29, Winter 1990: 38–57; Jeffrey Pfeffer and A. Konrad, "Do You Get What You Deserve? Factors Affecting the Relationship between Productivity and Pay," *Administrative Science Quarterly* 35, June 1990: 258–285.
7. Robert H. Frank and Robert M. Hutchens, "Wages, Seniority, and the Demand for Rising Consumption Profiles," *Journal of Economic Behavior and Organization* 21, 1993: 251–276; and Robert H. Frank, *What Price the Moral High Ground?* Princeton, NJ: Princeton University Press, 2004, chapter 6.
8. David Neumark and Andrew Postlewaite, "Relative Income Concerns and the Rise in Married Women's Employment," *Journal of Public Economics* 70, 1998: 157–183.
9. Erzo Luttmer, "Neighbors as Negatives," *Quarterly Journal of Economics: Relative Earnings and Well-Being* 120, 2005: 963–1002.
10. 我在1985年出版的一本书的第6章中对这一主张做了更详细的辩护，具体请参见 *Choosing the Right Pond: Human Behavior and the Quest for Status*, New York: Oxford University Press。

09 : 成功与运气：成功对于运气的依赖比多数人想象的要大

1. 相关总结详见Joseph LeDoux, *The Emotional Brain*, New York: Simon and Schuster, 1998。
2. 同上, p. 165。
3. National Taxpayers Union, “Who Pays Income Taxes and How Much?”ntu.org/tax-basics/who-pays-income-taxes.html.
4. Ken Rudin, “Dreading a Tie in the Electoral College,” National Public Radio.
5. Instapundit.com, Comment of the Day, June 27, 2010.
6. Leonard Mlodinow, *The Drunkard's Walk: How Randomness Rules Our Lives*, New York: Vintage, 1990, chapter 10.
7. Malcolm Gladwell, *Outliers*, New York: Pantheon, 2008, pp. 54, 55.
8. 同上, Chapter 1。
9. K. Anders Ericsson, Michael J. Prietula, and Edward T. Cokely, “The Making of an Expert,” *Harvard Business Review*, July–August, 2007.
10. 同上, p. 3。
11. Robert H. Frank and Philip J. Cook, *The Winner-Take-All Society*, New York: Free Press, 1995, chapter 4.
12. *Business Week* Annual CEO Compensation Survey, April 9, 2001.
13. Frank and Cook, op. cit.
14. Sherwin Rosen, “The Economics of Superstars,” *American Economic Review* 71, December 1981: 845–858, quote on p. 845.

15. Xavier Gabaix and Augustin Landier, “Why Has CEO Pay Increased So Much?” *Quarterly Journal of Economics* 123(1), 2008: 49–100.
16. James Quirk and Rodney D. Fort, *Pay Dirt: The Business of Professional Team Sports*, Princeton, NJ: Princeton University Press, 1992.
17. 关于IBM转型的历史，具体参见Doug Garr, *IBM Redux: Lou Gerstner and the Business Turnaround of the Decade*, New York: Harper Paperbacks, 2000。
18. Frank and Cook, op. cit.
19. 根据1988年的一项研究，近一半的美国公司的首席执行官在他们领导的组织中任期最少为20年，具体请参见Louis Boone and David Kurtz, “CEOs: A Group Profile,” *Business Horizons*, July–August 1988, pp.38–42。
20. Paul Pierson and Jacob Hacker, *Winner-Take-All Politics*, New York: Simon and Schuster, 2010.
21. Amy F. Woolf, “Nonproliferation and Threat Reduction Assistance: U.S. Programs in the Former Soviet Union,” Congressional Research Service, February 4, 2010.

10：伟大的权衡：为何降低高收入者的税率会抑制经济增长

1. “Business Bulletin,” *Wall Street Journal*, March 26, 1997, p. A1.
2. Alberto Alesina and Dani Rodrik, “Distributive Politics and Economic Growth,” *Quarterly Journal of Economics*, 109(2), May 1994: 465–490.
3. Michael Lewis, *Liar's Poker*, New York: W. W. Norton, 1979.

4. 关于过度自信的研究，请见Thomas Gilovich, *How We Know What Isn't So*, New York: Free Press, 1991, pp. 77–78。
5. Garrett Hardin, “The Tragedy of the Commons,” *Science* 162, 1968:1243–1247.
6. 更多关于导致赢家通吃市场人满为患的动机的详细讨论，请见Robert H. Frank and Philip J. Cook, *The Winner-Take-All Society*, New York: Free Press, 1995, chapter 6。
7. Jenny Anderson and Andrew Ross Sorkin, “Bill Is Offered to Increase Tax on Private Equity,” *New York Times*, June 23, 2007.
8. Robert D. Yaro, “An Investment We Have to Make,” *New York Times*, October 14, 2010.
9. U.S. Department of Energy, “The Smart Grid: An Introduction,” 2008.

11 ⚡ 对自由竞争的规制：对有害行为征税

1. A. C. Pigou, in *The Economics of Welfare*, 4th ed., London: Macmillan, 1932.
2. 相关精彩案例请参见Gary W. Dorris, “Redesigning Regulatory Policy: A Case Study in Urban Smog,” PhD dissertation, Department of Applied Economics and Management, Cornell University, 1996。
3. Joseph Kruger and Melanie Dean, “Looking Back on SO₂ Trading: What's Good for the Environment Is Good for the Market,” *Public Utilities Fortnightly*, August 1, 1997.
4. A. P. Sokolov, P. H. Stone, C. E. Forest, R. Prinn, M. C. Sarofim, M. Webster, S. Paltsev, C. A. Schlosser, D. Kicklighter, S. Dutkiewicz, J. Reilly, C. Wang, B. Felzer, J. Melillo, and H. D. Jacoby, “Probabilistic

Forecast for 21st Century Climate Based on Uncertainties in Emissions (without Policy) and Climate Parameters,” MIT Joint Program on the Science and Policy of Global Change, Report 169, January 2009.

5. “Climate Change Policy and CO₂ Emissions from Passenger Vehicles,” Congressional Budget Office, October 6, 2008.
6. Jonathan Leape, “The London Congestion Charge,” *Journal of Economic Perspectives* 20(4), Fall 2006: 157–176.
7. Henry Goldman, “New York City Council Approves Manhattan Traffic Fees,” Bloomberg.com, April 1, 2008.
8. Keith Bradsher, *High and Mighty*, New York: Public Affairs, 2002.
9. Jonathan Gruber and Sendhil Mullainathan, “Do Cigarette Taxes Make Smokers Happier?” *Advances in Economic Analysis and Policy* 5(1), 2005: 1–43.
10. T. K. Greenfield and J. D. Rogers, “Who Drinks Most of the Alcohol in the U.S.? The Policy Implications,” *Journal of Studies on Alcohol and Drugs* 60(1), January 1999: 78–89.
11. Philip J. Cook, *Paying the Tab: The Costs and Benefits of Alcohol Control*, Princeton, NJ: Princeton University Press, 2007.
12. N. Gregory Mankiw, “Can a Soda Tax Save Us from Ourselves?” *New York Times*, June 5, 2010.

12 : 重新思考自由主义者的反对主张

1. John Rawls, *A Theory of Justice*, Cambridge, MA: Belknap Press of Harvard University Press, 1971.

2. Harriet Rubin, "Ayn Rand's Literature of Capitalism," *New York Times*, September 15, 2007.
3. Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, State College, PA: Penn State University, 2005, book 1, chapter 1 (originally published in 1776).
4. 同上, book 1, chapter 3。
5. Robert Nozick, *Anarchy, State, and Utopia*, New York: Basic Books, 1974.
6. www.sirclisto.com/cavalier/spain.htm.

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

湛庐阅读App：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本
怎么读是读者面临的最大阅读障碍
“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”App。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。

从单一到复合，从知道到精通，从理解到创造，湛庐希望建立一个“与最聪明的人共同进化”的社区，成为人类先进思想交汇的聚集地，与你共同迎接未来。

与此同时，我们希望能够重新定义你的学习场景，让你随时随地收获有内容、有价值的思想，通过阅读实现终身学习。这是我们的使命和价值。

湛庐阅读App玩转指南

湛庐阅读App 结构图:



三步玩转湛庐阅读App:



App获取方式:

安卓用户前往各大应用市场、苹果用户前往 App Store
直接下载“湛庐阅读”App，与最聪明的人共同进化！

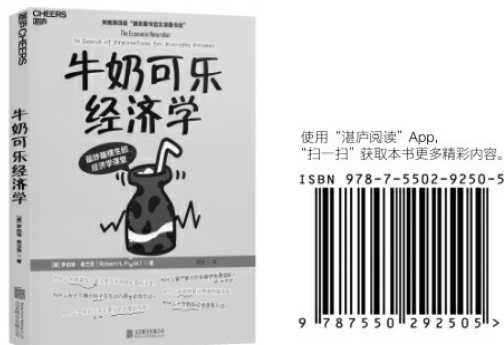
使用App扫一扫功能， 遇见书里书外更大的世界！



延伸阅读

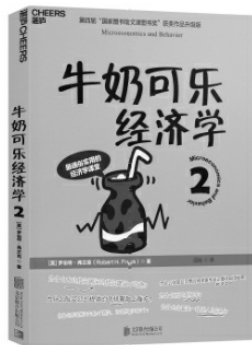
《牛奶可乐经济学》

- ◎ 《牛奶可乐经济学》是用经济学的原理和方法来解释我们在现实中司空见惯而又未注意到的现象，并通过这些事例和解释来加深人们对经济学的理解，可以说是通俗经济学的开山之作。
- ◎ 荣膺第四届“国家图书馆文津图书奖”。著名经济学家梁小民、诺贝尔经济学奖获得者罗伯特·索洛倾情推荐。



《牛奶可乐经济学2》

- ◎ 针对日常生活中“有限理性”的人们和在商界苦心经营的企业家们，相对应地打造了消费者行为学和组织行为学两块核心内容，读者完全可以把自已的问题在书中对号入座，我们为你一一破解行为背后的经济学谜题。



使用“湛庐阅读”App,
“扫一扫”获取本书更多精彩内容。

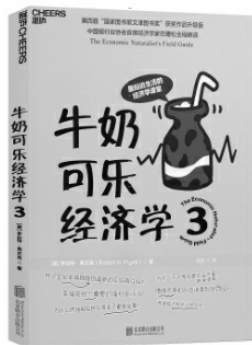
ISBN 978-7-5502-9251-2



9 787550 292512 >

《牛奶可乐经济学3》

- ◎ 经济学家到底能做什么？如果你觉得这个问题离你过于遥远，那我们不妨换个话题：经济学到底能做什么？罗伯特·弗兰克的《牛奶可乐经济学3》，实际上也是从特定的角度，试图给予一个解释。也许这本著作的观点不一定很深刻——事实上，全书用的都是经济学教科书里非常基本的原理；然而，罗伯特·弗兰克的立场更客观、方法更独特。



使用“湛庐阅读”App,
“扫一扫”获取本书更多精彩内容。

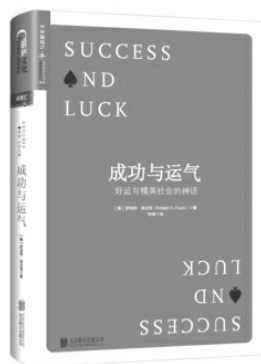
ISBN 978-7-5502-9771-5



9 787550 297715 >

《成功与运气》

- ◎ 《成功与运气》是一本媲美《异类》的成功动力学力作。在书中，罗伯特·弗兰克颠覆了人们对成功与运气的固有认知，提出“运气是可以放大和累积的，竞争越激烈，运气越重要。”



使用“湛庐阅读”App,
“扫一扫”获取本书更多精彩内容。

ISBN 978-7-5596-0757-7



-
- (1) 通过实践的途径确认工具（手段）的有用性，从而追求事物的最大功效，为人的某种功利目的服务。——编者注
 - (2) 对人民生活、经济等控制严厉政府。——译者注
 - (3) 美国保守主义者的一项财政手段，“野兽”一词指美国政府和投资项目，特别是社会项目。——编者注
 - (4) 萨克拉门托为美国加利福尼亚州首府，奥尔巴尼为纽约州首府。——编者注
 - (5) 涓滴理论是指在经济发展过程中并不给予贫困阶层、弱势群体或贫困地区特别的优待，而是由优先发展起来的群体或地区通过消费、就业等方面惠及贫困阶层或地区，带动其发展和富裕，或认为政府财政津贴可经过大企业再陆续流入小企业和消费者之手，从而更好地促进经济增长的理论。——译者注
 - (6) 一个大型免费分类广告网站。——译者注
 - (7) 滑坡谬误是一种逻辑谬论，即不合理地使用连串的因果关系，将“可能性”转化为“必然性”，以达到某种意欲而为之结论。——译者注

CHEERS
湛庐

OUT

The

巴菲特 最推崇的8大 企业家特质

Eight Unconventional CEOs and
Their Radically
Rational Blueprint for Success

巴菲特在致股东的信中盛赞
查理·芒格亲自审定
杨天南领衔翻译

马斯文 杨天南 译

【美】威廉·桑代克 著
William N. Thorndike, Jr.

SRH



中国财经出版传媒集团
中国财政经济出版社

版权信息

书名：巴菲特最推崇的8大企业家特质

作者：〔美〕威廉·桑代克（William N. Thorndike, Jr.）

译者：马斯文 杨天南

出版社：中国财政经济出版社·湛庐

出版时间：2022年10月

ISBN：9787522316215

字数：216千字

版权所有·侵权必究

推荐序 中国企业最不喜欢做的几件事

张化桥
慢牛资本董事长

2016年，我到伦敦GIM基金公司餐叙，有幸获赠了威廉·桑代克在2012年出版的*The Outsiders*一书，非常有趣！

此书简单易读，它讲了美国8家上市公司在过去40～50年(1)内持续为股东创造超高收益率的故事。这8家公司有以下几个共同点。

长期地、大量地回购自己的股票。比如，特利丹公司（Teledyne）在1972—1984年，从公开市场回购了自己公司90%的流通股。华盛顿邮报公司仅在2009—2011年，就通过回购股票把股票总数缩减了20%。宠物食品巨头罗尔斯顿·普瑞纳公司（Ralston Purina）回购了60%的流通股。其首席执行官（CEO）说，投资自己是把握最大的投资，所有资本全部在自己掌控之中。其实，任何并购和资本支出都必须跟回购股票的收益进行比较，这是一道最低的门槛。减少股份，可以增加净资产收益率，提高资金使用效率，也可以增加股票的每股收益和含金量。现金太多，就会造成浪费，容易投资愚蠢的项目。

埃克森美孚公司曾在2007—2012年回购了其净发行股票的25%。另外，思科、苹果、IBM和甲骨文这4家公司仅在过去的12个月，就回购了总计900亿美元的股票。

高额分红当然好，但由于分红需要缴纳两次税（公司层面、股东层面），所以回购股票（不涉及税收）其实是回馈股东的最好办法：可以一边回购股票，一边借款扩张（利息费用能够在税前扣除）。

在美国，只要价格合适，管理层就会随时乐意把主营业务卖掉，做减法。在中国，把主营业务卖掉会涉及一些额外问题：管理层的职位安排问题、做假账问题、历史遗留问题、权力问题以及面子问题等。

在美国，许多大集团随时准备并购其他企业。1973—1989年，美

国的有线电视公司TCI做了482项并购，基本上每两周并购一家公司。当然，要买，就必须有人愿意卖才行。中国人认为“只有撑不下去了才卖”，所以，想买的人也就买不成。

在美国，许多大集团不到万不得已，不愿意增发新股。几十年下来，这些公司股数要么大减，要么不变。大众影院公司（General Cinema）的CEO迪克·史密斯（Dick Smith）认为增发新股是非常痛苦的事情。在史密斯的任期内，他回购了公司33%的流通股。他们做回购有时是在市场内买，有时是拍卖式定额回购。比如，在2012年，公司准备回购8亿美元股票，价低者得，收完为止。过一段时间，现金多了，再重复回购。

在CEO约翰·马隆（John Malone）的任期内，TCI回购了净发行股票的40%。马隆非常强调节省，拒绝更新该公司在美国各地有线电视业务中落后的（或者说是已被证明落后了的）技术：让“先烈”们去发明、开拓和验证吧！“等一等再投资，我们不会丢失主要的地盘，因为有线电视技术的先行者们总是敢于冒险。”马隆如是说。

这让我想起巴菲特的一句话：“去太空旅游当然激动人心，不过，我很乐意把这份激动让给他人。”

译者序 斯文在兹信达雅

文 / 杨天南
北京金石致远投资管理公司CEO

《巴菲特最推崇的8大企业家特质》一书终于面世了，它讲述了8家企业在杰出企业家领导之下取得巨大成功的故事。

这本书英文原名为*The Outsiders*，之前的译本名为《商界局外人》。对于这样贴近字面的“直译”，一直以来存有争议，因为在中文语境中，“局外人”一般指不相干的人，如果用“局外人”来形容书中这些杰出企业家，多少有些令人费解，因为他们明明就是企业的主导者，属于“局内人”啊！

对此，巴菲特理论资深研究专家刘建位先生认为，将outsiders理解为“异类、另类”更切合实际。我认同他的观点，因为书中提到的这些人即便不是天赋异禀，也绝对属于特立独行之人，他们敢于不走寻常路，敢为人先，不从众、不随大溜儿，所以，他们看起来像是芸芸世界的“局外人”。

读到这里，读者或许心生疑惑：“怎么可以这样自卖自夸，毫不避嫌呢？”答案是，虽然由我来作这篇译者序，但本书的主力译者却是马斯文先生。斯文是我太极班和投资人生班上的学生，早年留学欧洲，归国之后曾在高校任教10年，工作是教外国人学好中文，这令他对于中英双语转换有着异于常人的敏锐。承蒙湛庐错爱，让我挂着领衔之名，但一切功劳多归于斯文，这也是我坚持名列其后的原因。

本书提出，一名优秀的CEO应该做好两方面的工作：一是提高企业的运营效率，二是做好资本配置。前一个方面覆盖人们通常熟知的企业管理内容，如研发、产品、市场、存货、供应链等，总之是提高自身效率；后一个方面是如何配置运营活动中产生的现金流。

作者研究发现，大凡杰出的企业家，往往是在后一个方面的工作上大放异彩，无论是亨利·辛格尔顿（Henry Singleton），还是沃伦·巴菲

特，本书提及的这8家企业的当家人都是资本配置的高手。因为一般而言，很多企业的CEO通常来自企业内部的岗位提拔，或是来自市场部门，抑或是来自财务部门。无论是哪个岗位出身，其多会落于原有的窠臼，目光向内，着眼于内部效率提升。而具有资本配置眼光的企业家看中企业自由现金流的外部安排。例如，是扩张跨界、进行分红，还是回购股票，等等。本书作者的结论是，具有同样运营成果的两家企业，如果执行不同的资本配置策略，其对股东价值的长期影响会有天壤之别。

这些杰出的商界运作固然令人开阔眼界、帮人打开思路，但对于众多中小企业而言，在未能拥有丰富资源之前，例如，对富余现金的安排是进行分红还是用来回购股票，企业筹资方式是银行贷款，还是发债、发股融资……这样的“高阶管理”问题恐怕暂时还不在考虑范围之内，因为它们大多处于资金缺乏、融资手段单一的发展阶段。

我认为，CEO工作涉及的两个主要方面对普通人也有很大启发，因为每个人都是自己的CEO。我在大学教工商管理硕士（MBA）投资学已有5年，这个M就是管理。虽然MBA讲的是工商管理，但我倒是觉得人生处处都是管理，如投资理财管理、社会关系管理、人际关系管理等。

比照CEO应该做好的两件事，即一是提高企业的运营效率，二是做好资本配置，我们见到市场上太多人偏重后者，而忽视前者。为了手中的三五万元现金，天天盯盘股市涨跌，美其名曰“资产管理”，实际上却荒废了自己的本职工作，忘记了提高自身效率，到头来往往往遭到巨大风险的暴击，导致金钱与青春双重损失，弊大于利，得不偿失。

本书中的CEO们因重视资本配置而成功，但这里有个前提，那就是企业本身运营良好，能够产生源源不断的现金以供配置。对现实中那些手中资金有限，却一心考虑走职业投资道路的人来说，荒废了主业其实是自我断绝了源源不断的现金流。这就是巴菲特喜欢股票下跌，而普通人畏惧下跌的秘密所在。

有人用“信达雅”来形容好的翻译，在一定程度上，好的翻译是一种文字的再创作。例如那句名言“You shape your houses and then your houses shape you”，可以翻译为“你塑造房子，然后你的房子也塑造你”，也可以翻译为“你塑造建筑，然后建筑塑造你”。但实际上，如果翻译为“你塑造环境，环境反过来也在塑造你”，恐怕更贴近原意，更容易理解。

如此细致推敲的工作需要时间、精力、经验，还必须具有一定的牺牲奉献精神，以及最重要的一点——热爱。无论是在出版界，还是译作者队伍，这样的人都是值得珍惜的。

本书的担纲译者马斯文先生就是这样一位有心人，他的工作背景使他尤擅文字上的推敲琢磨。“斯文”一词常常用来形容有修养的人，有涵养、有礼貌，文质彬彬，懂得尊重人。但这样的人往往是不会主动要求别人点赞的，因为他们内心认为，这样的恳求也有点“低三下四，有辱斯文”。

不久前，我到曲阜孔庙拜访时，见大殿前的匾额上书“斯文在兹”，我拍了照片发给斯文，他回复表示“一定好好努力，不辱斯文”。各位读者阅后如果有所喜爱，不妨多多支持一下，塑造一个利人利己的环境，以免将来“斯文不在”的遗憾。

前言 谁比杰克·韦尔奇更伟大的CEO

沃伦·巴菲特 人称“股神”	给真正优秀的 CEO 付再多薪酬 也不为过……但是，他们属于 稀有物种，踪迹难觅。
比尔·帕斯尔斯 (Bill Parcels) 传奇足球教练	你是谁，由你过往的记录说了算。
约翰·邓普顿 (John Terpleton) 著名基金经理人	成功总是有迹可循的。

谁是过去50年来最伟大的CEO？

你如果和大多数人想的一样，那么可能脱口而出的回答是“杰克·韦尔奇”。原因显而易见，1981—2001年，韦尔奇经营美国通用电气公司20年，而通用电气被认为是美国最具标志性的公司之一。通用电气的股东在韦尔奇任内赚翻了天，年复合回报率达到20.9%。如果你在韦尔奇就任CEO时购买了1美元的通用电气股票，那么当他把权力移交给继任者杰夫·伊梅尔特（Jeff Immelt）时，你的股票价值会达到惊人的48美元。

韦尔奇既是一位活跃的管理者，也是一位重要的企业形象大使。他传奇般地四处巡视，不停地出差造访通用电气遍布各地的分支机构，不知疲倦地给经理人评级并让他们在不同的业务部门间轮岗。韦尔奇制订了全公司的战略计划，并用“六西格玛”（Six Sigma）和“全面质量管理”（TQM）等充满异域色彩的名字来命名。韦尔奇性格活泼好斗，经常周旋于华尔街和商业媒体之间。韦尔奇在聚光灯下挥洒自如，在通用电气任职期间，经常成为《财富》杂志的封面人物。即使退休以后，韦尔奇依然不时就继任者的表现等各种商业话题发表激进言论，进而常常占据新闻的头条。韦尔奇还写了两本管理建议方面的书，书名通常都充满了攻击性，比如他的自传就叫作*Straight from the Gut*，这个书名直译

为“直截了当”，中文版名为《杰克·韦尔奇自传》。

关于韦尔奇的争议和他出色的业绩交织在一起，他创设的一套特殊管理方法已然成为衡量CEO业绩的黄金标杆。韦尔奇强调要积极监督公司运营，要定期与华尔街沟通以及要高度关注公司股价。然而，韦尔奇是过去50年来最伟大的CEO吗？

答案是一个大写的“不”。

要明白这么说的原因，重要的是从一开始就使用一套全新的、更精确的方法来衡量CEO的能力。CEO和职业运动员一样，在一个高度量化的领域里竞争，却没有一个简单的、公认的度量标准来衡量他们的表现，就好像对棒球投手没有ERA(2)的考核，对外科医生没有并发症发生率的考核，或者对曲棍球守门员没有平均失球数的考核一样。

那些商业媒体并不试图以任何严格的方式来确定最佳选手，相反，它们通常只关注最大、最知名的公司，比如《财富》100强的上榜者。这就是这些公司的高管经常出现在顶级商业杂志封面上的原因。媒体通常关注的指标是公司收入和利润的增长，然而，衡量CEO是否伟大的最终晴雨表，不是销售额、收入或员工人数的增长，而是公司股票每股价值的增长。此种情形像极了《体育画报》(*Sports Illustrated*)，它只将身材最高大的投手和最壮硕的守门员放在封面上。

评估公司的业绩时，重要的不是绝对回报率，而是与同行和市场比较的相对回报率。评价一位CEO是否伟大时，你也仅需参考三件事：他的任期内给股东带来的年复合回报率，同行同期的回报率，通常由标准普尔500指数来衡量的市场大盘同期回报率。

时代背景非常重要，开始和结束的时点都会对一名CEO的业绩产生巨大的影响。韦尔奇的任期几乎正好与1982年底开始的史诗级的牛市持续的时间相吻合，这一波行情基本上无间断地持续到2000年初。在这段令人瞠目的时期，标准普尔指数的平均年回报率为14%，约为长期平均数值的两倍。在牛市行情中实现20%的回报率是一回事，而在横跨数个大熊市后实现20%的回报率，则完全是另一回事。

我们可以用棒球做类比来说明这一点。在20世纪90年代中后期，29记全垒打的进攻成绩只能算是平庸水平，因为佼佼者可以打出60记以上。然而，当贝比·鲁思(Babe Ruth)在1919年打出29记全垒打时，他

打破了自1884年以来保持的纪录，开创了以力量为导向的打法并永久改写了棒球运动的历史。所以再强调一遍，时代背景非常重要。

评估CEO历史业绩的另一个重要因素是其相对于同行的表现，而评定这点的最佳方法是将CEO与广泛的同业者进行比较。就像复式桥牌比赛一样，选手们持有的手牌通常大致相同，因此长期业绩的差别更可能来源于CEO的管理能力，而不是外部因素。

让我们来看一个采矿业的例子：2000年，黄金价格在每盎司(3)400美元上下徘徊，2011年黄金价格最高达到每盎司1 900美元。我们绝不可能把这两种情况与金矿企业CEO的表现相提并论。金矿行业的CEO们无法控制标的商品的价格，他们必须利用市场赋予的“手牌”，为股东做出最佳选择。因此在评估业绩时，最有用的方法是，将金矿企业的CEO与其他相同环境下运营的公司的管理者进行比较。

当一位CEO带来的回报率远远高于同行和市场时，他就堪称“伟大”。按照这个定义，韦尔奇任职期间通用电气的业绩超过标准普尔500指数的3.3倍，他是当之无愧的伟大的CEO。

然而，他和辛格尔顿完全不可同日而语。

力压韦尔奇，了不起的辛格尔顿

如今，只有少数投资者和业内人士知道辛格尔顿是个了不起的人物。作为一名CEO，辛格尔顿有着不凡的背景：他是一位喜欢下盲棋的世界顶级数学家，在获得电气工程博士学位的同时，还曾为麻省理工学院的第一台计算机编程。第二次世界大战期间，辛格尔顿开发了一种“消磁”技术，使盟军舰艇能够避开雷达探测；20世纪50年代，他开发了一种惯性制导系统，至今仍为大多数军用和商用飞机所使用。以上这些都是过去式。60年代早期，辛格尔顿创建了综合企业集团特利丹公司，自此成为历史上最伟大的CEO之一。

60年代的综合企业集团就好比之后的网络公司，当时大批综合企业集团纷纷上市。然而，辛格尔顿经营的是一家极不寻常的企业集团。早在众人蜂拥而至之前，辛格尔顿就积极地回购了自己的股票，最终买回了特利丹公司90%以上的股权；辛格尔顿避免分红，看重现金流而不是账面盈余，用去中心化(4)的方法让集团良好运转。辛格尔顿从不分拆股票，在70年代和80年代的大部分时间里，特利丹公司是纽约证券交易所

（NYSE）交易价格最高的股票。辛格尔顿因不愿与分析师或记者交谈而被称为“斯芬克斯”（Sphinx）[\(5\)](#)，他也从未在《财富》杂志的封面上出现过。

辛格尔顿是一个反传统主义者，他的特立独行在华尔街和商业媒体上引发了广泛的质疑和恐慌。结果证明，辛格尔顿完全有理由对这些怀疑者置之不理。那些更知名的同行的长期回报率都很平庸，他们的平均年化回报率只有11%，略微高于标准普尔500指数。

相比之下，辛格尔顿经营特利丹公司近30年，其投资者的年复合回报率达到了惊人的20.4%。如果你在1963年向辛格尔顿投资1美元，那么到1990年，当辛格尔顿在熊市的撕咬中退休时，这笔投资的价值会达到180美元。同样的1美元投资于其他大多数综合企业集团，只能得到27美元；如果投资于标准普尔500指数，只能得到15美元。令人瞩目的是，辛格尔顿的表现是标准普尔500指数的12倍。

根据我们对成功的定义，辛格尔顿是一位比杰克·韦尔奇更伟大的CEO。辛格尔顿的各项数据明显更好：相对于市场和同行，他的每股回报率不仅更高，而且保持高回报率的时间更长（28年，韦尔奇是20年），而且这期间市场还经历了数次漫长的熊市。

辛格尔顿的成功并非缘于特利丹公司拥有什么独特的、快速增长的业务。辛格尔顿与同行的主要区别在于，他在一个关键但有些神秘的领域，即资本配置方面游刃有余。资本配置过程决定了如何配置公司资源，从而使得股东获得最好的回报。因此，让我们花少许篇幅来解释：什么是资本配置？为什么资本配置如此重要？为什么很少有CEO精通此道？

重新定义职责，**CEO**的超级工具箱

CEO要想取得成功，需要做好两件事：一是让企业高效运营，二是合理配置运营产生的现金。大多数CEO以及他们撰写或阅读的管理书籍都专注于对运营的管理，不可否认，运营非常重要。辛格尔顿却与众不同，他把大部分精力都放在了第二件事上。

通常而言，CEO们在资本配置上有5种基本的选项，即投资现有业务、收购其他业务、发放股息、偿还债务、回购股票；而获得资本有3种选项，即发掘内部现金流、发行债券、发行股票。

想象一下：如果我们把所有的选项都放入CEO的工具箱里，那么长期而言，股东的回报将主要取决于CEO在种种工具间的取舍。简单来说，两个经营业绩相同但资本配置方式不同的公司，会给股东带来两种截然不同的长期回报。

从本质上讲，资本配置就是投资，因此，所有的CEO既是资本配置人，也是投资人。事实上，这可能是任何一位CEO最重要的职责，然而尽管它很重要，顶级商学院却没有任何关于资本配置的课程。正如巴菲特所观察到的，极少有CEO能为这项关键任务做好准备，他说：

许多公司的负责人并不擅长资本配置，这种缺憾不足为奇。大多数领导能晋升到高层，是因为他们在市场营销、生产、工程、行政管理，有时甚至是职场政治中表现出色。他们一旦成为CEO，就必须当即做出资本配置的决策，而他们此前可能从未涉足过这项至关重要的工作，因此往往会显得手足无措。举个夸张一点儿的例子，就好比一位极具才华的音乐家，最后的归宿居然不是在卡内基音乐厅演出，而是被任命为美联储主席。¹

这种经验不足对投资者的回报有着直接而重大的影响。巴菲特强调了这种技能差距的潜在影响，他指出：“如果一家公司每年留存的收益保持在净资产的10%，并且这项资本配置工作持续10年，它的CEO就等于配置了这家企业总计60%以上(6)的资本。”²

辛格尔顿是资本配置大师，对于不同的资本配置选项，他的起领航作用的决定和同行明显不同，并对股东的长期回报产生了巨大的积极影响。具体而言，辛格尔顿将特利丹公司的资金集中在严格筛选后的并购和一系列大规模股份回购上。对发行股票，辛格尔顿的态度较为保守，却善于利用债务杠杆，直到20世纪80年代末才派发股息。对比之下，其他综合企业集团奉行的资本配置策略就好像它的镜像世界，即积极发行股票去并购公司，派发股息，避免股票回购，一般很少利用债务杠杆。简而言之，其他综合企业集团在CEO的工具箱里配置了一套不同的工具，结果也自然不同。

如果你从更广泛的角度来思考，资本配置就是资源配置，包括人力资源的配置，那么你会再次想起辛格尔顿那高度差异化的方法。具体来说，辛格尔顿相信极致的去中心化的组织形式，即公司总部结构单一、人员精简，日常运营的责任和权力集中在各个业务部门的总经理身上。这与同行们的做法截然不同，同行的总部通常部门林立，充斥着各种副职和MBA的头衔。

事实证明，真正伟大的CEO都深谙资本配置之道。实际上，他们的方法和辛格尔顿惊人地相似。

局外人视角，特立独行的CEO

1988年，巴菲特撰写了一篇关于一群投资者的文章，这群投资者不仅有出色的业绩记录，而且信奉传奇的哥伦比亚商学院的教授本杰明·格雷厄姆（Benjamin Graham）和戴维·多德（David Dodd）的价值投资原则。格雷厄姆和多德颠覆了传统的投资策略，他们主张买入那些净资产价值被远远低估的公司。

为了说明非凡的投资回报与格雷厄姆-多德原则之间的紧密关联，巴菲特用了全美掷硬币大赛的比喻。在这场比赛中，2.25亿个美国人每天下1美元的赌注掷硬币猜正反。每一天的输家都会退出，而赢家赢的钱都将继续押在第二天的赌注上，因此赌注不断加码。20天后，剩下215人，每人赢得了超过100万美元的奖金。巴菲特指出，这个结果纯粹是运气使然，就算是2.25亿只猩猩参与也会产生同样的结果。然后，巴菲特谈及一个有趣的现象：

如果发现有40只猩猩来自奥马哈的某个特定的动物园，你将会非常肯定这意味着什么……科学研究通常也遵循这样的模式。对一种罕见的癌症，如果你试图找出其背后可能的病因，发现此类癌症的400个病例都发生在美国蒙大拿州的某个矿业小镇上，你就会对那里的水、那些患者的职业或者其他变量非常感兴趣。我想你会发现，在投资领域，极高比例的掷硬币赢家来自一个极小的知识部落，它可以被称为“格雷厄姆-多德部落”。³

历史学家劳雷尔·乌尔里克（Laurel Ulrich）曾写道，循规蹈矩的人很少创造历史。按照这个理论，传统的CEO也很少能战胜市场或同行；投资界也是如此，表现出众的管理者人数稀少。如果你把掷硬币的赢家的名单列出来，那么毫不奇怪，你会发现他们都是打破传统的人。

这些杰出的管理者正是本书描述的对象，他们经营的公司历经市场的繁荣与衰退，涉及制造业、媒体、国防、消费品和金融服务等不同的行业，他们的公司在规模和成熟度上差别巨大。没有一家公司拥有火爆的、可迅速复制的零售模式或知识产权壁垒，但其表现却远超同行。

与辛格尔顿一样，在各自的业务上，他们创立了独特的、明显不同的方法，这通常会招来同行以及商业媒体的大量评论与质疑。但更有趣的是，尽管他们彼此独立地提出了各自不同的方法，事实却证明，他们

骨子里都是不走寻常路的同路人。换言之，只有打破旧的束缚，才能成就新的模式，成功的潜在蓝图才能就此展开，出色的回报才能随之而来。

他们的经营活动似乎发生在和我们的时空平行的另一个宇宙，那个宇宙中的人们奉行着同样一套特立独行的原则和世界观，并因此获得了这个小小部落的居民身份。我们称之为辛格尔顿部落。抛开别的不说，这个部落里的男女精英都明白：

- 资本配置是CEO最重要的工作。
- 从长远来看，重要的是每股价值的增长，而不是总收入或企业规模的增长。
- 决定长期价值的是现金流，而不是账面盈余。
- 去中心化的组织形式释放了企业的能量，降低了成本以及管理层之间的“积怨”。
- 独立思考对长期成功至关重要，与外部顾问（华尔街、媒体等）打交道会分散精力且浪费时间。
- 有时最好的投资机会就是回购你自己公司的股票。
- 对于收购，耐心是一种美德……偶尔也得当机立断。

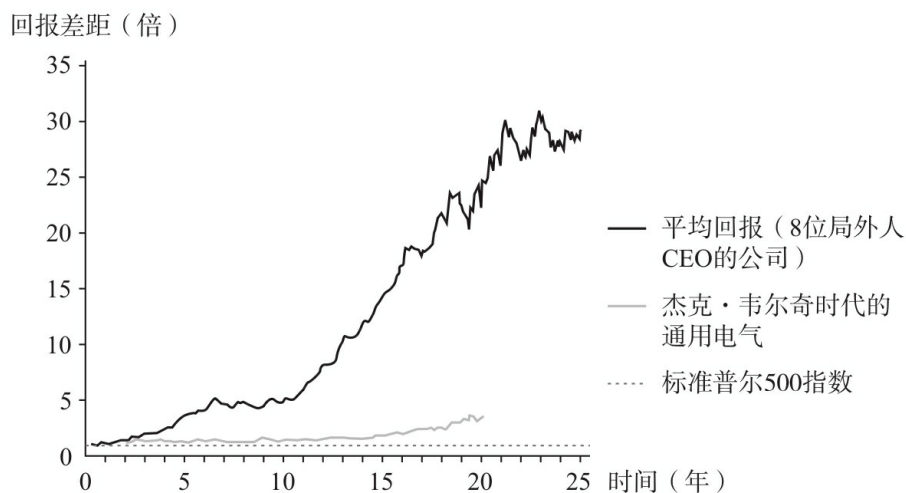
有意思的是，他们的不落俗套在很多情况下因地理位置而得到强化。大部分人的公司分布在丹佛、奥马哈、洛杉矶、亚历山大、华盛顿和圣路易斯等城市，从而远离波士顿—纽约走廊的金融地震带。这种距离有助于他们远离华尔街世俗认知的喧嚣。有两位CEO的办公室在美国东北部，他们都偏爱不起眼的位置：史密斯的办公室位于一个郊区购物中心后面；汤姆·墨菲（Tom Murphy）的办公室在曼哈顿中城的一处旧住宅区里，距离华尔街足有60个街区。

辛格尔顿部落的居民和身为“局外人”的CEO们都有一些有趣的个人特征：他们普遍节俭，这一点常常被世人调侃；他们为人谦逊，善于分析且为人低调。他们忠于家庭，常常提早离开办公室，去参加子女学校的活动。他们通常不喜欢CEO角色带来的抛头露面，无意在商会发表演

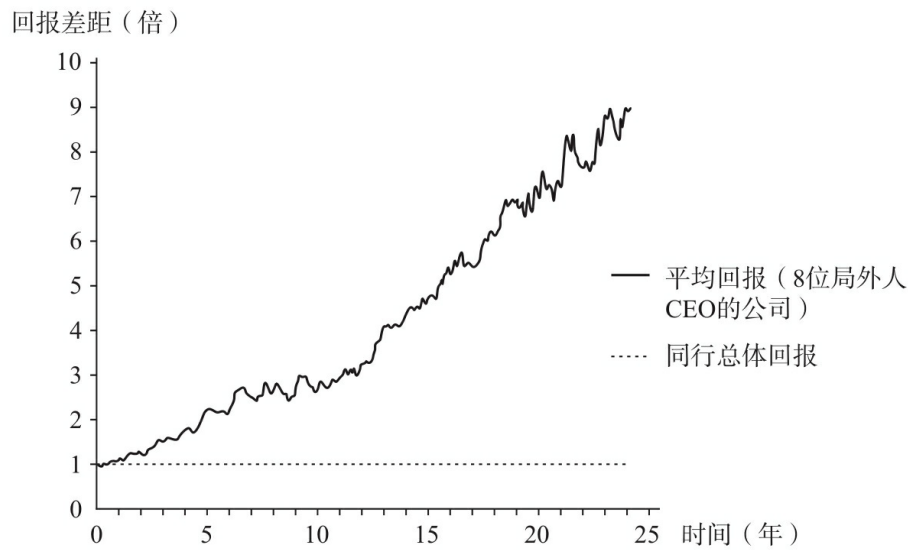
讲，也不出席达沃斯论坛。他们很少出现在商业出版物的封面上，也不撰写管理建议方面的书。他们不是啦啦队队长，不是营销员，不会和陌生人自来熟，他们似乎和号召力无缘。

与史蒂夫·乔布斯、山姆·沃尔顿（Sam Walton）、美国西南航空的赫布·凯莱赫（Herb Kelleher）、马克·扎克伯格等高调的CEO相比，他们显得格格不入。这些天才是商界的牛顿，就像是被苹果砸中而被激发出石破天惊的想法，继而以疯狂的专注和决心去努力实践。然而，他们面临的处境和环境与绝大多数企业CEO没有什么迥异之处，那些职业生涯中的经验教训也并非不可借鉴。

身为局外人的CEO们既没有沃尔顿和凯莱赫那样的感召力，也没有乔布斯和扎克伯格那样的营销或技术天赋。其实，他们的处境就是众多美国企业高管曾经历过的。但是，他们带来的回报却极不寻常。如图P-1和图P-2所示，平均而言，他们所带来的平均回报超过标准普尔500指数的20倍，超过同行的7倍。在接下来的章节里，我们将持续关注这些回报是如何实现的。正如水门事件的线人“深喉”所言，本书将“循着钱的印记”，仔细探究这些身为局外人的CEO为实现股东回报最大化而做出的关键决策以及得到的教训，从而给今天的管理者和企业家一些启示。



图P-1 8位局外人CEO带来的平均回报与标准普尔500指数的回报对比



图P-2 8位局外人CEO带来的平均回报与同行回报的对比

测一测 你了解一位出色的**CEO**都做对了什么吗？

- 一名优秀的**CEO**只把功夫下在企业管理方面就够了吗？（ ）
 - A. 是
 - B. 否
- “首席执行官就是首席投资官。”——这句话出自：（ ）
 - A. 巴菲特
 - B. 查理·芒格
 - C. 凯瑟琳·格雷厄姆
 - D. 迪克·史密斯
- **CEO**大量曝光有助于提高企业在公众媒体的影响力吗？（ ）
 - A. 有助于
 - B. 无助于

扫描下面二维码查看本书更多测试题

扫码鉴别正版图书
获取您的专属福利



扫码获取全部测试题及答案
看看你是否了解出色的CEO都做对了什么

引言 像做投资一样做管理

约翰·邓普顿

除非做些与众不同的事情，否则你不可能有卓越的表现。

《纽约客》杂志的供稿人阿图·葛文德（Atul Gawande）[\(7\)](#)用“积极型偏离者”这个词来描述医学领域中效率奇高的行动者。在葛文德看来，我们应该研究这些离群值，以便从中吸取经验并提高工作效能。¹

令人颇感意外的是，商业领域中的那些佼佼者，却没有像在医学、法律、政治、体育等其他领域那样被深入研究。研究了亨利·辛格尔顿的案例之后，在一群天资聪颖的哈佛大学MBA学生的帮助下，我开始寻找那些在其所处市场中轻松击败同行和杰克·韦尔奇的其他案例。事实证明，正如本书前言所引用巴菲特的那句话所说的，这些公司和它们的CEO确实属于稀有物种。在哈佛商学院贝克图书馆的数据库中进行了大量搜索之后，我们只找到了另外7家公司符合两个“击败”的条件。

十分有趣，像特利丹公司一样，这些公司和它们的CEO通常都不是很有名。很多当下风光的CEO，若论业绩，都和这些CEO判若云泥。

颠覆式创新，“局外人”视角引发的方法革命

在媒体的笔下，以杰克·韦尔奇为代表的当代成功的CEO是充满魅力、注重行动的领导者，他们在宽敞明亮的办公楼里工作，身边簇拥着勤奋的MBA毕业生。他们乘坐公司商务机出差，大多数时间在巡视公司运营状况、会见华尔街分析师并出席各种各样的会议。“摇滚明星”这个词经常用来形容这些在公司间来去匆匆的高管，新公司经常通过一番搜索就招募他们入职，而他们的上一份工作往往是另一家知名公司的高管。

自2008年9月雷曼兄弟公司倒闭以来，这类备受瞩目的CEO一直饱受诟病。这也不难理解，因为在大众的眼中，他们通常贪婪（可能涉嫌欺诈）且无情，他们坐着公司的飞机满世界飞，解雇工人，做一些损害股东价值的大额交易。简而言之，他们非常像美国真人秀节目《飞黄腾达（第一季）》（*The Apprentice 1*）中的唐纳德·特朗普（Donald Trump）。在这档节目中，特朗普毫不掩饰自己的贪婪和傲慢，毫无顾忌地自我推销。这些都与美国开国元勋本杰明·富兰克林所提倡的价值观背道而驰。

好在辛格尔顿部落的居民改变了人们对CEO的刻板印象。他们都是首次担任CEO，大多数几乎没有什么管理经验，没有人是从高位空降得到这份工作的。除了一个人外，其他人都是所在行业和公司的新人。他们中只有两人拥有MBA学位。作为一个群体，他们没有吸引或寻求聚光灯的照耀，反而在默默无闻的氛围中工作，通常只有屈指可数的内行投资者和狂热爱好者对他们赞赏有加。

作为一个群体，他们分享着传统的、前现代的价值观，包括节俭、谦逊、独立，以及保守与大胆不同寻常的结合。他们通常在简陋的办公室里工作并对此无比自豪，对诸如使用公司飞机等额外福利避之不及，尽可能躲开聚光灯，很少与华尔街或商业媒体打交道。他们还主动远离银行家和其他外部顾问，更喜欢自己的智囊和身边特定的小圈子。富兰克林肯定会喜欢他们的。

这群婚姻美满的中年男性（有一位是女性）过着看似平淡无奇、乐善好施的生活，然而在他们的商业生涯中，他们既不墨守成规，也不沾沾自喜。他们是“积极型偏离者”，骨子里都是不走寻常路的猛士。

“iconoclast”这个词源于希腊语，本意是“粉碎偶像者”。这个词演变到今天，更广泛的含义是指决意与众不同、傲慢地背离世俗规范的人。最初的“粉碎偶像者”来自崇拜偶像的社会或庙宇之外，他们是社会规范和习俗的挑战者，并让古希腊的人们感到异常恐惧。本书中描述的CEO们并没有那么可怕，但确实与他们那些远古先祖有着奇妙的相似之处：他们也是局外人，蔑视那些诸如支付股息或避开股票回购等习以为常的传统方法，享受离经叛道的乐趣。

和辛格尔顿一样，这些CEO总是做出与同行截然不同的决定。然而，他们并不是盲目的逆向思维者，而是通过仔细分析和用不寻常的财务指标来做出决定，这与行业或华尔街的惯例迥然不同，他们的决定是明智的打破传统的行为。

从这个角度而言，他们的反传统与迈克尔·刘易斯（Michael Lewis）在《魔球》（*Moneyball*）一书中所描述的比利·比恩（Billy Beane）的行为很相似。²比恩是常年缺钱的奥克兰棒球队的总经理，以统计分析为基础，他获得了超越那些有钱的竞争对手的优势。比恩的方法聚焦新指标，比如上垒率和长打率，与传统统计全垒打、打击率和打点“三驾马车”相比，前者与球队的胜率相关性更高。

比恩在分析上的洞察力影响了球队管理的方方面面，如从起草和改变策略到是否在比赛中偷垒或牺牲短打，而他选择不偷垒也不牺牲短打。在所有这些领域中，比恩的做法都非常不正统，但非常成功。尽管比恩的球队在联赛中的薪资总额排名倒数第二位，但在他任职的第一个6年中，球队4次打进了季后赛。

像比恩一样，辛格尔顿和本书中其他7位CEO对各自的业务开拓出了独特的、打破常规的方法，这引起了同行和商业媒体的大量评论与质疑。仍然像比恩一样，他们的业绩非常出色，轻轻松松就超越了号称传奇的杰克·韦尔奇和诸多同行对手。

他们的背景五花八门：一个是曾经绕月飞行的宇航员，一个是此前没有商业经验的女性，一个继承了家族企业，两个是量化分析博士，还有一个是以前从未经营过公司的投资人。对CEO这个角色来说，他们都是新兵，但是他们却共有两项重要的特性：新颖的视角，以及对理性根深蒂固的坚守。

以赛亚·伯林（Isaiah Berlin）在一篇关于列夫·托尔斯泰的著名文章中，引出了关于“狐狸”和“刺猬”这个富有启发意义的对比：“狐狸”通晓很多事，“刺猬”只知道一件事，但了解得很深入。多数CEO是“刺猬”，他们是在某个行业中深耕从而成长起来的，到担任最高职位时，他们对这个行业已经了如指掌。“刺猬”类型的CEO有许多积极的属性，包括专业技能的纯熟、专业知识的精通以及对事业的专注。

不过“狐狸”也有许多引人注目的特质，包括跨领域的连接和创新能力。毫无疑问，本书中的CEO们都是“狐狸”。他们熟悉其他公司、行业和学科，这种广度转化为新的视角，又反过来帮助他们形成新方法，最终转变成卓越的成果。

破除同侪压力，理性是最好的解毒剂

在1986年伯克希尔-哈撒韦公司的年报中，巴菲特回顾了他担任CEO的前25年，得出的结论是：迄今为止他职业生涯中最重要、最惊人的收获是发现了一种神秘的力量，这种力量好比我们在青春期时面对同龄人的压力，这种力量迫使CEO们模仿同行的种种行为。巴菲特将这股强大的力量称为“惯性驱使”[\(8\)](#)，并指出它几乎无处不在。巴菲特警告说，想取得成效的CEO们必须找到抗拒这种力量的某种方法。

本书中的CEO们都设法避开了这种强大驱使力量的潜在影响。怎么做呢？他们在共同的管理哲学中找到了解药，这是一种世界观，渗透在他们的组织和文化中，并推动他们做出企业经营和资本配置的决策。他们彼此独立地提出了自己的管理理念，但令人吃惊的是，尽管行业和环境千差万别，这群高管的经历却惊人地相似。

他们各自经营着高度去中心化的组织，至少进行了一次大规模的收购，制定了基于现金流的不寻常指标，并回购了大量股票。他们中无人支付大量股息或听从华尔街的指导。所有人都受到了来自同行和商业媒体的嘲笑、不解与质疑，所有人在很长的任期内（平均20多年）都有令人瞠目结舌、难以置信的表现。

传统的商界分为两大阵营：经营公司的和投资公司的。这些打破传统的CEO的经验表明，他们对CEO的工作有了新的、更加微妙的定义，这个定义不再强调有号召力的领导能力，而是更加重视对公司资源谨慎配置的能力。

归根结底，这些CEO更像投资者，而不是管理者。从根本上说，他们对自己的分析能力很有信心；在极少数情况下，当他们看到公司价值和股票价格之间的巨大差异时，他们已经做好大干一场的准备。当公司的股票价格便宜时，他们就买进，且通常是大量买进；当公司的股票价格高企时，他们就借机收购其他公司或筹集廉价资本，以备未来之需。他们如果不能确定项目是否有吸引力，就会安心地等待，有时会等待很长时间，例如大众影院公司的迪克·史密斯就整整等了10年。从长远来看，这种系统化的、有条不紊的低买和高卖相结合，为股东带来了非凡的回报。

·遥远的镜子：1974—1982年·

在评估这些局外人CEO对当下世界的意义时，有必要看看第二次世界大战后他们各自如何掌舵，对比当下的经济衰退，1974—1982年的危机持续时间更长且极为残酷。

1974—1982年，外部石油危机、灾难性的财政和货币政策以及美国历史上最严重的国内政治丑闻交织在一起。一系列负面消息导致了长达8年的严重的通货膨胀、与熊市相伴而行的两次大萧条、高达18%的利率、上涨3倍的油价，以及100多年来美国首位在任总统辞职。1979年8月，在这个黑暗时期的中间点，美国《商业周

刊》刊登了一篇著名的封面文章，题为“股票已亡？”

那次经济衰退期充满了不确定和恐惧，以致大多数管理者都袖手旁观。但对于所有的局外人CEO来说，这是他们职业生涯中最活跃的时期，每个人要么施行重大的股票回购计划，要么发起一系列的大型并购，或者像汤姆·墨菲一样二者兼而有之。作为一个群体，用巴菲特的话来描述，当他们的同行深深“恐惧”之时，他们自己却十分“贪婪”（2006年7月24日，巴菲特在与作者的访谈中提到的）。

这种对工作的重塑源于8位CEO共同的不寻常的背景。他们都是商界的局外人，所有人都是第一次担任CEO，其中一半上任时还不到40岁，除一人外其他人都是行业新人。他们不受以往经验或行业惯例的束缚，其业绩的综合记录显示了新视角的巨大力量。作为自古有之的催化剂，新鲜的视角激发着许多领域的创新。在科学领域，范式转换概念的提出者托马斯·库恩（Thomas Kuhn）发现，最伟大的探索几乎总是由行业新兵和奋发有为的年轻人做出的，例如，曾是印刷工人的本杰明·富兰克林在他年近不惑时驯服了闪电，曾是专利审查员的爱因斯坦在27岁这一年推导出质能方程 $E=mc^2$ 。

这种“狐狸”类型局外人的视角帮助这些CEO制定了差异化的方法，并影响了他们整体的管理理念。作为一个群体，他们非常独立，通常避免与华尔街来往，不屑于使用外部顾问，他们更喜欢去中心化的组织结构，各部门各司其职。

CEO的首要目标是追求长期每股价值的最大化

在最近的畅销书《异类》（*Outliers*）中，作者马尔科姆·格拉德威尔（Malcolm Gladwell）提出了一条经验法则，即要形成跨领域的专业技能需要1万小时的实践。³那么，这群新任CEO了不起的成功如何与上述法则相吻合呢？在被委以重任之前，这群CEO中显然没有人拥有近1万小时的管理履历，或许他们的成功表明了专业技能和创新之间有着重要的区别。

格拉德威尔的法则是技能达到精通水准的指南，这与创新不一定是一回事。正如引言开篇所引用的邓普顿的话所表明的那样，相对卓越的行为需要做些与众不同的事，而这些CEO共同世界观的核心是对理性的

坚守、对数据的分析、对自我的反思。

这8位CEO既不是富有魅力的梦想家，也不会为华而不实的战略宣言所迷惑，他们具有务实和求真的性情。通过培养极度的专注力，以及在文化和交流上严谨的态度，他们成功地屏蔽了传统观念的噪声。科学家和数学家经常谈到“繁”的背面是“简”，而这8位CEO都是量化分析的高手，很多人拥有工程学学位，而不是MBA学位；他们拥有化繁为简的天赋，屏蔽了同行和媒体的聒噪，专注于自身企业的核心经济特征。

在本书所有案例中，“化繁为简”使得局外人CEO们重视现金流，摒弃盲目追求报告期盈利的所谓华尔街“圣杯”。大多数上市公司CEO关注的重点是让企业每季度净利润最大化，这也可以理解，因为这正是华尔街崇尚的标准。但是净利润是一个会失灵的工具，企业的负债水平、税收、资本支出和过去的收购历史的差别，都可能会导致净利润的严重扭曲。

因此，局外人的公司经常有着复杂的资产负债表、积极的收购计划和高负债水平。这些局外人认为，创造长期价值的关键是优化自由现金流。这种对现金的重视影响着他们运营公司的方方面面，如从支付收购费用和管理资产负债表的方式，到会计政策和薪酬体系等。

这种对现金的专心致志是局外人CEO们打破传统的基础，这通常与业内同行的方向背道而驰。他们始终如一地专注在少数几个变量上，这些变量塑造了每一家公司的战略。对20世纪七八十年代的辛格尔顿来说，它是股票回购；对约翰·马隆来说，它是对有线电视用户数量的不懈追求；对比尔·安德斯（Bill Anders）来说，它是剥离非核心业务；对巴菲特来说，它是保险浮存金的产生和运用。

他们共同世界观的核心是，相信所有CEO的首要目标都是追求长期每股价值的最大化，而不是组织规模的增长。这看起来似乎是一个明确的目标，然而在美国商界，有一种根深蒂固的冲动——做大企业。大企业能得到媒体更多的关注，它们的高管往往年薪不菲，并更有可能被邀请加入著名的董事会和俱乐部。因此，很少有企业会主动压缩自己的规模。然而，本书中几乎所有的CEO都通过股份回购来大幅缩减自己的股本基数。他们中的大多数还通过资产出售或分拆来缩减业务，对出售或关闭业绩不佳的部门也毫不手软。事实证明，企业规模的增长与股东价值的最大化往往并无关联。

1979年，辛格尔顿罕见地接受了《福布斯》杂志的采访，他的谈话展现了对现金务实的关注和与之相伴的打破传统的精神，语气中还透露出一丝强硬。辛格尔顿说：“在收购了一些企业之后，我们对企业进行了反思。我们得出的结论是，现金流就是关键.....我们对创造现金和管理资产的态度，来自我们自己的思考。”辛格尔顿还补充了一句，似乎他认为这个补充很有必要：“这一点思考绝非照搬他人。”⁴

特质**1** 效率决定生死
汤姆·墨菲和首府广播公司

The Outsiders

- 他是29年创造204倍收益的商界传奇，他是巴菲特的管理学导师。
- 他是伯克希尔-哈撒韦公司、德士古、强生、IBM公司董事，纽约大学基金受托人及荣誉副主席。
- 他为他的股东协商到了一个天价，足足190亿美元，是公司现金流的13.5倍、净收入的28倍。



他是

汤姆·墨菲

TOM MURPHY

沃伦·巴菲特

汤姆·墨菲和丹·伯克 (Dan Burke) 可能是管理领域最伟大的二人组，他们的组合前无古人，后无来者。

在商学院课堂上的演讲中，巴菲特经常将墨菲的公司（大都会通信公司，以下简称大都会）和哥伦比亚广播公司（CBS）之间的竞争，比作皮划艇和“伊丽莎白二世”邮轮之间横跨大西洋的竞赛，以说明管理对长期回报的巨大影响。

“内卷”，细嚼慢咽式兼并

1966年墨菲出任大都会CEO时，传奇人物比尔·佩利（Bill Paley）经营的哥伦比亚广播公司是美国媒体行业的霸主，它的电视台和广播电台占据了市场的最大份额，它还拥有收视率最高的广电网络，以及价值不菲的出版和音乐版权。相比之下，当时的大都会只有5家电视台和4家广播电台，占据的市场份额很小。哥伦比亚广播公司的市值是大都会的16倍。然而，30年后，当墨菲将公司出售给迪士尼时，大都会的市值是哥伦比亚广播公司的3倍。换句话说，皮划艇赢了，还赢得毫不含糊。

那么，这两家公司之间看似不可逾越的鸿沟是如何弥合的呢？答案在于完全不同的管理方法。

哥伦比亚广播公司在20世纪六七十年代的大部分时间里，通过其网络和广播业务获取了巨大的现金流，同时投资了激进的收购项目，包括收购一家玩具公司和纽约扬基棒球队，这让它进入了全新的领域。哥伦比亚广播公司发行股票，为一些收购项目提供资金，并在曼哈顿市中心斥巨资建造了一座地标性的办公大楼，建立了一个由42位总裁和副总裁组成的公司架构。以上这些总体印证了巴菲特的合伙人查理·芒格所说的一句话：“为了面子，丢了里子。”¹

佩利在哥伦比亚广播公司的战略与综合企业集团时代的传统认知是一致的，它支持“多元化”和“协同效应”那难以捉摸的好处，以证明收购不相关业务的合理性，并幻想这些业务一旦与母公司合并，会神奇地产生更多利润并抵御经济周期的影响。佩利的核心战略是将哥伦比亚广播公司的规模做大。

相反，墨菲的目标是让他的公司更有价值。正如他对我所说：“我们的目标不是拥有最长的火车，而是使用最少的燃料率先抵达车站。”²在墨菲和他的副手丹·伯克的领导下，大都会摒弃了多元化，他们创建了一个异乎寻常的新型综合企业集团，专注此前熟悉的媒体业务。墨菲收购了更多的广播电台和电视台，经营得顺风顺水。他还定期回购公司股票，最终收购了公司的竞争对手——美国广播公司（ABC），其结果不言而喻。

让墨菲超越佩利的“伊丽莎白二世”邮轮的公式看似简单：专注于具有诱人的经济特性的行业，偶尔利用杠杆有选择地购买大型资产，改善运营，偿还债务，周而复始。正如墨菲在接受《福布斯》采访时简明扼要地提到的：“我们只是不断地看准时机购买资产，明智地利用杠杆发展公司，改进运营，然后我们……就只等着坐享其成。”³然而有趣的是，其他媒体公司的同行并没有走这条路。相反，他们倾向于像哥伦比亚广播公司一样紧跟时尚，向不相关的业务领域延伸以发展多元化，建立庞大的员工队伍，并为无尽的媒体版权支付过高的费用。

墨菲领导下的大都会是一个非常成功的例子，我们现在称之为“上台阶”。典型的“上台阶”就是一家公司收购了一系列业务并试图改善其运营，然后继续收购。随着时间的推移，公司从规模优势和最佳管理实践中获利。这一概念在20世纪90年代中后期开始流行，并在21世纪初随着许多头部企业在过重的债务负担下倒闭而消失。这些企业的失败通常是因为它们收购得太快，但对于收购整合以及运营改善的难度与重要性都估计不足。

墨菲“上台阶”的方法不太一样。他稳扎稳打，让公司形成了真正的运营优势，对少数几次大型收购，他胸有成竹且全神贯注。在墨菲的领导下，大都会企业运营和资本配置的结合达到了不同寻常的卓越程度。正如墨菲告诉我的：“生意的诀窍就是将每天很多的小决定和几个大决定混在一起。”

专注战略，CEO的工作新重心

墨菲1925年出生于纽约的布鲁克林区，第二次世界大战期间曾在美国海军服役。得益于罗斯福总统签署的《退伍军人安置法案》（GI Bill）的规定，战后他就读并毕业于康奈尔大学，他是大名鼎鼎的哈佛商学院1949级的杰出成员，这一届的毕业生包括一位未来的美国证监会

主席、众多成功的企业家以及《财富》500强的CEO。从哈佛商学院毕业后，墨菲曾担任消费品包装巨头利华兄弟（Lever Brothers）的产品经理。老天爷和他开了个玩笑，1954年，滴酒不沾的他在父母位于纽约州斯克内克塔迪市（Schenectady）的家中参加一个夏季鸡尾酒会，他的生活从此发生了不可逆转的变化。墨菲的父亲是当地著名的法官，他还邀请了自己的老朋友弗兰克·史密斯（Frank Smith）来参加鸡尾酒会。弗兰克是著名广播记者洛威尔·托马斯（Lowell Thomas）手下的业务经理，拥有丰富的创业经验。

弗兰克见到墨菲后，立刻把他引到隔间，告诉墨菲自己最近启动的商业项目——WTEN，这是一家在奥尔巴尼市（Albany）苦苦挣扎的超高频电视台，弗兰克刚刚在破产危机中收购了它。该电视台坐落在一个废弃的修道院旧址内。夜幕降临之前，年轻的墨菲同意辞去他在纽约那份受人尊敬的工作，搬到奥尔巴尼市去管理电视台。墨菲既没有广播电台方面的工作经验，也没有任何相关的管理经验。

从一开始，弗兰克就在曼哈顿市中心的办公室管理公司，日常业务则由墨菲负责。在经历了几年的运营亏损之后，墨菲通过改进节目和积极的成本管理，把电视台变成了一个创造源源不断的现金的引擎。这个方法在此后几年被该公司反复运用。1957年，弗兰克和墨菲在北卡罗来纳州的罗利郡（Raleigh）买下第二家电视台，那里以前是一个疗养院。在罗得岛州的普罗维登斯（Providence）增设第三家电视台后，公司更名为大都会通信公司。

1961年，墨菲聘请了30岁的丹·伯克。虽然丹·伯克毕业于哈佛大学并拿到了工商管理硕士学位，但他此前也没有广播电视方面的经验，墨菲还是将自己在奥尔巴尼电视台的工作交给了丹·伯克。20世纪50年代末，丹·伯克最初是由他的哥哥吉姆·伯克（Jim Burke）介绍给墨菲的。吉姆·伯克是墨菲在哈佛商学院的同学，也是强生公司一位处于上升期的年轻高管，后来他成为强生公司的CEO，并因20世纪80年代中期对“泰诺危机”的处理而饱受赞誉。丹·伯克曾在军队中服役，然后进入哈佛商学院并于1955年毕业。之后丹·伯克加入了通用食品公司，担任果冻部门的产品经理。丹·伯克在1961年与大都会签约，墨菲很快向他灌输了精简机构和去中心化的经营理念。在这个领域，丹·伯克最终也成为一代宗师。

墨菲随后搬到纽约与弗兰克合作，通过收购来建构公司。在接下来的4年里，在弗兰克和墨菲的领导下，大都会精挑细选，收购了更多的

广播电台和电视台并发展壮大，直到1966年弗兰克去世。

弗兰克去世后，40岁的墨菲成为大都会的CEO，该公司上一年的收入仅为2 800万美元。墨菲上任的第一件事就是将丹·伯克提拔为公司总裁兼COO（首席运营官）。墨菲和丹·伯克的合作非常出色，他们的分工极其明确：墨菲负责收购和资本配置，以及偶尔与华尔街互动；丹·伯克负责日常运营管理。正如丹·伯克所说：“我们的关系建立在相互尊重的基础上。墨菲不感兴趣的事情，我正好有兴趣也愿意去做。”⁴丹·伯克的例子说明了本书中那些异常强大的COO所扮演的核心角色，他们对公司运营的密切关注，使他们的合作伙伴——那些CEO能够集中精力去应对长期战略和资本配置的问题。

一旦坐上CEO的位子，就没有什么能够阻止墨菲建功立业了。1967年，墨菲以2 200万美元收购了美国广播公司在休斯敦的分支机构KTRK，这是当时广电史上最大的一笔收购。1968年，墨菲以4 200万美元收购了贸易杂志的头部出版商仙童通信（Fairchild Communications）。1970年，墨菲以1.2亿美元从沃尔特·安娜伯格（Walter Annenberg）手中收购了广播公司三角洲通信（Triangle Communications），这是墨菲最大的一笔收购。这笔交易之后，大都会拥有5个甚高频电视台，达到了当时美国联邦通信委员会允许的上限。

接下来，墨菲将注意力转向了报纸出版业。作为广告驱动型行业，报纸出版业具有诱人的利润和强大的竞争壁垒，这与广电业非常相似。20世纪70年代初，在购买了几家小型日报后，墨菲在1974年以7 500万美元收购了《沃斯堡电讯报》（*Fort Worth Telegram*），1977年以9 500万美元收购了《堪萨斯城星报》（*Kansas City Star*）。1980年，为了寻找相关业务的其他增长渠道，墨菲以1.39亿美元收购了Cablecom公司，从而进入了刚刚起步的有线电视领域。

效率传导，“蛇吞象”背后的基础逻辑

在20世纪70年代中期至80年代初漫长的熊市期间，墨菲积极地购买了自己公司的股票，最终接近总股本的50%。当时，其买入的大部分股票的市盈率都是个位数。1984年，美国联邦通信委员会放宽了电视台所有权的管理规定。1986年1月，墨菲从朋友巴菲特那里获得融资，完成了他的得意之作，以近35亿美元收购了美国广播公司及其相关广电资

产，包括纽约、芝加哥、洛杉矶重要市场的电视台。

美国广播公司的交易是截至当时商业史上最大规模的非石油和天然气交易，对墨菲而言也是孤注一掷，收购价超过大都会市值的100%。这次收购震惊了传媒界，《华尔街日报》头条将其戏称为“小鱼吞鲸”。在交易结束时，丹·伯克对媒体投资人戈登·克劳福德（Gordon Crawford）说：“我穷其一生做的准备，就是为了等待这场收购。”⁵

这次交易的核心经济理由是，墨菲坚信他可以把美国广播公司的利润率（30%以下）提高到大都会所代表的行业领先水平（50%以上）。在丹·伯克的管理下，美国广播公司的管理团队从60人减少到8人，美国广播公司旗下电视台之一纽约WABC旗舰电视台的人数从600人减少到400人。仅用了两年时间，利润率的差距就弥合了。

精简机构，削减一切不必要的费用

在贯彻大都会精简机构、去中心化的方法上，丹·伯克和墨菲几乎没有浪费一分钟，他们立即削减不必要的额外补贴，比如高管专用电梯和私人餐厅，并迅速采取行动裁撤冗余的职位，在交易结束后的头几个月里裁掉了1 500名员工。他们还整合了办公室，出售了无用的不动产，比如通过出售曼哈顿闹市区的总部大楼就回笼了1.75亿美元。正如美国广播公司新闻部的鲍勃·泽尔尼克（Bob Zelnick）所说：“20世纪80年代中期以后，我们坐飞机就不能坐头等舱了。”⁶

从那个时候开始流传的一则故事说明了美国广播公司的CEO与崇尚精简、更具创业精神的收购者之间的文化冲突。不仅美国广播公司，事实上整个广播电视行业都崇尚豪华轿车文化，对于一个行业内的高管来说，最珍视的特权之一就是能坐着豪华轿车穿越几个街区去吃一顿午餐。然而和这些高管相比，墨菲一直搭乘出租车，从很早的时候起，他就乘出租出席美国广播公司的所有会议。不久，这一做法在美国广播公司高管队伍中激起涟漪，大都会的风气逐渐渗透到美国广播公司的文化中。当被问及这是否是一个以身作则的案例时，墨菲回答说：“难道还有别的说法吗？”

在并购交易之后的9年时间里，美国广播公司的每一条主要业务线，包括电视台、出版资产和娱乐与体育电视网（ESPN）的收入及现金流都大幅增长。即便是在收购时排名倒数第一位的广播电台，黄金时

段的收听率也攀升到第一，盈利能力超过了哥伦比亚广播公司和美国国家广播公司（NBC）。

在美国广播公司交易之后，大都会再没有进行过任何大规模收购，而是专注于整合、小规模收购和持续的股票回购。1993年，丹·伯克过完65岁生日之后就从小都会退休了，这让墨菲感到很惊讶。丹·伯克随后收购了波特兰海狗棒球队（Portland Sea Dogs），在他的监管之下，特许经营权再度重生。这支队伍现在是小联盟中最受尊敬的球队之一。

1995年夏天，艾伦公司（Allen & Company）在爱达荷州太阳谷举行年会，传媒业大亨相聚一堂，巴菲特建议墨菲与迪士尼CEO迈克尔·艾斯纳（Michael Eisner）坐下来谈一谈。墨菲当年70岁，没有合适的继任者，他同意与艾斯纳会面，后者曾表示有意收购大都会。几天后，墨菲为大都会的股东们谈妥了190亿美元的售价，相当于公司现金流的13.5倍，净利润的28倍。迪士尼董事会为墨菲保留一席，他随后从一线退休。

墨菲身后是一群欣喜若狂的股东。如果你在1966年墨菲担任CEO时向他投资1美元，那么当他把公司卖给迪士尼时，这1美元已经升值到204美元。在29年时间里，内部回报率达到惊人的19.9%，大大超过同期标准普尔500指数的回报率（10.1%）和传媒板块领先指数的回报率（13.2%）。巴菲特也从中获利颇丰，他的伯克希尔-哈撒韦公司在对大都会10年的持股期内年复合回报率超过20%。如图1-1所示，在墨菲任职的29年里，大都会的回报惊人地达到标准普尔500指数的17.7倍，达到同行水平的近4倍。



注：一篮子媒体股包括塔夫特通信（1966年9月—1986年4月）、都市传媒（1966年9月—1980年8月）、时代明镜（1966年8月—1995年1月）、考克斯通信（1966年9月—1985年8月）等。

图1-1 大都会的股票价格走势

The Outsiders

汤姆·墨菲的方法

本书的主题之一就是资源配置。

任何一位CEO都需要配置两种基本的资源：财务和人力。我们已经谈及了前者，然而后者也极为重要。那些局外人CEO对后者的配置也共享了一条非传统的路径，即强调扁平化的组织形式和公司员工的精简再精简。

去中心化，经营哲学的基石

局外人CEO们对去中心化有一种由衷的认同，即承认公司总部并不总是拥有所有问题的答案，许多真正的价值是当地具体业务部门的管理人员创造的。没有一家公司像大都会那样把去中心化放在企业价值观的核心位置。

在大都会每年年报封面的背页都有一段简短的文字：“去中心化是我们经营理念的基石。我们的目标是尽可能雇用最优秀的员工，并赋予他们履行工作所需的责任和权利。所有的决定都是分公司属地的业务部门做出的.....我们希望我们的经理.....永远注重成本，认准并开拓产品的销售潜力。”这段文字代表了大都会文化的特点，即运营经理拥有充分的自主权。

公司总部员工经过精简，其主要职能变为支持运营部门总经理的工作。没有市场营销、战略规划或人力资源等职能领域的副总裁，没有公司法律顾问，也没有公共关系部门。当时，墨菲的秘书负责处理所有来自媒体的电话。在大都会的企业文化中，出版商和电台管理者在内部拥有权力和威望，他们只要达到了各自的业绩目标，就几乎不会收到纽约总部的指示。这种环境是为那些独立自主、锐意进取的总经理设置的，同时环境也在促进他们的成长。公司在人力资源方面的指导理念，被墨菲无数次反复强调，即“雇用你能找到的最佳人选，然后放手让他们干”。丹·伯克曾经告诉我，公司去中心化的极致做法“降低了成本和积怨”。

在去中心化理念的发展过程中，丹·伯克自己就曾经充当过实

验的小白鼠。那是1961年，在接任WTEN总经理后，伯克按照他之前在通用食品公司接受的培训，开始每周向墨菲发送工作报告。好几个月都没有收到任何回复后，丹·伯克停止了发送，并意识到与其花时间向总部汇报，不如花时间在当地用心经营。正如丹·伯克描述他早年在奥尔巴尼市的经历时所说：“墨菲代表了无政府状态。”⁷

节俭、防范收益“肿块”

节俭也是大都会企业文化的核心。墨菲和丹·伯克很早就意识到，虽然你不能控制电视台的收入，但你可以控制成本。这一点已深植于公司文化之中，他们认为，作为广告驱动型商业，电视台收入不稳定是结构性的，最好的防御措施是时时刻刻对成本保持警惕。

事实上，在那些最早的、最常被提及的企业传闻中，有一个是这样的：墨菲甚至仔细审查了公司在涂料方面的支出。在墨菲抵达奥尔巴尼市后不久，弗兰克为了向广告商展示专业形象，请墨菲把工作室所在的破旧的修道院粉刷一番。墨菲立刻实施，但是仅仅粉刷了面向道路的两面，其他面保持原样（真正的“永远注意成本”）。这栋建筑物的照片至今仍挂在墨菲纽约的办公室里。

墨菲和丹·伯克认为，即使是最小的经营决策，特别是那些与员工人数有关的决策，都可能会产生无法预见的长期成本，因此需要持续关注。出版部门负责人菲尔·米克（Phil Meek）将这一点牢记在心，他的总部只有3个人，其中还有一名是行政助理。这3个人担负起全部的出版业务，包括6家日报、几家杂志集团和每周购物指南。

丹·伯克追求经济效益的热忱为他赢得了“红衣主教”的绰号。为了管理公司分散在各地的业务，丹·伯克制定了年度预算流程，其详细程度堪称传奇。每一位总经理每年都会到纽约参加预算扩大会议。在这些会议上，管理层提出了下一年的运营和资本预算，丹·伯克和他的首席财务官（CFO）罗恩·德夫勒（Ron Doerfler）则逐条详细讨论这些预算的细节。有趣的是，丹·伯克对砍掉个别招聘缺口和对砍掉过高成本的态度一样强硬。

预算会议绝不是走走过场，而是几乎总能带来实质性的变化，资本支出和各项费用在会议中得到特别关注。公司希望经理人的业绩超越同行，并时刻关注利润率，丹·伯克将其视为“某种成绩单”。在这些会议之外，经理们拥有绝对的自主权，有时几个月都不会接到总部的通知。

然而，公司并不是简单地抄近路奔着高回报率而去的，而是同样重视为了长期增长而进行的投资。墨菲和丹·伯克意识到，他们大部分业务盈利的关键驱动力是用户增长和广告市场份额的扩大，于是他们做好了对企业投资的准备，以确保在当地市场的领导地位。

例如，墨菲和丹·伯克很早就意识到，在当地新闻节目中排名第一的电视台，其广告收入的市场份额最终会大到夸张的地步。因此，大都会的电视台总是在新闻人才和技术方面投入巨资。引人注目的是，大都会的几乎每一家电视台都在当地市场处于领先地位。又如，由于丹·伯克意识到彩色印刷对保持《沃斯堡电讯报》竞争地位的重要性，他坚持花大笔资金来升级沃斯堡印刷厂，金额比米克所要求的还多。正如早期雇员菲尔·贝斯（Phil Beuth）所说：“公司花钱很谨慎，但绝不小气。”⁸

寻找“狐狸”

同时，大都会的招聘政策也不同寻常。墨菲和丹·伯克在加入大都会之前没有任何广播行业的经验，他们明显更看重员工的智力、能力和职业驱动力，而非直接的行业经验，他们在寻找有天赋的、有全新视角的“年轻狐狸”。当公司完成收购或进入新行业时，大都会会指派一位高层管理人员来管理这项新的资产，这位高管通常来自一个并不相关的部门。正是在这样的情况下，曾经经营着底特律旗舰电台WJR的比尔·詹姆斯（Bill James），被任命为有线电视部门的负责人；而以前是出版部门领导的约翰·赛厄斯（John Sias）则接手了美国广播公司电视网。两人都没有任何行业经验，最终却取得了非凡的业绩。

墨菲和丹·伯克也乐于把责任托付给年轻有为的经理人。正如墨菲向我描述的那样：“我们很幸运能拥有他们，并且知道他们可以胜任。”接手WJR电台时，詹姆斯35岁，没有电台工作经验；从

福特汽车公司转行执掌庞蒂亚克出版社（Pontiac Press）时，米克32岁，没有出版经验；从纽约搬到好莱坞负责美国广播公司的娱乐节目时，鲍勃·伊格（Bob Iger）37岁，他此前的职业领域是体育类广播节目。

大都会的员工流失率也低得出奇。罗伯特·普赖斯（Robert Price）是大都会广播业务的竞争对手，他曾经提到：“我们总能看到大量的求职简历，但其中从来没有来自大都会的员工。”⁹丹·伯克向我提及他与弗兰克曾讨论过的这种用人哲学的有效性，他记得弗兰克这样说：“体制让你拥有如此充分的自主权和自决权，以致你深陷其中无法自拔。”

收购，墨菲的地盘

在资本配置领域，墨菲的做法与同行截然不同。他摒弃多元化经营，支付最低股息，很少发行股票，积极利用债务杠杆，定期回购股票，经过很长的蛰伏期才偶尔进行一次大规模的收购。

大都会的两个主要资金来源是内部经营现金流和举债。正如我们看到的，大都会持续产出大量的、行业领先水平的经营现金流，为墨菲的资本配置提供了可靠的资金来源，让他可以进行收购、回购、偿债和其他投资选择。

墨菲也经常借助债务融资来筹集收购资金，他曾经这样总结自己的方法：“通常来说，我们付清款项后就获得了这些资产，然后再以这些资产为基础，获得杠杆资金去购买其他资产。”¹⁰在完成收购后，墨菲会积极配置自由现金流以降低债务水平，负债通常能提前还清。收购美国广播公司的大部分债务在交易后三年内就已经偿还。有趣的是，墨菲从不借钱进行股份回购，而更喜欢利用债务杠杆来收购正在运营的业务。

墨菲和丹·伯克极力避免增发股票对股权的稀释。除了出售股票给伯克希尔-哈撒韦公司以获得收购美国广播公司的资金以外，在和迪士尼公司交易之前的20年内，大都会都没有发行过新股。在这段时间内，由于反复回购，大都会的总股本缩水47%。

在墨菲的任期内，收购无疑是大都会最大的出路。最近的研究

显示，市面上大约2/3的收购行为，实际上有损于股东利益。那么，为什么大都会的收购却创造了如此巨大的价值呢？收购属于墨菲的职责范围，他将大部分时间都花在这个领域。他从不把收购决策委托他人，也从不聘请投资银行家。随着时间的推移，在那些显著并且意义重大的方面，墨菲总结发展出一种独特的方法，既有效又不同于竞争对手。

对墨菲而言，作为资本配置者，公司极致的去中心化策略有其重大好处：它使得公司盈利能力高于同行，在每一条业务线上，大都会的利润率都是最高的，这反过来又让墨菲在收购上占据了优势。墨菲能够买入资产，并且知道这些资产在丹·伯克的管理下很快就可以看到利润，从而变相地降低了实际的收购价格。换句话说，大都会的经营和专业领域的整合，有时候会在商业领域给墨菲带来罕见的商品：确定性。

在获得确定性后，墨菲便做好了积极行动的准备。在他的领导下，大都会的收购近乎激进，在三个不同的时期，分别完成了广播行业史上最大规模的交易，并在对美国广播公司的大型交易中达到顶峰。在这段时期内，除了广播、有线电视、杂志出版等行业的交易，大都会还参与收购了全美规模最大的几家报纸。

墨菲愿意为了一次诱人的收购而等待很久，他曾说：“我拿薪水不光是为了做成买卖，而是为了做好买卖。”¹¹然而，当看到自己喜欢的东西时，墨菲会毅然决然地押上重注。在近30年的CEO任期内，墨菲创造的价值在很大程度上是少数几次重大收购决定的结果，每一次都带来了良好的长期回报。每一次收购的成交金额都占到大都会市值的25%甚至更多。

对于交易前景的展望，墨菲堪称大师。他以风趣幽默和诚实正直而著称。与其他媒体公司的CEO不同的是，墨菲选择远离公众的视野，尽管在收购美国广播公司之后，离开公众视野在实际上变得更加困难。但这些特质有助于他发现潜在的收购机会。墨菲知道自己想买什么，他会花费数年时间与他中意的理想资产的所有者建立联系。墨菲从未参与过恶意并购，大都会完成的每一笔重大交易都是通过与卖方的直接联系而获得的，如三角洲通信的安娜伯格以及美国广播公司的伦纳德·戈登森（Leonard Goldenson）。

墨菲通过公平对待员工并认真经营，使那些被他收购的企业在

市场上始终处于领先地位，他努力工作让自己成为被优先考虑的买家。1984年，当墨菲与戈登森接洽收购美国广播公司时，他的好名声起到了积极的促进作用。墨菲以典型的自嘲风格开始他的演讲：“戈登森，请不要把我扔出窗外，即使我真的想收购你的公司。”

然而，这副和善友好的外表下，潜藏着一颗极其敏锐的商业头脑。墨菲是一位非常自律的买家，对回报有着严苛的要求，不会为收购放宽标准。他曾经因为500万美元的价差，放弃了一笔极大规模的报纸业收购，涉及三家得克萨斯州的公司。与本书中的其他局外人CEO一样，他在评估交易时依赖简单但十分有效的法则。对墨菲来说，这一基准是：企业在没有杠杆的情况下经营，10年内的税后回报率达到两位数。由于这种定价原则，尽管他参加了许多竞拍，但从未中标。墨菲告诉我，他的拍卖出价通常只是最终成交价的60%~70%。

墨菲有着一种不寻常的谈判风格，他认为“让卖家赚点儿”无可厚非，并觉得最好的交易是能让每个人高高兴兴地离开。墨菲经常先问卖家，他们认为自己的资产值多少钱，如果认为卖家出价合理，他就会接受。就像安娜伯格告诉他，三角洲通信的价值是税前利润的10倍一样。如果觉得他们的价格太高，墨菲会报出自己的最高出价，如果卖家拒绝了他的提议，墨菲就会走开。墨菲认为这种直截了当的方法既节省时间，又避免了双方不必要的相互挑刺儿。

股份回购是墨菲另一个重要的渠道，为其提供了重要的资本配置基准，多年来被他频繁使用。如果公司的市值低于企业全部私有化后的估值，墨菲就会在公开市场买回自己的股票。多年来，墨菲在回购上投入了超过18亿美元，大多数的交易发生之时，公司市值相较于现金流的倍数仅停留在个位数。总体而言，这些回购对公司来说是一场豪赌，资金规模仅次于和美国广播公司的交易。但它们为股东创造了丰厚的回报，19年来累计复合回报率为22.4%。正如墨菲今天所言：“我只希望当初多买一些。”

跨界吞并：墨菲的方法引爆出版业

1970年，在大都会完成和三角洲通信的交易后，美国联邦通信委员会禁止大都会拥有更多的电视台。因此，墨菲将注意力转向报

业。1974—1978年，墨菲发起了当时报业史上最大的两笔交易，收购了《沃斯堡电讯报》和《堪萨斯城星报》，并在美国各地购买了若干家规模较小的日报和周报。

公司报业出版部门的业绩犹如一块有趣的试金石，以检验其经营技巧。在吉姆·黑尔（Jim Hale）和米克的领导下，大都会运营电视台的经验在对报刊业务的管理上持续进化。这种方法强调谨慎的成本控制和最大限度地扩大广告市场份额。

引人注目的是，从公司4家主要报纸的经营情况来看，其收入和经营现金流持续逐年增长。尤其让人惊讶的是，这些于1997年卖给奈特·里德公司（Knight Ridder）的资产，在平均20年的持有期内，总体来说产生了25%的复合收益率。据《堪萨斯城星报》的出版商鲍勃·伍德沃思（Bob Woodworth，他随后担任了普利策公司的CEO）所说，作为公司发行量最大的报纸，《堪萨斯城星报》的营业利润率从20世纪70年代中期的个位数提升到1996年的35%，而现金流则从1 250万美元增长到6 800万美元。

大都会惊人的长期表现赢得了美国国内顶级媒体投资人的赞赏。巴菲特和马里奥·加贝利（Mario Gabelli）分别用他们各自时代的扬基棒球队和大都会做类比，用曾经传奇的击球手来类比墨菲和丹·伯克的管理业绩。巴菲特就好比贝比·鲁思和美国历史上的传奇棒球运动员卢·格里克（Lou Gehrig），加贝利就好比米奇·曼托

（Mickey Mantle）和罗杰·马里斯（Roger Maris）。戈登·克劳福德是美国最有影响力的媒体投资人之一，从1972年开始一直是《纽约时报》的股东，直到大都会和迪士尼完成交易，他是其最大的股东之一。克劳福德认为墨菲和丹·伯克实现了资本配置技能和企业运营技能的完美结合，造就了一台“产生回报的永动机”。¹²大都会的追捧者还包括鲁安坎尼夫（Ruane, Cunniff）公司的比尔·鲁安（Bill Ruane）、道富研究公司（State Street Research）的戴维·沃戈（David Wargo）。

引领全行业变革：大都会的企业文化和商业模式

20世纪90年代中后期，赛厄斯将大都会运营和人力资源方面的独特做法成功移植到西海岸一家名为年鉴（Chronicle）的出版集团上。赛厄斯是大都会出版事业部和美国广播公司新闻网的前负责

人，他于1993年担任年鉴出版集团的CEO，这是一家多元化经营的家族式企业，总部位于旧金山。

年鉴出版集团旗下有报纸《旧金山纪事报》（*San Francisco Chronicle*）、美国国家广播公司旧金山分部KRON、30万名有线电视用户以及一家图书出版公司。在赛厄斯到来之前，家族矛盾已经让公司四分五裂，公司运营也受到了影响。赛厄斯和他年轻的CFO艾伦·尼科尔斯（Alan Nichols）不失时机地采用了大都会的制度，彻底改变了公司的运营模式。他们立即辞退了公司总部整整一层楼的高管，制定了严格的预算流程，并赋予了总经理很大的自主权和自决权，这期间有许多人在新的、要求更高的企业文化中感到不适应，在一年内纷纷离职。

令人震惊的是，在这期间KRON的利润率提高了2 000个基点，从30%提高到50%，并最终在2000年6月以7.3亿美元的价格出售；《旧金山纪事报》通过与《旧金山观察家报》（*San Francisco Examiner*）建立的新型联营模式，营业利润率翻了一番多，从4%上升到10%，最终赫斯特集团（Hearst）在1999年以6.6亿美元的天文数字买下了这家报社。赛厄斯和尼科尔斯还以免税的方式将有线电视用户并入TCI，并以诱人的价格将图书部门出售给一位家族成员。在为公司股东创造了数亿美元的价值后，赛厄斯于1999年退休。

美国国家橄榄球联盟（NFL）中许多成功的教练曾经为比尔·沃尔什（Bill Walsh）工作过；很多成功的外科医生20世纪五六十年代曾在弗朗西斯·穆尔（Francis Moore）领导下的波士顿彼得·本特·布里格姆医院（Peter Bent Brigham Hospital）工作过。同样，传媒界到处都是大都会的门徒。大都会的文化和运营模式广受赞誉，除了年鉴出版集团的赛厄斯外，公司的前高管们还在各种媒体公司担任高层管理职位，数量多到令人目眩，其中首开先河的就是如今由伊格执掌的华特迪士尼。大都会的门徒占据的管理者岗位还包括林氏广播公司（LIN Broadcasting）CEO、普利策公司CEO、赫斯特集团公司CFO和斯克里普斯公司（E. W. Scripps）报业部门负责人等。丹·伯克的儿子史蒂夫曾任康卡斯特公司（Comcast）的COO，现任美国国家广播公司环球公司的CEO。

尽管本书的焦点是可量化的企业绩效，但更值得称道的是，墨菲把大都会缔造成了一家广受赞誉的公司，使其拥有异常强大的企

业文化和团队精神（至少有两个不同的高管群仍在定期聚会）。除了华尔街分析师外，大都会还受到员工、广告商和社会贤达的广泛尊重。米克给我讲过一个故事：在一次管理层团建活动中，他遇到一位调酒师，该调酒师通过在20世纪70年代初购买大都会股票获得了丰厚的回报。当一位高管后来问他为什么要进行这笔投资时，调酒师回答：“这些年来，我为很多公司的会务活动工作过，但大都会是唯一一家你不知道谁是老板的公司。”¹³

Transdigm：一个当代复制品

Transdigm可能是大都会当代的复制品，这家上市公司知名度不高，主营航空器配件。从1993年开始，通过内部增长和一次非常出色的收购，Transdigm的现金流复合增长率超过25%。与大都会一样，这家公司专注于具有特殊经济特征的具体业务。

就Transdigm而言，其专业领域是高度精细的航空零部件。这些零部件一旦组装到军用或商用飞机上，就无法轻易被替代，需要定期维护和更换，它们对飞机的性能至关重要，但其成本相对于飞机的总成本而言微不足道。因此，他们的客户，最大的军用和商用飞机制造商更注重产品性能而不是价格，于是该公司同时拥有了定价权和出色的盈余。按照税息折旧及摊销前利润（EBITDA）计算，其现金流利润率高达40%。

由CEO尼克·豪利（Nick Howley）领导的Transdigm管理团队，在20世纪90年代初就意识到了企业自身这些卓越的经济特征，为了优化这些专业零部件业务的盈利能力，形成了高度去中心化的公司结构和运营体系。豪利和大都会的墨菲一样，知道自己的团队将能够快速、显著地提高被收购公司的盈利能力，从而变相降低实际支付的收购价格，这都为未来的收购提供了令人信服的逻辑支撑。

自上市以来，Transdigm公司还采取了不寻常的、激进的资本配置策略，这一策略在华尔街引发了相当多的评论和困惑：一直保持高杠杆率，回购股票，在曾经一次金融危机巅峰时刻宣布发放高额的通过举债来融资的特别分红。毫无悬念，公司股东的回报也非常出色，自2006年公司首次公开发行股票以来，其股价已经翻了4倍。

特质2 聚焦资本配置

亨利·辛格尔顿和特利丹公司

The Outsiders

- 他用27年创造180倍收益，被誉为巴菲特的“孪生兄弟”。
- 他是世界级数学天才，美国电子行业早期领袖，苹果公司初始风险投资人之一。
- 在1972—1984年的8次独立的投标报价中，他令人震惊地买回了特利丹公司90%的流通股。在回购上，他总共花费了令人难以置信的25亿美元。



他是

亨利·辛格尔顿

HENRY SINGLETON

沃伦·巴菲特，
1980 年

亨利·辛格尔顿有着美国商界最好的企业运营和资本配置纪录……如果有人拿出 100 名顶尖商学院毕业生的综合成就，其成绩不会比辛格尔顿更好。

约翰·梅纳德·凯恩斯
(John Maynard Keynes)

我的想法随着事实变化而变化，你会怎么做呢？

1987年初，素以特立独行著称的中型综合企业集团特利丹宣布分红，这个看起来平淡无奇的事件引发了商业媒体的格外关注，《华尔街日报》甚至用头版文章加以报道。它们发现了什么具有新闻价值的东西呢？

在20世纪的大部分时间里，人们期待上市公司将其年度利润的一部分作为分红支付给股东。许多投资者（尤其是老年人）把这些分红当作收入，他们在做出投资决策时会密切关注公司分红的水平和政策。然而，在20世纪60年代的综合企业集团中，只有特利丹一家公司坚定地拒绝分红，它认为分红就要缴税，即分红资金需要在公司层面和个人层面纳两次税，因此降低了资金利用效率。

事实上，在深居简出的创始人兼CEO辛格尔顿的领导下，正如我们所见，这次分红决定只是特利丹一系列决策之一，这些决策非同寻常并和市场逆向而为。除了拒绝分红之外，辛格尔顿去中心化的经营也为世人所诟病，他回避与华尔街分析师沟通，也不分拆股票，他对公司股份的回购力度前无古人，后无来者。

所有这些做法都不同凡响甚至鹤立鸡群，但真正让辛格尔顿一骑绝尘并最终成为传奇的，是让市场和同行们相形见绌的回报率。从20世纪60年代不断上扬的牛市发端，到90年代初惨烈的熊市收尾，在近30年变化无常的宏观经济环境中，辛格尔顿让特利丹的市值以惊人的速度增长。

辛格尔顿通过不断地适应市场变化和牢牢地坚守资本配置来实现上述增长。辛格尔顿的方法与同行明显不同，这种反传统的根源可以追溯到他此前的背景。对于一个位列《财富》500强的公司的CEO来说，其

背景非常传奇.....

低买高卖，价格落差带来的收购良机

1916年，辛格尔顿出生于得克萨斯州的小城哈斯莱特（Haslet），他从未获得过MBA学位，却曾是一位成就卓著的数学家和科学家。他曾经就读于麻省理工学院，在那里获得了电气工程专业学士、硕士和博士学位。辛格尔顿为麻省理工学院的第一台学生电脑编程并把这件事写进博士论文，他在1939年作为全美数学尖子生获得了普特南奖章

（Putnam Medal）。普特南奖章后来的获奖者包括诺贝尔物理学奖得主理查德·费曼（Richard Feynman）。辛格尔顿还是一名狂热的国际象棋手，其积分比特级大师还高出100分。

1950年，辛格尔顿从麻省理工学院毕业后，先后在北美航空公司和休斯飞机公司（Hughes Aircraft）担任研究工程师。他后来被利顿工业公司的传奇少年特克斯·桑顿（Tex Thornton）招至麾下，并于20世纪50年代末在利顿工业公司开发了一种惯性制导系统，目前该系统仍在商用飞机和军用飞机上使用。辛格尔顿后来被提升为利顿工业公司电子系统集团的总经理，在他的领导下，到20世纪50年代末，该部门以超过8000万美元的营收成为公司最大的部门。

辛格尔顿很清楚他不会接任桑顿的CEO职位，于是在1960年离开利顿工业公司，那一年他43岁。他的同事、利顿工业公司电子元件组的负责人乔治·科兹梅茨基（George Kozmetzky）和他一同离开，并于1960年7月共同创建了特利丹公司。他们首先收购了3家小型电子公司，并以此为基础成功地中标了海军的一项大订单。特利丹公司于1961年上市，当时正值美国综合企业集团时代的黎明。

综合企业集团就是公司集团，拥有众多互不关联的业务部门，它在当时受到市场追捧，犹如后来的网络股。综合企业集团利用股价飙升的机会，贪婪且近乎盲目地收购五花八门的公司来实现业绩增长。这些收购项目最初带来了高额利润，并导致股票价格继续上涨进而收购更多公司。大多数综合企业集团相信自己能够在不同的下属企业间找到并建立协同效应，它们建立了员工众多的庞大总部，并积极讨好华尔街和商业媒体，以提振公司股票价格。然而，它们的太平盛世在20世纪60年代末戛然而止，当时最大的综合企业集团（美国国际电话电报公司、利顿工业公司等）盈利不达预期，股价随之暴跌。

今天的主流观点认为，综合企业集团是一种低效的企业组织形式，缺乏“单一业务”公司的灵活和专注。但是在20世纪60年代的大部分时间里，人们却并不这样认为。当时的综合企业集团享有高市盈率的估值，人们把高涨的股价当作通货，以此进行持续疯狂的收购。在这段头脑发热的时期，围绕收购的竞争远远弱于今天（那时私募股权公司还未出现），相比在股市买入对方股票进而取得控股权，综合企业集团支付的收购价格（以市盈率来衡量）往往低廉很多，这就为收购提供了令人信服的逻辑。

辛格尔顿充分利用了这一绝佳的套利机会，发展了多元化的业务组合。1961—1969年，他收购了130家公司，涉及航空电子、特种金属和保险等行业。除了两家公司外，其他所有公司都是特利丹在自身股价高企时完成收购的。

然而，辛格尔顿的收购方式不同于其他综合企业集团。辛格尔顿不会盲目收购，也极力避开那些陷入危机的公司，他将重点放在可盈利的、成长型的公司身上，这些公司往往在利基市场占据行业领先的地位。正如特利丹公司特种金属部门负责人杰克·汉密尔顿（Jack Hamilton）向我总结其所在部门的业务时所说：“我们专注于可以按千克销售的高利润产品，而不是那些按吨销售的产品。”¹辛格尔顿是谨慎而自律的买家，从不支付超过利润12倍的价格，他收购的大多数公司的市盈率都非常低。这与特利丹公司股票的高市盈率形成了对照。这一时期，特利丹的市盈率在20倍和50倍之间徘徊。

1967年，辛格尔顿完成了截至当时最大的一笔收购，他以4 300万美元收购了瓦斯科金属公司（Vasco Metals），并把其总裁乔治·罗伯茨（George Roberts）提拔为特利丹公司总裁，他自己则出任CEO和董事长。罗伯茨曾是辛格尔顿在美国海军学院的室友，他16岁时被美国海军学院录取，成为该校历史上最年轻的新生。后来教育经费和补助因为大萧条而削减，罗伯茨和辛格尔顿随之转学。罗伯茨也有理科背景，他毕业于卡内基-梅隆大学，获得冶金学博士学位，之后在多家特种金属公司担任各种行政职位，最终于20世纪60年代初加入瓦斯科金属公司并担任总裁。

自从罗伯茨加入特利丹公司，辛格尔顿就开始从公司经营中抽出身来，腾出大部分时间专注于公司战略和资本配置的问题。

集中每一分钱，“去中心化”模式的核心支柱

此后不久，辛格尔顿在综合企业集团中第一个叫停了收购。1969年中期，随着特利丹股票市盈率回落、收购目标价格上涨，辛格尔顿断然解散了收购团队。作为训练有素的买家，辛格尔顿意识到，随着市盈率的降低，特利丹公司股票的流通性在收购时不再有吸引力。从此以后，公司再也没有进行过重大收购，也再没有增发过股票。

从表2-1中可以看出这一收购策略的成效。作为上市公司，特利丹最初10年的每股收益（EPS）增长了惊人的64倍，而总股本增长不到14倍，这就为股东创造了丰厚的回报。

表2-1 特利丹公司前10年财务业绩(9)

项目	1961 年 (百万美元)	1971 年 (百万美元)	同比增长 (倍)
销售收入	4.508 0	1 101.90	243.40
净利润	0.058 1	32.30	553.90
每股收益	0.131 9	8.55	63.80
总股本	0.481 8	6.60	12.70
债务	5.220 0	151.00	27.90

注：每股收益和总股本根据股票分拆和分红进行调整。

资料来源：本表由长期关注特利丹公司的投资者汤姆·史密斯（Tom Smith）提供。

辛格尔顿出生的时候，市场对定量分析信心满满。20世纪四五十年代是“天才少年”的时代，一群出类拔萃的年轻数学家和工程师使用先进的统计分析方法，改造了一连串的美国标志性机构，从第二次世界大战中的陆军航空队开始，到20世纪50年代的福特汽车公司，不一而足。这一改造运动于1961年达到顶峰，当时五角大楼任命天才少年罗伯特·麦克纳马拉（Robert McNamara）担任国防部部长。

这些机构的权力都集中在其总部，那里有一批精锐部队，他们年轻、聪明异常，精通定量分析。这些高管实施集权化控制，在运营操作中推行新的、基于数学运算的系统。他们的分析才能广泛应用于高度离散的、混乱的运作，使其井然有序，他们试图通过施加指令来提升效率，无论是突袭轰炸还是建造工厂。

许多综合企业集团都采用了这种以总部为中心的方式来管理下属公司，并构建了庞大的员工队伍，其间充斥着众多的副总裁和企划部门。有趣的是，曾与最早的“天才少年”桑顿并肩作战的辛格尔顿，却为自己的公司设计了不同的模式。

与桑顿和美国国际电话电报公司的哈罗德·吉宁（Harold Geneen）等同行不同，辛格尔顿和罗伯茨避开了风行一时的企业“整合”和“协同”的概念，转而强调极致的去中心化。他们把公司分拆为最小的单元，将管理责任不断下放。一家员工总数超过4万人的公司，其总部员工不足50人，也没有人力资源、投资者关系和业务发展部门。具有讽刺意味的是，这个时代最成功的综合企业集团，在其运营中，却完全不像一家综合企业集团。

去中心化给特利丹公司带来了一种客观的、没有内斗的文化。几位公司的前总裁曾提到，不用钩心斗角的公司文化令人耳目一新，管理者只操心业务数据是否更好，数据糟糕就直接走人。其中一位说：“没有人关心辛格尔顿在和谁共进午餐。”

1969年，收购的引擎慢慢熄火，罗伯茨和辛格尔顿将注意力转向了公司现有的运营上。辛格尔顿的做法再一次颠覆了传统认知，他摒弃了当时华尔街的关键指标——账面盈余，转而以优化自由现金流为目标来运营公司。辛格尔顿和他的首席财务官杰里·杰尔姆（Jerry Jerome）设计了一个独特的指标，他们称之为“特利丹回报”：通过计算每个业务部门的人均现金流和净利润，强调现金的产出，并以此作为所有业务部门总经理奖金薪酬的基础。正如辛格尔顿曾经对《金融世界》杂志

（*Financial World*）说的那样：“如果有人想模仿特利丹，他们就应该接受这样的事实，即我们的季度收益会上下波动。我们的财务目标是实现现金流的最大化，而不是账面盈余的最大化。”²今天，在那些华尔街骄子和《财富》500强公司的CEO口中，你不可能再听到这样的话。

辛格尔顿和罗伯茨迅速提高了利润率，并大幅减少了特利丹公司的运营资本，这一过程释放了大量的现金。结果显而易见：特利丹的运营业务保持着较高的资产回报率，在20世纪七八十年代，公司的平均资产回报率在20%以上。巴菲特的合伙人查理·芒格将这些非凡的业绩形容为“任何人都难以企及……简直好到了荒谬的地步”。³

·帕卡德·贝尔公司（**Packard Bell**）：罕见的失误·

有一家下属公司没有达到辛格尔顿严格的标准，即帕卡德·贝尔公司，其业务是制造电视机。我们很有兴趣去看看辛格尔顿和罗伯茨是如何处理这个少有的、表现不佳的业务部门的。他们意识到，与成本较低的日本对手相比，帕卡德·贝尔公司的竞争劣势难以扭转。当帕卡德·贝尔公司不再获得满意的回报时，他们立即关闭了它。帕卡德·贝尔公司成为第一家退出该行业的美国制造商。在接下来的10年中，所有其他的美国制造商纷纷关门。

这些极具创造性的做法带来了最终结果：从1970年开始，特利丹公司在各种市场环境中都保持了非常稳定的盈利能力；涌入的现金被送到总部，由辛格尔顿进行配置。不足为奇的是，他在配置资本时所做的决定极其不同寻常，而且非常有效。

大规模回购，投资自己才是最靠谱的买卖

1972年初，由于现金余额不断增长，可收购对象价格畸高，辛格尔顿在曼哈顿市中心的电话亭给他的一个董事会成员打了电话，这个人是传奇的风险投资家阿瑟·罗克（Arthur Rock，他后来投了苹果公司和英特尔公司）。辛格尔顿开口便说：“罗克，我一直在想，我们的股票实在太便宜了。在这样的价位买入自己的股票，比做其他任何事情都能获得更好的回报。我想宣布一个股票的回购要约，你怎么看？”罗克沉思片刻，说：“我喜欢你的点子。”⁴

这几句话开启了资本配置史上一个最具开创性意义的时刻。自1972年开始的回购要约(10)持续了12年，辛格尔顿进行了一场史无前例的股票回购狂欢，对特利丹的股价产生了巨大的影响。他几乎单枪匹马就推翻了华尔街长期秉持的信条。

如果仅仅说辛格尔顿是股票回购的先驱，那就大大低估这个案例的意义。也许更准确的说法是，辛格尔顿是回购领域的贝比·鲁思——在企业金融学分支的早期历史中，辛格尔顿是如奥林匹亚众神一般高耸入云的人物。在20世纪70年代之前，股票回购并不常见且存在争议。传统观点认为，回购意味着企业内部缺乏投资机会，因此被华尔街视为公司疲软的表现。辛格尔顿无视这一教条，1972—1984年，他在8次回购要约中购入了特利丹公司总股本90%的股票。正如查理·芒格所说：“从未有人如此凶猛地回购股票。”⁵

在辛格尔顿任期内的大部分时间里，税率都很高。辛格尔顿认为，与分红相比，回购以高效避税的方式把资金还给了股东。他还认为，用低价回购股票是一种自我催化，类似于上紧一个发条，在未来某个时刻，它会突然发力，实现全部价值。换句话说，这个过程中会产生出非凡的回报。这些回购提供了一个有用的资本配置基准。相对于其他投资机会，每当购买自家股票的回报看起来更有吸引力时，辛格尔顿就会出价回购。

回购在20世纪90年代开始流行，近年来，CEO们经常使用回购来提振低迷的股价。然而，回购只有在价格合适的情况下才能增加股东价值。不出所料，辛格尔顿买得漂亮，历次回购为特利丹公司股东创造了令人难以置信的42%的年复合回报率。

这些回购要约几乎都被超额认售。辛格尔顿做过分析，知道这些回购理由充分。带着这种坚信，他总是将回购份额用尽。对特利丹公司来说，这些回购下了重注：宣布回购计划时，其规模仅是公司账面价值的4%，后来实际回购的规模令人震惊地达到公司账面价值的66%。辛格尔顿总共花费了25亿美元进行回购。表2-2对这一成就进行了展示。1971—1984年，辛格尔顿在特利丹公司收入和净利润持续增长时，于市盈率处于低点时回购了大量股票，导致每股收益惊人地增长了40倍。

表2-2 特利丹公司股票回购计划的结果

项目	1971 年 (百万美元)	1984 年 (百万美元)	同比增长 (倍)
销售收入	1 101.90	3 494.30	2.20
净利润	32.30	260.70	7.10
每股收益	8.55	353.34	40.30
总股本	6.60	0.90	-0.90
债务	151.00	1 072.70	6.10

注：每股收益和总股本根据股票分拆和分红进行调整。

资料来源：本表由长期关注特利丹公司的投资者汤姆·史密斯提供。

然而必须认识到，这种对回购的痴迷代表辛格尔顿观念上的演变。在职业生涯早期创建特利丹公司时，他曾是活跃而高效的股票发行者。伟大的投资者和资本配置者必须能高卖低买：特利丹公司股票发行的平

均市盈率超过25倍；相比之下，辛格尔顿回购的平均市盈率倍数不到8倍。

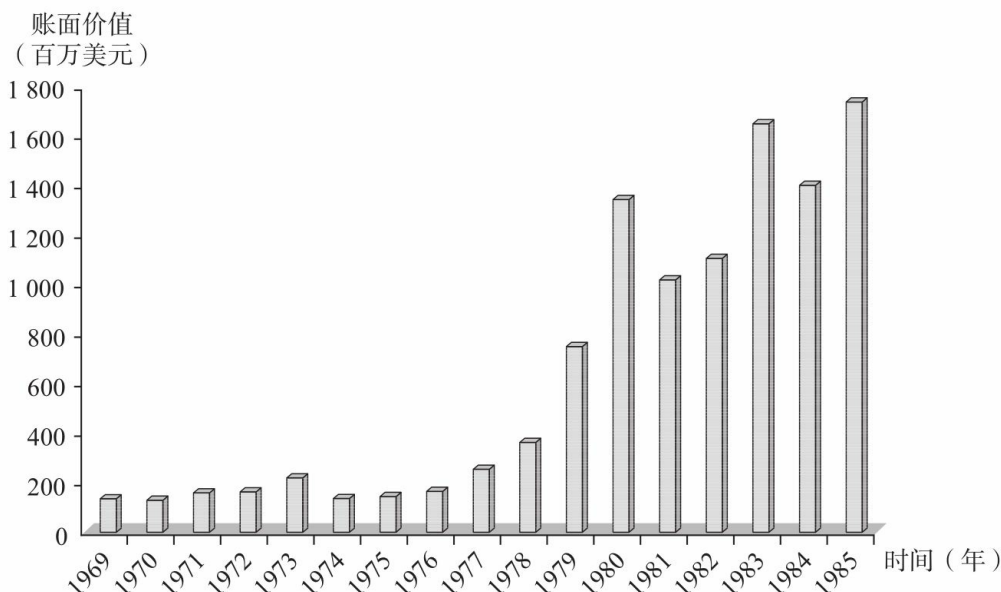
辛格尔顿从青少年时期就对股市着迷。罗伯茨曾经向我透露：第二次世界大战期间，在纽约度假的辛格尔顿在一家证券公司的橱窗前盯着代表股票价格的纸带从报价机上闪过，一站就是好几个小时。

20世纪70年代中期，辛格尔顿终于有机会实现这一终生的梦想。他在熊市最严重的时候直接负责特利丹保险子公司的股票投资组合，当时市盈率处于大萧条以来熊市的最低点。在投资组合的管理方面，与收购、运营和回购一样，辛格尔顿独辟蹊径，成果显著。

在一次重要的逆势操作中，辛格尔顿大胆地重新配置了这些保险投资组合中的资产，使股票配置的占比从1975年的10%增加到1981年的77%。辛格尔顿对投资组合的具体调整更不寻常，颇具戏剧性。股票组合中超过70%的资金仅仅投资于5家公司，其中，对他的前雇主利顿工业公司的投资就不可思议地占到全部资金的25%。这种非同寻常的集中投资（通常一只共同基金拥有100多只股票）引起了华尔街一片哗然，许多观察人士认为辛格尔顿正在为新一轮收购做准备。

辛格尔顿其实并无此意，然而更近距离地观察他是如何管理这些投资组合的，对我们会有所启发。辛格尔顿重仓的无一例外都是他非常了解的公司，包括柯蒂斯·赖特（Curtiss Wright）这样较小的综合企业集团以及如德士古（Texaco）和美国安泰（Aetna）等大型能源和保险公司。这些公司的市盈率在辛格尔顿投资时处于或接近历史低点。正如查理·芒格在谈到辛格尔顿的投资方法时所说：“就像巴菲特和我一样，他喜欢集中持股，并且只买入几家他非常了解的公司的股票。”⁶

与回购特利丹公司股票一样，辛格尔顿在这些保险投资组合中的回报也非常出色。这些回报的指标可以在图2-1中看到，图中显示，1975—1985年，当辛格尔顿开始对他的公司分拆上市时，特利丹保险子公司的账面价值大约增长了8倍。



注：图中数据是特利丹保险子公司 Unitrin 和 Argonaut 股票账面价值的总和。

图2-1 特利丹保险子公司账面价值

价值释放，通过分拆深挖公司潜能

1984—1996年，在特利丹运营部门业绩停滞不前的情况下，辛格尔顿将工作重点从投资组合管理转向管理权移交（1986年，他聘请罗伯茨接替他担任CEO，自己保留了董事长头衔），以及优化股东价值。为了实现这些目标，辛格尔顿采取了新的策略，再次震惊了华尔街。

辛格尔顿是对公司进行分拆上市的前驱，他相信通过分拆可以降低公司的复杂性，进而简化特利丹公司继任者面对的问题，并为股东释放公司大型保险业务的全部价值。用长期担任董事会成员的法耶兹·萨罗菲姆（Fayez Sarofim）的话来说，辛格尔顿认为“联合有其时，分拆亦有其时”。⁷ 1986年，分拆特利丹公司的时间终于到来，首秀是Argonaut公司的上市，它主要负责公司员工的薪酬保险。

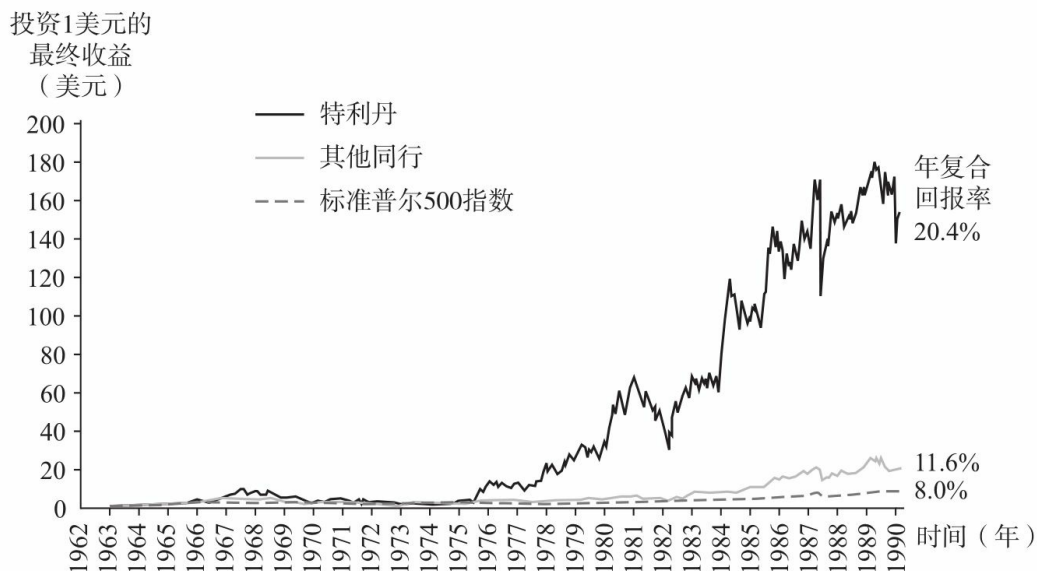
接下来，在1990年，辛格尔顿剥离了公司最大的保险子公司Unitrin，当时它的CEO是杰尔姆。考虑到子公司Unitrin的价值占据了当时特利丹公司价值的很大部分，这个动作就显得意义重大。Unitrin上市以后，在杰尔姆和他的继任者迪克·维（Dick Vie）的领导下，获得了优异的回报。

从20世纪80年代中后期开始，由于能源和特种金属市场的周期性低

迷及其军工业务受到欺诈指控，特利丹公司的非保险业务增速放缓。1987年，当收购价和股价（包括辛格尔顿所拥有的自己公司的股票）都处于历史高位时，辛格尔顿得出结论：他没有更好的、回报更高的选择来配置公司的现金流，并宣布公司上市26年来的首次分红。对那些长期关注特利丹公司的人来说，这是一则爆炸性的消息，标志着公司由此进入新的历史阶段。

在这些分拆上市取得成功以及罗伯茨担任CEO之后，辛格尔顿于1991年从董事长的职位上退休，专注于经营他那广袤的牧场。辛格尔顿和他同代出生于得克萨斯州的成功企业家一样，对牧场情有独钟，他后来在新墨西哥州、亚利桑那州和加利福尼亚州共买下100多万英亩(11)的牧场。然而1996年辛格尔顿复出了，亲自就特利丹公司留存的制造业务与阿勒格尼工业公司（Allegheny Industries）的合并进行了谈判，并阻止了“野蛮人”贝内特·勒博（Bennett LeBow）的恶意收购。据当时特利丹公司的总裁比尔·鲁特里奇（Bill Rutledge）说，在这次谈判中，辛格尔顿只专注于获得尽可能的高价，而忽略如管理头衔和董事会组成等其他边缘问题。⁸辛格尔顿再一次为特利丹公司的股东赢得有利结果：比公司之前的交易价格高出30%。

辛格尔顿保持着一个非凡的投资纪录，使他的同行和市场相形见绌。从1963年（这是我们掌握其可靠股票数据的最早时间）到1990年辛格尔顿辞去董事长一职，他为股东带来了20.4%卓越的年复合回报率（包括多次分拆上市）。相比之下，同期标准普尔500指数的回报率为8.0%，其他大型综合企业集团股票的回报率为11.6%（见图2-2）。



注：纵坐标根据股票分拆、转增和现金分红进行了调整；其他同行包括利顿工业公司、美国国际电话电报公司、海湾西部公司 (Gulf & Western) 和德事隆集团 (Textron)。

图2-2 辛格尔顿时代，特利丹股票与标准普尔500指数、其他同行的比较

1963年投资于辛格尔顿的1美元，到1990年价值达到180.94美元，比同行高出近9倍，超过标准普尔500指数12倍，而杰克·韦尔奇在辛格尔顿高速列车后视镜里成了一个遥远的小点儿。

The Outsiders

亨利·辛格尔顿的方法

任何一位CEO最重要的决定之一就是如何分配时间，具体来说，是他在下面三个非常重要的领域分配多少时间：运营管理、资本配置和投资者关系。毫不奇怪的是，辛格尔顿的时间管理方法与桑顿和吉宁等同行截然不同，却与其他局外人CEO异常相似。

去计划，做思想上的特立独行者

正如1978年辛格尔顿在接受美国《金融世界》杂志采访时所说：“我不让自己卷入每日例行的事务，所以就不会陷入任何特定的窠臼，我也不用任何精确的术语来定义自己的工作，而是可以在任何时候自由地做任何看起来符合公司最大利益的事情。”⁹辛格尔顿拒绝事无巨细的计划，他更愿意保持灵活性和开放性的备选方案。正如他曾经在特利丹公司年会上解释的那样：“我知道很多人都有非常坚定和明确的计划，他们觉得自己已经考虑了各种各样的可能性，但是我们要承受大量外部因素的干扰，其中绝大多数都无法预测。因此，我的想法是随机而动。”¹⁰在为数不多的面对《商业周刊》记者的一次采访中，他解释得更加简明扼要：“我唯一的计划就是继续上班……我喜欢每天掌舵去调整航向，而不是提前去规划未来。”¹¹

与其他综合企业集团的高管不同，比如桑顿、吉宁或海湾西部公司个性鲜明的查尔斯·布卢多恩（Charles Bluhdorn），辛格尔顿不会去讨好华尔街的分析师和商业媒体。事实上，辛格尔顿认为维护投资者关系是对时间的低效利用，所以他干脆拒绝提供季报，也不出席行业会议。当时，那些随和的同行高管经常在顶级商业杂志的封面亮相，辛格尔顿的举动显然不合潮流。

·特利丹公司与《萨班斯-奥克斯利法案》·

特利丹公司打破传统的做法是企业管理领域持续至今的热门话题。按照《萨班斯-奥克斯利法案》（*Sarbanes-Oxley*）的现行标准，特利丹的董事会是不及格的。辛格尔顿和本书中其他局外人

CEO一样，是小型董事会的支持者。特利丹的董事会只有6名董事，其中包括辛格尔顿，一半是内部人士。然而，这是一个才华横溢的团体，每个成员的身家都和公司的经济利益绑在一起。除了辛格尔顿、罗伯茨和科兹梅茨基（他1966年从特利丹公司退休后，负责管理得克萨斯大学商学院）之外，董事会成员还包括辛格尔顿的麻省理工学院同学、信息理论之父克劳德·香农（Claude Shannon）；传奇的风险投资家阿瑟·罗克；以及休斯敦基金的管理人、亿万富翁法耶兹·萨罗菲姆。截至这一时期末，董事会成员共持有公司近40%的股份。

即使在一本充斥着CEO如何积极回购股票的书中，辛格尔顿的方法也能自成体系。考虑到他对特利丹公司股票的胃口之大，以及与其他局外人CEO相比，其回购率之高，我们很有必要深入观察辛格尔顿的回购方法。他的方法与当今大多数CEO的做法有很大不同。

辛格尔顿式回购

从根本上讲，回购股票有两种基本方法。当代最常见的方法是，公司批准一定数量的资金（通常是其资产负债表上相对较小比例的多余现金）用于回购股票，然后在一个季度（有时是几年）内逐步在公开市场上购买股票。这种方法谨慎保守，当然也不太可能对股票的长期价值产生任何有意义的影响。让我们把这种谨慎、有条不紊的方法称为“小酌”。

另一种方法则大胆得多，它也是本书中的CEO们喜欢的由辛格尔顿开创的方法。这种回购方法的特点是：频率较低且在股价低迷时大规模回购，通常在非常短的时间内完成，往往使用要约的方式，有时还通过债务融资的方式来收购。辛格尔顿至少8次采用这种方法，他鄙视“小酌”，而更喜欢“牛饮”。

1980年，辛格尔顿的股票回购就是他敏锐把握资本配置机会的绝佳范例。同年5月，由于特利丹公司的市盈率接近历史低点，辛格尔顿发起了该公司截至当时最大规模的收购要约，结果要约被超额认售了3倍。辛格尔顿随即决定购买所有被认售的股票（超过流通股总额的20%）。考虑到公司自由现金流充沛并且当时利率下行，辛格尔顿用固定利率的债券融资为整个回购提供资金。

回购之后，利率大幅上升，新发行债券价格下跌。辛格尔顿认为利率不可能继续上升，所以他开始回购这些债券。不过，他是用公司养老基金的现金赎回的这些债券，以避免缴纳投资收益税。

在这一系列复杂的交易之后，特利丹公司成功地用廉价的融资为大规模的股票回购提供了资金。之后，随着利率下行，养老基金在债券的投资上又实现了可观的免税收益。哦，对了……特利丹公司股票也实现了大幅升值，10年复合回报率超过40%。

如果每家公司都这么做，那么肯定会出问题

终其一生，辛格尔顿思想上强烈的独立意识成为他的显著特征。辛格尔顿82岁时死于脑癌。在其去世前两年的1997年，辛格尔顿与特利丹公司的长期投资人莱昂·库珀曼（Leon Cooperman）进行了一次座谈。库珀曼问辛格尔顿怎么看待当时很多《财富》500强企业都宣布大规模回购股票的事。辛格尔顿很有预见性地回答：“如果每家公司都这么做，那么肯定会出问题。”¹²

巴菲特和辛格尔顿：失散的孪生兄弟？

巴菲特管理伯克希尔-哈撒韦公司有一套独特方法。对于巴菲特的那些真知灼见，辛格尔顿早在对特利丹公司的管理中就已采用了。事实上，辛格尔顿可以被看作是巴菲特的原型。这两位杰出的CEO之间有着不可思议的相似之处，如下所示：

- 作为投资者的CEO。巴菲特和辛格尔顿设计的公司架构，使他们能够专注于资本配置，而不是企业运营。他们认为自己主要是投资者，而不是管理者。

- 经营要去中心化，投资决定要集中化。用员工极少的总部管理着高度去中心化的公司组织，运营层和管理层间几乎没有（如果有的话就尽量减少）中间阶层。他们二人决定各自公司所有的重大资本配置决策。

- 投资理念。巴菲特和辛格尔顿都将投资重点放在他们熟悉的行业，并喜欢在投资组合中集中持有上市公司证券。

- 处理投资者关系的方法。他们都没有向分析师提供季报，也不出席各种会议，但是他们都提供了包含详尽业务部门信息的年报。

- 分红。在综合企业集团中，特利丹公司只在第26年支付过一次股息，而伯克希尔-哈撒韦公司从未分红。[\(12\)](#)

- 股票分拆。特利丹公司是20世纪七八十年代纽交所发行股票中价格最高的，巴菲特从未分拆过伯克希尔-哈撒韦公司的A股（2011年股价超过12万美元）。

- CEO持股非常重要。辛格尔顿和巴菲特在各自的公司都拥有大量的股权（辛格尔顿持股13%，巴菲特持股30%以上）。他们像主人一样思考，因为他们自己就是主人。

- 保险子公司。辛格尔顿和巴菲特都认识到投资保险公司的“浮存金”在创造股东价值方面的潜力。对这两家公司来说，保险是最大的，也是最重要的业务。

- 关于餐厅的类比。著名的投资者菲利普·费雪（Phil Fisher）曾通过一系列政策和决策将公司与餐厅进行类比。每家餐厅的风味、价格和氛围都不同，也自然会吸引到不同的客户前来就餐。按照这个标准，巴菲特和辛格尔顿都有意经营着不同寻常的餐厅，随着时间的推移，这些餐厅吸引了志同道合、目光长远的客户（股东）。

特质**3** 绑定股东利益
比尔·安德斯和通用动力公司

The Outsiders

- 他是全球唯一漫步月球表面的CEO，美国核能管理委员会第一任主席。
- 3年时间，他就从负债6亿美元变为手握50亿美元现金。
- 入职第一年，他替换了公司全部25名高管中的21名，在头两年裁掉了60%的员工（工厂裁员人数占比则达到了80%），卖掉了集团旗下一半的公司和业务部门。



他是

比尔·安德斯

BILL ANDERS

爱默生

目光短浅的哥布林^①只会愚蠢地
一以贯之。

①哥布林是西方神话中的一种小精灵。——译者注

30年来，柏林墙一直作为冷战时期紧张与焦虑的国际象征，它终于在1989年倒塌了。紧随其后，美国军工行业长期以来的商业模式也因此崩塌。军工行业传统上依赖于诸如导弹、轰炸机等大型武器系统的销售，它是第二次世界大战后美国军事战略的核心。似乎一夜之间，美国军工行业就陷入了混乱。那些行业的领军高管曾一直被卸任的将军和海军将领视为亲密兄弟，而当时的这些公司高管却要争相寻找新的伙伴。柏林墙倒塌不到6个月，上市的头部军工企业景气指数就下跌了40%，其中一家公司的处境似乎尤其糟糕。

这家公司就是通用动力（General Dynamics），它曾是军工行业的先锋。其前身可以追溯到19世纪末，有着向五角大楼出售大型武器的漫长历史，包括在第二次世界大战期间出售传奇的B-29轰炸机和现代空军的F-16主力战机；作为潜艇的主要制造商，出售船舶；以及作为坦克和其他作战车辆的主要供应商，出售陆地车辆。多年来，通用动力已经在导弹、航天系统、民用飞机和建筑材料等非国防业务领域实现了多元化发展。20世纪80年代，美国联邦调查员发现该公司高管滥用公司飞机和其他额外收入，通用动力的发展因此丑闻而备受打击。

1986年，通用动力公司引进了在五角大楼享有盛誉的新任CEO斯坦·佩斯（Stan Pace）。佩斯改善了通用动力公司与美国参谋长联席会议的关系，但公司业务仍然停滞不前。屋漏偏逢连夜雨，1990年，该公司获悉其最大的新飞机项目A-12可能被取消。1991年1月，当另一位新CEO上任时，通用动力公司负债达6亿美元，现金流入不敷出，并被外界猜测可能面临破产。通用动力公司年收入曾达到过100亿美元，而当时的市值仅为10亿美元。用高盛集团军工防务分析师朱迪·博林杰（Judy Bollinger）的话来说，该公司“差到不能再差”，是衰退的行业中处境最糟的公司。¹

换句话说，这是一次转型。陷入财务困境的公司通常会聘请重组“顾问”。这些顾问从天而降，大幅削减成本，与贷款机构和供应商谈判，并希望尽快出售公司，好赶去处理下一单业务。这些被雇来的“干将”往往忽视了对公司文化、资本配置和组织架构的长期考虑，只是把重点放在短期的现金需求上。重组结局通常是将困境中的业务出售给一

家大公司，如此的转型往往能带来诱人的眼前利益。这一过程就好像捡起地上的烟头，猛嘬最后几口。

一家进入转型的企业，历经多任CEO，在长时期内仍能维持高回报的情况非常罕见，但这恰恰发生在通用动力公司身上。通用动力公司的故事表明，即使在整个行业严重混乱的局面下，那些局外人CEO所采用方法的核心要素依然有效。

资产剥离，卖掉低效业务的三个标准

这一切都始于1991年1月比尔·安德斯执掌通用动力公司，当时正值20世纪90年代初，军工市场在海湾战争后陷入极度低迷的熊市。安德斯可绝对不是普通的CEO。在加入通用动力之前，以主流观点来看，安德斯有着杰出的职业生涯：1955年，他从海军学院毕业，获得电气工程学士学位，并在冷战期间担任空军战斗机飞行员；1963年，安德斯获得了核工程领域的研究生学位，并从数千人中脱颖而出，入选仅有14名成员的美国国家航空航天局精英宇航员团队。

作为1968年阿波罗8号任务中的登月舱驾驶员，安德斯拍摄了那张当代经典照片《地球升起》（*Earthrise*），照片最终刊登在《时代》、《生活》和《美国摄影》等杂志封面上。一位著名的防务分析师认为，这些早年的成就赋予安德斯在今后的风浪中承担风险的能力：“绕月飞行之后，寻常的业务问题就不再会成为他的困扰。”

安德斯以少将的军衔离开了美国国家航空航天局，然后短暂出任美国驻挪威大使，之后被任命为核能管理委员会（Nuclear Regulatory Commission）第一任主席。以上这些都发生在安德斯45岁之前。安德斯在五角大楼享有盛名且受人尊重，离开政府部门后，他加入了通用电气，在那里他接受了通用电气管理方法的培训，是与杰克·韦尔奇同时代的人。正如安德斯所说：“那里有一大群了不起的经理人，他们都是优秀的游泳教练……尽管偶尔他们也想淹死你。”²

1984年，安德斯终于被聘为领先的综合企业集团德事隆集团的商业运营负责人，然而这段经历却令他感到沮丧。安德斯个性独立，不随声附和，有直截了当的沟通风格。安德斯不能接受碌碌无为的业务部门和官僚主义作风的结合，所以毫不意外，他与德事隆集团的时任CEO发生了冲突。

1989年，在一次行业协会会议上，安德斯遇到了通用动力公司的一位高管，对方承诺如果安德斯加入通用动力公司担任一年副董事长，就有机会接任CEO。安德斯欣然受命。在一年过渡期里，安德斯花时间了解公司的业务和文化，并在贝恩公司的协助下，研究了席卷整个国防工业的巨大变局，结论是军费高昂开支的时代似乎戛然而止。这一年的学习使安德斯在正式被任命为CEO后旗开得胜。

虽然安德斯是本书中年龄最长的CEO，也是唯一一位50多岁时才掌舵公司的CEO，且在接受通用动力公司这一职位时，只有10年的私营公司工作经验，但他的视角仍然新颖独到。

由工程师和退役的军官管理的军工行业，历来有点俱乐部或联谊会的感觉。安德斯既是一名工程师，又曾是一名将军，天时地利人和，他就是用扫帚清除行业蛛网的那个人，他的看法以及随后的行动将动摇军工界的核心层，把他们赶出舒适区。

安德斯为通用动力公司制定的转型战略根植于一个核心的战略洞见：随着冷战的结束，军工行业的产能严重过剩。因此，安德斯认为，行业参与者需要积极行动，要么收缩业务，要么通过收购来实现增长。在这种新的环境中，行业整合者和被整合的对象比比皆是，公司需要迅速判断自己属于哪个阵营。安德斯在其最初的年度和季度报告中勾勒了他的战略，并开始积极实施。

这一战略基于三个关键原则：

- 安德斯从通用电气的前同事杰克·韦尔奇那里学来一招，即只有在确保能处于市场第一或第二的位置时，通用动力才会进入该业务领域。这与同时代的鲍威尔主义惊人地相似，后者要求美国只参与能够取得决定性胜利的军事冲突。
- 该公司将退出那些回报低到无法接受的商品业务。
- 坚守自己所熟悉的业务领域。具体而言，军转民领域似乎是军工企业新的利润点，但对于这尊“长久以来令军工人垂涎的圣杯”要保持警惕。

通用动力公司将退出不符合以上战略原则的业务领域。

此外，安德斯认为，通用动力公司需要一场剧烈的企业文化变革。在担任通用动力公司CEO之前，通过与公司高管的深入访谈，安德斯发现了一种根深蒂固的工程师思维定式，这种思维定式始终围绕“更大、更快、更致命”的武器研发，却很少关心股东利益，这与通用电气形成了鲜明的对比。安德斯积极地纠正了这一点，不断强调对股东利益和净资产收益率等指标的重视。

安德斯还认为，经营活动需要大幅精简，以提高回报。为了达成这一目标，安德斯需要一个新团队，于是他立刻着手组建。第一步是把吉姆·梅勒（Jim Mellor）提升为总裁兼COO。梅勒负责通用动力公司的造船业务时，曾取得优异的业绩。安德斯对我说：“梅勒是这种人，他会追踪最后一分钱的去向，并追究相关责任人。”1991年上半年，安德斯和梅勒二人联手，换掉了公司25名高管中的21名。

除了新的运营人才外，安德斯还聘请了金融奇才哈维·卡普尼克（Harvey Kapnick）担任副董事长，与转型相关的各种法律和战略工作则由一位才华横溢的律师尼克·查布拉贾（Nick Chabraja）来协助完成。在团队各就各位后，安德斯就马不停蹄地实施了一项非同寻常的重组。

现金的生成与配置

安德斯任职通用动力公司短暂的3年可以分为两个基本阶段：现金的生成和配置。每一个阶段，公司的做法都非常独特。

我们从现金的生成开始讲起。当安德斯和梅勒开始实施他们的计划时，通用动力公司的情况是，过度使用融资工具且现金流入不敷出。然而在接下来的3年里，公司获得了50亿美元的现金流。这一惊人的资金流主要来自两个方面：一是大幅缩减业务；二是出售安德斯战略框架下那些无关紧要的业务。

在经营方面，安德斯和梅勒发现了公司的遗留问题，即在存货、资本设备和研发方面大量过度投资。他们一起迅速采取行动，将过分的做法排除在框架之外。在参观一家F-16工厂时，四处巡视后，他们发现有大量昂贵的F-16座舱盖（驾驶舱的透明玻璃罩）闲置，但是这家工厂每周只能生产一架飞机。于是梅勒制定了新的规定：工厂最多保留两块座舱盖。在邻近的两家坦克工厂，他们发现了重复购置的两套价格不菲的机械设备都未得到充分利用，于是梅勒合并了这两家工厂。巡查中更常见的情况是，他们发现工厂经理手中的存货太多，但是经理们在要求追

加资金时，却从不考虑投资的回报。

在梅勒的管理下，情况很快发生扭转，他和安德斯果断行动，创造了坚决强调回报的企业文化。具体来说，正如长期担任通用动力公司高管的雷·刘易斯（Ray Lewis）所言：“资本的现金回报率成为整个公司的重要指标，也时刻萦绕在我们的脑海中。”³这种扭转对整个行业来说是第一次，因为军工行业过去只会短视地关注收入增长和新产品开发。

重要的是，这一新的原则影响了该公司对政府合同的投标方式。在安德斯到来之前，通用动力公司和其他同行一样，使尽浑身解数去参与各种各样的政府招标。相比之下，安德斯和梅勒则坚持，只有在回报丰厚且中标概率很高的情况下才会投标。因此，投标次数急剧下降，投标成功率反而上升。正如资深的行业分析师彼得·阿塞里蒂斯（Peter Aseritis）所说：“安德斯和梅勒带来了新的关注点，即股东利益……这在军工行业还是第一次。”⁴

在他们执掌公司的头两年，安德斯和梅勒将总人数减少了近60%（行政人员减少了80%），他们将公司总部从圣路易斯迁至弗吉尼亚州北部，并建立了正式的投资审批程序，同时大幅削减营运资金的投入。正如梅勒所言：“头几年，我们不需要花任何钱，我们只需清空前几年积累的库存和资本支出就可以了。”⁵

这些举措为通用动力公司带来了海啸般的巨额现金，高达25亿美元，公司在资产回报率方面迅速成为同行中毋庸置疑的领头羊，这一地位一直保持到今天。

从资产出售到资本配置

资产出售把我们带到了另一个远超想象的公司现金源头。当梅勒从运营中攫取冗余现金时，安德斯开始剥离非核心业务，并通过收购来扩大他的主营业务阵营。有趣的是，当安德斯与业内同行会面时，他发现，整体上同行们购买资产的兴趣大于出售资产的兴趣。安德斯还发现，同行往往愿意为购买资产支付溢价。最终的结果是，通过一系列高溢价的资产剥离，通用动力公司的规模大幅缩水。

大规模的资产剥离对通用动力公司和整个行业来说都是第一次。在就任CEO的前两年，安德斯出售了通用动力公司的大部分业务，包括其

IT部门、民用飞机业务以及导弹和电子业务。

其中最大的一笔资产剥离值得我们深入研究，这就是安德斯出售了占公司主导地位的军用飞机业务，这让安德斯的战略框架遭受了意外的挑战。这笔交易实际上始于安德斯试图收购洛克希德公司（Lockheed）的小型战斗机部门，但是洛克希德公司的CEO拒绝出售其业务，并针对通用动力公司F-16的战机业务提出了价格夸张的反要约收购。

我们有必要在这里暂停一下，先来探讨一个宽泛的话题。本书中大多数局外人CEO都拒绝制订过于细节的战略计划，而更倾向于保持灵活性和见机行事。相比之下，安德斯有着非常明确和具体的战略愿景，不仅要求出售竞争力较弱的业务部门，而且要求强化主营业务部门。在出售业务部门的工作取得初步进展后，安德斯将注意力转向了军用飞机部门的并购，这是公司的主营业务，也是合乎逻辑的并购起点。作为退役的战斗机飞行员和航空爱好者，安德斯去拓展这个庞大的业务部门，除出于经济方面的考虑之外，还有对这个行业的热爱。因此，当洛克希德公司的CEO报出15亿美元的收购价时，安德斯倍感震惊。对通用动力公司的F-16业务来说，这是个令人目眩神迷的高价。这一刻，安德斯面临着关键的抉择。

安德斯最后的所作所为很能说明问题，他毫不犹豫地当场同意出售军用飞机业务，尽管带有一丝遗憾。安德斯做出了理性的商业决策，而这种决策与每股价值的增长相一致，尽管这使得公司的规模缩减到原来的一半以下，并剥夺了安德斯作为CEO最喜欢的特权：有机会驾驶公司尖端的喷气式战斗机。但就是这个决定勾勒出本书中局外人CEO们一个重要的共同点：作为一个群体，他们的内心理性且务实，他们不揣测未来，同时目光敏锐，他们不会恪守教条。即使价格合适，安德斯也不会出卖自己的母亲，但他会毫不犹豫地卖掉自己最喜欢的业务部门。

在业界，这种出售行为是前所未有的，因此安德斯备受争议，尤其是在五角大楼内部。然而，安德斯以其杰出的军事生涯在华盛顿赢得了无出其右的声誉，凭借着这个声誉，激进的行动方针在他手中得以实施。又是那位资深防务分析师阿塞里蒂斯，他对我说：“安德斯有点像尼克松，后者长期歌颂资本主义，却开启了与中国的关系。所以说，除了安德斯，很少有人有如此魄力。”⁶结果是，剥离这些资产使通用动力公司获得了额外的25亿美元现金收入，同时它还保留了两项占据市场主导地位的业务：坦克和潜艇。

资本返还，回报股东的三次特殊分红

随着资产出售和运营改善带来的大量现金，安德斯将注意力转移到了资本配置上。由于价格高企，安德斯选择不进行更多的收购，相反，他决定将公司的大部分现金返还给股东。为了以最有效的方式做到这一点，安德斯去向卡普尼克寻求建议。

卡普尼克是会计业巨头安达信会计师事务所（Arthur Andersen）的前董事长，作为一名律师，他久经考验且深谙税务法规。由于在Chicago Pacific这家综合企业集团的巨大成功转型中工作出色，卡普尼克声名鹊起。对于通用动力公司积累的大量现金，卡普尼克提出了两个创造性的建议，用非常高效的合法避税方式将大部分现金返还给了股东。

第一步，卡普尼克向股东派发了三次特别分红，总额刚好低于公司总市值的50%。由于通用动力公司很大一部分业务已经被安德斯剥离，这些分红可以被视为剥离业务产生的“资本回报”，既不需要缴纳资本利得税，也不需要缴纳普通所得税；第二步，安德斯和卡普尼克宣布了一项高达10亿美元的要约，以回购公司30%的股份。正如我们看到的，股票回购是非常高效的避税手段，而传统的分红在公司和个人两个层面都要缴税。

再怎么夸大这些举措的不同寻常也不为过：在不到3年的时间里，安德斯大幅精简了运营业务，卖掉了半数以上的子公司，整个过程创造了50亿美元的现金。安德斯并没有将现金重新配置到研发或新的收购中，而是使用创新的避税技巧将大部分资金返还给了股东。这些举措在军工行业中都是史无前例的，并为股东创造了巨大的价值。

一家上市公司系统性地缩减自己的规模是极其罕见的，安德斯对我总结道：“大多数CEO都会根据公司规模和增长给自己打分.....很少有人真正关注股东回报。”同样极其罕见的是（本书中其他局外人CEO除外），一家公司系统性地以特别分红或股份回购的形式将公司收益返还给股东。两个“罕见”的结合更是闻所未闻，尤其是在传统的军工行业。

华尔街为这一系列戏剧性的、陡然而起的举动所震惊，通用动力公司股价也随之一路飙升。这也引起了巴菲特的关注。巴菲特发现，在安德斯的领导下，通用动力公司正在剥离资产，并专注于创造性的、利于股东的资本配置策略。巴菲特随即在1992年以平均每股72美元的价格买

下了通用动力公司16%的股权。值得一提的是，尽管二人只有一面之缘，但巴菲特还是把伯克希尔-哈撒韦公司在通用动力公司的投票权委托给了安德斯代为行使，从而帮助安德斯实施其战略规划。

提升公司股价，保证“翻四倍”的三大措施

根据原定计划，安德斯的任期于1993年7月结束，离开时他把大权交给了梅勒。巴菲特在安德斯离开时出售了股票并获得了丰厚回报，然而今天，巴菲特对卖出通用动力公司的股票感到遗憾。在退隐美国西北部一个幽静的岛屿前，安德斯还担任了一年通用动力公司的董事长。安德斯笃信海军的接班模式，为了不影响继任者的权威，退休的船长会避免返回他们的舰艇。安德斯自豪地告诉我，自1997年起，他只跟梅勒的继任者查布拉贾有过一次对话。

在1981年加入通用动力公司之前，梅勒也曾在休斯飞机公司和利顿工业公司担任过工程师，并最终成为造船部门的主管。梅勒所在的部门帮助公司巩固了市场主导地位，他个人因为和安德斯志趣相投引起了安德斯的注意，并成为安德斯的副手和未来的接班人。

梅勒在安德斯离任后接任通用动力公司CEO，梅勒继续专注于优化运营，并出售了包括空间系统在内的最后几个小型非核心业务部门。然而，1995年他变守为攻，以4亿美元收购了美国最大的海军舰艇建造商之一巴斯钢铁公司（Bath Iron Works）。这次收购具有巨大的象征意义，它向员工和五角大楼表明，通用动力公司当时已经为再次成长做好了准备。正如梅勒所说：“收购巴斯钢铁公司终结了公司将被完全清算的传闻。”⁷1997年，梅勒到了退休年龄，并将接力棒传给查布拉贾。

查布拉贾毕业于美国西北大学法学院，在芝加哥顶级律师事务所简博律师事务所（Jenner & Block）从事公司法方面的事务近20年。20世纪80年代通用动力公司陷入麻烦时，查布拉贾为其提供法律服务。当安德斯到来时，查布拉贾已经成为公司的核心顾问。安德斯很快看到了查布拉贾的潜力，称其为“我见过的最有效率、最具商业头脑的律师”。1993年，查布拉贾以总顾问和高级副总裁的身份加入通用动力公司，不言而喻，他将成为梅勒的继任者。

成为CEO后，查布拉贾为自己设定了雄心勃勃的目标。具体而言，查布拉贾希望在任职CEO的第一个10年让公司股价翻两番（复合收益率

达到15%)。在回顾了标准普尔500指数的记录之后，查布拉贾发现这是一个难以企及的目标。记录显示：此前10年，位列《财富》500强的所有公司中只有不到5%达到了这一基准。查布拉贾冷静地展望着公司未来10年的前景并得出结论，他可以通过市场的增长和提高营业利润率来实现大约2/3的目标，其余的部分需要借助收购。这明显背离了安德斯的战略框架。

查布拉贾的收购方式很独特，对通用动力公司来说，一种新的资本配置重心应运而生，即围绕已有产业链，主要聚焦于小公司的收购。正如查布拉贾所说：“我们的战略是积极追求与公司核心业务直接相关的目标，并将我们的产品线扩展到相邻领域。”⁸上任第一年，查布拉贾收购了12家小公司。

雷·刘易斯将查布拉贾的收购方式描述为“即使在我们非常了解的市场中，每次也只前进一小步”。⁹逐次叠加的收购带来了公司的快速增长，最终引导通用动力公司进入了快速增长的军事信息技术市场。到2008年，军事信息技术领域成长为通用动力公司最大的业务单元。此外，通过查布拉贾一系列的收购，通用动力公司坦克部门得以成功研制出史崔克装甲车（Stryker Attack Vehicle）；同时，作为水下潜艇制造历来的领导者，收购也让通用动力公司业务从水下上升到水面，船舶集团开始建造更多的水上舰艇。

不过，查布拉贾时期最辉煌的成就当属1999年大手笔收购全球最大的商用飞机制造商湾流公司（Gulfstream）。这笔交易押上了公司的命运，总共价值50亿美元，相当于通用动力公司56%的市值。

因为价格看似高昂，而且明显偏离了安德斯“只关注军工行业”的战略，当时这宗交易遭到了广泛批评，然而它并不像看上去那么冒进。湾流公司是民用航空领域无可置疑的领导者，这个市场具有巨大的长期增长潜力。湾流公司已经在私募股权投资公司福斯特曼-利特尔公司（Forstmann, Little）手中运营了5年，但其对新产品开发的投资却相对滞后。

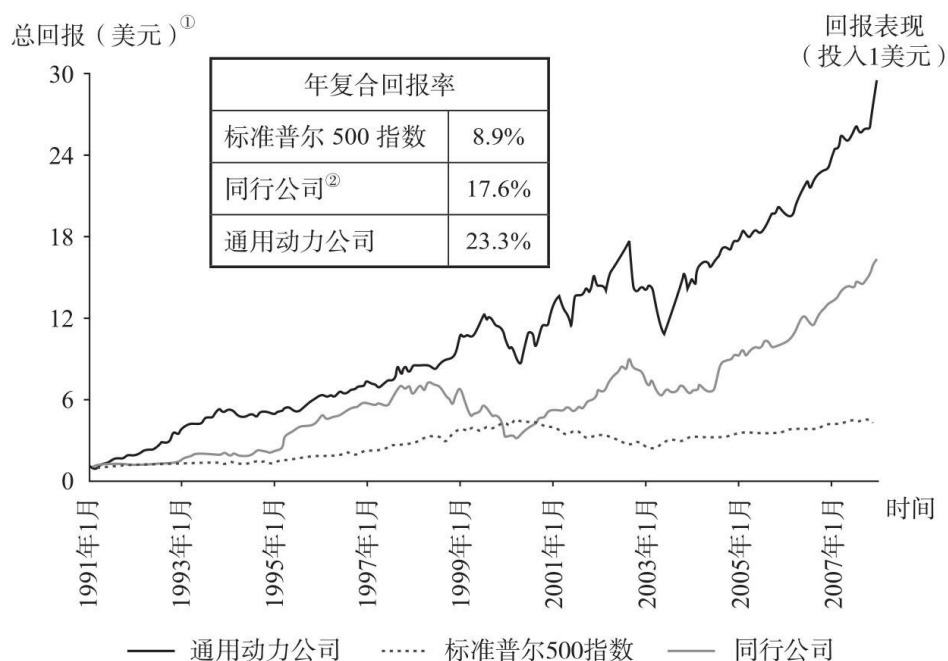
多年来，通用动力公司拥有民用飞机公司塞斯纳（Cessna），并为空军制造军用飞机，它在民用和军用飞机公司运营的管理方面积累了丰富的经验。查布拉贾认为，他将能够利用这一隐性的专业技能大幅提高湾流公司的业绩。查布拉贾还认为，民用航空将为通用动力公司提供宝贵的多元化收入来源，以抵御国防开支的波动。其后的公司收益证明了

查布拉贾的逻辑是正确的。在过去的很多年里，随着国防开支减少，湾流公司不断增长的业务为通用动力公司提供了根本性的支持，让它免受国防开支变幻莫测的周期影响。

至关重要的是，我们必须承认这样一个基本观点：周围环境始终在变化，你能否打好手中的牌，最终成为作为管理者的你能否成功的关键。尽管查布拉贾和安德斯共享着理性的、以股东利益为导向的思维定式，但是他们的具体行动却因情势不同而异。不同的举措，比如查布拉贾任期内的资产收购、安德斯时期的资产剥离，在他们任期内的不同时期都自有其道理，尽管他们都对股票回购充满同样的热情。

到2008年中期，查布拉贾离开CEO的岗位时，他创造的收益已经远远超越了自己当初那个雄心勃勃的目标。这就向我们提出了一个基本问题：由“三驾马车”管理的通用动力公司，整体回报到底如何呢？他们是如何齐心协力打败同行，并超过杰克·韦尔奇那看似高高在上的业绩指标的呢？

从1991年1月安德斯到任至2008年7月查布拉贾离任，在长达17年半的时间里，安德斯和他钦点的两位接班人造就了惊人的结果，他们为通用动力公司的股东创造了23.3%的年复合回报率。相比之下，标准普尔500指数的年复合回报率为8.9%，其他同行公司为17.6%，如图3-1所示。



注：① 投资回报包括股票分拆和股票分红。

② 同行公司包括洛克希德 - 马丁公司（LMT）和诺斯罗普 - 格鲁曼公司（Northrop Grumman）数据基于 1991 年 1 月 1 日各公司股票市值的加权平均值。

图3-1 通用动力公司“三驾马车”造就的历史

资料来源：证券价格研究中心（CRSP）和通用动力公司年报。

安德斯掌舵时投入的1美元，17年后的价值达到了30美元；如果投资于同行公司指数，同样的1美元价值达17美元；如果投资于标准普尔500指数，则价值达6美元。CEO三人组的表现分别是同行的1.8倍和标准普尔500指数的5倍，超过了被视为标杆的杰克·韦尔奇。

今天，通用动力公司看起来与安德斯卸任时大不相同，但它的基本原则并没有改变。通用动力公司在每条业务线上都是不容置疑的市场领导者，拥有业界最高的利润率和资产回报率，以及稳如泰山的资产负债表。2007年底，查布拉贾指定了杰伊·约翰逊（Jay Johnson）作为继任者。约翰逊有着令人印象深刻的履历：在成为通用动力公司副董事长和指定接班人之前，他是有史以来最年轻的海军作战司令，也是电力巨头多米能-弗吉尼亚公司（Dominion Virginia Power）的CEO。

与此同时，安德斯在圣胡安（San Juan）的一个幽静的小岛上定居下来，过上了典型的、老有所为的退休生活。安德斯在西雅图城郊建造了一所广受好评的航空博物馆，到了70多岁，仍然能驾驶喷气式飞机。尽管安德斯与通用动力公司保持距离，但公司的股票却一直在握。

The Outsiders

比尔·安德斯的方法

“三驾马车”取得惊人回报的关键在于，通用动力公司在人力和资本方面实施了高效的配置方法，即使比照军工行业标准，这也很不寻常。在运营领域，安德斯和继任者专注于两个主要的优先事项：去中心化的组织以及管理层薪酬与股东利益相一致。

去中心化，将权力下放

鉴于许多CEO的军人背景，传统军工行业的特点是集中化和组织结构官僚化，这一点不足为奇。然而，在安德斯及其两位继任者的领导下，通用动力公司推行了一种截然不同的组织策略。20世纪90年代初，安德斯和梅勒在收紧业务、大幅削减总部人员的同时，开始积极推动去中心化，并将权力在组织内部逐级下放，取消了管理的中间层级。查布拉贾延续了这种去中心化的做法并将其发扬光大。

到查布拉贾任期结束时，公司的雇员人数比安德斯上任时要多，但是公司总部的人数只有原来的1/4。CEO和任何利润中心负责人之间只隔了2个人，而以前隔4个人。总部所有的人力资源、法务和会计人员都被裁撤或下放到业务部门，同时公司竭尽所能地把总部对各业务部门的干涉降至最低，以防止查布拉贾所谓的“总部员工总在业务人员周围晃悠”。运营经理要对他们的预算超支承担责任，用查布拉贾的话来说“这是底线”，如果预算超支，他们就得走人。¹⁰

从安德斯开始，通用动力公司开始强调基于绩效的薪酬。20世纪90年代初，安德斯心知肚明，只有大幅提高薪酬，才能吸引新的管理人员加入通用动力公司。安德斯本想制订一个传统的股票期权计划，但董事会告诉他，股东们对他上任前几年的股票表现感到不满，不会批准这一计划。然而，安德斯希望将管理层薪酬与股东利益绑定在一起，于是制订了一项薪酬方案，让管理层能够在持续上涨的股价中获益。

问题在于，这项薪酬方案几乎刚刚付诸实施，华尔街就开始认识到安德斯不同寻常的举动将带来的影响，股价随之迅速上扬，导致管理层很早就收到大笔的奖金。这些奖金立即被媒体揪住，引发了巨大的争议。然而，通用动力公司仍然坚守承诺，推行绩效薪酬方案。如今，通用动力公司高管薪酬的核心组成部分仍然是奖金和股票期权。

管理层薪酬与股东利益绑定在一起

在安德斯及其继任者的领导下，通用动力公司筹集和配置资本的方式与军工同行截然不同。凭借安德斯早期资产剥离带来的巨额收益，以及持续、健康的经营现金流，通用动力公司无须动用巨大的财务杠杆或者发行股票的方式来融资，然而仍然发生了一次重大例外。

不过，这次例外有助于揭示资本配置的一个重要问题。查布拉贾任期内的重大事件就是收购湾流公司。那么，查布拉贾到底是如何支付这笔巨额交易费用的呢？与安德斯的做法背道而驰，查布拉贾的做法属于见机行事且不循规蹈矩，查布拉贾增发了通用动力公司的股票，而且增发了很多。这似乎是一个稀释股东权益的举动，然而仔细研究，我们就会发现其中的奥妙：查布拉贾的做法与安德斯的原则殊途同归。

如图3-2所示，此次股票发行恰逢通用动力公司股票市盈率达到历史新高。与巴菲特大举收购通用再保险公司（General Re Corporation）没什么不同，此次也是以发行股票来获取现金，然后在历史高位完成溢价交易。

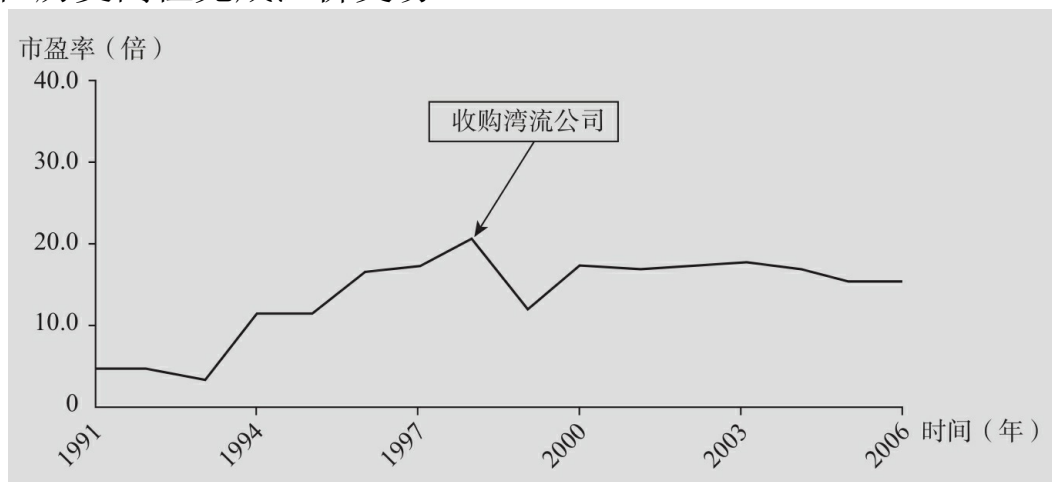


图3-2 通用动力公司市盈率曲线（以年均数值为基准）

资料来源：证券价格研究中心和通用动力公司年报。

查布拉贾向我描述过：“驱使我这么做的原因是，我意识到该股的交易价格比历史正常水平高出了很多：股价相当于公司来年期利润的23倍，而历史平均市盈率是16倍。这么高的股价，你会怎么做呢？当然是利用这次机会，收购相关领域低市盈率的溢价资产，并从中套利获益。”¹¹雷·刘易斯做出这样的总结：“查布拉贾卖出了相当于公司1/3市值的股票，然后买入了一项业务，这项业务提供的现金流占据了企业整合后全部现金流的一半。”¹²

和安德斯出售公司的F-16业务一样，查布拉贾的股票增发也强调这样一个重点：最好的资本配置者都是务实的，他们见机行事、随形而变，不受意识形态或既定战略的束缚。1999年，查布拉贾看到了极其难得的机会，遂利用独特的廉价通货（股票）使得公司实力大幅增强，公司业务也得到多元化的发展。查布拉贾抓住了机会，为他的股东实现了可观的经济增值。

“三驾马车”的资本配置法则

那么，三位CEO究竟用全部现金做了什么呢？

在资本配置方面，安德斯和继任者的种种决策一直和军工行业的巨头们有着根本的不同。我们已经看到，当同行们陷入收购的狂热时，安德斯却成为一个积极的卖家。安德斯没有进行任何收购，在资本支出上也花费很少，并且精明地利用了分红和股票回购的策略，这一切对军工行业来说都是新鲜事儿。

在延续了安德斯收缩业务的策略之后，梅勒在资本配置舞台的主要贡献是，通过1995年与巴斯钢铁公司的大笔交易重新开启了收购之门，梅勒同时秉承了安德斯对分红和资本支出的节俭做法。与安德斯一样，查布拉贾也采取了反常规的分配策略：资本支出上的花费明显少于同行，支付的股息也少于同行，同时他将主要资源用于收购，并时不时地回购股票。

这些回购的回报率非常高，使得通用动力公司的平均回报率达到17%。三位CEO都致力于回购，安德斯和查布拉贾在这一重要领域尤为活跃。如我们所见，在卡普尼克的帮助下，安德斯在回购方

面的贡献是1992年的要约收购，那一次公司30%的股份被回购。

如我们所见，查布拉贾也是通用动力公司股票回购的“大胃王”。与他的法学背景不符，查布拉贾更像投资者那样去思考，不断地将通用动力公司的股票价格与其内在价值进行比较，并在看到二者的差异时果断出击。谈及查布拉贾的时代，雷·刘易斯这样说：“当认为可以好好利用市场对公司股票的错误定价时，我们就狠狠地买入。”¹³

颠覆传统，不与愚蠢为伍

安德斯和查布拉贾有着相似的、不讨喜的性格，他们都热衷于颠覆传统，而且都不乐意与愚蠢为伍，这种态度在局外人CEO群体中绝非个例。他们对待华尔街分析师的方式与足球教练比尔·帕斯尔斯对待记者的方式大致相同，总体上的态度往往近乎鄙视。他们鄙视投资分析师的那点雕虫小技，虽然华尔街完全无法欣赏他们的冷傲，但通用动力公司的股东们肯定欣赏他们。

超级模仿者：诺斯罗普-格鲁曼公司

如果模仿是最真诚的恭维，那么通用动力公司最近无疑得到了诺斯罗普-格鲁曼公司的大力赞扬。诺斯罗普-格鲁曼公司是美国军工行业最大的公司之一，也是战斗机和导弹系统的领先制造商。自20世纪90年代初开始，该公司股票市场的回报率大大落后于通用动力公司。

2009年，新任CEO韦斯·布什（Wes Bush）接手诺斯罗普-格鲁曼公司并宣布了重大的战略转型，其特征是出售非核心资产，重新强调净资产收益率和股票回购，以及大幅减少企业员工人数。听起来是不是似曾相识？

正如一位华尔街分析师所言：“诺斯罗普-格鲁曼公司正在采取的步骤，让人回想起20世纪90年代初通用动力公司的变革……而且与军工企业的通常做法相悖，传统的军工企业总是过分强调收入增长……在过去的20年里，通用动力公司的表现明显优于其他继续追求规模的军工企业。”

自从韦斯·布什宣布这项新战略并让人们看到初步落实的迹象后，诺斯罗普-格鲁曼公司的股票价格就开始疯涨。今天，即便面临截然不同的外部政策环境，“安德斯公式”的每个部分仍然像倒塌前的柏林墙那样坚如磐石。

特质**4** 高效运用杠杆

约翰·马隆和**TCI**

The Outsiders

- 他是“美国有线电视业教父”，时代华纳公司大股东，仅次于默多克的新闻集团第二大股东。他是巴菲特最欣赏的商业奇才。
- 他是美国最大的地产拥有者，名下拥有8 900多平方千米的土地。
- 20世纪90年代后期，其战略性的非有线电视项目结出硕果。1997年，Teleport以110亿美元被卖给了美国电话电报公司；1998年，他将Sprint/PCS合资公司出售给Sprint公司，获得了作价90亿美元的Sprint公司股票；1999年，通用仪器公司被以110亿美元的价格出售给摩托罗拉。



他是

约翰·马隆

JOHN MALONE

约翰·马隆

至少，无论如何……他们还没有摒弃算术法则。

布兰奇·瑞基
(Branch Rickey)

运气可遇不可求。

到1970年，马隆在麦肯锡咨询公司工作的经历，足以让他一眼便知哪个行业更具吸引力。对有线电视行业了解得越多，马隆就越喜欢它。特别是这个行业有三件事引起了马隆的注意：类似公用事业那样可精确预测的收入；税收的特别优惠；行业的发展好像疯长的野草。在麦肯锡咨询公司工作的那些年里，马隆从未见过哪个行业能将这三个特点结合在一起，他很快下定决心，要在有线电视行业成就自己的事业。

财务杠杆，举债扩张的巨大魅力

高增长性和可预测性的结合尤其有吸引力。在20世纪六七十年代，有线电视行业呈现出快速增长的态势，全美各地的乡村社区都在为他们喜爱的频道和节目寻求更好的电视信号接收效果，有线电视用户数量随之增长了20多倍。有线电视用户按月付费，很少要求停网，这使得有线电视业务可以高度量化，经验丰富的高管从而能够非常精确地预测客户增长和盈利能力。马隆拥有异于常人的定量分析能力，干这一行简直就是如鱼得水。套用诺曼·梅勒（Norman Mailer）的话来说：“马隆加入有线电视业务，犹如超人来到了超市。”

1941年，马隆出生于康涅狄格州的米尔福德（Milford）。他的母亲是一名教师，父亲是一名研发工程师，每周有5天都在走访通用电气的工厂。马隆很崇拜自己的父亲。十几岁时，马隆在机械制造方面的才能初见端倪，他通过购买、翻新、出售二手收音机来赚取零花钱；高中时，马隆参加过击剑、足球和田径比赛，是个运动健将。马隆毕业于耶鲁大学，获得了经济和电子工程的双学位，毕业后几乎立刻就与高中时的心上人莱斯莉步入婚姻殿堂。

从耶鲁大学毕业后，马隆在约翰·霍普金斯大学获得了运筹学硕士和博士学位。马隆研究的两个学术领域即工程学和运筹学都要进行高度的定量分析，且都专注于系统优化，聚焦于最小化“噪声”、最大化“输出”。事实上，马隆未来的整个职业生涯可以看作是高效“价值工程学”的延伸，最大化输出的是股东价值，最大限度减少的是各种噪声，

比如减少税收、管理费用和各种法规的影响。

获得博士学位后，马隆就职于美国电话电报公司享有盛誉的研究机构贝尔实验室。在那里，他专注于研究垄断市场下公司的最优策略。经过大量的财务建模，马隆得出结论：美国电话电报公司应该提高债务水平，并通过股票回购来大量减少股本总数。美国电话电报公司董事会表面上欣然接受了这个非常规的建议，但随之将其束之高阁。

几年后，马隆认定自己不适应美国电话电报公司的官僚文化，于是去麦肯锡咨询公司找到了一份差事，并向妻子保证不会向父亲那样频繁出差。但是很快，马隆就发现这份工作要服务于形形色色的《财富》500强公司，他每周有4天在路上奔波。1970年，马隆的客户通用仪器公司（General Instrument）为他提供了一个职位，让他去管理旗下正在迅速发展的有线电视设备部杰罗尔德（Jerrold）。29岁的马隆欣然接受了这个机会。

在杰罗尔德，马隆积极培养与各大有线电视公司的关系，两年后，他同时收到了两家最大的有线电视运营商老总的邀请：华纳通信公司的史蒂夫·罗斯（Steve Ross）和TCI的鲍勃·马格内斯（Bob Magness）。尽管TCI给的薪水比华纳通信公司低了60%，但是因为马格内斯为马隆提供了一个更好的股票期权方案，而且他的妻子更喜欢相对宁静的丹佛而不是快节奏的曼哈顿，所以马隆最终选择了TCI。

马隆决定加入的TCI曾有过长期高速发展的历史，然而当时却面临着破产危机。1956年，马格内斯抵押了房产，然后在得克萨斯州孟菲斯市买下了自己第一个有线电视网络，TCI得以创立。马格内斯此前是一个四处兜售棉花籽的推销员和牧场主，在一次搭顺风车的时候了解到有线电视业务，就像15年后的马隆一样，他立刻就意识到这个行业诱人的经济特性，并在极短的时间内抓住了该行业税收优惠的特点。

精明的有线电视运营商通过举债去建设新的电视网络以及对固定资产进行大幅折旧，成功地利用税盾效应去减少税收对现金流的影响。大额的折旧支出和债务利息费用都能降低应纳税收入，因此，运营良好的有线电视公司账面上很少显示营业利润。最后的结果就是，尽管现金流非常充裕，但公司却很少缴纳税款。如果运营商不断地利用举债去购买或建造新的有线网络，进而对新获得的资产进行折旧，它们就能够继续无限期地利用税盾效应进行避税。马格内斯是最早充分认识到这些特性的先驱之一，并积极利用税务杠杆把公司做强做大。马格内斯的名言

是“付息胜于纳税”。

TCI于1970年上市，到1973年马隆加入时，它已经成为美国第四大有线电视公司，用户达到60万个。当时公司的债务已经达到惊人的水平，是营业收入的17倍。马格内斯意识到，在下一个成长阶段，他需要新的管理人才来带领公司。经过漫长的“求爱”之后，马格内斯终于俘获了麦肯锡咨询公司的这位青年才俊。马隆给TCI带来了不同寻常的卓越组合，包括卓越的分析能力、丰富的财务经验、对技术的精通和过人的胆识。然而，马隆的任期开始时却十分不顺。

1972年末，有线电视行业的股票市场十分火爆，TCI计划公开增发新股，以减轻巨额的债务负担。但是，马隆到任后的数月内，整个业界被行业新规弄得晕头转向，有线电视股也迅速降温，迫使公司取消了增发计划，TCI因此陷入了摇摇欲坠的债务危机。

1973—1974年的阿拉伯石油禁运导致了有线电视市场的流动性突然蒸发，整个行业处于风雨飘摇之中。拥有32岁新任CEO的TCI，其债务负担远远超过任何同行，蹒跚在破产的边缘。“简直一塌糊涂。”马隆以其素有的直率评价了他在TCI的开局。¹

马隆不得不采用强硬手段，他和马格内斯在接下来的几年里一直在阻止债权人行使权力，让公司免于破产。他们不停地与银行家见面。在一次剑拔弩张的债权人会议上，马隆把钥匙扔在会议室的桌上，边走出房间边说：“如果你们想要这一堆电线，拿去便是。”惊慌失措的银行家最终让步，同意放宽TCI的债务条款。

规模杠杆，用户人数决定未来成败

在此期间，马隆引入了新的财务和运营规定，他告诉各运营经理，如果他们在保持利润的同时每年能增加10%的用户，他就给予他们独立经营及独立决算的权力。崇尚节俭、锐意创新的企业文化经过多年发展在公司内蔓延开来，逐渐从公司总部渗透到运营一线。

作为一家正在重新定义美国传媒业格局的公司来说，TCI的总部怎么看也不像是行业龙头的总部。公司办公室装修简朴，只有为数不多的高管人员和极少的秘书，塑胶地板上摆放着金属涂层已经开始剥落的办公桌。公司只有一位接待员和一部自动应答电话。TCI的高管们出差都是合住，而且通常住的是汽车旅馆。COO斯帕克曼回忆说：“那些日

子，对我们来说，能住上假日酒店是罕见的奢侈。”²

马隆把自己视为投资者和资本配置者，他将日常运营的职责委托给斯帕克曼。作为马隆的左膀右臂，斯帕克曼通过严格的预算程序管理着公司遍布各地的业务。这位从空军退役的军官，以近乎军事化的管理要求运营经理们必须达成现金流的预算目标。只要他们达成目标，这些一线的管理者就拥有高度的自主权。而那些月度预算不达标的经理人会经常受到巡视的COO的拜访，表现不佳的人很快会被淘汰。

·高楼大厦情结·

精心建造的新总部大楼与投资者回报之间存在着明显的负相关。例如，纽约时报、IAC集团和时代华纳这三家媒体公司都曾斥巨资在曼哈顿市中心建造了像泰姬陵一样精致的总部大楼。与此同时，这些公司中没有一家进行过重大的股份回购，也没有获得超过市场预期回报。与之相反，局外人CEO中没有任何一位建造过豪华的总部。

由于厉行节俭，TCI在很长一段时间内有着业界最高的利润率，并在其投资人和债权人中赢得了美誉。长期以来，公司“承诺得很少，但是做得很多”。翻阅公司早期历史上那些分析报告，可以看到一个模式在反复循环：一个季度又一个季度，公司实际的现金流和用户数总是略高于预期。

1977年，TCI终于发展到足以吸引保险公司财团，从而以较低成本的负债取代银行债务。随着资产负债表日益稳健，马隆终于能够发动攻势，实施他对TCI的战略规划。这项规划另辟蹊径，它源于马隆的核心战略洞察力，自从加入公司的那一刻起，就已经在马隆心中萌芽。

作为工程师兼企业优化师，马隆很早就意识到，在有线电视业务中创造价值的关键是最大化财务杠杆与供应商杠杆，尤其是那些节目供应商，而这两种杠杆的关键就是规模。这就是看似简单但实则强大的洞察力，马隆一心一意地追求着这个目标。正如他在1982年告诉TCI资深投资者戴维·沃戈的：“有线电视未来盈利和成功的关键，将取决于通过杠杆规模来控制节目成本的能力。”³

在一个有线电视网络中，占总运营费用40%的最大的成本项目是支付给节目供应商的费用，比如HBO电视网、音乐电视网、娱乐与体育电

视网等。较大的有线电视运营商能够通过谈判来降低每个用户的节目成本，并且有线电视公司的用户越多，每个用户的节目收看成本就越低，有线电视公司的现金流也就越高。随着公司规模不断扩大，这些费用折扣也在累积，从而为最大的行业玩家提供了强大的规模优势。

因此，与小规模公司相比，支付最低节目费用的反而是规模最大的公司，它们在进行新的并购时具有可持续的优势：它们能够支付更多的费用以买入一家有线电视公司，并购之后仍能获得相同或更高的回报。这就创造了扩大规模的良性循环，就像这样：你如果买入更多的有线电视网，就会降低节目成本并增加现金流，这又允许你使用更多的财务杠杆去买入更多的网络，从而进一步降低节目成本。这个反馈回路的逻辑及其能量现在看来一目了然，但当时没有其他人像马隆和TCI那样疯狂地追求规模。

与追求规模这一核心思想相关的是，马隆认识到，当时大多数上市公司所追求的是每股收益最大化，这与新兴的有线电视行业追求规模的做法不一致。对马隆来说，净利润越高意味着税收就越高，他认为有线电视公司的最佳财务策略是：使用一切有效手段来最小化账面盈余和各种税项，并用税前现金流为企业内部成长和外部并购提供资金。

很难找到合适的词来形容这种不合常规的方法。当时，华尔街以一段时间内的每股收益来评估公司效益。在很长一段时间内，马隆是电视行业里唯一关注现金流的人；其他大型有线电视公司最初是以每股收益为目标来经营公司的，后来它们意识到发展有线电视业务时很难兼顾每股收益，才转而关注现金流，而康卡斯特公司到20世纪80年代中期才最终转向关注现金流。正如长期跟踪有线电视行业的分析师丹尼斯·莱博维茨（Dennis Leibowitz）告诉我的那样：“忽视每股收益，使TCI在与其他上市公司的竞争中获得了重要的先发优势。”⁴

尽管关注现金流战略现在看来理所当然，并最终为同行们所复制，但在当时，华尔街并不知道该如何利用它。摒弃每股收益的概念，马隆向债权人和投资人强调的是现金流，并在这个过程中，发明了一个为今天的经理人和投资人所熟知的新词。这个术语和概念就是EBITDA，由马隆首度引入商业词汇中。EBITDA是一个全新的概念，它比以往任何概念都更进一步地阐释了利润表，从而得出了企业在支付利息、税项、折旧或摊销费用之前产生现金的能力。如今，EBITDA在整个商业领域得到了广泛应用，特别是在私募股权业和投资银行业。

有线电视行业的股票行情在整个20世纪70年代和80年代早期一直起伏不定。马隆和马格内斯担心公司有可能被恶意收购，于是利用市场偶尔的下跌，把握机会进行回购，从而增加了他们联合控股的比例。1978年，他们创立了拥有超级投票权的B股，并实施了一系列复杂的回购和交易，到1979年总共控制了56%的B股，从而确保了资深高管谢诚刚（John Sie）所描述的状态：对TCI的“铁腕控制”。

自那时起，由于掌握了控股权和更强劲的资产负债表，马隆秉持坚持不懈、不断创新的精神开始追求扩大规模。利用公司新债权人的出资、公司内部现金流和偶尔的股票增发，马隆开始了一项异常活跃的收购计划。1973—1989年，TCI完成了482宗收购案，平均每周一次。对马隆来说，完成一宗收购案就增加了一群用户，再完成一宗再增加一群，再完成再增加。正如老牌投资人里克·赖斯（Rick Reiss）所说：“为了追求规模，马隆愿意去考察海滨的物业，即便那里靠近一个臭气熏天的垃圾场。”多年来，马隆从各种各样的卖家那里购买电视网络，比如卡车司机工会（Teamsters）和美国前第一夫人伯德·约翰逊（Bird Johnson）。⁵

然而，马隆并没有盲目地收购。20世纪70年代末和80年代初，随着HBO电视网和音乐电视网等卫星频道的出现，有线电视行业进入了新阶段。它的服务对象突然从主要面向接收效果不佳的乡村客户转变为城市客户，后者正对内容匮乏的电视节目感到厌倦，高度期待新的频道。随着有线电视行业进入新阶段，许多大型有线电视公司开始集中精力争夺在大都市的特许经营权。这些特许经营权的竞标很快升温并价格不菲。

但是马隆并未随波逐流。有线运营商往往受制于市政当局，后者会强加一些极其苛刻的经济条款，马隆对此感到十分不满。在特许经营权的竞标之战中，马隆是唯一拒绝参战的大型有线电视运营商，他转而关注对低价的乡村和郊区用户的争取。1982年，TCI拥有250万个用户，已然成为业内最大的公司。

在过高的债务和政府苛刻的合同条款的综合作用下，许多早期获得了城市特许经营权的公司纷纷破产，马隆却大步前进，仅以当初竞标价格的零头就获得了经营权。通过这种方式，TCI获得了匹兹堡、芝加哥、华盛顿、圣路易斯和布法罗等市的有线电视特许经营权。

整个20世纪80年代，美国联邦通信委员会的法规都非常宽松。得益于此，TCI以激进的方式不断购入电视网络，稳定的小额收购中也夹杂

着几笔大额交易，比如收购西屋公司（Westinghouse）和斯托勒通信公司（Storer Communications）。除此以外，TCI还通过合资企业的方式继续积极布局，与那些业界传奇的企业家合作，如比尔·布雷斯南（Bill Bresnan）、鲍勃·罗森克兰茨（Bob Rosenkranz）和利奥·欣德里（Leo Hindery）。在和这些企业家合作创建的有线电视公司里，TCI仅持有少量股权。1987年，TCI的规模是其第二大竞争对手时代公司下属公司ATC的两倍。

20世纪70年代末和80年代初，通过与那些年轻有为的节目制作人和有线电视行业的创业者合伙，马隆的创造力在合资企业的浪潮中得到了进一步证明。读一读这些合伙人的名单，就像进入了有线电视领域的名人堂，这些人包括泰德·特纳（Ted Turner）、谢诚刚、约翰·亨德里克斯（John Hendricks）和鲍勃·约翰逊（Bob Johnson）。在建立这些合伙关系的过程中，马隆实际上就是极具创造力的风险投资家，他积极网罗那些年富力强的创业者，并把TCI庞大的用户数和极低的节目成本等规模优势嫁接入这些创业公司。作为回报，马隆会要求获得这些企业的少数股权。如此一来，马隆为股东创造了巨额回报。一旦发现自己中意的创业者或者好点子时，马隆就立即采取行动。

强强联合的开端当属1979年，马隆和黑人娱乐电视（BET）的创始人鲍勃·约翰逊第一次见面结束时，就给对方签下了一张50万美元的支票，一时间名声大噪。马隆由此开始积极追求对节目制作方的参股，作为回报，他提供了让对方很难拒绝的组合方案：TCI提供创业启动资金和链接公司数百万家庭用户的机会。1987年，特纳的特纳广播网（Turner Broadcasting System）濒临破产，其节目频道包括美国有线电视新闻网（CNN）和卡通电视网（The Cartoon Network），马隆带领一个有线电视公司财团对其进行了援助。到20世纪80年代末，TCI的节目组合除了特纳公司的频道外，还包括探索频道、昂科拉电影频道（Encore）、QVC购物频道和BET娱乐频道。至此，马隆已成为有线电视网络和有线电视节目的重要拥有者。

业务再梳理，如何把自己的公司卖到天价

20世纪90年代初，有线电视行业迎来了一场不能更糟的厄运风暴，先是1990年出台的针对高杠杆交易（HLT）的新法规，它限制了行业的债务融资渠道；然后更糟糕的是，美国联邦通信委员会在1993年收紧了行业监管政策，降低了有线电视费率。马隆无视这些不利因素，继续选

择性地收购大型有线电视网络维亚康姆和联合艺术家有线电视公司（Viacom and United Artists Cable），并推出新的节目网，包括星光电视台的昂科拉电影频道（Starz/Encore）以及与鲁伯特·默多克、福克斯公司合作出品的一系列地区性体育节目网络。

1993年，马隆提出了一项惊人的发展举措，他与电话巨头大西洋贝尔公司（Bell Atlantic）达成协议，将TCI的股票作价340亿美元出售给对方。然而，收紧后的行业法规不允许这项并购，加上TCI的现金流和股价齐跌，这项交易最终被取消。随后10年，马隆把更多的时间花在了有线电视业务之外的非核心项目上。他带领一个由不同的有线电视公司组成的财团，创建了两个颇具规模的新实体：Teleport和Sprint/PCS，前者是一家极具竞争力的电信服务公司，后者是与Sprint共同竞标移动电话特许经营权的合资企业。

在达成这些新目标的过程中，马隆把公司资本和自己的时间都分配到这些项目上，他认为这些项目会极大地提升公司的市场主导地位，并将提供可观的潜在回报。1991年，马隆将占TCI小部分权益的节目类资产分拆出来，并入新成立的自由媒体公司，并最终获得了该公司的大量个人股权。这是马隆创建的一系列追踪股票（tracking stocks）中的第一只，其他的追踪股票包括用于投资Teleport、Sprint/PCS、其他非有线资产的TCI风投和用于投资各种各样外国有线电视资产的TCI国际。

马隆是公司分拆和股票追踪领域的先驱，他认为二者实现了两个重要目标：

- 提高透明度，使投资者能够对公司不同部分的资产进行估值，而这些资产之前为TCI迷宫般的公司结构所掩盖。
- 促成TCI核心的有线电视业务与其他利益相关方特别是节目制作方的进一步分离。如果不分离，就有可能构成关联交易，从而导致监管机构的审查。

1981年，马隆开始分拆西部远程通信公司（Western Telecommunications）的微波业务。当这项业务出售给美国电话电报公司时，TCI已经分拆了14家不同的实体，股东由此获得了非凡回报。像亨利·辛格尔顿和比尔·斯蒂利茨（Bill Stiritz）一样，在分拆过程中，马隆为了让股东获得最佳的经济效益，有意识地增加了交易的复杂性。

1995年斯帕克曼退休后，马隆将公司有线电视业务的运营权交给了新的管理团队，团队由前营销主管布伦丹·克劳斯顿（Brendan Clouston）负责。在克劳斯顿的领导下，TCI开始重视客户服务，并大举投入资金升级老化的有线网络基础设施。但是，在1996年第三季度，TCI严重偏离预期，有史以来首次出现用户流失，季度现金流也出现了下滑。马隆对此结果感到十分失望，他收回帅印，并一反常态地直接管理运营：迅速裁员2 500人，停止所有资本设备订单，并积极地对节目采购合同重新谈判。马隆还解雇了一些顾问，他们此前被雇来为有线网络的升级提供服务。马隆还将客户服务的责任交还给了负责地区网络的经理们。

业务逐渐稳定下来，现金流也得到改善，马隆邀请利奥·欣德里主抓运营。欣德里是TCI大型合资企业联合媒体合伙人公司（InterMedia Partners）的CEO。马隆自己将注意力重新转回到那些战略目标上。欣德里继续推进重组进程：任命TCI的老职员马文·琼斯（Marvin Jones）为首席运营官，赋予区域经理更多职权，他还积极追求并购交易，从而使得用户集群更加密集且成本更加低廉。

欣德里刚到任，马隆就将注意力集中在开发数字机顶盒上，这项技术使得有线电视行业能够有效抗衡那些新的卫星电视运营商。马隆曾向微软寻求合作，但最终与该行业最大的设备制造商通用仪器公司达成协议，以每台300美元的价格订购1 000万台机顶盒。作为回报，马隆要求拥有通用仪器公司的大量股权，并最终获得了16%的股权。

在1996年和1997年的经营危机中，马隆的导师和长期合作伙伴马格内斯去世，这使公司的控股权陷入不确定状态。通过一系列惯常的复杂交易，马隆与公司一道，买下了马格内斯那些具有超级投票权的股票，从而确保在终局阶段保有对TCI的“铁腕控制”。

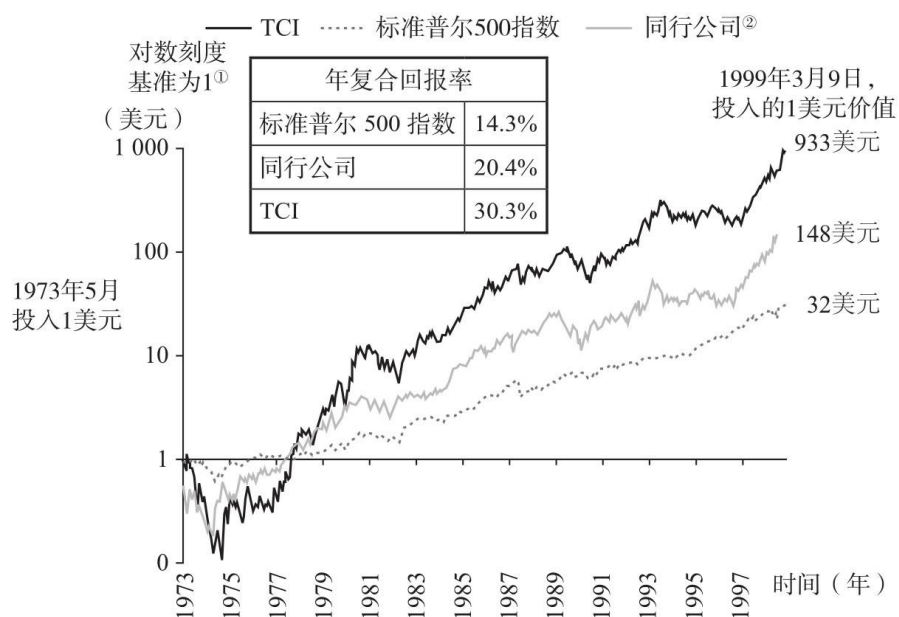
20世纪90年代后期，马隆的几个战略性的、非有线电视项目开始结出丰硕果实，他此前对项目潜在的回报预期也被证明是正确的：1997年，Teleport以惊人的110亿美元的价格出售给美国电话电报公司，投资回报率高达28倍。1998年，Sprint/PCS合资公司出售给Sprint公司，获得了作价90亿美元的Sprint公司股票；1999年，通用仪器公司被以110亿美元的价格出售给摩托罗拉。

同样是20世纪90年代后期，马隆的重心有所转移，希望为TCI找个好买家。尽管马隆的内心热爱有线电视业务，但作为一位极度理性的

CEO，早在1981年，他就曾对分析师戴维·沃戈表示：“我觉得TCI的股票每股值48美元，如果有人向我们开这个价，我们就卖。”⁶这个目标价持续提升，而且在很长一段时间内无人问津。然而，在20世纪90年代的发展中，马隆看到了一大片影响TCI未来的阴云：来自卫星电视的竞争持续加剧、公司乡村网络升级的巨大成本以及继任管理层的不确定性。当他收到美国电话电报公司雄心勃勃的新任CEO迈克·阿姆斯特朗（Mike Armstrong）的询价信时，马隆便热切地开始与对方讨价还价。和往常一样，马隆亲自主持这些谈判，常常要面对桌子对面那一大堆来自美国电话电报公司的律师、银行家和会计师。

随着两家公司谈判逐渐展开，马隆证明了自己 在出售公司方面和收购公司一样娴熟。正如里克·赖斯所说：“马隆把美国电话电报公司的董事会翻了个底朝天，晃荡出他们口袋里的每一分钱，然后才让他们回到董事会的位子上去。”⁷最终的结果是：交易价格相当于12倍的EBITDA，即每个用户作价2 600美元。这个结果超乎想象，更惊人的是，公司那些如同打满补丁的棉被一般的、老化的乡村网络一点也没有打折。当然了，马隆一贯对不必要的税收保持高度警觉，他把这笔交易架构成股票买卖，从而允许他的投资人延迟缴纳资本利得税。

此外，马隆通过获得董事会9个席位中的6个，保留了对子公司自由媒体的有效控制，并为其频道在美国电话电报公司有线电视网络中播出签署了一份价格诱人的长期合同。这笔交易是马隆在TCI独特战略的最后一次有力验证，为他的投资人创造了令人难以置信的丰厚回报：在马隆掌舵TCI的25年里，整个有线电视行业迅猛发展，行业里所有的上市公司都兴旺发达。然而，与马隆作为股东创造的巨大价值相比，其他所有有线电视公司的高管只能望洋兴叹。从1973年马隆首次亮相到1998年把公司出售给美国电话电报公司，TCI股东的年复合回报率达到了惊人的30.3%，而同期其他上市交易的有线电视公司的年复合回报率为20.4%，标准普尔500指数的年复合回报率为14.3%（见图4-1）。



注: ① 包括股票分拆和股票分红。

② 同行公司包括 Adelphia Communications 公司、American Television & Communications 公司、Cablevision Systems 公司、Century Communications 公司等。

图4-1 马隆任职期间, 超高的总回报

资料来源: 证券价格研究中心和TCI年报。

假如在马隆时代的开端, 向TCI投资1美元, 到1998年中, 其价值会超过900美元。如果投资于其他上市交易的有线电视公司, 这1美元价值180美元; 如果投资于标准普尔500指数, 这1美元价值22美元。因此, 在马隆任期内, TCI的表现是标准普尔500指数的40多倍, 是其他同行的5倍。

The Outsiders

约翰·马隆的方法

马隆任期内的有线电视业务是高度的资本密集型的，建设、购买和维护有线电视网络需要大量的现金。当马隆试图通过扩大用户基数来实现规模扩张时，除了TCI强劲的经营现金流之外，他还有三个主要的资金来源：举债、发行股票和出售资产。马隆对每一项来源的使用都有其独到之处。

避税是关键

在有线电视行业，马隆率先积极使用债务手段进行融资。他认为金融杠杆有两个重要属性：放大财务收益，以及通过支付利息来抵扣税款，从而帮助TCI的现金流避税。马隆将总负债与EBITDA的比率定为5倍，并在20世纪80年代和90年代的大大部分时间里保持这一比率不变。庞大的规模使TCI能够将其债务成本降至最低，而马隆具有在70年代中期煎熬的惨痛经历，所以他非常谨慎地构建自己的债务结构，一来降低成本，二来避免交叉担保。这样，如果某个系统出现债务违约，也不会影响整个公司的信用。这种所谓“密封舱”（这个术语源于马隆对所有与航海相关事项的痴迷，他有时也会提到TCI的“折旧涡流”）的做法进一步加大了TCI结构的复杂程度，却为公司在出现下行风险时提供了实质性的保护。

在发行股票方面，马隆都力求十拿九稳，公司偶尔增发股票的时机都正好与其股票创纪录的市盈率先相吻合。正如马隆在1980年的一次接受采访中所说：“我们公司股价最近的上涨为这次增发提供了一个很好的机会。”⁸马隆有理由为自己在发行股票方面的小气感到自豪，并且认为这是他和同行不同的另外一个因素。

马隆偶尔会抓住机会出售资产。对于那些有线电视公司，马隆会冷静地评估：如果上市，其股票价值是多少；如果没有上市，其企业价值又是多少；当二者背离时，他就积极地在两个市场中进行交易。马隆谨慎地让公司处于净运营亏损（NOL）状态，这些亏损来自多年来折旧款和利息抵扣税款的累积，这使得他可以出售公司资产而无须纳税。由于这种税收方面的防护，只要价格够高，马隆

就可以放心地出售部分网络，为未来的增长筹集资金。正如马隆早在1981年对戴维·沃戈所说的那样：“卖掉我们的部分网络也许是有意义的……以相当于10倍现金流的价格卖掉，再用7倍市盈率的价格回购我们的股票。”⁹

公司的另一个主要资金来源是免缴的税款。正如我们所见，税负最小化是马隆在TCI战略的核心组成部分，他将马格内斯应对税收的传统方法提升到了全新的水平。马隆憎恶税收，他认为税收是对自由主义精神的亵渎。马隆将工程思维模式应用到如何实现税负最小化的问题上，就像他在电子工程考试中让信号丢失最小化一样。在马隆任期内，公司的现金流增长了20倍，却从未缴纳过巨额税款。

事实上，马隆不惜投下重金雇用了一些公司内部税务专家。这个内部税务团队每月召开会议，由马隆本人亲自主持，制定最佳的税务策略。马隆出售公司资产，几乎总是为了购买股票，这就是今天自由媒体公司仍持有新闻集团（News Corp.）、时代华纳、Sprint和摩托罗拉公司大量股票的原因；抑或是马隆出售公司资产的目的在于，通过累积的净运营亏损来掩盖收益；同时，马隆还不断地使用最新的税务策略。正如丹尼斯·莱博维茨所说：“除非基于税收方面的考虑，否则TCI不会去处置资产。”¹⁰没有其他任何一家有线电视公司像TCI那样在税务上投入如此多的时间和精力。

鉴于20世纪七八十年代有线电视行业的非凡增长，就资本配置来说，马隆拥有阵容奢华的高回报选项，他以此优化了TCI的结构。对待资本配置，马隆拥有一种冷静理性、近乎外科手术式的方法，他乐于研究任何能提供诱人回报的投资项目，而无视其复杂性或独特性。马隆把工程思维运用其中，寻找那种投资潜力十分明显的标的，并只专注于有高额回报的项目。有趣的是，马隆不使用复杂的电子表格，而喜欢可以通过简单算术就能确认回报率的项目。正如马隆曾经说过的：“计算机需要大量的细节……而我是数学家，不是程序员。我可能正确，但绝不精确。”

收购的简单法则

在决定如何配置TCI的资本时，马隆做出了与同行截然不同的选择。他从不分红，甚至不考虑分红，也很少偿还债务。马隆在资

本支出方面颇为吝啬，在企业收购方面却跃跃欲试，对待股票回购则随机应变。

在90年代中期卫星电视竞争潮流到来之前，马隆发现，除非带来新的收入，否则改善有线基础设施并不会带来任何可量化的利润。对马隆来说，其中的数学逻辑无可辩驳地清晰：如果资本支出越低，现金流就会越高。因此，多年来马隆无视华尔街的请求，一直拒绝升级他的乡村网络。带着一如既往的坦诚，马隆曾经提到：“这些乡村网络是我们的糟粕，我们不会试图重建它们。”¹¹这种态度与其他有线电视公司的领导人迥然不同，后者经常鼓吹自己对新技术的广泛投资。

具有讽刺意味的是，马隆这位最懂技术的CEO，通常是有线电视公司里最后一个实施新技术的人，他更喜欢扮演技术“传播者”的角色，而不是“拓荒者”。马隆深知新技术的应用是多么困难并且代价高昂，他更倾向于等待，让同行们先去证明新技术在经济上的可行性。在谈及20世纪80年代初延缓引入新安装盒的决定时，马隆说：“等一等再投资，不会让我们失去任何主战场。遗憾的是，有线电视技术的拓荒者们常常会被射成刺猬。”TCI是最后一家推出付费节目的上市公司，并且，马隆说服节目供应商由其支付设备费用。

然而，在需要投资的时候，马隆却绝不手软。90年代中期，来自卫星电视的竞争逐渐凸显，马隆是业内最早支持引入昂贵的新型机顶盒的人士之一。这种机顶盒有助于提高频道容量并丰富客户的选择。

毋庸置疑，TCI最大的资本配置渠道当然是收购。如我们所见，作为有线电视网络的买家，马隆既积极进取又恪守原则。这看似矛盾，其实不然。马隆收购的公司比其他任何人都多，事实上，他收购的公司数目比三四个最大的竞争对手收购公司的总和还要多。总的来说，面对长期以来监管的不确定性和潜在竞争威胁的行业特征，这些收购是对有线电视行业的未来押下了重注；1979—1998年，TCI平均每年花在收购上的金额相当于企业市值的17%，其中有5年超过20%。

不过，马隆这个买家会基于企业价值买卖，他很快就开发出一条简单的法则，成为公司收购计划的基石：对节目类资产打折计算

并扣除日常开销，在其现金流易于量化并得到确认的基础上，才会用不超过现金流5倍的价格去收购公司。这种分析可以在一张纸上完成，如果必要的话，甚至可以在餐巾纸的背面完成，而不需要大量的建模或推论。

法则能否发挥作用，重点在于一系列推论的质量和预期协同效应的实现能力。马隆和斯帕克曼训练了他们的运营团队，让团队可以高效地消除新收购带来的不必要成本，从而实现协同。从陷入困境的华纳通信公司手中接下匹兹堡市的特许经营权后，TCI立即将员工削减了一半，关闭了华纳通信公司在匹兹堡市建造的豪华演播室，并将其总部从市中心的摩天大楼搬到了一个轮胎仓库。几个月内，以前没有盈利的华纳通信公司有线电视网络就产生了可观的现金流。

马隆的简单法则让他在机会来临时能够迅速出手。1987年，拥有百万有线电视用户的霍克家族（Hoak Family）决定出售其业务，马隆在一小时内就与其达成了交易。对于不符合法则的交易，马隆也能轻松地弃之而去。资深行业分析师保罗·卡根（Paul Kagan）回忆道：“马隆放弃了规模可观的夏威夷业务，那笔交易只比他的目标价高出100万美元。”

在大型有线电视上市公司的CEO中，只有马隆会在周期性的市场低迷时期，抓住机会购买自己公司的股票。正如莱博维茨所说：“在这一时期内，其他上市的多系统运营商（MSO）中，没有人进行过任何重大的股票回购。”¹²与此形成对照的是，在马隆任期内，TCI回购了40%以上的公司股份。马隆对回购时机的把握非常到位，带来的平均复合回报率超过了40%。

20世纪80年代初与戴维·沃戈的交易，清晰地阐释了马隆关于回购的机会主义哲学：“我们正在评估所有的可能，以便以当前价格回购股票之后能就其差价进行套利，这个差价源于当前市盈率下企业市值和企业真实价值之差。”¹³这些回购提供了一个有用的基准，以评估包括收购在内的其他资本配置方案。如马隆在1981年对戴维·沃戈所说：“我们的单只股票价格只有20美元出头……买下它看起来比购买其他有线电视公司更有吸引力。”¹⁴

在5种资本配置选项的标准菜单中，马隆又增加了第6项：投资

合资企业。没有一位CEO对合资企业的利用像马隆那样积极，或通过合资企业为股东创造如此多的价值。马隆很早就意识到，他可以借助公司规模产生的杠杆效应，获得节目制作方和其他有线电视公司的股权；同时只需要付出很少的增量投资，这些股权就可以为股东增加可观的价值。在出售给美国电话电报公司时，TCI拥有41项独立的股权投资权益，它的大部分长期回报都归功于这些有线电视或非有线电视领域的合资企业。

由于这些合资企业经营业务各不相同，TCI是出了名的难以分析，而且常常将其业务以折扣价出售给有线电视同行。正如戴维·沃戈所说：“要了解TCI，你就不得不去阅读它们所有的‘脚注’，但很少有人这么做。”¹⁵不过，马隆认为，考虑到这些业务多年来创造的巨大价值，这种复杂性只是很小的代价。正如马隆的许多举措一样，入股这些合资企业事后看来都合乎逻辑，但当时它们都有违常理：业界没有其他人利用合资企业来扩大电视网络的所有权，很长时间以后，其他多系统运营商才开始寻求获得节目制作方的股权。

“牛仔”文化

几乎就像《星际迷航》（*Star Trek*）中的斯波克船长一样，马隆特立独行并且精于算计，同时他也成功地创造了强大的企业文化，并造就员工极高的忠诚度。马隆通过激励和自主权的强有力组合做到了这一点。TCI有一个积极的员工持股计划，该计划匹配员工对公司做出的贡献，并且公司内部所有级别的员工都能参与。许多早期员工，据说包括马隆的长期秘书，都成了百万富翁。这种文化培养了高度的员工忠诚度，在马隆掌舵的前16年里，没有一位高级主管离开公司。

TCI的运营奉行高度的去中心化，直到1995年斯帕克曼退休时，公司总部只有17名员工，同时却拥有1 200万个用户。正如马隆以特有的直率所言：“我们不相信中间管理人员，他们都是事后诸葛亮。”公司没有人力资源主管，直到20世纪80年代末才聘用了一名公关人员。莱博维茨将TCI的文化描述为一群崇尚节俭、注重行动的“西部牛仔”，说到那些经营其他大型有线电视公司的美国东部企业家，他们保守和官僚的作风都是TCI人所不能容忍的。

不断追求极致

马隆为快速增长的资本密集型企业打造了一个有效的资本配置模型，它为诸如移动电话、档案管理和通信塔等不同行业的高管们所仿效。本书中的局外人CEO中，和他最像的另一位是顶级数学家辛格尔顿博士。对数学家来说，当变量趋于极限时，往往会出现洞见，马隆也认同这一点。TCI的特点就是不断追求极致。论有线电视行业的公司，它是最大的；论节目成本，它是最低的；论设备维护费用，它是最少的；论公司结构，它是最复杂的；论回报，它无疑是最高的。

马隆对TCI的管理特质是一种苦行僧式的。从追求规模到最小化税负，再到积极利用财务杠杆，公司战略的每一个要素都旨在优化股东回报。很多人在研究是什么方法驱动着TCI的发展，马隆总结道：“至少，无论如何……我们还没有摒弃算术法则。”马隆的股东们对此将永远心存感激。

特质5 勇于逆向而行

凯瑟琳·格雷厄姆和《华盛顿邮报》

The Outsiders

- 她是“美国报业第一夫人”“世界上最有权势的女人之一”。在凯瑟琳的领导下，公司股东在22年中获得了89倍的惊人收益。
- 通过回购40%的流通股，她是为其股东增加了巨额财富。在麦肯锡专家的建议下，公司停止收购，而其子唐纳德估计，麦肯锡的建议让公司股东少赚了数亿美元。



她是

凯瑟琳·格雷厄姆

KATHARINE GRAHAM

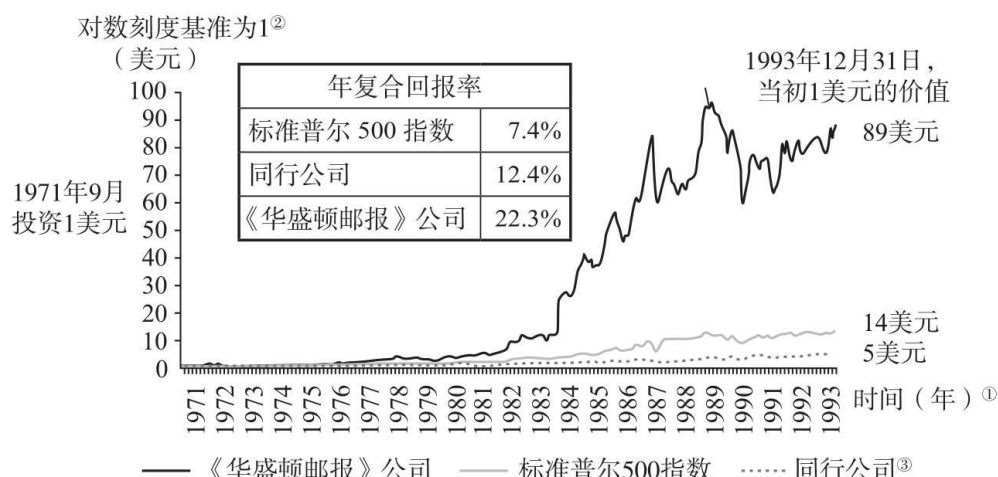
戴维·斯文森
(David Swensen)
耶鲁大学投资基金
首席投资官

不因循守旧需要打破常规……
但是在世俗智慧的眼中，它们
却经常变成彻头彻尾的无厘头。

凯瑟琳成为《华盛顿邮报》董事长兼CEO的道路极不寻常。凯瑟琳是著名金融家兼《华盛顿邮报》公司的老板尤金·迈耶（Eugene Meyer）的女儿，在一个有仆人、寄宿学校、乡村别墅和国际旅行的优越环境中长大。1940年，凯瑟琳嫁给了菲利普·格雷厄姆（Philip Graham）。菲利普是一位从哈佛毕业的杰出律师，同时也是美国最高法院大法官费利克斯·弗兰克福特（Felix Frankfurter）的助理。1946年，迈耶任命菲利普管理公司，其间菲利普在经营公司的过程中不时展露出才华，直到他1963年自杀身亡。在菲利普去世后，凯瑟琳发现自己被意外地推上了CEO的位置。

不用说，凯瑟琳对接手这个职位毫无准备。凯瑟琳当时46岁，是4个孩子的母亲，差不多20年前从第一个孩子出生后，就没有朝九晚五地上过班。在菲利普去世后，凯瑟琳突然发现自己成为《财富》500强公司CEO里唯一的女性，天性腼腆的她被吓坏了。凯瑟琳的故事奇异非凡，但也已经家喻户晓，迄今为止最好的版本是凯瑟琳的自传。凯瑟琳的自传《我的一生略小于美国现代史》（*Personal History*）于1997年出版，曾获得普利策奖。

相比之下，凯瑟琳为股东所做的一切却鲜为人知。从1971年《华盛顿邮报》公司首次公开募股到1993年凯瑟琳卸任董事长，股东的年复合回报率高达22.3%，令标准普尔500指数（7.4%）和同行公司（12.4%）相形见绌。公司首次公开募股时投资的1美元，到凯瑟琳退休时价值89美元，而同样的1美元投资标准普尔500指数和其他同行分别得到5美元和14美元。如图5-1所示，凯瑟琳超越标准普尔500指数18倍，是同行的6倍以上。在这22年的时间里，凯瑟琳以绝对优势位列全美最佳报业高管。



注：① 出于制图的考虑，假定凯瑟琳的任期从 1971 年《华盛顿邮报》公司上市时开始，而不是 1963 年。

② 包括股票分拆和股票分红。

③ 同行公司包括甘尼特公司 (Gannett Co.)、奈特·里德公司、美国媒介综合集团 (Media General)、纽约时报和时代镜报公司 (Times Mirror)，基于各公司股票市值的加权平均值。

图5-1 凯瑟琳执掌公司期间，《华盛顿邮报》股东的总回报

资料来源：证券价格研究中心。

耐心的价值，“不为”才是管理者的最大考验

1963年9月20日，在凯瑟琳的挚友肯尼迪总统遇刺前两个月，她就任《华盛顿邮报》公司总裁。这家公司曾在菲利普领导下发展迅速，并拥有一系列媒体行业的资产，包括日益增长的华盛顿市场中的3家报纸之一《华盛顿邮报》、《新闻周刊》杂志以及佛罗里达州和得克萨斯州的3家电视台。

在接下来的几年里，凯瑟琳花时间适应新岗位，并熟悉公司、董事会和管理团队。1967年，当凯瑟琳首次做出重大的人事决定时，人们才感觉到她的存在，凯瑟琳启用《新闻周刊》时年44岁、性情急躁、经验稍显不足的助理总编本·布拉德利 (Ben Bradlee) 取代了长期担任《华盛顿邮报》主编的拉斯·威金斯 (Russ Wiggins)。

1971年，在董事会的建议下，凯瑟琳批准了公司的上市决议，以期收购筹集资金。在新股发行的那周，《华盛顿邮报》公司卷入了五角大楼文件危机。这份文件是五角大楼内部对越南战争做出的极具争议且负面的评估报告，法院此前已经禁止《纽约时报》刊登这份文件。对

《华盛顿邮报》公司来说，这是机遇更是挑战。尼克松政府担心反战宣传会激起新的浪潮，威胁说《华盛顿邮报》如果刊登这一报告，其广播电视执业许可证就可能被吊销。如果广播电视执业许可证被吊销，新股认购就会被终结，并威胁到公司一个主要的盈利点。面对公司法务部门含混不清的建议，凯瑟琳孤立无援，不得不独自做出决定。凯瑟琳下令按计划行事并刊登这篇报道，由此奠定了《华盛顿邮报》在全行业的美誉。最终尼克松政府没有吊销其广播电视执业许可证，股票认购也取得成功，共募集到1 600万美元。

1972年，《华盛顿邮报》公司在凯瑟琳的全力支持下，开始对共和党总统竞选中的不端行为展开深入调查，最终导致水门事件的爆发。布拉德利和两位年轻的调查记者卡尔·伯恩斯坦（Carl Bernstein）和鲍勃·伍德沃德（Bob Woodward）率先报道了这一惊人丑闻，最终导致尼克松总统于1974年夏天辞职。这场由新闻业掀起的政变为《华盛顿邮报》赢得了普利策奖（布拉德利担任编辑期间获得的18次普利策奖之一），使其成为唯一可以和《纽约时报》比肩的媒体。尼克松政府的咆哮和威胁在整个水门事件中持续不断，凯瑟琳坚决不予理睬。

凯瑟琳用公开募集的一部分资金收购了总部位于新泽西州的《特伦顿时报》（*Trenton Times*），这次收购后来被证明成绩平平。与《特伦顿时报》激烈竞争、一决雌雄的是另一家晚报《泰晤士报》，它们为了生存而竭尽全力。凯瑟琳从这段经历中吸取了宝贵的教训，在未来的收购中更加小心谨慎。

1974年，一位名不见经传的投资者开始不断购入《华盛顿邮报》公司的股票，最终购买了总股份的13%。凯瑟琳不顾董事会的建议，会见了这位新人巴菲特，并邀请他加入董事会。巴菲特很快成为凯瑟琳的商业导师，并帮助她领导公司走上了一条非凡之路。

1975年，《华盛顿邮报》公司面临由强大的印刷工会领导的大规模罢工，他们放火烧毁了印刷厂并宣布罢工开始。与巴菲特和其他董事会成员协商后，凯瑟琳决定抵制罢工。仅仅错过了一天的出版时间，她就和布拉德利以及她27岁的儿子唐纳德组织起骨干成员继续生产，他们连续发行报纸长达139天，直到印刷工会最终同意做出重大让步。

这次罢工对所有参与的人来说都是一次痛苦的经历。有一次，一名工会纠察队成员穿着一件衬衫，上面写着“菲利普误杀了格雷厄姆”[\(13\)](#)。但这些让步妥协却极大地提高了《华盛顿邮报》的盈利能力，

也意味着整个行业的转折点：一家大都市的报社粉碎了一次罢工，这绝对是第一次。对凯瑟琳而言，就像水门事件对公司业务的影响，此次罢工也成为她个人的转折点。从此往后，再也无人质疑谁才是《华盛顿邮报》公司的当家人。

在巴菲特的指导下，凯瑟琳做出了另一个非常规的决定，她开始大举购买自己公司的股票。除了亨利·辛格尔顿和汤姆·墨菲，当时很少有人能有如此胆识和远见。在接下来的几年里，凯瑟琳以最低价回购了公司近40%的股份。值得注意的是，其他主要报业公司的同行中没有一家效仿凯瑟琳的做法。

1981年发生了两起重大事件。首先，《华盛顿邮报》的长期竞争对手《华盛顿星报》（*Washington Star*）在发行量多年下滑后最终停刊。这使得经过罢工洗礼、成本结构精简的《华盛顿邮报》垄断了美国首都的日报发行市场，发行量和盈利能力大幅增长，并持续了整整10年。

第二起事件更为重要。在20世纪70年代，经过4次尝试，凯瑟琳终于找到了一位强有力的COO迪克·西蒙斯（Dick Simmons）。西蒙斯曾是另一家多元化媒体公司邓白氏（Dun & Bradstreet）的COO。西蒙斯就任《华盛顿邮报》公司的COO后，立即理顺了公司的运营部门，而此前这些部门的利润率低于同行。西蒙斯的到来开启了报社盈利能力大幅提升的时代，而且在局外人CEO们的成功中，进一步凸显了那些强大的运营副手所起的关键作用。

在凯瑟琳的支持下，西蒙斯引进新的管理人才，他剥离了《特伦顿时报》，改变薪酬结构以提升奖金的权重，并保持领先于同行的强劲表现。几年内，《华盛顿邮报》公司报纸和电视业务的利润率几乎翻了一番，盈利能力飙升。

20世纪80年代报刊行业的并购交易达到新的历史高度。随着公司利润和市盈率倍数的急剧上升，并购价格如火箭般直插云霄。在各大报业公司高管中，只有凯瑟琳一人置身事外。《华盛顿邮报》公司密切关注了包括艾奥瓦州、得克萨斯州和肯塔基州大型报业公司在内的多起交易，但只进行了两次收购，而且规模都很小。必须认识到，在20世纪80年代中后期炙热的并购气氛中，这种克制是多么难能可贵。这是一段孤独的旅程，凯瑟琳的独善其身招来了同行和媒体的大量非议。在这个掌控着舆论、实行俱乐部制度并由男性主导的行业中，作为唯一的女性高管，凯瑟琳几乎无立身之所。

大出击，市场低潮往往意味着巨大机会

更为重要的是，在凯瑟琳的领导下，《华盛顿邮报》大部分的并购将其引入了与报纸或广播业无关的新领域。1983年，在西蒙斯的一位新雇员、前管理顾问艾伦·斯普恩（Alan Spoon）进行广泛调研的基础上，《华盛顿邮报》公司成功进军移动电话业务，以2 900万美元买下了包括底特律、华盛顿和迈阿密在内的6个大都市市场的特许经营权。1984年，凯瑟琳收购了斯坦利·卡普兰公司（Stanley Kaplan）的考前辅导业务，在教育市场站稳了脚跟。最终，在1986年，由于巴菲特的及时引荐，凯瑟琳进行了她有史以来最大的一笔收购：以3.5亿美元收购了大都会通信公司的有线电视业务资产。接下来的几年中，对《华盛顿邮报》公司来说，以上每一笔收购都将被证明是至关重要的。

1988年初，随着移动电话业务价值的飞涨，凯瑟琳开始意识到扩展移动电话系统需要巨额的资本支出。于是凯瑟琳进行了一次罕见的资产剥离，决定以1.97亿美元的价格出售公司电话业务的资产，这为其带来了非凡的投资回报。

在20世纪90年代初的经济衰退期，当那些同行因为过度负债而被迫抛售副业时，《华盛顿邮报》公司却一反常态开始了收购。利用极端下跌的价格，凯瑟琳不失时机地买入有线电视网络、表现不佳的电视台以及一些教育业务。

1993年，当凯瑟琳辞去公司董事长时，《华盛顿邮报》公司无疑是报业同行中经营最为多元化的公司，它几乎一半的营收和利润都来自非印刷业务。这种多元化经营，为公司未来在凯瑟琳的儿子唐纳德的领导下再创辉煌奠定了坚实的基础。

凯瑟琳在安排接班人方面做得很出色，这在家族企业里并不多见。从20世纪80年代末到90年代初，凯瑟琳一直在为公司的下一代领导人接班做准备，他们包括：唐纳德，他在1991年接替凯瑟琳担任CEO；斯普恩，1991年接替他的导师西蒙斯担任首席运营官。随着凯瑟琳在76岁时辞职，那些成为经理人的青年才俊也将在公司日益重要的部门担任领导角色，包括由汤姆·迈特（Tom Might）负责的有线电视部门及由乔纳森·格雷尔（Jonathan Grayer）负责的教育部门。他们的专业学识和领导能力也为《华盛顿邮报》公司在接下来的15年里取得远超同行的优异业绩奠定了基础。

The Outsiders

凯瑟琳·格雷厄姆的方法

在巴菲特的指导下，凯瑟琳证明了自己是效率很高但绝非正統的资本配置者。在这一重要领域中，凯瑟琳的做法有如下特点：分红和负债方面行业排名靠后，股票回购方面行业排名靠前；相对来说，并购较少，对待资本支出十分谨慎。现在我们对以上特点逐一分析，首先从公司的资金来源开始。

在凯瑟琳的任期内，《华盛顿邮报》产生了持续强劲的现金流，盈利能力在整个20世纪80年代大幅提高，这得益于报纸收入在《华盛顿星报》倒闭后猛增，以及西蒙斯提升了所有运营部门的利润率。除了这一波现金流，《华盛顿邮报》偶尔也会通过另外两个渠道来获得资金：债务杠杆和出售资产。

凯瑟琳通常对使用债务杠杆持谨慎态度，在她的任期内，《华盛顿邮报》不仅始终拥有着同行中最保守的资产负债表，而且借入巨额资金的次数屈指可数。其中，最大的一笔借债是1986年为了买下大都会通信公司的有线电视网络。不过，《华盛顿邮报》强劲的现金流使得这笔债务的大部分款项在不到3年的时间内就还清了。

在凯瑟琳的领导下，就像巴菲特的伯克希尔-哈撒韦公司一样，《华盛顿邮报》很少出售其运营业务，也主动避免公司分拆，它更倾向于拥有长期直接所有权的模式。唯一一次重大例外是1988年初，《华盛顿邮报》决定出售移动电话业务的资产，由此得到了非凡的投资回报。

极度谨慎，严守准则

对现金的配置，凯瑟琳非常谨慎。由于认识到分红就要纳税从而导致资金效率低下，所以在整个任期内，凯瑟琳一直保持着最低水平的分红。同样值得强调的是，这种方法可谓“冒天下之大不韪”，特别是在报业领域，创始人家族通常高度控股，一些家族成员就依靠分红作为收入。在凯瑟琳的领导下，《华盛顿邮报》的分红水平一直是行业最低，但也因此拥有行业最高的留存收益。

凯瑟琳配置收益资金的方法受到了西蒙斯、巴菲特和大都会通信公司高管丹·伯克的影响。所有资本支出的决定都要遵循严格的审批程序，支出的资本必须带来有吸引力的投资回报。正如斯普恩总结的那样：“严格的审批流程将资金配置原则提升到一个新的高度，所有多余的现金都汇给了总公司，经理人必须为所有资本支出项目给出充分的理由。其中的关键问题是：下一美元最好用在哪儿？在回答这个问题时，公司态度严谨而又审慎。”¹

谨慎配置资金原则使凯瑟琳对实体工厂的投资采取了比同行更为谨慎的态度。在20世纪80年代，为了缩短交货期并实现彩色印刷，其他大型报业公司花费数亿美元安装了新的印前和印刷设备。凯瑟琳在各大报业的CEO中独树一帜，她“抱残守缺”，最终在主要出版商中，《华盛顿邮报》成为最后一家依赖老式凸版印刷设备的公司。各种成本下降之前，凯瑟琳一直推迟对新工厂的巨额投资，其经济效益得到了同行的明确认可。

凯瑟琳的收购方法具有双重主题：耐心和多元化。凯瑟琳开始掌舵《华盛顿邮报》的那几年，从首次公开募股开始，传媒公司的盈利和市盈率普遍上升，但为其后两次严重的熊市所打断：一次是在20世纪70年代中期，另一次是在90年代初。换句话说，媒体资产的价值在凯瑟琳任职期间有着大幅的波动。在翻滚的潮流中，凯瑟琳证明了自己是一位精明的领航者。

和宏观的经济图景相比，凯瑟琳的活动如同镜像世界，即所谓“人弃我取”。职业生涯开始和结束时正逢两次大熊市，凯瑟琳也分别在这两个重要时期完成了股票回购和公司收购，而其职业生涯的大多数时间则奉行“无为而治”。

在董事会，凯瑟琳对所有的交易预案都进行了严格的分析测试。汤姆·迈特对此总结道：“在无杠杆的情况下，所有的收购需要在10年的持有期内每年获得至少11%的现金回报。”同样，这个看似简单的测试被证明是一个非常有效的筛选器。迈特还说：“极少有交易通过这个筛选器。公司的整体收购风格就是：等风来，再借东风。”²

我们看到的是，在80年代的收购狂潮中，面对大大小小为数众多的报业并购机会，凯瑟琳通常选择置身事外。凯瑟琳的儿子唐纳德曾经说过：“那些没有进行的交易其实非常重要，否则，昨天的

另一家大型报业公司就是今天卡住我们脖子的船锚。”³在接下来的10年中，《华盛顿邮报》唯一的报业投资只是收购了考尔斯传媒（Cowles Media）的少数股权，考尔斯传媒是《明尼阿波利斯明星论坛报》（*Minneapolis Star Tribune*）和几家小型日报的发行商。基于凯瑟琳和考尔斯家族之间的长期关系，这笔交易是在一次拍卖会的场外以低价转让的方式来完成的。

在对“一美元”原则的贯彻中，巴菲特发挥了关键作用，他仿佛凯瑟琳资本配置法院里的上诉法官，帮助凯瑟琳权衡所有涉及资本投资的重大决策，尤其是在并购事务上。然而，巴菲特的风格并不是直接插手，根据资深董事会成员、顶级律师事务所克拉瓦斯斯韦恩（Cravath, Swaine）的合伙人乔治·吉莱斯皮（George Gillespie）所说的：“他从来不会说‘别那样做’，而会婉转地表达：‘出于这些理由，我可能不会那样做，但我会支持你的任何决定。’”⁴巴菲特话虽如此，但是他给出的理由总令人信服，并鼓励凯瑟琳在资本配置方面保持克制。

在凯瑟琳的主导下，《华盛顿邮报》公司的大部分收购引领公司进入了与报纸或广播无关的新业务领域，这些领域的竞争不那么激烈，估值也更为合理。在这些多元化的收购中，最重要的是收购斯坦利·卡普兰公司的考前辅导业务以及1986年对大都会通信公司的大型收购，前者让公司在教育市场站稳脚跟，后者为公司进入高速增长有线电视业务领域提供了切入点。

《华盛顿邮报》公司与大都会通信公司的这笔有线电视业务交易不仅规模庞大，而且时机恰到好处，就像与考尔斯传媒的交易一样，它们揭示了凯瑟琳严守准则的收购之法。在大都会通信公司收购美国广播公司之后，美国联邦通信委员会责令其必须剥离有线电视业务。获知消息的巴菲特遂安排《华盛顿邮报》独家报道了这笔交易。凯瑟琳意识到这是一个潜力巨大、外人难得的机会。经过一个周末疯狂的准备，凯瑟琳和她的团队同意以每名用户1 000美元的低价收购该业务。更为重要的是，没有投资银行家参与其中。

不失时机的收购

在20世纪90年代早期的经济衰退期，凯瑟琳一反常态地大举买入行为也颇具深意。凭借异常强劲的资产负债表，在同行们由于杠

杆过高而被迫离场观望时，凯瑟琳却成为积极的买家。得益于极度下跌的价格，《华盛顿邮报》不失时机地收购了一系列乡村有线电视网络、得克萨斯州几家表现不佳的电视台以及一些教育业务。所有这些收购，事后都被证明对股东有极大的好处。

如我们所见，股票回购是凯瑟琳另一个主要的资本配置渠道。巴菲特对股票回购原理曾有过令人信服的解释。一旦时机成熟，凯瑟琳就启动回购计划，并尽力推行下去。通过购买大量股票最终几乎达到40%，为股东增加了巨大的价值，这些股票大部分是在20世纪70年代和80年代初在市盈率保持个位数时回购的。

正如东南资产管理公司（Southeast Asset Management）的专业投资者罗斯·格洛茨巴赫（Ross Glotzbach）所说：“并非所有的回购都一样，它们的区别在于，凯瑟琳买入的量并且时机合适。”⁵如图5-2所示，凯瑟琳是报业高管中唯一积极回购公司股票的人，为此，她必须克服董事会最初的巨大阻力。正如吉莱斯皮所说：“在那些日子里，回购股票的公司绝对是异类。”⁶巴菲特原本持有《华盛顿邮报》13%的股份，但从未卖出过一股。由于回购，2011年时，巴菲特拥有《华盛顿邮报》超过22%的股份。

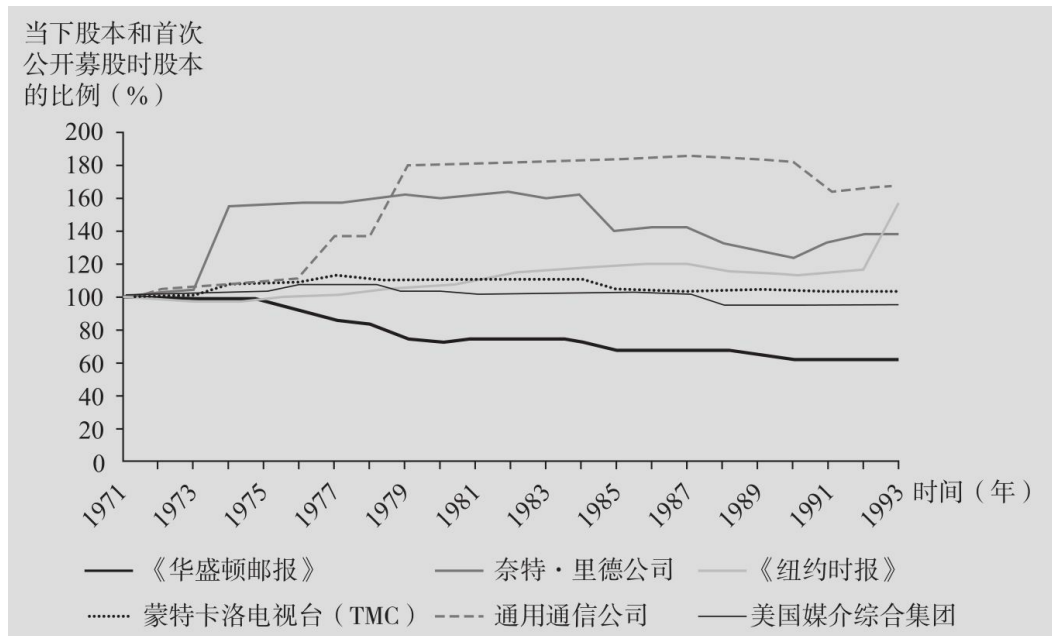


图5-2 《华盛顿邮报》是其同行集团中唯一大比例回购（38.5%）已发行股份的公司

资料来源：证券价格研究中心。

具有讽刺意味的是，20世纪80年代初，管理咨询公司麦肯锡建议《华盛顿邮报》停止股票回购计划。凯瑟琳在两年多一点的时间

里听从了其建议，后来在巴菲特的帮助下恢复了理智，并于1984年重启了回购计划。据凯瑟琳的儿子唐纳德估计，麦肯锡的“智慧”策略让《华盛顿邮报》的股东损失了数亿美元，他称之为“有史以来最昂贵的咨询”。

独特的识人能力

让我们从更广泛的角度去看待资源配置的问题。凯瑟琳拥有一项标志性的管理特质，即独特的识人能力，她能够识别和吸引人才加入她的公司或董事会。虽然看起来冷漠清高，但凯瑟琳有识人天赋，正如斯普恩对我所说：“凯瑟琳是一个完美的‘召集人’。”⁷凯瑟琳选择的这些人显然非同小可，其中一些人已经是功成名就的典范，包括美国国防部前部长罗伯特·麦克纳马拉等董事，以及克拉瓦斯斯韦恩律师事务所的一批律师。这其中有两位最为突出。

首先，在1967年，凯瑟琳聘请了本·布拉德利，让这位年轻且相对而言没什么名气的《新闻周刊》助理总编接替了长期担任《华盛顿邮报》主编的拉斯·威金斯。这个决定太不可思议了！一家报业公司的负责人做出的最重要的人事决定，可能就是选择一名对报纸内容和基调全权负责的总编。凯瑟琳得出结论，她需要一位更年轻的总编，一位更能适应20世纪60年代末迅速变迁的政治格局和文化的总编。当凯瑟琳第一次在午餐时向一本正经的布拉德利提及这个想法时，布拉德利的回答堪称经典：“我愿意把余生献给这份工作。”布拉德利之前没有报纸管理经验，他和杰出的、有教授风范的威金斯有着天壤之别。

然而，布拉德利证明了自己是一个行事果断且直觉敏锐的总编，他用独家新闻带领《华盛顿邮报》度过了70年代的辉煌和动荡，独家新闻也成为吸引年轻的、顶尖新闻才俊加入报社的磁石。随着人才的涌入，加上第一次引入美式风格的版面等诸多创新，70年代到80年代的《华盛顿邮报》发行量持续上升，成为公司利润增长的关键引擎。

凯瑟琳的第二个用人决策发生在1974年熊市最低迷的时期，当时一位名不见经传的投资人累计买入《华盛顿邮报》公司大量的股票，引起了公司上下广泛的焦虑，董事会尤其怀疑这位新人和他买入股票的意图。凯瑟琳承袭的董事会由当地富有经验的商人及其丈

夫菲利普的亲信所组成，尽管她在新闻事务上越来越自信，但仍然要经常听命于董事会对商业事务的判断和建议。

但是这一次，凯瑟琳独自一人决定与这名新人进行会面。值得称赞的是，凯瑟琳立刻意识到了这名新人的独特才能，并不顾董事会的建议，邀请他加入董事会，从而确保他成为生意上的重要密友和导师。凯瑟琳对这名新人只有一点要求：“要绅士，不要辜负我的信任。”⁸

当时，欢迎巴菲特加入董事会的决定是力排众议和不同寻常的决定。20世纪70年代中期，几乎没人知道巴菲特是谁。再强调一次，对任何一位高管来说，导师的选择都是至关重要的决定，而凯瑟琳的选择既不落俗套又出类拔萃。正如凯瑟琳的儿子唐纳德所言：“提拔这个默默无闻却又天赋异禀的家伙，是母亲做过的最好的但也是相对不太受欢迎的举措之一。”

第三次人事决策偏向传统，但同样重要，凯瑟琳于1981年聘用了西蒙斯。鉴于凯瑟琳缺乏运营经验，首席运营官职位的人选就是一个特别关键和困难的问题。凯瑟琳花了很长时间去寻觅合适人选。在进行人事变更时，凯瑟琳坚决果断不拖泥带水，找到西蒙斯之前，在整个20世纪70年代她就解雇了4名首席运营官。西蒙斯在接手首席运营官后的10年里，着手大幅收紧运营，并提高报纸和广播部门的利润率。

西蒙斯和凯瑟琳一样倾向于去中心化的管理。在凯瑟琳的支持下，西蒙斯几乎没有浪费时间，就把合适的人员安排在运营部门的关键岗位上。西蒙斯还赋予经理人很大的自主权，并将他们的薪酬与绩效挂钩。

《华盛顿邮报》一直有去中心化和总部员工精简化的历史传统，这可以追溯到尤金·迈耶和菲利普时期，他们二人对自己的判断都非常自信，觉得没有必要让自己身边环绕着大群顾问。在凯瑟琳和西蒙斯的领导下，这种趋势变得更加专业化和越发明显。凯瑟琳努力找出最优秀的人选，然后非常放心地交给他们独自打理。在巴菲特和丹·伯克的影响下，凯瑟琳的儿子唐纳德继承并发扬了母亲对去中心化一贯的重视。唐纳德告诉我：“我相信，今天的《华盛顿邮报》是全美管理权最为分散的企业之一。”

曾经目睹了父亲把《华盛顿邮报》的日常管理交给年仅31岁的丈夫菲利普，凯瑟琳也就不再惧怕把责任交给那些朝气蓬勃、前途可期的经理人。后来出色地管理着公司有线电视业务的迈特，在负责管理公司有史以来最大的资本项目时年仅30岁。该项目是在弗吉尼亚州斯普林菲尔德（Springfield）建造一个新印刷厂，迈特用低于预算30%的支出完成了项目。斯普恩被任命为公司首席运营官时39岁，乔纳森·格雷尔被任命为教育部门CEO时年仅29岁，而唐纳德被任命为《华盛顿邮报》出版部门的负责人时也只有33岁。

乐于做出有争议、非常规的决策

尽管取得了成功，但是在担任CEO近30年的任期内，凯瑟琳偶尔也会受到自我怀疑的困扰。幸运的是，凯瑟琳意志坚强、性格独立，并且敢于做出那些有争议的、非常规的决策，比如拒绝顽固的罢工者的要求，或是拒绝屈从于尼克松政府的一再威胁，抑或是当其他报业高管质疑她沉迷于股票回购却在收购上畏首畏尾时，她都置之不理。最后，这位初出茅庐的CEO以她那不凡的风格和气派，成就了足以令同行羡慕的新闻和金融传奇。布拉德利带着怅然若失(14)的微笑对我说：“凯瑟琳是那么有趣！”⁹

两家公司的故事

在过去的20年里，很难找到一个行业像报刊业那样迅速地从宠儿变为弃儿。报刊业曾经是巴菲特心中坚不可摧的、拥有“宽阔护城河”的企业典范，在本土广告市场中拥有无可撼动的竞争优势。但这个行业长期以来一直处于衰退之中，面对诸如谷歌等互联网广告参与者带来的新一轮竞争，最大的报社也难以保持盈利。在过去的几年中，几家大型连锁报业公司宣布破产，该行业的股价也反映了这种长期、剧烈的下跌，过去8年下跌超过60%。

即使在唐纳德的领导下，《华盛顿邮报》也未能幸免于这些行业的逆流影响。不过，由于非出版业务总体表现强劲，《华盛顿邮报》的表现仍然成功超越同行，同期市值仅缩水40%。

历史再次证明，CEO们只能根据手里的牌来见机行事。作为大型报业公司CEO的唐纳德拿到了一副烂牌，尽管多亏了他的母亲凯瑟琳在多元化方面的努力，这副手牌明显好于他的同行。不过，对

于这副烂牌，唐纳德却打出了远远超越同行的水平，这得益于他对母亲管理方法的坚守：有选择地进行收购，积极地不失时机地回购股票，包括2009—2011年已发行股票的20%，以及保持相对较低的分红水平。

相比之下，同一时期，另一家知名的、公开上市的东北部报业公司，由苏兹伯格家族（Sulzbergers）控制的《纽约时报》却高价收购了一家互联网门户网站，又在曼哈顿中心建造了一座精致的公司新总部大楼……然后，它的市值下跌了90%。

特质**6** 数据思维制胜

比尔·斯蒂利茨和罗尔斯顿·普瑞纳公司

The Outsiders

- 他是靠玩扑克牌攒够大学学费的数据狂人，是“化腐朽为神奇”的品牌猎手。
- “计算赔率”是他心中永远的行为罗盘。在掌权的19年时间中，他为罗尔斯顿·普瑞纳公司股东创造了57倍的收益。
- 2001年，雀巢公司同意为罗尔斯顿·普瑞纳公司支付创纪录的价格，104亿美元，相当于现金流的14倍。



他是

比尔·斯蒂利茨

BILL STIRITZ

沃伦·巴菲特

如果你发现自己身处一条老旧的破船上，那么更有成效的做法是换一条船，而不是花精力去补漏。

在过去50年的大部分时间里，包括金宝汤、亨氏和家乐氏在内的家喻户晓的大型包装商品公司，因其兼具增长可预期、抗经济衰退和分红可靠等诱人特征，而被认为是蓝筹股中的蓝筹股。这些公司长期以来一直是财务保守主义的楷模，几乎不使用杠杆，稳定地支付股息，同时很少回购股票。在20世纪60年代和70年代，为了追求不切实际的协同效应，它们中的大多数公司追随时尚，积极进行多元化经营，许多公司最终进入了餐饮和农产品行业，以寻求“垂直整合”那莫须有的好处。

罗尔斯顿·普瑞纳公司在这个群体中堪称典型。80年代初，罗尔斯顿·普瑞纳公司曾名列《财富》100强，在农产品饲料领域有着悠久的历史。70年代，在CEO哈尔·迪安（Hal Dean）的领导下，该公司走的是与同行一样的道路，利用传统饲料业务提供的巨额现金流从事多元化项目，这让公司的运营部门变成了一锅大杂烩：从蘑菇农场、大豆农场到Jack in the Box的快餐连锁店，从圣路易斯蓝调曲棍球队（Blues Hockey Team）到科罗拉多州的基斯通滑雪度假村（Keystone）。当迪安在1980年宣布退休时，罗尔斯顿·普瑞纳公司的股价已经原地踏步了10年。

在迪安宣布消息后，罗尔斯顿·普瑞纳公司董事会为了找到迪安的接班人展开了全面的搜寻，候选者包括公司内部和外部的一大批人。随着遴选工作的展开，候选名单上出现了全美的一些顶尖人才，包括后来担任哥伦比亚广播公司CEO的汤姆·怀曼（Tom Wyman）。在这个过程接近尾声的时候，一位不太知名的候选人主动向董事会提交了一份备忘录，详细阐述了对公司的战略构想。这位候选人只是公司的一位老员工，甚至在内部候选人中都排名靠后，但是这一举动却成为他极大的加分项。玛丽·韦尔斯·劳伦斯（Mary Wells Lawrence）当时在董事会内颇具影响力，她还是韦尔斯-里奇-格林（Wells, Rich, Greene）广告公司的创始人，读完备忘录后她回电4个字“直击要害”。几天之后，候选人斯蒂利茨就得到了这份工作。

数据狂人，分析力就是领导力

斯蒂利茨的职业轨迹与本书中其他局外人CEO不同。斯蒂利茨是个

内行，在47岁成为CEO之前，他已经在罗尔斯顿·普瑞纳公司工作了17年。然而，这种看似中规中矩的背景，掩盖了斯蒂利茨内心对于独立思考的强烈渴求，尽管谁都觉得不太可能，但他就是成了一个高效的改革推动者。当斯蒂利茨就任CEO时，无人能预见他将在罗尔斯顿·普瑞纳公司实现根本性的转变，也无人能预见他会对食品和包装制品行业同行产生更广泛的影响。

作为一名CEO，斯蒂利茨有着不同寻常的教育背景。他本科阶段有过休学的经历，在阿肯色大学只上了一年就由于囊中羞涩不得不离开学校去海军服役。在海军服役的4年里，斯蒂利茨苦练扑克技艺，凭着精湛的技能最终凑足了学费。服役期满，斯蒂利茨回归校园，在美国西北大学主修商科并获得学士学位。斯蒂利茨从来没上过MBA，尽管同事和华尔街分析师反复给他贴上“理工男”的标签，但他35岁从圣路易斯大学获得的却是欧洲历史学硕士。

从美国西北大学毕业后，斯蒂利茨曾在美国品食乐食品公司（Pillsbury Company）工作，最初是一名店面理货员，主要负责密歇根州北部商店里公司谷类产品的货架陈列，其中一个最大的客户位于印第安人自治区。斯蒂利茨认为，这种一线的“草根”经验对于帮助他理解分销渠道的细节至关重要。随后，斯蒂利茨被提升为产品经理，这个岗位使他对快消品营销有了更深入的了解。两年后，为了更好地理解媒体和广告，斯蒂利茨离职并加入了位于圣路易斯的加德纳广告公司（Gardner Advertising）。在加德纳广告公司，斯蒂利茨很快就对营销量化产生了兴趣，并成为新兴的尼尔森评级服务（Nielsen Ratings Service）的首批用户。这些都有助于斯蒂利茨深刻领会市场份额与促销费用之间的关系。

1964年，30岁的斯蒂利茨加入了罗尔斯顿·普瑞纳公司，并被分配到杂货产品部门，负责宠物食品和谷类食品的销售工作。在罗尔斯顿·普瑞纳公司庞大的业务体系中，杂货产品部门历来被认为是“后妈生的”而备受冷落。斯蒂利茨在那里工作了几年，职位逐渐提升，1971年成为该部门的总经理。斯蒂利茨通过不断推出新产品以及扩展产品线使得杂货产品部门的业务量飞速增长，营业利润增长了50倍。

斯蒂利茨亲自策划了普瑞纳狗粮（Purina Puppy）和妙多乐猫粮（Cat Chow）的推介，由此成就了宠物食品行业历史上最成功的两款产品。作为营销人员，斯蒂利茨非常善于分析，对数字天生敏感，并敢于质疑，总是能激怒对手。这些性格特质曾在牌桌上帮过他，也会在CEO

的岗位上继续发挥作用。

品牌大收购，超级分销撬动利润暴涨

1981年就任CEO后，斯蒂利茨没有停下前进的步伐，而是在第一时间积极地重组公司。斯蒂利茨充分认识到公司拥有众多的消费品牌，其经济效益极其诱人，并迅速围绕这些业务对公司进行了重组。在斯蒂利茨眼中，这些业务提供了高利润率和低资本投入的绝佳组合。斯蒂利茨立即调整了前任制定的战略核心，第一步是对那些不符合其盈利和回报标准的业务进行主动剥离。

上任不久，斯蒂利茨就出售了快餐店Jack in the Box、蘑菇农场和圣路易斯蓝调曲棍球队。尤其是蓝调曲棍球队的出售，让华尔街和当地商界警觉到这位新任CEO将采取截然不同的方式来管理罗尔斯顿·普瑞纳公司。

斯蒂利茨继续出售其他非核心业务，包括该公司的大豆业务、大杂烩一样的餐厅和食品服务业务，最终使罗尔斯顿·普瑞纳公司成为一家纯粹的品牌产品公司。在这方面，斯蒂利茨与伯克希尔-哈撒韦公司早期的巴菲特没有什么不同，当时巴菲特把资本从低回报的纺织业务中抽取出来，将其配置在回报更高的保险和媒体业务中。

20世纪80年代初，斯蒂利茨克服了董事会最初的阻力，发起了一项积极的股票回购计划。在主要的品牌产品公司中，只有罗尔斯顿·普瑞纳一家在不断地进行回购。斯蒂利茨认为，回购可以产生确定无疑的回报。在余下的任期内，回购一直是斯蒂利茨资本配置计划的核心原则。

20世纪80年代中期，在首轮资产剥离完成之后，斯蒂利茨进行了两次大规模收购，总金额达到罗尔斯顿·普瑞纳公司市值的30%。两次收购主要的资金来源都是债务融资。

第一次收购为罗尔斯顿·普瑞纳稳定的品牌阵营中添加了大陆烘焙公司（Continental Baking），该公司生产Twinkies品牌的饼干和Wonder品牌的面包。斯蒂利茨从多元化综合企业集团国际电话电报公司手中收购了大陆烘焙公司。作为国际电话电报公司唯一的包装商品公司，大陆烘焙公司此前的业务一直萎靡不振。后来在罗尔斯顿·普瑞纳公司的管理下，大陆烘焙公司欣欣向荣，其分销规模得到扩大，冗余成本得到削减，新产品不断推出，现金流也大幅增长，从而为股东创造了巨大价

值。

接下来的第二次收购发生在1986年，斯蒂利茨进行了他有史以来最大规模的收购，以相当于罗尔斯顿·普瑞纳公司市值的20%，即15亿美元，收购了联合碳化物公司（Union Carbide）的劲量电池业务

（Energizer）。当时，联合碳化物公司在博帕尔灾难事件(15)的余波中苦苦挣扎。联合碳化物公司的电池业务尽管拥有强大品牌，但是长期以来被公司忽视。与国际电话电报公司一样，联合碳化物公司缺乏消费品营销的专业知识，而且急于脱手。斯蒂利茨最终在拍卖中获胜，并当机立断地付清了全款。斯蒂利茨觉得这项资产属于日益增长的双寡头垄断市场，且在前任手中运营管理不善，这二者的结合对他来说极具吸引力。

与对大陆烘焙公司所做的一样，斯蒂利茨立即着手改进劲量电池的产品和营销，包括发起著名的以同名兔子为标识的广告宣传，同时加强分销并削减冗余成本。通过上述一系列行动，斯蒂利茨完成了改造罗尔斯顿·普瑞纳公司的第一步。到了20世纪80年代末，罗尔斯顿·普瑞纳公司的营业收入中来自快消品的占比已经上升到近90%。

这一转变对公司关键的运营指标产生了显著的影响。随着罗尔斯顿·普瑞纳公司的业务组合转向品牌产品，它的税前利润率从9%增长到15%，净资产收益率从15%增长到37%，翻了一番有余。如果再加上股本总数的减少，那么其每股收益和股东回报都有了惊人的增长。

在整个80年代，通过有选择地在剥离资产和企业收购间保持平衡，斯蒂利茨持续优化着他的品牌组合。如果不能带来令人满意的回报，该业务就会被出售或关闭。这些被剥离的资产包括：表现不佳的食品品牌，其中的范德坎普（Van de Kamp）冷冻海鲜部门是斯蒂利茨一次少有的收购败笔；公司历史遗留的农业饲料业务普瑞纳磨坊（Purina Mills），昔日辉煌的它在当时长期回报率低，增长空间也很有限。斯蒂利茨的收购重头戏集中在电池和宠物食品的核心品牌上，特别是在那些还没有发掘国际市场的品牌。以上所有的决定都以提升股东潜在回报为导向，并且做决策之前都要进行仔细的分析。

衍生企业，非主营业务的最佳归宿

在整个20世纪90年代，为了让罗尔斯顿·普瑞纳公司的品牌组合更加合理，斯蒂利茨把握时机专注于下列行动：股票回购、偶尔的并购以

及企业分拆。分拆非常重要，它是一种相对较新的公司结构化工具。斯蒂利茨开始意识到，即使公司结构相对分散，一些业务却依然没有得到公司内部或华尔街应有的关注。为了纠正这一点以及尽量减少税负，斯蒂利茨成为企业分拆的先行者。

在分拆过程中，业务部门从母公司剥离出来，成为新的公司实体。母公司股东在新公司中享有同等比例的股票所有权，并可以自行决定是持有还是卖掉这些股份。更为重要的是，分拆使得小型业务部门的价值得到凸显，部门利益和管理层激励绑定得更加紧密，同时它延缓了资本利得税的缴纳。

斯蒂利茨的分拆计划始于1994年，当时他将包括切克斯谷类食品（Chex）和滑雪度假村在内的一系列小品牌分拆出来，成立了新实体拉尔考普公司（Ralcorp）。该公司有着独立的董事会和两位联席CEO，斯蒂利茨仍然出任新公司的董事长。他继续致力于让罗尔斯顿·普瑞纳公司的业务组合更加合理化，1998年将公司包括快速增长的蛋白技术业务在内的剩余的农产品业务以创纪录的高价出售给杜邦公司，并置换成对方的股票，同时再次避免了资本利得税。

斯蒂利茨最后的一次也是截至当时最大的一笔分拆是2000年对劲量电池业务的分拆，劲量电池当时的企业价值相当于公司总市值的15%。对于独立的上市公司而言，分拆举动后的表现异常出彩，其中拉尔考普公司最初只是集合了一众不被看好的资产，但其在2012年的企业价值已达50亿美元。

在千禧年的曙光中，这一系列的举动让罗尔斯顿·普瑞纳公司成为一家专注于经营宠物食品的公司，在美国同类产品市场上占据着绝对的主导地位。削减不相关的业务让那些战略并购者对公司核心的宠物食品品牌蠢蠢欲动，这一切都没有逃过斯蒂利茨的眼睛。2001年，斯蒂利茨和雀巢公司进行了接洽。经过广泛的磋商，并由斯蒂利茨按照自己的风格亲自主持双方的洽谈会，这家瑞士的企业巨头同意为罗尔斯顿·普瑞纳公司支付创纪录的104亿美元，相当于罗尔斯顿·普瑞纳公司现金流的14倍，这太令人意外了。这桩交易成为斯蒂利茨在罗尔斯顿·普瑞纳公司任职期间的巅峰之作。

在此期间，如果说所有同行的回报为优秀，那么斯蒂利茨的数字当属卓越。在执掌公司的19年中，斯蒂利茨将罗尔斯顿·普瑞纳公司改造成一家高效率的包装商品公司，也推动着公司股价一路飙升。如图6-1

所示，在斯蒂利茨出任CEO时向他投资1美元，19年后价值57美元，年复合回报率为20.0%，轻松超过了同行的17.7%和标准普尔500指数的14.7%。



图6-1 1美元价值变化

The Outsiders

比尔·斯蒂利茨的方法

迈克尔·莫布森（Michael Mauboussin）现在是美盛集团（Legg Mason）备受尊崇的投资人，而在20世纪80年代中期，他在德崇证券（Drexel Burnham）的首份研究任务就是斯蒂利茨和罗尔斯顿·普瑞纳公司。莫布森被罗尔斯顿·普瑞纳公司这位特立独行的CEO迷住了，在导师艾伦·格雷迪特（Alan Greditor）工作的基础上，针对罗尔斯顿·普瑞纳公司撰写了一份早期的全面研究报告。格雷迪特是少有的赢得斯蒂利茨尊敬的华尔街分析师，在其指导下，莫布森得以领会斯蒂利茨那独特的资本配置方法。

当莫布森被要求总结斯蒂利茨的与众不同之处时，他告诉我：“有效的资本配置.....需要一种特定的性情。要想获得成功，你必须像一个投资者那样去心平气和地思考问题，不意气用事而依概率行事。斯蒂利茨具备这种心态。”¹

斯蒂利茨自己把资本配置比作扑克游戏，在那里，取胜的关键技能是计算赔率、洞察人性，以及在己方有压倒性胜算的时候押上重注。斯蒂利茨既是活跃的买家，也乐于出售或剥离那些他认为已经成熟或华尔街不感兴趣的业务。

正如高盛资深分析师诺米·盖兹（Nomi Ghez）向我强调的，食品行业传统上是一个利润丰厚、可预测性强的行业，但是通常维持着低速增长。在该行业上市公司的CEO中，只有斯蒂利茨看清了这些特点的全貌，并找到了提升股东价值的新方法。事实上，他通过一系列举措从根本上改变了行业范式：积极利用债务杠杆来实现更高的每股收益，削减盈利较少的业务，收购关联业务的企业，并积极回购股票。在实践中，他的做法暗合了那些私募股权公司的先锋所采用的技术，比如，KKR公司在一些早期的大型杠杆收购中

（LBO），把业绩不佳的包装商品公司成功地收入囊中，包括比特阿里斯食品公司（Beatrice Foods）和后来的雷诺兹-纳贝斯克公司（RJR Nabisco）。事实上，斯蒂利茨也参加过这两家公司的竞标，但是出价过低。斯蒂利茨还对吉列公司和佳得乐公司出过价，但是也未成功。

积极获取现金

在斯蒂利茨任职期间，罗尔斯顿·普瑞纳公司的主要资金来源是内部现金流和债务；早期的时候，资金来源还包括出售资产所得。

在斯蒂利茨掌舵期间，经营现金流一直是重要的且不断增长的资金来源。在斯蒂利茨的管理下，公司从奉行混合经营到认准品牌战略，同时也认准精简的、去中心化的经营哲学，最终反映为公司利润率的稳步提高。被雀巢公司收购之前，罗尔斯顿·普瑞纳公司的利润率是包装商品行业里最高的。

在快消品行业的CEO中，斯蒂利茨是使用债务融资的先驱。快消品行业长期以来以极端保守的财务管理为特征，对它们来说，债务融资的做法简直就是异端行为。但是，斯蒂利茨却认为谨慎地使用债务杠杆可以显著地提高股东回报率。斯蒂利茨一方面相信那些拥有稳健现金流的行业应该利用债务来提高股东回报；另一方面积极实践，利用债务杠杆为股票回购和企业收购筹措资金，其中就包括他最大的两笔收购，劲量电池和大陆烘焙公司。如图6-2所示，罗尔斯顿·普瑞纳公司在斯蒂利茨的任期内负债对现金流(16)的比例一直高于行业平均水平。

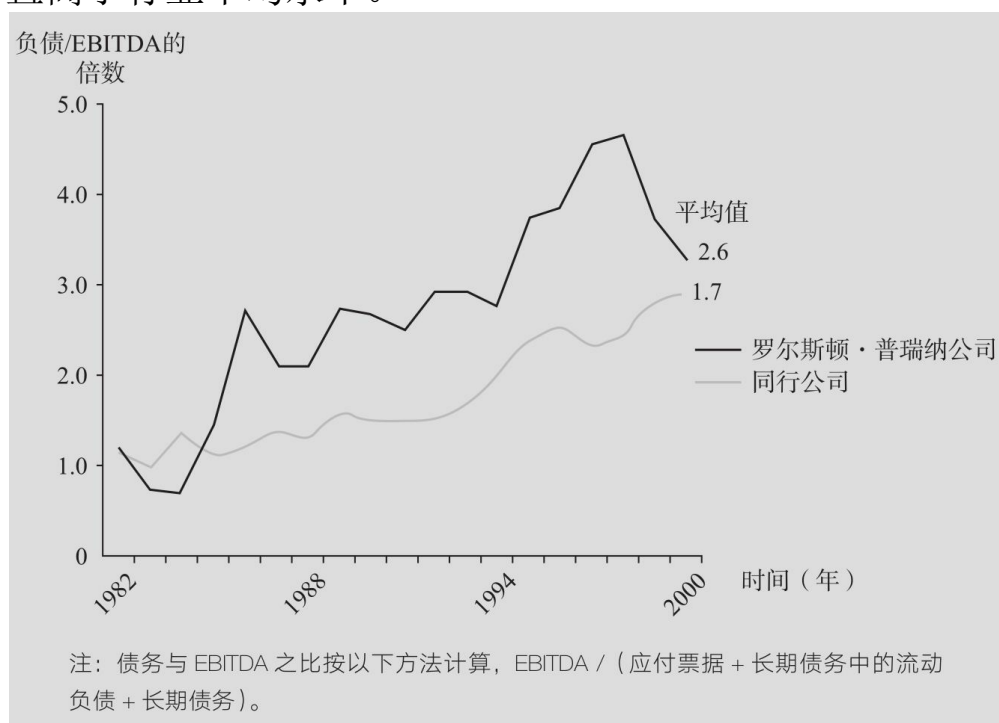


图6-2 罗尔斯顿·普瑞纳公司的负债水平持续高于同行

资料来源：证券价格研究中心和罗尔斯顿·普瑞纳公司文件。

斯蒂利茨的资产出售和剥离方法随着时间的推移而演变。开始时，斯蒂利茨出售非核心业务，比如由于达不到他对盈利能力和回报的要求，蘑菇农场和蓝调曲棍球队被出售，这些资产出售是公司早期重要的现金来源。对斯蒂利茨来说，即使是公司最初发迹的饲料生意也可以拿来出售。“斯蒂利茨知道什么东西是值得的，只要价格合适，他会出售任何资产。”莫布森充满赞许地告诉我。²在此期间，斯蒂利茨专注于以尽可能理想的价格剥离非核心资产，并将资本重新配置到高回报的包装商品业务上，比如买下劲量电池和大陆烘焙这两个品牌。

不过，斯蒂利茨最终意识到，从税负的方面来考虑，出售资产无效率可言。我们也观察到，后期斯蒂利茨开始使用企业分拆的方法，他认为这样既释放了企业家的活力和创造力，同时也延缓了资本利得税的缴纳。从一开始，斯蒂利茨就信奉去中心化的管理，致力于减少公司官僚机构的层级，并将公司关键业务的管理责任和决定权交给紧密协作的管理团队。在去中心化的方向上，斯蒂利茨认为企业分拆是行动的下一步，这种“极致的分权”为管理者和股东提供了信息透明、决策自主的强劲组合。与那些受制于母公司复杂结构的公司相比，分拆后的企业允许管理者更直接地根据其经营业绩来获得报酬。

斯蒂利茨也被证明是一个非常精明的卖家。在20世纪80年代初首轮资产剥离完成之后，斯蒂利茨只进行了两次大手笔的资产出售。首先是向杜邦公司出售了罗尔斯顿·普瑞纳公司的蛋白业务。杜邦公司为这项业务支付了高昂的价格。斯蒂利茨选择了股票作价的方式进行交易，从而延缓了资本利得税的缴纳。其次是向雀巢公司出售宠物食品业务，众所周知，这次交易的价格创下了纪录，超过100亿美元。后来，斯蒂利茨虽然承认这个价格非常划算，但考虑到雀巢公司的实力以及股东缴纳的资本利得税，他十分后悔当初没有选择雀巢公司的股票。

回购股票与收购

除了每年固定的债务偿还、内部资本支出和最低限度的股息支付外，斯蒂利茨把现金主要用在两个方面：回购和收购。在这两个

方面，斯蒂利茨对时机的把握都已经到了炉火纯青的地步。

在快消品行业，斯蒂利茨是股票回购方面的先锋。20世纪80年代初，当斯蒂利茨开始回购股票时，回购行为并不多见且备受争议。罗尔斯顿·普瑞纳公司的一位董事当时就质问斯蒂利茨：“你为什么要缩减公司规模，就没有其他任何有价值的、值得投资的增长点了吗？”相比之下，斯蒂利茨认为回购是他所能进行的最有潜力的投资，在说服董事会之后，他立刻化身为活跃的回购者。斯蒂利茨最终令人震惊地回购了罗尔斯顿·普瑞纳公司60%的股份。在本书的CEO中，这仅次于亨利·辛格尔顿。从这些回购中，斯蒂利茨获得了非常丰厚的回报，长期来看，其平均内部回报率为13%。

同时，斯蒂利茨还是一位精于算计的买家，更喜欢瞅准机会在公开市场上买入，而不是发起大型收购要约，后者在交易之前股票价格可能已经抬升。斯蒂利茨的买入都发生在市盈率处于周期性低点时，为了节省交易佣金，斯蒂利茨甚至亲自就这些回购进行谈判。

斯蒂利茨认为，股票回购的回报率为其他内部资本投资决策特别是收购决策，提供了一个十分便捷的参考基准。斯蒂利茨的长期副手帕特·马尔卡希（Pat Mulcahy）说：“在制定投资决策时，我们经常遇到的拦路虎就是股票回购的回报率。如果收购的回报率在一定程度上能够超过股票回购的回报率，那就值得去做。”³相反，如果收购潜在的回报率没有明显超过回购回报率，斯蒂利茨就会放弃收购。

斯蒂利茨的收购方法是一直寻找优势企业，他专注收购的企业总有以下特征：可以通过罗尔斯顿·普瑞纳公司的营销专长和分销能力对其进行改善。斯蒂利茨更喜欢在前任所有者手中经营不善的公司，而且并非巧合的是，斯蒂利茨最大的两次收购——劲量电池和大陆烘焙公司，同属于大型企业集团内部被忽视的小部门。这两次收购的长期回报都非常出色，劲量电池在14年内实现了21%的年复合回报率；大陆烘焙公司在11年的持有期内实现了13%的年复合回报率。

斯蒂利茨专注于通过与卖家直接接触的方式来进行收购，尽可能地避免参与竞争性拍卖。大陆烘焙公司的收购是斯蒂利茨直接给国际电话电报公司（ITT）董事长兰德·阿拉斯科格（Rand

Araskog) 写了一封信，从而避开了一场竞拍。

斯蒂利茨认为，罗尔斯顿·普瑞纳公司致力于抓住那些在保守估计下仍能获得确定收益的机会。对于详细的财务模型提供的虚假精确性，斯蒂利茨不屑一顾，反而是把重点放在少数几个关键变量上：市场增长、竞争力、改善运营的潜力，以及自始至终非常重要的、产生现金的能力。斯蒂利茨对我说：“我真的只关心模型中代入的关键假设。首先，我必须了解市场的基本趋势，即它的增长和竞争态势。”⁴

斯蒂利茨的门生马尔卡希后来负责这项业务。在谈及产生深远影响的劲量电池收购案时，马尔卡希对斯蒂利茨的策略做出描述：“当时收购劲量电池的机会就在眼前，下午1点我们一小群人见了面，拿到了卖家的标书。在一个信封的背面，我们对杠杆收购模型进行了推演。下午4点，我们再次会面并决定出价14亿美元，就这么简单。我们知道自己需要关注什么。整个过程中没有大量的研究，也没有成群的银行家。”⁵同样，斯蒂利茨的方法类似于汤姆·墨菲、约翰·马隆、凯瑟琳·格雷厄姆和其他局外人，只需要一张纸就可以写满。斯蒂利茨高度关注关键假设，而不是写满40页的各种预测。

账面价值毫无意义

了解到包装商品行业早期的杠杆收购后，斯蒂利茨有意识地采用了类似于私募股权的思维模式。马尔卡希精妙地总结了斯蒂利茨的管理世界观：“斯蒂利茨经营罗尔斯顿·普瑞纳公司有点类似于杠杆收购。他是最早洞察到以下真谛的人之一：只要现金流强劲且稳定，高杠杆率对股东就有利.....他直接舍弃了那些现金不断流失的企业，不管它们来自何处.....并通过大规模的股票回购或偶尔的企业收购来实现对已有强势业务的深度投资，从而达到我们的收益目标。”⁶

斯蒂利茨既有包装商品营销专长又有金融敏锐性，这是一对难得的精妙组合。斯蒂利茨凝神于新兴私募股权行业的通用语言，比如EBITDA和内部收益率（IRR）；不屑于当时华尔街视为圭臬的传统会计指标，比如账面盈余和账面价值。斯蒂利茨视账面价值如草芥。资深分析师约翰·比尔布塞（John Bierbusse）回忆：斯蒂利

茨在难得出席的行业论坛上曾宣称“在我们的业务中，账面价值没有任何意义”，这一声明令在场听众顿时呆住了。莫布森补充道：“在忽略账面价值、每股收益和其他标准的会计指标方面，你必须胆略过人，因为它们总是与经济现实相脱节。”⁷

拒绝外部建议

斯蒂利茨极其独立，对外部顾问所提建议的蔑视不加掩饰。斯蒂利茨认为，个人魅力作为一种管理特质，其作用被过分高估，而分析能力才是成为CEO的先决条件和独立思考的关键。对此，斯蒂利茨直言不讳：“没有分析能力，CEO们将任由那些银行家和财务官员摆布。”斯蒂利茨发现，许多CEO来自诸如法务、营销、制造、销售等职能领域，而这些领域中不太用到这种分析能力。没有分析能力的CEO，被斯蒂利茨认为是严重残疾。斯蒂利茨的忠告很简单：“领导能力就是分析能力。”

这种独立的思维模式转化为与生俱来的怀疑，把那些外部顾问特别是投资银行家排除在信任之外。斯蒂利茨曾将投资银行家描述为“寄生虫”。在使用顾问时，斯蒂利茨目标明确：尽可能少的人数，通常每名顾问的方向都有清晰的针对性。同时，斯蒂利茨在洽谈咨询费时十分较真儿，在觉得那些银行家要价太高时，他停止了和雀巢公司数十亿美元的交易。斯蒂利茨有个观点，即不同的交易要聘请不同的银行家，这样就没有哪个银行家会觉得自己的业务是铁饭碗。

面对重要的尽职调查会议或谈判，斯蒂利茨以独自一人出席而闻名，而谈判桌的另一边却挤满了银行家和律师。对这种有违常规的做法，斯蒂利茨怡然自得。曾在高盛集团就职的一位银行职员告诉我：在并购雷诺兹-纳贝斯克公司的过程中，有一次尽职调查会议安排在深夜，斯蒂利茨一个人进入高盛集团办公楼的会议室，手里只拿着一个黄色的便签簿，他逐一浏览那些关键的运营假设，然后给出最终报价就回去睡觉了。斯蒂利茨热切地享受投资过程中的乐趣，出售罗尔斯顿·普瑞纳公司之后，他精心管理着一家投资合伙企业，资金来源也主要是其自有资金。

斯蒂利茨深知“一寸光阴一寸金”，对曝光度高、耗时长的慈善理事会避之不及，也不参加“主要是浪费时间”的休闲午餐会。斯蒂

利茨解释道：“这些活动发展至今，已经到了占用太多时间的地步，所以我彻底屏蔽了它们。”然而，斯蒂利茨一定会抽出时间出席其他公司的董事会，把这看作是一个让自己接触新情况和新思想的难得机会。

斯蒂利茨是“狐狸”型的人，不断吸收着新的想法，而不问其出处。资深的行业分析师约翰·麦克米林（John McMillin）曾经略带恭维地写道：“有些人是创新者，有些人从别人那里借鉴，斯蒂利茨则兼而有之。”⁸斯蒂利茨有意识地在日程表中挤出一块块时间，不受干扰，独自一人思考业务中的关键问题，无论是在佛罗里达的海滩还是在圣路易斯家中的办公室。

斯蒂利茨避免与华尔街进行耗费时间的沟通，并保持着分析师比尔布塞所说的“影星嘉宝在无声电影中的气质”，斯蒂利茨很少与分析师交谈，几乎从不出席会议，也从不发布季度指南。⁹

到20世纪90年代中后期，斯蒂利茨的另类做法已经变成时代主流，几乎所有的同行都实施了与他大同小异的战略：剥离非核心资产、回购股票、收购与其核心产品线互补的业务。不出所料，2001年，正当斯蒂利茨的战略获得了业界广泛认可，包装商品公司的市盈率达到历史最高水平时，斯蒂利茨突然改变方向，将罗尔斯顿·普瑞纳公司以创纪录的价格卖给了雀巢公司，此举再次让摸不着头脑的竞争对手抓耳挠腮。

最近的一个例子：美国莎莉集团

斯蒂利茨执掌罗尔斯顿·普瑞纳公司30多年后，业界同行的“模仿秀”仍络绎不绝。最近一次“模仿秀”的情形是这样的：在CEO布伦达·巴恩斯（Brenda Barnes）和马塞尔·斯米茨（Marcel Smits）的领导下，消费品公司美国莎莉集团（Sara Lee）追随斯蒂利茨的脚步，在过去5年里抛售了非核心业务，回购了13%的股份，保持了相对较高的杠杆率，为股东带来的回报让同行黯然失色。在撰写本文时，一个由私募股权公司构成的财团对美国莎莉集团发起收购要约，收购价较集团当前的股价有相当的溢价，但是美国莎莉集团却拒绝了。相反，美国莎莉集团宣布将剥离其利润丰厚的咖啡和茶品业务，并一次性派发可观的股息。这是不是似曾相识？

特质7 永远关注现金流
迪克·史密斯和大众影院公司

The Outsiders

- 史密斯的疯狂业绩：34年684倍收益，足以“秒杀”众多同行，笑傲伟大CEO之巅。
- 杰克·韦尔奇本人对史密斯的成就大为震惊，因为其业绩足足是自己的11倍多。
- 史密斯长期稳定地回购大众影院公司的股票，最终买回了公司1/3的股票。在这些回购上，他的长期内部收益率高达16%。



他是

迪克·史密斯

DICK SMITH

戴维·沃戈
普特南投资公司
(Putnam Investments)

由一小群真正的精英创造出的
价值是惊人的。

1962年，菲利普·史密斯因心脏病突发而去世。菲利普于1908年从俄罗斯来到美国波士顿，在打过各种零工之后，在新兴的“五分钱电影院”行当里找到了自己的事业方向。刚开始，菲利普是一名引座员，后来晋升为售票员，最后被提拔为波士顿市中心一家电影院的总经理。1922年，菲利普向亲朋好友借了钱，在波士顿北区开了一家电影院。5年后，由艾尔·乔尔森（Al Jolson）主演的有声电影打破了默片的声音屏障。

在接下来的40年里，菲利普成功地建立了连锁影院，从美国的新英格兰地区发端，然后扩展到中西部。菲利普是汽车影院的开拓者，以精明的经营者而著称。在去世前一年，也就是1961年，菲利普曾将公司上市，以筹集资金建造更多的汽车影院。62岁的菲利普突然离世后，儿子迪克·史密斯立即接任大众汽车影院的CEO一职，当时汽车影院也被称作免下车影院。那一年，迪克37岁。

在之后的43年里，从这家不寻常的连锁影院起步，这位新生代CEO的持续表现相当抢眼，超过标准普尔500指数11倍，直接完胜市场大盘和杰克·韦尔奇。迪克取得这些成果的背景是：这是一家由创始人家族控制的上市公司。迪克却把它当作私人控股公司来经营，并对分配公司的现金流表现出难得一见的耐心。成熟的汽车影院业务带来的现金流被迪克进行了多元配置，首先是投资购物中心影院，然后再投资完全不同的业务线。

迪克会长期蛰伏，偶尔发现机会则倾巢出动。在其任期内，迪克在不相关的业务领域共进行了3次重大收购，分别是：20世纪60年代末收购经营瓶装软饮料的美国饮料公司（American Beverage Company），20世纪80年代中期收购经营零售业务的卡特-霍利-黑尔公司（Carter Hawley Hale, CHH），90年代初收购哈考特-布雷斯-朱万诺维奇出版公司（Harcourt Brace Jovanovich, HBJ）。这一系列交易让这家地区性汽车电影公司转型为非常成功的经营消费品的企业集团。

商业世界到处布满了企业的“遗骸”，它们跨行收购但惨遭失败。这类多元化收购执行起来是出了名的困难，时代华纳和美国在线

（American Online）莫不如是。然而，迪克依靠家族关系上位，相对缺乏经验，却成了这个方面的大师。事实上，迪克在大众影院公司的任期可以看作连续不断的革新过程，也可以被视为关于及时撤离的故事集合：一次是在20世纪80年代末，然后2003年和2006年各有一次撤离。多元化收购和资产剥离导致企业像手风琴一样时而扩张时而收缩，这是极不寻常的，虽然在某些方面和亨利·辛格尔顿的特利丹公司有点类似。同时，多元化收购和资产剥离也为大众影院公司的股东带来了巨大的利益。

轻资产战略，“售后租回”的资本游戏

迪克1924年出生于马萨诸塞州牛顿市一个温暖有爱的大家庭里，他是四个孩子中的老大，从小就在周末和假期时帮助打理家族生意。迪克在剑桥大学读完预科班，1946年从哈佛大学毕业，并获得工程学学位。第二次世界大战期间，迪克是一名海军工程师，战后没有去读MBA而是回到家族企业工作。1956年，迪克32岁，他的父亲让他成为正式合伙人。

父亲去世后，迪克积极地将公司院线扩展到郊区的购物中心，在那里它是毫无疑问的先驱。迪克是业内第一个认识到，郊区影院的业务将受益于强大的、潜在的人口增长趋势。为了抓住这个机会，迪克开创了两种革命性的做法。

首先是新剧院的融资。在影院开发中，传统方法强调取得标的土地的所有权，这样可以保证对资产的长期控制，以及通过抵押土地来获得贷款。然而，迪克意识到，一个位置优越的影院可以迅速产生可预测的现金流。迪克率先使用融资租赁来建造新影院，大大减少了前期投资。这一创新使迪克以最小的资本投入，实现了其院线的迅速发展。

其次是在每家影院增加更多银幕，以吸引更多观众，并优化高收益的影院特许商品销售。双重革新带来的结果就是，在整个20世纪60年代到70年代初，大众影院公司通过投资新影院获得了丰厚回报。然而，到了60年代末，迪克觉察到影院的增长不可能无限期地持续下去，于是开始探索多元化经营，进入那些长期前景更看好的新业务领域。

迪克的转型发生在1968年，他收购了总部位于俄亥俄州的美国饮料公司，这是全美最大的百事可乐独立灌装公司。迪克通过影院特许商品经营而熟悉了饮料行业，当得知美国饮料公司可能要出售时，他便迅速

采取行动。迪克谈成的这笔生意因为收购价格相当于公司现金流5倍数额而引人注目，其业务量相当于公司当时企业市值的20%以上的规模。迪克把自己对不动产的理解发挥到极致，通过对美国饮料公司生产设施及场地的出售或返租，创造性地为收购融得了资金。今天，迪克仍然有理由为这一战略而感到自豪。

无形资产，小运营资金带来超高现金流

电影院是由一砖一瓦搭建的现实世界，迪克在这里长大，美国饮料公司使他第一次接触到诸如饮料品牌等企业无形资产的价值。迪克渐渐喜欢上了饮料行业，这个行业属于卖方市场，有着非常高的资本回报率和诱人的长期增长趋势。迪克特别喜欢百事可乐灌装公司内部的动态机制，这些公司的所有权很分散，许多公司二代和三代的所有者都是潜在的经销商。这不同于可口可乐的体系，可口可乐的灌装由屈指可数的几家大型独立公司把持。由于百事可乐是饮料行业排名第二的品牌，其特许经营权的交易价格往往低于可口可乐。

通过收购美国饮料公司，迪克收购了一家合法的平台公司——其他公司可以轻松高效地并入其中。随着美国饮料公司发挥出规模优势，迪克意识到他能以高价购买新的特许经营权。虽然这个价格可能相当于卖家现金流的很多倍，但是通过削减支出、合理避税以及专业营销，实际的购买价格就会大幅降低。在这一观念的基础上，迪克积极行动，去收购其他特许经营权，包括1973年买下百事可乐美国公司、1977年买下百事可乐瓶装公司以及1977年买下华盛顿特区的特许经营权。

与经营影院业务一样，迪克和他的团队是创新型的营销人员和高效的运营商。公司致力于不断降低成本，例如，利用其规模优势降低易拉罐价格；为了避免母公司加价，直接在国际市场采购食糖。这种致力于降低成本的经营之道所带来的结果就是，美国饮料公司获得了行业领先的利润率。除了百事可乐，迪克偶尔也会买入其他饮料的特许经营权，比如七喜和胡椒博士（Dr Pepper）。1976年，美国饮料公司与最大的橘子种植商合作，推出了新奇士（Sunkist）橘子苏打汽水，并通过其分销网络进行推广。为了推出新奇士橘子苏打汽水，美国饮料公司最终投资了2 000万美元，1984年却以8 700万美元的价格将其出售给了加拿大的公司Canada Dry，非凡的投资回报由此而来。

第三支点，“非收购入股”缔造完美公司构架

在饮料业务走上正轨后，迪克开始寻找下一个目标，一项可以作为大众影院公司“第三条腿”的新业务。沿着这一思路，在20世纪70年代末和80年代初，大众影院公司在广播媒体业务方面进行了许多规模较小的收购，买下了数家电视台和广播电台。然而，迪克的出价原则不允许大众影院公司在市盈率达到两位数的情况下去收购，但这是当时广播业普遍的市盈率倍数。尽管大众影院公司在媒体业的小型投资组合获得了良好回报，但是从未成为该行业的主角。“这是一次错失的机会。”普特南投资公司的老牌投资人鲍勃·贝克（Bob Beck）这样说。¹

随着时间的推移，迪克的收购方式也在与时俱进。从20世纪80年代初开始，迪克和他的收购团队专注于等待，偶尔在时机成熟时下重金，同时还对那些他们认为被低估的上市公司进行非控股型投资。这些投资旨在帮助大众影院公司进行多元化的尝试，迪克把这一战略命名为“参与式投资”，就是进行大量的非控股型投资，以此在其他公司的董事会中获得一席之地，然后与其管理层合作去改善公司运营、增加公司价值。

20世纪80年代前半期，迪克亲自介入3项“参与式投资”的尝试：哥伦比亚影业、休伯莱恩公司（Heublein）和吉百利公司。在后两个案例中，对方公司在职的管理团队对大众影院公司的投资持怀疑态度，甚至是公开的敌视，结果迪克没有得到董事会席位，在最初交易后的一到两年内，迪克把它们都剥离了。这些投资的回报很不错，但没有实现多元化这一深层次的目标。1985年4月，大众影院公司首席财务官伍迪·艾夫斯（Woody Ives）接起电话，发现另一头是摩根士丹利公司的投资银行家埃里克·格里切尔（Eric Gleacher），“第三条腿”终于出现了，情势因此骤变。

这天是周三，格里切尔的电话是关于CHH的事，这是一家上市零售企业集团，拥有多家百货公司和专业连锁零售店。利明特公司（The Limited）的CEO莱斯利·韦克斯纳（Leslie Wexner）最近对其提出了恶意收购要约。而格里切尔受雇CHH为其寻找一位“白衣骑士”，友好的投资者应该有能力买下CHH绝对大比例的股票，以此来挫败恶意收购。

起初，艾夫斯对格里切尔的描述反应冷淡，但他在放下电话的瞬间，便意识到这是一个潜在的重大机会。时间紧得不能再紧了，他们需要在下周二之前做出回应。艾夫斯意识到，任何能够按照这个如此紧迫的时间表行事的买家，都必须在交易谈判中借助巨大的杠杆。放下电话，艾夫斯就向迪克以及管理团队的其他高层成员做了汇报，当天下午

5点，他们登上了飞往CHH公司洛杉矶总部的飞机。

整个周末都在紧张的尽职调查和谈判中度过，周日晚间就达成了协议。周一是美国的爱国者节，波士顿的银行放假，迪克他们紧急召集了一个由3家银行组成的财团为交易提供资金支持。到了周四，也就是格里切尔第一次致电后的第8天，交易完成。迪克和他的管理团队对他们的收购标准和流程做了极致的精简，所以才能够跟得上这个异常紧急的时间表。几乎没有上市公司能在这种大规模的交易中行动如此迅速。

对CHH的投资是迪克奉行机会主义原则的极好例证，只要情势所需，迪克就愿意下重注。这次交易规模相当于大众影院公司市值的40%以上，交易规模可见一斑，而且过程非常复杂，但同时利益也相当诱人。艾夫斯以可转换债券的形式谈下了大笔资金，保证债权人能获得大众影院公司10%的收益；如果业绩表现良好，大众影院公司还允许债权人将其收益转换为CHH公司普通股的40%，并且能以固定价格购买CHH全资子公司华登堡书业（Waldenbooks）的期权。艾夫斯向我总结道：“那天结束的时候，我们借到了钱，其中6%~7%的部分全额免税，同时获得10%的税前利息抵扣额度。此外，我们还有一项转换期权，该转换期权允许我们最终剥离内曼·马库斯集团（Neiman Marcus）和对华登堡书业的认购期权。”²这个周末的加班成果着实不错。

最终，大众影院公司将其在CHH 40%的股份换为CHH专业零售部门60%的控股权，该部门的主要资产是内曼·马库斯连锁商店。大众影院公司对CHH投资的长期回报率高达51.2%。通过CHH的交易，大众影院公司决定性地转向零售业，这项新业务诱人的增长前景与饮料或影院业务都没有什么关联。

20世纪80年代末，迪克发现了两个令人不安的趋势：一个是重整旗鼓的可口可乐开始在当地市场蚕食百事可乐；另一个是因为饮料行业的经济效益变得广为人知，其特许经营的价格大幅飙升。作为回应，虽然很不情愿，但迪克决定试着出售饮料业务，最终在1989年以创纪录的高价将美国饮料公司出售给百事可乐的母公司。这次出售之后，大众影院公司的资产负债表上留下了超过10亿美元的现金，迪克开始再次寻求能实现多元化的收购机会。

没过多久，迪克就找到了。1991年，经过18个月的曲折过程，迪克完成了他最大的也是最后一笔收购，在一次情况复杂的拍卖中收购了出版商HBJ，由此也为大众影院公司安装了最后的“第三条腿”。HBJ是一

家领先的教育和科学类读物出版商，还拥有一家考试机构和一家职业介绍公司。自60年代中期以来，这家公司一直由CEO威廉·朱万诺维奇（William Jovanovich）个人独揽经营。1986年，HBJ收到了“变节”的英国出版商罗伯特·麦克斯韦尔（Robert Maxwell）的敌意收购要约。作为回击，朱万诺维奇大举借债，出售了HBJ的游乐场业务，还向股东派发了大量的股票。

这一系列举措让麦克斯韦尔望而却步，但也让HBJ背负上无法承受的债务。由于HBJ债务违约并无法付款，其债券都在打折出售，公司的结构化债券犹如多层蛋糕一样复杂，一些专门针对危机企业的秃鹫型投资者，比如阿波罗投资公司（Apollo Investments）的莱昂·布莱克（Leon Black）开始在其中积累头寸。

由于业务停滞不前，朱万诺维奇于是退休，由长期担任HBJ高管的儿子彼得继任。1990年底，该公司聘请史密斯·巴尼（Smith Barney）执行拍卖程序。在大众影院公司内部，执行团队对HBJ复杂的资本结构进行了深入分析，尽管他们厌恶拍卖，但是管理团队的结论是：HBJ的业务非常符合大众影院公司的收购标准。于是他们决定积极推进。

管理团队还认为，HBJ迷宫般的被投资银行家西泽·斯韦策（Caesar Sweitzer）描述为“企业金融预修课程”的资产负债表，可能会吓跑其他买家，从而为思想灵活、资本充足的独立买家提供机会，并可能以惊人的估价来达成协议。³与HBJ众多债权人进行广泛磋商后，迪克同意以15.6亿美元收购该公司，这相当于大众影院公司当时企业市值的62%，绝对是“押上了身家性命”。这一价格相当于HBJ核心出版资产现金流的6倍，与可比交易相比，这个价格相当诱人。迪克后来以11倍现金流的价格出售这些业务。表7-1列举了交易资金的来源和用途，从中可以看出交易的复杂性和众多的参与方。

表7-1 大众影院公司收购HBJ的金融数据

HBJ 的资产类别	初始报价	最终报价
	14.6 亿美元	15.6 亿美元
普通股	每股 1.30 美元现金	每股 0.75 美元（以大众影院公司的股票支付）
优先股	每股 1.30 美元现金	每股 0.75 美元（以大众影院公司的股票支付）
主要票据	本金的 93%	本金的 100%
主要次级债	本金的 77%	本金的 91%
次级债	本金的 45%	本金的 47.5%
无息债券	本金的 32.4%	本金的 40.97%
实物付息债券	本金的 40%	本金的 47%

资料来源：General Cinema / Harcourt Brace Jovanovich Joint proxy statement, pp.32, 40, 46-47.

1991年收购HBJ后，大众影院公司将其成熟的影院业务剥离出来，成立独立的上市实体影院集团（GC Companies），这样管理层就能够把注意力集中在规模更大的零售和出版业务上。在接下来的10年里，迪克和他的管理团队继续经营着零售和出版业务。2003年，迪克将HBJ的出版资产出售给了里德·爱思唯尔集团（Reed Elsevier）；2006年，迪克又将内曼·马库斯集团大众影院投资组合里的最后一部分出售给了一个私募股权基金组成的买家财团。这两项交易的估值都创下了各自行业的纪录，为迪克及其股东的超常表现画上了完美句号。

由于父亲去世，迪克出人意料地被推上了CEO的宝座，他在执掌大众影院公司的43年里创造了惊人的业绩（见图7-1），为其股东创造了16.1%的年复合回报率，把同期标准普尔500指数（9.0%）和通用电气（9.8%）的回报率甩在身后。1962年初，投资1美元给迪克，43年后其价值达684美元；同样的1美元投资于标准普尔500指数，43年后其价值达43美元；如果投资于通用电气，则其价值达60美元。



注：内曼·马库斯集团于1989年10月分拆给大众影院公司的股东，哈考特通用公司（Harcourt General）于2001年7月出售给里德·爱思唯尔集团。为了获得2005年10月出售内曼·马库斯集团给德州太平洋集团和华平基金（TPG and Warburg Pincus）的回报数据，假设出售哈考特通用公司的收益再投资于证券，该证券每年的增长率取莱卡迪亚公司（Leucadia）和标准普尔500指数当年的平均增长率（2001年7月至2005年10月期间，莱卡迪亚公司和标准普尔500指数的复合年增长率分别为17.7%和0.7%）。

图7-1 大众影院公司股东总收益远超标准普尔500指数和同行

资料来源：证券价格研究中心。

The Outsiders

迪克·史密斯的方法

迪克发展出一种独特的方法来管理大众影院公司的运营，他通过与3位高管——CFO艾夫斯、COO鲍勃·塔尔（Bob Tarr）和公司顾问萨姆·弗兰肯海姆（Sam Frankenheim），组成的小圈子密切合作来实现对公司的管理。迪克正式命名这个组织为“董事长办公室”，或简称“董办”（OOC）。董办每周召开一次会议，会上迪克积极鼓励高管之间进行辩论。大众影院公司资深的投资银行家斯韦策将这些会议描述为“以建设性的、合议庭的方式进行的摔跤比赛”。⁴

迪克甚至希望自己的提案被董办的其他成员否定。后来离开大众影院公司并领导东方资源公司（Eastern Resources）成功转型的艾夫斯是公司才华出众的CFO，他还记得自己在大众影院公司最自豪的时刻之一：在迪克准备与康卡斯特公司和哥伦比亚广播公司组建合资公司开展有线电视业务而发表言论时，艾夫斯发表了反对意见，迪克的提案因此被董办否决。“他允许我在董事会面前公开反对他。没有几个CEO会这么做。”⁵

大众影院公司的总部只有非常精简的运营人员，公司总部位于马萨诸塞州栗树山市（Chestnut Hill）一家不起眼的购物中心后方，紧挨着自家影院。迪克自豪地指出，这个位于郊区的办公空间实际上是免费的，因为影院支付了整个建筑群的租金。“迪克将日常运营管理委托给了董办和各部门负责人，他几乎没有时间和投资者进行沟通，偶尔的沟通也‘只是意思一下’。”普特南投资公司分析师鲍勃·贝克这样描述道。⁶相反，迪克将大部分时间花在了公司战略和资本配置问题上。

这一小群管理者对公司的管理十分严谨。资深媒体投资人戴维·沃戈给我分享过一份报告，当时大众影院公司刚刚完成对HBJ的收购，这份报告是他与管理层会晤后撰写的。报告中引人注目的是其简明扼要的陈述：清晰地列举了交易理由，然后是附带的具体基准和回报目标清单，没有多余措辞。事后来看来，同样值得注意的是，大众影院公司实现或超过了当初制定的每一个目标。正如沃戈在谈到大众影院公司的团队时所说：“由一小群真正的精英创造出的价

值是惊人的。”⁷

迪克在30岁的尾声时就任CEO，他很乐意让高管们在职业生涯的早期就承担起管理责任。1974年，迪克聘请没有任何企业运营经验的37岁的投资银行家艾夫斯担任公司CFO；同样，1978年，迪克任命34岁的退役潜艇指挥官、哈佛大学MBA塔尔为饮料部门总裁；后来迪克又雇用了35岁的保罗·德尔·罗西（Paul Del Rossi）来经营公司的影院业务。

用迪克的话来说，这些高管的薪酬“有竞争力，但不是特别高”。⁸然而，大众影院公司确实通过期权和慷慨的员工股票购买计划向核心管理人员配置了股权，其比例达到法律对员工持股要求的上限。艾夫斯认为，这些创举的直接效果就是，高管团队“感觉自己像是老板……我们都是股东，工作自然卖力”。⁹

把现金流视为第一要务

迪克资本配置的历史表现堪称优异。在迪克的漫长任期内，3个主要的现金来源是：经营现金流、长期负债和偶尔出售大型资产的所得。

电影院业务在现金流方面有其独特之处：一方面，其运营资本为负数，下游客户需提前支付费用，同时上游电影制片厂的电影版权费可以拖欠90天；另一方面，电影院业务拥有低资本要求的特点，一旦影院建成，维护费用的投资可谓九牛一毛。这些诱人的经济因素对迪克的商业世界观产生了巨大影响，从公司历史的早期开始，他就专注于现金流最大化，而不是传统的每股收益最大化。

我在迪克的办公室见到他时，他给我看了1962年公司的年报，那是他第一次担任CEO，报告中他反复提到现金收益即净收益加折旧是评估公司业绩的关键指标，而不是净利润。这很可能是美国商业用语中第一次使用现金收益这个词，但其现在已成为标准术语。正如公司CFO艾夫斯所说：“我们的重点始终是现金。”在迪克的任期内，公司一直创造着高水平的经营现金流。¹⁰

迪克不愿意发行股票。事实上，从大众影院公司首次公开募股开始，迪克差不多完全避免了增发股票，直到1991年才增发了少量

股票，也是为了帮助公司在HBJ的交易中获得优惠的税收待遇。迪克对我说：“我们从未增发过一股。我就像一个封建领主，坚守着祖先的土地！”

另外，为了给收购提供资金，大众影院公司的战略是利用债务。它最大的两次收购，即收购CHH公司和HBJ公司，其资金完全来源于债务融资。因此，从20世纪80年代中期开始，公司债务与现金流的比率一直保持在至少3倍，以此来放大股票回报并将税负降至最低。

税负最小化是大众影院公司另一个重要的资金来源，也是该公司与众不同的另一个领域。尤其是迪克，他是这方面的先行者。该公司长期的税务顾问迪克·丹宁（Dick Denning）告诉我：“他们非常老练.....在探索和利用新的税收理念方面绝不保守。”在迪克任期内，大众影院公司的平均税率是33%，而同期的企业税率接近50%，我们从中便可以看出这种税收筹划的有效性。

通常而言，在没有外部股东压力的情况下，一家公司的CEO很少会出售公司的大型部门或大型业务。然而，迪克是一位杰出的商人。与本书中的几位CEO一样，包括比尔·安德斯和比尔·斯蒂利茨，迪克曾经分别3次以创纪录的高价出售了公司的大型部门和业务：1989年的饮料业务、2003年HBJ的出版业务和2006年的内曼·马库斯集团。每一次，当看到业务增长、前景黯淡和市场高估值并存时，迪克便大举抛售，即使这意味着公司规模的大幅缩水。

在2006年的业务出售案例中，迪克看到，内曼·马库斯连锁商店的进一步扩张需要密集的资金投入（每开设一家新店需要投入5000万美元）以及他们面临运营方面的挑战。迪克还发现，拥有大量低成本债务的私募股权公司，愿意为一流的零售业资产支付创纪录的价格。于是迪克聘请了投行高盛集团策划了一次全面竞拍，最后由得州太平洋集团牵头的财团买下内曼·马库斯集团，支付价格对比现金流的倍数创下了历史纪录。

影院集团负责公司历史悠久的影院业务，迪克没有出售这个成熟的业务部门。据称，迪克曾在20世纪80年代末动过出售该集团的心思。到90年代末，电影放映行业的竞争变得更加激烈。大众影院公司没有出售影院集团，而是试图优化已有院线，以期与全国各地涌现的新影院展开竞争。大众影院公司关闭了一些同时扩建了另一

些影院，投资新的放映技术，但一切于事无补。到了90年代末，影院集团无力偿还债务，遂宣布破产，这对迪克来说是罕见的挫败。

花好每一分钱

迪克将这些来源不同的现金配置到3个主渠道：收购公司、股票回购和资本支出。公司支付的股息微乎其微，值得注意的是，它保有大量的现金余额，同时静待有吸引力的投资机会浮出水面。

迪克收购的公司有几个共同特点：它们都是市场领导者，具有稳定的增长前景和受人尊敬的品牌。迪克通常会充分利用时机来完成此类交易，并在其他潜在的买家迟疑时积极出手。在CHH的收购案例中，没有其他买家能够像迪克这样迅速地反击利明特公司的收购要约；在收购HBJ的案例中，没有其他买家愿意花时间来厘清复杂的资本结构，并与多层次的债权人进行谈判。相对于公司规模来说，这些收购也押上了身家性命，其收购价格占公司企业市值的22%到62%不等。

随着时间的推移，迪克一直在稳步回购大众影院公司的股票，最终回购了该公司1/3的股份，这些回购的长期内部收益率达到诱人的16%。1982年，作为对迪克“参与式投资”的回应，休伯莱恩公司买下了大众影院公司的大量股份。迪克随机应变，回购了10%的自家股份，这也是他最大的单笔回购。

大众影院公司对资本支出保持严谨态度，所有投出去的资本都要求有较高的现金回报。公司早期的郊区影院创造了非凡回报，饮料部门也有着众多诱人的内部投资机会，公司的其他业务则遵循已有标杆业务的高标准。HBJ出版业务的实物资产很少，因此其资本需求很低，而内曼·马库斯集团的连锁零售业务却需要大量资金。

尽管零售业比大众影院公司的其他业务更具资本密集性，但迪克认为内曼·马库斯集团拥有独一无二的品牌，只是前任所有者经营不善。有些时候，迪克愿意进行大笔投资开设新的内曼·马库斯连锁商店，因为他相信，只要显示出增长潜力，将来公司就能实现溢价出售。在拥有内曼·马库斯集团的20年里，大众影院公司只开了12家新店，而内曼·马库斯集团的开店计划则是这个数字的许多倍。内曼·马库斯集团最后的转让天价，充分证明了迪克的经营逻

辑。

与采访大都会通信公司一样，对大众影院公司前任高管的访谈中弥漫着一种富有感染力的热情，一种同事间的友情和冒险精神。这群管理者齐心协力，带领公司进入了一系列兼收并蓄的新行业。在每一个案例中，他们都证明自己是卓越的运营者，拥有行业领先的利润率和卓越的回报。迪克·史密斯成功地创造了一个环境，让这群才华横溢的管理者获得超乎寻常的自主权，并让他们感觉自己就是企业的主人。“我们所有人都玩得如此开心。”说到这里时，迪克的眼中闪烁着光芒。伍迪·艾夫斯最后谈到他在公司的个人股权时说：“我只希望自己永远不会卖出一股。”¹¹

特质**8** 首席执行官就是首席投资官
沃伦·巴菲特和伯克希尔-哈撒韦公司

The Outsiders

- 他是举世公认的投资之神，并以数百亿美元财富总额高居全球富豪榜前三位。
- 伯克希尔-哈撒韦公司全资子公司的利润，已经远超其在二级市场投资的利润。他的管理者身份与投资者身份越发难以区分！
- 伯克希尔-哈撒韦公司拥有27万名员工，其总部却只有区区23人。



他是

沃伦·巴菲特

WARREN BUFFETT

温斯顿·丘吉尔	你塑造环境，环境反过来也在塑造你。
爱因斯坦	复利是宇宙间最强大的力量。
沃伦·巴菲特	做一个 CEO 让我成为更好的投资者，反之亦然。

伯克希尔-哈撒韦公司的前身是一家拥有上百年历史的纺织公司，位于马萨诸塞州的新贝德福德地区（New Bedford）。这家公司一直由蔡斯和斯坦顿这两个当地家族所拥有，时间长达几代人。1965年，这家新英格兰地区的企业在度过了黄金时代之后渐渐没落，此时的它原本不会成为恶意收购的目标。虽然公司业务日薄西山，但即便是被收购，至少也不应沦为恶意收购的对象。但年届七旬的CEO西伯里·斯坦顿（Seabury Stanton）性格固执，由于他拒绝与一位心怀不满的职业投资人会面，硬生生导致后者成为令他意想不到的可怕对手。

其后的股权之争持续了一段时间，最终，伯克希尔-哈撒韦纺织公司居然被来自内布拉斯加州的青年才俊沃伦·巴菲特收入囊中，而巴菲特最不像一个掠食者。这一年，巴菲特35岁，名不见经传，长着一双招风耳。他在奥马哈一座不起眼的办公楼里经营着一家小型投资合伙公司，此前也没有公司管理经验。

然而，巴菲特与20世纪80年代那些声名狼藉的杠杆收购狂徒迥然不同。首先，巴菲特没有十足的敌意，收购之前他就和蔡斯家族建立了紧密的联系。其次，巴菲特没有利用任何负债，而利用负债是戈登·盖柯（Gordon Gekko）和亨利·克拉维斯（Henry Kravis）常用的伎俩。

巴菲特被伯克希尔-哈撒韦纺织公司吸引，是因为相对于公司账面价值，其股票价格低廉。当时，该公司在激烈的商品市场（西装衬里）竞争中处于劣势，市值仅为1 800万美元。巴菲特从这个平淡无奇的开端起步，创造了史无前例的业绩。以股票的长期表现来看，相对于其他CEO，这个内布拉斯加州曾经的平头小子简直是天外来客。在巴菲特手中，那家昔日破败的新英格兰纺织公司取得了惊人的回报，2011年时这家公司的市值已达1 400亿美元，而股本几乎没变。巴菲特第一次买入伯克希尔-哈撒韦纺织公司股票时，其股价是7美元，如今股价已经超过12万美元。

巴菲特是如何从这样一个难以置信的平凡起点实现了上述非凡的转变呢？巴菲特的投资人背景又怎样帮助他形成了管理伯克希尔-哈撒韦公司的独门秘籍？精彩故事自此开始。

垃圾股淘金，捡小便宜的烟头策略

巴菲特1930年出生在内布拉斯加州的奥马哈，这也是他家族深深扎根的地方。祖父在当地经营着一家人尽皆知的杂货铺；父亲是奥马哈市中心的股票经纪人，后来成为国会议员。巴菲特继承了他们的淳朴性格，他很早就表现出创业倾向并进行了一系列的商业尝试，从6岁开始到高中毕业，他干过的工作包括送报纸、出租弹子游戏机和售卖软饮料。大学阶段，巴菲特曾在沃顿商学院短暂求学，后来，20岁时从内布拉斯加大学毕业并开始申请哈佛商学院。

巴菲特一直对股票市场很感兴趣。像保罗在大马士革接收到神的启示一样，19岁的巴菲特读到本杰明·格雷厄姆写的《聪明的投资者》

（*The Intelligent Investor*）一书继而顿悟，一夜之间皈依为价值投资的信徒。巴菲特遵循格雷厄姆的公式：买入那些从统计数据来看很便宜、公司市值远低于净运营资产或净资产的企业。明白道理后，巴菲特知行合一，把早期商业活动的收益都投入了股票市场，当时大概1万美元。申请哈佛商学院失败后，巴菲特去了哥伦比亚大学跟随格雷厄姆学习，成为班上耀眼的明星，并且得到了格雷厄姆在哥伦比亚大学从教20多年里给出的第一个A+。

1952年从哥伦比亚大学毕业之后，巴菲特想在老师的投资公司谋个职位，但遭到拒绝，他于是返回奥马哈找了一份经纪人的工作。工作期间，巴菲特向客户推荐的第一家公司就是盖可保险（GEICO），这家汽车保险公司直接向政府雇员销售保单。盖可保险最初吸引了巴菲特的注意，是因为老师格雷厄姆担任其董事长。随着研究的深入，巴菲特发现盖可保险既有重要的竞争优势，又符合安全边际的要求。按照格雷厄姆的定义，安全边际就是一个拥有充分知情权的精明买家愿意支付的价格远低于公司内在价值。巴菲特把自己的大部分净资产投资于盖可保险，并试图说服他的客户也买入这家公司的股票，然而并没有人买账。此时巴菲特意识到，总体来说，经纪人的工作和他钟爱的投资事业相去甚远。

回到奥马哈之后，巴菲特一直和格雷厄姆保持着联系，把自己对于

股票的理解不断地发送给老师。到了1954年，格雷厄姆终于被打动，并在盖可保险为巴菲特谋得一个职位，巴菲特得以回到纽约，并在接下来的两年里一直为老师工作。后来，巴菲特用“雪茄烟蒂”这个生动的比喻来形容这个时期的研究对象，即那些价格低廉、质量低下的公司。1956年，格雷厄姆解散了公司，开始追逐生活中的“诗和远方”，比如翻译古希腊作家埃斯库罗斯的作品。巴菲特也回到奥马哈，从朋友和家人那里募集到10.5万美元，开设了他的小型投资合伙企业。此时，巴菲特的净资产达到14万美元，相当于2011年的100万美元。

现金流折现，只买最值的企业

在接下来的13年里，巴菲特合伙公司（Buffett Partnership）取得了非凡的业绩，在不使用杠杆的情况下每年都大幅击败标准普尔指数（见表8-1）。这些成绩，总体上是通过格雷厄姆深度价值方法的运用而取得的，但也有例外，巴菲特在20世纪60年代中期进行的两次大额投资，即投资美国运通公司（American Express）和迪士尼公司都没有再遵循格氏理念。这预示着巴菲特的投资哲学将发生翻天覆地的变化，从“烟蒂”转向那些具有强大的竞争壁垒的高质量公司。

表8-1 巴菲特合伙公司投资结果

年份	巴菲特合伙公司 (%)	道琼斯指数 (%)	差距 (%)
1957	10.4	-8.4	18.8
1958	40.9	38.5	2.4
1959	25.9	20.6	5.3
1960	22.8	-6.2	29.0
1961	45.9	22.4	23.5
1962	13.9	-7.6	21.5
1963	38.7	20.6	18.1
1964	27.8	18.7	9.1
1965	47.2	14.2	33.0
1966	20.4	-15.6	36.0
1967	35.9	19.0	14.9
1968	58.8	7.7	51.1
1969	6.8	-11.6	18.4
平均值	30.4	8.6	21.8

1965年，巴菲特通过巴菲特合伙公司收购了伯克希尔-哈撒韦公司的控股权。在接下来的4年多时间里，巴菲特带领合伙企业继续取得优异的业绩。1969年，面对20世纪60年代末期牛市中高涨的股价，巴菲特突然解散了合伙企业。这并非巧合，同年亨利·辛格尔顿的特利丹公司也停止了收购。不过，巴菲特坚持保留手中的伯克希尔-哈撒韦公司的股权，把它当作开辟未来投资事业的潜在工具。

取得伯克希尔-哈撒韦公司的控股权后，巴菲特立即任命了肯·蔡斯（Ken Chace）为新的CEO。上任头3年，通过降低库存、出售冗余的厂房和设备，另外叠加纺织行业罕见的周期性利润爆发，蔡斯不辱使命，为伯克希尔-哈撒韦公司创造了1 400万美元的现金。这笔资金中的绝大部分被用于收购国民赔偿保险公司（National Indemnity Insurance）。这家公司在保险损失和赔付产生之前就能先行收取保费，因此以浮存金的形式获得了大量现金。巴菲特非常高效地使用这些浮存金进行投资，同

时购买公开交易的证券和非上市公司，包括《奥马哈太阳报》（*Omaha Sun*）这样一家周报以及伊利诺伊州罗克福德（Rockford）的一家银行。

与此同时，在伯克希尔-哈撒韦公司的事务之外，巴菲特开始与查理·芒格密切合作。查理·芒格也是土生土长的奥马哈人，是一位杰出的律师和投资家，当时的主要业务集中在西海岸。二人惺惺相惜，相见恨晚。当20世纪80年代初查理·芒格出任伯克希尔-哈撒韦公司的副董事长时，二人开始了正式的合伙，一直到今天。

整个70年代和80年代早期的伯克希尔-哈撒韦公司年报，持续弥漫着对通货膨胀的担忧。当时人们普遍认为，只有黄金、木材等硬通货才是对抗通货膨胀的有效工具。但是，在查理·芒格的影响下，巴菲特从格雷厄姆的传统方法中蜕变并提出不同见解。巴菲特独辟蹊径地认为：实际上那些资本需求低同时有提价能力的公司，才是抵御通胀侵蚀的利器。

基于以上理解，巴菲特投资了很多拥有“特许经营权”的消费品牌和媒体产业，这些生意或处于市场支配地位，或品牌名称广为人知。伴随着上述投资标准的变化，另一个重要变化是，巴菲特的持股周期也变得越来越长，并由此实现了税前利润的长期复利回报。

巴菲特的这种投资标准和理念的变化，其重要性再怎么强调都不为过。此前，巴菲特注重资产负债表和有形资产，此种投资方法已经得到市场验证并为他带来丰厚利润。但是巴菲特在职业生涯中期就完成了跃迁，找到了另一种完全不同的方法，那就是：着眼未来，强调利润，关注品牌和市场份额等难以量化的资产的价值。为了确定安全边际，巴菲特摒弃了格雷厄姆钟爱的净运营资本计算方法，转而依靠现金流折现法和非公开市场的企业估值。此举引发的争议和同时代鲍勃·迪伦（Bob Dylan）的转变毫无二致，后者完成了从木吉他到电吉他的跃迁。

这种结构性转型也表现在伯克希尔-哈撒韦公司的保险投资组合中，整个20世纪70年代，投资组合中媒体和品牌消费品公司的占比不断上升。70年代结束时转型得以完成，那时，巴菲特投资组合已经拥有喜诗糖果（See's Candies）和《布法罗新闻》（*Buffalo News*）的全部股权，以及《华盛顿邮报》、盖可保险和通用食品的大量股票头寸。

1981—1985年，巴菲特专注于在投资组合中增加非上市公司，比

如，1983年以6 000万美元收购内布拉斯加州的家具百货公司，1985年以3.15亿美元收购了一家企业集团斯科特·费策尔公司（Scott Fetzer）。1986年，巴菲特进行了截至当时的最大一笔投资，投入5亿美元帮助朋友——大都会通信公司的CEO汤姆·墨菲收购了美国广播公司。巴菲特和伯克希尔-哈撒韦公司最终持有合并后的大都会通信公司18%的股份，大都会通信公司成为巴菲特所谓的“永久”持有的第三家公司，另外两家是盖可保险和《华盛顿邮报》。

·转折点：喜诗糖果·

1972年，巴菲特和查理·芒格出价2 500万美元收购了喜诗糖果，这是巴菲特将投资重点从“烟蒂”转向“特许权”的关键。当时，喜诗糖果的有形账面净资产为700万美元，税前利润为420万美元。看上去收购价居然是账面净资产的3倍多，不过只有税前利润的6倍。以格雷厄姆的标准来看，收购代价如此之高的公司，他永远不会碰。然而，巴菲特和查理·芒格却意识到：喜诗糖果作为一个深受消费者喜爱的品牌，拥有优异的资本回报率和尚未启动的定价权。为了发掘这个机会的价值，他们立即任命查克·哈金斯（Chuck Huggins）为新任CEO。

在喜诗糖果被收购之后，其单位产量并没有显著增长，但由于其品牌的力量，产品可以持续提价。最终，伯克希尔-哈撒韦公司的此笔投资在前27年里获得了惊人的年化32%的复利回报。1999年后，喜诗糖果的业绩不再单独公布。

在过去的39年中，最初投资于喜诗糖果的2 500万美元为伯克希尔-哈撒韦公司提供了16.5亿美元的自由现金流。这些现金流被技艺高超的巴菲特成功赋能，进行再投资。喜诗糖果注定是伯克希尔-哈撒韦公司成功的关键基石。有趣的是，当初的收购价对高回报率贡献无多。即使巴菲特和查理·芒格当时付出两倍的价格，回报率仍会达到非常诱人的水平，即年化21%。

1987年，在10月美股崩盘之前，巴菲特已经卖出了其保险投资组合中的所有股票，仅保留3家核心公司。完成大都会通信公司的交易后，巴菲特没有在公开市场再进行过投资。直到1989年，巴菲特宣布伯克希尔-哈撒韦公司历史上的投资规模再创新高：用相当于伯克希尔-哈撒韦公司账面资产1/4的资金买下可口可乐公司7%的股份。

20世纪80年代末期，巴菲特配置了少量的上市公司可转换债券，包括所罗门兄弟（Salomon Brothers）、吉列、美国航空和冠军产业公司（Champion Industries）。这些可转债的利息享有税收优惠，为伯克希尔-哈撒韦公司提供了诱人的利润，并且如果公司运营良好，把债券转换为普通股后还存在溢价空间。

1991年，所罗门兄弟公司成为一起重大财务丑闻的焦点，该公司被指控在政府国债竞标中进行价格操纵，巴菲特临危受命成为所罗门兄弟公司的CEO。巴菲特全心投入，花了9个多月的时间与政府斡旋，任命新CEO，并试图厘清所罗门兄弟公司那拜占庭皇宫般复杂的薪酬体系。最终，所罗门兄弟公司支付了一笔相对数额较小的罚金并恢复了昔日的繁荣。1996年末，桑迪·威尔（Sandy Weill）的旅行者集团（Travelers Corporation）以90亿美元的价格买下了所罗门兄弟公司，巴菲特赚得盆满钵满。

20世纪90年代初，巴菲特继续在公开市场对精选标的进行大额投资，他持股较多的有富国银行（1990年）、通用动力公司（1992年）和美国运通公司（1994年）。10年时光流逝，巴菲特将注意力再次转回到并购上，并在保险领域的两项重大交易中达到高潮：1996年以23亿美元收购了盖可保险剩余的一半股权；1998年以价值220亿美元的伯克希尔-哈撒韦公司股票收购了通用再保险公司，最大交易额也再度改写。

无界，臻于化境的投资之神

从90年代后期到21世纪初，巴菲特逮住机会就买入非上市公司，比如肖氏地毯公司（Shaw Carpets）、本杰明·摩尔涂料公司（Benjamin Moore Paints）和克莱顿家居公司（Clayton Homes）。“9·11”恐怖袭击使得很多非上市公司都被市场打入冷宫，因此收购价格较为低廉。巴菲特还通过中美洲能源公司（MidAmerican Energy）在电力行业进行了一系列重要投资。中美洲能源公司是巴菲特与奥马哈老友沃尔特·斯科特（Walter Scott）的合资企业，后者也是吉威特建筑公司（Kiewit Construction）的前任CEO。

在此期间，巴菲特还活跃在传统股市以外的各种投资领域。2003年，巴菲特在垃圾债券上下了70亿美元的重注，获得丰厚利润后随即抛售；2003年和2004年，巴菲特用200亿美元押注美元贬值。2006年，巴菲特宣布了伯克希尔-哈撒韦公司的第一次国际并购：以50亿美元收购

以色列的艾斯达公司（Istar）。艾斯达公司是行业领先的刀具和刀片制造商，被收购后，在伯克希尔-哈撒韦公司的羽翼下日渐繁荣。

在接下来的几年里，巴菲特按兵不动，直到雷曼兄弟公司破产引发了金融危机他才再度出山。出山后横刀立马，而这一时期也成为巴菲特职业生涯中最活跃的时期之一。这一时期的决战发生在2010年初，伯克希尔-哈撒韦公司购买了美国最大的铁路公司伯灵顿北方圣达菲铁路公司（Burlington Northern Santa Fe），这家公司的总估值达到342亿美元。

我们来看看这些云泥之别的数字。从1965年6月巴菲特接管伯克希尔-哈撒韦公司到2011年，公司股价实现了惊人的年化20.7%的复合增长率，使得同期标准普尔500指数年化9.3%的回报率相形见绌。在巴菲特接管时，如果投资1美元给他，45年后会变成6 265美元；如果在他首次进行股票交易时投资1美元，那么其价值会变成1万美元；同样的钱投资标准普尔500指数则只能得到62美元（见图8-1）。

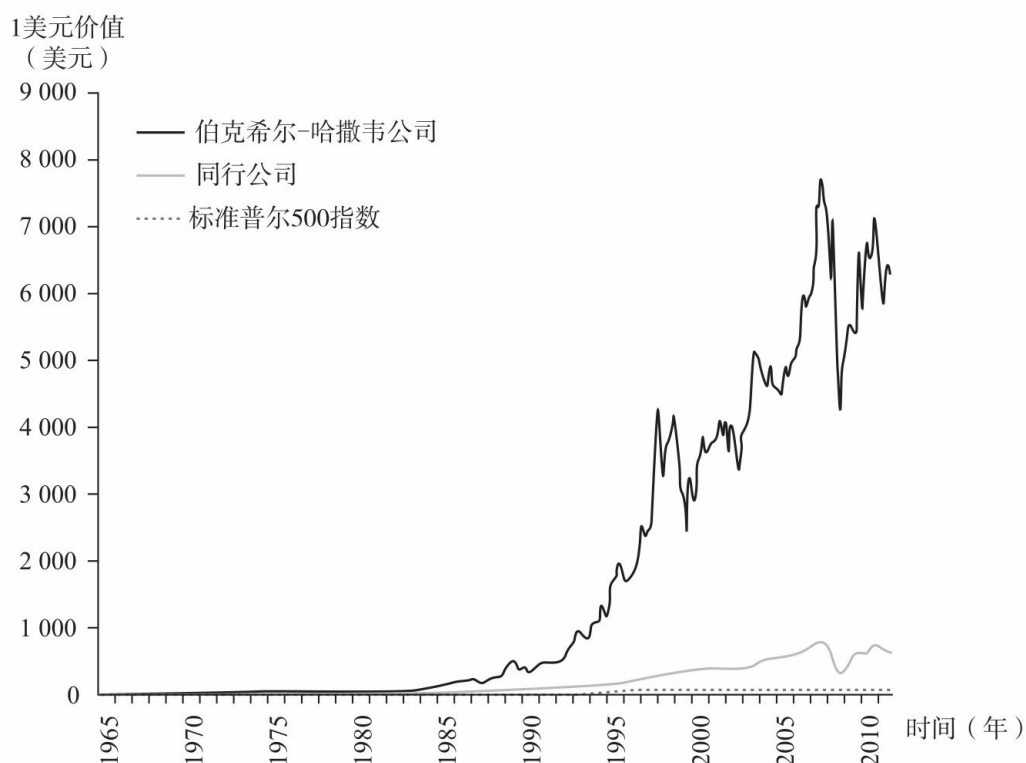


图8-1 伯克希尔-哈撒韦公司的复合收益

资料来源：根据证券价格研究中心数据制作。

在巴菲特的漫漫任期中，伯克希尔-哈撒韦公司的回报率超过标准

普尔500指数百倍之多，这是杰克·韦尔奇时代的通用电气以及任何一位同行所无法企及的。

The Outsiders

沃伦·巴菲特的方法

巴菲特的卓越业绩来自他在3个重要且内在关联的领域采用了独特的方法，这3个领域是资本生成、资本配置和运营管理。

查理·芒格曾说过，伯克希尔-哈撒韦公司长期成功的秘诀在于，它有能力“以3%的成本获得资金，然后用它们投资，获得13%的收益”，这种持续获得可用于投资的低成本资金的能力，帮助该公司在金融领域取得了成功。但是一直以来，这种能力的作用都被人们低估了。¹值得注意的是，巴菲特几乎完全回避了借债和发行股票的方法来获得资金，伯克希尔-哈撒韦公司差不多所有的投资资本都是内部生成的。

低成本的“浮存金”

尽管全资子公司的运营和偶尔出售投资标的也为伯克希尔-哈撒韦公司提供了大量现金，但是其主要资金来源却是那些保险子公司的浮存金。实际上，巴菲特在伯克希尔-哈撒韦公司创造了一个资本“飞轮”：从上述来源获得的资金被用来收购其他企业的全部或部分股权，这些企业自身又能产生现金，这些现金继续被用于新的投资，如此往复。

在很大程度上，保险业务是伯克希尔-哈撒韦公司最重要的业务，也是公司业绩非凡增长的关键所在。巴菲特对保险业务形成了独特的管理方法，与他在资本配置和运营管理方面的举措相比，有很多奇妙的相似之处。

在1967年收购国民赔偿保险公司时，巴菲特是最早认识到保险公司有能力产生低成本浮存金的内在杠杆的人物之一。用巴菲特的话来说，这次收购是伯克希尔-哈撒韦公司历史上的“分水岭”。巴菲特解释道：“浮存金是我们持有但不拥有的资金。保险公司在运营过程中，在损失和赔付之前就收取了保费，所以有了浮存金。保费收取和损失赔付之间有时会经过很多年，在此期间，保险公司用这笔钱去进行投资。”²这是另一个不走寻常路的例子，当时业内的

其他人士却根本无视这一点。

随着时间的推移，巴菲特逐渐形成了运营保险业务的独特战略，他一方面强调承保必须有利可图，另一方面强调浮存金的生成而非保费收入的增长。这种方法和其他大多数保险公司迥然不同，它有赖于对待业务的不同意愿：在保险定价较低时，即使损失短期盈利也要避免承保；相反，在保险定价有吸引力时，则应该放开手脚承接大量业务。

此种方法的运用，导致了保险业务虽有起伏，但最终的盈利却很可观，例如，1984年，伯克希尔-哈撒韦公司旗下最大的财产和意外险公司国民赔偿保险公司的保费收入达到6 220万美元。两年后，这个数字惊人地增长了约5倍，达到3.662亿美元。1989年该数字又回落了73%，保费收入为9 840万美元，并且在接下来的12年里再也没有回到1亿美元的水平。到了2004年，该公司的保费收入超过6亿美元。在此期间，国民赔偿保险公司的平均年化承保利润是保费收入的6.5%；相比之下，同时期的财产和意外险公司的平均年化数据却是亏损7%。

对于一家独立的上市保险公司来说，这种犬牙交错的收入模式（见图8-2）是无法向华尔街交差的。然而，由于多家保险子公司只是伯克希尔-哈撒韦公司这个巨型多元化公司的分支，从而让它们躲过了华尔街的审查。由此促成了一个重要的竞争优势：国民赔偿保险公司和伯克希尔-哈撒韦公司旗下的其他保险子公司得以专注于盈利能力的提升，而不是保费的增长。正如巴菲特所说：“在回报方面，我和查理·芒格总是宁可要起伏的15%，也不要平稳的12%。”³

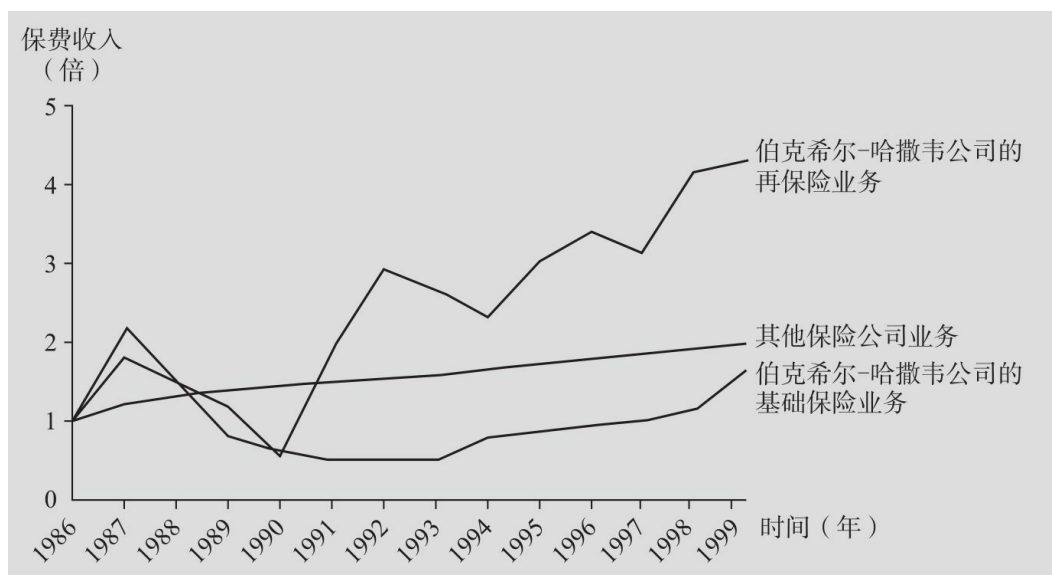


图8-2 伯克希尔-哈撒韦公司的保费增长曲线比整个行业波动幅度更大

资料来源：其他保险公司保费数据来源于“贝斯塔保险合计与平均”（Best's P&C Insurance Aggregates and Averages）中的承保保费总额。伯克希尔-哈撒韦公司的数据来源于其年报。

这一时期，伯克希尔-哈撒韦公司所有保险业务的浮存金都实现了巨额增长，从1970年的2.37亿美元增长到2011年的700多亿美元。这种不可思议的低成本资金就像火箭燃料，推动着伯克希尔-哈撒韦公司业绩的惊人飙升。而且，正如我们将看到的那样，按兵不动和果断出手交相辉映，也反映了伯克希尔-哈撒韦公司种种投资活动背后的模式。无论在保险还是投资领域，巴菲特认为长期成功的关键在于“性情”，性情决定意愿，即“别人贪婪时恐惧，别人恐惧时贪婪”。⁴

伯克希尔-哈撒韦公司的另一个重要资金来源是旗下全资公司的经营利润。在过去20年里，随着巴菲特大力扩张伯克希尔-哈撒韦公司的企业版图，这些利润变得越来越重要。1990年，全资公司的税前利润为1.02亿美元；2000年这一数字为9.18亿美元，复合增长率为24.5%；到2011年，这个数字更是达到了惊人的69亿美元。

投资选项越多，竞争优势越大

伯克希尔-哈撒韦公司运营中产生的资金就如同不断涌出的喷泉。现在我们把注意力转向巴菲特是如何调度这些资金的。每收购一家公司，巴菲特做的第一件事就是控制它的现金流，并坚持把多余的现金送回奥马哈进行配置。正如查理·芒格指出的那样：“与非

常分散的经营权不同，伯克希尔-哈撒韦公司的资本配置权是高度集中的。”⁵这种松散与紧密、授权与集权混合的政策，其他商界局外人的公司也都在采用，但通常情况下没有像伯克希尔-哈撒韦公司这样用到极致。

来到伯克希尔-哈撒韦公司之前，巴菲特已经是一位极其成功的投资者，他正是配置伯克希尔-哈撒韦公司资本的最佳人选。大多数CEO只面对过本行业内的投资机会，因此都如同刺猬一般陷入了经验的窠臼。相比之下，巴菲特是最典型的“狐狸”。得益于此前对各种证券和行业投资评估的经验，巴菲特的优势在于拥有更多的资本配置选项，包括收购非上市公司和购买公开交易的股票。简而言之，一个CEO拥有越多的投资选项，他就越有可能做出高回报的投资决定。就伯克希尔-哈撒韦公司而言，投资选项的多样性已经转化为它的一个重要竞争优势。

巴菲特的资本配置方法天下无双：他从不派发股息或大量回购股票。取而代之的是，由于伯克希尔-哈撒韦公司旗下的公司通常不需要资本投入，巴菲特则关注对上市交易的股票进行投资以及收购非上市公司。大多数CEO没有巴菲特那样丰富的投资经验，所以对他的投资选项也只能望尘莫及。不过，在研究这两个选项之前，我们先来审视他早期的一个关键决策。

和纺织业短暂的卿卿我我之后，巴菲特很早就做出选择：不再进一步投资伯克希尔-哈撒韦公司低回报的历史遗留业务西装衬里，而是聚集纺织部门所有的冗余资本，然后将其部署到其他领域。与此形成对照的是，当时和现在都是纺织业中最大龙头的伯灵顿工业公司（Burlington Industries）则选了另一条道路：1965—1985年，它将所有获得的资金都投入到已有业务中。20年间，伯灵顿工业公司股票的年增长率只有微不足道的0.6%；伯克希尔-哈撒韦公司股票的复合回报率则达到惊人的27%。二者的不同结果告诉我们关于资本配置的一个重要寓言：遇到回报诱人的生意，就与它结伴而行；遇到回报不佳的生意，逃之夭夭也很重要。

“与君结伴或一拍两散”是伯克希尔-哈撒韦公司的一个关键决策，由此也引申出对待资本配置的基本观点。相比收购等更吸引人的行为，资本配置不太受人关注，而它的重要之处在于，资本配置决定着企业的哪些业务因回报率低而不值得未来继续投资。关闭或出售前景不佳的业务，并将它们的资本集中在回报能够满足内在增

值目标的业务上时，那些作为CEO的商界局外人通常都很决绝。正如巴菲特在1985年最终关闭伯克希尔-哈撒韦公司的纺织业务时所说：“如果你发现自己身处一条老旧的破船上，那么更有效率的做法是花精力换一条船，而不是去补漏。”⁶

集中且长期持有

巴菲特最有名的还是在股票市场的投资，在他担任CEO的前25年里，股票市场是伯克希尔-哈撒韦公司资本配置的主要渠道。无论用什么标准衡量，巴菲特的公开市场回报的战绩都很骄人，这可以从不同的角度来观察。正如我们所看到的，1957—1969年，巴菲特合伙企业的平均回报率为30.4%；而《货币周刊》（*Money Week*）杂志的一项研究表明，1985—2005年，伯克希尔-哈撒韦公司的投资回报率高达25%。⁷

股票投资对伯克希尔-哈撒韦公司总体收益的重要性毋庸置疑，并且以它为窗口我们才得以窥探巴菲特更广泛的资本配置哲学。因此，我们有必要仔细观察一下巴菲特在公开市场投资方法中的一个特殊方面：投资组合管理。投资组合管理就是一个投资者持有多少只股票以及持有多长时间，它对收益有着巨大的影响。两个投资理念相同但组合管理方法不同的投资者，其结果是截然不同的。巴菲特在伯克希尔-哈撒韦公司股票投资的管理方法上，有两个重要特征区别于他人：高度集中和极长的持有期。在这两个方面，巴菲特的思考都与传统理念相悖。

巴菲特认定集中的投资组合会带来超常的回报，也认定优秀的投资理念是很稀有的。巴菲特曾多次告诫学生，如果在职业生涯之初他们就拿到一张穿孔卡片，上面的20个孔洞代表着投资生涯中可以进行的投资总数，那么他们的投资结果会比现在有所改观。正如巴菲特在1993年年报中做出的总结：“我们相信，在买入之前，如果投资者对一项生意深思熟虑，如果他们对生意的商业特性感到更加安心（他们本应如此），那么投资组合的集中策略可以很好地降低风险。”⁸

巴菲特在伯克希尔-哈撒韦公司的投资模式与其保险子公司的承保模式相类似，他长期按兵不动，但是一动就是大手笔的投资。在伯克希尔-哈撒韦公司的投资组合中，排名前5位的持仓总和通常

显著地占据公司总资产的60%~80%，而在典型的共同基金投资组合中，这一比例为10%~20%。巴菲特至少有4次将伯克希尔-哈撒韦公司账面价值15%以上的资产投资于一只股票，他还曾将巴菲特合伙公司40%的资金投资于美国运通公司。

巴菲特投资组合管理方法的另一个显著特点是超长的持有期。2011年，除了于当年购买的IBM股票外，巴菲特的前五大持仓股票平均持有时间超过20年；相比之下，典型的共同基金的平均持有期还不到一年。也就是说，巴菲特投资的活跃度极低，他称之为“无所事事，近乎懒惰”。

高度集中和极长的持有期这两条投资原则结合在一起，形成了一个强有力的、具有高度甄别性的过滤器，极少有公司能顺利通过筛选。

有趣的是，尽管巴菲特一直倡导股票回购，但他是本书中唯一一位没有大量回购自己公司股票的首席执行官（早期有过几次例外，买回了少量股票）。尽管巴菲特欣赏并鼓励其他首席执行官进行回购，自己却一直认为，对伯克希尔-哈撒韦公司而言，回购有悖于其独特的、合作伙伴式的企业文化，并可能会破坏多年来通过诚实直率的沟通和出色的回报所建立的信任纽带。

话虽如此，面对机会却坐以待毙，那绝不是巴菲特。曾经有两个时间段，伯克希尔-哈撒韦公司股票的交易价格远低于其内在价值，此时巴菲特打破常规尝试回购。一次是在2001年初，当时互联网泡沫破灭，伯克希尔-哈撒韦公司股价暴跌；更近的一次是在2011年9月，当时巴菲特宣布将以不高于账面价值1.1倍的价格回购大量股票。在这两次事件中，伯克希尔-哈撒韦公司的股价都迅速上涨，从而阻止了更大数量的回购。

伯克希尔-哈撒韦公司资金的另一个主要出口是收购非上市公司，在过去的20年里，这已悄然成为巴菲特的主要投资渠道，并在2010年初大规模收购伯灵顿北方圣达菲铁路公司时达到顶峰。针对收购，巴菲特自创了一套举世无双的方法体系。表8-2将该方法体系与传统的私募股权公司的方法进行了比较。

表8-2 巴菲特与私募股权公司对非上市公司收购方法的比较

比较项目	巴菲特	私募股权公司
持有期	永远	5 年以内
管理层	现任 CEO 留任	新任 CEO 上任（经常）
杠杆	从不使用	大量使用
交易方式	直接购买	拍卖获得
收购后的管理互动	偶尔	经常
成本削减	从不	经常
尽职调查	粗略	详尽
咨询外部顾问	从不	总是
薪酬体系	简单	复杂

·两种有趣的模式·

对于那些有兴趣深入研究巴菲特在股票市场投资的人来说，还有两种模式值得注意。

第一个是根深蒂固的逆势而为。巴菲特经常引用本杰明·格雷厄姆关于“市场先生”的比喻：“有一个叫‘市场先生’的家伙总是很热情，他每天都会出现，要么从你那买入，要么卖给你.....这家伙越是狂躁抑郁，投资者就越能获得更大的投资机会。”（伯克希尔-哈撒韦公司1977—2011年年报。）当格雷厄姆的“市场先生”抑郁到极点之时，巴菲特却在稳步地大举买入。伯克希尔-哈撒韦公司在公开市场主要的大笔投资都以某种行业或公司危机为契机，这些危机掩盖了有巨大潜力的公司业务的价值，表8-3表明了这一点。

表8-3 巴菲特的逆势而为

公司	初始投资时间（年）	逆势而为的背景
美国运通	1964	色拉油丑闻
《华盛顿邮报》	1973	政府广电执照危机
盖可保险	1976	潜在的破产危机
富国银行	1989	南加州衰退和房地产危机
房地美	1989	经济萧条和存贷危机
通用动力	1992	后冷战时期国防工业衰退

第二种模式在于，投资时机要与重大的管理或战略转变相契合。当一家拥有优秀的“特许经营权”业务的公司投资了其他低回报的业务时，就会出现投资机会。巴菲特用业余选手和职业选手混搭的高尔夫赛事对这种投资机会做出比喻：“即使所有的业余选手都是毫无得分希望的傻瓜，凭借职业选手独占鳌头的技术依然可以得到最佳击球分数，整个球队也会因此备受尊崇。”（伯克希尔-哈撒韦公司1989年年报。）当巴菲特看到新的管理团队将业余人士从赛场上的4人组中除名，并将注意力重新放在公司的核心业务上时，他就会密切关注，如表8-4所示。

表8-4 巴菲特密切关注混搭模式

公司	投资时间（年）	管理或战略转变
盖可保险	1976	聚焦核心保险业务，新 CEO 上任，已经持有其股份
通用食品	1979	聚焦核心品牌，新 CEO 上任，进行回购
可口可乐	1988	剥离非核心业务，进行回购
通用动力	1992	剥离非核心业务，进行回购，新 CEO 上任
美国运通	1994	剥离雷曼兄弟公司，新 CEO 上任，已经持有其股份

巴菲特为大型非上市公司的卖家创造了一个极具吸引力且差异化程度高的新选项，它介于首次公开募股和出售给私募股权投资者之间。出售给伯克希尔-哈撒韦公司可以享受无与伦比的待遇，它允许公司原来的所有者或经营者获得流动资金，同时无投资人干扰和无华尔街监管的情况下继续经营公司。巴菲特提供了一个完全

摒弃官僚作风的公司环境，并为有价值的项目提供无限的资金支持。这套方案与私募股权公司的方案有很大区别，后者的方案中允许投资人深度介入公司管理，并且通常在5年的持有期后会再次卖出公司。

巴菲特从不参与拍卖。正如中美洲能源和利捷航空（MidAmerican Energy and NetJets）的CEO戴维·索科尔（David Sokol）告诉我的那样：“我们绝不会被竞价的兴奋冲昏头脑。”⁹相反，引人注目的是，巴菲特创造了一种机制，在这种机制中，顶尖的非上市公司的所有者都会给他致电。巴菲特拒绝谈判桌上的价格拉锯战，而是让有意出售的卖家与他联系并报价。巴菲特承诺“通常在5分钟或更短时间内”给出答案。¹⁰这一承诺逼迫潜在的卖家第一时间给出他们的底价，并确保巴菲特在这短短的5分钟之内做出决策。

巴菲特没有在传统的尽职调查上花费大量时间，而是以超乎寻常的速度，通常是在第一次接触后的几天之内就达成交易。在决定收购之前，巴菲特从不参观企业的运营设施，也很少与其管理层会面。汤姆·墨菲告诉我：“大都会通信公司是伯克希尔-哈撒韦公司有史以来最大的投资之一……只花了15分钟就谈妥了交易，并就条款达成一致。”¹¹

作为授权管理的大师，巴菲特却从未就资本配置决策进行过授权。伯克希尔-哈撒韦公司没有商务拓展团队或投资委员会；除了查理·芒格之外，巴菲特从不依赖投资银行家、会计师或律师的建议，他自己做分析工作并亲自处理所有谈判。巴菲特从不看中介机构提供的预测，而更愿意关注历史财务报表，然后做出自己的判断。巴菲特之所以行动迅速，是因为只收购那些自己熟知领域的公司，如此一来便能迅速聚焦关键的运营指标。查理·芒格在谈到伯克希尔-哈撒韦公司的收购方式时说：“我们不试图进行收购，我们只是坐等无脑买入的机会。”¹²

良性的无政府状态

除了是他同代人中最伟大的投资者外，事实证明，面对伯克希尔-哈撒韦公司经营业务的日益增长和多元布局，巴菲特还是一位极有效率的管理者。在过去的10年里，伯克希尔-哈撒韦公司的每

股收益大幅增长，尽管公司规模庞大、业务多元，但公司的运营效率非常高，它的有形资产回报率一直位居《财富》500强公司的前1/4。

那么，巴菲特是如何实现这些经营业绩的呢？将他的方法与杰克·韦尔奇的进行比较（见表8-5），不难看出巴菲特和蔼可亲的外表之下，呈现的是一个非同寻常的CEO。杰克·韦尔奇在通用电气推行的体系风风火火，它强调集权的战略举措，强调CEO轮换制以及马不停蹄的出差和会议节奏。尽管巴菲特非常钦佩杰克·韦尔奇的能力，但是他们管理风格的反差极富戏剧性。

表8-5 杰克·韦尔奇和巴菲特管理方法的比较

比较项目	杰克·韦尔奇	巴菲特
盈利模式	平稳	震荡
雇员人数	40 万	27 万
总部人员	数千人	23 人
出差	非常多	几乎不
主要活动	会议	阅读
投资者关系维护时间	大量	无
工作日节奏	狂热、繁忙	安静、无固定日程
更换管理层	大量	几乎不
场外会议	经常	从不
战略规划	定期	从不
股票分拆	是	否

巴菲特在没有任何相关运营经验的情况下承担起CEO的角色，他有意识地建构伯克希尔-哈撒韦公司，让其允许自己把时间集中花在资本配置上，而尽可能少地花在运营管理上，对于运营他觉得自己无所作为。这样做的结果就是，运营管理的极度分权成为伯克希尔-哈撒韦公司体系的标志。如果说特利丹公司、大都会通信公司和本书中的其他公司拥有分权管理的风格与理念，那么相比之下，伯克希尔-哈撒韦公司奉行的就是彻头彻尾的无政府主义。

在拥有超过27万名雇员的公司中，奥马哈公司总部只有23个人。伯克希尔-哈撒韦公司旗下的公司不用召开定期的预算会议，

子公司的CEO们除了主动给巴菲特打电话寻求建议或者为业务争取资金外，几乎从没接到过巴菲特的电话。巴菲特把这种管理方法总结为“雇得好，管得少”，并且相信这种极致的分权方式可以削减开支并释放企业活力，从而提高组织的整体效率。¹³

避免同侪压力的影响

在1986年的伯克希尔-哈撒韦公司的年报中，巴菲特描述了一个现象，正如我们在引言中看到的那样，这就是拥有惊人力量的“惯性驱使”，它导致管理者盲目地模仿同行。巴菲特经常引用丘吉尔的名言并洞悉其中的深意，他有意识地构建自己的公司和生活并避免它们对自己的驱使。与其他《财富》500强公司的CEO的时间安排不同，巴菲特自己管理自己的日程以免无谓的干扰，并且保留不被干扰的阅读和思考时间（每天要阅读5份报纸和数不清的年度报告）。巴菲特为保有一张空白的日程表、完全没有例会而感到自豪，他的办公室里也没有电脑，也没有股票报价机。

巴菲特处理投资者关系的方法也很独特，而且是原创的。巴菲特估计CEO们平均花费20%的时间与华尔街沟通。相比之下，巴菲特从不花时间见分析师，从不出席投资会议，也从不提供季度的盈利报告。巴菲特更喜欢通过详尽的年报和年度大会与投资者沟通，这两种方法独步天下。

伯克希尔-哈撒韦公司的年报印刷在普通的无涂层纸上，配以简洁的单色封面，看起来就与其他公司的年报不同。这份报告的核心是巴菲特撰写的详细回顾公司过去一年中各项业务的长篇文章，其中，卡罗尔·卢米斯（Carol Loomis）提供了编辑方面的协助。文章风格直接、形式活泼、内容简洁明了，堪称典范；报告包括每个运营部门的详细信息以及一本“所有者手册”，手册清晰地勾勒出巴菲特和查理·芒格与众不同的运营理念。

伯克希尔-哈撒韦公司的年度大会也是独一无二的。会上管理团队代表发言的部分一般不超过15分钟，之后在长达5个小时的时间里，巴菲特和查理·芒格回答股东提问。年会吸引了大量的人群，其中超过3.5万人参加了2011年的会议。巴菲特喜欢把年度大会称作“资本主义的伍德斯托克(17)狂欢节”。¹⁴年度报告和大会强化了重视节俭、崇尚独立思考和长期管理的强大的企业文化。除此

以外，年度大会还充满着奇思妙想和幽默感。巴菲特在20世纪90年代初做了一件出格的事，购买了一架公务飞机，他戏称其为“不可原谅号”，并在年报中以滑稽的小号字体进行了披露。

另一种非传统股东的做法与股票分拆有关。众所周知，巴菲特一直避免分拆伯克希尔-哈撒韦公司的A股，2011年A股的股价超过12万美元，是纽约证券交易所股票中价格第二高者的50多倍。巴菲特认为股票分割纯粹是为了粉饰门面，并把这个过程比作将一个比萨四刀分成八块：比萨分割前后的卡路里没有发生变化，股票分拆前后资产的价值也没有变化。避免股票分拆形成了另一个过滤器，它有助于伯克希尔-哈撒韦公司自主选出其股票的长期持有人。1996年，巴菲特勉强同意创建一个价格较低的B类股，其交易价格为A股的1/30，并成为纽约证券交易所价格第二高的股票。与此相关的是，在2010年初与伯灵顿北方圣达菲铁路公司的交易中，顾及该铁路公司的小额投资者，巴菲特同意将B股按照50:1的比例进一步分拆。

以上所有这些加起来，比商业或投资战略更强大，巴菲特已经形成了一种世界观，其核心是强调与优秀的人和企业建立长期关系，避免不必要的变化。这种变化会中断经济复利增长的强大链条，而这种复利链条正是创造长期价值的精髓。

·巴菲特和《萨班斯-奥克斯利法案》·

巴菲特治理公司的方法也有别于传统，并与《萨班斯-奥克斯利法案》中的许多条款相冲突。巴菲特认为，最好的董事会是组成人数较少（伯克希尔-哈撒韦公司仅有12名董事），董事们拥有丰富的商业经验并持有公司大量的股票。巴菲特要求所有董事把大笔的个人资金投资于伯克希尔-哈撒韦公司的股票。巴菲特坚信董事们应该对糟糕的决策承担后果（伯克希尔-哈撒韦公司不为其董事提供保险），坚信董事们不应该依赖董事费收入。因此，伯克希尔-哈撒韦公司的董事费是最低的。

按照《萨班斯-奥克斯利法案》的标准，这种做法带给巴菲特的是一小撮“内部人”，这与大多数上市公司的董事会形成鲜明对比。后者的成员很少同股东一起向公司投入大笔个人资金，他们的损失有保险托底，董事费通常在总收入中占比很高。哪种方法能更好地让董事与股东结成同盟呢？

实际上，理解巴菲特最好的方式是把他当作一个经理人、投资者或哲学家，其主要目标是减少变化。伯克希尔-哈撒韦公司许多打破常规的政策都有一个共同的目标，那就是选择与最优秀的人和企业为伍，以此降低由于管理者、投资人或股东流失带来的巨大财务成本和人力成本。对巴菲特和查理·芒格而言，在选择与优秀者同行及避免不必要的改变之间，存在令人信服的、禅宗般的逻辑。这不仅是一种可以获得可观经济回报的方式，而且是一条可以获得更为平和生活的大道；在他们可传授的众多经验中，这些长期关系背后的力量可能是最重要的。

结语 伟大的CEO都善于打好手中的牌

本杰明·格雷厄姆

你之所以正确，不是因为别人的赞同，而是因为你的事实正确以及推导合理。

威廉·德莱赛维茨 (William Deresiewicz),
2009 年 10 月西点军校新生讲座

一个人之所以成为领导者，那是因为他能够独立思考。

通过图9-1，让我们来回顾一下，投资局外人CEO的1美元其最终价值与投资同行、大盘和杰克·韦尔奇的不同。

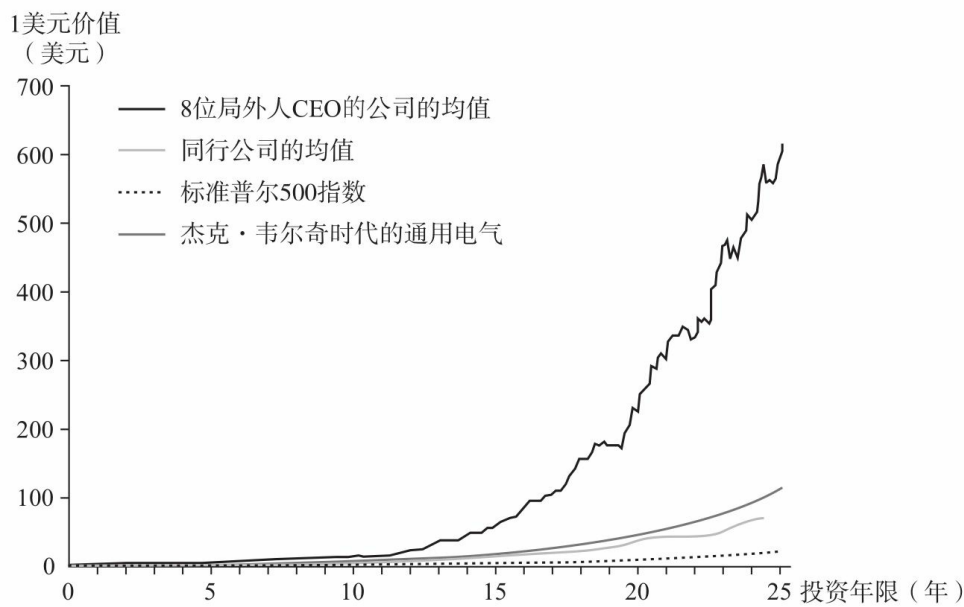


图9-1 1美元价值变化

图9-1令人印象深刻，数字也不言自明，它很好地概括了这些杰出高管的成就。然而，这些惊人的记录大部分是在20世纪收集的。这就引出一个问题：在当今快速变化的竞争环境中，这些局外人CEO的经验和

教训对管理人和投资人是否仍然有意义呢？答案在两家更为年轻的公司中得到了证实：一家是小公司，预付费法务公司（Pre-Paid Legal）；另一家是大公司，埃克森美孚公司（ExxonMobil）。

预付费法务公司直到2011年还是一家上市公司，它向个人和企业出售法律服务方案。这种法律服务方案实际上是一种保险产品：客户提前缴纳年费后，未来一年发生包括诉讼、房地产、信托和遗嘱等方面的法律需求时，就可以得到免费的法律服务而无须缴费。这种服务方案的推出时间是20世纪70年代，而在整个80年代和90年代，预付费法务公司的业务都在迅猛增长。有趣的是，该公司在经历了最初一轮强劲增长后，其营业收入在之后的10年中几乎停滞。

预付费法务公司业绩快速增长之后随之长期停滞，从历史上看，这样的模式是股市大幅下跌的罪魁祸首。但是，在同一时期内，预付费法务公司的股价却上涨了4倍，大大超过了市场和业内同行。预付费法务公司是如何取得如此战绩的呢？从1999年底开始，该公司CEO哈兰·斯通西弗（Harland Stonecipher）意识到，他所处的市场已经饱和，为了增长而追加投资已经不太可能获得高收益的回报。在董事会的敦促下（其董事会里有几位大投资人，这对于一家上市公司来说不太常见），斯通西弗开展了一系列的、积极优化自由现金流的项目，并通过积极的股票回购计划，有步骤地将资本返还给股东。在接下来的12年里，斯通西弗回购了50%以上的流通股，赢得了股东和股票市场持续的掌声，2011年6月他同意将公司以高额溢价出售给一家私募股权公司。

预付费法务公司属于少数人控股的小公司，可能不足以说明问题，因此在那些大公司中找到另一个例子就显得尤为重要。接下来是个绝对的大公司，它就是埃克森美孚公司，一度是全世界市值最大的公司。自1977年以来，埃克森美孚公司以及后来的埃克森石油公司为其投资者创造了惊人的15%的年复合回报率。鉴于其规模，这确实是一个很了不起的纪录，令市场和同行相形见绌。当探究其管理者是如何取得这些成果时，我们会发现他们的做法与局外人CEO们有很多惊人的相似之处。

局外人视角的8个特点

特点1，总是做数字运算

局外人CEO们总是先询问：回报是多少？每一项投资都会产生回

报，涉及的运算其实只是小学五年级的算术，但这些CEO始终如一地做着小学数学题，他们使用保守的假设，只推进那些能够提供令人信服的回报的项目。他们专注于关键性的假设，不相信过于详尽的电子表格；他们自己进行分析，也不依赖于下属或顾问。这些局外人CEO认为，财务分析的价值取决于假设的质量，而不是报告的页数。他们中的许多人开发了简洁的、一页纸就能写完的财务分析模板，让员工们关注公司发展中的关键变量。

在丹尼尔·卡尼曼（Daniel Kahneman）[\(18\)](#)的优秀著作《思考，快与慢》（*Thinking, Fast and Slow*）中，他阐述了一个人类决策的模型，该模型是他历时30年的研究成果，并因此获得诺贝尔经济学奖。¹卡尼曼的模型中有两个不同的系统。系统1是纯粹本能化的认知模式，以便人们在任何情况下都能瞬间参与决策过程，并使用经验法则快速做出决策；系统2相对迟缓，反射弧也更长，它允许人们做出更复杂的分析。系统2的决策质量碾压系统1，但问题在于，系统2的生效需要花费更多的时间和精力，因此我们中的许多人并没有充分利用它。

按照卡尼曼的观点，让系统2发挥作用的关键通常需要催化剂或触发器，而对于局外人CEO们来说，这些看似简单的“一页纸”分析往往就起到了这一作用。这“一页纸”确保了对经验数据的关注，并防止了盲目从众。因此，这些“一页纸”就像一剂预防世俗观点的疫苗，在局外人的公司中广泛播种。特利丹公司CEO亨利·辛格尔顿曾对《福布斯》杂志说过：“我们的管理者严格遵循投资纪律，我们极少接到低回报的投资建议。”²

埃克森美孚公司在其CEO雷克斯·蒂勒森（Rex Tillerson）及其脾气暴躁的前任李·雷蒙德（Lee Raymond）的领导下，也有类似的纪律，他们要求所有投资项目至少要达到20%的回报。在2008年的金融危机期间，由于能源价格下跌，蒂勒森和他的团队随即降低产量，此举受到华尔街分析师的批评。然而，他们只是拒绝向回报不足的项目注入更多的资金，即使这意味着短期利润的下降。

特点2，分母很重要

这些CEO都非常关注每股价值的最大化。要做到这一点，他们不是只简单地关注作为分子的公司总市值。总市值可以通过任何方式来得到增长，包括支付过高的收购费或注资那些没有经济效益的内部资本项

目。这些CEO还一心一意地通过投资项目的谨慎融资以及不失时机的股份回购来管理分母。这些回购与目前流行的两种回购理由相向而行，即它们并不是为了支撑股价或抵消期权的影响，而是因为它们能提供有吸引力的回报，它们本身就是投资。

在大型能源公司中，只有埃克森美孚公司一直积极地回购自家股票，在过去5年中回购了已发行股份的25%以上。在雷曼兄弟公司破产后的危机中，埃克森美孚公司实际上还加快了回购步伐。

特点3，独断专行

局外人CEO们是授权大师，他们管理着高度去中心化的组织，并将运营决策权下放到组织中最低的、最区域化的层级。不过，他们并没有下放对资本配置的决策权。正如查理·芒格向我描述的那样，他们公司是“对日常运营分散管理和对资本配置高度集中的奇怪组合”，这种松散与紧密、授权与专制的混合体，被证明能够有效地对抗“惯性驱使”的力量。³

他们除了独立思考之外，还对外部顾问不屑一顾，并觉得这是天经地义。一切就像西部电影《正午》（*High Noon*）里的场景那样出乎意料：约翰·马隆独自一人面对美国电话电报公司企业开发部员工、律师和会计师构成的方阵；斯蒂利茨带着一个黄色的便笺簿独自出席可能高达数十亿美元交易额的尽职调查；巴菲特在一天之内就做出了对伯克希尔-哈撒韦纺织公司的收购决定，甚至没有再去公司看看……

特点4，个人魅力无关紧要

这些局外人CEO也明显不会自我推销，在维护投资者关系方面所花的时间比同行少得多。他们不提供收益指南，也不参加华尔街论坛，作为一个群体，既不张扬，也没有超凡的魅力。他们为人谦逊，吉姆·柯林斯（Jim Collins）在他那本《从优秀到卓越》（*Good to Great*）中也强调过这种品质。他们不寻求通常也不吸引聚光灯，然而他们的回报却大大弥补了这种低调。

蒂勒森也一样，他参与了埃克森美孚公司的所有重大资本配置决策，很少参加电话业绩说明会或出席行业论坛。在华尔街分析师口中，蒂勒森以简明扼要的沟通风格而著称。

特点5，像鳄鱼般的性情，极具耐心

对回报的要求是他们耐心的来源，他们中的绝大多数都愿意为了合适的机会等待漫长的时间，其中大众影院公司的迪克·史密斯整整等待了10年，只有马隆是个例外，他为了追求规模不断收购有线电视公司。与凯瑟琳·格雷厄姆一样，他们中的许多人只是在收购狂热时期保持观望，避免定价过高的“战略性”收购，从而创造了巨大的股东价值。

埃克森美孚公司在超过10年的时间里还没有完成一笔重大收购，直到2011年才有所改变。

特点6，该出手时，绝不手软

有趣的是，我们发现，这种对经验主义和分析的偏好并没有导致怯懦。事实恰恰相反：在极少数情况下，当发现项目有确定性的高额回报时，他们就以迅雷不及掩耳之势果断行动。每一家公司都至少进行了一次大额收购或投资，金额相当于他们公司企业市值的25%或更多。汤姆·墨菲对美国广播公司的交易金额甚至大于他整个公司的市值。

1999年，当时油价处于历史低点，埃克森公司以一笔轰动一时的交易收购了竞争对手埃克森美孚公司，交易总额超过其公司市值的50%。

特点7，在大大小小的决策中始终贯彻理性分析的方法

这些资本行业的高管像外科医生那样操作精准，总是把可用的资本投向效率最高、回报最高的项目。提升价值的决策不断累积，同样重要的是有损价值的决策被有效避免。经过长期发展，以上原则对股东价值就产生了巨大的影响。事实证明，这种非传统的思维模式本身构成了局外人公司实质性的、可持续的竞争优势。这种思维模式的作用相当于偏光镜片，让局外人CEO们能够穿透同行的行为和世俗智慧的光芒，发现核心的经济现状，并据此做出决策。

在这本书中，有大量的例子展示了这种务实的分析方法所带来的简洁和高效。这些CEO很清楚自己在寻求什么，他们的员工也很清楚。他们没有过度分析或过度建模，也没有通过那些不请自来的外部顾问或银行家来证实自己的想法。比尔·斯蒂利茨在罗尔斯顿·普瑞纳公司的长期

副手帕特·马尔卡希说过：“我们知道需要关注什么，就这么简单。”⁴

2009年，《巴伦周刊》的一篇文章描述埃克森美孚公司“与众不同的”企业文化时，称其“以牺牲自我为代价，冷血地关注回报”。⁵一切并非偶然，这种节俭的企业文化产生了非凡的结果：在过去的25年里，埃克森美孚公司的股东回报率在石油和天然气行业一直居于领先地位。

·一个预言·

如今，创纪录的企业现金水平，加上普遍较低的利率和市盈率，为积极的资本配置提供了历史性机遇。这种情况在思科、微软和戴尔等最大的核心蓝筹股技术公司中尤为明显，很多这样的公司仍由创始管理团队的成员所运营，它们拥有数字庞大的现金余额，其股票交易的市盈率正处在前所未有的个位数。我认为，这其中的某一家公司有可能会扭转过去对研发投资的重视程度，转而通过大幅增长的股票回购或分红来提高股东收益。如果真是这样，那么市场的反应可能会欣喜若狂，人们想象着其中一家公司的CEO是亨利·辛格尔顿复生，面对如此良机，众人摩拳擦掌，跃跃欲试。

特点8，一种长远视角

尽管天性节俭，但局外人CEO们依然愿意投资自己的企业，以获得长期价值。要做到这一点，他们需要忽略像跑步机一样枯燥无味的季度收益，不去理会那些华尔街分析师的声音，不看诸如《扬声器》

（*Squawk Box*）和《我为钱狂》（*Mad Money*）之类聒噪的不停地强调短期思维的有线电视节目。墨菲坚持要对新印刷厂增加大笔资本支出，马隆在20世纪90年代末购买昂贵的尖端机顶盒，他们这么做是有意识地牺牲短期收益，从而改善客户体验并捍卫公司长期的竞争地位。

这种长远视角往往导致和主流观点背道而驰的行动。埃克森美孚公司的减产决定备受争议，但是在金融危机期间，主要的能源公司中只有它一家坚决保持能源勘探方面的投入，以期提高企业的长期价值。2009年初能源价格暴跌后，其他大型石油公司纷纷撤出加拿大油砂市场，但埃克森美孚公司依然在加拿大艾伯塔省（Alberta）开展了一个大型勘探项目，尽管他们知道这样做会影响短期收益。

关键是，打好手中的牌

与预付费法务公司和埃克森美孚公司的成就形成鲜明对比的是美国最大的金融机构花旗集团。2005年前后，当抵押和杠杆导致最严重的金融海啸之时，花旗集团的CEO查克·普林斯（Chuck Prince）曾说过一句名言：“只要音乐还在播放，你就必须站起来跳舞。”⁶在世俗观点的诱惑和“惯性驱使”下，普林斯无法自拔，他和他的股东们很快就会跳下悬崖，因为花旗集团的股票从2007年每股最高40美元暴跌到2009年初的不到3美元。同期的大盘和行业表现已经令人不寒而栗，但是普林斯比标准普尔500指数和同行表现得更差。

把这些CEO以及他们公司的业绩和其他人区分开的是两种截然不同的思维模式。像斯通西弗和蒂勒森这样的局外人CEO，当其他人在一旁围观时，他们往往下场起舞；而当音乐最欢畅时，他们却羞涩地退居角落。他们是聪明的逆向投资者，当回报没有吸引力时，他们愿意无限期地靠在墙边等待。

在不断变化的行业和市场条件下，局外人群体不约而同地围绕着一套极为相似的核心原则起舞。从根本上说，斯通西弗、蒂勒森和他们的局外人伙伴相比同行取得了非凡业绩，是因为在同行们后知后觉的急转弯时，他们已经先人一步地扭转了方向。如表9-1所示，他们在转向中遵循了几乎完全相同的蓝图：他们鄙视分红，严守收购底线，选择性地使用杠杆，大量回购股票，最小化税负，施行去中心化的组织管理，并专注于现金流而不是年度净利润。

表9-1 局外人共享的世界观

局外人 CEO	初次担任 CEO	分红	回购 30% 以上	收购金额超过其市值 25%	去中心化的组织架构	听从华尔街指导	个性化指标	对税负的关注
亨利·辛格尔顿	✓	否	✓	✓	✓	否	特利丹回报	高
沃伦·巴菲特	✓	否	—	✓	✓	否	浮存金	中等偏高
汤姆·墨菲	✓	低	✓	✓	✓	否	现金流回报率	中等偏高
约翰·马隆	✓	否	✓	✓	✓	否	EBITDA	高
迪克·史密斯	✓	低	✓	✓	✓	否	现金收益	高
比尔·安德斯	✓	低 / 特殊分红	✓	✓	✓	否	现金投资回报率	高
比尔·斯蒂利茨	✓	低	✓	✓	✓	否	内部收益率	高
凯瑟琳·格雷厄姆	✓	低	✓	✓	✓	否	现金内部收益率	中等偏高

再强调一遍，这些局外人高管的手牌都截然不同，所以重要的是，如何打好自己的手牌。他们的境遇千差万别：柏林墙倒塌后，安德斯面临的境遇，与20世纪70年代早期有线电视爆发期马隆接手TCI时面临的境遇截然不同。在给定的环境下进行优化是其中的关键所在，类似的例子比比皆是：一位高中足球教练每年都要调整他的策略以适应球队中不断变化的球员组合，如果四分卫很弱，他就必须带球跑动；一位剧团负责人则必须选择排练的剧目，从而把不同演员独特的演艺才能嫁接到一出戏里。

这里没有严格的公式，也没有硬性的规定：回购自家股票、进行并购或在一旁观望，并不是放之四海而皆准的真理，正确的资本配置决策总会因时而变、因势而变。这就是为什么辛格尔顿认为学会变通是如此重要。作为一个群体，这些CEO在面对商业世界固有的不确定性时，靠的是耐心、理性、务实地把握时机，而不是靠一套详细的战略计划。

他们的具体行动源于共同的、更广泛的思维模式，归总起来无异于一种CEO取得成功的新模式，这种模式以对企业资源进行优化管理为中

心。尽管局外人群体中的CEO都很有天赋，但相比同行，他们的优势在于性情，而非智力。归根结底，他们认为重要的是要头脑清晰地制定决策，他们认同的文化强调那些看似过时的美德，比如节俭和耐心、独立和偶尔的勇敢、理性和逻辑思维能力。

事实证明，在五花八门的行业和市场条件下，他们那些打破常规的做法是其竞争优势的重要来源。如表9-2所示，从根本而言，这些高管践行着一种激进的理性。与大多数CEO截然不同，他们拥有投资者或所有者才具备的长远视角，而不是那些高薪雇员的短期视角。

表9-2 偶像颠覆主义概览

对比项目	局外人 CEO	其他同行 CEO
经验	初次担任 CEO 时只有非常少的管理经验	拥有格拉德威尔所说 1 万小时经验的管理者
主要活动	资本配置	运营管理和外部沟通
目标	长期每股价值最大化	扩大规模
核心指标	边际效益、回报、自由现金流	营业收入、年度净利润
个人品质	善于分析、节俭、独立	具有个人魅力、外向型
目标	长期	短期
类似何种动物	狐狸	刺猬

那么，让我们回到本章开头的问题：这些CEO的经验和教训会对谁有所裨益呢？简而言之，答案就是：几乎所有的经理人或企业主都适用。好消息是，要成为一位高效的CEO，你不必是营销或技术天才，也不必成为魅力超凡的预言家。然而，你的确需要明白资本配置，并认真考虑如何最好地配置公司资源，从而为股东创造价值。你必须有意识地经常反问自己“回报是什么”，并且只推进即使基于保守估计也能产生诱人回报的项目。你必须有信心偶尔做一些与同行不同的事情。遵循以上原则的管理者和企业家，他们致力于理性思维和独立思考，我们有理由相信他们会把手中的牌用到极致，并使股东感到心情愉悦。

尾声：老马，旧途

拉迪亚德·吉卜林 (Rudyard Kipling)
《如果》(If)
如果你能在周围所有人失去冷静时保持理智……

诺贝尔化学奖得主路易斯·巴斯德 (Louis Pasteur) 说过, “机遇偏爱有准备的头脑”。说到有准备的头脑, 让我们来看看两位仍然活跃在舞台上的局外人CEO巴菲特和马隆, 看看他们在2008年9月雷曼兄弟公司倒闭后的那场金融危机中是如何乘风破浪的。

如你所料, 两人追求的方向与同行截然不同。当几乎所有的美国公司都紧握现金, 守着糟糕的资产负债表作壁上观时, 这两只狮子正在寒冬里积极地四处觅食。

2001年“9·11”事件后, 巴菲特很长一段时间处于蛰伏状态, 但是2008年9月后他却进入了其漫长职业生涯中最活跃的一个时期。自2008年第四季度以来, 巴菲特在种类极其繁多的投资活动中投入了800多亿美元, 其中150多亿美元是在雷曼兄弟公司倒闭后的25天内投入的:

- 从高盛集团和通用电气购买了80亿美元的可转换优先股。
- 购买了合计90亿美元的大量普通股, 比如美国联合能源公司 (Constellation Energy) 的股票。
- 向玛氏箭牌 (Mars / Wrigley, 65亿美元) 和陶氏化学 (Dow Chemical, 30亿美元) 提供夹层融资。
- 在公开市场购买了各种次级债券, 89亿美元。
- 伯克希尔-哈撒韦公司以265亿美元收购了伯灵顿北方圣达菲铁路公司77.5%的股份, 这是伯克希尔-哈撒韦公司有史以来最大的一笔交易。
- 以87亿美元收购了行业领先的上市润滑油公司路博润 (Lubrizol) 。
- 宣布对IBM股票进行新的大规模投资, 109亿美元。

与此同时，马隆一直在悄无声息地进行着扩张试验，他从TCI最初的自由媒体公司等节目类资产中分拆出来不同的公司实体，以此进行激进的资本配置。在金融危机最严重的时候，马隆做了下面这些：

- 对卫星节目巨头DIRECTV实施了“杠杆式权益成长”战略，提高负债水平并积极回购股票，24个月里回购已发行股票的40%以上。
- 对前自由媒体公司的下属实体企业发起了一系列“攻坚战”，包括：剥离有线电视节目商星光电视台的昂科拉电影频道（Starz/Encore）；以及自由都市公司（Liberty Capital）和自由互动公司（Liberty Interactive）之间的债转股交换。自由都市公司是马隆控股的公共和私有资产集合体；自由互动公司是QVC购物网和其他在线实体的所有者。
- 2009年初，在市场最低谷时进行了一次过程烦琐却颇有吸引力的交易，通过自由都市公司一举收购了卫星广播服务公司天狼星广播（Sirius Broadcasting）的控股权。2010年第二季度，马隆还回购了自由都市公司11%的股份。
- 通过旗下国际有线电视公司Liberty Global，宣布了该公司历史上最大的一笔收购，即以不到该公司现金流7倍的超过50亿欧元的价格收购了德国有线电视公司Unitymedia，同时以超过9倍现金流的价格出售了该公司在日本最大的有线电视业务的大量股份。所有收益都可以通过公司巨大的净运营亏损来实现避税。马隆还让国际有线电视公司继续执行声势浩大的回购计划，该公司在过去5年里回购了一半以上的股份。

看到了吗？当美国企业界普遍站在场外不知所措时，这两位老谋深算的CEO却加入凯恩斯“动物精神”的投资狂欢之中。他们正应了巴菲特的那句名言：在同行因前所未有的恐惧而颤抖时，自己则非常贪婪。

后记 一个案例和一份清单

让我们用一个案例来收尾，这个案例展示了局外人在不同情境下的工作方式。

假定你拥有一家运营良好的高档面包店，专门售卖法式面包和新鲜糕点，成功的关键是你使用了一种意大利制造的特殊烤箱，而面临的进阶难题是店面无法满足更多的市场需求。

你要发展业务就会面临两种选择：盘下隔壁的店铺并购买第二台烤箱，或者在城镇的其他地方再开一家新店，当然这也需要购买一台新烤箱。此时，你在市区另一个地方的竞争对手刚刚扩展了店面并获得巨大成功，并且你也读到了一家烘焙公司上市的故事，该公司通过谨慎地扩张现有店铺而成长。传统观点认为，扩张店面是正确的选择，但是你觉得需要坐下来做一做算术题。

首先，基于自己认为的保守假设，你计算了每个场景的前期成本以及可能的营收、利润。其次，从不同的场景展开，去计算每个选项的回报。

你已经设定了个人对回报率的最低要求：只有当项目产生至少20%的回报时才会继续前进。你也做了如下计算：一个新烤箱需要花费5万美元，扩张现有店铺需要建造费5万美元，在扣除劳动力、材料和其他运营成本后，可能会产2万美元的年利润增量。因此，你有10万美元的前期成本（烤箱加上扩建）和2万美元的预期年利润，预期回报率为20%——正好达到你对回报率的最低要求。

然后将注意力转向开设新店的选项，前期费用包括一个5万美元的烤箱和15万美元的新建费用。新店的销售情况很难预测（比如它在城镇的另一端），但你估计潜在的年利润为5万~7.5万美元。因此，你计算出前期成本为20万美元，预期回报率为25%~37.5%。这个回报率，即使是考虑最低的水平，也明显高于前一种选择。但在决定走哪条路之前，你要问自己一些重要的、定性的问题：

- 新店位于城镇的另一端，你对其销售情况预测落空的风险更大，那么对自己的预估，你究竟有几成把握？
- 更高的回报率是否足以弥补这一额外的不确定性？
- 新店所需资金是扩张店面所需资金的两倍，你能为开设新店筹集到额外的10万美元吗？（如果能，资金的成本是多少？）
- 反之，开设新店是否具有潜在益处？比如，它是否有可能让你的经营实现多元化，即使在现有门店销售额下降时，你的利益也可以拥有某种保障？
- 开设第二家店是否能赋予你商业洞察力，随着时间的推移，凭借这种洞察力你有可能建立一家更大的公司吗？

无论企业规模大小，以上这些资本配置的问题和决策就是管理者和企业家每天都必须面对的（尽管大公司通常会聘请顾问和投行来帮助他们找出答案），而且无论面包店主还是《财富》500强CEO，以条理和分析为导向的思维过程对于做出有效决策都至关重要。

对于本地的小商铺或是大型公司的董事会来说，局外人的做法似乎并没有那么复杂，但是，为什么没有更多的人去效仿他们的方法呢？答案是，实际运用比看起来更难。与同行的做法背道而驰，忽略“惯性驱使”都不是什么容易的事情，在许多方面，商界就像一个被青春期同龄人压力所笼罩的高中食堂。特别是在危机时期，人们自然的、本能的反应是，同伴做什么自己就做什么，行为学者称其为社会认同。当今世界充斥着社交媒体、即时讯息和嘈杂的有线电视节目，信息的噪声越来越难以消除，我们需要后撤一步采用丹尼尔·卡尼曼教授所提到“系统2”来思考。“系统2”是一个近来广受关注的思维工具，在思考时确实能派上用场。

局外人的清单

事实证明，检查清单是航空、医疗和建筑等领域极其有效的决策工具，它们简单的形式下蕴含着强大的力量，多亏了阿图·葛文德（Atul Gawande）的新作《清单革命》（*The Checklist Manifesto*）才让它们的应用成为当前的热门话题。检查清单是特别有效的“选择架构”形式，旨在促进分析和理性思维，并消除干扰，后者常常让复杂决策蒙上阴影。

清单为激活“系统2”提供了体系化的路径，对CEO们来说，它堪比高效的疫苗，让接种者对世俗观点和“惯性驱使”有所免疫。加万德建议清单的选项最好保留10项或更少，于是我们从局外人CEO们的经验中整理出一份清单，以帮助人们做出有效的资源配置决策，希望以此来避免损毁价值的决策。那么，请开始吧：

1. 配置过程应当由CEO主导，而不能委派给财务或业务开发人员。
2. 首先确定收益率目标，即投资项目可接受的最低回报率，这是CEO最重要的决策之一。注意：收益率目标应参考公司可获得的一系列投资机会来确定，并且通常应超过股权和债务融资的混合成本，通常在十几个点或更高。
3. 计算所有内部和外部投资方案的回报，并根据回报和风险对其进行排序（计算不需要非常精确）；采用较为保守的假设。注意：诸如收购等高风险项目应当要求高回报。要非常小心“战略”这个名词，它通常是低回报公司的代码。
4. 计算股票回购的回报率，收购的回报率应当显著超过这一基准。注意：虽然股票回购是局外人CEO们创造价值的重要来源，但它并不是万灵药，如果回购价格过高，也会破坏价值。
5. 关注税后利润，进行所有交易前都要咨询税务顾问。
6. 确定可承受的、保守的现金和债务水平，并使公司运营保持在该水平之内。
7. 考虑采用去中心化的组织模型，例如：公司总部员工占员工总数的比例是多少？这与你的同行相比如何？
8. 只有在有信心长期获得高于收益率目标的回报时，你才能把资本保留在当前业务中。
9. 如果没有潜在的高回报投资项目，请考虑支付股息。然而请注意，支付股息的决定一旦做出则很难撤回，并且支付股息从税务的角度来说是不划算的。
10. 当市场价格非常高时，考虑出售企业或股票都是可行的方案；

如果业务部门表现不佳，不再产生可接受的回报，那么关闭它们也是可行的。

无论是回顾过去还是展望未来，商界局外人CEO们对资源配置提供了行之有效的方法，凭借于此，他们得以在难以预测、千头万绪的商界中航行，并在五花八门的行业和市场条件下都创造了卓越的成果。这份清单是一个工具，可以帮助包括邻家的面包店和跨国集团在内的任何企业，让我们采用这种行之有效的方法吧，让我们张开双臂、睁大好奇的双眼，让我们去拥抱、去见证商业世界固有的不确定性吧.....

致谢

2003年，我在和家人一起度假时，萌发了写这本书的念头。当时我在一家私募股权公司工作，正在为公司每半年召开一次的CEO论坛准备发言稿。发言的主题是介绍一位杰出的CEO，把他职业生涯中的经验教训分享给我们投资组合中的公司经理人。我选择了亨利·辛格尔顿。在研究的起始阶段，从阿利姆·乔杜里（Aleem Choudhry）开始，我很幸运地结识了哈佛商学院一群非常有才华的二年级研究生。乔杜里是全美优等生联谊会成员，也是斯坦福大学物理学专业的毕业生和校网球队队员，他在该项目研究的过程中表现非常出色，我们的合作为随后的全部工作创建了研究模板。项目结束时，乔杜里向我介绍了另一位杰出的学生约翰·吉利根（John Gilligan），当时吉利根刚刚完成哈佛商学院第一年的课程。吉利根也是优等生联谊会成员，本科毕业于哈佛大学的化学专业，他同意对大都会通信公司做一个类似的独立研究，于是我们一拍即合。

与这些学生合作，我详细分析了每一位局外人CEO和他们的同行。每个项目用时整整一学年来完成，并遵循相同的调研流程和时间表：第一学期致力于详细分析局外人公司及其同行的历史财务报告。我们总共查看了1 000多份公司年度财务数据，对包括文章、书籍和视频等其他数据来源也做了深入回顾；第二学期的重点是对公司前任高管、投资人、华尔街分析师、银行家、同行公司高管进行广泛的访谈，当然还包括CEO本人。我们采访了每一位健在的局外人CEO。这些访谈通常会牵出失散已久的同事和投资人，由此需要进一步深入，最后，总访谈人数超过100人。

所以现在，我要感谢很多人，首先要感谢天才般的哈佛商学院学生所做的数据整理和收集工作。在过去的8年里，我十分荣幸与他们一道完成了这个项目。这一天才群体包括保罗·布塞尔（Paul Buser，负责罗尔斯顿·普瑞纳公司）、乔杜里（负责特利丹公司）、埃琳·艾森伯格（Erin Eisenberg，负责伯克希尔-哈撒韦公司）、马特·埃斯特普（Matt Estep，负责通用动力公司）、约翰·吉利根（负责大都会通信公司）、

布赖恩·赫斯曼和莫里茨·乔布克（Brian Hersman, Moritz Jobke, 负责TCI）、克里斯蒂娜·米勒（Christina Miller, 负责《华盛顿邮报》），以及康斯坦蒂诺斯·帕帕康斯坦蒂努（Konstantinos Papakonstantinou, 负责大众影院公司）。

我同样要感谢来自哈佛商学院的纳比勒·哈格（Nabil El-Hage）和迈克·罗伯茨（Mike Roberts）两位教授的支持，以及一位在读学生马特·克利特斯（Matt Klitus）。在成稿的最后阶段，克利斯特为分析提供了重要帮助。

我还要感谢那些愿意花时间阅读手稿并给出灼见的人，他们包括比尔·凯里（Bill Carey）、阿特·查彭蒂埃（Art Charpentier）、威尔·加德纳（Will Gardner）、贾尔斯·古德海德（Giles Goodhead, 他阅读得特别仔细）、艾尔夫·格罗斯贝克（Irv Grousbeck）、鲍勃·格鲁斯基（Bob Grusky）、迈克·杰克逊（Mike Jackson）、克里斯·金博尔（Chris Kimball）、萨姆·麦考斯兰（Sam MacAusland）、乔·尼豪斯（Joe Niehaus）、布兰登·尼克松（Brandon Nixon）、戴维·西蒙斯（David Simmons）、凯文·塔维尔（Kevin Taweel）、兰尼·桑代克（Lanny Thorndike）、汤姆·特里福罗斯（Tom Tryforos）、埃利奥特·沃兹沃思（Eliot Wadsworth）、拉姆齐·沃克（Ramsey Walker）和史蒂夫·沃尔曼（Steve Wallman）。特别感谢戴维·沃戈，他为TCI和大都会通信公司这两章提供了额外的支持。

我还要感谢我的父亲尼克·桑代克（Nick Thorndike）的仔细阅读和意见；感谢我的孩子夏洛特和尼古拉斯对这个项目付出的耐心，感谢他们让我在研究路上的关键时刻能偷闲片刻，也感谢他们偶尔超出年龄的智慧。从早期的研究助理伊丽莎白和艾默里·桑代克（Amory Thorndike）那里，我也获益匪浅，一并致谢。

我要特别感谢巴菲特那无可替代的合作伙伴查理·芒格先生，在目前期他就给出鼓励和深刻的洞见，特别是对特利丹公司、大都会通信公司、《华盛顿邮报》和通运动力公司的章节。

我还要特别感谢丹尼斯·埃亨（Denise Ahern），她是镇定自若、无所不能的助手，在过去近9年里给予了我非凡的帮助、支持以及耐心。

还要感谢夏洛特·麦克唐纳（Charlotte MacDonald）多年来的大力支持。

我的经纪人戴维·米勒（David Miller），对整个项目而言，他一直是洞察力和新观点的来源。

我的编辑杰夫·基欧（Jeff Kehoe）和艾利森·彼得（Allison Peter），以及哈佛商业评论出版社的整个团队，他们做得非常出色，他们的呵护和指导让这部手稿及其毫无经验的作者顺利走完了出版流程。

最后，我要感谢吉妮，她是我不可或缺的“贴身”编辑和妻子，感谢她在各个阶段对手稿提出了极其宝贵的充满理性和启发的解读，感谢她对我多年来始终如一的支持。

诚然，本书中如有任何错误，责任都在我，但以上这群人的智慧和支持对本书的最后成型做出了不可估量的贡献。谢谢你们，谢谢你们。

参考文献

前言 谁是比较杰克·韦尔奇更伟大的**CEO**

1. Berkshire Hathaway annual reports, 1987.

2. Ibid.

3. Warren E. Buffett, “The Superinvestors of Graham and Doddville,” *Hermes Magazine*, April 1984.

引言 像做投资一样做管理

1. Atul Gawande, “The Bell Curve,” *The New Yorker*, December 6, 2004. See also Richard Pascale, Jerry Sternin, and Monique Sternin, *The Power of Positive Deviance: How Unlikely Innovators Solve the World’s Toughest Problems* (Boston: Harvard Business Press, 2010).

2. Michael Lewis, *Moneyball: The Art of Winning an Unfair Game* (New York: W.W. Norton and Company, 2003).

3. Malcolm Gladwell, *Outliers: The Story of Success* (New York: Little, Brown and Co., 2008)

4. Robert J. Flaherty, “The Singular Henry Singleton,” *Forbes*, July 9, 1979.

特质**1** 效率决定生死，汤姆·墨菲和首府广播公司

1. Charles T. Munger memo, January 1, 1983.

2. Unless otherwise noted, all quotations from Tom Murphy come from a telephone interview on March 23, 2005, and an in-person interview on July

25, 2005.

3. “Tom Murphy’s Pleasant Cash Problem,” *Forbes*, October 1, 1976.
4. Author interview with Dan Burke, April 1, 2005.
5. Author interview with Gordon Crawford, April 20, 2005.
6. Author interview with Bob Zelnick, March 23, 2005.
7. Author interview with Dan Burke, April 1, 2005.
8. Author interview with Phil Beuth, April 28, 2005.
9. Author interview with Dan Burke, April 1, 2005.
10. “Tom Murphy’s Pleasant Cash Problem.”
11. Author interview with David Wargo, April 28, 2005.
12. Author interview with Gordon Crawford, April 20, 2005.
13. Author interview with Phil Meek, April 1, 2005.

特质2 聚焦资本配置，亨利·辛格尔顿和特利丹公司

1. Author interview with Jack Hamilton, April 20, 2004.
2. Robert J. Flaherty, “The Sphinx Speaks,” *Forbes*, February 20, 1978.
3. Author interview with Charles T. Munger, September 10, 2004.
4. Author interview with Arthur Rock, April 15, 2004.
5. Author interview with Charles T. Munger, September 10, 2004.
6. Ibid.
7. Author interview with Fayez Sarofim, March 2, 2004.

8. Author interview with William Rutledge, February 23, 2004.
9. James P. Roscow, "The Many Lives of Teledyne," *Financial World*, November 1, 1978.
10. Robert J. Flaherty, "The Singular Henry Singleton," *Forbes*, July 9, 1979.
11. Flaherty, "The Sphinx Speaks."

12. Author interview with Leon Cooperman, November 20, 2003.

特质3 绑定股东利益， 比尔·安德斯和通用动力公司

1. Author interview with William Anders, April 5, 2008.
2. Unless otherwise noted, all quotations from William Anders come from telephone interviews on April 15 and April 24, 2008.
3. Author interview with Ray Lewis, March 20, 2008.
4. Author interview with Peter Aseritis, March 7, 2008.
5. Author interview with James Mellor, March 12, 2008.
6. Author interview with Peter Aseritis, March 7, 2008.
7. Author interview with James Mellor, March 12, 2008.
8. Author interview with Nicholas Chabraja, April 2, 2008.
9. Author interview with Ray Lewis, March 20, 2008.
10. Author interview with Nicholas Chabraja, April 2, 2008.
11. Ibid.
12. Author interview with Ray Lewis, March 20, 2008.

13. Ibid.

特质4 高效运用杠杆，约翰·马隆和TCI

1. Unless otherwise noted, all quotations from John Malone came from an inperson interview on April 30, 2007.

2. Author interview with J.C. Sparkman, April 30, 2007.

3. David Wargo TCI analyst report, 1981.

4. Author interview with Dennis Leibowitz, April 17, 2007.

5. Author interview with Rick Reiss, April 26, 2007.

6. Author interview with David Wargo, April 17, 2007.

7. Author interview with Rick Reiss, April 26, 2007.

8. David Wargo TCI analyst report, 1980.

9. David Wargo interview/analyst report.

10. Author interview with Dennis Leibowitz, April 17, 2007.

11. David Wargo TCI interview/analyst report, 1982.

12. Author interview with Dennis Leibowitz, April 17, 2007.

13. David Wargo TCI analyst reports, 1981, 1982.

14. Ibid.

15. Author interview with David Wargo, April 17, 2007.

特质5 勇于逆向而行，凯瑟琳·格雷厄姆和《华盛顿邮报》

1. Author interview with Alan Spoon, April 2, 2009.

2. Author interview with Tom Might, April 30, 2009.

3. Unless otherwise noted, all quotations from Donald Graham came from an in-person interview in his office on April 3, 2009.

4. Author interview with George Gillespie, April 8, 2009.

5. Author interview with Ross Glotzbach, March 30, 2009.

6. Author interview with George Gillespie, April 8, 2009.

7. Author interview with Alan Spoon, April 2, 2009.

8. Author interview with Dick Simmons, April 22, 2009.

9. Author interview with Ben Bradlee in his office, April 3, 2009.

特质6 数据思维制胜，比尔·斯蒂利茨和罗尔斯顿·普瑞纳公司

1. Author interview with Michael Mauboussin, February 25, 2009.

2. Ibid.

3. Author interview with Pat Mulcahy, April 23, 2009.

4. Unless otherwise noted, all quotations from William Stiritz came from numerous telephone interviews in April, May, and June, 2009.

5. Author interview with Pat Mulcahy, April 23, 2009.

6. Ibid.

7. Author interview with Michael Mauboussin, February 24, 2009.

8. Author interview with John McMillin, April 2, 2009.

9. Author interview with John Bierbusse, February 24, 2009.

特质7 永远关注现金流，迪克·史密斯和大众影院公司

1. Author interview with Bob Beck, February 26, 2008.

2. Author interview with Woody Ives, December 15, 2007.
3. Author interview with Caesar Sweitzer, April 22, 2008.
4. Ibid.
5. Author interview with Woody Ives, December 15, 2007.
6. Author interview with Bob Beck, March 26, 2008.
7. Author interview with David Wargo, April 22, 2008.
8. Unless otherwise noted, all quotations from Dick Smith are from an inperson interview on April 23, 2009, at his office in Chestnut Hill, MA.
9. Author interview with Woody Ives, December 15, 2007.
10. Ibid.
11. Author interview with Woody Ives, December 15, 2007.

特质8 首席执行官就是首席投资官，沃伦·巴菲特和伯克希尔-哈撒韦公司

1. Author interview with Charles T. Munger, February 24, 2006.
2. Berkshire Hathaway annual reports, 1977–2011.
3. Ibid.
4. Unless otherwise noted, all quotations came from interviews with Warren E. Buffett, July 24, 2006.
5. Author interview with Charles T. Munger, February 24, 2006.
6. Berkshire Hathaway annual reports, 1977–2011.
7. “The World’s Top Investing Stars,” *Money Week*, July 6, 2006.

8. Berkshire Hathaway annual reports, 1977–2011.
9. Author interview with David Sokol, April 15, 2006.
10. Berkshire Hathaway annual reports, Owner’s Manual.
11. Author interview with Tom Murphy, March 9, 2006.
12. Author interview with Charles T. Munger, February 24, 2006.
13. Berkshire Hathaway annual reports, 1977–2011.
14. Berkshire Hathaway annual reports, 1977–2011.

结语 伟大的**CEO**都善于打好手中的牌

1. Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011).
2. Author interview with George Roberts, April 8, 2004.
3. Author interview with Charles T. Munger, September 10, 2004.
4. Author interview with Pat Mulcahy, April 23, 2009.
5. Andrew Barry, “What a Gusher,” *Barron’s*, November 16, 2009.
6. Michiyo Nakamoto and David Wighton, “Citigroup Chief Stays Bullish on Buy-outs,” *Financial Times*, July 9, 2007.

后记 一个案例和一份清单

1. Atul Gawande, *The Checklist Manifesto: How to Get Things Right* (New York: Metropolitan Books, 2010).

附录 巴菲特的测试

巴菲特曾针对配置资本的能力提出过一个简单的测试方法：一位CEO在任期内，每1美元的留存收益是否至少创造了1美元的价值？仅用一个数字，巴菲特的衡量标准就综合反映了CEO整个职业生涯中决策制定的智慧和愚蠢。可悲的是，要通过测试比听上去的更为艰难。当然我们也毫不意外，局外人CEO都以优异的成绩通过了测试，如表10-1所示。

表10-1 局外人CEO的测试成绩

所属公司	目标时间段		巴菲特测试段（年）		累积“巴菲特回报”（倍）
伯克希尔－哈撒韦公司	1965-6-30	2010-9-30	1965	2007	2.3
哈考特－布雷斯－朱万诺维奇公司 / 大众影院公司	1966-1-31	2001-7-31	1962	2000	3.6
特利丹公司	1963-5-31	1990-6-30	1966	1989	2.0
大都会通信公司 / 美国广播公司	1966-9-30	1995-12-31	1966	1994	2.7
《华盛顿邮报》	1971-6-30	1993-12-31	1971	1993	1.9
TCI	1973-5-31	1999-3-31	1973	1997	不适用
罗尔斯顿·普瑞纳公司	1980-1-31	2001-12-31	1981	2000	2.4
通用动力公司	1990-12-31	2007-12-31	1990	2007	3.5

注：①大众影院公司1991年收购哈考特通用公司导致数据突变，因此进行了调整，否则会扭曲结果。

② 马隆管理TCI期间，有意识地让账面盈余最小化，因此，这个指标对他不适用。

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

而真正的阅读，应该在书中与智者同行思考，借他们的视角看到世界的多元性，提出比答案更重要的好问题，在不确定的时代中领先起跑。

湛庐阅读 App：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本

怎么读是读者面临的最大阅读障碍

“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”App。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、电子书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者为伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。



下载湛庐阅读 App，
坚持亲自阅读，
有声书、电子书、阅读服务，
一站获得。

CHEERS

本书阅读资料包

给你便捷、高效、全面的阅读体验

本书参考资料

湛庐独家策划

- ✓ 参考文献
为了环保、节约纸张，部分图书的参考文献以电子版方式提供
- ✓ 主题书单
编辑精心推荐的延伸阅读书单，助你开启主题式阅读
- ✓ 图片资料
提供部分图片的高清彩色原版大图，方便保存和分享

相关阅读服务

终身学习者必备

- ✓ 电子书
便捷、高效，方便检索，易于携带，随时更新
- ✓ 有声书
保护视力，随时随地，有温度、有情感地听本书
- ✓ 精读班
2~4周，最懂这本书的人带你读完、读懂、读透这本好书
- ✓ 课程
课程权威专家给你开书单，带你快速浏览一个领域的知识概貌
- ✓ 讲书
30分钟，大咖给你讲本书，让你挑书不费劲



湛庐编辑为你独家呈现
助你更好获得书里和书外的思想和智慧，请扫码查收！

(阅读资料包的内容因书而异，最终以湛庐阅读App页面为准)

(1)本书初作于2012年，书中所有涉及时间推算的内容都以2012年为基准。——编者注

(2)ERA (Earned Run Average) 是棒球比赛中的技术统计术语，可译为“投手自责分率”或“防御率”。——编者注

(3)1盎司≈28.35克。——编者注

(4)英文原文为decentralize，汉语中也可以用“权力下放”或“分权”来对应，但是为了和原文尽量保持一致，书中采用了“去中心化”这个译法。——译者注

(5)斯芬克斯，指狮身人面像，意思是说辛格尔顿像谜一样令人费解。——译者注

(6)计算公式是 $(1.1^{10}-1) / 1.1^{10} \times 100\% = 61.5\%$ 。——译者注

(7)葛文德是曾推动奥巴马医改的白宫顾问，他的著作《清单革命》将其在医疗领域的革命风潮推广到建筑、金融、行政等与我们生活息息相关的领域。这本书的中文简体字版已由湛庐引进，北京联合出版公司于2017年10月出版。——编者注

(8)英文原文为institutional imperative，也可译为“制度性强制力”。——译者注

(9)原书数据由于四舍五入的原因无法对应，翻译时根据“1971年”和“同

比增长”两栏对“1961年”的数据稍加修订。——译者注

(10)甲公司欲以每股1元（市价0.5元）的价格向市场回购1 000股，这就是回购要约。拥有甲公司股票的人，觉得回购价很划算，就会向甲公司发出承诺，愿意按照1元的价格出售手中股票。——译者注

(11)1英亩≈4 047平方米。——编者注

(12)伯克希尔-哈撒韦公司在1967年进行过一次分红，每股10美分。——译者注

(13)这里的意思是罢工者十分讨厌凯瑟琳，只希望菲利普当初朝她开枪而不是自己。——译者注

(14)作者2009年和布拉德利对话时，凯瑟琳已经去世。——译者注

(15)博帕尔灾难事件是指1984年12月3日发生在印度博帕尔市特别重大的一次工业事故，这次事故导致数千人死亡。——译者注

(16)原文中分母为现金流，但是图表中分母又变成了EBITDA，这两者在现实中并不完全等同。对罗尔斯顿来说，这两个数字很接近，所以作者混同使用。——译者注

(17)伍德斯托克（Woodstock）是位于美国纽约州北部的一个城镇，因伍德斯托克音乐节而著名。——编者注

(18)丹尼尔·卡尼曼的新作《噪声》探讨了影响人们做出判断的敌人之一“噪声”，并告诉人们应该怎样重塑自己的决策框架。这本书的中文简

体字版已由湛庐引进，浙江教育出版社于2021年9月出版。——编者注

CHEERS
湛庐

媲美西奥迪尼《影响力》
锤子科技创始人 罗永浩
鼎力推荐



无价

[美] 威廉·庞德斯通 (William Poundstone) 著 阎佳 译

浙江教育出版社
ZHEJIANG EDUCATION PUBLISHING HOUSE

版权信息

书名：无价

作者：〔美〕威廉·庞德斯通（William Poundstone）

译者：阎佳

出版社：浙江教育出版社·湛庐

出版时间：2022年10月

ISBN：9787572243950

字数：280千字

版权所有·侵权必究

中文版序 价格蕴含着人类心灵的奥秘

来想想上海的房价之谜吧！从2000年到2004年，上海房地产的价格将近翻了一番。尽管房价的变化速度着实叫人眼花缭乱，但房地产开发商们仍然费了不少心血，以便给房子定一个正确的价格。对新建的高层楼房来说，开发商要“照顾”到买家的偏好：低楼层潮湿且容易滋生蚊子，窗外看不到好风景，大众觉得它不够有面子；13这个数字不吉利，14听起来像是“去死”，所以，这两层不受买家的欢迎；此外，购房者们一般也不想要最顶层，因为到了夏天实在太热了。

在确定价格时，开发商要顾及所有这些因素。房子太贵，卖不出去；太便宜，开发商赔钱。定价的标准是，“坏”楼层卖得跟“好”楼层一样快。可惜，开发商们在这一点上失败透顶：“坏”楼层的房子要卖出去耗费的时间总是更久，定价再便宜也不管用。

2008年，奚恺元、让-皮埃尔·杜贝（Jean-Pierre Dubé）和张岩在《营销研究》杂志上撰文论述了这一矛盾。它是价格心理学上的绝佳例子。价格心理学是一个超有趣的研究领域。这个领域有一条戒律：人们并非总有明确的偏好。还有另一条戒律：人们提出选项的方式方法，会影响他们做出的决定。正如阿莫斯·特沃斯基（Amos Tversky）所言：“与其说我们挑选的是选项本身，倒不如说我们挑选的是选项的说法。”

举个例子，人们现在了解到，做出选择的潜在认知机制有别于估计价格或其他定量数字。这是一点重要甚至烦人的发现，因为大部分传统经济学都假设：做出选择（买还是不买）和确定公平价格是配套的，前者相当于你的手，后者相当于你的手套——瞧，伸出手戴上手套，就是这么合适。确定公平价格的能力是市场经济的基础。

可有时候，我们设定的价格却跟我们做出的选择完全是两码事。奚恺元的研究小组向潜在的购房者提出以下问题：“你是愿意花50万元买一楼的房子，还是愿意用56万元买八楼的房子？”绝大多数购房者表示，他们倾向于购买更贵的八楼。这没什么奇怪的，因为购房者说，他

们认为楼层比价格更重要。人在做选择的时候，考虑的因素越是重要，这种因素所得的权重就越多。故此，那些购房者强烈地偏爱“更好”的楼层，只有很大的价格差异才能驱使他们选择楼层不佳的房子。

另一组购房者回答的问题稍有不同。研究小组告诉他们，一楼房子的卖价是50万元，请他们给八楼房子安排一个公道的价格。所有购房者定的价格，全都低于56万元。指定价格和做出选择有所区别，两者其实分属于大脑不同的部位，是两种思维过程。尽管人们努力想给较好的楼层分配一个公平的“贴水”，但他们力有未逮。结果，较好楼层的价格出现了系统性低估——在实验室研究里是这样，在研究小组采访的地产销售部里也是这样。好楼层总是最先卖光，于是晚来的买家可选范围变小了，开发商少赚了钱。

多年来，经济学家一直不愿承认决策心理学在商业世界的重要地位。而奚恺元等人的实验极有影响地论证了价格心理学的相关作用。蓬勃发展的经济和新兴的市场使中国成了此类研究的重地。《无价》的读者遍及全球，着实让我欣喜。我希望中国读者也能看出它的启示意义。正如我们所发现的：价格不只是数字，在很多方面，它蕴含着人类心灵的奥秘。

威廉·庞德斯通
William Poundstone

测一测 你了解价格的魔力吗？

- 是谁把经济冲突比喻为博弈？（ ）
 - A. 冯·诺伊曼
 - B. 卡尔·马克思
 - C. 亚当·斯密
 - D. 丹尼尔·卡尼曼
- **SKP**公司的座右铭是：价格是一道危险的杠杆。这是指价格波动能大幅度影响：（ ）
 - A. 员工流动率
 - B. 用户满意度
 - C. 股票价格
 - D. 利润率
- 很多超市把主要入口设在店面右侧，这更可能是因为：（ ）
 - A. 右侧风水好
 - B. 顾客沿逆时针方向逛商店会让店家赚得更多
 - C. 巧合而已
 - D. 大部分顾客习惯靠右行走

扫描下面二维码查看本书更多测试题

扫码鉴别正版图书
获取您的专属福利



扫码获取全部测试题及答案，
一起了解价格的本质

引言 一杯近290万美元的咖啡

“你要的越多，得到的就越多”

1994年，新墨西哥州阿尔伯克基（Albuquerque）地区陪审团判定，麦当劳应赔偿斯特拉·利柏克（Stella Liebeck）近290万美元，原因是利柏克自己打翻了麦当劳卖给她的一杯热咖啡。尽管热咖啡造成利柏克三度烫伤，但美国公众可不怎么同情她。深夜模仿秀和电台主持人把她变成了人肉靶子。谈话类节目的客座嘉宾们把这场官司当成重点案例，展示“我们的法律制度出了什么问题”。《宋飞正传》（*Seinfeld*）有一集的剧情就是克莱默(1)起诉被打翻的咖啡，一家网站设立了“斯特拉奖”——专门给司法制度下最疯狂的判决颁发雷人奖。

利柏克受伤并不是闹着玩的。她的孙子开车载她来到麦当劳的“得来速”窗口(2)。他们买了咖啡，孙子把车开到路边停下，方便利柏克老太太往里面加奶加糖。老太太把咖啡杯夹在两膝之间，掀开盖子。就在这时，咖啡洒了。利柏克花了11 000美元的医疗费，用于腹股沟、臀部和大腿的皮肤移植。棘手的问题是，你该怎样给利柏克的痛苦和麦当劳的过失定价呢？起初，利柏克要求麦当劳赔偿20 000美元。麦当劳拒绝了，并且提出赔偿800美元。

利柏克的律师，在新奥尔良出生的里德·摩根（Reed Morgan）擅长打这种官司。1986年，他代表休斯敦的一位妇女向麦当劳提起了诉讼。那位妇女也是因为咖啡洒了，三度烫伤。摩根用他迷人的美国南部腔男中音，提出了一个高明的法律理论：麦当劳的咖啡存在“缺陷”，因为它太烫。麦当劳负责质量控制的人说，咖啡应当介于180~190华氏度(3)，而这个温度，比其他连锁店的咖啡要烫些。休斯敦一案以27 500美元的赔偿额达成和解。

摩根跟踪监控了随后的咖啡官司。他知道，1990年，加利福尼亚州的一位妇女因为麦当劳的咖啡造成三度烫伤，以麦当劳赔偿23万美元达成和解，只不过这件事没有大张旗鼓地宣传。但这里有一个很大的区别：在加利福尼亚州的案子里，是麦当劳的员工把咖啡洒到妇女身上

的。

由于利柏克是自己把咖啡弄洒的，按理说这件案子的索赔金额应当远远少于23万美元。但摩根根本不管这一套，他对陪审团使用了一种颇具争议性的心理战术。等会儿我再详细介绍它，此刻先暂时用一连串的美元符号“\$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$”来指代它。

这种技术奏效了。陪审团就好像被催眠了一般，判给利柏克近290万美元的赔偿费。这包括16万美元的补偿性赔偿，外加270万美元的惩罚性赔偿。为做出这一判决，陪审团用了4个小时。据报道，一些陪审员报出的赔偿额高达960万美元，是其他人好说歹说才把他们劝下来的。

虽说赔偿额减少了，但麦当劳上诉是免不了的。81岁的利柏克可不会返老还童。她很快就跟麦当劳达成了庭外和解，和解费保密，据说低于60万美元。她肯定意识到自己已经打出了一记漂亮的“本垒打”，这种事恐怕是可一而不可再。

说完麦当劳的咖啡，让我们再来看看那些换包装的秘密。“四季宝”（Skippy）花生酱重新设计了包装用的塑料罐。“以前用的罐子，底部是平的，”西蒙顾和管理咨询公司（Simon-Kucher & Partners，下文简称SKP公司）的定价顾问弗兰克·卢比（Frank Luby）解释说，“现在的则向内凹进去，这样可以少装几盎司(4)花生酱。”老包装一罐能装18盎司，而新包装只能装16.3盎司。原因嘛，再明显不过：方便四季宝减少分量而价格不变。

花生醬罐子底部的凹陷設計，跟一種定價理論有着很大關係，在心理學文獻里，它被叫作“任意連貫性”（coherent arbitrariness）。

你不可不知的价格术语

任意连贯性

该理论说，消费者其实并不知道什么东西该值多少钱。他们茫然地穿过超市货架，根据种种线索判断价格。

任意连贯性首先是一种相对理论。买家的主要敏感点是

■ 相对差异，而非绝对价格。

四季宝换了新包装，相当于花生酱涨价10%。可要是厂家直接涨价

10%（比如涨到3.39美元），消费者肯定会注意到，有些说不定还会转而购买其他品牌。根据这一理论，消费者会很乐意支付给四季宝3.39美元，只要他们不知道涨价了。四季宝花生酱的消费者往往是些小孩子，他们隔一段时间就会买一罐，并且记得上次是什么价。对于这样的产品，顾问们建议采用“隐性”缩小包装的创新方法。2008年夏天，家乐氏（Kellogg's）旗下生产的可可亚脆米花、香果圈、爆米花、苹果酒和蜂蜜麦片逐步采用细长型纸包装盒——几乎没有人注意到。消费者只看到盒子摆在货架上的高度和宽度，等他们伸手取下盒子时，其实已经做出购买决定，转头开始想别的事儿了。

价格实验

大雅（Dial）和激爽（Zest）肥皂把模具改小了，这样肥皂的重量少了差不多0.5盎司（约14克），而外包装盒子丝毫没变。北棉（Quilted Northern）把自家生产的超柔卫生纸缩窄了0.5英寸（约1.27厘米）。泡芙（Puffs）纸巾的制造商把纸巾长度从8.6英寸（约21.84厘米）缩短到8.4英寸（约21.34厘米）。由于包装盒大小不变（约24.13厘米宽），所以盒子里多出来的空间足有1英寸多（约2.54厘米）。你看不出来，是因为纸巾盒的开口在中间。不管怎么说，消费者根本不会注意到这些缩水，除非他们有旧款泡芙纸巾，来仔细量一量。

这套把戏也就能耍到这个地步。再怎么缩水，麦片盒子也不能扁得像信封，罐子也不能做成实心塑料疙瘩。总有一天，制造商们要采取一项人人都能注意到的大胆行动。它会引入一种全新的经济型包装。不管是在大小、形状，还是在其他设计方面，新包装（和新价格）都难以跟以前的比较。消费者困惑不已，无从判断换了新包装的产品是否划算。于是，他们干脆两眼一闭，把产品扔进了手推车。包装缩水的戏码就这么永无休止地循环上演。

如果你觉得这套把戏很蠢，那你不是第一个这样认为的人。好多人嘟囔说，他们宁可付高价，也要搞清楚所买产品的重量。还有些人赌咒发誓说，他们会观察商场的比较标签，精确计算每盎司的价格，断不会受此愚弄。但价格顾问们知道一件事：消费者说的和做的并非一回事。对大多数人来说，价格记忆是短时记忆，外包装记忆则更是如此。

很多公司曾经只能用经济学入门教科书上的需求曲线来制定产品定

价策略。上一代，波士顿咨询、罗兰贝格等事务所，靠着为企业提供更复杂、更复杂的定价心理技巧服务而赚了个盆满钵满。但抢先专攻定价咨询的，还要数前文提到的SKP公司。正是SKP公司把定价策略彻底发展成了一门高深的学问。

关于SKP公司对日常生活所需种种商品的定价的惊人影响，我们意识到的很少。适用于其他类型咨询工作的规则，放到价格上就不成立了。一家广告公司不可能同时代理可口可乐和百事可乐——SKP公司却可以。在不少行业，SKP公司招揽了一多半的顶尖公司。目前，在它的客户名单里，是这样一些如雷贯耳的名字：宝洁、雀巢、微软、英特尔、德州仪器、沃达丰、诺基亚、霍尼韦尔、蒂森克虏伯、华纳音乐、贝塔斯曼、默克、拜耳、强生、瑞银、巴克莱银行、汇丰银行、高盛、道琼斯、希尔顿、英国航空、汉莎航空、阿联酋航空、宝马、奔驰、大众、丰田、通用、沃尔沃、卡特彼勒、阿迪达斯和多伦多蓝鸟队。不管你是给短信、卫生纸，还是给机票设定价格，同样的心理技巧都适用。对SKP公司的顾问们来说，价格具有最普遍的无形说服力。

虽说价格只是一个数字，但它能唤起人们复杂的情感——这种东西，现在靠着大脑扫描，咱们都能看见了。在不同背景下，消费者对同样的价格会产生不同的感知：既可能觉得是捡了天大的便宜，也可能觉得是被狠狠敲了一记竹杠；还可能完全没上心。有些技巧是普遍适用的，比如包装缩水，价格以神奇数字“9”结尾。但价格咨询，可不仅仅是在扁平世界里用来叫卖的新招数。它利用了心理学近些年来最重要、最新的一些科研成果。设定价格的行为稀松平常，可在这一过程中，商家把内心的欲望变成了人人都看得见的数字语言。事实证明，这一过程，狡猾得惊人。

1984年，阿莫斯·特沃斯基获麦克阿瑟奖时开玩笑说，自己的工作不过是确定了“广告人和二手车推销员早就知道”的东西。这可不光是自谦的客套话。当时，钻进钱眼儿里的从业者们对特沃斯基的接受程度，可比大多数经济学家和CEO要高多了。营销人员早就在做价格心理学方面的试验了。在邮购的鼎盛时期，厂家为了测试定价策略的效果，经常印制不同版本的产品目录或传单。你还以为价格是固定不变的？这些研究成果一定可以打消你所有的幻想。营销人员和销售人员清楚地知道客户愿意付多少钱，并且随时都可以变；凭借这一事实，钞票滚滚而来。经济学家唐纳德·考克斯（Donald Cox）甚至说，行为经济学基本上是“营销专家们的老把戏，他们早就把‘经济人’概念从焦点小组里踢出去

了”。

到了今天，研究价格的心理学家与营销顾问及价格顾问群体产生了互相依赖的关系。不少顶尖的理论家，包括特沃斯基、丹尼尔·卡尼曼（Daniel Kahneman）[\(5\)](#)、理查德·泰勒（Richard Thaler）、丹·艾瑞利（Dan Ariely），都在营销期刊上发表过重要研究成果。SKP公司有来自三个大洲的学者当学术顾问。当今的营销工作者大声地赞美锚定和任意连贯性，以及它们令人不安的力量。“许多像我这样从事营销教育的人都会在开课时说：‘我们讲的不是要操纵消费者，我们是要发现并满足他们的需求。’”哥伦比亚大学的埃里克·约翰逊（Eric Johnson）说，“接着，等你入行工作一阵子，你会意识到，嘿，我们分明能够操纵消费者嘛！”



第一部分

价格背后的心理奥秘

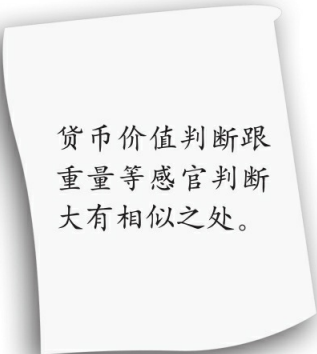
- 为什么百老汇剧场里价格越高的位置卖得越火？
- 为什么 100 万美元带来的愉悦感，却要用 400 万美元才能翻倍？
- 为什么有人说得出张国荣的老唱片该比蔡依林的贵 2 倍，却说不出它到底该卖 10 块还是 1 毛？

第1章 价格没头脑

假设有人让你拎起一个手提箱，猜猜它多重。你能猜得多准确呢？大多数人都承认自己猜不了多准。手臂肌肉、大脑和眼睛不是用来掂量物体具体有多重（精确到千克）的。这就是为什么超市的公平秤和体重秤前总是挤满了人——大家用秤一称，个个目瞪口呆。

现在再假设这个行李箱无人认领，被摆上了拍卖台。它的锁被撬开了，里面装着一些休闲衣物、一台高级相机，还有一些没怎么用过的物品。这一回，你的任务是猜多少钱能中标——也就是手提箱及其中物品的市场价值。你认为你能猜得多准确？会比你猜手提箱重量更准吗？

你可能会说，拍卖这事儿有时料不准。好吧，问题再简单些。假设你是拍卖场上的一个投标人，你所要做的就是判断你的最高出价。不是猜别人会怎么做，而只需说出这个手提箱对你来说值多少钱，以美元和美分来表示就行了。你的估价会精确到什么程度呢？为没有明确市场价值的东西标上一个价格并不难，但在本例中却要稍微难一些。你恐怕总会怀疑自己的最高出价比另外两个投标人高出一大截。



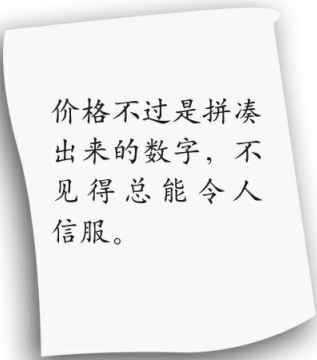
货币价值判断跟
重量等感官判断
大有相似之处。

价格心理学上的一个常见主题是，货币价值判断跟质量（或者亮度、音量、温度、气味的强度）等感官判断大有相似之处。研究感官知觉的学科，叫作心理物理学。早在19世纪，心理物理学家就发现，人们对相对差异非常敏感，对绝对数值则不怎么敏感。给你两个外观上一模

一样的手提箱，一个重15千克，一个重16千克，你轻轻一提，就很容易说出谁重谁轻。可要是没有公平秤，你恐怕很难判断手提箱是否超过了航空公司20千克的质量限制。

人们对价格也是一样找不着头绪。可人们在很大程度上对这一重要的事实没有正确的认识。这是因为，人们生活在一个媒体虚报价格和市场价值的氛围下。人们由于记得什么东西的成本“应该”是多少，所以接受了假象：我们的价值感觉没问题。视力不好的人照样能在家里穿梭自如，因为他记得家具摆放的位置。消费者就是这样。这叫补偿作用，不叫“视力敏锐”。

我们时不时地能得到一些线索，看出自己的价格感觉有多不靠谱。摆过旧货摊的人都知道，要给家里的废旧物品标上一个有意义的价格会是多么困难。“这张张国荣的老唱片该比那张蔡依林的贵2倍，我百分百肯定。但它到底该卖10块还是1毛，这我可就说不准了。”



价格不过是拼凑
出来的数字，不
见得总能令人
信服。

相对估值稳定而一致，具体的数额却无比随意。旧货摊的例子揭示了一个我们或许不愿在商业交易中承认的真相：价格不过是拼凑出来的数字，不见得总能令人信服。

简单的发现带来了深远的影响。推动我们周遭世界运作的数字并不像看起来那么牢不可破、有着符合逻辑的基础。在新的价格心理学当中，价值短暂而又不可靠，就好像哈哈镜里的样子一样变化多端。

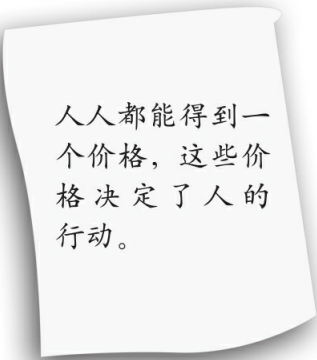
它挑战了人们在商业意识和常识中根深蒂固的信条：“人人都能被收买，只不过各有不同的价码。”特里·索泽恩（Terry Southern）1959年的小说《奇妙基督徒》（*The Magic Christian*）对这一民间智慧来了一次即兴演奏。

亿万富翁盖伊·格兰德（Guy Grand）爱玩恶作剧，他花一辈子想要证明，只要价格足够高，所有人就都能被收买。他的典型做法是，买下芝加哥的一栋办公大楼，把它拆了，立上一口热腾腾的大桶，里面装着从猪圈运来的屎、尿、血。而在这堆臭烘烘的大杂烩里，炖着一张张百元大钞，总数100万美元。桶上贴着告示：“此处美元可免费自取。”

照格兰德的信条，只要给上足够多的钞票，什么样的事人们都肯做。人们未必个个都是贪婪积财的物质主义者，但在当今社会，相信金钱是战无不胜、超越一切的人不在少数。

“人人都有价码”理论认为，人们对自己的估价是稳定的，只要稍稍耍点手腕，你就能发现它。倘若有人报上一个价格（要你把自己卖给魔鬼什么的），你会把它跟心里的底价做个比较，以决定是否接受。毫不夸张地说，所有传统经济学理论都建立在格兰德的简单前提之上：人人都能得到一个价格，这些价格决定了人的行动。

现在，我们掌握了压倒性的证据，足以证明这个想法是错的。至少，把它拿来当成人类真实行为的模型绝不靠谱。早在20世纪60年代末，心理学家萨拉·利切坦斯泰因（Sarah Lichtenstein）和保罗·斯洛维克（Paul Slovic）便证明了价格具有极度的模糊性。在他们的实验中，被试无法为想要的东西或所做的选择设定一致的价格。自那以后，心理学家们研究出了这种现象的前因后果。按照新的观点，要在内心“构建”价格，人得从周围的环境中获得“线索”。特沃斯基和卡尼曼的主要研究领域是心理学下一个名为“行为决策理论”的年轻分支。它的研究内容，是人们如何做决策。乍一看，这一主题似乎颇有意义，但稍显沉闷。事实上，人类的悲喜瞬间尽在其中。生活的一切，无非就是做决定。他们所做的“联合国”实验，示范了它是如何运作的。



人人都能得到一个价格，这些价格决定了人的行动。

价格实验

他们使用了一件道具——类似嘉年华会上的幸运大转盘，边上标着从1到100的数字。一组大学生看着转盘转动，并随机选中一个数字。你可以自己试着玩一玩——想象转盘旋转，转出的数字是65。现在，请回答以下两个问题：

（a）联合国中非洲国家所占的比例，是高于还是低于65%（就是刚刚出现在转盘上的那个数）？

（b）非洲国家在联合国中所占的比例是多少？

请把你的答案写在此处，或者稍微想一下，给出一个具体的数字。

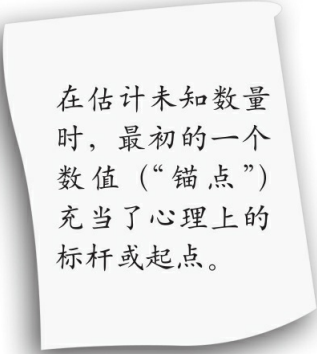
和许多实验一样，这个大转盘也是动了手脚的。转盘只能转出两个数字，10或65。做这个手脚，只是为了简化结果分析。不管怎么说，特沃斯基和卡尼曼发现，这一所谓的随机数字，影响着被试对第二个问题的回答，而且影响极大。

如果转盘停在10，学生们对非洲国家在联合国中所占比例的估计，平均值是25%。可要是转盘停在65，平均估计值就成了45%。后一估计值几乎是前者的两倍。在这个实验中，唯一的变量就是估计比例的人看到了一个不同的“随机”数字，而且他们知道这个“随机”数字是毫无意义的。

你兴许会说，美国人的地理知识向来糟糕。美国大学生们不知道正确答案，只好靠猜，凭空幻想一个数字。你大概觉得，他们是因为对答案拿捏不准，所以才只好把刚刚凑巧出现的数字生拉硬扯上去，然而实际情况并非如此。被试并不是简单重复他们刚才碰到的数字（10或65）。他们给出了自己的数字，但在这一过程中，他们受了先前数字的巨大影响。

特沃斯基和卡尼曼把这称为“锚定与调整”。1974年，他们在《科学》杂志上发表了一篇经典论文——《不确定状态下的判断：启发式和偏差》（*Judgment Under Uncertainty: Heuristics and Biases*）。他们建立的理论认为：在估计未知数量时，最初的一个数值（“锚点”）充当了心理上的标杆或起点。在上面的实验中，幸运转盘上的数字就是锚点。第一个问题要被试将锚点与待估计的数量进行比较。特沃斯基认为，在回答第二个问题时，被试在心理上将锚点上下调整，以此得出自己的答

案。这一调整往往并不充分。结果，答案离锚点太近。如果你只看最后的结果，那就好像是锚点施加了一种磁铁般的吸引力，把估计值拼命往自己附近拉。



在估计未知数量时，最初的一个数值（“锚点”）充当了心理上的标杆或起点。

请容我顺便问一句，你刚才猜的答案是多少呢？跟“65”对照组的平均估计值45%相差多少？此外，事实上，非洲国家在联合国中所占的比例，接近28%。

第2章 黑就是白

一位女士怀着7个月的身孕，穿着高跟鞋摇摇晃晃地走着。她看到报纸上刊登的广告称，埃斯基尔森医生的诊所可以免费检查视力，觉得这挺划算。埃斯基尔森医生的诊所设在俄勒冈州尤金市（Eugene）中心法院的对面。门上一板一眼地挂着招牌：俄勒冈研究所视觉研究中心。那时是20世纪60年代，诊所里候诊室的样子跟其他小镇上的验光师诊所差不多：什么东西都不太贵，可样样整洁崭新。家具贴面是菲律宾红木镶板，地毯是墨绿色的。几幅海报为房间增添了几分色彩，其中一幅是旅游海报，上有“神奇的哥本哈根”字样——兴许埃斯基尔森医生是丹麦人？接待员接待了患者，引她上了三级台阶，进了验光室。

埃斯基尔森医生一脸严肃，看不出具体多大年纪。他下巴中间有点儿向内凹，秃顶前肯定是个帅哥。他戴着眼镜，让人觉得好像心情不大好——仿佛这份工作在跟他唱反调似的。

“请您走过来，站在这个有标志的地方，”他温和地说，“我会在墙上投射一些三角形，请您目测一下它们的高度。”

患者照做了，很快就进入了视力检查的状态。过了几分钟，埃斯基尔森医生注意到患者的动作有几分异样。

“您怎么了？”他问。“怪了，”患者道，“我怎么有点儿晕啊？”“大概是怀孕的缘故？”医生不怎么确定地问。“我从来没这样过，”患者坚持道，“我感觉连站都站不住了。”这位怀孕的女士穿着高跟鞋走了几步，用手扶着墙：“你是在催眠我吗？这可太卑鄙了！”埃斯基尔森医生冲着墙上的对讲机说了话：“好啦，吉姆，咱们的被试发火了。”

埃斯基尔森医生的合作者保罗·霍夫曼（Paul Hoffman）曾在南太平洋空军当过领航员。告别军队后，他攻读了实验心理学博士，后来在俄勒冈大学当小助教，这期间他发现自己不怎么喜欢教书，反倒冒出了一个想法：建立智库，专门研究人类的决策。1960年，他终于碰到了机会。他凭着美国国家科学基金会发的60 000美元补贴，以及房子抵押

款，在11号码头买下一栋房子，改名为“俄勒冈研究所”。霍夫曼认为，有些研究，不搞大学那一套繁文缛节最好。

1965年，纽约一栋办公大楼的设计师给霍夫曼出了一道难题。那栋建筑顶楼的租客们支付的租金是最高的。建筑师和工程师们担心，曼哈顿刮大风时顶层会晃悠——他们可不愿让宝贵的租户们觉得大楼摇摇欲坠。为避免此种情况，他们需要知道多大幅度的水平晃动会引起人们的注意。当时似乎完全没有这方面的数据。

霍夫曼意识到，他们需要做一个心理物理学实验。所谓“人刚好能注意到的区别”，就是一种刺激的最小可感知量（在上述案例中，就是房间的晃动幅度）。就如何测量可察区别，自19世纪以来已经有了大量的心理物理学文献。修些活动的小隔间本来是很容易的，不过霍夫曼知道，要是他一开始就告诉人们实验的目的，被试会满心盼着房间晃动，而这种期盼会让他们提前感觉到晃动。“于是我想，”霍夫曼回忆说，“该怎么邀请一个人来办公室，出于这样或那样的目的，请他待在某个房间里，然后开始让房间摇晃呢？”

霍夫曼在尤金市珍珠大街800号的办公大楼租了个房间，改装成验光师诊所。验光室地板下有轮子。通过一套静音液压装置（本来是锯木厂用来拖原木的），房间会逐渐加速前后摇摆，晃动幅度也会递增。晃动幅度介于1英寸（约2.54厘米）到12英尺（约3.66米）之间。心理学家埃斯基尔森凑巧是个注册验光师，他答应扮演验光师这个角色。他们总共做了72例假验光测试。验光的时候，他们慢慢加快房间的摇摆速度，直到被试“发火”，即说出一些表示自己对晃动有所察觉的话。埃斯基尔森和霍夫曼关心的数据是，房间的晃动幅度要多大，患者才能察觉。他们详细记录被试的生理特征（如怀孕、穿高跟鞋等），以及被试察觉时说了什么话：

我觉得站不稳。我觉得自己好像在船上。当初在宾夕法尼亚州的时候，我们做过醉酒驾驶测试，要求被试沿着一条直线走……

真不舒服。你是不是在给我照X光一类的东西啊？我好像被偷拍了……

我觉得你偷走了我的地心引力……

埃斯基尔森自己也没能免疫。每天他都头晕，回到家调养好，第二天早晨回来，又晕了。

实验结果表明，可察觉晃动临界值比建筑工程师们预计的要小10倍。虽说这不是客户想要听到的消息，但霍夫曼的研究方法引起了他们

的兴趣。建筑师山崎实和工程师莱斯利·罗伯逊（Leslie Robertson）来到俄勒冈，坚持要试一下实验装置——然后，他们信了。

之前签的保密协议使得霍夫曼无法出版甚至无法公开自己的发现。建筑开发商可不愿跟负面宣传一类的消息沾边，但俄勒冈测试的确让工程师们采用了刚性更强的外部支撑部件。1970年，该建筑以“世界贸易中心”（以下简称“世贸中心”）之名，大张旗鼓地开张了。2001年，两架飞机在恐怖分子的劫持下撞向了世贸中心的双子塔。由于采纳了霍夫曼的建议，塔楼维持了足够长的挺立时间，使14 000多人得以安全逃生。

今天，俄勒冈研究所备受推崇，因为它是行为决策理论诞生的摇篮。长久以来，它是利切坦斯泰因和斯洛维克的工作场所。在这里，两人首次清晰地揭示了人们对价格和基于价格的决策是多么地摸不着头脑。还有一年，俄勒冈研究所成为当时最具影响力的两位心理学家——特沃斯基和卡尼曼的根据地。那可真是硕果累累的一年啊！

在谈到这个杰出群体之前，有必要先谈谈他们的前辈，以及心理物理学这一奇特的科学。“心理物理学”这个词是19世纪中期德国心理学家古斯塔夫·西奥多·费希纳（Gustav Theodor Fechner, 1801—1887）普及开来的。靠着心理物理学，费希纳遇上了哲学史上的一个古老问题：主观体验能够比较或沟通吗？一般而言，颜色体验就是个很方便的例子。

人们是以同样的方式体验颜色的吗？有没有这样的可能：同为红色的“停止”标识，一个人看到的是红色，另一个人看到的却是绿色？有没有什么办法能把它说出来呢？看到绿色的人仍然会把标识叫作红色，因为别人一直教他“停止”标识的颜色就叫作“红色”。

本着纯粹的哲学精神，类似这样的问题是无法回答的。可它为下面的问题打开了口子：感觉的强度能否测量呢？19世纪的德国心理学家威廉·冯特（Wilhelm Wundt）对此表示怀疑：

一种感觉比另一种强多少或弱多少，我们永远说不出来。不管太阳是比月亮亮一百倍还是一千倍，大炮的声音比手枪响一百倍还是一千倍，要进行估算，都远非我们的能力所能及。

我们要搞懂冯特说的是什麼。他并不是说，物理学家无法测量日光和月光的客观亮度。冯特的那个时代，物理学家们已经开始做这件事了。他也不是说，你问太阳和月亮哪一个更亮，人们会回答不出来。事

实上，人们会百分百同意，太阳要亮得多。

冯特只是在说，主观上的比率是没有意义的。在这一点上，他犯了一个弥天大错。在接下来的一百多年里，跟冯特同时代的人及他的接班人们（大多顶着“物理学家”的头衔），收集了令人信服的证据，证明了人们很擅长做冯特心目中“不可完成的任务”。

“心理物理学”最切实的定义是“对物理量（声音、光、热和质量）和主观感受之间联系的研究”。就算只在德国莱比锡市，费希纳也不是第一个探讨它的人。早在1834年初，莱比锡的生理学家厄恩斯特·韦伯（Ernst Weber）就确定了迄今为止该领域最了不起的一大成果。

价格实验

韦伯让被试蒙上眼睛，要他们判断不同的砝码组合有多重。韦伯谨慎地增加小砝码，直到被试说他感觉自己负担的物体质量明显沉了些（这就是“可察差异”）。他认为，质量的相对变化（百分比）比绝对克数的变化更为重要。壮汉举起的杠铃上飞来一只苍蝇，当事人感觉不到质量的明显变化。可要是人蒙着眼睛，手掌上放了一枚硬币，同一只苍蝇飞到硬币上，他恐怕能察觉出不同。

在电灯泡和扬声器时代到来之前，心理物理学是甚为原始的一档事儿。早期的一位研究者朱利叶斯·默克尔（Julius Merkel）曾做实验要人们判断金属球落到乌木块上的声音大小。若是想让声音大些，默克尔只能从更高的地方把球抛下。另一位先驱，比利时物理学家约瑟夫-安东尼·费迪南德·普拉托（Joseph-Antoine Ferdinand Plateau）要求8名画家精确地画出位于黑与白正中间的灰色。为了不让“黑”与“白”的含义出现混淆，普拉托提供了色样。画家们取了色样，回到画室调灰色。尽管每个画室的光线有所不同，普拉托记录说，8人得出的灰色基本上相同。有人拿它当证据，想借此说明感觉也并非那么主观。他还做了一件跟费希纳的“苦命实验”极为相似的事情，直盯太阳看25秒，致使视力永久性受损。他逝世于比利时根特市（Ghent），死时双目失明，他的身后之所，距离画家凡·艾克（Van Eyck）名作《根特祭坛画》中所画的那座教堂只有咫尺之遥。

20世纪，心理物理学的发展主要得益于更好的视听设备。装配了最新的幻灯机、变阻器和音频振荡器之后，心理物理学的整个领域蓬勃发展。它的范围不仅涉及感官世界，还涵盖了道德、美学和经济价值判

断。研究人员让大学生们观察倾斜的线条、色彩或现代画作的复制品；用鼻子闻毒油，用耳朵听白噪声；比较暴行、薪水和香味。接着发问就开始了：跟水平线相比，斜线的倾斜度是多少？请按1到7级给你刚才听到的声音响度分级。哪一种犯罪程度更严重？你觉得照片里的这个孩子智力怎样？

S. S.史蒂文斯（S. S. Stevens）之所以出名，是因为他确立了物理强度与主观感觉的关联曲线。长久以来，人们就知道那不会是一条直线。

价格实验

设想有一个完全黑暗的房间。打开一盏60瓦的灯。然后打开第二盏60瓦的灯。此刻的光线会比只开一盏灯时亮1倍吗？不会（几乎人人都这么说）。它看起来会更亮，但不会亮上1倍。严谨的实验揭示，要想在主观上令光线看起来亮上1倍，光源点的强度在物理上必须是先前的4倍。

这其实是一条典型的幂曲线。不用数学公式来说的话，这里有一个把握大意的办法：你正用圣诞彩灯装饰房子，想要房子的灯光亮过你邻居的。具体而言，你想要自己的灯看起来比他家亮1倍。据史蒂文斯说，你得买上瓦数差不多是邻居家灯泡瓦数4倍多的灯泡。

不管你的邻居是个环保主义者，只点一盏彩灯，还是他喜欢出风头，把自己的房子弄得五彩缤纷，这条规则都成立。要想让主观效果翻倍，灯泡瓦数就得增加到4倍（你家当月的电费账单一定很可观）。

史蒂文斯满意地注意到，他的幂曲线可以简单地概括为：相同的刺激比率可以得出相同的主观比率。这通常被称为史蒂文斯定律（幂定律），或心理物理学定律。史蒂文斯和同一时期的研究者用了整整一代人的时间确定：幂定律具有普遍性，不光适用于光线的亮度，也适用于对温度、滋味、气味、震动和电击的感觉。

两个比率之间的因数，根据刺激类型的不同而有所不同。换言之，不一定总是“4倍的刺激带来2倍的反应”。比方说，在液体软性饮料中，要感受到2倍的甜度，糖量只需是原先的1.7倍。比率还取决于刺激的呈现方式。比如，一小片金属片接触皮肤，或是热源照射小范围皮肤，抑或是像桑拿一样的热浪包裹全身，在这三种情况中，热的感觉遵循不同的幂曲线。但在给定的实验中，曲线是非常一致的。在1965年，史蒂文

斯的两位同事指出：“从实验事实的角度来看，幂定律是毫无疑问地建立起来了，它恐怕比心理学领域提出的其他任何理论都更为牢不可破。”

史蒂文斯想解释为什么感觉会遵循幂定律。他注意到，物理定律（比如 $E=mc^2$ ）大部分是幂定律。通过调整为物理定律的形式，感觉能更清楚地“告诉我们实际情况是怎么一回事”。在史蒂文斯逝世后出版的文稿中曾写到，心理物理学：

举例来说，感觉上需要保持恒定的，是差数，还是比例，抑或是比率？显然，是比例——比率。

价格实验

走向一栋房子，房子的相对比例保持不变：从任何距离看，三角形的屋顶始终都是三角形。不管是在明亮还是黯淡的光线下看，照片上显示的都是同一幅画面：虽然光线不同，可照片光亮和阴影部分的比例看起来大致相同……不管刺激水平出现多大幅度的变化，知觉比例的效用及其关系几乎不变。请想想看，如果对话只在单一的强度范围内才能为人所理解；或者随着距离拉远，物体的外观比例就发生变化；又或者一有乌云遮蔽光线，照片就辨识不清了，那我们熟悉的生活会发生怎样天翻地覆的变化啊！

这样说来，我们的感觉以比率为基础，简直再合情合理不过了！但这里却有一个致命的弱点。对比率、对比例如此敏感的代价是，对绝对数量相对迟钝。

心理学家史蒂文斯用他典型的文风指出了这一点：

本书中的字体看似是黑色，实则不然，它只是因为没有来自黑色区域的光线进入你的眼睛罢了。事实上，黑色释放了大量的光，要是我们把黑色周围的白纸全都扯掉，你会发现黑色本身似乎散发着明亮的光芒，就像黑夜里的霓虹灯一样。

感觉以比率为基础，这一性质带来了许多后果。比较微不足道的一点是，它影响了心理物理学实验的设计。人们发现，实验结果在很大程度上取决于反应量表。这就是“测验”——以前是纸质的，现在是在线形式的。有两种最流行的反应量表：等级和量值。这两者你应该都很熟悉。

等级量表用于消费者调查和互联网投票。你如何评价你家的洗碗机？请选择：

1. 很差
2. 一般
3. 良好
4. 非常好
5. 绝佳

等级量表会列出数量固定的可能回答，并以文字标注。它有一个最低分（即最差选项），和一个最高分（最优选项）。

另一种方法是量值量表。它要求你用数字给某种东西打分。最低分数是0，最高分是——这里没有最高分。为什么没有最高分呢？诸如响度或质量这类物理量是没有上限的，对其的主观感受也没有明显的限制。

有时，量值量表会提供一个比较标准——模数（modulus）。它可能会给你展示一个投影出来的光圈，告诉你它的亮度是100。接着，它要你估计其他圆圈的亮度。有前者一半亮的，就是50；两倍亮的，是200；当然了，完全看不见的，是0。

模数应该能有所帮助，它的作用类似于地图上的比例尺。但史蒂文斯的妻子杰拉尔丁·斯通（Geraldine Stone）却说，他原打算省掉估计模数这一环节。史蒂文斯发现，没有它，被试们反而能给出更前后一致的判断。此后，他更青睐的技术是指导被试给出一个数字（任何数字都可以），来跟亮度、甜度和不快程度相对应。

这一套实验听起来乱哄哄的。从某种意义上来说，它确实乱——不同的人会给同样的东西分配完全不同的数字。但这并非必然。在中世纪，各地商人使用不一样的度量衡。但倘若一头牛比另一头牛重2倍，那么不管到哪儿，这个事实都不会发生变化，哪怕不同地方称量出的克数略有不同。在史蒂文斯的实验中，被试的绝对判断并不一致，但其比率却很有意义。让被试自己创造精神标杆，衡量史蒂文斯给出的答案，

这样更为合理。

为什么模数没用呢？有了模数，被试害怕“犯错”。没有它，他们只好跟着最初的冲动走，而这通常更准确。“我喜欢这样，我可以放松下来，思考音调。”一名被试告诉史蒂文斯，“要是有了固定的标准，我会感到更多的限制，老想着对声响做加减乘除，但这很难；而没有标准的时候，声响的数据值看起来该在哪儿，我只需要把它放过去就行了。”

追溯20世纪30年代的心理物理学文献，人们有时会用“锚点”来指代模数或等级量表中的两个端点。文中说，要根据这些比较标准“锚定”判断。然而，“锚点”似乎就像玻璃窗上的气泡，会歪曲判断。

在学生们的记忆里，史蒂文斯可能并不是一个好老师，可他做过几次令人印象深刻的课堂示范。

价格实验

一次，他给学生们看了一个被白色包围的灰色纸盘子。漆黑的房间里，聚光灯照在灰色盘子上，灰色看起来是白的。接着，史蒂文斯又把灯光打到了盘子周边的白色上。在周围耀眼的白色包围下，先前还是“白”色的盘子变成了“黑”色。

许多错觉就建立在类似原理的基础之上。由麻省理工学院认知科学家爱德华·H.阿德尔森（Edward H. Adelson）提供的例子，在本质上同史蒂文斯的示范很类似。在图2-1中，方块A和方块B的灰度是一样的。这个错觉很有意思，连酒吧都经常用它打赌。要想证实它，不妨先去找点便利贴（大概需要6小块）。

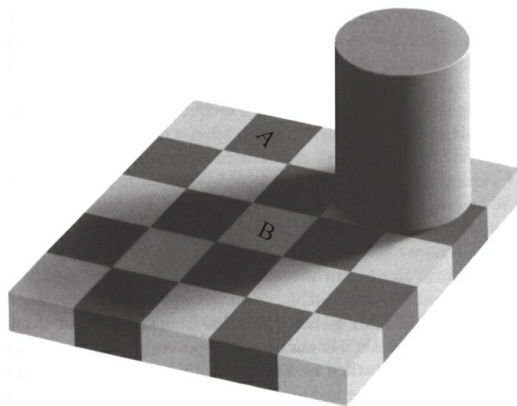


图2-1 这些方块颜色一样吗？

价格实验

把跳棋盘其他的格子仔细地遮起来，只留下方块A和方块B两处可见。还没把最后一块便利贴贴上去之前，你恐怕会惊疑：方块A和方块B两处的颜色怎么会是相同的呢？可突然之间，它们就“变成”了同一种中灰色。

要理解这一错觉的原理并不难。圆柱体投下了一道影子，让“白色”的方块B变“黑”了（实际上就成了灰色）。从纸张印刷的角度来看，方块B的灰度值和“黑色”的方块A一样。但人的眼睛和大脑有比判断绝对灰度值更重要的事情要做。它们正试着理解世界（在本例中，就是图2-1）的意义。这就意味着大脑会专注于对比。我们看到了一张棋盘，所有的“白”方块都是同一颜色，还有一道边缘模糊的影子。光与影之间的对比干扰不了棋盘黑白方块之间的对比，反过来也是一样。

所以，史蒂文斯说过这么一句名言：“黑就是周围有一圈光环的白。”这句话的口气挺像作家奥威尔，但确确实实不是乱说。史蒂文斯再清楚不过了，稍稍耍些小花招，你就能让人相信任何有关自己感觉的事情。在主观上，没有什么绝对的东西，只有对比。

而心理物理学家赫尔森在暗房里的经历，也为我们提供了宝贵的理论。

价格实验

赫尔森正在摄影暗房里亮着的一盏红灯底下工作，突然间他注意到一件怪事——他的烟头闪着绿光。

当然，在普通光线下，燃着的烟草放出的光看上去应该是余烬红。这一经验帮赫尔森明确了一个重要的概念——一个有关适应水平的概念。很明显，当时赫尔森的眼睛适应了暗房里不同寻常的红色灯光。较之安全灯发出的红色光芒，燃烧的烟头发出的是一种调子更冷、更黄的红色光。两相比较，后者便显得像是绿色了。赫尔森的眼睛和大脑登记的不是绝对色彩（这是数码相机用的方法），而是香烟和房间基准颜色之间的色差。

赫尔森最终得出结论，所有的感觉都要适应一定的刺激水平，接着记录从基准开始的变化。他在一系列著名的质量实验中示范了这一观

点。

价格实验

他先让被试一个接着一个地举起两枚小砝码，然后再让他们描述第二枚砝码有多重。他发现，被试会因为第一枚砝码而产生偏差，因为第一枚砝码充当了比较的锚点或基准。倘若锚点砝码比第二枚砝码轻，它会让第二枚砝码感觉重些；要是锚点更重，它则会让第二枚砝码感觉更轻。这种感觉的相对性会带来彻底的矛盾。赫尔森可以做些安排，同样一枚砝码，放在轻砝码之后，被试感觉重；放在重砝码之后，被试感觉轻。

从概念上讲，这没什么好大惊小怪的。如果你想显得苗条点儿，就结交些胖朋友！我们都注意到了对比效应。你是否有过这样的经历：一边想着要喝一口咖啡，实际上喝的却是茶？在那短短的一瞬间，你喝的东西味道怪不可言。它尝起来既不像茶，也不像咖啡。你所喝的，是介于期待和现实之间的落差。

第3章 价格效用：谁在决定价格

每个人的生活中都有一套非常重要的量值量表，它叫作价格。大概在公元前3000年，美索不达米亚人就认识到，他们用的质量单位谢克尔也可以用来表示大麦的量——或者能用同等大麦数量交换的其他任何东西的价值。这就是金钱和价格的萌芽。

经济学用保留价格来探讨市场力量对支付价格有着什么样的影响。但对于这事儿，有一种很是不同的看待方式。你可以把保留价格想成一套量值量表。对买方来说，价格是对拥有某物渴望程度的量化指标。对卖方来说，价格衡量的是保住自己一方也已拥有的某物（包括时间、精力和自尊这类绝对重要的东西）的欲望程度。

你不可不知的价格术语

保留价格（又称最低价格、底价）

对经济学家来说，指的是买家愿意支付的最高数额，或卖家愿意接受的最低数额。交易价格便位于这两个极值之间。

按日常生活的常识来看，价格是单一维度的，就像尺子上的刻度。每一种商品在量表上都对应着单一的一点。这些点整整齐齐地把世界上所有东西的价格都排列了出来。然而，价格的心理现实并非如此简单。

史蒂文斯为哈佛大学实验室里的人上了好多免费的金钱心理学课。他在班上提出过这样一个谜题：

假设我告诉你说，我有了一个特别基金，我会给你10美元。这会令你高兴，对吧？现在你来仔细想一想：要让你感到双倍开心，我得给你多少钱呢？

哲学家大可对此表示反对，所谓“双倍开心”这样的说法没有意义。但史蒂文斯的学生们似乎能毫不费力地按出题者的本意来回答这个问题。不过，他们的回答不见得能叫哲学家大吃一惊，倒是会大出经济学

家们的意外：平均答案是40美元上下。

出乎意料地得到10美元是个挺不错的小惊喜。接下来的一两天，你偶尔会想起钱包里多出来的这笔钱，感觉良好。但从现在开始的一个星期之内，你会用掉这笔钱，然后忘了它。

这下子，你真能拍着胸口说，获得20美元会让你感觉好1倍吗？我刚才所说的10美元的结果，换成20美元也是一样。这样推理下来，要让人感到双倍开心，必须多于20美元，事实也正是如此：课堂上，学生们给出的平均回答介于35美元和50美元之间。

金钱的收益递减很难说是什么新消息。史蒂文斯发现，100万美元带来的愉悦感，400万美元才能让它翻倍。没有哪个经济学家会为此感到一丁点儿的吃惊。这是些足以改变人生的大数目。金钱能买到的绝大部分东西，100万美元都够了（至少在史蒂文斯生活的那个年代是这样的），没人会觉得第二个100万会像第一个那么有意义。

这就是所谓的财富效应。但它解释不了史蒂文斯的小实验。他的研究对象是哈佛大学的学生，他们中有不少来自富裕家庭，大部分人都有望得到终身的经济保障。从一辈子的角度来看，区区几十美元应该毫无意义才对。唯一有关系的事情应当是，这笔钱能买些什么。不管金钱跟开心的兑换率是多少，10美元能买到的量，20美元都能买上两倍。“正确”的答案应该是20美元。

为什么史蒂文斯的学生们不这么想呢？显然，他们想的不光是这笔钱可以买什么。金钱本身是一种能产生感觉的“刺激”——其运作原理和史蒂文斯研究的其他刺激差不多。

价格实验

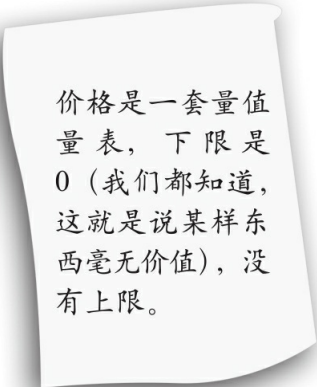
1959年，日本心理物理学家印藤太郎向127名大学生展示了手表的图片和简介。他要学生们评价每只手表的合意度，并给它指定一个公平的日元价格。根据实验结果，学生们认为，一只表要达到两倍的合意度，必须为它支付8.7倍高的价格。

拿现在的价格来说，天美时手表大概在40美元左右，而你可以选择150美元上下的斯沃琪手表，3 000美元的卡地亚手表，或者30 000美元的劳力士手表。它们全都是功能良好的手表，都起计时的

作用。唯一的区别是档次。戴卡地亚手表，表明你有钱，而且不怕露富。戴劳力士手表也一样，而且更张扬。劳力士手表大概比卡地亚手表更加金光闪闪，但没有哪个地方能贵出将近10倍的价格来。正如印藤太郎的学生们所指出的，价格大幅上涨，可能只买来社会地位的小幅提高。

还有研究发现了收入与社会地位之间存在的幂定律关系，以及盗窃钱财严重程度内含的幂定律关系。根据史蒂文斯引用的一项研究，要让你的社会地位翻番，你的收入必须是先前的2.6倍。盗窃的严重程度与所窃财物价值之间只存在微不足道的递增关系。要让罪行严重程度翻倍，盗窃金额需要增加到60倍。乍看起来，这似乎很奇怪。但大多数人都认为，凡偷东西都是错的，所窃金额反倒是次要问题。故此，根据盗窃幂曲线，偷6 000块钱只比偷100块钱情况严重2倍。

总体来讲，这一研究证实了史蒂文斯的观点：人对金钱的看法，跟其他感官知觉很像。价格是一套量值量表，下限是0（我们都知道，这就是说某样东西毫无价值），没有上限。不同的特征比（礼物、盗窃等）也是量值量表里很典型的东西。



价格是一套量值量表，下限是0（我们都知道，这就是说某样东西毫无价值），没有上限。

由于我们的文化对金钱追求得过于疯狂，比起很多东西来，在实际中我们对它并没那么敏感。许多感觉的增加速度比刺激本身的量提高的速度要快。要让人感觉重两倍，质量本身只需增加到1.6倍（所有举重选手都明白这是啥意思）；要带来双倍的冲击感，电流量只需增加到1.2倍（这就是为什么用它来折磨人很管用）。而要让金钱带来双倍快感，所需的金钱要更多才行。相对而言，根本就没有物美价廉这档事儿。

价格当然是一套独特的量值量表。我们当然很看重绝对价值——物品索取的绝对价格。然而，看重绝对值，并不能带来准确感知它们的能力。锚定的花招以及利用对比和暗示营造幻觉的手法，很容易动摇人们对货币价值的评估。很少有人预料到，这项研究揭示了一只看不见的手在操纵、误导全世界的金融决策。除了搞心理物理学的，几乎没人对此有所留意。

本尼·高福斯坦（Benny Goffstein）相当看重家庭。等他有机会自己开赌场时，他便为它取名“四女王”，讨4个女儿的欢心。和他经营的第一家赌场“里维埃拉”比起来，“四女王”坐落在城里，规模较小，利润也更为丰厚。

而“四女王”的一位投资者，跟高福斯坦在“里维埃拉”碰到的那些黑帮分子截然不同。此人名叫查尔斯·B. G.墨菲（Charles B. G. Murphy），是一位品位有些“独特”的马萨诸塞州贵族。墨菲过去是耶鲁大学的足球运动员，他是J.斯特林·洛克菲勒（J. Sterling Rockefeller）的朋友，他本人还是个非洲探险家、大冒险家、律师兼赌徒。他去世前的最后几年是在拉斯维加斯度过的。墨菲带着一个问题来找高福斯坦。为了避免税，他先前成立了一个慈善基金会。政府向墨菲施加压力，要他把基金会的一部分资金真正投入到慈善事业上，不然别想合理避税。墨菲决心把钱投到他真心喜爱的一项科研主题上：赌博。

墨菲找周围的人打听了一圈，想找个是赌博专家的科学工作者。他想到了密歇根大学的心理学家沃德·爱德华兹（Ward Edwards）。爱德华兹提出了一个非同寻常的请求。他和自己从前的几个学生，现在为一家名为俄勒冈研究所的机构工作，想在拉斯维加斯的赌场做一些实验。他们很喜欢在实际环境下对真人做实验。能在“四女王”做这事儿吗？身为主要支持者，墨菲身上有足够的痞气，他暗示得很清楚：这个要求，高福斯坦必须接受，没有拒绝的余地。

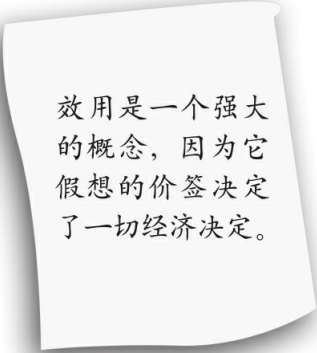
爱德华兹从事科研工作，最喜欢的便是提难题。他出生于新泽西的莫里斯敦（Morristown），父亲是个经济学家。他是在餐桌上听着父亲和同事们聊天长大的。这令他对经济学产生了一种叛逆的怀疑态度。在斯沃斯莫尔学院和哈佛大学求学期间，爱德华兹决心走心理学的道路。也正是在哈佛大学，他读了约翰·冯·诺伊曼（John von Neumann）和奥斯卡·莫根施特恩（Oskar Morgenstern）的著作，可他对所读到的一切并不怎么热衷。

在匈牙利出生的冯·诺伊曼是20世纪伟大的数学家。在普林斯顿大学经济学家莫根施特恩的大力劝说下，冯·诺伊曼把自己天才的头脑转向了经济问题。这才有了1944年《博弈论与经济行为》（*Theory of Games and Economic Behavior*）一书的问世。冯·诺伊曼把经济冲突比喻成“博弈”，一种类似打扑克的东西，认为它理应经得起数学分析。

经济博弈里的筹码是美元、英镑和日元。不过，这么说还不够确切。冯·诺伊曼和一般的经济学家一样，坚持要使用一种名叫“效用”的主观性货币。你可以把效用想成是一种个人“价签”，人们会把它贴在各

种事情和结果上。重要的是，人们会尝试积聚最大效用，但这并不一定意味着最多的金钱。谁死的时候获得的效用最多，谁就是大赢家。

“效用”一词可追溯至18世纪。瑞士数学家丹尼尔·伯努利（Daniel Bernoulli）指出，金钱的价值是相对的。过生日时得到一张百元大钞，在5岁的孩子看来是做梦也梦不到的一笔财富，可在45岁的百万富翁看来则毫无意义。为预测人们会怎么用钱，有必要根据这些不同的评价进行调整，就好像有时必须根据通胀调整美元币值一样。



效用是一个强大的概念，因为它假想的价签决定了一切经济决定。

经济学家们认同伯努利的想法，原因有二：一，它承认了一个显而易见的事实——心理状态（不光是简单的贪欲）决定着经济决策；二，有了效用概念，经济学家不用在心理学上花太大功夫。令经济学家们感兴趣的是，建立一门如数学般精密的科学。除了少数例外，他们可不想费心思衡量金钱的心理效用，只要假设它原则上能完成就行了。

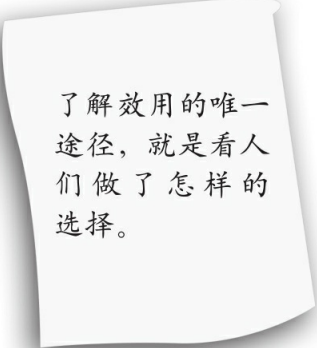
效用是一个强大的概念，因为它假想的价签决定了一切经济决定。麻省理工学院经济学家保罗·A.萨缪尔森（Paul A. Samuelson）把这一概念融入他的“显示性偏好”理论。该理论看起来极为合理，它指出：了解效用的唯一途径，就是看人们做了怎样的选择。选择揭示了我们对外用所知的一切，反过来，效用决定了消费者愿意支付的价格。

假设有人可以在A和B之间自由选择，他只需对照自己无形的价签，选中效用较高的那一个即可。这样一来，做决定就简化成了数字对比。这种假设很自然地成了大多数经济理论（从需求曲线到纳什均衡）的基础。

这就又把我们带回了冯·诺伊曼的贡献上。按照冯·诺伊曼的说法，每一结果的主观价格乘上它的出现概率，即可得出最终答案。许多经济

选择就是赌博。这个世界如此反复无常，棘手和有趣的选择必然会是这样或那样的赌博。因此，有必要为赌博分配价格。

在冯·诺伊曼提出博弈论后的第三年，赫伯特·西蒙（Herbert Simon）发表了重要作品《管理行为》（*Administrative Behavior*）。它展示了一幅全然不同的画面，来说明人们“博”的是什么“弈”。西蒙分析了企业和各层级机构如何做决定的案例。他提出了一个不朽的观点，人类是“有限理性的”。信息不足，他们太忙，偶尔也很愚蠢，没法像冯·诺伊曼说的那样把事情思考透彻。对冯·诺伊曼理论中所要求的结果，现实生活中的人似乎并没有表现出绝对的赏识。相反，决策者往往依靠灵感或者心理捷径，迅速做出符合直觉的选择。



了解效用的唯一
途径，就是看人
们做了怎样的
选择。

在决定付多少钱买车或跟谁结婚的时候，总会有些取舍权衡——按爱德华兹的同事库姆斯的话来说，也就是“比较没法比的东西”。赌博让你在以下两种情况之间权衡取舍：可以赢到的钱是多少，赢到的机会是多大。故此，库姆斯和爱德华兹会让被试选择赌博的方式，看他们是选奖金多的那种，还是选获胜概率大的那种。心理学家筛选了一定的偏好，试图洞悉人是如何做决定的。1960年，库姆斯和D. G.普鲁特（D. G. Pruitt）在研究中发现，大多数选择可以按一条简单的规则加以解释——“永远选择获胜回报最高的赌博方式”。

欢迎来到有限理性的世界。遵循这一规则的人，一定是忘了概率这档事——总是选择风险大的东西来赌。这一策略在赛马场上不怎么合适，在别处也并不见佳。爱德华兹最著名的一个实验，用上了两个背包。

价格实验

背包里装满了数目相等的扑克筹码。一个背包里装的主要是红色筹码——比如，70%的红筹码，30%的白筹码。另一个背包里主要是白筹码——比例刚好相反，70%的白筹码，30%的红筹码。你并不知道哪个背包里装的是哪样。你的任务是判断哪一个背包主要装的是红筹码。为了做出判断，你每次可以从一个背包里抽取一枚筹码。你必须根据自己抽取的次数判断概率。这就好比你是庄家，必须报出当前下注的盘口。爱德华兹让学生来做实验，自己则谨慎地记下所抽筹码的颜色。

假设你正从一号背包里抽筹码，抽出的第一枚是红色。问题来了——这个背包里主要装红色筹码的概率是多少？

正确答案比你想的要简单：不多不少正好是70%。但这个实验的本意并不是设计一道数学难题。大多数决定是靠直觉做出来的，爱德华兹是想看看这种直觉的本能到底有多准确。他发现，人们的猜测往往比正确数值要低。人们没有意识到，单独的一枚红色筹码也可以蕴含着有价值的信息，可事实偏偏恰好如此。

1952年，伦纳德·“吉米”·萨维奇（Leonard “Jimmie” Savage）碰到了这辈子最难熬的一顿午餐。萨维奇是个35岁的美国人，来巴黎参加学术会议。桌子对面，坐着一位一脸惊诧的男士。他是莫里斯·阿莱（Maurice Allais），40岁的法国经济学家。阿莱告诉萨维奇，自己有些东西要给他看。法国人希望美国人做个小测试。重要的一点在于，萨维奇竟然没通过这个测试。

萨维奇是个急性子的统计学家，有心为人如何做决策创建一套理论。他关心的决策大多和钱有关系。他很想知道，人们如何为商品及服务分配价格，如何在其间加以选择。萨维奇想表明有关金钱的决策是（或者说，有可能是）完全合乎逻辑的。美国著名经济学家米尔顿·弗里德曼（Milton Friedman）指望的就是这样一种理论，因为它能为他的自由市场经济乌托邦提供一个坚实的基础。

可有一个大问题，阿莱告诉萨维奇：他的理论错得离谱。[\(6\)](#)他轻而易举地证明了萨维奇的理论有错。阿莱就像童话故事里的巨魔怪一样，提出了三道谜题：

谜题一：你愿意选择以下哪种情况？

(a) 稳得100万美元。

(b) 赌一把：旋转幸运转盘。该转盘分为100个刻度。有89%的机会赢100万美元，有10%的机会赢250万美元，1%的机会什么也赢不到。

阿莱认为，大多数人会选择(a) 稳赢100万，因为选(b)的话，有可能什么也得不到——虽说可能性很小。显然，萨维奇同意这一看法。

谜题二：这次你的选择是：

(a) 11%的机会赢100万美元。

(b) 10%的机会赢250万美元。

阿莱认为，大多数人会选择(b)。两者的概率没有太大的差别。你当然会选奖金更高的(b)。萨维奇再次表示同意。此时，他落入了阿莱的陷阱。

这样，我们就来到了谜题三面前。你面前摆着一个密封的盒子。你会选择哪一个呢？

(a) 89%的机会赢取盒子里的东西，11%的机会赢得100万美元。

(b) 89%的机会赢取盒子里的东西，10%的机会赢250万美元，1%的机会什么也赢不到。

这一招准确地切中了萨维奇逻辑中的要害。诚如阿莱所知，萨维奇理性决策的一条公理指出，（基本上）决定选择汉堡加汽水还是比萨加汽水时，你可以把汽水忽略，因为两个选项中都有它。唯一起作用的一点是，你更喜欢汉堡还是比萨。按萨维奇的说法，一般而言，决策者应该忽略各个选项中的共同要素，根据差异进行选择。

几乎每个人都觉得这听起来很合理。阿莱却发现了一个微妙的漏洞。照萨维奇的逻辑，谜题三中的选择跟盒子里有什么应该是没关系的。无论是选(a)还是选(b)，你都有89%的机会赢得同一个盒子。

这并不意味着盒子里的内容无关紧要。盒子里可能装着10亿美元、一只致命的狼蛛，甚至是你在地铁上碰到的帅小伙的电话号码。但根据萨维奇的观点，盒子跟选(a)还是选(b)无关。人们的选择，只应当取决于是有11%的机会赢到100万美元更好，还是有10%的机会赢到250万美元更好。

换言之，谜题三的答案应当和谜题二的相同。这还没完。假设我们打开盒子，发现里面有100万美元。那么谜题三中的选择，最终跟谜题一相同。一句话，三道谜题的答案应当都一样，不管是(a)还是(b)，都该死磕到底，犯不着换来换去地折腾。阿莱捉弄得萨维奇违背了自己的原则。

几个月后，阿莱向弗里德曼提出了类似的小测试。弗里德曼没像萨维奇那样上当，他给出了一致的回答。

1953年，阿莱在法国出版的《计量经济学》（*Econometrica*）上发表了一篇文章，跟萨维奇和弗里德曼的理论公开叫板。萨维奇说，所有东西，每个人都能得到一个价格（效用）。这些主观价格左右了所有的决策。阿莱认为，人类比这要复杂得多。选择取决于背景，只是一个数字，无从表示人对不确定结果有着什么样的感受。自此以后，这一论证就被称作了“阿莱悖论”。

如果你还没搞清楚阿莱的用意以及它的重要性，别担心。让我给你举一个混合悖论的例子——这是哈佛大学理查德·济科豪瑟（Richard Zeckhauser）想出来的。

价格实验

你是一档热门游戏节目《金钱还是生活》的参加者。和大多数游戏节目一样，它不过是客厅老游戏的翻版罢了。可你运气不太好，碰上的老游戏是俄罗斯轮盘。

每次节目最开始，都会由“子弹夫人”蒂凡尼转动幸运转盘。转盘等分为6块。蒂凡尼将依照转出的结果给左轮手枪上子弹（从1颗到6颗），然后交给节目主持人布莱恩。短暂的广告时间过后，布莱恩转了转手枪的转轮，并用它直指着你的左太阳穴。他扣动扳机前，提出了一套你无疑会很感兴趣的财务安排。

你可以买1颗子弹。你可以和布莱恩商量一个价格，他随机从左轮里取出1颗子弹，你则递给他一叠钱。届时，他会再次转动左轮，指着你的太阳穴扣动扳机。

这里有件怪事儿。如果转轮里只有1颗子弹，你说不定会愿意为它支付更高的价格。买下那颗子弹，你就能100%地活下来（要不然，你就有1/6的机会熬不到下一轮广告时间）。你会为它出上一大笔，对不对？

为了便于比较，假设现在手枪转轮里有4颗子弹。你愿意付多少钱买一颗子弹呢——让枪里的子弹剩下3颗而不是4颗？不知为什么，你好像觉得不愿为这颗子弹拼了老命地出钱了。你甚至会想，

不妨碰碰运气，就试试4颗子弹算了。

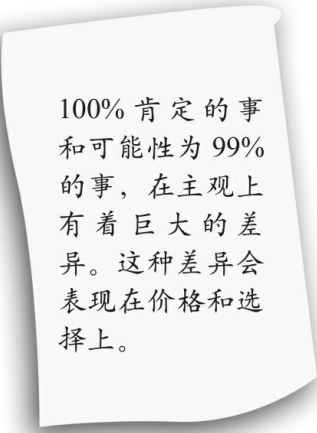
这是不是人意识上的一件趣事？就是1颗子弹，死了就是死了。前后两种情况下，你减少的死亡概率都是完全一样的。为什么你出的价格却不一样呢？

又或者假设枪里有6颗子弹。除非你出钱买1颗子弹，不然就死定了。这下子，说不定你又会反复折腾，并得出结论：这颗子弹是无价的，你愿意付出自己的一切去交换。

这个游戏和阿莱最初的谜题解释的都是同一种效应。100%肯定的事和可能性为99%的事，在主观上有着巨大的差异。这种差异会表现在价格和选择上。与此同时，10%和11%的概率差异就可以忽略不计了。

自此以后，阿莱悖论成了一群经济学家、心理学家和哲学家眼中的试金石。杰出的头脑们对它反复进行自我测试，但逃过其魔爪的寥寥无几。随后几年，阿莱本人也围绕谜题做了大量的思考，并写了很多东西。他试图真正按照经济学家的方式，阐明人类决策的公理。他指出，它们总有着微妙的不相容，并最终导致矛盾。

这里，我只想稍微谈谈为什么阿莱悖论如此棘手。从本质上看，这里的绊脚石并不是确定性效应，而是聪明人受了语言（即框定选项的方式）的影响。正如特沃斯基之后所写：“我们是在选项的描述之间进行选择，而不是单纯地在选项之间做选择。”在很大程度上，经济学家还没有做好准备接受这样的事实。



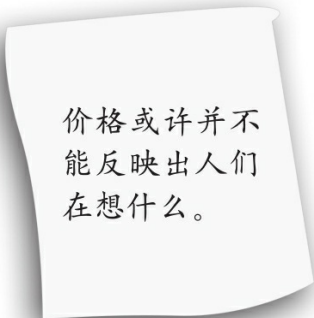
100% 肯定的事
和可能性为 99%
的事，在主观上
有着巨大的差
异。这种差异会
表现在价格和选
择上。

第4章 偏好逆转：为什么喜欢A却给它的定价低

假设你有1/8的机会能赢得77美元。你愿意出多少钱打这个赌？

最明显的办法是计算你每次出钱平均能赢多少。算出来是77美元的1/8，也就是9.63美元。当然，这个数字你心算起来挺麻烦。心理学家们关心的是直觉判断，他们观察到，被试分配给简单打赌的价格一般都太高。较之获胜的概率，人们似乎更注重奖金总额。

这可以解释为什么彩票会那么受人欢迎。举例来说，一张彩票能带来几千万分之一的机会赢上5 800万美元。基本上，买家买的只是幻想中了大奖的权利。“几千万分之一”这个数，只存在于纸面上和买家的脑海里。想招徕生意的时候，彩票主办机构会提高奖金，但这并不会增加中奖机会。



价格或许并不能反映出人们在想什么。

类似现象也适用于风险的规避。假如有1/12的机会损失63美元。你愿意出多少钱来避免这种情况的发生？人们愿意给的钱往往高于平均损失。做决定时，潜在的损失总额比损失概率要重要得多。

这暗示了人们为什么会购买保险。他们愿意为保险支付“高价”，因为较之风险渺小的发生概率，他们更担心灾难带来的损失。

价格或许并不能反映出人们在想什么。我们可以设计两种打赌方式——就叫A和B吧！大部分人会说他们偏爱A，可要是让他们给两者定

价，他们又会赋予B更高的价值。具体举例：

假设盒子A和盒子B是用漂亮包装纸包着的有趣礼品盒。盒子里到底有什么，我不敢肯定。我有机会摇一摇它们，大概估计一下里面装着什么。最后我决定愿意为盒子A付40美元，为盒子B付70美元。但同时我又判断，我宁可要盒子A。

这可真是疯透了！我的价格竟然跟我的愿望或行动不一致！利切坦斯泰因和斯洛维克还发现了更疯狂的一点：对某些类型的赌博，大多数人都会这样估价。他们把这称作“偏好逆转”，这里有个例子。如图4-1所示，两个圆环代表投掷飞镖的靶盘。你任选其一；之后，会有一名发牌员朝你选中的靶盘投出一只飞镖，飞镖有可能落在圆环中的任一位置。它将决定你能赢多少钱（要是你赢得了的话）。你会选择哪个靶盘呢？

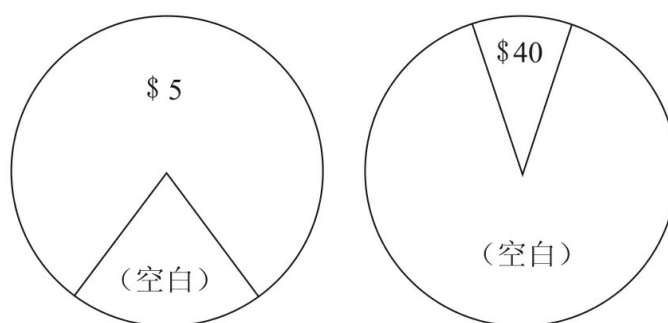


图4-1 偏好逆转示例

左边的靶盘有80%的机会赢5美元，20%什么也没有；右边的靶盘则有10%的机会赢40美元，90%什么也没有。

两个赌注的预期价值恰好一样，都是4美元，它不足以成为做出选择的理由。然而，绝大多数人都会选择左边的靶子。利切坦斯泰因和斯洛维克把类似左边的赌称作P(7)类赌。P类赌提供了很高的获胜概率。他们把右边的赌称作\$类赌，提供更高的奖金，但获胜概率要小得多。倘若你要人们在两者中进行选择，大多数人会选P类赌。

这没什么可奇怪的。选择P类赌，提高了你获胜离场的概率。奇怪的是，同样的被试总是会给右图中的\$类赌分配更高的价格。价格跟偏好相矛盾。

真正的实验中使用了12种不同的赌注。它们比上面的例子要复杂得多，还涉及被试可能输钱的情况。(8)研究人员先让被试一次看两种赌，要他们选择喜欢的一个。接着，又让他们每次只看一种赌（赌的条件跟

之前一样），请他们定价。在这一部分，研究人员告诉被试，他们“拥有”讨论中的赌注，可以把它按原价卖回赌场。他们愿意接受的最低回收价格是多少？

173个被试当中，127个一直选的是P类赌，同时又总是给\$类赌分配更高的价格。几乎人人都会出现几次偏好逆转。他们不一定意识到自己在做什么。要人记住自己先前的反应，并在执行时保持一致，这是很难的。被试跟着直觉走，而这些直觉表现出了奇怪的模式。

“很明显，这些逆转构成了前后矛盾的行为，违背了目前所有的决策理论。”1971年，两位心理学家在《实验心理学杂志》（*The Journal of Experimental Psychology*）上写道。这一文章的署名是“萨拉·利切坦斯泰因和保罗·斯洛维克”。

实验表明，大多数人分配的价格跟做出的选择不是一致的。心理学家精心设计的研究方法，更突出了这一惊人发现。

在一轮尝试中，利切坦斯泰因和斯洛维克用尽全力，确保被试仔细想过之后再做出回答。

价格实验

这组被试玩的是轮盘赌，可以真正赢钱（尽管数目不大）。(9) 每一组赌，研究人员都要向被试展示三次，并提醒他们前一轮做过的选择。被试可以改变主意。只有第三次做的选择才是铁板钉钉的。尽管有了这样的预防措施，被试还是给自己抛弃不要的赌注分配了更高的价格。

另一组测试，研究人员改变了打赌定价的方式。他们要被试假装每一个赌都想买，并报出自己愿意支付的最高价格。从逻辑上说，一个简单的赌，买入价和卖出价应该是一样的。赌注价值几何，这赌就值多少钱。但利切坦斯泰因和斯洛维克发现，人们在买进赌注的时候，给\$类赌分配较高价格的概率会小许多。偏好逆转的次数大幅减少。

这就是早期文献中所谓的“禀赋效应”(10)，倘若不存在市场价格，卖出价格一般是买入价格的两倍（绝非夸张的策略性讨价还价）。故此，利切坦斯泰因和斯洛维克检验了三种评估价值的方式，发现它们全

都有着潜在的抵触之处。

自1971年以来，心理学家和经济学家都试图解释偏好逆转。很明显，所有被试都使用了心理捷径。不管是给赌博定价还是在赌博之间进行选择，心理捷径都把事情简化了。

这里有一道利切坦斯泰因和斯洛维克测试过的选择题：

[P类赌]：10/12的机会赢9美元，2/12的机会输3美元。

[\$类赌]：3/12的机会赢91美元，9/12的机会输21美元。

依照设计，你很难一看就直觉地发现哪个赌“更好”。那你如何选择呢？早期实验的一位被试解释说：“如果获胜的概率大，我会付预期获胜金额的3/4。如果输的概率大，我会要实验人员付给我所输钱的一半。”

听到这样的话，当庄家的都会有种不寒而栗的感觉。这名被试忽视了大部分的所得信息。其实我们都这么做过。分摊餐馆账单、猜测买多长的停车时段合适……人人都会来个简单的四舍五入。因为首先，事情涉及的总金额不多，就算估计出错也没什么大不了的，我们的时间和精力恐怕更值钱；其次，可能是记忆的限制。据说，短期记忆，大概是指此刻就能进入你意识的可回忆概念仅限于7个元素左右。或许你对数字有很棒的长期记忆，笔记本电脑上有数十亿字节的金融模型，可它们只能供你参考。在决定性时刻（假设有一个“决定性时刻”），你能想得起的只有差不多7个数字或概念。

偏好逆转实验中要做的选择必然会跟这种限制产生摩擦。被试会碰到6个明明白白的数字：两组打赌的获胜率、获胜金额和损失金额。尽责的被试说不定还会计算额外的数字，比如每次下注的损失概率或预期值。可是意识一瞬间里只能停驻7个数字。思考计算出来的数字，意味着至少要暂时忘掉一部分最初的数字。用利切坦斯泰因和斯洛维克的话说：“把不同类型的信息合并到整体决定当中，人在这么做的时候会产生一种紧张感，这种紧张感或许会迫使人采用一种歪曲潜在价值系统的判断策略。”

你以为人只有在面对心理实验室里虚构的小赌注时会这样？错。大多数重要的价格决策同样为我们展现了太多的信息。

为二手车、房子、企业收购设定价格时，我们会考察数十个，甚至成百上千个有关数字。你必须抛弃大部分数据，提炼出少数几个最具说服力的数字和原因。这样做，就意味着要进行直觉的判断：哪些信息可以安全地省略掉。同样，公司同事在论证新的供货商、广告活动是否合适时，也总会掺杂着不少半真半假的直观推断。“我采纳韩国方面的报价，因为它十拿九稳。”“我总是给出预期付款的75%——有时这挺管用。”“这样，我们保证能赚回本钱，并且很有可能赚到更多。”我们把事情过分简单化，是因为要在这个世界上过活，没有别的路可走。

实验过后，利切坦斯泰因会让发生了偏好逆转的被试进行陈述。每一次，她都试图让他们相信是自己“出了错”，想看看被试是会坚持原则，还是会放弃主张。俄勒冈研究所保存了这些谈话的录音。对话中，利切坦斯泰因的开场白近乎完美。我从1968年的一段录音中做了少许摘录（建议你到网站上收听完整的音频）：

利切坦斯泰因：我明白了。那么，A赌的出价是怎么回事？你现在有没有更深入些的认识，怎么会选择了其中一个，却又给另一个出更高的价格？

被试（男大学生）：的确挺奇怪。但要说这到底是怎么回事，我并没有什么深入的认识。不就是这么一回事嘛！它表明我的推理过程不怎么样，但，除此之外，我……问心无愧吧！

利切坦斯泰因：问心无愧。好吧！有些人会说，这种反应模式并不理性。

被试：是的，我看得出来。

利切坦斯泰因：嗯，假设我要你把它做得理性些。那么，你会说它现在就挺理性的，还是会做些更正？

被试：事实上，它现在就挺理性。

利切坦斯泰因：我能不能说服你，这是个不理性的模式？

被试：我想你恐怕做不到……

你大概想知道，我们会不会放这些偏好逆转的可怜被试一马。“固执己见无非是促狭鬼的小心眼儿”，爱默生曾这么写过，从那以后，固执己见的人就爱上了他。不过，说到固执己见，这里还有几句额外的话。价格上的前后矛盾，和音乐品味的前后矛盾不一样。每个角落里都站着些精明的人物，准备从歪曲的价格里大捞一把。事实上，几乎所有人正常的、深思熟虑的定价模式都给套利提供了机会。让我们来看一个叫作“钱泵”的有趣骗术。

价格实验

利切坦斯泰因握有A、B两种赌，邀请被试对其进行定价，被

试对A赌定价550点（\$类赌），对B赌定价400点（P类赌）。

第一阶段

利切坦斯泰因：如果你认为A赌值550点，那么要是我给你这个赌，你应该愿意给我550点吧？这听起来是否合理呢？

被试：如果我要给你.....是的，这挺合理。

利切坦斯泰因：也就是说，你先用550点买到了A赌，而我有B赌以及你的550点。对吧？

被试：是的。

此时，被试付出550点，持有A赌；利切坦斯泰因得到550点，持有B赌。

第二阶段

利切坦斯泰因：你有A赌，但是如果我们真正进行赌的话，你宁愿要B赌，对吧？

被试：当然了，毫无疑问。

利切坦斯泰因：好，那我用B赌跟你换A赌。这下.....

被试：我亏钱了。

此时，被试持有B赌，利切坦斯泰因持有A赌及先前得到的550点。

第三阶段

利切坦斯泰因：接下来，我会慷慨地付给你400点以上从你那儿买B赌。付你401点，你愿不愿意把B赌卖给我呢？

被试：好的，当然行。

利切坦斯泰因：好了，把B赌给我吧！

被试：嗯。

利切坦斯泰因：我给你401点，你会发现，我本就留有你的550点.....

被试：是的。

利切坦斯泰因：给你401点后我还剩149点，也就是说，我比最开始多了149点。

此时，被试持有401点，利切坦斯泰因持有149点以及A、B赌。

第四阶段

被试：在我这边看来，推理没问题呀！（笑）这个我们还要玩多少次？

利切坦斯泰因：嗯.....

被试：好吧，我看到你赚到点数了.....

利切坦斯泰因：瞧，只要继续遵照你告诉我的反应模式，我们可以无限制地这么做下去。现在，你在“钱泵”游戏里看到，这个反应模式不.....

被试：不适合。

利切坦斯泰因：不适合。

被试：不够好。

利切坦斯泰因：对于你在上面做出的三种反应，你还是一种都不想改吗？

被试：我得花些时间想想。

“钱泵”游戏的确可以无限制地重复下去。利切坦斯泰因和被试不断交换A和B，每一轮，利切坦斯泰因都捡回149点。这简直就像从小孩儿

手里骗糖吃那么简单！只不过，这出“诡计”和街头骗术有一点区别：这个把戏货真价实，童叟无欺。每一步，被试都明白发生了什么事，并根据自己所谓的价值做出了选择。

利切坦斯泰因的苦苦追问并没能让这位被试屈服。有一次，他打趣地说，放弃先前立场“只不过能让自己看起来理性些”，但他无法这么做，因为“理性”意味着否定内心的感觉。面对宗教裁判所，倔强的伽利略坚持“日心说”——地球围着太阳转；这位先生也一样，他知道，哪怕被花招玩死，自己的估价还是老样子。

“轮盘赌或将决定人类命运”，1969年3月2日，《拉斯维加斯评论报》（*Las Vegas Review-Journal*）上打出一个奇怪的标题。大叔模样的爱德华兹上了报纸，报道说他要搞一次“科学家们设计用来探测人类内心活动”的赌博。

拉斯维加斯轮盘赌桌上25美分的一场赌局，可能涉及人类有史以来面临的重大决策。把世界投入核战争，是不可想象的灾难性决定。可总有一天，在某个地方，只要有人能把手指头放在核按钮上，就有这个可能。

利切坦斯泰因认为，偏好逆转实验“在拉斯维加斯进行简直太完美了”。有人对最初的研究提出批评，说被试或许没有动力做出合理的决定。大学生为了一星半点儿钱反复做实验，很快就会生厌。过上一阵子，他们可能会连试都懒得试了。走出实验室，碰到更高的赌注，人们才有动力投入更多时间和精力去关注决策。在拉斯维加斯进行尝试会是一次严峻的考验，它将证明偏好逆转现象是否真实存在。

1969年有10个星期，“四女王”开出了全拉斯维加斯最优厚的赔率——公平的赌博，庄家不占优。实验名叫“赌注和赔率”，赌台管理员由约翰·庞迪塞罗（John Ponticello）担任。

利切坦斯泰因和斯洛维克只在拉斯维加斯待了几天。利切坦斯泰因亲自赌了几次，检查发牌手。严格地说来，“赌注和赔率”并不是轮盘赌，而是纸牌赌。由于赌博规则完全陌生，需要在40次投注当中做出选择并定价，庞迪塞罗不得不提醒每一位被试，完整地赌一回要花1到4个小时。出于科学有效性的考虑，他们要求被试必须完成整个过程。不愿花这么多时间的人，他们会劝其退出。

价格实验

一开始，每名被试要购买250枚筹码。他可以自由指定每枚筹码代表的价值，从5美分到5美元都可以。参加的没有老赌棍，没有人指定的筹码价值高过25美分。赌博的第一阶段，被试选择电脑屏幕上显示的成对赌注，选中后按赌桌上的按钮确定。接着，他把选中的赌注放到轮盘的布局上。由于轮盘分为36格，所有的获胜概率先用12除过。庞迪塞罗转动轮盘，投出小球，叫出号码。（出现0的话不算数。庞迪塞罗会重新转轮盘。）

赌博的第二阶段，由被试指定投注的价格。价格可以是正数，也可以是负数，因为一半的赌注有利于庄家，另外一半有利于玩家。（从整体上来看，庄家并不占优。）让赌徒定出一个诚实的价格很讲技巧。我们都习惯于讨价还价，会本能地要价高、出价低，以便之后适当地调节价格。这有可能给此类实验带来严重的问题。利切坦斯泰因和斯洛维克需要被试报出一个诚心的价格 X ，这样，若出价在 X 或 X 以上，他们会乐于卖出自己的赌注，要是出价低于 X ，则宁肯不卖。

为此，他们采用了BDM机制（Becker-DeGroot-Marschak system），它其实是个很好的交易，没有听起来的那么复杂。这是一种实验室经常使用的协议。它要求卖方（给赌注，或者任何其他东西）报出一个诚心的价格。发牌手旋转轮盘，生成随机的“投标价”。要是“投标价”比底价高，交易就会按随机选中的这个标价完成。（卖家会很高兴，因为他得到的价格比自己的最低价高。）如果投标价低于卖家的价格，就不进行交易。（卖家还是会很高兴，因为要是价格低于他报出的最低价，他并不愿意卖。）在这个过程中，最优策略是诚心地报出自己的价格。

“这个实验的结果，”利切坦斯泰因和斯洛维克报告说，“和先前的实验（大学生们以虚拟赌注或小额金钱打的小赌）惊人地相似。”做选择的时候，拉斯维加斯人首选P类赌，可总给S类赌定价更高。这一回，被试掏的可是自己的钱包。赢得最多的是83.50美元，输得最多的是82.75美元。[\(11\)](#)虽然赌博是公平的，但平均起来，是玩家输钱给了庄家。这就是现实中的钱泵。

“人们自然而然地关心实验的结果能否在实验室之外的地方重复。”两位心理学家写道。在一篇谦逊低调的经典论文中，两人表述了

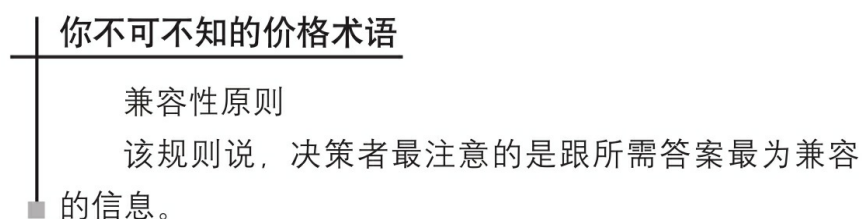
在拉斯维加斯学到的东西：“本次研究并不支持普遍的看法，即事关切身利益时，决策者能做出最优行动。”

以今之眼光看来，利切坦斯泰因和斯洛维克掀起了一场革命。我们不妨将偏好逆转实验与物理学上极为经典的迈克尔逊-莫雷实验（以下简称迈-莫实验）做个比较。迈-莫实验驳斥了19世纪物理学盛行的绝对速度，为爱因斯坦的相对论奠定了基础。我们禁不住要在物理学家的“以太”和经济学家的“效用”之间画上等号。两者都看不见、摸不着、闻不到，它们“存在”只是因为人人都假设它们必然存在。利切坦斯泰因和斯洛维克向人们揭示，并没有什么看不见的估价左右着一切的经济决策，从而吹响了价格相对论的号角——这就是当今行为经济学的基石。

利切坦斯泰因和斯洛维克为偏好逆转提出了一个简单的解释：锚定。给赌注定价时，玩家把注意力放到了奖金数额上。最可能出现或者最高的奖金数额成了一个起点或锚点。玩家们知道自己必须考虑概率或其他奖金及处罚，在锚点的基础上进行调整。这种调整需要强大的心理学为后盾。普通人不过是图方便瞎蒙罢了，结果调整幅度往往不够，最终答案跟锚点靠得太近。这就好像苹果从苹果树上掉了下来，可掉不了多远。

要人们在赌注中进行选择，启动了一个不同的思维过程。金钱的数额关系不大，因为许多赌博都没多少获胜的指望。当然了，人都喜欢赢。故此，人们强烈地倾向于选择最可能带来愉快结果的赌注。当然了，这里，被试同样会为金钱数额和其他复杂的细节预先保留些余地。可跟之前一样，调整大多不足。

不久以后，特沃斯基和斯洛维克把这一概念归纳成了所谓的“兼容性原则”。



每当要指定价格的时候，你都会把焦点放在该问题中出现的价格或其他金钱数字上。比如判断买一辆二手车要花多少钱，相关网站上的估

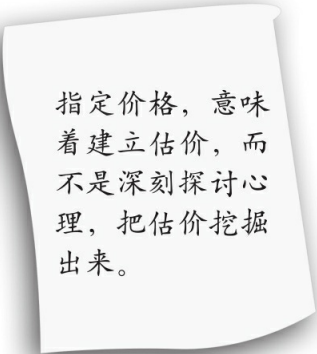
值和报价会吸引你的注意力。其他因素（车况、维修记录、颜色、选配件等）得到的关注则远远不够。因为后一类因素不那么容易反映在价格上。

利切坦斯泰因和斯洛维克靠转移注意力设计了一件“不可能完成的任务”。被试相信自己的选择和给出的价格经过了理性的思考，他们没上当，没说过任何有违自己本意的话。然而，事实表明，他们的估价颠三倒四。庄家最后用的一招是钱泵——“咻”的一声，你的钱没了。

可不管魔术师的戏法耍得多么天花乱坠，我们都知道：箱子里的女人没被切成两半，喷气式飞机并未凭空消失。当感觉跟物理定律发生矛盾时，理性总会告诉你：物理定律是对的，感觉出了错。观众们看罢魔术表演，回家的时候信心满满：事情还跟从前一样，现实的牢固基础并未土崩瓦解。

对偏好逆转而言，不存在这样的自我宽慰。我想要什么，我愿意给它出多少钱，没有人比我自己更了解。当事人对自己深信不疑——偏好逆转的“幻觉”是“真实”的，是这类事情唯一可能的根本立足点。

魔术只是人们用来描述利切坦斯泰因和斯洛维克发现的诸多比喻之一。另一种通俗的比喻说，估价是构建产生的，而并不是揭示出来的——价格是建筑学，不是考古学。指定价格，意味着建立估价，而不是深刻探讨心理，把估价挖掘出来。



指定价格，意味着建立估价，而不是深刻探讨心理，把估价挖掘出来。

1990年，特沃斯基和理查德·泰勒发表相关论文，从美国人最常用的比喻库——棒球里找出了一个有趣的说法。它引用了三个裁判的老笑话。

第一个裁判说：“我一看到犯规行为就吹了哨子。”第二个裁判说：“因为它们是犯规，我才吹了哨子。”第三个裁判却表示反对：“除非我吹了哨子它们才叫犯规，要不然就不算。”

类似地，我们可以描述对价值本质的三种不同看法：

第一种，价格存在就跟体温差不多，人们感知到它们，并尽可能最好地报告它们，尽管可能有些偏见（相当于“我一看到犯规就吹了哨子”）；第二种，人们直截了当地知道自己的价值和偏好——就好像他们懂得九九乘法表（相当于“因为它们是犯规，我才吹了哨子”）；第三种，价值或偏好是伴随诱发的过程构建起来的（相当于“除非我吹了哨子它们才叫犯规，要不然就不算”）。本文的观点与第三种态度最接近，它把偏好视为一个跟背景相关的、构建性的过程。

价格的相对性得到了确凿的支持。人们想要什么，愿意付多少钱，取决于问题如何措辞的微小细节。跟寓言故事里的盲人摸象差不多。摸到了象鼻子的人说大象像一条蛇，摸到了躯干的人说大象像一堵墙，摸到了象腿的人说大象像一根柱子。“每位盲人都说对了一部分。”漫画家沃尔特·凯利（Walt Kelly）笔下的一个人物说。“是啊，”他的朋友接嘴道，“但他们在整体上全错了。”

第5章 锚定效应：随锚点上下浮动的价格

你有过这样的经历吗？你刚买了一辆新车，突然之间你发现，高速公路上“所有人”都开着和你相同的一款车？你学到一个新名词（或者新听说一种难记的海上哺乳动物或民族舞蹈），接着却一连几天都碰到它？你在新闻里听到它，在电车里或电台里听到它，无聊时翻阅的过期《国家地理》杂志上正好有一篇讲它的文章……

这就是启动（一些莫名其妙的巧合强化了它）。在你浏览报纸，半听半看着电视，或者在高速公路上开车的时候，你会忽略周围发生的大部分事情。只有少数几件事引起了你的注意。诡异的地方在于，选择把哪种刺激传递到意识中，是一个无意识的过程。先前接触过的某东西（启动），降低了它吸引到注意的门槛，于是该东西更容易得到关注。这也就是说，其实你以前可能早就多次碰到过学会的新单词或刚买的新车了。只不过，现在你才注意到它罢了。

“启动”不仅影响着你注意到了什么，还影响着你怎么做。后一种情况中，“启动”可以用暗示的力量来实现确认。比如，沉闷的会议上，大厅里响起哈欠声，人人就都受了传染，开始打哈欠；去了苏格兰或者美国亚拉巴马州之后，你的口音突然变成了当地腔。

“说到我们每时每刻的行为，最大的问题是‘接下来做什么’。”耶鲁大学心理学家约翰·巴奇（John Bargh）说。他写过大量有关启动的论文。这个问题不大可能有一个明确的、符合逻辑的答案。相反，巴奇说：“我们发现，我们这些无意识的行为指导系统，整天都在不断调整有关接下来做什么的建议，还没等我们的知觉意识到，大脑就已开始考虑，并照它们去做了。”

你不可不知的价格术语

锚定

按照目前对启动的认识，文字和其他刺激激活了相关的精神过程。一旦“开关接通”，这一认知机制就会保持一段时间的易感性，从而影响随后的思考和行动。启动在影响数值估算的时候，心理学家们就把它叫作“锚定”。

锚定效应（大部分）是这样形成的：只要我问你，最高大的红杉树是不是有200多米高，我就启动了你去思考非常高的树，你从记忆里抽出的树木样本就产生了向上的偏差。你想到了红杉、巨杉和桉树，还有隐约记得的从“发现”频道上看到过的一切超高树种。你想到了200多米高的东西以及一棵树应不应该有那么高的原因。这一连串的思考保持活跃，使你高估了红杉树的最大高度。与此同时，其他想法也在和锚点争夺注意力。最终的答案在某种程度上权衡了诸多考虑过的数值，并做了折中处理。就算你（正确地）得出结论——地球上没有什么树能有200多米高，你也无法完全无视刚刚进行过的思考。

“谎言重复3遍就成了真的。”刘易斯·卡罗尔（Lewis Carroll）在诗集《猎蛇鲨》（*The Hunting of the Snark*）中写道。想就是信——至少，在一定程度上是这样。有些决定是逻辑思考的结果。有些决定则完全是无意识产生的，比方你要不要咳嗽一类。最重要的决定介乎于这两个极端之间。

尽管有关价格的决定在本性上是数字的，可往往含有组合强烈的直觉成分。价格不是数学问题的答案，而是欲望的表达，或对其他人怎么做的猜测（如接受或拒绝你的出价）。你给出一个“感觉”对的价格。读者们将在本书的其余部分看到，意识知觉认为不相干、不理性或政治不正确的因素，会影响到价格数字。

在电影《非常嫌疑犯》（*The Usual Suspects*）中，凯文·斯派西（Kevin Spacey）扮演一名骗子，他供称自己犯了罪，可内容全是他编出来的。审讯他的警察转动椅子，瞅见了背后的公告牌，这才发现他是在骗人。他意识到，斯派西故事里的每一个名字或细节，全是从贴在公告牌上的一张张备忘录里抽取的。侦探惊讶得连手里的咖啡杯都掉了。他拾起杯子的碎片，注意到杯子的制造商“小林瓷器”，跟斯派西刚才提到的律师的名字一样。

虚构的能力——打磨掉真实经历的毛刺儿编出故事，是人类的一种

本性。意识营造着持久的虚构，告诉你它比实际上懂得更多，行为更符合逻辑、更崇高。我们相信这是虚构。锚定是它的一小部分。在把感觉投射到数字或金钱上时，我们捏造准确性。我们总是随手抓起环境中的数字，把它们变成估计值或价格。

这个有点儿令人不安的想法，提出了另一种极端的可能性，弗吉尼亚大学心理学家蒂莫西·威尔逊（Timothy Wilson）称之为“基本锚定效应”。

任意一天里，我们脑海中都有许多任意的数字，比如电台里播报过的温度，我们刚在电脑键盘上输入的数字，我们刚刚校对过的时钟的数字，我们读过的一本书或问卷上的页码，表面上，这些一闪而过的数字恐怕不会被用来做无关的判断。

威尔逊和同事们试图找出“背景”锚点可以微妙到何种程度。有一个实验是下面这样的。

价格实验

被试们得到一份粘贴有额外小纸条的问卷。每张小纸条上写着一个介于1928和1935之间的四位数“ID号”。实验人员要一组被试把这个数字复制到问卷上，而后就请他们估计当地电话簿上医生的的人数。结果，平均的估计值是221。

这里的关键在于，该ID号只是一个偶然出现的数字，对问题没有任何意义。其他小组得到了稍有不同的指示，让他们对该ID号码多留了点儿心。实验人员要部分被试注意ID号是用红色还是用蓝色写的（表面上的借口是，这将决定被试填写问卷的哪一页）。结果，这部分人给出的平均答案是343名医生。只是对该数字多瞟了那么一眼，估计值就提高了55%。（所有的ID号都很大。从锚点的意义上来说，它们将提高估计值。）

另一个小组得到的要求是，看看ID号是否介于1920和1940之间（它们都属于这一范围）。与有关墨水颜色的问题不同，这要求被试把数字当成数字来看。结果，这一组被试估计电话簿上的医生人数平均为527名。

还有一个小组要回答两个问题。实验人员先要他们猜测电话簿上医生的人数是大于还是小于ID号，接着又请他们做出具体的估计。结果，这个小组的平均答案为755名。

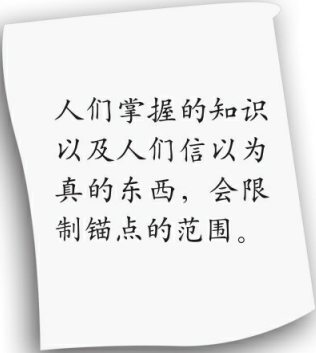
到目前为止，在人们对锚点和估计值进行的明确比较中，锚定效应表现得最为强烈。锚点数字即使是相当迂回地出现在外围，也会影响答案。

之后，研究人员还询问了一些被试，对于他们的判断是否受了ID号的影响。绝大多数人的回答是“没有”。正如斯派西在《非常嫌疑犯》中所说，“魔鬼耍的最大把戏，是要人相信它并不存在”（引自法国诗人波德莱尔的诗作）。

- 旧金山的平均气温是高于还是低于290摄氏度？请估计旧金山的平均温度是多少？
- 披头士乐队发行过多少张打入排行榜前10名的唱片？是多于还是低于100 025张？请你估计披头士乐队到底发行了多少张排行榜前10名的唱片。

你肯定以为它们不可能会影响到人们对旧金山气温的猜测吧？不可能影响到对披头士热门唱片的估计吧？可惜，它们真的影响到了。被试对这些荒谬得离谱的高位锚点先入为主，给出的估计值比看到低位锚点的人要高得多。

当然不会有人猜旧金山的气温接近290摄氏度。人人都应该知道它是个两位数，介乎常温和零点之间。人们掌握的知识以及人们信以为真的东西，会限制锚点的范围。如果一个地理迷早就知道非洲国家在联合国里所占比例为多少，他就能不受随机数字的左右，给出正确答案了。锚点是用来辅助猜测的工具。



人们掌握的知识
以及人们信以为
真的东西，会限
制锚点的范围。

威尔逊领导团队进行了一次实验，让被试估计本地电话黄页上有多

少名医生。估计最准的人有奖励：可带一名朋友到某热门餐馆吃饭免单。这回的问题照样分为两个部分，而且针对不同群体给出了或高或低的锚点。威尔逊和同事们以为，有了免费大吃一顿的诱饵，被试或许能全力以赴给出最佳答案，而不是随随便便地写上个脑袋里冒出的愚蠢数字。结果，他们发现，锚定效应跟没有奖励时差不太多。

威尔逊的研究小组甚至尝试预先提醒被试，随意设定锚点有误导作用。一组被试得到的指示上说：“人们脑袋里的某个数字，会影响他们对随后问题的回答……在你回答以下问题时，请千万小心，别让这种污染效应影响到你。我们希望你能得出最准确的估计。”

这个警告完全没起作用。被试的估计还是受了毫无意义的数字的影响。威尔逊的团队认为，得到提醒的被试很可能是想纠正锚点，却做不到。这就好像你提醒别人，脑袋里别想大象哦！对方诚心诚意地照着你说的做，可还是满脑子大象。

威尔逊的团队认为，由于锚定效应是无意识下不知不觉发生的，所以人们很难知道锚点数值对其估计值要造成多大的影响，故此，他们对锚定效应的感染性怀着天真的念头。

举个现实生活中锚定效应的例子：去看看百老汇和拉斯维加斯演出的门票价格。

价格实验

1999年，百老汇一位坦率的制片人在博客“谈谈百老汇”（Talk in Broadway）上说：“便宜座位根本不卖——知道他们为什么不卖吗？因为要是你把靠近舞台的座位或者包厢座位的价格定得太便宜，人们会觉得里头肯定有鬼。”

百老汇的生意主要靠游客维持。游客们选剧码的时间有限，对自己买的票值不值价没什么了解。几乎没有谁知道该如何判断哪个座位值多少钱。在评估座位价值时，游客们除了从票价上找线索就没别的法子可想（“一分钱一分货”嘛）。一张票的感知价值跟其价格成正比——不管那价格到底是多少。不少人相信，看一场戏剧，花480美元买票并不亏，对制作人来说，要维持该剧目长期上演且赚到钱就得卖这个价。游客们以为，一场演出门票敢卖480美元，那一定值得看，于是径直跑去售票厅。

还有重要的一点：即使是绝不会付480美元买票的戏剧迷们，也还是会受这一价格的影响。跟这个价比起来，他们不管出多少钱都显得特别划算（毕竟，这可是同一场戏剧）。

配价指的是给剧院或音乐厅不同部分的座位分配价格。这是生意的重要一环，它甚至能决定演出是座无虚席，还是冷冷清清。一位匿名制作人透露说：

现在，所有的前排座位和大部分包厢，我都给配最高的价。如果你这么做，你就能很快就把它们销售一空。我也可以按若干不同的价格给演出配价，最高到最低的都有，可结果呢，只有价格最高的位置卖光了，大部分便宜座位都空着。我还可以给70%到80%的座位都配上最高价。你知道，要是大部分座位都是最高价，那就算只把40%的票放到售票厅出售，我也还是能赚到更多的钱。

多年来，好莱坞露天剧场（Hollywood Bowl）的夏季音乐会最便宜的门票才1美元。露天剧场由洛杉矶县经营，1美元的票以服务公众为目的。麻烦的是，从没买过这种票的人总是暗暗以为它糟糕透顶。露天剧院很大（共有17 376个座位），1美元的票离舞台最远。但观赏体验基本上差不多。1美元的票的座位欣赏落日和城市景观效果也更好。可大部分时候，100美元的座位挤得满满当当，一票难求，1美元的票的座位却无人问津。好些音乐爱好者们错过了机会——就因为价格太低。

把个人价值翻译成可以跟别人沟通的数字，很多时候都是极为必要的。为了实现这一目的，我们的心理软件提供了锚定的功能（当然，说是程序缺陷也不错）。每当要估测一个没法计算的未知数量，我们都很容易受刚刚提及或考虑的其他数字的影响。我们意识不到这一点（为了从统计学上论证它，研究人员做了无数次实验），但它却的确存在。为了帮助我们做出大胆的猜测、依照直觉行事，为了帮助我们在餐巾纸上给出报价、讨价还价，为了帮助我们给餐馆和性伴侣打分，以及更概括地说，为了帮助我们在一个迷恋数字和金钱的社会里运作自如，锚定是整体心理过程的一环。锚定适用于各种各样的数字——包括前头带有金钱符号的那种。

“事实上，在所有需要判断的领域，都存在着共同的原则。”美国心理物理学家威廉·亨特（William Hunt）言之凿凿地说。在若干实验中，他让被试“按罪行违反道德的严重性”给犯罪行为打分。他给出的是如下一道难题。

价格实验

第一部分：先想想谋杀亲生母亲这一罪行，它是“蓄意的，找不到任何站得住脚的开脱理由”，然后，想一种恰好相当于这桩罪行的一半的坏事，把它写出来。

第二部分：再次回到谋杀亲生母亲这一罪行，现在想想“自己玩蜘蛛纸牌游戏时作弊”，最后，设计一种罪行，它的严重性恰好介于上述两种罪行的中间，把它写出来。

从罪恶量表上来看，玩纸牌时作弊几乎只能得零分。你大概会以为，第一部分的答案会跟第二部分的答案差不多。事实却不然，在14个被试中，12个人在第一部分写出的罪行，比第二部分要严重得多。

亨特的结论是，他在问题中给出的例子影响了被试的回答。在第一部分，参考框架只有谋杀亲生母亲这一个。这暗暗怂恿人联想到了其他残酷的罪行。第二部分则有两个例子，一个是重罪，另一个却不是。不会有什么人觉得在纸牌游戏里作弊会是“犯罪”。但问题本身却称之为“罪行”，这就鼓励被试把无伤大雅的小过失想成是“罪行”，从而拉低了答案里罪行的平均严重性。

亨特把这一效应称为“锚定”（这个词是另一个意思）。他区分了两种不同情况。

你不可不知的价格术语

对照锚点与同化锚点

如果你比较两种刺激，出现的是对照锚点。路灯的光芒让天上的星星看起来黯淡了不少；跟在别人后头的小丑碰到了倒霉事儿，比先前好笑40%。倘若给出一种或若干种可能的反应，非要你想一个答案，出现的则是同化锚点。在人们把一种罪行说成相当于另一种罪行的一半，或是在陪审员听完律师的索赔要求后，裁定损失赔

偿的时候，同化锚点就来了。这两种类型的锚点起着相反的效果。使用对照锚点，人的主观感受会偏离锚点。使用同化锚点，反应则会受锚点所吸引。

赫尔森花了大量的精力，尝试理解一种经历怎样才能成为锚点，得以影响判断。他的回答是：“近因、频率、强度、范围、时限，以及诸如意义、熟悉程度和自我参与等较高层次的特性。”这句话听起来挺绕的，但没你想的那么难理解。

价格实验

从近因开始。你先把3盎司的砝码举了几秒钟，再举5盎司的砝码就显得重了。可要是举这两枚砝码，中间间隔了一个小时，对照效应就消失了。你忘了前一个砝码感觉起来有多重。

频率也很重要。你一连举了好几个3盎司的砝码，于是对这一程度的重量产生了适应。要是你接着举5盎司的砝码，便会感觉重。若干个3盎司砝码产生的锚定效应，强于一个锚点产生的效应。

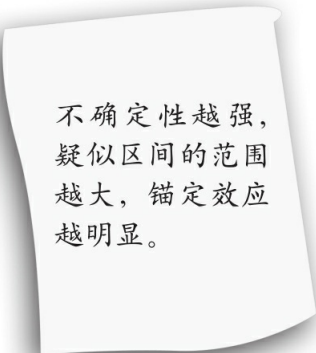
赫尔森最有趣的发现在于诸如意义等“高层次的特性”。他对某些被试设了个套。在实验的中途，他要求被试把一托盘砝码挪开。托盘（外加上面的砝码）比之前实验里用过的所有砝码都要重。但沉重的托盘并没有让被试接下来举的物体显得更轻。被试全神贯注于小的金属砝码，根本没注意到托盘，所以下意识地把托盘给“屏蔽”了。这说明，锚定不是肌肉的反应，而是精神上的反应。

早就有人提出，可以把锚定和调整视为偏好逆转的成因之一。在利切坦斯泰因和斯洛维克的实验里，锚点（奖金数额）至少是跟赌博价值挂钩的。特沃斯基和卡尼曼却怀疑，哪怕锚点根本不相关，锚定效应也仍会发挥作用。为了检验这一假设，他们设计了联合国实验。幸运转盘是道具，用来强调锚点数字完全是随机的，没有任何意义，可它还是照样管用。在心理学对理性提出的所有挑战当中，“锚定效应是最容易验证的，可最难以解释。”

联合国实验已成为锚定效应的经典例子。但问题在于，它到底是调整，还是（用现代的术语来说）启动。特沃斯基更喜欢认为它是实际调整。

特沃斯基的想法是这样的。让被试猜测非洲国家占联合国成员国的百分比，他们会从一个锚点值（从幸运转盘上得来的数字）开始，向上或向下进行调整。他们会一直调整，直到进入一个模糊、宽泛的疑似值

区间范围。此时，他们停了下来。休止值将位于疑似区间靠近锚点的一侧。不确定性越强，疑似区间的范围越大，锚定效应越明显。



不确定性越强，
疑似区间的范围
越大，锚定效应
越明显。

这就好像我请你去帮我买个汉堡，你可能碰到第一家汉堡店就停了下来，在那儿买回汉堡带给我，而不会跑遍全城去买最好吃的汉堡。

按特沃斯基的理论，人们从锚点开始做调整，停下得太早了。他们不会梳理大脑寻找“最佳”答案，而是满足于得出的第一个差不多的答案。倘若锚点本身就高，该答案也会过高；而要是锚点低，该答案也会过低。

最初，特沃斯基指示被试从幸运转盘得出的数字开始，在心底默默向上或向下调整。这一指示，具体表达了特沃斯基对实际情况的想法。现在，人们看得很清楚，根本没必要给予指示。重要的是，锚点和待估数量之间存在某种心理上的比较。这一过程会很自然地出现在偏好逆转实验当中。当锚点毫无信息意义时（随机数字或明显错误的数字），你只要随便提个问题，比如“非洲国家占联合国成员国的比例是在65%上下吧”，对比就自动弹了出来。

特沃斯基的调整论，解释不了先前给出疑似值的锚定。

(1) 爱因斯坦第一次访问美国，是在1939年之前还是之后？

(2) 爱因斯坦第一次访问美国，是在1905年之前还是之后？

看到前一个问题的德国学生，给出的年份晚于看到后一个问题的学生。前述两个锚点年份看起来都有道理（真正的时间是1921年）。照理说，倘若数字已经合乎情理了，就没必要再调整它了。

针对锚定现象，人们又提出了许多其他的解释。有人认为，锚定是合乎逻辑的，被试接受了实验者给出的“对话提示”，相当于抓住了救命稻草。除非有个合理的答案，不然，实验者不会问爱因斯坦是在1939年之前还是之后访问的美国。故此，给出一个靠近1939年的年份，不会令被试显得太笨。

锚定也可应用在法律上。早在利柏克诉麦当劳案发生前，就有人发表过相关研究，探讨锚定效应对陪审团裁定赔偿金额造成的影响。1989年，心理学家约翰·马洛夫（John Malouff）和尼古拉·舒特（Nicola Schutte）做了一次研究。

价格实验

安排4组模拟陪审员，让陪审员阅读一份真实人身伤害案的材料，在该案件中，被告罪名成立。两位研究者告诉所有的4个小组，被告辩护律师建议赔偿50 000美元。4个小组所接触的信息，唯一不同的是原告律师要求的赔偿数额。一个小组听到的是，原告律师索要100 000美元，这一组陪审员裁定的平均赔偿额是90 333美元。另一个小组得知律师索要700 000美元的赔偿额，他们裁定的赔偿额平均为421 538美元。

要是陪审员们能够裁定“正确”的赔偿额，那所有小组给出的数字应该一样。案件的事实没有变化。当然了，法律赔偿并没有标准的计算公式。于是陪审员们很容易受到他人建议的影响。把马洛夫和舒特的4个数据点（他们向另两组展示的索赔额是300 000美元和500 000美元）绘制成图，就可以看出它们完全是一条直线。陪审员们判定的数额虽然总是比原告律师要的低一些，但基本上是跟着要求的索赔额一路上涨。

就算是痴心妄想，也不会有几个律师觉得陪审员能任自己摆布。这个研究，再加上另一些研究，提出了以下问题：你在法庭上可以把锚点抛到多远？一个聪明的律师会提出数亿美元的庞大索赔额吗？

传统的智慧说，不行。据说有个叫“反弹”的神奇效应：高得过火的索赔额，会让原告或律师显得太贪婪，从而带来事与愿违的结果。出于报复，陪审团的裁定金额会很低。较之一开始就提出合理索赔额的做法，漫天要价是不划算的。

1996年，在利柏克诉麦当劳案频频成为新闻焦点期间，心理学家格

雷切·查普曼（Gretchen Chapman）和布莱恩·伯恩斯坦（Brian Bornstein）对这种说法做了测试。

价格实验

他们向伊利诺伊州80名大学生提出了一个假想案例：一名年轻姑娘说，自己因为服用避孕药患上了卵巢癌，对医疗保险提供商提出控告。参与的学生被分为4组，每一组听到的索赔额都不一样，分别为100美元、20 000美元、500万美元和10亿美元。研究人员要求模拟陪审员裁定补偿性赔偿金。所得的结果，凡是信任陪审团制度的人看了恐怕都会吃惊不小。

在500万美元的限度之前，陪审员们惊人地好说话。超低的索赔额100美元平均只得到了可怜兮兮的990美元。这可是害得原告“疼痛不断”的癌症啊！而且“医生觉得她没几个月可活了”。当索赔额增加到这个数字的20倍，到达20 000美元时，陪审员的裁定额则增至36倍多，到了36 000美元；索要500万美元时，则裁定额又是36 000美元的12倍多。

查普曼和伯恩斯坦的实验不能排除反弹效应，但也没有找到任何支持它的证据。相反，它发现收益递减。要求10亿美元（一个超级疯狂的数字），所得裁定赔偿仍比要求500万美元时更多。只不过多得没那么离谱罢了。

道听途说的证据会造成误导。律师们总记得自己要了个天价，所得却比期望值要少的时候。真的疯得够彻底、胆敢索要10亿美元赔偿的律师，听到裁定赔偿才区区490 000美元，当然会倍感失望，怪罪于反弹效应。然而，上述实验结果表明，10亿美元的数字在4个受测索赔额中表现最好。

照理来说，陪审员应当根据病人痛苦和受难的程度做出赔偿裁定。查普曼和伯恩斯坦要陪审员们用数字给原告的痛苦程度打分。他们发现，陪审员对痛苦的估计和判定的赔偿之间没什么有意义的关联。换句话说，理应有关系的变量事实上没关系，照理说不相关的变量（原告的索赔要求）实际上却有关系。

两位心理学家还询问陪审员：“原告的伤害是被告的产品造成的，这种情况的可能性有多大？”陪审员们报告的可能性跟他们裁定的赔偿额存在一定的关系（可能性越大，裁定赔偿额也会相应提高），但并不

特别突出。因此，没有证据能证明，10亿美元的索赔损害了原告案件的可信程度。

负责利柏克诉麦当劳案的里德·摩根，说自己这样的律师叫“创业家”。这类诉讼律师们专找巨额赔偿官司，迫使大公司关注产品的安全性。不太同意这套说辞的观察人士则叱责这是“把打官司当成抽彩票”。从这两个角度来看，面对法律的财富大转盘，有时律师们会不愿向陪审员提出太高的赔偿额。他们觉得，提出合理的数字说不定反倒能够带来一笔意外之财，要得太多则可能引起陪审员反感。查普曼和伯恩斯坦的实验表明，事情并非如此。他们的论文题目说得很直白：“要得越多，你得到的就越多。”

锚定研究令一些人相信，不应由陪审员直接裁定损害赔偿额。卡尼曼认为，陪审员试图用前后矛盾的金钱语言表达他们对被告行为的愤慨之情。就好像陪审员从火星来，不知道金钱在这个星球上价值几何一样。究其本质，他们是用1到10的分值来给被告的过失打分。他们通过律师寻找线索，判断它值多少地球币。

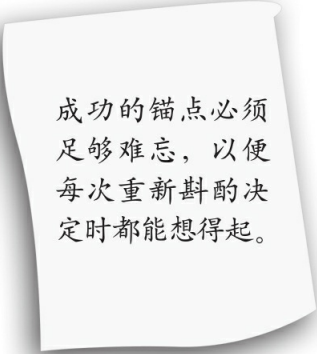
在利柏克诉麦当劳案中，摩根成功地激起了陪审员们的义愤。他连打两记重拳：麦当劳的咖啡比许多竞争对手要烫；麦当劳对利柏克受的伤害无动于衷。在审判的裁决阶段，摩根要求陪审员们罚麦当劳一两天的全球销售额：麦当劳咖啡一天的销售额大概是135万美元。

什么？事故的肇事者是咖啡？摩根没有过多解释为什么这一要求是合理的，或许是因为它并没有特别合理的地方。你对这个要求想得越多，就越会觉得它没道理。为什么是一两天的销售额？为什么是全球的销售额，而不仅仅是美国境内的销售额，或者新墨西哥州的销售额，甚至是利柏克那天买的那杯咖啡的售价（49美分）？

这里，“想”就是关键。一个有效的锚点必须是存储在做决定当时的短期记忆当中的。表面上看，这是一个很大的限制。短期记忆，就是我们用来拨打陌生电话号码的那种，仅仅能持续大约20秒钟。这就是为什么许多人都觉得锚定效应恐怕并不适用于实验室外面的环境。陪审团可能会讨论好几天。陪审员会感到无聊，大部分时间都在做白日梦。天知道他们要接触到多少个数字。

现场研究表明，现实当中，锚定效应可以持续相当长的时间。像陪审团裁决这样重要的事情，不见得有什么做决定的关键一刻。每名陪审

员都会在陪审席上多次考虑这一问题，其间间隔着不经意的分心。每当控辩双方提出新的论点，或确定新的证据，他们就会重新考虑自己的决定。成功的锚点必须足够难忘，以便每次重新斟酌决定时都能想得起。



成功的锚点必须
足够难忘，以便
每次重新斟酌决
定时都能想得起。

摩根提出的索赔要求，哪怕不合情理，也足够叫人过目不忘。麦当劳一两天的咖啡销售额设定了理想的惩罚标准。它框定了人们的思考，鼓励陪审员暗暗寻思两个问题：

(a) 一两天的咖啡销售额公平吗？

(b) 多少天的咖啡销售额才公平？

其实，陪审员并不擅长根据罪行或问题的严重程度来设定赔偿金额。1992年，W. H.德斯冯格斯（W. H. Desvousges）和同事主持了一次调查。

价格实验

他们告诉被试，由于炼油厂的蓄油池没盖盖子，鸟儿们掉进去溺死了。这个问题（当然是虚构出来的）可以通过在蓄油池上设置防护网来解决。他们询问被试，为了拯救小鸟，他们愿意出多少钱来架设防护网。研究人员对不同小组所说的禽鸟死亡数量不同：有的说一年2 000只，也有的说一年200 000只。结果，人们的答案跟鸟死掉的数量并不相关！不管怎么说，被试平均愿意掏腰包的钱数在80美元上下。显然，人们只记住了一点：“死了很多鸟儿。我们应该做些事。”

摩根肯定希望利柏克诉麦当劳案的陪审员们从财大气粗的麦当劳钱

包里狠狠地掏出一笔钱（因为咖啡太烫打起官司的，小餐馆碰到的很少）。“一天的咖啡销售额”能充当有效的赔偿尺度，还有另一个原因。一旦陪审员在多少天上达成了一致，递增幅度也就一目了然了。

你可能想知道为什么摩根会要求“一两天的咖啡销售额”。为什么要说得这么含糊呢？当碰到三种价格（不妨把它想成是小杯咖啡、中杯咖啡和大杯咖啡的售价），自身又没有强烈偏好时，人们一般倾向于选择“中间”价格。摩根说不定估计到，被告或者没什么同情心的陪审员兴许会提出一个低得多的数字。于是，他引入了“中间”选项，给了迟疑不决的陪审员们一条轻松的出路，同时又有利于自己的客户。

利柏克诉麦当劳案的陪审团裁定270万美元的惩罚性赔偿，正好是摩根估计的两天咖啡销售额。不容否认，摩根的索赔请求具有令人信服的影响力。按照前面的研究，摩根唯一的失误之处，或许就是他没干脆要求麦当劳赔一两年的咖啡销售额。

第6章 记忆价格和前景理论：我们对价格很愚钝

电视剧《巨蟒》（*Monty Python*）描述过阿尔冈星球上的一个任务。

阿尔冈星是金牛座的毕宿五(12)，跟1972年的英国很像——只不过那儿的的价格都是天文数字。剧中演员约翰·克利斯（John Cleese）说：“这里一杯普通的热巧克力售价400万英镑，烧开水用的浸入式加热器售价60多亿英镑，一条内裤的价格高到了无法形容的地步……一个像这样的电热水壶新零件，价格几乎相当于美国自1770年到2000年的国民生产总值。而即便如此，也还不够买个热水壶上的小小固定环。

过了一会儿，迈克尔·佩林（Michael Palin，该剧的另一位演员）插嘴说：“旋转式割草机的附件相对便宜！在9 000万到1亿英镑之间，不过这似乎说明，阿尔冈星对那些有着大花园的家伙来说是个很好的地方。”

你或许会想，其实阿尔冈星跟地球又有多大的不同呢？我们出生在太阳系的第三行星，对东西的价格没什么概念。说不定我们怎么学都学不会。我们能做的，就是从周围人身上找线索。我们假装他们神智健全，提出的价格合乎情理。

要是让笛卡尔来推断价格，他或许会说：我们唯一知道的就是相对价值。从深层意义上来说，我确实无法判断割草机的附件卖1亿英镑是不是个好价钱，但我知道它比其他价格要便宜。短短几年，靠着一些出色的实验，这种相对价值重要、绝对值没什么意义的愚蠢价格观已经得到人们的普遍接受。你可能会说，它们表明，我们都生活在阿尔冈星上。

阿雷利是另一个深入思考定价心理的杰出以色列裔美国人。他的好多研究都可以追溯到他头一次走进一家巧克力商店的经历。

他眼前摆着一列列美不胜收的巧克力，价格也高得令人咋舌。“我正盘算着想要哪一块，”他说，“接着就意识到两件事。一是我迅速适应了那里的价格水平。我没去想超市里的巧克力卖多少钱。另一件事是我非常容易受影响——不管商店标注了什么样的建议零售价，我都把它们当成入手的合适价格。”

阿雷利是杜克大学的行为经济学教授，他做了好多引人注目的实验，证明价格的灵活性有多大。其中一个跟乔治·鲁温斯坦（George

Loewenstein) 和德拉佐·普雷莱克 (Drazen Prelec) 合作完成的, 实验内容是无声拍卖比利时巧克力、红酒和电脑设备。

价格实验

投标的是麻省理工学院斯隆商学院的MBA考生, 研究人员要他们写下自己社会安全号码的末尾两位。接着, 每名投标者要说出他的心理投标价是高于还是低于这个两位数 (按美元计)。最后, 投标人写出自己愿意出的价格 (诚实的心理保留价)。中标的人自己掏腰包出钱, 拍得的东西就归他们了。

拍卖物品里有一瓶1998年的罗纳谷村酿红酒。例如, 我的社会安全号码末尾两位数是23, 所以, 我要回答的第一个问题是: “这瓶酒, 你愿意出的价是高于还是低于23美元?” 第二个问题则是: “你愿意出多少钱呢?”

结果显示出惊人的锚定效应。社会安全号码“低” (末尾两位数为00~19) 的投标人, 对这瓶酒愿意出的价格平均为8.64美元。社会安全号码“高” (末尾两位数为80~99) 的人, 愿意出的价格平均为27.91美元。不光红酒是这样, 在巧克力、无线键盘鼠标和设计书的拍卖中也有类似的差异存在——全都是因为社会安全号码。大多数的情况都是碰巧有着较“高”社会安全号码的学生拍走了商品。号码较“低”的人则一无所得。至于谁是真正的赢家, 以及谁是真正的倒霉蛋, 你自己来判断吧!

阿雷利开辟了一个影响广泛的论题: 记忆价格遮掩了人类不准确的价格判断力。假设让买东西的人猜测一台椭圆踏步机的价格, 他会努力回想从前购买健身器材的价格, 或是他看见广告里的椭圆机卖多少钱。他会根据产品的质量和性能做调整, 得出一个差不多的数字。然而, 从某种意义上来说, 他跟奥斯卡·王尔德的讽刺语录里说的一样: 对样样东西的价码都知道, 对价值却一无所知。

记忆价格遮掩了
人类不准确的价格
判断力。

麻省理工学院斯隆商学院的拍卖，其设计目的是消除记忆价格的部分影响。他们挑选的要么是学生们不大可能购买的东西，要么就是众所周知价格波动幅度大的东西。（人们喜欢把葡萄酒和比利时巧克力当成礼物，部分原因就在于收礼的人很难猜到送礼人花了多少钱。）问题不是“你还记得它的价格吗”，而是“它对你来说值多少钱”。

拍卖结果跟史蒂文斯的量值量表实验很相像。相对比值趋于一致，绝对值却大相径庭。图6-1显示，平均投标额按社会安全号码的末尾两位数为5个区间。

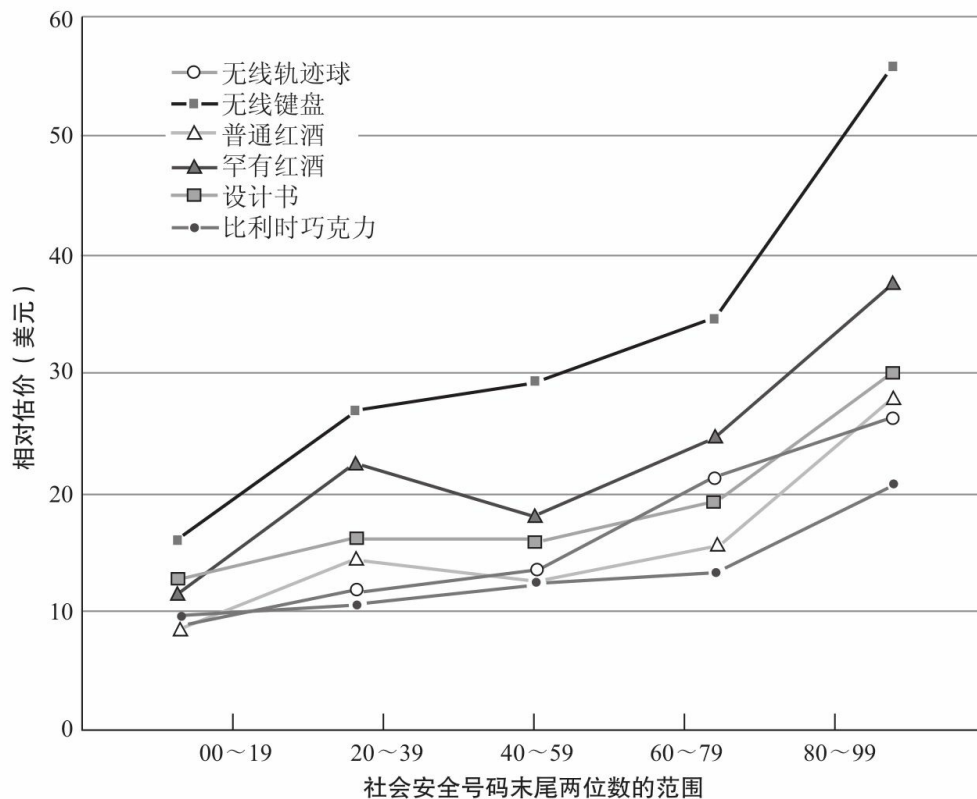


图6-1 不同物品的投标情况

每一条线段代表不同物品的投标情况。就这个实验的目的而言，社会安全号码可以视为一个随机数。一般情况下，5个随机小组的平均估价应该差不多。结果，所有的线段都呈上升趋势。社会安全号码末尾两位数“低”的人（左）低于号码末尾两位数“高”的人（右）。这暗示了锚定效应的存在。

在每组社会安全号码区间之内，不同物品的相对估价跟其他小组大致相当。所有小组都认为无线键盘是最值钱的东西，比利时巧克力的估价最低或差不多最低。罕有红酒的估价总是比普通红酒的更高，而且，在每个小组里，前者的价格都差不多是后者的1.5倍左右。

阿雷利、鲁温斯坦和普雷莱克的理论认为，被试是用后推法来建立自我一致性的。他们写道：

假设一个被试的社会安全号码末尾两位数是25，对标示为“普通”的红酒，他的先验（愿意支付价）范围是5~30美元，对“罕有”红酒则为10~50美元。故此，这两种红酒，以25美元的价格都有可能买得到，也都有可能买不到。假设被试出于某种原因说自己会按25美元的价格购买普通红酒。那么，倘若我们过一会儿问他，愿不愿意用25美元购买“罕有”红酒，他的回答肯定是“当然愿意”，因为从他的角度来看，这个特殊的“选择问题”已经解决了，方法也明摆着：既然普通红酒价值25美元，“罕有”红酒必然要高于25美元！而且，被试之后在报出对红酒的愿意支付价时，这个问题也有了足够的限制条件：价格必须符合这样的排列顺序——两者的价格都得高于25美元，且“罕有”红酒的价格更高。

阿雷利的研究小组将这些结果写成论文，发表在2003年的《经济学季刊》（*The Quarterly Journal of Economics*）上，题为《“任意连贯性”：没有稳定偏好的稳定需求曲线》（“*Coherent Arbitrariness*”: *Stable Demand Curves Without Stable Preferences*）。这篇论文还为记忆理论提供了更惊人的证据。

“我们想找一种人们没有强烈参考价格的东西，”阿雷利解释说，“他们需要一种有待标价的全新产品，这种产品是——痛苦。”

价格实验

麻省理工学院的132名学生通过耳机，听到了恼人的高亢声音（类似救护车警报器发出的3 000赫兹声音）。屏幕上的说明写着：

几分钟后，我们将在您的耳机中播放一种新的烦人声音。我们想知道你认为它有多讨厌。等你听到该声音后，我们会立刻问你是否愿意重复该体验，换取10美分的报酬（另一小组为50美分）。研究人员要被试分别为听烦人声音10秒、30秒和60秒的选择标价。

看到低锚点（10美分）的人标出的价格，总是比看到高锚点（50美分）的人标出价格的要低。所有人的价格都是按照倾听时长等比例增减的。此外，对同一被试重复试验，也不能消除初始锚点造成的影响。大多数人坚持原价，丝毫没有意识到它是由一个毫无意义的锚点提示出来的。

价格实验

研究人员又列出若干小烦恼，要一些被试把这些和烦人的声音进行排序。小烦恼包括“发现自己买了一罐坏牛奶”“忘了还录像带，必须付罚金”“冰激凌掉地上了”及其他7种项目。总体而言，烦人的声音排在第二位，位于“只迟到了几秒就错过了公车”之后。

神奇之处在于，10美分和50美分的价格锚点对烦人声音的排名没有任何影响，大家有着基本一致的看法。

价格实验

另一组被试答应把一根手指放进老虎钳里。研究人员收紧老虎钳，直到被试开始感到疼痛。接着，研究人员将老虎钳再收紧1毫米，并要被试记住疼痛的程度。研究人员把手指头从老虎钳里松开以后，会提出以下问题：你愿意忍受哪一种折磨呢？是被老虎钳夹30秒，还是听烦人的声音30秒？

大多数人选择了烦人的声音。这里，锚定效应对人们倾向于烦人的声音或老虎钳同样不存在统计上的影响。锚点只影响价格。

经济学家早就阐述过，对金融事务来说，决断性和前后一致是最理想的状态。显然，这不只是学者们提出来的抽象概念，也是一般人努力想要达到的普遍理想。我们都假装有着前后一致的理论及常识。但事实却是，我们知道的不过是相对估值。我们对比率敏感，对价格却很愚钝。

还记得阿莱及阿莱悖论吗？跟特沃斯基合作之初，卡尼曼就认定阿莱悖论是决策心理学没能解决的首要问题。要能找出一个合理的解释，必将得到莫大的奖励。为了做到这一点，他和特沃斯基开始设计“有趣的选择”。要是选择一个选择看起来挺有意思，他们的家人就成了实验专用“小白鼠”。有下面这样一个例子。

价格实验

你是愿意稳拿3 000美元，还是80%赢取4 000美元的机会（以及20%两手空空的可能性）？

几乎人人都偏爱稳拿3 000块的选项，而不选择冒风险。这没什么好奇怪的。为了让事情变得更有意思，特沃斯基出了个巧妙的主意。在每一笔数额前头摆上一个“负号”，“反射”问题。这么说你会更容易懂些。

假设有人起诉你，要你赔偿4 000美元。你是愿意现在花3 000美元调解（固定的损失，-3 000美元）呢，还是去打官司？假设你事先就知道有80%的可能性会输，需要全额偿付（-4 000美元），有20%的机会赢，一分钱都不出（不考虑法律费用）。

要是人们对风险的态度一致，对第二个问题的回答似乎应该跟前一个相同。可大多数人并不是这么想的。绝大多数人拒绝确凿的损失，倾向于上法庭打官司搏一把。他们的思路是这样的：“我可不想损失几千美元，只有打官司才有可能避免。损失3 000美元还是4 000美元，区别没那么大。”

靠着提出形形色色的此类问题，卡尼曼和特沃斯基得以从各个方向观察人们对待收益、损失和风险的态度。他们不再麻烦家里人，而是开始找学生被试进行谨慎的调查。1979年，他们发表了文章《前景理论：风险条件下的决策分析》（*Prospect Theory: An Analysis of Decision Under Risk*）。卡尼曼说，标题里的新名词是他们故意创造的：“我们觉得，要是这套理论能广为人知，有个与众不同的名字会很占优势。”

前景理论建立在若干简单、有力的概念之上。其第一个关键概念是金钱（或一般而言的得失）的相对性。卡尼曼和特沃斯基提出了和心理物理学相似的主张：

我们的感知系统，对估计绝对数量不如变化量在行……接触相同的问题，我们既可能觉得热，也可能觉得冷，这完全取决于我们先前适应的温度。

同样，人们会习惯于一定水平的财富或收入，并主要对变化做出反应。假如你指望有钱的阿姨送给你1 000美元的支票做结婚礼物，因为你的兄弟姐妹们都享受了这个待遇。结果，她却送给你一张价值25美元的蹩脚贺卡！你很容易觉得自己损失了975美元，而非获得了25美元。

按卡尼曼和特沃斯基的术语，期待中的1 000美元是一个基准点。这很像心理物理学上的“适应水平”。基准点被用于判断进入心理账户的某样东西属于收益还是损失。这可以造成行为上的巨大差异。

前景理论的第二个关键概念是厌恶损失。损失金钱（或任何有价值的东西）带来的伤害，远远大于获得同样东西带来的喜悦。抛硬币打赌就能看出损失厌恶来。抛出正面你输100元，抛出背面你赢X元。X要多大，你才愿意接受这个赌？

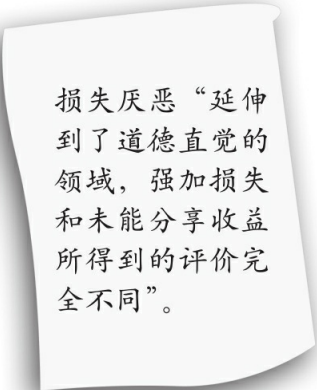
调查显示，很少有人愿意接受X=100的“公平”打赌。X=110的情况也很少有人接受，哪怕它提供了挺不错的预期收益。（接受这个价格的人往往是赌徒或经济学家。）一般人会要求200美元左右的奖金，来平衡100美元的预期损失。

不管收益还是损失都不是简单的累加。获得20美元意外之财的惊喜，少于获得10美元惊喜的两倍。这跟史蒂文斯的小谜题是一个意思，它说要差不多40美元才能让人感觉比10美元“好两倍”。经济学家一直都知道，大规模的收益和损失不是累加的，但前景理论把这一规律扩大到了看似完全微不足道的数目上。从人们的行动来看，“财富效应”似乎也适用于小变化。

有一个通俗的比喻表达了前景理论背后的部分概念：“金钱是毒品。”对大把钞票上瘾的人，会对滥用金钱达到一定的适应程度。之后，要想再次“高”起来，他必须超过原来的基准线。要是上瘾者没能达到基准线，他就会体验到“戒药”的痛苦。戒药的痛苦远甚于“高”起来的愉悦。

卡尼曼观察到，损失厌恶“延伸到了道德直觉的领域，强加损失和未能分享收益所得到的评价完全不同”。有反盗窃的法律，可没有反吝啬的法律。

前景理论的第三个关键概念是确定性效应。卡尼曼和特沃斯基的调查肯定了阿莱的谜题，确定性和仅仅是非常高的可能性之间（举例来说，类似100%和99%的概率区别）存在着主观上的鸿沟。这一发现也可以“反射”：非常不可能和保证不会发生之间（1%和0%的概率区别）同样有着巨大的心理区别。



损失厌恶“延伸到了道德直觉的领域，强加损失和未能分享收益所得到的评价完全不同”。

一方面是收益和损失，另一方面是可能和不可能，这期间涉及行为的4个领域。我们可以用一幅简单的四格图来概括它（如图6-2所示）。前景理论的四重模式不仅解释了阿莱悖论，还解释了为什么大赌徒会买保险这一类的怪事。

以阿莱的第一道谜题为例。你可以（a）稳得100万美元；或者（b）参与一场很诱人的冒险，你有1%的可能性两手空空地离场。不管怎么做，你几乎肯定可以获得100万美元甚至更多。换句话说，你处在一个很好的位置上，从诸多可能的收益中做出选择。此时，你处在图6-2中左上角的单元格。

这一格的特点是“行为厌恶风险”。你可能觉得100万美元伸手可得——只要选择（a）就行了。要是你选择（b）却输了，你会把肠子都悔青的。这样一来，（b）的风险完全无法叫人接受。

阿莱的第二道谜题是，选择有11%的机会赢100万美元，还是选择有10%的机会赢250万美元。这仍然属于收益，唯一的重要区别是获胜的可能性大大降低了。你会告诉自己，别太兴奋——你恐怕不会赢。这就改变了心理，触发了冒险意愿，如图6-2右上角的单元格所示。你愿意赌一把拿更高的奖金，1%的概率区别似乎并不那么重要了。

	可能发生的事件	不大可能发生的事件
收益	行为厌恶风险 （“一鸟在手胜过双鸟在林” ——投资债券，不投股票；接受一个过得去的报价，而不是待价而沽，等待更好的出价）	行为追求风险 （“不入虎穴，焉得虎子” ——购买彩票；想加入职业球员大联盟）
损失	行为追求风险 （“孤注一掷”——赌博时压下双倍的赌注，要么输双倍，要么翻盘；设局骗人，受害者已经投下了太多赌本，退出的话会一无所有，索性跟到底）	行为厌恶风险 （“先求稳妥好过事后后悔” ——购买保险；使用安全带；不吃生鱼片，以免沾染寄生虫）

图6-2 收益和损失，可能和不可能所涉及的4个领域

从收益翻个面到损失，行为的类型也随之改变。要是很可能遭受损失，鲁莽的赌博变得可以接受了（如图6-2左下单元格）。赛马日快结束时，赌徒们愿意“继续花钱打水漂”，希望借此挽回损失。而要是存在概率较小的损失可能性（如图6-2右下单元格），人们愿意买保险来对抗损失。

金融顾问告诉客户，在做金钱决定时要考虑风险承受能力。麻烦的是，行为的这4个领域存在于我们所有人身上。在这个环境中厌恶风险的人，换个环境就成了冒进家。只要换个基准点就行了。

投资者认为债券“安全”，股票则有风险，尽管它能提供更高的平均回报率。由于两种投资都承诺能带来收益，许多投资者厌恶风险（如图6-2左上单元格），在投资组合中选择债券。但还有其他的看待方式。要是你把通货膨胀和税收因素考虑在内，债券的实际收益可能是0，甚至是负数。“把钱买成债券，你保准会损失购买力！”对指望卖股票出去的人，这可是一个非常拿得出手的论点！

房地产泡沫破裂的时候，卖家们还记得当初市场高峰时自己可以大捞一笔的。这成了基准点，按当前市价卖出房子就变成了“损失”（如图6-2左下单元格），而不是接受一个合理的当前市场提供的产品。要卖家们调整基准点以适应新的现实，可能要用好几年的时间。这一期间，交易成交的很少。

卡尼曼说，他相信厌恶损失的概念是自己和特沃斯基对决定论做出的最大贡献。对这一点的基本认识其实已经存在好长一段时间了。埃德

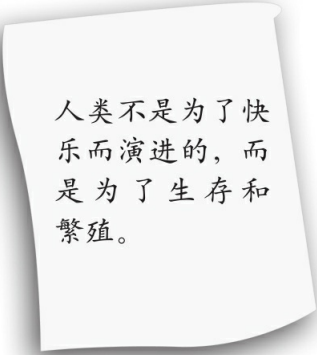
蒙·伯克（Edmund Burke）在1757年的《关于我们崇高与美观念之根源的哲学探讨》（*A Philosophical Enquiry into the Origin of Our Ideas on the Sublime and Beautiful*）一书中写道：“痛苦的观念总比那些触及愉悦部分的观念要强大得多，我很满意。”卡尼曼和特沃斯基的贡献，在于实现了科学的缜密性，并将这一概念覆盖到了前所未有的广大范围。哈佛大学的马克斯·巴泽曼（Max Bazerman）说：“前景理论的要点不难用文字说明，但（他们）增加了数学，提高了接受度，这很重要。”特沃斯基自学了数学，为前景理论做了完整的数学处理，以便让它获得经济学家的严肃对待。

他们在《计量经济学》杂志上发表了这一理论，该杂志差不多是所有经济学期刊里难度最大的。经济学家一贯回避解释人类的不合理性，就好像鸭子把头埋在水里一样。他们甚至只用一个词来表达自己的不屑一顾——“心理学”。言下之意是，心理学并不是什么很严肃或重要的课题。“前景理论”做了大量工作来改变这种心态。有人估计，截至1998年，它已经成了《计量经济学》刊载的文章中被引用次数最多的一篇。

2009年，由于金融风暴带来的冲击，德国亿万富翁阿道夫·默克尔（Adolf Merckle）撞火车自杀。但他的净资产仍有10亿美元之巨。

传统的经济理论涉及的是财富的绝对状态。10亿美元就是10亿美元，你应该高兴才对。人类的现实是，失去一半财产的亿万富翁会感到穷困潦倒，走投无路，中了5 000美元的彩票得主却会兴高采烈。一切无非对比罢了。

为什么损失带来的痛苦甚于收益带来的愉悦，这仍是一个尚未得到解决的问题。为什么牌面对我们不利呢？卡尼曼和特沃斯基的论文发表之后的岁月里，进化学解释冒出了头，日益流行开来。“人类不是为了快乐而演进的，而是为了生存和繁殖。”科林·卡默勒（Colin Camerer）、鲁温斯坦和普雷莱克写道。请想象寒冬里有一只快要饿死的动物。出去寻找食物有风险，因为这会让自己暴露在天敌之下。可待在洞里，表面上看起来安全，最终结果却是慢慢饿死。故此，动物赌上一把，出去找食物是合理的。夏天，这只动物有着丰富的食物来源，它的策略也会相应改变。它应该不会冒丢掉性命的风险去寻找它不需要的浆果。



人类不是为了快乐而演进的，而是为了生存和繁殖。

把“食物”换成“金钱”或其他任何的收益，你就得到了前景理论。从我们的行动来看，打扑克输掉500块就像是一个生死攸关的问题。卡默勒认为，损失厌恶是一种不理智的恐惧形式，就跟恐高症患者从高楼层的窗户望出去差不多。

卡默勒写道：“人们最害怕的许多损失并没有生命危险，但对过度适应传输恐惧信号的情绪系统来说，这不具说服力。把损失厌恶想成是恐惧，也意味着存在这样的可能性：诱发的情绪可以摆布买卖价格。”

第7章 最后通牒博弈：讨价还价中的心理状态

五金店的雪铲卖15美元一把。一场暴风雪之后的早晨，店家把价格提高到了20美元。82%的被试说这不公平。“供需原则”不是涨价的借口。

某商店的卷心菜娃娃一个月前就卖光了。圣诞节前一个星期，有人在商店储藏室里发现唯一一个剩余的娃娃。商店经理知道，好多客户都想买这个玩具。他们通过商店的广播系统宣布，这个玩具将通过拍卖售出，出价最高者得。74%的公众认为这种做法不公平。

某足球队正为一场大赛出售门票，座位有限。球队有三个选择：拍卖门票；随机抽选球迷，允许其买票；排队，按先到先得的原则出售门票。绝大多数被试认为排队的办法最公平。在大多数人眼里，拍卖是最不公平的。

某社区“美味红”苹果严重短缺，没有一家杂货店或超市的货架上有这种苹果，但其他种类的苹果多得是。一家商店以常规批发成本收到了一批“美味红”苹果，并按比通常零售价高25%的价格出售。这个问题赤裸裸地点出了“价格欺骗”的奇怪禁忌。25%的价格涨幅，低于产品一般的季节性波动。孩子们不会因为缺了“美味红”苹果就饿肚子——他们还可以吃“澳洲青”苹果。然而，仍有63%的人认为“美味红”提价的做法不公平。

公众的态度足够现实，他们明白，有时价格不得不涨。成本提高了，商店将它转嫁给消费者，没问题；公司亏损了，削减员工的工资，没问题；但利用市场力量占便宜（比如，在短缺的时候提高现有存货的价格），这不可以。公平的关键原则似乎是“别通过对我敲竹杠来提高你的利润”。

· 如果这样说：

某公司赚取的利润十分微薄。它所在的社区，正经历经济衰退，虽然尚未通货膨胀，但失业人口庞大。许多人都急于去该公司工作。公司决定今年减薪7%。62%的人

认为这不公平。

· 如果这样说：

该社区失业人口众多，通货膨胀率达12%.....该公司决定今年只提薪5%。78%的人说可以接受。

但事实上，在上述两个版本中，工人所得的结果几乎是一样的。其他商品价格提高12%，却只加薪5%，这跟购买力减少7%没什么两样。

由此得出一个结论：通货膨胀是滑头老板最好的朋友。同样的原则也适用于奖金。大多数人的判断，一家陷入困境的公司可以不发10%的例行奖金，但减薪10%就不妥。（碰到不稳定的市场行情时，华尔街上的东家们一贯利用这一点。）

卡尼曼、尼齐（Knetch）和泰勒写道：

按照传统的经济分析，商品的额外需求理所当然地是供货商涨价的机会，市场上也确实存在这样的涨价。在这种观点看来，市场上充斥着以逐利为目的进行的价格调整，自然得就像是水总往低处流一样——它是无关道德的。普通公众却无法秉持这种漠然的立场.....人认为是公平的行为，跟人期待市场上存在的行为，两者之间的区别其实很小。

令人惊讶的地方在于，公平的民间原则大多是为个人利益服务的。不管政治立场偏向如何，哲学家们总觉得有必要保持逻辑上的一贯性。公众不存在这样的自我约束。压倒多数的公众拒绝自由放任资本主义的财产观和自由企业观，但同样也拒绝工人权利或公共利益等前后一致的概念。公众充分表现出了安·兰德（Ayn Rand）[\(13\)](#)笔下的自私态度：它判断自由市场不公平，因为自由市场很可能做出不利于他们私利的事情来。

想象一下，在末日后的未来，美国文化除了几部法拉利兄弟（Farrelly Brothers）[\(14\)](#)拍的电影，什么也没留下来。早期的罗马文学基本上就遭遇了这一幕：什么都没了，除了古罗马喜剧作家普劳图斯的低俗闹剧。多亏了它的意外留存，我们看到西方世界对讨价还价最早的描述竟然挺滑稽。它出现在普劳图斯《绳子》（*The Rope*）一剧的关键场景。一个名叫格里普斯的奴隶，在海里无意发现了一箱金子，想靠它买回自由。格里普斯在路上碰到了特拉察里奥，后者认出金子是某大皮条客的财产，觉得这是勒索的好机会。

特拉察里奥：好吧，好吧，听着。我看到一个强盗在抢东西，我认识他抢劫的人，我走过去，给他开了个价，“我认识你抢劫的那人，”我说，“你给我一半，我就当不知道这事儿。”他不听我的。好啦，我问你，要一半够不够公平？

格里普斯：你还该多要点儿呢！要是他不给你，我觉得你应该去告诉失主。

特拉察里奥：谢谢，我会的。说到这儿，你跑出来了。

格里普斯：什么意思？

特拉察里奥：你手里有口箱子。我知道它属于谁。我认识他好长时间了。

用现代术语来说，这就是“最后通牒博弈”。一个人（格里普斯）手里有些战利品，另一个人（特拉察里奥）有办法把它给弄没了。这是否让后者获得了分享权呢？在普劳图斯的故事里，这一点毫无疑问。特拉察里奥威胁道，除非把金子分他一半，不然，就去告诉合法所有者。这样的话，两个人都什么也得不到。格里普斯怒道：“你能分到的只有麻烦，我保证。”他发誓，一定要让特拉察里奥落个竹篮打水一场空。

为了比喻人类处境的荒诞性，普劳图斯发现，他只需要两个演员，外加一连串的最后通牒就够了。按格里普斯的说法，箱子落在了渔网里，自己用一根绳子把它捞了起来（这出剧目的名字便打这里来）。观众们一定也看出两名奴隶的言语滑稽戏逐渐激化成了针锋相对的角力。永恒的信息蕴含于此：“讨价还价”不过是文雅一点儿的勒索，逻辑和最终结果没什么关系。

卡尼曼、尼齐和泰勒在芝加哥大学的学术会议上公布了他们关于公正的研究。他们的谈话刊登在1986年的《商业周刊》上，文章里涉及的邪恶小实验，就是如今众所周知的最后通牒博弈。

你得到10美元可以分给陌生人，你来提议如何分配这笔钱——比如“我6美元，对方4美元”。关键是，由对方决定是接受你的分配方案，还是拒绝。只要他接受，钱就完全照你说的分。可要是他拒绝，你也得不到一分钱。正如博弈的名称所示，你要么接受，要么走人，没有任何还价的余地。

你不承担任何“公平”的义务。只要你觉得自己有办法搞定对方，你要求独吞10美元也没问题。当然，你会想停在某个点之前，免得合作伙伴感到过分“不公平”，彻底否决这笔交易。

在往下读之前，你可以想一想自己会怎么玩这个游戏。

首先，假装你是负责分钱的人（下称“提议者”或“分配者”）。你会把10美元奖金的多少分给一个完全陌生的人呢？（你将永远不会知道这个人的身份，他也不知道你是谁。）请把这个数字写下来。

我把**10**美元分出去____美元

接下来，你是对方那个人，“响应者”。既然你是一个人在玩，有必要确定一下你将如何响应可能碰到的每一个提议。这些提议说不定跨度颇大，0美元到10美元都有可能。为简单起见，提议者大多提的金额都是整数。圈出你愿意接受的最小数额（凡等于或大于该数的提议你都愿意接受，低于此数的则拒绝）。

我愿接受

\$0 \$1 \$2 \$3 \$4 \$5 \$6 \$7 \$8 \$9 \$10

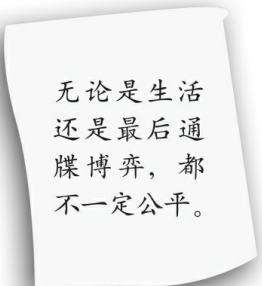
对一个追求最大化的理性人，最后通牒博弈简直不费脑子。响应者绝不应该拒绝“白来的钱”。再少他都应该接受，而不是行使否决权。反过来，理性的提议者应该料到这一点，并提出一个象征性的微小数额，信心满满地等着对方接受。

实际情况并非如此。泰勒在康奈尔大学让学生进行这个游戏时发现，提议者最常见的分配方案是“公平”地对半分。他还发现，分得太吝啬，响应者宁肯拒绝。平均来说，响应者愿意接受3美元，而拒绝2美元。

这是怎么回事不难理解。提议者有足够的情商，知道得给响应者合适的份额，好让他们满意。所有人都想得到对半分是“公平”的。这就使得对半分成了常例，康奈尔大学的大多数学生也正是这么做的。

问题在于，无论是生活还是最后通牒博弈，都不一定公平。两名被试有着不同的选择，也有着不同的权力。除非响应者不爽得想割了自己的喉咙，否则提议者有权利也有动机在平分的基础上稍微多给自己留些。为什么不建议四六分、三七分……甚至是一九分呢？

你看得出来这会走向何方。对任何响应者来说，总有一个限度会让他火冒三丈，行使否决权。贪婪而谨慎的提议者会希望尽量靠近这个限度而又不超过它。那么这个限度到底在哪儿呢？这就是最后通牒博弈所要提出的问题。



无论是生活
还是最后通
牒博弈，都
不一定公平。

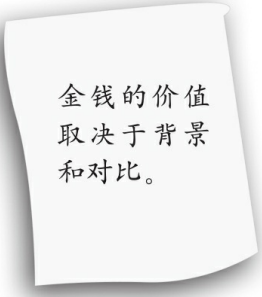
你很容易从自己的生活里辨识出最后通牒博弈的回声。提出无理要求的人会成功，因为其他人只不过是叹了口气，就纵容他们行事了——除非到了某个限度。最后通牒博弈探讨了一种并非不合理的焦虑：公平交易能让我们在这世上走多远？为了弄清这一点，它创建了一个模糊的道德空间。提议者并没有做什么理应得到10美元的事。响应者也没有做任何有权获得分享的事。博弈剥离了所有传统的社会道德、法律、财富权利，把任何社会都要面对的不平等问题赤裸裸地摊在了桌面上。

从某种意义来说，最后通牒博弈就是史蒂文斯“黑就是白”课堂示范的货币版本。金钱的价值取决于背景和对比。什么也没做就得到100美元，你感觉如何？大概感觉很不错。在一笔1 000美元的意外之财里，你只分到了100美元，你的“伙伴”单方面决定自己留下900美元，你又会感觉如何？肯定不怎么好。和900美元比起来，100美元渺小到了侮辱人的地步，尽管换个背景看它还是挺不错的一笔小财富。对比引发了情绪，情绪影响了行动。总会有人利用优势，因为他们觉得自己能蒙混过去，其他人却发现，自己唯一的筹码就是玉石俱焚的否决权。在这一意义上，我们所有人都在参与最后通牒博弈。

其实，1977—1978年，德国博弈理论家维尔纳·居特（Werner Güth）和同事罗尔夫·施密特伯格（Rolf Schmittberger）、贝思德·施瓦茨（Bernd Schwarze）就首次尝试了最后通牒博弈实验。

居特说，他丝毫无意证明人类行为跟经济学家的假设不是一回事。“那不过是痛打落水狗、再踏上三脚罢了。”他想做的是，设计“最简单的双人版最后通牒博弈”，看看真正的人会怎么玩它。

他想出了两种博弈，称其一为“复杂博弈”，其二为“简单博弈”。



金钱的价值
取决于背景
和对比。

价格实验

在前一种博弈中，被试要把一些黑白色的筹码分成两摞，让另一名被试为自己挑选一摞。复杂的地方是，对前一名被试来说，每一枚筹码值2马克(15)；可对后一名被试来说，白色的筹码仅仅值1马克。科隆大学的学生们不怎么擅长找出最佳分配方案。

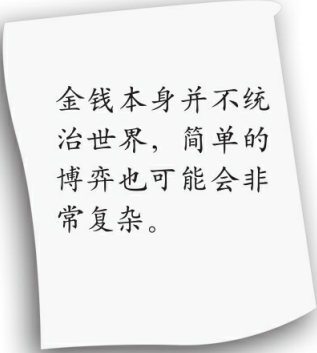
故此，居特又尝试了“简单博弈”，亦即现在说的“最后通牒博弈”。第一次实验时，42名经济系学生两两配对。在每一对学生中，一人负责分配4~10马克不等的现金奖励。分配提议转达给他的合作伙伴，这个人只有两种选择：答应或拒绝。最常见的提议是五五分配，21名提议者中有7人这样做。据居特说，经济学家们听到这样的研究结果，最爱说的是：“这些科隆大学的学生难道是蠢货吗？”

由于大部分提议者的分配方案都很接近对半分，所以居特率领的德国小组没有太多机会观察响应者面对不公平分配的反应。后来卡尼曼、尼齐和泰勒对响应者更感兴趣。“我们对公平的所有问题都跟这个有关，‘你认为那家伙，那个有权力的家伙，做得公平吗？’”卡尼曼解释说，“身为心理学家，我喜欢这么想：人们渴望公平。但泰勒是个经济学家，他知道得把响应者当成关键。”

故此，他们折磨响应者，非要他们说出愿意接受什么样的提议。这牵涉到一系列的“是”或“否”问题。（“如果对方提议给你0.50美元，你接受还是拒绝？”）这种方法叫作最后通牒博弈的询问完整策略，现已得到广泛使用。实际上，它揭示了响应者的保留价格。

结果跟半个地球之外的居特类似。平分是最普遍的提议，分配数额平均为4.50美元上下。对低于2.30美元的提议，响应者宁肯拒绝。响应者行使否决权，是该博弈对经济理论提出的最大挑战。“那是不满，是不顾一切也要惩罚的意愿，就这么回事。”卡尼曼解释说。

被试行使否决权，意味着他抛开了“白拿钱”的逻辑，基于情绪做出了经济决定。违背经济理论行事的并非只是一两名被试，几乎人人都会这么做。“事情真正怪的地方在于，”卡尼曼说，“这套理论屹立了几百年，没人挑战它，直到有人说：‘嘿，看那，国王没穿衣服！’反面例子微不足道。”



金钱本身并不统治世界，简单的博弈也可能会非常复杂。

“最后通牒博弈是终极实验吗？”2007年，尤拉姆·哈勒维（Yoram Halevy）和迈克尔·彼得斯（Michael Peters）在论文的标题中发问。他们半开玩笑地指出，这一博弈已经成了一门学术产业。据称，最后通牒博弈是当今重复次数最频繁的人类实验之一。教授常把它布置给心理学和经济学的研究生，训练他们招募被试，签署同意书，完成卡方检验。但它盛行多年的主要原因还在于，人们相信该博弈揭露了价格和讨价还价中的诸多心理状态。

博弈意味着什么，为什么我们应该重视它呢？正如居特现在的看法，该博弈蕴含着两条信息：金钱本身并不统治世界，简单的博弈也可能会非常复杂。卡尼曼把这一博弈视为确立如下观点的里程碑：哪怕是极为简单的经济决定，心理状态都是关键。“必须来点特别的东西，才能让经济学家们注意，”他解释说，“最后通牒博弈有这样的效果。”

经济学家需要关注的原因之一是，它跟价格设置存在明显的相似之处。10美元可以代表销售的潜在利润（“盈余”）。分钱的人是“卖家”，响应者是潜在的“买家”。卖家可以选择自己保留所有的利润（定一个高价），或是把所有的利润都拱手让给买家（“按成本卖”），又或是和买家分享利润。买家来决定是接受这一价格，还是因为它太高而拒绝。这个博弈还可以看成是谈判准模型。

莱缪尔·博尔韦尔（Lemuel Boulware）是20世纪50年代通用电气的劳资谈判代表，名声恶劣得很，因为他一贯把工资方案往桌上一扔，之后便拒绝让步。谈判不是通用电气的政策。博尔韦尔出的价，一般是经过大量研究才选出来的。该方案显然是工会领袖能够接受的底线——哪怕是咬牙切齿才接受的。博尔韦尔（以及许多试图效仿他的劳资谈判代表）其实相当于最后通牒博弈中的策略性提议者一角。

现实生活中更常见的一来一回地讨价还价，则可以视为一系列的最后通牒博弈。买卖房子的出价就采用了最后通牒的结构：这个报价，必

须在星期二下午6点前接受，否则它就作废。除非你接受最新的报价，不然对方说不定会退出。

讨价还价往往是社会认可的礼貌性仪式。我降低我的报价，你把你的出价往上抬一抬。我们在中间的某个地方达成一致。有时，假装的“最后通牒”也属于仪式的一部分。“这是我最后的出价了，你要么接受，要么就算了。我反正是不能再让步了……说真的，我走了……”双方恐怕都知道情况没那么严重。

谈判的关键是如何应对强硬的交易者提出过分倾斜的要求。最后通牒博弈以最浓缩的形式，展现了谈判的真正难点。不管参与谈判的是一个还是多个强硬的交易者，真正关键的一刻总会出现，所有假动作、虚张声势和预留的讨价还价余地都扔到了一边，只留下一道最后通牒。那时，你会怎么做呢——忍气吞声地受利用，还是把钱留在桌上走人？

价格实验

纽约州州长纳尔森·洛克菲勒（Nelson Rockefeller）在第五大道上有一处绝佳的公寓，能够尽览中央公园的景色。可他也碰到了一个难题。在房子的西边，市政府做了公共住房项目的规划，要修建一幢摩天大楼。这样的话，大楼就把洛克菲勒欣赏日落美景的视线给挡住了。住房项目的发起人是米德·埃斯波西托（Meade Esposito），民主党的幕后大老板。洛克菲勒邀请埃斯波西托来自己的公寓，绅士般地讨论这个问题。“要是你停建那幢大楼，”洛克菲勒说，“我就把那幅毕加索画作给你。”

他指了指墙上挂的那幅现代主义油画。埃斯波西托答应尽力而为。摩天大楼没能修起来，洛克菲勒也履行了自己的诺言。埃斯波西托得到了毕加索的画作，洛克菲勒得到了一个余生里一讲再讲的好故事。之后的岁月里，洛克菲勒总是充满感情地回忆那次“馈赠”的每一个细节，并在结束时揭开谜底：“那只是一幅复制品，印刷的！”

谈判可不是一幅漂亮的图片。大部分时候，熟练的谈判者就是最能误导价值的人。驱动价格心理的公平，并不像表面上显得那么“公平”。（本着新闻公平的原则，我必须补充说明，不是只有洛克菲勒是坏人。埃斯波西托或许不怎么懂艺术，但他很清楚交易的艺术。1987年，他因散布谣言影响交易，被罚50万美元，获缓刑2年。）

最后通牒实验的一些早期评论提到了利他主义。提议者并不为难响应者。从统计上来看，他们通常会开出比对方底线稍微高一点的条件，以便让响应者答应交易（尽管是咬牙切齿地答应的）。故此，该博弈展现了人高贵慷慨的内在本性。在某些专题报道里，你依然能看到这类解释。说来悲哀，这美好的概念，几乎被后来的研究撕了个粉碎。交易者其实并不如旁人想的那么在乎公平。

卡尼曼、尼齐和泰勒在其探讨博弈的第一篇文章里就着手研究利他问题。他们设计了现在所说的“独裁者博弈”。

价格实验

他们给康奈尔大学心理学系的学生每人20美元，让学生们分给未知的陌生人（被试）。钱怎么分配，全听提议者（“独裁者”）指派，被试完全没有发言权。

在第一轮实验中，“独裁者”按规定只有两种选择。他们可以贪心地自己留18美元（把2美元给被试），或是公平地对半分。76%的人者选择了对半分。

卡尼曼的研究小组认为这似乎是含糊地“抗拒不公平”的结果。“独裁者”避免不公平。利他主义可以算是一个解释，但并不是唯一的解释。

研究人员在另一博弈“利他惩罚”中进一步探索这种抵抗态度。他们又安排了一轮“独裁者”博弈，并给新的被试如下选择：

（a）他们可以与上一轮“独裁者”博弈实验中的一位被试平均分享12美元。但那位被试前一轮是个“贪婪”的“独裁者”（给自己留了18美元，只给了搭档2美元）；

（b）他们可以与“独裁者”博弈中一位不同的被试均分10美元。这名被试是“公平”的（选择了对半分）。

绝大多数人选了（b）。他们宁愿自己少得1美元，来“惩罚”一个对自己没有做过任何错事的人，全因为知道那人是个“不公平”的被试。

到目前为止，上述结果听起来还挺让人欣慰的。“独裁者”大多是公

平的，不公平的人遭了报应。可惜康奈尔大学的原始“独裁者”博弈有其限制。它只有两个可选项，一种做法是平分，另一种做法则极度贪婪（独吞奖金的90%）。如此的贪婪，看得顺眼的人不多。自那以后，其他研究人员重新做了实验，允许“独裁者”拥有更多的分配方式。这些实验大多发现，“独裁者”并没那么慷慨。要是可以自由地进行分派，他们平均分给弱勢的搭档30%左右。大约1/5的“独裁者”一毛钱都不给。

伊丽莎白·霍夫曼（Elizabeth Hoffman）和她亚利桑那大学的同事们做了一轮权威性的独裁者博弈实验。霍夫曼怀疑，“独裁者”表现慷慨只是因为知道有人在看。实验者往往是被试的老师，过几个月，老师可会给学生打分咧！难道为了几美元，就值得让教师知道你是个贪婪的混账东西吗？

故此，霍夫曼的研究小组煞费苦心地想办法，保证没人会知道被试具体会如何行动。

价格实验

研究人员交给每名“独裁者”一个白信封，并让他们到房间后面去。那儿有个纸箱子，“独裁者”把信封在箱子里打开，免得有人偷窥。

大多数信封里装着10张一美元的钞票，以及10张裁成一美元大小的空白纸条。“独裁者”可以任意拿走若干张钞票，把信封里剩下的东西留给搭档。此时，他需拿走相应张数的空白纸条，好让信封里只留下10张纸（钞票加纸条的总数）。做完之后，他把信封交给监督人（此人不是被试，他无法从信封的重量或感觉上推断任何事情）。监督人带着信封去另一个房间，交给“独裁者”的搭档。

上述设计最关键的地方是，所有被试事先都知道，有些信封一开始就没装钞票，只有20张纸条。得到这种信封的倒霉蛋必须抽走10张纸条，把另外10张留给搭档。这样一来，哪怕是搭档一分钱没分到，也无法断定自己被故意“独裁”了。

基于这样的条件，约有60%的“独裁者”自己独吞了10张钞票，只在信封里留下了空白纸条。

没什么好吃惊的（美国国税局的查税员们绝不会吃惊）。“贪

婪”和“慷慨”这类概念总是依赖于参照框架的。就在这一刻，你便有机会跟一家声誉卓著的慈善机构，比如“无国界医生”，分享你钱包里的钱。你真的应该捐点儿钱……但要是你把钱全自己留着，那也再明智不过了。没人会知道你钻了空子，什么也没捐。

对霍夫曼实验做个悲观的阐释：它表明人是何等伪善。要是没人看（当然，也只有在没人看的时候），被试几乎全像经济学家们假设的那么自私。卡默勒和泰勒则提出了另一种解释：最后通牒博弈和独裁者博弈的结果跟利他主义的关系不大，但跟态度很有关系。公平竞争的社会规范没那么容易摆脱。就算“伪善”，也不一定是坏事。有时，不管你最初怀着什么样的心机和目的，单单是假装做个更好的人，你最终都会真的表现为一个更好的人。

最后通牒博弈演变成了所谓的“ur-experiment”，相当于科学里的即兴弹奏，本身可以进行无尽的取样和混合。它在全球各种文化里被人们实践过：儿童做过，自闭症患者做过，高智商人士做过，雄性激素超高的男性也做过，有些被试还注射过一种提高对陌生人信任感的激素；甚至，黑猩猩也做过，它们要分配10枚葡萄干的奖励。该博弈持久的魅力来自行为将如何随背景变化或不变。实验设计的经济选择就像一支上好了油的风向标，总能敏感地察觉影响我们的微妙压力。这些压力，我们平常很难注意得到。

许多简单的变量极大地影响着人的行为。霍夫曼在亚利桑那大学的小组做了一连串的尝试，只要被试答对一道小小的智力问答，就能赢得当“提议者”的权利。提议者的慷慨度由此降低。他们显然觉得自己的特权地位是应得的，所以有权趾高气扬。而且，响应者似乎也答应。要是提议者的地位是当事人光明正大赢回来的，响应者就愿意从提议者那儿少拿点。大多数提议者建议给3美元或4美元，从来没有人拒绝这样的出价。

伊丽莎白·霍夫曼的小组还尝试把该博弈设置成零售交易的样子。提议者叫成“卖家”，响应者是“买家”，买家决定是否接受卖家的价格。买卖双方都收到一份表单，说明各种价格能带给他们怎样的好处。相应的报酬跟标准最后通牒游戏一样。

对理性行为者来说，这应该不存在什么区别，但区别可大了。卖家变得更贪婪了，一般只给买家分派3~4美元。然而，后者大多都会买。显然，被试认为，卖家有权设定价格。在人们的判断里，较之标准模式

中的不公平分派，高价格不怎么值得惩罚。

霍夫曼实验最有趣的一点发现是，提议者和响应者的反应基本上是同步的。面对博弈的新改款，提议者能立刻感觉出自己应该给多还是给少，响应者亦随之调整自己的期待。此时，双方并未进行沟通。

在最后通牒博弈中，经济学家普利什·塔尔沃克（Presh Talwalkar）挖苦说：“我的以色列博弈论教授自豪地指出，以色列是世界上少有的几个开出低报价并得到接受的地方。”

价格实验

确认“以色列神话”的是1991年在匹兹堡、卢布尔雅那、耶路撒冷和东京进行的一次比较行为研究。以色列提议者最常见的分派比例是四六分，美国人则为五五分。其实区别真的不算大，真的（等会儿我们就能看到）。但它禁不住叫人联想到以色列人果然是神选出来的，够理性——要不然，就是像《威尼斯商人》中的夏洛克那样的奸商。四城研究的一位合作者——希伯来大学的什穆埃尔·扎米尔（Shmuel Zamir）回忆说，来找他的有个年轻的以色列人“明显气呼呼的”。他抱怨说：“我没有赚到一毛钱，全因为对方的玩家们是蠢蛋！你怎么能够拒绝一笔白白得来的钱，却宁肯选择零呢？他们根本没明白这游戏！您应当停止实验，给他们做做解释。”

在人类学家看来，上述4个城市都是一种同质化全球文化的一部分罢了。等到博伊德的研究生乔·海因里希（Joe Heinrich）对来自秘鲁东部的马奇根加（Matsigenka）人做了最后通牒博弈之后，故事变得更有意思了。

“他回来之后问我，你能不能来看看我的数据？”卡默勒回忆道，“于是我去了加利福尼亚大学洛杉矶分校，乔说，‘我想我是哪里弄错了，因为他们开了许多低价，而且都被接受了。只有一个拒绝的，但那更加可疑，因为我是个说西班牙语的助手，他还能说当地方言，那家伙有点儿吓唬人对被试说，我想你不该接受这样的东西。所以，他们其实都算是接受了。’”

海因里希的发现极具讽刺意味。最终，在秘鲁的内陆地区，他发现有人完全像传统经济学家假设的那样做事。可这些人完全没什么经济可言。

他们发现，凡有市场经济的地方，均可见到欧洲或北美洲大学生表

现出的那种行为。它不要求工业化。肯尼亚的奥勒莫（Orma）人靠贩牛为生。他们的平均出价是44%，与西方文化持平。不管非洲牛贩子和美国股票交易员有多大区别，两种文化都会奖励做成最好买卖的成员。这就意味着要给出一个足够公平、能被接受的价格，并在看到敲竹杠的时候明白过来。

而在相对孤立的小范围文化中，博弈行为有着很大不同。相邻文化（也就是说这些民族在基因上有可能很接近）中的博弈也会存在极大的区别。这证明，最后通牒博弈是一种文化上的X射线（卡默勒这么说），它是理解社会如何对待经济不平等的一种方式。

许多非市场文化建立在精心构建的社会合作守则之上。对印度尼西亚拉马莱拉（Lamalera）捕鲸人和巴拉圭东部的阿切猎手来说，他们最尊重打猎中贡献最大、分肉最慷慨的成员。这些人在玩最后通牒博弈时“超公平”。提议者会把50%以上的奖金分给响应者。

巴布亚新几内亚的奥族人和格努族人也是“超公平”的，对50%以上的分派提议，响应者一般会拒绝。在这两种文化中，礼品和好意都是附加了条件的。它们带来了回报的义务，而大多数人宁肯不承受这种负担。“给太多的钱，并不是极为慷慨，反而是一种吝啬。”“亚当·斯密有句名言，‘我们的晚餐，不是出自屠户、酿酒师或面包师的恩惠，而是出于他们自身利益的需要。’”这就引出了一种说法：倘若人们只顾着自己，市场会兴旺发达。本次研究透露出的信息是：人们从事大量交易的文化似乎具备公平分享的准则；而在成员不怎么从事交易的文化里，准则是给你多少你就拿多少，此外，我不指望你给我任何东西，所以哪怕是一点点也能让我满意。

最接近智人的物种不是传说中的经济人，而是类人猿黑猩猩。2007年，莱比锡马克思·普朗克考古人类学研究所的基思·詹森（Keith Jensen）、约瑟·考尔（Josep Call）和迈克尔·托马塞洛（Michael Tomasello）在研究中发现，黑猩猩比人类更自私（也就是更“理性”）。他们设计了一个巧妙的实验。

价格实验

相邻的两个笼子里关着两只猩猩，它们面前摆着一个小柜子，柜子里有两格抽屉。每格抽屉里装着两盘葡萄干，每只猩猩都可分得一盘。负责提议的猩猩要选择一格抽屉，扯动一条绳子，把它拉

到响应猩猩可以拿到的范围。响应猩猩要用一根伸出的棒子，把抽屉里的盘子勾到自己能拿到的地方。这样，两只猩猩都可以吃到各自盘子里的葡萄干。

在典型的设置里，一格抽屉的两盘葡萄干是对平均分的。另一格抽屉里提议猩猩的盘子里装着8枚葡萄干，响应猩猩的盘子里只有2枚。

75%的概率，提议黑猩猩会选择贪婪的分法。而它们卑微的合作伙伴，95%的概率都会纵容了事。它们接受了2枚葡萄干，而不是惩罚提议者。“这样看来，”基思·詹森的小组总结说，“.....跟人类血缘关系最近的动物是按照传统经济学家的自私模型行事的，它们跟人类不同，并不像人类那样对公平敏感。”

第8章 一掷千金：丰厚的奖金是否影响判断

1996年以后，主流经济学家甚至商人，大多都接受了行为决策理论。基本上，柯恩、吉仁泽等人提出的各种反对意见，锐减到了几乎看不见的程度。近年来，廉价奖品成了一个更迫切的关注点。

在美国，最后通牒游戏通常只提供10美元的奖金，都不够在曼哈顿买张电影票。然而，进行此类研究的心理学家和行为经济学家却相信它们的结果跟实验室以外广阔天地里的结果差不多。这样看的话，最后通牒博弈的响应者（事实上，也就意味着我们所有人）都应当特别在乎自己的份额跟提议者相比起来怎样，而对绝对的金額相对迟钝。

那么，来设想一个1 000万美元的博弈吧！

价格实验

提议者自己保留900万美元，只分给你区区100万美元。你会放弃自己的这100万美元，让他尝尝价值900万美元、永生都难以忘记的苦头吗？

大概不会。有鉴于此，提议者恐怕会给自己保留一个更大的比例.....

针对100万数量级的最后通牒博弈会有怎样的不同，研究者们已经做了大量的思考工作。部分经济学家认为，在奖金背后添上若干个零的话，人们的理性就会占上风。

伊丽莎白·霍夫曼、凯文·麦凯布（Kevin McCabe）和弗农·史密斯（Vernon Smith）听厌了这类说法。批评他们的经济学家甚至根本不必想到100万美元。有些人说，100美元的博弈就能面目全非。人们当然会拒绝一两美元，可没人会拒绝白白到手的10美元或者20美元。

霍夫曼和同事们凑钱进行了一些100美元的最后通牒博弈。这就是

说，至少得筹措5 000美元，才能多做几次博弈，让所得结果具有统计学意义。从这些在亚利桑那大学完成的实验来看，不管是100美元的博弈，还是10美元的博弈，在行为上它们并没有什么显著的差异。

实验里出现了这样一段小插曲。一个提议者在自己的开价单上笔迹潦草地给响应者留言（这么做是不合规矩的）：“别太贪多，这35美元，你挣得恐怕最轻松了。”这名提议者把自己的100美元三七开，只分出去“区区”30美元，另外5美元是所有被试都有的。响应者拒绝了这30美元，并在留言后面加上了自己的话：“贪婪正把这个国家带往地狱。你跟它一样，活该跟它一块儿下地狱。”

2002年，荷兰播出了一套名为《追逐百万》的电视游戏节目，一炮而红，从毛里求斯到美国的60多个国家都制作了本地版。美国版叫《一掷千金》（*Deal or No Deal*）。节目呈现的困境跟决策理论学家们研究的东西差不多，只是金钱数目要大得多，而且还全部兑现。2008年，蒂里·波斯特（Thierry Post）、马蒂·冯·登·阿森姆（Martijn van den Assem）、吉多·巴尔图森（Guido Baltussen）和泰勒指出，《一掷千金》这个节目“设计上不像是电视节目，倒更像是一出经济学实验”。

除了长腿的漂亮女模特，《一掷千金》里根本没有什么娱乐性的东西，而只有经济学实验。

价格实验

《一掷千金》里有26名女模特，每人拿着一个公文包，里头装着金额未知、数量不等的现金（1美分到100万美元）。参赛者先从26个公文包里挑选一个——“拥有”所选包里的一切。但主持人豪伊·曼德尔（Howie Mandel）并不立刻揭晓奖金，反而玩起了延时版的猫鼠游戏。他从参赛者没有挑中的公文包里随便选一个露底给观众看。通过如此这般逐一排除，参赛者选中的公文包里有什么也就间接地暴露了出来。所有的奖金金额都贴在一块记分牌上，参赛者和观众们都能看到，每排除掉一个金额，板子上也相应将其划去。

接着，一位“银行家”向参赛者提出一笔交易。“银行家”从俯瞰舞台的漆黑办公室里打来电话，提议以某个价格买下参赛者手里的公文包。玩家必须选择是接受银行家的报价，还是拒绝（留着公文包，继续参加游戏，承受一切可能的结果）。银行家最开始的出价

总是很低。倘若参赛者拒绝，主持人又打开更多的公文包，参赛者对公文包里到底有什么、没有什么也就更加心里有数了。银行家进一步开出价格。最终，等舞台上只剩最后两个公文包没打开的时候，参赛者还应当拒绝银行家的出价吗？银行家报上最终出价。如果参赛者拒绝了，他的公文包便当场打开，不管里头有什么，都归他所有。

所有公文包都不曾打开的时候，《一掷千金》中的26个包里，平均奖金为131 477.54美元。随着公文包的打开，这一平均数字开始变化。比如，一个没选中的公文包里装着100万美元，这对参赛者来说是个坏消息，他的前景相应黯淡下来。

节目唯一不透明的环节是银行家计算自己出价的方法。在开始的若干轮，他的出价只占预期数值的很小一部分，你要是接受那才叫疯了！可等游戏进行下去，出价相较于期待值就显得越来越慷慨了。最终出价几乎等于期待值（美国版），甚至略微偏上（其他国家的版本）。

波斯特和同事们找来了该节目的荷兰版、德国版和美国版好几年的录像带。他们辛苦地分析了这三个国家151名参赛者所做的每一个选择。举个例子：弗兰克是荷兰版节目里最倒霉的一名选手。一眨眼的工夫，弗兰克就拒绝了银行家高达75 000欧元的出价，这相当于他一整年的收入。弗兰克最终只得到了10欧元，勉强够买上一瓶烈酒。

就在弗兰克打开公文包之前，场上只剩下两笔奖金：10欧元和10 000欧元。银行家开出了高于弗兰克公文包的预期价值，比美国版节目里的出价要优厚一些。老妈、会计师，或者任何拿固定薪水的财务顾问都会告诉弗兰克：接受吧！可他偏不，他打定主意要拿到那10 000欧元。

任何假设选择都仅仅取决于最终财富水平的理论，解释不了这种行为。任何目睹了整个事件的人都明白可怜的弗兰克是怎么走上这一步的。他只是在跟先前落在自己头上的所有坏运气对着干罢了。和其他所有选手一样，弗兰克一开始是满怀希望，有着很高基准点的。荷兰版节目的最高奖金是500万欧元，比美国版节目多得多。弗兰克眼睁睁地看着三项最大笔的奖金在头两轮里就出局了两个。从这以后，他就觉得自己是个失败者。他没把银行家的报价看成是意外之财，而当成了损失（相较于他期待中的财富）。这令他愿意冒险碰碰运气。

1979年的展望理论论文探讨的困境和弗兰克并无不同。卡尼曼和特沃斯基报告了这一选择。

价格实验

除了你自己拥有的东西以外，你已经得到了1 000镑。现在，有人要你做出选择：是有50%的机会再多赢1 000镑，还是选择所谓的500镑“银行家出价”。卡尼曼和特沃斯基称，84%的被试说：“选后一个，成交。”较之赌博，他们更倾向于有把握的事情。然后，他们改写了问题，测试另一组的被试。你已经得到了2 000镑，但你必须在以下两种情况中做出选择：要么选50%的机会损失1 000镑，要么选“银行家出价”，100%地损失500镑。此时，69%的被试说，不干，还是赌运气吧！较之确定的损失，他们宁愿赌一把。

单从你离开时能带走多少钱这一点来看，前后两个版本的问题是等价的。后一种提法只是一开始就给你加上了额外1 000镑，然后再从中减去一部分，得出跟前一种提法里完全相等的结果。第二个问题的措辞则怂恿你把最初的2 000镑看成是基准点。接着，它又把选项框定成“损失”，鼓励你冒险。

弗兰克一连串的坏运气起到了相同的作用。银行家的最终出价被记入了亏损栏——虽说换一种更喜庆的背景，弗兰克会把它看成是一笔横财。这令他宁愿奋力一搏：要么翻本儿，要么输光。

波斯特的团队比较了预期效用模型和前景理论预测《一掷千金》参赛者们决策的成绩。他们发现，预期效用的正确率是76%，前景理论的为85%。当涉及庞大数额的时候，前景理论对行为的预测击败了效用理论。

《一掷千金》里的选择当然是在一时冲动下做出的。然而，波斯特和同事们猜想，节目里所作的决定，跟选择抵押贷款或退休投资组合时做出的决定一样是经过谨慎考虑的。就跟调动庞大资金的人一样，《一掷千金》的参赛者们会向亲密的家人和朋友寻求建议。大部分参赛者都是节目的铁杆粉丝，说不定在报名出场之前他们就早早规划过自己的策略。（好多人都觉得抵押贷款和投资枯燥无味，没到最后关头，他们根本不会去想。决定往后推了又推，最后才终于一时冲动地拿定主意。）

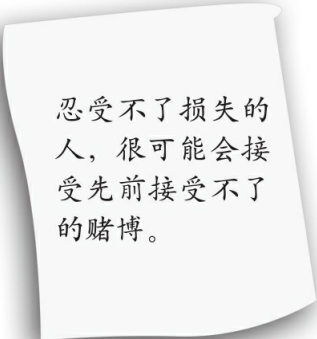
研究人员还在鹿特丹伊拉斯姆斯大学复制了两次《一掷千金》节目

的家庭版。

价格实验

他们“尽可能地在教室里”复制了该节目，找来主持人（一位受欢迎的讲师）和现场观众，更好地“制造参赛者在电视演播室里必然会碰到的那种压力感”。他们尽可能地原样照搬电视节目的脚本，使用相同的银行家报价，随机选择打开哪个公文包。这样的形式方便他们对比学生和电视参赛者的行为。不同的地方在于奖金的多少。在一个版本里，奖金是荷兰电视台里的万分之一，另一个是千分之一。后者意味着最高奖金是5 000欧元，平均能赢的金额是400欧元左右。这在所有的行为经济学实验里算得上是出手最阔绰的了。

如果把金钱看成一个量值量表，你大概能料到行为会一样——果然如此。参与奖金低的博弈实验的学生，跟参与奖金高10倍的电视参赛者做起事来一个样。两个小组的行为都类似电视参赛者，哪怕奖金多上1 000甚至10 000倍。参赛者判断银行家的出价是否“公平”都跟他过往的遭遇有着极大的关系。像弗兰克那样多次失望的被试，接受一个好价格的可能性较低。卡尼曼和特沃斯基曾写道：“忍受不了损失的人，很可能会接受先前接受不了的赌博。”这真像是在说《一掷千金》啊！



忍受不了损失的人，很可能会接受先前接受不了的赌博。



第二部分

魔术般的价格骗局

- 为什么重达 72 盎司的牛排会免费？
- 为什么同样一瓶啤酒，小卖部和旗舰店售价相同，而小卖部却被顾客认为是敲竹杠？
- 为什么一双 800 元的鞋子旁要摆一双 1 200 元的鞋子？

第9章 免费的72盎司牛排

美国历史最悠久的一出盛大闹剧，每天都在得克萨斯州的阿马里洛市（Amarillo）上演。就在刚下40号公路的地方，立着一块巨大的广告牌，上面写着：免费的72盎司牛排。这是“大得州牛排牧场”餐厅的招牌菜，内容包括色拉、基围虾、烤土豆、肉卷、黄油和一块超大牛肉。餐厅的要求是：你必须在一个小时里把所有东西全都吃完。要不然，你得掏72美元埋单。

在如今这个好打官司的时代，这样一笔交易当然会有些详细的规定条款。顾客必须预先支付72美元，吃完以后店方全额退还。第五条规则：“你不用吃肥厚的脂肪，但是否属于肥厚的脂肪则由我们判断。”任何第三者均不得碰触食物（难道是为了预防他们暗中偷了烤土豆吗）。此外，就餐者必须签署一份弃权书，声明所有健康风险由自己承担。点72盎司大牛排的人，其实成了进行餐厅例行表演的人：他们必须坐在特殊的舞台上，当着所有人的面吃，且吃饭期间不得离开桌子。还有，再提醒一句：要是有人吃吐了，就算自己还想继续吃下去，也没资格了。店方提供呕吐专用桶。

自1960年（当时的价格是9.95美元）以来，约有6万个食客参与过这个挑战。餐厅报告说，有8 500人全部吃完，整体成功率达到14%。勇于尝试的妇女并不多，但登台的妇女，50%都成功了。点72盎司牛排大餐的人大概觉得，不管怎么说，这都是笔划算的交易。每盎司牛肉才1美元，而且，跟那种“能吃多少吃多少”式的促销活动不一样，顾客可以把吃剩下的部分打包带回家。表面上看，你好像没啥损失，但好好想想：你在阿马里洛吃了一顿72美元的晚餐啊！

“免费”的72盎司牛排大餐，是民间智慧和专业定价的共同产物。1960年，大得州牛排牧场餐厅老板鲍勃·李（Bob Lee）独自想出了这个点子，那时可根本没什么点餐顾问。他的促销手法料中了如今学术界和营销专业人士都信奉的几条原则。最重要的一点是，72盎司牛排是一个锚点。一来到大得州牛排牧场餐厅，你必然会反反复复地听人提到、

自己看到“吃72盎司牛排”的事儿。虽然绝大多数顾客不会点这道菜，可它巧妙地提高了就餐者对自己食量的估计，也提高了他们的支付意愿。在这方面，卡尼曼做的一个锚定实验值得一提。他和凯伦·雅可维茨（Karen Jacowitz）提问道：

（a）普通美国人每年吃下的肉食，是多于还是少于50磅(16)？

（b）普通美国人每年吃多少磅肉？

答案的中间值为100磅。他们又问另一组被试：

普通美国人每年吃下的肉食是多于还是少于1 000磅？

结果，这组被试的估计中间值成了500磅。

大得州牛排牧场餐厅的促销，其实就是一个简单的非线性定价的例子。

你不可不知的价格术语

“非线性”定价
指价格（或单位重量的价格）不是一条直线，而是
■ 随着消费量的变化而变化的。

72盎司牛排售价72美元，可要是你吃完所有的东西，价格就直线下降为0。

这类定价手法施出了一记魔法咒语。它是价格顾问们最常用的一种手法，从话费账单到飞机票，样样东西上都有它的身影。大得州牛排牧场餐厅里饥肠辘辘的顾客并不知道自己是将要付出72美元，还是一分钱都不必出。这种不确定性大大削减了72美元的真实性。判断这类交易还有另一种方法：按每盎司的价格来看。在图9-1中，每盎司的价格曲线是一条一开始剧烈下斜的曲线，它慢慢靠近1美元，最终降到零。点了72盎司牛排的就餐者，要是只能吃下1盎司，就得付每盎司72美元的离谱价格。可要是有人能吃下2磅（32盎司）肉，那他就只需出每盎司2.25美元；几乎把整块牛排都吃完的人，所付则仅为每盎司1美元多一点点。这相当合理嘛！顾客的脑袋里想的全是进行一笔划算的交易，而实际上，那个划算的价格需要他撑爆肚子才换得来。

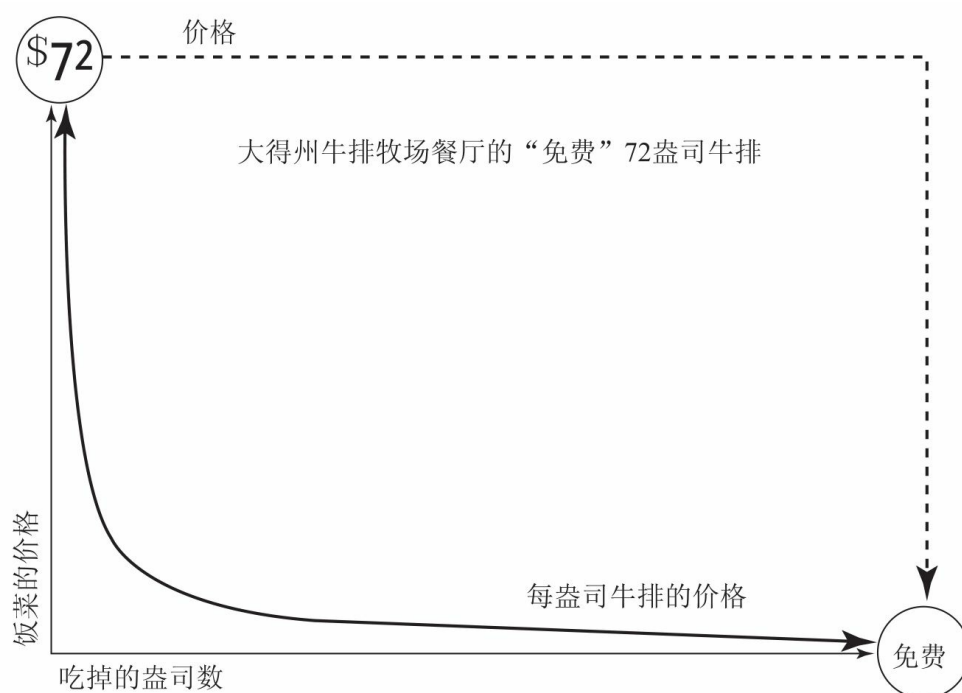


图9-1 牛排的价格曲线

赫尔曼·西蒙（Hermann Simon）是个怪人，买相机时，要是人家给他35%的折扣，他会烦躁不安（最近这一次就是这样）。他本来正高兴兴地买相机，可销售人员坚持要给他打6.5折，这违背了西蒙的经营理念。他写过一篇论文，名字便明明白白地说道：“务必充分利用（消费者的）支付意愿。”

这种命令式的语句，再搭上西蒙的日耳曼口音，第一次碰到难免会叫人大吃一惊。然而，西蒙“支付意愿”的魅力却极具传染性。过去二十多年里，他一直躲在定价这一专门职业的背后。20世纪80年代初，西蒙是比勒费尔德大学的企业管理和市场营销学教授，偶尔向企业提供咨询服务，只不过，他的建议往往没人理会。定价心理学逐渐成为热门话题有多方面原因。另外，卡尼曼和特沃斯基研究工作掀起的波澜层层向外蔓延，点燃了营销商和零售商的兴趣。来看看特沃斯基和卡尼曼在1981年发表的一个调查问题吧。

价格实验

假设你打算用125美元买件夹克，用15美元买台计算器。计算器销售员告诉你，你想买的那种型号在该商号的另一家分店10美元就能买到，但你得开车20分钟过去。你会去那家分店吗？

大多数被试表示他们会去。

另一组随机选择的被试听到的则是问题的另一个版本：夹克仅售15美元，计算器要125美元。另一家店里的同款计算器卖120美元。值得到那家店去一趟吗？

这次，大多数人说“没必要”。

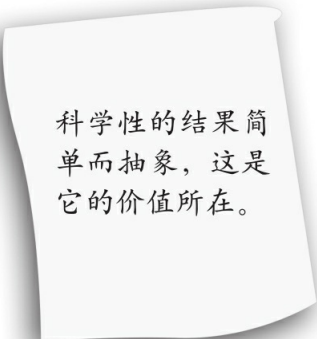
零售商绞尽脑汁想要搞懂寻常百姓为何宁愿在这儿出高价，也不愿意费点儿功夫去别处买便宜货。这样的结果相当具有挑衅性（假设这一调查也适用于真正的消费者），而且完全超出了标准经济学的理解范畴。在问题的两个版本中，买家都打算花140美元整，开20分钟车也都能省下5美元。“为什么我们愿意为了买个小件东西开车穿过城区省点儿钱，换成贵重东西反倒不肯了呢？”泰勒说道，“显然，这里有心理物理学的作用。对15美元的东西来说，省下5美元显得挺多；而对125美元的东西来说，5美元算不了什么。”

过去10年来，心理学家和行为经济学家对自己工作的实践意义甚为乐观。泰勒设想，不久的将来，会有“选择工程师”利用这门科学帮助人们做出更能准确反映其内在价值观的决定——这里的内在价值观，是推定价格及偏好这个勇敢新世界的内在价值观。泰勒从前的一个学生，森德希尔·穆拉伊纳丹（Sendhil Mullainathan）曾提到用决策理论帮助第三世界国家打破贫困的恶性循环。“我们这里说的是，一种技术正从行为经济学中显现出来，”卡尼曼说，“它不光是一件抽象的东西。你可以用它做事情。我们才刚刚起步。”

同样是这10年，该学科还衍生出另一种平行的、不那么理想主义的概念：靠它赚钱。一些学者，包括泰勒和特沃斯基在内，在营销杂志上大量发表文章，帮商人们打通任督二脉。围绕古老的营销技巧，比如价格以数字“9”结尾、回扣、折扣、类似“免费”牛排大餐的噱头，人们产生了新的科学好奇心，想知道它们到底管不管用，如果管用的话，又是怎么运作的。成立于1984年的专业定价协会（Professional Pricing Society）开始联合《财富》1 000强公司的商人们来交流想法。

西蒙虽然沉浸在行为学家的工作当中，却对它能否应用在商业世界有些怀疑。科学性的结果简单而抽象，这是它的价值所在。前景理论的论文就是一个例子。《科学美国人》杂志（*Scientific American*）的读者们喜欢它，因为它是一个能解释一切的简单概念。它本身对公司帮不上

什么忙，因为公司面对的是复杂的细节性问题，对具体的解决办法更感兴趣。



科学性的结果简单而抽象，这是它的价值所在。

技术上的发展成了关键。1974年6月26日，俄亥俄州特洛伊市的马什超市（Marsh Supermarket）扫描了一包黄箭口香糖的条形码。这可是收银台扫描的第一件消费产品。IBM设计的产品扫描的条形码至此大功告成。这之后的几年，扫描器迅速推广至大西洋两岸。它们带来了满坑满谷的信息。人人都觉得这些数据应该有用，可没一个人知道该怎么用。

西蒙的一个博士研究生，埃克哈德·库彻（Eckhard Kucher），写了一篇有关扫描数据的论文。库彻和西蒙意识到，数据能在价格心理学和现实实践当中架起一座桥梁。他们让分析家反向进行决策“实验”，观察消费者面对涨价、推销或打折会做出什么样的反应。跟传统经济学的数据一样，扫描数据捕捉到了个别和整体的行为影响，而对公司和消费者来说，其结果又足够具体。库彻建议开办一家咨询事务所，为企业如何微调价格当顾问。西蒙本来就在考虑这方面的事情，于是立刻就同意了。1985年初，两人的事务所开张了。

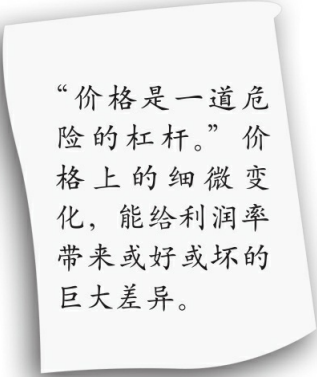
西蒙和库彻认为，一家公司对成本能想的招数并不多，但通常有定价的自由。两人发现，没有几个商人明确知道自己的主顾愿意付什么样的价格，或是价格对利润有什么样的影响。这是扫描数据能帮忙解决的事情之一。

自20世纪80年代以来，价格咨询业蓬勃发展。软件是工具箱里的一个重要组成部分。超市、百货公司或网上零售商有那么多价格，只有软件才能对其一一进行管理。按瑞沃尼斯（Revionics）公司CEO托德·P. 米肖（Todd P. Michaud）的说法，定价软件如今已经发展到了第四代。

客户公司输入扫描数据，便会为每一条码的产品得到利润最优化的价格，此外还有画得整整齐齐的图表，解释为什么价格应当调得高些或低些。“其实，零售定价软件现在已经能够自我教学了。”米肖夸口道。

当然，不光软件如此。价格也变得比以往任何时候都更具创造性。西蒙把他的顾问们看成是“价格结构”的建筑师。这些账单方案（想想你的手机费用套餐）以这样那样的方式，详细说明了大得州牛排牧场餐厅非线性定价的花招。它们怂恿消费者对低价开展荒谬的追求，让他们付出超乎本意的价钱。为什么这套把戏能够管用，决策心理学可以给出解释——可要是有人苦恼于它的伦理后果，倒没什么可说的。许多人觉得，顾客是上了当，所以才买下了多于自己真正需要的肉或通话时长。新的心理学反驳说，“顾客想要什么”并不是一种特别明确的东西。它是在现场构建的，许多在意识思维看来无关紧要的细节都可以对其产生影响。毫无疑问，价格顾问的工作，就是设计有利于自己客户的环境。

SKP公司的座右铭为：“价格是一道危险的杠杆。”价格上的细微变化，能给利润率带来或好或坏的巨大差异。西蒙估计，优化一家公司的价格，一般能把利润率提高两个百分点，比如，从5%提高到7%。米肖的说法同西蒙的差不多，他认为典型的增幅是1~4个百分点。由于初始的利润空间小，增加一两个百分点能极大地提高利润。能给绩效带来如此明显效果的干涉措施并不多。在数以百计的公司看来，这样的宣传说辞简直无法抗拒。

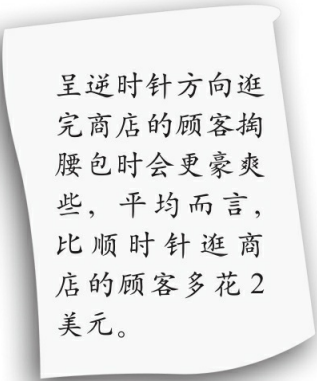


“价格是一道危险的杠杆。”价格上的细微变化，能给利润率带来或好或坏的巨大差异。

第10章 价格吸引的诱饵

说到对任意连贯性最狡诈的应用，去收银台看看便知道了，那就是超市发的“打折积分卡”（也叫“忠诚卡”，以下简称积分卡）。使用这种卡的消费者，是自己把自己当成了吝啬鬼。一想到错过了便宜5毛钱买条毛巾的好机会，他们简直要抓狂。所以，一站到收银台跟前，他们便摸索着掏出积分卡来。这些人，就是能为了节省5块钱开车穿过整座城市的顾客。

积分卡的数据能告诉市场，对价格最敏感的顾客最定期地购买哪些品牌和物品。按超市咨询事务所“威拉德·毕晓普”的吉姆·赫特尔（Jim Hertel）的说法，连锁店一般会安排500多样最频繁出售的商品，给它们特殊对待。超市知道，可乐、牛肉或者咖啡一涨价，顾客就能注意到。所以，它们总是尽可能地找最不容易受人注意的商品提价。细叶芹或者其他顾客购买不频繁的东西，比如甜面酱、石榴、奶酪、鲜榨橙汁等提了价，几乎没人会有意见。“靠这些东西，有机会赚点儿利润。”赫特尔解释说。这是因为顾客记不得上回买它的时候付了多少钱，或是对这些东西的成本没有准确的概念。



呈逆时针方向逛完商店的顾客掏腰包时会更豪爽些，平均而言，比顺时针逛商店的顾客多花2美元。

在确定什么能调动消费者支付意愿的道路上，超市顾问们尚未搬动的石头不多了。最近的一项调查其结果耐人寻味：呈逆时针方向逛完商店的顾客掏腰包时会更豪爽些，平均而言，比顺时针逛商店的顾客多花

2美元。

这一发现，是购物车运动研究提出来的。索伦森公司（Sorensen Associates）的赫伯·索伦森（Herb Sorensen）在手推车上安装了RFID电子标签，它们每隔5秒就发射一道无线电波。这叫“路径跟踪”（Path Tracker）技术，通过它，传感器能三角定位到每一辆手推车的位置，绘出它的运动状态，统计推车人以什么价格买了什么东西。没人清楚为什么逆时针运动的购物者买得更多。Envirosell公司的CEO帕可·安德希尔（Paco Underhill）提到一种挺热门的推测：北美人把购物车看成是“汽车”，要靠公路右侧行驶。“要想让我注意到你，”安德希尔说，“最好到我右边去。”根据这一理论，要是货架或墙壁在右边，习惯用右手的人（占了大多数）便更容易冲动购物。索伦森的研究结果已经得到广泛采用，超市把主要入口设在店面的右边，鼓励消费者沿逆时针方向购物。

泰勒最著名的思想实验跟杂货店有关。

价格实验

热天的午后，你躺在海滩上，迫切地想来上一瓶冰镇啤酒。一位朋友主动说，他愿意到附近唯一能买到啤酒的小杂货店去弄点儿来。他提醒你，那儿的啤酒可能挺贵，所以他问你愿意付多少钱。只有店里的价格不超出你的限额，他才会买。如果超过了，他就空手回来。

20世纪80年代初，泰勒向企业高管们提出这个谜题，人们报出的保留价格平均是1.50美元。他又把同一个故事讲给另一组听众，只不过把卖啤酒的地方换成了豪华度假酒店下设的酒吧。这组受众报出的平均价格是2.65美元。

故事的两个版本都明确指出，朋友买的是你最爱的啤酒品牌。不管在哪儿买，它都是同一种产品。酒店的氛围与此无关，因为啤酒是买回来在沙滩上喝的。然而，高管们愿为豪华饭店的啤酒出2.65美元的价格，但对小杂货店卖的同种啤酒却不愿出这个价儿。酒店附设的酒吧要价2.65美元，人们觉得那是个公平的价格，小杂货店要这个价，就成了敲竹杠。

泰勒建议小杂货店“投资一些表面上看似多余的豪华设施，或是增设一间酒吧”，他认为这将提高人们对啤酒恰当价格的心理预期，进而

带来更多的销量。泰勒给小杂货店的另一项建议是，卖超大容量的啤酒，不是通常的12盎司装，而是16盎司装。因为消费者记得一罐12盎司装啤酒卖多少钱，却恐怕不知道16盎司装的啤酒该卖多少钱（他们可以算一算，但大多数人不会这么做）。此外，大容量装啤酒要比小罐装更容易带来额外利润。

泰勒的这两个想法都可以在如今的超市行业中看到。像全食市场（Whole Foods Markets）这样的高档超市，就充分利用了大部分“多余的豪华设施”。这样一来，他们得以开出在其他情况下顾客接受不了的高价格。每一家“全食市场”的分店都有一个引人注意的特色生产部。“这些土豆有多小呢？”曼哈顿时代华纳中心分店摆卖的俄罗斯手指土豆旁竖着告示牌，上面和气地这样提问道。显然，比你的小指头还小巧可爱——你还好意思比较价格吗？

诸如开市客和山姆会员制仓储式超市出售按千克卖的蓝奶酪沙拉酱、32卷一包的卫生纸。你以为你批量买进很划算——有时确实如此，可另一些时候，这笔交易并没你想得那么划算。好多消费者都不知道6磅重的菠萝块成本应该是多少。

“有机”和“绿色”的标签让高低档超市都大赚了一笔。不管这些术语到底是什么意思，它们都别有用意：偏高的价格不再显得像是在宰客了。

再来一个有关啤酒的问题。

价格实验

老乔到货架上选啤酒。有一种高级啤酒，售价2.60美元，另一种是廉价品牌，只卖1.80美元。高级啤酒更“好”（不管它指的是哪方面）。品酒的行家们给高档品牌的质量打了70分（百分制），廉价品牌则只有50分。老乔应该买哪种啤酒呢？

杜克大学商学院的教授乔尔·休伯（Joel Huber），还有他的研究生克里斯托弗·普多（Christopher Puto），向一群商学院的学生提出了这道难题。学生们首选高档啤酒，选择高级啤酒和廉价啤酒的人数比是2:1。

另一组学生则可从三种啤酒中做选择，除了前述的两种，还有

另一种超低价的劣质啤酒，售价1.60美元，品质得分也最低，40分。

没有任何一个学生想要买这种超低价啤酒。即便如此，它仍然影响了人们的选择。选择先前那种廉价啤酒的学生比例从33%增加到了47%。超低价啤酒的存在让廉价啤酒变得名正言顺了。

还有一组被试，他们面对的三种选择是最初的廉价啤酒和高级啤酒以及一种超一流啤酒。和不少高档货一样，这种超一流啤酒要贵得多（3.40美元），但质量只稍微好一点点（75分）。

10%的学生表示，他们会选超一流啤酒。令人吃惊的是，其余90%的学生全选了高级啤酒。这下没人想要买廉价啤酒了。

这简直就像牵线木偶被绳子拉着走嘛！休伯和普多发现，只需增加一种很少有人会选的第三个选项，他们就能摆布学生的选择。

选择一种美国产啤酒应该是件挺简单的事情。好多“味道盲测”（指测试味道的人事先不知道喝的是什么品牌的饮料）都说，贪杯客们分不出百威、米勒或者银子弹啤酒的味道。由于面向大众的各种啤酒味道都差不多，所以你只需在价格和质量之间做个权衡就可以了（当然了，你必须好好考虑一下“质量”是不是市场营销制造出来的幻觉）。

如图10-1所示，理想中的啤酒应该既便宜又质量好，落在图的左上角。用不着我说你也知道，价格和质量之间往往有着关联（尽管不那么紧密）。这就是说，各种品牌大多是落在一条从左下角到右上角的斜线上。

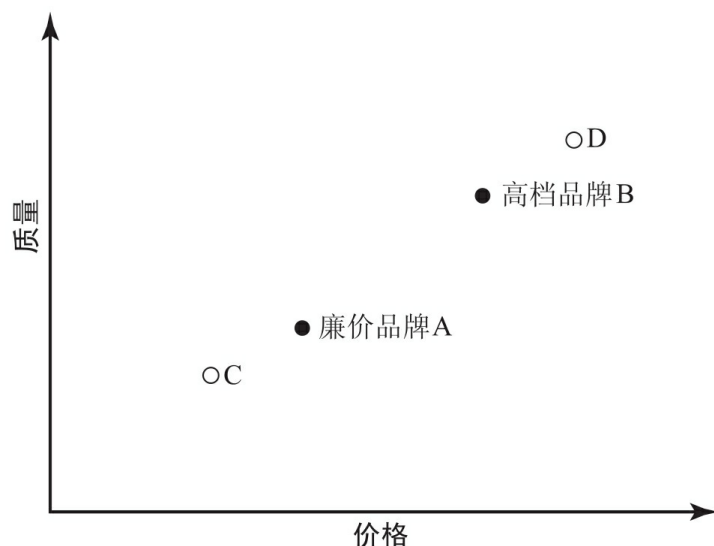


图10-1 价格和质量的关系

休伯和普多发现，为了提高廉价品牌A的市场份额，你只需提供一种更便宜的选项C。C成了“诱饵”，它本身可能不会得到太多市场份额，但它会产生一种吸引效应，把消费者的选择转移到原先的廉价品牌A上。类似地，增加一种高价诱饵D，会把消费者的眼光往上拉，扩大高档品牌B的市场份额。

等被试做好了选择以后，休伯和普多问他们为什么要那样选。答案听起来也不无道理。选择三者间中等价位选项的人说，他们这样决定“比较安全”，是一个“折中”的选择。最便宜的啤酒说不定非常难喝，最贵的那种则有敲竹杠的嫌疑，中档的啤酒应该没啥问题。

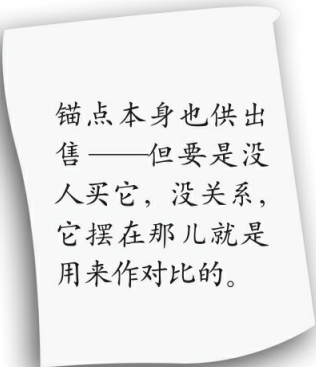
休伯和普多的论文发表在1983年的《消费者研究期刊》（*The Journal of Consumer Research*）上，如今已是当代市场营销的奠基之作。然而，他们评论说，企业其实早就对这些概念有着直观上的认识。20世纪60年代，安海斯-布希（Anheuser-Busch）公司在全美范围内积极推广它旗下的超高档啤酒米狮龙（Michelob），百威却借力登上了销量第一的位置。要是啤酒消费者们真的知道他们想要什么、该出多少钱，那米狮龙肯定会冲击百威的市场。可结果，百威和米狮龙的总销量却都增加了。休伯和普多认为，米狮龙让百威显得“不那么极端，不那么昂贵，又不那么精英”。一部分百威的买家改喝了更高档的米狮龙，但平常买更便宜品牌啤酒的人却有不少换成了百威，两者扯平了。换句话说，米狮龙的广告让一些米勒啤酒的买家改喝百威了。总体而言，安海斯-布希突围成功。

吸引效应的使用方式很灵活。1961年，宝洁公司推出了纸尿裤品牌帮宝适。最初，帮宝适是跟传统的布质尿布竞争。那时的人们觉得纸尿裤更方便，但也太贵。1978年，宝洁公司推出了另一种高价的一次性尿布露肤（Luvs）。它不光占据了高档纸尿裤的市场，还带来了一种对比，让购买布质尿布的家长们相信：帮宝适并不是那么奢侈的东西。到20世纪90年代中期，时代已经变了。家长们早就改用一次性纸尿裤了，只有少数环保分子还坚持用布质尿布。宝洁公司决定，这一回该借助低档产品的诱饵了。从1994年开始，它把露肤重新定位成一个廉价品牌。

第11章 普拉达的“托儿”

普拉达门市经理们的嘴皮子上经常挂着“锚点”这个心理物理学术语。在奢侈品行业，它指的是一种价格高得令人咋舌的东西，展示它的主要目的是摆布消费者。锚点本身也供出售——但要是没人买它，没关系，它摆在那儿就是用来作对比的。跟它一比，其他所有东西就都显得负担得起了。“这种手法最早可以追溯到17世纪，”帕可·安德希尔说，“你卖一样东西给国王，但朝廷上的每一个人都必须另有一件稍微差些的配套货。橱窗里有500美元的皮包，你选件新T恤走总不算过分吧？”

如今，这一手法或许会用在5位数的手袋和7位数的手表上。置身自20世纪30年代以来最严峻的经济衰退时期，美国知名时装品牌拉尔夫·劳伦（Ralph Lauren）却叫卖起了14 000美元的鳄鱼皮手袋。爱马仕有一种售价33万美元的手表，如果你有100万美元整，则可以买到宇舶（Hublot）的“黑色鱼子酱创世纪”表，上面密密麻麻地镶嵌了322颗黑色钻石，让你看不到一点儿金属底盘（被钻石遮住的金属是18K白金）。谁会花100万美元买只手表呢？这正是你该好好关心的问题。接下来的问题则是，你愿意花多少钱买一只非常好的手表呢？这非常类似锚定实验里提出的问题，说不定也会有着同样的结果。

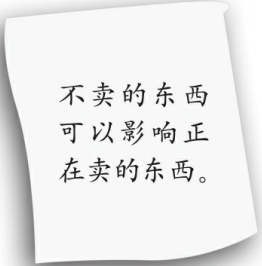


锚点本身也供出售——但要是没人买它，没关系，它摆在那儿就是用来作对比的。

锚点的价格就像史蒂文斯实验里的耀眼白色光环。跟它一比，购物

狂们相中的东西就显得像是便宜货了。高价格还能起到跟“托儿”一样的作用。它们叫顾客相信，肯定有人会掏出那么多钱（要不然干吗把它们摆出来呢）。这不一定是正确的结论。宇舶只生产了一只100万美元的手表（并小心地把它标成是“特殊订单”）。爱马仕只生产了两只33万美元的表，而超贵的手袋一般连旗舰店里也只有一个。尽管这类东西的供需市场是彻头彻尾的幻觉，在《罗博报告》（*Robb Report*）和名人街拍的推波助澜下，它们却显得极其逼真。有人拍下了伊娃·朗格丽亚（Eva Longoria）[\(17\)](#)背着亮蓝色蟒蛇皮米兰达教练包的照片！当然了，她到底有没有出过这笔钱跟这里的论点没有任何关系。

就算是在经济最好的时候，奢侈品商店也不过是面子工程，是用来叫物质主义者们相信，这个世界是比实际上更为富裕、更爱挥霍的。感官逻辑（Sensory Logic）的营销顾问丹·希尔（Dan Hill）说，成功的商店利用高价物品来创造“混合着愤怒与幸福的复杂感受”。中产阶层消费者感到愤怒，因为他们买不起店里陈列的、穿在名人身上的东西，但他们又因为买了其他东西而下意识地高兴起来。



不卖的东西
可以影响正
在卖的东西。

行为定价理论的重要见解之一是，不卖的东西可以影响正在卖的东西。特沃斯基喜欢讲下面这个故事。

价格实验

在以所售产品质量好、价格高而出名的威廉姆斯-索诺玛（Williams-Sonoma）厨具连锁店，有一种神奇的烤面包机，售价279美元。他们后来新增的一种稍微大些的型号，售价429美元。猜猜后来怎么样？

429美元的型号滞销得一塌糊涂。顾客又不是开寄宿学校的，要一台更大的面包机来干吗？可279美元的型号销量差不多翻了一倍。

显然，有人早就对威廉姆斯-索诺玛厨具连锁店卖的高质量面包机念念不忘了。只可惜价格叫他们望而却步。279美元看起来有点贵。可等店里新增了429美元的型号，279美元的机器就不再那么像奢侈品了。人们可以给自己找理由说：这是一台有用的机器，429美元型号能做的事情它全能做，而且它的价格还更便宜。新增了另一个价格的产品，尽管几乎没人会选它，它却提高了消费者愿意为面包机支付的价格。

按特沃斯基的判断，威廉姆斯-索诺玛厨具连锁店并不是预先安排好这么做的。打那时起，零售商对价格的对比效应就渐渐开窍了。1992年，特沃斯基和伊塔玛·西蒙森（Itamar Simonson）发表论文，扩展了休伯和普多的研究成果，指出了操纵性零售的两大原则。

第一条原则：避免极端。他们通过调查（对象包括照相机、钢笔、微波炉、轮胎、电脑和擦手纸）发现，在消费者不确定的时候，他们会避免购买价格最贵或者最便宜、质量最好或最差、型号最大或最小的物品。大多数人倾向于走中庸之道。故此，要想卖出800美元的鞋子，方法就是在它旁边摆双1200美元的鞋子。

西蒙森和特沃斯基写道，就感知和判断领域而言，对比效应无处不在，同样的圆圈，如果周围围的是小圈，它会显得大一些；而要是周围围的是大圈，它又会显得小一些。同样道理，同一种产品，倘若有吸引力较为逊色的产品映衬，它会显得更迷人；而在吸引力更强的产品映衬下，它则会显得没什么特色。我们建议，对比效应不光要用在大小或吸引力等单一的属性上，还要用在各种特性的权衡上。

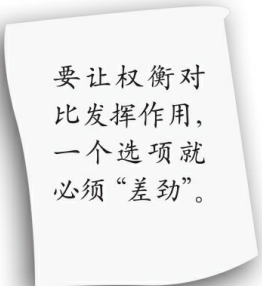
由此，他们提出了第二条原则：权衡对比。走进一家批发店，店里会摆有数十种手袋，按任何人的标准，这些货色都不是什么最高水准的精品。一种比较实用，一种更为时尚，另一种颜色则要有趣些，还有一种打6折。厌恶损失的消费者，对如此复杂多样的选择感到不舒服，她担心自己选了甲又后悔没选乙……

权衡对比原则认为，倘若甲物明显比较差的乙物要好，消费者会倾向于购买甲——哪怕还有许多其他选项，哪怕根本没办法判断甲是不是所有选项中最好的。光是甲比乙好这个事实就是一个卖点，它承载着远比理性更大的分量。很明显，购物者想选择一个（对自己、对朋友、对仔细盘查她信用卡账单的配偶来说）合乎情理的东西，缓解自己的焦虑。他可以告诉自己，买甲物是因为它比乙物要好得多。

权衡对比在奢侈品贸易中尤其重要，各品牌都有只卖自家商品的旗舰店。此外，拥有强势品牌的零售商在价格上有很大的灵活性。一个只

买周仰杰（Jimmy Choo）高跟鞋的购物者，对其他品牌卖什么根本不在乎。西蒙-库彻事务所的顾问们发现，他们老是责备客户把价格定得太低。“奢侈品的价格和任何成本都没有直接的关系，”SKP公司的一份市况报告干巴巴地宣告，“奢侈品定价的艺术，就是要量化产品对消费者的价值，不考虑成本、竞争对手或市场价。”

蔻驰对自己的每家旗舰店只分配一两个超昂贵的手袋。它们陈列得漂漂亮亮，并用一种端庄、好认的大号字体标示其价格。这些手袋蔻驰卖不了许多，但就算一个都卖不出去，它说不定也挺高兴。



要让权衡对比发挥作用，
一个选项就必须“差劲”。

举个例子，蔻驰旗舰店有一种7 000美元的鳄鱼皮包，还有一种样子很相似的包，只不过是鸵鸟皮的，卖2 000美元。大多数消费者很难猜出哪种是7 000美元的，哪种又是2 000美元的，甚至还有人觉得鸵鸟皮的比鳄鱼皮的更高级。

要让权衡对比发挥作用，一个选项就必须“差劲”。因为几乎所有人，甚至是蔻驰的客户也在乎价格，所以，一种价格似乎高得没道理的东西，在价格这个尺度上，无疑成了一个“差劲”的选项。7 000美元的包让样子类似的2 000美元的包更受人欢迎了（它便宜这么多，而且是出自同一个设计师的品牌）。于是，2 000美元的鸵鸟皮手袋销量提高了——若非如此，消费者说不定会觉得它价格太贵、太过分而放弃购买它。

时尚界的现实情况完全吻合西蒙森和特沃斯基的两条原则。真正的时尚总是那么昂贵、不舒服，叫人吃惊又离谱。只有少数拥有精挑细选的无瑕身材和足够充裕的钱包的人才穿得起它。其他人买些更舒服、价格更合理的东西就满足了。这几样近乎遥不可及的东西可以操纵绝大多数的消费者。

普拉达最为推崇环境的营造。它找著名建筑师雷姆·库哈斯（Rem Koolhaas）设计自己位于纽约苏荷区的门店，费用是每平方英尺1 700美元，房租另付，也是1 700美元。它绝不会用一楼的空间来放置几乎卖

不出去的东西，除非别有内情。权衡对比属于这门买卖的一部分成本，跟广告、橱窗展示或者“建筑师”设计一个样。要是有什么东西跟高价的锚点商品类似，售价又不超过前者的1/10（这可不常见），买不起高价锚点的人，总可以买副300美元的太阳镜试试。再不然，110美元的手机挂件也行啊！英国的普拉达网站会提示你去哪儿捡便宜（至少是在网上）。它会提供10款女鞋、23款手袋和54样“礼物”——类似钥匙串、手链和高尔夫球钉架等小玩意儿。一串手链卖60英镑，利润率何其惊人啊！

第12章 菜单标价心理战

“丹尼尔·布吕得（Daniel Boulud）开过一家餐厅，香菇加神户牛肉的汉堡卖100美元，”餐厅顾问布莱顿·奥德尔（Brandon O'Dell）说，“兴许每个星期只会有人点一回100美元的汉堡，但要点不是靠卖这种汉堡狠狠地赚钱，而是为了让菜单上的其他东西显得比较便宜。有人看见菜单上一个汉堡就卖100美元，再看到50美元一份的牛排，就会觉得选后者占便宜了。”

经济衰退降临之前，曼哈顿的菜单贵得离谱，布吕得便是开创这股潮流的“功臣”。

2001年，布吕得的DB现代小酒馆（DB Bistro Moderne）便出售起28美元的汉堡（内夹红烧排骨和鹅肝）了，那时候这个价儿真是叫人瞠目结舌。它吸引了媒体的大量关注，一时间仿效者众。布吕得又提高了赌注，改卖一种夹着20块黑蘑菇（时令菜）的汉堡，售价150美元。随后，仿效者之一，华尔街汉堡专柜（Wall Street Burger Shoppe）卖起了夹25块香菇和神户牛肉的汉堡，售价175美元一个。

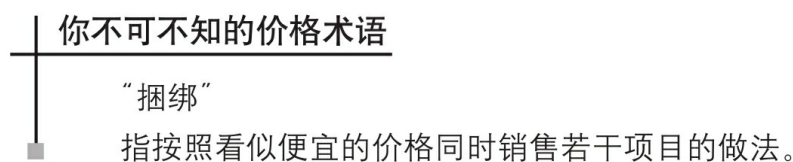
酒店餐厅接受了这套理念，毕竟，从理论上说，凡在曼哈顿住得起酒店的人，都有钱可烧。在美丽殿（Le Parker Meridien）酒店的诺玛餐厅，菜单上出现了1 000美元的鱼子酱和龙虾煎蛋卷；威斯汀酒店拿出了1 000美元的香菇和枸杞子。把这些东西列在菜单上，餐厅花不了多少成本。如果哪天真有人点，大厨可就中头彩了。但1 000美元的百吉饼主要是向客户施魔法的，好叫他们无意当中花更多的钱，效果奇佳。

“像红辣椒（Chili）和苹果蜂（Applebee）这样的地方，都把这做成一门学问了，”从前是餐馆老板，后来转职为菜单顾问的吉姆·劳布（Jim Laube）说，“到这些地方，留心看看他们是怎么给想要卖的东西争取关注的。听我说的错不了，他们对想卖什么产品再清楚不过了。”

时时乐（Sizzler）、猫头鹰（Hooters）、星期五（TGI Fridays）和橄榄花园（Olive Garden）等美国连锁餐厅都表示：不管烹饪技法何等高明，他们靠的都是一套前沿的菜单科学。菜单心理设计的目标是把顾客的注意力引导到盈利（也就是“标价过高”的）项目上。按餐饮行业的

传统，菜单上的项目分成明星、问题、耕马和瘦狗。明星指的是受人欢迎的高利润项目，换言之，这种东西，顾客愿意付出比制作成本高得多的价格买它。问题产品的利润高，但不受欢迎；耕马则相反，受人欢迎但无利可图。瘦狗既不受欢迎，又赚不上钱。顾问们想办法把问题变成明星，怂恿顾客远离耕马，叫所有的顾客都相信菜单上的价格比表面上看起来更合理。

“包围”就是一种常见的伎俩。把像牛排这样的昂贵项目分成两种分量。顾客看不出小份菜品到底少了多少，但没关系。他认为小份菜品的价格更有吸引力，因为它花费得少。其实，餐馆方面本来就想卖出小份牛排，所谓的“低价”也正是他们想要开的价。如果菜单上有三种这样的东西，真的很能提高利润呢！



快餐里的“套餐”和“份饭”就属此类。人人都晓得，捆绑销售是要诱使你选购额外的东西。汉堡加薯条和苏打水的套餐，价格只比单点汉堡加苏打水贵几分钱，而你还能得到薯条。价格顾问泰普·卡尔玛（Tepper Kalmar）说：“给第三个项目打个小小的折扣，总毛利就上去了。”

有效捆绑还有另一个原因，它混淆了视线。因为有了份饭的价格，所以就餐者花13美元买两枚扇贝（这是泰勒在旧金山的“扎格特就餐指南”上找出的例子）也不觉得愤怒了。你没法儿肯定成本到底是多少以及它是不是太贵了。

等回头客对常见套餐的价格日渐熟悉后，捆绑效应也随之减退。出于这个原因，连锁餐厅的菜单总是不停地在变。新的主菜冒出来，旧的要么改名，要么消失。套餐可以变成超大型。想买扭扭薯条？你没法买到跟上回完全一样的东西，自然也没法准确地比较价格了。

要是所有这些招数都不奏效，卡尔玛便要餐馆利用“机会”涨价。这应用的是公平研究的发现：凡涨价，卖家都该把责任归咎到其他因素上。卡尔玛建议，若有必要，餐馆应张贴告示，解释说是能源成本或是粮食成本（诸如此类的原因）逼得他们被迫“暂时”涨价，转移额外成本。

Pastis的菜单上还用了省略号。省略号的目的是叫用餐者把视线从菜品转到价格上。他们自然那么做了。可餐厅并不该这么想，因为就餐

者照价格点菜，餐馆就赚不了钱了。为了把价格敏感度降到最低点，西雅图顾问格雷格·拉普（Gregg Rapp）告诉客户，把省略号、美元符号、小数点和分位数全部去掉。联合广场咖啡厅如法炮制。它的菜单采用中央对齐方式，这样价格栏就没法整整齐齐地排成一行。倒不是说顾客没法核对价格了，只不过，大多数人会跟着看得到的微妙线索走。菜单上的线索说：“请注意食物，别看重价格。”

图12-3是纽约巴萨泽餐厅最新菜单上的一页。尽管有太多的价格排成了一列，可巴萨泽餐厅还是采用了菜单心理学里的一些复杂技巧。



图12-3 纽约巴萨泽餐厅的菜单

典型的就餐者打开菜单，最先看到的是右手页的上部。巴萨泽餐厅没在这儿抢机会：它在右上方放了一幅图片——另一种抓眼球的方式。从那里，人们的视线大多会向下移动到右手页的中央部分。菜单顾问利用这部分最佳菜单空间列出高利润项目和价格锚点。本例中，价格锚点是巴萨泽餐厅的海鲜拼盘，售价110美元。心理物理学研究称，在刺

激的邻近区域对比效应最为强烈。普通人对它是否适用于菜单上的价格还有点儿拿不准，可顾问们却认为它能行。他们建议把高利润的项目紧紧地挨着高价锚点。110美元这个价格的真正用意是，吸引顾客在冲动之下点它左边65美元的豪华大拼盘，或是它下面价格更“温柔”的各类海鲜。

在菜品周围围上方框能吸引注意力，通常来说，也能吸引顾客点它。基围虾15美元很奢侈吗？跟隔壁页的110美元狂欢比起来它算个啥？带花边的方框更好。底部的奶酪恐怕属于利润颇高的问题产品。

文字说明和照片也能带动盈利项目的销售。食物照片是最有力的刺激，可同时也是菜单上的一种特大禁忌。红辣椒和苹果蜂这两家连锁餐厅大量使用照片，所以，对其他任何心怀美食壮志的餐厅而言，用照片就意味着死定了。连红龙虾连锁店（Red Lobster chain）后来在做形象升级时，也觉得有必要把照片从菜单里拿掉。巴萨泽餐厅的菜单印着美味的海鲜拼盘插画，这个档次的餐厅最多也只能做到这一步了，它是用来给两种最昂贵的菜品吸引注意力的。

拉普认为，自己的任务并不是从菜单上除掉不赚钱的主菜。“我们不想把它从菜单上取消，因为这么做可能会失去顾客。”他解释说。其实，只要把上述建议反过来用，就能把菜品“最小化”——取消方框或文案，把它放到菜单最不起眼的地方。巴萨泽餐厅对它的汉堡和神秘的招牌鳕鱼用上了这一招，前者小得叫你容易把它看漏，后者则全然未加任何评注。

第13章 超级碗门票的价格

每年，美国职业橄榄球大联盟都会以“票面价值”卖出500对超级碗门票。如今它大概是每张票400美元（一对800美元），对门外汉来说，这蛮便宜的。在出售二手票的网站上，超级碗的门票值2 000到6 000美元呢！

想得到一张票面价值门票的机会微之又微，你必须突出重围才行。申请者必须把申请函打印出来，并在每年2月1日至6月1日，以挂号信等方式邮寄到美国职业橄榄球大联盟的纽约办事处。纽约办事处的工作人员会在10月进行随机抽选。近年来，差不多每年都有36 000名球迷申请购票，也就是说，抽中的机会大概是1/70。有些摇奖彩票的中奖率都比这高。为什么美国职业橄榄球大联盟要搞这套把戏呢？按美国职业橄榄球大联盟公共关系副总裁格雷格·艾洛（Greg Aiello）的话来说，抽选的目的是确定一个“公平、合理的价格”。

这话听起来虚伪透了，但其实还好。美国职业橄榄球大联盟的门票抽选制度完全吻合公平研究所得出的结论：在人们心里，抽选和排队要比高到没边的市场价更公平。按那样的市场价，就只有富人才买得起超级碗的门票了。SKP公司的一份报告发现，体育赛事的门票“几乎是尖叫着扑向非线性定价结构”，也就是说，对相同的门票，不同的人付不同的几个价格。

普林斯顿大学的经济学家艾伦·克鲁格（Alan Krueger）给2001年第35届超级碗的门票打了分，并设法对球迷搞了一次快速调查。他发现了一件令人吃惊的事：接受调查的球迷里竟有40%是免费拿到票的，只有20%的人掏了比门票票面价更多的钱。

这怎么可能呢？美国职业橄榄球大联盟说，超级碗75%的门票被分发给参加联赛的各支球队，其中的大部分更是给了打比赛的两支队伍。这些球队可以按自认合适的方法处理门票。大多数球队会举行门票抽选，一般是为了不让它们又落到拿季票看比赛的球迷手里。另外25%的门票由美国职业橄榄球大联盟自己分派。大部分给了贵宾、媒体和慈

善机构。这么慷慨，美国职业橄榄球大联盟给得起，因为它60%的收入来自电视台的授权转播费。

克鲁格最值得注意的发现是，几乎没人愿意按市场价购买或出售超级碗门票。卡伦·麦克利恩（Karen McClearn）持有巴尔的摩乌鸦队的季票，她告诉克鲁格，她跟丈夫能来看比赛是因为他们抽中了门票，付了一个远比市场价要低的数儿。克鲁格问她愿不愿意按每张票4 000美元的价格把票卖了。“没门儿！”麦克利恩说。后来，乌鸦队跟巨人队比赛，以17:0的分数领先，她干脆说：“哪怕每张票5 000美元也不卖。”

克鲁格后来又对按票面价（当时是325美元）买了票的球迷做了一次更正式的调查，他问球迷愿不愿意把票以3 000美元的价格卖掉。93%的人说不愿意。门票显然比这个数更值钱。就算要他们在3 000美元和门票之间选，他们也会选择门票。克鲁格又要球迷们想象，要是把按票面值买到的票弄丢了，他们会花3 000美元去买吗？球迷们一致说不。这样说来，门票并不值市场价了。在某种感觉上，超级碗的门票是无价的，单一尺度的估价方式根本没办法解释球迷对克鲁格问题的回答。

对于如何向过分热情的球迷分配过分稀缺的门票，美国职业橄榄球大联盟很有经验。把它跟2007年全程免费的汉娜·蒙塔纳（Hannah Montana）巡演比比看吧！

价格实验

在巡演途经的55个演出点，疯狂的家长们都在拼抢演唱会的门票。在每一个城市，定价在25~65美元间的门票，通过官方渠道几分钟就卖个精光。其中很大一部分落到业余或专业的“黄牛”（票贩子）手里。价格翻10倍的情况比比皆是。麦莉·赛勒斯（Miley Cyrus）歌迷俱乐部甚至提起了诉讼，说官方告诉他们凭每年29.95美元的会员资格就能拿到门票，结果俱乐部会员根本没得到票。

网上一张海报拿此事打趣道：“妈咪，给我票，要不我就一辈子不喘气了！”广播电台拿门票当比赛的奖品。一位妇女在短文比赛里赢到了门票，她在文中说自己女儿的父亲在伊拉克被路边炸弹给炸死了，当然，那可怜的男人并没遇到这么一回事。

从涨价幅度来看，票贩子在整体上挣得应该比赛勒斯多。可一张票到底值多少呢？没得到票的父母们说eBay上的价格太高了；得到票的幸

运家庭说它是无价之宝。

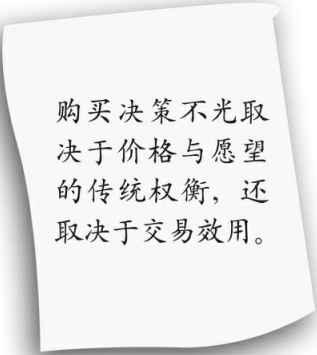
倒卖门票的家伙们冒险打破了公平原则。2009年布鲁斯·斯普林斯汀（Bruce Springsteen）巡演期间，票务网站Ticketmaster居然把歌迷们重新转到了二手票网站TicketsNow，这家网站又刚好是Ticketmaster的全资子公司。“卖完了”的门票转眼就出现在了Ticketmaster上——售价高达1 600美元。一名歌迷说，票刚开始卖的时候，她就从两台电脑上登录了Ticketmaster，结果立刻被定位到“黄牛票”网站。斯普林斯汀勃然大怒，逼得Ticketmaster的发言人公开道歉。美国新泽西州的总检察官答应进行彻底调查。其实这有点儿不合常理——票价太高（高于歌迷的支付意愿）竟然比错过演出更能激怒歌迷。

这种矛盾并不仅仅体现在娱乐活动的门票上。想想酒店下设的迷你酒吧。那儿的好吃的琳琅满目，可价格真是疯子才愿意给。倘若价格是单一向度的，你根本就不会光顾这种迷你酒吧。问题在于，有时你一个人走在陌生的城市，忙活完一天又累又饿，还有什么比一大块巧克力饼干更美妙呢？这样一块饼干，在迷你酒吧里卖8美元，还得上税。你大概会进行一番激烈的心理斗争。一来，不管那饼干卖多少钱，你都想来上一块；二来，一块饼干卖8美元，真该有法律来管一管。

聪明的朋友会说，买了那该死的饼干吧！要是你太过节俭，对想要又买得起的东西都出不了手，那就叫吝啬啦！可连平素里乱花钱的浪荡子也不见得会采纳这一建议。买卖总该按道理来啊！

泰勒用“交易效用”的概念来解释这一点。当消费者认为物品的真实价值高于其销售价格时，购买它会带来正交易效用。换用通俗的语言来说，这是一个便宜货，人人都喜欢便宜货。可要是物品的感知价值低于价格，人们就觉得卖方是在敲竹杠，交易效用为负。泰勒的论点是，购买决策不光取决于价格与愿望的传统权衡，还取决于交易效用。

交易效用会造成两种我们很熟悉的后果。有时，消费者觉得买卖划算，会一时冲动买下完全没用的垃圾。在这种心理的驱使下，电视购物节目、厂家直销、亏本甩卖、免税商店欣欣向荣。硬币翻个面儿，则是迷你酒吧和超级碗门票面临的困境。有时，消费者会克制自己，不去购买想要而又买得起的东西，因为内心的声音告诉他们，那是“敲竹杠”。而且，他们还会对自己压根儿不埋单的价格抱怨个不停。你可以说自由市场就是这样——那是空洞的逻辑，而这是热辣辣的情绪。



购买决策不光取
决于价格与愿望
的传统权衡，还
取决于交易效用。

按泰勒的模型看，消费者有两种意识。最近有证据显示，真实情况就是这样。有人对最后通牒博弈做了一些独创性的大脑扫描研究。

响应者要面对类似超级碗、迷你吧困境中的低出价。假设提议者的出价是10美元里只分给响应者1美元。一方面，即便1美元也是白捡的。我们一生下来就接受过教育，只要是钱，牢牢抓住准没错。另一方面，10美元里只能分到1美元，这样的待遇太不公正。对大多数西方人来说，交易不公的想法会胜过白捡钱的想法，他们会对这样的分配方案说不。

2003年，艾伦·桑菲（Alan Sanfey）和同事们做了一个实验。实验者用磁共振成像扫描进行最后通牒博弈的被试头部。扫描图像揭示，公平的出价（五五分或四六分）激活的大脑部位，有别于极不公平的出价（一九分或二八分）激活的大脑部位。不公平的出价激活了脑岛皮质（痛苦和恶臭也能激活这一区域）和背外侧前额皮质（这一区域跟规划及决策有关）。这似乎表明，人一方面发自内心地拒绝低出价，一方面又想留下白得的金钱，两者产生了冲突。正如一篇调查文章评论此项研究时所说，“不公平出价激活了脑岛，说明‘我痛恨有人这么对我’一类的口头表白并非夸张的比喻，而是事实——他们是真的觉得那很讨厌。”

第14章 电视购物全是套装且还有赠品

1978年，广告人阿瑟·希夫（Arthur Schiff）接了一档没啥指望的工作：为美国俄亥俄州弗里蒙特（Fremont）生产的廉价刀具设计电视商业广告。希夫为产品取了个亚洲化的名字“金厨”，写了一段两分钟的现场广告词，它成了以后电视购物节目的通用模板。照希夫的想象，你不光是卖产品，还“免费”搭配售卖一大堆额外的东西。

“你会为这样一把刀具出多少钱？”金厨广告的播音员问，“在回答之前，请先听：它送一把配套的叉子，锋利无比，能帮你把切菜变成享受。别急，还有很多其他的……”接下来，播音员抛出一套六合一厨房工具，一套牛排用刀，还有一把独特的螺旋形切片刀。

“广告播到最后，”金厨的合伙人之一，埃德·范伦蒂（Ed Valenti）说，“你都搞不清自己得到了哪些东西，可你知道，它花不了几个钱。”

金厨刀具原价9.95美元——外加其他各种免费的东西，另外广告还列举了从电视上买东西的各种不确定性。范伦蒂甚至说，是金厨的订购热线最先创造了“打订购电话免费”的说法。到沃伦·巴菲特的伯克希尔-哈撒韦公司于1984年收购它的时候，金厨公布的销售额已达5 000万美元。

电视购物节目的风格有点儿像歌舞伎表演。成功的电视购物节目总是最能打动消费者的心。毕竟不管产品有多么不同，人性始终大同小异。电视购物行业的重心，就是泰勒所谓的“别把所有圣诞礼物包在一个盒子里”原则。1985年，泰勒在《营销科学》（*Marketing Science*）期刊上发表论文《心理账户和消费者选择》，对消费者如何判断东西值不值得买、该在什么价格上买提出了一个独创的观点。

泰勒把前景理论应用到典型的交易上，一方付出一个价格（损失）以获取某种有价值的东西（收益）。收益和损失都呈现出报酬递减的趋势。3万块奖金挺好，但它并不比1万块的奖金好上三倍。故此，较之一次获得3万块奖金，分别获得三次1万块奖金带来的喜悦更多。得到三次奖金，你会高兴三次。意外之财的实际金额并不像你想的那么重要，

得到的次数才更能影响你的情绪。

泰勒在康奈尔大学的学生们身上试验了这一原则。他问学生：

A先生分别赢了两次彩票，一次50美元，一次25美元，B先生只赢了一次，但金额是75美元。谁更高兴一些呢？

大多数人觉得A先生更高兴，因为他赢了两次。基于这一点，泰勒推断：营销工作者与其宣传产品本身有多么美妙，倒不如把产品分散开来，一个特性一个特性地进行宣传，或是捆绑销售若干产品。电视购物节目20世纪80年代就这么做了，如今还在这么做。在电视购物节目里，你唯一买不到的一件东西，就是单纯地只买一件东西。

“买一件Snuggie(18)，免费赠送价值19.95美元的阅读台灯，邮费7.95美元；买两套，邮费全免。”Snuggie的推销词说。可要是你只想要一件Snuggie呢？对不起，这样不行。

有一种在电视上打广告的黏合剂，最小的起订量差不多是三瓶半，推销词说：“通常，一瓶强力胶卖19.99美元，外加8.95美元的手续费。可要是你今天就订购，我们会送您三大瓶强力胶。而且，您还会收到一份旅行套装作为特殊奖励。现在订购一套强力胶，让我们为您省钱到家！”

形状就像一发炮弹的魔弹搅拌机把泰勒提出的规则用得淋漓尽致。

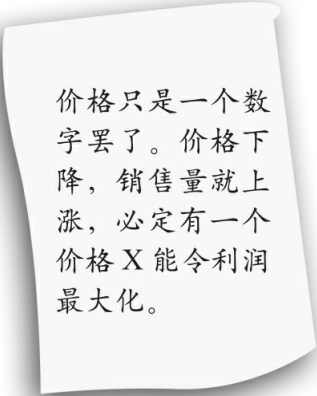
“您将得到哪些东西呢？高扭矩的基座、十字形和一字形的刀片、大号和小号的子弹杯……”他们喋喋不休地列出了21种零配件，仿佛每一件都是单独而又有用的产品。此外还附送“4个大号宴会马克杯，加上舒适的杯套，把您的魔弹搅拌机变成终极宴会利器……魔弹搅拌机‘10秒速成菜谱’和‘特惠项目’——魔弹搅拌机配盖子……魔弹搅拌机……”接着，就在你以为它们是单独销售的时候，广告又说：“用买一套魔弹搅拌机的价格，获得两套完整的21件‘魔弹’套装！只要您现在就下订单，我们还将送您30天的减肥燃脂特效保健品！”

显然，和广告词那噼里啪啦的节奏比起来，这些玩意儿的价值根本没那么高。每一个功能、每一样免费赠品或是三合一套装，都是一轮新的享乐冲击。魔弹搅拌机的“每一发炮弹”打出来，都激得消费者的支付意愿噌噌上涨，直到价格（不管价格具体是多少）看上去恰到好处。

第15章 移动电话资费套餐的秘密

价格变得越来越烦人了。2007年，苹果的iPhone问世，消费者对电话账单的长度感到震惊。匹兹堡的博主贾斯汀·伊萨瑞克（Justine Ezarik）8月的电话账单是拿盒子寄来的——有整整300页。贾斯汀干脆给它拍了一段视频，上传至视频平台YouTube，成了大热门。她的iPhone一连上互联网，电话公司就要对传输数据计费，可使用数据又是免费的。也就是说，这份罗列了上千个项目的账单说的只不过是：您的数据使用费是0.00美元。

在过去的一代，我们大多数人逐渐接受了这样的看法：我们的电话消费、有线电视费、上网费，机票、租车费、酒店费；医疗保险费、汽车保险费、人寿保险费，健康俱乐部和乡村俱乐部会费，信用卡消费和抵押贷款等，是我们永远没法完全搞懂的。运算法则取代了价格。要是你能得到一个简简单单的价格，那可就贵了。



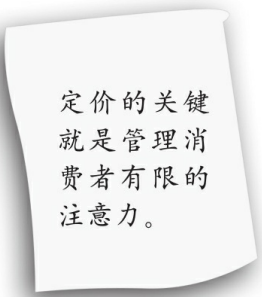
价格只是一个数字罢了。价格下降，销售量就上涨，必定有一个价格X能令利润最大化。

电话账单那么复杂，SKP公司至少要承担一部分责任（说是荣誉归它也行）。他们负责给沃达丰、德国电信、瑞士电信等公司提供定价咨询。如今纷繁复杂的电话账单，是以前景理论为基础演化出来的一套精妙哲学。按普通商学院的说法，价格只是一个数字罢了。价格下降，销售量就上涨，必定有一个价格X能令利润最大化。故此对于解得X，SKP公司的顾问所受的训练则是，要从价格结构的角度去想。孤零零的

一个价格不行，要有一整套的公式（“收费方案”），说明每一项消费行为的费用是多少。

消费者通常都会选择这一公式。表面看来，价格结构很慷慨：“如果你的通话时间较长，这里有一种不限时套餐。”选项更多，意味着可以自由选择，常识告诉我们这是一件好事。其实，消费者既是刀俎，也是鱼肉。由于偏好是根据罗列的选择来构建的，所以额外的选项也意味着受人操纵。较之没有选项的情况，提供额外的计费方案可能让消费者愿意支付更高的价格，或是买得更多，甚或两者皆然。

“最优化”的价格一般指的是更复杂的价格。赫尔曼·西蒙提到了德国铁路采用过的一套成功的宣传手法。他们引入了铁路卡，一种价值400欧元的打折卡——有了这种卡，消费者全年买火车票都能打半价。打折卡本身并不值钱。你不能用铁路卡本身来乘车出行。



定价的关键
就是管理消
费者有限的
注意力。

这张卡值400欧元吗？全看情况。唯一可以确定的是，经常出行的旅客可以节省很多钱。“铁路卡每年有300多万个用户，这一手段大获成功，”赫尔曼·西蒙说，“但只有少数消费者知道跟正常票价相比的盈亏平衡点在哪儿。”

不知道盈亏平衡点在哪儿，正成为后现代的一种现实环境。SKP公司的一份公告说，定价的关键就是管理消费者有限的注意力。公司需要回答如下几个问题：

在消费者眼中，哪些定价因素最为重要？在消费者查看报价单时，他的眼睛会往哪儿看？他们是更注意一次性收费、月租费，还是每次下载的价格、硬件补贴，或者其他什么因素呢？

要为消费者看重的要素定个具有吸引力的价格，好把他们拉拢过来，至于消费者不怎么看重的东西，则可以维持较高的、没什么吸引力

的价格水平。移动电话的定价要素实在是花样繁多：有一次性的开户费、月租费、每分钟的通话费（繁忙时段、非繁忙时段、周末时段），还有结算周期。这样一种复杂的定价挑战，又能给消费者留下什么样的自由度呢？

心理定价中最有力的一种工具是统一费率偏爱。消费者喜欢统一的费率，哪怕他们会出更多的钱。

2009年非营利组织效用消费者行动网络（Utility Consumers' Action Network，简称UCAN）在一份研究报告中声称，圣地亚哥地区的手机用户每分钟通话费平均为3.02美元。这个价格，是用计费总价除以通话的总分钟数得到的。每分钟的通话费这么高，是因为有许多消费者尽管打不了那么多电话，也还是选了统一费率的套餐。

泰勒解释说，这是前景理论带来的结果。电视购物节目把产品切成许多小奖品，相对的，你也应该把所有的损失打包成一大堆。票价90元并不见得比票价30元糟糕三倍，但一次性地出90元却比三天里每天都花30元感觉要好。

由于所有产品的费用都是一种损失，所以把费用弄成统一的会显得没那么痛苦。每个结算周期只支付一次，其他的就不用管了，所有人几乎都倾向于这种选择。“免费”食物是度假游轮的一大卖点，尽管有点儿莫名其妙。度假者知道自己的船票是包含了食物价格的，而且并不便宜，但他们还是感觉它们是免费的。偏爱统一费率，还有助于界定美国中产阶层。美国人喜欢有自己的房子和车子，讨厌租房或搭乘公共交通工具。其实，倒不见得是自己有车比租车更便宜，只不过是租用车的成本更明显。很多美国的城市居民发现，把SUV卖了，去哪儿都打车反倒更便宜。但一想到打车去超市要花15美元，就难免觉得不合理。没有人会喜欢盯着出租车的计价器噌噌地跑。

不少人认为，视频租赁服务商奈飞公司（Netflix）大获成功在很大程度上是定价的结果。它以邮寄的方式提供影碟租赁服务，月租费多种多样，在4.99美元到47.99美元不等。除了最便宜的月租套餐，其他收费方式都允许无限制地租用。要是奈飞公司按次收费，那它就免不了要和传统音像店拼价格。理性的消费者在考虑是否选用奈飞公司服务的时候，必须估计自己每个月会看多少部电影。可学术研究却表明，消费者总是倾向于极大地高估自己对各项服务的使用程度。出于这一倾向，奈飞公司的顾客们大多希望改变自家的收视习惯。有了“免费”的DVD，以及退还时便捷的收方付邮费制度，人人都会看更多的电影——至少会以为自己会这样。电影爱好者可能会得出结论：不管价格到底是多少，奈

飞公司的定价都合理。消费者的支付意愿如此模糊，以至于奈飞公司的价格波动范围非常大，传统音像店却不是这样。赫尔曼·西蒙还有一个例子：

价格实验

某家连锁电影院为消费者提供免费会员卡。该卡记录顾客的访问次数。每个月第一次来看电影，收取一个价；第二次来，收一个较低的价；第三次来，价格更低。这样的定价十分巧妙，它鼓励消费者靠看更多次电影来“省钱”。但顾客根本没省钱。有了上述收费安排之后，电影票销售量增加了22%，平均票价上涨了11%，该电影院的利润增长了37%。

赫尔曼·西蒙说，这样的进步，靠单向度的涨价或降价是办不到的，只有靠经过精心研究的全新定价结构才能做到。

按如今价格顾问的说法，他们的目标是设计出一种价格结构，提炼出每个顾客的最低支付意愿。消费者欲望的柔软线条是什么样儿，定价方案就比照它裁剪成什么样儿。

第16章 折扣券的秘密

折扣券根本没有意义嘛！与其在买东西之后再打折，何不一开始就给个较低的价格呢？这个问题注重实际的消费者问了好多年了。企业和大多数其他人却对此置若罔闻。较之以往，现在市场上的折扣太多了。大约1/3的电脑设备可以打折，20%的液晶电视和数码相机也一样。乘坐你最喜欢的航线，获得免费机票和升级成头等舱的贵宾；使用信用卡，可得现金返还或是航空公司的飞行里程数；汽车有“经销商奖励”；一些房地产开发商甚至向买家提供免费汽车。想在收银台来个折上折，不需要你有多么精明的头脑：使用制造商的优惠券，再掏出积分卡打个折，之后用信用卡埋单，就可以享受几个百分点的提成返还。

早在20世纪初，折扣券就是大热门了。1896年，托马斯·斯佩里（Thomas Sperry）和谢利·哈钦森（Shelly Hutchinson）创办了一家公司，发行S&H绿色酬宾赠物券。这家公司把赠物券卖给商场和加油站，后者在销售自家产品时免费赠送它。消费者把这些赠物券积攒起来，贴在“免费”的册子里，再拿这些册子兑换商品。这就创造了被后人委婉地称为“忠诚度”的东西。消费者不想换别家商场，因为他们想积攒更多的赠物券，好换回一台烤面包机或者体重秤。赠物券在20世纪60年代大行其道，斯佩里和哈钦森的公司印刷的赠物券比美国邮局发行的邮票都多了整整两倍，价值约8.25亿美元。该公司还经营“兑换中心”，这是一家迷你百货商店，这家店不收现金，只收赠物券。这桩买卖在70年代走入低迷期，进入80年代，又被诸如常飞里程数、超市积分卡等当代打折方案所取代。斯佩里和哈钦森创办的公司仍在运营网上购物的“绿色点数”，不过，在如今轰轰烈烈的打折浪潮里，它已经风光不再了。

斯佩里和哈钦森留给当今打折喜爱者的一样东西就是“赠物券综合征”。把赠物券贴在本子上要花好多功夫。美国人有满满几抽屉的赠物券，却从来没时间去兑换。没兑换的赠物券就是S&H公司的纯利润。

两家彼此独立的公司，年轻美国（Young America）和帕拉果（Parago）经手全美大部分的折扣券兑付。在消费圈里，它们的名声也

就稍微比“狗咬狗”的促销员好一些。顾问宝拉·罗森布鲁姆（Paula Rosenblum）告诉《商业周刊》：“这个游戏太明显不过了，只要折扣券不是100%的兑现，那就是白捡钱。”

你不可不知的价格术语

“损耗”和“延误”

折扣券行当里的术语，前者指消费者从来没把打折券寄出去，后者指返还折扣的支票寄还了，但不知道为什么消费者没去兑现。两者都是提高利润的强力发电机。

从理论上讲，折扣券没兑换，处理机构就无法从中得利。但它们的客户却能。“TCA兑换服务”的处理机构吹嘘说自家发出的兑现支票少，实际兑现的支票比例更是极低——10美元的折扣券，兑现率不到10%。“如果你使用其他服务公司，兑现费至少要上涨20%。”TCA的宣传小册子上说。（TCA把客户清单卖给了帕拉果，而帕拉果又推翻了上述说法。）

按行业惯例，兑换折扣券要有商店的收据和流通价格。要把包装盒上的条形码剪下来（这样一来，该商品就无法退换了），还得完完全全、一丝不差地填完申请表格。稍有疏漏，兑换机构就会做“进一步研究”，要你填更多的表格和文件，并把你的情况转入“特别小组”。有人开脱说，这是为了预防欺诈，可它也令不少消费者打消了兑现折扣券的念头。该行业提高“延误”率的手法之一是，用没有任何标记的信封邮寄支票，把自己打扮成垃圾邮件的样子。猜猜看，这些支票最后到哪儿去了？不消说，当然是被消费者扔进了垃圾堆。

听起来挺滑稽，可折扣券始终是一笔大生意。最近的一项估计显示：

商家每年大约要提供4亿美元的折扣券，面值为60亿美元。只要从这个数上稍微砍几个百分点下来，那就是九位数了。据说，有40%的折扣券从来没被人收集。

这大概能够解释公司何以这么热衷提供打折券项目。但还有一个更棘手的问题：为什么消费者对折扣券那么着迷呢？实验和实践都证明，折扣券抛出了一道魔咒。人们更倾向于用200美元买一台打印机，得到25美元的折扣券；而不是直接买标价175美元的同款打印机。

泰勒解释说，折扣券就像是一种心理物理学上的套利。首先，200美元好像比175美元高得并不多。但打了折和没打折，在心理上的差异可就大了。大多数人喜欢打折。泰勒把它叫作“一线希望”原则。他用以下问题做了演示：

A先生的车在停车场给人弄坏了。他要花200美元修车。汽车受损的同一天，他在办公室同仁发起的足球博彩中赢了25美元。

B先生的车在停车场给人弄坏了。他要花175美元修车。请问，谁更郁闷一些呢？

绝大多数的康奈尔大学学生认为B先生更郁闷。尽管经济上他并没有比A先生多出钱，可他错过了25美元的天降横财。

用常识来解读折扣券，有一句俗话说对了：天下没有免费的午餐。每种提供折扣券的产品，都会因为要打折而卖得更贵一些。这不会对销售造成影响。由于消费者对价格充其量只有个模糊的概念，所以他们会从标价里寻找线索，以为200美元的打印机比175美元的那种要好——哪怕前者唯一的额外“功能”就是能打折。

第17章 浮云般的定价依据

花上大概2 400美元，任何一个经营加油站的人都能买到一台甚为荒唐的机器。不，何止荒唐，说它是史上最厚颜无耻的诈骗也不过分。这台机器一头卖“空气”，另一头卖“真空”。

一家制造商的网站上说，真空抽气机“寿命长，维修费低，能为您带来更多利润”。空气和真空的价格显然跟该机器的推销成本没多大关系，反倒跟心理学息息相关。我们为空气掏腰包，因为我们逐渐接受了你得为它埋单的概念，不管它的价格是个什么样。这种糊里糊涂的消费态度让营销人员如登天堂。

我们其实早就以各种各样的方式为空气掏了腰包。就拿电池来说，你买的是电池的寿命：你可以用满电的数码相机拍多少张照片；你多久就得给烟雾报警器换电池，手电筒能坚持多长时间直至没电。可电池包装上哪儿也找不到相关的说明。你只能看到电压的标注，而电压这种度量标准只是意在限制电池的用途。这就好像你去加油，油泵上只告诉你汽油的等级是多少，却不说明你加了多少升油。

要是各种品牌的电池寿命都一样，那倒也无所谓，可惜不然。2008年的《消费者报告》测试了13种一次性5号电池在数码相机里的寿命。最好的电池能拍637张照片，最差的只能拍95张。一些电池比另一些含有的电解液更多，可消费者只能靠猜来选择。

图17-1列出了12种测试电池的比较情况（作者把《消费者报告》（*Consumer Reports*）测试的一种锂电池给省略了，它比普通电池贵得多，寿命约为后者的4倍）。电池（单位是“对”，也就是两枚）的价格列在底部的横轴上，而它们照的照片数则列在左边的纵轴上。要是出什么钱就买什么质量的东西，那么黑点应该呈直线排列，从左下方斜到右上方。可这些黑点的排列毫无规则，乱七八糟。平均来看，所有受测电池差不多能拍150张照片，但这跟价格不挂钩。最便宜的电池（超市里卖的一种自用品牌“柯克兰”）寿命最长。

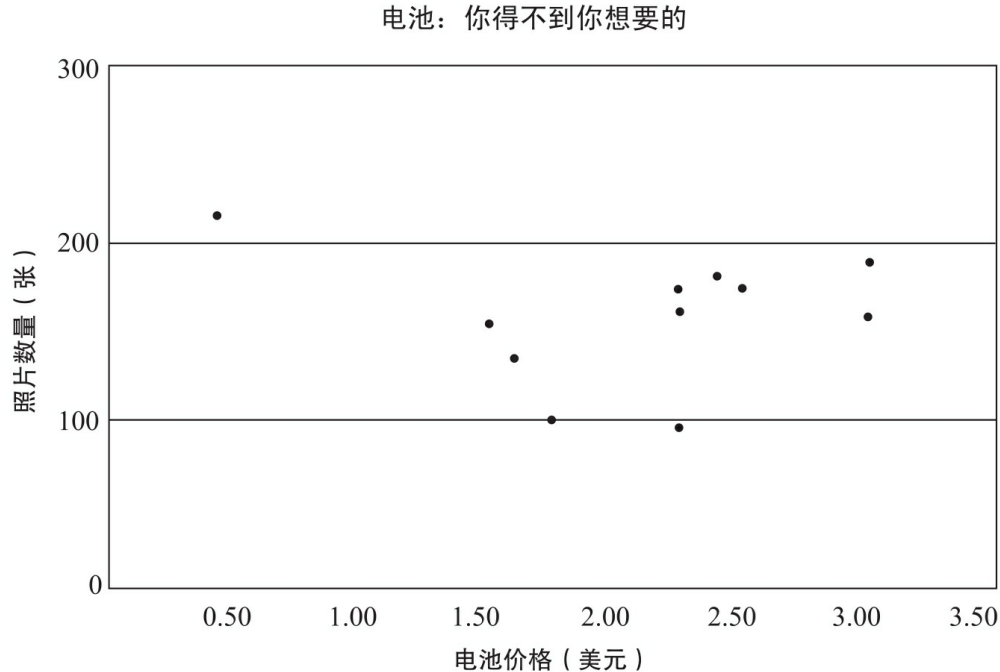


图17-1 12种测试电池的比较情况

这下可好：整整一货架的电池摆在你眼前，你却记不得《消费者报告》里的结论了。该怎么选呢？我对自己说，虽说没法判断电池的寿命，但我们总可以判断价格吧！我倾向于购买最便宜的电池，可我也超喜欢有折扣。看到金霸王电池或劲量电池的价格几乎降到了我通常买的廉价货的水平，我会买它，就像普通的美国人所做的那样，总觉得自己该买在电视上打了广告的真正品牌。我为这种做法找借口说，劲量电池的那些广告肯定有点儿道理，而且我给的价钱并不比那些没名气的牌子多。可劲量和金霸王的管理者们正希望我这么想。

电池并不是唯一一种你难以判断质量高下的产品。顾名思义，洗衣液多少都得“掺水”。你知道自己买了多少水，却不知道自己买了多少肥皂浆。近来，制造商们已经开始说自家品牌的洗衣液是两倍浓缩的了。可他们不说是相对于什么浓缩的。

香水、酒吧里的酒精饮料以及所有装在喷雾压缩罐子里的东西都是这样。对不经常购买的耐用消费品，情况几乎也一样糟糕。消费者很难知道一台冰箱、热水器或者传真机的性能如何、能用多久。我们一辈子只买得了几次，每一次去选购，它们的品牌、型号和功能都会变得面目全非。倘若消费者有责任心，愿意为节能家电支付高价，那他更会头大如斗。2008年，美国消费电子产品协会（Consumer Electronics Association）公布的调查数据显示，89%的消费者有意在选择下一台电

视机时考虑节能问题——同时超过半数的人承认自己搞不清楚节能标签的意思。

目前正在进行的最大骗局大概要算手机短信了。手机短信所谓的市场价格，跟带宽或其他任何技术层面的东西都没有关系。它是由多少消费者（或消费者的家人）入了圈套愿意给它付费决定的。

短信是一个非常小的带宽包。它仅限于160个字母，每个字母需要一个字节。跟它相比，多媒体信息服务（彩信）或电子邮件可以插入图片，占用几兆字节。SKP公司调查发现，消费者认为彩信的价值是文字短信的3.5倍。可按数据来衡量，普通彩信要比文字短信大100万倍。

对照单点菜付费的手机用户来说，传输数据的零售价在1兆字节1美元左右。按这个费率，要是一条短信含10个字母，对应的价格应该是0.001美分才对，再来个四舍五入，那就只有0美分。

其实，0.001美分这个数都极度夸大了一条文字短信的真实成本。和电子邮件、互联网及语音数据不同，文字短信是附加在手机网络上的。它们占用的是原本供网络维护的受控频道，这个频道如果不拿来发短信的话，就根本没别的用处。

由于消费者对文字短信的成本根本没有概念，所以他们只好从电话公司找线索。短信业务计划大获成功。从2005年至2008年，美国承运商的短信价格翻了一倍，从10美分涨到了20美分。这期间，短信通信量增加了大约10倍。

第18章 你真的相信便宜和更便宜吗

“廉价”的英文单词在越途网（CheapTickets）网站的主页上出现了45次。我可以亲自证明越途网商标的催眠魔力。我真的相信他们的票价更便宜吗？不……呃……说不清。

航空公司是率先倡导区别定价的——基于不同的支付意愿，对不同的消费者收取不同的价格。在美国航空公司担任过CEO的罗伯特·克兰德尔（Robert Crandal）曾说：“要是某条航线上我有2 000名旅客和400种价格，那我显然还少了1 600种价格。”互联网能让消费者快速、便捷地进行价格比较，本来事情应该变得容易。可它并没有起到这个作用，越途网就是个很好的例子。它的网站不显示美国西南航空公司或捷蓝航空公司等廉价航线的票价。由于这些航空公司素来有着票价最低的名声，越途网的承诺就成了空头支票：你得到了最低的票价，可它来自最昂贵的航班。

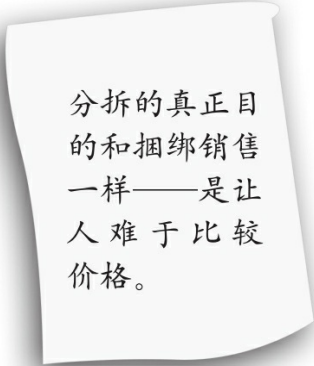
也难怪不少人会直接上美国西南航空公司或捷蓝航空公司的网站。我刚才核对了从洛杉矶到菲尼克斯一线的票价，这是美国西南航空公司最繁忙的一条航线。美国西南航空公司网站上最低的往返票是98美元另加税费。在越途网上，英国和美国的航空公司也为同一行程提供了相同的票价，98美元另加税费。

廉价承运商的收费确实更便宜，可惜它们也有不便宜的时候。有时，它们的价格比普通航空公司的最低票价要高得多。这些更贵的飞机票是廉价承运商赚钱的关键。讽刺的地方就在这儿，由于美国西南航空公司和捷蓝航空公司建立起了票价最低的好名声，所以它们能够收取更高的费用。便宜是相对的，它取决于环境。这是美国西南航空公司和捷蓝航空公司不在主流旅行网站上的一个重要原因：它们宁愿顾客们不做比较。

其实，所有的航空公司都是这么想的。这是一门有趣的生意：大多数旅客会根据价格和时段任意选择一趟航班。（想想看，酒店行业要是也像这样会是怎样一番情形：旅客只按如家连锁酒店的价格付钱，多一

个子儿也不肯出。)机票的价格敏感性令航空公司采用了分拆的做法:托运行李收钱,充气枕头收钱,餐食收钱,咖啡收钱,电话预订收钱,纸质机票收钱,选择座位收钱,还有其他所有惯例是免费的设施,只要你用就都收钱。“三四年前,航空公司厌倦了自己的机票像菜市场卖的大白菜一样收费,所以它们设计了一套策略,好让自己的价格不那么透明。”美国机票比价网站FareCompare的CEO里克·西恩尼(Rick Seaney)说。率先搞分拆的承运商是欧洲的航空公司。在美国,到2008年5月,航空公司报复性地开始对首件托运行李收取15美元的费用。愤怒的旅客立即咬牙切齿地说再也不飞去美国了。可他们的决心没能坚持太久。随着其他航空公司也开始加收托运行李费,并为从前免费的设施收费,这种决心马上土崩瓦解了。

消费者把分拆服务看成不相干的琐碎小事,以为这部分费用是净利润。实际上并非如此,至少竞争航线上不是这么回事。分拆的真正目的和捆绑销售一样——是让人难于比较价格。附加费用差异很大:一家航空公司托运行李最便宜,另一家则提供划算的充气枕头和软性饮料,还有一家允许你免费打电话预订机票。要比较一趟行程的真实价格,现在的收费项目太多,却又没有方便的电子表格(有些网站能帮上点儿忙)。但大多数旅客的做法正中航空公司的下怀:他们把杂费抛到一边,单纯根据某种东西来做出选择——我说的“某种东西”是什么都可以,反正不是最低的价格。



分拆的真正目的和捆绑销售一样——是让人难于比较价格。

第19章 神奇数字9

“我来告诉你什么叫广告的辉煌，”罗杰·斯特林（Roger Sterling）在电视剧《广告狂人》（*Mad Men*）里说道，“0.99美元。”调查声称，大约30%~65%的零售价格都是以数字9结尾的。不管在什么数量级，它都站得住脚。有时，9的单位是千或者万，有时它代表分，而在汽油上，它指的是厘。苹果公司的史蒂夫·乔布斯坚持iPod下载歌曲的价格是0.99美元（视频下载则为1.99美元），被誉为天才。2009年，苹果公司发了慈悲，歌曲下载新增了0.69美元和1.29美元这两种价格。

这一现象的典范，是九毛九杂货店[\(19\)](#)。20世纪60年代，大卫·戈尔德（David Gold）在洛杉矶开了一家酒品店，想清空店里那些走货速度太慢的低价红酒。他贴了一张横幅，上面写着：“红酒世界。你的选择：0.99美元。”这招超管用，只要是0.99美元的东西，顾客们便照单全收。

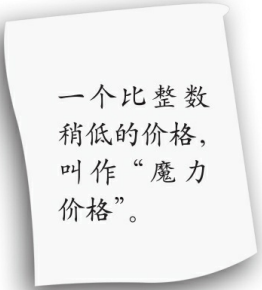
有趣的是，红酒以前的标价在0.79美元到1.49美元不等。“原价0.79美元的，标成0.99美元后卖得更好；原价0.89美元的，标成0.99美元后卖得更好；当然了，原价1.49美元的，标成0.99美元也卖得更好。”戈尔德说。九毛九效应如此惊人，戈尔德禁不住开玩笑说，真该开一家商店，里头所有的商品全卖这个价儿。

笑话成了现实，1982年，戈尔德开了第一家九毛九商店。该商店目前约有277家分店，受它的鼓舞，从东海岸到西海岸，起着类似名字的商店雨后春笋般地冒了出来。要是有人没看店标就走进一家典型的九毛九商店，那他保准会犯嘀咕。店里陈列着方便面、长筒袜、扑克牌、洗涤剂、万圣节道具服、女性卫生用品、金属丝和软糖。样样东西都没什么联系，而且样样东西都不怎么靠谱。

2008年，《纽约时报》上有一篇文章调查了纽约地区九毛九杂货店的繁荣状况（没有一家是戈尔德九毛九商店的分店）。哈勒姆区（Harlem）最繁忙的商业地带弗雷德里克·道格拉斯大道上有“新富塔九毛九商店”，还有竞争对手“巴布九毛八折扣店”。9的魔力大大方方地四处洋溢。招牌上承诺商品只卖0.99美元或0.98美元，或是各种以9结尾的价格。布鲁克林的迪特玛斯大道上开着大把的“五毛九、七毛九、九毛

九”和“六毛九、八毛九及九毛九”商店。

在为数众多的仿效者里，规则难以辨别，而且随时都在变。“0.99美元的承诺越来越像是空头支票了，”《纽约时报》得出结论，“这些商店堕落成了下诱饵的骗子，靠招牌把顾客引诱过来，实际上却是为了卖那些更贵的东西。真正摆在0.99美元货架上的只有廉价的发夹、闪光贴纸和一小卷粗糙得更适合打磨家具的厕纸。”



一个比整数
稍低的价格，
叫作“魔力
价格”。

通货膨胀不断冲击着这种商业模式，这早就不是什么秘密。1982年卖0.99美元的东西，现在得卖约2美元。对正宗的九毛九商店而言，2008年是它的转折点。经过多年的变通，比如只卖半打鸡蛋、牛奶盒子越缩越小，它终于撑不下去了，把最高价格提到了99.99美元。对总裁杰夫·戈尔德（Jeff Gold）来说，做出这样的无奈之举，简直就像是家里有人过世了那么难过。“99是神奇的数字——如今我们驶离了它的轨道，对谁来说都绝对不轻松，”他说，“做出这样的改变，我心里很难过。”

一个比整数稍低的价格，叫作“魔力价格”。这通常意味着价格以9或者99结尾，但98和95同样被看成是魔力价格。没人知道这种做法始于何时、何地，为什么会出现。有种理论提到了英国硬币。直到南北战争，美国的分币都很稀少，美国国内流通着英国先令和英国便士。纽约的商店经常标示两种价格：英国货币价和美国货币价。把先令转换成美国的分币，往往会弄出一个奇数来。于是，以奇数结尾的价格就这么跟英国进口货拉上了关系，当时的美国人觉得英国货比美国货要好些。精明的商家便把国产货的价格也标成奇数，好让人觉得它沾染了些英国味儿。

一个稍微可信一点儿的故事扯上了收银机。

价格实验

1879年，酒吧记账员詹姆斯·瑞提（James Ritty）发明了第一台收银机。瑞提知道，要拿酒吧的酒品库存核对现金收入几乎是不可能的，他对自家酒吧柜台上的员工也不那么信得过。所以，他设计了一种机器，要员工输入价格才能打开钱匣子。员工这么做的时候，会有铃铛声响起，提醒老板。在午餐时段，老板应该能听到源源不断的铃铛声，要是什么时候耳边清静下来，他便很有必要去查一查。瑞提设计的机器还能记录输入的金额，拿记录核对现金也相对简单。梅西百货是采用收银机的第一家大型商店。由于梅西百货的价格多是整数的“元”，所以他们开始采用奇数价，强迫员工输入价格，打开抽屉找零钱。实际上，如图19-1所示，梅西百货刊登在《纽约时报》上的广告打从19世纪90年代就出现“魔力价格”了，广告上显示的价格有一大半都是以9结尾的。

R. H. MACY & CO.,

SIXTH AVE., 13TH TO 14TH ST.

UPHOLSTERY.

(Third floor—reached by elevators.)
IMPORTED CHENILLE CURTAINS.
 In Red, Cream, Olive, Salmon, and Copper, figured all over,

\$5.99; worth \$8.
 6.99; worth 9.

Imported FRENCH TAPESTRY CURTAINS
 and PORTIERES, double faced, assorted colors,
 \$7.79, \$7.99, and \$8.49; worth \$10 to \$15.

FURNITURE COVERINGS.
 Rich Jute TAPESTRIES, (Renaissance),

54c.; worth \$1.00
 Medallion TAPESTRY, \$1.49; worth 2.50
 French BROQUETTE, 3.49; worth 5.00
 Titian VELOUR, (6 colors,) 1.79; worth 3.00
 All-wool TERRY, (6 colors,) 99c.; worth 1.50

Large variety of WOOL PANELS, for Wall Decorations, at very low prices.

Silk PLUSHES, \$1.26; worth \$2.00
 Crushed Mohair PLUSHES, 99c.; worth 1.50

SILK SLUMBER ROBES,
 \$1.99 each; worth \$3.50.

1,000 SMYRNA RUGS, 16x31 in.; 61c. each;
 worth \$1.25.

1,000 Japanese Painted and Embroidered PANELS,
 Painted—

16x53 in. 24x24 in.

16c.; worth 25c. 27c.; worth 50c.

Painted and Embroidered—

24x24 in. 19x50 in.

49c.; worth 75c. 54c.; worth \$1.

A FEW ODD PIECES OF FURNITURE, SUCH
 AS TOILET TABLES, MUSIC AND WRIT-
 ING CABINETS, DIVANS, &c., AT A RE-
 DUCED OF 25 PER CENT. FROM
 FORMER PRICES.

ORIENTAL GOODS.

(Third floor, reached by elevators.)
 A LARGE VARIETY OF RARE COLORINGS
 AND DESIGNS IN TURKISH RUGS
 AND CARPETS, ALSO CHOICE SPECI-
 MENS OF ORIENTAL EMBROIDERIES
 AT OUR POPULAR PRICES. THESE
 GOODS ARE FROM 20 TO 50 PER CENT.
 LESS THAN IS USUALLY CHARGED BY
 OTHER HOUSES.

CLOSING OUT AN ODD LOT OF ORIENTAL
 RUGS AT \$4.99, \$7.99, AND \$17.49, RE-
 DUCED RANGING FROM 20 TO 30 per
 cent.

PLUSHES.

WE HAVE NOW ON SALE A COMPLETE LINE
 OF THE CELEBRATED LISTER'S ENG-
 LISH PLUSHES AT VERY ATTRACTIVE
 PRICES.

24-in. Plush at 97c. per yd.; is worth \$1.25.

" " " \$1.29 " " " " 1.50

" " " 1.49 " " " " 1.85

" " " 1.59 " " " " 2.00

" " " 1.99 " " " " 2.80

**THREE SPECIAL BARGAINS IN 52-IN. SILK
 SEAL PLUSHES,**

No. 1, \$5.69 per yd.; worth \$7.

No. 2, 7.99 " " " 10.

No. 3, 11.49 " " " 15.

LACES.

Special sale of Black DRAPEY NETS,

39c. per yd.; worth 50c.

Striped SILK GAUZE, all colors, for evening wear,
 45 in. wide,

94c. per yd.; worth \$1.25.

Job lot of ORIENTAL LACES, 1½c. per yd.; worth
 25c.

Large variety of LACE VEILINGS, in all colors,
 12c. per yd.; formerly 24c. to 69c.

Full assortment of Real and Antique LACE TIDIES,
 8c., 14c., 18c., 31c., up to \$5.88.

BLACK SILK.

SPECIAL SALE OF SU-
 PERB GROS GRAIN, RHA-
 DAME, and ROUND CORD
 FAILLE FRANÇAISE, 24
 INCHES WIDE, IN
 LENGTHS VARYING FROM
 10 TO 20 YARDS, MADE
 FOR TRAVELERS' SAM-
 PLES, AND USUALLY RE-
 TAILED FROM \$1.50 TO \$2
 PER YARD.

AT THE UNI-
 FORM PRICE
 OF

99c. PER
 YD.

图19-1 1890年11月2日梅西百货的一则广告

不管是英国硬币还是收银机，都解释不了神奇数字9。1先令价值

0.125美元，所以只能带来类似0.125美元、0.25美元、0.375美元一类的价格。它们都不是以9结尾的。至于说收银机的解释，只要一个价格不是偶数元，就都需要找零。

不管到底是怎么开始的，总之魔力价格得到了广泛的应用，不只在美国，不只是营销专业人士，也不只是对便宜的东西。在eBay上，以9结尾的价格无所不包，二手转让的三卧室殖民风格自住房，标价是59.9万美元。除了房地产，我碰到的最昂贵的魔力价格是罗迪欧大道路易威登专卖店里的镶钻手表：14.9万美元。他们是真的以为这会比15万美元整更好卖吗？这只手表是墙壁展示柜里最昂贵、最显眼的东西。奇怪的是，路易威登稍微便宜些的手表，却并未标以魔力价格——有一只表标价是7 540美元。

除了价格以9结尾（加上0），还可以在9的右边加上非0的数字，比如19.7万美元，3.95美元。对餐厅顾问布莱顿·奥德尔来说，后一种价格是他最讨厌的东西。“他们完全可以定成3.99美元，”他说，“这两者对消费者来说几乎没有区别，可那是4美分呢！”在餐饮业，一份订单多4美分，加起来可就颇为可观了。

快餐行业把魔力价格看得就跟自己的亲生骨肉一般，甚至把它们当成进行自嘲性营销活动的道具。2008年，塔可贝尔（Taco Bell）总裁格雷格·克里德（Greg Creed）致信歌手50美分[\(20\)](#)，要他把名字改成“八毛九”或“九毛九”，以便给该连锁店的低价搞搞公关。乐手的回应是提起诉讼，为这些毫无魅力的数字索赔400万美元——于是，双方都获得了大把的免费曝光。

魔力价格开辟了心理定价这一研究领域。1936年，哥伦比亚大学的伊莱·金兹伯格（Eli Ginzberg）对他所谓的“惯常价格”发表了一份笔记，篇幅只有一页纸。“多年来，我国引用的零售价格总是比十进制单位少那么一两美分——比如0.49美元、0.79美元、0.98美元、1.49美元、1.98美元，可惜这是个编出来的故事。”金兹伯格报告了在一家大型零售商处进行的非正式实验。该公司的好奇心颇强，它把产品目录印了多个版本，有些使用按9结尾的惯常价格，有些则采用相对应的整数价。

据金兹伯格所说，结果“很有趣，但也叫人困惑”。有些产品以魔力价格卖得更好，有些却卖得更糟了。他简短的文章并未提供统计细节。“主管销售的副总裁大胆猜测，收益把损失扯平了。他完全清楚地意识到，重复实验或许能带来更确定的结论。可因为事关金钱，哪怕是

大胆的商人，也会克制住自己的实验热情。”他说。

近半个世纪以来，许多知情人士认为，魔力价格只是个无伤大雅的迷信罢了，没什么坏处（自然也没什么作用），但这拦不住零售商使用它们。到了20世纪80年代，卡尼曼和特沃斯基的革命又把心理定价的热情给点了起来。1987年至2004年发表的8篇论文报告显示，较之相邻的整数价，魔力价格平均提高了24%的销售量。

别对这个数据太当真。销量增长的百分比差异可大了，有些几近于无，有些却高达80%。以芝加哥大学的埃里克·安德森（Eric Anderson）和麻省理工学院的邓肯·西梅斯特（Duncan Simester）所做的实验为例。他们找到一家邮购公司，把产品目录印刷出多个版本。该公司销售价格适中的女装，通常使用以9结束的整数价。两人选中了一个39美元的东西进行测试。在实验版本的产品目录（如表19-1所示）中，同一产品的标价分别是34美元和44美元。依照公司的邮件列表，这些目录寄送到了随机抽选的规模相当的地区。

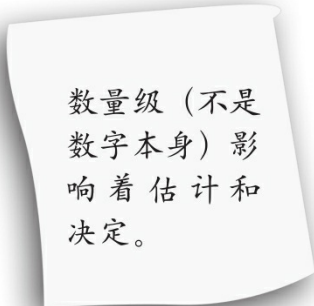
表19-1 产品目录

价格（美元）	卖出数量（个）
34	16
39	21
44	17

魔力价格39美元的销售量比其他两种价格的都多。最重要的一点发现是，为39美元埋单的人比为34美元埋单的人多。魔力价格带来了更大的销售量，单笔交易的利润也更高。

这正吻合资产负债表上的数据。2002年，《福布斯》杂志得出结论，九毛九商店的毛利润达到了惊人的40%，比沃尔玛高两倍。平均而言，该连锁店卖0.99美元的东西，成本仅在0.60美元上下。

举个典型的手法：大卫·戈尔德把内衣品牌“水果织布机”（Fruit of the Loom）公司拍卖的70万包《星球大战前传：幽灵威胁》（*Star Wars: The Phantom Menace*）主题内衣统统买下，等到下一集《星球大战前传：克隆人的进攻》（*Star Wars: Attack of the Clones*）电影上映时再卖掉。不明内情的购物者或许会想，怎么这内衣的主题是前一集“幽灵威胁”，而不是现在正播的“克隆人进攻”呢？但0.99美元的价格让所有的疑惑都烟消云散了。



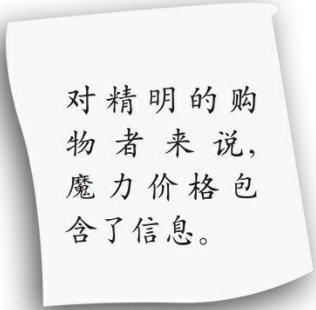
数量级（不是
数字本身）影
响着估计和
决定。

魔力价格为什么能发挥作用呢？因为购物者肯定是顺着整数看起来，省去零头，注意力总是集中在第一个重要数字上。比如29.99美元这样的价格，在心理上被归入了20多美元的范畴，而30.00美元（或以上）的价格，则被看成是30多美元的东西。20多美元比30多美元似乎低得多。

营销和心理学界对这一解释争论得很激烈。魔力价格其实提出了一些有关意识如何运作的有趣问题。在数量级这条无尽的高速公路上，数字不过是随意标注出里程而已。大脑对数量的含义有深刻的理解吗？又或者，它只是以肤浅的方式处理数字？大量心理研究暗示，人，哪怕是小孩子，都对数量级有着相当不错的认识。他们完全明白29只比30小一点点。锚定实验还表明，数量级（不是数字本身）影响着估计和决定。

可单单是精神上的四舍五入，还无法解释安德森和西梅斯特得出的结果。要是购物者只注意第一个数字，照理说，34美元和39美元都应当被理解成30多美元。这两个价位的销售量会大致相当。可实际情况却是买家更倾向于买39美元这一更高的价格上的东西。9的确是一个神奇的数字。

另一种理论认为，魔力价格传递出“价格已经打过折了”的信息。以前，有个小镇上的加油站1加仑(21)汽油卖20美分。不久，街对面开了个新加油站，把价格减了1美分：每加仑19美分。头一家加油站为了报复，又减了1美分：18美分。对很久以前价格战的文化记忆，或许能让我们把19这样的数字跟竞争定价联系起来，把类似20这样的整数跟垄断和高价联系起来。毫无疑问，诸如此类的事情现在都有。哈勒姆的“九毛八折扣店”起这个名字，就是为了抄“九毛九商店”的底，但它很快就迎来了一家叫“九毛七商店”的竞争。



对精明的购物者来说，魔力价格包含了信息。

对精明的购物者来说，魔力价格包含了信息。洞悉餐馆或酒店意图（有时还要加上质量）的一个好办法，就是看它的价格是以整数结尾，还是以“.99”或“.95”结尾。诺斯通百货商店就声称自己不使用魔力价格，还做了解释。大意是说：“我们不是沃尔玛，顾客来这儿图的就是质量好，理应为此掏腰包。”这或许可以解释为什么有时魔力价格不管用。价格顾问弗兰克·卢比为一家汽车制造商提供服务，这家厂商原想把旗下的一款车定价为19 999美元。卢比的研究发现，要是价格在2万美元以上的话，这车会卖得一样好。大概是因为买车的顾客们不愿意觉得自己在买一辆“廉价”车吧。一些零售商，比如艾迪·鲍尔（Eddie Bauer）和J.克鲁（J. Crew），只在残次品的价格上使用“.99”的结尾。开市客用“.97”的价格示意某商品打折或滞销。对知道这套代码的人而言，魔力价格就像是吹响了集结号，更不必说还有下意识回应一类的规律。当然，这方面的事情，店方可不希望消费者搞得太清楚。

在安德森和西梅斯特的实验里，外套标价34美元或44美元，销量并没有显著差异。这进一步证明，买家没有强大的内在价值感。刺激销量的，是39美元这个价。有人提出假设，从心理比较上看，魔力价格显得比整数价格要便宜。

邮购公司有个习惯，把正在销售的品目列出新老价格：“原价X，现价Y。”研究人员要它们重新印刷一些商品目录，不在销售价里提示打折信息。如你所料，他们发现，销售价加上前后对比时，销售量更高。这也就是说，要是产品目录不显示，买家就不知道Y是打折价。

对比标注的售价比魔力价格更能有力地带动销量。请看图19-2中两种价格标示方法。采用左边标示方法的商品，消费者购买它的可能性会更大，哪怕右边是个魔力价格。

原价 \$ 48

\$40 **\$39**

大减价

图19-2 消费者倾向选择的价格

安德森和西梅斯特把上述两种手法结合到了一起，采用了对比标示的魔力价，比如“原价48美元，现价39美元”。这发挥出了最强的效果。只不过，效果好得也并不太多，只比单列售价高出了一点点。这或许意味着，销售价格和魔力价格利用的是同一种心理原理。魔力价格本身似乎就在暗示一种并不存在的折扣，这么说吧，它像是用哑剧的形式来表现一堵玻璃墙，以无形再现无形。价格的受众对虚拟折扣和实际折扣做出了同样的反应。

以下事实支持了这种解释：对商品目录以前没刊登过的新产品，魔力价格的效果更大。顾客对新产品的价值概念最弱，于是更多地依赖价格线索。

喜欢打折货（我是说，真正打了折的商品）没什么可奇怪的。可标价19.99美元就意味着它是从20.00美元减下来的吗？就算标价19.99美元的东西真是从20.00美元减价下来的，0.99美元也比1美元少了1美分，打了1%的折扣（如图19-3所示）。

原价 \$ 1.00

99¢ = 99¢

大减价

图19-3 打折价

以理性标准来看，这太微不足道了，不足以给行为造成很大的影响。然而，这符合消费者选择和权衡对比研究。在有許多难以评估的选项时，人的注意力是漂浮不定的。它会寻找易于对比的项目。要是有的选项明显比另一种更优越（哪怕差异很小），那它就能轻易地吸引到关注。想象中的整数价格成了“.99”价格的陪衬，给它洒上了一轮莫名其妙

的诱人光辉。

第20章 分不清的免费和低价

终极折扣是“免费”——就像“免费赠品”一样。“零价格”是营销工作者的宠儿，它能触发一些独特的心理。阿雷利、尼娜·马扎（Nina Mazar）和克里斯蒂娜·尚姆潘尼尔（Kristina Shampanier）做过一次实验：

价格实验

他们设了一个小摊，供应好时巧克力和瑞士莲巧克力。就算你不是巧克力迷，也该知道好时巧克力是一般的货色，瑞士莲巧克力则要好一些。他们对好时巧克力索价1美分一块，瑞士莲巧克力则为1.5美分一块。大大的标牌上写着：每位顾客仅限一块。

在所有购买者中，73%的人选了瑞士莲巧克力。虽说这有点儿对不住好时巧克力，但也没什么好奇怪的。接着，他们把两种巧克力的价格都减了1美分。瑞士莲巧克力卖1.4美分，好时巧克力则免费（每位顾客仍仅限一块）。此举立刻扭转了人们的偏好。69%的顾客选择了免费的好时巧克力，只有31%的人买了瑞士莲巧克力。

阿雷利和同伴们卖的瑞士莲巧克力，价格只相当于批发价的一半。为了获取一颗自己不怎么喜欢、说不定只值1美分的免费糖果，大多数顾客放弃了对半的折扣。

阿雷利认为，这主要是确定性效应造成的。只要是买东西，买家都要承受后悔的风险。我买的巧克力味道不像想的那么好；说不定从别处买更便宜些；它热量太高，我正在节食……免费的东西却不然。你没法后悔把钱花在了免费的东西上，因为你根本就没花钱。由于高估了确定性，我们也就高估了免费的东西。

据说，心理物理学的量值量表包含着有意义的0。在声音响度的量表上，宁静的“声音”应该是0。可实践起来，事情没那么简单。把真正的沉默和勉强能听得到的声音区分开来，可是一项蛮难的挑战。人们往

往会基于暗示性力量说自己经历过（或未经历）某事，并捏造一致性。

价格量表大体上也一样。消费者根本识别不出来真正的没价值。在低价值物品当中，有一片颇大的混乱地带，说不清什么东西值得出钱，什么东西不值得。2006年，阿雷利、鲁温斯坦和普雷莱克发表论文《汤姆·索亚和价值构建》（*Tom Sawyer and the Construction of Value*，这篇文章如今已经很有名），对此进行了证明。

文章的标题暗指美国文学中的经典偏好逆转案例。马克·吐温笔下的骗子兼英雄汤姆·索亚，受命完成粉刷围栏的讨厌活计。汤姆更希望让别人来做。为达到这一目的，他假装享受这份工作，弄得朋友也想一起来找些乐子。他们恳求汤姆让自己来帮忙，至少刷上几刷子。汤姆先是拒绝，最终让了步，条件是，朋友们得付钱给他，获得粉刷围栏的特权。马克·吐温想要讽刺的是，生活里没有绝对的东西，要是有人另外编一套，那他就像汤姆一样的大骗子。

在2006年进行的一次“汤姆·索亚”实验中，研究人员试图吸引学营销的学生去听阿雷利的校园诗歌朗诵会（朗诵惠特曼的《草叶集》）。

价格实验

他们问第一组学生，愿不愿意出2美元去听阿雷利朗诵。同学们非常坚定地回答说：“不干。”只有不到3%的学生说愿意出钱去。

等收集完所有人的回答后，研究人员告诉学生们，其实阿雷利的朗诵会不要钱，问他们愿不愿意收到通知时间和地点的电子邮件。这下，35%的人说“愿意”，他们希望收到通知。

他们向第二组学生提了一个不同的问题：如果我们付你2美元，你愿意去听阿雷利背诗吗？这一回，59%的人说愿意。接着，和第一组学生一样，研究人员又对他们说：其实去听朗诵会是不会收到钱的（忘了那2美元的报酬吧）。此时再问学生们愿不愿意收到通知，只有8%的学生仍旧表示有兴趣。

他们再次调整了问题，找来麻省理工学院的两组学生，问其中一组愿不愿意花10美元去听阿雷利背10分钟的诗，问另一组是否愿意拿10美元的报酬做同一件事。接着，他们又请同一批学生给听6

分钟、3分钟和1分钟的诗朗诵定个价。和烦人声音实验一样，平均价格是跟持续时间挂钩的。可这一次，一组学生给的是正数价格

（也就是他们愿意出多少钱来享受听阿雷利朗诵诗歌），另一组给的是负数价格（忍受听阿雷利背诗的薪水）。整体而言，麻省理工学院的学生们对于去听诗歌该出钱还是该拿钱，根本搞不大清楚。

只要活动是免费的，愿意参加的人就会更多。

第一组学生认为免费朗诵会值得去的有35%——这是对大于零的价值的正面体验。第二组这么想的人只有8%。唯一的区别只在于第一组受了引导，以为诗朗诵得花钱才能去听；第二组则以为那是一件苦差事，得花钱才能请人去听。

汤姆·索亚实验驳斥了所谓“每一种体验都可以归类为积极或消极”的常识。是的，这世上有可怕的经历，也有辉煌的经历。可大多数的体验是好坏参半的。去巴黎旅游是件好事吗？哦，当然，人人都立即表示肯定。这是因为所有人都那么说，也因为它要花好大一笔钱。假设去巴黎旅行免费，而且自此以后永远免费。那你这个周末会去吗？下个周末呢？

汤姆·索亚天真的小把戏成了21世纪的头等商业模式，它叫作Web 2.0。靠着尊称为“用户生成内容”的东西，谷歌、YouTube、Facebook和Twitter变成了价值数十亿美元的企业。所有这一切都建立在如下前提下：用户愿意免费做值得的“工作”（新闻、电影、政治评论）。有些人赚了很多钱——我是说有些人，但不是那些忙着粉刷互联网围栏的人。

第21章 房价一定要定得比市值高

玛格丽特·尼尔（Margaret Neale）最有名的一个实验把房地产经纪人们气坏了，连她自己的亲妈也坐不住了。尼尔想看看锚定效应在房地产市场上管用不管用。

1982年，她来到亚利桑那大学，对讨价还价的心理产生了兴趣。她说：“当时的谈判奄奄一息，心理学家和经济学家互相都不说话了。”尼尔把自己埋进了卡尼曼、特沃斯基、希勒尔·艾因霍恩（Hillel Einhorn）和罗宾·贺加斯（Robin Hogarth）的著作里。她意识到，决策心理学可以成为有力的谈判工具。

她解释说：“我们当时的论点是，谈判里没太多可以改动的东西，你面对的形势改不了，它铁板钉钉一样地存在着。我们知道，倘若有未来，人的行为会有所不同。”（也就是说，当他们知道跟谈判对手还会有进一步接触的时候）。“可你谈判的时候，没法选择是否还有未来，也没法选择谈判对手的个性。它就是那样，已经确定了。你能改变的只是自己采用的认知视角。”

“从前，尼尔和我每天都一起吃午饭，”同事格雷戈里·诺斯克拉夫特（Gregory Northcraft）说，“我们坐下来，看到了日常生活和研究工作存在的千丝万缕的联系。”联系之一是锚定效应和房屋价格。诺斯克拉夫特和尼尔分别买了自己的第一处房子。诺斯克拉夫特说：“我们都有这样的体验：最初看到房子的时候，不知道对这房子该怎么想，要等看了报价之后才有个念头，要是价格比较高，我们往往会着重去看它标价之所以那么高的地方，要是房价比较低，我们的着眼点大多会放在能够解释它为什么价格低的地方上。”

他们意识到这就是锚定。他们还知道，对卡尼曼和特沃斯基的发现能否应用到重大经济决定上，经济学家们持怀疑态度。据说，市场力量将规定出合理的价格。

诺斯克拉夫特告诉我：“看待这一点，有两种方式，一是，当信息

很少的时候，试探和偏见会产生重大影响。如果你没有任何其他信息，就只好往自己的魔法口袋里掏一掏，找出个东西来顶事儿。不过，当你进入一个细节丰富的真实环境，那儿有好多其他东西可注意，所以不需要思维上的捷径。二是，当你进入一个细节足够丰富的环境，可用的信息多得压死人。这为试探和偏见登上舞台提供了第二条路线。要是你有太多的信息，它们就帮你筛选信息。”

诺斯克拉夫特和尼尔从美国国家科学基金会申请资金，以便在现实世界里对试探和偏见进行检验。他们勾勒了三个可能的研究方向：房地产、商务谈判和法律判决。等申请一获准，他们就从房地产入了手。

两人的目标是看看锚定效应是否会影响图森市场上真实房产的感知价值。要做到这一点，他们得找个房地产经纪借所房子用在实验里。尼尔让当房地产经纪的母亲给些建议。母亲把人脉关系网的作用渲染了一下。她说，经纪们很乐意结交教师。图森市房地产信托公司的经纪马丁同意两人使用她手里的一处房产。

价格实验

被试是商学院54名大三、大四的学生，还有47名当地房地产经纪。对于房地产专业人士来说，图森市场是他们的衣食父母。他们平均每年购入或卖出16处物业，在本地有着8年以上的从业经验。

诺斯克拉夫特把被试带到家里，让他们随意检查。被试将得到普通买家能接触到的所有信息，比如周围最近出售的房子的清单，一个装着本处房子以及周围所有待售房屋的“多重上市服务”（Multiple Listing Service）[\(22\)](#)表单的包裹。接下来，被试要估算房子的价值。实验里的一个变量是标价。4组被试得知的标价都不一样。

在诺斯克拉夫特开车把被试们带去房子的路上，碰巧来了一场倾盆大雨，被试们拒绝下车。

他们找了个阳光明媚的日子打算再试一次。可惜，数据还没收集齐全，房子已经卖掉了。两人只好请求换一处房子。两所房子的结果类似。前一年，这处房子的鉴定价是13.5万美元，挂牌价是134 900美元。不过，实验里没人看到这个价。被试们听到的4个虚构价格是：119 900美元、129 900美元、139 900美元和149 900美

元。

两位研究者要房地产专业人士和学生以4种不同的方式给房子报价。他们要扮演住房鉴定师，给一个公正的评估值；假装成代理经纪，提出一个合适的挂牌价；把自己想成是买家，提出一个愿意支付的合理价格；最后，扮演卖家，报出一个自己愿意接受的最低价格。所有的4种做法都表现出类似的锚定效应。

诺斯克拉夫特说：“人们往往把科学说成是一种非常系统化、干净、无菌无毒的过程，可这次研究证明，良好的科学并非如此。”

表21-1是买家对合理购买价的估计值。

表21-1 买家对合理购买价格的估计值

估计购买价（平均）		
挂牌价（美元）	学生（美元）	房地产专业人士（美元）
119 900	107 916	111 454
129 900	120 457	123 209
139 900	123 785	124 653
149 900	138 885	127 318

请记住一点，所有这些数字都是针对同一所房子的。对业余组的学生们来说，挂牌价提高了30 000美元（从119 900美元提高到149 900美元），他们对房子价值的平均估计值也提高了将近31 000美元。他们明白买入价会比挂牌价低。但挂牌价每增加1美元，他们对房子价值的估计也就增加1美元。

信任执牌专家的人想必会很高兴：房地产专业人士受虚假挂牌价的影响较小。对专业人士来说，挂牌价提高30 000美元，他们的估计值只提高了16 000美元。挂牌价应该完全不影响专业人士才对。最先说市场决定价值、价值跟卖家没关系的不就是房地产经纪吗？卖家大多不是专家，可能有着完全不切实际的期待。经纪工作的一部分就是了解市场价，（买家的经纪）并指点客户远离价格太高的房产。

职业的房地产经纪怎么可能这么容易犯错呢？诺斯克拉夫特说：“我觉得，在很多领域，有经验的人就觉得自己是专家了，实际上两者的不同之处在于，专家有预测模型，有经验的人则只有不一定能

预测的模型。”

只有在存在反馈的情况下，经验才是管用的。就算房地产经纪把房子以过高或过低的价格出售，也很少会有确凿的证据能够证明他弄错了该产业的价格。

诺斯科拉夫特和尼尔认为，“对这类判断”专业人士或许掌握了更多相关处置方式的知识，与其说反馈纠正了判断本身的精准性，倒不如说它纠正的是对判断过程的描述（以便描述跟惯例相吻合）。就这类判断任务，我们或许能指望专业人士比业余选手做得更好一些，但平均而言，他们的判断是类似的。”

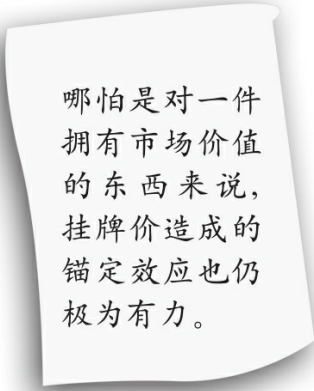
房地产专业人士和学生之间有一个明显的区别。37%的学生承认，他们考虑到了挂牌价。但这么说的房地产专业人士只有19%。诺斯科拉夫特和尼尔顽皮地评论说：“房地产专业人士否认在对房产估价时考虑到了挂牌价，既有可能是因为他们没意识到这一点，也有可能是因为他们不愿当众承认自己依赖了一件不恰当的证据。但目前，这一点还是个悬而未决的问题。”

实验开始之前，一个经纪顾问小组告诉诺斯科拉夫特和尼尔说，的确存在“可信地带”。跟评估价相差超过5%的挂牌价一眼就看出来了，因为它“明显存在异常”。实验的两个中间价格（129 900美元和139 900美元）刚好处在这个区间之内。它们跟评估价刚好差了近4%。而最高和最低的两个价格，跟评估价相差近12%，房地产专业人士理应拉响警报器。

可惜没人这么做。跟更可信的挂牌价139 000美元相比，经纪们在看到异常挂牌价149 900美元时，对房子价格的想法提高了将近3 000美元。学生们的估价更是高了将近15 000美元。

“这里的问题是，决策过程的弹性到底能有多大，又是否存在某种现实的制约，能控制这类过程受影响的程度，”诺斯科拉夫特和尼尔写道，“举例来说，是任何挂牌价都能影响一处房产的感知价，还是挂牌价必须可信，人们才会考虑它，从而影响自己的价值估计？本次研究对现实制约一说的支持极为有限……”

该实验发表在1987年《组织行为和人类决策过程》（*Organizational Behavior and Human Decision Processes*）的一期上，引起了强烈反响。该实验为锚定效应的现实性提供了所需的证据，引用它的学术论文超过了200篇。



哪怕是对一件
拥有市场价值
的东西来说，
挂牌价造成的
锚定效应也仍
极为有力。

最无动于衷的倒是房地产经纪。研究人员向一组参与此事的经纪说明了研究的结果，“他们完全不接受，”尼尔回忆说，“他们反驳说，‘你拿统计数据怎么说都可以。可它不是真的。’”

要是你因此得出个“房地产经纪都是骗子”的结论，那你可就错了。经纪确实比缺乏专业知识的学生更不容易受锚定效应的影响。这个实验真正想探讨的是人类的意识如何从现实生活中无处不在的丰富数据里得出数字。它说的不光是房地产经纪，而是我们所有人。（我提醒诺斯科拉夫特，有人对他1987年那篇论文里的房地产经纪做了些尖酸苛刻的评论，他说，“如果是我，我大概会说，做个人没什么好羞耻的”。）

美国亚利桑那州的实验提出了一个重要主张：哪怕是对一件拥有市场价值的东西来说，挂牌价造成的锚定效应也仍极为有力。诺斯科拉夫特得出结论：可信价格区间的范围比大多数经纪想的要宽泛得多。此外，没有任何理由认为，高于估价12%就能限制策略化的锚定。它纯粹只是研究人员敢在本次实验中尝试的最高数字而已。

尼尔开玩笑说，她的同事们都跑来读这篇论文，为自己卖房子寻求建议。但它并不是一幅完整的蓝图。众所周知，卖家得在要价和挂牌叫卖的时间中做出权衡。挂牌价较高固然有好处，可要价过高，找买家的时间也长，这方面的成本自然相对较高。房地产经纪毫不介意指出这一现实情况。

“打那以后，我们开始着手研究这样一个问题：价出到什么程度，就变得荒唐了？”尼尔说。在精神分析实验室，荒唐不一定有什么关系。哪怕数字看起来荒谬，锚定效应也照样管用。可买房子的人很少会去看价格超出自己承受范围的物业。“对于‘你总可以降下来’的说法，市

场可不买账。”纽约房地产经纪戴安娜·萨奇（Diane Saatchi）谈及2008年房地产市场崩盘时如此解释。定价合理的房子几个月都无人问津，还想把价格定得远远高于市值？没多大可能吧？

第22章 卖了安迪·沃霍尔的海边别墅

2000年，电影导演、美国近代艺术家安迪·沃霍尔（Andy Warhol）的业务伙伴，保罗·莫里西（Paul Morrissey）把自己在美国蒙托克（Montauk）的房产挂上了出售的招牌。

1971年，他和沃霍尔以22.5万美元的价格，买下了一处22英亩(23)大小、名为“恩索”（Eothen）的物业。沃霍尔从没去过那儿——海风轻轻一吹，他的假发就要上天了。在其鼎盛时期，恩索招待了从杰奎琳·奥纳西斯到滚石乐队的各路豪杰。沃霍尔于1987年去世后，他的基金会把3/4的空地捐给了大自然保护协会，作为动植物保护区。莫里西卖的是剩下的部分，占地5.6英亩的5座房子，能停3辆车的车库，一处长达600英尺(24)面朝大西洋的平整海滩。

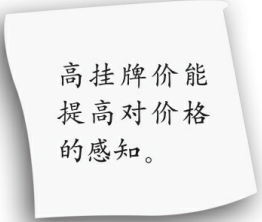
按美国东区房地产的标准，恩索是一处独特的物业。它建于1931年，主人是爱好打猎的艾禾美（Arm & Hammer）公司的继承人。当时，恩索古朴的木头墙壁上挂满了鹿头鱼骨标本。沃霍尔和莫里西从没改变过这夸张的装饰。房间小，从没升过级——有人说那房间是给“小矮人”住的。对买得起它的人来说，这地方可没什么吸引力。

但这些东西也可以从另一种角度来看。“想必大家都可以理解，跟沃霍尔挂了钩，是这地方最大的资本。汉普顿的买主们，跟以创纪录的价格买下沃霍尔画作的是同一批人。”莫里西若有所思地说，“希望能找到一个保存这地方的买家。”

由于正负因素混杂，评估恩索的市场价成了一件棘手的事情。莫里西要价5 000万美元。地产行业的人觉得这也太夸张了。东区的买家们想要的是“绸缎床单、制冰机、左右对开的高级冰箱、液晶平板电视、室内游泳池”，房地产经纪保罗·布伦南（Paul Brennan）告诉《纽约时报》，“要是他愿意以2 500万美元卖的话，我可以帮他来卖。”

这样看来，莫里西要了一个两倍于现实价的数目。这个锚点比诺斯克拉夫特和尼尔实验中用的那些可太多了。莫里西并非没耐心的人。他把恩索的出售牌在市场上挂了7年——这么长的时间，没有哪个心理学家耗得起。莫里西不着急，因为他每年夏天都用着恩索，还把房子租出一部分以弥补支出。随着时间的推移，他降价到4 500万美元，而后又降到了4 000万美元。显然，直到后一次降价（2006年夏天），他才开

始有了些真心的主顾上门。4 000万美元的价格仍然离可信区间比较远（比布伦南建议的挂牌价高了60%），但在旁观者眼里，它已经不那么吓人了。2007年1月9日，莫里西跟时任J.克鲁的总裁米奇·德雷克斯勒（Mickey Drexler）交易成功，成交价是2 750万美元。“他看起来是个了不起的家伙，立刻明白了它的价值，”莫里西评价德雷克斯勒说，“他有心让这里保持原样。”



高挂牌价能
提高对价格
的感知。

为了得到一个好价格，客户把房产一直搁在市场里——这是房地产经纪的噩梦。房地产经纪的薪水不是按工作时间来算的，他们宁肯早出手，也不愿干耗着。他们甚至弄出好多吓人的故事，以向买家们证明这样的偏好有道理。有人说，房产报价过高，就会成为破损商品。等它真正被卖出去，也肯定是贱卖，不会卖出什么高价。

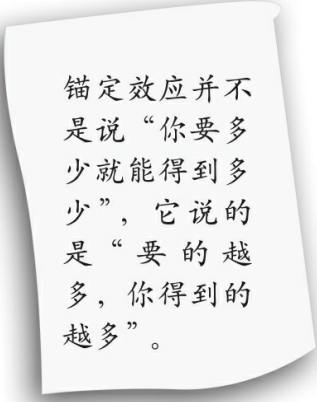
有人会说，莫里西把价格定得那么高太蠢了，全能又聪明的市场会让不切实际的卖家吃一脑门子的灰。可恩索这笔买卖似乎跟诺斯科拉夫特和尼尔的实验结果挺一致：高挂牌价能提高对价格的感知。恩索的售价比房地产经纪布伦南2 500万美元的建议挂牌价（这个报价只比房子卖出早了4个月）高了8个百分点。报价2 500万美元，意味着卖家可能在2 300万美元左右就愿意卖掉。故此，莫里西用一个高到荒谬的锚点，多卖了450万美元（20%）。

大多数把价格定得很高的卖家完全是真心希望卖到那个价。但他们注定是要失望的。锚定效应并不是说“你要多少就能得到多少”，它说的是“要的越多，你得到的越多”。要想锚定成功，卖家必须设定一个高价格，但并不指望真能卖到那么高。

能待价而沽整整7年，让辛苦为自己忙活的房地产经纪望而生畏，这样的家庭不会太多。这里有一种能让你鱼与熊掌兼得的办法。那就是使用被叫作广告参考价（Advertised Reference Pricing，简称ARP）的小把戏。

长久以来，折扣商店就使用广告和价格标签，把本店的价格跟另一地方或另一时段更高的参考价格相比较。更高的价格其实就是锚点，它提高了产品的感知价格，并形成了有利的对比。出于同样理由，商店在清仓打折的时候会留着原来的价签，故意让顾客看见。

科罗拉多大学专攻心理定价的营销研究人员唐纳德·利切坦斯泰因（此人跟萨拉·利切坦斯泰因没血缘关系）解释说：“今年夏天，我去买网球拍，我去体育用品商店看了许多球拍。比较价格的时候，我把广告参考价看得跟买入价一样重。这套把戏我知道得很清楚，可我就是没办法控制自己。”



锚定效应并不是说“你要多少就能得到多少”，它说的是“要的越多，你得到的越多”。

这就是参考价格的阴险之处。人人都知道它们不可能行得通！可诚如唐纳德在2004年的一次讲演中所说的：

广告参考价管用，大量研究表明它们管用，零售实践和反馈也表明它们管用。这不是什么新鲜事儿——而是一件众所周知的事儿。要是我宣传自己的售价是，比如29.95美元吧，再配上个广告参考价39.95美元，这样的话，大多数情况下，较之没有ARP的情况，销量会有相应的提高。随着我把广告参考价提高到49.95美元，59.95美元，再到69.95美元，销量可能还会增加。可要是ARP设在129.95美元呢？329.95美元又如何？好吧，再增加点儿乐子，5 000美元怎么样？

唐纳德和其他人做过实验，研究参考价能提到多高。1988年的一份报告说，消费品的参考价格和感知价格几乎呈直线关系，哪怕参考价比正常的市场价值高2.86倍也是如此。这就相当于一件279美元的东西，广告里说它在别的地方卖799美元。唐纳德说：“就算广告里的价格完全不可信，你对这样东西该值多少的认识也会受其影响。”

7年来，恩索臭名远扬。在蒙托克的房屋买卖聚餐和酒会上，总有

人对着那头5 000万美元的“大笨象”议论纷纷。莫里西把价格减到4 000万美元之后，5 000万美元的价格并没有随之消失。我可以向你保证，每个买家都听说过这处房产最初的挂牌价是5 000万美元。在有意无意之间，这成了一个广告参考价。最初的报价仍然拉高了估计价值。买家德雷克斯勒显然清楚5 000万美元和4 000万美元都是吹牛讲大话的。但要是他跟实验里的被试一样，那他肯定觉得自己做了一桩好买卖。在房地产市场上就跟在J.克鲁的商店里一样，谁都难以对45%的折扣视而不见。

倒卖房子的人有个惯用手法，先在短期内把房子挂一个非常高的价格，然后在卖家和经纪的耐心容忍范围内，把它降到合理的水平。之后，挂牌价上便可以“诚实”地提到最初的价格（原价为多少多少）。利用这种手法，只需把房子的挂牌周期稍微延长几天，就很有可能斩获锚点价格的大部分好处。

这类做法是否合乎道德，留给你自己去判断。这里还有一种略微迂回的伎俩：

价格实验

卖家A把房子挂牌出售，然后说服邻居B在若干网站上发帖，说邻居B的房子正由房主自售。邻居B并不真心想卖，她标出一个高到荒唐的价格（要是真有人愿意出那么多，她倒也会欣然接受）。关键是要让卖家A的房子显得像是笔划算的交易。

Zillow网站有一个“打动我”的功能，房主们可以为自己的产业标上幻想中的价格，哪怕它并不是为了招揽生意。凡是用过Zillow的人都知道，这些“打动我”的价格荒唐得很。然而，买家搜索在售住房时，这些幻想出来的“打动我”价格也会列在地图上。这不禁叫人很好奇，它们是不是有什么对比效应，能帮忙出售附近的房产。

卖房子的人使用锚点或参考价格的并不多，因为他们觉得买家肯定很精明，不会这么容易上当。唐纳德把参考价效应跟某种都市传说做了比较。有一回出现了这样一个流言：麦当劳的汉堡里有蚯蚓。部分地区麦当劳的销量随之猛跌了30%之多。其实，几乎没有任何人会相信这样的谣言，30%的公众并不相信一家大企业会为了在牛肉上省几个小钱就拿自己价值数十亿美元的品牌冒风险。问题是，哪怕是没人相信的事情，也仍然会影响人们的行为。



第三部分

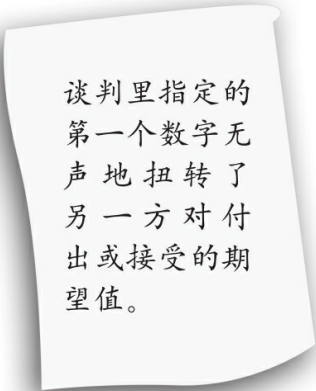
挥舞价格的魔棒

- 为什么议价时一定要抢先报价，而且一定要“狮子大开口”？
- 为什么如此强硬的杰克·韦尔奇会在离婚大战中败下阵来？
- 为什么一颗头骨会拍出上千万的高价？

第23章 抢先报出你的价格

“锚定不是什么稀奇事儿，”卡尼曼说，“它在谈判中很管用，先报出你的数目能带来优势。”对于交易者来说，这个简单的规则或许是价格心理学中最重要也最容易应用的规则了。谈判里指定的第一个数字无声地扭转了另一方对付出或接受的期望值。实地研究和学术实验都为这一论点提供了证据。决策心理学家似乎毫不怀疑它适用于真实世界的可行性。然而，把这条规则传达给商人却是难上加难。按尼尔的说法，问题在于管理人员对锚定信不过，觉得它是骗小孩儿的东西。

说到底，商业谈判人员要比购买花生酱的消费者对价格考虑得多得多。在坐下来谈判以前，他们会长时间地仔细思考自己的底价是多少。他们还尝试估计对方的底价以及夹在两者之间的差额。这么想下去，他们就觉得保留底价很现实，自己的底价牢不可破。



谈判里指定的
第一个数字无
声地扭转了
另一方对付
出或接受的期
望值。

未必如此。著名的棘手谈判家塞缪尔·冈帕斯（Samuel Gompers）曾被人问到，工人罢工到底想得到什么呢？他的回答是“更多”。谈判人员的一个普遍信念是，只要是有可能要到的东西，他们就会尽可能地多要。价格不是某人单方面地表达想要什么，而是某人认为自己能得到什么。这必然是一种粗略的猜测。大量证据表明，你可以对这种猜测加以操控。

尼尔是斯坦福大学商学院的教授，为《财富》500强企业与各国政府提供谈判技巧咨询。

尼尔说：“我们花了很多时间向这些真正做决定的人讲述锚点的力量，人们却很抵触。他们说，我不可能受它的影响；我已经做了很多工作，我了解情况。我说，我可以把玛姬放到某个环境下，不设锚点；也可以把她放到完全相同的环境下，再设上锚点。我们的研究能办到这样的事情，你却做不到。这时，我可以比较她行为的差异。两者存在差异，而且是系统性的，所以，锚点有着强大的效果。

显微镜的发明引发了强烈的情绪反应。安东尼·范·列文虎克（Antonie van Leeuwenhoek）发现，能让干渴的旅客重振精神的清澈湖水当中却有着不停蠕动的可怕东西。他看到血液里有微小的粒子（红细胞和白细胞），精液里有像蝌蚪一样的精子细胞，而人类的口腔更是聚集着形形色色的怪物。人人都知道，这不可能是真的。任何一个视力良好的荷兰公民，都可以把手举在面前，从这儿看到那儿。没有人看到过什么细菌。

行为决策理论家向我们展示了一些无法通过其他任何方式了解的东西。生活没有重播键，它从来没法回放，让我们看看要是情况稍有不同我们会做出什么样不同的决定，或是答应什么不同的价格。这就需要实验。

而这些实验的结果往往对自由意志的概念提出了挑战。从本性上看，高层管理者都是强硬的人。当你告诉他们，谁先报出一个数字，就能对他们施加下意识的作用力，对公司的收益造成影响，他们肯定会勃然大怒：“像我这么意志强硬的人，绝不会被催眠！”他们确信自己会从经验里了解谈判里什么做法行得通，什么做法行不通。

卡默勒把这叫作“土拨鼠日”。《土拨鼠日》（*Groundhog Day*，又译《偷天情缘》）是1993年上映的一部电影，比尔·默里（Bill Murray）扮演的男主角每天早晨醒来都发现当天是2月2日，他把这一天过了一遍又一遍。默里得以进行一连串的“决策实验”：勾搭妇女，酒后驾驶，甚至自杀之后安然无事。经过多次灾难性的尝试，他终于让生活回到正轨。卡默勒说，这部电影和真实生活之间的区别在于：生活很少给人机会，让人把复杂的因果搭配起来。

以色列本·古里安大学的伊拉娜·里托夫（Ilana Ritov）做过一个实验：

价格实验

148名工程和管理系的学生进行模拟谈判。一半的被试得到买

家证章，另一半被试得到卖家证章。游戏的目标是为一种假想商品谈交易，赚取尽量多的利润。每起交易必须指定价格、交货期限和折扣水平。被试们商谈出一套盈利方案，通过它来判断自己依照特定的协议能赚多少钱。为了显得真实些，谈判双方并不是简单地把8 000美元的蛋糕对半分掉了事。依据交易的构建方式，买卖双方每人最高可以得到5 200美元。为了得出双赢的解决办法，人们往往要来来回回地磋商。

现实世界里的讨价还价是没有规则的，所以里托夫也没有故意设定规则。任何人都可以跟自己喜欢的人搭伙，只要两人愿意一人当买方，一人当卖方就行。两人均可首先报价。他们可以为自己的报价提出任何理由，并使用任何谈判策略。倘若有人觉得谈了半天是在白费劲儿，随时都可以退出，另寻谈判伙伴。一旦做成一笔交易，两名被试就可以去寻找其他被试，继续做交易，直至用完分配的时间。被试可以尝试让每一笔交易都实现最大利润，要不就“用数量来弥补”。

里托夫发现，卖家往往会主动接近买家。从某种意义上说，这不禁叫人吃惊，因为游戏极为抽象。没有需要搬来搬去的实际商品，除了名字标签之外，两组人没什么根本上的区别。不过，词语可以为行为钉起框框，被试迅速投入了买方和卖方的熟悉角色。通常，卖方会叫价，买方则还价。对大多数人来说，事情就是这样。

只见树木不见森林，是人们常犯的错误。里托夫的实验得以揭示一些谈判者本身看不清楚的东西：抢先报出己方数字所蕴含的力量。平均而言，先报价者赚的钱更多，初始报价越高，赚钱越多。

这一点可以清楚地从里托夫1996年发表在《组织行为与人类决策流程》中的论文所附的图中看出。图23-1以横轴标示的是初始报价，纵轴上标示的则是最终约定价。纵横两轴的出价都以发起者的利润来表示。每一个点均为达成的交易。重要的不是个别的“树木”，而是“森林”的形状。点群大致上顺着一条向上倾斜的直线分布。换句话说，你要的越多，得到的也越多。

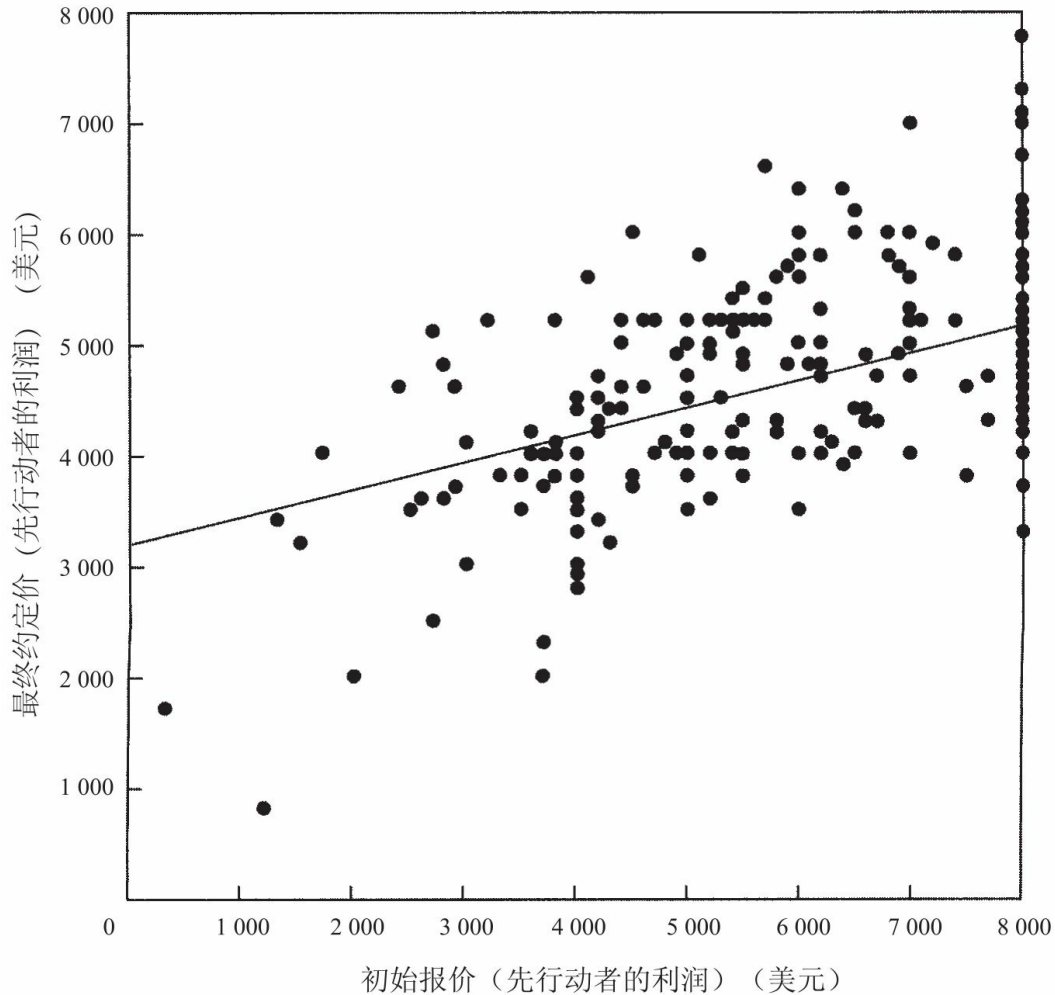
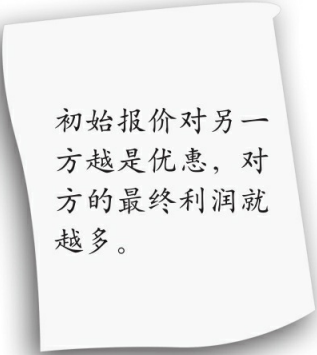


图23-1 里托夫实验

对任何一桩交易的任何一方被试来说，最高利润是8 000美元。最低利润是0。不少发起人一开始就来了个狮子大开口，直接就要8 000美元。上图右侧密集的点群便表示这种情况。“贪婪”的开价给对方留下的利润空间很小，甚至干脆就没有。

可是，狮子大开口却没有什么明显的负面因素。尽管最终没有人赚到8 000元的利润，可最初开出这个价的人赚到的和开价较低的人一样多，说不定还更好一些。

另一点发现更叫人惊讶。非发起者一方的利润表跟图23-1看起来差不多。初始报价对另一方越是优惠，对方的最终利润就越多。这就凸显出发起者对双方命运所具有的决定性意义。



初始报价对另一方越是优惠，对方的最终利润就越多。

房地产业是由卖方要价的。买方对这一点没什么办法可想。在不少其他环境中，待价而沽是一种先发优势。工资谈判往往如此。

大多数提出薪水谈判要求的员工都觉得自己处在不利的地位上。大公司每年要面试上千名申请工作的人。它会开出数百种薪资条件，看看有多少人接受。这令雇主对当前市场环境有着很好的直觉。普通的求职者同企业面谈的次数并不多，他只能猜测自己的当前市场价值。要价太低，你是在自欺欺人；要价太高，又显得自己太愚蠢（还有可能错过自己喜欢的工作）。难怪许多求职者会采用以下策略：

- 让雇主先开价；
- 不管雇主开价多少都说它不够；
- 要求多加20%；
- 要是雇主增加10%（或者你心目中的理想百分比）以上，就答应下来。

如果采用这种方法，那么，无论雇主最初的开价是多少，只要最后增加了10%，你就能满意。这意味着，较之典型的被试，你加倍地成了锚点的奴隶。

先报出数字的人所确立的锚点是最有力的。任何人都不应主动放弃这样的机会。幸运的是，现在求职者要调查自己的价值比以前容易多了。Salary等网站会问你几个问题（岗位名称、教育水平、经验年限、邮政编码），而后为你生成薪水钟形曲线。比如，你可以从中得知，90%的同类工人的收入少于73 415美元。诚实地回答网站的问题，然后

找出百分比排名为90的那个数字（即90%的人拿的薪资是多少），它会是一个很得体的锚点（初始出价）。你可能拿不到那么多，但至少招聘的家伙不会在办公室里嘲笑你。

谈判中可能出现的最糟糕情况是对方开出一个完全无法接受的价码。此时，巴泽尔曼和尼尔认为，有必要“重抛锚点”——要求从头开始。《理性谈判》（*Negotiating Rationally*）是他们为MBA课程合写的通俗教材，二人在书中警告道：“面对初始报价，以建议调整作为回应，就给了锚点某种程度的可信性……以退出谈判作为威胁，比答应一个无法接受的起始点更合适。”

反对价格锚定概念最常见的说法大概是，只有傻瓜才上它的当，我这么聪明，肯定不会中招，跟我打交道的那些人也一样。

2008年，德国劳动研究协会（German Institute for the Study of Labor）的乔戈·欧切斯勒（Jorg Oechssler）、安德烈·罗伊德（Andreas Roider）和帕特里克·W.施米茨（Patrick W. Schmitz）检验了这种看法。他们找来1 250名被试回答一份“认知反射测试”（Cognitive Reflection Test，简称CRT），这是一种迷你IQ测试，只有3道题。问题都是典型的脑筋急转弯。你也不妨试试看，回答以下3道题（先别看答案）：

1. 球棒和球一共是110美分。球棒比球贵100美分。请问球是多少钱？
2. 倘若5台机器生产5个零件要5分钟，那么100台机器生产100个零件要多少分钟？
3. 湖里开着一片百合花。百合花丛的面积每天都翻倍。假设它覆盖满整个湖面要48天，覆盖湖面一半需要多少天？

欧切斯勒的团队把被试分为两组。答对两道题及以上的人属于慎思组，得零分或只答对一道题的人是冲动组。（要想知道你属于哪个组，我在这里附上正确答案：5美分；5分钟；47天。）

两组人还回答了关于锚定的问题。慎思组和冲动组对锚定的感受性没什么区别。事实上，研究人员发现，较之冲动组，慎思组受锚定效应的影响还稍微多一些，只不过这没什么统计上的意义。

对聪明又慎思的人来说，一个数字或一个假设性问题就能引发丰富的联想。他们对答案思考的时间越长、越用力，初始想法的曝光量也就越大。这似乎抵消了额外思考带来的准确性优势。

第24章 让对手注意力不集中

“每当我给谁修什么东西的时候，”地产商人唐纳德·特朗普（Donald Trump）有一回吐露内情道，“我总是在原价上多加5 000万美元或6 000万美元。我的手下进来，他们说成本是7 500万美元。我说成本是1.25亿美元，但给我1亿美元我就答应修。基本上，我做的是桩下流事儿，可他们却认为我做得很棒。”

不是只有特朗普意识到了两个数字所具有的魔力。来看看巴泽尔曼、萨莉·布朗特·怀特（Sally Blount White）和鲁温斯坦设计的新型最后通牒博弈吧！

价格实验

他们要一组响应者指出10美元里自己最低愿意接受的是多少钱。平均答案是4美元，这很典型。

接着，他们向第二组响应者出示两个报价（比如3美元和2美元），而非通常的一个。响应者可以接受其中任何一个，或是两个都拒绝。

这极大地改变了人们的行为。有了选择余地之后，响应者更容易接受较高的那个出价（本例中为3美元），而不是否决它。请记住，第一组里的绝大多数人都说自己会否决3美元的出价（或是这么说，凡是低于4美元的出价都不接受）。这就是说，当3美元是唯一的出价时，人们会否决它；可当2美元的出价也摆上桌以后，人们却愉快地接受了3美元。

巴泽尔曼的团队检验了不同的两两出价。他们发现，人们接受的最低报价平均为2.33美元（只要它跟另一个报价比起来更高）。在这种情况下，两两出价造成了很大的影响。响应者愿意接受的价格低了40%，只因为它是两个报价里较高的那个。

为什么会这样呢？显然是对比和误导造成的。在标准的最后通牒博弈中，要是提议者对响应者开价3美元，响应者会用它跟提议者自己留下的7美元做比较。跟7美元一比，3美元显得很少，于是触发了不公平，甚至愤怒的感觉。可当桌子上有两个出价时，人们的注意力就转移到甲比乙好上面去了。至于说这两个报价跟提议者能得到的份额相比起来差很远，人们已经没有什么心力去顾及了。到了做选择的时候，当事人会这么想：要哪个呢——3美元，2美元，还是什么也得不到？

卡默勒、鲁温斯坦和普雷莱克写道：“不管是认知还是感情，自动流程都是大脑的默认操作模式，它们随时都在呼呼作响，甚至连我们做梦的时候也照常运转，构成了大脑里的大部分电化学活动……例如，注意力便主要有自动化流程控制，反过来，注意力又决定了我们吸收了什么样的信息。”兴许，在你报税的时候，棒球飞来砸烂了窗户，你并不是“决定”要抬起头来看看是怎么回事，它是自动的反应。

神经科学正开始描绘细节情况。大脑的底部有一团灰色的小肉球，叫作杏仁核。它的作用之一是充当“看门狗”，专门监测可能的威胁，哪怕你当时的关注焦点放在别的地方。在实验室研究中，杏仁核能“看到”周边视觉范围里的东西，这些东西，大脑更审慎的部位反而是看不见的。

很早以前，魔术师就在利用这种潜意识机制来指引飘忽不定的注意力关注点。他们知道，观众很快就能适应所见所闻，并主要根据对比或变化做出反应。按变魔术的学问，想实现误导，最好是趁着物体在移动而非静止的时候；趁着它活生生而非死气沉沉的时候；趁着它刚刚出现而非先前就在舞台上的时候；趁着它陌生而非熟悉的时候。一阵青烟腾起，打扮夸张的助手突然出现在舞台上，此刻，魔术师借着掩护把兔子装进了帽子。耍魔术的一大真谛就是“大动作掩盖小动作”。想让人把注意力从一个可疑的小动作上偏离开去，就要做一些加倍可疑的大动作。两相对比，小动作似乎就没那么可疑了，于是就会被忽略。这个简单的诡计竟然管用，因为人的意识总会把飘忽的知觉融入对周遭世界的幻想中，将它们整合成一幅完美的、无缝的、实时的地图——整个过程类似谷歌地图将卫星在晴朗日子里拍下的数千幅照片拼合成世界地图一样。谷歌地图中万里无云的地球只是一种幻想，同样道理，以为我们眼睛看到的東西都可信，也是一种幻想。

魔术就是在自由意识的幻觉上做买卖。因为观众没有觉察到心理操控令他们把注意力放在A、B、C而不是X、Y、Z上，他们相信自己看到

了所有重要的东西。当然了，只要他们选择不去看助手被刀劈成了两半，他们本来是能够看到一切的。当今的行为决策理论家倾向于以同样的道理看待讨价还价和价格设定。擅长此道的人，其实也擅长利用伙伴受限的注意力和有限度的理性。

双重出价的最后通牒博弈，就类似“桌上有死狗”和“红脸白脸”一类的古老技术。精明的谈判者有时会提出一个他明知对方绝不会接受的出价（俗称的“死狗”）。他坚持一会儿，然后重新考虑，提出另一个于对方更有利的出价。比较起来，新的出价好极了，于是对方迫不及待地接受了它。搞定！其实，新出价是谈判者早就盘算好的——要不是这么做，对方是断然不会接受的。

还有一种做法，谈判小组里的一名成员（也就是“白脸”）开出了“死狗价”。等他去洗手间的时候，他的合伙人，“红脸”向对方表示同情，并透了口风，应允给些更慷慨的条件。白脸回来了，跟红脸闹起了矛盾。最终，红脸胜出。对方高兴地接受了他的提议（这正中红脸、白脸的下怀）。

巴泽尔曼的研究小组发现了这些选择的类似效果。以下选项，你愿选哪个呢？

（a）自己留400美元，给对方400美元；

（b）自己留500美元，给对方700美元。

若将上述二选项单独列出，看到（a）的被试认为它可以接受；看到（b）的被试认为它不怎么好。（b）选项中对方比己方更好这个事实，是搞砸交易的元凶。

可要是把两个选项一并列出，绝大多数人（78%）选了（b）。通过直接对比，人们看出，凡是选（a）的人，就等于自己罚了自己100美元。选项（b）对双方都更有利。

在这种抽象情境中，除非跟别的东西做比较，否则被试无从判断某个报价是否“可以接受”。增加的第二个选项改变了他们的注意焦点。

巴泽尔曼请西北大学凯洛格商学院的MBA给一桩假想中的工作机会打分（20世纪90年代中期的美元币值）：

工作甲：4号公司提供一年75 000美元的薪资。众所周知，对所有刚从顶尖商学院毕业的MBA，这家公司都开价75 000美元；

工作乙：9号公司提供一年85 000美元的薪资。众所周知，对其他从凯洛格商学院毕业的学生，这家公司给95 000美元。

对雄心勃勃的MBA来说，工作乙无异于赤裸裸的侮辱。若一前一后地罗列甲、乙两个工作机会，大多数MBA都拒绝工作乙而倾向于工作甲。可要是同时罗列两者，他们又都倾向于工作乙。不管公正的价值是多少，放弃额外的10 000美元薪水都不值得。

这一发现颇值回味，因为我们其实很少有机会选择，而只能一前一后地评价工作机会。它们是一个接着一个来的。接到一个工作机会，你会有几天时间来做决定：薪水够不够好呢？我接受它，还是继续找呢？在这种情况下我们所做的决定，跟工作机会同时出现、能从两者或三者中挑挑选选时所做的决定不一定一样。

巴泽尔曼、怀特和鲁温斯坦认为，所谓的公平也有黑暗的另一半：“归纳到一起，我们的研究表明，在孤立评价结果时，人们往往更注重人际比较，而不是最大化个人所得的结果，上述实验结果暗示，倘若是一一做出政治决策，人们就有可能把公正感当成决策的基础。不管是对他们自己，还是对整个社会来说，这样做都不尽理想。”

第25章 喝喝小酒，好做生意

黎塞留公爵（Duc de Richelieu）[\(25\)](#)说过，多来一瓶约翰山堡（Johannisberg）顶级白葡萄酒往往能改变帝国的命运。美国企业显然同意这个说法。为了拉拢、款待客户和商业伙伴，每年它们大概要花掉200亿美元。这大概占了酒类零售市场12%的份额。此举绝非白白慷慨。酒精能以自己的方式带来回报：让客户和供应商达成原本难以达成的交易，给出比清醒时更优惠的价格。所谓的“试探气球”多是靠着酒精相伴才提出来的，有时，交易大纲干脆是在鸡尾酒会的餐巾纸上勾勒的。只要是“正常且必要的”，美国国税局允许企业和个人从应缴税额里扣除酒精饮料“款待费用”。而对于酒精“正常且必要”，似乎没有人表示怀疑。

在经济衰退的时候，靠酒精润滑做成交易是万万少不了的。2008年，纽约房地产市场沉了底，普天寿道格拉斯地产公司（Prudential Douglas Elliman Real Estate）在带高净值客户看房的时候，为他们免费提供泰斯卡和拉加维林威士忌——这两种酒的市场价，前者是60美元一瓶，后者是77美元一瓶。一位销售员工说，“一小杯波旁酒”对销售很有好处，他颇有信心地认为，酒类预算能够获得补偿，甚至还有富余。房地产记者克里斯汀·霍赫尼（Christine Haughey）写道：“几杯小酒，能让胆小的交易者站上舞台，也能让他们鼓足勇气，买下数百万美元的公寓。”

这些价格，跟人们在完全清醒时愿意接受的价格一样吗？英国利兹大学和牛津大学的一个研究小组做了一次实验：

价格实验

让社交性饮酒者在喝酒精或安慰剂后完成一组心理测试，测试中包括若干带有博彩性质的选择题。实验组饮用的鸡尾酒是奎宁水、塔巴斯哥辣酱油和酒精的浓烈混合品，对照组喝的则是纯净水。酒精的饮用量为三大口（因被试的体重而略有差异）。（是的，很难相信被试无法判断自己是在酒精组还是对照组。这是酒精

实验长久存在的老问题了：它没法设计出一种可信的安慰剂。）

按民间智慧的说法，酒精让人爱冒险。喝醉了的朋友坚称自己能开车，你非得跟他打上一架才能抢下他手里的车钥匙；赌场里的“免费”酒精鼓励消费者更肆无忌惮地下注。然而，从英国的这次实验来看，在很多方面，酒精组跟对照组并不存在太大的区别。酒精并未抹杀前景理论。不管是喝了酒还是没喝酒，人们面对收益都厌恶损失，面对损失都乐于冒险。两个组人都青睐“P类赌”，即胜率更大的赌注。

不过，英国的研究人员发现了一个非常显著的差异。它出现在被试面对涉及巨大损失的利切坦斯泰因和斯洛维克式“艰难”选择时。

如果你想自己试试这个实验，那你需要调三杯饮料。每一杯饮料都用高脚玻璃杯装3.6盎司80%高纯度伏特加，再用奎宁水掺满至10盎司。（3.6盎司适用于体重150磅的人。请根据这一比例适度调整。）在15分钟里把这三杯酒都喝下去。接着等上10分钟，回答以下两个问题（我把英国心理学家用的“点数”改成了“美元”）：

问题1：你会选择哪一个呢？

- (a) 50%的机会赢10美元，不然就输10美元。
- (b) 66%的机会赢20美元，不然就输80美元。

问题2：你会选择哪一个呢？

- (a) 50%的机会赢10美元，不然就输10美元。
- (b) 66%的机会赢80美元，不然就输80美元。

写出你的答案！

现在是最重要的部分：在接下来至少两个小时里，别开车、别骑车、别操作机器，也别干其他任何事。

从设计本意上说，两个问题都颇难选择。这里没有最佳答案。两道题的选项（a）是一样的，相当于公平地投了硬币。由于人们对损失的遗憾之情超过了对收益的欣喜之情，所以，从主观上来说，人人都会觉得打这样的赌是损失。两道题的（b）选项均为P类赌，我们知道大家喜欢P类赌。选择（b），你或许会轻松地赚到一些钱，高高兴兴地抽身而退。

且慢！这儿有一个“包袱”：两个（b）选项都包含了一笔令人担心的80美元罚款，这减少了它们的吸引力。故此，对第一个问题（说不定第二个问题也是一样），被试被迫要两相其害择其轻。

两道题只有一个区别。那就是选项（b）中赢取的金额。问题1里是20美元，问题2中则是更为慷慨的80美元。依照逻辑，你大概会认为问题2里选（b）的人会比问题1要多。

事实也正是如此。大多数被试（不管是清醒的还是喝醉的）都选了问题1的（a）和问题2的（b）。然而，清醒组对两道题做出不同选择的人更多。较之酒精组，他们对赢取金额的变动更为敏感。

英国研究小组的整体结论便建立在这一微妙的区别之上。在权衡巨额损失的前景时，喝了酒的人对收益额的权衡能力有所衰减。（b）选项中的收益从20美元提高到80美元，许多喝了酒的人却并没有改变自己的选择。这好像是因为他们对此并没有注意到。

酒精把人本来就有限的注意范围进一步缩小了，人们给这种现象取了个绰号，叫“酒精近视”。它把理性的界限再度收紧。问题1和问题2对被试的注意力提出了竞争需求。既然被试要进行选择（而不是指定一个价格），那他们就必须把注意力集中在获胜率上（50%对66%）。其次，他们务必要关心不利的风险。选项（b）中有可能损失的最大数额是80美元，这比选项（a）中的10美元损失要多得多，所以，80美元的损失吸引了人们的注意。和魔术师的“大动作”一样，80美元的损失造成了误导。被试首先权衡的是1/3的概率输掉80美元会有多糟糕，它是否足以推翻对P类赌的固有偏好。

不管是对喝醉了酒的人还是清醒的人来说，用在考虑选项（b）收益面上的认知资源都远远不够。因为这里有太多可以变戏法的数字了。喝了酒的被试尤其不堪重负。最终，他们没能给予收益足够的关注。这有时会让他们的决策显得冒险，有时还会让他们做出异常保守的选择。

做生意时有太多类似的事情。向潜在客户报价就是一场赌博。你永远没法准确地知道它会涉及多大的工作量：客户的要求有多苛刻，哪些地方可能出岔子，相关的机会是什么样。午餐时来上三杯马天尼就能把复杂的问题变得简单。如果未能反映出所有的相关信息，任何报价都可能是“错”的。要是报价太高，卖方说不定会失去一单有赚头的好生意，

可要是报价太低，那就给自己穿上小鞋了。

第26章 通货膨胀背景下的货币错觉

10亿津巴布韦元(26)买不了与从前一样多的东西。

2008年7月，罗伯特·穆加贝（Robert Mugabe）政府发行了一种100亿津巴布韦元的钞票。它立刻成了收藏家们的追捧对象。可作为钱本身，几个星期里它就差不多毫无价值了。2009年1月，津巴布韦储备银行又推出了一种100万亿津巴布韦元的新钞票。钞票上画有水牛和维多利亚瀑布的图案，据说值30美元。那个时候，几乎已经没人使用津巴布韦币了。据报道，津巴布韦币的通货膨胀率最高达到每年百分之5 000亿。除了印刷的面额越来越大，政府还定期砍掉钞票后面的“0”，2008年底就一口气砍掉了13个。不知怎么回事，津巴布韦的货币设计师居然能跟科学记号的禁忌对着干。

津巴布韦到底是如何管理的呢？全世界都在问这个问题，津巴布韦的记者却发现很难向外国人给出一个直白的答案。津巴布韦的经济乱成一团，失业率高达80%，饥荒蔓延。通货膨胀只是普通津巴布韦人要面对的最微不足道的问题。那些侥幸保住工作的人便能应付。他们坚忍克己地接受了自己国家的钱币保质期像牛奶一样短。时间一天天地过去，尽管绝对价格变了又变，价格的相对比率却保持稳定。

第一个研究恶性通货膨胀心理的伟大学者是欧文·费雪（Irving Fisher）。目前，人们重新提起了对这位经济学家的关注。泰勒称赞费雪是行为经济学的先驱，这么做的还不止他一个。特沃斯基在后期写的一篇论文中认为，费雪提出的“货币错觉”概念其实就是通货膨胀时期玩弄的一套认知诡计。

其实对行为经济学家这个群体来说，费雪是怎么也不该当上英雄的。1892年他发表论述，抱怨古斯塔夫·费希纳对经济学界造成了恶劣的影响。他写道：“把心理学强加于经济学之上，在我看来这不恰当而且有害。”在20世纪的数十年里，费雪大概是美国最有名的经济学家。公众最初认识他，是他与人合著的那本自我救助畅销书，书名很是真挚：《如何生活》（*How to Live*）。费雪还是一个成功的发明家，设计了一种索引卡系统，它是旋转抽出式索引卡（Rolodex）的前身。后来，他的索引卡公司跟另一家公司合并成了雷明顿兰德公司（Remington Rand）(27)，费雪自己也赚了钱。打从在耶鲁大学执教的

时候，他就对当今时代的诸多议题发表过意见。他倡导素食和优生学，支持禁酒，凡是你想得到的健康养生法，他都赞成。1919年，为了治疗女儿的精神分裂症，他叫一个庸医切掉她一部分结肠，这一尝试并不成功，他女儿死了。

费雪的光辉事业在1929年戛然停止。“黑色星期一”到来之前，他都在尝试安抚投资者们焦虑不安的神经。他说：

市场近期的波动，只不过是“震出丧失理性的狂热”。等市场把狂热分子都踢走了，价格肯定一飞冲天。“股价已经到了一个像是永恒不变的高原上。”

可股价并没有稳住，市场把索引卡带来的财富蒸发殆尽，费雪的声音也伴随他的声明烟消云散。

费雪相信，靠着物理学家的严谨态度，应该可以预测价格。这一点，肯定是受了他的博士生导师、隐居的物理学家约西亚·威拉德·吉布斯（Josiah Willard Gibbs）的鼓舞。正如气体体积可以靠压力和温度计算得出一样，费雪渴望根据供给和需求预测价格。他的论文描述了如何做到这一点，他甚至还建了一台价格产生机（如图26-1所示）。这是一口大水缸，里面有若干入水一半的“蓄水池”跟另一套杠杆系统相连，调整“停止器”，把收入、边际效用和供给等数据输入杠杆，接着就可以从量表上读出价格了。吉布斯一定很高兴。该设备预示了（但愿不是嘲讽）20世纪经济学的发展方向。（“按下1号停止器，抬起3号，”费雪的操作手册上这么写着，“现在，1号、2号和3号分别代表一个富裕的中产阶层男士和一个穷苦人……”）

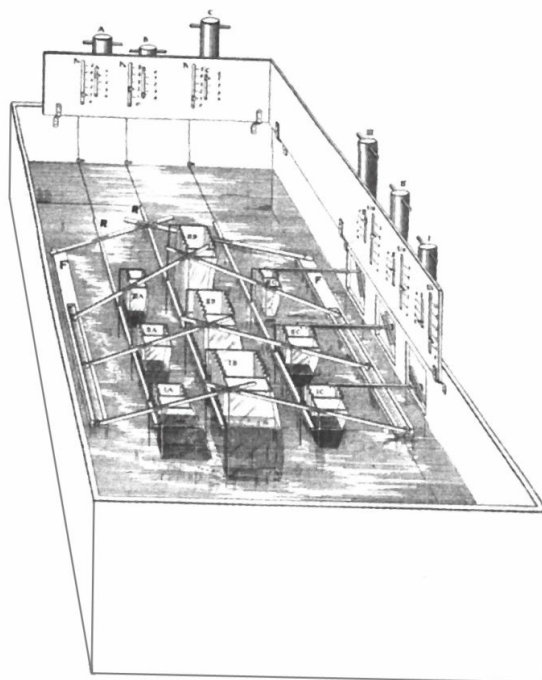


图26-1 价格产生机

跟同时代的一些人不同，费雪对不适合他机器的异常情况特别感兴趣。1928年，他在《货币错觉》（*The Money Illusion*）中对通货膨胀做了史诗般的论述，迄今未有超越之作。1922年，费雪去了魏玛共和国，看到了普通民众如何应付肆虐的通货膨胀。为了偿还惊人的战争债务，德国的印钞机片刻不息地印着马克，“一战”结束后，商品价格已经翻了50倍。在柏林的一家商店，费雪挑了一件衬衫，并按店主的报价付了钱。“因为店主担心我会觉得她是在赚取暴利，她说：‘我卖给你这件衬衫，再进新货的成本跟你付给我的价是一样的。’我还来不及问她为什么要以这么低的价格卖给我，她就自顾自地说了下去：‘可因为我买这件衬衫的价格低，所以我也算赚了钱了。’”

当然，从任何意义上来说，店主都没有赚钱。她提前为衬衫预付了若干马克，当时它们尚具有相应的购买力。从进货到把衬衫卖给费雪，马克的购买力已经降低了。她的确抬高了标价，但此时，马克的币值早已大大缩水。

费雪的看法是，金钱只是获得东西的一种工具。价格稳定的时候，我们可以把金钱和购买力画上等号。可要是货币的购买力不停地变化，就有必要对两者加以区分。

不管怎么说，经济学家就是这么看的。普通人，比如店主，往往会忽视通货膨胀。德国通货膨胀的高峰是1923年，当时的价格每两天翻一倍。在一张新闻照片上，一位德国妇女用铲子把钞票往炉子里铲。那时，烧一堆钱所产生的热量要比用它能买到的一堆木柴所产生的热量还多。即便如此，费雪发现，德国人还是设法活了下来，并且对现实抱着部分否定态度。他们把关注点放在价格上，而不是东西上。

货币错觉的概念几乎总是在通货膨胀的背景下引入的。事实上，美元或津巴布韦元等货币缩水，不必跟它有任何关系。只要价格变化，都可能出现货币错觉。它的立足基础是，消费者太看重价格，而对价格所代表的购买力缺乏足够的重视。符号变得比符号所代表的东西更加重要了。

你正打开一瓶很好的波尔多葡萄酒，跟朋友共进晚餐。你在期货市场（即收获季节之前）上买了一大箱这种酒，当时它的价格是每瓶20美元。结果，这一年的葡萄品质非常好。你碰巧得知（并且克制不住地告诉了客人们），同样的酒现在每瓶要卖75美元。你觉得今天晚上款待客人的这瓶酒花了你多少钱？

- (a) 一文不花（因为你是一年前付的款，现在可能都不记得价格了）；
- (b) 20美元（因为原来你花的就是这么多）；
- (c) 20美元，另加利息；
- (d) 75美元（因为如今你要是再买一瓶就得花这么多）；
- (e) -55美元（因为你花20美元就得到了一瓶现在卖75美元的葡萄酒）。

1996年，泰勒和埃尔德·沙菲尔（Eldar Shafir）向一群订阅红酒消息的收藏家提出了这个问题。许多人以前肯定碰到过这种情况。当然了，这里没有所谓的“正确”或“错误”答案。泰勒和沙菲尔只是在问当事人觉得红酒的成本是多少。他们的具体措辞是这样的：“关于你喝掉的这瓶红酒成本是多少，你认为以下哪种说法最符合你的感觉？”

经济学家差不多全都力撑答案（d）。要替换你喝的葡萄酒，现在就得花这个价。你以前为它出了多少钱，只不过是一个适合在晚宴时讲述的精彩故事，历史价格不顶用。

会计师或许觉得答案（b）很自然。对库存估价的进销存管理法用的是已支付价格。这合乎情理，因为零售商知道已经付了多少钱。他并不一定非得知道当前的市场价值，说不定都不值得为判断它花费精力。

答案（a）认为，价格历史不仅毫不相关，还有可能完全被遗忘，答案（e）则完全否定了“历史不顶用”的看法，并为一瓶绝佳的红酒得出了负成本！经济学家和会计师对这两个选项必定是嗤之以鼻的。可答案（a）和答案（e）恰恰是最受人欢迎的答案，选前者的人占30%，选后者的占25%。只有20%的红酒爱好者选择了经济学家的答案（d）。过往价格像幽灵一样缠住了绝大多数人。

原因之一是票面数额太难以否认了，我们受了它们的狂轰滥炸。“就算是在熟悉的环境里，在一些某种程度上所知甚多的人里，共同话语和新闻报道也往往表现出货币错觉。”沙菲尔、彼得·戴蒙德（Peter Diamond）和特沃斯基写道。要不，翻翻《吉尼斯世界纪录大全》吧，里面满是金钱的纪录——收入最高的运动员、最高拍卖价、最昂贵的餐点等。没有几个条目是根据通货膨胀做了调整的。没错儿，安德烈·阿加西（Andre Agassi）[\(28\)](#)挣到的美元钞票的确比阿诺·帕玛（Arnold Palmer）[\(29\)](#)要多。但至于说谁更富裕，你还是得靠猜的。

《纽约时报》、美国有线电视新闻网的记者们也并不得比《吉尼斯世界纪录大全》的编辑好多少。看看历年来的货币价值新闻排行榜吧！就算是在最权威的媒体上，也没有多少新闻是根据通货膨胀做了调整的。这大概是因为新闻稿总是热衷“最”字号的东西。“最大手笔的兽医捐款”永远比“捐款实际数额是第八大”听起来响亮得多。

是什么造成了货币错觉呢？最简单的答案是，数学太麻烦。但这并不是故事的一切。研究人员找来擅长做数学题的学生，用与通货膨胀有关，或是价格变化显而易见、方便计算的“容易”题目考他们。总的来说，这些学生还是成了货币错觉的牺牲品。

沙菲尔、戴蒙德和特沃斯基在纽瓦克自由国际机场和美国新泽西州北部的两家大型商场调查了一个多样化的群体。

价格实验

有个问题说，安和芭芭拉是出版公司的两名员工。有一年，安得到了2%的加薪，当时并非通货膨胀期。另一年，芭芭拉得到了5%的加薪，可同期的通货膨胀率是5%。

研究人员问第一组被试，“从经济的角度”来看，提薪后谁的日子更好过，是安还是芭芭拉？大多数人选安。这是“正确”的答案。

安的加薪让她的购买力提高了2%。由于通货膨胀，芭芭拉的加薪并未让她提高1%的实际购买力。

现在有趣的部分来了。研究人员从美国新泽西州游客和购物者随机选出了第二组被试，问他们谁加薪后会更开心。大多数人选择了芭芭拉。第三组被试的问题是，谁离职的可能性更大。他们选了安（意思是芭芭拉更可能留下来）。总体而言，\$\$\$ = 快乐 = 未经通货膨胀调整的实际美元数。

对第一个问题的回答表明，被试是有能力考虑通货膨胀因素的。碰到有“从经济角度”这样的提示时，他们会这么做，可其他时候就不然了。研究人员把原因归结到“多重表示”上。人在心理上用两种方式来表示金钱，一种基于实际货币数额，另一种则基于购买力。几乎人人都知道，一旦出现通货膨胀，第一种方式就用不得了。可两种表示方法都会有人给予注意，也都会影响决策（有时这种影响是不自觉的）。这暗示货币错觉可能是锚定的一种形式。美元票面数额是一个锚点，（根据通货膨胀）调整往往不足。

普通老百姓是货币错觉的真正受害者。雇主利用通货膨胀削减他们的薪资，还美其名曰“加薪”。工会谈判代表拍着他们的肩膀，恭喜他们“取得了胜利”。他们把积蓄存进账户，买成实际回报很低甚至干脆没有回报的房产、债券和养老金。政府对他们根本没有利润的房子和储蓄课以“利润”税。

也不是说货币错觉总不好。2008年《洛杉矶时报》的一篇文章指出，“2000年以来加利福尼亚州房价飙升，其实为少数族裔打开了房地产市场的大门，减少了社区因为害怕少数族裔到来而让房产贬值的恐惧情绪。”不管怎么说，经过一辈子的训练，货币错觉必然会得到强化。很多时候，我们的社会就是一条巴普洛夫实验里用的小狗，金钱就是唤起它反应的小铃铛。多次的重复之下，我们学会了对着空洞的符号猛流口水，反倒把真真正正的肉骨头抛到了脑后。

第27章 兜售货币错觉

假如你是一家电脑公司新加坡分号的负责人。这时是1991年，你正跟当地的一家公司谈合同，你向它出售电脑，交货时间是从现在开始算的两年以后。你现在的售价是每台电脑1 000美元。到1993年交货时，新加坡的电脑价格有望上涨20%。当然了，这只是猜测而已。安排这笔交易，有如下两种方式。

合同A：你同意（1993年时）按每台1 200美元的价格出售电脑，不管那时的电脑到底多少钱。

合同B：你同意按1993年的价格出售电脑。

你更倾向于采用哪种合同呢？沙菲尔、戴蒙德和特沃斯基在一份调查里提出了上述选择。他们发现，受访者在这两份合同的选择上分歧很大。46%的人选A，54%的人选B。心理学家们还发现，只要对两份合同的描述方法稍做改变，人们的回答就会发生戏剧性的变化。这个发现“有可能给讨价还价和谈判带来重大后果”。

上文的描述措辞尽可能保持了中立的态度。另一组被试看到的合同文本却加入了“真实”（根据通货膨胀加以调整）的限制条件。

合同A：你同意（1993年时）以每台1 200美元的价格出售电脑，不管那时的电脑到底多少钱。因此，如果通货膨胀率低于20%，你的所得将高于1993年的市价；反过来，要是通货膨胀率超过了20%，你的所得则将少于1993年的市价。由于你同意了一个固定的价格，所以你的利润将取决于通胀率。

合同B：你同意按1993年的价格出售电脑。因此，如果通货膨胀超过20%，你的所得将超过1 200美元，如果通货膨胀率低于20%，你的所得将低于1 200美元。由于生产成本和价格都跟通货膨胀率联系在一起，所以你的“真实”利润基本上能保持不变，无关通货膨胀率。

一旦做了这样的说明，绝大多数被试（81%）都选了B。在这个版本里，B保证了真实利润，A却是一场赌博。

还有另一组被试看到的合同则以美元价值做了表述——也就是故意挑起了货币错觉：

合同A：您同意（1993年时）按每台1 200美元的价格出售电脑，不管当时电脑的价格是多少。

合同B：您同意按1993年的价格出售电脑。因此，你无法担保每台一定能卖到1 200美元，倘若通货膨胀率超过20%，你的所得会更多，可要是通货膨胀率低于20%，所得就会比它低了。

这样的措辞把A说成了一件保准的事情，而B则像是看运气了。此时，59%的人选了B。按沙菲尔研究小组的看法，这意味着两件事：一是，人们会“自然地”根据美元数额来看待事情，在措辞“中立”和故意挑起货币错觉的时候，人们对问题的反应差异并不太大；二是，人的选择弹性非常大。厌恶损失是一种强大的动力源。为了回避风险，人们愿意出更多钱（是的，正如实验所示）。为了淡化风险的措辞多出钱，人们也愿意出更多钱。

沙菲尔、戴蒙德和特沃斯基认为，不管限定什么样的框架，人们往往都会接受。工会领导人想向会员兜售一份合约，或是管理层想向工会兜售一份重大协议，都应当谨慎考虑介绍它的措辞。诀窍是尽量把合约的风险说到最低。不管合约的实际内容如何，这都是有可能做到的。

假设合同要求把工资提高到每小时20美元，那么侧重点应当是它保证了20美元的时薪。可调式薪资有可能少于每小时20美元，甚至还会减薪。

要是它要求每年加薪3%，那侧重点则应放在它保证了薪资的提高上。工资肯定上涨，你不用担心通货紧缩，倘若合同与生活指数挂钩，通货紧缩说不定会导致减薪。

要是合同签订的是与生活成本挂钩的指数型薪资（费雪认为所有明智的合同都该这样），那么你可以说，它保证了最重要的一件事：购买力。讽刺的地方刚好也在这儿，这一点也是抨击指数型合同的要害。调查显示，除非你专门指出，否则人们是不会这样想问题的。

营销专家随时都在利用通胀的力量。互联网营销大师马琳·詹森（Marlene Jensen）建议客户使用这种机灵的盗窃手法。假设你有一种100美元的产品。你不卖100美元，而是说它本来卖149美元，打折后的售价为99美元。随着时间的推移，通货膨胀挤掉了你的利润，你不得不提价。詹森建议，别提价——少打些折扣就行了。

官方价格依然是149美元（当然，没人会付这个价），但现在你把

它打折而后卖119美元。对很多东西来说，比方说订阅电子报，大部分消费者根本不会注意。他们记不得原来的价格，对这种产品应该卖什么价也毫无头绪。于是，他们抵挡不了这样的诱惑：只用119美元就买到149美元的产品。

这才只是詹森的设计的前一半。时间不断推移，通货膨胀却片刻也不消停。你告诉顾客，由于成本提高了，你必须从149美元加价到179美元。可对忠实的优质客户（其实就是所有人）你继续打折，还是照老价格卖给他们，一个子儿也不多——119美元。没人能抗拒这个——支付的价格保持不变。

这为下一轮的狮子大开口奠定了基础。最后，你找到客户说，由于成本上涨，你只能调低折扣——同时抬高官方价格。如有涨价的需要，从头再来就成。

第28章 利用性别差异的最后通牒博弈

简·比斯利·韦尔奇（Jane Beasley Welch）拿起电话分机，听到许多根本不想听到的内容。她的丈夫，通用电气公司刚刚退休的CEO杰克·韦尔奇（Jack Welch）正在跟一个陌生女人说话。简悄悄放下了电话。她看了杰克黑莓手机上的短消息，确认了自己的怀疑。简跟丈夫正面对质时，最大的冲击来了。他毫不否认自己有了婚外情，连借口都不找。他爱上了苏西·韦劳佛（Suzy Wetlaufer）。苏西42岁，长相和身段都可以跟模特相媲美，是《哈佛商业评论》的编辑。该杂志说要给杰克·韦尔奇做个形象侧写，他同意了，却不料从此便一头栽进了“史上最昂贵的幽会”当中。

双方的离婚律师很快就围绕杰克·韦尔奇悬殊的身价估值争论起来。简的律师认为它是8亿美元（并想要它的一半），杰克·韦尔奇的团队则说只有4.56亿（给简的部分不低于30%）。谈判陷入僵局，在此期间，杰克·韦尔奇给简每月35 000美元的临时津贴。对像简这样具有权利意识的女性来说，这简直不算什么。轮到简来玩最后通牒博弈了。

2002年的夏天，人们的话题全围绕着贪婪的CEO们。安然、世通、泰科和阿代尔菲亚（Adelphia）的领导者们，接二连三传出丑闻。6月14日，腐化堕落的泰科CEO丹尼斯·科兹洛夫斯基（Dennis Kozlowski）为妻子的40岁生日举办了一场狂欢会。科兹洛夫斯基说这是股东会议，以此为借口让泰科负担了200万美元花销的一半。短短几个星期，这些丑闻就让科兹洛夫斯基成了众矢之的，他别无选择，只能辞职。讽刺的是，此前科兹洛夫斯基经常被人比作杰克·韦尔奇——这是当时对CEO的最高赞誉了。凡是翻看报纸商业版的人都知道，杰克·韦尔奇恐怕是最后一个平安下台的CEO了，没人怀疑他的正直、坦率和公平。

简有能力改变这一切。她知道，通用电气公司背着媒体和股东，供给杰克·韦尔奇多到令人吃惊的额外待遇。例如，通用电气公司答应为杰克·韦尔奇在任期内和退休后提供特朗普大楼（Trump Tower）里的一处公寓，每月租金80 000美元。简的律师告诉她，她有权使用该处公寓，因为它其实就是杰克·韦尔奇的资产。这类特殊待遇杰克·韦尔奇多得是。简的律师们要她一一道出，并用彩色图表做成了一份供词。

这成了谈判里的一个重要筹码。在那充斥着企业丑闻的年头，这些信息一曝光，（至少）会把杰克·韦尔奇从宝座上扯下来，说不定他还被迫放弃这些特权待遇。简这边提出的要求很简单：要么把这些特权待遇公平地分给我一份，要么谁都别想得到。

杰克·韦尔奇是通用电气公司最后通牒谈判传统的重要组成部分。因为他一贯炒掉10%绩效最差的经理人，所以便得了个“中子弹杰克”的绰号。他把人全部轰走了，“房子”却屹立不倒。可要是他以为简是在虚张声势，那他就错了。

简的律师于9月5日提交了供词。第二天一早，《纽约时报》就登出了它的所有细节。媒体不再围绕着“世纪大离婚案”喋喋不休，而是惊叹于杰克·韦尔奇的薪资方案是何等多元化。杰克·韦尔奇在通用电气公司的退休金大约是每年800万美元，比他当年的最高薪水翻了一倍。关键是他还什么都不必做。杰克·韦尔奇还为通用电气公司提供咨询，为此可得86 000美元的永久年俸。

较之终生享受的特别待遇，他的薪水简直微不足道。他可以随意使用公司的波音737飞机，飞行员免费，燃料也免费。每逢通用电气公司赞助的球队——红袜队、洋基队和尼克斯队打比赛，黄金位置他可以随便坐；杰克·韦尔奇在餐馆吃饭，公司埋单；汽车、电话、鲜花、干衣、红酒和维生素，公司统统包办。杰克·韦尔奇每年800万美元的退休金要如何花，倒成了真正的谜题了。“他好像谈出了一套退休方案，”供职于《纽约时报》的约瑟夫·诺塞拉（Joseph Nocera）说，“买任何东西都不必他亲自掏腰包。”

消息曝光后，杰克·韦尔奇大怒，很快，人们就把他跟科兹洛夫斯基说成了一路货色——这可不是恭维。简把他的老底爆料给媒体之后，不到10天，铺天盖地的批评声就把杰克·韦尔奇打趴下了。他宣布放弃通用电气公司付给他的所有特殊津贴。有人算了算，简的最后通牒一出，这对昔日夫妇余生里每年就少了250万美元的收入。

萨拉·索尔尼克（Sara Solnick）之所以对性别与谈判产生兴趣，起因是一件T恤。她年轻时修读经济，报名参加了卡尼曼和泰勒主持的夏日学院。在那儿，她看到了学院印的T恤，上面问道，“经济人是否存在”。“他们批评了经济人的现有模型，但仍然觉得‘经济人’（economic man）是个‘男人’（man），”索尔尼克回忆说，“我说，这个人的身份也应该有所不同。”

索尔尼克研究了劳务经济学，知道性别差距是该领域的一大难题。人们很早就知道，哪怕排除了有可能影响结果的所有明显因素，妇女的收入也仍比同等素质的男性要低。等索尔尼克了解了最后通牒博弈之后，她认为，可以从全新的角度来阐释性别角色。她想知道，在博弈对

价格设定的最低限度模拟中，性别差异是否仍然存在。导师告诉她，这是个有趣的研究课题，因为不管结果如何，它都会很有意思。她申请了5 000美元的研究补助金，就开始行动了。

价格实验

在索尔尼克巧妙的设计中，提议者和响应者坐在隔离物的两侧，彼此看不见。对照组的被试只知道伙伴的代码。另一组被试则知道伙伴的名字。第二组里的每个人必然都意识到了伙伴的性别，只不过没人知道实验的目的就是要研究“性别”。（有几个被试的名字很中性，比如“凯西”或“乔丹”。他们的结果不计算在内。）

不知道伙伴性别的提议者平均开价是4.68美元（总数为10美元）。可要是提议者知道伙伴是男性，则平均开价是4.89美元。倘若他们知道自己是在跟女性打交道，平均开价便只有4.37美元了。

最容易想出来的一种解释是，人人都认为男性报复心强，女性却可以被随便踩在脚下。奇怪的是，提议者是女性时，性别差距反而更大。女性对男性响应者的平均出价是5.13美元，比对半分还要高，可对同性响应者的平均出价却只有4.31美元。要么是女性对男性更慷慨，要么就是她们更害怕男性抓狂。一位女性提议者把10美元全给自己的男性伙伴，这种事情，哪怕是在新几内亚也没有发生过。当事人这么解释说：“我希望我们中至少有一个得到了点什么。”

索尔尼克让响应者报出自己愿意接受的最低出价。当他们知道提议者是女性时，最低值会更高。不管扮演哪一方角色，女性总会遭受不公平待遇。

这里，“性别歧视”一类的字眼或有误导之嫌。因为索尔尼克的被试是宾夕法尼亚大学的学生，一个个年轻得都记不得女权主义闹起来之前的事儿。尽管他们或许会有意识地拒绝双重标准（就好像锚定实验的被试否认自己受了随机数字的影响一样），可性别的确带来了差异。光是改了伙伴的名字，就触发了潜意识的性别行为模式（这一点可以用美元数字衡量出来）。

总体而言，索尔尼克研究中的男性提议者比女性提议者出价要高14%。这个数字，跟现实世界里的薪资性别差距报告很接近。索尔尼克指出，薪资是谈判出来的，“对于薪资待价而沽的部分，女性说不定只

能得到较小的一个份额”。

对我们这个讲究男女平权的社会来说，这些发现真叫人心烦意乱。个体薪资靠讨价还价决定，“同工同酬”就成了个滑稽的概念。要是雇主（不管男女）对女性提出较低的薪资报价，而女性接受了，那还能怎么做呢？索尔尼克发现，不少老板丝毫也不担心。根据她对雇主的研究，人们的一个常见反应是：“要是女性接受了我们最初开的薪资，那可真是遗憾。男人会讨价还价，得到一个更优厚的起薪。”

机会平等和结果平等诚然是两个不同的概念。人人都支持机会平等。可基本上，我们倾向于认为，机会平等会自然地带来结果的平等。索尔尼克的研究对这种想法提出了挑战。“如果你真的想要公平，”她说，“你不能光假设自己是公平的。你必须预先安排好一套程序。”

离婚争议就是一个很好的例子，究其本质，它是一种因为性别而复杂化的最后通牒博弈。韦尔奇夫妇的离婚大战就是其中的典型。男方有着更强的赚钱能力，他扮演提议者的角色，要求己方所得的比例多于50%。女方的优势主要在于她能否定任何提议（律师是按工作时间收费的，他们的计价器可没停下）。简的做法是她能够采取的最有效举措。她叫两人都牺牲掉了每年数百万美元的特殊待遇，一举成为爆发力超强的“中子弹筒”[\(30\)](#)。她表明了自己拒绝不公平出价的意愿，大出杰克·韦尔奇及其律师团队的意料。这一招可能真的管了用。2002年10月临时赡养费听证会召开之前，杰克·韦尔奇说：“让我们谈谈吧。”几个小时之内，他们就达成了协议。按《华尔街日报》的说法，“双方都说，简所得的数额远远超过了每月35 000美元。”

第29章 漂亮的人薪资更高

漂亮的人，无论男女，总能享受特权。经济学家慢慢接受了这一所有人早就知道的事实。近年来，劳务经济学家发现，长得漂亮的员工薪资更高。不管当事人是什么职业，是天桥上的模特，还是格子间里写代码的程序员，事实似乎都是如此。天生好看的人，是有“美丽奖金”的；其他人等，则只有“平凡罚款”。

索尔尼克认为，说不定可以用最后通牒博弈来调查外貌对价格和工资的影响。通常，这是一个复杂的问题，因为雇主付给漂亮员工更高的薪水，原因可能很多。干销售或前台接待，外貌是整个薪资中的一部分。雇主可以合理地辩解，公众喜欢富有吸引力的面孔。最后通牒博弈至少可以消除这些因素里的一部分。“这里没有生产力问题，没有期待，被试之间没有联系。”索尔尼克和合作者莫里斯·舒维泽（Maurice Schweitzer）写道。要是外貌在最后通牒博弈里都能发挥作用，那它在人们设定价格或谈判工资时恐怕也会发挥作用。

价格实验

索尔尼克和舒维泽在实验中，招募了70名被试拍摄个人照片。这些照片被送到另一所大学，由一群完全陌生的被试给每张照片打分，总分是11分，从-5（非常缺乏吸引力）到+5（非常有吸引力）。索尔尼克和舒维泽分别从男女里各选了6名最具吸引力和最缺乏吸引力的被试。由此得到24张照片，编成一本相册。

两人在第二所大学选出被试，进行最后通牒博弈。研究人员给被试每人看一张从相册里抽出的照片，并说这个人就是他们的伙伴。有吸引力和没吸引力的人表现如何，不存在有意义的区别，区别出在其他人的反应上。

提议者对有吸引力的人出价稍微多于对没吸引力者的出价（对前者是4.72美元，对后者是4.61美元）。然而，面对有吸引力的人，响应者要求获得更多份额。伙伴若属于最具吸引力的一组，响

应者的最低要求平均是3.53美元，而要是伙伴属于最没有吸引力的一组，响应者的最低要求平均为3.32美元。由于大多数的出价响应者都是接受的，所以，较之前一种情况带来的优势，后一种情况带来的劣势就显得不那么重要了。总体而言，有吸引力的人赚到了更多的钱。

在要求公众出钱、外貌有价值的职业中，比如房地产经纪、汽车经销商、拍卖员、销售员，这一实验的结果或许能得到有限的支持。但索尔尼克和舒维泽还考察了性别，发现它比外貌更要紧。从这个实验的目的来看，身为男性比女性更具有吸引力，更容易获利。不管是男是女，人们对男性的出价都更高，预期却更低。故此，男性比女性多得15%以上。外貌非常具有吸引力则是一把双刃剑。总的来说，最有吸引力的被试小组比最没有吸引力的被试小组多得近10%的薪资。尽管后一数字夸大了外貌的作用——它只比较了最好看和最难看的群体，而忽略了长相普通的路人甲乙丙丁。

第30章 巧妙利用“傻瓜蛋理论”

认为妇女不善讨价的想法，有可能是一套自圆其说、越说越圆的预言。在汽车经销商里，这种看法似乎颇为普遍。曾是汽车推销员的达雷尔·帕里什（Darrell Parrish）回忆道：

推销员把买家归类成各种“典型”。我当推销员那年头，会把最常见的那种叫成是“典型的不知情买家”。他们或是优柔寡断、小心翼翼，或是十分冲动，因而也非常容易受误导。现在，猜猜哪个性别被放在这种“超容易受误导”类型的首位？你猜对了，女性。

有关性别、种族和车价的最著名实验，大概还是1991年到1995年耶鲁大学法学院的伊恩·艾尔斯（Ian Ayres）和美国律师基金会的彼得·席格曼（Peter Siegelman）所做的研究。不过，这次研究引来颇多争议。

价格实验

他们将一支由38名被试组成的小队派往在美国芝加哥地区随机选出的153家汽车经销点。被试的年龄介于28至32岁，都受过3至4年的大学教育。他们按指示着装：男性，上身是马球衫或按钮式衬衫，下装是休闲裤和便鞋；女性，不怎么化妆，短上衣，直筒裙，平底鞋。所有人都开租来的车去经销店，等着经销商接触他们，开始讨价还价买新车。

就像为大型辩论会做准备的候选人一样，每名被试都要接受为期两天的谈判“脚本”训练。每个人都要等待5分钟，让经销商先开价；要是没人出招，被试要催促经销商开价。被试应给出一个接近经销商边际成本的还价（包括选配件），那是根据《汽车价格服务消费者报告》（*Consumer Reports Auto Price Service*）和《埃德蒙新车价格指南》（*Edmund's New Car Prices*）算出来的。

之后，他们要照搬一套谈判策略。一种方法是“差额对分”。不管经销商报价多少，被试都把他的新数目和自己之前还价的差额对半分，再抬高这么多。如此反复，直到经销商接受，或是拒绝进一

步谈判为止。此时，被试说他要再想一想，并不买车，径直走掉。

实验发现了种族偏见的惊人证据。对黑人的最终报价，平均比白人多1100美元。这是在几乎相同的时间里，在同一经销商处采用同一模式所得的结果。事实上，在将近44%的案例中，白人男性得到的最初报价，比白人女性或黑人所得的最终谈判价还要低。

对被人叫成“甜心”和“美妞儿”，女性需要保持极大的容忍态度。“你是个漂亮姑娘，所以我会给你一笔超优惠的交易。”一名经销商嘴上这么说，实际却没这么做。尽管如此，性别偏见的证据也并不怎么确定。白人女性比白人男性出的钱要多一些，但从统计上看并不明显。黑人女性比黑人男性得到的交易则稍稍优惠一些。

设计这样的现场实验，细节决定了成败。有太多微妙的东西可能会给结果造成偏差（虽说这是一个有关“偏见”的实验）。艾尔斯和席格曼的实验引人注目，因为它做了许多深思熟虑的保护措施。这是一个真正的双盲实验，交易商和被试都不知道发生了什么事。艾尔斯和席格曼告诉被试，他们在“研究卖家如何谈判汽车生意”，完全没有透露性别或种族方面的内容。每名实验者都配了对（他们自己并不知情）。配对里的一名成员是白人男性，另一名则为黑人男性或白人女性（或黑人女性）。经过安排，配对里的两名成员都会在几天内拜访同一家经销商，对同一型号的汽车讨价还价。

艾尔斯和席格曼的结果在媒体上掀起轩然大波。尽管价格歧视是一件复杂的事情，很难简单地归结到一点上。艾尔斯就指出，他的结果并不一定像许多人假想的那样——经销商有偏见，欺负黑人和妇女。经销商经常是让跟被试性别、种族相同的推销员接待他们，“可接下来这些推销员却给了他们最糟糕的交易”。实际上，黑人从白人推销员手里得到的交易更优惠，女性从男性手里得到的交易更优惠。

1996年皮内洛皮·科坚诺·戈德堡（Pinelopi Koujianou Goldberg）发表了另一篇有关价格歧视的研究，似乎推翻了艾尔斯和席格曼的结论。他并没有做实验，而是用消费者支出调查的方法检验全美的顾客在1983年到1987年购买新车的花销。戈德堡发现，从统计上看，黑人和白人、男性或女性之间并不存在明显的价格歧视。这样一来，自由主义者们援引艾尔斯和席格曼的研究，视之为歧视继续存在的证据；保守主义者则援引戈德堡的文章，说事情并没有那么糟糕；更主要的是，两项科学研究所得结论针锋相对，令许多人迷惑不解。

戈德堡认为，这一表面上的矛盾是可以调和的。首先，有必要理解汽车经销商为什么要讨价还价。汽车是可互换的、大规模生产的产品，工厂提供质保。一两辆同款车没有理由卖不同的价钱。美国绝大多数的买家都说自己痛恨讨价。一种流行车型，丰田旗下的赛恩（Scion）就迎合了消费者的这种喜好，采用全国统一的售价。

艾尔斯引用一位经销商的话说，讨价还价的理由很简单。买汽车其实就是“寻找傻瓜蛋”。有些消费者会出于傲慢，或对讨价还价有着神经质的反感，而照标价全额支付。这样的客户并不多，但他们在经销商利润中占了很大的一块。艾尔斯报告说，有些经销商靠10%的此类消费者就赚到了一半的利润。

今年早些时候，我曾在一次采访中询问某汽车经销商，他的利润是否绝大部分都集中在少数买卖上。他告诉我，他的代理销售店，“傻瓜蛋”和“非傻瓜蛋”生意都做得很多。他又补充说：“不过，我的表亲在一个黑人社区开了一家代理经销店。他卖的车不如我多，可他按标价卖出的车特别多。你知道，有时候好像是这样，往往是买不起它的人，却被迫出了最多的钱。”

戈德堡在黑人和女性所付的价格中找出了更多的差别。较之白人男性，这些群体里会有更多的极端客人出高价，尽管所有群体的平均销售价几乎相等。这应该可以解释两次研究的不同结论了。在艾尔斯和席格曼的实验里，人人都要求使用相同的谈判策略。这是为了揭示经销商是否会区别对待少数族裔——确实如此。但实验在设计上并无意检验黑人和女性买家在讨价还价上是否跟白人男性有所不同。

有人合理地猜测说，很多经销商都信奉“傻瓜蛋理论”。故此，他们向少数族裔的初始价就开得很高（如艾尔斯和席格曼的实验）。跟公道价格相比，初始价开得高，买家讨价还价的时间会更长，过程会更艰难。这就抹杀了性别或种族偏见的大部分证据（如戈德堡的研究）。

别的不说，这足以表明价格歧视有多么复杂。有可能，一些经销商根本没意识到“傻瓜蛋理论”。他们的报价存在统计上的种族或性别歧视，但他们本心上并无此种意图。

艾尔斯发现，不管性别和种族为何，一点点信息对买家就值319美元。声称自己已经试驾过真车的被试，比没有这么说的被试平均少了319美元，这在统计上看起来极为突出。不难理解，为什么经销商听买家这么说了以后会急于完成交易。

第31章 利用性别差异定价

2003年秋天，包括森德希尔·穆拉伊纳丹（Sendhil Mullainathan）和沙菲尔在内的一个研究小组做了一次极为大胆的实验。

价格实验

他们从南非某大型消费者信贷机构得到许可，在垃圾贷款推销邮件里测试一系列心理诡计。该信贷机构提供的服务相当于美国的发薪日贷款——为穷苦的工人提供高利率的短期现金贷款。

该信贷机构给自己过去的53 194名客户发了电子邮件，声称为他们提供特别利率。穆拉伊纳丹和沙菲尔的团队测试了在电子邮件中附加照片的效果。他们把某些人的笑脸照片放在信件的右下角，签名的旁边。这含蓄地暗示，照片上的人是一名银行职员，说不定电子邮件就是他写的。

一半的照片是男性，另一半是女性。一部分客户收到的照片上的人跟自己同一性别，另一部分客户收到的照片上的人则是异性。由于种族在南非社会事关紧要，所以他们也测试了这一点。他们用的照片里包括黑人、白人、印度人和混血。

经济学家会说，理性的人在决定是否接受昂贵贷款的时候，不会受一张照片的影响。广告界人士和骗子的看法不同，他们认为漂亮脸蛋或种族诱饵都是有价值的。穆拉伊纳丹和沙菲尔想看看照片的影响到底有多大。为了实现这一目的，每封信会随机分配贷款利率。按南非消费贷款生意的惯例，一般是报出每月的单一利率（非复利）。测试利率的范围是每月3.25%至11.75%。对于这些客户而言，3.25%是真正的便宜货，还不到该公司正常最低值的一半。11.75%通常是提供给信用最差借款人的最高利率。

正如你所料，客户接受低利率贷款的可能性要比高利率大。通过跟踪特定电子邮件的反应，研究人员得以判断哪些因素促使客户提出了贷

款申请。他们发现，性别有关系，种族却不要紧。严格地说，性别效应是“男人的事”。在给定利率下，要是电子邮件里有女性的照片，男性客户会更容易接受贷款。对于女性客户来说，照片则没影响。

靠性诱惑卖东西，让穿着比基尼的女模特躺在车上，这类做法古已有之，这或许会让你对本次实验产生错误的认识。事实上，电子邮件里附的照片只是简单的职场年轻女性黑白照。较之电子邮件里没有照片或附加的是男性照片的男客户，收到电子邮件中附加普通女性照片的男人，提出贷款申请的可能性更大。在附加男性照片和没有照片的电子邮件之间，响应率上不存在有意义的差别。这暗示了女性照片是关键，并不单纯是所谓“为没有人情味的企业放上一张人性的面孔”在起作用。

这儿还有一件真正令人难以置信的事情。穆拉伊纳丹和沙菲尔的研究小组计算出，在致男客户的信件中附加女性照片，由此带来的贷款申请量跟将利率降低4.5个百分点所带来的申请量是一样的。这可是每月4.5个百分点的区别啊！一年就有54%的利率了。

“为什么会这样呢？”这是对心理学提出的最棘手问题。南非人能看到卡尔文·克莱恩（Calvin Klein）的广告，能看到《恶搞之家》（*The Family Guy*）动画片，赤裸裸的限制级影片也并不罕见。故此，我们还是这样假设比较保险：顾客并没有想着“我要见见这火热的辣妹，为了它我值得承担高得吓人的利率”。他们不可能是有意识地在根据照片做经济决定。

是把垃圾邮件删掉，还是读读看，一般不过是一瞬间的决定。一个可能说得通的解释是，男人就是喜欢女人照片，所以读了邮件的人比删掉它的人稍微多一些。男人的照片对女性起不到同样作用，要不然结果上就该表现出来。

另一种假说跟前一种说法并不排斥，即照片启动了性别行为的自动模式。男性以为女人不擅长讨价还价（或者是做事时表现得如此——这些全都可以是无意识的）。把一张女性的脸放在特定利率的旁边，让男人觉得这是一笔划算的交易。这就好像女性是一种人形参考价，它仿佛在无声地说：“这个利率比你从男性那里拿到的要低哦！”类似说得通的心理学解释还有很多。在许多社会环境和实验中，男性对其他男性持竞争态度，对女性则不然。看到女性的照片，可能会叫男性放下通常状况下想要得到最优惠利率的焦虑。

一个小照片=大笔的新增利润。这里肯定有陷阱。沙菲尔的研究小

组梳理数据，想寻找出证据证明些什么，什么都好，反正别是它表面上看起来的那样就行。比方说，他们想知道，“落入”照片“圈套”的客户是不是信贷风险更大的群体。他们对照片的反应说不定是系统化的，因为他们一贯做出糟糕的财务决策。这样一来，指望靠照片获得额外利润的念头也就可以打消了。

研究人员没有找到统计上的证据，说明容易受女性照片影响的客户在收入水平、教育、贷款偿还率上有什么与众不同的地方。相反，用照片增加利润比单纯提高利息要更有效。至少，利率高的贷款公司客户也会比较少，而支付较高利率的客户通常更容易赖账。依靠照片，公司可以用高利率吸引更多客户，而又不会增加赖账不还的风险。这就是性别的力量。

第32章 全是睾酮和催产素惹的祸

哈佛大学进化动力项目（Program for Evolutionary Dynamics）的工作人员特伦斯·伯纳姆（Terence Burnham）做了一次有关睾酮水平和谈判的实验，这引起了广泛讨论。

价格实验

它是一场最后通牒博弈，每名提议者有40美元可供分配。按要求，他们可以选择自己留15美元（另外25美元给响应者）或留35美元（另外5美元给响应者）。这就强迫提议者要么表现得比自己本心里更慷慨，要么显得过于吝啬，以致遭到对方的否决。这一实验还有其他一些新奇的改动：所有被试均为男性，都要提交唾液样本，测试睾酮含量信息。这样一来，伯纳姆就可以分析睾酮水平对博弈行为的影响了。

7名睾酮水平高的响应者，5人拒绝了侮辱性的5美元分配提议。其他的19名男性，睾酮水平处于平均值或偏低，只有1人拒绝了5美元的提议。拒绝提议的人里，睾酮水平高者占了八成。

这样的结果极具煽动性，因为否决权是最后通牒博弈的情绪核心。其他的一切不过是靠逻辑推动的。提议者表现“慷慨”，是为了保证自己不遭否决。博弈中的性别差异或许反映出了一种常见的信念：女性行使否决权的可能性较小。每当人们设定价格时，类似的行为也就运作起来了。谈判者怒气冲天地离开谈判桌，对面街区那个愤怒的伙计跑去取消了有线电视，因为他们提高了他的费用；抗议税收过高的人，为打官司付出的律师费比税款高得多，很快就把自己拖得破了产……这些都是响应者行使否决权的活生生的例子。究其根本，他们是对价格敏感的消费者，宁肯自己付出高昂的成本，也要拒绝过分的高价。这样的人，通常是男性。

伯纳姆认为，响应者否决出价，是为了避免显得太顺从。纵览人类进化史，金钱、逻辑和公正等概念是后来才出现的。最后通牒博弈和现

实世界价格设定中看到的情绪行为，有可能立足于更基本、更生物性的动机。在市场社会里，金钱是社会主导地位的媒介，是“雄性带头大哥”打动配偶的一种方式，它仪式性地削弱竞争对手。说一个支付太高价格的男性“被搞定了”，显然不是什么巧合。

单单威胁要否决，就能起到一定的阻慑作用，故此它也影响着实际行使否决权的频率。人人都不太愿意发起不公平的出价、设定不公道的价格，因为他们知道自己可能会吃不了兜着走。从这个意义上来说，睾酮水平高的少数人帮助创造了我们的价格世界。

睾酮决定着男性的性发育和两性情欲。它在社会主导行为中扮演着重要的角色。“睾酮中毒”这个表述证明了，大众普遍相信睾酮太多会导致冲动、不当的侵害的发生。但对这一观点的怀疑也是古已有之的。大多数实验都是在动物身上完成的，而对人类的研究偶尔会卷入性别政治话题。有人担心，把睾酮水平和攻击性联系起来，是为男性做出不可原谅的行为找借口。（另一种观点认为，它为“男人就是猪头”提供了科学证据。）

在动物和人类身上，睾酮主要跟面对他人挑衅所做的进攻反应存在联系，跟主动挑起事端的联系并不明显。瑞士研究者在20世纪80年代做过研究，没有找到说明睾酮水平高的年轻人爱挑起事端，跟其他男孩打架的相关证据。但他们同时发现，这类年轻人确实更容易跟教师顶嘴。

有人做了这样的实验，在最后通牒博弈里给被试注射睾酮。“基本上，我们是在研究里创造了‘雄性带头大哥’。”神经经济学家保罗·扎克（Paul Zak）说。注射睾酮的作用跟天生睾酮水平高一样，它令被试更可能否绝对方的低出价。

哈佛大学心理学家埃琳娜·科里（Elena Kouri）和同事们设计了一种博弈。

价格实验

每名被试单独坐在一个按钮前面。研究人员告诉他们，按下按钮，就可以减少付给看不见的伙伴的金钱数额。他的合作伙伴面前也有类似的按钮，他可以采取报复行动，减少当事人的回报。这样一来，博弈就跟核僵局差不多了。谁都不应该首先按下开始键。

为了让事情变得更有意思，科里的研究小组告诉被试，他们的合作伙伴按下了按钮。注射了睾酮的被试更倾向于采取大规模的报复行动，反复按下按钮。不过，较之对照组，注射了睾酮的被试并不见得更容易首先按下按钮。

社会主导地位是一件相对的事。雄性带头大哥就是比周围其他雄性拥有更多金钱、权力的人。绝对数字并不那么重要。两头雄鹿争夺一头雌鹿，贪图的并不是什么双赢结局，而是要比其他雄性做得更好。把这一点放到最后通牒博弈的背景下：要是对手得到95%，你赢到一毛钱是没好处的。双方什么都得不到，这样反倒更好。睾酮推动的就是这样的看法。

有个笑话这样说道：

在北方的树林里，一名猎人穿上了一双昂贵的耐克跑鞋。“嘿，干吗穿跑鞋呀？”同伴问。

“免得我们碰到大灰熊。”猎人说。

“你觉得你能跑过熊瞎子？”

“我用不着跑得比熊快，”猎人解释说，“只要跑得比你快就行了。”

好吧，你正要去跟人谈价格，你想必很清楚自己的攻击性有多强。那么，了解对方那家伙的睾酮水平会有点帮助，对吧？看看他的无名指，你兴许能得到些线索。

戴没戴结婚戒指？研究显示，已婚男性的睾酮激素水平比单身男性的要低。

比比他的无名指和食指的长度。无名指跟食指的长度比，是由产前对雄激素（这一激素决定了胎儿的性别）的接触量所决定的。新近有大量研究指出，无名指比食指长的男性，擅长竞技体育和做交易，在最后通牒博弈中拒绝低报价的可能性更高。剑桥大学约翰·科茨（John Coates）领导的研究小组检验了金融交易员，发现无名指跟食指的长度比和交易成功率存在联系，睾酮和交易成功率也存在联系。

大体上可以这么说，在棘手的谈判里，要是对方是男性且无名指比食指短，又戴着结婚戒指，那你的运气是再好不过了。

不少激素是以阴阳成对的形式出现的。对价格决定来说，睾酮似乎是跟催产素配对的。1953年，文森特·迪维尼奥（Vincent du Vigneaud）

分离出了催产素，这种激素是在分娩和哺乳期间自然分泌的。它有助于促成母子之间建立起直接的感情纽带。男女两性都会产生催产素，和睾酮一样，它也会影响行为。在进行性相关的亲密活动时，催产素水平会提高。在实验中为被试注射催产素，能增加金钱决定中的信任行为。

“神经经济学”（neuroeconomics）这个术语，据说是保罗·扎克提出的。他发现，在最后通牒博弈里注射了催产素的被试，慷慨程度有大幅提高。2007年，他和安吉拉·斯坦顿（Angela Stanton）、希拉·阿曼迪（Sheila Ahmadi）共同做了一次实验，注射了催产素的提议者开价多了21%（安慰剂组的开价平均为4.03美元，催产素组则为4.86美元）。

催产素不影响响应者可接受的最低报价（跟睾酮不一样）。扎克的研究小组还尝试了独裁者博弈，发现催产素对其没有明显影响。他们的结论是，催产素影响战略性的金钱决定，也就是当事者务必要考虑自己行为会给对方造成怎样感受的决定。当催产素水平高的时候，提议者的同情心更强，这成了慷慨做法的幕后推手。

扎克早在高中生时代就对价格决策发生了兴趣。他当时在圣巴巴拉城外的一处加油站工作，有顾客走进来说，他在洗手间的地板上发现了一串珍珠项链。不久，电话铃响了，打电话的人说他掉了一串珍珠项链，那是他买给妻子的礼物。扎克告诉他，项链找到了。太好了，那家伙许诺要给200美元的奖励。他在半小时内就赶过来。与此同时，捡到项链的人说自己要参加一场重要的面试，必须离开。他同意与扎克对半分奖金，一人一半。扎克答应了，从收银机里拿出100美元，换回了项链。不用说，“丢失”项链的家伙根本没现身。这串项链只值两美元。这出“最后通牒博弈”也叫作“凑份子骗局”，是史上最古老的骗术之一。

在评估受扎克研究启发而出现的一些产品时，有必要把这个故事记在脑袋里。网络零售商现在会向指望拿下更划算交易的推销员叫卖催产素喷剂（当然，扎克跟这类产品是没有任何关系的）。有一种这样的喷剂，美其名曰“信任液”，卖49.95美元，只够用两个月的（“如达不到效果，保证全额退款！”）。该网站罗列了五花八门的证明信，其中有一封信是一名酒保写来的，他说这种喷剂让他的小费翻了5倍。网站上的建议如下：

- 销售人员如何使用信任液；
- 每天早晨洗澡后使用信任液；

- 和你喜欢的古龙水或香水一起使用；
- 在整整一天里，无味的催产素将从你身上释放出来；
- 在销售会议上发挥你的正常表现，看看潜在客户会有多么渴望从你那儿买东西。他们不会怀疑自己，只是莫名地受了你和你产品的吸引；
- 你会向潜在客户和现有客户寄送答谢卡吗？往信封上喷一些信任液，见证奇迹的诞生。尽管人们闻不到它，但信任液却提高了他们对你的信任。

照我猜，催产素在最后通牒博弈里的影响如此巨大，恐怕没有几个研究人员会怀疑它也会影响商业决策。喷剂的这部分宣传并不见得疯狂。疯狂的是把它当成产品卖。扎克在实验里是把40国际单位(31)的催产素直接喷进了被试的鼻腔。想想看，要是你这么做，该怎么向大老远赶来的客户解释呢？卖信任液的商家暗示你可以把它当成召唤金钱的香水使用。事实并非如此。催产素并非挥发性质的。往自己的身上喷，或者往答谢卡上喷，并不会对别人造成多大的影响。大部分的建议用法，对使用者本身的影响要大于那些毫无戒心的“受害者”。要是喷剂真的管用，使用者本人必定是头一个受它影响的，没准儿慷慨得连内裤都会想要送给别人呢！

第33章 百万美元俱乐部

1997年，通用电气公司的一家子公司一反常态，给出了一笔极端慷慨的工资。杰里·宋飞（Jerry Seinfeld），美国全国广播公司电视台热门喜剧的明星，宣布自己打算辞职。他那时每拍一集《宋飞正传》，能挣到100万美元，这已经是个天价了。美国全国广播公司听说他萌生去意之后，主动提出把每集的报酬涨到500万美元——只要他答应再多拍一季。

这个提议相当够意思。美国全国广播公司每年能从《宋飞正传》的广告和各地联合播放版权费里挣大约两亿美元。也就是说，这出电视剧每季20多集，每集能带来大约900万美元的利润。为了不丢掉这棵摇钱树，美国全国广播公司宁肯选择“超公平”的做法——把半数的利润都拱手献给大明星。

宋飞不屑一顾。尽管他的节目依然搞笑，他还是铁了心地要辞职。不可避免地，美国全国广播公司开价的消息传了出去，很快就上了各地的娱乐新闻。广播电视台的大佬们肯定指望所有人都以为宋飞是个特殊的例子，遭到拒绝、未能兑现的500万美元开价并未成为先例。

各层食物链上的演员们却不这么想。接下来的几年里，明星以及各色配角的加薪要求前所未有地白热化了。

2002年，《老友记》（*Friends*）的主角们集体开价，要求薪水涨到每集100万美元。这可是6位“老友”每人100万美元啊！《人人都爱雷蒙德》（*Everybody Loves Raymond*）中的主角雷·罗曼诺（Ray Romano）要求每集80万美元，《欢乐一家亲》（*Frasier*）的主演凯尔希·格兰莫（Kelsey Grammer）要求每集160万美元。詹姆斯·甘多费尼（James Gandolfini）拒绝继续出演《黑道家族》（*The Sopranos*），致使剧集草草收尾，因为他发现自己的收入才跟《欢乐一家亲》里的管家一样多。

一个电视明星值多少钱？从这个意义出发，建筑工头、棒球手或者美国总统又该值多少钱？劳务经济学把工资看成人才的供给和需求之间理性权衡的结果，要不，就当成是就职者在闲暇的欲望与购物金钱的欲望之间深思熟虑的选择。可到了最近，行为经济学家收集了一批案例，

说工资有可能跟价格一样，是随心所欲定下来的。“我们怀疑，挣工资的人面对消费与闲暇的权衡，对自己的时间值多少钱并没有准确的认识，甚至也不清楚自己在别家公司能挣多少钱，”阿雷利、鲁温斯坦和普雷莱克写道，“换句话说，工人关心工资的变化，但对工资的绝对水平以及它相对于其他公司同类工人的薪资水平如何，他们是比较迟钝的。”他们指出，有一句老话心照不宣地说明了工资的任意连贯性：所谓富人，就是总能比自家连襟多挣100块的家伙。

按通货膨胀调整后，收入最高群体的所得相差极大。以CEO为例，经济政策研究所算出了一个广泛采用的比例数。根据他们的数据，2005年，美国高层企业管理人员的收入是英国同行的1.8倍，是日本同行的4倍。另一个指标是美国CEO收入与普通工人的比率（如图33-1所示）。2007年，这个数字是275。这变化可真够大的，里根时代才不过50上下，20世纪60年代是25。

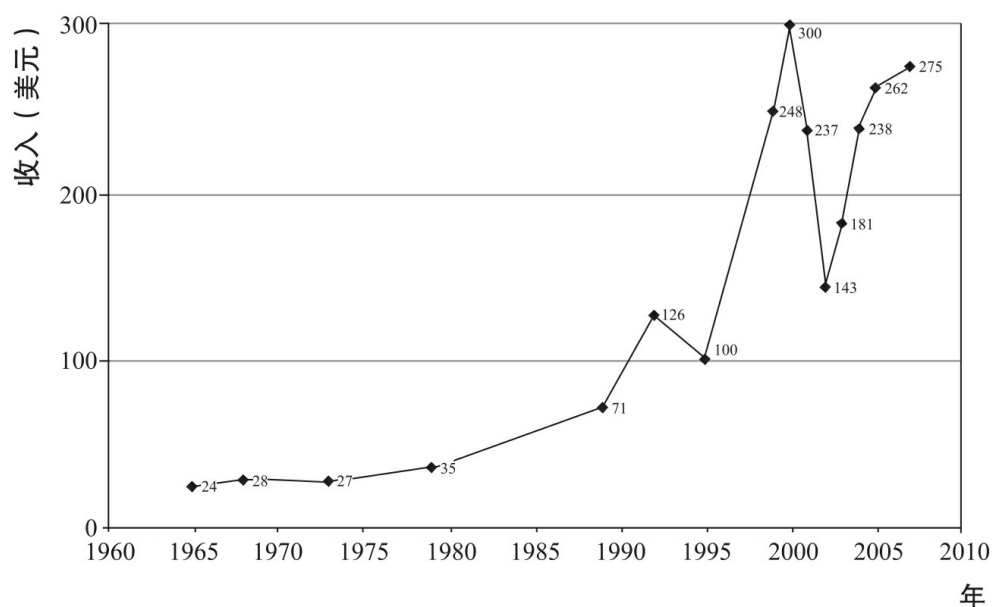


图33-1 CEO与普通工人收入之比（美国）

20世纪90年代初，美国参议员特德·肯尼迪（Ted Kennedy）带头哀叹这一变化趋势。20世纪80年代时，普通工人的收入涨幅不过勉强跟上了通货膨胀的速度，CEO的收入却翻了一倍。美国国会做出反应，在1993年制定了一部法律，取消了百万美元收入门槛之上的若干税收减免政策。

此举非但没有起到限制CEO薪酬的作用，反倒把百万美元的门槛弄

得像是锚点了。它对着企业世界脚步稍慢的大部队高声嚷嚷：7位数的薪水是可以的（为什么我没挣到呢）。1989年，也就是该法律颁布的4年之前，CEO与普通工人的收入之比是71。2000年，它已经蹿上了300。“要是评选无意识后果的排行榜，”管理监督小组“企业图书馆”（Corporate Library）的内尔·米诺（Nell Minow）说，“这部法律肯定排名最靠前。”

劳动者和管理层之间的阶级斗争打开了新战线——管理层和股东闹开了。为回应股东对不合理的CEO高薪的关注，证券和交易委员会颁布新规定，要求对高管薪酬曝光。“我绝对认为‘高管薪酬’会下降，因为曝光多叫人尴尬啊！”曝光规定的设计师格拉芙·克里斯托（Graef Crystal）回忆说，“可事实证明，要是有人年收入高高盘旋在2亿美元之上，他是绝不会尴尬的。”

身为苹果公司的CEO，史蒂夫·乔布斯的薪水是1美元。他的真正报酬主要来自公司给他的限制性股票。2006年，这笔钱达到了6.47亿美元，约占苹果56亿美元利润的11.6%。

“独行侠理论”（Lone Ranger Theory）认为，公司的股票市值，CEO负主要责任。对乔布斯和苹果公司，这一点倒不难相信；在公众心目中，他俩简直是同义词。2008年冒出一连串的谣言，开了好多电话会议，说是乔布斯的健康状况堪忧，于是苹果公司的股价遭受重挫。一些统计学研究声称CEO与股票价值存在有力的联系，就算是公众并不熟悉的普通CEO也一样。

接受独行侠理论，CEO的薪水支票开得再大也可以想象了。

典型的例子是杰克·韦尔奇。他执掌通用电气公司20多年，公司的市场价值从140亿美元膨胀到5 000亿美元。“取得了如此成就的CEO值多少钱呢？”乔治梅森大学的经济学家沃尔特·E.威廉斯（Walter E. Williams）最近问道，“倘若韦尔奇只拿相当于通用电气公司1%价值增幅的一半，那他的报酬总额也差不多有25亿美元了，但他实际上只拿了寥寥几千万美元。”

独行侠理论的问题在于，你很难判断通用电气公司的发展有多大部分归功于韦尔奇，多大部分归功于运气（这里提到“运气”，是因为我一时找不到更合适的字眼）。

1981年到2001年，单单是通货膨胀，就让通用电气公司的价值增加了1倍。同一时期碰上股票牛市，其标准普尔股票增加了9倍。在这两点上，韦尔奇大概说不上有什么功劳吧！韦尔奇还在合适的时间退了休——就在“9·11”事件前5天，运气实在好得叫人

没法相信。“9·11”事件之后，世贸中心向保险公司提出赔偿，而相关的保险公司恰为通用电气公司所有。通用电气公司一下子就损失了6亿美元，接下来的几年里更是损失了几十亿美元。不过，这可就不是韦尔奇要头疼的问题了，21世纪初惨淡的股市自然也跟他毫无关系。

到韦尔奇的继任者杰夫·伊梅尔特（Jeff Immelt）手上时，通用电气公司的市场价值已减少到960亿美元左右。你可能会说，伊梅尔特是个倒霉蛋。他一上任，通用电气公司80%的股东财富就化为了泡影。按独行侠理论来说，这一切全是伊梅尔特的错。

这当然十分荒唐。伊梅尔特是个有天赋又苦干的经理人，有人说他和韦尔奇一样优秀。伊梅尔特恐怕接受不了这样的提议：别从通用电气公司领工资了，直接赔偿股东们的损失吧！他会坚决地指出，通用电气公司股票价值下跌是时机不好，运气糟糕。那么，韦尔奇的成功又有多大的运气成分呢？有什么办法能判断出来吗？

2006年，韦尔奇在MSNBC电视台的《硬式棒球》（*Hardball*）节目上亮相，援引了企业界最爱的平民风格比喻。“CEO们就好像棒球选手，”韦尔奇说，“这些人有什么比率吗？人人都掏出支票本和钱包想签下个明星球手，经纪人乐得开了花。他们有三个星期的斡旋时间。”

主持人克里斯·马修斯（Chris Matthews）也很帮忙，他回忆了美国20年代最著名的棒球选手贝比·鲁斯（Babe Ruth）说的俏皮话。人家问鲁斯为什么挣得比总统还多，他回答道：“我去年干得比他漂亮。”

其实，球员的薪水跟CEO的一样叫人觉得神秘。1922年，鲁斯成为头号球手，每年赚5万美元。这差不多相当于现在的64万美元。2000年，亚历克斯·罗德里格斯（Alex Rodriguez）签下一份10年期合同，队里每年给他2 500多万美元。倘若按通货膨胀加以调整，罗德里格斯也比鲁斯的收入高49倍。为什么会这样呢？不可能只是因为类固醇。罗德里格斯和棒球比赛在流行文化上的脚印，也比不上鲁斯那个时代的辉煌。自20世纪20年代以来，美国人口增加了差不多3倍，电视让更多人可以观看棒球比赛，也增加了棒球业的广告收入。然而，如今要打发闲暇时间，方式真是太多了。

假设我们把棒球球手的薪水根据通货膨胀加以调整，再除以鲁斯1922年按通货膨胀调整后的薪水。所得的结果，我们就叫它“贝比·鲁斯比”好了。表33-1显示，尽管棒球在体育比赛和娱乐方式里所占的比例越来越小，球手们的薪水却在一个劲儿地飞涨。

表33-1 贝比·鲁斯比

年	球手	薪水 (票面值) (美元)	贝比·鲁斯比
1922	贝比·鲁斯	50 000	1.00
1947	汉克·格林伯格	100 000	1.49
1979	诺兰·莱恩	1 000 000	4.63
1991	罗杰·克莱蒙斯	5 380 000	13.27
2000	亚历克斯·罗德里格斯	25 200 000	49.17

工资结构存在很大的一致性。大联盟球手比小联盟球手挣得多，伊梅尔特挣得比他手下的副总裁多，副总裁挣得比生产线上装灯泡的家伙多。薪水的任意性到底有多大，我们都不太了解。我们都喜欢这么想：勇气胜过运气，“明星”人才能扭转球队或跨国企业的败局。但这很难被证实，给它设置什么价格就更难确定了。

在实践中，顶级薪水是留待少数人判断的，剩下的人只能耸耸肩，觉得那数字恐怕不会不合理。这不光是供给和需求，更是锚定和调整。

第34章 淘气的市场先生

深夜，你翻来覆去地换着电视频道，看到一种神奇新产品的广告。一个黑色小方盒子，每年一次地往外喷一张崭新的钞票。推销员向你保证，它完全合法，你想怎么花这笔钱都可以。这个盒子年年都能产出1美元——今年，明年，后年，大后年……直到永远！你愿意为这样的产品出多少钱呢？

评估这个盒子的价值，方法之一是想象你一年可以怎样花掉1美元。你可以把它当成小费，在圣诞节送给某个你不怎么喜欢的人，或者明年夏天把一份快餐订单改成加大份。你大概会得出结论，这个盒子至少值1美元。你第一年就能回本，之后，它的产出都算是你的外快。

你也有理由认为，这盒子的价值应该低于你目前的预期寿命年限数，因为这个年限决定了你还能从它那儿得到多少美元。顺便说一下，原来的主人去世以后，盒子还将继续工作，你可以把它转赠给你喜欢的人。

你对这个盒子的估价应该跟你延迟享受的能力有些关系。也就是说，你现在得放弃辛苦挣来的一部分钱，以购入价格的形式换回之后的一条收入流。把焦点放在眼前这一刻的人（也就是老把信用卡刷爆的人），恐怕对这个盒子完全不感兴趣。看重长期持有它的人兴许愿意付相对较高的价格。

有一点很清楚。这里没有什么所谓绝对“正确”的价格。要是你想做一个锚定实验，那你恐怕会发现自己可以操纵价格。要是购物节目里的现场观众高声呼叫要以2美元的价格购买盒子，大多数观众兴许会觉得这是一个合理的价格。而要是现场群众觉得它值60美元，那60美元这个价格也就成了合理的了。

本杰明·格雷厄姆（Benjamin Graham）[\(32\)](#)，价值投资的传奇创始人，对每年产生1美元的黑盒子报出了一个简单的估价：8.5美元。格雷厄姆说的其实是股票。一份股票能在将来产生收益流。把股价除以每股

的盈利，你就得到了市盈率（价格与盈利之比）。它能说明买家为将来的1美元收入付出多少钱。由于黑匣子每年产生1美元的收入，所以你对它的出价，按美元来算，就等于它的市盈率。根据格雷厄姆的分析，一家公司的股票，倘若没有收入增长，应在8.5的市盈率上卖出。

格雷厄姆用“市场先生”的形象讽刺了投资者的价格心理学。市场先生是个好心肠的怪叔叔，每个工作日都会出现在你家门口，提议你购买或卖出股票。市场先生的价格每天都不一样。尽管市场先生很固执，你却不必担心是否得罪了他。不管你接不接受他的提议，第二天他总会带着新的价格卷土重来。

按格雷厄姆的说法，市场先生其实并不知道股票值多少钱。聪明的投资者是可以靠这一点赚钱的。有时，市场先生会建议以超出其价值的价格购买股票。你应该趁机卖掉！有时，市场先生的出价低于股票的价值。你应该赶紧买进！

这一招对格雷厄姆管用，对他的少数弟子——比如沃伦·巴菲特，也管用。可遵循格雷厄姆的建议，说起来容易做起来难。牛市（不客气地说，也就是泡沫）的时候，市场先生每天都给出天价，就好像股票永远也不会跌似的。大多数投资者都没办法抵御这样的诱惑。可市场先生怎么可能一天又一天地犯下这么离谱的错误呢？

早在1982年，斯坦福大学经济学家肯尼斯·阿罗（Kenneth Arrow）就指出，特沃斯基和卡尼曼的研究兴许可以解释股市泡沫。劳伦斯·萨默斯（Lawrence Summers）[\(33\)](#)在1986年发表的一篇论文中进一步探讨了这一主题，文章名为《股市是否合理地反映了基本价值》。萨默斯是第一个对现在所谓“股价的任意连贯性”加以扩充说明的。市场日复一日地对最新的经济消息做出迅速反应，由此产生的价格“随机漫步”，就成了市场了解真正价值的证据。由于股票价格已经反映了对公司未来收益所知的一切，所以只有突如其来的财经消息（无论好坏）才能改变价格。

萨默斯敏锐地指出，这种证据是站不住脚的。随机漫步是有效市场模型的一种预测，就好像你错过了火车是“黑色星期五要倒霉”理论的一种预测一样。你没办法通过它证明任何东西，因为有很多其他原因都可以造成同样的后果。萨默斯勾勒了其中的一种：股票价格包含了强大的任意成分，不过会通过当天的财经新闻进行前后一致的调整。

萨默斯的想法很可怕。它认为股票价格可能是一种集体幻觉。一旦投资者不再盲信，它就会一个跟头跌进谷底。“谁知道道琼斯工业平均

指数的价值应该是多少呢？”1998年，耶鲁大学的罗伯特·希勒（Robert Shiller）问道，“它今天真的‘值’6 000点吗？还是5 000或7 000？2 000或10 000？如今没有什么一致的经济理论可以回答这些问题。”图34-1是历年来标准普尔指数股票的市盈率。

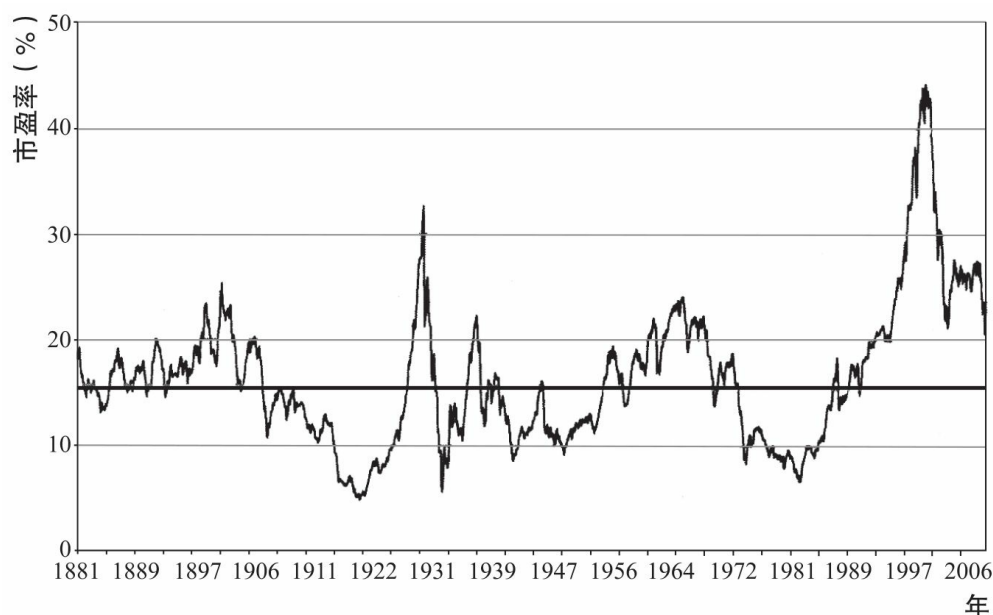


图34-1 1881—2006年标准普尔指数股票的市盈率

这500家公司的股票，目前占美国国内股市总投资的3/4。和黑盒子的价格一样，市盈率代表了延迟享乐的能力。你兴许认为这种能力是人性中的恒定量，或是随着美国消费文化缓缓波动的变量。图表却讲述了一个不同的故事。曲曲折折波动的线就是市盈率（使用前10年的平均盈利，这是希勒所用的指标）。黑色的粗直线是参考线，表明了市盈率的历史平均值是16%。在过去的70多年里，标准普尔的市盈率从不到5%（20世纪20年代）增加到了44%以上（1999年）。

部分变化是合理的。市场正试着预测未来的收益。倘若盈利增长前景不错，市盈率就应当高些，要是前景不容乐观，它就应当低些。利率和税率也会影响市盈率。但观察者，从格雷厄姆到希勒，都认为市盈率的大多数波动都是由投资者情绪不稳所导致的。要是能扫描输入市盈率和股票销量的数据，价格顾问们会得出结论：企业收益的“消费者”们，其需求明显缺乏弹性。这大体就是格雷厄姆的评价。他认为，大多数投资者冲进或退出市场时都是靠情绪在做决定，他们并不怎么关心价格。

很多人都对市场价格心理学做了实验研究。卡默勒就用加利福尼亚

州理工学院的实验经济学和政治学实验室设计出了超简化的股票市场。该实验室是经济学家查尔斯·普洛特（Charles Plott）创办的（他重复过偏好逆转实验）。实验室由若干小格子间组成，每间都配备了电脑。每一次敲击键盘，每一个鼠标动作，软件都会记录并存档。实验结束的时候，研究者可以回放格子间里发生的一切，就好像视频点播电影一样。卡默勒做过这样一次实验。

价格实验

被试得到一种虚拟证券的两股，又得到一些真正的钱。75分钟之内，他们可以自己购买或出售股份，只需在软件里输入买入或卖出命令就行了。软件把买卖双方匹配起来，执行交易。被试明白，倘若在实验中赚了钱，他们是可以带走的。

由于证券是虚构的，被试没法翻查价格。他们只能自己出价、问价。卡默勒把这一切弄得尽量简单。在实验的整个过程中，每一股每隔5分钟就发24美分的股息。故此，在实验中一直持有卡默勒股票的人，能得到总计15次股息分红，即3.6美元。根据严格的价值投资者的标准，股票一开始时值3.6美元，每分一次红就少24美分。随着时间的推移，股票价值图会像下降的台阶一样。

实验开始后，股票的初始交易价格是3美元左右。10分钟后，它已上升到3.50美元附近。之后的绝大部分时间，它都在3.5美元上下徘徊。一直到最后10分钟，现实才控制了局面。随着结束的时刻步步临近，股价暴跌。

卡默勒向被试询问情况。“他们会说，当然了，我知道价格太高，可我看其他人都在高价买进卖出，就觉得我应该买进，收取一两次股息，之后以同样的价格买给其他某个傻瓜。没错儿，有些说法是对的。只要他们能在崩盘前脱身，他们就挣了很多钱，因为另外有些可怜的家伙没能及时脱手，被迫埋了单。”

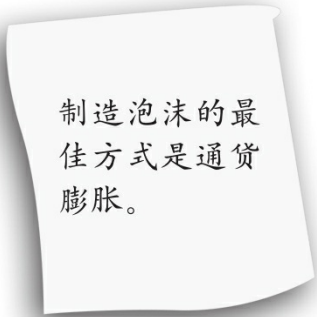
这就是所谓的“大傻瓜理论”。人们在20世纪90年代末买了科技股，在2000年买了房地产，并不一定是因为他们以为价格合理，而是因为他们相信自己能把手里的东西卖给更大的傻瓜，获取利润。

那么，价值投资者（少有的几个不受他人愚弄的人）又是怎样的呢？在卡默勒的实验中，他们在赛场边上坐起了冷板凳。等股票“真

正”的价值降下3.5美元以后，价值投资者就早早出空了自己手里的两股。此后，他们手里就没有更多的股票可卖了，更无意在卖家孜孜以求的价格上买进。所以，价值投资者对市场价格没有影响。

多次重复实验后，卡默勒弄清了如何制造而后再打消泡沫。制造泡沫的最佳方式是依靠通货膨胀。卡默勒曾在实验里不断往虚拟经济里投钱，就跟政府印刷钞票的做法差不多。钱越来越多，股票数目却保持不变，股价便上涨了。卡默勒发现，此时，他可以找回同一组被试，再做一次实验，这回不再有通货膨胀。“如果他们体验过了通货膨胀，”卡默勒解释说，“那么我们就在他们脑海里种下了价格会涨的信念，就像人工降雨一样。”结果“价格真的涨了，全因为这种建立在他们共同经历上的自我应验的预言”。

共同经历同样也是打消泡沫的关键。运行实验，接着找回同一组人重复。这一次，投资者还记得前一轮实验里的崩盘，于是态度更谨慎了。他们不把价格哄抬得太高，很快就开始退出。崩盘来得温和了一些，时间也提前了。第三次重复实验时，完全没有崩盘。价格基本上没有偏离价值投资者的底线。



制造泡沫的最佳方式是通货膨胀。

遗憾的是，对现实中的市场来说，记忆短暂，而泡沫与泡沫之间隔得又太久。作为一个整体，投资大众从未有机会做出决定，看看自己造成的后果，并相应改变自己的行为。没有“土拨鼠日”，故此投资者注定要重蹈黑色星期一的覆辙。

第35章 看在上帝的分上

2007年6月，英国艺术家达米安·赫斯特（Damien Hirst）推出了世界上最昂贵的艺术作品。作品名为“看在上帝的分上”，是一颗镶嵌了8601枚钻石的白金头骨，每一枚钻石的来源都清清白白。它要价5 000万英镑——当时差不多是1亿美元，相当于基里巴斯共和国一年的国内生产总值。“这头骨非同凡响，”流行艺术家彼得·布莱克（Peter Blake）说，“价格看起来还蛮对头的。”

赫斯特靠着创造性定价闯出了一番事业。早先，收藏家查尔斯·萨奇（Charles Saatchi）委托赫斯特制作泡在福尔马林里的大鲨鱼时(34)，艺术家要了个高得离谱的数目——5万英镑。这个数充当的是宣传噱头，为当时还籍籍无名的艺术家造造舆论。这一招奏效了。英国《太阳报》的头条是“5万英镑买条鱼”。2004年，萨奇把鲨鱼以800万美元的价格卖给了对冲基金经理史蒂夫·柯恩（Steve Cohen）。要是鲨鱼的形状再好点，价格无疑还会再高些。2007年，赫斯特其他的活体牲畜和药箱子作品按惯例都要卖7位数的高价。2007年头骨的价格恰好是1991年鲨鱼的1 000倍。赫斯特说，光是头骨上镶嵌的钻石，就花了他2 400万美元。“我们想把钻石弄得到处都是，”他解释道，“连里面都是，就在鼻子的内侧。只要是能放钻石的地方，我们都放上了。”

“是不是很漂亮？”供职于《纽约时报》的艾伦·莱丁（Alan Riding）问，“跟什么比呢？”批评界对赫斯特爱恨不一，头骨把恨他的那批人招了出来。“为了彰显人类的愚蠢和贪婪，这颗闪闪发亮的头骨真是太累了。”《时代周刊》的理查德·拉卡约（Richard Lacayo）抱怨说。伦敦评论家尼克·科恩纠结地说：“价签才是艺术品。”

赫斯特的支持者认为，价签的确是艺术品。作品正是为了折射艺术市场的愚昧与荒唐。赫斯特选择使用钻石并非偶然，毕竟，钻石这种矿物质的价格，就是靠卡特尔人为抬高的。与此同时，批评赫斯特的人不怀好意地预测，头骨在这个世界上留存的时间不会太长。钻石比赫斯特的名气更值钱。不管是谁最后买下了这东西，总有一天钻石会被一颗颗

挖下来卖掉。

对于价格这难以捉摸的幽灵，头骨的短暂历史已经讲述了一个精彩的传说。该作品在伦敦白立方画廊展出后没几天，赫斯特就宣布：“差不多已经卖了……有人很感兴趣。”英国媒体分析，过气的流行偶像乔治·迈克尔（George Michael）可能是潜在买家之一。接下来事情就悄无声息了。看起来，画廊要做成这笔买卖有点儿艰难。到了8月底，据称头骨已经按全价——5 000万英镑，卖给了某个投资集团。画廊的发言人拒绝透露买家身份，也不肯进一步说明细节，只说买家打算稍后转售这件艺术品。



价签才是艺术品。

以更高的价格转售？不管怎么说，这些金融怪才们竟然按全价掏了腰包。按传统，画廊是会给大收藏家打折的。购买世界上最昂贵艺术品的人，似乎够资格享受这个待遇。

买家的身份泄露了。无非是达米安·赫斯特、白立方画廊的老板杰伊·乔普林（Jay Jopling），还有弗兰克·邓菲（Frank Dunphy）——赫斯特的会计。不难理解这是怎么回事。头骨的价格是个锚点，一种推动赫斯特其他作品价格的巧妙方式。不管5 000万英镑的头骨能不能卖出去，维护这一价格的信誉都是最要紧的事。头骨作为宣传噱头，再成功不过了。它没能卖掉会变成新闻，于是，赫斯特和公司炮制了一份财务安排，让他们宣布头骨按全价卖掉了。

这步棋兴许下得更聪明。2008年9月15日，苏富比拍卖行开始对赫斯特（及其工作室）的223件新作品进行空前大拍卖。这一天也是雷曼兄弟公司申请破产的日子，可98%的作品都卖掉了。拍卖成交价最高的是“金牛犊”，牛的犄角和蹄子都用18K金叶制成。它卖到了1 030万美元，创下了赫斯特作品的拍卖纪录。为期两天的大拍卖，拍卖总金额达1亿多英镑，约合2亿美元。

至于头骨，会计邓菲承认它还没卖出去：“不过，它现在的价格翻番了。”

第36章 锚定解毒剂

托马斯·穆斯维勒、弗里茨·施特拉克（Fritz Strack）和蒂姆·菲佛（Tim Pfeiffer）在维尔茨堡大学做了一项实验。

价格实验

把一辆开了10年的欧宝Kadett E给60名德国机械师和汽车经销商看。研究人员之一说，车主的女朋友砸凹了汽车，车主的心里很矛盾，不知道它还值不值得修。他提到，他认为这辆车值2 800马克。“照您看来，这个价是太高还是太低呢？”接下来，他请专家们估计这辆车的现值和维修成本。专家们对汽车价值的平均估计是2 520马克。

研究人员又把同一辆车拉到另一组机械师面前，这回，他们说自己估计这车值5 000马克。结果，第二组专业人士的平均估计是3 563马克，高了足足40%。

到此为止，实验只不过再次示范了现实世界中锚定效应对专业人士的影响。机械师们眼前实实在在地摆着那辆车，一个随意提及的价格却仍然左右了他们。

维尔茨堡实验的目的是测试一种锚定解毒剂。这是一种叫“反向思考”的技术。听到高价值，机械师就开始寻找能证明高价正当性的理由。这些理由存储在记忆里、唾手可得，受此影响，估价就往锚点的方向偏转了。

这意味着只要能让机械师思考锚点数字可能出错的理由（“反向思考”），就能轻易削弱锚定。为了对此进行测试，研究人员又找来两组机械师做进一步的实验。

价格实验

研究人员先说自己认为车值5 000马克，接着便补充道，“昨天我的一位朋友说，他觉得这个价太高（太低）了。对这个价，你怎么看？”

如此罗列正反双方的理由提醒机械师。接着，和之前一样，研究人员询问机械师对汽车的估价。

这下子，被下了高锚点的机械师给出了3 130马克的平均估价（没服下“解毒剂”的机械师报出平均值为3 563马克）。被下了低锚点的机械师给出的平均估价是2 783马克（之前为2 520马克）。

上述两种情况中，“反向思考”都削弱了锚点的力量，令估价不再那么极端。此外，找到更多理由的机械师不容易受锚点的影响。

“反向思考”不是一个新概念。1650年，奥利弗·克伦威尔（Oliver Cromwell）向苏格兰教会的长老发出了那句著名的呼吁：“我恳求你们，看在基督的面上，想一想你们有没有可能是错了！”克伦威尔想要说服他们：以支持查尔斯二世为王来威胁英联邦是错的。他的话长老们置若罔闻，却留在了青史之上。三个世纪后，美国法官勒尼德·汉德（Learned Hand）说，克伦威尔的呼吁应该“写在每一座教堂、学校、法院的门口，还有，我或许可以这么说，也该写在美国每一个立法机构的门口”。

汉德法官提出的观点是，得出结论之前要三思。事先考虑你的判断可能在哪里出错，兴许会让你想到一种之前被忽视的原因，从而改变你的看法。克伦威尔和汉德谈的都是下决定前有意识思考的一面。穆斯维勒的研究小组认为，“反向思考”也能影响决定直观和本能一面。它可以减少锚点对价格的效力。这一点在谈判中会很有用，因为你不可能总是抢先提出一个数。

这些结果还有另一种解释。研究人员所找的机械师，有一半都对汽车的价值持极端乐观的看法。机械师或许并不希望打破客户的泡沫（“顾客永远是对的”）。研究人员曾表示，要是汽车还值那么多钱，他会修好凹痕。凡是想接下这份活的机械师都有理由给出较高的估价。这会造成跟锚定相同的效果，无从判断无意识试探的成分是多少，有意识推销的成分又是多少。

同样道理，研究人员提到自己的朋友持怀疑态度，机械师可以把它

视为谈话里的暗示：听到不同的意见，这个人不会觉得受了冒犯。对坦率的鼓励，也可以解释如上所见的行为模式。

为了解决这个问题，穆斯维勒的研究小组做了第二次实验。

价格实验

他们要维尔茨堡大学的学生估计德国政治家赢取下届选举的概率。比如，总理赫尔穆特·科尔（Helmut Kohl）的胜率是大于还是小于80%。接着又问学生自己对科尔获胜概率的估计是多少。这显示了通常的锚定效应。之后，研究人员又请另一组学生说出科尔可能败北的三个原因，这一回，锚定效应大大减弱，就跟之前在机械师身上看到的一样。

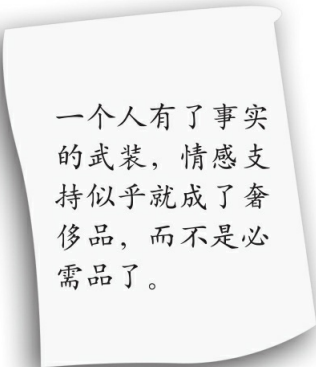
“反向思考”很容易应用。要是碰到经销商、供应商、中介或雇主向你报出一个数字，你深深吸一口气，在有机会思考为什么这个价格有可能不合理之前，不做任何表态。就把它当成游戏好了：尽量想出更多的原因。

许多头脑清醒的企业家对正向思考（或反向思考）的练习不屑一顾。但锚定效应是真实的，只要是能用到的帮助，我们都应该利用起来。

第37章 带上你的好朋友

如今的新型买家会拿着打印好的资料跑去汽车专卖店，这可真叫经销商抓狂。凡是能接入互联网的人都能解开卖车人利润的奥妙。只要花上几美元，就能买到一套经销商财务报告，当中包含最新车型及其可选配件的成本，如送抵目的地的费用、预付定金以及其他没做广告宣传的促销赠品。售卖此类信息的组织一般会建议买家说，经销商在真实成本上加价5%，应算是“公平”的利润。故此，拿着打印资料来的买家，清楚地知道自己打算付多少。他跟“傻瓜蛋”截然两样，需要商家采用不同的应对策略。

有了这样的客户，谈判不再是随口说个价，而更多的是向事实发起挑战。经销商成了抵赖的大师。他们坚称《消费者报告》是错的，互联网是错的，买家算出来的数字是错的。打印小册子上的信息过时了（情况天天都在变）；买家心仪的车型本地已经没货了；不管小册子怎么说，其他买家愿意付经销商报的价，等等。做了功课的买家兴许不怎么相信这些说辞，但不可能完全置之不理。总会出现某个点，买家会厌倦：他提出的每一点事实和合理的推断，经销商都抵赖个一干二净。总之，买家让步了，他付出的价格超过了成本加价5%。也许他相信了经销商的辩白，也许他只是不想再纠缠下去罢了。



一个人有了事实的武装，情感支持似乎就成了奢侈品，而不是必需品了。

很多地方建议购车者使用“好友系统”：带上配偶或朋友，以提供支

持和备选意见。好友系统是“反向思考”的社交形式。你的朋友在必要的时候向经销商提出了相反的意见。不过，我怀疑掌握信息最详尽的买家利用好友系统的可能性恐怕是最小的。一个人有了事实的武装，情感支持似乎就成了奢侈品，而不是必需品了。

社会心理学的经典实验之一就牵涉到了好友系统。1951年，当时在斯沃斯莫尔学院的所罗门·阿希（Solomon Asch）发表了一篇文章论决策的群体压力。

价格实验

被试均为本科男生。被试坐在一张桌子边，桌边还有其他8个人，被试以为这些人都是自己的伙伴。可事实上他们是阿希安排的“内鬼”。研究人员展示了一套“视力测试”，由18张简单的图表构成。图37-1就是一份摹本，按阿希使用的实际尺寸复制而成。仔细看看最左边的直线，再看看右边的三条直线，它们跟左边的直线一样长吗？



图37-1 视力测试摹本

研究人员事先告诉内鬼，头两张图要给出正确答案，其后的图表则交替报出错误和正确答案。真正的被试坐的位置是安排好的，他最后一个作答。在关键题目上，被试要在听到多名内鬼给出同样的错误答案后才能发言。

总体而言，被试有32%的时候会给出错误答案。74%的人至少

给出一次错误答案，还有相当多的人75%的时候都屈服于同辈压力。由于题目无比简单，所以这个数字十分惊人。在没有内鬼的对照组，几乎人人都100%地给出了正确答案。

阿希试图揭示被试在屈从于群体意见时是怎么想的。他听到了三类解释。有些人说，集体报出的错误答案看起来的确是错的，但他们寻思，集体可能是对的。另一组被试告诉阿希，他们知道自己是对的、集体错了，可他们不想引人注目、惹起风波。还有极少数人是真正被洗了脑，就算阿希对实验做了解释，他们也坚持说自己的确看到左边的直线跟1号线一样长。

给出正确答案的被试大多承认他们心里是没底的。“你们或许是对的，但也可能错了！”有个人在实验里对内鬼小组这么说。后来，听说真相之后，他感到“欢欣与宽慰”，他告诉阿希，“实话实说，那时我曾这么想过，‘管它的，我跟其他人一样就行了’。”

阿希又尝试在实验里加入表示同情的“哥们儿”。

价格实验

在一次实验里，桌边坐了两名不知情的被试（其余的人还是内鬼）。这给比较直线长短的那道题带来了显著影响。给出错误答案的比例从32%降到了10.4%。

第一个回答的不知情被试“享受”不到听其他人给出正确答案的“福利”。他有时会屈服，随大流给出错误的答案。这令第二名被试提出异议变得更难了。故此，阿希又使用了另一种安排：先回答的“哥们儿”其实也是内鬼，只不过他得给出正确答案。这再次令错误率降了一半。真正的被试给出错误答案的百分比只有5.5%了。

阿希想弄清多大的群体能左右孤独的被试。答案是，三个人就可以。

倘若只有被试一个人，他们几乎全都会给出正确的答案。当被试一对一地“单挑”报出错误答案的内鬼时，情况也没有太大的不同。可要是一对二的话，被试的错误率就提高了。一旦有三个内鬼，他人影响差不多就达到最大程度了。人数再多，它的变动也不大了。此时，三人为“众”。

在汽车专卖店，“真理”是可以商量的。带一名好友，这是个好主意；带两名好友——达到了三人为众的神奇门槛，就更加不赖。

第38章 义愤理论

琼并不是一个真实存在的孩子，她是卡尼曼、大卫·施卡德（David Schkade）和凯斯·阿桑斯坦（Cass R. Sunstein）在一次有趣实验中虚构出来的人物。他们想看看自己能不能诱导陪审员，像利柏克诉麦当劳案那样，就并不太严重的损伤裁定疯狂的赔偿金。他们还想测试一种简单的、实际的纠正措施，一种能把理智和公正带回民事诉讼制度的办法。

价格实验

琼，一个好奇的6岁孩子，打开了一瓶过敏药的盖子。按说明书来看，这盖子本该是“儿童打不开”的。她吞下了太多药丸，以致在医院躺了好些天。琼的父母把制药公司告上了法庭。审判时，证据中提交的公司文件显示，制药商很清楚这种“儿童打不开”的瓶盖尽管“通常有效”，其失效率却比“本行业任何其他厂家都高得多”。可怜的琼“对所有的药丸都产生了巨大的心理畏惧。每当父母叫她服用维生素、阿司匹林或感冒药等药品，她都会不受控制地哭喊，说她害怕。”

猜猜得克萨斯州奥斯汀的陪审员们认为琼应当获得多少赔偿。
2 200万美元！

1998年，卡尼曼、施卡德和桑斯坦发表文章《公愤和出离的赔偿：惩罚性赔偿心理学》（*Shared Outrage and Erratic Awards: The Psychology of Punitive Damages*），其中他们描述了陪审团裁决的“义愤理论”。他们说，陪审团其实就是一场心理物理学实验，陪审员打分评价自己对被告的行为有多大的义愤。问题是，他们需要把义愤换算成美元，而美元这一量值量表是没有比较标准的。“单纯的美元赔偿具有不可预测性，”文章作者们写道，“它主要是由个体在使用美元量表时的差异造成的（很可能是些毫无意义的差异）。”

他引述了史蒂文斯（从前法律学者们根本不知道有这么一位权威人士）的研究，表明陪审团裁决具有量值量表的许多特点。心理估值上的

误差或“噪声”，等比例地随估值本身的大小而提高。不管你是让一名被试反复做估计，还是比较不同人士所做的估计，情况都是这样。对陪审团来说，这意味着最高额的赔偿裁决有可能最是离谱。此外，陪审团是小样本。12人的样本太少，不足以准确表明公众的意见。这会带来异常高的赔偿，也会造成低得可笑的赔偿（尽管后者很少见诸报端）。

价格实验

被试是得克萨斯州奥斯汀的899名居民，奥斯汀是施卡德所在的得克萨斯大学的总部。被试是从选民名册上招募的，跟履行陪审团义务的居民是同一批人。他们在市中心的一家酒店碰面，读了虚拟诉讼的介绍：一个蒙冤的人状告一家公司。案件中的公司被告已经被裁定有罪，并负担20万美元的赔偿损失费。被试们的任务是确定惩罚性赔偿的数额。

研究人员要一组被试指定美元金额，另一组被试则只需根据一份“义愤”量表给被告的行为打分。这张量表的数值是从0（“完全可以接受”）到6（“绝对无耻”）。还有一组被试要按0（“不处罚”）到6（“极严重的处罚”）分来评价惩处力度。

所有的模拟陪审员都要单独填写问卷，不得与其他人商量。人们对义愤量表和惩罚力度量表（即等级量表）的反应有着极强的联系。可美元赔偿这一量值量表却完全不同。在心理物理学上，这样的结果是可以料到的。

按平均数来看，可怜的琼得到了最高的赔偿金。这一点很荒谬。首先，它并不代表共识。尽管2 200万美元是平均赔偿额，但中位赔偿值只有100万美元。一半的被试认为赔偿数额应为100万或更少，甚至有少数陪审员（2.8%）认为赔偿金应该是0。

这些迥异的美元金额代表陪审团意见不同吗？不。看看等级量表评分，你就能发现一个相当像样的共识。在义愤量表上，陪审员给制药公司的行为平均评分为4.19（总分为6），在惩处量表上平均评分为4.65。回答在中间值上下散乱分布，但大致呈钟形。

但等到指定赔偿金额时，陪审员们的共识就分崩离析了。每个人都不一样。你可能会碰到这样两个人：他们完全同意对被告“严厉处罚”。可一个人认为，严厉处罚意味着10万美元；另一个人却认为它意味着1

亿美元。琼诉制药公司一案的平均数字之所以高，原因在于有好几个人判出了天文数字。在计算平均值时，他们的评价产生了巨大影响。

当然，真正的陪审团不会把每名陪审员的数字都平均计算。他们彼此争论金额，并尝试对外阐明原因（正如对利柏克诉麦当劳案的报道）。不过，也有研究显示，审议群体，尤其是陪审团，在做判断时不见得比单个成员更出色。当人人都独立进行判断时，“群众智慧”的效果最好。陪审团甚至可能放大成员的偏见。倘若最先发言的陪审员报出了一个高得惊人的金额，上述情况就有可能出现。“陪审团裁定赔偿金额的不可预测性和典型失真性，在实验室条件下可以很容易地再现出来，”研究小组写道，“在这类情况下，我们认为裁决是极不稳定的，故此容易受审判或陪审团审议过程中出现的任意锚点所影响。”

琼一案中2 200万美元的平均赔偿额，跟其他假设场景里的美元金额并不吻合。对此，最好的证据是，研究人员还检验了琼场景的另一版本。

价格实验

一些陪审员读到的案情简介是，琼过量服用药物，永久性地损伤了她的呼吸系统，“这会让她的余生更容易患上与呼吸有关的疾病，如哮喘或肺气肿。”这些陪审员给出的平均赔偿额是1790万美元——低于我们最开始介绍的那个场景（那里面，琼不过是害怕药丸而已）。这并不意味有谁真的以为呼吸系统的永久性损伤不严重。没有哪名陪审员看过故事的两个版本；每一回，陪审员都是从奥斯汀的选民中随机抽选的（跟真正的陪审团一样）。显然，在原告病情并不太严重的那个场景里，判处极端高赔偿金额的陪审员碰巧多了些。

还是老样子，等级量表的评分更为一致。较之害怕药丸版本，琼呼吸系统永久受损的版本得到的义愤度更高，惩罚分值也更高，跟逻辑相吻合。这类判断并不随收入、年龄或种族的不同而变化。（在处罚力度评分中，妇女比男人更为严厉。）研究者的结论是，“处罚力度评分的基础是，社会普遍共有的道德直觉。”

美元金额却不是。陪审团裁断出疯狂赔偿金，这个问题的根源在于，就如何把义愤转换成美元，人们没有共识。

卡尼曼、施卡德和桑斯坦利用上述研究发现来解决一些哲学问题。他们写道：正义需要一致性。相同的罪行需要得到相同的惩处。然而，在实践中，所有的情境都不尽相同。这就是为什么我们需要陪审团确保惩处吻合社会的情绪。

文章概述了几种可行的改革方式。其中大多数都涉及让陪审员在评定损害时使用等级量表而非美元量表。他们评分选出惩处力度，而不是美元金额。之后用“转换功能”把惩处评分换算成罚金。这一转换功能可以由法官或立法机构来制定。还有一个更民主的想法是让人民来判断。司法区或整个国家，可以做类似奥斯汀的这种实验，确定公众的惩处意图该换算成多少钱。之后，这种通过实践得出的转换功能就可以用来确定损失赔偿费。每隔几年可以重复实验，确保该功能与公众的想法与时俱进。卡尼曼、施卡德和桑斯坦写道：“我们该如何对社会情绪做出最准确的估计呢？一旦提出这个问题，新的可能性就展现在我们面前了。”但是，目前的制度根本不这么考虑。

第39章 环境影响判断

埃里克·约翰逊（Eric Johnson）是哥伦比亚大学商学院的教授，热情又带点孩子气，年纪不算小了：他在赫伯特·西蒙手下修了博士学位，又跟特沃斯基合作过。约翰逊的学生娜奥米·曼德尔（Naomi Mandel）读到了“启动”概念，想知道它能不能用在网页上。“我说，这是一个很可爱的想法，”约翰逊回忆自己还补充了一句，“它不可能成事儿的。”不管怎么说，曼德尔还是想做一些尝试性研究。“我们不停地做，它回回不落空，”约翰逊说，“我完全没料到数据会这么干净，效果是这么显著。”

曼德尔和约翰逊的实验，发表在《消费者研究期刊》上，在营销界和网络设计群体中造成轰动。长久以来，宣传都说互联网为购物者提供了公平的环境。面对实体店的价格，消费者用不着非接受不可了。买家可以在宽广的网络世界比较各家商店，不再受高压推销手法的操纵……呃，请把最后一句话划掉。曼德尔和约翰逊发现，操纵买家极为简单，就跟调整一段超文本代码差不多。

价格实验

76名大学生参与了这场所谓的在线购物测试。每人访问两个（假的）网站，一个卖沙发，另一个卖汽车。利用网站上的信息，他们要在每一产品的两类款式中选择其一。每人都要进行熟悉的价格与质量权衡，并确定哪个更重要。

实验的一个变量是每个网站主页的背景图片。沙发网站的一些访问者看到的是一张硬币铺在绿底色上的墙纸。其他人看到的则是蓬松云彩的背景（暗示舒适）。汽车站点也有不同的背景图：一种是绿色美元符号，另一种是红橙两色的火焰。

曼德尔和约翰逊认为“请务必注意，我们的启动操纵不是潜意识上的”，“我们的所有被试都能明明白白地看到第一页上的背景，在问到的时候，许多人也记得起墙纸是什么。”可要是问他们墙纸是否影响了其

决定，86%的人说没有。“这种自觉的匮乏，暗示着.....电子环境有可能向消费者提出巨大的挑战。”

价格实验

第二次扩大实验找来了385名互联网用户。这些被试是来自全美各地的成年人，平均年龄和收入都接近普通的互联网用户。一份问卷调查了每名用户在买车和买沙发上有经验。这一回，网站跟踪了用户在每一网页上停留的时长。新手买家的浏览历史清楚地表现出了启动效应。用金钱图片“启动”时，他们花了更长时间比较价格。

专业买家的浏览行为受墙纸图片的影响较小。可他们的选择却照样受了左右。曼德尔和约翰逊怀疑，经验丰富的消费者会觉得，判断哪张沙发柔软舒适、哪辆汽车便宜反倒更简单。“启动”影响了专业人士从记忆中检索的事实。新手必须依靠超文本页面构建类似水准的能力。最终的结果大致相同。背景图片把购物者的思维定式从“价格重要”推到了“质量重要”上。

营销人员已经开始利用这一科学。约翰逊正帮助德国的一家汽车厂商（他不能透露是哪家）重新设计网站。这些应用提出了广告与生俱来的伦理问题。在决策研究面前，我们的伦理，还有我们的经济，全都成了过时货。很大程度上，我们仍然赞同以下观点：人有着固定的价值观。凡是想偷偷摸摸改变这些价值观（“无形的说客”）的事情，都被看成是侵犯个人自由。现实情况是，消费者想要什么，大多是在鼠标的一次次点击之间构建出来的。各种内容细节施展着可测量出来的统计效应。没有哪个消费者愿意觉得自己“受了操纵”。但就某种程度而言，这无异于说鱼不想感受水。

想想看：曼德尔和约翰逊的实验还找了一个对照组，被试看到的网站是中性版本的，完全没有背景图片。他们的选择跟看到金钱背景的被试没有太大不同。这就提出了一个可能性：美国消费者的默认焦点就在价格上。需要“操纵”才能让他们注意到别的东西。

我们这个痴迷于利润的熙攘社会，很少花时间仔细思考“金钱到底有多重要”。但这并不会叫问题自动消失，只是把它推进了潜意识和自动的领域。2004年，克里斯琴·惠勒（Christian Wheeler）和同事们在斯坦福大学做了一次实验，他们找来被试做“视力测试”，之后再行最后

通牒博弈。视力测试包括按大小将照片排序。这只是一个借口，目的是向被试展示一些照片又不引起他们的怀疑。

价格实验

一组被试看到的是跟商业相关的照片（会议桌、西装、公文包），另一组看到的是和商业或金钱无关的图片（风筝、鲸鱼、电源插座）。这使得被试在之后进行的最后通牒博弈里表现迥异。跟对照组比较，看到商业图片的提议者对响应者的出价少14%。看到风筝或鲸鱼图片的被试更倾向于对半分，而不是自己多贪个几分钱。

“对相对轻微的操纵而言，这是很大的影响了，”惠勒说，“人们总是想弄清楚在特定环境下如何采取行动，他们会寻找外在线索指引行为，尤其是在搞不清楚他人期待的时候。要是没有太多明确的线索来帮忙确定环境，我们更有可能根据随手找到的含蓄线索行事。”

价格实验

多年来，纽卡斯尔大学的休息室一直使用“诚实箱”给茶和咖啡收费。任何人都可以自主取用热饮，并按标价把钱投入诚实箱。这样就免去了雇用收银员——收银员的工资说不定比收来的钱更高呢。诚实箱其实就是独裁者博弈的白话版。按道理，人人都应当公平地投入自己的那份钱。但他们也可以少投，甚至完全不投。根据对独裁者博弈进行的研究，人们是否往诚实箱里如约投钱，跟周围是不是有人看大有联系。2006年的实验发现更惊人的内情。

心理学家梅丽莎·贝特森（Melissa Bateson）、丹尼尔·内特（Daniel Nettle）和吉尔伯特·罗伯茨（Gilbert Roberts）自己做了标价签，换下了“诚实箱”上原来的标签。标签上的价格完全一样，只是顶部多了一幅图片。有些标签上印的是一双直愣愣地看着你的眼睛。另一些则印的是鲜花的图片。贝特森的研究小组每周更换标价签，并计算每个星期收到的钱，以观察付费行为是否存在差异。（他们用牛奶的消费量来比较实际使用了多少咖啡和茶。）平均而言，他们发现，当价签上是眼睛图片时，人们投入的钱是有鲜花图片的2.6倍。

“我们本来以为影响会微不足道，结果却是这么大，我很惊讶。”贝

特森说。职场的诚实就像台灯一样能开能关。

独裁者博弈的被试显然有意识地不想显得太自私。可对一幅图片来说，这种解释却是说不通的。“我们的大脑设了默认的程序要对眼睛和面孔做出反应，不管我们是否有意识地晓得它是真的还是假的。”贝特森提出。另一项实验发现镜子也有类似的效果。虽说镜子能改变行为不是什么新闻（想想蜜月套房天花板上的大镜子吧），可它的效果兴许比我们想象中还要大。C.尼尔·马克雷（C. Neil Macrae）、盖伦·V.博登豪森（Galen V. Bodenhausen）和艾伦·B.米尔恩（Alan B. Milne）发现，人在有镜子的房间里，欺骗或表现性别歧视及种族歧视的可能性较低，更愿意助人为乐、努力工作。“通过安排让人们意识到自我，”博登豪森说，“他们更可能停下来思考自己在干什么。”这反过来导致了“更可取的行为方式”。

第40章 金钱，巧克力，幸福

查尔斯·达罗（Charles Darrow）享有游戏《1935年：大萧条中的垄断》（以下简称“垄断”）的专利权。其实他并没有发明这个游戏，而是盗用了他人的想法。游戏讽喻了自由市场资本主义，可它到底是赞成还是反对这种制度，从来没有人说得清。尽管游戏赞美利润，可“垄断”从来都是个贬义词。

“垄断”最终获得了成功，因为它卓有成效地创造出了一个叫人身临其境、内部又和谐统一的世界。玩家们忘记了自己钱包里的真钱，使用“垄断币”——这个词逐渐有了引申意义，象征脱离实际的价格决策。“垄断”里的价格没有任何意义（房子一套100美元），但价格比率把被试需要知道的一切都说清楚了。“垄断”的世界就其本身是合理的——正如你我居住在这个地球上，总想着要搞懂它。

2006年，凯瑟琳·沃斯（Kathleen Vohs）、妮可·米德（Nicole Mead）和米兰达·古德（Miranda Goode）在实验中用到了“垄断”游戏。她们用它来“启动”被试想到金钱（当然，它只是诸多的操纵途径之一）。一组被试玩“垄断”；另一组人坐在电脑显示器边上，屏幕保护程序不停飘动着美钞；一组人要看一幅外国货币的海报了；还有一组人要想象自己贫穷或富裕的样子。沃斯的研究小组发现，所有这些金钱“启动”都有着类似效果。它们会让人变得不好社交，不愿合作。受了金钱启动的被试有以下想法。

想要更多“个人空间”。研究人员告诉每一名被试，他要跟另一名被试进行谈话、互相了解。研究人员要他从房间的角落里搬过一张椅子，放在他自己的椅子旁边。接着，研究人员离开，去接另一名被试。这么做的目的是看被试会把两张椅子隔多远。受了金钱启动的被试总会在两张椅子之间隔开更多的距离。

想要单独工作。研究人员分配给被试一件小差事，允许其选择是单独工作还是跟人合作。绝大多数看了金钱屏幕保护程序的人选择单独工作。绝大多数看了鱼或空白屏幕保护程序的人希望跟人合作。不选择团

队工作其实毫无理由。工作量是一样的，不管一个人做还是两个人做。

想一个人消遣。被试填写一份调查问卷，要求从若干组活动中选出自己喜欢的那一种。每道题都包含两个选项：一种单独的消遣（读一本小说），一种跟家人或朋友共同进行的消遣（跟朋友去咖啡馆）。看到金钱的被试选择单独活动的可能性更大。

不太愿意帮助陌生人。被试从一个房间走到另一个房间，“恰好”碰到一起“人为制造”的小事故：有人（当然是安排好的同伙）掉了27支铅笔。受了金钱启动的人不大可能帮忙捡起铅笔，平均而言，他们拾起的铅笔少得多。

不向人提出帮助请求。被试得到一件“不可能完成的”任务。它目的是看被试要花多长时间才去找人帮忙。受了金钱启动的人在开口之前的挣扎时间比其他入长48%。

对慈善机构的捐赠少。研究人员给被试一个私人机会去向本校的大学生基金会捐款。被试没有理由会察觉这也是实验的一部分。金钱启动组的捐款数额只相当于对照组的58%。

“有些人认为我们的发现表明，金钱让人变得自私，”沃斯和同事们写道，“金钱导致贪婪或自私，这似乎成了当代西方文化必不可少的一部分。”但他们进而指出，实验结果并不适合采用这样过分简单化的阐释。

他们要被试描述自己的情感状态。接受或未接受金钱启动的被试并不存在什么有意义的区别。想到金钱并不会让人“不信任他人，焦虑或骄傲”，这些状态都可以用来解释研究的部分发现。

面对艰巨的任务，一个自私的人可能会立刻会找人帮忙，或是跟伙伴一起分摊工作，免得自己单独操劳。相反，金钱启动让人想要当个独行侠。他们就像是老套的男司机，迷路了也不愿意找人打听。

沃斯的研究小组认为，把金钱启动触发的行为称为“自给自足”更合适。和“垄断”游戏一样，自给自足是松散地起源于市场经济特点的一种“游戏”。游戏的规则说，你应当一个人玩，金钱是你的计分方式。跟其他玩家的互动要遵循公平、互惠的原则。（你不能从别人那里偷“垄断币”，虽说人人都知道它是假的。）玩这个游戏，不是要相信金钱万

能，人际关系不重要，而是以金钱作为临时性共享拟制（temporary shared fiction）。

自给自足只是人类可以玩的诸多游戏之一。它在美国文化和全球强大的市场经济中扮演了一个重要角色。“启动效应或许提供了一种文化可以运作的机制，”卡尼曼提出，“有些文化里包括了一些相当于持久的类似金钱提醒器的东西。其他文化提醒你，有眼睛在看着你。有些文化让你想到‘我们’，有些让你想到‘我’。”

在行为决策实验里，巧克力恐怕是第二受欢迎的激发因素了。人们对巧克力的反应方式，跟他们对金钱的反应方式非常类似。他们试着理性地追求巧克力的最大化，构建巧克力力量值量表。有时人对巧克力的贪恋会叫他们做出奇怪的事情。让我们来看看这些巧克力“经济学”实验。一种怪诞的“辨识感”浮现出来，就好像是看到黑猩猩“模仿”那些再熟悉不过的人类弱点。

奚恺元和张岩做了一次实验，让中国大学生在以下两个选项中进行选择：

（a）回忆并写下生活里一件失败的事，并吃一大块（15克）德芙巧克力。

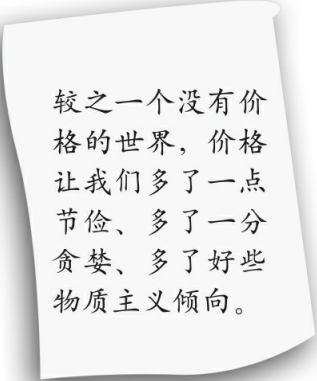
（b）回忆并写下自己生活里一件成功的事，并吃一小块（5克）德芙巧克力。

学生们必须在写的时候吃，不能把巧克力省下来带回家。你可能猜到了，大部分学生（65%）选择了更大的巧克力。精神上的戒律似乎是这样：“对巧克力，能选更多的时候，绝不选更少。”

奚恺元和张岩并未让所有被试都有所选择。对另一组被试，他们只是简单地说，写一件生活里失败的事情，同时吃15克巧克力。随后，他们要给这次经历（边吃巧克力边写作）打分，量表总计9分，从极端不快乐到极端快乐。还有一个小组按要求完成选项（b），也在同一套9分量表上打分。分配到选项（b）的人压倒性地比分配到（a）的更快乐。完成（b）的人得到了一件愉快的任务，得以边做边吃巧克力。他们并不知道自己吃的巧克力比别人的要小。

一旦知道了内情（也就是知道能吃更多巧克力），事情就糟糕了。

人们没办法让自己接受小一点的巧克力。奚恺元和张岩把自己的实验视为“生活的一种缩影”。在现实当中，金钱是苦乐参半的巧克力。我们穷尽一生来寻找最低的价格，最高的薪水，最多的金钱——用这些数字来确认自己的幸福。套用我们熟悉而又浅显的一个观点，金钱买不到幸福，你不能给人际关系标上价格。奚恺元和张岩让这些老套的说教焕发了全新的光彩，像是给这些颂歌镀上了全新的光泽。把金钱当成量值量表并不是一切罪恶的根源。因为金钱是一个数字，数字最容易比较，在跟其他东西进行比较以便做出决定的时候，它得到了过多的权重。较之一个没有价格的世界，价格让我们多了一点节俭、多了一分贪婪、多了好些物质主义倾向。



较之一个没有价格的世界，价格让我们多了一点节俭、多了一分贪婪、多了好些物质主义倾向。

行为决策理论中最难于回答的问题是：人真正想要的是什么？你不能以为价格或选择反映了真正的价值。问题似乎出在问题本身。它假定人在心理上存在一种虚构的精密尺度，里面有着定义严格、跟背景无关的“真正价值”。越来越多的证据证明，事情并非如此。（最广义上的）偏好逆转就是人类的条件与处境。

多年来，行为决策理论家已经很擅长制作设计巧妙的偏好逆转实验了。就拿奚恺元做的一个实验来举例吧。你可以在两块同样精美的巧克力中挑选。一块小一些，呈心形；另一块大，可形状像蟑螂。你会选择哪一块？

奚恺元拿这个难题去问学生和朋友，他发现大多数人选择了蟑螂形的巧克力。可令人哭笑不得的一点是，奚恺元接下来又问人们更喜欢哪块巧克力，大多数人承认是心形的、小的巧克力。

湛庐文化特别奉献

PRICELESS
精读指导

无价：潘多拉星球的故事

精读指路人 孙路弘
营销与销售行为专家

一家外企招聘管培生。要求应聘者上午10点到达指定地点。原来指定地点是一个会议室，会议室墙壁上挂满了一个月以来的剪报，标题有：“应届生愿意接受月薪800元”“硕士生愿就1 000元月薪”“3 000人争抢一个职位”“外企在中国特殊待遇将取消”等。先到的约100多名应聘者陆续开始阅读着墙上的这些剪报。有一些剪报是通过复印机放大了贴到墙上的，会议室的桌子上还放了一些剪报的复印件，公司组织者鼓动前来应聘求职者们随意翻阅。

作为《无价》的读者，你应该不难看出这家外企的人力资源招聘主管在玩什么把戏吧？如果你真的看懂了《无价》，请指出，该公司这么做依赖的是书中的哪条原理，并详细指出该原理在第几页，以及作者选用的是什么例子。

读书就应该这么读。不仅读懂作者用文字描述的例子以及提炼出来的原理，还应借此看透生活中点点滴滴的精明花招。

一个毕业了的高级管理工商管理硕士（EMBA），在投递简历后，被邀请参加面试。面试后，人力资源的招聘主管觉得他确实不错。他是EMBA，在学习这个专业前就有多年的企业高管的经历，与企业要招聘的职位匹配度非常高。于是，第二次面试时，业务部门的主管出席，就许多业务问题同他进行了交流。当然，业务主管也非常满意。最后，业务主管问出了一个关键问题：就你应聘的这个市场总监的职位，你期待的收入是多少呢？该应聘者是这么回答的：在决定全职学习EMBA专业之前，我在公司最后的职位是华东区市场经理，当时的年薪是38万元。入学后，我发现同学们个个都很不俗，最厉害的一个入学前年薪居然都达到七位数了。不过，现在的就业市场对我们这些EMBA的毕业生来说，也不是买方市场，我的一个室友，他上周得到的职位年薪43万元，其实我的期待并不高，还是根据企业自己现有的标准吧！

如果你掌握了精读一本书的方法，你就肯定可以发现书中与这个例子类似的例子在第几页，并且能够提炼出当中核心的本质原理，当然，你也可以预测这个应聘者的年薪范畴大约是多少。

一本《无价》，区区200多页，汇集了50多个经典的心理学实验。每一个实验都有详细的描述，也有因此而产生的经济学结论。其中不少实验的设计者获得了诺贝尔奖，后来的一些学者甚至仅仅是通过对这些实验进行多角度分析就摘得诺贝尔奖的桂冠。你能对这50多个实验进行分类吗？比如，将所有与赌博相关的实验归为一类；将感觉偏差的实验也归到一起；再将社会群体影响类别的实验都挑出来。只要你能将对这些实验的分类落实到书面上，这本书你就读懂65%了。你也就应该能读懂作者创作该书的思维脉络了。何不尝试一下呢？不妨看看我阅读本书后拍得的两张图片，它们可以作为你的参考。

要是—本书中都是实验，那它大概要被归类为实验手册了。难能可贵的是，作者并没打算编写一部社会心理学、经济学实验方面的专著，而是要彰显自身的思考力度和逻辑实力。书中有许多作者根据实验提炼出来的自己的观点，而这些观点在许多经济学家的专著中也屡见不鲜。

一本好书并不一定是一本学术价值多么高的书。学术价值高的通常都是论文，而论文变成图书从而成功地被社会大众所接受的例子并不多见。《无价》是一本让人看得懂的书，当中既有大家看得懂的哲理，也有大家看得明白的逻辑。而这些通常都是阻挡大众学到实用知识的障碍。当然，说此书没有任何阅读障碍，那是不可能的，毕竟《无价》并非小说。所以，要想从本书当中学到真正的价格、价值、社会、心理等珍贵的道理，仍然需要你采用精读的方法阅读本书。

有关精读这本书的实用建议如下：

1. 将书中所有的社会实验汇总一下。
2. 对汇总的结果进行分类。
3. 将书中所有的“不可不知”的内容都抄写下来。
4. 将所有的实验与“不可不知”的内容相匹配。
5. 将书中提到的所有的经济学家罗列出来。
6. 再将这些经济学家的核心观点罗列出来。
7. 将书中提到的价格方面的咨询公司的名单列出来。

8. 在互联网上搜集一下这几个公司的网站及其主要业务范围和成功客户的名单。

9. 将书中测试出来的人类的愚蠢表现罗列一个清单。

10. 将书中人们被价格策略弄晕的情节罗列出来。

然后，你将发现《无价》当中可能隐藏了太多充满邪恶诡计的故事。说的就是源自希腊神话中的故事：

为了惩罚人类，上帝造了一个女人，她的名字叫潘多拉，她的特点就是凡事不经思考，凭感觉去做，概括地说，冲动、盲从、感性就是潘多拉的特点。在好奇心的驱动下，她打开了一个盒子，从盒子里出来了贪婪、虚无、诽谤、嫉妒和痛苦，看到如此不详的景象后，她立刻将盒子关得紧紧的。盒子确实关住了一个没能跑出来的东西，名字叫“希望”！

想想卡梅隆斥巨资打造的影片《阿凡达》中星球的名字吧，就是“潘多拉”。影片结尾，在将来自地球且背负无数罪恶的人类驱出之后，星球留下了“希望”，给“潘多拉”留下未来的光明。

你已经打开了《无价》，你已经看到了太多人为设计的诡计，不要匆忙地掩卷沉思，打开了就多读几次吧，用精读图书的方法放飞希望吧！

以上10个具体的行动内容就是放飞希望的途径，带上你的思维，打开你的笔记本，任由行走的文字报告希望的所在，我在这里等你来相会：

yes4you@gmail.com。

精读拓展

精读《无价》之前有一本书绝对不能错过，它会让你在理解《无价》的原理时一帆风顺。作为心理学畅销书，《影响力》20年来雄踞消费者行为学和应用心理学畅销榜榜首。对于人们为什么会顺从高昂的价格、为什么会接受无理的要求，《影响力》不仅给出了答案，还对你如何运用它提出了指导。

做完《无价》的读书笔记，你最好放松一下大脑，看看那本被全球顶级思想家誉为新锐经典的《看不见的大猩猩》，测一下你是

否属于那50%，这个在哈佛大学完成的实验是有史以来最受关注的心理学实验之一。它会给你不少启示，从而让你的思维更加敏锐。人们为什么愿意为一条家庭影院数据线掏500美元的巨款？为什么经济模型预测不了经济危机？那是因为我们有6种错觉，其中的任何一种都可以把我们耍得团团转。我们以为自己看见的就是真实的世界，但是看见不等于看到，了解不等于知道，错觉就在潘多拉的盒子里，随时检视一下，人类心灵的奥秘就可随之揭开。

另外，面对现代社会铺天盖地的纷乱选项，你最好读一下与丹尼尔·卡尼曼齐名的心理学、经济学跨界大师巴里·施瓦茨（Barry Schwartz）所著的《选择的悖论：用心理学解读人的经济行为》（*The Paradox of Choice: Why More is Less*）——它用心理学理论解读了人类的经济行为。在该书中，施瓦茨提出了一个革命性的观点：幸福意味着拥有自由和选择，但更多的自由和选择并不能带来更强烈的幸福感，相反，选择越多，幸福感越少！这本书用11条实用建议，帮助你在更短的时间里做出更明智的选择，从而指引你一步步走向充实美好的人生！

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

而真正的阅读，应该在书中与智者同行思考，借他们的视角看到世界的多元性，提出比答案更重要的好问题，在不确定的时代中领先起跑。

湛庐阅读 App：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本

怎么读是读者面临的最大阅读障碍

“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”App。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、电子书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。



下载湛庐阅读 App，
坚持亲自阅读，
有声书、电子书、阅读服务，
一站获得。

CHEERS

本书阅读资料包

给你便捷、高效、全面的阅读体验

本书参考资料

湛庐独家策划

- ✓ 参考文献
为了环保、节约纸张，部分图书的参考文献以电子版方式提供
- ✓ 主题书单
编辑精心推荐的延伸阅读书单，助你开启主题式阅读
- ✓ 图片资料
提供部分图片的高清彩色原版大图，方便保存和分享

相关阅读服务

终身学习者必备

- ✓ 电子书
便捷、高效，方便检索，易于携带，随时更新
- ✓ 有声书
保护视力，随时随地，有温度、有情感地听本书
- ✓ 精读班
2~4周，最懂这本书的人带你读完、读懂、读透这本好书
- ✓ 课 程
课程权威专家给你开书单，带你快速浏览一个领域的知识概貌
- ✓ 讲 书
30分钟，大咖给你讲本书，让你挑书不费劲



湛庐编辑为你独家呈现

助你更好获得书里和书外的思想和智慧，请扫码查收！

（阅读资料包的内容因书而异，最终以湛庐阅读App页面为准）

(1) 剧集中的人物，宋飞的怪邻居。——译者注

(2) 顾客待在车里就能点餐的便捷服务。——译者注

(3) 相当于82~88摄氏度。——译者注

(4) 1盎司≈28.35克。——编者注

(5) 诺贝尔经济学奖得主，行为经济学之父。其最新著作《噪声》中文简体字版已由湛庐引进，由浙江教育出版社2021年出版。——编者注

(6) 这里，我采用的版本取自阿莱次年发表的论文，我对问题做了简化，并把货币单位改成了美元。虽说它跟阿莱向萨维奇提的谜题并不完全相同，但足以让你领略阿莱的论点了。

(7) Probability，英语“概率”一词的首字母大写。

(8) 类似于赌场博彩：你拿出一笔钱来赌，就要冒损失它的风险。

(9) 做实验的时候，心理学家大多要拼了命地省钱。被试赌的是“点数”，点数可以兑换成美元，最高赢取金额是8美元。

(10) 芝加哥大学经济学家理查德·泰勒在1980年起的名字。

(11) 换算成如今的币值，大概是587元人民币。

(12) 毕宿五（Aldebaran，意为“追随者”）即金牛座α，是全天第14亮

星。——编者注

(13)20世纪美国知名作家、思想家。——编者注

(14)美国导演，多拍爱情轻喜剧。

(15)德国旧货币单位。1马克 $\approx$ 4.19元人民币。——编者注

(16)1磅 $\approx$ 0.45千克。——编者注

(17)美国演员，曾出演美国电视剧《绝望主妇》。——译者注

(18)一个可穿戴的宽松毛毯。——编者注

(19)类似国内的一元店。——译者注

(20)50 Cent，美国黑人说唱乐手。——编者注

(21)1加仑 $\approx$ 3.78升。——编者注

(22)美国的一种房地产销售模式。

(23)1英亩 $\approx$ 4 046.86平方米。——编者注

(24)1英尺 $\approx$ 0.30米。——编者注

(25)法国贵族，在路易十三时代担任过首相。

(26)成书时，1津巴布韦元=0.020 7元人民币。——编者注

(27)一家打字机公司，后来又与竞争对手合并为斯佩里·兰德公司。

(28)美国顶尖网球运动员。

(29)美国家喻户晓的高尔夫球运动员，辉煌时期是20世纪60年代到70年代。

(30)杰克·韦尔奇曾有绰号叫“中子弹杰克”，暗指他的铁腕管理，下狠手裁员。这里把外号安在他前妻头上，也是暗指其颇具“杀伤力”。——译者注

(31)所谓的国际单位（I. U.）是评估维生素A和D两种营养素含量的计量单位。1微克视网醇当量=1微克视网醇=6微克β-胡萝卜素=3.33国际单位视网醇=10国际单位β-胡萝卜素与其他维生素A先质（1 000微克=1毫克，通常市售β-胡萝卜素一颗约15mg，即25 000国际单位维生素A）；1微克的维生素D相当于40国际单位。——编者注

(32)其经典著作《证券分析》中文简体字版已由湛庐引进，由四川人民出版社2019年出版。——编者注

(33)曾任奥巴马政府的国家经济委员负责人。

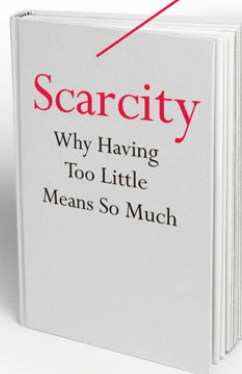
(34)也就是著名的后现代艺术作品《生者对死者无动于衷》，成品创作于1991年。

CHEERS
湛庐

稀缺

我们是如何陷入贫穷与忙碌的

[美] 塞德希尔·穆来纳森 (Sendhil Mullainathan) 著
埃尔德·沙菲尔 (Eldar Shafir) 著
魏薇 龙志勇 译



浙江教育出版社
ZHEJIANG EDUCATION PUBLISHING HOUSE

版权信息

书名：稀缺

作者：[美] 塞德希尔·穆来纳森（Sendhil Mullainathan） [美] 埃尔
德·莎菲尔（Eldar Shafir）

译者：魏薇 龙志勇

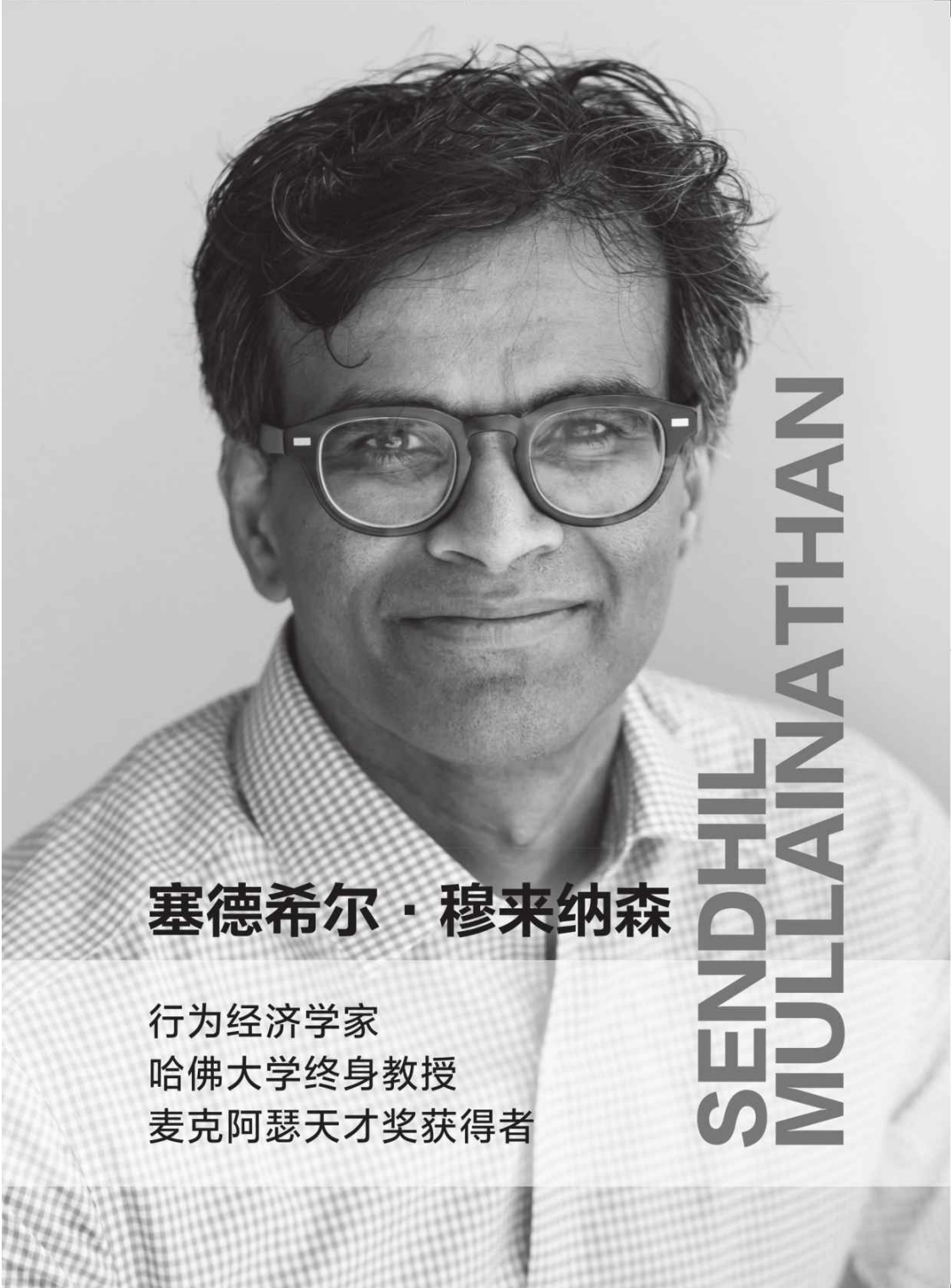
出版社：浙江教育出版社·湛庐

出版时间：2022年10月

ISBN：9787572243394

字数：312千字

版权所有·侵权必究

A black and white portrait of Sendhil Mullainathan. He is a middle-aged man with dark, wavy hair, wearing round-rimmed glasses and a light-colored, patterned button-down shirt. He is looking directly at the camera with a slight smile. The background is a plain, light gray.

塞德希尔·穆来纳森

行为经济学家
哈佛大学终身教授
麦克阿瑟天才奖获得者

SENDHIL
MULLAINATHAN



SENDHIL MULLAINATHAN

麦克阿瑟天才奖获得者

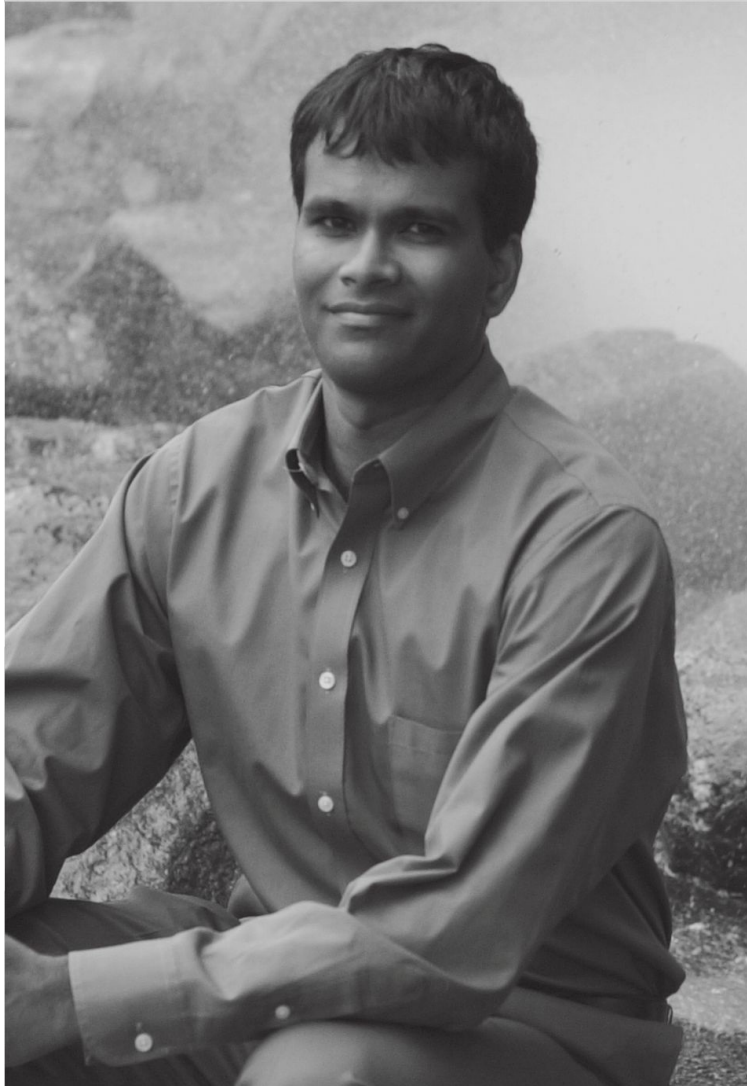
哈佛大学终身教授

塞德希尔·穆来纳森出生于印度的一个小村庄，7岁时随父母移民美国。1993年，他获得了康奈尔大学计算机科学、数学和经济学三个学士学位，1998年，获得哈佛大学经济学博士学位。

从哈佛大学取得博士学位后，穆来纳森来到麻省理工学院，开始了他的行为经济学探索之旅。在麻省理工学院，穆来纳森与经济学教授、《贫穷的本质》作者阿比吉特·班纳吉（Abhijit V. Banerjee）等人联合创立了麻省理工学院贫困行动实验室（Poverty Action Lab），通过随机控制实验，研究世界上的重大问题并取得了丰硕成果。2002年，由于在行为经济学领域的卓越贡献，穆来纳森荣获麦克阿瑟天才奖。

2004年，穆来纳森被哈佛大学聘为终身教授，成为哈佛大学行为经济学领域的重要领头人。

2008年，知名网站Edge主办了“行为经济学”大师班（Edge Master Class 2008），穆来纳森和诺贝尔经济学奖获得者丹尼尔·卡尼曼、理查德·塞勒一同担任主讲人，由此可见穆来纳森在行为经济学领域的重要地位。



《外交政策》杂志“全球**100**位思想家”之一

《连线》杂志“**50**位改变未来世界的人”之一

2018年，穆来纳森从哈佛大学来到芝加哥大学布斯商学院，担任计算机与行为科学教授。同年，因在行为经济学领域的突出贡献获得印孚瑟斯奖，这是印度表彰卓越科学和研究的最高奖之一。

目前，穆来纳森研究的主要方向是使用机器学习来理解人类的行为以及社会政策。

穆来纳森获奖无数。除了麦克阿瑟天才奖，他还获得了美国国家科学基金会、欧林基金会、斯隆基金会和塞奇基金会等组织的奖励。

穆来纳森是美国国家经济研究局的副研究员，取得过很多可喜的成绩。他还与美国消费者金融保护局合作，所取得的研究成果可圈可点。

2012年，穆来纳森被英国《连线》杂志评选为“50位改变未来世界的人”之一。

2013年，入选达沃斯世界经济论坛“全球青年领袖”（Young Global Leader）和《外交政策》杂志“全球100位思想家”之一。



行为经济学家

研究“稀缺”的世界权威

2013年，穆来纳森在《科学》杂志上发表了《贫困阻碍认知功能》（*Poverty Impedes Cognitive Function*）的论文，这是他关于“贫穷影响智力”的重要研究成果。

穆来纳森对印度的农民进行了实验研究，试图探究贫穷仅仅是“缺钱”的问题，还是影响人们智力的大问题。他细心地选择了甘蔗农作为研究对象，尽可能剔除了其他因素的影响。研究发现，同样一群农民，收割前承受“缺钱”压力的日子和收割后卖掉农产品的富足时段，在认知能力和执行控制力方面表现出了明显的差别。贫穷带来的压力，会影响他们在其他行为上的表现，他们的“心智带宽”（Mental Bandwidth）明显不够用了，心智能力也出现了全面降低。

他在《忙碌之人所犯的错误》（*The Mistake Busy People Make*）一文中给出了自己对“时间稀缺”的研究结论：人们以为只要学会“时间管理”就够了，但实际上，我们还需要管理好我们的“心智带宽”。时间可以分解，但“带宽”却无法分解。分心加剧了焦虑，进而导致无法专心处理重要任务。

穆来纳森在研究贫穷和忙碌等问题上所取得的成就，让他成为“稀缺”领域当之无愧的世界权威。

作者演讲洽谈，请联系
BD@cheerspublishing.com

更多相关资讯，请关注



湛庐文化微信订阅号

湛庐CHEERS 特别制作

A black and white portrait of a man with short dark hair, smiling slightly. He is wearing a dark jacket over a light-colored collared shirt. The background is a textured wall, possibly made of bricks or stone.

埃尔德·沙菲尔

行为经济学家
普林斯顿大学心理学教授
美国艺术与科学院院士

ELDAR SHAFIR



ELDAR SHAFIR

普林斯顿大学心理学教授

美国艺术与科学院院士

《外交政策》杂志“全球**100**位杰出思想家”之一

1988年，埃尔德·沙菲尔获得麻省理工学院认知科学博士学位。他不仅是普林斯顿大学心理学教授，还是哈佛大学社会科学定量研究院的研究员、判断与决策学会前会长、美国艺术与科学院院士。

沙菲尔的主要研究方向是行为经济学，主要探索人们在日常生活中是如何决策的以及社会、认知和情感因素对经济决策的影响。

沙菲尔是判断与决策学会希勒尔·艾因霍恩新研究者奖（Hill Einhorn New Investigator Award）获得者。2012年，沙菲尔荣获古根海姆奖，他将奖金继续用于研究稀缺心理。2013年，沙菲尔被《外交政策》杂志评选为“全球100位杰出思想家”之一。



阿莫斯·特沃斯基的学术搭档

我们都或多或少听说过卡尼曼和特沃斯基的合作与友谊，实际上，沙菲尔和特沃斯基也有多年的合作经历。在工作上，除了卡尼曼，沙菲尔也是特沃斯基的学术搭档。

早在1992年，沙菲尔和特沃斯基就在《认知心理学》杂志上联合发表了论文。后来，沙菲尔又与特沃斯基、彼得·戴蒙德一起开始研究货币幻觉（Money Illusion）。

货币幻觉是这样一种心理错觉：人们倾向于按名义价值而不是实际价值来看待货币。如果人们经常把货币的面值误认为它的购买力（实际价值），他们就不会做出理性的财务决定，这对经济理论和公共政策有至关重要的影响。

沙菲尔坚信，货币幻觉是存在的，并与戴蒙德、特沃斯基一起通过实验证明：认知偏差等许多因素确实会影响决策。此外，他们在现实世

界中也收集了许多证明货币幻觉存在的证据。

1997年，沙菲尔与戴蒙德、特沃斯基关于货币幻觉的研究论文在《经济学季刊》上正式发表。



行为经济学家

研究“稀缺”的世界权威

沙菲尔的研究重点是贫穷对决策的影响，也就是稀缺心理学。

关于贫穷，曾经存在两个思想学派。一个学派认为，穷人在行为上也是理性的，不过，过度的贫穷导致了“穷人文化”。另一个学派认为，由于错误态度和心理问题，穷人做出了糟糕的选择。

沙菲尔和穆来纳森提出了第三种观点：在计算决策的后果方面，穷人与常人并没有什么差别。跟其他人一样，穷人在决策时也会犯错。不过，穷人的错误决定会导致非常糟糕的后果。

沙菲尔的研究表明，穷人的心理跟那些因其他原因而承受压力的人的心理是类似的，比如说工作忙碌的人、缺少朋友的孤独者或者因过度节食而缺少营养的人。沙菲尔说：“我们认为，是稀缺心理让人们过度专注于自己。这导致人们除了自己专注的事情之外，对其他事情完全视而不见。人们倾向于过度借用和缺乏计划。

2008年，沙菲尔与穆来纳森等人联合创立了非营利性组织ideas42，致力于利用行为科学帮助人们解决社会问题。

沙菲尔在研究贫穷和忙碌等问题上所取得的成就，让他成为“稀缺”领域当之无愧的世界权威。

作者演讲洽谈，请联系
BD@cheerspublishing.com

更多相关资讯，请关注



湛庐文化微信订阅号

湛庐CHEERS 特别制作

赞誉

《稀缺》是行为经济学领域一部了不起的著作。经济学家穆来纳森和心理学家沙菲尔用心理学的实证研究重新解读了经济学的核心概念——“资源稀缺”，发现贫穷和忙碌并不是简单的因为缺少金钱和时间，而是一种心态和能力的匮乏。一本《稀缺》，会让忙碌的人警醒，让贫穷的人振作。物随心转，境由心生，有容乃大，命由己造！

彭凯平
清华大学社会科学学院院长

把“稀缺”界定为一个心理问题而不是资源问题，这真是一个奇妙且深刻的视角。正如本书的两位作者所说，稀缺会俘获一个人的大脑，使他过于关注稀缺的对象从而失去正常的判断力，这恰恰会导致稀缺状态进一步恶化而不是使它得到改善。因此，真正可怕的不是资源的稀缺，而是俘获我们大脑的稀缺心态。我想，这个结论对所有读者都将是一个终身受用的告诫。

叶航
浙江大学经济学教授、跨学科社会科学研究中心主任

没有什么就嚷嚷什么，越没有就越嚷嚷，这是我们在日常生活中经常会看到的现象。两位作者提出的稀缺会导致认知障碍的观点，将我们的直观感受上升到了理论高度。他们的研究给我们带来的启发是：如果你想要什么，不要等急需的时候才去要，要提前积累。

李淼
物理学家，南方科技大学物理系教授

为什么穷人会更穷、富人会更富，而拖延症患者永远没时间？针对这一现象，社会学家称之为“马太效应”，经济学家称之为“稀缺现象”，但是所有这些说法都不如心理学的解释来得更妙趣横生又直抵人心，让人时常产生“原来如此”的恍然大悟感。

周 濂
中国人民大学哲学院教授

为什么稀缺心态会让我们无法摆脱稀缺？因为处于稀缺中的人，其心智容量和认知能力会受到很大限制，从而导致其习惯于应付匮乏的行为模式。个人如何才能摆脱财务困境，组织如何才能重获时间，而社会又如何才能帮助穷人？《稀缺》一书以独特的行为经济学视角，不仅从理论上进行了开创性的分析，而且还提供了大量的心理学与社会学实验，为摆脱稀缺提供了有力支撑。

周健工
未尽研究创始人

天底下只有四种事：重要且急迫的、不重要且不急迫的、不重要但急迫的、重要但不急迫的。人们不大会在前两者上犯错误，但经常会在后两者上犯错误：被不重要但急迫的事情占去有限的注意力带宽，而忽视了重要但不急迫的事情。读完《稀缺》后，就可以避免这个错误。

王 烁
财新传媒总编

我们的肉身刚刚步入富足时代，但我们的精神还滞留在稀缺的恐惧之中。穷人思维植根于人类的基因，率先用理性击碎它的人，也将率先获得身与灵的富足。

罗振宇
得到App创始人

塞德希尔·穆来纳森是著名经济学家，埃尔德·沙菲尔是著名心理学家，而《稀缺》就是他们强强联手的佳作。他们将科学的严谨和人文的独到巧妙地融为一体。在《稀缺》一书里，直觉与理性实现了完美契合，这是一本不可多得的好书！

丹尼尔·卡尼曼
诺贝尔经济学奖得主，畅销书《思考，快与慢》《噪声》作者

这是一个必胜组合：一位行为经济学家，一位认知心理学家，他们都是各自领域著名的领导者。让他们富有创意的头脑相互碰撞，得到的就是一本充满原创思想和精彩观点、具有高度可读性的著作。一位想要在大型律师事务所升迁为合伙人的单身母亲，与一位用一半收入支付利

息的农民有何共同之处？答案就是稀缺。读一读这本书，你将能惊奇地了解到稀缺是如何影响我们每一个人的。

理查德·塞勒
诺贝尔经济学奖得主，芝加哥大学行为决策研究中心主任

《稀缺》是一本引人入胜的书，充满了新思想、精彩的故事和深入浅出的道理，很可能会改变你的生活方式。

史蒂芬·列维特
畅销书《魔鬼经济学》作者

《稀缺》一书将故事与研究成果自然而然地合为一体。它告诉我们：感觉自己拥有的比需要的少，这种情况会缩小我们的视野，扭曲我们的判断力。这本著作对个人发展和公共政策制定都有着极大的参考意义。

丹尼尔·平克
畅销书《驱动力》《全新思维》《时机管理》作者

塞德希尔·穆来纳森和埃尔德·沙菲尔告诉我们，稀缺的逻辑，无论贫富、受教育程度高低，亚洲文化、欧美文化、拉美文化还是非洲文化，都同样适用。两位作者的独到观点可以帮助我们改变个体行为，同时也能为公共政策的制定开拓出一片新视野。真是令人振奋的成就啊！

安妮·玛丽·斯劳特
普林斯顿大学荣誉退休教授，新美国基金会总裁兼**CEO**

《稀缺》一书思想深刻，原创且内容独特，令人手不释卷。本书是那些没有太多时间读书人士的必读之作。

丹尼尔·吉尔伯特
哈佛大学心理学教授，畅销书《哈佛幸福课》作者

埃尔德·沙菲尔访谈录

本文是美国心理学会面向大众的杂志《心理学追踪》对埃尔德·沙菲尔的采访，2014年2月发布，采访记者是艾米·诺维妮（Amy Novotney）。在这次采访中，沙菲尔回顾了他和穆来纳森的研究历程与成果，以及这些研究的政策意义。

问：您何时开始对稀缺感兴趣的？

答：大概在8年前，穆来纳森想要和我一起合作，研究贫穷状态下的决策问题，因为这个主题当时还没有人研究。历来有两种思考贫穷的方式：有一半人认为，穷人受环境所迫，从而是完全理性的，可以做出完美的、合乎情理的成本收益决策；另一半人致力于研究贫穷的文化，研究穷人的价值观，以及穷人不善于规划的能力。我们认为还有第三种视角，我们不把任何人看作是完全理性的，也没有理由认为穷人是严重病态的，或者在所有方面都不寻常。和其他人一样，穷人在生活中也是充满困惑的，也有自己的偏差，这样来看的话，比起在舒适状态下，我们在贫穷状态下犯错带来的后果要严重得多。

随着时间推移，收集的数据越来越多，观察到的案例也越来越多，我们发现，似乎穷人会犯更多极端的错误。这就渐渐让我们意识到，陷入贫穷状态时，人们会涌现出一种特殊的心理状态，正是这种心理状态带来糟糕的结果。

问：稀缺是怎样带来糟糕的结果的？

答：每个心理学家都理解，人的认知空间和心智带宽很有限。当你极其专注于某件事情的时候，就没多少精力投入到其他事情上去了。我们把这个现象称为“管窥”，当你把越来越多的精力投入在稀缺问题上时，在其他地方投入的精力就会越来越少，而你投入很少的事情里，有些事情甚至比稀缺问题更重要。有大量文献表明，穷人在生活中的很多事情都没做好，和有钱人相比，他们通常是不那么上心的家长，生病时也是不大会按时吃药的人，甚至贫穷的农民在种植方面也不如人。

问：请说说让你们获得这些结论的研究过程。

答：一开始，我们连续观察了印度金奈的一个大型商场里的水果、鲜花商贩。没人可以说那些卖东西的妇女是懒惰的或说是目光短浅的，她们工作非常卖力，每天都规划得井井有条，从早到晚进货，用1 000卢比的价格进鲜花或芒果，但先不用付款，转手以1 100卢比的价格卖掉后，再给供货商1 050卢比。日复一日，她们以这种很高的利息进货，每个人平均做了10年这样的买卖。但如果她们多存一点点钱，或者从供货商那儿少借一点钱，她们早就还清债务了，而且收入还会翻倍。这个买卖有它自己的一个逻辑，那就是一天就管一天的事情，没有长远考虑、调整的能力。

我们之后又完成了一系列研究后发现，处理稀缺对人的认知能力有重要影响。首先，我们和当时还是研究生的赵家英（音译）一起去了新泽西州的一个商场，请路人做测试，去度量他们的认知控制力和流体智力，流体智力是智商的一个组成要素。我们在请路人做测试时，让他们想象一个容易操作的财务场景，是用150美元去修一辆破车，或者更难一点的场景，花1 500美元在车上。我们根据路人的家庭收入分层发现，在这两个场景里，商场里的富人在认知测试方面表现得一样好，而穷一点的人在150美元的场景里，认知能力和流体智力表现得和富人一样好。但是在1 500美元的场景下，穷人获得的测试分数就下降了，仅仅是处理更严格的财务挑战就会让他们表现得不好。

很明显，在这个实验里，我们尽可能控制了实验条件，但最后依然出现了富人和穷人两种不同的表现，你可以说在健康和教育这类事情上，穷人和富人考虑的不一样。所以，我们又跑去印度研究甘蔗农，他们每年都在收获以后才有一大笔收入，他们需要保证这笔收入足以支撑到第二年收获的时候。通常甘蔗农们在收获前是穷人，在收获后是富人，所以我们对同一批甘蔗农进行了认知测试。我们是在收获之前和之后两个月进行测试的，结果发现，同一个农民，同样的教育程度和价值观，但比起收获之后，在收获之前这个农民的智商测试分数要低10分。

问：这种认知上的差异对他们的行为和决策有什么影响？

答：美国的穷人常被人诟病的一点就是“工薪日贷款”（Payday loan），这种贷款在当下可能是个好办法，但是两周之后就会产生很高的利息。所以，我们决定在普林斯顿大学的本科生里做个研究——没人会说这些本科生不够聪明吧。和阿努依·沙阿一起，我们设计了一个类

似于《家庭问答》那样的电脑游戏，并且随机把实验对象分成两组，根据要求他们回答问题的时间，分成时间富足的（每一轮问答50秒）和时间不多的（每一轮问答15秒）。还有一半参与者有权去借答题时间，但每借1秒钟，在整个游戏里的答题时间就要减少2秒。

我们发现，时间很充足的人表现得很明智，不用那么担心时间问题，只在很偶然的情况下才会借时间。但时间不多的参与者里，那些借时间的学生最后赚到的钱要少于没借时间的学生。尽管这些人都是普林斯顿大学里很聪明的学生，但我们观察到的行为和那些穷人一样。

问：对稀缺的研究有什么让您惊讶的地方？

答：最让我惊讶的是这些研究支持了一个观念，那就是在稀缺状况下表现糟糕的人，其实他们的能力和其他人并没有什么不同，只要稀缺没有占据他们的心智。关于很多行为研究，有意思的地方在于我们并没有充分的直觉意识到那些结论。比如说，对驾车时打电话的研究一直就很令人震惊，因为很多人都有一个错觉，认为我们理所当然可以在开车时打电话。但是研究结果很清楚，在车里打电话，即使没有把电话拿在手上，你的反应速度和醉驾时是一样的。同样地，大家都知道自己很忙，常被各种事情干扰，但是干扰的影响和结果比起我们意识到的更加严重。

问：稀缺状态对美国人有什么影响？

答：有很多美国人整日忧心忡忡，在财务问题上挣扎着，从而可能缺少足够的带宽。他们一遇到波澜就会满面愁容，失去了应对其他事情的认知能力。这些发现甚至可能意味着，2008年的金融危机之后，美国人可能失去了大量流体智力。人们为了生活中的一个要素而四处奔波，没有为其他事情留出空间。

问：有没有什么解决办法？

答：在你负担得起的程度之内，留给自己一点余闲。当你步履不停地忙于生计时，没有余闲，哪怕一丁点的突发事件都会让你陷入泥潭。你不知道会发生什么，但总有意料之外的事情会发生，比如开车的路上，车里的水管爆了，车就坏了，你就会拿到一张停车罚单，又或者，如果你太忙，时间太紧，你会突然接到一个意料之外的电话，或者在赶着开会的路上碰到堵车的情况。

如果你给自己创造一点余闲呢？当你在处理时间稀缺问题的时候，规划出一些余闲，比如空闲半个小时，这样如果有突发事件，你就可以用这半个小时来处理。我把这称作是“与自己相聚的时间”。当然，如果你很贫穷，这样做就不容易。但是为穷困期存钱，可以让你有能力处理意料之外的空头支票或违章停车的罚单，你可以从存款里取钱，这样日子可以继续过下去。

我们还有很多想法用来减少世界上的贫困，可以用来制定政策帮助穷人。我们不会收穷人两三百美元，让他们去加入某个福利项目，因为他们根本就没钱。或者给穷人们设计一个很复杂的项目，要求他们连着三天必须准时到达某个地方，这就是在给他们施加带宽税。所以，这就是在设计一个他们势必无法完成的任务。我们提议，政策制定者们要尽可能做到，即使人们因为没处理好稀缺问题而一时失足，但依然有机会可以重新爬起来，而不是越陷越深。

问：您最希望其他心理学家从你们的工作中获得什么？

答：随着白宫逐渐意识到行为研究的重要性，现在是时候让心理学家们真正影响政策的制定了。我之所以被任命到总统的财务咨询委员会，就是因为我们的研究成果。之前很难把心理学里关于有限认知、有限注意力和行为的根本洞见在那样一个论坛上提出来，这种事情很少发生，因为通常那些论坛里的人根本不是这样想问题的。行为科学家的影响力正在壮大，虽然发展缓慢，但远胜从前，而且对这些研究感兴趣的人越来越多。

测一测 关于“稀缺心态”，你了解多少

· 一旦时间不够用且无法逃避，我们就会开始集中精力。这是对的吗？

A. 对

B. 错

· 专注于某一事物就意味着我们会忽略其他事物吗？

A. 是

B. 否

· 人们总是很难攒下钱、很难坚持去健身房锻炼，容易高估短期利益，忽略长远利益，这体现了行为经济学中的什么观点？

A. 无意识偏差

B. 现时偏差

C. 归因偏差

D. 沉没成本谬误

扫描下面二维码查看本书更多测试题

扫码鉴别正版图书
获取您的专属福利



扫码获取全部测试题及答案，
一起来了解“稀缺心态”

引言 资源稀缺不可怕，就怕有稀缺心态

假如蚂蚁真的是一天到晚忙个不停，又怎么还有时间去四处野餐呢？[1](#)

——玛丽·德雷斯勒（Marie Dressler）

奥斯卡金像奖获奖女演员

之所以写这本书，是因为我们已经忙碌到了不得不写的程度。塞德希尔一直在向埃尔德抱怨：要做的事情太多，而能用来做事情的时间又太少。许多事情早就过了原定的截止日期，一直在延期，越来越让人担心：他不得不厚着脸皮去重新安排会议时间；收件箱里堆满了等待处理的邮件；很久没有打电话问候母亲了，他甚至都能想象到她老人家那伤心失落的样子；车子也早就过了年检期限。而且，事态还越拖越严重：转趟飞机就能去参加的那场会议，6个月前还令他很兴奋，而现在却让他兴趣全无。拖延，已经演变成了恶性循环，烦心事层出不穷。现在又多了一件，那就是去给年检过期的车子重新办理登记手续。而因为迟发了一封邮件，就导致整个项目的前进方向发生了错误，若想让它回到正轨，又要去做更多的工作。最后，日常生活中应做而未做的事情渐渐堆积成山，随时都有崩塌的危险。

颇具讽刺意味的是，对于埃尔德来说，他完全没有因时间太少而一筹莫展的经历。对于受此困扰的塞德希尔来说，这种状况虽然令自己有些不知所措，但他还是毅然决然地制订了一个“走出困境”的计划。

首先，他要制止事态的进一步恶化。以前积压下来的事情优先去做，而新找上门的事情则一概拒之门外——他要对所有的新要求说“不”。他决定要一丝不苟地完成之前的未竟之事，以避免拖延状况进一步恶化。最终，他预期自己的辛劳一定能获得成效：未竟事项的数量会减少到易于管理的程度。只有这样，他才会考虑接手新事项。而且，从今往后，他要以更加谨慎的态度去做事——在点头答应之前，一定要反复斟酌才行。虽然这样做并不容易，但却很有必要。

计划终于制订出来了，塞德希尔感觉很满意。就像伏尔泰很久之前说过的那样，“幻觉是所有乐趣之首”[2](#)。

一星期之后，塞德希尔又打来电话：“我的两位同事正在构思一部关于美国低收入人群的著作。这是一个好机会，我们应该贡献一章内容。”据埃尔德回忆，塞德希尔当时的语气里完全没有一丝讽刺的意味。

显然，“贡献一章内容”这个机会太过宝贵，不容错过。我们自然应承了下来。但也不出所料，应承下这件事情本身就是一个错误——我们写得匆匆忙忙，还拖了很久的稿。而令我们没有预料到的是，这个错误犯得很值得。因为正是这个错误为我们创造了意想不到的机会，并最终催生了本书。

现摘录两段关于那一章内容的背景说明：

肖恩在克利夫兰工作，是一位办公室经理。他正因财务状况青黄不接而苦恼：几份账单还等着他去还；信用卡又刷爆了；好不容易等到发工资，没想到一下子就花了个精光。境况正如他所说的：“月末总是没钱花。”有一天，他过于乐观地估计了账户余额，忘记还有一笔22美元的支出，结果开出去的支票就被退了回来。那时，连电话铃响都会令他紧张，因为很有可能是某位债主“提醒”他又要还钱了。缺钱的日子着实不好过。有时和朋友聚餐，他只能拿出比别人少的份子钱。朋友们都能体谅他的难处，但这样的情形实在令他汗颜。

而且从短期来看，这种处境也不会有所改变。5个月前，他刚赊账买下了一部前半年无须付款的蓝光影碟播放机，令他头疼的是，等到下个月，自己该拿什么去支付这笔额外的账单？为了还清过去欠下的债务，他已经搭上了越来越多的钱——被退回的支票要收取一笔可观的透支费；迟还的账单还要附上滞纳金。肖恩深陷于债务泥潭中，无力自拔。

肖恩就像许多与他有着相同处境的人一样，从许多地方得到了理财建议，而这些建议都大同小异：

不要再越陷越深了。不要再借贷了。节约开支，能省则省。有些支出实在难以放弃，但你还是要想办法学会忍痛割爱。尽快还清以前欠下的债务。最终，在没有新债务的情况下，你的财务状况才会出现好转。达到这一目标之后，你还要继续保持警惕，谨防重蹈覆辙。支出和借贷时，请三思而后行。不要考虑那些买不起的奢侈品。如果你不得不借贷，那么就一定要搞清楚还债所要付出的代价。

对于肖恩来说，这些建议中看不中用。抵制诱惑是一件很难做到的事情，而抵制所有诱惑则更是难上加难——他觊觎已久的皮夹克，现在终于打出了千载难逢的优惠价；女儿的生日快到了，用廉价礼物搪塞过去，似乎又不太合适。超支的方面太多了。最后，肖恩再一次掉进了债务深渊。

显而易见，塞德希尔和肖恩的行为存在着相似之处：截止日期之前

未完成的工作，就像逾期没还的账单一样；约了又约的会议（草率地答应下来，又没时间履行约定），就像被退回来的支票一样（草率地花了钱，又没钱还账）。你越是忙碌，就越需要拒绝他人的请求；你的债务越多，就越需要不出去购物。用逃避的方式解决问题，看似合理，但实践起来却颇有难度。一旦你放松了警惕，稍微一点有关时间或金钱的诱惑就会让你陷得更深。后来，肖恩的债务越欠越多，而塞德希尔应承下来的工作也越积越多。

上述两种看似无关的情况竟会如此相似，倒是令人多少有些意外。一直以来，我们都认为时间管理与金钱管理是两个截然不同的问题。而且，时间管理上的失败，与金钱管理上的失败，两者所造成的后果也各不相同：拙劣的时间管理会令人陷入难堪的境地，以致很难取得理想的工作成效；而拙劣的金钱管理则会导致更多的债务，甚至遭遇物品被没收或人被赶出出租屋的厄运。另外，两者所处的文化环境也不同：对于日理万机的专业人士来说，错过工作上的截止日期是一回事；而对于低收入人群来说，没钱还债又是另外一回事。而且，这两种情况的周遭环境不同，主人公的受教育程度不同，就连他们心中的愿望也各不相同。但尽管有诸多不同之处，两类人最终的行事方式却非常相似。

塞德希尔和肖恩有一点是相同的，他们都感受到了稀缺所造成的影响。稀缺，是“拥有”少于“需要”的感觉³：塞德希尔感觉心烦意乱，觉得自己的时间太少，而要做的事情又太多；肖恩感觉手头拮据，有永远都还不完的账单。这一点相似之处，是否可以用来解释他们的行为？我们是否可以认为，“稀缺”就是导致塞德希尔和肖恩有如此相似行事方式的根源？

找出稀缺的内在通用逻辑，有着很大的意义。稀缺是一个宏大的概念，远远超越了上面事例中的个人范畴。

举例来说，失业问题同时也是资金稀缺的问题——如果你丢了工作，那么你手头就会一下子变得拮据起来，因为收入太少，你也就没办法支付住房的按揭贷款、车子的分期付款和日常生活中的各种开销；日渐增多的社交孤立问题——“独自打保龄球”，也是社交稀缺、人际交往纽带缺乏的一种表现形式⁴；坚持节食的人就需要面对饥饿的挑战，面对热量减少或稀缺的挑战；而世界贫困问题，也就是世界各地大量人口每天只能靠一两美元勉强度日的问题，则是另一种资金稀缺的表现形式。不过，贫困与失业不同，后者导致的囊中羞涩事发突然，可以通过寻找新工作来解决，而前者则意味着持久的拮据与艰难。

稀缺，不仅仅是塞德希尔和肖恩所面临问题的共同点。在现今诸多的社会问题中，我们都能看到稀缺的存在。这些社会问题发生在不同的

文化背景、经济环境和政治体制中，但都逃不开稀缺的范畴。那么，是否存在一个关于稀缺的通用逻辑，在各种背景与环境下都适用呢？

我们必须找到这个问题的答案，因为我们已经忙碌到了不得不这样做的地步。

稀缺俘获大脑

对稀缺的兴趣，让我们了解到了半个多世纪以前的一个研究项目。当时的研究者并不认为他们是在研究稀缺问题，但在我们看来，他们研究的问题正是稀缺的一种极端表现形式——饥饿。当时正值第二次世界大战的尾声，盟军遇到了一个麻烦。在进入德军占领区后，盟军发现生活在那里的大多数人都处于饥饿的边缘。不过，真正的麻烦不在于食物的短缺，因为美军和英军拥有充足的食物供给，完全可以养活那里的战犯和平民，而是在于要为他们提供食物颇具技术含量：面对这些长期处在饥饿边缘的人们，应该怎样开始为他们提供食物？应不应该每次都给他们一顿饱饭？是让他们想吃多少就吃多少，还是应该开始时少给一点，然后慢慢地增加分量？若想挽救他们，什么做法才是最安全的？[5](#)

面对这些问题，专家们当时并没有给出妥善的解决办法。为了找到答案，明尼苏达大学的一支研究团队做了一项实验。为了搞清楚应该如何为人们提供食物，首先要做的就是让他们感到饥饿。[6](#)

稀缺实验室

SCARCITY

实验招募了一批身体健康的男性志愿者。在受控环境下，研究人员们为志愿者们提供的食物一直在减量。这些食物的热量仅够维持生命，不过还不至于对志愿者们的身體造成永久性伤害。这一阶段持续了几个月时间。随后，真正的实验开始了：研究人员开始观察志愿者们的身体会对不同的食物供给量产生怎样的反应。作为这一实验的研究对象，他们实在不轻松，但时值第二次世界大战，没有去前线奋勇杀敌的人们，即使对这样的实验颇有微词，也只能安守本分，敢怒而不敢言。

参与研究的36位志愿者生活在集体宿舍中，接受严格的监控，每个行为都会被研究人员观察和记录下来。虽然研究人员最关心的是食物供

给量的问题，但他们也对饥饿造成的影响进行了观测。

研究人员发现，长期挨饿的人，从体型上就能看出来：实验过程中，研究对象的臀部脂肪会大大减少，以至于坐着都会感觉到疼痛，他们不得不垫上坐垫。另外，实际的体重减轻情况，则因水肿而难以进行估算，有人还曾因为饥饿而在体内额外累积了多达6千克的液体。研究对象新陈代谢的速度也减缓了40%。他们开始感觉到有气无力，缺乏耐性。一位实验对象说：“在淋浴头下洗头时，我感觉到手臂瘫软。仅是洗头这一件事情，就令我的双臂疲劳到了极点。”

饥饿不仅让志愿者们身体变得虚弱，而且还让他们的思想发生了变化。沙曼·萝赛（Sharman Apt Russell）在其著作《饥饿》（*Hunger*）一书中讲述了午餐时的一个场景：[7](#)

如果服务速度很慢，排队等候取餐的人就会急不可耐。他们对食物的占有欲很强。有些人甚至会用胳膊环住餐盘，以保护盘中的食物。绝大多数情况下，他们都十分安静，专心致志地进餐……本来不爱吃某些食物的人，现在也不挑剔了。他们会将所有食物都吃个精光，之后还会将盘子舔一遍。

挨饿的人，想必会有上述表现。而与此同时，这些人也出现了一些预料之外的精神变化：

人们开始迷上菜谱和当地餐厅的菜单。有些人甚至会用好几个小时的时间，对两份报纸上水果蔬菜的价格进行对比。有些人开始产生进入饮食行业的打算：他们梦想着去开餐厅，从此开启一番新事业……他们丧失了攻克学术挑战的斗志，而对菜谱产生了浓厚的兴趣……而在看电影时，只有与食物有关的情节才会引起他们的注意。

人们将注意力全部集中在了食物上。当然，如果你也整天忍饥挨饿的话，那么找食物必然会第一要务。但这些人关注的方面却远远超越了自身的实际利益——开餐厅、对比食物的价格、研究菜谱，这些想法都不能缓解饥饿。而所有这些关于食物的执着，只能进一步加剧饥饿带来的痛苦。他们对此毫无招架之力，但又别无选择。参与了明尼苏达大学研究的一位志愿者在回忆起当时的情形时，对因大脑中一直充斥着有关食物的想法而给自己带来的苦恼记忆犹新：

这辈子没有几件事情像这次实验那样，让我恨不得赶快结束。其实身体上的不适并没有多么严重，真正令人无法忍受的是，在实验过程中，食物成了人生中最重要东西，成了人生的中心和唯一。而如果食物成了唯一，那么生活就会变得十分无趣。看电影时，男欢女爱的场面不会令你产生多大兴趣，而一旦那些男男女女开始吃东西，你就会眼前一亮。

饥饿的实验对象并不是有意选择忽略浪漫情节而关注进餐场景的，

也不是有意选择将食物置于大脑中最重要的位置的，出现这一问题的原因是，饥饿俘获了他们的思想和注意力。在明尼苏达大学进行的研究中，这些行为只是研究记录中的一段脚注，并不是研究人员的关注点。而对于我们来说，这些行为恰恰反映了稀缺是如何改变我们的。

稀缺会俘获大脑。就像饥饿的研究对象日思夜想着食物一样，当我们经历任何一种形式的稀缺时，都会对稀缺的事物全神贯注。我们的思想会自动而强有力地转向未得到满足的需要：对于饥饿的人来说，他们需要食物；对于忙碌的人来说，他们需要完成某项工作的时间；对于缺钱的人来说，他们需要想办法支付每个月的房租；而对于孤独的人来说，他们需要他人的陪伴。稀缺造成的后果不仅仅是因为我们会因拥有的太少而感到不悦，而是因为它会改变我们的思维方式，强行侵入我们的思想之中。

稀缺俘获大脑

(Scarcity Captures the Mind)

稀缺造成的后果不仅仅是因为我们会因拥有的太少而感到不悦，而是因为它会改变我们的思维方式，强行侵入我们的思想之中。

仅从一项研究中就得出这么多结论，未免武断。饥饿是稀缺的极端形式，它既包括稀缺，但也包括其他一些生理变化。看上去，该项研究仅有36位实验对象；而我们引用的证据，大部分又都是饿着肚子的人所发的牢骚，而非科学的数据。但是，其他许多更为严格的研究也得出了相同的结论。不仅如此，这些研究还为我们打开了一扇窗，让我们看到了稀缺究竟是如何俘获大脑的。

稀缺实验室

SCARCITY

在一项研究中，实验人员请研究对象在午餐时分（之前禁食了3~4个小时）赶到实验室。其中一半研究对象被允许去吃午餐，而另一半不能去。这样，一半人大快朵颐，另一半人则继续忍饥挨饿。实验对象在研究中所要执行的任务非常简单，就是盯着屏幕看

——屏幕上会有一两个字一闪而过，实验对象的任务就是认出这些字。比如，当屏幕上闪过“住”这个字时，实验对象就要指出刚刚闪过的字是“住”还是“往”。这看似轻而易举的任务，如果不是因为屏幕上显示出来的字消失得太快，就真的是小菜一碟了。一闪而过的速度真的很快：字体在屏幕上停留的时间仅有33毫秒，也就是一秒钟的1/30。[8](#)

你也许会认为，那些饿肚子的实验对象的成绩会比较差，因为他们已经精疲力竭了，而饥饿感会令他们无法集中注意力。但实验结果却让人大吃一惊：他们的成绩与那些吃饱喝足的实验对象一样优秀。有趣的是，一个特例除外。当与食物有关的词汇闪现出来时，饿肚子的实验对象会比饱餐后的实验对象认得更快、更好——他们能更加准确地认出“蛋糕”这个词。这类实验旨在帮助我们了解一个人大脑中的关注点——当某个想法整天在我们的脑海中盘旋时，我们就能更加迅速地看到与此相关的词汇。因此，当饿肚子的人能更快地认出“蛋糕”时，我们便能一眼看出他们此刻心中所想的的就是食物。在此，我们并不是依赖于翻阅菜谱或打算开餐馆等奇怪举动来推断他们的执念，他们作出回应的速度和准确性就能直接告诉我们：稀缺已经俘获了这些饿汉们的大脑。

稀缺对人类大脑产生的影响，存在于潜意识之中。之所以将词语的闪现时间定在以毫秒计的一瞬，是为了观察人们的快速反应能力，而这个速度需要快到可以超越人类意识的控制范畴。[9](#)如今，我们对人类大脑已经有了足够多的了解，知道了时间长度的意义：复杂的高阶运算，需要300毫秒以上的时间；更为快速的回应，则依赖于更为自动自发的潜意识处理过程。因此，当饿肚子的人能以更快的速度识别出“蛋糕”这个词语时，并不是因为他们“选择”将注意力集中在这个词语上——识别的速度太快，根本来不及进行选择。这就是为什么我们用“俘获”这个词来描述稀缺会令大脑产生专注现象的原因。

这一现象并不局限于饥饿。一项研究发现，当实验对象口渴时，他们会更加快速地识别出“水”这个字（速度以几十毫秒计）。[10](#)在所有这些情况中，稀缺都在潜意识的层面发挥着作用。无论大脑的主人是否愿意，稀缺都会牢牢地俘获他的注意力。

口渴与饥饿都是生理需求。而其他与人类本能关系不大的稀缺形式，同样也会俘获我们的大脑。

稀缺实验室

SCARCITY

一项研究要求儿童凭记忆，通过调整实物模型的方式，来估计硬币的大小，其中包括从1分到5角的各种硬币。[11](#)在穷人家孩子眼中，硬币“看上去”很大，他们会明显高估硬币的尺寸。美元中价值最大的硬币，也就是两角五分和五角的硬币，其尺寸被扭曲得最为严重。正如食物会俘获饥饿之人的大脑一样，硬币也会俘获贫困儿童的注意力。这一强化了注意力会使这些硬币“看上去”更大。之所以看上去更大，不乏一种可能性，那就是穷人家的孩子并不善于记住物体的大小。于是，研究人员们干脆直接让孩子们估计摆在他们眼前的硬币的大小，这样就能更加直观、更加简单地得出结论了。结果，这些贫困儿童估计出的大小数值误差比之前更大。[12](#)

可见，比起记忆中的抽象硬币，真正的硬币唤起了更强的注意力。（在没有硬币的环境下，孩子们在估计相似大小硬纸片的尺寸时，准确度很高。）[13](#)

对注意力的俘获，会改变人的体验。在诸如车祸和抢劫等短暂而注意力高度集中的事件中，注意力的增强会引发“主观时间延展”现象——人们会感到这类事件的持续时间更长，因为在短时间内人们需要处理大量的信息。[14](#)同样，稀缺对注意力的俘获，不仅会影响我们的所见和所见的速度，而且也会影响我们对周遭世界的认识。一项针对孤独者的研究中，实验人员在实验对象面前用一秒钟展示了带有表情的人脸图片，然后要求实验对象描述出表情的含义。[15](#)这些面孔表达的是愤怒、恐惧、快乐，还是悲伤？这一简单的任务可以对一项关键的社交技能进行衡量——理解他人感受的能力。有意思的是，孤独者的成绩更好。你也许会认为，孤独者无法回答这类问题，因为他们之所以会孤独，就是由不善于社交或缺乏社交经验所导致的。但当我们把稀缺心态考虑在内时，就能理解这一现象了：因为孤独者会去关注自身的稀缺状况，关注为数不多的亲人朋友；他们会对他人表情中所流露出来的情绪非常敏感。[16](#)

这也意味着，孤独者能记忆更多的社交信息。一项研究要求实验对象阅读某人的日记，并在心中形成对此人的印象；之后，再要求实验对象去回忆日记中的细节。实验结果表明，孤独者的成绩与社交达人的一样好。不过有一个特例除外：孤独者非常善于记住与他人互动等社交内容的细节。[17](#)

此项研究的设计者以一则轶事很好地诠释了孤独对人们注意力的改变：一位名叫布莱德利·史密斯（Bradley Smith）的人，不幸陷入了一段痛苦的婚姻，他又没有什么亲密的朋友。离婚之后，他觉得自己对周遭世界的看法发生了改变：

突然之间，布莱德利总会不由自主地去关注人们之间的关系，其中有夫妻，也有家庭，而这种关注往往是精细而又令人痛苦的。也许在人生中的某一时刻，我们也曾遭遇像布莱德利一样的困境。也许像布莱德利那样，失恋的经历会让我们去关注公园中牵手漫步的亲密情侣。或者，在进入新学校、入职新岗位的头一天，我们被陌生人包围时，每一个投向自己的微笑、蹙眉或凝望，都会让自己感觉充满了特殊意味。[18](#)

当然，你也可以将布莱德利的行为视为社交饥渴。但是，他手中自有一本被翻阅了千遍的“菜谱”。

经济学研究稀缺，但不触及心态

当我们向一位经济学家同事谈到我们正在研究稀缺这个问题时，他说：“有一门学科就是关于稀缺的，你可能也曾听说过，它就是经济学。”他当然没有错。经济学是一门研究如何利用有限方式去实现无限愿望的学科，一门研究人与社会应该如何对实物稀缺进行管理的学科。如果你花钱买了一件新外套，就没有那么多钱外出用餐；如果政府在前列腺癌的实验性手术研究中投入过多资金，就没有那么多资金用于公路安全领域。许多人都会忽略“权衡”的存在（我们的理论有助于对这种忽略进行解释）。还有一些经济学观点认为，价格是对实物稀缺的反映，有时这种反映会以令人意想不到的方式表现出来。

19世纪时，欧洲的古生物学家就为此付出了代价——为了寻找稀有的恐龙骨骼化石，古生物学家们为可以提供化石碎片的村民发放金钱奖励。结果，化石碎片的供给量大幅上升：村民们一找到化石，就会将其敲成碎片，以增加可以销售的碎片数量。而这种结果并不是古生物学家们的初衷。[19](#)

我们对稀缺的研究方法是不同的。在经济学领域，稀缺无处不在。我们每一个人所拥有的金钱数量都是有限的，就算是最富有的人也买不起世界上的所有东西。但我们认为，虽然稀缺是无处不在的，但对稀缺的感觉却并非如此。

稀缺实验室

SCARCITY

试想，某一天上班时，你的日程安排中只有零零散散的几个会议，需要去做的几件事情也完全在你的掌控之中。于是，你利用空闲时间慢悠悠地吃了顿午餐，从容地开了个会，还给同事打了个电话聊了聊最近的新闻。现在请再想象一下，另一天上班时，你的日程表里排满了一大堆会议，唯一一点自由时间，你还必须投入到过了截止日期还未完成的项目之中。

上述两种情形，从实质上讲，都是一种时间稀缺——用来工作的小时数不变，而在工作时间里需要去做的事情却有很多。在其中一种情况下，你会十分敏锐地觉察到时间的有限性；而在另一种情况下，时间这个概念似乎又与你无关，你根本感觉不到。可见，对稀缺的感觉，与稀缺的实质存在是不同的。

稀缺的感觉来自哪里？实质性的限制起到了一定的作用——有限的存款、欠下的债务、等待我们去完成的工作等。而同时，我们对事物重要性的主观理解也发挥了作用——我们需要完成多少工作？购买某件东西的意愿究竟有多急切？诸如此类的愿望，是由我们所处的文化环境、成长环境甚至遗传基因所决定的：我们可能会非常渴望得到某样东西，也许是出于生理原因，也许是因为看到了邻居的炫耀；我们对寒冷的感知，不仅取决于绝对温度，而且也与我们自身的新陈代谢有关。同样，对稀缺的感觉，取决于可用的资源和我们自身的体验。²⁰包括社会学家、心理学家、人类学家、神经学家、精神病学家和营销专家在内的各类学者，都曾尝试对这类体验进行解析。我们避开对这一领域的讨论，将偏好置于一旁，转而去关注稀缺的内在逻辑及其带来的后果。当我们感觉自己拥有的太少时，大脑会发生什么样的变化？这种变化又会如何塑造我们的选择和行为？

粗略来看，包括经济学在内的绝大多数学科，在面对这一问题时都会给出同样的答案。拥有的比需要的少，结果很简单：人们会变得不幸福。²¹我们越是贫穷，就越是买不到什么好东西，大到昂贵的学区房，小到给食物提味的盐和糖；我们越是忙碌，就越享受不到休闲时光，无论是看电视还是与家人欢聚一堂；我们的体重越是超标，就越不敢品尝美食，如此种种。拥有的比需要的少，的确会令人不快，而且会产生其他不良影响，比如我们的健康、安全和受教育水平都会受到影响。稀缺，会导致不满，甚至抗争。²²

面对这样的事实，我们感觉到其中缺少了一些关键因素。稀缺不仅仅是实质上的约束，也是一种心态。当稀缺俘获了我们的注意力时，就

会改变我们的思维方式，无论它的存在是以毫秒计，还是以小时、天或星期计。对稀缺的感觉会萦绕在我们的脑海中挥之不去，从而影响我们所关注的事物，影响我们对选择的权衡，影响我们的思维方式，并最终影响我们的决策和行为。当我们在稀缺的影响下做事情时，会以不同方式去描述、管理并处理问题。曾有学者对特定的稀缺情况所引发的心态问题进行了研究：节食会对情绪造成什么样的影响？某种文化语境会对当地贫困人群的民意造成什么样的影响？[23](#)而我们的立场将会更具有普遍性：稀缺，无论以什么样的形式表现出来，都会创造出类似的心态。对这种心态的理解，有助于解释许多行为和稀缺所导致的后果。

稀缺心态 (Scarcity Mindset)

稀缺是一种心态。当它俘获我们的注意力时，就会改变我们的思维方式，影响我们的决策和行为方式。

当稀缺俘获大脑时，我们的注意力就会变得更加集中，做事就会更有效率。在生活中的许多情况下，保持注意力高度集中都颇具挑战：工作上的拖延，是因为我们总会被杂七杂八的事情所干扰；从超市买回了标价过高的东西，是因为我们心不在焉。而当我们心神专注时，就不会那么容易犯下粗心的错误了。这一点很好理解：稀缺之所以会俘获我们，是因为稀缺很重要，值得我们投以关注。

但究竟应该何时锁定心神，我们却无法完全自主选择。手头上那件尚未完成的工作，不仅让我们在工作时间内全情投入，而且也让我们在家辅导孩子做功课的时候心不在焉。这种不由自主的俘获，虽然会使我们集中注意力，但同时也是伴随一生的负担。因为我们总是将关注点放在稀缺上，思想总是围着稀缺打转，所以就没有那么多精力投入人生中的其他事物。这不仅仅是一个比喻。关于心智的容量，也就是我们所谓的“带宽”，是可以直接进行测量的。我们可以对影响自己处理信息、做出决策的“流体智力”进行测评，也可以对影响我们行为冲动与否的“执行控制力”进行测评。我们发现，稀缺会降低所有这些带宽的容量，致使我们缺乏洞察力和前瞻性，还会减弱我们的控制力。其影响力非常大。举例来说，贫穷状态会比彻夜不眠对人的认知能力产生的影响更大。产生这个问题的原因并不是穷人的带宽不及富人，而是因为贫困的经历会降低任何一个人的带宽。

带宽 (Bandwidth)

即心智的容量，包括认知能力和执行控制力。稀缺会降低所有这些带宽的容量，致使人们缺乏洞察力和前瞻性，还会减弱人们的执行控制力。

每当想到“贫穷”这个词语时，我们就会自然而然地联想到缺钱；每当想到日理万机的人或孤苦无依的人时，我们就会联想到没时间与没朋友。我们的研究显示，各种类型的稀缺都会导致带宽变窄。由于带宽会对各个方面的行为产生影响，因此它的变窄也会引发一系列负面的后果。在塞德希尔和肖恩的例子中，我们看到了这种影响：难以坚持按计划行事，无法抗拒新皮衣的诱惑，无法拒绝新项目的吸引，忘性大（忘记验车、打电话、还账单），脑子不好使（错误地估计了银行账户中的余额，发错了邀请函），这些都是因为带宽变窄才发生的。在稀缺所引发的后果中，有一项尤为重要：稀缺会进一步延续并加剧稀缺。塞德希尔和肖恩落入稀缺陷阱而无力逃脱，并非巧合。稀缺会创造出更大的陷阱。

贫穷之人会一直穷困潦倒，忙碌之人会永远日理万机，孤独之人会终日形单影只，而节食之人也总是会以失败而告终。关于这些现象的原因，我们找到了一个与众不同的解释。现有的一些理论在回应这些问题时，会从文化、个性、偏好或制度因素中寻找答案：负债之人对金钱和债务怀着什么样的态度？忙碌得不可开交的人，他们有着什么样的工作习惯？肥胖人士的饮食选择会受到什么样的文化和偏好的影响？而我们的研究给出了更加根本的结论：许多这类问题都能从稀缺心态的角度进行解释。当然，这并不是说文化、经济影响力和个性不重要。虽然这些影响因素的确会发挥重要作用，但稀缺存在着自身的内在逻辑，而这一逻辑会先于其他影响力发挥作用。

将这些稀缺陷阱综合在一起进行分析，并不意味着所有稀缺表现形式都拥有同样的重量级。在某种情境下，稀缺心态可能会产生比在另一种情境下更大的作用。比如，我们可以利用人类的记忆结构来理解问题，小到为什么会忘记带钥匙，大到目击证人的可信度和阿尔茨海默病的发病率等。[24](#)同样，虽然稀缺在不同领域有着类似的逻辑，但其影响却大相径庭。在对贫困问题进行研究时，我们尤其深刻地体会到了这一

点。贫穷的境遇可能会非常极端，其周遭环境会极其恶劣，让人难以忍受。而带宽负担对于穷人所造成的影响，很可能会大于对工作忙碌之人或节食者所造成的影响。出于这一原因，我们随后会重点关注贫困人群。

从某种角度来看，本书所要表达的观点十分简单。稀缺会俘获我们的注意力，并带来一点点好处：我们能够在应对迫切需求时，做得更好。但从长远的角度来看，我们的损失更大：我们会忽视其他需要关注的事项，在生活的其他方面变得不那么有成效。这一观点不仅有助于对稀缺塑造行为的现象进行解释，同时也可以让我们从中得出一些令人意想不到的结论，并对我们应如何对自身的稀缺状态进行管理这个问题提供新答案。

正在发展中的稀缺科学

稀缺是一门“正在形成中的科学”。这门科学旨在揭示稀缺的心理学基础，并利用这一知识去解释各式各样的社会现象和行为方式。本书的绝大部分内容都以原创研究为依托。这些研究开展于各种不同的环境，有大学实验室、购物中心，还有火车站等。从新泽西州的施粥厨房到印度的甘蔗田，我们的足迹遍布世界各地。同时，我们也从这一新型假设的角度出发，回顾了前人的一些研究成果（譬如前面讲过的饥饿研究），并以前人始料未及的方式去加以理解。我们利用这些成果，构建起了一个个案例，以作为新观点的论据支持。

在一个全新领域进行研究的优势之一就是，既可以与专家展开讨论，也可以将看法展示给非专业人士。由于我们的观点建立在一系列科学的基础之上——从认知科学到发展经济学等，因此很少有人是横贯全部这些领域的通才。对于大多数人来说，本书的内容中至少有一部分是新鲜的。为了适应广大读者的需要，我们已经努力将整本著作写得深入浅出。同时，书中也包罗了大量的小故事、小段子。虽然这些故事和段子不能作为严谨事实证据的替代品，但却可以让一些概念变得更加直观，令一些思想变得更加生动。最终，我们的观点是否能站得住脚，是否有足够的说服力，还得依赖于我们所列举的那些事实论据。

本书并非一锤定音之作，而是用一种全新的观点去审视一个古老的问题——它值得我们用严肃的态度去细细品味。每当出现一种新的思维方式时，就会从中衍生出新的含义、解析出新的影响力、发展出新的成

果。还有太多事情等待我们去完成，而从这个角度来看，这本书就是一份邀请，请你坐到影院的首排，来欣赏这部发现之旅的每一个精彩瞬间。



SCARCITY

WHY HAVING TOO
LITTLE MEANS
SO MUCH

第1章 专注的“得”与管窥的“失”

当稀缺俘获大脑时，人们的注意力会集中在紧急的事情上，并将其他事物排除在外。这种专注会让人们从稀缺中获益，让人们获得“专注红利”。由于“目标抑制”的作用，人们在专注于某项重要事物的同时，就不容易想到其他重要事物。因此，专注也会导致管窥，让人们的视野变窄，从而付出沉重代价。

霍布斯：你构思好要写的小说了吗？

卡尔文：创造力可不像水龙头，说开就开。你得找到适当的情绪才行。

霍布斯：什么样的情绪是适当的？

卡尔文：紧要关头的恐慌。[1](#)

——比尔·沃特森（Bill Waterson）
《卡尔文与霍布斯虎》作者

不久前的一个晚上，我们去了一家名叫“土之糖”（Dirt Candy）的素食餐厅。这家餐厅之所以起这样一个名字，是因为老板兼主厨阿曼达·科恩（Amanda Cohen）认为，蔬菜是生于土地的“糖果”。餐厅的招牌菜是西兰花脆皮豆腐配橘子酱，很多顾客对它赞不绝口。[2](#)我们也点了这道著名的菜肴，发现果然名不虚传，十分美味。它无疑是全桌最受欢迎的一道菜。

我们来得还真是时候。因为第二天我们就得知，阿曼达·科恩即将在一档收视率很高的料理综艺节目《铁人料理》（*Iron Chef*）中露面。《铁人料理》的比赛规则是，几位厨师要在短时间内完成3道菜，然后决出胜负。节目一开始，厨师们会在现场得知每道菜必须使用的一种神秘食材，然后用几个小时的时间进行菜品的设计和烹饪。这档料理大赛十分受欢迎，是烹饪界新人、美食鉴赏家以及美食爱好者们必看的节目。

在看这期节目时，我们都觉得科恩真是太幸运了——她得到的神秘食材是西兰花，所以她烹饪了自己的拿手好菜，也就是我们前一天刚刚品尝过的菜肴。评委们都对这道菜的口味赞赏有加。但事实上，科恩其实并非像我们所想的那样幸运——西兰花这种神秘食材，并没有让她显示出自己得心应手的烹饪技艺。恰恰相反，这期节目是在一年前录制的。就像她所说的那样：“现在菜单上的西兰花脆皮豆腐，就是在录制《铁人料理》时创作出来的。”就在那天晚上，科恩创造出了这道令她名声大振的拿手好菜。如果我们可以将这种现场灵感称为“幸运”，那么这种幸运就显得更加“卓尔不凡”。一位专家级人物花费了数年时间潜心积累了精湛的厨艺，而其最引以为傲的作品，却是在紧迫的时间和高度的压力下创造出来的。[3](#)

当然，这道菜并非是凭空创造的。诸如此类的创意大爆发，是建立在成年累月的经验和辛劳的基础上的。[4](#)紧迫的时间会令人精神集中，迫使我们之前积累的努力浓缩成即时产出的成果。设想你正在准备几天后就要在会议上做的报告。会议之前的这几天，你埋头苦干，但许多

想法却总是摇摆不定。事实上，报告的大方向你可能已经确定好了，但如何将诸多思想组织起来却令你十分苦恼。眼看开会的日子一天天逼近，你再也没有磨蹭的时间了。稀缺，会迫使人做出选择。所有抽象的事物都会变得具体起来。如果没有最后的这一点推动力，即使你的大脑中已经充满了创意，却拿不出最终的计划。参加《铁人料理》节目之时，科恩已经拥有了属于自己的“秘密调料”——潜心钻研了数月甚至数年的烹饪技艺。稀缺并不能凭空创造出这些技艺，而是给了她一股推动力，让她将这些技艺整合起来，做出了一道人人称赞的美味佳肴。

人们常常会将稀缺与其最可怕的后果联系在一起——有人变得穷困潦倒，还欠了一屁股债；有人累死累活却永远赶不上任务进度。我们就是由此构思出这本书的。不过，科恩的经历还体现了稀缺的另一面，而这一面并不易被人察觉——稀缺能令人们收获更大的成效。想必所有人都曾有过这样的经历，在手头资源不足、感觉紧迫之时会取得突出的成绩。在科恩的例子中，正是因为她充分地意识到时间的紧迫性，才能心神专注，将多年的经验融会贯通，最终烹饪出一道美味佳肴。所以，我们认为，当稀缺俘获大脑时，人们的注意力会集中在如何以最有成效的方式去利用自身资源上。虽然这种现象存在着负面影响，但也意味着人们可以从稀缺中获益。本章会先讲述稀缺带来的利益，之后再讨论我们为稀缺所付出的代价，其中预示着稀缺会以失败而告终的命运。

资源稀缺换来了专注与回报

很多人都讨厌开会。著名组织行为学学者康妮·格西克（Connie Gersick）对此进行过深入的研究。为了搞明白会议的进程以及进程中工作与交流的变化模式，她进行了多次详细的定性研究。她对各种类型的会议均有研究——学生之间的会议、管理者之间的会议、为了在几种选择中做出权衡从而进行决策的会议与以头脑风暴形式设计产品推销方案的会议等。这些会议五花八门，但从某种角度来讲，它们又是异曲同工的：所有会议一开始都是没有重点的，而讨论的内容也非常抽象，离题甚远——大家的谈话总是漫无边际，经常会脱离会议的主题，原本简单的要点总是要费尽周折才能说明白；每个人都有自己的见解，却没有谁能给出解决方法；时间都花费在了与主题无关的细节上。⁵

在会议开到一半时，情况发生了转变。用格西克的话说就是，出现了“中途修正”。与会者开始意识到时间快不够了，于是变得认真起来。就像格西克所说的那样：“会议开到一半，当人们开始担心会议结束时

间和会议进度时，就是‘进程大幅跃升’的起点。（此时）人们会开始变得齐心协力起来，完成工作的意愿和能量会突然增强。”就这样，不同意见纷纷被铲除，关键的细节问题成了焦点话题，其他一切靠边站。会议的后半段总是会出现很多实质性的进展。

中途修正反映了稀缺俘获大脑所引发的一种结果——一旦时间不够用的现实变得无法逃避，我们就会开始集中精力。

就算我们是独自一人做事，也会出现这种行事规律。试想你正在撰写一部著作，正在写作的章节需要在几周之内完成。于是，你坐下来准备埋头苦干。刚写了几句话之后，你突然想起来还有一封邮件没回。打开收件箱之后，你又发现收到了其他几封需要回复的邮件。一眨眼的工夫，半个小时过去了，你还在回邮件。你并没有忘记自己的任务——写作，于是又回头看了看刚才写出来的那几句平淡无奇的话。而就在你继续“写作”这一任务的时候，却发现自己脑子里想着的竟是以下这些事情——午餐要不要吃比萨？最近一次检查胆固醇指标是什么时候？有没有将自己的人寿保险更新到新地址？从一个想法神游到另一个基本不搭边的想法已经耗费多长时间了？所幸马上就到午餐时间了，于是你决定提早收工。你与许久不见的朋友共进午餐，然后又慢悠悠地品尝了一杯咖啡——反正还有几周时间可以用来完成这章内容呢。就这样，一天时间被一点一点地耗尽了，虽然你写出了一些东西，但远比所希望的要少得多。

现在，请再想象一下一个月之后的情形：还有几天时间就得完成这一章节的写作了，此时，你坐在办公桌旁准备干活时的心态就多了几分紧迫。当你想起同事发过来的邮件时，毅然决然地继续写作，没有半点儿分心；也有可能你太过专注于写作大业，根本没有想起来还有邮件这回事；你的思想没有漂移到午餐、胆固醇检查或人寿保险的条款上；与朋友共进午餐时（假设你没有将午餐推迟），你也没有为了喝杯咖啡而磨磨蹭蹭；就算身处餐厅，你满心想着的也是尚未完成的章节和步步紧逼的截稿日期。一天结束之时，专注为你带来了回报：你终于写出了一大段内容。

稀缺实验室

SCARCITY

心理学家通过控制实验的方法，对设定截止日期的益处进行了

研究。在一项研究中，实验人员雇用了一组本科生来校对3篇文章，并给了他们很长时间——只要在3周内完成任务就可以了。学生们的收入以校对出的错误数量以及是否按时完成任务为依据。他们要在第3周结束之时上交所有文章。针对另一组实验对象，研究人员定下了更为严格的截止日期，从而创造出了更大的稀缺：总时间依然是3周，但这组学生需要每周上交一篇校对完成的文章。结果如何呢？就像之前预想的情形一样，截止日期更为紧迫的一组更有成效。虽然这组学生需要去应对的截止日期更多，但迟交的情况更少，找到的错误数量却更多，因此赚到的钱自然也更多。[6](#)

截止日期的作用不仅仅在于可以提高工作成效。对于大学毕业生来说，大二的第二学期就是一个截止日期。能尽情地享受大学生活的日子已经屈指可数了。心理学家杰米·库尔茨（Jaime Kurtz）曾针对大四学生如何面对毕业这个“截止日期”的问题进行了一项研究：研究从毕业前6周开始。6周时间，距离大学生活正式结束还不是那么紧迫，并不能引起学生们充分的重视；而6周时间也不算长，足够令他们意识到毕业之日即将来临。[7](#)

库尔茨对其中一半学生强调毕业迫在眉睫（剩余时间以小时计），对另一半学生强调时间还很充裕（6周相当于一年时间中很可观的一部分）。对于稀缺的理解和意识产生的变化，也改变了学生们管理时间的方式：当他们感觉时间所剩无几时，会尽量充分地利用余下的每一天；他们会花更多的时间参加各种活动，沉浸在大学生活最后的一点点美好时光中。这样一来，他们觉得自己更加开心了，也许是因为他们要抓紧时间享受大学生活的丰富多彩。

时间稀缺所带来的影响体现在各种各样的领域中。

稀缺实验室

SCARCITY

在一项大规模的营销实验中，实验人员为一批客户邮寄了一张写有截止日期的礼券，而为另一批客户邮寄了不会过期的礼券。虽然不会过期的礼券一直都有效，但其使用率很低。可见，如果没有时间稀缺的影响，礼券就不会吸引人们的注意力，甚至有可能被丢弃在一旁，被人渐渐遗忘。[8](#)在另一个领域中，组织研究者发现，在销售周期的最后几周（或几天）中，销售人员的工作努力程

度达到了顶点。[9](#)而我们进行的一项研究发现，距离发薪日越近，数据录入员工作得就会越卖力。[10](#)

英国著名记者马克斯·黑斯廷斯（Max Hastings）曾在了一本关于丘吉尔的著作中这样写道，“英国人头脑最好用的时候，就是快要来不及的时候”。[11](#)每个要在截止日期之前完成工作的人，都会觉得自己像个英国人。截止日期是一种很有效的手段，之所以这么说，是因为它产生了时间稀缺，可以让人集中精力去做事。

在第二次世界大战时的饥饿研究中，饥饿感令实验对象对食物念念不忘；同样，截止日期的存在，也令当下的工作在人们的大脑中挥之不去。无论是会议结束前的几分钟，还是大学毕业前的几周，截止日期都会被我们强化和放大。我们会为了完成工作而投入更多的时间，各种干扰因素也会变得没那么富有吸引力。文章的截止日期临近时，我们就不会在午餐时磨磨蹭蹭；会议的结束时间临近时，我们就不会再纠结于偏离主题的讨论；毕业之日快要到来时，我们会抓紧时间享受大学生活的美好。

专注红利（Focus Dividend）

稀缺会自动将干扰和诱惑等因素推至一旁，让我们做到凭一己之力很难做成的事情。

无论是有关工作还是娱乐，只要时间有限，我们都会尽量将其利用得淋漓尽致。我们将这种现象称为“专注红利”，也就是稀缺俘获大脑时所产生的积极成果。

专注红利

各种类型的稀缺都会产生专注红利，而不仅仅局限于时间稀缺。比如，当牙膏越用越少时，我们就不会像刚拿到一管新牙膏时那样一下挤出来很多；在享受一盒价格不菲的美味巧克力时，我们总会由衷地去仔细品味最后那几块，有时还舍不得吃掉；假期的最后几天，我们会马不停蹄地游览各个景点，争取一个不落；如果有字数的限制，在文章快收尾时，我们就会以更加谨慎的态度去写作，而且还会出其不意地写出更加精彩的文章。

在与心理学家阿努伊·沙阿（Anuj Shah）的合作中，我们找到了一种可以利用专注红利的广泛影响来测试我们所得理论的方法。如果说我们的理论适用于所有类型的稀缺，而不仅限于金钱或时间稀缺，那么就应该同样适用于人工创造出来的稀缺。那么，诞生于实验室的稀缺是否也能够产生专注红利？在实验室中，我们能够去研究人们在受控条件下的行为，从而揭示出思想和行动的发生机制。这种方法与心理学领域的长期传统是一致的——心理学家们常常利用实验室来研究诸如从众、服从、策略互动、帮助行为甚至犯罪行为等重要的社会问题。

稀缺实验室

SCARCITY

出于研究目的，我们开发了一个灵感来源于《愤怒的小鸟》（*Angry Birds*）的视频游戏。在这个被我们称作《愤怒的蓝莓》（*Angry Blueberries*）的游戏中，玩家们需要用虚拟弹弓将蓝莓瞄准华夫饼——需要在射击时决定将弹弓拉多远，以及拉成什么角度。蓝莓被射出之后，会在华夫饼之间回弹，摧毁撞击到的所有华夫饼。这是一场关于目标、精准度和物理学的游戏。玩家们需要猜测蓝莓子弹的弹道轨迹，并估算子弹在华夫饼之间的弹射路线。

这一实验要求玩家至少玩20轮游戏，然后用赚到的点数来换取奖励。每轮游戏开始时，他们都会得到一些蓝莓——可以一次将所有蓝莓都射出，也可以储存一部分，留到后面几轮使用。如果玩家玩完20轮还有剩余的蓝莓，那么就可以继续玩下去，继续积累点数，直到蓝莓用光为止。游戏中，蓝莓决定了玩家的财富：拥有的蓝莓越多，意味着射击的次数越多，也就意味着玩家能够获得更多点数，可以换取更大的奖励。而实验的下一步就是创建蓝莓稀缺：我们会给一些实验对象较多的蓝莓（每一轮获得6个），而给另一些实验对象较少的蓝莓（每一轮仅获得3个）。[12](#)

玩家们表现得如何呢？拥有较多蓝莓的玩家无疑获得了更多的点数，因为他们手里的“弹药”很充足。但从另一个角度来看，拥有较少蓝莓的玩家表现得更好，因为他们的投射更加精准。出现这种现象的原因并不是后者的视觉敏锐度出现了神奇的提升，而是因为他们每次投射时都要花上更长的时间进行瞄准。（实验对游戏用时并没有限制。）由于子弹数量较少，他们就会做出更为明智的判断。而另一方面，拥有较多蓝莓的玩家们在射击时则会更加随意——这并不是因为他们可以玩更

多轮游戏，觉得无聊，于是决定少花点时间；当然，也不是因为他们玩累了。就算是第一次射击，与拥有较少蓝莓的玩家相比，拥有较多蓝莓的玩家就已经表现得没有那么专注和细心了。这种现象与我们的预测相符：拥有较少蓝莓的玩家得到了专注红利。

从某种角度来看，蓝莓稀缺产生了与截止日期引发的时间稀缺相似的效果，这一现象令人感到有些意外。在游戏中少拿几颗蓝莓，与开会时仅有几分钟时间或工作时只剩下几个小时，两者之间似乎没有多少相似之处；而专注于每次射击中应该将弹弓拉到多长以及何时射出子弹的问题，与决定工作之中的对话及进度这样的复杂选择，也不存在什么相似之处。但当我们剥去了真实世界中的所有复杂性，仅剩下稀缺一项时，相似之处就会出现。蓝莓实验的初步研究结果显示，无论真实世界中发生了什么，稀缺都能够创造出专注红利。

在受控环境下观察到的稀缺效应还可以告诉我们另外一个道理。众所周知，真实世界中的穷人与富人在诸多方面都存在着较大差异：不同背景和经历使他们产生了不同个性、能力、健康状况、受教育程度以及偏好。所以说，那些为了在截止日期前完成任务的人们，本身很可能就不是同一类人。当他们表现出不同的行为方式时，稀缺可能就是导致这些行为的一个原因，但其他一些不同之处也可能会发生作用。

在《愤怒的蓝莓》游戏中，我们用掷硬币的方式来决定谁是“蓝莓富翁”、谁是“蓝莓穷光蛋”。在这种情况下，如果这些人表现出不同的行为方式，那么我们就不能将其归因于任何有所传承的个体差距，而一定是由于蓝莓的稀缺，这是将他们区分开来的唯一因素。通过用这种方式在实验室中创造出稀缺，我们就能洞悉稀缺问题的关键。因为我们知道，稀缺本身就是原因所在。

专注红利是人们在面临截止日期时有所提高的工作效率，是在玩家们手中蓝莓数量较少时有所提高的射击精准度。专注红利来源于我们提出的核心原理：稀缺俘获大脑。“俘获”这个词在这里很重要，因为它形象地描绘了事情的发生是不可避免的，而且超越了我们的控制范畴。稀缺可以帮助我们做成一些凭借自身力量很难做到的事情。

同时，这场游戏实验也让我们重新认识了稀缺。理论上，《愤怒的蓝莓》游戏中的富翁玩家可以利用“模拟穷光蛋”的策略：他们可以像穷光蛋一样，每轮只射出3个蓝莓，将剩余的留下来备用。这种策略可以使“富翁们”拥有比真正的“穷光蛋们”多一倍的游戏轮数，从而获得两倍

的点数。而在实际操作中，蓝莓富翁们赚得的点数与两倍相差甚远。[13](#) 这可能是因为玩家们没有想到这一策略；但即便他们想到了，也基本上不会为此采取什么行动。

佯装稀缺是很难做到的。之所以会产生“专注红利”，是因为稀缺会将自身的影响强加于我们，在纷繁的琐事中俘获我们的注意力。我们发现，稀缺的影响超越了意识控制的范畴，其发生速度以毫秒计。这就是为什么迫近的截止日期会使我们能如此义无反顾地避开干扰和诱惑的原因——稀缺会自动将这些因素推至一旁。

我们不能通过挠痒自己来取乐，同样，也很难以提高工作绩效为目的，用假装有个截止日期的方式来自我欺骗。空想出来的截止日期不过是一个空想而已，它永远不会像真正的截止日期那样俘获我们的大脑。[14](#)

实验数据显示，稀缺会在许多不同的时间单位中俘获我们的注意力。在引言中我们提到，稀缺能在几毫秒间俘获我们的注意力——在几毫秒间，饥饿的人就能辨认出“蛋糕”这个词。同样，稀缺也能在几分钟（用蓝莓瞄准目标）、几天和几周的时间内发挥作用（大四学生在毕业之前充分享受大学时光的行为）。稀缺的牵引力始于最初的那几毫秒，然后逐渐积累，最后转化为表现在更长时间段之中的行为。

这样看来，其反映的就是稀缺是如何通过潜意识和有意识的行为这两个方面来俘获人类的大脑的。用心理学家丹尼尔·卡尼曼（Daniel Kahneman）[①](#)的话说就是，“在思考快与慢的两个节奏上，稀缺都会俘获大脑”。

管窥

2005年4月23日晚上10点，美国阿马里洛消防队的布莱恩·亨顿（Brian Hunton）接到了他这辈子的最后一次报警。[15](#)

有些报警不过是虚惊一场。而有一些，就像这次南波克街一处民房发生火灾的情形，却危险得令人害怕。一般情况下，消防队员们事先都无从了解事态的严重情况，所以他们必须认真对待每一次报警。消防队员们必须立刻从放松状态调整到应对火灾的战斗状态。他们不仅要迅速赶往火灾现场，而且还要全副武装，准备与熊熊火焰决斗。对于其中的每一个步骤，消防队员们都要不断演习，不断优化。就连穿防护服这件

事情，都要经过反复训练，以保证能在最短时间内做好。不过，所有努力都能得到回报。收到报警后的60秒内，亨顿和其他消防队员就已经穿戴整齐——裤子、外套、面罩、手套、头盔和靴子一件不落，并跳上了消防车。

如果不是消防业内人士，就很难理解亨顿的死因。他并没有因火焰灼烧而殉职，也不是因为吸入了烟雾，更不是因为房屋倒塌。事实上，亨顿在去往火灾现场的路上就出了事。消防车飞驰向南波克街时，遇到了一个急转弯。在全速转弯时，消防车的左后门忽然被掀开了。亨顿摔出了车外，头部撞上了马路牙子，巨大的冲击力导致他头部遭受了严重的外伤。两天之后，亨顿离开了人世。

亨顿的去世充满了悲剧色彩，因为这种事情本来是可以避免的——如果在车门不小心被掀开时，他身上系着安全带，那么他最多也只是会受点惊吓，不会因此而丧命。

亨顿的悲剧令人扼腕，但类似的悲剧还有很多。据统计，美国消防队员在车辆事故中丧生的概率排行第二位，仅次于心脏病。[16](#)1984—2000年期间，机动车相撞引起的事故，在消防队员的死因中占到了20%~25%。[17](#)在这些事故中，79%的消防队员死于没有系安全带。虽然我们无法确保消防队员的人身安全，但如此看来，只要随手系上安全带，就能挽救许多人的生命。

消防队员们对这些统计数据并不陌生，因为安全培训上总是在反复强调。亨顿在去世前一年，刚刚拿到了安全课程的毕业证。[18](#)美国消防管理局（U.S. Fire Administration）副局长查理·迪金森（Charlie Dickinson）于2007年写道：“据我了解，消防队员们在驾驶机动车时，没有不系安全带的。而且我也了解到，没有哪位消防队员不会去提醒自己的家人系安全带的。所以说，消防队员们竟会因为从消防车中被抛出去而丧失宝贵的生命，这到底是怎么回事？”[19](#)

原因就是，接到报警之后，消防队员们直面的是时间稀缺问题：他们不仅要迅速地跳上消防车，赶赴火灾现场，而且还要在抵达现场之前做好诸多准备工作——他们要在路上制定出消防策略；他们需要利用消防车上配备的计算机来研究起火建筑的结构和布局；他们需要制定出进出火灾现场的路线；还要计算出所需水龙头的数量。所有这些，都要在抵达现场之前的短暂时间内完成。虽然消防队员们非常善于管理这种时间稀缺——他们能在几分钟之内就赶到较远的火灾现场，收获巨大的专

注红利，但想获得这种红利却要付出代价。

专注于某一事物就意味着我们会忽略其他事物。想必我们都有过这样的体会：当我们沉浸在一本书或一部电视剧中时，身旁朋友提出的问题我们往往会充耳不闻。可见，专注同时也会将其他事物排除在外。所以，我们可以说稀缺令人“专注”，也可以说稀缺导致我们有了“管窥”之见——只能一门心思地专注于管理手头的稀缺。

管窥 (Tunneling)

专注于某一事物就意味着我们会忽略其他事物，也叫“隧道视野”。

“管窥”这一词语指的是人们的视野会因稀缺问题变得狭窄，形成管窥之见，即只能透过“管子”的孔洞看清少量物体，而无视管外的一切。[20](#)关于摄影技术，苏珊·桑塔格 (Susan Sontag) 曾写过这样一句名言：“摄影就是将景物装入框内，而框入一些东西就意味着其他景物会被排除在外。”[21](#)管窥就是人们对这种体验的总结与融汇。

我们了解到，消防队员们不仅要专注于如何全副武装地准时赶赴火灾现场，而且还要用管窥心态来对待任务——与任务不相关的想法都会被忽略。在亨顿的例子中，被忽略的就是消防车上的安全带。管窥现象会出现在所有人身上，并不局限于消防队员，而消防队员没有系安全带，也可能是有其他原因。但从未在你脑海中闪现过的安全带，是永远不会被系上的。

专注具有积极意义：稀缺会迫使我们专注于迫近的重要事情上；管窥则是消极的：稀缺导致我们有了管窥之见，让我们忽略了其他可能更重要的事情。

管窥所导致的忽视

管窥改变的是我们做出决策的方式。比如，某天早上，你没有像往常一样去健身房锻炼身体，而是将时间用在了工作上。这项工作的截止日期没剩几天了，现在成了你的当务之急。这种选择是如何形成的？很可能是你在经过理性思考之后，才在事物之间做出了合理的权衡。你可能是算了一下最近去健身房的频率，在多去一次健身房能给你带来的收

益以及完成工作的迫切需要之间，进行了比较，然后决定不去健身。对你来说，当天早上能多工作几个小时，比锻炼身体更加重要。在这种情境下，就算稀缺没有对你产生心理影响，你依然会认为不去健身是最佳选择。

相比之下，当我们有了管窥之见时，就会做出不一样的选择。截止日期会让我们产生短视。一觉醒来，你满脑子想的都是最为迫切的需要。你可能根本想不到要去健身，因为你那满满当当的脑子已经容不下任何其他想法了。于是，你不假思索地将健身弃于一旁。就算你想到了，此时的你也会用不同眼光去评价健身所需要付出的成本和可能得到的收益。管窥会放大其他事物的成本（在上面所花费的时间更少），缩小其他事物的收益（那些长期坚持锻炼才能看到的健康收益，现在也不那么迫切了）。无论健身是否是正确的选择，无论中立的成本收益权衡是否会让你得出同样的结论，你都会放弃健身。我们在面临截止日期时会更有工作效率，侵入大脑的分心之事会更少。出于同样的原因，我们也会做出不同的选择。

管窥之所以能发挥作用，就是因为它会影响并指挥我们去想哪些事和不去想哪些事。

如果想要切身地感受一下这种作用，就请尝试一下下面这个小任务：尽量多地列出白色的东西。我们先给出几个简单的例子来帮你开动脑筋。请花点时间想一想，除此之外，还有哪些白色的东西：

雪、牛奶.....

你能想出多少种白色的东西？这个小任务是否比你最初设想的要难一些？

研究显示，有个方法可以帮你将这项任务变得简单：不给你“雪”和“牛奶”这两个例子。通过实验我们了解到，有了这些“小引子”的帮助，人们反而只能想到更少的相关事物，就算加上那些显而易见的例子也一样。[22](#)

这种看似违背常理的结论，就是心理学家们常说的“抑制”（Inhibition）发挥作用的结果。[23](#)一旦你在脑海中建立起了“白色”与“牛奶”的链接，每次当你想到“白色的东西”时，这条被激活的链接就会将你直接带回到“牛奶”上，然后产生更为深化的激活。结果，所有其他白色的东西就都会受到抑制，让我们更不容易想到。就这样，你的脑子里一片空白。就连现在的我，也都想不到什么其他例子。“牛奶”是一个非常典型的白色事物，一旦激活与“白色”之间的链接，就会将其他所有白色的事物排除在外。这也是大脑的一个基本特点：对一项

事物的专注会抑制竞争意识。当你生某人的气时，抑制作用会令你忽略他的好处：专注于他那些让你反感的方面，就会抑制你记忆中他积极的一面。

大脑并不只针对词汇或记忆而发挥抑制作用。

稀缺实验室

SCARCITY

一项研究要求实验对象写下一项个人目标——一个他们想要获得的具有某种品质的特征（比如“受欢迎”或“成功”）。研究人员要求其中一半实验对象写下一项重要的个人目标，要求另一半实验对象随便写下一些目标。完成这一步之后，就像上面提到的牛奶实验一样，研究人员会要求两组实验对象写下尽可能多的目标（无论重要与否）。结果显示，如果实验对象一开始就写下了重要目标，那么他们列出的目标总数就会减少30%。正如“牛奶”会在脑海中将其其他白色物体排除在外一样，激活一项重要目标也会将其他竞争目标排除在外。专注于某项重要事物会令你不那么容易想到其他你所关注的事物。心理学家们将这种现象称为“目标抑制”（Goal Inhibition）。[24](#)

目标抑制是在管窥现象背后动作的机制。稀缺创造出了一个强大的目标——应对紧迫的需要，而这一目标会抑制其他目标和想法。消防队员们心中有一个目标：迅速赶赴火灾现场。而这一目标成功抑制了其他想要进入大脑的想法。这种现象可以说是一件好事，因为他们的思想此刻不会开小差，不会想到晚餐吃什么或是退休金拿多少钱，而是专心致志地准备应对前方的火灾。但同时，这种现象也可能变成坏事，因为他们想不到与直接目标不相关的事物（比如系安全带），就算想到了，更为迫切的关注点也会将其排除在外。从这个角度上理解，我们就明白了为什么消防队员们会忽略安全带的存在和发生事故的原因了。

稀缺既能为人们带来收益（专注红利），也能让人们付出代价，而抑制机制就是这两者出现的根源。抑制作用限制了分心的想法，使得人能够专注于某一项事物。引述之前举过的例子，为什么截止日期会大大提高我们的工作效率？因为我们不会那么容易分心。我们不会想到同事发来的邮件，就算想到了也会很快忘掉。而目标抑制机制就是我们不那么容易分心的原因所在。“完成本章内容”这个主要目标会俘获我们的大

脑。目标抑制机制限制了所有会导致拖延的分心事物，比如电子邮件、视频游戏或者零食；但同时也限制了我们本该投以关注的其他事物，比如去健身或回复一通重要的电话。

我们专注、管窥、着手做事、疏忽其他事，都出于同样的原因：那些存在于“管子”视野之外的事物被抑制了。在面对工作的截止日期时，我们不去健身的决定有可能是合理的，也有可能是不合理的。只不过我们因为想要在截止日期前完成工作而决定不去健身时，并不会以这种方式进行足够的思考。我们会对“管子”之内的想法反复斟酌，而“管子”之外的想法或多或少都会被忽视。比如负责对飞机进行管制的空管员。如果一架大型客机报告发动机出了问题，那么空管员就会全身心地投入解决这个问题之中。此时，空管员不仅想不起来要去吃午餐，而且也无暇顾及由他管制的其他几架飞机，这样一来飞机就有可能飞到与其他飞机相冲突的航线上。

通过《愤怒的蓝莓》这一游戏，我们了解到了专注红利。在实验室中，我们也能够看到管窥所导致的负面后果。如果因稀缺导致的忽视对得失权衡不敏感，我们就可以认为，即使忽视会产生不利影响，但稀缺还是会造成忽视。为了对此进行测试，我们与阿努伊·沙阿一同进行了另一项研究。研究中，我们为实验对象布置了简单的记忆任务，每个任务包括4张图片（见图1-1）。



图1-1 任务中的4张图片

实验对象会先记下这些图片，之后再进行回忆。研究人员为实验对象提供了这4张图片中的一张，要求他们试着回忆起剩下的3张。比如，在实验对象看过上面这些图片之后，可能会这样提问：

请重建包括图1-2的场景。



图1-2 蜘蛛图案

如果想要跳到下一轮请点击[此处](#)查阅可供选择的图片（见图1-3）。

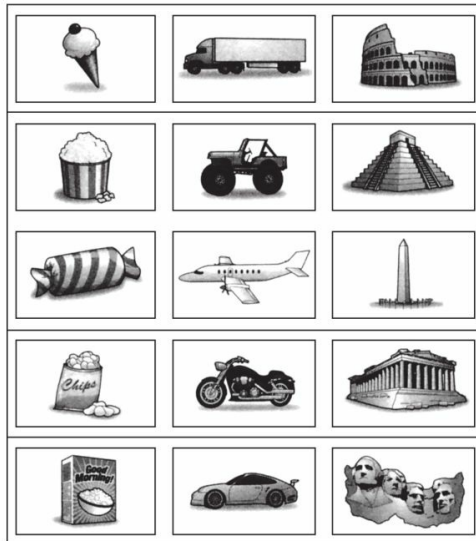


图1-3 可供选择的图片

稀缺实验室

SCARCITY

实验对象要通过记忆来判断其他几幅图片中的物体——食物、车辆、纪念碑，与给定的蜘蛛图片是如何排列的。如果回答正确，就能得分（判断所需的时长没有限制）。在这里，虽然不存在时间稀缺的问题，但是却有猜测稀缺（Guess Scarcity）的问题——实验对象只能进行固定次数的猜测。与之前的实验一样，我们设计了猜测次数较少的“贫穷”小组和次数较多的“富裕”小组。[25](#)

为了对管窥的成本进行衡量，我们还增加了一个细节。我们让实验对象两两并排而坐，进行两轮游戏。研究人员展示两组图片，请实验对象进行回忆和重建。在其中一轮游戏中，实验对象是贫穷的（猜测次数较少）；而在另一轮游戏中，实验对象则是富裕的

（猜测次数较多）。这样，他们在尝试对一组图片进行重建时，就会体验到稀缺，而另一组图片的重建则不存在稀缺。所有实验对象的总收益取决于他们在两轮游戏中的表现，他们需要将获得的总分数最大化。当然，我们也可以这样理解这场游戏：实验对象有两项工作任务，一项的截止日期是第二天，另一项的截止日期是一周以后。如果实验对象出现了管窥心态，那么他们在一轮游戏中获得的分数就会被另一轮游戏所抵消。

与专注红利实验中的情况一样，人们在贫穷时反而能够做出更有成效的猜测。但同时他们也会出现管窥心态，从而忽略另一组图片。这种现象会使他们的做事效率变得更低。在被忽略的一组图片上，实验对象的表现差了许多，他们赚得的总分数在两轮游戏中都比较贫穷的实验对象要低。虽然他们的总猜测次数更多，但得分却更少。[26](#)在两轮游戏中都面临猜测次数稀缺的问题，意味着实验对象不能忽略任何一组图片；而在一轮游戏中拥有充分的猜测机会，则容易令人们忽略这组图片，将注意力投入猜测次数贫乏的那一轮游戏之中。于是，他们对这轮游戏产生了过度关注。

如果实验对象能够在对游戏投以关注时稍加思考，就不会表现得如此极端。很明显，他们没有对管窥心态的成本和收益进行评估，只是单纯地有了管窥心态，而在游戏所创建出来的环境中，管窥心态对实验对象造成了不利影响。

我们将这种负面后果称为“管窥负担”。关于管窥负担与专注红利孰轻孰重的问题，就要具体问题具体分析了。若将游戏稍加改动，红利就会超过负担。我们的研究目的不是证明管窥的成本永远大于专注的收益，而是要证实我们产生管窥心态并非是由得失利弊的权衡来决定的，稀缺会自动俘获我们的大脑。当稀缺将我们的大脑俘获时，我们不会通过对得失利弊的谨慎思考来进行权衡。可见，我们用管窥的方式来应对稀缺，既有利，又有弊。

管窥负担 (Tunneling Tax)

因没有对管窥心态的成本与收益进行评估而产生的心智负担与不利影响。

管窥负担

我上了速读课，用20分钟读完了《战争与和平》，然后知道了这本书与俄国有关。

——伍迪·艾伦²⁷（Woody Allen）

在本章的尾声，我们会再次讲述管窥负担在日常生活中的表现，以帮助读者从直觉上加以理解。这些小例子并不是为了说明人们如何犯错误，而是为了让我们明白，管窥会导致我们忽视某些因素。首先，我们来看看《华尔街日报》上讲到的省钱攻略：

好吧，你想在下次感恩节到来之前多省出一万元钱。如何才能做到呢？关于那些千篇一律的节省方法，想必你都听了无数遍了。而且你已经有了很大进步，比如尽量少出门喝咖啡、提高保险免赔额和躲开那些商品标价不菲的店面。²⁸

提高保险免赔额，究竟是好是坏？对于手头紧的人来说，的确很难回答这个问题。没错，这样能省钱，但是却要为此付出代价：你一开始可能是省下了些钱，但一旦事故发生，就存在支付更多成本的风险。所以，对免赔额进行理性的选择，才能做好权衡。而在“管子”的视野里，只有一个被扩大的想法：现在要想办法省钱。提高免赔额就像少出门喝咖啡与看电影一样，现在就可以省些钱，于是这个想法在“管子”视野中建立了稳定的地位。其他考虑因素，比如一旦车子坏掉了如何支付维修费的问题，早已被你忽略到“管子”之外了。

这种想法不仅会让有些人提高免赔额，而且还会让有些人有了干脆连保险都不买的决定。

研究人员发现，在贫困国家，他们很难说服穷苦的农民购买各类保险，包括健康险、农作物险等。²⁹比如，降水险能在降水量过低或过高时，保护农民的生计免受损失。但是就算政府提供大量补贴，绝大多数（甚至90%之多）农民都不会购买降水险。³⁰健康险也是一样。³¹当被问到为什么不买保险时，穷困的农民总是回答说他们买不起。而这一困扰正让他们遭遇了其最不想看到的事实：因为买不起保险，反而要承担经受不了的风险。

在这个例子中，保险就是因管窥而被忽视的。对于连糊口都成问题的农民来说，下一季的干旱或身体健康问题此刻就会显得非常抽象，就会被排除在“管子”的视野之外。对于农民现在迫切需要的东西——食物、房租、孩子的学费，保险都不能予以解决，而且还会令本已揭不开锅的家庭更加拮据。

管窥的另一个表现就是，我们会因此做出同时处理多项任务的决定。我们可能会在“听着”电话会议的同时查收邮件，或者在吃晚餐的时候抽空回两封邮件。这样做的确能够节省出一些时间，但是却要付出代

价。我们很可能会错过电话会议或晚餐中某人提供的某个信息，或是回了一封词不达意的邮件。而开车时同时处理多项任务就更有可能会酿下大错了。每次想到有人开车时在做别的事情，我们首先就会联想到边开车边打电话的司机。的确，研究显示，开车时通过非手持移动电话进行通话的情况，可能比酒驾还要危险。[32](#)我们也会想到边开车边吃三明治的司机。研究显示，开车时吃东西的后果同样很严重。[33](#)很多人都有过边开车边吃东西的经历：一项研究发现，41%的美国人都曾经在开车时吃过一顿饭，包括早餐、午餐和晚餐。[34](#)边开车边吃东西虽然可以节省下一些时间，但却让车主们承担了额外的风险，如把车弄脏、出交通事故，甚至在分心时摄入更多的卡路里。[35](#)管窥会令人们想要同时做几件事情，因为这样可以节省时间，这是在“管子”视野之内获得的收益；而节省时间的同时可能会出现的问题，则处在“管子”视野之外。

有时，当我们有了管窥之见时，就会完全忽略其他事物。当我们手头有一项急需完成的工作时，就会缩短与家人共度的时光，将理财计划向后推，甚至例行的体检也要改期。时间特别紧张时，我们会想：“下周也能陪孩子玩。”而不会想：“孩子现在真的需要我，要等到什么时候我才能真的有时间呢？”对于我们来说，“管子”视野之外的事物并不容易看清，更容易被我们忽略，而且也更容易被我们遗忘。

公司也会受到稀缺心态的影响。比如，许多公司都会在淡季削减营销预算。但有些专家认为，这种做法并非明智的商业决策。事实上，这与管窥心态十分相似。正如一位中小企业顾问所说的：

一到经济衰退时，许多小企业就会犯同样的错误——大刀阔斧地削减营销预算，有的甚至将其削减为零。但是恰恰相反，经济衰退期正是小企业最需要进行营销的时刻。消费者永远都不会满足，总是想要改变他们的购买决策。你要将自己的品牌打出去，帮助消费者找到自己的产品和服务，才能在众多竞争对手中脱颖而出。因此，不能放弃营销。而如果有可能的话，中小企业更应该在经济低迷时增加营销力度。[36](#)

经济低迷时削减营销支出的决策究竟是否明智，关于这一争论，还需要大量的实证工作。我们在此想要说明的是，营销所带来的利益很像那类在管窥视野之外被忽略的事物，因为你在考虑营销问题时，心中所想则是如何削减这一季度的预算。营销就像保险一样，在“管子”视野里面是成本，而其收益则被排除在视野之外了。

在许多这类案例中，我们都会质疑这些选择是否明智。我们如何得知，边开车边吃东西所节省下来的时间，不值得为之承担因此而增加的事故风险呢？确定某个选择正确与否，永远是个难题。如果为了赶上工

作进度而忽略孩子，那么这是不是一个错误的选择？谁能做出客观的评价呢？答案取决于草率工作而引发的不利后果，没有陪伴孩子而造成的负面影响，还有你对人生的期望。局外人若要替你将这些因素全部考虑在内，是需要费一番周折的。但是，对稀缺心态的研究能够告诉我们管窥之见是如何发挥作用的，某些考虑因素是如何被我们所忽略的，从而在不需要得出这场辩论孰胜孰负的前提下，令我们对问题产生更加深入的理解。

在搞清楚稀缺心态之后，我们就会在通过某人的行为推断其偏好时更为谨慎。我们可能在看到某个大忙人总是对孩子不管不顾时，就会认为他的心思全放在了事业上，不顾家。但这样的结论可能是错误的，这就像认为不上保险的农民不在乎庄稼遭受旱涝之灾一样。事实是，大忙人很可能产生了管窥心态——他虽然也非常珍视与孩子共度的时光，但更着急去完成工作，所以将其他所有事情都推到了“管子”的视野之外。很可能多年之后，他在追忆当初时会无比悔恨，希望曾将更多时间花在孩子身上。[37](#)这种悔恨是发自内心的。产生管窥之见的任何人，日后都会产生或多或少的失望。工作必须现在完成，孩子明天还在这里。当回顾一下我们在稀缺心态下是如何利用时间和金钱时，肯定会觉得失落。眼前的稀缺问题总是会被放大，而与此无关的事物则会被忽略。当我们一次又一次地经历稀缺时，这些被忽略的事物就会越积越多。这与“没兴趣”完全不是一回事儿。毕竟，当事人会因自己的选择而后悔。

本章一开始讲述了稀缺是如何俘获我们的注意力的。而现在我们了解到，这种原始机制通过融合会产生更大的影响力：稀缺会改变我们看待事物的眼光，让我们做出不同选择。这种现象会让我们变得更有成效；但同时也会让我们为此付出代价：这种“专心致志”会令我们忽略原本很重要的其他事物。

第2章 带宽负担会降低人的智商

钱的问题会永远纠缠着穷人，而时间问题则会永远烦扰着忙碌之人。若仅仅把稀缺现象视作压力和忧虑的结果，那就是忽视了深层次的影响因素。一方面，稀缺会导致人们的认知能力下降，削弱人们的分析、判断和逻辑推理能力；另一方面，稀缺也会导致人们的执行控制力下降，削弱他们抑制行为和控制冲动的能力。

现在，我们先讲3个关于稀缺的小故事，其中就反映了专注所造成的不同结果。

故事1：管窥会造就不合格的家长

一位最重要的大客户通知你，他要将业务转移到其他公司去。你请求客户的主管经理来最后听一次你的推介。她应承了下来，但却说只有第二天有时间。于是，你取消了所有会议，暂缓了其他工作。你将所有时间都用在了推介的准备上，但有一件事情你没有办法推掉：女儿当天晚上要参加城市冠军杯垒球赛。不去观看比赛的念头划过脑海，但一番挣扎后，“好爸爸”的一面还是战胜了这个念头：对于女儿来说，她的这场比赛就像第二天的推介会一样重要。在赶往比赛场地的路上，女儿突然想起自己的幸运符落在了家里。还没掉头回家取幸运符，你就开始冲女儿大发雷霆。等你情绪稳定下来时，已经太晚了。女儿本就对当天晚上的比赛感到惴惴不安，现在你的脾气让她更害怕了。本来很有乐趣的一件事情，现在却让你们之间充满了紧张气氛。看比赛时，你一点儿也不享受。脑子里一直想着第二天要做的PPT。你也知道，你现在根本没有办法做PPT，可就是不能集中精神观看比赛——你很心不在焉。女儿在比赛间歇偷瞄了你一眼，她看出了你的心不在焉。但你很幸运，女儿的球队获得了胜利，庆祝的欢乐气氛掩盖了你犯下的错误。但不可否认，当晚的你，肯定算不上是一位合格的家长。

故事2：没有余闲就无法应对突发事件

大学生约翰第二天有个考试。虽然父母为所有孩子都存了教育基金，但还是不够。父母从来没有想过，学费竟会上涨到如此昂贵的地步。约翰是4个孩子中年龄最小的一个。等到他该上大学时，家里的教育基金已经所剩无几了，而学费则继续水涨船高。尽管如此，约翰依然选择了一所声名显赫、学费高昂的大学。他认为，如果要为学位投资，不如投在最值钱的大学上。他东拼西凑，用上了助学贷款、学院提供的补助和奖学金。虽然过程复杂周折，但他还是想尽办法凑齐了学费。一直以来，他都觉得这个选择是正确的，但现在他的心里却泛起了波澜——两项本应发放的奖学金突然被取消了。这是因为发放奖学金的基金会备受经济萧条的打击，不得不节省开支。那么，约翰该如何凑齐下个学期的学费呢？还有不到一个月的时间就要交学费了，银行能再为他批准一部分助学贷款吗？他有能力偿还这笔贷款吗？当然，他也可以向叔叔婶婶借钱，但父亲知道了一定会不高兴，但他又能怎么办呢？要不干脆转学到当地一些学费较低的普通大学？

此时，约翰实在无法集中心思准备第二天的考试。他一直在思考如何解决学费的问题。心事重重的他错过了一场本想参加的学习小组聚会。他根本不在迎战考试的状态，但他没得选。考试终于开始了。他试着集中精力，但脑子里乱糟糟的，本来很简单的题也做错了。考试结束后，他的心情变得更糟了。现在，他不仅要为学费的事情烦心，还要为这场考砸了的考试而郁闷。

故事3：表现不好都是带宽惹的祸

一家快餐店的经理将他所面临的麻烦归咎于（低收入）员工。“他们实在不靠谱。”他说道。他抱怨自己的大部分时间都用在了劝导这些员工用更友善的态度与客人打交道上。他对员工说：“服务业本来就是这样的，你要面带微笑、态度友好。客人与你说话时，陪他们聊聊天；客人无理取闹时，别跟他们一般见识。你的工作就是要有礼

貌地对待客人。”另外，这位经理余下的时间都用在处理因员工的粗心而犯下的错误上。“客人点中号炸薯条时，按下‘炸薯条’那个键，就这么简单一个动作，真的有那么难吗？”他用难以置信的口气问道。很明显，他对手下的员工感到很不满意。“可能就是因为他们心里不在乎这份工作，或者是因为这个国家的教育体制有问题，再不然，就是因为这些员工从小缺乏管教。”他说。

这3个小故事反映了稀缺俘获注意力的不同后果。在第1章中，我们了解到管窥之见会令我们所做的权衡发生扭曲——若要着眼于迫近的财务吃紧问题，我们就无法考虑提高保险免赔额对未来造成的影响。而上面的3个小故事则恰恰相反，讲的都是人们专注于与当下的稀缺问题无关的事情——那位心烦意乱的公司主管在准备推介时并没有出问题，而是在做家长、照顾女儿时出了问题；那位学生在想办法凑齐学费，以致没有办法集中精力准备考试；低收入员工并不是因为缺钱而发愁，而是因为他们无法像经理要求的那样很好地为客人服务。

这些小故事反映了一个核心假设：因为人们对稀缺的专注是无意识的、非自愿的，也因为稀缺会俘获注意力，所以就使得人们对其他事物投入了过多的关注——公司主管想要专心地观看女儿的比赛，但稀缺总是时不时地将他的思绪带回到工作上。就算我们想要试着做点别的事情，稀缺的“管子”还是会将我们的目光吸引进来。稀缺意味着在人生的另一个角度上，我们的注意力更少，投入于此的心思也更少。

心理学家们对“心思更少”的概念进行了大量的研究。详细的心理学研究对这一思想进行了细化和区分，而我们则用“带宽”一词涵盖所有这些分支概念。¹带宽，是对我们计算能力、关注能力、英明决策能力、坚持计划能力和抵制诱惑能力的衡量。从智商到高考成绩，再到冲动控制和成功节食，都与带宽相关。本章做出了一个大胆的推断：稀缺会不断将我们吸入“管子”的视野之中，形成带宽负担，并由此抑制我们的各项基本能力。

“噪声”干扰了我们的思想

试想你正坐在一个距离铁轨很近的办公室里。每小时都会有好几列火车呼啸而过。火车的声音虽然并非震耳欲聋，不会干扰人们的对话，但的确会影响工作。就在你想要集中精力做事时，一列火车轰隆隆地驶过，将你的思路又打断了。干扰本身为时很短，但其影响却很大。你需要时间，才能重新集中注意力，重新整理被打断的思路。更糟糕的是，就在你再次准备做事时，又一列火车驶了过来。

纽黑文市的一所学校位于繁忙的铁道旁边。为了评估噪声对学生学习成绩的影响，两位研究人员经过调查发现，学校的建筑物只有一面朝向铁道，而这一面教室中的学生就最直接地暴露在了火车的噪声之中，这些学生其他方面则与别的同学无异。调查显示，位于学校两侧教室中的学生有着很大的不同：铁道一侧的六年级学生，比安静一侧的同年级学生在学业表现上落后了整整一年。²研究结果出来后，市政府为此还专门安装了噪声隔离设备——这又为研究人员提供了新证据。他们发现，在噪声隔离设备安装好之后，学生之间的差距消失了。现在，建筑物两侧学生的学习成绩基本相当。随后进行的一系列研究显示，噪声会影响专注力和学习成绩。就算噪声会造成负面影响这一点不足为奇，那么这一负面影响的程度（六年级学生整整落后一年）也会令人感到惊讶。事实上，许多研究都得出了微量干扰即可产生巨大影响的结论，而这一案例只是其中之一。³

现在，请想象你正坐在一个环境优雅、安静的办公室里工作，没有干扰，也没有火车经过。但你自然也有令自己头疼的事情，那就是如何凑钱偿还按揭贷款，而且作为一位自由职业者，现在接活越来越难了。虽然你与妻子二人都外出工作，但你只有1/4的时间在赚钱。此时，你坐下来，准备干活。但没过多久，你就开始神游。“要不把家里的第二辆车卖掉？我有没有能力再承担一份贷款？”突然之间，本来安静的办公环境也变得不再安静了。这些在思想中轰隆隆驶过的“火车”，同样实实在在地存在。这些思想噪声以更快的频率造访，来得猝不及防。而且，这些思想中的“火车”还会神不知鬼不觉地把你拉上车——“要不把家里的第二辆车卖掉”这个想法最后发展成了“这样手头能多点现钱，但会让上下班麻烦许多，而且我现在还要多多努力工作赚钱，一定要尽力保住唯一一份稳定的工作”。你可能会乘着这些“思想列车”神游很久，才能跳下来，回到最初专注的工作上。虽然办公室看似安静，但却充满了各种干扰。而这些干扰，都来自你思想的深处。

这就是稀缺为带宽造成负担的过程。那些干扰我们、占据我们思想的事物，不一定来自外界，反而常常是我们自己产生的，而它们往往比真正的火车更能打乱我们的思绪。思想中轰隆隆驶过的“火车”，都带有个人特色。关于按揭贷款的想法之所以会挥之不去，是因为这个想法对于当事人来说很重要。该想法并非一次性的干扰，而是当事人非常关注的问题。正是因为这个想法导致我们产生了管窥心态，所以才称得上是一个干扰。持续的关注对我们的思想产生了影响，将我们吸入“管子”之内。正如外部噪声会干扰我们进行清晰而有序的思考一样，稀缺也会让我们产生内部干扰。

在认知科学和神经科学领域，“内部干扰”的概念众所周知。大量研究都记录并证实了“内源性思考”对总体认知能力的深远影响，就连在大脑中重复一连串数字这种小事也被囊括在内。多年来，通过核磁共振成像技术进行的实验室研究，我们从中也了解到了大脑集中注意力的方式

以及被干扰的表现形式。一个主要的区别存在于“自上而下”处理和“自下而上”处理之间：“自上而下”处理指的是思想会对所关注的事物进行有意识的选择；而“自下而上”处理则指的是注意力会被很难为我们所控制的刺激所俘获。在引言部分，有一个例子讲的是与食物有关的词汇能俘获饥饿之人的注意力。我们每个人都对这种感觉再熟悉不过：任何时候，一个快速的动作或突发的声音都能将我们的注意力从正在做的事情上吸引开来。一种特别值得注意的分心形式，它并不需要外部干扰因素的影响就会存在，那就是“神游”。

内源性思考 (Internal Thought)

即便在大脑中重复一连串数字这种小事也会对人的总体认知能力产生深远影响。

大脑处于休息状态时，默认网络会将我们的思想从正在做的事情上拉走，而我们根本意识不到。“神游”名副其实，即在意识的控制之外，我们的心神会四处游走。由此可见，虽然我们有能力直接控制大脑活动，但有些时候它们还是会失去控制。对于铁道旁边学校里的学生而言，在自下而上分心事物存在的情况下，保持专注的能力就取决于大脑当时的工作量，取决于其“装载”了多少东西。行为和神经成像研究显示，在大脑装载量较多时，引起分心的事物一旦出现，分心程度以及大脑的活动量都会有所增加。⁴而且，自上而下的注意力无法阻止自下而上的侵入。比如，在热闹的聚会场所中，当房间的另一边有人叫你的名字时，无论你正多么专心致志地关注某件事情，你的注意力都会有所转移。

稀缺也会通过自下而上的方式来俘获注意力。这就是为什么我们称之为非自愿行为的原因——行为的发生不会再受意识的控制。因此，稀缺与火车或突如其来的噪声一样，就算我们主观地想要关注一事物，心神也会因此而被扰乱。

早年进行的一项研究对这一观点进行了测试。

稀缺实验室

SCARCITY

研究人员为实验对象布置了非常简单的任务：当他们看到屏幕上出现红点时，就按下按钮。有时，就在红点出现之前，屏幕上会闪过一幅图片。对于非节食者来说，这幅图片对他们是否会看到红点并不会产生影响；而对于节食者来说，就出现了非常有意思的现象。如果他们刚刚看到了一幅有关食物的图片，就不会那么容易看到紧接着就出现的红点。举例来说，在屏幕上闪过一幅蛋糕的图片之后，节食者看到红点的概率就会有所降低，就好像蛋糕图片让他们瞬间失明了一样。只有出现有关食物的图片时，才会产生这种现象，而与食物无关的图片则没有效果。当然，节食者并不是真的失明了，他们只不过是分心了。[5](#)

心理学家将这种现象称为“注意瞬脱”（Attentional Blink），即一闪而过的食物图片会使节食者的心智“眨一下眼”。红点出现之时，他们的心思在别处，还想着刚才看到的食物。所有这一切，都发生在不到一秒钟的时间里，因为太快而无法控制。其速度之快，甚至都让人反应不过来。研究题目很好地诠释了其主题——《我看到的全部，只是蛋糕》（*All I Saw Was the Cake*）。

注意力“眨眼”是瞬间发生的现象。但是，我们认为稀缺所引发的分心效应，能维持更长的时间。为了进行测试，我们与心理学家克里斯·布莱恩（Chris Bryan）共同进行了一项研究。研究中，我们请实验对象参与一项寻找词语的游戏（见图2-1和图2-2）。[6](#)

稀缺实验室

SCARCITY

D	N	O	V	I	G	Z
I	T	J	M	S	F	U
Q	L	E	W	O	X	N
K	W	C	E	P	B	X
H	R	E	B	R	X	J
W	P	D	S	W	T	A
N	U	X	K	R	Z	S

STREET

TREE

PICTURE

CLOUD

CARPET

LAMP

DAYTIME

RAIN

VACUUM

DOOR

图2-1 寻找重点标注的词语

实验对象要找到重点标注的词语（这个例子中是STREET一

词)。找到并点击之后，会出现新的字母拼图，然后接着再去找下一个词语。

另一组实验对象接受了相同的任务，但要找的词语有些许不同：

O	Q	M	V	T	W	A	CAKE
J	O	R	G	T	M	G	TREE
R	M	X	H	T	D	K	DONUT
N	A	R	E	E	E	C	CLOUD
T	O	E	K	F	P	Z	SWEETS
Q	X	G	T	P	I	V	LAMP
J	C	A	K	E	Q	P	INDULGE
							RAIN
							DESSERT
							DOOR

图2-2 寻找两组词语中通用的偶数位中性词

对于两组实验对象来说，列表中位于偶数位置的词语是相同的。第一组实验对象的奇数位词是中性词，而第二组实验对象的奇数位词则具有一定的吸引力：街道（STREET）换成了蛋糕

（CAKE），图片（PICTURE）换成了甜甜圈（DONUT），以此类推。之后，我们对实验对象需要多长时间找到相同的词语进行了观察（这里说的相同词语指的就是那些两组通用的偶数位中性词）。

对于大多数实验对象来说，改变奇数位词并没有什么效果；而对于节食者来说则有所不同。节食者刚刚找过DONUT一词后，需要用比其他人多30秒钟的时间才能找到CLOUD一词。总体来讲，节食者并不比别人迟钝。如果CLOUD一词之前出现的是PICTURE一词，那么他们找到CLOUD的速度就会与别人一样快。关键区别就在于DONUT这个词上，这个影响是显而易见的。[7](#)这就是心理学家所谓的“前摄干扰”现象。节食者看到DONUT一词后，会将这一概念置于思想的中心。非节食者不过是简单地去寻找这个词语，找到这个词语，然后继续下一项任务。而节食者则没那么容易放下与这个词语相关的想法，继续执行任务。就算已经开始寻找下一个词CLOUD，DONUT的影子还是会像呼啸而过的火车一样，把他们的注意力带走。而当你心里想着其他事物时，自然会很难找到CLOUD一词。

前摄干扰

(Proactive Interference)

当我们心里想着其他事物时，就会将其置于思想的中心，而产生忽略行为。

你一定经历过类似的事情。就算与食物无关，也会与时间有关。比如，你有一项需要在短时间内完成的工作，但此时你又要去参加一个与此无关的会议。开会时，你能听进去多少内容？你坐在会议桌边，努力集中注意力，但就算竭尽所能，你还是会时不时地想到那份尚未完成的工作。你身在会议室，而思想却在别处。就像节食者看到DONUT一词一样，未完成的工作总是会将自己的注意力转移至别处。

试想你正在用笔记本电脑上网。在速度较快的电脑上，你可以轻松地从一个网页跳转到另一个网页。但电脑现在同时运行着好几个程序：你边播放音乐，边下载文件，还同时开着好几个浏览器窗口。突然间你意识到，自己不是在网上冲浪，而是在网上蠕动。这些同时运转的程序占据了处理器的大量空间，而浏览器的速度也因此变得极其缓慢。

对于我们的心智处理器来说，稀缺也会产生类似的效应。稀缺会不断地给大脑装载其他处理内容，而留给处理当前任务的“心智”就没有那么多了。这就引出了我们本章所要呈现给读者的核心假设：稀缺会直接减少带宽——不是减少某人与生俱来的带宽容量，而是减少其当下能用得上的容量。

为了对这一假设进行验证，我们需要对“带宽”的定义进行细化。在此，我们会将带宽这个概念作为几项具有细微差异的心理学研究成果的总和，但我们需要非常谨慎。因为一方面，作为心理学家，我们应该对这些成果及其对应的大脑功能之间的细微差异非常重视，但带宽则是一个将差异做了模糊化处理的通用说法。而另一方面，作为社会科学家，我们对稀缺效应很感兴趣，希望在此能将其中的差异按下不谈，就像我们能用“民主”或“亚原子粒子”等概念来表达思想，而不用探究二者之间的细微差异一样。就概念细节的问题做出让步之后，我们就能利用“带宽”这个概括性的说法来指代心智功能中宽泛而彼此相关的两类元素。随后，我们将对此进行深入的解析。

认知能力 (Cognitive Capacity)

它是我们解决问题、获得信息、进行逻辑推理等能力背后的心理学机制。认知能力中最突出的就是“流体智力”，即在抽象思维和推理时，在无须特定学习或体验的情况下解决问题的能力。

第一类元素可以被宽泛地称为“认知能力”，它是我们解决问题、获得信息、进行逻辑推理等能力背后的心理学机制。认知能力中最突出的就是“流体智力”（Fluid Intelligence），也就是在进行抽象思维和推理时，在无须特定学习或体验的情况下解决问题的能力。第二类元素是“执行控制力”，其作用存在于我们管理自身认知能力的过程中，包括计划、关注、发起并抑制行为和控制冲动等。执行控制力就像中央处理器一样，是我们进行日常活动的必要能力，可以决定我们的专注力和转移注意力的能力、从记忆中提取信息的能力、同时处理多项任务的能力以及自我监督的能力。[8](#)认知能力和执行控制力拥有多个层面和丰富的细节。稀缺对以上两者都会产生影响。

执行控制力 (Executive Control)

其作用存在于我们管理自身认知能力的过程中，包括计划、关注、发起并抑制行为和控制冲动等。

带宽负担1——认知能力

认知能力的一个核心特性就是流体智力。为了就稀缺对人们认知能力的影响进行测试，我们应用了流体智力最显著也是普遍为学界所接纳的衡量标准——瑞文推理测验（Raven's Progressive Matrices test）。[9](#)这一测验开发于20世纪30年代，以开发者、英国心理学家约翰·瑞文（John Raven）的名字命名。下面两幅图与典型的瑞文测验非常相似，我们的要求是：从图2-4中进行选择，以填补图2-3中空缺的内容。

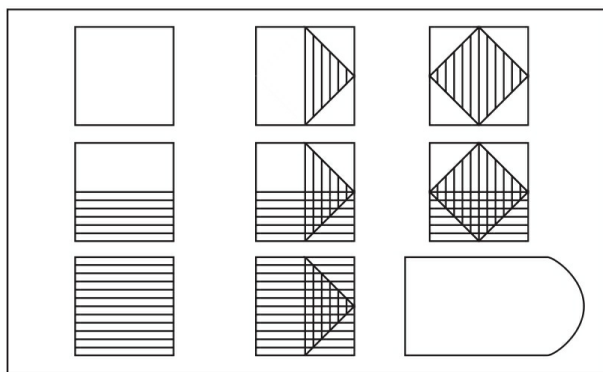


图2-3 需要填补的内容

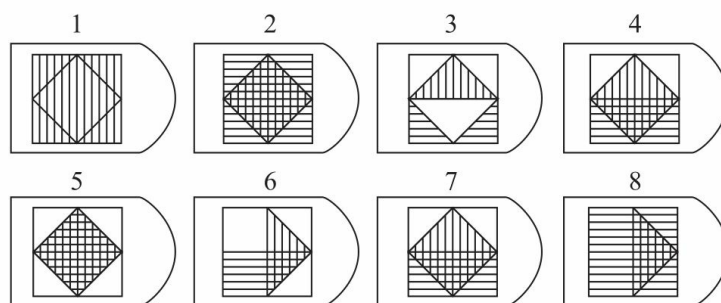


图2-4 可供选择的填补内容

在学生时代，你很可能也做过类似的测验——这是智商测验中都会包含的测试题。智商测试非常复杂，形式各不相同，但绝大多数人都认为，瑞文推理测验是其中最重要、最可靠的一个。做瑞文推理测验，被试不需要拥有任何关于世界的认识，也不需要正式的学习背景。它是心理学家、教育界人士以及军方等对流体智力进行测试的最常用的方法，可以在独立于背景知识的条件下对某人的逻辑思维能力、分析能力和解决新问题的能力进行评价。[10](#)面对一个无法启动的发动机，机械师会利用汽车制造的背景知识和推理能力来进行分析；而这位机械师在做瑞文推理测试时，却只能将其推理能力运用在一个他并不擅长的环境中。面对瑞文推理测试，机械师与来自印度的农民无异。这就使得瑞文推理测试在衡量一般智力时非常有用，而一般智力又是不受文化和国界限制的。但依然有一些人对此持怀疑态度：对测试和答案非常熟悉的人无疑能得高分；学习过几何学的人，可能得分更高；上学时间的长短也会对分数产生影响——读书时间更长的孩子比那些年龄相仿而读书时间较少的孩子得分更高。[11](#)关于流体智力，也存在智商究竟是衡量何物的争论。幸运的是，这些争论与我们的研究目的没有多大关系。我们不会在个体之间或文化之间进行流体智力的对比。我们感兴趣的是，稀缺是如何对同一个人的认知能力产生影响的。一个人的能力竟会如此轻易地受到影响，这一点也许会令你始料未及，而这正是意义所在，因为我们总

是认为，认知能力是固定不变的，但事实上，它会随着环境的变化而变化。

为了看到稀缺对流体智力所产生的影响，我们与研究生赵佳英（Jiaying Zhao，音译）共同进行了一系列研究。我们给新泽西州一家商场里的路人发放了瑞文推理测验的试卷。一开始，我们为其中一半实验对象提供了简单的假设性场景。[12](#)

稀缺实验室

SCARCITY

请想象一下，你的汽车出了问题，需要300美元的维修费。你的保险能支付一半费用。你需要决定是现在就去修理，还是等一等，盼着车子能再多开一阵？你将如何做决定？从经济学角度看，这对你将是个轻松的决定，还是艰难的决定？

在这个问题之后，我们会再次给出一系列瑞文推理测验的问题，用以测试实验对象的智商。我们依据实验对象提供的家庭收入总数，将实验对象分为富有与贫穷两类。在富有与贫穷的购物者之间，我们没有发现明显的统计学差异。虽说差异或多或少都会存在，但在我们的取样中并不是很大。可见，富人与穷人一样聪明。

我们对另一半实验对象也进行了同样的研究，只不过对给出的问题做了一点点改动。（不同的地方已加黑）。

稀缺实验室

SCARCITY

请想象一下，你的汽车出了问题，需要**3 000**美元的昂贵维修费。你的保险能支付一半费用。你需要决定是现在就去修理，还是等一等，盼着车子能再多开一阵？你将如何做决定？从经济学角度看，这对你将是个轻松的决定，还是艰难的决定？

我们只是将300美元换成了3 000美元，却对贫富两组的实验对象产生了不同影响。对于富人来说，拿出300美元或3 000美元的一半，不过是轻而易举的事情，只要用存款或信用卡解决就行了。而对于没有那么

富有的人来说，用150美元满足非常重要的需求也不难，这样一笔钱还不足以让他们考虑稀缺或自身的财务问题。

但3 000美元的修车费用则完全不同。对于低收入者来说，凑齐1 500美元的修车费存在着一定的难度。2011年进行的一项调查显示，近半数美国人没有能力在30天内凑齐2 000美元，就算是急用也没有办法。¹³我们向商场里的路人问出的这个问题虽然是虚构的，但却具有真实性，能使人联想到自身的财务状况。实验对象的车子也许没有坏，但金钱稀缺的问题一旦出现，就意味着他们会将金钱问题视为一件大事。因此，一旦触碰到大脑中的这个部分，关于稀缺的所有真实、非虚构的想法就会浮现出来：拿出1 500美元的确很难；我的信用卡早就刷爆了；马上就要到还款日了，最低还款额也让我觉得很有压力；这个月我如何才能凑齐最低还款额呢？这次还不上，我还能不能维持自己的信用？要不要借一笔工薪日贷款以渡过难关？可见，一点小涟漪就能引起脑海中的轩然大波。

这场轩然大波也会影响我们的表现。富有的实验对象面对此种稀缺并不会产生过多的想法——他们面对3 000美元和面对300美元时的心情一样，测验成绩都不错；而贫穷的实验对象，其测试成绩却会与之之前相比差很多。仅仅是稍稍涉足稀缺问题的地盘，他们的智商就会一下子呈现出很大幅度的下降。他们因稀缺而心神不宁，而其流体智力的水平也会随之跌落。

这项研究我们做过很多次，每次得到的结果都相同。这并非是因为3 000美元这一数字从数学角度来看更具挑战性。在进行非财务问题的测试时，我们发现给出绝对值较小或绝对值较大的数字，并不会对测试结果产生任何影响。只有难以处理的、与财务有关的问题（尤其是对于那些缺钱的人来说），才会产生这种结果。同时，这种结果也不是因为缺乏积极性。在研究过程中，实验对象每答对一道瑞文推理测验题，我们就会支付其一定的报酬。想必那些实验对象中的低收入者更希望题答得好，因为金钱对于他们来说更重要。但结果是，他们答得不好，而且要比之前的表现更差。在面对3 000美元的修车费用问题时，对答题酬劳满怀期待的低收入实验对象们却没能拿到多少钱；而经济状况良好的实验对象得到的钱反而更多。

在所有这些稍做修改的研究中，我们都得到了同样显著的效果。¹⁴为了更好地把握这一效果的显著性，我们以一项关于睡眠的研究为基准。¹⁵研究中，一组实验对象的作息很正常，晚上一到点就会上床睡

觉；而另一组实验对象则被迫整夜都要保持清醒。彻夜不眠会令人疲惫不堪（试想一下，你一夜没睡后的状态）。第二天早上，研究人员叫醒了熟睡中的第一组实验对象，然后请两组实验对象同时进行瑞文推理测验。结果与大家所预料的一样，一夜未眠实验对象的成绩差了许多。

与睡眠实验相比，我们在商场对路人进行研究的成绩差距到底有多大？结论是比睡眠实验的差距更大。如果一夜没睡，你第二天会觉得自己的大脑很灵活、很机智吗？我们的研究显示，只是单纯地引起低收入者对财务的顾虑，就会严重地影响他们的认知表现能力，其严重性甚至超越了失眠所造成的影响。

我们还可以利用另一种方法来理解这一研究结论的意义。瑞文推理测验是用来衡量流体智力的，是对智商的直接模拟。在典型的智商研究中，假设的是：人们的智商分数是正态分布的，平均值为100分，标准差为15分。（标准差体现的是围绕平均值所形成的分数离差。在正态分布中，约70%的分数都分布于平均值两侧的一个标准差之内。）我们可以将影响效果与标准差进行对比，从而对干扰的程度进行评价。举例来说，如果某个干扰的影响效果为标准差的1/3，那么该效果就等于智商打分中的5分。[16](#)

利用这种方法进行衡量，我们的研究所产生的效果相当于智商评分中的13分和14分。如果借用最常用的智商分类来描述，那么13分就能将你的智商从“普通”提升到“出众”。请记住，这些区别并不存在于穷人和富人之间，而是同一个人在不同影响条件下的差异化表现。同一个人在将注意力集中于稀缺问题时，他的智商分数就会比不存在稀缺问题时要低。这一点非常重要。如果车辆的维修费用不高，那么穷人的成绩就会与富人一样好，因为此时稀缺问题并没有显现出来。这或与生俱来的认知能力无关。正如电脑同时运行着多个应用程序时，中央处理器就会降低运转速度一样，穷人在此时会表现得更差一些，因为他们的一部分带宽正在想着其他事情。

带宽负担2——执行控制力

带宽的第二类元素是执行控制力。执行控制力拥有多个层面，自我控制是其中的一个重要方面。20世纪60年代后期，沃尔特·米歇尔（Walter Mischel）[2](#)与同事就“冲动”进行了一项非常有趣的棉花糖实验。[17](#)

稀缺实验室

SCARCITY

研究人员每次会让一名四五岁的小朋友坐在房间里，并在他面前放上一块棉花糖。在实验中，有的孩子会心驰神往地盯着棉花糖看，而有的孩子会激动得坐立不安。所有小朋友都很想吃，也可以吃。但在吃到棉花糖之前，他们要了解游戏规则。研究人员会离开房间一小会儿。如果在研究人员回来之前，孩子们没有将棉花糖吃掉，就可以得到第二块棉花糖。可见，这些孩子们面对的是人类所知的最古老的问题之一。

这个被社会学家托马斯·谢林（Thomas Schelling）称为“自制力在个人内心的斗争”的问题，也正是自我控制力的问题。[18](#)

自我控制是心理学研究中颇有难度的领域。我们知道自我控制力的形成因素有很多，它们取决于我们对未来的评价。而自我控制力并不会一成不变：眼前的奖励（现在就能吃到棉花糖）非常有诱惑力，能获得很高的权重；遥远未来的奖励（之后得到两块棉花糖）的诱惑力就相对要小一些，权重也会低一些。当我们在抽象的未来概念中，思考关于一块棉花糖还是两块棉花糖的问题时，两块肯定比一块好。但当一块棉花糖现在就摆在我们面前时，就会一下子将未来的两块棉花糖打败。自我控制力也取决于意志力。关于意志力的作用原理，我们并不完全了解，但它会在一定程度上受到个性、疲劳度和专注力的影响。[19](#)

自我控制力对执行控制力有很强的依赖性。我们利用执行控制力来引导注意力、发起行动、避免直觉反应或抵御冲动。米歇尔研究中的一部分内容经常为其他研究人员所重复，但却不太为公众所知。这一部分内容对我们的讨论具有很强的指导意义。能够成功抵制棉花糖诱惑的孩子，是因为他们将注意力转移到了别处——他们没有盯着棉花糖，想着将其吃到嘴里，而是在想其他事情。[20](#)他们想办法不去关注棉花糖，这样就不用与想吃的欲望做直接抗争了。正如米歇尔所说的，“一旦你意识到意志力不过是去学习如何控制注意力和想法时，就能开始提高自身的意志力了”。[21](#)

这就为执行控制力和自我控制力的关联提供了一个富有说服力的论据。由于执行控制力能够帮助引导注意力的走向并控制冲动，因此，执行力的减弱就会对自我控制力产生影响。一些实验形象地反映了这种关

系。一项实验为参与者布置了记忆任务：有些人要记住一个两位数，有些人要记住一个七位数。之后，研究人员会引导实验对象来到一个大厅，在这里等候进一步测试。等候区里摆放着蛋糕和水果。这才是真正的测试：在实验对象边等待边在大脑中重复数字的时候，他们会选择哪种食物。要求记住两位数的那些记忆量不是很大的实验对象，多数都选择了水果；而那些满脑子都是七位数的实验对象，选择蛋糕的概率则比前者超出了50%——蛋糕是冲动之下的选择。[22](#)可见，有意识的行为才能够制止自动选择。当我们的心智带宽用在记住诸如不断重复的一组数字等工作上时，用来制止我们去吃蛋糕的欲望就会减弱。

在另一项研究中，实验人员为澳大利亚的白人学生呈上了一道令他们反感的菜：整个维持原貌的卤鸡爪。这是一位华裔研究人员进行的实验，目的在于为实验对象制造出一些在饭桌前保持文明风度的压力。就像蛋糕研究一样，研究人员们会先请一些实验对象记住一组八位数——他们要在脑子里不断重复才能保证记忆的正确性。实验结果表明，那些不用记住数字的实验对象，在面对鸡爪时基本能保持风度——就算有些想法也只是默默忍受；而认知负荷量很大的那组实验对象则不然——就算他们极力想压抑内心的感受，却会情不自禁地说出一些粗俗之语。[23](#)

无论是我们本想克制自己不去吃的蛋糕，还是我们本不愿说出口的粗话，当心智带宽存在负担时，都会令我们更加难以控制冲动。稀缺会形成带宽负担，而这就意味着，稀缺不仅会降低流体智力，而且会降低自我控制力。因此，澳大利亚的学生会对着鸡爪说粗话，满脑子想着第二天推介会的公司主管会对着女儿大发雷霆，因为工资低而心事重重的员工会怠慢客户。

稀缺实验室

SCARCITY

为了解决稀缺是否会减弱执行控制力的问题，我们对新泽西州一座商场中随机挑选的实验对象进行了一项测试。这一测试经常被用来测量执行控制力和人们抑制本能的能力。首先，研究人员为实验对象呈上了假设性的财务问题，与之前一样，分为难易两类。之后，实验对象会看到如下图片（见图2-5）。

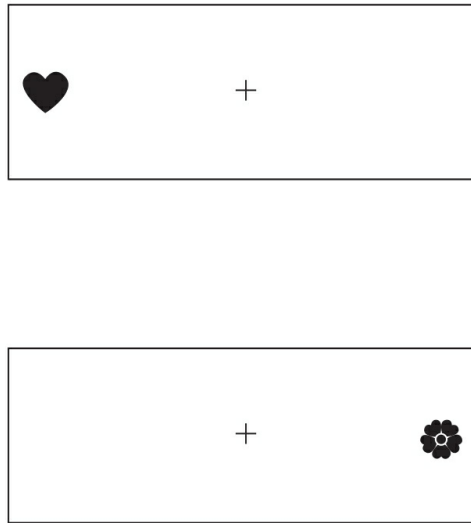


图2-5 执行控制力测试实验

图片以非常快的速度在电脑屏幕上连续播放。实验对象将双手放在键盘上，如果看到心形图案，就按下同一侧的按键；如果看到花朵图案，就按下另一侧的按键。也就是说，如果心形图案出现在右侧，就要按下右键；如果花朵图案出现在右侧，就要按下左键。

花朵图案的出现会造成需要抑制的自我冲动——按下与心形图案同侧的按键很简单，但要按下花朵图案另一侧的按键就有些难度了。若要做好，就需要战胜快速按下同侧按键的冲动。你的执行控制力越强，就能做得越好。这一测试衡量的是你抑制第一冲动的能力，无论是抵御蛋糕的诱惑、缄口不言，还是上面这个示例中避免按下与花朵图案同侧的按键。

虽然这一任务测量的是与流体智力完全不同的执行控制力，但结果却是一样的。在面对数额较小的财务问题时，穷人与富人的测验成绩会保持同步。此时，他们控制冲动的程度相同，犯错误的次数也基本相同。但数额较大的财务问题严重改变了穷人的表现。而当富裕的实验对象在面对数额较大的财务问题时，能继续表现良好，状态与遇到数额较小的财务问题时无异，也表现出了同样水平的执行控制力。而经济能力较差的实验对象，面对同等情况时的成绩却差了许多：他们变得更加冲动，更容易在按下与花朵图案同侧的按键上出错。在遇到数额较小的财务问题时，他们看到花朵图案后选对按键的概率是83%，而在遇到数额较大的财务问题时，他们选对的概率会下降到63%。可见，仅是轻轻触动一下稀缺的念头，人们就会突然间变得更加冲动。这样看来，除了流体智力之外，稀缺还会削弱我们的执行控制力。

收割前后的印度农民

在商场进行的实验检测了我们的假说。从某种角度上讲，这些实验及其结果是人为创造出来的，它们会告诉我们，当某一事物触动大脑中关于稀缺的念头时，人们会做何反应。而我们触动稀缺想法的方式，是给出能引起关于经济困难遐想的假设性问题。但我们真正的兴趣点是人们在日常生活中的表现。我们想知道，当没有研究人员在商场角落里偷偷为人们设下“圈套”时，稀缺是否还会成为人们认知能力的累赘？

这个问题的答案是我们这项研究的关键，但想找到它却并不容易，因为我们无法单纯地观察穷人与富人的认知能力或自我控制力，并在二者之间做对比。穷人与富人之间还存在着太多的不同之处，健康状况、朋友圈、受教育程度等，因此我们无法将观察到的差异都归因于稀缺。很多学者都曾尝试进行类似的对比，但都无法解决与这种对比同时存在的统计学问题。我们如何才能抛开所有这些错综复杂的因素，观察稀缺带给人们的影响？

就在此时，我们与经济学家阿南迪·曼尼（Anandi Mani）前往印度，对那里的农耕问题进行实地考察。在印度，我们注意到一些非常有意思的现象：农民在收获时节会一次性拿到一笔钱，而这就是他们的全部收入。这就意味着，从财务上看，农民与绝大多数工作者有着很明显的差异：一般工作者都是定期领取薪金的（每天、每周或每月）；而农民每年却只能拿到两次钱，有时甚至只能拿到一次。试想一位农民六月份时刚拿到一笔钱，所以在随后的几个月里过得还算不错。但就算他再节俭，再未雨绸缪，到了第二年四五月份时，他手头上的资金还是会紧张。这样看来，这位农民在收获结束后的几个月内是富有的，而在收获之前的几个月内却是贫穷的。

这与我们所要寻找的情况非常相近，所以我们能够在收获之前以及收获之后，对同一位农民的带宽进行检测。我们不用在穷人和富人之间做对比，而只要看同一个人在缺钱和不缺钱的情况下会产生什么样不同的行为方式就可以了。但这里有一个小问题需要解决，那就是农民是否需要在收获季节承担不同的财务责任？比如，收获季节有许多节日和婚宴庆典活动，这与此时人们手头资金比较充裕也有关系。这样看来，我们看到的可能不是稀缺的影响，而是其参与的庆祝活动的影响。

为了绕过这个难题，我们决定选择甘蔗农作为研究对象。甘蔗种植

很特殊。甘蔗采收之后，需要运抵大型工厂，进行甘蔗的挤榨和取汁（甘蔗汁经过蒸发后，就形成了蔗糖）。工厂的处理量是有限的，而收获后的甘蔗无法久置。因此，甘蔗的采收一般分布在4~5个月间进行。在有的地区，甚至全年的任何时候都可以进行甘蔗采收。这样，左邻右舍的甘蔗田就可能拥有完全不同的收获周期——一位甘蔗农正在采收时，他农田一侧的邻居在几个月前就采收的，而另一侧的邻居则要等几个月后才能采收。这一特殊情况正好可以帮我们绕开需要规避的问题。现在，我们就可以放手去研究同一位甘蔗农在富裕和贫穷时的表现，而且能够大胆地认为，收获前与收获后的月份并无差异。因为某个月份对于一位甘蔗农来说是收获前，而对于他的邻居来说很可能就是收获后。

正如我们所预料的那样，测试得出的数据显示，甘蔗农收获前手头上的现金会比较紧张。78%的农民在收获前的一个月都曾去当铺当过东西（99%的农民都借了钱），但只有4%的农民在收获后的一个月曾去当铺当东西（只有13%的农民借过钱）。调查显示，收获前有很多农民在支付日常开销时遇到了资金紧张的难题。[24](#)

与在商场进行的实验一样，我们对执行控制力和流体智力进行了测量。

稀缺实验室

SCARCITY

我们同样也为这些甘蔗农们布置了瑞文推理测验，但由于野外作业难度较大，无法进行“心形—花形”按键判断测验。于是，我们针对执行控制力，选择了一个比较相近的测验，也就是学界所称的斯特鲁普测验（Stroop Task）。测验中，实验对象会看到一连串标识，比如F、F、F、F，实验要求就是他们需立即说出共有几个标识。（这个例子的答案为4。）当看到2、2、2、2时，脱口而出答案“4”还是有些难度的，此种情况与看到花朵图案迅速敲击反侧键盘的道理一样。[25](#)

通过这一测验我们发现，甘蔗农们在收获前的表现比收获后差很多。这是因为，与富有时（收获后）相比，同一位甘蔗农在贫穷时（收获前）的流体智力和执行控制力更弱。这与商场里的实验对象一样，同一个人的智慧在贫穷状态时会下降，他也会变得更冲动。但是，在印度

农田中，并不是我们这些实验人员引起了与稀缺相关的想法，我们也没有利用任何手段将这些想法诱发出来。当甘蔗农们囊中羞涩时（收获时赚到的钱已为数不多），这些想法就会自然而然地存在于他们的思想之中；而当甘蔗农们吃喝不愁时（收获时赚到的钱还有很多），他们的脑子里就不会存在这些想法。

甘蔗农们在收获前后的表现也很悬殊。在瑞文推理测验中，收获后的甘蔗农比收获前的甘蔗农能多答对25%的问题。像之前提到的商场实验一样，换作智商分数来看，这就相当于智商测试的9分或10分。虽然差距没有商场实验那么明显，但也在我们的预期之中。毕竟，我们并没有利用任何方法让农民在做测验的那一刻想到钱的问题。我们只不过选定了时间，然后对他们的心智状态进行了测量，而选定的这个时间，具有因稀缺而产生带宽负担的可能性。在执行控制力的测验中，甘蔗农们处在贫穷状态时的响应速度要比富裕状态时慢**11%**，而且还会犯下比后者多出**15%**的错误，这与商场实验的结果相仿。[26](#)如果我们在一位甘蔗农贫穷时遇到他，可能会将他有限的的能力归因于个人品质。但通过研究我们发现，他的局限性跟他与生俱来的能力几乎无关。正因为收获前几个月甘蔗农们口袋中的余钱不多，所以他们才会表现得没有那么充满智慧，而其认知控制力也会相对减弱。

在宣布我们的理论成立之前，还有几个问题需要解决。我们知道，稀缺（贫穷）在收获前后会发生变化，但其他事物是否会随之一起发生改变呢？如果有，这些其他事物是否会成为心智变化的驱动力？我们考虑了以下3种可能性。

首先，如果说甘蔗农们在收获前更加贫穷，那么他们是否会吃得更少？如果是这样，那么吃得少会不会是认知能力减弱的原因之一？营养不良与饥饿确实会令人的大脑运转变慢。[27](#)但对于我们所研究的农民来说，并不存在这种情况。这些农民在手头现金不足时，也没有穷到吃不上饭的程度。从食物消费的角度来看，他们在收获后反而会在购买食物上花更少的钱。虽然他们在收获前的总体消费额减少了，但并没有在食物上节俭，而是在其他东西上省钱。举例来说，他们可能会在表兄弟结婚时买一份便宜的礼物。在印度文化中，送礼不仅是礼节性的来往，而且是一种责任，是为了偿还之前收到的赠予。由此来看，在礼物上省钱，对于他们来说，也是一个相当艰难的决定。

其次，这些甘蔗农们会不会在收获前增加劳动强度呢？为采收做准备会很辛苦。身体劳累很容易会引起心智上的疲乏。而事实上，我们的

研究时间比实际的收获日期提前了4周，而4周在农业领域已经属于相当长的时间了，所以当时他们并没有正式开始为作物采收做准备。就我们的研究时间来看，这些甘蔗农们在收获前后的劳动量和劳动强度相当。

最后，收获时节不仅是这些甘蔗农们赚钱的时候，也是其知晓自己究竟能赚到多少钱的时候。

耕种的收成浮动很大：有时丰收，有时歉收。不知道自己能收到多少回报而造成的焦虑，是否会影响他们的心智状态？对于诸如大米等农作物来说，这的确是一个很重要的问题，但甘蔗种植却不存在这种问题。只要检查一下田中作物的生长情况，甘蔗农们就能估算出自己的收入。

甘蔗的生长，其实早在收获之前的几个月就已经定型了。最后的几个月主要是增加甘蔗中的糖分，而不是增加作物的产量。糖分多少的问题，就是压榨厂所关心的问题。甘蔗农们按甘蔗的产量获得收入，而收入与甘蔗的糖分含量并无关联。他们之所以不在更早的月份采收甘蔗，只是因为压榨厂不允许。简而言之，在收获前几个月的时间里，甘蔗农们就已经能准确地估计采收时节的收入了。对于能赚多少钱，他们的心理状态在收获前后是一样的。

还有其他一些余下的小问题有待讨论，[28](#)但归根结底，有一点可以明确：贫穷会成为心智的负担。就算没有实验人员去提醒稀缺的存在，贫穷状态也会削弱流体智力和执行控制力。回到我们最初讨论的话题，这种结论就意味着，针对穷人认知能力的讨论有了重大转变。我们现在认为，穷人的“有效”能力比富人要弱一些。这并不是因为穷人的能力差，而是因为他们的一部分大脑早已被稀缺俘获了。

其他形式的稀缺

那个时候我意识到，我生存在这个世界上，而只有一部分大脑处于工作状态。其中10%用在了学习上，10%用在了女儿身上，可能还有10%用在了处理家庭危机和疾病上，剩下的70%则一直在关注食物：一颗葡萄的卡路里含量，爆米花的吸收比例，如何用喝水的方法来让自己有饱足感。我那时想：“如果将这70%的大脑以更加明智的方式投入真实的世界中，我会有什么样的作为？”[29](#)

——纳塔莉·克兹（N.Kusz）

《肥妞的歌唱》（*The Fat Lady Sings*）作者

我们都知道节食不容易：抵制美食的诱惑，对于所有人来说都有难度。而带宽负担让我们洞悉的是，节食不仅仅是有难度而已。与此同

时，它还会让心智产生负担。节食者在做事情时，会发现自己的心智资源正在变少，因为他们满脑子想的都是食物。[30](#)一些研究也证实了这一点。研究人员通过多种不同的认知测量手段，对节食者和非节食者进行了比较。这些测量手段，通常是心理学家评估有效认知能力的标准。有时，研究人员会在节食者与非节食者之间进行对比；有时，研究人员又会对同一个人的节食状态和非节食状态进行对比。无论以什么样的形式展开研究，他们都发现了同样的效应——不同的认知测验得出了相同的结论。研究人员发现，人们在节食状态下表现得更差。心理学家们在采访实验对象时发现了一个普遍的规律：节食者脑子里最重要的一件事情，就是与节食有关的种种顾虑，而整日想着这些顾虑，对他们的表现自然产生了一定的负面影响。

这种结果并不是单纯由实验对象身体中缺乏卡路里所造成的，它同样会出现在体重并未减轻的节食者身上。[31](#)这不足为奇，因为许多节食计划都以失败告终。另外，直接的生理检查结果显示，营养缺乏并不是导致这些认知减损的原因。我们可以这样来认识这个问题：在减肥的同时，你一心只想着节食，背负着带宽负担。但如果你能够找到饮食与消耗的内在平衡，不再需要限制吃喝，那么带宽负担就会消失。当然，我们也能从研究数据中找到漏洞：节食者与非节食者之间可能存在着其他方面的差异。我们还需要进行更多的研究，才能对节食者所背负的带宽负担进行量化。但卡路里稀缺的研究结论竟与我们对收入稀缺进行研究得出的结论如此相似，这倒是令人颇为吃惊。

在孤独者身上，类似的现象也会发生。一项针对孤独者和非孤独者的研究分别进行了不同类型的带宽测量，这种测量手段十分简单，叫作“双耳分听测验”（Dichotic Listening Task）。[32](#)

稀缺实验室

SCARCITY

实验对象要用两只耳朵分别倾听两种不同的声音：可能一边是女声，另一边是男声。这一测验衡量的是人们追踪一边耳朵听到的声音，同时阻断另一边干扰的能力。测验所依赖的是大脑中一个很有意思的功能：大脑侧化（Brain Lateralization）。大多数人在收听语音信息上都是右耳占优势，也就是说，人们更容易接受其他人对着自己的右耳发送语音信息。[33](#)在不接收指令的情况下，人们会专注于从右耳听到的声音。当研究人员询问实验对象其右耳听到了什

么时，孤独者与非孤独者回答得同样好。相比之下，实验对象若让自己专注于非占优势的左耳，就需要带宽的支持，需要执行控制力去超越用右耳聆听的自然倾向，从而将注意力转向左耳。这次，孤独者的成绩明显落后了——他们将关注点从右耳转到左耳的能力，比非孤独者要差。[34](#)换句话说，孤独者表现出了受损的带宽，而在这一案例中，其指的就是更弱的执行控制力。

在其他研究中，研究人员进行的实验与我们的商场实验有些类似。实验对象拿到一份测验题，以为是个性测试。做完测试之后，实验对象会随机拿到反馈结果，其会让实验对象认为，测试结果明确指出了他们之后会成为善于交际的人或者非常孤独的人。[35](#)这样，研究人员就通过引导实验对象对孤独产生预期，当场随机创造出了假想中的稀缺。待实验对象进入角色之后，研究人员会为他们布置瑞文推理测验，结果显示，预期会变得非常孤独的人的测验成绩比善于交际的人要差得多。而当研究人员为实验对象做脑部扫描时又发现，令人们产生自己将会孤独这个想法，会降低脑部负责执行控制区域的活力。在另一项有关冲动控制的研究中，研究人员拿出巧克力曲奇饼让所有人品尝，而预期将会变得孤独的实验对象吃的是其他人的两倍。[36](#)还有一项研究得出了同样的结论：一项针对中老年人饮食健康的调查发现，现实生活中的孤独者会吃掉更多的高脂食物。[37](#)

就算是人工创造出来的稀缺，也能引发类似的效应。请回忆一下第1章中讲到的《愤怒的蓝莓》游戏。通过类似的游戏我们发现，“穷光蛋”（在游戏中得到较少蓝莓的人）玩完游戏后，在“心形—花形”测验中的成绩更差。[38](#)虽然由于拥有的蓝莓较少，玩游戏的时间也短得多，但因为他们太过专注，游戏结束时，他们的带宽负担会比别人更重。与节食者、穷人和孤独者一样，拥有较少蓝莓的玩家们也会承受稀缺所引发的带宽负担。

稀缺，完全不同于压力和忧虑

稀缺并不是唯一会增加带宽负担的事物。试想，某天早上，你与妻子吵了一架。当天工作时，你的效率也不高，一整天都比平时看起来“迟钝”。而本该缄口不言的时候，你说错了话。原因就是，你的一部分带宽用在了与妻子吵架时的牢骚、烦躁和愤怒上，而能用在其他事务上的带宽就会相应变少。从这个角度来看，每个人都会产生对大脑造成负担的顾虑和需求。

那么，稀缺又有何不同呢？

稀缺是几类重要关注点的聚合。婚姻中的争执有可能在任何时间发生在任何人身上。与此不同的是，金钱问题会永远纠缠着穷人，而时间问题则会永远烦扰着忙碌之人。穷人必须要面对持续存在的金钱顾虑；而忙碌之人则必须要面对持续存在的时间顾虑。在其他所有顾虑之上，稀缺又会为这些人带来无法推卸的额外重负。稀缺会持续不断地为带宽增加负担：每个人都可能有一肚子的心事；无论贫富，谁都可能与妻子吵架；谁都会被老板搞得措手不及。但是，什么都不缺的人，其中只有一部分会满腹心事；而所有经历着“稀缺”的人，每一位都会心事重重。

这就引发了另一个重要问题。所有这些关于稀缺的讨论，是否只是在围绕着“压力”这个话题打转？

日常生活中，我们总会将“压力”这个词挂在嘴边，用它来指代许多事物。而从科学的角度来看，研究人员对压力的研究也取得了可观的进展。³⁹如今，我们已经能更加准确地了解一般性压力回应机制的生化指标，也能识别参与其中的几类分子，包括糖皮质激素（如皮质醇等）、去甲肾上腺素、血清素等，以及这些分子所发挥的作用。⁴⁰这些相关知识使得我们能够更加仔细地思考，压力是不是稀缺影响大脑所引发的生理机制。

从我们获得的数据中可以得知，压力发挥了某些作用。可见，稀缺的体验一定是有压力的。举例来说，在收获研究中我们发现，收获后的农民比收获前所承受的压力要小。我们也发现农民的心率起伏出现了明显的降低，而心率起伏是衡量压力的常用指标。

同时，压力也不太可能是我们观察到的诸多现象的主要驱动因素。有些与带宽负担相关的现象，与稀缺之间存在着必然联系。相较之下，压力却无法引发这些现象。有些研究发现，压力会增强记忆力。⁴¹还有一些研究发现了参差不齐的证据，包括在压力下出现执行控制力提升的现象。⁴²长期压力会对人造成不同的影响，而我们对稀缺的研究显示，其效应是即时的。在商场实验中，我们只不过是提醒了一下关于金钱的问题，就立刻对人们的心智能力产生了影响。另外，我们也看到了人们的表现由增强（专注红利）和减弱（带宽负担）组合而成的特定规律，而单凭焦虑和压力是无法对这一规律进行解释的。

最后，若将所有这些现象视作压力和忧虑的结果，就会忽视深层次的原因。带宽负担并不是孤立存在的研究成果，它根源于与专注红利或

塑造选择的管窥之见相同的核心机制。仅将关注点放在压力上，就会忽略这些深层次的联系，并进一步限制我们对稀缺心态的理解。

稀缺让人变笨和更加冲动

本章开头讲到的3个小故事，若从带宽负担的角度去理解，是颇为浅显的。如果收银员在火车驶过时没有听到顾客点了炸薯条，那就完全在情理之中。因此，如果收银员满脑子想的都是如何筹钱来支付当月的房租，而没有听到顾客点了炸薯条，那么你（还有经理）也不必大惊小怪。因为收银员并不是粗心疏忽，而是心事重重。“这次我能不能迟些日子再还信用卡账单？”这种想法与附近轰隆驶过的火车一样吵闹；第二天就要在客户会议上做推介的经理人，也想集中精力观赏女儿的比赛，但他自己都不知道心思什么时候又回到了那份推介材料上；想要集中精力准备考试的学生，心里却一直惦记着没有解决的学费问题。因此，当心里有负担时，就连微笑、对人友善这样简单的事情，难度都会增加。餐厅员工对粗鲁的顾客没有耐心的表现，超越了她的本意；家长也会不经意地对孩子发火。带宽有了负担，就会让人变得粗心大意——学生会忘了参加学习小组的聚会，而服务员则会点错菜。

带宽负担会以强大的力量改变我们，这些改变本身及其程度令人震惊。几十年来，心理学家们就认知负担对行为所引发的诸多效应进行了记录。我们给出的小故事反映了其中一些最重要的效应：分心、爱忘事、缺乏冲动控制能力。这些效应的规模也体现了带宽负担对各类行为所产生的实质性影响，甚至包括那些一般被归于“个性”或“才华”的耐心、忍耐力、关注力和奉献精神。许多被我们认为个性或才华的特质，都要依赖于认知能力和执行控制力。餐厅经理会用常见的原因来解释员工的行为，比如技能不足、积极性不够、受教育程度低等。而当带宽有了沉重负担时，人们的行为就会呈现出以上3种形式中的一种：满腹心事的经理人在冲着女儿发火时，似乎不是一位合格的家长；为学费发愁的学生连简单的试题都答错了，好像是能力不足、为人懒惰的表现。但这些人既非一无所长，也非情感冷漠，而是背负着沉重的带宽负担。所以，问题的根源不在于当事人，而在于稀缺。

我们之前说过，同时运转着多个程序的计算机的速度会降下来。请试想你正在操作这样一台计算机，但你并不知道还有其他程序正在运行。浏览器打开的速度出奇地慢，而这时，你很可能会得出错误的结论。你会想，这台计算机真慢，你以为这部机器出了问题。同样，我们

也很容易误解一个人，觉得他能力不足，而并不了解他的大脑正因稀缺而负重累累。这就是餐厅经理对员工做出错误判断的原因。我们与餐厅经理的不同之处是，我们了解到穷人并不是天生就有着较少的带宽。事实恰恰相反。我们的观点是，所有人一旦身陷贫穷，其有效带宽都会变窄。

现实情况告诉我们，我们需要拓宽对稀缺的认识。当我们想到所拥有的东西开始变少时，无论是时间、金钱，还是热量，都会使自己专注于稀缺的实际含义。没空，就意味着用来享乐的时间变少了；没钱，就意味着用来消费的钞票变少了。而带宽负担则说明，还存在着一个更加重要的问题：稀缺心态一旦出现，我们就必须应对心智资源更加紧缺的问题。

稀缺，不仅仅会令我们入不敷出，不知如何分配资源，而且还会让我们在生活的其他方面手足无措。稀缺会使人变笨，变得更加冲动。我们不得不在流体智力和执行控制力减弱的情况下，依靠更为有限的脑力去勉强度日。生活，就这样变得艰难起来。



第二部分

贫穷和忙碌是如何
让“带宽”变窄的

SCARCITY

WHY HAVING TOO
LITTLE MEANS
SO MUCH

第3章 装箱、余闲和权衡式思维

把东西装进行李箱，是对资源管理问题的最好比喻。我们每个人都会有一个时间箱，要在里面装上工作、休闲和与家人共度的时光。我们也都有一个金钱箱，要在里面装上住房、服饰和其他所有支出。资源的稀缺和富足，会改变我们装箱的方式。如果没有余闲存在，我们在装箱时就不得不进行权衡。可见，稀缺的本质就是没有余闲。

你马上就要出差了。试想你拖出了一只大行李箱，准备收拾行李。首先，你将所有必需品装进去：洗漱用品、正装、数码设备。装完这些物品后，你发现行李箱里还有空地，于是就又装了几件可用可不用的东西——你装上了一把雨伞，以备下雨天使用；又装进去了一件毛衣，以备冷空气突袭；还带上了运动服和跑鞋。（说不定这次真的能抽空去锻炼一下呢。）看着行李箱里的这些东西，你心满意足地拉上了拉链。行李箱还没有完全装满，还可以再装一些东西，但你觉得这些已经足够了。

现在请想象一下，同样的一次旅程，如果你只有一只小行李箱，怎么办？与之前一样，你一开始便将必需品装进行李箱，但很快行李箱就满了。于是你将东西全部拿了出来，试图利用更合理的布局重新装箱。你小心翼翼地将物品堆叠、拼凑，用颇具创意的方法腾挪有限的空间——你将袜子和手机充电器塞进鞋子，还将皮带展开，挤在行李箱侧壁上。这样一来，行李箱里就只剩下一点点空间了。此时你会想：要不要带上毛衣？要不要装上运动服？雨伞呢？要不要冒着被雨淋的风险？要不要至少给自己一个开始锻炼身体、恢复身材的机会？用小行李箱打包时，就不得不做出权衡。斟酌一番后，你决定带上毛衣，然后费了点力气才将行李箱拉上。

小行李箱迫使我们认清了一个事实：装进一样东西的同时，就必须拿出另一样东西。大行李箱的主人在思考是否带上跑鞋时，只不过是在想他需不需要这双鞋；而小行李箱的主人则是在考虑，若想带上跑鞋，就需要从行李箱中拿出某样物品，以腾出空间。

无论是大行李箱还是小行李箱，都是一种限制：无论行李箱大小，你怎么也不能将每一件可能用得上的物品都装进去；无论行李箱大小，都需要你去选择装入什么和放下什么。而从心理学的角度考虑，只有小行李箱会令人觉得这是个问题：用大行李箱收拾行李时，人们总是十分随意；而用小行李箱收拾行李时，人们便会变得小心翼翼，思索再三。

以上关于行李箱的叙述是对生活中许多其他问题的比喻。我们每个人都有时间行李箱，需要在里面装上工作、休闲和与家人共度的时光。我们也都有一个金钱行李箱，需要在里面装上住房、服饰和其他所有支出。而且，有些人还为自己制定了“热量行李箱”，要在里面装上所有食物。

这一比喻告诉我们，当稀缺令我们专注时，同时也会改变我们收拾行李箱的方式。也就是说，稀缺会改变我们花每一分钱、度过每一个小时、吃掉每一份食物的方式。稀缺会给我们一个不同的结果，装出一个内容不一样的行李箱：大行李箱总是装得随心所欲，还剩下不少空间；而小行李箱却总是装得十分谨慎，满满当当。

理解我们装箱方式的不同，是理解稀缺创造稀缺的关键。

权衡式思维

一架现代重型轰炸机的成本是这样计算的：它相当于30多座城市的现代化砖砌学校；相当于两座发电站，每座能为拥有60 000人的城镇供电；相当于两家设备精良的医院；相当于约80 450米的水泥公路。我们为每架战斗机支付的费用，相当于1 360万千克的小麦。我们为每艘驱逐舰建造的船坞，相当于8 000多人的住所。¹

——德怀特·艾森豪威尔

设想你与几位朋友正在餐厅吃晚餐。服务员介绍了特色菜，并询问你们是否需要喝点什么。你平时不怎么喝鸡尾酒，但菜单上的一款鸡尾酒却吸引了你的目光。是否要尝试一下？你该怎么决定？你可能要算一算，到晚餐结束驾车离开还有多长时间；你可能要等一等，看看几位朋友是否会点酒喝；你可能还要考虑如何买单的问题，是不是需要将账单平均分摊给每个人；你也可能会考虑，10美元一杯的价格是否合理。而真正值得关注的却是你没有想到的问题。你不会去想，如果我买了这杯酒，那么我会因此而放弃购买其他什么东西？之所以不会这样自问，是因为你感觉似乎没有这个必要。感觉上，你就算买下了这杯鸡尾酒，也不会影响其他消费。感觉上，这里并不存在权衡问题。

请思考一下以下这个有趣的问题。

从基础会计学的角度来看，权衡无处不在。无论你多么富有，金钱的数量总是有限的。如果你用10美元买了一件东西，那么口袋里就会少了用来购买其他东西的10美元，就算这个“其他东西”是你留给孩子的遗产也一样。这10美元一定有个出处，但通常却并不会给人以这种感觉。

我们许多人并不觉得花10美元还要做什么权衡，因为它通常都不需要牺牲其他消费来完成。极端一点来看，我们似乎永远有花不完的10美元。如果就这个问题深究下去，我们知道，10美元当然不是无穷无尽的，只不过我们并不以这种认识为基础而采取行动。

而有时，我们会意识到权衡的存在。请想象你在节食时，遇到了同样的“鸡尾酒问题”。就算10美元的价格不会让你考虑到权衡问题，但酒

中的“热量价码”也会让你斟酌一番。突然之间，你要将这杯酒所带来的额外的300卡路里热量考虑在内。一旦喝下这杯酒，你就要忍住吃其他东西的冲动。那么，这杯酒值不值得我们放弃甜点或明天早上的面包？只要开始节食，我们每个人都会成为精打细算的“卡路里会计”，手中的“热量账本”必须要实现收支平衡。此时，我们意识到：选择吃一样东西就只能不吃另一样东西。就这样，我们产生了权衡式思维。

权衡式思维(Trade-off Thinking)

它是由稀缺所引发的一种思维方式。在稀缺状态下，因为所有没有被满足的需要俘获了我们的大脑，以致我们开始对之念念不忘，开始产生决策难题。

10美元对于囊中羞涩的人来说，与300卡路里热量对节食者的影响是一样的。在收拾行李箱的比喻中，我们也体会到了相同的道理。

稀缺迫使我们产生了权衡式思维。所有那些没有被满足的需要俘获了我们的大脑，成了我们时时刻刻念念不忘的事情。

当我们手头缺钱时，就会非常关注那些需要支付的账单。当我们想要购买某样东西时，一直被我们惦记着的账单就会让权衡变得显而易见起来。当我们要在紧迫的截止日期之前完成某项工作时，所有那些待办事项就会凸显在我们的大脑之中。当我们想要用一个小时的时间做点别的事情时，权衡就会再一次出现。如果时间和金钱都有富余，那么我们就不会如此专注，而权衡也就不会那么明显。从这个角度来看，权衡式思维是稀缺引发的固有结果。

稀缺实验室

SCARCITY

为了对这一理论进行更为严格的测试，我们在波士顿的一个火车站对上班族展开了调查。[2](#)我们请这些乘客设想自己正在考虑购买一台电视机，并请他们列出所有考虑因素。一开始，大家都只列出了电视机大小、屏幕分辨率、价格的合理性等比较明显的因素。

而当我们把实验对象分为低收入者和高收入者两组后，就得出了一些规律：只有一部分人表现出了权衡式思维，他们主动道出了一些想法，比如具有“如果买下电视，我要放弃些什么”这样想法的人，其收入比平均水平要低很多。表现出权衡式思维的穷人，几乎是富人的两倍（15:8）。³这一差异非常显著，尤其是考虑到我们所利用的收入分隔线，最多只能粗略地将其视为稀缺的替代性表现。一些被我们归类为富人的人，很可能同样面临着稀缺问题，比如一定有人正承担着按揭贷款、信用卡的债务、大学贷款，或是要养一家人的压力。

我们在印度进行同样的调查研究时，发现了一个值得注意的地方。我们了解到，可支配金钱与物品大小之间存在着联系，而稀缺就是由这种联系决定的。与之前进行的调查一样，面对考虑购买搅拌机的问题，只有不到30%的富有实验对象提到了权衡，而超过65%的穷人都产生了权衡式思维。而当我们问到更为昂贵的物件——电视机时，无论穷人还是富人都会表现出权衡式思维。⁴可见，我们是否会联想到权衡，取决于相对可支配金钱能购买的物品的大小。对于穷人来说，搅拌机代表了其手头现钱的很大一部分，而富人则不然。相比之下，就算是印度的富有家庭，购买电视机的钱也是比较大的一笔支出。换句话说，搅拌机仅引发了一部分人的稀缺，而由于电视机的价格对于每个人的可支配金钱数量来讲都很可观，因此就引发了所有人的稀缺，就好像考虑购买汽车的问题会引发绝大部分美国家庭产生权衡式思维一样。

余闲

收拾行李箱的比喻让我们了解到稀缺为什么会创造出权衡式思维——我们会随随便便地往大行李箱里装东西，而角落里还有没利用上的空间。我们将这部分空间称为“余闲”。所谓余闲，就是我们装箱剩余下来的可利用空间。面对大行李箱，我们经常会产生余闲。余闲，就是我们在拥有很大空间，不存在稀缺心态时的产物，也是我们在资源丰富时进行资源管理的特定方式。余闲的概念，可以用来解释我们考虑（或未考虑）权衡、关注（或不关注）价格的倾向。

余闲 (Slack)

就是我们在拥有很大空间、不存在稀缺心态时的产物，也是我们在资源丰富时进行资源管理的特定方式。

试想，当你刚收拾好一只大行李箱，但又想往里面放件东西，此时你完全可以直接打开箱子扔进去，而不需要拿东西出来。你不需要对行李箱里的物件进行重新整理，因为行李箱本来就有多出来的空间——余闲。但如果是只小行李箱，要想往里面放件东西就需要拿东西出来。余闲的存在，令我们感觉不到权衡的存在。10美元一杯的鸡尾酒，钱从哪里来？如果你很有钱，那么喝杯酒就好像没花钱一样，因为从某种角度讲的确是如此。余闲替你买了单。可见，余闲可以将我们从做权衡的苦差事中解脱出来。

我们都体会过时间上的余闲。如果这周工作不太忙的话，我们在日程安排中就会留出空闲。此时，你可以在两次会议之间留出15分钟的休息时间，但在工作繁忙时，你可能会利用这点时间回个电话。现在，时间就在这里，仿佛散落于房间四处的零钱，但是你却没有要去利用这些时间的紧迫感。你不会为了完成各项事务而马不停蹄。当同事说他会于10点到11点之间给你打电话时，你没有跟他约定具体时间，而是将整整一个小时都留给了这通30分钟的电话。

许多富人会十分享受金钱上的余闲。研究显示，高收入购物者不记账的比例是普通人的两倍多，因为他们“没有必要记录，根本不差钱”。⁵

稀缺实验室

SCARCITY

荷兰学者进行的一项研究发现，富人不会在大脑中盘算自己有多少可支配的金钱，而财务规划师则会直接假定客户拥有金钱上的余闲——他们会对大笔支出精打细算，然后将余下的部分留给规划师随意支配。⁶比如，MSN公司的理查德·詹金斯（Richard Jenkins）就建议将资产的10%作为“乐趣资金”。这就是可支配金钱的余闲，即用来休闲与享受的钱。⁷

当然，对积蓄的精打细算是一个明智之举。手头留有余钱，以备不时之需，是非常英明的做法，就像是用来应对人生变化无常的保险。就算去机场只要25分钟时间，以防万一，你也应该提前45分钟出发。但是，我们并不会用“余闲”这个说法来指代那些为了应对突发事件而精心规划、刻意预留的空间。你也许会在行李箱里留出余地，以备到罗马出差时采购一番。但请注意，这里的余地是属于计划之内的，就像行李箱里的其他物品一样，在你的分配与规划之下。

我们所谓的余闲不是刻意预留的空间，而是因为装箱时空间充裕而产生的“副产品”。在我们工作顺利、收入稳定时，并不会仔细到1美元都计较的份上；买房买车这种大事，只要支出后能剩下供我们支配的充足资金，就没问题；我们大概会了解自己能以什么样的频率光顾什么档次的餐厅，这样就能基本保证不超支；我们选择的度假产品的价格只要心理价位左右就行，而不需要考虑银行账户里还有多少钱，然后再选择一款与预算分毫不差的产品。这种心态就是“充裕”的一个特点，而余闲便是其结果。[8](#)

为什么穷人的余闲更少，而富人更多呢？从大自然中，我们可以找到答案。

穷蜜蜂与富泥蜂

任何一座人工建筑都不如蜂巢那般精致。[9](#)蜂巢是这样建造起来的：年轻的工蜂会吃下大量蜂蜜，然后分泌出一点点蜂蜡——蜂蜜与蜂蜡之间的比例悬殊，1千克蜂蜡需要8.07千克蜂蜜，而为了采集这8千克多的蜂蜜，蜜蜂需要往返于花丛9万次。然后，蜜蜂会将蜂蜡聚集成小堆，用体温将之软化，以便塑形。一点一点，一块一块，蜜蜂将蜂蜡连接起来，就制成了筑造蜂巢的“砖瓦”。这些工作是在各处分散进行的，并没有统领整个项目的监工。试想，你要用一粒粒细沙建造沙堡，但却从未停下手去看看建造工作的进展情况，也没有人给你任何指示。然后再试想一下，同样一份建造沙堡的工作，你需要与几百位朋友在伸手不见五指的黑暗中完成。但尽管如此，蜂巢还是建成了。蜂巢巢壁之间的角度是精准的120度，形成了完美的六边形。每层巢壁的厚度都不到0.1毫米，偏差仅在0.002毫米之内。也就是说，误差为2%。从建筑标准的角度来看，它的建造水平很高。因为对于建筑中使用的预制内层层压板的宽度，美国国家标准与技术研究院（NIST）允许的误差为10%。

与蜜蜂一样，泥蜂也要筑巢，但它们的原材料是泥巴。[10](#)它们会对蜘蛛发起进攻，将20多具蜘蛛尸体塞进巢穴，产卵并封存。孵化出来的泥蜂幼虫以这些猎物为食，在密封的巢穴中过冬。泥蜂并不会像蜜蜂那样精心地筑巢：巢室基本是柱状的，以不规则的方式用泥巴黏合在了一起，完全比不上蜂巢的精美。

为什么蜜蜂会建造出如此精美的蜂巢，而泥蜂却马虎行事？答案在于稀缺。泥蜂采用的建筑材料——泥巴，是充裕的；而蜜蜂的建筑材料——蜂蜡，是稀缺的。蜂蜡就像塞满东西的行李箱里的空间或经济不景气时的钞票一样，必须得到妥善的保存和利用。如果草草筑巢，就会浪费蜂蜡，而为了不至于浪费，就必须提高工作效率，明智地分配资源。相反，泥蜂则拥有相对充裕的建筑材料——大量可以浪费的泥巴，它们有能力享受余闲，可以粗心地筑巢，因为它们的建筑材料实在是太廉价了。但蜜蜂不能这样做，因为它们筑巢材料的成本实在是太昂贵了。

穷人和富人也会有类似的体验。试想一下，在行李装箱前，你将所有需要带上的东西都摆在了床上——最有价值的东西放在左边，不那么重要的东西放在右边。对于为期3天的行程来说，最左边的应该是第1套换洗内衣，而第5套换洗内衣则应放在最右边。你开始将东西以从左向右的顺序装入行李箱。在行李箱装满之前，你可以放进去许多东西，而若想将行李箱装满，你就会放进去一些原本不需要的东西，比如第5套换洗内衣。所以说，空间富裕的大行李箱最终抛弃的是重要性较低的物品；而空间拥挤的小行李箱，在主人还没装完需要的东西时就满了。对于小行李箱来说，空间尤为宝贵，而对于大行李箱来说则不然。

经济学家们称这种现象为“边际效用递减”，即你拥有的越多，对你而言，每一样额外增加的物品的价值就会越低。

我们的叙述存在着这样一个经济学逻辑：穷人之所以没有那么多余闲，是因为他们没有那么多能力负担。装箱所需要的空间，对于拥有大行李箱的人来说是廉价的，就像泥巴一样；但对于小行李箱的主人来说则是昂贵的，就像蜂蜡一样。因此，大行李箱的主人会像泥蜂那样，以低效的方式随随便便地装箱——因为行李箱里还有余闲；而小行李箱的主人则会像蜜蜂那样细心，尽量不留余闲。

同时，导致这种现象出现的原因还有深层次的心理学因素。当大行李箱和小行李箱的主人在装箱时停下手来，每个人都有尚未被装进行李箱的物品时，二者就会表现出不同行为：对于小行李箱的主人而言，许

多十分重要的东西装不进行李箱的情况会唤起他们的关注和焦虑。就这样，他们会对这些装不进去的东西产生管窥心态，会不由自主地想：“难道我不能重新整理一下箱中的物品，然后将这些东西也塞进去吗？”装箱这件事俘获了他们的注意力，因为装不进行李箱里的物品的确十分重要。但当大行李箱的主人停下手来时，就说明没有装进行李箱的物品确实没那么重要——装上也好，不装也罢。[12](#)大行李箱的主人之所以会产生余闲，是因为他们并不那么在意装箱这件事情。

余闲心态下的购买行为

房子不过是一堆东西上面加了个盖子。[13](#)

——乔治·卡林（George Carlin）
美国脱口秀演员、作家

余闲都用来做什么了？如果你与大多数人一样，那么就可以自己找到这个问题的答案。去厨房打开橱柜看看吧！里面可能装满了很久以前你买回来的东西。其实许多人都会这样。一般来说，许多美国家庭的厨房里都会堆满好多年没有人碰过的汤包、果酱和罐头食品。这一现象非常普遍，食品研究人员甚至专门为这些东西起了一个名字——“橱柜遗弃品”（Cabinet Castaways）。[14](#)据估计，在从超市买回来的10件物品中，就有一件会成为橱柜遗弃品。

实际上，我们许多人家里都称得上是“遗弃品博物馆”。记不记得上次你清理衣柜时发现了一件很陌生的衣服，然后自言自语道：“我怎么不记得还有这件衣服！”这些衣柜遗弃品实在是太多了，以至于让你觉得自己真正稀缺的不是金钱，而是空间。甚至，有些人还需要租用自助式仓库，才能储存所有这些东西。有人估计，每年人们花费120亿美元用于租赁自助式仓库，是音乐类消费的3倍多。[15](#)在美国，自助式仓库共占用1.86亿平方米的面积。自助式仓储协会（Self Storage Association）曾得出这样的结论：“自助式仓库顶棚的面积可供所有美国人同时站于其上。”[16](#)

自助式仓储业之所以能得到如此兴盛的发展，主要依赖于资源充裕所形成的余闲。这与一位作家在《纽约时报杂志》（*New York Times Magazine*）上所讲的观点不谋而合。

提供仓储营销解决方案的咨询集团总裁德里克·内勒（Derek Naylor）曾对我说：“人们的懒惰情结一直是自助式仓储业从业者的‘密友’。因为只要将东西装进仓库，就没有人会想要再花上一整天时间将其搬出。只要人们有能力去支付费用——从心理上感觉他们能够付得起费用，那就会一辈子将东西放在那里。”而在2008年开始的

经济大萧条时期，他曾说：“现在，有些人要比以往更加关注信用卡账单。他们真的开始关注储存在仓库里的东西，他们开始意识到，这些东西根本不值得每个月100美元的储存成本。于是，他们就把东西丢掉了。”[17](#)

因为余闲的存在，我们才可以随意购买这些最终被丢掉的物品。有了余闲，我们会随随便便地买下特色风味的汤罐头，或者遥控飞机模型。有了余闲，我们就不会觉得一定要去思考“这个东西是否真的有用”这样的问题。我们不会想到：“我是否能常常用到这台榨汁机？这部机器值得我花这些钱吗？”或者：“我真的有机会穿上这双‘恨天高’（鞋子）吗？干脆用同样的钱来买条裤子吧？”因为不存在权衡，所以我们只能想到“为什么不呢”。余闲令我们脱离了权衡的困扰，允许我们去购买那些具有一定吸引力的东西，而不去考虑其他因素。

这样一来，低效和浪费就会产生。当我们拥有很多空闲时间时，就会无所事事，然后时间就这样被耗尽了。从各项工作的间歇中挤出来的零散时间，加在一起也会有好几个小时。原本一天可以工作16个小时，但结果只做了6个小时的事情。我们常常会用一周时间去完成一项本来只需要两天时间就能完成的事。而且，我们现在指的并不是那些你明确认为“没事可做”的时间段，而是那些没有被你合理利用的时间。当我们拥有一些自由时间时，就会散漫地将其一点点浪费掉。当我们不费什么力气就能赚到钱时，就会购买一些之后会被弃置一旁的物品。就这样，我们浑浑噩噩地过着日子，不知道时间是怎么渐渐流逝的；橱柜里装满了我们永远不会去吃的汤罐头，而自助式仓库里还存放着一大堆早就被我们忘却了的东西。

但是，余闲并不仅仅意味着低效。我们曾在一组大学生中进行过此类调查。

稀缺实验室

SCARCITY

你计划今晚在图书馆里完成第二天就要交的论文。但当你穿过校园时，忽然发现你仰慕已久的一位作家要在校园里办讲座。此时，你是坚持去图书馆，还是去听讲座？

随后，另一组大学生也拿到了同样的问题，不过里面加上了一个新选择（加黑部分），以进一步诱惑他们放弃图书馆之行：

你计划今晚在图书馆里完成第二天就要交的论文。但当你穿过校园时，忽然发现你仰慕已久的一位作家要在校园里办讲座。而且，在另一个礼堂里，一部你一直想看的外国电影也要放映。此时，你是坚持去图书馆，去听讲座，还是去看电影？[18](#)

当只有“讲座”这一个富有吸引力的选择时，60%的学生都会坚持去图书馆。而当同时面对两个富有吸引力的选择时，更多的学生（80%）会选择去图书馆。这一结果可能会让人觉得有一些诧异：虽然人们有了更多能引起兴趣的选择，却更坚持最初的决定。之所以会产生这种现象，是因为做选择实在是一件不容易的事情。当选择存在于讲座和图书馆两者之间时，你可以决定在那一天里哪件事情更重要，是学习还是休闲？但当两种休闲活动同时出现时，你就又多了一个选择：究竟哪个休闲活动更适合我？面对这个多出来的选择，人们常常会说：“算了，我还是去图书馆吧。”所以，人们常常会通过坚持最初的计划来避免选择的负担，实际上就是选择了不去做选择。

余闲能让我们轻松地避免选择负担。你之所以必须在讲座和电影之间做选择，是因为你可以利用的时间是有限的。如果你有余闲，就可以两件事情都做。当你逛街买衣服时，同时看到两件你喜欢的衣服，如果口袋里的钱有限，你就不得不二选一。如果你同时喜欢上两款口味的冰激凌，而此时你正在节食，就不得不从中选择一款。无论是金钱余闲、时间余闲还是热量余闲，都会给予你不去做选择的权利。此时，你就可以说：“我两个都要。”与米尔顿·弗里德曼（Milton Friedman）“自由选择”的理想相反，[19](#)余闲给了我们无须做选择的自由。

没有余闲时犯错，后果很严重

余闲还给了我们另一类重要收益。先来看看下面这则小故事。

亚历克斯和本路过一家服装店。两人都看上了一款皮夹克——他们俩都没有皮夹克，而且一直都想买一件。这款皮夹克非常适合，只不过太贵了，要价200美元，而且也不是非常实用。对于他们来说，正确的选择是掉头走开，但一直想要皮夹克的愿望却实在难以放弃。他们俩说：“为什么不买呢？”于是向欲望屈服，稀里糊涂地花了钱。

此时，亚历克斯的现金很充裕。他回到家才想明白：“真是花了冤枉钱！”

而本正缺钱。他回到家也想明白了：“真是花了冤枉钱！”紧接着他又想到：“现在我没有钱修车了。车子坏了，我上班就会迟到，上班迟到就会……”

本所处的环境比亚历克斯更具挑战性。显然，他们两人都没能抵挡

住皮夹克的诱惑，花了冤枉钱，自作自受。两人虽然都为这件皮夹克支付了同样的费用，但亚历克斯没过多久就淡忘了自己犯下的这个错误；而本却不能。可见，同样的错误会产生不同的后果。本所处的环境之所以更具挑战性，不是因为他遇到了更有本事的销售员，或是需要负担更高的利息，而是因为他根本就没有金钱上的余闲。

这200美元的帐单该如何支付？对于富有的亚历克斯来说，他可以用余闲来支付。就算没有头脑发昏买下皮夹克的这个插曲，他的可支配资金也没有用完——200美元可以从没有用完的部分里支出。而对于囊中羞涩的本来说，他并没有余闲。为了支付这200美元，他必须挪用计划内的其他支出，而计划内的支出都是他认为必要的。他的错误让自己付出了实实在在的代价。可见，余闲不仅会令你免于权衡，而且还意味着：即便你犯了错误，也不会让你遭受实实在在的痛苦。

现在来看一个关于时间的例子。在一项研究中，心理学家们请了一些大四学生来对完成毕业论文所需要的时间进行预估。[20](#)最后，他们预估的平均时长是34天。当研究人员问到所需要时间的最好情况和最差情况时，学生们普遍认为，27天（如果一切顺利）到48天（如果诸事不顺）都有可能。而实际上，学生们完成论文的平均时间是55天。当然，这并不是缺乏经验的大学生们所犯下的无知错误，而是每个人都有可能会经历这样的计划谬误，从经理人到制片人，谁也逃不掉。我们所有人都会对未来的计划抱有乐观的态度。就连顶尖的象棋选手都会给开始的几个回合分配过多的时间，以致后来时间紧迫。[21](#)

虽然计划谬误对于许多人来说都已经司空见惯，但并非所有人都会得到同样的结果。假设一项工作需要你在月底前完成——这项工作实际上需要40个小时的工作量，而你却错误地认为只需要30个小时，并错误地实施了工作计划。随着截止日期一天天邻近，你的这一错误也变得明显起来。你该怎么弥补漏掉的这10个小时呢？

如果你并非忙到不可开交的地步，那么这10个小时不过是一个小麻烦。你看了看工作计划，然后想办法将这10个小时加进去：几件事情可以改一下时间，几件需要完成的任务可以缓一缓，而且最重要的是，你的日历中本来就四处散落着空闲的时间。随便改一改，你就能找到需要的10个小时。

如果你这星期已经排满了各项工作，那么这10个小时就不仅仅是一个小麻烦而已了。你看了看工作计划，感觉被压得喘不过气来。实在是太糟糕了。工作好像一座摇摇欲坠的积木塔一样，如果少一件，或挪动一件，整座塔就会在顷刻间崩塌。在没办

法的情况下，你一百个不愿意地做出了几个艰难的选择：你暂缓了另一项工作（这项工作的紧迫程度仅稍稍有一点缓和），主要是害怕耽误手头上的这项工作，所以并没有认真思考。但是你要知道，你是借用了其他工作的时间，而借了总是要还的。随后的一周，你将会遭遇一场更加可怕的噩梦。

对于不那么忙的人来说，余闲会将错误消化，将错误所引发的后果最小化。而忙碌之人就没有那么容易逃避犯错的后果了。每加上一个小时，都要以其他一些事情为代价。甚至，同样的错误会引发更为严重的后果。我们只看到了余闲会导致低效——我们去购买那些注定会成为橱柜遗弃品的东西的行为，是低效地利用了时间和金钱。而在这里我们了解到，余闲还能够提供一个隐藏的功效——在我们犯错误时，它给了我们周转的空间和失败的空间。

余闲也以另一种方式将我们隔离开来。亚历克斯和本为那件皮夹克支付了同等的费用，但从某种意义上讲，本为这件皮夹克支付了更高的成本。200美元，不过是亚历克斯收入的一个小零头，是他余闲的一个小零头，但却是本收入中相当可观的一部分。同样的钱数，从比例上来讲，其代价对于本来说更加高昂。正如经济学家阿比吉特·班纳吉（Abhijit Banerjee）所说的，“诱惑负担是递减的，对于资产较少的人来说更为沉重”。[22](#)

稀缺实验室

SCARCITY

一位名叫丹·比约克格伦（Dan Bjorkegren）的经济学研究生曾对这一概念进行了测试。在印度尼西亚，他针对人们的消费规律展开了一项大调查。他将某些消费品归类为诱惑性商品。（这种分类当然是主观的，有探讨的空间。）之后，他会请受调查者自己划分哪些商品属于诱惑性商品。但鉴于这是第一次调查，也算是值得借鉴的实践活动，而且诱惑性商品清单中的东西也是经过深思熟虑后列出来的，其中包括香烟、酒类和其他成瘾性商品等。比约克格伦就人们在购买这些商品时的消费比例进行了研究，并对诱惑性商品的赋税进行了量化。研究发现，对于最为贫穷的人群，诱惑性商品的赋税高达10%，而比例会随着人们富有程度的上升而下降，直到降至消费总额的1%。不可否认，富人会在诱惑性商品上投入更多金钱，但赋税比例却会更低。

如果失误的成本更加高昂，失败的可能性更大，那么稀缺难道不会

令我们变得更加谨慎吗？这句话说起来容易做起来难。若想减少失误，仅凭努力是不够的。许多错误之所以会出现，并不是因为我们不够谨慎，而是要从我们心理过程的根源处找原因。努力和专心还不足以令我们避免计划谬误，也无法提醒我们那些根本想不到的事情，更不能给予我们钢铁般抵制诱惑的意志力。偏差是大脑运转的直接结果，而且并不总能对事件的后果作出响应。我们也许会屈服于瞬间的诱惑，在身体健康时随意吃零食，而在患有糖尿病时，也同样有可能屈服。我们在玩简单的视频游戏时，有可能会走神，而在高速公路上开车时，也有可能会走神。就算会导致更加极端的后果，偏差也会持续。[23](#)

不仅如此，稀缺还会令我们犯下更为严重的失误：带宽负担让我们身陷险境，很容易出错；忙得不可开交的人更有可能对未来的工作做出错误的规划——他此时需要将精力投入手头的工作，所以在顾及其他事项时会心神涣散、不知所措，而在这种状态下，制订出来的工作计划十有八九会存在失误。在带宽受到负面影响的情况下，我们更有可能向冲动屈服，向诱惑低头。在没有余闲的时候，我们也没有允许失误的空间。

从这个角度出发，我们就能够从新的视角来审视稀缺。滞纳金是对规划失误或忘记做事的惩罚，而其无疑也是对那些生存于稀缺之中的人们火上浇油。随处可以买到的垃圾食品会让穷人和忙碌之人患上肥胖症，而这些人更容易暴露在垃圾食品的诱惑之下，会更加不注重饮食健康。相对来说，富人和悠闲自得的人则更有可能躲避垃圾食品的威胁。低成本按揭贷款表格上那些难以看清楚的披露内容，特别容易为人们所误解（也会导致更为严峻的后果），尤其是对于那些生存于财务稀缺状态下的人。有些环境会创造出允许失误的空间，而随后这些失误又会给予人们惩罚。这种情况对于我们所有人来说都是一个难题，而对于面临稀缺的人来说更是雪上加霜。

稀缺不仅意味着人们没有失误的空间，也意味着人们更有可能出现失误。在亚历克斯和本的例子中，皮夹克就是他们面对的诱惑。买下皮夹克对于两人来说都是一个错误。但如果我们将故事写成下面这样，又会是另一种情况：

亚历克斯和本路过一家服装店。两人都看上了一款皮夹克——他们俩都没有皮夹克，而且一直都想买一件。这款皮夹克非常合适，只不过太贵了，要价200美元，而且也不是非常实用。亚历克斯很有钱，他想：“为什么不买呢？”反正他的钱也没有什么其他用途。而本有些囊中羞涩，他意识到购买皮夹克的行为不是明智之举，所以他必须抵制诱惑。

买下皮夹克对于本来说是一个错误，而对于亚历克斯来说则不是。这就是充裕的资源为我们提供的自由，它可以让我们去购买更多东西。同样一件商品，在你身无分文时是一个诱惑，而在你不差钱时，不过就是随手的挥霍。就像非节食者毫不犹豫地大吃曲奇饼，节食者必须忍住不吃；忙碌之人必须避开各种干扰，不能去与朋友喝酒，更不能坐在沙发上漫无目的地看电视，而悠闲自得的人则完全可以毫无顾虑地享受这些。

可见，稀缺不仅提高了失误的成本，也为人们制造了更多犯错、做出不明智选择的机会。稀缺状态下，将事情做对做好会变得更加艰难，因为无论是忙碌之人的时间，还是贫穷之人的金钱，都必须在十分有限的范围内精打细算。为了更好地理解这一点，我们不妨再回忆一下装行李箱的比喻。

请想象塞德希尔与埃尔德应邀去参加野餐。塞德希尔负责带水果做沙拉，埃尔德则负责带软糖。塞德希尔必须好好考虑一下如何将水果装进包里，因为他要带一个大西瓜，而他的包中已经装了许多其他水果。在塞进菠萝之后，背包实在装不下其他东西了。此时他想：也许我可以将一串香蕉掰开，沿着背包的边缘码放整齐，要不就在苹果和梨的缝隙中塞进一点点葡萄或草莓。将水果装进背包的事情看似简单，实际上却很难找到其最佳的排列方法。与此相比，埃尔德的任务就轻松多了。他将西瓜味的软糖扔进包里，又随手丢了几个橙子口味的。之后他拿起背包晃了晃，背包中瞬间就又多出了一些空间，于是他又在包里装了其他几种口味的软糖。

埃尔德也许同样面临着权衡问题，那就是他可能无法将所有口味的软糖都装上。但只要做出了选择，他的打包行动从本质上讲就会比塞德希尔简单许多——装软糖这件事情根本不需要开动脑筋来想办法。将塞德希尔和埃尔德的任务区分开来的，就是物品的“粒度”（Granularity）。水果个头相对较大，而软糖则小了许多，像沙粒一样可以随意堆积。随着物品的粒度增大，打包的复杂程度也在提高。

生活中，你是在往行李箱里装小物件还是大物件，这取决于你可支配的资金。如果你拥有的资金较少，那么一部iPod就会让你感觉到贵重，因为它可能会占据你整月开销的一大部分。而随着资金的增加，一部iPod所占的比例也会变得越来越小。在你可支配的收入中，购买iPod的资金会变得越来越微不足道。可支配资金越多，就越能令决策不那么容易引发直接后果，而且也能降低打包与装箱的复杂性。当可支配空间变少时，装箱的物件体积就会变得相对较大，装箱难度也会相应增加；而当可支配空间变多时，装箱物件的体积就会相对变小，装箱难度也自然会减轻。

当然，就算拥有大量可支配的空间，如果物件的体积真的很大，还是会令装箱变得复杂：在一起重大而漫长的刑事案件审判过程中担任陪审员，就算对于拥有大把时间的人来说都不是一件易事；购买一处度假别墅的决策，就算对于有钱人来说，也需要投入精力。但如果资源充裕，你所做出的诸多选择，平均来看，“粒度”就会变小一些，也不会占用你全部的可支配资源或计划。

这就引发了更深一层的思考。虽然我们在此关注的是稀缺所引发的心理变化，但其影响可能不仅仅存在于心理层面。我们可以从数学角度来理解稀缺——稀缺可以创造出运筹学上难度更大的装箱问题；人们的大脑会受到稀缺心理的挑战，让人们发现自己身处一个计算难度更大的环境中。[24](#)

稀缺的本质就是没有余闲

我们在本书一开头就讲到了稀缺的定义，即一种需求比资源多的主观感受。这就引出了主观感受与实际限制之间的对比，因为我们所有人都必须面对有限的金钱、时间等资源，而装箱的概念则令这种差异变得更加尖锐。实际限制和权衡永远存在：行李箱不管有多大，体积都是固定不变的。但事实上，只有小行李箱才能让我们感受到稀缺的存在。这时，我们才会注意到权衡，才会意识到我们所拥有的空间实在是太小了。同时，从客观上讲，小行李箱也令稀缺管理变得更加复杂。大行李箱不仅给予了我们更多的空间，而且也消除了稀缺的感受——我们不仅会觉得自己拥有了足够的空间，而且也想不到要去做权衡。虽然实际限制和权衡无处不在，但人们的感受却并非如此。

从这个意义上讲，余闲的概念深入到了稀缺心理的核心。拥有余闲，我们就会拥有充裕的感受。余闲不仅仅意味着低效，更是一种奢侈的心理享受。充裕不仅让我们购买更多的商品，也会给予我们随意装箱的奢侈享受、不用动脑筋的奢侈享受以及犯了错误无所谓的奢侈享受。就像亨利·戴维·梭罗曾说过的那样，“一个人的富有程度，与他所能舍弃之物的数量成正比”。[25](#)

第4章 行为经济学告诉我们的道理

1美元的价值，在穷人眼中和富人眼中是不一样的。环境条件会影响富人对1美元的价值判断。当穷人在评估1美元的价值时，会用上大脑中内化的衡量标准，而不会依赖环境进行判断。穷人是金钱价值方面的专家，他们更接近于“经济人”。人们对货币价值的衡量是相对的，这是行为经济学中的经典结论。

几年前，塞德希尔和他的一位博士生亚历克斯在印度金奈城郊，想找一辆电动三轮车载他们去下一个开会地点。他们所在的地方很少会有电动三轮车经过，所以他们等候的时间不仅漫长，而且痛苦。那一天天气湿热，空气中飘浮着灰尘和沙土。温度计上显示的37℃高温，根本不能说明任何问题。（印度南部的夏季气温需要进行温度修正，就像北方冬季刮起凛冽的寒风时，体感温度会更低一样，印度夏季的高温也会令人感觉更加闷热。）焦灼中等待了10分钟之后，一辆电动三轮车停了下来。塞德希尔松了一口气，但他这口气松得为时过早了。

金奈的每一样东西都需要讨价还价。一般情况下，这样一段行程的费用是40卢比（80美分）左右，但司机看到亚历克斯这位外国人在场（塞德希尔是印度裔美国人），发现宰客的机会到了。他开价100卢比，一番你来我往之后，司机一步步让到了60卢比，然后就一口咬定，不再让步。塞德希尔准备跳上车去——天气实在太热了，而且他们还要赶着去开会。[1](#)

亚历克斯却非常固执，他不想支付60卢比，还对塞德希尔说自己不想上车。“还会有三轮车路过的，咱们再等会儿。”此时，塞德希尔真想抽自己一个嘴巴——真不该用英文讨价还价，要是用泰米尔语的话，说不定现在已经上路了。但他实在累得不想争辩，于是只能看着三轮车绝尘而去。他们又在烈日下干巴巴地等了10分钟，终于等到了另一辆三轮车。幸运的是，这位司机同意用40卢比的价格载他们去往目的地。亚历克斯终于决定上车了。塞德希尔随着他也上了车，暗自发誓今后一定要带个靠谱点的博士生出差。

为什么亚历克斯不肯接受第一位司机的出价？他之所以会拒绝，一部分原因在于公平原则——谁也不想被欺骗。但另一方面，亚历克斯已经在印度生活了一段时间，已经适应了当地的环境，了解要价偏高不过是日常生活中的一种普遍现象，并不针对他个人。所以，他是从纯粹的金钱角度来审视这个交易的。亚历克斯说道：“让我多给钱没问题，但不能多出50%！”亚历克斯明确地做出了选择：他宁愿在烈日和尘土中多等10分钟，也不愿意多支付50%的费用。

现在，假设在另一个场景中，塞德希尔提出了这样的建议：“亚历克斯，我想让你穿着衣服蒸10分钟的桑拿浴，听着频频作响、震耳欲聋的汽车喇叭声。对了，我还会时不时地扔一把土在你脸上。但为了让你觉得这一切是值得的，我会给你50美分。”

对于此，亚历克斯很可能不会接受，而更大的可能性是他会放弃塞

德希尔，转而投奔另一位导师。但事实上，这与他在金奈所接受的权衡并无差异。而现实中，他不仅接受了，还非常坚持。这是为什么呢？

在另一次相似的情况中，塞德希尔帮一个外国人跟另一个三轮车司机因为几个卢比而讨价还价。这次，司机没讲英文，而是讲泰米尔语。他问道：“你为什么为了这点小钱计较？这点小钱对这个外国人来说小菜一碟！”从某个角度来看，司机说的没错。这点小钱对于富人来说真的不算什么。但从另一个角度来看，司机说的又是不对的，因为人们就是会在某些时候表现得好像小钱很重要一样。

对于研究判断与决策的心理学家来说，亚历克斯的行为具有高度的可预测性——我们不用去印度就能在身边见到。这种行为方式与有关选择的最古老的结论相吻合。在下面的实验中，我们分别为两组调查对象展示了下面这些文字。

稀缺实验室

SCARCITY

请想象你要花一天时间去购物，[2](#)购物清单中的一样物品是DVD播放机。在一天要结束时，你在一家商店里找到了自己想要的品牌与型号，价格是100美元。这个价格还算合理，但不是你当天看到最优惠的——还有一家商店仅卖65美元，但到它所处的位置需要你在回家途中绕路30分钟。那么，你是会在这家商店里花100美元买下DVD播放机，然后直接回家，还是会绕路去另一家商店用65美元买下同一款播放机？请思考一下你会怎么做。

请想象你要花一天时间去购物，购物清单中的一样物品是笔记本电脑。在一天要结束时，你在一家商店里找到了自己想要的品牌与型号，价格是1 000美元。这个价格还算合理，但不是你当天看到最优惠的——还有一家商店仅卖965美元，但到它所处的位置需要你在回家途中绕路30分钟。那么，你是会在这家商店里花1 000美元买下笔记本电脑，然后直接回家，还是会绕路去另一家商店用965美元买下同一款笔记本电脑？请思考一下你会怎么做。

我们可以看到，以上这两个场景都需要你为了省下35美元而绕路30分钟。但调查结果显示，大多数人都会为了购买便宜点的DVD播放机而绕路，却不会为了购买便宜点的笔记本电脑而绕路。这与标准的经济学

模型产生了矛盾，因为在经济学模型中，时间与金钱的兑换率应该是恒定的。但在这里，兑换率出现了极大的浮动。为了精确起见，我们请调查对象明确说明绕路能省下多少钱，这样就能计算出人们为时间而标上的隐含价值。得出的结论让人震惊：通过改变商品的价格，一小时的价值从5.64美元（想购买3美元一支的钢笔）到1 364美元（想购买3万美元一辆的汽车）不等。³这说明，我们所谓的节俭存在着一个违背常理的推论——我们在小物件上连几毛几分钱都会计较，而在大物件上却挥金如土。这样看来，我们所谓的节俭根本派不上用场。我们会花好几个小时在网络上东查西找，就只是为了从一双标价为150美元的鞋子上省下50美元；但我们却不会为了从一辆价值两万美元的汽车上省下几百美元，而花费几个小时的时间去做信息搜集工作。

这一推论颇具现实意义与重要性，因为其说明了人们的惯常行为与经济学家所提出的“理性”的行为标准是相悖的。如果说人们理解中的1美元价值会如此轻易地发生变化，那么对经济行为的传统分析就需要进行极大程度的延展。这些现实情况以及相关结论将心理学与经济学融合在了一起，促成了“行为经济学”的诞生与发展。这些理论的影响力是深远的，因为其分析结果具有广泛的实用性，不仅能解释亚历克斯在印度金奈的古怪行为，也可以用来解释大学生、MBA、职业赌徒以及各领域业内人士的行为。⁴我们一直认为，这些基本结论存在于每个人的行为之中。

50美元，在穷人和富人眼中的价值不一

我们与博士生克丽斯朵·霍尔（Crystal Hall）曾一起进行了一项实验，它与DVD播放机 / 笔记本电脑购买实验很类似。⁵

稀缺实验室

SCARCITY

请想象，你的朋友想要购买一台价格为100美元的家用电器。虽然商店中的这个价格已经很优惠了，但店员告诉你的朋友，一家距离此地45分钟路程的商店售卖的同一款电器还要再优惠50美元。此时，你会建议朋友前往另一家商店，从100美元中省下50美元吗？

与最初的DVD播放机 / 笔记本电脑购买问题一样，我们在叙述

中做了点手脚：第一组调查对象看到的家用电器价格是100美元，第二组调查对象看到的价格是500美元，而第三组调查对象看到的价格是1 000美元；不过，这3组商品都优惠了50美元。

一开始，我们对比较富裕的人群进行了测试。当我们在新泽西州普林斯顿火车站对过往的旅客展开调查时，我们也得出了许多学界前辈曾得出的结论：当电器的价格为100美元时，54%的人建议去另一家商店；当价格为500美元时，39%的人建议去另一家商店；而当价格为1 000美元时，只有17%的人建议去另一家商店。由此看来，电器标价越高，50美元的优惠就会显得越少。对于价格较高的大件商品来说，为了节省50美元而在路上奔波就显得太不值得了。

稀缺实验室

SCARCITY

我们在近20公里以外的新泽西州特伦顿慈善食堂进行了同样的调查。这家慈善食堂和美国其他地方的慈善食堂没有什么两样，来这里吃饭的人在年龄、性别、种族上各有不同，但都有一个共同点：没钱。在此事实基础上，我们假设，他们更愿意为了省钱而在路上奔波。实际调查结果与假设相符——当电器价格为100美元时，76%的人认为应该为省下50美元而赶赴另一家商店。

当然，不是所有人都做出了这种选择，但原因却各不相同。也许他们的时间也很紧迫，也许他们有其他需要处理的事务，也许他们因为贫穷而买不起车，所以去另一家店的交通问题没法解决。又或者，也许慈善食堂的食客们与所有人一样，为时间设定了价值。

这项研究值得注意的地方在于，我们将商品价格提升之后所收集到的调查对象的反应：当电器价格是500美元时，愿意赶路的人的比例基本没变，为73%；当价格上升到1 000美元时，愿意赶路的人的比例就会稍微上升到87%。出现这一增长的原因可能是因为购物者认为既然自己要花这么多钱，那么就一定要尽量节省。[6](#)

对于绝大多数人来说，如果能从价格为100美元的DVD播放机中省下50美元，就是很大一笔钱（相当于打了五折），但对于价格1 000美元的笔记本电脑来说，这50美元就显得不那么多了（只有九五折）。相反，慈善食堂的食客们却对这种变化无动于衷。那么，稀缺（这一案例

中特指缺钱）为何会对这一规律发挥倒置作用呢？

为了明白个中原委，我们要先了解一下关于认知的行为心理学知识。

知觉的相对性

实验心理学的奠基人之一德国物理学家恩斯特·韦伯（Ernst Weber），发现了关于人类感觉如何发挥作用的重要事实。在一项由他首创的实验中，实验对象被蒙上双眼，一只手托着装有砝码的盘子。⁷研究人员会向盘中悄悄地一点点增加金属屑，并请实验对象在感觉到重量发生变化时告知研究人员。这个实验的测量目的是：在增加多少重量的情况下，人们才会发现；多少重量是“正好能感觉到的变化”。韦伯发现，“正好能感觉到的变化”相对于背景重量的比例是恒定的。对于重量来说，这一恒定比例约为1/30——如果你举着大约1.5千克的重物，就需要至少增加大约0.05千克的重量；但如果你举着大约15千克的重物，那么则要增加大约0.5千克的重量。

韦伯的实验说明，认知具有高度的相对性。举例来说，人眼并非曝光表^③，对光度的判断与背景的明暗相关：如果你身处黑暗的洞穴中，即使一根火柴的光亮都足够照亮你的四周；如果是在一个阳光明媚的午后，你在户外咖啡厅点燃了同样一根火柴，那就可能根本察觉不到它所发出的光亮。同样，我们所能感知的物体大小也是相对的，在现实生活中常常会有这样的体验。洗衣液生产商很早以前就发现，如果瓶盖变大，洗衣液的用量就会变多。⁸大瓶盖里面的倾倒口只会占整个瓶子的很小一部分，因为我们体会到的是相对量而非绝对量，所以倾倒口就显得很小。就这样，人们每次都会多倒一些洗衣液，而洗衣液的销量也会因此增长。在某种程度上，人们对金钱的判断也是与背景相对的。这就是为什么我们会更在乎为购买一本20元钱的书去节省4元钱，而不会在乎为购买一台1 000元钱的冰箱去节省100元的原因。在印度金奈，亚历克斯不过是用看火柴光亮的眼睛去看待金钱，是在与背景相对比——当公允价格为40卢比时，60卢比看起来的确是太贵了。

虽然相对性认知是大脑处理信息时的固有特征，但经验与专业技能还是能让我们摆脱这一限制。

稀缺实验室

SCARCITY

心理学家西蒙·格朗丹（Simon Grondin）与彼得·基利恩（Peter Killeen）曾进行过一项研究。两组实验对象参与了这项研究：一组是非音乐家，另一组是接受过11~23年正规音乐培训的音乐家。研究人员请两组实验对象分别弹奏6秒、12秒、18秒、24秒的八音阶间隔的反复高低音。[9](#)非音乐家的表现如最初所料——错误与音程长度成正比。也就是说，音程越长，错误越多。可见，他们是从相对角度去估计音程的。相比之下，接受过大量音乐训练的那组实验对象，随着音程增长，其表现出了逐渐缩小的变化性。也就是说，音程越长，音乐家们出错的比例就越小。可见，他们的判断力更接近于绝对标准。

从这一实验中我们了解到，专业技能，也就是对某一领域知识的深度理解，能够对认知形成改变。对音程有着专业理解的音乐家们，大脑中存在着内化的度量标准，他们不需要依靠对时间长度的直觉性进行启发式的估计。研究显示，经验丰富的酒保在倒酒时更富技巧，在客人要求倒一定量的酒时，他们的判断力不会那么容易受到酒瓶高度的影响。[10](#)

稀缺也会迫使我们成为专家——装箱专家。在没有任何余闲时，我们深谙行李箱中每一寸空间的价值。穷人珍视每一元钱的价值，日理万机的人珍视每一个小时的价值，节食者珍视每一卡路里的价值。

营销研究人员以特定方式对专业技能问题进行了研究。他们在超市门口，对采购之后正准备走出大门的顾客进行了快速访问。[11](#)他们拿到了顾客的购物清单，并询问了一些问题：“你刚刚购买的高露洁牙膏是多少钱？”针对这一问题，生活富裕的采购者基本都答不上来。“高露洁牙膏的价格？差不多3美元？要不就是5美元？”大多数富人甚至连他们刚才一共花了多少钱都不知道。但低收入购物者却对一切消费都心中有数：他们不仅会记得采购费用的总数，而且也会记得各项物品的价格。在我们进行的一项研究中，同样的现象也出现了。对于这项研究，我们进行了精心的设计，从而将知识与频繁的经验区别开来。我们针对波士顿的旅客进行了调查，并询问他们出租车计价表上的起步价是多少。[12](#)富人中只有12%给出了正确答案，而没有那么富有的人，答对的人数是富人的3倍。但事实上，富人打车的频率要多得多。

对价格心知肚明，并不仅仅是留意价签的结果。同时，你还需要一

定的警觉性，因为你看到的价格通常不是你所支付的货款。举例来说，香烟收取的税金存在两种分类：消费税会显示在价签上，而销售税则不会。总价只有在收银台结算时才会显示。如果你只看到价签上标明的价格，就是忽略了销售税。当标识于明处的消费税发生变化时，穷人与富人都会作出回应——他们的吸烟量会有所下降。¹³而当藏于暗处的销售税发生变化时，就不会出现以上现象。此时就只有低收入消费者才会作出回应，因为只有低收入消费者会同等地看待消费税和销售税（他们应该这样做）：他们不仅会注意到标明的价格，而且也会更加明白，总价不仅是价签上标注的数字。¹⁴

低收入消费者在其他方面也非常精明。去超市购买薯片或金枪鱼罐头时，他们会自然而然地假定，大包装商品的单价一定更便宜，所以购买大包装商品能省钱。而事实上，这种假定常常是错误的。大包装商品的单位价格有时更高，因为可能存在“大件商品附加费”。一项调查发现，在不同型号包装的品牌商品中，其中25%的商品存在某种形式的大件商品附加费。¹⁵这些附加费并非定价失误，《消费者报告》

（*Consumer Reports*）称其为“消费品的卑鄙伎俩”。¹⁶在那些不留意价格，只是假设大包装商品会更合算的消费者身上，这一伎俩非常奏效。

（你是否也经常落入这种圈套？）一项研究对哪些超市善于玩弄这些“花招”进行了调查，结果与我们所讨论的话题不谋而合——位于低收入社区的超市较少出现大件商品附加费。¹⁷当消费者十分在意花出去的每一元钱时，商家就很难从他们身上榨出什么油水了。

简而言之，穷人是金钱价值方面的专家。他们在评估1元钱的值时，会用上大脑中内化的衡量标准。他们并不会凭借环境去判断物品值多少钱。需求的压力始终存在于穷人心中，挥之不去，从而造就了他们自身的内化尺度。拥有内部标尺的意义在于，他们不像普通人那样容易受到背景环境的影响，如同音乐家能准确地掌握节奏一样。慈善厨房的食客们不会表现出与亚历克斯在金奈时表现出的同样的偏差，或者那些高收入实验对象表现出的同样的偏差，因为他们不那么容易利用环境中随处获取的特征作为价值衡量标准。

这一现象值得我们深思。研究中，穷人表现得更加“理性”。从这些案例中可以看出，穷人更贴近于理性的经济学理想，更接近于“经济人”。从中，我们不仅了解到了一些关于贫穷的内涵，也了解到了一些行为经济学的内涵。人们对货币价值的衡量是相对的，这一点是行为经济学中的经典结论。行为经济学假设每个人的思维中都包含这样的特

征。但在这里，我们了解到稀缺的概念已经冲淡甚至推翻了这一经典结论。事实上，稀缺也改变了许多其他结论。

1美元的真正价值

记得塞德希尔还在念本科时的一天，他正在考虑是不是要买一部随身听。（可能有人不知道什么是随身听。随身听就像iPod，只不过里面装的是磁带。不过也可能有人不知道磁带是什么，对于这些朋友，此问题暂且搁置。）

随身听的价格是70美元。那么，随身听值不值70美元呢？他应不应该买呢？其实，这个价格还算公允，因为塞德希尔四处逛了一下，发现没有比这更便宜的了。但他到底更倾向于选择哪一个？是70美元现金，还是一部随身听？深究下去，70美元到底是什么？1美元的真正价值又是什么？面对这种决策难题，塞德希尔有一套自己的方法。当时，他总吃（基本上只吃）塔可贝尔（Taco Bell）的豆沙馅玉米煎饼，于是决定以豆沙馅玉米煎饼作为衡量一切事物价值的基本标尺。他没有思考是买随身听还是存下70美元现金的问题，而是思考他是想要随身听还是买78个豆沙馅玉米煎饼的问题。因为与钞票比起来，豆沙馅玉米煎饼更加有形，更加真实。

为什么塞德希尔需要去构建一个比较基准，从而为70美元赋予意义？因为余闲的存在。资源充裕意味着不需要做权衡。当在资源充裕的情况下发生购买行为时，我们感觉不到自己需要放弃任何东西。从心理上讲，这种感觉很开心。但在做决策时，却是一种障碍。如果你不清楚自己为了获得某件东西而需要放弃什么，那么就很难想清楚这件东西的成本和价值。余闲以及权衡的缺失意味着我们无法拥有对物品进行价值判定的简单、直接的方法。

对于塞德希尔来说，豆沙馅玉米煎饼这个衡量标准也并没发挥多大作用。但这种方法与专家给出的建议有一定的相似性。一位研究决策的心理学家认为，一款iPhone的APP就起到相同的作用。¹⁸“你可以在APP中输入‘我想去巴哈马群岛度假，还想买鞋子、拿铁咖啡和书籍’。之后，当你想要购买某样东西时，这样东西的价值就能转换成一定数量的你感兴趣的東西。APP会告诉你，‘这件物品的价值相当于在巴哈马群岛度假半天、两双鞋子和一杯拿铁咖啡’。”还有专家建议利用“时间价格”作为衡量标准。假设你的工作每小时能赚20美元（扣除通勤成本、各项税款等之后的净收入），如果你想购买一台价值80美元的冰激凌机，那么就需要多工作4个小时；如果你想购买每个月60美元的有线电视套餐，那么就需要从此之后每个月多工作3个小时。（如果你每天要喝一大杯拿铁咖啡，那么每年就要多工作大约50个小时。）

在思考关于随身听的问题时，塞德希尔意识到这种推理具有误导性。他已经吃够了豆沙馅玉米煎饼。假设他选择不购买随身听，也不会再出去额外多吃78个玉米煎饼。所以，他并不是在随身听和玉米煎饼之间做权衡。为了让这种思维方式能真正发挥作用，他需要知道，不购买随身听而省下来的钱能花到哪里。这些钱肯定不会用来购买豆沙馅玉米煎饼，而管着自己不乱花钱，也不会因此就能去巴哈马群岛度假。若想让权衡具体化，就需要对节省下来的金钱进行追踪，搞清楚这些钱之后会有何用。同时，我们还需要思考另一个问题：我们如何挑选不同的物品进行对比，才具有切实的意义？

人们总是在价格相当的物品之间进行对比，而这种对比又非常具有误导性，因为许多这类物品是他们在任何情况下都不会购买的。同样，时间价格（“这样东西的价格相当于4个小时的工作所得”）也具有误导性，因为即便你管住自己没有买某样东西，也不会因此少工作几个小时；而即便你买了某样东西，也不会因此多工作几个小时。寻找金钱的最佳用途同样会产生误导作用。

如果我花40美元吃了一顿大餐，那么并不能说我花出去的每40美元都能给我带来同等的愉悦。就算我花出去的每笔钱都无比正确，也只有很少的40美元能与这次大餐带给我的享受相提并论。我每天能享用多少顿这样的美味大餐呢？依据收益递减原则，我最后花出去的40美元，也就是我正在思考的准备做权衡的40美元，其带来的体验根本无法与那种愉悦同日而语。

所有这些比较基准存在的问题都要归因于它们本身就是不现实的。在稀缺状态下进行权衡，就像是鱼和熊掌两个都想要一样。由于现实中我们并不会进行太多真正的权衡，因此权衡不过是被虚构出来的。在没有这些权衡的情况下，少量金钱并不会让你费多少脑力。如果你多出来20元钱，那么你会用这点钱去购买什么你尚未买到的东西？如果你的经济状况良好，就永远不需要回答这个问题，连想都不用想。如果你想要某件只需花20元钱就能买到的东西，你早就买了。

这些问题之所以存在，是因为我们在资源充裕的条件下并不会理解10元钱的真正价值，而这种理解上的模糊会令我们更容易落入他人的操控之中。通过精心设计的对比，购买行为可以看起来更富有吸引力，或者没那么富有吸引力。度假时将酒店房间升级为更好的房型，如果将多出来的费用与你所支付的房租做对比，就会显得微不足道。但如果将多出来的费用换成可以购买到的美味甜点，你就会觉得这是一大笔钱。营销公司和非营利性机构常常会使用这一策略：资助非洲儿童或者购买吸尘器，每天只需花个几分钱。而对于那些拥有余闲的人来说，这几分钱

就像是白捡来的一样。

有些富人同样也会保持节俭的作风。当我们与这些富人聊起自己的工作，他们总会点头附和道：“这就是我的做人风格——对金钱非常专注。”但是，节俭与稀缺的体验并不是一回事。节俭之人对金钱怀有一种保证富余的责任感，而穷人则需要随时进行权衡。发生购买行为时，节俭之人会思考的是，该商品的价格是否“优惠”。相比之下，穷人则会思考，为了付出相应的金钱他们要放弃什么。所以说，节俭之人并不需要进行真正的权衡，他们与所有生活在资源充裕条件下的人一样，很难体会到1元钱的真正意义。因此，他们就需要依赖周遭环境做判断。亚历克斯与三轮出租车司机打交道时就是这样：廉价（而且前后不统一）出售自己的时间，就因为他是根据当时所处的环境来确定三轮出租车的“合理”价格的。可见，亚历克斯很节俭，但并不贫穷。

我们的一位朋友也是行为学专家，他最近花3美元买了一块白兰地松露巧克力。之后，我们问他这块巧克力买得值不值。为了进行判断，他想到了用同样的价钱能买到的其他东西：“6块士力架、一本《体育新闻》杂志，或者晚餐时的一杯高档红酒。”他还想到自己可以将这几元钱节省下来，虽然钱少，但如果其他地方也节省一点，说不定明年就能买一个大一点的公寓。他还想到每个月的卫星电视收费是49美元，而且他最近并不怎么看电视，如果能省下这49美元，就能买好多松露巧克力。最后他不得不承认：“我不知道这块巧克力买得值不值。”所以说，充裕的资源使得我们不那么容易了解金钱的价值。[19](#)

行为经济学所研究的许多偏差与矛盾，其实都是因为人们无法真正体会金钱的价值。在我们的研究中，实验对象无法清晰地理解节省下50美元的意义，于是只能利用背景信息来对之进行评估。相比之下，因为穷人真的面临着50美元的权衡问题，所以早已拥有了对50美元的价值进行内化衡量的尺度（可能并不精确）。因此，穷人比较不容易采取前后不一致的矛盾行为。在知道了这一点之后，我们就可以认为，一定存在着某些情形会令生存于稀缺状态下的穷人对事物的价值进行更加明确的判断，而这种认识是生存于充裕资源中的富人所缺乏的。缺乏清晰价值观会导致可以预见的失误：同样的情况下，富人会犯错误，而穷人却能够避免。

识解

对认知的研究让我们从另一个角度了解了人们是如何对不清晰的价值赋予意义的。在认知过程中，大脑会利用大量的环境线索去理解视觉数据。我们只要了解了大脑所利用的线索，就能予以操控，但有时这种

操控甚至可能会引发反常的结果。由麻省理工学院的泰德·阿德尔森（Ted Adelson）所设计的方块阴影错觉图充分利用了这一知识，也就是我们常用的视觉错觉展示图（见图4-1）。[20](#)

识解（Construal）

在认知过程中，大脑会利用大量的环境线索去理解视觉数据。

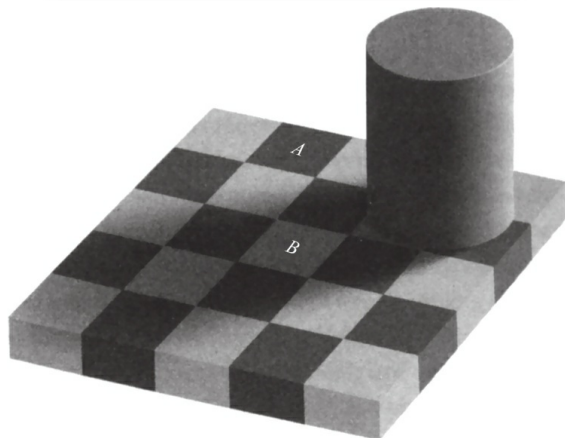


图4-1 视觉错觉展示图

这幅图会给人以错觉，方块A的颜色看起来比方块B要深一些。之所以称之为错觉，是因为方块A和方块B的灰度是完全相同的。你可能不相信，就连我们有时也觉得需要再去仔细查一下两个方块的颜色，因为方块A和方块B看起来真的不一样。如果你不想轻信这个说法，就请拿出一张纸，盖在图上剪两个洞——只露出方块A和方块B。这样你就能看到，这两个方块的颜色完全一致。那么，我们的双眼为什么会遭遇如此的误导？

一般来说，视觉系统会利用图像中的背景线索为物体赋予意义，而背景线索则会影响前景物体在人们视野中的呈现。我们可以看到，方块B的背景与方块A不同：方块B的周围都是颜色更深的方块，而且正好位于圆柱体的阴影之中。因为阴影中的物体看起来颜色更深，所以我们的眼睛就会针对阴影进行修正，让物体看起来颜色更浅。可见，感知颜色与感知距离一样，取决于背景环境。感知价值也是同理。

行为经济学家理查德·塞勒（Richard Thaler）所领导的一次经典实验，堪称这一视觉错觉现象的“金钱”版本。我们和阿努伊·沙阿一起重新进行了这场实验。我们请实验对象思考以下两个场景，它们的区别只

存在于文字——一个场景是超市，另一个场景是度假酒店。

稀缺实验室

SCARCITY

请想象在一个炎热的日子里，你正躺在沙滩上。你能喝到的只有冰水。最后一个小时，你特别想喝一瓶自己最喜欢的冰啤酒。一位朋友提出可以从附近能买到啤酒的地方带酒回来，要么是从一个简陋的小超市，要么从在一家高档的度假酒店。朋友说，啤酒可能有点贵，问你愿意为啤酒付多少钱。他说，如果啤酒的价格与你所说的价格一样或低一些，他就会买；但如果啤酒的价格高于你说的价格，他就不会买。你信任这位朋友，而且也确实不存在与酒保讨价还价的可能性。此时，你会告诉他什么样的价格？[21](#)

与塞勒最初的实验报告一样，富有的实验对象表现出了经典的决策偏差。在高档度假酒店中，他们会为同一款啤酒支付更高的价格。与亚历克斯的行为相同，支付意愿的差异也是一种自相矛盾。一瓶啤酒就是一瓶啤酒（而且他们也会在同样的沙滩上享用同样的啤酒），无论啤酒是来自简陋的小超市还是来自高档的度假酒店，都同样能解渴。而富有的人并不清楚应该支付多少钱，所以只能利用环境来做判断。

穷人的表现就非常不一样了。在两个环境中，他们愿意支付的价格非常接近。关键不在于他们愿意支付价格的高低，而在于他们给出了更加统一的价格。请注意，我们向实验对象询问的并不是他们预期支付的价格。向实验对象提问时，无论是穷人还是富人都会给出同样的答案：高档度假酒店中啤酒的价格一定会更高。两组实验对象的不同之处只在于，他们愿意支付的价格变低了。实验结果与我们的预测相符：穷人对所支付的金钱有着更加清晰的认识，他们不会为环境所动，而是会依赖于自身对金钱价值的内化衡量尺度做出判断。

这样我们就掌握了一个秘诀，可以从中找到办法“推翻”传统行为经济学的结论，也就是那些依赖于从任意局部环境中进行价值分析而得出的结论。在这些结论中，人们会认为自己可以将金钱划分到不同的账户中。

研究发现，当汽油价格上升时，[22](#)人们会以消耗更少的汽油这种方法来进行弥补。我们表现出来的就好像自己变得“更加贫穷了”，虽说多

出来的汽油成本并不一定会对我们总体可支配的资金构成任何实质性的影响。但就算是这样，我们的行为表现也好像是自己在“汽油”上变得更加贫穷了。（想想看，如果问题出在钱上，那么你完全可以通过购买便宜一些的饼干或少去打高尔夫球来省钱。）出现这种情况是因为金钱被我们划分到了具体的账户中：更高的汽油价格对汽油账户形成了负面冲击，导致我们要去节省该账户里的资金，消费更少的汽油。

心理账户的想法有许多含义。举例来说，我们消费2 000美元退税的途径，会与消费2 000美元股票收益的途径非常不同，而原因就在于心理账户。在这两种情况下，我们都多拥有了2 000美元，但我们将这两个账户（“意外之财”和“退休账户”）完全区分了开来，用不公平的眼光去对待它们，因此，对账户内的资金进行消费的倾向也会不一样。相反，穷人则不太会表现出这种行为。[23](#)

机会成本

我们之所以会对物品的价值感到模糊和混淆，是因为我们在资源充裕的情况下是不会进行权衡的，而且我们可能根本不知道应该如何进行权衡。为了对这一观点进行直观的分析，我们请实验对象试想下面这个场景。

稀缺实验室

SCARCITY

你买下了最喜欢球队的季票套餐。[24](#)套餐中包括8场固定场次的比赛门票。单场比赛的门票价格是30美元，而季票套餐只需要160美元，算下来观赏每场比赛你只需要花费20美元。你很喜欢季票套餐中包含的那些场次，于是决定买下来。

现在请想象一下，这一赛季即将结束，你只剩下一场比赛还没看。这最后一场比赛大受球迷追捧，门票价格已经被炒到了75美元。你就要去观看比赛了。现在，请说出你对这场比赛观赏成本的感受。

我们列出了两句话，请实验对象选择，哪个说法更贴近他们对观赛成本的感受：

- 我感觉观看比赛花了我**75**美元，也就是门票的现价，即假设我选择将门票出售时的价格。

- 我感觉观看比赛花了我**20**美元，也就是我购票时所支付的费用。

正确答案是什么？经济学家将75美元视作真正的成本：如果你不去观看比赛，就能卖掉门票，换得75美元。（这里不包括时间权衡。）经济学家称之为机会成本，也就是同样的钱作他用时所进行的权衡。相反，富人答不对这道题，他们会觉得答案是20美元。还有许多人甚至会选择第三种情况：0美元。因为门票的费用早已支付过了。[25](#)我们可以理解为什么富人会有这种感受：当你拥有余闲时，就会觉得0元（或20美元）都没错；有了余闲，你就不会为了观看比赛而放弃任何东西，就算卖掉那张门票，你也不会用换来的钱去买尚未购买的东西。

相比之下，穷人对75美元的用途了解得非常清楚。我们发现，穷人更倾向于认为票价的成本是75美元。可见，他们的答案更接近于经济学的理想情况。

每年，来自世界各地的经济学家都会汇集一处展示自己的研究成果。2005年，经济学家保罗·费拉罗（Paul Ferraro）和劳拉·泰勒（Laura Taylor）决定反客为主，[26](#)他们向在场的200多位职业经济学家提出了与上述问题十分相似的一个问题。可以想见，经济学家们的回答远远脱离了经济学的理想情况。正如经济学家亚历克斯·塔巴罗克（Alex Tabarrok）在博客中写的一样：“我实在不敢相信这是真的，78%的经济学家都答错了！这并不是什么难题，也没有玩花招。机会成本是经济学的核心思想，回答问题的经济学家们是全世界最优秀的经济学者，其中许多人都在大学里教授经济学入门课程，但他们却没有给出正确答案。”[27](#)

世界顶尖的经济学家却不以经济学理想情况作为思维标准，这的确令人有些惊讶。实际上他们都收入颇丰，可支配资源中存在大量余闲，既然他们都不习惯面对日常生活中的小权衡，又怎么会愿意计算小钱的机会成本呢？如果按经济学教科书的标准来看，这些经济学家的回答是错误的；但如果以人们的惯常行为方式作为标准，他们的答案就是正确的。许多富人包括这些经济学家在内，都不会为小钱去做权衡。

我们不妨这样理解：身陷贫穷会让人比职业经济学家更懂经济

学。我们也可以这样总结：如果经济学家的收入降低一些的话，就会更懂经济学。但本书两位作者中至少有一位对这一总结存在异议。

行为经济学的诞生，是因为实证观察发现人们的行为会违反几条基本的经济学预测：人们不会去考虑机会成本；人们购买物品的意愿也会轻易发生动摇。但经济学的本意还是遵从稀缺的内在逻辑的。而经济学的预测，对于那些真正拥有稀缺心态的人来说，会更加准确。

我们并不是说与富人比起来穷人永远会更理性，而是说穷人所拥有的是一种特殊技能：他们更善于维持当下的生计。他们能让1美元发挥更大的价值，他们是金钱价值方面的专家。这种专业技能会令他们在一定的环境中表现得更加理性，更不容易做出自相矛盾的选择。因为专注，所以拥有专业技能；因为专业，所以有了管窥之见；因为管窥，所以产生了一系列负面后果。

第5章 借用与短视

在稀缺状态下，我们会产生管窥心态。当稀缺成为带宽负担时，我们会对当下更加关注，从而导致我们产生借用行为。而当我们借用时，就是给自己的将来挖下了更深的坑。今天的稀缺，将造成明天更大的稀缺。当我们为了解决眼下的难题而极度专注时，就无法有效地规划未来，这样一来，向前看的能力就很可能会因管窥负担而丧失。

为获得生活必需品而不停地艰苦奋斗，这样的日子毫无希望可言。若鼓起勇气展望未来，我们能想到的每一件事情都会让自己气馁。¹

——雅各布·里斯（Jacob Riis）

《另一半人怎样生活》（*How the Other Half Lives*）作者

美国车主贷款责任贷款中心颁布了一份报告，其中讲到了桑德拉·哈里斯（Sandra Harris）的故事。

当桑德拉还是学生时，曾参与了一项针对低收入家庭的“领先”儿童发展项目。²现在，她在新汉诺威小镇的“领先”项目理事会中负责管理工作。2003年，她被北卡罗来纳大学威明顿分校评选为年度优秀员工，还在WMNX电台主持了一档广播节目，算是威明顿当地的名人。但光鲜的外表不能代表全部：她丈夫曾是一名行政主厨，最近却失业了；夫妻俩本来一直能提前一个月支付房租和各项账单，但现在手头却有些紧；他们的汽车保险眼看就要到期了，却没钱续保。

后来，桑德拉想到了一个解决办法：申请工薪日贷款。这个贷款的理念很简单：先通过贷款拿到现金，几周后发薪水时，再偿还贷款，并支付手续费。这正是她所需要的服务。她用贷款续了汽车保险。到了下一个发薪日时，桑德拉就有能力还清这笔小额贷款，并支付50美元的手续费。

工薪日贷款顾问对桑德拉说：“你可以接着贷款。”于是桑德拉想到她还没有支付电费账单，便说：“你说的没错，我的确有这个需要。”

桑德拉就这样拉起了一条贷款链条。第二个月也不比第一个月轻松。她手头可以周转的钱变得更少了。因为需要支付手续费，所以她签下的贷款数额变得越来越大。接下来的几个月中，她不断去贷款，然后再用新的贷款来偿付之前的贷款。有的月份，她甚至连手续费都支付不起。

几次延期还款后，第一位借款商提出了贷款全额偿付的要求。桑德拉没有能力偿还，于是找到了另一家工薪日贷款商“急钱服务公司”，贷了一笔钱，以偿还第一位贷款商的贷款。就这样，她越陷越深。短短6个月时间，桑德拉就要同时为6笔不同的工薪日贷款支付滞纳金。2003年6月，桑德拉和她丈夫差点就被房东从他们居住了6年的公寓中驱逐出去。桑德拉写道：“我们就是在用一笔贷款偿还另一笔贷款，后来手续费涨到了每个月495~600美元，我们就只能支付手续费，再没有能力偿还本金了。”

这种情况一直持续了至少6个月时间。这些钱并不能让夫妻俩过上奢侈的生活。桑德拉说：“人们总觉得你是挥霍无度才去借钱的。”但她并没有用这些钱买衣服，而是在支付账单。一直以来，桑德拉都勤勤恳恳，在艰难的日子里努力管理着家庭的收支。

后来，桑德拉的支票开始被退回，汽车也被收回了。她提高了免税额度，这样就能有多一点的钱去支付账单，结果却欠下了数千美元的退缴税。最后，她破产了。在广播电台值班的时候，她总在休息时悄悄擦眼泪。

“我不是一个轻易掉眼泪的人。”桑德拉说。

调查数据显示，桑德拉的情况非常典型。2006年，全美共有23 000多家工薪日贷款的分支机构，3比全球麦当劳（12 000家）4和星巴克（差不多9 000家）5加在一起的分店还要多。桑德拉延期偿付贷款、积压手续费的这类行为也十分常见。所有工薪日贷款的3/4都来自延期偿付，最终形成了每年高达35亿美元的手续费。6

为什么那些缺钱的人会选择他们根本偿还不起的极端贷款类型？为什么他们一开始就放任自己走上这条不归路？这类问题一不小心就会演变成一场辩论，不是关于个人责任的重要性的讨论，就是关于商家如何不择手段地将低收入人群作为捕猎目标，随后引发出关于穷人目光短浅和财务知识教育必要性的讨论。支持消费者的一方会为工薪日贷款行业的存在而痛心疾首，他们认为这一行业掠夺成性，并敦促相关机构叫停这些贷款。还有一些人认为，如果真的有需求，如果真的能贷到款，无论利息多高，总比没有要好。我们讲桑德拉的例子，并不是因为我们想参与这场激烈的讨论，而是因为这种现实与稀缺有着重要关联。

问题不仅仅在于工薪日贷款。缺钱的人会想尽各种方法借钱，不仅限于工薪日贷款。他们通过迟还账单来“借钱”。在收入最低的20%的家庭中，6户人家中就有1户会在任意一年至少迟还一份账单。这种情况的极端表现就是重装费。一项研究发现，最贫穷的家庭中有18%都曾有过电话停机的经历，10%曾在12个月内停水停电。7因无法支付账单而被停机，之后为了重新开通电话服务而缴纳了40美元的费用，就像一开始为了在拿到薪水之前紧急周转而贷款，并为此缴纳40美元手续费一样。1997年进行的一项研究认为，在穷人的年收入中，近5%都用在了重装费、滞纳金等费用上。8而我们估计，与当时相比，现在这一费用的比例早已出现了大幅提升。桑德拉·哈里斯也“借钱”了，一开始是通过减少扣缴税款，后来就演变成了迟交税款。世界各地的穷人都会借钱，通常都是向那些非正规的放贷机构借钱，而这些机构的贷款利息和工薪日贷款商的一样高，有时甚至更高。9尽管如此，穷人们依旧要支付这一利息，且不只一次，而是持续地支付，他们就这样踏上了滚雪球般的债务不归路。

这种现象不仅限于贫困人群。工作繁忙的人同样会借用时间，而借用时间的“利息”也同样高昂。为了尽快完成迫近截止日期的工作，他们会放下手中的其他工作。就像工薪日贷款一样，账单迟早都会来，延迟的工作也迟早都要完成。而且，借用时间也要支付“费用”：将工作暂时放下，会增加完成这项工作所需的时间。用挂号信邮寄纳税申报单，只

要几分钟时间，但如果是最后一天，邮局门口就会排起长龙。因为截止日期即将到来，你就没时间将手写的会议笔记输入计算机里。之后，你必须重新研究这份手写笔记，而所花费的时间要比会议刚刚结束、头脑清晰时多出许多。就像工薪日贷款人一样，繁忙之人也会将他们的债务雪球越滚越大。一些本来今天要做的事情必须先放一放，因为你还有一些本应昨天完成的事情。多少工作会在最后完成之前一遍又一遍地被延期？而且每次延期的原因都大同小异：下次准备着手做事时，你会发现时间并不比之前多。

借用 (Borrowing)

当人们面临资源稀缺时，就会通过借用相应的时间或金钱来应对突发事件。从长远来看，借用会进一步加剧稀缺。

借用与稀缺同时存在。

管窥与借用

在稀缺状态下，我们为何会产生借用行为？之所以去借用，是因为我们有了管窥心态。而当我们借用时，就是给自己的将来挖下了更深的坑。也就是说，今天的稀缺将造就明天更大的稀缺。

以桑德拉为例，她第一份没有能力偿还的账单，最终产生了稀缺。之后，她有了管窥心态，一心想着当月如何才能收支相抵。在这种心态下，工薪日贷款就有着巨大的吸引力。贷款的好处落在了“管子”之中，桑德拉看到贷款能帮助她渡过这个月的难关。而贷款的成本，也就是还款和手续费，都落在了“管子”之外。贷款看似好像能为她纠结的问题给出解决方案。

管窥之见会令工薪日贷款显得尤为吸引人。我们进行的量化实验也支持同一观点。如果你问贷款人打算如何偿还债务，总能得到这样草率的答复：“没事，我一周之后就能领工资了。”如果你继续盘问：“到你没有其他支出吗？”当事人就会怒火中烧：“你怎么还不明白？我这个月必须付房租啊！”言外之意就是：“我正专注于现在需要做的事情。下个月的钱从哪里来，是一个抽象概念，之后再想办法也来得及。”这就

像你要着急赶赴医院时，所有的重要目标都失去了意义一样，工薪日贷款所引起的长期财务问题，在当时看来也并不重要。这就是为什么工薪日贷款如此富有吸引力的原因——当人们用管窥心态盯着眼下的燃眉之急时，就会求助于工薪日贷款。而工薪日贷款最好的一点就是，它能够快速有效地解决人们的燃眉之急。而这场大火迟早都要烧起来，以后很可能会摧毁更多东西。但是，工薪日贷款所导致的恶劣后果则往往会被人忽视。

这种现象并不仅限于工薪日贷款或与金钱有关的问题。请想象你现在有一封邮件迟迟未回。当我们借下这笔时间债时，一直在专注于借债的好处：“现在我有其他工作要完成。”而没有停下来自我反省：“我之后从哪里找时间回邮件？”这么做并不是说我们无视这种行为的成本，只不过是成本得不到我们的关注罢了。

关于我们管窥的事物，这里有一个重要的隐喻：桑德拉今天缺钱，下个月照样会缺钱；日理万机的人这星期很忙，下星期依旧会很忙。处在稀缺状态下的人，不仅现在有这样的体验，一般之后的生活也会持续有这样的体验。但人们还是会管窥他们当下的稀缺，就算你知道下个月会饿肚子，这种认识也不会像今天这样饿肚子让你全神贯注。即刻到期的账单摆在眼前，会激发你的危机感，但两个月之后到期的账单则完全不会引起你的重视。就算你能细心地想到明天将要面对的稀缺，但也只能抽象地“知道”其存在，却无法切实地感知、体会。因此，明天的稀缺注定无法像今天的稀缺一样俘获你的注意力。其中一个原因就是带宽负担——当下的一切会自动加载在你身上，但未来却不然。若想关注未来，则需要带宽，而稀缺已经占据了带宽中很大的一部分。当稀缺成为带宽负担时，我们会对当下更加关注。我们需要通过对资源的认知来估测未来的需求，需要利用执行控制力来抵御当下的诱惑。在稀缺成为带宽负担时，会让我们关注当下，从而导致我们产生借用行为。

我们已经了解了支持这一假设的相关数据。请回忆第1章中有关截止日期的研究：一组学生有3周时间来完成任务，而另一组学生每周都要面临一个截止日期。最后的结果是，第二组学生的表现更佳。我们将这一优势归因于专注红利。但第一组学生也要面对截止日期，只不过是在3周之后，而不是每周一次。我们从中了解到，3周之后的截止日期并不能给人以紧迫感。事实上，在实验开始时，每周一次的截止日期也不会让人感觉到压力重重。但我们可以猜测之后发生的事情：只有截止日期即将到来时，这个日期的重要性才能得以体现。在此之前，它只不过

是一个抽象的概念而已，无法激起学生们的稀缺心态。对于那些面对一周一次截止日期的学生来说，3周就会产生3次稀缺心态；而对于在3周时间结束时才面对截止日期的学生来说，仅会产生1次稀缺心态。我们所有人对此都不陌生。这就是为什么我们会在截止日期即将到来时发现自己的工作效率会出现大幅提升的原因。

以这种方式形成的管窥之见，就会造成针对借用的偏差。只有最迫切的稀缺才能进入“管子”视野，也正是因此，工薪日贷款才显得尤为富有吸引力。[10](#)

当然，贷款本身并不一定是一个错误的选择。如果你下周会拥有更多的时间，那么现在将手头上的其他事情放一放，以尽快完成当务之急，也是明智的做法。在面临被房东驱逐的情况时，如果马上就能领到工资，那么借钱交房租也是可行的。

当今天的资源，无论是时间还是金钱，能在当下为你带来比未来更大的利益时，借用就是行得通的做法。而当我们有了管窥之见时，借用就会超越得失均衡的限度。当面临稀缺状况时，我们的借用，从长期来看通常是不合理的。

《家庭问答》的启示

这一关于借用的解释，与寻常的说法有所不同。若想说明穷人为什么会过度借贷，其实并不需要在财务教育、掠夺成性的贪婪借款商或自我放纵等方面找原因。若想解释忙碌之人为什么会将事情一拖再拖，工作进度一直落后，也不需要自控能力弱、理解能力低或时间管理能力不足等方面找原因。借用，不过是管窥心态所引发的后果而已。为了对这一理论进行测试，我们又用上了最得心应手的工具：在实验室中建立人造稀缺。

这一次，我们将目标瞄准了美国一档电视游戏节目《家庭问答》（*Family Feud*）。我们的同事阿努伊·沙阿是这档节目的忠实粉丝。（对于当时还在普林斯顿大学攻读博士学位的阿努伊来说，时间异常宝贵，而他竟然会有看电视的时间，实在让人有点摸不着头脑。）

《家庭问答》的选手要说出哪些东西属于特定的物品类别，比如“如果芭比娃娃急需钱，就会将这件东西拍卖”等。节目播出前，节目组会随机挑选100位美国人，让他们看这些分类，并给出自己心中的最佳答案。选手要猜出最为普遍的答案，答案越是为人所熟知，就越能拿到更高的分数。回答“芭比娃娃的梦想汽车”能获得35分，因

为在100个人中，35个人给出了这个答案；回答“芭比娃娃的朋友‘肯’”能获得21分。

许多竞猜类节目都会问一些看似微不足道的问题，若想给出这些问题的答案，却需要挑战者掌握深奥的知识。因此，看节目的观众总会忍不住猜想：这些选手是不是以阅读年鉴为乐？相比之下，《家庭问答》提出的问题完全是普通人所能理解的，而且非常引人入胜，因为这些问题没有正确答案，只有更为人所熟知的答案。这个节目将真理变得更加民主化，我们不妨称之为第一部后现代电视节目。

阿努伊意识到，《家庭问答》的参赛选手会体验到稀缺：他们必须在有限的时间内给出答案，这样一来，用来思考的时间就会变得非常少。普通的知识竞赛需要选手在记忆中搜寻答案，要么知道，要么不知道；而在《家庭问答》中，回答问题需要完全不同、更富创意的方法。当问题是“给出芭比娃娃能出售的物品名称”时，就要在脑海中分类整理各种不同的答案。你可能会想到与芭比娃娃有关的东西，想一想这些东西是否能出售；你也可能会想到人们一般会出售的物品，想一想芭比娃娃是否拥有这些物品。每一种思路都会得出不同答案，可能是“肯”，也可能是“汽车”。这些答案纯属猜测，每个答案的普及程度还需要进一步分析。时间上的压力，意味着选手没有办法拥有太多思路，而分析每个可能答案的时间就变得更少了。忙碌之人用几个小时或几天为单位来衡量稀缺，而《家庭问答》中的选手则以几秒钟为单位，选手没有时间思考要先做哪项工作，必须迅速决定如何给出他们认为最普遍的答案。

我们招募了普林斯顿大学的一些本科生，在受控环境下玩起了《家庭问答》游戏。[11](#)参与者在固定时间内进行几轮游戏，而分配给他们的时间决定了他们的“财富”：“富人”时间多，“穷人”时间少。每一轮都有一个新问题。在几轮游戏结束后，他们得到的总分会被兑换成美元作为奖励。

在将实验对象分为“贫、富”两组之后，我们在游戏中加入了自己真正感兴趣的元素：给他们一个选择，可以借时间，但需要支付利息。如果参与者选择在一轮游戏中增加一秒钟时间，就会在总时间中减去两秒钟。我们也会给他们“存储时间”的机会，如果他们提早完成了一轮游戏，剩余时间就可以被存入总时间内。

“穷人”十分专注。他们每一秒钟都会比“富人”更有成效，他们猜测的次数更多，赚得的分数也更高。尤其是到了后面几轮，时间快要用完的时候，“穷人”平均每秒钟的猜测次数比“富人”多50%，每次猜测也能

获得更高的分数。如果“富人”能像“穷人”一样专注，他们就能赚取多得多的分数。我们给了“富人”3倍的秒数，他们完全可以玩3倍的轮数，赚3倍的分数。但事实上，他们的分数只是“穷人”的1.5倍。后来经分析证实，我们所有能想到的其他原因都无法对这一结果进行解释，包括“富人”因为玩的时间长而觉得无聊以及最佳猜测出现在刚开始玩的时候等。

“穷人”之所以更富成效，是因为他们拥有管窥心态。所以，他们比“富人”借用的更多。虽然利息很高，但从管窥心态来看，他们会觉得贷款极富吸引力，其吸引力要比普通眼光看起来大得多。因此，“穷人”常常会采取借用手段，为当下寻求帮助。但最终，他们会反受其害。这就是管窥心态的负面作用。为了证明这一负面作用，我们在新一轮实验中取消了借用功能——玩家只能竭尽所能地玩好当前的一轮游戏，之后再继续下一轮。结果，穷人比上次能借用秒数的时候多赚到了60%的分数，而“富人”的得分则没有受到什么影响。

在这一实验的另一个版本中，我们重建了与桑德拉的体验非常类似的工薪日贷款陷阱。《家庭问答》中的“穷人”就像工薪日借贷人一样，将债务雪球越滚越大。欠下的时间债从下一轮开始偿还，使得下一轮游戏的时间比这一轮要稍短一些。就这样，每轮游戏的时间越缩越短，玩家们越发感觉到了借用更多秒数的需要。从游戏之初就开始借用的“穷人”，就这样陷入了恶性循环。因为时间紧迫，他们来不及给出什么高明的答案，于是就只能去借用更多的时间。而且，他们绝大部分的时间都用来偿还之前借用的时间和利息。与之前一样，在拥有借用能力时，“穷人”的成绩会差许多，根本无法与没有借用能力时的成绩相提并论；而“富人”在借与不借时的发挥都不会受影响。

这一研究显示的是稀缺状态下成败之间的紧密联系。《家庭问答》的参赛选手在成效最高、精神最投入、感觉最需要更多时间时，会借用最多的时间。从某个角度讲，他们的借用行为没有错，因为他们之后有能力去偿还现在多出来的时间；但从另一个角度讲，他们的借用行为却是错误的，因为偿还借用时间的同时，他们还需要偿还利息。他们用管窥心态注意到的利益，是现在真的需要那多出来的几秒钟时间，所以借用是正确的。他们的错误在于忽略了“管子”视野之外的事物——现在多用几秒钟，会让我在接下来的游戏中付出什么样的代价？值得注意的是，“穷人”和“富人”在非常有成效和感觉时间紧迫时，都会采取借用手段。只不过因为“穷人”拥有的秒数少，所以会比“富人”更加频繁地借用

罢了。

为什么“穷人”会更加频繁地借用呢？这是由于管窥心态使然，还是有其他影响因素？也许时间压力会导致人们在恐慌之下采取借用手段。毕竟，没人会每天都面临15秒钟内给出答案的紧张状况。我们在其他一些环境中对这些结论进行了复制。在第1章中讲到的《愤怒的蓝莓》游戏中，我们也赋予了实验对象借用的能力。我们发现，“蓝莓穷光蛋”们虽然没有时间压力，但也借用了更多蓝莓，并因而反受其害。不过，专注也发挥了一定的作用：每次射击所用时间更长的人，借用的可能性更高。越是投入，借用的越多。我们在许多类似的游戏中都尝试过这种实验，得到的结果都一样：无论稀缺以什么形式呈现，都会导致借用。

抑或，我们的结论源于普遍的短视。举例来说，研究人员发现，人们会就“此时此刻”怀有偏差，并将这种现象称为“双曲贴现”（Hyperbolic Discounting）或“现时偏差”。¹²人们会将未来的利益作为代价，过高地估计即刻的利益，这就是为什么坚持储蓄、去健身房锻炼、提早报税存在难度的原因。现时偏差也会引起借用。穷人之所以会借用，也许就是因为他们有着更强的现时偏差。事实上，有人曾利用这一理论来解释现实世界中的实际借贷行为。在我们进行的研究中，实验对象是被随机分配成“穷人”的，他们与实验中“富人”的唯一区别就在于，掷硬币时一个是正面，一个是反面。研究中的贫富两组应该表现出同等的现时偏差。事实上，若想在贫富两组的个体差异层面对短视思维进行分析，无论是分析现时偏差还是分析其他因素，都需要对稀缺如何在当下的环境中导致借用这个问题进行解释。而在这个环境中，“穷人”与“富人”是随机分配的，彼此之间并无实质性的差异。

现时偏差（Present Bias）

人们会将未来的利益作为代价，
过高地估计即刻的利益。

这些研究也支持了我们关于世界的一个更普遍的假设：穷人之所以会产生借用行为，是因为贫穷本身造成的。所以，我们不必在短视、财务知识缺乏等因素上寻找解释。掠夺成性的贷款商无疑在这类借贷上起了煽风点火的作用，但他们并不是问题的根源。想要借钱的强烈冲动与对滚雪球般高利率恶性循环式贷款的需求，是管窥心态的直接后果。

稀缺致使我们借用，并且让我们在稀缺的泥潭中越陷越深。

忽视未来

请想象一项工作的截止日期马上就要到了。在反复思考了几周之后，仿佛一瞬间，这份报告就得交了。交稿日期是第二天，而你根本没有做好准备。你施展开浑身解数，通宵工作，但就是有几本参考书目找不到——至少第二天到来之前是找不到了。于是，你只好破罐破摔地将报告交给了老板，暗自祈祷不要露出马脚。随后的一周，就在你准备出差之前的几个小时，你收到了老板的邮件：“报告里缺少参考书目。我立刻就要！”上周匆忙中的权宜之计像回旋镖一样，在最不恰当的时刻飞回了你的手中。与借用一样，这类行为从“管子”中窥视时会非常吸引人，但从“管子”外来看，却很有可能会导致恶性循环式的后果，将你拉进稀缺的泥潭。

两位组织研究专家用钢丝制造商的故事来说明了这一点。

机器的运行时间非常重要，公司鼓励维修工程师对故障尽快采取措施。尽管如此，总体业绩还是没有得到提高。后来，公司改变了以人为单位的做法，开始以机器为单位进行日常工作的记录和分析，才意识到问题出在了哪里。工程师会在匆忙中随便修理，然后就赶赴下一台机器。每次机器故障都至少要维修3次才行。[13](#)

从某种角度讲，公司要求维修工程师迅速解决问题，而维修工程师的做法完全满足了这一要求。你可能会认为，管理层犯下了一个经典错误——用组织研究专家的话说，管理层就是在“幻想种瓜得豆”。他们对速度提出要求，但又同时要求质量。这并非是一个对员工激励不足的问题。案例中的员工，就算自己是老板，也会同样在匆忙间解决问题。在“管子”视野中，匆忙维修是他们需要做的事情，而偷工减料就是完美的解决办法，至于为此要付出的代价就只有到以后再考虑了。与高息贷款一样，从“管子”视野中看出去，匆忙维修的解决办法非常具有吸引力。这种办法会对未来产生更严重的影响，却能为现在节省一些资源。而到了未来，需要我们去做的事情会变得更多，需要我们解决的问题会变得更多，需要我们支付的账单会变得更多。匆忙修补与借用有许多相似之处，都是在当下不能投入足够的资源去解决问题。

缺钱的人也会将短期解决办法匆忙地拼凑在一起。你需要一台洗碗机，可是手头的钱不够？那就买最便宜的一款好了。便宜当然没好货，但这就是“管子”视野之外的问题了。遇到轮胎漏气的情况时，你可能会找家便宜的汽车维修站随便补补，而不是换一个新胎。你知道补过的轮

胎肯定没有新的那样安全、耐用，而补胎的做法也并不明智。但眼下从“管子”视野中看出去，补胎更省钱省事。就像能为眼下节省出时间的权宜之计一样，所有这类权宜之计都能为眼下省出一些钱。随着补丁越积越多，那些工程师、报告者、穷人们的长期成本也会变得越来越高的。

史蒂芬·柯维（Steven Covey）认为，将任务根据重要性和紧迫度进行分类，可以有所助益。¹⁴他认为，繁忙之人会将时间用在既紧急又重要的事务上，这就是在截止日期前埋头赶工的意思。当我们为了紧急而重要的事务辛勤工作时，工作成效就会突飞猛进。我们称之为专注红利。

同时柯维也认为，繁忙之人此时会容易忽略那些重要但并不紧急的事务，即那些永远都能往后推的事情。就这样，我们一推再推。从办公室和家中的整洁程度，就能清楚地了解这一点。当我们忙得不可开交时，办公室和家里就会变得凌乱不堪。总有一些事情比整理东西更紧迫，而整理东西这件事情则永远不会紧急到火烧眉毛的程度。

我们并不是有意选择生活在一片狼藉之中，而是在我们忙于处理紧急事务时，周围的一切就这样凌乱了起来。乱七八糟的家庭或办公环境是一系列小选择的结果，而这些选择绝大多数是被动的、容易的，是不为我们所留意的。比如，你急着赶赴一场会议，于是就将一摞信件随手放在了另一摞文件上面；为了接电话，你随手将看到一半的书扔在了沙发上。诸如此类的许多小事，最终让你的环境乱七八糟。虽然这不是什么紧要的事情，但却十分重要。在凌乱的环境中生活和工作，你一定不会很开心，而你的工作效率也一定不会很高。

将一件重要但不紧急的事务搁置一旁，与借用无异。你通过不去做这件事情，而为今天赚得了时间，但却增加了未来的成本，因为你早晚要借用其他时间（很可能是更多时间）来处理这件事。与此同时，你也许会因为没做这件事情而付出成本，或损失这件事情可能带来的收益。办公环境混乱，一定会令你的工作效率有所下降。你会花上许多时间，去寻找压在信件下面的那摞文件。每天，你都会给未来增加一点成本。这些成本永远不会大到将此事升级为紧急要务的程度，永远不会像截止日期那样奏效。被你忽略的办公室环境，会像小刀一样，一刀接一刀，让你遍体鳞伤，流血不止。

稀缺，尤其是管窥心态，会让你搁置那些重要但不紧急的事务，比如整理办公室、做结肠镜检查、写遗嘱等，这些都是我们容易忽略

的事情。做这些事情的成本是立即显现的，用管窥心态来看的确十分高昂，而且这些事情可以轻松地往后推，其好处都落在了“管子”的视野之外。于是，人们会等到所有紧急事务都处理完毕后才会计做这些事。就算现在做这些事情能为今后带来很大的好处，人们也常常不愿意在现在就付出这些小小的努力。

将重要却不紧急的事务搁置，这种倾向不仅有关于时间，也有关于金钱。印度的拾荒者在城中四处寻找可以出售、重新利用的旧衣服和旧布条。可以想见，这份工作的收入很低——拾荒者一般每天的人均收入还不到1美元。同时，这也是一份低投入的工作，除了劳动力之外，唯一的设备就是一辆售价为30美元的手推车。而大部分拾荒者都没有自己的手推车，他们每个月只能花5~10美元去租赁手推车。许多拾荒者都希望能省下钱来买一辆属于自己的手推车，但很少有人能做到。

在手推车上投资，对于拾荒者来说，是重要但不紧急的事务。就像整理办公室可以给今后带来好处一样，这种事情总是可以再等一等。讽刺的是，如果拾荒者拥有了属于自己的手推车，租赁手推车的日常支出就会节省下来，也就意味着他不用那么艰难地去应对其他紧急的费用问题。而办公室整理也一样，如果将办公室整理干净，你就能省出时间，用不着凡事都那么匆忙。（也就有了更多整理办公室的时间。）手推车的案例，是贫穷问题研究专家经常谈及的例子之一：就算购买手推车的回报很高，但最需要这些回报的穷人却总是没有能力攒钱购买，而这一现象是贫乏的财务知识或技能所无法解释的。

如果你觉得这些理论似曾相识，那么很可能是因为你在政治讨论中听到过类似的言论。在政府的工作中，工作人员经常会为了处理紧急事务而牺牲重要事务；在几十年的财政紧缩过程中，政府常常会削减基础设施建设的支出。举例来说，桥梁维修是一项事关重大的投资，而在预算紧张，需要进行经费裁减时，政府总会将桥梁维修暂时搁置。这是因为桥梁的老化问题很重要，但并不紧急。美国土木工程师学会2009年公布的一份报告显示，美国约有1/4的郊外桥梁和1/3的城市桥梁存在缺陷。[15](#)

无力规划

所有这些不同行为都有着同样的明显特征：短视。这就引出了管窥心态最基本的含义：当我们为了解决眼下的难题而极度专注时，就无

法有效地规划未来。当然，研究也显示，规划对于所有人来说都是一个**问题**，而稀缺则令这一问题变得更为严峻。[16](#)

某一天你可能感觉还不错，于是就看了看日历，花了点时间计划了当天需要做的事情，可能还想了想整个星期需要完成的事项。对未来将要发生的事项心知肚明，可以让你从心理上有所准备，对富有挑战性的对话有所预期，还可以自我提醒一些有关事务的细节，不至于开会时跟不上进度。而相较之下，如果这一天你特别忙，就会不假思索地直接投入工作。你不会退后一步，思量一下一整天的安排；你不会清楚与会**人员**都有哪些，会议会讲什么内容。这不仅仅是由时间紧迫所引起的。你可能会有一点时间投入工作，但心思却全都放在了那些需要你完成的事务上，以至于你的视野开始变得模糊。你已经没有能力考虑几次会议之后的事情了。

退一步讲，从当下脱离出来向前看，你需要具备宽广的视野和一定的认知资源。你需要考虑下个月到期的账单、其他可能的收入来源和有可能接手的新任务，所有这些都需要你的认知能力拥有一定的余闲。当大脑专注于当下的稀缺时，向前看的能力就很可能因管窥负担的存在而丧失。

我们是否也能在《家庭问答》游戏中对此进行重建？与以往一样，实验对象要玩几轮游戏。同样，一部分是“富人”（每一轮都有许多秒钟时间），另一部分是“穷人”（每一轮只有几秒钟时间）。但这次，我们为实验对象提供了“向前看”的选项，即为接下来的游戏做准备的机会：其中一半的人会看到下一轮问题的预览，他们可以在思考当前问题的同时，去思考下一个问题；他们也可以在看到问题后，决定是节约时间还是借用时间，因为他们可以考虑自己在下一个问题上会花更少的时间还是更多的时间。

预览的确会有所帮助。说得更准确些，是会对“富人”有所帮助。他们向前看，利用提前给出的信息，获得了更多的分数。而“穷人”就算在有预览的情况下，也不会表现得更好——他们太过专注于当前进行的这轮游戏，无法投入向前看所需要的心智资源。可见，稀缺会令“穷人”拘泥于当下，以致他们不能从未来的信息中获益。

许多形式的管窥负担都拥有相同的内在逻辑。稀缺所引发的行为会令我们产生短视现象：在忙碌不堪时，我们会忽略在外用餐可能给自己造成健康负担；当手头缺钱时，我们想不到将来偿还工薪日贷款的难度；在截止日期到来之前忙着赶工时，我们不会想到保持办公室整洁能给自己将来的工作带来好处。当然，例外总会存在，总有一些事情在任何时候都会萦绕于我们心头。你心里可能会一直惦记着一年之后将要举办的婚礼，而忘记了当天的会议。这就是人类心灵的有趣之处。但总的

来说，稀缺问题会给当下的我们带来压力。明天，我们可能依然贫穷（无论是时间上还是金钱上），但那是明天的问题，可以留到明天再去解决。俘获我们的稀缺，就存在于当下，它所产生的管窥负担，令我们带着短视的眼光做人做事。

从这个角度讲，短视并非个人问题，管窥也并非由个人性格所导致。毕竟，如果我们说桑德拉是短视之人，也是不妥当的——她从“领先”项目中脱颖而出，还是大学里的年度优秀员工，同时也是“领先”项目的理事会成员。同样，我们也不能说那些日理万机的人是短视的——参与我们实验室研究的学生们，不可能带着短视的大脑考进普林斯顿大学；许多借用时间的繁忙人士，都在竞争激烈的职业发展上投入了多年的努力，也对如何力争上游进行了精心的规划。事实上，从个人性格上讲，这些人远非短视，而是在稀缺环境下，都表现出了短视的行为。

可见，管窥心态会限制每一个人的视野。

第6章 稀缺陷阱

稀缺陷阱就如同“杂耍”，太多的“最后关头”让人们持续从一项紧要任务转移到另一项紧要任务。这是管窥所引发的后果。我们可以在当下尽己所能，但这样的做法会给未来带来新问题。稀缺陷阱源于人们将可预期事件当作突如其来的事件处理。避免落入稀缺陷阱的唯一方法就是要拥有余闲，尤其是要建立起应对突发事件的缓冲机制。改变心态，才是人们逃离稀缺陷阱的唯一希望。

如果你有时间，每个地方都会在步行距离之内。¹

——史蒂夫·赖特（Steven Wright）

印度金奈的克延比都（Koyambedu）蔬菜市场总是一派热闹繁忙的景象。它占地约243亩，坐拥2 500家店铺，销售的商品从芒果到金盏花，应有尽有。熙熙攘攘的数万名顾客攒动在各色商品间，仿佛高峰时段的地铁。吸引眼球的东西很多，但最有意思的事情，也是最容易为人所忽视的。

在黎明前的几个小时，小贩们就会来到市场。如果你去过世界上任何地方的贫穷城市，就很可能见过这类街头小贩，也很可能从他们那里买过东西。在金奈，小贩们会坐在路边，有的摆个货摊，但大多数人只有一张毯子，兜售着蔬菜、水果或鲜花。他们的商业模式很简单：小贩们一般会在清晨购买价值为1 000卢比（20美元）的货物；²当天卖完之后，收入1 100卢比，总利润100卢比（2美元多一点）。³这种买卖需要两类投入：小贩们的劳动力和每天用来进货的1 000卢比。有些小贩拥有1 000卢比，而大部分（我们得到的数据是65%）小贩得去借钱。这笔贷款一点也不便宜，每天的利息是5%。换句话说，小贩在卖完东西后，所赚的100卢比总利润中要拿出一半来支付利息。小贩们为贷款支付的利息，可能是克延比都市场上最有意思的事。

你可能会觉得，只有经济学家才会用“有意思”这个说法来形容利息，但让我们先来了解一下这个故事的原委。

在每位小贩的可支配金钱中，都有一小笔余闲，可以用来消费。他可以去买杯茶喝，买点零食吃，也可以给儿子或孙子买些糖果。假设小贩们没有将这笔小钱花掉，而是每天都省下5卢比，就可以用这5卢比进货。这样，他们每天就可以少借5卢比。如此看来，经过200天的努力，他们就可以免于1 000卢比的借贷。而事实上，只消50天就可以实现这个目标。这就是复利的作用（尤其是在利率很高时），每天5%的利率会以很快的速度积累复利。

这样做的意义十分重大。如果每天都能节省一点钱，50天之后，小贩们就会无债一身轻。没有债务之后，小贩们每天的收入就能翻番。如果在不到两个月的时间内，就能让穷人的收入翻番，这样的社会福利项目简直美好得让人难以置信。虽然每位小贩都能享受“项目”带来的好处，但却无法加以利用，而且对借贷依然执迷不悟。在我们提取的样本中，典型的街头小贩已经借贷达九年半之久。

小贩们落入了稀缺陷阱。值得思考的是，小贩们是怎样落入这个陷阱的？我们总是认为，稀缺就是摊派到自己身上的现实。在某些情况

下，这种认识是没错的。生活在发展中国家，每天挣1美元的人，与生活在发达国家，每天挣100美元的人相比，其收入差别的原因不在于行为差异上，而在于出生的地理位置上。但某些稀缺的确可以在行为上找原因，比如印度的街头小贩们：如果他们采取了不一样的行为，就可以不那么贫穷。

小贩们的处境就是我们所称的稀缺陷阱。稀缺陷阱，就是某人的行为有助于稀缺形成的一种情况。处于稀缺陷阱中的人，就像小贩们一样，可能会继承一些稀缺元素，而这些元素其实早已远远超出了他们的控制范畴。如果小贩们生于纽约，就会富裕得多。而令我们感兴趣的是，由我们自身的行为而导致的稀缺。让我们更感兴趣的是，稀缺是如何激发出这类行为的，它是如何经久不衰的，而我们又是如何在稀缺心态中通过行为来使其进一步强化的。

稀缺陷阱 (Scarcity Trap)

就是某人的行为有助于稀缺形成的一种情况。

请想象有两位学生，一位名叫菲力克斯，一位名叫奥斯卡。菲力克斯每个周末都有一份作业要交，于是花费了许多时间写作业，也能按时交作业。他很忙，但心态很放松。奥斯卡也很有天赋，与菲力克斯同班，却总觉得时间不够用。他用来写作业的时间比菲力克斯多，感觉心烦意乱。他每周都匆忙赶着交作业，还经常迟交。那么，什么让奥斯卡变得更加忙碌？事实上，他并没有选修更多的课程，也不是一个效率低下的人。只不过他比菲力克斯晚了一步：他在做的永远是上一周的作业。菲力克斯在写作业时，记忆中的相关内容十分清晰，因为他刚刚听完课；而奥斯卡总要多花时间去回忆上周课堂上讲过的内容，还要将上周的内容与当天早上听过的课程区分开来。虽然奥斯卡更加忙碌、更加奋力地写着作业，却没能完成更多的作业。

金钱上，我们也可能会落后一步。试想，菲力克斯和奥斯卡两人都是农民，拥有同等面积的土地，每一季都种植同样的作物。菲力克斯用存款去购买种子和肥料，在收获时节到来之前，还能用存款养活一家老小；而奥斯卡为了购买种子和肥料以及养活一家老小，不得不去借钱。与心态放松的学生菲力克斯一样，农民菲力克斯在经济上也显得更加富裕；而奥斯卡能用来支出的钱则更少。虽然菲力克斯和奥斯卡的收入相同，但后者的一部分收入却用来支付贷款利息。同样，农民奥斯卡的问题也在于他落后了一步：菲力克斯的收入用来为下一季度投资，而他的收入却用来偿还上一季度的贷款。

可见，稀缺并不仅仅有关于实物资源。在两个案例中，菲力克斯和奥斯卡都拥有同样的资源，而在奥斯卡体会到稀缺的同时，菲力克斯

却没有。在第一个案例中，菲力克斯和奥斯卡的作业量和时间是相同的；在第二个案例中，他们的土地面积和收入也是相同的。之所以最后的结果会出现不同，是因为他们使用资源的方式不同。

菲力克斯和奥斯卡之间的对比，使稀缺陷阱的概念变得清晰起来。两人都面对着明确的约束，但奥斯卡却因为自身的行为而陷入了稀缺中不能自拔。更宽泛地说，稀缺陷阱并不限于实物资源的短缺。它的根源在于对资源的错误利用，从而出现了实际的短缺。稀缺陷阱就是一直在落后一步，一直在偿还上个月的支出。用这种方法管理并使用资源会让我们感觉所拥有的资源变得更少了。对稀缺进行放大的行为，就像复利一样，会使最初的稀缺变本加厉。[4](#)

我们在审视稀缺时，总是会忽略这一特征。当我们看到农民奥斯卡一直在借钱时，就会想：“他花钱太多了，省不下钱来。”当我们看到学生奥斯卡每天通宵达旦地写作业，还不能按时完成时，就会想：“他太累了，要放慢脚步。”但只要我们明白了稀缺陷阱的内在逻辑，就会说：“奥斯卡支出的太少了（请记住，他与拥有同等面积土地的菲力克斯相比，支出要少一些）。”或：“奥斯卡的努力还不够（他用更多的时间去写作业，但完成量并没有比菲力克斯多）。”所以，问题不在于花了多少钱，用了多少时间，而在于如何花钱和利用时间。一直靠借贷维生的人，并没有在他想要的东西上花什么钱，许多收入都用在偿还贷款上；工作一直赶不上进度的人，并没有用足够的时间来完成需要完成的工作，而是将许多时间用来追赶早就应该完成的工作。更确切地说，我们在了解了街头小贩们的境况之后，很可能会认为他们因收入太少而存不下钱。这一点无疑是正确的。但稀缺也同时会因为另一个原因使他们落入陷阱而无法自拔。

我们将在本章中讨论稀缺陷阱：这种陷阱是如何发挥作用的，我们为什么会陷入其中以及为什么我们不会去做那些能将我们从陷阱中拯救出来的事情，就像每天连5卢比都省不下来的街头小贩们一样。

停不下来的杂耍

为了了解我们为什么会落入稀缺陷阱中而不能自拔，首先就要明白稀缺陷阱常常会被人所忽视的一个特征。

在我们的工作过程中，第一次意识到这一特征，是在与经济学家迈克尔·法耶（Michael Faye）于印度泰米尔纳德邦（Tamil Nadu）的村庄中研究珠宝借贷问题的时

候。在当地，珠宝借贷就像首饰典当一样。我们与一座贫困村庄的银行合作，那里提供年利率为13%的珠宝贷款。⁵而令人不解的是，客户更愿意选择当地放贷人年利率高达70%的贷款。村庄中普遍流行的看法是，珠宝贷款是紧急情况下的选择，是紧要关头才会考虑的办法。而放贷人就在村民们身边，他们的工作时间非常灵活——周末也可以去他们家敲门借钱，而银行只有工作日和星期六半天开放。在面临真正的紧要关头时，谁也不想等。只要有了管窥心态，我们就会做出这种选择。道理还是讲得通的，至少一开始便是这样。

但之后，我们了解到这些“紧急事务”的内容：排名第3位的原因比较合理，是医疗费用；而排名第2位和第1位的原因则让人有些摸不着头脑——学费和购买种子的费用。按理说，人们应该提早就知道他们要在什么时候交学费、什么时候买种子，那么为什么这两件事情会成为紧急事务呢？事实上，我们深入调查后发现，就连一些医疗费用都算不上他们真正的紧急事务，有些钱是花在了治疗白内障或生孩子等早就有计划的外科手术上。为什么人们会在最后时刻才对这些事情采取应对措施？为什么他们对待例行的预定事件就好像对待突发事件一样？

你以前肯定也有过这种经历。当你专注于解决本周的吃饭问题时，就不会考虑下一周即将发生事情的细节。之后，当下一周到来之时，一些你有所预期的事情很可能会存在一些你意想不到的地方。你很可能早就准备买一张机票，却错过了提前一周购票才能享受的优惠。或者，你万分惭愧地向妻子低头认错，因为你早就答应陪她去看一场表演，但临时却买不到票了。工作上，在你脚打后脑勺地完成了一项任务之后，突然意识到另一项任务的完成时间只剩两天了。仿佛就在前不久，这项任务的截止日期还有好几周时间。你一直“知晓”的事实，现在却成了突如其来的惊诧。

如果长时间地经历这种“最后关头”，我们就会一直处在所谓的“杂耍”状态，也就是持续从一项紧要任务转移到另一项紧要任务的状态中。杂耍是管窥的必然结果。当有了管窥心态时，我们就只能局部地、暂时性地解决问题。我们可能会在当下尽己所能，但到头来却会出现新问题：今天的账单迫使我们去借钱，但就会给未来创造另一份数额更大的账单；廉价的医疗服务虽然能缓解一时的问题，但之后还是需要更为昂贵的治疗。在同时抛起许多球时，带有管窥心态的我们只能将目光集中在即将落入手中的那个球上。有时，我们能一次性地解决问题，而更多时候，我们只能在最后关头接住球，然后再将之抛入空中。

杂耍，就是人们将可预期事件当成突如其来的事件进行处理的根源。杂耍时，你从“管子”视野中只能看到即将落入手中的球，而会忽视

抛在空中尚未落下的其他几个球。当那几个球突然降落时，对于用管窥心态看问题的杂耍玩家来说，这些球就是新的“惊诧”。用旁观者的眼光看，你可能早就知道球快要落下来了。作为不偏不倚的局外人，我们也能想到学费应该什么时候交；而对于忙着处理财务问题的人来说，只有当财务问题迫近时，他们才会觉得债务是实实在在存在的。

这种管理稀缺的方法会使收支变得乱七八糟。如果我们一次又一次地用应急心态去处理最紧迫的问题，长此以往，这种临时修补的做法就会织起一张由各种承诺构成的大网。结果，就形成了用各项资源和责任东拼西凑成的凌乱的补丁图。对于繁忙人士来说，这就意味着我们在本章开头讨论过的那种压力重重的扭曲计划，里面充满了堆积如山、摇摇欲坠的待办事项和定了又改、改了又定的会议。对于穷人来说，这就意味着烦琐的财务人生。《穷人的投资组合》（*Portfolios of the Poor*）一书中进行的详细研究显示，平均来讲，穷人会使用10款完全不同的金融产品。⁶在孟加拉国，有42户家庭在一年间使用了一种短期零利率贷款产品高达300多次。任何时候，调查中的穷人都会同时欠与被欠着好几笔钱，而这一借贷行为在经年累月的管窥心态下，为了解决当下最紧要问题而被织成了一张凌乱的补丁图。

决策，无论是购买一样东西，还是进行一项投资，都要依据这套越来越复杂的大补丁图来进行。之前进行的选择留下了尚未解决的问题，而这些问题则令新的选择更加棘手。雪上加霜的是，杂耍通过我们自身的行为使问题变得更为复杂。稀缺陷阱中那乱七八糟的收支，更是增加了手头这些问题的复杂性和挑战。

杂耍，并不一定是由时间紧迫、任务繁忙造成的。在有些地方，穷人们会同时打好几份工，他们是真的很忙；而在另一些地方，穷人们虽然有很多自由时间，但依然以杂耍的方式处理事务。在农耕区，每到邻近收获时，都是杂耍最为频繁的时候。此时，正是上一次收获得来的收入用光时。我们的研究显示，正是此时，人们的流体智力表现得更低，执行控制力也有所减弱。同时，此刻的农民们除了等待农作物成熟之外，整天都无所事事。时间利用的相关数据显示，农民们在这一阶段虽然每天只工作很短的时间，但依然会出现许多杂耍现象。⁷杂耍，并不是因时间不够用而心烦意乱，而是因为人们的大脑中装了太多需要做的事情——大部分带宽都用来为抛向空中尚未坠落的球而焦虑。

落后一步和杂耍这两个特征清晰地定义了稀缺陷阱。生活在稀缺陷阱之中，你所拥有的就会比实际应拥有的更少。这种生活总是需要我

们紧赶慢赶，在每个球就要落下的刹那开始应对，结果就在匆忙间织出了一张乱七八糟的补丁图。所有这些问题大部分都是我们处于稀缺状态中所采取的行动的后果，而这就引出了一个显而易见的问题：如果说管理固定资源有几种不同的方法，那么我们为什么会不断地重复其中非常低效的一种？为什么我们无法逃离这个陷阱？

逃离

我们已经了解到陷于稀缺陷阱而无法自拔的一个主要原因：管窥心态会致使我们产生借用行为。当利息很高时，比如街头小贩们所面临的境况，这种极为本能的冲动就会创造出更多的稀缺。这不仅仅是小贩们的经历，也是第5章中提到的借了工薪日贷款的桑德拉的经历。虽然借用机制十分强大，但出于其他原因，稀缺心理也会让我们难逃其中。

逃离稀缺陷阱，首先就需要制订计划，而在稀缺心态的控制之下，我们很难集中精力做到这一点。制订计划很重要，但并不紧急，而这类事情正是管窥心态导致我们所忽略的事情。做计划，需要退后一步来审视自己的生活，而杂要则会让我们始终处于当下——将关注点集中在快要落下的球上，使我们很难顾全大局。你也很想停下脚步，不再追赶，但你需要去做的事情太多了，以至于根本不知道如何才能停下来。现在，你必须付房租；现在，你必须在截止日期之前完成一项工作。长期计划，肯定是在你的“管子”视野之外的。

最重要的是，对未来进行计划需要带宽，而稀缺则是带宽的沉重负担。印度金奈市场上的小贩们每天都要反复思考许多事情：每样蔬菜、水果要进多少，进什么质量的？前一天还有哪些货物没有卖完？这些东西是否能过夜？为什么这一天的货物卖得比平时慢？会不会一整天都这样？每位商家都会有这类想法：富有的商家能承受得起偶尔的起伏，有能力从容决策，然后接着去做别的事情；但对于街头小贩们来说，这种想法会一直盘旋于他们的脑海中。这些问题会成为带宽的沉重负担，就算小贩们觉得自己已经做出了选择，还是会反复琢磨：是不是真的应该为下周的节日储备点货？我是不是承担了巨大的风险？这样的想法会一直占据他的大脑。我们已经了解到，这些想法会成为实实在在的重负。在这种环境下，集中精力去制订逃离稀缺陷阱的计划确实很难。

雪上加霜的是，真正的计划比我们上面描述的情形要复杂得多。对于小贩们来说，每天省下5卢比，是不是正确的做法？是不是应该在某

些日子省下更多的钱？如果真的需要花钱怎么办？同样，类似的问题也不局限于街头小贩们。在引言部分，我们讲述了一个可以让塞德希尔和肖恩逃脱困境的简单“计划”：拒绝所有新任务和购买欲望。而若想形成真正的计划，难度则要大得多：肖恩真的不能买任何东西吗？那些从长期来看能省钱消费，是不是可以考虑？比如定期去看牙医，给车子换个新轮胎。那么，他需要首先偿还哪些债务？是最迫近的、借贷时间最长的，还是数额最大的？杂耍和稀缺陷阱形成了由各类责任拼凑而成的凌乱的补丁图。从中找出最佳的解决方案，可不是一件轻而易举的事情。

最后，就算制订了计划，我们真正执行起来也会有难度。我们了解到，美好的意图总是无法落实。在面对非常有吸引力的工作任务或商品时，我们很难拒绝。按计划行事，需要带宽和执行控制力，而稀缺会削弱我们在这两方面的能力。

杂耍也会使逃离稀缺陷阱的难度变得更大，因为意想不到的事情总会发生。比如，你终于制订好了逃出陷阱的计划，突然之间，因为车子年检过期，一张罚单横空飞来。你曾将年检的事往后推，就像重新抛入空中的球一样，现在，这个球又落回了你的手中，你又多了一件事要做。就这样，你再一次回到了稀缺陷阱的底部。

所有这些又会因为余闲的缺乏而变得更加复杂。如果街头小贩们能做到日复一日、谨小慎微地攒钱，时刻保持警醒和自我控制，那么就能像预想的那样积少成多。但很可能在几周之后，就有那么一天，他大意了，在冲动下买了样东西——有些东西看起来就是很值得购买的样子，毕竟他手里有钱。就这样，几周时间的精打细算和自我忍耐全都付之东流了。所以，逃离稀缺陷阱需要的不是偶尔表现出来的警醒，而是持续、永恒的警醒——一种在任何时间里都能抵御所有诱惑的能力。

那么，意志力是否会在不断的磨炼中变得更加强大？穷人们不得不持续考验自己的意志力，他们是否能拥有更为坚定的意志呢？没有证据显示，意志力会随着经历困难次数的增加而提升。⁸（人们总是认为，穷人们的意志力会更强一些，而事实与普遍的认识相悖。）退一步讲，就算贫穷会增强人们的意志力，我们还是有理认为，单凭增强的意志力并不足以让人们拥有从一而终的坚忍。而且，我们还有明确的证据可以反驳上述认知。

研究显示，在我们不断进行自我控制时，自我控制力很可能会被消

耗掉。

稀缺实验室

SCARCITY

一项研究让节食者进入一个满是美味零食的房间，里面放有薯片、彩虹糖、巧克力和咸味花生豆，然后让他们在计算机上完成一项任务。对于有些人来说，食物就摆在他们旁边非常显眼的位置；而对于另一些人来说，食物摆在他们看不见的地方。在计算机上完成任务后，实验对象可以吃到大桶的冰激凌。那些坐在食物旁边的人，刚才一直在抵御美食的诱惑，现在终于忍不下去了——他们所吃的冰激凌量要比远离零食诱惑的另一组实验对象多得多。[9](#)

这一领域的研究人员将意志力比喻为肌肉，以此说明意志力在频繁使用后会变得疲劳。从这个角度理解，如果某个人需要持续地抵御诱惑，那么他的意志力就会被耗尽，想要逃离稀缺陷阱就会变得更难。

问题的根源

稀缺陷阱尤其会令人感到辛酸的就是，它会让人产生一种感觉：只要再多一点资源，就能摆脱欠下的所有债务，就能逃脱这个恶性循环。所以，工作一直赶不上进度的人会说：“如果我再多那么一点点时间，就能赶在进度要求之前做完这些事。”街头小贩会想：“如果我拥有买水果的钱，而不用一点一滴地攒钱，就能摆脱债务，我的收入也能翻番。”在所有这些情况中，一次性的资源补充似乎是解决问题的关键。

我们想要一探究竟，于是为印度市场的小贩们提供了他们所需要的现金。我们与经济学家迪恩·卡兰（Dean Karlan）合作，针对数百名小贩展开研究。对于其中的一半，我们进行了为期一年的跟踪调查，对他们的财务状况进行了记录。对于另一半，我们为他们指明了一条逃离陷阱的道路，帮他们还清了所有债务。一夜之间，我们将他们从借债者转变成了有能力存钱的人。于是，他们的收入翻番了。

我们想要了解稀缺陷阱的细节和形成原因，所以参考了对街头小贩们落入债务陷阱这种现象的惯常解释。第一种解释认为，小贩们宁愿借债也不愿意存钱，因为他们没有地方储存省下来的钱——一方面，银行拒绝为他们提供服务；另一方面，他们也会担心随处放置的现金的

安全，因为这些钱很容易被偷，或者被家里的其他成员悄悄花掉。如果真的是这样，那么在拿到我们提供的现金后，他们应该会迅速地去购买一些耐用品，然后将这些东西保存好，并继续借债。最终，他们会再次落回稀缺陷阱之中。

第二种解释认为，小贩们缺乏长远的目光，他们之所以会陷入债务陷阱之中，是因为他们对未来没有长远打算。在我们看来，这种观点不符合实际情况。这些小贩们每天凌晨3点钟起床，在拥挤的三轮车上忍耐45分钟，前去进货。接下来的一整天，他们都要在烈日下看守摊位。这些行为不是短视之人能做到的。还有人说，至少小贩们在财务问题上并不能着眼于未来。如果真的是这样，那么只要拿到我们所提供的现金，他们就会立刻挥霍一空。可以想象，短视之人花起钱来相当快。这样，小贩们很快就又会落回稀缺陷阱之中。

第三种解释认为，小贩们理解不了复利的强大力量。毕竟，对于我们来说，50天就能无债一身轻的事实也会颇令人惊讶。也许当小贩们听说这样的事时，也同样会感到惊讶。对于选择借债的小贩们来说，如果其无法理解借债行为带来的叠加成本，那么每天去借债，在感觉上就会比这种行为本身更廉价。既然为小贩们提供现金也无法改变他们对复利的认识，他们会继续借债，很快就又会落回稀缺陷阱之中。

我们认为，为小贩们提供一次性的资金补充，让他们逃离稀缺陷阱，可以从中获得许多新发现。于是，我们在随后的一年中对这些无债一身轻的小贩们进行了跟踪调查。

在最初的几个月里，无债一身轻的小贩们并没有重归债务陷阱之中。他们没有将现金随意花掉，没有将现金换成物品以便保存，也没有再次开始借债。好像他们现在了解到了身陷债务陷阱的害处，坚持要远离以前的生活方式似的。这种现象符合定性数据：小贩似乎完全理解了落后一步的高昂成本。与那些忙着做不完事情的人一样，小贩们似乎也深刻地认识到了处在稀缺陷阱中所要付出的沉重代价。

这并不是事情的全部。在随后的几个月中，小贩们一点一点地又回到了最初的状态，其实说一个接一个地回到了最初的状态更为贴切。到了一年结束之时，他们所有人累积的债务量都等同于没有得到现金帮助的控制组。

由此可见，实验数据驳斥了上述那些惯常解释，小贩们没有立即落回债务陷阱中，而帮助他们脱离债务的一次性现金补给也不能让他们免于稀缺陷阱的威胁。

那么，对这种行为我们应该如何解释？小贩们为什么会最终落回稀缺陷阱之中？稀缺陷阱究竟有何神通广大之处，能在人们拿到足够的金

钱，将收入翻番之后，还能对其生活再次产生如此巨大的影响？

余闲，应对突发事件的利器

问题的核心在于余闲的缺乏。就算拿到我们资助的现金，小贩们每天也只能靠不到两美元的金钱来维持生计。毕竟，他们的收入不只用来养活自己一个人。当眼前的行李箱如此之小，而需要容纳的东西很多时，他们一旦遇到点意外，比如亲戚结婚要给礼物等，就会出现资金周转困难——印度的社会习俗要求其在亲戚结婚时购买一件大礼物。那么，小贩们如何处理这件意外事件，这部分取决于其是处于身负债务还是尚有余钱的状态中。

身负债务时，小贩就要面对这个难题。他必须做出权衡：若想买礼物，我要放弃什么？或者干脆买件小一点的礼物。此时，他有了管窥心态，但借钱并不是一个选择，因为其已经从放贷人那里借钱去进货了。他只能尽自己的微薄之力去平息这场风暴。他可能会因为购买礼物而放弃许多，而放弃令其感受到了切实的痛苦。他也可能购买了一件拿不出手的礼物，并因此而感到羞愧不已。

现在请想象，我们帮助这个小贩解决债务问题后，他手头尚有余钱。当他突然需要去购买礼物时，就会产生管窥心态。他必须去应对这一紧迫的需要。对于小贩来说，摆在眼前的是一个“轻松”的解决办法：用手里的现金去解决问题。当然，这些现金只是用来应对紧急事件的，而亲戚的婚礼也属于紧急事件。他可以借用自己的运营资本，用手头的现金去购买结婚礼物。那么，他将如何才能从另一轮债务的恶性循环中走出来？这样做的成本是什么？我们似乎能轻易地想到这些问题的答案：“目前，我考虑不了那么多。”可见，这些顾虑全部都落在了“管子”视野之外。

从这个角度来看，这名小贩之所以落回了稀缺陷阱之中，是因为在他的可利用资源中没有足够的余闲，也就使他无法平息自己所面临的各种突发事件。如果他现有的余闲不足以应对大的突发事件，那他就会重新回到稀缺心态之中。只要回到这种心态，最先被牺牲的就是存款。虽然我们没有得到直接证据，但从小贩们那里搜集到的数据完全支持这种解释。小贩们并非立刻落回到陷阱之中，而是逐渐地、一个接一个地掉了回去，就好像有一只无形的大手将他们推了回去一样。而当突发事件偶然发生在他们身上时，就会呈现出这种现象。在许多情况下，小贩们报告称，突发事件就是他们重新开始借贷并最终跌落回债务陷阱之中的导火索。

当问题不是金钱而是时间时，我们会更加感同身受。

请想象，有个日理万机却总是完不成任务的人获赠了一份时间礼物：过期未完成的任务凭空消失，所有需要花时间解决的问题都被解决了。这位曾经不堪重负、如今依旧忙碌的人，可能会在一段时间内保持积极状态。但最终，他还是会因一项重大工作中的突发问题、因病休假、工作效率低下等突然之间又落后。

任何一点小小的不稳定都会威胁到生存于稀缺陷阱边缘的人，因为他们没有足够的余闲去吸收这些不稳定因素，只能任由其影响自己的生活。在《穷人的投资组合》一书中，作者指出穷人们的生活中充满了许多不稳定因素和突发事件。那些每天靠不到两美元维持生计的人，不可能每天只用两美元——可能有时要3美元，有时要1美元。¹⁰社会底层的生活是无常的。在美国等发达国家，这种震荡幅度可能会小一些，但依然足够大。穷人们会从许多地方获得收入，他们常常要打好几份工，而且基本都是临时性的。许多人的工作按小时计价，每天工作几小时也说不准。另外，他们还随时都有可能失业。突然出现的费用问题，比如车子坏了或家人生病了，这也是一道难关。请看下面这段引自新墨西哥州一所社区大学的访谈内容：

汽车修理费本来就是不可预期的。访谈参与者称，修理费一般都在数百美元，而这个数字占据了他们月收入的很大比例。为了支付汽车修理费，访谈参与者会向朋友和亲戚借钱，寻求财务支持……或等待预期的大笔收入到手，比如学费补助等。¹¹

最重要的是，有可以用来平息突发事件的可用余闲。这就是为什么不稳定性会有如此巨大影响力的原因。如果没有充足的余闲，在车子坏掉时，你要去哪里寻找汽车修理费？如果你有活期存款，就会使用这些存款。如果你很富有，就会削减其他消费：你也许会临时取消本周末计划之中的奢侈晚餐；如果你有第二辆车，可能会推迟修理这辆车的时间，先细心筹备出修车的钱。这些都是简单廉价的可选方案。但当你存款不够，没有第二辆车，也没有晚餐可以取消时，修车这件事情就会变成一个十分严峻的考验：你上哪里去找钱？就在那一刻，你有了管窥心态。于是你去借用，开始踏上回归稀缺陷阱之路。

所有这些严酷的事实说明：我们要深化对稀缺的认识。稀缺并不仅仅是资源和寻常欲望之间的鸿沟。就算在许多日子里都能享受到余闲，但真正重要的，还是那些忍耐稀缺的日子，就像街头小贩们的例子一样。为了远离稀缺陷阱，仅拥有比寻常欲望更多的资源是不够的。我们要拥有足够的余闲（或其他一些机制），以去应对任何时候都有可能突然出现的重大事件。社会科学家，尤其是经济学家，很早就了解到不确定性会对结果产生巨大的影响。我们知道，不确定的回报会减少投资，不确定的收入会让人焦虑、迟疑。而我们目前的讨论是，从另一个

视角在稀缺的大环境中看待不确定性和不稳定性。这种视角认为，阶段性的稀缺会引发一些行为，而这些行为最终会将我们拖回稀缺陷阱之中。在稀缺陷阱之中，如果某个资源充裕的阶段不时地被稀缺所打断，那么其很快就会演变成一种永恒的稀缺。

顺便说一下，我们并不是说避免稀缺陷阱的唯一方法就是拥有足够平息所有突发事件的大笔财富，而帮助小贩们解决问题的唯一方法也并不是给予他们更多的钱。我们的讨论强调的是，我们要建立起针对突发事件的缓冲机制。如果小贩们能得到低成本贷款或建立活期储蓄账户，而且只在紧急事件发生时才使用，那么他们就有了关键时刻所需要的余闲。同样，针对一些突发事件的保险也能解决问题。当然，许多人都能意识到这些缓冲机制的好处。这些好处实际上远比我们预期的要大，它们不仅能用来进行风险管理，还可以用来保护我们免于再次跌落到稀缺陷阱之中。

贫穷始于富足

我们可以认为，小贩们之所以会重归稀缺陷阱，是因为突发事件，但我们也能将问题归咎于缺乏缓冲机制。既然他们知道自己生存于动荡之中，那么为什么不在经济宽裕的日子里存点钱以备不时之需呢？印度街头的小贩们并非唯一犯下这种错误的人，实际上世界各地的穷人们的活期存款都太少了。我们之前讲到过，研究报告显示，美国人一般认为，如果他们遭遇了紧急情况，那么自己根本无法在30天内筹到2 000美元。[12](#)数据也显示，手头活期存款较少的穷人更容易遭受突发事件的打击。

我们可以这样来看：小贩们的问题远在突发事件发生之前就已经存在了，稀缺陷阱的种子也早在资源相对充裕的阶段就已经种下了。而且，同样情况也会随着事件的发展而出现。你热火朝天地赶着工作，但还是赶不上进度。你的生活非常苦闷，你暗自发誓，这辈子再也不做这类事情了。截止日期过去了，你终于爬到了陷阱的边缘，喘了口气。此时，距离下一个截止日期还有几周时间。谢天谢地，现在你终于可以放松一下了。几周之后，你想不明白时间都去哪儿了。再一次，你开始加班加点地工作。就像小贩们面临的稀缺一样，你的稀缺来自资源相对充裕阶段时所犯下的错误。

在资源充裕阶段，我们太过松懈了，会浪费时间与金钱。在第2章

中提到的收获研究中，农民们在收获之前很穷，但他们本可以避免这种境况。如果他们能在收获之后更好地管理金钱，就不会在收获之前发现自己兜里没钱了。他们之所以在收获之前穷困潦倒，只是因为他们在金钱充裕时没能很好地进行财务管理。这与穷人借用的问题不同，而是关于金钱充裕时的浪费行为。浪费的结果就是，资源充裕的阶段紧跟着备受稀缺威胁的阶段，如此周而复始。但是，这种情况是完全可以避免的。

到目前为止，我们一直在讨论由稀缺心态所导致的问题。我们会产生管窥之见，我们会有所忽视。我们的带宽有了负担，缺乏远见，更为冲动。所有这些可能都会让人认为，在资源充裕的阶段，我们会精打细算，充满远见卓识。事实当然并非如此。数十年的研究显示，尤其在资源最为充裕的时期，我们更容易拖延。拖延，就是对眼下之事的过分关注，是阵发性的盲目乐观。我们将需要去做的事情搁置起来；我们将本应存下的金钱挥霍出去。我们错误地分配了自己充裕的资源，存储下来的金钱和做完的事情太少，无法免于那些可能威胁到我们的稀缺。富人和穷人一样，都会犯拖延症：富人因为拥有余闲，所以不会受到太大打击；而穷人和忙碌之人由于拥有的余闲太少，只要一遇到突发事件，就会立刻落回稀缺陷阱之中。

远离稀缺陷阱的威胁，需要的不是充裕的资源，而是足够充裕的资源——就算我们乱花钱，做事一拖再拖，还是有足够的余闲去应对绝大部分突发事件；如果拥有足够充裕的资源，就算长时间的拖延，我们还是有时间去管理预期之外截止日期的到来。远离稀缺陷阱的威胁，需要拥有足够的余闲去应对这个世界中林林总总的突发事件，去应对我们给自己制造出来的诸多麻烦。

将所有这些内容联系起来，我们就能明白，稀缺陷阱的出现，其背后有着几个相互关联的原因，而最核心的问题还是稀缺心态。管窥致使我们借用，这样我们就可以用更低效的方式使用同样的实体资源。因为管窥和忽视，我们最终落入了杂耍的境地。稀缺陷阱成了一场无比复杂的事件，一张由到期未完成的任务和代价高昂的短期解决方案拼凑而成的补丁图，而其中的诸多细节，需要我们不断地回顾、修正。我们没有足够的带宽去制订挣脱陷阱的方案。当我们做计划时，也没有足够多的带宽去抵御诱惑、坚持不懈。而且，余闲的缺乏，也意味着我们没有能力去平息突发事件。由于我们无法在资源充裕阶段利用宝贵的时间为未来创建出缓冲机制，所以事态就会进一步恶化。

孤独就是社交稀缺

请想象某人刚来到一座陌生的城市。在老家，他有许多朋友，但在这里，他谁也不认识。几天之后，孤独的处境开始让他有所感触。他给还在老家的朋友们打电话，但感觉却不如之前那样亲密了。他觉得一个人出去吃饭有些尴尬，于是就坐在电视机前随便吃了一顿。这样怎么能认识新朋友呢？于是他决定在约会网站上碰碰运气，几封邮件来往之后，他定下了一次约会。但随着日子一天天邻近，他越来越紧张，之前的约会从来没有像这次这样让他紧张。约会当天，情况也不怎么理想。他想讲个笑话，但笑话说出来一点也不好笑，整个晚上都过得平淡乏味。他一直在琢磨接下来要说什么，以至于根本没有办法集中精力听对方讲话。他意识到，自己做得有点过犹不及。这次约会真是噩梦一场。

我们可能会认为，这个人陷入了社交稀缺的陷阱。他的孤独导致自己很难交到新朋友，而他表现出来的行为也会让孤独一直延续下去。但这种稀缺陷阱，与我们之前所讨论的所有稀缺陷阱都有所不同：这里没有借用，也没有为应对突发事件而建立缓冲机制的问题。笑话讲得不好笑，无法倾听对方讲话等问题，源自此人太想被其他人喜欢而做得过犹不及，源自太过专注于自身的稀缺。

稀缺实验室

SCARCITY

研究显示，孤独者会对一件事情过于专注。[13](#)一项研究中，实验人员请那些将自己评判为“孤独者”的人录一段音。他们没有具体的任务，只需进行自我描述，并尽量做到有趣。他们唯一了解的是，事后会有人倾听他们的录音，给他们打分。可以想见，打分的人在倾听这些孤独者的讲话时，都觉得没有什么新鲜之处，认为这些孤独者比其他人要无趣得多。这不足为奇，我们多数人可能都认为，这就是这些人如此孤独的原因。

实验的另一个版本显示，这种理解没有考虑到某些重要元素。这次，孤独者同样会对着录音机讲话，但重要的是，有一点不同之处：他们不知道之后会有人倾听他们的录音，并对他们进行评价。所以，他们录下的内容都是轻松状态下的心声。

针对这些录音，评判人员认为，孤独者讲话的有趣程度和普通人相

仿。所以，孤独者的问题不在于他们本身是无趣或缺乏魅力的，而在于他们在自认为重要的事情上表现得很糟糕。这也不是缺乏知识所造成的。不知读者是否还记得引言中提到的研究项目：孤独者在理解他人情感方面更擅长，这正是他们得到的专注红利。而当利害关系增大时，他们就无法灵活地运用这方面的技能。我们可以认为，孤独者发挥失常了。请回忆一下，你以前是否也曾经历过语塞或言语笨拙的情况？如果你和我们一样，就很可能依然会记得在某些社交场合，你特别希望一切进展得很顺利，但却在那一刻表现得十分笨拙。

发挥失常，并不是孤独者的专属。体育场上，这一现象比比皆是。篮球比赛中，罚球十分简单——篮球运动员与篮筐之间的距离并不远，不仅能自己掌握速度和节奏，还不会受到其他人的盯防。罚球的英文，直译过来就是自由投篮，这个说法本身就表达了投篮的轻松程度。自由投篮的世界纪录保持者是一位72岁的老人，他连续投进了2 750次球。[14](#)只要经过训练，以90%的命中率进行自由投篮，对任何人来说都不是什么难事。但有些运动员却认为罚球的难度很大。在2002年到2003年的赛季中，职业篮球运动员布鲁斯·鲍文（Bruce Bowen）就遇到了这一问题。那一年他自由投篮的命中率只有40%。[15](#)鲍文的问题不在于缺乏技巧，因为难度大得多的投篮对他来说都不在话下。那一季，他三分球的命中率达44%——三分球的距离要比自由投篮远得多，而且角度通常都十分诡异。投三分球时，动作必须要快，因为总会有对方球员挡在你面前或朝你冲过来。但就是这样，鲍文在那个赛季三分球的命中率还是要比自由投篮的命中率高。

无论是哪种体育比赛的粉丝，都能讲出许多运动员发挥失常的故事。篮球运动员会因错失自由投篮的分数而输掉一场本该获胜的比赛；高尔夫选手会在最关键时刻，简单的入洞一击，却不幸跑偏。不管运动员的水平有多高，在那些时刻，他们也会紧张。而正是因为我们的恐惧，因为我们可能会预期到自己发挥失常，才使事情的进展更富戏剧性。

如今，研究人员已经对人们发挥失常的心态有了更好的了解。许多体育运动中的动作，可以在有意识或无意识的自动状态下进行。你可以在进行自由投篮时，想着胳膊的动作；可以在打高尔夫球时，专注于随势挥杆的动作。你也可以自发地进行这些动作，但却什么也不想。对于职业运动员来说，这些动作早已司空见惯，自发地做到也完全不在话下。[16](#)（下次跑着下楼梯时，可以想一想自己双脚的动作。但如果你险

些跌倒，可别让我们负责。虽然你是走楼梯的行家里手，但用心思考一下上下楼这个任务，也会降低你的成效。）对于新人来说，在心中铭记自由投球时要收紧肘部，或打网球时记得在击球后完成整个动作，就会让自己的表现有所提升。有意识地关注是有帮助的。对于职业运动员来说，所有这些动作都可以自动自发地完成。只要具备了这个级别的技能，过度的关注就会对最快速、最自然的肌肉协调产生干预。所以，运动员常常会因为注意力过度集中而发挥失常。

发挥失常，是一种更宽泛现象的冰山一角而已。心理学家们在研究各类任务的过程中发现，表现与关注度或称“觉醒”是相互关联的，其关联性呈现为倒转的U型弧线——如果关注度过低，表现就会变弱；而若关注度过高，表现同样会被削弱（见图6-1）。[17](#)

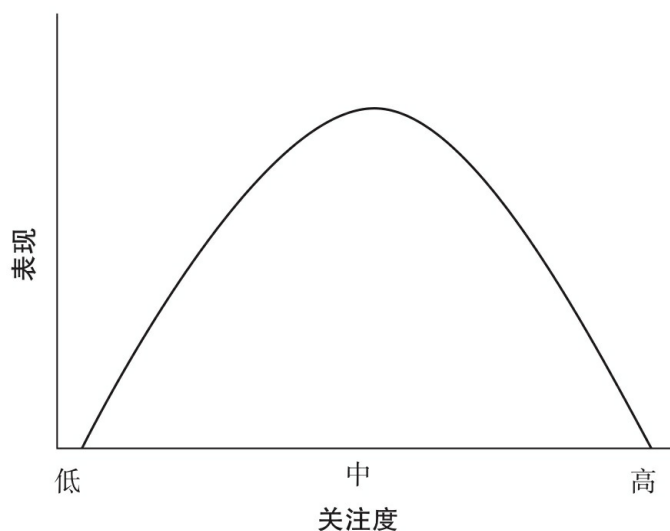


图6-1 表现与关注度的关系

如果我们所从事的任务位于波峰的左侧，那么过多的关注就是好事。对于其他任务来说，比如自由投篮，如果你是职业选手，就有可能因为在动作上投入了过多的精力而到达波峰的另一侧。自由投篮对于一些优秀运动员来说颇有难度，因为他们总是太过专注。布鲁斯·鲍文没有时间去思考他将要投出的三分球，但自由投篮却给了他太多思考的时间。更麻烦的是，你越是尽力不去想一件事情，就越是会想到它。心理学家们将这种现象称为“矛盾历程”（Ironical Process）——研究者请实验对象不去想“白熊”这种动物时，实验对象反而想不到什么别的东西。[18](#)

我们现在将目光放到孤独者身上。我们终于了解到他们为什么会表现得如此糟糕了。他们发挥失常，正是因为稀缺让他们太过专注。对于

人们之间的对话效果，同样存在一条倒转的U型弧线。注意力不集中地与人对话，肯定会让你显得很无趣；太过专注地与人对话，又会让人觉得你特别黏人、特别不独立。孤独者之所以会表现得很糟糕，正是因为他们除了对自身的孤独进行管理外，无法思考其他任何事。他们之所以表现得很糟糕，是因为他们越过倒转U型弧线的波峰，做得过犹不及。他们无法倾听对方的讲话，无法从容地闲聊，而是聚精会神地专注在“他们是否喜欢我”或“讲这个故事会不会好笑”等想法上。就像专业自由投篮选手将注意力转移到别处，随手一抛就能命中一样，孤独者如果将注意力从他们的社交需求上转移开来，或许就能做得更好。但是，稀缺却阻碍了这种变化，将孤独者的思想拉回了本该避免的状态。

节食者也面临着类似的问题。节食最大的挑战就是自我控制，而抵御冲动的最简单办法就是让冲动夭折于摇篮中。如果某样美食没有出现在你的大脑里，那么你就更容易避免在冲动下去吃这种东西。如果你确实想到了这种美食，那么将其遗忘的速度越快，就越容易抵御诱惑。一直想着某种美味的甜点，只能让节食变得更加艰难。节食造成了热量稀缺，而稀缺又牢牢地将甜点摆在了你大脑中最明显的位置。研究显示，节食者成天想着各种美味，不仅仅是因为他们肚子饿，同时也是因为他们所面临的稀缺问题。一项研究指出，刚刚吃过巧克力的节食者，联想到食物的概率更大。生理上，他们拥有了更多热量；但心理上，他们需要做出的权衡却进一步加剧了。节食很难，正是因为节食令我们将专注力放在了那些本想避免的事物上。

上述两个例子中，稀缺锁定注意力的这一关键特征，成了一种阻碍。节食者和孤独者面对稀缺时，内心充满了挣扎，因为稀缺会令他们专注于每一个小细节。

改变心态，逃离稀缺陷阱的唯一希望

穷人会一直贫穷下去，而孤独者也注定会继续形单影只；繁忙之人永远会日理万机，而节食者的计划也总是会以失败告终。稀缺造就了一种心态，而这种心态会令稀缺长存。如果这种说法让你觉得太过残酷，那么就来看看另一种观点：穷人之所以贫穷，是因为他们缺乏技能；孤独者之所以孤独，是因为他们不受人欢迎；节食者之所以失败，是因为他们缺乏意志力；繁忙之人之所以整天忙碌，是因为他们没有能力过上井井有条的生活。这种观点认为，稀缺是深层次个人问题所引发的后果，很难改变。

相比之下，稀缺心态是环境造就的结果，更有可能通过一些措施来加以改变。稀缺并非个人特质，而是自身创造的环境条件所引发的结果，而这些条件是可以进行管理的。我们越是深入了解稀缺在大脑中的发展历程，就越有可能找到办法去避免稀缺陷阱，或至少去减轻稀缺陷阱的影响程度。

第7章 穷人为什么穷

无能可以导致贫穷，贫穷也可以导致无能。穷人的稀缺心态，是导致他们无能的主要原因。研究表明，穷人的认知能力和执行控制力更弱。他们的大脑中装满了稀缺，就没那么多心思去想其他事了。穷人不仅缺钱，更缺带宽。反过来，带宽负担会致使他们的智力下降。

在批评别人之前，你首先要穿上他的鞋子走上一里地。这样，当你批评他时，就是穿着他的鞋子站在了离他一里远的地方。[1](#)

——杰克·汉迪（Jack Handey）
《周六夜现场》节目作家

毫无疑问，贫穷是最为普遍、最为重要的稀缺形式。现代世界中，贫穷的深度与广度令人震惊：联合国儿童基金会估计，每天都有约22 000名儿童因贫穷而死亡；[2](#)近10亿儿童是文盲，不会写自己的名字；[3](#)全世界约有一半儿童的生活水平在全球贫困线以下；[4](#)约16亿儿童的生活环境中没有供电。美国的贫穷问题也十分显著。在美国，几乎50%的儿童在某些特定时候领取过食物补贴；[5](#)15%的美国家庭在一年中的某个时间段连糊口都成问题。[6](#)

截至目前，我们看待各种形式稀缺的方式是，仿佛它们彼此之间可以互换一样。我们的话题从节食转到贫穷，再转到时间压力，而没有关注其中的区别。毕竟，稀缺才是我们的主题。如果稀缺能够唤醒一种独特的心理，而这种心理又与其源头无关，那么我们就能将各种形式的稀缺同等对待。如果存在稀缺的共同心理，那么我们在穷人身上所观察到的现象，难道不应该同时适用于繁忙之人和节食者吗？

不同形式的稀缺拥有共同的元素，但并不意味着这些不同的稀缺类型会产生同样的结果。在化学领域，同样的基础元素按不同的结构比例组合起来，会生成不同的化学物质——碳元素和氧元素能形成二氧化碳这种生命存活所必备的要素，也可以形成一氧化碳这种致命的污染物。[7](#)同样的元素却组合成了完全不同的东西。我们对稀缺的分析遵从的是同样的逻辑。各种类型的稀缺都有着相同的元素：管窥、借用、余闲缺乏、带宽负担。但这些元素会依据所在环境，以不同的方式表现出来：在金钱稀缺的情况下，借用是主要特征；而在孤独的情况下，借用就没有什么意义了。借用这种特定的元素，就像一氧化碳不需要额外多出来的氧原子一样，在特定的环境中也会完全不存在。贫穷所包含的元素会创建出一种环境，这一环境对稀缺心态怀有特别的敌意。[8](#)

一位家境富裕的专业人士非常忙碌，他之所以会陷入这种境况，是因为他接手了许多工作。如果他少做一些，就不会那么忙了。事实上，他可以选择拥有更少的稀缺。从某种程度上来讲，他所拥有的稀缺是可以自由裁定的。

这种自由裁定能力可以为这位专业人士提供一个重要的安全阀，对

稀缺带来的压力和造成的损害起到限制作用。执意要在一周时间内遍览意大利风光的游客，最后只会因为时间的稀缺而感到焦躁。到了某一时刻，他很可能会干脆地说：“算了，我下次再去看罗马竞技场好了。”或是决定：“不如在罗马多玩一天，在南部少花些时间。”这种安全阀很好地限制了稀缺陷阱所带来的压力和损害。

对于那些具有自由裁定能力的人来说，稀缺陷阱会威胁到他们，但其程度是有限的：接手太多工作的人，可能会赶不上几个截止日期；节食者可能会在节食过程中停下来休息几天；繁忙之人可能会去休个假。

但人们无法逃离贫穷去休假。穷人们无法选择不去过贫穷的生活。决定破罐破摔、肥胖到底的节食者和决定放弃雄心壮志的忙碌人士，根本不能与穷人所遭遇的困境相提并论。没有人会认为，印度乡下的穷人只要少一些欲望，就能适应金钱稀缺所带来的窘境。一些最基本的欲望，比如穿衣、远离疾病的威胁以及为逗孩子笑而买一两件廉价的玩具，都是很难抛开不顾的。在强制性稀缺中，穷人并非是独自在战斗。被严重疾病困扰的节食者、极度孤独的人以及那些打两份工才能付得起房租的繁忙人士，面对稀缺都别无他选。可见，自由裁定能力的缺乏会导致一种极端形式的稀缺。

上述讨论进一步澄清了我们所谓的贫穷。贫穷，在我们看来，就是一种经济稀缺的表现形式。在贫穷状态下，若想改变或思考你所需要的事物，是根本不可行的。有些难以改变的需求是生理需求，比如自给自足型农民所面临的饥饿问题，而有些需求则是由社会造成的。我们对需求的感受，取决于他人所拥有的东西和我们的习惯。如今，拥有室内卫生间并不能让生活在发达世界中的人感觉自己特别幸运。其实，室内卫生间直到19世纪后期才出现，现在依然是许多地方的人们可望而不可即的。对于自给自足的农民来说，室内卫生间太过奢侈；而对于生活在新泽西州的美国人来说，则是生活必需品。拥有汽车，在20世纪50年代还是身份的象征，直到如今，在世界上的许多地方也依然如此；而在另一些地方，汽车就是生活必需品。这就引出了一个深层次的复杂问题：我们如何在这些需求之间进行对比？从感受上讲，没有能力拥有室内卫生间的美国穷人，真的与买不起衬衫的印度贫民或买不起汽车的欧洲穷人很像吗？绝对贫穷与相对贫穷，这两者从心理上是否可以对比还尚无有力的证据。而在我们的讨论中，绝对贫穷与相对贫穷都是贫穷的表现形式。

从另一个角度来看，贫穷也十分极端。想一想新生儿的父母，他们

突然之间就感受到了时间的稀缺。

这些新生儿的父母们无法选择“提出更少的需求”，因为要带孩子去看医生，要喂孩子吃奶，要给孩子换尿布，要抱着孩子，要给孩子洗澡，还要不停地哄孩子睡觉。面对这一系列事情，家长们没有自由裁定权，只能用杂耍的心态一样一样地应对。但如果你是一位有钱的家长，时间稀缺的问题就可以用其他方式来缓解：你可以请位育儿嫂或保姆；从外面叫外卖而不亲自做饭；招个会计；雇用一位园丁。所有这些人人都能帮你节省一些时间。同样，如果你在节食，而你很有钱，那么就可以去购买一些味道鲜美的健康食品。金钱是用来交换的，所以我们完全可以用它来解决其他形式的稀缺。

如果你想反过来解决金钱稀缺问题，那么难度就大得多了。你可以尝试多工作几个小时，但绝大多数情况下，你并没有那么多富余的时间，而且就算你多工作几个小时，也只能带来有限的额外收入，却会让你变得更加忙碌、更加疲惫。没钱，就意味着没时间；没钱，就意味着更难开展社交；没钱，就意味着食物的质量很可能会变低。贫穷，就意味着在那些决定生活方方面面的事物上，你更可能处于稀缺状态。

我们利用稀缺心理架起了一座共情之桥，即用一种稀缺形式的体验（比如时间）来与另一种形式的稀缺（比如金钱）体验相连。我们已经了解到迫切需要更多一点时间的感受，可能就能借此想象出迫切需要多一点金钱或多几位朋友的感受。我们利用这座桥梁，在整日抱怨着时间不够、无法在截止日期之前完成工作的繁忙经理人和抱怨没钱付房租的穷人之间建立了联系。

但这座共情之桥也只能发挥有限的作用。毕竟，经理可以说：“算了。我不那么努力也罢，还是去寻找工作和生活的平衡吧。”而手头缺钱的人却不能说：“算了。反正我也不需要这所公寓。”因此，虽然时间和金钱都会成为带宽负担，但这些负担的量级或严重程度却是非常不同的。

“最后一公里”问题

大多数关于贫穷的言论，都会包含一个避而不谈的重要问题。

全世界约有2.85亿人身患糖尿病，[9](#)这种疾病十分严重，可能会致人昏迷、失明、截肢甚至死亡。幸运的是，这种疾病如今已经在人类的可控范围之内了。如果定期以口服或注射的形式用药，就能预防糖尿病对人体造成的伤害。但尽管如此，糖尿病依然是一个大问题：问题的一部分原因出在药理上，因为医药尚无法彻底地治愈糖尿病；但更重要的

原因出在患者的心理上——若想让药物发挥作用，患者必须按时服药，但糖尿病患者通常仅会在50%~75%的时间服药，这严重地影响了药效的发挥。[10](#)

这一事实令人震惊。几十年的医疗研究，终于将一种令人日渐衰弱的致命性疾病转化成了可控疾病，但我们却在“最后一公里”、最不起眼的一小步上跌了跤——这一小步就是吃粒药或打一针这么简单。事实上，医疗界的诸多领域都存在“最后一公里”问题。

20年前，如果我们拥有如今可供治疗艾滋病的抗反转录病毒药物，一定会无比高兴。但是如今，数百万人却因不能坚持服药而死亡。对于肺结核病，不按时服药的问题非常严重，发展中国家为此还专门制定了标准的送药规程——直接观察疗法，即专门的工作人员会每天上门监督患者服药。在有些国家，我们没有能力提供肺结核治疗，并不是因为药片本身价格高昂——这些药不过是廉价的抗生素，而是因为直接观察疗法的成本实在是太高了。

一项又一项伟大的医疗成就，就是因为人们不遵从医嘱而不能充分发挥效用。患病又不遵从医嘱，这是多么奇特的人类行为啊！

许多人都会不遵医嘱，但尤以穷人为甚。虽然任何收入阶层的人们都有可能忘记吃药，但穷人不服药的频率最高。艾滋病、糖尿病、肺结核等一项又一项疾病，却重复着同样的规律。无论是地理位置、医药种类还是副作用，只有一样东西没有发生改变，那就是：穷人最不能持之以恒地服药。

我们再看看农业收成的例子。一块土地上可以种植的农作物量会影响社会的方方面面，可以决定食品价格、世界贸易、环境状况甚至全球可生存人口总数。而对于农民来说，这个问题显得尤为重要，因为其全部收入都来自地里的收成。农业和医学一样，凭借着先进的科技，产生了更优质的种子、更高超的耕种技术以及有机种植的方法，大幅提升了农作物产量和农业的可持续性。像医生一样，致力于解决这些问题的农业科学家也总是为一件事而苦恼：农民们出乎意料的行为。

自古以来，所有农民都深知除草可以大幅度提高农作物的产量——杂草会和农作物争抢养料和水分。除草不需要什么专业技能或机械设备，只需要勤劳肯干就行了。但在世界上的很多贫穷地区，农民们却往往不除草。

据估计，非洲某些地区因不除草而损失的农作物产量，至少占到总产量的28%。[11](#)在亚洲，不受控制的杂草影响到稻米总产量的50%。[12](#)这些数据不乏夸张之

嫌，但就算能将产量提高10%，几天的辛苦劳作也是值得的。而且，因为除草可以在不花钱或不占用土地的情况下增加农作物的产量，产量增加10%就意味着收入会增加20%~30%，这样的增长可不容小觑。但尽管如此，许多农民还是做不到定期除草，以至于让本该到手的收入白白溜走。在这些地区，受影响最大的就是那些最为贫困的农民。

我们再来看看父母育儿的问题。研究人员花费了大量时间对人们养育子女的方式进行了研究。那么，结果是什么呢？父母会在没有必要的时候提高嗓门对孩子嚷嚷吗？孩子有需要的时候，父母会表现出爱意和支持吗？父母在制定规矩的时候，是从一而终，还是会随意地、武断地变更要求？孩子们表现好的时候，家长们能否作出积极的回应？父母与孩子是频繁地互动，还是把孩子往电视机前一放就不管了？家长们是否会给孩子们辅导功课？

经过数十年的研究，专家们发现了一个普遍现象：穷人做家长时更加不称职。他们对孩子更严厉，更无法从一而终，而且情感上也会更加疏离，因此也就显得没那么富有爱心。[13](#)另外，穷人更容易将自身的怒气发泄在孩子身上，前一天还在因一件事情而训斥孩子，第二天又会因相反的另一件事而责备他们。[14](#)穷人无法充分地与孩子进行互动。[15](#)他们也不会经常辅导孩子们做功课。他们还会放任孩子们看电视，而不会给他们念故事书。我们现在已知的良好家庭环境所需要的组成部分，都是贫穷的父母没有能力为孩子们提供的。[16](#)

穷人在许多方面都会显现出不足：美国的穷人普遍更加肥胖；[17](#)在许多发展中国家，穷人不怎么会让孩子接受教育；[18](#)穷人不懂得进行储蓄；穷人也不怎么会带孩子去打疫苗；[19](#)村中最穷的人家，往往就是最不爱洗手、喝水前也不加以处理的人家；[20](#)穷人女性在怀孕后，往往不懂得合理饮食的重要性，也不参加产前护理。[21](#)类似的例子还有很多，我们就不在此一一列举了。

道出这样的事实，好像是怀着不可告人的目的而做的沉闷陈述（滥用一下T.S.艾略特的诗句）。而贫穷问题，徘徊在我们心中挥之不去的问题，竟是如此古老，甚至让人感到厌倦。为什么穷人会无能到如此地步，而且会以如此多样的方式表现出来？

这就是我们一直避而不谈的重要问题。

无能可以导致贫穷，贫穷也可以导致无能

当我们直面这一令人困扰的事实时，首先自然而然地就会去质疑人们对事实的理解。也许穷人们并不是因为“无能”才没有服药，而是因为这些药物的价格太高了。他们为什么不除草？因为太忙了。他们为什么不能更好地养育子女？因为他们就是在类似的环境中成长的，根本不了解其他养育方法。诚然，所有这些关于获取能力、成本以及方法的问题，都发挥了一定的作用。但当你一次又一次地去研究相关数据时，仅凭这些因素是无法对穷人的“无能”作出解释的。

在美国，享受医疗补助计划的穷人不需要为医药花一分钱，但他们还是做不到按照医嘱服药。调查中，生活在农村地区的穷人称，他们在两次收获之间总会有充裕的时间，但就是不想去除草。

我们不能将这些现象看作偶然，而对其不闻不问。所有这些现象的核心，都要在行为上找原因。

另一个直觉上的本能反应就是，我们会去质疑这些事实本身。穷人“无能”与否，实际上只是旁观者的看法。也许他们并非无能；也许那些搜集数据的人本身就怀有偏差。对于这样的说法，许多有趣的心理学理论都可以作为佐证。

稀缺实验室

SCARCITY

在一项研究中，实验对象观看了一段视频，这是一段关于一位名叫汉娜的女孩参加测试的视频。她在测试中的表现有些让人摸不透，她答对了一些难题，却做错了一些很简单的题。其中一组实验对象拿到的材料说，汉娜来自穷人家庭，另一组实验对象拿到的材料说她来自上等中产阶级家庭。两组实验对象都看了她参加测试的视频，之后对她的表现和能力进行了评价。那些看过“穷人汉娜”材料的实验对象，会找到她表现中的更多失误之处，对她的评判较差。与看过“富人汉娜”材料的实验对象给出的评价相比，他们还认为她是低年级学生。[22](#)

我们似乎很容易对关于穷人的数据在理解上产生偏差。我们都对穷人怀有非常负面的刻板成见，而且这种成见还包含“无能”的因素，因此就会自然而然地将穷人的无能归咎于个人特质。那么，研究人员针对贫困人群无能的“发现”，岂不是有些“明知故犯”？而当你近距离观察他们时，就会看清这一众人避而不谈的重要问题，没有那么容易解决。绝大

多数相关数据都与现实情况有着真实的关联，而非偏差下的理解。

研究人员们也无法因为政治偏差而对数据进行特殊处理。通常，这些数据的搜集工作都是由客观的研究人员所完成的——就算他们有些看法，实际的调查结果一般也会与他们的看法相悖。在另一些情况下，得出的研究结论是偶然的，而且并不是他们最初设定的目标。农学家和医疗工作者会搜集大量数据，而收入只是诸多变量中的一项。这些研究人员会将收入情况连带许多其他相关内容一起写在报告中。他们一开始没有抱着对贫困人群展开研究的目的去工作，也没有对数据进行夸大。而且，就算研究人员以贫困为关注点进行研究，也常常会带着同情的态度去接近研究对象。在诸如肥胖或其他与贫困相关的诸多领域中进行研究的学者，总会自然而然地怀有一种与研究对象的亲近感，在总结研究结论时，也常常会颇为伤感。也许，最引人注意的就是，事实证据的广度与深度：许多证据并非来源于孤立的研究或争议性很强的结论，而是经过长时间的数据搜集工作。将这些数据整合起来，就形成了我们所面对的这个重要问题。

如果我们无法对这一重要问题视而不见，那么应该如何理解它呢？一种办法是假设“无能”为因，“贫困”为果。认为穷人之所以贫穷，是因为他们的能力比常人差。如果你的收入取决于你是否能做出正确的选择，那么就自然能得出这样的结论：无法进行正确选择的人们最终会变得贫困。但事实上，这一观点背后又蕴含着复杂的成因。比如，每个人出生的情况——你生于哪个大洲都会对你是否贫穷造成很大的影响。尽管如此，还是有一种非常流行的观点认为：无能导致了贫困。

我们搜集到的数据显示，无能与贫困之间的因果关系，反过来也一样讲得通。贫困，也就是稀缺心态，是导致无能的原因。

大脑自由才能成为合格家长

有一项关于父母养育子女的研究，是以空管员为研究对象的。[23](#)研究之所以会对空管员感兴趣，是因为他们的工作内容每天都有所不同，有时还会处于高度紧张的状态中。某些日子里，空中会有许多需要管制的飞机——天气条件非常糟糕时，等待起飞的飞机就会排起长队，频频延误。每当赶上这种日子，空管员的认知负荷就会变得非常高，他们会长时间地专注于让所有飞机安全降落这一件事上。在其他日子中，空管员的工作节奏可能会稍微慢下来——空中没有那么多飞机，脑子里也没

有装那么多“飞机”。研究人员发现，根据某一天空中飞机的数量，他们就可以预测当晚空管员养育子女的水平。飞机越多，为人父母的水平就会越低。用一个世俗一点的比喻来讲就是，同样一位空管员在轻松地工作了一天之后，就会表现出“中产阶级”的样子；而在辛苦工作了一天之后，就会落入“贫困人民”的队伍。

你肯定也曾有过类似的体会。累得要死的一天结束后，你就想在家里安安静静地休息一会儿，可孩子们却在兴致勃勃地看动画片。电视机的声音没有那么大，但足够吵到你脆弱的神经。你恳求孩子们把电视机关了，心里暗自庆幸自己的语调还算柔和。但孩子们却回答说，现在是他们看动画片的时间，你曾经答应过，如果他们做完作业，就可以在这个时间段看电视，而且他们的确也已经把作业做完了。你迟疑了片刻，但噪声还是太大了。于是你大吼道：“赶快把这个破玩意儿关了！”后来，你对此感到懊悔。你也不想这样对待可爱的孩子们，可就是无法控制自己。

而且，你不开心的理由也很充分。研究者对父母如何养育子女有各种各样的建议，其中最重要的就是始终如一。如果家长在教育孩子们的说法和行动上前后不统一，那么孩子们要面对新事物时就会觉得很难，很容易焦虑，无论是在学习规矩和行事方式上，还是在寻找安全感上。但这些话说起来容易，做起来却有难度。做一位合格的家长不容易，就算你知道自己应该做什么，有时也很难做到。始终如一、前后一致、持之以恒，需要父母对孩子们给予持续的关注、不断的努力以及坚定不移的信念。

良好的教养水平需要足够的带宽、复杂的决策能力和自我牺牲精神。你需要调动孩子们的积极性，去做他们不喜欢做的事情；你需要带着他们去看医生，为他们规划各种治疗活动；你需要和老师谈话，贯彻学校的建议；你还需要为孩子们辅导功课，或者请其他人帮忙辅导功课，自己在一旁监督。不管对谁来说，这些事情做起来都不轻松——无论资源多少都一样。而当你的带宽减少时，这些事情的难度就会加倍。在面对孩子的那一刻，你没有办法表现得耐心，做不到你深知的正确之事，因为你的大脑此时根本就不够自由。这也就是空管员在一天的繁忙工作之后，回到家里就再也没有办法做一位称职的家长的原因。

穷人有自己“领空”里需要管制的“飞机”。他们周旋于房租、贷款、迟还的账单之间，还要算着时间等待下一笔收入。他们的带宽在管理稀缺的过程中，已经用尽了。就像脑子嗡嗡作响的空管员一样，穷人也有

着相同的体验。一位旁观者如果不了解空管员白天忙成什么样，肯定会认为这位家长缺乏教养能力。

最近一项研究也给出了一些证据。我们知道，贫穷的家长每个月都会收到食品补贴，但每个月末，他们还是会陷入没钱买食物的困境。月底，正是穷困家长带宽负担最重的时候，而他们此时对子女的教育也很可能是最严厉的。经济学家丽莎·杰内蒂安（Lisa Gennetian）和她的同事发现，月末时段是家长们领取食品补贴的时候，也是孩子们在学校里最容易犯错被管教的时候。[24](#)

若想成为一位合格的家长，需要许多方面的能力，但最重要的是大脑的自由。而自由的大脑，对于穷人来说，常常像是无法企及的奢侈品。

穷人缺钱又缺带宽

穷人不仅缺钱，也缺带宽。我们在商场研究和收获研究中，看到的就是这种现象。体验贫困或者在诱导下想到金钱问题的同一个人，与不贫困或没有想到金钱问题时相比，几项测试的成绩要差得多。此人会表现出更低的认知能力和更弱的执行控制力。所以说，当大脑中装满了稀缺时，就没有那么多心思去想其他事情了。

这一点之所以重要，是因为我们的许多行为都依赖于带宽，不仅限于为人父母。举例来说，带宽负担过重会让人更容易忘事。也就是说，你不太会忘记那些自己知道的事情〔心理学家所称的“陈述性记忆”（Declarative Memory）〕，比如你第一辆车的品牌，而是会忘记心理学家所称的“前瞻记忆”（Prospective Memory），即那些你计划要记住的事情，比如给医生打电话或者在截止日期之前还账单。这些任务必须存在于你的大脑中，而一旦带宽减少，它们就会被忽略。这样说来，穷人会忘记吃药的问题就很好理解了。有些人会觉得这种说法难以置信：你怎么能忘掉如此重要的事情？但记忆并不会以人们想象的方式工作。你的记忆并非为长期价值服务：肯定没有人会忘记吃止痛药，因为疼痛会一直提醒病人服药；但像糖尿病这样的疾病是“沉默”的，其后果不会立即被病人察觉。对于一位带宽承受着沉重负担的病人来说，没有什么能比持续提醒他服药效果更好的了。

缺乏带宽的另一个后果就是，工作的成效会变低。所有类型的工作，无论是处理快餐订单还是整理超市货架，都需要工作记忆，也就是

在使用信息之前，同时使几部分信息在大脑中保持鲜活的能力。贫穷会为工作记忆加上负担，并导致我们的工作成效降低。我们的工作成效之所以会降低，是因为我们的心智处理器正忙着处理其他需要注意的事项。于是就形成了富有悲剧色彩的境况：最需要用劳动换取薪水的穷人，他们的生产力却承受着最为沉重的负担。

负担过重的带宽，意味着处理新信息的能力有所减弱。如果你在听课时总是开小差，那还能学到多少知识？现在，请想象一位低收入的大学生一直在为如何凑齐房租的事而苦恼，此时，他上课能吸收多少知识？我们在上述研究中得出的数据显示，收入和课堂表现的相关性，可以用带宽负担来解释。许多公共健康项目的实施都要靠穷人吸收新信息才能实现。现在，许多宣传活动都致力于帮助大众去了解许多事情的重要性，比如健康饮食、戒烟、产前护理、艾滋病筛查等。在贫穷国家中，会有专门的工作人员去田间地头，帮助农民了解最新的农作物种类，或最近出现的害虫。这些活动在贫困人群中收效不大，我们很难劝服他们少抽烟、吃健康食品或采纳最新的农业技术。因为想要让某人吸收新信息，就必须让其拥有相关的工作记忆。

同时，带宽负担也意味着用来进行自我控制的心智资源变得更少了。辛苦忙碌了一天之后，你还会用牙线清洁牙齿吗？还是会想：“算了，我明天再用牙线也不迟。”而且，我们也深知，长时间挣扎在贫困（稀缺）之中，会进一步令自我控制力变得枯竭。如果你只能承担很少的东西，那么就要去抵御更多事物的诱惑。这样，你的自我控制力会一点点被用尽。假设你是农民，整日为下星期如何糊口的问题而烦恼。上床睡觉时，你脑子里想的只有儿子牙疼的事情——不知道要到哪里凑钱带孩子去看牙医；你与朋友早就约好的聚会，可能去不成了；而且，你最近还要去除草。一觉醒来，你疲惫不堪，更加焦虑。就像懒得用牙线洁牙一样，我们不费吹灰之力就能想到你接下来的想法：“我明天再去除草也不迟。”

我们在搜集到的与吸烟相关的数据中也看到了同样的现象。²⁵经济压力较大的烟民，比较不容易在戒烟上持之以恒。而且，穷人往往会变得更胖，因为坚持选择健康食品需要强大的自我控制力。一项研究发现，低收入女性搬迁到收入水平较高的小区后，极度肥胖和患糖尿病的比率都会大幅下降。²⁶也许其他因素也发挥了作用，但减轻压力无疑是最关键的。如果想要做一个合格的家长，需要自我控制力；如果想要在患病时去上班，需要自我控制力；不冲着老板或客户发火，需要自我控

制力；时常去参加工作培训，需要自我控制力；生活在农村，保证孩子每天去上学，需要自我控制力。环绕在贫困周围的所有这些“无能”，都可以从带宽负担的角度去加以理解。

最后，请思考下面这个例子。

稀缺实验室

SCARCITY

第二天，你要做报告。为了这份报告，你已马不停蹄地准备了许久。你知道休息很重要，于是每天确保在下午5点钟停止工作。下班后，你回到家中，与家人共享温馨的晚餐，然后早早上床休息。但你脑子里依然满是关于这次报告的各种想法。虽然你非常需要睡眠，但就是睡不着。睡眠研究显示，你这种情况其实并不罕见。在一项研究中，实验人员要求38位擅长睡觉的人尽快入睡。[27](#)其中一部分实验对象了解到，小睡之后他们要进行演讲——绝大多数人都不喜欢当众演讲。因此，这组实验对象入睡的难度要大得多，睡眠质量也差得多。其他关于失眠症患者的研究显示，经常睡不着觉的人，很有可能是心里装着烦心事。[28](#)简而言之，一旦你的脑子里装满事情，就很难睡好觉。

关于稀缺的想法会妨碍睡眠，而从长期来看，这有可能是稀缺对带宽造成负担的最有害的方式。针对孤独者的研究显示，他们的睡眠质量更差，睡眠时间也更短。[29](#)这些影响在穷人身上表现得尤为明显，他们也睡不好。[30](#)睡眠不足会带来灾难性的后果。美国陆军发现，缺乏足够的睡眠会令士兵向自己的队伍射击。[31](#)1989年埃克森·瓦尔迪兹号（Exxon Valdez）油轮在阿拉斯加州失事，原因很可能就在于船员睡眠不足。[32](#)如此的负面影响会逐渐累积。研究显示，如果每晚仅睡4~6个小时，这样持续两周时间后，工作表现就会受到影响——相当于连续两天晚上不睡觉。[33](#)睡眠不足会进一步危及带宽。

穷人最缺的就是带宽。整日为了糊口而奔波，就会令这种至关重要的资源日渐减少。这种短缺现象，与营养不良或幼年时代遭受压力从而阻碍大脑发育等因素无关，并不能从生理上找原因。事实上，它是由眼下为糊口而产生的认知超载所导致的。随着收入的增加，认知能力也会相应提高。[34](#)同理，农民的带宽，在用农作物换来钞票时就能得到恢

复。可见，贫穷会成为带宽的负担，并削弱我们的各项能力。

带宽支撑着我们各种各样的行为。我们利用带宽计算扑克牌游戏中获胜的概率，评判其他人的面部表情，控制我们的情感和冲动，读书或者用富有创意的方式进行思考。几乎每一类高级认知能力都要依靠带宽。但是，带宽负担却常常为人们所忽视。这样一则类比可以很好地对其进行诠释：请想象你与正在忙着做另一件事情的人讲话，假设此人边与你说话边上网。如果你并不知道他正忙着做什么，那么在你眼中，他会是什么样的？没礼貌？冷淡？不太正常？带宽负担就会给人留下诸如此类的印象。

因此，如果你想对贫困人群有更加深入的了解，就请想象自己的思想总是在别处游离的状态吧！你晚上总是睡不好；发现自己很难想清楚一些事情；很难自我控制；心不在焉，烦躁不安……而且这些情形每天都会发生。可见，贫困在给人们带来物质挑战的同时，也会带来心智上的挑战。

从这个角度理解，贫穷这一重要问题就不会再令我们困惑。穷人的无能，其主要原因就在于贫穷本身。只要生存于贫困之中，我们所有人都可能变得无能。

带宽负担导致智力下降

本章伊始，我们就举了几个小例子，将矛头直指贫穷这个避而不谈的重要问题。在各种不同的环境中，贫穷都与无能有着直接的关联。我们对这一结论给出了一种解释：带宽负担。但我们从何而知，带宽负担真的能对这一问题进行解释？比如，你可能会想，带宽负担的覆盖面是否够大，能将包括不能坚持服药和忘记除草在内的诸多问题统统囊括进来？我们认为，答案是肯定的。第2章中讲到的商场研究，其中的低收入群体根本算不上真正的贫穷，但他们的带宽负担却已经十分可观了，相当于IQ成绩的13~14分，而且他们的带宽负担对执行控制力的负面影响也很大。在印度进行的收获研究中，我们得出了带宽负担的影响效应在IQ成绩的9~10分的结论，而其对执行控制力的影响则更大。我们曾指出，带宽负担会对认知能力产生相当可观的影响。从标准IQ分类上看，这种影响及其相应的分数可以将你的智商从“正常”升级到“高级”，也可以从“正常”降级到“低级”，甚至降到“缺陷级”的边缘。带宽负担不容小觑。而且，我们在两种完全不同环境中均对此有所发现的事实，更

是强有力的证据。印度农村的穷人和新泽西州商场中的低收入购物者有着太多的不同之处，但他们在很多方面都拥有同样的带宽负担。因此，带宽负担在世界各地贫困人群的生活中都扮演着极其重要的角色。

带宽负担这种解释之所以具有吸引力，是因为它可以将许多现象都包罗在内。对穷人无能的解释，一般都是零碎而非整体的：农民不除草，可能是由于文化水平低；糖尿病患者不在服药上持之以恒，可能是害怕药物的副作用；贫穷的家长不合格，可能是因为其缺乏养育子女的知识。这种解释是不系统的，因为穷人所处的环境差异极其巨大。生活在新泽西州特伦顿市的人们不知道的事情，对于生活在肯尼亚内罗毕市的人们可能就是常识；而在内罗毕市令人习以为常的事情，对于生活在菲律宾农村的人们可能就会是完全陌生的。相比之下，带宽这种单一的基础机制，就能跨越行为、时间和地理位置的藩篱，解释所有这些不同的实证事实。具体的环境在理解穷人生活时同样重要，但带宽具有从基础层面理解时的重要性，而且适用于解释与穷人生活有关的各类现象。

对带宽作用的理解也能帮助我们更好地理解穷人所处的特定环境。疾病、噪声和营养不良不仅仅是痛苦的根源，同时也是带宽负担的额外表现形式。有一种看法认为，穷人缺乏某些基础技能。现在，我们除了将这种看法视为无可辩驳的事实之外，也会思考：带宽负担也许是导致穷人缺乏技能的原因之一。任何形式的技能习得，无论是去学习社交技巧还是养成良好的消费习惯，都需要带宽。如果穷人缺乏带宽，那么他们就无法很好地掌握这些实用技能。

以上论述为我们提供了理解贫困的新视角。我们要从认知的视角，带着稀缺的思维，去审视那些关于规范服药、除草、养育子女等行为的数据。这些行为与需要逐个分析的孤立行为不同，是可以综合到一起，将其视作带宽负担过重所导致的可预见后果的。这一观点也让我们在搜集新数据时有了新的关注点。我们在研究贫困问题时，总是会将目光专注于物质条件上，但同时，我们也应该去了解人们的心理情况，了解他们的带宽情况。这样，许多曾经解释不清的问题就会变得明朗起来。为了加深对贫困人群的了解，我们必须认识到：他们会专注于自身的问题，会形成管窥心态，会犯错误；他们缺乏的不仅是金钱，还有带宽。



第三部分

如何从稀缺
走向富足

SCARCITY

WHY HAVING TOO
LITTLE MEANS
SO MUCH

第8章 如何让穷人摆脱贫穷

从稀缺心态的角度出发，我们就不难理解发生在穷人身上的众多失误。这些失误的不可避免，不是因为他们缺少积极性，而是因为缺少带宽。因此，我们要包容穷人的不当行为。我们要设计有效的预警机制，让穷人提早为未来做准备。所以，为穷人提供的培训要简单实用——节省带宽的教育方法才是好方法。

第二次世界大战期间，美国军方为“起轮事故”（Wheels-up Crash）的频发而头疼不已。¹当时，飞行员在降落后，总会在收起侧翼时，误将轮子收起。可以想见，飞机在陆地上滑行时收起轮子，多么可怕！为了解决这个问题，军方请来了一位专家——阿方斯·查帕尼斯中尉（Lieutenant Alphonse Chapanis）。查帕尼斯中尉是一位接受过正规培训的心理学家，是研究这些飞行员想法的理想人选。为什么这些飞行员会如此粗心？他们太过疲劳了，还是他们太早就放松警惕，觉得可以在完成一项艰巨任务后就此“放手”，或者是在培训他们时出了什么岔子？

很快，一条线索就浮出了水面：问题只出在驾驶B-17和B-25轰炸机的飞行员身上，而运输机飞行员却不会犯这样的错误。这条线索帮助查帕尼斯中尉打破了自身的偏差：他决定不再从飞行员的大脑中找原因，而是去看一看他们的驾驶舱。在这些轰炸机里，机轮控制杆和侧翼控制杆紧挨在一起，看上去根本就是一模一样！而相比之下，运输机的控制杆布局就非常不同。所以，轰炸机飞行员和运输机飞行员表现出不同行为的根源，就在驾驶舱内——其中一种控制杆布局很容易让飞行员犯错。

这一问题被发现后，轰炸机的驾驶舱便从此改头换面。查帕尼斯中尉等研究人员意识到，飞行员的许多过失其实都是由驾驶舱布局不合理而引起的。但在此之前，军方的关注点一直在飞行员培训和如何确保飞行员时刻保持警觉上，并且着重于培养那些不会出错的“优秀飞行员”。查帕尼斯中尉的结论彻底改变了这种情况。当然，飞行员还是要经过培训才能上岗，军方也肯定要去挑选最优秀的人才。但无论你怎么培训，如何甄选，飞行员还是会犯错，尤其是让他们置身于容易糊涂的驾驶舱内。

失误在所难免，灾难却并非如此。优秀的驾驶舱设计，不仅不应该诱导飞行员犯错，而且还要避免将失误酿成灾难。查帕尼斯中尉在着陆变速杆的末端安装了一个小橡皮圈，这样飞行员就能通过触摸来得知自己操纵的是哪根控制杆，从而解决轰炸机存在的问题。优秀的驾驶舱设备，能在飞行员犯错误时发出警报。测高仪旁边的低海拔警示灯，可以有助于提醒飞行员，在航空高度较低时，是飞行员有意而为之，而非无心之失。如今，飞机的安全性得到了很大的提升，不仅是因为我们有能力生产更为优质的机翼和发动机，而且是因为我们对人为失误有了更好的了解与掌控。

包容穷人的不当行为

最初，查帕尼斯中尉因飞行员的行为而感到困惑不已，一心想从他们身上找原因。相比之下，许多分析人士也同样因为穷人的某些失当行为而感到困惑不解。美国针对低收入人群的培训项目，总是有人旷课、中途退课，就连劝服他们报名都很难。²发展中国家的小额信贷项目负责人，总是会为客户不在高回报产品中做充分的投资而感到惋惜——这些客户宁愿将贷款用在偿还其他债务上，用去“救火”（比如马上要交的学费），或是去购买耐用品。³疫苗接种项目实施起来也是困难重重，因为人们总是不去打疫苗，所以项目的实际收效甚微，某些本可以通过接种疫苗而得到预防的疾病，现在依然在一些发展中国家传播。

我们在研究工作中也看到了同样的现象。我们曾在美国的一个福利就业项目中担任顾问，这个项目旨在帮助依赖公共救济的人们找工作。项目遇到的最大问题之一就是，服务对象本身。虽然工作人员反复告知人们要着正装去上班，但他们还是常常穿着不合时宜的衣服露面。很多人的简历漏洞百出，格式混乱，错别字一大堆。虽然这些错误是因为人们缺乏相应的知识或技能，但绝大多数错误都是由人们缺乏按计划行事的能力导致的。就算有工作人员的指导，人们还是不会去使用服务站的计算机来修改简历格式，或利用服务站提供的资源找到更适合的着装风格。面试时，他们常常会两手空空地到场，连简历都不带，也不知道如何展现出最佳状态。而在许多情况下，他们甚至连安排好的面试都不会参加。

而这一社会项目的设计者，并没有从查帕尼斯中尉的视角出发。相反，他们假定问题就在这些服务对象身上，而不去“驾驶舱”内找原因。他们假定问题在于人们缺乏理解能力和动力。于是，他们便尝试通过教育去改变这些人，为他们提供激励机制。在发达国家，这种现象很可能会引发一场关于“福利文化”的讨论。其中一种解决办法就是，对一个人一生中能接受福利救济的总年数加以限制。这种方法的目的很简单，就是要激励失业者去找工作。这种做法导致政府对一系列援助项目加以惩戒，还时不时地取消一些福利待遇，比如改变免费政策，向人们收取纯净水使用费。而且，这种做法还引发了一些具有强大激励机制的计划，比如条件性现金补助计划，该计划服务的人群得到的补助金额，取决于他们所表现出来的各种“良好”行为。

但是，为什么不将这些“飞行员”的工作能力放在一边，去“驾驶

舱”内一探究竟呢？为什么不将人们表现出来的无能暂且搁置，去审视一下这些项目和计划的具体结构呢？如果我们接受了飞行员会犯错、驾驶舱的布局需要进行精心改造，才能够避免错误发生的事实，那么，我们为什么不能为了贫困人群而接受同样的事实呢？为什么不将这些援助项目设计得更具有包容性呢？

稀缺实验室

SCARCITY

针对扶贫项目，我们也能提出同样的问题。以培训项目为例，在培训过程中，学员总是不出勤，中途退课的人也不在少数。但是，在学员满脑子都是烦心事儿和心力交瘁时，错过一堂课又会怎么样？如果他来上课，却在课堂上想着别的事情，又会怎么样？答案就是下一节课的难度会变大。在错过一两堂课后，人们会自然而然地选择中途退课——也许中途退课才是最佳选择，因为再继续听下去也听不懂老师在讲什么。将前一堂课的内容作为基础来讲这一堂课的方法，如此严格的课程设置，并不能包容那些带宽负担过重的学员。偶尔错过一堂课，学员的成绩就会走下坡路，而单凭其一己之力是无法再次跟上课堂进度的。这种项目设计假设的就是，如果人们拥有足够的积极性，就不会犯错误。我们也可以这样理解：那些不愿意准时上课的人，一定是不把上课当回事儿，根本“不值得”培训。

但从稀缺心态的角度出发，我们可以预见到这些失误的频发和不可避免，而失误与否，与人们是否有积极性根本无关。请想象自己在一天的辛苦工作后回到家中，心中忧虑重重，不知道如何才能赚足够的钱来交这个月的房租、支付所有的账单和为女儿办个像样的生日派对，而且你想休息却怎么也睡不好。几周之前，你报名参加了计算机技能培训课程。你的想法是，只要掌握了新技能，说不定哪天就能找到更好的工作。但就在当天晚上，参加培训能给你带来好处却显得那么抽象而遥远。你精疲力竭，快要被眼前的这些问题击垮了。而且你知道，就算你现在去上课，也听不进去什么东西。现在，我们可以将目光跳转到几周之后。那时，你又落下了一堂课。再去上课时，你能理解的比以前更少了。终于，你决定中途退课，等财务问题解决之后再来上课，毕竟你现在根本没有心思学习。所以，这个课程不仅无法包容你的失误，反而将你表现出的这些可预见的错误进一步放大了，而且还强硬地把你推到了教室之外。

其实，我们完全可以不必如此。与其总是强调人们不犯错误或改善行为的重要性，还不如去对“驾驶舱”进行重新设计和布局。举例来说，我们可以更改课程设置，将之转换成模块化的课程单元，然后在不同的时间段里交错开课，共同向前推进。错过一堂课，落下了一段内容怎么办？不要紧，只要去参加比这个课程模块“迟”一两周的同个课程就行了。如果整个模块的内容都错过了，那就等下一轮开讲时再重新学起也不迟。虽然这种方法会让结业的时间推迟，但学员们至少能真正学到知识。目前的培训项目，在设计之初并没有考虑到学员可能出现的失误，就好像学员们不应该，也不允许犯错误一样。但是那些贫困人群，尤其是在失业的状态下，还有许多其他需要处理的事情。而这类事情的存在，致使他们无法安心坐在课堂上听讲。在稀缺的窘境之中，错过培训项目的一堂课，与中学生旷课的概念完全不同。线性设计理念指导下的课程，不允许学生中途缺课，这种课程设置比较适合全日制学生，但对于脑子里有一大堆烦心事的穷人来说，则完全行不通。

值得强调的是，对失误的包容并不能取代个人责任。相反，对失误的包容，可以确保穷人在勇于承担责任的同时，拥有提升的空间。对失误的包容，使人们获得的机会与他们付出的努力和所处的环境保持对等。这并不是说人们不需要努力工作，而是说让那些敢于迎接挑战并为此付出了辛勤汗水的人们能获得更多的回报，正如只要改进驾驶舱内的操纵杆设计，就会让专心致志的飞行员表现得更加完美一样。带宽负担不可避免地会让人们犯下小过失、小错误，而对失误的包容，则能确保这些小过失、小错误不会让人们的努力付之东流。[4](#)

让“警报”来得更早些

还记得前面讲到的对福利救济年份加以限制的政策吗？之所以这么做，是因为有些人认为，频频接受福利救济的穷人会缺乏自力更生的动力。这种观点认为，穷人之所以时而申请福利救济，时而退出福利系统，是因为这一系统太容易让他们产生不去工作的念头。为了解决这个问题，针对主要的福利项目（现在更名为“贫困家庭临时救助项目”），美国对其公民一生中能领取福利救济的年份设定了上限——如今，每人一生中总共只能享受5年的福利救济。[5](#)

规定福利救济的年限，并非莽撞之举。从理论上讲，有了限制，就产生了稀缺，而稀缺又可能让人们资源的“利用”进行更好的管理。这种说法看上去貌似是以稀缺心理学为基础的，却存在缺陷。我们知道，

截止日期在即将到来、成为人们的心头大事时，才会发挥作用。而限制期限较长，就像很久之后的截止日期一样，只有在临近时才会让人产生紧迫感。对于那些眼下有一大堆事情需要处理、抱着管窥心态看待事物的人来说，几年之后的期限无疑会落到“管子”视野之外，而只有截止日期临近时才会有所察觉。在这一限制变成紧迫威胁之前，就会为人所忽视。等人们想到了，也为时已晚。那些制订出这一计划的人们的初衷肯定不是这样的，他们肯定不希望人们在数年间完全忽视救济年限问题，然后在最后一刻才发现自己没有办法继续接受救济。从某种角度来看，这是所有方案中最不理想的一种，既是一种惩罚，也没有起到任何激励作用。

在了解管窥心态之后，我们就可以让限制手段变得更加有效。若想让限制手段影响人们的行为，就必须令其进入人们的“管子”视野之内。其中一种方法是，定期发送剩余月份的提醒。通过唤起人们的关注，我们可以试图将这个遥远的问题主动推入“管子”视野之中。另一种方法是，改变限制的结构。我们知道，频繁的点式截止日期，会比单一的远期截止日期对人产生更大的影响。因此，更好的解决办法就是，制造出范围更小、更加频繁的限制。（比如在给定的几年期限内，只有固定数量的月份，而非一生中的总年份。）为了让人们能立刻感受到超越限制的后果，同时又令这种后果易于为人们所觉察，不会一步置人于死地，也许可以考虑对救济金额予以降低，而非彻底停发。

关于如何设计激励机制（以及如何避免不恰当的设计），我们可以借鉴一个通用原则。落在“管子”视野之外的激励机制，不太可能发挥作用。请想象你在想办法让孩子们接种疫苗，而这些孩子的家长正为这个月的青黄不接而发愁。一两个月之后的支付，和现在就支付，哪种情况对于家长来说更富有吸引力？在印度拉贾斯坦邦农村地区进行的一项研究显示，只要以一千克扁豆作为奖励，就能很有效地吸引人们前来接种疫苗。⁶对于产生了管窥心态的人来说，在遥远的未来才会发生的奖惩措施，并不会产生多大效用。储蓄计划所提供的慷慨补贴，若只能在几年之后才能领取，那么就等于将储蓄这件事划归到了“重要而非紧急”的事件之列，这样事情就会落到“管子”视野之外，被人无限期地忽略。为了让激励机制发生作用，人们必须在“管子”视野中看到它的存在。而大多数激励机制都有落在“管子”视野之外的危险，这就等于是无形且无效的。可见，我们需要对激励机制进行精心的设计。

节省带宽的方法才是好方法

有条件式现金补贴（Conditional Cash Transfers）是为穷人提供救济的一种越来越流行的办法。某人收到的救济金额，取决于他表现出来的良好品行。⁷研究显示，这种计划可以发挥作用，人们会对现金激励作出回应。但这只是事情的一面。另一方面，许多本可以加入计划的人们，却没能积极地作出回应。

这里存在的问题与之前提到的一样，激励机制落在了人们的“管子”视野之外。现金奖励只有到了未来才能领取，而所需表现出来的良好品行却并非是在“管子”视野中就能看到的。这就引发了另一个问题：就算我们有能力将这些激励机制拉入“管子”视野之内，我们是否应该这样做呢？每一个额外的动机，都会成为带宽的负担。带孩子去体检能获得奖金，但家长们必须为孩子们预约体检，记住约定下来的时间，腾出时间去医院，还要督促孩子们去体检（没有哪个小孩喜欢看医生）。做到上述每一步都需要带宽，而且这还只是一种行为。有条件式现金补贴致力于鼓励数十种甚至数百种类似的良好品行，所以仅仅是搞清楚这些激励机制以及所需做出的必要权衡——哪件事值得做，哪件事不值得做，都需要带宽的支持。

我们从来没有扪心自问过：这是否是我们希望穷人使用他们的带宽的方式？我们从来没有在确定哪些行为最值得提倡时，将这么做的成本考虑在内。我们在设计贫困救助计划时，意识到了穷人缺钱的事实，于是就以这一事实为基础展开计划，但却没有想到，穷人除了缺钱以外，还缺带宽。

在我们为穷人提供受教育机会的行动中，尤其深刻地感受到了这一点。在面对许多问题时，我们的第一个反应总是想要针对人们缺乏的技能，为他们提供受教育的机会：针对教育子女的种种问题，我们提供子女养育技能培训；针对财务管理中的种种错误，比如以过高的利率借贷的行为，我们提供财务教育课程；针对缺乏社交技巧的员工，我们提供“软技能”培训。我们将教育看作最温和、最不具侵犯性的解决办法，看作一种最纯粹的善行。但在穷人带宽有限的情况下，教育并非这么简单。无疑，教育是一件好事，但我们对待教育的方式，就好像穷人受教育时并不会为此而付出代价一样。而事实上，带宽需要付出高额的成本：要么就是受教育者本人不会专心致志地接受教育，我们的努力全部白费；要么就是他们会专心地接受教育，但却要为此承受带宽负担。当人们真正专注在培训项目或激励机制上时，其没有关注到的事物是什么？多上这一堂课，是否真的值得他们为此付出本可以用来阅读或陪伴

孩子的时间？可见，带宽负担的增加存在着隐性成本。

就算我们坚信提供受教育机会是正确的做法，但也还是有许多种方法可以在达到这一目的的同时让穷人节约使用带宽。经济学家安托瓦内特·施格尔（Antoinette Schoar）与其著作的合著者共同进行的一项研究，对此就有所体现。他们与多米尼加共和国的一家名为ADOPEM的小额信贷机构展开合作。⁸这家机构的客户是经营小商店、美容院、小吃摊等小本生意的商贩，他们通常都不雇用员工。ADOPEM发现，客户们的账簿总是错误百出，他们对财务知识的掌握并不合格。于是，施格尔与其合著者给出了一个看似简单的解决办法：为这些人提供财务知识方面的培训。施格尔从ADOPEM处得到了一套标准的财务知识培训教程，是世界各地针对小微企业主所进行的一般培训内容。她看到这套教程后的第一反应是：我的天呐！太乏味了！（她是麻省理工学院的金融学教授。）完成全部课程的学习需要几周之久，而课程的主要内容还主要偏重于传统的会计技能，主要是教授学员们如何对现金收入和费用支出进行记账、如何进行存货管理、什么是应收账和应付账以及计算利润和投资方法等。

在没有带宽限制的世界中，上述所有知识都是值得学习的。但在现实世界中，施格尔认为，她可以给出更好的解决办法。她请来当地最优秀的一群创业人士，对他们的理财方法进行了研究。这些创业人士并没有学过复杂的会计方法，而 they 与那些未能成功的创业人士之间的区别就是，他们做到了一件事情：遵从经验法则。

举例来说，几位小老板会将现金收入放入店中的一部收银机里，还会给自己发放固定数额的薪金。这样，他们就能避免把家里的钱和生意上的钱混在一起，从而搞清楚家里究竟花了多少钱，生意上究竟赚了多少钱。（有些女老板会将一沓钞票放在胸罩左边的罩杯里，另一沓钞票放在右边的罩杯里。）虽然这种办法称不上复式记账法，但却十分简单而有效。这种方法既节约了带宽，又保留了记账的绝大部分好处。

施格尔对这些经验法则进行了整理，并以此为基础设计了一套与之前截然不同的“财务教育”课程。她的这套课程时长很短，内容很容易理解。因此，课程内容所需要的带宽也比较少，这一点通过学员出勤率大幅上升这一数据就能感受到。通常情况下，若想让人们回到财务教育课堂上来，总是要费上一番周折。在这一经验法则课程结束时，学员们都非常激动，想要得到更多类似的学习机会。许多人甚至说，他们愿意为了上课而自掏腰包。

学习财务知识所需的带宽减少了，就意味着课堂知识更容易被吸收

了，也就意味着授课变得更富有成效了。课后的跟进调查发现，学生们更愿意运用这些经验法则，而非复杂的会计学规则。这种做法也带来了实际利益：经验法则课程班毕业生的业务销售收入有了提升，尤其是在淡季，因为淡季最能体现出良好的管理习惯所带来的好处。平均来看，上过经验法则课程的小微企业主，他们的淡季收入上涨了25%；而传统的财务知识培训，则对实际收入没有任何影响。我们可以从中清楚地了解到：节约带宽能带来高额回报。

无论是人们在引导之下做出的权衡与选择、教育手段的设计与布局和激励机制的设立，还是对“无能”的把握，我们都能利用对稀缺心理的理解来根本性地改变社会服务项目的设计方案。当然，任何一种方法都不能够奇迹般地彻底解决贫困问题。贫困问题有着深层次的原因，但把握住稀缺心理以及稀缺心理所导致的行为特征，对提升扶贫措施的实施成效能起到一定的积极影响。

带宽可以构建

试想你是一位单亲母亲，正同时打两份工。你手头有许多事情需要应对：除了我们讲过的财务问题之外，你还要处理孩子们白天的看护问题，而日托服务的价格又十分昂贵。虽然社会提供了一个享受高额补贴的日托服务项目，但只能接收你的一个孩子，而且关门时间太早，没办法解决你打第二份工的问题。于是，你东拼西凑地想出了一套解决办法：你请母亲帮忙照顾孩子。你要为一个孩子安排放学后从学校到母亲家的路线，为另一个孩子安排从日托所到母亲家的路线。从事服务业工作，能照顾孩子的时间取决于主管如何安排你的工作。你的主管人很好，总是尽力帮助你，但工作时间上的变动是不可避免的。

现在，请想象我们为你介绍了一家享受高额补贴的日托服务机构。你具体能从中得到什么？答案就是你不用再往返接送孩子，可以从中节省时间；也可能会省钱，无论是显而易见的（这家机构的价格比之前一家低），还是暗含的（我们将你母亲的时间成本也考虑在内）。但我们还给你提供了另外一样东西，而它比时间和金钱都更为重要——你可以将这样东西应用于许多事情上。以前，你因孩子和工作上的种种安排而烦恼、焦虑、忙不开，并因此消耗了许多带宽。现在，我们会将这部分带宽统统还给你，帮你卸下了很大一部分认知负担。我们之前提到过，认知负担一旦减轻，你的执行控制力和自我控制力就会有所提升，养育子女的方法和态度也都会得到改善。你的认知能力会提高，同样获得提

高的还有专注力、工作质量和你选择去关注的每一种能力。从这个角度看，在育儿方面提供协助的意义远远超越了帮助本身，因为这种帮助能构建起带宽这种最深层次的人力资本。

一般情况下，专家在评估这类日托服务机构时，只会看到一小部分效果：母亲们是否能工作更长时间？她们迟到的次数是否减少了？但这种眼光未免太过短浅了。母亲们从中能获得的是，无法量化的大脑自由以及带宽。如果某个项目是成功的，那么其所带来的利益就会体现在许多方面。在所有其他因素相同的情况下，我们能在考察成功项目时，了解到项目对人们心智产生的影响：工作记忆力是否有所增长？冲动控制力和自我控制力是否有所提高？某些对现存服务项目的悲观看法，可能是因为没有认识并考虑到上述影响的存在。如果我们用太过短浅的眼光去看待这些日托服务项目，就会遗漏许多其他方面的好处。如果将所有这些好处均考虑在内，那么成功的服务项目带给人们的就不仅仅是尚可接受的回报。而如果我们不能看到人们最深层的需求，没有了解这些利益会给人们带来什么样的好处，那么我们就注定会低估其影响力。

除了在照看孩子方面，世界各地都有许多实例可以向我们证实：带宽是可以被构建起来的。首先，我们来看看财务领域。在穷人每日忙于应付的事务中，很大一部分都是为了解决临时出现的生计问题。如果我们能帮助他们解决这些问题，就能为之腾出一部分带宽。而这些问题通常都是“急性”的，需要马上弄到钱才能解决。这种需要一般并非大笔投资，而是小额的，比如为孩子买校服之类的消费。换种说法就是，穷人最需要的往往是放贷人最容易提供的，即为解决紧急需要而快速借出并偿还的小额现金。但实际上提供给穷人的金融支持，一般都与此恰好相反：谨慎而缓慢地提供中高额度贷款。这类贷款在做投资时可能会派上用处，但如果人们忙着解决临时出现的生计问题，就不会有足够的带宽去考虑做投资。这样看来，虽然信誉很高的小额信贷机构为穷人敞开了大门，但他们还是更愿意去找私营的放贷人。

在印度时，我们与KGFS这家为农村贫困人群提供全方位金融服务的机构合作，对一种极短期的小额贷款产品进行了测试。测试结果显示，人们对平均不到10美元的贷款有着大量需求，而这种结果令我们颇为吃惊。⁹因为这款金融产品并没有办法帮助客户积累财富，更没有办法将人们变成创业家。表面看来，这一数额也无法让人们的生活发生什么变化。但实际上，它很有可能会取得以上种种成效。

稀缺陷阱的出现就源于不断的救火和管窥心态，源于人们做的事情，在“管子”视野之外都潜藏着需要付出的巨大成本。如果能对此进行改变，我们就能改变贫穷的逻辑根源。

我们也可以去寻找问题的源头。在发展中国家，由于缺乏正式、稳定的雇佣关系，人们的收入总是起伏不定；就算在发达国家，许多低收入职工也要面对同样的问题。我们之前了解到，收入的不稳定性是人们被迫“杂耍”于多项事务之中的主要原因。那么，为什么不试着予以缓解呢？对于全世界的贫困人群来说，若有关部门能够加强对工作可靠性和收入稳定性的关注，就会极大地改变人们的心理状态。

我们还可以走得更远。我们总是倾向于去关注那些较为重大的突发事件，比如医疗费或旱涝保险。虽然这些事情非常重要，但当人们忙于处理多项事务时，即使很小的突发事件也会产生极其重大的影响。对于一位贫困农民来说，如果牛生病了，那么因为这头牛不能干活而引起的收入减少，就足以将他拖入稀缺陷阱之中。这样看来，我们应该想办法确保贫困人群有能力去抵御这些表面上“微不足道”的突发事件。在美国，如果无法保证统一的工作时长（比如这周你工作50个小时，下周就只需工作30个小时），那么就会导致人们忙于杂耍，陷入持久的稀缺状态之中。对于穷人来说，工作时长不稳定的问题可能比失业更可怕。因此，一种解决办法就是，创建出类似于失业保险的保障机制。

我们了解到，杂耍时所遇到的以及带着管窥心态所看到的大多数突发事件，其实都是可以预见的。一方面，突然间需要购买肥料的钱，可以算是一件突发事件。另一方面，这种事完全可以被预见。每年都要买肥料，但当你忙着在多项事务间杂耍时，就可能会忘记买肥料这件事情。我们从中可以看到，如果能找到方法为这些突发事件提供缓冲，那么其潜在的价值将会变得非常可贵。其中一种办法是，向穷人提供能构建起储蓄余闲的金融产品。我们可以利用之前讨论过的稀缺管理手段来实现这一点。举例来说，我们可以利用人们的管窥心态，为之提供用来救急的高利息贷款。从“管子”视野中来看，这些贷款非常有吸引力，这样，我们就能帮助人们将同时收取的高额费用储存起来。

更理想的一种方法是，创造出能够避免救急行为的产品。我们了解到，稀缺陷阱和杂耍状态通常都是因为在资源相对充裕时管理不善所造成的。那么，为什么不在资源充裕时提供帮助呢？比如开发出一款金融产品，将农民在收获时赚得的收入进行储蓄，然后再平摊到每个月，从而有效地将一次性收入转换成月收入。这只是诸多方案中的一个例子。从长远来看，我们也为退休财务规划耗费了大量资源。同理，帮助穷人摆脱生活中持续不断的杂耍和救火，也能改变其人生。

上述内容反映了看待贫困问题的一种更为深层次而且与众不同的角

度。这个角度不仅关注显而易见的资源和收入稀缺问题，而且也关注不那么容易觉察但同样重要的资源——带宽。针对带宽的研究显示，就连在“正确的时候”给予现金支持这样简单易行的办法，都会为穷人带来相当可观的利益。如果方法正确、时机得当，那么100美元就能为之换来心灵上的平静。而心灵上的平静能让人将许多其他事情做好，避免犯下代价高昂的错误。马拉维的一个现金补助项目显示，现金支持能为低收入成员减少40%的心理负担。[10](#)如何在正确的时机提供现金补助以及对其深远的影响进行衡量，针对这些问题的研究，能为以带宽为考虑因素的政策制定提供更多构想。

这就需要我们全新定义扶贫政策。我们应认识到人们许多不同行为之间的内在联系。我们知道，房租、食物和学费都是家庭预算的一部分。现在，我们不应该再继续将教育、医疗、财务和子女养育问题分开来看，而是应该认识到这些因素都是带宽容量的一部分。正如财务负担会让我们的预算不堪重负一样，带宽负担同样也会导致我们无法顾及上述所有因素。相反，对某些瓶颈问题的修复会产生深远的影响——儿童的日托服务提供的不仅仅是日托服务，适时的金融产品也远非未雨绸缪而已。这类协助中的每一种，都能释放带宽，提高智商，强化自我控制力，提高思维清晰度，甚至还能改善睡眠质量。听起来有些牵强？我们可是有实际数据作为依据的。

扶贫是个长期项目

一直以来，与贫困问题作斗争都是艰苦的历程。各种扶贫项目，要么以失败告终，要么仅获得了部分成效。社会安全保障似乎总会让人产生惰性。在美国，人们一旦落入社会安全保障的救济范围，就注定会一次又一次地回归接受救济的窘境。[11](#)实践证实，培训项目的效果也很一般。研究人员对这些培训项目所发挥的作用进行了测评，认为培训项目的投入是值得的，却仍然无法改变贫困问题的发展趋势。改善贫困人民的居住环境，也只能发挥一点点作用。美国进行了一项实验，将数千户家庭从低收入社区迁到了收入水平较高的社区，发现改善居住环境的确能产生一定的作用，主要是能减轻人们的压力并提高其生活质量，但潜在的贫困发展规律却无法得到改变。[12](#)

其他国家与贫困问题的斗争，结果也是大同小异。为启动小微企业而提供贷款的小额信贷服务，据称有着很强的变革性。虽然小额信贷的影响是积极的，但研究显示，仅凭小额信贷本身很难改变贫困的内在逻辑。

辑。[13](#)食品补助项目，在儿童学习领域获得了一些成效。虽然教育手段总能搞得风生水起，但能带来的回报是有限的。多年以来，非营利性组织试图通过提供各种周到的扶助措施，以帮助贫困人民解决各种需求和问题。非营利性组织的做法值得赞赏，但同时它们也仅获得了有限的成效。

我们并不是想要对目前的扶贫项目进行批评，毕竟，贫困是一个非常棘手的难题。就算是有限的成效，也值得我们进行社会投入。我们的目的是希望能给出一些建议，将扶贫工作做得更好。每当我们看到那些成效并非十分显著的扶贫项目时，可能就会认为，这些项目所提供的扶助并非人们的真正需求，并不重要。但也许问题并不出在这些项目所提供的扶助内容上，而在于提供扶助的方法。就像第二次世界大战期间的轰炸机驾驶舱一样，也许我们只要通过更加优质的设计，就能提高这些项目的成功率。而更为优质的设计，需要将稀缺心理中所体现的专注力和带宽问题等根本性见解囊括进来。

第9章 如何解决组织中的时间稀缺

时间表排得最满的计划，不是好计划。如果没有构建起余闲，我们就会专注于当下必须完成的工作，预测不到未来可能发生的所有事情。当然，余闲不能太多也不能太少：太多是浪费，太少又不能应对突发事件。只有存在余闲，人们才不会全神贯注在迫近的截止日期上，也会去关注那些重要但并不紧急的任务。

圣约翰医疗中心（St. John's Regional Health Center）是位于密苏里州的一家提供急诊服务的医院。¹这家医院在手术室的安排和使用上遇到了一些问题。每年，医院里的32间手术室要进行的手术多达3万台，而安排手术室就成了一个难题，因为手术室永远都被排得满满当当的。2002年，手术室的使用率达到了100%。当出现需要紧急医治的患者时，医院不得不将很早以前安排好的手术时间往后推。急诊手术在所有手术中的概率大约为20%。一份研究报告称：“由于存在这种现实情况，医生有时要在凌晨两点钟做手术，他们经常会为进行一台两个小时的手术而等待数个小时之久，而其他员工也经常会在毫无准备的情况下加班。”同时，这份报告也对随后发生的那些令人瞩目的变化进行了总结。

手术量超越了手术室的接纳能力，可谓稀缺的经典案例。显然，圣约翰医疗中心掉进了稀缺陷阱而无法自拔。正因为这种落后于时间表的状态，医院不得不将本已安排好的手术往后推，员工们也因睡眠不足和工作规定而烦恼不已，这些都致使其工作效率进一步下滑。在这种情况下，重新安排手术的代价会十分高昂。而且，至少从短期来看，重新安排手术的行为会进一步加剧稀缺问题，因为医院又要“浪费”一部分已经不够用的资源在重新安排手术这件事情上。就像负担过重的人发现完成工作任务的时间太长而想要重新安排时间所带有的顾虑，之所以有此顾虑，其中一部分原因就在于此人的负担太过繁重，无法想象在这种条件下退后一步重新安排时间的代价——这样做就等于额外增加了一项工作，而且还会消耗时间。

但圣约翰医疗中心必须想到应对之法。医院管理层从美国医疗卫生改善协会（Institute for Healthcare Improvement）请来了一位顾问。这位顾问没有带着医院每日压力所导致的管窥心态去看待问题，而是换了一种眼光对整件事情进行分析。他给出的解决办法颇为令人惊讶：留下一间手术室待用。圣约翰医疗中心负责普通外科和创伤外科的肯尼斯·拉森（Kenneth Larson）医生后来在回忆当初自己的想法时这样说道：“我们已经很忙了，他们还要拿走我们的东西。简直是疯了。”

但是，这一建议背后存在着深刻的内存逻辑，对稀缺管理有着指导性意义。从表面来看，圣约翰医疗中心缺乏的是手术室——无论如何重新安排手术，都没有办法解决这个问题。但如果你看到了问题的深层次原因就会发现，医院所缺乏的是另一种东西。医院的手术分为两种：计划之中的和计划之外的。目前，计划之中的手术已经占满了所有手术

室；计划之外的手术一旦出现（而且肯定会出现），医院就需要对现有的手术时间表进行调整。为了解决急诊手术问题而将计划之中的手术改期，需要付出一定的成本：一部分是经济成本，比如加班费；一部分是医疗成本，因为工作人员可能会出现医疗失误；还有一部分是效率成本——让员工在没有准备的情况下加班到很晚，效率一定不会高，而且他们在加班状态下肯定不会表现出最佳状态，每台手术的时间都会比前一台长。

如果没有因急诊手术而导致的重新安排，每个人都能在指定的时间内做好工作，以更高的效率完成计划内的任务，那么手术室的数量就足以容纳所有手术。因此，手术室的稀缺，并非真的是医院缺乏做手术的空间，而是因为其没有能力用现有的手术室来处理急诊手术。这种情况与负债累累的穷人十分相仿，因为穷人如果能安安稳稳地在没有突发事件的情况下使用自己的金钱，那么就完全能够过得更好一些。但事实是，他们的大部分钱都用来还债了。问题不仅在于手头的钱少，而且在于很大一部分金钱要用来偿还之前的债务。在圣约翰医疗中心的案例中的问题不是手术室数量太过“贫乏”，而是当急诊手术突然出现时，紧张的手术室不仅要用来接纳这些急诊手术，还要用来弥补那些为此而延后的手术。

领导这次行动的急救中心副总裁克里斯蒂·登普西（Christy Dempsey）称：“每个人都认为，因为无法对计划外临时出现的急诊手术进行预期，所以将一间手术室独立出来专门处理这些‘临时增加’的手术，是对空间的浪费。”其实，“计划外”或“预期外”手术的说法，存在一定的误导性。这种说法意味着，这些急诊手术的发生是无法预测的。虽然每一次急诊手术的发生都无法提前知晓，但这类手术存在的事实是完全可以预测的，就像那些将穷人或繁忙之人一举击垮的突发事件一样。“预期外”手术出现的规律很稳定，那么，为什么不将一间手术室专门开辟为计划外手术之用呢？这样一来，其他所有手术室就都能顺顺利利地排满日程，不受突发手术的影响，而所有计划外手术则全部被归入了一间特别设立的手术室。

毫无疑问，这种方法发挥了实际成效。

医院意识到改变带来的好处后，也发现了随之而来的一些新变化。医生们开始将手术尽量安排在每周更早一些的时候，以便手术后的查房不排在周末。因为如果手术后查房总是在周末，就会导致非急需施行的手术时间分配不均。没有了急诊手术的干扰，这种不均衡现象就会变得明显起来。没过多久，圣约翰医疗中心就开始将非急需实施的手术平摊到整周的时间里，随之，医院手术室紧张的状况也得到缓解。[2](#)

稀缺实验室

SCARCITY

一间手术室专门用来接纳急诊手术之后，医院的手术接诊率立刻上升了5.1%。每天下午3点之后进行的手术数量下降了45%，医院的收入也出现了增长。实验仅进行了一个月，医院就将这种操作方法纳入了正规流程。随后的两年间，医院的手术接诊量分别上涨了7%和11%。

余闲的重要作用

圣约翰医疗中心的例子，反映了稀缺陷阱的基本特点。医院手术室的缺乏，实际上就是余闲的缺乏。许多系统的正常运转，都以余闲的存在为前提。³过去的卷带式磁带录音机，需要额外多放一节磁带，以确保整段磁带不会被扯断；咖啡研磨机如果装得太满，也无法正常工作；道路在占用率为70%时达到最佳运转状态，堵车就是因为缺乏余闲。理论上，如果道路占用率为85%，所有车辆又均能保持同样的速度匀速前进，那么车辆之间还是能轻松地留下一点空间的。但如果一位司机踩了一下油门后紧跟着又是一脚刹车，那么这位司机后面的每辆车的司机就都要踩刹车。这样，这些车的速度就会全部降下来，而提速远比减速难。又或者，某辆车稍稍偏离了正确的前进速度，然后刹了下车，这么一桩小小的突发事件就会导致交通速度的大幅下降。如果再多出那么几起突发事件，整个路面就会陷入瘫痪状态。道路占用率为85%时，虽然道路的面积足够多，但却没有充裕的余闲去吸收小型突发事件。

就算是那些深谋远虑的人，也常常会低估余闲的重要性。

试想你曾雇用过一位非常出色的助理，他总是时刻准备着去做你安排下去的事情。就算是你临时告知的事情，他也能精神饱满、情绪高昂地完成。但后来，一位管理顾问发现，你的助理拥有许多可随意支配的自由时间。所以，公司的各部门进行了重组，现在，你要与其他两个人共用一位助理。办公室的时间使用数据显示，这种做法更为高效，而助理的时间也排得和你一样紧凑。如今，你的临时事务无法再得到即刻处理。这就意味着，你的日程安排是满满当当的，而只要一个不起眼的突发事件，就能让你落后。落后的你，开始用杂耍的心态去应付各项事务，于是变得更加落后。可见，助理是余闲的重要来源：在你的所有时间都被占满的情况下，助理可以帮助你处理“突发事件”。就像圣约翰医

疗中心空出的一间手术室一样，余闲没有得到充分利用，也自然无法保证其在关键时刻能发挥价值。

面对太多等待去做的事情，人的第一反应就是将时间表排得尽可能紧凑，以便将所有事情都安排进去。如果你安排得不够紧凑，就会感觉好像没有尽力去做事情一样。当效率专家发现员工们没有得到充分利用时，总会敦促他们“更有效率”地利用时间，但这样做就意味着他们会丢掉余闲。⁴时间表排得很满时，如果因交通堵塞而耽误在路上，就会将所有安排全盘打乱，而交通堵塞对一般人来说不过是有点心烦罢了，不会真的误事。你因为交通堵塞，而在一号会议上迟到。因为一号会议和二号会议之间没有预留时间，所以二号会议你也无法准时出席。继而，三号任务也受到了影响。最后，你束手无策，只得将当天日程表中的一项任务推迟到第二天去做，可第二天的日程表也安排得相当“高效”，于是乎，一项任务的推迟最终让你付出了高昂的代价。这听起来似曾相识？没错。因为你也是这样低估余闲的价值的。最不起眼的小差错都会酿成你无力承担的重大责任，将明天的预算提前借用到今天，是要付出昂贵的代价的。

我们之所以没有能力构建起余闲，是因为专注于当下必须完成的工作，预测不到未来可能发生的所有事情。当下的工作迫近而清晰，而未来的可能性并不给人以紧迫感，也难以想象。当无形的未来与具体的当下产生冲突时，余闲就成了奢侈品，这就是你觉得自己没有资格去挥霍的原因。那么，你应该怎么做？你是不是要在时间太少、任务太重的情况下，还刻意在日程安排上留下空白，比如周一到周三的下午3~4点，以此来预防意料之外事件的发生？答案是肯定的。同理，原本30分钟就能赶到的地方，你应该预留出40分钟；在每个月的家庭开支之外，你也应该存点钱，以备不时之需。面对稀缺的威胁时，余闲是必需品，而我们却总是不能将其列入计划之内。很大一部分原因在于，稀缺令我们很难预留出余闲。

余闲，不能太多也不能太少

余闲的处理失当，不仅与个人有关，与组织也有着密不可分的关系。20世纪七八十年代早期，有一种观点认为许多公司太过于“臃肿”。⁵某些行业拥有的现金太过充裕，致使主管人员挥霍无度。在房地产和商业并购交易上，他们支付的价钱远超实际价值，从不讨价还价，对预算的底线也毫不在意。乱花钱的结果就是，有些石油公司的市值比

他们拥有的石油价值还要低。（市场假定，这些公司只会浪费自身的资产。）20世纪80年代的融资收购风潮，其初衷正是为了解决这一问题。风潮背后的逻辑很简单：收购这些公司，然后使其陷入债务压力。将这些公司从资源充裕的状态转化到稀缺状态。债务所带来的惩戒（即我们所谓的稀缺引发的关注力）会改善公司的绩效。主管们会因此开始投入关注，在花钱时更加谨慎，并为公司带来更高的收益。

事实上，大量实证研究显示，融资收购虽然饱受争议，但的确能提高公司的业绩。⁶一个原因在于，公司的“赘肉”加剧了职业经理人所面临的激励问题的严重性。他们之所以乱花钱，是因为他们花的是别人的钱。这种“赘肉”实际上就是让管理层挥霍享受的天降横财，但从股东的角度来看，这种享受对公司的发展毫无益处。通过融资手段减少“赘肉”，职业经理人就能在花钱时更加明智。

由于稀缺心理的存在，融资还有另外一个作用。公司实现精益化、节约化背后的原因，与截止日期能提高生产力、低收入旅客更了解出租车费的道理相同。职业经理人需要时刻保持高度警惕，努力降低成本，提升认知能力。他们必须与供应商进行艰苦的谈判，仔细检查每一个细节，以确定每笔支出是否必要。这种专注力在稀缺状态下是自然而然的，而在资源充裕的情况下则相对困难。就算在民营企业，职业经理人自掏腰包时，如果现金充裕，他们也会花钱大手大脚的。

正如我们所了解的一样，余闲既是一种浪费，也能带来收益。采取精简措施时，我们很难将真正的浪费从有用的余闲中剔除出去。事实上，在融资并购风潮过后，许多公司都被逼到了破产边缘。⁷面对这种现实，公司主管们产生了管窥心态。如果说人们在20世纪80年代习得了企业精简的经验教训，那么21世纪则给人们在管理目光短浅带来的危险方面上了一课。不过，两者之间也存在着某种联系：如果太过精简——削除了太多余闲，那么经理人就只能用按揭未来的手段来弥补当下的青黄不接。

4.45倍！火星探测器的教训

1998年12月，美国国家航空航天局（NASA）将火星气候探测器发射升空。⁸火星任务负载着人类数百年来对这个星球的向往。火星与我们的距离如此之近，其体积与地球非常相似（就连一天的时间长短都差不多），虽然存在生命的可能性不大，但让人很想一探究竟。探测器本

身不太可能取得什么重大发现，但它起到了先锋作用，能为将来载人登陆火星的任务提供宝贵数据。火星探测器的发射，是这项价值1.25亿美元项目的顶点，是专家们数万个小时辛勤努力的成果。发射升空后，火星探测器按计划进入了临近火星的固定轨道，在那里执行数据搜集工作。

将卫星置于某个行星的固定轨道，是一项颇为棘手的工作。卫星越接近行星，重力作用就会越大，越有可能被吸进去。如果卫星速度过慢，重力作用就会令其坠毁于行星表面；而如果卫星速度过快，重力作用就会变弱，则会导致卫星在行星周围逆方向滑行。只有在正确的速度和角度上，重力作用才能够刚好将卫星安置于固定轨道上。可以想象，确定恰当的速度，需要极为复杂而精密的计算工作。随着探测器接近火星，点燃反向推进器的任务就变得紧急起来（以便将速度降下来），而减速的时机和尺度又要能恰好让探测器落入火星轨道之中。信号从地球传到火星需要10分钟，因此所有程序都要事先被安装好。地面控制人员能做的，只有等待和（延时）倾听。所幸，寂静的宇宙空间并不会出现什么意外。天体物理学的计算结果要求非常精准，地球上的工程师们再怎么仔细都不为过。

1999年9月23日，火星探测器发射升空九个半月后，终于到达了火星，开始执行进入程序。执行程序时，探测器位于火星背面，这种位置关系会致使地面和卫星失联几分钟。但随后，真正的麻烦出现了：时间一秒秒地过去，按计划探测器早就应该重新绕到火星前面，但地面人员却收不到卫星的任何信号。人们的希望也随着时间的流逝而泯灭。最终，地面工作人员只得假定火星探测器已经坠毁。

如此重大的事故发生后，必然要进行详尽的调查。究竟发生了什么？为什么NASA将之定性为坠毁事故？如果事前做了哪些工作，就有可能避免事故的发生？应该将责任归咎于谁？故障，尤其是复杂系统的故障，通常存在许多方面的原因。而在这场事故中，罪魁祸首既有很高的报道价值，又非常显而易见。真正的元凶在于，反向推进器的“火力”太过强劲。但最值得深思的是，反向推进器点火的级别——NASA计算出来的点火期望值与实际值的比率，是一个为人所熟知的奇特数字：4.45。这个数字是公制与英制换算时的系数。所以，令人难堪的错误很快就浮出了水面。

火星探测器这种卫星通常是由几家分包商逐块完成的。负责制造推进器的公司，以英制的磅为单位理解接收到的数据；而负责制造中央处

理器的另一家公司，则以公制的牛顿为单位发送数据。所以，每次处理器发出“X”时，推进器制造商都将之理解为“4.45倍的X”。（比如处理器发出的数据为10，意思就是10牛顿，但推进器却将之理解为10磅，相当于44.5牛顿。）结果，探测器减速过度，没能逃过火星引力的吸附。对于如此规模的项目来说，这种荒唐的错误不仅可笑，更会带来严重的后果。

失误是无可避免的，NASA的工程师当然了解这一点。这就是为什么在火星探测器发射前他们要进行无数的检验和测试工作的原因。那么，究竟发生了什么？原来，在探测器发射之前的几个月，喷射推进实验室的整个团队都落后于工作进度一大截。实验室的人手太少，无法将关注力全部投入项目细节中，等他们发现问题时，已经太晚了。每个人的工作进度都落在时间表之后，而处于救火状态的公司，总是会为新项目分配小团队——大部人马都在忙着扑救最近的一场大火。而航空航天领域的工程师们所面对的截止日期是确定的，不像其他行业那样可以往后推。天体轨道决定了发射日期、火星和其他天体的位置，只能给卫星发射留下很短的窗口期。在天文历法问题上讨价还价根本是不可能的。

迫近的截止日期使得工程师们只能加班加点地赶工，同时也使他们产生了管窥心态。人们专注的问题是，确保发射日期前完成所有工作。与此目标没有直接关系的事项，全部被搁置下来。而事实证明，这些事项后来也没有人再去过问。4.45的错误，就这样发生了。在发射之前，工程师们自身搜集到的数据就表现出了一些不对劲——他们注意到了前后不一致的现象，但要想找到原因，就意味着庞杂的待办事项中会再平添一项。他们还没有去做的事情太多了，不是每件事情都能做到。对明显的前后不一致进行源头追溯，就是给自己多增加一份新任务。另一件受牵连的事项，就是推进器和处理器的联合模拟。如果在发射之前进行了联合模拟，就能使问题直接暴露出来。为了赶时间，团队牺牲了常规的检查与测试，忽略了提示问题可能存在的信号。现在，我们能清楚地看到，悲剧就是由管窥心态所引发的，而这种关联具有紧密的逻辑关系。

这并非事后之见。一份于坠毁发生之前提交的NASA喷射推进实验室报告，就反映出了问题的关键。报告显示，项目一开始的落后（可能是由于人手不够），会导致团队在遇到问题时选择走捷径。团队成员每天都要加班加点，错误肯定会存在。最初的落后会导致更多的低效现象。而更严重的是，关键的监测也会被忽略，因为这些活动在当时看来

根本没有那么紧迫。显然，这份报告预见了导致力学单位混淆和坠毁事故的工作模式。

这不仅仅是落后于工作进度表所引发的征兆。火星探测器事故调查人员在完成技术分析之后，转而继续去寻找事故的组织因素。他们发现，其中一个原因就是，NASA之前一直标榜的是“更快、更好、更廉价”的口号。这一口号将重点放在了成本节约和进度缩减上。正是这种工作作风，致使团队工作的时间出现短缺，并产生了管窥心态。之后，工作人员便开始有意识地忽略一些事项。在火星探测器的案例中，工程师们忽略的是关键的检测，因为这些检测虽然重要，但并不紧急。当时，工程师们手头最紧迫的任务就是将探测器准时发射升空，考虑到这项任务的紧迫程度，检测工作就只能靠边站了。

用余闲对抗救火陷阱

圣约翰医疗中心和NASA都掉进了救火陷阱中。组织学专家罗格·博恩（Roger Bohn）和拉姆钱德拉·贾库玛（Ramchandran Jaikumar）认为，[9](#)救火型组织有几条共同特征：第一，这种组织“问题太多，时间不够”；第二，这种组织会解决紧迫问题，但对于非紧急事务，无论其重要性有多高，都会被搁置；第三，这种行为会导致层叠效应，致使要去完成的工作量变得越来越大。简而言之，它们的时间都用在扑救最迫近的火灾上了，而四处又会不时地冒出新的火焰，因为没有人采取预防火灾的措施。在圣约翰医疗中心，外科医生们都忙着处理急诊患者，无法腾出时间去审视病人群体的结构；在NASA，工程师们都忙着在截止日期之前完成每一个部件的准备工作，却没有去研究这些部件是否能融为一体。可见，救火陷阱是稀缺陷阱的一个特例。

一项为期5年的研究以4家美国顶尖制造公司为研究对象，其间记录了许多次救火事件。就像一位经理人讲到的一样：“看一看我们在传统项目上的资源分配就会发现，我们总是起步迟缓，无法将工作人员快速安排到项目团队中……之后，我们又会在项目上投入大量人力……在项目启动之时，资源分配达到了顶峰。”[10](#)研究人员在多年研究的基础上给出了如下总结：“现今关于研发管理的讨论中，最为司空见惯的话题就是，工作量超负荷的工程团队会在项目启动前的最后阶段没日没夜地赶工。”

救火不仅仅会导致失误，而且还会导致完全可以预见到的失误类型：重要但并不紧急的任务会被人们所忽略。顾名思义，救火行为就是忙着去扑救最紧急的问题（火灾），其他问题，无论有多重要，都会被最紧急的事务所掩盖（如赶赴火灾现场时不系安全带）。结构性问题很重要，但却可以等待，于是就永远被搁置了。微软公司发布Windows

2000软件时，已知的漏洞有28 000个。[11](#)项目团队知道他们推出的产品存在许多问题，但他们已经落后于产品发布的最后期限了。结果，他们立即开始投身于第一个补丁的研制，以修正产品发布时已知的所有漏洞。此时，关于新漏洞的报告已经开始源源不断地出现，而团队还在修补之前的问题。

救火陷阱意味着大量的杂耍。你全神贯注在迫近的截止日期上，当终于完成手头的工作时，才突然意识到，下一个项目马上也要到期了。我们很多人都曾经有过这种切身体验，凭直觉就能知道，救火陷阱之所以是“陷阱”，其原因与稀缺陷阱相同——只要开始救火，就很难毫发无伤地走出火灾现场。当团队在一项早该完成的工作上疯狂赶工时，下一项任务的起步就必然会被延迟，而落后就意味着团队成员同样要去扑救下一项任务的火势，并从此一直落后下去。[12](#)

在了解了稀缺的内在逻辑之后，我们就能明白：余闲能降低我们进入救火陷阱的概率。而我们也知道，管窥心态会令人很容易忽略其他考虑因素。对组织而言，有一种解决办法就是，确保余闲的存在，并对余闲进行明确的管理。从银行管理风险的教训中，我们能更深刻地领悟到这一点。一直以来，银行都深知抱着管窥心态看待账目底线的经理，没有办法充分地将风险因素考虑在内。2008年的金融危机证实，这种认识太过保守。最近，许多银行都引入了“首席风险官”（Chief Risk Officers, CRO）这一职位。首席风险官与管理团队的其他成员相互独立，直接向CEO汇报工作。他们会用风险评估的眼光，去审批金融产品、贷款和其他交易。这一职能与其他银行家不同，许多银行经理人都专注于（用管窥心态看待）最具吸引力的交易，总是想赚到最多的利润，实现最大的销售目标；而首席风险官的唯一目标，就是监管风险。

组织“赘肉”一点一点地被削掉，随之一同被带走的，还有余闲。在这种情况下，组织很希望能拥有一位干将，不是整天想着如何将一分钱掰成两半花，而是能超脱于每日的管窥心态，确保组织拥有足够的余闲。他所关注的问题，不是今天需要完成什么工作，而是明天可能会出现什么样的突发事件，而这些突发事件又会如何影响紧张的工作计划。一定要有这样一个人，他的存在，能够保证专注于当下工作目标的人们，不会向未来的项目借用资源，并因此耗尽余闲，将组织拉入未来的带宽陷阱。

亨利·福特的时间管理智慧

真正有效率的劳动者，不会整天马不停蹄地工作，而是闲庭信步般轻松愉悦地处理事务。[13](#)

——亨利·戴维·梭罗

NASA的案例中，还有另一点值得我们学习。喷气推进实验室的工作人员刚开始落后于计划的进度时，管理人员做了绝大多数管理者都会做的事——他们延长了工作时间。他们看到了时间的稀缺——火星探测器马上就要发射了，只好用更多的时间解决问题。这是人们应对时间稀缺的一种惯常方法。项目落后于进度了？那就在项目里多安排几个人，以解决问题。如果组织内部能灵活调遣的员工数量有限，时间很紧迫，而聘请、培训新人又是件很花时间的的事，那么就不妨在新员工入职之前，让现有的项目组成员加班加点。表面看来，这似乎是显而易见的解决办法，是在拥有定量资源的前提下取得更多成果的最简单的途径，但是，如此的应对手段，并非最明智的做法。因为这种做法只认识到了“完成项目所需时间”这一种形式的稀缺，而忽略了另一种形式的稀缺——带宽，完全忽视了被削弱了的带宽对工作成绩造成的负面影响。

我们以手机的使用为例。美国现在有10个州禁止开车时使用手持电话。[14](#)该规定颇有道理，其他州早晚也会效仿。毕竟，在只有一只手控制方向盘的情况下，我们的驾驶效率和反应速度肯定会有所下降。但同时，这里也隐含着一个重要假设，即手持手机打电话的司机比正常驾驶的司机更容易出交通事故。[15](#)但事实上，用蓝牙耳机打电话的司机也同样难逃厄运。[16](#)问题其实并不出在手上，而出在大脑中。一项模拟研究证实，使用非手持电话通话的司机，他们闯的红灯数量，是不打电话司机的两倍多。[17](#)我们总是自然而然地认为，驾驶是一项体力劳动，但安全驾驶需要的不仅仅是两只手而已，更需要带宽（脑力）的支持。

我们在安排时间时，也总是会忽略带宽。我们自然而然想到的是完成待办事项所需要的时间，而非完成这些事项所需要或占用的带宽。请想一想实验室的工程师们在卫星探测器发射日期迫近的压力下，是如何对存在的问题作出回应的——他们针对这些问题，投入了更多的工时。但这样做并不能提供更多的带宽，而且很有可能的一种情况是，虽然工作时间变长了，但疲劳不堪的工程师们在工作上投入的总带宽却变少了。

大约一个世纪之前，亨利·福特（Henry Ford）就认识到了工作时长与带宽之间的关系。他为工人定下了每周40个小时的工时规定，不仅是出于利润方面的考虑，而且也饱含了人文主义情怀。[18](#)而他进行的为期12年之久的实验，其结果明确显示，将工作日的工作时间从10个小时缩短到8个小时，将每周的工作天数从6天减少到5天，既能增加工人的总产出，又能减少企业的生产成本。福特积极宣扬缩短工作周能为社会带

来的益处，坚定地认为增加工人的消费时间对所有人都是一件好事。而他的看法中最核心的一点就是，缩短工作时长能带来更多产出。

若想找到福特实验的原始数据，难度会比较大。但自从福特实验之后的一个世纪以来，学者们发起了几项类似的研究。其中一项是以建筑项目为研究对象，其结果显示：每周工作时间在60个小时以上，持续超过两个月时，员工的生产力会下降，导致完工日期拖延；而同样规模的工程团队，如果每周工作40个小时，则能按时完工。[19](#)在与建筑行业搭不上边的软件业，一位软件开发师注意到，当他的员工刚开始一周工作60个小时，前几周时间总是能比以前完成更多的工作，但从第5周开始，员工们能完成的工作量，会比他们每周工作40个小时的时候少得多。[20](#)

另一项研究以医院的胸心外科诊室为对象，调查的是随着每位医护人员接待患者数量的增多而产生的现象。[21](#)同样，短期来看，单人接待更多患者会提高生产力，医生也能以更快的速度治疗患者。但这样做是要付出代价的：在忙着处理病人的时候，医生会疏忽大意。为了更快地诊治更多患者，医生的诊疗质量会下降。由此，患者的死亡率就会上升。而事实上，就连短期所见的好处也不能持久——工作量的持续增加，最终会导致管理每一位患者所用的时间也都有所延长。

超时工作对生产力的负面影响，同样能表现在其他方面。下面是一位研究人员对创新的论述。

每次面试结束时，我都会向面试者提问：“如果你突然间变得无所不能，那么首先会在自己所在的组织中采取什么样的做法来鼓励创新？”截至目前，最常见的答案就是时间。但给出答案的人们还会对时间有所明确，他们不需要同样类型的更多时间，而需要更多的非结构化时间，或者没有具体的产出或流程与之相对应的时间。常务董事所渴望得到的东西，最能生动地体现这一点：“玩乐的时间……凝视窗外的时间……让事情尘埃落定的时间……观察、理解和作出反应的时间。”[22](#)

从某种角度来看，这并不新奇。因为我们的身体会疲劳，需要休息。同样，心智资源也会耗竭，需要恢复。而稀缺的长期存在，则会导致带宽负担越积越重。为了对这一机制加深了解，我们不妨以睡眠为例。时间稀缺的人，每天都工作很多个小时——他们总是想要在有限的时间里尽可能多地做一点事。他们会忽略问题，遇到了就打个“补丁”了事。睡眠就是受到影响的一大因素：当时间不够用时，人们就会选择少睡点觉，多工作几个小时。但事实上，睡眠对生产力的影响非常惊人：科学研究多次证实，睡眠不足的员工，其积极性会变得更低，犯的错误会变得更加多，而且还经常会出现头昏脑涨的现象。[23](#)一项研究对夏令时

的起止进行了调查，结果就真实地反映了这一点。因为时间的变化，人们晚上会失眠。在失眠的夜晚，人们每个小时花在网上闲逛的时间就会多出20%。[24](#)而这只是一个晚上，长期睡眠不足所累积的效应会变得更加严重。随着工作时间越来越长，睡眠时间越来越少，生产力最终会一落千丈。

然而，大多数公司依然会对工作时长进行管理，而非对带宽进行管理。一支研究团队曾针对一位38岁、已婚、有4个孩子的大型会计师事务所合伙人做出了如下描述：

一年前我们认识他时，他每天工作12~14个小时，总是感到精疲力竭，觉得自己很难在晚间时光全心投入与家人的交流，并因此感到愧疚和难过。他的睡眠质量很差，没有时间锻炼身体，很少吃健康食品，总是随便买点什么在路上边走边吃，或是坐在办公桌旁边工作边进餐。他的经历并没有什么特别之处——许多人都用延长工作时间的的方式来应对工作上不断增加的压力和需求。但是，这种做法会不可避免地使我们在身体、精神和情感上遭受负面影响，而这又会导致员工的工作积极性下降、分心度提高、离职率增长和医疗费大幅上升。[25](#)

这支研究团队尝试了一个叫作“能量管理”的实验项目，项目包括工作间歇外出散步，将关注点集中在睡眠等重要因素上等。[26](#)在研究过程中，他们发现，12家银行的106位员工在多项绩效参数上都有所提升。乍听起来，我们也许会觉得能量管理与绩效提高之间的联系有些牵强附会。但这与我们管理自己身体的方式又有什么不一样呢？为了避免重复性劳损，经常使用计算机的员工要进行强制性的休息；为了缓解计算机视觉综合征，专业人士建议人们每工作20分钟就转头看向别处20秒钟，让眼睛得到休息。[27](#)从这一点来看，认知系统其实与我们的身体系统并无太大差别。

我们从中领悟到的更深层次的信息就是，虽然稀缺会给我们带来很大压力，但我们依然需要专注于能量管理和带宽培养。增加工作时长、让员工更加卖命地干活和放弃度假机会等，都是管窥心态的表现。就像以高利率借钱的行为一样，这些行为都忽略了长期后果。精神病学家做出的报告称，越来越多表现出急性应激症状的患者都“将自己逼到了极限，甚至超越了极限，没有一点儿余地，生活中没有一点儿空间能用来休息、放松和反省”。[28](#)无论是一周工作40个小时，还是一周工作50个小时、60个小时，都无法产生任何神奇的效果。但让自己暂时换换脑子休息一下，却非常重要。因为真正重要的是，将有效的带宽最大化，而非将工作的小时数最大化。

当然，从救火行为到带宽维护的失败，所有这些错误都是个人问

题。每个人都有可能会犯下这种错误，但组织却有可能将这些问题放大。当团队中的一位成员开始落后，或进入救火模式时，就会加剧团队中其他成员感受到的稀缺。当一个人的带宽产生了负担，特别是沉重的负担时，他所做出的一系列拙劣决策，就会导致更为严重的稀缺，继而为他人的带宽增加负担。组织很有可能会产生一种多米诺骨牌管理效应，每一位个体成员都会将整个团队拉向陷阱，使带宽锐减。不过，组织同样可以富有洞察力，以创建出一种环境，使管理稀缺的挑战更容易获得成功。

“红花”日本料理的启示

青木广彰（Hiroaki Aoki）与许多美国创业者一样，有着一段狂野的青春岁月。20世纪50年代，生活在日本的他是一名放荡不羁的少年。在学校，他向同学们贩卖黄色录像带，还创建了一支名为“劲歌之声”（Rowdy Sounds）的摇滚乐队。同时，他也表现出了严谨的一面。他是一位轻量级摔跤选手，通过艰苦不懈的努力，赢得了参加1960年夏季奥运会的资格，获得了去美国读大学的体育奖学金，并最终赢得了轻量级摔跤奖牌，在摔跤名人榜上有了一席之地。随着日渐成熟，他将丰富的创造力、充沛的精力和勤奋努力的精神转移到了商业领域。在摔跤比赛的间歇，他获得了餐饮管理的准学士学位，还利用闲暇时间在纽约哈林区开着冰激凌卡车四处兜揽生意。[29](#)

青木广彰最为成功的创业史，也经历了从无到有的过程。利用从冰激凌买卖中赚到的10 000美元，他开办了一家名为“红花”的日本牛排馆。餐厅位于纽约城的西56街。最初几年，生意起起落落，但餐厅慢慢因美味的食物和特别的气氛而吸引了越来越多的客人，最终成了名人的聚集所。（拳王阿里和甲壳虫乐队的成员都在那里用过餐。）青木广彰利用第一家餐厅的成功，将红花品牌扩展成了连锁店，一开始在纽约城开办，后来推广到了全美乃至全世界。如今，红花日本料理已经遍及全世界17个国家。2008年，青木广彰与世长辞，而据估计，他的商业帝国价值上亿美元。他的个人色彩十分鲜明，与其赫赫大名并存的，还有其生父确认诉讼、家族内诉讼、一大堆古董车、许多稀奇古怪的癖好，还有富有民族色彩、略带神秘感的连锁店名字的由来（第二次世界大战期间，美国轰炸东京时，青木的父亲在废墟中看到了一朵红花）。

每个去过红花日本料理的人，都能感受到这家餐厅的独特之处。厨师们会在你面前为你烹饪菜肴，而事实上，“烹饪”一词远不能形容厨师

们的精湛技艺。红花日本料理的厨师们简直堪称艺术大师：刀具在空中飞舞，用小铲将食物直接抛向你的盘中，还能创造出洋葱圈火山！在红花餐厅，一餐的结束会以掌声收场。如果在网上搜索“红花”或“日本烤肉厨师”，能看到与之相关的数百段视频，点击量数以万计，其中展示了烹饪技艺的精妙。所有这些，都间接地为红花日本料理的成功做出了贡献。青木广彰做到的，远非创造出了一点点娱乐效应而已。他深刻地理解了餐饮行业所面临的稀缺问题，并且想出了解决之道。

人们总是认为，餐厅的制胜法宝就是菜品、装潢和服务。毕竟，我们作为客人所体验到的，无非也就是这些。但我们都知道，有些在这些方面表现得非常不错的餐厅还是无法逃脱倒闭的命运。让客人走进门，并不能确保餐厅的成功。驱动餐厅盈利的，是枯燥乏味的物流和运营决策。餐厅所面临的问题在于，大多数成本都是固定的。当然，餐厅要将钱花在食物上，但原料的成本并没有日常开销那么多。日常开销一般包括员工的工资、店面的租金、电费、保险等。无论你是给一大群客人服务，还是只有零星几位客人，日常开销的总数都不会有太大出入。做生意就是要想办法捞“油水”。当收入的数额能够抵消固定成本之后，剩余部分中很大一块就可以被直接划入收益。如此看来，我们就能算出一笔很有意思的账。繁忙的周六晚上，如果有3个座位，那么这3个座位的收益能力就远远不只超过2个座位的50%而已。如果2个座位就能抵消固定成本，还能给你留下一点点收益，那么第3个座位就是“油水”——从第3个座位处赚到的收入，基本上全是利润。

青木意识到，餐饮行业真正稀缺的是座位资源。作为管理者，你能在餐厅中摆放多少座位？如果能布置更多的餐桌，就能有更多的座位。如果每张餐桌能容纳更多人，就能有更多的座位。如果能以更快的速度实现食客的翻桌，每天晚上每张餐桌招待4轮客人而非3轮，那么也就意味着有更多的座位。

红花日本料理上演的厨艺表演，实际上是解决稀缺问题的一种非常聪明的办法。厨师的表演，需要食客们坐在公用餐桌边观赏。供8人进餐的公用餐桌，就能更加有效地容纳客人。如果4人同去用餐，就不用等待临近的两对人用餐结束。有了公用餐桌，只要让客人彼此挨着坐下就好。4人餐桌，不过是一张桌子周围摆放4把椅子。而且，红花日本料理的翻桌率也变得更高了。厨师在你面前上演精妙绝伦的厨艺，而且动作麻利。你走进餐厅找个座位坐下，厨师就会站在中间时刻待命。菜单十分简洁，点菜的时间也有限制。然后，厨师会欢快地一步步将菜肴为

你烹饪好，将食物抛向你的盘中。你必须很快吃完，因为这样才能看到下一道即将抛向你盘中的菜是怎样做好的。就连甜点冰激凌也是怀着让客人快速吃完的心机而设计的，因为在红花日本料理这样的开放式烹饪环境中，冰激凌很快就会融化。烹饪表演结束时，厨师鞠躬，客人鼓掌，用餐结束。之后你还想做什么，继续坐在座位上咬筷子？厨师就站在那里，一切结束，服务员早已将桌子清理干净，其他人也都准备离开，在这种情况下，实在没有人愿意磨蹭。而这就意味着红花日本料理每晚从每张桌子上能赚到更多钱。有人估计，红花日本料理与其他餐厅相比，每1美元收入中的盈利都要多出10美分。[30](#)

“舍维斯”的餐桌管理经验

除了精心设计的餐点以外，红花日本料理的案例还为许多组织提供了重要的借鉴经验。就算商业组织拥有足够的洞察力，能够识别出哪些才是真正稀缺的资源，也常常无法真正理解稀缺管理的复杂性以及做好稀缺管理所能带来的利益。

谢里尔·吉姆斯（Sheryl Kimes）是康奈尔大学的运营研究专家，她曾为一家名为“舍维斯”的连锁墨西哥餐厅提供咨询，帮助其提升盈利能力。[31](#)首先，她与餐厅员工交谈，以便更好地了解餐厅当前面临的挑战。其中一个问题显现了出来，那就是排队问题。从某种角度讲，餐厅门口排起长龙是一件好事，因为这说明餐厅的生意很红火。但排队也不一定是一件好事，因为餐厅需要的是人们在餐厅里用餐，而非在餐厅外等待。客人们很可能会因此而忍无可忍，从此不再光顾这家餐厅。而你也自然不希望人们引用尤吉·贝拉（Yogi Berra）的话这样谈论自己的餐厅：“没有人再去那里吃饭了，那里人太多。”[32](#)究竟是应该提高菜品的价格，扩张餐厅的面积，还是利用其他办法？吉姆斯进行了深入的统计学分析，从中了解到了比员工议论更为精确的要点：每张桌子的收入是多少？哪些桌子最常被客人占用？餐厅的翻桌率是多少？等等。

答案使吉姆斯也很震惊。虽然一眼望去总能看到这家餐厅门口排起了长龙，但实际数据却显示出餐桌的使用率很低——每周只有5个小时时间，餐厅里面会有一半以上的餐桌被客人占用。但店门口排队的时间可远远不止每周5个小时！到底发生了什么？数据中的两个线索提供了问题的答案。

首先，用餐时间上存在着巨大差异，而最大的差异发生在一餐结束

和下一餐开始之前的这段时间。就算在繁忙时段，一张餐桌连续接待的两场晚餐之间，都会有长时间的空白。其次，虽然像舍维斯这样的餐厅已经被人们公认为朋友和同事聚会的场所，但实际数据却给出了另一番解释：70%的“聚会”都是一位客人或两位客人，但餐厅并没有为它所招待的聚会布置下适宜的餐桌。为了验证这一结论的正确性，吉姆斯将聚会数据输入了一个特定算法，以为舍维斯餐厅寻找最高效的餐桌布局，看看不同大小的餐桌各需要多少张。她最后得出了一个清晰的结论：餐厅中应该布置更多适合两人用餐的餐桌。餐厅管理层采纳了这一建议，结果收入突飞猛进——销售额增长超过5%，仅一家分店一年的收入就增加了12万美元。当然，购买新餐桌和改造餐厅布局结构等，并非不需要成本，但最终从会计提供的数据上来看，采纳新建议的第一年，利润就超过了成本，从第二年开始，就转化成了纯利润。可见，在稀缺管理上进行投资，可以赢得很高的回报率。

在吉姆斯介入舍维斯连锁餐厅的业务之前，这家餐厅都没能很好地对稀缺进行管理，因为领导层低估了稀缺所带来的挑战。这些挑战并非微不足道的琐事，因为仅一家餐厅的问题就需要动用大量计算机分析工作，需要进行严谨的研究。餐饮业并非唯一被稀缺问题所累的行业，很多商业机构常常会因为稀缺管理方式的优劣而沉浮于商海。

第10章 如何应对日常生活中的稀缺

通过外力将人们所处的环境进行小小的改变，把重要的事情拉入“管子”视野，就能缓解稀缺带来的不良后果。当带宽有限时，让“疏忽”等同于“默许”就是一个不错的办法。稀缺会产生带宽负担，因此节约利用带宽是对稀缺进行管理的重要内容。应对稀缺，要在带宽充裕时开始行动。更重要的是，要留有应对突发事件的余闲。

医生和有线电视安装工人有一个共同之处：约在某个时间要做的事情，很少会真的兑现。按照约定时间做事情，确实很有难度。如果之前出现了一些疏忽，比如有点磨蹭，或某件事情所花的时间比预期的要长，而同时又没有余闲去抵消这些意外，那么影响就会被放大。这样一来，一开始看似在你掌控之中的紧密安排，就会引起连锁反应，导致接下来的事情也一件件地被推迟。你赶赴的每一场约定，也都会匆匆忙忙。于是，你产生了管窥心态，一心只想着赶快完成手头的事情，然后接着去做下一件事情。可以想见，你会从未来的安排中借用时间。就这样，你给自己挖了一个越来越深的时间债陷阱。而排满各项任务的日程表，很有可能会让你在赶赴每项任务时迟到。而且，在大多数情况下，你过不了多久就会开始迟到（而客户为什么会容忍你这一点，就是另外一回事了）。

我们的一位同事，是一家基金会的理事长。他对排满各种事项的日程表再熟悉不过了：基本上每一天，他都要马不停蹄地赶赴各种会议。他很容易就会像医生或有线电视安装工人那样陷入落后状态——每一次会议都会比前一次更晚。（但因为人们都是来找他投资的，所以不得不忍耐。）事实上，他的会议从来不会拖延——会议原定的结束时间之前5分钟，他的助理就会露面宣布：“会议还剩5分钟。”会议结束的时间一到，助理会再次宣布。这种比较直接的干预方式，能防止事态的进一步恶化，也能预防稀缺陷阱的出现。这种方式为许多主管人员所用，而这些主管人员都配有经验丰富且敬业的助理为他们提供帮助。

助理去敲门，并非什么特别有创意的干预方法，但我们却能从中体会到某些深刻的意义。对某人所处的环境进行小小的改变，就能缓解稀缺所带来的不良后果。稀缺心理是原生的，若想从“内在”进行改变，难度很大。但我们若想获得更大的成效，其实并不需要对这种心理进行改变。事实上，这位基金会理事长的管窥心态一点也不比别人少。但他的技巧在于对环境进行了改变，从而对管窥心理产生了反作用。而且，这种改变的程度也比较缓和——助理没有创造出额外的余闲：会议的安排依然一场接着一场，理事长在赶赴一场场会议时，依然会抱着管窥的眼光看待事物。助理做到的，不过是采取了干预手段，阻止稀缺心理对现状造成实质性的破坏。我们可以将其视为类似于公路旁边的减速带。减速带只是公路上的一点点小变化，但却能保护司机免受走神或疲劳驾驶的危害。减速带的设置，比想方设法地让司机专心开车或是多睡点觉要容易得多。

同样，我们也能对自身所处的环境进行“稀缺防御”。我们可以引入类似于减速带或助理这样的工具，利用我们对事情本身的洞察力，来构建出更为理想的成果。关键在于你所采用手段的内在逻辑，在于你是否真正理解了稀缺能帮助我们用不同的方式进行思考，以及帮助我们管理长久以来得不到解决的问题。

把重要事情拉入“管子”视野中

对稀缺进行管理，有一种简单而常常为人们所低估的方法，那就是对“管子”视野之内的事物施加影响。这就是助理能做好的事情：当主管的“管子”视野还停留在当前的会议上时，他可以将下一场会议带入主管的视野之中。我们与经济学家迪恩·卡兰、玛格丽特·麦康奈尔

（Margaret McConnell）和乔纳森·辛曼（Jonathan Zinman）合作，尝试将储蓄的理念带入生活在玻利维亚、秘鲁和菲律宾的穷人的“管子”视野之中。¹

我们认为，穷人之所以储存不下钱，一部分原因就在于他们的“管子”视野——储蓄是一件重要但并不紧急的事务，而这类事务基本上总是会被穷人遗留在“管子”视野之外。无论何时，总会有比储蓄更为紧要的事情等着被处理。于是，我们将储蓄设定为人们关注的重点问题，将其带回“管子”视野之内。每个月月底，我们都会发一条短信或一封邮件作为提醒，询问人们的储蓄目的以及金额。仅凭如此微不足道的提醒，我们就将人们的储蓄额度提高了6%——如果想到我们提醒的频繁程度仅为每月一次，而且并不会打扰人们的日常生活，你就会感觉这6%已经是非常显著的成效了。（毕竟，短信和站在门口的助理相比，远没有那么显眼和活灵活现。）我们没有通过教育或强化人们意志力的方式来改变现状，而是仅仅通过提醒人们在管窥心态之下容易忽略的重要事项，就实现了储蓄额度的提升。

管窥心态让我们可以从一个全新角度去思考金融产品。有些财务决策会自然而然地出现在“管子”视野之中。某人会专门负责确保你偿还贷款、付清房租。该人或组织，就像助理一样，会将还钱的信息带入你的“管子”视野之中，无论你当时的管窥心态有多么严重。而储蓄，却没有设置一位专职助理去敦促你。如果我们没有提供这类介入性的行为，关于储蓄的想法大多数时候都会落在“管子”视野之外。

当然，我们也可以利用一些对管窥心态的理解。我们可以设定高额

的滞纳金，而且不提醒人们要在什么时候还钱。而许多这类手段的效果，无论是提醒还是滞纳金，都会对穷人产生更大的影响，也会让穷人承受更大的后果，因为他们正是那些拥有管窥心态的人。

我们能提醒的，不仅限于金钱。日理万机的人随时都有可能忘记去健身，而健身就是重要但并不紧急的事务。事实上，聘请一位私人教练，就能缓解这一问题。私人教练会提醒你，从而将健身这件事情拉回“管子”视野之中。这样，去健身就成了无法被你忽略的事情，因为教练会主动闯入你的“管子”中，询问你这周计划什么时候去锻炼。私人教练会永远在那里，以确保你一直都记得去健身房锻炼这件事情。

冲动而非提醒，也很容易将人带入“管子”视野之中。超市正是这方面的专家。它们发现了一种赚钱的好办法：将巧克力摆放在收银台边上。巧克力会立即勾起人们的“馋虫”，闯进人们的“管子”视野中。它们会令人突然想到：我想吃巧克力。其实许多冲动都是这样，无论某样东西多么重要或多么令人渴望，只要没在眼前，让人看不到，人们也就想不起来，因为这些冲动并不那么迫切。但当这些东西出现在眼前时，人们就会对之自我强化，并将其他冲动推出“管子”视野之外。（在巧克力的例子中，被推出去的就是你维持体重的冲动。）

在有了这些观察和理解后，我们自然会想：为什么不针对储蓄问题采取同样的策略呢？于是，我们在另一个项目中将这个想法落实到了行动之中，创建了一款叫作“冲动储蓄”（Impulse Savings）的产品。²就像巧克力一样，标有冲动储蓄的卡片被挂在了显眼的位置，比如收银台旁边。卡片上有一些图片，描绘了人们的储蓄目标，比如为孩子存钱上大学、置房购车等。卡片的设计初衷就是想与巧克力一样，去创造出一种冲动。只不过人们不是去“购买”这些卡片，而是去储蓄：人们为卡片而支付的现金，会转到他们的储蓄账户之中。

卡片能够有效地对抗管窥心态，因为它不仅能挖掘出人们潜藏于心中的目标，而且也可以在目标幻灭之前，提供一种为目标而行动的简易方法：购买卡片。我们与IFMR信托公司（为穷人提供金融服务的大型机构）合作进行了这个小型实验项目，惊喜地发现，许多人都十分愿意采纳这种储蓄方式。对于事务繁忙的人来说，家人的照片即使只是偶尔在办公桌上出现也能起到作用（照片不是一直摆在桌上成为背景，而是以不规律的频率出现，但足以吸引人的注意力），这样就可以使本来会被忽略的家庭得到重视。

提醒非常有效，但因其太过于明显和直接，所以重要性反而更容易为人们所低估。2008年，马萨诸塞州机动车登记处想出了一个节约成本的办法。³以往，登记处会寄出提醒车辆登记即将到期的信函，成本颇为高昂。于是，他们决定取消信函提醒。从某种角度来看，这个决定完全讲得通。但用我们的分析眼光来看，你就能理解这种做法的愚蠢之处。登记到期的时间可以是一年中的任何一天，这完全取决于上次你的登记时间。如果没有提醒，人们很难会记住这个日期。对于贫困和繁忙之人来说，信函提醒很可能是确保他们的车辆登记不过期，让他们免于罚款的唯一方式。事实上，与这条简单的政策同时出台的，还有州政府施加的递减税。

虽然提醒这种手段简单至极，但又常常被人们所忽略。政策制定者会在储蓄态度的培养上投入数百万美元，但却不懂得将提醒人们储蓄的方法融入其中。我们总是会花很多钱去办健身房的会员卡，但却不知道应该如何让自己将健身这件事情记在脑子里。

让“疏忽”等同“默许”

我们早已把储蓄这件事情忘到了九霄云外。事实上，我们两人已经有一段时间没有想到过储蓄这件事情了。到底是什么导致了这样的粗心大意？（我们其中一个人家里还有孩子！）其实，这也算不上什么不可救药的粗心大意。我们的储蓄账户，从退休存款到孩子们的大学教育存款，数额一直在稳步上升。那么，我们是如何在不主动储蓄的情况下存钱的呢？许多人都在采用这种方法：很早以前，我们两人就加入了一个储蓄计划，它每个月会自动从我们的工资中转存10%。虽然我们每日的行为似乎是对储蓄的完全忽略，但存款余额却在增长。我们在花钱的时候，根本没有想过储蓄这个问题。自动转存功能使得我们可以在完全忽略储蓄的情况下进行储蓄。

这个例子反映了一个简单的理念。当忽略问题存在时，改变行为的结果往往比直接与其对抗更有效。下面是一个关于退休存款的例子。

稀缺实验室

SCARCITY

美国人开始接受新工作时，需要填写一份参加401（k）计划

（美国的退休储蓄计划）的表格。一般情况下，如果没有正确填写表格，人们就无法参加计划，也就意味着后半生要承受灾难性的后果。而人们在刚刚接受一份新工作时，常常会经历诸多新变化，心理上也要经受紧张焦虑的考验，在这种情况下，人们会产生管窥心态，就有可能忘记填写表格。在一项设计新颖的研究中，研究人员改变了忽略表格这件事情的次序。⁴新员工会收到一张改版了的表格，上面写着：“你同意以3%的额度参加401（k）计划。如果你选择不参加该计划或决定以不同的额度参加计划，请归还表格。”现在，当人们忽略表格时，就等于是同意进行储蓄。而且更好的一点在于，对于所有想要参加退休储蓄计划的人来说，需要去做的事情全部都已经安排好了，不会因为疏忽而遗漏什么。毫无疑问，研究成果非常令人惊喜。就算到了3年之后，401（k）计划的参与率依然能反映出这项研究实施与否带来的巨大差异。在新员工需要主动选择退出计划的公司，超过80%的员工参加了401（k）计划；而在新员工需要主动选择参与计划的公司，只有45%的员工参加了该计划。可见，当我们对决策遭受忽略的默认后果稍加改变时，就能获得显著的成效。

当然，由“别人”来设定你的默认选择，肯定会遇到不少棘手的政策问题。但是在许多情况下，你可以自行设定自己的默认选择。自动还款就是一个很好的例子。公务繁忙的人，在工作的“管子”视野中，如果选择自动还款，就不会再承担忘记还账单的风险。他完全可以忘记支付账单这回事，而当他忘掉时，账单依然能按时还。这样看来，对于繁忙之人来说（至少对于那些能利用现代技术的人来说），最为顽固的管窥问题是那些无法自动被处理的任務，比如车辆登记、驾照更新、税款支付等。更不易解决的问题是那些既不能自动处理，也没有截止日期或提醒的事务，比如写遗嘱或体检。

这种思路有着广泛的适用空间，尤其适用于那些重复性高、可预测的事件。请想象某人在家工作，正因一项任务的截止日期迫近而产生了管窥心态。我们知道，此人一定会忽略一日三餐的质量，会吃掉手边能找到的任何食物。事实上，此人的心思和精力全用在了工作上，很可能会选择最不健康的食品，因为这些不健康食品最能满足冲动之下的欲望。橱柜里装满了各种美味，因此这位大忙人长了几斤体重。而相比之下，如果橱柜里只有健康食品，就能让这位大忙人的腰围免受工作的负面影响。

美国银行近期推出了一项名为“留下零钱”（Keep the Change）[5](#)的活动，进一步证实了我们可以建设性地利用人们的忽略行为。银行对这一活动给出了以下解释：

在参与了“留下零钱”活动后，你就能够自动增加储蓄额。我们会将你所有银行借记卡上的购买金额四舍五入到整数，然后将活期存款户头上的差额转存到储蓄账户上。每喝一杯咖啡，每加一次油，每去一次超市，都能增加你的储蓄额度。还有什么比这种做法更简单？

“留下零钱”活动（此活动的其他方面也受到了批评，包括低利率和高收费）的好处是：通过引导，而不是通过压抑人们的购买冲动来帮助人们储蓄。[6](#)人们的确会忽略储蓄这件事情，于是，这项活动就让人们通过最自然的“消费”方式来进行储蓄。

保持警觉

对于事务繁忙的职场人士来说，定期去健身房锻炼身体要比办张会员卡难得多。其中一个原因显而易见：办会员卡会给人们带来的痛苦，远远不及锻炼腹肌或在椭圆机上踏步半小时的痛苦。但同时还有另一个原因：办会员卡，只是一次性选择；而定期去健身房，则需要一直对身材和健康保持警醒，才能一次又一次地去做正确的事情。由此，我们可以认为，选择以两种形式存在，其中一种需要时刻保持警醒，而另一种则是一次性的。警醒型选择需要我们持续不断地进行重复选择，比如去健身房、为不时之需而储蓄、吃健康食品，或与家人共度不受干扰的家庭时光等。有些选择甚至还需要特别的警醒。错过一次去健身房锻炼身体的机会，不过就是少付出一些努力而已。但若是没有按照医嘱服药，问题就会严重许多。仅仅一次疏忽大意——用存款买了一件皮夹克，也会使几个月的勤俭节约付之东流。一次性选择只需要采取一次行为（或仅是偶尔为之），就能得到预期的成果：如果申请自动还款，你就不用再担心还账单的问题；如果买下一台洗衣机或烘干机，你就可以在未来几年里不再去洗衣房；如果参加通信服务提供商的折扣活动，你就能在活动取消前一直省钱。

特别是在你产生了管窥心态的情况下，一次性地去做正确的事情，要比重复去做正确的事情容易得多。而很多良好行为都需要警醒，比如做一位优秀的家长、省钱和吃健康食品等。与此相比，许多错误行为只需犯一次，便能导致诸多问题的发生，比如借贷、许下不明智的承诺和胡乱买东西等。只消一次挥霍或欠下一笔贷款，你就会给自己的未来挖

下陷阱。所以，只有持续的警醒才能使你有望从中逃脱。

这些现象给了我们一个提示：只要有可能，就要将警醒型行为转变成一次性行动。若不想每次经过厨房都要保持警醒，以免去拿零食，不如在去超市时保持警醒，不买零食。许多日常事务都适用于这一原则：保持房间的整洁，需要警醒意识，要么就一次性地请位清洁工定期打扫（假设你有能力承担这一费用）；每个月支付账单时需要警醒意识，而设置自动还款只需一次；开车上高速路时提前准备好零钱，需要警醒意识，而办理ETC速通卡仅需一次。从长远来看，由于管窥心态会导致忽略，若能为这些容易被人们所忽略的事务找到一次性解决方案，它们就会发挥非常大的作用。坚持抽出时间陪伴孩子，如果单纯凭借你自身的警醒意识，效果一定不理想。但如果你每周带着孩子报名参加一次活动，那么这种一次性行为就能确保你每周至少能有一小段时间是全心全意陪伴孩子的。

相反，把存在问题的一次性行为转化为需要时刻保持警醒的行为也十分必要。有些政策制定者提出在消费者购车时设置“冷却期”，²而类似的方案也同样适用于各类借用行为（金钱、时间和热量等）。其实，这一举措就等于为你设定了一套系统——要求你在采取实际行动之前，对自己的决策进行多次确认。（请想象，每当你收到一份富有吸引力的邀请时，电子邮箱都会自动回复道：“谢谢。我也许可以参加，一周后给你答复。”）

偶尔，你也可能会想要将一次性行为转换为需要警醒意识的行为。多年之前，你经过反复研究最终选定了汽车保险，而现在是否有了更优惠的选择？上一次你打听汽车保险的行情，是什么时候？选择总是在变化，某些一次性选择也可能存在误导：我们在购买影片租赁年卡时，总会假设自己每个月要看好几部电影，看完后也会立即归还碟片。而事实上，如果我们回想一下自己看过的影片数量以及为办年卡而花的钱，就会为观赏每一部电影的高昂成本而惊诧。除了自动更新之外，也许偶尔确认一下之前所做出的一次性选择的合理性，也是十分明智的。

那么，我们应如何看待贷款这个问题？我们是否应该取缔快速贷款这种可能带来恶劣后果的一次性选择？在第5章中讲到的《家庭问答》实验中，我们了解到，不给穷人借用的机会，就能提高他们的整体表现水平。现实生活要比实验室环境复杂得多。某些贷款的确不好，但还是有一些贷款产品能够为人们提供帮助。我们该如何确定哪个好、哪个不好？就算运用我们自己提出的理论进行分析，有些贷款也能为人们提供

他们所需要的余闲。当车子出现故障，你需要现金去修车时，如果能得到一笔贷款，就算费用很高，也可以防止一系列更严重问题的发生，比如上班迟到或者因迟到而失业等。

矛盾的是，稀缺虽然会使你选择权宜之计，但同样也会提高你因这些权宜之计而受到伤害的可能性。

在了解了稀缺心理后，我们就能从中领悟到：我们需要为管窥心态做准备，需要防止忽略情况的发生。我们要提前做好充分的调查，这样就不会在产生管窥心态时做出错误的选择。而且，我们要安排好各项事务，这样就不需要时刻保持警醒才去做那些值得鼓励的事情，还能不定期地对这些行为进行重新评估。

找准关联行动时机

在充满管窥心态和忽略现象的世界中，许多事情都取决于时机。在对未来进行决策时，我们会犯下许多最为严重的错误，因为未来不在我们的“管子”视野之中，是遥远而模糊的概念。有些事情我们当天完全不可能去做（今天太忙了），却想当然地认为会在一个月之后做（没问题！日程安排还空着呢）。我们当天的需要是紧迫的，而一个月之后的需要则是抽象的、不现实的。我们了解到，这正是我们接手过多任务的原因，也正是囊中羞涩的人购买他们无力承担的商品的原因。6个月前，那台洗衣机实在是吸引人，因为买下之后可以在半年后再付款；而现在，这台洗衣机却成了家里的主要负担。

我们一旦了解了稀缺心理，就能对其加以利用。对未来可能出现的稀缺没有充分的理解，会给我们带来麻烦，但我们也能利用它。

稀缺实验室

SCARCITY

一项叫作“为明天储蓄更多金钱”的计划，就了解到了人们愿意为资源相对充裕的未来许下承诺的愿望。⁸通过这项计划，那些认为自己目前没有能力进行储蓄的人们，会同意在未来某一天涨工资时，提高他们的储蓄扣除额度。也就是说，他们当下并不需要做出任何新的牺牲，而只有在今后那个模糊而遥远的未来，才会被从工资卡里扣钱。这项计划的效果令人惊叹。在一家公司中，超过75%

的员工选择了该项计划，只有少部分员工后来退出了计划，选择自己去储蓄。到了第三次涨工资时，员工的储蓄率早已翻了三番之多。

该计划的明智之处就在于，恰当地利用了人们所期待事物（涨工资）与其所希望事物（增加储蓄额度）之间的关系。这一计划自动地将两者联结为了一体。对于借用问题，我们也能以此为解决方案。请看下面这项思维实验。

稀缺实验室

SCARCITY

甲州政府为了抑制掠夺性借贷业务的发展，强制工薪日贷款商降低收费标准，比如对200美元的贷款，将收费从50美元降至25美元。（前提是贷款行业继续保持盈利和生存状态。）乙州政府创建了一个与此不同的项目。收费标准依然是50美元，但贷款商却只能收到25美元。余下的25美元以借款人的名义存入账户。只要账户总额累积到200美元（在这个例子中，就是8次贷款），此人就无须再借钱。当他需要贷款时，可以直接使用账户里的储蓄。事实上，从50美元的贷款费用中省下25美元，借款人很快就能成为“自己的贷款商”。

一言以蔽之，你打算在之后日子好过一点时再去做的那些明智决策，很有可能不会兑现，因为随着未来一天天临近，你会发现，现在也不比以前好过多少。因此，我们需要先发制人，做好过渡工作。就当你专注于健身这件事情的那个时刻，抓紧机会，买张健身卡，请一位私人教练，让朋友时刻监督，尽你所能地做到最充分的准备。这样，当你的“管子”视野中装进其他事务时，健身这件事情也依然不会消失。如果你去超市购物时能足够专注于健康，那么就要确保家中的橱柜里装满了你应该吃的健康食品，因为当你的心思不在饮食这件事情上时，就要让自己除了健康食品外别无他选。而当某样东西（一本书或一段广告）碰巧让你在那个时刻想到了年迈之后的生活，不妨抓紧时间采取行动：为工资账户设定自动转存、给律师打电话安排一次会议以及为写遗嘱做准备。否则，你总会想着过两天再去做，而到那时，你又会进入另一个“管子”视野中。

节约利用带宽

因为稀缺会产生带宽负担，所以在对稀缺进行管理时，一个主要的关注点就是如何节约利用带宽。繁忙之人总是会争分夺秒，而贫穷之人也总是会一分钱掰成两半花，每一位处于稀缺状态之中的人，都会因带宽如何分配和使用而受到极大影响。

利用带宽，就是要去分配我们有限的信息处理能力。从这个角度讲，那些需要进行更多信息处理的决策，会占用更多的带宽。每一位时间有限的管理者，都需要一位助理，这位助理应该擅长对决策进行综合处理，对选择进行简化，并能清晰地将其呈现出来。如果一位下属只会给出大量未经处理的数据，那么这些数据也不会有多大价值。清晰而简洁的综合，是节约利用认知资源的理想方法。

而我们在给出信息时，常常不能充分认识到这一点的重要性。经济学家玛丽安·伯特兰（Marianne Bertrand）和亚戴尔·莫斯（Adair Morse）进行的一项关于工薪日贷款的研究，就反映了这一现象。[9](#)

稀缺实验室

SCARCITY

研究人员将准备借入工薪日贷款的客户分为两组：一组看到表格里列出的是他们将要支付的年有效利率（443%）与可比较贷款品种（信用卡利率为16%）之间的对比；另一组看到的是类似的数据，但并非利率，而是他们将要为贷款支付的费用（美元）——如果客户可以在两周内还清贷款，就要支付45美元，一个月还清则要支付90美元，以此类推。而如果他们用信用卡借贷同样数额的钱，为此支付的费用（美元）为两周2.5美元，一个月5美元，以此类推。换句话说，类似的数据以不同方式展现在了两组实验对象面前：一组看到的是利率这种抽象的衡量标准，而利率的确切含义可能令人很难揣摩；另一组看到的是要从自己兜里掏出来的实实在在的钱，而人们对美元这种货币单位早已再熟悉不过了。

伯特兰和莫斯发现，看到美元成本的那一组客户，选择工薪日贷款的人少了许多。为工薪日贷款而来的人们，已经习惯了以美元为单位进行思考，他们眼里看到的是美元，需要的也是美元。相比之下，利率在日常生活中却很少被人用到，因为这需要大量的智力资源，才能将其转换为更直接、更真实的数据。当人们产生了带宽负担时，具体的数额就

会比抽象的概念带有更丰富的意义。

营养标识也存在类似的问题，它们向人们灌输了大量陌生的信息。消费者现在不仅能看到卡路里数据，而且还能看到来自不同种类脂肪的卡路里数据——好脂肪与坏脂肪的对比，还有重要营养物质的信息（是否能从产品中摄取欧米伽3脂肪酸），一份产品能提供的几种维生素和矿物质的每日所需摄入量百分比，等等。所有这些都需要人们进行认真的信息处理，在没有简单处理方式的情况下，这些数据实在让人如堕雾中。吃掉一个面包圈究竟会对身体产生什么样的不良影响？真的不好说。

有时，就连做出权衡这种简单的事都会给人以负担。请想象你自己手头上现在有许多尚未完成的工作。此时，一位好朋友准备去另一座城市定居，为此组织了一场欢送会。虽然你工作繁忙，但真的应该去参加。于是，你决定将朋友的聚会挤进繁忙的日程安排里，但不打算在聚会上花太长时间。你想到了那里再根据当时的气氛和感受决定什么时候离开。因此，你出席了聚会，但一个小时之后就开始琢磨：“我是不是能走了？”聚会充满了乐趣，如果你现在离开，很可能被朋友误解，但工作在召唤你。一个小时足够了吗？如果离开会不会显得很没礼貌？你开始发愁。你又等了一会儿，但你的心思早已不在聚会上了。你所做出的权衡——为了与朋友聚会而放弃的工作，使得你很难全身心地投入到娱乐活动之中。你以为保持灵活度是在帮自己，但却让自己产生了拖延和分心。

繁忙之人非常渴望能挤出时间来陪伴家人和朋友。而在一项接一项的待办事务中挤出时间的难度很大，最终还会导致可以预见的忽略现象以及因忽略而引发的后果。而且，就算你勉强将时间挤出来，也无法体会其中的乐趣，因为你的心思在别处，会寻思着如果将时间用来做别的，可以完成哪些事情。犹太人的安息日，是应对稀缺权衡的一种独具智慧的干预手段。安息日是一个古老的传统。人们在安息日不工作，不发邮件，不写字，不做饭，甚至不能开车。这一天，是宁静而安详的，是供人们来恢复活力的。我们许多人可能很多年都没有机会享受这种体验。安息日的妙处至少有两个方面：其一，没有选择，没有困境。这一天什么也不能做，只能休息，不存在权衡问题；其二，安息日固定在每周的同一时间，周五一结束，无论你有多忙，都要进入休息状态。没有提问，无须计划。犹太学者亚伯拉罕·约书亚·赫舍尔（Abraham Joshua Heschel）曾写了一本关于安息日的著作，他在书中提到：“安息日是上

帝赐予的时间礼物。”[10](#)

阿氏食谱（Atkins diet）总会使人联想到安息日。大多数节食计划都会鼓励人们去做权衡，它们会给出固定的卡路里摄入量、固定的碳水化合物克数以及其他一些限制。节食者要去选择他们喜爱的食品组合，同时满足总体的限制要求。这样，节食者就有了“灵活性”，可以将自己的偏好考虑在内。但就像前面讲到的那位与朋友聚会的大忙人一样，这种灵活性只能使背负着带宽负担的节食者花费大量的时间去权衡。权衡式思维不仅会使人分心，而且对节食计划也尤为不利，因为将思想专注在食物上，会让人更加难以抗拒食物的诱惑。一项研究将参与者随机分配到规则复杂性互不相同的节食计划中，并给出了如下总结：“体重管理计划需要大量认知能力，而可感知的规则复杂性则是放弃节食行为的概率最大的影响因素。”[11](#)

阿氏食谱的诸多版本都有助于解决这一问题。阿氏食谱不需要人们一直去做出权衡，而是强制性地给出了碳水化合物的限额。这就使得选择变得十分简单：某些食物的碳水化合物含量很低，你没得权衡，只能吃它们；而诸如大份甜品等其他选择，就是你根本不可能考虑的，因为它们碳水化合物的含量太高了。阿氏食谱给人很小的权衡空间，你可以尝尝小甜点，吃两口面包，但与标准的节食计划相比，选择就少了许多。有些人并不认可阿氏食谱节食法，但从心理学角度来讲，阿氏食谱有着明显的优势，因为你不用想着去分配卡路里摄入量，去精心安排每一餐的饮食。所以说，阿氏食谱更接近安息日——简单的禁止，很少的权衡。

在带宽充裕时行动

带宽的另一个重要特征是，它不会一直保持同一状态。请回忆我们在第2章中讲到的甘蔗农。他们在临近收获时较为贫困，而在收获后较为富裕。但更重要的是，临近收获时，他们的带宽更少，而反之则带宽更多。同样，由于低收入工作者很难做到消费均摊，所以那些按月领工资的人和领取粮食补贴的人，很可能在月底时带宽最少，而在月初时带宽最多。若能在政策推行和项目设计时将时机因素考虑在内，将是十分明智的做法。如果你想要为人们提供受教育的机会，而所教授的内容，无论是保健知识还是会计学知识，都需要一定的带宽，那么什么时候提供这些教育内容最为有效？如果你的目标对象是农民，你是准备在收获之前开课，还是在收获之后？穷人是在圣诞节之前还是在圣诞节之后东

拼西凑地为买礼物发愁？只要懂得了带宽变化的时间规律，就能在日历上将最恰当的日期标注出来：在某段时间，人们能真正地去听课并吸收知识；而在另一些时间，人们只会心不在焉。带宽变化时机的重要性，在于能让人们将计划与更为理想的带宽变化时机联系起来。下面这则研究就生动地反映了这一点。

稀缺实验室

SCARCITY

肥料能为农民带来很高的经济回报，比如，在肯尼亚的经济回报率就高达75%。[12](#)但尽管如此，肯尼亚的很多农民依然不为田地施肥。问题不在于他们缺乏知识，大多数农民都称，他们有购买肥料的计划，但实际上只有不到1/3的农民真正将购买行为落在了实处。他们总是以钱不够为理由。而他们真正想说的是，他们在需要肥料的时候，才发现自己的钱不够了。他们在收获之后立刻就能获得收入，而肥料是好几个月之后才要买的，而此时，农民们手头已经没有什么现钱了，带宽也承受了沉重的负担。

为了填补金钱与肥料之间的缺口，研究人员发明了一种简单而新颖的干预手段：他们让农民们在现金充裕的收获之时提前购买肥料，然后在种植阶段再将肥料送达。[13](#)仅仅凭借这一点改变，肯尼亚的农民们购买和使用肥料的比例就从29%一下子升到了45%，效果非常显著。可见，将重要决策的制定时机从农民缺钱、缺带宽的时间段转移到有钱、有带宽的时间段，就能避免问题的出现。[14](#)

对带宽的自然变化规律有所了解，也能帮助那些事务繁忙的人。大忙人总是以时间表为参考来安排活动，考虑到某项任务需要多少时间，比如正好在周三上午11点能有合适的时间。

除了时间以外，工作任务同样需要带宽，只不过有些需要的多，有些需要的少。参加一场电话会议，确保会议按计划推进，这项工作所需的带宽，要比与老板或客户面谈所需要的带宽少得多。而我们却总是将关注点放在时间上，没有对带宽予以充分的认识。不可否认，我们的带宽在一天中也会有所波动。所以，我们是否能够明智地进行任务规划，确保高带宽任务能安排在高带宽时间段上呢？

对带宽加以利用，不仅需要对各项任务的执行时间进行安排，而且

也要设置最佳的顺序。

在撰写本书时，我们每天早上都会专门留出一段时间用于写作，并坚持了很久。我们不惜一切代价去保护这段时间，有时甚至会为此忍受痛苦，比如即将召开一场6人会议，而自己是唯一一位负责安排时间的人。我们不仅仅是在保护这段时间，同时也是在保护高带宽的时间。但这种做法的效果并不理想，比如我们为写作专门预留了时间，但也并不是很有效。后来我们才意识到，之前的做法是不对的。我们拼尽全力去保住写作的时间，而在坐下来投入写作之前，我们会快速查一下邮件，处理一下紧要事务。到了9点钟，我们会强迫自己放下手里的一切事务，就算采取极端手段也在所不惜，比如关掉无线路由器。但事实上，我们并没有完全放下。一封关于项目进程滞后的邮件，强调了我们落后的严重程度；另一封邮件，又提醒了我们筹钱的紧迫性。我们虽然坐在办公桌旁准备写作，但内心并不平静。我们的大脑中已经展开了一个接一个喧嚣吵闹的想法。我们就像是两个节食者，每天早上先遍览色香味俱全的甜点，然后再坐下来思考其他事情。

为带宽不足的人提供便利

许多低收入高中毕业生都没有上大学，[15](#)所以，许多慷慨的助学项目都假设其原因在于缺钱，并希望从这个角度为低收入高中毕业生提供帮助。然而，这些项目的功效远远没有得到充分的发挥，结果根本没有几个申请人。这种现象让人有些想不通，于是一群研究人员决定想办法找出其中的原因。

稀缺实验室

SCARCITY

他们将因报税而前来求助的高中毕业生（及其家人）分为3组，将所有申请大学助学金的表格发放给他们。对于第一组，研究人员只是去观察他们的申请倾向。对于第二组，研究人员尝试着去填补其信息缺口。也许高中毕业生不了解他们有资格申请的助学金数额，于是税务工作人员将这些信息提供给了他们。对于第三组，研究人员做了点颇有意思的事情。税务工作人员不仅告诉高中毕业生们能申请到的数额，而且还替他们将表格填好。最后的结果显示，仅仅告诉人们他们有资格获得的具体利益，并不能收到明显的成效，但帮助其填好表格，却起到了巨大作用。学生们申请助学金

的比例提高了，而且上大学的比例也增加了29%。[16](#)

对于任何人来说，填写表格都是一件麻烦事，是个障碍，人们总会拖延和遗忘。而对于低收入人群来说，他们的带宽已经承受了负担，而且很可能因自身的经济条件而觉得脸上无光，所以，填表就成了一个更加麻烦的障碍。来自没有大学经历的高中毕业生，在得到填表帮助的情况下，大学入学率是之前的3倍。

关于如何对稀缺进行管理，我们从中还能得到更深的体会。看起来不足挂齿的小步骤，在错误规划、拖延和遗忘的影响下，就有可能演变成巨大的障碍。而我们在规划人生或为他人制定政策时，常常会忽略这些障碍的存在。给某人发放一张表格，让他带回家填写，此人就有可能遗忘，而如果让他当场填好，参与率就会直线上升。无疑，填表这件事情是“不足挂齿”的小步骤，但人们就是很容易在这一小步上摔跤，就像计算利率和为车子做年检一样。当我们的带宽承受负担时，最简单的障碍也会造成巨大的损失。

享受公共福利的人每年都要填写一大堆表格，去“重新认证”，以证实他们依然有资格享受福利待遇。可以想见，就是在进行重新认证的时间段，人们会从福利项目中脱离出去。而且，填写表格的要求总是将最需要福利救济的人排除出去。带宽负担最为沉重的人，也是最容易延迟重新认证的人，而这些人却正是最需要福利救济的人。

为了进一步了解带宽负担的内在逻辑，我们可以这样思考：请想象我们会为填写财务补助表格而收取一大笔费用。我们很快就会意识到，这笔钱收得实在有点傻——以穷人为目标的项目本就不应该向他们收取一大笔费用。而现实中，我们却常常会设计出一些项目，这些项目以带宽承受沉重负担的人为服务对象，但却需要他们拥有大量带宽才能申请。用另外一个比喻来说，这就像是走到一位需要帮助的杂耍演员身边，然后向空中抛起一个球，让他的杂耍负担进一步加重一样。

我们之所以讲到这些，并不是想要取缔所有的障碍。有时，障碍的存在自有它的道理：申请财务资助的表格之所以复杂，是因为提供财务资助的组织的确需要许多信息；之所以要去做重新认证，是因为人们的境遇总是在不断变化，而资助项目需要将目标瞄准在那些真正需要帮助的人身上。但同时，我们也能找到替代方案，比如许多表格可以利用税务数据来自动填写。稀缺管理过程中出现的错误，在于我们在低估带宽负担的同时，过多地将目光放在了移除障碍的高额成本上。而实际数据

告诉我们，障碍所引发的带宽负担很可能会异常沉重。扶助项目的成功与否，很可能就取决于那些看似不起眼的小障碍。而这些小障碍会使人们无法获得他们理应获得的资助，或者让高中毕业生们没有机会去大学深造。

从富足起步

在完善稀缺管理方法的同时，我们也应该记住稀缺通常是从资源充裕期开始的：截止日期临近之时的忙乱，常常源于几周之前没有对充裕的时间进行有效的利用；收获之前的几个月，农民们手头的钱总是不够用，这是因为上次收获之后，他们没有对现金进行很好的规划。

不知读者是否还记得第1章中讲到的研究，实验对象在更为紧迫的截止日期之下，能更好地完成文章校对工作。虽然很多人都能认识到截止日期对工作的积极作用，但它还是没有得到人们的充分重视。在该实验的另一个版本中，研究人员允许一部分实验对象自行选择截止日期。[17](#)这种办法发挥了效果，实验对象自行设定了截止日期，帮助他们比没有截止日期的一组实验对象赚到了更多的钱。但是，他们自行选择的截止日期还不够严格和紧迫。自行选择的一组，比强加截止日期、没有选择的一组赚到的钱少了25%。我们在学生们的身上也观察到了同样的现象。一次，我们让学生们自行选择期末论文的截止日期。有些学生很明智地选择了比学期结束时提早一些的时间。相反，很多学生没有这样做，而这就使得他们在论文的截止日期即将到来时，还在忙不迭地写论文。

在充满稀缺的世界中，漫长的截止日期就是酿造麻烦的问题所在。早期的资源充沛会变相鼓励人们浪费，而等到截止日期临近时，管窥心态和忽略行为又会出现。所以，将漫长的一次性截止日期改为渐进式的阶段性截止日期，有助于问题的改善。对于金钱来讲，道理也一样。一次性拿到一笔钱的农民，很容易会落入同样的循环——一开始钞票一大把，最后却身无分文。而处理方法也一样，我们可以将一次性付款改为渐进式多次付款。如果农民们不是一次性收款，而是更为频繁地收到报酬，问题就可以得到有效的改善。食物补贴也一样。接受食物补贴救济的人，同样无法将收入均摊到整个月份。因为他们需要许多带宽，才能做到计划、记忆、控制和权衡。那么，为什么不将救济改为按周发放呢？或者在需要的情况下，将两种发放方式结合起来：一开始发给他们一笔能够保证大额月度花费的钱，然后每周再发一小笔钱。打

破“充裕—稀缺”循环的一种办法，就是将资源均摊，从而创造出长期的节制状态，而非突如其来的资源充裕，然后紧跟着阶段性的稀缺。

以余闲应对突发事件

正如我们所了解的一样，“充裕—稀缺”循环之所以对我们非常不利，是因为稀缺会让我们落入陷阱之中。这不仅是因为我们无法在资源充裕的状态下将资源均摊，也是因为我们做不到为未来预留余闲。我们在第6章中讲到了关于印度街头小贩们的例子，并从中了解到余闲缺乏给他们带来的困扰。一旦发生突发事件，他们就会立刻落回债务陷阱之中，而如果能对之前的充沛资源进行合理的分配，就能避免这一后果。这就是不预留足够余闲和不为可能发生的突发事件准备足够缓冲所引发的危险。突发事件不仅仅会对我们造成伤害，而且还会将我们推入稀缺心理的泥潭中。我们会因此开始产生管窥心态，开始向未来借用，过不了多久，我们就会落后一大步，然后开始紧追慢赶。

尽管如此，我们依然不能构建起一种缓冲储备机制。虽然就此问题直接进行的研究并不多，但我们可以从其他研究中找到很好的线索。举例来说，实际数据显示，人们总是会低估许多低概率事件发生的可能性。¹⁸这就是为什么我们不会为洪水和地震购买足量保险的原因。当事情的进展一切顺利时，我们都能做到居安思危，却会低估其发生的可能性，做不到充分的准备。而在许多可能发生的突发事件中，任何一件事情都有可能让我们从此一蹶不振。从学术上讲，我们面对的是诸多彼此不相关的低概率事件。可能打乱你的计划的，不仅有洪水、地震，还有你可能会生病，你的家人可能会生病，或者你家里进了贼、车子被偷、战争爆发、失业、闪婚、计划外生育等状况。上述所有事件都有可能发生，不过可能性很小。但问题就在于，上述任何一件事情都是突发事件，而为了应对突发事件，我们应该构建起某种缓冲机制。

缓冲机制的构建需要在资源充裕时进行。如果你早已预期到了自己的时间稀缺，那么就应在日程安排上留出富余的空间，不一定要找到什么明确的理由，只不过是为了在不浪费成本的情况下，去周转你手头的诸多项目和责任。如果是关于金钱的，那么就应储存一份存款以备不时之需，就算你手头没那么阔绰，也要尽量储蓄。做到这些事情并不轻松，也不会让人感觉自然而然，因为就算你知道突发事件和稀缺状态迟早会到来，但在资源充裕之时，也不会有这种感觉。

稀缺给人造成的负担非常沉重。而只有了解了稀缺的内在逻辑，我们才能尽量缓解其负面影响。我们可以在一定程度上对周遭环境进行“稀缺防御”。就像在房间里安装烟感器或者为新生儿设立教育基金一样，这些一次性的努力很可能会让我们收获持久的回报。

结语 规划和管理好“带宽”，稀缺问题就会迎刃而解

人们总是忽略带宽的重要性。当决定做某件事情时，人们会想到他们能利用的时间以及这件事需要花费的时间，但很少会想到带宽。迫近截止日期的稀缺体验，起因于人们在充裕时所采取的时间管理方式，我们在“富足”时往往想不到为将来留有余闲。所以，只要规划和管理好“带宽”，稀缺问题就会迎刃而解。

随着知识之岛的开拓，无知的海岸线也将蔓延。¹

——约翰·惠勒（John A. Wheeler）

本书的目的在于邀请读者们一同来了解一门正在成型的学科。我们希望，读者在对稀缺科学有了初步认识后，能改变对许多事情的看法——从偶尔感觉工作过度，到孤独、贫穷等持续无法解决的问题，其实背后都潜藏着深刻的原因。

带宽负担人人都有

从一个新角度去审视熟悉的事物，能让我们于不经意间获得意想不到的新发现。我们两人经常在手机玩一个叫作《争抢》（*Scramble*）的游戏。玩这个游戏可以让我们在工作间隙稍事休息，打发时间，有时还能借机拖延。游戏很简单，我们玩得得心应手。但我们发现，在撰写本书的同时，我们在游戏中获得的分数也出现了大幅下滑。在截止日期的压力下，我们连游戏都玩不好。这个例子生动地反映了带宽负担的无孔不入。虽然我们正进行着相关研究，也搜集了诸多数据，但在自己身上发现的成绩下降幅度，依然令我们感到惊讶。我们能隐约感觉到自己有些“认知疲劳”，但30%~40%的成绩下降完全超出了我们的预料。

（这个游戏本身其实非常简单有趣。）我们想，自己的大脑可能没有全力运转，但却没有认识到带宽负担的沉重程度。

读者可能会试着回忆自身的类似经历。那么，你生活中的什么活动可能会产生大量的带宽负担？带宽负担会在什么方面造成值得你重视的影响？背负沉重的负担时，你是不是连驾车水平都有所下降？你知道，人在困倦时不应该驾车，但你是否曾经想过，一整天的辛劳工作后也不应该驾车？负担过重时，你讲的笑话是不是都不那么好笑了？你是不是也变得不那么友善了？你做出的决定是不是更不明智了？你是否曾经这样说过：“我现在不想做这个重要决定，因为我的带宽超负荷了。”

人们总是会忽略带宽的重要性。当你事务繁忙，需要决定接下来要做哪件事情时，就会想到你能利用的时间，这件事需要花多少时间，但很少会想到自己的带宽。你可能会说：“我只有半个小时，来把这件小事做了吧。”而你很少会说：“我的带宽有限，来把这件轻松就能完成的事情做了吧。”当然，有时你意识不到，但也做到了这一点，比如你在一项任务上怎么努力都停滞不前时，会换一件事情来做。但这不过意味着，你是为已经稀缺的带宽又增加了一份负担。

我们会对时间进行规划和管理，却不会对带宽进行规划和管理。我们很少会关注自身起伏不定的认知能力。与饮食、睡眠和锻炼身体等生理能力相比，就更体现了我们对认知能力的忽略。与许多生活在现代社会中的人们一样，我们以大脑谋生，却对大脑每日的节律知之甚少。如果我们手头的任务是将一堆箱子从一处运往另一处，那么我们会比较清楚如何将工作效率最大化、什么时候用多大力气以及什么时候休息。但如果手头的任务是传达思想而非实实在在的箱子，那么我们就不知道如何才能将有限的认知能力最大化了。

从个人层面来看，我们对自身不断变化的带宽情况不甚了解；而从社会角度来看，我们也对不断变化的社会带宽十分陌生。科学家们总是会量化那些科学理论让他们去量化的事物。由此，社会科学家们就能量化稀缺的实际规模，比如多少人处于失业状态、某一季度生产了多少产品、收入是多少，等等。

我们对经济认知方面一无所知。就像我们自身的个人带宽会不断波动一样，社会带宽也存在着同样的波动规律。我们是否会发现，2008年的经济萧条同样造成了影响深远的认知萧条？也许在那段时间里，人们的带宽也出现了大幅下降。倘若失业率不断上升，决策质量出现下滑，又会如何？我们没有能回答这些问题的数据。虽然重拾2008年的历史为时已晚，但我们依然有机会去搜集未来经济繁荣与萧条时的相关数据。近些年来，人们开始关注社会幸福水平，力图建立起一个与国民生产总值并行的国民幸福总值。那么，为什么不同时对国民带宽总值进行量化呢？

从这个起点开始，我们不仅能了解美国的整体面貌，而且也能知晓不同子群体的情况。当失业率从5%上涨到10%时，就意味着每20人中又有一个处于劳动年龄的人在经济上出现了困难。若从带宽的角度去看就会发现，失业率增长所造成的影响覆盖面变得更广了。在这种时期，更多的人会因此而整日为钱的事情发愁。也许就连那些收入稍有缩减的人，都会因为少了余闲而体会到了稀缺的困扰。而且，也许那些与新近失业之人有关系的朋友、亲属、邻居，也会对这种影响有所表现。认知影响很可能会比经济影响产生更大的覆盖面。

我们的话题不仅仅有关于经济萧条。生产力作为经济增长的驱动力在很大程度上取决于带宽：工人必须有效地工作；管理者必须做出明智的投资决策；学生们必须通过学习构建起人力资本。所有这些都需要带宽，今天带宽的下降很可能会造成未来生产力的缩减。

我们的话题也不仅仅有关于经济。带宽是一种核心资源，我们在为人父母、学习、强迫自己去健身、思考自身的人际关系时，都会用到。带宽会影响我们的思维方式，影响我们做出的选择。当经济陷入萧条时，我们能购买的商品会变得更少；而当我们陷入认知萧条时，生活中的方方面面都有可能受到影响，从子女教育到健身、储蓄、婚姻等。

当然，带宽测评不应该仅限于国家范畴。公司同样也可以进行带宽检测：员工们的情况如何？每个员工都可以自行测试。在做出重要决定之前，最好还是确认一下自己所有的带宽可以全部被投入其中。我们已经了解了几个相关测试，研究人员们可以就此进行更多的测试与实验。某些研究以稀缺为关注点：怎样才能最好地对余闲进行测评？怎样才能最有效地确定人们是否产生了权衡式思维？我们也可以在此基础上更深一步，从更为普遍的角度去测评不断浮动的认知能力。

我们还可以利用这些测评数据，对社会项目和公共政策做出更为公允的评价。在一项针对失业者的项目中，我们将重点放在了重新找工作上。无疑，这一点十分重要。但为什么不同时估测一下重新雇用对带宽的影响呢？如果失业者拥有了更多带宽，就能体会到更多好处。数据显示，失业家长的孩子在学校的成绩要比其他孩子差许多。如果说带宽是罪魁祸首，我们可以采取某些措施来予以缓和，那么这些项目所能带来的利益就会远超其初衷。

带宽和稀缺心理，解决稀缺问题的关键

将关注点集中在带宽上，不仅仅能带来更为优质的测评结果。不知读者们是否还记得我们在第2章中提到的快餐店经理，这位经理因为总要去管理他手下不称职的员工而抱怨不休。他可以怎么做呢？他是应该花时间和精力去动员手下，还是采取以解雇相威胁的手段，抑或应该实施更多的激励机制，进行更多的培训，或者多与员工谈谈话？这位经理的境遇没有什么特别之处。许多低收入员工的雇主都面临着生产力低下、缺勤等问题，他们也尝试过上述各种干预手段。

如果将关注点放在稀缺心理上就会发现，这位经理应该去解决另一个不同的问题。他不用激励，不用培训，更不必威逼利诱，而是应去关注如何增加带宽。低收入员工的经济情况总是不稳定的。我们也了解到了这种状况所造成的影响。在这种条件下，激励机制的作用也会大打折扣。

扣。当你产生了管窥心态时，许多奖励就会落在“管子”视野之外。那么，为什么不去帮助员工们应对经济情况不稳定问题，帮他们释放一部分带宽，想一想什么样的金融产品适合他们，应该为他们提供什么样的后期干预机制和工作条件呢？

我们从第5章中了解到，为了解决金钱稀缺问题，许多员工都选择了工薪日贷款。而值得注意的是，工薪日贷款不过是以已经完成的工作为抵押而借出来的钱。在某个月的中旬去借工薪日贷款的人，实际上都已经完成了一半工资所支付的工作量。之所以需要贷款，很大一部分原因就在于薪水总是在工作完成之后才发放。雇主为什么会让员工去借这种贷款，让他们因此面对落入稀缺陷阱、背负贷款负担的威胁，并由此降低生产力，而同时雇主又完全有能力以较低的成本向员工提前支付工资呢？如果雇主可以通过提供合理的金融产品和创建带宽来提高生产力，那将是一件多么有价值的事情啊！[2](#)

雇主的例子告诉我们，若从带宽角度进行思考，我们就可以用不同方式来回答并解决不同问题。从遵从医嘱这个简单的例子来看，穷人更容易忘记服药。我们可以说，“现实情况就是这样”，然后放任不管。或者，我们也可以创建一款像GlowCaps智能药瓶这样的产品。[3](#)这种药瓶，如果一天没有按正确的次数打开，就会作出某种反应：药瓶一开始会发光，如果还是没有被打开，就会开始发出哔哔声，最后还会给用户的手机发短信。药瓶“得寸进尺”地去烦扰它的主人，并以这种方法防止伴随管窥心态而产生的忽略行为。数据显示，自从有了GlowCaps智能药瓶后，穷人遵守医嘱的程度得到了大幅度的提高。通过对稀缺心理的了解，类似的产品和干预手段同样能解决类似的问题。GlowCaps智能药瓶的例子反映出的是：我们可以用廉价而有效的科技应对带宽所导致的问题，而类似的方法也能在其他领域中创造出同样富有成效的成果。

当我们想到增加世界各地的农作物产量问题时，也许不应该将关注点放在开发新型农作物或农民的培训上。也许，我们应该思考如何才能让农民去做那些小事，比如除草这件所有农民都了如指掌，却常常落在其“管子”视野之外的事情。农民们手中的“GlowCaps”是什么样的？又该如何提醒他们去除草和杀虫呢？

解决稀缺问题要从富足着手

在思考关于稀缺的问题时，我们遇到了几个新难题。举例来说，本

书的写作就没有按时完成。原因为何？除了所有显而易见的因素以外，我们在回顾过去几年的工作时，找到了两点主要原因：第一，一部分工作是在我们面对紧张的截止日期时完成的；第二，当我们面对紧张的截止日期写作时，就会体验到稀缺。正如理论所示，许多日子，我们都从稀缺中获得了利益，可以更加专注、有效地工作。

但大多数时间，我们都没有将紧张的截止日期时刻记在心中。很长一段时间里，我们在工作时总是觉得自己的时间很多。可以想见，在这段日子里，时间就一点点地流逝了。这虽然算不上是浪费时间，但我们每天的生产力，如果以写作字数来计算的话，远远不及我们应该达到的水平。你可以说，这是因为我们没有遭遇稀缺问题。但这句话就能概括全部吗？如果换个角度来看，是不是与充裕心理有关呢？

我们一直都将“充裕”视为稀缺不存在时的状态，将其视为“标准”，也就是一切顺利时的情况。但回顾以往，我们发现，有些时候我们能感觉到真实的充裕，而有些时候感觉到的是与稀缺或其他时间紧张的感受都不相同的独特体验。有时，我们会产生充裕心理。而充裕心理之所以耐人寻味，就在于其中似乎埋藏着终将使我们落入稀缺陷阱的种子。

我们许多人都曾在截止日期之前感觉时间不够用，因为我们浪费了曾经充裕的阶段。我们的学生总是在交作业截止日期前的两天（许多情况是一个晚上）才动笔写，而这种情况总是在好几周的无所事事之后才发生。在学期之初，学生们也并不是抱着这种想法来上课的。他们在最后时刻的东拼西凑，与临时抱佛脚的经理人以及不知道时间花在哪里的度假者所存在的时间管理问题如出一辙。

之所以会出现临近截止日期的稀缺体验，与人们在充裕时所采取的时间管理方式有关。稀缺与充裕之间的紧密联系，在许多情况下都会出现。农民之所以会在收获之前囊中羞涩，是因为他们在收获之后、现金充裕时花了太多钱。充裕时期的行为会酿成最终的稀缺：我们在现金一大把时总是不知道要去存钱；而我们在截止日期距离很远时，也总是会游手好闲。

请回顾2008年的金融危机。许多人都认为，导致这场危机的原因之一是认知盲点。从20世纪90年代到21世纪早期，房价一直在上涨。在房地产繁荣时期，人们很难想象房价可能会下跌，而且会认为根本不值得为此担忧。这种想法影响了人们的许多选择。如果房价注定不断上升（至少不下降），那么高杠杆交易就是有道理的，高住房贷款担保比率

就是更安全的。但是，实际上房价必定会下跌，有时下跌幅度非常惨烈。因此，所有以房价不会下跌为假设而做出的投资决定，导致了一场雪崩般的金融崩溃，险些就拖垮了全球金融体系。在这个例子中，金融危机所导致的巨大稀缺，其根源正是之前几年充裕阶段时的松懈行为。

当然，我们也可以将上述种种都看作普遍的行为方式：人们总会浪费时间；人们总会过于自信。但金融危机之前的好时光和充裕资源，将人们的这些倾向进一步放大，促成了过分自信的心理，强化了他们的自满心理。

沿着稀缺的线索向上追溯，我们就能看到充裕：萧条，是由我们在繁荣阶段的行为所导致的；最后一刻的拼凑，是由我们之前几周的不作为所造成的。虽然说稀缺在许多重要问题中都扮演了主角，但充裕却是其根基。

充裕，是否也存在着内在的通用逻辑，可以用来解释所有这些不同的问题，就像稀缺一样？

我们需要找到这个问题的答案。而现在，本书的写作已告尾声，我们终于又有了许多时间，可以不必为此而烦恼。

注释

引言 资源稀缺不可怕，就怕有稀缺心态

1. Marie Dressler—Biography. IMDb.

2. T. Smollett and J. Morley, eds., *The Works of Voltaire: The Maid of Orleans (La Pucelle d'Orléans)*, vol. 41 (New York: E. R. DuMont, 1901), 90.

3. 这一对稀缺的定义实质上是主观的。如果一个人拥有很多财富的同时拥有很多欲望的话，这个人也会体验到稀缺感，和另一个没那么多财富、也没那么多欲望的人相比，二人的稀缺感会一样。这一对稀缺的主观定义对于理解心理很重要。当然，结果依赖于心理和物理的现实情况。采用主观的方式只是为了理解心理。分析问题时，比如分析第7章的贫困问题，我们就会采用主观和客观两种方式。

4. 在罗伯特·普特南开创性的著作中，通过广泛的数据分析，他论证了美国人参与公共组织的趋势。可参考Robert D. Putnam, *Bowling Alone: The Collapse and Revival of American Community* (New York: Simon & Schuster, 2000)。此次之后，随着关于社交互动的数据大量涌现，这一领域已经蔚为大观了，可参考Jim Giles, “Computational Social Science: Making the Links,” *Nature* 488 (August 23, 2012): 448–50。当然，从经济发展问题到城市价值的问题，社交资本（社交稀缺的反面）的重要性已经被广泛讨论了。

5. Todd Tucker, *The Great Starvation Experiment: Ancel Keys and the Men Who Starved for Science*. Minneapolis: University of Minnesota Press, 2008.

6. A. Keys, J. Brožek, A. Henschel, O. Mickelson, and H. L. Taylor, *The Biology of Human Starvation*, 2 vols. Oxford: University of Minnesota Press, 1950.

7. S. A. Russell, *Hunger: An Unnatural History*. New York: Basic Books, 2006.

8. R. Radel and C. Clement-Guillotin, “Evidence of Motivational Influences in Early Visual Perception: Hunger Modulates Conscious Access,” *Psychological Science* 23, no. 3 (2012): 232–34. doi:10.1177/0956797611427920.

9. B. Libet, C. A. Gleason, E.W.Wright, and D.K.Pearl, “Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential): The Unconscious Initiation of a Freely Voluntary Act,” *Brain* 106, no.3 (1983): 623–42.

10. H. Aarts, A. Dijksterhuis, and P. de Vries, “On the Psychology of Drinking: Being Thirsty and Perceptually Ready,” *British Journal of Psychology* 92, no.4(2001): 631–42. doi:10.1348/000712601162383.

11. P. Saugstad and P. Schioldborg, “Value Size and Perception,” *Scandinavian Journal of Psychology* 7, no.1 (1966): 102–14. doi:10.1111/j.1467-9450.1966.tb01344.x.

12. 在视觉的知觉上，注意力更集中并不意味着看得更精确。某些研究表明，动机和注意力都可以渗透并指导初始的视觉处理过程。心理学的、神经心理学的和对行为的研究的证据都表明，通过增强一个刺激的显著性可以改变刺激的强度，从而强化知觉表征，改善或减弱视觉表现的各个方面。比如说，有观察表明，知觉到的事物与现实事物可以完全相反。Marisa Carrasco, Sam Ling, and Sarah Read, “Attention Alters Appearance,” *Nature Neuroscience* 7 (2004), 308–13; Yaffa Yeshurun and Marisa Carrasco, “Attention Improves or Impairs Visual Performance by Enhancing Spatial Resolution,” *Nature* 396 (Nov. 5, 1998), 72–75; Rémi Radel and Corentin Clement-Guillotin, “Evidence of Motivational Influences in Early Visual Perception Hunger Modulates Conscious Access,” *Psychological Science* 23, no.3 (2012), 232–34.

13. 在这一研究中，穷人家的小孩对硬币的估值更大。当然，穷人和富人的小孩在其他很多方面有不同。还有研究实验内生价值的问题，而不仅仅只是研究价值在不同人群中的差异。可参考Brian A. Anderson, Patryk A. Laurent, and Steven Yantis, “Value-driven Attentional Capture,”

Proceedings of the National Academy of Sciences 108, no.25 (2011):10367–71。

14. P. U. Tse, J. Intriligator, J. Rivest, and P. Cavanagh, “Attention and the Subjective Expansion of Time,” *Attention, Perception, and Psychophysics* 66, no.7, (2004):1171–89.

15. W. L. Gardner, Valerie Pickett, and Megan Knowles, “On the Outside Looking In: Loneliness and Social Monitoring,” *Personality and Social Psychology Bulletin* 31, no.11 (2005): 1549–60.
doi:10.1177/0146167205277208.

16. 这不是说孤独者拥有更强大的社交技能。恰好相反！我们必须严格定义我们所说的“社交技能”。这一研究度量了解读社交暗示的能力。另一方面，大量研究表明，孤独者在社交环境下的行为自治能力更弱。我们将在第6章论证，在社交环境中行为自治能力的减弱，是稀缺的一个可预测的结果。更多详细解读可参考John T. Cacioppo and William Patrick, *Loneliness: Human Nature and the Need for Social Connection* (New York: W. W. Norton, 2008)。

17. W. L. Gardner, C. L. Pickett, and M. B. Brewer, “Social Exclusion and Selective Memory: How the Need to Belong Influences Memory for Social Events,” *Personality and Social Psychology Bulletin* 26, no.4 (2000): 486–96. doi:10.1177/0146167200266007.

18. W. L. Gardner, Valerie Pickett, and Megan Knowles, “On the Outside Looking In: Loneliness and Social Monitoring,” *Personality and Social Psychology Bulletin*, 31, no. 11 (2005): 1549–60.

19. K. Vitasek, M. Ledyard, and K. Manrodt, *Vested Outsourcing: Five Rules That Will Transform Outsourcing* (New York: Palgrave Macmillan, 2010).

20. A. F. Bennett, “Structural and Functional Determinates of Metabolic Rate,” *American Zoologist* 28, no. 2 (1988): 699–708.

21. “稀缺”这个词也被用来描述心理学里的另一个效应。通常所说的“稀缺原则”抓住了这样一个观念：越少越想要。市场营销人员就在广

泛应用这一观念，比如“限量供应”，在线商城上显示“只剩3件”。Robert B. Cialdini, *Influence: Science and Practice*, vol. 4 (Boston, Mass.: Allyn and Bacon, 2001).

22. 在经济学里，这就是效用递增原理。对一种资源拥有更多，效用会增加，人会感到更幸福。在大量经济分析中，这些偏好或者说效用函数都是给定的。

23. Peter J. Rogers, “A Healthy Body, a Healthy Mind: Long-Term Impact of Diet on Mood and Cognitive Function,” *Proceedings—Nutrition Society of London* 60, no.1 (CABI Publishing, 1999, 2001). Doris Stangl and Sandrine Thruet, “Impact of Diet on Adult Hippocampal Neurogenesis,” *Genes and Nutrition* 4, no.4 (2009): 271–82. David J. Harding, Michèle Lamont, and Mario Luis Small, eds., *The Annals of the American Academy of Political and Social Science* 629 (May 2010).

24. E. R. Kandel, *In Search of Memory: The Emergence of a New Science of Mind* (New York: W. W. Norton, 2007).

第1章 专注的“得”与管窥的“失”

1. MOOD—Calvin and Hobbes—Full Story. Retrieved from <http://web.mit.edu/manoli/mood/www/calvin-full.html>.

2. *Dirtcandy*. Retrieved from <http://www.dirtcandynyc.com/?p=731>.

3. *Dirtcandy*. Retrieved from <http://www.dirtcandynyc.com/?p=2508>. 有人可能觉得科恩把脆皮豆腐这道菜放在比赛菜单上，只是为了把她在《铁人料理》上的名气变现，人们去她的餐馆品尝节目上的这道名菜，但其实这期节目是在录制一年以后才播出的。所以这可不是营销伎俩。

4. 创造力与时间压力之间的关系远比这个故事复杂。在很多案例中，时间压力会阻碍创造力。还有一个直觉想法对我们有用。当所要完成的任务需要发散性思维、产生新想法的时候，时间压力就会变成阻力；当所要完成的任务需要收敛思维、把大量想法凝聚成一个想法时，就像科恩的例子那样，时间压力就会成为助力。可参考Teresa M. Amabile, Constance N. Hadley, and Steven J. Kramer, “Creativity Under the Gun,” *Harvard Business Review* (August 1, 2002)。

5. Connie J. Gersick, “Time and Transition in Work Teams: Toward a New Model of Group Development,” *Academy of Management Journal* 31, no.1 (1988): 9–41.在这一原创性研究中，格西克参加了8个群体的每一次集体会议。我们只是简要介绍了她参加的一次会议，其实她为了这项研究参加过好几次。另外可参考Ruth Wageman, Colin M. Fisher, and J. Richard Hackman, in “Leading Teams When the Time Is Right” (*Organizational Dynamics* 38, no.3 [2009] 192–203)，这一文献探讨领导可以怎样去利用这些研究中的洞见。他们发现，在会议中途，整个群体都特别适合做出改变，这一点可以被领导所用。

6. D. Ariely and K. Wertenbroch, “Procrastination, Deadlines, and Performance: Self-Control by Precommitment,” *Psychological Science* 13, no. 3 (2002): 219–24.更早有研究发现，比起距离截止日期还剩3周，大学生更有可能在距离截止日期只剩1周时交出最优的带薪工作表。可参考A. Tversky, and E. Shafir, “Choice under Conflict: The Dynamics of Deferred Decision,” *Psychological Science* 3, no.6 (1992): 358–61。经济学家们一直在对截止日期的力量进行理论探索，利用不同框架进行研究，比如双曲线折现：我们总倾向于不成比例地衡量现在在未来的价值。可参考Shane Frederick, George Loewenstein, and Ted O'Donoghue, “Time Discounting: A Critical Review,” *Journal of Economic Literature* (2002)。按这种观点来看，设立中间期限的截止日期可以把未来的回报转化成更近期的回报，从而让人做事变得更有效率。

7. J. L. Kurtz, “Looking to the Future to Appreciate the Present: The Benefits of Perceived Temporal Scarcity,” *Psychological Science* 19, no.12 (2008):1238–41. doi:10.1111/j.1467-9280.2008.02231.x.

8. J. J. Inman and L. McAlister, “Do Coupon Expiration Dates Affect Consumer Behavior?” *Journal of Marketing Research* (1994): 423–28; A. Krishna and Z. J. Zhang, “Short or Long-Duration Coupons: The Effect of the Expiration Date on the Profitability of Coupon Promotions,” *Management Science* 45 no.8 (1999): 1041–56.

9. Paul Oyer, “Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality,” *The Quarterly Journal of Economics*, 113, no.1 (1998): 149–85.这份研究的诠释方法没有像我们的研究这样引入更多的心理学因素，而是聚焦于不同时期的替代效应。

[10.](#) S. Kaur, M. Kremer, and S. Mullainathan, “Self-Control and the Development of Work Arrangements,” *American Economic Review Papers and Proceedings* (2010).

[11.](#) M. Hastings, *Finest Years: Churchill as Warlord, 1940–45* (London: Harper Press, 2009).

[12.](#) Shah, Mullainathan, and Shafir, “Some Consequences of Having Too Little,” *Science* 338, no.6107 (November 2012): 682–85.

[13.](#) 并不是蓝莓富翁们不耐烦或者不想花太多时间在这一游戏上。如果真是这样，蓝莓富翁们完全可以玩更少的轮数，早早地结束游戏。

[14.](#) Sarah-Jayne Blakemore, Daniel Wolpert, and Chris Frith, “Why Can't You Tickle Yourself?” *Neuroreport* 11, no. 11 (2000): R11–R16.流行的观点认为，自我创造的举动可以被预测，它的效果也会被减弱。但关于空想的截止日期或时间压力，我们并不知道有什么更仔细的实证研究。不过，自我重新协商的问题倒是经常被研究。空想出来的截止日期并不会让人感到压力，因为在人们的内心里，总可以与自己重新协商。

[15.](#) *State Fire Marshal's Office Firefighter Fatality Investigation*, no.05-307-04, Texas Department of Insurance, Austin, Texas.我们感谢杰西卡·格罗斯（Jessica Gross）为这一案例提供的有用研究，也感谢伯顿·克拉克（Burton Clark）博士的通信。

[16.](#) P. R. LeBlanc and R. F. Fahy, *Full Report: Firefighter Fatalities in the United States—2004* (Quincy, MA: National Fire Protection Association, 2005).

[17.](#) *Firefighter fatality retrospective study*, April 2002. (Prepared for the Federal Emergency Management Agency, United States Fire Service, National Fire Data Center, by TriData Corporation, Arlington, Virginia).

[18.](#) C. Lumry (January 21, 2010). *Amarillo Firefighter Fatality—COFT* | *Council On Firefighter Training*. Retrieved from <http://www.coft-oklahoma.org/news-updates/m.blog/21/Amarillo-firefighter-fatality>.

[19.](#) C. Dickinson, *Chief's Corner* (February 27, 2007), retrieved from

<http://www.saratogacofire.com/seatbelt.htm>.

[20](#). L. J. Williams, *Human Factors* 27, no.2 (1985): 221–27. 科学家们研究的管窥，原来是指很具体地通过管子来看东西，他们研究了很长时间，有时候就是指真实的眼睛视觉层面。科学家让被试的视网膜中间的中心凹专注在目标事物上，而目标事物放在中心凹周围的区域，这个时候视觉灵敏度就降低了。他们研究人们在中心凹周围区域探测事物的能力。他们的研究成果很显著，比如，他们让所有被试看同一个字母A，有些被试要回答所看到的是不是A，这个任务很简单，而有些被试要回答所看到的字母是不是元音字母，这个任务就难一点。他们研究发现，尽管所有人的视觉经验一样，但那些要回答难一点问题的人，在看的时候视觉就没有那么灵敏，这些人要更加专注于完成这一任务，从而形成了管窥，没法看到中心凹周围的区域。这些都是在物理意义上研究的管窥，管窥还可以指认知意义上的视觉经验，就是专注于某个事物，而看不到其他事物。

[21](#). Susan Sontag, *Regarding the Pain of Others* (New York: Farrar, Straus and Giroux, 2002), 46.

[22](#). N. J. Slamecka, “The Question of Associative Growth in the Learning of Categorized Material,” *Journal of Verbal Learning and Verbal Behavior* 11, no.3 (1972): 324–32. 另一项研究要求被试列出美国的各个州名，研究发现，如果“帮助”被试，先告诉他们几个州名，反而减少了他们所能想到的州名数量，可参考Raymond Nickerson, “Retrieval Inhibition from Part-Set Cuing: A Persisting Enigma in Memory Research,” *Memory and Cognition* 12, no.6 (November 1984): 531–52。

[23](#). C. M. MacLeod, “The Concept of Inhibition in Cognition,” in *Inhibition in Cognition*, ed. David S. Gorfein and Colin M. Macleod (Washington, D.C.: American Psychological Association, 2007), 3—23.

[24](#). J. Y. Shah, R. Friedman, and A. W. Kruglanski, “Forgetting All Else: On the Antecedents and Consequences of Goal Shielding,” *Journal of Personality and Social Psychology* 83, no.6 (2002): 1261.

[25](#). 我们在书中描述的这个实验有点简化，实际的实验有两点不同：第一，被试看到的物体更多；第二，物体的颜色都不一样，而且要求实验对象回忆物体的颜色。

26. 这些实验结果来自我们写书时还没公开发表的研究报告。

27. Woody Allen—Biography, IMDb,
<http://www.imdb.com/name/nm0000095/bio>.

28. B. Arends, “How to Save \$10 000 by Next Thanksgiving,” *Wall Street Journal*, November 20, 2011.

29. Michael J. McCord, Barbara Magnoni, and Emily Zimmerman, “A Microinsurance Puzzle: How Do Demand Factors Link to Client Value?” *MILK Brief*, no.7.

30. X. Giné, R. Townsend, and J. Vickery, “Patterns of Rainfall Insurance Participation in Rural India,” *The World Bank Economic Review* 22, no.3 (2008): 539–566.

31. A. Aizer, “Low Take-Up in Medicaid: Does Outreach Matter and for Whom?” *The American Economic Review* 93, no.2 (2003): 238–41.

32. D. L. Strayer, F. A. Drews, and D. J. Crouch, “A Comparison of the Cell Phone Driver and the Drunk Driver,” *Human Factors: The Journal of the Human Factors and Ergonomics Society* 48, no.2 (2006): 381–91; D. Redelmeier and R. Tibshirani, “Association Between Cellular-Telephone Calls and Motor Vehicle Collisions,” *New England Journal of Medicine* 336, no.7 (1997), 453–58.注意，有一份大规模的调查报告的结论很惊人，开车时打电话对撞车的概率影响很低。可参见Saurabh Bhargava and Vikram Pathania, “Driving Under the (Cellular) Influence” (2008), available at SSRN 1129978。这一研究避开了醉驾问题，虽然引起了很多人的兴趣，但是与其他大量数据相抵触，还需要后续的调查。

33. 我们不知道有关边开车边吃东西的实验。我们所拥有的最好的数据来自“100-car study”，这个项目利用监控设备跟踪了12~13个月。他们发现，开车时吃东西导致撞车的概率增加57%，开车时打电话发生事故的概率增加29%，但开车时拨电话号码，风险增加279%，这也论证了这项研究的关键结论，视觉注意力分散带来的结果是致命的。Sheila G. Klauer et al., “The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data,” no.HS-810 594 (2006).

[34.](#) Paul Taylor and C. Funk, “Americans and Their Cars: Is the Romance on the Skids?”(2006).

[35.](#) B. Boon, W. Stroebe, H. Schut, and R. Ijntema, “Ironic Processes in the Eating Behaviour of Restrained Eaters,” *British Journal of Health Psychology* 7, no.1 (2002): 1–10.

[36.](#) “Recession-Proof Your Business,” *About.com Small Business*: Canada, retrieved October 22, (2012).

[37.](#) 对悔恨的研究源远流长，通常悔恨被视作自我控制问题的结果。T. C. Schelling, “Self-Command in Practice, in Policy and in a Theory of Rational Choice,” *American Economic Review* 74 (1984): 1–11.

第2章 带宽负担会降低人的智商

[1.](#) 对带宽，或者说计算能力，其实一直在各种各样的名目下进行研究，包括智能测量、推理能力、短期记忆能力、工作记忆能力、流体智商、认知控制、执行控制、注意力控制和冲突调控等等。还有些专业的研究者抓住了相关的不同点，但这些远远超出了本书的范围。R. W. Engle, “Working Memory Capacity as Executive Attention,” *Current Directions in Psychological Science* 11 (2002): 19–23.

[2.](#) A. L. Bronzaft, “The Effect of a Noise Abatement Program on Reading Ability,” *Journal of Environmental Psychology* 1, no.3 (1981): 215–22; A. L. Bronzaft, and D. P. McCarthy, “The Effect of Elevated Train Noise on Reading Ability,” *Environment and Behavior* 7, no.4 (1975): 517–28. doi:10.1177/001391657500700406.

[3.](#) 认知心理学关注的一个重点是干扰对认知表现的影响，特别是它与注意力和认知负荷的相互关系。虽说微量干扰被证明会造成巨大影响，但这种说法其实不是直觉所能想到的。干扰影响相关的实验范围很广，从反应时间实验到模拟装置の利用，再到田野调查，考察了各种各样的任务，包括视觉、听觉、疼痛感知、开车、手术、工作表现和学习成果。

[4.](#) 莱维和同事们的一些研究记录了在记忆负荷很重时，显著的干扰会增加对注意力的影响。这项研究把视觉注意力和工作记忆这两项不相

干的任务联合起来。可以预测，当工作记忆的负荷增加时，我们防止视觉被干扰的能力会减弱。请想象你参与了这个非同寻常的实验。你注视着电脑屏幕，看到了一连串数字，比如0、3、1、2、4，你需要记住这些数字。然后，你又在电脑屏幕上看到了一些名人的名字，你需要回答这些名人是流行歌手还是政治家。这些名字旁边还配有他们的脸部照片，实验人员要你忽视这些照片。之后，屏幕上又出现一个数字，比如2，你的任务是回答你记忆里这个数字后面的一个数，在这个例子中就是4。为了让这个实验更有趣，做了两处变动。第一个是负荷操纵：当记忆负荷很重时，要记忆的数字串在每一次实验时是不一样的，当记忆负荷很轻时，每次出现的数字串是固定的：0、1、2、3、4。很明显，你可以很轻松地记下固定的数字串，而需要很积极地记下变动的数字串。另外，忽略脸部照片的条件变了：在干扰很小时，面部照片和名字是统一的，比如陈奕迅的照片对应的是陈奕迅的名字，周杰伦的照片与名字也是对应的。但是当干扰很大时，陈奕迅的照片对应的是周杰伦的名字，反过来了，这样干扰就大了！实验表明，当工作记忆负荷增加时，注意力面对的干扰更大。当记忆负荷很重时，这种名字和照片的不一致带来的影响会更大。N. Lavie, “Distracted and Confused?: Selective Attention under Load,” *Trends in Cognitive Sciences* 9, no.2 (2005): 75–82.

5. R. M. Piech, M. T. Pastorino, and D. H. Zald, “All I Saw Was the Cake: Hunger Effects on Attentional Capture by Visual Food Cues,” *Appetite* 54, no.3 (2010): 579.在注意力研究领域，某些心理和物理事件会俘获注意力是一个经久不衰的话题，这有助于理解知觉和认知里目标导向和刺激导向的交互过程。

6. 本书出版时这篇研究报告还没公布：C. J. Bryan, S. Mullainathan, and E. Shafir, “Tempting Food, Cognitive Load and Impaired Decision-Making,” invited talk at the United States Department of Agriculture, Economic Research Service, Washington, D.C., April 2010.

7. 389人参与了这项研究。节食者找词所用时间的区别很显著，在看到食物的词与看到中性词之后的差别很大($p = 0.003$)。节食者与非节食者在这两项试验中的区别也很显著($p = 0.047$)。实验激励参与者努力去找出尽可能多的词。

8. 认知科学家和脑科学家们一直专注于研究认知机制和脑的结构，由此揭开指导行为的执行或认知控制之谜。比如，可以参考G. J. DiGirolamo, “Executive Attention: Conflict, Target Detection, and Cognitive

Control,” in *The Attentive Brain*, ed. Raja Parasuraman (Cambridge, Mass.: MIT Press, 1998), 401–23。

9. J. Raven et al., *Manual for Raven's Progressive Matrices and Vocabulary Scales*, research supplement no.3, 2nd/3rd edition (Oxford: Oxford Psychologists Press/San Antonio, Tex.: The Psychological Corporation, 1990/2000)。

10. J. Raven, “The Raven's Progressive Matrices: Change and Stability over Culture and Time,” *Cognitive Psychology* 41, no.1 (2000): 1–48.

11. 参考文献同上条。值得注意的是，研究者们论证说教育只能解释智商测试分数的一小部分，可参考J. R. Flynn, “Massive IQ Gains in 14 Nations: What IQ Tests Really Measure,” *Psychological Bulletin* 101 (1987): 171–91；关于环境和文化对智商的影响，可参考Richard Nisbett, *Intelligence and How to Get It: Why Schools and Cultures Count* (New York: W. W. Norton, 2010)。

12. Anandi Mani, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao, “Poverty Impedes Cognitive Function” (working paper, 2012)。

13. A. Lusardi, D. J. Schneider, and P. Tufano, *Financially Fragile Households: Evidence and Implications* (National Bureau of Economic Research, Working Paper No. 17072, May 2011)。

14. 科恩d值（Cohen's d）表示的效应范围在0.88到0.94之间。科恩d值可以算作不同样本标准差的均值之间的差异。

15. L. Linde and M. Bergstrom, “The Effect of One Night without Sleep on Problem-Solving and Immediate Recall,” *Psychological Research* 54, no. 2 (1992): 127–36。一般而言，大量研究已经表明缺觉对认知过程的有害影响，不论是对注意力和记忆力，还是对计划和做决策，都有危害。可参考Gerard A. Kerkhof and Hans Van Dongen, *Human Sleep and Cognition: Basic Research* 185 (Amsterdam: Elsevier Science, 2010)。

16. “What Is a Genius IQ Score?” *About.com*, retrieved October 23, 2012.

[17.](#) W. Mischel, E. B. Ebbesen, and A. Raskoff Zeiss, “Cognitive and Attentional Mechanisms in Delay of Gratification,” *Journal of Personality and Social Psychology* 21, no.2 (1972): 204.在后续几年的研究中，之前参加实验的孩子们长大了，米歇尔和同事们发现认知和社交能力具有可预测性，这丰富了科学家们对决定行为的因素的思考。W. Mischel, Y. Shoda, and P. K. Peake, “The Nature of Adolescent Competencies Predicted by Preschool Delay of Gratification,” *Journal of Personality and Social Psychology* 54, no.4 (April 1988): 687–96.

[18.](#) Thomas C. Schelling, *Choice and Consequence* (Boston: Harvard University Press, 1985).

[19.](#) 罗伊·鲍姆斯特（Roy Baumeister）及合作者们已经进行了很多研究，研究包括自我损耗、执行和自控的持续与减少。R. F. Baumeister and J. Tierney, *Willpower: Rediscovering the Greatest Human Strength* (New York: Penguin Press, 2011).

[20.](#) Mischel, Ebbesen, and Raskoff Zeiss, “Cognitive and Attentional Mechanisms.”

[21.](#) J. Lehrer, “DON’T!” *New Yorker*, May 18, 2009.

[22.](#) B. Shiv and A. Fedorikhin, “Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making,” *Journal of Consumer Research* 26, no.3 (1999):278–92. doi:10.1086/209563.

[23.](#) W. von Hippel and K. Gonsalkorale, “‘That Is Bloody Revolting!’: Inhibitory Control of Thoughts Better Left Unsaid,” *Psychological Science* 16, no.7 (2005): 497–500. doi:10.1111/j.0956-7976.2005.01563.x.

[24.](#) Mani, Mullainathan, Shafir, and Zhao, “Poverty Impedes Cognitive Function.”

[25.](#) 标准的斯特鲁普测验要求实验对象回答标识的字体颜色。如果用蓝色字体写下XKYD标识时，实验对象就要回答“蓝色”。斯特鲁普测验的挑战在于某些标识本身是表示颜色的标识，比如用蓝色字体写的RED（这是“红色”的英文单词）。可参考Colin M. MacLeod, “Half a Century of Research on the Stroop Effect: An Integrative Review,”

Psychological Bulletin 109, no.2 (March 1991): 163–203。关于斯特鲁普测验，常被说起的一件逸闻就是，它在冷战时曾被用来检测苏联间谍，比如用红色字体来写俄语里的“蓝色”，对于普通美国人来说，看到这个俄语词没有影响，但是对于苏联间谍而言，就会造成认知阻力。

26. Mani, Mullainathan, Shafir, and Zhao, “Poverty Impedes Cognitive Function.”

27. K. Alaimo, C. M. Olson, and E. A. Frongillo Jr., “Food Insufficiency and American School-Aged Children's Cognitive, Academic, and Psychosocial Development,” *Pediatrics* 108, no.1 (2001): 44–53.

28. 一个问题是收获后的实验对象还会接受第二次实验。收获后的农民的表现改善只会被归因为实验经历。为了控制实验本身的影响，我们对100位随机挑选的农民隐瞒了实验真相，然后让他们接受第一次实验。由于他们是被随机挑选的，我们就拿他们作参照，对比收获前的农民，但得出的结论一样，这就说明实验经历并不影响农民们的表现。我们也研究了一个收获后的农民样本，由于他还有尚未还清的债务，因而依然处于贫穷状态。那些农民在收获前后的行为一样，这就说明收获机制并不影响我们的研究结论。

29. N. Kusz, “The Fat Lady Sings,” in *The Bitch in the House: 26 Women Tell the Truth About Sex, Solitude, Work, Motherhood, and Marriage* (New York: William Morrow, 2002).

30. D. Borchmann, *Fasting, Restrained Eating, and Cognitive Performance—A Literature Review from 1998 to 2006*.

31. 一项研究发现，给节食者一块巧克力棒确实会恶化他们的认知表现。这被归因于，节食者拿到巧克力棒以后要更专注于食物（“我需要什么东西来让自己不吃巧克力棒？”），可参考N. Jones and P. J. Rogers, “Preoccupation, Food, and Failure: An Investigation of Cognitive Performance Deficits in Dieters,” *International Journal of Eating Disorders* 33, no.2 (March 2003): 185–92。

32. J. T. Cacioppo, J. M. Ernst, M. H. Burleson, M. K. McClintock, W. B. Malarkey, L. C. Hawkley, R. B. Kowalewski et al., “Lonely Traits and Concomitant Physiological Processes: The MacArthur Social Neuroscience

Studies,” *International Journal of Psychophysiology* 35, no.2 (2000): 143–54.

[33.](#) 同上条。

[34.](#) John T. Cacioppo and William Patrick, *Loneliness: Human Nature and the Need for Social Connection* (New York: W.W.Norton, 2009).

[35.](#) R. F. Baumeister, J. M. Twenge, and C. K. Nuss, “Effects of Social Exclusion on Cognitive Processes: Anticipated Aloneness Reduces Intelligent Thought,” *Journal of Personality and Social Psychology* 83, no. 4 (2002): 817.

[36.](#) R. F. Baumeister, C. N. DeWall, N. J. Ciarocco, and J. M. Twenge, “Social Exclusion Impairs Self-Regulation,” *Journal of Personality and Social Psychology* 88, no.4 (2005): 589.

[37.](#) W. Lauder, K. Mummery, M. Jones, and C. Caperchione, “A Comparison of Health Behaviours in Lonely and Non-Lonely Populations,” *Psychology, Health and Medicine* 11, no.2 (2006): 233–45.
doi:10.1080/13548500500266607.

[38.](#) Mani, Mullainathan, Shafir, and Zhao, “Poverty Impedes Cognitive Function.”

[39.](#) L. E. Bourne and R. A. Yaroush, “Stress and Cognition: A Cognitive Psychological Perspective,” unpublished manuscript, NASA grant NAG2-1561 (2003); Bruce McEwen, *The End of Stress as We Know It* (New York: Joseph Henry Press/Dana Press, 2002).

[40.](#) Robert M. Sapolsky, *Why Zebras Don’t Get Ulcers* (New York: Henry Holt, 1994).

[41.](#) S. Vijayraghavan, M. Wang, S. G. Birnbaum, G. V. Williams, and A. F. T. Arnsten, “Inverted-U Dopamine D1 Receptor Actions on Prefrontal Neurons Engaged in Working Memory,” *Nature Neuroscience* 10, no.3 (2007):376–84. doi:10.1038/nn1846.

[42.](#) Robert and J. Hockey, “Compensatory Control in the Regulation of Human Performance under Stress and High Workload: A Cognitive-Energetical Framework,” *Biological Psychology* 45, no.1 (1997):73–93.

第3章 装箱、余闲和权衡式思维

[1.](#) Dwight D. Eisenhower, *The Chance for Peace* (U.S. Government Printing Office, April 16, 1953).

[2.](#) 我们只调查了100位上班族: $p < 0.05$ 。

[3.](#) Stephen Spiller, “Opportunity Cost Consideration,” *Journal of Consumer Research* (forthcoming).

[4.](#) 2009年在印度的泰米尔纳德邦调查了274个人。在这里用收入来比较城乡居民之间的差异,在收入上,二者存在6倍的差异。在购买搅拌机的问题上,差异很显著, $p < 0.01$ 。而在购买电视机的问题上,就没有经济和统计上的显著差异(分别是58.6%和60.8%)。

[5.](#) K. Van Ittersum, J. Pennings, and B. Wansink, “Trying Harder and Doing Worse: How Grocery Shoppers Track In-Store Spending,” *Journal of Marketing* (2010).

[6.](#) G. Antonides, I. Manon de Groot, and W. Fred van Raaij, “Mental Budgeting and the Management of Household Finance,” *Journal of Economic Psychology* 32, no.4 (2011): 546–55.
doi:10.1016/j.joep.2011.04.001.

[7.](#) Simpler saving: “The 60% Solution,” *MSNMoney*, October 24, 2012.

[8.](#) 另一种对可预测的余闲的研究方式可见G. Zauberman and J. G. Lynch, “Resource Slack and Propensity to Discount Delayed Investments of Time Versus Money,” *Journal of Experimental Psychology: General* 134, no.1 (2005): 23–37。

[9.](#) J. M. Graham, *The Hive and the Honey Bee* (Hamilton, Ill.: Dadant & Sons, 1992).

[10.](#) *Plywood Standards*, Voluntary Product Standard PS 1-09, National

Institute of Standards and Technology, U.S. Department of Commerce.

11. H. J. Brockmann, “Diversity in the Nesting Behavior of Mud-Daubers (*Trypoxylon politum* Say; Sphecidae),” *Florida Entomologist* 63, no. 1 (1980):53–64.

12. 这一对余闲的解释让我们联想到赫伯特·西蒙的论证，他说人们并不要最大化效用：他们满意就行，得到足够的效用就行。可参考 Herbert A. Simon, “Rational Choice and the Structure of the Environment,” *Psychological Review* 63, no. 2 (1956): 129。在西蒙看来，人们缺乏最大化效用的认知资源。用他的话来说就是，稀缺带来了不那么令人满意的行为。这一观点确实抓住了余闲的某些要素，但稀缺的影响更加具有自主性，也没有他说的那么好控制。我们会在下文中论证，不可控性对于理解稀缺至关重要。

13. George Carlin, *Brain Droppings* (New York: Hyperion, 1997), 37.

14. Brian Wansink, S. Adam Brasel, and Stephen Amjad, “The Mystery of the Cabinet Castaway: Why We Buy Products We Never Use,” *Journal of Family and Consumer Science* 92, no. 1 (2000): 104–8. 对于我们会有这么多的橱柜遗弃品，经济学家们给出的一个解释是“期权的价值”（Option Value），虽然我们买那么多不一定用得着的东西，但买来以后我们就有了选择的机会以防万一，这就是期权的价值。心理学的解释要比这复杂。我们会说，在稀缺情况下，人们会更加仔细、更加专注地思考会不会用得上，仔细地评估选择的价值，而不仅仅只是为了以防万一那么简单。

15. SSA | 2012 SSA Fact Sheet.

16. 同上条。

17. J. Mooallem, “The Self-Storage Self,” *New York Times*, September 6, 2009.

18. D. A. Redelmeier and E. Shafir, “Medical Decision Making in Situations That Offer Multiple Alternatives,” *JAMA—Journal of the American Medical Association* 273, no.4 (1995): 302–5.

[19.](#) M. Friedman and R. Friedman, *Free to Choose: A Personal Statement* (Orlando, Fla.: Mariner Books, 1990).

[20.](#) R. Buehler, D. Griffin, and M. Ross, “Exploring the ‘Planning Fallacy’: Why People Underestimate Their Task Completion Times,” *Journal of Personality and Social Psychology* 67, no.3 (1994): 366.

[21.](#) M. Sigman, “Response Time Distributions in Rapid Chess: A Large-Scale Decision-Making Experiment,” *Frontiers in Neuroscience* 4 (2010). doi:10.3389/fnins.2010.00060.

[22.](#) A. Banerjee and S. Mullainathan, *The Shape of Temptation: Implications for the Economic Lives of the Poor* (Working Paper No. w15973, National Bureau of Economic Research, 2010).

[23.](#) 当风险很高时，人们的行为表现其实会不一样，这一事实很早被用来反驳心理学对社会现象的解释。在过去20年里，已有研究表明人们的心理偏差会影响决策，对退休或健康方面的决策影响是决定性的。

[24.](#) 可以通过对比线性规划和整数规划来理解计算复杂性这一概念。在线性规划中，事物可以被划分为无穷小，这是对粒度概念的逻辑延拓。而在整数规划中，事物必须被放入一个固定的单位里，这是对大体积概念的逻辑延拓。计算科学家们在精确的数学意义上论证了，整数规划比线性规划本质上要困难得多。可参考Alexander Schrijver, *Theory of Linear and Integer Programming* (West Sussex, England: John Wiley & Sons, 1998)。

[25.](#) Henry David Thoreau, *Walden* (Yale University Press, 2006), 87.梭罗本人在这一观察中得到的是另一个教训。他不是鼓励增加财富，而是鼓励减少欲望。用我们的话来说，有两种获得余闲的方式，要么买个更大的行李箱，要么减少你想装的东西。

第4章 行为经济学告诉我们的道理

[1.](#) 在这本书里，我们只是简单地使用名义汇率来描述国外货币的价值（在这个例子中就是以美元来衡量卢比的价值）。对于某些情形，名义汇率就够用了，比如这个例子里亚历克斯用美元的名义汇率来衡量卢比的价值。但在有些情形中，名义汇率就有误导性，因为名义汇率无法

衡量不同国家的价格水平的差异。比如，卢比走高，是由于印度的很多商品价格都更低了。要评估不同国家的收入差异，大多数经济学家不仅会调整汇率，也会调整购买力平价——对各国物价水平差异的衡量方式。由于本书的主旨不在于详细比较各国的收入水平，我们为了方便读者阅读，就只是简单地使用名义汇率了，但是读者要在自己头脑里意识到这个区别。

2. 这就是特沃斯基和卡尼曼提出的著名的“夹克和计算器问题”。可参考A. Tversky and D. Kahneman, “The Framing of Decisions and the Psychology of Choice,” *Science* 211, no. 4481 (1981): 453–58。也可参考R. Thaler, “Mental Accounting Matters,” *Journal of Behavioral Decision Making* 12 (1999):183–206。

3. Ofer H. Azar, “Relative Thinking Theory,” *The Journal of Socio-Economics* 36, no.1 (2007): 1–14.

4. 一些研究利用激励理论得出了类似的效应。在某个研究案例中，要调查对象去回答代数问题，答对了可以获得6美分，并对调查对象进行分组，只要参加这个实验，就可以获得1美元、3美元或10美元的“出场费”，按照出场费的不同进行分组。对于出场费1美元的组来说，答对一道题得6美分看起来比较多了，而对于出场费10美元的组来说，这看起来就不算多少。实验结果表明，确实是出场费1美元的组更加努力去答题，也回答了更多问题。某些有幽默感的研究人员去美国经济学会的夏季会议上做了这个实验，让那些职业经济学家当调查对象，得出的实验数据也相同。这就表明，经济学家并没有比其他人更善于做理性决策。

5. 相比较于低收入水平的调查对象（特伦顿慈善食堂），高收入水平的调查对象（普林斯顿火车站）在100美元和1000美元两种情况下，表现出来的差异更大，这个实验N=123。可参考C. C. Hall, *Decisions Under Poverty: A Behavioral Perspective on the Decision Making of the Poor* (PhD diss., Princeton University, 2008)。

6. 很可能这个结果是由于“天花板效应”，因为对于穷人而言，要提高赶路意愿的余地不大。即使价格上升了，依然不是100%的人都愿意赶路。

7. H. E. Ross, “Weber Then and Now,” *Perception* 24, no. 6 (1995):

599.

8. G. Trotta, “Some Laundry-Detergent Caps Can Lead to Overdosing,” June 5, 2009.

9. S. Grondin and P. R. Killeen, “Tracking Time with Song and Count: Different Weber Functions for Musicians and Nonmusicians,” *Attention, Perception, and Psychophysics* 71, no.7 (2009):1649–54.

10. B. Wansink and K. Van Ittersum, “Bottoms Up! The Influence of Elongation on Pouring and Consumption Volume,” *Journal of Consumer Research* 30, no.3 (2003): 455–63.

11. I. M. Rosa-Díaz, “Price Knowledge: Effects of Consumers’ Attitudes Towards Prices, Demographics, and Socio-cultural Characteristics,” *Journal of Product and Brand Management* 13, no.6 (2004): 406–28. doi:10.1108/10610420410560307.

12. 高收入和低收入的顾客答对率的差异具有统计显著性, $p < 0.05$, $N=104$ 。

13. Jacob Goldin and Tatiana Homonoff, “Smoke Gets in Your Eyes: Cigarette Tax Salience and Regressivity” *American Economic Journal: Economic Policy* 5, no.1, (February 2013): 302–36.

14. 如果标价更高, 穷人会付出更多注意力, 但这一点不是我们想说的一部分内容。这里有意思的内涵在于, 付出更多的注意力是怎样改变决策过程的, 是怎样改变那些已被记录下的“偏差”的。

15. J. K. Binkley and J. Bejnarowicz, “Consumer Price Awareness in Food Shopping: The Case of Quantity Surcharges,” *Journal of Retailing* 79, no.1 (2003): 27–35. doi:10.1016/S0022-4359(03)00005-8.

16. Consumer Reports, “Sold Short? Are You Getting Less Than You Think? Let Us Count the Ways,” February 2000, 24–26.

17. 同上条。

18. 丹·艾瑞里 (Dan Ariely) 说明白了对权衡思维的挑战, 可参

考: <http://bigthink.com/ideas/17458>.

[19](#). Shane Frederick, Nathan Novemsky, Jing Wang, Ravi Dhar, and Stephen Nowlis, “Opportunity Cost of Neglect,” *Journal of Consumer Research* 36, no. 4 (2009): 553–61.

[20](#). 有大量证据表明了知觉的背景依赖性。阿德尔森设计的“方块阴影”错觉图是我们最喜欢的例子之一。在此使用也获得了授权。读者也可以自己去做这个实验，登录 <http://web.mit.edu/persci/people/adelson/checkershadowillusion.html>。想要了解更多关于错觉背后的认知机制，可以参考Edward H. Adelson, “Lightness Perception and Lightness Illusions,” *The New Cognitive Neurosciences* (1999): 339。

[21](#). 这是基于理查德·塞勒的研究，见“Mental Accounting and Consumer Choice,” *Marketing Science* 4, no.3 (1985): 199–214。我们和阿努伊·沙阿一起在2012年搜集的数据。富人在不同的框架下具有显著的差异，而穷人并没有 $p < 0.01$ ($N = 0148$)。

[22](#). J. Hastings and J. M. Shapiro, *Mental Accounting and Consumer Choice: Evidence from Commodity Price Shocks* (Cambridge, Mass.: National Bureau of Economic Research, Working Paper No. 18248, 2012).

[23](#). 我们和阿努伊·沙阿在2012年搜集的数据支持这一预测。我们给实验对象提供了退税和股票收益不同的场景。富人对于这两种框架显示出了不同的倾向，而穷人并没有； $p < 0.05$ ($N = 141$)。

[24](#). 我们和沙阿在2012年搜集的数据。富人更倾向于选择历史成本，而穷人更倾向于处置成本；在两种情况下 $p < 0.05$ ($N = 98$)。

[25](#). E. Shafir and R. H. Thaler, Invest Now, Drink Later, Spend Never: On the Mental Accounting of Delayed Consumption, *Journal of Economic Psychology* 27 (2006): 694–712.

[26](#). Paul J. Ferraro and Laura O. Taylor, “Do Economists Recognize an Opportunity Cost When They See One? A Dismal Performance from the Dismal Science”(2005).

[27.](#) 这来自博客*Marginal Revolution*。
<http://marginalrevolution.com/marginalrevolution/2005/09/opportunitycos.html>.

第5章 借用与短视

[1.](#) J. A. Riis, *How the Other Half Lives* (Boston, Mass.: Bedford/St. Martin's, 2010).

[2.](#) 关于桑德拉对这个项目的描述可以参考
<http://www.responsiblelending.org/payday-lending/tools-resources/victims-2.html>。

[3.](#) M. Fellowes and M. Mabanta, *Banking on Wealth: America's New Retail Banking Infrastructure and Its Wealth-Building Potential* (Washington, D.C.: Brookings Institution, 2008).

[4.](#) 关于麦当劳跨国比较的数据可以参考*NationMaster*。

[5.](#) *Loxcel Starbucks Store Map FAQ*, retrieved from
<http://loxcel.com/sbux-faq.html>.

[6.](#) *Fast Facts*, retrieved October 24, 2012, from
<http://www.responsiblelending.org/payday-lending/tools-resources/fast-facts.html>.在这个行业里“回头客”特别多，有98%的贷款都流向了反复贷款的人。

[7.](#) Michael Barr, *No Slack* (Washington, D.C.: Brookings Institution Press, 2002)里面有精彩的论述。

[8.](#) K. Edin and L. Lein, *Making Ends Meet: How Single Mothers Survive Welfare and Low-Wage Work* (New York: Russell Sage Foundation Publications, 1997).关于美国穷人的经济生活，可参考Sarah Halpern-Meekin, Kathryn Edin, Laura Tach, and Jennifer Sykes, *It's Not Like I'm Poor: How Working Families Make Ends Meet in a Post-Welfare World* (Berkeley: University of California Press, forthcoming)。

[9.](#) Abhijit Banerjee, “Contracting Constraints, Credit Markets, and

Economic Development,” in *Advances in Economics and Econometrics: Theory and Application, Eighth World Congress of the Econometric Society*, vol.3, ed. Mathias Dewatripont, Lars Hansen and S. Turnovsky (Cambridge: Cambridge University, 2004), 1–46.

10. 另一个常用来解释过度借贷的理由是目光短浅，但这一叙述有意思的地方在于，这里所说的目光短浅——管窥，并不是普遍的人格特征。当面对稀缺的时候，每个人都会管窥。回忆一下，不同于目光短浅，造成管窥的力量同样也会带来专注红利，管窥还会带来积极的后果。

11. Anuj Shah, Sendhil Mullainathan, and Eldar Shafir, “Some Consequences of Having Too Little,” *Science* 338 (2013): 682–85.

12. 关于“现时偏差”和时间贴现的其他模型，可参考一篇常用的概述，见Shane Frederick and George Loewenstein, “Time Discounting and Time Preference: A Critical Review,” *Journal of Economic Literature* (2002)。

13. R. E. Bohn and R. Jaikumar, *Firefighting by Knowledge Workers* (Information Storage Industry Center, Graduate School of International Relations and Pacific Studies, University of California, 2000), retrieved from <http://isic.ucsd.edu/pdf/firefighting.pdf>.

14. S. R. Covey, *The Seven Habits of Highly Effective People* (New York: Free Press, 2004).

15. *Bridges—Report Card for America’s Infrastructure*, retrieved from <http://www.infrastructurereportcard.org/fact-sheet/bridges>.

16. Roger Buehler, Dale Griffin, and Michael Ross, “Inside the Planning Fallacy: The auses and Consequences of Optimistic Time Predictions,” in *Heuristics and Biases: The Psychology of Intuitive Judgment*, ed. Thomas Gilovich, Dale Griffin, and Daniel Kahneman (Cambridge: Cambridge University Press, 2002), 250–70; D. Lovallo, and D. Kahneman, “Delusions of Success,” *Harvard Business Review* (2003): 1–8.但这些研究都没有明确研究稀缺，不过很自然，在稀缺情境下，尤其是管窥情况下，更容易出现错误规划。

第6章 稀缺陷阱

1. Steven Wright. *In W. Way, Oxymorons and Other Contradictions*. (Bloomington, Ind.: Author House, 2005).

2. 这些数据来自Dean Karlan and Sendhil Mullainathan, “Debt Traps” (working paper, 2012)。

3. 在本书中，我们换算成美元的等值额时，我们只是简单地利用名义汇率。但很多专家认为这会带来误解，因为不同国家的人要面对不同的物价水平。所以，比如这个例子中的小贩要购买其他食物或商品的价格也会更低。因此，小贩的名义收入并没有确切地反映其购买力水平。经济学家们建议使用购买力平价来代替名义汇率。在印度的例子中，按照购买力平价来计算，小贩的收入大约会高出2.5倍。

4. 经济学家特别是发展经济学家已经关注到“贫困陷阱”，就是开始变穷的人会停留在贫穷状态。一个常被讨论的机制原理就是，能赚钱的投资机会需要一定数量的资本投入。富人拥有足够的资本去投资，而穷人很难节省出钱来投资。还有些机制研究包括抱负和目光短浅的问题。相关的参考文献可见于Debraj Ray, “Development Economics,” *The New Palgrave Dictionary of Economics*, ed. Lawrence Blume and Steven Durlauf (2007)。

5. Michael Faye and Sendhil Mullainathan, “Demand and Use of Credit in Rural India: An Experimental Analysis” (working paper, Harvard University, 2008)。

6. Daryl Collins, Jonathan Morduch, Stuart Rutherford, and Orlanda Ruthven, *Portfolios of the Poor: How the World's Poor Live on \$2 a Day* (Princeton, N.J.: Princeton University Press, 2010)。

7. 尽管发展中国家的时间利用数据很难获得，但依然有一些精彩的研究，可见于Quentin Wodon and Mark Blackden, *Gender, Time Use, and Poverty in Sub-Saharan Africa* (Washington, D.C.: World Bank Press, 2006)。

8. M. Muraven and R. F. Baumeister, “Self-Regulation and Depletion of Limited Resources: Does Self-Control Resemble a Muscle?” *Psychological*

Bulletin 126, no. 2 (2000): 247–59. doi:10.1037//0033-2909.126.2.247.

9. K. D. Vohs and T. F. Heatherton, “Self-Regulatory Failure: A Resource-Depletion Approach,” *Psychological Science* 11, no.3 (2000): 249–54.

10. D. Collins et al., *Portfolios of the Poor*.

11. New Amsterdam Consulting, “Stability First Pilot Test: Pre-Test Interviews Narrative Report” (March 2012).

12. A. Lusardi, D. J. Schneider, and P. Tufano, *Financially Fragile Households: Evidence and Implications* (National Bureau of Economic Research, 2011), retrieved from <http://www.nber.org/papers/w17072>.

13. John T. Cacioppo and William Patrick, *Loneliness: Human Nature, and the Need for Social Connection* (New York: W.W.Norton, 2009).

14. J. Friedman, “How Did Tom Amberry Set the World Free Throw Record?” *Sports Illustrated*, October 17, 1994, retrieved from <http://sportsillustrated.cnn.com/vault/article/magazine/MAG1005796/index.ht>

15. Bruce Bowen, *Basketball-Reference.com*, retrieved October 31, 2012, from <http://www.basketball-reference.com/players/b/bowenbr01.html>.

16. S. L. Beilock, A. R. McConnell et al., “Stereotype Threat and Sport: Can Athletic Performance Be Threatened,” *Journal of Sport and Exercise Psychology* 26, no.4 (2004): 597–609.

17. R. M. Yerkes and J. D. Dodson, “The Relation of Strength of Stimulus to Rapidity of Habit-Formation,” *Journal of Comparative Neurology and Psychology* 18, no.5(1908): 459–82.

18. Daniel M. Wegner, David J. Schneider, Samuel R. Carter, and Teri L. White, “Paradoxical Effects of Thought Suppression,” *Journal of Personality and Social Psychology* 53, no.1(1987): 5–13; D. M. Wegner, *White Bears and Other Unwanted Thoughts: Suppression, Obsession, and the Psychology of Mental Control* (New York: Viking, 1989).

第7章 穷人为什么穷

1. J. Carr and L. Greeves, *Only Joking: What's So Funny About Making People Laugh?* (New York: Gotham Books, 2006).

2. *Levels and Trends in Child Mortality* (Washington, D.C.: The UN Inter-Agency Group for Child Mortality Estimation [IGME], 2010).

3. <http://www.globalissues.org/article/26/poverty-facts-and-stats>.

4. 世界银行划定的贫困线是每天2.5美元。这里关注的是“绝对”贫困。照此计算，美国不存在处于贫困状态的儿童。关于全球贫困问题的更多研究可参考Anup Shah, “Poverty Facts and Stats,” *Global Issues* 26 (2008)。也可参考Abhijit Banerjee and Esther Duflo, *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty* (New York: PublicAffairs, 2011)。

5. Mark R. Rank and Thomas A. Hirschl, “Estimating the Risk of Food Stamp Use and Impoverishment during Childhood,” *Archives of Pediatrics and Adolescent Medicine* 163, no.11 (2009): 994.

6. Alisha Coleman-Jensen et al., “House hold Food Security in the United States in 2010,” *USDA-ERS Economic Research Report* 125 (2011).

7. B. Ritz and F. Yu, “The Effect of Ambient Carbon Monoxide on Low Birth Weight Among Children Born in Southern California Between 1989 and 1993,” *Environmental Health Perspectives* 107, no.1 (1999): 17.

8. 关于贫困背后的因素，另一项富有原创性的研究可参考Charles Karelis, *The Persistence of Poverty: Why the Economics of the Well-Off Can't Help the Poor* (New Haven: Yale University Press, 2009)。

9. *International Diabetes Federation, Atlas*.
<http://www.diabetesatlas.org/content/some-285-million-peopleworldwide-will-live-diabetes-2010>.

10. 之所以有这么大的估算范围，因为按时服药的人群比例取决于研究样本所包含的人数。计量按时服药的方式——比如患者自己报告、

药物补充率、电子设备监控——也会影响估算结果。基础研究可参考 Eduardo Sabaté, ed., *Adherence to Long-Term Therapies: Evidence for Action* (Geneva: World Health Organization, 2003)。这本书也囊括了其他疾病的按时吃药数据。

11. 很难从这些研究中概述除草对所有农民的好处，这些研究依赖的是模型或横截面数据。要测度这个好处，设计出随机化控制实验就特别有用了。关于非洲农民的估算，可参考L. P. Gianessi et al., “Solving Africa's Weed Problem: Increasing Crop Production and Improving the Lives of Women,” *Proceedings of “Agriculture: Africa’s ‘engine for growth’—Plant Science and Biotechnology Hold the Key,”* Rothamsted Research, Harpenden, UK, October 12–14, 2009 (Association of Applied Biologists, 2009)。

12. D. E. Johnson, “Weed Management in Small Holder Rice Production in the Tropics,” Natural Resources Institute, University of Greenwich Gatham, Kent, UK 11 (1996), retrieved from <http://ipmworld.umn.edu/chapters/johnson.htm>.

13. J. Lexmond, L. Bazalgette, and J. Margo, *The Home Front* (London: Demos, 2011).

14. 早期的研究见J. Garbarino, “A Preliminary Study of Some Ecological Correlates of Child Abuse: The Impact of Socioeconomic Stress on Mothers,” *Child Development* (1976): 178–85。近期利用更多数据研究这一问题的有Christina Paxson and Jane Waldfogel, “Work, Welfare, and Child Maltreatment,” *Journal of Labor Economics* 20, no.3 (July 2002): 435–74。

15. J. S. Lee and N. K. Bowen, “Parent Involvement, Cultural Capital, and the Achievement Gap Among Elementary School Children,” *American Educational Research Journal* 43, no.2 (2006): 193–218.

16. A. T. Clarke and B. Kurtz-Costes, “Television Viewing, Educational Quality of the Home Environment, and School Readiness,” *Journal of Educational Research* (1997): 279–85.

17. A. Drewnowski and S. E. Specter, “Poverty and Obesity: The Role

of Energy Density and Energy Costs,” *The American Journal of Clinical Nutrition* 79, no.1 (2004): 6–16.

[18.](#) R. Tabberer, “Childhood Poverty and School Attainment, Causal Effect and Impact on Lifetime Inequality,” in *Persistent Poverty and Lifetime Inequality: The Evidence—Proceedings from a Workshop Held at HM Treasury, Chaired by Professor John Hills, Director of the ESRC Research Centre for Analysis of Social Exclusion* (1998).

[19.](#) N. Adler, J. Stewart, S. Cohen, M. Cullen, A. D. Roux, W. Dow, and D. Williams, “Reaching for a Healthier Life: Facts on Socioeconomic Status and Health in the U.S.,” *The John D. and Catherine T. MacArthur Foundation Research Network on Socioeconomic Status and Health* 43 (2007).

[20.](#) 很多地方都在观测收入与洗手、处理饮水之间的关系。在秘鲁，有一项研究关注母亲或其他人照顾小孩的行为，结果表明，只有46%照顾小孩的人在上厕所后洗手。甚至在这一数据内，也表现出与收入很强的关联：最高收入的人群里有56.5%的人会在上厕所后洗手，而底层收入人群里只有34%的人会洗手。在清洗小孩屁股或饭前洗手方面，他们也证实了类似的差异。可参考Sebastian Galiani and Alexandra Orsola-Vidal, “Scaling Up Handwashing Behavior,” *Global Scaling Up Handwashing Project, Water and Sanitation Program* (Washington, D.C., 2010)。

[21.](#) Adler et al., “Reaching for a Healthier Life.”

[22.](#) John M. Darley and Paget H. Gross, “A Hypothesis-Confirming Bias in Labeling Effects,” *Journal of Personality and Social Psychology* 44, no.1(1983): 20–33.

[23.](#) R. L. Repetti, “Short-Term and Long-Term Processes Linking Job Stressors to Father–Child Interaction,” *Social Development* 3, no.1 (2006): 1–15.

[24.](#) L. A. Gennetian, G. Duncan, V. Knox, W. Vargas, E. Clark-Kauffman, and A. S. London, “How Welfare Policies Affect Adolescents’ School Outcomes: A Synthesis of Evidence from Experimental Studies,”

Journal of Research on Adolescence 14, no. 4 (2004): 399–423.

[25.](#) M. Siahpush, H. H. Yong, R. Borland, J. L. Reid, and D. Hammond, “Smokers with Financial Stress Are More Likely to Want to Quit but Less Likely to Try or Succeed: Findings from the International Tobacco Control (ITC) Four Country Survey,” *Addiction* 104, no.8 (2009): 1382–90.

[26.](#) Jens Ludwig, et al. “Neighborhoods, Obesity, and Diabetes—A Randomized Social Experiment,” *New England Journal of Medicine* 365, no. 16 (2011): 1509–19.

[27.](#) R. T. Gross and T. D. Borkovec, “Effects of a Cognitive Intrusion Manipulation on the Sleep-Onset Latency of Good Sleepers,” *Behavior Therapy* 13, no.1(1982): 112–16.

[28.](#) F. N. Watts, K. Coyle, and M. P. East, “The Contribution of Worry to Insomnia,” *British Journal of Clinical Psychology* 33 no.2 (2011): 211–20.

[29.](#) J. T. Cacioppo, L. C. Hawkley, G. G. Berntson, J. M. Ernst, A. C. Gibbs, R. Stickgold, and J. A. Hobson, “Do Lonely Days Invade the Nights? Potential Social Modulation of Sleep Efficiency,” *Psychological Science* 13, no. 4(2002):384–87.

[30.](#) N. P. Patel, M. A. Grandner, D. Xie, C. C. Branas, and N. Gooneratne, “Sleep Disparity in the Population: Poor Sleep Quality Is Strongly Associated with Poverty and Ethnicity,” *BMC Public Health* 10 (2010): 475–75.

[31.](#) G. Belenky, T. J. Balkin, D. P. Redmond, H. C. Sing, M. L. Thomas, D. R. Thorne, and N. J. Wesensten, “Sustaining Performance During Continuous Operations: The U.S. Army's Sleep Management System,” in *Managing Fatigue in Transportation. Proceedings of the 3rd Fatigue in Transportation Conference*(1998).

[32.](#) Alaska Oil Spill Commission, *Spill: The Wreck of the Exxon Valdez*, vol.3 (State of Alaska, 1990).关于睡眠方面的文献一项整体探讨可见于 William C. Dement and Christopher Vaughan, *The Promise of Sleep: A Pioneer in Sleep Medicine Explores the Vital Connection Between Health,*

Happiness, and a Good Night's Sleep (New York: Dell, 1999)。

33. Hans PA van Dongen et al., “The Cumulative Cost of Additional Wakefulness: Dose-Response Effects on Neurobehavioral Functions and Sleep Physiology from Chronic Sleep Restriction and Total Sleep Deprivation,” *SLEEP* 26, no.2(2003): 117–29.关于长期的睡眠紊乱的研究综述可参考D. F. Dinges, N. L. Rogers, and M. D. Baynard, “Chronic Sleep Deprivation,” *Principles and Practice of Sleep Medicine* 4(2005): 67–76.

34. 越来越多的研究论证，早期的童年经历会影响大脑发育。最近的研究可见Clancy Blair et al., “Salivary Cortisol Mediates Effects of Poverty and Parenting on Executive Functions in Early Childhood,” *Childhood Development* 82, no.6 (November/December 2011): 1970–84。我们的研究表明，除了这些影响，贫困对认知功能具有很大的直接影响。

第8章 如何让穷人摆脱贫穷

1. A. Chapanis, “Psychology and the Instrument Panel,” *Scientific American* 188(1953): 74–82.

2. Burt S. Barnow and Christopher T. King, eds., *Improving the Odds: Increasing the Effectiveness of Publicly Funded Training* (Washington, D.C.: Urban Institute Press, 2000).

3. 最近两项关于小微金融的影响评估从量化上证明了这些潜在的问题，见Dean Karlan and Jonathan Zinman, “Microcredit in Theory and Practice: Using Randomized Credit Scoring for Impact Evaluation,” *Science* 332, no.6035(2011): 1278–84; Abhijit Banerjee et al., “The Miracle of Microfinance? Evidence from a Randomized Evaluation” (MIT working paper, 2010)。

4. 有些论据不需要借助于稀缺心理学就可以得出。大多数政策制定都假设了人是完全理性的。仅仅只要假设人具有天然的心理局限，就可以提高政策制定的水平。这一观点最近有很精彩的研究，见Richard H. Thaler and Cass R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (New Haven, Conn.: Yale University Press, 2008)。也可参考Eldar Shafir, ed., *The Behavioral Foundations of Public Policy* (Princeton, N.J.: Princeton University Press, 2012)。我们之前用了这一逻辑。

辑去论证，只要理解穷人有和其他人一样的心理缺陷，就可以更好地理解贫困。可参考Marianne Bertrand, Sendhil Mullainathan, and Eldar Shafir, “A Behavioral-Economics View of Poverty,” *American Economic Review* (2004): 419–23。通过挤压带宽，稀缺会进一步放大这些论据。在贫困的情境里，富有心理学洞见的政策特别重要。

5. D. Ellwood and R. Haskins, *A Look Back at Welfare Reform*, *IPRNews* (Winter 2008), retrieved from <http://www.ipr.northwestern.edu/publications/newsletter/iprn0801/dppl.html>.

6. A. Cappelen, O. Maetad, and B. Tungodden, “Demand for Childhood Vaccination—Insights from Behavioral Economics,” in *Forum for Development Studies* 37, no.3 (November 2010): 349–64. (Oslo: Routledge, 2010).

7. L. B. Rawlings and G. M. Rubio, “Evaluating the Impact of Conditional Cash Transfer Programs,” *The World Bank Research Observer* 20, no. 1 (2005):29–55.

8. A. Drexler, G. Fischer, and A. Schoar, *Keeping It Simple: Financial Literacy and Rules of Thumb* (London: Centre for Economic Policy Research, 2010).

9. *Emergency Hand Loan: A Product Design Case Study*, Financial Access Initiative, ideas42 and IFC. Discussion and document at <http://www.financialaccess.org/blog/2011/05/product-design-poor-emergency-hand-loan>.

10. S. Baird, J. De Hoop, and B. Ozler, “Income Shocks and Adolescent Mental Health,” *World Bank Policy Research Working Paper Series*, no. 5644 (2011).

11. 关于美国的反复接受救济的比例已经被广泛研究了，比如参考J. Cao, “Welfare Reciprocity and Welfare Recidivism: An Analysis of the NLSY Data,” *Institute for Research on Poverty Discussion Papers* 1081–96, University of Wisconsin Institute for Research on Poverty (March 1996)。

12. “走向机会”（Moving to Opportunity）这个项目对幸福指数有积

极影响，但对经济上的自给自足没有影响。可参考J. Ludwig, G. J. Duncan, L. A. Gennetian, L. F. Katz, R. Kessler, J. R. Kling, and L. Sanbomatsu, “Neighborhood Effects on the Long-Term Well-Being of Low-Income Adults,” *Science* 337 (September 21, 2012): 1505–10, online edition。

13. 关于小微金融的影响的研究综述可参考M. Duvendack, R. Palmer-Jones, J. G. Copestake, L. Hooper, Y. Loke, and N. Rao, “What Is the Evidence of the Impact of Microfinance on the Well-Being of Poor People?” (London: EPPICentre, Social Science Research Unit, Institute of Education, University of London, 2011)。

第9章 如何解决组织中的时间稀缺

1. 可参考S. Crute, “Case Study: Flow Management at St. John's Regional Health Center,” *Quality Matters* (2005)。也可参考“Improving Surgical Flow at St. John's Regional Health Center: A Leap of Faith,” Institute for Healthcare Improvement。这一案例的研究也可见于E. Litvak, M. C. Long, B. Prenney, K. K. Fuda, O. Levzion-Korach, and P. McGlinchey, Improving Patient Flow and Throughput in California Hospitals Operating Room Services, Boston University Program for Management of Variability in Health Care Delivery. Guidance document prepared for the California Healthcare Foundation (CHCF), 2007。

2. 圣约翰医疗中心并非个例，详细的案例分析可参考Mark van Houdenhoven et al., Improving Operating Room Efficiency by Applying Bin-Packing and Portfolio Techniques to Surgical Case Scheduling, *Anesthesia and Analgesia* 105, no.3 (2007): 707–14。关于改善医院的病床管理的文献可参考Brecht Cardoen, Erik Demeulemeester, and Jeroen Belien, Operating Room Planning and Scheduling: A Literature Review, *European Journal of Operational Research* 201, no. 3 (2010):921–32。

3. 可参考John Gribbin, *Deep Simplicity: Bringing Order to Chaos and Complexity* (New York: Random House, 2005)。

4. 汤姆·德马可（Tom DeMarco）对组织的余闲的重要性有精彩的论述，“不优化组织也有可能让组织变得更有效率，方法就是让组织拥有余闲。也有可能让组织在效率上降低一点，但是极大地改善组织。为

了实现这一目的，你需要重新引入足够的余闲，让组织喘口气、再造自身，做出必要的改变”。见Tom DeMarco, *Slack: Getting Past Burnout, Busywork, and the Myth of Total Efficiency* (New York: Broadway, 2002)。

5. 关于融资收购，可参考Steven N. Kaplan and Per Stromberg, *Leveraged Buyouts and Private Equity*, *Journal of Economic Perspectives* 23, no.1 (Winter 2009): 121–46。

6. F. R. Lichtenberg and D. Siegel, The Effects of Leveraged Buyouts on Productivity and Related Aspects of Firm Behavior, *Journal of Financial Economics* 27, no.1 (1990): 165–94.

7. 当经济冲击来临时，关于融资并购导致企业进入危险境地的可能性已经被广泛讨论了，比如参考Krishna G. Palepu, “Consequences of Leveraged Buyouts,” *Journal of Financial Economics* 27, no. 1 (1990): 247–62。

8. Arthur G. Stephenson et al., “Mars Climate Orbiter Mishap Investigation Board Phase I Report, 44 pp.,” NASA, Washington, D.C. (1999).另一份可读的文献是James Oberg, “Why the Mars Probe Went Off Course,” *IEEE Spectrum* 36, no. 12 (1999): 34–39。

9. Roger E. Bohn and Ramchandran Jaikumar, *Firefighting by Knowledge Workers* (Information Storage Industry Center, Graduate School of International Relations and Pacific Studies, University of California, 2000)。

10. N. P. Repenning, “Reducing Cycle Time at Ford Electronics, Part II: Improving Product Development”。

11. 这一数据来自Bohn and Jaikumar, *Firefighting by Knowledge Workers*。但与这一数据矛盾的是，有数据表明微软发布的软件当时存在61 000个已知的漏洞，可参考<http://www.computergripes.com/Windows2000.html#28000bugs>。

12. 最近有研究论证，法官是怎样从试图同时接手很多案件到最后却一团乱麻的。可参考Decio Coviello, Andrea Ichino, and Nicola Persico, “Don’t Spread Yourself Too Thin: The Impact of Task Juggling on

Workers' Speed of Job Completion" (National Bureau of Economic Research Working Paper No. 16502, 2010)。

[13.](#) H. D. Thoreau, *A Week on the Concord and Merrimac Rivers*.

[14.](#) 相关法律见http://www.ghsa.org/html/stateinfo/laws/cellphone_laws.html。

[15.](#) 这些事实可见于edgarsnyder.com, 2012年11月2日的报告, 见<http://www.edgarsnyder.com/car-accident/cell-phone/cell-phone-statistics.html>。

[16.](#) 见J. Wilson, M. Fang, S. Wiggins, and P. Cooper, "Collision and Violation Involvement of Drivers Who Use Cellular Telephones," *Traffic Injury Prevention* 4, no.1 (2003):45–52.

[17.](#) D. L. Strayer, F. A. Drews, and D. J. Crouch, "A Comparison of the Cell Phone Driver and the Drunk Driver," *Human Factors: The Journal of the Human Factors and Ergonomics Society* 48, no.2 (2006): 381–91.追踪调查已经利用高保真的驾驶模拟器, 去对比打电话驾车的司机和醉驾的司机, 得出的结论是, 开车时打电话增加的风险, 比酒驾的风险还要大。

[18.](#) 可参考E. Robinson, "Why Crunch Mode Doesn't Work: 6 Lessons," *IGDA* (2005), retrieved February 17, 2009.还有Sara Robinson, "Bring Back the 40-Hour Work Week," *Slate*, March 14, 2012。两份文献都有清晰的立场, 比如减少工作日, 工人会受到激励, 并且把自己的工作做得很好。

[19.](#) Robinson, "Why Crunch Mode Doesn't Work."

[20.](#) "Scrum & Overtime?", 2008年7月9日发在博客*Agile Game Development*上。

[21.](#) Diwas S. Kc and Christian Terwiesch, "Impact of Workload on Service Time and Patient Safety: An Econometric Analysis of Hospital Operations," *Management Science* 55, no. 9 (2009): 1486–98.

[22.](#) Seonaidh McDonald, "Innovation, Organisational Learning and

Models of Slack,” *Proceedings of the 5th Organizational Learning and Knowledge Conference* (Lancaster University, 2003).

[23.](#) D. T. Wagner, C. M. Barnes, V. K. Lim, and D. L. Ferris, “Lost Sleep and Cyberloafing: Evidence from the Laboratory and a Daylight Saving Time Quasi-Experiment,” *Journal of Applied Psychology* 97, no.5 (2012): 1068.

[24.](#) 同上条。

[25.](#) “Manage Your Energy, Not Your Time,” *Harvard Business Review*, November 3, 2012.

[26.](#) 同上条。

[27.](#) 这就是所谓的“20-20-20法则”，比如可参考
<http://www.mayoclinic.com/health/eyestrain/DS01084/DSECTION=preventic>

[28.](#) J. De Graaf, D. Wann, and T. H. Naylor, *Affluenza: The All-Consuming Epidemic* (San Francisco, Calif.: Berrett-Koehler, 2005).

[29.](#) 关于“红花”日本料理的简要介绍可见
<http://www.rockyhaoki.com/biography.html>。

[30.](#) 这一讨论来自HBS关于红花日本料理的商业模型的精彩的案例分析，见W. Earl Sasser and J. Klug, *Benihana of Tokyo* (Boston: Harvard Business School, 1972)。也可参考Ernst Ricardo and Glen M. Schmidt, “Benihana: A New Look at an Old Classic,” *Operations Management Review* 1 (2005): 5–28.

[31.](#) S. E. Kimes, “Restaurant Revenue Management Implementation at Chevys Arrowhead,” *Cornell Hotel and Restaurant Administration Quarterly* 45, no.1 (2004): 52–67.

[32.](#) Y. Berra, *The Yogi Book* (New York: Workman Publishing, 1997).

第10章 如何应对日常生活中的稀缺

[1.](#) D. Karlan, M. McConnell, S. Mullainathan, and J. Zinman, *Getting to*

the Top of Mind: How Reminders Increase Saving (National Bureau of Economic Research, Working paper No. w16205, 2010).

2. “Impulse Savings,” ideas42 case study.
3. *Snopes.com*: Massachusetts License Renewal (November 4, 2008).
4. J. J. Choi, D. Laibson, B. C. Madrian, and A. Metrick, “For Better or for Worse: Default Effects and 401(k) Savings Behavior,” in *Perspectives on the Economics of Aging*, ed. D. A. Wise (Chicago: University of Chicago Press, 2004), 81–26.
5. http://www.bankofamerica.com/promos/jump/ktc_coinjar/.
6. Bank of America's Keep the Change program: “Keep Your Savings Elsewhere,” *BloggingStocks*, November 1, 2012.
7. L. E. Willis, “Against Financial Literacy Education,” (2008).
8. R. H. Thaler and S. Benartzi, “Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving,” *Journal of Political Economy* 112, no. S1 (2004): S164–87.
9. M. Bertrand and A. Morse, “Information Disclosure, Cognitive Biases, and Payday Borrowing,” *The Journal of Finance* 66, no. 6 (2011): 1865–93.
10. R. Levine, *A Geography of Time: The Temporal Misadventures of a Social Psychologist, or How Every Culture Keeps Time Just a Little Bit Differently* (New York: Basic Books, 1997).
11. J. Mata, P. M. Todd, and S. Lippke, “When Weight Management Lasts. Lower Perceived Rule Complexity Increases Adherence,” *Appetite* 54, no. 1 (2010): 37–43. doi:10.1016/j.appet.2009.09.004.
12. E. Duflo, M. Kremer, and J. Robinson, *Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya* (No. w15131, National Bureau of Economic Research, 2009).

[13.](#) 同上条。

[14.](#) 研究者在双曲线贴现率模型里来解释这一现象，作为对推迟满足这一普遍挑战的解决方案。我们收集到的数据，即农民在丰收时带宽增加，可以更好地解释这一现象，这些数据表明当农民的带宽到达最大值时，可以改善决策质量。

[15.](#) K. Haycock, “Promise Abandoned: How Policy Choices and Institutional Practices Restrict College Opportunities” (Washington, D.C.: Education Trust, 2006).

[16.](#) E. P. Bettinger, B. T. Long, P. Oreopoulos, and L. Sanbonmatsu, *The Role of Simplification and Information in College Decisions: Results from the H&R Block FAFSA Experiment*. National Bureau of Economic Research, (2009).

[17.](#) D. Ariely and K. Wertenbroch, “Procrastination, Deadlines, and Performance: Self-Control by Precommitment.” *Psychological Science* 13, no. 3 (2002): 219–24.

[18.](#) C. F. Camerer, and H. Kunreuther, “Decision Processes for Low Probability Events: Policy Implications,” *Journal of Policy Analysis and Management* 8, no.4 (1989): 565–92.

结语

[1.](#) J. Horgan, “The New Challenges,” *Scientific American* 267, no. 6 (1992): 10.

[2.](#) 这个想法是一个非营利组织ideas42在探索的，我们帮助创立了这个组织，它旨在利用行为研究获得的洞见去创造产品或政策。感兴趣的读者可以访问网站www.ideas42.com。

[3.](#) *Vitality-About GlowCaps*. Retrieved from <http://www.vitality.net/glowcaps.html>.

译者后记

稀缺和理性经济人是传统经济学的两大基本假设：资源是稀缺而有限的；人具有完全的理性，可以做出利益最大化的选择。

有意思的是，如果把本书的观点总结为一句话，正是对这两大基本假设的一次重新审视——过度的稀缺会导致人的非理性，从而进一步加剧这种稀缺。再深究下去，这是不是意味着传统经济学两大假设之间存在着明显的冲突呢？

人的“有限理性”理论早在20世纪兴起的行为经济学中就已出现，并在社会生活中得到了应用。而《稀缺》一书的许多观点和案例也可以看作行为经济学的一种应用：缺钱（或时间）会对人的心理产生影响，让人做出错误的决策，从而变得更缺钱（或时间）。既然我们知道这个规律，何不针对性地设计制度，来防止人们继续犯错呢？例如，将“每个人一生中只能享受5年福利救济”的政策改为“每个人一生中只能享受6年救济，其中每年里能享受8个月”。再如，从手术室资源严重不足的医院里再留出一间手术室，不安排计划内手术。站在完全理性的假设角度，这些举措都是低效且不合理的，但如果将稀缺对人的心理影响考虑在内，这些做法却能取得更好的成效。

不仅如此，本书的观点和案例还能改变人们对“穷人”的一些偏差。穷人之所以会长期贫困，不一定是由于智力或努力程度的问题，而有可能是稀缺造成的恶性循环。这在当前强调消除贫富差距、“走共同富裕”的中国，也有非常积极的社会意义。

最后，还要感谢康洁、王晓明、张晓英、张萌、张翰、魏宁、李维、赵晓光、魏雯雯和王朝阳对我们的帮助。对于译者而言，时间与金钱的稀缺状态一时难以改变，幸而还有富足的友谊，一路支撑着我们完成了本书的翻译工作。

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

而真正的阅读，应该在书中与智者同行思考，借他们的视角看到世界的多元性，提出比答案更重要的好问题，在不确定的时代中领先起跑。

湛庐阅读 App：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本

怎么读是读者面临的最大阅读障碍

“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”App。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、电子书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。



下载湛庐阅读 App，
坚持亲自阅读，
有声书、电子书、阅读服务，
一站获得。

CHEERS

本书阅读资料包

给你便捷、高效、全面的阅读体验

本书参考资料

湛庐独家策划

- ☑ 参考文献
为了环保、节约纸张，部分图书的参考文献以电子版方式提供
- ☑ 主题书单
编辑精心推荐的延伸阅读书单，助你开启主题式阅读
- ☑ 图片资料
提供部分图片的高清彩色原版大图，方便保存和分享

相关阅读服务

终身学习者必备

- ☑ 电子书
便捷、高效，方便检索，易于携带，随时更新
- ☑ 有声书
保护视力，随时随地，有温度、有情感地听本书
- ☑ 精读班
2~4周，最懂这本书的人带你读完、读懂、读透这本好书
- ☑ 课 程
课程权威专家给你开书单，带你快速浏览一个领域的知识概貌
- ☑ 讲 书
30分钟，大咖给你讲本书，让你挑书不费劲



湛庐编辑为你独家呈现
助你更好获得书里和书外的思想和智慧，请扫码查收！

(阅读资料包的内容因书而异，最终以湛庐阅读App页面为准)

(1)丹尼尔·卡尼曼是诺贝尔经济学奖获得者，其著作《噪声》中文简体字版已由湛庐引进、浙江教育出版社于2021年出版。——编者注

(2)沃尔特·米歇尔是著名的人格心理学家，其著作《棉花糖实验》中文简体字版已由湛庐策划、北京联合出版公司于2016年出版。——编者注

(3)曝光表，亦叫测光表。摄影时计量光的强弱并借以控制摄影曝光时间和光圈大小的仪表。——编者注

CHEERS
湛庐

影响力

INFLUENCE, NEW
AND EXPANDED

更新多达10万字
增加5年来近50个
新商业案例
升级6大武器为7种武器

Robert B. Cialdini
[美]罗伯特·西奥迪尼 著

全新升级版

风靡全球 35 载，畅销
数百万册的“影响力法则”

The Psychology
of Persuasion

闫佳 译



北京联合出版公司
Beijing United Publishing Co., Ltd.

版权信息

本书纸版由北京联合出版公司于2021年11月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

书名：影响力（全新升级版）

著者：罗伯特·西奥迪尼

电子书定价：89.99元

INFLUENCE: NEW AND EXPANDED by Robert Cialdini.

Copyright ©1984, 1994, 2007, 2021 by Robert Cialdini.

Published by arrangement with Harper Business, an imprint of HarperCollins Publishers.

All rights reserved.



INFLUENCE NEW AND EXPANDED 各方赞誉

彭凯平 | 清华大学社会科学学院院长、心理学系主任

《影响力》绝对可以称得上是社会学心理学领域的经典之作，它不仅阐释了影响他人行为的科学原理，还提炼出了人人都能用、而且好用的提升影响力的方法，难怪全世界的读者都会为其买单。我很高兴西奥迪尼教授在这本书出版 30 多年后依然传来了这样重要而有价值的更新，《影响力（全新升级版）》必定会成为心理学领域的一部名著，让更多人从中受益。

刘润 | 润米咨询董事长、微软前战略合作总监

管理者都希望自己拥有影响力，影响力究竟来自哪里？从管理学角度来说，影响力来自对各方利益诉求的把握和协调。西奥迪尼教授则从心理学角度解读影响力，他认为影响力来自建立在敏捷反应和胆识对比决策基础之上的 7 种“影响力武器”，掌握了这 7 大武器，管理者将获得一种与众不同的视角，拥有掌握全局的能力。

金鹏远 | 环时互动首席知识官

换历了社交网络在中国迅猛发展的这 10 多年，最早一版的《影响力》也被我翻遍了封皮。如果说环时互动取得了一些小小社交网络传播的成就，那这本书就是我们的知识启蒙的老师之一。

熊浩 | 复旦大学法学院副教授、谈判专家、《我是演说家》双季冠军

很多人有误解，认为学习说服、增进影响力是在学习表达，而我说不然。在我看来，影响力的本质是去构建对人的基本思维方式与行为模式的科学理解与准确预言。正是在这个意义上，《影响力（全新升级版）》不是一本关于表达的书，而是一本驾驭商业、理解销售、了解传播，乃至穿透人性的书。相信读完这本书，你一定会认同我所说的话。

张建新 | 中国社会心理学会会长、中国科学院心理研究所研究员

此前，西奥迪尼对 6 大武器的娴熟使用，令《影响力》的影响力强悍。如今，《影响力（全新升级版）》特别增加了联盟原则，将此功内底升级为 7 大武器。制此武器被发明者列为机密，他却将之公开。这等气魄或正是使其影响力倍增的另类武器。但“人后影响”深究文化和国情的特约，7 大武器能否在中国成为法宝，尚需读者熟读此书后用实际行动予以回答。

雷富 | 金蝶奖年度 CMO

《影响力》向来是市场营销专业人士的必修书，也是影响我职业生涯最深刻的书之一。不管是在市场中影响消费者，还是在谈判中影响客户，我都建议你认真地读这本书。《影响力》是真正把社会学、心理学和市场营销学紧密结合的著作，是对市场营销学的一次里程碑式的提升。

吴晨 | 《经济学家·商论》执行总编辑

这是一本深刻剖析人性的书，希望能助你用理性来规避感性思维可能被人利用的危险。礼尚往来和小恩小惠在西方语境中的解读，和“天下没有免费的午餐”一样，给人警醒；以貌取人和墨屋及乌，是人性中常见的弱点，中外概然。以退为进可能是求人办事的优秀策略，概不掩瑜可能是增强权威的小伎俩，打肿脸充胖子可能带来同知上当受骗还得死撑面子活受罪的结果……当然，这也是一本矛盾的书，一方面希望每个人都能顺势而为，善用他人的感性思维，施加正向的影响力；一方面又提醒每个人，在社交媒体圈化的当下，商家可能运用各种制造影响力的手段。

成甲 | 《好好学习》《好好思考》作者

为什么历史总是惊人地相似？因为人性深处有一些特性是穿越时空也不变的。这也是《影响力》在出版 30 多年后仍然有生命力的原因。西奥迪尼先生在《影响力（全新升级版）》中加入了他人性的新理解，也增加了一个重磅影响力武器——联盟。在当今环境下，如何持续影响他人变得越来越困难，而带来的精神意义层面的影响也就越来越重要。企业与组织都可以从联盟这一武器中找到提升自身使命、愿景与价值观的有效方法。

丹尼尔·卡尼曼 | 诺贝尔经济学奖得主、《噪声》作者

西奥迪尼博士完成了一件几乎不可能完成的事：他把一件杰作两次推向了新的高度。《影响力（全新升级版）》对人们相互施加的微妙影响力进行了极为丰富和引人入胜的叙述。

B.J. 福格 | 斯坦福大学行为设计实验室创始人、《习惯行为模型》作者

《影响力（全新升级版）》相比此前的版本，其影响力更加实际且强大。仅靠联盟这一原则本身就能足以改变商业和生活中的游戏规则。实在是太好了！

多利·克拉克 | 美国杜克大学富卡商学院兼职教授、《深潜》作者

《影响力》的过往版本已经是现代商业经典，深刻影响了市场营销和心理学领域。全新升级版包含了大量强有力的新见解和例子，让这本精彩的书变得更加耀眼。

凯蒂·米尔克曼 | 沃顿商学院教授

如果能为你推荐一本对工作与生活都有帮助的书，我会选《影响力（全新升级版）》。这是一部杰作，西奥迪尼这次将它变得更加精彩了。

乔·波兰 | 知名企业家网站 Genius Network 创始人

《影响力（全新升级版）》是一部杰作，它既有趣又值得一读。

贝奇·莉薇·帕鲁克 | 普林斯顿大学卡尼曼-特里斯坦行为科学和公共政策中心副主任

和我认识的所有心理学家一样，和成千上万好奇世界是如何运作的人一样，我是从《影响力》开始学习说服技巧的。这本全新升级版在我桌子上那本老旧的经典版的基础上进行了如此有意义的构建，它将在未来几年里继续帮助和启发发展的艺术和科学。

盖伊·川崎 | 苹果公司前首席营销官、《纽约时报》畅销书作家

这本书给每个想要改变世界的人创造了绝佳的机会，全新升级版更是如此。

汤姆·彼得斯 | 知名管理学大师、“商界教皇”

如果在谈判中碰上了棘手的大麻烦，我一定会找罗伯特·西奥迪尼咨询。在这方面没有人比他更合适了。

理查德·塞勒 | 诺贝尔经济学奖得主

在这部既精又简的新作中，世界上最优秀的社会心理学家西奥迪尼博士在分享其智慧的同时，也展示了十足的个人魅力，你将在本书中学到的东西无异于一粒重磅炸弹，请小心使用它！

蒂姆·哈福德 | 牛津大学经济学教授、畅销书《试错力》作者

《影响力（全新升级版）》作为这一主题的权威著作实至名归。我从中学到了很多，相信你也一样。

亚当·格兰特 | 知名心理学家、沃顿商学院终身教授

这是有史以来关于说服科学最重要的书，而且它正随着作者的不间断代而日臻完美。我想不出比这更有趣、更实用的读物了。

杰弗里·普费弗 | 斯坦福大学商学院组织行为学教授

这是一部非凡的智慧和成就！《影响力》作为一部关于人类行为基本原则的精彩著作，这次恰逢其时地加入了新的原则——联盟。

巴里·里索尔兹 | 里索尔兹财富管理公司董事长兼首席投资官

这是心理学和行为经济学的经典著作，对任何想详细了解决策过程的人来说，《影响力（全新升级版）》都是一本必读书。

安杰拉·达克沃思 | 知名教育网站 Character Lab 创始人兼首席执行官

如果你只读一本书，只信任一个专家的建议，那就应该是《影响力（全新升级版）》和西奥迪尼。

安妮·杜克 | 认知心理学博士、知名风险决策专家

西奥迪尼是构建复杂的科学有趣、易于消化的形式表达的先驱，每个人都可以理解并从中受益。在全新升级版中，作者用前沿、尖端的研究报告和最新案例更新了原本那部强大的一本书，巧妙地吸引读者。《影响力》一直是必读，现在更是如此。

劳伦斯·霍夫 | 艾德凡塔公司关系咨询副总裁

西奥迪尼博士把本书的素材直接跟我们反对客户的具体细节联系起来，让我们得以做出重大的改变。他的作品给我们带来了突出的竞争优势。

查理·芒格 | 伯克希尔·哈撒韦公司副总裁

《影响力》能帮助管理者做出更好的决策和明智地使用自身的影响力。西奥迪尼在这个话题上对性的洞察比其他任何科学家都大——这本杰出的畅销书揭示了 7 大影响力原则，在这些原则中，你无意识的想法往往能和你自己的最佳利益相吻合。

丹尼尔·平克 | 畅销书《驱动力》《全新销售》《全新思维》作者

西奥迪尼让这部经典之作更精彩了。《影响力（全新升级版）》再次捍卫了它作为过去 30 多年来最重要的商业和行为科学书籍之一的地位。新增加的内容太棒了。

乔纳·伯杰 | 沃顿商学院市场营销学教授

做好为之准备的准备吧！西奥迪尼博士被誉为“影响力教父”，这本书的最初版本已是经典之作，如今的《影响力（全新升级版）》又带来了更大的惊喜。无论你是试图影响他人，还是想更理解他人如何影响你，这本书都会给你切实的实践指南。

丹尼尔·夏皮罗 | 哈佛大学国际项目创始人

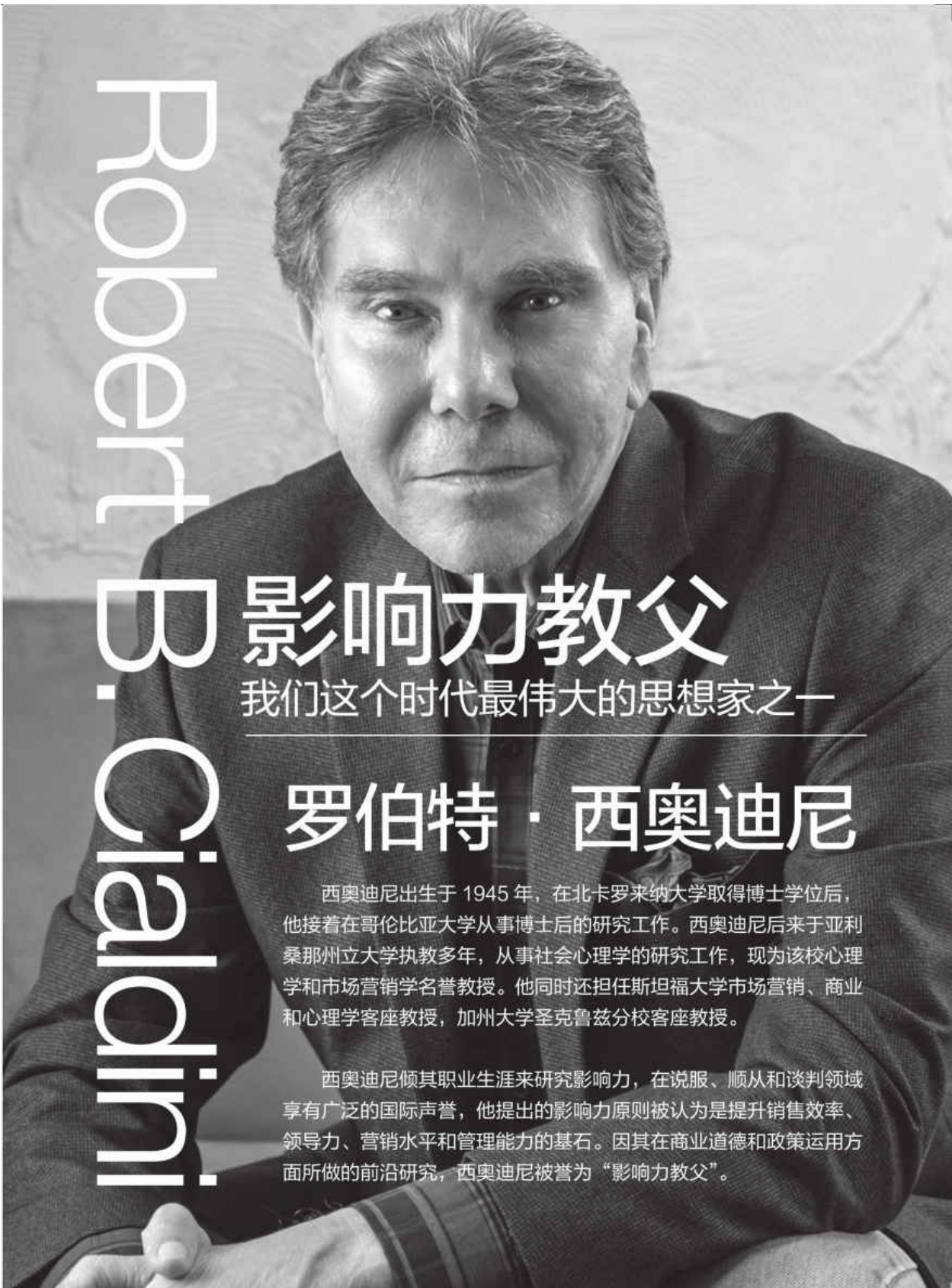
一本非凡的书！无论你是想提高销量、达成一笔更好的交易，还是想改善人际关系，《影响力（全新升级版）》都提供了经过科学验证的原则，它足以改变你的生活。

罗伯特·萨顿 | 斯坦福大学管理学教授

过去 25 年里，我在斯坦福大学组织行为学专业给学生布置过的唯一一次阅读任务就是阅读这本书。学生们很喜欢它，几年后还在赞赏《影响力》给他们整个职业生涯的帮助。全新升级版更有趣、更幽默，读起来也更有趣。

克里斯·沃斯 | FBI 前国际危机谈判专家、世界 500 强谈判咨询公司“黑天鹅集团”创始人

任何想让自己的沟通或谈判能力达到最高水平的人都必须阅读这本书。没有它，你的知识库就是不完整的。

A black and white portrait of Robert B. Cialdini, a middle-aged man with short, slightly graying hair, wearing a dark suit jacket over a light-colored shirt and a patterned tie. He is looking directly at the camera with a slight smile. The background is a textured, light-colored wall.

Robert B. Cialdini

影响力教父

我们这个时代最伟大的思想家之一

罗伯特·西奥迪尼

西奥迪尼出生于 1945 年，在北卡罗来纳大学取得博士学位后，他接着在哥伦比亚大学从事博士后的研究工作。西奥迪尼后来于亚利桑那州立大学执教多年，从事社会心理学的研究工作，现为该校心理学和市场营销学名誉教授。他同时还担任斯坦福大学市场营销、商业和心理学客座教授，加州大学圣克鲁兹分校客座教授。

西奥迪尼倾其职业生涯来研究影响力，在说服、顺从和谈判领域享有广泛的国际声誉，他提出的影响力原则被认为是提升销售效率、领导力、营销水平和管理能力的基石。因其在商业道德和政策运用方面所做的前沿研究，西奥迪尼被誉为“影响力教父”。



影响力科学的奠基人



自《影响力》1984年首次出版以来，西奥迪尼就一直是营销领域最好的心理学家。

——美国心理协会

西奥迪尼曾将他对社会影响的兴趣归因于其多元文化碰撞的成长背景：他在一个完全是意大利人的家庭中长大，生活在一个主要是波兰人的社区，身处以大量德国移民著称的密尔沃基，属于原本是地道乡村的威斯康星州。正因如此，西奥迪尼对影响力科学的兴趣与研究几乎贯穿了他的整个职业生涯。

《影响力》作为西奥迪尼30多年研究成果的结晶，自1984年出版以来，三度登上《纽约时报》畅销榜，被译为40多种文字，累计销量突破500万册，受到全球读者的喜爱，多年来雄踞亚马逊网站消费者行为学类图书排行榜首位。

《影响力》是心理学和管理学专业学生的必读书，是《财富》杂志

鼎力推荐的“75本商业必读书”之一、是亚马逊推荐的“人生必读100本书”之一，也是巴菲特和芒格向他们的股东推荐的唯一一本书。畅销书作家丹尼尔·平克称赞道：“任何关于说服和影响力的写作都是站在西奥迪尼肩膀上的。”

这部开创性著作最初是基于西奥迪尼三年的“卧底生涯”创作而成的。他通过参加电话营销公司、二手车经销商、筹款组织等机构的培训项目，观察、记录、营销、销售等行业说服他人的技术，结合自己对社会心理学的潜心研究，先后总结出说服他人的6大原则，即互惠、喜好、社会认同、权威、稀缺、承诺与一致。

《影响力》改变了很多人和组织，被引述率高居当今社会心理学之冠。然而，西奥迪尼并未止步于既有成就，而是在对社会影响力的不断研究和实践中，不断迭代《影响力》。他推出了更适合高校作为教材使用的教材版，该书被斯坦福大学列为经典权威教材；又出版了着重在人们做出决定之前提前对其施加潜在影响的《先发影响力》。经过近10年的研究，西奥迪尼又为影响力增加了“终极武器”——联盟原则。

如今，影响力的7大原则首次集结于《影响力（全新升级版）》一书，该书相比此前版本，更新多达10万字，新增了“线上影响力”专栏，指出互联网的蓬勃发展对人们生活的深刻影响，并收录了最新研究、案例和读者报告。西奥迪尼还根据不同的动机，列出了使用影响力武器的不同组合与先后顺序。

西奥迪尼理念的影响范围包括政界、商界、学界、文艺界甚至军界，巴菲特、芒格、丹尼尔·卡尼曼、福格等各界翘楚都曾在“影响力”系列作品中得到不少启发。正如斯坦福大学商学院组织行为学教授奇普·希思所说：“西奥迪尼就是影响力研究领域的本杰明·富兰克林。”

在全球践行影响力

西奥迪尼对影响力的研究并不限于教学和对大众的科学普及，他还专门成立了职场影响力咨询公司（INFLUENCE AT WORK, IAW），带领企业从事影响力的研究与培训工作，提高组织和个人的绩效和道德影响力，旨在帮助管理人员更好地决策，巧妙地使用他们的影响力。

西奥迪尼的客户除了谷歌、微软、思科公司、可口可乐、爱立信、IBM、葛兰素史克、奥美等世界500强企业，还包括哈佛大学肯尼迪政治学院、美国司法部等教育、政府部门。



2012年，西奥迪尼作为行为科学团队的一员，帮助奥巴马在当年的大选中获胜。西奥迪尼说，无论是政治领域还是职场，只要学会在不同的场合灵活使用不同的影响力策略，你便能成功赢得他人的支持。

因在社会心理学、消费者心理学以及社会影响力方面所做的贡献，西奥迪尼赢得了众多赞誉和奖项：

- 1996年至1997年担任美国人格与社会心理学会主席。
- 2000年获得消费者心理学会颁发的杰出科学成就奖。
- 2003年获得“唐纳德·坎贝尔奖”和首届“皮托奖”。

- 2008年获得人格与社会心理学会颁发的杰出服务奖。
- 2009年获得实验社会心理学会颁发的科学家奖，以及西部心理学会颁发的杰出教学奖。
- 2016年获得西部心理学会终身成就奖。
- 2019年亚利桑那州立大学为纪念西奥迪尼在行为研究领域所做的杰出贡献，特将原营销实验室（Marketing Lab）更名为西奥迪尼行为研究实验室（Robert B. Cialdini Behavioral Research Lab）。

西奥迪尼以科学为依据，将说服打造为多面体，告诉人们什么是影响力以及如何根据不同的情况，在不违背道德的情况下，有技巧地使用影响力武器。作为说服领域的权威，西奥迪尼以人类的基本本能为基础，帮助人们更好地掌握了影响他人的方法。

作者演讲洽谈，请联系

speech@cheerspublishing.com

更多相关资讯，请关注



湛庐文化微信订阅号

湛庐CHEERS

特别
制作

西奥迪尼

“影响力”系列作品



《影响力》的最初版本是想写给大众读者的，所以在写它的时候，我试着像讲故事那样娓娓道来。我承认，其实有些担心这么做会让学术界的同事把这本书视为“通俗”心理学。正如法律学者詹姆斯·博伊尔（James Boyle）所说，“你如果没听过学者们是怎么说‘大众读物’（popularizer）这个词的，就不算见识过真正的傲慢”。而这也正是我所担心的。出于这个原因，在我最初撰写《影响力》的时候，大多数社会心理学同行都觉得，从职业发展的角度而言，为非学术界的通俗读者写作不怎么安全。事实上，如果说社会心理学是一家公司，那它大概会以卓越的研发部门出名，配送部门则不太拿得出手。除了偶尔在学术期刊上发表一些普通读者不乐意看的文章，其他我们概不配送。

幸运的是，尽管决定采用通俗风格，但我担心的事情无一应验，因为《影响力》并未因立足于“通俗”心理学而受到贬低。^①于是，之后的各版保留了这种叙述风格。除此之外，我还为自己的陈述、建议和结论加入了经研究得来的证据。虽说我用了采访、引述、系统化的个人观察等方式来表达结论、巩固结论，但归根结底，本书的结论是建立在严格控制变量的心理学研究基础之上的。

对《影响力（全新升级版）》的评价

怎样改动全新升级版，让我颇觉棘手。一方面，本着“没坏就不用修”的公理，我不愿意做大规模的改动。毕竟，之前的若干版本，再加上41种外语版，《影响力》卖出的册数比我想象中要多得多。从这一点来说，我的波兰同事威廉明娜·沃辛斯卡（Wilhemina Wosinska）对本书的价值做出了肯定性的评论，同时也给我浇了一瓢冷水以使我保持清醒。她说：“罗伯特，你知道吗，《影响力》在波兰可出名了，我的学生都以为你早就死了。”

另一方面，我那来自西西里的祖父爱说这样一句话：“如果你想让事情保持现状，那事情就必须改变。”进行及时的修订工作，恰如此言所说。^②距《影响力》上一个版本的出版已经有些日子了。在此期间发生的一些事情，有必要补充到全新升级版中。首先，我们现在对影响过程比从前有了更深入的认识。对说服、顺从和改变的研究有了新进展，相关内容亦依此做了更新。除了对整体素材进行全面更新，我还特别注意增加了影响力在日常人类互动中的作用，即施加影响的过程怎样在现实世界环境而非实验室环境中运转。

我还扩展了一个专栏，这是先前读者的反馈让我想到的。这个专栏的内容主要来自之前读了《影响力》的读者。这些读者意识到某个原则在特定的情况下是怎样对自己发挥作用或为自己所用的，并写信告诉了我。他们的描述（即每章的“读者报告”专栏）说明，在日常生活里，我们是多么容易又多么频繁地成为影响过程的“受害者”。所以，如何将本书的原则应用于商业和个人生活，我们现在有了许多更新的第一手资料。

我想要感谢以下为《影响力》此前的“读者报告”专栏做出过贡献的诸君，他们要么是亲自来信贡献了案例，要么通过自己的课程导师提交了例子，他们是：帕特·鲍勃斯（Pat Bobbs）、哈特纳·比伯克（Hartnut Bock）、安妮·卡托（Annie Carto）、迈克尔·克洛伊（Michael Conroy）、威廉·库珀（William Cooper）、艾丽西亚·弗里德曼（Alicia Friedman）、威廉·格拉齐亚诺（William Graziano）、马克·哈斯廷斯（Mark Hastings）、乔纳森·哈里斯（Jonathan Harries）、恩戴耶·肯迪（Endayehu Kendie）、凯伦·克罗尔（Karen Klawer）、丹努塔·鲁尼卡（Danuta Lubnicka）、詹姆斯·迈克尔斯（James Michaels）、斯蒂芬·莫塞（Steven Moysey）、凯蒂·缪勒（Katie Mueller）、保罗·内尔（Paul Nail）、丹·诺里斯（Dan Norris）、山姆·奥玛（Sam Omar）、艾伦·雷斯尼克（Alan Resnik）、达里尔·雷兹拉夫（Daryl Retzlaff）、杰弗里·罗森博格（Geofrey Rosenberger）、乔安娜·斯派查拉（Joanna

Spychala)、丹·斯威夫特(Dan Swift)、罗伯特·斯陶瑟(Robert Stauth)和卡拉·瓦斯科(Karla Vasks)。也特别感谢为全新升级版贡献了全新“读者报告”的各位:劳拉·克拉克(Laura Clark)、杰克·埃普斯(Jake Epps)、胡安·格梅兹(Juan Gomez)、菲利普·约翰斯通(Phillip Johnston)、宝拉(Paola)、乔·约翰(Joe St. John)、卡罗尔·托马斯(Carol Thomas)、杨斯·特拉波特(Jens Trabolt)、卢卡斯·维曼(Lucas Weimann)、安娜·罗布鲁斯基(Anna Wroblewski)、安格玛·亚达夫(Agrima Yadav)。我还想邀请读者为未来将要出版的新版提供类似的“报告”,“报告”可发送至邮箱 ReadersReports@Influenceatwork.com。

本次更新除了对此前版本的一些更新与扩展,还有以下三个重大变动。

- 其一是探索社会影响力策略在互联网上的应用。很明显,社交媒体和电子商务网站已经投入了说服科学的怀抱中。因此,全新升级版每章新增了“线上影响力”专栏,说明怎样将相关手法应用于当前的技术平台。
- 其二是增加了一些新的脚注,方便读者找到正文所引用的研究及相关工作的介绍,对手头议题做更广泛的探讨。
- 其三,也是最重要的一点是,我在全新升级版增加了关于社会影响力的第7条普遍原则——联盟原则。在相关章节里,我介绍了这样一种情况,即如果受众相信沟通者跟自己有着共同且有意义的个人或社会认同,就更容易被沟通者的说服诉求所打动。

前言

我现在可以坦白了，我这一辈子，一直是个容易上当的家伙。在记忆所及的岁月里，我总是容易被小贩、筹款人、这样那样的运营商当成好捏的软柿子。是的，这些人里只有一部分动机不大光彩，其他人，比如慈善机构的代表，都有着崇高的目的。我发现自己老是会订些根本不想要的杂志，或是买下环卫工人舞会的门票，这种事情出现的频率之高，让我自己都感到吃惊。或许，我这种一贯的傻瓜蛋状态，正解释了我为什么会对研究说服感兴趣：到底是什么因素让一个人向另一个人说“行”？哪些技术最有效地利用了这些因素，使人们产生了顺从心理？我想搞清楚，为什么相同的请求，按某种方式说出来会遭到拒绝，稍微换种方式说结果却会不一样呢？

所以，我干上了实验社会心理学这一行，开始研究说服心理学。起初，我的研究大多以实验的形式开展，基本都在实验室里进行，被试则是大学生，我想找出哪些心理原则会对人们答应一个要求产生影响。现在，关于这些原则是什么、如何发挥作用，心理学家已经有了很多认识。我认为这些原则是影响力的武器，并将在本书重点讨论其中的一部分。

不过，过了一段时间，我逐渐意识到：做实验尽管有必要，但还不够。一旦走出心理学大楼，走出我研究它们的校园，我就没办法再靠实验来判断这些原则在真实世界里是有多重要。很明显，要想彻底了解顺从心理，我得放宽自己的调查范围。我需要去观察专门利用人们顺从心理的职业老手，也就是套用这些原则左右了我一辈子的那些人。优胜劣汰的生存法则让他们知道哪些原则可行，哪些不可行。他们的生计就是说服别人，他们是靠这一套吃饭的。他们当中，那些不知道怎样才能让别人顺从的人很快就会被淘汰掉，而知道的人则能留下来，并将买卖做得风生水起。

当然，了解并利用这些原则帮助自己的人，不光只有上述那些专业人士。在跟邻居、朋友、爱人或其他家人进行日常交往的过程中，我们或多或少都会用到它们，或者成为它们的受害者。但哪些原则最管用，我们大多数人最多只有些模糊的认识，以让他人顺从为业的说服专家们却懂得更多。我仔细想了想，要了解哪些顺从原则对我适用，这些专家是最丰富的信息源。于是，近年来，我把自己的实验研究跟一个更有趣的项目结合了起来。我系统地让自己深入说服专业人士，如销售员、筹款者、广告商的工作中。

我的目的是从内部观察诸多说服专业人士最常使用且效果最好的技术和策略。在这一观察项目实施的过程中，我有时会访问从业者本人，有时则会调访某些从业者的“天敌”，如警方的反诈骗调查人员、消费者权益保护机构等。还有些时候，我会对一代代传承说服策略的书面资料，如销售手册等进行大范围的检验。

我们最常采用的手段是参与式观察（participant observation）。参与式观察是一种由研究者充当各种间谍的调研方法。研究者会利用一个伪装的身份，捏造一套意图，渗入感兴趣的环境，成为所调查群体里的一员。所以，如果我想了解百科全书、真空吸尘器、肖像摄影、舞蹈课销售者的说服手段，我就会去应征报纸上招募销售实习生的广告，直接让他们教给我方法。使用类似但不尽相同的办法，我得以渗透到广告、公共关系和筹款机构内部去考察。因此，这本书里列举的许多证据，都来自我在各类致力于让你点头称是的组织里假冒说服专业人士或是怀有远大抱负的专业人士的亲身体验。

我通过这几年的参与观察，掌握了一些极具启发意义的信息。尽管说服专业人士会使用上千种不同的策略来让人顺从，但绝大部分策略都能分为7个基本类型。每一个类型都是从一种能指导人们行为的基本心理原则衍生出来的，正因为如此，这些策略才具有左右人们行为的力量。本书将围绕这7大原则来组织结构，一章介绍一种原则。我将分别

按照互惠、喜好、社会认同、权威、稀缺、承诺与一致、联盟的社会功能来展开讨论，看看说服专业人士们是如何将之整合到请求别人购买、捐赠、让步、选举或赞成的行为中，使之发挥出巨大力量的。

值得注意的是，我提出的7大原则中，并不包括物质利己主义的简单原则，即人们希望通过自己的选择得到更多，付出更少。我没有这么做，并不是因为我认为利益最大化、成本最小化的欲望在推动我们做决策方面无关紧要，也不是因为我找到了什么证据，说明说服专业人士忽视了物质利己主义的力量。恰恰相反，在调查中，我经常看到从业者使用“我能给你一笔划算的交易”这样的方法（有时他们说的是真心话，有时则不然）。在本书中，我不会对物质利己主义做单独的讨论，因为我认为这是一项给定的动机，一个不言而喻的事实，所以人人都理解它，无须深入描述。

本书所讨论的每一种原则，均能使人产生不同的自动、无意识的顺从行为，即一种不假思索就答应的冲动。有证据表明，现代社会日益加快的步伐和大量信息带来的冲击，会使人们的无意识顺从在将来变得越发普遍。因此，理解无意识影响是怎么回事，为什么会这样，对当今社会来说是非常重要的。

最后，在全新升级版中，我按照同事格雷戈里·奈德特（Gregory Neidert）博士的洞见，对相关章节的顺序进行了重新排列。他认为，基于交流者希望通过信息实现的说服目标，有一些原则比其他原则更为适合。诚然，任何想要施加影响力的人都希望促成他人的改变，但根据奈德特设计的社会影响力核心动机模型，沟通者在进行沟通时的首要目标，将影响他选用这7种原则的优先顺序。举例来说，该模型断言，劝导者的主要动机（目标）之一和培养积极的关系有关。研究表明，如果信息的接收方先对发送信息的人产生了好感，那么，后者就更有可能达成目标。而互惠、喜好和联盟这三条原则似乎尤其适用于这样的任务。

还有些时候，受众和沟通者之间的美好关系已经确立，那么，减少

不确定感就成为当务之急。毕竟，双方的积极关系并不能确保说服的成功。在可能改变主意之前，人们总是希望自己所做的决策是明智的。此时，奈德特的模型指出，绝不能忽视社会认同和权威原则——把证据指向同行或专家认可的选项，能明显提升这一选择在受众心目中的明智性。

但是，哪怕培养起了积极的关系，减少了不确定感，为提高说服效率，还有一个目标需要达成，即激发行动。一个朋友曾用充分的证据向我指出，按照专家的建议，每天锻炼是件好事，顶尖的医学专家们也都完全认同每天锻炼给健康带来的益处。但这恐怕还是不足以让我坚持每天锻炼。为了劝说成功，这位朋友还可以提醒我，我过去曾公开说过保持健康的重要性（承诺与一致原则），而要是我不锻炼的话，就无从享受锻炼带来的独到乐趣（稀缺原则）。这样的信息，才最有可能让我从仅仅是做出行动的决定，过渡到根据决定采取行动步骤。因此，也正是这样的信息，才最有可能让我大清早起床去锻炼身体。

因此，本书各章考虑到了哪些原则特别适合达成说服者的三种动机：

如果培养关系是当务之急，互惠、喜好和联盟最为适用；

如果减少不确定感最为重要，那么社会认同和权威原则最适用；

倘若激发行为是基本目标，那么用到的原则就是承诺与一致原则与稀缺原则。

读者朋友们有必要认识到，我并不是说这些互有关联的原则，是实现相应目标的唯一选择。相反，我想说的是，如果碰到了与这些原则所对应的目标却忘了运用这些原则，那将是你很大的损失。

你的影响力水平有多高？

- 你的公司刚刚推出一款新产品，老板让你决定采用以下哪种策略来激发公众最大的兴趣：在“有限时间”内降价，或者在“有限数量”内降价。那么你认为___。
 - A. 应选“有限时间”
 - B. 应选“有限数量”
 - C. 任何一种选择都会产生同样积极的结果
 - D. 这两种选择都不会产生积极的结果
- 假设有一位政治候选人刚刚失去民众的信任。不幸的是，你是这位候选人竞选班子的负责人。如果这位候选人欲借严厉打击犯罪重树他的声望，你认为在他开始下一站宣传时，以下哪个选项是最好的方式？
 - A. 我的对手在打击犯罪方面做得很不够.....
 - B. 很多民众支持我打击犯罪的意感，而且他们相信我有这个能.....
 - C. 虽然我的对手在打击犯罪方面有着不俗的表现.....
 - D. 打击犯罪是一个关键问题
- 最近你带领的部门人员流失情况严重，所以你组织了一次团建来鼓舞士气。你想给参加团建的每位员工送一份礼物，并由此强化员工回馈组织的承诺。以下哪种策略可能产生最好的

结果？

- A. 给他们同样的、昂贵的礼物，并刻上公司名
- B. 不要送礼物，但要感谢他们的参加
- C. 给每位员工一个有意义的个性化礼物，即使它不贵
- D. 只给完成相关评估表格的员工送礼物

扫码鉴别正版图书
获取您的专属福利



扫码获取全部测试题及答案
看一看影响力武器是如何
影响他人、识破套路的

目 录

[序](#)

[前言](#)

[测一测 你的影响力水平有多高？](#)

[第1章 影响力的武器](#)

“专家的话肯定对”“贵的基本等于好的”“销量高的店铺更可靠”。这样的心理捷径很常见，它能帮助我们更高效地解决问题，但也很可能让不法分子有可乘之机。掌握影响力武器的精髓，能让我们做出更明智的决策，同时大幅减少不必要的损失。

[按一下就播放](#)

[把赌注押在抄捷径上](#)

[渔利的奸商](#)

[以柔克刚](#)

[本章小结](#)

[第2章 互惠](#)

为什么商家爱提供试吃品？为什么问卷调查信中常附有小礼物？为什么埃塞俄比亚人不顾自己的困境也要向墨西哥灾区捐款？

- 在达成更大的目标之前，不如先送出一个小恩惠。
- 不请自来的恩惠很可能让你损失惨重。

[互惠原则怎样起作用](#)

[互惠式让步](#)

[如何防范](#)

[本章小结](#)

[第3章 喜好](#)

为什么特百惠如此重视家庭聚会式的销售？世界上“最伟大的汽车

推销员”如何通过一张小小的贺卡实现销量巅峰？为什么我们更容易顺从和自己相似的人？

- 想达成目标，最好能让你和你的合作者彼此喜欢。
- 警惕对请求者的过度好感，并从感性上把请求人和请求本身分开。

[通过喜好赚钱](#)

[我喜欢你的理由](#)

[如何防范](#)

[本章小结](#)

[第4章 社会认同](#)

商家为什么要在菜单上注明哪道菜“最受欢迎”？为什么以高度保密信息著称的奈飞公司会突然大量公布自己最成功影视产品的信息？危急之中如何求救最有效？

- 想说服他人去做一件事，只需证明已有更多的人正在做。
- 警惕明显伪造的社会证据，随大溜之前先综合评估当前环境下的其他证据源。

[社会认同原则](#)

[最佳条件](#)

[负面事件效应](#)

[如何防范](#)

[本章小结](#)

[第5章 权威](#)

著名的米尔格拉姆实验是何原理？为什么人们爱给自己的产品找专家推荐？为什么巴菲特要在不必要的情况下自曝公司的问题？

- 想办法让权威替你说话。
- 参考权威的意见做决策时，要辨明对方是不是真正的专家，以

及和当前领域是否匹配。

[权威高压的力量](#)

[盲目服从的诱惑和危险](#)

[内涵不是内容](#)

[让权威更加可信](#)

[如何防范](#)

[本章小结](#)

[第6章 稀缺](#)

苹果、小米等品牌的饥饿营销有何心理学原理？为什么父母的许诺前后不一最容易让孩子叛逆？为什么人们在恋爱中会倾向于让自己显得更独特？

- 合理地展示稀缺，让你的东西更受欢迎。
- 识破“数量有限”的真正目的再行动。

[物以稀为贵](#)

[逆反心理](#)

[最佳条件](#)

[如何防范](#)

[本章小结](#)

[第7章 承诺与一致](#)

为什么亚马逊公司要推出高价鼓励员工离职的政策？为什么赛百味要把开10 000家分店的目标印在餐巾纸上？如何利用“登门槛”手法实现看似难以完成的目标？

- 尽量让他人公开对你的承诺，以保障落实。
- 在不确定你最初的承诺是否正确时不妨自问，如果回到从前，你是否还会做出同样的选择。

[言出必行](#)

[承诺是关键](#)

[发自内心地认同](#)

[如何防范](#)

[本章小结](#)

[第8章 联盟](#)

诈骗犯麦道夫漏洞百出的庞氏骗局靠什么维持了数十年之久？为什么破冰团建中常包含一些让成员共同吃苦的项目？如何利用联盟原则有效化解伴侣间的冲突？

- 让你的合作对象建立深刻的“我们”感。
- 在注重联盟感的集体中，不要放任个体的不道德行为。

[“我们”的力量](#)

[联盟1：身心合一](#)

[联盟2：行动合一](#)

[跨群体大联盟](#)

[如何防范](#)

[本章小结](#)

[第9章 即时的影响力](#)

何种情况更容易使人忽略多方信息、依据单一因素做出决定？如何判断哪些信号是以刺激我们的捷径反应为目的的虚假信号？既然采用便捷方法是现代生活中的必然选择，那么有哪些已被证明可靠的单一因素？答案就在这7大影响力原则中：互惠、喜好、社会认同、权威、稀缺、承诺与一致、联盟！

[原始的自动反应](#)

[现代的自动反应](#)

[捷径应受到尊重](#)

[本章小结](#)

[参考文献](#)

Influence

第1章

影响力的武器

文明的进步，靠的是人们不假思索就可以做的事情越来越多。

——阿尔弗雷德·诺斯·怀特海德
(Alfred North Whitehead)

简约是复杂的最终形式。

——列奥纳多·达·芬奇 (Leonardo da Vinci)

NEW AND EXPANDED

本书收录的不少研究结果，乍看起来令人困惑，但其实可以通过对人类自然倾向的认识来加以解释。不久前，我在阅读一篇研究时就碰到了这样的情况：该研究让志愿者喝一种旨在提高其智力的能量饮料。一些志愿者按饮料的零售价（1.89美元）付了款；而研究者告诉另一些志愿者说，因为饮料是实验室批发来的，他们只需要支付0.89美元即可。接下来，两组志愿者都要在30分钟内回答尽量多的智力测试题。我以为，第二组人会对优惠价格产生好感，因此也会更努力地尝试，解决更多的问题。但结果跟我想的恰恰相反。^③

这个结果让我想起了多年前接到的一通电话。电话来自我的一个朋友，她新近在亚利桑那州开了一家印度珠宝店。朋友有点前言不搭后语地告诉我她碰到了一件不可思议的事情，她认为，我这个心理学家或许能够为她解释清楚。

故事是这样的，朋友手里有一批绿宝石首饰，一直不大好卖。当时正值旅游高峰期，商店里挤满了客人，绿宝石首饰的质量着实对得起她开的价钱，可就是卖不出去。为了卖掉它们，我朋友尝试了若干种销售技巧：把它们放到更显眼的展示区以引起人们的注意，没用；让销售人员疯狂“推售”，也没用。

最后，她在要出城采购新商品的前一晚，给雇员潦草地写了一张破罐子破摔的字条：“本柜的所有物品，价格乘以 $1/2$ 。”本意是哪怕亏本也得把这批倒霉的货给弄出去。几天后，她回来了，竟发现所有的东西都销售一空，当然了，这也在她的预料之中。可随即她发现，由于自己的字迹太潦草，雇员把“ $1/2$ ”误当成了“2”，所有的首饰都是按原价的两倍卖出去的！这下子，她彻底惊呆了。

就这样，她给我打来了电话。我想我知道是怎么回事，但我告诉她，要解释清楚这件事，她也得听我讲一个故事。其实，这不是我的故事，而是关于雌火鸡的。它属于动物行为学——也就是研究自然环境下

的动物。

影响力研究

雌火鸡是很合格的母亲，它充满关爱，警惕性高，会全心全意地保护小宝宝。雌火鸡会花很多时间来照料小火鸡，做好保暖和清洁工作，把孩子们收拢在身子底下。可有个奇怪的现象，上述一切母爱行为几乎都是靠一样东西触发的：小火鸡发出的“叽叽”声。在雌火鸡的照料过程中，火鸡宝宝的其他特点，比如气味、运作和相貌等，都扮演着非常轻微的角色。要是一只小火鸡发出了“叽叽”声，那么火鸡妈妈就会照料它；要是小火鸡不出声，火鸡妈妈就会完全忽视它，有时甚至还会误杀它。

动物学家M. W.福克斯（M. W. Fox）在1974年做了一个实验，生动展现了雌火鸡对“叽叽”声的极度依赖性。^[4]实验用到了一只雌火鸡和一个臭鼬充气玩具。对雌火鸡来说，臭鼬是天敌，只要它一出现，雌火鸡就会“叽叽”大叫，并用喙啄它，用爪子抓它。事实上，实验发现，哪怕是仅用绳子将一只臭鼬充气玩具拉到雌火鸡面前，前者也会立刻遭到猛烈的攻击。然而，要是在充气玩具里装上一台播放火鸡宝宝“叽叽”声的小型录音机，那么雌火鸡不仅会接受臭鼬，还会把它收拢到自己的翅膀底下。录音机一关掉，臭鼬玩具又会立刻遭到雌火鸡猛烈的攻击。

按一下就播放

雌火鸡的不同举动看起来是何等荒谬啊：它热烈地拥抱起了天敌，仅仅因为对方发出了“叽叽”声；它虐待甚至害死了自己的宝宝，仅仅因为小火鸡没有发出“叽叽”声。雌火鸡像是一台机器，它的母性本能全受

一种声音的自动控制。动物行为学家告诉我们，这种情况并不是火鸡独有的，他们已经确认有大量物种都具有这种规律、盲目、机械的行为模式。

这其实就是所谓的固定行为模式，其中甚至包括极为复杂的行为序列。固定行为模式的基本特点是：构成模式的所有行为每一次几乎都是按相同的方式、顺序发生的。这些行为就好像是被记录在动物身体内置的磁带上一样。在一个完整的求偶或交配过程中，每当出现适合求偶的环境，求偶磁带就会播放；每当出现适合抚养、生育的环境，母爱磁带就会播放。只要按个键，动物体内相应的磁带就会被激活。“哗啦啦”，固定的行为按顺序依次展开。

这其中最有意思的是磁带的激活方式。举例来说，当一种动物要采取行动保护自己领地时，说明此时是同一物种另一动物的侵入启动了前者捍卫领地的磁带，触发了它严阵以待、威胁甚至战斗（如有必要）的行为。然而，这套系统里有个很怪的地方，触发者并不是对手这个整体，而是对手具备的一些特征。通常，触发特征只是整体上不足挂齿的一小方面。有时，颜色就是触发特征。比如，动物行为学家的实验指出，雄性知更鸟只要看到一丛红色知更鸟的胸羽，就会做出有敌人侵犯自己领地的样子，凶猛地攻击胸羽。可只要你拿走那丛红色的羽毛，哪怕是摆上一只惟妙惟肖的雄知更鸟玩具，雄性知更鸟也会对它不理不睬。研究者在蓝喉鸟身上也观察到了类似的现象，只不过，触发蓝喉鸟捍卫领地的是一种特殊的蓝色胸羽。

看到触发特征轻而易举就能欺骗低等动物做出不恰当的反应，我们难免会有点自鸣得意。且慢，有两件事我们千万要搞清楚。第一，这些动物自动化的固定行为模式在大部分时间都是运作良好的。例如，因为只有正常、健康的雏鸟才能发出小火鸡特殊的“叽叽”声，所以雌火鸡根据这种声音做出照料行为是合乎情理的。只对这一种刺激产生反应，普通的雌火鸡做出的行为基本上都会是正确的。只有当像科学家这样的人

故意捉弄它的时候，它那磁带式的反应才会显得傻乎乎。第二，我们人类也有早已预设好程序的磁带，尽管这些自动行为一般是对我们有好处的，可激活它们的触发特征也有可能愚弄我们，让我们在错误的时候播放磁带。

心理学家埃伦·兰格（Ellen Langer）^⑤通过一个实验，巧妙地揭示了人类跟动物相似的自动反应模式。一个众所周知的人类行为原则认为，我们在请别人帮忙的时候，要是能给出一个理由，成功的概率会更大。因为人就是单纯地喜欢做事有个理由。为了证明这点，兰格做了以下研究。

影响力研究

人们在图书馆里排队用复印机，兰格这时想请别人帮个小忙，于是便说：“真不好意思，我有5页纸要印。因为时间有点赶，我可以先用复印机吗？”提出要求并说明理由真是太管用了：94%的人答应让兰格排在自己前面。兰格也试过只提要求：“真不好意思，我有5页纸要印。我可以先用复印机吗？”这么说的效果就差多了。在这种情况下，只有60%的人同意了她的请求。乍一看，两次请求的关键区别似乎在于，前一次的请求给出了额外的信息，即“时间有点赶”。然而，当兰格尝试了第三种请求方式后，却证明发挥作用的地方并不在这儿。关键点并非原因，而是原因之前的“因为”两个字。

兰格的第三轮请求里并没有包含一个足以让人顺从的原因，只用了“因为”，接着便把明显的事实又重复了一遍。她是这么说的：“不好意思，我有5页纸要印。我能先用复印机吗？因为我必须印点儿东西。”结果，差不多93%的人都同意了。虽说这个请求里并没有真正的原因，也没有提供什么

信息来说明人们照着兰格的话去做是合理的。

正如小火鸡的“叽叽”声能触发雌火鸡的自动哺育反应，哪怕声音是从充气臭鼬玩具里发出来的。“因为”这个词触发了兰格实验里被试们的自动顺从反应，哪怕兰格根本没有给他们一个说得通的理由。按下按钮，磁带就“哗啦啦”地播放起来。

此外，当家长要小孩子解释自己的行为时，小孩子最常用的回答是“因为.....就是因为”。说不定，孩子们已经敏锐地察觉出，这个词在成年人的世界里有着不同寻常的分量。“因为”暗示了理由，而人们总是希望自己的行动有理由。兰格还探讨了复印机实验蕴含的更广泛的意义，并说明了自动反应在人类行为中的普遍存在。

人类和低等动物的这种自动行为有重要的相似之处与区别。人类的自动行为模式往往是后天习得的，并不是天生的，它比低等动物固定步骤的模式更灵活，并可对大量的触发事件做出反应。

兰格的另一些调查结果显示，在很多环境下，人类的行为并不会完全按机械化的磁带激活方式展开，但她和其他许多专家也相信：大多数时候，人类行为真的跟机械播放的磁带没什么两样。就拿本文开头举的例子来说：绿宝石起初总也卖不掉，但在售货员误把它们的价格抬高一倍后，顾客们却一拥而上买了个一干二净。为什么他们的行为这么古怪呢？这么说吧，除非你从“按一下就播放”的角度来看这个问题，不然还真没法理解这种行为。

READER'S REPORT | 读者报告

致西奥迪尼先生：

镇上有开古董珠宝店的人给我讲了一个故事，说他是如何学到“东西贵=质量好”这一课的。他有个朋友想给

未婚妻送一份特别的生日礼物，于是，珠宝商选了一串项链，跟他的朋友说这在店里本来要卖**500**美元的，可既然是这位朋友要，给**250**美元就行了。起初，朋友对这串项链非常满意，但听到珠宝商报价**250**美元之后，他脸色沉了下来，不乐意成交了，因为他想给未婚妻买一件“真正好”的东西。

隔了一天，珠宝商终于反应过来是怎么一回事了。他打电话给朋友，让朋友再来店里看另一串项链。这一回，珠宝商给他看了一串新项链，并报上了市场价：**500**美元。他的朋友很喜欢，当场就想买下来。但不等他掏钱出来，珠宝商就告诉他说，因为这是结婚礼物，自己愿意亏损一点，将项链降价到**250**美元。朋友激动万分。这一回，他再也不觉得**250**美元有什么不好了，相反，他欣喜若狂，高高兴兴地把项链买了下来。

这位管理学博士生：

注意，与绿宝石买家的例子相同，真正想买好东西的人对价格低的东西是看不上眼的。我相信，除了“东西贵=质量好”的原则之外，在我们的思维里，“便宜无好货”的原则也是适用的。毕竟，在英语里，“cheap”这个词不仅意味着价格低，也有“次等货”的意思。

顾客大多是生活富裕来度假的游客，对绿宝石认识不多，且他们总用一套标准原则，即“范式”来指导自己买东西：一分钱一分货，价格高就等于东西好。许多研究表明，要是人们对物品的质量拿不准，便经常会用到这一范式。因此，想买“好”珠宝的度假者，一看到绿宝石的价格

涨了上去，就觉得它们更贵重了，也更值得拥有了。价格本身成了质量的触发特征，绿宝石在渴望质量的买家中销量激增，完全是由价格暴涨这一点带来的。

把赌注押在抄捷径上

挑剔游客们的愚蠢购买行为很容易，但要是再仔细想想，我们也许会不那么苛刻。人都是在“一分钱一分货”的教导中长大的，更何况，这在人们的生活中一次又一次地应验过。过不了多久，人们就会把这条原则提炼成“价格高=东西好”。“贵=好”的公式向来是管用的，因为一般而言，物品价值高，价格也会高。较高的价格通常反映了较好的质量。因此，当人们发现自己想要质量好的绿宝石首饰，但对绿宝石懂得却不多的时候，便自然而然地会用上价格这个一贯的衡量标准。

尽管买家本人可能并没有意识到，但只参考绿宝石的价格，实际上就是抄了条捷径。他们并没有煞费苦心地了解每一点能暗示绿宝石首饰价值的特点，力争稳操胜券，相反，他们只把宝押在价格这一点上。因为他们知道，在通常的情况下，这一点是跟物品的质量相关的。他们下注打了赌：光凭价格这一点就能告诉他们需要知道的一切。可这一回，由于售货员把“1/2”误当成了“2”，他们赌错了。而就长期而言，综合他们过去、未来整整一辈子遇到的所有情况，把赌注押在抄捷径上却可能是最为理性的方法。

我们现在可以解释本章开篇研究里那令人困惑的结果了——该研究显示，给人们一杯据说能提高智力的饮料，为饮料支付费用高的人，解决的问题也更多。研究者将这一发现追溯到“贵=好”的刻板印象上：饮料标价为1.89美元时，人们对其效果的预期高于饮料标价为0.89美元时。值得指出的是，仅仅是这种预期就能影响结果。类似现象也出现在另一项研究中，该研究让被试接受轻微电击，并在电击前为之提供止痛药。研究者告诉一半的被试，止痛药的价格是0.10美元，告诉另一半被

试，它的价格是2.50美元。但实际上，所有人服用的止痛药都是一样的，但认为它更贵的人对该药减轻电击疼痛的效果给出了更高的评价（见图1-1）。



图1-1 鱼子酱与工匠精神

这一广告要传达的信息显然是贵等于好。

值得一提的是，如果人们不熟悉一种产品或服务，就尤其可能会套用“贵=好”的原则。在市场营销领域，这种现象的经典案例来自芝华士苏格兰威士忌，这一度是个默默无闻、苦苦挣扎的品牌，但等到管理者决定将其价格提高到远超竞争对手的水平之后，尽管产品本身没有任何改变，销量却直线上升。

事实上，模式化的自动行为在大部分人类活动中的出现频率都是相当高的，因为很多时候，它是最有效的行为方式，而另一些时候，它则是必要的。你我生活在一个极端复杂的环境中，它说不定是地球有史以来变化最为迅速的环境。为了对付它，我们需要捷径。哪怕就是短短的一天当中遇到的每一个人、每一件事，我们也不可能把与之相关的方方面面都辨识、分析出来。我们做不到，因为我们没有足够的时间、精力和能力。相反，我们必须频繁地利用范式和首选经验，根据少数关键特

征把事情分类，一碰到这样或那样的触发特征，就不假思索地做出反应。

有时候，人的行为不适合所处的情境，因为即便是最准确的范式和触发特征都不可能回回管用。我们之所以会容忍这样的不完美，实在是因为没有其他的选择。若没有了这些特征，我们就只能傻站着慢慢分类、鉴别和校准，而错过本该采取行动的时机。种种迹象表明，将来我们会更严重地依赖这些典型范式。充斥在我们生活里的刺激会更为复杂、变数更大，我们必然要越来越多地依赖捷径来应对、解决它们。

心理学家最近发现，我们在日常判断中会使用大量心理捷径。他们把这些捷径叫作“启发式判断”（judgmental heuristics），其发挥作用的方式跟“价格高=东西好”的原则是一样的，由此带来的简化思维在大多数时候都行之有效。可同时，我们也可能因此偶尔犯下错误，付出沉重的代价。跟本书内容最相关的一种“启发式判断”告诉我们，什么时候该相信别人，照着别人的吩咐去做。就拿“专家说的话肯定是对的”这一条捷径来说吧，在第5章我们会看到，我们的社会出现了一种令人不安的发展趋势，人们往往会轻率地接受权威人士的说辞，照着他们的指点去做事。这也就是说，我们不是先思考专家的论点，看看值不值得相信，而是直接忽视论点，仅仅因为专家是“专业人士”，就选择相信他们。这种在某种环境下机械地回应某一信息的倾向叫作自动化反应或“按一下就播放”式反应；对所有相关信息进行彻底分析后做出反应的倾向，则叫作可控式反应。

相当多的实验研究表明，在有愿望也有能力仔细分析信息时，人们会更多地按可控的方式去处理；可要是条件不允许，人们则更可能采取较为容易的“按一下就播放”式方法。例如，有这么一项研究：

影响力研究

密苏里大学的学生听了一次录音演讲，该演讲认为，所

有的高年级大学生必须通过综合考试才能毕业。这跟在场的一些学生是密切相关的，因为录音里说，综合考试明年就生效，也就是在大四学生毕业之前。自然，这个消息让学生们想要仔细地分析其论点。然而，本次研究另外还找了一些学生，他们被告知考试制度要在他们毕业之后才生效，所以这个消息对他们来说没什么重要性，因此，他们没有仔细考虑其中的论点正确与否的强烈需求。

研究结果一目了然：自身不受影响的学生基本上被演讲人在教育领域的专业身份给说服了，他们使用了“专家说的话肯定是对的”的原则，对演讲者的论点是否合理并未给予太大的关注。反过来，那些自身要受影响的被试，并不在乎演讲者的专业地位，而更在乎他的论点质量。

这样看来，在使用“按一下就播放”这种危险的响应方式时，我们其实是给自己架了一张安全网：要是议题对我们自己很重要，我们就会拒绝根据单一特征来触发反应，因为这太奢侈，我们玩不起。毫无疑问，这种情况经常发生。不过，我并没有完全放下心来。各位读者或许还记得，我们刚才提到过，只有在人既有欲望也有能力的时候，才会以深思熟虑的可控方式做出反应。现代生活形形色色、节奏紧张，有证据表明，哪怕是许多跟我们切身相关的主题，我们也没法彻底想清楚后再做决定。也就是说，有时候问题太复杂、时间太紧张、分心的东西太多、情绪唤起太强烈或心理极度疲劳，从认知条件来看，我们都无法谨慎思考，不管议题是否重要，我们都只能抄捷径。

最能生动说明这一点的，是航空界存在的一种现象，可称之为“机长综合征”，它关系到许多人的生命安危。美国联邦航空局的空难调查员发现，很多时候，机长犯了非常明显的错误，但其他机组人员却不做纠正，结果导致坠机事故。看起来似乎是这样：尽管事情对每一个人来

说都性命攸关，可机组成员还是使用“既然专家都这么说，那肯定没错”的思维捷径，忽视了机长所犯的灾难性错误，或者即使注意到了也没采取行动。

影响力研究

国际商用机器公司（IBM）前任主席，小托马斯·沃森（Thomas Watson）记录的一个案例为这种现象提供了生动的证据。第二次世界大战期间，沃森受命调查有高级将领遇难或受伤的飞机事故。有一起事故牵涉到著名的空军将领乌扎尔·恩特（Uzal Ent）。起飞之前，恩特的副驾驶员生病了。新分来的副驾驶员听说能和传奇将领一起飞行，感到很光荣。起飞时，恩特轻声哼着歌，并随着脑海里的节奏点头打着拍子。新的副驾驶员以为这是恩特要他把飞机的助跑轮收起来。尽管那时候他们的速度还非常缓慢，根本没法飞上天，副驾驶员还是把助跑轮给抬了起来。结果，飞机的腹部立刻贴在了地上。混乱中，螺旋桨叶片切入了恩特的背部，割断了他的脊椎，导致他下半身瘫痪。按照沃森的描述，副驾驶员是这样解释自己的行动的：

“副驾驶员作证的时候，我问他：‘既然你知道飞机不能起飞，为什么还要把助跑轮升起来呢？’

“他说：‘我以为是将军要我这么做的啊。’他真是个蠢货。”

蠢吗？在这一起事件中，副驾驶员的确是够蠢的。但现代生活的迷宫确实需要捷径，从这个角度看，这就完全可以理解了。

事实上，自动响应在我们生活中有极大的必要性和存在价值。除非人们有仔细检查输入信息的动机和能力，否则他们就会依赖直观判断来对信息做出反应。

其中有一点颇具启发意义，即尽管我们并不经常采用深思熟虑的复杂方式应对那些事关自身的重要议题，我们却希望别人，比如顾问、医生、会计、律师、经纪人这样做。当我们为一个复杂而必然的选择心力交瘁时，我们仍然想对它进行全面、细致的分析。然而，这样的分析是我们力所不能及的，只能通过捷径实现：依靠专家。真是讽刺。

图1-2中的客机坠入华盛顿国家机场附近的波托马克河的几分钟前，机长和副机长就如何在机翼结冰的情况下起飞交换了意见。飞机的“黑匣子”记下了他们的对话。



图1-2 “机长综合征”的灾难性后果

副机长：读数似乎不大对劲儿。

机长：是对的。

副机长：不，我想不对。（停顿了几秒钟）好吧，也许是对的。

副机长：拉里，我们要降落了。

机长：我知道。

（传来一阵撞击声，在撞击中，机长、副机长和67名乘客丧生）

渔利的奸商

奇怪的是，尽管自动化行为模式用处十分广泛，且将来还会变得越来越重要，可我们大多数人还是对它知之甚少。或许这正是因为它们总是以机械化、不假思索的方式发生和出现的。不管怎么说，它有一点特性，我们务必清晰地意识到。不然，要是碰到知晓它们奥妙的人，那我们可就门户大开、任其摆布了。

为了充分理解我们的脆弱本性，让我们再来看一眼动物行为学家的的工作。原来，不仅只有这些录下小火鸡“叽叽”声、拿鲜艳胸羽做伪装的学者通晓如何激活各物种的行为磁带。有一种通常叫作“拟态体”的生物也会模仿其他动物的触发特征，企图诱骗后者在不恰当的时间错误地播放原本正确的行为磁带。此时，拟态体便会抓住机会，利用对方的错误行为，达到自己的目的。

影响力研究

有一种雌萤火虫（*Photuris*属）会对不同类（*Photinus*属）的另一种雄萤火虫设下致命的陷阱。出于可以理解的原因，这种雄萤火虫从来都小心翼翼地不去接触嗜血的雌萤火虫。然而，通过数百年的进化，雌性猎人锁定了猎物的一个弱点：受害物种的雌性成员在到了交配期的时候，会发出特殊的闪光求偶代码来通知雄性同类。靠着模仿这一闪光求偶信号，雌性杀手触发了猎物们的交配磁带，诱得这些可怜的

家伙们不由自主地飞过来，投入死亡的怀抱，成了雌性杀手的一顿美餐。

很明显，雄性容易受强大的交配信号迷惑，连人类也不例外。维也纳大学的两位生物学家——阿斯特丽德·朱伊特（Astrid Juette）和卡尔·格拉默（Karl Grammer）偷偷给年轻的男性闻一种叫“性诱引剂”的化学气体，这模仿的是人类女性阴道的味道。之后，研究者又让男性被试给女性被试的面部吸引力打分。这些男性被试闻了性诱引剂后，觉得所有女性被试的吸引力都提高了，再也无暇关注她们外貌上的真正差异了。一些最原始的病原体还会通过模拟化学物质进入健康的身体。

在生存斗争中，几乎每一种生命形式都有着对应的拟态体，就连一些最原始的病原体也不例外。这些聪明的细菌和病毒通过模拟对身体有益的激素或营养素的部分关键特征，进入健康的宿主细胞。健康的细胞受了欺骗，热切又天真地把狂犬病、单核细胞增多症和普通伤风的病源接纳下来。

不足为奇，人类世界也有着极其相似的事情。有些逐利的奸商也会模仿触发特征，激起我们的自动响应机能。不过，人类之外的生物大多都是本能的响应序列，而人类却不同。我们的自动磁带通常来自通过经验习得的心理学原理或范式。尽管效力各有高下，但有些原则也能够极为强烈地左右人的行动。我们从小就接触这些原则，它们对我们的影响也十分普遍，所以，寻常人很少能察觉它们的力量。可有些别有用心的人却看穿了这些原则，把它们当成了触手可及的武器，即自动影响力武器。以社会认同原则为例，该原则认为，人们往往会相信或去做自己身边人相信以及正在做的事。每当我们在网购之前查看评论或星级时，都是在奉行这一原则。

然而，点评网站也存在许多虚假评价。2019年，美国联邦贸易委员会投诉了一家化妆品公司，指控其评价造假。投诉引用了经营者对员工

所说的一句话，以说明虚假评价的炮制者有多么清楚它们的效力：“如果你发现有人说‘我不喜欢产品的某某方面’之类的话，就写一篇意见相反的评价。评价有着强大的力量，人们会通过他人所说的话来自我说服，并从别人的意见中寻找解开自己疑惑的答案。”至于如何识别它们，接下来的“线上影响力”专栏提供了一些方法。

EBOX | 线上影响力

识别虚假在线评价的方法

杰西卡·斯蒂尔曼 / 文

一款新的计算机程序能以令人难以置信的准确率识别虚假评价，《科学》杂志指出，其准确率高达90%。

当你上网购买产品时，评价可能会在你的决定中发挥很大的作用。我们更乐意选择亚马逊网站上评分为五星而非四星半的酒店，也更可能会在前房客的热情推荐下预订民宿。

当然，我们也都知道这些评价有可能作假——要么是卖家花钱买的，要么是竞争对手恶意投放的。因此，康奈尔大学的一支研究团队决定开发一套能够识别虚假评价的计算机程序。

那么，该怎么判断一家评分为五星的酒店的房间其实可能又小又窄且潮湿发霉呢？怎么识别出一台评价很高的面包机质量低劣，用不了多久就会坏掉，说不定你甚至连一块完整的面包都没能用它做出来呢？根据康奈尔大学的研究，你应该小心以下这样的评价：

- 缺少细节。人很难描述自己没有真正体验过的东西，这就是为什么假评价往往只会做整体性的赞美，而不会深入具体细节。“比如，真实的酒店评价更有可能是跟酒店相关的具体词语，如‘浴室’、‘入住’或‘价格’。骗子写的评论则更多地跟

场景铺垫有关，如‘度假’、‘出差’或者‘我丈夫’。”

- 包括更多的第一人称代词。如果担心自己表现得不够真诚，你往往会更多地谈论自己。这就是为什么虚假评价里常常会出现更多的“我”字。
- 动词比名词多。语言分析表明，假货中往往包含更多的动词，因为它们的作者经常用听起来很愉快或令人震惊的故事来代替实际的见解。真正的评价则更注重名词。

当然，光靠这些标记还不足以让你鉴别出真假评价。但结合其他方法来评估其可信度，如查看不同类型的卖家，留心可疑的时间戳等，你或许就能做出更明智的决定。

作者点评：请当心这些虚假评价，在线评论网站正跟它们进行着一场旷日持久的战争。我们应该加入战斗。稍做比较，我们就能理解：从2014年到2018年，每个类别的在线评价好评率都有所上升。例如，在购买前就进行阅读评价的用户从88%上升到92%。但也有例外：信任企业的用户从72%下降到了68%。这些虚假评价似乎正在破坏我们对自己所寻求的快捷信息价值的信心。

有些人很清楚自动影响力武器藏在哪儿，他们娴熟老练地使用它们，借此达成自己的意图。他们出没于各种社交场合，要求别人顺从自己的愿望，其成功的概率令人目瞪口呆。他们之所以能屡战屡胜，奥妙就在于他们对所提要求的结构体系做了设计，他们利用社交环境中这样那样的自动影响力武器把自己武装了起来。要实现目的，有时候只需要选择一个恰当的字眼，就足以调用某条强大的心理学原理，按下我们自动行为磁带的播放键，让我们机械化地响应这些原理。他们如何从中得利呢？这一点你可不用担心，他们学得可快了。

还记得我那个开珠宝店的朋友吗？虽说她第一次是意外捡了便宜，

可没过多久，她就开始有意识地定期利用“价格高=东西好”的公式了。如今，每逢旅游旺季，若碰到有什么东西不好卖，她就会先来上一轮大加价，以此加快销售速度。她说，这一手操作不仅成本低，效果还出奇地好。不知情的游客频频中招，因此她获得了极为可观的利润。

就算最初的提价不成功，她还可以祭出“特价”的大旗，按原先的标价卖给那些喜欢买打折货的人。面对虚高的标价，消费者仍然会产生“贵=好”的心理，于是又被她暗中宰了一刀。

“价格高=东西好”原则的后一种用法，即吸引淘便宜货的买家，并非我朋友的原创。近年来的研究表明，为一种商品打上“原价xxx元；现价xxx元”的标签，这种促销手法的效果极佳。实际上，早在研究者确认它的效力之前，零售商们就已经把这一招运用得炉火纯青了。教育家兼作家利奥·罗斯滕（Leo Rosten）讲过一个例子。

影响力研究

20世纪30年代，德瑞贝克兄弟西德和哈里在罗斯滕所住的街区开了一家男装裁缝店。每当西德有新客户对着店里的三开大镜子试衣服，他就会告诉对方，自己听力有些问题，并反复让客户说话时提高音量。只要客户喜欢上了哪套衣服，问起价格，西德就会大声叫他兄弟。哈里是首席裁缝，在店堂后面。“哈里，这套衣服多少钱？”哈里抬头看看自己做的衣服，并大大地抬高了真实的价格，他高喊着回答说：“那件漂亮的纯羊毛西服，要42美元。”西德假装没听见，用手在耳边做个敞口杯子的形状，再问了一遍。哈里再次回答：“42美元。”此时，西德转过身，对客户说：“他说要22美元。”好多人这时都会急急忙忙地付钱，抢在可怜的西德发现自己“搞错了”之前，抓着西服狂奔出店，并且还满心以为自己捡了天大的便宜。

以柔克刚

学过日本柔道的女性在抗击对手时，自己往往用力很少。相反，她会尽可能地利用重力、杠杆作用、动量和惯性等物理原理中天然蕴含的力量。要是她知道怎样调动这些原理、从哪儿去调动，便能轻轻松松地击败体格比自己壮硕的对手。倘若有人掌握了我们周边世界天然存在的自动影响力武器，事情也是一样。奸商利用这些武器的威力对付他们的靶子，而自己却不费吹灰之力。这些蓄意牟利的人因此得到了莫大的好处，他们虽然一点儿也不显得像是在刻意操纵，却实实在在地操纵了对方。就算是受害者本人也大多认为自己的顺从是自然而然的，并不是对方为贪图好处而刻意设计的。

举个现成的例子。人类在认知方面有个原理，叫“对比原理”。当两样东西一前一后地展示在你面前时，我们要怎样看待其间的区别，对比原理又会对我们产生怎样的影响。简单地说，要是第二样东西跟第一样东西相当不同，那么，我们往往会认为两者的区别比实际上要大。这样一来，如果先搬一种轻的东西，再拿一件重的东西，我们会觉得第二件东西比实际上更沉；而要是我们一开始直接就搬这件重东西，反倒不会觉得它有这么沉。对比原理是心理物理学领域确立的，它不仅适用于重量，还适用于其他各种感官知觉。假如我们正在减肥，午餐时试图估算一份芝士汉堡的热量。要是我们先估算一份沙拉的热量，那我们会觉得芝士汉堡的热量比实际要高得多。一项研究表明，人们的估计值比实际高38%。因为跟沙拉相比，芝士汉堡的热量似乎要高得多。同理，聚会时，要是我们先跟一个非常有魅力的人聊天，接着插进来一个相貌平平的家伙，那么我们会觉得第二个人简直没劲儿透了，而他可能其实并没有那么索然无味。

同样，研究表明，观看了媒体上性感的身体后，现有伴侣的性吸引

力会下降。一些研究者发现，艺术作品的吸引力也存在类似影响，即如果观众在看了较高质量的抽象画之后再观看质量普通的抽象画，后者对其的吸引力就会明显降低。有证据表明，对比效应在没有认知识别的情况下仍可发挥作用，甚至对老鼠也有效。

为了向学生们介绍该原理，我会在课堂上用到知觉对比的另一个例子。

影响力研究

每名学生依次坐在三桶水跟前：一桶冷水，一桶常温水，一桶热水。学生把一只手放进冷水里，另一只手放进热水里，之后，教授要他们把两只手同时放进常温的那桶水里。学生脸上立刻便露出好笑的困惑表情：尽管两只手放在同一桶水中，可刚刚放在冷水里的手觉得它是热水，刚刚放在热水里的手却觉得它是冷水。

实验想要说明的是，基于先前所发生事件的性质，相同的东西，即常温的水，会在之后显得极为不同。此外，对其他事物的感知，比如大学课程成绩，也有可能受到类似因素的影响，具体可见图1-3。

一封高明的信

亲爱的爸爸妈妈：

我上大学后就一直疏于写信回家，真是不好意思。我要向你们报告我的近况啦，但在你们读下去之前，请先坐好了。没坐好之前千万别往后面读，好吗？

啊，我如今一切都好。我的颅骨骨折和脑震荡差不多已经好了。那时我才到这里不久，因为宿舍失火而跳出窗外摔伤造成的。我在医院里只待了两个星期，现在基本上恢复了正常，烦人的头痛每天也只会来上一次。幸好，宿舍着火后我跳出去时，隔壁加油站的伙计看见了我，他马上打电话找了消防队，叫来救护车，他还到医院看望我。由于火灾之后我无处可住，他好心地邀我去他公寓住。其实那就是一个地下室啦，不过是很可爱的那种。他是个非常棒的小伙子，我们深深地坠入了爱河，还打算结婚呢。我们还没有订好具体的日子，但肯定是在别人看出我怀孕之前吧。

是的，爸爸妈妈，我怀孕了。

我知道你们有多想升级当外公外婆，也知道你们肯定会欢迎我的宝宝，给他无私的爱护与关怀，就像我小时候那样。我们婚期延后的原因是，他有点小小的感染，所以我们没通过婚前血检，因为我也不小心被他传染了。

好啦好啦，我要给你们讲我真正的近况啦。我想说的是，我的宿舍没有着火，我没有摔成脑震荡，也没有摔断骨头，我没有住院，没有怀孕，没有订婚，没有感染，我现在连男朋友也没有哦。不过，我的美国历史得了个“D”，化学得了个“F”，我希望你们能从正确的角度看待这些分数。

深爱你们的
雪伦

作者点评：雪伦的化学可能学得不怎么样，但她的心理学能得个“A”。

图1-3 巧用知觉对比

请放心，这个由对比原理带来的小小影响力武器，早就有人用过了。这一原理带来的巨大好处，在于它不光管用，还几乎能让人无法察觉。利用它的人尝够了甜头，因为你根本察觉不到整个环境是他们早就布置好的。服装零售商就是个很好的例子。

影响力研究

假设有人走进一家时尚男装店，说自己想买三件套的西服和一件毛衣。如果你是售货员，你该先给他看哪样东西，好让他花最多的钱呢？服装店指点销售人员，要先给顾客看贵的东西。但依照人们的常识，顺序应该反过来才对：要是人们买西服时就花了大把的钱，他恐怕不愿再多花钱买毛衣了。但服装零售商们对此却是心知肚明的。他们完全依着对比原理来设计销售策略：先卖西服，因为顾客接下来买毛衣的时候，哪怕它再贵，价格跟西服一比，也会显得不怎么贵。要是这位顾客还想给自己的新西装买些配件，如衬衣、皮鞋和皮带，同样的原理也适用。尽管对比原理的预测违背常识，证据却支持它。

售货员先展示昂贵的物件对他们更有利；不这么做，不仅会白白放弃对比原理造成的影响，还会使这一原理掉转枪口，对准他们自己。先拿出便宜的东西，再拿出昂贵的东西，会使昂贵的东西显得更昂贵，而这定会给大多数销售组织造成不良后果。因此，正如同一桶水会因为之前手接触的水温高低而显得冷一些或热一些，卖东西的人也可以让同样东西的价格显得高一些或低一些，它完全取决于最先展示的价格。

对知觉对比的巧妙应用，绝不仅限于服装商。我在卧底调查房地产公司采用的说服手段时，就碰到过一种采用了对比原理的手段。

影响力研究

为掌握窍门，一个周末，我陪同销售员菲尔跟有意买房的主顾去看房子。菲尔负责指点我，帮我度过实习期。我很

快注意到一件事，每当菲尔带顾客去看他们想买的房子时，总是会先给他们看几套不甚合意的房子。我问菲尔为什么，他笑了。原来，这些房子是所谓的“垫底货”，公司手里总会留一两套破房子，还会标上虚高的价格。这些房子并不打算卖给客户，只是用来给他们看看，有了破烂房子做比较，公司手上真正要卖的房子就显得更加合适了。不是所有销售员都采用“垫底”法，但菲尔会用。他说，当他带着顾客看了破房子，再带他们看公司真正想卖的房子时，顾客总会“眼睛一亮”。菲尔喜欢看到他们的那种表情，“等他们看了几处‘垫底货’后，我给他们找的房子就会显得相当妙不可言”（同类手法可见图1-4）。



“希望在座的各位今晚在回家的路上，都停下脚步仰天想想，
我们比赛第二节的失利，算得了什么呢。”

图1-4 “一流的想法”

对比原理的用处非常多。

汽车经销商也会用对比原理，他们要等到跟客户谈妥了一辆车的价格后，再一一报上备选配件。几万块都花了，再多花个几百块升级音响

系统这样的小玩意儿，似乎太微不足道了。车商随后会建议给车窗贴膜、换用更好的轮胎，或是做些特别的车内装饰，道理也是一样的。总之，诀窍在于，要单独地提出各个选项，这样每一项的小价目跟已经确定的大数目比起来就会不值一提。买过车的人都可以作证，把这些看起来没什么大不了的配件价格加起来一算，整车价格就好像注水猪肉一般涨了上去。顾客目瞪口呆，手里拿着签好的合同，不知道是怎么回事，好不容易才琢磨出是自己上了当，可谁也怪罪不得。经销商站在一边窃窃私笑，他们的柔道功夫又精进了。

READER'S REPORT | 读者报告

致西奥迪尼先生：

在奥黑尔机场等着登机的时候，我听到工作人员说飞机超载了，要是乘客自愿改搭下一班飞机，他们愿意赔偿价值**10 000**美元的代金券！这个夸张的数字当然是开玩笑的，它是拿来逗乐的，乘客们也真的笑了。但我注意到，等那人报出真正的赔偿金**200**美元之后，没有人愿意接受。于是他不得不提了两次价，先是**300**美元，然后是**500**美元，这才找到了自愿的人。

我当时正在读您的书，我意识到：虽说他把人们逗乐了，但根据对比原理，他也把事情搞砸了。他先提到的**10 000**美元成了人们比较的标杆，这一下子就让几百美元相形见绌。这个玩笑可真是代价不菲，他的航空公司要为每名自愿改乘的人多付**300**美元。

芝加哥大学商学院的这名学生：

这名工作人员该怎样使对比原理为己所用呢？他可以先用5美元讲笑话，然后再报出真正的赔偿数额——200美元。这下子，后一个数听起来就有魅力多了。如果他这么做，我敢肯定他既能逗人发笑，也能超额完成目标。

本章小结

- 动物行为学家注意到许多动物都存在机械刻板的行为模式。这种机械的行为序列叫作固定行为模式，它们跟人类的某些自动响应，即按一下就播放，有很大的相似性。对人类和近似人类的动物，自动行为模式往往是由该环境下相关信息的单一特征所触发的。这种单一特征或触发特征，大多数时候是极其可贵的，有了它们，个体无须对每一条信息进行详尽彻底的分析，就能决定正确的行动方针。
- 捷径反应的优点在于它的效率和经济性。只根据一种常见的信息触发特征自动做出反应，能节省人们宝贵的时间、精力和心智能量。而这种反应的缺点在于它易受愚弄，一犯错代价就很大。只根据可用信息的一个片段，甚至是一般都能预测出的某个片段做出反应，尤其是按无意识的自动方式做出反应，人们出错的概率就会大增。要是其他人通过故意的安排（如对触发特征加以操纵）让自己牟利，刺激对方在不恰当的时机做出合乎他们心意的行为，那么对方犯错的概率就更大了。
- 大部分的顺从技巧，即刺激一个人顺从另一个人的要求，都可以从人们自动化的捷径反应这一角度来理解。我们文化中的大部分个体都对顺从建立起了一套触发特征，这是

一些具体的信息片段，通常情况下，它们可以告诉我们什么时候顺从他人的请求会是正确的、有利的。这些顺从的触发特征，每一条都可能会变成影响力武器，刺激人们同意他人的请求。

- 知觉对比是把两种不同事物之间的差异看得比实际上更大的倾向，是一些说服专业人士（销售、广告商等）喜欢利用的影响力杠杆。例如，房产经纪人或许会先向有意向的购房者展示一两处缺乏吸引力的房子，再给他们看另一处更有吸引力的房子，这时后者在购房者眼里看来会显得吸引力更强。这种影响力杠杆的优点之一是，人们通常很难意识到有人在战术性地使用它。

Influence

第2章

互惠

如果你压根儿不想还，那就干脆别伸手要。

——《便西拉智训》

NEW AND EXPANDED

几年前，一位大学教授做了个小实验。他给随机抽选的陌生人寄圣诞贺卡，虽说他料到此举会掀起些波澜，但所得到的回应还是让他吃了一惊：回寄的贺卡从四面八方涌了过来，全是从没见过，也没听说过他的人给他寄的。这些回寄了贺卡的人，绝大多数根本没打听这位不知名的教授到底是谁。他们收到了教授的贺卡，一下子便启动了自动反应，给他回寄了贺卡。

尽管这是一项小范围研究，却指出了我们身边最有效的影响力武器之一——互惠原则，发挥着什么样的作用。这条原则说，要是人家给了我们什么好处，我们也应当尽量回报。假设有位女性帮了我们的忙，我们应当也帮她一回；倘若有位男性送给我们一份生日礼物，我们也应当记得在他生日时献上小小心意；要是有对夫妇邀请我们参加聚会，我们下次务必也要记得邀请他们参加我们的聚会。回赠的贺卡、生日礼物和聚会邀请，似乎不足以证明这条原则，但事实上，它们能带来相当大的改变。研究者考察了英国的慈善筹款人员，他们的工作是去找投资银行家募捐，要求对方捐赠一笔很大的慈善捐款——相当于捐赠者一天的薪水，有时甚至超过1 000美元。很值得注意的是，如果筹款员在提出请求之前先送一小包糖果作为礼物，募捐成功率就会提高一倍以上。

这一原则甚至可扩展到国家行为上。1215年，英国《大宪章》用它定义了战争爆发时各国应怎样对待来自敌国的商人：“如果我们的人在他們那里安全，那他们的人在我们的土地上也应该是安全的。”因此，依照互惠原则，我们有义务在将来用送礼物、发出邀请等手段回报别人的好意。由于接受行为往往伴随着回报的义务，“承蒙美意”（much obliged）一类的短语几乎成了“谢谢你”的同义词。不仅英语是这样，其他不少语言也是如此，如葡萄牙语里的“obrigado”、日语里的“済みません”（谢谢你，字面直译的意思是“尚未结束”，这更好地表现了接受者将来的义务）。

深刻烙印在脑海中的互惠原则及伴随其而来的亏欠还债感，在人类文化中十分普遍。阿尔文·古德纳（Alvin Gouldner）和其他社会学家在对此深入研究后报告说，所有的人类社会都认同互惠原则。在每一个独特的社会文化里，它都无处不在，甚至可以说它已经渗透到了每一种交换形式当中。说起来，源自互惠原则的成熟的礼尚往来系统，甚至有可能是人类文化的一个独有特征。著名考古学家理查德·利基（Richard Leakey）认为：正是因为有了互惠体系，人类才得以被称为人类。他说：“由于我们的祖先学会了在‘有债必还的信誉网’里分享食物和技巧，我们才变成了人。”文化人类学家利昂内尔·泰格（Lionel Tiger）和罗宾·福克斯（Robin Fox）认为这种“欠债网”是人类的一种独特适应机制，有了它，人类才得以实现劳动分工，交换不同形式的商品和服务，让个体相互依赖，凝结成高效率的单位。

为产生泰格和福克斯所描述的那种社会进步，对未来的还债义务感至关重要。这种普遍认同的强烈感觉，给人类社会的进化带来了巨大的影响，因为它意味着一个人给了另一个人某种东西（如食物、精力或照料等），却不用担心它会变成损失。人可以把各种资源给予他人，却又不是无偿的，这在进化史上还是头一遭。由此而来的结果是降低了一对一资源交换的天然门槛。救助、送礼、防卫和贸易等复杂而协调的体系有了实现的基础，这为具备这些体系的社会带来了莫大的好处。既然互惠原则对文化有着这么明显的适应性结果，那么，靠着人人都经历过的社会化过程，它便根植到了我们每一个人的大脑深处。

事实上，有些社会已经把互惠原则变成了正式的仪式和典礼。在巴基斯坦和印度的部分地区，有一种交换礼物的常见风俗，叫“瓦尔坦班吉”（Vartan Bhanji）。古德纳对“瓦尔坦班吉”评论道：

值得我们注意的是……这套制度煞费苦心地把所有人情债都结清。比如，举办婚礼时，宾客离场，女主人会送他们糖果当礼物。她一边送客，一边说：“这5颗是你的。”意思是：这

是偿还你从前给我的。接着她又会额外添一些，并说：“这些是我送你的。”等到下一次婚礼，她便会拿回这些礼物，此外又会收到一些以后要还的新礼物，如此周而复始。

对节日贺卡的研究是从菲利普·昆兹（Phillip Kunz）开始的。后来，他的女儿、行为科学家珍妮弗·昆兹（Jenifer Kunz）对此做了扩展（这是人类行为存在连续性的绝佳例证），她发现，如果发送第一张贺卡的人属于地位更高的人，回赠率还会更高。

早在古德纳等社会学家、利基和罗杰·卢因（Roger Lewin）等考古学家，以及泰格和福克斯等文化人类学家发现之前，社会科学家就已经认识到社会内部和社会之间互惠交换的可取性。新近的证据表明，互惠原则同时适用于积极交换和消极交换，这也正印证了W. H.奥登（W. H. Auden）的著名诗句：“我和大家都知道 / 每个人从小就懂得/做恶事者有恶报。”可以说，互惠原则向我们做了这样的保证：种什么样的因，就会得什么样的果。这甚至适用于人机交流。从一台计算机上获得高质量信息的用户，回报给这台计算机的信息也会更好，反之亦然。总的来说，互惠是人类行为的驱动力。

虽说人情债事关未来，但时间跨度也并非没个限制。尤其是相对较小的恩惠，偿还的愿望似乎会随着时间的推移变得越来越淡。可要是礼物真的非常贵重、令人难以忘怀，那么这种偿还的义务感也可以延续得相当久远。要说明互惠义务会给将来造成多么深远的影响，用墨西哥和埃塞俄比亚之间发生的一件怪事来解释再恰当不过了。这个故事，跟一笔5 000美元的救灾款有关。

影响力研究

1985年的埃塞俄比亚，说是饿殍遍地、贫困潦倒绝不放过。经济萧条，连年的干旱和内战彻底摧毁了食物供应链，

成千上万的国民因疾病、饥饿而死去。在这样的困境下，要是墨西哥向它捐出5 000美元的救灾款，我肯定不会感到惊讶。可报上的一条简讯居然说，捐赠的方向刚好反了过来，我不免大吃了一惊。埃塞俄比亚红十字会的官员决定向墨西哥捐赠5 000美元，帮助当年遭受墨西哥城地震的灾民。

每当人类行为的某个方面让我摸不着头脑时，我就会有一种想去一探究竟的冲动。这个嗜好虽说常给我自己惹麻烦，但对我从事的职业却颇有好处。对上面这件事，我找到了更为详细的资料。恰好有个记者跟我一样，被埃塞俄比亚的举动搞得全无头绪，也想刨根问底知道为什么。他得到的回答为互惠原则做了强有力的证明：尽管埃塞俄比亚当前急需援助，可送到墨西哥的这些钱，是1935年意大利人入侵埃塞俄比亚的时候，墨西哥援助给后者的。听了这样的说法，我虽说还是免不了惊讶，却不再困惑了。巨大的文化差异、千山万水的阻隔、严重的饥荒、几十年的岁月、眼前的私利……这么多的不利因素，都没能阻止埃塞俄比亚人报恩的需求，偿还人情债的义务感战胜了一切。

倘若持续了半个世纪之久的人情债，可以用埃塞俄比亚文化的特殊性质来解释，那么再来看看另一件起初显得怪异的事情是怎么解决的。

2015年，94岁的英国著名出版商亚瑟·乔治·韦登菲尔德勋爵（Arthur George Weidenfeld）找到了一种回报的方式。他组织并资助了“避风港行动”（Operation Safe Havens），把叙利亚和伊拉克的基督教家庭从战争地区转移出来。人们虽然对这一善举赞誉有加，但也批评这一行动的范围太窄，并好奇为什么勋爵不将救助之手伸向该地区同样受到威胁的其他群体。

也许有人会想，勋爵只想帮助信奉基督教的同胞。但如果你意识到韦登菲尔德勋爵是犹太人，就会明白这个简单的解释站不住脚。1938年，少年韦登菲尔德搭乘着一列满载犹太孩子的列车来到英国，躲过了

纳粹在欧洲对犹太人的迫害。这一趟列车之旅，是由基督教人道主义团体联合组织的，他们通过这种途径救出了数千名犹太儿童。韦登菲尔德从互惠原则优先的角度，对自己的行为做了解释：“我无法拯救整个世界，但站在犹太人和基督徒的角度，我有一笔人情债要还。”显然，互惠的力量可以拯救生命，并持续一生。

埃塞俄比亚对墨西哥恩情的惦念和韦登菲尔德勋爵对基督徒的亏欠感，跟下面这个例子比起来或许显得黯然失色：一群法国孩子，隔了几个世代之后仍强烈地想要帮助一群自己从没见过的澳大利亚孩子。1918年4月23日至24日，第一次世界大战即将结束时，为从德国军队手中解救法国村庄维莱布勒托讷，澳大利亚军团牺牲了几个营的兵力。2009年，维莱布勒托讷的一群小学生听说，一场森林大火摧毁了澳大利亚小镇斯特拉瑟文，便筹集了21 000美元，帮助重建当地小学。一家报纸报道：“他们对自己要帮助孩子知之甚少。他们只知道，自己的曾祖父辈曾在91年前许下承诺，绝不忘记澳大利亚，以及为解放自己村庄而牺牲的1 200名澳大利亚士兵”。

尽管上文的救命之恩能让人产生持久的亏欠感，但并非所有这类行为都有同样的效果。有充分的证据表明，随着时间的推移，日常恩惠会随着时间的流逝让人失去亏欠感。有一组研究甚至发现，在施恩行动完成之前，受惠者对施恩者最为感激。结论是什么呢？小小的帮助行为遵循“百吉饼法则”：人们更喜欢新鲜热乎的百吉饼，而不是硬邦邦的陈饼子。

READER'S REPORT | 读者报告

致西奥迪尼先生：

受训期间，一位前辈告诉我，我会喜欢跟我的上司合作的，因为他是一个非常好又非常慷慨的人。她说，上司

曾在各类场合送给她鲜花和其他礼物，她是因为要生孩子才决心辞职的，要不然，她一定还会在这个岗位上干好多年。

我在这位上司的手下工作了**6**年，也经历了同样的事情。他给我和我的儿子送圣诞礼物，还会给我送生日礼物。两年前，我在这个岗位上的薪水已经到最高级了，而这个工作是没有升职机会的，所以我要么选择参加加利福尼亚州政府的考试，然后重新申请到另一个部门，要么在私人企业重新找一份工作。但我发现自己并不想找新工作，或是换到其他部门。我的上司就快到退休年龄了，我想等他退休以后再换地方，因为他对我那么好，我觉得有义务留下来。

俄勒冈州的这位政府雇员：

你在谈及自己目前的就业形势时用的字眼给我留下了很深的印象，你说“想等”上司退休以后再换职位。看来，上司小小的善意之举已经使你无意识地产生了一种亏欠感，让你没法去换一个收入更好的岗位。对于希望培养员工忠诚感的管理者来说，这值得借鉴。这个故事还给了我们更大的启发，即不要轻视小事情，它一旦跟生活中的大原则，如互惠原则联系起来，照样做得了大文章。

互惠原则怎样起作用

没错，人类社会从互惠原则中得到了一项重大的竞争优势，由此，他们必须保证社会成员全都接受这一训条，遵守并信任这一原则。

我们每个人从小听人教导，也都知道不能辜负它，凡是有人敢违背它，必然都会受到社会的制裁和嘲笑。由于普通人大多讨厌一味索取、从不回报的家伙，我们往往会想方设法地避免自己被别人看成是揩油鬼、忘恩负义的家伙，或是不劳而获的懒虫。但这样一来，我们的煞费苦心又会容易遭到那些一开始就想从这种知恩图报的做法中谋取好处的人利用。

为了理解互惠原则如何被把它当成影响力武器的人利用，我们或许可以仔细来看看心理学家丹尼斯·里根（Dennis Regan）做的一项实验。

影响力研究

研究者告诉被试其参加的是所谓的“艺术鉴赏”实验，被试要跟别人一起为几幅画作的质量打分。与被试一起打分的人，我们就叫他“乔”吧，他只是假装作为被试的同伴，其实他是里根教授的助手。为了达成实验目的，研究者采用了两种不同的环境来进行。有几回，乔都主动帮了真正的被试一个小忙。在短暂的休息时间，乔离开了房间几分钟，回来时带了两罐可口可乐，一罐给被试，一罐给自己，乔说：“我问他（实验员）能不能弄瓶可乐喝喝，他说没问题，于是我就给你也带了一罐来。”另外几回，乔没有帮被试这个小忙，他只是到房间外休息了两分钟，然后就回来了。除此以外的各个方面，乔的表现都是相同的。稍后，等所有的画作都评分完毕，实验员暂时离开了房间，乔便请被试帮他一个忙。乔表示，他正帮一款新车卖抽奖彩票，要是他卖掉的彩票最多，就能得到50美元的奖金。乔请被试以每张25美分的价格买些彩票：“帮帮忙，买一张也行，当然越多越好。”

实验的主要目的之一，就是研究在上述两种情况下被试从乔手里买

的彩票数量的差异。毫无疑问，先前接受了乔好意的被试，买起彩票来更慷慨。显然，他们觉得自己欠了乔一点儿人情，因此所买彩票的数量比另一种情况下的多了一倍。尽管里根的研究只对互惠原则的运作做了很简单的阐释，但它勾勒出了该原则的若干重要特点。这里，我们不妨来仔细地分析一下，看看互惠原则是怎样被人当成牟利手段的。

互惠原则所向披靡

互惠原则能用作获取他人顺从的有效策略，原因之一在于它的效力实在是太强了。有些要求，要是没有亏欠感，本来是一定会遭到拒绝的。可靠着互惠原则，你很容易让别人点头答应。从里根实验的第二项结果中可以看到一些证据，说明互惠原则的力量有多么大，以至于一些通常情况下决定了当事人是否顺从的因素，碰到它也只有认输的份儿。

里根除了对互惠原则的说服作用感兴趣外，还调查了个人好感对顺从他人要求会产生什么样的影响。为衡量被试对乔的好感跟买乔的彩票之间有什么样的关系，里根让被试们填写了几份评分表，从这些评分便可以看出被试对乔的感觉如何。之后，里根将好感指数与购买彩票的张数做比较，发现人们表现出了明显的倾向性：越喜欢乔，买他彩票的数量也就越多。但若光是这一点，倒还算不上什么大发现，毕竟，我们大多数人都更乐意帮助自己喜欢的人。

然而，里根实验发现的有趣的地方是，一旦被试接受了乔的可乐，好感对顺从来说就完全退居二线了。对那些欠了乔人情的被试来说，不管他们喜不喜欢乔，事情都没了区别，他们觉得自己有义务报答乔，而且他们也正是这么做的。说自己不喜欢乔的人，买的彩票和说自己喜欢乔的人买的一样多。互惠原则效力之强，压倒了通常会影响顺从决策的另一个因素，即是否喜欢提出要求的人。

想想这意味着什么吧！我们通常都不怎么喜欢的人，比方说不请自

来的讨厌的推销员、不愿交往的熟人、名字都没听说过的奇怪组织的代表，只要他们在向我们提出请求之前，先对我们施个小小的恩惠，最终都能极大地提高我们依其言行行事的概率。

影响力研究

几年前，在一趟横跨美国的航班上，我被分到了连排三个座位里靠过道的那个。虽说我自己喜欢靠着过道，但还是跟另一个被分到靠墙座位的男子换了位置。他说，靠墙连坐5个小时，可能会让他发作幽闭恐惧症。他郑重地向我道了谢，而我也并没有像从小接受的教导那样认为这是一个不值一提的小忙，所以我说：“要是你，你也会跟我换座位的。”他无比赞同。

接下来的航程好极了。坐在我旁边的两个人开始聊天，发现彼此有许多共同点。过去，两人都曾住在亚特兰大附近，都是纳斯卡赛车迷，都喜欢收藏枪支，政治观点也相同。看得出来，一段友谊正在萌芽。然而，每当空乘人员提供零食、口香糖、报纸体育版的时候，坐在过道的那人总是会先递给我，有时甚至从他新伙伴的面前推给我。我还记得自己当时想：“哇，他跟谁坐得更近不重要，他跟谁有更多的共同点、爱跟谁说话也不重要，最重要的是，我才是他觉得欠了情的人。”

所以，如果你刚刚因为帮了别人一把而收获感谢，请别说“没什么大不了的”“别想太多”“我对谁都会这么做”一类的老套客气话，这样只会削弱互惠原则带来的影响力。相反，你最好保留（应得的）影响力，不妨这么说：“如果我们的立场互换，我相信你也会这么做。”由此带来的好处应该相当可观。

此外，小孩子在上学之前就已经渐渐理解了收到东西后有义务做出回应并回馈。研究表明，与剥夺、拷打、折磨等强制刑讯手段相比，能诱发互惠的善意等“软性”方法能带来更好的效果。

READER'S REPORT | 读者报告

致西奥迪尼先生：

我是纽约州罗切斯特市一家公司的秘书，平常总是白天上班，但有一天晚上，我在公司待到很晚，处理一些未做完的重要工作。等干完工作，开车驶出停车位的时候，我的车碰到冰碴打了滑，卡在了路旁的阴沟里。当时已经很晚了，天气很冷，四周黑乎乎的，我办公室的所有同事都走了。但另一个部门的员工此时却跑过来，帮我把车弄了出来。

我是从事人事工作的，大概两个多星期以后，我得知这名员工被“打了报告”，据说他严重违反了公司政策。我并不真正了解这个人的品性，可我还是把这件事揽了下来，替他去找了公司总裁。直到今天，尽管质疑此人品性的人越来越多，我仍然觉得自己欠了他的情，愿意为他撑腰。

纽约的这位职业女性：

跟里根的实验一样，你决定帮助该员工的行为，跟这名员工的个人道德似乎不怎么相关，而只在于一点简单的事实：那名员工也帮过你的忙。

各种类型的组织都学会了这一套：利用一份小礼物，促使目标答应本来会拒绝的请求。研究者发现，在邮寄调查问卷的信封里附上一份现金礼物，如一枚银币或是一张5美元的支票，能大大提高问卷的完成率。反之，答应事后奖励同等金额的做法，效果就差多了。事实上，经研究表明，在邮寄保险问卷的同时奉送5美元支票作为礼物，效果跟完成问卷后寄送50美元的报酬同样好。类似地，餐馆的服务员也早就知晓，在向顾客送上账单的同时奉送一颗糖果或口香糖，能大幅提高小费的金额；而在一家国际游客经常光顾的餐厅里，不管客人来自哪个国家，互惠原则都将发挥作用。

影响力研究

我的同事史蒂文·J.马丁（Steven J. Martin）和海伦·曼金（Helen Mankin）进行过一项小型研究，他们在巴西和哥伦比亚的几家快餐厅里展示了率先给予的力量。在半数的快餐厅，成年顾客的孩子在离开时收到了一个气球。在另外半数的快餐厅里，孩子们一进门就得到了一个气球。结果发现，先给气球的快餐厅，其家庭总账单金额增加了25%。值得注意的是，购买咖啡的金额增加了20%——而孩子们不太可能会点咖啡。为什么会这样呢？至少我可以证明，给我孩子的礼物，就等于是给我的礼物。

总体而言，企业经营者发现：接受礼物后，客户会愿意购买本来不愿买的产品或服务。

同理，和前文中邮寄调查问卷时附赠5美元的策略一样，一项关于酒店的研究也证实了这种“事先施惠”的威力。酒店在房间里放一张倡导重复使用毛巾的卡片。卡片上的附带说明分为两种，一种说酒店已经以客人的名义向环保组织捐了款，另一种则说客人重复使用毛巾后会以其

名义向环保组织捐款。统计表明，事前捐款的卡片，其劝说效力比事后捐款的卡片高得多。

EBOX | 线上影响力



作者点评：2011年，星巴克为了庆祝成立40周年，向顾客提供了免费的电子代金券。为了增强顾客与礼物相关的亏欠感，任何接受代金券的顾客都必须在社交媒体上明确感谢该公司。

值得注意的是，免费的代金券利用了互惠原则，而后来发放的代金券越来越少则利用了稀缺原则——我们将在第6章单独考察其力量。

政治

政治是互惠原则发挥威力的另一个舞台。在各个级别上，我们都能看到这套手法。

在高层，民选代表“互投赞成票”，互施小恩小惠，这使得同床异梦的伙伴比比皆是。民选代表对一项法案或措施投出跟自己一贯主张完全相悖的支持票，大多都是为了回报法案发起人的人情。林登·约翰逊刚刚当上总统的时候，他提出的项目在国会里总是轻轻松松就通过了，政治分析家们不免对此感到惊讶，就连照理说该强烈反对这些提案的议员也投了赞成票。分析家们（如罗伯特·卡诺在他极具影响力的约翰逊传记当中）做了一番认真的研究后发现，与其说这是因为约翰逊在政治上特别善于钻营取巧，还不如说是因为他长年在众议院和参议院里摸爬滚打，帮了其他议员许多忙。当上总统以后，议员们纷纷偿还从前欠下的人情债，这才使他得以在短短的时间内就通过了大量的立法提案。有意思的是，其后的一些总统（比如卡特、克林顿、奥巴马）碰到的问题，也是这个原因在作祟。他们是从国会山外头赤手空拳打进白宫当总统的。竞选期间，他们也狠狠地利用了这一点来搞宣传，自己是华盛顿圈子外头的人，不欠任何人的人情债。这些总统就任之初难于让立法提案在议会通过，说不定正是因为——没有人欠他们的人情债。

在另一个层面上，我们可以看到：一方面，企业和个人愿意向司法和立法官员赠送礼物，施加恩惠；另一方面，国家又制定了一系列法律，禁止官员接受此类礼物和恩惠。这充分说明人们意识到了互惠原则的力量。就算是合法的政治捐款，表面上说是支持自己喜欢的候选人，其实大多还是为了囤积人情债。碰上重大选举，看看有多少公司和组织是同时给民主党和共和党的竞选活动捐款的，你就能意识到其中的奥妙了。要是你心存怀疑，非要看看政治捐赠者换回了什么好处，那不妨听听商人罗杰·塔马兹（Roger Tamraz）在国会召开的选举活动经费改革听证会上说的话。人家问他捐款30万美元后是否收到了良好回报，他笑着回答：“我想下一次我会捐60万美元。”这番话说得是足够厚颜无耻了。

这种坦率在政治上可不多见。大多数时候，不管是送礼的也好，收礼的也好，都会齐声说道：“政治献金、免费旅行和橄榄球超级碗的门票，都不足以撼动政府官员‘严肃、正直’的秉性。”一家游说组织的领导

人坚持认为，没有必要担心，因为“这些政府官员是聪明、成熟、老练的职业人士，长年的训练让他们目光敏锐，时刻保持批判态度，而且十分警觉”。当然，政客们肯定是同意这种说法的。我们经常听到他们说，会影响其他人的那种亏欠感，对他们是全无作用的。我所在的这个州，当有位议员谈到自己是否对送礼的人有什么义务时，他斩钉截铁、不容置疑地说：“他们送了礼物，可我对他们的态度跟对其他人是完全一样的：别想从我这儿捞到什么好处。”

身为科学家，听到这样的话，我忍不住笑了。亏欠感对这些政客到底有没有作用，我们这些严肃、正直、谨慎的科学家再清楚不过了。之所以这么说，原因之一在于，我们这些“聪明、成熟、老练，正值科学事业巅峰”的家伙发现，自己跟所有人一样，也会很容易就被互惠原则左右。举个例子，有一种治疗心脏病的药叫钙通道阻滞剂，围绕它的安全性曾引发过许多争议。有人做了研究，对这种药，凡是持支持态度，并发表了正面文章的科学家，100%都接受过医药公司从前给的好处，如免费旅行、研究资金或是工作机会；而对这种药持批评态度的科学家里，只有37%的人曾接受过好处。身为科学家，“长年的训练让他们目光敏锐，时刻保持批判态度，而且十分警觉”，可要是连他们都难免为长年的恩惠所左右，政客们就更不在话下了。

在这一点上，我们的看法完全正确。例如，2002年美国中期选举期间，美联社记者通过观察发现，在6大主要政治议题上，收了特殊利益集团金钱的国会议员，对有利于这些团体的主张投赞成票的概率比一般情况下高了7倍之多。这样一来，这些团体赚回本钱的概率是83%。不管是民选官员还是上级指派的官员，总觉得自己不是普通人，无须遵守规章条例的限制，如他们停车从不按规矩，诸如此类。但如果他们纵容自己在互惠原则上也这么妄想，那么这不光是可笑的，更是不负责任的。^⑥

在谈判中，一方的主动帮助能让另一方更愿意进行合作以回报善

意，哪怕这样做会违背后者的经济利益。

影响力研究

国际谈判史上也有许多利用互惠原则将潜在危险冲突转化为和平解决方案的例子。其中一项给予一妥协式的协议，或许曾拯救了全世界，只是它出于政治原因，无法获得应有的荣誉。1962年10月22日，美国和苏联的“冷战”几乎就要变成“热战”。在电视讲话中，肯尼迪总统称，俄罗斯核弹已被证实在重重掩护下被运到了古巴并对准美国。为要求苏联领导人赫鲁晓夫下令卸载导弹，肯尼迪命令美国海军舰队封锁海面，禁止更多的导弹进入古巴，直至已安装的导弹得到移除。赫鲁晓夫则回应说，苏联方面前往古巴的船只将无视这种“公然的海盗行为”；此外，苏联会将任何试图实施封锁的行为视为引发战争的侵略行为。这可不是普普通通的战争，而是一场几乎能毁灭1/3人类的核战争。在长达13天的时间里，两位领导人恶狠狠地彼此对视，而全世界的人民则提心吊胆地看着他们，指望其中一位让步。最终，此事以赫鲁晓夫屈服于肯尼迪不屈不挠的谈判风格、同意将导弹带回苏联而告终。至少按我一直以来听说的故事，古巴导弹危机是这么结束的。

但最近解密的文件提供了一个完全不同的说法。肯尼迪“获胜”不是由于他固执的谈判姿态，而是因为他愿意从土耳其和意大利撤除美国的木星导弹，以换取赫鲁晓夫从古巴撤除导弹的条件。考虑到自己的政治声望，肯尼迪将导弹交易保密作为最终协议的条件之一，因为他不希望被人看到美国对苏联做了任何让步。具有讽刺意味也令人遗憾的是，哪怕直到多年后的今天，互惠交换（这是“拯救世界”的真正因

素)的力量仍未被认定为化解此次冲突的重要因素,功劳反而落到了另一个很可能毁灭世界的因素上——不愿妥协(见图2-1)。[\[7\]](#)



图2-1 “在卡斯特罗峡谷的让步”

这幅当时的政治漫画描述了人们对古巴导弹危机结束的普遍看法——面对肯尼迪总统强硬的态度,赫鲁晓夫屈服了。可事实恰好相反。这个迫在眉睫的世界性核危机得以解除,靠的其实是一项重大的妥协,即双方对等地撤除核导弹。

社会心理学家李·罗斯(Lee Ross)讲过一个两兄弟的故事,也证明了给予—妥协式谈判相较于绝不让步式谈判的优势。这两兄弟在加拿大开着一家大型折扣宠物商品店。两人必须在其位于多个城市的产品分销点就仓储空间问题进行谈判。其中一人说:“我很清楚每个城市合理的仓储价格,唯一的策略就是报出并在谈判中永不偏离这个价格。而讨价还价的工作全都是我兄弟来做。”

免费样品不免费

当然,生意场上也能见识互惠原则的威力。虽说例子不胜枚举,但在这里,我们还是来看一下人人都熟悉的一两种。赠送免费样品这种营

销技巧历史悠久，也很管用。通常的做法是向潜在客户送上少量的相关产品，看看他们是否喜欢。不必说，制造商是想借此举让公众知道自己产品的质量，这种愿望合情合理。然而，免费样品的真正妙处在于，它同时也是一份礼物，能把互惠原则应用起来。推销的人提供免费样品，表面上不过是为了让消费者知晓他们的商品，暗中却是把礼物天然具备的亏欠感给释放了出来，这完全是借力打力、四两拨千斤的柔道手法。

在南加利福尼亚州的一家糖果店里，研究者考察了顾客的购买模式。进店时，有些顾客会收到一块免费的糖果。收到赠品后，顾客下单购买的可能性增加了42%。当然，购买率的上升，也可能不完全是因为互惠原则的拉动。也许这些顾客只是单纯地喜欢自己尝到的东西，所以买了更多同类的糖果。但仔细观察后就会知道，这种解释站不住脚。接受了免费糖果的顾客并不是购买了更多的品尝过的糖果，而是买了更多其他类型的糖果。看起来，就算他们并不特别喜欢店家赠送的糖果，也仍然觉得自己有义务购买一些东西作为回报。

超市是赠送免费样品的绝佳场合，消费者经常在那儿得到某种产品的少量试用装（见图2-2）。服务人员总是会微笑着递上样品，好多人都觉得光是还回牙签或杯子就走开太过分了。于是，他们购买了一些产品，哪怕自己并不是十分喜欢。据来自零售巨头开市客（Costco）的销售数字，其他类型的产品，如包括啤酒、奶酪、冷冻比萨、口红，也会因为免费样品而出现销量的大幅提升。万斯·帕卡德（Vance Packard）在经典作品《隐形的说客》（*The Hidden Persuaders*）中介绍了这一营销手法特别管用的一种变体形式：印第安纳州的超市经营者把奶酪摆在货柜外面，让消费者自己切一小块免费品尝。如此一来，有一天他们在短短几个小时内就卖掉了四五百千克的奶酪。还有一种与互惠原则吻合的购买模式是，超市购物者获得了专门购买某类商品的惊喜礼品券后，随之从商店购买了更多其他的商品，使得总购买量增加了10%。



图2-2 免费发放玉米脆片的工作人员

一些食品制造商会人们在进入超市之前就派发免费样品供人们品尝。

EBOX | 线上影响力



作者点评：观察上图后，我们可以将免费样品有效的原因归纳为两点：一是免费的章节让顾客有能力就是否购买整本书做出更明智的决定；二是作为赠品，免费章节可能会让顾客觉得自己有义务购买完整版。我碰巧认识本书的作者，我问他做这则广告看中的是上述哪个原因，他说完全是出于第一个原因。我知道他是个诚实的人，但作为心理学家，我也知道人们相信的往往是自己乐意相信的事，所以我并

不完全相信他的说法。

影响力研究

安利是一家制造家用和个人护理产品并在全球范围内上门推销的公司，他们用的是另一种免费试用策略。公司早先的办公室还只是在地下室里，但很快，公司每年的销售额就达到了15亿美元，这多亏了一种名叫“臭虫”（BUG）的免费试用手法。“臭虫”由一系列的安利产品组成，包括若干瓶家具抛光剂、清洁剂或洗发水，要不就是喷雾式除臭剂、杀虫剂或玻璃清洁剂。销售员用一种特别设计的托盘或塑料袋把它们提到消费者的家里。安利公司的机密《操作手册》指导销售员说：“把‘臭虫’留在消费者那里，‘一天、两天甚至三天，全部免费使用，也不让消费者承担任何义务。只要告诉他，你希望他试试这些产品。没人能拒绝这种请求的’。”等试用期结束，安利的客户代表就会回来，顺利地拿到客户希望购买的产品订单。由于消费者不大可能在这么短的时间里把“臭虫”组合产品中的任意一瓶用完，于是销售员又可以将剩下的部分带到对门或隔壁的下一位潜在客户家里，整个过程再从头来上一遍。安利的不少客户代表通常会在自己负责的销售区域里同时使用若干“臭虫”套装。

现在，你我当然知道接受并使用了“臭虫”产品套装的消费者中了互惠原则的套。对自己试用并部分消耗了的产品，许多客户都觉得有义务订购它，当然了，安利公司也早就知道会这样。就算是在安利这样一个有着卓越绩效的企业，“臭虫”手法也掀起了一场大轰动。各州的分销商向母公司提交的报告记下了它的神奇功效，下面这两段话分别来自伊利诺伊州和马萨诸塞州的分销商：

真是不可思议！我们从来没有见过这么令人兴奋的事儿。产品正以飞快的速度卖出去，而我们才刚刚开始呢……地区经销商一采用“臭虫”模式，我们的销售量就会出现飞速增长。这真是我们最非凡的一套销售理念呀！……平均而言，销售员取回“臭虫”的时候，消费者总会购买总量的一半……一句话，太了不起了！在我们整个组织里，还从没见过这样的情况呢。

“臭虫”的惊人威力，让安利分销商有点莫名其妙，虽说他们非常高兴，但的确是不知所以。当然了，你我现在再也不会这样摸不着头脑了。

在无关金钱也无关商业交易的单纯人际关系中，很多情况也受互惠原则的控制。在下面这个故事里的妇女，因为拒绝接受一份礼物，躲开了随之而来的强烈亏欠感，得以侥幸逃生。

影响力研究

1978年11月，在圭亚那的琼斯镇，邪教组织“人民圣殿教”的头领吉姆·琼斯（Jim Jones）要所有信众集体自杀。绝大多数人在顺从地喝下了有毒的饮料后就这么死了。然而，有个叫黛安·路易（Diane Louie）的信众却拒绝服从琼斯的命令，她逃出了琼斯镇，在丛林里躲了起来。她说自己之所以会这么做，是因为先前在有困难的时候拒绝了教主琼斯的特殊照顾。当时她生病了，琼斯派人送来食物，可她没接，因为“我知道一旦我接受了他给的这些好处，他就完全控制了我，我什么也不想欠他的”。

或许，琼斯错在让路易女士把《旧约》学得太好了，尤其是《出埃及记》：“不可受贿赂，因为贿赂能叫明眼人变瞎了，又能颠倒义人的

话。”

量身定制的礼物

尽管互惠原则的力量已不容小觑，但如果提供的礼物是根据接受方当前的需求、偏好量身定制的，即个性化的，那么它还能产生更大的力量。

一位朋友曾对我讲过自己怎样利用个性化礼物的手法让一个付款慢得出名的客户（业内人人都知道他一般要6个月才付款）加快结账速度。不久前，她在寄出发票的同时会附上一份小礼物，可能是一包高质量文具、一小盒巧克力或一张星巴克咖啡优惠卡。她发现，客户结账付款的时间缩短了一半。更晚些时候，她又附上了一张来自当地艺术博物馆的个性化明信片，明信片上印着一件现代艺术品——它正是客户自己会收藏的艺术品品类。这位朋友拍着胸口说，她的发票现在几乎可以立刻兑付。她的同行深感惊讶，还想知道她是怎么做到的。她说，自己一直保密到现在。

除了按照接受方的爱好定制礼物，根据对方当前的需求进行定制，也能放大礼物的力量。研究者在一家快餐店所做的研究，解释了这类量身定制礼物的效力。在顾客进门时，服务员会为一些顾客送上热情的问候，而为另一些顾客送上热情的问候和一个钥匙链。跟没有得到钥匙链的顾客相比，收到钥匙链的顾客的花销提高了12%，这恰恰跟一般性的互惠原则相吻合。还有一些顾客得到了热情的问候和一小杯酸奶。虽然酸奶的零售价跟钥匙链相同，但它让顾客的消费额提高了24%。为何如此呢？因为顾客是为食物而来，面对一份满足自己需求的礼物，人们会感到更为强烈的回报义务。

钥匙链与酸奶带来的购买率的变化情况和一项超市研究颇为相似，后者向进店顾客赠送非食品礼物（钥匙链）或与食品相关的礼物（品客

薯片），使整体购买量分别增加了28%和60%。而一家连锁酒店的完美客户体验项目，在提高客户满意度方面的结果却令人失望。按需定制礼物的效力并不仅适用于商业环境，在人际关系中同样如此，只有当“礼物”符合接受者当下的需求时，它才能带来更大的关系满意度。

不久前，我的同事布赖恩·埃亨（Brian Ahearn）发来一篇销售杂志上刊登的文章，该文章介绍了一家全球连锁酒店的高管在审核公司耗资巨大的“完美客户体验”项目之后所感受到的震惊。满意度和未来忠诚度最高的并非那些在住店时没有遇到问题的客人，而是那些碰到了服务上的问题、酒店员工立刻为其解决的客人。要理解为什么会出现这样的情况，方法很多。例如，或许是客人知道酒店能有效补救错误之后会更加确信，万一自己将来的住宿碰到类似的局面，也会得到这样的待遇。我也深信还有一个因素同样在发挥作用：客人很可能是把补救措施看成了酒店竭尽全力为其提供的“特殊个性化服务”。根据互惠原则，酒店理应得到一些特殊的回报，比如更高的评价和忠诚度。

影响力研究

在商务会议上演讲时，我经常提及这位酒店高管的意外发现，以及我对此事的解释。有一次，我演讲所在的度假酒店总经理正好也是听众之一，他讲到当天发生的一件事，正好证实了我的互惠解释。一位客人想跟自己的两个小孩打网球，但度假酒店里的儿童球拍已经有其他人在用了，客人感到很失望。出乎意料的是，20分钟后，一副新球拍就送到了客人手里，这是总经理让一名员工开车去当地的体育用品商店买的。事后，这位客人来到总经理的办公室说：“因为你为我做的这件事，我刚刚预订了下个周末再带全家到这里度假的日程。”

如果度假酒店一开始就额外预备了两副儿童球拍以确保客人获得“完美体验”，那客人就不会把它视为一份重要的礼物或服务，认为这值得自己付出特别的感激和忠诚，从而为酒店带来更多的业务。这是不是很有趣呢？实际上，在那位客人回想度假体验时，这些球拍说不定根本就没机会出现。

我相信，正是因为“纠正失误”是一种量身定制的行为，它才代表了个性化的礼物或服务。这一特点撬动了运用互惠原则的杠杆，使我们得以理解为什么失误反而带来了更高的满意度和忠诚度。简而言之，没碰到问题的人，或许不如“碰到问题但又得到了解决”的人感觉那么好。

互惠原则适用于强加的恩惠

早些时候我们指出，互惠原则的威力大到了这样的地步：其他人，不管有多奇怪、多讨厌、多不受欢迎，只要先给我们点小恩小惠，就能提高我们照着其要求去做的概率。然而，除了威力大，该原则还有另一个特点，它居然允许这种情况的发生：一个人靠着硬塞给我们一些好处，就能触发我们的亏欠感。回想一下，互惠原则只是说，当别人帮了我们的忙，给了我们好处，我们就应当回报他；可它并没有说，我们滋生出的偿还亏欠感，一定来自我们主动请人家帮忙，要人家给好处。举个例子，美国伤残退伍军人组织报告指出，只寄出一封请求捐款的信，回应率大概是18%；可要是在信里附上一份小礼物，如带不干胶的个性化地址标签，成功率就能翻上近乎一倍，达到35%。当然了，倘若要求是我们先提出来的，我们偿还的义务感会更强烈，但这里，我想要说明的关键点在于：要让我们产生亏欠感，不一定需要这样，哪怕是被人硬塞了些好处，我们也会产生亏欠感。

稍微回想一下互惠原则的社会目的，我们便能看出这其中的原因。原则确立起来，是为了推动个人之间互惠关系的发展，如此一来，首先发起这种关系、第一个表示善意的人就不必担心会有损失。倘若原则是

为这样的目的服务的，那么不请自来的恩惠就必然具备创造亏欠感的能力。再回想一下，互惠关系给孕育它的文化带来了意想不到的优越性，因此，文化里也就存在着强大的压力，确保原则为其最初的目的服务。这也就难怪著名的法国人类学家马塞尔·莫斯（Marcel Mauss）在描述人类文化围绕赠礼过程产生的社会压力时说：“人有送礼的义务、接受的义务，更有偿还的义务。”

尽管偿还义务构成了互惠原则的实质，可使原则那么容易遭到利用的，却在于接受的义务。有了接受的义务，欠谁的人情就不归我们选择了，反过来我们还会落到对方的手里。让我们重新回顾一下先前举的几个例子，看看这个过程是怎么运作的。

在里根的研究中我们发现，乔通过给被试提供恩惠使得后者买他彩票的数量翻了一倍，但这份恩惠并不是后者要求的。乔自己主动离开，回来时带了两瓶可乐，一瓶给自己，一瓶给被试。没有任何一个被试拒绝了他递过来的可乐。很容易看出，拒绝乔的好意实在是太尴尬了。乔已经花钱买了饮料，在该情形中，一瓶饮料是很恰当的善意举动，因为乔也给自己买了一瓶，拒绝乔这么善解人意的举动肯定不礼貌。然而，等乔说明自己卖彩票的愿望时，接受可乐这一行为所带来的亏欠情绪就变得很清晰了。

注意，这里存在一种重要的不对称性，乔掌握了所有真正的选择。他选择了最初施恩的形式，又选择了回报这种恩惠的形式。当然，你可以说，乔的这两项提议，被试都可以选择拒绝，但这样的选择未免太艰难了。拒绝乔两项提议中的任何一项，都要求被试跟文化里天然有利于互惠的力量对着干。

许多组织都意识到人们会因为出乎意料的赠礼而产生亏欠感。我们每个人，是不是都收过许多慈善机构发来的信件，信件里还会附上一些小礼物给我们，如个性化地址标签、贺卡、钥匙环，但同时还会附着一张字条，要我们捐款？光是去年，我就收到了5封，两封来自伤残退伍

军人团体，剩下的来自教会、学校和医院。每一封信里附带的信息都有一个共同点：信封里的东西可以看成是该组织送的礼物；而我希望捐助的金钱，则不应看成是捐款，而应算作是一种还礼。正如一家教会组织寄来的信中所说，我收到的贺卡并不需要直接付钱，而是旨在“激发您的善意”。就算不看税收上的好处，我们也应当明白，把贺卡当成礼物赠送，而不当成商品贩卖，对组织来说都是很有好处的。社会上有着强大的文化压力，让你收到礼物后要还礼，哪怕这份礼物你并不想要；可社会并没有强迫人们购买不想要的商品。

关于主动施惠，有一个真实的例子：2013年，波士顿城市公交司机大罢工之际，优步租下了公交车，并向该市所有公立学校提供免费服务。结果表明，这一举措明显增加了优步在波士顿的客流量。

READER'S REPORT | 读者报告

致西奥迪尼先生：

去年，在感恩节放假回家的路上，我的车胎爆了，之后发生的事让我亲身感受到了互惠带来的压力。一位穿着护士制服的女司机停下车来，并主动送我回家。我告诉了她好几次，我家在40公里以外，而且跟她前进的方向相反。但她仍坚持帮我的忙，而且无论如何也不收钱。她拒绝收我的钱，让我产生了您在《影响力》一书中讨论的那种不舒服的令人不安的感觉。

这件小小的意外发生之后，连我的父母也焦虑起来。互惠原则，再加上没有回报善意造成的不安，令我全家人都有点神经衰弱了。我们一直想要找到她，好给她送去鲜花或礼物，却遍寻不着。要是我们能找到她，我相信，这位护士要什么我们都会给她的。由于找不到其他办法可以

缓解亏欠感，我母亲最终祭出了她的撒手锏。在感恩节晚餐上，她带领我们祈祷，请天主赐福这位女士。

这位男大学生：

这个故事除了表明他人主动提供帮助能调动互惠原则之外，还透露出另外一点有必要了解的事实。伴随互惠原则出现的亏欠感，并不仅限于最初提供和接受援助的个体，该个体所归属的群体中的其他成员也会受到影响。不光你的家人会因为得到的帮助而产生亏欠感，新的研究揭示，要是你和家人能帮助这名护士的某个家庭成员，亦可了却这笔人情债。另一项研究表明，这种基于群体的“互惠”也可能扩展到虐待上。如果我们遭到另一个群体成员的伤害，但又没法伤害那个人，我们便很有可能通过虐待该群体的其他成员来进行报复。

互惠原则可触发不对等交换

互惠原则还有另一个特点，也很容易遭人利用。尽管它的确立是为了促进伙伴之间的平等交流，但也可以用来实现完全不平等的结果，这可真够自相矛盾的。互惠原则要求，某一种行为需要以与其类似的行为加以回报。人家施恩于你，你必以恩情报之，不理不睬是不行的，以怨报德更加不可以。但这里面也有很大的灵活性，别人最初给予的小小恩惠，能够让当事人产生亏欠感，最终以大得多的恩惠回报。正如我们所见，在互惠原则当中，最初让他人产生亏欠感的行为，以及缓解亏欠感的回报行为，都可以由最初的发起者来选择，这样一来，那些打定主意要利用互惠原则的人就能轻易地操纵我们，让我们完成一种不公平的交

换。

让我们再次回到里根实验中来。请记住，在这项研究中，乔给了被试一瓶可口可乐作为最初的赠礼，稍后，又要求所有被试以每张25美分的价格购买他的抽奖券。有一点我之前没提过，进行这项研究的时间是在20世纪60年代末，那时一瓶可口可乐的价格是10美分。平均而言，喝了乔10美分可乐的被试买了他两张抽奖券，也有人一买就买了7张。就算只看平均数，我们也可以判断出乔的买卖做得相当划算：他得到的回报是最初投资的整整5倍！尽管在乔的例子中，整整5倍的回报也就是50美分罢了。

互惠原则真的能影响换得恩惠的大小吗？只要环境合适，的确可以。这里举个我学生的例子吧。每当她回想起这件事，总是后悔不已。

影响力研究

大约一年前的一天，我的车发动不了了。我正束手无策地坐着的时候，停车场有个人走过来，最终帮我把车发动了起来。我说“谢谢”，他说“不客气”。他离开的时候，我说：“要是遇到什么我能帮忙的事儿，请随时开口。”

过了一个来月，他来敲我的门，要求借用一下我的车，两个小时就够了，他自己的车送去店里修理了。虽然我觉得欠了他的情，但又不太好做决定，因为我的车还相当新，他又是个特别年轻的小伙子，但不管怎么说，当时我还是把车借给了他。结果呢？我的车自然是毁在了他的手里。后来，我才晓得他没成年，也没有保险。

一个聪明的年轻姑娘，怎么会因为陌生的小伙子在一个之月前帮了她小小的忙，就答应把自己的新车借出去呢？这是怎么发生的呢？更概

括地说，为什么最初的小小善意往往会刺激人们以大得多的恩惠回报？亏欠感让人觉得很不舒服，是一个很重要的原因。我们大多数人都会觉得亏欠别人是很不愉快的，这种感觉沉沉地压在我们身上，要求我们尽快将之除去。不难看出这种感觉的起源在哪里，由于互惠在人类社会体制下极其重要，一旦欠了人情债，我们就会条件反射般地感到不舒服。要是我们忽视了回报他人首发善意的需求，互惠的循环就会终止，我们的恩人将来也就不大可能再做这种好事了，这不符合社会整体的最佳利益。所以，根据我们从小接受的教导，只要亏欠了别人，我们在情绪上就会感到烦躁不安。单从这个原因看，光是为了卸下心理上的债务包袱，我们说不定就会乐意还以比先前所受更大的恩惠。日本有句谚语一针见血地指出：“再没有什么比免费的东西更昂贵的了。”

还有另外一个原因：违背互惠原则，接受而不试图回报他人善举的人，是不受社会群体欢迎的。当然，要是客观条件或能力限制使得他无法偿还恩情，例外也是可以的。然而，大多数情况下，不照着互惠原则做事的人，人们都是普遍嫌恶的。谁都不愿被贴上“揩油鬼”“忘恩负义之徒”这样的讨厌标签。为了躲开这样的标签，人们有时也会答应不平等的交换。

内心的不舒服，加上丢脸的可能性，足以让人产生沉重的心理负担。从这个角度来看，我们经常打着互惠的旗号，给出比自己获得的还要多的东西，也就不足为奇了。此外，以下的情况也不奇怪：要是我们觉得无法回报，哪怕自己真的需要，也会尽量避免找人帮忙，心理负担说不定比物质损失更让人难以忍受（见图2-3）。

我给小费一贯不大方，所以，你就把最粗鲁的服务员派过来吧，这样我也不觉得那么愧疚了……



图2-3 心怀愧疚的交换

就算是最小气的人也会感受到互惠原则造成的压力。餐馆服务员也会用这一原则来多收小费。研究发现，在送上账单的时候，服务员给就餐者送去一块糖果，他们得到的小费就能提高3.3%；要是他们给每位顾客赠送两颗糖，小费便可增加14%。

最后，还有一类损失可能会使人们婉拒某种礼物和好处。女性经常提起，要是男人给她们赠送了昂贵的礼物，或是带她们享受了奢侈的夜生活，她们就会有一种欠了情的不舒服感觉，并想要回报这男人的恩惠。就算是买一杯饮料这样微不足道的小事，也会给她们带来亏欠感。我班上有个学生在一篇论文里直截了当地指出了这一点：“有了一些恼火的经历之后，我总算学乖了。我再也不会让任何一个在夜总会遇到的男人给我买饮料了，因为我可不想我们中有哪一方认为我该以肉体偿还这份人情债。”研究表明，她的看法其实是有现实基础的。要是姑娘不自己买单，而是让男人给她买饮料，那么，不管是男的还是女的，都会立刻判断这个女人很有可能会跟这个男人上床。

互惠原则适用于大多数关系，然而，在家庭或稳定的友谊这类长期关系中，纯粹的互惠交换并无必要，也不受欢迎。在这类“共有”式

关系中，人们互惠交换的是一种意愿：只要你需要，我就帮你的忙。基于这种互惠形式，只要双方都遵照一般性原则，算计谁给的多、谁给的少是毫无必要的。

我们当然讨厌接受而不回报他人善举的人，而有趣的是，跨文化研究表明，要是有人反方向破坏互惠原则，即施予了恩惠却不让接受的人有机会回报，也是不受人喜欢的。有充分的证据表明，人们常常不去寻求帮助，以避免产生社会亏欠感。一项持续了10年的研究考察了许多人都面临的一个困境：搬家时，是请朋友和家人帮忙，还是把整个任务交给搬家公司。研究发现，人们避免向自己认识的人求助，往往不是因为担心这些非专业人士会破坏宝贵的财产，而是担心这种帮助会让自己产生“负债感”。

READER'S REPORT | 读者报告

致西奥迪尼先生：

不久前，我们搬到了澳大利亚，5岁的女儿一直在努力适应新文化并寻找新朋友。最近，她和我妻子在附近散步时，试着在邻居的信箱里留下“礼物”。这些礼物其实只是简单的蜡笔涂鸦，对折后粘成一封信的样子。我认为这么做基本上没什么害处，只不过挺担心人们把它看成是讨厌的骚扰，害怕全家人会被叫成“在信箱里乱丢垃圾的幽灵”。可没过多久，有趣的事情发生了。我们开始在家里的邮箱里发现别人写给我女儿的贺卡——而且是本地品牌的贺卡，每张价格在3到5美元不等。后来，信箱里又出现了一包包的糖果和小玩具。要不是读过你的书，我真没法理解这件事。但互惠的力量令人难以置信，现在，我女儿每天都和一群朋友在马路对面的公园玩耍。

这位从美国到澳大利亚的移民：

我喜欢这篇文章，因为它巩固了我们已经谈到的互惠原则的两个特征：它不仅能激发不对等的交换，还可发起持久存在的社会安排。更重要的是，就连小孩子也把它认为是实现此种安排的一种方式。

互惠式让步

用互惠原则使他人依从要求行事，还有第二种办法。这比直接给人恩惠再索取回报的方式更微妙，但从某些方面来看，它也更为有效。几年前的一次亲身经历给了我第一手证据，说明了这种驯服技巧是多么管用。

影响力研究

我在街上走着，碰到了一个十一二岁的男孩子。他做了自我介绍，并说童子军一年一度的马戏表演就要在这个星期六晚上举行了，而他正在卖门票。他问我是否愿意购买5美元的门票。我可不想把大好的周末时间耗在看童子军马戏表演上，于是婉言拒绝了。“好吧，”他说，“要是你不想买门票，买我们几块巧克力如何？一块才1美元。”我买了两块，但立刻意识到发生了点怪事：第一，我对巧克力没什么兴趣；第二，我喜欢钞票；第三，我正拿着两块巧克力傻站在那里；第四，男孩拿着我的钞票走开了。

为了弄明白到底是怎么回事，我回到办公室，召集研究助理们开了个会。在讨论当时的情况时，我们逐渐意识到，在我照着小男孩的要求买巧克力的过程中，互惠原则在

其中发挥了怎样的作用。互惠原则的一般性原则指出，要是有人以某种方式对我们行事，我们理当对他还以类似的行为。我们已经看到，这一原则造成了两个后果：第一，面对接受的善意，我们感到有义务偿还；第二，倘若有人对我们让了步，我们便觉得有义务也退让一步。经过思考，我的研究小组意识到，小童子军对我来的就是这一手。他要我购买1美元的巧克力，是以让步的形式提出的，在我眼里，这是他对头一次请求，即要我购买5美元的门票的让步。如果要遵守互惠原则的规范，我也必须有所让步。正如我们所见，我的确让了步：他从大请求退让到小请求，我则从不顺从变成了顺从，尽管我对他卖的门票和巧克力都没什么兴趣。

这是一个如何将影响力武器植入顺从要求的典型例子。我“被”说动购买了某样东西，不是因为我对这样东西产生了什么好感，而是因为购买请求的设计方法调动出了互惠原则的力量。我不喜欢巧克力，不要紧，小童子军对我让了步，我的磁带立刻播放起来，致使我自己也让了步。当然了，以让步来互惠的做法尚未强大到在所有环境下对所有人都管用，本书提到的任何一种影响力武器都没有这种功效。然而，在我和小童子军的交换当中，这一做法足够让我莫名其妙地出了高价，买了两根自己并不想要的巧克力。

为什么我会觉得人家让了步，我就也应该让步呢？答案仍然是，这种倾向对社会有好处。成员为实现共同的目标而一起努力，对任何社会群体来说，这都符合其利益。可在很多社会互动当中，参与者往往一开始就会提出一些其他成员无法接受的要求或条件。因此，为了完成对社会有益的合作，整个社会都必须设法解决这些互不相容的初始欲望。这就要借助有助于双方达成妥协的程序，而互相让步是这类程序里十分重要的一种。

互惠原则通过两条途径来实现互相让步。第一条很明显：它迫使接受了对方让步的人以同样的方式作出回应；第二条尽管不那么明显，但更为关键：由于接受了让步的人有回报的义务，于是人们都乐意率先让步，从而启动有益的交换过程。归根结底，要是不存在回报让步的社会义务，谁乐意头一个牺牲利益呢？你有可能放弃了某种东西，却得不到任何回报。不过，有了互惠原则的影响，我们就可以安安心心地率先向合作伙伴让步，因为他有义务也牺牲自己的一些利益，以此回报我们的善意。

拒绝-后撤策略

由于互惠原则决定了妥协过程，因此你可以把率先让步当成一种高度有效的顺从技巧来使用。这种技巧很简单，一般叫作“拒绝-后撤策略”，也叫“留面子法”。假设你想让我答应你的某个请求，为了提高获胜的概率，你可以先向我提一个大些的要求，对这样的要求，我保准是要拒绝的，等我真的拒绝以后，你再提出一个稍小的要求，而这个要求才是你真正想提的。倘若你的要求设置巧妙，我会把你的第二个要求看成是一种对我的让步，并有可能感到自己这一边也该让让步，于是最终顺从你的第二个要求。

这就是小童子军让我买他巧克力的做法吗？他放弃了5美元的要求，重新提出1美元的要求，这是他为卖巧克力而故意设计出来的吗？我活到这把年纪，至今还留着自己得到的第一枚童子军奖章，所以，我真心希望这不是真的。不管“先提大要求，后提小要求”的顺序是不是他安排好的，效果总归一样：它管用了！因为管用，某些人便有意识地利用拒绝-后撤策略来实现其目的。首先，我们来看看怎样把这一策略当成可靠的说服手段使用；其次，我们会看到现在的人们是怎么使用它的；最后，我们会讲一讲这一策略鲜为人知的几点特性，因为正是靠着这些特性，它才成了一种最有效的说服手段。

还记得我在遇到小童子军之后，立刻把研究助理们召集到一起，尝试弄清楚这是怎么回事吗？最终的结果是，我们把巧克力给吃掉啦。其实，我们做的不止这些。我们还进行了一个实验，检验拒绝-后撤策略，即先提出较大的要求，遭到拒绝后再提出较小的真正要求的有效性。实验的目的有二。

其一，我们想看看这套手法对别的人管不管用。很明显，它对我是管用的，但我一贯容易掉入各种说服手段的陷阱。所以，我们的问题是：“拒绝-后撤策略能不能用到足够多的人身上，成为一种获得顺从的有效手段？”如果回答是肯定的，那将来我们可要小心提防着它了。其二，我们想确定这一顺从策略的力量到底有多大。它能让人顺从相当够分量的要求吗？换言之，提出要求的人退而求其次的“较小”要求，一定是一项真正“够小”的要求吗？我们探讨了这一策略管用的奥妙，要是这个想法没错，第二项要求不一定要有多小，只要比第一项要求小一点就行了。我们怀疑，提出要求的人从较大的要求退到较小的要求的行为之所以能达成目的，最关键的地方就在于它显得像是一种让步。这样一来，第二项要求在客观上就可以是一项很大的要求，其只需要比第一项要求小就可以了。经过一番考虑，我们决定运用这一策略提出一个我们认为大多数人都不会予以满足的要求。

影响力研究

我们假装成“县青年辅导项目”的代表，接近在校园里走动的大学生，问他们是否愿意花一天时间陪伴一群少年犯去游览动物园。对这些大学生来说，花好几个小时去陪一群年龄各异的少年犯出现在公共场合，又没有报酬，这样的要求实在没什么吸引力。不出所料，绝大多数的人（83%）都拒绝了。可是，当我们从相同的样本里抽选了另一群大学生，对提问手法稍加调整，然后向他们提出同样的请求后，却得

到了截然不同的结果。在请他们无偿陪伴少年犯逛动物园之前，我们先要他们做一件更大的“善举”：每星期花两个小时为少年犯当辅导员，为期至少两年；等他们拒绝了这个极端的要求之后（实际上所有人都拒绝了），我们再提出去动物园的小要求。把动物园之请打扮成对最初请求的让步之后，我们的成功率有了大幅上涨，答应陪同去动物园的大学生比之前增加了两倍多。

面对一项实质性请求，凡是能将顺从率在原先的基础上增加接近两倍（我们的实验将顺从率从17%提高到了50%）的手段，现实中必然早就有人变着花样用它了，这是毫无疑问的。比如，劳工谈判就经常采用以下策略：先提出极端的要求，但并不指望对方能同意，只不过从这一立场，他们可以更方便地向后撤，并让对方做出真正的让步。这样看来，最初的请求越大，这套做法的效果就越好，因为让人产生错觉的空间也相应较多。

但这么做是有限度的。根据以色列巴兰大学对拒绝-后撤策略的研究，要是最初的请求极端到了不合情理的地步，那便会产生事与愿违的结果。此时，首先提出极端要求的一方便会被认为是缺乏诚意。对方并不会觉得从完全不切实际的立场后退是真正的让步，因此也不会回应它。真正有天分的谈判人员只会把最初的立场稍做夸张，够他讨价还价、来上一连串的小小让步，最终能够使他从对方那里得到理想的结果，就足够了。

关于证明拒绝-后撤策略的有效性，我最喜欢来自法国的一项研究。当地3家餐馆的女服务员一边清理餐桌，一边问顾客是否需要甜点。如果顾客说“不”，女服务员就会立即改为询问是否需要咖啡或茶。结果使得这类订单的比例几乎增加了两倍。我从该研究的另一种情况中发现了特别具有启发意义的事情，也就是女服务员没有立刻给出是否需

要咖啡或茶的提议，而是隔了3分钟左右才这么做。结果这类订单只增加了一倍。显然，回报小恩惠的亏欠感会随着时间的流逝而衰减，这一道理也适用于回报小后撤的亏欠感。

READER'S REPORT | 读者报告

致西奥迪尼先生：

在大学完成电气工程的学业，又到能源部门工作了4年之后，我决定辞职，跟随自己的内心，投入软件开发这一行业重新开始。因为我所有的软件知识都是自学的，所以一开始的起点很低，是在一家只有**10**个人的小公司里做软件工程师。两年后，我决定要求加薪。问题是：我们老板是出了名地不爱给员工加薪。于是，我采取了以下做法。

首先，我向老板汇报了我在公司的加班时间，最重要的，是我为公司带来的利润。接着我说，“我认为自己不是普通员工，我做得比普通员工多，所以我希望能达到这一职位的市场平均薪资水平，也就是每年×××欧元”（当时，我的薪资比平均水平低**30%**）。他严厉地回答，“不行”。我沉默了**5**秒钟，说：“好吧，那么你能每月多给我×××欧元，并允许我每星期在家工作一天吗？”他答应了。

我知道他不会给我市场平均工资。我真正想要的是得到公平的加薪，然后一天的在家工作时间，这样我就能多陪伴我的未婚妻了。离开他的办公室时，我实现了两个结果：一是加薪**23%**；二是对拒绝-后撤策略重燃热情。

这位德国软件工程师：

请注意，通常情况下，使用拒绝-后撤策略也调动了对比原理。最初较大的数额让第二次较小数额的加薪请求显得像是后撤，还让它的数目显得更小了。

另外，这位读者的名字没有出现在本书最开头“读者报告”的作者名单里，因为他要求只提及他名字的首字母（M. S.）。

相互让步、知觉对比和“水门事件”之谜

我们已经讨论了拒绝-后撤策略管用的原因之一：它调用了互惠原则。不过，“先提大要求后提小要求”的策略之所以能发挥作用，还有另外两个原因。一是我们在第1章中碰到的知觉对比原理。人们买了西服之后容易在毛衣上多花钱，便是知觉对比原理在搞鬼：先接触到大件商品的价格，之后再看到不那么贵的商品时，后者的价格便会在对比中显得更加低廉。同样道理，“先大后小”地提要求，也是使用了对比原理：小要求跟大要求一比，更显得微不足道了。要是我想找你借10美元，我可以先提向你借20美元的要求，以使前一项要求显得小一些。这么做的好处之一就在于同时调用了互惠原则和对比原理的力量。先要20美元再要10美元，10美元的要求不仅会被看成是一种让步，还会显得数目更小一些。

两相结合，互惠原则和知觉对比原理就能产生一种令人生畏的强大力量。拒绝-后撤策略便是把它们捏在一起，发挥出惊人效用的。依我看，20世纪70年代最令人费解的一起政治举动——共和党成员擅自闯入民主党全国委员会办公室，最终导致尼克松总统下台的“水门事件”，大概便可以从这方面找找原因。该事件的决策参与者之一，杰布·斯图尔特·马格鲁德（Jeb Stuart Magruder）一听说擅闯水门的窃贼被抓住了，便大惑不解地说：“我们怎么会这么蠢哪？”确实，怎么会这么蠢哪？

为了说明尼克松行政当局破门而入的举动是多么有欠考虑，让我们来回顾几点事实：

- 这个主意是戈登·利迪（G. Gordon Liddy）出的，他在“总统竞选连任委员会”负责情报收集工作。行政当局的高层都知道利迪这个人“疯疯癫癫”，因此对他情绪的稳定性和判断力并不看好。
- 利迪的建议非常费钱，他要求获得25万美元现金的预算，以防追查。
- 3月下旬，委员会主任约翰·米切尔（John Mitchell）和他的助手马格鲁德（Magruder）、弗雷德里克·拉鲁（Frederick LaRue）开会通过了这一提议。此时，尼克松在11月大选中获胜的前景可谓一片光明。在前几轮投票中唯一有可能击败尼克松的候选人埃德蒙·马斯基（Edmund Muskie）初选表现不佳；看起来，最容易打败的候选人乔治·麦戈文（George McGovern）会赢下民主党的内部提名，共和党大选获胜，似乎已经十拿九稳了。
- 破门计划本身的风险性极高，它需要10个参与人员，而且人人都必须守口如瓶。
- 要偷偷溜进去并安装窃听装置的地点是民主党全国委员会及其主席劳伦斯·奥布莱恩（Lawrence O'Brien）设在水门大厦里的办公室。而这里并没有什么信息会对现任总统造成损害。民主党似乎也根本没有这样的信息，除非行政当局做了一件非常非常愚蠢的事情。

尽管上面说到的无一不是显而易见的事实，可这么一个众所周知的欠缺判断力的家伙提出的昂贵、冒险、毫无意义、具有潜在毁灭性的建议却还是被批准通过了。像米切尔和马格鲁德这样聪明、见多识广的

人，怎么会做出如此愚蠢的决定呢？答案或许就藏在一个少有人谈及的事实：他们批准通过的这个25万美元的计划并不是利迪的头一项提议。事实上，利迪先前还提过两个计划，而25万美元的这个计划于利迪而言已经是做出重大让步了。头一个计划是两个月前，利迪跟米切尔、马格鲁德、约翰·迪恩（John Dean）开会时提出来的。它需要耗费100万美元，除了要在水门大厦民主党全国委员会办公室安装窃听器之外，还需要一架装有特殊通信器材的“跟踪飞机”、一支负责绑架和抢劫的小分队和一艘载有高级应召女郎以便勒索民主党政客的游艇。一个星期后，利迪又在同一群人（米切尔、马格鲁德、迪恩）参与的会议上提出了第二项计划，这个计划削减了部分方案，把成本降低了50万美元。等这两项计划都被米切尔否决了之后，利迪才提交了最终付诸实施的25万美元“精简”计划，这次是在米切尔、马格鲁德和拉鲁参加的会议上提出的。这个计划还是很愚蠢，但比之前那两个要好些，于是得到了批准。

米切尔这位强硬、精明的政客，是不是跟我这样素来爱吃亏上当的傻子一样，中了同一套顺从手法的招儿，落入糟糕的交易里面了呢？只不过，哄我上当的是兜售糖果的小童子军，哄米切尔上当的却是兜售政治灾难的“神经男”。

让我们来看看马格鲁德的证词吧。“水门事件”的大多数调查员都认为，这一证词最忠实地描述了利迪计划最终得以通过的那次会议。这里面有一些发人警醒的线索。马格鲁德说：

没有谁对这个项目特别感兴趣，但跟利迪之前提出的100万美元的荒唐数目相比，我们觉得25万美元是个可以接受的数字……我们都不愿让他空手而归。米切尔认为我们总该给利迪一小点儿……米切尔签字的时候就好像在说：“好啦好啦，我们就给他25万美元，看看他能折腾出什么来吧！”

与利迪开始的两个极端方案相比，“25万美元”成了“一小点儿”，变

成了回赠利迪妥协的让步之举了。事后，马格鲁德脑袋清醒了，他意识到利迪采用的方法就是经典的拒绝-后撤策略。

要是利迪一开始就跑来对我们说：“我有个计划，我们偷偷潜入奥布莱恩的办公室，在那儿装上窃听器。”我们肯定会直截了当地拒绝。相反，他先拿给我们看的是什么应召女郎、绑架、抢劫、破坏、窃听这样荒诞不经的复杂方案。表面上，利迪要一整条面包，可在他心里，只要给他一半甚至1/4他就满足了。

有趣的是，小组里只有一名成员——拉鲁直接对这个提议表示了反对，尽管他最终还是服从了老板的决定。拉鲁凭借显而易见的常识说道：“我不觉得有必要冒这个险。”他肯定好奇为什么米切尔和马格鲁德这两位同事不这么想。当然，拉鲁和另外两人就利迪的方案是否可取看法不一致的背后肯定有各方面的原因。但最突出的一点应该是：在三个人里面，只有拉鲁没有出席前面两次会议，没听到利迪勾勒他那更加野心勃勃的方案。或许，也正是因为这样，拉鲁才得以从客观的角度做出评价，而不像其他两人那样会受互惠原则和知觉对比原理的双重影响。

进退两难

之前我们说过，除了互惠原则，拒绝-后撤策略还借助了另外两点有利因素。我们已经讨论了第一点，即知觉对比原理。跟其他原理不同，拒绝-后撤策略的额外优势算不上真正的心理学原理。相反，它只不过是一种请求顺序上的安排。让我们再做一次先前的假设：我想找你借10美元。我先请求你借给我20美元，对我而言，这样做是稳赚不赔的。如果你同意的话，我得到的钱将是我所需的两倍；如果你拒绝了，我还可以让步到一开始就想要的10美元，而靠着互惠原则和知觉对比原理的作用，这回我成功的概率将大大增加（见图2-4）。无论哪种方

式，对我都是有好处的。这就好像我们玩投硬币定输赢游戏，正面，我赢；反面，你输。



图2-4 戈登：危险人物？

相似的风格能带来相似满足的微笑吗？似乎的确如此。

拒绝-后撤策略的优点这么多，有人或许会觉得，它肯定也存在很大的缺点。受害者被这一手法逼得只能顺从，说不定还会充满怨恨。怨恨或许会以若干种形式表现出来：其一，受害者可能会否认跟请求者达

成的口头协议；其二，受害者可能会对操纵自己的请求者产生怀疑，并决定永远不再跟此人打交道。倘若这样的情况发生，请求者在使用拒绝-后撤策略时必然要三思而后行。然而，研究表明，采用拒绝-后撤策略时，上述受害反应的发生频率并未增加。更令人惊讶的是，其实际发生频率好像反倒降低了！要弄清为什么会这样，让我们先来看看下面的证据。

这是我的血，请千万记得打电话

拒绝-后撤策略的受害者是否会履行承诺，按照请求者的后一项要求做呢？加拿大的一项研究有助于解答这个问题。这次研究除了记录了目标对象是同意还是拒绝了为社区心理健康诊所无偿工作两小时这一真正的请求之外，还记录了他们是否如约履行了自己的职责。和通常一样，一开始提出较大的请求（要求被试在至少两年时间内，每周到该诊所工作两小时），再后退到较小的请求，口头答应的人会更多，有76%；而直接提出较小的请求，答应的人则只有29%。不过，答应来的志愿者里有多少人真正出现了呢？这个结果显然更为重要。此时，拒绝-后撤策略仍然更为有效，前者是85%，后者仅为50%。

还有一个实验检验了受害者是否会因受了拒绝-后撤策略的操纵，而对进一步的要求一概加以拒绝。这项研究的被试是大学生。

影响力研究

研究者要学生们每人在学校一年一度的献血活动中献血400毫升。一组被试最初听到的要求是，每6个星期献血400毫升，为期至少3年，之后才改为只献血400毫升；另一组被试听到的要求则从一开始就是只献血400毫升。接下来，对这两组被试里口头答应并真正来到献血中心的人，研究者都

问了“是否愿意留下电话号码，以便下一次献血时再联系”的问题。凡是因为拒绝-后撤策略而献血400毫升的学生，几乎全都答应再来献血（84%）；其他来到献血中心的学生答应再次献血的却不到一半（43%）。

实验证明，就算考虑到将来的合作，拒绝-后撤策略也仍然更胜一筹。

甜甜蜜蜜的副作用

这可真是够怪的：拒绝-后撤策略似乎不仅会刺激人们答应请求，还会鼓励人们切实践践承诺，甚至人们还会自愿履行进一步的要求。这个手法到底有什么奥妙，能糊弄得人们一再地上当受骗呢？要想知道答案，我们或许可以观察一下请求者的退让行为，即该手法的核心环节。我们已经看到，只要对方不觉得它是个一眼就能洞穿的骗局，让步便有可能刺激对方也退让一步。不过，让步举动还有一项我们尚未着手研究的、少有人知的积极附加作用：对方会对这种安排产生更强的责任感和满意感。靠着这种甜蜜的附加作用，拒绝-后撤策略推动受害者履行了协议，痛快地答应了之后的约定。

通过研究人们讨价还价的方式和方法，我们可以清晰地看到让步给人际交往带来的积极附加作用。加州大学洛杉矶分校的社会心理学家们曾经做过一项实验，为此做出了很好的示范。

影响力研究

研究者答应提供一定数额的金钱，并告诉被试，他有一个“谈判对手”，两人必须讨价还价，商量如何分配这笔钱。被试还被告知，要是过了一段时间，两人还没有达成协议，

那么谁都分不到钱。但被试并不知道，所谓的“对手”其实是研究助理，此人将按预先的指示，用以下三种方式中的一种跟被试讨价还价。对第一组被试，“对手”会首先提出极端的要求，把差不多所有的钱都留给自己，而且在整个谈判过程中都顽固地坚持这一主张；对第二组被试，“对手”一开始提出的要求只是稍微有利于自己，但在谈判过程中，他也始终坚守这一立场，拒绝让步；对第三组被试，“对手”最开始提出的是极端的要求，之后逐渐退让到略微有利于自己的要求。

实验中的三点发现有助于我们理解为什么拒绝-后撤策略如此管用。首先，相较于另外两种方法，先提极端要求再退让到适度的要求，能让使用此方法的人获得最多的钱。我们先前就看过了相关的证据，知道这种手法能强有力地达成有利于己方的协议，所以上述结果尚在情理之中。但研究的另外两点发现很是惊人。

责任感 请求者运用拒绝-后撤策略主动让步，不仅提高了目标对象答应的概率，也令他们觉得最终协议是自己“说了算的”，进而产生了更多的责任感。人们要是对于协议的条款感到负有责任，自然也更乐意遵守这一契约，因此，目标对象履行承诺也就不足为奇了。

满意感 尽管平均而言，被试分给采用后撤策略的对手的钱最多，但这部分人对最终安排是最为满意的。这样看来，人们或许是这么想的：靠着自己的努力，对手“后撤”了。如此达成的协议自然分外圆满。根据这一点，我们就可以解释拒绝-后撤策略的第二个神奇特点了——受害者对之后的请求居然也会照样答应。这一策略利用后撤来使人顺从，因此受害者可能会对最后的安排感到更为满意。毫无疑问，对特定安排感到满意的人，更乐意答应类似的安排。

消费者研究者罗伯特·辛德勒（Robert Schindler）在零售领域做过两

次研究，发现倘若消费者觉得做成划算的交易有自己的一份功劳，那么他们就会对整个过程中感到更满意，并会回头购买更多的产品。

也就是说，拒绝-后撤策略能让他人更有可能实际执行我们接下来的请求，并进而答应执行类似的请求，这一发现与加州大学洛杉矶分校的实验所发现的责任和满意感是一致的。该实验还证明了另一个结果：从一个极端的立场开始，接着退回一个温和的立场，比一开始就从温和的立场开始并坚持到底更有效。这一结果与前文所介绍的加拿大折扣宠物商品店老板的谈判策略一致。

READER'S REPORT | 读者报告

致西奥迪尼先生：

有好长一阵子，我都在一家大型零售店的电视和音响器材部门工作。零售商会为客户提供延期保修合同，而售货员卖出这种合同的能力，决定了他能不能继续受聘上岗。我一听他们给我解释了这一点，就设计了以下采用了拒绝-后撤策略的方案，虽说那时我还不知道该手法的名字。

客户购机时可以挑选一年到三年不等的质保服务，但不管卖出的合同是哪一种，我所得的积分都是一样的。我意识到多数人都不愿意购买三年质保，所以，一开始我总是劝他们购买这种时间最长、价格也最高的质保。这样一来，要是客户拒绝了我真诚推销的三年质保，我就得到了一个绝妙的机会，后退到相对便宜的一年质保上。只要能把这种质保服务卖出去，我照样很高兴。事实证明，这种手段非常有效，因为平均下来，**70%**的顾客都买了延期质保，而部门内其他销售人员卖出的比例才**40%**左右。此

前，我还从没对人透露过这个小秘密呢。

这位前电视机和音响器材销售员：

拒绝-后撤策略不光提高了顾客接受协议的概率，也提高了他们对所达成协议的满意度。

如何防范

要抵挡应用了互惠原则的请求者，你我面对的将是一个可怕的敌人。靠着向我们首先示好、主动让步，请求者招募到了一支强大的同盟军，势要争取我们的顺从。乍看之下，我们的前景显得不怎么光明：我们可能会屈从于互惠原则，顺从请求者的愿望；我们也可能拒绝顺从，这样一来，我们心底的公平感和义务感就会承受互惠原则带来的猛烈冲击。要么举手投降，要么死伤惨重，两者都不是什么好结果。

幸运的是，我们不是只有这些选择。只要我们对对手的性质有了正确的认识，就能安然无恙地撤离顺从战场，有时甚至还能捞回些战果。关键是要意识到，请求者并不是我们真正的手，他不过是在借助互惠原则或其他任何影响力武器来争取我们的顺从罢了。他选择了四两拨千斤的柔道手法，让自己跟互惠原则的强大力量站到了同一阵线，之后，又靠着抢先施恩或让步，释放出这种力量。我们真正的手是互惠原则，要想不受它蹂躏，我们必须采取措施，化解它的威力。

拒绝互惠原则

要怎么做才能抵消互惠这种社会原则的影响呢？一经激活，它的力量便会铺天盖地地压下来，强大到让我们根本没法抵挡。这样看来，不

让它激活似乎是个好办法。或许，抢先出手，不让请求者借用它的力量，我们便能避免跟互惠原则发生正面冲突。据此，拒绝请求者最初的善意或让步，大概可以让我们成功回避这一问题。可是一概拒绝请求者的最初善意或让步，理论上看起来不错，但实践起来却颇为棘手。最主要的问题在于，当最初碰到一个请求时，你很难判断它到底是出于真诚，还是打算利用你。这是个“不给糖就捣蛋”式的问题：要是我们总是戒备森严（总想着防备别人的捣蛋），碰到无意利用互惠原则的人向我们做出的让步，或是施予的好意（接受别人的糖果），我们也就收获不了它所带来的果实了。

我有个同事愤怒地对我说过一件事。有个男人为了免遭互惠原则的“毒手”，非常粗暴地拒绝了这位同事10岁女儿的好意，小女孩的心灵因此受到了深深的伤害。

影响力研究

女孩班上的孩子们在学校里组织招待会，欢迎自己的爷爷奶奶前来参观，女孩的任务是给所有进入学校操场的家长送花。她碰到的头一个男人看到她递花上来，立刻吼叫起来：“你自己留着吧！”女孩不知道该怎么办，再次拿着花朝他走过去，这个男人却厉声喝问她到底有什么企图。女孩无力地回答：“什么也没有，这只是一份礼物。”这个男人还是不信任地瞪着她，坚持说自己早就识穿了她的“把戏”，接着便推开她走掉了。

这次经历给小女孩造成了莫大的伤害，她再也没法接触其他人了，只好放弃了自己的任务。起初她可是满怀期待来做这件事的。很难说这该怪到谁头上，是那态度粗暴的男人，还是之前依靠互惠原则利用他、弄得他最后只知道一味拒绝别人的人。不管你觉得应该怪谁，这里的教

训很明显。我们总会遇到真正慷慨的人，还有按互惠原则公平游戏、不利用它占便宜的人。要是有人不分青红皂白地拒绝了他们的努力，那么他们肯定会觉得受了侮辱，社会摩擦和孤立也会由此而生。因此，一概排斥的策略似乎并不合适。

另一种解决方案成功的把握更大。倘若别人的提议我们确实赞同，那就不妨接受它；倘若这一提议别有所图，我们就置之不理。比如，有人给了我们一个恩惠，我们大可以接受下来，同时认识到将来有回报他的义务。跟别人达成这样的协议，并不意味着这个人能通过互惠原则利用我们。相反，自从人类来到这个世界，公平地参加“义务信誉网”之后，我们就在个人和社会层面上得到了许多好处。然而，要是最初的善意其实是个圈套、机关或诡计，即专门设计来刺激我们回报以更大的恩惠，那么情况就不一样了。我们的合作伙伴并非心肠好，而是想牟取暴利。基于这样的条件，我们当然应该采取相应的措施。一旦我们确定最初的恩惠并非出于善意，而只是一个顺从伎俩，我们就不必受它影响了，该怎么做就怎么做。只要我们能准确地判断、界定顺从伎俩，不再把它们看成是恩惠，施予者也就没法再跟互惠原则站到同一阵线了。互惠原则只说要以善意回报善意，可没说要用善意来回报诡计。

明辨敌友

为了说得再具体些，我们举个实际的例子吧。

假设有一天，一位妇女打来电话，说她是城里居民消防安全协会的会员。她问你有没有兴趣了解一下家庭防火安全知识，检查一下房子是否存在安全隐患，并称之后还会送你一套家用灭火器，一切全都免费。若你对这些都挺感兴趣，并且你还答应让协会的安检人员晚上到访你家。安检员到了之后，送给你一小罐手提灭火器，并开始检查你家的火灾隐患。之后，他给你讲了一些有关火灾的综合信息，这些信息有些意思，也挺吓人的。他还评估了你家发生火灾的可能性。最后，他建议你

安装家庭火警系统，接着就离开了。

这样的事儿并不少见。好多城市都有非营利的协会，通常，专职消防员会在业余时间提供此类免费的住宅防火检查。要是你真的碰到了这样的事情，显然是得到了安检员的好心帮助。根据互惠原则，倘若将来你看到他有什么需要帮忙的地方，你也应当主动施以援手。这一类的善意交换，是完全符合互惠原则的优良传统的。

可类似的事件也有可能出现不一样的结局：安检员推荐火警系统之后并未离去，而是展开了一场推销陈述，想要说服你购买他公司生产的一套热感应报警系统，价格自然十分昂贵。上门推销家用火警系统的公司经常采用这种做法。通常他们的产品虽说足够管用，但价格都是虚高的，他们知道你不熟悉这种系统的市场价。要是你打算安装这一系统，你会觉得，既然这家公司为你提供了免费的灭火器，还检查了你家的安全情况，那么你就欠了他们一个人情，应当从他们那儿买一套。而这些公司正是利用了你的这种心情，向你施压，要你马上就买。靠着这种手法，卖防火装置的企业把生意做得红红火火。

若你发现自己碰到了这样的情形，可又意识到安检员来你家的首要动机是向你推销昂贵的报警系统，那下一步该怎么做呢？最有效的做法既简单也不会惊动他人。你只要在心理上重新下个定义就行了：把从安检员那里得到的一切，即灭火器、安全信息、隐患检查，全都当成是销售手法而不是礼物，这样一来，你就能轻松地拒绝他要你买东西的提议了。而且，你还可以免受互惠原则的影响：善意自然应当以善意回报，可对销售策略却没这个必要。

倘若你乐意，你甚至还可以让安检员的影响力武器掉转枪口。回想一下，互惠原则告诉你，人家怎样对你，你就有权怎样对他。要是你确定“防火安全检查员”的礼物并非真正意义上的礼物，而是用来从你那儿赚钱的工具，那么你也可以用它们为自己赚得好处。安检员给你的东西，如安全信息、家用灭火器，你照单全收，然后礼貌地道个谢，把他

送出门去。毕竟，互惠原则说了，公正的意思就是：盘剥的行为要还以盘剥。

READER'S REPORT | 读者报告

致西奥迪尼先生：

我对行为心理学很感兴趣，所以买了《影响力》这本书，并在昨天读完了关于互惠的那一章。

今天我去超市，一个自称是瑜伽爱好者的家伙拦住了我。他主动要解读我的气场，并说能看出我是一个冷静和乐于助人的人。接着，他从口袋里拿出一颗小珍珠，作为礼物送给我。过了一会儿，他要求我捐款。我说我是个穷学生，没有多余的钱能捐款，他便强调他给了我珍珠，我只有以捐赠作为回报才公平。因为我刚在不到**24**小时前读过关于互惠的内容，清楚地知道他提及珍珠是为了达到什么目的，所以我拒绝了他。一番争吵后，他只得走开了。

苏黎世的这名化学工程专业的学生：

“知识带给我们自由”这句古老的格言放在本例中很合适。知道了如何防范利用互惠原则投机取巧的人，你得以摆脱不请自来的虚假礼物的牵绊。此外，我敢肯定，在这个故事里，瑜伽爱好者给的并不是真正的珍珠，不过他的故事或许为我们提供了一颗智慧的珍珠。

本章小结

- 社会学家和人类学家指出，互惠原则体现了人类文化最普遍、最基本的一种规范。该原则主张，只要他人给了我们东西，我们就应当设法偿还。原则赋予了接受者将来偿还的义务，这样人们就可以大胆地把东西给别人，而无须担心自己会遭受损失。有了该原则内置的未来责任感，人类才得以发展出多种持久的关系、交易和交换行为。因此，所有社会成员从小就被告知应该遵守该原则，不然就会遭到严重的社会排挤。
- 互惠原则常常影响人们做出是否顺从他人要求的决定。部分说服专业人士最喜欢也最有利可图的手法之一，就是在索要好处之前先给他人点什么东西。这一手法之所以能被别有用心的人所利用，原因就在于互惠原则的三个特点。第一，互惠原则极为强大，通常情况下会影响决定人们是否顺从他人要求的因素，人们都不是它的敌手；如果礼物、好意或服务是个性化的，或是根据接受方当下偏好或需求量身定制的，这条原则尤其有效。第二，即使最初的恩惠是强加于人的，该原则也照样管用，这样一来，我们就没法主动决定自己愿意欠谁的人情了，因为选择权已经落到了他人手里。第三，该原则能推动不平等的交换，为了消除令人不安的亏欠感，人们往往会答应一个比先前所得人情大得多的要求。
- 还有一种利用互惠原则提高他人顺从概率的方法，该方法对基本主题做了个小调整：不是抢先给好处，推动他人回报，而是抢先让步，刺激他人也让步。这种顺从程序，叫作拒绝-后撤策略，也叫留面子法。该方法主要依靠的就是回报让步的压力。提出请求的人以肯定会遭到拒绝的极端要求拉开序幕，之后再后撤到较小的要求，即他原本就想达成的目标，而后一要求很可能会被对方接受，因为它

显得像是一种让步。经研究表明，拒绝-后撤策略除了能提高对方答应要求的概率以外，还能提高对方将来履行这一要求的概率。这是因为在参与让步的互惠交换之后，人们觉得负有更多的责任，对结果也更满意了。

- 要防备别人用互惠原则向我们施压、要我们顺从，最好的办法不是一概拒绝他人的最初善意。相反，我们应当大方接受最初的恩惠或让步。可一旦事实证明对方并非出于善意，我们就要重新定义其行为。只要不再把这些行为看成是恩惠或让步，也就不会觉得有必要以善意或让步做出回应。

Influence

第3章

喜好

最有效的销售手法是，让顾客真心地相信你
喜欢他们。

——乔·吉拉德（Joe Girard）

NEW AND EXPANDED

我们大多数人总是更容易答应自己认识和喜欢的人所提出的要求。对于这一点，恐怕不会有什么人感到吃惊。令人吃惊的地方在于，有些我们完全不认识的人，却想出了上百种方法利用这条简单的原则，让我们顺从他们的要求。这种倾向让科学传播工作者解决了困扰他们数十年的问题。这个问题是：怎样让更多的人接受达尔文的进化论。进化论认为，包括人类在内的所有生物，其当前的形式完全是由自然选择等系统进化过程塑造而得的。科学传播工作者的困难在于，由于进化论的主张与大量宗教团体坚信是上帝之手决定了人类模样的信仰背道而驰，后者往往难以接受进化论。最近一项关于这一主题的民意调查显示，只有33%的美国人认为人类纯粹是通过自然进化发展起来的。

为解决这一问题，相关工作者希望通过以下方法提高进化论在民众中的接受度：描述科学家就进化论有效性达成的共识态度；介绍数以千计的证实了进化思想的研究；强调应用进化论给医学、遗传学、农业和药理学方面带来的进步；通过更密集的教学工作，提升民众对进化逻辑的认同态度。然而，这些工作都收效甚微。例如，最后一种方法对构建进化论的信念完全没有作用，因为研究表明，一个人对进化的信念与他对进化逻辑的理解并无关联。这种脱节的理由也很充分：对进化论的抵制并非缘于对进化逻辑的不认同，而是因为进化论有悖于人们基于情感的偏好、信仰和价值观（这些往往建立在宗教信仰的基础上）。

因此，试图用逻辑论证来推翻基于信念、因情感而保留的信仰是徒劳的，因为两者代表着截然不同的认知方式。300年前，英国作家乔纳森·斯威夫特（Jonathan Swift）就指出了这一点：“想用理性去说服一个人放弃相信一件他从不是因为理性而相信的事情，根本没有用。”⁽⁸⁾

这些科学传播工作者的误区在于，他们总是把思考视为高于一切的认知方式，始终坚信事实最终会说服那些把对观念的感受摆在第一位，而非优先接受进化论的事实受众。有什么方法能拯救这些误入歧途的

科学传播工作者呢？

这就需要喜好原则了。一支加拿大的心理学家团队认为，找一个广受欢迎的进化论支持者传递简单的消息，就能提升人们对进化论的态度。他们锁定了美国电影明星乔治·克鲁尼（George Clooney）来充当达尔文理论的倡导者。

在研究中，当人们在引导下相信克鲁尼对一本支持进化论的书做出了正面评价时，他们就明显变得更能接受该理论了。更重要的是，这种变化与被试的年龄、性别和宗教信仰程度无关。为了确保这一结果不是克鲁尼本人（或男明星）所独有的特质导致的，研究者又找了因主演“哈利·波特”系列电影而广受欢迎的女明星艾玛·沃特森（Emma Watson）重新做了研究，也发现了相同的结果。对渴望说服别人的人来说，答案显而易见：要想改变感觉，就要用其他的感觉来抵消它；对科学传播工作者来说，喜好是一种能够抵消其他感觉的有用感觉。

有关克鲁尼和沃特森的研究还包含另外两个极具启发意义的点。第一点是，它证明受欢迎的名人的观点，既能降低也能提高人们对进化论的接受程度，从而扩展了基本效应适用的范围。如果被试相信克鲁尼或沃特森对一本反进化论的书发表过正面评论，那这些人对进化论的支持度便会大幅下降。因此，喜好的影响力并不是单向的，它可以将态度引导到积极或消极的方向上。第二点是，它证明借助讨人喜欢而非权威的沟通人员来改变人们的态度更为有效。在另一个实验中，研究者向其中一组被试展示了一条正面评语。这条评语针对某本支持（或反对）进化理论的书，据说出自一位著名大学的生物学教授，即专家的意见。研究者发现，不管专家是支持还是反对进化论，对被试是否接受进化论都没有太大影响。此时证据再明显不过：为什么这么多年来，科学传播工作者为提高进化论的支持率而发起的一次次艰苦努力都失败了？因为他们选错了战场。

想要了解喜好的感觉能多么强有力地引导他人的选择，可以看看顶

尖的医疗事故律师爱丽丝·贝尔金（Alice Burkin）对以下记者问的回答：

记者：每个医生都会偶尔犯错误。但大多数错误不会演变成医疗事故诉讼。为什么有些医生就更容易遭到起诉呢？

贝尔金：在许多案件中，除了疏忽本身，最重要的因素是医患关系的质量。这么多年来，从来没有哪个潜在客户走进来说：“我真的很喜欢这位医生，但我想起诉他.....”人们是不会起诉自己喜欢的医生的。”

通过喜好赚钱

据我所知，“特百惠家庭聚会”就是一个专门利用喜好原则做生意的最明显的例子，我认为它属于典型的顺从环境。凡是熟悉特百惠聚会那套手段的人都认得出它运用了本书中涉及的几种影响力原则：

- 互惠。一开始，去参加聚会的人玩游戏、赢奖品。没中奖的人可以从一个袋子里摸奖，这样所有人还没买东西就得到了一份礼物。
- 权威。特百惠产品的质量和安全得到了专家的认证。
- 社会认同。买卖拉开序幕之后，每一笔做成的生意都在强化、巩固以下观点：其他类似的人都想要这种产品，因此它一定很不错。
- 联盟。顾客一旦购买产品，就成了“特百惠大家庭的一员”。
- 承诺与一致。聚会要求参与者当众介绍自己发现从已有的特百惠塑料器皿上找到了怎样的用途，带来了哪些好处，

这也就是对特百惠做出了当众承诺。

- 稀缺：聚会上会介绍产品的独特好处和限时优惠。

所有的影响力武器全都上场了，为的是让事情进行得顺顺利利。但特百惠聚会最厉害的一招还在于，它根据喜好原则做了一种特殊的安排。虽说特百惠的推销员讨人喜欢，说服能力也很强，但这个陌生人并不会向潜在买家提出真正的购买请求，负责做这件事的是房间里所有人的一位共同朋友。也就是说，特百惠的销售代表可以要求参加聚会的人下订单，可真正给大伙儿造成更大心理压力的却是坐在一边说说笑笑、端茶送水的那位。她就是主办聚会的女主人，是她叫朋友们到家里来看特百惠的产品展示的，而且，人人都知道这次聚会上每卖出一件产品，她都能从中抽成。

通过让女主人分成的做法，特百惠家庭聚会公司让客户看在朋友的面子上从朋友手里买东西。这样一来，友谊的吸引力、温情感、安全感和义务感全都被带到了销售环境当中。事实上，消费研究者曾检验过家庭聚会销售环境下女主人和参加者之间的社会纽带，肯定了该公司策略的有效性。在决定是否购买该产品时，社会纽带的影响比对产品本身的好恶强两倍（见图3-1）。



图3-1 特百惠通过家庭聚会销售产品

特百惠在类似这样的家庭聚会上展示一系列的环保清洁产品。参加聚会的人和女主人之间的情感纽带，往往能让交易顺利进行下去。

这一招的确效果惊人。根据最新的估计，特百惠的日销售额已经超过了550万美元！而且，特百惠还把这套做法搬到了欧洲、拉丁美洲和亚洲，同样取得了巨大的成功。在这些地方，人们在朋友和家庭关系网里的地位要比在美国重要得多。所以，如今特百惠在北美的销售量还不到其总销售量的25%。

有趣的是，客户们似乎完全明白特百惠聚会中喜欢和友谊所造成的压力。对此有些人似乎并不介意，有些人虽然介意，但似乎并不知道如何避免这样的压力。我曾跟一位女士聊过，她用一种相当沮丧的声音描述了自己的反应：

现在我已经到了痛恨受邀去参加特百惠聚会的地步。我早就有了我所需要的各类容器。要是我真的想要，我也会到商店里选其他更便宜的品牌。可当朋友打电话来时，我却觉得必须得去。等我到了那儿，又觉得必须买点什么才好。我能怎么做

呢？那可是我朋友！

由于找到了友谊这个无比强大的盟友，不足为奇，特百惠公司放弃了零售网点，转而大力推动家庭聚会的概念。2003年，特百惠公司做了一件大事，挑战了几乎所有的商业逻辑，它断绝了跟零售商美国塔吉特公司（Target）的关系。因为他们的产品在塔吉特卖得太疯狂了，这样一来，塔吉特的零售就对举办家庭聚会的次数造成了威胁，所以，特百惠不得不中止双方的合作。

研究表明，是社交联系的质量而非实体产品的质量决定了特百惠聚会上的购买行为。2020年2月新型冠状病毒在全球范围内出现后，特百惠品牌在纽约证券交易所的股价大幅下跌。股价下跌（与2019年2月相比下跌了90%）在很大程度上是由于消费者认为线下聚会不再安全，就连朋友聚会也不安全。这足以证明特百惠的成功仰赖社交基础。

有趣的是，尼尔森公司的调查显示，人们对好友的推荐更信任。但如果喜欢的朋友变成讨厌的人（如前女友或前男友），这种模式就会发生逆转。在这种情况下，与网络评论相比，消费者相信这类人关于产品的意见的可能性要比前者低66%。不管是哪一种情况，喜好似乎都是关键。

统计显示，特百惠聚会每1.8秒就会举办一次。当然，其他各类说服专业人士也意识到了友谊给人带来的压力。比方说，越来越多的慈善组织开始让志愿者到自家附近的地区拉票、募捐。他们完全知道，我们有多难拒绝来自邻居或朋友提出的慈善请求。

其他说服专业人士还发现，朋友哪怕不在场也能发挥作用，很多时候，稍微提一下朋友的名字就够了。专攻上门推销各类家居产品的嘉康利公司（Shaklee）就建议销售人员采用“无穷链”方式寻找新客户。只要客户承认自己喜欢某件产品，就可以向他施加压力，问他还有哪些朋友可能会喜欢这种产品。之后销售人员就去找他的朋友们，他的朋友们又

推荐其他朋友，其他朋友再推荐更多的潜在客户，如此形成一条“无穷链”。

这套方法成功的关键是，销售员每次上门拜访新的潜在客户，总会报出此人一位朋友的名字，说：“是他建议我来找您的。”人们很难在这种情况下把销售员拒之门外，因为这简直就像是在拒绝自己的朋友。嘉康利公司的销售手册主张，员工们务必要使用这种方法：“它的价值说得再高也不为过。打电话或拜访潜在客户时，要是你能说是他的一位朋友建议你来找他的，那简直相当于进门之前就成功了一半。”

尼尔森公司的一项调查揭示了嘉康利集团的“无穷链”技术如此成功的秘密：92%消费者信任自己认识的人（比如一位喜欢的朋友）所推荐的产品，这比其他任何推荐源带来的成交率都要高得多，带来第二高效率的产品推荐源是网上的其他用户评论，信任度比前者低22%。朋友之间的这种高度信任，给被推荐公司带来了“惊人的利润”。研究者对一家银行的“朋友引荐计划”进行分析后发现，和普通新客户相比，经朋友推荐来的客户在3年里对银行的忠诚度提高了18%，其带来的利润提高了16%。

READER'S REPORT | 读者报告

致西奥迪尼先生：

尽管我从没参加过特百惠聚会，可最近我受到了一种与之类似的友谊压力。当时，我接到一个长途电话公司的推销员打来的电话。她告诉我，我的一位好朋友把我的名字列入了“**MCI**朋友和家人通话圈”。

我的这位朋友名叫布拉德，从小跟我一起长大，去年为了工作搬到了新泽西州。他经常跟我通电话，打听老朋友们的最新消息。推销员说，他给“通话圈”名单上的人打电

话，能省下**20%**的电话费，但前提是这些人都得是**MCI**电话公司的用户。她问我想不想换用**MCI**电话公司，享受这家公司提供的各种服务，这样布拉德给我打电话时也可以节省**20%**的话费。

呃，我才不在乎**MCI**服务能带来的好处呢！我对现在用的这家长途电话公司非常满意，但“布拉德给我打电话能省钱”这部分真的打动了我。要是我说不想加入布拉德的“通话圈”，不希望布拉德省点钱，那么布拉德听到了肯定会觉得很不好受。所以，为了不让布拉德受伤害，我告诉电话推销员，我愿意换用**MCI**。

我以前很好奇，为什么妇女们去了特百惠聚会，会只因为主办者是自己的朋友就买下一大堆并不想要的东西，这下我算是明白了。

芝加哥的这位先生：

能够证明“**MCI**通话圈”概念里蕴含着强大友谊压力的不光只有你，《消费者报告》杂志曾调查过这一做法，他们采访的**MCI**公司销售人员简洁明了地给出了回答：“这样做十次有九次都能成功。”

虽然**MCI**和“通话圈”服务均已过时，我仍然选择保留了你这个例子，因为它的启发意义太大了。许多公司的“向朋友推荐”方案都出现了新版本。事实证明，这些方案相当有效。据说，一位特斯拉车主向自己社交网络里的**188**个人推荐了该车，他因此赚到了**13.5**万美元的奖励，而特斯拉则获得了**1 600**万美元的销售额。我也碰到过类似的例子，健身房的一个朋友最近收到了他的互联网供应商**Cox**

Communications的“向朋友推荐”促销活动，如果他成功介绍一位新客户给Cox，他的账单就会减免100美元。当他向我介绍时，因为我知道Cox公司在搞什么把戏，所以拒绝了他。可即便如此，事后我仍然有好几个星期看到他就觉得不好意思。

利用交朋友来影响人

说服专业人士对朋友间喜好纽带的普遍利用，说明喜好原则在促人答应请求时的力量是多么强大。事实上，我们发现，哪怕是在根本没有现成的友谊可供利用的时候，这些专业人士也能靠这条原则获得好处。在这类情况下，他们会采用一条相当直接的顺从策略来利用喜好纽带：先让人们喜欢他们。

影响力研究

底特律有个叫乔·吉拉德的人，专门利用喜好原则销售雪佛兰轿车。他凭这个发家致富，每年能赚好几十万美元。听到这么高的薪资，我们大概会猜他是通用汽车的高层主管，或是雪佛兰经销店的老板。事实并非如此，他就是一个基层推销员，但他的业绩惊人。在整整12年里，他连续夺得“头号汽车销售员”的称号，他平均每天能卖掉5辆汽车和皮卡，《吉尼斯世界纪录》称他是世界上“最伟大的汽车推销员”。

吉拉德取得了这么大的成功，采用的办法却出奇地简单（见图3-2）。无非是向顾客提供两样东西：一个公平的价格，一个人们乐意从他那儿买东西的家伙。“就是这样，”吉拉德在接受采访时说，“找个你

喜欢的推销员，再加上优惠的价格。要是你两者皆有，那生意就成了。”

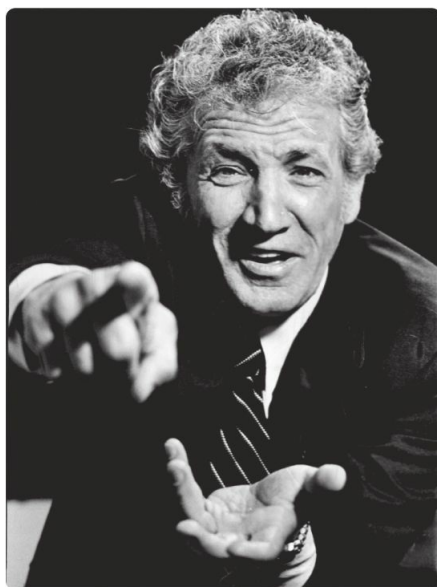


图3-2 乔·吉拉德：“我喜欢你。”

吉拉德先生透露了他每年12次通过邮寄贺卡告诉自己13 000名客户的话，这帮助他成了“最伟大的汽车推销员”。

吉拉德的策略告诉了我们喜好原则对他的业务有多重要，但他说得还不够清楚。至少，他并没有告诉我们，较之其他报出了公道价格的销售员，他为什么更受顾客喜欢。这里有一个关键、迷人的概括性问题，吉拉德的策略未做解答，是什么因素让人喜欢上某个人的？要是能够知道这个问题的答案，我们也就朝前迈出了一大步，就有希望搞懂吉拉德这样的人是怎么让人们喜欢上他们的了。同样地，我们也会明白该怎样做好安排，才能令别人喜欢上我们。幸运的是，社会学家们几十年来一直在研究这个问题。他们收集了很多证据，确认了一系列能使人喜欢上某人的可靠因素。从中我们可以看到，为了让我们走上答应他们请求的道路，每一个因素都被说服专业人士巧妙地利用上了。

我喜欢你的理由

外表魅力

虽然人们普遍承认长得好看的人在社会交往中占有优势，可最近的调查结果显示，对这种优势到底达到了何种程度，我们恐怕低估得太厉害了。碰到漂亮的人，我们似乎同样会报之以“按一下就播放”的反应，而且这种反应同样是不假思索地自动做出来的。这种反应属于社会学家口中的“光环效应”。光环效应指的是，一个人的某个正面特征就能主导其他人看待此人的眼光。现有证据清楚地表明了，大多数时候外表魅力就是这样的一个特征。

研究表明，我们会自动给长得好看的人添加一些正面特点，比如有才华、善良、诚实、聪明、随和与值得信任等。而且，我们在做出这些判断的时候并没有意识到外表魅力在其中发挥的作用。“好看就等于好”这种无意识假设造成的部分后果把我吓了一跳。

影响力研究

例如，一项针对加拿大联邦选举的研究发现，外表富有魅力的候选人得到的选票比缺乏吸引力的候选人要多2.5倍。除了这些英俊政治家受偏爱的证据，后续研究还表明，选民们并没有意识到自己的偏爱。事实上，在受访的加拿大选民中，73%的人都措辞强硬地否认了自己投票会受到候选人外表吸引力的影响，只有14%的人认为有可能存在这种影响。选民们尽管可以否认外表魅力对选举结果的影响，但越来越多的证据证明，这种烦人的倾向的确存在。

招聘的时候也存在类似的效应。一项研究模拟了招聘面试，发现应聘者能否获得聘用，打扮是否得体比工作资历占的比重更大，但面试官承认外表对他们的选择只有小小的影响。外貌优势不仅表现在招聘决定

上，还表现在薪水上。经济学家检验了美国和加拿大的样本后发现，长得好看的人的薪水平均比没什么魅力的同事的薪水高12%~14%。科学家丹尼尔·哈默梅什（Daniel Hamermesh）就这一主题写了一本书，他估计，在职业生涯中，富有魅力的人能多挣23万美元。哈默梅什向我们保证，他的发现并非出于自夸，因为他说，按10分制的评定标准，“我只能得3分”。

其他实验还证明，长相好看的人更容易在需要的时候获得帮助，在改变听众意见时也更具说服力。也就是说，在我们的文化里，长得好看的人明显占有极大的社会优势。他们更招人喜欢，更有说服力，更频繁地得到帮助。在他人眼里，他们还具备更理想的人格特质，更高的知识水平。而且好看的社会效益从很早就开始积累了，对小学儿童的研究表明、长得好看的小孩子即使做出好斗的行为，成年人也不会觉得他太淘气。另外，教师还相信长得好看的孩子比不好看的孩子更聪明。

这也就难怪外表魅力的光环会为说服专业人士利用了。销售培训课程里总会包含教人打扮的环节，时尚服装商会挑选好看的人充当基层销售员，男女骗子大都长得比较好看。因为我们喜欢漂亮的人，也因为我们容易顺从自己喜欢的人，所以这一切就不足为奇了。

外表魅力会给其他判断带来光环效应，这个想法并不新鲜。早在120年前，列夫·托尔斯泰就说过，“美即为善这一错觉是如此根深蒂固，真叫人难以置信”。有充足的研究成果表明，外表魅力在各种社会、职业和政治领域都有着广泛、即时和先入为主的效应。新近的一篇文献不仅更新了相关的证据，还从进化的角度对这一基本效应进行了解释：我们对富有身体魅力的人所产生的正面感受和有利行为，来自对他们的自动的、笼统的浪漫感觉。

相似性

如果外表占的分数不多呢？毕竟，大多数人的长相都挺普通的。还有其他因素能让人产生好感吗？没错。正如研究者和说服专业人士所知，这样的因素有好几个，相似性就是其中影响力最大的一个。

我们喜欢与自己相似的人，不管相似之处是在观点、个性、背景还是生活方式上，我们从小（小至9个月的婴儿）都有这样的倾向。测量婴儿对与自己相似的东西有好感的研究，借助的是口味偏好跟婴儿类似或不同的木偶（饼干对豆类）。一家在线约会网站对4.21亿潜在的约会对象进行了大规模研究，发现能最准确预测对方好感度的因素是相似度。研究者说：“两个人越是相似，就越有可能发现彼此好看，并选择线下见面。”

因此，一些别有用心的人可以假装在若干方面跟我们相似，有意识地讨我们喜欢、要我们顺从。穿着打扮是个很好的例子。很多研究都表明，我们更喜欢帮助那些衣着跟我们类似的人。

有一项研究显示，我们对跟自己衣着相似的人产生的积极反应，来得是多么不假思索。参加反战示威游行的人更愿意签署跟自己穿着类似的人递过来的请愿书，而且，他们签名的时候往往连请愿书的内容都懒得读一下。好一个“按一下就播放”啊！

请求者还有另一种利用相似点提高好感、增加顺从概率的办法，即假装跟目标人群有着相似的背景和兴趣。例如，训练汽车销售员的时候，公司会要他们注意观察客户旧车上的蛛丝马迹。要是货箱里有野营器材，过一会儿，销售人员或许就该说起自己总是一有空就到远离城市的地方去；要是车子后椅上放着高尔夫球，那就不妨说但愿今天别下雨，因为自己下班还安排了打18洞的球呢。

这些相似之处看起来微不足道，但很管用。在学习到“相似指纹类型”之后，人会更乐于帮助对自己的“指纹模式伙伴”。如果一种产品的品牌名和自己名字的首字母相同，人们会有更大可能购买该产品。一位研究者在进行问卷调查时，对问卷稍微做了一项小小的更改，他把信封

上调查员的名字改得跟接受调查者的名字相仿，而这极大地提高了人们回复调查问卷的概率。也就是说，罗伯特·格里尔收到的问卷来自调查员罗伯特·格瑞格，辛西娅·约翰斯顿接到的问卷来自调查员辛迪·约翰逊。结果，光是加上了这么一点点的相似之处，回复调查问卷的概率就翻了几倍。

就连组织都容易高估名称中包含与自己有关的元素的东西。为了庆祝摇滚乐诞生50周年，《滚石》杂志发布了一份摇滚时代500首最伟大歌曲榜单。根据编辑的考量和权衡，两首排名最高的歌曲是鲍勃·迪伦（Bob Dylan）的《像一块滚石》，和滚石乐队的《无法满足》。撰写本书期间，我比较了其他10份同类的摇滚歌曲榜单后发现，没有一个榜单将《滚石》杂志选中的这两首歌列为第一名或第二名。

不止如此。在学校里，决定青年辅导项目成功与否的最大因素是学生与导师之间是否有着共同的爱好。此外，如果初三学生收到自己与教师之间存在相似之处的信息，那么学生在这些教师的课程里成绩就会有显著的提高。类似地，在谈判中，谈判员一旦了解到自己跟对方谈判员存在相似之处，就更有可能达成协议（哦，你喜欢跑步呀？我也爱跑步！）。那么，不足为奇的是，选民会喜欢跟自己有着类似面部特征的候选人，此外，语言风格（即对话双方所用的词语类型和口头表达方式）和短信风格的相似都提高了浪漫的吸引力，甚至提高人质谈判和平收场的概率（最后这一点尤其令人惊讶）。

由于很小的相似之处也能有效地带来他人积极的回应，又因为编造一个相似之处很容易，所以我建议要特别当心那些声称“跟你一样”又对你有所求的人。老实说，提防跟你看起来相似的推销员是很明智的。好多销售培训项目都敦促学员“模仿和迎合”顾客的身体姿态、语气和口头表达风格（见图3-3），因为这些方面的相似之处全都能带来积极的结果。

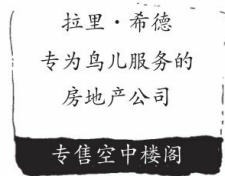


图3-3 叽喳房地产公司

说服专业人士早就摸清了相似性对销售的潜在影响。

我们不妨来看看这些事实：第一，受过模仿顾客言语训练的餐馆服务员能收到更多的小费；第二，接受指导模仿顾客的语言和非语言行为的销售人员能卖出更多的电子设备；第三，不管是美国、荷兰还是泰国的谈判员，经培训模仿对方语言或肢体动作后，均可得到更好的结果。商业同行也不甘示弱，关系顾问关于借助人为共性的建议，收到了良好效果：在速配互动中，经指导模仿约会对象的口头和肢体语言的女性，会被评价为更有身体魅力，进而获得了更多的后续联系请求。^[9]

EBOX | 线上影响力

我相信你自己就曾体会过不少次喜好原则。如果一个请求来自我们的朋友，你会发现，我们更难拒绝来自自己朋友的请求。你可以通过一些简单的技巧让某人喜欢你：经常出现在他们身边以制造熟悉感；指出你们之间的相似之处，模仿他们的行为，帮他们一些小忙，

表现出你喜欢他们。

怎样在线上营销中使用它：使用你的受众自己的语言。团体的内部词汇、短语和俚语效果甚至更好。反过来说，如果你使用了受众并不使用或不理解的词语，那你就在你们之间制造出了距离感，让他们无法与你建立联系。社交媒体和电子邮件是你跟受众互动的完美方式。注意，在最初联系的时候，千万不要对他们提出任何要求，要像对待朋友一样对待他们。

作者点评：网络说服从业人员常常获得的建议是，采用与面对面时同样的影响力实践来提升喜好度。因此，当此类手段出现在电子商务平台上时，我们应有所觉察。例如，营销人员心理学（Psychology for Marketers）网站就建议网络销售人员通过相似和友谊这两点来利用喜好原则。

如果你觉得人为制造共性有违道德，做作的模仿是一种欺骗，我不会反对。渴望得到他人喜欢是人类的基本目标，但实现这一目标并不意味着一定要伪造事实，就像人为制造相似点一样。但话说回来，策略性地让别人喜欢你，甚至通过努力发现并交流和他人真正的相似之处，也未必不妥。事实上，我认为这在很多时候是一种促进和谐互动的途径，是值得称道的。当然，这也并不是一个容易实现的目标，因为人在本性上总是容易注意到区别而非相似之处。

一般来说，人们更乐于寻找并关注不同而非联系。对物理维度的重量和大小也是如此，观察者看到的差异总是多于共性。在互动各方是否存在和谐一致的地方等社会层面上更是如此。莉·汤普森（Leigh Thompson）博士对32项独立的谈判研究进行分析后发现，竞争的谈判人员在很大程度上无法识别和提及双方共同的利益和目标以增加喜好并取得互利结果——哪怕这些共性是真正存在且显而易见的。

这种令人遗憾的倾向，可能正是种族或族裔群体成员与其他同类群

体成员之间存在社会距离的部分原因。他们主要着眼于不同群体之间的差异，而低估了与群体外部人士互动的积极面，从而降低了实际的互动频率。研究者的一系列研究，为这一推论提供了支持。期待与黑人学生对话的白人学生实际参与对话时，会低估对话本身蕴含的真正快乐，因为在此之前，他们太过关注与谈话对象之间存在的感知差异。而在完全相同的实验情境中，研究者事先请另一群白人学生注意与未来谈话对象之间的一切相似之处，情况就大不相同了。策略性地关注真正的相似点，纠正了白人学生在谈话中带入的消极观点。在这种情况下，白人学生现在的积极期待，与他们从黑人学生身上实际获得的积极体验是一致的。

这个结论为我们带来了一种扩大令人满意的交际范围的方法。对看似不一样的人，我们可以寻找并关注跟他们的相似之处，以消除对他们期待过低的错误。[\(10\)](#)

赞美

1713年，作家斯威夫特在一段著名的诗句中宣称：“学校里有句古老的格言，赞美奉承是愚人的食物。”可惜他没有告诉我们，人们是多么渴望吞下这些虚幻的食物。喜剧演员麦克莱恩·史蒂文森（McLean Stevenson）在描述他妻子是怎么“骗”他结婚时，说了一句颇具启发意义的幽默话：“她说她喜欢我。”如今，“喜欢”[\(11\)](#)频繁出现在网络上，并对积极情绪产生着类似影响。在一项大脑成像研究中，研究者发现，当青少年在社交媒体上发布的照片获得大量“喜欢”时，他们大脑中的奖励区域会像圣诞树一样亮起来，而这些奖励区域一般会被吃巧克力或赢钱等令人满意的事件激活。

利用“喜欢我们”这微不足道的信息，人们就能有效地诱使我们还以好感、答应请求。所以，很多时候别人的赞美和亲近只是因为有求于我们。反过来说，人们也乐于得到赞美。当得到餐馆服务员（“这道菜选

得好”)或理发师(“你留什么样的发型都好看”)的赞美时,顾客往往会进行明显更慷慨的消费。同样,在求职面试时,如果求职者在跟面试官的互动中称赞了面试官,他们也可能会从面试官这里获得更有利的评价,并最终赢得工作机会。

就连科技设备也能从表达赞美中受益。执行线上任务的人要是收到计算机发来的奉承反馈,比如“你似乎有一种罕见的能力,能按逻辑方式构建数据”等,就会对计算机产生一种更为正面的情绪,哪怕他们知道这些反馈是程序里预先编制的,并不能反映出自己实际的工作绩效。更值得注意的是,在得到这种空洞的表扬后,人们对自己的表现会更为自豪。很明显,我们总会相信各种各样的赞美,而且还喜欢赞美我们的人(见图3-4)。^[12]



图3-4 赞美会产生自动的吸引力

有趣的是,研究表明,人类表达的赞美会刺激人们产生大量的喜好反应。计算机赞美研究的作者认为,不管是人还是计算机发出的赞美,其结果都缘于同样的心理倾向,所以,设计师应该在软件程序中预置频繁的诸如“你的细心工作令人印象深刻!”或是“好想法!”此类的赞美,哪怕在“评估没有太多依据”的时候也应该这么做。

还记得世界上“最伟大的汽车推销员”吉拉德说他成功的秘诀就是让客户喜欢他吗?他会做一些表面上看起来愚蠢又麻烦的事情。每个月,他都会给自己13 000多位老客户寄送印了字的节日贺卡。节日贺卡的名目每个月都不一样,如新年快乐、情人节快乐、感恩节快乐等,但印在

封面上的信息却从不改变，上面写着“我喜欢你”。吉拉德解释说：“贺卡上什么也没有，除了我的名字。我只不过是告诉他们，我喜欢他们。”

“我喜欢你”这句话被印在贺卡上，每年向13 000多人寄出12次，次次不落，像钟表一样精准。这么一句话，说得丝毫没有人情味，显然是专门用来卖车的。它真的有用吗？吉拉德认为的确有用。既然一个像他这么成功的人都这么想，那么我们恐怕有必要重视起来。吉拉德搞懂了人性中的一个重要事实：我们特别喜欢被人赞美奉承。

北卡罗来纳州对一群男士所做的实验就能说明我们面对赞美的时候是多么容易丢盔弃甲、溃不成军。

影响力研究

参加实验的人会听到另外一个人对自己的评价，而后者需要前者给予帮助。一些被试只听到了积极的评论，一些只听到了消极的评论，还有一些好坏评论都听到了。实验发现了三件有趣的事情。首先，只给了称赞的评估者最为被试们喜欢；其次，哪怕被试完全明白那人拍马屁只是为了讨好他们，也还是最喜欢那个人；最后，和其他两种情况不同，单纯的赞美无须准确。积极的评价，不管是真是假，都能让人产生对赞美者同等程度的喜欢。

既然我们面对赞美会做出这样无意识的正面反应，难怪会被靠赞美交换好感的人利用。从这个角度来看，每年打印、邮寄13 000多份“我喜欢你”的贺卡，似乎就显得不像先前我们认为的那么蠢，那么无用了。

事实上，人们容易受不真诚或别有用心赞美的影响，后来又有许多研究结果支持了这一说法。我自己也并不比别的人做得好。入选某科学

协会之后，我收到了当时所在州的一位议员的贺信，她称赞我“贡献卓越”。尽管我知道贺信是一种为讨好选民而设计的选举策略，但这以后，我还是更喜欢她了。旁观者看到不真诚的赞美，会认为赞美者这么做是出于不可告人的动机。因此，虽然接受奉承的人对真诚或者不真诚的赞美都容易相信，但不真诚的奉承是有代价的，旁观者会注意到其真实目的，并进而讨厌奉承者。

幸运的是，就像虚假的相似性一样，我们并非只有假意逢迎这一个办法可以用。至少，诚实的赞美跟虚假的赞美一样能产生良好的结果。话虽是这么说，但我本人却是一个鲜少表达赞美的人。可能是因为我的成长经历，也可能是出于其他原因，我总是羞于给出合情合理的表扬。数不清有多少次，我在研究会议上评论自己的研究生说：“杰西卡（或者布拉德、琳达、弗拉德、诺亚、查德或罗莎娜）刚才的话真的很有见地。”只可惜，我是在肚子里对自己说的！由于从不曾把赞美从脑海中转移到舌头上，我常常会损失这种转移带来的所有善意。

但我现在不会这样了。我会有意识地跟这种不利倾向做斗争，找出心中萌发的赞许之意，并把它大声说出来。结果让所有人都受益了。这些话的效果太好了，我也开始探索在什么环境下真诚的赞美对表达赞美的人最有利。有一种情况很明显：那就是赞美提升了接受方对自己在某个时候或某种层面上弱点的信心。对此不再赘述。不过还有另外两种策略值得关注。

第一，在值得称赞的人背后称赞他。在研究会议上当众称赞学生的新习惯很适合我的部分原因在于，我是会议的负责人。可在其他很多场合，你可能并不是领导，表扬别人这件事或许也并不适合由你来做。比如开会时老板说了一些你认为很高明的话，如果你当众称赞，可能就会很尴尬，而且显得过于油滑。那该怎么做呢？需要说明的是，我的学生很少会遇到这个问题。但不管怎么说，我有个好的解决办法：到喝咖啡的休息时间或会议结束时，把你的意见告诉老板助理：“你知道吧，

我觉得老板关于×××的那些话，太精彩了。”

这么做可能会带来若干种结果。首先，由于人们希望与他人心目中的好消息联系在一起，并主动进行安排，所以助理很可能会把你的话告诉老板。其次，你并不是当着老板的面说出积极评价的，没有人（不管是旁观者还是老板）会认为你别有用心。最后，出于我们对接受赞美心理机制的了解，老板会相信你（真诚）的赞美，因此也更喜欢你。⁽¹³⁾

第二，找到并就你希望对方奉行不渝的地方做出真诚的赞美。人们接受赞美后会自我感觉良好，并对为自己带来称赞的特质或行为感到自豪。因此，如果别人做了一件我希望他继续做下去的事情，就不妨真诚地表达赞美。如此一来，对方将来会有动力做更多类似的事，以求不辜负我们送上的良好声誉。这一理念与影响力手段“角色设定”（altercasting）不谋而合，即为当事人分配一种特定的社会角色，并希望此人的行为符合这一角色。例如，保险经纪会强调家长的“保护人”角色，以让家长更乐于为自己的家人购买保险。

在为本书做初步研究期间，我偶然目睹了这种策略的影响力。当时，我想跳出实验室研究，转而从社会上了解销售员、营销员、广告商、招聘员和慈善筹款员等说服专业人士的洞见。毕竟，他们的收益取决于其所用手法的成功与否，我相信经过数十年的反复尝试，他们肯定早就找出了成效最好的做法。但我也明白，他们不会轻易把这些宝贵的经验昭告天下。影响力从业人士出了名地爱藏招，最有效的手法他们只留给自己用。

影响力研究

于是，我开始回复他们的广告，匿名参加他们的培训项目。在培训班上，他们总是渴望把自己学到的各种经验教训传授给学员。不出所料，在这些环境下假扮成一个有志从业者，让我接触到了大量原本接触不到的信息。不过我也担

心，一旦我在培训结束后透露自己的真实身份和目的，请他们允许我使用自己收集到的数据，他们会严词拒绝。因为这么做对我来说是百利而无一害，对他们来说则完全相反。

大多数的情况都是，我红着脸，承受着他们怀疑的眼光，承认自己不叫罗伯·考尔德（Rob Caulder），不是真正的学员，而是打算写一本书披露我所收集的信息，还想让他们书面同意我使用他们的专有信息……直到从某一次开始，我多加了一点额外的信息，虽说我那时还不知道它有什么作用。我对这些执业专家说，我是一名研究社会影响力的大学教授，想“在这方面从你们身上学些东西”。通常，他们都会这样问：“你是说，你是个大学教授，专门教授这一主题，而我们是你的老师？”当我向他们保证我所言非虚后，他们往往会挺起胸膛，挥挥手回答说：“没问题，你当然可以分享我们的智慧结晶。”

现在回想起来，我终于明白了他们对我如此宽容的原因。我的“最终告白”让执业者们扮演了教师的角色，而老师是不会囤积信息的，老师只会传播信息。

从此以后，我意识到了该怎样把角色设定手段与真诚的赞美成功结合起来。也就是说，不要仅仅把保护者或老师等角色分配给对方，而是可以诚实地赞美那些表现出乐于助人、富有责任心等值得赞许的品质的人。接着，我们就可以期待将来从对方身上看到更多此类的特质了。研究结果也支持这一假设。如果孩子因为认真负责地完成任务而受到表扬，那在之后的相关任务中，他们就会表现得更认真。类似地，因乐于助人而受到赞美的成年人，也会变得更加乐于助人。

我最近在家里尝试了这一手段。邮递员卡尔几年来一直负责派送我家的报纸，他每天都会开着车经过我家，把卷成一卷的晨报扔在我的停

车道上。大多数时候，报纸都离车道中央很近，不会在左右两侧的草坪浇水系统调整浇水范围时被弄湿。每到假期，卡尔都会在送达的报纸里附上一个写好地址的信封，以提醒我给他寄张感谢他服务的支票，而我也一直照做。但最近，我在寄出支票的同时，还附上了一张便条，夸奖他经常把我的报纸放在不会被弄湿的地方，表现出了很强的责任心。过去，卡尔的报纸卷落在车道中央区域的概率是75%，而今年达到了100%。

这意味着什么呢？如果一个人通常的表现值得称赞（比如认真负责的同事经常早到为会议做准备，乐于助人的朋友经常为你的想法给予有益的反馈），那不光要称赞他的行为，还要赞美他所表现的特质。之后，你很可能会看到他表现出更多同样的特质。[\(14\)](#)

READER'S REPORT | 读者报告

致西奥迪尼先生：

我还在波士顿工作的时候，有个叫克里斯的同事总是给我忙里添活儿。我一般还挺善于推托这类要求的。但克里斯在要我帮忙之前，总会先说上一大堆赞美话。开始他会说：“我听说你做过一个这样的项目，干得特别好，我有个类似的项目，希望能得到你的帮助。”要不就是：“你是个xxx的行家，能帮我解决这件事情吗？”其实我根本不怎么喜欢克里斯，可就在这短短的几秒钟里，我总是会改变主意，觉得他或许是个好人，所以一般都会答应他的请求，帮他的忙。

亚利桑那州的这位MBA学员：

克里斯可不光是个马屁精。他对自己的赞美之词做了巧妙安排，给了听者一顶愿意戴的高帽子。

接触与合作

大多数时候，我们都喜欢自己熟悉的东西。不妨做个小实验来证明这一点。找一张拍下了你正脸的照片底片，然后一正一反冲出两张照片。一张是你实际的样子，一张是你的脸反过来的样子，也就是左右脸换了个位置。现在选一张你更喜欢的照片出来，再让你的好朋友选一下。研究者曾对密尔沃基的一群女性做了这个实验，要是你跟她们一样，你应该会注意到一件怪事，你的朋友更喜欢你正脸的那张照片，而你却喜欢左右脸反过来的那张。为什么会这样呢？因为你们俩都是在积极回应自己更熟悉的面孔，你的朋友平常看到的都是你的正脸，而你每天从镜子里看到的则是左右脸反过来的自己。

通常，我们根本意识不到自己对某种东西的态度是受了先前接触它次数多少的影响。曾有人做了如下实验：

影响力研究

屏幕上飞快地闪过几个人的面孔，因为速度太快，看到这些面孔的被试根本不记得自己见过。然而，一张面孔在屏幕上闪现的次数越多，被试在随后的互动交流中真正遇到这个人时，就越是喜欢他。又因为越是喜欢，社会影响力也就越强。所以，一个人的面孔在屏幕上闪现的次数越多，他的意见观点也就越是容易说服被试。

在一个互联网高度发达、“假新闻”泛滥的时代，人们越发对接触得

最频繁的消息信以为真，这不免令人忧心。正如纳粹宣传部部长保罗·约瑟夫·戈培尔（Paul Joseph Goebbels）那句断言：“谎言重复千遍就成了真理。”而且相关研究发现，哪怕最牵强附会的主张（也就是“假新闻”炮制者最喜欢的那类道听途说），也会随着重复显得更为可信。

有研究显示，人们会逐渐相信自己接触了很多遍的消息。这种效应甚至适用于那些难以置信的主张（“难以置信”正是“虚假新闻”的一大特点）。有人对“谎言重复千遍就成了真理”现象做了总结，将它归结为“流畅”效应，即重复使得一个设想更容易检索、想象和处理，让人在心理上“感觉”像是真理。至于为什么更多地接触某物会让它显得更有价值，另一些研究者承认“流畅”有其作用，但更强调“显著性”（即物品吸引注意力的程度）的作用。

既然我们对自己接触过的东西会更有好感，基于这样的证据，有人建议用“接触”法来改善种族关系。他们认为，只要让不同种族背景的人多跟其他族群平等接触，大家就会很自然地逐渐喜欢上彼此。

很多研究都与这一观点相一致。然而，当科学家对学校的融合教育（这是检验“接触”法的最佳场合）进行考察的时候，却发现了完全相反的模式。黑人白人同校就读，并未减少两个族群之间的偏见，反而起到了反作用。

让我们谈谈学校种族融合这个问题吧。不管那些倡导用单纯的接触来实现种族和谐的人用心是多么良苦，这个方法都不会收到良好的成效，因为它的论点建立在错误的基础上。首先，研究表明，学校环境并不是一个孩子们乐意和其他种族成员交流互动的大熔炉。学校正式取消种族隔离制度之后，社会融合并没什么进展。学生们还是只跟同一种族的孩子玩耍嬉戏，基本上不跟其他种族待在一起。其次，研究表明，即便种族交流的机会更多了，通过反复接触熟悉某样东西，也并不一定会带来更多的好感。事实上，在不愉快的条件下（如挫折、冲突和竞争）持续接触某人或某物，反而会减少好感。⁽¹⁵⁾

典型的美国课堂恰恰孕育了这些不愉快的条件。心理学家艾略特·阿伦森（Elliot Aronson）应得克萨斯州奥斯汀学校管理部门之邀，就当学校存在的问题写了一篇发人深省的报告。他所描述的课堂教育方式，几乎可见于美国的每一所公立学校：

一般而言，事情是这样运作的：老师站在教室前面，提出一个问题；6~10个孩子伸直了背，迫切地朝老师挥手，渴望老师点他们的名，借此显示自己有多聪明；其他的孩子则安安静静地坐着，垂着眼睛，竭力想变成隐形人。老师叫一个孩子起来回答问题的时候，你能看到其他举手学生脸上的失望和沮丧神情，因为他们又错过了一个获得老师表扬的机会；你也能看到其他不知道答案的孩子一脸如释重负的表情……这场比赛竞争激烈，风险极大，因为孩子们的世界里无非只有两三个最重要的人，而他们正在争夺其中之一爱和赞许。

更何况，孩子们保准无法从这样的教学过程里学会如何彼此喜欢和理解。回想一下你自己的经历吧！如果你知道正确的答案，老师却叫了别人，你或许指望起来回答的这个孩子出错，好让你有机会展示自己的知识。如果老师叫了你，你却没回答正确，或是你根本没有举手参与这轮竞争，你恐怕会嫉恨知道答案的同学。在这套体制里，失败的孩子会嫉妒、怨恨成功的同学，说他们是老师的跟屁虫，甚至在操场上用暴力欺负他们。反过来，成功的学生也大多对不成功的学生怀有蔑视态度，说他们是“呆瓜”或“猪脑子”。

这样看来，学校严格地执行种族融合政策，不管是靠校车跨区运送学生、重新划分学区，还是关闭部分学校，总是会恶化而非改善种族偏见，也就不是什么奇怪的事情了。既然孩子们在各自的族群里享受着愉快的社会交往、与他人建立起友谊，却只在竞争激烈的课堂里反复接触

到其他种族的孩子，那么我们也实在不能指望得到什么更好的结果。

那要怎样才能解决这个问题呢？幸运的是，根据教育专家对“合作学习”概念的研究，真正有望消除敌意的方法正一步步浮现出来。课堂种族融合之所以会加剧种群偏见，大多是因为学生把其他族群的成员当成了竞争对手。于是这些教育工作者开始尝试一些新的学习形式，让孩子们通过合作而非竞争来学习。⁽¹⁶⁾

要理解合作法的逻辑，不妨来重温一下土耳其出生的社会学家穆扎费尔·谢里夫（Muzafer Sherif）及其同事在40年前完成的一个有趣的研究项目。出于对群体冲突的好奇，研究小组决定到男生夏令营去做一番调查。男孩们本身并未意识到自己参与了一场实验，谢里夫和同事们不断以巧妙的手法操纵着夏令营里的社会环境，并观察它给群体关系造成的影响。

研究者发现，要让男孩们对彼此产生某种敌意很容易。只要把男孩们分到两个宿舍就足以激发出一种“我们对他们”的感觉；再让男孩们给两间宿舍起个名字（“老鹰”和“响尾蛇”），竞争意识便进一步加剧，男孩们很快就开始贬低对方一组人的素质和成绩。不过，这一阶段的敌意还算不了什么。等研究者有意识地引入一些竞争性活动后，两组人之间的敌意就更深了。宿舍之间的寻宝、拔河、体育比赛，造成了孩子之间的谩骂和对抗。在竞争的过程中，男孩们称对方宿舍的成员为“骗子”“小偷”和“讨厌鬼”。之后，男孩们又频频翻抄对方的宿舍，偷走、烧毁对方的旗帜，张贴威胁性的字条，午餐时打架斗殴。

此时，谢里夫明显看出了问题的症结所在。想引发不和简单得很，只要把被试分组，让他们自发形成小圈子意识；之后，再把他们混在一起，用竞争的火焰烤上一烤。这样不同群体之间的恨意就会像烧开了的水一样沸腾起来。

接着，一个更具挑战性的问题摆到了研究者面前，如何消除眼下双方根深蒂固的敌意呢？他们先试着让两组人加深接触。可即便是进行令

人愉快的联谊活动，如看电影、社交等，最终结果也不尽如人意。去野餐，男孩们为争夺食物打起了架；娱乐活动变成了吵闹竞赛；午餐排队时推推搡搡。谢里夫和研究小组开始担心自己是不是创造出了一种没法控制的“科学”怪物。不过，在争斗进入白热化阶段的时候，他们尝试了一种简单而又有效的策略。他们设计了一系列的环境，在这种环境下，两组人要是继续竞争，每个人的利益都会受损，只有相互合作才对大家都有好处。

在一次郊游中，唯一能载人进城买食品的卡车“坏掉”了。男孩们集合起来，又是拉又是推，直到卡车上了路。还有一次，研究者切断了夏令营的供水管道。夏令营的水是来自远处的蓄水池，要靠管道把水输送过来。面对这场共同的危机，男孩们意识到了团结行动的必要性，于是融洽地组织起来，在夜幕降临之前修好了管道。另一次要求合作的情况则是营地方面告诉男孩们，有一部很好看的电影正在出租，可费用太高，组织者负担不起。男孩们意识到唯一的解决办法就是整合资源，于是凑钱把电影租下来，共同度过了一个美好的夜晚。

这些合作活动的效果虽说过了一段时间才显现出来，却相当惊人。为了成功实现共同的目标而齐心协力，这样的体验慢慢地弥合了两组人之间的裂痕。没过不久，男孩们的口头叫骂就消停了，排队时也不再推推搡搡了，就餐时也开始混着坐了。研究者此时又要男孩们列出自己最好朋友的名单，好多人的单子上都出现了另一组成员的名字。而最初，所有人列的名单上都只有自己这组人的名字。一些孩子甚至为有机会重新评价朋友而向研究者表示感谢，因为现在他们的想法有所改变，跟第一次列名单的时候不一样了。还有一个小插曲很能说明问题。篝火晚会之后，男孩们是搭乘同一辆公车回的宿营地。要是在过去，他们肯定会吵成一团，但这一次却是孩子们主动要求的。公车停在一处饮料摊时，一组男孩还拿出仅剩的5块钱公费，买来冰激凌奶昔款待另一组人。就在不久之前，他们还是互相恨得牙痒痒的仇敌呢！

我们可以找出这种惊人转折的根源，那就是男孩们不再把彼此视为敌人，而是看成盟友。这其间的奥妙又在于研究者为两个群体设定了共同的目标，而实现这些目标需要合作，于是这些竞争群体的成员便不得不把彼此视为理性的同伴、重要的帮手、朋友或是朋友的朋友。等大家通过共同努力成功达成目标之后，任何一个人都很难再以敌意对待这些曾跟自己一同战斗过的队友了。

谢里夫和同事们的课堂研究，得到了另一些研究者的支持。后者证实，从竞争转向合作，是有可能让对手变成朋友的。研究表明，在谈判开始前握手，能提高谈判各方的共同成果，而且我猜如果谈判各方在午餐休息后再次握手，还有可能强化此种效应。尽管有相当多的证据表明，合作式的方法在谈判中比其他形式的人际定位更具优越性，但要是认为合作行为始终是最好或最有效的，那也未免过于天真。例如，要是谈判员在谈判中每隔几分钟就发起握手邀请，那这种策略很容易会让对方产生怀疑，也会给谈判带来不良影响。另一些研究表明，合作学习项目也并非总能带来成功，有时也需要利用竞争的作用。一味地选择合作，有时也会适得其反。

不过总的来说，“合作是解决人类问题的完美方式”这一寓意，在每一种文化的传说里都有所体现。

学校废除种族隔离制度之后反而导致种族局势更为紧张，一些教育心理学家开始注意到课堂教学和谢里夫等人研究的关系。倘若通过修正学习体验，让全班同学至少可以偶尔通过跨种族合作实现共同的成功，那么跨种族友谊或许能找到生长的土壤。好多州都开展了类似的项目，其中以阿伦森及其同事在得克萨斯州和加利福尼亚州采用的方法最为有趣，它叫作“拼图教室”。

拼图学习法的本质是要求学生们一起合作，掌握考试里将会出现的问题。为此，老师先把学生们分成合作的小组，每个学生只获得信息的一部分，即“拼图”的一块。而要通过考试，学生必须掌握全部的信息

才行。这样一来，每个学生都必须互相帮助，互相指导。要想考得好分数，人人彼此需要。跟谢里夫实验里的夏令营成员们必须联手合作才能取得成功一样，学生们也变成了盟友而非敌人。

研究者在新的种族混编课堂上采用这种新方法后，效果相当突出（见图3-5）。研究表明，与同一学校使用传统竞争教学法的其他班级相比，拼图学习让不同种族的同学们结下了更深的友谊，减少了种族偏见。除了敌意的明显减少，少数族裔学生的自尊心、对学校的好感和考试成绩也都提高了。白人学生同样受益，他们的自尊心、对学校的好感亦有提高，他们的考试成绩至少跟传统班级里白人学生的一样好了。



图3-5 拼图教室

研究显示，拼图教室不仅有效地培养起了各种族学生之间的友谊和合作精神，还提高了少数族裔学生的自尊心、对学校的好感和考试分数。

看到拼图教室取得了这么积极的成效，人们很容易对这种方法怀有过高的热情，以为光靠它就能解决一个大难题。但经验告诉我们，要解决这样的难题光靠一个简单的补救办法是远远不够的。毫无疑问，种族融合也是如此。就连合作学习法本身也还存在着不少复杂的问题有待探讨。在我们真正适应拼图法或类似改善学习、加强好感的方法之前，我

们还需要做很多的研究，以确定合作策略的适用范围。它的适用频率是多高？适用规模是多大？它适合哪个年龄段的孩子？它适合哪些群体？如果教师们都愿意采用新方法，我们还需要了解他们采用新方法的最佳途径。毕竟，合作学习法跟大多数教师更熟悉的传统教学方法不一样。它把大部分传道授业的指导工作交给了学生，这有可能对教师在课堂里扮演的主导角色造成威胁。最后，我们还要认识到，竞争也有其积极的一面。它可以激发学生做出恰当的行为，培养学生树立自我意识。因此，我们的任务不是要消除学业竞争，而是要打破它在课堂上的垄断地位，定期采用合作学习法，让各族群的学生都参与进来，并取得成功的结果。

比如，让我们想一想犹太教师拉比·哈伊姆（Rabbi Haim）对天堂和地狱的定义。

地狱：宴会厅里摆满了美味佳肴，但人人饥肠辘辘，因为他们的肘关节被直直地固定着，没法喂东西给自己吃。

天堂：所有条件都一样，只不过，人们互相喂东西吃。

也许这个故事为在课堂上采用合作教学技术提供了一个有用的思路。选择此类技术的出发点，应当是让所有人都有最大的机会获得营养。值得指出的是，一如哈伊姆的故事所描述，最好的合作行为，不光能产生良好的人际感情，还能为共同的问题提出共同的解决方案。例如，研究告诉我们，在谈判开始时主动与对方握手的议价者，提前释放出了自己的合作意向，最终能为各方带来更好的经济结果。

我们跑题跑了这么远，大谈学校种族融合及种族关系问题，有什么特别的用意呢？用意有两点。

第一，虽然接触带来的熟悉往往能导致更大的好感，可要是接触本身蕴含了让人反感的体验，就会适得其反。因此，当不同种族的儿

童被投入标准美国课堂里那种连续不断的严酷竞争中时，我们肯定会看到敌意的加深，事实也正是如此。

第二，有证据表明，以团队为导向的学习能缓解这种混乱状态。通过这一点，我们可以看出合作对喜好过程有着强大的影响力。

在我们认定合作是导致好感的强力因素之前，不妨用一个我眼中的严峻考验来测试一下。说服专业人士是否系统化地使用了合作以令我们喜欢他们，答应他们的请求呢？要是环境中自然地存在着合作关系，他们是否会向我们指出来呢？他们是否会竭力放大原本甚为薄弱的合作关系呢？最重要的一点是，要是不存在合作关系，他们会不会硬生生地蓄意制造呢？

事实证明，“合作”顺利地通过了这场考验。说服专业人士从来都在努力建立一种“我们和他们在为了同一目标而奋斗”的氛围，这样，我们就必须为了共同的利益“团结一致”，同时也让我们明白，他们其实是我们的“战友”。这里可以举出很多很常见的例子，比如新车销售员会站在我们这一边，向老板力争给我们一个优惠的价格。事实上，销售员进了经理的办公室后，根本不会有什么“力争”的举动。通常，销售员很清楚他能给出的价格底线，所以他跟上司甚至话都用不着说。在为本书做调查时，我曾打入一家汽车经销店。在这家店里，情况大多是这样，销售员跑到经理办公室静静地喝上一杯饮料，或者抽上一支烟，而上司照常工作。过上一段合适的时间之后，销售员会松开领带，回到客户身边，露出一脸疲惫的样子，拿出他刚刚“力争”来的好价钱。其实，这个价钱在他走进老板的办公室时就想好了。

还有一个更夸张的例子，能一眼把它看穿的人就比较少了，因为本例中的说服专业人士是警方的审问员，他们的任务是让犯罪嫌疑人如实招供罪行。近年来，法院对警察接触犯罪嫌疑人的方式方法做出了诸多限制，对获取口供的要求尤其严格。过去可以让嫌疑人招供的许多做法现在都不能用了，因为警察们担心整个案件都会被法院驳回。然而，对

于警察们在审问过程中使用微妙的心理学方法，法院并不觉得有什么不合法。基于这个原因，所谓的“好警察-坏警察”这套手法在刑事审讯中用得越来越多了。

“好警察-坏警察”的工作原理如下。假设有个年轻的抢劫犯在听过自己的权利之后一直说自己是清白无辜的，那么他在被带入一个房间里后，房里会有两名警官负责审问他。一位警官扮演“坏警察”的角色，姑且不管是这角色适合他，还是仅仅因为这回轮到了他。犯罪嫌疑人还没坐下来，“坏警察”就对着他来上一大堆“你这个狗娘养的”之类的咆哮。在接下来的审讯里，这个警官叫骂不断；狠踢嫌犯的椅子，加强语气吓唬嫌犯；看嫌犯的时候，用的是看“社会垃圾”的眼神。要是嫌犯反驳他的指责，或是拒绝回答，“坏警察”就会满脸铁青、怒火中烧。他赌咒发誓，说要想尽办法让嫌犯获判最高刑期。他还会说，他在地方检察官办公室有朋友，而那朋友要是知道嫌犯这么不合作，一定会按重罪提起诉讼。

在“坏警察”开始表演的时候，他的伙伴，“好警察”坐在后面，并不怎么说话。但接下来，“好警察”就要逐渐开始插嘴了。他先是试着宽慰“坏警察”，平息后者的怒火：“冷静，弗兰克，冷静。”但“坏警察”吼着说：“这小子当着我的面对我撒谎，别跟我说什么冷静！我痛恨这些说谎的混蛋！”过了一会儿，“好警察”开始帮犯罪嫌疑人说话了：“轻松些，弗兰克，他只是个孩子。”尽管和支持还差得远，但跟“坏警察”的咆哮相比，“好警察”的这些话在嫌犯听来简直像是动听的音乐。可“坏警察”一点也不给他面子：“孩子？他才不是什么孩子。他是个无赖！他根本就是无赖！我还要说，他已经满了18岁了。就凭这个，我就能一脚把他踢进监狱里去，叫他们打着灯笼也找不着这小子！”

现在“好警察”开始直接跟嫌犯说话了，叫他的名字，并指出案件里对嫌犯有利的细节：“我要告诉你，肯尼，你的运气不错，没人受伤，而且你没有携带武器。这样上法庭的时候，你会显得挺不错。”如果犯

罪嫌疑人还是坚持自己无罪，“坏警察”就会开始另一轮的咒骂和威胁。这一次，“好警察”阻止了他。“好啦，弗兰克，”他塞给“坏警察”一些钱，“去给我们弄点咖啡来，一人一杯，买三杯怎么样？”

等“坏警察”走了，就轮到“好警察”演大戏了：“你看，老兄，我也搞不懂为什么我的同事这么不喜欢你。他会想方设法地针对你的。他得到这一点，因为现在我们手里掌握了足够的证据。而且，他说检察官会对不合作的家伙提起最严厉的起诉，那可不是假的。你恐怕会被判上5年，伙计，整整5年哪！我并不想你落个这么惨的下场。所以，要是你趁他还没回来，承认在案发地点抢劫了，我会负责你的案子，给地方检察官说些好话。要是我们合作的话，5年说不定能减成2年，甚至1年。肯尼，帮咱俩一个忙吧。只要告诉我你是怎么抢的，我们就一起想法渡过难关。”这之后，嫌犯大多会一五一十地全部交代。

“好警察-坏警察”的做法之所以管用，有若干原因。靠着“坏警察”的威胁，嫌犯的心里很快就注满了对长期监禁的恐惧情绪；知觉对比原理（见第1章）也发挥了作用，相较于满嘴胡言乱语的“坏警察”，“好警察”显得像是个特别讲道理的好人；又因为“好警察”屡次帮嫌犯说话，甚至还自己掏钱为嫌犯买咖啡喝，互惠原则使嫌犯感到了压力，让他想要回报“好警察”的好意。然而，这种刑讯手法见效的主要原因还是在于它让嫌犯感觉有人站在自己这一边，有人为自己着想，有人愿意跟自己合作。就算在正常的环境下，这样的人也会显得特别好心肠。更何况此时抢劫犯陷入了大麻烦，这样的人简直就是大救星了。用不了多久，在嫌犯眼里，“好警察”就会从大救星变成值得信赖的神父，连所做的坏事都可以向他忏悔了。

条件反射和关联

“为什么他们要怪我呢，博士？”本地电视台的一位气象播报员声音颤抖着给我打来了电话。他对这个问题已经迷惑不解很久了，近来更是

为了它倍感困扰和沮丧。他打电话向我所在的大学心理系求助，想知道谁能解开这个谜，人们便把我的号码给了他。

“我的意思是，这太疯狂了，对不对？人人都知道，我不过是在预报天气，又不是在吩咐天气，对吧？所以，天气糟糕的时候怎么会有那么多人怪罪我呢？去年发洪水的时候，我收到了满怀恨意的邮件！有个家伙还威胁我说，要是我不让雨停下来，就要开枪打死我。老天爷，就为了这个，我现在还提心吊胆呢！连我在电视台的同事们也这样！有时候，就在我做现场直播的时候，他们也会因为热浪来袭一类的事儿嘘我。他们显然知道这跟我无关，可他们还是这么做。你能让我搞懂这一点吗，博士？我真的很沮丧。”

我们约好在我的办公室聊聊。我试着向他解释说，人们很容易觉得事物之间只存在单一的联系，这是一种古老的“按一下就播放”式反应，而他则不幸成了这种反应的受害者。这类例子在现代生活里不胜枚举。为了宽慰这位沮丧的天气预报员，我想起了历史上的一个例子。我要他想想古波斯帝国信使的悲惨命运。倘若信使的任务是传递军事信息，那他一定怀着私心，希望波斯王这一方取得胜利。因为要是他带来的是捷报，到了宫殿后便能享受英雄一般的待遇，美食美酒都任他选。可要是他带来的是失利的消息，结局就完全不同了，他会立刻被杀掉。

我希望天气预报员别误解这个故事的中心思想，也希望他明白一个不管是在如今还是在古波斯都存在的事实。糟糕的消息会让报信人也染上不祥。人们总是自然而然地讨厌带来坏消息的人，哪怕报信人跟坏消息一点关系也没有。光是两者之间存在联系，就足以引发我们的厌恶了（见图3-6）。

天气预报员成了气候变幻莫测的替罪羊

大卫·兰福德 / 文 美联社

电视气象预报员靠着谈论天气能过上蛮不错的生活，可当大自然母亲投出变幻莫测的曲线球时，他们也会成为替罪羊。

这个星期，我跟美国各地几个老资格的天气预报员聊了聊，听他们说了好些有趣的故事：天气不好的时候，会有老妇人用雨伞打他们；他们在酒吧里会遭到酒鬼的挑衅；会莫名地挨雪球和雨靴的“空袭”；收到死亡威胁；被人控诉“假扮上帝”。

大部分天气预报员说，当天的天气预报能有80%~90%的准确率，但长期预报就难说了。大多数人称他们只是照本宣科罢了，真正的信息源头，是电脑、国家气象局或者其他私营机构某个不具名的气象学家。可上电视抛头露面的，毕竟是这些天气预报员。

汤姆·博纳，35岁，在阿肯色州的小石城做了11年的天气预报工作。他记得有一回，在酒吧里，一个从洛诺克来的魁梧农民，喝得醉醺醺地走向他，用指头戳着他的胸口说：“你就是派龙卷风卷走我房子的家伙，我要把你的脑袋拧下来。”

博纳说，他当时到处找酒吧的保安，可一时没见着人影，只好急中生智地回答说：“你说得没错，而且我还要告诉你点别的消息。要是你不

退后，我就再派一场龙卷风来。”

几年前，一场大洪水把圣地亚哥的使命谷给淹了，城里的水足有3米深。一位妇女走到KGTV的麦克·安布罗斯的汽车前边，用雨伞猛砸他的挡风玻璃，说：“这场大雨都怪你。”

印第安纳州南本德WSBT电视台的查克·惠特克说：“有位小个子老太太打电话报警，要他们把我逮起来，因为是我这个天气预报员引来了这么大的雪。”

有个妇女对女儿婚礼碰到下雨感到很不愉快，打电话给纽约州水牛城WKBW电视台的汤姆·乔尔斯，把他狠狠地骂了一顿。“她觉得我应该为这事儿负责，要是她碰到我，很可能会揍我呢！”乔尔斯说。

WJBK电视台的桑尼·艾略特在底特律地区播报了30年的天气。他记得几年前，自己播报了城里要下5~10厘米厚大雪的消息，结果雪下了整整20厘米厚。为了报复，电视台的同事们设计了一套机关，趁他正要播报第二天天气的工夫，往他头上“淋”了两百多只雨靴。

“我说的句句是真，不信你瞧，现在，我脑袋上的肿包都还没消呢！”艾略特说。

图3-6 饱经风霜

请注意到我办公室来的气象预报员和其他同行的共同遭遇。

在一个包含11项内容的系列研究中，研究者发现，被指派大声朗读

坏消息的人会变得不受欢迎。有趣的是，听到消息的人还认为朗读者心怀恶意且能力低下。前文介绍过，一个人诸如身体魅力这样的可取特征可以带来“光环效应”，观察者会认为具备这些特征的那个人在其他各方面也都很好。而现在的情况反了过来：带来坏消息的人似乎创造了一种与光环效应相对的“尖角效应”。传达负面消息让送信人长出了一对魔鬼的“尖角”，在接收消息的人看来，这些尖角也适用于其他许多方面。

我还希望天气预报员能从这段历史故事中了解一点别的东西。他的困境不仅几百年来的“报信人”都遇到过，而且，跟其他一些人比起来，比如古波斯的信使，他算是很幸运的了。在我们会面结束时，他说了一句话，我知道他已经完全明白了这一点。“博士，”他一边说一边往门外走，“我对自己的工作感觉好多了。我是说，幸好我待的地方是菲尼克斯，每年有300天都艳阳高照。谢天谢地，我不是在水牛城播报天气。”

天气预报员临别的一席话说明他彻底搞懂了影响观众对他的好感的原理。跟坏天气联系在一起会带来负面影响，跟好天气联系在一起却能提高他的声望和人气。没错。关联原理是一条普遍性的概念，好坏联系都归它管。不管是好事还是坏事，只要跟我们偶然地联系在了一起，都会影响人们对我们的感觉。

我们对负面联系留下的最初印象，似乎主要是父母教的。还记得他们总是警告我们别跟街上的坏孩子玩吗？还记得他们是怎么说的吗？我们做没做坏事无关紧要，可在邻居眼里，我们只要跟坏孩子玩，就跟坏孩子是一伙的。父母把关联原理带来的负面效应教给了我们，他们说得没错，人们的确有“物以类聚，人以群分”“近朱者赤，近墨者黑”的想法。

正面的关联则是说服专业人士教会我们的。他们不断尝试把自己或自己代理的产品跟我们喜欢的东西联系在一起。你有没有想过，为什么汽车广告里总站着一群漂亮的女模特？广告商希望她们把自己漂亮、性感的积极特性投射到汽车身上。广告商认为，只要漂亮模特跟自己的汽

车联系在了一起，我们对汽车的反应就会变得跟对女模特的反应一样。果不其然，我们的反应正中他们下怀。

有这样一项研究：同一款汽车打广告，一个广告里有性感的女模特，另一个广告里没有性感的女模特。男性普遍觉得前一种广告里的车速度更快、更讨人喜欢、显得更名贵、设计更精致。可事后问起的时候，男人们拒不相信漂亮的姑娘影响了他们的判断力。

类似的例子还有不少，但这方面最有趣的一个证据，大概要数一系列来自信用卡及消费的调查，调查结果证明：关联原理能潜移默化地影响我们花钱的方式。在现代生活中，信用卡具有一种突出的心理特征，它们能让我们立刻享受到商品和服务带来的好处，而几个星期以后才需要付钱。因此，我们很容易把信用卡及其徽记、符号和标志跟消费的积极方面联系起来，而不去想它的负面因素。

消费者行为研究者理查德·范伯格（Richard Feinberg）想搞清楚信用卡及其相关因素对我们的消费倾向有什么样的影响。他在印第安纳州的西拉斐特做了一系列研究，获得了一些有趣又令人不安的结果。

影响力研究

第一项研究发现，在用信用卡付款时，饭店就餐者给的小费更多。第二项研究是让大学生在附有万事达信用卡徽记的房间里翻阅邮购目录，结果这些学生花在邮购商品上的钱平均多出了29%，当然，他们并没有意识到信用卡徽记也是实验的一部分。最后一项研究是要大学生向一家慈善机构（国际联合劝募协会）捐款，倘若他们所在的房间里附有万事达信用卡的徽记，跟在没有徽记的房间里比起来，他们会更乐意捐款。前者捐款的比例是87%，后者仅为33%。从关联原理带来顺从的角度来看，最后这一项发现最令人不安，也最具启示作用。尽管信用卡本身跟慈善捐款并无关系，但

信用卡徽记以及伴随而来的积极关联的存在却促使人们花了更多的钱。研究者又在餐厅里重复了这一做法，结果也是类似的。餐馆服务员用小费盘把账单递给顾客，有一些小费盘上有信用卡徽记，有一些没有。面对有徽记的小费盘，顾客给小费明显要慷慨许多，哪怕他们是用现金结账的。

范伯格所做的后继研究进一步证明，用关联原理来解释上述结果是很有道理的。他发现，房间里的信用卡徽记，只会刺激那些过去对信用卡有正面印象的人多花钱；而那些对信用卡曾有过负面体验的人，比如他们去年支付的利息高出了平均水平，则并未表现出这一倾向。事实上，当后一种人在看到信用卡徽记时，他们的消费反而更趋保守。[\(17\)](#)

由于关联原理的效果如此之好，又如此神不知鬼不觉，制造商们总是急着把自己的产品跟当前的文化热潮联系起来。随着“自然”成为流行的文化概念，“自然”的大部队也是浩浩荡荡。有时候，跟自然的联系根本就不靠谱，如一个大受欢迎的电视广告上说“自然地改变你的发色”。2019年，一群学者对此主题发表了如下看法：

考虑到当今存在的自然产品和服务无比丰富，那些喜欢自然产品的人赶上了好时代。夏天里，人们坐在用第七代自然清洁剂清洗过的地板上，吃着来自阿普尔盖特公司的自然牛肉热狗，夹热狗的面包来自佛蒙特面包公司的纯自然面包，面包上涂着“自然承诺”公司出品的番茄酱和芥末。他们还把热狗跟自然乐事薯片搭配到一起，就着汉森自然苏打水吞下肚。过了一会儿，他们可能会选择抽上一支来自自然美国精神的香烟，看着美国自然草坪公司的技术人员打理自己的草坪。到了夜里，要是消化不良，他们还可以服用“自然好”牌自然胃药。

美国第一次登月期间，从早餐饮料到除臭剂，所有商品都忙不迭地

跟太空项目攀亲戚。而且，这种联系的感知价值经受住了时间的检验：2019年是美国登月50周年，欧米茄手表、IBM和吉米·迪恩香肠都发布了整版广告，宣称自己跟这一著名事件有关。

每逢开奥运会的年份，美国体育代表团指定用什么发胶和纸巾我们都会知道。想获得进行这类关联的权利，代价可不菲。企业会花上数百万美元竞夺奥运会赞助权。而说到宣传自家产品使其跟奥运赛事联系所需的费用，几百万美元最多只算得上个零头。这些企业赞助商卖东西所得的利润比上述数字要大得多。《广告时代》的一项调查发现，1/3受访的消费者都说自己更愿意购买一种跟奥运会联系在一起的产品。

与此类似，1997年美国“拓荒者”号火箭在火星放出探测器后，火星探测器玩具销量猛增，其实这倒没什么好奇怪的，可连“马尔斯”棒棒糖这种跟太空项目根本毫无关系的东西（它的名字来自棒棒糖公司的创始人，富兰克林·马尔斯⁽¹⁸⁾），其销量也来了个大跃进，这就让人感到不可思议了。2016年星球大战系列电影之一《星球大战外传：侠盗一号》（*Rogue One: A Star Wars Story*）上映后，日产汽车旗下的“侠盗”SUV的销量也出现了类似大幅增长（如果没想到是这个原因的话，该车销量的增长简直叫人感到费解）。此外还有一种相关的效应，研究者发现，打着“特卖”的促销标语之所以能增加购买量（哪怕价格并没有降多少），不仅是因为购物者有意识地想到“我买这个可以省钱”，还因为这类标语跟购物者过去所得的便宜价格产生了关联。因此，凡是跟“特卖”标语联系起来的产品，购物者都会下意识地觉得更划算。

把产品跟名人联系在一起，是广告商利用关联原理赚钱的另一种办法。他们付钱给职业运动员，把他们跟运动员角色直接相关（如运动鞋、网球拍、高尔夫球等）或不相关（如饮料、爆米花、连裤袜等）的东西联系起来。对广告商来说，重要的是把联系建立起来，合不合逻辑无关紧要，只要是正面、积极的关联就行了。毕竟，老虎伍兹对别克汽车真能说得上有什么了解吗？

当然，制造商还乐意花大价钱让自己的产品跟当红明星联系起来（见图3-7）。最近，政治家们也意识到，跟名人拉好关系更方便拉选票。总统候选人总是会找来一大堆跟政治无关的知名人物，这些人有些是该政客竞选活动的积极参与者，有些则只是借出自己的名号罢了。就算是州和地方一级的政治造势活动，也要玩这类把戏。这里有个好玩的证据，洛杉矶的一位妇女曾向我表达了她对加利福尼亚州限制在公共场所抽烟政策的公投的矛盾心情：“真难决定啊！有些大明星支持这么做，有些却又反对它。简直不知道该怎么投票了。”⁽¹⁹⁾



图3-7 让产品和积极的因素联系起来

你能发现这则广告将百年灵手表与积极实体联系起来的两种方法吗？第一种显而易见：跟富有魅力的明星挂钩。第二种关联不那么明显，但很可能仍然有效：广告中手表指针的位置使手表看起来像是在微笑。这种微笑的形式会带来各种有利的联想，几乎已经成为所有钟表广告的标准做法。这么做的原因很简单：上述位置的指针会让人在观看广告时体验到更多的乐趣，并对手表产生更强烈的购买意愿。

尽管政治家们素来不遗余力地让自己跟母亲、祖国、苹果派等东西靠拢，但最后一种联系，即跟食物的联系，恐怕才是他们最擅长设计

的。举例来说，白宫一直有个传统，靠一顿美餐来拉到摇摆不定的议员的选票，它可以是一顿室外午餐、一顿丰盛的早宴或是一场优雅的晚宴。总之，每当重要的法案需要拉选票的时候，精致的银质餐具就会摆出来。近来政治筹款活动也照例要吃喝喝。还要注意，在典型的筹款晚宴上，呼吁人们进一步捐款、再接再厉的演说从来不会在餐点还没上桌前开始，而只会出现在宴会当中或众人吃喝完毕的时候。这么做大有好处：一来节省时间，二来利用了互惠原则。但这其中还有最不为人所知的一点好处。来看看20世纪30年代杰出心理学家格雷戈里·拉茨兰（Gregory Razran）在研究中的发现。

拉茨兰把这套手法叫作“午宴术”，他发现，人们对就餐期间接触到的人或事物会更为喜爱。跟我们的论述最相关的一个例子是这样的：

影响力研究

研究者给被试看一些他们从前批评过的政治声明。在实验结束的时候，所有的声明都罗列完毕，拉茨兰发现，被试只对很少的一部分声明改变了看法，也就是那些他们吃饭时过目的声明。被试似乎是在无意识中改变态度的，因为他们根本不记得自己在就餐期间看过哪些声明。为了说明关联原理也适用于不愉快的体验，拉茨兰还在实验里加入了一项条件：在给一部分被试看政治口号时，他会往房间里喷腐烂的气味。此时，这些口号的支持率下降了。新近的研究发现，就算气味轻微得不足以引起人们的注意，但它还是能发挥影响。当看照片上的面孔时，人们要是闻到潜意识里觉得愉快的味道，那么对这些面孔就更容易产生好感；要是闻到潜意识里觉得不快的味道，好感也会随之下降。

拉茨兰是怎么想出午宴术这一招的呢？是什么让他觉得这一套能管

用呢？答案可能跟他在职业生涯里扮演的双重学者角色有关。他不仅是一位受人尊敬的独立研究者，还是最早把俄国开创性心理学文献带入英语世界的翻译之一。这部分文献恰好跟关联原理的研究有关，并主要来自杰出的学者伊万·巴甫洛夫。

巴甫洛夫是一位兴趣广泛的天才科学家，还曾因研究消化系统得过诺贝尔奖，但他最重要的贡献还在于他通过一些极为简单的实验一目了然地论证了他的理论。巴甫洛夫证明，他能让动物冲着一些跟食物完全无关的东西（铃铛）产生对食物的典型反应（分泌唾液），只要把这两样东西在动物的体验中关联起来就行了（见图3-8）。倘若端食物给狗的时候总是伴随着铃铛的响声，过不了多久，狗一听到铃铛响就会分泌唾液，哪怕根本没见着食物。



图3-8 等等，听起来像是食物的味道

图中为巴甫洛夫的一条狗，它嘴上挂着唾液收集管，用来测量它对食物产生的唾液反应，能在多大程度上转移到铃铛声上（经典条件作用）。

从巴甫洛夫的经典示范转到拉茨兰的午宴术，用不着花多大的工夫。显然，对食物的正常反应可以通过原始的关联过程转换到其他东西上。拉茨兰认为，除了分泌唾液，对食物的正常反应还有许多，其中之

一就是赞许、舒服的感觉。因此，把这种愉快的感受、这种积极的态度转向跟美食紧密相连的东西（政治声明只是其中之一）上是有可能的。

值得一提的是，巴甫洛夫的经典条件作用研究先于拉茨兰的“午宴术”研究，前者是后者的基础。后来的研究者又将拉茨兰的成果扩展到被试根本无法察觉的微弱气味。有充分的证据表明，人就和巴甫洛夫的狗一样，很容易受这种策略性配对的影响，而且对此毫无觉察。例如，把一种比利时品牌的啤酒跟愉悦活动的照片，如航海、冲浪和拥抱等重复显示5次，就能提升观众对该啤酒的积极感觉。同样，把一种品牌的漱口水跟美丽的自然风光重复显示6次，会让观众立刻对该品牌产生更多好感，而且这种好感能持续到3个星期以后。在向口渴的观众提供新款饮料之前，向一部分人展示了8次幸福面孔的照片，向另一部分人展示了8次愤怒面孔的照片，统计显示，前者消费了更多饮料，而且在商店购买该饮料时愿意花费的金额也提升了3倍。值得注意的是，这些研究里的被试无一意识到自己受到了配对展示的影响。我们常常在不知不觉间受到单纯关联的影响，但这并不意味着我们不知道它们是怎么发挥作用的。研究表明，人会强烈地倾向于将自己跟好消息联系起来，并竭力避开和坏消息的关联。

从午宴术过渡到说服专业人士的实践应用也并非难事。那就是，各种美好的东西都可以拿来替换食物的角色，把它们讨人喜欢的特质“出借”给人为地跟它们联系在一起的东西，如观念、产品和人民。总而言之，这就是为什么杂志广告里总站着漂亮的模特，电台播音编排师总会在播放热门歌曲前插入本电台的主题音乐。这也就是为什么在特百惠家庭聚会上，妇女们玩“疯狂农场”游戏冲到房间中央拿奖品时，嘴里会不喊“中奖啦”，而喊“特百惠啦”！其实玩家喊的是“特百惠啦”，可公司却“中奖啦”。

我们经常在无意识里成为说服专业人士和关联原理的“受害者”，但我们并非不懂它是怎么回事，事实上，我们自己也经常用它。有充分的

证据表明，我们完全明白古波斯帝国信使或当今气象预报员传来坏消息时面对的尴尬处境。事实上，我们甚至会采取措施，避免自己碰到类似的麻烦。佐治亚大学做过一项研究，看人们碰到传递好坏消息的任务时会如何操作。

影响力研究

实验一开始，等候的学生会接到一项指派的任务，去通知一名同学，说有一通重要的电话。电话有一半的可能带来的是好消息，一半的可能带来的是坏消息。研究者发现，根据消息的好坏，学生传达信息的方式会很不一样。倘若碰到的是好消息，传话的学生保准会提到这一点：“快去接电话，你有好消息啦，快去找实验负责人打听详情吧！”倘若消息不好，传话的学生就闭口不提了：“有电话找你。详细情况你最好是去找实验负责人问问。”显然，学生们早就明白，要想得到他人的好感，必须把自己跟好消息联系起来，而躲开坏消息。

从新闻、天气预报到体育

人们深明关联原理的奥妙，并会努力把自己跟积极的事情联系起来，跟消极的事情保持距离，哪怕他们并非事情的起因。好多奇怪的行为都可以用这一点来解释。这类行为里最怪异的一部分，发生在竞技体育的大世界。不过，这里要说的不是运动员们的行动。毕竟，在激烈的比赛交手当中，偶尔做出些古怪举动无可厚非。更多的时候，暴躁狂怒、失去理性又热情无限的体育迷们才显得怪不可言。欧洲爆发的体育骚乱，南美疯狂的足球迷杀死了球员和裁判，还有本地球迷在特殊日子给本来就很有钱的球员赠送奢侈的礼品。这些事情要怎么理解呢？用理

性的眼光来看，这一切毫无道理。只是一场比赛而已，对不对？

非也非也。体育和狂热粉丝之间的关系，远远不止一场比赛那么简单，这种关系严肃至死。以安德烈斯·埃斯科巴（Andres Escobar）一案为例。埃斯科巴是哥伦比亚国家队的一员，1994年足球世界杯比赛期间，他误将一球送入己方大门。因为这个乌龙球，对手美国队获胜了，本来很被人看好的哥伦比亚队断送了晋级前程。回到家的两个星期之后，埃斯科巴在一家餐馆里被两名枪手打死。一记乌龙球让他付出了生命的代价，而且挨了整整12颗子弹。

因此，我们想要自己支持的运动队赢得胜利，是为了证明自己的优越性，但是我们是想向谁证明呢？当然是向我们自己，也是向其他所有人。根据关联原理，倘若我们能用一些哪怕是非常表面的方式，比如我们的居住地，让自己跟成功联系起来，我们的公共形象也将显得光辉起来。

这一切说明，我们会有意识地操纵我们跟输赢双方的联系，这样，在目睹这些关联的人眼里，我们会显得更好看些。我们展示积极的联系，隐藏消极的联系，努力让旁观者觉得我们更高大、更喜欢我们。我们这样做的方式方法多种多样，但最简单也最常用的一种就是巧妙地选择代名词。举个例子，你有没有注意过，主队胜利后，总有狂热的球迷冲进摄像机镜头，高高地伸出食指，大声叫道：“我们是第一！我们是第一！”请注意，球迷喊的不是“他们是第一”，也不是“我们队是第一”。这里用的代名词是“我们”，意在尽可能地拉拢跟得胜球队的距离和认同（见图3-9）。



图3-9 体育迷们

团队精神可不光是穿着相同的校队队服，阿拉巴马大学的学生们在自己球队获胜后，直接把学校名字写在了身上。

我们还要注意，碰到输球的时候，可不会出现类似的事情。没有哪位电视观众听到球迷大叫：“我们是倒数第一！我们是倒数第一！”要是主队失利，我们最好是跟它保持距离。这时，代名词“我们”就不如疏远的“他们”好了。为了证明这一点，我曾在亚利桑那州立大学的学生中做过一个小实验。

影响力研究

我给他们打电话，要他们描述一下几个星期前校队打橄榄球赛的结果。有些学生被问到的是一场校队输了的比赛；有些学生被问到的则是赢了的比赛。我和同事阿夫里尔·索恩（Avril Thorne）把他们说的话录音，记录学生使用代词“我们”的比例。统计出来的结果很明显，学生会在自己校队获胜的时候使用代词“我们”，把自己跟成功联系起来，“我们打败了休斯敦，17:14”，或是“我们赢了”。可输球的时候，学生就很少用“我们”了。相反，他们会有意识地让

自己跟输掉比赛的校队保持距离，“他们输给了密苏里队，30:20”，或是“我不知道具体的比分，但亚利桑那队输了”。有个学生的评论绝妙地把亲近赢家、疏远失败者的双重愿望暴露无遗。他先是冷冰冰地报出了主队失利的分数：“亚利桑那输了，30:20。”接着他分外痛苦地脱口说道：“他们断送了我们夺取全国冠军的机会！”

不仅体育赛场上存在这种吹嘘自己跟胜利者联系的倾向。比利时大选后，研究者想看看居民要过多久才会取下插在自己门口的支持这方或那方政党的标示牌。调查结果显示，要是标示牌上写的是获胜一方的政党，沉浸在积极关联中的户主保留它的时间会更长。

虽说所有人都或多或少想沾染一点荣耀的光彩，但有些人似乎走得太远了些。是什么样的人呢？倘若我猜得没错，这些人不仅是热情的体育迷，也是一些有着隐性人格缺陷——自我意识太差的人。他们内心深处的个人价值感过低，没办法靠推动或实现自身成就来追求荣誉，而只能靠着吹嘘自己与他人成就的关系找回尊严。我们的文化中有好几类这样的群体。爱跟名人套近乎，每次说起名人总是假装很熟的样子，这是典型的一类。不管采取什么样的形式，这类人的行为都具有一个共同的特征：他们的成就并不来自本身。这可真是可悲啊！

这类人里有些还会以一种略有不同的方式来运用关联原理。他们并不会尽力抬高他人成功与自己之间的可见联系，而是会尽力抬高与自己有着明显联系的人的成功。最明显的例子大概要数那些一心想让自家孩子变成大明星的“星妈”了。当然，这么做的也不光是女性。几年前，在艾奥瓦州的达文波特，一位妇产科医生中止了对三名学校官员妻子的服务，据说是因为他的儿子在学校篮球比赛里没有获得足够多的上场时间。[\(20\)](#)

致西奥迪尼先生：

我是个超级电影迷，又在这一行工作。对我来说，每年最重要的日子就是奥斯卡颁奖之夜。我甚至会把颁奖仪式录下来，重放我喜欢的演员的得奖感言。我最喜欢的一段感言是**1991**年凯文·科斯特纳在《与狼共舞》夺下最佳影片之后说的话。我喜欢这段感言，是因为科斯特纳回应了那些说电影没什么重要性的批评家。确切地说，我真是太喜欢它了，干脆把整段致辞抄了下来，但致辞里有一段话我以前一直搞不懂。科斯特纳对于获得最佳影片奖是这么说的：

“尽管这项奖或许不像世界上的其他事情那么重要，但它对我们来说却始终是最为重要的。我的家人永远不会忘记此刻发生的这一幕；我的美洲原住民兄弟姐妹们，尤其是我的苏族同胞们，永远不会忘记；还有，当年跟我一起上高中的同学们，也永远不会忘记。”

好吧，我明白为什么科斯特纳永远不会忘记这巨大的荣誉，而且我也知道为什么他的家人永远不会忘记，我甚至知道为什么美洲原住民会记得它，因为这部电影就是讲他们的，但我怎么也搞不懂为什么他会提到他的高中同学们。

后来，当我读到体育迷们怎样觉得自己“分沾”了主队球星、球队获胜带来的荣耀时，才明白是怎么回事。所有跟科斯特纳一起上过高中的人，在他得奖之后都会马不停蹄地告诉自己认识的每一个人：我跟科斯特纳是同学！他们觉得自己也分享了这至上的荣耀，哪怕他们没为电影出过一分力。而且，他们这么想也是对的，因为事情就是

这样的。要想得到荣耀，你用不着非得当个大明星，有时候，你只需跟某个明星有点联系就成了。这多么有趣啊！

洛杉矶某电影工作室的这名员工：

这类事情，在我自己的生活里也有不少。我曾经告诉我的建筑师朋友，我跟了不起的弗兰克·劳埃德·赖特（Frank Lloyd Wright），就是那位设计“流水别墅”的建筑师，是在同一个地方出生的。实际上，我本人连条线都画不直，可我总能从朋友眼里看到我想要的那种赞许的反应。“哇，”他们似乎是在说，“你和赖特是老乡？”

如何防范

用很多手段都可以增加好感，而要招架采用喜好原则的说服专业人士，方法反倒十分简单，短短一条就够了。

影响好感的途径多种多样，逐一设计反击策略是毫无必要的，指望这种一对一的策略把每条路都堵死简直不现实。此外，一些导致好感的因素，如外表魅力、熟悉感和关联，都是在潜意识中影响我们的，我们不太可能找出一种合适的防御措施。

我们需要的是一种通用的方法，所有借助好感因素来影响我们做出顺从决策的手腕，都能靠它挡在门外。这种方法的奥妙在于使用的时机。我们不需要识别出所有导致好感的因素，严防死守地不让它们对我们发挥作用。恰恰相反，我建议你听之任之，顺其自然。说服专业人士用来诱使我们产生好感的東西，我们不必提防，只要当心它们带来的过度好感就行。一旦我们觉得自己对说服专业人士的好感超出了该场合下的正常程度，那就到了唤出防御机制的时机了。

把注意力放在效果而非成因上，我们就用不着去辨别、转移针对好感的多种心理影响力了，这本来也是一个近乎不可能完成的烦琐任务。在跟说服专业人士接触的时候，我们只需关注跟好感有关的一件事就行，即我们是不是觉得自己超乎寻常的迅速、热烈地喜欢上了对方？只要发现这种感觉，我们就该警惕了，因为对方可能采用了某种手法，而这时我们就可以采取必要的反击对策。

请注意，我建议使用的策略其实就是说服专业人士自己最青睐的社交柔道术：不去压抑好感因素产生的影响力，听凭这些因素发挥力量，然后用这股力量反过来对付那些想从中获利的人。这股力量越大，其反作用也就越明显，对我们的戒备防御也就越有帮助。

假设我们正跟交易员丹商谈一辆新车的价格。丹是继吉拉德之后有望摘取“最伟大的汽车推销员”称号的谈判高手。谈了一会儿，磋商了一阵，丹想要结束交易了，他希望我们打定买车的主意。在做出这类决定之前，我们首先要问自己一个关键的问题：“认识这家伙才不过25分钟，我是不是有点超乎预期地喜欢他了呢？”如果答案是肯定的，我们或许还会回想一下丹在这期间的表现。我们大概会想起他给我们递了吃的（咖啡、甜甜圈），赞美我们选择的配件和颜色，逗我们笑，帮我们一起对付销售经理，给我们争取更优惠的价格。

这样把事情从头到尾地回顾一番，或许能给你提供不少信息，但这并不是保护自己不为喜好原则所动的必要步骤。一旦我们发现自己对丹产生了超乎预期的好感，我们并不见得非得知道为什么。知道这种好感来历可疑，就足够我们做出反应了。反应之一是逆转这一进程，主动地讨厌丹，但这对他或许不大公平，也不符合我们的利益。毕竟，总有些人天生就讨人喜欢，丹说不定就是其中之一。非要讨厌那些碰巧很招人喜欢的说服专业人士，好像不大对头。再说，为了自己考虑，我们并不想断绝了跟这些好人的商业关系，尤其是他们能给我们一桩划算生意的时候。

我推荐采用另一种反应。倘若我们对关键问题的回答是：“没错，在当前这种情况下，我挺喜欢这家伙的。”那么这应该是一个迅速采取还击策略的信号，是时候在心智上把丹和他销售的丰田或雪佛兰区分开来了。这时你务必要记住，要是我们选择了丹的车，把这辆车从经销商店里开出去的人可是我们，而不是丹。不管出于什么原因，我们喜欢上了丹。他长得好看，他好玩，他对我们的个人爱好感兴趣，他有亲戚住在我们的老家，凡此种种，跟我们是否做出了明智的购车决策完全不相干。

因此我们恰当的反应，就是有意识地把注意力放在这笔生意的好处上，放在丹推荐给我们的这辆车的优点上。在我们做出顺从决定时，把提出请求的人和请求本身从感性上分开，这是很明智的。可一旦我们跟提出请求的人有过亲身的社交接触，哪怕时间十分短暂，也会很容易使我们忽视这其中的区别。在我们对请求者还没什么感觉的情况下，忘掉两者的区别不会造成什么太大的偏差。可要是我们喜欢那个提要求的人，分不清两者的界限，那就有可能酿成大错了。

这就是为什么有必要当心对说服专业人士的过度好感。意识到这种好感，能提醒我们把交易者和交易分开，只根据生意本身的好坏做决定。如果我们都能遵循这样的做法，我敢保证，我们会对与说服专业人士的交易结果更为满意，当然，我猜交易员丹就不那么满意了。

本章小结

- 人们倾向于答应自己认识和喜欢的人提的要求。说服专业人士已经意识到这条原则的威力，他们会通过强调几个能提升个人全面吸引力和好感的因素，来增加他人顺从自己的效力。
- 影响好感和顺从的第一点因素是外表魅力。虽说人们早就

意识到外表漂亮能带来社交优势，但研究表明，这种优势恐怕比我们想得要大得多。外表魅力似乎能造成一种光环效应，把好印象延伸到诸如天赋、善意、智力等其他特点上去。因此，长得好看的人在提请求和改变他人态度方面说服力更强。

- 影响好感和顺从的第二点因素是相似性。我们喜欢跟我们相似的人，更愿意轻率地答应他们的请求。还有一个能带来好感的因素是赞美。一般而言，赞美有助于我们获得他人的好感和顺从。真诚的赞美有两种特别有用的类型，其一是在背后说出来；其二是，让接受你赞美的人获得一种他渴望恪守的声誉，从而继续做出对你有利的行为。
- 通过与他人反复接触增强熟悉感，在正常情况下，这也是一种增加好感的方式。尤其是当双方在积极而非消极的氛围下进行接触时，事情更是如此。互助和成功的合作是最为有效的一种积极氛围。
- 跟好感有关的还有一个因素是关联。广告商、政治家和厂商靠着把自己或自己的产品跟积极的东西联系起来，力求通过关联的过程分享他人积极的观感。其他一些人如体育迷也似乎认识到了这种简单的联系带来的效果，努力在他人眼中把自己跟有利的事情关联起来，疏远不利的事情。
- 要想在做出顺从决定时消除好感带来的不必要影响，有效策略之一是警惕对提要求者的过度好感。只要发现自己在当前情况下对请求者产生了不相称的好感，我们就应该从社会互动中退后一步，在心智上把请求者和他所提的请求分开，而只根据这个请求是否对自己有好处来做出决定。

Influence

第 4 章

社会认同

人们在可以自由行事时，往往互相模仿。

——埃里克·霍弗（Eric Hoffer）

NEW AND EXPANDED

影响力研究

几年前，北京一家连锁餐厅的经理与研究者合作完成了一项研究，旨在找出在既有效又不花钱的情况下增加某些菜品销量的办法。他们想看看，能不能在不降低菜品价格、不使用更昂贵的食材升级菜谱、不新增更有经验的大厨也不聘请顾问对菜品的品质大肆宣传的情况下，让顾客更频繁地选择特定菜品。他们想通过直接给菜品贴上一个标签就搞定这件事。经理和研究者的确找到了一种效果特别好的标签，不过和之前预想的诸如“本店特色菜”或“厨师今日特别推荐”这样的标签不同，他们只是给该菜品标上了“最受欢迎”字样。

结果令人印象深刻。每道菜的平均销量都增长了13%~20%。显然，这些菜因为受欢迎而变得更加受欢迎。值得一提的是，菜品的销量增长，靠的是一种既廉价又完全合乎道德（这些菜品的确是最受欢迎的品目）的说服手法，它便于执行，只是此前从来没有经理采用过。类似的事情也发生在伦敦。当地一家啤酒厂的酒吧与研究者一起进行了一项实验。酒吧入口处的告示牌如实指出，该啤酒厂当周最受欢迎的酒品是波特。统计发现，波特的销量立刻翻了一番。

这样的结果让我很好奇，为什么其他零售商不提供类似的信息呢？在冰激凌店或冻酸奶店，顾客通常可以点完甜品后再自选巧克力粉末、椰子片、饼干渣等配料。既然受欢迎程度的带动力那么强，你或许以为，经理们应该知道贴出海报，说明当月哪些配料或配料组合最受欢迎。很遗憾，他们并没有这么做。其实尤其是对那些不点配料或是只点一种配料的顾客来说，受欢迎度的真实信息会让他们更多地选择加入配料。举个例子，麦当劳的不少门店都提供一种叫“麦旋风”的甜点。当服

务员告诉一组顾客：“来份甜点怎么样？麦旋风是最受欢迎的。”它的销量就激增了55%。当顾客点了麦旋风，店员接着说，“某某口味的配料顾客最喜欢”，那么，顾客购买的额外配料还会再提升48%。值得一提的是，北京的一家连锁餐厅眉州东坡酒楼受此启发，将“最受欢迎菜品”策略纳入了所有的分店，以测量该策略的可行性和影响力度。

EBOX | 线上影响力

奈飞为何要“泄露商机”

虽然不是所有的零售商都理解怎样利用受欢迎度来赚钱，但媒体巨头奈飞（Netflix）从自己的数据中了解到了这一点，并立刻付诸实践。据科技与娱乐记者尼克尔-拉波特（Nicole LaPorte）报道，该公司“素以高度保密观看时长和收视率等方面信息为傲，它高兴地坦言，由于奈飞无须回应广告客户，所以不需要透露任何数据”。但到了2018年，奈飞的态度突然来了个大逆转，它开始大量公布自己最成功的影视产品的相关信息。拉波特说：“奈飞在致股东的信中一口气说出了那些成功影片的名称及其播放量，那感觉就像一名喝醉了的水手接管了一艘平常戒备森严的战舰，一夜间泄露商业机密。”

为什么要这样做呢？因为这时公司高层已经看出，人气会提升人气。首席产品官格雷格·彼得斯（Greg Peters）透露了内部测试的结果。奈飞会根据测试结果告知其会员哪些节目受欢迎，从而这使受欢迎的节目变得更受欢迎。公司的其他高管也很快领悟到了这一奥妙。内容负责人泰德·萨兰多斯（Ted Sarandos）宣称，日后奈飞将更乐于公布“世界各地的人们正在观看些什么内容”。奈飞董事长兼首席执行官里德·哈斯廷斯（Reed Hastings）证实了这一承诺，他表示：“我们才刚开始分享这些数据，而且将一个季度一个季度地朝这一目标迈进。”

作者点评：奈飞高管的这些发言表明，该公司的领导层里没有笨蛋。但萨兰多斯还有一句话让我印象深刻：“受欢迎度是一个人们可以选择使用的数据点.....如果它对会员有帮助，我们就不会秘而不宣。”这里的关键在于，像过去那样对公司产品真实的受欢迎度秘而不宣，不仅无益于创造直接利润，也无益于用户的审慎选择和由此带来的满意度——对公司的长期利润同样没有好处。

社会认同原则

要揭示为什么受欢迎度有这样的作用，我们首先需要了解另一种强大影响力武器的性质：社会认同原则。该原则指出，在判断何为正确时，我们会根据别人的意见行事。这一原则尤其适用于我们对正确行为的判断。特定情形下在判断某一行为是否正确时，我们的看法往往会取决于其他人是怎么做的。因此，广告商乐于让我们知道某种产品“销量增长最快”或“销量最大”，因为这样一来，他们就没必要直接说服我们这种产品的质量有多好，而只需要说其他很多人是这么认为的就足够了。

看到别人正在做，就觉得这种行为是恰当的，这种倾向通常都运作得挺不错。以符合社会规范的方式行事，总比跟它对着干犯的错误少。大多数时候，很多人都在做的事情，也的确是应该做的事情。社会认同原则的这一特点，既是它的强项，也是它的主要弱点。和其他影响力武器一样，它为我们判断如何行事提供了一条方便的捷径，与此同时，选择这条捷径的人，也很容易遭到沿途伺机出手的牟利者的攻击。

问题出在这儿：我们对社会认同的反应方式完全是无意识的、条件反射式的，这样一来，偏颇甚至伪造的证据也能愚弄我们。我们利用其他人的行为来帮助自己判断在某种情况下该怎么做；这完全吻合证据确凿的社会认同原则。但遗憾的是，大多数人在面对明显是牟利者伪

造出来的证据时也自动做出了同样的反应。

例子太多了。一些夜总会的老板在会所里还很空的时候，故意让门口排起长队，为自家夜总会的品质制造可见的社会认同感。销售员也受到指点，要多多提到已经有多少多少客户购买了自家的产品。每晚开始营业前，调酒师常常会在自己的小费罐子里放上几张之前客户给的票子，给后来的客人留下印象，即把钱折起来当小费是酒吧里司空见惯的礼貌行为。出于同样的原因，教会募款员也会在筹款箱里放上一些钱，以期产生同样的积极影响。基督教传教士也有一套广为人知的做法：他们在听众当中会安排“托儿”，到了特定的时间，这些“托儿”就会走上台做见证或捐款。当然，产品评论网站也常常受到虚假评论的污染，这些虚假的评论要么是制造商自己编的，要么是花钱请人提交的。

同样，对音乐下载选择的研究也表明流行会带来流行。在音乐网站上，当一首此前从来没有人听过的歌被研究者随机设定为流行歌，它就变得更加流行。这一结果证实了人们的普遍认识：跟着人群走一般是正确的。

榜样的力量

为什么这些牟利者这么乐意利用社会认同来捞钱呢？因为他们知道，人们在许多情况下都强烈地倾向于认为，如果一件事别人也在做，那它就更正确了。销售兼励志顾问卡维特·罗伯特（Cavett Robert）在对销售学员的建议中准确地总结了社会认同原则：“95%的人都爱模仿别人，只有5%的人能率先发起行动，所以，要想说服别人，我们提供任何证据的效果都比不上别人的行动。”证据简直不胜枚举，我们不妨来看看。

道德：在一项研究中，被试得知大多数同龄人支持军队使用酷刑来获取信息之后，80%的人在公开意见和私人观点里都

表示对酷刑更能接受、更支持了。

犯罪行为：如果犯罪嫌疑人认为其他人经常酒后驾车、在残疾人专用车位停车、在商店行窃、交通肇事后逃逸，那么，他们就有更大的可能性做出上述行为。

个人行为问题：如果当事人认为对亲密伴侣的暴力行为很普遍，那他日后更可能亲自施暴。

健康饮食：在得知绝大多数同龄人都尝试靠吃水果保持健康之后，荷兰高中生的水果消耗量提高了35%（虽然他们秉持典型的青少年态度，声称自己无意改变）。

网络购物：虽然产品推荐不是什么新鲜事物，但互联网改变了游戏规则，它让潜在客户随时都能接触到此前众多用户对产品的评价。于是，98%的网络购物者表示，真实的顾客评论是影响他们购买决定最重要的因素。

支付罚单：肯塔基州路易斯维尔市向收到停车罚单的人发送了一封信，指出大多数此类罚单都在两个星期内被缴清，此举使付款人数增加了130%，让该市停车罚款收入翻了两倍多。

科学建议：2020年新型冠状病毒肺炎疫情暴发期间，日本研究者调查了决定国民戴口罩频率的原因。研究考察了各种原因，如疾病的严重性、保护自己免受感染、保护他人免受感染等，但只有一个原因对戴口罩的频率造成了重大影响，那就是看到其他人戴口罩。

环保行为：如果当事人认为许多人都在通过回收垃圾、节约水电等行为来保护环境，那他们自己也会采取类似的行动。

就环保行为而言，社会认同同样适用于组织。许多国家的政府耗费

了大量资源调控、监管、制裁那些污染空气和水源的企业。可对一些违规公司来说，这些功夫都是白费劲，它们要么彻底无视法规，要么就甘愿支付比合规成本略低一些的罚款。但也有国家设计出了有效的方案，通过发动社会认同引擎来促使企业停止污染。该方案组织各个行业内部对排污企业的环境绩效进行评估后公布分数，让相应行业所有企业都能看到自己跟同行比较起来做得如何。此举带来的整体效果非常明显，改善幅度高达30%，几乎所有的变化都来自那些相对来说最严重的污染大户，因为他们意识到了自己跟同行比起来成绩是多么糟糕。[\(21\)](#)

研究者还发现，基于社会认同原则的手法，在小孩身上就能发挥作用了，有时甚至能取得相当惊人的效果。心理学家艾伯特·班杜拉（Albert Bandura）一直在运用此类手法消除不受欢迎的行为。班杜拉和同事们介绍了一套十分简单的方法，让患有恐惧症的人摆脱极端的害怕情绪。

影响力研究

在初始研究中，班杜拉和同事选出了一些害怕狗的学龄儿童，让他们看一个小男孩快乐地跟狗玩耍，每天看20分钟。通过这样的展示，害怕狗的孩子们在反应上出现了明显变化，仅过了4天，67%的孩子都愿意爬进围栏跟狗玩耍了，等其他人都离开了房间，他们也仍然会继续抚摸、逗弄狗。此外，一个月后，研究者再次测试了孩子们对狗的害怕程度。他们发现，孩子们对狗的恐惧情绪的好转并未随着时间的推移而有所减退，孩子们甚至比从前更乐意跟狗互动了。

研究者又对另一组极端怕狗的孩子做了实验，发现了一项极具实践价值的重要情况：要减少孩子们的恐惧情绪，不一定非得让另一个孩子在现场跟狗玩耍，播放电影片段也具

有同样的效果。若电影片段里有许多孩子在跟狗互动，效果会更好。显然，此时社会认同原则发挥了作用。

关于“人多”的放大作用，我们稍后会进行更多探讨。[\(22\)](#)

READER'S REPORT | 读者报告

致西奥迪尼先生：

我为俄克拉何马州最大的汽车零售商工作，目前面临的最大挑战是如何获得高质量的销售人才。我和同事发现在报纸上刊登广告的回复率很低，所以决定在下班后的通勤时段，在电台上播放招聘广告。我们播放的一则广告重点介绍了顾客对我们所销售车辆的巨大需求，以及有多少人在购买，所以我们需要扩大销售队伍跟进这些需求。如我们所愿，申请加入我们销售团队的人数大幅提升。

但这些广告带给我们最大的影响是客户现场人流量的增加，新车和二手车的销量都在增加，客户态度也有了明显变化。最疯狂的是，当月的销售总量比去年1月增加了**41.7%**！在汽车市场萎缩**4.4%**的同时，我们的业务量却几乎是去年的**1.5**倍。当然，我们的成功可能还有其他原因，比如管理层的变动、设施的更新。可即便如此，不管我们什么时候发布招聘广告，如果在广告中说明我们需要人手来满足市场对我们汽车的需求，那几个月的汽车销量总会出现明显的增长。

俄克拉何马州塔尔萨市丰田汽车经销商的招聘和培训主

管：

消费者对该经销商的汽车需求量大的说法，极大地影响了顾客对其汽车的态度与行动，这跟我们本章所介绍的社会认同原则一致。但这里还有一个重要信息，有助于解释经销商所目睹的现象。需求量大的信息，是招聘销售人员的广告“无意间”透露的。它的大获成功与以下证据相吻合：如果人们认为信息（包括社会认同信息）不是用来说服自己的，就更有可能被这些信息所说服。我敢肯定，如果经销商的广告直接呼吁顾客购买，并说：“人们正在疯狂购买我们的车！你也快来买吧！”它的效果一定没这么好。

末日过后

为了说明社会认同的影响力有多大，这里有一个我最喜欢的例子。它的吸引力表现在以下几个方面：第一，它为参与式观察提供了一个绝佳的样本。所谓参与式观察，指的是科学家亲身涉足某件事的自然发生过程当中，观察整个过程，但这种方法尚未得到普遍应用。第二，这个例子能为历史学家、心理学家和神学家等不同群体提供各自感兴趣的信息。第三，也是最为重要的一点，它表明我们自己，是的，就是我们自己，不是别人，会怎样使用社会认同来进行自我宽慰，把幻想当成事实。

这是一个古老的故事，不妨翻翻从前的数据，它在过去数千年的宗教活动里从来就没绝迹过。各教派和邪教都有过预言：到了一个特定的日子，信奉本教教义的人就能获得拯救，享受极乐。所有的预言都说，这一天会发生一件重要得不容置疑的大事，通常，它指的是世界末日的到来。不过，让信徒们感到绝望的是，事实证明，此类预言统统是假的。

可是，从历史记录来看，预言落空之后紧接着就会出现一种令人难以理解的情形。大多数时候，信徒们并不会因为幻想破灭而如鸟兽散，反倒会越发坚定信仰。他们顶着普通人的嘲笑，走上街头，公开宣讲他们的教条，怀着强烈的热情发展皈依者。总之，哪怕教派的基本教义都成了泡影，信徒的狂热也丝毫不曾减退。公元2世纪的土耳其孟他努教派是这样，19世纪的美国米勒教派是这样，16世纪的荷兰再洗礼教派是这样，17世纪的伊兹密尔沙巴泰教派还是这样。所以，有三名对此感兴趣的社会学家认为，当代芝加哥的一支末日邪教恐怕也会这样。这三名社会学家是利昂·费斯廷格（Leon Festinger）、亨利·雷根（Henry Riecken）和斯坦利·沙克特（Stanley Schachter），他们当时都就职于明尼苏达大学。在听说了芝加哥邪教之后，他们觉得有必要仔细做一番研究，于是他们决定改名假扮成新信徒，加入该邪教做调查，还额外出钱安插了一些观察者到里面去。这样一来，对预言末日到来前后所发生的一切，他们掌握了丰富的第一手资料。

影响力研究

这支邪教的信徒很少，成员从未超过30人。为首的是一对中年男女，为方便日后发表论文，研究者将两人分别称为托马斯·阿姆斯特朗医生和玛丽安·基奇夫人。阿姆斯特朗医生在一所大学的学生健康中心就职，对神秘主义、超自然现象和飞碟一直很感兴趣。所以，说到这些议题，他是邪教组织里受人尊重的权威。不过，基奇夫人才是大家关注的焦点，教派活动也以她为中心。这一年，她开始收到一些来自其他星球神明（被她称为“守护神”）的信息。这些信息，通过所谓的“自动书写”设备从基奇夫人手里源源不断地流出来，构成了该教派宗教信仰体系的主体。“守护神”的教诲是“新时代概念”⁽²³⁾的集合，跟传统的基督教教义有着或多或少的联系，就好像“守护神”在到访北加利福尼亚州时刚好把

《圣经》读了一遍。

“守护神”传来的信息，本来就是教派信徒有诸多讨论和阐释的主题，可当它们预言即将发生一场大灾难，暴发一场始于西半球、最终将淹没整个世界的特大洪水时，它们立刻获得了新的意义。出于可以理解的原因，教众们一开始非常惊恐，可随后的信息宽慰他们说，凡相信基奇夫人所传教诲的人，都能幸免。灾难降临之前，会有太空人出现，用飞碟把信徒带去安全的地方，那地方有可能是另一个星球。有关营救行动的其他细节很少，但信徒们要做好登上飞碟的准备，预先排练特定的口令，如“我把帽子留在家里了”“你有什么问题吗”“我就是自己的挑夫”。要把衣服上的所有金属配件取下来，因为携带金属物品会使飞碟在飞行时“极其危险”。

费斯廷格、雷根和沙克特研究者观察了教众在洪水暴发日到来之前数个星期的准备，注意到他们的行为有两个特别重要的方面。第一，信徒们对邪教信仰体系的投入程度极高。因为觉得就要离开即将毁灭的地球了，信徒们采取了一些无法挽回的举措。大多数信徒的家人和朋友都反对他们的信仰，可这些人却固执己见，哪怕失去旁人的关爱也义无反顾。有好几名成员的邻居和家人甚至威胁要对他们采取法律行动，宣告他们精神失常。比如，阿姆斯特朗医生的姐姐就曾向法院提出诉讼，要求取消阿姆斯特朗医生对两名年幼孩子的监护权。许多信徒辞掉了工作，中断了学业，全职投入邪教活动中。有些人还觉得个人财产很快就沒用了，于是要么将其送人，要么扔掉。这些人分外肯定真理掌握在自己手里，所以顶住了来自社会、经济和法律等各方面的巨大压力。而在跟所有压力对抗的同时，他们对教条的虔诚信奉也随之深化。

第二，洪水到来之前，信徒们的行为是一种古怪的无所作为。毫无

疑问，他们深深相信自家的教义，但他们很少对外宣扬。虽说一开始他们公布了大难即将到来的消息，但并未积极尝试转化别人的观念。他们愿意发出警报，劝告那些主动做出回应的人，但至多如此。

这个小群体不愿招募新信徒的体现多种多样。除了不愿劝说别人，他们还想方设法地保密，如将多余的教义复印件统统烧毁；制定密码和秘密手势；规定部分私人录音带的内容不得和外人讨论，这些录音带非常隐秘，连老信徒都不得对其做笔记。他们想方设法地避免外界关注他们。随着灾难日的临近，聚在该教派总部，也就是基奇夫人家外的报纸、电视台和电台记者越来越多。大多数时候，信徒们都把记者拒之门外，或是置之不理。要是有人问问题，最常见的回答是“无可奉告”。有一段时间，记者们感到很是气馁，但等阿姆斯特朗医生因为搞宗教活动被大学医疗中心开除之后，他们又报复地般蜂拥而至。有个特别固执的记者，甚至还受到了吃官司的威胁。洪水来临的前一天晚上，记者们再度挤过来要信徒们透露信息，照例也被驱散了。事后，研究者总结了教派在洪水到来之前对公众曝光和吸纳新成员的立场：“面对声势浩大的媒体宣传，他们想尽了一切办法保持低调，不为名声所动。明明有无数的机会可以发展信徒，他们却回避、保密，呈现一种近乎超然的漠不关心的态度。”

终于，赶走了所有的记者和想来投奔的信徒，教众们开始为登上预计在午夜到达的飞碟做最后的准备。在费斯廷格、雷根和沙克特眼里，这场面简直是一出荒诞剧。本来是普普通通的人，包括家庭主妇、大学生、高中生、出版商、医生、五金店伙计和他的妈妈，现在却认认真真地参加着演出。他们从两名定期与“守护神”联系的成员那里接受命令，除了基奇夫人的书面信息，那天晚上还有伯莎做补充。伯莎从前是个美容师，通过她的嘴巴，“造物主”下达指示。他们勤奋地排练台词，齐声呼喊飞碟到来前的口号：“我就是自己的挑夫。”“我就是自己的指南针。”此时来了一名访客，自称是“录像船长”，这是当时一出电视剧里的虚构人物，此人说他带来了些口信。信徒们居然认真讨论起这到底是

恶作剧，还是该把他说的话当成救援飞碟捎来的密报。

因为不能带任何金属物品上飞碟，信徒们把衣服上的所有金属零件都取下来了。鞋子上的金属扣眼挖掉了；妇女们要么不戴胸罩，要么就扯掉了胸罩里的金属撑架；男人们使劲拽掉了裤子上的拉链，皮带也不用了，改用绳子系住裤腰。

信徒们去除金属物件的狂热情形，有一位研究者亲身体验过。还有25分钟就到午夜了，他忘了取下裤子上的拉链。这位研究者这样说：“周围的人知道这件事后，立刻恐慌起来。他被推进卧室，阿姆斯特朗医生双手颤抖，眼睛每隔几秒就看看时钟，他用一把刀片把拉链割掉，又用钳子拧掉了金属扣。”紧张的行动结束之后，研究者回到客厅，他身上少了些金属器件，脸色更加苍白了。

很快就要到预言里离开地球的时间了，信徒们安静下来，无声无息地期待着。幸运的是，训练有素的科学家们详细地记录下了这一重要时间段里发生的事情。

最后10分钟，客厅里的信徒分外紧张。他们什么也不做，只是把外套放在膝盖上，坐着等待。在紧张的沉默当中，两座时钟嘀嗒嘀嗒地走动，声音十分响亮，有一座钟比另一座快了10分钟。走得快的那座钟指向了12点5分，一位观察员，即在场的科学家大声地指出了这个事实。人群齐声答道，午夜还没到呢。鲍勃·伊斯曼（Bob Eastman）肯定地说，走得慢的那座钟计时准确，因为他下午才亲手校准过。而从时钟上看，还有4分钟就到午夜了。

这4分钟在一片死寂中流逝，只有一丝杂音冒出来。壁炉架上走得慢的时钟显示，离飞碟预计降临的时间只有1分钟了，基奇夫人紧张地高声叫了一嗓子：“还没有哪次安排落了空！”时钟敲了12下，每一声敲击都清晰得令满怀期待的信徒

们痛苦万分。他们一动不动地坐着。

人们或许指望着出现一些明显的反应。午夜过去了，什么事也没有发生。大洪水还有不到7小时就要来了，但此时却看不见屋里的人有什么反应。没人说话，没有声响。信徒们呆若木鸡地坐着，脸部似乎凝固了，毫无表情。马克·波斯特是唯一还能动弹的人。他躺在沙发上，闭着眼睛，但没有睡觉。过了一会儿，跟他说话的时候，他能哼哼出几句单音节的回答，可还是躺着一动不动。其他人表面上没有什么变化，可后来才知道他们受了沉重的打击。

渐渐地，一种绝望和混乱的气氛在痛苦中笼罩了整个小团体。他们再三检查着预言和相关信息；阿姆斯特朗医生和基奇夫人又重申了信念。信徒们仔细审视了当前的困局，提出了一个又一个的解释，但都不满意。到了凌晨4点，基奇夫人崩溃了，痛哭起来。她抽泣着说，她知道有人开始起疑心了，但整个教派必须向最需要它的人给予支持，人们必须紧紧地团结起来。其余的信徒也渐渐沉不住气了，他们颤抖着，好多人都快哭出来了。现在将近凌晨4点30分了，解决困境的办法却还没找到。这时候，午夜飞碟没能前来营救的事，大部分人都都公开地承认了。教派似乎面临解体。

在一片怀疑的阴云中，信徒们的信心马上就要破碎开来，研究者却接连目睹了两件不同寻常的事情。第一件事出现在凌晨4点45分左右，基奇夫人的手突然开始“自动书写”，记下了天上传来的神谕。她大声地朗读出来，事实证明，这段信息为当晚发生的事情做出了优雅的解释：“小教团，你们独坐了整整一晚，你们散发出的光芒，让上帝拯救了世界，使之免于毁灭。”虽然这个解释简洁有力，但本身还不够令人满意。比方说，听了神谕以后，一名信徒站起身来，戴上帽子，穿上外衣，一走了之，再也没有回来。要恢复信徒们的信心，还需要来点额外

的东西才行。正是为了满足这一需求，第二件值得注意的事情出现了。在场的研究者再一次为读者们提供了生动的描述：

教派里的气氛顿时为之一变，信徒们的行为也跟着变了。宣读了解释预言不中的信息之后，短短几分钟里，基奇夫人就收到了一段信息，要她将该条神谕公之于世。她拿起电话，拨给一家报社。在等电话接通的时候，有人问：“基奇夫人，这是你第一次亲自打电话给报社吗？”她立刻回答：“哦，是的，这是我第一回给他们打电话。以前我从来没什么可以告诉他们的，但现在我觉得事出紧迫。”整个小团体都对她的感觉有了共鸣，因为所有人都觉察出了一种紧迫感。基奇夫人打完电话后，其他成员也轮流给报社、通讯社、电台和全国范围发行的杂志社打电话，告知洪水未曾降临的原因。因为想要迅速而又轰动地把消息传播开来，信徒们把以前严加保密的事情公开了。就在几个小时之前，他们还使劲躲着报社记者，觉得媒体的关注让人痛苦，现在却热心地想要曝光了。

不仅长久以来的保密政策来了个大转弯，教派对转化潜在信徒的态度也完全不同了。之前，对拜访总部、有意投靠的新人，信徒们要么是置之不理、拒之门外，要么就是马马虎虎地敷衍了事。可预言落空之后的这一天，一切都变了。他们接待了所有的访客，回答了所有的问题，想努力改变来访者的信仰。信徒们招纳新成员的愿望空前高涨，最能说明这一点的，是在第二天晚上9名高中生来找基奇夫人谈话的时候。

这几名高中生发现，基奇夫人正跟人在电话里讨论飞碟的事，过了一会儿才知道，基奇夫人认为对方是个外星人。基奇夫人既想继续跟他说下去，又急着招待新客人，于是就让他们也来参加谈话。在一个多小时里，她跟客厅里的客人和电话那头的“外星人”轮流说着话。她太想劝人改变信仰了，简直容不

得放过任何机会。

信徒们在态度上发生了180度的大转弯，原因何在呢？一开始，他们沉默寡言、对上帝旨意严加保密，可仅过了几个小时，他们就成了热心的宣传家，到处传播神的福音。预言里的大洪水根本没来，不信他们那一套的人反而会觉得这个教派及其教条全是闹剧。是什么使他们选择了这样一个完全不合适的时机呢？

一切的关键就发生在“洪水”之夜的某个时候，所有人都越来越清楚地意识到，预言不会成真。奇怪的是，驱使信徒们宣扬其信仰的，并不是先前的确定感，而是一种逐渐扩散的怀疑。他们稍微摸到了点头绪：要是飞碟和洪水的预言根本是错的，那么整个信仰体系恐怕都站不住脚。对蜷缩在基奇夫人客厅里的人们而言，这种很快就要变成现实的前景太恐怖了。

信徒们已经走得太远了。为了自己的信念，他们放弃了太多东西，要是信仰破产了，他们也完了。由此而来的耻辱感、经济损失和旁人的嘲弄，都让人承受不起。从他们自己的话里，就可以看出坚持信仰是信徒们的关键需求。一位带着三岁小孩的年轻妇女说：“我非得相信洪水会在21日来袭不可，因为我已经用光了所有的钱。我辞了工作，从计算机学校退了学，我不能不信。”

在外星人预计到达的时间之后，过了4小时，阿姆斯特朗医生对一位研究者说：“我已经走了好长的一段路了。我几乎放弃了一切，我跟所有人都断了交。我拆掉了回头的每一座桥，我背弃了世界。我绝不能怀疑，我只能相信。事实只有这一个，其他的全是假的。”

想象一下阿姆斯特朗医生及其追随者在清晨来临时陷入了什么样的窘境吧！他们为自己的信仰许下了太多承诺，以至于已经容不下其他的真相了。然而，这套信仰刚刚遭到了现实的无情冲击，没有飞碟降落，没有外星人敲门，预言说的一切都没发生。既然唯一可以接受的真理被

现实证据彻底否定了，教众们要摆脱困境就只有一条路可走了。它必须为信仰的有效性建立另一种证明形式：社会的认同。

这样就解释了他们为什么会从保守秘密变成狂热宣传，同时也说明了他们态度的转变为什么会发生在这个关头。信仰直接遭到否定，最难以说服外人的时候，他们必须冒险直面外界的嘲笑和轻视，因为宣传和吸纳新人是他们唯一的指望。要是他们能传播神谕，告诉那些不知情的人，说服持有怀疑态度的人，并在这个过程中拉拢新人，那么，他们遭到威胁的宝贵信仰便能更加真实。社会认同原则这样说，认为一种想法正确的人越多，持有这种想法的人就越会觉得它正确。教派的任务很明确，既然事实证据无法更改，那就只有改变社会证据了。你能说服别人，自己也必然信服。

不过，大概是因为接下任务时太过绝望，信徒们招募新人的努力彻底失败了，他们没拉拢到一个愿意皈依的人。到了这时候，整个教派在物质和社会上面临着双重失败，很快就解体了。洪水预言破灭后不到三个星期，教派成员就如鸟兽散，只在彼此间维持着零星的沟通。预言的最后一次失灵，就使整个教派毁在了这场从未发生的洪水里，真有些讽刺。

预言失灵的世界末日教派也不一定全是这个下场。要是他们能行之有效地招募到新人，为自己的信仰建立起社会认同，那就还能继续繁荣发展下去。举例来说，荷兰的再洗礼教派预言1533年世界会毁灭，但那一年太平无事地过去了，之后，教徒狂热地招募皈依者，为此投入了前所未有的精力。据说，有个叫雅各布·范·康朋的传教士口才超凡，一天之内就给100人施了洗礼。社会认同的力量如同滚雪球般越来越大，巩固了再洗礼教派的地位，并很快逆转了他们预言落空的不利局面，荷兰大城市里2/3的人口都成了该教派的成员。

最佳条件

本书讨论的所有影响力武器都有适用的条件，在有些条件下效果好些，有些条件下效果差些。倘若我们要保护自己免受这类武器所伤，那么了解它的最佳适用条件，明确我们在什么时候最容易受它影响，毋庸置疑是最为重要的。社会认同原则最适用的条件有三项：当我们不确定怎么做才最好时（不确定感）；当最佳做法的证据来自大量其他人时（从众）；当证据来自跟我们类似的人时（相似性）。

不确定感使人从众

关于社会认同原则最适用的条件，我们已经可以从芝加哥教派的例子中看出一点端倪。信心的动摇，引发了他们吸纳新人的需求。一般来说，在自己不确定、情况不明或含糊不清、意外性太大时，我们最有可能觉得别人的行为是正确的。

对环境不熟悉，是不确定感扩散的另一个原因。在这种情况下，人们特别容易跟着别人走。还记得前文介绍过的北京连锁餐厅实验吗？经理给菜单上的某些菜品贴上“最受欢迎”的标签后，大大增加了这些菜品的点单率。尽管菜品“最受欢迎”标签提高了各类就餐者（男性、女性、各年龄段顾客）的点单率，但有一种顾客最有可能按受欢迎程度来下单，即不常光顾、对餐厅并不熟悉的顾客群体。顾客此时无法依赖现有经验来做决策，便会强烈地倾向于寻求社会认同。

有人光凭这一点洞见就变成了千万富翁，此人名叫西尔万·戈德曼（Sylvan Goldman）。

影响力研究

1934年，戈德曼收购了几家小型杂货店。他注意到，要是顾客觉得手提购物篮太重，就会停止购物。通过这一点，他得到启发，发明了购物车。购物车最早的样子是一把带有

轮子的折叠椅，上面载着一对沉重的金属篮。起初，这套装置太标新立异了，戈德曼的顾客们都不愿使用，哪怕他在店里提供充足的购物车，把它们放在显眼的地方，还竖起告示牌，说明其用途和好处也无济于事。戈德曼十分沮丧，但在打算放弃的时候想出一个通过社会认同原则减少顾客不确定感的办法：他请店员推着购物车在店里四处走。结果，他的顾客们都竞相效仿。这一发明很快就席卷全国，到戈德曼去世以前，他已经变得非常富有了，资产超过了4亿美元。

除了对特定环境的不熟悉，另一种不确定感来自我们对自己在某一问题上的既有偏好缺乏信心。这时，我们同样特别容易受社会认同的影响。我的同事马丁和曼金在拉丁美洲麦当劳餐厅所做的另一项研究可为此提供证据。大多数来到麦当劳的顾客点餐时并不买甜品，这使他们对自己的甜品偏好没有信心。因此，当社会认同信息告诉他们麦旋风是最受欢迎的选择时，他们购买麦旋风的概率便明显提高了。不过，因为大多数麦当劳顾客更熟悉店里的汉堡，也就会对自己选择汉堡的决定很有信心，所以就算告诉他们哪一种汉堡最受欢迎，也不会影响他们对汉堡的选择。

最后，在一项研究中，研究者为被试接入了大脑成像设备，通过观察大脑与感知价值相关的区域——前额叶皮层的背侧区来监测大脑反应。研究者要求被试浏览亚马逊上各种产品的评价，然后发现，如果被试对自己关于产品的观点缺乏信心，就极有可能在看到越来越多他人的评价后向后者靠齐。

READER'S REPORT | 读者报告

致西奥迪尼先生：

有一次，我要坐火车去伦敦找我女朋友。那列火车没能准时发车，而且没有说明原因，我就坐在停着的火车里。不一会儿，站台对面的另一列火车也停了下来。接着发生了一件奇怪的事。一些人离开了我坐着的那列火车，登上了另一列，这引发了一种自发的、放大的效应，让包括我在内的大约**200**人都从我坐着的那列火车走下来，登上了另一列。过了几分钟，一件更奇怪的事情发生了：一些人开始离开第二列车，而整个流程以相反的顺序又来了一次，让包括我在内的每个人回到了原来的火车上，哪怕没有任何通知能证明这么做的合理性。

不用说，这给我带来了一种相当愚蠢的感觉，觉得自己是一只愚蠢盲目的火鸡，追随着社会认同的每一波集体冲动。

这名丹麦的大学生：

在不熟悉、缺乏客观线索进行判断的情况下，人们会产生不确定感。例如，在这个例子中，因为火车站工作人员没有发布通知，所以社会认同便开始接手引导行为，不管这有多可笑。

在审视他人反应，消除不确定感的过程中，我们很容易忽视一点微妙而重要的事实，即其他人有可能也在寻找社会证据。尤其是在局面模糊不清的时候，人人都倾向于观察别人在做什么，这会导致一种叫作“多元无知”的有趣现象。深入理解“多元无知”现象，能帮我们解释一道在全美国频频出现的谜题：受害者迫切需要帮助，全体旁观者却无动于衷。

看客袖手旁观的经典报告，始于《纽约时报》的一篇文章，它在新闻界、政治界和科学界掀起了轩然大波。20多岁的姑娘凯瑟琳·吉诺维斯（Catherine Genovese）深夜下班回家，在其住所所在的街道上被杀害。令人吃惊的是，她的38名邻居只是透过窗户眼睁睁地看着这一切，而没有一个人动动手指打电话报警。这起案件引发了一场全美的骚动，并使得学术界做了一系列的科学研究，以调查旁观者什么时候会在突然事件中出手相助，什么时候不会。近些年来，研究者已经深入探讨了邻居们之所以无动于衷的细节，以及他们是否真的无动于衷。就这一案件而言，研究者甚至揭穿了劣质新闻夸大社会矛盾等无良的报道方法。然而，由于此类事件频繁发生，旁观者在紧急情况下何时出手干预的问题，仍然有其重要意义。答案之一就是“多元无知”效应有能导致悲剧性后果。芝加哥合众国际社发表的一篇报道对此进行了沉痛展示：

警方星期六说，一名女大学生在本市最受欢迎的一处旅游景点，于光天化日之下遭到殴打并被勒死。

星期五，一名12岁男孩在艺术学院围墙边的浓密灌木丛中玩耍，发现了23岁的李·亚历克西斯·威尔森的裸尸。

警方推论，威尔森遭到袭击时，有可能正坐在或站在艺术学院南广场的喷泉附近，而后，行凶者将她拖进了灌木丛中。警方说，她明显遭到了性侵犯。

警方表示，有好几千人从案发地点经过，一名男子曾报告说在下午两点前后听到一声尖叫，但并未深究，因为似乎没有其他人注意到。

很多时候，紧急情况乍看起来并不会显得十分紧急。倒在小巷里的男人，是心脏病发作了，还是只是喝醉了酒？隔壁的喧闹是需要报警的暴力打斗，还是只是无须外人干涉的夫妻俩的小打小闹？到底是怎么回事？碰到这种不确定的情况，人们很自然地会根据周围其他人的行动

来加以判断。我们可以根据其他目击者的反应方式，得知事情到底够不够紧急。

可我们很容易忘记，其他旁观该事件的人恐怕也正在寻找社会证据。因为我们所有人都喜欢装出一副镇定自若、从容不迫的样子，所以我们可能只是暗中瞟着周围的人，不动声色地寻找着证据。这样一来，在每一个人眼里，其他的人全都是镇定自若的，没打算采取什么行动。于是，在社会认同原则的作用下，人们会觉得这起事件没什么紧急的。

科学方法

关于旁观者什么时候会提供紧急援助，社会学家们已经有了明确的结论：只要消除了不确定感，并且目击者们确信出现了紧急情况，那他们就很可能会出手相助。在这种情况下，自己帮忙或叫人帮忙的旁观者数目，都是令人颇感欣慰的。以佛罗里达州的4次独立实验为例。研究者模拟事故现场，叫一位维修工人出演。他们前后做了两次实验，当该名男子明确表示受了伤，需要帮忙时，100%的旁观者都伸出了援手。在另外两次实验里，要帮忙，就有可能会触电，但仍有90%的旁观者伸出了援手。而且，不管旁观者是一个人还是一群人，主动帮忙的概率都很高。

可要是旁观者无法肯定看到的是不是紧急情况，局面就大不相同了（见图4-1）。



图4-1 他们是受害者吗？

类似图中的这种情况，无法肯定当事人是否需要紧急救助，因此就算此人真的需要帮助，人群中恐怕也很难会有人伸出援助之手。想想看，如果你是图中的第二个过路人，恐怕你也会受到第一个过路人的影响，觉得倒在地上的人并不需要帮助。

别让自己成为受害者

以不那么险恶的术语解释了现代都市生活的危险性，并不意味着这种危险就消除了。好在我们对旁观者的“冷漠”有了新的理解，带来了真正的指望。靠着这一科学认识，紧急事件的受害者可以极大地提高自己获得他人救助的概率。这里的关键是要意识到，旁观者群体没能帮忙，不是因为他们无情，而是因为他们不能确定。他们不帮忙，因为他们无法确定紧急情况是否真的存在，也无法确定此时是否需要自己采取行动。只要他们明确地意识到自己有责任插手干预紧急事件，他们便大有

可能采取行动。

既然我们现在知道，“敌人”只是单纯的不确定状态，那么紧急事件的受害者就有可能减少不确定状态，保护自己。设想一个夏天的午后，你在公园里参加音乐会。音乐会结束时，人们开始离场，你发现自己的一只胳膊略感麻木，但你以为没什么好大惊小怪的。然而，当你随着人群朝着远处的停车场走去时，你感觉麻木的范围在扩展，手也僵了，半边的脸也硬了。你困惑不解，决定靠在树下休息片刻。很快，你意识到麻烦大了。坐下来没有丝毫帮助，事实上，你的肌肉控制力和协调力越来越差，你逐渐连嘴巴都难以动弹，快说不出话来了，你想站却站不起来。一个可怕的想法冒了出来：“啊，天哪，我中风了！”人们成群结队地经过你身边，可大部分人都没注意到你。少数几个人注意到了你跌倒在树下的古怪样子，或是你脸上难看的表情，但他们看了看周围的人，想寻找社会证据，却发现他人并没露出关心的样子，于是也觉得没出什么事儿，径直走开了。这样一来你就只能靠自己了。

要是你发现自己正处在这样的困境中，你该怎么做才能获得救助呢？由于你的身体能力正在退化，抓紧时间至关重要。要是你在找到救助之前丧失了说话能力、行动能力，甚至连意识也没了，那你获得帮助、恢复健康的概率必然大幅下降。因此，赶紧得到帮助非常重要。那么什么样的求救形式最有效呢？呻吟、叹息或喊叫恐怕无济于事，它们能让别人注意到你，但其信息量还不足以让路人确信你存在真正的紧急情况。

如果光是呼喊难以引来过路群众的援手，你恐怕应该更有针对性一些。事实上，你要做的不仅仅是吸引别人的注意，你应当清楚地喊出你需要帮助。你不能让旁观者来判断你的情况，不能让他们认为你没什么要紧的，要用“救命”这样的字眼来表现你需要紧急救助的事实，别担心会不会是自己搞错了。这里，尴尬是你要镇压的头号敌人。如果你认为你中风了，那你可没工夫去管自己是不是把问题说得太严重了。你是愿

意承受一时的尴尬，还是愿意就这么死掉，或是终身瘫痪？

就算是高声呼救，也不见得是最有效的手法。它或许可以让旁观者不再怀疑你此刻是否存在紧急状况，但无法消除他人心目中另外几个重要的疑点：你需要什么样的救助？我应该上前帮忙，还是让其他更有资格的人来做？是否已经有其他人去找专业人士来帮忙了，还是该我去找？旁观者怀着这些困惑呆呆地望着你，对你来说生死攸关的时间就这样转瞬即逝。

很明显，身为受害者，除了提醒旁观者你需要紧急救助之外，你还必须多做点什么，你要消除他们的不确定感，告诉他们怎样提供救助，谁该提供救助。那么，什么样的办法才最有效、最可靠呢？

根据我们已经看过的研究结果，我建议是从人群里找出一个人来，盯着他，直接指着他说：“你，穿蓝夹克的那位先生，我需要帮助。请叫救护车来。”这样一句话，消除了一切有可能阻碍或拖延救助的不确定状态。你把穿蓝夹克的先生放在了“救助者”的位置上。他现在应该明白，紧急救助是必要的；他也应当理解，负责提供救助的不是别人，而正是他本人；最后，他还应该很清楚自己要如何提供救助。各项科学证据表明，只要你这样做，就应该会得到快速、有效的帮助。

READER'S REPORT | 读者报告

致西奥迪尼先生：

有一次我从一条灯火通明的大街上走过，看到似乎有人掉到了施工中的地沟里。地沟的保护措施做得很好，我并不肯定自己当真看到了，或许只是我眼花，只是在胡思乱想呢。要是在一年前，我会继续走我的路，以为离得更近的人会看得更清楚。但现在我看过了您的书，所以，我停下脚步，回过头去看看自己是不是真的眼花了。结果真

的是有人出事了，一个男人掉进了沟里，昏迷不醒。地沟很深，所以旁边走过的行人很难发现他。我正努力想做点什么的时候，两个途经此地的人已经帮忙把他拉了上来。

今天的报纸上说，在过去**3**个严寒的星期里，全波兰冻死了**120**个人，我们救出来的那家伙有可能是第**121**个，那天晚上的温度是零下**21℃**。

他应该感谢您的书救了他的命。

波兰弗罗茨瓦夫的这位女士：

几年前，我在十字路口发生了一场严重的车祸。我和对方司机全都受了伤。他倒在方向盘上昏迷不醒，我脚步踉跄地下了车，血滴滴答答地洒在路上。周围的汽车慢慢从我们身边开过去，司机们目瞪口呆，却并未停车。和这位波兰女士一样，我也读过自己的书，所以，我知道该怎么办。我径直指着一辆车的司机：“打电话报警。”又指着第二辆和第三辆车的司机：“停车，我们需要帮助。”不仅他们飞快地过来帮了忙，人们的善意也传染开来，更多的司机自发停了车，去照看另一位伤者。

因此，一般而言，在需要紧急救助的时候，你的最佳策略就是减少不确定状态，让周围的人注意到你的状况，搞清楚他们的责任。尽可能精确地说明你需要什么样的帮助，不要让旁观者自己判断，因为尤其是在人群里，社会认同原则以及由此产生的“多元无知”效应很可能会使他们认为你的情况并不紧急。在本书提到的所有顺从技巧里，这一条恐怕最为重要，必须记住。毕竟，要是没得到紧急救助，说不定你就没命了。

除了这一宽泛的建议外，还有一种形式特殊的不确定状态，只有女性需要在特殊紧急情况下加以消除，那就是女性在公共场合受到男性肢体攻击时的对抗。相关研究者怀疑目击者会因为不确定这两人关系的性质而不加以援助，因为他们觉得，干预情侣吵架是一件费力不讨好的事。为检验这种可能性，研究者让实验对象观看了一场男女之间的公开斗殴。如果没有明确线索表明两人之间的关系，近70%的男女被试都认为这两人是恋爱关系，只有4%的人认为他们完全是陌生人。另一些实验提供了一些线索定义两人之间的关系，比如让女性大叫“真不知道我为什么会嫁给了你”或是“我不认识你”，研究者发现旁观者群体产生了并不乐观的反应。虽然打架的严重程度一样，但旁观者不太愿意帮助已婚女性，因为他们认为这是夫妇之间的私事，妄加干涉只会让大家都觉得尴尬和不快。

因此，一名女性如果与一名男性发生肢体冲突，绝不应该指望通过简单的喊叫来获得旁观者的帮助。旁观者很可能会把这件事看成家庭争吵，因而觉得出手帮忙会不合时宜。幸运的是，研究数据为解决这一问题提供了一条思路：要大声地给攻击者贴上“陌生人标签”——“我不认识你！”，如此一来，女性获得帮助的机会将大大增加（见图4-2）。⁽²⁴⁾



图4-2 想得到帮助，你就必须正确地大喊

一男一女发生冲突，旁观者通常会认为这两人是恋爱关系，外人出手干预没有必要，或是不恰当。为了消除这种感知，获得帮助，当事女性应该大喊：“我不认识你！”

有样学样

正如我之前所说，跟其他影响力武器一样，社会认同原则也有最适用的条件。我们已经探讨了这些条件中的一个：不确定感。毫无疑问，在人们不确定的时候，更容易根据其他人的行为来判断自己该怎么做。除此以外，还有一个重要的适用条件：人多。同一种行为，做的人越多就显得越正确。如果你怀疑这一点，就不妨来做个小实验（见图4-3）。请站在川流不息的人行道上，找一处天空或高层建筑，凝视它整整一分钟。这期间可能不会发生什么，大多数人会头也不抬地从你身边走过，几乎没人会停下来跟你一样凝视高空。第二天，请你到同一个地方，带上4个朋友跟你一起抬头望天。60秒内，肯定会有一群路人停下来，跟你们一起朝上看。行人就算不跟你们一起发傻，也抵挡不了暂时朝上看一看的压力。纽约的3名社会心理学家曾经做过这个实验，要是你尝试的结果跟他们一样，那你和朋友们能使80%的路人一起抬头朝那空空如也的地方看。而且，如果人数多到一定的程度（20人左右），那你带来的朋友越多，加入的路人就越多。近半个世纪后，牛津的调查人员观察到了跟纽约市人群里抬头望天的研究类似的模式。

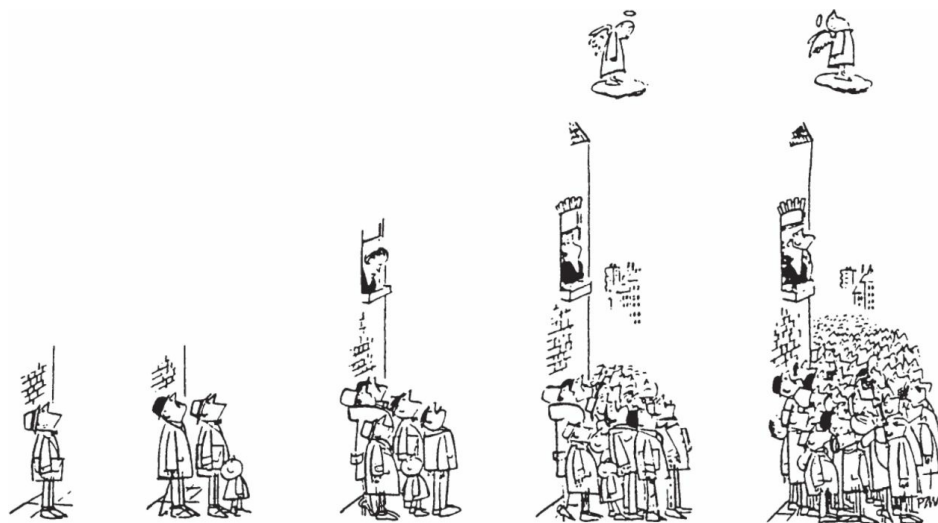


图4-3 追随人群“向上看”

群体的吸引力太强大了！

社会认同信息不一定非得是视觉形式才能影响人们。让我们来看看，在大型歌剧这一人类最古老、最庄严的艺术表现形式的历史里，从业者是怎样大肆利用社会认同原则的吧。他们用的就是所谓的“捧场”现象。

影响力研究

“捧场”是1820年巴黎歌剧院的一对常客首创的，这两人名叫索通和波奇尔。不过他们可不光是歌剧看客，还是生产掌声的商人，并且知道怎样构建社会认同来诱发更多的掌声。

他们成立了一家叫作“歌剧演出成功保险公司”的组织，向希望得到观众积极反馈的歌手和剧院经理提供服务。索通和波奇尔极为成功地靠着通过预先安排的反应激发出了观众的真实反应，没过多久，“捧场”（一般会有个“首席喝彩”，再加上几个附和的）就变成了整个歌剧世界的惯例和传统。

音乐历史学家罗伯特·萨宾（Robert Sabin）指出：“到了1830年，捧场制度达到全盛时期。各家机构白天收钱，晚上鼓掌，一切都在光天化日之下进行……但不管是索通，还是他的盟友波奇尔，都没想过自己设计的这套付费鼓掌的做法，会在歌剧界蔓延到如此轰轰烈烈的程度。”

随着捧场制度的发展，从业者们开始提供形式和强度不一的各种服务项目。有能一声令下说哭就哭的“哭娘”；有能用狂喜的声调高喝“再来一个”的“喝彩人”；还有如今录制笑声音轨的演员的老祖宗，笑声极具感染力的“笑匠”。不过，在索通、波奇尔及其接班人所创造的商业模式里，我们可以观察到最具启发意义的一点：散坐在观众里的捧场客越多，给人留下“还有许多人喜欢这场演出”的印象越强，也就越具有说服力。所以，托儿们，喝彩吧！

同样地，当今政治事件（比如美国总统辩论）的旁观者也有可能受到观众反应程度的极大影响。在美国总统竞选辩论中，观众对候选人表现的感知对选举结果至关重要，政治学者也注意到了辩论的重大影响。为此，研究者调查了造成辩论成功或失败的因素。其中一个因素是，辩论现场观众的反应，会影响远程观众（一般是通过电视、广播或流媒体⁽²⁵⁾视频）的反应。研究者展示候选人真实的表现，但借助技术调整了现场观众的反应（鼓掌、欢呼、大笑），进而考察远程观众对候选人的看法会受这些人为调整过的反应怎样的影响。研究结果显而易见：从辩论表现、领导才能和好感度的角度来看，在1984年里根对蒙代尔的辩论中、在1992年克林顿对布什的辩论中，哪个候选人收获了现场观众最强烈的反应，就能赢得当天远程观众的心。一些研究者甚至有些担心总统辩论中出现的一种倾向：候选人故意在现场辩论的观众里插派喜欢大声闹腾的追随者，这些人的热烈反应会给人一种大厅里的支持度大于实际情况的印象。只要允许，“捧场”的做法远不会消亡。

致西奥迪尼先生：

读到关于社会认同原则这一章时，我想起了一个有趣的例子。在我的祖国厄瓜多尔，你可以雇用一个人或一群人（传统上均由女性组成）来出席家庭成员或朋友的葬礼。这些人的工作是在死者下葬时号啕大哭，其目的是让更多的人哭起来。几年前，这种工作相当热门，干这行最出名的人会得到“乌拉纳斯”（**Uoronas**）的称号，也就是“哭丧人”的意思。

中美洲的这位营销高管：

我们可以看出，在不同的时代和不同的文化中，都有人可能利用伪造的社会证据来渔利。在今天的情景喜剧节目里，我们不再需要“捧场人”来引导观众笑得更大声更用力了。但我们会用音频技术工作人员来放大录影棚里观众的笑声，让节目的喜剧效果在真实观众（也就是像你我这样的电视观众）里显得比实际上更可乐。说起来可悲，我们挺容易受这类把戏愚弄的。实验表明，使用加工过的欢乐效果让观众笑得更频繁、更持久，并认为幽默素材更有趣。

为什么“人多”的效果这么好

几年前，英国埃塞克斯的一家购物中心遇到了一个问题。在正常的午餐时间，它的美食广场会变得非常拥挤，顾客们需要等待很长时间，餐桌也不够用。商场经理们向一组研究者寻求了帮助。这组研究者安排了一项研究，根据“人多”带来的心理吸引力提供了一套简单的解决方

案。这套解决方案包含了社会认同原则的三大因素：有效性、可行性和社会接受度。

在研究中，研究者制作了两张海报，敦促购物中心的顾客早早前往美食广场享用午餐。其中一张海报上是一个人这样做，另一张海报则将一个人换成了好几个人。事实证明，提醒顾客尽早去吃午餐的第一张海报很成功，美食广场中午前的顾客活跃度增加了25%。但真正成功的是第二张海报，它使美食广场中午前的顾客活跃度提高了75%。

有效性 遵从身边大多数人的建议或行为，往往被视为做出良好决策的捷径。我们往往会利用他人的行为来做出正确选择并加以验证。如果人人都对一家新餐馆赞不绝口，那它大概也会是我们喜欢的好餐馆。如果大多数在线评论都在推荐一款产品，那我们也或许会更干脆地买下它。以购物中心的海报为例，当顾客看到一张许多人提早吃午餐的照片时，似乎就更容易顺从，并认为这是个好主意。更多的研究表明，如果广告表现出喜欢某个品牌的消费者的比例越来越高（“7个人里有4个”、“7个人里有5个”或“7个人里有6个”），那么偏好该品牌的旁观者也会越来越多。出现这种情况的原因是，旁观者认为，受较多消费者喜爱的品牌，肯定是正确的选择。此外，统计数据表明，广告效力越强，人们对广告品牌的偏爱程度就越高。

通常，借助他人的选择来做决策并不需要复杂的认知操作，这个过程可以更加自动化。例如，果蝇没有复杂的认知能力，当雌果蝇看到另一些雌果蝇跟研究者做了特殊染色（粉色或绿色）的雄果蝇交配，它们就会更乐于选择另一只同样颜色的配偶，这一概率高达70%。无须认知指导就对社会认同做出反应的可不只有果蝇。著名旅行作家道格·兰斯基（Doug Lansky）承认，在参观英国皇家阿斯科特赛马会时，他看见了不列颠王室就准备拍上一张照片。“我的镜头对准了女王，以及坐在她身边查尔斯王子和菲利普王子。可我突然想到：我干吗要拍这张照片？世界上又不缺王室照片。没有哪家小报会为这张照片付给我大价

钱，而且我也不是狗仔队。但我周围的快门如同冲锋枪一般‘哒哒哒’地响了起来，我也加入了这股狂热之中，完全是情不自禁。”

让我们继续留在英国，从具有启发性的历史案例中，看看“人多”在验证选择和引发传染性效应上的力量。几个世纪以来，人们都遭受着各式各样非理性的癫狂、狂躁和恐慌的折磨。在经典作品《大癫狂》一书中，作者查尔斯·麦凯（Charles MacKay）列举了在该书1841年首次出版之前发生的数百件事。大多数事件都有一个极具启发性的特点——传染性。他人的行为传播给了旁观者，旁观者因此做出行动，进而在其他旁观者眼里验证了该行为的正确性，于是后者也相应地做出行动。

1761年，伦敦经历了两次中等规模的地震，时间恰好相隔一个月。这一巧合使得一个名叫贝尔的士兵坚信，一个月后的同一天将发生第三次规模更大的地震，于是他便开始散布自己的预言：4月5日，伦敦将毁于一旦。起初，没有几个人拿它当真。但相信他预言的人却开始采取预防措施，把家人和财产转移到周围地区。看到这一小群人逃亡，其他人也相继效仿。在接下来的一个星期里，人们一拨一拨地出逃，引发了近乎恐慌的一场大规模疏散。大量伦敦人拥入周边村庄，花了天价争抢住所。按麦凯的说法，惊慌失措的人群里，有不少“一个星期前还在嘲笑这一预言的人，但当他们看到别人这样做，也收拾好行装，慌张出逃”。

等4月5日平安过去后，逃亡的市民们回到伦敦城里，对误导自己的贝尔异常愤怒。可麦凯的叙述清楚地表明，他们的愤怒找错了对象。疯子贝尔并没有什么可信性，只是伦敦人自己向彼此证实了贝尔的说辞。

还有其他证据可以证实社会认同的有效性。在一项研究中，研究者对一群6~11岁的孩子说，参与研究的其他孩子吃了很多胡萝卜。听到这一消息后，这些孩子也吃了更多的胡萝卜——因为这个消息让他们相信，吃胡萝卜是个好选择。一项在线消费者选择实验也表现出了类似的结果。如果参与者知道一款红酒已经售出了总量的2/3，那么跟那些得

知同一款酒只售出了总量的1/3的人比起来，前者更乐意购买此种酒。为什么呢？因为消费者觉得，如果销量可观，那红酒的质量也一定好。

EBOX | 线上影响力

“可怕的”白色货柜车司机

因社会认同引发的毫无根据的恐慌，我们用不着到18世纪的英国去找例子。实际上，由于互联网的特点和能力，我们完全可以从身边看到无数个例子。

2019年底到2020年初，令人担忧的谣言在网上迅速传播，该谣言说有驾驶白色货柜车的男子绑架女性进行性交易，甚至出售被绑女性的身体器官。社交媒体巨头Facebook的算法会使被广泛分享或成为热门话题的帖子出现在其页面最明显的位置。于是，这个始于巴尔的摩的故事滚雪球般传遍了美国乃至世界各地。结果，很多城市的白色货柜车司机报告说，随着谣言在自己社区传播开来，他们受到了居民的威胁和骚扰。一名白色货柜车司机在Facebook的一篇帖子里被攻击，最终丢了工作。另一名白色货柜车司机则被两名男子枪杀，而后者声称自己是在面临绑架威胁时才开的枪。即便如此，当局也从未发现过哪怕一起与白色货柜车司机传言相关的案件。

但这不重要。巴尔的摩市市长伯纳德·扬（Bernard Young）被故事影响，向该市女性发布了一则令人不安的电视警告：“不要把车停在白色货柜车附近。一定要随身带好手机，以防有人绑架。”伯纳德市长的这些说法有证据吗？他自己所辖市内的警察可没这么说。

可他却说，“Facebook上到处都是这样的传言。”

作者点评：这生动地说明，人们相信谣言往往是因为心中毫无根据的恐慌，而且谣言通过人们频繁的社交媒体动态变得更具传染性。很多时候，人们认定真相靠的并不是实体证据，而是社会认同。

有句老话很切题：如果一个人说你长了尾巴，你会笑那人说蠢话；可要是有三个人都这么说，你会忍不住摸摸自己的屁股。

可行性 如果我们看到很多人在做某件事，这并不意味着它或许是个好主意，还意味着我们或许也可以这么做。就英国购物中心的例子而言，游客们看到一张许多人提前吃午餐的海报后很可能会对自己说：“这个想法似乎可行。重新安排一下购物计划或工作时间去提前吃午餐，好像也不是什么大不了的事。”所以，除了感知有效性，“人多”会影响人们决策的另一个原因在于它传达出了可行性：如果很多人都能那么做，那它应该不难做。一项针对意大利几座城市的居民的研究发现，如果居民相信自己有很多邻居会在家里回收垃圾，那他们自己也会更乐于回收垃圾，因为他们看到回收垃圾并没有那么难。意大利居民回收生活垃圾意愿的数据，具体来自卡利亚里、特尔尼和马科梅尔等城市。

影响力研究

在一些值得尊敬的同事的引导下，我曾做过一项研究，想看看什么样的说辞能最好地动员人们节约家庭能源。我们在一个月内，每星期向住户发送四条号召减少能源消耗的信息。前三条信息中包含着常见的节能理由：造福环境、对社会负责、节省电费。第四条信息打的是社会认同牌：社区内的大多数居民都在努力节约能源。我们在月底进行统计后发现，以社会认同为基础的信息带来的节能量，是其他信息的3.5倍。二者差异之大，让参与研究的我和同事乃至业主都吓了一跳。事实上，业主们原本以为基于社会认同的信息应该是效果最差的。

我向政府部门的官员们汇报了这一研究，可因为“人类行为的最强动力来自经济利益”这一根深蒂固的观念，他们总是不相信我的结论。他们说：“得了吧，跟业主说他们的邻居比他们节能3倍，居然比说节能可以大幅减少电费更管用？这叫我们怎么相信呢？”面对这一合理的质疑，能予以回击的方式不少，但有一种对我来说说服力最强，即基于社会认同的信息效果好的第二个原因：可行性。如果我告诉业主，节能降耗可以让他们省下不少钱，这并不意味着他们真的能做到。毕竟，如果把家里所有的电器都关掉，在黑屋中的地板上蜷缩一个月，我能把电费降到零。可我并不愿意这么做。基于社会认同的信息的一个优势在于，它解决了“能否实现”这一不确定的问题。如果人们得知许多跟自己差不多的人都节约了能源，那这么做无疑是可行的，因为它显得现实且具有可执行性。

值得一提的是，这个节能研究是我和同事们在加利福尼亚州圣马科斯的中产阶级社区进行的，除了上述结果，我们还发现了另一个值得一提的点。我们的研究包括两个对照组：一组居民收到敦促他们节约能源的信息，但该信息并未说出明确的理由；第二组居民则没有收到任何信息。结果表明，这两个对照组随后的能源使用情况并无不同。换句话说，光是劝诫人们节约和什么都不做，效果是一样的。人们需要付诸行动的理由。当然，最重要的问题在于，“什么原因的动力效果最好”。在我们的研究中，家庭节能最具说服力的原因是，大多数邻居都在这么做。

社会接受度 成为“众人”中的一员，我们会感觉更被社会所接纳。原因很简单。还是就英国购物中心的研究来说。顾客看到的海报内容，要么是一个购物者到购物中心的美食区提前吃午餐，要么是多个购物者这么做。效仿第一张海报，旁观者有可能会承受社会不赞成的风险，被视为“独狼”、怪人或局外人。效仿第二张海报的则恰恰相反，它会让旁观者因为感觉置身在人群中而觉得自如。两种体验之间的情绪差异非常明显。较之保持与集体相同的观点，持有出格观点会使人产生心理上

的痛苦。

在一项研究中，被试的大脑被接入一台大脑扫描仪，同时从其他人那里接收与自己意见相左的信息。冲突信息有时来自其他4名被试，有时来自4台电脑。尽管被试认为两种判断同样可靠，但如果冲突信息来自其他人而非计算机，他们就更愿意顺从。如果被试认为两种信息源的可靠性相同，那是什么使得他们更乐于顺从其他被试的选择呢？答案在于他们违背他人共识时会出现的生理反应。此时，接收冲突信息的被试大脑中与负面情绪相关的区域——杏仁核会被激活，这就是研究者所称的“独立之痛”。违背他人似乎会产生一种痛苦的情绪状态，从而强迫被试顺从。而违背一组电脑就没有同样的行为后果，因为它没有相同的社会接受后果。说到群体动力，老话说得好：“要想处得好，就得跟着走。”

例如，心理学家欧文·贾尼斯（Irving Janis）记录了一群老烟民去诊所接受治疗的事。在小组第二次见面时，几乎所有人都表示，抽烟太叫人上瘾了，没人能一次性戒烟。但一名男子反驳了小组的观点，他声称自己从上个星期加入小组以来，彻底戒了烟，所以其他人也能做得到。此话一出，他从前的“战友”们禁不住联合起来反对他，并对他的立场发起了一连串愤怒的攻击。到下一次开会时，这个发表不同意见的人报告说，斟酌了其他人的观点之后，他做出了一个重要决定：“我又恢复了每天抽两包烟的状态。从上一次开会之后，我再也不会努力戒烟了。”其他成员立刻对他的回归表示欢迎，并为他的决定鼓起掌来。

促进社会接受、逃避社会排斥的双重需求，有助于解释为什么邪教组织可以高效地招募和挽留成员。最初对潜在成员倾注感情，被称为“关爱轰炸”，这是邪教诱导的典型做法。“关爱轰炸”在一定程度上解释了这些团体在吸引新成员，尤其是那些感到孤独或被隔绝的人方面的成功。随后，邪教组织又威胁撤回这种关爱，这一做法解释了一些成员愿意留在团体中的原因：当成员们在邪教组织的鼓动下切断自己与外界

的联系后，他们除了该组织，就再也没有可以寻求社会接纳的地方了。⁽²⁶⁾

相似性：同侪说服

我们在观察与自己类似的人的行为时，社会认同原则能发挥出最大的影响力。这类人的行为让我们意识到自己该怎样做才正确。和“人多”一样，如果一种行为出自跟我们相似的其他人，我们会对它更有信心，认为自己照着它做也会有效、可行，并得到社会的接受。因此，我们更倾向于效仿与自己相似的人，这种现象叫“同侪说服”。

研究表明，担心自己的成绩或适应学校环境能力的学生，在得知有许多学生跟自己有着同样的担忧并最终得以克服后，他们的成绩会明显提高。如果消费者在购买太阳镜时听说其他许多消费者都买了某一品牌，那他们更有可能随大溜地购买该品牌的太阳镜。如果教室里的青少年好斗行为也会传染，但几乎仅限于同侪群体。例如，班里的男孩子的好斗性，对女孩子的好斗性几乎没有影响，反之亦然。如果员工看到其他同事（而不是经理）会分享工作信息，他们就更有参与进来。喜欢过度开具抗生素或抗精神病药物处方的医生，不太可能彻底改变自己的行为，除非他们听说自己的相关处方率比同侪高。经济学家罗伯特·H.弗兰克（Robert H. Frank）⁽²⁷⁾对环保行为改变做了大范围回顾后指出：“到目前为止，能够最有力地预测我们是否安装太阳能电池板、购买电动汽车、吃得更负责（如不吃野生动物）、支持气候友好政策的指标，是采取这些措施的同侪百分比。”⁽²⁸⁾

而且值得一提的是，我们往往认为，青少年思想叛逆且独立。可事实上，这种态度其实只是针对家长。在同龄人群体中，青少年同样要根据社会认同来判断自己怎么做才合适（见图4-4）。



图4-4 “独立思考的年轻人”

我认为，近来电视上普通人做的广告越来越多，原因就在这里。广告客户现在知道，要向普通观众（这些人构成了最大的潜在市场）推销一种产品，最好的办法就是表现其他“普通人”喜欢它、爱用它。所以，我们频繁看到电视上的老张、老李、老王在盛赞某个品牌的软饮料、止痛药或汽车。

有充足的证据表明，相似性在决定我们是否效仿他人行为中的重要意义。有人在校园里对筹款活动做了研究，提供了极为切题的例证。要是筹款人声称自己跟筹款对象一样，说“我也是这里的学生”，那么筹款数额就能翻不止一倍，因为这句话暗示对方也会支持同样的事业。这些结果揭示了社会认同原则的一个重要限定条件。我们会用他人的行为判断自己做出怎样的行为才恰当，尤其是在我们觉得别人跟自己一样的时候。

我曾在欧维达（Opower）公司担任了3年首席科学家，当时它尚属初创阶段。在工作中，我考虑到了同侪说服的因素。在我的建议下，欧维达与公用电力公司合作，向居民发送本户与邻居的能源消耗信息。这一信息有一项关键的特征，即比较并不是在任意两户邻居中进行，而是在房子相近、住房面积也差不多的邻居中进行的。换句话说，这是“与

你家情况类似的住房”。这一做法在节能方面带来了极佳的结果。如果用户的能耗大于同侪近邻，那他们就会努力减少自己家的能耗。到最后一次统计时，同侪说服手法减少了360多亿磅的二氧化碳排放量，电力消耗每小时减少了23万多亿瓦。而且，这一手法每年为电力用户节省了近7亿美元的账单。

同侪说服不光适用于成年人，也适用于孩子。例如，健康研究者发现，在学校开展反吸烟活动时，只有当同龄的孩子王以身作则，活动才能实现持久的效果。另一项研究发现，看过表现一个小孩主动去看牙医的影片之后，跟片中小孩同龄的儿童去看牙医时的焦虑感会大幅降低。几年前，我曾想方设法地减少我儿子克里斯的另一种焦虑情绪，要是那时候我就知道这一项研究结果就好了，虽说那时它还没发表。

影响力研究

我住在亚利桑那州，那里家家户户的后院都有泳池。遗憾的是，每年总会有几个孩子跌入无人看管的泳池溺水而亡。所以，我下定决心，要趁早教会克里斯游泳。克里斯倒是不怕水，他还很喜欢水；可问题在于，只要不带游泳圈，他就不肯进泳池。不管我怎么劝他、哄他，他就是不肯放弃游泳圈。我白忙活了两个月，什么进展也没有。于是我雇了我的一个研究生来帮忙，虽说他当过救生员和游泳教练，但也跟我一样失败了，他甚至没法劝克里斯脱下游泳圈划一下水。

就在这个时候，克里斯参加了一次户外野营，组织者提供了一系列的活动，包括到一个大型游泳池去游泳，克里斯因为很害怕，所以没敢去。有一天，就在我的研究生无功而返之后没多久，我去营地接克里斯，惊讶地看到他从跳板上纵身一跃，栽进了深深的游泳池里。我惊慌失措地快速脱了

鞋，想跳进去救他，却看见他顺顺当当地从水里露出脑袋，划到了泳池边上。

“克里斯，你会游泳了！”我冲过去兴奋地说，“你会游泳了！”

“是呀，”他若无其事地回答，“今天刚学会的。”

“太了不起啦！太了不起啦！”我兴高采烈，手舞足蹈，“但你怎么今天不再需要游泳圈了呢？”

“我都3岁了，汤米也3岁。既然汤米可以不带游泳圈游泳，我想我也可以！”

我真想狠狠踹自己一脚。毫无疑问，在寻求自己能做什么、应该做什么的相关信息时，克里斯会去参考小汤米，而不是一个身高一米八的研究生。要是我好好想想，早点让小汤米来给克里斯当榜样，就用不着这么徒劳无功地忙活好几个月了。我应该早点注意到野营里的小汤米会游泳，然后跟他的父母安排一下，找个周末让男孩们在我家的游泳池玩上一下午。我猜如果那样的话，在那天结束之后，克里斯应该就能甩掉游泳圈了。

READER'S REPORT | 读者报告

致西奥迪尼先生：

在我上大学时，每年放暑假，我都会到田纳西州、密西西比州、南卡罗来纳州和堪萨斯州挨家挨户地卖《圣经》。我想出了一个办法，在卖书时向女性潜在客户提及先前女客户的名字或女推荐人的名字，向男性潜在客户提及男客户的名字或男推荐人的名字，如果潜在客户是夫

妇，则提及先前买了书的夫妇的名字。有趣的事情发生了，我的销售业绩有了极大改善。

在最初推销《圣经》的**15**个星期里，我严格按照公司教我们的统一推销说辞进行销售，即强调我们图书的特点，那时每个星期的平均销售额是**550.80**美元。但是，一个新的销售经理教我们在陈述里提及从前客户的名字，比方说：“苏·约翰逊就买了一套，这样她就好给孩子们讲《圣经》故事了。”从第**16**个星期开始我采用这套做法，我发现，在第**16~19**个星期，每星期的销售额提高到了**893**美元，增加了**62.13%**！当然，不只有这些改变，还有些别的事儿。我清楚地记得，在第**19**个星期我醍醐灌顶般地意识到，尽管引述他人名字的做法提高了我的总销售额，但也令我损失了一些生意。有一天，我向一位家庭主妇做陈述，她似乎对书很感兴趣，但没法决定该不该马上下订单。这时，我提到她的一些已婚朋友已经买了书，结果她这样说：“玛丽和比尔买了？那我还是跟哈罗德谈谈再决定好了，我们一起拿主意会更好。”

接下来的好几天，我都在琢磨这件事，并有了些认识。我告诉一位家庭主妇另外一对夫妇已经买了书，这无意中给了她一个不马上买书的好理由，即她得先跟丈夫谈谈。但要是许多跟她一样的家庭主妇都买了，她买肯定就没问题了。从这一刻起，我下了决心，在向家庭主妇做陈述的时候，我只提及其他家庭主妇的名字。接下来的一个星期，我的销售额飙升到了**1 506.1**美元。我很快把这一战略扩展到丈夫和夫妻身上，向男性做陈述的时候只提男客户的名字，向夫妇做介绍的时候只提夫妇客户的名字。在接下来做销售的**20**个星期里，我每个星期的平均销售额达到了**1 209.15**美元。到了末期，我的销售额有所下降，是因

为我已经挣了太多钱了，我发现没办法再激励自己天天出门努力干活了。

毫无疑问，在整个销售过程中，我还学到了其他一些有助于提高销售额的知识。然而，在亲身体验了这些快速的变化之后，我确信，要使销售额提高**119.67%**，没有其他任何一个因素的作用能比得上来自相似者的社会认同这一条。

阿肯色州的这位大学教师：

作为我的朋友，你曾当面对我讲述这个神奇的故事。我觉得当时你肯定察觉到我对此有所怀疑，于是把4年暑假里每个月的销售记录拿给我看（你当时就记下来，而且保存了几十年），用来支持自己的说法。不愧是在大学里教统计学的啊！

自杀模仿

尽管我们已经看过了社会认同对人类决定的强大影响，可在我看来，最能说明这一影响的例子，来自一则看似荒诞的统计：每当自杀上了头条新闻，飞机，不管是私人小飞机、企业的喷气机，还是商业航班，就会以惊人的速度从空中掉下来。

例如，研究显示，每当自杀新闻炒得热火朝天时，死于商业客机坠毁事故的人数便会激增10倍！更惊人的是，不仅坠机而死的人会增多，发生致命车祸的人数也会急剧增加。这到底是怎么一回事呢？

有这样一种现成的解释马上冒了出来，即能让某些人自杀的社会条

件，同样也能让另一些人死于意外。比方说，面对紧张的社会压力（经济衰退、犯罪率上升、国际局势紧张等），某个人可能会以自杀来结束这一切。而面对同样的事情，其他人会做出不同的反应，他们可能会愤怒、急躁、紧张或分心。要是这样的人在驾驶、修理我们的汽车或飞机，那这些交通工具的安全性必然会有所降低，死于空难和车祸的人数就会猛增。

按照这种“社会条件论”的解释，使人自杀的社会条件同样也会导致事故发生，这也是为什么自杀新闻和致命事故之间会存在强烈的联系。可惜另一组有趣的统计数据表明，这种解释并不正确。致命事故激增的现象仅仅局限于自杀事件广为报道的地区，其他地方，就算存在类似的社会条件，可当地媒体如果没有曝出自杀的消息，类似的致命事故激增现象就不会出现。而且，在报道了自杀事件的地区，曝光力度越大，其后发生的致命事故增幅也就越大。因此，刺激自杀又引发致命事故的原因，并不是社会条件，而是对自杀事件的宣传，这导致了更多的车祸和坠机。

于是，又有人提出了所谓的“丧亲说”来解释自杀新闻报道和致命事故之间存在的强烈联系。它认为，由于能上报纸头条的自杀消息往往涉及广受尊敬的知名公众人物，也许他们自杀的消息让许多人感到既震惊又悲痛。如此一来，这些人因为惊讶和分神，在开车和开飞机的时候就会粗心大意。所以，每当有自杀新闻上了头条，交通工具发生致命事故的情况就会急剧增多。诚然，“丧亲说”能解释自杀事件曝光力度和致命事故之间的联系，知悉自杀消息的人越多，悲痛和粗心大意的人也就越多，可它无法解释另一个触目惊心的事实：要是新闻报道的是一个人自杀的消息，之后增加的也大多是一个人出事的事故；要是新闻报道的是一个人自杀并导致多人死亡的消息，之后增加的往往就是导致多人死亡的车祸或坠机事故。倘若人们单纯是因为丧亲、悲痛、出神而发生了事故，那就不可能有这样一种事实出现。

这样看来，自杀新闻对车祸和坠机事故的影响还具有令人难以置信的针对性。单纯的自杀，只有一个人死掉，会带来只有一个人死掉的事故；自杀的同时又导致多人死亡的新闻，则会带来多人致死事故。倘若“社会条件”和“丧亲”理论都无法解释这种诡异的联系，那么它到底是怎么回事呢？加利福尼亚大学圣地亚哥分校的一位叫大卫·菲利普斯（David Phillips）的社会学家认为自己找到了答案，他认为罪魁祸首是所谓的“维特效应”（Werther Effect）。

“维特效应”的故事既惊悚，又有趣。两个多世纪以前，德国大文豪歌德出版了一本小说，叫《少年维特的烦恼》。该小说轰动一时，但故事以书中的主人公维特自杀而告终。它让歌德名声大震，也在欧洲引发了一阵自杀浪潮。因为影响太过强烈，好几个国家都把这本书列为禁书。

菲利普斯对当代的“维特效应”做了跟踪研究。他的研究证明，只要报纸头版一登出自杀新闻，在新闻曝光率高的地区，自杀率就会激增。菲利普斯认为，这是因为一些内心饱受折磨的人读了别人自杀而死的报道后，就效仿了这种做法，了断了自己。这是社会认同原则的一个病态例证：这些人根据其他陷入困境的人的行动，决定了自己该怎么做。

菲利普斯考察了美国整整20年间的自杀统计数据，为当代“维特效应”找到了证据。他发现，每次自杀事件上了头版，其后两个月里自杀的平均人数就会比通常情况下的多58个。从某个角度上来说，每条自杀新闻会“杀掉”58个本来能够活下去的人。菲利普斯还发现，自杀诱发自杀的倾向主要集中在大肆报道第一起自杀事件的地方。他指出，头一起自杀事件的曝光率越高，其后自杀的人数也就会越多（见图4-5）。又有研究表明，这种模式不仅限于报纸报道。2017年3月31日，奈飞公司推出了系列剧网《13个原因》（*13 Reasons Why*）。在这部剧中，一名自杀的高中生留下了13盘录音带，详细讲述其自杀的原因。网剧播出后

的30天里，青少年自杀率上升了28.9%，这个数字比研究者分析的5年内的任何一个月都要高，并因此排除了是“社会条件”导致自杀率上升的可能性。

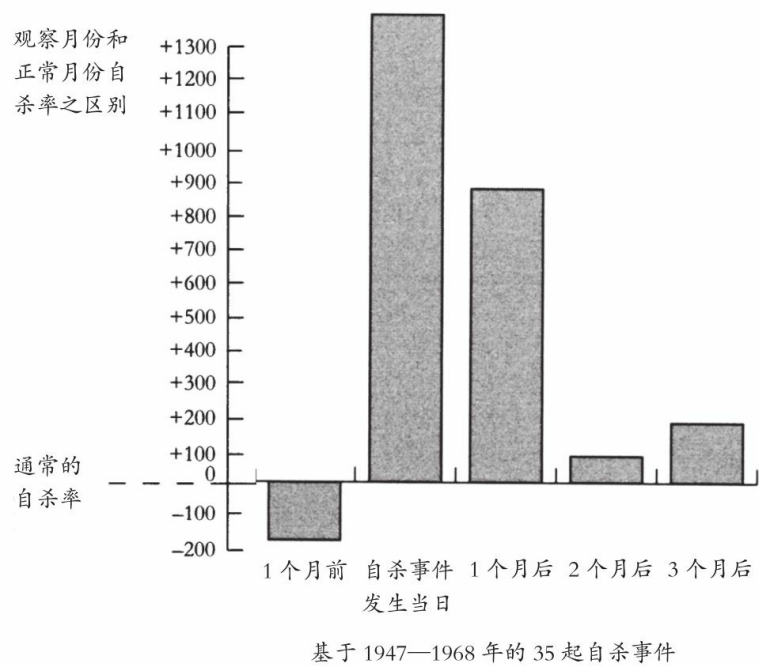


图4-5 自杀新闻出现之前、期间和之后数个月的自杀人数

上图提出了一个重要的伦理问题。仿效新闻自杀而死的人，本来是不该死的。自杀率在最初的猛增过后，虽然有所下降，但再也不会回到传统的水平。看到这样的统计，大肆炒作自杀消息的新闻编辑们应该停下来反思一下，因为相关报道有可能导致数十人死亡。最新的数据表明，除了报纸编辑，电视从业人员也该思考一下自杀新闻造成的负面效应。不管自杀消息是出现在新闻报道、信息特写还是虚构的电影里，它都能引发自杀狂潮，有模仿倾向的敏感青少年尤其容易成为受害者。

你是否觉得“维特效应”跟自杀新闻导致陆空交通事故激增的现象未免太过相似？菲利普斯也注意到了这种巧合。事实上，他声称，轰动的自杀新闻之所以会引发事故热潮，原因其实是同一个：模仿自杀。知道别人自杀之后，一大批人觉得，对自己来说，自杀也是一条不错的出路。于是，有些人采取了直接的行动，不加掩饰地自杀了，随之而来的便是自杀率猛增。

然而，其他一些人就不那么直白了。出于这样那样的原因，如保护自己的声誉、不让家人蒙羞受伤害或让家人享受到意外保险金，他们不希望自己显得像是自杀死的，而更想显得像是死于意外。所以，他们暗中故意让自己驾驶或乘坐的车辆、飞机出事。要实现这一目的，有太多方法可以选择：商业航班的飞行员可以在起飞的关键时刻倾斜机头，违背控制塔的指示，令人费解地把飞机降落在停着其他飞机的跑道上；开车的司机可以冷不防地撞到树上，驶入对面的车道；搭乘汽车或飞机的旅客可以弄晕驾驶员，造成致命冲撞；私人飞机的飞行员可以不顾所有通信设备里传来的警报，栽向另一架飞机。因此根据菲利普斯的说法，轰动的自杀新闻过后，致命交通事故的发生率陡增，很有可能是“维特效应”在偷偷搞鬼。

我认为菲利普斯的见解十分精辟。首先，它完美地解释了所有数据。倘若这些致命事故真的是隐性的模仿自杀，我们也就能够理解为什么自杀新闻见报之后事故率会升高了。我们同样也能理解，为什么自杀新闻闹得沸沸扬扬之后，事故率的增幅会最大。至于为什么只有在自杀新闻炒作得最厉害的地区，车祸、坠机次数才会出现猛增，单人自杀只会导致单人事故，多人死亡的自杀则会导致造成多人死亡的事故，这些谜题全都解释得通了。模仿是其中的关键。

其次，菲利普斯的认识还有另一点重要的地方。它不仅能帮我们解释已经存在的事实，还能帮我们预测从来没出现过的全新事实。举例来说，倘若轰动的自杀新闻之后频频出现的反常事故真的是模仿而非意外，那么它们必然会更为严重。也就是说，想要自杀的人很可能会把事故后果尽量安排得更为致命。开车时一脚踩在油门而不是刹车上，飞机下降时机头冲下而不是朝上，这样才能死个痛快。根据这一预测，菲利普斯核查了事故记录，发现了这样一种模式：较之发生在轰动性自杀新闻一个星期前的商业航班事故，发生在轰动性自杀新闻一个星期后的事故，其平均死亡人数要高出3倍之多。从交通统计数据中也可以看到类似的现象，轰动性自杀新闻过后的车祸更为致命，这一类车祸受害者的

死亡速度比一般情况下的要快4倍。

最后，菲利普斯的想法还引出了另一些有趣的预测。要是自杀新闻后增多的事故真的意味着有人在蓄意模仿，那么，这些跟风模仿者最可能效法的是跟自己类似的人。社会认同原则指出，我们会参考他人的行为方式，判断自己该怎么做。校园慈善捐款实验也表明，我们最容易受跟自己类似的人（同侪说服）的影响。

因此，菲利普斯推断，要是整个现象背后藏着社会认同原则，那么，轰动性自杀新闻的主角应该跟之后导致事故的人存在显而易见的相似之处。菲利普斯意识到，要验证这一可能性，最简单的方法是看只涉及单车单人的车祸记录，于是他比较了自杀新闻里自杀者的年龄，和新闻见报后死于单人车祸的司机的年龄。他的预测再一次准得出奇，每当报纸详细报道了一名年轻人的自杀事件，其后就会有年轻司机开着车撞到树上、栽进洞里、滚下河堤；而若新闻报道的是老年人自杀，那么死于这类车祸的就是年龄较长的司机。

最后一项统计把我彻底说服了。我完全相信了，同时也为它感到十足的震惊。很明显，社会认同原则这么普遍又有力，连是生是死这样的问题都要受它左右。菲利普斯的研究结果说明了一点令人痛心的倾向，报道自杀的消息，促使一部分跟自杀者类似的人走向了绝路，因为现在他们发现自杀的念头更加站得住脚了。更可怕的是，数据显示，在这个过程中，往往会有许多无辜者白白死掉（见图4-6）。

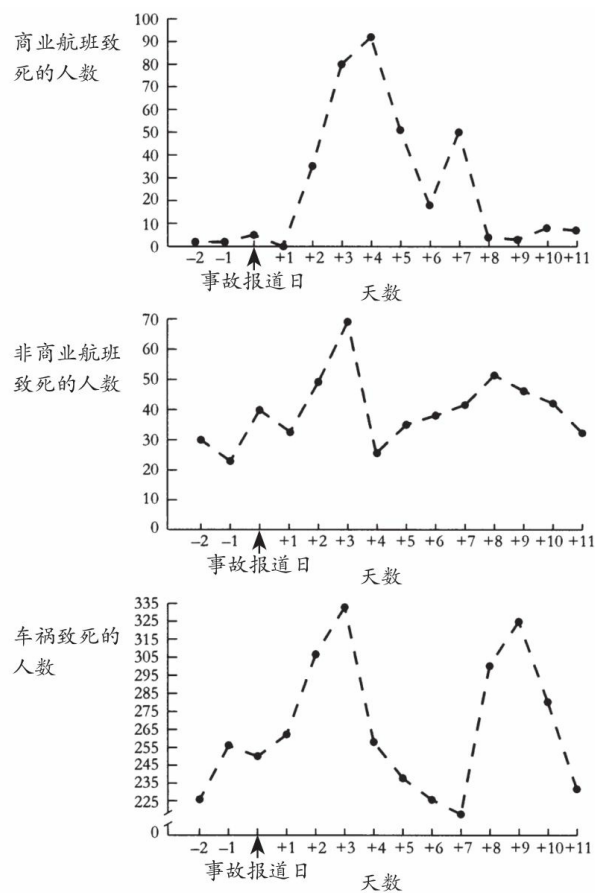


图4-6 自杀新闻上报之前、当天和之后的事故死亡人数波动情况

从上图可以清楚地看出，新闻报道刊出后的3~4天最为危险。短暂回落之后，过上大约一个星期又会出现一波高峰。到了第11天，“维特效应”就消失了。好几种类型的数据都表现出了同样的模式，暗示了隐性自杀有些值得注意的地方。想用意外来掩盖自己效仿自杀行为的人，要隔上几天才会采取行动，或许是为了积蓄勇气、策划事故或是安排后事。不管这一模式到底为什么呈现出这样的规律性，我们可以知道，在自杀并导致他人死亡的新闻报道刊出后的3~4天里，乘客的安全受到了最大的威胁。之后的几天也会很危险，但程度稍低。因此，我们建议，在这些时候外出旅行要特别当心。

让我们来看看地方媒体报道的一起自杀事件带来的致命后果吧。在这起事件中，一名青少年撞上了高速行驶的火车。接下来的6个月里，来自同一所高中的第二、第三、第四名学生相继追随他的脚步，以同样的方式死去。第五名学生的母亲注意到儿子不见了，并怀疑他有意自杀，这才阻止了另一场惨剧的发生。而她是怎么知道要去哪儿出手干

预，拦住孩子求死的行动呢？她直接去了他同学丧命的铁道路口。

说到社会认同原则的负面影响，最戏剧性的事件恐怕要数模仿犯罪了。20世纪70年代，我们注意到劫机事件就像病毒一样，能通过空气传染。20世纪80年代，在产品里下毒又成了热门，比如往泰诺胶囊里注入氰化物，往嘉宝儿童食品里掺玻璃碴等著名案件。据美国联邦调查局的法医专家所说，每发生一件全国曝光的大案，便平均会多冒出30件同类案件。自此以后，具有传染性的大规模枪杀事件又把我们给吓坏了。这类暴力事件先是发生在职场，之后又不可思议地蔓延到了学校里。

例如，在科罗拉多州里特尔顿的两名高中生制造了血腥的校园枪击案之后，警方紧接着又处理了数十宗类似的威胁、密谋和尝试犯罪事件，全都是饱受困扰的学生干的。有两名学生尝试成功了，就在里特尔顿大屠杀之后不到10天，艾伯塔省泰伯的一名14岁男孩和佐治亚州科尼尔斯的一名15岁男孩打死、打伤了共8名同学。美国弗吉尼亚理工大学可怕的杀人并自杀案件发生之后一星期内，全国各地的报纸纷纷报道了更多类似事件，光是在休斯敦就发生了3起。值得注意的是，弗吉尼亚理工大学大屠杀之后，接下来的同类事件就不光发生在高中了，北伊利诺伊州的一所大学也遭了殃。近些年来，大规模枪击事件甚至蔓延到了剧院和夜总会等娱乐场所。

这些事件数量如此之多，必须对其加以解释。为了搞清楚其中的奥妙，有必要理出一些共同的线索。就职场杀人案而言，观察家们注意到，杀戮现场大多是在美国邮政局的库房。所以，美国邮局“难以承受的紧张”环境成了千夫所指。而对于校园大屠杀，评论家总会谈到一個奇怪的共同点，所有受波及的学校都位于农村或郊区，城里那些总是吵吵闹闹没个消停的学校反倒平安无事。所以，媒体告诉我们，美国小城镇或郊区的青少年在成长期有着“难以承受的紧张压力”。根据这些说法，正是邮局工作和小镇生活的压力使得那里的人们做出了过激的爆炸式反应。这种解释很直白，类似的社会条件造就了类似的反应。

但此前，在理解致命事故的异常模式时，你我已经驳倒了“同类社会条件”这个说法。特定环境下相同的社会条件能解释自杀现象的猛增吗？菲利普斯想过这种可能性，但这个解释并不令人满意，我认为，用它来解释接二连三发生的枪击事件恐怕也不怎么样。我们还是再回到现实，看看能不能找到更好的答案吧。在邮局、郊区工作和生活，真的存在“无法承受的紧张压力”吗？跟在煤矿里工作相比，跟在黑帮横行、到处藏着危险的城里生活相比，这里的“压力”到底有多无法承受呢？好好想想吧，毫无疑问，发生大规模枪击事件的特定环境里的确有这样那样的紧张压力，但它们并不比没有发生此类事件的其他诸多环境更严重，而且往往也没有那么严重，“同类社会条件”论是站不住脚的。

那到底是怎么回事呢？我赞同社会认同原则，它认为，人们会效仿跟自己类似的其他人，尤其是在对自己不确信的时候（见图4-7）。

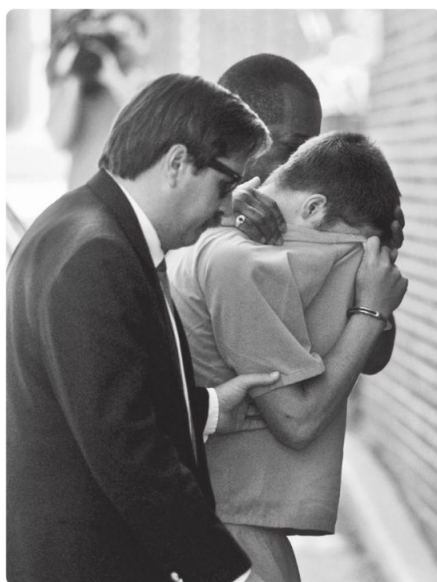


图4-7 机能失常的模仿者

影响力研究

1999年5月20日，学校开始上课前5分钟，15岁的托马斯·所罗门（Thomas Solomon）朝同学们开了枪。在击中了6名

同学之后，他才被一位勇敢的老师拦下。为了理解这其中的根本原因，我们必须意识到，在为期1年的时间里，一连串的同类事件曝光给托马斯造成的影响。先是阿肯色州的琼斯伯勒，接着是俄勒冈州的斯普林菲尔德，再接着是科罗拉多州的里特尔，最后是两天前艾伯塔州的泰伯。有人问托马斯的一个朋友，为什么好好的学生突然之间会变成学校里的杀人狂，此人回答说：“像托马斯这样的孩子，一直都听到、看到同样的事情。校园枪击像是成了他们的一条新出路。”

对一个心怀不满的邮局员工来说，有谁能比另一个心怀不满的邮局员工更像他自己呢？对一个饱受困扰的美国小镇青少年来说，有谁能比另一个饱受困扰的美国小镇青少年更像他自己呢？在现代生活中，总有很多人生活在心理痛苦当中。他们如何应对这些痛苦，取决于多方面的因素，其中之一是观察类似的其他人怎样应对。正如我们从菲利普斯的数据中看到的，大肆报道自杀新闻，会让其他类似的人竞相仿效，走上自杀之路。我相信，大肆报道群体枪击案也有同样的效果。和自杀案件中一样，媒体从业人员必须仔细想想该如何报道凶杀案件，该把它们摆在什么样的高度上进行报道。目前的报道方式太精彩、太具轰动性、太有新闻价值，也太过有害了——大量研究表明，它们具有传染性。

模仿之岛

诸如菲利普斯等人的工作可以帮助我们理解相似者行为的可怕影响力。一旦你意识到它的强大，也就有可能理解当代最惨绝人寰的顺从悲剧——圭亚那琼斯镇的集体自杀事件是怎么发生的了。让我们来回顾一下这起事件的几点重要特征。

影响力研究

“人民圣殿教”是以旧金山为基地的一个邪教组织，成员主要来自该市的穷人。1977年，该组织无可争议的政治、社会和精神领袖琼斯动员大部分信徒跟他一起搬到了南美圭亚那的一处丛林里。在那里，“人民圣殿教”的活动相对低调。直到1978年11月18日，去圭亚那调查该教派的加利福尼亚州众议院的里奥·瑞安（Leo Ryan）、调查小组随行的三名成员，还有一名邪教的叛逃者，在准备搭乘飞机离开琼斯镇的时候遭到杀害，“人民圣殿教”才成了大众目光的焦点。琼斯确信自己肯定会被逮捕，牵连到凶杀案里，到时候，“人民圣殿教”自然也难免解散。于是，他决心照自己的方式来控制圣殿教的结局。他把教派的所有信徒全召集到身边，要大家采取统一行动，集体自杀。

头一个响应者是个年轻的女性，她平静地走到带有草莓味的毒药桶跟前（现在这一点是尽人皆知了），先给孩子喂了一勺，自己也喝了。之后，她坐在地上，和孩子抽搐了4分钟，死了。其他人也一一效法。尽管有少数人逃跑了，据说还有几个人做了抵抗，可按幸存者的说法，910名死者里，绝大部分人是心甘情愿、秩序井然地赴死的。

这一事件震惊了全世界。广播媒体和报纸提供了全方位的报道、跟进和分析。好些天，人们的谈话都围着这个主题：“他们现在发现死了多少人了？”“有个逃掉的家伙说，他们像是受了催眠一样把毒药喝了。”“他们到底在南美洲干吗呀？”“真是太难以置信了，怎么会这样呢？”

是的，“怎么会这样”的确是最关键的问题。我们该如何说明如此惊人的顺从行为？人们给出了各种各样的解释，有的侧重于琼斯的个人魅力，说他有一种风度，能让信徒像救世主一样爱戴他、像父亲一般信任

他、像帝王那般对待他；另有解释指出，“人民圣殿教”吸引的教徒是有共同特点的，他们大多数是穷人，没受过什么教育，愿意放弃他们思考和行动的自由去一个安全的地方，让琼斯代他们做一切决定；还有的解释强调“人民圣殿教”的准宗教本质，这样的组织最推崇的就是对邪教领导者不容置疑的效忠。

毫无疑问，琼斯镇之所以发生惨剧，上述每一个特点都或多或少起到了推波助澜的作用，但我认为，光是它们还不够。毕竟，全世界到处都有领导者魅力惊人、追随者盲从盲信的邪教组织。更何况，这一类组织在历史上也从不少见。然而我们却几乎找不到证据，说明这类团体发生过哪怕是与琼斯镇事件类似的惨剧，肯定还有点别的什么关键因素。

有个特别能说明内情的问题给了我们一点线索：“要是这个团体依然留在旧金山，人们还会服从牧师琼斯下的自杀命令吗？”显然，这完全是个推测性的问题，但最熟悉“人民圣殿教”的专家却确凿地知道答案。时任加利福尼亚大学洛杉矶分校精神病学和生物行为学系主任兼神经精神病学研究所主任的路易斯·乔利恩·韦斯特博士（Louis Jolyon West），是研究邪教的权威人士。在琼斯镇惨剧发生之前，他曾观察“人民圣殿教”8年时间。事后不久我采访了他，他指出：“这件事不会发生在加利福尼亚。他们是住在一个怀有敌意的国家，他们所处的丛林与世隔绝。”

尽管惨剧发生后，各路评论铺天盖地，几乎叫人没了头绪，可在我看来，韦斯特的意见，再加上我们对社会认同原则的现有认识，似乎能为教众们顺从地走向自杀提供一个满意的解释。我以为，信徒们之所以能盲目顺从琼斯，最大的一点原因就在于“人民圣殿教”于一年之前搬到了一个人地两生的丛林国家。倘若我们相信琼斯真的具有邪恶的天分，那么，他必然完全理解这样的举动会给信徒们造成多大的心理冲击。突然之间，他们来到了一个一无所知的陌生地方。南美洲，尤其是圭亚那的热带雨林，跟他们在旧金山的人生经历完全不同。他们落入的环境，

包括自然和社会环境必然使人极具不确定感。

不确定感，这可是社会认同原则的左膀右臂！我们已经看到，人们在不确定的时候，会根据他人的行动来指导自己的行为。因此，在圭亚那的陌生环境里，教派成员会很乐意追随别人的领导。也正如我们所知，人们最容易效仿的是一种特别的人，即跟自己类似的人的行为。这才是琼斯搬迁战略里最可怕的地方。在圭亚那这样的国家，琼斯镇的居民们没有类似的人可以效仿，能效仿的只有琼斯镇的其他居民。

教派成员要怎样做才正确，这在很大程度上由受琼斯影响极深的其他教派成员所决定。从这个角度看，这些人井然有序、毫不惊慌、镇定自若地走向毒药桶、走向死亡，似乎更容易理解了一些。他们并不是被琼斯催眠了，只不过，琼斯和社会认同原则（这一点更为重要）说服了他们：自杀是正确的做法。最初听到死亡命令，他们肯定感到疑心重重，也必然会东张西望，想从他人身上知道自己怎样做才恰当。

值得特别注意的是，他们找到了两条令人印象深刻的社会证据，每一条都指着相同的方向。

第一，有些人立刻心甘情愿地喝下了毒药。在所有由强势领导人掌控的群体里，总会有少数狂热追随的人。在本例中，不管这些人是事先听从了专门的指示来充当榜样的，还是本来就最顺从琼斯的意愿（我们现在已经很难知道了），这已经无关紧要了，这些人的行为，给他人造成的心理影响不容小觑。既然新闻报道里同类人自杀的消息都能让陌生人自寻死路，可以想象，在琼斯镇那样的地方，眼见自己的邻居毫不犹豫地踏上了黄泉路，会给人带来怎样巨大的心理冲击。

第二，来自人群本身的反应。从当时的情况看，我怀疑那儿是出现了一场大规模的“多元无知”效应。琼斯镇的每个人都通过观察周围人的行动来评估当前局势，在他们眼里，其他所有人都挺平静，其实这些人也不过是在不动声色地暗暗观察、评估罢了。于是，他们“得知”，耐心地排队等死是正确的行为。显然，正是因为人们对社会证据做出了以上

可信却错误的解读，才导致了圭亚那丛林里的可怕一幕：所有人异常镇定地、有条不紊地走向死亡（见图4-8）。



图4-8 琼斯镇的成排尸体

琼斯镇躺着的成排尸体，向我们展示了顺从行为在当代酿造的一幕最大惨剧。

依我看，对琼斯镇惨剧的大多数分析，都太过强调琼斯的个人素质了。毫无疑问，他的确是个罕见的煽动家，可他运用的力量，与其说是来自他独特的个人风度，倒不如说是来自他对心理学基本原理的认识。身为领导者，琼斯真正的天赋体现在他意识到了个人领导的局限之处。没有哪个领导者能单枪匹马地说服群体里的所有成员，然而，一个强有力的领导者应该能说服群体里占相当大比例的一部分成员。“大量群体成员已经被说服了”，这一信息本身就足以令剩下的人信服。所以，影响力最强的领导者是那些知道怎样安排群体内部条件、让社会认同原则朝对自己有利的方向发挥作用的人。

琼斯似乎正是从这一点上汲取了灵感。他的神来之笔是决定把人民圣殿教总部从旧金山市搬到遥远的南美洲，在那个地方，不确定感和教派成员独一无二的相似性可以让社会认同原则发挥出最大的功效。光靠一个人的个性力量一般是难以掌控上千名居民的，可因为圭亚那的特殊

地理环境，上千名居民从追随的信徒变成了一群牲口。经营屠宰场的人早就知道，靠着从众心理，管理牲口很容易。只要你能让一部分牲口朝着预期的方向走，其他牲口就能机械地跟上去。与其说这些牲口是在跟随打头阵的，倒不如说它们是随波逐流，让大部队拖着走罢了。因此要理解琼斯的神奇力量，不仅要看到他独特的个人风度，还要看到他对同侪说服术炉火纯青的掌握与运用。

尽管没有那么令人痛心，但也有类似的证据显示，有同类居住的地方给人们带来的强大影响。一项在全美范围内针对影响品牌市场份额因素的分析显示，时间的流逝对品牌绩效的影响力小得出奇，3年的数值波动还不到5%。反过来说，地理因素却造成了巨大的差异。其中，对市场份额影响最大的是占80%的地理区域。人们对品牌的选择与其周围的人是一致的。不同地区的影响甚至大到让研究者怀疑起了“全国性品牌”这个概念及其相关性。“营销经理或许需要考虑在更大的程度上，针对不同地区采用去中心化策略，因为研究表明，同一地区的人们在态度、价值观和个性特征上是类似的，这可能是由于传染效应吧。”⁽²⁹⁾

负面事件效应

我居住的亚利桑那州自称“大峡谷之州”，得名于位于该州北部边缘著名又壮观的旅游景点。大峡谷看起来就像是一条向下倾覆的山脉。该州境内还有其他许多鬼斧神工的景观。其中之一是石化森林国家公园，这一地质奇观由2.25亿年前三叠纪晚期形成的数千个石化圆木、碎片和晶体构成。当时，被水卷携着倒下的树木和注入二氧化硅的火山沉积物结合到一起，最终石英和氧化铁取代了这些东西内部的有机物质，使之变成壮观、五彩斑斓的化石。

该公园的生态既强健又脆弱。它结构坚固重达数吨，却又容易受经常从地面上挪动、搬运和偷窃石化岩石碎片和晶体的游客的损毁（见图4-9）。虽说挪动和搬运化石这两种行为看似微不足道，但也颇让在公

园中研究树木远古运动模式的科学家感到头疼。当然，对公园构成持续性根本威胁的行为仍然是偷窃，这也是最让人担忧的问题。作为回应，公园管理人员在景区入口处立了一块巨大的禁止游客移动化石的告示牌。



图4-9 偷石头

游客到石化森林国家公园本是为了拍摄石化木化石的照片，但有些游客会把化石拿走。

不久前，我从前的一名学生带着自己的未婚妻到公园游览，他说，未婚妻是自己认识的最诚实的人，哪怕是借来的曲别针或橡皮筋都不会不还。然而，他们在公园的入口阅读“请勿偷窃”的大告示牌时，牌子上的一些措辞激起了这个未婚妻完全不符合其性格的反应，这让我的学生大感震惊。告示牌上诚恳地写道：

这里的遗迹每天都在遭到破坏。虽说一次往往只丢失一小块，但每年因盗窃损失的石化木材多达14吨。

于是，这叫那位本性诚实的未婚妻低声说道：“我们最好也拿上一块。”

告示牌上的措辞，怎么就把一位可敬的年轻女士变成了一个存心要偷窃国家宝藏的环境罪犯了呢？这显然是社会认同的力量，只可惜放错了地方。措辞里包含了一个弥天大错，而它又是公共服务沟通人员经常犯的误差。为动员公众避免一种不可取的行为，他们控诉这种行为发生得太过频繁。例如，在一份长期使用、名为“国民生产总值”的印刷广告里，美国林务局的吉祥物“森林猫头鹰”宣称：“今年，美国将生产比以往任何时候都多的垃圾和污染。”在我的家乡亚利桑那州，交通部把每个星期收集到的路面垃圾堆在高速公路旁的“垃圾之塔”里，好让所有人都看到。在一次为期6个星期的宣传“亚利桑那州的污染”的系列活动中，该州发行量最大的报纸向居民征集各地垃圾堆放最为严重地点的照片，用以刊发。

这种误差并不仅限于环保领域。各类宣传运动强调，滥用酒精和毒品的情况严重得让人无法忍受，青少年的自杀率令人震惊，行使投票权的公民太少……尽管这些说法可能是真实且出于善意的，但这些活动的发起者忽略了一件至关重要的事情：在“看看所有做这些不可取之事的人哪”的哀叹中，也暗含着削弱性的信息——“看哪，有这么多人都在做这件事”。公共服务沟通人员为了提醒公众注意一个广泛存在的问题，反倒因为社会认同原则而让问题更加恶化了。

影响力研究

为探索这种可能性，我和同事们在石化森林国家公园进行了一项实验，当时，那里平均每天有2.95%的游客会盗窃化石。我们在公园里盗窃行为频发的地方交替放置了两块告示牌，想通过这些告示牌来记录两种号召的效果，一种号召告诉游客有许多人都在公园里偷东西，另一种号召暗示只有少数人这么做。我们的第一块告示牌呼应了公园入口告示牌上的信息，它敦促游客不得盗走化石，同时又描绘了两三名

小偷行窃的场面。它几乎使盗窃行为翻了3倍，达到7.92%。另一块告示牌也号召游客不可盗走化石，但与前述适得其反的社会认同信息相反，它表现的是一名鬼鬼祟祟且孤零零的小偷，暗示只有极少人从公园偷走化石。这块告示牌将偷窃行为边缘化而非正常化了，从而将偷窃率降到了1.67%。

一些研究同样记录了试图让人们远离有害行为的宣传，反而带来了意想不到的负面后果。在一个教育项目中，几名年轻女性描述了自己饮食失调的情况，事后，项目中的其他被试也出现了失调症状。一个自杀预防项目告知被试，新泽西州的青少年自杀人数惊人，被试被发现事后有更大可能把自杀视为解决个人问题的潜在方法。初中生们参加了一个劝阻滥用酒精的项目，该项目要求被试扮演被同伴反复怂恿喝酒的角色，然而，被试却借此意识到，同龄人滥用酒精的情况比他们原以为的更普遍。简而言之，劝导性沟通不可在信息里把不良行为渲染得过分普遍，反而让它变得“正常”起来。

谴责不可取活动的严重性，这种倾向从另一重意义上来说同样是充满误导性的。通常，很多活动有可能根本就不普遍，只有在生动而激烈地描述此事发生时，它才会显得突出。以石化森林国家公园偷窃化石的行为为例。一般而言，从公园拿走化石的游客并不多，只占了不到3%。然而，由于公园每年接待的游客总数是60多万人次，所以盗窃行为的总数还是很高的。因此，公园入口的广告牌上说有大量化石被游客带走，并没有什么错。可即便如此，公园管理方让游客聚焦于盗窃行为经常发生这一事实，或许仍然犯下了两重错。他们不仅使社会认同的力量跟公园的目标对着干（即错误地暗示偷窃行为十分普遍），还错过了驾驭社会认同力量来支持公园目标的机会（即未能把遵纪守法的游客定义为大多数人）。真是个弥天大错。⁽³⁰⁾

遗憾的是，我们向公园管理方报告了研究结果，但他们却决定不对

告示牌上的相关内容做更改。因为他们做了一项问卷调查，询问几名游客的意见，游客表示，有关公园存在严重盗窃问题的信息，并不会增加自己盗窃化石的可能性，反而会降低它。我们对此感到很失望，但也并不意外，公园方面更看重游客对假设问题的主观反应，而不是我们基于实验证据所得的结果，这证明了一点：社会公众对什么构成了可信赖的研究结果，是缺乏认识的。

利用未来的社会认同

还有一种我自己也经常犯的社会认同失误。我在针对这一原则发表公众讲演时，一两位听众问了我一个重要的问题：“如果我没有切题的社会认同，该怎么做呢？”如果我开办了一家没什么名气的初创公司，或者我有一款新产品，在市场份额、销售数据或整体受欢迎程度方面都没有什么令人印象深刻的东西，那该怎么办呢？我总是回答：“好吧，如果你的确没有社会认同加持，当然不应该在这件事上撒谎。但你可以利用一些其他对你有利的原则，比如权威或喜好。稀缺性也是个好理由。”

然而，新近的研究表明，我关于如果尚未获得充分的社会认同证据那就回避它的建议是错的。与其仅仅依赖于现有的社会认同，不如通过未来社会认同的证据来达到同样目的。

研究者发现了一种人类在感知上的相应偏好。如果我们留心到一种变化，并发现它像是一种趋势，那我们就会期待该变化继续朝着同样的方向发展下去。这种简单的假设引导了历史上每一轮的金融投资牛市和房地产泡沫。观察到估值不断上涨的人预测，未来估值将进一步抬升。连续赢了几次的赌徒们会想象自己正处于好手气的连胜状态，下一场赌局还会再赢。像我这样的业余高尔夫玩家也可以证明同样的现象：看到自己前两轮的成绩提高了，我们便期望自己下一次的还会再提高，而完全罔顾概率问题和个人历史成绩。事实上，人们相信各种各样的行

为趋势都会顺着相同的轨迹继续下去，包括那些如今只有少数人在做的事，比如节约用水、吃素、无偿完成调查问卷等。

如前文所说，如果人们得知，只有少数人在做一件事，那他们也不愿去做。但要是他们得知，虽然现在做这件事的人还属少数，但会有越来越多的人参与其中，他们就会跳上潮流的大车去做这件事。以我最熟悉的一项研究为例。我们邀请了一些大学生作为被试，一部分人读到的信息暗示，同学中只有少数人在家里节约用水。而另一部分人读到的信息暗示，虽然节约用水的学生现在还是少数，但这么做的人的比例在过去两年里一直在提高。最后，最后一部分人（对照组）没有得到任何有关节约用水的信息。

此时，我们暗中准备好测试，观察这3种信息将对被试的用水量造成怎样的影响。我们要求所有人都参加一款新品牌牙膏的消费者偏好测试，在实验室水槽前刷牙后对牙膏进行评价。他们不知道，我们在水槽里安装了水表，用以记录他们测试新牙膏时用了多少水。

结果很清楚。与对照组的被试相比，知道只有少数同学努力节约用水的人，现在使用了更多的水。实际上，他们用的水是最多的。他们大概是暗自盘算了一番，认为既然只有少数人费心节约用水，那大多数人就不必费心了，所以他们听从了大多数人的引领。但另一些被试听说，虽然只有少数人节约用水，但其人数在不断增多，所以他们也加入了节约用水的行列。结果显示，该组被试刷牙时用水最少。

我们该怎样理解最后一项发现呢？这似乎跟我们前文所说的人们倾向于随大溜的研究结果背道而驰。这是否暗示，如果一种趋势明显可见，社会认同就不再是万能的？是，也不是。现在的社会认同或许不再占上风，但未来的社会认同或许会乘势而起。因为我们认为趋势将朝着同一个方向发展，它不光能告诉我们别人的行为过去怎么样、现在怎么样，还能告诉我们其他人的行为未来将变成什么样。因此，趋势让我们知道了另一种特殊而有力的社会认同——未来的社会认同。我们请研究

对象预测同事中有多少人会在未来6年内节约家庭用水，结果显示，只有那些了解节约用水趋势的人预测这一比例会提高。而且，这类研究对象中有不少人预测，节约用水在未来将成为主流行为。

基于以上结果，对那些不能在当下提供太多受欢迎的新东西的人，我的建议跟从前不一样了。我不再劝他们回避社会认同原则，转向别的原则，而是问他们：在一段合理的时间内，他们能否拿出诚实的证据，证明自己会越来越受欢迎。如果答案是肯定的，那我会建议他们把这一事实作为其所传达信息的核心点，因为他们的受众会假定，这些证据表明了他们所提供东西的真正的价值，也暗示了其将来的受欢迎程度。如果经过一段合理的时间之后，他们的答案是不，那我会请他们重新思考自己提供的东西，要么做大幅的调整，要么彻底放弃使用社会认同原则。[\(31\)](#)

如何防范

本章以餐厅对菜单进行小小调整的故事说起，接着介绍了成功推销《圣经》的手法，进而讲到了凶杀和自杀的例子，它们全都可以用社会认同原则来解释。如此之多的行为里都渗入了这种影响力武器，我们怎么才能抵挡它呢？更叫人头痛的是，大多数时候，我们并不愿防范社会认同提供的信息。它提供的行为方式信息，通常是正当且有价值的。靠着它，我们可以自信满满地在生活里穿梭自如，做出无数决定，而不必仔仔细细地考察每个决定的优劣利弊。从这个角度来看，社会认同原则为我们配备了一种奇妙的自动导航仪，就跟大多数飞机上装的那种差不多。

但自动导航偶尔也会出问题。倘若输入操控机制的飞行信息是错的，那这些问题就会冒出来。这时候，我们会偏离航线。误差的大小决定了结果的严重程度，但由于社会认同原则提供的自动导航仪更多的时候是朋友而非敌人，我们并不想彻底切断它。因此，我们面对的是一个

经典的问题，怎样使用一台既对我们有好处又对我们有坏处的设备呢？

好在这里有一条走出困境的道路。由于自动导航仪的弊端主要出在控制系统输入错误数据的时候，那么识别错误数据，就是我们对抗其弊端的最佳方式。要是我们能敏锐地察觉出社会认同自动导航仪采用失准信息时的状态，那就可以在必要时关闭它，自己接管控制权。

捣鬼的方式

不正确的数据会导致社会认同原则导航失准，这分为两类情况。一类情况出现在社会证据是蓄意伪造出来的时候。这种情况全是牟利者制造出来的，他们想给我们留下一种印象：很多人都在怎样怎样做（实际上可能根本不是这么一回事），而他们正好也希望我们那样做。情景喜剧节目里放的“罐头笑声”，就属于此类伪造的数据之一。同类的伪造数据还很多，大部分都假得一目了然。

由于自动导航仪可以随意切换和取消，我们可以按照社会认同原则设计的路线巡航游荡，但要是发现它使用了不正确的数据，我们完全可以接管控制权，对错误信息做出必要的修正，重启自动导航仪。由于伪造的社会证据大多一目了然，我们很容易知道该在什么时候执行这一简单的调整。面对明显是伪造的社会证据，我们只要多保持一点警惕性，就能很好地保护自己。回想一下在我们第1章提到的，虚假的在线产品评论有一些诸如缺乏细节、使用大量第一人称代词和动词多于名词等特征，把这些特征结合起来，我们就能发现它们是假货。

我们还可以利用另一些信息源来保护自己。例如，2019年，美国联邦贸易委员会成功指控某家化妆品公司自己发布客户对其产品的正面评价，而这些评价都是在公司领导的施压下由员工撰写的。此案被多家媒体广泛报道。因此，我们应该多注意这种有关虚假产品评论的新闻报道。

再举一个例子。前阵子，我注意到在街上拉普通人做广告的做法很是流行，也就是找许多普通人大谈一种产品的好处，不过这些普通人一般并不知道自己的话会被录下来。图4-10就是个幽默的例子。根据社会认同原则，我们可以预料到，这些来自“跟你我一样的普通人”的推荐能发挥相当有效的广告作用，但实际上它们对现实做了较为微妙的歪曲。我们只听到了那些喜欢该产品的人的意见，因此我们所得的印象也存在极度的偏差。

伙计们，别换台：接下来是来自火星的消费者

戴维·巴里 / 文

最近我看电视，总会插进一个广告，播音员用一种“波斯湾又发现大油田了”的兴奋声音说道：“现在消费者可以向安杰拉·兰斯伯里提出有关巴菲林镇痛药的问题啦！”

身为一个正常的人，听到这样的说法，自然的反应大概是：“啊？”也就是说：“安杰拉·兰斯伯里跟巴菲林有哪门子的关系？”但在这段广告里，几个一看就是从大街上随便拦下的消费者，居然都有关于巴菲林的问题要问兰斯伯里。基本上，他们问的问题是这样：“兰

斯伯里小姐，巴菲林这种产品，我买合适吗？”诸如此类。

这些消费者看起来很热切，就好像他们好几个月以来都坐立不安，死死绞着双手说：“我有一个关于巴菲林的问题！让我问问兰斯伯里！”

我们在这里看到的是一个日益严重的问题。长久以来，全美国都把这个问题用地毯遮起来，假装看不见——这就是火星消费者入侵的问题。火星看起来像是人类，但不像人类那么做事，而他们现在正接管了地球！

图4-10 大街上只是普通的火星人

更粗糙、更赤裸裸的伪造行为也出现了。广告制片人或许根本懒得费心去找真正的消费者做评价，他们会直接雇人扮演普通消费者，让他们在接受采访时假装成没彩排过的样子。索尼影业在一则广告中安排员工假扮影迷，称赞索尼拍摄的电影《爱国者》（*The Patriot*），广告随

后在电视网上播出，但被明眼人揭穿了真相。这些员工的老板开脱说，雇用演员或员工做假广告的欺骗性做法是“业界标配”，甚至并不仅限于索尼影业或者娱乐行业。此类造假还有一种形式是，雇用演员在电影院或商店外排队，以引起路人的兴趣。第一代苹果手机在波兰的发布活动，就是牟利者有时会用人手段捏造自己产品受欢迎程度的证明。负责代理苹果业务的广告公司承认伪造了有利于这款手机社会认同的证据。他们是怎么做的呢？一位发言人表示，在发布当天，“我们在全波兰20家商店前有偿找人员排长队，以求激发人们的兴趣”。

每当碰到这一类有人想愚弄我的情况，就会有一道清晰的指令叫我转入警惕状态：“当心！当心！遭遇假冒社会认同，暂时切断自动导航仪。”这是很容易的，我们只需要有意识地警惕造假的社会证据即可。平常的时候我们尽可以放松心情，等识别出奸商显而易见的造假，我们就可以还击了。

还击的时候要咄咄逼人。我说的可不光是无视对方的误导，尽管这种防御性策略很有必要，但我说的是主动反击。只要有可能，我们就应该狠狠戳一下那些伪造社会证据的人。我们不仅应该拒绝打虚假“即兴采访”广告的产品，还应向这些产品的制造商写信抗议，建议他们别再跟做这类广告的机构合作。

尽管我们并不总需要依靠他人的行动来指导自己怎么做，尤其是在事情很重要，必须亲自权衡优劣得失的时候，或者我们本身就是该领域专家的时候。但还有很多场合，我们的确需要把别人的行为当成有效的信息来源。要是我们发现自己处在这种证据经人蓄意篡改、信息不可靠的情况下时，我们便应做好还击的准备。我个人碰到这类事情，会觉得比单纯的上当受骗还讨厌。一想到有人利用我应对繁忙现代生活的决策捷径把我逼得无路可走，我就会怒发冲冠。只要有人胆敢尝试，我一定会厉声怒斥，并从中感到一种伸张正义的崇高感。如果你跟我一样，那你也应当这么做。

随大溜

除了人为伪造社会证据的情况，社会认同原则在另一个情况下也会导向失误。此时，一个纯属无心的失误会像雪球一样越滚越大，逼我们做出错误的决定。“多元无知”效应就是这种情况下的一个例子，明明出现了紧急事件，可人人都觉得没什么好惊慌的。比如下面这个例子。

影响力研究

几年前，新加坡一家银行的客户突然开始疯狂挤兑。这家银行本来声誉很好，出了这种事他们也觉得莫名其妙。直到后来，研究者采访挤兑者的时候才发现了出事的原因。那天，公交系统突然罢工，银行大门前的汽车站一反常态地站了一大群人。路人误以为这里聚集的人是为了从即将倒闭的银行里取钱，于是他们也慌乱地赶紧排队取起钱来，见此情形，越来越多的过路人产生了误解，纷纷加入了挤兑的行列。于是这家银行开门后不久，就被迫关了门，以免彻底倒闭。

这个例子有助于我们理解人们响应社会认同的方式方法。第一，我们似乎持有这样的假设：要是很多人在做相同的事情，他们必然知道一些我们不知道的事情。尤其是在我们并不确定的时候，我们很乐意对这种集体智慧投入极大的信任。第二，人群的反应很多时候都是错的，因为群体成员并不是根据优质信息才采取行动的，而只是基于社会认同原则在做反应。

这里有一点教训，人们绝对不应该完全信任类似社会认同这样的自动驾驶装置，哪怕没有坏分子故意往里面添加错误信息，有时候它自己也会发生故障。我们需要不定时地检查这台机器，用该环境下的其他证

据来源，如客观事实、先前的经验及我们自己的判断等与之进行对比，确保它没有出乱子。

幸运的是，以上预防措施并不需要花很多精力，也不需要太多时间，只要抽空迅速打量一下周围就行了。花点工夫采取这个小小的预防措施，是物有所值的，一根筋地依赖社会证据，有可能导致极为可怕的后果。比方说，许多坠机事故的发生，都是因为飞行员在天气条件不合适的时候做出了错误的着陆决定。航空安全研究者对此类飞机事故进行了深入的分析，并做出了解释。飞行员在决定着陆时，对物理证据考虑得并不充分，他们把重点放在了社会证据上，因为以前好多飞机全都这样安全着陆了。

显然，跟着其他飞机驾驶的飞行员最好还是偶尔看看仪表盘和窗外。同样的道理，在采纳群体证据时，我们也有必要周期性地四处看看。面对误导的社会认同，如果不使用这种简单的防护措施，我们很可能会跟这些不幸的飞行员和先前的那家新加坡银行落得一样的下场：坠毁。

值得一提的是，这件事发生在新加坡或许并非偶然，因为研究表明，东方社会的公民比来自西方文化背景的人更容易根据社会认同信息做出反应。其实，看重群体多于个体的文化都表现出了类似的敏感性。几年前，我和一些同事的研究显示了这类倾向在波兰是怎样发挥作用的。波兰这个国家的价值观正慢慢西方化，但民众的心态仍比普通美国人更依赖集体。我们问波兰和美国的大学生是否愿意参加一个市场调查。要预测美国学生的决定，主要是看他们以前是否经常答应同类市场调查的请求，这符合大多数美国人的个人主义偏好；可要预测波兰学生的决定，就主要得看他们的朋友过去是否经常答应同类市场调查的请求，这也符合该国更强调集体主义的价值观倾向。当然，本章的证据表明，社会认同在以个人主义为主的文化（如美国）中同样发挥着有力的影响。例如，表明社会认同对飞行员决策造成致命影响的数据，就来自

美国的航班。

READER'S REPORT | 读者报告

致西奥迪尼先生：

在一家跑马场工作的时候，我洞悉了一种伪造社会证据牟取私利的方法。为了降低风险，赚更多的钱，有些投注者能煽动公众把赌注押在劣马上。

跑马场的赔率是根据马身上下的赌注来确定的。一匹马身上押的钱越多，赔率就越低。好多赌马的人对赛马或下注策略的知识少得可怜。因此，尤其是当他们对参赛的马匹没什么了解的时候，他们就会把注下在最受欢迎的那匹马上。由于计分板每分钟都会更新赔率，公众随时都能判断出目前哪匹马最受欢迎。赌马老手改动赔率的手法其实非常简单，这家伙早就看准了哪匹马赢的机会大。接着，他挑选一匹赔率很大（比如**15:1**）、根本没机会赢的马，等下注的窗口一打开，这人就把**100**美元投在这匹劣马上，于是计分板上显示的赔率一下就降到了**2:1**，创造出“这匹马很受欢迎”的假象。

现在，社会认同原则开始发挥作用了。不确定把钱押到哪匹马上的人会观察计分板，根据先前赌客的投注判断哪匹马最受欢迎，然后自己再跟进。等其他人继续把钱押在这匹“最受欢迎”的马的身上时，滚雪球效应就出现了。此时，赌马老手可以回到投注窗口，在他真正看中的马身上下重注，现在，这匹马的赔率会比较高，因为“新的最受欢迎赛马”已经出现在计分板上了。要是这家伙赢了，先前的**100**美元投资就能赚回好多倍来。

我亲眼见识过这套把戏。记得有一回，一个人把**100**美元押在了赛前赔率是**10:1**的一匹马上，把它弄成了初期的大热门。赛场上流传起了谣言，说最早下注的家伙有内幕消息。接下来的事情你应该猜得到，人人（我也在内）都在这匹马上押钱。结果它跑了最后一名，还把腿给跑瘸了。很多人亏了大把的钱，可有人却赚了个盆满钵满。我们永远不会知道那人是谁，但他一个人把所有的钱都捞走了。这家伙把社会认同原则吃得很透。

这位前跑马场雇员：

我们再一次看到，对那些在特定环境下感到不熟悉、不肯定的人来说，社会认同最有说服力，因为这些人必须观察周围，寻找自己该怎么做的证据。本例中，我们可以看到牟利者是怎样利用趋势的。

本章小结

- 社会认同原则指出，人们用来判断自己在某个环境下该相信什么、该怎么做的一条重要途径，就是看其他人在这个环境下相信什么、怎么做。不管是成年人还是儿童，身上都存在强大的从众效应，而且这一效应适用于购物决定、慈善捐款、缓解恐惧等多种活动。社会认同原则可以用来使一个人顺从要求，只要告诉他其他很多人（人越多越好）都这么做了就行。因此，光是指出一样东西的受欢迎程度，就能提升它的受欢迎程度。
- 社会认同在三种情况下影响力最强。第一种是不确定感。

如果人们不确定，掌握的情况模糊，他们就更有可能关注他人行为，并认为这些行为是正确的。例如，较之一目了然的紧急情况，倘若形势暧昧不明，这时旁观者是否出手帮忙的决定会受到其他旁观者行为的极大影响。

- 第二种与“人多”有关：人们更倾向于依据群体人数的多少选择追随目标。如果人们看到很多人在做某事，就往往愿意追随，因为该行为显得更正确或有效、更具可行性、社会接受程度更高。
- 第三种是相似性：我们更倾向于顺从与自己相似者的信念和行为，这种现象叫“同侪说服”。从社会学家菲利普斯汇集的自杀统计数据中，我们可以清楚地看到同类人的行为对人们会产生多大的影响。这些统计数字显示：倘若其他陷入困境的人自杀的消息高度曝光，跟这个人相似的人也会决定走上绝路。有人对圭亚那琼斯镇的集体自杀事件做了分析，认为该组织的领导者琼斯利用了不确定感和相似性这两个因素来引诱当地的绝大部分民众顺从地自杀。
- 许多沟通人士在社会认同方面犯下的一个弥天大错是：声讨一种不可取行为（酒后驾驶、青少年自杀等）的发生频率，希望借此阻止它。然而，他们没有意识到，谴责声中（看看做这些事的这些人！）暗含着另一条削弱自己立场的信息：“看啊！居然有这么多人在做这件事。”因为社会认同原则的力量，情况可能会变得更加糟糕。
- 如果你由于观念、事业或产品尚未得到广泛支持而无法借助现有的社会认同，或许可以通过诚实地描述发展趋势（受众希望持续下去的趋势）的支持，借用将来社会认同的力量。
- 要想避免太过依赖不完善的社会认同，我的建议是：警惕

明显是伪造的社会证据，比如与自己类似的其他人在做什么，并意识到我们的决定不应该仅仅建立在“其他人怎么做”的基础之上。

Intelligence

第 5 章

权威

跟着权威走。

——维吉尔 (Virgil)

NEW AND EXPANDED

不久前，一位韩国记者问我：“为什么行为科学现在这么火？”我认为原因很多，但其中之一与全球政府、商业、法律、医疗、教育和非营利组织行为科学研究部门的运作有关。最新统计显示，在不到10年的时间里，大约有200个此类研究单位扎根下来，每个都致力于测试如何将行为科学原理用于解决各种现实世界的问题。它们中的第一家是英国内阁行为透视团队（Behavioural Insights Team），这个团队的研究颇有成效。

例如，为了研究怎样提高对有价值事业的捐赠额，尤其是针对那些经济条件宽裕、能够给出大手笔捐赠的群体，英国内阁行为透视团队的研究者比较了激发投资银行家将全天工资捐赠给慈善事业的各种手法的成功率。在一家大型国际银行的伦敦办事处，银行家们收到了一份捐款请求，请他们为两家主旨为“帮帮资本的孩子”和“英国脑膜炎研究”的慈善机构举办的筹资活动捐款。对照组里的银行家收到的是一封请求捐赠的标准信函，结果只有5%的人应允。针对第二组银行家，英国内阁行为透视团队则邀请了一位受人尊敬、支持该项目的名人去拜访，这种基于喜好的策略让应允率提高到了7%。第三组银行家面对的是基于互惠的请求。他们一走进大厦，一名志愿者就走上去先给了他们每人一包糖果，接着请他们参与该项目，此时应允率提高到了11%。第四组银行家收到了来自自己组织首席执行官所写的信件，信中强调这一项目对银行的重要性，以及该慈善机构对社会的价值。这一手法结合了权威原则，带来了12%的应允率。最后一组银行家得到的是基于互惠原则和权威原则的混合技术：志愿者送糖果，并送上来自首席执行官的私人信件，应允率此时飙升到了17%。

很明显，首席执行官的信，不管是单独使用，还是跟另一条影响力原则搭配使用，都对捐款者的决定产生了重大影响。这是因为，信的源头在收件人心目中同时具备了两种权威性。第一，首席执行官是掌权者，他可以影响收件人在组织内所获的成就，而且因为是针对收信人写

的私人信件，他也会知道这些收件人是否遵从了自己的要求。第二，他还是所议主题上的权威，他表明了自己懂得该慈善活动对银行具有什么样的价值，也懂得该慈善机构的内在意义。如果提出请求的人结合了这两种权威特征，我们便能预料到人们会表现出明显的顺从。实际上，行为科学史上最令人震惊的一种反应模式，就可以用这两种权威特征来加以解释。^[32]

假设你在翻看本地报纸的时候，留意到一则招募志愿者到附近大学的心理系参加“记忆研究”的广告。又假设你觉得这个实验会很有趣，于是你联系到了研究负责人斯利·米尔格拉姆教授（Stanley Milgram），安排好去参加一小时的实验。你刚到实验室，就碰到了两个男人。一个是负责实验的研究者，他身上穿着的灰色实验室大褂和带着的记录板都清楚地表明了这一点；另一个则是跟你一样普普通通的志愿者。

一番问候和寒暄后，研究者开始向你们解释要遵照的程序。他说，该实验的目的是研究惩罚对学习和记忆有什么样的影响。一位被试要完成的任务是学习一张超长清单上的成对单词，直到能把每对单词完美地记住，这个人是“学生”。另一名被试的任务则是检验学习者的记忆进度，每当后者犯错就加强电击力度，这个人是“老师”。

听到这个消息，你有点紧张。跟你的搭档抽签之后，你更恐慌了，因为你抽到了“学生”的角色。你从没想过参加这个实验还要挨电击，所以有那么一瞬间你想到了离开。但你又转念一想，真有必要的话，你随时都能走，再说一道电击能有多强呢？

等你学了一阵要记忆的单词组之后，研究者用皮带把你绑到了一张椅子上，让“老师”看着，将电极插入了你的胳膊。这下子你真的担心起电击的强度了，你提出了严厉的质询。研究者的态度更令你恐慌了，他说，电击确实有可能让你感到非常疼痛，但不会给你“造成永久性的组织损伤”。说完，他就和“老师”离开了，只剩你一个人在房间里。他们在隔壁，“老师”通过对讲装置向你提问，每当你回答错误，就电击你，

算是惩罚。

随着实验的进行，你很快认识到“老师”使用的模式。他提出问题，并等你回答，一旦你出了错，他就会宣布你要接受的电击伏特数，并拉下电闸惩罚你。最让人担忧的是，你每多犯一次错，电击强度就会增加15伏。

测试的第一部分进行得还算顺利。电击让你感到有点不适，但还可以容忍。可没过多久，你犯的错误越来越多，电击强度也随之攀升，此刻的惩罚强度扰乱了你的注意力，令你犯下更多错误，承受越来越痛苦的电击。等电击强度到了75伏、90伏和105伏时，你已经痛苦得呻吟起来。到了120伏，你冲着对讲系统痛苦地大叫，电击让你感到难以忍受。你又呻吟着接受了一轮惩罚，你拿定主意不再忍受下去了。等“老师”按下150伏的电压后，你向对讲机叫喊道：“够了！让我离开这里，请让我离开这里！让我出去！”

“老师”并没有如你预期的那样宽慰你，说他和研究者马上就来放你出去，相反，他只是提出了下一个问题，要你回答。你在惊讶和困惑中念叨出最先冲进脑子的答案。当然，它是错的。“老师”放出了165伏的电击。你尖叫着要他住手，放你出去。他却只用下一道试题来回应你，你在崩溃的边缘给出的答案当然是不正确的，于是他就又给了一轮电击。

你再也压抑不住恐慌了，此刻的电击已经强得让你全身扭曲并厉声尖叫了。你踢墙，你要人来放你，你求“老师”帮帮忙。但测试却照常进行下去，可怕的电击强度也在继续攀升，195伏，210伏，225伏，240伏，255伏，270伏，300伏。你意识到自己绝无可能正确地回答问题，于是你向“老师”大吼，你不再回答他的问题了，可一切还是照旧。“老师”把你的默不作声视为回答错误，又放出了一道电击。

酷刑就这样持续着，直到电击的强度令你几乎昏迷过去。你再也叫不出声了，也不再挣扎了，你只能感觉到每一轮可怕的电蛇撕咬。也

许，你会想，“老师”看到你都不动弹了，总该会停止了，没有理由再把实验进行下去了吧！可他却毫不留情地照章办事，提出问题，报出可怕的电击强度，并拉下电闸，这时已经在400伏以上了。你在混乱中想，这个人到底是什么样的魔鬼啊？他为什么不帮我？他为什么不停下来？

权威高压的力量

对于我们大多数人来说，以上的场景听起来十足是一场噩梦。可要说到它到底有多可怕，我们只要明白一点就行了：大体而言，这一切是千真万确的。心理学教授米尔格拉姆的确做过这么一个实验，确切地说，是整整一个系列的实验（见图5-1）。实验里扮演“老师”一角的被试真的会向尖叫、挣扎、哀求的“学生”施加连续不断的强烈电击，电击强度足以致人死亡。实验只有一个地方跟前面的描述不同，“学生”并不是真的挨了电击，那个痛苦呼喊、哀求怜悯的人不是真正的被试，而是假装遭到电击的演员。米尔格拉姆研究的真正目的，并不是惩罚对学习记忆的影响，相反，它和一个完全不同的问题有关：一个普普通通的人在履行职责的时候，会愿意向完全无辜的其他人施加多大的痛苦？



图5-1 米尔格拉姆实验

图为“学生”（“受害者”）被绑在一张椅子上，身着实验室大褂的研究者和真正的被试（他将担任“学生”的“老师”）正把电极安装到他身上。

答案极其令人不安。处在跟前面那场“噩梦”一模一样的情形下，大部分“老师”都会把折磨施加到可用的最大限度。

影响力研究

米尔格拉姆实验里2/3的被试都没有听从“学生”的请求，而是把面前整整30挡强度的电闸全部按了个遍，直到按下最后一挡450伏、研究者结束实验为止。更可怕的是，从“学生”最开始要求放了自己，到稍后的苦苦哀求，再到每一次电击都令“学生”“发出绝对的厉声惨叫”（这是米尔格拉姆的原话），该研究中的40名被试都几乎无动于衷，不曾主动放弃“老师”之职。

这样的结果把参与该项目的每个人都吓坏了，米尔格拉

姆本人也不例外。事实上，研究开始之前，他分别让同事、研究生和耶鲁大学（实验就是在这里进行的）主修心理学的学生看该实验的程序说明，并估算有多少被试会按下最后一挡电击强度为450伏的电闸。人们给出的答案基本上都在1%~2%。米尔格拉姆甚至找了39名精神科医生，根据他们独立做出的预测，1 000个人里也只会有一个人愿意把电闸按到底。也就是说，被试在实验中真正表现出来的行为模式是谁也没有预料到的。

我们该如何解释这种惊人的行为模式呢？有些人认为原因可能跟几种因素有关。所有被试都是男性，而众所周知，男性全都有侵略倾向；又或者，被试们没有意识到那么高的电压会带来什么样的潜在伤害；再不然就是，被试是一群变态的疯子，很享受折磨人的过程。但证据充分驳斥了这几种可能性。首先，后继实验表明，被试是否愿意一次次地电击“受害者”，跟性别无关，女性“老师”采取相同做法的可能性跟米尔格拉姆实验中的男性“老师”们一样高。

另一项实验针对“被试没有意识到电击对‘受害者’的危险性”这一看法做了研究。在本轮实验中，“受害者”会说自己有心脏病，他的心脏受不了这么强的电击：“够了！快让我出去。我告诉过你，我有心脏病。我的心脏开始难受了，我拒绝继续进行实验。快放我出去。”可结果还是一样，65%的被试忠实地履行了自己的职责，依次按到了电闸的最大挡。

最后，认为米尔格拉姆找来的被试全是心理扭曲的虐待狂，不能代表普通公民，这个借口也是不成立的。看了米尔格拉姆广告来报名参加“记忆”实验的人涵盖了各个年龄、职业和教育程度。而且，实验之后他们还做了一系列的人格测试，测试显示这些人心理十分正常，并不是一群疯子。事实上，他们就跟你我一样。套用米尔格拉姆的说法，他们

就是你和我。倘若他的看法没错，那么，我们每个人都有可能做出实验中那样可怕的事情来。这就引出了一个更加令人感到不舒服的问题：“什么能让我们做出这样的事情来呢？”

米尔格拉姆确信自己知道答案。他说，这必然跟我们对权威根深蒂固的责任心有关。据米尔格拉姆的观点，实验的真正罪魁祸首出在以下这一点上：被试没办法公然违抗自己的上级，也就是那些穿实验室大褂的研究者。研究者竭力吩咐被试履行职责，完全不管这么做是否会给他造成情绪和身体上的伤害。

米尔格拉姆的这套“服从权威”解释，得到了强有力的证据支持。首先，事情很清楚，若没有研究者的指示，被试们很快就会中止实验。他们痛恨自己的所作所为，并为“受害者”遭受的折磨感到痛苦。他们恳求研究者让自己停下来。研究者拒绝之后，他们继续照吩咐做，可在此期间，他们会颤抖着，汗流不止，并结结巴巴地提出抗议，还请求把“受害者”放掉。他们的指甲挖进了自己的肉里；他们的牙齿紧紧咬着嘴唇，都咬出血来了；他们用手托着脑袋；还有些人克制不住地紧张大笑起来。在米尔格拉姆最初进行的实验里，一名旁观者对有一名被试做了以下描述：

我看到一个成熟稳重的商人面带微笑，自信满满地走进了实验室。短短20分钟，他完全成了另一副模样。他颤抖不停，说话结结巴巴，似乎马上就要精神崩溃了。他一会儿一个劲儿地扯耳垂，一会儿握紧双手。有一回，他甚至用拳头砸自己的脑袋，并喃喃自语：“哦，天哪，让我们停下来吧！”可研究者所说的每一句话，他都照着做了，而且一路服从到底。

除了上述观察，米尔格拉姆还提供了更确凿的证据，证明被试的行为是顺从权威的结果。例如，在稍后的一轮实验中，他要研究者和“受害者”交换台词，也就是研究者告诉“老师”，停止电击“受害者”，而“受

受害者”却勇敢地让“老师”继续。结果再清楚不过了，面对另一名被试提出的要求，100%的人都拒绝再按下电闸。另一个版本的实验也出现了同样的结果，研究者和扮演“学生”的被试互换角色，研究者被绑在椅子上，“学生”让“老师”不听研究者的抗议，继续电击。这一回，仍然没有任何一个被试按下更高一挡的电闸。

还有一轮改编版的实验，记录下了米尔格拉姆研究中的被试顺从权威到了何等极端的程度。这一回，有两名研究者向“老师”下达互相矛盾的命令：一人要“老师”在“受害者”大叫“放我出去”的时候中止电击，另一人却认为实验应该继续进行。这些矛盾的指示成了整个研究项目唯一可笑的地方，被试悲喜交加，眼光困惑不解地从这个研究者身上转到那个研究者身上，不知道该服从谁才好。“等一下，等一下。这是怎么回事？一个说停，一个说继续。到底该怎么办？”趁着研究者争执不休的时候，被试拼命想要判断谁的级别更高。最后，由于没法找出到底该服从哪一个权威，所有的被试都听从了直觉，放弃了电击。和其他改编版实验一样，倘若被试本身就是虐待狂，或者具有神经质的攻击性，这样的结果是不可能出现的。

在米尔格拉姆看来，一个令人心寒的现象反复在他积累的数据中出现。“我们的研究发现，在权威的命令下，成年人几乎愿意干任何事情。”早有一些人担心另一种形式的权威，也就是政府，有能力从普通公民身上压榨出可怕的百依百顺，对他们而言，这样的结果显然具有更为严肃的引申意义。更何况，这个结果还告诉我们，权威的压力能够全然控制我们的行为。看到米尔格拉姆实验中的被试因为良知而汗如雨下、备受折磨，但还是不折不扣地执行任务时，还有谁会怀疑这一点呢？

要是还有人表示怀疑，布莱恩·威尔逊（Brian Willson）的故事大概会有些启发意义。

影响力研究

1987年9月1日，为抗议美国向尼加拉瓜运输军事装备，威尔逊先生和另外两名男子跑到加利福尼亚州康科德的海军兵站，躺到了铁轨上面。示威者满心以为这么做能让当天开出的列车取消，因为他们在三天前就向海军和铁路官员说明了自己的用意。可驾驶火车的非军方工作人员却接到命令，说不能停车。所以，哪怕他们在不到200米的地方就看到了抗议者，他们也并未减慢速度。另外两名抗议者及时爬出了铁轨，没有受伤，可威尔逊先生却慢了一步，两条腿从膝盖以下都被火车硬生生地压断了。由于现场的海军医护员拒绝治疗他，也不用军方的救护车把他送到医院去，围观者们，包括威尔逊先生的妻儿，只好自行给他止血，直到45分钟之后，私立医院的救护车才赶到。

出人意料的是，曾经当过4年兵的威尔逊先生却并不怪罪火车司乘人员，也没有责怪海军医护兵。相反，他谴责的是施加压力、让人盲目顺从的制度。“他们做的事情，跟我在战场上做的没什么不同。他们只是在服从一项疯狂政策下达的命令罢了。他们也是牺牲品。”尽管火车司乘班组也同意威尔逊先生的看法，觉得自己是牺牲品，却并不像他那么有雅量。事实上，整个事件里最令人瞠目结舌的地方也出在他们身上，司乘班组对威尔逊先生提出了控诉，要求他赔偿他们的损失，因为他们“不压断他的腿就没法执行命令”，因此承受了“羞辱、精神上的痛苦和肉体上的压力”。多亏了美国司法系统，这起诉讼很快遭到驳回。

米尔格拉姆最初进行这个实验的目的，是想弄明白为什么在纳粹统治期间，德国公民竟然会参与害死了上百万无辜者的集中营大屠杀。在

美国测试了实验程序之后，米尔格拉姆曾打算去德国进行实验，他认为，这个国家的民众肯定会表现出十足的顺从，好让他对整个概念进行完整的科学分析。然而，在纽黑文进行的第一轮实验便让米尔格拉姆大跌眼镜，他一下子明白：到德国去的经费完全可以省下，在家乡做实验就够了。“顺从的被试太多了，”他说，“完全没必要到德国去做实验。”可顺从权威也不是美国人的专利。后来，米尔格拉姆的基本实验程序曾在荷兰、德国、西班牙、意大利、澳大利亚和约旦等地反复进行，结果相差无几。而且，哪怕过去了几十年，米尔格拉姆所得的结论仍然成立。最近的一项效仿米尔格拉姆实验几点特征的研究，发现米尔格拉姆的被试和当今样本之间并无明显差异。

长达数十年的米尔格拉姆传奇，有个侦探故事般的结局。记者吉娜·佩里（Gina Perry）得以进入耶鲁大学档案馆，那里保存着米尔格拉姆论文。她在那里发现了米尔格拉姆此前从未发表过的一项实验的流程和结果。在此实验中，研究者要求每名“老师”自己视为朋友或邻居的学生施加电击。结果，人们对研究者命令的服从程度截然不同。在米尔格拉姆公开结果的实验中，65%的被试自始至终都服从研究者，而在这份未公开的研究中，只有15%的人始终服从。这一结果与我们将在第8章中看到的证据十分吻合，即较之陌生人或普通的熟人，人们更有可能站在自己有着联盟感的一方，如朋友、邻居或亲属。除了佩里的长篇记述，还有其他学者撰写过介绍米尔格拉姆的“秘密研究”的学术文章。⁽³³⁾

盲目服从的诱惑和危险

每当面对人类行为背后的一种强力推动因素时，我们便会很自然地想到，这种推动因素的存在是有充分理由的。就顺从权威一例而言，稍微思考一下人类社会的组织方式，我们就能找出许多说得过去的解释。被人类普遍接受的多层次权威体制能赋予社会巨大的优势，有了它，适于资源生产、贸易、国防、扩张和社会控制的成熟社会结构才得以建立

并巩固。倘若没有它，就会导致无政府状态。无政府状态对文化群体是没什么好处的，社会哲学家托马斯·霍布斯（Thomas Hobbes）就言之凿凿地说，它必然会让生活变得“孤独、贫乏、污秽、粗野和短暂”。因此，自出生之日起，社会就教导我们，顺从权威是正确的，违抗权威是不对的。父母的教诲、校舍里风传的小曲、故事和儿歌，甚至我们成年后碰到的法律、军事行动和政治制度，都无不充斥着这条信息。而所有这些“教化”，无不将服从和忠于正当规则摆到了极高的地位。

宗教教义也是一样。比如《圣经》开篇就提到，因为不服从至上的权威，亚当、夏娃还有整个人类失去了乐园。如果说这个隐喻太微妙了，那么让我们再来看看《旧约》。《旧约》用恭敬的叙述讲述了一个跟米尔格拉姆实验最为接近的故事，只因为上帝有了吩咐（哪怕没有半点儿解释），亚伯拉罕就愿意把利剑插入自己小儿子的肝脏。通过这个故事，我们知道判断一个行为的正确与否，跟它有没有意义、有没有危害、公不公平、符不符合通常的道德标准没有关系，只要它来自更高权威的命令，那就是对的。亚伯拉罕遭受的痛苦折磨，是考验他是否服从上帝，而他则像米尔格拉姆实验里的被试们一样通过了这场考验，说不定，被试们就是从这个《旧约》故事里学会服从权威的。

从亚伯拉罕和米尔格拉姆实验被试的故事里，我们可以看出顺从的能量和它在文化中的价值。然而，在另一种意义上，这样的故事又有可能造成误导。现实当中，我们其实很少对权威的要求痛苦地举棋不定。确切地说，我们往往没怎么思考，就下意识地顺从了，颇有“按一下就播放”的势头。来自公认权威的信息能为我们判断如何行动提供宝贵的捷径。

毕竟，正如米尔格拉姆所说，服从权威人士的命令，总是能给我们一些实际的好处。从小开始，这些人（家长、老师）就比我们懂得更多，我们发现，采纳他们的建议是有益的，部分是因为他们更有智慧，部分是因为他们手里攥着对我们奖惩的权力。成年之后的情况也是一

样，只不过此时的权威人士变成了老板、法官、政府领袖。他们因为所处地位更高的缘故，得以接触到更多的信息并掌握更多权力，因此按照正当权威的愿望去做是有道理的。正是因为它太有道理了，很多时候，哪怕权威人士说的完全没道理，我们也会照着去做。

当然，这种矛盾，正好也是影响力武器要借助的东西。一旦我们意识到服从权威在大多数情况下是有好处的，就会很容易不假思索地去服从。盲目服从这种机械做法，既有好的一面，也有糟的一面。盲目服从，我们就用不着思考了，省心又省力。可尽管它大部分时候都能让我们做出适当行为，例外的情况也有不少。

就从生活里找个权威压力强大又明显的方面吧：医疗领域。健康对我们极为重要，因此在医疗领域里掌握了丰富知识、具备强大影响力的医生，往往拥有着受人尊重的权威地位。此外，医疗机构本身也有着等级分明的权力和威望结构。各类医疗工作者很明白自己的工作处在这个结构里的什么位置，也很明白“医生”处在这个结构的最高层。没人能驳回医生对病例的判断，除非是另一个级别更高的医生。因此，医务工作人员素来有一套历史悠久的传统——自动服从医生的指示。

这么一来，就有可能出现这样的情况：医生犯下了明显的错误，层级较低的人却没想过要提出质疑。这是因为一旦正统的权威下达了命令，下属就用不着再思考，只管照着做就是了。这种“按一下就播放”的下意识反应跟医院这样复杂的环境结合到一起，出错简直不可避免。事实上，据为美国国会提供医疗政策咨询的“医学研究会”所说，就诊的病患每天至少会碰到一次用药失当。其他数据同样令人震惊：美国每年因医疗失误死亡的人数超过了所有事故的死亡人数，而全世界每年有40%的轻症患者和门诊患者因医疗失误而受到伤害。

患者拿到错误的药物，原因有很多。不过，人类学家、坦普尔大学药学教授迈克尔·科恩（Michael Cohen）和内尔·戴维斯（Neil Davis）却在《用药失误：成因和预防》中把大部分问题归结到患者对主治医师的

盲目服从。据科恩所说：“在一起又一起的案例中，（多是因为）患者、护士、药剂师和其他医护人员没有对方提出怀疑。”科恩和戴维斯报告了一起“肛门耳痛”的奇怪病例。一名患者右耳感染发炎，医生给他开了滴剂，让他点入右耳。但他在处方上并未把“right ear”（右耳）这个词写完整，而是写了个缩写“Rear”（“rear”在英文中有“屁股”的意思）。看到处方，值班护士立刻把规定的药水剂量点入了患者的肛门。

显然，耳朵痛却对肛门猛下药，实在不合情理，但不管是患者还是护士都没有丝毫的怀疑。这个故事给我们上的重要一课是：很多情况下，只要有正統的权威说了话，其他本来该考虑的事情就变得不相关了。这种时候，我们并不会从整体上来审视局面，而只会对其中的一个方面给予关注，做出反应。[\(34\)](#)

READER'S REPORT | 读者报告

致西奥迪尼先生：

我在宾夕法尼亚州沃伦的一个意大利人聚居区长大的。长大后，我偶尔会回家探访亲戚朋友。跟如今的大多数地方一样，小型的意大利商店消失了，取而代之的是较大的超级市场。有一回，我妈妈派我去超市买番茄罐头，我发现，几乎所有的“法玛诺”意大利切块番茄罐头都卖光了。我在货架上翻拣了一番，很快就在空货架的最下一层，找到了整整一排（还摆得满满当当的）“法玛”切块番茄罐头。我仔细看了看标签，意识到“法玛诺”和“法玛”是同一家公司。这家公司在经销部分产品的时候，会在自己的名字后面加个“诺”。我猜，这肯定是因为在销售意大利风格的食物时，名字里带个“诺”或者“奥”的厂商会让人觉得更权威。

得克萨斯州的这位大学教授：

感谢来信，我必须把你的这段评论也分享给读者：“从影响力触发因素的角度来看，新加的‘诺’完成了双份的任务。它不仅让厂商在‘意大利聚居区’显得更有权威，也让消费者感觉这家公司跟自己更类似，从而调动了喜好原则。”

每当我们的行为受这种不假思索的态度控制时，有一点保准没错：说服专业人士会跑来利用它。比如，在我们的文化中，广告商经常利用我们对医生的信任，找演员来假扮医生，宣传他们的产品。我最喜欢的例子是息可舒止咳糖浆的一则电视广告，该广告由曾在著名连续剧《综合医院》中扮演重要角色里克·韦伯医生的演员克里斯·罗宾逊出演。这段广告一开头，罗宾逊就说，“我不是个医生，我只是在电视里扮演了医生”（见图5-2），接着，他向一位年轻的母亲介绍息可舒止咳糖浆有些什么好处。这则广告非常成功，大幅提高了息可舒止咳糖浆的销量。



图5-2 我不是医生，我只是在电视里扮演了医生

治疗日常健康问题（如头痛、过敏和感冒）的药物广告中，经常使用类似上图这样用演员扮演医生的照片。广告无须明确声称该演员是医生，便可使用广告中的描述（包括展示医生用的许多配件，如白大褂、听诊器等）。

为什么这则广告的效果会这么好呢？人们到底为什么这么相信演员罗宾逊所说的话？因为聘请他的广告公司知道，他在观众的心目中是跟广受欢迎的电视剧里的重要角色联系在一起的。客观来看，我们明明知道这个人不过是个扮过医生的演员罢了，没道理受他的意见左右，可正是因为我们对权威产生的不假思索的反应，才让止咳糖浆在这个人的引导下卖得热火朝天。

作为对则段广告效力的旁证，1986年，罗宾逊因逃税入狱，息可舒止咳糖浆并没有停止播放它，而是找来另一位著名的电视演员彼得·伯格曼重新出演了这则广告，伯格曼曾在连续剧《我的孩子们》里扮演医生。除了更换了扮演医生的演员，新版广告的内容跟之前几乎毫无二致。值得注意的是，尽管遭到定罪，罗宾逊仍按监狱工作假释制度，获准继续在《综合医院》一剧中出演自己的角色。该怎么解释他得到的宽大处理呢？要知道，其他任何正在服刑的演员都不曾享受过这样的法网开恩。或许，是因为他在电视上扮演了一名医生吧。

内涵不是内容

第一次看到息可舒止咳糖浆的广告我就觉得很有趣，因为它利用了权威原则带来的影响力，却根本不曾拿出一个真正的权威，光是看起来像权威就足够了。这说明，我们对权威人士的下意识反应有一点很重要的特性：一旦处在“按一下就播放”的模式，只要拿出权威的象征就能将我们降服。

有几种象征权威的符号能十分有效地触发我们的顺从态度，因此，缺乏实质内容来证明自己的说法的说服专业人士对这几种符号爱不释手。比如，骗子大多喜欢给自己冠上各种头衔，做权威人士的打扮，带相关的身份标识；衣冠楚楚地从高档汽车里钻出来，自我介绍说是某某医生、法官或专员，这是他们最常用的把戏了。他们明白，一旦如此包装了自己，对方顺从的概率便会立刻大大增加。让我们分别来探讨一下

以上三种象征符号：头衔、衣着和身份标识，它们各自都有一大堆故事可讲。

头衔

头衔是最难却又是最容易得到的权威象征。正常来说，要得到真正的头衔必须付出多年的艰苦努力。然而，也有人毫不费力地就给自己贴上了个标签，轻松得到了他人的自动顺从。正如我们所见，电视广告里的演员和骗子们全是这么做的，而且做得都很成功。

最近，我跟一位朋友聊天时，从他那儿听到了一个很有说服力的故事，这个故事说明：头衔比当事人本人更能影响他人的行为。

影响力研究

我的这位朋友是一所著名大学的教授，他的出行行程总是很满，经常在酒吧、餐馆和机场跟陌生人闲聊。他说，从他这么多年的经验来看，最好别在谈话里透露“教授”的头衔。每当他一说自己是教授，人际交流的气氛马上就变了。前半个小时里风趣自然的话友就像是换了个人似时，他们对他毕恭毕敬，言听计从，乏味透顶。他的看法，先前可能会引出一场热烈的讨论，现在却只能带来一连串文绉绉的附和。这个现象让他感到有些懊恼和不解，因为他说：“我还是同一个人吗？我都跟他们聊了30分钟了，对吗？”总之，如今他再碰到类似的情况，他对自己的职业就守口如瓶了。

像他这种情况其实挺罕见的，更典型的模式是，某些说服专业人士谎称自己具有某某头衔，但事实上根本就没有。不管怎么说，上述两种

做法都表明，权威的象征照样能影响人的行为。

头衔除了能让陌生人表现得更恭顺，还能让有头衔的那个人在旁人眼里显得更高大。要是我的朋友知道这一点，不晓得他还会不会那么热衷于隐瞒自己的头衔，他个头可有点矮。有人研究过权威地位会怎样影响他人对当事者身材的判断。调查表明，头衔越是显赫，别人就越觉得有头衔的这个人高大威猛。以下是在澳大利亚一所大学进行的实验。

影响力研究

英国剑桥大学的一名访客先后来到5个不同的班级。在不同的班上，研究者对这个人的身份做了不同的介绍。在第一个班，这名访客是剑桥的学生；在第二个班，他是剑桥的助教；在第三个班，他是讲师；在第四个班，他是高级讲师；在第五个班，他是教授。当此人离开教室之后，研究者要班上的学生评估他的身高。结果发现，他的地位每往上提升一级，他在同学们眼中的身高也平均会提高1厘米，他是“教授”的时候足足比是“学生”的时候高出了6厘米。另一些研究发现，政治家选举获胜之后，在公众的眼中也会显得更高大；大学生被分配到扮演“经理”（与“员工”相对）这一地位更高的角色后，会认为自己个子也更高。

既然我们觉得身材体格与地位是挂钩的，那么肯定有人会靠偷天换日来获利。在一些动物群落，动物的地位来自它是否有能力控制其他动物。此时，动物的体格就成了决定它能否获得地位的重要因素了（见图5-3）。通常，跟对手打斗时，块头更大、力量更强的动物会赢。然而，为了避免这类肢体冲突给整个群体带来有害影响，许多物种都会采用一些打斗之外的办法来分出高下。两个对手相遇，必定会做出一些炫耀性的攻击姿态，这其中便包括一些增大块头的小伎俩。有些哺乳动物

会拱起背，把毛根根竖起来；鱼会张开鳍，用水把自己胀得鼓鼓的；鸟则会打开翅膀，并使劲拍动。很多时候，光是这么虚张声势一番，就能把对手吓跑，而人人都觊觎的地位，则落到了那个看起来个头更大、更强壮的战士手里。(35)



图5-3 高度期望

漫画家斯科特·亚当斯（Scott Adams）的描述并不见得有多夸张。研究表明，高个子男性比矮个男性的收入更高，也更有可能晋升到领导岗位。而且，虽说没有数据能直接证实，但我认为亚当斯对银色头发的看法也没错。

皮毛、鳍和羽翼，这些最脆弱的部位却能用来给人留下强壮、有分量的印象，难道不是挺有趣的吗？从这里面，我们可以吸取两点教训。

第一，体格和地位之间存在联系。因此，肯定会有人通过伪造前一种特征，营造具备后者的假象，从中渔利。这就是为什么哪怕骗子本来的身材在中等或中等偏上，也往往穿着增高鞋的原因。

第二点教训更具概括性：权力和权威的外部象征，说不定是靠假冒伪劣的材料编造出来的。让我们再举一个例子，回到“头衔”这个话题上。据我所知，从若干角度来看，这个例子都相当可怕。有一群由医生和护士组成的研究者，他们跟美国中西部三家医院有些联系。他们十分关注护士机械刻板地服从医生的现象。这些研究者发现，即便是高度训练有素的护士也不曾充分地应用自己的经验和技能，核对医生的判断。相反，每当医生给出指示，他们便一概照做。

正因为这样，才出现了前文谈到的往肛门里点耳用药水的案例。不

过，这群中西部的研究者把事情更深入地推进了几步。首先，他们想知道这种情况是孤立的个案，还是代表了一种普遍现象；其次，他们想看看在后果更为严重，比如说，对住院患者开出过量的未批准药物的时候，会不会出现这种问题；最后，他们还想试试，要是权威人士根本不在现场，而只是电话里一个自称是“医生”的陌生声音（这种时候，能表明对方是权威的证据是最为薄弱的），情况会怎样。

影响力研究

有一位研究者给外科、内科、儿科和精神科等22个不同的护士站打去电话，内容都一样。他说自己是医院的医生，并要接电话的护士向指定病房的某个患者用20毫克的药。面对这样的吩咐，护士有4个很好的理由保持谨慎：一是通过电话传达处方，直接违反了医院的规定；二是这种药本身是没通过批准的，不得大范围使用，而且也不在病房用药的清单上；三是处方的剂量明显超标，药品的容器上清楚地写着，“每日最大剂量”仅为10毫克，医生说的剂量却是它的一倍；四是护士从来没见过下命令的医生，甚至从未在电话里跟他交谈过。然而，在95%的例子中，护士都径直奔到药房，按吩咐拿出了相应剂量的药物，准备去病房里给药了。直到这个时候，在一旁不动声色的观察员才出手阻止了她，并说出实验的真正目的。

这样的结果真够吓人的。95%的正规护士毫不犹豫地服从了一道漏洞百出的指示，我们实在很有理由关注一下。毕竟，人人都可能住院看病啊！这项研究表明，用药失误并不仅限于往屁股里点耳用滴剂一类的琐碎事情，还可能造成更严重、更危险的后果。

在阐释上述令人不安的调查结果时，研究者得出了如下振聋发聩的

结论：

在实验对应的现实环境中，从理论上讲，应该是两种专业知识（医生的和护士的）结合到一起，确保采取的治疗措施对患者有好处，或至少不会有害处。可实验却清楚地表明，护士的专业知识实际上不起作用。

看起来，面对医生的指示，护士们放下了自己的“专业知识”，进入了“按一下就播放”的响应模式。工作的时候该做些什么，他们接受的医学训练和所学的知识完全没派上用场。由于在他们的工作环境中，顺从正统权威总是最有效率、最受青睐的做法，因此他们宁愿自动顺从，不惜犯错。更发人深省的是，他们朝这个方向走得太远了，连一个最容易假冒的头衔（根本不是什么真正的权威），也能诱使他们做出错误的响应。[\(36\)](#)

EBOX | 线上影响力

五年来，一支安全系统黑客团队对美国近1 000家地方银行和信用合作社的计算机网络发起了协同攻击，他们的成功率高得惊人。其中有963起案件，黑客得以入侵银行的安全系统，盗走诸如受保护的内部文件、贷款申请和客户数据库等资料。既然银行已经借助了先进的技术和软件，严密防范、侦测和防止数字入侵，为什么黑客还能在96%的案件中得手呢？答案在于黑客使用的方法。这些黑客并没有采用更为先进的数字技术来入侵银行先进的数字安全系统。事实上，他们根本就没有使用数字技术，而是利用了人们对权威的盲从心理。

事实上，这些黑客是受银行雇用而来的，我们也因此得以了解他们的手法为什么这么有效。他们给自己置办了消防检查员、政府安全检查员和灭虫人员诸如制服、徽章和标志这样的装备，没有预约就进

入银行机构，由银行员工陪同进入受限区域，之后员工还会离开、让黑客们自己开展工作。只不过，他们做的并不是银行人员期待中的工作。相反，黑客们从无人值守的计算机上下载了敏感的程序和数据，有时还会携带着硬盘、笔记本电脑甚至大型计算机服务器离开银行。黑客团队的老板吉姆·斯蒂克利（Jim Stickley）对此振聋发聩地指出：“人们总觉得，随着互联网的崛起，安保注意力和经费应该转移到计算机网络和来自黑客的威胁上。可‘此事’对此提出了一些不同意见。人们可能忘了一些基本的东西。”在顺从领域，服从权威就是最基本的原则。

作者点评：在获准进入银行机构的权威里，不光有本身被视为权威的角色，如消防检查员或政府安全监察员，也有那些只能称之为手握职权的人，如害虫控制专家。具有启发意义的是，这两种形式的权威都能发挥作用。

衣着

第二种可以触发我们机械顺从的权威象征是衣着。虽说相较于头衔，这种权威的外套更实在，可要伪造起来也很容易。警方的诈骗犯罪档案里，以换装为行骗手段的例子比比皆是。这些骗子像变色龙一样，一会儿穿上医院的白大褂，一会儿穿上神职人员的黑长袍，一会儿穿上军人的国防绿，一会儿又穿上警察的蓝制服。他们根据形势，穿上对自己最有利的服装。等受害人意识到衣着打扮只是像个权威但并没有实质意义时，往往为时已晚。

社会心理学家伦纳德·比克曼（Leonard Bickman）进行的一系列研究，表明了抵挡身着权威装束的人物提出的要求是有多困难。比克曼实验的基本流程是请街上的行人照着一些古怪的要求做，比如，拾起一个废弃的纸袋，或是站在公交站牌的背面。在半数的情况下，请求方是个

穿着普通的年轻男性；在另一半的情况下，他都穿着警卫制服。不管请求方提出的是哪一类请求，穿着警卫制服时，顺从他要求的人更多。倘若把提要求的人换成女性，同样是穿制服的效果会更好。有一个版本的实验最能说明问题。

影响力研究

请求方拦下路人，指着站在15米开外一处停车计费器前的男人。请求方不管是穿着便服还是警卫制服，向行人说的话都是一样的：“你看到站在计费器前的那个人了吗？他停车超时了，可没有零钱。给他一毛钱吧！”之后，请求方便转过街角走开了，这样等行人走到计费器那里的时候，他已经从行人的视线里消失了，不过，哪怕他的人早就没影了，制服蕴含的力量却延续了下来。在请求方穿警卫制服时，几乎所有的行人都照着他的请求做了，而当他穿便服时，照做的行人还不到一半。有趣的是，后来，比克曼要大学生评估实验中路人顺从的百分比。请求方穿便服时，大学生们的判断相当准确，他们猜测路人顺从的概率是50%，实际概率为42%；可大学生们却大大低估了请求方穿制服时的成功率，他们猜测的概率为63%，而实际概率为92%。

在我们的文化中，还有一种衣着打扮，尽管内涵不如制服那么一目了然，但照样也能暗示出权威的地位，那就是剪裁合体的西装，它也能唤起陌生人的顺从与尊重。例如，在得克萨斯州曾进行过一项研究，研究者安排一名31岁的男人在好几处地方违反交通法规，横穿马路闯红灯。有一半时间他身着笔挺的西装，打着领带；另一半时间他穿的是工装衬衫和长裤。研究者在远处观察，并记下有多少站在路边等候的人跟着这个人横穿了马路。结果他穿西服时跟着他横穿马路的人是他穿便装

时的3.5倍。

值得注意的是，诈骗犯在一种叫“银行核查员方案”的骗局中，将这两类经研究证明的具有强大影响力的权威装束天衣无缝地结合在了一起。诈骗的目标可以是任何人，但最受骗子偏爱的是独居老人。

骗局是这样开始的：一名穿着保守得体的三件套西服的男子出现在受害者的门口，骗子衣着的方方面面都表明他是个有地位、受敬重的人，白衬衫是上了浆的，皮鞋尖闪闪发亮。他的西装并不时髦，却是经典款。翻领不宽不窄正好；料子沉甸甸的有分量（哪怕当时是盛夏）；色调是蓝色、灰色或黑色等商务色系。

骗子向受害者（或许是他一两天前在银行盯上，之后尾随其回家的寡妇）解释说，他是银行的专业核查员，正在审计她银行的账目，并发现了一些异常。他认为自己已经找出了幕后黑手，有个银行工作人员经常篡改某些账户的交易报告。他说，寡妇的账户或许就是其中之一，但除非找到确凿的证据，否则他也拿不准。因此，他来征求她的合作，希望她能帮忙把账户里的存款都取出来。这样等嫌疑人经手交易时，核查员和银行的负责人员就能顺着线索追查了。

通常情况下，“银行核查员”的外貌和言谈都给受害者留下了深刻的印象。她根本没想过打一通电话问问真假，就立刻开车前往银行，把所有的钱都取了出来，然后回到家跟核查员一起，静待陷阱捕到猎物。最后，一名穿着制服的“银行警卫”在银行关门之后带着消息回来了，他说，账户没问题。显然，她的账户并没有被“坏人”篡改。“核查员”大大松了一口气，他亲切地道谢，并吩咐警卫把寡妇的钱送回金库去（因为现在银行已经关门了），免得人家第二天麻烦。警卫面带微笑，跟众人一一握手，离开了。“核查员”又道了几分钟的谢，之后也消失了。

自然，受害者最后终于发现，“警卫”并不是真的警卫，“核查员”也不是真的核查员，他们只是一对骗子，可他们意识到了精心穿着的制服所具有的魔力，能用“权威”的姿态轻松开启我们的顺从大门。

致西奥迪尼先生：

有了“医生”这个头衔，如果再穿上一件白大褂，就显得更加权威了。起初，我讨厌穿白大褂，但后来我在职业生涯中逐渐意识到了这件衣服蕴含的力量。有很多次，我到新的医院轮岗，总会穿上白大褂，交接班无一例外进行得顺顺当当。有趣的是，医生们也很明白这一点，甚至还创造出了一套等级排序。医学院的学生穿的白大褂最短，在实习期的住院医生的白大褂为中等长度，主任医师的白大褂最长。医院里的护士也察觉到了这种高低等级，很少质疑“长大褂”医生下的命令，而在跟“短大褂”医生交流互动时，医护人员会公开做出不同的医疗诊断，提出不同的治疗建议，有时甚至很粗暴。

佛罗里达州的这位医生：

你的来信提出了一个重要观点：在等级制组织中，享有权威地位的人会得到尊重对待，而没有权威地位的人则会得到无礼对待。我们在本文和接下来的内容中都会看到，人的衣着能发送信号，告知他人该用什么态度对待他们。

身份标识

衣着除了可以发挥制服的作用，还可以作为装饰，表现更广义上的权威。风格精致、昂贵的服装承载着地位和身份的光环，珠宝和汽车等类似的身份标识也一样。如果提出请求的人穿着带有昂贵设计师品牌

标志的衬衣或毛衣，商场购物者更乐于参与一项无报酬的调查，房主们会向门口的慈善募捐人捐更多钱，面试官会给求职者打更高的分、提供更高的起薪。而且，顺从意愿带来的差异大得惊人：答应参加调查的人多了79%，乐于慈善捐款的房主多了400%，面试官开给求职者的起薪高了10%。另一组研究指出了上述面试结果的原因：人们会认为穿着高品质衣服（哪怕仅仅是高质量T恤）的人，比穿着低劣质量衣服的人更有能力，而且，这种判断会在不到一秒的瞬间里自动做出。

另一些有关装饰的例子，比如高价珠宝和汽车，也能发挥类似作用。在美国，汽车作为地位象征尤其有趣，因为“美国人对汽车的爱恋”赋予了它非同寻常的意义。旧金山湾区的一项研究显示，名车车主能得到他人的特殊尊重。

影响力研究

研究者发现，当碰到一辆崭新的豪车堵在绿灯亮起的路口时，后面的驾驶员会等更长时间才按响喇叭；而如果是一辆旧款的经济型轿车，人们按喇叭可就按得此起彼伏了，驾驶员对经济车主没什么耐心。几乎所有人都按响了喇叭，而且大多数按了不止一次；有两次后面的车干脆顶上了它的后保险杠。而名车带来的光环就吓人了，50%的驾驶员会恭敬地等在它后面，等上足足15秒，直到它终于开动，也没人按过喇叭。

后来，研究者又询问大学生碰到这种情况会怎么做。相对于实验所得出的实际结果，学生们同样低估了自己碰到豪车按喇叭的等待时间。男学生们的估计尤其不准，他们觉得自己碰到豪车会比碰到经济车更快地按响喇叭，当然，研究结果恰恰相反。请注意，在其他许多关于权威压力的研究中，这种模式也同样存在。在米尔格拉姆的研究里、美国中

西部医院护士的研究里、警卫制服的实验里，人们都无法正确预测自己或他人面对权威的影响力会做出什么样的反应。每一次，人们都严重低估了权威的影响后果。权威地位的这种性质或许可以说明把它当成顺从策略为什么会如此成功。它不仅对我们很管用，而且还躲在我们的意识之外。[\(37\)](#)

READER'S REPORT | 读者报告

致西奥迪尼先生：

我干的这一行面临的一个大问题，那就是在客户的个人境遇或经济状况等条件变化时，让他们改变自己长期坚持的财务目标和战略，并让这些做法成为正确的事情。读完了你书中关于权威的章节后，我从以往只根据自己的观点向客户提供建议的基础上，新增了来自金融专家关于这一问题的观点。很多时候，这位专家是我们公司的首席经济学家，因为我的工作单位是一家大型经纪公司，在全国各地设有数百家办事处。但他们有时也可能是来自彭博社或美国消费者新闻与商业频道（**CNBC**）等金融频道的专家，或是发表过相关文章的作者。这一招很管用，我收获的认同比从前多出了**15%~20%**。不过老实说，按你那一章中的说法，我本以为结果会更好一些。我是做错了些什么吗？如果这种错误得到纠正，我能获得更好的结果吗？

密歇根的这位财务顾问：

这是一份不同寻常的读者报告。出于多种原因，我很少回应读者就个人事务寻求建议的请求。这些建议有可能是帮助一位大学生完成与影响力相关的家庭作业，也可能是求教

怎样“一劳永逸”地说服任性的配偶结束一段婚外情。但你的请求不一样，主要是因为你提到了两个跟其他读者普遍有关的问题。首先，如果你想要改变的人，对特定目标或方法有着长期承诺，那很难从他们身上取得任何进展。顺从性提高了15%~20%，我已经觉得很不错了。关于承诺与一致，我们在第7章会做更深入的探讨。其次，我可以推荐一种方法来增加专家建议的影响力：提供来自多名专家的建议。较之来自一位专家的建议，受众对来自一群专家的建议更信任、更乐于遵从。因此，如果沟通人士做好收集工作，借用多名专家的意见来支持自己的观点，会比只借助一位专家的意见获得更大的成功。

让权威更加可信

我们已经看到，在他人眼里手握职权或本身就是权威，会带来更多的顺从。但对于权威地位是因为手里掌握职权这种情况来说，以上方法还存在一些问题。一般而言，人们不习惯听吩咐去做事，这经常会带来抵触和怨恨情绪。出于这个原因，大多数商学院会告诉未来的管理者们要避免“命令与控制”的领导方式，而采用旨在促进自愿合作的方式。第二种权威，即在他人眼里见多识广非常有用。人们一般会很高兴，甚至会热切地渴望听从对自己正在从事的事情有更多知识的人提出的建议。

现代艺术专家米歇尔·斯特劳斯（Michel Strauss）讲过一个故事，很好地说明了人们追随专家的强烈倾向。

影响力研究

在著名表现主义画家埃贡·席勒（Egon Schiele）一幅画作的拍卖会上，斯特劳斯陷入了一场竞价之战。虽然这幅作

品最初的估价是20~25万美元，但为了跟另一位著名的席勒专家对着干，斯特劳斯的出价远高于此数，他认为，对方想要拍下这幅画，是因为知道一些自己不知道的事情。最终，竞价上升到62万美元时，斯特劳斯退出了。后来，他向自己的竞拍对手问起这幅画的情况，对方承认，自己之所以出价这么高，是觉得斯特劳斯可能知道一些自己不知道的事情。

那么，就让我们来看看，有什么方法能变成他人眼里的专家，这又会带来什么样的结果。

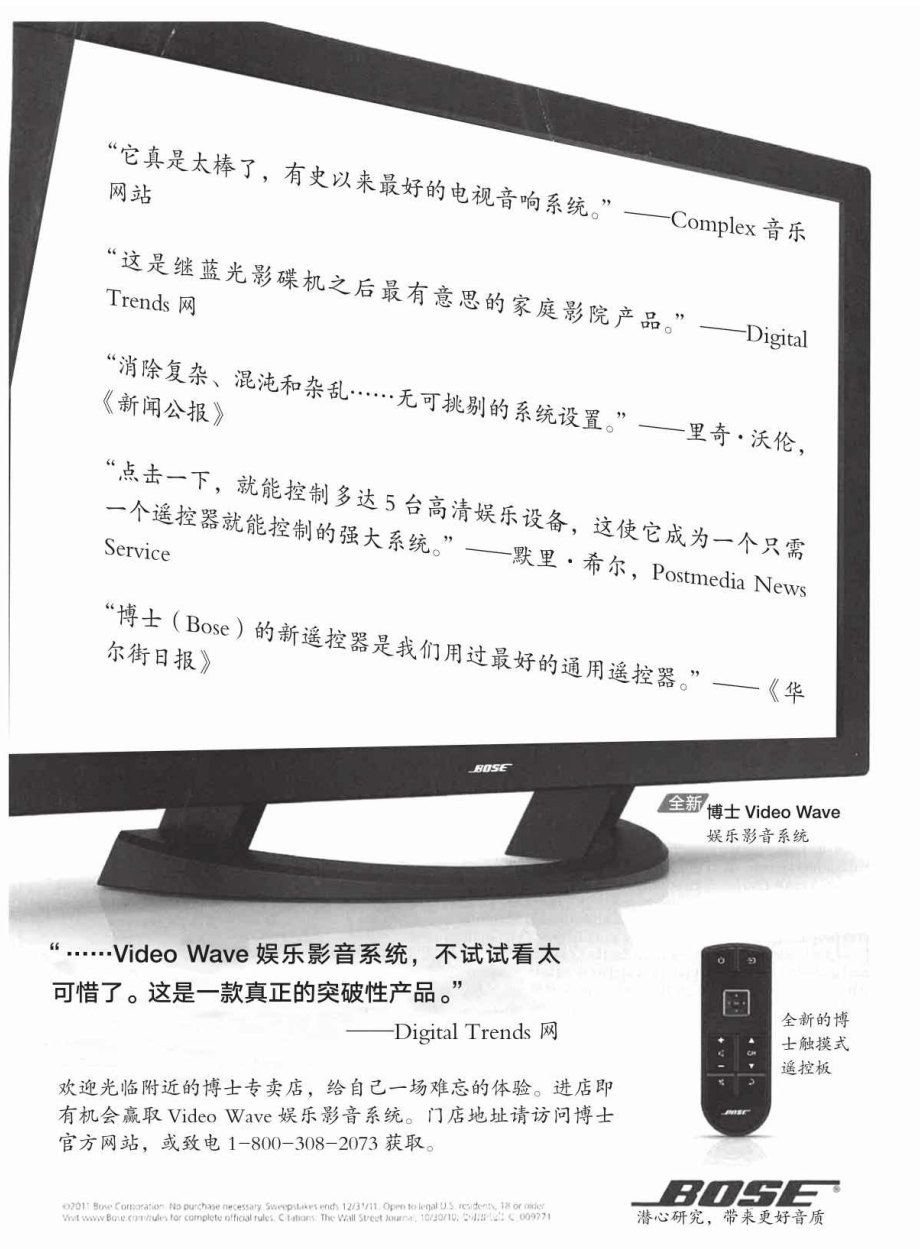
专业知识

研究区分了一种特别令人信服的权威，叫可信权威。在受众心目中，可信权威同时掌握了两种极具说服力的素质：专业知识和可信赖感。虽然我们已经讨论过专业知识发挥重大影响力的能力，但为确保可信赖感这第一根支柱得到应有的重视，我们可以额外补充一些有益的证据。例如，专业知识似乎能为拥有它的人造就光环效应：如果治疗师诊室的墙上挂着多种文凭和专业证书，那它不仅能为治疗师的能力带来更高的评价，也会为他的和善、友好、对客户的兴趣等带来更高的评价。此外，专家在报纸上发表的文章也能对读者的观点造成持久的巨大影响。在一项研究中，它将普通读者对相关观点的认同度提高了近20%，而且这种影响力无关读者的性别、年龄和政治倾向。

可信赖感

除了希望权威人物为我们提供专家信息之外，我们还希望他们是值得信赖的信息来源。我们想要相信，他们诚实而公正地在提供专家意见，即他们在竭力准确地描绘现实，而不是为个人利益服务（见图5-

4)。



“它真是太棒了，有史以来最好的电视音响系统。”——Complex 音乐网站

“这是继蓝光影碟机之后最有意思的家庭影院产品。”——Digital Trends 网

“消除复杂、混沌和杂乱……无可挑剔的系统设置。”——里奇·沃伦，《新闻公报》

“点击一下，就能控制多达 5 台高清娱乐设备，这使它成为一个只需一个遥控器就能控制的强大系统。”——默里·希尔，Postmedia News Service

“博士（Bose）的新遥控器是我们用过最好的通用遥控器。”——《华尔街日报》

全新 博士 Video Wave 娱乐影音系统

“……Video Wave 娱乐影音系统，不试试看太可惜了。这是一款真正的突破性产品。”
——Digital Trends 网

欢迎光临附近的博士专卖店，给自己一场难忘的体验。进店即有机会赢取 Video Wave 娱乐影音系统。门店地址请访问博士官方网站，或致电 1-800-308-2073 获取。

全新的博士触摸式遥控板

BOSE
潜心研究，带来更好音质

©2011 Bose Corporation. No purchase necessary. Sweepstakes ends 12/31/11. Open to legal U.S. residents, 18 or older. Visit www.Bose.com/rules for complete official rules. Citation: The Wall Street Journal, 10/20/10, 24/20/10; © 0909791

图5-4 外包的可信度

这则广告的说服力完全来自：该主题权威人士，他们申明了自己的知识；无须向公司尽忠的人，这确立了他们评论的可信赖感。

多年来，我参加过许多有关影响力技能的培训项目。几乎每一个项目都强调，让人感觉到值得信赖，是提升自身影响力的有效途径，而要

培养起这种感觉，得花一定的时间。尽管前一点尚有待证实，但越来越多的研究表明，后一点是存在明显特例的。只要采用一种聪明的策略，就有可能获得即时的可信赖感。如果沟通者在陈述一开始就早早提到自己的提议或观点的缺陷，而不是只描述优点、在陈述末尾才少许提到缺点甚至完全不提，受众会立刻认为他很诚实。采用这一介绍顺序的优点在于，听众产生信赖感之后，沟通者再提出自己观点的主要优势，前者会更容易相信。毕竟，这些信息是由一个值得信赖的人介绍的，这个人的诚实以先发影响力的形式体现在他不光愿意谈到积极面，也愿意谈到消极面。

这种方法的效力有案可查：在法庭上，抢在对手之前承认己方弱点的律师会显得更可信，打赢官司的次数更多；在竞选活动中，先赞赏对手积极方面的候选人，比如说“我相信我的对手拿出这份提案用心良苦……”更能赢得选民的信任感和投票意愿；在广告信息中，在强调产品优势之前先承认缺点的商家往往能带来大幅提升的销售量。2009年，达美乐在“新达美乐”的宣传活动中承认自己比萨质量不佳，结果反而使得比萨销量飙升，其股票也应声大涨。

在受众已经知道缺点存在的时候，这一手法尤其管用。在这种情况下，沟通者提到缺点所造成的额外损害很小，因为它并不是什么新消息，反倒能因此得到受众的信赖。在使用“然而”“但”“却”等转折词时，还能得到额外的收益：把受众的注意力从弱点引开，转到自身的强项上。求职者可以说：“我在这个领域的经验不够丰富，但我学得很快。”信息系统销售员可以说：“我们的架设成本并不是最低的，但我们的系统效率极高，你很快就能收回成本。”

沃伦·巴菲特被公认为我们时代最成功的金融投资家，他和合作伙伴查理·芒格带领伯克希尔-哈撒韦公司创造了惊人的价值。巴菲特不满足于自己在专业方面取得的成就，还不断提醒现在和潜在的股东，自己拥有可信度的另一个要素：值得信赖。他通常会在年报的前两页介绍自

已犯的某个错误，或是公司在过去一年里出现的问题，并探讨它可能对公司的未来造成什么样的影响。巴菲特从不像其他公司年报里经常做的那样，掩盖困难或者大事化小，他往往会表明：第一，他充分认识到了公司的问题；第二，他完全愿意揭露问题。随之而来的优势是，等他接下来介绍伯克希尔-哈撒韦公司的强大实力时，受众已被预先说服，对公司有更深刻的信任。毕竟，这些信息的源头有着显而易见的可信度。

巴菲特热衷于通过承认缺点来展示自己的坦荡透明，最清晰的例证或许来自他2016年所做的年报。这一年，他的公司股价比标准普尔500指数⁽³⁸⁾的涨幅要高一倍，因此没有必须公开报告的投资失误。那么，他做了些什么，来证明自己仍然拥有股东最关心坦率和诚实呢？在年报正文的第二页，他介绍了前一年的一项投资失误，他说：“尤其惊人的失误是1993年以4.34亿美元收购德克斯特鞋业（Dexter Shoe），该公司的价值很快跌到了零。”之后，他立刻详细解释了自己从这场惨败中学到的东西：他不仅错误判断了德克斯特的未来价值，还用伯克希尔-哈撒韦公司的股票支付了收购款项。他向股东们承诺，再也不会犯这个错误：“今天，我宁可准备做结肠镜检查，也不再发行伯克希尔-哈撒韦公司的股票了。”在我看来，巴菲特先生知道的远不止如何做一名超级成功的投资家，他还知道怎样与人交流，并在做一名超级成功的投资家方面给人留下深刻的印象。⁽³⁹⁾

更值得一提的是，“主动揭露弱点”手法的有效性，在法律诉讼中已经得到反复证实。同样，该手法对那些主动透露自己负面信息的企业也是有效的。研究表明，政客可以通过反对私利来提高自己的可信赖感和选票价值。在政治演讲中，还有一个相关效应，使用负面词汇来为信息建立框架（15%的人失业），跟使用正面词汇建立框架（85%的人获得了就业）相比，前者的说服力更强，因为它们显得更值得信任。恒美广告公司（DDB）率先通过承认弱点来制衡优点，制造出了大获成功的广告，比如为大众甲壳虫汽车设计的“丑，仅仅是表象”“它很丑，但它能带你去目的地”等广告，以及为安飞士（AVIS）租车公司设计的改变游

戏规则的“我们是第二，但我们更努力”广告。从此之后，加拿大的巴克利止咳糖浆等产品也设计了类似的促销用语（“真难喝，但真管用”），并取得了极佳效果。不过研究者证实，只有在负面情况发生之后，“承认负面情况”策略才有效。

EBOX | 线上影响力

网络评论的说服力也受到感知可信赖感的影响。西北大学斯皮格尔研究中心为营销传播的有效性提供了一些研究成果，它发表的一篇证据摘要，提到了网络评论在塑造顾客行为上的力量，其中有3项与感知信赖感直接挂钩：

- 五星好评显得太假。一种产品的星级性价越高，顾客购买它的可能性就越高，但这里的“高”仅限于一定程度之内。如果平均评价超过了4.2~4.7的最优分值区间，买家就会怀疑评分是假的，从而不愿意购买。
- 负面评价能建立可信度。研究中心认为，接近完美的评价会削弱可信赖感，负面评价的存在则会增加产品评估的可信度。事实上，如果网站收录一定的负面评论，购买转化率会跃升67%。
- 经过验证的买家是黄金评价员。也就是说，如果潜在顾客确认评价人是从前的买家而非店家收买的评价者，就会觉得这些评价更可信。因此，网站上存在这种用户，会大幅提高销量。

作者点评：除了斯皮格尔研究中心的结论，另一组研究者对网络评论“先承认错误”（也即将巴菲特的的手法搬到了网上）的做法同样表示支持。如果评论者承认自己从前购物时犯过错误，那么潜在顾客就

有更大概率去购买该评论者所推荐的产品。

值得注意的是，我并不推荐这样的做法：市场营销人员或销售人员一开始就说：“让我先把我、我的组织、我们的产品和服务的所有问题都给你讲一讲。”相反，我建议的是以下两点。第一，如果真有需要承认的缺点，那应该在信息较靠前的地方就表现出来，这样，它带来的可信度便能为信息的其余部分提供吸引力。第二，在说服交流中，有一个可以呈现最有力论点或特征的点。在这个地方，缺点会显得不值一提，或至少显得没那么重要。此时，在可信度的支持下，信息的接收方会最深刻地处理、最充分地接受对沟通方高度有利的要素。

如何防范

为免受权威地位的误导，防御策略之一是提前做好心理准备。因为我们一般都会低估权威及其象征对自己行为的影响，一旦它出现在要求顺从的场合，我们往往会来不及提防。因此，解决这一问题的基本方法，就是提高对权威力量的警惕性。有了这种警惕性，再认识到权威符号伪造起来很容易这一点，当碰到有人试图用权威的影响力左右我们的时候，我们就会更为谨慎。

听起来很简单，对吧？在某种程度上的确就这么简单。对权威影响力的运作有了更清晰的认识，理应能帮我们抵御它。可随之而来可能会有一种并发症，这是对付所有的影响力武器时都会碰到的，就是我们并不想彻底拒绝权威的影响，或者说，大多数时候都不想。一般来说，权威人士说的话都是很有道理的。医生、法官、企业主管、立法领袖等，这些人绝大部分都是通过丰富的知识和卓越的判断力获得当前地位的。因此，他们的意见确实具有极佳的指导作用。

这也就是说，权威大多是专家。事实上，“权威”一词，在字典里也

真的有“专家”的意思。在大多数场合，用我们自己欠缺知识和信息的判断来替代专家、权威的判断，未免太过自不量力。与此同时，我们又看到，不管是在街头巷尾还是在医院，随时都依赖权威的指点也很不明智。所以这里的关键就是要用不太费劲的方式判断什么时候该遵循权威指示，什么时候该保持独立的见解。不妨借助可信权威中专业知识和可信赖感这两个要素，向自己提出两个问题，帮助我们判断什么时候才应该听从权威的劝诫。

真正的权威

碰到貌似权威的人物试图发挥影响力的时候，我们要问的第一个问题是：“这个权威是真正的专家吗？”这个问题能让我们把焦点放在两点关键信息上：权威的资格，以及这些资格是否跟眼前的主题有关。通过这种简单的办法，着眼于权威地位的证据，我们就能避免自动顺从带来的大部分问题。

让我们从这个角度审视一下息可舒止咳糖浆的电视广告。倘若人们不是看到演员就想起“哦，这是电视里的那位医生”，而是好好看看他们的真实身份，我相信，这则广告不会播放这么长时间，效果又这么好。显然，扮演医生的演员并未接受过医生的培训，也不具备相关的知识。他拥有的只不过是“医生”的头衔，而且还是个空头衔，是通过他在电视中扮演的角色得来的。我们完全明白这一点，可一旦按下播放键，显而易见的事实就变得不重要了，非得特别注意才能意识到。这一点挺有趣的，不是吗？

“这个权威是真正的专家吗？”这一问题的宝贵之处便在这里。它让我们把注意力放在该放的地方上。它轻轻松松地让我们从或许毫无意义的权威符号上挪开视线，转到真正的权威资格上。而且，这个问题还逼得我们去搞清楚权威跟事情到底相不相关。在繁忙的现代生活里，碰到权威的压力，我们很容易忽视这一点。得克萨斯州的行人跟在衣冠楚楚

的绅士后面乱穿马路就是个典型的例子，事实上就算这个人像他的穿着打扮暗示的那样，的确是个生意场上的权威，他也无权横穿马路。在这件事上，他跟身后的行人没什么区别。

尽管如此，行人还是跟了上去，就好像绅士身上那套权威的标签抹去了相关和不相关的权威之间的区别似的。要是他们稍微费心问问自己，他是不是当时情形下真正的专家，他的行动是否暗示他掌握了丰富的知识，我想答案会完全不同。息可舒止咳糖浆广告里演员医生的例子也是一样。他并非不具备专业知识，他长期在一个艰辛的行业讨生活，取得了许多成就，可他掌握的是演员的技能和知识，而不是医生的。如果观看那段著名的广告时，我们把焦点放在演员们真正的能力上，就会很快意识到，他对息可舒止咳糖浆的赞许是靠不住的，他不过是一个有名的演员罢了。

我和同事们在一个研究项目中证明，通过训练，让被试专注于广告代言人的真实资质，能让前者变得比过去更擅长对广告进行评估。如果广告里的代言人没有相关的资质，比如，让演员阿诺德·施瓦辛格推广一种互联网技术，让游戏节目主持人亚历克斯·崔贝克宣传牛奶有益健康的特点，被试就不再那么容易顺从。与此同时，如果代言人具备相关资历，比如让疼痛研究所的主任推荐止痛药，让首席执行官介绍自己的公司对一款商业保险多年来积累的经验，也会让被试更容易顺从。

这里给我们的启发是什么呢？为了保护自己不受虚假权威的误导，我们应该经常问：“这个权威是真正的专家吗？”不要自信地认为自己不会被权威的象征所蒙蔽，这些象征会自动在我们身上发挥作用。在我们的研究中，只有当被试认识到自己是多么容易陷入这一自动过程，他们才能够通过质疑沟通者的相关专业知识来中止这一过程。而且，只有这样的被试，才不曾遭受愚弄。

油滑的真诚

接下来，假设我们碰到了——一个权威，而且他的确是相关的专家。在屈从他权威的影响力之前，我们应该问出第二个简单的问题：“这个专家说的是真话吗？”哪怕是知识最丰富的权威，也不见得会诚实地把信息告知我们。因此，我们必须考虑一下他们在当前情形下的真实可信度。大多数时候我们都会这么做。我们一般更愿意听从那些看似公正的专家，而对那些能通过说服我们得利的专家保持戒心。研究表明，全世界的人都是这样，哪怕小学二年级的孩子也是这么做的。多想一想专家会不会因为我们的顺从而得到好处，我们就为自己又设立了一道安全网，防御权威不必要的影 响。即便是某个领域的资深权威也无法说服我们，除非我们确信他们阐述的信息如实反映了真相。

在评估权威可信度的时候，我们应当牢牢记住，说服专业人士经常使用一种小策略，他们会偶尔说些有违自己利益的话来让我们相信他们真诚可靠。如果运用得当的话，这种方法可以微妙而有效地“证明”当事人的诚实。他们或许会提到其立场或产品存在某个小小的不足，但这一缺点必定是次要的，其他更突出的优点轻松地就能克服它（见图5-5）。看看这些广告词吧。汽车租赁公司安飞士：我们是第二，但我们更努力；欧莱雅：我们价格高，但你值得拥有。说服专业人士们靠着暴露自己的一些小缺点来建立基本的诚实感，这样等之后强调更重要方面的时候，他们就显得更可信了。



图5-5 药苦糖更甜

在很多场合，缺点都能变成优点。例如，各科成绩全是A，这很容易让人觉得分数掺了水。可是若有几个“C”，反倒把其他的“A”映衬得更地道了。有研究发现，送到大企业人事处的推荐信里要是除了夸奖，还能有一两句稍带负面的评价，那便能给求职者带来最有利的结果。

这里有一点很重要：要区分这种对错误的披露是否诚实。沟通者在传递信息的早期就揭露自己的缺点或先前的错误，从而收获信赖感带来的回报，这本身并没有什么错，但绝不应该为此捏造信息。回想一下，谨慎正直的巴菲特在年报一开始的时候是怎么做的。定期让读者提前了解自己的实际情况，在我看来这并不是一种伎俩。相反，我认为这说明了值得信任的沟通者也可以有足够的社交智慧，通过及时、真实的披露来暗示可信的信任。

我们要警惕的是欺骗性地使用这一做法。我曾在一个餐厅里，见到有人借助一套迂回手法取得了辉煌战果。餐厅的服务员因为底薪太低，要靠客人给的小费补足收入，这早就不是什么秘密了。除了提供良好的服务之外，大部分成功的服务员都有一些多收小费的诀窍。他们知道，顾客的账单金额越高，自己得到的小费就可能越多。于是为了提高顾客的总消费额，并提高他们给小费的比例，服务员也会定期使用说服专业人士的手段。

因为想搞明白他们是怎么操作的，几年前，我曾到几家相当高档的餐厅应征当服务员。可惜我没有经验，只能当杂役。结果这份工作反而给了我一个很有利的位置观察和分析服务员的做法。没过多久，我就了解了其他员工都知道的一件事，这里最厉害的服务员是文森特，经他手点的菜单总是价目较高，他收到的小费也总是很丰厚，其他服务员的周薪都比不过他。

所以我开始在干活的时候溜到文森特负责的桌子边观察他采用的手法。我很快发现，他的手法的确不拘一格。

文森特有一大堆方法，他会根据不同的情况有针对性地使用。如果顾客来的是一家人，他会表现得分外活泼，甚至稍带滑稽。他招呼大人，也不忘孩子。如果来的是一对在约会的年轻男女，他会变得很讲礼仪，甚至有点专横地要挟男方点大餐，多给小费。这种时候，他只跟男方说话。如果来的是一对年纪较长的已婚夫妇，他仍然是彬彬有礼，但姿态会放得较低，并尊重地对待夫妇双方。要是顾客是单独一人来就餐，文森特则会选用一种友好的态度，诚恳、健谈、热情。

面对8~12个人的大聚餐时，文森特的保留曲目就来了，他会说些看似有违自己利益的话。具体手法精彩至极。在第一个，通常是女性点餐的时候，他就开始行动了。不管她选什么，文森特都会做出同样的反应：眉头紧锁，手在点餐单上打转，之后飞快地扭过头去瞅一眼经理在哪儿。这一番表演之后，他会稍稍朝餐桌倾过身子，用不高但整桌人都能听见的声音说：“今晚这道菜恐怕不怎么好。我可以向您推荐这个或者那个吗（此时，文森特推荐了菜单上两道比顾客最初点的那个稍微便宜些的菜品）？它们今晚都不错。”

这套把戏调用了几条重要的影响力原则。第一，就算是没采纳他建议的人也会觉得文森特帮了自己的忙，提供了有助于点餐的宝贵信息。人人都很感激，因此，等到了顾客决定给多少小费的时候，互惠原则就会有利于文森特。第二，除了能提高小费的百分比，他的态度还可能增加这桌人点菜的总金额。他把自己打扮得像是这家店里的权威人士，清楚今晚哪样菜好，哪样菜不好。而且，他看似违背自己利益的做法在这里发挥了作用，它向顾客证明文森特是个值得信赖的内线，因为他推荐的菜比原本点的菜要稍微便宜一些。看起来，文森特并非只顾着往自己兜里揣钱，而是把顾客的最大利益放在了心上。

综合以上几方面，文森特立刻显得见多识广又诚实起来，这就给他带来了很高的可信度。文森特很快就利用起这个新形象，等聚餐的众人点完菜，他会说：“很好，你们愿意让我帮你们选些红酒来搭配吗？”这

一幕差不多每天都要来上一回，顾客们的反应也几乎一模一样，他们微笑着点头，基本上都表示赞同。

就算站在我的位置，我也可以从顾客们脸上读出他们的想法。“当然，”他们好像是在说，“你知道这里有什么好东西，显然你是站在我们这一边的。那就告诉我们该点些什么吧！”文森特看起来很高兴，他的确知道哪些东西好，便推荐了几种出色的红酒（当然也很贵）。到了饭后甜点的阶段，他也同样很有说服力。有些顾客本来并不想点甜点，或是打算跟朋友共吃一份的，此刻也被文森特打动了，他对火焰冰激凌和巧克力慕斯的描述真是叫人垂涎欲滴。毕竟，有谁能比一位确凿诚实的权威更可信呢？

文森特通过优雅的态度把互惠原则和可信权威原则结合到了一起，既提高了顾客的消费总额，又提高了自己的小费在其中所占的比例。他靠这一手赚到的钱相当可观。

READER'S REPORT | 读者报告

致西奥迪尼先生：

我在一个商学院为雄心勃勃的**CEO**们上课，我教他们如何通过承认失败，来促进自己的职业生涯。一名学生把这个教训牢牢记在了心底，他把自己从前在一家网络公司失败的经历详尽地写进了自己的简历，并说明自己从这次经历中学到了什么。在此之前，他总想着把失败藏到没人知晓的地方，但这并没有给他带来真正的事业成功。可这么做了以后，好几个名声很大的岗位都选中了他。

某《财富》**500**强公司的前首席执行官：

主动承担失败的责任，这种策略不仅适用于组织中的人，也适用于组织本身。研究表明，在年报中自愿为业绩不佳负责的公司，来年的股价会比不这么做的公司更高。

本章小结

- 社会中存在着顺从权威要求的强大压力。很多正常、心理健康的人宁肯违背自己的意志，也要向他人施以危险又残忍的痛苦折磨，只因为旁边有权威人士要他们这么做。这种顺从正统权威的强烈倾向，来自一种系统化的社会实践，它的目的是向社会成员灌输“这么做是正确的”印象。一般来说，按照真正权威的指示去做，大多数时候都是合适的，因为这些人往往拥有更多的知识、智慧和权力。基于这些原因，人们在做决定的时候往往会盲目地选用这一思考捷径。
- 自动响应权威的时候，人们容易对权威的表面象征（而非实质）做出反应。研究发现，头衔、衣着和汽车这三种象征能有效地起作用。研究者针对这些象征符号的影响力做了独立研究，证明拥有其中之一（无须正规的资质）的人，会得到他人更多的顺从和尊重。此外，在所有案例中，服从的人都低估了权威压力对其行为的影响。
- 权威的影响力来自被认为手握职权或本身是权威。第一类的权威存在问题，因为吩咐别人去做事往往会引起抵触和怨恨情绪。第二类权威则避免了这一问题，因为人们认为这类权威掌握了充分的信息，而他们又往往乐于听从对自己正在从事的事情有更多知识的人提出的建议。
- 如果在他人眼里是权威的人，不光是专家（在相关主题上

知识渊博），还显得值得信赖（诚实地介绍自己的知识），那他就会有最大的说服效力。为确立可信度，沟通者或许会承认自己在某方面存在不足（多为极小的缺点），但这一不足，很快就为随后所介绍的优点所弥补。

- 要保护自己免受权威影响力的负面影响，不妨问自己两个问题：这个权威是真正的专家吗？这个专家说的是真话吗？第一个问题把我们的注意力从权威的象征转到证明他资历的真正证据上；第二个问题建议我们不要光着眼于权威的知识，还要从当时的环境出发，考虑他的可信度。想过后一点的话，我们应该会对以下用于提高他们自身可信度的手段提高警觉：有些别有所图的人会先透露一点有关自己的负面消息，这样一来，他们就能给人留下一种诚实的印象，好让随后的信息显得更可信。

Intelligence

第 6 章

稀缺

不管是什么东西，只要你晓得会失去它，自然也就会上它了。

——G. K. 切斯特顿 (G. K. Chesterton)

NEW AND EXPANDED

我有个朋友叫桑迪，是一位非常成功的离婚律师。通常，她会充当离婚双方的调解人，因为双方都希望在不耗费大量时间、招惹无数麻烦、付出高昂法庭审判费用的情况下，就离婚条款达成一致。在桑迪着手进行调解前，她会把伴侣双方及各自的律师带到不同的房间，以避免双方因置身同一空间而出现面红耳赤、大吵大闹的局面。双方事先都已向桑迪提交了书面提议，她负责在两个房间中穿梭调停、寻求妥协并最终达成协议。她说，这个过程更多的是靠她对人类心理而非离婚法律的认识。所以她希望我能试着帮忙预防谈判接近尾声时常常出现的致命僵局。僵局一旦出现，就很难化解，有时甚至会破坏整个调解过程，让夫妇不得不走上离婚法庭。

导致僵局的问题可能至关重要，例如，涉及孩子的监护权和探视协议的条款，或是以同等激烈的程度争夺一条圣伯纳宠物狗；它也可能是一件相对微不足道的事情，比如一个人要花多少钱买断对方的分时度假合同。不管怎么说，是伴侣两人固执己见，拒绝在协议的最后一项上做出任何有意义的让步，妨碍了所有的进展。我问桑迪，碰到这种情况，她一般会对争执者说些什么。她说，她会把涉及问题的最后一轮提议从一个房间拿到另一房间说，“只要你同意这一条，我们就成交了”。我感觉自己发现了问题所在，并提议将措辞做一处小小的改动：“我们成交了。你只要同意这一条就行。”

几个月后，在一次聚会上，桑迪满面笑容地走过来告诉我，调整后的措辞非常成功。“每次都管用。”她说。我有些怀疑，问道：“什么？每一次？”她用手拍了拍我的胳膊说：“是的，每一次都管用。”

尽管我仍然对100%的成功率有所怀疑（我们讨论的是行为科学，而不是魔法），但毫无疑问，我仍为自己的提议所产生的效力感到高兴。不过老实说，我也并非大感意外。我提出这样的建议，是因为知道两件事。其一是我对行为科学相关研究工作的认识。例如，我知道有一

项针对佛罗里达州立大学学生的研究结果显示，和大多数受调查的本科生一样，他们对校园食堂的伙食质量并不满意。9天后，研究者做了第二次调查，本科生们改变了主意。这是为什么呢？很有意思，导致他们改变看法的事情，跟食堂的服务质量没有关系，而且食堂的服务质量也丝毫没有改变。原因在于，进行第二次调查的那一天，学生们得知，由于发生火灾，他们在接下来的两个星期都没有机会在学校食堂里吃饭了。

其二是我从本地电视台里看见的一件事，跟桑迪找我求助大约是一时期。这幅场景现在很常见了：新一代苹果手机上市后，购买者会绕着门店所在的街区排起长龙，有些人还会带着睡袋守候一整晚，只等到商店开门后冲进去抢购（见图6-1）。在iPhone 5推出的那天早晨，我所在城市的一家电视台派出记者前去报道这一现象。记者走到一位比她早到很多、排在队伍第23位的女士，问她是怎样熬过这段漫长的等待时间的，以及她是否曾跟周围的人闲聊。这位女士回答说，大家花了很多时间谈论iPhone 5的新功能，也聊了不少关于彼此的情况。她甚至透露，她排在队伍的第25位，到了半夜，她偶然跟排在第23位的人聊起天来。对方喜欢她价值2 800美元的单肩包。于是她抓住这个机会，提出做一笔交易：“用我的包换你的排位吧。”她成功了。在这段让那位女士颇为自得的故事结束时，目瞪口呆的记者问：“可是……为什么这么做呀？”那位女士给出了一个生动的答案：“因为我听说这家店货源不足，我不想失去这个买到手机的机会。”



图6-1 传染性竞争

这名男子买到了新一代的苹果手机，忍不住欢呼起来。为了保证自己能买到，他整晚都在排队等候苹果专卖店开门。

我当时一听到那位女士的回答就忍不住坐直了身子，因为它完美印证了长期研究所得的结果。这些研究表明，尤其是在存在风险和不确定感的条件下，人们更愿意去做出避免价值损失的选择，而不是选择去获得某种东西。一想到有可能买不到一部人人都想要的手机，排在第25位的那位女士意识到了其中蕴含的不确定风险，由此安排了一笔代价昂贵的交易，以免错失良机。诺贝尔经济学奖得主丹尼尔·卡尼曼

（Daniel Kahneman）⁽⁴⁰⁾提出的“前景理论”，就以整体上的“损失厌恶”，即较之获得某物的前景，损失某物的前景更能驱使人们采取行动这一设想为核心。来自多个国家、多个领域（商业、军事和职业体育）的研究结果，均普遍支持前景理论。例如有研究发现，商业世界的管理者在做决定时更看重潜在损失而非潜在收益。同样的道理也适用于体育行业，与可能获得收益的情况相比，可能招致损失的情况会让决策者思考更长时间。因此，参加美国职业高尔夫巡回赛的球手，会花更多的时间和精力在推杆上，以防止标准杆失球（避免打出柏忌），而不是标准杆得球（打出小鸟球）。

科学研究告诉我什么是损失厌恶心理，前文提到的原本排在第25位的那位女士身上看到它发挥了强大的作用，这两点认识何以激发了我向桑迪提出那条建议呢？我认为，如此调整措辞能让客户一开始就意识到，自己拥有了想要的某种东西，“我们成交了”。如果他们不愿妥协，就会失去这笔交易。而在桑迪原本的措辞里，交易还有待达成，“只要你同意这一条，我们就成交了”。鉴于我所了解的情况，调整措辞是我很容易想到的做法。

READER'S REPORT | 读者报告

致西奥迪尼先生：

有一年，在采购圣诞礼物时，我看到一条自己挺喜欢的黑裙子。但我没钱买它，因为那时我是在帮别人买礼物。我问商店店员能不能把它先放到一边，等星期一放学之后，我带着妈妈一起来看看。但是店方称他们不能这么做。

我回家就对妈妈说了这件事。她告诉我，要是我真的喜欢那条裙子，她可以借钱给我买，等我有了钱再还给她。星期一放学之后，我去了那家商店，却发现裙子没了，别人已经买走了。直到圣诞节我才知道，当星期一我还在学校的时候，妈妈就跑去了商店，把我对她描述的那条裙子给买了。

尽管那年的圣诞节已经过去很久了，可我仍然觉得那条裙子是我的最爱，因为我曾一度以为自己失去了它，这种失而复得的感觉让它变成了我的一件宝贝。

住在纽约上城区的这位妇女：

我们有必要想一想，为什么损失能如此有力地影响人类的行为。从进化的角度来看，“避免损失”优先于“获得收益”。有个理论可以解释这一点：如果人本来有足够的资源可以生存下去，这时，资源增加只不过是锦上添花罢了，但资源减少却可能致命。因此，在适应性进化中，人变得对损失极其敏感。

虽然损失厌恶是稀缺原则的一个核心特征，但它只是这一原理包含的若干因素之一。我们不妨更完整地看看这条原理。

物以稀为贵

几乎人人都会受到稀缺原则的影响。各类收藏家，不管是收藏棒球卡的，还是收藏古董的，他们都很清楚稀缺原则在决定物品价值上的影响力。如果一样东西少见，或越来越少见，那它就更贵重，规律就是这样。最能说明稀缺性在收藏市场重要地位的是一种叫“珍贵的错误”的现象。有时候，瑕疵品（模糊的邮票、冲压过重的硬币）的价值最高。因此，倘若一张邮票上的乔治·华盛顿是三只眼，虽然这在解剖学上不正确，在美学上也缺乏吸引力，那它一定备受收藏家的追捧。让人感到讽刺的地方就在这儿，倘若瑕疵把一样东西变得稀缺了，垃圾也能化身成值钱的宝贝。

亲身经历了机会越少见，价值似乎就越高这一稀缺原则之后，我开始注意到它对我行为方方面面的影响。举例来说，我经常中断有趣的面对面交谈，接听一通未知来电。此时，来电者拥有一种坐在我对面的客人所不具备的重要特点：如果我不接电话，就可能跟电话中的信息失

之交臂。哪怕眼下的交谈是多么热烈、多么重要，又哪怕我通常接到的电话讲的都是些没什么大不了的事，可电话铃每多响一声，我接起它的机会就少一分。出于这个原因，在那一刻，我总是会优先拿起电话。

如我们所见，对失去某种东西的恐惧，似乎要比对获得同一物品的渴望，更能激发人们的行动力。比如，大学生们在想象恋爱关系或考试中所失的时候，情绪波动会比想象所得的时候更强。在英国，较之节省能源，居民们更有可能为了避免账单金额过高而更换电力供应商。在后一种情况下，居民更换供应商的可能性比前一种高45%。在开展任务时，人们为了避免损失而作弊的可能性比为获得利益而作弊的可能性更大。这种倾向并不仅限于与金钱有关的事件。在一项研究中，团队成员为了避免在团队中地位下降而作弊的意愿，比为地位获得同等程度提升而作弊的意愿强82%。此外，相较于收益，损失对关注度（凝视）、生理唤醒（心率和瞳孔扩张）和大脑激活（皮层刺激）所产生的影响更大。

处在面临风险和不确定感的条件下，遭受潜在损失的威胁能强有力地影响人们的决定（见图6-2）。健康研究者亚历山大·罗斯曼

（Alexander Rothman）和彼得·沙洛维（Peter Salovey）把这点发现应用到了医疗领域。医生常常鼓励人们进行身体检查，如及早进行胸部肿瘤X光透视、艾滋病毒筛查、癌症自检等，以便及时查出疾病。由于体检可能会查出某些疾病，甚至可能发现一些很难治愈的疾病，所以，宣传时着重强调潜在损失最合适不过了。例如，建议年轻妇女自检乳房癌症的小册子，站在不这么做就有可能损失什么的立场上宣传，效果明显要比强调这么做能带来什么要好得多。商业世界的研究发现，管理者对潜在损失比潜在收益看得更重。就连我们的大脑似乎也是为保护我们免遭损失而进化的，阻挠着眼于损失所做出的明智决定，要比阻挠着眼于收益的决定难得多。[\(41\)](#)

天伦之乐，美好瞬间。
别让它消逝。

老年性黄斑变性是 55 岁以上人士失明的主要原因。别因为你或爱人未能发现视力模糊、阅读需要强光等预警信号而错过生活的美好瞬间。及时检查是保护视力的关键。
保护视力，别让它消逝。

请联系抗盲基金会
(Foundation Fighting
Blindness)，免费获取
预防及管理老年
性黄斑变性的信息。



抗盲基金会

治疗方法触手可及，
请致电 800-610-XXXX

图6-2 不要损失（视力）

这则广告的发布方是一家资助老年性黄斑变性相关研究工作的慈善基金会。它明智地借助互惠（免费提供应对病变的信息）和损失厌恶（描绘不容错失的美好瞬间）这两种心理，激发捐赠者慷慨解囊。

数量有限

由于稀缺原则在我们确定事物价值时会产生强大的影响力，很自然地，说服专业人士就会搞些类似的小把戏。最直截了当的做法是所谓的“数量有限”策略，即告诉顾客，某种商品供不应求，不见得随时都有。在打入各类组织研究顺从策略期间，我曾亲眼见识过它在多种环境下的应用：“使用这种发动机的敞篷汽车，我们州只有不到5辆。而且等它们卖完了就真正没货了，因为这种车型已经停产了。”“整个楼盘里，

只剩两个户型还没卖掉，这就是其中之一。另一个户型你肯定不想要，因为它是东西朝向的。”“你大概得认真考虑一下今天要不要多买一箱，因为工厂那边已经忙不过来了，我们不知道什么时候才能再进货。”

READER'S REPORT | 读者报告

致西奥迪尼先生：

我一直在一家叫布克曼的转售商店运用稀缺性原则。这家店收售二手图书、唱片和玩具。我家有**20世纪90年代**儿童电视连续剧《理查德·斯凯瑞》（*Richard Scarry*）的一些角色玩偶，我把它们带到了布克曼二手店销售，但一个都没卖出去。于是，我决定每次只带一款玩偶单独寄卖。果然，每次都被买走了。现在，我已经把它们都卖掉了。稀缺性原则实在好用！

我爸爸用棒球队的烈酒杯在易趣上也做过同样的事。他总共花了**35美元**买了一盒**24支**的烈酒杯，然后放到易趣上单独出售。第一支杯子就卖出了**35美元**，赚回了所有成本。等了一段时间后，他又把第二支酒杯放上网，卖了**26美元**。他又等了更长时间以后，以**51美元**的价格卖出了第**3支**酒杯。接着，他有些贪心，很快放上第**4支**酒杯，这一回只卖了**22美元**。他吸取了教训。他至今仍留着几支酒杯，准备等它们足够稀缺的时候再出手。

亚利桑那州菲尼克斯的这位女士：

把原本充足的物品分多次出手，每次只卖一个，这一招聪明的地方在于，使用者意识到了充足是稀缺的对立面，因

此，一件商品如果数量多，会降低人们对它的感知价值。

有些时候，数量有限的信息是真的，可有时完全是假的。不过，不管信息是真是假，卖家的用意都一样：让顾客相信一样东西很紧俏，从而提高它们在顾客眼中的价值。说服专业人士能把这种简单的手法耍出各种式样和风格来，我不得不承认，他们很叫人佩服。不过，令我印象最深的还是基本方法的一个加强版，它把这种逻辑延伸到了极端的程度，把商品在它最稀缺的点（即卖掉之后就没有了）上卖出去。我曾暗访过一家把这一手法应用得出神入化的家电商场，那儿30%~50%的存货都打着降价出售的旗号。

影响力研究

假设从远处看，商店里有对夫妇似乎对某样待售的电器感兴趣。有很多线索能说明他们感兴趣，如近距离地检查电器，翻看与该电器相关的说明书，在它前面展开讨论，但他们并不打算找销售员进一步了解情况。观察到该夫妇的兴致颇高之后，销售员可能会走过去说：“我看到你们对这台电器很感兴趣，我大概猜得出是为什么，它质量好，价格又很优惠。但很遗憾，20分钟之前，我已经把它卖给另一对夫妇了。而且，要是我没记错的话，这是最后一台了。”

顾客听后一脸失望。因为买不到了，这台电器突然变得更具吸引力了。一般来讲，夫妇中有一个会问，店里的库房或其他分店是否还有这一型号的电器。“嗯，”销售人员会说，“有可能。我去查查吧！但我不敢确定你们是否真的很想要它，如果我帮你按这个价格找到多余的一台，你愿意买下来吗？”

这套手法最漂亮的地方就在这儿。根据稀缺原则，务必要在一件商品最不可得、因此也显得吸引力最大的时候要顾客承诺购买。在这个奇妙的脆弱时刻，许多顾客也当真会答应购买。因此，销售员通常会带着好消息回来（从无例外），说找到了额外的库存，同时他们手里还握着笔和销售合同。得知想要的机器货源还很充足时，实际上有可能再次让一些顾客觉得它没那么大的魅力，可这时，双方的交易已经进入大多数人都没法食言的程度了。在先前的关键点，他们已经下定了购买的决心，又当众做了承诺，所以他们不买也不行了。

在向商业群体介绍稀缺性原则时，我会强调，不要提供“数量有限”的虚假信息来使用这一手法，这很重要。人们一般会问我：“但要是我们卖的东西并不短缺呢？如果我们的商品数量完全能满足市场需求呢？那该怎样利用稀缺性的力量？”解决办法是要意识到，稀缺性不光适用于物品的数量，也适用于物品的特征或组成要素。从你的产品或服务里找出一种独特或不常见的要素，该要素从其他地方根本无法以同样的价格买到。接着，在此基础上诚实地进行营销，告诉买家如果错过了机会，就会丧失随之而来的好处。如果所售物品没有这样的特征，但它很可能拥有一套竞争对手无法匹敌的独特的组合特征。如果是这样，诚实地宣传该独特组合特征的稀缺性即可。

时间有限

亚利桑那州的梅萨市最引人注意的地方，大概要数当地数量可观的摩门教徒了。这里摩门教徒的人数仅次于全世界最大的摩门教定居点盐湖城。此外，市中心环境优美的地方，还耸立着一座宏伟的摩门教堂。虽说我从远处欣赏过那儿的风景和建筑，我却从没想过要走进教堂看一看，因为我对那儿不怎么感兴趣。可有一天，我从报上读到了一篇文章，说凡是摩门教堂，都设有一处特别的暗室，只有虔诚的教徒方可入内。不过，这个规矩也有例外。教堂新修好之后的那几天，整座建筑都

允许人参观，连那处暗室也不例外。

报道上还说，梅萨教堂刚刚翻新，按教会的标准，“翻修”也属“新建”之列。因此，未来几天里，不是摩门教徒的游客也可以看到平常禁止入内的暗室。这篇文章对我产生的影响至今让我记忆犹新，我立刻决定去看一看，可等打电话约朋友盖斯一起去的时候，我才逐渐醒悟，明白了这是怎么一回事，随即放弃了这个念头。

盖斯拒绝了我的邀请，他还很好奇我为什么会这么急切地想去参观教堂。我不得不承认，我从前从没想过要去教堂参观，我对摩门教没什么渴望解答的问题，对宗教建筑也丝毫不感兴趣，我也不指望在梅萨教堂看到比本地区其他教堂更精彩、更具刺激性的东西。我一边说，一边回过了神，那儿吸引我的唯一原因是，要是我不赶快去看一眼禁区，以后就不会有机会了。一样本来对我毫无吸引力的东西，仅仅因为以后恐怕看不到了，就立刻变得迷人起来。

EBOX | 线上影响力

在一份令人印象深刻的在线商业网站实验综述中，两名研究者汇总了超过6 700个A/B测试的结果。所谓A/B测试，检验的是同一个电子商务网站在包含或不包含某一特点（有时对比的是包含甲特点或含乙特点）时效力如何。接受评估的29种功能，即包括纯粹的技术（如是否包含搜索功能，是否有回到页面顶部的按钮，是否有默认设置），也涉及动机（如免费配送、产品徽章和行动呼吁）。在调查报告的结尾，两位研究者总结道：“我们测试中的最大赢家都有着行为心理学的基础。”本书读者可以感到欣慰的是，测试中显示的最有效的6个特点，都涵盖于我们介绍过的影响力原则中：

稀缺——突出显示库存较低的品目。

社会认同——描述最流行、最热门的品目。

紧急（稀缺）——时间有限，通常带有倒计时。

让步（互惠）——为网站访客提供折扣。

权威 / 专业知识——告知访客可选的替代产品。

喜好——在网站上展示产品受欢迎的信息。

作者点评：早在电子商务出现之前，数量有限和时间有限这两种强调稀缺的做法就很盛行了，而且，两者都排在上述综述里总结的前三大特点当中。这很能说明问题。我们再一次看到，尽管施加影响力的平台可能发生了巨大变化，但这些原则给人类反应带来的影响并未改变。同样具有启发意义的是，这两种可操作的稀缺手法的效力大小，与其他研究结果吻合，即一般而言，数量有限比时间有限的效力更大。在接下来关于竞争的内容里，我将介绍其中的原因。

人往往会因为剩下的时间不多而变得更想要做某事。说服专业人士往往会利用这种倾向，通过“时间有限”的手法来牟利。就跟我想去摩门教堂密室的情形差不多，人们经常发现，仅仅因为剩下的时间不多了，自己就会跑去做本来并不太喜欢做的事情。熟练的商家会利用这一倾向，对顾客设置最后期限并广而告之，激起顾客本来没有的兴趣。这种方法主要集中在电影宣传上。最近我看到一家电影院在只有短短一句话的传单上就运用了整整三次稀缺原则（这么做是什么目的，我们当然都知道）：“专场放映，座位有限，欲订从速，过时不候！”

一些面对面施加高压推销的卖家十分青睐“时间有限”的各式手法，因为它规定了拿主意的最后时限——现在。他们经常告诉顾客，要赶紧下决心买，要不然之后的购买价会更高，甚至根本买不到了。

销售员或许会对有意参加健康俱乐部或购买汽车的主顾说，眼下他报出的交易价格很优惠，而且机会仅此一次，如果顾客放弃这个机会，交易就结束了。一家儿童肖像摄影公司力劝家长，对孩子拍下的各种姿

势的照片要尽量多买，因为“我们的存储空间有限，你家孩子没卖出去的照片，24小时内就会销毁”。上门卖杂志的推销员可能会说，他们只在顾客所处的地区待一天，之后他们就走了，顾客买这份杂志的机会也就没了。

我曾打入一家卖家用吸尘器的公司，他们要销售学员说：“我还有很多别的客户要去拜访，每一家我只来一次，这是公司的规定。就算你之后决定要买这台机器，我也不能再回来卖给你了。”这当然是无稽之谈，既然这家公司和它的销售代表是做买卖的，那么只要客户肯请他们再来，他们定会很乐意。正如销售经理对受训学员所说，说自己不能再来，实际上真正目的跟日程安排太紧毫无关系。这是“为了防止潜在客户思前想后而吓唬他们，要他们相信现在不买，以后就买不到了”（见图6-3）。

骗局

彼得·克尔 / 文 《纽约时报》

丹尼尔·居尔班已经记不得自己一辈子的储蓄是怎么消失的了。他只记得电话里传来一个温柔可人的声音，他只记得自己关于石油和白银期货的发财梦。但时至今日，这位81岁的电力局退休工人还没搞明白骗子是怎么说服他花掉18000美元的。

“我只是想在我余下的日子里过得更好些，”家住加利福尼亚州霍尔德的居尔班说，“可等我知道真相后，我吃不下，睡不着，瘦了整整13公斤。我还是不相信我竟然会做出这样的事情来。”

居尔班是执法官所谓的“锅炉房行动”的受害者。这种骗局通常会找上几十个电话推销员，每天给上千个客户打电话。美国参议院下属的一支调查委员会在去年就此事发表过报告。据报告说，这种公司每年能从上当的客户那里捞到数亿美元。

“他们会使用一个令人印象深刻的华尔街地址，说一番天花乱坠的好听的话，骗人把钱投到各种冠冕堂皇的迷人计划里去。”纽约州总检察长罗伯特·艾布拉姆斯说。过去四年里，他经手了十多宗“锅炉房”案件，“有些受害者被他们灌了迷魂汤，把一辈子的积蓄都投丢了”。

据投资者保护和安全局负责人、纽约助理总检察长奥列斯特斯·米哈伊说，这类公司的操作大多分为三个阶段。首先是“开场白”，推销员说自己代表一家名字和地址听起来就很气派的公司，他只简单地询问一下潜在客户是否愿意接收自己公司的宣传资料。

第二通电话里就包含有推销的手法了。推销员先把能赚多少多少钱吹嘘一番，接着就告诉客户，可惜这个项目已经不再接受投资了。第三通打给客户的电话说，现在还有机会参加这笔交易，但时间极为紧迫。

“他们的用意是在买家眼前挂上一根胡萝卜，然后又把它拿走，”米哈伊说，“目的是让人不假思索地赶紧投资。”有时，第三通电话里的推销员会一副上气不接下气的样子，告诉客户自己“刚刚才从交易大厅回来”。

就是这种策略说服了居尔班，让他把毕生的积蓄交了出去。居尔班说，一个陌生人反复打电话来，让他电汇1756美元到纽约买白银。之后，销售员又打了好几轮电话，哄骗居尔班再电汇6000美元买原油。他最终又汇出了9740美元，可到头来一分钱利润也没看到。

图6-3 稀缺骗局

请注意，第二通和第三通电话里应用到了稀缺原则，销售员要居尔班先生“别想太多尽快买”。

多项实验结果都表明，拥有独特元素的产品和体验对消费者具有强

烈的吸引力。稀缺物品恢复充分供应后，它就失去了对人们的吸引力。我曾收到一位来自明尼阿波利斯的妇女的读者报告，它强有力地说明了另一个相关的观点：面对某一稀有物品，我们以为自己是因其固有特质而喜欢它，可一旦它失去稀缺性，我们便会大感意外，并因此对它不再感兴趣。报告中说：“虽然我来自美国，但一直很喜欢玩伦敦大本钟的拼图。在美国，这种拼图很少见，我每次碰到都倍觉兴奋。但易趣出现之后，我能在网上搜到很多此类拼图。一开始，我每一款都买了，但买完就失去了兴趣。你的书让我意识到，大本钟拼图的稀缺性，是我对它产生迷恋的主要原因。我喜欢玩大本钟拼图足足有23年了，但等我能找到大量大本钟拼图之后，我再也没有拼图玩的欲望了。”

逆反心理

至此，证据已经很确凿了。说服专业人士经常把稀缺性当成影响力武器，其方式和用途多样，而且自成体系。一种原理会成为影响力武器，那么可以肯定它左右人行为的力量不容小觑。就稀缺原则而言，它的力量主要来自两个方面。

第一点我们应该很熟悉了，和其他影响力武器一样，稀缺原则钻了我们思维捷径上的漏洞。这个漏洞本来也自有道理。我们都知道，难于得到的东西，一般都比能轻松得到的东西好。因此，我们基本可以根据获得一样东西的难易程度，迅速准确地判断它的质量高低，因为我们往往不愿意失去质量高的东西。这也就是说，稀缺原则成立的一个原因在于，根据它来做出判断，大部分时候是正确的。

因此，“稀缺的东西更珍贵”的观念在人们心中越发根深蒂固，而人们也会逐渐认为它反过来也成立，即珍贵的东西就稀缺。20世纪60年代中期，杰克·布雷姆（Jack Brehm）提出了逆反理论，随后的研究为它提供了相当大的支持。

此外，稀缺原则的力量，还有第二种独特的来源：机会越来越少的话，我们的自由也会随之减少，而我们又痛恨失去本来拥有的自由。保住既得利益的愿望，是心理逆反理论的核心。心理学家杰克·布雷姆（Jack Brehm）提出了心理逆反理论，以此解释人类在丧失个人控制权时做出的反应。根据这个理论，只要选择自由的行为受到限制或威胁，保护自由的需求就会使我们想要它们（以及与其相关的商品和服务）的愿望愈发强烈。因此，一旦短缺或其他因素妨碍了我们获取某物，我们就会比从前更想要得到它，更努力地想要占有它，跟这种妨碍对着干。

心理逆反理论的核心看似简单，但它却盘根错节地交织在各种社会环境当中。不管是年轻人在花园里谈情说爱，丛林里爆发的武装革命，还是市场里的水果交易，人类的大量行为都可以用逆反心理来解释。不过，在讨论开始之前，我们最好来看看人最初是从什么时候表现出跟自由限制对着干的欲望的。

孩提逆反

儿童心理学家把这种倾向追溯到了两岁，家长们早就知道有这个问题，还给它起了个名字“可怕的两岁”。大多数家长都可以证明，自己的孩子在这一时期会有更多的逆反行为。两岁的孩子似乎掌握了抵挡外界压力，尤其是父母的技术。告诉孩子这样做，他们却偏要那样做；给孩子一样玩具，他们却偏想要另一样；把孩子抱起来，他们却拼命挣扎着要你放下；把他们放下来，他们又声嘶力竭地要你抱。

弗吉尼亚州进行的一项研究巧妙地捕捉到了一群男孩在“可怕的两岁”时的行事风格。

影响力研究

男孩们在妈妈的陪同下，进入了一个房间。房间里摆着

两件同样好玩的玩具，但一件玩具放在透明有机玻璃屏障旁边，另一件则放在屏障后面。实验里的有机玻璃板有一部分只有30厘米高，算不上真正的屏障，因为孩子们能很容易地翻过它拿到后面的玩具；另一部分却有60厘米高，足够拦着孩子不让他们拿到玩具，除非他们绕过屏障。研究者希望看一看，置身如此条件之下，小宝宝们要多快才能接触到玩具。研究结果很清楚，在屏障矮得拦不住后面的玩具时，男孩们对两件玩具并未表现出特别的偏好，平均而言，接触屏障旁边玩具的速度跟接触屏障后面玩具的速度差不多；可一旦屏障高得成了真正的障碍，孩子们便会立刻奔向被拦着的玩具，接触它的速度比接触没拦着的玩具时要快上3倍。总之，在这项研究中，面对限制了自己自由的东西，男孩们表现出了典型的“可怕的两岁”做法：直接挑衅。

为什么逆反心理会出现在两岁的时候呢？或许答案跟这一时期大多数孩子经历的一种关键变化有关。到了这个年纪，他们才首次意识到自己是个体。他们不再把自己仅仅视为社会环境的延伸，而是把自己视为有自我意识的、独立的个体。伴随着这种自主意识的发展，自由的概念也形成了。独立的个体应当有选择的余地，一个刚刚发现这一点的孩子，迫切想要探索这种选择余地的深度和广度。

因此，看到两岁的小孩总是跟我们对着干，我们其实没必要惊讶，也不必烦恼。他们正在体验一件最令人兴奋的事情：他们是独立的人类个体。在他们小小的心灵里，有那么多有关选择、权利和控制的关键问题要去问，要去答。将他们力争自由、反抗限制的倾向理解成对信息的求索，恐怕是最合适的。测试过自由的极限在哪里，顺便也测试了父母有多大的耐心，孩子们就能弄明白，在他们的世界里，自己能控制多大的地方，又有哪些地方必须受制于人。稍后我们就会看到，明智的父母总会尽力提供前后高度一致的信息。

研究者揭示了两岁男童对障碍物表现出的行为倾向，在研究中，面对同样的大障碍物，两岁的女孩并未像男孩那样表现出逆反心理。另一项研究表明，女孩们不这么做，并不是因为她们不爱跟自由限制对着干，而是因为她们主要会对来自他人的限制产生逆反心理，对物理障碍的反应则稍逊。然而，不管是男孩还是女孩，当他们长到18~24个月，第一次认识到自己的“认知自我”时，便开始把自己视为独立个体。

尽管两岁可能是逆反心理最明显的时期，但面对限制行动自由的举措，我们一辈子都会表现出强烈的反抗倾向。另一个反抗倾向最为突出的年纪是青春期。一个深明此理的邻居曾经给我建议：“要是你很想做什么事情，那么有三种选择：自己做；出大价钱找人做；要不就故意禁止你家十几岁的孩子做。”和两岁一样，个性意识萌芽也是青春期的特点。对青少年而言，这种萌芽意味着走出儿童的角色，摆脱与此角色相伴的家长控制，迈向成人一角，获得随之而来的一切权利和义务。青少年对义务想得比较少，他们更关注的是身为一个年轻成年人应有的权利，这没什么出奇。更不奇怪的是，在这些时候对其施加父母的权威，往往会适得其反。要是家长企图控制他们，那么青少年不是会阳奉阴违，就是会公然对抗。

最能说明家长施压造成青少年反抗的例子，大概要数所谓的“罗密欧与朱丽叶效应”了。大家都知道，罗密欧与朱丽叶是莎士比亚笔下的悲剧人物，两人相爱，但两个家族却是世仇。为了反抗父母拆散他们的企图，他们双双自杀殉情，用这种最极端的悲剧方式来伸张自由意志。

这对年轻人感情和行为的强烈程度，一直为该剧的观众迷惑不解。他们这么年轻，怎么会在如此短的时间里发展出如此非同一般的浓烈感情呢？浪漫的人或许会说他们之间是一种少有的理想爱情，可社会学家大概会说，这是父母的干涉及其带来的逆反心理所致。或许罗密欧和朱丽叶最初的感情也没有强到能超越家人设置的庞大障碍的地步。相反，正是因为这些障碍的存在，这段感情才发展到了如此白热化的程度。要

是听凭这对青年男女自由恋爱，他们的浓情蜜意最后说不定只是初恋时短暂的冲动罢了。

因为罗密欧和朱丽叶的故事是虚构的，这类问题当然纯属假设，所谓的回答也只是妄加揣测。不过，对当代的“罗密欧”与“朱丽叶”们，提出并解答类似的问题倒是有可能的。少年男女碰到父母干涉，是不是会发展出更坚定的恋爱关系，彼此爱得更深呢？有人对科罗拉多州140对少年情侣做了研究，发现他们确实如此。研究者发现，尽管家长干涉会令感情关系出现某些问题：如情侣中的一方以更挑剔的眼光看待另一方，更多地报告另一方的负面行为，但干涉同时也让情侣双方觉得彼此更加相爱、更想结婚了。在研究过程中，随着父母的干涉越来越多，爱的体验也越来越强。而当干涉减少的时候，浪漫的感觉也会慢慢冷却。

出现“罗密欧与朱丽叶效应”，并不意味着父母必须接受、包容自家处在青春期的孩子谈恋爱。爱情是一种微妙的游戏，新入场的玩家很可能会频频犯错，所以让更有经验、更有眼光的成年人指点一二是有好处的。家长在提供这类指导时应该意识到，青少年认为自己已经成年了，面对典型的家长—孩子式控制关系，他们不会做出很好的回应。尤其是在择偶这种显而易见的成人舞台上，用成年人的影响力武器（偏好和劝说）要比传统的家长控制（禁止和惩罚）的效果更佳。罗密欧与朱丽叶的家族固然是极端的例子，但对年轻爱侣施以铁腕控制，也很有可能把它搞成一场炽热的地下恋爱悲剧。

成人逆反

对两岁的孩子和十几岁的青少年而言，逆反心理贯穿多种体验和经历，而且总是狂躁有力的。而对我们其余大多数人来说，逆反的能量池是平平静静、隐藏起来的，只是偶尔才像火山一样爆发一次。尽管如此，这些爆发仍然以各种各样有趣的方式表现出来，不仅研究人类行为的学者感到好奇，制定法律和政策的人也能从中得到些启示。例如，超

市购物者听说有联邦官员反对支持联邦价格管制的请愿书之后，会有更大概率对此请愿表示支持。如果违背交通规范的司机以今天是自己的生日为借口请求获得宽大处理，有权处罚他的警官很可能偏偏就要在这一天惩罚他。为什么会这样呢？因为警官们会觉得自己决定是否惩罚的自由受到了环境的限制——这就是典型的逆反举动。

统计数据表明，警官有更大可能在违规者生日当天，尤其是违规者以过生日为由替自己开脱求情的情况下对其做出处罚。研究者对这一现象开展了6项研究，其中一项调查了华盛顿州13.4万名因醉驾被捕的司机，发现警察会在违规者生日当天对司机做出更严厉的处罚。

再来看看佐治亚州的肯纳索发生过的一件古怪例子。

影响力研究

肯纳索镇政府制定了一项法律，要求本地所有成年居民都持有枪支和弹药，违者入狱6个月，并罚款200美元。肯纳索持枪法所具有的每一个特点，都让它成了逆反心理的大靶子。持有枪械本来是大多数美国公民长久以来认为自己享有的一项重要自由，但这项法律却硬性规定人人非得持枪不可，因此冒犯了自由。而且，肯纳索议会在通过这项法律时，并未动员公众广泛参与。逆反理论预测，在这种情况下，镇上5 400名成年公民恐怕没几个会服从。然而，报道该消息的报纸却证实，该法通过以后的3~4个星期，肯纳索的枪支销售的确很火爆。

我们该如何理解这个明显有违逆反原理的现象呢？答案要到那些在肯纳索买枪的客人身上去找。采访了肯纳索几家商店的老板之后，我们发现，原来买枪的人并不是镇上的居民，而是游客，不少人都是看到消息之后，禁不住诱惑跑来购买自己的第一支枪的。唐纳·格林是一家商店的经营者，

报上形容她的这家店“简直就是个军火库”。唐纳说：“生意棒极了。但大部分枪械都是镇外的人买去的。按法律规定来买枪的本地人只有两三个。”法律通过之后，购买枪支成了肯纳索的频繁活动，但买枪的人却并不是那些法律有意管辖的人。当地居民大多都在消极抵抗，只有那些自由并未受到此地法律限制的人才热衷在这里买枪。”

十多年前，在离肯纳索几百公里以外的佛罗里达州迈阿密也出现过类似情况。

当时，为保护环境，迈阿密地区制定了一项条例，禁止使用和拥有含磷酸盐的洗衣剂或清洁剂。一项研究发现，对于这项法律带来的社会影响，迈阿密居民同时做出了两种反应。

首先，不少迈阿密人按本地的传统开始走私。有时候，他们会跟邻居和朋友一起开着“大篷车”，到相邻的县大宗买入磷酸盐洗衣剂。囤积现象迅速蔓延，好些人甚至为此着了迷，吹嘘自己已经囤积了足够使用20年的磷酸盐洗衣剂。

第二种反应比走私和囤积这种蓄意挑衅更为微妙而普遍。因为人们总是想要得不到的东西，迈阿密绝大多数的消费者逐渐对磷酸盐洗衣剂这种产品有了质量更好的印象。跟相邻的坦帕市居民相比（坦帕市不受迈阿密条例的管辖），迈阿密市民认为磷酸盐洗衣剂更温和，在冷水中使用效果更好，增白更佳，还能整旧如新，强力地清除污渍。等条例通过后，他们甚至觉得磷酸盐洗衣剂更容易从瓶子里倒出来。

这类反应，对那些丧失了既得自由的人是很典型的。要理解逆反心理和稀缺原则的运作，理解这一点十分重要。每当有东西获取起来比从前难时，说明我们拥有它的自由受到了限制，那么我们会越发地想要得到它。不过，我们很少意识到，是逆反心理带来了这种想要的迫切感，我们只知道自己就是想要。为了解释这种莫名的渴望，我们开始

给它安上各种积极的特质。在迈阿密禁用磷酸盐洗衣剂以及其他类似的例子中，人们的渴望与东西本身具备的优点并没有什么因果关系。磷酸盐洗衣剂的清洁、漂白和流动性并没有在遭禁止以后变得更好。我们之所以会这么想只是因为我们更想得到它了。

READER'S REPORT | 读者报告

致西奥迪尼先生：

去年圣诞节，我碰到一个**27**岁的男人，那时我**19**岁。他其实并不是我喜欢的那种，但我还是跟他约会了，大概是因为跟年纪较大的男人约会显得很有面子吧。我本来对他没多大兴趣的，可我的家人却纷纷对他的年纪表示异议。结果他们越是插手我的私生活，我就越是爱他。这段关系只持续了**5**个月，可要是我爸妈没说这说那的，能维持**1**个月就算好了。

弗吉尼亚州布莱克斯堡的这位女士：

罗密欧与朱丽叶虽然早就殉情离去了，但“罗密欧与朱丽叶效应”却延续至今，连在弗吉尼亚布莱克斯堡这样的地方都时不时地会露个面。

审查造就的逆反

因为一样东西遭到禁止而觉得它更有价值，这种倾向不仅限于洗衣剂等消费品，连信息的获取也是这样。当今时代，获取、存储和管理信息的能力越发影响到了财富和权力的分配，因此，我们有必要理解，面

对审查或限制人们获取信息的做法，我们会做出什么样的反应。我们对有可能受审查的资料会做出怎样的反应，这方面的数据很多，可对于我们对审查本身会有什么样的反应，相关证据却惊人地少。幸运的是，几项有关审查的研究得出了异常一致的结果。某类信息遭到禁止之后，我们无一例外地更想得到这种信息，相较被禁之前，我们会对其给出更有利的评价。

说到信息审查对受众的影响，最耐人寻味的一点倒不是受众比从前更渴望得到这些信息了，而是人们对得不到的信息变得更易接受、更包容了。例如，北卡罗来纳的大学生们在得知一场反对男女混住宿舍的演讲遭禁之后，对男女混住的想法越发地反对。因此，尽管从来没能听到这场讲演，学生们对它的论点却更支持和同情了。这样一来，就有可能出现以下的情况，要是一些特别聪明的人在某问题上持有不受欢迎的立场，或其观点根本站不住脚，他们或许会做一些刻意的安排，让自己的信息遭到禁止，从而博取我们的赞同。

具有讽刺意味的是，对这类人（如一些边缘群体的成员）来说，最有效的策略恐怕不是大肆宣传他们不受欢迎的意见，反倒是让这些观点遭到官方的审查，再告知公众自己遭到封杀的消息。由此看来，美国国父们在起草宪法的时候，把保护言论自由写进了宪法第一修正案，不只体现了他们坚定地倡导自由主义，还体现了他们对社会心理的透彻理解。因为他们拒绝对言论自由加以限制，也就减少了新的政治概念通过逆反心理这一非理性过程赢得支持的概率。

当然，政治理念并不是唯一容易遭到限制的东西，跟性有关的素材也一样。除了警方偶尔会打压“成人”书店和电影院之外，家长和公民团体也频频施压，要求审查教育资料（性教科书、生理卫生教科书，甚至学校图书馆收录的书籍）里有关性的内容。斗争的双方似乎都有着良好的用意，但整件事情却很不简单，因为它牵涉了道德、艺术、家长对学校的控制、宪法第一修正案保障的自由等方方面面的议题。不过，从单

纯的心理学角度出发，支持严格审查的人不妨来仔细看看在普渡大学本科生中完成的一项研究。

影响力研究

研究者向大学生们出示了一本小说的若干广告。一半学生看到的广告文案中说“本书仅限21岁以上的成人阅读”，另一半学生看的广告里没有提及这样的年龄限制。稍后，研究者询问学生们对这本书有什么感受。学生们的反应就跟对其他禁令一样，较之以为能够随意阅读该书的学生，看到了年龄限制的学生更想读这本书，并觉得自己会更喜欢它。

支持从学校课程里正式取缔性内容的人或许会说，他们的目标是减少社会，尤其是年轻人的色情倾向。可从普渡大学以及其他有关禁令效应的研究来看，官方审查的手段恐怕并不能达到这一目标。倘若研究的结果是可信的，那么，审查反倒有可能提高学生对性资料的渴望度，令他们以为自己就是偏爱这一类内容。

“官方审查”这个词通常会让我们想到对政治性材料或露骨色情材料的检查，但还有一种常见的官方审查容易为我们忽视，大概是因为它出在事发之后。法庭议案时经常有这种情况，律师提出一份证据或证词，但法官却裁定它无效，还要陪审员们忽略这项证据。从这个角度看，法官就是审查员，尽管他审查的方式比较特别，他并不禁止把信息呈交给陪审团（因为已经太晚了），而是禁止陪审团使用该信息。法官下这样的规定管用吗？有没有这样的可能，陪审员认为自己有权考虑所有可用的信息，而法官要他们忽视某一证据的做法，激起了他们的逆反心理，于是，陪审员对遭禁的证据反而给予了更大的关注？很多时候，事实正是如此。

我们看重受到限制的信息，一旦认识到这一点，稀缺原则就能被应

用到物质商品之外的领域，信息、沟通和知识都适用这条原理。从这个角度考虑，我们可以看出，想让信息变得更宝贵，不一定非要封杀它，只要把它变得稀缺就行了。根据稀缺原则，要是我们觉得没法从别处获取某条信息，就会认为它更具说服力。

据我所知，对“独家信息最能说服人”这一观点最有力的支持，来自我从前的一名学生所做的小实验。当时，这名学生是个成功的商人，开着一家牛肉进口公司。他又回到学校学习是想在市场营销方面再接受一些深入的训练。一天，在我的办公室里，我们谈起了信息的稀缺性和独占性，他决定用自己的销售员来做个研究。

影响力研究

销售员照常给公司的牛肉购买客户打电话，但采取了三种不同的陈述方式。在下订单之前，第一群客户听到的是标准销售陈述；第二群客户则除了听到标准销售陈述，还知悉了未来几个月进口牛肉有可能短缺的信息；第三群客户除了听到标准销售陈述以及牛肉短缺的消息外，还了解到供应短缺消息的来源可不一般，是公司靠某条专门渠道获得的。所以，最后一组顾客不仅认为产品的供应有限，还以为相关信息也只有少数人知道，这样就形成了一种双重稀缺状态。

实验的结果很快就出来了。公司的销售人员要老板赶紧多采购一些牛肉，因为接到的订单太多，库存供应不了了。较之只听到标准销售陈述的顾客，听说牛肉即将短缺的客户买下了双倍的牛肉。但真正推动销售量的，还要数那些通过“独家”信息知道牛肉供应吃紧的客户。他们购买的牛肉量，是第一种客户的6倍。很明显，供应短缺的独家消息显得特别具有说服力。

在商品稀缺性加信息独家性的影响的研究中，出于道德上的原因，销售员向客户提供的消息是真实的。进口牛肉的确即将出现短缺，而且这个消息也的确是公司通过独家渠道得来的。

消除逆反心理

当人们碰到一条信息，如果他们认为这是为了说服自己，就不太可能立刻接受它。原因之一在于，他们产生了逆反心理，觉得说服诉求试图减少自己做决定的自由。因此，所有想要说服听众做出改变的人，必须先战胜逆反心理。有时候，说服者试图提供证据以证明改变是正确举动，哪怕听众本人并不情愿。为此，他们会暗示接受信息的人过去欠了说服者一笔人情债（互惠），或是说明说服者是个好人所以值得认同（喜好），或是其他人也做了同样的改变（社会认同），或是专家推荐做出改变（权威），以及采取行动的机会就快要没了（稀缺）。

不过，这里还有另一种方法可以战胜逆反情绪，它不需要借助更强大的动机，而是靠减少逆反情绪的力量来战胜它们。比如，沟通者可以从一开始就提及拟议改变存在的缺陷。这种做法不仅能提高沟通者的可信度，还给了接收方正反两方面选项的信息，减少了后者被人朝着一个方向推的感觉。

在施加影响力时，为了让目标接收者产生“自由选择”的感觉，人们专门设计了一种影响策略。这叫“但随便你”技巧，强调接收者享有说“不”的自由。在一组包含了42轮独立实验的系列实验中，在请求后加上一句“要不要拒绝，随便你”，或是类似的“没问题，你想怎么做都行”，能明显提高顺从度。这一技巧适用于各种请求，不管是要求给海啸救济基金捐款，参加当面、电话或邮寄形式的无偿调查，为路上的行人购买公交车票，为乞讨者购买食物，甚至答应对家庭垃圾分类并做一个月的记录。这种让人重新获得自由选择感的措辞影响相当大，它的成功率往往比不包括此类措辞的标准请求高出一倍以上。

最佳条件

和其他有效的影响力武器一样，稀缺原则也有最为适用的条件。因此找出什么时候它对我们最起作用就是一种重要的防御措施。通过心理学家斯蒂芬·沃切尔（Stephen Worchel）和同事设计的一个实验，我们可以了解不少这方面的信息。

影响力研究

沃切尔和他的研究小组使用的基本程序很简单。他们给一群消费者偏好研究的被试从罐子里拿了一块巧克力饼干，并让被试给饼干的质量打分。有一半的被试看到罐子里装着10块饼干；另一半则看到罐子里只有2块。正如稀缺原则预测的，被试吃的饼干属于仅有的2块之一时，给出的评价更高。较之供应充裕的饼干，人们觉得短缺的饼干味更美、吸引力更大、价格更高，虽说两种饼干根本就是一模一样的。

照我的直觉看来，1985年，可口可乐公司在犯下《时代》杂志称为“这个十年最沉痛的营销惨败”时，一定很希望自己能及时知道上述科学发现。那一年的4月23日，该公司决定从市场上撤下传统的可口可乐配方，以“新可乐”取而代之。此举激怒了粉丝们。用一则新闻报道的话来说：“可口可乐公司没能预见到这一行动在消费者心中掀起的全然狂怒和抵制。东起缅因州的班戈，西至加利福尼亚州的伯班克，从底特律到达拉斯，数十万的可口可乐爱好者们团结起来，痛斥新可乐的味道，要求老可乐回归。”

说到这种因为失去老可乐而产生愤怒与渴望交织的例子，我最喜欢的故事是关于西雅图一位叫盖伊·穆林斯（Gay Mullins）的退休投资者的。他因创办“美国老可口可乐饮用者协会”而成了全美的名人。这是一

家成员广泛分布在各行各业的组织，他们不懈地努力，利用一切可以利用的民事、司法甚至立法手段，要求传统配方重回市场。例如，穆林斯先生威胁要对可口可乐公司提起集体诉讼，以公开老配方；他向数万人分发了反对新可乐的胸针和T恤；他设立了一条电话热线，以便愤怒的市民发泄怒火、表达感受。两次单独进行的盲品测试表明，穆林斯其实更喜欢新可乐而非老可乐的味道，但他觉得这无关紧要。穆林斯先生更喜欢的东西，对他而言反而不如他遭到剥夺的东西有价值，这是不是很有意思呢？

值得一提的是，哪怕在屈服于顾客的要求，将老可乐重新上架之后，可口可乐公司的高管们还是被自己惹出来的麻烦刺痛了一下，感到颇为茫然。时任公司总裁的唐纳德·基奥（Donald Keough）说：“这是一道奇妙的美国谜题，一个可爱的美式费解。你没法衡量它，就像你无法衡量爱、自豪感或爱国主义一样。”对此，我并不认同基奥先生的看法。首先，只要你了解稀缺原则背后的心理学，这就完全没什么神秘的了。一种产品与一个人的历史与传统紧紧地联系在一起（一如可口可乐在这个国家一贯以来的地位），那么，一旦得不到这种产品，这个人也就更加想要它。其次，这种冲动是可以衡量的。事实上，我认为可口可乐公司在做出这个蹩脚的决定之前，已经在自己的市场调查中衡量过它了，但因为他们的不了解影响力原则，所以没能看到它。

可口可乐公司在营销调研方面并不吝啬；它愿意花费数十万美元，甚至更多的费用，以确保为新产品做出正确的市场分析。在决定将产品改为新可乐之前，他们同样是这么做的。从1981年到1984年，他们在25座城市对近20万人进行了口味测试，仔细测试了新老配方。他们在盲品测试中发现，偏好新可乐口味的人占了55%，偏好老可乐的人占了45%。然而，有些测试不是用无标记样品进行的，即被试事先就知道哪一种是老可乐，哪一种是新可乐。在这样的条件下，对人们新可乐的偏好还要增加6%。

你或许会说，“真奇怪，测试明明显示出人们更喜欢新可乐，可当公司最终推出新可乐，人们却对老可乐表达出了不容置疑的偏爱。他们的态度为何前后不一呢？”要解开这个谜，唯一的办法就是应用起稀缺性原则来：在口味测试期间，人们无法购买到的是新可乐；所以当人们知道哪种样品是新可乐时，便对自己无法得到的东西表现出了特别强烈的偏好。可随后，公司用新配方取代了老配方，这次人们得不到的成了老可乐，于是，老可乐又成了人们的最爱。

我想说的是，公司从盲测结果和已知口味测试结果中看到了人们对新可乐的偏爱增加了6%这一点没问题，但他们的阐释出了错。他们自言自语地说：“呀，太好了，这就意味着当人们知道自己得到了某种新东西后，他们对它的欲望便直线上升了。”而实际上，这6%增幅的真正含意是，当人们知道自己不能拥有某种东西时，对它的欲望便会直线上升。

尽管这样的结果再次为稀缺原则提供了有力的证明，但并没有透露出什么我们还不知道的东西。我们只是又一次看到稀缺的东西更招人喜欢，更贵重。事实上，饼干实验的真正价值来自两点额外的发现。让我们来逐一看看。

新的稀缺：昂贵的饼干和民权冲突

第一点值得注意的结果，是在实验基本程序稍加改动之后出现的。此时，饼干的稀缺状态并不是一成不变的。

我们前面提到，研究者先给部分被试看装着10块饼干的罐子，之后又换成装有2块饼干的罐子。因此，在张嘴吃饼干之前，有些被试眼睁睁看到饼干从供应充裕变成了短缺。而另一些被试从最开始就知道供应短缺，因为他们的罐子里本来就只有2块饼干。靠着这样的做法，研究者想解答一个有关稀缺类型的问题：我们觉得新近变得短缺的东西更宝

贵，还是一直就短缺的东西更宝贵？在饼干实验里，答案一目了然。较之一贯短缺，对从充裕变到短缺的饼干，人们的反应更为积极。

新出现的稀缺更能让人觉得迫切这个概念的适用范围，远远不只饼干实验。举例来说，社会科学家已经确定，这种稀缺是造成国家政治动荡和暴乱的主要原因。这一观点最重要的支持者詹姆斯·戴维斯（James C. Davies）曾指出，一个国家经济和社会条件改善后，要是在短期内出现剧烈逆转，最有可能爆发革命。而且，最容易起义的，不是那些传统上最受压迫的底层人民，这些人已经把自己的贫困潦倒看成社会的自然秩序了。相反，走上革命道路的，往往是至少品尝过了更美好生活的人。他们经历并习以为常的经济和社会进步突然之间可望不可求了，于是他们对进步产生了更为迫切的渴望，甚至不惜采取暴力来保护既得的进步。例如，很少有人知道，在美国独立战争时期，移民们的生活标准在西方世界是最高的，纳税也最低。据历史学家托马斯·弗莱明

（Thomas Fleming）的说法，要不是英国人想通过加税从这种普遍繁荣里分上一杯羹，北美移民根本不会造反。

戴维斯为这一新颖的论点从各地的革命、暴乱和内战中收集了一系列令人信服的证据。从法国、俄国和埃及的革命，以及19世纪罗得岛州的多尔叛乱、美国南北战争，到20世纪60年代的城市黑人反抗斗争……在上述每一个例子当中，较长的安定发展时期后都出现了一连串紧张的倒退，并最终引发了武力斗争事件。

20世纪60年代中期，美国城市的种族冲突，是一个我们许多人还能回想起来的好例子。当时经常能听到人们问：“这种事情为什么会出现在这当口呢？”在美国300多年的历史里，黑人大部分时间都饱受奴役，生活贫困潦倒，可他们却选择在社会最为进步的60年代发出了反抗的怒吼。的确，正如戴维斯指出的，第二次世界大战结束后的20年给黑人群体带来了巨大的政治和社会利益。40年代，美国黑人在住房、交通和教育领域还都会受到严格的法律限制，而且，就算教育程度相同，黑人家

庭的平均收入也只有白人家庭的一半多一点。15年后，情况有了很大的改观。联邦已经立法废除了学校、居民区、公共场合和工作环境下的正式及非正式种族隔离；经济状况也有了很大的改善，黑人家庭的收入已经提高到同等教育程度的白人家庭的56%~80%。

根据戴维斯对社会状况的分析，之后出现的一连串事件阻碍了前些年因为飞速进步所带来的昂扬乐观的氛围。首先，跟政治和法律上的变化比起来，社会现实的变化落在了后面。尽管20世纪四五十年代通过了一系列的进步立法，可大部分的街区、就业岗位和学校仍然把黑人摒弃在外。因此，跟政治上的胜利相比，身处的现实显得更像是一种挫败。例如，1954年，最高法院裁定，所有公立学校都必须取消种族隔离，可在这之后的4年里，为了阻挠学校的种族融合，出现了530起以黑人为目标的暴力事件，如直接威胁黑人儿童和家长的爆炸和纵火等。暴力行为让人感觉，黑人民众的权益再一次出现了倒退。第二次世界大战之前，每年平均会出现78起针对黑人的私刑。但在60年代，黑人群体头一回担心起家人的基本安全来。新的暴力事件也并不仅限于教育领域，和平的民权示威游行也频频跟充满敌意的人群和警察发生对峙。

其次，还有一种形式的倒退出现在黑民众的经济状况上。1962年，黑人家庭的平均收入下滑至同等教育程度的白人家庭的74%。按戴维斯的说法，本来，74%这个数字代表的是第二次世界大战结束后长期发展的结果，可在当时的黑人眼里，它意味着50年代中期兴旺繁荣后的短期倒退。1963年爆发了伯明翰骚乱，之后又断断续续地发生了多起后继的反暴力示威，最终累积导致了瓦茨、纽瓦克和底特律的大规模反暴运动。

跟革命历来的鲜明模式一样，长期的进步一旦遭到阻碍，美国黑人的反抗情绪就会比进步开始之前还要强烈。这种模式为统治者提供了一条宝贵的经验：自由这种东西，给一点又拿走，比完全不给更危险。倘若政府想要从政治和经济上改善传统中受压迫群体的地位，问题就来

了：在这么做的过程中，该群体得到了以前从来没有的自由，一旦有人想要夺走这些自由，那政府就注定要付出惨痛的代价。

这个道理不仅适用于国家政治，家庭也是一样。父母随随便便地许诺权利、设定规矩，有可能在无意之间给了孩子一些自由，之后再想夺走这些自由的话，孩子们必然会反抗不休。有些家长只在想起来的时候才禁止孩子正餐后吃零食，于是孩子觉得吃些零食是理所当然的。这时，再想定规矩不让他们吃，事情可就棘手、麻烦多了，因为孩子并非少了一种从来没享受过的权利，而是丧失了一种既得的权利。正如我们在政治自由以及（尤其切合当前讨论的）巧克力饼干的例子中所见，跟一贯的稀缺比起来，一样本来有后来没有了的东西，会叫人更想要。既然如此，以下研究结果也就没什么好奇怪的了：管教前后不一的父母，最容易教出反叛心强的孩子。

研究表明，父母在许诺自由时前后不一，最容易让孩子行为叛逆。要避免此问题，家长也无须过分严格、面对规矩毫不通融。举例来说，如果孩子老是错过午餐，那么晚餐之前可以来上一份饭前甜点，因为这并不违反孩子禁吃零食的正常规矩，也就不会确立起孩子吃零食的自由。可要是有些日子准许孩子享受特权，有些日子又不让他这么干，而且还给不出做法前后不一的理由，那事情就难办了。这种武断的做法最容易让孩子觉得先享有而后又失去了某种自由，他们当然会反抗。

READER'S REPORT | 读者报告

致西奥迪尼先生：

我最近在《华尔街日报》上读到了一则报道，很好地阐释了稀缺原则以及人们对得不到的东西的渴望。文章里说，宝洁公司在纽约北部做了一个实验，取消了各类产品的优惠券，直接降低了产品的价格。此举招来了消费者的

极大反感：他们抵制，抗议，写了无数的投诉。尽管宝洁公司的数据表明，发出去的优惠券只有**2%**得到了使用；而且在取消优惠券期间，消费者为宝洁产品支付的价格跟之前大体是相当的。更何况，无须使用优惠券，对消费者来说其实更方便快捷。按文章的说法，消费者反感是因为宝洁公司没有意识到一点：“对很多人来说，优惠券其实是一项不可剥夺的权利。”令人吃惊的地方是，一旦你想把东西拿走，哪怕人们从来不用它，人们的反应也会异常强烈。

纽约的这位投资经理：

虽然宝洁公司的负责人或许为消费者看似非理性的反应感到困惑，但归根结底，公司本身也要负一定的责任。一个多世纪以来，优惠券已经成了美国人生活的一部分，宝洁也主动为自己的产品发放了几十年的“优惠券”。这就难怪消费者觉得优惠券是自己天然的权利，面对长久以来确立的权利，人们总是会为了保护它而激烈抗争。

竞争稀缺资源：狂热愚行

让我们回到饼干实验，从另一个角度来探讨我们对稀缺的反应。从实验的结果当中我们已经看到，较之供应充足的饼干，稀缺的饼干会得到较高的评价；从充足转为稀缺的饼干，所得评价更高。那么，哪些饼干得到了最高的评价呢？就是那些由于需求造成供应短缺的饼干。

请记住，实验里，新出现的稀缺是如何出现的。研究者先给被试装着10块饼干的罐子，然后再换成装有2块饼干的罐子。而这种稀缺，研究者是通过以下两种方法之一创造的。他们告诉一部分被试，饼干还要

拿给其他评分者，以便确保本次研究中对饼干的需求。而另一部分被试听到的说法是，饼干数量必须减少，因为研究者犯了个错误，最开始拿来的罐子不对。结果表明，前一种被试对饼干的喜爱程度明显高于后一种被试。事实上，因为社会需求导致稀缺的饼干，在整轮研究中最招人喜欢。

这一发现凸显了在追求有限资源时竞争的重要性。我们不仅在物品稀缺时想要它，而且，要是碰上有人竞争，那我们就更想要了。广告商经常利用我们的这种心理。在广告里，我们得知某样东西的“大众需求”极高，必须“赶紧来买”；我们看到商店开始营业之前，就有人群乌压压地围在门口；我们看到无数双手伸向货架，把产品一扫而空。这类图像不仅是想让社会认同原则发挥作用，它传递的信息是，其他人都觉得这件产品很好，所以它肯定很好；除此以外，我们还要跟其他人竞争，才能得到这件产品。

参与竞争稀缺资源的感觉，有着强大的刺激性。对情人的态度本来不咸不淡，可听说有对手出现，那便会立刻热情四射。因此，恋爱中的男女常用的一个小手段，就是透露或编造自己有了新的爱慕者。销售人员也会对举棋不定的客户玩弄同样的手法。比方说，要是地产经纪想把房子卖给一个犹豫不决的潜在客户，他可能会告诉客户说，有新的买家来看了房子，很是喜欢，还打算第二天来商谈细节条款。这个新买家当然是完全捏造出来的，销售人员一般会说他们是很有钱的外地人，为了避税的外州投资客，或者刚搬到镇上的一位医生和他妻子。有些圈子把这一手法叫作“赶鸭子上架”，效果好到出奇。因为不想败给对手，好些迟迟不出手的买家立刻就会果断地拍板（见图6-4）。

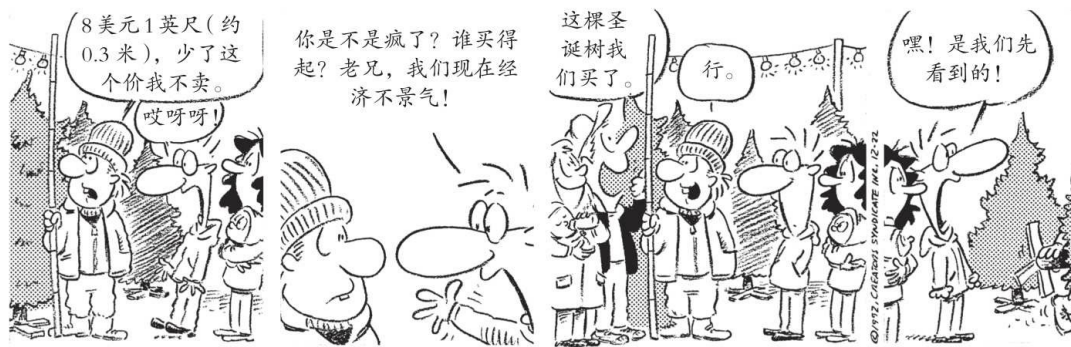


图6-4 有了竞争就来劲

从这幅漫画中可以清楚地看出，争夺有限资源的竞争才不会因为假期而停止呢。

渴望拥有一件众人竞争的东西，几乎是本能的身体反应。按购物者的说法，碰上商场结束营业前的大规模抛售或大降价特卖会，他们简直情不自禁地就给卷进去了。在热火朝天的竞争氛围下，人们一窝蜂地拼抢平时根本不想买的商品。这类行为让我想起野生动物的“狂喂滥吃”现象。商业捕鱼者喜欢利用这种现象。他们先冲着鱼群撒下大量的鱼饵，用不了多久，水里就满是扑腾着鱼鳍、大张着嘴巴夺食的家伙们了。等它们陷入了疯癫的状态，连光秃秃的金属鱼钩也往肚里吞的时候，渔民们便把空鱼钩投入水中，顺顺当当地钓起鱼来，既省时间又省钱。

蓄意在想钓上钩的对象当中掀起狂热的竞争，商业捕鱼者和百货公司的做法颇有异曲同工之处。为了吸引鱼群，渔民撒下切成小块的松散鱼饵。出于类似的目的，百货公司举办大减价特卖会，抛出一些价格特别优惠的商品，大肆宣传，招揽顾客。倘若这两种形式的诱饵发挥了作用，一大群饥渴的鱼或顾客很快就会聚集起来。受现场你争我夺的氛围影响，鱼群和人群会变得急躁不安，盲目、迫切地想要争抢。不管是人还是鱼，此时都会忘了自己想要的是什么，只要别人在抢的东西，自己也要冲上去争（见图6-5）。到了最后，金枪鱼嘴巴里含着没有鱼饵的空钩子躺在甲板上挣扎，购物者提着大包小包的打折货回到家，此时，他们脑袋里说不定有着同样的困惑：我到底是中了什么邪？



图6-5 竞争会传染

这位闷闷不乐的员工刚经历了一场运动鞋清仓销售，据报道，顾客们“为争夺有时连尺码都没看到的鞋子而发狂、扭打和拼抢”。

别以为竞争有限资源的狂热症只出现在金枪鱼或贪图小便宜的购物者这类头脑简单的生命体上。来看看发生在美国广播公司副总裁、后来又执掌了派拉蒙电影公司和福克斯电视网的巴里·迪勒（Barry Diller）身上的一个小故事吧。

影响力研究

1973年，迪勒在美国广播公司（ABC）负责黄金时段的节目时做了一个非同凡响的采购决策。他答应为电影《海神号历险记》支付330万美元的一次性电视播放费。这个数字大大超过了有史以来一次性电视播放费的最高纪录，即《巴顿将军》的200万美元。事实上，这笔钱高得太过离谱，美国广播公司估计要亏100万美元。时任美国全国广播公司（NBC）特别节目副总裁的比尔·斯托克（Bill Storke）当时

就断言：“他们绝无可能赚回本钱，想都别想。”

一个像迪勒这样经验丰富又老到的商人，怎么会做一笔铁定会亏掉100万美元的生意呢？答案恐怕藏在这笔交易中另一个值得注意的特点上。这是电影首次以公开拍卖的方式把播映权卖给电视广播公司，三大电视网从来没有这样被迫地争夺过一项稀缺资源。

想出拍卖这个新颖点子的，是《海神号历险记》浮夸的制片人欧文·艾伦（Irwin Allen）和20世纪福克斯公司的副总裁威廉·谢尔夫（William Self）。如此出人意料的结局，铁定令他们大喜过望。但我们何以肯定，是拍卖的壮观形式而不是电影本身的质量带来了这惊人的成交价格呢？

一些拍卖参与者的言论为我们提供了有力的证据。首先夺下拍卖的迪勒为自家电视台定下规矩。他似乎是有点咬牙切齿地说：“美国广播公司将来决不再参与播映权的拍卖活动。”迪勒的对手，时任哥伦比亚广播公司（CBS）电视台总裁的罗伯特·伍德的评论更发人深省。在拍卖中，他几乎失去理智，开出了比美国广播公司和美国全国广播公司还要高的价钱：

开始的时候，我们都很理性。我们先根据电影能带给我们的利润定好了价格，再在这上面留出了一点回旋的余地。

之后拍卖开始了。美国广播公司开出了200万美元，我还以240万美元；美国广播公司提高到280万美元，这时候我们的脑袋都发起热来，我就像是失去了理智一样，不断开出高价投标；最后，我把价格抬到了320万美元。这时候，我内心冒出了一个声音：“老天爷！万一是我中了标，我该怎么办呀？”还好，美国广播公司的出价最终高过了我，我这才松了一口气。

这件事很有教育意义。

据采访记者鲍勃·麦肯齐（Bob MacKenzie）的观察，伍德在说“这件事很有教育意义”的时候，脸上挂着欣然的笑容。我们可以保证，美国广播公司的迪勒在发誓“决不再参与播映权的拍卖活动”的时候，肯定是笑不出来的。两人显然都从“伟大的海神号拍卖”事件中吸取了教训。但之所以一个人能笑出来，另一个人笑不出来，是因为后者付出了100万美元的代价。幸运的是，我们也能从这里吸取宝贵却不昂贵的一点教训。请注意，笑到最后的人居然是那个从竞争中败下阵来的家伙。作为一条一般性的规律，倘若尘埃落定之后，输家看起来像是赢家，说起话来也像是赢家，我们就该对掀起尘埃的条件多留个心眼。在本例中，也就是对稀缺资源的公开竞争留心眼。正如电视台的高级管理者们所学到的，在碰到稀缺资源加竞争的魔鬼组合时，务必要小心谨慎。

相比数量有限的条件，广告商在信息中设定的时间有限的限制条件出现的频率更高，有研究考察了13 594则报纸广告，发现时间有限的出现次数是数量有限的3倍。然而，又有研究暗示，如果可以选择的话，使用数量有限作为限制条件的效果会更好，因为只有数量有限的安排，才包含了人际竞争的因素并很可能导致疯狂竞争。

“独特性”的妙用

由于人们都重视稀缺资源，所以我们往往更乐意成为他人眼里拥有与众不同特点的人，有些时候尤其如此，比如处于恋爱心境下时。倘若我们有可能跟他人发展出一段浪漫关系，那我们更希望自己与众不同（例如展示自己有着更大的创造力），以吸引对方兴趣。在此种心境下，我们甚至更愿意去那些能让我们脱颖而出的地方。我曾和其他研究者一起设计了一则广告，以敦促人们参观旧金山艺术博物馆。广告里包含了博物馆的名字和一张照片。如果广告受众刚刚看了一段浪漫电影的

剪辑片段，那么“让你从人群中脱颖而出”的措辞会使他们参观博物馆的意向大幅提升。而如果他们之前并未看到浪漫电影的片段，参观（“让你从人群中脱颖而出”的）博物馆的想法也就没有那么富有吸引力。

在有谈恋爱机会的情况下，人们会努力将自己和他人区分开来，这一观点已经在动物研究和人类研究中得到了验证。在后一种研究中，大学生置身恋爱心境时，会明显表现出更强烈的创造性。这种对人类的影响，并不仅限于大学生。例如，毕加索每一个高产的艺术创作时期（蓝色时期、立体派时期和超现实主义时期），背后都有着一样不变的事情——“每一个新时期都绽放着与一位新女性相关的作品，这位女性不是模特而是情人，她们每一位都被吹捧为毕加索炽热（虽然只是暂时）的缪斯”。

另一种让我们感到强烈需要表达独特性的场合，和品位有关。为了保持正确，我们通常会改变自己的信念和观点，以求与他人相符。但在诸如服装、发型、香味、食物、音乐等品位方面的问题上，出于对独特性的追求，我们往往会产生一种相反的动机，即跟群体保持距离。不过，即便在品位方面，群体（尤其是来自一个小圈子）压力也有可能极大。一项研究考察了此类群体成员会怎样做，来平衡“保持一致”与“展现个性”这两种欲望。如果小圈子里大部分人都喜欢某一品牌的物品，那我们很可能也会这么做，但同时又会通过一个可见的维度（如物品的颜色）来区分自己。在确保团队成员遵从核心工作目标的同时，领导者最好将这种对独特性的渴望考虑在内，让成员不必用完全相同的方式去实现核心工作目标。

通过努力赢得的荣誉象征，是另一种能提升个性的因素。我们来看这样一个例子：如果领导者出于善意来消除某个群体的独特性，会发生些什么。

影响力研究

2001年6月14日，几乎所有的美国士兵都将自己的标准野战头盔换成了黑色贝雷帽，而在以前，只有美国陆军游骑兵（这是一支经过特殊训练的精锐作战部队）才会佩戴黑色贝雷帽。原来，为了提升军队士气，陆军参谋长埃里克·新关（Eric Shinseki）下令进行这一调整，以团结军队，成为“军人卓越性的象征”。除了得到一顶黑色贝雷帽，没有任何证据证明此举在成千上万名士兵中发挥了类似作用。相反，它激起了现役和退役游骑兵们的不满，因为他们感觉贝雷帽所代表的独特荣誉遭到了剥夺。游骑兵米歇尔·海尔（Michelle Hyer）中尉称：“这是一场闹剧。黑色贝雷帽是游骑兵和其他特种作战人员付出辛苦努力以示自己与他人区别的东西。现在……戴贝雷帽没有任何意义了。”

新关的命令存在两方面的误导，而这两方面，也都说明了独特性标记是怎样起到作用的。与黑色贝雷帽挂钩的自豪感，来自它的独特性。消除它的独特性，对那些数万名收到黑色贝雷帽的人的自尊，几乎没有影响。而对那些通过努力才赢得了贝雷帽特殊意义的人来说，独特性的丧失深深地刺痛了他们，也招来了他们猛烈的批评。新关将军是怎样解决这个问题的呢？他不能简单地撤回命令，因为他已经一次次地当众强调贝雷帽对全军团结和部队精神的价值。再说，被迫收回说过的话对将军而言也不是什么光彩的事。

新关的解决办法很有启发意义。他允许游骑兵在黑色之外重新选择一种颜色，以彰显其精英部队的成员身份。游骑兵选择了鹿皮古铜色，此后，这种色调就由他们的贝雷帽所独有，而且他们至今仍骄傲地戴在头上。干得漂亮！如新关所愿，他把黑色贝雷帽授予了绝大多数士兵，他们喜欢这种讨人喜欢的新风格；此外，游骑兵也在这种大范围的变化中保留了自己的独特性。双倍的漂亮！

新关将军这么做的理由，以及由此引发的问题和他的解决办法，最出色的报道来自美国军队的官方报纸。

如何防范

面对稀缺压力产生恰当的警觉还算容易，但根据警觉采取行动就难得多了。一部分原因在于，我们对稀缺的典型反应阻碍了我们的思考能力。一看到想要的东西就要得不到了，我们的身体就会亢奋起来。尤其是在涉及直接竞争的环境下，我们更是会血脉偾张、眼光短浅、情绪激昂。这种内在的冲动一旦冒出来，我们知性、理性的一面就会后退。处在此种亢奋状态，人们很难平静下来思量对策。哥伦比亚广播公司总裁伍德在自己的“海神号历险”之后评论说：“你完全陷入了狂热，你的逻辑早就飞到了爪哇国。”

既然这是我们的困境，那么了解稀缺压力的成因和运作原理，或许还不足以保护我们，因为了解是一种认知行为，而认知过程是抵挡不了我们面对稀缺压力产生的情绪性反应的。事实上，这恐怕就是稀缺战术卓有成效的原因所在。只要稀缺战术运用得当，我们防御蠢行的第一道防线，即对形势加以深入分析，就会溃不成军。

倘若因为糊涂虫钻进了脑瓜子，我们无法依靠对稀缺原则的知识来激发妥当谨慎的行为，又该怎么做呢？不妨还是用四两拨千斤的柔道手法，把情绪高涨本身当成重要线索吧！这样一来，我们便可以让敌人的力量为己所用。不靠对整个形势做深思熟虑的认知分析，而是倾听来自内心的警告信号。一旦在顺从环境下体验到高涨的情绪，我们就应该提醒自己，说不定有人在玩弄稀缺手法，必须谨慎行事。

不过，就算我们确实利用高涨的情绪信号，提醒自己平静下来，当心陷阱，那接下来又该怎么办呢？还有没有其他信息能帮助我们在碰到稀缺的时候做出明智决定呢？毕竟，光是意识到自己应该谨慎行事，并

不能说明下一步该往哪个方向走，它只是提供了一个必要的背景，让你多想想再拿主意。

好在的确有这么一种信息，能让我们据此对稀缺物品做出合理决定，它仍然是从饼干研究里得来的。研究者在实验过程中发现了一件看似奇怪却千真万确的事情：尽管人们明显更想要稀缺的饼干，但并不认为它比供应充足的饼干味更美。因此，哪怕渴望程度会随着稀缺而提高（评分人表示，他们将来想要更多稀缺的饼干，还愿意多出钱），它也不能让饼干变得更好吃。

这里隐藏着一点重要的洞见：喜悦并非来自对稀缺商品的体验，而来自对它的占有。我们千万不能把两者混为一谈。每当碰到某种稀缺压力时，我们也会面对一个问题，我们到底想从这样东西里获得什么呢？如果答案是占有这件稀缺的东西能让我们享受来自社会、经济或心理上的好处，那就去占有它吧！稀缺压力能相当准确地衡量我们愿意为它承担的价格，它越是难以得到，对我们来说也就越是宝贵。可更多时候，我们想要一样东西，并不是单纯地想要占有它。我们想要它，只是因为它的实用价值；我们想看它、喝它、摸它、听它、开它或者以各种方式用它。在这样的情况下，我们务必记住：稀缺的东西并不因为难以弄到手，就会变得更好吃、更好听、更好开、更好用。

尽管这一点非常简单，但在经历对稀缺物品强烈的渴望感时，我们常常会忽视它。我举个自己家里人的例子吧！我哥哥理查德就曾靠着大多数人的此种倾向，利用顺从技巧挣得了学费。确切地说，他的策略管用得出奇，只需每个周末干上几个小时就能挣够钱，其他时间都可以拿来学习。

影响力研究

理查德的工作是卖车，但并不是在经销处或者专营摊点上卖。他会按报纸上的消息，在这个周末私人买进几辆二手

车，再把车子用肥皂和水洗干净，下个周末在报纸上打出广告卖掉，从中赚取可观的差价。要做这桩买卖，他得了解三件事情：

- 他必须对汽车有足够的知识，以便按蓝皮书上的价格底线买进车辆，同时在合法的范围内以较高的价格卖出；
- 一旦买下了车，他必须知道如何编写一则能刺激潜在买家兴趣的报纸广告；
- 等买家到了，他必须知道如何利用稀缺原则，刺激对方对汽车产生物超所值的欲望。

这三件事，理查德都很精通。不过，就本处讨论的目的而言，我们只需来看看他在第三点上的高超技艺。

他会为自己前一个星期买下的车在星期天的报纸上登一则广告。因为广告写得特别高明，通常星期天一早，他就能收到一大堆潜在买家打来的电话。他会安排一个时间，让每个想看车的人上门来，事实上，他让所有人都在同一个时间来。这也就是说，如果有6个人想看车，他会让这6个人全在当天下午两点来。

这种在时间安排上的小小技巧，为之后的顺从铺平了道路，因为它创造了一种竞争有限资源的氛围。

一般而言，头一个到的潜在客户会对车况做一番详细的检查，展开标准的买车程序，比如指出车有哪些缺陷或不足，问价钱还有没有商量的余地。可等第二个买家一来，整个环境的心理氛围马上就变了。因为对方的存在，第一和第二个买家把车买到手的可能性突然受到了限制。

很多时候，较早来的那个会在无意间煽动竞争意识，宣称自己有先到先得的权利。“只要一分钟就好，我先来的。”如果他不这么说，理查德也会帮他说。他会对着第二个买家解释说：“真不好意思，但这位先生先来的。所以，麻烦你到对面去稍等几分钟，等他看完车再过来好吗？如果他决定不要，或是暂时拿不定主意，我就请你看车。”

理查德说，有时候，你能亲眼见到第一个买家脸上现出激动的神情。他本来正悠闲地评估着车子的好坏，突然之间，这车就成了一项有人竞争的资源，他得分秒必争地做出决定。如果他不按理查德的要价把车买下来，再过几分钟，他就有可能败在那个在一旁窥视的新到者手里。竞争和资源有限的组合，也同样搞得第二个买家十分激动。他会在路旁来回踱步，明显很紧张地想得到这坨突然变得紧俏起来的金属疙瘩。要是一号买家不买车，或是不能当场拿定主意，二号买家便会立刻扑上前来。

倘若说光有这些条件还不够让买家立刻拿定主意，等三号买家赶到的时候，理查德的圈套就彻底收紧了。据他说，此时累积起来的竞争压力，往往会让最先来的买家没法承受。他要么答应理查德的要价，要么立刻走人，以此结束压力。如果碰到后一种情况，二号买家会为前一位买家没把车买走而长舒一口气，但同时也感觉到了在一旁窥视的三号买家带来的压力。

所有这些为我哥哥的大学学费做了贡献的买家，都没有意识到他们下决心买车的一点基本事实，那就是刺激他们出手买车的强烈欲望，跟车子自身的价值毫无关系。他们之所以没意识到这一点，有两个原因：

- 理查德设计的买车环境让他们产生了情绪反应，买家的脑袋不清醒了。
- 出于前一个原因，他们从来没停下来想过自己买车是为了使用，而不光是拥有它。理查德利用稀缺资源营造的竞争

压力，只影响了他们对车的占有欲。但从买家买车的真正目的来看，这种压力并不会影响到车子的实际价值。

READER'S REPORT | 读者报告

致西奥迪尼先生：

几个星期前，我被你提到的技巧给坑了。这令我很吃惊，因为我并不是那种容易顺从的人，而且我又刚刚读了《影响力》，对那些策略相当敏感。

超市里有个小型试饮活动，一个漂亮女孩递给我一杯饮料。我尝了一下，口感的确不错。然后她问我喜不喜欢，我做了肯定回答，于是她建议我买**4**罐（一致性原理：既然喜欢，我就应该买；互惠原则：她先免费给了我东西）。我可没那么天真，我拒绝了她的请求。可这位售货员并不放弃，她说：“那么只买**1**罐如何呢？”（用的是拒绝-后撤策略）我还是不松口。

漂亮女孩说，这种饮料是从巴西进口的，不知道以后超市里还能不能买到。这时，稀缺原则发挥作用了，我买了**1**罐。等回家再喝，我觉得味道还好，但也就是一般而已。幸运的是，大多数销售员可真的不会这么耐心又执着。

波特兰的这位女士：

尽管知道是稀缺原则，但你还是买了一件自己并不想要的东西。这是不是很有趣呢？最佳的防御做法是提醒自己：这跟饼干一样，饮料稀缺也不会变得更好喝。而且，事实上

它也并不好喝。

如果我们发现自己在顺从环境中受到了稀缺压力的包围，那么，我们最好是采用一套两步应对法。

第一步，一旦我们觉得自己在短缺影响下产生了高度的情绪波动，我们就应该把这种波动当成暂停的信号。要做出明智的决定，恐慌、狂热的反应是不合适的。我们需要冷静下来，重拾理性的眼光。只要做到了这一点，我们就可以转入第二步，问问自己，为什么我们想要那件东西。如果答案是我们想要它主要是因为想拥有它，那么我们应当利用它的稀缺性来判断该为它出多少钱。可倘若答案是我们想要它主要是为了它的功能，即想要驾驶它、喝它或吃它，那么我们必须牢记一点，该物品不管是稀缺还是充足，其功能都是一样的。简而言之，稀缺的饼干并没有变得更好吃。

值得一提的是，对伴随限制而来的情绪激发和焦点缩小，相关证据足以令人信服。同样，对产品进行欺骗性限制的营销手法（即“人为稀缺”），往往是不会对外公开的，但棉花糖米酥（Rice Krispies Treats）却选择在广告中（广告中的售货员故意将成堆的棉花糖米酥藏起来，只在柜台上放一个，以造成产品紧俏的假象）公开此类手法之一，以证明其产品的价值。

本章小结

- 按照稀缺原则，人们认为难以获得的机会价值更高。从“数量有限”和“时间有限”等顺从手法中，可以看到利用这一原理牟利的方式。此时，说服专业人士试图让我们相信，如果我们现在不采取行动，就会损失某种有价值的东西。这调动了人类的损失厌恶倾向，即想到要损失某物比想到

获得同等价值的某物的激励作用更强。

- 稀缺原则的成立有两个原因。首先，由于难于获得的东西通常更有价值，因此一样东西或一种体验的稀缺性，便可以视为判断其质量的捷径。其次，随着东西越来越难于获取，我们丧失了自由。根据逆反心理理论，面对自由的丧失，我们的反应就是比从前更想要得到它们（以及与自由相关的商品和服务）。
- 在一生中的绝大部分时间里，人们都存在逆反心理这种行为动机。但有两个年龄段逆反心理表现得最为明显：“可怕的两岁”和青春期。这两个年龄段都以个性意识觉醒为特点，控制、权利和自由等问题显得极为突出。因此，处在这两个年龄段的人，对限制特别敏感。
- 除了对商品价值的影响，稀缺原则也适用于对信息的评估。研究表明，限制对某一信息的获取，会让人更想得到它或对它产生更多好感。以审查为例，哪怕是人们根本得不到信息，也要受到它的影响。反过来说，倘若人们感觉得到的信息中包含独家（稀缺）内容，它的说服力会更强。
- 稀缺原则最适用于以下两种条件。其一，倘若物品是新近才变得稀缺的，它的感知价值会更高。也就是说，我们觉得新近受限的东西比素来受限的东西更宝贵。其二，当有人跟我们竞争时，稀缺资源对我们的吸引力最大。
- 人们很难用认知来强化自己免受稀缺压力的影响，因为这种影响能使人情绪高涨，让人无法思考。要想抵御它，我们或许可以试着在稀缺环境中对高昂的情绪保持警惕。一旦得到警告，我们可以采取步骤平息高涨的情绪，从“为什么会想要它”这个角度来评估机遇的价值。

Intelligence

第 7 章

承诺与一致

今日之我，确立自昨日或更久之前。

——詹姆斯·乔伊斯（James Joyce）

NEW AND EXPANDED

亚马逊公司每年在全球最富有和绩效最佳公司类的排行榜上都名列前茅。然而，它每年都会向自己物流中心的员工（正是这些员工，帮助亚马逊取得了这样的成就）提供高达5000美元的奖金，鼓励他们离职。这种员工离职可获现金奖励的做法，让许多观察人士感到困惑，因为员工离职后公司付出的成本非常高。包括招聘、雇用和培训替代人员等在内的与员工离职直接相关的费用可占到员工年薪的50%。此外，如果考虑到制度记忆流失（离职的老员工往往很熟悉公司的制度和流程）、生产中断、团队剩余成员士气低落等间接成本，代价甚至会进一步提高。

亚马逊如何从商业角度解释“付费离职”项目的合理性呢？公关发言人梅勒妮·埃切斯（Melanie Etches）对此说得很清楚：“我们只希望愿意留在亚马逊的人留在这里工作。长远来看，待在一个你不愿意待的地方，对我们的员工和公司都不健康。”亚马逊认为，为不开心、不满意或士气低落的员工提供一条诱人的退出道路，其实能节省资金，因为这些员工实际上医疗成本更高、生产力也更低。这一逻辑没有什么可怀疑的。但我认为，这大概不是亚马逊推出这一项目的唯一理由。还有另外一个重要原因——员工留下即意味着向公司做出了认同公司的承诺。关于这个原因的效力，我是从行为科学研究的结果中得知的。此外，我还亲眼见到（并且至今仍能看到）它在我身边有力地运作着。

举个例子，让我们来看看我的邻居莎拉和她同居男友蒂姆之间发生的故事吧。

影响力研究

莎拉和蒂姆相遇后约会了一段时间，最终搬到了一起。在此期间，蒂姆丢了工作。对莎拉而言，事情进展得一点儿也不顺畅：她希望蒂姆跟她结婚，希望他戒掉酒瘾，可这两

件事蒂姆都不同意。两人冲突了好一阵子，之后，莎拉结束了这段关系，蒂姆也搬走了。这时，莎拉的前任男友给她打来电话。他们又开始约会，并很快就订了婚，做了结婚的计划。等具体的日子都订好了，请帖也都发出去了，蒂姆却又打来电话。他后悔了，想跟莎拉和好。莎拉告诉他自己马上就要结婚了，蒂姆恳求再给他一次机会，他希望两人再续前缘。莎拉拒绝了，说她不想再那样生活。蒂姆甚至提出要跟她结婚，可莎拉说，那她倒情愿跟前男友结婚。最后，蒂姆主动说，只要她松口，他就戒酒。莎拉觉得这样挺好，而且蒂姆也有他的优点，于是她取消了婚礼，又跟蒂姆重归于好。

可还不到一个月的时间，蒂姆就告诉莎拉，他觉得自己没有必要戒酒。又过了些日子，他决定结婚的事儿也得“再等等看”。一眨眼，两年过去了，蒂姆和莎拉的生活还是老样子。蒂姆仍然酗酒，一点儿结婚的打算也没有，可莎拉对他却比从前更投入了。她说，正因为有过在前男友和蒂姆之间做选择的经历，她才知道蒂姆在自己心里其实排第一。所以，尽管当初蒂姆所做的承诺从未兑现，莎拉却还是觉得更幸福、更快乐了。

请注意，支撑莎拉承诺的东西，来自她做了一个艰难的个人选择（她在前男友和蒂姆之中选择了后者）。我相信，这也是亚马逊希望员工为公司做出一个同类的选择的原因。面对辞职奖金时选择留下还是离开，不仅是为了高效地识别出那些投入不足的员工，让他们把自己筛选出局，还有助于让那些主动选择留下的人变得更加忠诚。

为什么说后一种结果，也是“付费离职”项目的目的之一呢？我们怎么会这么肯定呢？不要只听公司公关发言人埃切斯女士对此事的说法，

我们不妨把注意力放在亚马逊首席执行官杰夫·贝索斯的说法上。这是一个靠着商业头脑让自己成为世界首富的人。在一封致股东的公开信里，贝索斯先生指出，该项目的目的无非是鼓励员工“花些时间，想想自己真正想要的是什么”。他还提到，年度提案备忘录关于此项目的标题是，“请不要接受这项提议”。也就是说，贝索斯希望员工在考虑辞职时，主动选择不这么做，而事实情况也正是如此：很少有人接受离职奖金。在我看来，这个项目的主要目的正是让员工主动做出留下来的决定，原因也很充分：员工的承诺与其生产力高度相关。

贝索斯对人类心理的敏锐认识，在一系列研究中得到了证实。这些研究表明，一旦做出了艰难的选择，人们总会很乐意地相信自己选对了。两位加拿大心理学家完成的一项研究，揭示了赛马场上人们的奇妙心理。赛马时，只要一下注，人们对自己所选之马获胜的信心就会立时大增。当然，这些马的实际获胜概率并没有发生任何变化，马还是原先那匹马，站在跟原先相同的赛道上，赛马场也还是原先那个赛马场。只不过，在下注者看来，一买下彩票，这匹马获胜的可能性就顿时变大了。类似地，在政治领域，投完票之后，选民们总是更加强烈地相信自己支持的候选人会获得胜利。在其他领域，如果人们主动当众做出了节约能源或水资源的决定，那么，他们会更专注于节约的理念，并想出更多理由来支持它，付出更多的努力以实现它。

一般而言，这种选择方向上的转变，跟另一种常见的社会影响力原则相关。和其他影响力原则一样，这种“武器”也深深地扎根在我们心中，无声无息地指引着我们的行动。这其实很简单：人人都有一种言行一致，同时也显得言行一致的愿望。一旦我们做出了一个选择，或采取了某种立场，我们立刻就会被来自内心和周围的压力，迫使自己按照承诺去行事。而且，在这样的压力之下，我们还会想方设法地以行动证明自己先前的决定是正确的。[\(42\)](#)

言出必行

心理学家早就认识到承诺与一致原则对人的行为有着强大的指引力量。许多杰出的早期理论家都把言行一致的欲望看成行为的一种重要驱动力。这种力求一致的观念，真的强大到能迫使我们做正常情况下不想做的事情吗？确实是这样。保持并显得前后一致的动力，是一种威力巨大的社会影响力武器，它经常令我们做出明显有违自己最佳利益的行为。

来看看下面这个实验吧！研究者在纽约市的一处沙滩导演了一起“偷窃”事件，观察旁观者是否会不顾个人安危来阻止犯罪。

影响力研究

研究者的助手在沙滩上随机选一个人，即实验的被试，在离他两米开外的地方铺上一块沙滩浴巾。助手躺在浴巾上，用便携收音机听了一小会儿音乐，然后站起身，离开浴巾到海滩上去散步。过了一会儿，研究者会假装成小偷，走过去拿起收音机，试着把它带走。你可能已经猜到了，正常情况下，由于自己可能会受到伤害，被试是不愿冒险去阻止小偷的。“偷窃”事件上演了20回，旁观者出手阻止却只有4回。但只要稍加调整，同样的过程再来上20回，结果却与之前大为不同。在后面的20回中，助手在离开浴巾之前要请被试“帮忙看着我的东西”，所有被试都答应了。这下，在承诺与一致原则的推动下，20个被试中有19个成了事实上的义务警员，他们主动阻止了偷窃行为，要求对方给出解释，甚至出手拦住小偷，不让他拎着收音机逃跑。

要想理解为什么人的一致性动机如此强大，我们应当意识到，在大多数环境下，言行一致都是一件既有价值又得当的事。依照人们的普遍感觉，言行不一是一种不可取的人格特征。一方面，信仰、言语和行

为前后不一的人，会被看成是思路混乱、表里不一，甚至精神有毛病；另一方面，言行高度一致大多跟个性坚强、智力出众挂钩，它是逻辑性、稳定性和诚实的核心。伟大的英国化学家迈克尔·法拉第说过一句话，暗示了人们非常看重一致性，有时甚至到了觉得它比做事正确还重要的程度。一次演讲之后，有人问法拉第，他的意思是不是说某个讨厌的学术对手一贯出错，法拉第瞪着提问者回答道：“他才没那么前后如一呢！”

所以，在我们的文化里，一个人高度的言行一致是备受称道的，也理应如此。大多数时候，要是我们在做事时始终如一、坚持不懈，肯定会做得很好。没有了一致性，我们的生活便会困难重重、混乱不堪。

捷径

由于承诺与一致一般来说符合我们的最佳利益，因此我们很容易养成自动保持一致的习惯，哪怕有时候这么做并不明智。不假思索地保持一致，有可能带来严重的后果。不过，就算是盲目地保持一致，也不乏迷人之处。

和大多数其他自动响应方式一样，保持一致为穿越复杂的现代生活提供了一条捷径。只要我们对事情拿定了主意，死脑筋地坚持到底就能给我们带来一种分外难得的好处：我们再不用苦苦地思考这件事了。我们不需要从每天接触的庞杂信息中挑挑拣拣来确定相关事实；我们不必再劳心费神地权衡利弊；我们也犯不着再做出任何棘手的抉择。相反，每当碰到同一类的事情，只需要按下我们的一致性磁带，让它哗啦啦地播放起来，我们就会立刻知道该去信什么、说什么和做什么。不管怎么样，我们的信念、说辞和行为，只要跟之前的决定保持一致就行。

千万别低估这种享受的吸引力。要知道，日常生活的纷繁复杂对我们的精力和能力都提出了苛刻的要求，可有了一致性我们就能以相对轻

松、高效的便利方式来应对这一切了。这么一来，人们很难克制保持一致的下意识反应，也就不难理解了：它让我们有了逃避连续思考这桩苦差事的捷径。一旦保持一致的磁带播放起来，我们就可以开开心心地去做事，不用想太多。正如乔舒亚·雷诺兹爵士（Joshua Reynolds）所说：“要是有什么办法能省掉动脑筋这档子真正的体力活儿，那人们断然不会放过它。”

愚昧的城堡

机械地保持一致还有另外一点迷人之处。有时候，我们逃避思考活动，不是因为它辛苦、要动脑筋，而是因为这么做了以后反而会招来严重的后果。有时候，只要稍加思考，就能得出一连串显而不受人待见的该死答案。就因为这个，我们才懒得去思考。所以有些烦人的事情，我们宁肯视而不见。由于自动保持一致是一种预先设置好的响应方式，所以一旦碰到麻烦事，它就会为我们提供一处安全的藏身之所：躲在保持一致的城堡里面，我们总算可以逃过理性带来的折磨了。

一天晚上，我参加了一次介绍超自然冥想的讲座，亲眼看到了人们是怎么藏在一致性的城墙背后，不愿承担思考带来的恼人后果的。

主持讲座的是两个热心的年轻人，他们想招募一些新成员。他们说，自己的协会提供一种独有的冥想术，能让人得偿所愿。获得内心的平静自不必说，等修炼到了高级阶段，甚至能让人掌握超能力，比如腾空飞行，穿越墙壁一类的。

我来参加这个讲座的目的是观察此类招募会里使用到的顺从手法。这天，我带了一个感兴趣的同行朋友，他是数据统计和符号逻辑学的大学教授。随着会议的进行，讲师们解释起他们冥想术背后的理论来，我发现身边的逻辑学家朋友越来越焦躁。他如坐针毡，换了好多姿势，终于再也忍不住了。讲座结束之后，讲师们要大家提问题，他迫不及待地

举起手，轻言细语但态度坚定地一一驳斥了我们刚刚听到的陈述。只用了短短两分钟，他便一针见血地指出讲师的复杂论证在哪些地方是互相矛盾的，以及为什么它们不合逻辑又欠缺证据。这对主讲人可真是不小的打击。他们手足无措地沉默了一阵之后，便开始尝试驳倒我朋友的观点。但他们的开脱苍白无力，而且说到中间还得跟伙伴商量一番。最终，他们无奈承认，我朋友的看法很好，他们“有必要进一步研究”。

不过，在我看来，更有趣的地方在于这对听众的影响。提问时间过后，大量的听众围着两位讲师，竞相掏出75美元，报名参加他们的冥想培训。讲师们一边收钱，一边用胳膊肘碰碰对方，耸耸肩膀，窃窃私笑，显然，他俩也搞不明白这是怎么一回事。毫无疑问，先前的尴尬一幕搞砸了他们的陈述，可不知为什么会议却取得了空前的成功，听众们就像被灌了迷魂汤一般，对他们言听计从。我也一头雾水，以为听众没搞懂我朋友的反驳逻辑。然而，事实证明，情况恰恰相反。

讲座结束后，三名听众找到了我们。他们都在听完讲座后立刻付费报了名。他们问我们为什么来听这个讲座。我们做了解释，同时也向他们提出了同样的问题。三个人中有一个是胸怀大志的演员，很想在表演方面取得成功，他来是想看看冥想术能否帮他实现必要的自我控制，以便他在演技上能更上一层楼。而负责招募的讲师向他保证，这没问题。第二个人说自己患了严重的失眠症，她希望依靠冥想术来放松心情，晚上能轻松入睡。第三个人在做非正式的代言人。他大学有几门课没通过考试，因为好像学习的时间总是不够用。他来听讲座是想了解冥想术能不能帮助他减少每晚的睡眠时间，这样多出来的时间就能用来学习了。讲师的说法当然跟先前告诉失眠者的一样：没问题。看起来，超自然冥想术似乎什么问题都能解决，不管这些问题是不是互相矛盾的。

这时，我仍旧以为这三个人报名是因为没听懂我的这位朋友的观点，于是就拿朋友所说的几点问了问他们。让我吃惊的是，我发现他们完全明白我朋友说的，实际上，是再清楚不过了，而且，正是因为朋友

的论点太有说服力了，才促使他们赶紧在现场报了名。那位代言人说：“我本来不会当场就掏腰包的，因为我现在穷得要命。我原本是打算等到下次听讲座时再说。可你的朋友一开始说话，我就明白了，最好还是现在就把钱给他们，要不然，只要一回家我就会想到你朋友说的话，今后便再也不会报名了。”

我这才明白过来。来听讲座的这些人都碰到了真正的问题，正拼命想办法解决这些问题。要是研讨班讲师的话值得一信，那么，超自然冥想术就不失为一种潜在的解决办法。在需求的驱动下，他们非常想要相信超自然冥想术就是他们的救星。

就在此时，我同事的理性之声传了出来，指出来听讲座的这些人认定的新办法似乎在理论上就不合理。他们惊慌起来！必须赶紧采取行动，要不然，等逻辑占了上风，他们的希望就又破灭了。赶紧，赶紧，筑起对抗理智的高墙来！即使修起来的是座愚昧的城堡，也无关紧要。“赶紧地，找个地方藏起来，再也不动脑筋了！来，拿着我的钱。好啦，这下子可就安全多啦。再也不想这些问题了。”既然决心已定，从现在起，一致性磁带就能在必要的时候播放起来了：“超自然冥想术？它当然能帮到我，我当然想要继续下去，我当然相信超自然冥想术。瞧，我已经投了钱进去了，不是吗？啊，不伤脑筋地保持一致真舒坦啊！我就在这儿休息一会儿好了。艰苦地寻觅太紧张、太焦虑了，现在感觉好多了。”

捉迷藏

要是下意识的保持一致真的是逃避思考的盔甲，那么那些想要我们不假思索便答应他们要求的人必然会利用它们。这没什么好奇怪的。面对他们的要求，倘若我们不假思索地做出机械反应，牟利的人可就有福了：我们下意识的一致性倾向根本就是一座金矿。所以，他们聪明地做了巧妙的安排，让我们一致性的磁带播放起来为他们赚钱，而我们自

己却浑然不知。他们用高明的柔道手段来设计跟我们的交流互动，利用我们保持一致的自身需求赚得盆满钵满。

一些大型玩具制造商就用这种方法来减少季节性购买模式带来的问题。当然了，玩具公司的销售旺季是在圣诞节假期之前。问题在于，接下来的几个月，玩具的销售情况便会陷入可怕的低迷期，消费者已经花光了原本计划买玩具的钱，再也不愿给孩子买更多的玩具了。

于是玩具制造商面临着一个两难境地：如何保住高峰期的销售旺势，同时，也在随后的几个月里维持健康的需求量。刺激孩子过完圣诞节之后还想要更多的玩具对玩具制造商们来说显然没什么可犯难的，可问题在于，怎么才能让过完节又花光了钱的父母给玩具到处都是的孩子再买新的玩具。玩具公司要怎么做才能完成这个近乎不可能完成的任务呢？有些公司试过大幅提高广告宣传力度，有些公司则在低迷期搞降价促销，但从实践来看，这些标准的销售策略都不怎么成功。这两种策略既费钱，又不能把销售量拉到一个理想的水平。家长的确没有买玩具的心情，广告或降价的影响不足以改变他们顽固的死脑筋。

可有些大型玩具制造商却觉得自己找到了解决的办法。这个办法很是巧妙，只需要正常的广告支出，外加理解人们保持一致的强大心理需求就足够了。我最初意识到玩具公司搞的这套操作手法，是在上过一回当之后，可还没等回过神来，我就又活生生地上了当，我可真是个傻瓜蛋。

那是在1月，我来到了城里最大的玩具店。就在一个月之前，我给儿子买了太多太多的礼物，我发誓，未来很长一段时间里，我再也不踏进这类商店了。然而，我不仅在之后不久再次来到了这个“残忍”的地方，还打算给儿子再买一件昂贵的玩具：一套大型电动赛车。在电动赛车的展柜前，我碰巧遇到了一位从前的邻居，他也正要给儿子买同样的玩具。奇怪的是，我们之前很少碰到对方。说起来，我俩上次见面已经是一年半之前了，那次也同样是在这家店，同样是在圣诞节之后，也同

样是在给儿子买一样的贵玩具：一个能走路、能说话，甚至能排便便的机器人。想到我们一年里总是在同一时间、同一地点、做同一件事的时候碰到对方，我们不禁笑了起来。当天晚些时候，我跟一位朋友提到了这一巧合，他以前在玩具行业工作过。

“才不是什么巧合呢。”他一副深知内幕的样子。

“不是巧合？你说的是什么意思？”

“瞧，”他说，“我来问你几个问题，是关于你今年买的那套玩具赛车的。第一，你是不是答应了儿子，圣诞节给他买一套？”

“嗯，是呀。克里斯托弗在星期六早晨的卡通节目里看到了这玩意儿的好多广告，他说圣诞节就要这个。我自己也看过几段广告，看上去挺好玩的，所以我说行。”

“中了一条，”他说，“现在我来问第二个问题。等你去买的时候，是不是发现所有的商店都卖完了？”

“太对了，真是这样！商店说他们已经下了订单，可不知道货什么时候才能到。所以我给克里斯托弗买了其他玩具来代替。可你怎么知道呢？”

“中了两条，”他说，“让我问完最后一个问题。去年你买机器人时是不是也发生过这种事儿？”

“让我想想……你说的没错，就是这样。太奇怪了，你到底是怎么知道的？”

“我可不会什么读心术。只是，我刚好知道几家大玩具公司是怎么拉动一月和二月的销量的。圣诞节前，他们开始在电视上做一些特别玩具的广告。显然，孩子们挺想要的，他们缠着父母答应圣诞节买来送给自己。好了，这些玩具公司的精明

之处就在于：他们故意不给商店提供足够的货品。这下子，大部分父母会发现这些玩具早就卖光了，所以只好被迫买下等值的其他玩具给孩子充数。当然了，对于这些充数的玩具，制造商们的货供应充足。接着，过完了圣诞节，公司又开始为前面那些特别的玩具打广告，这使得孩子们越发想要了。他们跑去跟父母哭诉：‘你答应过的，你答应过的。’于是父母只好无奈地跑去玩具店兑现自己的诺言。”

“我算是明白了，”我气呼呼地说，“这就是为什么家长们总能在玩具店碰到一年多没见的老朋友，因为对方也落入了同一个圈套，对吧？”

“是呀。咦，你要去哪儿？”

“我去把这套赛车给退了。”我火冒三丈地吼道。

“别着急呀。你再考虑一分钟，你今早为什么去买它？”

“因为我不想让克里斯托弗的希望落空啊，还因为我想教育他，人得言出必行。”

“好了，现在有什么与今早不一样的地方吗？听我说，要是你把他的玩具退了，他是搞不懂什么的，他只知道他老爸说了话却做不到。你想要这样的结果吗？”

“不，”我叹了口气说，“我不想。可想想看，你告诉了我，过去两年里玩具公司在我身上赚了双倍的钱，可我却一无所知。好了，现在我明白了，但还是爬不出陷阱，而且还被自己说的话给套住了。照这样说来，我可真是三条全中啊。”

他点点头说：“没错。所以你玩不过那些玩具商呀！”

从那以后，每年的圣诞节假期，类似我经历的那种爹妈忙着买玩具的热闹景象，我观察到了不少：芭比娃娃、挠我痒痒埃尔莫、菲比娃

娃、Xbox游戏机、Wii游戏机等。但回顾历史，最符合上述销售模式的玩具还要数卷心菜娃娃。

影响力研究

20世纪80年代中期的圣诞节期间，制造商使劲为这种25美元一个的娃娃打广告，但给商店的供货却少得可怜。结果，政府以虚假广告为由起诉了生产商，控告他们打广告卖一种买都买不到的玩具。而为了给孩子买到一个娃娃，疯狂的父母们在商店里大打出手，甚至把价格抬到了每个700美元。这家玩具制造商赚疯了，当年的销售总额高达1.5亿美元。甚至在圣诞节之后，卷心菜娃娃的热潮也没有消退。

1998年的圣诞节期间，每个孩子都想要，却最难买到的玩具是绒毛菲比娃娃，它是玩具制造巨头孩之宝的一家分公司设计的。有人问孩之宝的女发言人，垂头丧气的父母们买不到菲比娃娃，该怎么告诉孩子才好。她建议的回答是：“我尽量吧，可要是我现在买不到，过一段时间总会给你弄到手的。”唉，几十年来，玩具商们不正是靠着那么多父母这样的承诺发大财的吗（见图7-1）？[\(43\)](#)



图7-1 不劳无获

漫画里的游戏玩家贾森，成功地利用了索要节日礼物的策略，但我认为他把成功的

原因搞错了。以我自己的经验来看，父母之所以会给贾森买许多其他的礼物作为补偿，不是为了缓解他的“痛苦”，而是为了缓解他们失信于儿子的痛苦。

承诺是关键

一旦我们意识到人类的行动不可避免地要受保持一致的强大力量所指引，那么一个具有实际意义的重要问题就会冒出来：这种力量到底是从哪里来的呢？是什么东西按下了播放键，激活了难以抵挡的一致性磁带呢？社会心理学家认为他们已经找到了答案：承诺。要是我能让你做出承诺，即选择立场、公开表明观点，就相当于帮你搭建好了舞台，促使你不假思索地自动照着先前的承诺去行事。只要立场站稳了，人们就自然地想要倔强地按照与该立场保持一致的方式去行事。

正如我们所见，明白承诺与一致之间联系的可不光只有社会心理学家，各行各业的说服专业人士都会拿承诺策略来对付我们。这些策略都有着这样的目的：诱使我们采取某种行动或做出某种表态，从而通过我们内心保持一致的压力逼我们顺从。诱使我们做出承诺的手法多种多样，有些非常直接，有些则十分微妙。

关于前者，可见阿尔伯克基一家汽车经销商的二手车销售经理杰克·斯坦科（Jack Stanko）的做法。在一次全美汽车经销商协会会议上，他做了一个名为“二手车经营手法”的演讲，对100多位渴望卖车的经销商提出了如下建议：“把约定写在纸上。只要客户说了好，就让他们写下来。先把价钱告诉他们。控制客户，控制交易。要问他们：‘要是价钱合适，是否即刻就买？’抢先把他们钉死，叫他们无法反悔。”显然，斯坦科先生是这方面的行家，他很清楚怎样通过承诺来实现“控制”，进而让客户顺从。

使用多重手腕获得顺从的做法同样有效。比方说，你正为自己最喜欢的慈善组织筹款，想提高本地区答应你上门收钱的居民人数。考虑一下社会心理学家史蒂文·谢尔曼（Steven J. Sherman）采用的方法吧！

影响力研究

谢尔曼给印第安纳州卢布明顿地区抽选出来的居民打电话说，自己正在做一项调查，想知道他们的意愿：要是美国癌症协会需要筹款，他们是否愿意花3个小时帮忙。当然，大家都不愿显得自己缺乏爱心，所以很多人都说他们愿意。在通过如此微妙的手法征得承诺之后，过了几天，美国癌症协会真的打电话来要求社区组织募捐团，结果，帮忙的志愿者比从前多了足足7倍。另有一批研究者也采用了同一策略，请居民预测自己是否会在选举日当天投票，这下子，在接受电话调查的人里面，投票率大幅提高。

律师们似乎也开始采用这种做法，依靠最初的承诺，刺激当事人在未来采取与承诺一致的行为。在为审判筛选合格的陪审员时，这一行里被公认为最佳咨询师的乔-艾伦·德米特里斯（Jo-Ellen Demitrius）问了一个十分巧妙的问题：“要是陪审团里只有你一个人相信我的当事人是无辜的，你能顶住其他陪审员要你改变主意的压力吗？”面对这样的问题，有哪一个自尊自爱的准陪审员会说“不能”？再者，做了这样的公开承诺之后，又有哪一个自尊自爱的陪审员会公然反悔？

电话募捐人员获取承诺的手法就更加狡诈了。不知道你注意过没有，如今他们打电话来募捐时，总会先问问你近况如何，身体好不好。他们会说：“你好，先生 / 女士今晚心情如何？”或者“今天过得如何？”这样子的开场白，可不光是为了显得亲切友善。这样做是要让你像平常听到这类客套话时那样给个礼貌的回答，如“挺好的”，“还不错”，或者“谢谢，还算行吧”。一旦你公开表明“还不错”，募捐员逼你资助那些过得不咋样的人就容易多了：“听您这么说我可真高兴，因为我打电话来是想问问，您愿不愿意捐款帮助某某不幸的受害者……”

这种手法暗含的逻辑是：人们要是刚刚说了自己感觉挺好或者过得不错，哪怕这么说只不过是出于社交时的客套，若之后马上就做出一副小气样，未免会显得很尴尬。倘若你觉得这似乎太过牵强，不妨来看看专门研究消费者行为的研究者丹尼尔·霍华德（Daniel Howard）的发现，他亲自检验了以上逻辑。

影响力研究

丹尼尔打电话给得克萨斯州达拉斯市的居民，问他们是否答应让饥荒救济委员会的代表上门兜售饼干，所得收益将用来给贫困家庭供应伙食。倘若电话工作人员只是提出这一要求，那这便是标准募捐法，只有18%的居民会答应下来。但要是他一开始先问“今晚您感觉如何”，并等对方回答之后再展开标准的募捐流程，这时就会出现好几件引人注意的事情。首先，在120名接到电话的居民里，大多数人（90%）都给了客套回复，如“挺好”“不错”“非常好”等；其次，32%被问了“今晚你感觉如何”的人答应在家接待卖饼干的销售人员，成功率比只采用标准募捐法高了差不多两倍；最后，根据承诺与一致原则，几乎所有答应销售员上门的人（89%）实际上都买了饼干。

还有一个用相对较小的口头承诺就可以产生实质影响的领域是避免出轨。心理学家提醒我们，出轨、欺骗伴侣是恋爱关系中激烈冲突的根源，往往会导致愤怒、痛苦和关系的结束。好在研究者找到了一种有助于预防上述有害后果的活动，那就是祈祷。不过这不是泛泛而言的祈祷，而是一种很特殊的祈祷。如果一个人的恋人答应每天为对方祈求幸福，那么过上一段时间之后，进行祈祷的这个人出轨的可能性就大大降低了。毕竟，不忠行为有悖于每天主动为伴侣的幸福祈祷这一做法。[\(44\)](#)

致西奥迪尼先生：

我从您书中学到的对我最有效的一课就是承诺了。几年前，我为一家电话销售中心培训电话保险推销员。可这里面最大的麻烦在于，我们没办法仅通过电话就切切实实地把保险卖掉，只能报个价，然后指导来电者到离家最近的公司办事处去做具体面谈。结果，答应要去办事处的来电者却根本没露面。

后来，我找了一组新的电话保险推销员，教给他们一种新的方法：只是把其他推销员采用的销售方法调整了一下。他们所用的“罐装”陈述跟别人的完全一样，只不过，在通话结束时，该方法多加了一个额外的问题。在客户确认了预约时间之后，我们指示销售员先别挂断电话，而是问道：“我希望您能告诉我一下，为什么您选择购买我们公司的保险？”

我最初只是想收集客户的服务信息，但这些新推销员却比同期的其他推销员多卖出了**19%**的保险。后来，我们把这个问题整合到所有人的推销陈述里，连熟练的业务员也比从前多完成了**10%**的业务量。没看过您的书之前，我还搞不清楚这是怎么回事呢。

得克萨斯州的这位销售培训师：

虽然是无意中用上的，但你的手法相当高明，因为它不仅让客户对自己的选择做了承诺，还让他们为自己所说的理由做了承诺。再者，诚如我们在第1章中所见，人们往往会

看在理由的面子上去做某件事。

我在亚特兰大的一位熟人，也肯定了这一策略的有效性。他曾遵循标准的建议，充分描述了自己应该获得聘用的所有理由，但在工作面试中仍一无所获。为了改变这样的结果，他开始主动采用一致性原则。他向评估人员保证自己会尽量完整地回答所有问题，接着又说：“不过，在我们开始之前，我想知道您能否回答我一个问题。我很好奇，我的哪些背景让你们把我加入候选名单呢？”这就使得评估人员亲口对他及其资历做出积极的描述，换言之，在他自己还没有陈情之前，招聘方就已经自我承诺要聘用他了。他发誓说，依靠这种手法，他已经连续拿到了三份更好的工作。

自设的牢笼

承诺为什么这么有效呢？原因很多。很多因素都会影响承诺对我们将来行为的限制程度。一个旨在让人顺从的大规模项目阐明了若干因素的运作详情。这个项目最令人啧啧称奇的一点是，它在近半个世纪前就系统化地应用了相关的因素，可在那时，科学研究还根本没有确认到底是哪些因素与此相关。

影响力研究

A、B两国战争期间，许多被俘的A国士兵被关在C国人管理的战俘营里。他们很快发现，C国人对待战俘的方式明显跟B国不同，后者喜欢用严刑拷打来迫使战俘顺从。C国方面则有意回避了这种残忍做法，他们采用的是“宽大政策”，实际上，这也是一种专门设计的复杂心理攻势。

战争结束后，A国心理学家对释放回国的战俘提出了连珠炮般的问题，想要搞清楚到底发生了什么，因为从某些方面来看，C国的战俘政策取得了惊人的成功。比方说，C国人非常有效地让A国战俘互相揭发，彼此监视。出于这个原因，当然也还有其他种种因素，要是有人想逃跑，计划很快就会暴露，逃跑的人几乎没有成功的。心理学家埃德加·沙因（Edgar Schein）是A国负责调查C国战俘改造项目的首席研究员，他写道：“要是真有人逃跑了，只要给告发他的人一袋子大米，C国人就能轻轻松松地把人给找回来。”事实上，据说几乎所有关在C国战俘营里的A国俘虏都曾以这样那样的方式跟C国合作过。

在对C国战俘营做了研究之后，人们发现，C国人大量依靠承诺与一致的压力来让战俘顺从。显然，C国人面临的第一个问题是要想办法让A国人与他合作，任何形式的合作都可以。A国士兵都受过训练，除了自己的姓名、军衔和编号之外，他们什么也不会说。C国人并不施以肉体暴力，那他们到底是怎么让这些A国士兵透露军事情报、告发同房战友，甚至公开谴责自己的国家的呢？答案很简单：以小积大。

C国人经常要战俘做一些态度温和地反对A国、支持和平的陈述，如“A国并不完美”“这里没有失业问题”等。表面上看起来，这些陈述没什么大不了的，可一旦顺从了这些小的要求，战俘们就马上发现，自己要被迫答应内容相关但更具实质性的要求。假设说，C国审讯员要一个A国战俘同意A国并不完美这一说法，战俘认可了；紧接着，审讯员就要他谈一谈，在他看来，A国有哪些地方不完美；等战俘做了解释，审讯员说不定又要他列一张“A国的问题”清单，并签上名字；之后，他们又要他跟其他战俘结成小组，讨论自己的

这张清单：“毕竟，你自己也相信这些问题，对吧？”随后，他们又要他写一篇文章，扩充清单，更详尽地探讨A国存在的问题。

此后，C国或许会在反对A国的广播电台上向整个战俘营，以及B国的所有战俘营，甚至驻扎在邻国的A国军队播报这个人的名字和他写的文章。突然之间，这个战俘就发现自己成了“合作者”，给敌人帮了忙。因为他知道自己写这篇文章并非出于他人的要挟或胁迫，于是这人便会不断调整形象，好让自己的行为符合“合作者”这个新标签，如此又带来了更广泛的合作举动。所以，根据沙因的研究，“只有极少数的人能完全不跟对方合作，绝大多数人都做过一些自己看来没什么大不了的事情，都合作过一两次。C国则把这些事情有效地利用了起来.....在审讯中获取口供、要战俘自我批评、透露情报，这么做尤其管用”。

这里有必要着重指出，在A、C两国交战期间的心理教化项目中，所谓的“合作”并不一定是有意识的。A国调查人员把合作定义为“所有帮助了敌方的行为”，所以它包括多种活动，如签署和平请愿书、跑跑腿、在电台里发出呼吁、接受特别的待遇、做虚假供述、告发同狱战友、泄露军事信息等。

其他对说服感兴趣的群体也知晓这种方法威力巨大又有效。例如，慈善组织常常要人们一步步做出越来越有分量的承诺，以最终让当事人捐甚为可观的数目。研究表明，最初看似琐碎的小承诺，如答应面谈，能打开“顺从动量”的闸门，诱使人们之后做出捐献器官和骨髓的行为来（见图7-2）。



图7-2 从小事做起

猪喜欢泥巴。可它们并不吃泥巴。为此，似乎有必要让承诺逐渐升级。

许多商业组织也经常采用这种方法。对销售人员来说，这就意味着从一笔小生意做起，最终拉到大生意。商人做小生意几乎都不是为了贪图利润，而是为了建立承诺。有了承诺，之后的生意自然而然地就来了，做成大生意也不是什么稀罕事。《美国销售员》杂志上有一篇文章写道：

总体而言，小订单为全面推销铺平了道路……这样来看：有人签了订单，购买了你的商品，尽管利润微薄得不足以弥补你打电话所花的时间和精力，但他已经不再是潜在客户了，他成了你真真正正的客户。

这种从小请求开始、最终要人答应更大请求的手法，叫作“登门槛”。1966年，心理学家乔纳森·弗里德曼（Jonathan Freedman）和斯科特·弗雷泽（Scott Fraser）公布了一批令人吃惊的数据，社会科学家们这才首次意识到这种手法是多么有效。这两位心理学家做了一项实验，派研究者假扮成义工，到加利福尼亚州的一处居民区，当面向业主们提出一个荒谬的要求。

影响力研究

研究者要业主们同意在自己的前院草坪上立一块公益告示牌。为了让业主们明白牌子是什么样子的，他们出示了一张照片，照片上的房子挺漂亮，可房子正面的视线完全被一块硕大的“小心驾驶”的告示牌给挡住了。出于可以理解的原因，该地区的绝大多数业主都拒绝了这个要求（只有17%的人答应下来），但有一组业主的反应却分外积极。这一组中，76%的人都答应把自家的前院贡献出来。

他们答应得如此爽快，主要是跟两个星期前发生的一件事有关：他们对保障驾驶员安全做了个小小的承诺。当时，有一位义工来到他们家，请他们在院前立一块长宽只有七八厘米的警示牌，上面写着“做一个安全的驾驶员”。这个要求实在是太微不足道了，几乎所有人都答应下来，但它给人造成的影响却极为惊人。由于几个星期前毫不知情地答应了一个有关安全驾驶的小小要求，这些业主对另一个分量重得多的要求居然也照单全收。

弗里德曼和弗雷泽并未止步于此。他们重新找了一组业主，尝试了一种稍有不同的程序。

影响力研究

起初，业主们收到了一份请愿书，要他们签名支持“保护加利福尼亚州的美丽环境”。显然，差不多人人都签了名，因为维持一个州的美丽环境，跟提高政府工作效率、进行合理的产前保健一类的议题一样，是不会有人反对的。过了大概两个星期，弗里德曼和弗雷泽派了一名新义工到这些家庭，请居民答应在自家前院草坪上竖一块硕大的“小心驾驶”的告示牌。从某些方面来看，这些业主的反应是本次研

究里最出人意外的：将近一半的人都同意在自家院子里竖立“小心驾驶”的告示牌，尽管几个星期之前，他们做出的小小承诺跟小心驾驶毫无关系，而是关于另一项公共服务议题：保护环境。

起初，连弗里德曼和弗雷泽也被这样的研究结果弄糊涂了。为什么在支持保护本州美丽环境的请愿书上签了名，就可以让人乐意做出全然不同、分量更大的善举呢？经过思索，在排除了其他原因之后，弗里德曼和弗雷泽提出了一种解释：签署保护环境的请愿书，使这些人对自身的看法发生了改变。他们把自己看成了具有公益精神、履行公民职责的好市民。这样一来，等到两周以后，当有人要他们履行另一项公益使命，即竖起“小心驾驶”的告示牌时，为了符合新塑造起来的自我形象，他们都乖乖地答应了下来。弗里德曼和弗雷泽这样说：

发生变化的大概是人们对参与或采取行动的感觉。一旦答应了某个请求，他们的态度就可能改变。在他们看来，自己成了做这种事情的人：答应陌生人提出的请求，对自己承诺的事情采取行动，配合有着高尚动机的善举。

弗里德曼和弗雷泽的发现告诉我们，在接受琐碎请求时务必小心谨慎，因为一旦同意了，它就有可能影响我们的自我认知。它不仅能提高我们对分量更大的类似请求的顺从度，还能使我们更乐意去做一些跟先前答应的小要求毫不相关的事情。正是后面这一种藏在小小承诺里的普遍影响力，叫我甚感惊恐。

它把我吓得都不怎么愿意在请愿书上签名了，哪怕请愿书的立场我原本就支持。因为这类行动不仅可能影响到我将来的形象，还可能让我按照自己并不想要的方式去改变自身的形象。再者，一旦人的自我形象发生了改变，那些想要利用这一新形象的人就有了各种微妙的可乘之

机（见图7-3）。



图7-3 只要签个名就行

你考虑过各类社会团体要你在请愿书上签了名以后，会拿这些签名来干什么吗？大多数时候，他们并不会拿签名派什么具体的用场，因为请愿书的主要目的在于，让我们对该团体的立场做出承诺，以便我们以后会更乐意采取与承诺相一致的进一步举动。

研究表明，哪怕是微不足道的承诺也能导致自我概念的改变。弗里德曼和弗雷泽实验里的业主们有谁想过，那个要他们签署保护本州环境请愿书的义工，其真正的目的竟然是在两个星期以后让他们展示一块安全驾驶的告示牌？他们中又有谁会疑心自己答应展示这块广告牌的决定，居然跟签一份请愿书有着很大的关系？我猜没人会这么想。倘若告示牌竖起来以后他们觉得有点后悔，除了自己，以及除了自己那强烈得活见鬼的公民精神，他们还能怪谁呢？他们恐怕绝不会想到那个拿着“保护加州美丽环境”请愿书的家伙吧，更不会想到什么社交柔道术。

心理学家休·弗朗茨（Sue Frantz）描述过巴黎街头这种手法的邪恶版。一名骗子走近游客，请后者签署一份请愿书“支持聋哑人”。紧接着，他们就会要签了名的人捐款，不少人都会照做，以与刚刚表示支持的态度保持一致。但这一招完全是个骗局，因为慈善机构不会收到任何捐款——它们全都落入了骗子的腰包。更糟糕的是，请愿人的同伙会在

一旁观察游客从什么地方掏出钱包（是口袋还是衣兜），并将其作为随后扒窃的靶子。

发自内心地认同

每当你做出一个选择，你都在把自己核心的部分，也就是进行选择的那部分，变成与此前稍有不同的东西。

——英国作家C. S.刘易斯（C. S. Lewis）

请注意，所有的“登门槛”专家似乎都对同一件事情感到兴奋不已：你可以利用一个小的承诺操纵一个人的自我形象；你可以利用它们把人们变成“公仆”，把潜在客户变成“客户”，把战俘变成“合作者”。只要你把一个人的自我形象设置在了你想要的位置上，那么这个人就会自然而然地遵从一整套与这一全新自我形象相一致的要求。

倒也不是所有的承诺都会影响自我形象。要想承诺达到这样的效果，必须满足一定的条件：它们得是当事人积极、公开、经过一番努力后自由选择的。A、B两国战争中，C国的主要意图并不单纯是从战俘身上索取情报，而是要教化他们，改变他们对自己、对A国政治制度、对A国在战争中所扮演的角色等一系列问题的态度和看法。战争结束后，心理评估小组的负责人亨利·西格尔博士（Henry Segal）考察了被释放回国的战俘后报告说，这些人对战争的信念发生了根本性的转变，他们的政治态度也受到了重大影响。

看起来，C国的真正目标是修正，至少是暂时性地修正战俘们的心灵和思想。西格尔得出结论，倘若我们从“变节、不忠、改变态度和信仰、败坏军纪、打消士气和团队精神、怀疑A国扮演的角色”这些角度来衡量他们的成绩，“C国干得非常成功”。让我们更仔细地看看C国是怎么做的。

神奇的行为

要判断人的真正感觉和信仰，光听他们怎么说是不够的，还要看他们怎么做。想通过观察判断某人是个怎么样的人，必须仔细考察他的行为。研究者发现，人们自己也是依靠观察行为这一同样的方式，来对自己加以判断的。行为是确定一个人自身信仰、价值观、态度以及他接下来想怎么做的主要信息源。网站常常希望访客注册，提供自己的信息。但86%的用户表示，他们有时会因为表单太长、太暴露隐私而退出注册过程。网站开发人员要怎么做来克服这一障碍、又不减少从客户那里所得的信息量呢？他们减少了表单第一页上请求信息的平均字段数，因此给用户一种已经开始并完成了注册过程初始环节的感觉。设计顾问迭戈·波扎（Diego Poza）指出：“下一页是否有更多字段要填写（的确有）变得不重要了，出于承诺与一致原则，用户有更大概率会注册到最后。”现有数据证明他是对的：只要将首页字段数从4个减少到3个，注册完成率便会提高50%。

行为会对人们的自我概念造成影响，反过来，自我概念又会影响到人们将来的行为。这种连锁推进现象，可以参见有关主动与被动承诺效应的研究。

影响力研究

有一项研究要大学生志愿到当地学校开展艾滋病教育。在研究者的组织下，有一半的大学生是自己主动报名的，他们填写了一张申请表，说自己想要参与；另一半的大学生则是被动登记的，因为他们没有填一张说自己不想参加的申请表。三四天之后，研究者要大学生们开展志愿活动，绝大部分（74%）真正来尽责的学生都是那些主动答应参加的。而且，那些主动报名的人更有可能用个人价值观、喜好和个性

来解释自己的决定。

总之，主动承诺似乎能给予我们用来塑造自我形象的合适信息，而自我形象又能塑造人们将来的行为，使之起到巩固自我形象的作用。

C国战俘营的管理层对自我认知的这种变化路线深知其详，他们对战俘营做了种种巧妙的安排，好让俘虏们总能按照他们想要的样子做事。C国人知道，过不了多长时间，这些行为就会起到作用，让战俘们把自我认识与自己做过的事情调整到一致。

写作就是C国人不断敦促战俘进行承诺的一种做法。战俘们光是安静地倾听或者口头认同他们的方针路线还远远不够，必须得把它写下来。沙因描述过C国采用的一种标准改造手法：让战俘把问题写出来，再要他自己给出（支持C国的）答案。要是他不愿自己写，从笔记本里摘抄也行，这似乎是个没什么恶意的让步。

唉，这些“没什么恶意的让步”啊。我们已经见识到，看似无关紧要的承诺能让人做出更深入的一致性行为。而书面宣言这种承诺方式的好处就更加明显了。首先，它成了一个行为业已发生的物证。只要战俘写下了C国人想要的东西，他就很难再剖白自己没这么干过。他或许可以忘记或否认自己的口头发言，但写下来之后，一切就成了白纸黑字。那上面是他自己的笔迹，再怎么开脱也没有用了。这样一份文档，会驱使他调整自身的信念和自我形象，使之与先前没法抵赖的行为达成一致。其次，书面自白可以拿给其他人看。当然，这就意味着可以用它来说服其他人，劝说别人朝着声明里的方向改变态度。最后，更重要的是，书面承诺能说服其他人：写这份东西的人，真心相信自己写下来的事情。

人们有一种天然的倾向，总认为声明反映了当事人的真实态度。出奇的地方在于，哪怕他们明知道当事人作的声明并非出于自愿，他们还

是这么认为。心理学家爱德华·琼斯（Edward Jones）和詹姆斯·哈里斯（James Harris）做过一次研究，为此提供了一些科学证据。

影响力研究

研究者给被试看了一份支持菲德尔·卡斯特罗的文章，并要他们推断作者的真实感受。爱德华和哈里斯告诉部分被试，作者是自愿写的；又告诉另一部分被试，作者是出于胁迫才写的。奇怪的是，后一部分被试明知道作者是被逼无奈的，还是觉得作者喜欢卡斯特罗。

看起来，一篇表明信念的声明能按下读者的自动反应播放键，除非另有强有力的反面证据，否则旁观者会自动假定写这份声明的人写的都是真心话。

想想看，一旦战俘写了支持C国或反对A国的声明，这对他的自我形象会造成什么样的双重效应？这份声明不仅会让他一直记着自己的行为，还能让他周围的人相信，这反映了他的真实信仰。我们会在第4章里看到，周围的人认为我们怎么样，对我们的自我认知起着十分重要的决定作用。例如，有人做了这么一项研究：康涅狄格州纽黑文市的一群家庭主妇听说人家觉得自己乐善好施，过了一个星期，当“多发性硬化症协会”的募捐员找她们捐款时，她们果然大方了许多。很明显，光是知道有人觉得自己乐善好施，就让这些主妇做出了与之一致的行为。

瑞典一家超市在果蔬区进行了一项研究，也得到了类似的结果。顾客们看到了两箱不同的香蕉，一箱贴有“生态种植”的标签，另一箱则没有。在这种情况下，只有32%的顾客会选择生态香蕉。另外两组样本里的购物者看到的是两口箱子之间插着一块告示牌。一组人看到的告示牌通过生态香蕉的价格进行促销，说“生态香蕉与普通香蕉价格相同”，这将生态香蕉的购买率提高到了46%。另一组样本顾客里，促销标语为购

物者赋予了环保的公众形象，“你们好，各位环保主义者，我们的生态香蕉就在这里”，它将生态香蕉的购买率提高到了51%。

给人贴上标签既然具有这样的特点，精明的政客自然不会放过这个手段。埃及前总统穆罕默德·安瓦尔·萨达特就精通此道。每当开始国际谈判之前，萨达特总会对谈判对手说，天下人都知道，你们的国家和人民素来是讲求合作与公正的。靠着这样的赞美话，萨达特不仅创造了一种积极正面的感觉，还把对手的身份跟有利于己方的行为联系了起来。按照谈判大师亨利·基辛格的说法，萨达特取得成功，主要是因为他先给了对方一个好名声，而对方为了维护这个好名声，总会采取对萨达特有利的行为。

一旦主动做了承诺，那么，自我形象就要同时承受来自内外的一致性压力。一方面，人们内心有压力要把自我形象调整得与行为一致；另一方面，外部还存在一种更为鬼祟的压力，人们会按照他人对自己的感知来调整形象。由于别人觉得我们相信自己写的东西，哪怕我们实际上是逼不得已才写的，这种力量也还是会再一次迫使我们把自我形象调整得跟书面声明一致。

在A、B两国战争中，C国并不靠直接的胁迫，而是采用了若干种微妙的方式来让战俘写出他们想要的东西。比方说，C国知道许多战俘迫切想要告诉家人自己还活着，同时，战俘们又知道C国会审查自己的往来信件，只有一部分符合规定的邮件才能寄出战俘营。为了让自己的信件能够寄出，有些战俘开始在信里穿插呼吁和平的信息，说自己在这里挺好的，还表达了对B、C两国人民的同情。他们希望C国人允许这些信件寄出去，送到自己家人的手里。C国人当然也乐得顺水推舟，因为这些信对他们的好处太大了。首先，A国军人支持B、C两国的言论极大有助于B、C两国开展的全球宣传攻势；其次，在战俘改造过程中，C国人毫不费力就得到了好些人支持他们事业的公开宣言。

类似的技巧还包括在战俘营里定期举办政治征文比赛。获胜的奖品

没什么大不了的，几支香烟、一些水果，但这些在战俘营里也相当稀罕，因此战俘们还是非常感兴趣的。通常，获奖的文章都确凿无疑地站在支持C国的立场上.....但也不一定。C国人很聪明地意识到，要是比赛只有靠写支持C国的宣言才能得奖的话，大多数囚犯是不会参赛的。此外，C国人也很明白，只要能在战俘心里埋上一颗对和平承诺的小小种子，靠着悉心培育，以后是会结出果实的。所以，也有一些整体上支持A国立场，但对C国看法稍微附和了一两处的文章获奖。

这种策略完全带来了C国人想要看到的结果。战俘们一次又一次地自发参赛，因为他们发现，写赞美自己祖国的文章也能获奖。然而，在有意无意之间，他们逐渐把文章的基调调整得更加偏向C国了一些，以便得到更大的胜算。但凡是向C国让了步的文章，C国人都很欢迎，因为这样他们才好向当事人施以保持一致的压力。从C国的立场来看，战俘自愿写的这些文章是一份完美的承诺。靠着它，战俘的合作和立场转变很快就能顺理成章地确立下来。

其他说服专业人士也深知书面声明的承诺力量。例如，极为成功的安利公司就有这么一套刺激销售人员实现越来越高目标的好办法。他们要员工拟定个人销售目标，而且得亲手写下来，建立起对这些目标的承诺感。

定下目标，把它写下来。无论你的目标是什么，关键是你定了这个目标，这样你就有了努力的方向。接着，把它写出来。把东西写下来，有种非凡的力量。所以，定下目标，把它写下来。等你达到了这个目标，再定另一个，也把它写下来。这样，你会进步如飞的。

既然安利的人都发现“把东西写下来有种非凡的力量”，其他的商业组织自然也知道这个奥妙。有些上门推销的公司利用书面承诺的魔法来对抗许多州的“冷静期”法，这些法律允许消费者在买了东西几天之后取

消交易，获得全额退款。起初，强买强卖的公司因为这种规定遭受了沉重的打击。由于强调高压手法，他们的客户买东西大多不是因为想要那种产品，而是因为受了骗，在胁迫之下才答应成交的。这些法律生效以后，这部分消费者成群结队地取消了交易。

自从此类企业学会了一套简单漂亮的手法以后，取消交易的数量就大幅下降了。方法很简单，只需要让消费者而不是推销员来填写销售协议就行了。根据某知名百科全书销售公司的销售培训课程：个人承诺是预防客户撕毁合同的一种重要心理机制。和安利公司一样，这些组织发现，只要让人们把承诺写到纸上，就会出现神奇的事情：人们当真会照着写的去做。

还有一种看起来没什么不良企图的促销方式，也利用了书面声明的神奇魔法。在开始研究社会影响力武器之前，我根本搞不懂宝洁和通用食品这类大公司为什么总是要举办“25、50或100字”的宣传征文比赛。比赛全都大同小异，参赛者以“我喜欢某某产品，因为……”开头，写一篇短小的个人声明，把当时在售的某种蛋糕粉、地板蜡的特点吹嘘一番。公司对参赛文章进行挑选，并对获奖者颁奖。

我很困惑，这些公司这么做能得到什么好处呢？参加这种比赛又不必真正购买产品，只要提交一篇短文就够了。可是，各家公司似乎都很乐意掏腰包举办比赛，还办了一次又一次。比如，图7-4就反映了同类问题。



写下我们的优点，
赢取 25 000 美元。

There are two ways to benefit from the Teacher's advantage. Trace it, and you may win \$25,000. Drink it, and you'll be instantly rewarded by the rich, assertive taste of Teacher's Highland Cream.

How to enter
Teacher's \$25,000 Sweepstakes

TEACHER'S SCOTCH OFFICIAL RULES

1. Trace not all entries but listings on the missing words. Winners may be found on the face label of a Teacher's Scotch bottle. Then complete your entry by clearly printing your name and address. Entries must be submitted on an official entry form (place 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000).
2. The winner will be selected in a random drawing from among all entries received and mailed before the deadline.

OFFICIAL ENTRY FORM
To enter, trace not all words but listings on the missing words. Winners may be found on the face label of Teacher's Scotch bottle.

Trace or print missing words:
Mail completed entry form to:
Teacher's \$25,000 Sweepstakes
P.O. Box 20, New York, N.Y. 10108

Name: _____
Address: _____
City: _____ State: _____ Zip: _____

图7-4 书写即相信

这则广告邀请读者手写出自己最喜欢的帝雀斯威士忌的特点，邀请其参与抽奖活动。

如今，我揭开了这个谜。征文比赛的目的是让尽量多的人写下对一种产品的表白，就跟C国人在战俘营搞政治征文比赛一样，这两种举动过程如出一辙。为了得到吸引人的奖品，尽管获胜的概率很小，参与者也自愿写文章。他们知道，要想文章胜出，就必须赞美相关的产品。于是，他们开始寻找该产品值得称道的地方，并在文章里加以描述。随之而来的结果是，成千上万的人以书面的形式证明了该产品具有这样那样的优点。在书面文字的神奇推动之下，他们真正相信了自己所写的东西。

主动、公开、付出了努力并自由选择的承诺能改变我们的自我形象的原因在于，其中的每个元素都给了我们必须真正相信的信息。如果你

曾为某个特定的立场采取过行动，觉察出自己对它的认同，那你很可能会认为自己对此立场持有强烈的个人信念。同理，如果你认为自己所持立场所有人都看得到，要付出很大的努力，且纯属你的自愿选择，那你同样会认为自己对此立场持有强烈的个人信念。人们对自我概念的影响可能会带来适应性和持久的转变。

反过来，又有研究表明，旁观者倾向于相信一段陈述的作者相信这段话，除非另有强有力的反证。

READER'S REPORT | 读者报告

致西奥迪尼先生：

20世纪90年代末，我问赛百味连锁餐厅的创办人兼CEO弗雷德·德鲁卡（**Fred DeLucca**），为什么他非要在每一家赛百味分店的餐巾纸上印上这样一句话：**2001年开10 000家分店**。它似乎没什么意义，一来因为我知道他离目标还有很长的一段路要走；二来消费者们并不怎么关心他的发展计划；三来加盟的特许经营者们也很为与这一目标相关的竞争担忧。德鲁卡回答说：“要是我把目标写下来，让全世界都知道，我就一定会很努力地去实现它。”不用说，他超乎预期地实现了这个目标。

某大型国际广告公司的创意总监：

到2021年1月1日，赛百味已经在全世界110个国家拥有38 000多家分店了。因此，把目标写下来，公开做出承诺，不仅能按你希望的方式影响其他人，还能影响到我们自己。

众目睽睽

书面声明能有效地使人发生改变，原因之一在于它们很容易公之于众。A、B两国战争中A国战俘的经历，说明C国清晰地意识到了一条重要的心理学原理：公开承诺往往具有持久的效力。C国不断把战俘支持他们的声明拿给别人看，把声明贴在战俘营里，让作者在战俘讨论小组里大声朗读，甚至通过战俘营的广播站加以宣传。反正，对C国来说，这些声明弄得越多人知道越好。这是为什么呢？

每当一个人当众选择了一种立场，便会产生维持它的动机，因为这样才能显得前后一致。我在本章前面的部分提到过，前后一致是一种很好的为人特点，不具备这一特点的人，会被视为浮躁、多变、优柔寡断、糊涂、欠缺稳定；具备这一特点的人，则会显得理性、自信、可靠、值得信赖。考虑到这样的因素，也就怪不得人们总是避免显得前后不一了。因此，出于观感上的原因，立场越是公开，人们就越不愿意对其做出改变。

由杰出心理学家莫顿·多伊奇（Morton Deutsch）和哈罗德·杰拉德（Harold Gerard）所做的一个著名实验，阐释了当众承诺是如何进一步带来与之一致的行为的。

影响力研究

实验的基本过程是给大学生们看一些直线，让他们先在脑子里估计一下直线的长度。此时，第一组学生需公开自己的最初判断，把估计值写出来，并签上名字，交给研究者。第二组学生也得对自己的估计做出承诺，但他们只需把数值悄悄写在一块磁性书写板上，并可趁没人注意时修改。第三组学生则完全不需要对自己做出承诺，只需记住最初的估计值就行了。

通过这样的方式，多伊奇和杰拉德巧妙地做了安排，让一部分学生对自己的最初决定当众做了承诺，一部分学生私下做了承诺，一部分学生完全不做承诺。多伊奇和杰拉德想知道，三组学生里哪一组在知悉自己的判断不正确的情况下，仍有很大可能会坚持到底。所以，研究者拿出新的证据告诉所有的学生，他们原来的估计是不对的，现在有机会可以更正自己的估计值。

没把自己的最初估计写下来的学生，对这些选择是最无所谓的。看到新的证据不支持他们脑子里做出的最初决定，这些学生便立刻受到影响，赶紧改成了看似“正确”的决定。较之这些没做过承诺的学生，把估计值写在磁性板上给自己看的学生便不那么乐意改变主意了。尽管他们是在谁都不知道的情况下对自己做出的承诺，但把最初的估计写下来的行为，仍然使得他们对新数据产生了抗拒情绪，他们坚持最初的选择不愿更改。最不情愿改变初始立场的，还要数那些把最初的估计值当众记录下来的学生，公开承诺把他们变成了最顽固不化的人。

就算是在准确远远重要过保持一致的环境下，也不乏这种死脑筋。有人做过研究，在实验里找来6~12人组成陪审团，裁断一桩陈年旧案。较之无记名投票的方式，陪审员举手投票表达意见时，固执己见的人更多。一旦公开了个人的最初看法，陪审员就不愿再当众做出改变。要是你有机会在陪审团里当领头的召集人，选择不记名投票而不是当众投票，能帮你减少碰上死脑筋陪审员的风险。

多伊奇和杰拉德的这项发现，即人们会更忠于自己的公开决定，可以善加利用。有些致力于帮助人们摆脱坏习惯的组织就做得很好。比如，不少减肥诊所就明白，一个人私下决定减肥，大多是经不住诱惑

的，经过面包房的橱窗、闻到饭菜的香味、半夜看到美食广告，意志力很容易就溃败了。所以，他们认为，减肥决定必须用公开承诺的大梁来加以支撑。他们要求客户写下短期内的减肥目标，拿给尽量多的朋友、亲戚和邻居看。诊所经营者报告说，很多时候，其他方法都失效了，这种简单的小策略却能成功。

其实，要想跟公开承诺缔结盟约，没必要专门掏钱去诊所。圣地亚哥的一位女士向我讲述了她是怎样靠着公开承诺戒掉烟瘾的。她找了些空白卡片，在每张卡片的背面写道：“我向你保证，再也不抽烟了。”她把这种签了名的卡片交给“生活中所有我真正想要得到其尊重的人”。她说，此后每当她想要抽烟，就会想起要是自己没能信守承诺，那些人会怎么看待自己。于是她再也没抽过一根烟。如今，一些与社交网络挂钩的行为改变应用程序，使得我们能在远比几张卡片能接触到的更大朋友圈中使用这种自我影响技巧。具体例子可参见“线上影响力”专栏。

EBOX | 线上影响力

怎样改变你的生活

艾丽西亚·莫格 / 文

欧文·托马斯（Owen Thomas）最近在《纽约时报》上带着惊讶的口吻写到，他通过一款手机应用程序成功减掉了38公斤。他使用的这款应用叫MyFitnessPal。应用程序的开发人员发现，把自己的热量数值透露给朋友的用户，比普通用户多减掉了50%的体重。

社交网络能帮你做出改变，看似是件显而易见的事，但它是怎么做到的，就不太为人所知了。许多人认为发挥作用的是社会认同（观察他人来判断自己应该怎样做），但承诺与一致，能更好地解释这一行为上的转变。

我们所做的承诺越是公开，我们就越是感到有压力要依照承诺行事，进而显得行为前后一致。按罗伯特·西奥迪尼的说法，它可以变成良性循环（但也可能变成恶性循环），“你可以利用很小的承诺来操纵一个人的自我形象”，一旦你改变了这个人的自我形象，你就能让对方去做任何与这一全新自我形象一致的事情。

想要改变你的生活吗？做出具体的承诺，利用社交媒体来宣扬它，借助你内心所感受到的压力坚持到底。这反过来会让你以全新的方式看待自己，从而让你继续坚持。

托马斯先生的经历证明了这一理论在节食方面的威力，不过，我认为它可以应用到任何地方。比如苦苦挣扎的西班牙裔高中生（他们的高中辍学率最高）。为什么不让他们做出上大学的公开承诺呢？或许能有更多人真的上了大学。应该为此开发一款应用程序。

作者点评：在这篇博客文章中，作者正确地判断出，尽管存在同侪压力，但让托马斯先生带来理想转变的原则，并非社会认同，而是承诺与一致。而且，有效的承诺是公开的，这与研究结果相吻合：减重目标的承诺越是公开，（无论长期还是短期的）成功率也就越高。

公开承诺往往会使之更长久，这一观点得到了充分的证据支持。在这些支持中，还有一种好玩的形式：研究表明，消费者对自己当众使用的品牌比私下使用的品牌更忠诚。而且有证据表明，我们既希望自己前后一致，也希望在他人面前保持一致。

然而，有一项研究的结果与我们针对当众承诺所得出的结论形成了鲜明的对比，该研究的数据表明，公开承诺目标，实际上会降低一个人达到目标的可能性。对现有文献进行回顾之后，一组研究者失望地说：虽然发现上述模式的只有这一个矛盾的数据集，但它得到了学术圈之外最广泛的报道——博客、畅销书以及一场有数百万人观看过的TED讲演。我们要怎样解释它的非典型模式呢？我认为，逆反心理可能（见第

6章）起到了一定作用。逆反理论认为，如果人们在认为是否采取行动代表了自己的一项重要自由的观念下，感受到了外部要其采取行动的壓力，就不太可能采取行动。

研究者要求被试说明自己将采取怎样的步骤来推进自己的教育目标。接下来，部分被试还要把步骤交给外部评估人员（研究者），以将步骤变成公开的信息。研究者会先对这些步骤进行判断，再让被试继续推进。处在私密条件下的其他被试，不需要研究者的批准即可推进自己的步骤。他们只需要提交计划步骤，不需要获得研究者的允许。结果表明，这样的流程使得被试不太可能采取特定步骤推进目标，除非：第一，目标对他们自己十分重要；第二，他们碰到了要让研究者批准自己的步骤这样的外部障碍。这和逆反理论的预期完全相符。

READER'S REPORT | 读者报告

致西奥迪尼先生：

我刚刚在报上读到一篇文章，说的是一个餐馆老板如何靠公开承诺来解决顾客预订了座位却没来的大问题。我不知道他事先有没有看过您的书，但他做的事情，却完全吻合您讲的承诺与一致原则。他告诉接线员，以后别再说些什么“要是您计划有变，请给我们来电”之类的话了，而要这样问：“要是您计划有变，请给我们打个电话来好吗？”然后等对方做出回答。这下子，顾客订了座位却没来的比例立刻从**30%**降到了**10%**，跌了足足**67%**。

加拿大的这位大学教授：

为什么这个微妙的变化竟能带来如此戏剧化的不同结果

呢？在我看来，关键在于接线员要求并等待来电顾客做出一个承诺。这种方法能刺激顾客做出公开承诺，提高他们之后履行承诺的可能性。顺便一说，这位精明的老板是芝加哥戈登餐厅的戈登·辛克莱（Gordon Sinclair）。接下来的“线上影响力”专栏介绍了这一手法的网络版。

EBOX | 线上影响力

兴奋吗？你的预订就在明天。
还打算来吗？

我会来的！

2019 年 8 月 31 日，星期六下午 6:30，4 人桌
确认 #：2019809112
看菜单 | 看路线
金水街北 4175 号
亚利桑那州斯科茨代尔
邮编 85251
电话 (480) 265-9814

 日历  修改  取消

已订妥

作者点评：如今，越来越多的餐馆会要求顾客在预约日期到来之前，在网上做出主动和公开的承诺，以减少爽约的情况。近来，我的医生所在的诊所也开始做同样的事情，并附带了额外一项提高顺从度的要素。在确认预约的电子邮件中，护士给了一个主动做出公开承诺的理由：“告诉我你能不能来，也是在帮忙确保其他所有患者能得到所需护理。”这一在线确认程序使爽约的患者减少了81%。

额外的努力

有证据清楚地表明，为一个承诺付出的努力越多，它对承诺者的

影响就越大。这样的证据比比皆是，近的就在身边，远的绕到地球对面也是一样。

让我们先从身边的案例开始：很多地方都要求居民将生活垃圾分类，这样有利于环保处理。而如何正确分类垃圾，各地的要求可能并不相同。中国杭州的情况就是这样，在该市的部分地区，正确的垃圾分类和处理步骤比其他地方更麻烦。在向居民介绍了正确处理垃圾所带来的环保效益后，研究者想看看，必须付出更多努力才能符合环保标准的居民，在整体上上会不会对环保更投入（减少家庭用电量）。情况果然如此。在生活垃圾分类上必须付出更多努力的居民，也会更努力地节约用电，支持环保行动。这一结果很重要，因为它表明，深化我们对一项使命的承诺（本例中指付出更多努力进行垃圾分类），可以激励我们从相关方面推进这一使命。

有关需要付出更多努力做出承诺会带来的更大影响，另一些研究发现，在购买商品或服务时使用带来更多心理不适感的支付方式（现金对支票，信用卡对借记卡），会让人对这笔交易和相关品牌都变得更忠诚，进而让人有更大可能重复购买。

付出努力公开承诺的例子，在遥远的原始部落同样存在。非洲南部有个叫“汤加”（Thonga）的部落，要求本族每一名男孩都要完成一套复杂的成年仪式，才能真正算是男人。跟许多其他原始部落的少年一样，汤加小伙子也要忍受许多折磨，方可得到族人的接纳，获得成年的资格。人类学家怀丁（Whiting）、克拉克洪（Kluckhohn）和安东尼（Anthony）简短而生动地描述了这场为期三个月的严峻考验。

影响力研究

等男孩长到10~16岁，父母就会把他送到“割礼学校”，这种学校每隔4~5年办一届。在这里，男孩跟其他同龄人一道承受部落成年男性的侮辱和折磨。仪式的第一道关卡由两

列手持棍棒的男人组成，男孩要从他们中间跑过去，接受他们的殴打。之后，男孩的衣服会被剥掉，头发也会被剪。接下来，男孩要坐在一块石头上，见一个全身覆盖着狮子毛的“狮人”。有人从背后偷袭他，等他转过头去看是谁在打他时，“狮人”便会抓住他的包皮，两下便割掉它。其后的三个月，男孩会被隔离在“神秘院”里，只有已经通过成年仪式的人才能看到他。

在整个成年仪式当中，男孩主要需经历6种考验：挨打、挨冻、挨渴、吃难以下咽的东西、受罚、承受死亡的威胁。只要稍微被人逮到一点借口，就会有刚通过成年仪式的人狠狠揍他一顿。而这个打他的人，是部落里的长者专门指派的。男孩不能盖东西睡觉，只能硬生生地忍受冬天的严寒。整整三个月里，男孩不准喝一丁点儿的水。他吃的东西上，通常会盖着一层从羚羊胃里掏出来的半消化的草，故意弄得十分恶心。违反仪式里任何一条重要规则，他都会遭到严酷的惩罚。比方说，有一种惩罚是在触犯者的手指头之间夹上木头棍，一个壮汉把手合在少年的手上使劲捏，几乎要把他的指头弄断。看管的人会告诉少年们，从前想要逃跑的，或是把秘密泄露给妇女和未成年男孩的人，统统已被吊死烧成了灰。少年们听了害怕，只能乖乖就范。

表面上看来，这些仪式显得十分怪异。不过，它们在原理和细节上，都跟学校兄弟会的入会仪式有着不少相似之处。大学校园里每年都要按传统举办“地狱周”，申请入会的新手，都要通过老会员设计的一连串活动，旨在试探他们生理、心理和社交上的承受极限。等这个星期过完了，只有坚持到底、通过了考验的男孩才能跻身正式会员之列。大多数时候，这些折磨只不过会让人感到分外疲倦虚弱罢了，可结果出格的例子也不时可见。

有趣的是，“地狱周”活动的特点跟部落成年仪式几乎别无二致。刚才我们说了，人类学家发现，汤加少年在“神秘院”里要承受6种考验。看看报纸上的新闻，每一种考验都能在兄弟会折腾人的入会仪式里找到。

- 挨打。14岁的迈克尔·卡罗格里斯在参加高中兄弟会奥米伽-伽马-德尔塔的“地狱之夜”入会仪式时受了内伤，在医院待了三个星期。他被自己将来的兄弟们投了“原子弹”：他们要他把手高举过头，之后兄弟们蜂拥而上，一起狠揍他的肚子。
- 挨冻。一个冬天的晚上，加利福尼亚的一名大学新生弗雷德里克·布朗纳被他今后兄弟会的前辈们带进了国家森林公园十五六公里深处的一处山坡上。弟兄们把他留在山上，让他自己找路回去。众人口里的“胖弗雷迪”只穿着单薄的衬衣和便裤，在严寒中瑟瑟发抖，跌下了陡峭的溪谷，摔碎了骨头，碰伤了脑袋。因为受伤无法走动，他只好缩着身子抵挡寒冷，直到被活生生冻死。
- 挨渴。俄亥俄州立大学的新生因为在“地狱周”里违反了“新申请入兄弟会者吃饭时必须爬着进餐厅”的规定，于是被关进了“地牢”。“地牢”的门关上之后，在将近两天的时间里，新生们只有咸菜可吃，什么喝的东西也没有，他们只有两个塑料杯子，那是为他们接尿喝用的。
- 吃难以下咽的食物。在南加州大学校园的卡帕-西格马兄弟会之家，11名新入会的成员看到眼前恶心的任务不禁瞠目结舌。他们每人的盘子里放着100多克的生肝。生肝切得厚厚的，浸满了油，男孩们必须将它一口吞下。年轻的理查德·斯旺森把生肝吞下去后又吐了出来，连续三次都没

吃下去。但他打定主意非把它吃下去不可，终于他把这块油浸过的肉塞进了喉咙里。可它却卡在了那里，怎么弄都上不来又下不去，斯旺森就这么被活活噎死了。

- 受罚。在威斯康星州，一个新申请入会的人因为忘了所有新人都必须在入会仪式上念的一段咒语，遭到了处罚。前辈们要他把脚放在一张折叠椅子的后腿下，然后让体重最重的兄弟会成员坐到椅子上喝啤酒。尽管这个人没有大喊大叫，但在接受惩罚的过程中，他两只脚的骨头都断了。
- 承受死亡的威胁。泽塔-贝塔-陶兄弟会的一名申请人被带到了新泽西州的一处海滩上，他被要求“自掘坟墓”。挖完坑以后，兄弟会的前辈们又要他躺进去，他照做了。几秒钟之后，大坑的侧面垮塌，把他给活埋了。等前辈们把他挖出来时，他已经没了呼吸。

部落的成年仪式和兄弟会的入会仪式之间还有另一个惊人的相似点：它们是不会消亡的。尽管人们想方设法地要取缔它们、打压它们，但这些仪式却异常顽强地存在着。当地政府或大学行政管理部门等权威机构什么办法都用过了，威胁、施加社会压力、采取法律行动、流放、收买、下禁令，想要劝说各方团体放弃入会或成年仪式里的这些既危险又羞辱人的做法，可是统统没用。权威机构严密监视的时候或许会有所改进，但那不过是做做表面文章，等周围的压力一过去，它们就又会立刻浮出水面，只是会进行得更为秘密，更加严厉罢了。

南非汤加成年礼自1958年报道以来，这套严格的仪式并无太大改变。例如，2013年，部落中的23名年轻人在为期9天的成年礼中丧命后，南非政府才被迫叫停包括汤加在内的多个部落的成年礼。类似的结论也适用于大学兄弟姐妹会的欺辱性入会仪式。这类仪式自1657年在哈佛大学首次留下记录后便一直存在，难于根治，且有出人命的可能。关

于入会的艰辛（无论其形式是让人感到尴尬还是痛苦）带给人积极影响的研究，已经扩展到了商业领域：如果消费者艰苦地（而非轻松地）得到了专享优惠，那他对这桩交易的态度就会更正面。

在有些大学，官方也试过用为社区服务的“帮忙周”来取而代之，甚至直接插手管控入会仪式。但兄弟会不是狡猾地规避这类管理，就是直接搞对抗。例如，在斯旺森窒息死亡之后，南加州大学的校长颁布新规定，要求所有入会仪式都必须由校方审定方可进行，在举办入会仪式期间，还要有成年辅导员在场。美国一份全国性杂志报道称：“新规定引发的骚乱相当暴力，连本市的警察和消防队都不怎么敢进入校园了。”

看到了这种必然后果，其他大学的管理者干脆打消了废除“地狱周”的念头。“既然折磨是一种普遍存在的人类行为，而且所有的证据都支持这一结论，恐怕没有一种有效的办法来禁止它。你不让它公开进行，它就干脆转入地下。你不可能禁止人进行性行为，不可能禁止人喝酒，恐怕也不可能消除折磨！”

折磨到底有什么迷人的地方，让这些兄弟会看重到了如此地步？每当有人想要取缔入会仪式里有辱人格的危险做法时，这些团体就会想方设法地逃避、破坏、抗议，这到底是出于什么原因呢？有人认为，这些群体本身就是由心理扭曲、社交紊乱的人构成的，他们就是想看到别人受到伤害和羞辱。但证据并不支持这一观点。例如，有人研究了兄弟会成员的人格特质，发现他们在心理调整方面比其他大学生还稍微健康一些；再者，兄弟会向来是出了名的积极参加社会上的公益活动，但他们就是不愿意把折磨人的环节从入会仪式里去掉。华盛顿大学进行的一项研究调查了许多兄弟会的章程，大部分兄弟会都有类似“帮忙周”的传统规定，但这种社区服务跟“地狱周”并行不悖。社区服务跟入会程序直接挂钩的只有一个例子。

这样看来，折磨仪式上作恶的那些家伙，大多是心理稳定、关心社会的正常人，只是到了某种特殊的时候，即新成员加入协会的时候，才

会跟周围的人一起，变得超乎寻常的严苛。因此，证据似乎是在说，仪式本身才是罪魁祸首。它那么严格，必定是因为里面有些东西对整个团体至关重要。折磨新人肯定起到了某种作用，正是这种作用令兄弟会拼死也要将它维持下去。那这种作用究竟是什么呢？

我个人认为，有一项社会心理学圈外鲜为人知的研究给出了答案。两名年轻的研究者，阿伦森和贾德森·米尔斯（Judson Mills）想要验证他们观察到的一个现象：“费尽周折才得到某样东西的人，比轻轻松松就得到的人，对这件东西往往会更为珍视。”他们的神来之笔是，选择兄弟会的入会仪式来检验这一猜想。他们发现，忍受了令人超尴尬的入会仪式才得以加入性学讨论小组的女大学生，会觉得自己新参加的这个小组及其讨论的东西是非常有价值的，尽管阿伦森和米尔斯预先安排好了，让其他小组成员“要多无聊有多无聊，要多无趣有多无趣”；另一些女生经历的入会仪式比较温和，甚至她们完全没有通过入会仪式就参加了讨论会，她们就觉得自己新加入的这个小组非常“没意思”。阿伦森和米尔斯两人又做了进一步的研究，结果也是一样。当女生需要忍受痛苦才能入会时，在入会仪式上越痛苦，她们后来就越容易说服自己：新加入的这个小组及其活动非常有趣、聪明、可取。

READER'S REPORT | 读者报告

致西奥迪尼先生：

我想告诉你上个月发生在我身上的一件事。我当时和男友在伦敦，看到一家文身工作室在招牌上自称“全伦敦最便宜的眉环穿孔”。想到那么疼，我有点惊恐，但还是决定要去穿孔。刺穿的疼痛让我几乎昏迷，我没法动弹，也睁不开眼。我难受极了，鼓足勇气说：“快送我去看医生。”医生来了，她对我说会变好的。过了**10**分钟，我感觉好些了，但我向你保证，那是我人生中最可怕的**10**分钟！

接着，我想到了父母，他们是不会为我做的这件事感到高兴的。但我想，或许我可以在去见他们的时候把眉环取下来。最终我还是决定算了，不取了，因为取下来太疼了。

意大利的平面设计师葆拉：

跟阿伦森和米尔斯研究中的年轻姑娘们差不多，你对自己忍痛获得的东西感到很满意，而且忠心不二。

这下子，入会仪式上的折磨、羞辱甚至殴打，就都变得有意义起来。汤加部落里的父亲眼里噙着泪水，眼睁睁地看着10岁大的儿子晚上躺在“神秘院”冰凉的地板上瑟瑟发抖；大学二年级学生在“地狱之夜”神经质地大笑打断兄弟会“小兄弟”的发言。这些并不是什么虐待狂的行为，他们这么做，是为了维持团体的生存。奇怪的是，这样的举动却使得未来的成员觉得自己加入的团体更具吸引力、更有价值。只要人们一直珍惜并相信自己经过一番努力才得来的东西，这些团体就会继续安排困难重重的入会仪式。团队成员的忠诚和奉献精神，能极大地提高团队的凝聚力和生存概率。有人研究了54种部落文化，发现内部最为团结的部落，都有着最严格、最戏剧化的成年仪式。依照阿伦森和米尔斯的解释，严格的入会仪式极大地强化了新成员对团体的承诺感。因此不足为奇，各团体必然会想方设法地维系这一事关组织将来存活的纽带，倘若有人想取消它，那可是万万不能应允的。

军事团体和组织也照样不能免俗。“新兵训练营”里的痛苦极具传奇意味，但又卓有成效。小说家威廉·斯蒂伦（William Styron）回忆了自己在美国海军陆战队新兵营的“集训噩梦”，同时也证明了它的效力：

在我认识的前海军陆战队队员里.....没有一个人不认为新兵训练是一口严酷的大熔炉，但从这个熔炉里熬出来以后，他们却变得更坚强、更勇敢、更能承受磨难了。

内心的抉择

不管是C国人改造战俘，还是大学兄弟会坚持入会仪式，只要对此类活动加以考察，就可以看出一些有关承诺的宝贵信息。能有效改变一个人自我形象和将来行为的承诺，似乎都是当事人当着众人的面，付出努力主动做的。然而，有效的承诺还有一个比上述三点，即公开、主动和付出努力更重要的特征。为了搞清楚它到底是什么，我们首先要解决C国战俘管理人员和大学兄弟会弟兄们所做的一些怪异举动。

头一桩怪事是兄弟会的章程无不拒绝把公共服务活动纳入入会仪式。前面我们提到过，有调查报告说，兄弟会经常开展社区项目，但社区服务跟入会仪式几乎总是独立开来的。这是为什么呢？倘若兄弟会的入会仪式追求的是付出了努力的承诺，那肯定可以在里头包含一些艰苦、麻烦的公益活动，如修葺一下破旧的房子，到心理健康中心扫院子，去医院帮忙倒痰盂，这些事儿都是足够累人又不好玩的。再说，这类公益活动能极大地改善兄弟会“地狱周”仪式在公众心目中和媒体上的负面形象。调查显示，报纸上每登出一则有关“地狱周”的正面新闻，就会登出与其相关的另外五则负面新闻。就算光从公共关系的角度出发，兄弟会也应该把社区服务活动纳入入会仪式里，但他们偏不。

要看第二桩怪事，我们得回到C国战俘营和那儿为A国战俘举办的政治征文比赛上。C国希望能让尽量多的A国人参加比赛，让他们不知不觉地写一些支持C国的文章。然而，既然想吸引更多的人参与，为什么奖品却这么小气呢？征文比赛的获胜者最多只能得到一些额外的香烟和少量的新鲜水果。从战俘营的环境看，尽管这些奖品还算有价值，但设些更大的奖励也是很容易达到的：保暖的衣物、通信时的特别待遇、

更多的行动自由，这些东西都可以用来吸引人参加征文比赛。可C国人却特意选择了小气的奖品，不选更大、更吸引人的奖品。

尽管背景全然不同，但兄弟会拒绝在入会仪式里纳入公益活动，跟C国人不为征文比赛设置更刺激的奖品的原因是一样的：他们希望参与者对自己的所作所为负责，一旦做了，就没有借口可找，没有退路可选。新会员在入会仪式上主动承受了非人的折磨，他不可能说自己这么做只是出于慈悲心肠；同样，也不能让写了反对自己国家的政治文章的战俘有机会在事后耸耸肩说，“我只是贪图那份大奖罢了”。绝对不行。兄弟会的章程和C国战俘营的征文比赛都是要让人做了以后就回不了头，光让兄弟们、战俘们写出承诺还不够，还得让他们发自内心地为自己做过的事承担责任。

社会科学家已经确定了一点：只有在认为外界不存在强大的压力时，我们才会为自己的行为发自内心地负责任。

优厚的奖品就属于此类外部压力，它可以让我们去执行某一行动，但并不足以让我们自觉自愿地对此行动负责任。顺理成章地，我们也不会觉得该对它有什么承诺感。强大的威胁也一样：它能叫人当场顺从，却不大可能带来长期的承诺感。事实上，优厚的物质奖励或强大的威胁甚至可能减少或“破坏”我们对某一行为的内在责任心，一旦奖励或惩罚不存在了，我们就不愿意再去做它。

这些认识对教育孩子具有重要意义。它表明，对于我们希望孩子真心相信的事情，绝不能靠贿赂或威胁让他们去做，贿赂和威胁的压力只会让孩子暂时顺从我们的愿望。倘若我们不仅希望他们暂时顺从，还希望孩子相信自己做的事情是正确的，就算我们不在现场提供外部压力，他们也会继续照着我们乐于见到的方式去做，那么，我们就得做一些安排，让他们为自己的行为负起责任来。心理学家弗里德曼就做过一个实验，为我们在这方面提供了一些启示。

影响力研究

弗里德曼找来一种诱人的玩具，对一群二到四年级的小男孩说，玩它是不对的。他想看看6个星期后自己说的话是不是还管用。熟悉7~9岁小男孩的人，想必都知道这项任务是何等艰巨，但弗里德曼有个计划。他觉得，倘若能够先把男孩们说服，让他们发自内心地觉得玩这种玩具是错的，或许他们之后真的不会再去玩它。麻烦的是，怎样才能让孩子们相信玩这种玩具（昂贵的遥控机器人）确实不对。

弗里德曼知道，让男孩暂时听话很简单。只需要威胁他们说，要是逮到他偷玩机器人，他会遭到很严厉的惩罚。之后，大人会待在附近假装严厉“执法”，这样就很少有孩子会冒险去碰机器人了。弗里德曼猜得没错。

弗里德曼给男孩儿依次看了5种玩具，并警告说：“玩机器人是不对的。要是你玩了它，我会很生气，那时候我做的事情恐怕会让你不好受。”之后，弗里德曼离开了房间几分钟。在此期间，他通过一面单向玻璃暗中观察男孩。他先后找了22个不同的男孩做此尝试，在他离开的那几分钟里，有21个孩子摸都没摸过机器人。

所以，只要孩子们觉得有可能被逮到并挨罚，强大的威胁就管用。当然，弗里德曼早就猜到了这一点。他真正感兴趣的是，等过上一段时间，当他不在周围的时候，威胁还能不能有效指导孩子们的行为。出于这一目的，6个星期之后，他派了一名年轻的姑娘又来到男孩们的学校。

姑娘把孩子们从班上逐一叫出来，参与一项实验。她并未提及自己跟弗里德曼有任何关系，只是带着孩子们回到那间放有5种玩具的房间，说要给他们做个画画的测试。她一

边给测试打分，一边告诉男孩，想玩哪种玩具都行。当然了，几乎所有的男孩都玩了玩具。有趣的地方在于，所有玩了玩具的孩子中，77%都选了先前禁止他们玩的机器人。因为弗里德曼不能回来执行惩罚，6星期前非常管用的威胁，这下子差不多完全没用了。

弗里德曼的实验并未到此结束。他另选了一组男孩，把程序稍微调整了一下。他仍然先给孩子们看了5种玩具，也对他们说，在自己离开房间期间，别玩机器人，因为“那是不对的”。这一次，他并未威胁孩子非要他们服从。他只是离开房间，通过单向玻璃观察，看看他的指示是否管用。结果指示同样管用。和前一组男孩一样，弗里德曼短暂离开期间，22人里只有1个孩子玩了机器人。

6个星期之后，弗里德曼不在那里了，孩子们有机会跟机器人玩了，这个时候，两组男孩的真正区别显现了出来。先前没有施以强烈威胁的男孩做出了一件令人惊讶的事情：他们明明被允许想玩哪种玩具都可以，可大多数人都没去碰机器人。尽管在5件玩具里，机器人的吸引力是最大的。其他4种玩具分别是：一艘便宜的塑料潜水艇、一只儿童棒球手套但没有球、一把没上子弹的玩具来复枪和一辆玩具拖拉机。孩子们被允许选择其中之一来玩时，只有33%的孩子选了机器人。

两组男孩身上都出现了戏剧性的结果。对第一组男孩来说，弗里德曼说玩机器人“是不对的”，为了支持这一说法，他向孩子们施以严厉的威胁。在弗里德曼有可能逮到孩子们犯规的时候，威胁很管用。可之后他不在现场观察孩子们的行为了，威胁就没用了，他定的规矩自然也就作废了。看起来是这样：威胁并未让男孩们懂得玩机器人是错的，只不

过，要是存在挨罚的可能性，玩它便不够明智。

对另一组孩子来说，带来这种戏剧性不同的原因来自他们内心，而非外部。弗里德曼同样曾告诉过他们，玩机器人是错的，但他并未施以额外的威胁，说要惩罚不照做的孩子。最终结果有两点很重要：第一，光靠弗里德曼的指示，就足以在他短暂离开房间的时候，阻止男孩们玩机器人了；第二，自那以后，男孩们为自己不玩机器人的选择负起了责任。他们认为，不玩机器人是因为他们不想那么做。毕竟，就算他们玩了玩具，也不会受重罚，所以不能用这一点来解释他们的行为。因此，几个星期之后，当弗里德曼不在周围时，他们仍然不玩机器人，因为他们已经从内心相信自己不想玩了。

成年人要想教育孩子，可以从弗里德曼的研究里提取一点心得。假设有对夫妇想告诉女儿“说谎不对”。要是家长在场，或女儿觉得会被发现，那么明明白白的严肃威胁会很管用，比如，“宝贝儿，说谎不好，要是我逮到你说谎，我会把你的舌头给割掉”。但威胁没法实现说服小姑娘的长远目标，也没法让她打心眼里认为：因为说谎是错的，所以自己才不想做。所以，家长们需要采用一种更奇妙而有效的方法。你得找一个有力的理由，足以让她在大多数时候保持诚实，可这个理由又不能强大到让孩子觉得，自己完全是为了它才保持诚实的。

这有点棘手，因为不同的孩子需要的理由也不一样。对有些孩子来说，可能光是请求就够了（“宝贝儿啊，说谎不好，所以我希望你别说谎了”）；另一些孩子，可能要加上一个稍微强烈些的理由（“.....要是你说谎，我会对你失望的”）；还有些孩子，或许还得给予适当的警告（“.....要是你说谎，我恐怕不得不做些我不想做的事情”）。明智的家长自然知道哪种理由对自家的孩子适合。要点是找出一个理由，能让孩子从一开始就照着家长的意愿去做，同时又让他对这一行为自觉自愿地负责。也就是说，这种理由里蕴含的可察觉外部压力越少，效果就越好。对家长而言，选择一个合适的理由并不容易，但这番努力应该是物

有所值的。它决定了孩子是会在短期内顺从，还是会做出长久的承诺。正如塞缪尔·巴特勒（Samuel Butler）300多年前说过的一句话：“人违心答应倒是不难，可他的看法还是老样子。”

值得注意的是，人们为自己采取的立场支付小额报酬（与丰厚报酬相对），反倒能激发他们更大的承诺感，这一设想自首次预测以来就得到了连续不断的支持。例如，在最近的一项实验中，向朋友推荐某个品牌后可获得小额金钱奖励的被试，会对品牌更加青睐和忠诚。出于类似的思路，让人们做出自由选择可带来更大的忠诚感，这一设想也自阐述之初便得到了不断的支持。自愿选择会强化承诺的原因之一在于，它们激活了我们大脑的奖励区。研究显示，出于外部压力（如丰厚金钱奖励或惩罚）所做的承诺的效力会受到削弱。最后，出于内部原因而非外部原因做出的承诺，会给当事人带来更高的心理健康水平。

自己长出腿来

出于上文探讨过的两个原因，说服专业人士很喜欢能带来内心变化的承诺。一来，内心变化一旦出现，就跟当前的环境不挂钩了，它能涵盖所有相关的环境；二来，变化能发挥持久的作用。因此，一旦人们受到诱导，采取能改变自我形象的行为，比如说变成具有公益精神的好市民，那他们在其他多种情况下就都有可能热心公益，此时，说不定正好有人想要他们这么做呢！最后，只要新的自我形象能维持下来，他们就很可能继续从事热心公益的行为。

导致内心改变的承诺还有另一点吸引力：它们能自己“长出腿来”。说服专业人士不需要费时费力地花功夫来不断强化人们的内心变化，靠保持一致的压力就足够了，它会搞定一切。等人们逐渐认为自己是热心公益的好市民，就会自觉自愿地从不同的角度来看问题。他们会说服自己，人就该这么做。跟社区服务有关的事情，以前他们根本注意不到，现在却会给予关注。他们会听取有利于公益行为的论点，并觉得这些论

点比以前更有说服力了。一般来说，由于人们的内心信仰系统需要保持一致，于是他们会宽慰自己：我选择的行为是正确的。在生出额外的理由来为承诺的正当性辩护的这个过程中，最重要的一点在于，人们找到的理由是新的。因此，就算采取公益行为的初始原因没有了，这些新发现的理由也足以让人们继续认为自己的行为是正确的。

这对肆无忌惮的说服专业人士们可是天大的美事儿，因为我们会建立新的论据来巩固内心做了承诺的选择。有人便利用这一点，诱惑我们做出这种选择。一旦我们做好决定，那人便会取消诱因，他知道，我们的决定应该已经自动长出了腿，足够站得稳稳当当了。汽车经销商经常通过一种叫“抛低球”的伎俩从这一过程中渔利。我头一回碰到这套手法，是在本地一家雪佛兰汽车经销商那里当销售学员的时候。经过一个星期的基本训练，商家允许我旁观正式的销售员上岗，我立刻注意到了他们“抛低球”的做法。

影响力研究

汽车经销商对某些客户会提供十分优惠的价格，比如某款车比竞争对手的价格要低上700美元。不过，这笔划算的交易可不是真的，经销商根本无意兑现，它的唯一目的是让潜在客户决定在本店买车。一旦客户做了决定，经销商就会采取一系列的活动，培养客户的个人承诺感：填写一大堆购车表；安排各方面的贷款条件；有时候，还鼓励客户试驾一整天的车，之后再签合同，“这样你就有了拥有这辆车的感觉，还开着它给邻居和同事看了”。经销商知道，在此期间，客户一般会找出大把的新理由来支持自己的选择，证明自己的投资很划算。

之后便会发生一些事情。有时，销售员会在计算中发现一个“错误”，比如忘了把导航设备算到成本里，倘若买家还

是要空调，那就得把700美元重新加到价格当中。为了撇清自己的嫌疑，有些经销商会让银行批贷款的工作人员发现错误。还有些时候，到了最后关头交易突然被驳回了。因为销售员跟老板汇报工作，老板唱了黑脸：“这样子卖车会亏钱的。”买一辆车要好几千甚至上万美元，多上700美元似乎也没那么肉疼，再说了，销售员会强调，价格跟竞争对手是一样的：“这可是你选的车呀，对吧？”

还有更阴险的“抛低球”。潜在客户开着旧车来买新车时，销售员答应以旧换新，故意抬高旧车的估价。客户觉得这笔交易太划算了，立刻就想成交。之后，等快要签订合同的时候，二手车经理说，销售员对旧车的估价高了700美元，并把换购补贴降到了正常水平。客户知道扣了钱之后的交易仍然是公平的，也就接受了，有时还会为自己想占销售员的便宜感到愧疚。我亲眼看到过一位妇女向对自己使用了“抛低球”伎俩的销售员道歉——而这时候，她正在签购车合同，销售员马上就能得到一大笔佣金呢！销售员装出有点受伤的样子，并努力挤出了一个宽容的微笑。

不管用的是哪种“抛低球”手法，顺序总是千篇一律。先给人一个甜头，诱使人做出有利的购买决定；而后，等决定做好了，交易却还没最终拍板时，卖方再巧妙地取消最初的甜头。在这种情况下，客户还会买车，看起来似乎不可思议，可它真的管用。当然，不是对所有人都管用，但效果也足够好了，许多汽车卖场都把它当成一项基本的顺从手法。汽车经销商意识到，个人承诺能建立起一套自圆其说的系统，能为最初的承诺找到新的理由。大多数时候，这些理由会像粗壮的腿一样，牢牢地支撑起最初的决定，就算经销商把最初的那根腿给抽走，决定也不会坍塌。面对损失，客户会耸耸肩一笑而过，甚至挺开心，因为还有那么多上佳的理由支持着他们的选择。买家们从来没有想过，要不

是最初先做了选择，这些额外的理由根本就不会出现。

看到汽车销售员先给甜头，得到客户有利于卖方的决定之后，再拿走最初给的甜头这一“抛低球”手法在汽车展厅里的威力，我决定换一种环境来检验它的效力。我想知道，对这一手法稍加调整之后，它的基本原理是否还站得住脚。倘若我的看法没错，我应该可以让它换一种不同的方式来发挥作用。如果我先给出一个优惠条件，让对方做出至关重要的决定，然后再给最初的协议加上一项令人不快的限制。由于“抛低球”手法的作用是使人们坚持先前的决定，那么哪怕情况发生了变化，变得不那么有利了，“抛低球”手法都应该管用。

为了验证我的猜想，我跟同事约翰·卡乔波（John Cacioppo）、罗德·巴塞特（Rod Bassett）、约翰·米勒（John Miller）在俄亥俄州立大学做了一次实验。

影响力研究

该实验是让心理系的同学们答应做一件不大愉快的事情：大清早起来参加早晨7点的“思维过程”研究。我们叫来第一组学生，当场就告诉他们早晨7点要出席研究会议的消息，只有24%的人愿意参加；接着，我们叫来第二组学生，抛出一记低球：先问他们是否想要参加“思维过程”的研究，等他们答应了之后（56%的人给出了正面回应），我们再提早晨7点出席的事实，并给他们反悔的机会。结果，没有一个人改变主意，而且，95%被抛了低球的学生都准时在早晨7点来到了心理系教学楼，履行了自己的承诺。我知道这一点是因为我招了两名研究助理到现场进行“思维过程”实验，并记下了到场的学生的名字。[\(45\)](#)

“抛低球”手法最令人印象深刻的一点在于，当事人明明做了一个糟

糕的选择，却还觉得挺高兴。那些没什么好选择给我们的人最喜欢这一套了。不管是在生意、社交还是私人场合，我们都能发现他们在“抛低球”。我的邻居蒂姆就是个真正的低球爱好者。大家想必还记得，他答应改变自己的行为方式，好让女朋友莎拉回心转意，取消跟别人的婚礼，让他回来住。自从莎拉决定选蒂姆以后，她对他更百依百顺了，尽管蒂姆当初的承诺从未兑现。莎拉解释说，她发现蒂姆身上有好些她从前没注意到的优点。

莎拉是个“抛低球”手法的受害者，我心知肚明。正如我在汽车展示厅看到买家们上了“给你甜头又拿走”策略的当一样，我知道莎拉也同样中了蒂姆的招。蒂姆从来就是老样子，丝毫没有有什么改变。可由于莎拉在他身上发现了不少对她来说实实在在的新优点，现在她对之前无法接受的安排深感满意。选择蒂姆的决定，从客观上看实在够糟糕的，可它已经长出了腿，站起来了，还让莎拉感到挺幸福。我从来没对莎拉提过“抛低球”的知识。我保持沉默，倒不是因为我觉得茫然无知对她更好。我个人从来都认为，多些信息总比少些好，这应该是一条基本的指导原则。只不过，要是我在他们俩的关系上多说一个字，莎拉铁定会恨死我，而且什么也不会改变。

为了公共利益挺身而出

本书中讨论的所有顺从技巧，都是既可为善也可作恶的，全看使用者的心术如何了。不足为奇，“抛低球”手法除了用在卖新车、跟老情人重建关系上之外，还能用在社会公益事业上。例如，社会学家迈克尔·波拉克（Michael Pallak）在艾奥瓦州主持过一个研究项目，该研究表明，“抛低球”手法能让居民注意节约能源。

影响力研究

研究项目是在艾奥瓦州的初冬时节开始的，相关人员到访使用天然气取暖的住户家，教给住户一些节能技巧，并请他们为了将来节约燃料。所有的住户都答应试试看。可一个月以后，研究者核对了这些家庭的气表，发现没人真正注意了节能问题。冬天过后又对比了一次，结果也相同。答应节能的居民的天然气使用量，跟随机抽选出来的、节能宣传员没上过门的住户一样多。看来，光有良好的意图和节能的信息，还不足以改变人们的生活习惯。

项目开始之前，帕拉克和研究小组就意识到，为了改变人们长期的能源使用模式，必须来点别的东西。所以，他们对艾奥瓦的另一组天然气用户采用了稍有不同的程序。宣传员同样到访了这些家庭，提供节能技巧，要他们注意节能。此外，宣传员还提出了一个额外的条件：要在本地报纸上登出这些答应节能的家庭的名字，表彰他们的公益精神和环保态度。这么一来，效果显而易见。一个月之后，燃气公司上门检查气表时，这一组居民平均每户节省了12立方米的天然气。有机会让名字见报，激励了这些居民在为期1个月的时间里努力节能。

接下来，研究者开始暗中动手脚，他们取消了最初促使人们节约使用燃料的原因。每户家庭都收到了一封信，其中说明出于种种原因，没办法再在报上登出用户的姓名了。

等冬天过去，研究小组开始调查这封信给天然气用户们带来了什么样的影响。没了上报的机会，他们会重新回到浪费能源的老路上去吗？完全没有。在当年冬天的其他月份，这些家庭节约的能源比头一个月，即以为自己可以受到登报表彰的奖励时的还要多！按百分比来看，在以为自己能见报的第一个月，他们节约使用了12.2%的天然气。然而，在收

到信、知道自己名字见不了报之后，他们并未“旧疾复发”，反而省下了15.5%的天然气。

虽说有些事情不一定完全靠得住，但要理解这些住户坚持节能的行为，有一个现成的解释。在登报表彰的“抛低球”诱惑下，这些人作出了节能的承诺。承诺一旦做出，就会开始长出腿来支撑自己。居民开始培养新的节能习惯，并对身体力行实践公益活动感觉良好；他们劝说自己，美国必须减少对外国能源的依赖；自家天然气使用费用降低，让他们很愉快；他们为自己的自制力深感骄傲；最重要的是，他们开始觉得自己很有节约精神。有了这些新的理由，先前节能的承诺就显得更为正当了。这样一来，哪怕最初的理由，即登报表彰没有了，承诺还是站得稳稳当当的（见图7-5）。

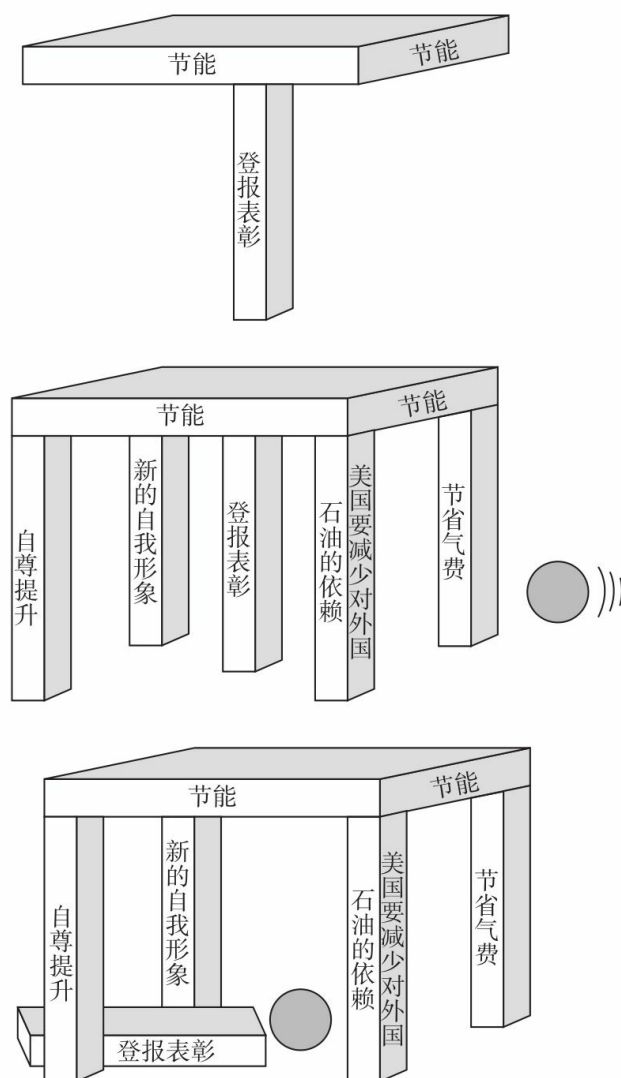


图7-5 抛低球带来的长期影响

在这个艾奥瓦能源研究的案例中，我们可以看到，人们最初的节能承诺是以登报表彰为前提的（上）。可没过多久，这种节能承诺就自己生出了新的支撑点，研究小组得以抛出低球（中）。它虽碰倒了一开始的登报表彰诱因，可新长出来的腿却牢牢地支撑住了节能的做法（下）。

最奇怪的一点是，明知登报表彰已经不可能实现了，这些家庭却反倒更加努力地节约能源。要解释此种现象，许多理由都说得通，但我最中意的是：在某种程度上，登报表彰的机会反而使居民没办法全力以赴地履行节能的承诺。他们做出节约能源决定的所有理由都发自内心，可登报表彰却是唯一一个外部因素。有了它，居民们没法说服自己，他们

节约用气是因为相信这么做是正确的。如此一来，收到信件、得知登报表彰一事取消之后，妨碍居民们塑造全新自我形象（具有节能意识、关注公益的好公民）的拦路虎也随之消失了，全情投入的崭新自我形象促使他们更努力地节约能源。就跟莎拉一样，居民们是因为引诱才做出了选择和承诺，可等引诱他们的东西没了时，他们对自己的选择却更加执着了。

还有其他几项实验也证明了抛低球手法在各种环境下的有效性，而且它适用于男女两性。研究者报告了一项确认了该手法有效的元分析，后来的研究也支持该手法的效力是基于承诺这一解释。

利用提示激发人们信守承诺

基于承诺的顺从手法，还有一项额外的优势。只要提醒当事人此前所做的承诺，就足以激发他在未来做出一致的反应，保持先前的立场、姿态或行为。让人回想起自己的承诺后，一致性需求便会让人们再次校准相关反应。为此，我们不妨举几个医学领域的例子。

每当向医疗管理群体介绍影响力过程时，我都会问：“在你们的系统里，什么人最难被影响？”无一例外，人们总是带着强调的意味回答说：“医生！”一方面，这种情况似乎理所应当。为了在医疗保健体系里得到崇高的地位，医生要经历多年的训练和实践，要接受医学专业教育、实习、当住院医师，由此掌握大量信息和经验以作为其决策的基础。可以理解，他们一旦做出选择，就很难动摇立场。可另一方面，如果医生不采纳他人的建议，拒绝做出有利于患者的调整，这种抗拒心理也很成问题。在最初走上这条职业道路之前，大多数医生都宣读过希波克拉底誓言，承诺要主要根据患者的健康采取行动，尤其是不得伤害患者。

那么，为什么医生在给患者做检查之前不按规定洗手呢？美国一家

医院进行了一项研究，就这一问题给我们带来了许多宝贵的见解。研究者亚当·格兰特（Adam Grant）和戴维·霍夫曼（David Hofmann）发现，不管如何强烈地建议医生在检查每一名患者之前洗手，大多数医生洗手的次数都不到规定的一半；而且，各种针对这一问题的干预措施都没有效果。因为忽视洗手这一环节，医生和患者都要承担更大的感染风险。造成这一问题的原因不在于医生放弃了对患者安全的承诺，也不是没有意识到该承诺与洗手的联系；相反，原因在于，进入检查室时，这种联系在意识上的地位不如患者的样貌、护士说了些什么、病例记录上写了些什么等其他因素那么高。

格兰特和霍夫曼认为，如果能让医生在进入诊室之前提醒自己对患者承诺、以及这跟手部卫生的关系，就能改善这个局面。研究者在诊室的清洁用品前放置了标语牌，上面写着“保持手部卫生能保护患者免于感染疾病”。这些标语牌将医生对肥皂和洗手液的使用率提高了45%。

以上研究还评估了另外两种贴在肥皂和洗手液上的标语牌的作用，它们都不是为了提醒医生对患者安全的承诺（“抹上清洁液，洗干净双手”和“保持手部卫生能保护患者免于感染疾病”），也都对肥皂和洗手液的使用量没有任何作用。

医生的另一项过失和过量开具抗生素处方有关。过量使用抗生素在美国是一个日益严重的健康问题，每年会导致23 000名患者死亡。和洗手的例子一样，几种旨在减少该问题的策略（如培训、电子提醒和现金奖励）都收效甚微。但有一群医学研究者在解决此问题上取得了显著的成功，他们对洛杉矶一组门诊诊所的医生运用了以承诺为中心的方法。医生在诊室里贴了一张展示周期为几个星期的海报。一半医生诊室中的海报为患者提供了有关使用抗生素的标准信息。另一半医生诊室中的海报里除了标准信息，还包括医生的照片，以及他对避免过度使用抗生素的签名保证信。接下来的统计表明，每天看到标准内容海报的医生，开

出的不恰当抗生素处方增加了21%。而面对海报不断提醒其承诺过减少这一问题的医生，开出的不恰当处方减少了27%。

提醒当事人想起自己已经做出的承诺，还有一点好处。人们不光恢复了承诺，似乎还通过增强相关的自我形象，巩固了承诺。跟从前采取过环保行动但未得到承诺提醒的消费者相比，得到承诺提醒的消费者认为自己更具环保意识，并有更大可能在购买灯泡、纸巾、除臭剂和洗涤剂等同类产品时选择对环境友好的版本。因此，要求人们回想之前的环保承诺，不仅能激发其做出与之前行为一致的后续反应，还强化了其“环保主义者”的自我形象。

如何防范

作家爱默生说过一句经常为后人引用的话：“一致性是渺小的心灵上的恶鬼。”这么说可能让人很费解，因为放眼周围，内在一致性显然是逻辑性强、智力高者的标志，而缺乏一致性则是知识零散、所学有限者的特征。那爱默生又为什么会认为一致性是头脑狭隘之人的特征呢？他指的到底是什么呢？回头看看他这句话的原始出处——其著作《自立》就能明白，问题不在爱默生，而在于这句话流传开来的版本。他实际上写的是：“愚蠢的一致性是渺小的心灵上的恶鬼。”岁月侵蚀了他原本准确的断言，把它的意思变得完全不同甚至完全错误了。

但我们绝不应该忽视“一致性”和“愚蠢的一致性”二者的区别。据我所知，要对抗结合了承诺与一致原则的影响力武器，唯一有效的防御措施是思想上的觉悟：尽管保持一致一般而言是好的，甚至十分关键，但我们必须避免愚蠢的死脑筋。我们必须警惕不假思索地自动保持一致的反应，因为有些耍花招的人正想利用它牟利呢。

不过，由于大多数时候自动保持一致都能让我们做出恰当又合算的行为，所以，我们在生活里不可能彻底摆脱它，否则可就坏事儿了。想想

看，要是我们不能依照从前的想法和行为顺理成章地做事，而是每次碰到新行动就要停下来想想它有什么好处，我们恐怕一辈子也做不成什么重要的事。机械地保持一致尽管存在风险，但我们还是离不开它。摆脱这一困境的唯一出路，是明白这种一致性有可能让人做出错误的选择。有两种独立的不同信号能给我们提示，面对其中任何一种信号，我们的身体都会有不同的部位会做出反应。

READER'S REPORT | 读者报告

致西奥迪尼先生：

我写信给您是想讲一件事。在这件事里，一致性原则逼得我做了个我一般不会那么做的决定。我去了一家购物中心的美食街，打算买一小杯可乐。

“请给我一杯可乐。”我对售货员说。

“中杯还是大杯？”他一边给另一位顾客结账，一边问。

“我已经吃饱了。不可能一口气喝完一大杯可乐。”我心想。“中杯。”我自信地边说边递上了信用卡。

“哎呀，对不起，”售货员好像真的犯了个口误似的，“小杯还是中杯来着？”

“哦，中杯。”我按照一致性原则说，接着拿起饮料走开了。这时，我才意识到自己上当了：我还是买了两种选项里较大的那一杯。

我被他打了个措手不及，为了和之前说的一致，我脱口而出地说要“中杯”，压根没有处理迎面而来的新信息。

愚蠢的一致性绝对是渺小的心灵上的恶鬼！

印度新德里的这名大学生：

碰到这种情况，你居然认为是自己心灵渺小，对自己未免太过苛刻。面对选择，人在匆忙中无法深入思考，机械地追求一致是很正常的。

在此我必须指出，哪怕是我们最熟悉的一些引言，在时间的流逝中也曾遭到篡改，意义大变。这种情况并不少见。例如，《圣经》中并没有说金钱是万恶之源，而说的是“对金钱的热爱是万恶之源”。为了不让自己犯同样的错误，我应该指出，爱默生的引言比我说的更长一些，而且更具文学性。它的全文是：“愚蠢的一致性是渺小的心灵上的恶鬼，受到小政客、小哲学家和小牧师的顶礼膜拜。”

肠胃信号

第一种信号很容易识别。每当我们意识到自己中了套、被迫遵从了一个并不想答应的要求时，我们的胃就会警铃大作。这种状况我碰到好几百回了。在我还没开始研究顺从技巧之前，有年夏天发生了一件特别令我难忘的事。

那天傍晚，门铃响了，我打开门，看到一个穿着短裤和吊带衫的漂亮姑娘。尽管如此，我还是注意到她带着笔记本，她是想要我参加一项调查。因为想给她留下个好印象，我答应了。我承认，为了让自己的形象显得特别高大光辉，我在回答问题时对实际情况稍微做了点夸张。我们的谈话如下：

漂亮姑娘：您好！我正在调查城市居民的娱乐习惯，不知您能否回答我几个问题？

西奥迪尼：请进。

漂亮姑娘：谢谢。我就坐在这儿开始好了。每个星期您出门吃晚餐的次数是多少呢？

西奥迪尼：哦，大概每星期三四次吧。说真的，只要有机会我都是出去吃，我喜欢高档餐馆。

漂亮姑娘：真棒。你晚餐时一般会点酒吗？

西奥迪尼：只点进口酒。

漂亮姑娘：明白。电影呢？您常去看电影吗？

西奥迪尼：电影院吗？好电影不多啊。我最喜欢看带字幕的外国文艺片了。你呢？你喜欢看电影吗？

漂亮姑娘：呃.....我喜欢。不过，我们还是回到调查上来吧。你经常去听音乐会吗？

西奥迪尼：当然。我最常听的是交响乐。不过，我也很喜欢有档次的流行乐队。

漂亮姑娘：（下笔如飞）很好！还有一个问题。歌剧或芭蕾舞团的巡回演出如何？城里有这类表演时你去看吗？

西奥迪尼：啊，芭蕾啊，动作、韵律和形式都是我喜欢的。写下来吧，我喜欢芭蕾。只要有机会我都会去看。

漂亮姑娘：很好。让我核对一下记录，西奥迪尼先生。

西奥迪尼：确切地说，是西奥迪尼博士。但这听起来太一板一眼啦，你叫我鲍勃多好。

漂亮姑娘：好的，鲍勃。根据您给我的信息，我很高兴地说，参加“美国俱乐部”您每年能省下1 200美元！只需交一点会员费，您刚才提到的大部分活动就都可以打折啦。像您这么喜

欢社交活动的人，肯定会利用我们公司的服务来享受特大优惠吧。

西奥迪尼：（像只落入陷阱的老鼠）呃……嗯……我……
呃……我猜会吧。

我记得很清楚，当我结结巴巴答应下来的时候，我的整个胃都缩紧了。这是一个明白无误的信号：“嘿，你上当啦！”但我看不到出路，我被自己说的话逼到了绝路上。此时拒绝她的提议，有可能出现两种结果，但两种结果都不怎么乐观。如果我想收回前言，说自己并不是刚才在采访里说的那种时髦的城里人，那我就成了骗子；可要是不收回前言就拒绝她，我又成了一个傻瓜蛋，居然会放弃能省下1200美元的机会。我买下了姑娘提供的娱乐套餐，尽管我知道自己是掉入了陷阱。为了跟先前说的话保持一致，我中了招。

不过打那以后我再也不上这种冤枉当了。如今，我会倾听肠胃的声音，也找到了一个办法来对付那些想要利用一致性原理坑我的人，即只需要一语道破他们在干吗就行了。这成了我的完美反击战术。每当我的肠胃告诉我，仅仅因为想要跟最初所做保持一致就顺从别人的要求是很愚蠢的时，我便会把这一点原原本本地告诉提要求的人。我并不否认保持一致的重要性，但我想指出，顽固地保持一致其实是荒谬透顶的。不管提要求的人听了以后是内疚地离开，还是困惑地撤退，对我都是件好事：我赢了，利用我的人输了。

我有时候会想，要是多年前的那位漂亮姑娘现在跑来向我卖娱乐套餐，那会是个什么情形。我已经准备好应付的办法了，之前的对话都一样，只不过结尾焕然一新。

漂亮姑娘：……像您这么喜欢社交活动的人，肯定会利用我们公司的服务来享受特大优惠吧。

西奥迪尼：（自信满满）大错特错。你瞧，我意识到这是怎么回事了。我知道，你表面上说是做调查，实际上是为了让人们告诉你他们有多喜欢出门，而在这种情况下，说得夸张些也很自然。但既然我知道这事儿动机不纯，我才不会让自己掉入“做了承诺就想保持一致”的自动陷阱里去呢。“按一下就播放”的把戏对我没用。

漂亮姑娘：啊？

西奥迪尼：好吧，这么说吧：第一，要是我在本来不想要的东西上花钱，那我就是个猪脑子；第二，最可靠的权威人士，也就是在下的肠胃，明确指出，我不想买你的娱乐套餐；第三，如果你还是觉得我会买它，那你就太天真了。当然喽，一个像你这么聪明的人，应该能理解吧。

漂亮姑娘：（像只落入陷阱的漂亮小老鼠）呃……嗯……我……呃……我猜可以吧。

心灵信号

肠胃不是感觉特别敏锐的器官。只有人明明白白是受骗了，它才能反应过来，向大脑传递这一信息。其他时候，如果我们挨宰挨得不那么明显时，肠胃就不会警铃大作。在这类情况下，我们必须到其他地方寻找线索。我的邻居莎拉是个典型的例子。她向蒂姆做了一项重要的承诺：取消跟前男友的结婚计划。承诺自己长出了腿，哪怕做出承诺的最初理由已经没了，她还是一如既往地坚守了承诺。她用新找到的理由说服了自己，相信自己做了正确的事，于是继续跟蒂姆在一起。可想而知为什么莎拉的肠胃没缩紧，只有自己都觉得做的事情不对，肠胃才能提醒我们。莎拉不是这样。照她想来，自己选得很正确，所以行为也该和选择保持一致。

不过，除非我彻底猜错了，莎拉身体上肯定有一部分已经意识到自己选错了，也知道她目前的生活方式是在冥顽不灵地保持一致。那一部分到底是哪儿，我们说不清，但在人类的语言里，早就给那里起好了名字：心灵。从定义上来说，在这个地方，我们没法自欺欺人。这个地方，我们自己找的那些所谓的理由和借口统统渗透不了。莎拉的真相也在那里，尽管她现在可能还没法清晰地听见它发出来的信号，因为她新找到的支撑腿脚不仅站得挺稳，还会发出噪声干扰她呢。

要是莎拉选蒂姆选得不怎么高明，她得死扛多久才能明确地意识到这一点，才会出现一场大规模的心灵动荡呢？谁也说不清。但有一点很肯定，随着时间的推移，能代替蒂姆的备胎男友也会越来越少。她最好赶紧弄清自己是不是犯了错。

自然，说起来容易做起来难。莎拉必须回答一个极度纠结的问题：“知道了我现在掌握的这些情况，要是能回到从前，我还会做出同样的选择吗？”问题就出在“知道了我现在掌握的这些情况”上，她现在对蒂姆的情况到底掌握了哪些呢？她对蒂姆的看法，有多少是为了想要证明自己所做承诺是正当的而绝望地编造出来的呢？莎拉说，自从她决定重新跟蒂姆在一起之后，蒂姆更关心她了；蒂姆努力地戒酒，还学会了做很好吃的鸡蛋卷。我尝过几次蒂姆做的鸡蛋卷，对此很是怀疑。不过，最重要的问题在于，莎拉是从心灵深处相信这些东西呢，还是只在理性上觉得是这样？

这里有个小办法，莎拉可以用它来弄清自己目前对蒂姆的满意度，有多少是发自内心的，有多少是出自死脑筋的一致性的。心理学证据表明，面对一样东西，我们总是先体验到感觉，过上短暂的一瞬间之后，才能将之理性化。照我猜测，心灵深处发出的信号是一种纯粹而基本的感觉。因此，如果我们多注意训练自己，应该可以在感觉十分轻微、认知器官还没来得及插手的时候就觉察到它。根据这一方法，倘若莎拉真的向自己提出了关键问题：“我会做出同样的选择吗？”她最好是

找到并信任自己在做出反应那一瞬间所感受到的灵光。这抹灵光很可能就是她内心深处发来的信号，它是趁着她为自己找的各种借口还没发挥作用之前偷偷溜出来的，一点也不失真。

我们对一件事的感受比对它的认知出现得更快，并不是说我们对事情的感觉总是会跟理性认识的不一样，也不是说前者更靠得住。然而，数据清楚地表明，我们的情感和理性往往不在同一个方向上。因此，倘若当前局面牵涉自己会产生合理支撑点的承诺，那么感觉提供的忠告通常更为准确。对诸如莎拉的幸福等事关情感的问题来说更是如此。

每当怀疑自己做事时犯了死脑筋保持一致的毛病，我就会试着使用这个办法。比方说，有一次，我停在加油站的自助加油泵跟前，这里汽油的广告价比本地区其他加油站的每加仑要低两分钱。但等我把加油泵拿起来，才看到泵上的标价其实比广告价要高两分钱。我向路过的服务员（后来我才知道他就是老板）提出异议，他低声说，价格几天前刚刚调整过了，还来不及把招牌上的价格改过来。当然了，这套说法不怎么可信。我试着决定该怎么做，脑袋里跳出了好些留下来加油的理由，如“我必须得加油了”“油泵空着，我又急着赶路”“我好像记得，我的车用这个牌子的汽油开起来更顺畅”。

我得判断这些理由到底是真的，还是我为就在这儿加油了事找的借口。于是我向自己提出了关键问题：知道了这里汽油的实际价格以后，要是能回到先前，我还会做出同样的选择吗？我集中注意力捕捉最先进发的感受，收到了一个毫不含糊的明确答案：我会马上开车走人，甚至停都不会停下来。这时我知道，要不是看在价格较低的份上，其他的原因根本就不会让我停下来加油。它们不是我做出决定的源头，是我做出决定之后才生编硬造出来的。

解决了这一点之后，我还得做出另一个决定。既然我已经把油枪拿在手里了，那到底是把油加上好，还是忍受不便，到别的地方出一样的价格加油呢？幸好，这时候加油站服务员兼老板走过来帮我拿定了主

意。他问我为什么迟迟不加油，我告诉他我不喜欢他在价格上弄手脚，他怒吼起来：“听着，我怎么做生意是我的事，别人少插嘴。要是你觉得我骗了你，赶紧把油枪放下，尽快滚出我的地盘，小子。”因为已经肯定他是个骗子了，我很高兴地照着自己的信念行了事，这同时也满足了他的愿望。我把油枪放下，开着车走了。有时候，保持一致也挺值得一做呢！

最容易受攻击的弱点

有没有哪些人必须跟从前的行为保持一致，特别容易中上述承诺策略的招呢？当然有。要了解这类人的性格特点，让我们来看看当今一位著名体育明星在生活里碰到的倒霉事。

影响力研究

2005年3月1日，高尔夫界的传奇人物杰克·尼克劳斯17个月大的孙子竟意外溺死在浴缸里。一个星期后，仍沉浸在悲伤中的尼克劳斯推掉了之后所有跟高尔夫相关的活动，包括即将举行的大师赛。他说：“我们家发生了这样的惨剧，我想，自己的时间得多花在其他的事情上了，跟高尔夫比赛有关的活动，我目前完全没有计划参与。”可就在发表这番声明的当天，尼克劳斯却做了两件自食其言的事：向佛罗里达州一家高尔夫俱乐部的准入会员发表了讲演；在老对手盖里·普莱耶主办的慈善锦标赛上打了球。

是什么东西的力量这么强大，竟把尼克劳斯从悲伤的家人身边拉走，去做了两件跟陪伴家人比起来完全无关紧要的小事呢？尼克劳斯说：“你做了承诺，就必须履行诺言。”他的回答就是这么简单。尽管考虑到背景情况，这两件小事没

什么重要的，但先前做出的承诺对尼克劳斯来说却很重要。但是，为什么尼克劳斯先生这么……嗯，一诺千金呢？是性格上的什么特点让他如此一板一眼地言出必行吗？

没错，的确有两个特点：第一，他65岁了；第二，他是美国人。

年龄 不足为奇，在态度和行为上倾向于强烈保持一致的人，最容易成为一致性策略的受害者。

事实上，我做了一项研究，用一张量表来衡量人们的一致性偏好。我发现，要是提要求者使用“登门槛”或“抛低球”策略，那些在一致性偏好上得分高的人最容易顺从。更奇怪的是，我们在后续研究中调查了18～80岁的被试，发现一致性偏好是随年龄而逐步增强的，一旦过了50岁，被试对先前做出的所有承诺都会表现出强烈的保持一致的意愿。

我相信这一发现有助于解释65岁的尼克劳斯对先前承诺的执着态度。此时，他若推辞不去，也是谁都能理解的，但为了忠于自己的特质，他需要跟承诺保持一致。我还相信，同一发现也可以解释为什么专门针对老年群体的诈骗犯大多会使用一致性手法来诱捕猎物。

美国退休人员协会对年龄在50岁以上的会员频频上电话诈骗的当感到颇为担忧，于是便进行了一次研究，提供了与此相关的证据。该协会派出调查员到12个州做了调查，试图揭露电话诈骗犯针对老年群体所采用的骗术，他们积累了骗子和受害者之间谈话的大量电话录音带。研究者安东尼·普拉卡尼斯（Anthony Pratkanis）和道格·沙德尔（Doug Shadel）对录音带进行了透彻分析，他们揭穿了骗子最常见的做法：先从目标那里得到一个小小的承诺（有时光是一段声明自述也行），再要目标履行诺言，榨取钱财。

下面我节选的部分录音对话表明，在骗子手中，一致性原理发挥出

了可怕的威力，就像是一根大棍子，把那些偏好保持一致的人打了个晕头转向。

“不，我们不光是在说这个。你预订了的！你说了‘好的’，你说‘订吧’。”

“嗯，你是上个月注册的，你不记得了？”

“你三个星期之前向我们做了保证的。”

“上个星期你对我做了承诺的。”

“你不能买了东西五个星期以后又反悔啊，你不能这么做。”

个人主义 尼克劳斯对言出必行的强烈意愿，除年龄之外，还有另外一个影响因素。我先前给过一点提示：他是个美国人，出生、成长在俄亥俄腹地。美国这个国家跟世界其他地方的最大区别就在于，它信奉“个人主义邪教”。在美国和西欧等个人主义盛行的地方，人们更注重“自我”；而在其他更强调集体主义的地方，人们更注重“群体”。举例来说，个人主义者决定某种情况下自己该干什么，主要是看自己的过往、观点和选择，并不过多考虑同辈的意见。这样一来，要是影响力策略借助了他从前行为的杠杆力量，他们就特别容易上当。

为了验证这一想法，我和同事找来一组大学生，试用了一套“登门槛”手法。

这些学生，一半来自美国，一半来自个人主义不那么盛行的亚洲国家。我们先要学生们花20分钟时间，完成“学校和社会关系”的在线测试。过了一个月，我们又要他们花40分钟时间完成同一主题的另一次相关调查。在完成了20分钟问卷调查的学生里，答应做40分钟问卷调查的美国学生差不多是亚洲学生的两倍，前者为21.6%，后者为9.9%。这是为什么呢？因为个人主义者是这样的：答应了之前的同类小要求，就觉

得应该根据这一点做后面的事情。

因此，在个人主义盛行的社会，社会成员，尤其是年纪较长的社会成员，必须警惕“先提小要求”的影响力策略。这些要求尽管小，却能叫人盲目地跳下危险的悬崖。[\(46\)](#)

本章小结

- 心理学家早就认识到大多数人都有一种让个人的言论、信仰、态度和行为保持一致的欲望。这种一致性倾向有三大源头：第一，社会高度重视良好的个人一致性；第二，除了对公众形象的影响，整体保持一致能为日常生活带来有利的应对方法；第三，面对当今复杂的社会，方向一致提供了宝贵的行为捷径，只要跟先前的决定保持一致，人们以后碰到类似的环境就不用处理所有相关信息了，只需要想一想先前的决定，再做出一致的反应就行了。
- 在说服业界，确保初步承诺最为关键。人做出了承诺，即站定了立场之后，会更乐意答应跟先前承诺一致的要求。因此，许多说服专业人士会设法诱使人们预先站到一个立场上，这个立场跟他们之后提出的要求、要人们做的事情是一致的。不过，并不是所有的承诺都能同样有效地带来前后一致的行动。主动、公开、付出努力、发自内心、没人强迫的承诺最为有效，这是因为，这些元素里的每一个都能改变自我形象。而它们之所以能改变自我形象，是因为每个元素都在向我们透露必须真正相信的信息。
- 做的承诺哪怕是错误的，也有一种自我延续的倾向，因为它们能“自己长出腿来”。也就是说，人们往往会寻找新的理由和借口，证明自己已经做下的承诺很明智。因此，即

使刺激人们做出承诺的条件发生了变化，承诺也能长久地维持下去。这种现象解释了为什么诸如“抛低球”等欺骗性顺从手法会有效。

- 基于承诺的手法还有一个优点：简单地提醒人们先前的承诺，可以重新带来指导行为的效果，哪怕情况已经跟此前完全不同。此外，提醒不仅能让承诺复活，还能通过巩固相关的自我形象来强化承诺。
- 要想发现并抵挡一致性压力对决定造成的不当影响，我们应当倾听来自身体两个部位，即肠胃和心灵发出的信号。当我们意识到承诺与一致的压力在迫使自己答应本来不乐意答应的要求时，肠胃就会发出信号。此时，最好是向提要求的人说明，要是我们答应了他们的要求，那就是在死脑筋地保持一致，我们不乐意这么干。而心灵深处发出的信号有所不同，它们最好是用在当我们搞不清最初的承诺是否正确的时候。这里，我们应该向自己提出一个关键的问题：知道了现在掌握的情况，倘若我能回头，还会做出同样的承诺吗？答案就藏在最先冒出来的灵光里。在个人主义盛行的社会，特别是对年纪在50岁以上的人来说，承诺与一致手法最为管用。

Intelligence

第 8 章

联盟

如果说，我们无法实现和平，那是因为我们忘记了自己属于彼此。

——特蕾莎修女 (Mother Teresa)

NEW AND EXPANDED



或许很多人都曾遇到过一位不走寻常路的室友，他的举止活动既让我们感到不安、困惑，也让我们对人类的能力范围有了全新认识。但没有人像人类学家罗纳德·科恩（Ronald Cohen）过去的室友那样，在这两方面都给人留下了不可磨灭的印象。

在一次深夜谈话中，这名曾经在纳粹集中营当过警卫的男子描述了一件让他自己和科恩都难以忘怀的事情。多年以后，科恩仍为室友讲述的故事感到困扰，还围绕它写出了一篇学术文章。

在纳粹集中营里，要是有一名囚犯违规，守卫常常会让所有人排队报数，从每10个人中杀1个人。在科恩的故事里，一个资深守卫跟往常一样做着这件事，可突然之间，他做了个奇怪的选择：倒霉的第10号囚犯走出了队伍，守卫却扬了扬眉毛，朝第11号囚犯开了枪。

稍后，我将揭示这名守卫一念之间定生死的原因。不过，为了圆满地达到这一目的，我们有必要先来看一条在人们心中根深蒂固的社会影响力原则，它才是该原因背后的力量来源。

“我们”的力量

几乎人人都会不假思索地立刻把别人分成属于“我们”和不属于“我们”的不同群体。这对影响力有着强烈的暗示，因为在“我们”这个群体之内，任何与影响力相关的事情都更容易实现。那些属于“我们”群体内的人，会得到更多的认同、信任、帮助、喜好、合作、情感支持、宽容，甚至获得更美好的评判：比如更具创造力、更有道德心、更富人情味。这种偏爱似乎不仅对人类行为有着广泛的影响力，还有着漫长的历史根源：婴儿很早就展现出了这种偏爱，连其他灵长类动物也不例外。“我们”这种部落意识，只要按下按钮就会被唤醒。⁽⁴⁷⁾

因此，成功的社会影响力，往往以“我们”式关系为基础。但这仍然

存在一个核心问题：怎样描述这种关系最为准确呢？答案在于一种微妙而关键的区分。“我们”式的关系，不是要让人们说“嘿，那个人跟我们很像”，而是让人们说“嘿，那个人是我们中的一员”。影响力的联盟原则可以这样表述：人们倾向于顺从他们认为属于自己群体的人。联盟并不是说人与人之间存在简单的相似之处，而是指彼此有着共同的身份认同。这是一种部落式的分类，人们用来界定自己和自己所属群体，比如民族、种族、国籍、家庭以及政治和宗教背景。举个例子，和同胞兄弟姐妹比起来，我跟同事在品位和偏好上有更多的共同点，但毫无疑问，我只会把同胞兄弟姐妹视为“自家人”，而同事只不过是“跟我很像的人”。这些类别的一个关键特点是，其成员往往会感到跟其他人合而为一。他们中一个人的行为，会影响其他成员的自尊心。简单地说，“我们”，就是扩大的“我”。

因此，在“我们”式关系群体里，人们往往无法正确区分自己的特点与群体内其他成员的特点，体现了自我与他人的混淆。神经科学家为这种混淆提供了一个解释：自我和亲密他人的概念是由相同的大脑回路进行心理表征的，激活其一，就可导致另一概念的神经元也被交叉激活，从而导致身份认同的模糊。神经科学证据提出很久以前，社会科学家们就在测量自我—他人重叠感以及由此产生的身份认同问题（见图8-1）。凭借这一测量结果，研究者考察了有哪些因素会带来更强烈的身份认同感，以及这些因素是怎么发挥作用的。

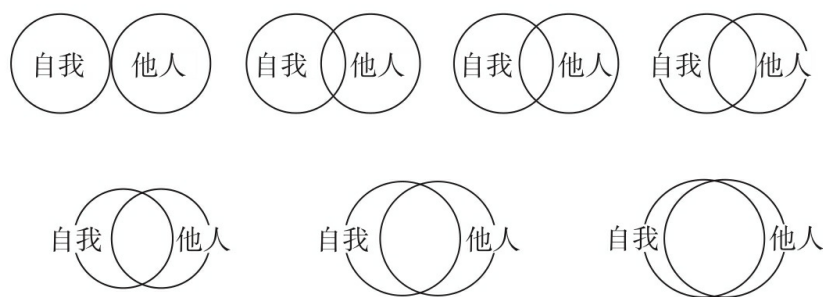


图8-1 重叠的圆圈，重叠的自我

自1992年公布以来，科学家们一直使用《自我涵盖他人量表》（*Inclusion of Other in the Self Scale*）来观察哪些因素可推动与他人“合而为一”的感觉。

群体成员之间的身份认同混淆，可以从以下倾向看出：将自己的特点投射到其他群体成员身上，记不清楚自己此前评价的特点是属于自己还是同群体成员的，需要较长时间才能确认自己和群体内成员之间的不同特点。

其他类型的认知混淆也是因为大脑在做不同的事时使用了相同的结构和机制。举例来说，一个人如果反复想象自己做了某事，就会逐渐相信自己真的做了某事，这一现象可以被部分解释为：执行一个行为和想象执行该行为，会使用相同的脑区。再举一个例子，在大脑中，感受社会排斥带来的伤痛跟感受身体的伤痛一样，使用的是同样的脑区。因此，某些感冒药可同时减少两者带来的不适感。

“我们”式关系对人类反应的影响，有着惊人多样化的适用环境和背景。但有三个常量是反复出现的。第一，建立在“我们”基础上的群体，其成员对群体内成员所带来的成果和幸福的偏爱，远甚于不属于本群体的成员所带来的成果和幸福。例如，互相竞争的工作小组的成员（每一小组包括两名人类队友和两个机器人）不光对自己的人类队友有着更积极的态度，对本团队中机器人的态度也比对手团队的机器人（及人类队员）更积极！第二，“我们”式群体的成员，极有可能用其他成员的偏好和行为来指导自己，以确保群体团结。第三，从进化的角度来看，这些偏爱和追随的抱团冲动，是为了让我们的群体获得优势，并最终让自己获得优势。事实上，一群学者回顾了数十年来有关这一点的相关科学研究之后得出结论：群体主义不光普遍存在，它还是人类的本性。稍微留意，就能发现社会环境中这种偏袒是多么普遍而强大，并且大多数“按一下就播放”。

商业

销售 还记得第3章提到的世界上“最伟大的汽车推销员”吉拉德所取得的惊人销售成绩吗？他在连续12年中，平均每天卖出5辆汽车。他是靠着善结客源做到这一点的，因为他也真心喜欢自己的客户。吉拉德经常给客户发去“我喜欢你”的明信片，总是确保他们购买汽车时得到快速而礼貌的服务，并一直给他们报出公平的售价。最近，有报道指出，吉拉德的销售数据已经被密歇根州迪尔伯恩汽车销售员阿里·雷扎（Ali Reda）所取代，后者的年销售量比吉拉德成绩最好的年份还要高（见图8-2）。雷扎先生在采访中承认，他密切关注了吉拉德的具体成功建议。但如果雷扎只是在模仿吉拉德，那他怎么可能超得过师父呢？他一定在“菜谱”里加入了一种与众不同的秘密成分。他的确加了，但那并不是什么秘密，而是从种族方面彻底地体现了“我们”式关系的力量。



图8-2 雷扎就位

雷扎是密歇根州迪尔伯恩阿拉伯社区的老面孔，他在此地创下了汽车销量的历史新纪录。

毫不意外的是，吉拉德的支持者对雷扎创下销售纪录的说法提出了质疑。不过，雷扎的销售经理支持了这一说法，因为他查看的经销记录与雷扎的说法吻合。

迪尔伯恩大约有10万人口，有着全美最多的阿拉伯后裔。雷扎本人就是阿拉伯裔美国人，并致力于成为阿拉伯社区（这是一个人际关系联系极为紧密的社区）里引人注目的活跃成员，他极力向这个群体推销自己。大部分客户来找雷扎，是因为他们知道他是自己中的一员，愿意信任他。如果从“种族”这个层面来划分“我们”，雷扎就彻底打败了吉拉德。吉拉德出生时的名字是吉拉迪，这暗示了他的西西里血统，所以在他的大多数顾客眼中，他不是“我们中的一员”。实际上，吉拉德说自己不得不改名换姓，恰恰是因为当时有些顾客不愿意跟“外国佬”做生意。

科学研究证实了共同的“我们”身份能对销售结果带来有利影响：如果潜在客户跟未来的教练出生在同一个社区，那么前者更愿意接受个人培训计划的销售游说。同样，如果客户知道自己和将要去看的牙医的出生地相同，那么牙科服务套餐对他的推销会更成功。

金融交易 如果共同的种族身份有助于解释雷扎为何能在紧跟吉拉德方法的同时超越吉拉德的绩效，或许，同样的因素还能解释另一个商业上的谜团。显而易见，我们这个时代最大的投资骗局要数华尔街业内人士伯纳德·麦道夫（Bernard Madoff）精心策划的庞氏骗局。分析师们聚焦于这起诈骗案意义深远的一些方面，比如其规模庞大（超过150亿美元）、持续时间很长（数十年都无人察觉），但另一个明显的特点给我留下了深刻的印象：许多受害者都是有着丰富金融知识的人，其中不乏头脑冷静的经济学家、经验丰富的基金经理、极为成功的商业领导者。事实上，从麦道夫哪怕在经济不景气的年份交付的利润也异常丰厚这一点而言，人们就应该很快对他产生疑心了。麦道夫一案不是“狐狸骗了鸡”，而是“老狐狸骗了狐狸”。

他是怎么做到的呢？人类反应中的重大事件几乎从不是由单一因素导致的，它们无一例外均由多种因素共同造就。麦道夫事件也不例外。他在华尔街的长期经历、他声称采用的基于衍生品的金融机制的复杂性、“获允”加入他基金的投资者圈子的有限性，这些因素都起到了一定

作用。但除此之外还有一个积极的因素，那就是共同的身份。麦道夫是犹太人，他的大多数受害者也是犹太人，他们往往是麦道夫同为犹太人的助手寻觅来的。另外，新加入的人和过去加入的人互相认识，又是同一种族，这些都充当了社会证明：参加麦道夫的投资必定是个明智选择。

当然，这类欺诈并不仅限于同一种族或宗教群体。这种被称为“杀熟”的投资骗局，几乎总是和群体内的成员对群体内其他成员的掠夺——拉丁人欺骗拉丁人，亚美尼亚裔美国人欺骗亚美尼亚裔美国人。查尔斯·庞兹（Charles Ponzi）是一名意大利移民，他在1919至1920年从其他意大利移民手里骗取了数百万美元。麦道夫的庞氏骗局，就得名自庞兹。

除了投资决策，建立在“我们”基础上的选择，在其他金融交易里也普遍存在。在美国金融咨询公司内部，如果一名顾问和另一顾问种族相同，那么，要是前者存在财务不当行为，后者效法的可能性会是一般情况的两倍。一项研究考察了印度一家大银行的记录，发现信贷员对来自同一教派的申请人批准了更多的贷款申请，给予了更为优惠的条件。而且，这种偏袒有可能双向发挥作用：如果信贷员和贷款人来自同一宗教，那么贷款人的贷款偿还率会显著增加。这里还有一个群体内偏袒的例子：在一家餐厅里，当顾客受到不良的服务待遇时，如果服务员与顾客姓氏相同，顾客就不太愿意责怪他。

如果说，这些来自世界各地的研究仍不足以证明群体内效应的跨文化影响，那就让我们再来看最后一个例子。在加纳，出租车司机和乘客通常会在出行前就车费进行协商。如果这两个讨价还价的人支持同一个政党，司机往往会以更低的车价接单——但这里有一个有趣的转折。价格让步仅发生在选举前后的几个星期，在这段时间里，选民对自己所属的党派更为看重。这一发现说明了“我们”式群体反应的一个重要方面：它会因突出群体身份的线索或环境而得到强化。因此，联盟或其他任何

影响力原则的吸引力，不同于随时都释放吸引力的普通磁铁，而更类似电磁铁，它的吸引力受当下输入的电流强度的调节。

政治

近年来，在旨在保护他人感情而说的无伤大雅的“白色谎言”（“不，这件衣服 / 这种发型 / 这枚鼻环真的很适合你”）和旨在伤害他人利益而说的“黑色谎言”（“要是你穿着它去跟我前男友约会，他会很高兴的”）之间，又出现了一种新的分类——“蓝色谎言”。这种谎言有两种核心元素：既要保护他人，也要伤害他人；但不同的“我们”式群体，选择保护的对象和选择伤害的对象不一样。它们是故意讲述的谎言，通常是群体内成员为保护自己群体的声誉而对群体外成员说的谎。在这些身份融合的群体内部，联盟胜过真相。用政治色彩不太浓的一种方式来说就是，在群体成员的眼中，旨在强化“我们”式群体的欺骗，会在道德优于削弱本群体的真话。

与“蓝色谎言”类似的一项研究发现，如果人们对某个群体有着强烈的身份认同感，就会更乐于遵循该群体的规范，哪怕他们知道这些规范与现实脱节。研究表明，对所属党派有着强烈认同感的人，更愿意包庇、偏爱和追随同党派成员，这个结论与新兴的学术研究结论吻合，即党派成员的许多政治决定与其说是基于意识形态，不如说是基于对此类定义了身份认同的党派及其成员的忠诚。科学还对对此类偏见产生的心理神经学机制做了有益的总结。

在基于“我们”的逻辑中，群体内成员更愿意掩盖其同伴的过错，这样的逻辑并不只适用于政党。在受到警察质询时，首先，人们会强烈地偏向于不报告好友或家人等与自己联系紧密之人的有害行为；其次，当有害行为程度严重而非轻微（如盗窃之于非法音乐下载，肢体性骚扰之于目光骚扰）时，人们尤其不愿意报告；最后，人们承认，这种不愿意是出于保护自己声誉的心理。我们看到，“我们”真实暗示着“我”。

在这个问题上，政党的表现更为严重。有人回顾了相关研究后总结道：“这种（出于政治利益的）谎言似乎正在愤怒、怨恨和两极分化的氛围中蓬勃发展。对党派的认同太过强烈，一听到他人对党派的批评，就感觉像是对自我的威胁，进而引发一连串的防御心理机制。”听上去耳熟吗？这种狂热的党派认同，不光认可那些促进、保护自己党派的谎言，还会触发其他的防御机制。身份认同跟所属政党“融合”的人报告说，他们更愿意向政党隐瞒党内政客税务欺诈的证据。哪怕已经看到证据表明，两党对城市幸福做出了同等投入，狂热的党派成员仍会相信自己所属党派的贡献更大。研究者请人们将肾病患者候选名单排序，以判断其是否有资格接受新一轮治疗。结果显示，人们会选择与自己党派一致的人。

人们不光会偏爱自己所属党派的成员，哪怕在令人困惑的情况下，也会选择更相信同党派成员。在一项在线研究中，研究者向被试展示了若干种体型，并请他们根据一套指导方针对体型分类。划分到正确类型下的体型越多，被试得到的钱就越多。在判断怎样最好地将体型分类时，被试可以选择参考另一名被试（根据研究的设定，前者可以从此前的信息里得知后者的政治倾向）的答案。

令人惊讶的是，被试会选择参考并照搬跟自己政治立场相似的人的答案，哪怕后者在这项任务中的表现相对糟糕。想想看：人们竟然更愿意在一项任务中求助于政治盟友的判断，然而，该任务与政治无关，盟友在任务中表现也较差，而且被试有可能因此受到金钱损失！总的来说，这些发现与新兴的学术研究结论吻合，即党派成员的许多决定并不是基于意识形态，是基于忠诚，而忠诚则来源于“我们”。

体育运动

几个世纪以来，体育比赛的组织者充分认识到，偏袒“自己人”是很自然的倾向，所以必须依靠独立的评判者（裁判）来维护规则，以不偏

不倚的态度评出获胜者。但对这些裁判员，我们能期待他们做到多公正呢？毕竟，如果说“部落主义是人类的本性”，那我们真的能合理地相信他们能做到不偏不倚吗？既然已经知道自己对“自己人”的偏爱，我们就应当对此保持怀疑的态度。此外，还有直接的科学证据能支持这种怀疑的合理性。

在国际足球比赛中，倘若球员来自裁判的祖国，那前者往往会多获得10%有利于自己一方的判罚。不管是精英裁判，还是经验不足的新手，都会产生这种偏向。在美国职业棒球大联盟赛事中，裁判和投手之间是不是同一种族，会影响到投出的球是否被判断为“好球”。在美国职业篮球比赛中，裁判对跟自己同一种族的球员判犯规的次数较少。研究者得出结论：由于上述偏差的程度太过严重，“球队取胜的概率，明显会受比赛执裁团队种族构成的影响”。因此，哪怕是经过专门挑选和训练要消除偏差的人所做出的判断，也会受到“我们”式群体偏差的影响。要理解为什么会出现这样的情况，我们必须了解影响体育裁判人员的力量，其实也就是那些能让球迷一边倒的声名狼藉的力量。

著名作家艾萨克·阿西莫夫（Isaac Asimov）描述过我们观看比赛时的反应：

倘若其他的条件全都一样，你铁定会支持跟自己同样性别、来自同一文化、同一地区的队伍……你想要证明自己比另一个人更优秀。你支持的一方就代表了你，他（她）赢了，你就赢了。

从这个角度来看，体育迷的狂热就变得有意义起来。我们观看比赛，并不是因为它固有的表现形式或艺术意义，而是我们投入了自我。这就是为什么球队主场获胜以后，粉丝们会报以那么强烈的崇拜和感激之情。这也是为什么球队主场失利之后，同一批粉丝会马上翻脸不认人，简直恨不得把球员、教练和官员生吞活剥了。在发表上述引言的

同一篇文章中，阿西莫夫还评论了美国各州对当年美国小姐选美比赛候选人的过度偏差。

我很喜欢的一段逸事，就是很恰当的例证。

影响力研究

第二次世界大战结束后，一名老兵退伍回到巴尔干的家乡，没过多久就不说话了。医生给他做了健康检查，却找不出毛病来。没有伤口，没有脑损伤，声带也没有受损。他能读、能写、能理解对话、能服从命令，但就是不说话。不跟医生说，不跟朋友说，也不跟苦苦恳求的家人说。

医生困惑又生气，于是把他转移到了另一座城市，安置到了一家退伍军人医院。在那里，老兵待了整整30年，从来不曾张口说话，过着与世隔绝的生活。后来有一天，他病房里的收音机刚好调在了一个转播足球比赛的波段，当时那场比赛又正好是他家乡的球队跟老对头打。在比赛的关键时刻，裁判判沉默老兵家乡球队的球员犯规，老兵气得从椅子上跳起来，瞪着收音机，30年来头一回开了口。“你这个蠢蛋！”他大叫道，“你是想让他们赢比赛吗？”说罢，他又坐回了椅子，重新回到了一贯的沉默当中，再也没开口。

从这个真实的故事里，我们可以了解到两件重要的事情。一是体育运动蕴含着惊人的力量。老兵希望家乡球队获胜的欲望是如此强烈，光是这一点冲动，就让他打破了自己多年来顽固坚持的生活方式。二是体育运动和粉丝之间的关系是非常个人化的。不管沉默老兵的个人认同残缺到了什么程度，足球比赛都能让他感同身受。经过病号房里30年的无声自我放逐，不管他的自我已经虚弱到了何等程度，比赛结果仍然牵动着他的心弦。为什么会这样呢？因为要是家乡队输了，他的自我会更加

消沉；要是家乡队赢了，他的自我则会提升。何以如此呢？是关联原理在搞鬼。他与故乡的关系，把他跟一场球赛的胜利或失败捆在了一起、包在了一块儿、系在了一处。

关于体育领域非理性群体内偏好的案例，我还可以举一个我本人的例子。我在威斯康星州长大，当地全美橄榄球联盟的主场球队一直是绿湾包装工队。不久前，我读到一篇介绍不同名人最喜欢哪支橄榄球队的新闻，得知艺人贾斯汀·汀布莱克（Justin Timberlake）和说唱乐手李尔·韦恩（Lil Wayne）都是绿湾包装工队的狂热球迷，我立刻改变了对两人音乐的看法。不止如此，我还希望他们未来取得更大的成功。沉默的二战老兵跟我在很多方面都不同（比如，没人会求着我开口说话），但从从不假思索地偏爱“自己人”这方面来看，我们是类似的。别否认了，没用的。

个人关系

情侣 所有的恋爱关系都会经历分歧，而如果不处理产生冲突的来源，就会导致不和与不满，同时损害双方的心理和身体健康。有没有什么特别有效的影响力方法，可以用来说服另一方做出改变，进而减少分歧呢？有，而且还易于实施。在一项研究中，在一起的时间平均已达21个月的夫妻答应讨论两人关系中一直存在的问题，并试图找到解决办法。研究者注意到，这种交流有两个重要的方面。第一，总是有一方在扮演说服者的角色，试图把另一方转移到自己的立场上。第二，说服者施加影响的方法分为3种形式，带来的结果完全不同。

第一种是高压式方法，依靠的是贬低性的评语以及“你最好做些改变，要不你会后悔的”之类的威胁。这种攻击不仅不会取得成功，反而会适得其反，让接收方进一步远离说服方的立场。第二种是“梳理事实或讲逻辑”的方法，说服方声称自己的立场在理性上更为优越，并使用“只要想一想就知道我是对的”这样的说辞。此时，接收方会简单地驳

斥主张，但完全不会改变。第三种是“伙伴关系提升”法，即在意识上突出两人的个体身份已融合为“伴侣”，它的成功率极高。说服方只需要使用“我们”或“我们的”这样的措辞，并提及两人共同的感受、在一起度过的时光（比如，“你知道，我们在一起很久了，我们彼此关心；如果你能为我们这么做，我会很感激”），就能达到理想的效果。研究结果表明，只有这一类说服者在另一方身上实现了自己想要的改变。

这种提升“联盟”意识的方法，不光证明了它的有效性，还有另外两个值得注意的特点。第一，它的功能本质是“证据推论失效”（*evidentiary non sequitur*）的一种形式。“你知道，我们在一起很久了，我们彼此关心”这句话不足以建立沟通者立场的逻辑或实证效用。相反，它为改变提供了一个完全不同的理由——对两人关系的忠诚。

第二，它并未提供任何未知的东西。一般而言，双方都很清楚自己处在伴侣关系当中。但这一含义丰富的信息，如果碰到正在争夺同一空间的其他因素，就很容易在双方的意识中退居其次。一如“证据推论失效”这个名字，伴侣关系提升法只是在提高双方对彼此关系的觉察。这种改变的建立基础，非常符合我最近看待许多社会影响力研究的方法。最可能指导一个人决策的因素，并不是全局中最有力或最具指导性的方面，反而是人们在做决定时意识里最突出的那个方面。

亲密友情 除了爱情伴侣，“我们”式关系还可以来自其他有力的个人连接形式，友谊便是其中之一。不足为奇，跟只是普通熟人（如同事）的人相比，朋友之间的体育锻炼活动更有可能互相匹配。

EBOX | 线上影响力

如今，基于友谊的群体经常会通过网络建立起来，这就构成了电子商务活动的一个子集，即“友谊型商务”。按照为各大品牌提供咨询服务的社交媒体软件提供商Awareness所说，“友谊型商务”的利润非

常丰厚。针对梅西百货公司和李维斯这两家传统实体企业所做的“友谊型商务”活动报告说：“梅西百货的‘时尚总监’活动允许用户自己设计一套衣服，并收集朋友们是否愿意购买此套衣服的意见和投票。通过‘时尚总监’活动，梅西百货的Facebook粉丝翻了一倍，达到180万，销售额提高了30%。李维斯的‘朋友商店’则支持用户创建个性化的商店，商店由用户的朋友们喜欢的物品组成。商店功能上线时，一举吸引了超过30 000名粉丝，李维斯的社交网络覆盖面也扩大到了900万名粉丝。‘朋友商店’的销售额提高了15%，平均订单价值提高了50%。”

作者点评：我对李维斯“朋友商店”的例子尤其感到信服，因为它的影响力并不来自朋友们说自己喜欢该店商品的风格，相反，它来自对朋友们当前风格偏好的了解，进而提高了这些风格的物品的购买量。

富有教育意义的一点是，友谊（以及伴随而来的联盟感）越紧密，朋友的行为对我们的影响就越强。在一项有6 100万人参与的大规模政治选举实验中，如果一条Facebook信息中包含已经投票的Facebook好友的照片，还有一张照片来自一名亲密朋友（这一点最为关键），那么，敦促人们投票的信息效力最强。

最后，最好的朋友之间存在一种比亲密朋友更强烈的联盟感。“我们是闺蜜”或者“我们永远都是兄弟”等特别的标签和主张便传达了这种纽带关系的力量。在一项关于大学生饮酒行为的研究中，一名学生每星期的酒精摄入量、饮酒频率以及酒精相关问题，均与最好朋友的程度一致。

表明朋友之间身体活动水平联系的研究还发现，被试低估了朋友对自己活动产生的影响，错误地认为与健康及个人形象相关的因素对自己影响更大。最好的朋友对大学生的饮酒情况有着强有力的影响，该研究

还表明，这一效应在白人学生和美国原住民学生身上都存在。一般而言，朋友之间可见的与实际的基因重叠程序，都比非朋友之间的重叠程度高。

宠物 人人都会打呵欠，通常是出于疲倦或无聊等原因。就此处的目的而言，导致打呵欠的另一个更有趣的原因与影响力过程有关：这就是传染性呵欠，即别人打呵欠之后自己也打起呵欠来。我们知道，联盟感会给人类的反应带来影响，因此，传染性呵欠的发生频率，跟后一个打呵欠的人与前一个打呵欠的人之间的个人依恋程度直接挂钩。传染性呵欠最容易发生在亲人之间，其次是朋友，然后是熟人，陌生人之间最少。类似的情形也发生在其他物种中（黑猩猩、狒狒、倭黑猩猩和狼），一只动物打呵欠主要是由亲属或朋友的接触引起的。

我们知道，同一物种（并主要是以“我们”为基础单位的物种）的成员之间会发生传染性呵欠的现象。这种影响力能跨物种存在吗？日本的一项研究给出了肯定的答案，而且证据能让人大吃一惊。该研究中的物种之一是人类，之二是狗（经常有人说狗是“人类最好的朋友”）。实际上，还常常有人形容人和狗之间的纽带超越了友谊甚至血缘。例如，经常会有人把狗当成自己的家庭成员，比如“我养了3个孩子和1只苏格兰狗”。

该项研究使用类似的程序检验了25只狗。在5分钟内，让每只狗看着研究者或自己的主人打了几次呵欠。研究者记录下狗的反应，并分析传染性呵欠的次数。结果一目了然：跨物种的传染性呵欠确实存在，但只存在于狗与主人之间。我们再一次看到，在以“我们”为基础的单位（单位的边界可扩展余地极大，在本例中甚至包括了跨物种成员）内，影响力尝试会更为成功。

研究者收集的数据还揭示了人类传染性呵欠发生的次数，与传染和被传染呵欠者之间个人连接程度的等比关系。他们发现，就算呵欠只通过声音传播，也存在同样的比例关系（如见图8-3）。



图8-3 会传染的呵欠

宠物及其主人打哈欠会互相传染。到目前为止，研究者只验证了主人打呵欠会传染给宠物。我不是个好打赌的人，但我敢打赌，这种现象是双向的。

各位爱猫人士，不要绝望。我没能提供表明宠物猫与其主人之间传染性打呵欠的数据，但这并不意味着这种影响不存在。缺乏证据可能只是因为研究者还没有检验这种可能性，这或许是因为研究者很难让猫长时间保持静止、集中注意力。

除了商业、政治、体育和人际关系，人类互动的其他重要领域也显示出“我们”式群体身份的偏见效应，而且偏误程度同样惊人。在健康方面，如果主治医生与新生儿属于同一种族，婴儿出生死亡率就会显著下降。在执法领域，如果警察和司机的种族相似，警察就不太可能拦下司机的车辆进行搜查。在教育领域，教师打分也表现出了同样的影响：教师与学生在种族、宗教、性别、民族或国籍方面的匹配，会提高学生的课堂评估水平和考试成绩。荷兰马斯特里赫特大学的一项研究提供了最明显的证据：该大学位于荷兰与德国边境附近，拥有大量来自荷兰和德国的师生。把学生的试卷随机分配给来自相同或不同国籍的老师评分，如果学生的名字跟打分老师的国籍匹配，那么该学生的分数就会更高。

很明显，行为科学家一直忙于检验联盟原则影响人类反应的广度和

深度。在这个过程中，他们发现了能带来联盟感的两类主要因素：一类是身心合一，一类是行动合一。我们来逐一看看。

联盟1：身心合一

唤起亲人意识

从遗传学的角度来看，属于同一个家庭，即血缘相同，是自我—他人统一的终极形式。事实上，进化生物学中被广为接受的一个观点是，个体会付出极大的努力，争取让自己的基因延续下去，而非单纯地让自己存活。这种主张的言外之意是：自我利益可以在当事人自身之外，在与自己有着相同遗传物质的他人身上。出于这个原因，人们特别愿意帮助血缘关系近的亲戚，尤其是在事关生死的决定上，比如美国人是否愿意捐献自己的肾脏，日本人是否愿意救出着火大楼里的人，委内瑞拉人是否愿意在丛林斧头大战里出手干预。脑成像研究给出了一个最可能的解释：在帮助了家庭成员之后，大脑的自我奖赏中枢会体验到高得非同寻常的刺激；这么做，就好像是帮助了自己一样，哪怕是对青少年，情况也是这样！

READER'S REPORT | 读者报告

致西奥迪尼先生：

最近，我进一家商店购买必需品时，在门口使用了保安提供的洗手液。我注意到，一个在该商场药店工作的人，进店时拒绝使用消毒液。这样的场面并非个例。比如，商场里有更多人都接受保持社交距离的规劝。

事后，我致电商场经理指出了这一现象，她说自己并未获得授权做出任何改变，但声称会向“公司”提出这个问

题，然而情况并无改观。接着，我联系了当地的国会议员，并给议员留下一条电话信息：“议员先生，请想象一下你的奶奶或妻子因为传染控制措施实施不到位而患病。良好的传染控制本可以让她们免遭此劫。请将这一想法传递给更多的人。”

两天后，我接到了议员的电话和电子邮件。他用我的方案联系了卫生部、卫生部长和两家全国零售连锁店的首席执行官。随后，我在新闻中发现，零售连锁店突然实施了新的手部消毒和社交距离限制。该新闻鼓励人们联系推动这一改变的议员。我想，是我促成了这种改变，虽然功劳落到了这位议员头上，但我并不在乎。

2020年新冠疫情肆虐期间悉尼的这名护士：

虽然很难知道到底是哪些因素导致了你所目睹的变化，但我怀疑，你在游说议员时动情地提到了对方的家人，还建议议员也采用这一影响力手法，这发挥了一定的作用（因你要求隐去自己的姓名，所以本书前言部分所列的“读者报告”撰稿人名单中并未收录你的名字）。

影响力研究

从进化的角度看，任何能让亲人得到优势的事都值得一做，哪怕是很小的优势。我在职业生涯里曾使用过一次非常有效的影响力技术，证实了这一点。

有一次，我想对比大学生及其父母对待一系列问题的态度有什么不同，于是安排这两个群体填写同一份冗长的问

卷。找一群大学生来完成这桩任务并不难，我在自己执教的心理学大课上把问卷当成课堂练习发了下去。但想办法让学生的父母也完成问卷就困难多了，因为我没有研究经费，而且成年人参与这类调查的比例很低，通常低于20%。一位同事建议我打打“亲人牌”试试看：如果学生的家长愿意完成问卷，在下一次随堂考试中，学生就能多得一分。

这个方法的效果好得惊人。全班一共有163名学生把问卷寄给了家长，159人（97%）在一个星期内寄回了填好的问卷，就为了让孩子在某一学期某一门课的其中一场考试上多得一分。身为影响力研究者，我还从来没碰到过这样的事。可从之后的经验来看，这事儿还可以做得更好：我该让学生把问卷发给祖父母。我估计，这样一来，163份问卷在一个星期里能返回162份。而且，缺的那一份还很可能是因为爷爷填完问卷一路小跑去邮局的路上心脏病发作了。

这类来自爷爷奶奶对孙辈的宠爱，我从幽默作家乔尔·斯泰因（Joel Stein）讲的一个故事里获得了部分证明。斯泰因想说服奶奶投票给某位总统候选人。一开始，奶奶对此并不热心。于是，斯泰因对奶奶展开了长篇大论的说服工作，然而，要么是他的论点缺乏说服力，要么是“奶奶”听不懂。但奶奶还是说，自己会去给那位候选人投票。斯泰因迷惑地追问原因，奶奶解释说，那是她孙子希望她做的事情呀。

进化思想的主要观点（即个体会付出极大的努力，争取让自己的基因延续下去，而非单纯地让自己存活）最初特指“广义适合度”这一概念。虽然多次有人提出反对意见，但它仍不断得到证据支持。从基因重叠的角度看，亲属关系越近（如父母或兄弟姐妹相较于叔叔或表亲），自我—他人重叠的感觉就越强烈。另一项研究为群体促进效应提供了解释：在意识里突出强调群体身份，会让人们强烈关注与这一身

份吻合的信息，这反过来也让人们认为这些信息更加重要，更有因果意义。研究者在我班级的心理课上说，如果家长回答一套47道题目的问卷，学生就可多得一分，结果96%的问卷都被完成并交回了。研究者还对“养育后代是范围更广的帮助行为的基础”这一观点进行了详细的分析。

虽然生物学家、经济学家、人类学家、社会学家和心理学家都通过研究了解到了这一点，但要意识到孩子对家长的巨大影响力，人们用不着非要先去当科学家。举例来说，小说家经常表现这种情感上的强大拉力。当代最伟大的小说家之一海明威讲过一个赌注的故事。海明威素以文字简练却动人心弦著称。他跟一位编辑在酒吧喝酒，海明威打赌说，自己能用6个单词写出一个人人能懂并产生深刻共鸣的完整故事。如果读完这个故事之后，编辑认同他的观点，就给整个酒吧的人买杯酒；如果不认同，那就由海明威来埋单。定好打赌的条件以后，海明威在餐巾纸背面写了几个字，把它给编辑看了，编辑什么话也没说，站起身就去了柜台，给所有人买了一轮酒。海明威写的是：“出售：婴儿鞋，未穿过。”

但对跟我们没有特殊基因联系的人，能不能借用亲人的影响力呢？使用能先发制人地在意识里唤起亲人概念的语言和意象，或许是一种办法。例如，使用有关家庭的图像和标签，如兄弟、姐妹、祖先、祖国、血脉，在成员里创造出集体感，能提升人为了集体利益牺牲个人利益的意愿。一支国际研究团队发现，人极其擅长感受象征的力量，因此这些想象出来的“虚构家庭”能提高人的自我牺牲精神。一般而言，自我牺牲精神是与血缘关系极近的氏族相关的（见图8-4）。西班牙曾有两项研究，用“同胞如家人”的概念唤起人们对本国人民的“联盟感”，立刻大幅提升了受访者愿为西班牙战斗至死的意愿。



图8-4 家庭第一

家庭关系的重要性，不仅体现在长辈对待子女的行为上，也体现在子女对待长辈的行为上。2016年，演员茱莉亚·路易斯-德瑞弗斯因电视剧《副总统》获得了艾美奖的最佳女主角称号。在颁奖仪式上，她把这份荣誉献给了自己刚刚去世的父亲，她的致辞生动地体现了父女关系的重要性：“我很高兴他喜欢《副总统》，因为对我来说唯一重要的就是他的看法。”

对我们现有集体之外的人，也可以采用这种方式吗？一个在基因上与对方无关的沟通者，能借助亲人概念获得顺从吗？我在金融服务公司年会上做讲演的时候，有时候会这么问：“在各位看来，我们时代最成功的金融投资家是谁？”人们一定会齐声给出答案：沃伦·巴菲特。自1965年以来，巴菲特和合作伙伴查理·芒格联手，让伯克希尔-哈撒韦公司为股东们创造了惊人的价值。

几年前，我收到一份礼物，是伯克希尔-哈撒韦公司的股票。这份礼物的意义不仅仅在于金钱，它为我提供了一个观察巴菲特和芒格的战略投资及沟通方法的黄金视角。对战略投资，我一无所知；但对战略沟通，我略知一二。在这个过程中，我所看到的技巧让我深为震撼。

具有讽刺意味的是，伯克希尔-哈撒韦公司的财务造诣超凡脱俗，反倒带来了沟通上的问题：怎样让人们相信，公司将来还会继续保持这

样的成功？如果没有这种信心，股东就会卖掉自己的股份，而潜在客户则会去购买其他公司的股票。

毫无疑问，伯克希尔-哈撒韦公司有着优秀的基础商业模式，还有着若干独特的规模优势，能为其将来的估值提供令人信服的理由。但令人信服的理由，跟令人信服地提出理由，并不是一回事。巴菲特每年发布公司报告时，都会结合诚实、谦虚和幽默感等方式来做到后一点。2015年2月，他似乎需要一种超乎寻常影响力的东西，因为当时他要发出一封特别的致股东的信，纪念公司成立50周年，概括历年来的业绩，证明公司将来还会继续保持活力。[\(48\)](#)

50周年隐含了一点言外之意，外界对此表示关心颇有一段时间了，网上的评论更是一边倒地苛刻：创办公司已经50年，巴菲特和芒格显然不再年轻，两人都不应继续领导公司，公司的前景和股价都有可能就此下滑。我还记得，刚读完评论，我深感困扰。我手里的股票在巴菲特和芒格管理期间翻了两番，可如果两人都由于年事已高离任，股票还能继续站在高位吗？是不是应该赶在他俩离开之前把股票卖了，拿走赚到的利润呢？

在信中，巴菲特正面谈到了这个问题，具体是在“伯克希尔的下一个50年”一节。他阐述了伯克希尔-哈撒韦公司的成熟商业模式所带来的深远的积极后果，它几乎牢不可破的金融资产护城河，以及公司已经为下任CEO找到了“合适的人选”，待时机成熟就会走马上任。但从说服专业人士的角度来看，我最受打动的是巴菲特在开始这最重要的一节之前所采用的影响力方法。他以自己典型的方式，正面迎击了潜在的弱点，重建人们对他的信赖感：“现在，让我们看看前进的道路。请记住，如果让我在50年前揣度未来会是怎样的情形，我一定错得离谱。”接着，他做了一件我从未见过，也从未听说他曾在公共场合做过的事情。他说：“请记住我上面的提醒，现在我再来告诉你，如果今天我的家人问及公司的未来，我会怎么回答。”

随之而来的是精心构建的论点，为伯克希尔-哈撒韦公司将来的经济健康做支撑：成熟的商业模式，金融资产护城河，对未来CEO的严格筛选。除去对这些优势部分进行可信的论述，巴菲特先发制人地做了一件我认为更具可信性的事情：他说，要像对待家人一样，把这些事情告诉我。基于我对他的认识，我相信这一说法。于是，我再也没有认真想过要把伯克希尔-哈撒韦公司的股票卖掉。在汤姆·克鲁斯出演的电影

《甜心先生》里，有一个令人难忘的瞬间：主人公冲进房间，向已经跟他分居的妻子多萝西打招呼，然后开启了长长的独白，列举了多萝西应该继续和自己在一起的种种理由。在这个过程中，多萝西抬起头来，用一句今天已经广为人知的台词打断了他：“在你说‘你好’的那一刻，我的心就属于你了。”而巴菲特先生在这封信里视我为家人的那一刻，我的心就属于他了。

有趣的是，巴菲特这封纪念公司成立50周年的信迎来了如潮的好评，有一些评论甚至以“巴菲特写出了他有生以来最精彩的致股东的信”“不投资伯克希尔-哈撒韦公司你就是个傻瓜”等为标题，但没有一个人提到巴菲特在论证中娴熟运用的家人逻辑。人们对这一点竟然这么缺乏认识，我倒也不是太吃惊。在态度强硬、看重事实的金融投资世界，默认的做法是把关注点放在有关优点的信息上。诚然，论据中的优点是一种信息。但与此同时，在另一些层面上，有效沟通本身也可以变成至关重要的信息。马歇尔·麦克卢汉告诉我们，媒介即信息；社会认同原则告诉我们，数量大是信息；权威原则告诉我们，信使即信息；现在，联盟概念又告诉我们，自我与他人的融合也是信息。所以我们有必要考虑，除了直接的血缘关系，在环境中还有哪些特点，有助于让人感受到身份的融合。

值得注意的是，这些特点有很多都缘于对血缘关系的强调。显然，没有谁能钻到别人身体里，判断两者有多少共同基因。这就是为什么，人必须本着审慎的态度，依靠对方身上某些可见的与基因重合相关的特点来进行判断——最明显的自然是外表相似了。长相类似所带来的牵引

力，会使得人们结成友谊群体、大学兄弟会或由长相近似的人所组成的棒球队。在家族里，人们最乐于帮助跟自己相像的亲人。在家族之外，人们用面部相似性来判断跟陌生人之间的遗传相关性，事实上这种判断方法相当准确。不过，人们也有可能遭到欺骗，错放好感。有人用数字技术修改了照片上的脸，使之与观察者更相像，结果极大地提高了观察者对照片中人的信任度。如果这张修改后的脸属于一位政治候选人，选民也会更乐意投票给他。说到底，不管是在家里还是家外，人们都会用相似性来判断基因重合，并偏爱与自己重合度更高的人。

除了身体特征和性格上的可比性，人们还以态度相似性为基础来评估遗传相关性，并进而构建小圈子，决定要帮助谁。但在这方面，不是所有的态度都有同样作用：在判断圈内人身份时，对性行为以及自由或保守意识形态等事物的基本宗教和政治态度，似乎有着最强大的力量。它还可以从另一个基于血缘上的理由来看：这一类态度是最有可能通过遗传来传递的，因此最能反映基因上的“我们”。这种高度可遗传的态度也是极难改变的，这可能是人们不愿意改变在自己眼中足以界定自身的立场。

场所

对于判断突出的遗传共性，还有另一条通常情况下靠得住的线索。它不是身体上的相似性，而是物理上的接近性。这就是察觉到跟别人来自同一个地方，对人类行为的影响大得惊人的原因。要说明这种影响，在我看来最好的办法莫过于把目光投向近现代历史上一段最黑暗的岁月，也就是第二次世界大战大屠杀期间，解开在那时浮出水面的一些重大的人类行为之谜。让我们先从一个人所处地点的最小形式开始，再转到更大的形式。

家园 人类和动物都会把成长期间生活在自己家的同类视为亲人。虽然这一认亲的线索偶尔也会造成误导，但通常都很准确，因为待在家

里的人一般就是“自家人”。此外，在家居住时间越长，对人们的家庭感影响就越大，对人们为彼此牺牲的意愿影响也越大。

但有一个相关因素，能带来相同的结果，又用不着花太长时间。人们在观察到父母照料家里其他人的需求时，同样会产生家庭般的感觉，也更乐于向那个人提供帮助。这种过程带来了一个有趣的结果：看到父母打开家门接待形形色色的人，很可能使孩子在成年之后也更愿意帮助陌生人。对这些孩子而言，归属感或许超出了家庭甚至家族，而适用于整个人类。

这一见解对解开大屠杀时期的重大谜题有什么用处呢？按历史的记录，当时最著名、最成功的救援者是：勇气十足的瑞典人拉乌尔·瓦伦贝格（Raoul Wallenberg），他英勇地救出了数千犹太人，自己最终因此丧生；奥斯卡·辛德勒（Oskar Schindler），他的“名单”救下了1 100名犹太人。然而，大屠杀期间最有力、最集中的一次救援行动，却几乎淹没在了历史长河里。

1940年夏季的一天，破晓时分，200名波兰犹太人聚集到立陶宛的日本领事馆门外，恳求帮助，希望能赶在纳粹大军压境之前，从东欧逃出去。为什么他们会选择向日本官员寻求帮助，本身就是一个谜。当时，纳粹德国和日本帝国政府有着紧密的关系、共同的利益。事实上，仅仅几个月之后，在1940年9月，日本、德国和意大利就签署了三方契约，正式宣布结盟。那么，为什么这些犹太人、第三帝国仇恨的目标，会向阿道夫·希特勒的外国伙伴求助呢？他们期望从日本人那里得到什么样的帮助呢？

在20世纪30年代末跟纳粹德国建立密切的战略联盟之前，日本允许流亡的犹太人借道本国领土，以求从国际犹太人社群那里得到一些财政资源和政治上的善意。由于日本国内的部分圈子里对这一计划的支持仍很强劲，政府并未完全取消向欧洲犹太人发放旅游签证的政策。由此带来了一个不可思议的结果：战前几年，包括美国在内的世界大多数国家

都抛弃了希特勒“最终解决方案”中绝望的“猎物”，反而是希特勒的盟友日本为他们提供了庇护：上海日租界以及日本神户市的犹太定居点。

所以，1940年7月，200名犹太人聚集在立陶宛的日本领事馆门外，他们知道大门背后的人手里，恐怕握着他们逃出生天的最后机会。这个人的名字叫杉原千亩（Sugihara Chiune），表面看来，他不像是个能安排这些犹太人获救的可靠人选。

杉原千亩正处在自己外交官生涯的中期，靠着16年来在一连串岗位上的恪尽职守，成了日本驻立陶宛总领事。良好的资历推动了他在外交官圈子里的晋升：他生于武士之家，是政府官员的儿子。他精通俄语，给自己设定了十分高远的事业目标，希望有朝一日能成为日本驻莫斯科大使。跟更出名的奥斯卡·辛德勒一样，杉原千亩热爱狩猎、音乐和聚会。所以，要说这位过得舒舒服服、喜好享乐、干了一辈子外交工作的外交官，会拿自己的事业、声誉和未来冒险，救助一大早就把他从美梦里吵醒的陌生人，从表面上真看不出什么迹象。然而，他恰恰这么做了，而且他完全知道这可能会给自己和家人带来什么样的后果。

他跟等候在大门之外的人群聊了一阵，了解了他们的困境，便打电话到东京，希望上级允许他为这些犹太人发放旅游签证。虽然日本对犹太人发放签证和允许其定居的宽松政策尚未发生变化，外交部的上司却担心，继续推行这些政策会损害日本与纳粹德国的外交关系。因此，杉原千亩的请求遭到了拒绝。随后，他更为迫切地提出了第二次和第三次请求，也同样遭拒。他此时40岁，在这以前，从未有过丝毫的不忠，也从不曾违抗命令。可这一天，这位在生活中声色犬马、在事业上雄心勃勃的外交官，做出了没人想得到的举动。他直接违背了上司重申了三次的命令，开始动手发放旅游签证。

这一选择摧毁了他的职业生涯。不到一个月，杉原千亩就从立陶宛总领事被贬为闲职，不得再独立执业。最终，他以“犯上”罪名被外交部开除。战后他一文不名，靠卖灯泡为生。但在被迫关闭立陶宛领事馆前

的几个星期，杉原千亩坚守个人选择，从清早到深夜，一直在面试申请人，为他们准备文件。甚至，在领事馆关闭之后，他住进宾馆，仍继续签发证件。哪怕这件事带来的压力让他形容消瘦、精疲力尽，哪怕同样的压力使他的妻子无法再为襁褓中的孩子哺乳，他仍片刻不停地签发证件。哪怕最后，他已经登上了离开立陶宛的列车，在站台上，在列车上，他仍在签发证件，并把写好的签证抛向车厢外求助的犹太人。他一共救下了几千条无辜的生命。等列车开动，要把他带走，杉原千亩朝着那些他来不及帮助的人深深鞠了一躬表示歉意，请他们原谅自己已经无能为力（见图8-5）。



图8-5 杉原千亩和他的家庭：内与外

在立陶宛领事馆办公室（上图）为犹太人发放了几千张旅游签证之后，杉原千亩被贬为闲职，来到纳粹控制的欧洲。在捷克斯洛伐克，他让家人站在一座公园外

面照了张相（下图，从左到右依次为他的儿子、妻子和妻妹）。公园门口挂了块德语标志牌：“犹太人不得入内。”这块牌子是偶然拍下的，还是暗含了有意识的辛辣讽刺呢？暗示性的证据来自妻妹的右手，你看到它在哪儿了吗？

杉原千亩决定帮助数千犹太人逃往日本的原因，恐怕无法用单独的某个因素来解释。一般来说，这种非凡的慈悲之心是来自各种合力的交互作用。但就杉原千亩而言，家庭因素表现得很明显。他的父亲是税务官，曾一度被派驻韩国，就把家搬到了当地，开了一家旅馆。他的父母愿意接纳形形色色的旅客，在家庭旅馆里照料其基本食宿需求，甚至为客人提供浴缸，并帮他们洗衣服，哪怕有些人穷困潦倒、付不起钱也没关系。杉原千亩曾回忆说，这些点点滴滴的小事带给他极大的影响。从这个角度，我们可以看出杉原千亩日后搭救数千欧洲犹太人的一个原因：因为经常在家接触不同的人，他的家庭感扩大了。45年之后，他在接受采访时说：犹太人的国籍和宗教并不重要；他们都跟他一样，是人类大家庭里的一员，他们需要他的帮助，这才是最重要的。如果父母们希望孩子培养起广泛的慈悲心，不妨参考他的经历：让孩子在家里接触到来自不同背景的人，并像家人一样对待他们。

传奇人道主义者特蕾莎修女也曾讲过一个来自自己童年的类似故事，而且，她双亲的行为也隐含着类似的暗示。她在塞尔维亚长大，起初生活优渥，但后来随着父亲的去世而家道中落。她看着母亲德拉比娜为任何有需要的人提供吃穿用度。从学校回来，她和兄弟姐妹经常发现陌生人坐在餐桌旁，吃着家里本来就不多的食物。特蕾莎问他们为什么在这儿，母亲会回答：“他们是我们的同胞。”请注意，在概念上，“我们的同胞”就等于“我们”。

事实上，从杉原千亩这个个案出发推广得出更宽泛的结论，哪怕再加上特蕾莎修女对小时候家庭环境的回忆，始终是很冒险的。不过就本例而言，我们知道，在那个时代伸出救援之手的人里，不光只有杉原千亩一个人的童年家庭生活中融合了多种多样的人。研究者从一个收容过犹太难民的欧洲人士的大规模样本当中发现了类似的经历。一如预期，

较之同时期未救助犹太人的人员样本，救助者样本里的人在成长期间，对多样化的人群有着更强烈的认同感。这种扩大的归属感，不光与决定在大屠杀期间救助他人的行为有关，甚至，在半个世纪后接受采访时，这些救助者仍在帮助更多样化的人群。

最近，研究者设计了一套人格量表，评估个人与全人类自发认同的程度。这一重要量表衡量了一个人使用代词“我们”的频率、把他人视为家人的观念、自我与他人概念的整体重合程度，可以预测出此人通过向国际人道救援活动捐款救助外国人的意愿。此外，对他国移民困境的同情反应，似乎来自当事人与他人之间的自我他人感知重合度。杉原千亩在第二次世界大战前的帮助行为的相关背景及个人因素信息，来自对当时日本和欧洲环境的历史记录以及杉原千亩接受的采访。

乡土 从拥有共同基因的个体组成小而稳定的群落开始，人类演化至今，由此获得了一种倾向，除家人之外，我们还偏爱跟自己有着紧密接近性的人。这种倾向甚至带来了所谓的地区主义。有时候，我们可以从邻里和社区关系上看出它的巨大影响力。让我们再次回头看看大屠杀时期几个扣人心弦的故事。这些故事很好地证实了这一点。

让我们首先来揭开本章开头那个故事的谜底。在纳粹集中营里，守卫让所有囚犯排队报数，每10人杀1人。然而，突然之间，他举动反常：跳过了第10号，而朝第11号囚犯开了枪。他这么做的原因可能有好几个。说不定，那个侥幸留得一命的囚犯过去的劳动表现一贯不错；也说不定，守卫注意到那人体力、智力或者健康状况良好，将来会是个好劳力。但另一名守卫问他为什么这么做的时候，他的答案却很明显跟这些实际的考量毫无关系。他所说的理由非常简单，但足够充分：他认出10号囚犯跟自己是同乡。

人类学家科恩在一篇学术文章里讲述了这个故事，对它所体现出的深刻矛盾做了评论：“守卫在大规模杀人时尽忠职守，却对受害者群体里的一名特定成员保持着怜悯与同情。”虽然科恩并未探讨相关的问

题，但辨识出是什么因素导致守卫从大规模行刑的冷血杀手，变得对具体的某个人有了怜悯和同情，这非常重要。这个因素就是地点的相关性。

值得一提的是，科恩对集中营故事的描述，来自他和一个前纳粹守卫的对话，该守卫在转述这个故事时正是科恩的室友。据估计，利尼翁河畔勒尚邦的群众在法国新教牧师安德烈·特罗克梅（Ander Trocme）及其妻子玛格达（Magda）的带领下，救下了3 500条人命。1940年12月，他在自己家门口发现了一个冻僵的犹太妇女，这是他帮助的第一个难民。至于他为什么决定帮助这位妇女，我们很难给出确定的答案。但在战争快结束时，特罗克梅被关起来，法国维希市政府官员要求他及其他居民报上犹太人的名字，他的回答仿佛直接出自杉原千亩之口，其核心和世界观如出一辙：“我们不知道什么是犹太人。我们只知道什么是人。”至于说亲属和邻居谁更容易答应特罗克梅的请求这个问题，其他证据来源暗示答案是前者。也就是说，肯定是亲戚的人提出的要求更难以拒绝。举个例子，在20世纪90年代中期的卢旺达大屠杀期间，胡图族人会向自己的图西族邻居发起攻击，鼓动发起攻击的人以部落为号：“胡图力量”既是战斗口号，也是屠杀的理由。

现在让我们来看看，完全相同的因素，在完全相同的历史时期，曾产生过怎样完全不同的结果。许多在大屠杀期间对犹太人施以援手的好心人的故事，都揭示出一个少有人分析但值得关注的现象：大多数情况下，掩护受纳粹迫害的犹太人并为之提供藏身之所和食物的人，并不是自发自觉地找到受害者给予帮助的。更值得注意的是，受害者本身通常也并没有直接向他们寻求帮助。相反，直接提出请求的人，大多是提供救助的好心人的亲戚或邻居，由他们代表受追捕的个人或家庭请求获得帮助。因此，从真正的意义上看，这些施救的好心人不是顺从了需要帮助的陌生人，而是答应了自己亲戚或邻居的请求。

当然，这并不是说，救助者出手相助不是因为同情受害者。前文提

到的特罗克梅先是收留了自家门外一名沦落多时的难民，继而劝说家乡小镇利尼翁河畔勒尚邦的其他居民收容、藏匿、协助数千犹太人逃跑。特罗克梅这一非凡故事的启发意义，不在于他怎样帮助第一个难民，而在于他怎样安排照料之后的其他许多难民：他先是请求那些很难拒绝自己的人，即他的亲戚、邻居帮忙，接着再敦促后者向亲戚、邻居提出同样的请求。策略化地利用现有人际纽带，不光让他成了一个富有同情心的英雄，还让他的营救行动大获成功。

影响力研究

其他非常成功的传播工作者也曾利用过同一地域内的“现成联盟”。2008年美国总统竞选期间，有研究显示，与选民直接进行特定类型的个人接触，可能会明显提升选票总数。基于此，奥巴马的战略师向“战场州”700多个地方办事处投入了前所未有的巨额资金。这些办事处的工作人员和志愿者的主要责任并不是说服附近的选民相信奥巴马适合担任总统，而是为了确保那些可能支持奥巴马的居民登记，并在选举日投下一票。为实现这一目标，办事处的志愿者被分配到各自的社区内进行密集的挨家挨户游说工作，战略师们知道，这会增加邻里间的接触，进而对选民后续投票产生更大影响。选举之后，研究者对这一战地办事处策略所做的影响分析表明，它效果很好，让奥巴马在三个摇摆州（佛罗里达州、印第安纳州和北卡罗来纳州）赢得了选举。该研究者分析，是办事处的策略让奥巴马在选举团投票中获得压倒性胜出。

人们尤其容易受到本地呼声的影响，这一现象被称为“本地优势效应”，转换到选举政治中，就意味着公民更容易顺从来自本地社区成员

的投票请求。顺便说一下，最后这一点认识，并不是选举团队阅读了大量行为科学文献，而是因为行为学家大卫·尼克尔森（David Nickerson）在奥巴马的竞选团队中担任了行为科学顾问。

你是否留意到，有些商业组织会把它的客户、订阅者或粉丝称为“xxx社群”的成员？还有些类似的组织会称为“xxx大家族”的一员，我想，两者道理是相同的，它们都调动了一种强大的原始的“我们”意识。

地区 就连来自同一个大致的地理区域，也能带来一种归属感。在全球范围内，体育锦标赛都会激起运动队所在周边地区居民的个人荣誉感：球队赢了，就像是居民自己赢了一样。仅在美国，研究证据就从多种多样的途径证实了这一观点：如果调查研究是家乡的州立大学发起的，公民答应参与的概率更大；在亚马逊购物，如果评价者跟购物者住在同一个州，购物者就更有可能听从其推荐；人们会过度高估自己的家乡州在美国历史上的作用；读到阿富汗战地阵亡的新闻报道，发现倒下的军人跟自己是同州人，读者会更反感战争；甚至，回到两百年前的美国南北战争时期，如果士兵们发现他们来自相同的地区，就不太容易疏远，对作战单元里的战友们会更加忠诚。

从球迷到士兵，我们都可以看出地区身份对类似归属感的反应的极大影响。但还有一个更具说服力的例子，同样来自大屠杀期间，而且表面上看有些扑朔迷离。

虽然杉原千亩的签证救下了数千犹太人，但当他们抵达日本控制地区时，就成了更大规模的犹太难民的一部分，这些难民群体集中在日本的神户以及上海的日租界。1941年日本偷袭珍珠港，促使美国参战之后，难民进出日本的所有途径就都中断了，犹太社区的安全也岌岌可危。说到底，日本当时正跟希特勒全面结盟，要维持双方关系的稳固，就必须支持希特勒的反犹主义。更重要的是，1942年1月，希特勒在柏林的万湖会议上正式发起了全面消灭犹太人的计划。随着轴心国“最终解决方案”的就绪，纳粹官员开始向东京施加压力，要求“解决”在日犹

太人。这次会议之后，东京方面陆续收到了德国有关死亡集中营、医学实验和海上集体溺杀的提议。然而，尽管有可能损害日德关系，但从1942年初到战争结束，日本政府始终顶住了这些压力。为什么会这样呢？

答案很可能跟几个月前发生的一系列事件存在联系。纳粹派出约瑟夫·梅辛格（Josef Meisinger）上校前往东京，梅辛格是盖世太保（即纳粹的秘密警察），曾在波兰下令处死了16 000人，史称“华沙屠夫”。1941年4月抵达日本后，梅辛格便着手施压，要求日本政府对犹太人采取残暴政策，并表示很乐于帮忙设计和执行这项政策。日本军政府的高级成员起初不知道怎样回应，想听听各方意见，就传召犹太难民派两名领袖去开会，这次会议对旅居日本的犹太难民的未来将产生重大影响。两名获选代表都是受人尊敬的宗教领袖，但他们受人敬重的方面不同。一个人是拉比摩西·沙提克（Moses Shatzkes），他饱读经文，战前曾是全欧洲最著名的塔木德学者。另一位是拉比希蒙·卡利什（Shimon Kalisch），他年纪更长，对人性有着非凡的洞察力，算得上是个社会心理学家（见图8-6）。



图8-6 留日期间的拉比们

第二次世界大战期间，日本没有屈服于纳粹的压力，残酷对待犹太人。原因之一可能是拉比们使用的说法调动了日本官员的归属感，将犹太族裔包含在内，而排斥了纳粹。图为关键会议当天所摄，另外两人为陪同。

两人进入会议室后，和翻译一起站在日本最高统帅部一群大权在握的权贵面前。这些能决定犹太人生死的人，径直问出了两个最致命的问题：为什么我们的盟友纳粹这么仇视你们？我们为什么要站在你们这一边，反对他们？沙提克认为这些问题涉及复杂的历史、宗教和经济事宜，没能立刻做出回答。但卡利什却根据自己对人性的洞察，说出了我研究影响力30年来印象最深的一句说辞。“因为，”他平静地说，“我们都是亚洲人，跟你们一样。”

话虽说得短，但可谓神来之笔。通过暗示纳粹自称的“优越的”雅利安人种跟亚洲民族有着天然的不同，它把日本官员心中占主导地位的群体认同，从战时的临时同盟转到了与地区、遗传相关的亲近性上。卡利

什靠着深刻的洞察力，把犹太人和日本人并列在一起，而把纳粹孤立开来。老拉比的回答对日本官员产生了强烈的影响。一阵沉默后，他们彼此商议，宣布休会。当他们返回时，最高军事长官站起身，向拉比们宣布了一个令人宽慰的好消息：“回去找你们的族人吧。告诉他们，我们会保证他们的安全与和平。在日本领土之内，你们无须担心。”日后发生的一切也果真如此。[\(49\)](#)

毫无疑问，技巧娴熟的沟通者能够驾驭亲人意识和乡土意识带来的联盟的力量。巴菲特和拉比卡利什对此作了证明。与此同时，对立志提升影响力的人来说，还有另外一种联盟效应可以运用。它不是来自血缘或地理位置上的在一起，而是来自同步或合作展开行动。接下来我们就要介绍它。

联盟2：行动合一

我的同事威廉明娜·沃辛斯卡（Wilhelmina Wosinska）教授回想起自己20世纪50年代到60年代在波兰的成长经历，总是百感交集。从消极面来说，除去基本物资的持续短缺，还有对个人自由方方面面的限制。然而，官方引导她和同学们积极地看待这些限制，把它们视为建立公正平等的社会秩序的必要条件。这些积极的情感，会通过庆祝活动定期表现出来，反过来，庆祝活动也助推了积极情感的高涨。在活动中，大家一起唱歌、游行、挥舞旗帜。她说，她深刻地感受到了活动带来的影响：身体的活跃，情绪的昂扬，心理上的肯定。在这些经过严格编排、有力协调的团体活动中，她前所未有地体会到了“我为人人，人人为我”的概念。我听沃辛斯卡教授说起这些活动，一概是在论述群体心理学的学术演讲当中。虽然那都是学术会议，但她形容自己参加活动的时候，总是会音量提高，血涌到脸上，眼睛闪闪发光。这些同步化的体验有着不可磨灭的本能色彩，从人类生存条件的角度来看，它们原始且重要。

事实上，考古学和人类学的记录在这点上十分明确：所有人类社会都发展出了协同一致的方式，齐声或同步唱歌、游行、举行仪式、吟诵、祈祷和舞蹈。更重要的是，从史前时代开始他们就在这么做了，比如新石器时代和红铜时代的岩画、洞穴壁画等常常表现集体舞蹈（见图8-7）。行为科学的记录同样清晰地解释了其中的原因。人们统一行动时，就真的统一了。集体联盟的合力感非常符合社会的利益，由此带来了通常只跟小家庭单位挂钩的高度忠诚和自我牺牲精神。因此，人类社会早在古代就已经发现了让反应协同一致的黏合“技术”。它带来的效果类似亲人关系：“我们”感、融合感、自我与他人界限的模糊、为集体牺牲的意愿。不足为奇，在部落社会，战士经常会在作战前有节奏地一起跳舞。



图8-7 新石器时代的集体舞

按考古学家优素福·加芬克尔（Yosef Garfinkel）所说，史前艺术所描绘的社会活动几乎总是舞蹈。图为印度比莫贝卡特石窟的壁画。

与他人的融合感听起来罕见，但事实并非如此。它可以很轻松地通过多种方式产生。在一组研究中，跟被试与同伴独立朗读故事相比，与同伴一起大声朗读故事（或通过轮流朗读故事中的举止来协调）产生的“我们”感和与同伴的联盟感更强。其他研究也显示出一同行动带来的

有利影响。研究者将被试安排为23人或24人一组，让一些小组的成员按相同的顺序一起说单词，让另一些小组的成员说相同的单词但采用不一样的顺序。结果显示，一起说单词的小组成员在同伴身上感受到了更多的“我们”感。后来，一起说单词的小组成员在玩集体电子游戏时，也因为更大程度地协调了彼此的努力而得到了更高的分数。关于这一现象的最后一项证明来自对大脑活动的研究。当密切参与联合项目时，被试的脑电波模式会彼此匹配、同起同落。因此，当人们同步行动时，大脑便会表现出相同的波长。

如果以运动、声音或认知方面一同行动，可以充当亲属单位彼此归属的替代品，我们应该从这些联盟形式中看到类似的结果。的确如此。上面结果中有两种，对想要获得更强影响力的人来说尤为重要：提升喜爱感、获得他人更多支持。

不同类型的行为科学数据都支持协同反应在联盟感（包括自我—他人身份混淆感）中发挥的作用。在我们的进化史上，甚至比新石器时代和铜器时代更早的时候，就出现了让运动与有节奏的声音相协调的趋势。黑猩猩会对音乐节拍做出反应，这表明大约600万年前我们与它的共同祖先也存在这种反应。一名研究者将人类之间协调行动形成的群体称为临时“邻里”，成员对彼此的指向发挥了很大的影响对旨在鼓励集体联盟的社会机制做了尤为可信的论证。渴望施加影响力的人，或许可以从协同效应中获得极大益处。世界著名历史学家威廉·H.麦克尼尔

（William H. McNeill）说过这样一句精辟的话：“一边一起发出声音，一边有节奏地运动，是人类所能想出来的创造并维持有意义社群的最可靠、最迅速、最有效的方式。”

提升喜好之情

当人们行动合一时，不光会认为彼此更相似，此后的互相评价也更积极。彼此之间的更相似，变成了彼此之间的更喜欢。在实验室里敲击

手指、在对话时微笑、师生互动中的身体调整.....所有这些行为，一旦同步，就能提高人们对彼此的评价。加拿大的一群研究者想知道，协调行动能不能用来做一件有着更重要社会意义的事情：协调行动将相似转化为喜好的能力，能不能用来减少种族歧视呢？研究者注意到，虽然我们通常会跟群体内成员相互协调产生“共鸣”，但跟群体外成员却难得如此。他们推测，随之而来的联盟感差异，或许至少能部分地解释人类偏爱自身所属群体的自动化倾向。如果是这样，想办法让人跟群体外成员协调行动，就有可能减少偏见。

影响力研究

为了验证这一想法，他们进行了一项实验：白人被试观看7段视频片段，内容是黑人从玻璃杯里喝了一口水，接着把玻璃杯放在桌子上。一些被试只观看剪辑片段和片中动作，其余被试则要模仿视频片段，以完全一致的动作从摆在眼前的玻璃杯里喝一口水。

稍后，被试接受检验，测量其隐性的种族偏好，只看了黑人演员视频的被试对白人表现出典型偏好，而曾跟那些黑人演员同步行动的被试没有表现出这种偏好。

在对实验结果做太多阐释之前，我们应该首先认识到，评价上的积极变化是在研究的协同环节完成短短几分钟之后测量出来的。研究者并未提供证据，表明这种变化在跳出研究的时间范畴和地点环境后还能持续。不过，就算考虑到这些因素，仍存在值得乐观的情况，因为我们已经找到了一种纠正群体内外偏好的方法，足以在某些具体情境，如工作面试、销售电话、初次会面中发挥作用了。

值得一提的是，喝水实验还包括第三道程序，即被试被要求模仿白人演员的喝水行为，这一程序令被试对白人产生的典型偏好达到了几乎

夸张的程度。

有趣的是，有一种协同活动还有一点额外的好处：把注意力引导到一条信息上的时候，人们如果发现周围同时还有别人一起这么做，就会以更高的强度去做这件事，为其分配更多的认知资源。不过，只有在当事人对另一个人有着归属感时，情况才是这样。跟一个关系亲密的人一起关注某件事的行为，似乎是这件事值得特别关注的信号。

增加支持行为

有充分的证据表明，跟他人，哪怕是陌生人一起行动，能产生联盟感，提升喜好之情。但来自协调反应的联盟和喜好，是否强大到了足以改变行为呢？毕竟，能否改变行为才是社会影响力是否有意义的黄金标准。要回答这个问题，不妨看看下面两项研究。其一考察的是一个个体向事先产生了联盟感的另一个个体提供帮助，其二考察的是一群事先产生了联盟感的团队成员之间的合作。两项研究所要求的行为，都要求被试做出一定的自我牺牲。

影响力研究

在第一项研究中，被试戴着耳机，听一连串录制好的曲调，同时按自己听到的节奏敲击桌子。有些被试跟搭档听到的是相同的曲调，因此会认为自己在跟对方一致地敲击；还有些被试跟搭档听的是不同的曲调，所以并未同步行动。事后，所有的被试均得知，他们可以自行离开，但他们的搭档则需要留下来回答一连串冗长的数学和逻辑题。被试可以自己选择是否留下来替搭档做一些任务。

研究所得的结果毫无争议地证明，协同行动具有提升自我牺牲式的支持行为的先发影响力。敲击桌子时跟搭档不同

步的被试，只有18%选择留下来帮忙；而跟搭档同步敲击桌子的被试，有49%放弃了自由时间，向搭档伸出了援手。

另一群不同的研究者进行了第二项有趣的研究，采用历史悠久的军事战术来培养团队凝聚力。研究者将被试编成小组，并要求部分小组齐步走；又要求另几支小组一起走相同的时间，但无须齐步，正常走即可。稍后，所有组员都参加了一个经济游戏，他们既可以最大化地提升自己获得个人收益的机会，也可以放弃这一机会，以确保自己的队友获得更大收益。先前曾齐步走的组员跟队友合作的概率，比之前正常走的队友高50%。

后续研究解释了其中的原因：最初的协同反应带来了一种联盟感，从而让人更乐意牺牲个人利益，换回集体的利益。

我说判断社会影响力的黄金标准是“改变行为”，并不是要否认在影响力过程里改变他人感受、信念、看法或态度的重要性。但在我看来，努力令这些因素发生改变，似乎总是为了实现行为上的改变。时至今日，军事训练中依然会采用齐步走的做法，哪怕从战场战术的角度来说它的价值早已丧失。研究者通过两项研究给出了一个令人信服的理由：齐步走之后，被试变得更乐意顺从队友的要求，伤害群体之外的成员；无论提出要求的人是长官还是仅仅是同伴，情况都是如此。

由此可见，集体可以预先安排协同反应，在多种情况下促进成员的联盟感、喜好和随后的支持行为。但到目前为止，我们考察的手法，如敲击桌子、喝水、轻抚脸庞等，好像都不能立刻执行，至少不能大规模地执行。从这个角度看，齐步走稍好，但也好得很有限。社会组织要想带来这样的协同，让成员为了团队目标而努力，有没有什么可以普遍运用的机制呢？有，那就是音乐。另外，就算是个人沟通者，也能借助音

乐这一影响力中介，推动他人朝着目标前进。

音乐共鸣

从人类有史以来，在全球多种多样的人类社会当中，为什么都有音乐的身影出现呢？对此有一个很合适的解释。音乐集合了多种可辨识的规律性：节奏、韵律、强度、节拍和时间，因此拥有极为稀有的协同力量。听众可以很容易地在肌肉运动、感官、声音和情绪维度上彼此达成一致。这种一致状态带来了我们熟悉的联盟标志：自我与他人的融合，社会凝聚力，支持行为。

影响力研究

就支持行为这最后一个方面，让我们来看看在德国进行的一项对4岁儿童的研究。孩子们参加一个游戏，有一些会跟搭档一起边唱歌边绕着圆圈走，他们的动作要跟上录好的音乐；另一些孩子要做的几乎相同，但没有音乐伴奏。后来，等孩子们有机会展现自己的助人精神时，一起和着音乐唱歌绕圈的孩子帮助搭档的概率，比未受音乐先发影响的孩子高3倍。

这项研究的作者对他们所观察到的乐于助人精神提出了两个具有启发意义的观点。首先，他们指出，帮助行为要求自我牺牲，帮忙的人要放弃个人的玩耍时间协助搭档。共同的音乐和动作体验，大幅提升了孩子们稍后的自我牺牲精神，这对想要改变4岁小孩自私选择的家长很有启示，让他们不必再一遍遍地大吼：“莉亚，该把玩具给哈利玩了。莉亚？莉亚！莉亚，你赶紧给我把那个玩具拿回来！”而在我看来，研究者提出的第二个观点，和第一个同样重要：孩子们的个人牺牲，并不来

自对是否应当提供协助的理性权衡。帮忙的意愿完全不是扎根在理性当中的。它是自发的，直觉性的，以跟搭档一起投入音乐所自然产生的情感联系为基础。这对我们控制社会影响力过程有着深远的意义。[\(50\)](#)

系统工程

行为科学家很早就曾指出，存在两种评估和认识世界的方式。最近获得了广泛关注的这一类主张，是经济学家卡尼曼对思维系统1和思维系统2所做的区分。前者是快速的、联想的、直觉的，往往还是情绪化的；后者则速度较慢，是蓄意的、分析的和理性的。这两种思考方式是相互独立的，支持的证据是：激活其中一方，另一方就会被抑制。充满感情地体验一件事时，你很难认真地思考它；反过来，使用逻辑分析一件事的时候，你也很难全身心地体验这件事。这对影响力有什么启示呢？说服方诉求的着眼点，最好跟接受方的思维方式相对应。也就是说，如果你主要是从与情感相关的特点来考虑买车，希望车子外观好看、提速令人振奋，那么销售人员就最好使用与情感相关的论点来说服你。研究表明，在这种情况下，哪怕只是说“我感觉这款车很适合你”，也能取得很大的成功。但是，如果你主要站在理性的基础上来考虑买车，关注燃油经济性和保值性，“我感觉这款车很适合你”的说法则很可能搞砸买卖。

音乐的影响作用，针对的是思维系统1。举例来说，英国创作歌手埃尔维斯·科斯特洛（Elvis Costello）曾说过一句话，论述在写作的结构下描述音乐是何等之难：“描写音乐，就如同用舞蹈描述建筑。”

说到爱情中认知与情感的不匹配，比尔·威瑟斯（Bill Withers）在1971年有过一首歌名叫《逝去的阳光》（*Ain't No Sunshine*），讲的是一位年轻姑娘再次离开了家，男人痛苦不堪：“我知道，我知道，我知道.....（重复26次）/嘿，我不能再想这件事 / 当她走了，就没有了一点阳光。”威瑟斯在一首流行歌曲的歌词里，用了我认为最纯粹的诗歌

形式，提出自己的观点：在浪漫爱情带来的阵痛里，人们在认知上意识到一件事并重复了26次，也不能缓解他们所感受到的强烈情绪。

人们跟着节奏，也跟着彼此一起唱歌、摇摆、扭动、挥手，这反应发自他们的感官和内脏。当音乐在意识里占主导地位的时候，人很少会分析性地思考。在音乐的影响之下，人很难进入审慎而理性的思维路径，因此也就无法采用这种思维方式。有两句名言点明了音乐的这一特点。一句来自伏尔泰，他曾不屑地说：“说出来愚蠢的话，唱出来就不一样了。”另一句则是广告行业的战术宝典：“如果你无法用事实向受众证明你的观点，那就唱给他们听。”因此，如果沟通者的观点理性力量不强，不妨放弃这条战线，改为侧向迂回。以音乐和歌曲为武器，可以把战斗带到一个理性无能为力的地方，最终胜出的将是和谐、同步和联盟感。

这一认识帮我解开了长久埋在心底的一个谜。我年轻时没什么音乐天赋，这个问题可把我愁死了：为什么年轻姑娘这么容易对乐手倾心钟情呢？真没有逻辑可言，对不对？的确如此。跟大部分乐手成功建立亲密关系的概率都低得惊人，可这无关紧要：概率是以理性为基础的。大多数乐手眼下和将来的经济前景都堪忧，这也无关紧要：这些都是经济上的原因。音乐跟这类实际的问题不相关。音乐讲究的是和谐。和谐的旋律，唤起了情感和亲密的关系。

此外，音乐和爱情因为有着情感和和谐性上的共同点，在生活里也是彼此强烈相关的。否则你怎么解释当代歌曲里有那么高的比例是以浪漫爱情为主题的呢？按最近的一轮系统性统计，这个比例是80%，可谓占了绝大多数。真是奇妙啊！在我们说话、思考或者写作的时候，绝大多数时间并不以浪漫爱情为主题，但在唱歌的时候却是如此。

所以，现在我明白为什么年轻姑娘抵挡不住乐手的诱惑了，她们正处在对爱情和音乐最感兴趣的年纪，这两种体验之间的强力联系让乐手势不可挡。想要些科学证据？且让我把法国一项研究的结果唱给你听

吧。研究者让一名男性去找年轻姑娘索取电话号码，此时，他要么背着吉他盒，要么背着运动包，要么什么也没背：

法国的科学家们呀，
开始还怀疑，
陌生人背着吉他盒去索要
姑娘的电话号码，
能不能提高成功概率呢？
其实不用太烦恼呀，
成功率翻了一倍，还要多，哟哟哟（见图8-8）。

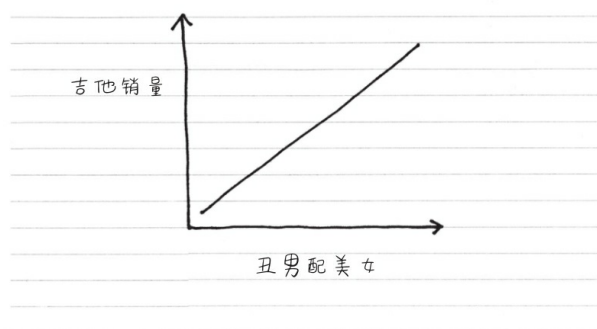


图8-8 吉他把笨蛋变英雄

对所有有意将说服成功率最大化的人来说，本节的关键信息不仅仅在于音乐与思维系统1的反应之间的相关性，或是一旦注意力被引导到这种反应方式上，人就会轻率行事。说服沟通选用的思维模式，应该跟受众的思维模式相搭配，这才是最重要的教训。

如果接受方的目标是不讲究理性的、享乐主义的，与之匹配的信息就应该包括非理性的元素，如音乐的伴奏等；而如果接受方的目标是理性的、务实的，那么信息则应包含理性元素，如事实等。营销专家斯科

特·阿姆斯特朗在其精彩作品《广告说服力》一书中指出，有人对1 513段电视广告做了分析，87%都配有音乐。但为信息增加音乐这一常规做法很可能存在缺陷。阿姆斯特朗还考察了相关研究，最后得出结论：只有在情绪性背景下宣传熟悉的、基于感受的产品时，也就是人们不太使用理性思考的时候，才应该使用音乐。对有着强大支持论据且会造成严重后果的产品，如安全设备、软件套装等，背景音乐反而削弱了广告效力，因为此时，潜在的买家很可能会进行仔细的思考斟酌。[\(51\)](#)

持续交换

2015年初，《纽约时报》上的一篇文章点燃了读者的兴趣，引来评论无数。它野火燎原般传播开去，成了《纽约时报》有史以来阅读量最大的文章之一。对一家像《纽约时报》这样的新闻媒体来说，这种事情似乎不算什么，毕竟，它是一份在美国国内和国际新闻界地位很高的报纸。但这篇文章并不是来自政治、商业、技术、科学、健康甚或时尚栏目。文章的标题是《想爱上一个人，请这么做》，作者是曼迪·莱恩·卡特伦（Mandy Len Catron），她声称发现了一种产生强烈的情感亲密性和爱情社会关系的有效途径，只需45分钟。她说，这种方法很管用，因为她亲身验证过。

这项技术来自心理学家阿瑟·阿伦（Arthur Aron）和伊莱恩·阿伦（Elaine Aron）夫妇，两人发起了一个考察亲密关系的研究项目。这个项目研究了协调行动的一种具体形式，即伴侣双方参与持续的交换。其他的心理学家早已证明，长期的交换，会让人向交换伙伴表现出额外的善意，谁是最后一个都无所谓。

影响力研究

阿伦夫妇和同事们揭示了互惠交换是怎样将参与交换的

双方结合在一起的，进而解释了这种顺从的意愿。他们采用了一种特别的交换，其作用强大得足以令人们彼此相爱。这种交换就是个人自我表露。

整个过程并不复杂：被试两人结为一组，向彼此提出问题，听到问题的人先回答，之后提问者自己也做出回答。问题共有36道，要被试逐步透露自己的个人信息，反过来说，两名被试也将一步步地了解到更多有关对方的信息。最开始，问题是这样的：“在你看来，完美的一天是什么样的？”而后来的问题则可能是这样的：“你在友谊中最看重的是什么？”接近尾声时的问题是：“家人里谁去世会让你最不安？”

关系的深化超出了所有人的预期。整个过程在45分钟之内带来了无可比拟的情感亲密和人际联盟，尤其这还是在情感真空的实验室环境里，互相搭档的都是陌生人。而且，实验的结果更让人意想不到。按伊莱恩在一次采访中所说，自此以后，数百次应用这一方法的研究都证实了它的效果，一些被试事后甚至结婚了。她还介绍了这种方法神奇效力的两个关键。首先，问题在个人信息表露的强度上逐渐提升。所以，被试在回答的时候，越来越信任地向对方敞开心扉，这种信任感，在关系亲密的伴侣里极具代表性。其次，被试是一起这么做的，也就是说，他们协调地、有来有往地提问和回答，互动始终保持同步。这一点恰好和本章的主题相符。[\(52\)](#)

一起吃苦

这里还有另一个解决途径，仍然和第二次世界大战期间纳粹大屠杀

时期的一道谜题有关。它为实现联盟提供了一条不同的途径。1940年夏天，当杜塞尔多夫的盖世太保正有序地识别犹太居民，以将之押送到欧洲各地的死亡集中营时，他们收到了来自纳粹头领、党卫队队长海因里希·希姆莱（Heinrich Himmler）一封不同寻常的信。该信指示他们不得迫害犹太居民以及该市法官恩斯特·赫斯（Ernst Hess），说有一位纳粹高官下令，“不得以任何方式纠缠他”。

我们在前文介绍的所有联盟感的源头，都无法解释赫斯受到的特别待遇。赫斯法官未曾在案件审理中对纳粹高官的家人做出过有利的裁决，也不曾跟纳粹高官一起在小镇长大，也不曾单纯地跟纳粹高官在军队里一起排队踢过正步（严格来说，多年前他们的确这么做过）。原因比这更复杂：在第一次世界大战期间，赫斯法官和纳粹高官一同经历了那场旷日持久的可怕冲突所带来的苦难、贫困和不幸。在长达141日天的索姆河战役（这场可怕的战役夺去了120万士兵的生命，其中50万来自德国军队）中，两人同一天在战场上相继受伤。或许，莎士比亚笔下的亨利五世在战前演讲中的那句话最能说明原因：“共浴血者，即吾手足。”

顺便说一句，希姆莱信中下令不得骚扰赫斯的“纳粹高官”，并不是普通的高级官员。这封信宣称，赫斯应该得到“救济和保护”，是“出于元首的意愿”——也就是希特勒，全世界最歹毒的迫害犹太人的大独裁者的意愿。

历史学家在盖世太保的官方档案中发现了希姆莱的“保护信”，并通过其他文件做了证实。至于是希特勒亲自指示希姆莱撰写并寄出这封信，还是希特勒的私人副官弗里茨·维德曼（Fritz Wiedemann）代表希特勒撰写并寄出这封信，学者们还存在一些争论。赫斯的贱民地位只持续了一年。随后，他在战争期间被安放在了几个强迫劳动站，包括一座劳动营、一家建筑公司和一家管道公司，但从未被送到死亡集中营。他家里的其他成员却未能幸免，如他的妹妹就被毒死在了奥斯威辛集中

营。战后，他成为一名铁路管理人员，最终晋升到法兰克福德国联邦铁路局的总裁位置，并于1983年在当地去世。

这个例子与前文提到的科恩的故事有着令人难以置信的相似之处。回想一下科恩的困惑：“守卫在大规模杀人时尽忠职守，却对受害者群体里的一名特定成员保持着怜悯与同情。”我们从卫兵和犯人在出生地上的共性（这是联盟感的特点之一）上解开了这个谜。就希特勒（是他发动了战争，并折磨和消灭了数百万犹太人）而言，对一个具体的人施以“怜悯与同情”，偏离了他设定的标准流程。这一次，把他们联结在一起的联盟因素不是共同的出生地，而是两人一起吃过的苦。

纵观人类历史，共同的痛苦一直是一种黏合剂，将身份融合到基于“我们”的情感联结当中。除了莎士比亚笔下亨利五世的例子还有更多的近代案例，也为这一现象提供了科学证据。2013年波士顿马拉松爆炸案发生后，报告直接卷入了负面事件（如当场听到或看到爆炸的任何方面）的居民，或仅报告遭受了巨大身心痛苦的居民，较之没那么痛苦的居民，更倾向于认为自己是波士顿的一员。此外，居民对这场悲剧思考得越频繁、越深刻，就越感觉自己与波士顿同胞们“融为一体”（如见图8-9）。



图8-9 借助“泥泞”的力量联盟

企业往往会在团队建设活动（包含逆境和风险元素）中，借助共同经历困苦磨难这

一点来催生团队的联盟力量。我参观了许多组织此类活动的场地，发现一些活动看起来既费力又可怕：激浪漂流、攀岩、在悬崖边吊绳索下降、桥上蹦极、赤脚踏在灼热的火炭上和雪地露营。在上面这张照片里，我们可以看到两名选手在帮助另一名选手，团队建设活动里的泥地赛跑似乎达到了刺激合作行为的预期效果。

研究者又做了检查，以确保共同吃苦带来的纽带效果，不是出于任何类型的共同活动体验。毕竟我们已经在前文看到，一起朗读故事、敲击手指或列队行军，都可产生“我们”感。在共同活动体验里加入痛苦，能否让最终结果表现得更为强烈呢？确实可以。一起完成了将手浸泡在冰水中90秒这一任务的小组成员之间的关系，比将手浸泡在常温水里的小组成员更紧密。随后，在跟小组同伴一起参与经济博弈时，一起吃过苦的组员更有可能做出旨在让整个小组共同赢钱（跟只实现个人赢钱的做法相对）的经济选择。

影响力研究

一起吃苦产生联盟和自我牺牲的强大力量，还可以从它在不同种族群体之间建立纽带的能力中看出。2020年，美国原住民特别是纳瓦霍人遭受了新冠疫情的蹂躏，一群意想不到的捐助人为他们提供了可观的援助。当地志愿者在慈善众筹平台GoFundMe上发布了消息，请求获得食物和其他必需品。意外的是，爱尔兰突然为他们发来了数十万美元的援助。爱尔兰人慷慨解囊的意愿，我们很容易就能从第2章有关互惠的内容里看出：这是一起跨越了数百年的时间、数千里的距离、也跨越了不同民族的互惠行为。在1847年爱尔兰的马铃薯大饥荒期间，一群美国乔克托族原住民筹集并寄去了170美元（约合今天的5 000美元）用以救助当地的灾民。现在，爱尔兰人报恩的时候到了。一位捐赠人在捐款时说道：“爱尔兰人永远不会忘记你们在爱尔兰饥荒期间饱含联盟和同情之情的义举。我们将和你们一同抗击新冠疫情。”

既然这个故事跟第2章的互惠原则以及其他许多惊人的互惠例子吻合，那我为什么还要放在“一起吃苦”这一小节呢？要找到答案，我们必须先跳出“为什么爱尔兰人在2020年采取行动提供帮助”这个问题，看一看为什么1847年乔克托人伸出了援手。在1847年之前的几年，乔克托人在政府的命令下，经历了一场痛苦的大规模迁徙，史称“血泪之路”。在这场迁徙中，6 000多名乔克托人丧生。美国原住民组织者凡妮萨·塔利（Vanessa Tully）解释说：“‘血泪之路’上死了那么多人，激发了乔克托人对急需帮助的爱尔兰人民的同情。这就是乔克托人伸出援手的原因。”值得注意的是，爱尔兰的许多捐赠者都在网上的留言中谈到，两个民族之间的纽带是由家族联结和共同的逆境铸就的并哀叹“我们的印第安兄弟姐妹”的苦难和共同的“血泪回忆”⁽⁵³⁾。

EBOX | 线上影响力

近年来，研究者开始通过分析社交平台上用户留下的痕迹，来挖掘关于人类行为的丰富信息脉络。2015年11月13日巴黎恐怖袭击之后，有人对推特活动的数量和特征进行了分析，让我们可以通过一种新颖的视角看到共同的困境给群体联盟造成的影响。从袭击发生之日起到之后的几个月里，行为科学家戴维·加西亚（David Garcia）和伯纳德·莱姆（Bernard Rimé）研究了法国62 114个推特用户的近1 800万条推文。他们搜索了推特上表达痛苦感受、痛苦的同步性（反映它的群体性质）、以及群体联盟和支持的措辞。该分析表明，恐怖袭击本身立刻激发了一波共同焦虑和悲伤的高峰，在2天到3天内有所下降。但在随后的几个星期和几个月里，推特上对这一情绪的声援和支持仍然高涨。此外，联盟和支持的表达力度和持续时间，跟最初事件导致的痛苦程度直接相关，而且两者是同步的。

作者的结论是：“我们的研究结果为集体情绪的社会功能提供了

新的解释，说明一个遭受集体创伤的社会不仅会在事件发生的同时产生负面情绪.....这些发现表明，人们不仅会在遭遇恐怖袭击后变得更加团结，而由于共同承受的痛苦，人们之间的纽带会变得更强大，社会也能因此迅速调整以面对下一次的威胁。”

作者点评：如果一种特定的行为模式在不同的观察方法之下都表现类似，我就会对它印象深刻。在我看来，共同的痛苦对随后的集体凝聚力与振作产生的巨大影响，就是那些激发自信的模式之一。

共同创作

早在美国大众接受保护自然的价值观以前很久，一位名叫奥尔多·利奥波德（Aldo Leopold）的男子就在宣传这项事业。20世纪30年代到40年代，他在威斯康星大学拿到了美国第一个野生动物管理教授职位，并针对这一主题设计出一种独特的伦理方法。这种方法挑战了环境保护的主导模式，即自然生态要根据人类的使用目的加以管理。

他在畅销书《沙乡年鉴》（*A Sand County Almanac*）里对自己提出的方法做过详细介绍，其主张是，要尽量让所有动植物都在自然状态下存活。他真心实意地坚持这一立场，直到有一天惊讶地发现，自己手里拿着斧子，做着完全与之相悖的行为：他砍倒了自己家里的一棵红桦树，只为了另一棵白松树得到更多的阳光和生长空间。

他想，为什么自己会做损害桦树而有利于松树的事情呢？按照他自己宣称的伦理，两者都有同样的权利在土地上自然生长呀。他怀着困惑，在头脑里搜索自己偏爱行为背后的“逻辑”，想到了两棵树之间有可能导致这种偏爱的各种差异。最终，他认为只有一点是主要因素。这一点跟逻辑毫无关系，是完全建立在感情上的：“好吧，松树是我亲手用铲子种下的，而桦树匍匐在篱笆下面，是自己扎的根。也就是说，我的偏爱在某种程度上来自家长作风.....”⁽⁵⁴⁾

利奥波德不是唯一一个怀有这种感情的人。人都对自己亲手创造的东西有着特殊的亲近感。这是人类的共性。举个例子，有一种现象，研究者称之为“宜家效应”（Ikea effect），那是指，对自己亲手造出来的东西，人会觉得“这业余水准的作品跟专业人士的作品有着同样的价值”。考虑到我们当前的关注点是共同行动的影响，这里有必要来探究另外两种可能性。如果一个人跟另一个人手把手地共同创作了某种东西，他会不会不光对作品产生特殊的亲近感，也对跟自己合作的人产生同样的感情呢？这种特殊的亲近感，是不是来自与他人的联盟感呢？

为寻求答案，让我们先来分析一个问题：为什么本节以共同创作为题，我却先介绍了利奥波德亲手种松树产生的影响呢？在这个过程中，利奥波德不是唯一的演员，这一点我敢肯定，他本人想必也同意。他跟大自然是共同的创作者：他种下了松树，大自然让松树成长。那么，有没有可能，跟大自然共同行动带来了这样的结果：他感觉与自然更为融合，从而对自己的合作伙伴更倾心、更尊重了？如果是这样的话，我们就得到了一点启发：共同创作也是一条联盟之路。

遗憾的是，我们没法就此询问利奥波德本人，因为1948年他就去世了。但我对这个答案有信心。

信心部分来自我参与的一项研究：调查管理者在开发产品中的个人投入度。我预期，管理者越是感到自己和员工一起做出了最终成品，对产品质量的评价就会越高。

结果的确是这样：首先，如果管理者受了引导，相信自己在开发设计最终产品，如一款新手表的广告的过程中扮演了重要角色，那么，较之相信自己没怎么参与开发的管理者，他们对广告的偏爱要强50%，哪怕两者最终看到的广告是一样的。其次，我们发现，认为自己参与得更多的管理者，同时也认为自己对广告的质量负有更大的责任，因为他们感到自己对手下员工有了更大的管理控制力（见图8-10）。这一点也如我所料。



图8-10 通过“归功于老板”避免阻碍

“做假账”是众所周知的商业诀窍，“假创作”也是。

但最后一点发现是我未曾料到的。管理者越是认为项目的成功缘于自己，也就越是认为项目的成功和手下员工的能力密不可分。我记得，拿到数据的时候，我不禁吃了一惊。或许不像利奥波德发现自己手里拿着斧子的那一刻那么吃惊，但总归也是狠狠地吃了一惊。监督者认为自己较多地参与了产品的开发，何以会让他们觉得自己和其他同事对最终产品的成功都有着更大的功劳呢？功劳的大饼不是只有一个吗？如果一方当事人认为自己的功劳更多，那么按照简单的逻辑来说，工作伙伴的功劳就应该减少才对呀。我当时始终没想明白，但现在明白了。如果共同创作带来了暂时性的身份融合，那么，适用于一个人的，也适用于另一个人，不管逻辑怎么规定。

征求建议是个好建议

对那些前来征求意见的人，我们都应该钦佩他们的睿智。

——本·富兰克林

共同创作不仅减少了项目合作中怎样让上级分给高效员工更多功劳的问题，还能缓解其他许多传统上难以消除的麻烦事。六七岁的孩子在分享奖励上往往很自私，很少跟玩伴平均分配。但如果他们是跟玩伴合作努力赢得奖励的，那么绝大多数时候连3岁的孩子也会均分奖励。在学校里，学生会根据种族、人种和社会经济地位结成小圈子，只在圈子内部寻找朋友和帮手。不过，当学生们和来自其他小圈子里的同学进行“合作学习”练习，即每一名学生都必须向其他学生传授一部分知识，好让所有人都得到好分数时，如果他们互相帮过忙，上述模式会明显缩水。公司总是在努力让消费者对自己的品牌产生情感联系，从而对品牌更加忠诚。如果公司在设计新款或升级款产品及服务时，能邀请消费者进行共同创作，让他们向公司提出自己对产品特点的要求，这场仗就算打赢了。

需要注意的是，在进行这类营销合作时，企业必须将消费者的意见框定为对公司的建议，而非观点表达，或对公司的期待。上述措辞差异看似微不足道，其实却是达成公司“身份统一”目标的关键。

提供建议，让人进入了联盟心态，刺激消费者认同公司的身份，与之产生连接。反过来，表达观点或期待，则让人进入了自省心态，把关注点放在自己身上。这些形式略有不同的消费者反馈，以及它们带来的截然不同的联盟或隔离心态，能给品牌的消费者参与度造成重大的影响。

有研究向一群来自美国各地的网络受访者展示了来自新快餐店“Splash!”的商业计划，快餐店希望通过强调菜品有益于顾客的健康来突出自己的竞争特色。受访者阅读了计划之后，按要求提供反馈。研究者要一部分人针对餐厅提出“建议”，另一些人则提出“观点”或表达对餐厅的“期待”。最后，受访者还要估计自己有多大的概率到这家餐厅去吃饭。提供建议的受访者想去就餐的愿望，明显高于提供其他两类反馈的受访者。而且，正如我们所期望的，提供建议是一种实现身份统一

的机制，支持餐厅的愿望更强，是因为受访者感觉跟品牌有了更深的关联。

调查里还有一项发现，在我看来是最确凿的“联盟”论据：被试认为，这三类反馈对餐厅的帮助是一样大的。所以，那些提建议的人感觉跟品牌的关联更紧密了，不是因为他们认为自己向餐厅提供了更多的帮助。事实上，提建议使得被试在发表对品牌的看法之前，进入了联盟心态，而不是隔离心态。我必须承认，这一发现让我感到很愉快，因为它表明，心理过程的先发制人特点，对这些提建议的被试产生了作用。

对我来说，这组结果还证明了在跟朋友、同事、客户面对面互动时，向对方征求建议这一做法的明智性。而且如果诚心诚意地用这种做法来寻找有用信息的话，它同时也是合乎伦理的。就算是跟上级互动，它应该也是有效的。当然，对它的潜在不利因素感到担心也合乎情理：让老板给自己提建议，他或许会认为你不称职、太依赖他人、欠缺安全感。虽然我认为这种担心合乎逻辑，但我同时也认为它是错的。因为，一如前文“上级评估与自己合作员工的贡献”的研究表明，理智、理性或者逻辑，并不能很好地捕捉到共同创作产生的影响，但联盟感可以，这是一种符合该情境的社会促进感，对你极为有利。小说家索尔·贝娄（Saul Bellow）曾经说过：“我们寻求他人的建议，往往也是在寻求同伙。”我只想为此补充一点科学证据：得到了建议，我们通常也就得到了同伙。要为项目找帮手，有什么做法比让对方负责更合适的呢？[\(55\)](#)

值得注意的是，关于不同类型的消费者反馈对随后这些消费者参与度的影响的调查发现，对消费者的建议提供高额奖金，会消除他们对该品牌的偏爱。研究者没有调查为什么会这样，但他们推测，意外的高额奖金让被试的关注点偏离了提建议的公共方面，转到了私人方面，在本例中就是被试在一次商业交换中所得的经济结果。此外，品牌利用共同创作等技巧，将消费者的身份认同与品牌联系起来，有着充分的研究证据支持。对品牌有着强烈认同感的消费者，在决定他们对该品牌的

态度和忠诚度时，更容易忽略关于该品牌产品存在瑕疵的信息。

跨群体大联盟

现在，是时候回顾一下我们从身心合一与行动合一中看到的最有利结果了，甚至还可以进一步做些更深入的思考。我们了解到，让人们参与相应的联盟体验活动，可以把选举推向对我们有利的方向，可以巩固来自公司股东及其客户的支持，可以保证战争中的士兵是并肩作战而不是仓皇逃窜，可以保护社群免遭毁灭。我们还发现，有两种联盟的体验，可以让玩伴、同学和同事彼此喜欢、互相帮忙、共同合作，可以让97%的家长无偿填写一份长长的调查问卷，甚至让人们在实验室里产生爱意。但这里有一个无法回答的问题：我们能否把这些经验教训应用到更大的舞台上，比如国家之间的世仇、长期的暴力宗教冲突、种族对立等？这些来自身心合一与行动合一的经验，能否让我们人类作为同一个物种联盟，合而为一？

这个问题回答起来很难，主要是因为这些令人痛苦、难以解决的分歧里，天然地带有大量的复杂因素。即便如此，哪怕是在这些矛盾重重的领域，我仍然相信，创造一种有联盟感的程序，也能为理想的变革造就适合的环境。虽然在所谓的挥手（hand-waving）理论^[56]里这个设想听起来很有希望，但如果就此以为这一理论能流畅地付诸实践，不免太过天真，因为实践里存在大量程序上和文化上的复杂因素。我们必须考虑到上述复杂因素，才能设计出最合适的促进联盟的程序并妥善执行。我想相关问题的专家也必会同意这样的看法，说不定这个主题还值得写一整本书来讨论。当然，我衷心欢迎专家们在这方面为我提供建议。

尽管最后一句话有些开玩笑的意味，但我们显然不能用一个简单的办法一刀切地对付牵连甚广、沉痾宿疾般的复杂问题，这一点再怎么强调也不过分。著名生物学家史蒂夫·琼斯（Steve Jones）观察到过一种与此有关的现象。他发现，科学家也会倚老卖老：到了某个年龄，科学家

他们就常常开始“为宏大问题大放厥词”，似乎他们在特定领域掌握的知识带给了他们足够的自信，于是敢于对这些专业边界的宏大问题指手画脚。史蒂夫的告诫似乎跟我此刻的处境有几分契合：首先，我已经到了他所描述的年龄阶段，其次，我要对自己全无专业知识的外交、宗教冲突、种族敌意领域给出结论。显然，我会在黑暗中“大放厥词”。

因此，如果我借助本章所给的经验教训去解决怎样合而为一的问题（如通过若干个影响力原则来展示），那就再好不过了。从人类大家族（而非部落）的视角来建立四海一家的“我们”感，来影响人们做出与大家族成员应有的回应也是最合适的，因为大家族的成员意识是人类大脑线路里本来就设置好的，“按一下就播放”。那么，我们就先从孩子的成长阶段和家长塑造他们的过程着手，接着转到有可能打动成年人的手法上吧。

如何让联盟更有效

对家庭影响的认识 在家里，有两种一定能成功的做法，可以让孩子把任何人（哪怕不是亲戚）都当成家人来对待。第一种做法是长期的共同居住。如果一个没有血缘关系的成年人（比如这家人的朋友）和全家人一起住了很长时间，那么，这个人常常会被称为某某“阿姨”或“叔叔”；如果是没有血缘关系的孩子跟全家人共同居住了很长时间，则会被贴上兄弟姐妹的标签。此外，在同一个屋檐下共同生活的时间越长，这个没有血缘关系的人能得到的血缘关系所特有的好处就越多，如来自家庭成员的自我牺牲式的帮助。第二种做法是孩子观察到双亲（尤其是母亲）照料无血缘关系的个体，此举会带来孩子类似血缘关系的行为。还记得我们这个时代两位最伟大的人道主义者杉原千亩和特蕾莎修女的自述吗？他们都说曾看到自己的父母无私地照顾家族之外的人。值得注意的是，无偿提供衣食住行等照顾行为通常仅限于家族成员。

行动启示 对那些希望提高孩子四海一家亲意识的家长来说，上述

发现也很有意义。为跨群体孩子提供长期住所这一做法固然令人钦佩，但对大多数家庭来说并不可行。领养或寄养父母必须达到的要求、付出的成本、承担的义务往往都太大了。

然而，在家里为跨群体孩子提供家庭式体验这一做法就容易管理多了。它可以分为两步来完成：首先，家长在自己孩子的课堂、运动队或舞蹈队中找到跨群体孩子，然后邀请他们（经其父母许可）到家里来玩或过夜。在我看来，邀请跨群体的孩子到家后，关键在于不要让他们感觉自己是在客人。邀请者家的孩子应该把来拜访的孩子视为自己家庭中的一员。

如果邀请者的孩子们有家务要做，来访的孩子也应被分派到任务。如果家里的衣物一般是妈妈负责清洗，而她注意到来访孩子的衣服上留下了在后院玩耍后蹭到的草渍，那她就应该帮忙清洗衣服。同样，她也应该当心来访孩子是否擦伤了皮肤，需要喷消毒剂、贴上创可贴。如果爸爸通常负责孩子们的体育活动，那么和所有的孩子们一起参与活动还不够，他应该成为每个孩子的老师：握住孩子们的小手，帮忙调整他们握球棒或高尔夫球杆的动作，让他们更好地击球；解释怎样投出橄榄球，好让球旋转出合适的弧度；展示怎样用假动作骗过守门员，把足球踢进球门。如果他的角色是进行和教授家庭或汽车维修，上述做法也同样适用。如果有机会为孩子们提供指导，家长就应该一视同仁，而不应该等到来访的孩子离开之后再去做。

当然，只要跨群体孩子来家里玩耍，就应该不停地重复这些做法。我认为最重要的一点是，别对来访的孩子偏心——有些好心的家长或许会这么做，以示自己不带偏见的榜样。相反，为了对来访的孩子们好，应该尽可能地把他们纳入家庭常规做法，而不是为他们额外“开小灶”。

如果邀请来访的家人共进晚餐，类似的建议也是适用的。如果是在室内吃饭，那家长应该等到访客到达之后再布置餐桌，方便访客像家人一样帮忙参与。如果是在后院野餐或烧烤，可以让访客帮忙摆放桌椅。

此外，用餐后的收拾和清理工作，应该让所有人都出力。

如果我妈妈还在世，她一定会这样说：“罗伯特，你怎么回事？这可不是待客之道。”或许在某种意义上她是对的。但我会回答：“妈妈，他们不是我们的客人，而是一些你希望他们立刻感到获得接纳，融入我们一家的活动，来自不同民族、种族或是有着不同身份的人。而且研究表明，一起布置餐桌、收拾厨余，还有在做这些事情时进行的闲聊，能让我们彼此感觉到更加团结。”

还有一件我一直在想却没说出来的事，因为妈妈教我不要“自作聪明”地和她争论。这件事是这样的：就算这么做的确如她所说有违晚宴标准礼仪，但款待是否恰当也并非问题的重点。邀请客人来的目的，是为自己的孩子灌输一种包含了各种各样其他人的宽泛“我们”感。我不会当着妈妈的面说出来，但我会想：“妈妈，你希望孩子记得你像对待外人一样招待客人，还是希望像家人一样对待客人？”而且，科学研究也支持我的建议，该研究表明，孩子能学习并效法成年人对社会群体成员的非语言信号。⁽⁵⁷⁾

我们对多元化邻里及友谊关系的认识 生活在多民族或多种族社区的人更容易产生全人类的身份认同，这让他们整体而言更乐于助人。此外，增加接触也一般会让他们对跨群体的其他人更友好，少有偏见。多元化的友谊也能带来类似的效果，它会让人们对朋友的民族或种族群体产生更积极、更支持的态度。这一效果不仅适用于多数群体成员，同样也适用于少数群体成员，如果这些少数群体成员有一个多数群体的朋友，那前者对多数群体的态度也会变得更积极。更值得注意的是，跨群体友谊提升了群体联盟感，也提升了双方的期待，使之认为跨群体成员之间的互动会是友好的。同时，跨群体友谊还能带来一种不为人知的间接影响：只要知道自己群体内的人有跨群体朋友，就会减少我们对另一个群体的负面情绪。

行动启示 住在多元文化社区里的孩子更有可能产生对人类整体的

身份认同，对于这样的发现，家长们应该怎样应对呢？哪怕非常重视这一思维的家长，也不太可能立刻搬到此类环境中去。但对重视这一思维的家长来说，把邻里多元化放到将来选择居住环境的需求列表中才更为合适。基于对这一思维的重视程度，家长们还可以选择它在清单里的位置是靠前还是靠后。

与邻里多元化相比，友谊多元化意味着更多的选择。其一跟之前的推荐做法一样：家长到学校里、体育赛事中或公园操场上去找一个跟自己孩子特别合适的跨群体小伙伴。邀请跨群体的孩子们去玩、来自己家过夜、参加生日聚会，都是推进孩子们友谊发展的自然做法，之后还可以邀请对方孩子的家长共进晚餐，为家长之间建立跨群体友谊奠定基础。这种成年人的联盟，可以通过在外面聚餐或者喝咖啡等一对一的方式加以巩固。

在家以外的场所见面很重要。第一个原因是，这种公开的见面也会让其他人观察到这段友谊，研究表明，这有助于降低旁观者的跨群体偏见，让他们更乐于自己也建立同类友谊。实际上，一对一的聚会越公开，其他人就越有可能受此影响去建立跨群体关系，并进而影响到更多的旁观者。新冠肺炎疫情发生期间，我们不幸地见证了群体传染数倍扩展的规律。不过，就公开的跨群体友谊而言，同一规律会向着有利于人类福祉的方向发展。

第二个原因是，与跨群体家长同伴（或任何跨群体成年人）安排一对一见面与其说是为了扩大友谊的影响，不如说是深化它。这种互动提供了另一种肯定能增进联盟关系的方式：相互的自我表露。我们在第2章讲过，互惠原则支配着各类行为，其一就是自我表露。如果谈话对象透露了一条个人信息，那么另一方几乎总会提供一条作为回报。如果采用前面介绍过的阿伦夫妇的36道问题程序，此类交流就有可能产生类似爱情的社会纽带。虽然一些研究者使用了这种方法来减少跨群体偏见，但按部就班地问完36道问题的做法并不适合在星巴克进行社交互动。我

们并不是在寻找不切实际的爱情反应。不过研究表明，哪怕是有限的自我表露，也有助于加深跨群体关系。做法很简单：如果你的目标是减少全世界跨群体地区中人们的敌意和偏见，那你可以从结交跨群体的朋友开始，模仿你身边跨群体的友谊，在公共场所跟跨群体朋友见面，并在谈话中透露一些个人信息以激发随后的对话。[\(58\)](#)

何种联系能带来联盟感 我们已经看到了各类通过行动合一（跳舞、唱歌、阅读、散步、工作）建立起来的联系，能够同步或合作地创造出一种广泛的“我们”意识。另一种不同类型的联系——意识到共同之处也能达到同样的效果。对那些希望激发他人心中联盟感的人来说，这类共同之处有一个特别有用的特点：只要简单地让人们意识到这些共同之处，就能将其调动起来。

采用这方面最有效的共性形式，即共同的身份往往能收到预期的效果。拉比卡利什通过向日本占领官指出双方“亚洲人”的身份拯救了自己的族人；一对夫妇发生争执，如果一方提醒另一方两人的共同伴侣身份，就更容易达成一致；想让美国的民主党人和共和党人提升对彼此的好感，只要提醒他们都是美国人即可。同样，犹太人和阿拉伯人读到双方在基因身份上的高度重合，就会减少对彼此的偏见和敌意，并更支持巴以和平努力。哪怕是重度自我导向型的精神病患者，也会在这种强烈偏爱的驱使下，对同属“我们”的群体成员表现出更多的关注。考虑到精神病患者是出了名地缺乏对他人的关心，那我们要怎样解释这一反常的结果呢？请记住，身份统一的过程，能让自我更多地与彼此相连的他人相融合。所以精神病患者的这种行为也并非“反常”。

其他形式的共同之处也有类似的作用。例如，传统上对立的群体会因共同的敌人而联盟。阅读了有关恐怖分子的介绍后，美国的白人和黑人会认为彼此没有那么不同了；以色列犹太人和阿拉伯人读到他们对癌症等疾病在易感性上类似时，也产生了同样的感受。更重要的是，这些变化是自动发生的，不需要认知反思。另一种共同之处，即基本的情感

体验发挥作用的途径则有所不同。一个群体的成员常常会通过对另一个群体的成员进行非人化，即否认后者完全拥有基本的人类情感和品质（如同情、宽恕、教养、道德和利他精神）来为自己对他们的偏见、歧视和虐待开脱和辩解。对此，我们可以用人类基本情感（它们为所有人所共同体验）的证据来驳斥这种非人化的观念。如果一个群体外的成员在同样的悲剧场面中和我们一起流泪、或在一个有趣的笑话里跟我们一起大笑，在同一起社会事件中跟我们同样愤怒，我们就很难不把他当成人来看待。当以色列犹太人得知，听到肇事逃逸现象的增加或工厂污水泄漏导致数千只海豚死亡等新闻时，巴勒斯坦人的愤怒程度与犹太人相当，那么此时以色列犹太人就会觉得巴勒斯坦人有了更多的人情味，也会对有利于巴勒斯坦人的政治政策表现出更多的支持。

最后一种值得强调的能产生联盟感的做法是采用对方的观点，即让自己站到对方的立场上去想象那个人在想些什么、有什么感觉，或正在经历些什么。在漫长的研究生涯中，我研究过促使人们帮助他人的因素，很快就知道了一点重要的事实：“如果你站到需要帮助之人的立场上，就很可能会对这些人伸出援手。”何以如此呢？我很快发现，把自己放在他人的处境中，可以提升自我—他人重叠感。因此，接受了澳大利亚土著居民观点的澳洲大学生，接受了跨性别观点的佛罗里达居民，都对有利于上述少数群体的政治政策变得更加支持了。有趣的是，知道有人在互动中努力接受自己的观点，会让我们与这些换位思考者产生更大的自我—他人重叠感，进而带来更多的喜欢和善意。显然，换位思考带来的结果是双赢的。研究者指出，借助换位方式来思考他人的人，其大脑中与思考自己相关的部分（腹内侧前额叶皮层）会出现更大程度的激活。

但有个问题需要注意。不同于与跨群体人士建立类似家庭、邻里和友谊的关系所带来的效果，由共同的敌人、大多数类型的共同身份认同、类似的情感反应以及换位思考等情境锻造出的关系在许多情况下无法发挥作用，就算能发挥作用，持续时间也不会太久。因为这种关系的

联盟目的，一般都与达尔文主义式的压力（推动本群体与其他竞争者斗争，获得生存空间，抢占优势）背道而驰。古罗马哲学家塞涅卡

（Seneca）曾说过：“我们是同一片海洋里的浪涛，同一棵树生长的叶子，同一座花园里盛开的芬芳花朵。”还有人认为此句出自巴哈伊教的创始人巴哈欧拉之口，这种情操毫无疑问是正确的，但它力量平平的推动力显然无法与自然选择的进化原理（后者往往断言的是正与之相反的真理）相提并论。每一道浪涛、每一片叶子和每一朵花，都在与其他的浪涛、叶子和花争夺资源、储备和生长空间，没有这些东西，它们就会萎缩甚或消失。

对支持联盟观的人来说，更糟糕的是，人性中还有另一种强大的特质把我们引向竞争和隔离：体验到威胁感。每当群体的福祉或声誉遭受威胁，我们会猛烈还击：贬低对立群体的价值与价值观甚至人性。因为大量证据表明，感知到威胁，会削弱促成联盟做法的效应。在我们这个时代，互相竞争的国家、种族和宗教实体都掌握着破坏性技术和毁灭性武器，能够对彼此造成大规模的恐慌和伤害，所以我们最好还是找到减少群体间敌意的办法，并将之导向和平。

行动启示 显然，要认同跨群体联盟这一使命的价值，就要面对一个得到了强大进化压力助推的可怕对手。它来自一种确保我们基因复制、存活与延续的无情拉力，在对我们意义重大的群体内部成员中，与我们复制相同基因的人有着极高的比例。科学分析表明，我们一般与那些跟自己有着家庭、友谊、地域、政治和宗教纽带联系的人有着更多的基因重叠。不足为奇，较之那些与我们基因关联度较弱的群体，我们更愿意帮助上述人群存活。面对自然选择这样强大的敌人，我们怎么能指望打赢这场旨在实现更大群体联盟的战斗呢？

也许我们可以再一次设法破解达尔文式的压力，并利用它的力量为我们服务。还记得在第1章中，一名修习日本柔道的女性，通过引导对手的力量（能量、重量、动量）打败更强壮对手的故事吗？借此启发，

我提议让跨群体成员更频繁地出现在我们的家庭、社区和友谊网络（这些网络已经进化为基因相似性这一可靠的线索，人们会本能地对该线索做出响应）中，以此构建联盟。要想改变进化压力的方向，把它引向联盟，我们的口号不应该是《星球大战》中的“愿原力与你同在”，而应该换成柔道版“愿对手的力量与你同在”。

那么，通过同样的普适方法，我们要怎样破解进化过程中通过共同敌人（我们易患癌症）、相对次要的共同身份（我们俩都是篮球迷）、人类共同情感（我家人人都为市长的决定感到愤愤不平）、换位思考（现在我已经设身处地站到了你的位置，能更好地理解你的处境了）所建立起来的连接、强化联盟时持续时间短、过于脆弱这一困局呢？如我们所见，尽管这些连接在当时或许有些影响，但它往往太过脆弱、容易遭到忽视，从而无法持久地指导行为。好在有一个因素可以提升其力量、增强其稳定性。这就是注意力的聚焦，它可以极大地促成有利的信念、价值观和选择。

一旦把注意力集中在某件事上，我们就会立刻认为它对我们更有意义。经济学家卡尼曼将这种现象称为“聚焦错觉”（focus illusion），即人们自动认为，如果自己在关注某件事，那此事必然事关利益。他甚至在一篇文章的标题中总结这一错觉说：“生活中没有什么事是非同小可的，除非你正在想它。”更重要的是，研究表明，如果焦点物品有着令人想要的特点，那它们会显得更重要而且更令人想要。所有认知错觉，都是由于通常而言运转良好的系统存在这样那样的缺陷而出现的。

卡尼曼聚焦错觉的研究，相关的数据来自下面这项研究，该研究考察了为什么放在商店货架中央的各品牌商品往往更容易被人购买。与左边或右边的商品相比，中间的商品得到了更多的视觉关注。此外，这种更多的关注，预测了购买决定。关于聚焦错觉的一般原理和后果，有证据表明，重要的东西会获得我们的关注，而我们关注的东西也会获得重要性。例如，在态度领域，研究者表明，人类的认知组织方式会让我

们会把最容易接触到的态度视为对自己最重要的态度。同样，任何我们可以轻易接触到的态度，在我们眼里都会显得更重要。甚至有证据表明，对一件消费品的视觉聚焦，会影响到大脑中支配感知价值的部位，进而增加个体对该物品的判断价值。

就聚焦错觉而言，为我们效力的系统在大多数时候都非常理性。对在特定情况下有着最重要意义和实用性的因素给予高度关注再合理不过了：黑暗中的异常声响，饥饿时闻到了食物的气味，以及看见首席执行官站起来要发言。这在进化上很有意义，因为如果做不到，就会在进化过程中显得适应不良。可它存在缺陷的地方在于，我们全神贯注的注意力并不总是会受环境最重要的方面吸引。有时候某些无关因素会把我们的注意力聚焦到某样东西上，使我们误以为它很重要。

有一项调查曾要受访的美国人说出两件自己眼里美国历史上“特别重要”的国家大事，他们提到“9·11”恐怖袭击的概率是30%。但在它10周年纪念日快到来前夕，随着媒体的大面积报道，人们对该事件的提名率飙升至65%。10周年纪念日过后不久，媒体对它的报道迅速减少，认为这场悲剧“特别重要”的受访者也随之减少——跌回了30%的水平。显然，媒体报道的力度，影响了旁观者对这一事件的关注，极大地改变了人们对它之于国家的重要性的感知。

尽管并非所有方法都获得了有效证明，但大量的研究表明，我们可以训练将注意力从造成威胁的目标转移到更积极的，或至少不那么可怕的目标上。除了训练自己把注意力从有时具有威胁性的外部群体身上转移开来，我们还可以用另一种方式来化解由此产生的焦虑。这需要注意力从焦虑本身转移到我们的力量上。当我们体验到这类威胁时，关键是要进行“自我肯定”，将注意力引导到自己珍视的某一方面，比如和家人、朋友和友谊网络的牢固关系；也可以是我们奖励的一种特质，比如创造力或者幽默感。它的作用是把我们的注意力焦点，从自己受到威胁的方面到伴随而来的防御反应（偏见、好斗、自我推销）上，重新导

向自身的价值方面和随之而来的自信反应（开放、镇定和自我控制等）。大量研究都记录了及时的自我肯定能够逆转外部群体威胁造成的负面影响。

影响力研究

一项研究对某在线家具店的访客进行引导，在浏览该店的产品之前，研究者使一半的访客浏览以蓬松云朵为背景的登录页面。这一注意力焦点使得访客把家具的舒适度视为一个较为重要的因素，他们也因此更偏爱购买更舒服的家具。另一半的访客并未表现出这种模式，他们认为价格更重要，更喜欢购买便宜的家具。为什么呢？因为他们看到的登录页面是一幅与成本相关的画面——一叠硬币。于是，在战术性引导之下，访客注意力所聚焦的概念，其权重发生了实质性的改变。最后，一项在线研究让被试把注意力放到他们自己的照片上，一些人看到的照片是他们现在的样子，另一些人看到的照片则是做过调整、年老后的样子。看到自己变老后照片的人，愿意为自己的退休金计划分配更多的资金。值得一提的是，如果他们看到的是其他人年老后的照片，情况就会不一样。年迈照片仅能影响到他们对自己将来经济状况的考量。这里，让受访者把注意力放在即将退休的自己身上时会提升其心中“照顾好这个人”的重要性。

如果新闻工作者、网页设计师和储蓄偏好研究者可以通过引导注意力，来提升人们对“9·11”恐怖袭击、家具卖点、退休账户资金的重要性的感知，为什么我们不能做一些类似的事情来促进联盟事业呢？为什么我们不能利用这一提升关注点重要性的力量来放大跨群体连接的感知价值呢？这意味着训练我们自己不再把视线放在对不同群体成员的怨恨、

敌意和偏见上，而是将它转移到共同的连接上。这种重新定向的行为，不光让我们的心态从割裂转到了连接上，还能消除群体间的分裂，放大聚焦之处的重要性，从而增强连接的力量。我是不是有点太天真了呢？谁知道呢。

如果你决心这么做了，首先，我们的使命中会获得一位强大的同伴。我们应该把聚焦视为自己的朋友、自己的燃料、力量的来源。接着，有证据表明，人们可以通过训练，将注意力从威胁性想法上转移到不那么险恶的想法上，从而减少对这些想法源头的焦虑。最后，每当我们遇到或仅仅听到不同群体时，都可以认真地尝试将注意力从割裂转移到连接上，要是它管用，那使命就完成了。但如果事实证明，我们把思路放在共同连接上的努力不成功（或许是因为，哪怕有聚焦带来的提升作用，这种联系仍然不够强大），其实我们仍有一张好牌可以打。我们只需要回想一下自己之前对跨群体“我们感”的尝试，把它作为我们的确对“我们感”有真实个人偏好的证据。这两种方式都可以让跨群体联盟在我们的自我概念中获得地位，也都可以发展跨群体联盟。

如何防范

大多数公司都有行为规范声明，员工入职之初便应该阅读这些声明，并在工作期间遵照执行。很多时候，这些行为规范是员工在组织中接受道德培训的基础。一项针对标准普尔500指数上市的制造业公司的研究发现，可以根据行为规范声明把公司分成两类：一类声明主要使用与联盟有关的语言，其条款用“我们”来称呼员工；另一类声明使用了更正式的语言，在条款中把员工称为“成员”或“雇员”。出人意料的是，在使用“我们”式语言来表达道德责任的组织中，人们就职期间从事非法行为的可能性明显更高。

为了理解其中的原因，研究者进行了8项系列实验，他们雇用被试阅读使用联盟语言（用“我们”来称呼员工）或非个人语言（用“成员”称呼员工）所写的道德行为准则指示，接着再让其执行一项工作任务。这里出现了若干让人大开眼界的现象。如果被试阅读的是用“我们”式语言所写的行为准则，就更有可能为获得绩效奖金而说谎或作弊，以牺牲组织为代价中饱私囊。另外两点发现提供了一种解释。第一，以“我们”为基础的措辞，让被试认为，这家组织不太可能进行监视并抓住违反道德政策的人。第二，接受这些指示的人，认为组织会对违规者较为宽容，最终原谅他们。

研究表明，企业的行为准则声明如果强调联盟，那么员工会更不诚实。替“我们”式群体成员的行为做开脱的倾向并不仅限于人类。在另一个例子中，如果年幼的黑猩猩是成年黑猩猩的血亲，后者就更能容忍前者偷取自己食物的盗窃行为。

我们在前面几章已经看到，每一条影响力原则都可能遭到奸商利用：赠送毫无意义的小礼品，迫使受赠人回报以更大的恩惠；利用统计数据说谎，让自己的产品营造出获得社会认同的虚假印象；伪造证书以暗示自己在某个主题上的权威性……联盟原则也不例外。一旦渔利者觉察到“我们”式群体里蕴含的联盟影响力（也即我们容易将同一群体成员的不当行为大事化小，给他们找借口开脱，甚至促成这些不当行为），他们便力争从人的这一原始行为倾向中牟取利益。会这么做的不单只有商业机构。我曾亲身遭遇过两件令人不快的事情，可以这么说，其他类型的工作单位也会产生同样的渔利者，以及对其加以纵容的联盟管理倾向。

其一是工会。这体现在，不管是警察工会、消防工会、制造业工会还是服务业工会，都愿意站在自己成员（哪怕是最恶劣的成员）一边。

工会不仅为下属成员提供了可观的福利，也为整个社会带来了各种好处，如改善安全法规、进行合理的工资调整、执行产假政策，以及扩大中产阶级队伍。但从在遵守职场恰当的道德行为方面看，它们存在明显的缺陷。就因为道德不端者是自己的一员，他们便为之提供保护、为之对外战斗，哪怕有清楚的证据证明这些道德不端者的违规行为极度恶劣、旷日持久。我家里有个现已过世的亲戚，就是这样一个典型的违规者。他在一家制造公司当焊工，经常装病旷工、干活偷懒、手脚不干净、涂改考勤，甚至伪造工伤记录。老板多次想要解雇他都不成功，反而被他嘲笑。他说，工会会费是他做过的最佳财务投资。他每一次做出有悖道德的事情后，工会都为他辩护。工会罔顾对错，仅仅是出于一种独特的道德义务——对工会会员保持忠诚而这么做。工会纵容他利用这种忠诚牟取私利，这种僵化态度总是让我感到不安。

其二是罗马天主教，其所辖教会神职人员的行为，对我有着类似的影响。我在天主教家庭长大，住在天主教社区，上天主教学校，参加天主教仪式，直至成年。虽然我已不再是身体力行的教会一员，但我与教会仍然保持着一种心理上的联系，我为教会的慈善拓展和减贫项目深感自豪。可同样的心理联系，也让我对教会领导层的一些做法（他们可耻地包庇了那些性侵孩子的豺狼牧师）感到羞愧。第一次传出教会高层管理不当的消息（赦免犯罪牧师，隐瞒其性侵行为，将之派到新教区，给他们第二次甚至第三次机会）时，我听到天主教内部的辩护者试图大事化小。他们主张，教会高层的神职人员也是牧师，宽恕罪孽恰恰是他们的职责。因此，他们只是在做与其宗教职责相符的事情。我知道这并不是真正的理由。教会当局不仅宽恕了这些性侵行为，还隐瞒了有关此事的消息。出于保护“圈内人”的原因，他们掩盖真相，使得惨剧再次发生在其他儿童身上，让这些孩子受到惊吓，承受永久的伤害。他们落入了道德的阴沟，其行为只能以“我们是一伙的”来开脱。

有证据证明，纵容组织内部的道德不端行为会损害自身的经济利益，进而也证明了对这种行为采取不容忍政策是极为明智的。同事和我

把此类后果称为“组织不诚实的三重肿瘤结构”。我们已经讨论过，如果一家组织经常允许员工使用欺骗手法（对同事、客户、客户、股东、供应商、分销商等），那它将承受三重代价高昂的内部后果：员工绩效下降、员工流动率居高不下、员工欺诈和渎职行为盛行。此外，这些后果会像恶性肿瘤一样，不断增长、扩散，并逐渐侵蚀组织的健康和活力。

在以“我们”为基础的工作群体（如商业单位、工会、宗教组织等各种联盟）中，有没有可能阻止成员的假公济私行为呢？我相信可以，但这需要组织管理者采取3个步骤：首先是要认识到，腐败者以为自己正受到“我们”式群体的保护，这类群体往往愿意为违背道德行为规范的成员开脱；其次，组织管理者要向所有相关人士宣布，在我们这个“我们”式群体里，不会出现这样的宽大处理；最后，组织管理者要确立起“不容忍”政策，凡违规者均以除名处理。

这种道德行为承诺，应该在什么时候、什么地方做出呢？应该从一开始就这么做（放在组织行为准则声明中，凡加入者均应恪守），之后要经常团队会议上对道德和不道德行为做出定义，反复重申和解释不容忍政策，并坚决执行。有人对第7章所述的研究做了综述，该综述指出，对重要价值观做出这种强化的书面承诺，可以从实际上增强价值观的推进落实。我曾在偶然中亲身体会过这种书面承诺是多么有效。

我曾在法律案件中担任专家证人，这些案件主要包括产品制造商的欺骗性广告和营销行为。我做了三年就不做了，因为这个工作的紧迫性太强了。在形成初步的意见书之前，我常收到成箱成箱的书面材料：陈述、证词、申请书、证据报告和之前的法庭判决。接着，我需要将这份意见书提交给法庭，随即在正式作证过程中接受对方一群律师的质询，同时，我需要就这一意见书展开辩护。在作证日来临之前，我需要与雇用我的律师团队成员多次会面，打磨并完善我的意见书，以期发挥最大影响。

在这些会面中自然而然发生的事情，给我带来了一个棘手的问题：

道德。我成了一个联盟的“我们”式群体的一员，该群体有着明确的目的：在官司里打赢另一方的律师和专家证人团队。合作期间，我将和同事们建立友谊，在讨论中理解他们的知识技能，逐渐了解彼此在食物和伴餐音乐上的共同品位，并在席间喝酒时（一般是酒过三巡之后）通过互惠式自我表露来形成更亲密的关系。在准备意见书期间，我知道我的意见是我方军火库里的一种重要武器，会拿来对付对手。我的意见对我方情况支持越充分，我越是能自信满满地发表意见，对我方就越是有利。

虽然人们很少明确地表达这些情绪，但我立刻就明白该怎样提高自己在团队里的地位。我可以尽量在声明中自信地强调证据（包括研究文献）与我方论点相吻合那部分内容的重要性，同时淡化与我方论点相悖的那部分内容的价值，这样一来，人们就会越来越认为我对这个小团体及其目标忠心耿耿。

我从一开始就感受到了这个位置存在道德冲突，而且给我带来了很大压力。身为科学家，我有义务对所见证据做最为准确的呈现，此外，在对证据的分析中，我也有义务以最真实的方式表明我对打赢官司的信心。可与此同时，我又是“我们”小群体中的一员，在道德上（基于职业责任）有义务为客户提供最有利于取胜的案例。尽管我不时向客户提及自己历来秉持的科学诚信价值观，但我始终拿不准，他们到底有没有充分地考虑这一点。过了一阵，我想最好还是在一份正式声明中，向对客户（也向我自己）明确地表示我对这些价值观（而非客户的价值观）所做的承诺。我开始在意见书的末尾加上一段话，指出我的观点是基于雇用我的律师们提供给我的信息和论证得出的，如果碰到新的信息和论证（哪怕来自对方律师所提供），我是有可能修改自己的观点的。这段话立刻带来了立竿见影的效果，队友们觉得我没那么忠诚了，可我觉得我强化了自己喜欢的角色。

这段话还在一起法律案件中带来了一个意想不到的好处。在该案件

中，我认为一家公司对其产品健康特性的广告宣传存在误导性。该公司利润丰厚，有足够的资源聘请律师团队，他们的首席律师可能是我碰到过的最老练的“审讯员”。在一份包含了我初步意见书的证词中，我的任务是捍卫自己的立场，而他的任务是想方设法地贬低我的观点、可信度和诚信度。他的批评极其尖锐，我必须时刻警惕，以避免他的锋芒。很奇怪，我还挺享受这种互动的，因为这是一种智力上的挑战，可他却突然做了一件我意想不到的事情。他提醒我，我曾写过一种叫“登门槛”（见第7章）的影响力手法，我还写过一篇研究文章，说屋主们答应在窗户上贴一块倡导安全驾驶的小小警示牌，这使得他们在过了几个星期之后，有更大可能性去做一件原本不可能去做的相关事情——答应在草坪上放置一块硕大的广告牌宣传同一内容。他问我，这是否意味着我认为，对一个做法的最初承诺，比如在窗户上贴一块安全驾驶的警示牌，会促使人在这一做法上采取更极端的立场。我回答说是的，他立刻扑了上来，举着我的初步意见书说：“在我看来，这份声明就像是你做出的最初公开承诺，用你自己的话来说，它会使得你无论如何都要与之严格保持一致，甚至采取更为极端的立场。那么，为什么我们要相信你现在说的话呢？显然，西奥迪尼教授，你已经在窗户上贴了警示牌了。”

我深为所动，朝椅子里向后一靠，承认道：“这可真是高明！”他挥了挥手，不理睬我的赞美，还催促我回答，脸上露出一抹猎人看到新猎物在陷阱中挣扎的微笑。幸运的是，我并没有被困住。我请他阅读意见书的最后一段话，表明我会接纳新的信息，并愿意做出调整，而不是刻板地保持前后一致。“实际上”，等他读完这段话抬起头来时，我告诉他，“这才是我窗户上的警示牌。”他倒没有靠回椅子，也没有大声说出来，但我几乎可以肯定，我看到他喃喃自语道：“这可真是高明！”

我很高兴他这样想，但其实这段话并不是为了反驳那天他对我观点的攻击，而是为了解决我身为专家证人所面临的另一个问题：来自“我们”式群体的内部压力（随着我们友谊的发展，我内心越来越大的压

力），使得我想要履行对群体尽忠的道德义务的同时呈现真相。我写这段话，就是在尝试以书面形式让所有人知道，我不会纵容自己朝着这个方向前进。我想这算是一次成功的尝试。

如果组织希望从以“我们”为基础的集体文化中获益（如更和谐、更富合作精神），但又不愿负担组织内部渔利者所带来的腐败成本，那上述故事能给我们什么启发呢？在组织的行为准则声明中，应该放上一块自我承诺的“窗户警示牌”，设定不容忍条款，说明一旦查明员工存在严重违背准则或多次轻微违规并屡教不改的行为，将做解雇处理。不容忍政策的基本原理，应当从与道德文化相关的职场满意度和工作自豪感出发，更重要的是，从保持工作场所联盟感的诚实愿望出发。为什么要用这个基本原理呢？因为，通过呼吁对“我们”的需求，将组织从“我们”式集体感所带来的缺陷中解救出来……那可是真的高明。

我不记得自己从前曾提倡过要在人际交往中冷酷无情，但在组织中，特别是在注重联盟的组织中，对违背道德准则的员工进行解雇的不容忍政策，看似冷酷无情，其实从我们的发现来看，这么做非常明智。当然，犯错是人之常情，应该给人第二次机会。对于对犯规者持宽容态度的人，我也很能理解。一如莎士比亚在《威尼斯商人》里提到怎样处置违背道德者时写道：“慈悲不是出于勉强，它像甘霖一样从天上降下尘世。”但具体到员工里的不道德行为，我（和莎翁不同）已经看到了大量记录它一系列腐蚀性和传染性后果的研究。如果我们低估这些后果，不免有些愚昧。

本章小结

- 如果人们将某人视为自己团队中的一员，就更容易答应此人提出的请求。和他人产生“我们”感，关键在于双方有着共同的身份，即人们用来定义自己和所属群体的部落式类别，如种族、民族、国籍、家庭以及政治和宗教联盟。

- 针对“我们”式群体所做的研究，有三条一般性结论。第一，这些群体的成员更关注内部成员（而非群体之外人士）的成果和幸福。第二，“我们”式群体成员还会利用其他成员的偏好和行动来指导自己，以确保群体联盟。第三，这种“联盟”倾向是在进化过程中出现的，它有利于“我们”的群体，最终也有利于我们自己。这三种“常量”经常出现在包括商业、政治、体育和个人关系在内的广泛领域。
- 和他人产生的“身心合一”感，是带来“我们”感的第一个基本因素。这种感知由是血缘共性（基因重叠度）和地域共性（包括家园、乡土和地区）产生的。
- 行动合一的体验，是带来联盟感的第二个基本因素。共同的音乐体验，能让人们行动合一，并感受到随之而来的联盟。其他途径还包括多次互惠交换、一起吃苦和共同创造。
- 利用身心合一和行动合一的联盟效应提高人类“合而为一”的概率，是有可能做到的。这需要人们主动和群体外成员在自己家里分享家庭体验，在社群里分享邻里体验，在社交互动中分享友谊体验。
- 包括国家认同、共同敌人、共同情感经历和共同观点等其他类型的联系，也能带来与群体外成员的联盟感。不过，通过此种方式带来的联盟感大多为时甚短。然而，把注意力多次集中到这些联系上，能提升其感知重要性，或许能延长其效力。

Influence

第9章

即时的影响力

每时每刻，我都变得更好了。

——埃米尔·库埃（Émile Coué）

每时每刻，我都变得更忙了。

——罗伯特·西奥迪尼

NEW AND EXPANDED

20世纪60年代，有个叫乔·佩恩（Jon Pyne）的人主持过一出相当有名的脱口秀节目，整个加利福尼亚地区都可收看。这个节目别具一格的地方在于，佩恩总是会用一种刻薄又挑衅的态度对待嘉宾。嘉宾们绝大多数都是渴望曝光的娱乐圈人士、快要成名的新人，以及各种社会组织的代表。主持人的粗暴采访风格，是为了刺激嘉宾们，怂恿他们展开争论，让他们在慌乱中承认自己做过的尴尬事儿。简单地说，就是要让他们显出一副蠢样。佩恩经常在介绍嘉宾的时候就接二连三地攻击对方的信仰、天赋或外貌。有些人说，佩恩尖酸刻薄的个人风格，一部分原因在于他做过截肢手术，饱受生活的磨难；而另一些人说，事实并非如此，佩恩只是天生牙尖嘴利罢了。

一天晚上，摇滚歌手弗兰克·扎帕（Frank Zappa）做了这个节目的嘉宾。那时候毕竟是20世纪60年代，男人留长发还不太常见，也引人争议。佩恩刚刚介绍完扎帕，不等他落座，下面的对话就噼里啪啦地开始了：

佩恩：“我猜，你的头发这么长，肯定是个姑娘。”

扎帕：“我猜，你有条木头腿，肯定是张桌子。”

原始的自动反应

佩恩和扎帕的这出唇枪舌剑的对话里除了有我最喜欢的即兴发挥，亦例证了本书的一个基本主题：很多时候，我们在对某人或某事做判断的时候，并没有用上所有可用的相关信息；相反，我们只用到了所有信息里最具代表性的一条。这条孤立的信息虽说通常都能给我们以正确的指导，但它也能让我们犯下显而易见的愚蠢错误。这样的错误，一旦遭到其他聪明人利用，就会让我们显得又笨又呆，甚至带来更糟的结果。

然而，还有一个复杂的平行主题贯穿本书：尽管只靠孤立数据容易做出愚蠢的决定，可现代生活的节奏又要求我们频繁使用这一捷径。在第1章的开头部分，我们曾将这一捷径与低等动物的自动反应相比较，单独的一点刺激特征，如“叽叽”的叫声、一撮红色的胸羽或特定序列的闪光，就能触发一整套复杂的行为模式。这些低等动物必须频频依赖环境中的孤立刺激，原因在于它们的智商有限。它们的小脑袋瓜无法将环境中的所有相关信息进行登记和处理。因此这些物种对信息的某个方面进化出了极度的敏感性。一般来说，这类片面信息足以提示它们做出正确的反应，所以这套系统基本上是极为有效的。每当雌火鸡听到“叽叽”的声音，就会自动搬出一套恰当的机械化母性行为。因为这么做能让它节省自己有限的脑力，应对每天必须面对的其他环境和选择。

当然，在这方面，我们的大脑有着比雌火鸡或者任何其他动物都更为有效、复杂的机制。考虑多方相关信息并据此做出正确的决定，我们在这方面的能力是其他动物比不了的。事实上，人类成了地球上的支配物种，靠的就是这种信息处理上的优势。

不过，我们的能力也是有限的。况且为了追求效率，有时候我们也必须放弃耗时、复杂、需整体把握的决策过程，转而使用更简单、原始、由单一特征触发的响应方式。例如，在判断是否答应请求者要求的时候，我们经常是只注意到了相关信息中的一条。在前面的章节里，我们已经探讨了触发人做出顺从决定的几种最常用的单一信息。它们之所以最为常用，完全是因为它们的可靠性高，一般都能指引我们做出正确决定。这就是为什么我们会这么频繁地利用互惠、喜好、社会认同、权威、稀缺、承诺与一致、联盟等方面的因素自动做出顺从决定。究其本源，上述每一方面的因素都是极为可靠的线索，能提示我们在何时说“是”要比说“不”更恰当。

在没有意愿、没有时间、没有精力或没有认知资源对情况进行全面分析的时候，我们最容易使用这些孤立的线索。倘若我们正赶时间，正

处于压力大、不确定、不在乎、心烦意乱或心力交瘁的时候，我们往往会把焦点放在一些片面的信息上。在这类环境下做决定，我们通常使用的都是原始而必要的“单一可靠证据”法。这一切带出了一个令人不安的结论：靠着成熟而精密的大脑，我们建立了一个信息繁多的快节奏复杂世界，使得我们不得不越发依赖类似动物的原始反应方式来应对它。

现代自动反应

英国经济学家、政治思想家、科学哲学家约翰·斯图亚特·穆勒（John Stuart Mill）离世大概有140多年了。他的去世（1873年）是一件历史大事，因为他是最后一个享有“掌握世上已知的一切知识”美誉的人。如今，这样的说法变得很可笑了。经过世世代代的逐步积累，人类的知识如同滚雪球一般，进入了一个靠惯性驱动、成倍扩张的大爆炸时代。在我们现在生活的世界上，大部分信息的存在历史不超过15年。在某些科学领域，如物理，据说知识的数量每8年就会翻一番。科学信息爆炸并非仅限于分子或量子物理化学等神秘领域，而是囊括了与我们日常生活息息相关的一切知识领域：健康、教育、营养等等。更重要的是，这种快速增长很可能会持续下去，因为研究者还在源源不断地把最新发现变成每年200多万篇的科学期刊文章。

EBOX | 线上影响力

“你愿意使用这部手机吗？”

“我愿意……随时随地带着它。”



作者点评：我们的数字设备不仅拥有前所未有的信息力量，还会使人上瘾。调查显示，人们平均每天查看手机的次数超过100次，84%的人说他们“没了手机一天都过不下去”。

除了科学的迅猛进步，与普通人的生活相关的东西也在飞速变化。每年的盖洛普民意调查显示，公众关心的事件越来越多元化，但关心的时间却越来越短。此外，我们的出行更多、更快捷了；我们更频繁地搬迁到新的住宅，因为新住宅的兴建和拆迁也越来越快了；我们接触更多的人，但跟他们之间的联系却越发短暂；在超市、汽车展厅、购物中心，我们会碰到各式各样的选择，看到前一年闻所未闻的产品，但到下一年，这些产品说不定就销声匿迹了。在论及当代文明的时候，最常用的描述词就是新颖、短暂、多元、速度惊人。

排山倒海的信息和选择是飞速的技术进步带来的。这当中首先受到影响的要数我们收集、存储、检索和传递信息的能力。起初，享受这类进步成果的仅限于大型组织，如政府机关或强大的企业。随着电子通信和数码技术的发展，哪怕个人也能接触到数量惊人的信息。有线和卫星

系统的普及，成了信息进入普通家庭和个人的渠道。如今，一部手机的信息能力，就已经超过了几年前整个大学系统的信息能力。

但有一点务必注意，当今时代，通常被叫作“信息时代”，从来没有人称它为“知识时代”。信息并不直接转化为知识，它首先必须经过处理，如获取、吸收、理解、整合和保留。

捷径应受到尊重

因为技术的进化速度远远快于我们，我们处理信息的天然能力有可能越来越难于应对当代生活中繁多的变化、选择和挑战。我们越来越频繁地发现自己陷入了跟低等动物一样的处境之中：外界环境的错综复杂超出了我们心智器官的处理能力。当然，也有不同的地方，低等动物的认知能力从来就比较欠缺，可我们的问题却类似作茧自缚：是我们自己创造了一个太过复杂的世界，最终却搞得自己都应付不了。我们新产生的这种缺陷，跟动物长久以来的缺陷一样：在下决定的时候，我们越来越难于对整个局面加以全盘考虑了。为解决这种“分析瘫痪”问题，我们只好更多地把注意力放到环境中通常靠得住的单一特点上。

倘若这些单一特征确实可靠，那么依靠聚焦注意力、自动响应特定信息的快捷方法做出决定并没有什么根本上的错误。问题在于，有些因素会让这些通常靠得住的线索误导我们，让我们做出错误的行为和糟糕的决定。正如我们所见，某些说服专业人士耍的手段就是误导因素之一。这些人利用我们不假思索的机械反应来获利。随着现代生活的节奏越来越快，复杂程度越来越高，假如我们利用快捷响应的频率也越来越高，那么可以肯定，别人对我们耍这类手腕的频率也会越来越高。

既然我们的捷径系统可能遭到密集的攻击，那么我们该怎么做呢？我认为，光回避是不管用的，还要进行有力的还击。但这里有一条重要的先决条件：倘若说服专业人士公平公正地利用我们的捷径响应方式，

我们就不应该把他们看成是敌人，事实上，他们是我们的盟友，有了他们，我们能更方便地开展高效率、高适应度的生意往来；只有那些通过弄虚作假、伪造或歪曲证据，误导我们做出快捷响应的人，才是正确的还击目标。

我们可以从最常用到的捷径里举个例子。根据社会认同原则，我们往往会效仿跟自己类似的其他人的做法。这么做是合情合理的，因为大多数时候，一种行为在特定环境下能得到众人的认同，往往说明它管用而且恰当。因此，广告商采用真实可靠的信息而不是欺骗性的数据告诉我们某个品牌的牙膏销量最大，这的确能说明这种产品质量好，我们很可能会喜欢上它。假设我们正在市场上寻找优质牙膏，其实就可以依赖“使用人数”这一单一信息来判断是否要购买它。这种策略给我们指明正确方向的可能性高于出错的可能性，我们可以把认知精力节省下来，去应付生活里铺天盖地的其他信息，做出更合理的决定。在广告商的帮助下，我们有效地利用了这一高效策略，所以，此时广告商不是我们的敌人，而是我们的合作伙伴。

可要是说服专业人士给我们虚假信号，试图刺激我们的捷径反应，那事情就完全不一样了。比方说，广告商拍摄了一系列现场“随机访谈”广告，让若干演员假扮普通市民盛赞该产品，试图营造出这种牙膏大受欢迎的假象，那么这样的广告商就是我们的敌人。因为在此受欢迎的证据是伪造的，我们、社会认同原则、我们对它的捷径反应都遭到了利用。在先前的章节中，我曾建议，凡是打虚假“随机访谈”广告的产品，我们都不要购买。不仅如此，我们还要写信向制造商投诉，并建议他们撤换广告代理公司。同时我还建议，只要碰到说服专业人士用这种方式滥用社会认同原则或其他影响力武器，我们都应采取这一强硬立场。我们应该拒绝收看使用“罐头笑声”的电视节目；如果我们看到酒保在小费罐里先装上一两块钱，以诱使人给他小费，那么我们就应当给这个酒保一分钱的小费；如果我们看到夜总会门外排着长长的人龙，可走进去一看，里头还有大把的空位，外面的长队只是为了制造生意兴隆

的假象来打动路人的，那么我们应当立即离开，并把这么做的原因告知外面还在排队的人。简而言之，我们要采取一切合理的方法，如抵制、威胁、对峙、谴责、抗议，来报复以刺激我们的捷径反应为目的的虚假信号。

我并不认为自己是个天生好斗的人，但我却建议大家主动还击。这是因为从某种意义上来说，我已经跟那些剥削者宣战了。确切地说，我们人人都应宣战。然而，有必要认识到，我们采取对抗做法，并不是因为他们想赚钱。说到底，我们人人都想赚钱。关键在于，他们赚钱的方式威胁到了我们捷径的可靠性，是真正的背叛，这才是我们无法容忍的地方。我们必须依赖可靠而合理的捷径和首选规则来应对现代生活的繁忙节奏。这不是什么奢侈品，而是不折不扣的必需品。随着生活节奏的加快，它们还会变得越来越重要。这就是碰到有人为谋求私利而误导这些规则时，我们应当还以颜色的原因，我们希望这些规则尽可能地有效。倘若投机客习惯性地耍手段利用它们，我们只好对其弃之不用，这样一来，我们也就没法卓有成效地应对日常生活的繁多抉择了。我们实在不能对此种局面听之任之，必须奋力一战来捍卫这些宝贵的捷径。这一仗，胜败的赌注太大了。

READER'S REPORT | 读者报告

致西奥迪尼先生：

不久前，我在一家电器商店买东西时，发现一台高质量的大屏幕电视机正在打折，价格很是诱人。我原本并不打算买新电视机，但这优惠的售价外加强势的产品评价，让我禁不住停下来翻了翻相关的宣传册。一位名叫布拉德的销售人员走过来说：“看得出你对这台电视机很感兴趣，我大概能猜到是为什么。它质量好，价格又很优惠。但我必须告诉你，这是最后一台了。”这立刻激起了我的兴趣。

接着他又告诉我，一位女士刚刚打电话来，说自己可能下午就来买。我研究了一辈子的说服与顺从，知道他是在对我使用稀缺性原则。

但这并不管用。**20**分钟后，我用购物车推着“奖品”走出了商店。西奥迪尼博士，我对布拉德的稀缺性故事做出这样的反应，是不是太傻了？

致广大读者：

读者们想必已经发现，报告中的罗伯特就是我，所以我对这个问题有一个知晓内情的视角。他是否应该产生受骗的感觉，取决于布拉德告知的稀缺性相关信息是否准确。如果是准确的，罗伯特应该感谢布拉德提供的信息。想想看，要是布拉德没有把真实情况告诉罗伯特，罗伯特回家想了想，当晚又回到店里打算把电视机买下，结果却发现最后一台已经卖掉了，他肯定会对销售员大发雷霆：“怎么搞的！我走之前你怎么不告诉我这是最后一台？”

现在，假设布拉德并未提供真实信息，与电视机相关的稀缺条件是他编造出来的。等罗伯特走掉之后，他去库房拿出一台相同型号的电视机放在货架上，借助同样的故事再把它卖给下一位顾客。（顺便说一句，几年前，百思买的员工就是这样做的，但被人逮了个正着）那对罗伯特来说，布拉德就不再是一个有价值的线人，而是个可恶的奸商了。

真是这样吗？罗伯特决定弄个明白。第二天早上，他又来到这家商店，想看看是否还有同款电视机在出售。没有了。布拉德对他说的是实话。于是，罗伯特回去就给这家店，尤其是布拉德写了一条高度赞赏的评价。如果布拉德撒

谎，这条评价将给予同样强烈的谴责。在面对影响力原则时，对那些企图伤害我们的人，我们应始终不遗余力地加以谴责，而对那些力争助我们一臂之力的人，我们也该投桃报李。

本章小结

- 当今有别于以往任何时代，技术的重大进步使得信息蓬勃发展，选择日益繁多，知识趋向爆炸。面对这种如雪崩般袭来的变化和选择，我们不得不进行调整，我们做决定的方式出现了一个根本性的变化。虽然我们在任何情况下都希望做出周到的决定，但当代生活变化多端，节奏又快，常常逼得我们没有条件对正反两方面的相关信息进行详尽分析。无奈之下，我们越来越多地采用另一种决策方法，一种便捷的方法，只根据基本可靠的单一信息做出决定。因此，在本书中，我们讨论了最可靠也最为人所接受的此类单一顺从触发因素：互惠、喜好、社会认同、权威、稀缺、承诺与一致、联盟。
- 由于我们当今社会的认知过载倾向日益严峻，抄捷径下决定的做法也随之普遍化。说服专业人士只要在提要求时加入一两个影响力触发因素，就很可能成功实现他们的目的。从业者使用这些触发因素，不一定是在利用我们。只有当触发因素不是环境中自然蕴含的特征，而是从业者捏造出来的时，才意味着赤裸裸的剥削。为了维护快捷反应的有益特质，必须采用各种恰当的方式来反抗这种弄虚作假。

- Aaker, D. A. (1991). *Managing brand equity*. New York: Free Press.
- Abrahamse, W., & Steg, L. (2013). Social influence approaches to encourage resource conservation: A meta-analysis. *Global Environmental Change*, 23, 1773–1785.
- Achen, C. H., & Bartels, L. M. (2017). *Democracy for realists: Why elections do not produce responsive government*. Princeton, NJ: Princeton University Press.
- Adarves-Yorno, I., Haslam, S. A., & Postmes, T. (2008). And now for something completely different? The impact of group membership on perceptions of creativity. *Social Influence*, 3, 248–266.
- Agerström, J., Carlsson, R., Nicklasson L., & Guntell L. (2016). Using descriptive social norms to increase charitable giving: The power of local norms. *Journal of Economic Psychology*, 52, 147–153.
- Aggarwal, P., Jun, S. Y., & Huh, J. H. (2011). Scarcity messages. *Journal of Advertising*, 40, 19–30.
- Albarracin, D., & Wyer, R. S. (2001). Elaborative and nonelaborative processing of a behavior-related communication. *Personality and Social Psychology Bulletin*, 27, 691–705.
- Alison, L., & Alison, E. (2017). Revenge versus rapport: Interrogation, terrorism, and torture. *American Psychologist*, 72, 266–277.
- Algoe, S. B. (2012). Find, remind, and bind: The functions of gratitude in everyday relationships. *Social and Personality Psychology Compass*, 6, 455–469.

Algoe, S. B., Gable, S. L., & Maisel, N. (2010). It's the little things: Everyday gratitude as a booster shot for romantic relationships. *Personal Relationships, 17*, 217–233.

Algoe, S. B., & Zhaoyang, R. (2016). Positive psychology in context: Effects of expressing gratitude in ongoing relationships depend on perceptions of enactor responsiveness. *Journal of Positive Psychology, 11*, 399–415.

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics, 95*, 1082–1095.

Allcott, H., & Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review, 104*, 3003–37.

Allison, S. T., Mackie, D. M., Muller, M. M., & Worth, L. T. (1993). Sequential correspondence biases and perceptions of change. *Personality and Social Psychology Bulletin, 19*, 151–157.

Allison, S. T., & Messick, D. M. (1988). The feature-positive effect, attitude strength, and degree of perceived consensus. *Personality and Social Psychology Bulletin, 14*, 231–241.

Al Ramiah, A., & Hewstone, M. (2013). Intergroup contact as a tool for reducing, resolving, and preventing intergroup conflict: Evidence, limitations, and potential. *American Psychologist, 68*, 527–542.

Ames, D. L., Jenkins, A. C., Banaji, M. R., & Mitchell, J. P. (2008). Taking another person's perspective increases self-referential neural processing. *Psychological Science, 19*, 642–644.

Ampel, B. C., Muraven, M., & McNay, E. C. (2018). Mental work requires physical energy: Self-control is neither exception nor exceptional. *Frontiers in Psychology, 9*, 1005.

Anderson, M. (2014). *After Phrenology: Neural Reuse and the Interactive Brain*. Cambridge, MA: The MIT Press.

Anderson, E., & Simester, D. (2003). Mind your pricing cues. *Harvard Business Review*, 81, 103–134.

Anderson, S. M., & Zimbardo, P. G. (1984). On resisting social influence. *Cultic Studies Journal*, 1, 196–219.

Andor, M. A., & Fels, K. M. (2018). Behavioral economics and energy conservation-a systematic review of non-price interventions and their causal effects. *Ecological Economics*, 148, 178–210.

Antonovics, K., & Knight, B. G. (2009). A new look at racial profiling: Evidence from the Boston Police Department. *Review of Economics and Statistics*, 91, 163–177.

Aramovich, N. P., Lytle, B. L., & Skitka, L. J. (2012). Opposing torture: Moral conviction and resistance to majority influence. *Social Influence*, 7, 21–34.

Arbuckle, N. L., & Cunningham, W. A. (2012). Understanding everyday psychopathy: Shared group identity leads to increased concern for others among undergraduates higher in psychopathy. *Social Cognition*, 30, 564–583.

Arizona Republic (1999, March 7). For women, all's pheromones in love, war, E19. Armstrong, J. S. (2010). *Persuasive advertising*. London: Palgrave Macmillan.

Arnocky, S., Bozek, E., Dufort, C., Rybka, S., & Herbert, R. (2018). Celebrity opinion influences public acceptance of human evolution. *Evolutionary Psychology*.

Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Self-relationships as including other in the self. *Journal of Personality and Social Psychology*, 60, 241–253. Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and Social Psychology Bulletin*, 23, 363–377.

Aronson, E. (1975, February). The jigsaw route to learning and liking. *Psychology Today*, 43–50.

Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *Journal of Abnormal and Social Psychology*, 59, 177–181.

Aronson, E., Stephan, C., Sikes, J., Blaney, N., & Snapp, M. (1978). *The jigsaw classroom*. Beverly Hills, CA: Sage.

Aronson, E., & Tavis, C. (2020, July 20). The role of cognitive dissonance in the pandemic. *The Atlantic*.

Ashmore, R. D., Ramchandra, V., & Jones, R. A. (1971, April). *Censorship as an attitude change induction*. Paper presented at the meeting of the Eastern Psychological Association, New York, NY.

Ashokkumar, A., Galaif, M., Swann, W. B. (2019). Tribalism can corrupt: Why people denounce or protect immoral group members. *Journal of Experimental Social Psychology*, 85.

Asimov, I. (1975, August 30). The Miss America pageant. *TV Guide*.

Atalay, A. S., Bodur, H. O., & Rasolofoarison, D. (2012). Shining in the center: Central gaze cascade effect on product choice. *Journal of Consumer Research*, 39, 848–856.

Aune, R. K., & Basil, M. C. (1994). A relational obligations approach to the

foot- in-the-mouth effect. *Journal of Applied Social Psychology*, 24, 546–556.

Australian. (2009, December 11). Coin by coin, B14.

Ayres, I., Raseman, S., & Shih, A. (2013). Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *Journal of Law, Economics, and Organization*, 29, 992–1022.

Bai, X., Ramos, M. R., & Fiske, S. T. (2020). As diversity increases, people paradoxically perceive social groups as more similar. *Proceedings of the National Academy of Sciences*, 117, 12741–12749.

Bailenson, J. N., & Yee, N. (2005). Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science*, 16(10), 814–819.

Bailenson, J. N., Iyengar, S., Yee, N., & Collins, N. A. (2008). Facial similarity between voters and candidates causes influence. *Public Opinion Quarterly*, 72, 935–961.

Balancher, S., Liu, Y., & Stock, A. (2009). An empirical analysis of scarcity strategies in the automobile industry. *Management Science*, 10, 1623–1637.

Balliet, D., Wu, J., & De Dreu, C. K. W. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, 140, 1556–1581.

Bandura, A., Grusec, J. E., & Menlove, F. L. (1967). Vicarious extinction of avoidance behavior. *Journal of Personality and Social Psychology*, 5, 16–23.

Bandura, A., & Menlove, F. L. (1968). Factors determining vicarious extinction of avoidance behavior through symbolic modeling. *Journal of Personality and Social Psychology*, 8, 99–108.

Bannan, N. (ed.). (2012). *Music, language, and human evolution*. Oxford:

Oxford University Press.

Barden, J., Rucker, D. D., & Petty, R. E. (2005). "Saying one thing and doing another": Examining the impact of event order on hypocrisy judgments of others. *Personality and Social Psychology Bulletin*, 31, 1463–1474.

Bargh, J. A., & Williams, E. L. (2006). The automaticity of social life. *Current Directions in Psychological Science*, 15, 1–4.

Barlow, F. K., Paolini, S., Pedersen, A., Hornsey, M. J., Radke, H. R. M., Harwood, J., Rubin, M., & Sibley, C. G. (2012). The contact caveat: Negative contact predicts increased prejudice more than positive contact predicts reduced prejudice. *Personality and Social Psychology Bulletin*, 38, 1629–1643.

Barnett, M. A., Sanborn, F. W., & Shane, A. C. (2005). Factors associated with individuals' likelihood of engaging in various minor moral and legal violations. *Basic and Applied Social Psychology*, 27, 77–84.

Barone, M. J., & Roy, T. (2010). The effect of deal exclusivity on consumer response to targeted price promotions: A social identification perspective. *Journal of Consumer Psychology*, 20, 78–89.

Bastardi, A., & Shafr, E. (2000). Nonconsequential reasoning and its consequences. *Current Directions in Psychological Science*, 9, 216–219.

Bastian, B., Jetten, J., & Ferris, L. J. (2014). Pain as social glue: shared pain increases cooperation. *Psychological Science*, 25, 2079–2085.

Bawden, D., & Robinson, L. (2009). The dark side of information: Overload, anxiety and other paradoxes and pathologies. *Journal of Information Science*, 35, 180–191.

Benjamin, J. (2015, June 22). Market research fail: How New Coke became

the worst flub of all time. Business 2 Community (website).

Benson, P. L., Karabenic, S. A., & Lerner, R. M. (1976). Pretty pleases: The effects of physical attractiveness on race, sex, and receiving help. *Journal of Experimental Social Psychology*, 12, 409–415.

Benton, A. A., Kelley, H. H., & Liebling, B. (1972). Effects of extremity of offers and concession rate on the outcomes of bargaining. *Journal of Personality and Social Psychology*, 24, 73–83.

Bergquist, M., Nilsson, A., & Schultz, W. P. (2019). A meta-analysis of field experiments using social norms to promote pro-environmental behaviors. *Global Environmental Change*, 58. doi.org/10.1016/j.gloenvcha.2019.101941.

Bernache-Assolant, I., Lacassagne, M-F., & Braddock, J. H. (2007). Basking in reflected glory and blasting: Differences in identity management strategies between two groups of highly identified soccer fans. *Journal of Language and Social Psychology*, 26, 381–388.

Berndsen, M., & McGarty, C. (2012). Perspective taking and opinions about forms of reparation for victims of historical harm. *Personality and Social Psychology Bulletin*, 38, 1316–1328.

Bernieri, F. J. (1988). Coordinated movement and rapport in teacher-student interactions. *Journal of Nonverbal Behavior*, 12, 120–138.

Berns, G. S., Chappelow J., Zink, C. F., Pagnoni, G., Martin-Skuski, M. E., & Richards, J. (2005). Neurobiological correlates of social conformity and independence during mental rotation. *Biological Psychiatry*, 58, 245–253.

Bickman, L. (1974). The social power of a uniform. *Journal of Applied Social Psychology*, 4, 47–61.

Binning, K. R., Kaufmann, N., McGreevy, E. M., Fotuhi, O., Chen, S., Marshman, E., Kalender, Z. Y., Limeri, L., Betancur, L., & Singh, C. (2020). Changing social contexts to foster equity in college science courses: An ecological-belonging intervention. *Psychological Science*, 31, 1059–1070.

Bizer, G. Y., & Krosnick, J. A. (2001). Exploring the structure of strength-related attitude features: The relation between attitude importance and attitude accessibility. *Journal of Personality and Social Psychology*, 81, 566–586.

Blake, R., & Mouton, J. (1979). Intergroup problem solving in organizations: From theory to practice. In W. Austin and S. Worchel (eds.), *The social psychology of intergroup relations* (pp. 19–32). Monterey, CA: Brooks/Cole.

Blanco, F., Gómez-Fortes, B., & Matute, H. (2018). Causal illusions in the service of political attitudes in Spain and the United Kingdom. *Frontiers in Psychology*, 28.

Blass, T. (2004). *The man who shocked the world: The life and legacy of Stanley Milgram*. New York: Basic Books.

Blass, T. (2012). A cross-cultural comparison of studies of obedience using the Milgram paradigm: A review. *Social and Personality Psychology Compass*, 6, 196–205.

Boen, F., Vanbeselaere, N., Pandelaere, M., Dewitte, S., Duriez, B., Snauwaert, B., Feys, J., Dierckx, V., & Van Avermaet, E. (2002). Politics and basking-in- reflected-glory. *Basic and Applied Social Psychology*, 24, 205–214.

Boh, W. F., & Wong, S-S. (2015). Managers versus co-workers as referents: Comparing social influence effects on within- and outside-subsidiary knowledge sharing. *Organizational Behavior and Human Decision*

Processes, 126, 1–17.

Bollen, K. A., & Phillips, D. P (1982). Imitative suicides: A national study of the effects of television news stories. *American Sociological Review*, 47, 802–809.

Bomey, N. (2017, July 3). Nissan Rogue gets a galactic sales boost from “Star Wars.” *Arizona Republic*, B4.

Bond, M. H., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch’s (1952b, 1956) line judgment task. *Psychological Bulletin*, 119, 111–137.

Bond, R., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489, 295–298.

Bonneville-Roussy, A., Rentfrow, P. J., Potter, J., & Xu, M. K. (2013). Music through the ages: Trends in musical engagement and preferences from adolescence through middle adulthood. *Journal of Personality and Social Psychology*, 105, 703–717.

Boothby, E. J., & Bohns, V. K. (2020). Why a simple act of kindness is not as simple as it seems: Underestimating the positive impact of our compliments on others. *Personality and Social Psychology Bulletin*.

Borgida, E., Conner, C., & Manteufal, L. (1992). Understanding living kidney donation: A behavioral decision-making perspective. In S. Spacapan and S. Oskamp (eds.), *Helping and being helped* (pp. 183–212). Newbury Park, CA: Sage.

Borman, G. D., Rozek, C. S., Pyne, J., & Hanselman, P. (2019). Reappraising academic and social adversity improves middle school students’ academic achievement, behavior, and well-being. *Proceedings of the National Academy*

of Sciences, 116, 16286–16291.

Bornstein, R. F., Leone, D. R., & Galley, D. J. (1987). The generalizability of subliminal mere exposure effects. *Journal of Personality and Social Psychology*, 53, 1070–1079.

Bouchard, T. J., Segal, N. L., Tellegen, A., McGue, M., Keyes, M., & Krueger, R. (2003). Evidence for the construct validity and heritability of the Wilson-Paterson conservatism scale: A reared-apart twins study of social attitudes. *Personality and Individual Differences*, 34, 959–969.

Bourgeois, M. J. (2002). Heritability of attitudes constrains dynamic social impact. *Personality and Social Psychology Bulletin*, 28, 1063–1072.

Brandt, M. J., & Reyna, C. (2011). The chain of being: A hierarchy of morality. *Perspectives on Psychological Science*, 6, 428–446.

Brehm, J. W. (1966). *A theory of psychological reactance*. New York: Academic Press.

Brehm, S. S. (1981). Psychological reactance and the attractiveness of unattainable objects: Sex differences in children's responses to an elimination of freedom. *Sex Roles*, 7, 937–949.

Brehm, S. S., & Weintraub, M. (1977). Physical barriers and psychological reactance: Two-year-olds' responses to threats to freedom. *Journal of Personality and Social Psychology*, 35, 830–836.

Brendl, C. M., Chattopadhyay, A., Pelham, B. W., & Carvallo, M. (2005). Name letter branding: Valence transfers when product specific needs are active. *Journal of Consumer Research*, 32, 405–415.

Bridge, J. A., Greenhouse, J. B., Ruch, D., Stevens, J., Ackerman, J., Sheftall, A.H., Horowitz, L. M., Kelleher, K. J., & Campo, J. V. (2019). Association

between the release of Netflix's *13 Reasons Why* and suicide rates in the United States: An interrupted times series analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*.

Briñol, P., Petty, R. E., & Wheeler, S. C. (2006). Discrepancies between explicit and implicit self-concepts: Consequences for information processing. *Journal of Personality and Social Psychology*, 91, 154–170.

Brockner, J., & Rubin, J. Z. (1985). *Entrapment in escalating conflicts: A social psychological analysis*. New York: Springer-Verlag.

Bronnenberg, B. J., Dhar, S. K., & Dubé, J.-P. (2007). Consumer packaged goods in the United States: National brands, local branding. *Journal of Marketing Research*, 44, 4–13.

Broockman, D. & Kalla, J. (2016). Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science*, 352, 220–224.

Brown, J. L., Drake, K. D., & Wellman, L. (2015). The benefits of a relational approach to corporate political activity: Evidence from political contributions to tax policymakers. *Journal of the American Taxation Association*, 37, 69–102.

Brown, S. L., Asher, T., & Cialdini, R. B. (2005). Evidence of a positive relationship between age and preference for consistency. *Journal of Research in Personality*, 39, 517–533.

Browne, W., & Swarbrick-Jones, M. (2017). What works in e-commerce: A meta-analysis of 6700 online experiments. *Qubit Digital LTD*.

Brownstein, R., & Katzev, R. (1985). The relative effectiveness of three compliance techniques in eliciting donations to a cultural organization. *Journal of Applied Social Psychology*, 15, 564–574.

Bruneau, E. G., Kteily, N. S., & Urbiola, A. (2020). A collective blame hypocrisy intervention enduringly reduces hostility towards Muslims. *Nature Human Behaviour*, 4, 45–54.

Buchan, N. R., Brewer, M. B., Grimalda, G., Wilson, R. K., Fatas, E., & Foddy, M. (2011). Global social identity and global cooperation. *Psychological Science*, 22, 821–828.

Budesheim, T. L., & DePaola, S. J. (1994). Beauty or the beast? The effects of appearance, personality, and issue information on evaluations of political candidates. *Personality and Social Psychology Bulletin*, 20, 339–348.

Bullack, A., Gass, C., Nater, U. M., & Kreutz, G. (2018). Psychobiological effects of choral singing on affective state, social connectedness, and stress: Influences of singing activity and time course. *Frontiers of Behavioral Neuroscience*. 12:223.

Burger, J. M. (2009). Replicating Milgram: Would people still obey today? *American Psychologist*, 64, 1–11.

Burger, J. M., & Caldwell, D. F. (2003). The effects of monetary incentives and labeling on the foot-in-the-door effect. *Basic and Applied Social Psychology*, 25, 235–241.

Burger, J. M., & Caldwell, D. F. (2011). When opportunity knocks: The effect of a perceived unique opportunity on compliance. *Group Processes & Intergroup Relations*, 14, 671–680.

Burger, J. M., & Caputo, D. (2015). The low-ball compliance procedure: a metaanalysis. *Social Influence*, 10, 214–220. DOI: 10.1080/15534510.2015.1049203.

Burger, J. M., Horita, M., Kinoshita, L., Roberts, K., & Vera, C. (1997). Effects of time on the norm of reciprocity. *Basic and Applied Social Psychology*, 20, 1–11.

Psychology, 19, 91–100.

Burger, J. M., Messian, N., Patel, S., del Prado, A., & Anderson, C. (2004). What a coincidence! The effects of incidental similarity on compliance. *Personality and Social Psychology Bulletin*, 30, 35–43.

Burger, J. M., & Petty, R. E. (1981). The low-ball compliance technique: Task or person commitment? *Journal of Personality and Social Psychology*, 40, 492–500.

Burgoon, M., Alvaro, E., Grandpre, J., & Voulodakis, M. (2002). Revisiting the theory of psychological reactance. In J. P. Dillard and M. Pfau (eds.), *The persuasion handbook: Theory and practice* (pp. 213–232). Thousand Oaks, CA: Sage.

Burnstein, E., Crandall, C., & Kitayama, S. (1994). Some neo-Darwin decision rules for altruism: Weighing cues for inclusive fitness as a function of the biological importance of the decision. *Journal of Personality and Social Psychology*, 67, 773–789.

Bushman, B. J. (1988). The effects of apparel on compliance. *Personality and Social Psychology Bulletin*, 14, 459–467.

Bushman, B. J. (2006). Effects of warning and information labels on attraction to television violence in viewers of different ages. *Journal of Applied Social Psychology*, 36, 2073–2078.

Buttleman, D., & Bohm, R. (2014). The ontogeny of the motivation that underlies ingroup bias. *Psychological Science*, 25, 921–927.

Buttrick, N., Moulder, R., & Oishi, S. (2020). Historical change in the moral foundations of political persuasion. *Personality and Social Psychology Bulletin*, 46, 1523–1537. doi:10.1177/0146167220907467.

Cadinu, M. R., & Rothbart, M. (1996). Self-anchoring and differentiation processes in the minimal group setting. *Journal of Personality and Social Psychology*, 70, 666–677.

Cai, H., Chen, Y., & Fang, H. (2009). Observational learning: Evidence from a randomized natural field experiment. *American Economic Review*, 99, 864–882.

Campbell, M. W., & de Waal, F. B. M. (2010). Methodological problems in the study of contagious yawning. *Frontiers in Neurology and Neuroscience*, 28, 120–127.

Cappella, J. N. (1997). Behavioral and judged coordination in adult informal social interactions: Vocal and kinesic indicators. *Journal of Personality and Social Psychology*, 72, 119–131.

Carducci, B. J., Deuser, P. S., Bauer, A., Large, M., & Ramaekers, M. (1989). An application of the foot-in-the-door technique to organ donation. *Journal of Business and Psychology*, 4, 245–249.

Carey, B. (2016, July 26). Mass killings may have created contagion, feeding on itself. *New York Times*, A11.

Caro, R. A. (2012). *The passage of power*. Vol. 4 of *The years of Lyndon Johnson*. New York: Knopf.

Carpenter, C. J. (2013). A meta-analysis of the effectiveness of the “But You Are Free” compliance-gaining technique. *Communication Studies*, 64, 6–17.

Carr, N. (2010). *The shallows: What the internet is doing to our brains*. New York: W. W. Norton.

Carter, S. E., & Sanna, L. J. (2006). Are we as good as we think? Observers’ perceptions of indirect self-presentation as a social influence tactic. *Social*

Influence, 1, 185–207.

Catapano, R., Tormala, Z. L., & Rucker, D. D. (2019). Perspective taking and self-persuasion: Why “putting yourself in their shoes” reduces openness to attitude change. *Psychological Science*, 30, 424–435.

Cavazza, N. (2016). When political candidates “go positive”: The effects of flattering the rival in political communication. *Social Influence*, 11, 166–176.

Čehajić, S., & Brown, R. (2010). Silencing the past: Effects of intergroup contact on acknowledgment of in-group responsibility. *Social Psychological and Personality Science*, 1, 190–196.

Čehajić-Clancy, S., Effron, D. A., Halperin, E., Liberman, V., & Ross, L. D. (2011). Affirmation, acknowledgment of in-group responsibility, group-based guilt, and support for reparative measures. *Journal of Personality and Social Psychology*, 101, 256–270.

Chagnon, N. A., & Bugos, P. E. (1979). Kin selection and conflict: An analysis of a Yanomano ax fight. In N. A. Chagnon and W. Irons (eds.), *Evolutionary biology and social behavior* (pp. 213–238). North Scituate, MA: Duxbury.

Chaiken, S. (1979). Communicator physical attractiveness and persuasion. *Journal of Personality and Social Psychology*, 37, 1387–1397.

Chaiken, S. (1986). Physical appearance and social influence. In C. P. Herman, M. P. Zanna, and E. T. Higgins (eds.), *Physical appearance, stigma, and social behavior: The Ontario Symposium* (vol. 3, pp. 143–177). Hillsdale, NJ: Lawrence Erlbaum.

Chajut, E., & Algom, D. (2003). Selective attention improves under stress. *Journal of Personality and Social Psychology*, 85, 231–248.

Chambers, J. R., Schlenker, B. R., & Collisson, B. (2013). Ideology and prejudice: The role of value conflicts. *Psychological Science*, 24, 140–149.

Chan, C., Berger, J., & Van Boven, L. (2012). Identifiable but not identical: Combining social identity and uniqueness motives in choice. *Journal of Consumer Research*, 39, 561–573.

Chan, E., & Sengupta, J. (2010). Insincere flattery actually works: A dual attitudes perspective. *Journal of Marketing Research*, 47, 122–133.

Cheng, L., Hao, M., Xiao, L., & Wang, F. (2020). Join us: Dynamic norms encourage women to pursue STEM. *Current Psychology*.

Chernyak, N., Leimgruber, K. L., Dunham, Y. C., Hu, J., & Blake, P. R. (2019). Paying back people who harmed us but not people who helped us: Direct negative reciprocity precedes direct positive reciprocity in early development. *Psychological Science*.

Christakis, N. A., & Fowler, J. H. (2014). Friendship and natural selection. *Proceedings of the National Academy of Sciences*, 111, 10796–10801.

Chugani, S., Irwin, J. E., & Redden, J. P. (2015). Happily ever after: The effect of identity-consistency on product satiation. *Journal of Consumer Research*, 42, 564–577.

Cialdini, R. B. (2003). Crafting normative messages to protect the environment. *Current Directions in Psychological Science*, 12, 105–109.

Cialdini, R. B. (1997). Professionally responsible communication with the public: Giving psychology a way. *Personality and Social Psychology Bulletin*, 23, 675–683.

Cialdini, R. B. (2016). *Pre-Suasion: A revolutionary way to influence and persuade*. New York: Simon & Schuster.

Cialdini, R. B. (2018). Why the world is turning to behavioral science. In A. Samson (ed.), *The behavioral economics guide 2018* (pp. vii–xiii).

Cialdini, R. B., & Ascani, K. (1976). Test of a concession procedure for inducing verbal, behavioral, and further compliance with a request to give blood. *Journal of Applied Psychology*, 61, 295–300.

Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R. Freeman, S., & Sloan, L. R. (1976). Basking in reflected glory: Three (football) field studies. *Journal of Personality and Social Psychology*, 34, 366–375.

Cialdini, R. B., Brown, S. L., Lewis, B. P., Luce, C., & Neuberg, S. L. (1997). Reinterpreting the empathy-altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology*, 73, 481–494.

Cialdini, R. B., Cacioppo, J. T., Bassett, R., & Miller, J. A. (1978). Low-ball procedure for producing compliance: Commitment then cost. *Journal of Personality and Social Psychology*, 36, 463–476.

Cialdini, R. B., Eisenberg, N., Green, B. L., Rhoads, K. v. L., & Bator, R. (1998). Undermining the undermining effect of reward on sustained interest. *Journal of Applied Social Psychology*, 28, 249–263.

Cialdini, R. B., Li, J., Samper, A., & Wellman, E. (2019). How bad apples promote bad barrels: Unethical leader behavior and the selective attrition effect. *Journal of Business Ethics*.

Cialdini, R. B., Petrova, P., & Goldstein, N. J. (2004). The hidden costs of organizational dishonesty. *MIT Sloan Management Review*, 45, 67–73.

Cialdini, R. B., Trost, M. R., & Newsom, J. T. (1995). Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications. *Journal of Personality and Social Psychology*, 69, 318–328.

Cialdini, R. B., Vincent, J. E., Lewis, S. K., Catalan, J., Wheeler, D., & Darby, B. L.(1975). Reciprocal concessions procedure for inducing compliance: The door- in-the-face technique. *Journal of Personality and Social Psychology*, 31, 206–215.

Cialdini, R. B., Wosinska, W., Barrett, D. W., Butner, J., & Gornik-Durose, M. (1999). Compliance with a request in two cultures: The differential influence of social proof and commitment/consistency on collectivists and individualists. *Personality and Social Psychology Bulletin* 25, 1242–1253.

Cikara, M., & Paluck, E. L. (2013), When going along gets you nowhere and the upside of conflict behaviors. *Social and Personality Psychology Compass*, 7, 559–571.

Cikara, M., & van Bavel, J. (2014). The neuroscience of inter-group relations: An integrative review. *Perspectives on Psychological Science*, 9, 245–274.

Cioffi, D., & Garner, R. (1996). On doing the decision: The effects of active versus passive choice on commitment and self-perception. *Personality and Social Psychology Bulletin*, 22, 133–147.

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, 17, 1003–1011.

Clark, C. J., Liu, B. S., Winegard, B. M., & Ditto, P. H. (2019). Tribalism is human nature. *Current Directions in Psychological Science*, 28, 587–592.

Clark, M. S., Lemay, E. P., Graham, S. M., Pataki, S. P., & Finkel, E. J. (2010). Ways of giving benefits in marriage: Norm use, relationship satisfaction, and attachment-related variability. *Psychological Science*, 21, 944–951.

Clark, M. S., Mills, J. R., & Corcoran, D. M. (1989). Keeping track of needs and inputs of friends and strangers. *Personality and Social Psychology Bulletin*, 15, 533–542.

Clark, R. D., III, & Word, L. E. (1972). Why don't bystanders help? Because of ambiguity? *Journal of Personality and Social Psychology*, 24, 392–400.

Clark, R. D., III, & Word, L. E. (1974). Where is the apathetic bystander? Situational characteristics of the emergency. *Journal of Personality and Social Psychology*, 29, 279–287.

Clarkson, J. J., Tormala, Z. L., & Rucker, D. D. (2011). Cognitive and affective matching effects in persuasion: An amplification perspective. *Personality and Social Psychology Bulletin*, 1415–1427.

Clifford, S., & Jerit, J. (2016). Cheating on political knowledge questions in online surveys. *Public Opinion Quarterly*, 80, 858–887.

Coghlan, T. (2015, July 14). Weidenfeld's crusade to save Christians of Syria. *The Times* (London), A30.

Cohen, R. (1972). Altruism: Human, cultural, or what? *Journal of Social Issues*, 28, 39–57.

Cohen, A. (1999, May 31). Special report: Troubled kids. *Time*, 38.

Cohen, G. L., & Sherman, D. K. (2014). The psychology of change: Selfaffirmation and social psychological intervention. *Annual Review of Psychology*, 65, 333–371.

Cohen, M., & Davis, N. (1981). *Medication errors: Causes and prevention*. Philadelphia: G. F. Stickley.

Coleman, N. V., & Williams, P. (2015). Looking for my self: Identity-driven attention allocation. *Journal of Consumer Psychology*, 25, 504–511.

Collins, J. (2018). Simple heuristics that make algorithms smart.

Combs, D. J. Y., & Keller, P. S. (2010). Politicians and trustworthiness: Acting contrary to self-interest enhances trustworthiness. *Basic and Applied Social Psychology*, 32, 328–339.

Converse, B. A., & Fishbach, A. (2012). Instrumentality boosts appreciation: Helpers are more appreciated while they are useful. *Psychological Science*, 23, 560–566.

Conway, A., & Cowan, N. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review*, 8, 331–335.

Cooper, J., & Fazio, R. H. (1984). A new look at dissonance theory. In L. Berkowitz (ed.), *Advances in experimental social psychology* (vol. 17, pp. 229–266). New York: Academic Press.

Coppock, A., Ekins, E., & Kirby, D. (2018). The long-lasting effects of newspaper opeds on public opinion. *Quarterly Journal of Political Science*, 13, 59–87.

Cornelissen, G., Pandelaere, M., Warlop, L., & Dewitte, S. (2008). Positive cueing: Promoting sustainable consumer behavior by cueing common environmental behaviors as environmental. *International Journal of Research in Marketing*, 25, 46–55.

Corning, A., & Schuman, H. (2013). Commemoration matters: The anniversaries of 9/11 and Woodstock. *Public Opinion Quarterly*, 77, 433–454.

Cortijos-Bernabeu, A., Bjørndal, L. D., Ruggeri, K., Alí, S., Friedemann, M., Esteban-Serna, C., Khorrami, P. R., et al. (2020). Replicating patterns of prospect theory for decision under risk. *Nature Human Behaviour*, 4, 622–

633.

Cosmides, L., & Tooby, J. (2013). Evolutionary psychology: New perspectives on cognition and motivation. *Annual Review of Psychology*, 64, 201–229.

Craig, B. (1985, July 30). A story of human kindness. *Pacific Stars and Stripes*, 13–16. Crespelle, J. P. (1969). *Picasso and his women*. New York: Hodder & Stoughton.

Cronley, M., Posavac, S. S., Meyer, T., Kardes, F. R., & Kellaris, J. J. (2005). A selective hypothesis testing perspective on price-quality inference and inference-based choice. *Journal of Consumer Psychology*, 15, 159–169.

Cullum, J., O’Grady, M., Sandoval, P., Armeli, A., & Tennen, T. (2013). Ignoring norms with a little help from my friends: Social support reduces normative influence on drinking behavior. *Journal of Social and Clinical Psychology*: 32, 17–33.

Cunningham, M. R. (1986). Levites and brother’s keepers: A sociobiological perspective on prosocial behavior. *Humboldt Journal of Social Relations*, 13, 35–67.

Curry, O., Roberts, S. G. B., & Dunbar, R. I. M. (2013). Altruism in social networks: Evidence for a “kinship premium.” *British Journal of Psychology*, 104, 283–295.

Dai, X., Wertenbroch, K., & Brendel, C. M. (2008). The value heuristic in judgments of relative frequency. *Psychological Science*, 19, 18–19.

Daly, M., Salmon, C. & Wilson, M. (1997). Kinship: The conceptual hole in psychological studies of social cognition and close relationships. In J. A. Simpson and D. T. Kendrick (eds.), *Evolutionary Social Psychology* (pp. 265–296). Mahwah, NJ: Erlbaum.

- Danchin, E., Nöbel, S., Pocheville, A., Dagaëff, A-C., Demay, L., Alphan, M., Ranty-Roby, S., et al. (2018). Cultural flies: Conformist social learning in fruitflies predicts long-lasting mate-choice traditions. *Science*, 362, 1025–1030.
- Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8, 377–383.
- Dauten, D. (2004, July 22). How to be a good waiter and other innovative ideas. *Arizona Republic*, D3.
- Davies, J. C. (1962). Toward a theory of revolution. *American Sociological Review*, 27, 5–19.
- Davies, J. C. (1969). The J-curve of rising and declining satisfactions as a cause of some great revolutions and a contained rebellion. In H. D. Graham and T. R. Gurr (eds.), *Violence in America* (pp. 547–644). New York: Signet.
- Davies, K., Tropp, L. R., Aron, A., Pettigrew, T. F., & Wright, S. C. (2011). Cross-group friendships and intergroup attitudes: A meta-analytic review. *Personality and Social Psychology Review*, 15, 332–351.
- Davis, M. H., Conklin, L., Smith, A., & Luce, C. (1996). Effect of perspective taking on the cognitive representation of persons: A merging of self and other. *Journal of Personality and Social Psychology*, 70, 713–726.
- DeBruine, L. M. (2002). Facial resemblance enhances trust. *Proceedings of the Royal Society, Series B*, 269, 1307–1312.
- DeBruine, L. M. (2004). Resemblance to self increases the appeal of child faces to both men and women. *Evolution and Human Behavior*, 25, 142–154.
- Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The truth about the

truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, 14, 238–257.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Deci, E. L., Spiegel, N. H., Ryan, R. M., Koestner, R., & Kauffman, M. (1982). Effects of performance standards on teaching styles: Behavior of controlling teachers. *Journal of Educational Psychology*, 74, 852–859.

De Dreu, C. K. W., & McCusker, C. (1997). Gain-loss frames and cooperation in two-person social dilemmas: A transformational analysis. *Journal of Personality and Social Psychology*, 72, 1093–1106.

De Dreu, C. K. W., Dussel, D. B., & Ten Velden, F. S. (2015). In intergroup conflict, self-sacrifice is stronger among pro-social individuals and parochial altruism emerges especially among cognitively taxed individuals. *Frontiers in Psychology*, 6, 572.

DeJong, C., Aguilar, T., Tseng, C-W., Lin, G. A., Boscardin, W. J., & Dudley, R. A. (2016). Pharmaceutical industry-sponsored meals and physician prescribing patterns for Medicare beneficiaries. *Journal of the American Medical Association: Internal Medicine*, 176, 1114–1122.

de la Rosa, M. D., Sanabria, D., Capizzi, M., & Correa, A. (2012). Temporal preparation driven by rhythms is resistant to working memory interference. *Frontiers in Psychology*, 3.

Dellande, S., & Nyer, P. (2007). Using public commitments to gain customer compliance. *Advances in Consumer Research*, 34, 249–255.

De Martino, B., Bobadilla-Suarez, S., Nouguchi, T., Sharot, T., & Love, B. C. (2017). Social information is integrated into value and confidence judgments according to its reliability. *Journal of Neuroscience*, 37, 6066–

6074.

Demuru, E., & Palagi, E. (2012). In Bonobos yawn contagion is higher among kin and friends. *PLoS ONE*, 7.

DePaulo, B. M., Nadler, A., & Fisher, J. D. (eds.). (1983). *Help seeking*. Vol. 2 of *New directions in helping*. New York: Academic Press.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, 51, 629–636.

Devlin, A. S., Donovan, S., Nicolov, A., Nold, O., Packard, A., & Zandan, G. (2009). “Impressive?” Credentials, family photographs, and the perception of therapist qualities. *Journal of Environmental Psychology*, 29, 503– 512.

DeWall, C. N., MacDonald, G., et al. (2010). Acetaminophen reduces social pain: Behavioral and neural evidence. *Psychological Science*, 21, 931–937.

Dhami, M. K. (2003). Psychological models of professional decision making. *Psychological Science*, 14, 175–180.

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., et al. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current Biology*, 27, 1375–1380.

Dillard, J. P., Kim, J., & Li, S. S. (2018). Anti-sugar-sweetened beverage messages elicit reactance: Effects on attitudes and policy preferences. *Journal of Health Communication*, 23, 703–711.

Dimmock, S. G., Gerken, W. C., & Graham, N. P. (2018). Is fraud contagious? Coworker influence on misconduct by financial advisors. *Journal of Finance*, 73, 1417–1450.

Dion, K. K. (1972). Physical attractiveness and evaluation of children's transgressions. *Journal of Personality and Social Psychology*, 24, 207–213.

Dixon, J., Durrheim, K., & Tredoux, C. (2005). Beyond the optimal contact strategy: A reality check for the contact hypothesis. *American Psychologist*, 60, 697–711.

Dolińska, B., Jarzabek, J., & Doliński, D. (2020). I like you even less at Christmas dinner! *Basic and Applied Social Psychology*, 42, 88–97.

Doliński, D. (2000). Inferring one's beliefs from one's attempt and consequences for subsequent compliance. *Journal of Personality and Social Psychology*, 78, 260–272.

Doliński, D. (2016). *Techniques of social influence: The psychology of compliance*. New York: Routledge.

Doliński, D., & Grzyb, T. (2020). *Social psychology of obedience toward authority: Empirical tribute to Stanley Milgram*. London: Routledge.

Doliński, D., Grzyb, T., Folwarczny, M., Grzybała, P., Krzyszycha, K., Martynowska, K., & Trojanowski, J. (2017). Would you deliver an electric shock in 2015? Obedience in the experimental paradigm developed by Stanley Milgram in the 50 years following the original studies. *Social Psychological and Personality Science*, 8, 927–933.

Dolnik, L., Case, T. I., & Williams, K. D. (2003). Stealing thunder as a courtroom tactic revisited: Processes and boundaries. *Law and Human Behavior*, 27, 267–287. Donaldson, S. I., Graham, J. W., Piccinin, A. M., & Hansen, W. B. (1995). Resistanceskills training and onset of alcohol use. *Health Psychology*, 14, 291–300.

Doob, A. N., & Gross, A. E. (1968). Status of frustrator as an inhibitor of horn-honking response. *Journal of Social Psychology*, 76, 213–218.

Dovidio, J. F., Gaertner, S. L., & Saguy, T. (2009). Commonality and the complexity of “We”: Social attitudes and social change. *Personality and Social Psychology Review*, 13, 3–20.

Drachman, D., deCarufel, A., & Inkso, C. A. (1978). The extra credit effect in interpersonal attraction. *Journal of Experimental Social Psychology*, 14, 458–467.

Driscoll, R., Davis, K. E., & Lipetz, M. E. (1972). Parental interference and romantic love: The Romeo and Juliet effect. *Journal of Personality and Social Psychology*, 24, 1–10.

Drolet, A., & Aaker J. (2002). Off-target? Changing cognitive-based attitudes. *Journal of Consumer Psychology*, 12, 59–68.

Drury, J. (2018). The role of social identity processes in mass emergency behaviour: An integrative review. *European Review of Social Psychology*, 29, 38–81.

Du, X. (2019). What’s in a surname? The effect of auditor-CEO surname sharing on financial misstatement. *Journal of Business Ethics*, 158, 849–874.

DuBois, D. L., Portillo, N., Rhodes, J. E., Silverthorn, N., & Valentine, J. C. (2011). How effective are mentoring programs for youth? A systematic assessment of the evidence. *Psychological Science in the Public Interest*, 12, 57–91.

Duguid, M. M., & Goncalo, J. A. (2012). Living large: The powerful overestimate their own height. *Psychological Science*, 23, 36–40.

Duguid, M. M., & Thomas-Hunt, M. C. (2015). Condoning stereotyping? How awareness of stereotyping prevalence impacts expression of stereotypes. *Journal of Applied Psychology*, 100, 343–359.

Dunbar, R. I. M. (2012). On the evolutionary function of song and dance. In N. Bannan(ed.), *Music, language and human evolution* (pp. 201–214). Oxford: Oxford University Press.

Dunfield, K. A., & Kuhlmeier, V. A. (2010). Intention-mediated selective helping in infancy. *Psychological Science*, 21, 523–527.

Eagly, A. H., Wood, W., & Chaiken, S. (1978). Causal inferences about communicators and their effect on opinion change. *Journal of Personality and Social Psychology*, 36, 424–435.

Easterbrook, J. A. (1959). The effects of emotion on cue utilization and the organization of behavior. *Psychological Review*, 66, 183–201.

Edwards, M. L., Dillman, D. A., & Smyth, J. D. (2014). An experimental test of the effects of survey sponsorship on internet and mail survey response. *Public Opinion Quarterly*, 78, 734–750.

Effron, D. A., Bryan, C. J., & Murnighan, J. K. (2015). Cheating at the end to avoid regret. *Journal of Personality and Social Psychology*, 109, 395–414.

Efran, M. G., & Patterson, E. W. J. (1976). The politics of appearance. Unpublished manuscript, University of Toronto.

Ellemers, N., & van Nunspeet, F. (2020, September). Neuroscience and the social origins of moral behavior: How neural underpinnings of social categorization and conformity affect everyday moral and immoral behavior. *Current Directions in Psychological Science*.

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why adjustments are insufficient. *Psychological Science*, 17, 311–318.

Epstein, S., Lipson, A., Holstein, C., & Huh, E. (1992). Irrational reactions to negative outcomes: Evidence for two conceptual systems. *Journal of*

Personality and Social Psychology, 62, 328–339.

Epstein, S., Donovan, S., & Denes-Raj, V. (1999). The missing link in the paradox of the Linda conjunction problem: Beyond knowing and thinking of the conjunction rule, the intrinsic appeal of heuristic processing. *Personality and Social Psychology Bulletin*, 25, 204–214.

Facci, E., L., & Kasarda, J. D. (2004). Revisiting wind-shear accidents: The social proof factor. Proceedings of the 49th Corporate Aviation Safety Seminar (pp. 205–232). Alexandria, VA: Flight Safety Foundation.

Faiman-Silva, S. (1997). *Choctaws at the crossroads*. Lincoln: University of Nebraska Press.

Fan, M., Billings, A., Zhu, X., & Yu, P. (2019). Twitter-based BIRGing: Big data analysis of English National Team fans during the 2018 FIFA World Cup. *Communication & Sport*.

Fang, X., Singh, S., & Ahulwailia, R. (2007). An examination of different explanations for the mere exposure effect. *Journal of Consumer Research*, 34, 97–103.

Farrow, K., Grolleau, G., & Ibanez, L. (2017). Social norms and pro-environmental behavior: A review of the evidence. *Ecological Economics*, 140, 1–13.

Fazio, L. K., Rand, D. G., & Pennycook, G. (2019). *Psychonomic Bulletin Review*.

Fazio, R. H., Sherman, S. J., & Herr, P. M. (1982). The feature-positive effect in the self-perception process. *Journal of Personality and Social Psychology*, 42, 404–411.

Fein, S., Goethals, G. R., & Kugler, M. B. (2007). Social influence on

political judgments: The case of presidential debates. *Political Psychology*, 28, 165–192.

Feinberg, R. A. (1986). Credit cards as spending facilitating stimuli. *Journal of Consumer Research*, 13, 348–356.

Feinberg, R. A. (1990). The social nature of the classical conditioning phenomena in people. *Psychological Reports*, 67, 331–334.

Feld, J., Salamanca, N., Hamermesh, D. S. (2015). Endophilia or exophobia: Beyond discrimination. *Economic Journal*, 126, 1503–1527.

Fennis, B. M., Janssen, L., & Vohs, K. D. (2008). Acts of benevolence: A limited-resource account of compliance with charitable requests. *Journal of Consumer Research*, 35, 906–924.

Fennis, B. M., & Stroebe, W. (2014). Softening the blow: Company self-disclosure of negative information lessens damaging effects on consumer judgment and decision making. *Journal of Business Ethics*, 120, 109–120.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, 58, 203–210.

Festinger, L., Riecken, H. W., & Schachter, S. (1964). *When prophecy fails*. New York: Harper & Row.

Fischer, P., Krueger, J. I., Greitemeyer, T., Vögrincic, C., Kastenmüller, A., Frey, D., Heene, M., et al. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin*, 137, 517–537.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation:

Influences of information and motivation on attention and interpretation. In M. P. Zanna (ed.), *Advances in experimental social psychology* (vol. 23, pp. 1–74). New York: Academic Press.

Fisman, R., Paravisini, D., & Vig, V. 2017. Cultural proximity and loan outcomes. *American Economic Review*, 107, 457–492.

Flade, F., Klar, Y., & Imhoff, R. (2019). Unite against: A common threat invokes spontaneous decategorization between social categories. *Journal of Experimental Social Psychology*, 85.

Fleming, T. (1997, November 23). 13 things you never knew about the American Revolution. *Parade*, 14–15.

Flynn, F. J. (2002). What have you done for me lately? Temporal adjustments to favor evaluations. *Organizational Behavior and Human Decision Processes*, 91, 38–50.

Foddy, M., Platow, M. J., & Yamagishi, T. (2009). Group-based trust in strangers. *Psychological Science*, 20, 419–422.

Foerster, M., Roser, K., Schoeni, A., & Rösli, M. (2015). Problematic mobile phone use in adolescents: Derivation of a short scale MPPUS-10. *International Journal of Public Health*, 60, 277–286.

Fogg, B. J., & Nass, C. (1997a). How users reciprocate to computers: An experiment that demonstrates behavior change. In *Extended Abstracts of the CHI97 Conference of the ACM/SIGCHI*. New York: ACM.

Fogg, B. J., & Nass, C. (1997b). Silicon sycophants: The effects of computers that flatter. *International Journal of Human-Computer Studies*, 46(5), 551–561.

Fombelle, P., Gustafsson, A., Andreassen, T. W., & Witell, L. (2010). *Give*

and thou shall receive: Customer reciprocity in a retail setting. Paper presented at the 19th Annual Frontiers In Service Conference, Karlstad, Sweden.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Research Systems*, 19, 291–313.

Fornara, F., Carrus, G., Passafaro, P., & Bonnes, M. (2011). Distinguishing the sources of normative influence on pro-environmental behaviors: The role of local norms in household waste recycling. *Group Processes & Intergroup Dynamics*, 14, 623–635.

Fox, C. R., Linder, J. A., & Doctor, J., N. (2016, March 27). How to stop overprescribing antibiotics. *New York Times*.

Fox, M. W. (1974). *Concepts in ethology: Animal and human behavior*. Minneapolis: University of Minnesota Press.

Frank, R. H. (2020). *Under the Influence: Putting peer pressure to work*. Princeton, NJ: Princeton University Press.

Fraune, M. R. (2020). Our robots, our team: Robot anthropomorphism moderates group effects in human–robot teams. *Frontiers in Psychology*, 11, 1275.

Fredman, L. A., Buhrmester, M. D., Gomez, A., Fraser, W. T., Talaifar, S., Brannon, S.M., & Swann, Jr., W. B. (2015). Identity fusion, extreme pro-group behavior, and the path to defusion. *Social and Personality Psychology Compass*, 9, 468–480.

Freedman, J. L. (1965). Long-term behavioral effects of cognitive dissonance. *Journal of Experimental Social Psychology*, 1, 145–155.

Freedman, J. L., & Fraser, S. C. (1966). Compliance without pressure: The foot- in-the-door technique. *Journal of Personality and Social Psychology*, 4, 195–203.

Freling, T. H., & Dacin, P. A. (2010). When consensus counts: Exploring the impact of consensus claims in advertising. *Journal of Consumer Psychology*, 20, 163–175.

Frenzen, J. R., & Davis, H. L. (1990). Purchasing behavior in embedded markets. *Journal of Consumer Research*, 17, 1–12.

Friedman, H. H., & Rahman, A. (2011). Gifts-upon-entry and appreciative comments: Reciprocity effects in retailing. *International Journal of Marketing Studies*, 3, 161–164.

Friestad, M., & Wright, P. (1995). Persuasion knowledge: Lay people's and researchers' beliefs about the psychology of persuasion. *Journal of Consumer Research*, 22, 62–74.

Fröhlich, M., Müller, G., Zeiträg, C., Wittig, R. M., & Pika, S. (2020). Begging and social tolerance: Food solicitation tactics in young chimpanzees (*Pan troglodytes*) in the wild. *Evolution and Human Behavior*, 41, 126–135.

Furnham, A. (1996). Factors relating to the allocation of medical resources. *Journal of Social Behavior and Personality*, 11, 615–624.

Gaesser, B., Shimura, Y., & Cikara, M. (2020). Episodic simulation reduces intergroup bias in prosocial intentions and behavior. *Journal of Personality and Social Psychology*, 118, 683–705.

Gal, D. & Rucker D. D. (2010). When in doubt, shout! Paradoxical influences of doubt on proselytizing. *Psychological Science*, 21, 1701–1707.

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing

stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, 78, 708–724.

Gallup, A. C., Hale, J. J., Sumpter, D. J. T., Garnier, S., Kacelnik, A., Krebs, J. R., & Couzin, I. D. (2012). Visual attention and the acquisition of information in human crowds. *Proceedings of the National Academy of Sciences*, 109, 7245–7250.

Gansberg, M. (1964, March 27). 37 who saw murder didn't call the police. *New York Times*, 1.

Garcia, D., & Rimé, B. (2019). Collective emotions and social resilience in the digital traces after a terrorist attack. *Psychological Science*, 30, 617–628.

Garcia, J. H., Sterner, T., & Afsah, S. (2007). Public disclosure of industrial pollution: The PROPER approach in Indonesia. *Environmental and Developmental Economics*, 12, 739–756.

Garner, R. L. (2005). What's in a name? Persuasion perhaps? *Journal of Consumer Psychology*, 15, 108–116.

Gawronski, B. (2003). Implicational schemata and the correspondence bias: On the diagnostic value of situationally constrained behavior. *Journal of Personality and Social Psychology*, 84, 1154–1171.

Geers, A. L., Rose, J. P., Fowler, S. L., Rasinski, H. M., Brown, J. A., & Helfer, S. G. (2013). Why does choice enhance treatment effectiveness? Using placebo treatments to demonstrate the role of personal control. *Journal of Personality and Social Psychology*, 105 (4), 549–566.

Gehlbach, H., Brinkworth, M. E., King, A. M., Hsu, L. M., McIntyre, J., & Rogers, T. (2016). Creating birds of similar feathers: Leveraging similarity to improve teacher–student relationships and academic achievement. *Journal of Educational Psychology*, 108(3), 342–352.

George, W. H., Gournic, S. J., & McAfee, M. P. (1988). Perceptions of postdrinking female sexuality. *Journal of Applied Social Psychology, 18*, 1295–1317.

Gerard, H. B., & Mathewson, G. C. (1966). The effects of severity of initiation on liking for a group: A replication. *Journal of Experimental Social Psychology, 2*, 278–287.

Gerend, M. A., & Maner, J. K. (2011). Fear, anger, fruits, and veggies: Interactive effects of emotion and message framing on health behavior. *Health Psychology, 30*, 420–423.

Gergen, K., Ellsworth, P., Maslach, C., & Seipel, M. (1975). Obligation, donor resources, and reactions to aid in three cultures. *Journal of Personality and Social Psychology, 31*, 390–400.

Ghosh, B. (2009, June 8). How to make terrorists talk. *Time*, 40–43.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*, 650–669.

Gilbert, D. T. (2006). *Stumbling on happiness*. New York: Knopf.

Gino, F., & Galinsky, A. D. (2012). Vicarious dishonesty: When psychological closeness creates distance from one's moral compass. *Organizational Behavior and Human Decision Processes, 119*, 15–26.

Gneezy, A., Imas, A., Brown, A., Nelson, L. D., & Norton, M. I. (2012). Paying to be nice: Consistency and costly prosocial behavior. *Management Science, 58*, 179–187.

Gold, B. P., Frank, M. J., Bogert, B., & Brattico, E. (2013). Pleasurable music affects reinforcement learning according to the listener. *Frontiers in Psychology, 4*.

Goldenberg, A., Garcia, D., Halperin, E., & Gross, J. J. (2020). Collective Emotions. *Current Directions in Psychological Science*, 29(2), 154–160.

Goldenberg, J. L., Courtney, E. P., & Felig, R. N. (2020, April 29). Supporting the dehumanization hypothesis, but under what conditions? A commentary on “Over.” *Perspectives on Psychological Science*.

Goldstein, N. J., Griskevicius, V., & Cialdini, R. B. (2011). Reciprocity by proxy: A new influence strategy for motivating cooperation and prosocial behavior. *Administrative Science Quarterly*, 56, 441–473.

Goldstein, N. J., Mortensen, C. R., Griskevicius, V., & Cialdini, R. B. (2007, January 16). I’ll scratch your back if you scratch my brother’s: The extended self and extradyadic reciprocity norms. Poster presented at the meeting of the Society of Personality and Social Psychology, Memphis, TN.

Goldstein, N. J., Vezich, I. S., & Shapiro, J. R. (2014). Perceived perspective taking: When others walk in our shoes. *Journal of Personality and Social Psychology*, 106, 941–960.

Gómez, Á., Dovidio, J. F., Gaertner, S. L., Fernández, S., & Vázquez, A. (2013). Responses to endorsement of commonality by in-group and outgroup members: The roles of group representation and threat. *Personality and Social Psychology Bulletin*, 39, 419–431.

Gonzales, M. H., Davis, J. M., Loney, G. L., Lukens, C. K., & Junghans, C. M. (1983). Interactional approach to interpersonal attraction. *Journal of Personality and Social Psychology*, 44, 1192–1197.

Goode, E., & Carey, B. (2015, October 7). Mass killings are seen as a kind of contagion. *New York Times*, A21.

Goodenough, U. W. (1991). Deception by pathogens. *American Scientist*, 79, 344–355.

Goodman-Delahunty, J., Martschuk, N., & Dhami, M. K. (2014). Interviewing high value detainees: Securing cooperation and disclosures. *Applied Cognitive Psychology*, 28, 883–897.

Gorn, G. J. (1982). The effects of music in advertising on choice behavior: A classical conditioning approach. *Journal of Marketing*, 46, 94–101.

Gould, M. S., & Shaffer, D. (1986). The impact of suicide in television movies. *New England Journal of Medicine*, 315, 690–694.

Grant, A. (2013). *Give and take*. New York: Viking.

Grant, A. M., & Hofmann, D. A. (2011). It's not all about me: Motivating hand hygiene among health care professionals by focusing on patients. *Psychological Science*, 22, 1494–1499.

Grant Halvorson, H., & Higgins, E. T. (2013). *Focus: Use different ways of seeing the world for success and influence*. New York: Penguin.

Green, F. (1965). The “foot-in-the-door” technique. *American Salesmen*, 10, 14–16. Greenberg, M. S., & Shapiro, S. P. (1971). Indebtedness: An adverse effect of asking for and receiving help. *Sociometry*, 34, 290–301.

Greene, J. (2014). *Moral tribes*. New York: Penguin.

Greenwald, A. F., Carnot, C. G., Beach, R., & Young, B. (1987). Increasing voting behavior by asking people if they expect to vote. *Journal of Applied Psychology*, 72, 315–318.

Greenwald, A. G., & Pettigrew, T. F. (2014). With malice toward none and charity for some. *American Psychologist*, 69, 669–684.

Greenwood, B. N., Hardeman, R. R., Huang, L., & Sojourner, A. (2020). Physician–patient racial concordance and disparities in birthing mortality for newborns. *Proceedings of the National Academy of Sciences*, 117, 21194–

21200.

Grey, K., Rand, D. G., Ert, E., Lewis, K., Hershman, S., & Norton, M. I. (2014). The emergence of “us and them” in 80 lines of code: Modeling group genesis in homogeneous populations. *Psychological Science*, 25, 982–990.

Griskevicius, V., Cialdini, R. B., & Kenrick, D. T. (2006). Peacocks, Picasso, and parental investment: The effects of romantic motives on creativity. *Journal of Personality and Social Psychology*, 91, 63–76.

Griskevicius, V., Goldstein, N. J., Mortensen, C. R., Sundie, J. M., Cialdini, R. C., & Kenrick, D. T. (2009). Fear and loving in Las Vegas: Evolution, emotion, and persuasion. *Journal of Marketing Research*, 46, 384–395.

Guadagno, R. E., & Cialdini, R. B. (2007). Persuade him by email, but see her in person: Online persuasion revisited. *Computers in Human Behavior*, 23, 999–1015.

Guéguen, N. (2009). Mimicry and seduction: An evaluation in a courtship context. *Social Influence*, 4, 249–255.

Guéguen, N. (2016). “You will probably refuse, but . . .”: When activating reactance in a single sentence increases compliance with a request. *Polish Psychological Bulletin*, 47, 170–173.

Guéguen, N., Jacob, C., & Meineri, S. (2011). Effects of the door-in-the-face technique on restaurant customers’ behavior. *International Journal of Hospitality Management*, 30, 759–761.

Guéguen, N., Joule, R. V., Halimi, S., Pascual, A., Fischer-Lokou, J., & Dufourcq-Brana, M. (2013). I’m free but I’ll comply with your request: Generalization and multidimensional effects of the “evoking freedom” technique. *Journal of Applied Social Psychology*, 43, 116–137.

Guéguen, N., Meineri, S., & Fischer-Lokou, J. (2014). Men's music ability and attractiveness to women in a real-life courtship contest. *Psychology of Music*, 42, 545–549.

Guéguen, N., & Pascual, A. (2000). Evocation of freedom and compliance: The “But you are free of . . .” technique. *Current Research in Social Psychology*, 5, 264–270.

Guéguen, N., & Pascual, A. (2014). Low-ball and compliance: Commitment even if the request is a deviant one, *Social Influence*, 9, 162–171.

Gyuris, P., Kozma L., Kisander Z., Láng A., Ferencz, T., & Kocsor, F. (2020). Sibling relations in patchwork families: Co-residence is more influential than genetic relatedness. *Frontiers of Psychology* 11:993.

Hadland, S. E., Cerda, M., Li, Y., Krieger, M. S., & Marshall, B. D. L. (2018). Association of pharmaceutical industry marketing with opioid products to physicians with subsequent opioid prescribing. *Journal of the American Medical Association: Internal Medicine*, 178, 861–863.

Hadnagy, C., & Schulman, S. (2020). *Human hacking: Win friends, influence people, and leave them better off for having met you*. New York: Harper Business.

Hagler, K. J, Pearson, M. R., Venner, B. L., & Greenfield, K. L. (2017). Descriptive drinking norms in Native American and non-Hispanic white college students. *Addictive Behaviors*, 72, 45–50.

Hakamata, Y., Lissek, S., Bar-Haim, Y., Britton, J. C., Fox, N. A., Leibenluft, E., Ernest, M., & Pine, D. S. (2010). Attention bias modification treatment: A metaanalysis toward the establishment of novel treatment for anxiety. *Biological Psychiatry*, 68, 982–990.

Halpern, D. (2016). *Inside the nudge unit: How small changes can make a*

big difference. London: Elbury.

Hamermesh, D. (2011). *Beauty pays: Why attractive people are more successful*. Princeton, NJ: Princeton University Press.

Hamermesh, D., & Biddle, J. E. (1994). Beauty and the labor market. *American Economic Review*, 84, 1174–1194.

Hamilton, W. D. (1964). The genetic evolution of social behavior. *Journal of Theoretical Biology*, 7, 1–52.

Hamlin, J. K., Mahajan, N., Liberman, Z., & Wynn, K. (2013). Not like me = bad: Infants prefer those who harm dissimilar others. *Psychological Science*, 24, 589–594.

Harmon-Jones, E., Harmon-Jones, C., & Levy, N. (2015). An action-based model of cognitive-dissonance processes. *Current Directions in Psychological Science*, 24, 184–189.

Harvey, A. R. (2018). Music and the meeting of human minds. *Frontiers in Psychology*, 9, 762.

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrated evolutionary model of cognitive biases. *Personality and Social Psychology Review*, 10, 47–66.

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, 10, 252–264.

Haslam, N., & Loughnan, S. (2014). Dehumanization and infrahumanization. *Annual Review of Psychology*, 65, 399–423.

Hassan, S. (2000). *Releasing the bonds: Empowering people to think for themselves*. Boston: Freedom of Mind Press.

Hatemi, P. K., & McDermott, R. (2012). The genetics of politics: Discovery,

challenges, and progress. *Trends in Genetics*, 28, 525–533.

Hattori, Y., & Tomonaga, M. (2020). Rhythmic swaying induced by sound in chimpanzees (*Pan troglodytes*). *Proceedings of the National Academy of Sciences*, 117, 936–942.

Häubl, G., & Popkowski Leszczyc, P. T. L. (2019). Bidding frenzy: Speed of competitor reaction and willingness to pay in auctions. *Journal of Consumer Research*, 45, 1294–1314.

Hehman, E., Flake, J. K., & Freeman, J. B. (2018). The faces of group members share physical resemblance. *Personality and Social Psychology Bulletin*, 44(1), 3–15.

Heijkoop, M., Dubas, J. S., & van Aken, M. A. G. (2009). Parent-child resemblance and kin investment. *European Journal of Developmental Psychology*, 6, 64–69.

Heilman, C. M., Nakamoto, K., & Rao, A. G. (2002). Pleasant surprises: Consumer response to unexpected in-store coupons. *Journal of Marketing Research*, 39, 242–252.

Heilman, M. E. (1976). Oppositional behavior as a function of influence attempt intensity and retaliation threat. *Journal of Personality and Social Psychology*, 33, 574–578.

Heinrich, C. U., & Borkenau, P. (1998). Deception and deception detection: The role of cross-modal inconsistency. *Journal of Personality*, 66, 687–712.

Henry, M. L., Ferraro, P. J., & Kontoleon, A. (2019). The behavioural effect of electronic home energy reports: Evidence from a randomised field trial in the United States. *Energy Policy*, 132, 1256–1261.

Hershfield, H. E., Goldstein, D. G., Sharpe, W. F., Fox, J., Yeykelis, L.,

Carstensen, L. L., & Bailenson, J. N. (2011). Increasing saving behavior through ageprogressed renderings of the future self. *Journal of Marketing Research*, 48, 23–37.

Higgins, C. A., & Judge, T. A. (2004). The effect of applicant influence tactics on recruiter perceptions of fit and hiring recommendations: A field study. *Journal of Applied Psychology*, 89, 622–632.

Higgins, E. T., (2012). *Beyond pleasure and pain: How motivation works*. New York: Oxford University Press.

Higgins, E. T., Lee, J., Kwon, J., & Trope, Y. (1995). When combining intrinsic motivations undermines interest. *Journal of Personality and Social Psychology*, 68, 749–767.

Higgins, E. T., Shah, J., & Friedman, R. (1997). Emotional responses to goal attainment: Strength of regulatory focus as moderator. *Journal of Personality and Social Psychology*, 72, 515–525.

Higham, P. A., & Carment, D. W. (1992). The rise and fall of politicians. *Canadian Journal of Behavioral Science*, 404–409.

Hildreth, J. A., & Anderson, C. (2018). Does loyalty trump honesty? Moral judgments of loyalty-driven deceit. *Journal of Experimental Social Psychology*, 79, 87–94. Hills, T. T. (2019). The dark side of information proliferation. *Perspectives on Psychological Science*, 14, 323–330.

Hills, T. T., Adelman, J. S., & Noguchi, T. (2017). Attention economies, information crowding, and language change. In M. N. Jones (ed.), *Big data in cognitive science* (pp.270–293). New York: Routledge.

Hobfoll, S. E. (2001). The influence of culture, community, and the nested-self in the stress process. *Applied Psychology: An International Review*, 50, 337–421.

Hodson, G. (2011). Do ideologically intolerant people benefit from intergroup contact? *Current Directions in Psychological Science*, 20, 154–159.

Hodson, G., Crisp, R. J., Meleady, R., & Earle, M. (2018). Intergroup contact as an agent of cognitive liberalization. *Perspectives on Psychological Science*, 13, 523–548.

Hofling, C. K., Brotzman, E., Dalrymple, S., Graves, N., & Pierce, C. M. (1966). An experimental study of nurse–physician relationships. *Journal of Nervous and Mental Disease*, 143, 171–180.

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin*, 136, 390–421.

Holmes, W. (2004). The early history of Hamiltonian-based research on kin recognition. *Annales Zoologici Fennici*, 41, 691–711.

Holmes, W. G., & Sherman, P. W. (1983). Kin recognition in animals. *American Scientist*, 71, 46–55.

Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, 27, 949–961.

Howard, D. J. (1990). The influence of verbal responses to common greetings on compliance behavior: The foot-in-the-mouth effect. *Journal of Applied Social Psychology*, 20, 1185–1196.

Howard, D. J., Shu, S. B., & Kerin, R. A. (2007). Reference price and scarcity appeals and the use of multiple influence strategies in retail newspaper advertising. *Social Influence*, 2, 18–28.

Howe, L. C., Carr, P. B., & Walton, G. W. (in press). Normative appeals are

more effective when they invite people to work together toward a common goal. *Journal of Personality and Social Psychology*.

Howe, M. L. (2003). Memories from the cradle. *Current Directions in Psychological Science*, 12, 62–65.

Hubbard, T. L. (2015). The varieties of momentum-like experience. *Psychological Bulletin*, 141, 1081–1119.

Hufer, A., Kornadt, A. E., Kandler, C., & Riemann, R. (2020). Genetic and environmental variation in political orientation in adolescence and early adulthood: A Nuclear Twin Family analysis. *Journal of Personality and Social Psychology*, 118, 762–776.

Hughes, S., Ye, Y., Van Dessel, P., & De Houwer, J. (2019). When people cooccur with good or bad events: Graded effects of relational qualifiers on evaluative conditioning. *Personality and Social Psychology Bulletin*, 45, 196–208.

Hugh-Jones, D., Ron, I., & Zultan, R. (2019). Humans discriminate by reciprocating against group peers. *Evolution and Human Behavior*, 40, 90–95.

Ilmarinen, V. J., Lönnqvist, J. E., & Paunonen, S. (2016). Similarity-attraction effects in friendship formation: Honest platoon-mates prefer each other but dishonest do not. *Personality and Individual Differences*, 92, 153–158.

Inzlicht, M., Gutsell, J. N., & Legault, L. (2012). Mimicry reduces racial prejudice. *Journal of Experimental Social Psychology*, 48, 361–365.

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, 76, 405–431.

Jabbi, M., Bastiaansen, J., & Keysers, C. (2008). A common anterior insula representation of disgust observation, experience and imagination shows divergent functional connectivity pathways. *PLoS ONE*, 3, e2939.

Jackson, J. C., Gelfand, M. J., Ayub, N., & Wheeler, J. (2019). Together from afar: Introducing a diary contact technique for improving intergroup relations. *Behavioral Science & Policy*, 5, 15–33.

Jacob, C., Guéguen, N., Martin, A., & Boulbry, G. (2011). Retail salespeople's mimicry of customers: Effects on consumer behavior. *Journal of Retailing and Consumer Services*, 18, 381–388.

James, J. M., & Bolstein, R. (1992). Effect of monetary incentives and follow-up mailings on the response rate and response quality in mail surveys. *Public Opinion Quarterly*, 54, 442–453.

Jenke, L., & Huettel, S. A. (2020) Voter preferences reflect a competition between policy and identity. *Frontiers of Psychology*, 11: 566020.

Jiang, L., Hoegg, J., Dahl, D. W., & Chattopadhyay, A. (2010). The persuasive role of incidental similarity on attitudes and purchase intentions in a sales context. *Journal of Consumer Research*, 36, 778–791.

John, L. K., Blunden, H., & Liu, H. (2019). Shooting the messenger. *Journal of Experimental Psychology: General*, 148(4), 644–666.

Johnson, D. W. (2003). Social interdependence: Interrelationships among theory, research, and practice. *American Psychologist*, 58, 934–945.

Jones, E. E., & Harris, V. E. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, 3, 1–24.

Jones, J. T., Pelham, B. W., Carvallo, M., & Mirenberg, M. C. (2004). How do I love thee? Let me count the J's. Implicit egoism and interpersonal

attraction. *Journal of Personality and Social Psychology*, 87, 665–683.

Jong, J., Whitehouse, H., Kavanagh, C., & Lane, J. (2015). Shared negative experiences lead to identity fusion via personal reflection. *PloS ONE*, 10.

Joule, R. V. (1987). Tobacco deprivation: The foot-in-the-door technique versus the low-ball technique. *European Journal of Social Psychology*, 17, 361–365.

Judge, T. A., & Cable, D. M. (2004). The effect of physical height on workplace success and income. *Journal of Applied Psychology*, 89, 428–441.

Judge, T. A., Hurst, C., & Simon, L. S. (2009). Does it pay to be smart, attractive, or confident (or all three)? Relationships among general mental ability, physical attractiveness, core self-evaluations, and income. *Journal of Applied Psychology*, 94, 742–755.

Jung, J., Busching, R., & Krahé, B. (2019). Catching aggression from one's peers: A longitudinal and multilevel analysis. *Social and Personality Psychology Compass*, 13.

Kahn, B. E., & Baron, J. (1995). An exploratory study of choice rules favored for high-stakes decisions. *Journal of Consumer Psychology*, 4, 305–328.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux. Kahneman, D., Slovic, P., & Tversky, A. (eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.

Kalmoe, N. P. (2019). Dueling views in a canonical measure of sophistication. *Public Opinion Research*, 83, 68–90.

Kaminski, G., Ravary, F., Graff, C., & Gentaz, E. (2010). Firstborns' disadvantage in kinship detection. *Psychological Science*, 21, 1746–1750.

Kandler, C., Bleidorn, W., & Riemann, R. (2012). Left or right? Sources of political orientation: The roles of genetic factors, cultural transmission, assortative mating, and personality. *Journal of Personality and Social Psychology*, 102, 633–645.

Kang, S. K., Hirsh, J. B., Chasteen, A. L. (2010). Your mistakes are mine: Selfother overlap predicts neural response to observed errors. *Journal of Experimental Social Psychology*, 46, 229–232.

Kardes, F. R., Posavac, S. S., & Cronley, M. L. (2004). Consumer inference: A review of processes, bases, and judgment contexts. *Journal of Consumer Psychology*, 14, 230–256.

Karim, A. A., Lützenkirchen, B., Khedr, E., & Khalil, R. (2017). Why is 10 past 10 the default setting for clocks and watches in advertisements? A psychological experiment. *Frontiers of Psychology* 8:1410.

Karremans, J. C., & Aarts, H. (2007). The role of automaticity in determining the inclination to forgive close others. *Journal of Experimental Social Psychology*, 43, 902–917.

Kauff, M., Asbrock, F., Thörner, S., & Wagner, U. (2013). Side effects of multiculturalism: The interaction effect of a multicultural ideology and authoritarianism on prejudice and diversity beliefs. *Personality and Social Psychology Bulletin*, 39, 305–320.

Kay, T., Keller, L., & Lehmann, L. (2020). The evolution of altruism and the serial rediscovery of the role of relatedness. *Proceedings of the National Academy of Sciences*, 117, 28894–28898.

Keil, F. C. (2012). Running on empty? How folk science gets by with less.

Current Directions in Psychological Science, 21, 329–334.

Keinan, A., & Kivetz, R. (2011). Productivity orientation and the consumption of collectable experiences. *Journal of Consumer Research*, 37, 935–950.

Kende, J., Phalet, K., Van den Noortgate, W., Kara, A., & Fischer, R. (2018). Equality revisited: A cultural meta-analysis of intergroup contact and prejudice. *Social Psychological and Personality Science*, 9, 887–895.

Kenrick, D. T. (2012). Evolutionary theory and human social behavior. In P. A. M. Van Lange, A. W. Kruglanski, and E. T. Higgins (eds.), *Handbook of Theories of Social Psychology* (pp. 11–31). Thousand Oaks, CA: Sage.

Kenrick, D. T. (2020). True friendships are communistic, not capitalist.

Kenrick, D. T., Gutierres, S. E., & Goldberg, L. L. (1989). Influence of popular erotica on judgments of strangers and mates. *Journal of Experimental Social Psychology*, 25, 159–167.

Kenrick, D. T., Neuberg, S. L., Cialdini, R. B., & Lundberg-Kenrick, D. E. (2020). *Social Psychology: Goals in interaction*. 7th ed. Boston: Pearson Education.

Kerr, N. L., & MacCoun, R. J. (1985). The effects of jury size and polling method on the process and product of jury deliberation. *Journal of Personality and Social Psychology*, 48, 349–363.

Kesebir, S. (2012). The superorganism account of human sociality: How and when human groups are like beehives. *Personality and Social Psychology Review*, 16, 233–261.

Ketelaar, T. (1995, June). *Emotions as mental representations of gains and losses: Translating prospect theory into positive and negative affect*. Paper

presented at the meeting of the American Psychological Society, New York, NY.

Kettle, K. I., & Häubl, G. (2011). The signature effect: Signing influences consumption-related behavior by priming self-identity. *Journal of Consumer Research*, 38, 474–489.

Keysar, B., Converse, B. A., Wang, J., & Epley, N. (2008). Reciprocity is not give and take: Asymmetric reciprocity to positive and negative acts. *Psychological Science*, 19, 1280–1286.

Khamitov, M., Wang, X., & Thomson, M. (2019). How well do consumer–brand relationships drive customer brand loyalty? Generalizations from a metaanalysis of brand relationship elasticities. *Journal of Consumer Research*, 46, 435–459.

Kimel, S. Y., Huesmann, R., Kunst, J. R., & Halperin, E. (2016). Living in a genetic world: How learning about interethnic genetic similarities and differences affects peace and conflict. *Personality and Social Psychology Bulletin*, 42, 688–700.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, 31, 354–364.

Kissinger, H. (1982). *Years of upheaval*. Boston: Little, Brown.

Klein, C. (2020, March 13). Why Coca-Cola’s “New Coke” flopped. History (website).

Klein, H. J., Lount, R. B., Jr., Park, H. M., & Linford, B. J. (2020). When goals are known: The effects of audience relative status on goal commitment and performance. *Journal of Applied Psychology*, 105, 372–389.

Klein, N., & O'Brien, E. (2018). People use less information than they think to make up their minds. *Proceedings of the National Academy of Sciences*.

Knishinsky, A. (1982). The effects of scarcity of material and exclusivity of information on industrial buyer perceived risk in provoking a purchase decision. Unpublished PhD diss., Arizona State University, Tempe.

Knouse, S. B. (1983). The letter of recommendation: Specificity and favorability information. *Personal Psychology*, 36, 331–341.

Knox, R. E., & Inkster, J. A. (1968). Postdecisional dissonance at post time. *Journal of Personality and Social Psychology*, 8, 319–323.

Koch, T., & Peter, C. (2017). Effects of equivalence framing on the perceived truth of political messages and the trustworthiness of politicians. *Public Opinion Quarterly*, 81, 847–865.

Koch, T., & Zerback, T. (2013). Helpful or harmful? How frequent repetition affects perceived statement credibility. *Journal of Communication*, 63, 993–1010.

Kouchaki, M., Gino, F., & Feldman, Y. (2019). The ethical perils of personal, communal relations: A language perspective. *Psychological Science*, 30, 1745–1766.

Koudenburg, N., Postmes, T., Gordijn, E. H., & van Mourik Broekman, A. (2015). Uniform and complementary social interaction: Distinct pathways to solidarity. *PloS ONE*, 10.

Krajbich, I., Camerer, C., Ledyard, J., & Rangel, A. (2009). Self-control in decisionmaking involves modulation of the vmPFC valuation system. *Science*, 324, 12315–12320.

Kranzler, D. (1976). *Japanese, Nazis, and Jews: The Jewish refugee*

community of Shanghai, 1938–1945. New York: Yeshiva University Press.

Kraut, R. E. (1973). Effects of social labeling on giving to charity. *Journal of Experimental Social Psychology*, 9, 551–562.

Kriner, D. L., & Shen, F. X. (2012). How citizens respond to combat casualties: The differential impact of local casualties on support for the war in Afghanistan. *Public Opinion Quarterly*, 76, 761–770.

Kristensson, P., Wästlund, E., & Söderlund, M. (2017). Influencing consumers to choose environment friendly offerings: Evidence from field experiments. *Journal of Business Research*, 76, 89–97.

Krizan, Z., & Suls, J. (2008). Losing sight of oneself in the above average effect: When egocentrism, focalism, and group diffusiveness collide. *Journal of Experimental Social Psychology*, 44, 929–942.

Kteily, N., Bruneau, E., Waytz, A., & Cotterill, S. (2015). The ascent of man: Theoretical and empirical evidence for blatant dehumanization. *Journal of Personality and Social Psychology*, 109, 901–931.

Ku, G. (2008). Before escalation: Behavioral and affective forecasting in escalation of commitment. *Personality and Social Psychology Bulletin*, 34, 1477–1491.

Kuester, M., & Benkenstein, M. (2014). Turning dissatisfied into satisfied customers: How referral reward programs affect the referrer's attitude and loyalty toward the recommended service provider. *Journal of Retailing and Consumer Services*, 21, 897–904.

Kunz, P. R., & Woolcott, M. (1976). Season's greetings: From my status to yours. *Social Science Research*, 5, 269–278.

Lack, D. (1943). *The life of the robin*. London: Cambridge University Press.

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S. Brauer, M., Devos, T., Calanchini, J., et al. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, 145, 1001–1016.

Lammers, H. B. (1991). The effect of free samples on immediate consumer purchase. *Journal of Consumer Marketing*, 8, 31–37.

Langer, E., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of “placebic” information in interpersonal interaction. *Journal of Personality and Social Psychology*, 36, 635–642.

Langer, E. J. (1989). Minding matters. In L. Berkowitz (ed.), *Advances in experimental social psychology* (vol. 22, pp. 137–173). New York: Academic Press.

Langlois, J. H., Kalakanis, A., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty: A meta-analytic and theoretical review. *Psychological Bulletin*, 126, 390–423.

Lansky, D. (2002, March 31). A day for stiffupperlipps, other nags. *Arizona Republic*, T4.

LaPorte, N. (2018). In a major reversal, Netflix is about to reveal how many people watch its most popular shows.

Latané, B., & Darley, J. M. (1968). Group inhibition of bystander intervention in emergencies. *Journal of Personality and Social Psychology*, 10, 215–221.

Law, S., & Braun, K., A. (2000). I’ll have what she’s having: Gauging the impact of product placements on viewers. *Psychology & Marketing*, 17, 1059–1075.

Lazarov, A., Abend, R., Seidner, S., Pine, D. S., & Bar-Haim, Y. (2017). The effects of training contingency awareness during attention bias modification on learning and stress reactivity. *Behavior Therapy*, 48, 638–650.

Leach, W. C., Ellemers, N., & Barreto M. (2007). Group virtue: The impact of morality(vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, 93, 234–249.

Leakey, R., & Lewin, R. (1978). *People of the lake*. New York: Anchor/Doubleday.

Lee, A. Y., & Aaker, J. L. (2004). Bringing the frame into focus: The influence of regulatory fit on processing fluency and persuasion. *Journal of Personality and Social Psychology*, 86, 205–218.

Lee, F., Peterson, C., & Tiedens, L. Z. (2004). Mea culpa: Predicting stock prices from organizational attributions. *Journal of Personality and Social Psychology*, 30, 1636–1649.

Lefkowitz, M., Blake, R. R., & Mouton, J. S. (1955). Status factors in pedestrian violation of traffic signals. *Journal of Abnormal and Social Psychology*, 51, 704–706.

Legate, N., Weinstein, N., Sendi, K., & Al-Khouja, M. (2020). Motives behind the veil: Women's affective experiences wearing a veil depend on their reasons for wearing one. *Journal of Research in Personality*, 87, 103969.

Leopold, A. (1989). *A Sand County almanac*. New York: Oxford University Press. Leotti, L. A., & Delgado, M. R. (2011). The inherent reward of choice. *Psychological Science*, 22, 1310–1318.

Lepper, M. R., & Greene, D. (eds.). (1978). *The hidden costs of reward*. Hillsdale, NJ: Lawrence Erlbaum.

Levendusky, M. S. (2018). Americans, not partisans: Can priming American national identity reduce affective polarization? *Journal of Politics*, 80, 59–70.

Levine, H. (1997). *In search of Sugihara*. New York: Free Press.

Levy J., Markell, D., & Cerf, M. (2019). Polar similars: Using massive mobile dating data to predict synchronization and similarity in dating preferences. *Frontiers of Psychology*.

Lewis, G. J., & Bates, T. C. (2010). Genetic evidence for multiple biological mechanisms underlying in-group favoritism. *Psychological Science*, 21, 1623–1628.

Li, W., Moallem, I., Paller, K. A., Gottfried, J. A. (2007). Subliminal smells can guide social preferences. *Psychological Science*, 18, 1044–1049.

Lieberman, J. D., & Arndt, J. (2000). Understanding the limits of limiting instructions. *Psychology, Public Policy, and Law*, 6, 677–711.

Lieberman, D., & Smith, A. (2012). It's all relative: Sexual aversions and moral judgments regarding sex among siblings. *Current Directions in Psychological Science*, 21, 243–247.

Lieberman, D., Tooby, J., & Cosmides, L. (2007). The architecture of human kin detection. *Nature*, 445, 727–731.

Lim, S., O'Doherty, J. P., & Rangel, A. (2011). The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *Journal of Neuroscience*, 31, 13214–13223.

Lin, J. S., & Sung, Y. (2014). Nothing can tear us apart: The effect of brand identity fusion in consumer–brand relationships. *Psychology & Marketing*, 31, 54– 69.

Linder, J. A., Meeker, D., Fox, C. R., Friedberg, M. W., Persell, S. D.,

- Goldstein, N. J., & Doctor, J. N. (2017). Effects of behavioral interventions on inappropriate antibiotic prescribing in primary care 12 months after stopping interventions. *Journal of the American Medical Association*, 318, 1391–1392.
- Liu, W., & Gal, D. (2011). Bringing us together or driving us apart: The effect of soliciting consumer input on consumers' propensity to transact with an organization. *Journal of Consumer Research*, 38, 242–259.
- Lloyd, J. E. (1965). Aggressive mimicry in *Photuris*: Firefly *femme fatales*. *Science*, 149, 653–654.
- Loersch, C., & Arbuckle, N. L. (2013). Unraveling the mystery of music: Music as an evolved group process. *Journal of Personality and Social Psychology*, 105, 777–798.
- Lokhorst, A. M., Werner, C., Staats, H., van Dijk, E., & Gale, J. L. (2013). Commitment and behavior change: A meta-analysis and critical review of commitment-making strategies in environmental research. *Environment and Behavior* 45, 3–34.
- Loschelder, D. D, Siepelmeyer, H., Fischer, D., & Rubel, J. (2019). Dynamic norms drive sustainable consumption: Norm-based nudging helps café customers to avoid disposable to-go-cups. *Journal of Economic Psychology*, 75, 102146.
- Lynn, M. (1989). Scarcity effect on value: Mediated by assumed expensiveness. *Journal of Economic Psychology*, 10, 257–274.
- Lytton, J. (1979). Correlates of compliance and the rudiments of conscience in two-year-old boys. *Canadian Journal of Behavioral Science*, 9, 242–251.
- MacGregor-Hastie, R. (1988). *Picasso's women*. London: Lennard.

Mack, D., & Rainey, D. (1990). Female applicants' grooming and personnel selection. *Journal of Social Behavior and Personality*, 5, 399–407.

MacKay, C. (1841/1932). *Extraordinary Popular Delusions and the Madness of Crowds*. New York: Farrar, Straus and Giroux.

MacKenzie, B. (1974, June 22). When sober executives went on a bidding binge. *TV Guide*.

Madanika, Y., & Bartholomew, K. (2014, August 14). Themes of lust and love in popular music from 1971 to 2011. *SAGE Open*, 4(3).

Maddux, W. W., Mullen, E., & Galinsky, A. (2008). Chameleons bake bigger pies and take bigger pieces: Strategic behavioral mimicry facilitates negotiation outcomes. *Journal of Experimental Social Psychology*, 44, 461–468.

Maestriperi, D., Henry, A., & Nickels, N. (2017). Explaining financial and prosocial biases in favor of attractive people: Interdisciplinary perspectives from economics, social psychology, and evolutionary psychology. *Behavioral and Brain Sciences*, 40,E19.

Maglio, S. J., & Polman, E. (2016). Revising probability estimates: Why increasing likelihood means increasing impact. *Journal of Personality and Social Psychology*, 111, 141–158.

Magruder, J. S. (1974). *An American life: One man's road to Watergate*. New York: Atheneum.

Mahajan, N., Martinez, M. A., Gutierrez, N. L., Diesendruck, G., Banaji, M. R., & Santos, L. R. (2011). *Journal of Personality and Social Psychology*, 100, 387–405.

Maheshwari, S. (2019, November 29). The online star rating system is flawed

. . . and you never know if you can trust what you read. *New York Times*, B1, B4.

Main, K. J., Dahl, D. W., & Darke, P. R. (2007). Deliberative and automatic bases of suspicion: Empirical evidence of the sinister attribution error. *Journal of Consumer Psychology*, 17, 59–69.

Maisel, N. C., & Gable, S. L. (2009). The paradox of received social support: The importance of responsiveness. *Psychological Science*, 20, 928–932.

Makary, M. A., & Daniel, M. (2016). Medical error—the third leading cause of death in the US. *BMJ*, 353.

Makurdi, W. (2013, May 26). 23 youths dead in South Africa during adulthood initiation rites. *Arizona Republic*, A32.

Malinowski, B. (1922). *Argonauts of the Western Pacific: An account of native enterprise and adventure in the archipelagoes of Melanesian New Guinea*.

Mallett, R. K., Wilson, T. D., & Gilbert, D. T. (2008). Expect the unexpected: Failure to anticipate similarities leads to an intergroup forecasting error. *Journal of Personality and Social Psychology*, 94, 265–277.

Mallon, B., Redies, C., & Hayn-Leichsenring, G. U. (2014). Beauty in abstract paintings: Perceptual contrast and statistical properties. *Frontiers of Human Neuroscience*, 8, 161.

Mandel, N., & Johnson, E. J. (2002). When web pages influence choice: Effects of visual primes on experts and novices. *Journal of Consumer Research*, 29, 235–245.

Manis, M., Cornell, S. D., & Moore, J. C. (1974). Transmission of attitude relevant information through a communication chain. *Journal of Personality*

and Social Psychology, 30, 81–94.

Mann, T., Nolen-Hoeksema, S. K., Burgard, D., Huang, K., Wright, A., & Hansen, K. (1997). Are two interventions worse than none? *Health Psychology*, 16, 215–225.

Mannes, A. E., Soll, J. B., & Larrick, R. P. (2014). The wisdom of select crowds. *Journal of Personality and Social Psychology*, 107, 276–299.

Manning, R., Levine, M., & Collins, A. (2007). The Kitty Genovese murder and the social psychology of helping: The parable of the 38 witnesses. *American Psychologist*, 62, 555–562.

Marcoux, J-S. (2009). Escaping the gift economy. *Journal of Consumer Research*, 36, 671–685.

Margulis, E. H. (2010). When program notes don't help: Music descriptions and enjoyment. *Psychology of Music*, 38, 285–302.

Markman, K. D., & Guenther, C. L. (2007). Psychological momentum: Intuitive physics and naive beliefs. *Personality and Social Psychology Bulletin*, 33, 800–812.

Markowitz, D. M., & Slovic, P. (2020). Social, psychological, and demographic characteristics of dehumanization toward immigrants. *Proceedings of the National Academy of Sciences*, 117, 9260–9269.

Marks, J., Copland, E., Loh, E., Sunstein, C. R., Sharot, T. (2019). Epistemic spillovers: Learning others' political views reduces the ability to assess and use their expertise in nonpolitical domains. *Cognition*, 188, 74–84.

Martin, S. J., Goldstein, N. J., & Cialdini, R. B. (2014). *The small BIG: Small changes that spark big influence*. New York: Grand Central Publishing.

Martin, S. J., & Marks, J. (2019). *Messengers: Who we listen to, who we*

don't, and why. New York: Public Affairs.

Mashek, D. J., Aron, A., & Boncimino, M. (2003). Confusions of self with close others. *Personality and Social Psychology Bulletin*, 29, 382–392.

Masket, S. E. (2009). Did Obama's ground game matter? The influence of local field offices during the 2008 presidential election. *Public Opinion Quarterly*, 73, 1023–1039.

Mateo, J. M. (2003). Kin recognition in ground squirrels and other rodents. *Journal of Mammalogy*, 84, 1163–1181.

Mateo, J. M. (2015). Perspectives: Hamilton's legacy: mechanisms of kin recognition in humans. *Ethology* 121, 419–427.

Mather, M., Shafir, E., & Johnson, M. K. (2000). Misremembrance of options past: Source monitoring and choice. *Psychological Science*, 11, 132–138.

Matthies, E., Klöckner, C. A., Preißner, C. L. (2006). Applying a modified moral decision making model to change habitual car use: How can commitment be effective? *Applied Psychology* 55, 91–106.

Maus, G. W., Goh, H. L., & Lisi, M. (2020). Perceiving locations of moving objects across eyeblinks. *Psychological Science*.

Mauss, M. (1990). *The gift: The form and reason for exchange in archaic societies*. Translated by W. D. Halls. Abingdon: Routledge.

Mauss, S. (December 4, 2012). "Hitler's Jewish Commander and Victim." *Jewish Voice from Germany*.

Mayer, N. D., & Tormala, Z. (2010). "Think" versus "feel" framing effects. *Personality and Social Psychology Bulletin*, 36, 443–454.

Mazis, M. B. (1975). Antipollution measures and psychological reactance theory: A field experiment. *Journal of Personality and Social Psychology*,

31, 654–666.

Mazis, M. B., Settle, R. B., & Leslie, D. C. (1973). Elimination of phosphate detergents and psychological reactance. *Journal of Marketing Research*, 10, 390–395.

McCall, M., & Belmont, H. J. (1996). Credit card insignia and restaurant tipping: Evidence for an associative link. *Journal of Applied Psychology*, 81, 609–613.

McDonald, M., Porat, R., Yarkoney, A., Reifen Tagar, M., Kimel, S., Saguy, T., & Halperin, E. (2017). Intergroup emotional similarity reduces dehumanization and promotes conciliatory attitudes in prolonged conflict. *Group Processes & Intergroup Relations*, 20, 125–136.

McFarland, S. (2017). Identification with all humanity: The antithesis of prejudice, and more. In C. G. Sibley and F. K. Barlow (eds.), *The Cambridge handbook of the psychology of prejudice* (pp. 632–654). Cambridge: Cambridge University Press.

McFarland, S., Webb, M., & Brown D. (2012). All humanity is my in-group: A measure and studies of identification with all humanity. *Journal of Personality and Social Psychology*, 103, 830–853.

McGuinnies, E., & Ward, C. D. (1980). Better liked than right: Trustworthiness and expertise as factors in credibility. *Personality and Social Psychology Bulletin*, 6, 467–472.

McKenzie, C. R. M., & Chase, V. M. (2010). Why rare things are precious: The importance of rarity in lay inference. In P. M. Todd, G. Gigerenzer, and the ABC Research Group (eds.), *Ecological rationality: Intelligence in the world* (pp. 81–101). Oxford: Oxford University Press.

McKeown, S., & Dixon, J. (2017). The “contact hypothesis”: Critical

reflections and future directions. *Social & Personality Psychology Compass*, 11.

McNeill, W. H. (1995). *Keeping together in time: Dance and drill in human history*. Cambridge, MA: Harvard University Press.

Meeker, D., Knight, T. K, Friedberg, M. W., Linder, J. A., Goldstein, N. J., Fox, C. R., Rothfeld, A., et al. (2014). Nudging guideline-concordant antibiotic prescribing: A randomized clinical trial. *JAMA Internal Medicine*, 174, 425–431.

Meeus, W. H. J., & Raaijmakers, Q. A. W. (1986). Administrative obedience: Carrying out orders to use psychological-administrative violence. *European Journal of Social Psychology*, 16, 311–324.

Meier, B. P, Dillard, A. J, & Lappas, C. M. (2019). Naturally better? A review of the natural is better bias. *Social and Personality Psychology Compass*.

Melamed, B. F., Yurcheson, E., Fleece, L., Hutcherson, S., & Hawes, R. (1978). Effects of film modeling on the reduction of anxiety-related behaviors in individuals varying in level of previous experience in the stress situation. *Journal of Consulting and Clinical Psychology*, 46, 1357–1374.

Melamed, D., Simpson, B., & Abernathy, J. (2020). The robustness of reciprocity: Experimental evidence that each form of reciprocity is robust to the presence of other forms of reciprocity. *Science Advances*, 6.

Mercer, A., Caporaso, A., Cantor, D., & Townsend, J. (2015). How much gets you how much? Monetary incentives and response rates in household surveys. *Public Opinion Quarterly*, 79, 105–129.

Meredith, J. (2020). Conversation analysis, cyberpsychology and online interaction. *Social and Personality Psychology Compass*, 14.

Meyerwitz, B. E., & Chaiken, S. (1987). The effect of message framing on breast self-examination attitudes, intentions, and behavior. *Journal of Personality and Social Psychology*, 52, 500–510.

Michelitch, K. (2015). Does electoral competition exacerbate interethnic or interpartisan economic discrimination? Evidence from a field experiment in market price bargaining. *American Political Science Review*, 109, 43–61.

Midlarsky, E., & Nemeroff, R. (1995, July). Heroes of the holocaust: Predictors of their well-being in later life. Poster presented at the American Psychological Society meetings, New York, NY.

Milgram, S. (1970). The experience of living in cities: A psychological analysis. In F. F. Korten, S. W. Cook, & J. I. Lacey (eds.), *Psychology and the problems of society* (pp.152–173). American Psychological Association.

Milgram, S. (1974). *Obedience to authority*. New York: Harper & Row.

Milgram, S., Bickman, L., & Berkowitz, O. (1969). Note on the drawing power of crowds of different size. *Journal of Personality and Social Psychology*, 13, 79–82.

Miller, C. B. (2009). Yes we did!: Basking in reflected glory and cutting off reflected failure in the 2008 presidential election. *Analyses of Social Issues and Public Policy*, 9, 283–296.

Miller, C. H., Burgoon, M., Grandpre, J. R., & Alvaro, E.M. (2006). Identifying principal risk factors for the initiation of adolescent smoking behaviors: The significance of psychological reactance. *Health Communication*, 19, 241–252.

Miller, G. F. (2000). *The mating mind*. New York: Doubleday.

Miller, J. M., & Krosnick, J. A. (1998). The impact of candidate name order

on election outcomes. *Public Opinion Quarterly*, 62, 291–330.

Miller, N., Campbell, D. T., Twedt, H., & O'Connell, E. J. (1966). Similarity, contrast, and complementarity in friendship choice. *Journal of Personality and Social Psychology*, 3, 3–12.

Miller, R. L., Brickman, P., & Bollen, D. (1975). Attribution versus persuasion as a means of modifying behavior. *Journal of Personality and Social Psychology*, 31, 430–441.

Miller, R. L., Seligman, C., Clark, N. T., & Bush, M. (1976). Perceptual contrast versus reciprocal concession as mediators of induced compliance. *Canadian Journal of Behavioral Science*, 8, 401–409.

Mills, C. M., & Keil, F. C. (2005). The development of cynicism. *Psychological Science*, 16, 385–390.

Mita, T. H., Dermer, M., & Knight, J. (1977). Reversed facial images and the mere exposure hypothesis. *Journal of Personality and Social Psychology*, 35, 597–601.

Mogg, K., Waters, A. M., & Bradley, B. P. (2017). Attention bias modification (ABM): Review of effects of multisession ABM training on anxiety and threat-related attention in high-anxious individuals. *Clinical Psychological Science*, 5, 698–717.

Monahan, J. L., Murphy, S. T., & Zajonc, R. B. (2000). Subliminal mere exposure: Specific, general, and diffuse effects. *Psychological Science*, 11, 462–466.

Moons, W. G., Mackie, D. M., & Garcia-Marques, T. (2009). The impact of repetition-induced familiarity on agreement with weak and strong arguments. *Journal of Personality and Social Psychology*, 96, 32–44.

Moore, C., & Pierce, L. (2016). Reactance to transgressors: Why authorities deliver harsher penalties when the social context elicits expectations of leniency. *Frontiers in Psychology*, 7, 550.

Moore, D. E., Kurtzberg, T. R., Thompson, L. L., & Morris, M. W. (1999). Long and short routes to success in electronically-mediated negotiations: Group affiliations and good vibrations. *Organizational Behavior and Human Decision Processes*, 77, 22–43.

Moreland, R. L., & Topolinski, S. (2010). The mere exposure phenomenon: A lingering melody by Robert Zajonc. *Emotion Review*, 2, 329–339.

Moriarty, T. (1975). Crime, commitment, and the responsive bystander: Two field experiments. *Journal of Personality and Social Psychology*, 31, 370–376.

Morris, M., Nadler, J., Kurtzberg, T., & Thompson, L. (2002). Schmooze or lose: Social friction and lubrication in e-mail negotiations. *Group Dynamics: Theory, Research, and Practice*, 6, 89–100.

Morrison, K. R., Plaut, V. C., & Ybarra, O. (2010). Predicting whether multiculturalism positively or negatively influences white Americans' intergroup attitudes: The role of ethnic identification. *Personality and Social Psychology Bulletin*, 36, 1648–1661.

Morrow, L. (1991, September 2). The Russian revolution, *Time*, 20.

Mortensen, C. H., Neel, R., Cialdini, R. B., Jaeger, C. M., Jacobson, R. P., & Ringel, M. M. (2017). Upward trends: A lever for encouraging behaviors performed by the minority. *Social Psychology and Personality Science*.

Mousa, S. (2020). Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq, *Science*, 369, 866–870.

Mrkva, K., & Van Boven, L. (2020). Salience theory of mere exposure: Relative exposure increases liking, extremity, and emotional intensity. *Journal of Personality and Social Psychology*, 118, 1118–1145.

Mulla, M. M., Witte, T. H., Richardson, K., Hart, W., Kassing, F. L., Coffey, C. A., Hackman, C. L., & Sherwood, I. M. (2019). The causal influence of perceived social norms on intimate partner violence perpetration: Converging cross-sectional, longitudinal, and experimental support for a social disinhibition model. *Personality and Social Psychology Bulletin*, 45, 652–668.

Murayama, K., & Elliot, A. J. (2012). The competition–performance relation: A metaanalytic review and test of the opposing processes model of competition and performance. *Psychological Bulletin*, 138, 1035–1070.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition and awareness. *Journal of Personality and Social Psychology*, 64, 723–739.

Murrar, S., Campbell, M. R., & Brauer, M. (2020). Exposure to peers' prodiversity attitudes increases inclusion and reduces the achievement gap. *Nature Human Behavior*.

Murray, D. A., Leupker, R. V., Johnson, C. A., & Mittlemark, M. B. (1984). The prevention of cigarette smoking in children: A comparison of four strategies. *Journal of Applied Social Psychology*, 14, 274–288.

Nai, J., Narayanan, J., Hernandez, I., & Savani, K. (2018). People in more racially diverse neighborhoods are more prosocial. *Journal of Personality and Social Psychology*, 114, 497–515.

Nakayachi, K., Ozaki, T., Shibata, Y., & Yokoi, R. (2020). Why do Japanese people use masks against COVID-19, even though masks are unlikely to offer protection from infection? *Frontiers in Psychology*, 11.

Naylor, R. W., Raghunathan, R., & Ramanathan, S. (2006). Promotions spontaneously induce a positive evaluative response. *Journal of Consumer Psychology, 16*, 295–305.

Nelissen, R. M. A., & Meijers, M. H. C. (2011). Social benefits of luxury brands as costly signals of wealth and status. *Evolution and Human Behavior, 32*, 343–355.

News. (1988). *Stanford Business School Magazine, 56*, 3.

Nijjer, R. (2019). 5 types of social proof to use on your website now. *Search Engine Journal*.

Nolan, J. M., Schultz, P. W., Cialdini, R. B., & Goldstein, N. J. (2021). The social norms approach: A wise intervention for solving social and environmental problems. In G. Walton and A. Crum (eds.) *Handbook of Wise Interventions*. (pp. 405–428). Guilford.

Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2008). Normative social influence is underdetected. *Personality and Social Psychology Bulletin, 34*, 913–923.

Noor, M., Brown, R., Gonzalez, R., Manzi, Jorge, & Lewis, C. A. (2008). On positive psychological outcomes: What helps groups with a history of conflict to forgive and reconcile with each other? *Personality and Social Psychology Bulletin, 34*, 819–832.

Norscia, I., & Palagi, E. (2011). Yawn contagion and empathy in *Homo sapiens*. *PLoS ONE, 6*.

Norscia, I., Zanolli, A., Gamba, M., & Palagi, E. (2020). Auditory contagious yawning is highest between friends and family members: Support to the emotional bias hypothesis. *Frontiers of Psychology, 11*, 442.

Norton, M. I., Mochon, D., & Ariely, D. (2012). The IKEA effect: When labor leads to love. *Journal of Consumer Psychology*, 22, 453–460.

Oesch, N. (2019). Music and language in social interaction: Synchrony, antiphony, and functional origins. *Frontiers of Psychology*, 10, 1514.

Oh, D., Shafir, E., & Todorov, A. (2020). Economic status cues from clothes affect perceived competence from faces. *Nature Human Behaviour*, 4(3), 287–293.

Ohadi, J., Brown, B., Trub, L., & Rosenthal, L. (2018). I just text to say I love you: Partner similarity in texting and relationship satisfaction. *Computers in Human Behavior*, 78, 126–132.

Oliver, A. (2019). *Reciprocity and the Art of Behavioural Public Policy*. Cambridge: Cambridge University Press.

O’Leary, S. G. (1995). Parental discipline mistakes. *Current Directions in Psychological Science*, 4, 11–13.

Oliner, S. P., & Oliner, P. M. (1988). *The altruistic personality: Rescuers of Jews in Nazi Europe*. New York: Free Press.

Olson, I. R., & Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion*, 5, 498–502.

Olson, J. M., & James, L. M. (2002). Vigilance for differences. *Personality and Social Psychology Bulletin*, 28, 1084–1093.

Onyeador, I. N., Wittlin, N. M., Burke, S. E., Dovidio, J. F., Perry, S. P., Hardeman, R.R., Dyrbye, L. N., et al. (2020). The value of interracial contact for reducing anti-Black bias among non-Black physicians: A cognitive habits and growth evaluation (CHANGE) study report. *Psychological Science*, 31, 18–30.

Oosterhof, N. N., Tipper, S. P., & Downing, P. E. (2012). Visuo-motor imagery of specific manual actions: A multi-variate pattern analysis fMRI study, *NeuroImage*, 63, 262–271.

Oriña, M. M., Wood, W., & Simpson, J. A. (2002). Strategies of influence in close relationships. *Journal of Experimental Social Psychology*, 38, 459–472.

Oskamp, S., & Schultz, P. W. (1998). *Applied Social Psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Ott, M., Choi, Y., Cardie, C., & Hancock, J. T. (2011). Finding deceptive opinion spam by any stretch of the imagination. Proceedings of the 49th Annual Meeting of the Association for Computer Linguistics, 309–319. Portland, Oregon. Otten, S., & Epstude, K. (2006). Overlapping mental representations of self, ingroup, and outgroup: Unraveling self-stereotyping and self-anchoring. *Personality and Social Psychology Bulletin*, 32, 957–969.

Over, H. (2020, April 29). Seven challenges for the dehumanization hypothesis. *Perspectives on Psychological Science*.

Over, H., & McCall, C. (2018). Becoming us and them: Social learning and intergroup bias. *Social and Personality Compass*.

Packard, V. (1957). *The hidden persuaders*. New York: D. McKay.

Paez, D., Rime, B., Basabe, N., Włodarczyk, A., & Zumeta, L. (2015). Psychosocial effects of perceived emotional synchrony in collective gatherings. *Journal of Personality and Social Psychology*, 108, 711–729.

Paese, P. W., & Gilin, D. A. (2000). When an adversary is caught telling the truth. *Personality and Social Psychology Bulletin*, 26, 75–90.

Page-Gould, E., Mendoza-Denton, R., Alegre, J. M., & Siy, J. O. (2010). Understanding the impact of cross-group friendship on interactions with

novel outgroup members. *Journal of Personality and Social Psychology*, 98, 775–793.

Page-Gould, E., Mendoza-Denton, R., & Tropp, L. R. (2008). With a little help from my cross-group friend: Reducing anxiety in intergroup contexts through cross-group friendship. *Journal of Personality and Social Psychology*, 95, 1080–1094.

Palagi, E., Leone, A., Mancini, G., & Ferrari, P. F. (2009). Contagious yawning in gelada baboons as a possible expression of empathy. *Proceedings of the National Academy of Sciences* 106, 19262–19267.

Palidino, M-P., Mazzurega, M., Pavani, F., & Schubert, T. W. (2010). Synchronous multisensory stimulation blurs self-other boundaries. *Psychological Science*, 21,1202–1207.

Pallak, M. S., Cook, D. A., & Sullivan, J. J. (1980). Commitment and energy conservation. *Applied Social Psychology Annual*, 1, 235–253.

Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology*, 96, 574–587.

Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *Annual Review of Psychology*, 60, 339–367.

Pane, L. M. (2019, December 29). Study: US mass killings reach new high in 2019. *Arizona Republic*, 8A.

Paolini, S., Hewstone, M., Cairns, E., & Voci, A. (2004). Effects of direct and indirect cross-group friendships on judgments of Catholics and Protestants in Northern Ireland. *Personality and Social Psychology Bulletin*, 30, 770–786.

Park, H., Lalwani, A. K., & Silvera, D. H. (2020). The impact of resource scarcity on price-quality judgments. *Journal of Consumer Research*, 46, 1110–1124.

Park, J. H., & Schaller, M. (2005). Does attitude similarity serve as a heuristic cue for kinship? Evidence of an implicit cognitive association. *Evolution and Human Behavior*, 26, 158–170.

Park, J. H., Schaller, M., & Van Vugt, M. (2008). Psychology of human kin recognition: Heuristic cues, erroneous inferences, and their implications. *Review of General Psychology*, 12, 215–235.

Parkinson, B. (2020). Intragroup emotion convergence: Beyond contagion and social appraisal. *Personality and Social Psychology Review*, 24, 121–140.

Parsons, C. A., Sulaeman, J., Yates, M. C., & Hamermesh, D. S. (2011). Strike three: Discrimination, incentives, and evaluation. *American Economic Review*, 101, 1410–1435.

Pavlov, I. P. (1927). *Conditioned reflexes*. Translated by G. V. Anrep. Oxford: Oxford University Press.

Peiponen, V. A. (1960). Verhaltensstudien am blaukehlchen [Behavior studies of the blue-throat]. *Ornis Fennica*, 37, 69–83.

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12), 1865–1880.

Perry, G. (2012). *Behind the shock machine: The untold story of the notorious Milgram psychology experiments*. Melbourne: Scribe.

Pettigrew, T. F. (1997). Generalized intergroup contact effects on prejudice.

Personality and Social Psychology Bulletin, 23, 173–185.

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90, 751–783.

Petrova, P. K., Cialdini, R. B., & Sills, S. J. (2007). Personal consistency and compliance across cultures. *Journal of Experimental Social Psychology*, 43, 104–111.

Petty, R. E., Briñol, P., Fabrigar, L., & Wegener, D. T. (2019). Attitude structure and change. In R. Baumeister, and E. Finkel (eds.). *Advanced Social Psychology* (pp.117–156). New York: Oxford University Press.

Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, 41, 847–855.

Pfeffer, J., & Cialdini, R. B. (1998). Illusions of influence. In R. M. Kramer & M. A. Neale(eds). *Power and Influence in Organizations* (pp. 1–20). Thousand Oaks, CA: Sage.

Phillips, D. P. (1974). The influence of suggestion on suicide: Substantive and theoretical implications of the Werther effect. *American Sociological Review*, 39, 340–354.

Phillips, D. P. (1979). Suicide, motor vehicle fatalities, and the mass media: Evidence toward a theory of suggestion. *American Journal of Sociology*, 84, 1150–1174.

Phillips, D. P. (1980). Airplane accidents, murder, and the mass media: Towards a theory of imitation and suggestion. *Social Forces*, 58, 1001–1024.

Phillips, D. P., & Cartensen, L. L. (1986). Clustering of teenage suicides after television news stories about suicide. *New England Journal of Medicine*, 315,

685–689. Phillips, D. P., & Cartensen, L. L. (1988). The effect of suicide stories on various demographic groups, 1968–1985. *Suicide and Life-Threatening Behavior*, 18, 100–114.

Philpot, R., Liebst, L. S., Levine, M., Bernasco, W., & Lindegaard, M. R. (2020). Would I be helped? Cross national CCTV footage shows that intervention is the norm in public conflicts. *American Psychologist*, 75, 66–75.

Pierce, J. R., Kilduff, G. J., Galinsky, A. D., & Sivanathan, N. (2013). From glue to gasoline: How competition turns perspective takers unethical. *Psychological Science*, 24, 1986–1994.

Pinsker, J. (2014, October 1). The psychology behind Costco's Free Samples: Mini pizza bagels? Now we're talking. *The Atlantic*.

Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences*, 105, 1050–1054.

Pope, B. R., & Pope, N. G. (2015). Own-nationality bias: Evidence from UEFA Champions League football referees. *Economic Inquiry* 53, 1292–1304.

Pope, D. G., & Schweitzer, M. E. (2011). Is Tiger Woods loss averse?: Persistent bias in the face of experience, competition, and high stakes. *American Economic Review*, 101. 129–157.

Poulin-Dubois, D., Brooker, I., & Polonia, A. (2011). Infants prefer to imitate a reliable person. *Infant Behavior and Development*, 34, 303–309.

Powers, N., Blackman, A., Lyon, T. P., & Narain, U. (2011). Does disclosure reduce pollution?: Evidence from India's Green Rating Project. *Environmental and Resource Economics*, 50, 131–155.

Poza, D. (2016). 7 simple hacks to supercharge your registration process.

Pratkanis, A. R. (2000). Altercasting as an influence tactic. In D. J. Terry and M. A. Hogg (eds.), *Attitudes, behavior, and social context* (pp. 201–226). Mahwah, NJ: Lawrence Erlbaum.

Pratkanis, A. R. (2007). Social influence analysis: An Index of tactics. In A. R. Pratkanis(ed.), *The science of social Influence: Advances and future progress* (pp. 17–82). Philadelphia, PA: Philadelphia Free Press.

Pratkanis, A., & Shadel, D. (2005). *Weapons of fraud: A sourcebook for fraud fighters*. Seattle, WA: AARP Washington.

Pratkanis, A. R., & Uriel, Y. (2011). The expert snare as an influence tactic: Surf, turf, and ballroom demonstrations of some compliance consequences of being altercast as an expert. *Current Psychology*, 30, 335–344.

Prelec, D., & Simester, D. (2001). Always leave home without it: A further investigation of the credit-card effect on willingness to pay. *Marketing Letters*, 12, 5–12. Preston, S. D. (2013). The origins of altruism in offspring care. *Psychological Bulletin*, 139, 1305–1341.

Price, R. B., Wallace, M., Kuckertz, J. M., Amir, N., Graur, S., Cummings, L., Popa, P., et al. (2016). Pooled patient-level meta-analysis of children and adults completing a computer-based anxiety intervention targeting attentional bias. *Clinical Psychology Review*, 50, 37–49.

Price, J., & Wolfers, J. (2010). Racial discrimination among NBA referees. *Quarterly Journal of Economics* 125, 1859–1887.

Priebe, C. S., & Spink, K. S. (2011). When in Rome: Descriptive norms and physical activity. *Psychology of Sport and Exercise*, 12, 93–98.

Provine, R. (2000). *Laughter: A scientific investigation*. New York: Viking.

Pryor, C., Perfors, A., & Howe, P. D. L. (2019). Even arbitrary norms influence moral decision-making. *Nature Human Behaviour*, 3, 57–62.

Putnam, A. L., Ross, M. Q., Soter, L. K., & Roediger, H. L. (2018). Collective narcissism: Americans exaggerate the role of their home state in appraising U.S. history. *Psychological Science*, 29, 1414–1422.

Qiu, C., Luu, L., & Stocker, A. A. (2020). Benefits of commitment in hierarchical inference. *Psychological Review*, 127, 622–639.

Rachlin, H., & Jones, B. A. (2008). Altruism among relatives and non relatives. *Behavioural Processes*, 79, 120–123.

Rao, A. R., & Monroe, K. B. (1989). The effect of price, brand name, and store name on buyer's perceptions of product quality. *Journal of Marketing Research*, 26, 351–357.

Raue, M., & Scholl, S. G. (2018). The use of heuristics in decision-making under risk and uncertainty. In M. Raue, E. Lerner, and B. Streicher (eds.), *Psychological perspectives on risk and risk analysis: Theory, Models and Applications* (pp. 153–179). New York, NY: Springer.

Razran, G. H. S. (1938). Conditioning away social bias by the luncheon technique. *Psychological Bulletin*, 35, 693.

Razran, G. H. S. (1940). Conditional response changes in rating and appraising sociopolitical slogans. *Psychological Bulletin*, 37, 481.

Regan, D. T., & Kilduff, M. (1988). Optimism about elections: Dissonance reduction at the ballot box. *Political Psychology*, 9, 101–107.

Regan, R. T. (1971). Effects of a favor and liking on compliance. *Journal of Experimental Social Psychology*, 7, 627–639.

Reich, T., Kupor, D. M., & Smith, R. K. (2018). Made by mistake: When

mistakes increase product preference. *Journal of Consumer Research*, 44, 1085–1103.

Reich, T., & Maglio, S. J. (2020). Featuring mistakes: The persuasive impact of purchase mistakes in online reviews. *Journal of Marketing*, 84, 52–65.

Reilly, K. (2017, October 23). A deadly campus tradition. *Time*, 57–61.

Reis, H. T., Maniaci, M. R., Caprariello, P. A., Eastwick, P. W., & Finkel, E. J. (2011). Familiarity does promote attraction in live interaction. *Journal of Personality and Social Psychology*, 101, 557–570.

Reiterman, T. (2008). *Raven: The untold story of the Rev. Jim Jones and his people*. New York: Tarcher Perigee.

Rentfrow, P. J. (2010). Statewide differences in personality: Toward a psychological geography of the United States. *American Psychologist*, 65, 548–558.

Rice, B. (April 24, 2000). How plaintiff's lawyers pick their targets. *Medical Economics*, 77, 94–110.

Richeson, J. A., & Shelton, J. N. (2007). Negotiating interracial interactions. *Current Directions in Psychological Science*, 16, 316–320.

Riek, B. M., Mania, E. W., & Gaertner, S. L. (2006). Intergroup threat and outgroup attitudes: A meta-analytic review. *Personality and Social Psychology Review*, 10, 336–353.

Riek, B. M., Mania, E. W., Gaertner, S. L., McDonald, S. A., & Lamoreaux, M. J. (2010). Does a common in-group identity reduce intergroup threat? *Group Processes & Intergroup Relations*, 13, 403–423.

Riley, D., & Eckenrode, J. (1986). Social ties: Subgroup differences in costs and benefits. *Journal of Personality and Social Psychology*, 51, 770–778.

- Ritts, V., Patterson, M. L., & Tubbs, M. E. (1992). Expectations, impressions, and judgments of physically attractive students: A review. *Review of Educational Research*, 62, 413–426.
- Rochat, F., & Blass, T. (2014). Milgram's unpublished obedience variation and its historical relevance. *Journal of Social Issues*, 70, 456–472.
- Rodafinos, A., Vucevic, A., & Sideridis, G. D. (2005). The effectiveness of compliance techniques: Foot-in-the-door versus door-in-the-face. *Journal of Social Psychology*, 145, 237–240.
- Roese, N. J., & Olson, M. J. (1994). Attitude importance as a function of repeated attitude expression. *Journal of Experimental Social Psychology*, 30, 39–51.
- Rollins, T. (2020). *The CEO formula*. McLean, VA: Rollins.
- Romero, T., Ito, M., Saito, A., & Hasegawa, T. (2014). Social modulation of contagious yawning in wolves. *PLoS ONE*, 9.
- Romero, T., Konno, A., & Hasegawa, T. (2013). Familiarity bias and physiological responses in contagious yawning by dogs support link to empathy. *PLoS ONE*, 8.
- Rosen, S., & Tesser, A. (1970). On the reluctance to communicate undesirable information: The MUM effect. *Sociometry*, 33, 253–263.
- Rosenthal, A. M. (1964). *Thirty-eight witnesses*. New York: McGraw-Hill.
- Roseth, C. J., Johnson, D. W., & Johnson, R. T. (2008). Promoting early adolescents' achievement and peer relationships: The effects of cooperative, competitive, and individualistic goal structures. *Psychological Bulletin*, 134, 223–246.
- Ross, J. R. (1994). *Escape to Shanghai: A Jewish community in China*. New

York: Free Press.

Rothman, A. J., Martino, S. C., Bedell, B. T., Detweiler, J. B., & Salovey, P. (1999). The systematic influence of gain- and loss-framed messages on interest in and use of different types of health behavior. *Personality and Social Psychology Bulletin*, 25, 1355–1369.

Rothman, A. J., & Salovey, P. (1997). Shaping perceptions to motivate healthy behavior: The role of message framing. *Psychological Bulletin*, 121, 3–19.

Rubinstein, S. (1985, January 30). What they teach used car salesmen. *San Francisco Chronicle*.

Rusbult, C. E., Van Lange, P. A. M., Wildschut, T., Yovetich, N. A., & Verette, J. (2000). Perceived superiority in close relationships: Why it exists and persists. *Journal of Personality and Social Psychology*, 79, 521–545.

Sabin, R. (1964). *The international cyclopedia of music and musicians*. New York: Dodd, Mead.

Sacarny, A., Barnett, M. L., Le, J., Tetkoski, F., Yokum, D., & Agrawal, S. (2018). Effect of peer comparison letters for high-volume primary care prescribers of quetiapine in older and disabled adults: A randomized clinical trial. *Journal of the American Medical Association Psychiatry*, 75, 1003–1011.

Sagarin, B. J., Cialdini, R. B., Rice, W. E., & Serna, S. B. (2002). Dispelling the illusion of invulnerability: The motivations and mechanisms of resistance to persuasion. *Journal of Personality and Social Psychology*, 83, 526–541.

Sagarin, B. J., & Mitnick, K. D. (2012). The path of least resistance. In D. T. Kenrick, N. J. Goldstein, and S. L. Braver (eds.), *Six degrees of social influence: Science, application, and the psychology of Robert Cialdini* (pp.

27–38). New York: Oxford University Press.

Salant, J. D. (2003, July 20). Study links donations, vote patterns. *Arizona Republic*, A5.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311, February 10, 854–856.

Santos, H. C., Varnum, M. E. W., & Grossmann, I. (2017). Global increases in individualism. *Psychological Science*, 28, 1228–1239.

Sasaki, S. J., & Vorauer, J. D. (2013). Ignoring versus exploring differences between groups: Effects of salient color - blindness and multiculturalism on intergroup attitudes and behavior. *Social and Personality Psychology Compass*, 7, 246–259.

Sassenrath, C., Hodges, S. D., & Pfattheicher, S. (2016). It's all about the self: When perspective taking backfires. *Current Directions in Psychological Science*, 25, 405–410.

Savage, P., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S., & Fitch, W. (2020). Music as a coevolved system for social bonding. *Behavioral and Brain Sciences*, 1–36.

Schein, E. (1956). The Chinese indoctrination program for prisoners of war: A study of attempted “brainwashing.” *Psychiatry*, 19, 149–172.

Schindler, R. M. (1998). Consequences of perceiving oneself as responsible for obtaining a discount. *Journal of Consumer Psychology*, 7(4), 371–392.

Schkade, D. A., & Kahneman, D. (1998). Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science*, 9, 340–346.

Schlenker, B. R., Dlugolecki, D. W., & Doherty, K. (1994). The impact of self-presentations on self-appraisals and behavior. The power of public commitment. *Personality and Social Psychology Bulletin*, 20, 20–33.

Schmidtke, A., & Hafner, H. (1988). The Werther effect after television films: New evidence for an old hypothesis. *Psychological Medicine*, 18, 665–676.

Schmitt, M. T., Mackay, C. M. L., Droogendyk, L. M., & Payne, D. (2019). What predicts environmental activism? The roles of identification with nature and politicized environmental identity. *Journal of Environmental Psychology*, 61, 20–29.

Schrange, M. (2004, September). The opposite of perfect. *Sales and Marketing Management*, 26.

Schrift, R. Y., & Parker, J. R. (2014). Staying the course: The option of doing nothing and its impact on postchoice persistence. *Psychological Science*, 25, 772–780.

Schroeder, J., Risen, J. L., Gino, F., & Norton, M. I. (2019). Handshaking promotes deal-making by signaling cooperative intent. *Journal of Personality and Social Psychology*, 116, 743–768.

Schultz, P. W. (1999). Changing behavior with normative feedback interventions: A field experiment on curbside recycling. *Basic and Applied Social Psychology*, 21, 25–36.

Schumpe, B. M., Bélanger, J. J., & Nisa, C. F. (2020). The reactance decoy effect: How including an appeal before a target message increases persuasion. *Journal of Personality and Social Psychology*, 119, 272–292.

Schwarz, N. (1984). When reactance effects persist despite restoration of freedom: Investigations of time delay and vicarious control. *European*

Journal of Social Psychology, 14, 405–419.

Schwarzwald, D., Raz, M., & Zwibel, M. (1979). The applicability of the door-in- the-face technique when established behavior customs exit. *Journal of Applied Social Psychology*, 9, 576–586.

Sechrist, G. B., & Stangor, C. (2007). When are intergroup attitudes based on perceived consensus information? The role of group familiarity. *Social Influence*, 2, 211–235.

Segal, H. A. (1954). Initial psychiatric findings of recently repatriated prisoners of war. *American Journal of Psychiatry*, III, 358–363.

Seiter, J. S. (2007). Ingratiation and gratuity: The effect of complimenting customers on tipping behavior in restaurants. *Journal of Applied Social Psychology*, 37, 478–485.

Seiter, J. S., & Dutson, E. (2007). The effect of compliments on tipping behavior in hairstyling salons. *Journal of Applied Social Psychology*, 37, 1999–2007.

Sengupta, J., & Johar, G. V. (2001). Contingent effects of anxiety on message elaboration and persuasion. *Personality and Social Psychology Bulletin*, 27, 139–150.

Shadel, D. (2012). *Outsmarting the scam artists: How to protect yourself from the most clever cons*. Hoboken, NJ: Wiley & Sons.

Shaffer, D., Garland, A., Vieland. V., Underwood, M., & Busner, C. (1991). The impact of curriculum-based suicide prevention programs for teenagers. *Journal of the American Academy of Child and Adolescent Psychiatry*, 30, 588–596.

Shah, A. J., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort

reduction framework. *Psychological Bulletin*, 134, 207–222.

Shah, A. M., Eisenkraft, N., Bettman, J. R., & Chartrand, T. L. (2015). “Paper or plastic?”: How we pay influences post-transaction connection. *Journal of Consumer Research*, 42, 688–708.

Sharot, T., Fleming, S. M., Yu, X., Koster, R., & Dolan, R. J. (2012). Is choice-induced preference change long lasting? *Psychological Science*, 23, 1123–1129.

Sharot, T., Velasquez, C. M., & Dolan, R. J. (2010). Do decisions shape preference? Evidence from blind choice. *Psychological Science*, 21, 1231–1235.

Sharps, M., & Robinson, E. (2017). Perceived eating norms and children’s eating behavior: An Informational social Influence account. *Appetite*, 113, 41–50.

Shayo, M. (2020). Social identity and economic policy. *Annual Review of Economics*, 12, 355–389.

Shayo, M., & Zussman, A. (2011). Judicial in-group bias in the shadow of terrorism. *Quarterly Journal of Economics*, 126, 1447–1484.

Shelley, M. K. (1994). Individual differences in lottery evaluation models. *Organizational Behavior and Human Decision Processes*, 60, 206–230.

Sheng, F., Ramakrishnan, A., Seok, D., Zhao, W. J., Thelaus, S., Cen, P., & Platt, M. L. (2020). Decomposing loss aversion from gaze allocation and pupil dilation. *Proceedings of the National Academy of Sciences*, 117, 11356–11363.

Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1961). *Intergroup conflict and cooperation: The Robbers’ Cave experiment*.

Norman, OK: University of Oklahoma Institute of Intergroup Relations.

Sherman, D. K., Brookfeld, J., & Ortosky, L. (2017). Intergroup conflict and barriers to common ground: A self-affirmation perspective. *Social and Personality Psychology Compass*, 11.

Sherman, L. E., Payton, A. A., Hernandez, L. M., Greenfield, P. M., & Dapretto, M.(2016). The power of the like in adolescence: Effects of peer influence on neural and behavioral responses to social media. *Psychological Science*, 27, 1027–1035.

Sherman, S. J. (1980). On the self-erasing nature of errors of prediction. *Journal of Personality and Social Psychology*, 39, 211–221.

Shi, L., Romić, I., Ma, Y., Wang, Z., Podobnik, B., Stanley, H. E., Holme, P., & Jusup, M.(2020). Freedom of choice adds value to public goods. *Proceedings of the National Academy of Sciences*, 117, 17516–17521.

Shiv, B., Carmon, Z., & Ariely, D. (2005). Placebo effects of marketing actions: Consumers may get what they pay for. *Journal of Marketing Research*, 42, 383– 393.

Shnabel, N., Halabi, S., & Noor, M. (2013). Overcoming competitive victimhood and facilitating forgiveness through re-categorization into a common victim or perpetrator identity, *Journal of Experimental Social Psychology*, 49, 867–877.

Shnabel, N., Purdie-Vaughns, V., Cook, J. E., Garcia J., & Cohen G. L. (2013). Demystifying values-affirmation interventions: Writing about social belonging is a key to buffering against identity threat. *Personality and Social Psychology Bulletin*. 39, 663–676.

Shook, N. J., & Fazio, R. H. (2008). Interracial roommate relationships: An experimental field test of the contact hypothesis. *Psychological Science*, 19,

717– 723.

Shotland, R. I., & Straw, M. (1976). Bystander response to an assault: When a man attacks a woman. *Journal of Personality and Social Psychology*, 34, 990–999.

Shrestha, K. (2018). 50 important stats you need to know about online reviews.

Shrout, M. R., Brown, R. D., Orbuch, T. L., & Weigel, D. J. (2019). A multidimensional examination of marital conflict and health over 16 years. *Personal Relationships*, 26, 490–506.

Shteynberg, G. (2015). Shared attention. *Perspectives on Psychological Science*, 10, 579–590.

Shtulman, A. (2006). Qualitative differences between naïve and scientific theories of evolution. *Cognitive Psychology*, 52, 170–194.

Shu, S. B., & Carlson, K. A. (2014). When three charms but four alarms: Identifying the optimal number of claims in persuasion settings. *Journal of Marketing*, 78, 127–139.

Siegal, A. (2018). *Transcendental deception: Behind the TM curtain—bogus science, hidden agendas, and David Lynch’s campaign to push a million public school kids into Transcendental Meditation while falsely claiming it is not a religion*. Los Angeles, CA: Janreg.

Silver, A. M., Stahl, A. E., Loiotile, R., Smith-Flores, A. S., & Feigenson, L. (2020). When not choosing leads to not liking: Choice-induced preference in infancy. *Psychological Science*.

Sinaceur, M., & Heath, C., & Cole, S. (2005). Emotional and deliberative reaction to a public crisis: Mad cow disease in France. *Psychological Science*,

16, 247–254.

Skinner, A. L., Olson, K. R., & Meltzoff, A. N. (2020). Acquiring group bias: Observing other people's nonverbal signals can create social group biases. *Journal of Personality and Social Psychology*, 119, 824–838.

Slavin, R. E. (1983). When does cooperative learning increase student achievement? *Psychological Bulletin*, 94, 429–445.

Smith, C. T., De Houwer, J., & Nosek, B. A. (2013). Consider the source: Persuasion of implicit evaluations is moderated by source credibility. *Personality and Social Psychology Bulletin*, 39, 193–205.

Smith, D. L. (2020). *On inhumanity: Dehumanization and how to resist it*. Oxford: Oxford University Press.

Smith, G. H., & Engel, R. (1968). Influence of a female model on perceived characteristics of an automobile. *Proceedings of the 76th Annual Convention of the American Psychological Association*, 3, 681–682.

Smith, R. W., Chandler, J. J., & Schwarz, N. (2020). Uniformity: The effects of organizational attire on judgments and attributions. *Journal of Applied Social Psychology*, 50, 299–312.

Sokol-Hessner, P., & Rutledge, R. B. (2019). The psychological and neural basis of loss aversion. *Current Directions in Psychological Science*, 28, 20–27.

Sorokowski, P. (2010). Politicians' estimated height as an indicator of their popularity. *European Journal of Social Psychology*, 40, 1302–1309.

Southgate, V. (2020). Are infants altercentric? The other and the self in early social cognition. *Psychological Review*, 127, 505–523.

Spangenberg, E. R., & Greenwald, A. G. (2001). Self-prophecy as a method

for increasing participation in socially desirable behaviors. In W. Wosinska, R. B. Cialdini, D. W. Barrett, and J. Reykowski (eds.), *The practice of social influence in multiple cultures* (pp. 51–62). Mahwah, NJ: Lawrence Erlbaum.

Sparkman, G., & Walton, G. M. (2017). Dynamic norms promote sustainable behavior, even if it is counternormative. *Psychological Science*, 28, 1663–1674.

Sparkman, G., & Walton, G. M. (2019). Witnessing change: Dynamic norms help resolve diverse barriers to personal change. *Journal of Experimental Social Psychology*, 82, 238–252.

Sprecher, S., Treger, S., Wondra, J. D., Hilaire, N., & Wallpe, K. (2013). Taking turns: Reciprocal self-disclosure promotes liking in initial interactions. *Journal of Experimental Social Psychology*, 49, 860–866.

Sprott, D. E., Spangenberg, E. R., Knuff, D. C., & Devezar, B. (2006). Selfprediction and patient health: Influencing health-related behaviors through self-prophecy. *Medical Science Monitor*, 12, RA85–91.

Staats, B. R., Dai, H., Hofmann, D., & Milkman, K. L. (2017). Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare. *Management Science*, 63, 1563–1585.

Stallen, M., Smidts, A., & Sanfey, A. G. (2013). Peer influence: neural mechanisms underlying in-group conformity. *Frontiers in Human Neuroscience*, 7.

Stanchi, K. M. (2008). Playing with fire: The science of confronting adverse material in legal advocacy. *Rutgers Law Review*, 60, 381–434.

Stanne, M. B., Johnson, D. W., & Johnson, R. T. (1999). Does competition enhance or inhibit motor performance: A meta-analysis. *Psychological*

Bulletin, 125, 133–154.

Stehr, N., & Grundmann, R. (2011). *Experts: The knowledge and power of expertise*. London: Routledge.

Stelfox, H. T., Chua, G., O'Rourke, K., & Detsky, A. S. (1998). Conflict of interest in the debate over calcium-channel antagonists. *New England Journal of Medicine*, 333, 101–106.

Stephan, W. G. (1978). School desegregation: An evaluation of predictions made in *Brown vs. Board of Education*. *Psychological Bulletin*, 85, 217–238.

Stern, S. M. (2012). *The Cuban Missile Crisis in American memory: Myths versus reality*. Palo Alto, CA: Stanford University Press.

Stephens, N. M., Fryberg, S. A., Markus, H. R., Johnson, C., & Covarrubias, R. (2012). Unseen disadvantage: How American universities' focus on independence undermines the academic performance of first-generation college students. *Journal of Personality and Social Psychology*, 102, 1178–1197.

Stevens, M. (2016). *Cheats and deceptions: How animals and plants exploit and mislead*. New York: Oxford University Press.

Stewart, P. A., Eubanks, A. D., Dye, R. G., Gong, Z. H., Bucy, E. P., Wicks, R. H., & Eidelman, S. (2018). Candidate performance and observable audience response: Laughter and applause-cheering during the first 2016 Clinton-Trump presidential debate. *Frontiers in Psychology*, 9, 1182.

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science*, 21, 349–354.

Strenta, A., & DeJong, W. (1981). The effect of a prosocial label on helping behavior. *Social Psychology Quarterly*, 44, 142–147.

Strohmetz, D. B., Rind, B., Fisher, R., & Lynn, M. (2002). Sweetening the till— the use of candy to increase restaurant tipping. *Journal of Applied Social Psychology*, 32, 300–309.

Stok, F. M., de Ridder, D. T., de Vet, E., & de Wit, J. F. (2014). Don't tell me what I should do, but what others do: The influence of descriptive and injunctive peer norms on fruit consumption in adolescents. *British Journal of Health Psychology* 19, 52–64.

Stone, J., & Focella, E. (2011). Hypocrisy, dissonance and the self-regulation processes that improve health, *Self and Identity*, 10, 295–303.

Stone, J., Whitehead, J., Schmader, T., & Focella, E. (2011). Thanks for asking: Self-affirming questions reduce backlash when stigmatized targets confront prejudice. *Journal of Experimental Social Psychology*, 47, 589–598.

Strauss, M. (2011). *Pictures, passions, and eye*. London: Halban.

Styron, W. (1977). A farewell to arms. *New York Review of Books*, 24, 3–4.

Suedfeld, P., Bochner, S., & Matas, C. (1971). Petitioner's attire and petition signing by peace demonstrators: A field experiment. *Journal of Applied Social Psychology*, 1, 278–283.

Sumner, S. A., Burke, M., & Kooti, F. (2020). Adherence to suicide reporting guidelines by news shared on a social networking platform. *Proceedings of the National Academy of Sciences*, 117, 16267–16272.

Surowiecki, J. (2004). *The wisdom of crowds*. New York: Doubleday.

Sutcliffe, K. (2019, November 25). How to reduce medical errors. *Time*, 25–26.

Swaab, R. I., Maddux, W. W., & Sinaceur, M. (2011). Early words that work:

When and how virtual linguistic mimicry facilitates negotiation outcomes. *Journal of Experimental Social Psychology*, 47, 616–621.

Swann, W. B., & Buhrmester, M. D. (2015). Identity fusion. *Current Directions in Psychological Science*, 24, 52–57.

Swart, H., Hewstone, M., Christ, O., & Voci, A. (2011). Affective mediators of intergroup contact: A three-wave longitudinal study in South Africa. *Journal of Personality and Social Psychology*, 101, 1221–1238.

Sweis, B. M., Abram, S. V., Schmidt, B. J., Seeland, K. D., MacDonald III, A. W., Thomas, M. J., & Redish, D. (2018). Sensitivity to “sunk costs” in mice, rats, and humans. *Science*, 361, 178–181.

Sweldens, S., van Osselaer, S. M. J., & Janiszewski, C. (2010). Evaluative conditioning procedures and resilience of conditioned brand attitudes. *Journal of Consumer Research* 37, 473–489.

Szabo, L. (2007, February 5). Patient protect thyself. *USA Today*, 8D.

Sznycer, D., De Smet, D., Billingsley, J., & Lieberman, D. (2016). Coresidence duration and cues of maternal investment regulate sibling altruism across cultures. *Journal of Personality and Social Psychology*, 111, 159–177.

Tadlock, B. L., Flores, A. R., Haider-Markel, D. P., Lewis, D.C., Miller, P. R., & Taylor, J.K. (2017). Testing contact theory and attitudes on transgender rights. *Public Opinion Quarterly*, 81, 956–972.

Tal-Or, N. (2008). Boasting, burnishing, and burying in the eyes of the perceivers. *Social Influence*, 3, 202–222.

Tan, Q., Zhan, Y., Gao, S., Chen, J., & Zhong, Y. (2015). Closer the relatives are, more intimate and similar we are: Kinship effects on self-other overlap.

Personality and Individual Differences, 73, 7–11.

Tarr, B., Launay, J., & Dunbar, R. I. (2014). Music and social bonding: “Selfother”merging and neurohormonal mechanisms. *Frontiers in psychology*, 5.

Taylor, R. (1978). Marilyn’s friends and Rita’s customers: A study of party selling as play and as work. *Sociological Review*, 26, 573–611.

Tedeschi, J. T., Schlenker, B. R., & Bonoma, T. V. (1971). Cognitive dissonance: Private ratiocination or public spectacle? *American Psychologist*, 26, 685–695.

Teger, A. I. (1980). *Too much invested to quit*. Elmsford, NY: Pergamon.

Telzer, E. H., Masten, C. L., Berkman, E. T., Lieberman, M. D., & Fuligni, A. J. (2010). Gaining while giving: An fMRI study of the rewards of family assistance among White and Latino youth. *Social Neuroscience*, 5, 508–518.

Tesser, A. (1993). The importance of heritability in psychological research: The case of attitudes. *Psychological Review*, 100, 129–142.

Teuscher, U. (2005, May). The effects of time limits and approaching endings on emotional intensity. Paper presented at the meetings of the American Psychological Society, Los Angeles, CA.

Thaler, R. H., Tversky, A., Kahneman, D., & Schwartz, A. (1997). The effect of myopia and loss aversion on risk taking: An experimental test. *The Quarterly Journal of Economics*, 112, 647–661.

Thompson, D. (2017). *Hit makers: The science of popularity in an age of distraction*. New York: Penguin.

Thompson, L. (1990). An examination of naive and experienced negotiators. *Journal of Personality and Social Psychology*, 59, 82–90.

Thompson, L., & Hrebec, D. (1996). Lose-lose agreements in interdependent decision making. *Psychological Bulletin*, 120, 396–409.

Tiger, L., & Fox, R. (1989). *The imperial animal*. New York: Holt.

Till, B. D., & Priluck, R. L. (2000). Stimulus generalization in classical conditioning: An initial investigation and extension. *Psychology & Marketing*, 17, 55–72.

Todd, A. R., & Galinsky, A. D. (2014). Perspective-taking as a strategy for improving intergroup relations: Evidence, mechanisms, and qualifications. *Social and Personality Psychology Compass*, 8, 374–387.

Todd, P. M., & Gigerenzer, G. (2007). Environments that make us smart. *Current Directions in Psychological Science*, 16, 167–171.

Tokayer, M., & Swartz, M. (1979). *The Fugu Plan: The untold story of the Japanese and the Jews during World War II*. New York: Paddington.

Tomasello, M. (2020). The moral psychology of obligation. *Behavioral and Brain Sciences*, 43, E56.

Tormala, Z. L., & Petty, R. E. (2007). Contextual contrast and perceived knowledge: Exploring the implications for persuasion. *Journal of Experimental Social Psychology*, 43, 17–30.

Toufexis, A. (1993, June 28). A weird case, baby? Uh huh! *Time*, 41.

Towers, S., Gomez-Lievano, A., Khan M., Mubayi, A., & Castillo-Chavez, C. (2015). Contagion in mass killings and school shootings. *PLoS ONE*, 10.

Trocmé, A. (2007/1971). *Jesus and the nonviolent revolution*. Farmington, PA: Plough Publishing House.

Turner, R. N., Hewstone, M., Voci, A., Paolini, S., & Christ, O. (2007). Reducing prejudice via direct and extended cross-group friendship. *European*

Review of Social Psychology, 18, 212–255.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.

Unkelbach, C., Koch, A., Silva, R. R., & Garcia-Marques, T. (2019). Truth by repetition: Explanations and implications. *Current Directions in Psychological Science*, 28, 247–253.

Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion*, 11, 262–266.

van Baaren, R. B., Holland, R. W., Steenaert, B., & van Knippenberg, A. (2003). Mimicry for money: Behavioral consequences of imitation. *Journal of Experimental Social Psychology*, 39, 393–398.

Vandello, J. A., & Cohen D. (1999). Patterns of individualism and collectivism across the United States. *Journal of Personality and Social Psychology*, 77, 279–292.

van den Berg, H., Manstead, A. S. R., van der Pligt, J., & Wigboldus, D. H. J. (2006). The impact of affective and cognitive focus on attitude formation. *Journal of Experimental Social Psychology*, 42, 373–379.

Van der Werff, E., Steg, L., & Keizer, K. (2014). I am what I am, by looking past the present: The influence of biospheric values and past behavior on environmental self-identity. *Environment and Behavior*, 46, 626–657.

van Herpen, E., Pieters, R., & Zeelenberg, M. (2009). When demand accelerates demand: Trailing the bandwagon. *Journal of Consumer Psychology*, 19, 302–312.

Van Overwalle, F., & Heylighen, F. (2006). Talking nets: A multiagent connectionist approach to communication and trust between individuals.

Psychological Review, 113, 606–627.

Verosky, S. C., & Todorov, A. (2010). Generalization of affective learning about faces to perceptually similar faces. *Psychological Science*, 21, 779–785.

Vonk, R. (2002). Self-serving interpretations of flattery: Why ingratiation works. *Journal of Personality and Social Psychology*, 82, 515–526.

von Zimmermann, J., & Richardson, D. C. (2016). Verbal synchrony and action dynamics in large groups. *Frontiers of Psychology*, 7.

Vorauer, J. D., Martens, V., & Sasaki, S. J. (2009). When trying to understand detracts from trying to behave: Effects of perspective taking in intergroup interaction. *Journal of Personality and Social Psychology*, 96, 811–827.

Vorauer, J. D., & Sasaki, S. J. (2011). In the worst rather than the best of times: Effects of salient intergroup ideology in threatening intergroup interactions. *Journal of Personality and Social Psychology*, 101(2), 307–320.

Waber, R. L., Shiv, B., Carmon, Z., & Ariely, D. (2008). Commercial features of placebo and therapeutic efficacy. *Journal of the American Medical Association*, 299, 1016–1017.

Wagner, T., Lutz, R. J., & Weitz, B. A. (2009). Corporate hypocrisy: Overcoming the threat of inconsistent corporate social responsibility perceptions. *Journal of Marketing*, 73, 77–91.

Walker, J., Risen, J. L., Gilovich, T., & Thaler, R. (2018). Sudden-death aversion: Avoiding superior options because they feel riskier. *Journal of Personality and Social Psychology*, 115, 363–378.

Wall, L. L., & Brown, D. (2007). The high cost of free lunch. *Obstetrics &*

Gynecology, 110, 169–173.

Wan, L. C., & Wyer, R. S. (2019). The influence of incidental similarity on observers' causal attributions and reactions to a service failure. *Journal of Consumer Research*, 45, 1350–1368.

Ward, A., & Brenner, L. (2006). Accentuate the negative. The positive effects of negative acknowledgment. *Psychological Science*, 17, 959–965.

Warneken, F., Lohse, K., Melis, P. A., & Tomasello, M. (2011). Young children share the spoils after collaboration. *Psychological Science*, 22, 267–273.

Warren, W. H. (2018). Collective motion in human crowds. *Current Directions in Psychological Science*, 27, 232–240.

Warrick, J. (2008, December 26). Afghan influence taxes CIA's credibility. *Washington Post*, A17.

Watanabe, T. (1994, March 20). An unsung “Schindler” from Japan. *Los Angeles Times*, 1.

Watson, T. J., Jr. (1990). *Father, son, & co*. New York: Bantam.

Waytz, A., Dungan, J., & Young, L. (2013). The whistleblower's dilemma and the fairness-loyalty tradeoff. *Journal of Experimental Social Psychology*, 49, 1027–1033.

Wears, R., & Sutcliffe, K. (2020). *Still not safe: Patient safety and the middlemanagement of American medicine*. New York: Oxford University Press.

Wedekind, C., & Milinski, M. (2000). Cooperation through image scoring in humans. *Science*, 288, 850–852.

Weidman, A. C., Sowden, W. J., Berg, M., & Kross, E. (2020).

Punish or protect? How close relationships shape responses to moral violations. *Personality and Social Psychology Bulletin*, 46, 693–708.

Weinstein, E. A., & Deutschberger, P. (1963). Some dimensions of altercasting. *Sociometry*, 26, 454–466.

Weisbuch, M., Ambady, N., Clarke, A. L., Achor, S., & Veenstra-Vander Weele, S. (2010). On being consistent: The role of verbal–nonverbal consistency in first impressions. *Basic and Applied Social Psychology*, 32, 261–268.

Weller, J. A., Levin, I. P., Shiv, B., & Bechara, A. (2007). Neural correlates of adaptive decision making for risky gains and losses. *Psychological Science*, 18, 958–964.

Wells, P. A. (1987). Kin recognition in humans. In D. J. C. Fletcher and C. D. Michener (eds.), *Kin recognition in animals* (pp. 395–416). New York: Wiley.

West, S. G. (1975). Increasing the attractiveness of college cafeteria food: A reactance theory perspective. *Journal of Applied Psychology*, 60, 656–658.

Westmaas, J. L., & Silver, R. C. (2006). The role of perceived similarity in supportive responses to victims of negative life events. *Personality and Social Psychology Bulletin*, 32, 1537–1546.

Wheatley, T., Kang, O., Parkinson, C., & Looser, C. E. (2012). From mind perception to mental connection: Synchrony as a mechanism for social understanding. *Social and Personality Psychology Compass*, 6, 589–606.

White, M. (1997, July 12). Toy rover sales soar into orbit. *Arizona Republic*, E1, E9. Whitehouse, H., Jong, J., Buhrmester, M. D., Gómez, Á., Bastian, B., Kavanagh, C. M., Newson, M., et al. (2017). The evolution of extreme cooperation via shared dysphoric experiences. *Scientific Reports*, 7, 44292.

Whiting, J. W. M., Kluckhohn, R., & Anthony A. (1958). The function of male initiation ceremonies at puberty. In E. E. Maccoby, T. M. Newcomb, and E. L. Hartley (eds.), *Readings in social psychology* (pp. 82–98). New York: Henry Holt.

Wicklund, R. A., & Brehm, J. C. (1974) cited in R. A. Wicklund, *Freedom and reactance*. Hillsdale, NJ: Lawrence Erlbaum.

Williams, K. D., Bourgeois, M. J., & Croyle, R. T. (1993). The effects of stealing thunder in criminal and civil trials. *Law and Human Behavior*, 17, 597–609.

Wilson, P. R. (1968). The perceptual distortion of height as a function of ascribed academic status. *Journal of Social Psychology*, 74, 97–102.

Wilson, T. D., Dunn, D. S., Kraft, D., & Lisle, D. J. (1989). Introspection, attitude change, and behavior consistency. In L. Berkowitz (ed.), *Advances in experimental social psychology* (vol. 22, pp. 287–343). San Diego, CA: Academic Press.

Wilson, T. D., & Gilbert, D. T. (2008). Affective forecasting: Knowing what to want. *Current Directions in Psychological Science*, 14, 131–134.

Wilson, T. D., & Linville, P. D. (1985). Improving the performance of college freshmen with attributional techniques. *Journal of Personality and Social Psychology*, 49, 287–293.

Wilson, T. D., Reinhard, D. A., Westgate, E. C., Gilbert, D. T., Ellerbeck, N., Hahn, C., Brown, C. L., & Shaked, A. (2014). Just think: The challenges of the disengaged mind. *Science*, 345, 75–77.

Wilson, T. D., Wheatley, T. P., Meyers, J. M., Gilbert, D. T., & Axson, D. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, 78, 821–836.

Wiltermuth, S. S. (2012a). Synchronous activity boosts compliance with requests to aggress. *Journal of Experimental Social Psychology*, 48, 453–456.

Wiltermuth, S. S. (2012b). Synchrony and destructive obedience. *Social Influence*, 7, 78–89.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, 20, 1–5.

Winkielman, P., Berridge, K. C., & Wilbarger, J. L. (2005). Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Personality and Social Psychology Bulletin*, 31, 121–135.

Wölfer, R., Christ, O., Schmid, K., Tausch, N., Buchallik, F. M., Vertovec, S., & Hewstone, M. (2019). Indirect contact predicts direct contact: Longitudinal evidence and the mediating role of intergroup anxiety. *Journal of Personality and Social Psychology*, 116, 277–295.

Wolske, K. S., Gillingham, K. T., & Schultz, P. W. (2020). Peer influence on household energy behaviours. *Nature Energy* 5, 202–212.

Woolley, K., & Risen, J. L. (2018). Closing your eyes to follow your heart: Avoiding information to protect a strong intuitive preference. *Journal of Personality and Social Psychology*, 114, 230–245.

Wooten, D. B., & Reed, A. (1998). Informational influence and the ambiguity of product experience: Order effects on the weighting of evidence. *Journal of Consumer Research*, 7, 79–99.

Worchel, S. (1992). Beyond a commodity theory analysis of censorship: When abundance and personalism enhance scarcity effects. *Basic and Applied Social Psychology*, 13, 79–92.

Worchel, S., & Arnold, S. E. (1973). The effects of censorship and the attractiveness of the censor on attitude change. *Journal of Experimental Social Psychology*, 9, 365–377.

Worchel, S., Arnold, S. E., & Baker, M. (1975). The effect of censorship on attitude change: The influence of censor and communicator characteristics. *Journal of Applied Social Psychology*, 5, 222–239.

Worchel, S., Lee, J., & Adewole, A. (1975). Effects of supply and demand on ratings of object value. *Journal of Personality and Social Psychology*, 32, 906–914.

Wright, S. C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology*, 73, 73–90.

Xu, L., Zhang, X., & Ling, M. (2018). Spillover effects of household waste separation policy on electricity consumption: Evidence from Hangzhou, China. *Resources, Conservation, and Recycling*, 129, 219–231.

Yang, F., Choi, Y-U., Misch, A., Yang, X., & Dunham, Y. (2018). In defense of the commons: Young children negatively evaluate and sanction free riders. *Psychological Science*, 29, 1598–1611.

Yeh, J. S., Franklin, J. M., Avorn, J., Landon, J., & Kesselheim, A. S. (2016). Association of industry payments with the prescribing brand-name statins in Massachusetts. *Journal of the American Medical Association: Internal Medicine*, 176, 763–768.

Yu, S., & Sussman, S. (2020). Does smartphone addiction fall on a continuum of addictive behaviors? *International Journal of Environmental Research and Public Health*, 17, art. no. 422.

Yuki, M., Maddox, W. M., Brewer, M. B., & Takemura, K. (2005). Cross-

cultural differences in relationship- and group-based trust. *Personality and Social Psychology Bulletin*, 31, 48–62.

Zellinger, D. A., Fromkin, H. L., Speller, D. E., & Kohn, C. A. (1974). A commodity theory analysis of the effects of age restrictions on pornographic materials. (Paper no. 440). Lafayette, IN: Purdue University, Institute for Research in the Behavioral, Economic, and Management Sciences.

Żemła M., & Gladka, A., (2016). Effectiveness of reciprocal rule in tourism: Evidence from a city tourist restaurant. *European Journal of Service Management*, 17, 57–63.

Zhang, Y., Xu, J., Jiang, Z., & Huang, S-C. (2011). Been there, done that: The impact of effort investment on goal value and consumer motivation. *Journal of Consumer Research*, 38, 78–93.

Zhao, X., & Epley, N. (2020). Kind words do not become tired words: Undervaluing the positive impact of frequent compliments. *Self and Identity*.

Zitek, E. M., & Hebl, M. R. (2007). The role of social norm clarity in the influenced expression of prejudice over time. *Journal of Experimental Social Psychology*, 43, 867–876.

Zuckerman, M., Porac, J., Lathin, D., & Deci, E. L. (1978). On the importance of self-determination for intrinsically-motivated behavior. *Personality and Social Psychology Bulletin*, 4, 443–446.

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

而真正的阅读，应该在书中与智者同行思考，借他们的视角看到世界的多元性，提出比答案更重要的好问题，在不确定的时代中领先起跑。

湛庐阅读 App：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本

怎么读是读者面临的最大阅读障碍

“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”App。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、电子书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者为伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。

从单一到复合，从知道到精通，从理解到创造，湛庐希望建立一个“与最聪明的人共同进化”的社区，成为人类先进思想交汇的聚集地，与你共同迎接未来。

与此同时，我们希望能够重新定义你的学习场景，让你随时随地收获有内容、有价值的思想，通过阅读实现终身学习。这是我们的使命和价值。

CHEERS

本书阅读资料包

给你便捷、高效、全面的阅读体验

本书参考资料

湛庐独家策划

- ✓ 参考文献
为了环保、节约纸张，部分图书的参考文献以电子版方式提供
- ✓ 主题书单
编辑精心推荐的延伸阅读书单，助你开启主题式阅读
- ✓ 图片资料
提供部分图片的高清彩色原版大图，方便保存和分享

相关阅读服务

终身学习者必备

- ✓ 电子书
便捷、高效，方便检索，易于携带，随时更新
- ✓ 有声书
保护视力，随时随地，有温度、有情感地听本书
- ✓ 精读班
2~4周，最懂这本书的人带你读完、读懂、读透这本好书
- ✓ 课程
课程权威专家给你开书单，带你快速浏览一个领域的知识概貌
- ✓ 讲书
30分钟，大咖给你讲本书，让你挑书不费劲



湛庐编辑为你独家呈现
助你更好获得书里和书外的思想和智慧，请扫码查收！

(阅读资料包的内容因书而异，最终以湛庐阅读App页面为准)

-
- (1) 《影响力》出版以来，我不曾面对博伊尔所预测的那种愤怒而居高临下的态度，哪怕同事里最硬派的学者也没这么做过。有必要对此稍做解释。我认为这主要出于两个原因。首先，和日报上“人文”文章里出现的通俗社会科学形式不同，我扎扎实实地引用了数百本出版物作为我陈述和结论的基础。其次，我想要争取推广的是一种研究人类反应的特殊方法，即实验行为科学的方法，而不是推广我自己的研究或特定类别的研究。我当时并非有意为之，但它消除了实验行为科学家同行们的戒心。这或许证实了我长久以来秉持的一种信念：人总不会把自己搭乘的船给弄沉。
 - (2) 唉，上网查了一下，我发现我不能把这句见解独到的名言归功于我祖父。它出自他著名的同胞、意大利现代作家朱塞佩·托马西·迪·兰佩杜萨（Giuseppe Tomasi di Lampedusa）。
 - (3) 读到能量饮料实验时，我正在购买能量饮料，好帮助自己赶在即将到来的截止日期前完成一个大型写作项目。没看到研究结果前，我从未想过按优惠价（我最喜欢买打折货了）购买能量饮料居然会削弱它们的效果。
 - (4) 对雌火鸡实验的完整描述可参见动物学家福克斯1974年的专著。巧的是，这位动物学家的名字就叫“Fox”（也有“狐狸”的意思）。
 - (5) 兰格是哈佛大学著名的心理学家、积极心理学奠基人，想了解其更多研究，欢迎阅读由湛庐策划、浙江人民出版社出版的兰格“专念三部曲”。——编者注
 - (6) 制药公司的礼物不光会影响科学家对药物效力的研究结果，还会影响医生对这些药开具处方的偏好。制药公司向医生支付的诸如教育培训费、演讲费、差旅费、咨询费、会议注册费等费用，跟医生对赞助药物开具处方的频率相关。哪怕是一顿便宜的午餐，也足以对此产生影响——当然，更昂贵的大餐跟更高的处方率相关。
 - (7) 关于古巴导弹危机如何结束的新论述，最全面的研究成果来自谢尔登·斯特恩（Sheldon Stern），他曾在约翰·肯尼迪图书馆担任了23年历史学家。
 - (8) 认为人类纯粹是通过自然进化发展起来的美国人占33%这一数据，来自皮尤研究中心的一项调查，它还记录了宗教信仰在抵制进化论方面所扮演的重要角色。研究表明，理解进化论和相信进化论之间没有关系。
 - (9) 相似性的广泛影响，可明显见于教育环境、讨价还价的结果、选民决定、浪漫感觉以及人质谈判。有证据表明，被影响的人往往会低估它的力量。通过培训可提高餐厅服务员的小费、电子产品销售员的利润、谈判人员的工作成效，以及速配约会者的成功率。

(10)研究表明，西方文化下的人们通常更关注差异而非共性。尽管我不清楚是否有人做过相关研究，但值得指出的是，东方文化在传统上强调和谐，所以不知道不会出现同样的模式。研究显示，人们最初会低估与群体外成员互动可取之处，而且男女两性都容易犯这个错误。显然，女性虽然有着公认的追求人际和谐的倾向，但如果对方来自其他群体，她们也并不会少犯这类错误。

(11)原文是“like”，它在社交网络上有“点赞”的意思。——译者注

(12)脑成像研究来自加州大学洛杉矶分校脑测绘中心的研究者。

(13)研究发现，人们乐于把自己跟好消息联系起来，而跟坏消息保持距离——哪怕他们并非事情的起因。这种倾向的出现似乎是因为，人们意识到自己会沾染所传递信息的特点。在别人背后表达赞美有很大的优势，它可以避免让第三者觉得你别有用心。因为如果别人怀疑你的赞美别有用心，那么赞美就会削减别人对你的信任。

(14)最初介绍角色设定这一影响力技巧的人是社会学家，此后，心理学家对它做了理论发展。不管是把值得称赞的特质赋予孩子还是成人，都能产生更多的同类行为。

(15)研究者不仅记录了积极接触对改善不同种族、民族或性取向的群体成员关系的益处，还指出了减少焦虑增加同理心等产生上述益处的原因。

(16)其他版本的合作学习模式在不同的学校体系甚至商业组织等不同类型的机构都产生了类似的结果。

(17)毫无疑问，不管是正面还是负面的消息都会受到这种无辜的牵连。例如，顾客听到的音乐，就会对与之相关的产品产生正向偏好，反之亦然。有证据表明，人们往往和自己的朋友有着同类特点，汽车广告里的漂亮模特也会让男性更喜欢那辆车。关于信用卡对支付意愿的影响效应，相关研究还将其扩展到了餐厅的小费多少上和体育比赛的门票价格上。在关于门票价格的研究中，为了看一场职业篮球比赛，球迷们使用信用卡时愿意比使用现金时为门票多支付一倍的价钱。

(18)在英语里，“马尔斯”（Mars）也是“火星”的意思。

(19)企业花大手笔赞助的体育赛事，不光只有奥运会。在2018—2019赛季，美国国家橄榄球联盟获得的企业赞助总额高达13.9亿美元。当华尔街投资者注意到棒约翰比萨不再是美国国家橄榄球联盟的赞助商后，前者股价立刻下跌了8%。当时的记者记录下了马尔斯棒棒糖和日产“侠盗”SUV这两个流行文化现象偶然影响了消费品销量的故事。研究者也发现，“特卖”标语同仅仅降低价格相比，能带

来更多的销量。

(20)我的研究团队最初是在美国橄榄球迷身上进行“沾染荣耀光彩”效应研究的，随后又在法国和英国足球球迷身上做了重复研究。另一项研究证实了这种做法的有效性：能够跟成功的运动队建立起联系的人会更受旁观者的青睐。不过关联原理也表明，如果旁观者对成功的运动队并无好感，那这种效应就会发生逆转。研究者发现，“沾染荣耀光彩”效应适用于一种特定的、来自他人的理想评价形式。在旁观者眼里，声称自己跟成功的篮球选手有紧密关系（“好朋友”）的人，本身也会显得更成功。

(21)利用社会认同减少企业排污的国家包括印度尼西亚和印度。

(22)研究表明，对他人犯罪频率的感知与潜在犯罪者自己实施犯罪的可能性有关。如果一个人认为对伴侣施暴是件常见的事，那他自己更有可能做出此种行为，但好消息是，如果他通过证据得知不良行为并非社会规范，便会自我克制。98%的网络购物者表示，真实的顾客评论是影响他们购买决定最重要的因素。

(23)这是一种去中心化的宗教及灵性概念，发源于20世纪70年代到80年代的西方社会。——译者注

(24)关于吉诺维斯邻居们的“冷漠”的文章最先被作为长篇报道发表在《纽约时报》的头版，这篇先红极一时后又臭名昭著的报道还被《纽约时报》当时都市版的编辑亚伯拉罕·迈克尔·罗森塔尔（Abraham Michael Rosenthal）收录进了自己的书中。一些研究者推翻了这一记叙的许多核心细节。另有证据表明，如果旁观者确信存在紧急状况，是不太可能无动于衷的。

(25)流媒体是指将一连串的媒体数据压缩后，通过网络分段发送数据，在网上即时传输影音以供观赏的一种技术与过程。——编者注

(26)人们如果渴望得到社会认同，就更有可能在一件事上与群体观念保持一致；更危险的是，他们还可能顺从群体的饮酒水平。

(27)弗兰克是通俗经济学鼻祖，想了解其更多观点，欢迎阅读由湛庐策划、北京联合出版公司等出版的通俗经济学开山之作“牛奶可乐经济学”系列。——编者注

(28)几支研究小组均已证实，如果告诉焦虑的学生，有很多跟他们一样的学生克服了自己类似的担忧，那焦虑的学生就能更好地适应学校环境。另外大学生对少数群体的态度，会因其同侪的态度而发生改变。

(29)对维特效应的研究是菲利普斯的系列研究的开始，现代的研究手法，可以在对奈飞的系列网剧《13个原因》的研究中看到。此后，他还继续考察了大规模报道飞机和汽车蓄意自杀事故给公众带来的影响。加利福尼亚州高中生传染性卧轨

自杀的报道出自《洛杉矶时报》的记者，其对媒体在传染性自杀中扮演的角色进行了回顾。美国大规模谋杀事件在近年来愈演愈烈。2017年丧命于此类事件的总人数创下了224人的纪录，而有记录以来发生此类案件最多的年份是2019年，达到了41起。关于影响品牌市场份额因素的相关研究显示，不同地区居民的受影响因素和程度存在很大的个性和态度差异。

(30) 在一项旨在减少刻板印象的研究中，研究者告知被试，“虽说让人感到遗憾，但刻板印象是普遍存在的”，这使被试表现出了更强烈的刻板印象。

(31) 我们对节约用水趋势影响的研究，还收录了一项关于无偿完成问卷调查的意愿的研究，两项研究结果类此外，研究者还证明了趋势对其他不太普遍的行为的积极影响，如吃素食、减少糖分摄入、在食堂选择可重复使用的杯子以及高中和大学里的女学生选择在理工科领域继续深造等。

(32) 英国内阁行为透视团队的慈善研究中，两条影响力原则结合起来产生了较大的影响，但如果认为在说服信息里加入多条原则总能增加其影响力，这就错了。把多种策略硬塞到同一次交流里，会让接受者意识到对方在努力劝说自己，从而造成适得其反的效果。

(33) 有人对米尔格拉姆研究之后进行的一系列关于服从的研究做了综述，得出结论认为：20世纪60年代米尔格拉姆在美国所做的研究中发现的服从程度，跟更晚近时代的研究结果有着惊人的相似性。假酷刑类的电视游戏节目同样揭示了人们在相关情境中的服从意愿，与同类研究的结果一致。

(34) 遗憾的是，自从20多年前美国国家卫生研究院发布第一份关于医疗失误严重程度的报告《人非圣贤》（*To Err is Human*）以来，情况并无好转。一如研究者指出的那样，大部分问题不在于医生的操作手法而在于其心理状况。

(35) 研究表明，教师、政客和任务中的被试会基于其感知地位而表现出身体上的“变高”。此外，比对手高的政客往往能获得更多选票。例如，在自1900年以来的美国总统选举中，90%都是长得更高的候选人胜出。因此，在人们心目中，地位不光能增加身高，身高同样能提升地位。

(36) 研究者在对护士研究结果表明，护士们本身可能并没有充分意识到，他们的判断和行为会受到“医生”头衔的多大摆布。研究者另外采访了33名护士和实习护士，问他们在相同的条件下会怎么做，他们的回答和实际调查结果相去甚远，只有2%的人认为自己会按电话里的吩咐去给药。

(37) 研究发现，如果员工穿着制服提供服务，那么得到糟糕服务的人就有更大可能责怪组织，而不是员工本人。人们几乎是一瞬间就认定了穿着高品质服装与低品质服装的人的能力，从而影响面试结果。这一研究的作者指出了其中令人感到

不安的一个方面：经济背景较差的人买不起昂贵的衣服，在面试中就明确地陷入了劣势，而且这一切完全是自动发生的。

(38)标准普尔指数是一个记录美国500家上市公司的股票指数。——编者注

(39)一篇专家文章对读者观点的重大影响，同时适用于普通读者和专业“精英”，如智库学者、记者、银行家、法学教授、国会工作人员和学者。追随那些似乎知道自己在做什么的人的倾向，人们从婴儿时期就表现出来了。

(40)卡尼曼是行为经济学之父，对人类行为的本质有深刻的洞察，想了解其更多观点，欢迎阅读由湛庐策划、浙江教育出版社出版的《噪声》。——编者注

(41)有关损失在心理上占主导地位的研究，如对大学食堂满意度、管理决策、职业高尔夫球手比赛、大学生情绪、更换电力供应商的偏好、执行任务者是否作弊的选择、个体身体反应的研究等，都揭示了前景理论的广泛适用性。来自多种背景的证据表明，如果风险很高，不确定感很强，损失厌恶情绪就会特别强烈，这同样适用于健康、医疗行业。然而，如果风险较低，不确定感较弱，那么，增进倾向相较于保护倾向就会占据主导地位，此时人们重视“得”甚于“失”。

(42)做出承诺之后，承诺就驱使人做出随后的反应，相关证据见于赛马、政治选举、努力节约资源等活动中。各类研究都普遍证实了保持一致倾向给人们带来的压力。

(43)虽然费斯廷格并非第一个提出“一致性需求是人类行为中心”的著名理论家，但他是人们最容易想到的理论家。他的认知失调理论始于如下假设：我们会对自己的失调感到不舒服，并采取措施减少或消除它们，哪怕这需要自我欺骗。我们不仅把前后不一致视为自己身上的一种负面特质，也不喜欢别人前后不一。有充分的证据证明，一致反应会以自动的方式发生，以求避免理性思考可能带来的不合意结论，又或许这种自动仅是为了避免严谨的思考，因为严谨的思考是很累人的。机械地保持一致的倾向除了具备上述好处，阐释或选择保持一致，往往还能带来准确的决策。

(44)在汽车销售、慈善志愿活动、选举日投票、居家购物、自我介绍、健康行为的选择和避免出轨等领域，相对较小的口头承诺能带来可观的行为改变，这既值得注意，也很有启发意义。

(45)这里，我要澄清一个谣言：有人说，我为这个任务招聘研究助理时，先问他们是否乐意主持“思维过程”实验，等他们答应了以后，才告诉他们开始时间是早晨7点。这个说法毫无根据。

(46)有充分的证据表明，美国居民有着较为强烈的个人主义倾向，这种倾向使他们往

往有着与先前选择保持一致的偏好。

(47)本章摘取并更新了我于2016年所著的《先发影响力》（该书已于2017年9月由湛庐策划、北京联合出版公司出版）一书中的部分资料，并获得了出版社的许可。

(48)巴菲特的那封致股东的信是伯克希尔-哈撒韦公司2014年年报的一部分，该年报于2015年2月公布。

(49)有研究者指出，杉原千亩的签证救助了上万犹太人，他们绝大多数在日本领土找到了避难所。很多历史学家都曾描述过日本决定庇护犹太人的相关事件，但最详尽的记述来自东京前首席拉比马文·托卡雅（Marvin Tokayer）。细心的读者或许已经注意到，在介绍大屠杀政策时，我将涉事方称为纳粹，而非德国人。因为在我看来，把德国的纳粹政权跟德国的文化或人民等同起来是不准确、不公平的。政府政权，往往是崛起于强烈而短暂的外界环境，不能公正地概括人民的特点。因此，在讨论德国的纳粹统治时期时，我并不把两者混为一谈。

(50)对于音乐是一种群体联盟的社会统一机制，能够促进自我与他人的融合这一概念，相关的支持证据越来越多，接受度也越来越高了。4岁孩子帮助行为的研究，在14个月大的婴儿身上得到了概念上类似的结果。一项针对成年人的研究，为乐于助人的态度提供了解释。一起唱歌会导致自我与唱歌同伴的融合。

(51)研究者的最新数据表明，年轻女性认为对她们自己来说，音乐比衣服、电影、杂志、电脑游戏、电视、体育等更重要，但不如爱情重要。充分的科学证据表明，音乐和节奏的加工是与理性加工过程相独立的。但看看音乐家们本身如何看待这个主题，恐怕更有启发意义。科斯特洛的引言，来自伊丽莎白·赫尔穆特·马古利斯（Elizabeth Helmuth Margulis）的一篇有趣文章，作者还补充了一段新的证据，如果在听众倾听贝多芬弦乐四重奏节选前先给他们看结构性信息，会减少他们享受到的音乐的乐趣。有人研究了40年来流行歌曲的内容，发现80%都以爱情或性为主题。法国吉他盒实验记录了成功索要到电话号码的概率：背吉他盒，31%；背运动包，9%；什么也不背，14%。

(52)双方互相问36道问题这一程序的修改版，已经被用来减少种族偏见，即使对初始态度存在高度偏见的人也适用。

(53)研究者在波士顿马拉松爆炸案后分析了共同的痛苦对群体身份融合的影响，并对北爱尔兰统一党及共和党之间的长期冲突所带来的后果做了类似分析，得到了相似的结果。研究表明，将手浸入冰水，以及其他能导致疼痛的做法（如吃辣椒、反复下蹲），都能增进小组成员的团结。“血泪之路”的苦难程度，仅见于一段鲜为人知的事实。它最初的标签是来自一名乔克托族酋长的描述：“充满泪

水和死亡的小路。”

(54) 利奥波德的《沙乡年鉴》首次出版于1949年，自此以后成为许多环保团体的必读书目，书中转述的有关桦树与松树的思辨也来自此书。他坚信，管理荒野最好是通过以生态为中心的方式，而非以人类为中心的方式。他反对政府控制自然环境中食肉动物的政策，狼群为他的立场提供了非常有力的支持证据。

(55) 对同事及两人共同创作的产品的评价研究，是我跟杰弗里·普费弗（Jeffrey Pfeffer）合作完成的，他给我留下了深刻的印象。研究者对合作学习技术的积极结果做了归纳，这一技术被称为“拼图教室”（Jigsaw Classroom），由阿伦森及其同事发明。

(56) 挥手理论是指试图解释或说服而使用无实质性的言词、论据、姿势或动作等，也可理解为一种粗暴、大而化之的断言。——编者注

(57) 如何判断不同物种成员的血缘关系，一直是无数科学调查的主题。关于人类怎样进行这一过程的科学调查不多，但它们对我们此处讨论的话题有着特别重要的意义。例如，有报告指出，“荣誉亲属”的概念，存在于所有人类文化中。“荣誉亲属”指的是住在家里但没有血缘关系、同时也得到了类似家人称呼的人。

(58) 研究者收集的数据，揭示了生活在多元化社区给陌生人之间的善意及全人类身份认同带来的积极影响。在多种族地区和国家也发现了概念类似的效应。跨群体友谊对多数群体和少数群体成员对待另一群体的态度、期望和行为的有利影响，相关证据来源很多。例如在南非，与白人建立跨群体友谊的“有色人种”初中生，总体上对白人持有更多的信任态度和更少的伤害意图，还有研究者设计了另一个版本的36道问题，旨在让持顽固偏见态度的个人之间减少偏见。

CHEERS
湛庐

特沃斯基 精要

[以色列]
阿莫斯·特沃斯基 著

李慧中 译

Amos Tversky

行为经济学奠基人
一生的十四篇精华之作
诺贝尔经济学奖得主
丹尼尔·卡尼曼倾情推荐

The Essential Tversky

 浙江教育出版社
ZHEJIANG EDUCATION PUBLISHING HOUSE

版权信息

书名：特沃斯基精要

作者：〔以色列〕阿莫斯·特沃斯基（Amos Tversky）

译者：李慧中

出版社：浙江教育出版社·湛庐

出版时间：2022年09月

ISBN：9787572243493

字数：426千字

版权所有·侵权必究

Amos Tversky

阿莫斯·特沃斯基

20 世纪最具影响力的 100 位心理学家之一
行为经济学奠基人



20世纪最具影响力的100位心理学家之一

1937年，阿莫斯·特沃斯基出生在以色列。特沃斯基从小就表现出强烈的好奇心，并总是尝试自学很多东西。特沃斯基的母亲称他是一个独立的人。在数学等很多领域，特沃斯基都是自学成才的。

1961年，特沃斯基在以色列希伯来大学获得学士学位；1965年，在美国密歇根大学安娜堡分校获得博士学位。博士毕业之后，特沃斯基回到以色列，开始在希伯来大学任教。1978年，特沃斯基加入了美国斯坦福大学，并在那里度过了他的职业生涯。

1980年，特沃斯基当选美国艺术与科学院院士。1983年，特沃斯基获得美国心理学会颁发的杰出科学贡献奖。1984年，特沃斯基同时获得了古根海姆奖与麦克阿瑟天才奖。

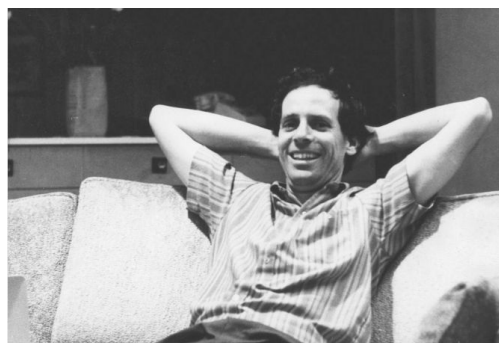
特沃斯基与夫人芭芭拉·特沃斯基于1963年结婚，两人育有三个子女。

1996年，特沃斯基因转移性黑色素瘤去世。直到生命的最后几天，他仍在家中工作。

2002年，特沃斯基被广受欢迎的心理学期刊《普通心理学评论》评为20世纪最具影响力的100位心理学家之一。

让我们把事情做好。

——阿莫斯·特沃斯基



行为经济学奠基人



特沃斯基多年来专注于行为决策领域的研究，主要集中于这三个方面：不确定状况下的判断、风险决策和理性选择。

特沃斯基和丹尼尔·卡尼曼在十几年间共同探索了人类决策中的缺陷。二人于1979年提出了“前景理论”，将心理学研究应用于经济学，在不确定状况下的判断与决策方面做出了突出贡献，揭示了人类非理性的经济选择，奠定了行为经济学的基础。前景理论认为，由于参考点的不同，个体会对风险表现出不同的态度。与类似的收益相比，个体更有可能规避损失，因为损失比收益会带来更强烈的刺激。

此外，特沃斯基还与诺贝尔经济学奖得主理查德·塞勒、心理学家保罗·斯洛维奇、心理学家托马斯·吉洛维奇等人合作开展研究。

卡尼曼曾经这样评价特沃斯基：“阿莫斯是我所认识的最自由的人，他之所以能够获得自由，是因为他同时也是最自律的人。”

你越早意识到特沃斯基比你聪明，你就越聪明。

——理查德·尼斯贝特
美国密歇根大学心理学家

因英年早逝错失诺贝尔经济学奖

我们本来可以一起来的。

——丹尼尔·卡尼曼

2002年，阿莫斯·特沃斯基最好的学术搭档丹尼尔·卡尼曼凭借与特沃斯基合作的研究而获得2002年诺贝尔经济学奖，但由于诺贝尔奖不会授予已经去世的人，所以特沃斯基未能获奖。

获奖之后不久，卡尼曼在接受采访时表示：“我觉得这是一个属于我们两个人的奖项，我们是十几年的‘孪生兄弟’。”

卡尼曼曾经这样评价特沃斯基：“他在选择问题上有着完美的品位，从不在注定不重要的事情上浪费时间。”



作者演讲洽谈，请联系
BD@cheerspublishing.com

更多相关资讯，请关注



湛庐文化微信订阅号

湛庐CHEERS 特别制作

测一测 你对“行为决策”了解多少？

· 一位飞行教练发现：对顺利降落的学员进行表扬，他未来的表现不会太好；但如果对勉强降落的学员进行批评，那这个学员下一次的表现则会有所提升。所以教练认为批评的作用大于表扬，这个结论可靠吗？

A. 可靠

B. 不可靠

· 在许多决策问题中，保持现状也是一种选择吗？

A. 是

B. 否

· 人们往往认为自己支持的球队会赢得接下来的比赛。这种现象体现了人们的：

A. 偏好

B. 理性判断

C. 外延性推理

D. 不确定性

扫描下面二维码查看本书更多测试题

扫码鉴别正版图书
获取您的专属福利



扫码获取全部测试题及答案，
一起了解“行为决策”

前言 纪念阿莫斯

迈克尔·刘易斯（**Michael Lewis**）

我可以确定地说：“特沃斯基测试”是最简洁的智力测试。该测试由美国密歇根大学心理学家理查德·尼斯贝特（**Richard Nisbett**）创建，其运行原理如下：当你遇到阿莫斯·特沃斯基时，你越早意识到特沃斯基比你聪明，你就越聪明。我从未见过阿莫斯·特沃斯基，当然也就未能亲身体验特沃斯基测试。但是我遇到过各种各样的人，他们对这个测试的设计都没有任何异议。阿莫斯的头脑并不仅仅是强大和有趣，还是他人用来衡量自己的标尺。即使在阿莫斯去世几十年后，人们仍然在这样做。

在一本关于阿莫斯·特沃斯基与以色列心理学家丹尼尔·卡尼曼（**Daniel Kahneman**）合著的书⁽¹⁾中，我遇到了很多此类现象。我发现许多人在阿莫斯去世后的几十年里仍然在问自己：“阿莫斯对此会怎么想呢？”或“阿莫斯会说什么？”我遇到过一位未婚女士，她说自己在面对求婚者时总会忍不住问自己：“这个家伙在阿莫斯的眼中会怎么样？”（答案一直是：不怎么样！）我遇到过几位有影响力的知识分子，他们都曾与阿莫斯合作过。他们表示总忍不住想象阿莫斯正在阅读自己准备发表的论文。如果想象中的阿莫斯没有认可他们的成果，他们就会重新打磨。我甚至见过阿莫斯之前的学生非常小心地保存着阿莫斯赠予自己的最平常的物品：一本旧书、一个订书机、一支激光笔，这些都成了神圣的遗物。

最重要的是，人们如此珍视阿莫斯身上那种令人震惊的不经意间释放出的智慧的光芒：

做好研究的秘诀在于保持一点儿松弛感。如果不舍得浪费几小时放松一下，那么最终可能会浪费几年的光阴。

如果有人要求你做某件事，而你当时并不知道自己是否想做的話，那就等一天。之后你会惊奇地发现，你答应做的很多事情其实并不是你真正想做的事情。

做好科学研究的一部分因素是：看到他人都能看到的，思考他们未说出口的。

有的人让人难以忘怀，阿莫斯就是其中之一。但是，许多挽留他的

尝试都失败了。丹尼尔·卡尼曼在对阿莫斯的悼词中指出了这个问题的困难所在。在阿莫斯的葬礼上，卡尼曼提到，阿莫斯曾让其为两人合著的论文集写一篇序言，阿莫斯给定的主题是：“相信你心中的那个我”。但是，谁能真正定义自己心中的阿莫斯呢？阿莫斯思想的奇妙之处在于，他想到的事情是别人永远都无法想象的。卡尼曼说，当人们想象阿莫斯的时候，他们脑海中出现的一切都只是阿莫斯本人的廉价仿制品。

阿莫斯自己并没有对此做更多的说明。事实上，阿莫斯强烈反对任何形式的纪念活动。当得知阿莫斯将要离世时，他的学生们询问是否可以为美国斯坦福大学的本科生录下他最后的演讲内容，阿莫斯拒绝了。临终前，阿莫斯要求斯坦福大学校长不要举办任何形式的纪念活动，如以他的名字来命名年度讲座（他不希望自己的名字与他无法控制质量的事情相关联）。阿莫斯向最亲近的人解释到，他认为自己在这一生中已经达成了目标，且很快就会被遗忘，而对此他并不介意。

也许如此。但是，真的，他并没有被遗忘。

引言 认知与决策科学领域的灯塔

埃尔德·沙菲尔 (Eldar Shafir)

阿莫斯·特沃斯基是认知与决策科学领域中灯塔般的存在。他的研究产生了巨大的影响；他开创了新的研究领域，并促进了其他相关学科的转型。他的研究引人入胜，具有审美价值且设计巧妙。同时，这些研究也改变了社会科学对人类行为和社会的研究方式。特沃斯基的研究被收录进迈克尔·刘易斯的《思维的发现》（*The Undoing Project*），该书记录了特沃斯基和他的朋友兼同事丹尼尔·卡尼曼长期以来的合作。在该书中，刘易斯提到，他发现特沃斯基和卡尼曼的很多思考已经为自己早期那本非常有名的书《魔球》（*Moneyball*）奠定了基础，而在当时，刘易斯并未意识到这一点。特沃斯基和卡尼曼所做的开创性研究也在卡尼曼的畅销书《思考，快与慢》（*Thinking, Fast and Slow*）中占据着核心地位。这些及最近出版的一些著作已经将特沃斯基的思想从专业学术领域推广到了更广泛的受众群体中。这本《特沃斯基精要》正是想借特沃斯基的一些原创文章使读者了解其思想。

本书包含特沃斯基的14篇学术文章，它们不仅具有巨大的影响力，而且可读性强，同时能充分体现特沃斯基思想和研究的特点。这些文章是从他在生命的最后几个月与他人合著的40篇文章中挑选出来的。即使是原创选集也只是特沃斯基已出版著作的一小部分，通过该选集，读者可以对其卓越成就稍作了解。想读到更多信息的人可以阅读1 000页的原始文献，或者阅读每章结尾部分的参考文献，其中包括了特沃斯基的完整参考书目。

本书所选择的文章涵盖了从医学、统计学到心理相似性和经济学等一系列主题，同时避免了特沃斯基著作中许多更具技术性的章节。本书的目的是让那些可能知道一些特沃斯基的观点却无缘跟随他学习或想要更深入地阅读他的作品的人，可以更加了解他非凡的思想。对于那些曾受其启发的人来说，这也是一份珍贵的纪念品。

特沃斯基于1937年3月16日在以色列海法出生。他的父亲是一名兽医，母亲是一名社会工作者，后来成为以色列第一届议会议员。特沃斯

基于1961年在耶路撒冷希伯来大学获得了文学学士学位，主修哲学与心理学，并于1965年在美国密歇根大学获得哲学博士学位。特沃斯基曾在希伯来大学（1966—1978年）和斯坦福大学（1978—1996年）任教，在斯坦福大学任教期间，他曾任行为科学领域的第一任戴维斯-布拉克教授（Davis-Brack Professor）和斯坦福冲突与谈判中心首席研究员。

特沃斯基早期在数学心理学方面的研究，重点关注个人选择行为研究和心理测量分析。几乎从一开始，特沃斯基的研究就在探索那些针对相关理论做出的简单且看起来可信的心理学假设的深刻含义；在此之前，这些理论似乎都是不言而喻的。在一项被大量引用的特沃斯基1969年的研究中，他通过心理学上的显著差异来预测违背传递性

（transitivity）的行为，传递性是选择规范理论最基本的原理之一。这项研究成果提出了很多问题，这些问题在特沃斯基后来的研究中起到了关键作用。这项研究结果向我们暗示了一些行为方式，即完全理性的人在做出看似合理的决定时，最终可能会违背最基本的理性决策原理。该研究还解决了这一问题：在没有对所涉机制和费用进行令人信服的分析的情况下，难以就这类违背理性决策原理的行为的合理性给出明确结论。研究显示，在选择过程中，简化可能是非常有用的方法，虽然它偶尔可能无法产生最佳选择。特沃斯基最感兴趣的是对规范性原则的系统性违背，以及这些违背所揭示的支配行为的心理机制。

特沃斯基早期作为数学心理学家，对相似性的形式化和概念化产生了浓厚兴趣。相似性的概念是心理学理论中的基础概念，它在学习、记忆、知识、感知和判断等理论中发挥着基础性作用。当特沃斯基开始这一领域的研究时，相似关系的理论分析一直被几何模型所主导，在该模型中，每个物体都由多维坐标空间中的点来表示，而点之间的距离对应着物体之间的相似性。特沃斯基改变了这一切。在其有影响力的相似性模型（第3章）中，特沃斯基做了一些简单的心理学假设：物体在心理学层面上表现为特征集合，它们之间的相似性随着共同性特征的增加而增加（递增函数），随着区别性特征的增加而减少（递减函数）。而且，相似性判断使我们更加重视共同性特征（例如，尽管儿子与父亲年龄不同但长得很像）而相异性判断则更注重区别性特征（例如，尽管父子二人长得很像，但年龄相差很大）。此外，这个简单的理论可以解释观察到的相似性判断中的非对称性，如A与B的相似性比B与A的相似性更高，以及选项A可能被认为与选项B非常相似，选项B和选项C非常相似，但选项A和选项C可能被视为非常不同（Tversky & Gati, 1978）。这些由特沃斯基基于特征模型预测得到的显著效应，与之前的几何方法

是不相容的。在许多方面，这些早期论文都预见到了后来出现的精彩研究。凭借对规范性理论相关技术的精通，以及对简单而令人信服的心理学原理的探索，他们在作品中提出的那些出人意料的理论才产生了越来越广泛的影响，常令读者产生醍醐灌顶之感。

特沃斯基与卡尼曼在1969年就开始了长期且极具影响力的合作，他们的研究横跨整个判断和决策领域。他们合著的第一篇论文的主题与“小数定律”有关（第10章），该论文表明无论是缺乏经验的被试，还是受过训练的科学家，都会对随机抽样有着强烈但错误的直觉。研究证明，这些直觉会导致对偶然事件的系统性误解，特沃斯基及其同事之后将这一发现用于分析普遍存在但显然具有误导性的信念，如篮球比赛中的“热手效应”，以及相信关节炎疼痛与天气有关的现象。

不涉及动机因素的认知和知觉偏差，成了他们极具创造力和影响力的“启发式和偏差”（heuristic and bias）理念的核心内容。由于已经认识到直觉上的预测和概率判断不符合统计和概率论原则，特沃斯基和卡尼曼开始进行偏差研究，并将其作为探究判断启发式的一种途径。他们在发表于《科学》上的那篇文章中提出了3种启发式，分别是代表性启发式、可得性启发式、锚定与调整，这些都是人们在评估概率和预估价值时会出现的情况。在一些简单概率规则的相关性显而易见的情境下，人们通常会表现出正确的统计直觉。但在相对复杂的情境中，人们往往会依赖那些稍加思考就会发现破绽的启发式，因此会做出错误的判断。例如，当我们依赖于代表性启发式时，选项A属于B类的可能性是通过A与B的相似程度来判断的。先验概率（prior probability）和样本大小这两个与可能性高度相关的因素，对一个选项的代表性没有影响，因此常常会被人们忽略。这会导致显著性错误，如“合取谬误”（conjunction fallacy），也就是说，人们会因为合取事件（conjunctive events）更具代表性，而认为合取事件的发生概率比其中包含的独立事件发生的概率更大（第2章）。

这项研究的巧妙之处在于，它清晰地展示了心理直觉和规范理论的相互作用，而且还给出了令人难忘的例证。这项研究表明，虽然人们对各种原则的规范性诉求很敏感，但他们的判断常常会违背这些基本原则。这种不可靠的直觉与对规范性判断的潜在认同的共存，向人们展示了概率推理的微妙之处。特沃斯基研究的一个重要主题就是，他认为人们并不是因为不够聪明或经验不够丰富，才无法掌握重要的规范性思维。相反，特沃斯基将这种反复出现的系统性错误归因于人们过于依赖

那些并不遵循规范性标准的启发式过程。例如，戴尔·格里芬（Dale Griffin）和特沃斯基运用人们关注证据的说服力（如推荐信的热情程度）的同时并不充分考虑其权重（如写信人对候选人的了解程度）的概念来解释概率判断中的各种系统性偏差，包括无法理解回归现象和人们的过度自信（证据具有显著性但权重低），以及偶尔的不自信倾向（证据不具显著性但权重高）。这种方法贯穿了特沃斯基的大部分研究。这些实验性演示值得我们关注，因为它们对流行且极具影响力的规范性理论发起了挑战。而且，这些演示具有显著性，因为表现出这种错误的人发现它们非常有说服力，但与自己设想的进行决策的方式惊人地不一致。

特沃斯基最初的著作与个人选择行为的研究相关，而这也贯穿了他的研究生涯。这项研究在很大程度上受到了经济学早期发展的启发，尤其是受冯·诺伊曼（von Neumann）和奥斯卡·摩根斯坦（Oskar Morgenstern）在其1947年发表的作品中对期望效用的规范性论述，随后这一研究成果也得到了进一步修正，它向人们展示了一些令人信服的定理，即在条件满足的情况下，一个人的选择最能体现其主观期望效用。20世纪70年代，特沃斯基和卡尼曼研究了一种有关风险决策的描述性理论，即“前景理论”（prospect theory；第4章）。前景理论包含了许多选择方面的基本心理原则，这些原则与规范理论在一些重要方面都有所不同。此外，前景理论假定了一个具有3种重要特征的效用函数：第一，它是根据收益和损失而不是总财富进行定义的，这表明人们通常将结果视为对于某个参考点的偏离，而不是最终资产；第二，损失给人们带来的刺激大于收益，因此财产损失带来的痛苦程度大于等量的财产收益带来的快乐程度，从而产生所谓的“损失厌恶”（loss aversion）；第三，收益曲线为凹形，损失曲线为凸形，这说明在收益中会产生风险规避，而在损失中会产生风险寻求（出现相反模式的概率极低）。这些特征是显著的，并且看起来是可以接受的。然而，它们会在规范性上导致一些问题。该理论预示的决策模式已经在很多研究中得到了证实，并与期望效用理论的基本假设形成了直接的对比。例如，损失厌恶在无风险的选择中扮演了非常重要的角色。除此之外，它也在很大程度上解释了我们观察到的人们放弃某个商品的最低意愿和为得到这个商品而花钱购买的最高意愿之间的巨大差距（第8章）。损失厌恶对人们的经济选择和摆脱现状的意愿有着深远的影响。数学上的巧妙设计和心理学上的洞见使得前景理论及其延伸对社会科学产生了重大影响。

特沃斯基对不确定性与偏好之间的关系特别感兴趣。在概率判断与

偏好的关系研究中，奇普·希思（Chip Heath）和特沃斯基提出了能力假说：在感觉自己有能力和有学识的情境中，人们会依赖自己的信念，而当他们感到自己没有能力或者无知时，则会赌运气（第9章）。有趣的是，这种模式与人们熟悉的“模糊规避”（ambiguity aversion）不一致。根据“模糊规避”，人们普遍倾向于把赌注押在已知的机会上，而不是概率模糊的信念上。沿着这个思路，克雷克·R.福克斯（Craig R. Fox）和特沃斯基提出了相对无知假设（第12章）。这一假设的内容是：对模糊的规避一般不会出现在没有比较的情境中，而是会出现在对相对不太模糊的事件进行比较的过程中，或者与更有学识的个体的比较中。

上述发现对由偏好来推断信念的基本概念提出了质疑。大多数关于在不确定性情境中进行决策的概念，包括规范性的和描述性的，都是结果论的。在这种意义上，人们认为决策是由对潜在后果及其感知概率的评估所决定的。然而，沙菲尔和特沃斯基记录了人们以非结果论的方式进行推理和选择的情况（第7章）。例如，那些无论不确定性消除至何种程度，都会做出相同选择的人，也会在情况仍不确定时做出不同的选择，这与结果论相反。

特沃斯基的研究最重要的一点是意识到偏好通常是由心理过程塑造的，而这些过程并不受人们经过深思熟虑形成的规范性考虑的约束。这些过程是“偶然偏好”（contingent preference）研究的基础。据此，描述、情境或程序中出现的一些非实质性的变化会改变人们赋予属性的权重，从而改变他们的偏好。唐纳德·A.雷德梅尔（Donald A. Redelmeier）和特沃斯基探讨了医学情境下的偶然偏好，他们认为从不同的角度看待问题可以改变属性的相对权重的分配，从而导致不同的选择（第6章）。他们发现，当医生将患者视为个体时，他们更重视后者的具体问题；而当医生将患者视为群体的一部分时，他们更重视效度和成本。因此，这些医生在评估单个患者时所做的决定与评估，可能与他们在评估一组患者时所做的决定不同，这种差异也存在于普通人的判断中。

决策者经常通过寻求和构建理由来进行艰难的决策，并为他们的选择进行辩护。不同的框架、情境和推导过程强调了各选项的不同方面，并提出了影响决策的不同理由和考虑因素。沙菲尔和特沃斯基研究了这些理由在决策过程中起到的作用（第14章）。基于理由的分析可以适应框架效应（framing effect）和诱发效应（elicitation effect），并且可以整合相对影响以及视角、冲突和情境等影响因素，而这些通常都未纳入

价值最大化的考虑范围。

心理常识构成了特沃斯基最深刻、最独到见解的基础。规范性理论背后的一个基本假设是“外延性原则”（extensionality principle）：具有外延等效性的选项会被赋予相同价值（无论如何描述，它们在世界范围内都是等价的，如“含1%脂肪”和“99%无脂肪”），外延等效性事件也会被赋予相同的概率。换句话说，这些理论是关于世界范围内的选项和事件的：对同一事物的不同描述并未改变这一事物的本质，因此得到了相似的评价。此外，特沃斯基的分析侧重于相关结构的心理表征。外延性原则在描述性上被认为是无效的，因为对相同的选项或事件的不同描述通常会产生具有系统性差异的判断和偏好。一个决策问题的描述方式，如收益或损失的不同角度，可能会引发相互矛盾的风险态度，从而导致对本质上相同的结果表现出不同的偏好（第5章）。类似地，对同一事件的不同描述会让人想起不同的案例，因此会产生不同的可能性判断（第13章）。在推导过程中，偏好和判断似乎是被建构出来的，而不仅仅是呈现出来的，它们的建构取决于问题的框架和推导的方法，以及这些所引发的评估和态度。

特沃斯基的研究表明，行为是人们在反思中认可的规范性理念的结果，并结合了可干预和塑造行为的心理倾向和过程，这些心理倾向和过程与是否经过深思熟虑无关。在掌握规范性研究技术，以及先凭直觉发现再以实验证明它们的奇特与影响方面，特沃斯基有着独特的能力。他是一位知识巨人，其著作具有极其广泛的吸引力。他的研究为经济学家、哲学家、统计学家、医学领域的学者、政治学家、社会学家和法律理论家等所熟知。特沃斯基的许多论文都具有开创性和权威性。阅读特沃斯基的论文就像在观赏大师级的作品：他为以前看起来令人困惑的领域绘制出了清晰的地图，并提供了一套新的工具和思路来思考这些问题。据统计，特沃斯基的著作在学术平台中被引用超过23万次！本书中的14篇文章每篇的平均引用次数都超过8 500次。

特沃斯基因其杰出的成就获得了许多奖项和荣誉。这些奖项包括：1956年因营救一名士兵的英勇表现而获得以色列最高荣誉；1984年获麦克阿瑟天才奖、荣誉学位等学术荣誉，并且被推选进入数个杰出社团。2002年10月，瑞典皇家科学院将诺贝尔经济学奖授予丹尼尔·卡尼曼，“因为他将心理学研究的见解整合到了经济科学中，特别是关于人们在不确定状况下的判断和决策”。诺贝尔奖颁奖词中还解释到，这些与阿莫斯·特沃斯基共同完成的研究，形成的多种理论能更好地解释现

实中人们的行为。瑞典皇家科学院没有设立追认奖项，但特别在颁奖词中介绍了特沃斯基在这方面的非凡贡献。“这个奖本来应该颁给我们两个人。”卡尼曼在消息宣布的当天说。两个月后，特沃斯基在去世后与卡尼曼一起获得了享有盛誉的2003年格文美尔教育奖（Grawemeyer Award）。该奖项的授予对象是在艺术和科学领域提出杰出理念的人。“在人文科学领域，我们很难找到比卡尼曼和特沃斯基提出的理念更有影响力的理念。”该奖项的颁奖词中这样写道。

特沃斯基以一种体现了强烈的自由和自主性的方式，成功地在自己的生活中将纪律和快乐结合在一起。夜以继日的工作习惯使他免受干扰，并让他有时间在研究中得到放松，也使他可以兼顾其他兴趣，包括对希伯来文学的毕生热爱，对现代物理学的痴迷，以及成为职业篮球比赛的专业球迷。他明智而坚定地拒绝了那些会分散注意力的事情，而“对于那些喜欢这类事情的人来说，”阿莫斯在拒绝各种各样的约会时总会带着他特有的微笑说：“这是他们喜欢的事情。”对于朋友和合作者来说，阿莫斯很有趣。他从与身边的人分享想法和经验中获得了极大的快乐，而这种快乐是可以传染的。他与许多朋友成了研究合作者，也和许多合作者成了亲密的朋友。他会花无数时间来构思一个观点，让它变得有趣，并不断打磨它。“让我们把这件事做好。”他会这样说，当然，他在这方面的能力是无与伦比的。

第1章 不确定状况下的判断——启发式和偏差

阿莫斯·特沃斯基
丹尼尔·卡尼曼

很多决策都是基于对不确定性事件发生概率的信心，如选举的结果、被告的罪行、美元未来的价值。人们通常会通过这些形式来表达这种信心：“我认为……”“概率是……”“……这不太可能”。通常，对于不确定性事件的信心会以概率或主观概率的数字形式显现。那么，这些信心的决定因素是什么？人们如何评估不确定性事件的概率或不确定量的价值？本章的主题就是人们如何基于有限的启发式规则，将评估概率和预测价值的复杂任务简化为决策判断。通常，这些启发式非常有用，但有时它们也会导致严重的系统性误差。

对概率的主观判断与对距离、大小等物理量的判断相似。这些判断都基于效度有限的数据，而这些数据都是通过启发式规则进行处理的。例如，一个物体的视觉距离与它的清晰度有关，清晰度越高，视觉上就越近。这种规则有一定的效度，因为在任何一个情景中，较远的物体在视觉上的清晰度都低于较近的物体。但是，对这种规则的依赖会导致在距离判断上出现系统性误差。具体来说，就是在视线较差的情况下，由于物体的轮廓变得模糊，距离往往会被高估。相对地，当视线较好时，由于物体可以被更清晰地看到，距离就容易被低估。因此，这种以模糊程度为线索的方法会在距离判断中导致特征性偏差。与启发式规则有关的系统性误差常常出现在对概率的直觉判断中。在下文中，我们将对3种用于概率判断和预测价值的启发式进行描述，同时还会列举这些启发式导致的偏差，并对这些观察结果的理论与现实意义进行阐述。

代表性

人们关注的很多概率问题都属于以下问题中的一种：A物属于B类的概率有多大？A事件是由B过程导致的概率有多大？A过程产生B事件的概率有多大？在回答这些人们通常依赖于代表性启发式判断的问题时，概率是由A在多大程度上可以代表B来决定的，即A与B之间的相似

程度。例如，当A对于B来说颇具代表性时，人们就会认为A由B产生的概率很高。而如果A和B不相似，人们就会认为A由B产生的概率很低。

为了说明人们在判断中使用了代表性启发式，我们可以参考这样一个事例。一个人的前邻居对他的印象是这样的：“史蒂夫非常害羞内向、乐于助人，但是他对现实世界和与人交往没有什么兴趣。他是一个谦恭且喜欢整洁的人，他需要的是秩序和组织结构，很注重细节。”那么，人们认为史蒂夫从事如农夫、销售员、飞行员、图书管理员、医生这样的职业的概率有多大呢？如果将这些职业以最可能到最不可能的顺序进行排列，人们会如何排序呢？根据代表性启发式，史蒂夫是一名图书管理员的概率取决于他自身的代表性程度或与图书管理员的刻板印象的相似性程度。事实上，针对这类问题的研究表明，人们在通过概率进行职业排序和通过相似性进行职业排序时使用的方法是一样的

（Kahneman & Tversky, 1973）。正如我们在下文中将要呈现的那样，这种判断概率的方法会产生严重的误差，因为相似性或代表性不会被那些应该会影响概率判断的因素所影响。

对结果的先验概率不敏感

对代表性没有影响但对概率判断来说非常重要的因素之一就是结果的先验概率，或可称之为基率的频率（base-rate frequency）。例如，在史蒂夫这个例子中，事实上在人口职业结构中，农夫的比例要比图书管理员大得多，因此对史蒂夫职业的合理推断应该是农夫而非图书管理员。但是，考虑到基率的频率，它并不影响史蒂夫与图书管理员和农夫的刻板印象的相似性。如果人们通过代表性进行概率判断，那么先验概率就会被忽略。这个假设在一个先验概率被显著操纵的实验中得以验证

（Kahneman & Tversky, 1973）。在这个实验中，被试看到了一些人的简短人格描述，研究人员声称这些人是从由工程师和律师组成的100人样本中随机抽取的。被试被要求根据每份描述来判断某个人是工程师而非律师的概率。在一个实验条件下，被试被告知这些人来自由70个工程师和30个律师组成的样本；而在另一个实验条件下，被试则被告知样本是由30个工程师和70个律师组成的。在工程师占多数的第一个实验条件下，任何特定描述属于工程师而非律师的可能性都高于律师占多数的第二个实验条件。具体来说，通过贝叶斯法则（Bayes' rule），这些可能性之比可以表示为 $(0.7/0.3)^2=5.44$ 。但这个实验的结果却明显违反了贝叶斯法则，两组被试的概率判断基本相同。很显然，被试在判断某一描述是工程师而不是律师的可能性时，依据的是这一描述在多大程度上

符合这两种职业的刻板印象，而很少甚至完全没有考虑两种类型的先验概率。

但是，如果人们完全没有其他信息，就会正确地运用先验概率。在没有人格描述的情况下，他们在两种基率条件下对某一个体是工程师的概率判断分别为0.7和0.3。然而，一旦引入描述，先验概率就完全被忽视了，即使这些描述完全没有信息含量。对以下描述的反应则说明了这种现象：

迪克是一位30岁的男性，已婚，没有孩子。他能力出众，能动性很强，他有能力在自己的领域获得巨大的成功。他的同事都很喜欢他。

在这一描述中没有包含任何可以用来判断迪克是工程师还是律师的信息。也就是说，如果没有给定任何相关描述，迪克是工程师的概率应该与该组中工程师的比例相等。但是，无论组内的工程师比例是0.7还是0.3，被试认为迪克是工程师的概率都是0.5。很显然，人们在毫无线索和面对毫无价值的线索时的反应完全不同。当没有具体的信息时，先验概率得到了正确的运用；当引入没有价值的信息时，先验概率完全被忽视了（Kahneman & Tversky, 1973）。

对样本量不敏感

为了评估从特定人群中抽取样本而得到的特定结果的概率，人们通常会运用代表性启发式。也就是说，他们会通过抽样结果和相关参数（如所有男性的平均身高）的相似性来对样本结果的可能性进行评估，如包含10个男子的随机样本的平均身高约为1.83米。样本统计量和总体参数的相似性并不取决于样本的大小。因此，如果通过代表性启发式来进行概率判断的话，那么样本统计量的判断概率就会基本上与样本量无关。的确，当被试评估不同大小的样本的平均身高分布时，他们得出的结论是一样的。例如，人们认为对于1 000人、100人和10人的样本来说，平均身高超过1.83米的概率是一样的（Kahneman & Tversky, 1972）。而且，即使在问题描述中强调了样本量的大小，被试也会忽略其影响。我们可以考虑以下问题：

某个小镇有两家医院。在较大的那家医院里，每天有45个孩子出生，而在较小的那家医院里，每天有15个孩子出生。已知大约50%的孩子是男孩。但是，每天出生的男孩的具体比例不同，有时高于50%，有时低于50%。在某一年，每个医院都记录了新生儿中男孩比例高于60%的天数。那么，你认为哪家医院记录的天数更多？

· 较大的医院（21）

- 较小的医院 (21)
- 两者大约相同 (即比例之差在5%以内) (53)

括号里的值表示选择这一选项的大学生人数。

大多数被试认为两家医院男孩的出生率高于60%的概率是一样的，可能是因为这些事件是以相同的统计量进行描述的，因此它们在总体参数上具有相同的代表性。相比之下，抽样理论表明，在较小的医院中，男孩出生率高于60%的天数大于较大的医院，因为相对较大的样本其偏离50%的可能性更低。这一统计学基本概念显然并不存在于人们的直觉之中。

这种对样本量的不敏感已经在“后验概率” (posterior probability) 的判断中得到了验证；也就是一个样本是从某个总体而非另一个总体中抽取的概率。请参考这个案例：想象一个罐子里装满了球，其中 $\frac{2}{3}$ 的球是一种颜色，另外 $\frac{1}{3}$ 的球是另一种颜色。某个人从这个罐子里取出5个球，发现其中4个是红色的，1个是白色的；另外一人从里面取出20个球，其中12个是红色的，8个是白色的。那么，这两个人中的哪一个人会更有信心认为：罐子里有 $\frac{2}{3}$ 的红球和 $\frac{1}{3}$ 的白球，每个人会给出什么样的概率？

在这个问题上，假设先验概率相等，那么对于红白球数量比为4:1的样本来说，正确的后验概率是8:1，而对于数量比为12:8的样本来说则是16:1。但是，大多数人认为第一个样本为罐子里多数是红球的假设提供了更有力的证明，因为第一个样本中红球所占的比例比第二个样本中的大。再强调一次，人们的直觉判断是由样本比例决定的，而基本不受样本量的影响，但后者在决定后验概率中起到了至关重要的作用

(Kahneman & Tversky, 1972)。而且，对后验概率的直觉判断远没有正确值那么极端。对证据产生的影响的低估已经多次出现在这类问题的判断中 (Edwards, 1968; Slovic & Lichtenstein, 1971)。这一现象被称为“保守主义”。

对概率的误解

人们总是希望由随机过程产生的一系列事件可以表示这一过程的基本特征，即使这一系列事件出现的时间很短。例如掷硬币，人们认为序列“正反正反反正”比看上去似乎并非随机出现的序列“正正正反反反”出现的可能性更大一些，并且也比序列“正正正正反正”出现的可能性更

大，因为后者并不能体现掷硬币的合理性（Kahneman & Tversky, 1972）。因此，人们总希望随机过程的基本特征不仅可以在整个过程中得到体现，而且也可以在每个部分得到局部体现。但是，一个局部具有代表性的序列会系统性地偏离人们对概率的预期：因为它的变化太多，但运行次数又太少。另一个相信局部代表性的结果就是大家熟知的“赌徒谬误”（gambler's fallacy）。例如，如果人们看到赌博轮盘上在很长的一段时间里总是出现红牌，那么大多数人会错误地认为下一轮就会出现黑牌。这可能是因为黑牌的出现比另一张红牌的出现更能形成具有代表性的序列。概率通常被认为是一种自我矫正的过程，其中一个方向上的偏差会产生相反方向的偏差来恢复平衡。事实上，偏差并没有随着随机过程的展开而得到“纠正”，而仅仅是被稀释了。

对概率的误解并非仅限于那些没有经过专业训练的被试。一项针对经验丰富的学术型心理学家的研究（Tversky & Kahneman, 1971）表明，存在着一种可以被称为“小数定律”（law of small numbers）的统计直觉，也就是即使是小样本也能高度代表其所在的总体。这些科研工作者的回答反映了一种预期：一个关于总体的有效假设可以由一个样本的具有统计学意义的显著结果来表示，而很少考虑这个样本的大小。因此，研究人员过于相信小样本的结果，并且在很大程度上高估了这种研究结果的可复制性。在实际的研究过程中，这种偏差会导致所选择的样本量不足，以及对研究结果的过度诠释。

对可预测性不敏感

人们有时会面临一些需要进行数值预测的情况，如预测未来的股票价格、某个商品的需求量或者某一场足球比赛的结果。这一类的预测通常会依赖于代表性。例如，假设一个人被要求通过一家公司的相关描述预测这家公司未来的利润。如果对这家公司的描述非常利好，那么高利润似乎更符合这种描述的代表性；如果描述相对一般，那么中等利润似乎最有代表性。现在，假设对这家公司描述的利好程度不受该描述可信度的影响，也不受其在多大程度上能形成准确预测的影响。如果仅凭相关描述的利好程度来进行判断，那么人们在预测时就会对线索的可信度和预测的准确性不敏感。

这样的决策模式违背了基本的统计理论，因为预测的极值和范围是由可预测性的考虑因素所控制的。当可预测性为零时，在所有情况下都应做出相同的预测。例如，如果公司的相关描述并不能提供与盈利有关

的信息，那么所有的公司利润预测值就应该是一样的（如均值）。当然，如果可预测性非常高，预测值就会与真实值相匹配，因此预测的范围就会与现实结果的范围相等。总的来说，可预测性越高，预测值的范围就越广。

一些关于数值预测的研究已经表明，直觉的预测会违背这一原则，而且被试很少或者根本没有考虑到可预测性这一问题（Kahneman & Tversky, 1973）。在其中一项研究中，被试被要求阅读几个段落，每个段落都描述了某个实践课程中的师生表现。一些被试被要求以百分数来“评价”描述中的课程相较于某一特定总体的质量。其他的被试则被要求“预测”每对师生在完成实践课程5年之后的发展状况，也是以百分数的方式进行打分。被试在这两种条件下所做出的判断其实是一样的。也就是说，对于长远标准的预测（5年后是否成功）与对这些预测所基于的信息（实践课程的质量）的评价相一致。这些进行预测的被试当然知道：仅凭一次5年前的课程来预测教学能力（好坏 / 影响）是远远不够的，其可预测性非常有限。但是，他们的预测却与他们的评价一样极端。

效度错觉

正如我们所看到的，人们经常通过选择最具代表性的信息（如对一个人的描述）来预测结果（如职业）。他们对自己做出的判断的信心主要是基于代表性的程度（如选择后的结果和输入信息的匹配程度），而很少考虑甚至不考虑那些会影响预测准确度的限制性因素。因此，当人们看到一段极为符合图书管理员刻板印象的描述时，就会非常自信地判断这个人的职业是图书管理员，即使这些描述提供的信息非常有限，甚至是不可信的、过时的。由于预测结果与输入信息十分相符而产生的毫无根据的信心，被称作“效度错觉”（*illusion of validity*）。甚至在做判断的人已经知道限制预测准确度的因素存在的情况下，这种错觉仍会出现。我们经常可以观察到，那些进行遴选访谈的心理学家通常对自己的预测非常自信，即使他们知道有大量的文献表明遴选访谈存在非常多的错误。尽管遴选访谈一再被证明有不足之处，但心理学家仍然依赖这一方式进行选择，这充分说明了效度错觉的影响之大。

输入信息形式的内部一致性（如一系列测试分数），是人们在基于这些信息进行预测时保持信心的主要决定因素。因此，人们在预测一个第一学年成绩全部为B的学生的期末平均成绩时，要比预测成绩中有A

和C的学生更有信心。当输入变量高度相关和冗余时，我们会观察到高度一致的形式。因此，人们常常会对基于冗余输入变量的预测更自信。然而，相关统计的一个基本结论表明：假如输入变量有稳定的效度，当它们之间相互独立而非相关时，基于这些输入的预测可以达到更高的准确度。因此，输入信息之间的冗余虽然能增强预测信心，却会降低准确度，而人们往往对很可能会出现错误的预测更自信！

（Kahneman & Tversky, 1973）

对回归的误解

假设有一大群孩子都接受了两个等效版本的能力测试（我们可称之为测试1和测试2）。如果研究人员从测试1中挑选10个表现优异的孩子，让其参加测试2，会发现他们在测试2中的平均表现有些令人失望。与之相反，如果你是从测试1中挑选10个表现最差的孩子参加测试2，会发现他们在测试2中的总体表现要更好。将上述情形推广开来，我们可以假设有两个变量X和Y，它们有相同的分布。如果挑选一组人，他们在X变量上的均值偏离了 k 个单位，那么他们在Y变量上的均值的偏离程度一定小于 k 。这些观察结果说明了一种普遍存在的现象，即均值回归，这一现象由英国统计学家弗朗西斯·高尔顿（Francis Galton）在一百多年前首次提出。

在日常生活中，我们会遇到很多均值回归的例子。例如，父子的身高、夫妻的智商，以及一个人连续多次考试的成绩等。但是，人们并没有形成对于这一现象的正确直觉。首先，在很多必然会出现回归现象的领域，人们并不期待看到它。其次，当人们意识到了回归现象的出现时，通常会得出荒谬的因果解释（Kahneman & Tversky, 1973）。我们认为回归现象仍然是很难捉摸的，因为它与人们秉持的“结果应该在最大程度上代表着输入信息”这一信念无法兼容，所以结果变量的值应该和输入变量的值一样极端。

意识不到回归的重要性会带来致命的后果，正如我们在之前提到的各种现象（Kahneman & Tversky, 1973）。在一次有关飞行训练的讨论中，有经验的飞行教练发现：如果在学员顺利降落后对其进行表扬，他接下来的表现通常都会不好；而在学员艰难降落后对其进行严厉的批评，他下一次的表現则会有所提升。因此飞行教练得出了结论：口头奖励对于训练来说没有好处，而口头批评则是有益的。这与公认的心理學原理完全相悖。飞行教练的这一结论并不可靠，因为存在均值回归现

象。与其他重复性测试一样，即使飞行教练对学员首次的训练不作反应，后者在一次糟糕的表现后通常都会有更好的表现，而一次出色的表现之后往往也会出现表现变差的情况。仅因为飞行教练在学员表现好时表扬了他们，而在表现糟糕时批评了他们，飞行教练就得到了错误且有潜在危害的结论：惩罚比奖励更有效。

因此，不理解回归的意义会让一个人高估惩罚的效用而低估奖励的效用。在社会互动，如有意识的训练中，表现好时通常才会得到奖励，而表现差时往往会受到惩罚。因此，仅通过回归分析就可预测，表现最有可能在惩罚之后有所提升，而在奖励之后有所下降。人类的状况就是这样：在得到奖励之后表现下降，在受到惩罚之后表现提升，这些都只是偶然事件。人们通常没有意识到这种偶然性。事实上，回归在决定奖惩结果中的作用非常隐晦，似乎没有引起这一领域的学生的注意。

可得性

在某些情况下，人们会通过回忆起一类或一件事的难易程度来评估这类事件发生的频率或可能性。例如，一个人可能会通过回想自己熟悉的人突发心脏病的情况来评估中年人突发心脏病的风险。类似地，一个人在预测某种投资行为失败的概率时，可能会通过想象这一过程中遇到的各种困难来进行评估。这种判断启发式被称为“可得性”。可得性是评估频率或概率的有用线索，因为一般来说，与频率较低的类别相比，人们更容易回想起频率较高类别中的案例。然而，可得性也会受到频率和概率之外的其他因素的影响。因此，依赖可得性会产生预测偏差，下面我们来举例说明其中的一些偏差。

实例的可检索性产生的偏差

当一类事件的发生频率要根据其实例的可检索性来判断时，在发生频率相等的情况下，人们会认为其实例更容易被想到的类别，比其实例更不容易被想到的类别更大。研究人员通过一个简单的实验展示了这一效应的影响。被试听到了一份包含知名人士的名单，名单中既有男性也有女性。他们随后被要求对这份名单是否包含更多男性进行判断。每组被试听到的名单内容不同。在一些名单中出现的男性比女性更知名，而在另一些名单中，情况恰好相反。但每一组被试都错误地认为由更知名的人物组成的类别中包含的男性更多（Tversky & Kahneman, 1973）。

除了熟悉度以外，还有其他因素会影响实例的可检索性，如显著性。例如，看到房子着火对此类事故的主观概率的影响，可能比在当地报纸上读到火灾新闻的影响更大。此外，人们会认为最近发生的事件可能比较早的事件更容易发生。当人们看到路边有辆车发生了侧翻，那么对于发生交通事故的主观概率会暂时提高，这可能是一种普遍存在的现象。

搜索集的效度带来的偏差

从英语文本中随机抽取一个由3个以及3个以上字母构成的单词。这个单词更可能以r开头，还是r更有可能为其第3个字母？人们的解决方式通常是先回忆以r开头的单词（如road）和第3个字母是r的单词（如car），然后根据能轻松回忆起这两种单词的数量进行判断。由于根据第1个字母检索单词比根据第3个字母检索单词容易得多，所以大多数人会做出判断：以某个辅音字母开头的单词比该辅音字母出现在第3个位置的单词更多。即使事实上这些辅音字母（如r或k）更常出现在第3个位置而不是开头，人们依然会做出同样的判断（Tversky & Kahneman, 1973）。

不同的任务会引发不同的搜索集。例如，假设有人要求你对抽象词（如思想、爱）和具体词（如门、水）出现在书面用语中的频率进行判断。回答这个问题的一种自然方法就是搜索该词可能出现的上下文。我们似乎更容易想到一个包含抽象词而非具体词（如门）的语境，如爱情故事中的“爱”。如果词语出现的频率是根据其出现的语境的可得性来判断的，那么人们就会认为抽象词多于具体词。最近的一项研究发现了这种偏差，该研究表明：人们认为抽象词出现的频率比具体词要高得多，虽然事实上两者具有相同的客观频率（Galbraith & Underwood, 1973）。人们也认为抽象词通常比具体词出现的语境范围更广。

可想象性偏差

有时候，人们必须评估脑海中没有存储相应实例的类别的频率，此时可以根据一些给定的规则进行评估。在这种情境中，人们通常会构造很多实例，并根据构造相关实例的难易程度进行频率或概率的估计。但是，构造实例的难易程度并不总能反映它们实际发生的频率，而且这种评估模型很容易出现偏差。为了对其进行说明，我们可以想象一个由10人构成的小组，该小组有一个由k个人组成的委员会，且 $2 \leq k \leq 8$ 。那么，

这个委员会会有多少种不同的组成方式？这个问题的正确答案是由二项式

$$\binom{10}{k}$$

系数给出的，当 $k=5$ 时为最大值252。显然，含有 k 个成员的委员会数量与含有 $(10-k)$ 个成员的委员会数量相等，因为如果其中的 k 个成员组成了一个委员会，那么唯一剩下的 $(10-k)$ 个成员自然而然也可以组成一个委员会。

不用计算就能回答这个问题的一种方法是，在心里构造一个由 k 个成员组成的委员会，并根据想到的案例的难易程度来评估其数量。由较少成员组成的委员会（如2个人），比由更多成员组成的委员会（如8个人）更容易被想到。构建这个委员会最简单的方法，就是将他们划分为不相交的多个小组。很明显，建立起5个由2个成员组成的委员会非常容易，但由8个成员组成的委员会最多只能有1个。因此，如果频率是通过可想象性或可得性来评估的话，小规模委员会显然比大规模委员会出现的次数更多，这与完全对称的钟形曲线（正态分布）相反。事实上，当被试被要求估计不同规模委员会的数量时，他们的估计是一个与委员会规模相关的单调递减函数（Tversky & Kahneman, 1973）。例如，2人委员会数量估计的中位数为70，8人委员会数量估计的中位数为20（两种情况的正确答案都是45）。

在现实生活中，可想象性在概率评估过程中起着重要的作用。例

如，探险活动存在的风险是通过想象无法应对的偶然事件来评估的。如果用这种方法描绘探险中遇到各种危险的概率，你就会发现探险会显得极其危险，但事实上，你能快速想到灾难事件并不能准确反映它实际发生的可能性。相反，如果一些可能的危险很难被想到或者根本想不到，那么该活动面临的风险就会被严重低估。

错觉相关

心理学家洛伦·查普曼（Loren Chapman）和琼·查普曼（Jean Chapman）描述了我们在判断两个事件同时发生的频率时的一个有趣偏差（Chapman & Chapman, 1967）。他们向没有统计知识的被试展示了几个假想的精神疾病患者的信息。每个精神疾病患者的数据都包含了一份临床诊断报告和一幅患者的画像。然后，研究人员要求被试判断每次的诊断结果（如偏执或多疑）中包含画像中各种特征（如奇特的眼睛）的频率。被试明显高估了自然联想事件同时发生的频率，如多疑和奇特的眼睛。这种效应被称为“错觉相关”（illusory correlation）。在对所接触数据的错误判断中，缺乏经验的被试“重新发现”了画人测验（the draw-a-person test）中存在的许多常见但毫无根据的临床知识。错觉相关效应对相互矛盾的数据具有极强的抵抗性，甚至当症状和诊断结果实际上负相关时，它也会令判断者难以发现其中存在的真实关系。

可得性为错觉相关效应提供了一种自然的解释。判断两个事件同时发生的频率可以依据它们之间存在的关联性的强度。当关联性很强时，人们可能会得出这样的结论：这些事件经常是成对出现的。因此，强关联体通常被判断为同时发生。根据这一观点，被试之所以会在多疑与画像中奇特的眼睛之间做出错觉相关，是因为与身体的其他部位相比，多疑更容易与眼睛相关联。

人生经验告诉我们，一般情况下，相比于低频率类别中的实例，人们更容易想到高频率类别中的实例；那些发生概率更高的事件也比不可能发生的事件更容易被想到；而事件之间的关联性在事件经常同时发生时会得到加强。因此，在估计某一类别的数量、事件发生的可能性或多个事件同时发生的频率时，人类有一套自己的评估程序，同时还有检索、构造或联想等心理操作。然而，正如前文案例所示，这种有价值的评估程序会受到系统性误差的影响。

调整和锚定

在很多情况下，人们会从一个初始值开始进行估算，然后通过调整得到最终答案。初始值或起始点，可以在问题描述中得到，也可以通过部分演算得到。无论初始值的来源是什么，调整通常都是不够的

（Slovic & Lichtenstein, 1971）。也就是说，不同的初始值产生了不同的估计值，这些估计值会偏向于初始值。我们称这种现象为“锚定”。

调整不足

在一个锚定效应的例子中，研究人员要求被试用百分比形式估计各类事物的概率，如联合国中非洲国家占比情况。每个问题的起始值为0~100，由被试面前转动的轮盘所决定。被试需要判断给定的（任意）初始值过高或过低，然后通过向上或向下调整来给出他们的最终估计值。对于每个问题，不同组得到的初始值不同。这些随机获得的初始值对人们的估计值有着显著影响。例如，在初始值为10%和65%的两个组中，他们对于非洲国家在联合国中所占比例的估计中位数分别为25%和45%。给予被试更多的奖励，也无法降低锚定效应。

锚定效应不仅发生在初始值给定的情况中，而且在被试根据一些不完整计算的结果进行估计时，也会产生锚定效应。对直观数值估计的一项研究说明了这种效应。两组高中生被要求在5秒内对黑板上的一组数字的乘积进行估算。其中一组需要估算 $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$ 的值，另一组需要估算 $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$ 的值。为了在5秒内快速回答这些问题，人们可能会通过推定和调整来进行计算和估计。因为调整通常是不足的，所以这个过程应该会导致被试得出过低的估算结果。此外，由于按降序排列的那组数字前几步相乘（从左到右）的结果，比按升序排列的那组数字前几步计算得到的结果更大，因此人们对前一组数字的估算结果应该比后一组数字更大。这两项预测都得到了证实。升序组估算结果的中位数为512，而降序组估算结果的中位数为2 250。正确答案为40 320。

合取与析取事件评估中的偏差

在最近的一项研究中，被试可以对几类事件中的一类进行投注（Bar-Hillel, 1973）。该研究使用了3类事件：（1）简单事件，如从一个红球和白球各占50%的袋子里取出红球；（2）合取事件，如从一个红球占90%白球占10%的袋子里，连续7次取出红球（每次取出球后再放回袋子里）；（3）析取事件，如从一个红球占10%白球占90%的袋子里，连续7次取球至少有一次取出红球（每次取出球后再放回袋子

里)。在这个问题上，绝大多数被试都倾向于在合取事件上投注（概率为0.48），而不是概率为0.50的简单事件。被试也更愿意投注简单事件而非析取事件，虽然后者的概率为0.52。也就是说，大多数被试在两两比较中都选择了概率更小的选项。这种选择模式说明了一种普遍现象。针对投注和概率判断的研究表明，人们往往会高估合取事件的概率

（Cohen, Chesnick & Haran, 1972）而低估析取事件的概率。这些偏差很容易被解释为锚定效应。基本事件的初始概率（如在任何一个阶段的成功）都为合取事件和析取事件概率的评估提供了一个天然的起点。因为从起点进行调整通常是不足的，在这两种情况下，最终评估都与基本事件的概率保持接近。请注意，合取事件的总体概率低于各个基本事件的概率，而析取事件的总体概率高于每个基本事件的概率。由于锚定效应，合取事件的总体概率会被高估，而析取事件的总体概率则会被低估。

评估复合事件时的偏差在制订规划的情境下尤为明显。成功地完成一项事业（如新产品的开发）通常有一个合取性的特征：为了取得成功，一系列事件中的每一步都要获得成功。即使这些事件中的每个子事件发生的概率都很高，如果子事件的数量巨大，获得整体成功的概率也会很低。人们往往会高估合取事件的概率，这会导致在没有依据的情况下做出过于乐观的评估，如对一项计划成功的可能性或一项工程准时完工的可能性的盲目乐观。相反，析取结构通常会用于风险评估。对于复杂系统（如核反应堆或人体）来说，任何基本组件的失效都会导致机能故障。即使每个组件失效的概率都微乎其微，如果涉及大量的组件，整个系统失效的概率也会很高。由于存在锚定效应，人们会倾向于低估复杂系统中的故障概率。因此，锚定偏差的方向有时可以从事件的结构中推断出来。合取事件的链状结构会导致高估，而析取事件的漏斗状结构则会导致低估。

主观概率分布评估中的锚定

人们出于很多目的（如后验概率计算、决策理论分析），需要以概率分布的形式来表达其对某个量（如某一天的道琼斯指数）的信心。这样的分布通常是以这样的方式构建的：让人们选择一个数值，让它与能代表自己的主观概率分布的具体百分位数相对应。例如，被试被要求选择一个数字 X_{90} ，来表示他认为这个数字高于道琼斯指数的主观概率为0.90。也就是说，他选择 X_{90} 即愿意以9:1的赔率在道琼斯指数不超过 X_{90} 时进行赔付。道琼斯指数的主观概率分布可以由这样几个对应着不同判

断的百分位数构造而成，如 X_{10} 、 X_{25} 、 X_{75} 、 X_{99} 等。

通过收集很多不同值的主观概率分布，可以测试人们在判断中进行的适当校准。在某一系列问题中，如果评估量的真实值的 $\Pi\%$ 恰好低于被试设定的 X_{Π} 值，那么这个判断就完成了适当的校准。例如，有1%的量的真实值应低于 X_{01} ，还有1%的量的真实值高于 X_{99} 。因此在98%的问题上，真实值的置信区间为 $X_{01} \sim X_{99}$ 。

研究人员已经从大量的判断中得到了许多量的概率分布（Alpert & Raiffa, 1969; Staeël von Holstein, 1971; Winkler, 1967）。这些分布表明其与合适的校准值之间存在很大的系统性偏差。对于大多数研究，在30%的问题中评估量的实际值小于 X_{01} 或大于 X_{99} 。也就是说，被试设定的置信区间过于狭窄，其反映的确定性超过了基于相应知识得出的合理范围。对于没有统计学知识和非常复杂的对象来说，这种偏差是很常见的，而且即使引入可以刺激外部校准的评分规则，这种偏差也是无法消除的。这种效应可以部分归因于锚定。例如，若要选择 X_{90} 作为道琼斯指数的值，首先人们会很自然地先想到自己对道琼斯指数的最佳估值，然后上调这个值。如果像大多数情况一样，调整不够充分，那么 X_{90} 将不够极端。类似的锚定效应也会发生在 X_{10} 的选择中，人们可能会通过向下调整最佳估值来得到这个值。因此，介于 X_{10} 和 X_{90} 之间的置信区间将会变得非常窄，且评估的概率分布会过于集中。为了支持这一解释，可以证明主观概率是在最佳估值无法作为锚的程序中被系统性改变的。

一个给定量的主观概率分布（如道琼斯指数的均值）可以通过两种不同的方法得到：（1）要求被试选择与其概率分布中某一特定百分位数相对应的道琼斯指数的值；（2）要求被试评估道琼斯指数超过某一具体值的概率。这两种方法在形式上是等价的，因此应该产生相同的分布。然而，它们却从不同的锚定点衍生出了不同的调整方式。在第一种方法中，自然起点是对某一量的最佳估值。而在第二种方法中，被试可能是以问题中设定的值为锚定点的。此外，他的锚定点有可能是正反概率相等，也就是似然估计的自然起点。无论哪种情况，第二种方法产生的极端概率都小于第一种方法。

为了对比这两种方法，我们向一组被试呈现了24个量（如从新德里到东京的空中距离），这些被试需要对每个问题进行 X_{10} 或 X_{90} 的评估。

另一组被试看到的则是第一组被试对24个量估计的中位数。他们被要求评估每个给定值超过其对应真实值的概率。在没有任何偏差的情况下，第二组应该重现第一组给定的概率，即9:1。不过，如果以平均概率和设定值作为锚定点，第二组的概率应该会不那么极端，也就是接近1:1。事实上，这组被试给出的所有问题的概率中位数为3:1。当对两组被试的判断进行外部校准测试时，我们发现第一组被试过于极端，而第二组被试则过于保守。

讨论

本章讨论了由于依赖判断启发式而产生的认知偏差。这些偏差不能归因于动机效应，如一厢情愿的想法或由于奖励和处罚而产生的判断扭曲。事实上，尽管被试被鼓励做出准确的回答，并且在回答正确时能获得奖励，但我们还是观察到了前文中出现的那几种严重的判断错误（Kahneman & Tversky, 1972; Tversky & Kahneman, 1973）。

对启发式的依赖和偏差的普遍性并不只发生在非专业人士身上。当经验丰富的研究人员进行直觉思考时也很容易产生同样的偏差。例如，我们依然可以在受过统计学训练的人身上观察到，他们在进行直觉判断时，在未充分考虑先验概率的情况下，总倾向于做出最拟合数据的预测（Kahneman & Tversky, 1973; Tversky & Kahneman, 1971）。虽然复杂的统计方法可以避免基本错误（如赌徒谬误），但他们的直觉判断也很容易在更复杂和不怎么一目了然的问题中产生类似的谬误。

即使偶尔会在预测和估计中导致错误，人们依然保留了有用的启发式，如代表性启发式和可得性启发式，这并不奇怪。令人惊讶的可能是，人们未能从长期的生活经验中推断出诸如均值回归或样本大小对抽样变异性的影响等基本的统计规则。尽管在日常生活中，每个人都遇到过无数个可以用来说明这些规律的案例，但很少有人能自行发现抽样和回归的原理。统计原理不是能从每天的经验中学得到的，因为相关的案例没有被适当地归纳总结。

缺乏适当的归纳总结也解释了为什么人们通常无法发现自己在概率判断中存在的偏差。可以理解的是，人们可以通过记录自己在分配了相同概率的事件中实际发生事件的总比例，来了解其判断是否得到了外部校准。然而，根据判断概率对事件进行分组是不自然的。例如，若没有这样的分组，人们不可能发现在自己预测的发生概率高达0.9及以上的

事件中，只有50%的事件会真的发生。

认知偏差的实证分析对于概率判断的理论和应用都具有重要意义。现代决策理论认为，主观概率是对一个理想化个体的意见进行量化的指标（Savage, 1954; de Finetti, 1968）。具体来说，针对某一具体事件的主观概率，是由个体能接受的关于这一事件的一系列赌注定义的。如果人们在投注中做出的选择符合某些原则（如理论定律），那么就可以为其推导出一个内在一致或连贯的主观概率测量方式。我们说推导出的概率是主观的，其含义是它允许不同的人对同一事件做出不同的概率判断。这种方法的主要贡献在于它提供了一种非常严格的对概率的主观理解，而这种理解适用于特定事件，并嵌入理性决策的一般理论之中。

需要指出的是，虽然主观概率有时候可以通过投注行为的偏好推断出来，但它们通常不会以这种方式形成。一个人投注于A队而不是B队，是因为他认为A队更有可能获胜；但他并没有从自己的投注偏好中推断出这一信念。因此，在现实中，主观概率决定了投注行为的偏好，而不是像理性决策的理论定律那样源于后者（Savage, 1954）。

概率固有的主观属性使许多学生相信，一致性或内在连贯性是评估概率的唯一有效标准。从主观概率形式理论的角度来看，任何一组内部一致的概率判断都是一样好。这个标准并不完全令人满意，因为一组内在一致的主观概率判断可能与个人持有的其他信念不相符。想象这样一个人，他对掷硬币游戏所有可能出现的结果的主观概率判断存在赌徒谬误。也就是说，随着前面连续出现正面的次数的增加，他对掷硬币中出现反面的概率估计也会增加。根据形式理论的标准，这个人的判断是内在一致的，因此可以被视作充分的主观概率判断。然而，这些概率判断显然忽略了硬币自身没有记忆，因此它无法自发形成一定的顺序这一事实。要使判断概率被认为是充分的或合理的，内部一致性是不够的，这些判断必须与个体持有的整个信念系统相一致。很可惜，我们没有一个简单的正规程序用来评估一系列的概率判断与判断者的整个信念系统是否兼容。尽管内部一致性更容易达到和被评估，但理性的判断者仍然会争取这种兼容性。具体来说，他会设法使自己的概率判断符合自身这3方面的知识：（1）主题；（2）概率定律；（3）自己拥有的判断启发式和偏差。

本章内容最初发表在1974年9月27日的《科学》杂志上，只做了少量修改。版权归美国科学促进协会所有，经许可出版。

参考文献

- Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment* (pp. 17–52). New York, NY: Wiley.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237–251.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430–454.
- Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance*, 6, 649–744.
- Tversky, A., & Kahneman, D. (1971). The belief in the law of small numbers. *Psychological Bulletin*, 76, 105–110.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207–232.
- Galbraith, R. C., & Underwood, B. J. (1973). Perceived frequency of concrete and abstract words. *Memory & Cognition*, 1, 56–60.
- Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology*, 73, 193–204; Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, 74, 271–280.
- Bar-Hillel, M. (1973). Compounding subjective probabilities. *Organizational Behavior and Human Performance*, 9, 396–406.
- Cohen, J., Chesnick, E. I., & Haran, D. (1972). A confirmation of the inertial- ψ effect in sequential choice and decision. *British Journal of Psychology*, 63, 41–46.

Alpert, M., & Raiffa, H. (1969). A report on the training of probability assessors. Unpublished manuscript, Harvard University.

Staeël von Holstein, C. (1971). Two techniques for assessment of subjective probability distributions—An experimental study. *Acta Psychologica*, 35, 478–494.

Winkler, R. L. (1967). The assessment of prior distributions in Bayesian analysis. *Journal of the American Statistical Association*, 62, 776–800.

Savage, L. J. (1954). *The foundations of statistics*. New York, NY: Wiley.

de Finetti, B. (1968). Probability: Interpretation. In D. L. Sills (Ed.), *International encyclopedia of the social sciences* (vol. 13) (pp. 496–504). New York, NY: Macmillan.

第2章 外延性推理与直觉推理：概率判断中的合取谬误

阿莫斯·特沃斯基
丹尼尔·卡尼曼

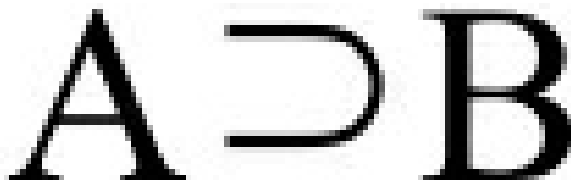
不确定性是人类生存环境中不可避免的一个方面。很多重要的选择都必须基于人们对这些不确定性事件所秉持的信念——嫌疑人的无罪推定、选举的结果、美元的未来价值、手术的成功率，甚至是一个朋友的反应。由于我们通常没有足够的形式模型（**formal model**）来计算这些事件的概率，因此直觉判断通常是评估这些不确定性的唯一方法。

关于人们和专家如何评估这些不确定性事件的概率这一问题，研究人员在最近几十年里已经开展了大量研究（Einhorn & Hogarth, 1981; Kahneman, Slovic & Tversky, 1982; Nisbett & Ross, 1980）。在这些研究中，很多都将直觉推论和概率判断与统计规则和概率定律进行了比较。进行判断的学生会将概率计算作为比较的标准，正如进行感知研究的学生会将物体的感知大小和物体的实际大小进行比较一样。但是，事件的“正确”概率与物体的实际大小不同，前者不容易确定。因为人们有着不同的知识体系或不同的信念，所以会对同样的事件做出不同的概率判断，没有一个对所有人都适用的确定值。而且，对某一个人来说，他也无法总是做出正确的判断。在随机抽样范畴之外，概率论不能决定不确定性事件的概率——它仅能对这些不确定性事件之间的关系加以界定。例如，如果A比B更有可能发生，那么A的补集出现的概率就低于B的补集。

概率定律源于外延性推理。一种概率测量方法是由一系列事件来定义的，而每个事件都会被理解为一系列可能性，如掷一对骰子可以得到10个点的3种方法。一个事件的概率等于它的所有不相容结果的概率之和。概率论传统上用于对重复的偶然进程进行分析，但是这种理论同样会被运用于一些本质上的单一事件，而这些事件的概率是无法还原到“理想”结果的相对频率的。飞机上坐在你身边的男性是未婚状态的概率等于他是单身的概率或他离异和丧偶的概率。可加性原则甚至适用于

没有频率解释和基本事件并非等概率事件的情形。

最简单和最基本的定性概率定律就是外延法则：如果A的外延包括



B的外延（如），那么A发生的概率就不会小于B发生的概率，即 $P(A) \geq P(B)$ 。因为与A & B合取相关的一系列概率是包含在与B有关的概率里的，而同样的原理也可以通过合取法则 $P(A \& B) \leq P(B)$ 来表示：一个合取事件的概率不可能比它所包含的单个事件的概率更高。无论A和B是否相互独立，该法则都成立，并且它对同一样本空间的任意概率分配都有效。同时，它不仅适用于标准概率计算，也适用于上限概率和下限概率（Dempster, 1967; Suppes, 1975）、信任函数（Shafer, 1976）、培根概率论（Baconian probability; Cohen, 1977）、理性信任（Kyburg, 1983）和可能性理论（Zadeh, 1978）等非标准模型。

与有关信任的形式理论相反，概率的直觉判断通常都不具有外延性。人们通常不会将日常事件发生的可能性编制成详尽的列表，或者通过整合各种基础事件来计算综合概率。相反，人们只是运用一系列启发式，如代表性启发式、可得性启发式（Kahneman et al., 1982）。我们对判断启发式的设想是建立在自然评估的基础上的，而自然评估通常是作为对事件的感知和对信息的理解的一部分而进行的。这种自然评估包括对相似性和代表性的计算、因果关系的推定，以及对关联事件和范例的评估。我们认为，在有具体任务的情况下，人们应任务的要求而做出这些评估，但在没有具体任务的情况下，人们仍会这样做。例如，只要提到“恐怖电影”人们就会想到一些恐怖电影，从而对它们做出可得性评估。同样，伍迪·艾伦（Woody Allen）的家人曾经希望他能成为一名牙医，这会引发人们将这个角色与刻板印象进行比较，以及对其代表性进行评估。可能是伍迪·艾伦的个性与人们对牙医的刻板印象不匹配，使得这个想法显得有点好笑。尽管这些评估都与频率或概率的估算无关，但当人们需要做出这样的判断时，它们很可能会产生主导作用。恐怖电影的可得性可用来回答这个问题：“在去年制作的电影中，恐怖电影占多大比例？”代表性可能会决定某个小男孩更可能成为一名演员还是一名牙医。

“判断启发式”这个术语指的是一种依靠自然评估进行估算或者预测的策略，而这种策略可能是有意也可能是无意的。启发式的一种表现就是对其他思路的相对忽视。例如，一个孩子和各种职业刻板印象的相似程度可能会严重影响人们对其未来职业选择的预测，从而使人们忽略了某种职业的基础比率等相关数据信息。因此，使用判断启发式会产生预测偏差。当启发式不适用时，自然评估会以其他方式影响我们的判断。首先，人们有时会误解自己的任务，所以难以将必要的判断与问题引发的自然评估区分开来。其次，自然评估可能会成为一个锚，使得所需的判断被同化了，而做出判断的人并不打算使用自然评估的结果来估算任务的结果。

上文中关于判断错误的讨论都集中在思维策略和对任务的误解上。而当前的处理需要特别关注锚定和同化机制，因为锚定和同化是在你意识不到的情况下发生的。有关知觉的一个案例可以很好地说明这一点：如果在一幅三维场景的图片中，两个物体的尺寸相同，那么较远的物体不仅看起来“确实”更大，而且在图中它的尺寸也更大一些。虽然观察者不可能将两者混淆或用前者衡量后者，但对于真实大小的自然计算显然影响了人们对图片（与现实相比不那么自然）中物体尺寸的判断。

对代表性和可得性的自然评估并不符合概率论的外延逻辑。具体来说，一个合取事件可能比它包含的子事件更具代表性，而某一具体类别中的案例可能比更复杂类型中的案例更容易被想到。下面这个例子可以证明这一点。当人们用60秒时间列出某种形式的包含7个字母的单词时，加拿大不列颠哥伦比亚大学的学生能说出的___ing形式的单词多于____n形式的单词，尽管后者显然包含了前者。在这两种条件下，学生们能想到的单词数量均值分别为6.4和2.9 ($t(44) = 4.70, p < 0.01$)。在这个针对可得性的测试中，记忆检索效率的提高足以抵消目标类型外延性的降低。

针对可得性启发式的实验说明：以ing结尾的单词和以“_n”结尾的单词在可得性方面的差异性，是通过频率方面的判断反映出来的（Tversky & Kahneman, 1973）。接下来的问题检验了这一预测。

在一篇4页长的英文小说中（总共约2 000个单词），请估计有多少个单词是___ing（以ing结尾的7字母单词）形式的？请在以下答案中圈出你的最佳估计：0, 1~2, 3~4, 5~7, 8~10, 11~15, 16+。

另一组被试读到的问题是估计有多少个____n形式的7字母单词。两组被试给出的估值中位数分别为13.4（ing组，共52人）和

4.7 (_n_组，共53人，中位数检验 $P<0.01$)，这违反了外延性原则。通过比较_____ly和_____l_两种形式的单词，也得到了同样的结果，两组估值的中位数分别为8.8和4.4。

这个案例说明了本章介绍的研究结构，该结构共由五部分组成。在第一部分中，我们提出了这样的问题：外延性的降低与可得性或代表性的增强有关，而且我们检验了频率和概率判断中的合取原则。在第二部分中，我们将讨论代表性启发式，并在个人知觉的情境中将其与合取原则进行对比。在第三部分中，我们将描述医学预测、运动赛事预测和投注中的合取谬误。在第四部分中，我们将对因果合取的概率判断进行研究，并描述在未来事件情境中出现的合取错误。在第五部分中，我们会给出能使被试避免出现合取谬误的操作方法。在最后一部分中，我们将主要讨论这些结果的意义。

具有代表性的合取事件

关于物体和事件类别的现代研究表明，信息通常会按照原型（prototype）和图式（schemata）等心理模式被记忆和加工（Mervis & Rosch, 1981; Rosch, 1978; Smith & Medin, 1981）。因此，根据一个事件在多大程度上能够代表它所匹配的心理模式来判断该事件的概率，是非常自然的，也是非常节省认知资源的行为（Kahneman & Tversky, 1972, 1973; Tversky & Kahneman, 1971, 1982）。由于本章中介绍的很多研究结果都可以归因为启发式，因此我们首先要对代表性的概念进行简短的分析，并对它在概率判断中的作用进行一些说明。

代表性是对一个样本与总体、实例与类别、行为与行为者之间一致程度的判断。更宽泛地说，这种一致是结果和模式之间的一致。模式，可以是一个人、一枚硬币或世界经济体系，而相应的结果可能是婚姻状况、硬币正反面出现的顺序，或者当前的黄金价格。针对代表性的研究可以通过直接询问某些人来进行，例如在掷硬币正反面出现的两种序列中，哪一种更具代表性？或者，在两种职业中，某种性格的人更符合哪种职业的刻板印象？这种关系和其他相近概念的不同之处在于，它具有明显的方向性。我们认为，如果说一个样本或多或少其所属总体或物种或者说一个物种（如知更鸟、企鹅）可以代表其上级类别（如鸟类），这是很正常的事情。但如果反过来说总体或多或少可以代表样本或类别或多或少可以代表实例，就显得太不准确了。

当可以用相同的术语描述模式和结果时，代表性就可简化为相似性。例如，一个样本和一个总体可以通过同样的属性（如集中趋势和变异）进行描述，如果样本的显著统计数据与对应的总体参数相匹配，那么这个样本就具有代表性。同样，一个人的性格与某一个社会群体的刻板印象相匹配，那么他在这个社会群体中就更具代表性。但是，代表性不总是等同于相似性，它也可以反映因果关系或相关信念（Chapman & Chapman, 1967; Jennings, Amabile & Ross, 1982; Nisbett & Ross, 1980）。我们之所以说某一具体行为（如自杀）代表着某个人，是因为我们认为行为人具有实施这一行为的倾向，而不是因为该行为与行为人有相似性。因此，如果在显著特征上相匹配或某种模式有产生某一结果的倾向，就可以说一个结果具有某种模式的代表性。

代表性会随频率产生共变：常见案例和高频事件比那些不寻常案例和罕见事件更具有代表性。具有代表性的夏天是炎热晴朗的，具有代表性的美国家庭有两个孩子，一个具有代表性的成年男性身高应该在178厘米左右。但是，在某些环境中，代表性是随实际频率和感知频率变化的。首先，一个非常具体的结果可能具有代表性，但可能并不常见。想象一个数字变量，如体重，某一个总体中所有人的身高数据应该呈一个单峰的正态分布。靠近分布模式狭窄区域内的个体，通常比靠近分布在尾部较宽区域里的个体更具代表性。例如，在斯坦福大学本科生这一样本中（ $N=105$ ），有68%的人认为该校“体重为124~125磅(2)”的女学生比体重“超过135磅”的女学生更具代表性。在另一个样本中（ $N=102$ ），78%的人认为在斯坦福女学生中，“超过135磅”的人要多于“体重为124~125磅”的人。也就是说，这些学生认为更窄的区间（124~125磅）比更宽的尾部区间（大于135磅）更具代表性，但出现的频率相对更低。

其次，如果某个属性具有很强的特征性，也就是说，如果某个类别中这一属性出现的频率相对于参照类别更高，那么对于该类别来说，这个属性就具有代表性。例如，有65%的被试（ $N=105$ ）认为对于一个好莱坞女明星来说，“离婚4次”比“投票给民主党”更具代表性。多次离婚对于好莱坞女明星来说更具特征性，一部分原因是人们对其离婚率高于其他女性群体的刻板印象。但是，在另一个样本中（ $N=102$ ），83%的人认为好莱坞女明星中“投票给民主党”的人数要多于“离婚4次”的人数。因此，与特征性更低的属性（投票给民主党）相比，特征性更高的属性被认为更具代表性，但频率更低。再者，某一类中不具代表性的个例完全可以具有更高级类别的代表性。例如，鸡在鸟类中不具有代表性，但可以是具有代表性的动物；水稻不是一种有代表性的蔬菜，但却

是有代表性的食物。

前文中的观察结果表明代表性是非外延性的：它不是由频率决定的，也不受类包含的约束。因此，对概率判断中合取原则进行的检验，使概率论的外延逻辑和心理学的代表性原则形成了鲜明对比。我们最初的合取原则研究是在1974年进行的，以职业和政治取向为目标属性，要求被试进行单独预测或通过简要的人格描述进行合取预测（Tversky & Kahneman, 1982）。在本节中描述的研究重复并拓展了之前的成果。我们设置了两个虚拟个体比尔和琳达的性格简述，并罗列了他们各自的职业和业余爱好。

- 比尔今年34岁。他很聪明，但缺乏想象力，有强迫症，而且总是缺乏活力。在学校的时候，他的数学成绩很优秀，但是社会研究和人文学科成绩较差。
- 比尔是一名医生，他的业余爱好是打扑克牌。
- 比尔是一名建筑师。
- 比尔是一名会计师。（A）
- 比尔的业余爱好是演奏爵士乐。（J）
- 比尔的业余爱好是冲浪。
- 比尔是一名记者。
- 比尔是一名会计师，并且他的业余爱好是演奏爵士乐。（A & J）
- 比尔的业余爱好是登山。
- 琳达今年31岁，单身，性格直率而且非常阳光。她主修哲学。在学生时代，她非常关心歧视和社会公平性问题，而且曾经参加过反核示威游行。
- 琳达是一名小学老师。
- 琳达在书店工作，并参加瑜伽课程。

- 琳达在女权运动中表现很活跃。（F）
- 琳达是一名精神病学社会工作者。
- 琳达是美国妇女选民联盟的成员。
- 琳达是一名银行出纳。（T）
- 琳达是一名保险销售员。
- 琳达是一名银行出纳，并且在女权运动中表现很活跃。（T & F）

正如读者已经猜到的那样：对比尔的描述被构造成具有会计师（A）的代表性，但不具有一个业余爱好为演奏爵士乐的人（J）的代表性。而对琳达的描述是一个具有代表性的女权运动者（F），但不是有代表性的银行出纳（T）。我们还预测，由属性合取定义的类别（比尔为A & J，琳达为T & F）的代表性评分要高于每个合取中代表性较低的组成部分（分别为J和T）的代表性评分。

加拿大不列颠哥伦比亚大学的88名本科生根据“比尔（琳达）与该类别的典型成员的相似性”，对与每种描述相关的8个陈述进行了排序。结果证实了我们的预测。符合预期排序（比尔为A>A&J>J；琳达为F>T&F>T）的被试的比例分别为87%和85%。这一发现既在意料之中也没有什么异议。如果与相似性和代表性一样，代表性既依赖于共同特征，也依赖于显著特征，那么应该通过添加共同特征来增强代表性（Tversky, 1977）。在一个面部图示上加上眉毛，会提高其与另一个有眉毛的面部图示的相似性（Gati & Tversky, 1982）。类似地，在银行出纳这一职业的基础上加入女权主义，就增加了琳达当下的活动与其个性的匹配度。更令人吃惊和难以接受的是，绝大多数的被试依然认为合取事件（A&J和T&F）比单独的代表性部分（J和T）更有可能发生。在下文中，我们会对这一现象进行描述和分析。

间接检验和隐性检验

我们可以将合取原则的实验检验分为3种：间接检验、直接-隐性检验（direct-subtle tests）和直接-外显检验（direct-transparent tests）。在间接检验中，一组被试评估合取事件的概率，而另一组被试评估单独的

子事件的概率。没有被试被要求对合取事件（如琳达既是一名银行出纳，也是一名女权主义者）与其子事件进行比较。在直接-隐性检验中，被试将对合取事件与其代表性较弱的子事件进行比较，但并不强调这些事件之间的包含关系。在直接-外显检验中，被试会评估或者比较合取事件及其子事件的概率，以强调它们之间的关系。

我们通过这3种实验程序对不同的假设进行了研究。间接程序检验的是概率判断是否符合合取原则；直接-隐性程序检验的是人们是否会应用偶然性对关键事件进行比较；而直接-外显程序检验的是人们在被要求对关键事件进行比较时，是否会遵循合取原则。这一系列的检验也体现了我们的研究进程，即一开始是在间接检验中，我们观察到人们违背了合取原则，然后我们越来越惊讶地发现，在几次直接-外显检验中，合取原则也都无一例外地遭到忽视。

有3组被试参加了这项主研究。在统计学意义上表现出“缺乏经验”的被试组由斯坦福大学和不列颠哥伦比亚大学的本科生组成，这些本科生没有任何概率或者统计方面的基础。“见多识广”被试组由心理学、教育学的一年级研究生和斯坦福大学医学院的学生组成，他们都修过一门或者多门统计学课程，因此对概率的基本概念非常熟悉。“资深”被试组由斯坦福大学商学院决策科学项目的博士生组成，他们已经学习过概率、统计和决策理论的很多高级课程。

在主研究中，被试首先会收到一个直接检验的问题（比尔或者琳达）。他们被要求对相关的8种描述（包括合取事件、组成合取事件的独立子事件及5个补充项）根据概率大小进行排序，1表示最有可能，8表示最不可能。然后，被试会收到以间接检验的形式呈现的剩余问题，而其中的选项列表包括了合取事件或其子事件。所有问题都使用了相同的5个补充项。

表2-1呈现了相对于补充项合取事件的排序均值 $R(A \& B)$ ，以及合取事件中代表性较低的子事件的排序均值 $R(B)$ 。在直接检验中，违背合取原则的百分比用 V 来表示。实验结果可以表述为：第一，在12次比较中，合取事件的排序均高于代表性较低的子事件；第二，直接检验和间接检验中的排序不存在一致性方面的差异；第三，在直接检验中，违背合取原则的总体发生率为88%，与代表性判断中相应模式的发生率基本一致；第四，无论是在间接检验还是直接检验中，被试在统计学方面的背景都没有对结果产生影响。

表2-1 针对可能性排序中合取原则的检验

被试类型	问题	直接检验			间接检验			总计 <i>N</i>
		<i>V</i>	<i>R</i> (A&B)	<i>R</i> (B)	<i>N</i>	<i>R</i> (A&B)	<i>R</i> (B)	
“缺乏经验” 的被试	比尔	92	2.5	4.5	94	2.3	4.5	88
	琳达	89	3.3	4.4	88	3.3	4.4	86
“见多识广” 的被试	比尔	86	2.6	4.5	56	2.4	4.2	56
	琳达	90	3.0	4.3	53	2.9	3.9	55
“资深”被试	比尔	83	2.6	4.7	32	2.5	4.6	32
	琳达	85	3.2	4.3	32	3.1	4.3	32

注：*V*为违背合取原则的百分比；*R* (A&B) 和 *R* (B) 是A&B事件和B事件的排序均值；*N*为直接检验中的被试人数；总计*N*为间接检验中的被试人数，他们被平均分成两组。

在直接比较B和A&B发生的可能性时，出现的对合取原则的违背现象被称为“合取谬误”。不同被试组在比较B和A&B发生的可能性时，间接比较时出现的对合取原则的违背现象被称为“合取错误”（conjunction errors）。表2-1中最令人惊讶的部分可能就是，间接检验和直接检验之间不存在任何差异。我们曾预计在间接检验中，人们会认为合取事件的概率高于其可能性较低的子事件，这与我们在代表性判断中观察到的现象一致。但是，我们也预计即使最缺乏经验的被试也能注意到一些属性会重复出现、单独出现或者和其他事件一起出现，因此他们应该会运用合取原则，并认为合取事件的概率低于它所包含的子事件。但这一预期并没有出现，不仅缺乏经验的本科生如此，就连统计学知识丰富的博士生也是如此。在直接检验和间接检验中，所有的被试显然是根据比尔（或者琳达）符合刻板印象的程度对陈述进行排序的。针对比尔和琳达的平均概率排序和代表性排序的相关系数分别为0.96和0.98。那么，在包含关系非常明显的情况下，合取原则还成立吗？我们摒弃了所有隐性部分，来使被试可以发现和鉴别目标事件之间的包含关系。

外显检验

在这一部分中，我们描述了一系列旨在引导被试遵循合取原则的极端操作。首先，我们向142名加拿大不列颠哥伦比亚大学的本科生提供了关于琳达的描述，然后让他们判断两个选项中哪一个更有可能是真的：

琳达是一名银行出纳。(T)

琳达是一名银行出纳，同时她在女权运动中表现很活跃。(T&F)

其中一半的被试看到的选项呈现顺序是相反的，但是这一操作没有对结果产生影响。总体来说，85%的被试认为T&F比T更有可能发生，这完全违背了合取原则。

这一发现在惊人，我们尝试从其他角度解释被试的回应。可能被试发现问题过于简单，认为不能仅从字面上对其进行解读，因而将具有包容性的T事件解读为“T&非F”，即“琳达是一名银行出纳且不是一名女权主义者”。当然，这样一来，我们观察到的判断就没有违背合取原则。为了检验这一解释是否成立，我们让一组新的被试(N=119)对T和T&F进行评估，评分范围为1~9分，1分为极不可能，9分为极有可能。因为即使在一个事件包含另一个事件的情况下，评估概率也是符合逻辑的，因此被试没有理由将T理解为“T&非F”。然而，被试的反应还是跟之前一样。T的概率均值为3.5而T&F的概率均值为5.6，而且82%的被试对T&F的概率判断要高于T。

尽管被试并不能自发地运用合取原则，但他们也许能意识到这一原则的正确性。我们向另一组加拿大不列颠哥伦比亚大学的本科生展示了带有T和T&F标注的有关琳达的两个描述，并要求他们指出以下两个论点中哪一个更有说服力。

论点 1：琳达是一名银行出纳的概率高于她是一名具有女权主义属性的银行出纳的概率，因为每个具有女权主义属性的银行出纳都是银行出纳，但是有些女性银行出纳却并非女权主义者，而琳达可能就是其中的一员。

论点 2：琳达是一名具有女权主义属性的银行出纳的概率高于她是一名银行出纳的概率；因为与银行出纳相比，她更像一个活跃的女权主义者。

大多数被试(65%，N=58)选择了无效的相似性论点2，而不是有效的外延性论点1。因此，尽管我们有意引导被试用反思的态度看待问题，仍没有消除代表性启发式对人们的影响。

接下来，我们进一步通过将T事件表示为析取事件来说明其包含性。注意：合取原则同样可以表示为析取原则，即 $P(A \text{ or } B) \geq P(B)$ 。我们再一次引入了关于琳达的描述，以9分制量表进行概率判断，但将T事件替换为：

琳达是一名银行出纳，无论她是否在女权运动中表现活跃。(T*)

这一形式强调了T中包含了T&F。尽管它们的关系已经表达得很明确，但被试对于T&F的可能性判断为5.1，而对于T*的可能性判断为3.8 (t 检验 $p < 0.01$)。而且，57%的被试 ($N=75$) 依然出现了合取谬误，认为T&F的可能性高于T*，只有16%的被试持相反观点。

在对T&F和T*进行的直接比较中，这样的合取谬误非常惊人，因为“琳达是一名银行出纳，无论她是否在女权运动中表现活跃”的外延明显包含了“琳达是一名银行出纳，同时她在女权运动中表现活跃”的情况。很多被试显然无法从“无论是否”这一短语中得出外延性推论，而这一表述可能会让人联想到软弱的性格。这一解释得到了组间比较的证实，在该比较中，不同组的被试在评估共同的补充项“琳达是一名精神病学社会工作者”之后，通过9分制量表对T、T*和T&F进行评估。结果显示，T的均值为3.3，T*的均值为3.9，T&F的均值为4.5，3组被试的均值之间都有显著性差异。T和T*的表述在外延性上肯定是一样的，但是人们赋予了它们不同的概率。因为女权主义符合琳达的形象，仅仅是提到这个属性都会让T*比T变得更有可能是，而对女权主义符合琳达这一信念的执着使得合取事件T&F的概率反而更高！

可以成功地使人们减少出现合取谬误的方法是让他们进行投注：选择T还是T&F？我们让被试看到了关于琳达的描述，并附带如下说明：

如果你在投注中赢了，可以获得 10 美元，你会对下面哪一个进行投注？

在这个任务中，只有56%的人 ($N=60$) 违背了合取原则，虽然比例还是很高，但是与比较两个事件的概率时通常会得到的数值相比，已经大大降低了。我们推断：进行投注的情境会让人们更关注哪个选项能为自己带来回报，这使得一些被试发现在T上下注比在T&F上下注更有可能获胜。

这一部分研究中的被试都是来自加拿大不列颠哥伦比亚大学的本科生，他们没有任何统计学知识背景。那么统计学教育是否能消除这种谬误？为了回答这个问题，我们组织了来自美国加州大学伯克利分校和斯坦福大学的64名社会科学专业的研究生（他们都已经获得了一些统计学课程的学分），以评定量表的形式就琳达的相关问题的合取原则进行直接检验。在这一系列的研究中，T&F的评价均值（3.5）首次低于T的评价均值（3.8），只有36%的人犯了合取谬误。因此，掌握一定的统计学知识使得大部分被试在外显检验中都遵循了合取原则，尽管在这些有相关知识储备和“老练”的被试中，出现合取谬误的比例依然不低。

而在另一项研究中，我们通过积极和消极两个方面的解释对违背规范性原则的判断和偏好进行了区分（Kahneman & Tversky, 1982a）。积极解释会关注产生某一具体反应的因素；而消极解释则试图解释为什么没有做出正确的反应。比尔和琳达问题中的积极分析引用了代表性启发式。但是，外显检验中难以消除的合取谬误使我们对消极分析中特有的问题产生了兴趣：为什么那些接受了良好教育的、聪明且理性的人没有意识到合取原则在外显问题中的适用性？在实验完成后对被试进行的访谈和课堂讨论，给这一问题带来了启示。无论是缺乏经验的被试，还是有丰富统计知识的被试，基本上都能在直接-外显检验中发现目标事件的包含关系。但是，缺乏经验的被试与具有统计学知识背景的被试不同，前者无法认识到这种关系在概率评估中的重要性。此外，大多数缺乏经验的被试并没有试图为自己给出的回答进行辩解。正如一名被试在承认了合取原则的效度之后说道：“我以为你们只是在问我的看法。”

这些访谈和外显检验的结果都表明：缺乏经验的被试不会自发地将合取原则视为决定性的。他们的态度让我们想起了皮亚杰的守恒实验中儿童的反应。处于前运算阶段的儿童并不一定完全理解不了基于总量守恒的观点，或者有关总量守恒的典型预测（Bruner, 1966）。他们只是不知道，守恒的论点是有决定性的，而且应当推翻“较高容器比较矮容器能容纳更多水”之类的感性印象。同样，缺乏经验的被试总是将合取原则视为抽象的事物，但是当将其运用到琳达问题中时，他们就会强烈感觉到T&F判断比T判断更具代表性，从而忽略了合取原则。在这个情境中，这些成人被试表现得就像他们还没有发展到形式运算（formal operation）的阶段一样。要完全理解物理学、逻辑学或统计学原理，需要知道它们能在哪些条件下克服相互矛盾的论点，如一个容器中液体的高度或一个结果的代表性。识别规则中具有决定性的本质的能力，是区分守恒定律（conservation）研究中不同发展阶段的重要依据，因此也可

以区分目前这一系列研究中的被试的统计学知识背景的“深厚”程度。

更具代表性的合取

之前提到的研究表明，在人类感知和社会刻板印象领域，人们一再违背合取原则。合取原则在其他判断中是否能更好地发挥作用呢？如果我们把目标事件的不确定性归因于概率，而不是一定程度的无知，合取原则是否成立？相关领域的专业知识是否有助于克服这种合取谬误？经济激励是否可以使受访者意识到这种错误？下面的研究就是为回答这些问题而设计的。

医疗判断

在这个研究中，我们让执业医师对临床证据进行一些直觉上的预测。[\(3\)](#)我们选择研究医疗判断，是因为医生拥有专业的知识，而且直觉判断在医疗决策中往往起到了非常重要的作用。两组医生参与了这个研究。第一组由来自大波士顿区的37名内科医生组成，他们正在哈佛大学攻读研究生课程；第二组由在新英格兰医疗中心颇有声望的66名内科医生组成。我们向其提供了如下问题：

一名55岁的女性在进行了胆囊切除术10天后发生了肺动脉栓塞。

请按照患者所经历的情况，对下列情况发生的可能性进行排序，其中1表示最可能发生，6表示最不可能发生。当然，患者可能会经历不止一种情况。

呼吸困难和轻偏瘫（A & B）	晕厥和心动过速
小腿疼痛	轻偏瘫（B）
胸炎性疼痛	咯血

每个问题的症状都包括一个经我们的咨询医生判断，为该类型患者的非代表性症状，记为B，还包括B与另一个高代表性症状A的合取。在上文中提到的肺栓塞（肺部血栓）的例子中，呼吸困难（气短）是一个典型症状，而轻偏瘫（部分瘫痪）是非典型症状。首先每个参与者会以间接的形式收到3个（或两个）问题，其中包括B或A&B，但不同时包括两者；然后是两个（或3个）直接问题，如上表所示。研究设计是平衡的，即每个问题在每种形式中出现的次数是相等的。我们请一个由32名斯坦福大学的医生组成的独立小组根据“每个症状的临床代表性程度”来对其进行排序。

这种设计本质上与比尔和琳达的相关研究是一样的。两个实验的结果也是相似的。在5个问题中，概率的均值和代表性程度的平均评分之间的相关性都超过了0.95。对于这5个问题中的每一个问题，一个不太可能出现的症状和一个可能出现症状的合取概率，都大于那个不太可能出现的症状的概率。在直接检验和间接检验中，症状的排序都是一样的。A&B和B的整体排序均值：直接检验中为2.7和4.6；间接检验中为2.8和4.3。在直接检验中，违背合取原则的概率为73%~100%，平均为91%。显然，即便是非常资深的专家，也难免会受到代表性启发式的影响，而且也无法避免合取错误的产生。

我们是否可以将这些结果理解为对合取原则的一致违背，而不是归因于这些专家？本研究中使用的说明旨在避免人们将症状B理解为对相关事实的详尽描述，这意味着不存在症状A。我们要求被试按照“这些症状将在患者身上出现”的概率对症状进行排序。同时，我们也会提醒他们“患者可能会出现不止一种症状”。为了检验这些说明的效果，我们在问卷的最后加入了以下问题：

在评估患者经历描述的 X 症状的概率时，你是否假设：

X 是该患者出现的唯一症状？

X 是该患者出现的多种症状之一？

在被询问的62名内科医生中，有60名选择了第二个选项，没有接受一种可以证实明显违背合取原则的解释。

另外24名内科医生（大部分是斯坦福医院的住院医生）参加了小组讨论，在这个讨论中，他们会在同一问卷中看到自己的合取谬误。这些被试没有对自己的合取谬误进行辩解，尽管有些参考文献指出这些都是“临床经验的本质”。大多数被试为自己犯了一个基本的逻辑错误而感到惊讶和失望。因为合取谬误很容易被发现，因此犯了这种错误的人会有一种自己本可以做得更好的感觉。

预测温布尔登网球公开赛

上述研究遇到的关于患者愈后情况或个人职业的不确定性，通常会被归因于信息的不完整而不是被随机过程所操纵。最近，理查德·尼斯贝特、D. H. 克兰茨（D. H. Krantz）、C. 杰普森（C. Jepson）和Z. 昆达

(Z. Kunda) 对日常事件进行的归纳推理的研究表明，统计原理（如大数定律）通常会被运用到诸如体育赛事、博彩等领域，因为这些领域中含有随机因素（Nisbett, Krantz, Jepson & Kunda, 1983）。接下来的两个研究检验了合取原则在预测体育赛事和碰运气游戏时的应用，而在这些赛事和游戏中，整个过程的随机性非常突出。

1980年10月，我们通过在美国俄勒冈大学报纸上刊登广告的方式招募了93名被试来回答下述问题：

假设瑞典网球运动员比约恩·博格（Bjorn Borg）在1981年打进了温布尔登网球公开赛（简称温网）决赛。

请将以下结果按概率从高到低排序：

- A. 博格将赢得比赛（1.7）
- B. 博格将输掉第一局（2.7）
- C. 博格将输掉第一局，但赢得比赛（2.2）
- D. 博格将赢得第一局，但输掉比赛（3.5）

每种结果后面括号里的数值为其排序均值（1为最有可能，2为第二有可能，依此类推）。这些结果代表着博格的实力的不同水平，选项A代表的水平最高；选项C代表的是一个相对较低的水平，因为它表明博格在第一局中处于劣势；而选项B也代表了较低的水平，因为它只提到了这种劣势；但选项D在所有选项中代表的水平是最低的。

博格在1980年赢得自己的第5个温网冠军之后，他看起来实力非常强大。因此，我们假设人们会判断选项C比选项B更有可能发生，因为与选项B相比，选项C代表着博格发挥得更好，但这种判断违背了合取原则。而排序的平均值也证实了这个假设；72%的被试对选项C的排序优于选项B，在直接检验中违背了合取原则。

被试是否有可能以非外延性方式来理解目标事件，而这种非外延的方式可以证明被观察到的排序的合理性，或可以对其进行解释？众所周知，连接词（如且、或、如果）在通用语言中经常以背离其逻辑定义的方式被使用。也许被试是将合取关系（A且B）理解成了析取关系（A或B）、指代关系（A暗指B），或条件关系（如果B，那么A）。或者，事件B可能被理解成了B和非A的合取。为了研究这些可能性，我们向另一组来自斯坦福大学的56个毫无统计学知识背景的被试展示了相关网球

赛事的假设结果，并以输赢的方式对这些结果进行编码。例如，序列LWWLW表示在一场5局的比赛中，博格输掉了第一局和第四局（L），但其他的几局都赢了（W），而且最终赢得了比赛。对于每个序列，被试被要求检查原来的博格问卷中的4个目标事件并通过符号+或-标记出给定的序列是否与每个事件相符。

除了极少的例外情况，几乎所有的被试都是根据对目标事件的标准（外延）理解进行标记的。当两个因素都满足时（如LWWLW），序列被判定为与合取事件“博格将输掉第一局但赢得比赛”相一致，当有一个或两个因素无法满足时则被判定为不一致。最终，这些被试都没有将合取事件理解为一种指代关系、条件关系或析取关系。此外，LWWLW和LWLWL被判定为与“博格将输掉第一局”的这一包含事件相一致，这与包含事件B在其他语境中被视为“博格将输掉第一局且输掉比赛”的假设相矛盾。因此，序列的分类在表明目标事件的外延时几乎不存在什么歧义。具体来说，所有被归为B&A的例子也可以被归到B中，但根据合取原则中应该满足的条件，一些被归于B的序列可能与B&A不符。

出现合取谬误还可能是因为，被试不是根据证据E（琳达的个性）来评估假设B（如琳达是一名银行出纳）的概率 $P(B/E)$ 的，而是评估所提出的假设证据的逆概率 $P(E/B)$ 。因为 $P(E/A\&B)$ 可能远远超过 $P(E/B)$ ，因此基于这种解释，被试的判断可能是合理的。无论这种可能性在琳达的例子中多么可信，它都不适用于当前的研究，因为在已知比赛结局的条件下评估博格进入决赛的条件概率是没有意义的。

风险选择

如果合取谬误不能被目标事件的重新解释证实，那么是否可以通过一个非标准性概率概念来将其合理化？基于这种假设，代表性被视为概率的一种合理的非外延性解释，而不是一种容易出错的启发式。因此，合取谬误可能被认为是对“概率”一词的含义的误解。为了研究这个假设，我们在下述决策问题中对合取原则进行了检验。在这些问题中，人们有选择最有可能发生的事件的动机，尽管我们并没有提到“概率”这个词。

想象一枚普通的 6 面骰子，其中有 4 个绿色面和 2 个红色面。将骰子滚动 20 次，记录出现绿色面（G）和红色面（R）的顺序。你需要从一组 3 个序列中选择一个序列，如果你选择的序列与骰子连续掷出的结果一致，将赢得 25 美元。请给出你的选择。

1. RGRRR
2. GRGRRR
3. GRRRRR

请注意，从序列2中拿掉第一个G就变成了序列1。因此，按照合取原则，序列1肯定比序列2更有可能发生。而且，我们可以看到，这3种序列都不具代表性，因为它们包含的R比G更多。然而，序列2看起来更像是序列1的改进版，因为序列2中包含可能出现的颜色的比例更高了。一组由50名被试组成的实验组被要求根据序列所具有的代表性程度对其进行排序。其中，88%的被试将序列2排到了最高位，而将序列3排到了最低位。因此，序列2最具代表性，尽管它是由序列1主导的。

来自加拿大不列颠哥伦比亚大学和斯坦福大学的260名学生回答了这个问题。不同人群给出的结果没有显著的差异性，因此我们将数据进行了合并。这些被试被分成30~50人的小组，并在教室中进行实验。约一半的被试（ $N=125$ ）通过投注获得了真正的报酬，另一半的被试则进行的是假设性实验。在进行真实投注的被试中，有65%的人选择了序列2；而在假设组中，有62%的被试选择序列2。在两组中，都只有2%的被试选择了序列3。

为了进一步说明两种关键序列之间的关系，我们向一组由59名被试组成的实验组提出了一个（假设性）决策问题，并将序列2替换为RGRRRG。这个新序列依然更受欢迎，有63%的被试选择了这个新序列而不是序列1——RGRRR，尽管这两个序列的前5个因素一模一样。这些结果表明：被试会根据G和R的比例对每个序列进行编码，而且按照这两种序列中不同结果（R和G）的比例（ $1/5$ 和 $1/3$ ）与预期值 $2/3$ 之间的差异对序列进行排序。

根据这些结果，合取错误显然并不局限于对“概率”这个词的误解。即使这个词根本没有出现，甚至没有出现在有大额报酬的选项中，我们

的被试都会遵循代表性启发式。而且，这些结果也表明合取谬误不仅局限于对连接词“且”的理解，因为这个连接词在问题中根本没有出现。根据之前的定义，对连接词的检验是直接的，因为被试被要求对两个具有包含关系的事件进行比较。然而，对几名被试的非正式访谈表明这种检验是很巧妙的：只有一小部分被试注意到了序列1和序列2之间的包含关系。很显然，人们并不习惯于检验事件之间的包含关系，即使这些关系已经被明显地呈现出来了。

假设被试注意到了序列1和序列2之间的包含关系，他们是否能立刻意识到它的影响，并将其视为选择序列1的决定性因素？最初的选择问题（不包括序列3）被呈现给了88名来自斯坦福大学的被试。但我们没有要求被试选择自己愿意投注的序列，而是要求他们在以下两个论点中选择自己认为正确的论点。

论点 1：序列 1（RGRRR）比序列 2（GRGRRR）更有可能出现，因为序列 2 只是在序列 1 的开头加了一个 G。因此，每当序列 2 发生的时候，序列 1 肯定会发生。因此，你可以在第一局赢、第二局输，但你不可能在第二局赢、第一局输。

论点 2：序列 2（GRGRRR）比序列 1（RGRRR）更有可能出现，因为与序列 1 相比，序列 2 中 R 和 G 的比例更接近原本骰子 4 个绿色面和 2 个红色面的预期比例。

结果显示，大多数被试（76%）选择了有效的外延性论点，而不是反映代表性直觉的论点。我们可以回想一下琳达的例子，其中也有一个类似的论点，但在阻止人们出现合取谬误方面的效度要低得多。目前这种操作的成功可以归因于机会设置和投注游戏的结合，这是在通过强调投注的回报条件来促进外延性推理。

谬误与误解

我们将直接检验中违反合取原则的现象称为谬误。“谬误”这个词在这里被理解为一种心理学假设，而不是一个评估性词语。当做出判断的大多数人在获得适当的解释后倾向于接受以下命题时，这个判断就被恰当地贴上了谬误的标签：第一，他们犯的这个错误非同小可，在类似的问题中可能还会犯这种错误；第二，这种错误是观念性的，而不仅仅是

字面上或技术上的；第三，他们应该知道正确答案或者寻找正确答案的方法。或者，如果被试误解了问题或者实验人员错误地解释了答案，那么相同的判断可以被描述为交流失败。那些因为误解而犯错的被试很可能会拒绝上述命题，并且会声称其实自己知道正确答案（就像学生在考试后的惯常表现），如果真的犯了错误，也只是字面上或技术上的，而不是观念上的。

心理学分析应该有清晰的解释，并避免将真正的误解当作谬误。而且，也应该避免用临场的理解将判断错误合理化。但是，谬误和误解之间的分界线并不总是清晰的。在我们早期的研究中，大多数的被试会认为某种描述更像是针对体育教师而非学科教师的。严格地说，后者包括了前者，但人们可能认为这个问题里的“教师”并不包括体育教师，就像“动物”在某种意义上不包括昆虫一样。所以，对于外延性原则的明显违背到底是应该被描述成谬误还是误解，其实是不明确的。在目前的研究中，为了避免产生歧义，我们专门将关键事件定义为具有明确界限的类别集合，如银行出纳和女权主义者。这些被试在实验结束后的相关讨论中发表评论支持了我们的结论，即我们在直接检验中观察到的违背合取原则的行为都是真正的谬误，而不仅仅是误解。

因果合取

前文中讨论的问题包括3个方面：一个因果模型M（琳达的人格）；一个基本的目标事件B，它是模型M的非代表性例子（琳达是一名银行出纳）；还有一个附加事件A，这个事件是模型M的高代表性例子（琳达是个女权主义者）。在这些问题中，模型M与A正相关，而与B负相关。这个结构被称为 $M \rightarrow A$ 范式，如图2-1的左侧部分所示。我们发现，当与琳达的人格相关的描述被省略，她被认为只是“一位31岁的女士”时，基本上所有的被试都会遵循合取原则，认为合取事件（是名银行出纳且是活跃的女权主义者）的概率比其子事件的概率更小。因此，初始问题中的合取错误可以被归因为M和A之间的关系，而不是A和B之间的关系。

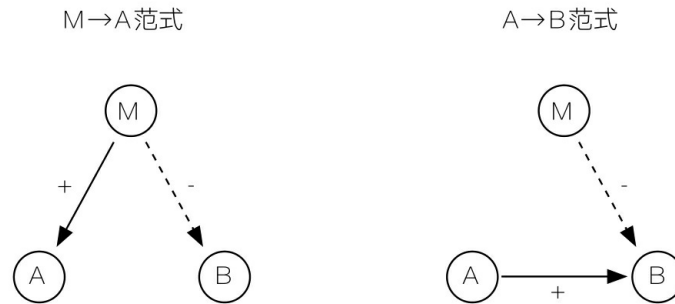


图2-1 因果关系的合取

注：实线箭头和虚线箭头分别表示模型M、基本目标事件B和附加目标事件A之间存在的强正性相关和强负性相关关系。

尽管银行出纳和女权主义者的刻板印象之间只是稍微不相容，但合取谬误在琳达问题中依然很普遍。当合取事件中的子事件之间高度不相容时，合取错误出现的概率会大大降低。例如，合取事件“比尔对音乐毫无兴趣，但他的业余爱好是演奏爵士乐”就会被认为比其子事件更不可能发生（也更不具代表性），尽管“对音乐毫无兴趣”会被认为是比尔的一个可能（具有代表性的）特征。因此，两种属性的不相容会降低它们合取的概率。

相容性对合取事件概率评估的影响不只限于子事件之间接近相互矛盾的关系。例如，一个学生的数学和物理两门课的成绩都名列前茅或都处于下游，比其中一门处于上游而另一门处于下游的情况更具代表性（也更有可能发生）。这就意味着一个合取事件的判断概率（或代表性）不能像其子事件的尺度值的函数（如乘积、和、最小值、加权平均等）那样得以计算。这一结论会排除掉一大批忽略了合取事件中子事件之间关系的形式模型。而这种合取概念模型的可行性也引发了激烈的辩论（Jones, 1982; Osherson & Smith, 1981, 1982; Zadeh, 1982; Lakoff, 1982）。

上文中的讨论提出了一种新的形式结构，即 $A \rightarrow B$ 范式，如图2-1右侧部分所示。合取错误发生在 $A \rightarrow B$ 范式中，因为A和B之间存在直接联系，哪怕附加事件A并不是模型M的代表性事件。接下来，我们将对一些问题进行研究，在这些问题中，附加事件A为事件B的发生提供了合理的原因或动机。我们假设因果关系的强度也会使合取事件的概率判断产生偏差（Beyth-Marom, 1981），前文已证实它会使条件概率判断产生偏差（Tversky & Kahneman, 1980）。正如人格和社会刻板印象会自然引发对其相似性的评价，一种结果和它可能的原因的想法也会激发人

们对其中的因果关系进行评价（Ajzen, 1977）。对喜好的自然评估将会使概率估计产生偏差。

为了说明 $A \rightarrow B$ 范式中存在的偏差，我们可以思考以下问题，这也是我们向115个加拿大不列颠哥伦比亚大学和斯坦福大学本科生提供的问题：

研究人员对不列颠哥伦比亚省（位于加拿大）所有年龄段、从事各类职业的成年男性进行了具有代表性的健康抽样调查。样本包括 F 先生。他是从被试列表中随机抽取出来的。

以下哪种描述更有可能发生（单选）：

F 先生经历过一次或者多次心脏病发作；

F 先生经历过一次或者多次心脏病发作；而且他已经超过了 55 岁。

在这些没有统计学知识背景的被试中，有58%的人在这个看似简单明了的问题中犯了合取错误。为了检验这种错误是由高龄和心脏病发作之间的因果（或者相关）关联引起的，而不是由每个子事件概率的权重均值引起的，我们在不改变其边缘概率的前提下通过拆解目标事件来消除这种关联性。

研究人员对不列颠哥伦比亚省所有年龄段和职业男性进行了具有代表性的健康抽样调查。样本中包括 F 先生和 G 先生。他们是从被试列表中被随机抽取出来的。

以下哪种描述更有可能发生（单选）：

F 先生经历过一次或者多次心脏病发作；

F 先生经历过一次或者多次心脏病发作；而 G 先生已经超过了 55 岁。

将关键属性分配给两个独立的个体，通过使事件（有条件地）保持独立性可以降低 $A \rightarrow B$ 的联结带来的影响。因此，合取错误的发生率下降到了29%（ $N=90$ ）。

$A \rightarrow B$ 范式可能会产生双重的合取错误，在这种情况下， $A \& B$ 会被

认为比A或B更有可能发生，正如下一个问题所示。

彼得是一名大学三年级的学生，他正在为地区运动会 1 英里赛跑进行训练。在本赛季前期的比赛中，彼得获得的最好成绩是 4 分 06 秒。请将以下结果按照最有可能到最不可能的顺序进行排序：

彼得会在 4 分 06 秒以内跑完全程。

彼得会在 4 分钟以内跑完全程。

彼得后会在 1 分 55 秒以内跑完后半程。

彼得后会在 1 分 55 秒以内跑完后半程，而且会在 4 分钟内跑完全程。

彼得会在 2 分 05 秒以内跑完前半程。

关键事件（在1分55秒以内跑完后半程且在4分钟内跑完全程）被清晰地定义为一个合取事件，而不是一个条件事件。然而，76%的被试（ $N=96$ ）认为这一合取事件比其中一个子事件的概率更高，而48%的被试认为它比两个子事件的概率都要高。人们对于子事件之间的关系的自然判断影响了他们对合取事件的判断。相比之下，没有人违背外延性原则，即认为第2个选项（4分钟以内跑完全程）比第1个选项（4分06秒以内跑完全程）更有可能发生。上述结果表明，对于合取事件的概率判断不能由平均模型来解释，因为在这种模型中， $P(A \& B)$ 位于 $P(A)$ 和 $P(B)$ 之间。但是，平均的过程可能会导致一些合取错误，尤其是当合取事件中的子事件的概率是以数值的形式呈现时。

动机与犯罪

动机-行动模式中的合取错误如下所示，我们对171名加拿大不列颠哥伦比亚大学的学生进行了以下考察：

约翰·P是个性格温和的男性，42岁，已婚，有两个孩子。邻居说他性情温和，但有点神秘。他在纽约市经营一家进出口贸易公司，经常去欧洲和远东旅行。P先生曾因走私宝石和金属（包括铀）而被定罪，被判6个月监禁，缓期执行，且被处以巨额罚款。P先生目前正在接受警方调查。

请将下列陈述按其将成为调查结论的概率进行排序。请记住，还存在其他的可能性，而且不止一项陈述可能是正确的。1代表可能性最高的陈述，2代表可能性次高的陈述，依此类推。

- P先生是个儿童猥亵犯。
- P先生参与了间谍活动并出售秘密文件。
- P先生是个瘾君子。
- P先生杀死了自己的一名员工。

一半的被试（ $N=86$ ）对以上陈述进行了排序。而其他被试（ $N=85$ ）看到的陈述列表是经过修订的，其中最后一项被替换为：

P先生杀死了自己的一名员工，以阻止其报警。

虽然增加了一种可能的动机显然会降低事件的外延性（P先生可能因为其他原因，如因报复或自卫而杀死自己的员工），但我们假设，提到一个看似合理但并不明显的动机，会增加人们对事件发生的可能性的感知。数据证实了这一预期。合取事件的可能性排序均值为2.90，而包含性陈述的排序均值为3.17（ t 检验， $p<0.05$ ）。此外，有50%的被试认为合取事件比P先生是个瘾君子这一单一事件更有可能发生，但只有23%的被试认为存在包含关系的目标事件比P先生是个瘾君子更有可能发生。我们在其他同类型的问题中发现，当动机为目标事件提供了一个合理的解释且其本身看起来可能性很高但同时又不明显时，提到这样一种原因或动机往往会提高某一行动的判断概率，从某种意义上来说，当提到结果时，这种原因或动机不会立刻浮现在人们的脑海中。

在 $A \rightarrow B$ 和 $M \rightarrow A$ 两种范式中，我们都在其他涉及犯罪行为的判断中观察到了合取错误。例如，一名有暴力倾向的警察参与毒品交易，这种假设的可能性（相对于标准的比较集）比他参与毒品交易且最近还袭击了一名嫌疑人的合取推论的可能性排序更低。在这个例子中，近期袭击嫌疑人的行为与毒品之间并没有因果关系，但它使合取事件中的描述与嫌疑人的性格更为相符。判断心理学对法律证据评估的影响值得深入研究，因为许多审判的结果取决于法官或陪审团根据部分和容易出错的数据做出直觉判断的能力（Rubinstein, 1979; Saks & Kidd, 1981）。

预测和情景

对未来事件的情景的构建和评估不仅仅是记者、分析员和新闻观察员的爱好。情景经常会在规划的背景下使用，它们的可能性影响着重要的决策过程。过去的情景在很多情况中也非常重要，包括刑法和历史编写。因此，评估现实事件的预测或重建是否会受到合取错误的影响是很

有意义的。我们的分析表明，一个包含可能原因和结果的情景可能比结果本身更有可能出现。我们分别在没有统计学知识背景和有统计学知识背景的被试中检验了这一假设。

1982年4月，我们请245名加拿大不列颠哥伦比亚大学的本科生对1983年的几起事件发生的概率进行了评估。我们采用了9分制量表，定义如下：低于0.1%、0.1%、0.5%、1%、2%、5%、10%、25%、50%和50%以上。每个问题都以两个不同的版本呈现给不同的被试：一个版本只有基本的结果，而另一个版本中包含了导致这一结果的更详细的情景信息。例如，一半的被试评估了1983年北美某地区发生巨大洪水的概率，在这场洪水中有1 000多人丧生。另一半被试评估了1983年加利福尼亚州发生地震的概率，这次地震引发的洪水导致1 000多人丧生。

结果表明，人们对合取事件（地震和洪水）的概率判断明显高于对洪水的概率判断（曼-惠特尼检验， $p<0.01$ ）。两个概率的几何平均值分别为3.1%和2.2%。因此，提醒人们灾难性的洪水可能是由事先已经预料到的加利福尼亚州地震造成的，这会让人们对地震和洪水的合取事件的概率判断大于只发生洪水的概率判断。在其他问题中，我们也发现了同样的模式。

这个研究的第二部分中的被试，是1982年7月土耳其伊斯坦布尔第二届全球预测大会的115名参会者。大部分被试都是职业分析师，受雇于企业、大学和研究机构。他们是预测和计划方面的专家，而且很多人也在自己的工作中采用了情景分析法。这个研究的设计和问卷量表都与上一个研究一样。其中一组预测者评估了两个国家于某个时期完全中断外交关系的概率。

其他被试评估了下述情景中出现相同结果的概率：在某个时期，世界爆发局部战争，两个国家完全中断外交关系。

尽管“断交”比“爆发局部战争且断交”发生的概率更大，但爆发局部战争为两个国家的断交提供了一个看似合理的情景。正如我们预期的那样，这两个问题的概率估值都很低，但“爆发局部战争且断交”这一合取事件的概率估值却显著高于“断交”（曼-惠特尼检验， $p<0.01$ ）。两者的几何平均值分别为0.47%和0.14%。在下述的比较中，也出现了类似结果：

1983 年，美国的石油消费量会下降 30%。

1983 年，石油价格急剧上涨，美国的石油消费量会下降 30%。

这两者的几何均值分别为0.22%和0.36%。我们推测在这个问题中这种效应会小一些（尽管仍然具有统计学意义），因为基本的目标事件（石油消费量大幅下降）使得附加事件（油价的大幅上涨）非常容易被人们想到，即使我们并没有提及这一点。

对于涉及假设性原因的合取事件，人们尤其容易犯错，因为在给定原因的情况下评估结果的概率比评估结果和原因的联合概率更为自然。我们不认为被试会刻意采用这种解释；相反，我们提出了一个更高的条件估计作为锚定点，使合取事件看起来更有可能一些。

尝试预测一些重大国际事件或争端通常都涉及情景的构建和评估。同样，一个关于受害者可能是由被告之外的人谋杀的这种似是而非的故事，可能会让陪审团对相关案件产生合理的怀疑。情景可以有效地激发想象力，建立结果的可行性，或为概率的判断设置界限（Kirkwood & Pollock, 1982; Zentner, 1982）。然而，将情景用作评估概率的主要工具，可能具有高度误导性。首先，这一过程会让人们更相信由一系列可能的步骤（如一个计划的成功执行）产生的合取结果，而不是同等概率的析取结果（如一个周密计划的失败），后者可能以多种不太可能的方式发生（Bar-Hillel, 1973; Tversky & Kahneman, 1973）。其次，通过情景应用来评估概率特别容易出现合取错误。一个由因果联系和代表性事件组成的详细情景可能比这些事件的子集更有可能出现（Slovic, Fischhoff & Lichtenstein, 1976）。这种影响有助于增强情景的吸引力以及这些情况带来的虚幻的洞察力。律师在对未知事实，如动机或操作方式进行推测时，可能会通过提高案件的连贯性来增加说服力，尽管这种补充只会降低案件发生的可能性。类似地，政治分析师可以通过添加可信的原因和具有代表性的结果来改善情景。

外延性线索

本文中提到的许多合取错误说明了人们对非外延性推理的偏好。很明显，人们可以理解并应用外延性规则。什么线索会引发外延性思考？什么因素会促进人们遵循合取原则？在这一节中，我们将重点关注单个估计问题，并报告几种会诱发外延性推理和降低合取谬误发生率的操作方式。本节所述研究的参与者是加拿大不列颠哥伦比亚大学没有统计学

知识背景的学生。括号中给出了平均估计值。

加拿大不列颠哥伦比亚省对当地所有年龄段、从事各类职业的成年男性进行了健康情况抽样调查。

请就下列问题做出最佳估计：

- 接受调查的男性中，经历过一次或多次心脏病发作的人的比例是多少？（18%）
- 接受调查的男性中，年龄在55岁以上且经历过一次或多次心脏病发作的人的比例是多少？（30%）

对于这个有关健康调查的问题，不具备统计学知识背景的被试犯下了大量的合取错误：65%的被试（ $N=147$ ）对第二个问题的概率估计比第一个问题的估计更高（Bar-Hillel, 1923）[\(4\)](#)。问题呈现的顺序对结果没有显著影响。

我们可以将在相对频率估计中观察到的对合取原则的违反归因于 $A \rightarrow B$ 范式。我们认为，对年龄与心脏病发作之间的因果或统计关联强度进行的自然评估，影响了人们对合取事件的概率判断。虽然问题的陈述看起来很明显，但我们考虑的假设是，犯了这种错误的被试实际上是把第二个问题理解成评估一个条件概率的前提。我们请另一组加拿大不列颠哥伦比亚大学本科学生也来回答这两个问题，但对第二个问题略做了修改：

接受调查的年龄在55岁以上的男性中，经历过一次或多次心脏病发作的人的比例是多少？

这组被试给出的平均估值为59%（ $N=55$ ）。这个值明显高于那些在原始问题中犯了合取谬误的被试给出的合取事件的平均估值（45%）。因此，违背了合取原则的被试并不是简单地用条件概率 $P(B/A)$ 代替合取概率 $P(A \& B)$ 。

问题中看似无关紧要的变化可以帮助许多被试避免合取谬误。一组新的被试（ $N=159$ ）被要求回答最初的问题，但需要在评估合取问题前先来评估“55岁以上男性受访者在样本中所占的比例”这一问题。这一操作将合取错误的发生率从65%降到了31%。在评估两类事件交集的频率之前，先对其相对频率进行估计的要求，似乎给了被试适当的暗示。

下面的构想也有助于进行外延性推理：

加拿大不列颠哥伦比亚省进行了一项针对当地所有年龄段和职业男性的健康情况

抽样调查。

请就下列问题做出最佳估计：

- 100名参与者中有多少人经历过一次或多次心脏病发作？
- 100名参与者中有多少人都超过了55岁，并且都经历过一次或多次心脏病发作？

在这个版本中，合取谬误的发生率只有25%（ $N=117$ ）。显然，明确标记出个案的数量会鼓励被试构建相关问题的表象，在这种问题的表象中，类别的包含关系很容易被感知和理解。我们已经在其他几个相同类型的问题中重复了这种结果。让一组被试（ $N=360$ ）在判断合取事件概率前，先估计55岁以上的参与者样本大小，此时他们的错误率进一步降低至11%。这个研究结果与R.贝斯-马罗姆（R. Beyth-Marom）的结果一致，他观察到在概率判断中，对合取事件的估计高于在频率评估中的估计（Beyth-Marom, 1981）。

这一部分的研究结果表明，即使是在目标事件的外延和评价量表的意义都完全明确，且针对相对频率的简单估计中，非外延性推理有时也会产生决定性的影响。此外，我们发现，用频率代替百分比并要求被试评估这两种成分类别，可以显著降低合取谬误的发生率。看起来似乎无关紧要的线索很容易让人做出外延性推理。这里有一个值得注意的对比，那就是外延性线索在健康调查问题中的效度，与用于消除琳达问题中的合取谬误的方法（论证、投注、“是否”）的相对无效性之间的对比。当合取由具体类别的交集而不是由属性的组合定义时，合取原则的影响力就更容易被理解。虽然从逻辑的角度来看，类别和属性是等价的，但它们会产生不同的心理表征，在这些心理表征中，不同的关系和规则是明确的。属性与类别在形式上的等价性显然无法通过程序化的方式灌输给外行。

讨论

在这个项目中，我们研究了多个领域的外延规则；我们对3 000多名被试进行了数十个问题的测试，并检验了这些问题的多种变体。本章所报告的结果是对这项工作的一个有代表性的总结，但并非详尽无遗。

数据显示，在间接和直接检验中，没有专业知识和具备专业知识的被试都普遍地违反了外延规则。这些结果在判断启发式的框架内得到了解释。我们认为，概率或频率的判断常常会受到问题引发的自然评估的影响。因此，估计某一类别的频率的要求会使人们开始进行范例搜索，

从性格概述中预测职业选择的任务会引发针对特征的比较，而关于事件共发的问题则会引发对其因果关系的评估。这些评估不受外延规则的限制。虽然任意减少事件的外延性通常会降低其可得性、代表性或因果一致性，但在许多情况下，对限制事件的评估要高于对合取事件的评估。自然评估可以通过3种方式使概率判断产生偏差：第一，被试可能会有意将自然评估用作一种评估策略；第二，被试可能会受到自然评估的引导或锚定；第三，被试可能没有鉴别出自然评估与所需评估之间的差异。

逻辑与直觉

合取错误以足够清晰的方式向我们展示了以最正式的概率概念为基础的外延逻辑与支配很多判断和信念的自然评估之间的对比。然而，概率判断并不总是由非外延性启发式所主导的。概率论的基本原理已经成为文化的一部分，甚至没有统计学知识背景的成年人也能在简单的概率游戏中列举出各种可能性和计算胜率（Edwards, 1975）。此外，一些真实生活的情境能帮助人们对事件进行分解。例如，可以对团队进入季后赛的机会进行如下评估：“如果我们击败团队B，就可以进入季后赛，我们应该能够做到这一点，因为我们有了更好的防守；或者B队输给C队和D队，我们也能进入季后赛，而这是不可能的，因为这两支队伍都没有很强的进攻能力。”在这个例子中，目标事件（进入季后赛）被分解成更基本的可能性，人们可以用一种直观的方式对其进行评估。

概率判断的不同之处在于它们遵循分解法或整体法的程度，以及对概率的评估和汇总是分析性或直觉性的程度（Hammond & Brehmer, 1973）。在一种极端情况中，可以通过计算“有利”结果的相对频率来回答一些问题，如在扑克游戏中击败选手牌的概率有多少？这种分析具有与扩展方法相关的所有特性：它是分解性的、频率性的和算法性的。在另一种极端情况中，人们通常以一种整体的、单一的和直观的方式对一些问题进行评估，如目击者说真话的概率是多少（Kahneman & Tversky, 1982b）。分解和计算可以防止合取错误和其他偏差的出现，但在随机抽样的领域外，直觉元素并不能从概率判断中被完全排除。

对合取原则的直接检验使直觉印象与概率论的基本定律对立起来。这一冲突的结果取决于证据的本质、问题的提法、事件结构的明显程度、启发式的吸引力，以及被试的专业背景。在任何一个特定的直接检验中，人们是否遵守合取原则取决于这些因素之间的平衡。例如，我们

发现很难引导那些没有专业背景的被试在琳达问题中应用合取原则，但是在健康调查问题中，只需要做出微小的改变就会对合取错误产生显著影响。这个结论和尼斯贝特等人的研究结果是一致的，他们发现外行人也可以将某些统计原则（如大数定律）应用于日常问题，这些原则的可得性会随着问题的内容发生变化，而且会随着被试的专业知识水平的提高而明显增加（Nisbett et al., 1983）。然而，我们发现，有专业背景的被试和缺乏经验的被试在间接检验中对琳达问题的回答是相似的，只是在包含关系非常明显的问题版本中出现了分歧。这些观察结果表明，统计的复杂性并没有改变人们对代表性的直觉，尽管它能使被试在直接检验中认识到外延性原则的决定性力量。

现实生活中的判断问题，通常不会以“被试间设计检验”或概率定律的直接检验的形式出现。因此，在被试间设计检验中，被试的表现可以为日常推理提供一个更现实的视角。在间接检验中，即使是经验丰富的判断者也很难保证一个事件中没有比其自身更有可能发生的子集，也不存在发生概率更低的超集。如果将相关集合中的所有事件都表示为基础概率的不相交并集，则肯定可以满足外延性原则，而无须直接将A&B与B进行比较。然而，在许多实际情况下，这种分析是不可行的。医生、法官、政治分析师或企业家通常只关注关键性目标事件，很少会发现可能会违反外延性原则的行为。

对推理和解决问题的研究表明，人们往往无法理解或应用抽象的逻辑性原则，即使他们在熟悉的具体环境中可以正确地使用，情况依然如此。例如，P.N.约翰逊-莱尔德（P.N. Johnson-Laird）和彼得·沃森（Peter Cathcart Wason）指出，在抽象形式中验证“如果……那么”（if then）语句时出现错误的人，通常在问题激活了一个熟悉的图式时就能获得成功（Johnson-Laird & Wason, 1977）。目前的结果则显示了相反的模式：人们通常能接受抽象形式的合取原则（B比A&B更有可能），但在具体的例子中却会违反合取原则，如在琳达和比尔问题中，这一规则就会与直观印象产生冲突。

违反合取原则的现象不仅在我们的研究中普遍存在，而且出现的次数也很可观。例如，在估计单词频率的时候，被试给出的以ing结尾的7个字母构成的单词出现的频率估计，是以_n_结尾的7个字母构成的单词出现频率估计的3倍。3倍的修正是消除两个估计之间不一致的最小改变。然而，被试当然知道有许多以_n_结尾但不是ing的单词（如present和content）。例如，如果他们相信N个单词中只有一半以ing结尾，那么

就需要6:1的调整来使整个系统一致。我们大多数的实验具有的常规性质不允许我们对一致性所需的调整系数进行估计。然而，这种影响的效应值往往是相当大的。例如，在琳达问题的评级量表版本中，T&F和T的评级分布之间几乎没有重叠。当然，我们给出的问题是为引出合取错误而设计的，它们不适用于针对这些错误的普遍程度进行的无偏估计。然而，请注意，合取错误只是一种更普遍的现象的症状：人们倾向于高估具有高度代表性（或可得性）的事件的概率或低估代表性较低的事件的概率。即使“真实”概率是未知的或不可知的，违背合取原则也显示了这种趋势。这种基本现象可能比它的极端现象更为普遍。

以往的关于合取事件的主观概率的研究主要集中于检验乘法规则 $P(A \& B) = P(B) P(A/B)$ （Bar-Hillel, 1973; Cohen & Hansel, 1957; Goldsmith, 1978; Wyer, 1976; Beyth-Marom, 1981）。这项法则比合取原则更严格；它还要求对概率进行基数而非序数的评估。结果表明，人们普遍高估了合取事件的概率，即 $P(A \& B) > P(B) P(A/B)$ 。一些研究人员，特别是怀尔（Wyer）和贝斯-马罗姆，也报告了与合取原则不一致的数据。

不确定状况下的对话

代表性启发式通常倾向于那些有好故事或好假设的结果。“在女权运动中表现活跃的银行出纳”的合取是比“银行出纳”更适合琳达的假设，我们可以通过将格莱斯关于合作的概念扩展到不确定状况下的对话中，来阐述一个好故事的概念（Grice, 1975）。对话规则的标准分析假设说话人知道真相。品质准则要求他只能说真话，数量准则要求说话人说出所有信息，而相关性原则是将信息限制在听话人需要知道的范围内。什么样的合作规则能适用于一个不确定的说话人？也就是说，什么样的合作规则可适用于一个不确定是否说了真话的人？这样的说话人只能保证同义反复的绝对质量（如“只要价格上涨，通货膨胀就会继续”），而这些同义反复的陈述不太可能在对话中产生重要意义。它们真正的意义在于传达说话人的相关信念，即使是在他们并不确定的情况下。因此，针对不确定的说话人的合作规则必须在进行信息评价时考虑到质量和数量之间的平衡。如果信息为真，那么其前景值可以由它的信息值来定义，并由它为真的概率来加权。一个不确定的说话人可能希望遵循价值准则：选择前景值最高的信息。

我们有时可以通过增加信息的内容来提高它的前景值，尽管这会降

低其概率。“到今年年底，通货膨胀率将为6%~9%”的预测可能比“通货膨胀率将为3%~12%”的预测更有价值，尽管后者更有可能得到证实。一个好的预测是在一个点估计（肯定是错误的）和一个99.9%的置信区间（通常太宽泛）之间做出的折中。科学假设的选择也面临着同样的取舍：假设必须有被证伪的风险才会有价值，但如果几乎肯定会被证伪，那么它的价值就会下降。好的假设可以平衡信息量和可能的真实性（Good, 1971）。在自然范畴的结构中也存在类似的折中。“狗”这一基础类别比范围更大的类别“动物”包含的信息量要大得多，但比范围较小的类别“小猎犬”包含的信息量略少。基础类别在语言和思想方面具有优势，这大概是因为它们为我们提供了范围和内容的最佳组合（Rosch, 1978）。不确定状况下的分类就是一个很好的例子。一个在黑暗中依稀可见的移动物体可能被恰当地贴上了“狗”的标签，而这时如果认定其属于细分类别的“小猎犬”会显得有些鲁莽，而如果贴上“动物”的标签则又过于保守。

请对这个问题的可能答案进行排序：“你认为琳达最近在忙什么？”在这个任务中，价值准则可以证明对T&F而不是T的偏好是合理的，因为增加“女权主义”的属性极大地丰富了对琳达当前活动的描述，这种可能的事实是可以接受的代价。因此，对不确定状况下的对话的分析确定了一个相关问题，可以通过将合取事件置于其子事件之上来对这个问题进行解答。然而，我们认为价值准则并不能完美地解释合取谬误。首先，我们的被试不太可能将“根据概率对语句进行排序的要求”解释为根据预期（信息）值进行排序。其次，我们在数值估计和投注选择中观察到了合取谬误，但对话分析则根本不适用于这些情况。然而，对具有高前景（信息）值的陈述的偏好可能会妨碍我们对外延性规则的理解。正如我们在讨论图片尺寸和真实尺寸的交互时所提到的那样，即使被试很清楚两者之间的区别，一个问题的答案也可能会因另一个同源问题的答案的可得性而产生一些偏差。

同样的分析也适用于概率的其他相关概念。惊讶的概念就是一个很好的例子。虽然惊讶与预期密切相关，但它并不遵循概率定律

（Kahneman & Tversky, 1982b）。例如，网球冠军输了第一局的信息比他输了第一局但赢得比赛的信息，更让人觉得惊讶；抛硬币连续4次出现正面比出现4次正面后接着出现2次反面更令人惊讶。然而，在每一次投注中，将赌注都压在不那么令人惊讶的选项上显然是荒谬的。我们与被试的讨论没有显示他们将概率判断的指令解释为评估惊讶程度的指令。此外，这种对惊讶程度的理解并不适用于频率判断中观察到的合取

谬误。我们的结论是，惊讶和信息价值并不能很好地解释合取谬误，尽管它们很可能有助于消除合取谬误的诱因，以及降低消除合取谬误的难度。

认知错觉

我们对归纳推理的研究集中在系统错误上，因为它对于那些通常支配着我们的判断和推理的启发式具有诊断作用。用德国生物学家亥姆霍兹（Helmholtz）的话来说：“正是那些不符合现实的情况，对我们发现一般知觉产生过程的规律特别有指导意义。”对偏差和错觉的关注是一种利用人为错误的研究策略，尽管它既不假设也不要求人们在感知或认知上是无能的。亥姆霍兹的观点意味着，知觉并没有被有效地分析为一个产生准确知觉的正常过程和一个产生错误和错觉的扭曲过程。认知与知觉一样，同样的机制既能产生有效判断，也可以产生无效判断。事实上，证据似乎并不支持“真理加错误”的模型，该模型假设存在一个连贯的、会受到各种扭曲和错误来源干扰的系统。因此，我们并不认同丹尼斯·林德利（Dennis Lindley）的乐观观点，即“每个语无伦次的人内心其实都有一个试图挣脱的，可以有序表达的人”（Lindley, 1980），而且我们怀疑语无伦次不仅仅是表面现象（Tversky & Kahneman, 1981）。

将关于一个领域的信念结构（如中美洲的政治前途）与对一个场景的看法（如从冰川点看约塞米蒂峡谷）进行比较是具有指导意义的。我们已经论证过，对相关边际概率、合取概率和条件概率的直觉判断不太可能是一致的，即无法满足概率论的约束条件。同样，对场景中距离和角度的估计也不太可能满足几何定律。例如，有些政治事件的概率 $P(A)$ 被判断为大于 $P(B)$ ，但 $P(A/B)$ 的概率却被判断为小于 $P(B/A)$ （Tversky & Kahneman, 1980）。类似地，在一个可能包含三角形ABC的场景中，其中A角看起来大于B角，尽管BC之间的距离看起来小于AC之间的距离。

在距离和可能性判断中违反几何和概率的定性规律对这些判断的解释和运用具有重要意义。不一致性极大地限制了可以从主观估计中得出的推论。我们不能从三角形的各个角的判断顺序中推导出各边的判断顺序，而且也不能从各个条件概率的判断顺序中推导出边际概率的判断顺序。本研究结果表明，假设 $P(B)$ 受 $P(A \& B)$ 的约束甚至是不安全的。此外，如果一个判断系统不遵守合取原则，我们就不能指望它遵守以合取原则为前提的更复杂的规则，如贝叶斯更新、外部校准和前景效

用最大化。偏差和不一致的存在并没有削弱这些规则的规范性，但它降低了其作为行为描述的效度，并阻碍了它们的应用。事实上，无偏差判断的产生和不连贯评估的调和造成了严重的问题，目前还没有令人满意的解决办法（Lindley, Tversky & Brown, 1979; Shafer & Tversky, 1983）。

一致性问题在偏好和信念的研究中比在知觉的研究中更为突出。对距离和角度的判断可以很容易地与客观现实进行比较，当准确度很重要时，我们可以用客观测量来代替。相比之下，对概率的客观测量往往是不可实现的，在进行存在风险的最重要的选择时都需要对概率进行直观的评估。在缺乏有效的客观标准的情况下，不确定性判断的标准理论将信念的一致性作为人类理性的试金石。在心理学、经济学和其他社会科学的许多描述性分析中也假定了一致性。这一假设之所以有吸引力，是因为概率定律强大的规范性使得那些反例显得难以置信。我们针对合取原则的研究表明，假设一致性的标准启发性理论在描述上是不充分的，而那些忽视规范性规则的心理分析也是不完整的。对人类判断的全面描述必须反映出令人信服的逻辑规则与引人入胜的非外延直觉之间的张力。

参考文献

- Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology*, 35, 303–314.
- Bar-Hillel, M. (1973). On the subjective probability of compound events. *Organizational Behavior and Human Performance*, 9, 396–406.
- Beyth-Marom, R. (1981). *The subjective probability of conjunctions* (Decision Research Report No. 81–12). Eugene, OR: Decision Research.
- Bruner, J. S. (1966). On the conservation of liquids. In J. S. Bruner, R. R. Olver, P. M. Greenfield, et al. (Eds.), *Studies in cognitive growth* (pp. 183–207). New York, NY: Wiley.
- Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology*, 73, 193–204.

Cohen, J., & Hansel, C. M. (1957). The nature of decision in gambling: Equivalence of single and compound subjective probabilities. *Acta Psychologica*, 13, 357–370.

Cohen, L. J. (1977). *The probable and the provable*. Oxford, England: Clarendon Press.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38, 325–339.

Edwards, W. (1975). Comment. *Journal of the American Statistical Association*, 70, 291–293.

Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology*, 32, 53–88.

Gati, I., & Tversky, A. (1982). Representations of qualitative and quantitative dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 325–340.

Goldsmith, R. W. (1978). Assessing probabilities of compound events in a judicial context. *Scandinavian Journal of Psychology*, 19, 103–110.

Good, I. J. (1971). The probabilistic explication of information, evidence, surprise, causality, explanation, and utility. In V. P. Godambe & D. A. Sprott (Eds.), *Foundations of statistical inference: Proceedings on the foundations of statistical inference* (pp. 108–127). Toronto, ON: Holt, Rinehart & Winston.

Grice, H. P. (1975). Logic and conversation. In G. Harman & D. Davidson (Eds.), *The logic of grammar* (pp. 64–75). Encino, CA: Dickinson.

Hammond, K. R., & Brehmer, B. (1973). Quasi-rationality and distrust: Implications for international conflict. In L. Rappoport & D. A. Summers (Eds.), *Human judgment and social interaction* (pp. 338–391). New York, NY: Holt, Rinehart & Winston.

von Helmholtz, H. (1903). *Popular lectures on scientific subjects* (E. Atkinson, Trans.). New York, NY: Green. (Original work published 1881).

Jennings, D., Amabile, T., & Ross, L. (1982). Informal covariation assessment. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 211–230). New York, NY: Cambridge University Press.

Johnson-Laird, P. N., & Wason, P. C. (1977). A theoretical analysis of insight into a reasoning task. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking: Readings in cognitive science* (pp. 143–157). Cambridge, England: Cambridge University Press.

Jones, G. V. (1982). Stacks not fuzzy sets: An ordinal basis for prototype theory of concepts. *Cognition*, 12, 281–290.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. New York, NY: Cambridge University Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430–454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237–251.

Kahneman, D., & Tversky, A. (1982a). On the study of statistical intuitions. *Cognition*, 11, 123–141.

Kahneman, D., & Tversky, A. (1982b). Variants of uncertainty. *Cognition*, 11, 143–157.

Kirkwood, C. W., & Pollock, S. M. (1982). Multiple attribute scenarios, bounded probabilities, and threats of nuclear theft. *Futures*, 14, 545–553.

Kyburg, H. E. (1983). Rational belief. *Behavioral and Brain Sciences*, 6, 231–245.

Lakoff, G. (1982). *Categories and cognitive models* (Cognitive Science Report No. 2). Berkeley: University of California.

Lindley, D. V., Tversky, A., & Brown, R. V. (1979). On the reconciliation of probability assessments. *Journal of the Royal Statistical Society. Series A (General)*, 142, 146–180.

Lindley, D., Personal communication, 1980.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, 32, 89–115.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review*, 90, 339–363.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9, 35–38.

Osherson, D. N., & Smith, E. E. (1982). Gradedness and conceptual combination. *Cognition*, 12, 299–318.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Rubinstein, A. (1979). False probabilistic arguments Vs. faulty intuition. *Israel Law Review*, 14, 247–254.

Saks, M. J., & Kidd, R. F. (1981). Human information processing and adjudication: Trials by heuristics. *Law & Society Review*, 15, 123–160.

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton, NJ: Princeton University Press.

Shafer, G., & Tversky, A. (1983). *Weighing evidence: The design and comparisons of probability thought experiments*. Unpublished

manuscript, Stanford University, Stanford, CA.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1976). Cognitive processes and societal risk taking. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior* (pp. 165–184). Potomac, MD: Erlbaum.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Suppes, P. (1975). Approximate probability and expectation of gambles. *Erkenntnis*, 9, 153–161.

Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327–352.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, 76, 105–110.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability.

Cognitive Psychology, 5, 207–232.

Tversky, A., & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in social psychology* (pp. 49–72). Hillsdale, NJ: Erlbaum.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453–458.

Tversky, A., & Kahneman, D. (1982). Judgments of and by representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 84–98). New York, NY: Cambridge University Press.

Wyer, R. S., Jr. (1976). An investigation of the relations among probability estimates. *Organizational Behavior and Human Performance*, 15, 1–18.

Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1, 3–28.

Zadeh, L. A. (1982). A note on prototype theory and fuzzy sets. *Cognition*, 12, 291–297.

Zentner, R. D. (1982). Scenarios, past, present and future. *Long Range Planning*, 15, 12–20.

第3章 相似性特征

阿莫斯·特沃斯基

相似性在认知和行为理论中发挥着基础性作用。它作为一种组织原则，能够帮助人们将对象进行分类，形成概念并做出概括。实际上，相似性的概念在心理学理论中无处不在。它是学习过程中的刺激和反应泛化（response generalization）的基础，被用来解释记忆和模式识别中的错误，也是分析隐含意义的核心。

相似性或非相似性数据以不同的形式出现：成对选项的评级、对象的排序，关联对象之间的共同性、替换性错误，以及事件之间的相关性。对这些数据进行分析是为了解释观察到的相似性关系，并捕获研究对象的基础结构。

相似关系的理论分析一直由几何模型主导。这些模型将对象表示为某些坐标中的点，使观察到的对象之间的非相似性能够与各个点之间相应的尺度距离对应起来。在实践中，对数据之间的距离的所有分析在本质上都可作为尺度，尽管一些数据（如分级聚类分析）是树状结构而不是以维度进行组织的空间结构。然而，大多数关于相似性的理论和实证分析都会假设其研究对象可以被充分地表示为某些坐标空间中的点，而非相似性的表现则类似于尺度距离函数。维度和尺度假设都是存疑的。

许多作者认为，维度表征适用于颜色、声调等刺激，但不适用于其他刺激。而诸如面容特征、国家或者人物个性等，以定性特征进行表征似乎比定量的维度更加恰当。因此，可以将对这些刺激之间的相似性的评估描述为特征的比较而不是各点之间的距离的计算。

作为尺度的度量距离函数 δ 根据以下3个定理为每一对点分配一个非负数，并将其称为它们的距离：

极小性： $\delta(a, b) \geq \delta(a, a) = 0$

对称性： $\delta(a, b) = \delta(b, a)$

三角不等式： $\delta(a, b) + \delta(b, c) \geq \delta(a, c)$

为了评估几何方法的充分性，我们检验了当 δ 被视为非相似性尺度时这些尺度定理的效度。极小定理意味着对于所有对象来说，一个对象与其自身之间的相似性都是相同的。然而，这种假设并不适用于某些相似性尺度。例如，两个相同的刺激为“相同”而不是“不同”的判断概率，并非对于所有刺激都是恒定的。此外，在识别中，非对角线条目通常会胜于对角线条目；也就是说，一个对象被识别为另一个对象的情况比它被识别为自身的情况更常见。如果识别概率被理解为一种相似性的尺度，那么这些观察结果就违背了极小定理，因此与距离模型并不契合。

哲学家和心理学家都认为相似性是对称关系的典型例证。实际上，对称性假设本质上是所有相似性理论处理方法的基础。与这种传统观念相反，本文提供了有关非对称相似性的经验证据，并认为相似性不应被视为对称关系。

相似性判断可以被视作相似性陈述的扩展，即“ a 就像 b ”这样的陈述。这种陈述是有方向性的；它有一个主体 a 和一个指示对象 b ；它一般不等同于相反的相似性陈述：“ b 就像 a ”。事实上，主体和指示对象的选择部分取决于对象的相对突出程度。我们倾向于选择更突出的刺激或原型作为指示对象，同时选择不那么突出的刺激或变体作为主体。我们说“肖像与人相似”，而不是“人与肖像相似”。我们说“儿子像父亲”，而不是“父亲像儿子”。我们说“椭圆形与圆形相似”，而不说“圆形与椭圆形相似”。

正如稍后将说明的，相似性陈述选择中的这种非对称性与相似性判断中的非对称性相关。因此，我们会说椭圆形更像圆形而不会说圆形更像椭圆形。显然，非对称性的方向取决于刺激的相对突出程度；变体与原型更相似而不是原型与变体更相似。

相似关系的方向性和非对称性在明喻和隐喻中尤为常见。我们说“土耳其人作战时如同老虎”，而不会说“老虎作战时如同土耳其人”。因为老虎以其战斗勇猛而闻名，因此它被用作指示对象而不是明喻的对象。诗人会写下“我的爱和海洋一样深”，而不是“海洋和我的爱一样深”，因为海洋体现了深度。有时候，这两种方向都会被使用，但它们具有不同的含义。“一个人就像一棵树”意味着人有根；“一棵树就像一个人”意味着这棵树有自己的生命史。“人生如戏”的意思是人们在生活中扮演着角色；“戏如人生”指的是戏剧是在捕捉人类生活的基本要素。

我将在最后一节简要讨论隐喻解释与相似性评估之间的关系。

三角不等式与极小性和对称性不同，因为它不能用序数来表示。由于三角不等式的内容是“三角形的任意两边之和大于第三边”，因此不能用次序或距离数据证明它不成立。然而，三角不等式意味着如果 a 与 b 非常相似，并且 b 与 c 非常相似，那么 a 和 c 不可能非常不相似。因此，它根据 a 和 b 及 b 和 c 之间的相似性设定了 a 和 c 之间的相似性的下限。在下面的例子中，威廉·詹姆斯（William James）对这种假设下的心理效度提出了一些质疑。考虑各国之间的相似性：牙买加与古巴存在相似之处，比如地理位置接近；而古巴与苏联(5)存在相似之处；但牙买加和苏联完全不同。

这个例子表明，正如人们所预料的那样，相似性不具有传递性。此外，它还表明：对牙买加和苏联之间的感知距离，超过了牙买加与古巴之间的感知距离和古巴与苏联之间的感知距离之和。这违反了三角不等式定理。虽然这些例子并不一定能完全反驳三角不等式定理，但它们表明三角不等式定理不应该被认定为相似性模型的基石。

应该注意的是，尺度定理本身不够有说服力。例如，当 $a=b$ 时， $\delta(a, b) = 0$ ；当 $a \neq b$ 时， $\delta(a, b) = 1$ 。为了确定距离函数，我们还进行了附加假设，如一维减法性和多维相加性，将对象的维度结构与其尺度距离相关联。关于这些假设的定理分析和批判性讨论，请参考R.比尔斯（R. Beals）、克兰茨和特沃斯基的相关研究（Beals, Krantz & Tversky, 1968）。

总之，尽管有许多卓有成效的应用，但是相似性分析的几何方法仍然面临着一些困境（Carroll & Wish, 1974; Shepard, 1974）。维度假设的适用性有限，尺度定理也存在着问题。具体而言，极小性在某种程度上存在着问题，对称性显然是错误的，而三角不等式又很难令人信服。

在下一节中，我们将开发另一种基于特征匹配的相似性理论方法，该方法本质上既不是用于衡量维度的也不是用于衡量尺度的。在后续章节中，这种方法将用于揭示、分析和解释几种经验现象，如共同和显著特征的作用，相似性和差异性判断之间的关系、不对称相似性的存在，以及情境对相似性的影响。在最后一节，我们讨论了当前研究的扩展和影响。

特征匹配

设 $\Delta=\{a, b, c, \dots\}$, Δ 表示研究中对象（或刺激物）的域或集合。假设 Δ 中的每个对象都由一组特征或属性表示，同时让 A 、 B 、 C 分别表示特征集，并与对象 a 、 b 、 c 相关联。这些特征可以对应眼睛或嘴巴之类的部分，可以代表大小或颜色等具体属性，也可以反映质量或复杂程度等抽象属性。将刺激的特征描述为特征集已经被用于许多认知过程的分析，如言语感知（Jakobson, Fant & Halle, 1961）、模式识别（Neisser, 1967）、知觉学习、优先选择（Tversky, 1972）和语义判断（Smith, Shoben & Rips, 1974）。

关于特征表示的两个初步解释是有序的。首先，注意到特定对象（如一个人、一个国家或一件家具）的总数据库通常内容丰富且形式复杂这一点非常重要。这些总数据库一般包括外观、功能及与其他对象的关系，还有从我们对世界的一般认识中推断出的该对象的任何其他属性。当面对特定任务（如识别或评估相似性）时，我们从数据库中提取并编译包含着有限相关特征的列表，在此基础上执行需要完成的任务。因此，将一个对象表示为特征集合，被视为提取和编译的最初过程的产物。

其次，术语“特征”通常表示二元变量（如浊辅音与清辅音）的值或标称变量（如眼睛的颜色）的值。然而，特征表示并不局限于二元或标称变量；它们也适用于序数变量（ordinal variable）或基数变量（cardinal variable），如维度。例如，仅在响度方面不同的一系列音调可以被表示为嵌套组的序列，其中与每个音调相关的特征集被包括在与响度更高的音调相关联的特征集中。这种表征与有方向性的一维结构是同构的。一个无定向的一维结构（如仅在音高上不同的一系列音调）可以由一系列重叠集合表示。雷斯特勒研究了定性和定量维度的集合理论表示（Restle, 1959）。

设 $s(a, b)$ 是对 Δ 中所有不同 a 、 b 定义的 a 到 b 的相似性的尺度。 s 为相似性的定序尺度。也就是说， $s(a, b) > s(c, d)$ 意味着 a 与 b 的相似性高于 c 与 d 的相似性。本理论基于以下假设。

1. 匹配

$$s(a, b) = F(A \cap B, A - B, B - A)$$

a 与 b 的相似性可通过有3个参数的函数 F 来表示，其中 $A \cap B$ 表示 a 和 b 共有的特征； $A - B$ 表示属于 a 但不属于 b 的特征； $B - A$ 表示属于 b 但不属于 a 的特征。这些要素的示意图如图3-1所示。

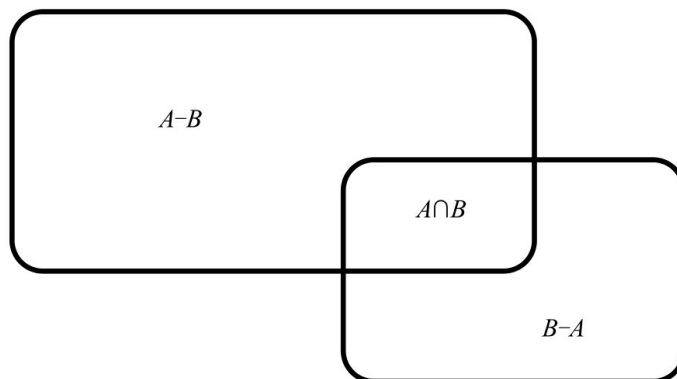


图3-1 两个特征集之间的关系示意图

2. 单调性

当

$$A \cap B \supset A \cap C, A - B \subset A - C$$

且

$$B - A \subset C - A$$

时

$$s(a, b) \geq s(a, c)$$

此外，只要其中任何一个包含关系成立，不等式就成立。

也就是说，相似性随着共同性特征的增加或区别性特征（属于一个对象但不属于另一个对象的特征）的减少而增加。如果我们用要素线（直线）来定义它们的特征，那么单调性就可以很容易地用印刷体大写字母来呈现。在这个假设下， E 应该更像 F 而不是 I ，因为 E 和 F 比 E 和 I 有更多的共同特征。此外， I 应该更像 F 而不是 E ，因为 I 和 F 的区别性特征比 I 和 E 更少。

任何满足假设1和假设2的函数 F 被称为“匹配函数”。它测量的是两个作为特征集的对象之间的匹配程度。在现有的理论中，相似性评价被描述为一个特征匹配过程。因此，相似性评价是根据匹配函数的集合理论概念而不是根据距离的几何概念来表述的。

为了确定匹配函数的函数形式，我们引入了关于相似性排序的附加假设。我们将在下文中介绍理论（独立性）的主要假设；其余的假设和表示定理（representation theorem）的证明详见附录。对形式理论不太感兴趣的读者可以浏览或跳过以下段落，直接阅读表示定理部分。

设 Φ 表示与 Δ 的对象相关联的所有特征集，并且设 X 、 Y 、 Z 等表示特征集（ Φ 的子集）。表达式 $F(X, Y, Z)$ 被定义为当 a, b 在 Δ 中存在时，则 $A \cap B = X$ ， $A - B = Y$ ， $B - A = Z$ ，其中 $s(a, b) = F(A \cap B, A - B, B - A) = F(X, Y, Z)$ 。接下来对于 X, Y, Z ，如果下列等式中的一个或多个成立，那么定义 $V \approx W$ ： $F(V, Y, Z) = F(W, Y, Z)$ ， $F(X, V, Z) = F(X, W, Z)$ ， $F(X, Y, V) = F(X, Y, W)$ 。

当以下一种、两种或三种情况成立时， (a, b) 和 (c, d) 对下面的1个、2个或3个子集是一致的： $(A \cap B) \approx (C \cap D)$ ， $(A - B) \approx (C - D)$ ， $(B - A) \approx (D - C)$ 。

3. 独立性

假设 (a, b) 和 (c, d) 及 (a', b') 和 (c', d') 在相同的两个子集上是一致的，而 (a, b) 和 (a', b') 及 (c, d) 和 (c', d') 在其余的（第三个）子集中达成一致，则当且仅当 $s(c, d) \geq s(c', d')$ 时， $s(a, b) \geq s(a', b')$ 。

为了说明独立性定理的影响力，可参考图3-2中所示的刺激物实验，其中

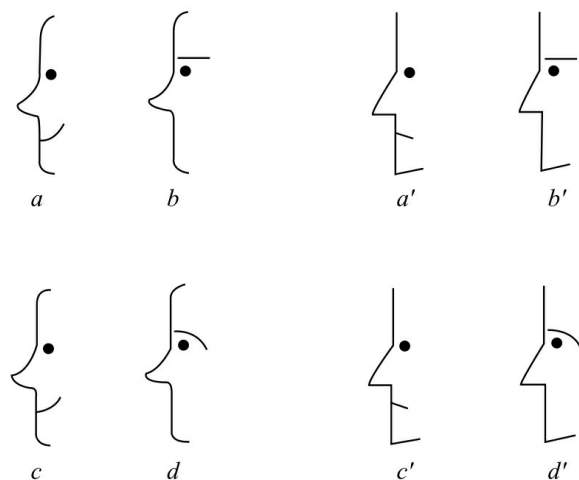


图3-2 独立性示例

$A \cap B = C \cap D = \text{圆形轮廓} = X$

$A' \cap B' = C' \cap D' = \text{尖锐的轮廓} = X'$

$A - B = C - D = \text{微笑的嘴} = Y$

$A' - B' = C' - D' = \text{哭丧的嘴} = Y'$

$B - A = B' - A' = \text{直眉} = Z$

$D - C = D' - C' = \text{弯眉} = Z'$

因此，通过独立性定理可得，当且仅当

$$s(c, d) = F(C \cap D, C - D, D - C) =$$

$$F(X, Y, Z) \geq F(X', Y', Z') =$$

$$F(C' \cap D', C' - D', D' - C') =$$

$$s(c', d')$$

时

$$s(a, b) = F(A \cap B, A - B, B - A) =$$

$$F(X, Y, Z) \geq F(X', Y', Z') =$$

$$F(A' \cap B', A' - B', B' - A') =$$

$$s(a', b')$$

因此，任何两个子集（如 X, Y 与 X', Y' ）的联合效应的排序与第三因素（如 Z 或 Z' ）的固定水平无关。

应该强调的是，对定理的任何检验都应以对特征的解释为前提。例如，独立性定理可能在一种解释中成立，而在另一种解释中则不成立。因此，对定理的实验检验同时检验了特征解释的充分性和假设的经验效度。此外，上述示例不能说明刺激物的特征（如大写字母、面孔示意图）可以由其子集进行适当的表征。为了形成充分的视觉特征表征，应引入更为广泛的属性，如对称性、连通性。对于这个问题的有趣讨论，可参见格式塔心理学传统（Goldmeier, 1972，最初发表于1936年）。

除了匹配性、单调性和独立性之外，我们还假设了可解性和不变性。可解性要求所研究的特征空间足够丰富，以便可以求解某些（相似性）公式。不变性确保了因素之间的区间的等价性。附录中给出了这些假设的严格表述，以及以下结果的证明。

表示定理

假设1、2、3、4和5成立，那么存在一个相似性指标 S 和非负指标 f ，使得 D 中的所有 a, b, c, d ：

(i) 当且仅当 $s(a, b) \geq s(c, d)$ 时， $S(a, b) \geq S(c, d)$ 。

(ii) $S(a, b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$ ，对于某些 $\theta, \alpha, \beta \geq 0$ 。

(iii) f 和 S 是区间尺度。

该定理表明，对于假设1至假设5，存在一个区间相似性尺度 S ，它能保证观察到的相似性顺序，并将相似性表示为测量共同性和区别性特征的线性组合或对比。因此，这种表示被称为“特征对比模型”。在接下来的进一步发展中，我们还假设 f 满足特征可加性。也就是说，只要 X 和 Y 不相交并且满足以上3个条件，那么 $f(X \cup Y) = f(X) + f(Y)$ 。(6)

请注意，对比模型并不是在定义单一相似性尺度，而是在定义以参数 θ 、 α 和 β 的不同值为特征的一系列尺度。例如，如果 $\theta=1$ 且 α 和 β 为零，

那么 $S(a, b) = f(A \cap B)$ ；也就是说，对象之间的相似性是它们的共同性特征的尺度。此外，如果 $\alpha = \beta = 1$ 且 θ 为0，那么 $-S(a, b) = f(A - B) + f(B - A)$ ；也就是说，对象之间的非相似性为各个特征集之间的对称性差异的尺度。雷斯特勒分别将这些公式列为相似性和心理距离的模型（Restle, 1961）。注意，在前一个模型中（ $\theta = 1, \alpha = \beta = 0$ ），对象之间的相似性仅由它们的共同性特征决定，而在后一个模型中（ $\theta = 0, \alpha = \beta = 1$ ），相似性仅由其区别性特征决定。对比模型将对象之间的相似性表示为其共同性特征和区别性特征的加权差异，从而允许在同一域中存在各种相似性关系。

本理论的主要概念是用于评估相似性的对比规则，以及评估反映各种特征的显著性或突出性的尺度 f 。因此， f 衡量了任何特定（共同性或区别性）特征对对象间相似性的影响。因此，与刺激物 a 相关联的标度值 $f(A)$ 被视为该刺激物的总体显著性的衡量标准。促成刺激物显著性的因素包括强度、频率、熟悉度、良好形式和信息内容。尺度 f 和参数（ θ, α, β ）的选择取决于所处的情境，具体的方式我们将在以下部分中讨论。

让我们概括一下表示定理中的假设和证明的内容。一开始，我们描述了特征集合的对象集，以及假设满足理论定理的相似性排序。根据这些假设，我们得出特征空间上的尺度 f ，并证明了对象组的相似性排序与其对比排序是一致的，并以线性组合来表征这些共同性特征和区别性特征。因此，尺度 f 和对比模型是从关于对象相似性的定性定理中推导出来的。

这种结果的本质可以用风险决策经典理论的一个类比来阐明（von Neumann & Morgenstern, 1947）。在风险决策理论中，人们首先会给出一组有关前景预期的选项，用来描述在某些结果空间上的概率分布，以及假设的满足其理论定理的偏好顺序。从这些假设中，人们可以得出有关结果的效用量表，证明前景之间的偏好顺序与其预期的效用顺序一致。因此，效用量表和期望原则都源于关于偏好的定性假设。现有的相似性理论与期望效用模型的不同之处在于，将对象的特征表达为特征集比将不确定选项的特征表达为概率分布可能存在着更多问题。此外，效用理论的定理是作为理性行为的（规范性的）原则被提出的，而当前理论的定理是描述性的而不是规定性的。

对比模型可能是匹配函数最简单的形式，但它并不是唯一值得研究的形式。另一个令人感兴趣的匹配函数是比率模型（ratio model）：

$$S(a,b) = \frac{f(A \cap B)}{f(A \cap B) + \alpha f(A - B) + \beta f(B - A)}, \quad \alpha, \beta \geq 0$$

其中相似性被正态化了，使得 S 介于0和1之间。比率模型概括了文献中提出的几个相似性集合理论模型。如果 $\alpha=\beta=1$ ，那么 $S(a,b)$ 将减少到 $f(A \cap B) / f(A \cup B)$ （Gregson, 1975; Sjöberg, 1972）。如果

$$\alpha = \beta = \frac{1}{2}$$

，那么 $S(a,b)$ 等于 $2 f(A \cap B) / (f(A) + f(B))$ （Eisler & Ekman, 1959）。如果 $\alpha=1$ 且 $\beta=0$ ，那么 $S(a,b)$ 减小到 $f(A \cap B) / f(A)$ （Bush & Mosteller, 1951）。因此，本框架包含各种相似性模型，这些模型在匹配函数 F 的形式和参数的权重分配上有所不同。

为了在任何特定领域中应用和检验本理论，我们必须对相应的特征结构做出一些假设。如果明确指定了每个对象的相关特征，我们就可以直接对理论进行检验，并根据对比模型对特征进行测量。然而，该方法通常仅限于测量从固定的特征集中构造的刺激物（如结构化面孔刺激物、字母、符号串）。如果无法轻易确定与研究中的对象相关联的特征（通常是自然刺激物的情况），我们仍然可以对对比模型的若干预测进行检验，这些预测仅涉及与对象的特征结构相关的一般定性假设。我和伊塔马尔·加蒂（Itamar Gati）都曾在一系列实验中使用过这两种方法。在以下3节中，我们回顾和讨论了我们的主要发现，侧重于定性预测的检验。关于刺激物 and 数据的更详细描述可参阅特沃斯基和加蒂的著作（Tversky & Gati, 1978）。

非对称性和焦点

根据目前的分析，相似性不一定是对称关系。实际上，无论对于对

比模型还是比率模型来说，当且仅当

$$\alpha f(A-B) + \beta f(B-A) = \alpha f(B-A) + \beta f(A-B)$$

以及

$$(\alpha - \beta) f(A-B) = (\alpha - \beta) f(B-A)$$

时

$$s(a, b) = s(b, a)$$

因此，如果 $\alpha = \beta$ 或 $f(A-B) = f(B-A)$ ，即 $f(A) = f(B)$ ，那么 $s(a, b) = s(b, a)$ ，该条件下的特征具有可加性。因此，当对象在测量中是相等的（即 $f(A) = f(B)$ ），或任务为非定向（ $\alpha = \beta$ ）时，对称性成立。要解释后一种情况，可比较以下两种形式：

(i) 评估 a 和 b 彼此相似的程度。

(ii) 评估 a 与 b 相似的程度。

在(i)中，任务是非定向的，因此 $\alpha = \beta$ 且 $s(a, b) = s(b, a)$ 。在(ii)中，任务是定向的，因此 α 和 β 可能不同并且不需要保持对称性。

如果 $s(a, b)$ 被解释为 a 与 b 相似的程度，那么 a 是比较的主体， b 是指示对象。在这样的任务中，人们自然会关注比较的主体。因此，主体的特征比指示对象的特征更重要（即 $\alpha > \beta$ ）。所以，相似性的减少更多的是由主体的区别性特征而不是指示对象的区别性特征引起的，即只要 $\alpha > \beta$ ，当且仅当 $f(B) > f(A)$ 时， $s(a, b) > s(b, a)$ 。

因此，焦点假设（即 $\alpha > \beta$ ）意味着不对称的方向是由刺激物的相对显著性决定的，使得较不显著的刺激物与显著刺激物更相似，反之亦然。特别需要指出的是，变体相似于原型而不是原型相似于变体，因为原型通常比变体更显著。

国家间的相似性

将21对国家作为刺激物，每一对国家中都有一个元素比另一个元素

更突出（如美国-墨西哥和比利时-卢森堡）。为了验证这种关系，我们要求一组69名被试(7)在每对中选出他们认为更突出的国家。除了一对国家，在其他每一对国家比较中，与先验顺序一致的被试比例都超过了2/3。第二组69名被试被要求在两个短语中选择他们更喜欢使用的短语：“国家a类似于国家b”或“国家b类似于国家a”。在所有21个案例中，大多数被试选择了更不突出的国家作为主语及更突出的国家作为指示对象的短语。这些结果表明，人们在选择相似性陈述时存在明显的非对称性，其方向与刺激物的相对显著性一致。

为了检验有关相似性的直接判断中的非对称性，我们向两组77名被试展示了上述21对国家的名单，并要求他们在20分制的评分体系中对相似性进行评分。这两组被试看到的21对国家的名单之间的唯一区别是每对国家的顺序。例如，一组被试被要求评估“苏联与波兰的相似性”，而另一组被试被要求评估“波兰与苏联的相似性”。我们采用了一些手段，将这些名单更突出的国家在第一个位置和第二个位置出现的次数相同。

对于任何一对 (p, q) 的刺激物，让 p 表示更突出的元素，让 q 表示不太突出的元素。在所有被试中， $s(q, p)$ 均值显著高于 $s(p, q)$ 均值：相关样本的 t 检验得到 $t(20) = 2.92$ ， $p < 0.01$ 。为了获得一个基于个体数据的统计性检验，我们计算了每名被试的定向不对称分数。这种定向不对称分数被定义为与显著指示对象进行比较的平均相似性，即 $s(q, p)$ 减去与显著主语进行比较时的平均相似性 $s(p, q)$ 。这种平均差异呈显著正相关： $t(153) = 2.99$ ， $p < 0.01$ 。

我们还通过差异性判断而非相似性判断重复了以上的研究结果。本研究有两组被试，每组23人。他们收到了相同的21对国家名单，第一组被试被要求判断国家a与国家b的差异程度，表示为 $d(a, b)$ ；而第二组被试则被要求判断国家b与国家a的差异程度，表示为 $d(b, a)$ 。如果差异判断遵循对比模型，并且 $\alpha > \beta$ ，那么我们预计显著刺激物 p 与不显著刺激物 q 不同；即 $d(p, q) > d(q, p)$ 。该假设用同一组21对国家和先前确定的显著性排序进行了检验。在所有被试和成对出现的国家中， $d(p, q)$ 的平均值显著高于 $d(q, p)$ 的平均值：相关样本的 t 检验结果为 $t(20) = 2.72$ ， $p < 0.01$ 。此外，上述每名被试的平均不对称得分显著大于零（ $t(45) = 2.24$ ， $p < 0.05$ ）。

图像的相似性

几何图像显著性的一个主要决定因素是形状的良好程度。因此，“好的图像”可能比“糟糕的图像”更为突出，尽管后者通常更为复杂。然而，当两个图像在形状的良好程度上大致相同时，更复杂的图像可能更为显著。为了研究这些假设并检验非对称性预测，我们绘制了两组8对几何图像。在第一组中，每对图像中的一个图像（表示为p）比另一个（表示为q）具有更好的形状。在第二组中，每对图像中的两个图像在形状上大致匹配，但是其中一个图像（p）比另一个图像（q）更丰富或更复杂。每组成对出现的图像的例子如图3-3所示。

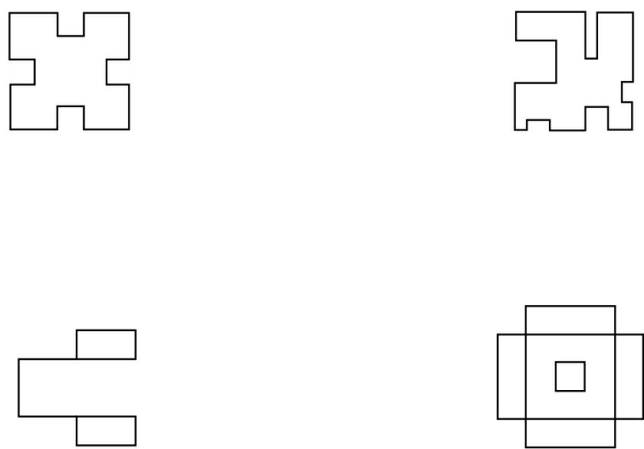


图3-3 用于检验非对称性预测的成对图像

注：上面两个图像是一对，它们在形式上有所不同；下面的图像在复杂性上有所不同。

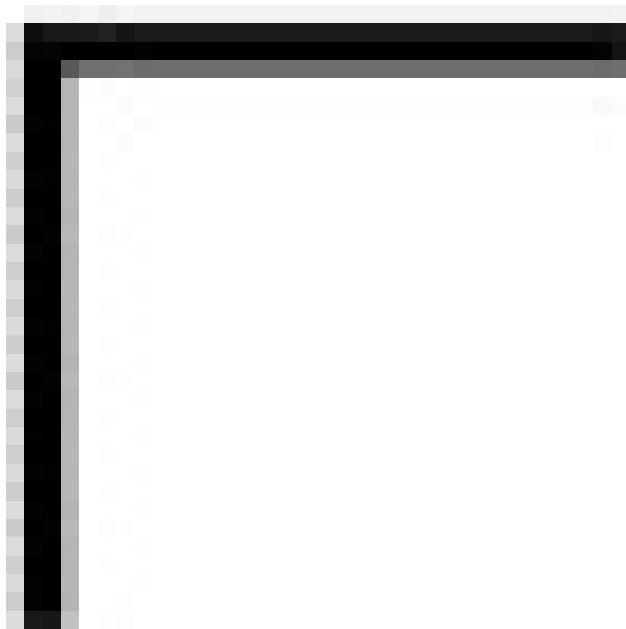
一个由69名被试组成的实验组都看了由16对图像组成的完整列表，其中每对中的两个图像并排显示。对于每对图像，被试被要求指出他们更倾向的一种陈述：“左图与右图相似”或“右图与左图相似”。刺激物的位置是随机的，因此p和q在左侧和右侧出现的次数相同。结果显示，在每对图像中，大多数被试都会选择“q与p相似”的句式。因此，人们通常选择更显著的刺激物作为参照物而不是相似性陈述的主体。

为了检验相似性判断的非对称性，我们选择了两组共67名被试参与，每名被试会看到相同的16对图像，并被要求评价左边图像与右边图像的相似程度（1~20分计）。两组被试收到的小册子基本相同，只有每对图像中的数字的左右位置是相反的。结果显示，所有被试对于s（q, p）的评价均值明显高于s（p, q）的评价均值。对相关样本进行的t检验结果为 $t(15) = 2.94$, $p < 0.01$ 。此外，在两组中，每名被试对于不同方向的相似性评价差异显著大于零：在第一组中 $t(131) = 2.96$, $p < 0.01$ ，第二组中 $t(131) = 2.79$, $p < 0.01$ 。

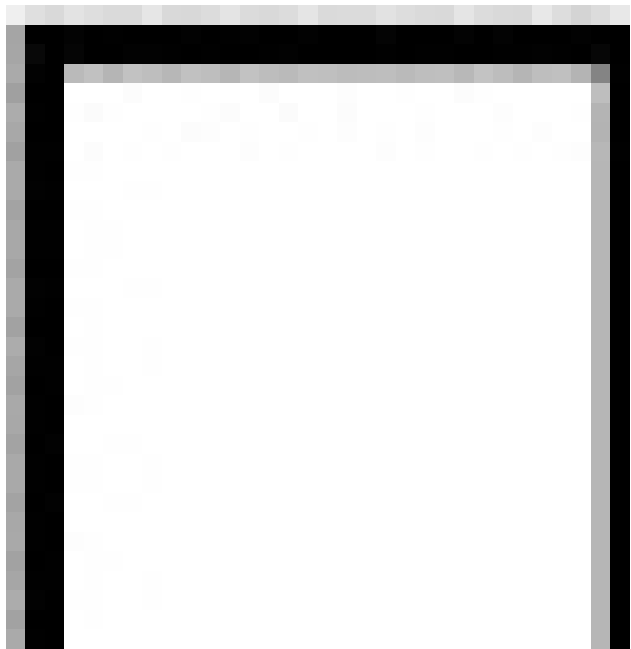
字符的相似性

一个常用的测量刺激物之间相似性的方法是检测它们在认知或识别任务中被混淆的概率：刺激物越相似，它们就越可能被混淆。而混淆概率通常是不对称的，即用b混淆a的概率与用a混淆b的概率不同，这种效应通常被归因于反应偏差（response bias）。为了消除对于非对称性的这种解释，我们可以引入一项实验任务，其中被试仅仅需要说明（按顺序呈现的或同时呈现的）两个刺激物是否相同。约阿夫·科恩（Yoav Cohen）和我采用这一程序研究了某些字符之间的混淆性。

以8个字符为刺激物：



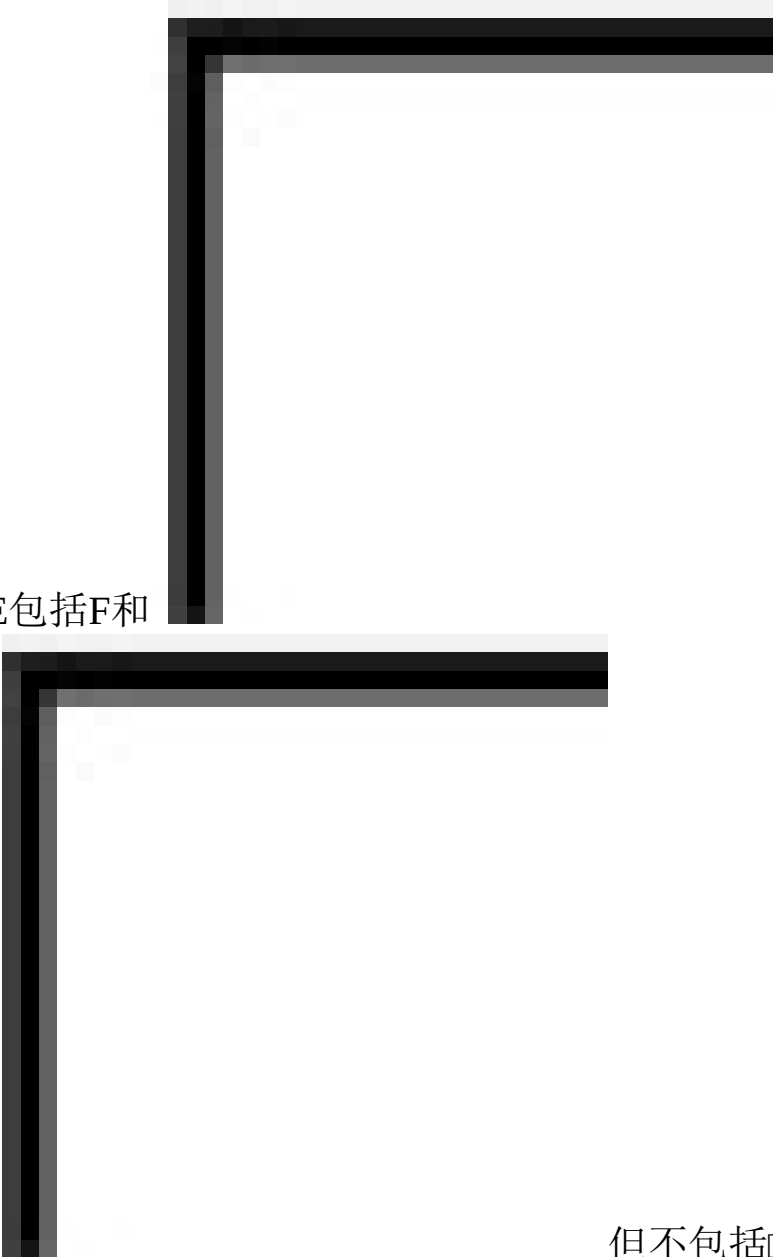
、匚、



、□、F、E、R、B。在嘈杂的背景下，成对的字符并排显示在阴极射线管（cathode-ray tube）上。这些字符按顺序呈现，每个字符大约出现1毫秒。右侧字符始终跟在左侧字符之后，间隔为630毫秒。在每次演示之后，被试可按下两个键中的一个来表示两个字符是否相同。

共有32名被试参与了该实验。每名被试单独进行测试。在每次实验中，一个已知字符都将作为标准刺激物。对于一半被试来说，标准刺激物总是出现在左侧，而对于另一半被试来说，标准刺激物总是出现在右侧。8个字符中的每个字符都会在某一组实验中作为标准刺激物出现。每组有10次实验，其中标准刺激物与其余的每个字符成对出现1次，与其自身成对出现3次。完整的实验设计由8组组成，每组做10次实验。每名被试重复参与3次完整的实验设计（即240次实验）。每次设计中各组的顺序和组内的字符顺序都是随机的。

根据目前的分析，人们会将作为主体的变量刺激物与标准刺激物（即参考对象）进行比较。因此，标准刺激物的选择决定了比较的方向。不同字母的显著程度，是根据字符之间的包含关系确定的。因此，如果一个字符包括另一个字符，我们就认为前者的测量值更大。例如，



E包括F和

但不包括□。在19对字符中，其中一个字符都包含另一个字符，用p表示更显著的字符，q表示不太显著的字符。此外，用 $s(a, b)$ 表示被试判断变量刺激物a与标准刺激物b相同的次数所占的百分比。

从对比模型得出， $\alpha > \beta$ ，当变量刺激物被包含在标准刺激物中时，“相同”的反应比例应该大于标准刺激物被包含在变量刺激物中时的比例，即 $s(q, p) > s(p, q)$ 。这一预测得到了数据的证实。所有被试和所有实验 $s(q, p)$ 的平均值为17.1%，而 $s(p, q)$ 的平均值为12.4%。为了进行统计学检验，我们计算了每名被试在所有实验中的 $s(q, p)$ 和 $s(p, q)$ 之间的差异。这种差异性显著大于零， $t(31)$

=4.41, $p<0.001$ 。这些结果表明, 从对比模型推导出的方向非对称性预测也适用于混淆数据而不仅仅是相似性。

信号的相似性

E.Z.罗思科普 (E.Z. Rothkopf) 组织了598名被试, 向其提供了36个摩斯电码信号的所有有序对, 并要求他们指出每对中的两个信号是否相同 (Rothkopf, 1957)。这些信号都以无固定标准的随机顺序呈现。每名被试对其中25%的序对进行了判断。

设 $s(a, b)$ 表示认为有序对 (a, b) “相同”的被试的百分比, 即将第一信号 a 判断为与第二信号 b 相同的被试的百分比。注意, a 和 b 在这里分别指的是第一信号和第二信号, 而不是前一节中的变量刺激物和标准刺激物。显然, 摩斯电码信号在一定程度上是根据时间长度进行排序的。对于时间长度不同的任意一对信号, 以 p 和 q 分别表示该对中较长和较短的元素。

在罗思科普对555对长度不同的信号进行的比较中, 有336个案例的 $s(q, p)$ 超过 $s(p, q)$, 有181个案例的 $s(p, q)$ 超过 $s(q, p)$, 有38个案例的 $s(q, p)$ 等于 $s(p, q)$, 通过符号检验, $p<0.001$ (Rothkopf, 1957)。所有案例中 $s(q, p)$ 和 $s(p, q)$ 之间的平均差异为3.3%, 这也是非常显著的。相关样本的 t 检验结果为 $t(554)=9.17$, $p<0.001$ 。

当我们只考虑其中的那些一个信号是另一个信号的特有子序列的比较时, 不对称效应得到增强, 例如, $\cdot\cdot$ 是 $\cdot\cdot-$ 及 $\cdot\cdot\cdot$ 的子序列。在这种类型的195个案例中, 有128个案例的 $s(q, p)$ 超过 $s(p, q)$, 有55个案例的 $s(p, q)$ 超过 $s(q, p)$, 有12个案例的 $s(q, p)$ 等于 $s(p, q)$, 通过符号检验可得 $p<0.001$ 。在这种条件下, $s(q, p)$ 和 $s(p, q)$ 之间的平均差异是4.7% ($t(194)=7.58$, $p<0.001$)。

M.威什 (M. Wish) 进行了一项后续研究, 该研究遵循的实验范式与上述相同但信号略有不同 (Wish, 1967)。他的信号由3个音调组成, 且由2个静音间隔隔开, 其中每个元素 (即音调或静音) 或短或长。研究人员向被试展示以这种方式形成的所有32对信号, 并要求被试判断每对信号中的两个元素是否相同。

上述分析也适用于威什的数据。在不同长度信号之间的386个比较

中，有241个案例的 $s(q, p)$ 超过 $s(p, q)$ ，117个案例的 $s(p, q)$ 超过 $s(q, p)$ ，28个案例的 $s(q, p)$ 等于 $s(p, q)$ 。这些数据显然是不对称的，在符号测试中 $p < 0.001$ 。 $s(q, p)$ 和 $s(p, q)$ 之间的平均差异为5.9%，这也是非常显著的， $t(385) = 9.23$ ， $p < 0.001$ 。

在罗思科普和威什的研究中，没有先验的方法来确定比较的方向性，或等效地识别主体和参考对象。然而，如果我们接受焦点假设（ $\alpha > \beta$ ）并且假设长信号比短信号更显著（明显），那么观察到的不对称方向表明：第一信号为主体，第二信号为参考对象。因此，根据目前的分析，刺激物的显著顺序和观察到的不对称方向决定了比较的方向性。

罗施的数据

埃莉诺·罗施（Eleanor Rosch）阐述并支持了这样的观点，即感知和语义类别是根据焦点或原型自然形成和被定义的（Rosch, 1973/1975）。由于原型在类别形成的过程中具有特殊作用，她假设：第一，在包含诸如“a本质上是b”这样的模糊限定的句子框架中，焦点刺激物（原型）出现在第二个位置；第二，从原型到变体的感知距离大于从变体到原型的感知距离。为了验证这些假设，罗施使用了3种刺激域：颜色、线条方向和数字（Rosch, 1975）。原型颜色很鲜艳（如纯红色），而变体颜色不鲜艳（如浅红色）或较透明。垂直、水平和对角为线的方向的原型，而其他角度的线为变体。将10的倍数（如10、50、100）作为原型数字，并将其他数字（如11、52、103）视为变体。

这3种刺激域都证实了第一个假设。当看到诸如“_____实际上是_____”的句子时，被试通常将原型放在第二个空白处，而将变体放在第一个空白处。例如，与“100实际上是103”相比，被试更喜欢句子“103实际上是100”。为了检验第二个假设，将第一个刺激物（标准物）放置在半圆形板的原点处，并且让被试将第二个刺激物（变体）放置在板上的某个位置以“表示其对刺激物与固定在原点处的刺激物之间的距离的感知”。正如假设的那样，在3种刺激域中，当原型（而不是变体）被固定在原点时，人们认为刺激物之间的距离明显更小。

如果焦点刺激物比非焦点刺激物更显著，那么罗施的研究结果支持目前的分析。模糊性语句（如“a大致为b”）可以被视为相似性陈述的特定类型。实际上，模糊性数据与相似性陈述的选择完全一致。此外，观

察到的放置距离的非对称性来自目前的非对称分析，以及标准物和变量物在距离放置任务中分别作为参考对象和主体的自然假设。因此，将b放到离a的距离为t的地方，被理解为从b到a的（感知）距离等于t。

罗施将观察到的非对称性归因于不同原型（如完美的正方形或纯红色）在信息处理中的特殊作用（Rosch, 1975）。此外，在本理论中，非对称性由刺激物的相对显著性来解释。因此，它意味着成对的刺激物样本的非对称性不包括原型（如相同形式的两个不同级别的失真）。然而，如果我们以相对意义而不是以绝对意义来理解原型的概念（即a比b更典型）的话，那么针对非对称的两种解释实际上是一致的。

讨论

对比模型和焦点假设的合取意味着存在非对称的相似性。这一预测在以判断方法（如评级）和行为方法（如选择）进行的知觉和概念实验中得到了证实。

在比较任务中，我们观察到了上文中讨论的非对称性，其中被试比较了两个给定的刺激物以确定它们的相似性。我们在生成任务中（即被试被给予单一刺激物并被要求产生最相似的反应），也观察到了非对称性。模式识别、刺激物识别和单词关联的研究都是生成任务的例子。我们在这些研究中观察到的一种常见模式是，更显著的对象更常发生在人们对不显著对象的反应中，而不是相反。例如，相对于“豹子”与“老虎”相关，更可能是“老虎”与“豹子”相关。同样地，W.R.加纳（W.R. Garner）让被试在给定的模式中，从一组给定的光点图形中选择一个相似但不相同的光点图形（Garner, 1974）。他的研究表明，“好”模式通常被选择作为对“坏”模式的反应，而不是相反。

生成任务中的这种非对称性通常被归因于反应的不同程度的可得性。因此，更可能是“老虎”与“豹”相关联而不是反过来，因为“老虎”更常见，因此比“豹”更有效。这个解释可能更适用于被试必须产生实际反应的任务（如在单词关联或模式识别中），而不是仅从某个指定集合中选择一种的情况（如加纳的任务）。

如果不考虑反应可得性的重要程度，现有的理论为我们提供了生成任务中观察到的非对称性的另一个诱因。例如，考虑将以下生成任务转换为问答方案。问题：a像什么？回答：a像b。如果这种解释是有效

的，并且给定的对象 a 是作为主体而不是作为参考对象的，那么我们观察到的生成非对称性就来自当前的理论分析，因为每当 $f(B) > f(A)$ ， $s(a, b) > s(b, a)$ 。

总之，来自比较和生成任务的数据似乎揭示了显著性的系统性不对称，其方向由刺激物的相对显著性决定。然而，我们不应该完全摒弃对称性假设。在许多情况下，我们都可以看到对称性假设成立的情况，并且它在许多其他情况下都可以作为有效的近似值。但是，它不能被视为心理相似性的普遍原则。

共同性特征与区别性特征

在现有的理论中，对象的相似性表示为其共同性特征和区别性特征的线性组合或对比。本章研究了这些元素的相对影响及其对相似性和差异性评估之间的关系的的影响。讨论仅涉及对称性的任务，其中 $\alpha = \beta$ ，因此 $s(a, b) = s(b, a)$ 。

特征的提取

第一项研究采用对比模型来预测由被试给出的特征之间的相似性。以12种载人工具作为刺激物：公共汽车、汽车、卡车、摩托车、火车、飞机、自行车、船、电梯、推车、筏和雪橇。一组48名被试按照从1（无相似性）到20（最大相似性）的等级评定所有66对工具之间的相似性。我们参照罗施和C. B.梅尔维斯（C. B. Mervis）的方法，让第二组40名被试在70秒内列出每辆车的特征（Rosch & Mervis, 1975）。不同的被试看到的刺激物的顺序不同。

每种工具的特征数量不等，飞机数量最多，有71种，雪橇数量最少，有21种。被试共列出了324个特征，其中224个是特有的，而另外100个特征是两个或更多种工具共有的。对于每种工具，我们计算了两者（至少一个主体）共有的特征数量，以及属于一种工具但不属于另一种工具的特征数量。计算每名被试列出的不同工具的共同性或区别性特征的频率。

为了从列出的特征中预测车辆之间的相似性，必须定义其拥有的共同性特征和区别性特征。最简单的衡量方法是计算被试给出的共同性特征和区别性特征的数量。刺激物的（平均）相似性与其共同性特征的数

量之间的积差相关性为0.68。物体相似性与其区别性特征数量之间的相关性为-0.36。相似性与共同性特征及区别性特征的数量之间的多重相关性（相似性与对比模型之间的相关性）是0.72。

无论被提及的频率如何，我们在计算中都赋予了所有特征相同的权重。考虑到这个因素，我们使 X_a 表示将特征 X 归因于对象 a 的主体的比例，让 N_x 表示共同性特征 X 的数量。对于任何 a 和 b ，用 $f(A \cap B) = \sum X_a X_b / N_x$ 定义它们的共同性特征，其中总和超过 $A \cap B$ 中的所有 X ，并且它们的区别性特征为

$$f(A-B) + f(B-A) = \sum Y_a + \sum Z_b$$

其中总和范围超过所有 $Y \in A-B$ 和 $Y \in B-A$ ，即 a 和 b 的区别性特征。相似性与上述计算中的共同性特征之间的相关性是0.84；相似性与上述计算中的区别性特征之间的相关性为-0.64。相似性与共同性特征及区别性特征之间的多重相关性是0.87。

请注意，用于定义 f 的上述方法仅基于提取的特征，并且根本没有使用相似性数据。在这些条件下，不应预计两者之间存在完美的相关性，因为与特征相关联的权重对于相似性的预测不是最佳的。某一特征可能会因为它易于标记或回忆而经常被提及，尽管它对相似性并没有很大的影响，反之亦然。实际上，当使用叠加树程序对特征进行衡量时（即特征尺度是由对象之间的相似性得出的），数据和模型之间的相关性达到了0.94（Sattath & Tversky, 1977）。

这项研究的结果表明：第一，从被试的语义刺激物中提取细节特征是可能的，如交通工具；第二，所列特征可用于根据对比模型预测相似性，并能获得合理的结果；第三，当我们考虑到提及频率而不仅仅是特征数量时，相似性的预测可以得到改善（Rosch & Mervis, 1975）。

相似性与差异性

一般情况下，人们会认为相似性和差异性的判断是互补的；也就是说，判断差异性判断相似性的线性函数，斜率为-1。一些研究证实了这一假设。例如，在J.霍斯曼（J. Hosman）和T.库纳帕斯（T. Kuennapas）对所有小写字母进行的以0~100为计量方式的描述性研究中，他们观察到了相似性和差异性的独立判断（Hosman & Kuennapas

1972)。判断之间的积差相关性为-0.98，并且回归线的斜率为-0.91。我们还使用20分评定量表收集了被试针对21组国家的相似性和差异性判断。在所有情况下，对每组国家的两个判断的总和非常接近20。评分之间的积差相关性也是-0.98。然而，相似性和差异性之间的这种反比关系并不总是成立的。

共同性特征的增加提升了相似性并且降低了差异性，而区别性特征的增加则降低了相似性并且提升了差异性。但是，分配给共同性特征和区别性特征的相对权重可能在两个任务中有所不同。在评估对象之间的相似性时，被试可能会更多地关注它们的共同性特征，而在评估对象之间的差异性时，被试可能更多地关注它们之间的区别性特征。因此，共同性特征在前一任务中的相对权重大于在下一任务中的相对权重。

设 $d(a, b)$ 表示 a 和 b 之间的感知差异。假设 d 满足本理论中单调性公理的反向不等式，即每当

$$A \cap B \supset A \cap C,$$

$$A-B \subset A-C \quad \text{和}$$

$$B-A \subset$$

$C-A$ 时， $d(a, b) \leq d(a, c)$ 。此外，假设 s 也满足本理论并假设（为简单起见） d 和 s 都是对称的。因此，根据表示定理，存在非负尺度 f 及非负常数 θ 和 λ ，使得对于所有 a, b, c, e ，当且仅当 $\theta f(A \cap B) - f(A-B) - f(B-A) > \theta f(C \cap E) - f(C-E) - f(E-C)$ 时：

$$s(a, b) > s(c, e)$$

当且仅当 $f(A-B) + f(B-A) - \lambda f(A \cap B) > f(C-E) + f(E-C) - \lambda f(C \cap E)$ 时：

$$d(a, b) > d(c, e)$$

在不损失一般性的对称情况下，与区别性特征相关的权重可以设置

为等于1。因此， θ 和 λ 分别反映了相似性和差异性评估中共同性特征的相对权重。

注意，如果 θ 非常大，那么相似性排序基本上由共同性特征决定。此外，如果 λ 非常小，那么差异性排序主要由区别性特征决定。因此，以下条件满足时，可以得到 $s(a, b) > s(c, e)$ 和 $d(a, b) > d(c, e)$ ：

$$f(A \cap B) > f(C \cap E)$$

且

$$f(A - B) + f(B - A) > f(C - E) + f(E - C)$$

也就是说，如果共同性特征在相似性判断中比在差异性判断中的权重更大，那么具有许多共同性或区别性特征的对象组可以被感知为比具有更少共同性或区别性特征的另一对象组更相似或更不同。

为了验证这一假设，我们在预实验的基础上进行了20组4个国家的测试。每组包括两对国家：一对显著的和一对不显著的。显著的国家对由我们所熟知的国家组成（如美国和苏联）。不显著的国家对包括被试知道但不突出的国家（如突尼斯和摩洛哥）。所有被试看到的20组对象是相同的。一组30名被试在更相似的国家组合中选择，另一组30名被试在差异性更大的国家组合中进行选择。

设 Π_s 和 Π_d 分别表示选择显著国家组更为相似和区别性更大的百分比。如果相似性和差异性互补的，即 $\theta = \lambda$ ，那么对于所有对象组， $\Pi_s + \Pi_d$ 应该等于100。此外，如果 $\theta > \lambda$ ，那么 $\Pi_s + \Pi_d$ 应该超过100。在所有集合中， $\Pi_s + \Pi_d$ 的平均值是113.5，显著大于100，其中 $t(59) = 3.27$ ， $p < 0.01$ 。

而且，平均而言，在相似性和差异性任务中，人们更多地选择具有显著性的国家组（相比于不突出的国家组）。例如，相似组中67%的被试选择联邦德国和民主德国（现已合为德国）比斯里兰卡和尼泊尔更相似，而差异组中70%的被试选择联邦德国和民主德国比斯里兰卡和尼泊尔更不同。这些数据表明共同性和区别性特征的相对权重随任务而变化，并验证了人们在相似性判断（相对于区别性判断）中会更多地关注共同性特征的假设。

情境的相似性

与其他判断一样，相似性取决于情境和参考框架。有时参考的相关框架是被明确指定的，比如：“英语和法语在发音方面有多相似？”“梨和苹果在味道方面有什么相似之处？”但是，一般来说，相关的特征空间未被明确指定，而是从上下文中被推断出来的。

例如，当被试被要求评估美国和苏联之间的相似性时，他们通常会假设相关情境是一个国家的集合，并认为相关的参考框架包括所有政治、地理和文化特征。当然，分配给这些特征的相对权重可能因人而异。对于国家、人物、颜色和声音等自然的、整合的刺激物，相关特征空间的模糊性相对较小。然而，对于人为的、可分离的刺激物，例如颜色和形状、不同的图形或者长度和方向不同的线条，被试有时难以评估整体的相似性，偶尔会倾向于评估一个因素或另一个因素的相似性或通过改变情境来改变属性的相对权重（Shepard, 1964; Torgerson, 1965）。

在目前的理论中，情境或参考框架的变化对应于特征空间的变化。例如，当被要求评估各国之间的政治相似性时，被试可能会只考虑这些国家的政治特征而忽略其他特征，或赋予其他特征零权重。除了由显性和隐性指令引起的对特征空间的这种限制之外，特征的显著性以及对象的相似性也会受到有效情境（如需要考虑的对象集）的影响。为了理解这个过程，让我们来检验一下决定特征的显著性的因素，以及这些特征对评估对象相似性的贡献。

诊断原则

特征的显著性（或尺度）由两种因素决定：强化因素和诊断因素。强化因素指的是增加强度或加大信号噪声比的因素，如灯光的亮度、音调的响度、颜色的饱和度、字母的大小、项目的频率、图片的清晰度或图像的生动性。诊断因素是指特征的分类意义，即基于这些特征进行分类的重要性或普遍性。与强化因素不同，诊断因素对研究中的特定对象高度敏感。例如，特征“真实”在真实的动物组中没有诊断性价值，因为所有真实动物都拥有这个特征，因此不能用来对它们进行分类。然而，如果将对象集扩展到传说中的动物，如半人马、美人鱼或凤凰，那么“真实”就具有了很高的诊断性价值。

当面对一组对象时，人们通常会将它们分别聚类以减少信息负载并利于进一步的处理。而人们选择聚类通常是为了最大化聚类内对象的相似性及不同聚类之间的差异性。因此，对象的添加或删除可以改变剩余对象的聚类。反过来，聚类的变化有望提高新特征的诊断性价值，也因此增加了共享这些特征的对象相似性。我们最好通过具体的例子来对相似性和分组之间的这种关系——诊断性假设进行解释。假设两组由4个结构化简笔笑脸图组成的集合（图3-4），只有一组元素（p和q）会导致它们发生变化。



图3-4 用于检验诊断性假设的两个集合

注：面孔下方显示为选择每张面孔（与目标最相似）的被试百分比。

25名被试需要从连续展示的两个集合中，将每个集合中的4张图片分成两对。集合1的最常见分区是c和p（笑脸），a和b（非笑脸）。集

合2的最常见分区是b和q（皱眉），a和c（非皱眉）。因此，用q替换p改变了a的配对方式：在集合1中a与b配对，而在集合2中a与c配对。

根据上述分析，微笑在集合1中比在集合2中具有更大的诊断性价值，而皱眉在集合2中比在集合1中具有更大的诊断性价值。因此，诊断性假设应该遵循分组的相似性。也就是说，a（中性表达式）与b（皱眉）的相似性应该在集合1中更大，在集合1中它们被分在了一起，而在集合2中它们被分开了。同样地，a和c（微笑）的相似性应该在集合2中更大，在集合2中它们被组合在一起，而在集合1中却不是。

为了检验这个预测，我们给50名被试（分为两个组）展示了集合1和集合2（如图3-4所示），并要求选择下面3个面孔刺激物中的1个（选择集），这些面孔刺激物最相似的部分在顶部（目标）。给出选择集中3个元素中的1个的被试百分比分别呈现在面孔下方。结果证实了诊断性假设：第1组被试比第2组被试更多地选择b，而第2组被试比第1组被试更多地选择c。两种差异在统计学上都是显著的， $p<0.01$ 。此外，用q替换p实际上颠倒了相似性排序：在集合1中，b与a更相似而不是c与a更相似；而在集合2中，c与a更相似而不是b与a更相似。

接下来，我们使用语义而不是视觉刺激物进行了更具有普遍性的诊断性假设检验。实验设计基本相同，只是用国家而不是面孔作为刺激物。实验构建了形式为{a, b, c, p}和{a, b, c, q}的以4个国家为一组的20对匹配组。两个匹配集的示例如图3-5所示。

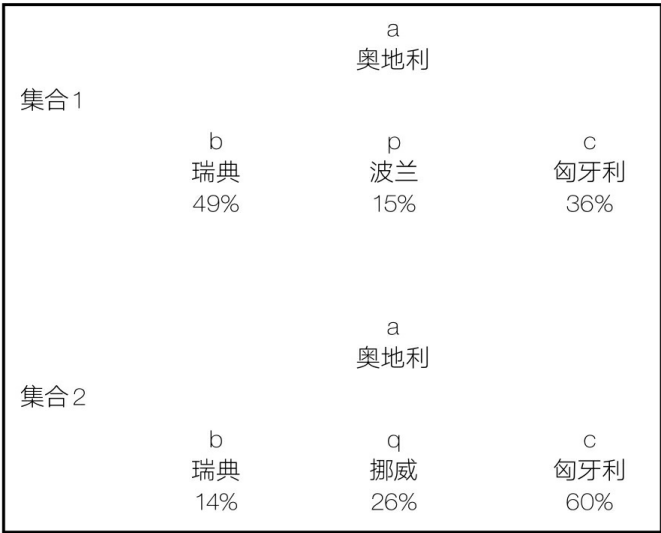


图3-5 用于检验诊断性假设的两个国家集合

注：国家下方的显示为选择与奥地利最相似的国家被试百分比。

请注意，两个匹配的集合（1和2）仅相差一个元素（p和q）。这些集合的构建使得a（奥地利）可能与第1组中的b（瑞典）和第2组中的c（匈牙利）分组到一起。为了验证这一假设，我们向两个由25名被试组成的实验组呈现了4个国家，并要求他们将每4个国家分为2对。每组接收4个国家组中两两匹配的一种情况，它们以随机顺序连续显示。结果证实了我们先前关于国家分组的假设。在每种情况下，用q替换p改变了目标国家在预测方向上的配对，通过检验发现 $p < 0.01$ 。例如，第1组中60%的被试将奥地利与瑞典配对，第2组中96%的被试将奥地利与匈牙利配对。

为了检验诊断性假设，我们向由35人组成的两组被试展示了由20个国家中的每4个为一组形成的刺激物，如图3-5所示。我们要求这些被试从每4个国家组成的集合中选择与目标国家最相似的国家。每组被试只看到由4个国家构成的组合中的一对国家。如果b与a的相似性与选择集无关，那么无论选择集中的第3个元素是p还是q，认为（b而不是c）与a最相似的被试比例应该是相同的。例如，选择瑞典（而不是匈牙利）与奥地利最相似的被试比例应该与选择集中的其他国家是挪威还是波兰无关。

相反，诊断性假设意味着由其他元素的替换引起的分组变化将以可预测的方式改变相似性。回想一下，在集合1中，波兰与匈牙利、奥地利与瑞典配对，而集合2中瑞典与挪威、奥地利与匈牙利配对。因此，集合1中选择瑞典而不是匈牙利（与奥地利最相似）的被试比例应该高于集合2。图3-5中的数据强烈支持这一预测，这表明在集合1中更多人选择了瑞典而不是匈牙利，而集合2中匈牙利被选择的频率高于瑞典。

当选择集中的单个元素为p时，设 $b(p)$ 表示选择国家b与a最相似的被试的百分比，依此类推。如上所示，选择符号使得b通常与q分为一组，c通常与p分为一组。因此， $b(p) - b(q)$ 和 $c(q) - c(p)$ 的差异反映了其他元素p和q，对b和c以及目标a的相似性的影响。在没有情境影响的条件下，两者的差异性应该等于零，而在诊断性假设下，两者的差异性应该是大于零的。例如，在图3-5中， $b(p) - b(q) = 49 - 14 = 35$ ， $c(q) - c(p) = 60 - 36 = 24$ 。所有4元组的平均差异为9%，显著大于零（ $t(19) = 3.65$ ， $p < 0.01$ ）。

实验形式的变化并没有改变实验的结果。当每个选择集包含4个元素而不是3个元素时；被试根据选择集中每个元素与目标的相似性对其进行排序，而不是选择最类似的元素，以及当目标由两个元素组成，并

且让被试指出选择集中的某个元素与两个目标元素最相似时，诊断性假设都得到了验证。有关详细信息，请参阅特沃斯基和加蒂的研究（Tversky & Gati, 1978）。

外延效应

回想一下，特征的诊断性是由基于它们的分类决定的。我们不能根据可被所有对象共享的特征来对这些对象进行分类，因此可以说这些特征缺乏诊断性价值。当通过扩展对象集来扩展情境时，原始情境中的所有对象共享的某些特征可能不会被更广泛的情境中的所有对象共享。此时，这些特征就获得了诊断性价值，并且增加了共享它们的对象之间的相似性。因此，原始情境中的一组对象的相似性通常小于它们在扩展情境中的相似性。

在研究不同动物和不同乐器之间的相似性时，萧伯格提出并验证了本质上相同的说法（Sjöberg, 1972）。例如，萧伯格指出，当一个管乐器（单簧管）被添加到弦乐器组（班卓琴、小提琴、竖琴、电吉他）时，它们之间的相似性增加了。由于弦乐器比起单簧管来说彼此更相似，上述结果至少可以部分归因于被试倾向于将反应量表标准化，也就是说，对任何一组进行比较都会产生相同的平均相似性。

在接下来的研究中，我们采用了稍微不同的实验设计以消除这种效应。所有被试看到的是几组接壤的国家，并让他们以20分制对这些国家的相似性进行评分。我们一共设置了4组，每组包含8对国家。第1组包含8对欧洲国家（如意大利-瑞士）。第2组包含8对美洲国家（如巴西-乌拉圭）。第3组包含的国家是第1组的4对国家和第2组的4对国家，而第4组包含的国家是第1组和第2组剩余的4对国家。4组国家分别被呈现给不同组的30~36名被试。

根据诊断性假设，“欧洲”和“美洲”这两个特征在第1组和第2组中没有诊断性价值，但它们在第3组和第4组中都具有诊断性价值。因此，我们推测差异组（第3组和第4组）中的总体平均相似性高于相似组（第1组和第2组）。该预测也得到了数据的证实（ $t(15) = 2.11$, $p < 0.05$ ）。

在本研究中，所有的相似性评估仅仅涉及同质组（即来自同一大洲接壤的一对国家）。与引入不同类别的比较组的萧伯格研究不同，我们的实验是通过由同质元素组成异质组来扩展情境的。因此，在本研究中

观察到的因情境扩大而增加的相似性，不能用被试倾向于将任意组评估的平均相似性等同起来进行解释。

相似性的两个方面

根据目前的分析，特征的显著性有两个组成部分：强度和诊断性。特征的强度由不同情境中的感知稳定性和认知因素决定。特征的诊断性价值取决于由它进行的分类的普遍性，而且这些分类会随着情境的不同而变化。因此，情境对相似性的影响被视为由对象分组的相应变化引起的特征的诊断性价值的变化。

实验发现支持了这一解释，即分组的变化（由替换或添加对象产生的变化）引起了对象相似性的相应变化。这些结果揭示了相似性和分类之间存在的相互的动态作用。我们通常会假设分类取决于对象之间的相似性。前面的讨论支持相反的假设：对象的相似性会被它们的分类方式改变。因此，相似性有两个方面，即因果和衍生。它是对象分类的基础，但也受到分类方法的影响。作为该过程的基础的诊断性原则，是分析情境对相似性的影响的关键。

讨论

在这一部分中，我们以聚类和树来表征对象，讨论了原型性和家族相似性的概念，并评论了相似性和隐喻之间的关系。

特征、聚类和特征树

对象的特征或属性与其所属的类之间存在着众所周知的对应关系。例如，红色花可以被表征为具有“红色”特征，或者被描述为红色物体类的成员。通过这种方式，我们将 Φ 中的每个特征与 Δ 中具有该特征的对象类相关联。特征和类之间的这种对应关系将本理论与近似数据的聚类表征方法直接联系起来。

在对比模型中，对象之间的相似性被表示为它们的共同性特征和区别性特征的函数。重叠集之间的关系通常用维恩图表示（Venn diagram；如图3-1所示）。然而，当对象的数量超过4或5时，这种表示就会变得很烦琐。为了获得对比模型的图形表征，我们采纳了两种替代性的简化方法。

首先，假设所研究的对象在显著性上是相等的，即对于 Δ 中的所有 a 和 b 来说， $f(A) = f(B)$ 。虽然这个假设一般而言并不是绝对有效的，但仍可以在某些情况下作为合理的近似假设。我们同时假设特征具有可加性和对称性，可以得到：

$$S(a,b) = \theta f(A \cap B) - f(A-B) - f(B-A) =$$

$$\theta f(A \cap B) + 2f(A \cap B) - f(A-B) - f(B-A) - 2f(A \cap B) =$$

$$(\theta+2)f(A \cap B) - f(A) - f(B) =$$

$$\lambda f(A \cap B) + \mu$$

由于 Δ 中的所有 a, b ， $f(A) = f(B)$ 。因此，在目前的假设下，对象之间的相似性是其共同性特征的线性函数。

由于 f 是附加尺度，因此 $f(A \cap B)$ 可表示为属于 a 和 b 的所有特征之和。对于 Δ 的每个子集 Λ ，令 $\Phi(\Lambda)$ 表示由 Λ 中的所有对象共享的特征集，并且不属于 Λ 的对象不得共享。因此

$$S(a, b) = \lambda f(A \cap B) + \mu =$$

$$\lambda (\sum f(X)) + \mu \quad X \in f(A \cap B) =$$

$$\lambda (\sum f(\phi(\Lambda))) + \mu \quad \Lambda \supset \{a, b\}$$

由于求和范围包括 a 和 b 两者在内的 Δ 的所有子集，对象之间的相似性可以表示为与包括两个对象的所有集合相关联的权重之和。

该形式与R. N.谢泼德（R. N. Shepard）和P.阿拉布尔（P. Arabie, 1975）提出的可加性聚类模型基本相同。这些调查人员开发了一个计算

机程序ADCLUS，它选择了一个相对较小的子集，并为每个子集分配了权重，以便最大化模型所占的（相似性）方差的比例。谢泼德和阿拉布尔将ADCLUS应用于几项研究，包括谢泼德、基尔帕特里克和J. P.卡宁哈姆（J. P. Cunningham）关于整数0到9之间的抽象数字特征的相似性判断（Shepard, Kilpatric & Cunningham, 1975）。具有19个子集的解决方案占方差的95%。表3-1显示了9个主要（权重最大的）子集及其建议性解释。所有主要子集都很容易解释，它们是重叠的而不是分层的。

表3-1 通过ADCLUS进行的整数0到9的相似性分析

级别	权重	子集的元素	解释子集
1	0.305	2 4 8	2 的次方
2	0.288	6 7 8 9	较大的数
3	0.279	3 6 9	3 的倍数
4	0.202	0 1 2	很小的数
5	0.202	1 3 5 7 9	奇数
6	0.175	1 2 3	小的非零数
7	0.163	5 6 7	中间的数（较大的）
8	0.160	0 1	加法和乘法的数
9	0.146	0 1 2 3 4	较小的数

以上模型仅表示共同性特征方面的相似性。或者，可以仅根据区别性特征来表达相似性。什穆埃尔·萨塔特（Shmuel Sattath）已经证明（Sattath, 1976），对于任何具有可加性的尺度 f 的对称性对比模型，在相同的特征空间上存在定义尺度 g ，使得：

对于某些 $\lambda > 0$

$$S(a, b) = \theta f(A \cap B) - f(A - B) - f(B - A) =$$

$$\lambda - g(A - B) - g(B - A)$$

该结果表明当特征空间 Φ 是特征树时，即每当 Δ 中的任意3个对象都可以被标记为 $A \cap B = A \cap C \subset B \cap C$ 时，就可以简单地表示不相似性。由什穆埃尔·萨塔特和特沃斯基根据库纳帕斯和詹森

（Kuennapas & Janson）提出的判断小写字母的相似性构建的特征树例子如图3-6所示（Sattath & Tversky, 1977）。主要分支被标注出来，以便对特征树进行解释。

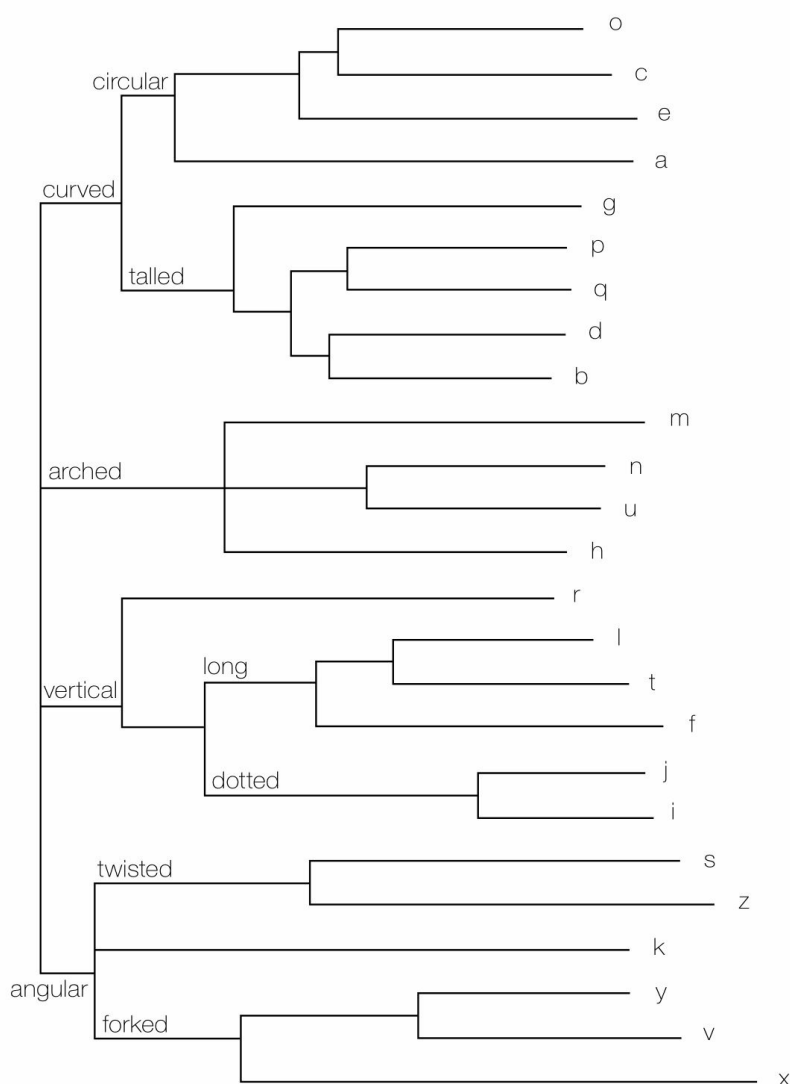


图3-6 用可加性（特征）树表示字母的相似性

图中的每段（水平）弧表示由顺着这条弧线的所有对象（字母）共享的特征集，并且弧长对应于该集的尺度。对象的特征是通向该对象的所有弧的特征，而且其大小就是它到根部的（水平）距离。对象a和b之间的树的距离是连接它们的路径的（水平）长度，即 $f(A-B) + f(B-A)$ 。因此，如果对比模型成立， $\alpha = \beta$ ，并且 Φ 是树形，那么非相似性（即 $-S$ ）可表示为树的距离。

特征树也可以被解释为分层聚类模型，其中每个弧长表示由该弧后面的所有对象组成聚类的权重。图3-6中的树与一般分级聚类树的不同之处在于分支的长度不同。萨塔特和特沃斯基描述了一个计算机系统 ADDTREE，该系统可根据相似性数据构造可加性特征树，并讨论了它与其他缩放方法的关系（Sattath & Tversky, 1977）。

从上面的讨论中可以很容易得出，如果我们假设特征集 Φ 是特征树，并且对于 Δ 中的所有 a, b , $f(A) = f(B)$ ，那么对比模型可简化为众所周知的层次聚类方法。因此，可加性聚类模型（Shepard & Arabie, 1975）、可加性相似树（Sattath & Tversky, 1977）和层次聚类方法（Johnson, 1967）是对比模型的所有特例。因此，这些缩放模型可用于发现研究对象的共同性特征和区别性特征。而我们的研究为分析近似表征的集-理论方法提供了理论基础。

相似性、原型性和家族相似性

相似性是两个物体之间的接近关系。在对象和类之间存在着其他接近关系，如原型性和代表性。直观地说，如果一个对象是它所属类别所有特征的例证，那么该对象就是该类别的原型。请注意，原型不一定是同类中最典型或最常见的。最近的研究已经证明了原型性或代表性在感知学习（Posner & Keele, 1968; Reed, 1972）、归纳推理（Kahneman & Tversky, 1973）、语义记忆（Smith, Rips & Shoben, 1974）以及类别形成（Rosch & Mervis, 1975）中的重要性。在接下来的讨论中，我们将用现有的相似理论来分析原型性与家族相似性的关系。

设 $P(a, \Lambda)$ 表示对象 a 相对于类 Λ 的原型性（程度），基数为 n ，由下式定义：

$$P(a, \Lambda) = P_n (\lambda \sum f(A \cap B) - \sum (f(A - B) + f(B - A)))$$

其中总和超过 Λ 中的所有 b 。因此， $P(a, \Lambda)$ 被定义为与 Λ 中的元素共享 a 的特征及与 Λ 中的元素不共享 a 的特征的线性组合（对比度）。如果 Λ 中的元素 a 使 $P(a, \Lambda)$ 最大化，那么它就是原型——一个类可能有多个原型。

因子 p_n 反映了类别大小对原型的影响，而常数 λ 决定了共同性和区别性特征的相对权重。如果 $p_n = 1/n$, $\lambda = 0$, 并且 $\alpha = \beta = 1$, 那么 $P(a, \Lambda)$

$=1/n \sum S(a, b)$ （相对于 Λ ， a 的原型性等于 a 与 Λ 中所有成员的平均相似性）。然而，与前面讨论的焦点假设一致，共同性特征在原型性判断中的权重比在相似性判断中似乎更高。

罗施和梅尔维斯发现了一些关于测量效度的证据（Rosch & Mervis, 1975）。他们从6个类别（家具、车辆、水果、武器、蔬菜、服装）中选取20个对象，让被试列出与每个对象相关联的属性。对象的原型性由其与该类别中每个成员共享的属性或特征的数量来定义。因此，关于 Λ 的原型 a 由 $\sum N(a, b)$ 定义，其中 $N(a, b)$ 表示由 a 和 b 共享的特征的数量，并且求和范围超过了 Λ 中的所有 b 。显然，罗施和梅尔维斯采用的原型尺度是建议尺度的一个特例，其中 λ 很大且 $f(A \cap B) = N(a, b)$ 。

研究者还通过指导被试根据其符合“类别含义的想法或图像”的程度，以7分制对每个对象进行评分，从而对原型性进行直接测量。这些评级与上述测量的排序关系在所有类别中都很高：家具，0.88；车辆，0.92；武器，0.94；水果，0.85；蔬菜，0.84；服装，0.91。因此，某一类别中成员的原型性可通过其与该类别中其他成员共享的特征数量来预测。

与认为自然类别是通过结合关键特征来定义的观点相反，维特根斯坦认为，一些自然类别（如游戏）并不具有所有成员共享的属性（Wittgenstein, 1953）。他提出，自然类别和概念通常是根据家族相似性来得以表示和理解的，即连接种类中不同成员的相似关系的网络。罗施及其合作者的研究有效地强调了家族相似性在类别形成和发展中的重要性（Rosch, 1973；Rosch & Mervis, 1975；Rosch, Mervis, Gray, Johnson & Boyes-Braem, 1976）。这项研究表明，对于自然类别和人工类别，人们通常以原型（或焦点元素）以及与原型的接近程度来进行理解和组织。此外，这一研究也实质性地证明，人们是基于最佳抽象水平的基本语义类别来构建自己的世界的。例如，椅子是一个基本类别；家具太笼统，而厨房椅子则太具体了。同样，汽车是一个基本类别；车辆太笼统，轿车则太具体。罗施认为选择基本类别是为了最大化由提示的效度定义的家族相似性。

目前的研究提出了以下衡量家族相似性或类别相似性的指标。设 Λ 是 Δ 的一些子集，基数为 n 。 Λ 的类别相似性表示为 $R(\Lambda)$ ，其由下式定义：

$$R(\Lambda) = r_n (\lambda \sum f(A \cap B) - \sum (f(A - B) + f(B - A)))$$

其总和超过 Λ 中所有a、b。因此，类别相似性是该类别中所有成对对象的共同性特征和区别性特征的线性组合。因子 r_n 反映了类别大小对类别相似性的影响，而常数 λ 决定了共同性特征和区别性特征的相对权重。如果 $\lambda = \theta$ ， $\alpha = \beta = 1$ ，并且 $r_n = 2 / n(n-1)$ ，那么

$$R(\Lambda) = \frac{\sum S(a, b)}{\binom{n}{2}}$$

总和超过 Λ 中所有的a和b。也就是说，类别相似性等于 Λ 内所有元素之间的平均相似性。虽然家族相似性的建议尺度与罗施提出的不同，但也捕捉到了其理论的基本概念，即家族相似性是最高的类别，这些类别“具有该类别内所有元素共有的最多属性，且与其他类别的元素可共享的属性最少”（Rosch et al., 1976）。

类别相似性的最大化可用于解释类别的形成。因此，只要 $R(\Lambda) > R(\Gamma)$ ，就应将集合 Λ 而不是 Γ 作为自然类别。同样地，每当 $R(\{\Lambda \cup a\}) > R(\Lambda)$ 时，可将对象a添加到类别 Λ 中。优选（基本）类别既非最具包容性的也非最特殊的这一事实，对 r_n 设定了某些限制。

如果 $r_n = 2 / n(n-1)$ ，那么 $R(\Lambda)$ 等于 Λ 所有成员之间的平均相似性。那么这一指数将使人们选择最小的类别，因为我们通常可以通过删减对象来增加平均相似性。例如，轿车之间的平均相似性肯定大于汽车之间的平均相似性；然而，汽车相对于轿车是一个更基本的类别。如果 $r_n = 1$ ，则 $R(\Lambda)$ 等于 Λ 所有成员之间的相似性之和。那么这一指数将使人们选择最大的类别，因为如果S是非负的，那么添加对象会增加总相似性。

因此，为了解释中间类别的形成，类别相似性必须是平均值和总和之间的折中。也就是说， r_n 必须是超过 $2 / n(n-1)$ 的n的递减函数。在这种情况下，每当平均相似性为常数时， $R(\Lambda)$ 随着类别大小而增加，反之亦然。因此，类别扩展的显著增加可能会超过平均相似性的小幅下降。

尽管相似性、原型性和家族相似性的概念密切相关，但它们之前并未以正式的外显方式相关联。目前的发展为在统一框架下的相似性、原型性和家族相似性提供了解释，在这个框架中，它们被视为共同性特征和区别性特征集的对比或线性组合。

明喻和隐喻

明喻和隐喻是创造性言语表达的基本要素。也许隐喻表达最有趣的特性在于，尽管它们具有新颖性和非文字性，但它们通常是可以理解的并且信息含量丰富。例如，“X先生像推土机”很容易被理解为X先生是一个勇敢的、有实力的人，他克服了工作中的所有障碍。对内涵的充分分析应该考虑到人们在先前没有具体学习过的特定情况下解释隐喻的能力。由于这些表达所传达的信息往往具有针对性和特殊性，因此不能用一些内涵意义的广义维度来解释，如评价或效力（Osgood, 1962）。人们通过扫描特征空间并选择适用于该主题的指示对象的特征来解释明喻（如选择特征适用于某个人的推土机）。这个过程的本质有待解释。

相似性评估与隐喻解释之间存在着密切联系。在相似性判断中，我们要假定特定的特征空间或参考框架，并评估主体和指示对象之间的匹配质量。在对明喻的解释中，人们假设主体和指示对象之间存在相似性，并寻找对空间的解释，以使匹配质量最大化。因此，相同的一对物体被视为相似或不同，取决于参照系的选择。

良好隐喻的一个特征是先验、字面解释与后验、隐喻解释之间的对比。过于明显的隐喻是无趣的；晦涩难懂的隐喻又是无法理解的。一个好的隐喻就像一个好的侦探故事。为了保持读者的兴趣，答案不应该提前给出，而应该在保持故事的连贯性之后给出。可以想象一下“一篇文章就像一条鱼”这个比喻。这句话令人费解。一篇文章不会具有腥、滑或湿这样的特点。但当我们想到一篇文章（像鱼一样的）有头有尾，并且偶尔还会在结尾处来一记翻尾时，那么这个困惑就会解除。

本章内容得益于约阿夫·科恩、伊塔马尔·加蒂、丹尼尔·卡尼曼、萧伯格的热烈讨论。

参考文献

Beals, R., Krantz, D. H., & Tversky, A. (1968). Foundations of

multidimensional scaling. *Psychological Review*, 75, 127–142.

Bush, R. R., & Mosteller, F. (1951). A model for stimulus generalization and discrimination. *Psychological Review*, 58, 413–423.

Carroll, J. D., & Wish, M. (1974). Multidimensional perceptual models and measurement methods. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception: Psychophysical judgment and measurement* (Vol. 2, pp. 391–447). New York, NY: Academic Press.

Eisler, H., & Ekman, G. (1959). A mechanism of subjective similarity. *Acta Psychologica*, 16, 1–10.

Garner, W. R. (1974). *The processing of information and structure*. New York, NY: Halsted Press.

Gibson, E. (1969). *Principles of perceptual learning and development*. New York, NY: Appleton-Century-Crofts.

Goldmeier, E. (1972). Similarity in visually perceived forms. *Psychological Issues*, 8, 1–136.

Gregson, R. A. M. (1975). *Psychometrics of similarity*. New York, NY: Academic Press.

Hosman, J., & Kuennapas, T. (1972). *On the relation between similarity and dissimilarity estimates* (Report No. 354). University of Stockholm, Psychological Laboratories.

Jakobson, R., Fant, G. G. M., & Halle, M. (1961). *Preliminaries to speech analysis: The distinctive features and their correlates*. Cambridge, MA: MIT Press.

Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika*, 32, 241–254.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237–251.

- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. 1). New York, NY: Academic Press.
- Krantz, D. H., & Tversky, A. (1975). Similarity of rectangles: An analysis of subjective dimensions. *Journal of Mathematical Psychology*, 12, 4–34.
- Kuennapas, T., & Janson, A. J. (1969). Multidimensional similarity of letters. *Perceptual and Motor Skills*, 28, 3–12.
- Neisser, U. (1967). *Cognitive psychology*. New York, NY: Appleton-Century-Crofts.
- Osgood, C. E. (1962). Studies on the generality of affective meaning systems. *American Psychologist*, 17, 10–28.
- Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, 77, 353–363.
- Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, 3, 382–407.
- Restle, F. (1959). A metric and an ordering on sets. *Psychometrika*, 24, 207–220.
- Restle, F. (1961). *Psychology of judgment and choice*. New York, NY: Wiley.
- Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 111–144). New York, NY: Academic Press.
- Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology*, 7, 532–547.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–603.
- Rosch, E., Mervis, C. B., Gray, W., Johnson, D., & Boyes-Braem, P.

(1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382–439.

Rothkopf, E. Z. (1957). A measure of stimulus similarity and errors in some paired-associate learning tasks. *Journal of Experimental Psychology*, 53, 94–101.

Sattath, S. *An equivalence theorem*. Unpublished note, Hebrew University, 1976.

Sattath, S., & Tversky, A. (1977). Additive similarity trees. *Psychometrika*, 42, 319–345.

Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1, 54–87.

Shepard, R. N. (1974). Representation of structure in similarity data: Problems and prospects. *Psychometrika*, 39, 373–421.

Shepard, R. N., & Arabie, P. Additive cluster analysis of similarity data. *Proceedings of the U. S. –Japan Seminar on Theory, Methods, and Applications of Multidimensional Scaling and Related Techniques*. San Diego, August 1975.

Shepard, R. N., Kilpatrick, D. W., & Cunningham, J. P. (1975). The internal representation of numbers. *Cognitive Psychology*, 7, 82–138.

Sjöberg, L. A cognitive theory of similarity. *Göteborg Psychological Reports* (No. 10), 1972.

Smith, E. E., Rips, L. J., & Shoben, E. J. (1974). Semantic memory and psychological semantics. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 1–45). New York, NY: Academic Press.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 81, 214–241.

Torgerson, W. S. (1965). Multidimensional scaling of similarity. *Psychometrika*, 30, 379–393.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, 79, 281–299.

Tversky, A., & Gati, I. (1978). Studies of similarity. In E. Rosch & B. Lloyd (Eds.), *On the nature and principle of formation of categories* (pp. 79–98). Hillsdale, NJ: Erlbaum.

Tversky, A., & Krantz, D. H. (1970). The dimensional representation and the metric structure of similarity data. *Journal of Mathematical Psychology*, 7, 572–597.

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Wish, M. (1967). A model for the perception of Morse Code-like signals. *Human Factors*, 9, 529–540.

Wittgenstein, L. (1953). *Philosophical investigations*. New York, NY: Macmillan.

附录：相似性的定理理论

令 $\Delta=\{a, b, c, \dots\}$ 是特征集对象的集合，并且让 A 、 B 、 C 分别表示与 a 、 b 、 c 相关联的特征集。设 $s(a, b)$ 是 a 到 b 的相似性的序数，并将其定义为所有在 Δ 中的不同 a 、 b 。本理论基于以下5个定理。由于论文中讨论了前3个，因此仅将其列出；同时简要讨论其余的定理。

1. 匹配。 $s(a, b) = F(A \cap B, A - B, B - A)$ ，其中 F 是在3个参数中的实值函数。

2. 单调性。每当 $A \cap B \supset A \cap C$,

$$A-B \subset A-C \quad \text{和}$$

$$B-A \subset C-A$$

时， $s(a, b) \geq s(a, c)$ 。此外，如果任何一个包含关系都是正确的，那么这种不等式就是严格的。

设 Φ 是与 Δ 的对象相关的所有特征的集合，并且令 X, Y, Z 等表示 Φ 的子集。可将表达式 $F(X, Y, Z)$ 定义为，当 Δ 中存在 a, b 时，使得 $A \cap B = X, A - B = Y, B - A = Z$ ，其中 $s(a, b) = F(X, Y, Z)$ 。如果以下等式中的一个或多个成立， $X, Y, Z: F(V, Y, Z) = F(W, Y, Z), F(X, V, Z) = F(X, W, Z), F(X, Y, V) = F(X, Y, W)$ ，那么定义 $V = W$ 。只要满足以下式子中的1个、2个或3个： $(A \cap B) \approx (C \cap D), (A - B) \approx (C - D), (B - A) \approx (D - C)$ ，则 (a, b) 和 (c, d) 在1个、2个或3个元素上分别一致。

3. 独立性。假设组合 (a, b) 和 (c, d) 及组合 (a', b') 和 (c', d') 在相同的两个分量上是一致的，而组合 (a, b) 和 (a', b') ，以及组合 (c, d) 和 (c', d') 与剩余的（第三个）元素达成一致。则

当且仅当 $s(c, d) \geq s(c', d')$ 时， $s(a, b) \geq s(a', b')$ 。

4. 可解性。第一，对于 Δ 中对象的所有组合 $(a, b) (c, d) (e, f)$ ，存在一对组合 (p, q) ，它们分别与第一、第二、第三个元素一致，即 $P \cap Q \approx A \cap B, P - Q \approx C - D, Q - P \approx F - E$ 。

第二，假设 $s(a, b) > t > s(c, d)$ ，则存在 e, f ，其中 $s(e, f) = t$ ，使得如果 (a, b) 和 (c, d) 在1个或2个元素上一致，那么 (e, f) 在这些元素上与它们一致。

第三，在 Δ 中存在成对的对象 (a, b) 和 (c, d) ，它们在任何元素上都不一致。

与其他定理不同的是，可解性不会对相似性顺序施加约束；它只是声称所研究的结构足够丰富，因此可以满足某些公式。定理4的第一部分类似于阶乘结构。定理的第二部分暗示了 s 的范围是实际区间：在 Δ 中

存在某对象，其相似性可匹配任何由两个相似性限定的实际值。定理4的第三部分确保了 F 的所有参数都是必不可少的。

令 Φ_1 、 Φ_2 和 Φ_3 分别作为 F 的第一、第二或第三自变量出现的特征集（注意， $\Phi_2=\Phi_3$ ）。假设 X 和 X' 属于 Φ_1 ，而 Y 和 Y' 属于 Φ_2 。每当两个区间匹配时，即每当 Δ 中存在同样类似的成对的对象 (a, b) 和 (a', b') 与第三个因子一致时，则定义 $(X, X')_1 \approx (Y, Y')_2$ 。因此，当

$$s(a, b) = F(X, Y, Z) = F(X, Y', Z') = s(a', b')$$

时

$$(X, X')_1 \approx (Y, Y')_2$$

该定义很容易扩展到任何其他因子中。接下来，对于某些 $(X, X')_j$ ， $j \neq i$ ，当 $(V, V')_i \approx (X, X')_j \approx (W, W')_i$ 时，定义 $(V, V')_i \approx (W, W')_i$ ，其中 $i=1, 2, 3$ 。因此，如果在另一个因子上匹配相同的区间，那么相同因子上的两个区间是等效的。下面的不变性定理提出，如果两个区间在一个因子上是等价的，那么它们在另一个因子上也是等价的。

5. 不变性。假设 V 、 V' 、 W 、 W' 属于 Φ_i 和 Φ_j ， $i, j=1, 2, 3$ ，那么

当且仅当 $(V, V')_j \approx (W, W')_j$ 时， $(V, V')_i \approx (W, W')_i$ 。

表示定理

假设定理1至定理5成立，那么存在相似性尺度 S 和非负尺度 f ，使得对于 Δ 中的所有 a, b, c, d ：

- (1) 当且仅当 $s(a, b) \geq s(c, d)$ 时， $S(a, b) \geq S(c, d)$ ；
- (2) 对于某些 $\theta, \alpha, \beta \geq 0$ ， $S(a, b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$ ；
- (3) f 和 S 是区间尺度。

虽然表示定理的自包含证明很长，但该定理可以很容易地被简化为

先前的结果。

回想一下 Φ_i 是 F 的第 i 个参数的一组特征，并令 $\Psi_i = \Phi_i / \approx$ ， $i = 1, 2, 3$ 。因此， Ψ_i 是 Φ_i 相对于 $\approx$ 的等价类集合。从定理1和公式3得出，每个 Ψ_i 都是经明确定义的，并且从定理4得出 $\Psi = \Psi_1 \times \Psi_2 \times \Psi_3$ 等价于 F 的定义域。我们希望证明根据 F 排序的 Ψ 是一个由3个元素组成的可加性合取结构（Krantz, Luce, Suppes & Tversky, 1971）。

然而，这一结果来自特沃斯基和克兰茨提出的可分解的相似性结构的分析（Tversky & Krantz, 1970）。特别要指出的是，文中定理1部分（c）的证明显示，在公理1、3和4中，存在定义在 Ψ_i 上的非负函数 f_i ， $i=1, 2, 3$ ，对于 Δ 中的所有 a, b, c, d ：

当且仅当 $S(a, b) \geq S(c, d)$ 时

$$s(a, b) \geq s(c, d)$$

其中

$$S(a, b) = f_1(A \cap B) + f_2(A - B) + f_3(B - A)$$

并且 f_1 、 f_2 、 f_3 是具有共同单位的区间尺度。

根据定理5，间隔的等价性在因子中保持不变。也就是说，对于 $\Phi_i \cap \Phi_j$ ， $i, j=1, 2, 3$ 中的所有 V, V', W, W' 来说，当且仅当 $f_j(V) - f_j(V') = f_j(W) - f_j(W')$ 时

$$f_i(V) - f_i(V') = f_i(W) - f_i(W')$$

因此，根据克兰茨等人的定理6.15的第一部分，存在标度 f 和常数 θ_i ，使得 $f_i(X) = \theta_i f(X)$ ， $i=1, 2, 3$ （Krantz et al., 1971）。最后，根据定理2， S 在 f_1 中增加并且在 f_2 和 f_3 中减小。因此，可以将其表达为：

对于一些非负常数 θ, α, β ，存在 $S(a, b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$ 。

第4章 前景理论：风险决策分析

丹尼尔·卡尼曼
阿莫斯·特沃斯基

引言

期望效用理论在风险决策的分析中已占据主导地位。它作为理性选择的规范性模型已被普遍接受（Keeney & Raiffa, 1976），并被广泛用作经济行为的描述性模型（Friedman & Savage, 1948; Arrow, 1971）。因此，我们假设所有理性的人都应该遵守该理论的相关定理，而且大多数人在多数情况下都是这样做的（Savage, 1954; von Neumann & Morgenstern, 1944）。

本章将描述几类选择问题，人们在这些問題中的偏好系统性地违背了期望的效用理论。基于这些观察结果，我们认为，人们通常理解的效用理论并不是一个完善的描述性模型，因此我们提出了另外一个理论，帮助人们在风险情况下做选择。

评价

我们可以将风险决策视为前景或博弈之间的选择。一个前景 $(x_1, p_1; \dots; x_n, p_n)$ 是产生结果 x_i 的概率 p_i 的一项合约，其中 $p_1 + p_2 + \dots + p_n = 1$ 。为了简化表示方法，我们省略了空值结果，并用 (x, p) 表示前景 $(x, p; 0, 1-p)$ ，其中 x 的概率为 p ，0 的概率为 $1-p$ 。 (x) 表示肯定能获得 x 的无风险前景。目前的讨论仅限于具有所谓的客观或标准概率的预期。

期望效用理论在前景选择中的应用以下面3个原则为基础。

$$(1) \text{ 前景: } U(x_1, p_1; \dots; x_n, p_n) = p_1 u(x_1) + \dots + p_n u(x_n)。$$

也就是说，一个前景的总体效用 U ，就是其结果的预期效用。

(2) 资产整合：当且仅当 $U(w+x_1, p_1; \dots; w+x_n, p_n) > u(w)$ 时， $(x_1, p_1; \dots; x_n, p_n)$ 在资产量为 w 时是可接受的。

也就是说，如果前景与某个人的资产整合所产生的效用超过了这些资产本身的效用，那么这个前景是可以接受的。因此，效用函数的取值是最终的状态（包括一个人的资产状况）而不是收益或损失。

虽然效用函数的使用范围不限于任何特定类别的结果，但该理论的大多数应用都与货币结果相关。此外，大多数经济学的应用都引入了下面这个额外假设。

(3) 风险规避： u 是凹函数（ $u'' < 0$ ）。

如果一个人更倾向于确定性前景（ x ）而不是具有前景值 x 的任何风险前景，那么他就是风险规避的。在期望效用理论中，风险规避等于效用函数的凹度。风险规避的普遍性可能是关于风险选择最广为人知的概括。早在18世纪，决策理论家就指出，效用是货币的凹函数，这种思想一直沿用至今（Pratt, 1964; Arrow, 1971）。

在接下来的部分中，我们列出了一些违反期望效用理论的现象。这些例子都是基于学生和大学教师对假设性选择问题的回答。被试被要求回答如下类型的问题：

你更喜欢下列选项中的哪一个？

A. 你有 50% 的机会赢得 1 000 新谢克尔，有 50% 的机会什么也得不到。

B. 你一定能得到 450 新谢克尔。

这里的结果指的是以色列货币。为了表明所涉金额的大小，这里要说明，一个家庭的月净收入中位数约为3 000新谢克尔。研究人员要求被试想象一下，他们实际面临这些问题时，会做出什么样的选择。这些回答都是匿名的，其说明中明确指出，对于这类问题没有“正确”的答案，这项研究的目的是观察人们在风险前景中是如何做出选择的。研究中的问题以问卷形式呈现，每本小册子最多有十几个问题。问卷有多种形式，这样可以使被试以不同的顺序看到问题。此外，每个问题都使用了两个版本，并互换了不同前景的位置。

本章提到的问题是我们选择出来的，用来说明一系列的效应。我们已经在一些结果和概率不同的问题中观察到了各种效应。这些问题也呈现给了斯德哥尔摩大学和密歇根大学的学生和教师。其结果与以色列被试的答题结果基本相同。

由于用到的是虚构的前景，所以被试难免会对方法的有效度和结果的普遍性提出质疑。我们敏锐地意识到了这些问题。然而，用于测试效用理论的其他所有方法也同样存在严重的缺陷。我们可以在真实情境下，或是在实验室中，对经济行为的真实选择进行自然的或统计学观察。实地研究只能提供相当粗略的定性预测检验，因为在这种情况下无法充分测量概率和效用。已经设计的实验从实际选择中获得了有关实用性和概率的精确预测，但是这些实验研究通常只适用于赌注较小的人为设定的博弈，以及大量非常相似的问题。实验博弈的这些特征使得其对结果的解释变得复杂，并限制了它们的通用性。

默认情况下，假设选择法是研究大量理论问题的最简单方法。这种方法的使用依赖于这样一种假设，即人们通常知道自己在实际情况下进行选择时会怎么做；以及进一步假设被试没有特殊的理由来掩饰自己的真实偏好。如果人们在预测自己的选择时是合理且准确的，那么在假设问题中普遍存在的系统性违反期望效用理论的行为，就为我们提供了否定该理论的假设证据。

确定性、概率和可能性

在期望效用理论中，计算结果的效用时，会将其概率作为权重系数。本节描述了一系列选择问题，其中人们的偏好系统地违反了这一原则。我们首先通过示例表明，相对于可能的结果来说，人们更看重确定性结果，这就是确定性效应现象。

法国经济学家莫里斯·阿莱斯（Maurice Allais）于1953年提出了期望效用理论的最著名的反例，并在其中探讨了确定性效应（Allais, 1953）。许多作者从规范和描述的角度讨论了阿莱斯的例子（MacCrimmon & Larsson, 1979; Slovic & Tversky, 1974）。以下两对选择问题是阿莱斯示例的变体，它与原始理论的不同之处在于，这里使用的是中等而非极大的增益。示例用 N 表示每个问题的回复人数，用中括号表示选择每个选项的被试百分比。

问题 1：请在以下选项中选择。

- | | | |
|----|---------------------------|------|
| A. | 赢得 2 500 (新谢克尔, 下同) 的概率 | 0.33 |
| | 赢得 2 400 的概率 | 0.66 |
| | 赢得 0 的概率 | 0.01 |

$N=72$ [18]

- B. 确定赢得 2 400
[82] *

问题 2：请在以下选项中选择。

- | | | |
|----|--------------|------|
| C. | 赢得 2 500 的概率 | 0.33 |
| | 赢得 0 的概率 | 0.67 |

$N=72$ [83] *

- | | | |
|----|--------------|------|
| D. | 赢得 2 400 的概率 | 0.34 |
| | 赢得 0 的概率 | 0.66 |
- [17]

数据显示, 82%的被试在问题1中选择了B, 83%的被试在问题2中选择了C。星号表示此偏好在0.01水平上是显著的。此外, 对个人选择模式的分析表明, 大多数被试(61%)在这两个问题上都做出了模式化的选择。这种偏好模式违反了阿莱斯最初描述的期望效用理论。根据该理论, 当 $u(0)=0$ 时, 第一个偏好表明:

$$u(2\,400) > 0.33u(2\,500) + 0.66u(2\,400), \text{ 即}$$

$$0.34u(2\,400) > 0.33u(2\,500)$$

而第二种偏好则反映了相反的不等式是成立的。请注意, 问题2是通过从问题1中的两个前景中同时减去“0.66的可能性赢得2 400新谢克尔”而得到的。显然, 当这种变化将前景特征从确定的利益变成可能的利益时, 与原始和减少的前景都不确定的情况相比, 其可取性会大大降低。

下面给出了相同现象的简单演示, 此处仅涉及两个结果的博弈。这个例子也是基于阿莱斯的示例(Allais, 1953)。

问题 3:

A. (4 000, 0.80) 或 B. (3 000)
N=95 [20] [80]*

问题 4:

C. (4 000, 0.20) 或 D. (3 000, 0.25)
N=95 [65]* [35]

在本节中的这对问题及所有其他对问题中，超过一半的被试违反了期望效用理论。为了表明问题3和问题4中的偏好模式与理论不相容，设 $u(0) = 0$ ，并认为B的选择意味着 $u(3\,000)/u(4\,000) > 4/5$ ，而C的选择意味着逆向不平等。注意，前景 $C = (4\,000, 0.20)$ 可以表示为 $(A, 0.25)$ ，而前景 $D = (3\,000, 0.25)$ 可以改写为 $(B, 0.25)$ 。前景效用理论的替代定理断言，如果B优于A，那么任何（概率）混合体 (B, p) 必须优于混合体 (A, p) 。我们的实验结果符合这个定理。显然，将获胜率从1.0降至0.25的效果比将获胜率从0.8降至0.2的效果更大。以下两对问题说明了不涉及经济收益情况下的确定性效应。

问题 5:

A. 你有 50% 的概率赢得英格 B. 你确定无疑能赢
兰、法国和意大利三周游 得英格兰一周游。
N=72 [22] [78]*

问题 6:

A. 你有 5% 的概率赢得英格 B. 你有 10% 的概率
兰、法国和意大利三周游 能赢得英格兰一
周游
N=72 [67]* [33]

确定性效应不是违反替代定理的唯一情况。以下问题说明了违反这个定理的另一种情况。

问题 7:

- | | | | |
|--------|---------------|----|---------------|
| A. | (6 000, 0.45) | B. | (3 000, 0.90) |
| $N=66$ | [14] | | [86] * |

问题 8:

- | | | | |
|--------|----------------|----|----------------|
| C. | (6 000, 0.001) | D. | (3 000, 0.002) |
| $N=66$ | [73] * | | [27] |

请注意，在问题7中，你获胜的可能性很大（0.45和0.90），并且大多数人选择更有可能获胜的情况。在问题8中，虽然有获胜的可能性，但获胜的可能性都很小（0.001和0.002）。在这种情况下，获胜是可能但不是一定的，大多数人选择了能带来更大收益的前景。K.R.麦克里蒙（K.R. MacCrimmon）和S.拉尔森（S. Larsson）也发表了类似的结果（MacCrimmon & Larsson, 1979）。

上述问题说明了在前景效用模型中人们对无法确定的风险或机会的一般态度。结果总结了关于违反替代定理的方式。如果 (y, pq) 等于 (x, p) ，那么 (y, pqr) 优于 (x, pr) ， $0 < p, q, r < 1$ 。该特性被纳入了替代定理，我们将在本章的第二部分对此进行阐述。

镜像效应

上一节中我们讨论了人们在正面前景，即不涉及损失的前景之间的偏好。当结果的符号变成负号，也就是前景由收益变成损失时，会发生什么？表4-1的左侧栏显示的是上一节讨论过的4道选择题，右侧栏显示了结果的符号变成负号的选择题。我们用 $-x$ 来表示损失 x ，用 $>$ 表示普遍的偏好，即大多数被试做出的选择。

表4-1 正面前景和负面前景之间的偏好

正面前景			负面前景		
问题 3: $(4\,000, 0.80) < (3\,000)$			问题 3': $(-4\,000, 0.80) > (-3\,000)$		
N=95	[20]	[80] *	N=95	[92] *	[8]
问题 4: $(4\,000, 0.20) > (3\,000, 0.25)$			问题 4': $(-4\,000, 0.20) < (-3\,000, 0.25)$		
N=95	[65] *	[35]	N=95	[42]	[58]
问题 7: $(3\,000, 0.90) > (6\,000, 0.45)$			问题 7': $(-3\,000, 0.90) < (-6\,000, 0.45)$		
N=66	[86] *	[14]	N=66	[8]	[92] *
问题 8: $(3\,000, 0.002) < (6\,000, 0.001)$			问题 8': $(-3\,000, 0.002) > (-6\,000, 0.001)$		
N=66	[27]	[73] *	N=66	[70] *	[30]

在表4-1所示的4个问题中，人们在负面前景之间的偏好是正面前景之间的偏好的镜像。因此，镜像0周围的前景的镜像颠倒了偏好顺序。我们将这种模式标记为镜像效应。

现在让我们谈谈这些数据的含义。首先，请注意镜像效应意味着正数区间中的风险规避将成为负数区间中的风险寻求。例如，在问题3'中，尽管博弈具有较低的前景值，但大多数被试愿意接受概率为0.80的损失4 000的风险，而不是确定损失3 000。A. C.威廉姆斯（A. C. Williams）早期就注意到了人们在负面前景之间进行选择时存在风险寻求（Markowitz, 1952）。威廉姆斯给出的数据显示，结果的转换产生了从风险规避到风险寻求的巨大转变（Williams, 1966）。例如，被试无视 $(100, 0.65; -100, 0.35)$ 和 (0) ，这表明他们厌恶风险。被试也无视 $(-200, 0.80)$ 和 (-100) ，这表明他们寻求风险。P. C.菲什伯恩（P. C. Fishburn）和G. A.柯亨伯格（G. A. Kochenberger）所做的综述，记录了人们在负面前景选择中寻求风险的普遍性（Fishburn & Kochenberger, 1979）。

其次，表4-1显示人们在正面前景之间的偏好与期望效用理论不一致。相应的负面前景之间的偏好也以同样的方式违反了前景原则。例如，问题3'和问题4'及问题3和问题4，均表明相对于不确定是否能赢得某物，确定能赢得某物更重要。在正数区间中，确定性效应导致了人们对确定性收益的风险规避偏好程度大于更大的不确定性收益。在负数区间中，同样的效应导致了人们的风险偏好，即相对于确定的较小损失而言，有可能发生的损失更容易受到风险偏好的影响。同样的心理原则，

即人们更看重确定性导致了人们在收益区间的风险规避和损失区间的风险寻求行为。

再次，镜像效应作为确定性效应的解释，消除了人们对不确定性或可变性的厌恶，如对 $(3\ 000)$ 大于 $(4\ 000, 0.80)$ ，以及 $(4\ 000, 0.20)$ 大于 $(3\ 000, 0.25)$ 的普遍偏好。为了解决这种明显的不一致性，可以引用人们更喜欢具有高前景值和小方差前景的假设（Allais, 1953; Markowitz, 1959; Tobin, 1958）。由于 $(3\ 000)$ 方差较小，而 $(4\ 000, 0.80)$ 方差较大，尽管前者前景值较低，但人们仍然会选择前者。然而，当前景减少时， $(3\ 000, 0.25)$ 和 $(4\ 000, 0.20)$ 之间的方差差异可能不足以克服前景值大小的差异。因为 $(-3\ 000)$ 具有比 $(-4\ 000, 0.80)$ 更高的前景值和更低的方差，根据这个解释，选择确定性损失应该更符合被试的偏好，而这种现象与数据结果背道而驰。因此，我们的数据与人们通常喜欢确定性这种说法并不相符。相反，似乎是确定性强化了人们对损失的厌恶，以及对收益的偏好。

概率型保险

为对抗损失而购买保险这一现象，已经被许多人视为货币效用函数具有凹性的有力证据。为什么人们会花这么多钱甚至以超过预期精算成本的价格购买保险呢？然而，对各种形式保险相对吸引力的检验，并不支持货币的效用函数具有凹性这一观念。例如，人们通常更喜欢可以提供有限保障但具有低免赔额或零免赔额的保险产品，而不是具有较大的保障范围和更高免赔额的保险产品，这与风险规避相反（Fuchs, 1976）。另一种类型的保险问题可被称为概率型保险，其中人们的反应与货币效用函数的凹性假设不一致。为了说明这个概念，请考虑以下问题，该问题已由95名斯坦福大学的学生进行了回答。

问题 9：假设你考虑为某些财产购买保险以防发生意外，如遭遇火灾或盗窃。在了解了风险和溢价之后，你在购买保险或不购买保险的选择之间没有明确的偏好。

然后，请注意保险公司将提供一项名为概率型保险的新计划。在这项计划中，你需要支付一半的常规保险费。如果财产遭受损坏，你有 50% 的机会支付另一半保险费，保险公司会承担所有损失；同时你还有 50% 的机会拿回保险费并承受所有损失。例如，如果在一个月的某个奇数日发生事故，你需要支付另一半保险费，你的损失将会得到赔偿；但如果事故发生在该月的偶数日，你的保险金将被退还，但你的损失将不会得到赔偿。

回想一下，全面保险的保险费就是这样的，你会发现这个保险几乎不值得购买。

在这种情况下，你会购买概率型保险吗？

	会	否
$N=95$	[20]	[80] *

尽管问题9看起来不怎么合理，但值得注意的是，概率型保险代表了许多形式的保障，其中，人们支付一定的成本来降低不良事件的可能性，而不是完全消除这种可能性。安装防盗报警器、更换旧轮胎及戒烟的决定都可被视为一种概率型保险。

对问题9和同一问题的其他几个变体的回答表明，概率型保险通常没有吸引力。显然，将损失概率从 p 降至 $p/2$ 的可能性比将损失概率从 $p/2$ 降至0的可能性要小。

与这些数据相反，期望效用理论（凹陷系数 u ）意味着概率型保险优于常规保险。也就是说，如果在资产 w 中，人们只愿意支付溢价 y 以确保护损失 x 的概率是 p ，那么一个人肯定愿意支付一个较小的溢价 ry 以将失去 x 的概率从 p 减小到 $(1-r)p$ ， $0 < r < 1$ 。也就是说，如果一个人认为 $(w-x, p; w, 1-p)$ 和 $(w-y)$ 无关紧要，那么相对于保险 $(w-y)$ ，人们应该更喜欢概率型保险 $(w-x, (1-r)p; w-y, rp; w-ry, 1-p)$ 。

为证明这一说法，我们证明了：

$$pu(w-x) + (1-p)u(w) = u(w-y)$$

从而推导出：

$$(1-r)pu(w-x) + rpu(w-y) + (1-p)u(w-ry) > u(w-y)$$

在不失一般性的前提下，我们可以设 $u(w-x) = 0$ 和 $u(w) = 1$ 。因此， $u(w-y) = 1-p$ ，我们希望得到

$$rp(1-p) + (1-p)u(w-ry) > 1-p$$

或

$$u(w-ry) > 1-rp$$

当且仅当 u 具有凹性时公式成立。

这是效用理论的风险规避假说的一个相当令人费解的结果，因为概率型保险看起来比常规保险更具风险性，常规保险完全消除了风险因素。显然，风险效用函数的假定凹度并未充分反映风险的直观概念。

人们对概率型保险的排斥很令人费解，因为从某种意义上说，所有保险都是概率型的。最狂热的保险买家仍然容易受到政策没有覆盖的许多金融风险和其他风险因素的影响。概率型保险与所谓的意外保险（contingent insurance）之间似乎存在着显著差异，意外保险为特定类型的风险提供了确定性。例如，将概率型保险与针对家庭内所有形式的损失或损害的意外保险进行对比。意外保险能消除因盗窃而造成的所有损失风险，但不包括其他风险，如火灾。我们推测，在没有保护的情况下，当损失的概率相等时，意外保险通常比概率型保险更具吸引力。因此，两个在概率和结果上等同的前景可能因其表述方式的不同而具有不同的值。下一节中我们将介绍一些普遍存在的此类现象。

隔离效应

为了简化在备选方案之间做出选择的过程，人们常常会忽略备选方案共有的成分，而专注于区分它们的内容（Tversky, 1972）。这种选

择问题的方法可能会产生不一致的偏好，因为一对前景可以被多种方式分解为共同性成分和区别性成分，而不同的分解方式有时会引发不同的偏好。我们将这种现象称为隔离效应。

问题 10：考虑以下包括两个阶段的游戏。在第一阶段，游戏结束却没有赢得任何东西的概率为 0.75，进入第二阶段的概率为 0.25。如果进入第二阶段，那么可以在 (4 000, 0.80) 和 (3 000) 之间进行选择。
你必须在游戏开始之前，即在知道第一阶段的结果之前做出选择。

请注意，在这个游戏中，人们可以选择以 $0.25 \times 0.80 = 0.20$ 的概率赢得 4 000，也可以选择以 $0.25 \times 1.0 = 0.25$ 的概率赢得 3 000。因此，就最终结果和概率而言，人们面临着 (4 000, 0.20) 和 (3 000, 0.25) 之间的选择，如问题 4 所示。然而，这两个问题的主导偏好是不同的。在回答问题 10 的 141 名被试中，78% 的被试选择了后者，这与问题 4 中的结果偏好相反。显然，人们忽略了游戏的第一阶段，其结果由两个前景共享，并将问题 10 视为在 (4 000, 0.80) 和 (3 000) 之间进行选择，如上文中的问题 3 所示。

我们把问题 4 标准形式和问题 10 顺序形式分别表示为决策树，如图 4-1 和图 4-2 所示。按照惯例，正方形表示决策节点，圆形表示机会节点。两个图之间的本质区别在于决策节点的位置。在标准形式（如图 4-1 所示）中，决策者面临着两个风险前景之间的选择；而在顺序形式（如图 4-2 所示）中，决策者面临的是风险前景和无风险前景之间的选择。这是通过在不改变概率或结果的情况下引入前景之间的依赖关系实现的。具体而言，事件“未赢得 3 000”被包括在顺序形式“未赢得 4 000”的事件中，而这两个事件在标准形式中是相互独立的。因此，“赢得 3 000”在顺序形式中具有确定性这一优势，而在标准形式中则没有。

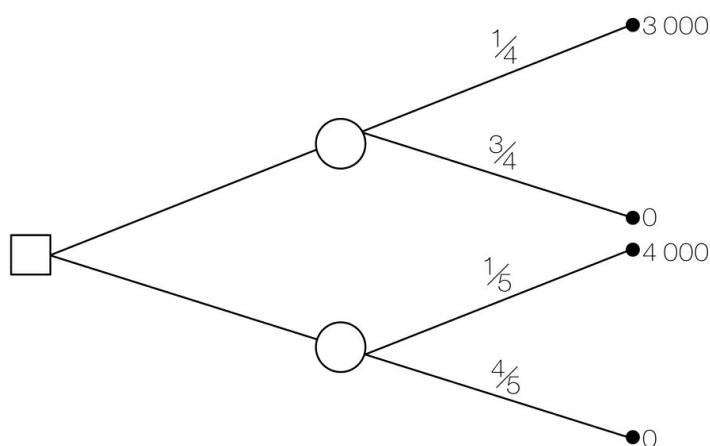


图4-1 问题4的表达式决策树（标准形式）

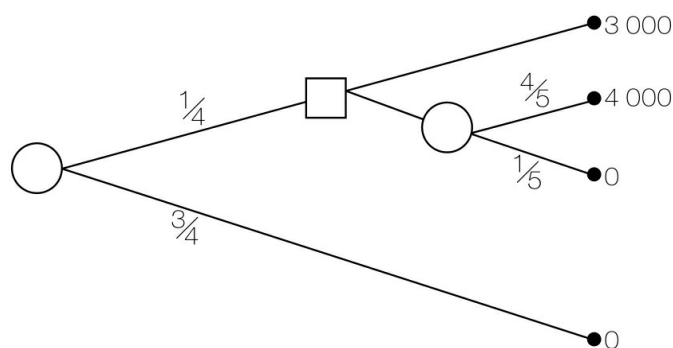


图4-2 问题10的表达式决策树（顺序形式）

由于事件之间具有依赖性，偏好逆转具有重要的意义，因为它违背了决策理论分析的基本假设，即前景之间的选择仅由概率的最终状态决定。

人们通常很容易想到以上述其中一种形式表示的决策问题。例如，我们可以在标准形式中查看两种不同风险的企业之间的选择。此外，以下问题最有可能以顺序形式表示。人们可以投资一个风险投资项目，如果投资失败，人们可能会失去个人资本；如果成功，人们可以在约定的固定收益和一定比例的收益之间做出选择。隔离效应意味着相对于具有相同概率和结果的风险投资而言，具有固定回报或确定性增强了该选项的吸引力。

前面的问题说明了如何用不同的概率表现形式来改变人们的偏好。我们现在阐述的是如何通过改变结果的表现形式来改变人们的选择。

考虑以下问题，这些问题被呈现给了两个不同的受试组。

问题 11：除了拥有的东西，你还会获得 1 000（新谢克尔，下同）。你需要在以下两个选项中进行选择。

A. (1 000, 0.50)

B. (500)

N=70 [16]

[84]*

问题 12：除了拥有的东西，你还会获得 2 000。你需要在以下两个选项中进行选择。

C. (-1 000, 0.50)

D. (-500)

N=68 [69]*

[31]

大多数被试在第一个问题中选择了B，在第二个问题中选择了C。这些偏好符合表4-1中显示的镜像效应，表现出了对正面前景的风险规避和对负面前景的风险寻求。但请注意，从最终状态来看，两个选择问题是相同的，具体如下：

$A = (2\,000, 0.50; 1\,000, 0.50) = C$

$B = (1\,500) = D$

实际上，问题12是通过在问题11初始奖金的基础上增加1 000，然后又从所有获得的奖金中减去1 000得到的。显然，被试没有将奖金与前景相结合。奖金没有进入前景比较中，因为每个问题的两个选项都很常见。

我们从问题11和问题12中观察到的结果显然与效用理论不一致。例如，在该理论中，不管之前的效用财富是9.5万美元还是10.5万美元，我们将相同的效用分配给财富为10万美元的人。因此，在拥有10万美元的总财富与拥有9.5万美元或10.5万美元的机会之间进行选择，应该与目前是否拥有这两个金额中较小或较大的金额无关。在附加了风险规避假设的情况下，该理论认为，拥有10万美元的确定性总是比博弈情况更可取。然而，被试对问题12和之前几个问题的回答表明，如果个人拥有较小的金额，则会出现上述结果，但如果他拥有较大的金额则不会。

在问题11和问题12的两个选项中，对奖金的明显忽视是常见现象，

这意味着价值或效用的载体是财富的变化，而不是包括当前财富的最终资产。这个结论是风险选择替代定理的基石，我们将在以下章节中进行描述。

理论

上文中，我们讨论和回顾了几项实证效应，这些效应似乎使期望效用理论成了无效的描述性模型。本章的其余部分介绍了风险条件下个人决策的另一种方法，即前景理论。该理论是针对具有货币结果和陈述概率的简单前景发展而来的，但它可以扩展到更多涉及选择的形式。前景理论区分了选择过程中的两个阶段：早期的编辑阶段和随后的评估阶段。编辑阶段包括对所提供的前景的初步分析，通常可以得到这些前景的简单的表现形式。在第二阶段中，人们对评估后的前景进行评价，并选择价值最高的前景。接下来我们将对编辑阶段进行概述，并进一步改进评估阶段的正式模型。

编辑阶段的功能是组织和重新设置选项，以简化后续的评估和选择。编辑包括若干操作的应用，这些操作可改变与所提供的前景相关的结果和概率。编辑阶段的主要操作如下所述。

编码。上文中的讨论表明，人们通常认为结果是收益和损失，而不是财富或福利的最终状态。当然，收益和损失是相对于某个中性参考点来定义的。这个参考点通常对应于当前资产，在这种情况下，收益和损失与收到或支付的实际金额一致。但是，这个参考点的位置及相应的损失或收益的结果编码，会受前景的描述形式及决策者的期望影响。

组合。有时可以通过将与相同结果相关的概率进行组合来简化前景。例如，前景 $(200, 0.25; 200, 0.25)$ 可简化为 $(200, 0.50)$ ，并以该形式来进行评估。

分离。某些前景包含一个无风险部分，因为该部分在编辑阶段风险成分组分离。例如，前景 $(300, 0.80; 200, 0.20)$ 自然地分解为200的确定性收益和 $(100, 0.80)$ 的风险前景。类似地，前景 $(-400, 0.40; -100; 0.60)$ 很容易被视作确定性损失100和前景 $(-300, 0.40)$ 。

前面的操作可分别应用于每个前景。以下操作适用于一组两个或更多前景。

消除。前面描述的隔离效应的本质是忽略前景的共享成分。因此，我们的被试显然忽略了问题10中呈现的连续游戏的第一阶段，因为这个阶段对于两个选项来说是共有的，他们会根据第二阶段的结果来评估前景（图4-2）。同样，他们忽略了问题11和问题12中增加到前景中的共同奖金。另一种类型的消除涉及舍弃共享成分，即结果-概率组合。例如， $(200, 0.20; 100, 0.50; -50, 0.30)$ 和 $(200, 0.20; 150, 0.50; -100, 0.30)$ 之间的选择可以通过消除而减少到 $(100, 0.50; -50, 0.30)$ 和 $(150, 0.50; -100, 0.30)$ 之间。

这里还要介绍另外两项操作，即对前景进行简化和对优势度进行检测。第一个操作是通过对概率或结果进行取整来简化前景。例如，前景 $(101, 0.49)$ 很可能被记录为赢得100的偶然机会。舍弃极不可能的结果是一种特别重要的简化形式。第二个操作涉及扫描所提供的前景，以发现占据优势地位的替代性方案，这些方案在没有得到进一步评估的情况下会被拒绝。

由于编辑操作可以使决策任务的执行更便利，所以我们假设它们是在可能的情况下被执行的。然而，有些编辑操作会允许或阻止其他操作的应用。例如，如果两个前景的第二个成分被简化为 $(100, 0.50)$ ，则 $(500, 0.20; 101, 0.49)$ 比 $(500, 0.15; 99, 0.51)$ 优势度更高。因此，最终的编辑前景可能取决于编辑操作的顺序，而编辑操作的顺序可能会随集合的结构和显示的格式发生变化。对这个问题的处理超出了目前研究的范围。在本文中，我们讨论的选择问题是，合理地假设前景的原始形式没有进一步编辑的空间，或可以明确地指定经过编辑的前景。

偏好上出现的许多异常情况都是对前景进行编辑的结果。例如，与隔离效果相关的不一致性是由于消除了共享成分。选择的某些不和谐性可以通过简化来解释，这种简化消除了前景之间的微小差异（Tversky, 1969）。说得更通俗一些就是，前景之间的偏好顺序在不同情况下不一定会保持不变，因为相同的前景可以根据其出现的情境以不同的方式被编辑。

在编辑阶段完成之后，我们假设决策者会对每个编辑后的前景进行评估，然后选择价值最高的前景。用 V 表示编辑后的前景的整体价值，用 π 和 v 表示其两个尺度。

第一个尺度 π 将每个概率 p 与决策权重 $\pi(p)$ 相关联，这反映了 p 对前景的总体价值的影响。然而， π 不是概率尺度，我们稍后将会证明

$\pi(p) + \pi(1-p)$ 通常小于1。第二个尺度 v 为每个结果分配一个数字 $v(x)$ ，它反映了该结果的主观价值。回想一下，结果是相对于参考点定义的，参考点是价值尺度的零点。因此，尺度 v 与该参考点的偏差值即增益和损失。

目前的表述涉及以 $(x, p; y, q)$ 为形式的简单前景，其最多有两个非零结果。在这样的前景中，人们接受 x 的概率为 p ，接受 y 的概率为 q ，而什么都不接受的概率为 $1-p-q$ ，其中 $p+q \leq 1$ 。如果结果全部为正数，即如果 $x, y > 0$ 且 $p+q = 1$ ，那么所提供的前景严格为正；如果结果全部为负数，那么所提供的前景严格为负；如果结果既不严格为正也不严格为负，那么所提供的前景为常规前景。

该理论的基本公式描述了 π 和 v 相结合来确定常规前景总体价值的方式。

如果 $(x, p; y, q)$ 是常规前景（即 $p+q < 1$ ， $x \geq 0 \geq y$ 或 $x \leq 0 \leq y$ ），那么

$$V(x, p; y, q) = \pi(p) v(x) + \pi(q) v(y) \quad (1)$$

其中 $v(0) = 0$, $\pi(0) = 0$, $\pi(1) = 1$ 。在效用理论中， V 是在前景上定义的，而 v 是在结果上定义的。两个尺度在确定性前景上重合，其中 $V(x, 1.0) = V(x) = v(x)$ 。

等式 (1) 通过放宽前景理论来概括期望效用理论。附录中描述了对这种表现形式的定理分析，同时也描述了确保存在唯一 π 和满足等式 (1) 的比例尺度的条件。

严格为正和严格为负的前景的评估遵循不同的规则。在编辑阶段，这些前景被分为两个成分：第一是无风险成分，即肯定会获得或支付的最低收益或损失；第二是风险成分，即风险的额外收益或损失。对这种前景的评估可参考如下等式。

如果 $p+q = 1$ 且 $x > y > 0$ 或 $x < y < 0$ ，那么

$$V(x, p; y, q) = v(y) + \pi(p) [v(x) - v(y)] \quad (2)$$

也就是说，严格为正或严格为负的前景的值等于无风险成分的值加上结果之间的值差与更极端的结果相关的权重的积。例

如, $V(400, 0.25; 100, 0.75) = v(100) + \pi(0.25) [v(400) - v(100)]$ 。等式(2)的基本特征是将决策权重应用于值差 $v(x) - v(y)$, 它代表着前景的风险成分, 但不代表 $v(y)$, 即无风险成分。注意, 等式(2)的右边等于 $\pi(p)v(x) + [1 - \pi(p)]v(y)$ 。因此, 如果 $\pi(p) + \pi(1-p) = 1$, 那么等式(2)可简化为等式(1)。我们将在下文中证明, 该条件通常无法满足。

评估模型的许多要素出现在前文对期望效用理论的尝试中。马科维茨第一个提出效用被定义为收益和损失而不是最终资产的假设, 这一假设在大多数实用性效用测量中已得到默认 (Daviclson, Suppes, & Siegel, 1957; Mosteller & Nogee, 1951; Markowitz, 1952)。马科维茨还指出在正前景和负前景中存在寻求风险的偏好。他提出了一个效用函数, 在正数区间和负数区间都有凸凹区域。然而, 他保留了前景原则, 因此, 他无法解释许多违反这一原则的行为, 详情如表4-1所示。

W.爱德华兹 (W. Edwards) 提出通过更常规的权重来替换概率, 并且在几个实证研究中对该模型进行了研究 (Edwards, 1962; Anderson & Shanteau, 1970; Tversky, 1967)。W.费尔纳 (W. Fellner) 开发了类似的模型, 他引入了决策权重的概念来解释模糊厌恶, 而范·达姆 (van Dam) 试图扩展决策权重的模型 (Fellner, 1965; van Dam, 1975)。对于期望效用理论和替代选择模型的其他关键分析, 可参见阿莱斯、C. H.库姆斯 (C. H. Coombs)、菲什伯恩和B.汉森 (B. Hansson) 的研究 (Allais, 1953; Coombs, 1975; Fishburn, 1977; Hansson, 1975)。

前景理论公式保留了期望效用理论的基础性一般双线性形式。然而, 为了适应本文第一部分中描述的效果, 我们必须假设值被附加到了变化上而不是最终状态上, 并且决策权重与所述概率不一致。这些与期望效用理论的偏离必然导致规范性的不可接受的后果, 诸如不一致、不和谐或违反支配性。当意识到自己的偏好不一致、不可过渡或不可接受时, 决策者通常会纠正这种偏好异常。然而, 在许多情况下, 决策者没有机会发现自己的偏好可能违反了其希望服从的决策规则。在这些情况下, 前景理论的异常是可以预见的。

效用函数

本理论的一个基本特征是价值的载体是财富或福利的变化，而不是财富或福利的最终状态。这种假设与感知和判断的基本原则一致。我们的感知器官适应的是对变化或差异的评估，而不是对绝对值的评估。当我们回应亮度、响度或温度等属性时，过去和现在的经验定义了适应水平或参考点，以及相对于该参考点的感知刺激物（Helson, 1964）。因此，一个物体被感知到的冷或者热取决于人们所适应的温度。该原则同样适用于非感官属性，如健康、声望和财富。例如，同样的财富水平对于一个人来说可能意味着赤贫，而对于另一个人来说则是巨额财富，这取决于他们目前的资产状况。

强调变化是价值的载体，并不意味着某一特定变化的价值与初始位置无关。严格地说，效用函数应该被视为有两个参数的函数：作为参考点的资产量，以及从该参考点开始（正或负）的变化幅度。比如，个人对金钱的态度可以通过一本书来描述，其中每一页都显示了特定资产量变化的效用函数。显然，不同页面上描述的效用函数并不相同：资产增加的形式可能是线性的。然而，前景的偏好顺序并未因资产状况的微小甚至相对变化而大幅改变。例如，对于大多数人来说，在通常的资产状况中，前景（1 000, 0.50）的确定性收益为300~400。因此，在一个参数中把值表示为函数通常可以为我们提供一个令人满意的近似值。

许多感觉和感知维度都有一个共同的特性，即心理反应是物理变化幅度的凹函数。例如，人们在室温下区分3℃和6℃的变化要比区分13℃和16℃的变化更容易。我们认为这一原则特别适用于对货币变化的评估。因此，收益100和收益200之间的差值似乎大于收益1 100和收益1 200之间的差值。同样，损失100和损失200之间的差值似乎大于损失1 100和损失1 200之间的差值，除非更大的损失是无法容忍的。因此，我们假设财富的效用函数变化通常在参考点之上是凹的（ $u''(x) < 0, x > 0$ ），在参考点下方是凸的（ $u''(x) > 0, x < 0$ ）。即收益和损失的边际价值通常随其幅度的增加而减小。加兰特和普利纳的报告支持了这一假设，他们量化了货币和非货币的收益和损失的感知程度（Galanter & Pliner, 1974）。

上述关于效用函数形状的假设基于对无风险情境下收益和损失的反应。我们发现从风险选择中得出的效用函数具有相同的特征，如下列问题所示。

问题 13:

	(6 000, 0.25)	(4 000, 0.25; 2 000, 0.25)
N=68	[18]	[82] *

问题 13':

	(-6 000, 0.25)	(-4 000, 0.25; -2 000, 0.25)
N=64	[70] *	[30]

将等式 (1) 应用于这些问题中的偏好模型可得:

$$\pi(0.25) v(6\,000) < \pi(0.25) [v(4\,000) + v(2\,000)]$$

和

$$\pi(0.25) v(-6\,000) > \pi(0.25) [v(-4\,000) + v(-2\,000)]$$

因此

$$v(6\,000) < v(4\,000) + v(2\,000) \text{ 和 } v(-6\,000) > v(-4\,000) + v(-2\,000)$$

这些偏好符合这样的假设: 效用函数对于增益是凹的, 而对于损失是凸的。

任何关于货币效用函数的讨论都必须为特殊情况对偏好的影响留出空间。例如, 需要60 000美元购买房屋的个人效用函数可能会异常地在临界值附近急剧上升。同样地, 个人对损失的厌恶可能会在损失临界值附近急剧增加, 这会迫使人们卖掉他的房子并搬到一个不太理想的社区。因此, 个人的衍生价值(效用)函数并不总是反映人们对货币的“纯粹”态度, 因为它可能受到与特定金额相关的附加后果的影响。这种影响可以轻易地在效用函数的收益中形成凸区域, 而在效用函数的损失中形成凹区域。后一种情况可能更常见, 因为大量损失通常会改变人们的生活方式。

对福利态度的变化有一个显著特征, 即损失大于收益。人们在失去一笔金钱时所经历的痛苦似乎比获得相同数额的金钱带来的快乐程度更大 (Galanter & Pliner, 1974)。实际上, 大多数人发现以 $(x, 0.50; -$

$x, 0.50$) 为形式的对称赌注明显没有吸引力。此外，人们对于公平赌注的厌恶通常会随着赌注的增加而增加。也就是说，如果 $x > y \geq 0$ ，那么 $(y, 0.50; -y, 0.50)$ 优于 $(x, 0.50; -x, 0.50)$ 。因此，根据等式 (1) 可得 $u(y) + u(-y) > u(x) + u(-x)$ 和 $u(-y) - u(-x) > u(x) - u(y)$ 。

设 $y = 0$ ，得到 $u(x) < -u(-x)$ ，让 y 无限接近 x ，得到 $u'(x) < u'(-x)$ ，且满足 u 的导数 u' 的存在。因此，损失的效用函数比收益的效用函数更陡峭。

总之，我们认为效用函数定义于参考点的偏差上；其次，收益通常为凹函数，损失通常为凸函数；再者，损失函数比收益函数更陡峭。满足这些属性的效用函数如图4-3所示。请注意，我们所提出的S形效用函数在参考点处最陡峭，这与马科维茨假设的效用函数形成了鲜明对比，该函数在该区域相对较平缓 (Markowitz, 1952)。

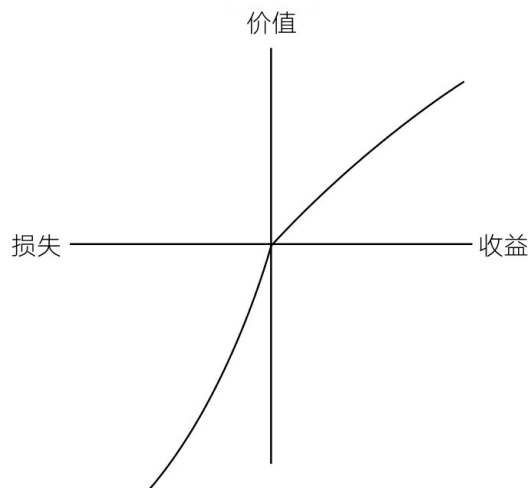


图4-3 假设的效用函数

尽管本理论可以应用于从前景间的偏好中推导出的效用函数，但由于决策权重的引入，实际衡量比实用理论要复杂得多。例如，即使应用线性效用函数，决策权重也可能产生风险规避和风险寻求。然而，值得关注的是，在冯·诺伊曼-摩根斯坦财富变化效用函数的详细分析中我们已经观察到了归因于效用函数的主要性质 (Fishburn & Kochenberger, 1979)。这些函数来自5项独立研究，这些研究的参与者是30位来自各个业务领域的决策者 (Barnes & Reinmuth, 1976; Grayson, 1960; Green, 1963; Halter & Dean, 1971; Swalm, 1966)。大多数收益效用函数是凹的，大多数损失函数都是凸的，只有3个人对收益和损失都表

现出了风险规避。除此之外，损失的效用函数比收益的效用函数要大得多。

加权函数

在前景理论中，我们用每个结果的价值乘以决策权重。决策权重是从前景之间的选择中推断出来的，主要概率是从拉姆齐-萨维齐（Ramsey-Savage）偏好方法中推断出来的。然而，决策权重不是概率，它们不服从概率定理，不应被解释为程度或信念的尺度。

我们可以思考这样一种下注，人们通过投掷硬币来决定可以赢得1 000或什么都赢不到。任何人在这种情况下获胜的概率是0.50。这个结果可以通过多种方式进行验证；例如，通过展示被试对硬币的正反毫不在意，或者通过被试口述他们认为这两个事件是等概率的。然而，从选择中得出的决策权重 $\pi(0.50)$ 可能小于0.50。决策权重能衡量事件对前景的可取性的影响，而不仅仅是这些事件的感知可能性。如果前景原则成立，那么两个尺度就会重合，即 $\pi(p) = p$ ，反之则不成立。

本文讨论的选择问题是根据明显的数值概率设置的，我们在分析时假设被试采用了 p 的规定值。此外，由于事件仅通过其陈述的概率来识别，因此在该情境中可以将决策权重表示为所述概率的函数。然而，一般而言，与事件相关的决策权重可能会受到其他因素的影响，如模糊性（Ellsberg, 1961; Fellner, 1961）。

我们现在转而讨论加权函数 π 的显著性质，其中我们将决策权重与所述概率相关联。当然， π 是 p 的递增函数，其中 $\pi(0) = 0$ 且 $\pi(1) = 1$ 。也就是说，我们忽略了不可能事件的结果，并且将标度正态化，使得 $\pi(p)$ 是与概率 p 相关联的权重与特定事件相关联的权重的比率。

我们首先讨论小概率加权函数的一些性质。问题8和问题8'中的偏好表明，对于较小的 p 值， π 是 p 的次可加性函数，即 $0 < r < 1$ 时， $\pi(rp) > r\pi(p)$ 。回想一下，在问题8中，（6 000, 0.001）优于（3 000, 0.002）。因此 v 的凹函数显示

$$\frac{\pi(0.001)}{\pi(0.002)} > \frac{v(3\,000)}{v(6\,000)} > \frac{1}{2}$$

问题8'中反映的偏好产生了相同的结论。然而，问题7和问题7'中的偏好模式表明，对于较大的 p 值，次可加性不需要保持不变。

此外，我们认为小概率通常较重要，即对于小 p ， $\pi(p) > p$ 。考虑以下选择题。

问题 14:

(5,000,0.001) 或 (5)
N=72 (72) * (28)

问题 14':

(-5,000,0.001) 或 (-5)
N=72 (17) (83) *

请注意，在问题14中，比起彩票的前景价值，人们更喜欢其实际价值。此外，在问题14'中，人们更喜欢小损失，这可以被视为支付小损失的保险费。马科维茨指出了类似的观察结果。在目前的理论中，问题14中对彩票的偏好意味着，假设增益的效用函数是凹的，则 $\pi(0.001) u(5000) > u(5)$ ，因此 $\pi(0.001) > u(5) / u(5000) > 0.001$ (Markowitz, 1952)。假设损失的效用函数是凸的，那么在问题14中支付保险的意愿也隐含着相同的结论。

重要的是要从评估罕见事件概率时常见的高估中区分权重，即指定决策权重的属性。注意，当前情境下不会出现高估的问题，其中假设人们采用 p 的规定值。在许多现实生活中，过高的估计和过高的权重赋值都可能会增加罕见事件的影响力。

虽然 $\pi(p) > p$ 的概率较低，但有证据表明，对于所有 $0 < p < 1$ ，存在 $\pi(p) + \pi(1-p) < 1$ 。我们将该属性标记为确定性。我们很容易看出阿莱斯版本中的典型偏好（如问题1和问题2）暗示了 p 的相关值的次确定性。我们将等式（1）应用于问题1和问题2中的普遍偏好，则分别得到

$$u(2400) > \pi(0.66) u(2400) + \pi(0.33) u(2500)$$

即

$$[1 - \pi(0.66)] u(2400) > \pi(0.33) u(2500)$$

和

$$\pi(0.33) \cup (2\ 500) > \pi(0.34) \cup (2\ 400)$$

因此

$$1 - \pi(0.66) > \pi(0.34) \text{ 或 } \pi(0.66) + \pi(0.34) < 1$$

将相同的分析应用到阿莱斯原来的例子中可得到 $\pi(0.89) + \pi(0.11) < 1$ ，而麦克里蒙和拉尔森研究的一些数据表明了 p 的附加值的次确定性 (MacCrimmon & Larsson, 1979)。

区间 $(0, 1)$ 中 π 的斜率可以被视为衡量偏好对概率变化敏感度的尺度。次确定性要求 π 相对于 p 是递减的，即偏好通常对概率变化的敏感度低于前景原则所要求的敏感度。因此，次确定性得到了人们对不确定性事件的态度基本要素，即与互补事件相关的权重之和通常小于与特定事件相关的权重。

回想一下，本文前面讨论的对替代定理的违反遵循以下规则：如果 (x, p) 等于 (y, pq) ，那么 (x, pr) 不优先于 (y, pqr) ，其中 $0 < p, q, r \leq 1$ 。通过等式 (1)

$$\pi(p) \cup (x) = \pi(pq) \cup (y)$$

可知

$$\pi(pr) \cup (x) \leq \pi(pqr) \cup (y)$$

因此

$$\frac{\pi(pq)}{\pi(p)} \leq \frac{\pi(qpr)}{\pi(pr)}$$

因此，对于固定的概率比，当概率低时，相应决策权重的比率更接近它们较高时的比率。 π 的这个属性为次比例，它对 π 的形状施加了相当大的限制：当且仅当 $\log \pi$ 是 $\log p$ 的凸函数时，它才成立。

值得注意的是，次比例及小概率的权重增加意味着 π 在该范围内是

次可加的。形式上，它可以被证明如果 $\pi(p) > p$ 且次比例保持不变，那么 $\pi(rp) > r\pi(p)$ ， $0 < r < 1$ ，条件是 p 是单调的并且在 $(0, 1)$ 上连续。

图4-4给出了一个假设的权重函数，它满足 p 的小值的权重增加、次可加性、次确定性和次比例性。这些特性要求 π 在开区间内相对平稳，并且在 $\pi(0) = 0$ 和 $\pi(1) = 1$ 的终点附近突然变化。在端点处 π 急剧下降或明显不连续的情况与一个概念是一致的，即如果给予任何权重，那么该事件的决策权重的最低值是有限的。类似的怀疑可能会对任何低于统一的决策权重施加上限。这种量子效应可能反映了确定性和不确定性之间的明确区别。此外，编辑阶段的前景简化会导致个人舍去极低概率事件并且将极高概率事件视为确定性事件。因为人们理解和评估极端概率的能力有限，所以极不可能的事件要么会被忽略要么会被过分夸大，也就是说，高概率和确定性之间的差异要么被忽视要么被夸大了。因此， π 在端点附近表现不佳。

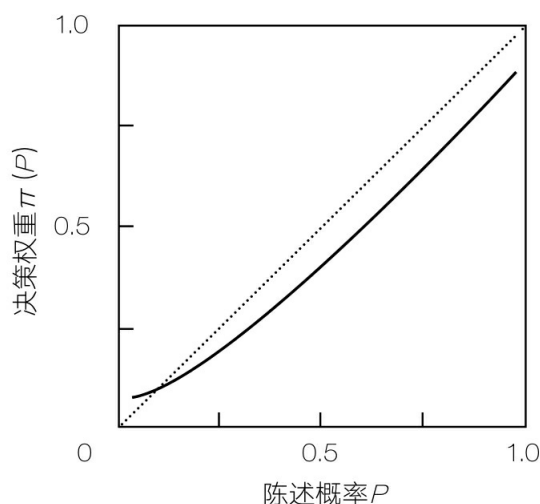


图4-4 假设的加权函数

根据理查德·泽克豪泽（Richard Zeckhauser）的研究，以下示例说明了 π 的非线性假设。假设你被迫玩俄罗斯轮盘，但你可以花钱从可能会射向你的上膛的枪中移除一颗子弹。你愿意花同样多的金钱将子弹数量从4颗减少到3颗，还是从1颗减少到0颗？大多数人认为自己愿意为将死亡概率从1/6减少到零而不是从4/6减少到3/6支付更多的钱。在后一种情况下，经济因素会导致人们支付更多的费用，此时金钱的价值大概会因为人们将无法享受它而大大降低。

对于 $\pi(p) \neq p$ 的假设，一个明显的反对形式涉及前景 $(x, p; x, q)$

和 $(x, p'; x, q')$ 之间的比较，其中 $p+q=p'+q'<1$ 。由于所有人都对这两个前景不感兴趣，我们可以认为这个观察需要条件 $\pi(p) + \pi(q) = \pi(p') + \pi(q')$ ，这种反向关系意味着 π 是恒等函数。这一论点在目前的理论中是无效的，该理论假设相同的结果概率在前景编辑中被合并了。关于 π 的非线性反对形式包含潜在的对支配原则的违背。假设 $x > y > 0$ 时， $p > p'$ ，并且 $p+q=p'+q'<1$ ，因此， $(x, p; y, q)$ 支配着 $(x, p'; y, q')$ 。如果偏好服从支配原则，那么

$$\pi(p) v(x) + \pi(q) v(y) > \pi(p') v(x) + \pi(q') v(y)$$

或

$$\frac{\pi(p) - \pi(p')}{\pi(q') - \pi(q)} > \frac{v(y)}{v(x)}$$

因此，当 y 无限接近 x 时， $\pi(p) - \pi(p')$ 接近 $\pi(q') - \pi(q)$ 。由于 $p - p' = q' - q$ ， π 必须基本上是线性的，否则必然会违反支配原则。

在目前的理论中，我们通过假设在评估前景之前检测并消除占支配地位的替代方案，可以防止直接违背支配原则。但是，该理论允许间接违背支配原则，如三元组前景， A 优于 B ， B 优于 C ， C 优于 A 。相关示例请参见 H. 雷法 (H. Raiffa) 的研究 (Raiffa, 1968)。

最后应该指出的是，目前的处理是在两个可用前景之间进行选择的最简单的决策任务。我们没有详细研究更复杂的生成性任务（如投标），其中决策者需要生成与给定前景价值相等的替代性方案。在这种情况下，两种选择之间的不对称可能会引入系统性偏差。事实上，S. 利希滕斯坦 (S. Lichtenstein) 和 P. 斯洛维克 (P. Slovic) 构建了一对前景 A 和 B ，人们通常更喜欢 A 而不是 B ，但愿意为 B 支付更高的费用

(Lichtenstein & Slovic, 1971)。这种现象已在几项研究中得到证实，包括假设和实际博弈 (Grether & Plott, 1979)。因此，通常不能假设招标程序可以恢复前景的优先顺序。

由于前景理论已被视为选择模型，因此博弈和选择不一致意味着价值和决策权重的衡量应基于特定前景之间的选择，而不是基于博弈或其他任务。这种限制使得对 v 和 π 的评估更加困难，因为生成任务与比较任务相比更便于量化。

讨论

在最后一节中，我们阐述了前景理论是如何解释人们对风险的态度，讨论了由参考点的变化引起的选择问题的替代性形式，并概述了该理论的几个扩展形式。

风险态度

在阿莱斯例子（问题1和问题2）中观察到的主要偏好模式来自现有理论，当且仅当

$$\frac{\pi(0.33)}{\pi(0.34)} > \frac{v(2\,400)}{v(2\,500)} > \frac{\pi(0.33)}{1-\pi(0.66)}$$

在这种情况下，对独立定理的违反可以归因于次确定性，更具体地说，可归因于不等式 $\pi(0.34) < 1-\pi(0.66)$ 。该分析表明，只要两个非零结果的 v -比率受相应的 π -比率限制，就会出现阿莱斯版本中的违背。

问题3至问题8具有相同的结构，因此只需要考虑其中的一对问题，如问题7和问题8。这些问题的选项是由下列理论暗示的，当且仅当

$$\frac{\pi(0.001)}{\pi(0.002)} > \frac{v(3\,000)}{v(6\,000)} > \frac{\pi(0.45)}{\pi(0.90)}$$

在这种情况下，对替代定理的违反应归因于 π 的次比例。因此，只要两个结果的 v -比率受各自的 π -比率的限制，就会以上述方式违反期望效用理论。同样的分析也适用于正数区间和负数区间中其他违反替代定理的情况。

我们接下来将证明，问题9中对定期概率型保险的偏好来自前景理论，即假设损失概率更重要。也就是说，如果 $(-x, p)$ 对 $(-y)$ 无关紧要，那么 $(-y)$ 优先于 $(-x, p/2; -y, p/2; -y/2, 1-p)$ 。为简单起见，我们定义 $x \geq 0$ 时， $f(x) = -v(-x)$ 。由于损失的效用函数是凸的，因此 f 是 x 的凹函数。应用前景理论，随着等式2的自然延伸，我们希望证明

$$\pi(p) f(x) = f(y)$$

这表示

$$f(y) \leq f(y/2) + \pi(p/2)[f(y) - f(y/2)] + \pi(p/2)[f(x) - f(y/2)] = \\ \pi(p/2)f(x) + \pi(p/2)f(y) + [1 - 2\pi(p/2)]f(y/2)$$

将 f 的凹度代入 $f(x)$ ，足以证明

$$f(y) \leq \frac{\pi(p/2)}{\pi(p)} f(y) + \pi(p/2)f(y) + f(y)/2 - \pi(p/2)f(y)$$

或

$$\pi(p)/2 \leq \pi(p/2)$$

这遵循了 π 的次可加性。

根据现有理论，风险态度是由 u 和 π 共同决定的，而不仅仅是由效用函数决定的。因此，研究前景出现风险规避或风险寻求的条件是有益的。考虑博弈 (x, p) 和前景值 (px) 之间的选择。如果 $x > 0$ ，那么每当 $\pi(p) > u(px)/u(x)$ 时意味着风险寻求，即如果增益效用函数是凹的，那么风险寻求大于 p 。因此，权重高估 $(\pi(p) > p)$ 是必要的，但是在收益领域寻求风险是不充分的。当 $x < 0$ 时，恰好相同的条件是必要的但不足以进行风险规避。这种分析将收益范围内的风险寻求和损失范围内的风险规避限制在了小概率中，其中权重高估将持续。实际上，这些是出售彩票和保险单的典型条件。在前景理论中，小概率事件的权重高估对于博弈和保险来说有利，而S形效用函数倾向于抑制这两种行为。

虽然前景理论预测了保险和博弈中的小概率事件，但我们认为目前的分析远不足以完全解释这些复杂现象。事实上，两项实验研究、调查研究及对经济行为（如服务和医疗保险）的观察结果都表明，购买保险通常会扩展到中等概率范围，而小的灾难概率有时会完全被忽视

（Slovic, Fischhoff, Lichtenstein, Corrigan & Coombs, 1977; Kunreuther et al., 1978）。此外，有证据表明，决策问题设置中的微小变化会对保险的吸引力产生显著影响（Slovic, Fischhoff,

Lichtenstein, Corrigan & Coombs, 1977)。除了对不确定性和金钱的纯粹态度之外，保险行为的全面理论还应考虑诸如安全价值、谨慎的社会规范、随时间推移的对大量小额支付的厌恶、关于概率和结果的信息等其他因素。这些变量的影响可以在本框架内得以描述，如随着参考点的变化、效用函数的变换，或对概率或决策的操纵权重增加。其他影响可能需要引入该方案中未考虑的变量或概念。

参考点的改变

到目前为止，我们提到的收益和损失是由前景被设置时所获得或支付的金额来定义的，参考点被视为现状或一个人的流动资产。尽管对于大多数选择题而言这可能是正确的，但是在某些情况下，得失是根据不同的前景或前景水平编码记录的收益损失情况。例如，从月度工资支票中意外退税是一种损失，而不是基于现状的前景收益的减少。同样地，一个比竞争对手更成功地度过了低迷期的企业家，相对于预估的更大损失来说，他可能会将一小笔损失解释为收益。

上述示例中的参考点对应于人们前景达到的资产量。参考点与流动资产所处的量之间的差异也会存在，因为人们尚未适应财富的变化（Markowitz, 1952）。想象一下，某人进行了一项商业风险投资，已经损失了2 000英镑，他现在面临着一个选择：是肯定能赚到1 000英镑，还是有机会赚到2 000英镑，或者什么都赚不到。如果还没有适应损失，他很可能将问题编码为（-2 000, 0.50）和（-1 000）之间的选择，而不是编码为（2 000, 0.50）和（1 000）之间的选择。正如我们所看到的，前者的表现形式比后者更具冒险性。

参考点的更改会改变前景的优先顺序。特别需要指出的是，目前的理论暗指选择问题的消极转化，如对近期损失的不完全适应，在某些情况下增加了风险寻求。具体来说，如果风险前景（ $x, p; -y, 1-p$ ）可以接受，那么（ $x-z, p; y-z, 1-p$ ）优于（ $-z$ ），其中 $x, y, z > 0$ 且 $x > z$ 。

为证明这一理论，请注意当且仅当 $\pi(p) u(x) = -\pi(1-p) u(-y)$ 时

$$V(x, p; y, 1-p) = 0$$

那么

$$V(x-z, p; -y-z, 1-p)=$$

$$\pi(p)v(x-z) + \pi(1-p)v(-y-z) > \quad (\text{由 } v \text{ 的特性得出})$$

$$\pi(p)v(x) - \pi(p)v(z) + \pi(1-p)v(-y) + \pi(1-p)v(-z)=$$

$$-\pi(1-p)v(-y) - \pi(p)v(z) + \pi(1-p)v(-y) +$$

$$\pi(1-p)v(-z)= \quad (\text{通过代入得出})$$

$$-\pi(p)v(z) + \pi(1-p)v(-z) >$$

$$v(-z)[\pi(p) + \pi(1-p)] > \quad (\text{因为 } v(-z) < -v(z))$$

$$v(-z) \quad (\text{由次确定性得出})$$

这一分析表明，如果一个人不能坦然接受自己的损失，他很可能会接受自己原本无法接受的博弈。在投注日期间投注远期投资的倾向性的增加，为假设未能适应损失或获得前景收益引发风险寻求提供了一些支持（McGlothlin, 1956）。再举一个例子，一个希望购买保险的人，也许是因为他过去拥有保险或者他的朋友这么做了。这个人可以将支付溢价 y 以防止损失 x 的决定编码为 $(-x+y, p; y, 1-p)$ 和 (0) 之间的选择而不是 $(-x, p)$ 和 $(-y)$ 之间的选择。前面的论证要求保险在前者中比在后者中更具吸引力。

参考点转移的另一种重要情况是，当一个人按照决策分析中所提倡的最终资产（而不是像人们通常所做的那样按照损益）来表述他的决策问题时，就会出现参考点转移。在这种情况下，参考点在财富规模上被设为零，效用函数很可能在任何一处都是凹的（Spetzler, 1968）。根据目前的分析，除了低概率的博弈，这个公式基本上消除了风险寻求。根据最终资产明确地设置决策问题，可能是消除损失领域中风险寻求的最有效方式。

许多经济决策都涉及这样一种交易：一个人为了得到一个理想的前

景而支付金钱。目前的决策理论分析了现状与替代状态之间的比较，包括从获得的前景中减去其成本。例如，人们把决定是否为博弈（1 000, 0.01）支付10美元视为在（990, 0.01; -10, 0.99）和（0）之间进行选择。在此分析中，购买正面前景的意愿等同于愿意接受相应的混合前景。

普遍存在的未能将无风险和高风险预期的失败进行整合的现象（在隔离效应中表现得更为突出）表明，人们在决定是否参与博弈时，不太可能执行从结果中减去成本的操作。相反，我们认为人们通常会分别评估博弈及其成本，如果综合价值为正，就决定购买。因此，如果 $\pi(0.01) v(1\ 000) + v(-10) > 0$ ，那么人们将以10美元的价格参与博弈（1 000, 0.01）。

如果这个假设是正确的，那么决定为（1 000, 0.01）支付10美元就不再等同于决定接受博弈（990, 0.01; -10, 0.99）。此外，前景理论还指出，如果一个人认为（ $x(1-p), p; -px, 1-p$ ）和（0）无关紧要，那么人们不会为购买前景（ x, p ）而支付 px 。因此，人们在决定是否接受一场公平博弈时，比在决定是否以公平的价格投注一场博弈时表现出了更为明显的风险寻求特点。参考点的位置及选择问题的编码和编辑方式成了决策分析中的关键因素。

扩展

为了包含更广泛的决策问题，前景理论应该向几个方向扩展。有些概括是直接的；其他的则需要进一步发展。我们将等式（1）和等式（2）推广到具有任意数量结果的前景中是直接的。但是，当结果的数量很大时，可能会调用额外的编辑操作来简化计算。如何将复杂的选择，如复合前景，简化为简单的选择，尚待研究。

虽然本章主要涉及货币结果，但这一理论能很容易地被应用于涉及其他属性的选择，如生活质量或由于一项政策性决定而可能丧失或挽救的生命数目。建议性货币效用函数的主要属性也应适用于其他属性。特别需要指出的是，我们预计结果将被编码为相对于中性参考点的收益或损失，而损失将大于收益。

该理论还可以推广到选择的典型情况中，即结果的概率没有明确给出时的情况。在这种情况下，决策权重必须附加到特定的事件上，而不

是指定的概率上，但是人们希望它们能显示出权重函数所具有的基本属性。例如，如果A和B是互补事件但又不是确定的， $\pi(A) + \pi(B)$ 应小于1，这是确定性的自然类似物。

与某一事件相关的决策权重主要取决于该事件的感知可能性，该可能性可能会受到主要偏差的影响（Tversky & Kahneman, 1974）。此外，决策权重可能会受到其他考虑因素的影响，如模糊性。事实上，丹尼尔·埃尔斯伯格（Daniel Ellsberg）和费尔纳的研究表明，模糊性降低了决策权重（Ellsberg, 1961; Fellner, 1965）。因此，次确定性更应该被视为模糊性而不是清晰的可能性。

目前对人们在风险期权之间的偏好的分析形成了两个主题。第一个主题涉及确定如何看待前景的编辑操作。第二个主题涉及支配得失评价和不确定结果权重的判断原则。这两个主题似乎为风险下的选择的描述性分析提供了一个有用的框架。

附录

在本附录中，我们描述了前景理论的定理分析。由于完整的独立处理步骤是漫长而乏味的，我们只是概述了基本步骤，并展示了建立等式（1）的双线性表现形式所需的关键序数属性。类似的方法可以扩展到公理化等式（2）中。

考虑形式 $(x, p; y, q)$ 的所有常规前景的集合，其中 $p+q < 1$ 。将 $p+q = 1$ 扩展到常规前景非常简便。设 $\succeq$ 表示毫无联系的关系， $(x, p; y, q) \simeq (y, q; x, p)$ 自然成立。我们还假设，正如我们的符号所暗示的那样， $(x, p; 0, q) \simeq (x, p; 0, r)$ 和 $(x, p; y, 0) \simeq (x, p; z, 0)$ 。也就是说，空值和不可能事件具有可乘零的属性。

注意，前景的表示形式 [等式（1）] 在概率-结果中是可加的。因此，可以应用可加性联合测量理论来获得保留偏好顺序的尺度V，其保留了偏好顺序以及在两个参数中的尺度f和g的区间，使得

$$V(x, p; y, q) = f(x, p) + g(y, q)$$

用于推导该公式的关键定理有：

独立性：当且仅当 $(x', p'; y, q) \succeq (x', p'; y', q')$ 时， $(x, p; y, q)$

$\succeq (x, p; y'q')$ 。

取消性：如果 $(x, p; y'q') \succeq (x', p'; y, q)$ 和 $(x', p'; y'', q'') \succeq (x'', p''; y', q')$ ，那么 $(x, p; y'', q'') \succeq (x'', p''; y, q)$ 。

可解性：对于某些结果 z 和概率 r ，如果 $(x, p; y, q) \succeq (z, r) \succeq (x, p; y'q')$ ，那么存在 y'' 和 q'' 使得 $(x, p; y''q'') \approx (z, r)$ 。

研究表明，如果偏好顺序符合阿基米德定理，这些条件足以构造所需的加法表示形式（Debreu, 1960; Krantz, Luce, Suppes & Tversky, 1971）。此外，由于 $(x, p; y, q) \approx (y, q; x, p)$ ， $f(x, p) + g(y, q) = f(y, q) + g(x, p)$ ，并且设 $q=0$ ，则可得到 $f=g$ 。

接下来，考虑 (x, p) 表示所有前景的集合，其中只有一个非零的结果。在这种情况下，双线性模型减少到 $V(x, p) = \pi(p) v(x)$ 。这是在相关实验中得到研究的乘法模型（Roskies, 1965; Krantz, Luce, Suppes & Tversky, 1971）。为了构造乘法表达式，我们假设概率与结果对的顺序满足独立性、可取消性、可解性和阿基米德定理。此外，我们假设要根据相关实验来确保符号的正确相乘方式（Krantz, Luce, Suppes & Tversky, 1971）。值得注意的是，在相关研究中使用的可解性公理必须被削弱，因为概率因子只允许有限的可解性（Roskies, 1965; Krantz, Luce, Suppes & Tversky, 1971）。

结合加法和乘法表达式可得

$$V(x, p; y, q) = f[\pi(p) v(x)] + f[\pi(q) v(y)]$$

接着，我们引入了一个新的分配定理，当且仅当 $(x, q; y, q) \approx (z, q)$ 时

$$(x, p; y, p) \approx (z, p)$$

将此定理应用于上述表达式，可得

$$f[\pi(p) v(x)] + f[\pi(p) v(y)] = f[\pi(p) v(z)]$$

则

$$f[\pi(q) v(x)] + f[\pi(q) v(y)] = f[\pi(q) v(z)]$$

假设在不失一般性的条件下, $\pi(q) < \pi(p)$, 并且设 $\alpha = \pi(p) \vee(x)$, $\beta = \pi(p) \vee(y)$, $\gamma = \pi(p) \vee(z)$ 和 $\theta = \pi(q) / \pi(p)$, 则 $f(\alpha) + f(\beta) = f(\gamma)$ 意味着, 对于所有 $0 < \theta < 1$, $f(\theta\alpha) + f(\theta\beta) = f(\theta\gamma)$ 。

因为 f 是严格单调的, 所以我们可以设 $\gamma = f^{-1}[f(\alpha) + f(\beta)]$ 。因此, $\theta\gamma = \theta f^{-1}[f(\alpha) + f(\beta)] = f^{-1}[f(\theta\alpha) + f(\theta\beta)]$ 。

该函数公式的解是 $f(\alpha) = k\alpha^c$ [1]。因此, 对于 $k, c > 0$, $V(x, p; y, q) = k[\pi(p) \vee(x)]^c + k[\pi(q) \vee(y)]^c$ 。前景的双线性模式通过重新定义标度 π 、 $\vee$ 和 V , 来引入常数 k 和 c 。

感谢斯坦福大学行为科学高级研究中心的支持。感谢戴维·克兰茨在设计附录方面提供的帮助。

参考文献

Aczél, J. (1966). *Lectures on functional equations and their applications*. New York, NY: Academic Press.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine. *Econometrica*, 21, 503–546.

Anderson, N. H., & Shanteau, J. C. (1970). Information integration in risky decision making. *Journal of Experimental Psychology*, 84, 441–451.

Arrow, K. J. (1971). *Essays in the theory of risk-bearing*. Chicago, IL: Markham.

Barnes, J. D., & Reinmuth, J. E. (1976). Comparing imputed and actual utility functions in a competitive bidding setting. *Decision Sciences*, 7, 801–812.

Coombs, C. H. (1975). Portfolio theory and the measurement of risk. In M. F. Kaplan & S. Schwartz (Eds.), *Human judgment and decision processes* (pp. 63–85). New York, NY: Academic Press.

- Davidson, D., Suppes, P., & Siegel, S. (1957). *Decision-making: An experimental approach*. Stanford, CA: Stanford University Press.
- Debreu, G. (1960). Topological methods in cardinal utility theory. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences* (pp. 16–26). Stanford, CA: Stanford University Press.
- Edwards, W. (1962). Subjective probabilities inferred from decisions. *Psychological Review*, 69, 109–135.
- Ellsberg, D. (1961). Risk, ambiguity and the savage axioms. *Quarterly Journal of Economics*, 75, 643–669.
- Fellner, W. (1961). Distortion of subjective probabilities as a reaction to uncertainty. *Quarterly Journal of Economics*, 75, 670–690.
- Fellner, W. (1965). *Probability and profit—A study of economic behavior along Bayesian lines*. Homewood, IL: Richard D. Irwin.
- Fishburn, P. C. (1977). Mean-risk analysis with risk associated with below-target returns. *American Economic Review*, 67, 116–126.
- Fishburn, P. C., & Kochenberger, G. A. (1979). Two-piece von Neumann–Morgenstern utility functions. *Decision Sciences*, 10, 503–518.
- Friedman, M., & Savage, L. J. (1948). The utility analysis of choices involving risks. *Journal of Political Economy*, 56, 279–304.
- Fuchs, V. R. (1976). From Bismark to Woodcock: The “irrational” pursuit of national health insurance. *Journal of Law & Economics*, 19, 347–359.
- Galanter, E., & Pliner, P. (1974). Cross-modality matching of money against other continua. In H. R. Moskowitz, et al. (Eds.), *Sensation and measurement* (pp. 65–76). Dordrecht, the Netherlands: Reidel.
- Grayson, C. J. (1960). *Decisions under uncertainty: Drilling decisions by oil and gas operators*. Cambridge, MA: Harvard University Graduate School of Business.

Green, P. E. (1963). Risk attitudes and chemical investment decisions. *Chemical Engineering Progress*, 59, 35–40.

Grether, D. M., & Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *American Economic Review*, 69, 623–638.

Halter, A. N., & Dean, G. W. (1971). *Decisions under uncertainty*. Cincinnati, OH: South Western Publishing Co.

Hansson, B. (1975). The appropriateness of the expected utility model. *Erkenntnis*, 9, 175–194.

Helson, H. (1964). *Adaptation-level theory*. New York, NY: Harper.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: preferences and value tradeoffs*. New York, NY: Wiley.

Krantz, D. H., Luce, D. R., Suppes, P., & Tversky, A. (1971). *Foundations of measurement*. New York, NY: Academic Press.

Kunreuther, H., Ginsberg, R., Miller, L., Sagi, P., Slovic, P., Borkan, B., et al. (1978). *Disaster insurance protection: Public policy lessons*. New York, NY: Wiley.

Lichtenstein, S., & Slovic, P. (1971). Reversal of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, 89, 46–55.

MacCrimmon, K. R., & Larsson, S. (1979). Utility theory: Axioms versus paradoxes. In M. Allais & O. Hagen (Eds.), *Expected utility hypothesis and the Allais paradox* (pp. 333–409). Dordrecht, the Netherlands: Springer.

Markowitz, H. (1952). The utility of wealth. *Journal of Political Economy*, 60, 151–158.

Markowitz, H. (1959). *Portfolio selection: Efficient diversification of*

investments. New York, NY: Wiley.

McGlothlin, W. H. (1956). Stability of choices among uncertain alternatives. *American Journal of Psychology*, 69, 604–615.

Mosteller, F., & Nogee, P. (1951). An experimental measurement of utility. *Journal of Political Economy*, 59, 371–404.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32, 122–136.

Raiffa, H. (1968). *Decision analysis: Introductory lectures on choices under uncertainty*. Reading, MA: Addison-Wesley.

Roskies, R. (1965). A measurement axiomatization for an essentially multiplicative representation of two factors. *Journal of Mathematical Psychology*, 2, 266–276.

Savage, L. J. (1954). *The foundations of statistics*. New York, NY: Wiley.

Slovic, P., Fischhoff, B., Lichtenstein, S., Corrigan, B., & Coombs, B. (1977). Preference for insuring against probable small losses: insurance implications. *Journal of Risk and Insurance*, 44, 237–258.

Slovic, P., & Tversky, A. (1974). Who accepts savage's axiom? *Behavioral Science*, 19, 368–373.

Spetzler, C. S. (1968). The development of corporate risk policy for capital investment decisions. *IEEE Transactions on Systems Science and Cybernetics*, SSC-4, 279–300.

Swalm, R. O. (1966). Utility theory—Insights into risk taking. *Harvard Business Review*, 44, 123–136.

Tobin, J. (1958). Liquidity preferences as behavior towards risk. *Review of Economic Studies*, 26, 65–86.

Tversky, A. (1967). Additivity, utility, and subjective probability.

Journal of Mathematical Psychology, 4, 175–201.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76, 31–48.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, 79, 281–299.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.

van Dam, C. Another Look at Inconsistency in Financial Decision-Making, presented at the Seminar on Recent Research in Finance and Monetary Economics, Cergy-Pontoise, March, 1975.

von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Williams, A. C. (1966). Attitudes toward speculative risks as an indicator of attitudes toward pure risks. *Journal of Risk and Insurance*, 33, 577–586.

第5章 理性选择和决策框架

阿莫斯·特沃斯基
丹尼尔·卡尼曼

现代风险决策理论不是从风险和心理分析中产生的，而是从针对机会博弈的逻辑分析中产生的。该理论被认为是一个理想化的决策者的规范性模型，而不是对真实人的行为的描述。用经济学家熊彼特的话来说，它“更应该被称为选择的逻辑，而不是价值心理学”（Schumpeter, 1954）。

使用规范分析来预测和解释实际行为有几个支持论据。第一，人们普遍认为自己在追求目标方面是有效的，特别是当他们有动力和机会从经验中学习时。因此，将选择描述为一个最大化过程似乎是合理的。第二，竞争对于理性的个人和组织有利。最优决策增加了在竞争环境中的生存机会，少数理性的个体有时可以将理性强加于整个市场。第三，理性选择定理的直觉吸引力使得由这些定理派生的理论似乎应该为我们提供一个可接受的有关选择行为的解释。

本章的论点是，尽管有这些先验性论点，但选择的逻辑并没有为描述决策理论提供充分的基础。我们认为来自规范性模型行为的实际偏差过于普遍，因此无法忽视；过于系统化，因此不能被视为随机错误；过于基础，因此无法通过放宽规范体系来适应。我们首先勾画出了对理性选择理论基础的分析和，然后表明决策者普遍违反了理论的最基本规则。我们从这些发现中得出结论：规范性和描述性分析无法协调。因此，我们提出了一个描述性的选择模型，它解释了规范理论中的异常偏好。

规范性规则的层次结构

现代风险决策理论的主要成就是从理性选择的简单原则中推导出了预期效用规则，而这些规则无法应用于长期（von Neumann & Morgenstern, 1944）。对期望效用理论基础的定理分析揭示了除可比性和连续性之外的4个实质性假设：消除性、传递性、占优性和不变性。

实质性假设可以通过它们的规范性诉求来排序，从被许多理论家质疑的消除条件，到被所有人接受的不变性。我们接下来简要讨论这些假设。

消除性。产生期望效用理论的关键定性属性是消除性或消除无论如何选择结果都相同的效应。这个概念已经被不同的形式特征所体现，例如冯·诺伊曼和摩根斯坦的替代定理（von Neumann & Morgenstern, 1944）、伦纳德·萨维奇（Leonard Savage）的扩展确定性原则（Savage, 1954），以及卢斯和克兰茨的独立条件（Luce & Krantz, 1971）。因此，如果 A 优先于 B ，那么明天下雨赢得 A 的前景就应该优先于明天下雨赢得 B 的前景，因为如果明天不下雨，这两个前景会产生相同的结果（什么也没有赢得）。在期望效用最大化过程中，为了将对不同前景的偏好表示出来，消除性就是有必要的。消除性的主要观点是，实际上只有一个情形会成为现实，因此单独评估每个情形的选择结果是合理的。因此，各种选项之间的选择应只取决于产生不同结果的情形。

传递性。风险选择模型和无风险选择模型的一个基本假设是偏好的传递性。这个假设对于用序数效用尺度 u 表示偏好来说是必要的且本质上是充分的，当 $u(A) > u(B)$ 时， A 优先于 B 。因此，如果可以为每个选项分配一个不依赖于其他可用选项的值，那么传递性就得到了满足。传递性在单独评估期权时可能成立，但在预期的后果取决于与之进行比较的备选方案时则不成立，如考虑到后悔这一因素时。传递性的一个常见论点是，循环偏好可以支持“金钱泵”，在这个泵中，非传递性的人会被诱导为一系列返回初始选项的交换进行支付。

占优性。这可能是理性选择中最明显的原则：如果一个选项在一个情形中优于另一个选项，并且前者与后者在所有其他情形中至少也是一样好的，则应选择这一占优选项。一个略微更强的条件，即随机占优条件，表明对于一维风险预期而言，如果 A 的累积分布在 B 的累积分布的右侧，那么 A 优于 B 。占优性比消除性和传递性更简单且更具吸引力，它是选择规范理论的基石。

不变性。主张规范性地位的选择理论的一个基本条件是不变性：同一选择题的不同表现形式可能会产生相同的偏好。也就是说，选项之间的偏好应该独立于对它们进行的描述。经过反思后，决策者会将两个特征视为对同一问题的可供选择的描述，即使没有这种反思带来的益处，这两个特征也应该导致相同的选择。不变性（或可拓性）原则是很基础的，因此常在描述选项时默认使用，而不是被明确地表述为一个可检验的定理（Arrow, 1982）。例如，将选择对象描述为随机变量的决策模

型都假定应该对相同随机变量的可选择形式进行相同的处理。不变性捕捉规范性的直觉，即不影响实际结果的形式的变化不应影响选择。哈蒙德对一个相关的概念进行了讨论，其被称为结果主义（Hammond, 1985）。

期望效用理论的这4个原则可以通过其规范性诉求来排序。不变性和占优性似乎是必不可少的，但传递性可能会受到质疑，消除性也可能被许多作者拒绝。事实上，阿莱斯（Allais, 1953）和埃尔斯伯格（Ellsberg, 1961）提出的巧妙的反例使得几位理论家放弃了消除性和前景原则，选择支持更一般的陈述。大多数模型都假定具有传递性、占优性和不变性（Hansson, 1975; Allais, 1979; Hagen, 1979; Machina, 1982; Quiggin, 1982; Weber, 1982; Chew, 1983; Fishburn, 1983; Schmeidler, 1984; Segal, 1984; Yaari, 1984; Luce & Narens, 1985）。其他人放弃了传递性但支持不变性和占优性（Bell, 1982; Fishburn, 1982, 1984; Loomes & Sugden, 1982）。这些理论家通过削弱规范理论来回应观察到的消除性和传递性的反例，以保持其作为描述性模型的地位。但是，这种策略不能扩展到我们要记录的占优性和不变性的失败中。由于不变性和占优性在规范上是必不可少的，并且描述性分析是无效的，因此理性决策理论无法为我们提供对选择行为的充分描述。

接下来，我们将通过几个例子说明不变性和占优性无效的情况，然后回顾描述性分析，将这些失败追溯到管理前景框架、结果评估和概率加权规则的联合效应中。我们将先对支持本解释的几种选择现象进行描述。

不变性无效的案例

在本节中，我们考虑了两个说明性案例，其中违反了不变性的条件；同时我们还对其产生的因素进行了讨论。

第一个案例来自对医学治疗之间的偏好的研究（McNeil et al., 1982）。被试看到的是两种肺癌治疗结果的统计信息。一部分被试得到的死亡率方面的统计数据是相同的，而其他被试得到的存活率方面的统计数据是相同的。然后被试被要求选择出他们的首选治疗方法。相关信息如下所示：(8)

问题 1（存活框架）

手术：手术期间有 100 名患者接受手术治疗，术后阶段有 90 名患者存活，其中 68 名患者在治疗结束 1 年后仍存活，34 名患者在治疗结束 5 年后仍存活。

放射治疗：100 名患者接受放射治疗，治疗期间无一例死亡，其中 77 名患者在治疗结束 1 年后仍存活，22 名患者在治疗结束 5 年后仍存活。

问题 1（死亡框架）

手术：在手术期间或术后，100 名接受手术的患者中 10 名患者死亡，其中 32 名患者在治疗结束 1 年后死亡，66 名患者在治疗结束 5 年后死亡。

放射治疗：100 名患者接受放射治疗，治疗期间无一例死亡，其中 23 名患者在治疗结束 1 年后死亡，78 名患者在治疗结束 5 年后死亡。

这些描述性语句中看似无关紧要的差异产生了显著的效果。更愿意接受放射治疗的被试的总体百分比从存活框架中的 18% ($N=247$) 上升到了死亡框架中的 44% ($N=336$)。从将死亡率由 10% 降低到 0 而不是将生存率由 90% 增加到 100% 来看，放射治疗的优势显然比手术的优势更明显。对于有经验的医生或具有较高统计学素养的商科学生而言，他们感知到的框架效应并不比一组临床患者感知到的框架效应小。

我们接下来列出一个涉及风险预期与货币结果的决策案例。每名被试要做出两个选择，一个选择在有利前景之间进行，另一个选择在不利前景之间进行（Tversky & Kahneman, 1981）。假设两个选定的前景是相互独立的。

问题 2 ($N = 150$): 想象一下, 你面临着下列并行决策。首先检查两个决策, 然后指出你喜欢的选项。

第一项决策:

- A. 确保赢得 240 美元 [84%]
- B. 25% 的概率赢得 1 000 美元, 75% 的概率没有任何收益 [16%]

第二项决策:

- C. 确定损失 750 美元 [13%]
- D. 75% 的概率损失 1000 美元, 25% 的概率没有任何损失 [87%]

用 N 表示被试总数, 方括号中的数字代表选择每个选项的被试百分比 (除非特别说明, 否则相关数据都来自斯坦福大学和加拿大不列颠哥伦比亚大学的本科生)。第一项决策中的多数选择是风险规避, 而第二项决策中的多数选择是风险寻求。这是一种常见的模式: 涉及收益的选择通常会表现为风险规避, 涉及损失的选择通常会表现为风险寻求, 除非输赢的概率很小 (Fishburn & Kochenberger, 1979; Kahneman & Tversky, 1979; Hershey & Schoemaker, 1980)。

由于被试同时考虑了两个决策, 他们实际上表达了对组合A和D及对组合B和C的偏好。但是, 被试没有选择的组合实际上比首选组合更占优势。合并后的选项如下所示。

A 和 D: 25% 的概率赢得 240 美元, 75% 的概率损失 760 美元

B 和 C: 25% 的概率赢得 250 美元, 75% 的概率损失 750 美元

当选项以这种并行的形式呈现时, 人们总是选择优势更为明显的选项。而在问题2中, 73%的被试选择了优势更为明显的组合A和D, 只有3%的被试选择了B和C。这两种对比违反了不变性。调查结果也支持一般性观点, 即不变性的无效可能会导致违背随机占优性, 反之亦然。

被试显然在问题2中分别评估了第一项和第二项决策, 他们表现出了在收益中规避风险, 在损失中寻求风险的标准模式。看到这些问题的人非常惊讶地发现, 他们以为非常合理的两种偏好的组合使自己做出了一种处于优势地位的选择。在用真实货币支付的由问题2按比例缩小得

到的模式中，研究人员也得到了相同的结果（Tversky & Kahneman, 1981）。

如前面的案例所示，决策问题框架的变化产生了系统性违反不变性和占优性的行为，这种行为无法在规范的基础上得到保障。对可以确保偏好不变性的两种机制：权威性描述和前景精算价值的使用进行研究是有益的。

如果将相同前景的所有公式转换为标准的权威性描述（如相同随机变量的累积概率分布），那么不变性将保持不变，因为随后人们将以相同的方式对不同版本进行评估。例如，在问题2中，如果在评估之前将这两个决策的结果进行汇总，那么不变性和占优性都将得到保留。同样，如果结果是根据一个主要框架（如生存率）编码的，那么在不同版本的医学问题中人们会做出同样的选择。在研究中观察到的不变性的失效表明，人们不会自发地合并前景或将所有结果转化为共同框架。

在决策问题中无法构建权威性描述，这与其他认知任务形成了鲜明的对比。在其他认知任务中，这种描述是自动、毫不费力地生成的。特别是我们的视觉体验主要由权威性描述构成：当我们在物体周围移动或光照发生变化时，物体的大小、形状、亮度或颜色似乎不会发生变化。在昏暗的光线下，从一个倾斜的角度看到一个白色的圆，它是圆形和白色的，而不是椭球体和灰色的。权威性描述也会在语言理解的过程中生成，听众会迅速将他们听到的大部分内容重新编排成一个抽象的命题形式，不再区分它们的主动和被动语态，并且通常不区分实际所说的内容，以及不再区分是否是从暗示或预设的角度看问题的（Clark & Clark, 1977）。很可惜，以感知和句子转换为标准形式的心理机制并不能自动适用于选择过程。

如果对前景的评估在概率和货币价值上分别是线性或接近线性的，那么即使在没有权威性描述的情况下也可以满足不变性。如果人们通过他们的精算价值来预设风险前景，那么不变性和占优性将永远存在。尤其是医疗问题的死亡率和生存率之间是没有区别的。因为结果和概率的评估通常是非线性的，并且因为人们不会自发地构建决策的权威性描述，所以不变性通常会失败。因此，我们假设不变性的选择规范性模型无法对选择行为进行充分描述。在下一节中，我们将介绍风险选择的描述性说明，即前景理论，并探讨其后果。不变性的失效可以由控制选项描述方式的框架效应来解释，并与价值和信念的非线性相结合。

结果的框架和评估

前景理论将选择过程区分为两个阶段：框架和编辑阶段，然后是评估阶段（Kahneman & Tversky, 1979）。第一个阶段包括对决策问题的初步分析，该分析包括有效行为、突发事件和结果。框架由选择问题的呈现方式及决策者的规范、习惯和前景来控制。在评估之前执行的其他操作包括消除共同成分和消除被其他人视为占优的选项。在第二个阶段中，我们对被包装过的前景进行了评估，最高价值的前景是经过选择的。该理论区分了两种选择前景的方法：通过检测一个对另一个的占优或通过比较它们的价值。

为简单起见，我们将讨论设定在具有数值概率和货币结果的简单博弈的范围内。设 $(x; p; y; q)$ 表示一个前景，得到 x 的概率为 p ，得到 y 的概率为 q ，保持现状的概率为 $(1-p-q)$ 。根据前景理论，在收益和损失上定义了值 $v(\cdot)$ ，并且在所述概率上定义了决策权重 $\pi(\cdot)$ ，使得前景的总值等于 $\pi(p) v(x) + \pi(q) v(y)$ 。如果前景的所有结果都具有相同的符号，那么需要稍做修改。[\(9\)](#)

效用函数

根据马科维茨的研究，在前景理论中结果被描述为来自中性参考结果的正偏差或负偏差（收益或损失），被赋值为零（Markowitz, 1952）。然而，与马科维茨不同的是，我们认为效用函数通常为S形，在参考点上方凹入，在参考点下方凸出，如图5-1所示。因此，100美元收益和200美元收益之间的主观价值差异大于100美元收益和1200美元收益之间的主观差异。价值差异之间的相同关系适用于相应的损失。我们所提出的函数的性质是，边际变化的影响随着其与参考点在任一方向上距离的缩短而减小。这些关于效用函数的典型形状的假设可能不适用于毁灭性损失或特定金额具有特殊意义的情况。

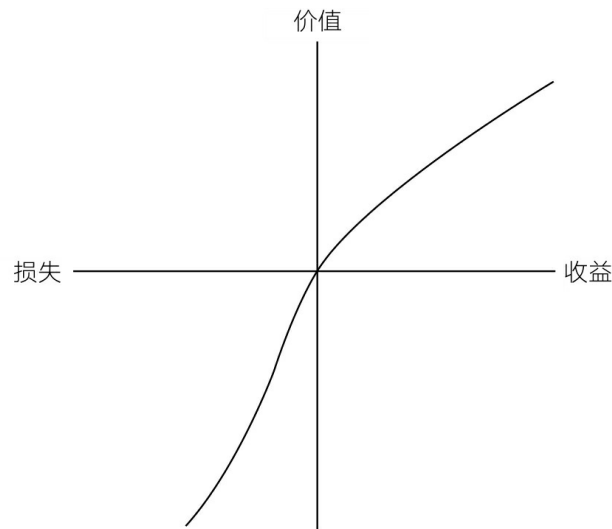


图5-1 典型的效用函数

效用函数的一个重要特性被称为损失厌恶，指的是人们对损失的反应比对收益的反应更极端。人们普遍不愿意接受投掷硬币的公平赌注，这表明人们对于失去一笔钱的痛苦超过了赢得相同数额金钱的乐趣。因此，我们提出的效用函数是在收益和损失上定义的，通常收益函数是凹陷的而损失函数是凸的，且损失函数比收益函数更陡峭。许多关于货币结果的风险选择（Fishburn & Kochenberger, 1979; Kahneman & Tversky, 1979; Hershey & Schoemaker, 1980; Payne, Laughhunn & Crum, 1980）和现实生活（Tversky, 1977; Eraker & Sox, 1981; Tversky & Kahneman, 1981; Fischhoff, 1983）的研究都支持效用函数的这些性质。损失厌恶也可能会使我们观察到人们愿意为了一件商品支付的金额与他们要求放弃的补偿之间的差异（Bishop & Heberlein, 1979; Knetsch & Sinden, 1984）。如果商品在前一种情况下被视为收益而在后者中被视为损失，那么效用函数表明了这种效应。

框架结果

我们将在下述问题中对结果框架，以及传统理论与当前分析之间的对比进行说明。

问题 3 ($N = 126$): 假设你比现在多了 300 美元。你必须从下列选项中做出选择:

100 美元的确定收益 [72%]

有 50% 的概率获得 200 美元, 50% 的概率未获得任何收益 [28%]

问题 4 ($N = 128$): 假设你比现在多了 500 美元。你必须从下列选项中做出选择:

肯定损失 100 美元 [36%]

有 50% 的概率没有任何损失, 50% 的概率失去 200 美元 [64%]

正如效用函数所表明的那样, 问题3中的多数选择都表现出了风险规避, 问题4中的多数选择表现为风险寻求, 尽管这两个问题基本相同。在这两种情况下, 人们都面临着在确定得到400美元和以均等的机会得到500美元或300美元之间做出选择。问题4是从问题3中得到的, 将最初值增加200美元, 同时从选项中减去这个增量。这种变化对偏好有很大影响。附加问题表明, 初始财富200美元的变化对选择几乎没有影响。显然, 偏好对财富的微小变化非常不敏感, 但对参考点的相应变化高度敏感。这些观察结果表明, 价值的有效载体是收益和损失, 或财富的变化, 而不是理性模型所暗示的财富状态。

在问题3和问题4中观察到的共同偏好模式具有特殊的意义, 因为它不仅违反了期望效用理论, 而且实际上违反了所有其他基于规范的选择模型。特别是, 这些数据与D. E. 贝尔 (D. E. Bell)、G. 卢默斯 (G. Loomes) 和罗伯特·萨格登 (Robert Sugden) 提出的遗憾模型 (Bell, 1982; Loomes & Sugden, 1982) 以及P. 菲什伯恩的定理 (Loomes & Sugden, 1982; Fishburn, 1982) 不一致。这是因为问题3和问题4产生了相同的结果和相同的遗憾结构。此外, 即使没有相应的赋值变化使问题在延伸上等价, 遗憾理论也不能适应问题3中的风险规避和问题4中的风险寻求的组合。

如上述问题所示, 可以通过将结果分解为风险成分和无风险成分来引发参考点变化。也可以通过只标记结果来改变参考点, 如以下问题所示 (Tversky & Kahneman, 1981)。

问题 5 ($N = 152$): 想象一下, 美国正在爆发一种不寻常的疾病, 预计将导致 600 人死亡。相关人员已经提出了两种防治该疾病的替代性方案。假设针对这些计划的后果的确切科学估计如下所示。

如果采用 A 计划, 将有 200 人获救。[72%]

如果采用 B 计划, 有 $1/3$ 的概率可以救活 600 人, 同时有 $2/3$ 的概率救不活任何人。[28%]

在问题5中, 结果以积极的术语 (挽救生命) 得以描述, 因此多数选择表现为风险规避。肯定能挽救200人的前景比具有同等前景值的风险前景更有吸引力。另一组被试获得了相同的起始故事, 并附有以下对替代性计划的描述。

问题 6 ($N = 155$):

如果采用 C 计划, 将有 400 人死亡。[22%]

如果采用 D 计划, 那么有 $1/3$ 的概率无人死亡, 同时有 $2/3$ 的概率 600 人全部死亡。[78%]

在问题6中, 结果由负面的词语 (失去生命) 来表示, 因此多数选择表现为风险寻求。确定400人死亡比有 $2/3$ 的概率死亡600人的结果更不容易被人接受。然而, 问题5和问题6基本相同。它们的不同之处仅在于前者以挽救生命的数量为框架 (相对于如果不采取行动, 预计会损失600人的生命), 而后者则以失去的生命数量为框架。

我们多次向相同的被试提供了两个版本, 并与他们讨论了两个框架引起的不一致性偏好。许多被试表示希望在“挽救生命”版本中保持风险规避, 并在“失去生命”版本中保持风险寻求, 尽管他们也表示希望自己的答案保持一致。在其吸引力的持续存在下, 框架效应更像是视觉错觉而不是计算错误。

折扣和附加费用

也许经济分析中最独特的智力贡献是对替代机会的系统性考虑。经济思维的基本原则是机会成本和自付费用应该被平等对待。偏好应仅取

决于选项之间的相关差异，而不是这些差异的标记方式。这一原则与使偏好容易受到表面形式变化影响的心理倾向背道而驰。特别需要指出的是，有利于结果A而不是结果B的差异，有时可以通过建议B或A作为中性参考点来构建A或B的劣势。由于存在损失厌恶，当A为中性且B-A被评估为损失时，比B为中性且A-B被评估为增益时的差异更大。在某些情境下人们已经注意到了这种框架变化的重要性。

2017年诺贝尔经济学奖得主理查德·塞勒（Richard Thaler）注意到了将两种价格之间的差异标记为附加费或折扣的效果。放弃折扣比接受附加费更容易（Thaler, 1980）。相同的价格差异在前一种情况下被视为收益而在后一种情况下被视为损失。实际上，据说信用卡销售人员坚持认为现金和信用卡消费之间的任何价格差异都应被标记为现金折扣而不是信用附加费。可以用类似的思路来解释为什么商家对弹性需求的价格反应通常采取折扣或特殊让步的形式（Stigler & Kindahl, 1970）。相对于直接的价格上涨，客户可能会对最终取消此类临时安排表现出较小程度的抵制。人们对公平的判断也表现出了相同的模式（Kahneman, Knetsch & Thaler, 1986a）。

2005年诺贝尔经济学奖得主托马斯·谢林（Thomas Schelling）描述了在税收政策背景下引人注目的框架效应（Schelling, 1981）。他指出，税表可以通过将有两个孩子的家庭或无子女家庭作为默认情况来构建。有两个孩子的家庭和无子女家庭之间的税收差异在第一个框架中表现为自然免税（有两个孩子的家庭），在第二个框架中表现为税收最大化（无子女家庭）。这种看似无害的差异对收入、家庭规模和税收之间所需关系的判断有很大影响。谢林指出，他的学生拒绝在第一个框架中赋予富人比穷人更大的税收豁免额度，但支持在第二个框架中赋予无子女的富人比无子女的穷人更高的税收优惠。由于税收豁免和税收优惠是两种情况下相同税收差异的可选择性标签，因此该判断违背了不变性。以正面或负面的方式构建公共政策的后果可以极大地改变其吸引力。

货币错觉的概念有时适用于在高通货膨胀期间，工人接受不能保护其实际收入的名义工资增加的意愿，尽管他们会在没有通货膨胀的情况下极力抵制相应的减薪。本质上，名义工资的削减被认为是一种损失，而名义工资的增加却不能保证实际收入的增加。在对经济行为的公平性研究中我们观察到了货币错觉的另一种表现（Kahneman, Knetsch & Thaler, 1986b）。被试在电话采访中评估了下述行为的公平性，它以两个版本呈现，只是括号中的条款有所不同。

一家公司利润微薄。它位于一个经济衰退社区，失业率很高（但没有通货膨胀 / 通货膨胀率为12%）。该公司今年决定（降低7%的工资和薪水 / 工资仅增加5%）。

虽然两个版本中实际收入的损失非常接近，但是判断公司行为“不公平”或“非常不公平”的被试中选择名义减少的比例是62%，而选择名义增加的比例只有22%。

应用行为心理学家马克斯·巴泽曼（Max Bazerman）[\(10\)](#)在有关交易的实验研究中记录了框架效应（Bazerman, 1983）。当将交易的结果表示为收益或损失时，他比较了实验对象的表现。对损失分配进行交易的被试往往未能达成协议，而且往往未能找到帕累托最优（Pareto-optimal）解决方案。巴泽曼将这些观察结果归因于人们在损失领域寻求风险的一般倾向，这可能会增加双方的参与者冒险承担负面后果的意愿。

每当参与者将自己的让步视为损失，将从另一方获得的让步视为收益时，损失厌恶就会成为交易的障碍。例如，在对导弹事宜进行谈判时，与拆除导弹有关的主观安全损失可能比对手采取类似行动所产生的安全增量更大。如果双方都对他们所做出的让步和获得的让步按2:1的比例进行价值分配，那么由此产生的4:1的差距可能是很难弥合的。无论手牌是什么，如果谈判者用价值相等的“筹码”进行交易，将更容易达成协议。这种交易模式在日常购买行为中很常见，其中损失厌恶趋于消失（Kahneman & Tversky, 1984）。

随机事件的框架和权重

在期望效用理论中，每种可能性结果的效用由其概率加权。在前景理论中，不确定性结果的值乘以决策权重 $\pi(p)$ ， $\pi(p)$ 是 p 的单调函数但不是概率函数。加权函数 p 具有以下属性。第一，丢弃不可能的事件，即 $\pi(0)=0$ ，并且将标度正态化，使得 $\pi(1)=1$ ，但是该函数在端点附近表现不佳（Kahneman & Tversky, 1979）。第二，对于低概率， $\pi(p) > p$ ，但 $\pi(p) + \pi(1-p) \leq 1$ （确定性）。因此，低概率被突显，中等概率和高概率被弱化，而后者的效果比前者更明显。第三，对于所有 $0 < p, q, r \leq 1$ （次比率性）， $\pi(pr) / \pi(p) < \pi(pqr) / \pi(pq)$ 。也就是说，对于任何固定概率比 r ，当概率低时，决策权重的比率比概率高时更接近统一，如 $\pi(0.1) / \pi(0.2) > \pi(0.4) / \pi(0.8)$ 。满足这些特性的假设加权函数如图5-2所示。我们将在下一节中对其结果进行讨论。[\(11\)](#)

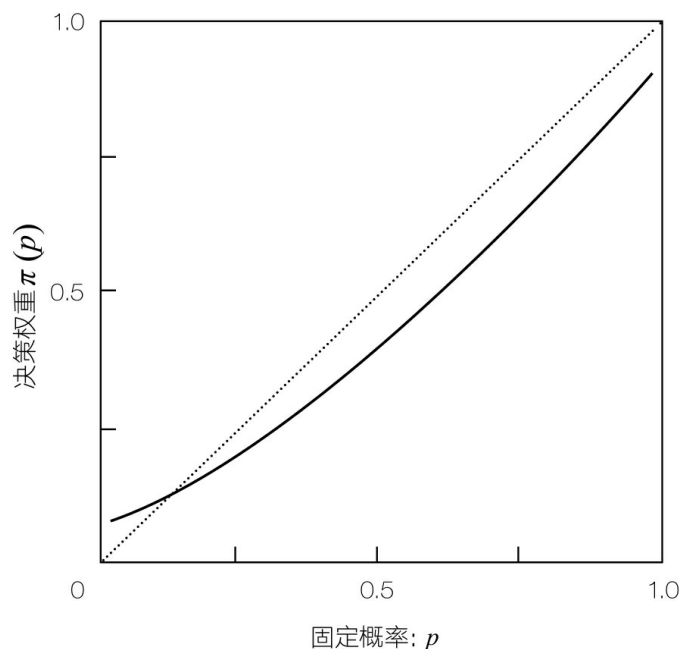


图5-2 典型的加权函数

非透明的占优性

加权函数的主要特征是相对于标度中部的可比较性差异，如 $\pi(0.3) - \pi(0.2)$ ，涉及确定性和不可能性的概率差异的权重过高，如 $\pi(1.0) - \pi(0.9)$ 或 $\pi(0.1) - \pi(0)$ 。特别是，对于小 p ， π 通常是次可加性的，如 $\pi(0.01) + \pi(0.06) > \pi(0.07)$ 。此属性可能会导致违反占优性，如下面两个问题所示。

问题 7 ($N = 88$)： 用每个盒子中不同颜色的弹珠的百分比来描述，且根据随机抽取的弹珠的颜色来判断你赢或输的货币金额。你更喜欢哪种彩票？

- | | | | | | |
|----|--------|-------|-------|-------|-------|
| A. | 90% 白色 | 6% 红色 | 1% 绿色 | 1% 蓝色 | 2% 黄色 |
| | 0 美元 | 赢得 | 赢得 | 输掉 | 输掉 |
| | | 45 美元 | 30 美元 | 15 美元 | 15 美元 |
| B. | 90% 白色 | 6% 红色 | 1% 绿色 | 1% 蓝色 | 2% 黄色 |
| | 0 美元 | 赢得 | 赢得 | 输掉 | 输掉 |
| | | 45 美元 | 45 美元 | 10 美元 | 15 美元 |

我们很容易看出选项B优于选项A：对于每种颜色，B的结果至少不比A的结果差。事实上，所有被试都选择了B而不是A。这种结果并不令人惊讶，因为优势关系是高度透明的，所以在没有进行进一步处理的情况下，占优势的前景就会被拒绝。下一个问题实际上与问题7完全相同，除了产生相同结果的颜色（B中的红色和绿色，A中的黄色和蓝色）被合并了。我们已经提出，如果没有检测到占优势的前景，那么决策者通常会执行该操作。

问题 8 (N = 124)：你更喜欢哪种彩票？

- | | | | | |
|----|--------|----------|----------|----------|
| C. | 90% 白色 | 6% 红色 | 1% 绿色 | 3% 黄色 |
| | 0 美元 | 赢得 45 美元 | 赢得 30 美元 | 输掉 15 美元 |
| D. | 90% 白色 | 7% 红色 | 1% 绿色 | 2% 黄色 |
| | 0 美元 | 赢得 45 美元 | 输掉 10 美元 | 输掉 15 美元 |

问题8的表述简化了选项，但掩盖了占优关系。此外，它还增强了选项C的吸引力。相对于具有两个正面结果和一个负面结果的选项C，选项D具有两个负面结果和一个正面结果。为了引导被试仔细考虑这些选项，我们告知他们，其中1/10是随机选择的博弈。虽然这个声明激起了一定程度的关注，但58%的被试选择了占优势地位的替代选项C。占优选项能否被发现，不仅取决于决策者的经验是否丰富，还取决于框架。在回答另一个问题时，大多数被试也分配给选项C比选项D更高的现金等价物。这些结果支持以下命题。第一，同一问题的两种表述方式引起了不同的偏好，违反了不变性。第二，如果规则是透明的，那么服从占优原则。第三，优势被一个框架掩盖了，在该框架中，较差的选项在确定状态下产生了更有利的结果（如绘制一个绿色弹珠）。第四，差异性偏好与决策权重的次可加性一致。透明的作用可以通过感知的例子来阐明。众所周知的缪勒-莱尔错觉（Müller-Lyer illusion）如图5-3所示：上面那条线段看起来比下面那条线段长，但它实际上更短。在图5-4中，相同的图案被嵌入矩形框架，我们就会发现突出边框的下面那条线段比上面那条线段更长。这种判断具有推理的性质，与图5-3所示的影响判断的感性印象形成了鲜明对比。类似地，在问题7中引入更精细的分区可以得出结论：在不评估它们的值的情况下，选项D优于选项C。问题8和问题1中的占优关系对于老练的决策者来说可能是显而易见的，但对于大多数被试来说并不明显。

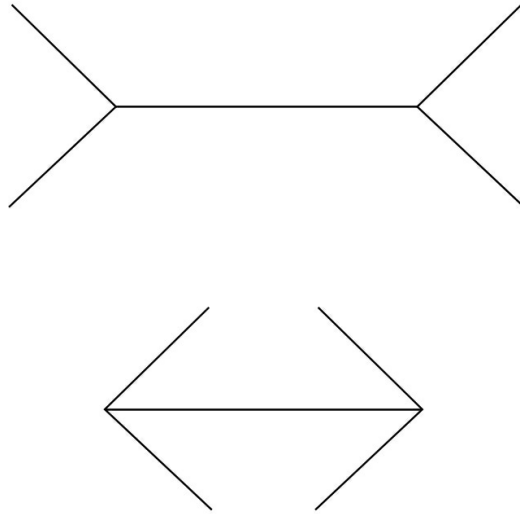


图5-3 缪勒-莱尔错觉

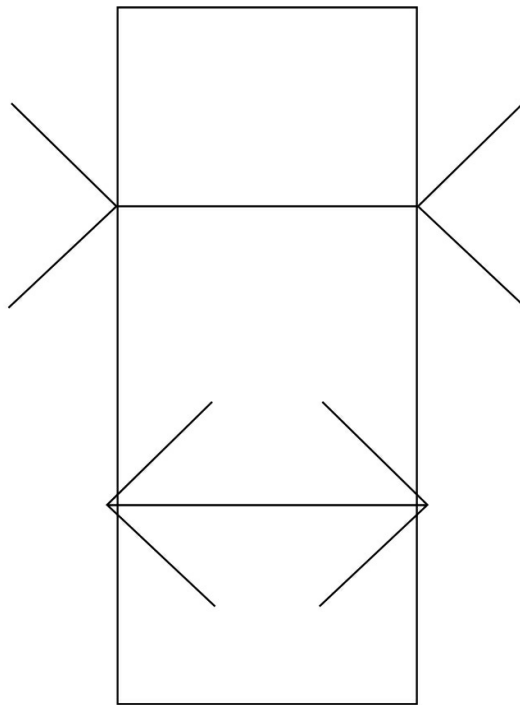


图5-4 缪勒-莱尔错觉的透明版本

确定性和假性

正如阿莱斯首次指出的那样，相对于可能的结果，确定性结果的权重过大会导致违反预期规则（Allais, 1953）。下列的一系列问题说明了阿莱斯发现的现象及其与概率权重和随机事件框架的关系（Tversky & Kahneman, 1981）。随机事件是通过从一个装有特定数量的有利和

不利弹珠的袋子里取出一颗弹珠来实现的。为了引导被试认真考虑选项，随机挑选的被试中只有1/10的人有机会参与他们选择的博弈。同样的被试按照这个顺序回答了问题9至问题11。

问题 9 ($N = 77$): 你更喜欢以下哪个选项?

- A. 肯定会赢得 30 美元 [78%]
- B. 有 80% 的概率赢得 45 美元, 有 20% 的概率未赢得任何奖金 [22%]

问题 10 ($N = 81$): 你更喜欢以下哪个选项?

- C. 有 25% 的概率赢得 30 美元, 有 75% 的概率未赢得任何奖金 [42%]
- D. 有 20% 的概率赢得 45 美元, 有 80% 的概率未赢得任何奖金 [58%]

注意, 问题10是通过将问题9中的概率降低1/4得到的。在期望效用理论中, 问题9中对选项A的偏好超过选项B意味着在问题10中对选项C的偏好超过选项D。与预测相反, 当获胜概率大幅降低时, 多数偏好从较低奖金(30美元)转为较高奖金(45美元)。我们把这种现象称为确定性效应, 因为获胜概率从1.0降低到0.25比从0.8降低到0.2的影响更大。在前景理论中, 问题9中的模式选项意味着 $v(45) \pi(0.80) < v(30) \pi(1.0)$, 而问题10中的模式选项意味着 $v(45) \pi(0.20) > v(30) \pi(0.25)$ 。然后, 如果以下条件成立, 那么我们观察到的对期望效用理论的违反是由 p 的曲率表示的(如图5-2所示)。

$$\frac{\pi(0.20)}{\pi(0.25)} > \frac{v(40)}{v(45)} > \frac{\pi(0.80)}{\pi(1.0)}$$

阿莱斯的问题引起了众多理论家的注意, 他们试图通过放宽消除性规则来构建确定性效应的规范性理论基础(Allais, 1979; Fishburn, 1982, 1983; Machina, 1982; Quiggin, 1982; Chew, 1983)。以下问题说明了一种相关的现象, 它被称为假确定性效应(pseudocertainty effect), 这种现象无法通过放宽消除性的方法来解决, 因为其中同样存在违反不变性的问题。

问题 11 ($N = 85$): 考虑下述二阶比赛。在第一阶段, 有 75% 的概率在没有获胜的情况下结束比赛, 并有 25% 的概率进入第二阶段。如果到达第二阶段, 你可以选择:

E. 肯定赢得 30 美元 [74%]

F. 有 80% 的概率赢得 45 美元, 有 20% 的概率未赢得任何奖金 [26%]

你必须在知道第一阶段的结果之前做出选择。

因为有 $1/4$ 的机会进入第二阶段, 在前景 E 中有 0.25 的概率赢得 30 美元, 前景 F 中有 $0.25 \times 0.80 = 0.20$ 的概率赢得 45 美元。因此, 问题 11 在概率和结果方面与问题 10 相同。然而, 这两个问题的偏好不同: 大多数被试在问题 11 中做出了风险规避的选择, 但在问题 10 中却没有。我们将这种现象称为假确定性效应, 因为实际上不确定性结果被加权了, 就好像它是确定性的一样。将问题 11 作为二阶比赛的框架来鼓励被试应用消除性: 未能达到第二阶段的事件在评估之前被舍弃了, 因为它在两个选项中产生了相同的结果。在这个框架中, 对问题 11 和问题 9 的评估是相似的。

虽然问题 10 和问题 11 在最终结果和概率方面是相同的, 但问题 11 有更大的诱发遗憾的可能性。例如, 一个决策者在问题 11 中选择了选项 F, 到达了第二阶段, 但未能赢得奖金, 而该决策者知道选项 E 会产生 30 美元的收益。此外, 在问题 10 中, 选择选项 D 并且未能获胜的被试无法确切地知道另一个选项的结果是什么。这种差异可能暗示了对遗憾方面的假确定性效应的另一种解释 (Loomes & Sugden, 1982)。然而, 确定性和假确定性效应在问题 9 至问题 11 的修订版本中同样强烈, 其中不同问题之间产生遗憾的概率是相等的。这一发现并不意味着遗憾机制在决策中不起作用 (Kahneman & Tversky, 1982)。它仅仅表明阿莱斯的例子和假确定性效应主要受决策权重的非线性和意外事件的框架的控制, 而不是由于遗憾的预期。 (12)

确定性和假确定性效应并不仅限于货币结果。以下问题说明了医学背景下同样存在这些现象。被试是参加加州医学会 (California Medical Association) 会议的 72 名医生。基本上相同的反应模式来自较大的一组 ($N=180$) 大学生样本。

问题 12 (N = 72): 在肿瘤的治疗中,有时需要在两种类型的治疗之间进行选择:第一,诸如广泛手术这样的根治性治疗,存在一些死亡风险;第二,中度治疗,例如有限的手术或放射治疗。以下每个问题描述了针对3种不同情况的两种替代性治疗的可能结果。在考虑每个病例时,假设患者是40岁的男性。假设如果不接受治疗,患者可能会很快死亡(一个月内);如果接受治疗,只能应用其中一种治疗方法。请说明你希望的治疗方法。

案例 1

治疗 A: 马上死亡的概率为 20%; 可以正常生活的概率为 80% 且预期寿命为 30 年。[35%]

治疗 B: 确定能正常生活, 预期寿命为 18 年。[65%]

案例 2

治疗 C: 马上死亡的概率为 80%; 可以正常生活的概率为 20%, 且预期寿命为 30 年。[68%]

治疗 D: 马上死亡的概率为 75%; 可以正常生活的概率为 25%, 且预期寿命为 18 年。[32%]

案例 3

考虑一个新病例, 其肿瘤可治疗的可能性为 25%, 不可治疗的可能性为 75%。如果肿瘤不可治疗, 将会很快死亡。如果肿瘤是可治疗的, 治疗结果如下:

治疗 E: 马上死亡的概率为 20%; 正常生活的概率为 80%, 且预期寿命为 30 年。[32%]

治疗 F: 确定能正常生活, 且预期寿命为 18 年。[68%]

该问题中的3种情况分别对应问题9至问题11, 并且我们观察到了相同的偏好模式。在案例1中, 大多数被试做出了一种风险规避的选择, 倾向于确定能存活但存活时间较短的选项。在案例2中, 中度治疗不再能确保生存, 大多数被试选择可提供更长预期寿命的治疗选项。特别

是，在案例1中选择B选项的64%的被试在案例2中选择了C选项。这是阿莱斯确定性效应的另一个例子。

案例2和案例3的比较为我们提供了假设的另一个例证。案例在相关结果及其概率方面是相同的，但偏好不同。特别是，在案例2中选择选项C的56%的被试在案例3中选择了选项F。条件框架使人们忽视了肿瘤不可治疗这一事件，因为在这种情况下两种治疗都是无效的。在这个框架中，治疗选项F拥有假性的优点。它似乎可以确保患者存活，但这种保证是以肿瘤的可治疗性为前提条件的。事实上，如果选择这个选项，只有0.25的机会存活一个月。

确定性和假确定性效应的结合对规范性和描述性选择理论之间的关系具有重要意义。我们的结果表明，人们在选择中，特别是在那些使其应用更透明的问题中，遵守了消除性。具体来说，我们发现人们在问题9、问题11以及问题12的案例3和案例1中做出了相同的选择。显然，人们会在二阶或嵌套结构中“删掉”一个会在所有选项中产生相同结果的事件。请注意，在这些示例中，消除性在一些问题中得到了满足，而在之前一些与其等同的问题中，却遭到了违背。因此，消除性的经验效度取决于问题的框架。

目前的框架概念起源于L.萨维奇和雷法对阿莱斯问题的分析，他们重新定义了这些案例，试图使消除性的应用更加引人注目（Savage, 1954; Raiffa, 1968）。萨维奇和雷法是对的：当应用程序足够透明时，缺乏经验的被试确实会遵守消除性定理。[\(13\)](#)但是，相同选项的不同版本（问题10和问题11以及问题12的案例2和案例3）的对比偏好表明当它的应用不透明时，人们不会遵循相同的定理。相反，他们会将（非线性）决策权重应用于所述的概率。因此，消除性与占优性的情况类似：两种规则都是直观且引人注目的抽象选择原则，在透明问题中始终被遵守，而在非透明问题中经常不被遵循。试图通过放弃消除性来合理化阿莱斯案例中的偏好面临着一个主要困难：他们没有区分遵守消除性的透明框架与不遵循消除性的非透明框架。

讨论

在前面的章节中，我们挑战了期望效用理论的主要原则的描述效用，并概述了风险选择的另一种说法。在本节中，我们将讨论替代定理，并为反对规范和描述性分析的调和进行辩解。经济学家对我们的分

析和结论提出了一些反对意见。

描述性和规范性考虑因素

在过去10年中，研究人员已经开发了许多风险选择的替代模型，用于解释观察到的对期望效用理论的背离现象。这些模型可被分为以下4类。第一，非线性函数是通过完全舍弃消除性条件而获得的（Allais, 1953、1979；Machina, 1982）。这些模型没有进行定理性效用的（基数）测量，但是它们对效用函数施加了各种限制（即可区分性）。第二，预期指数模型用较弱的替代定理代替消除性，并用两个线性函数的比率表示前景的值（Chew & MacCrimmon, 1979；Weber, 1982；Chew, 1983；Fishburn, 1983）。第三，具有非可加性概率的双线性模型假设了消除性（或替代性的）各种限制版本，并构建了双线性表达式，其中结果的效用通过非可加性概率测量或概率尺度的一些非线性变换进行加权（Kahneman & Tversky, 1979；Quiggin, 1982；Schmeidler, 1984；Segal, 1984；Yaari, 1984；Luce & Narens, 1985）。第四，通过二元效用函数表示偏好的非传递模型。菲什伯恩对这些模型进行了定理化，而贝尔、卢默斯和萨格登将其解释为预期后悔（expected regret；Fishbum, 1982, 1984；Bell, 1982；Loomes & Sugden, 1982）。有关进一步的理论发展，请参阅菲什伯恩的研究（Fishbum, 1985）。

表5-1总结了模型和数据之间的关系。子栏列出了期望效用理论的4个主要原则。第1列是实证中违背这些原则的主要现象，并列举了一些具有代表性的偏好。第2列是上面讨论的与观察到的违背现象相一致的模型子集。

表5-1 违背理论的实证现象和解释性模型汇总

原则	违背理论的实证现象	解释性模型
消除性	确定性效应（Allais, 1953, 1979; kahneman & Tversky, 1979）[问题 9、问题 10 和问题 12（案例 1 和 2）]	所有模型
传递性	词典半序（Tversky, 1969） 偏好逆转（Slovic & Lichtenstein, 1983）	双变量模型
占优性	对比风险态度（问题 2）；次级决策权重（问题 8）	前景理论
不变性	框架效果（问题 1、3、4、5、6、7、8、10、11 和 12）	前景理论

表5-1的结论可概括如下。首先，所有上述模型（以及其他一些模型）与确定性效应产生的违背消除性的行为一致。[\(14\)](#)因此，阿莱斯的“悖论”不能用于比较或评估竞争的非前景模型。其次，需要二元（非传递性）模型来解释观察到的敏感度较低的情况。再次，只有前景理论可以解释观察到的对（随机）占优性和不变性的违背。虽然有些模型允许一些有限的对不变性的违背，但它们没有考虑到本文中描述的框架效应应的范围（Loomes & Sugden, 1982; Luce & Narens, 1985）。

由于框架效应和相关的不变性失调无处不在，没有足够的描述性理论可以忽略这些现象。此外，由于不变性（或延伸性）在规范上是必不可少的，所以没有足够的规范性理论允许这种违背。因此，构建一个描述性和规范性上都可接受的理论，似乎是不可实现的梦想（Tversky & Kahneman, 1983）。

前景理论与上述其他模型的不同之处在于其不加掩饰的描述性，以及没有提出规范性主张。无论是否可以被合理化，它的主要目的都在于解释偏好。M.马奇纳（M. Machina）声称前景理论“作为风险行为的描述模型是不可接受的”，因为它意味着对随机的占优地位的违背（Machina, 1982）。但是，由于实际上已经观察到对理论预测的占优地位的违背（见问题2和问题8），因此马奇纳的反对意见似乎是无效的。

本文的主要发现是，理性选择定理通常在信息透明的情况下成立，但在非透明的情况下不成立。例如，当随机占优关系透明时（如问题2和问题7的聚合版本），实际上每个人都倾向于选择处于优势地位的前景。然而，当这些问题陷入困境时，占优性不再透明（如问题2和问题8的分离版本），大多数被试就会违反占优性，正如预测的那样。这些结果与所有暗示随机占优性的理论，以及那些预测人们在透明和非透明背景下，会做出相同选择的其他理论相矛盾（Machina, 1982）。同样的结论也适用于消除性，如虚假性讨论。看起来消除性和占优性都具有规范性吸引力，尽管两者在描述上都是无效的。

目前的结果和分析，特别是针对透明度的作用和框架重要性的分析，与赫伯特·西蒙（Herbert Simon）最初提出的有限理性概念是一致的（Simon, 1955, 1978; Nelson & Winter, 1982）。实际上，前景理论试图阐明一些限制选择合理性的感知和判断原则。

心理因素（如框架）的引入既丰富了对于选择的分析，又使其复杂

化了。因为决策的框架取决于用于表达的语言、选择的背景以及展示的性质，我们对过程的处理必然是非正式的和不完整的。我们已经确定了几个常见的框架规则，并且已经证明了它们对选择的影响，但还没有得出框架的正式理论。此外，本分析未考虑所有观察到的传递性和不变性的实效。虽然一些不和谐（Tversky, 1969）可以通过抛弃框架阶段中的小差异来解释，而其他不和谐（Raiffa, 1968）是由透明和非透明比较的组合产生的，但是存在的循环偏好和背景效应（Slovic, Fischhoff & Lichtenstein, 1982; Slovic & Lichtenstein, 1983）需要额外的解释机制，如多个参考点和可变权重）。在对选择进行充分考虑时，不能忽视框架和背景的这些影响，即使它们在规范上令人反感，且在数学上难以处理。

支撑性假设

理性假设在经济学中占优势地位。它具有所有方法论的优点：一个不言而喻的真理、一个合理的理想化处理、一个重言式逻辑和一个零假设。其中的每一种解释要么提出了不容置疑的理性行为假设，要么将举证责任直接放在对信念和选择的任何替代性分析上。理性模型的优势是复杂的，因为没有其他判断和决策理论可以在范围、权力和简单性方面与之匹配。

此外，合理性假设受到了一系列强大防御措施的保护，其形式包括支持任何观察到的违反模型的重要性假设。特别是，人们通常认为，严重违反标准模型的行为仅限于无关紧要的选择性问题，这些问题能通过学习迅速被消除，或由于市场力量具有纠正功能而与经济学无关。实际上，激励措施有时能提高决策质量，经验丰富的决策者往往比新手做得更好，套利和竞争的力量可以抵消错误和错觉的某些影响。这些因素能否确保任何特定情况下的理性选择是一个经验性问题，需要通过观察而不是假设来解决。

观察到的理性模型的失败可归因于思考的成本，因此应通过适当的激励来消除（Smith, 1985）。实验结果几乎没有给出支持这个观点的论据。经济学和心理学的研究表明，即使存在重大的货币收益，在针对假设性问题的回答中普遍存在的错误仍然存在。尤其是概率推理的基本错误（Grether, 1980; Tversky & Kahneman, 1983）、选择的严重不一致性（Grether & Plott, 1979; Slovic & Lichtenstein, 1983）以及违反非透明问题中的随机占优性（上文中的问题2）几乎都没有通过激励得以

减少。关于高风险并不总能改善决策的证据并不仅限于实验室研究。在涉及高风险和认真审议的实际决策中，记录了重大的判断性和选择性错误。例如，小企业的高失败率不易与理性前景和风险规避的假设保持一致。

激励不是通过魔术来发挥作用的：它们是通过集中注意力和延长审议来运行的。因此，他们更有可能防止由于误解或错误的直觉引起的错误，这种错误来自注意力和努力不足。视觉上的错觉的案例很有启发性。没有明显的可以仅仅引入激励（没有额外的机会进行测量）就能减少图5-3中观察到的错觉的机制，并且即使没有激励，当图5-4中的呈现方式发生改变时，错觉就会消失。激励的纠正力量取决于特定错误的性质，不能将其视为理所当然。

决策的理性假设常常会受到人们会学着做出正确的决策这一观念的支持，有时还会受到演化论证的支持，即非理性决策者会被理性决策者所取代。毫无疑问，学习和选择确实会发生，而且往往能提高效率。然而，就激励而言，不涉及任何魔法。有效的学习只有在某些条件下才能进行：它要求对情境条件与适当反应之间的关系进行准确和即时的反馈。管理者、企业家和政治家做出的决定往往缺乏必要的反馈，原因如下：第一，结果通常是延迟性的，而且不容易归因于某一特定行动；第二，环境的变化会降低反馈的可靠性，特别是在涉及低概率的结果时；第三，如果已经做出另一项决定，往往没有关于结果将会怎样的信息；第四，最重要的决定是独特的，因此所能提供的学习机会很少

（Einhorn & Hogarth, 1978）。组织学习的条件也不会更好。对于个人和组织来说，学习肯定会发生，但必须通过证明有效学习的条件已得到满足来支持通过经验可以消除特定错误的说法。

有时人们认为，由于市场具有纠正功能，因此个人决策中非理性是无关紧要的（Knez, Smith & Williams, 1985）。由于竞争力和套利者的行为，经济主体往往可以免受自身非理性偏好的影响，但在某些情况下，这种机制会失效。D. B.豪施（D. B. Hausch）、W. T.津巴（W. T. Ziemba）和M. E.鲁本斯坦（M. E. Rubenstein）记录了一个有启发性的案例：赛马场上的中奖投注市场是有效的，但是场地和表演中的投注市场则不然（Hausch, Ziemba & Rubenstein, 1981）。投注者会低估最受欢迎的一方最终会排在第二位或第三位的概率，并且这种影响足以维持具有正面前景价值的逆向投注策略。尽管有很高的激励机制、赛道参与者无可置疑的奉献精神和专业知识和专业知识，以及有明显的学习和套利机会，但

仍存在低效现象。

约翰·霍尔蒂万格（John Haltiwanger）和M.瓦尔德曼（M. Waldman）以及T.拉塞尔（T. Russell）和塞勒分析了市场不太可能纠正许多人常见的错误的情况（Haltiwanger & Waldman, 1985; Russell & Thaler, 1985）。此外，诺贝尔经济学奖得主乔治·阿克洛夫（George Akerlof）和美国现任财政部长珍妮特·耶伦（Janet Yellen）提出了他们近乎理性的理论，其中在对经济变化（如惯性或货币错觉）的反应中，一些普遍存在的错误对个人影响不大（从而消除了学习的可能性），也没有提供套利的机会，但其经济效应却很大（Akerlof & Yellen, 1985）。如果没有确凿的证据，就不能相信市场能够纠正个人的非理性行为的影响，而提出这种主张的人应承担详细说明合理的纠正机制的责任。

本文的主题是规范性和描述性的选择分析应被视为独立的部分。该结论提出了一个研究议程。为了使理性模式保持其惯有的描述性作用，必须验证相关的支撑性假设。在这些假设失败的情况下，追踪描述性分析（如损失厌恶、虚假性或货币错觉的影响）对公共政策、战略决策和宏观经济现象的影响都是有益的（Arrow, 1982; Akerlof & Yellen, 1985）。

本章从新的角度回顾了我们在风险决策方面的成果，主要在第一部分和最后一部分进行了讨论。大多数实证分析已经发表在了之前的出版物中。问题3、问题4、问题7、问题8和问题12是首次发表。

参考文献

Akerlof, G. A., & Yellen, J. (1985). Can small deviations from rationality make significant differences to economic equilibria? *American Economic Review*, 75, 708–720.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'Ecole Américaine. *Econometrica*, 21, 503–546.

Allais, M. (1979). The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the American School. In M. Allais & O. Hagen (Eds.), *Expected utility*

hypotheses and the Allais paradox (pp. 27–145). Dordrecht, the Netherlands: Reidel.

Arrow, K. J. (1982). Risk perception in psychology and economics. *Economic Inquiry*, 20, 1–9.

Bazerman, M. H. (1983). Negotiator judgment. *American Behavioral Scientist*, 27, 211–228.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research*, 30, 961–981.

Bishop, R. C., & Heberlein, T. A. (1979). Measuring values of extra-market goods: Are indirect measures biased? *American Journal of Agricultural Economics*, 61, 926–930.

Chew, S. H. (1983). A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. *Econometrica*, 51, 1065–1092.

Chew, S. H., & MacCrimmon, K. 1979. Alpha utility theory, lottery composition, and the Allais paradox. Working Paper no. 686. Vancouver, BC: University of British Columbia.

Clark, H. H., & Clark, E. V. (1977). *Psychology and language*. New York, NY: Harcourt Brace Jovanovich.

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review*, 85, 395–416.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75, 643–669.

Eraker, S. E., & Sox, H. C. (1981). Assessment of patients' preferences for therapeutic outcomes. *Medical Decision Making*, 1, 29–39.

Fischhoff, B. (1983). Predicting frames. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 103–116.

Fishburn, P. C. (1982). Nontransitive measurable utility. *Journal of Mathematical Psychology*, 26, 31–67.

Fishburn, P. C. (1983). Transitive measurable utility. *Journal of Economic Theory*, 31, 293–317.

Fishburn, P. C. (1984). SSB utility theory and decision making under uncertainty. *Mathematical Social Sciences*, 8, 253–285.

Fishburn, P. C. (1985). Uncertainty aversion and separated effects in decision making under uncertainty. Working paper. Murray Hill, NJ: AT&T Bell Labs.

Fishburn, P. C., & Kochenberger, G. A. (1979). Two-piece von Neumann–Morgenstern utility functions. *Decision Sciences*, 10, 503–518.

Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *Quarterly Journal of Economics*, 95, 537–557.

Grether, D. M., & Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *American Economic Review*, 69, 623–638.

Hagen, O. (1979). Towards a positive theory of preferences under risk. In M. Allais & O. Hagen (Eds.), *Expected utility hypotheses and the Allais paradox* (pp. 271–302). Dordrecht, the Netherlands: Reidel.

Haltiwanger, J., & Waldman, M. (1985). Rational expectations and the limits of rationality: An analysis of heterogeneity. *American Economic Review*, 75, 326–340.

Hammond, P. (1985). Consequential behavior in decision trees and expected utility. Institute for Mathematical Studies in the Social Sciences Working Paper no. 112. Stanford University.

Hansson, B. (1975). The appropriateness of the expected utility model. *Erkenntnis*, 9, 175–193.

Hausch, D. B., Ziemba, W. T., & Rubenstein, M. E. (1981). Efficiency of the market for racetrack betting. *Management Science*, 27, 1435–1452.

Hershey, J. C., & Schoemaker, P. J. H. (1980). Risk taking and problem context in the domain of losses: An expected utility analysis. *Journal of Risk and Insurance*, 47, 111–132.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986a). Fairness and the assumptions of economics. *Journal of Business*, 59, S285–S300.

Kahneman, D., Knetsch, J. L., & Thaler, R. (1986b). Fairness as a constraint on profit seeking: Entitlements in the market. *American Economic Review*, 76, 728–741.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.

Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American*, 246, 160–173.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39, 341–350.

Knetsch, J. L., & Sinden, J. A. (1984). Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. *Quarterly Journal of Economics*, 99, 507–521.

Knez, P., Smith, V. L., & Williams, A. W. (1985). Individual rationality, market rationality and value estimation. *American Economic Review*, 75, 397–402.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal (Oxford)*, 92, 805–824.

Luce, R. D., & Krantz, D. H. (1971). Conditional expected utility. *Econometrica*, 39, 253–271.

Luce, R. D., & Narens, L. (1985). Classification of concatenation measurement structures according to scale type. *Journal of Mathematical Psychology*, 29, 1–72.

Machina, M. J. (1982). “Expected utility” analysis without the independence axiom. *Econometrica*, 50, 277–323.

March, J. G. (1978). Bounded rationality, ambiguity, and the engineering of choice. *Bell Journal of Economics*, 9, 587–608.

Markowitz, H. (1952). The utility of wealth. *Journal of Political Economy*, 60, 151–158.

McNeil, B. J., Pauker, S. G., Sox, H. C., Jr., & Tversky, A. (1982). On the elicitation of preferences for alternative therapies. *New England Journal of Medicine*, 306, 1259–1262.

Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Cambridge, MA: Harvard University Press.

Payne, J. W., Laughhunn, D. J., & Crum, R. (1980). Translation of gambles and aspiration level effects in risky choice behavior. *Management Science*, 26, 1039–1060.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3, 323–343.

Raiffa, H. (1968). *Decision analysis: Introductory lectures on choices under uncertainty*. Reading, MA: Addison-Wesley.

Russell, T., & Thaler, R. (1985). The relevance of quasi-rationality in competitive markets. *American Economic Review*, 75, 1071–1082.

Savage, L. J. (1954). *The foundations of statistics*. New York, NY: Wiley.

Schelling, T. C. (1981). Economic reasoning and the ethics of policy. *Public Interest*, 63, 37–61.

Schmeidler, D. (1984). *Subjective probability and expected utility without additivity. Preprint Series no. 84*. Minneapolis, MN: University of Minnesota, Institute for Mathematics and Its Applications.

Schumpeter, J. A. (1954). *History of economic analysis*. New York, NY: Oxford University Press.

Segal, U. 1984. Nonlinear decision weights with the independence axiom. Working Paper in Economics no. 353. Los Angeles, CA: University of California, Los Angeles.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69, 99–118.

Simon, H. A. (1978). Rationality as process and as product of thought. *American Economic Review*, 68, 1–16.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1982). Response mode, framing, and information processing effects in risk assessment. In R. M. Hogarth (Ed.), *New directions for methodology of social and behavioral science: Question framing and response consistency* (pp. 21–36). San Francisco, CA: Jossey-Bass.

Slovic, P., & Lichtenstein, S. (1983). Preference reversals: A broader perspective. *American Economic Review*, 73, 596–605.

Slovic, P., & Tversky, A. (1974). Who accepts Savage's axiom? *Behavioral Science*, 19, 368–373.

Smith, V. L. (1985). Experimental economics [Reply] . *American Economic Review*, 75, 265–272.

Stigler, G. J., & Kindahl, J. K. (1970). *The Behavior of industrial prices*. New York, NY: National Bureau of Economic Research.

Thaler, R. H. (1980). Towards a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1, 39–60.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*,

76, 105–110.

Tversky, A. (1977). On the elicitation of preferences: Descriptive and prescriptive considerations. In D. E. Bell, R. L. Keeney, & H. Raiffa (Eds.), *Conflicting objectives in decisions* (pp. 209–222). New York, NY: Wiley.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453–458.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90, 293–315.

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Weber, R. J. (1982). The Allais paradox, Dutch auctions, and alpha-utility theory. Working paper. Northwestern University.

Yaari, M. E. (1984). *Risk aversion without decreasing marginal utility*. *Report series in theoretical economics*. London, England: London School of Economics.

第6章 个体患者和群体患者医疗决策之间的差异

唐纳德·A.雷德梅尔
阿莫斯·特沃斯基

在很多情况下，卫生政策和医疗实践之间的关系很紧张。例如，尽管有广泛的共享信息，但实践模式的区域性差异仍然存在（Chassin, Brook, Park et al., 1986; Wennberg, 1986; Iglehart, 1984），患者每天接受的护理都与公认的指导方针有很大偏差（Woo, Woo, Cook, Weisberg & Goldman, 1985; Kosecoff, Kanouse, Rogers, McCloskey, Winslow & Brook, 1987; Eddy, 1982; Lomas, Anderson, Domnick-Pierre, Vayda, Enkin & Hannah, 1989），并且给予了选定个体不成比例的护理（Woolley, 1984; Bunker, 1976; Levinsky, 1984）。这些观察结果表明，临床领域中侧重于个体患者的决策可能会与一般的医疗政策不一致，一般性医疗政策应基于更广泛的考虑。我们对这种差异进行了调查研究。

想象一名患者向医生提出一个特定的问题。通常情况下，医生会把每名患者作为一个独特的病例来对待，并选择对该患者来说似乎是最好的治疗方法。然而，随着时间的推移，医生可能会遇到许多类似的患者。当一个病例被认为是独一无二的，而不是一组可比较的病例中的一个时，医生会做出不同的判断吗？有证据表明，人们在面对单一而非重复的情况时，会在与金钱相关的冒险活动中做出不同的选择

（Samuelson, 1963; Keren & Wagenaar, 1987; Montgomery & Adelbratt, 1982）。此外，对经济和医疗决策的研究表明，从不同的角度看待问题可以改变其属性的相对权重，导致不同的选择（McNeil, Pauker, Sox & Tversky, 1982; Eraker & Sox, 1981; Tversky & Kahneman, 1986）。

我们假设，当医生将患者视为个体时，他们更重视患者的个人问题；当医生将患者视为群体的一部分时，他们更重视一般的效度标准。更具体地说，我们认为，在将患者视为个体而不是群体的一员时，医生更有可能做出以下行为：建议做花费不多且可能有帮助的额外的检查，对患者进行直接检查而不是通过电话跟进，避免讨论器官捐献等复杂性

问题，以及推荐一种成功率很高但有可能出现不良后果的治疗方法。在这项研究中，我们希望通过对这些问题进行探讨来解决下面这个问题：与考虑一组相似的患者相比，医生在评估单个患者时会做出不同的判断吗？我们的数据表明确实如此，接受过健康服务研究培训的医生能够认识到这种差异，而非专业人士也可以做出这种区分。

方法

在第一个实验中，我们邀请执业医师参与了一项医学决策研究。我们使用的问卷包含对患者管理问题中的临床情景的描述，高水平的医生可能不会认可问卷中列出的这些情景。我们要求每位医生选择最合适的治疗方法。

我们在两个版本的问卷中都提出了这些问题，每个版本都从不同的角度对其进行了讨论。个体版本涉及一名患者的治疗；总体版本涉及一组同类可比较的患者的治疗。在其他方面，这两个版本包含的信息相同。例如，个体版本的情景如下所述。

文献几乎没有提供任何关于使用电话作为医疗工具的信息。例如，海格是一名年轻女性，家庭医生很熟悉她的身体状况，她没有任何严重疾病。有一次她通过电话联系了家庭医生，因为她发烧了 5 天，且没有任何其他的局部症状。此时家庭医生做出的初步诊断是病毒感染，并根据症状给出了建议性措施，并告诉她要“保持联系”。大约 36 小时后，她打电话回来，说感觉和之前差不多：没有好转，没有恶化，也没有出现新的症状。这时医生必须做出选择，是继续通过电话跟踪一段时间，还是告诉她现在就来检查。你认为医生应该选择哪种方式？

在这种情况下的总体版本与之类似，不同点在于我们用表示一组患者的术语替换了对个体患者的所有引用。

文献几乎没有提供任何关于使用电话作为医疗工具的信息。例如，设想有这样一些年轻女性，她们的家庭医生熟知她们的身体状况，并且她们没有任何严重疾病。她们可能会因为发烧5天但没有任何其他的局部症状而通过电话联系各自的家庭医生。通常情况下，医生会将其初步诊断为病毒感染，给出相应的症状诊断措施，并告诉她们“保持联系”。假设在大约36小时后，她们打电话回来，说感觉和之前差不多：没有好转，也没有恶化，并且没有出现新的症状。这时医生必须做出选择，是继续通过电话跟踪她们一段时间，还是告诉她们现在就来检查。你会建议选择哪种方式？

4组医生参与了这项研究：斯坦福大学医院医学系的住院医师、在区域健康维护组织工作的全职医生、斯坦福大学内科学术医生，以及在县级医疗中心工作的全职医生。在每组中，我们随机分配医生来填写个体版本或总体版本的问卷。之后，我们使用曼-惠特尼检验来对比他们对两个版本的反应（Moses, 1986）。

在第二个实验中，我们展示了与第一个实验类似的情景，并要求被试直接比较这两个视角。在这份调查问卷中，我们调查了一组内科医生、精神科医生和儿科医生，他们在临床医学和卫生服务研究方面都接受过高级培训。对于每种情景，被试需要表明他们是否认为医生更可能从个体患者的角度或一般性政策的角度推荐特定的治疗措施。例如，我们提出了以下情况：“一名骑摩托车的25岁男子正在接受常规医疗检查。从哪个角度来看，医生更有可能与其讨论器官捐献？”

在第三个实验中，我们要求斯坦福大学的本科生考虑一个假设性医学案例，这个案例不需要技术性知识就可以理解。与第一个实验一样，一半的学生看到的是个体版本，另一半看到的是总体版本。这3个实验的被试都收到了调查问卷，他们可以在空闲时完成，并匿名提交。

结果

实验1

在第一个实验中，59名住院医师、94名大学附属医院医生、75名健

康维护组织的医生，以及128名县级医院的医生提交了完整的问卷。问卷总体回复率为78%。正如预料的那样，之前收到的不同版本问卷的两个群体在年龄、性别、经历和反应率方面很相似。下面我们来讨论之前提出的4个问题。

血液检测。为了探究第一个问题，我们请医生考虑一个大学生表现出疲劳、失眠和注意力不集中的情况。除了通常的评估外，我们还描述了一项额外的血液检测，通过这个血液检测可能会检测出罕见的、可治疗的疾病，但这需要花费20美元，且学生必须自付。当面对一个患者的个体版本时，医生会比面对一组患者的总体版本时更加频繁地进行血液检测（30%对17%； $p<0.005$ ）。4组不同的医生之间的差异性比较明显，分别为住院医师（26%对4%； $p<0.05$ ）、健康维护组织的医生（28%对7%； $p<0.10$ ）、学术医生（40%对19%； $p<0.01$ ），以及县级医疗中心的医生（43%对22%； $p<0.05$ ）。

电话医学。为了探究第二个问题，我们让医生考虑一个健康的年轻女性的情况，她因为持续的轻微发烧而打电话给家庭医生。医生在面对总体版本的情况下会比面对个人版本时更常建议通过电话跟进，而不是要求患者接受检查（13%对9%； $p<0.005$ ）。学术医生和县级医院医生间的差异性很明显，学术医生（15%对6%； $p<0.01$ ），县级医院医生（12%对2%； $p<0.05$ ），但在健康维护组织医生中并不明显（14%对24%； p 不显著）。住院医师没有参与该情景下的实验。

实验2

与实验1中只评估一个问题版本的医生相比，实验2中的医生直接比较了总体和个体两个角度。本次实验共收回89份已完成的问卷，回复率为77%。该结果证实了实验1的结果。对于疲劳的大学生，81%的被试（ $p<0.005$ ）认为如果从个体角度而不是总体角度考虑，他们将更频繁地推荐额外的检查。对于发烧的年轻女性，87%的被试（ $p<0.005$ ）认为从群体角度出发他们多数情况下会选择通过电话跟进。

器官捐献。为了探究第三个问题，我们还向卫生服务研究人员展示了一名健康的摩托车手的情景，他因为一个小的医疗问题而接受检查。当被问及讨论器官捐献问题时，93%的被试（ $p<0.005$ ）认为从总体角度来看，多数情况下他们会推荐器官捐献。

不良反应。为了探讨第四个问题，我们向卫生服务研究人员提供了一个患有血液病的女性的情景。我们描述了一种可以添加到她的治疗中的药物，这种药物有可能会延长她的寿命，但也有可能会使病情恶化。这种药物有85%的概率让她多活两年，有15%的概率让她少活两年。在这种情况下，59%的被试（ $p<0.10$ ）认为从个体角度来看，他们在多数情况下会推荐使用这种药物。

实验3

这个实验检验了不同视角之间的差异性，在普通人的判断中是否同样明显。我们共向327名学生展示了不良反应的情景，之所以选择这一方案，是因为即便被试不具备专业的医学知识也不影响判断。与第一个实验一样，每名学生都收到了个体版本或总体版本。根据我们之前的研究结果，面对个体版本的医生会比面对总体版本的医生更推荐使用药物（62%对42%； $p<0.005$ ）。

讨论

我们的研究结果表明，在评估个体患者时，医生做出的决定与考虑一组同类患者时不同（实验1）。这种差异性被认为是一种职业规范（实验2），它也会出现在非专业人士的判断中（实验3）。我们探讨了能突出这种差异性的4个问题。从个体角度来看，与总体角度相比，医生更有可能要求患者做额外的检查，花时间直接评估患者，避免提出一些令人不安的问题，并推荐一种成功率很高但可能会出现不良后果的治疗方法。

这些实验中所显示的总体角度和个体角度之间的差异不能被归因于医疗信息或经济激励的差异，因此，我们很难从规范的角度对其加以解释（Sox, Blatt, Higgins & Marton, 1988; Raiffa, 1968）。我们的研究结果与以下观点一致，即医生在将患者视为个体时更重视患者的个人问题，而在将患者视为群体中的一员时更重视一般的效度标准。例如，被试对我们给定的不良后果情景的反应表明，当只考虑一个案例时，医生对小概率事件的重视程度较低。赋予一个问题不同方面相应权重时，这些差异可能有助于解释为什么反映群体角度的一般性原则在临床实践中并不总是被遵循，这种临床实践是在逐个案例分析的基础上进行的。因此，即使卫生决策者和医务人员都接受相关事实，总体角度和个人角度之间的差异也可能让两者之间的关系紧张。医学决策的一些特点

可能会放大不同角度之间的差异。谢林讨论了统计生活和现实生活之间的区别，强调社会对现实中的人的生活来说具有更高的价值

（Schelling, 1968）。经济学家V. R.富克斯（V. R. Fuchs）在医患关系中提出了“技术要求”，这反映了医生在治疗个体患者时，更愿意做与他们所接受过的训练相关的所有事情（Fuchs, 1974）。R.埃文斯（R. Evans）谈到了医生作为患者的完美代理人和社会保护者之间的矛盾（Evans, 1983）。当然，财政奖励也可能导致政策与实践之间的紧张关系（Bock, 1988; Scovern, 1988）。

虽然总体角度和个体角度之间的差异需要解决，但我们并不建议放弃任何一种角度。个体角度强调患者的特殊性，更符合医患关系的个性化特征。但从总体角度来看，随着时间的推移，医生将治疗很多类似的患者。医生和决策者可能希望从两个角度来研究问题，以确保治疗决策适用于一名或多名患者。了解这两种角度可以提高临床判断的质量，同时也可以丰富卫生政策。

致谢

在此，我们要感谢塔米·滕（Tammy Teng）、琼·埃斯普林（Joan Esplin）、马库斯·克鲁伯（Marcus Krupp）、爱德华·哈里斯（Edward Harris）、艾略特·沃尔夫（Elliott Wolfe）和帕特里克·卡恩斯（Patrick Kearns）帮助设计本章采用的问卷。感谢霍尔斯特德·霍尔曼（Halsted Holman）、黛安娜·达顿（Dianna Dutton）、艾伦·加伯（Alan Garber）、罗伯特·瓦赫特（Robert Wachter）和米切尔·威尔逊（Mitchel Wilson）在原稿准备中给予我们的帮助；感谢我们的受访者。

参考文献

Chassin, M. R., Brook, R. H., Park, R. E., et al. (1986). Variations in the use of medical and surgical services by the Medicare population. *New England Journal of Medicine*, 314, 285–290.

Wennberg, J. (1986). Which rate is right? *New England Journal of Medicine*, 314, 310–311.

Iglehart, J. K. (Ed.). (1984). Variations in medical practice. *Health Affairs (Project Hope)*, 3(2), 6–148.

Woo, B., Woo, B., Cook, E. F., Weisberg, M., & Goldman, L. (1985). Screening procedures in the asymptomatic adult: Comparison of physicians' recommendations, patients' desires, published guidelines, and actual practice. *Journal of the American Medical Association*, 254, 1480–1484.

Kosecoff, J., Kanouse, D. E., Rogers, W. H., McCloskey, L., Winslow, C. M., & Brook, R. H. (1987). Effects of the National Institutes of Health Consensus Development Program on physician practice. *Journal of the American Medical Association*, 258, 2708–2713.

Eddy, D. M. (1982). Clinical policies and the quality of clinical practice. *New England Journal of Medicine*, 307, 343–347.

Lomas, J., Anderson, G. M., Domnick-Pierre, K., Vayda, E., Enkin, M. W., & Hannah, W. J. (1989). Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. *New England Journal of Medicine*, 321, 1306–1311.

Woolley, F. R. (1984). Ethical issues in the implantation of the total artificial heart. *New England Journal of Medicine*, 310, 292–296.

Bunker, J. P. (1976). When the medical interests of society are in conflict with those of the individual, who wins? *Pharos*, 39(1), 64–66.

Levinsky, N. G. (1984). The doctor's master. *New England Journal of Medicine*, 311, 1573–1575.

Samuelson, P. A. (1963). Risk and uncertainty: A fallacy of large numbers. *Scientia*, 98, 108–113.

Keren, G., & Wagenaar, W. A. (1987). Violation of expected utility theory in unique and repeated gambles. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 382–391.

Montgomery, H., & Adelbratt, T. (1982). Gambling decisions and information about expected value. *Organizational Behavior and Human Performance*, 29, 39–57.

McNeil, B. J., Pauker, S. G., Sox, H. C., Jr., & Tversky, A. (1982). On the elicitation of preferences for alternative therapies. *New England Journal of Medicine*, 306, 1259–1262.

Eraker, S. A., & Sox, H. C., Jr. (1981). Assessment of patients' preferences for therapeutic outcomes. *Medical Decision Making*, 1, 29–39.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, 59, S251–S278.

Moses, L. E. (1986). *Think and explain with statistics*. Reading, MA: Addison-Wesley.

Sox, H. C., Jr., Blatt, M. A., Higgins, M. C., & Marton, K. I. (1988). *Medical decision making*. Boston, MA: Butterworths.

Raiffa, H. (1968). *Decision analysis*. Reading, MA: Addison-Wesley.

Schelling, T. C. (1968). The life you save may be your own. In S. B. Chase, Jr. (Ed.), *Problems in public expenditure analysis* (pp. 127–176). Washington, DC: Brookings Institution.

Fuchs, V. R. (1974). *Who shall live? Health, economics, and social choice*. New York, NY: Basic Books.

Evans, R. W. (1983). Health care technology and the inevitability of resource allocation and rationing decisions. *Journal of the American Medical Association*, 249, 2208–2219.

Bock, R. S. (1988). The pressure to keep prices high at a walk-in clinic: A personal experience. *New England Journal of Medicine*, 319, 785–787.

Scovern, H. (1988). Hired help: A physician's experiences in a for-profit staff-model HMO. *New England Journal of Medicine*, 319, 787–790.

第7章 不确定状况下的思考：非结果主义推理和选择

埃尔德·沙菲尔
阿莫斯·特沃斯基

很多日常思考和决策都涉及客观世界状态及我们的主观情绪和欲望方面存在的不确定性。我们可能无法确定未来的经济状况、即将到来的考试后的心情以及假期是否想去夏威夷度假。当然，不同的状态往往会导致不同的决定。如果考得好，我们可能会觉得应该休息一下，想去夏威夷；如果考得不好，我们可能宁愿待在家里。当在不确定的情况下做出决定时，我们需要考虑现实中可能的状态及其对我们的愿望和行动的潜在影响。不确定性情况可以被认为是可能状态的析取：要么获得一种状态，要么获得另一种状态。例如，一位不确定考试表现如何的学生会面临两种结果：通过考试或考试不及格。在决定是否计划去夏威夷度假时，该学生需要考虑如果通过了考试自己是否想去夏威夷，以及如果未通过考试自己是否想去夏威夷，具体情况如图7-1所示。按照惯例，决策节点用正方形表示；机会节点用圆形表示。

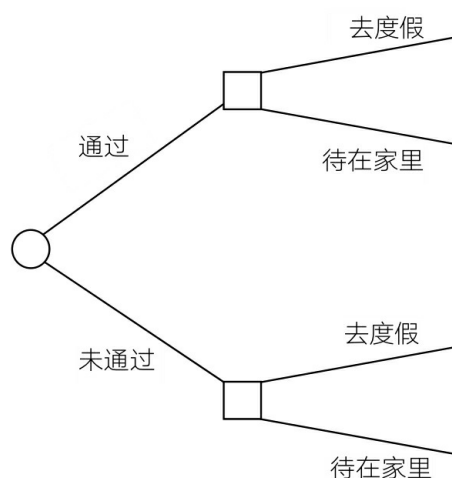


图7-1 夏威夷度假问题的树形图

大多数不确定状况下的决策概念，不管是规范性的还是描述性的，都是结果主义的，因为决策是由对潜在结果及其感知的可能性的评估决

定的。根据这种观点，学生去夏威夷度假的决定取决于他在考试通过和未通过的情况下待在家和去度假的主观价值，以及他认为考试通过和未通过的主观概率。(15)基于对预期结果的结果主义评估，我们预测选择满足不确定性决策的基本公理，即萨维奇的确定性原则（Savage, 1954）。确定性原则认为，如果我们在给定的任何可能状态下都更喜欢x而不是y，那么即使不知道现实的确切状态，我们也应该更喜欢x而不是y。图7-1所示的情况表明，如果该学生在通过和未通过考试时都更倾向于去度假，那么即使考试结果未知，他也应该更倾向于去度假。确定性原则是结果主义观点的一个重要内涵，它抓住了由预期结果决定决策的基本直觉。它是期望效用理论的基石，并且在其他对合理性标准不那么严格的模型中也成立。

但是，如果人们并不总是以结果主义的方式进行选择，那么有时可能会违反确定性原则。例如，我们在其他地方已经证明，很多不管考试通过或未通过都会选择去夏威夷度假的人，当考试结果未知时，在析取情况下他们会决定推迟度假（Tversky & Shafir, 1992）。如果通过考试，假期大概会被认为是一个成功的学期后的庆祝时间；如果考试不及格，度假则是安慰和治愈的方式。我们认为，由于不知道考试结果，决策者缺乏去度假的明确理由，因此在决定去度假之前，他们可能宁愿等待结果的揭晓，而这与确定性原则相反。

对于另一个非结果论推理的例子，假设你已经同意通过掷硬币来进行投注，那么你有相同的机会赢200美元或者输100美元。假设已经掷了硬币，但不知道输赢，你还想再一次投注吗？或者，如果知道自己在第一次投注中输了100美元，你会接受第二次投注吗？最后，如果你已经知道自己在第一场投注中赢了200美元，你会进行第二次吗？我们已经证明，与确定性原则相反，大多数被试不管第一次投注是赢是输，他们都接受了第二次投注；但在第一次结果未知时，大多数被试拒绝进行第二次投注（Tversky & Shafir, 1992）。这种赢时接受、输时拒绝、结果未知时拒绝的模式是被试表现出的最常见的偏好模式。我们已经发现，在获得收益后人们有充分的理由接受第二次投注，即“我已经赢了，无论发生什么事情我都不会有损失”；对于在失败后接受第二次投注，他们还有一个尽管不同但令人信服的理由，即“我失败了，这将是赢的机会”。但是，当第一次投注的结果未知时，人们不知道自己是领先且不会有损失，还是落后且需要挽回损失。我们认为，在这种情况下，他们可能没有明确的理由接受附加的投注，而就其本身而言，这种投注并不特别有吸引力。我们把上面的偏好模式称为析取效应（disjunction

effect)。当人们知道事件A成功或失败时都更喜欢x而不是y，但当不知道事件A是否成功时他们更喜欢y而不是x，这时就会发生析取效应。析取效应是对确定性原则的违反，因此也是对结果主义的违反。

在本章中，我们将探讨几个推理和决策任务中的非结果主义行为。我们认为，在存在和不存在不确定性的情况下，各种理由和考虑因素的权重不同，从而导致违反确定性原则。我们之前的研究探讨了一些情况，一旦不确定性得到解决，选择特定选项的理由（如去夏威夷或投注）比结果不确定时更令人信服。目前的研究集中在这些情景上，即那些在结果不确定时看似有吸引力的论据，一旦不确定性得到解决就会失去大部分效力。本文认为，由不确定性得到解决引起的观点转变，可能有助于解释一些非结果主义行为的令人困惑的表现。在本章的第一部分中，我们探讨了只有一次机会的囚徒困境，以及与计算机程序对弈的纽科姆问题的一个版本。接下来，我们将分析从决策扩展到推理。我们认为非结果主义推理在沃森的选择任务中扮演着重要的角色，然后我们会描述一个美国金融市场似乎表现出了非结果主义行为的情景。最后，我们探讨了目前的研究结果对不确定性思维分析的意义，并考虑了它们与自然和人工智能之间的比较的相关性。

游戏和决策

囚徒困境

博弈论分析了根据特定规则行事的参与者之间的互动。一个特别的双人游戏受到了极大的关注，那就是囚徒困境（Prisoner's Dilemma）

（Rapoport & Chammah, 1965; Rapoport, 1988）。一个典型的囚徒困境如图7-2所示。单元格条目表示每个玩家获得的收益（如分数）。因此，如果你和对手都合作，每人得到75分。若对方合作而你参与竞争，你会得到85分，而对方得到25分，依此类推。囚徒困境的特点是，无论对手的选择如何，每个玩家在竞争中都比在合作中表现得更好。如果他们都参与竞争，其表现就会比合作时差。虽然重复游戏会出现许多有趣的策略，但本次讨论仅限于只玩一次的囚徒困境（Axelrod, 1984; Kreps & Wilson, 1982; Luce & Raiffa, 1957）。

		他人	
		合作	竞争
你	合作	你: 75 他人: 75	你: 25 他人: 85
	竞争	你: 85 他人: 25	你: 30 他人: 30

图7-2 典型的囚徒困境

注：单元格条目表示你和对方玩家根据你们的选择获得的积分。

这是一种最简单、最尖锐的困境形式。因为对手之间只会遇到一次，所以没有机会传达战略信息、诱导互惠、建立声誉或影响其他玩家的策略选择。因为不管对方在这一场比赛中做了什么，如果参与竞争你就会比参与合作得到更多的分数，所以最主要的策略是竞争。然而，有些人可能会出于道德原因选择合作。当侯世达（Douglas Hofstadter）[\(16\)](#)向一些专家提出这个问题时，大约1/3的人选择了合作（Hofstadter, 1983）。研究人员在许多实验研究中也观察到相似的合作概率（Rapoport, Guyer & Gordon, 1976; Rapoport, 1988）。哲学家丹尼尔·丹尼特（Daniel Dennett）[\(17\)](#)抓住了这种道德动机，他说：“我宁愿做买下布鲁克林大桥的人，也不愿做卖掉它的人。同样，比起通过背叛获得10美元，通过合作获得3美元的感觉更好。”显然，有些人愿意放弃一些收益，以便做出倾向合作的、合乎道德的决定。

我们之前对非结果主义推理的讨论提出了在一次囚徒困境游戏中观察到的合作的另一种解释。一旦玩家知道对方选择了竞争或合作，显然竞争对他来说比合作更有利。但只要对方没有做出决定，相互合作就会成为对双方都有吸引力的解决方案。虽然每个玩家不能影响对方的决定，但他可能会试图尽自己最大的努力（在这种情况下是合作）来达到双方都想要的状态。当然，一旦结果出现，这种推理就不再适用。投票行为就是一个很好的例子。我们知道，个人的投票不太可能影响选举的结果。然而，我们中许多人在选举结果确定后不愿意投票，却倾向于在选举结果尚未确定时进行投票。如果这种对囚徒困境游戏中合作的解释是正确的，我们预测在析取条件下，当另一个玩家的策略未知时有更高的合作率。下面的研究检验了这一假设。

方法。我们组织80名普林斯顿大学本科生进行囚徒困境游戏，电脑屏幕上每次会显示一个囚徒困境游戏，游戏格式如图7-2所示。在每次实验中，被试都通过按下适当的按钮来选择是竞争还是合作。被试按照自己的节奏做出反应，一旦他们选择了自己的策略，屏幕就会被清空，下一个游戏就会出现。每名被试进行40个游戏，其中只有6个是囚徒困境游戏。其他双人游戏（具有不同的收益结构）被穿插在囚徒困境游戏中，以保证被试可以重新考虑每一个游戏，而不是采用“标准”策略。被试会被告知这些游戏是和其他学生一起玩的，而游戏的结果将由他们和每个游戏的新参与者的选择决定。他们的选择不会被透露给任何和他们一起玩的人。因此，被试要玩一系列的一次性游戏，每次都要面对不同的对手。此外，被试被告知他们被随机分配到一个奖励组，该组中的玩家偶尔会在选择自己的策略之前获得其他玩家已经选择的策略的信息。这些信息会随着游戏一起出现在屏幕上，因此被试可以应用它来做出决定。被试将根据他们在整个过程中积累的分数来获得报酬。他们平均获得了6美元，整个过程持续了约40分钟。完整的说明见本章附录。

我们现在关注的是被试玩的6个囚徒困境游戏。在整个过程中，每个囚徒困境游戏都出现了3次：第一次出现在不知道对方战略标准版本中；第二次出现在知道对方选择竞争的信息中；第三次出现在知道对方选择合作的信息中。每个囚徒困境游戏的标准版本先出现，其他两个游戏的顺序在不同被试之间是平衡的。每个游戏的3个版本之间都被一些其他游戏分隔开。我们将每个囚徒困境游戏的3个版本称为囚徒困境“三元组”。前18名被试参与4个囚徒困境三元组游戏，其余被试每人参与6个囚徒困境三元组游戏，总共产生444个三元组游戏。

结果和讨论。被试对囚徒困境三元组的反应如表7-1所示。表7-1A项总结了在所有的444个游戏中，当对方竞争和合作时，被试选择的策略。表7-1B显示了在析取条件下，当对方的策略未知时，这些相同的被试选择的策略。当被告知对方已经选择竞争时，绝大多数的被试都会以竞争作为回报。合作就意味着掉以轻心和扣分。在444个游戏中，当被试被告知对方选择竞争时（表7-1A项），只有3%的被试选择合作。当被告知对方选择合作时，则有更多的被试选择合作。这证实了一种普遍的观点，即当对方合作时，被试会产生一种道德上的回报倾向。在444个游戏中，当被试被告知对方已经选择合作时，有16%的被试选择合作。那么，当对方的策略未知时，被试应该怎么做？由于3%的被试在对方竞争时选择合作，16%的被试在对方合作时选择合作，所以我们预测在对方策略未知的情况下有一个合作率。相反，在444个游戏中，当

对方的策略未知时（表7-1B项），有37%的人选择合作（3个版本中的合作率都有显著差异，在所有情况下 $p<0.001$ ）。当不确定对方选择的策略时，合作倾向的增加不能归因于丹尼特所阐述的那种道德上的必要性。任何基于道德考量的解释都表明，当知道对方已经选择合作时，合作率应该最高，这与表7-1显示的情况相反。

表7-1 囚徒困境

		对方竞争		
		S 竞争	S 合作	
A. 对方的策略已知 ^a				
对方合作	S 竞争	364	7	371 (84%)
	S 合作	66	7	73 (16%)
		430	14	444
		(97%)	(3%)	
B. 对方的策略未知 ^b				
对方合作	S 竞争	113 合作	3 合作	
		251 竞争	4 竞争	
	S 合作	43 合作	5 合作	
		23 竞争	2 竞争	

注：a. 当对方竞争或合作时，被试（S）策略的联合分布。

b. 当对方的策略未知时，被试的策略分布，如A中所述，根据被试在其他玩家选择竞争和合作时的策略选择进行细分。

正如所料，竞争是所有条件下最受欢迎的策略。因此，最常见的选择模式是在所有3个版本中都选择竞争。然而，下一个最经常出现的模式是：在对方选择竞争时选择竞争，在对方选择合作时选择竞争，但在不知道对方策略时选择合作，这种模式占有所有响应三元组游戏的25%（表7-1B项中的444个三元组游戏中的113个），65%的被试对他们参与的6个囚徒困境三元组游戏的至少一个表现出了这种析取效应。在对方策略未知的情况下产生合作的三元组游戏中，69%的被试在对方选择竞争和合作的情况下选择竞争。此模式可以使用图7-3中的树形图来说明。大多数被试选择在上部分支（当对方合作时）及下部分支（当对方

竞争时)中选择竞争。然而,与确定性原则相反,当不知道自己在哪个分支时,许多被试会选择合作。

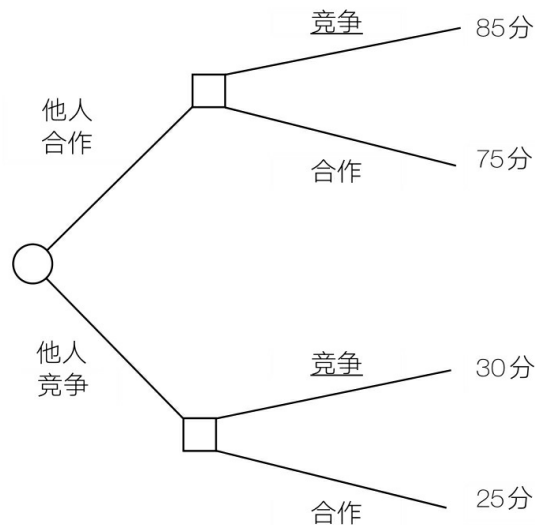


图7-3 囚徒困境的树形图

注:决策节点和机会节点分别由正方形和圆形表示;下划线是模式选择。

一种违反简单规范规则的行为模式既需要正面分析来解释引发观察到的反应的具体因素,也需要反面分析解释为什么没有做出正确的反应(Kahneman & Tversky, 1982)。合取谬误就是一个很好的例子

(Tversky & Kahneman, 1983)。对这种现象的实证分析引发了判断性启发式,例如可用性和代表性,而反面分析将合取错误归因于人们未能发现一个事件包含在另一个事件中的事实,或者他们未能理解这一事实的含义。类似地,对析取效应的反面分析表明人们没有适当地评估所有相关结果。这可能是由于人们有时没有考虑到相关决策树的所有分支,特别是当结果数量很多时。或者,人们可能会考虑所有相关的结果,但由于存在不确定性,可能无法很清楚地看到自己的偏好。考虑一下前文中描述去夏威夷度假的情景。一个刚刚参加完考试但不知道结果的人,在没有特别考虑成功和失败的影响下,可能会觉得现在不是选择去夏威夷的时机。另一种选择是,这个人可能会考虑结果,但不确定会发生什么结果,她可能会对自己的偏好感到不确定。例如,如果她通过了考试,可能十分想去夏威夷度假,但不确定如果考试没有通过是否还会想去度假。只有当她把注意力完全集中在考试不及格的可能性上时,对于去夏威夷的偏好才会变得清晰起来。类似的分析也适用于当前的囚徒困境游戏。由于不知道对手的策略,玩家可能会意识到如果对方竞争,他希望参与竞争,但如果对方合作,那么他可能不确定自己的偏好。在专

注于后一种可能性时，玩家现在更清楚地看到自己也希望在这种情况下竞争。我们认为，不确定性的存在使我们很难将注意力集中在任何一个分支上；扩大注意力的焦点会导致人们丧失敏锐度。因此，没有认识到确定性原则的力量是因为人们不愿意考虑所有的结果，或者是由于他们不愿意在这些结果存在不确定性的情况下制订一个明确的偏好。这一解释与研究结果一致，即一旦人们意识到他们对每种可能结果的偏好，就不再会违反确定性原则（Tversky & Shafir, 1992）。

这里有几个因素可能有助于囚徒困境博弈中析取效应的实证分析。游戏的特点是，每个玩家的个人理性决策导致的结果总体上不是最优的。我们的被试似乎表现出了一种视角的变化，这种变化可以被描述为从个人理性向集体理性的转变。一旦知道了对方的策略，玩家就变成“孤军作战”了。囚徒困境表中只有一列是相关的（对手所选择的策略），游戏的结果取决于自己。当然，个人的理性策略是竞争。此外，在析取条件下，表中的所有4个单元格都在发挥作用。博弈的结果取决于双方的集体决策，而合作是双方的最优集体决策。因此，囚徒困境中观察到的行为模式至少可以部分被解释为，在游戏的析取版本中人们更倾向于采用集体视角。[\(18\)](#)顺便提一句，请注意，集体行动虽然不协调，却是十分可行的。在某种程度上，当我们的非结果主义被试相互竞争时，他们获得的分数（达到合作-合作单元）会比结果主义被试获得的分数（达到竞争-竞争单元）高。罗宾·道斯（Robyn Dawes）和J.欧贝尔（J. Orbell）讨论了在社会困境中合作的潜在好处（Dawes & Orbell, 1995）。

通常情况下，在对方选择竞争和合作时选择竞争的结果主义被试也应当在对方的决定未知时选择竞争。相反，不确定性可以促进合作的倾向，一旦确定了对方的决定，这种倾向就会消失。似乎许多被试没有适当地评估每种可能的结果及其影响。相反，当对手的反应未知时，许多被试倾向于合作也许是作为“诱导”对方合作的一种方式。因为被试很自然地认为对方是同学，对方将会以与他们一样的方式参与游戏，无论他们的决定是什么，看起来对方也可能会做同样的事情。沿着这些思路，L.梅瑟（L. Messe）和J.西瓦切克（J. Sivacek）认为人们高估了其他人在混合动机游戏中会用与他们一样的方式行动的可能性（Messe & Sivacek, 1979）。这种态度可能导致被试选择合作，希望实现联合合作，从而获得最大的互利，而不是竞争和风险带来的联合竞争。如果他们能够协调出一份具有约束力的协议，那么被试肯定会就相互合作达成一致。由于无法在囚徒困境游戏中获得具有约束力的协议，被试仍然倾

向于按照双方都认可的协议行事。尽管实际上不能影响对方的决定，但是被试选择“尽自己的一份力量”来达到双方都喜欢的状态。这种解释与G.夸特隆（G. Quattrone）和特沃斯基的发现一致：人们经常选择对有利结果具有诊断意义的行为，即使这些行为并没有导致这些结果

（Quattrone & Tversky, 1984）。下一节中我们将继续讨论因果推理和诊断推理之间的关系。

我们将上述囚徒困境游戏中对确定性原则的违反现象解释为人们没有以结果主义的方式评估结果。我们现在来考虑针对上述发现的两种不同解释。首先，被试可能在游戏的析取版本中选择合作是因为他们担心自己的选择会在对方做出决定之前传递给对方。当然，一旦对方已经做出决定，这种担忧就不会出现。回想一下，被试被明确告知，他们的选择不会传达给任何与他们玩游戏的人。尽管如此，他们本可以持怀疑态度。然而，实验后的采访显示，虽然有少数被试怀疑系统中同时存在其他玩家，但没有人担心自己的选择会被秘密泄露。在目前的实验中，怀疑不太可能解释被试的策略。

其次，我们认为目前的结果可以通过以下假设来解释：竞争倾向随着实验的进行而增加。为了在标准的囚徒困境中观察被试未受干扰，我们在进行结果已知的游戏之前先展示了这些游戏。因此，随着实验的进行，竞争倾向的增加可能促进了观察到的模式的出现，因为通常在实验的早期，合作程度最高的析取问题就会出现。然而，我们并没有观察到这种时间上的变化。遇到的第一个析取（第6局）、第四个析取（第15局）和最后一个析取（第19局）中的合作率分别为33%、30%和40%。同样，当被试被告知对方已经选择合作时，此时的合作率在前3次发生时平均为13%，在后3次为21%。实际上，随着实验的进行，合作趋于增加将减少观察到的析取效应发生的频率。

回想一下，像侯世达这样的专家一样，近40%的被试在只有一次机会的囚徒困境中选择合作。然而，一旦他们知道了对方的策略，这些合作者中将近70%的人会在对方竞争和合作时都选择竞争。这些参与者遵循了康德绝对命令的一个变体：按照你希望别人怎么做的方式去做。然而，他们觉得没那么有必要按照别人已经采取的方式行事。这一模式表明，在只有一次机会的囚徒困境游戏中观察到的一些合作可能并非源自丹尼特所描述的那种道德上的需要，而是一种一厢情愿的想法和与非结果主义评估的结合。类似的分析也适用于我们接下来要讨论的相关决策问题。

纽科姆问题

纽科姆问题最初由罗伯特·诺齐克（Robert Nozick）提出，之后引发了一场涉及理性决策本质的生动的哲学辩论（Nozick, 1969）。问题的标准版本大致如下。

假设你有两个选择：取出你面前封闭盒子里的物品；或者取出封闭盒子和另一个打开的盒子里的物品，你可以看到打开的盒子里面有1 000美元现金。封闭的盒子里面要么有100万美元要么什么东西都没有，这取决于一个具有神奇预见能力的人，也就是所谓的预言家在你做出决定之前有没有把100万美元放在那里。你知道如果预言家预测你会只选封闭盒子，他会将钱放在封闭盒子里；如果他预测你会同时选择两个盒子，便不会在封闭盒子里放任何东西。你还知道，几乎所有选择两个盒子的人都会发现封闭盒子是空的，他们只得到了1 000美元，而几乎所有只选择了封闭盒子的人都会发现里面有100万美元。你的选择是什么？

许多作者对纽科姆问题与囚徒困境的逻辑关联性进行了评论（Brams, 1975; Lewis, 1979; Sobel, 1991）。在这两种情况下，结果取决于你做出的选择及另一个人的选择，即囚徒困境中的另一个玩家和纽科姆问题中的预言家。在这两种情况下，一个选项（竞争或同时拿两个盒子）支配着另一个选项，但另一个选项（合作或只拿一个盒子）似乎更好，如果预言家或其他玩家知道你将做什么，或许他们会做出和你一样的选择。

纽科姆问题产生的相互矛盾的直觉大致如下（Nozick, 1969）。

论证 1（一个盒子）。如果我选择两个盒子，几乎可以肯定的是，预言家会预测到这一点，并且不会将 100 万美元放入封闭盒子中，因此我只能获得 1 000 美元。如果我只选择封闭盒子，几乎可以肯定的是，预言家会预测到这一点，并且会将 100 万美元放进那个盒子里，所以我会获得 100 万美元。因此，如果我拿两个盒子肯定会得到 1 000 美元，如果我只拿封闭盒子几乎肯定会得到 100 万美元。因此，我应该只选择封闭的盒子。

论证 2（两个盒子）。预言家已经做出了预测，并且已经将 100 万美元放入封闭盒子中或者没放。如果预言家已经将 100 万美元放入封闭盒子中，我拿两个盒子会得到 100 万美元 + 1 000 美元，而如果我只拿走封闭盒子就只能得到 100 万美元。如果预言家没有把 100 万美元放进封闭盒子里，我拿两个盒子可以得到 1 000 美元，而如果我只拿走封闭盒子则拿不到钱。因此，无论 100 万美元是否存在，我都可以通过拿走两个盒子而不是仅拿封闭盒子来获得 1 000 美元。所以我应该拿两个盒子。

马丁·加德纳（Martin Gardner）在《科学美国人》上发表了纽科姆问题，并邀请读者把他们的问答寄给自己，大约70%表现出偏好的读者发现论证1更具吸引力，选择只拿走封闭盒子，而30%的读者认为论证2更有吸引力，因此拿走两个盒子（Gardner, 1973, 1974）。论证2依赖于让人回想起确定性原则的结果主义推理，即无论盒子的状态是什么，被试都会选择两个盒子而不是一个盒子。此外，论证1有更多的问题。从预期效用的角度来看，它似乎预设了预言家预测的内容，事实上，预言家的确这样做了，但在某种程度上，这取决于被试现在决定做什么。除了诡计之外，预言家的特异能力有两种解释。根据第一种解释，预言家只是一个善于判断人性的判断者。通过应用某些数据库（如性别、背景和外表），预言家可能能够预测决策者的反应并取得显著成功。如果这种解释是正确的，那么被试没有理由只拿一个盒子：无论预言家的预测多么有洞察力，拿两个盒子都会比只拿一个盒子更好。第二种解释是预言家确实具有超自然的洞察力。如果不愿意忽视这种可能性，那么被试就可能有可能有理由遵从预言家的神秘力量而只拿一个盒子（Bar-Hillel & Margalit, 1972）。这个谜题引起了许多人的想象。R.坎贝尔（R.

Campbell)和L.索登(L. Sowden)为我们提供了一系列关于纽科姆问题及其与囚徒困境之间关系的有趣文章(Campbell & Sowden, 1985)。

像加德纳的读者一样,许多人在纽科姆问题中只选择一个盒子,这与论证2的逻辑相悖。选择单个盒子可能源于相信预言家的超自然能力。或者,它可能反映了对相关选项的非结果主义评估。为了区分这些解释,我们创建了一个可信的纽科姆问题版本,它不涉及超自然因素。在我们的实验中,由虚构的计算机程序扮演预言家的角色,其对被试所做选择的预测基于先前建立的数据库。实验大致过程如下。在完成前面描述的囚徒困境研究部分后,被试($N=40$)在计算机屏幕上看到如下场景:

你现在还有一次机会获得附加分数。在整个过程中,研究人员会应用麻省理工学院最近开发的一个程序来分析你的偏好模式。基于这个分析,程序已经预测出你在接下来这个问题中的偏好。

20 分

?

盒子 A

盒子 B

考虑上面的两个盒子。盒子 A 确定有 20 分。盒子 B 可能有也可能没有 250 分。你的选择是?

(1) 选择两个盒子(并收集两个盒子中的分数)。

(2) 仅选择盒子 B(并仅收集盒子 B 中的分数)。

如果程序根据你之前的偏好进行预测,你将同时拿走两个盒子,那么它令盒子 B 为空。然而,如果它预测你只拿盒子 B,那么它会在盒子里放 250 分。为了确保在表明你的偏好后该程序不会改变其预测,请向研究人员表明你是选择两个盒子还是只选盒子 B。在表明你的偏好后,按任意键来了解得到的分数。

到目前为止,该程序取得了非常大的成功:仅选择盒子B的被试中,92%的人发现里面有250分,而同时选了两个盒子的被试中,只有17%的人得到了250分。

这个情景为我们提供了一个纽科姆问题的可信版本。虽然计算机程

序非常成功，但它并不是绝对可靠的。(19)此外，任何对反向因果关系的怀疑都已被移除：假设研究人员没有以某种复杂的方式作弊（我们实验后的访谈表明没有被试认为他会作弊），虽然程序的预测已经完成，并且可以在任何时候进行观察，而无须对被试的决定做出进一步反馈。这个问题有一个明确的“共同原因”结构：正如在之前的囚徒困境游戏中所观察到的那样，被试在这类游戏中的策略倾向可以用来预测他在下一场比赛中的首选策略以及程序的预测（Eells, 1982）。虽然选择单个盒子可以判断其中是否有250分，但两件事情之间不存在相关的因果影响。在这种情况下，似乎没有理由只拿一个盒子。正如诺齐克指出的那样，“如果行动或决定……不影响 / 帮助实现 / 影响……获得的状态，那么无论条件概率如何……，都应该执行支配性行为”，即拿走两个盒子（Nozick, 1969）。在这种情况下，看起来人们应该选择两个盒子，因为无论如何两个盒子都比一个盒子好。

结果显示，35%（40名被试中的14名）的人选择两个盒子，而65%的人仅选择盒子B。目前的情景，即从纽科姆问题的原始表述中去掉了所有超自然元素后得到的一个和两个盒子的选择比例，与加德纳从《科学美国人》的读者那里获得的结果大致相同。该怎样解释大多数被试只选择一个盒子的行为呢？如果他们确定250分在那个盒子里（并且可以看到20分在另一个盒子里），他们肯定会选择拿走两个盒子而不是仅仅拿走一个盒子。当然，如果他们知道250分不在那个盒子里，肯定会拿走两个盒子而不是只拿走那个空盒子。换句话说，如果他们知道盒子B不是空的，这些被试就会拿走两个盒子，如果他们知道盒子B是空的，也会拿两个盒子。即使不知道盒子B是不是空的，结果主义者应该也会选择两个盒子。然而，大多数人在盒子里的东西未知时选择了盒子B。请注意，前面讨论过的将析取效应归因于被试无法预测自己的偏好的假设，不能解释目前的发现。如果被试考虑了未知盒子可能的状态，那么他预测自己对更多而不是更少分数的偏好就不会有任何困难。很明显，许多被试没有单独考虑程序预测的结果，因此没有抵挡住选择单个盒子的诱惑，这恰好与较高的奖金相关。

准神奇式思考。“神奇式思考”指的是错误的信念，即人们可以通过某种象征性或其他间接性的行为（如想象一个特定的数字）来影响结果（如骰子的数字），即使该行为与结果并没有因果关系。我们引入了“准神奇式思考”（quasi-magical thinking）这个术语来描述人们的行为，就好像错误地认为自己的行为会影响结果，尽管他们并不真的持有这种信念。正如我们在囚徒困境和纽科姆问题中观察到的偏好模式，都

可以被描述为准神奇式思考。当程序的预测已知时，结果完全取决于被试的决定，显而易见的选择是选择两个盒子。但只要程序的预测结果未知且最终结果取决于被试和程序的行为，就会有诱惑存在，好像一个人的决定会影响程序的预测一样。正如A.吉伯德（A. Gibbard）和W.哈珀（W. Harper）试图解释人们对单个盒子的选择一样，“一个人可能.....想要创造一种他所希望的有关世界状态的迹象，即使他知道创造这种迹象的行为绝不会带来所期待的状态本身，他依然会这样做”（Gibbard & Harper, 1978）。当然，大多数人实际上并不认为他们能够改变程序或改变对方做出的决定。尽管如此，他们还是觉得有必要“尽自己的一份力”，以达到预期的结果。夸特隆和特沃斯基也为我们展示了这种准神奇式思考的例子，他们的被试实际上通过选择行为（如将手放在非常冰冷的水中一段时间）在医院检查中“作弊”了（Quattrone & Tversky, 1984）。尽管他们肯定知道自己的行为不可能产生预期的结果，但仍认为这样做有助于诊断出好的结果（如拥有一颗强壮的心脏）。

我们认为，准神奇式思考是与自欺欺人和控制错觉有关的几种现象的基础。例如，夸特隆和特沃斯基指出，加尔文主义者的行为似乎将决定他们是去天堂还是地狱，尽管他们相信神的预先决定，即他们的命运在自己出生之前就已经注定了，也依然如此（Quattrone & Tversky, 1984）。一些作者，尤其是E. J.兰格（E. J. Langer）指出，人们通常表现得好像自己可以对随机事件施加控制，因此与投掷硬币之后相比，人们在投掷之前会表现出不同的态度，并在投掷时下更大的赌注

（Langer, 1975; Rothbart & Snyder, 1970; Strickland, Lewicki & Katz, 1966）。(20)然而，大多数人并不真正相信他们可以控制掷硬币，也不认为纽科姆问题中选择单个盒子会影响程序已经做出的预测。在上述和其他类似情况中，人们都可能知道自己并不能影响结果，但他们表现得好像可以。据说物理学家尼尔斯·玻尔（Niels Bohr）曾被一位记者问及悬挂在他家门前的马蹄（据说是为了带来好运）时，他解释说自己当然不相信这种无稽之谈，但即便不相信，也会觉得它对自己有所帮助。

当然，要弄清人们真正相信什么是极其困难的。上文中的讨论表明，我们不能总是从行动中推断信念。人们可能会表现得好像自己能够影响无法控制的事件，尽管他们实际上并不相信自己能够做到这一点。例如，轻掷以获小数、重掷以获大数的掷骰子玩家可能并不一定相信投掷的性质会影响结果（Henslin, 1967）。那些迷信的人，比如佩戴幸运符或避免穿过黑猫走过的路，可能并不真的相信他们的行为能影响未

来。在某种意义上，准神奇式思考似乎比神奇式思考更理性，因为它不会让人相信明显荒谬的信念。此外，准神奇式思考似乎更令人困惑，因为它破坏了信念和行动之间的联系。虽然神奇式思考涉及站不住脚的信仰，但准神奇式思考会引发令人费解的行为。我们认为，存在不确定性是准神奇式思考的主要诱因；很少有人表现得好像自己可以通过执行一个诊断替代事件的动作来改变已经确定的事件。在这种情况下，夸特隆和特沃斯基实验中的被试如果知道自己心脏的强弱情况，而不是对“诊断”结果未知，他们就不太愿意将手放在冰冷刺骨的水中了（Quattrone & Tversky, 1984）。

综合讨论

如上一节所述，人们往往没有考虑不确定性事件可能的结果。在不确定状况下进行思考的困难在各种情况中都显露了出来，它们包括推理和决策任务，并且在实验室内外都可以观察到。在本节中，我们会将非结果主义评估的分析扩展到演绎推理和经济预测中。

华生选择任务

人类推理研究中最著名的任务之一是由P. C.华生（P. C. Wason）设计的选择任务（Wason, 1966）。在任务的典型版本中，被试面前有4张卡片，每张卡片的一面有一个字母，另一面有一个数字。每张卡片只显示一面。例如：

E	D	4	7
---	---	---	---

被试的任务是指出要验证这一规则，需要翻开哪些卡片：“如果卡片的一面是元音字母，那么卡片的另一面必定是偶数。”这个问题的简单性具有欺骗性。绝大多数被试都无法解决它。[\(21\)](#)大多数人只选择翻开写有字母E的卡片，或写有字母E和数字4的卡片，而正确的选择是翻开写有字母E和写有数字7的卡片。选择任务的难度令人费解，特别是因为人们通常可以毫无困难地评估可能隐藏在每张卡片的另一面的结果的相关性。华生和约翰逊-莱尔德评论了被试评估潜在结果相关性的能力（即理解规则的真实条件）与相关卡片的不恰当选择之间的差异（1970；Wason, 1969）。例如，被试明白4张卡片的另一面不管是元音字母还是辅音字母都不会导致规则本身出现矛盾，但是当另一面未知

时，他们选择翻开4张卡片。被试同样明白，如果写有数字7的那张卡片的另一面是辅音字母，那么不会使规则本身出现矛盾，而如果卡片的背面是元音字母，那么规则就出现了矛盾，但他们却忽略了翻开写有数字7的卡片。上述模式类似于析取效应，当可以轻易地评估特定结果的相关性的被试，在面对析取结果而未能应用这一知识时，就会出现上述模式。正如J.埃文斯（J. Evans）所指出的那样：“这有力地证实了这一观点，即卡片的选择不是基于对翻转卡片结果的任何分析。”（Evans, 1984）就像那些在考试结果未知时推迟去夏威夷度假的人，以及那些在囚徒困境的析取版本中选择合作的人一样，执行选择任务的被试未能考虑每个事件的后果。在不考虑卡片另一面每种特定符号的结果的前提下，当卡片的另一面未知时，它们似乎仍然隐藏在不确定的面纱之后。

大量的研究探索了难以捉摸的思维过程，这是被试在选择任务中的表现的核心。事实上，一个复杂的内容效果模式已经从原始任务的许多变化中显现出来（Johnson-Laird, Legrenzi & Legrenzi, 1972; Griggs & Cox, 1982; Wason, 1983; Evans, 1989; Manktelow & Evans, 1979）。为了解释这些发现，研究人员提出了验证偏见（Johnson-Laird & Wason, 1970）、匹配偏见（Evans & Lynch, 1973; Evans, 1984）、特定领域经验的记忆（Griggs & Cox, 1982; Manktelow & Evans, 1979）、语用推理模式（Cheng & Holyoak, 1985, 1989），以及寻找骗子的内心倾向（Cosmides, 1989）等解释。这些解释的共同之处在于，它们都是对选择任务中的表现的描述，但并不涉及形式推理。相反，人们被假定专注于已明确提及的项目，应用预先存储的知识结构，或记得过去的相关经验。华生总结道：“在这些情况下发生的推理过程，并不是‘逻辑’推理的实例。”（Wason, 1983）因此，人们发现对每一个孤立的结果进行逻辑推理相对容易，但结果的析取导致他们暂停了逻辑推理。这让人想起了D. N.奥谢森（D. N. Osherson）和E.马克曼（E. Markman）针对8岁儿童的研究，当被问到“我手中隐藏的单色筹码是或不是黄色”是否正确时，他们回答“不知道”，因为他们看不到它（Osherson & Markman, 1974）。虽然大多数成年人发现筹码的析取是无关紧要的，但更微妙的析取可能导致判断的暂时中止。

金融市场中的析取效应。非结果主义决策的一个结果是人们有时会寻求对其决策没有影响的信息。例如，在前文描述的去夏威夷度假问题中，被试愿意为那些不会改变他们的选择的信息付费，但正如我们所解释的那样，这仅仅是为了澄清他们的选择的理由。在早期囚徒困境实验的一个变体中，我们给一组新的被试提供了同样的囚徒困境游戏，但

这一次，被试不会被告知对方的决定，而是在做出自己的选择之前可以用非常少的费用知道对方的决定。无论对方是选择竞争还是合作，绝大多数被试选择竞争，但81%的被试会先选择花钱了解对方的决定。虽然这种行为可以归因于好奇心，但我们猜想如果人们意识到这并不会影响他们的决定，那么他们愿意为这些信息付费的意愿就会降低。在不确定的情况下，人们可能会频繁地搜索对决策没有影响的信息。例如，我们可能会在预订酒店之前打电话询问海滩酒店是否有游泳池，尽管事实上不管是否有游泳池，我们最终还是会去的。下面这段关于美国金融市场的描述为我们提供了一个有趣的非结果主义信息评估案例。

在1988年美国总统大选前的几个星期里，“由于总统选举前的谨慎”，美国金融市场相对不活跃且稳定（《纽约时报》，11月5日）。“投资者不愿在离总统选举还有7天时有重大举动，这周充满了经济不确定性”（《华尔街日报》，11月2日）。据《华尔街日报》的报道，市场正在“消磨时间”。选举后的第二天，席拉森-拉曼·休顿投资公司的交易台负责人指出，“几乎什么都没有发生，可能至少要到周三才会有改变”（《华尔街日报》，11月8日）。一家投资公司的总裁评论道：“一旦选举的不确定性被消除，投资者就可以开始对经济、通胀和利率前景有更好的感知。”（《纽约时报》，11月2日）投资组合策略师解释说：“最近几天，选举结果给市场带来了明显的阴影（《纽约时报》，11月9日）。其在未来几天内可能会迅速发展。”事实上，在大选之后，一个明确的方向浮出水面。美元大幅下跌至10个月以来的最低水平，股票和债券价格也会下跌。在布什获胜后的一周内，道琼斯工业平均指数总共下跌了78个点。[\(22\)](#)高盛公司投资委员会联合主席解释说：“选举后的真实情况正在形成。”（《华尔街日报》，11月21日）分析师解释道，美元的下跌“反映了市场对美国贸易和预算赤字的持续担忧”，“选举的兴奋期已经结束，蜜月期已经结束，现实经济已经回归”（《华尔街日报》，11月10日）。《纽约时报》11月12日的头版称金融市场“普遍支持布什先生的当选，并已经预料到他的胜利，但在选举后的3天内，他们对其未来表示了担忧”。当然，如果迈克尔·杜卡基斯（Michael Dukakis）先生当选，金融市场可能至少也会引起同样的担心。大多数交易员都同意，在选举日写道：“如果民主党候选人杜卡基斯从落后走向胜利，股市将大幅下跌。”事实上，市场对布什获胜做出的反应就像它对杜卡基斯获胜的反应一样。大选后一位交易员解释说：“当我走进去看屏幕时，我以为杜卡基斯赢了。”（《纽约时报》，11月10日）

在大选前长时间的沉寂之后，股市在布什获胜后立即下跌，如果杜卡基斯获胜，股市肯定至少也会出现同样程度的下跌。当然，对金融市场行为的全面分析可能会揭示出许多复杂的问题。例如，有一种可能性是，出乎意料的胜利优势（令人惊讶的最后时刻的结果）可能促成了这种矛盾的效果。然而，凑巧的是，“报纸和电视网络与民调专家所认为的大选结果最接近”（《纽约时报》，11月10日）。在大选前一周，尽管一些人认为杜卡基斯的支持率仍有可能“大幅上升”，但盖洛普咨询公司（Gallup）、《华盛顿邮报》、《华尔街日报》和《纽约时报》进行的民调显示，布什将以平均9个百分点的优势获胜，这比最终的8个百分点只差1个百分点。同样，预计民主党将继续控制国会（预计他们将在国会获得一两个席位）和众议院，而事实正是如此。选举结果似乎并不令人意外。至少从表面上看，这一事件具备了造成析取效应的所有条件：如果布什当选，市场将会下跌；如果杜卡基斯当选，市场将会下跌；但在知道两人中谁当选之前，市场不会做出任何改变。处于这样一个重大析取的节点，似乎阻止了华尔街认真对待选举的后果。虽然任何一位候选人当选总统都会让金融市场“担心它的走向”，但不确定的临时局势突显了“选举前要谨慎”的必要性。毕竟，我们怎么能在知道是谁在做之前就担心“他会从这里走到哪里”呢？

结论性意见

我们在包含不确定性的简单环境中观察到了违反确定性原则的决策和推理模式。我们认为，这些模式反映出了人们只是未能发现并应用这一原则，而不是缺乏对其规范性吸引力的认识。当我们要求被试在每种结果下表明他们的首选方案，然后在析取状态下做出决定时，大多数在每种结果下都选择了相同选项的被试在结果不确定时也选择了该选项（Tversky & Shafir, 1992）。换句话说，当确定性原则的逻辑变得明显时，析取效应的频率显著降低了。与有关决策的其他规范性原则一样，确定性原则在其应用透明时通常会得到满足，而在应用不透明时有时会被违背（Tversky & Kahneman, 1986）。

导致人们不愿意进行结果性思考的因素有很多。通过事件树进行思考需要人们暂时将其假设为真，而实际上它可能是假的。人们可能不愿意做出这样的假设，特别是当另一种可能性选择（树的另一分支）很容易得到时。显然人们很难将全部注意力集中到事件树的每个分支上（Slovic & Fischhoff, 1977）。因此，人们可能不愿意接受各种假设性分支。此外，他们可能缺乏详细研究树形结构的动机，因为通常情况

下，他们假定单独评估分支很难解决问题。我们通常倾向于通过筛选无关紧要的析取来表述问题：那些遗留下来的东西通常被认为涉及真正的冲突。

本章研究的析取情景相对简单，只涉及两种可能的结果。多种结果的析取更难以思考，因此更有可能引起非结果主义推理。对经济、社会或政治决策来说尤其如此，其中情况的严重性和复杂性可能掩盖了这样一个事实：所有可能性结果最终可能由于不同的原因导致类似的决定。

推理中的缺点通常可归因于人类作为信息处理者的数量限制。“困难问题”的典型特征是“所需的知识量”、“记忆负荷”或“搜索空间的大小”（Kotovsky, Hayes & Simon, 1985; Kotovsky & Simon, 1990）。这些限制在许多问题中起着关键作用。它们解释了为什么我们不记得扑克游戏中之前出过的所有牌，或者为什么我们在国际象棋游戏中可以提前规划的步数受到了严重限制。然而，这些限制不足以解释与思考有关的所有困难。与人们相对容易执行的许多复杂任务相比，本文研究的问题在计算上非常简单，它涉及两个明确定义状态的单个析取。目前的研究强调了逻辑复杂性 with 心理困难之间的差异。例如，与对人来说微不足道但对于人工智能来说极其困难的“框架问题”相比，通过析取进行思考的任务对人工智能（通常应用“树搜索”和“路径查找”算法）来说微不足道但对人来说却非常困难。由此导致的推理失败可能构成了自然智能和人工智能之间的根本区别（McCarthy & Hayes, 1969; Hayes, 1973）。

附录：在囚徒困境游戏中给予被试的指示

欢迎来到校际计算机游戏比赛。该游戏将在电脑上进行。在这个游戏中，你将看到自己和另一个玩家的情况。每种情况都需要你做出战略决策：与其他玩家合作或竞争。其他玩家也必须做出类似的决定。

每种情况都会有一个收益矩阵，该矩阵将根据你们是竞争还是合作来决定每个人的得分。矩阵如下所示：

	其他合作者		其他竞争者	
你合作	你	20	你	5
	其他玩家	20	其他玩家	25
你竞争	你	25	你	10
	其他玩家	5	其他玩家	10

根据这个矩阵，如果你们两个都选择合作，你们将获得较多的积分（每个人20分）。如果你选择合作而另一方选择竞争，另一方将获得25分，你只能获得5分。同样，如果你选择竞争而另一方选择合作，你将获得25分，对方只能获得5分。最后，如果你们都选择竞争，每人只能获得10分。

你将看到上面显示的众多类型的矩阵。在每种情况下，你都会被要求表明选择竞争还是合作。如上面的矩阵所示，如果你们两个选择合作，你经常会做得非常好，如果你们两个都选择竞争，你会表现得差一些；如果一个人选择竞争而对方选择合作，竞争的人往往会比对方表现得更好。

你将与当前使用电脑的其他学生一起参与游戏。对于每个新矩阵，你将与另一个不同的人匹配。因此，你永远不会与同一个人对战两次。

你已被任意分配到奖金组。随机奖励计划有时会告知你对方已经做出的选择。

因此，例如在呈现新矩阵时，可能会告诉你对方已选择竞争。你可以随意使用附加信息来帮助自己选择策略（你的策略不会被透露给与你一起玩的任何人）。

在游戏结束后，你积累的积分将转换为支付给你的实际金额（根据预先确定的算法）。积累的积分越多，你赚的钱就越多。

当然，不存在“正确”的选择。人们通常发现某些情况更有利于合作，而其他情况则更有利于竞争。矩阵差异性很大，其结果取决于你的选择及每个回合中不同玩家的选择。请仔细观察每个矩阵，并在每个特定情况下分别给出你的首选策略。此外，请务必注意奖励计划告知你的对方的选择情况。如果有任何疑问，请咨询负责人。否则，跳转到最后

并开始。

参考文献

Axelrod, R. (1984). *The evolution of cooperation*. New York, NY: Basic Books.

Bacharach, M., & Hurley, S. (1991). Issues and advances in the foundations of decision theory. In M. Bacharach & S. Hurley (Eds.), *Foundations of decision theory: Issues and advances* (pp. 1–38). Oxford, England: Basil Blackwell.

Bar-Hillel, M., & Margalit, A. (1972). Newcomb's paradox revisited. *British Journal for the Philosophy of Science*, 23, 295–304.

Brams, S. J. (1975). Newcomb's Problem and Prisoner's Dilemma. *Journal of Conflict Resolution*, 19(4), 596–612.

Campbell, R., & Sowden, L. (Eds.). (1985). *Paradoxes of rationality and cooperation: Prisoner's Dilemma and Newcomb's Problem*. Vancouver, BC: University of British Columbia Press.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391–416.

Cheng, P. W., & Holyoak, K. J. (1989). On the natural selection of reasoning theories. *Cognition*, 33, 285–313.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? *Cognition*, 31, 187–276.

Dawes, R. M., & Orbell, J. M. (1995). The potential benefit of optional play in a one-shot prisoner's dilemma game. In K. Arrow et al. (Eds.), *Barriers to conflict resolution* (pp. 62–85). New York, NY: Norton.

Eells, E. (1982). *Rational decision and causality*. Cambridge, England: Cambridge University Press.

Evans, J. St. B. T. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology*, 75, 451–468.

Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Evans, J. St. B. T., & Lynch, J. S. (1973). Matching bias in the selection task. *British Journal of Psychology*, 64, 391–397.

Gardner, M. (1973). Free will revisited, with a mind-bending prediction paradox by William Newcomb. *Scientific American*, 229(1), 104–108.

Gardner, M. (1974). Reflections on Newcomb's problem: A prediction and free-will dilemma. *Scientific American*, 230(3), 102–109.

Gibbard, A., & Harper, W. L. (1978). Counterfactuals and two kinds of expected utility. In C. A. Hooker, J. J. Leach, & E. F. McClennen (Eds.), *Foundations and applications of decision theory* (Vol. 1, pp. 125–162). Dordrecht, the Netherlands: Reidel.

Gilhooly, K. J. (1988). *Thinking: Directed, undirected, and creative* (2nd ed.). San Diego, CA: Academic Press.

Greene, S. B., & Yolles, D. J. (1990). *Perceived determinacy of unknown outcomes*. Unpublished manuscript, Princeton University.

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, 73, 407–420.

Hammond, P. (1988). Consequentialist foundations for expected utility. *Theory and Decision*, 25, 25–78.

Hayes, P. (1973). The frame problem and related problems in artificial intelligence. In A. Elithorn & D. Jones (Eds.), *Artificial and human thinking* (pp. 45–59). San Francisco, CA: Jossey-Bass.

Henslin, J. M. (1967). Craps and magic. *American Journal of Sociology*, 73, 316–330.

Hofstadter, D. R. (1983, June). Dilemmas for superrational thinkers, leading up to a luring lottery. *Scientific American*. Reprinted in Hofstadter, D. R. (1985). *Metamagical themas: Questing for the essence of mind and pattern*. New York, NY: Basic Books.

Hurley, S. L. (1989). *Natural reasons: Personality and polity*. New York, NY: Oxford University Press.

Hurley, S. L. (1991). Newcomb's Problem, Prisoner's Dilemma, and collective action. *Synthese*, 86, 173–196.

Johnson-Laird, P. N., Legrenzi, P., & Legrenzi, S. M. (1972). Reasoning and a sense of reality. *British Journal of Psychology*, 63, 395–400.

Johnson-Laird, P. N., & Wason, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, 1, 134–148.

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. *Cognition*, 11, 123–141.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology*, 17, 284–294.

Kotovsky, K., & Simon, H. A. (1990). What makes some problems really hard: Explorations in the problem space of difficulty. *Cognitive Psychology*, 22, 143–183.

Kreps, D., & Wilson, R. (1982). Reputations and imperfect information. *Journal of Economic Theory*, 27, 253–279.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32, 311–328.

Levi, I. (1991). Consequentialism and sequential choice. In M. Bacharach & S. Hurley (Eds.), *Foundations of decision theory: Issues and advances* (pp. 92–122). Oxford, England: Basil Blackwell.

Lewis, D. (1979). Prisoner's Dilemma is a Newcomb Problem.

Philosophy & Public Affairs, 8, 235–240.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions*. New York, NY: Wiley.

Manktelow, K. I., & Evans, J. St. B. T. (1979). Facilitation of reasoning by realism: Effect or non-effect? *British Journal of Psychology*, 70, 477–488.

McCarthy, J., & Hayes, P. (1969). Some philosophical problems from the standpoint of Artificial Intelligence. In B. Meltzer & D. Michie (Eds.), *Machine intelligence* (pp. 463–502). New York, NY: American Elsevier.

Messe, L. A., & Sivacek, J. M. (1979). Predictions of others' responses in a mixed-motive game: Selfjustification or false consensus? *Journal of Personality and Social Psychology*, 37(4), 602–607.

Nozick, R. (1969). Newcomb's problem and two principles of choice. In N. Rescher(Ed.), *Essays in honor of Carl G. Hempel* (pp. 114–146). Dordrecht, the Netherlands: Reidel.

Osherson, D. N., & Markman, E. (1974–75). Language and the ability to evaluate contradictions and tautologies. *Cognition*, 3(3), 213–226.

Quattrone, G. A., & Tversky, A. (1984). Causal versus diagnostic contingencies: On self-deception and on the voter's illusion. *Journal of Personality and Social Psychology*, 46(2), 237–248.

Rapoport, A. (1988). Experiments with n-person social traps I: Prisoner's Dilemma, weak Prisoner's Dilemma, Volunteer's Dilemma, and Largest Number. *Journal of Conflict Resolution*, 32(3), 457–472.

Rapoport, A., & Chammah, A. (1965). *Prisoner's Dilemma*. Ann Arbor, MI: University of Michigan Press.

Rapoport, A., Guyer, M. J., & Gordon, D. G. (1976). *The 22 game*. Ann Arbor, MI: University of Michigan Press.

Rothbart, M., & Snyder, M. (1970). Confidence in the prediction and

postdiction of an uncertain event. *Canadian Journal of Behavioural Science*, 2, 38–43.

Savage, L. J. (1954). *The foundations of statistics*. New York, NY: Wiley & Sons.

Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 544–551.

Sobel, J. H. (1991). Some versions of Newcomb's Problem are Prisoner's Dilemmas. *Synthese*, 86, 197–208.

Strickland, L. H., Lewicki, R. J., & Katz, A. M. (1966). Temporal orientation and perceived control as determinants of risk-taking. *Journal of Experimental Social Psychology*, 2, 143–151.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90, 293–315.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, 59(4, part 2), 251–278.

Tversky, A., & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological Science*, 3, 305–309.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology* (Vol. 1, pp. 135–151). Harmondsworth, England: Penguin.

Wason, P. C. (1969). Structural simplicity and psychological complexity: Some thoughts on a novel problem. *Bulletin of the British Psychological Society*, 22, 281–284.

Wason, P. C. (1983). Realism and rationality in the selection task. In J. St B. T. Evans (Ed.), *Thinking and reasoning: Psychological approaches* (pp. 44–75). London, England: Routledge & Kegan Paul.

Wason, P. C., & Johnson-Laird, P. N. (1970). A conflict between

selecting and evaluating information in an inferential task. *British Journal of Psychology*, 61, 509–515.

第8章 无风险选择中的损失厌恶：参考依赖模型

阿莫斯·特沃斯基
丹尼尔·卡尼曼

标准的决策模型假定偏好不依赖于现有资产。这一假设大大简化了对个人选择的分析和对交易的预测：绘制无差异曲线时没有参考当前持有的股票，而且科斯定理认为除了交易成本，初始权利不影响最终分配。但真实的情况更为复杂。有大量证据表明，初始权利确实很重要，即使在没有交易成本或收益影响的情况下，货物之间的交换率也会因获得和放弃而大不相同。根据对价值的心理学分析，参考水平在决定偏好方面起着很大的作用。在本文中，我们回顾了这一命题的证据，并提出了通过引入参考状态来概括标准模型的理论。

目前对无风险选择的分析拓展了我们在不确定情况下选择的处理方法，其中风险前景的结果由一个具有3个基本特征的效用函数来评估（Kahneman & Tversky, 1979, 1984; Kahneman & Tversky, 1991）。第一个是参考点依赖：价值载体是相对于参考点定义的收益和损失。第二个是损失厌恶：函数在负数区间比在正数区间更陡；损失大于相应的收益。第三个是敏感度递减：收益和损失的边际价值随着规模的减小而减小。这些特征形成了一个不对称的S形效用函数，在参考点上方凹陷并在其下方凸出，如图8-1所示。

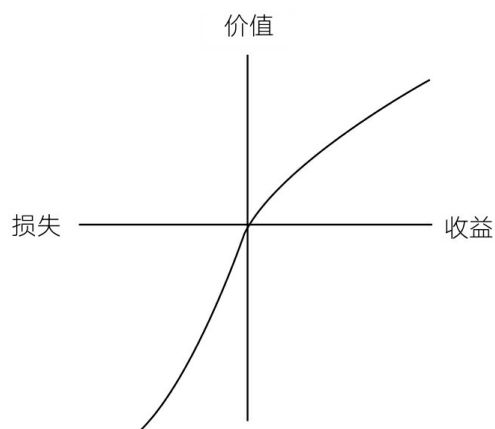


图8-1 效用函数示意图

在这里，我们应用了参考点依赖、损失厌恶和对无风险选择分析的敏感度递减现象。为了推动这种分析，我们首先回顾一些选定的实验演示。

经验性证据

本节讨论的例子可参考图8-2进行分析。在每种情况下，我们都会考虑两个选项 x 和 y （它们在两个有价值的维度上有所不同），并说明它们之间的选择是如何受到评估它们的参考点的影响的。这些偏好逆转的常见原因是维度1和维度2上 x 和 y 之间的差异的相对权重随着这些属性的参考值的位置而变化。损失厌恶意味着当差异被评估为损失时，它对维度的影响通常大于将相同的差异评估为收益时的影响。敏感度递减意味着当两个选项远离相关维度的参考点时，差异的影响会减弱。这个简单的方案可用于组织大量的观察。虽然孤立的发现可能会有其他解释，但整个证据体系为损失厌恶现象提供了强有力的支持。

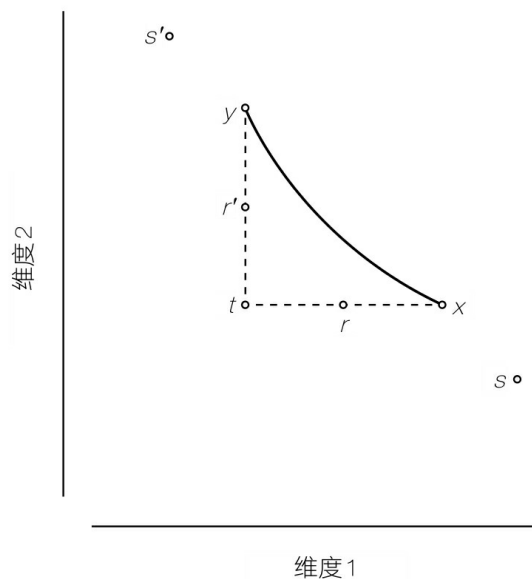


图8-2 x 和 y 之间的选择具有多个参考点

即时禀赋。损失厌恶的直接后果是放弃有价值的商品带来的效用损失大于接受它带来的效用增益。塞勒将这种差异称为禀赋效应

（endowment effect），因为当一种商品被纳入一个人的捐赠时，价值似乎会发生变化（Thaler, 1980）。卡尼曼等人在一系列实验中测试了禀赋效应，这些实验是在课堂环境中进行的（Kahneman, Knetsch &

Thaler, 1990)。在其中一个实验中，学生们选好座位后，在1/3个座位前放置了印有图案的杯子（零售价约5美元）。所有参与者都收到了一份问卷。给予杯子接收者（卖家）的表格显示“你现在拥有这件物品；如果你觉得之后的出价可以接受，你可以选择卖掉它；对于之后每种可能的价格，你都可以决定自己希望（ x ）出售你的物品并接受此价格或（ y ）保留你的杯子并带回家”。被试为杯子确定的价格为0.50~9.50美元，以50美分递增或增减。一些未收到杯子的学生（“选择者”）会收到一份类似的调查问卷，告知他们可以选择收到一个杯子或一笔待确定的钱。该调查显示了他们更倾向于选择杯子，还是一笔0.50~9.50美元的钱。

选择者和卖家面临着完全相同的决策问题，但他们的参考状态不同。如图8-2所示，选择者的基准状态为 t ，他们在两个主导 t 的选项之间进行正面选择：得到一个杯子或一笔现金。卖家从 y 评估相同的选项，他们必须在维持现状（留下杯子）或放弃杯子以换取金钱之间做出选择。因此，选择者认为它是一种收益，而卖家认为它是一种损失。损失厌恶意味着在两种情况下，杯子与货币的兑换率会有所不同。事实上，在一次实验中，杯子的中位数对卖家而言为7.12美元，对选择者而言为3.12美元；在另一次实验中这些数据分别为7.00美元和3.50美元。这些价值之间的差异反映了一种禀赋效应，这种效应显然是通过赋予个人对一种消费品的产权而在瞬间产生的。

禀赋效应的解释可以通过下面的思维实验来阐明：

想象一下，作为一个选择者，你更喜欢4美元而不是一个杯子。你知道大多数卖家更喜欢这个杯子而不是6美元，并相信如果自己有这个杯子，也会这么做。有了这些知识，你更有可能选择杯子而不是5美元吗？

如果确实是这样，大概是因为你改变了对拥有杯子带来的快乐的评估。如果你仍然更喜欢4美元而不是杯子，我们认为这是一种更有可能的反应，这表明你将禀赋的影响解释为一种对放弃杯子的厌恶情绪，而不是拥有它带来的意想不到的快乐。

现状偏好。在许多决策问题中，保持现状也是一种选择。正如对杯子例子中卖家问题的分析所示，损失厌恶引发了想要维持现状而不是其他选择的偏好。在图8-2中，在 t 的角度下对 x 和 y 不关心的决策者在 x 的角

度下会倾向x多于y，在y的角度下则会倾向y多于x。W.萨缪尔森（W. Samuelson）和理查德·泽克豪泽引入了“现状偏好”一词来描述参考位置的这种效应（Samuelson & Zeckhauser, 1988）。

J. L.尼奇（J. L. Knetsch）和J. A.辛登（J. A. Sinden）为现状偏好提供了令人信服的实验论证（Knetsch & Sinden, 1984, 1989）。在后一项研究中，两个本科班级的学生被要求回答一份简短的问卷。其中一个班级的学生会立即获得一个装饰好的杯子作为补偿；另一个班级的学生则会收到一大块瑞士巧克力。填写结束后，两个班级的学生都会看到另一样礼物，并且他们可以通过提交一张写有“交换”字样的卡片来选择交换自己收到的礼物。虽然交换行为的交易成本非常小，但大约90%的被试保留了他们收到的礼物。

尼奇和辛登记录了在大量决策中存在的现状偏好，包括工作、汽车颜色、金融投资和政策问题方面的假设性选择（Knetsch & Sinden, 1988）。每个问题的替代版本都被呈现给不同的被试：每个选项被指定为其中一个版本的现状；一个（中立的）版本没有列出任何选项。为每个问题列出的选项数量是有系统性区别的。我们通过回归被试对中性版本 $P(N)$ 中相同选项的选择比例，来对选择现状 $P(SQ)$ 或现状替代品 $P(ASQ)$ 的比例进行分析。下述公式很好地描述了分析结果：

$$P(SQ) = 0.17 + 0.83P(N) \text{ 和 } P(ASQ) = 0.83P(N)$$

$P(SQ)$ 和 $P(ASQ)$ 之间的差异（0.17）是该实验中的现状偏好的尺度。

尼奇和辛登也于1988年在哈佛员工医疗计划选择的实地研究中获得了现状偏好的证据。他们发现，尽管每年都有审查和改变决策的最低成本的机会，但通常情况下新员工比在该计划实行之前雇用的员工更有可能选择新的医疗计划。此外，现状的微小改变比大的变化更受青睐：从原本最受欢迎的蓝十字 / 蓝盾计划转来的入会者倾向于支持该计划的新版本而不是其他新的替代方案。尼奇和辛登还观察到，尽管回报率存在很大差异，但美国教师退休基金会的养老金储备分配往往每年都非常稳定。他们引用现状偏好作为品牌忠诚度和先锋企业优势的解釋，并指出，忽视现状效应的理性模型“将得出过于激进的结论，夸大个人对不断变化的经济变量的反应，并预测出比观察到的更强的不稳定性”。

损失厌恶意味着现状偏好。然而，正如尼奇和辛登于1988年所指出

的那样，即使在没有损失厌恶的情况下，也有一些因素会导致现状偏好，比如思考成本、交易成本和对先前选择的心理承诺。

改进与权衡。考虑从参考点 r 和 r' 评估图8-2中的选项 x 和 y 。当从 r 的角度评估时，选项 x 仅仅是维度1上的增益（改进），而 y 将维度2的增益与维度1的损失结合了起来。当从 r' 的角度评估相同的选项时，这些关系则相反。对损失厌恶的考虑表明，从 r 的角度而不是从 r' 的角度来看， x 更有可能为首选。

90名本科生参加了一项旨在检验这一假设的研究。他们收到的书面指示说明会有一些随机选择的被试将收到礼品包。给一半被试（晚餐组）的礼物包括“一顿麦克阿瑟公园餐厅的免费晚餐和一份斯坦福月历”。给另一半被试（照片组）的礼物是“一张8×10的专业肖像照片和一份斯坦福月历”。所有被试都被告知，再次随机选出的一些获奖者将有机会用最初的礼物交换下列选项之一：

x ：麦克阿瑟公园餐厅的两顿免费晚餐。

y ：一张8×10的专业肖像照，加上两张5×7的冲印照，以及3张钱包大小的照片。

被试被要求表明他们是否倾向于：第一，保留原始礼物；第二，将其换成 x ；第三，将其换成 y 。如果人们不愿意放弃原来的礼物，那么正如损失厌恶所表明的那样，原始礼物是一顿晚餐（ r ）的被试比原始礼物是单张照片（ r' ）的被试更喜欢两顿晚餐（ x ）而不是多张照片（ y ）。结果证实了这一猜想。只有10名被试选择保留原始礼物。在其余的被试中，81%的晚餐组和52%的照片组选择了选项 x （ $p < 0.01$ ）。

优势和劣势。在下一个演示中，研究人员将小增益和小损耗的组合与大增益和大损耗的组合进行了比较。损失厌恶意味着，如果将两个选项之间的相同差异视为两个缺点（相对于参考状态）之间的差异，那么其权重将大于将其视为两个优点之间的差异。图8-2表明，比起从 s' 的角度看，从 s 的角度来看 x 更可能优于 y 。因为维度1中的 x 和 y 之间的差异涉及相对于 s 的劣势和相对于 s' 的优势。类似的论证也适用于维度2。在对这个猜想进行的检验中，被试回答了以下问题的两个版本之一。

想象一下，作为职业培训的一部分，你被分配到一份兼职工作。培训现在结束了，你必须找工作。你考虑了两种可能性。它们和你的培训

工作在很多方面都很像，除了社交的数量和通勤的便利性方面。为了比较这两份工作和现在的工作，你制作了以下表格：

	社交联系	每日通勤时间
现在的工作	长时间独自一人	10 分钟
工作 x	与他人有限的接触	20 分钟
工作 y	适度社交	60 分钟

此问题的第二个版本包含相同的选项 x 和 y ，但参考工作（ s' ）不同，其描述如下：“非常愉快的社交互动和80分钟的每日通勤时间。”

在第一个版本中，两个选项在社交联系的维度上优于当前的参考工作，但在通勤时间方面都比当前的参考工作差。工作 x 和工作 y 中的不同程度的社交联系被评估为优势（收益），而通勤时间则被评估为劣势（损失）。这些关系在第二个版本中是相反的。损失厌恶意味着当两个选项之间给定的差异被评估为两个损失（劣势）之间的差异而不是两个收益（优势）之间的差异时，通常会产生更大的影响。这一猜想得到了证实：在版本1中70%的被试选择了工作 x ，而在版本2中只有33%的被试选择了工作 x （ $N=106$ ， $p<0.01$ ）。

参考依赖

为了解释由参考变换引起的偏好逆转，我们引入了一个原始概念，即反映给定参考状态的偏好关系。与标准理论一样，我们从选择集合 $X=\{x, y, z, \dots\}$ 开始，并且为了简单起见，假设它与现实平面的正象限（包括其边界）同构。 X 中的每个选项 $x=(x_1, x_2)$ ， $x_1, x_2 \geq 0$ 可以被解释为一个捆绑，它提供 x_1 单位的商品1和 x_2 单位的商品2，或者是一种以两个价值维度上的水平为特征的活动。扩展到两个以上的维度是很简单的。

参考结构是一组索引偏好关系，其中 $x \succeq_r y$ 从参考状态 r 的角度被解释为 x 弱优于 y 。关系 $>_r$ 和 $=_r$ 分别对应于严格偏好和无差异偏好。在整篇文章中，我们假设每个 $\succeq_r$ ， $r \in X$ ，满足经典理论的标准假设。具体来说，我们假设 $\succeq_r$ 是完整的、传递性的和连续的；也就是说， $\{x: x \succeq_r y\}$ 和 $\{x: y \succeq_r x\}$ 对任何 y 都是接近的。此外， $x \succeq_r y$ 和 $x \neq y$ 意味着 $x >_r y$ ，从这个意义上

讲，每个偏好顺序是严格单调的。在这些假设下，每个 $\geq_r$ 可以由严格增加的连续效用函数 U_r 表示（Varian, 1984）。

由于标准理论没有关注参考状态的特殊作用，所以它隐含地假设了参考的独立性；也就是说，对于所有 $x, y, r, s \in X$ ，当且仅当 $x \geq_s y$ 时， $x \geq_r y$ 。然而，在前面的实验中，这种性质一直是被违背的。为了适应这些观察，在描述个人选择时，我们不用单一的偏好顺序，而是用一个家庭或一本索引的偏好顺序 $\{\geq_r: r \in X\}$ 。为方便起见，我们使用字母 r, s 来表示参考状态，用 x, y 表示选项，尽管它们都是 X 中的元素。

参考依赖性选择的研究提出了两个问题：什么是参考状态？以及它是如何影响偏好的？本分析集中在第二个问题。我们假设决策者在 X 中有一个明确的参考状态，并且针对它对选项之间选择的影响进行了调查。参考状态的起源和决定因素的问题不在本文的范围内。虽然参考状态通常与决策者的当前位置相对应，但它也可能受到志向、前景、规范和社会比较的影响（Easterlin, 1974; van Praag, 1971; van de Stadt, Kapteyn & van de Geer, 1985）。

在本节中，我们首先根据偏好顺序 $\geq_r, r \in X$ 定义了损失厌恶和敏感度递减。接下来我们将介绍可分解参考函数的概念，并描述恒定损失厌恶的概念。最后，我们会讨论损失厌恶系数的一些经验估计。

损失厌恶

关于损失厌恶的基本直觉是损失（低于参考状态的结果）比相应的增益（高于参考状态的结果）的影响程度更大。参考标准的转移可以将收益转化为损失，反之亦然；它可以引起偏好的逆转。如下面的定义所示。

如果以下条件适用于 X 中的所有 x, y, r, s ，那么参考结构满足损失厌恶。假设 $x_1 \geq r_1 > s_1 = y_1, y_2 > x_2$ 且 $r_2 = s_2$ ，如图8-3所示。那么 $x =_s y$ 意味着 $x >_r y$ ；如果下标1和2在整个过程中互换，那么同样成立（注意，关系 $>$ 和 $=$ 是指选项的数值分量；而 $>_r$ 和 $=_r$ 是指参考状态 r 中选项之间的偏好）。损失厌恶意味着当从 r 对 y 进行评估时，通过 y 的无差异曲线斜率

比从s评估时更陡。换句话说， $U_r^*(y) > U_s^*(y)$ ，其

$$U_r^*(y)$$

中

是 U_r 在y处的边际替代率。

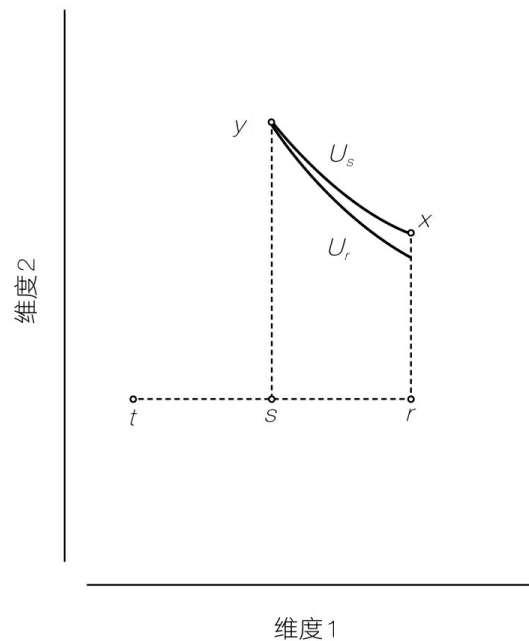


图8-3 损失厌恶示意图

为了推动对损失厌恶进行定义，有必要根据相对于参考点 r 的优势和劣势对其进行重新表述。如果 $x_i > r_i$ 或 $x_i < r_i$ ，那么有序对 $[x_i, r_i]$ ， $i=1, 2$ ，分别被定义为优势或劣势。我们使用括号来区分有序对 $[x_i, r_i]$ 和二维选项 (x_1, x_2) 。假设存在实值函数 v_1 、 v_2 使得 $U_r(x)$ 可以表示为 $U(v_1[x_1, r_1], v_2[x_2, r_2])$ 。

为了简化问题，假设 $x_1=r_1$ 且 $x_2>r_2$ ，如图8-3所示。因此， $x=y$ 意味着相对于参考状态 s ， $[x_1, s_1]$ 和 $[x_2, s_2]$ 这两个优势的组与优势 $[y_2, s_2]$ 和零间隔 $[y_1, y_1]$ 的组合有相同影响。类似地， $x>_ry$ 意味着优

势 $[x_2, r_2]$ 和零间隔 $[x_1, x_1]$ 的组合比优势 $[y_2, r_2]$ 和劣势 $[y_1, r_1]$ 的组合影响更大。因此，当参考状态从 s 变为 r 时，劣势 $[y_1, r_1] = [s_1, r_1]$ 成为 y 的评估的一部分，并且优势 $[x_1, s_1] = [r_1, s_1]$ 会被从对 x 的评估中删除。但由于 $[s_1, r_1]$ 和 $[r_1, s_1]$ 仅仅是符号不同，因此损失厌恶意味着引入劣势比删除相应的优势影响更大。类似的论证适用于 $x_1 > r_1 > s_1$ 的情况。

目前的损失厌恶概念解释了前一节中描述的禀赋效应和现状偏好。考虑不同参考点对 x 和 y 之间偏好的影响（如图8-2所示）。根据损失厌恶，一个从 t 角度在 x 和 y 之间中立的决策者，从 x 角度来说更喜欢 x 而不是 y ，从 y 角度来说更喜欢 y 而不是 x 。也就是说， $x =_y y$ 意味着 $x >_x y$ 且 $y >_y x$ 。这解释了卖方和选择者对商品的不同估值，以及现状偏好的其他表现形式。

敏感度递减

回想一下，根据图8-1的效用函数，边际价值随着距参考点的距离的增加而减小。例如，年薪60 000美元和年薪70 000美元之间的差距在当前年薪为50 000美元时比在年薪为40 000美元时的影响更大。如果以下条件适用于 X 中的所有 x, y, s, t ，那么参考结构满足敏感度递减（DS）。假设 $x_1 > y_1, y_2 > x_2, s_2 = t_2$ ，并且 $y_1 \geq s_1 \geq t_1$ 或 $t_1 \geq s_1 \geq x_1$ （如图8-3所示）。那么 $y =_s x$ 意味着 $y \geq_t x$ ；如果下标1和2在整个过程中互换，那么同样成立。如果相同的假设意味着 $y =_t x$ ，那么满足恒定的敏感度。根据DS，当参考点距离较远时，对给定维度上的差异的敏感度小于参考点距离较近时的敏感度。根据DS，从 s 评估时，通过 x 的无差异曲线的斜率

比从 t 评估时更陡，或者是 $U_s^*(x) > U_t^*(x)$ 。重要的是要区分目前有关参考状态影响的敏感度递减的概念和边际效用递减的标准假设。虽然这两个假设在概念上相似，但它们在逻辑上是相互独立的。特别是，敏感度递减并不意味着无差异曲线在参考点以下是凹的。

每个参考状态 r 将 X 划分为以 r 为原点的四个象限。只要 $x_i \geq r_i$ ，当且仅当 $y_i \geq r_i, i=1, 2$ ，则一对选项 x 和 y 属于 r 的同一象限。每当 x 和 y 属于关于 r 和 s 的相同象限，以及 r 和 s 属于关于 x 和 y 的相同象限时，如果所有在 X 中

的 x, y, r, s 都符合 $x \geq_r y$ ，当且仅当 $x \geq_s y$ ，那么参考结构满足符号相关性。这种情况意味着只有当参考变化将增益转化为损失或将损失转化为增益时，才会违反参考独立性。验证符号依赖性等同于恒定敏感度是很容易的。虽然符号依赖性通常不成立，但只要参考状态引起的曲率不是很明显，它就可以作为有用的近似值。

敏感度递减（或恒定）的假设允许我们将损失厌恶的含义扩展到在任一维度上与 x 或 y 不一致的参考状态。考虑图8-4中 x 和 y 之间的选择。请注意， r 由 x 支配而不是由 y 支配，而 s 由 y 支配而不是由 x 支配。令 t 是 r 和 s 的集合，也就是说， $t_i = \min(r_i, s_i)$ ， $i=1, 2$ 。它从损失厌恶和敏感度递减中产生，如果 $x=y$ ，那么 $x >_r y$ 和 $y >_s x$ 。因此，比起从 s 评估，从 r 评估时 x 比 y 更可能被选择。我们之前的观察结果证实了这一主张，即当一件礼物在某一属性上得到适度改善时，比其在某一属性上得到较大改善和损失的组合更有吸引力。

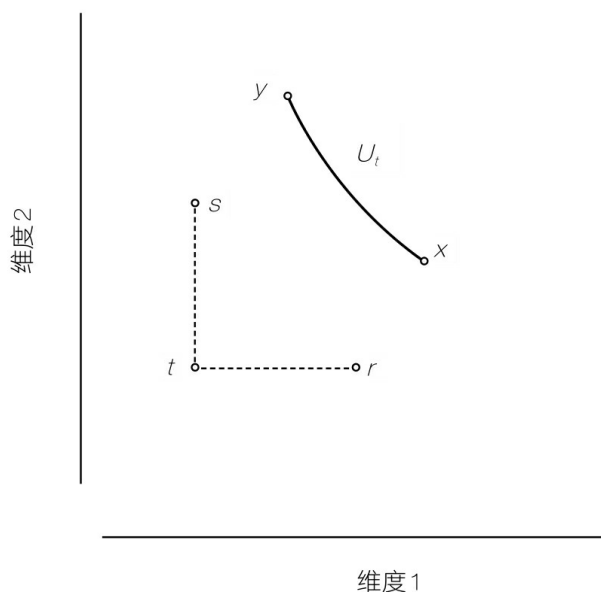


图8-4 参考依赖偏好示意图

考虑两个可交换的个体（即享乐主义双胞胎），他们每个人的职位都是 t ，地位和工资都较低（如图8-4所示）。假设两者在职位 x （非常高的地位、中等工资）和职位 y （非常高的工资、中等地位）之间无差异。想象一下，现在两个人都换到了新的职位，这成为他们各自的参考点；一个人换到 r （高地位、低工资），另一个人换到 s （高工资、低地位）。损失厌恶和敏感度递减表明换到 r 的人现在更喜欢 x ，而换到 s 的人现在更喜欢 y ，因为工资或地位都是他们不愿意放弃的。

恒定损失厌恶

本节介绍了限制从不同参考点评估的偏好顺序之间的关系的附加假设。如果存在每个参数都增加的实值函数 U ，那么参考结构 $(X, \geq_r)$ ， $r \in X$ 是可分解的，使得对于每个 $r \in X$ ，存在递增函数 $R_i: X_i \rightarrow$ 实数， $i=1, 2$ ，满足

$$U_r(x_1, x_2) = U(R_1(x_1), R_2(x_2))$$

函数 R_i 被称为与参考状态 r 相关联的参考函数。在该模型中，参考点的效应通过两个坐标轴的单独的单调变换来确定。可分解性具有可检验的特性。例如，假设 U_r 是可加的，也就是说， $U_r(x_1, x_2) = R_1(x_1) + R_2(x_2)$ 。因此，对于任何 $s \in X$ ， U_s 也是可加的，尽管各自的尺度可能不是线性相关的。

在本节中，我们将重点讨论可分解性的一个特殊情况，在这个案例中参考函数呈现为一种特别简单的形式。如果存在函数 $u_i: X_i \rightarrow$ 实数，常数 $\lambda_i > 0$ ， $i=1, 2$ 和函数 U 使得 $U_r(x_1, x_2) = U(R_1(x_1), R_2(x_2))$ ，那么参考结构 $(X, \geq_r)$ 满足恒定损失厌恶。其中

$$R_i(x_i) = \begin{cases} u_i(x_i) - u_i(r_i) & \text{当 } x_i \geq r_i \text{ 时} \\ (u_i(x_i) - u_i(r_i)) / \lambda_i & \text{当 } x_i < r_i \text{ 时} \end{cases}$$

因此，我们用两个常数 λ_1 和 λ_2 来描述参考位移引起的偏好顺序的变化，它们可以分别被解释为维度1和维度2上的损失厌恶系数。图8-5给出了恒定损失厌恶，其中 $\lambda_1=2$ 且 $\lambda_2=3$ 。为简单起见，我们选择了线性效用函数，但这并不是必需的。

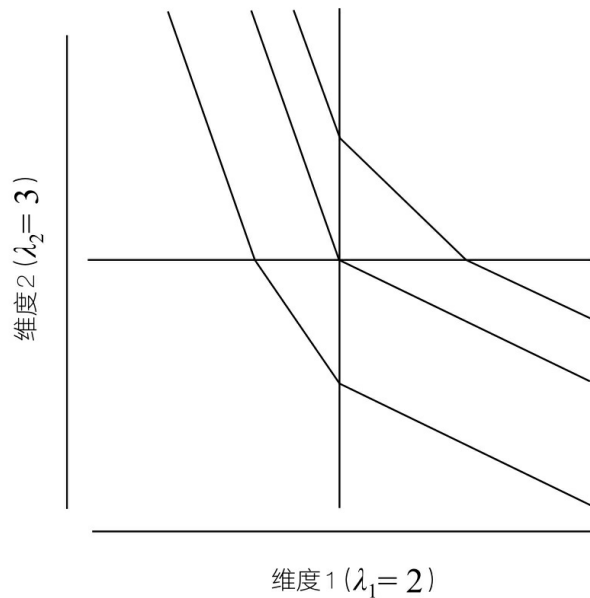


图8-5 一组说明恒定损失厌恶的无差异曲线

虽然我们没有一个关于一般情况下恒定损失厌恶的公理性表征，但我们在下文中描述了 U 可加的特殊情况，这被称为可加性恒定损失厌恶。这种情况很重要，因为在许多情况下，可加性是一个很好的近似值。实际上，一些常用的效用函数（如柯布-效用道格拉斯函数或CES）都是可加的。回想一下，如果坐标轴可以单调变换，使无差异曲线成为平行线，那么无差异曲线族就是可加的。以下的消除条件，也称为汤姆森条件，在目前的情况下是可加性的充分必要条件（Debreu, 1960; Krantz, Luce, Suppes & Tversky, 1971）。

对于所有的 $x_1, y_1, z_1 \in X_1$; $x_2, y_2, z_2 \in X_2$ 和 $r \in X$ ，当 $(x_1, z_2) \geq_r (z_1, y_2)$ 且 $(z_1, x_2) \geq_r (y_1, z_2)$ 时，则 $(x_1, x_2) \geq_r (y_1, y_2)$ 。

假设取消每个 $\geq_r$ 的值，我们获得了每个参考状态的可加性表示。为了将单独的加法表示为相互关联，我们引入了以下公理。考虑 X 中的 $w, w', x, x', y, y', z, z'$ 属于 r 和 s 的相同象限，并且满足 $w_1=w'_1, x_1=x'_1, y_1=y'_1, z_1=z'_1$ 及 $x_2=z_2, w_2=y_2, x'_2=z'_2, w'_2=y'_2$ ；如图8-6所示。根据上述假设，如果 $w=_r x, y=_r z$ 且 $w'=_s x'$ 意味着 $y'=_s z'$ ，那么参考结构 $(X, \geq_r)$ ， $r \in X$ 满足参考关联（reference interlocking）。基本上，特沃斯基、萨塔特和斯洛维克在偏好逆转的处理中都引用了相同的条件（Tversky, Sattath & Slovic, 1988），彼得·韦克尔（Peter Wakker）、特沃斯基和卡尼曼在不确定状况下的决策分析中引用了相同

的条件（Tversky & Kahneman, 1991; Wakker, 1989）。

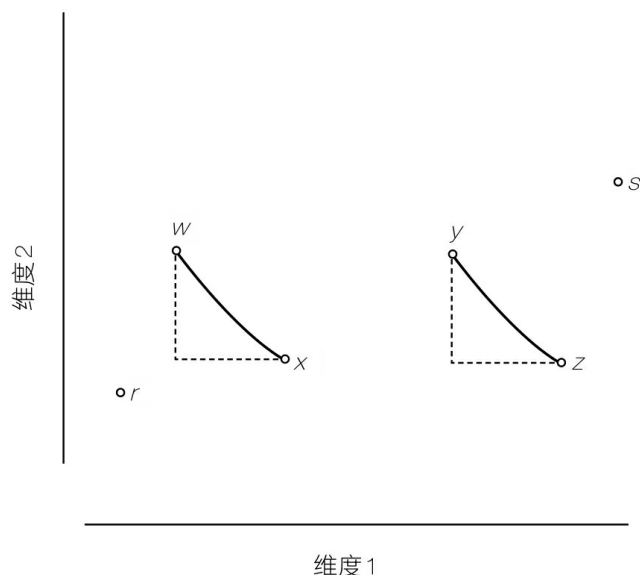


图8-6 参考关联的示意图

为了理解参考关联的内容，请注意在存在可加性的情况下，无差异可以被解释为一维上的区间与第二维上的区间的匹配。例如，观察值 $w \sim x$ 表示在第一维上的区间 $[x_1, w_1]$ 与第二维上的区间 $[w_2, x_2]$ 匹配。类似地， $y \sim z$ 表示 $[z_1, y_1]$ 与 $[y_2, z_2]$ 相匹配。但由于 $[w_2, x_2]$ 和 $[y_2, z_2]$ 在结构上是相同的（如图8-6所示），我们由此推断 $[x_1, w_1]$ 与 $[z_1, y_1]$ 相匹配。通过这种方式，我们可以将同一维度上的两个区间匹配为另一个维度上的一个区间。根据参考关联，如果两个维度区间被匹配为增益，那么它们也可以被匹配为损失。验证参考关联是否来自附加的恒定损失厌恶很容易。此外，以下定理表明，在存在抵消和符号依赖性的情况下，参考关联对于可加的恒定损失厌恶是必要且充分的。

定理：当且仅当参考结构 $(X \geq r)$ ， $r \in X$ ，满足消除性、符号依赖和参考关联时，它才满足可加的恒定损失厌恶。

该定理的证明见附录。损失厌恶系数的估计可以从上文中描述的实验中得出，实验的两组被试为同一消费品标出了货币价值：得到这一物品的卖家可以选择出售商品，而选择者可以选择接受这一物品或一笔钱（Kahneman, Knetsch & Thaler, 1990）。在两次独立的重复实验中，卖家对杯子估价的中位数分别为7.12美元和7.00美元；选择者对杯子的

估价为3.12美元和3.50美元。根据目前的分析，卖家和选择者的区别仅在于前者将杯子评估为损失，后者则将其评估为收益。如果货币价值在该范围内是线性的，那么在這些实验中杯子的损失厌恶系数略大于2。

这种对损失厌恶系数的估计与基于风险决策的估计之间有一个有趣的趋同之处。这样的估计可以通过观察G / L的值来获得，该比率使得收益或损失的机会均等，刚好可以被接受。我们在几次实验中观察到的比率略高于2:1。例如，在一次有实际收益的赌博实验中，赢得25美元或输掉10美元的概率均等的赌局是勉强可接受的，其产生的比率为2.5:1。类似的值来自关于较大赌博的可接受性的假设选择，范围超过几百美元（Tversky & Kahneman, 1990）。尽管对估计值的趋同的解释应谨慎，但这些研究结果表明，约为2的损失厌恶系数也许可以解释涉及货币结果和消费品的风险选择和无风险选择。

回想一下，损失厌恶系数可能会因维度而异（如图8-5所示）。我们推测，不同维度的损失厌恶系数反映了这些维度的重要性或显著性（Tversky, Sattath & Slovic, 1988）。例如，损失厌恶似乎在安全方面比在金钱方面更明显，在收入方面比在休闲方面更明显（Viscusi, Magat & Huber, 1987）。

损失厌恶的含义

损失厌恶是近年来一直在被讨论的一种现象的一个重要组成部分：我们经常能观察到人们愿意接受的放弃他们拥有的商品的最小数量（WTA）和他们愿意支付的最大数量（WTP）之间的巨大差异。这种差异的其他潜在来源还包括收入效应、战略行为和交易的合法性。买卖差异最初出现在涉及公共产品的假设性问题中，但它也已经在实际交易中得到了证实（Cummings, Brookshire & Schulze, 1986; Heberlein & Bishop, 1985; Kahneman, Knetsch & Thaler, 1990; Loewenstein, 1988）。在试图通过市场经验法则消除它的实验中，它也存活了下来，尽管有所减少（Brookshire & Coursey, 1987; Coursey, Hovis & Schulze, 1987; Knetsch & Sinden, 1984, 1987）。卡尼曼、尼奇和塞勒的研究表明，所有者和潜在买家对消费品的不同估值会抑制贸易的进行（Kahneman, Knetsch & Thaler, 1990）。他们为一半的被试提供了消费品（如一个杯子）并为这种商品建立了一个市场体系。由于杯子是随机分配的，因此标准理论预测，一半的卖家应该将杯子卖给更重视它们的买家。持续观察发现，实际交易量约为预测量的一半。在控制实验

中，被试用可兑换现金的代币进行交易，产生了近乎完美的效益，买方和卖方分配的价值之间没有差异。

交易涉及两个维度，而损失厌恶可能在其中一个维度上运行，也可能同时在两个维度上运行。因此，目前的分析表明，损失厌恶可能通过两种方式使WTA和WTP之间产生差距。声称愿意接受一个产品的WTA的人考虑放弃它；声称愿意购买的人考虑采用WTP。如果该物品存在损失厌恶，所有者将不愿意出售。如果买家认为花在购买上的钱是一种损失，他们就不愿意购买。通过将买卖双方与选择者进行比较，可以估计出这两种效应的相对程度。选择者可以在商品和现金之间做出选择，因此不容易受到损失厌恶的影响。几次比较的结果表明，不愿意出售的程度远远大于不愿意购买的程度（Kahneman, Knetsch & Thaler, 1990）。这些市场中的买家似乎并不认为他们在交易中放弃的钱是一种损失。这些观察结果与消费者选择的标准理论是一致的，在该理论中，是否购买一件商品的决定被视为在该商品和其他可以购买的商品之间的选择。

在5美元与5美元的交易中当然不涉及损失厌恶，因为交易是根据净值来得以评估的。同样，在日常的商业交易中，不愿意出售的情况肯定是不存在的，在这些交易中，待售商品具有代币的性质。然而，目前的分析表明，相对于先前交易中建立的参考水平，收益和损失的不对称评估将影响买卖双方对利润值变化的反应（Kahneman, Knetsch & Thaler, 1986; Winer, 1986）。当变化是不利的（损失）时，人们对变化的反应会比变化是有利的时候更加强烈。D. S.普特勒（D. S. Putler）对需求进行了分析，其中包含了价格上涨和下跌的不对称效应

（Putler, 1988）。相对于之前一系列价格估算的参考价格，他通过估算鸡蛋零售价格上涨和下跌的弹性需求来检验该模型。价格上涨的弹性系数为-1.10，价格下跌的弹性系数为-0.45，这表明价格上涨对消费者决策的影响更大。该分析假设，替代品的可用性消除了人们对鸡蛋消费减少的厌恶。在咖啡市场的扫描面板数据中也观察到了类似的结果

（Kalwani, Yim, Rinne & Sugita, 1990）。不愿接受损失也可能会影响卖方：一项针对股市的研究表明，价格上涨时的交易量往往高于价格下跌时的交易量（Shefrin & Statman, 1985）。

损失厌恶会使谈判复杂化。实验证据表明，当谈判者讨价还价的属性被定义为损失时，他们达成协议的可能性比被定义为收益时要小

（Bazerman & Carroll, 1987）。如果人们对负域的边际变化更敏感，那

么这个结果是可以预料的。此外，对一方做出的让步和另一方做出的让步的评价之间存在着一种天然的非对称性；后者通常被评估为收益，而前者被评估为损失。对让步的差异性评估显著减少了多议题谈判中的协议区域。

人们对价格或利润的有利或不利变化的反应存在明显的不对称现象，这一现象在一项研究中被注意到了，该研究对决定价格或工资的行为的公平判断规则进行了规范（Kahneman, Knetsch & Thaler, 1986）。特别是，大多数人拒绝接受是因为无法通过成本增加来证明明显不公平的价格上涨，以及无法通过破产威胁来证明工资削减。此外，经济公平的常用规范并不绝对要求公司与客户或其员工分享降低成本或增加利润的好处。与不区分损失和放弃收益的经济分析相反，公平的标准在将损失强加于他人的行为和不共享利益的行为（或不采取行动）之间做出了明显的区分。一项关于法院判决的研究记录了在处理损失和放弃收益方面的类似区别；例如，在疏忽的情况下，自付费用比未实现的利润更有可能获得补偿（Cohen & Knetsch, 1990）。

由于被视为不公平的行为往往受到抵制和惩罚，因此对公平的考虑被引用为工资粘性的解释之一，以及市场明显缓慢的其他案例之一

（Kahneman, Knetsch & Thaler, 1986; Okun, 1981; Olmstead & Rhode, 1985）。例如，对损失和放弃收益的评价的不同，意味着对减薪和在可能增加工资时未能增加工资的反应也有相应的不同。以前的合同条款规定了集体和个人谈判的参考水平；在谈判的背景下，对损失的厌恶形式表现为对让步的厌恶。损失厌恶引起的僵化可能导致效率低下的劳动合同，这种合同无法对不断变化的经济环境和技术发展做出充分的反应。因此，新公司可以在不受之前协议影响的情况下与员工讨价还价，从而获得竞争优势。

损失厌恶是不合理的吗？这个问题引出了一些难以解决的规范性问题。对决策者赋予结果的价值提出质疑需要一个评估偏好的标准。结果的实际经验为我们提供了这样一个标准：在决策环境中赋予结果的价值可以被证明是对该结果经验质量的预测（Kahneman & Snell, 1990）。根据这种预测立场，图8-1显示的效用函数（最初是用来解释风险选择的模式）可以被解释为对享乐体验的心理物理学的预测。效用函数恰当地反映了3个基本事实：有机体习惯于稳定状态，对变化的边际反应正在减弱，而痛苦比快乐更为紧迫。痛苦和快乐的非对称性是在选择中出现损失厌恶的最终理由。由于这种非对称性，一个寻求最大化结果的经

验效用的决策者最好将更多的权重分配给消极的结果，而不是积极的结果。

本文第一部分的论证比较了相同的两个客观状态之间的选择，这是从不同的参考点出发进行评估的。参考水平对决策的影响只能通过这些参考水平对结果性经验的相应影响来证明。例如，如果能更敏锐地体验到任何改变的劣势而不是优势，那么支持现状的偏好就是合理的。但是，在决策环境中自然采用的一些参考水平与之后的结果性经验无关，这些参考水平对决策的影响通常是可疑的。例如，在评估具有长期影响的决策时，如果适应最终会引起参考的转变，那么对这些结果的初始反应可能相对不重要。另一个案例涉及委托代理关系：委托人可能不希望代理人的决定反映出代理人对损失的厌恶，因为代理人的参考水平与委托人的结果性经验无关。我们的结论是，关于损失厌恶或其他参考效应的规范性地位，没有一个普遍的答案，但有一种原则性的方法可用来检查这些效应在特定情况下的规范性地位。

附录

定理。当且仅当参考结构满足取消、符号依赖和参考关联时，参考结构 $(X, \geq_r)$ ， $r \in X$ 满足可加的恒定损失厌恶。

证明。证明必要性很简单。为了证明充分性，请注意在目前的假设下，取消意味着可加性（Debreu, 1960; Krantz et al., 1971）。因此，对于任何 $r \in X$ ，存在连续函数 $R_i: X_i \rightarrow \text{实数}$ ，它是唯一的一个正线性位移，使得 $R(x) = R_1(x_1) + R_2(x_2)$ 代表 $\geq_r$ 。也就是说，对于任何 $x, y \in X$ ，当且仅当 $R(x) \geq R(y)$ 时， $x \geq_r y$ 。接下来我们证明以下两个引理。

引理1。设 A 是一组属于 r 和 s 的相同象限的选项。存在 $\lambda_i > 0$ ，使得对于 A 中的所有 x, y ：

$$R_i(y_i) - R_i(x_i) = (S_i(y_i) - S_i(x_i)) / \lambda_i, \quad i=1, 2$$

证明。我们希望表明对于所有 $r, s, w, x, y, z \in X$

$$R_i(z_i) - R_i(y_i) = R_i(x_i) - R_i(w_i)$$

即

$$S_i(z_i) - S_i(y_i) = S_i(x_i) - S_i(w_i), \quad i=1, 2$$

每当所讨论的*i*区间可以通过另一维度上的区间匹配时，该命题遵循连续性、可加性和参考关联。如果不能进行这种匹配，我们通过连续性将这些*i*区间划分为足够小的子区间，这些子区间可以通过另一维度上的区间进行匹配。因为 R_i 差异的相等意味着 S_i 差异的相等，所以引理1遵循连续性和可加性。

引理2。假设 $r, s \in X$, $s_i < r_i$ 且 $s_2 = r_2$ 。设 S 代表 $\geq_s$ ，满足 $S_1(s_1) = 0$ 。如果保持符号依赖和参考关联，那么存在 $\lambda_1 > 0$, $\lambda_2 = 1$ ，使得

$$R^*(x) = R_1^*(x_1) + R_2^*(x_2) \text{ 表示 } \geq_r, \text{ 其中}$$

$$R_1^*(x_1) = \begin{cases} S_1(x_1) - S_1(r_1) & \text{当 } x_1 \geq r_1 \\ S_1(x_1) - S_1(r_1) / \lambda_1 & \text{当 } s_1 \leq x_1 \leq r_1 \\ S_1(x_1) - S_1(r) / \lambda_1 & \text{当 } x_1 \leq r_1 \end{cases}$$

及 $R_2^*(x_2) = S_2(x_2) - S_2(r_2) / \lambda_2$ 。如果指数1和指数2在整个过程中互换，那么同样成立。

证明。符号依赖性 $\geq_r$ 和 $\geq_s$ 对 $\{x \in X: x_1 \geq r_1, x_2 \geq r_2\}$ 和 $\{x \in X: x_1 \geq r_1, x_2 \leq x_2\}$ 中的所有元素对一致。为了证明 $\geq_r$ 和 $\geq_s$ 也在它们的集合上一致，假设 y 属于前者， z 属于后者。足以证明 $y =_r z$ 意味着 $y =_s z$ 。通过单调性和连续性，存在 w 使得 $y =_r w =_r z$ 且 $w_2 = r_2 = s_2$ 。由于 w 属于两组的交集， $y =_r w$ 意味着 $y =_s w$ ，且 $z =_r w$ 意味着 $z =_s w$ ，因此 $y =_s z$ 。

所以，我们可以选择相应的尺度，使得 $R_i = S_i$, $i=1, 2$ ，在 $\{x \in X: x_1 \geq r_1\}$ 中。接下来，我们分析表明 $R^*(x) + S(r) = R(x)$ 。我们分别考虑了每个维度。对于 $i=2$ ， $R_2^*(x_2) + S_2(r_2) = S_2(x_2)$ 。我们证明了 $S_2(x_2) = R_2(x_2)$ 。选择 $x_1 \geq r_1$ 。通过构造 $S(x) = R(x)$ ，因此 $S_2(x_2) = R_2(x_2)$ 。

对于 $i=1$ ，如果 $x_1 \geq r_1$ ，通过构造我们得到

$$R_1^*(x_1) + S_1(r_1) = S_1(x_1), \text{ 和 } R_1(x_1)$$

$$R_1^*(x_1)$$

$$= S_1(x_1)。因此, \quad + S_1(r_1) \\ = S_1(x_1) = R_1(x_1)。$$

对于 $s_1 < x_1 < r_1$ ，我们希望证明存在 λ_1 ，使得

$$R_1(x_1) = S_1(x_1) + (S_1(x_1) - S_1(r_1)) / \lambda_1$$

或者

$$R_1(x_1) - R_1(r_1) = (S_1(x_1) - S_1(r_1)) / \lambda_1$$

这些由引理1推断得出。

对于 $x_1 \leq s_1$ ， $\geq_r$ 和 $\geq_s$ 一致，通过符号依赖可得 $R_1 = \alpha S_1 + \beta$ ， $\alpha > 0$ 。因为 $R_2 = S_2$ ， $\alpha = 1$ ，并且因为 $S_1(s_1) = 0$ ， $\beta = R_1(s_1)$ ，因此 $R_1(x_1) = S_1(x_1) + R_1(s_1)$ 。所以

$$\begin{aligned} R_1(x_1) - R_1^*(x_1) &= S_1(x_1) + R_1(s_1) - (S_1(x_1) - S_1(r_1)/\lambda_1) \\ &= R_1(s_1) + S_1(r_1)/\lambda_1 \end{aligned}$$

这足以证明该表达式等于 $S_1(r_1)$ 。考虑 $s_1 < x_1 < r_1$ 的情况，通过 s_1 的

连续性可得

$$R_1(s_1) - R_1(r_1) = (S_1(s_1) - S_1(r_1)) / \lambda_1$$

因此

$$R_1(s_1) + S_1(r_1) / \lambda_1 = R_1(r_1) = S_1(r_1)$$

至此我们就完成了引理2的证明。

接下来，我们证明 λ_i ($i=1, 2$) 与 r 无关。选择 $r, s, t \in X$ 使得 $r_2=s_2=t_2$ 和 $s_1 < r_1 < t_1$ 。根据之前的定理可知存在 R^* 和 T^* ，其由 S 定义，常数分别为 $\lambda^{(r)}$ 和 $\lambda^{(t)}$ 。因为 $\geq_r$ 和 $\geq_t$ 在 $\{x \in X: x_1 \leq r_1\}$ 上一致，所以根据符号依赖性， $\lambda^{(r)} = \lambda^{(t)}$ 。当指数1和指数2互换且 $r_1 < s_1$ 时，同样适用。

为了证明一般情况下的充分性，考虑 $r, s \in X$ ，其中 $r_1 > s_1$ ， $r_2 \leq s_2$ ， $t = (r_1, s_2)$ 。通过应用两次先前的（一维）结构，一次用于 (s, t) ，另一次用于 (t, r) ，我们得到了预期的结果。

致谢

感谢肯尼斯·阿罗（Kenneth Arrow）、彼得·戴蒙德（Peter Diamond）、戴维·克兰茨、马修·拉宾（Matthew Rabin）和理查德·泽克豪泽的评论。我们要特别感谢什穆埃尔·萨塔特和彼得·韦克尔（Peter Wakker）提出的有益建议。

参考文献

Bazerman, M., & Carroll, J. S. (1987). Negotiator cognition. In B. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. 9, pp. 247–288). Greenwich, CT: JAI Press.

Brookshire, D. S., & Coursey, D. L. (1987). Measuring the value of a public good: An empirical of elicitation procedures. *American Economic Review*, 77, 554–566.

Cohen, D., & Knetsch, J. L. (1990). Judicial choice and disparities between measures of economic values. Working paper. Burnaby, BC: Simon Fraser University.

Coursey, D. L., Hovis, J. L., & Schulze, W. D. (1987). The disparity between willingness to accept and willingness to pay measures of value. *Quarterly Journal of Economics*, 102, 679–690.

Cummings, R. G., Brookshire, D. S., & Schulze, W. D. (Eds.). (1986). *Valuing environmental goods*. Totowa, NJ: Rowman and Allanheld.

Debreu, G. (1960). Topological methods in cardinal utility theory. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences* (pp. 16–26). Stanford, CA: Stanford University Press.

Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. In P. A. David & M. W. Reder (Eds.), *Nations and households in economic growth* (pp. 89–125). New York, NY: Academic Press.

Heberlein, T. A., & Bishop, R. C. (1985). Assessing the validity of contingent valuation: Three field experiments. Paper presented at the International Conference on Man's Role in Changing the Global Environment, Venice, Italy.

Kahneman, D., Knetsch, J. L., & Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *American Economic Review*, 76, 728–741.

Kahneman, D., Knetsch, J. L., & Thaler, R. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy*, 98, 1325–1348.

Kahneman, D., & Snell, J. (1990). Predicting utility. In R. Hogarth (Ed.), *Insights in decision making* (pp. 295–310). Chicago, IL: University of Chicago Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of

decision under risk. *Econometrica*, 47, 263–291.

Kahneman, D., & Tversky, A. (1984). Choices, values and frames. *American Psychologist*, 39, 341–350.

Kalwani, M. U., Yim, C. K., Rinne, H. J., & Sugita, Y. (1990). A price expectations model of customer brand choice. *JMR, Journal of Marketing Research*, 27, 251–262.

Knetsch, J. L. (1989). The endowment effect and evidence of nonreversible indifference curves. *American Economic Review*, 79, 1277–1284.

Knetsch, J. L., & Sinden, J. A. (1984). Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. *Quarterly Journal of Economics*, 99, 507–521.

Knetsch, J. L., & Sinden, J. A. (1987). The persistence of evaluation disparities. *Quarterly Journal of Economics*, 102, 691–695.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. I). New York, NY: Academic Press.

Loewenstein, G. (1988). Frames of mind in intertemporal choice. *Management Science*, 34, 200–214.

Okun, A. (1981). *Prices and quantities: A macroeconomic analysis*. Washington, DC: The Brookings Institution.

Olmstead, A. L., & Rhode, P. (1985). Rationing without government: the West Coast gas famine of 1920. *American Economic Review*, 75, 1044–1055.

Putler, D. S. (1988). Reference price effects and consumer behavior. Unpublished manuscript. Economic Research Service, U. S. Department of Agriculture, Washington, DC.

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision

making. *Journal of Risk and Uncertainty*, 1, 7–59.

Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance*, 40, 777–790.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1, 39–60.

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, 95, 371–384.

Tversky, A., & Kahneman, D. (1991). Advances in prospect theory: Cumulative representation of uncertainty. Unpublished manuscript, Stanford University, Stanford, CA.

van de Stadt, H., Kapteyn, A., & van de Geer, S. (1985). The relativity of utility: Evidence from panel data. *Review of Economics and Statistics*, 67, 179–187.

van Praag, B. M. S. (1971). The individual welfare function of income in Belgium: An empirical investigation. *European Economic Review*, 20, 337–369.

Varian, H. R. (1984). *Microeconomic Analysis*. New York, NY: Norton.

Viscusi, W. K., Magat, W. A., & Huber, J. (1987). An investigation of the rationality of consumer valuations of multiple health risks. *Rand Journal of Economics*, 18, 465–479.

Wakker, P. P. (1989). *Additive representations of preferences: A new foundation of decision analysis*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Winer, R. S. (1986). A reference price model of brand choice for frequently purchased products. *Journal of Consumer Research*, 13, 250–256.

第9章 偏好与信念：不确定条件下选择的模糊性与竞争力

奇普·希思
阿莫斯·特沃斯基

我们在现实生活中遇到的不确定性并不容易被量化。我们可能会觉得自己最喜欢的足球队有很好的机会赢得冠军赛、黄金的价格可能会上涨、现任市长不太可能会连任，但通常不愿意为这些事件分配数值概率。然而，为了促进沟通并加强对选择的分析，我们通常需要对不确定性进行量化。量化不确定性最常见的方法是用概率来表示信念。例如，当我们说不确定性事件的概率是30%时，我们认为这个事件与从含有30个红球和70个绿球的盒子中拿出一个红球的可能性一样。衡量主观概率的另一种方法是通过期望效用理论推断偏好的信念程度。这种方法由弗兰克·拉姆齐（Frank Ramsey）开创（Ramsey, 1931），并由伦纳德·萨维奇（Savage, 1954）及F. J.安斯科姆（F. J. Anscombe）和R. J.奥曼（R. J. Aumann）进一步发展（Anscombe & Aumann, 1963），它是从投注之间的偏好中得出主观概率的。具体而言，如果决策者认为在不确定性事件E发生的情况下收到 x 美元（否则没有钱）的情景和“从一个装有红球的比例为 p 的盒子中抽到红球则收到 x 美元”的情景之间并无差异，那么我们可以认为不确定性事件E的主观概率为 p 。

用于衡量信念的拉姆齐方案及其所依据的理论受到了丹尼尔·埃尔斯伯格的挑战，他构建了一个令人信服的演示，被称为模糊效应，尽管“含糊”一词可能更合适（Daniel Ellsberg, 1961; Fellner, 1961）。这种效应最简单的演示涉及2个盒子：一个盒子中有50个红球和50个绿球；而另一个盒子中红球和绿球的总数也为100，但比例不详。你从盒子里随便拿出一个球并猜测它的颜色。如果你的猜测是正确的，就会赢得20美元，否则什么也得不到。你愿意在哪个盒子上投注？埃尔斯伯格认为，即使人们没有颜色偏好，也更愿意在比例相同的盒子上投注而不是在未知比例的盒子上投注，因此在任何一个盒子中投注红色或绿色都无关紧要。这种偏好模式在之后的许多实验中得到了证实，它违反了主观概率的可加性，因为它暗示红色和绿色的概率总和在比例相同的盒子

中比在比例未知的盒子中更高。

埃尔斯伯格的研究引起了人们的极大兴趣，原因有两个。首先，它为机会博弈背景下的（主观）前景效用理论提供了一个有益的反例。其次，它提出了一个普遍的假设，即人们倾向于选择明确而非模糊的事件，至少对于中等概率和大概率事件是这样。埃尔斯伯格认为，对于小概率事件，人们可能更喜欢选择含糊不清的情况。这些观察结果为期望效用理论和其他风险选择模型提出了一个很严肃的问题，因为除了机会博弈之外，现实世界中的大多数决策都依赖于不确定性事件，而这些事件的概率是无法被精确评估的。如果人们的选择不仅取决于不确定的程度，还取决于评估的精确度，那么风险选择的标准模型的适用性就会受到严重限制。实际上，一些作者通过引用非可加性的信念尺度

（Fishburn, 1988; Schmeidler, 1989）或二阶概率分布（Gärdenfors & Sahlin, 1982; Skyrms, 1980）来扩展标准理论，以便纳入模糊效应的影响。这些模型的规范性地位是一个被激烈讨论的话题。一些作者，尤其是埃尔斯伯格认为，在规范的基础上，模糊厌恶是合理的

（Ellsberg, 1963），尽管雷法已经证明它会导致不连贯（Raiffa, 1961）。

埃尔斯伯格的例子，以及随后大多数有关模糊性或含糊性反应的实验研究都局限于随机过程，如从盒子中抽出一个球，或者给决策者提供概率估计的问题。然而，模糊性的潜在重要性源于其与现实世界中证据评估的相关性。模糊厌恶是否局限于机会博弈和陈述概率，或者它是否也适用于判断概率？我们在文献中没有找到这个问题的答案，但有证据表明人们对模糊厌恶的普遍性产生了一些疑问。

例如，D. V.布德斯库（D. V. Budescu）、史蒂文·温伯格（Steven Weinberg）和T.瓦尔斯滕（T. Wallsten）比较了被试给出的投注的现金等价物，其概率分别用数字、图形或口头形式表示（Budescu, Weinberg & Wallsten, 1988）。在图形形式中，概率表示为一个圆的阴影区域。在口头形式中，概率通过诸如“非常可能”或“非常不可能”的表达方式来描述。因为语言和图形形式比数字形式更模糊，所以模糊厌恶意味着对数字形式的偏好。这一预测未得到证实。在所有的3种表示形式中，被试赋予投注的价格大致相同。在不同的实验范例中，J.科恩（J. Cohen）和M.汉塞尔（M. Hansel）以及W.豪厄尔（W. Howell）研究了被试在涉及技能和机会成分的复合投注之间的选择（Howell, 1971; Cohen & Hansel, 1959）。例如，在后一个实验中，被试必须用

飞镖击中目标（其中被试的命中率为75%）以及旋转轮盘，使其落在一个有标记的区域里，该区域占整个区域的40%。成功包括了75%的技能成分和40%的机会成分，总获胜概率为 $0.75 \times 0.4 = 0.3$ 。豪厄尔改变了投注的技能和机会成分，保持了总获胜率不变。因为被试已知机会水平但不知道技能水平，所以模糊厌恶意味着被试会尽可能多地将不确定性转移到投注的机会成分上。相比之下，87%的选择意味着被试对技能的偏好超过了对机会的偏好。科恩和汉塞尔得出了基本相同的结果（Cohen & Hansel, 1959）。

能力假说

前面的观察结果表明，在机会设置中观察到的对模糊性的厌恶（涉及偶然不确定性）并不容易扩展到判断问题中（涉及主观不确定性）。在本文中，我们研究了不确定性偏好的替代解释，即能力假说，它适用于机会和证据问题。我们认为投注不确定性事件的意愿不仅取决于该事件的可能性的估计和该估计的精确度，还取决于一个人的相关背景常识或理解力。更具体地说，我们提出设判断概率为常数，人们更愿意在他们认为自己很了解或胜任的情境中投注，而不是在他们感到不了解或不知情的情境中投注。假定我们在特定情境下的能力感是由我们已经知道的信息与所能知道的信息的对比来决定的。[\(23\)](#)因此，常识、熟悉度和经验可以增强能力感。注意，当决策者无法获得相关信息，特别是当其他人可以获得这些相关信息时，他们的能力感会降低。

对于能力假说，存在着认知和动机两种解释。人们可能从终身经验中意识到，自己在理解的情况下通常比在知识较少的情况下做得更好。这种期望可能会延续到在熟悉环境中获胜的机会不再高于陌生环境的情况中。也许能力假说的主要原因是动机而不是认知。我们认为，除了金钱收益之外，每次投注的结果还包括与结果相关的信用或责任。满足或尴尬的心理收益可能来自自我评价或他人评价。在任何一种情况下，我们认为与结果相关的荣誉和责任取决于成功和失败的原因。在机会领域，成功与失败主要可归因于运气。当一个人进行投注或判断时，情况就不同了。如果决策者对手头的问题理解有限，那么失败将归因于不了解，而成功很可能归因于偶然。相反，如果决策者是“专家”，那么成功可归因于知识，而失败有时可归因于偶然。

我们不想否认，在专家应该知道所有事实的情况下，他们的失败可能比新手更尴尬。然而，在需要有根据的猜测的情况下，专家有时没有

新手那么脆弱，因为他们可以更好地证明自己的投注，即使他们没有获胜。例如，在投注足球比赛的冠军时，自认为是专家的人可以因为正确的预测而获得赞誉，将不正确的预测视为发挥不稳定。此外，对足球知之甚少的人不会因为正确的预测获得太多的赞誉（因为他们是在猜测），而会因预测不正确而受到指责（因为他们不了解）。

因此，当人们成功时，能力或专业知识会帮助他们获得荣誉，当他们失败时，有时也会为他们提供保护，使其免受指责。此外，不了解或无能不仅使人们无法获得成功的信誉，而且会在失败的情况下使其受到指责。因此，我们认为信誉和指责的平衡在一个人进行专业领域的投注中最有利，对偶然事件投注次之，在了解有限的领域投注最不利。本文从知识或能力所导致的表扬和批评的非对称性出发，对能力假说进行了解释。

前面的分析适用于埃尔斯伯格的例子。我们认为，人们不喜欢在未知的盒子上投注，因为盒子中红球和绿球的比例信息原则上是可知的但是他们却并不知道。这些数据的存在使人们感觉自己知识不足、能力较弱，并减少了相应赌注的吸引力。D.弗里希（D. Frisch）和J.巴伦（J. Baron）对埃尔斯伯格的例子给出了密切相关的解释（Frisch & Baron, 1988）。能力假说也与S.柯利（S. Curley）、J. F.耶茨（J. F. Yates）和R.艾布拉姆斯（R. Abrams）的发现一致，即对之后会被展示给其他人的未知盒子中的内容的预测增强了被试对模糊性的厌恶（Curley, Yates & Abrams, 1986）。

本质上相同的分析也适用于对未来而非过去进行投注的偏好。M.罗思博特（M. Rothbart）和N.斯奈德（N. Snyder）要求被试掷骰子，在骰子滚动之前或之后（结果揭示之前）进行投注（Rothbart & Snyder, 1970）。在掷骰子前预测结果的被试比在掷骰子后预测结果的被试（“后测”）对自己的猜测更有信心。前一组也比后一组投入了更多的钱。作者将这种现象归结为神奇式思考或控制错觉，即相信人们可以在骰子滚动之前而不是之后对结果进行一些控制。然而，即使控制错觉没有为我们提供合理的解释，也会在实验中观察到对未来而不是过去事件投注的偏好，如以下问题所示，其中被试在两个投注之间进行选择：

1. 股票是从《华尔街日报》上随机挑选的。请猜测明天是上涨还是下跌。如果猜对了，会赢得 5 美元。
2. 股票是从《华尔街日报》上随机挑选的。请猜测它昨天是涨了还是跌了。你不能看报纸。如果猜对了，会赢得 5 美元。

67%的被试（ $N=184$ ）倾向于在明天的收盘价上下注。在随机抽取的被试中，只有10%的被试真的投注了自己的选择。因为过去与未来不同，它在原则上是可知的，但他们并不知道，所以被试更喜欢在不了解程度相对更低的未来上投注。同样，W.布伦（W. Brun）和K.泰甘（K. Teigan）观察到被试更喜欢在事件发生之前而不是之后猜测掷骰子的结果、孩子的性别或足球比赛的结果（Brun & Teigan, 1990）。大多数被试发现在事件发生之前猜测“如果正确会更令人满意”“如果错了也没那么不舒服”。在预测中，只有未来可以证明你错了；在后测中，你现在可能就是错的。同样的论证适用于埃尔斯伯格的问题。在球的比例相同的盒子中，只有在取球之后猜测才会出错。此外，在球的比例未知的盒子中，如果结果证明盒子中的大多数球都是相反的颜色，那么在取球之前，猜测就可能被证明是错误的。值得注意的是，投注于未来而不是过去事件的偏好无法用模糊性来解释，因为在这些问题中，未来与过去一样都是模棱两可的。

简单的随机事件，例如从具有已知构成的盒子中取球，不会产生模糊；获胜的机会众所周知。如果等概率事件之间的投注偏好是由模糊性决定的，那么人们应该倾向于投注机会，而不是自己的模糊判断（至少对于中等和高概率事件是这样）。相比之下，上面描述的归因分析意味着当人们感觉已很了解并且能够胜任时，会更倾向于投注自己的判断，而不是其他相匹配的随机事件。人们更喜欢投注于他们的技能而不是机会，这一发现证实了这一预测。这与J.马奇（J. March）和Z.沙皮拉（Z. Shapira）的观察结果一致，许多高层管理人员一直在高度不确定的商业主张上投注，而拒绝将商业决策与机会的博弈进行类比（March & Shapira, 1987）。

我们认为，目前的归因分析可以解释关于不确定性偏好的现有证据，无论它们是否涉及模糊性。这包括：第一，在埃尔斯伯格问题中投注已知而非未知盒子的偏好；第二，投注未来而不是过去事件的偏好；第三，投注技能而非机会的偏好。此外，能力假说意味着选择-判断差异，也就是说即使B被判断为至少与A的可能性一样，也倾向于在A而不是B上投注。在下面的一系列实验中，我们检验了能力假说，研究了选

择-判断差异。在实验1中，我们为人们提供了两种选择，即是对常识项目的判断概率投注，还是在相匹配的随机抽奖上投注。在实验2和实验3中，我们通过研究现实中的事件并引发对知识的独立评估来扩展检验。在实验4中，我们根据他们的专业领域对被试进行了分类，并比较了他们在专业领域、非专业领域和随机事件上投注的意愿。最后，在实验5中，我们在不涉及概率判断的定价任务中对能力假说进行了检验。本文的最后一部分讨论了信念与偏好之间的关系。

实验1：投注知识

被试回答了取自两个不同领域的30个知识性问题，如历史、地理或体育。针对每个问题设有4个备选答案，被试首先选择一个答案，然后根据25%（纯猜测）到100%（绝对肯定）的等级评定自己对该答案的自信程度。研究人员向被试详细说明了量表的使用方法和校准的概念。被试在使用该量表时接受了相关指导，因此60%的信心将对应于60%的命中率。他们还被告知，这些评级是赚钱游戏的基础，且不自信和过度自信都会降低收入。

在回答问题并评估信心之后，被试有机会在自己的答案和抽奖中进行选择，抽奖获胜的概率与他们表示出的自信程度相等。例如，对于75%的信心等级，被试可以选择对自己的答案的正确率投注，或对75%的抽奖进行投注，即从装有100个带编号的筹码的袋子里拿出一个编号范围在1~75的筹码。对于一半的问题，抽奖直接等同于信心等级。对于另外一半的问题，被试可以在答案的补充（对“选择的答案以外的答案是正确的”这一命题投注）或信心等级的补充中进行选择。因此，如果被试选择答案A的信心为65%，那么被试可以选择投注剩下的答案B、C或D中的一个的正确性，或者投注 $100\% - 65\% = 35\%$ 的抽奖。

两组被试参与了该实验。第一组（ $N=29$ ）包括心理学学生，他们获得了参与课程的学分。第二组是（ $N=26$ ）从入门经济学课程中招募的，并且他们参与实验可以获得现金收益。为了确定被试的收益，研究人员随机选择了10个问题，并且要求被试对他们之前的选择进行投注。如果被试选择投注自己的答案，若答案正确，他们可以得到1.50美元。如果被试选择在随机抽奖上投注，他们可以从袋子中抽出筹码，如果筹码上的数字落在恰当的范围内，他们可以得到1.50美元。该实验的平均收入约为8.50美元。

有偿的被试在选择答案和评估信心方面比无偿的被试花费了更多时间，因此前者的准确率稍高一些。两组都表现出了过度自信：有偿的被试回答对了47%的问题，他们的平均信心为60%。无偿的被试回答对了43%的问题，他们的平均信心为53%。我们首先描述简单抽奖的结果，补充（析取）抽奖将在后面讨论。

通过绘制有利于判断投注而非抽奖的选择（ C ）的百分比作为判断概率（ P ）的函数来对结果进行总结。在讨论实际数据之前，检查由5个备选假设所隐含的几个对比预测是有意义的，如图9-1所示。

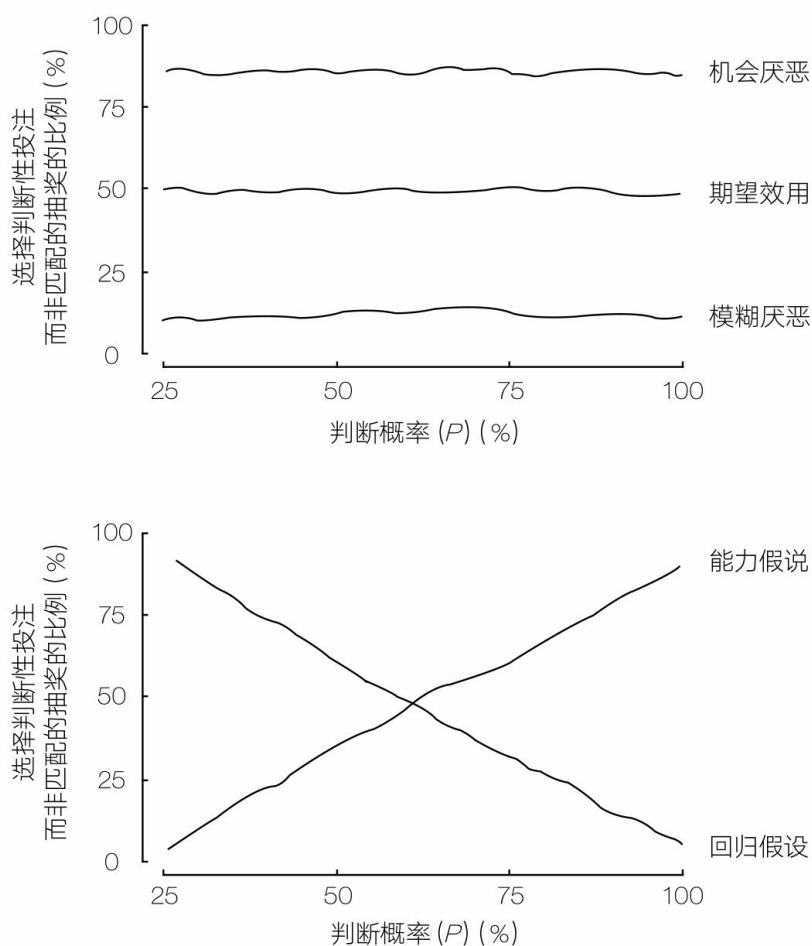


图9-1 对不确定性偏好实验结果的5种不同预测

图9-1的上半部分展示了3个假设的预测，其中 C 独立于 P 。根据期望效用理论，决策者在投注他们的判断或投注随机抽奖之间无差异；因此 C 应始终等于50%。模糊厌恶意味着人们会倾向于在一个概率很明确的随机事件上投注，而不是在总是很模糊的判断概率上投注；因此 C 应

该每次都低于50%。与此相对的假设，被称为机会厌恶，预测人们会倾向于投注他们的判断而不是与之相匹配的随机抽奖。因此，对于所有 P ， C 应超过50%。与上半部分显示的 averages 的预测相比，下半部分中的两个假设意味着 C 取决于 P 。回归假设表明，控制选择的决策权重相对于所述的概率是递减的。因此， C 对于小概率相对较高而对于大概率则相对较低。这一预测也来自H.艾因霍恩（H. Einhorn）和罗宾·贺加斯（Robin Hogarth）提出的理论，他们提出了一个基于心理模拟、调整和锚定的特定过程模型（Einhorn & Hogarth, 1985）。然而，该模型的预测与回归假设一致。

最后，能力假说意味着当人们感到对某件事很了解时会倾向于在自己的判断上投注，而当他们感到不了解时则倾向于在随机抽奖上投注。因为较高的概率通常需要更多的知识，所以 C 将是 P 的递增函数，除非100%的随机抽奖是确定的。

实验结果如表9-1和图9-2所示。对于 P 的3个不同范围，表9-1显示了自己的答案上投注而不是在匹配的抽奖上投注的有偿和无偿被试的百分比。回想一下，每个问题都有4个可能的答案，因此最低信心为25%。图9-2显示了支持投注自己的判断而不是抽奖的选项 C 的总体百分比是判断概率 P 的函数。[\(24\)](#)该图表示当 P 为低值或中等值（低于65%）时，被试选择抽奖；当 P 很高时，他们会选择自己的答案。有偿和无偿被试的结果模式相同，但后者的效果略强。这些结果证实了能力假说的预测，并排除了4个备选解释，特别是二阶概率模型隐含的模糊厌恶假设，以及艾因霍恩和贺加斯模型所暗示的回归假设（Gärdenfors & Sahlin, 1982; Einhorn & Hogarth, 1985）。

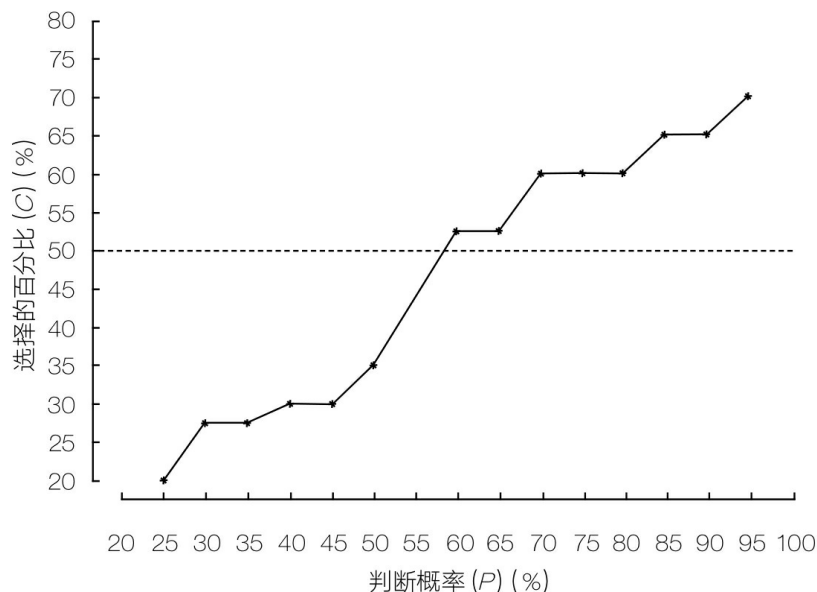


图9-2 在实验1中，倾向于优先考虑在判断上投注而不是在匹配的抽奖上投注的选择的百分比（C）是判断概率（P）的函数

表9-1 优先考虑在判断上投注而不是在低、中、高抽奖上投注的有偿被试和无偿被试百分比

	$25 \leq P \leq 50$	$50 < P < 75$	$75 \leq P \leq 100$ (%)
有偿	29 (278)	42 (174)	55 (168)
无偿	22 (394)	43 (188)	69 (140)

注：括号中是观察到的数量。

为了获得能力假说的统计检验，我们分别针对每名被试计算了选择（判断赌注与抽奖）和判断概率（中位数以上与中位数以下）之间的二元相关系数（ ϕ ）。判断中位数为0.65。72%的被试得到的是正系数，平均值为0.30，（ $t(54)=4.3, p<0.01$ ）。为了研究观察到的模式的稳健性，我们重新进行了一次实验，并对实验进行了较大程度的更改。我们没有建立概率与被试所述值相匹配的随机抽奖，而是建立了获胜的概率比被试的判断概率高出6%或低6%的抽奖。对于非常了解的问题

（ $P \geq 75\%$ ），大多数被试的回答（70%）都支持对自己的判断投注而不是抽奖，即使抽奖有更高的获胜概率（6%）。类似地，对于低信心问题（ $P \geq 50\%$ ），即使前者提供较低的获胜概率（6%），大多数被试（52%）也都支持投注抽奖而不是自己的判断。

图9-3显示了实验1中数据的校准曲线（标定曲线）。从图中可以看出，从整体上看，人们对低概率的判断是合理的，但对高概率的判断则表现出明显的过度自信。因此，在精算的基础上，倾向于投注自己的判断而不是高概率的抽奖是不合理的。

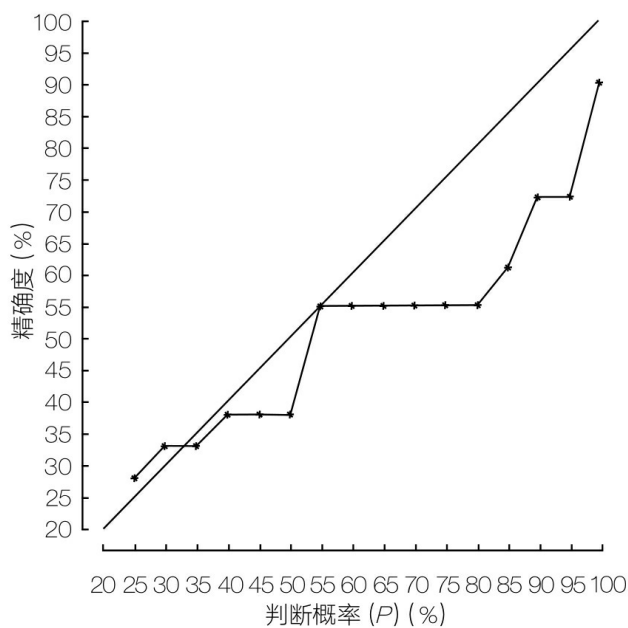


图9-3 实验1的校准曲线

对互补投注的分析得到了一种截然不同的模式，在这种投注中，被试被要求对他们选择的答案是错误的进行投注。在所有被试中，40.5%的人支持对判断投注，这表明他们对随机抽奖有统计学意义上的显著偏好 [$t(54) = 3.8, p < 0.01$]。此外，我们发现 C 和 P 之间没有系统性关系，这与图9-2中显示的单调关系形成了鲜明对比。根据我们的归因解释，这一结果表明人们更愿意投注自己的信念而不是相反。然而，这些数据也可以通过以下假设来解释，即人们更愿意投注于简单假设，而不是析取假设。

实验2：足球和政治

我们的下一个实验与上一个实验有3个方面的不同之处。首先，它关注的是对现实中的未来事件的预测，而不是对一般知识的评估。其次，它处理的是二元事件，这使得最低信心水平为0.5，而不是上一个实验中的0.25。最后，除了概率判断外，被试还评估了他们在每个预测中的知识水平。

一组20名学生连续5周每周14场的足球比赛的结果进行了预测。对于每场比赛，被试选择他们认为将赢得比赛的队伍，并评估他们所选队伍获胜的概率。被试还以5分制评估自己在每场比赛方面的知识水平。在评级之后，被试被问及他们是愿意在自己选择的球队上投注还是在相匹配的随机抽奖上投注。图9-4中显示的结果证实了先前的发现。对于高知识水平和低知识水平项（由知识评定量表中的中位数来分别定义）， C 是 P 的递增函数。此外，对于任何 $P > 0.5$ ，高知识水平项的 C 都高于低知识水平项的 C 。只有5%的被试在 C 和 P 之间产生负相关，平均系数为0.33 [$t(77) = 8.7, p < 0.01$]。

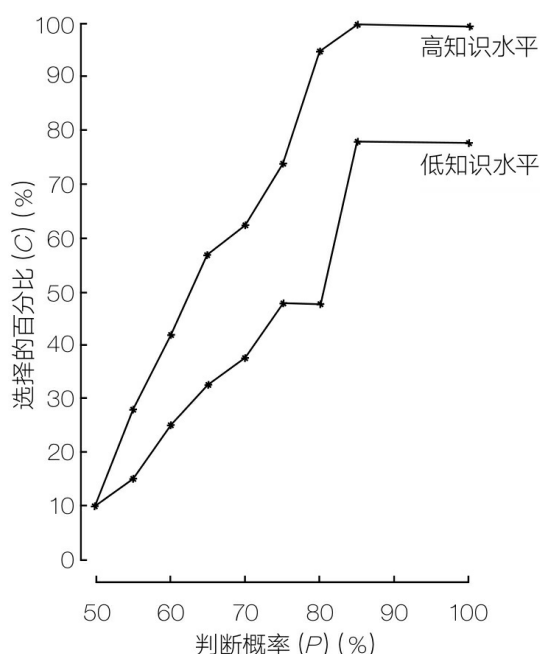


图9-4 对于足球预测任务中的高知识水平项和低知识水平项（实验2）倾向于对判断进行投注而不是对匹配的抽奖进行投注选择的百分比（ C ），是判断概率（ P ）的函数

我们接下来在1988年8月于新奥尔良举行的共和党全国代表大会上提出了能力假说。被试是大会的志愿工作者。他们收到了一份一页长的调查问卷，其中包含指导性说明和答卷。我们选出了13个州代表不同地理区域的横截面，包括在选举中最重要的州。被试（ $N=100$ ）由0（布什肯定会输）到100（布什肯定会赢）来评估布什在1988年11月的选举中拿下这13个州的概率。就像在足球实验中一样，参与者用5分制给自己对每个州的了解程度进行打分，并指出是愿意投注自己的预测还是投注随机抽奖。结果如图9-5所示，两种知识水平的 C 均随 P 增加，且在 P 水平上，高知识项的 C 高于低知识项。当被问及家乡时，70%的参与者

选择了 C 对判断进行投注而不是对抽奖进行投注。只有5%的被试在 C 和 P 之间产生负相关，平均 ϕ 系数为0.42 [$t(99) = 13.4$, $p < 0.01$]。

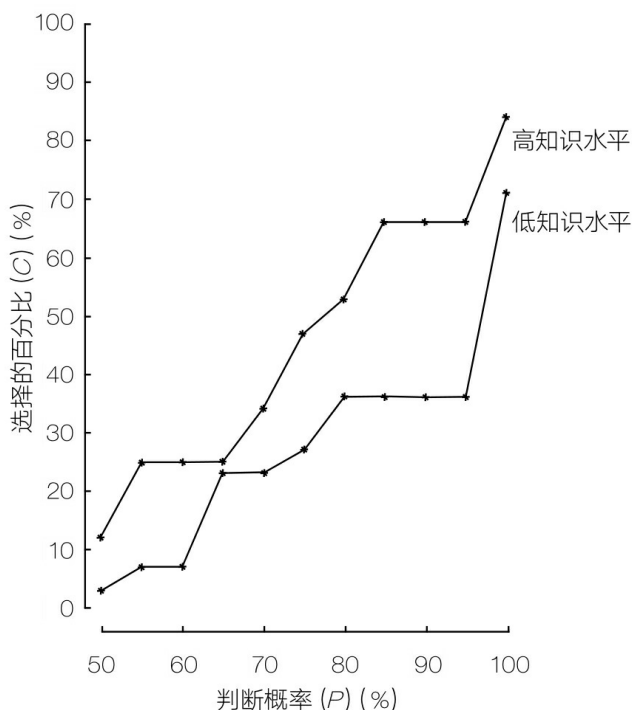


图9-5 对于实验2中的高知识水平项和低知识水平项（选举数据），倾向于对判断而不是匹配的抽奖投注的被试百分比（ C ）是判断概率（ P ）的函数

图9-4和图9-5中显示的结果支持预测真实事件的能力假说：在两个任务中， C 随 P 的增加而增加，如实验1所示。然而，在该研究中，概率和知识水平完全相关；因此，选择-判断差异可归因于判断任务中概率尺度的失真。该解释不适用于本实验的结果，因为本实验的结果显示了被评估的知识水平的独立影响。从图9-4和图9-5中可以看出，对于所有级别的判断概率，被试在高知识水平项中比在低知识水平项中更倾向于投注判断而不是随机抽奖。值得注意的是，在两个数据集中，投注判断的策略不如投注机会的策略成功。前者的命中率分别为64%和78%，而后者的命中率分别为73%和80%。因此，我们观察到的投注选择判断的倾向不会产生更好的表现。

实验3：远射

前面的实验表明，人们往往更喜欢投注自己的判断而不是匹配的随机事件，即使前者比后者更加模棱两可。图9-2、图9-4和图9-5显示的这

种效应是在概率量表的高端观察到的。这些数据也许可以用一个简单的假设来解释，即当获胜概率超过0.5时，人们更喜欢对判断投注，而当获胜概率低于0.5时，人们更喜欢随机抽奖。为了验证这一假设，我们采用了获胜概率较低的高知识水平项，因此被试的最佳猜测不太可能成立。在这种情况下，上述假设意味着对随机抽奖的偏好，而能力假说意味着对判断投注的偏好。这些预测在以下实验中进行检验。

由108名学生组成的被试被要求回答12个关于未来事件的开放式问题，如哪部电影将赢得今年的奥斯卡最佳影片奖？哪支球队将赢得下一届超级碗？下个季度你哪一科的成绩最好？他们被要求回答每个问题，并估计自己预测正确的可能性，同时还要指出自己在相关领域的知识是高水平还是低水平。开放式问题的使用消除了在先前实验中使用二分预测所设定的50%的下限。在被试完成这些任务后，他们被要求分别针对每个问题考虑是愿意投注于自己的预测还是相匹配的随机抽奖。

被试平均回答了12个问题中的10个。表9-2给出了在高知识含量和低知识含量项目中，以及在判断概率低于或高于0.5的情况下，相对于随机抽奖，支持判断投注的被试的百分比（C）。每个单元格中的回答数量在括号中给出。结果表明，对于高知识含量项目，无论P是高于还是低于50%（两种情况下 $p<0.01$ ），判断投注都优于随机抽奖，这与能力假说相一致。实际上， $P<0.5$ 时低知识和高知识水平项目之间的差异大于 $P\geq0.5$ 的时候。显然，即使预测不太可能正确，人们也更愿意投注自己对高知识水平项目的预测。

表9-2 倾向于选择对判断投注而不是对高知识水平和低知识水平，以及对判断概率低于或高于0.5的匹配抽奖进行投注的被试百分比（C）

知识评级	判断概率	
	$P < 0.5$	$P \geq 0.5$
低	36	58
	(593)	(128)
高	61	69
	(151)	(276)

注：括号中给出的是回答数量。

实验4：专家预测

在前面的实验中，我们使用被试对特定项目的评级来定义高知识水平和仅仅知道的知识水平。在这个实验中，我们根据专业知识对被试进行分类来控制知识或能力。为此，我们邀请110名学习心理学课程的学生，以9分制评估他们对政治和足球的了解状况。在两个领域中的知识水平的评分在中点两侧的所有被试都被要求参加实验。有25名被试符合这一标准，除了其中两名之外，所有人都同意参加。他们包括12名政治“专家”和11名足球“专家”，这是根据他们擅长的领域来定义的。为了使被试给出认真的答复，我们向他们提供了详细的说明，包括校准的讨论，并且我们采用了布列尔评分规则，旨在激励被试给出他们的最佳预测（Lichtenstein et al., 1982）。被试的平均收入约为10美元。

实验包括两个阶段。在第一阶段中，每名被试对40个未来事件（20个政治事件和20个足球比赛）进行了预测。所有的事件均在第一阶段起5周内得到了确切的结果。这些政治事件是关于1988年总统选举中各州的获胜者的。20场足球比赛包括10场职业比赛和10场校际比赛。对于每场比赛（政治或足球），被试通过圈选其中一个参赛者的名字来选择获胜者，然后评估自己的预测将成为现实的概率（从50%到100%）。

根据第一阶段的结果，我们为每个参与者构建了20个投注“三元组”。每个三元组包括3个匹配的投注，投注者的获胜概率相同，它们由以下3项产生：第一，机会设置；第二，被试在其擅长的领域中的预测；第三，被试在其不擅长的领域中的预测。显然，有些事件不只出现在一个三元组中。在第二阶段中，被试对20个投注三元组进行排序。在实验1中，我们参照一个装有100个编号筹码的盒子来定义机会投注。被试被告知他们实际上会在每个三元组中做出选择。为了鼓励他们认真排序，我们告知被试他们在投注游戏中的表现将由第一选择的80%和第二选择的20%联合决定。

数据汇总如表9-3和图9-6所示，其中绘制了3种类型投注对判断概率的吸引力（平均排名顺序）。结果表明人们明显倾向于在擅长的领域进行投注。在所有三元组中，擅长领域的平均等级为1.64，随机抽奖的平均等级为2.12，不擅长领域的平均等级为2.23。通过威尔科克森秩和检验，等级之间的差异性非常显著（ $p < 0.001$ ）。根据能力假说，在能力范围内人们更愿意投注自己的判断，但在他们没有充分知情的领域里更愿意投注随机事件。正如预期的那样，抽奖比100%的高知识水平的投注更受欢迎。这种结果模式与基于模糊性或二阶概率的解释不一致，因为高知识水平和低知识水平的投注都是基于模糊的判断概率，而随机抽

奖有明确的概率。模糊厌恶可以解释为什么低知识水平的投注不如高知识水平的投注或机会投注有吸引力，但它无法解释这个实验的主要发现，即模糊的高知识水平投注比清晰的机会投注更受欢迎。

表9-3 专家研究的排序数据

投注类型	排名			
	第一	第二	第三	平均排名
高知识水平	192	85	68	1.64
机会	74	155	116	2.12
低知识水平	79	105	161	2.23

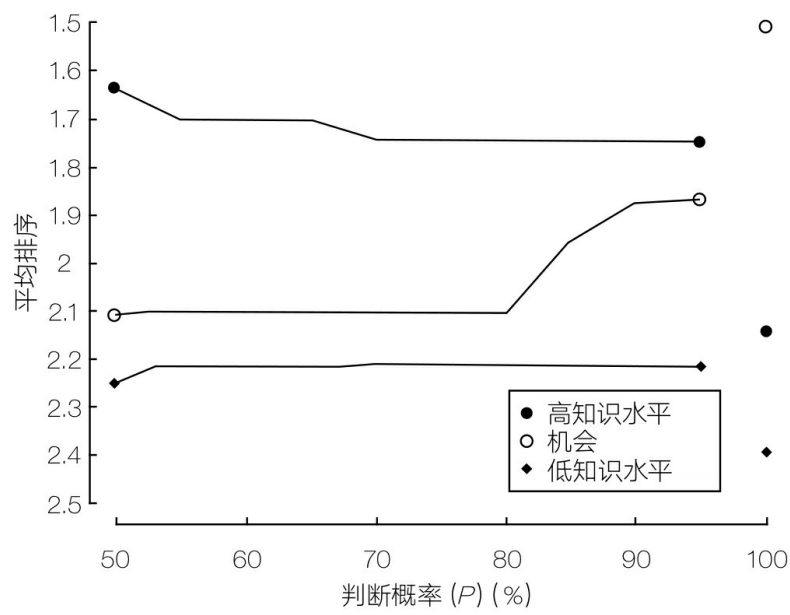


图9-6 实验4中作为P函数的高知识水平、低知识水平和机会投注的排序数据

图9-6与前几幅图的一个显著区别是，偏好在本质上是独立于P的。显然，在这种情况下，通过类别间的对比，充分体现了能力效应；因此，判断概率所暗示的附加知识对投注的选择几乎没有影响。

图9-7显示了实验4的平均校准曲线，分别针对高知识水平和低知识水平类别。由这些图表可知，判断通常都表现出过度自信：被试的信心超过了他们的命中率。此外，过度自信在高知识水平类别中比在低知识水平类别中更为明显。因此，投注的顺序并不反映判断的准确性。总结所有的三元组，投注随机抽奖的概率为69%，投注新手类别的概率为

64%，投注专家类别的概率仅为60%。通过投注专家类别，被试实际上会损失15%的预期收益。

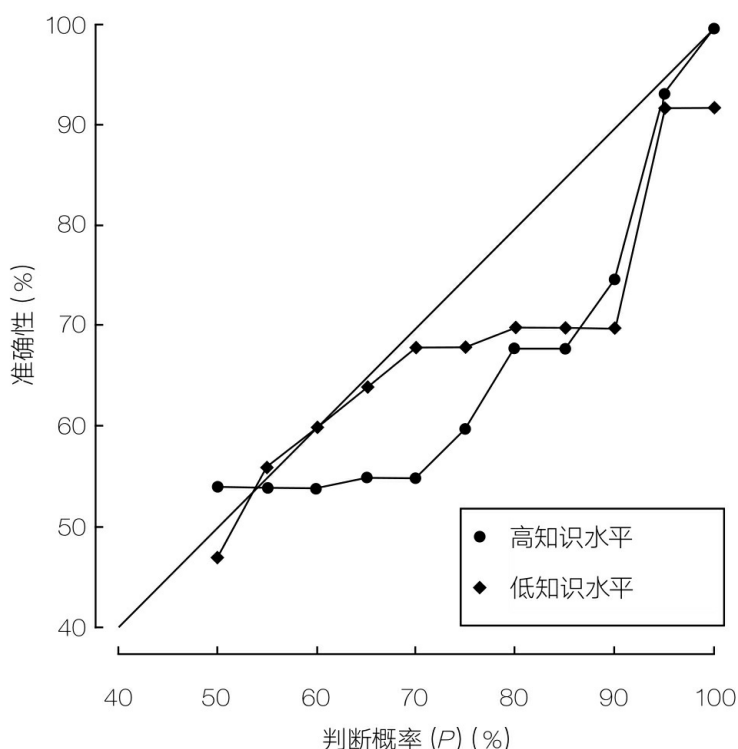


图9-7 实验4中高知识水平类别和低知识水平类别的校准曲线

对知识而非机会的偏好不仅体现在分类事件的概率判断（胜利、损失）上，也体现在数值变量的概率分布上。被试（ $N=93$ ）有机会为各种数据设置80%的置信区间（如斯坦福大学入学新生的平均SAT分数；圣弗朗西斯科到洛杉矶的行驶距离）。在设置置信区间后，被试被要求在以下两项中进行选择：一是其置信区间中含有真实值；二是有80%胜率的抽奖。在大多数情况下，被试更倾向于投注置信区间，尽管这种策略仅在69%的情况下得到了回报，因为他们设定的置信区间通常过于狭窄。同样，被试支付了近15%的溢价来投注自己的判断。

实验5：补充投注

前面的实验依赖于概率判断与随机抽奖和判断投注相匹配。为了控制判断过程中可能存在的偏差，我们对能力假说进行的最后检验是基于一个不涉及概率判断的定价任务。在该实验中我们还要求被试对为投注高知识水平项目而支付的额外费用进行了估计。

由68名学生组成的受试组被要求在12次投注中都要说明其现金等价物（预订价格）。他们被告知选择一对投注目标，一些随机选择的学生将对他们所说的更高额度的现金等价物进行投注（Tversky, Slovic & Kahneman, 1990）。如果给定的命题是真的，那么这个实验中的所有投注都会有15美元的奖金，否则将什么也没有。我们针对不同的被试给出了互补命题。例如，如果纽约和圣弗朗西斯科之间的空中距离超过2 500英里，那么有一半的被试被要求支付15美元的投注定价。如果纽约和圣弗朗西斯科之间的空中距离小于2 500英里，那么另一半的被试被要求为支付15美元的补充投注进行定价。

为了研究不确定性偏好，我们将高知识水平命题和低知识水平命题配对。例如，我们假设被试更了解纽约和圣弗朗西斯科之间的空中距离，而不是东京和曼谷之间的空中距离。我们还假设被试（斯坦福大学的学生）更了解在斯坦福大学而不是拉斯维加斯内华达大学接受校内住宿的本科生的百分比。与之前一样，我们将这些命题分别称为高知识水平项目和低知识水平项目。注意，不确定数量的所述值的选择（如空中距离和学生的百分比）控制了被试对所讨论命题的效度的信心，而与他拥有的与该主题相关的常识无关。我们构建了12对互补命题，并且每名被试评估由每对定义的4个投注中的一个。例如，在空中距离问题中，4个命题是 $d(\text{SF}, \text{NY}) > 2\,500$ ， $d(\text{SF}, \text{NY}) < 2\,500$ ， $d(\text{Be}, \text{Ba}) > 3\,000$ ， $d(\text{Be}, \text{Ba}) < 3\,000$ ，其中 $d(\text{SF}, \text{NY})$ 和 $d(\text{Be}, \text{Ba})$ 分别表示圣弗朗西斯科和纽约之间及东京和曼谷之间的空中距离。

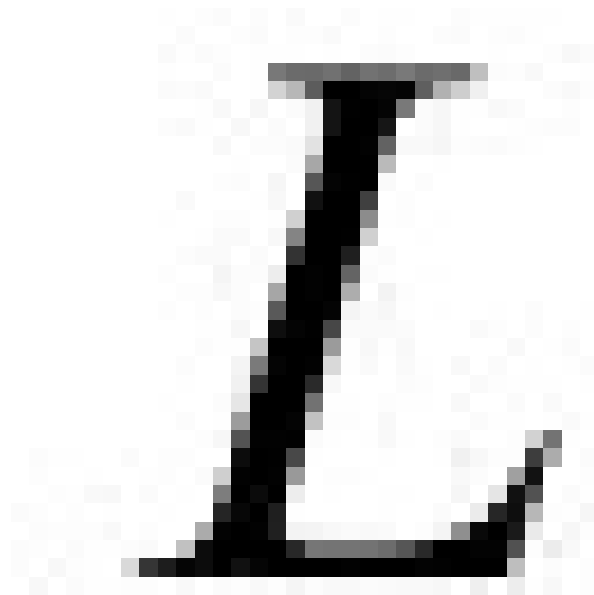
请注意，根据预期值，每对互补投注的平均现金等价物应为7.50美元。对所有12对互补投注进行总结得出，被试平均为高知识水平投注支付了7.12美元，对低知识水平投注仅支付了5.96美元（ $p < 0.01$ ）。因此，人们实际上支付了近20%的能力溢价，以便对更熟悉的命题进行投注。此外，在12个问题中，有11个问题的（互补的）高知识水平投注的平均价格高于低知识水平投注的平均价格。相比之下，投掷硬币赢15美元的平均现金等价物为7美元。与我们之前的研究结果一致，随机抽奖的价值高于低知识水平投注，但不高于高知识水平投注。

接下来，我们将用期望效用理论来检验能力假说。设 H 和

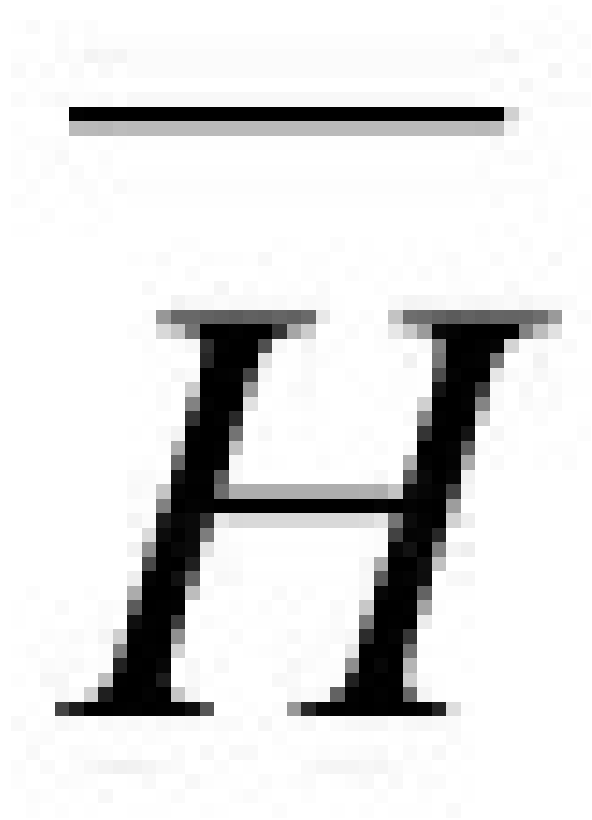
—

H

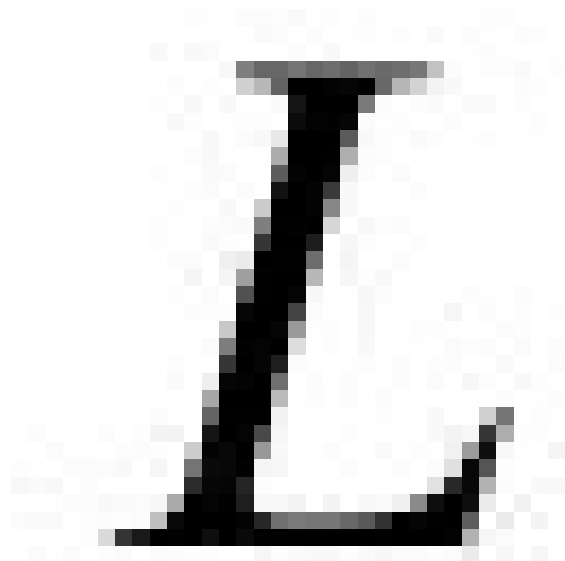
表示两个互补的高知识水平的命



题，并设 L 和 H 表示相应的低知识水平命题。假设一个决策者对投注 H 的偏好大于 L ，对投注



的偏好大于



。这种模式与期望效用理论不一致，因为它暗示了 $P(H) > P(L)$ 且

$$P(\bar{H}) > P(\bar{L})$$

，与可加性假设 $P(H) + P(\bar{H}) = P(L) + P(\bar{L}) = 1$ 相反。此外，如果高知识水平的投注比低知识水平的投注更受欢迎，那么可能出现这种模

式。因为4个命题 $(H, \bar{H}, L, \bar{L})$ 由4个不同的受试组进行评估，我们采用了被试之间的可加性测试。设 $M(H_i)$ 为高知识水平命题 H_i 的中位数价格。当 $M(H_i) > M(L_i)$ 且 $M(\bar{H}_i) \geq M(\bar{L}_i)$ 时，对问题 i 的反应违反了能力假说所暗示的可加性。

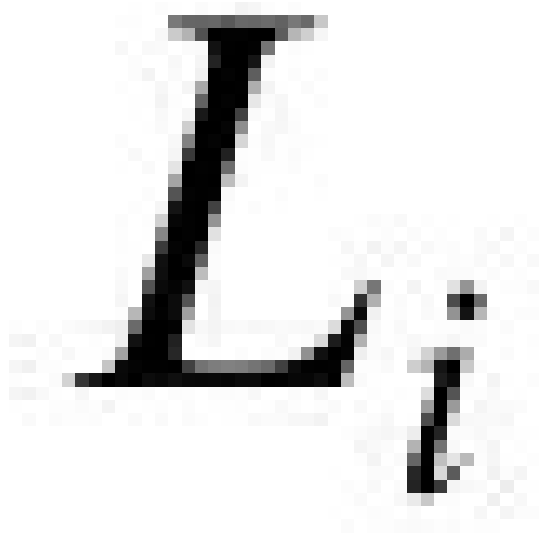
被试在12对问题中的5对中表现出了对高知识水平投注的偏好，并且没有一对表现出相反的模式。例如，对“斯坦福大学85%以上的本科生选择了校内住宿”这一命题投注的中位数为7.50美元，而对互补命题下注的现金等价物中位数为10美元。相比之下，投注“超过70%的拉斯维加斯内华达大学本科生选择校内住宿”这一命题的现金等价物中位数为3美元，互补命题投注的中位数为7美元。因此，大多数被试愿意支付更多金额来投注高知识水平项目的任意一方，而不是低知识水平项目的任何一方。

基于中位数的前述分析可以扩展如下。对于每对命题 (H_i, L_i) ，我们计算了 H_i 的现金等价物超过了 L_i 的现金等价物的相对比例，用 $P(H_i > L_i)$ 表示。我们还计算了互补命题的

$P(\bar{H}_i > \bar{L}_i)$ 。所有关系都被排除在外。根据期望效用理论

$$P(H_i > L_i) + P(\bar{H}_i > \bar{L}_i) = P(L_i > H_i) + P(\bar{L}_i > \bar{H}_i) = 1$$

因为概率的可加性意味着对于支持 H_i 而不是 L_i 的每一个比较，都应



该有另一个支持

而不是



的比较。此外，如果人们更喜欢高知识水平的赌注，正如能力假说所暗示的那样，我们期望

$$P(H_i > L_i) + P(\bar{H}_i > \bar{L}_i) > P(L_i > H_i) + P(\bar{L}_i > \bar{H}_i)$$

在12对互补命题中，10对满足了上述不等式，1对满足相反的不等式，1对为等式，这表明在能力假说所暗示的方向上显著违反了可加性（通过符号检验， $p<0.01$ ）。这些发现证实了检验中的能力假说，该检验不依赖于概率判断或者判断投注与匹配抽奖的比较。因此，目前的结果不能归因于在判断过程中或在高知识水平项目和低知识水平项目的匹配中存在的偏好。

讨论

本文中提到的实验在概率判断和投注选择之间建立了一致且普遍存在的差异性。实验1表明，对知识而不是随机抽奖投注的偏好随着判断信心的增加而增加。实验2和实验3在真实世界的未来事件中同样得到了这一结论，并证明了知识效应与判断概率无关。在实验4中，我们将被试分到擅长和不擅长的领域，并证明了人们喜欢投注自己擅长的领域且不喜欢投注自己不擅长的领域；机会投注介于两者之间。这种模式不能通过模糊性或二阶概率来解释，因为机会是明确的，而判断概率是模糊的。最后，实验5证实了在不依赖概率匹配的定价任务中对能力假说的预测，并且证明了人们会为投注高知识水平项目支付近20%的溢价。

这些观察结果与我们的归因解释一致，该解释认为知识导致了表扬和批评的内在平衡的非对称性。我们认为，有能力让人们在正确的时候获得赞扬，而缺乏能力则让人们在错误的时候受到指责。因此，人们更喜欢在高知识水平上投注而不是匹配的抽奖，并且他们更喜欢在匹配的抽奖而不是在低知识水平上投注。这个说明解释了文献中提到的其他不确定性偏好的例子，特别是在机会设置中偏好清晰的概率而不是模糊的概率（Ellsberg, 1961），偏好对未来的事件投注而不是对过去的事件投注（Rothbart & Snyder, 1970; Brun & Teigen, 1989），偏好对技能投注而不是对随机事件投注（Cohen & Hansel, 1959; Howell, 1971），以及在知识渊博的人面前表现出更强烈的模糊厌恶情绪（Curley, Yates & Abrams, 1986）。一个强有力的发现是，在能力范围内，人们更喜欢为自己（模糊）的信念投注而不是为相匹配的偶然事件投注，这表明知识或能力的影响超过了模糊的影响。

在实验1至实验4中，我们使用概率判断来建立信念，通过选择数据以建立偏好。此外，我们将选择-判断差异解释为偏好效应。相反，可以认为选择-判断差异可归因于判断偏差，即低估高知识水平项目的概率且高估低知识水平项目的概率。但是，现有证据并不支持这种解释。

首先，它意味着人们对高知识水平项目不像对低知识水平项目那样过度自信，这与事实相反（如图9-7所示）。其次，概率判断不能被认为是无关紧要的，因为在存在评分规则的情况下，如实验4中使用的评分规则，这些判断代表着另一种形式的投注。最后，判断偏差无法解释实验5的结果。实验5显示，在不涉及概率判断的定价任务中，人们倾向于投注高知识水平项目。

偏好和信念之间的区别是贝叶斯决策理论的核心。对该理论的标准解释假定：第一，个体表达的信念（即概率判断）与可加的概率尺度一致；第二，个体的偏好与期望原则一致，因此从选择中得出了（主观）概率尺度；第三，从判断和选择中获得的主观概率的两个尺度是一致的。请注意，前两点在逻辑上是独立的。例如，阿莱的反例违反了第二点但没有违反第一点。实际上，许多作者引入了源于偏好的非可加性决策权重，以适应观察到的违反期望原则的情况（Kahneman & Tversky, 1979）。不过这些决策权重无须反映决策者的信念。一个人可能会认为从一副洗好的牌中抽出黑桃A的可能性是1/52，但是在对这个事件进行投注时人们可能会赋予它更高的权重。同样，埃尔斯伯格的例子并不能证明人们认为明确事件比相应的模糊事件更有可能发生；它只表明了人们更愿意在明确的事件上投注。遗憾的是，“主观概率”一词在文献中被用来描述来自偏好的决策权重及信念的直接表达。在贝叶斯理论的标准解释下，这两个概念是一致的。然而，当我们超越这一理论时，对两者进行区分是至关重要的。

控制模糊性

信念和偏好之间的区别对模糊效应的解释尤为重要。一些作者得出的结论是，当获胜的概率很小或失败的概率很高时，人们宁愿它是模糊不清的（Curley & Yates, 1989; Einhorn & Hogarth, 1985; Hogarth & Kunreuther, 1989）。然而，这种解释可能会受到挑战，你很快就会看到，数据可能反映的是信念的差异，而非不确定性偏好。在本节中，我们研究了用于控制模糊性的实验程序，并认为它们往往将模糊性与感知概率混淆了。

也许控制模糊性最简单的程序是改变决策者对给定的概率估计的信心。贺加斯和他的合作者使用了这个程序的两个版本。艾因霍恩和贺加斯基基于“独立观察者的判断”，向被试展示了概率估计，并改变了与该估计相关的信心（Einhorn & Hogarth, 1985）。贺加斯和H.昆鲁斯（H.

Kunreuther) 赋予了被试对给定事件“概率的最佳估计”，并通过改变与该估计相关的信心来控制模糊性 (Hogarth & Kunreuther, 1989)。但是，如果我们希望将人们对这类事件的投注意愿解释为模糊寻求或模糊厌恶，就必须首先验证控制模糊性不会影响被试对事件的感知概率。

为了对这个问题进行研究，我们首先复制了贺加斯和昆鲁斯所做的模糊性控制 (Hogarth & Kunreuther, 1989)。我们组织了一组被试 (N=62)，称其为高信心组，并向他们提供了以下信息：

想象一下，你是一家大型保险公司的部门负责人。一家小公司的老板来找你，要求为一笔10万美元的损失投保，这笔损失可能是由于客户对一件有缺陷的产品的索赔而造成的。你已经考虑过制造过程、所用机器的可靠性及业务记录中包含的证据。在考虑了现有的证据之后，你对出现残次品的可能性的最佳估计值是0.01。在这种情况下，你对这个估算的准确性充满了信心。当然，随着对情况进行更深入的思考或收到了其他信息，你会更新自己的估计。

第二组被试 (N=64) 为低信心组。他们收到了相同的信息，只是“你对这个估算的准确性充满信心”被替换为“你认为这个估算的准确性有很大的不确定性”。

然后所有被试被问到：

你认为新的估计是 (单选):

高于 0.01 _____

低于 0.01 _____

正好是 0.01 _____

这两个受试组还被要求评估另一个案例，其中所述的损失概率为0.90。如果所述值 (0.01或0.90) 被解释为相应二阶概率分布的均值，那么被试对更新后的估计的期望应与当前“最佳估计”一致。此外，如果控制信心影响了模糊性而没有影响感知概率，那么高信心组和低信心组的反应之间应该没有差异。表9-4中“你给出的概率”标题下的数据明显与这些假设不一致。低信心条件下的答案分布比高信心条件下的分布偏差更大。此外，概率为0.01时，偏度为正，而概率为0.90时，偏度为负。告知被试他们的最佳估计“有相当大的不确定性”会引发逆转：第一个问题的预期损失概率高于0.01，第二个问题低于0.90。信心 (高-低) 和方向 (高于-低于) 之间的相互作用在统计学上是显著的 ($p<0.01$)。

表9-4 在高信心和低信心指令下对0.01和0.90的所述概率的主观评估

给定值	答案	你给出的概率		他人的估计	
		高信心	低信心	低信心	低信心
0.01	高于 0.01	45	75	46	80
	恰好是 0.01	34	11	15	6
	低于 0.01	21	14	39	14
0.90	高于 0.01	29	28	42	26
	恰好是 0.01	42	14	23	12
	低于 0.01	29	58	35	62

注：表中数据是选择3个响应中的每个响应的被试百分比。

我们还重复了艾因霍恩和贺加斯所用的方法，其中被试被告知“独立观察者已经表明有残次品的概率是0.01”（Einhorn & Hogarth, 1985）。高信心组中的被试（ $N=52$ ）被告知“你可能会对估计值感到自信”，而低信心组的被试（ $N=52$ ）被告知“你可能会认为估计值有相当大的不确定性”。然后两组被试都被问到他们对遭受损失概率的最佳猜测是0.01、低于0.01，还是恰好是0.01。两组被试还评估了另一个损失概率为0.90的案例。表9-4中“他人的估计”标题下的结果，揭示了上面观察到的模式。在高信心条件下，答案的分布非常对称；但在低信心条件下，分布在0.01时呈现正偏态，在0.90时呈现负偏态。同样，信心（高-低）和方向（高于-低于）之间的相互作用在统计学上是显著的（ $p<0.01$ ）。

这些结果表明，控制信心不仅影响了所讨论事件的模糊性，还影响了其感知概率：它们增加了极不可能事件的感知概率，并降低了可能事件的感知概率。这种类型的逆转根本不合理，甚至可以通过适当的先验分布加以合理化。由于概率的转变，当损失概率高（0.90）时，投注于更模糊的估计应该更有吸引力，而当损失概率低（0.01）时，投注于更模糊的估计应该就不再有吸引力了。当损失概率很高（0.90）时，对模糊估计的投注应该更具吸引力，而当损失概率较低（0.01）时，则不那么有吸引力。这正是艾因霍恩和贺加斯及贺加斯和昆鲁斯所观察到的偏好模式，但它并不需要模糊寻求或模糊厌恶，因为不仅在模糊性方面，事件的感知概率也不同（Einhorn & Hogarth, 1985; Hogarth & Kunreuther, 1989）。

表9-4的结果及贺加斯和他的合作者的发现可以通过以下假设来解释：被试将所述概率值解释为二阶概率分布的中位数（或模式）。如果与极端概率相关的二阶分布偏向于0.5，那么平均值不如中位数极端，并且当模糊度高时，它们之间的差异大于模糊性低时的差异。因此，控制贝叶斯模型中的选择的二阶概率分布的均值在低信心下比在高信心下更具有回归性（即更接近0.5）。

即使当模糊性被机会过程的信息控制时，也可能会出现模糊性和信念程度的混淆。与埃尔斯伯格对50/50盒子与未知盒子的比较不同，其中对称性排除了一个方向或另一个方向上的偏差，不对称问题中类似的控制模糊性的操作可能产生逆转，如帕雷尔（Parayre）和卡尼曼未发表的研究所示。[\(25\)](#)

这些研究人员比较了一个明确的事件（由盒子里红球比例决定）和一个模糊的事件（由指定颜色的球所占的比例范围决定）。对于一个模糊事件 $[0.8, 1]$ ，被试被告知红球的比例为0.8~1，而明确的事件则为0.9。表9-5列出了低、中、高3个概率水平的选择和判断数据。与先前的工作一致，选择数据显示，当获胜概率较低且失败概率较高时，被试倾向于在模糊事件上投注，并且在其他所有情况下他们都倾向于投注明确事件。帕雷尔和卡尼曼实验的新特征是使用基于长度判断的感知评定量表，它提供了对概率的非数值评估。通过应用这个量表，研究人员发现判断概率是递减的。也就是说，模糊的低概率事件 $[0, 0.10]$ 被认为与明确事件相比更可能为0.05，且模糊的高概率事件 $[0.8, 1]$ 被认为与明确事件相比更不可能为0.9。对于中等概率，模糊事件 $[0, 1]$ 与明确事件之间的判断没有显著差异，为0.5。这些结果与表9-4的数据一致，这表明对模糊事件的偏好（在低端观察到是正投注而在高端观察到是负投注）可以反映出对感知概率上的倒退性转变，而不是对模糊的偏好。

表9-5 在判断和选择中偏好明确事件和模糊事件的被试百分比

概率（赢 / 输）		判断 $N = 72$	选择	
			赢 100 美元 $N = 58$	输 100 美元 $N = 58$
低	0.05	28	12	66
	[0,0.10]	47	74	22
中	0.5	38	60	60
	[0, 1]	22	26	21
高	0.9	50	50	22
	[0.8, 1]	21	34	47

注：每种条件下两个值的总和小于100%，其余的反应等同。在选择任务中，低概率为0.075和 [0, 0.15]。N表示样本大小。

结束语

关于能力效应、偏好与信念之间关系的研究结果对假设偏好和信念的独立的选择模型的标准解释提出了挑战。结果也与后贝叶斯模型不一致，后贝叶斯模型应用二阶信念来解释模糊性或一部分知识的影响。此外，我们的研究结果对用偏好来定义信念的基本观点提出了质疑。如果对不确定性事件投注的意愿取决于该事件的感知概率和对该估计的信心，那么从投注之间的偏好中得出潜在的信念（如果有可能的话）是非常困难的。

除了挑战现有模型外，能力假说可能有助于解释不确定性决策的一些令人困惑的方面。这可能会让我们看到这样一个观察结果：许多决策者并未将其能力范围内可计算的风险视为赌博（March & Shapira, 1987）。这也可能有助于解释为什么投资者有时愿意放弃多元化的优势，而把注意力集中在他们可能熟悉的少数几家公司上（Blume, Crockett & Friend, 1974）。能力假说对整个决策的影响还有待探索。

致谢

我们从与马克斯·巴泽曼、丹尼尔·埃尔斯伯格、理查德·冈萨雷斯（Richard Gonzales）、罗宾·贺加斯、琳达·金泽尔（Linda Ginzel）、丹尼尔·卡尼曼和埃尔德·沙菲尔的讨论中受益匪浅。

参考文献

- Anscombe, F. J., & Aumann, R. J. (1963). A definition of subjective probability. *Annals of Mathematical Statistics*, 34, 199–205.
- Barlow, R. E., et al. (1972). *Statistical inference under order restrictions: The theory and application of isotonic regression*. New York, NY: John Wiley.
- Blume, M. E., Crockett, J., & Friend, I. (1974). Stock ownership in the United States: Characteristics and trends. *Survey of Current Business*, 54, 16–40.
- Brun, W., & Teigen, K. (1990). Prediction and postdiction preferences in guessing. *Journal of Behavioral Decision Making*, 3, 17–28.
- Budescu, D., Weinberg, S., & Wallsten, T. (1988). Decisions based on numerically and verbally expressed uncertainties. *Journal of Experimental Psychology: Human Perception and Performance*, 14(2), 281–294.
- Cohen, J., & Hansel, M. (1959). Preferences for different combinations of chance and skill in gambling. *Nature*, 183, 841–843.
- Curley, S., & Yates, J. F. (1989). An empirical evaluation of descriptive models of ambiguity reactions in choice situations. *Journal of Mathematical Psychology*, 33, 397–427.
- Curley, S., Yates, J. F., & Abrams, R. (1986). Psychological sources of ambiguity avoidance. *Organizational Behavior and Human Decision Processes*, 38, 230–256.
- Einhorn, H., & Hogarth, R. (1985). Ambiguity and uncertainty in probabilistic inference. *Psychological Review*, 93, 433–461.
- Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics*, 75, 643–669.
- Ellsberg, D. (1963). Risk, ambiguity, and the savage axioms: Reply.

Quarterly Journal of Economics, 77, 336–342.

Fellner, W. (1961). Distortion of subjective probabilities as a reaction to uncertainty. *Quarterly Journal of Economics*, 75, 670–689.

Fishburn, P. (1988). *Nonlinear preference and utility theory*. Baltimore, MD: Johns Hopkins University Press.

Frisch, D., & Baron, J. (1988). Ambiguity and rationality. *Journal of Behavioral Decision Making*, 1, 149–157.

Gärdenfors, P., & Sahlin, N.-E. (1982). Unreliable probabilities, risk taking, and decision making. *Synthese*, 53(3), 361–386.

Hogarth, R., & Kunreuther, H. (1988). *Pricing insurance and warranties: Ambiguity and correlated risks*. Unpublished manuscript, University of Chicago and University of Pennsylvania.

Hogarth, R., & Kunreuther, H. (1989). Risk, ambiguity, and insurance. *Journal of Risk and Uncertainty*, 2, 5–35.

Howell, W. (1971). Uncertainty from internal and external sources: A clear case of overconfidence. *Journal of Experimental Psychology*, 89(2), 240–243.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.

Lichtenstein, S., Fischhoff, B., & Phillips, L. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). New York, NY: Cambridge University Press.

March, J., & Shapira, Z. (1987). Managerial perspectives on risk and risk taking. *Management Science*, 33(11), 1404–1418.

Raiffa, H. (1961). Risk, ambiguity, and the savage axioms: Comment. *Quarterly Journal of Economics*, 75, 690–694.

Ramsey, F. (1931). Truth and probability. In F. P. Ramsey (Ed.), *The foundations of mathematics and other logical essays* (pp. 156–198). New York, NY: Harcourt, Brace and Co.

Rothbart, M., & Snyder, M. (1970). Confidence in the prediction and postdiction of an uncertain outcome. *Canadian Journal of Behavioural Science*, 2(1), 38–43.

Savage, L. (1954). *The foundations of statistics*. New York, NY: Wiley.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, 57(3), 571–587.

Skyrm, B. (1980). Higher order degrees of belief. In D. H. Mellor (Ed.), *Prospects for pragmatism: Essays in memory of F. P. Ramsey* (pp. 109–137). Cambridge, England: Cambridge University Press.

Tversky, A., Slovic, P., & Kahneman, D. (1990). The causes of preference reversal. *American Economic Review*, 80, 204–217.

第10章 对小数定律的信念

阿莫斯·特沃斯基
丹尼尔·卡尼曼

“假设你在一个由20个人组成的受试组中进行了一项实验并得到了一个显著性结果，证实了你的理论（ $z=2.23$ ， $p<0.05$ ，双侧检验）。你现在需要再对另外一组的10名被试进行实验。那么通过单侧检验，这个小组的结果为显著性结果的概率是多少？”

如果你觉得概率大约是0.85，你可能会很开心地知道自己属于多数派。事实上，这是两个被试小组答案的中位数，他们很友好地填写了在数学心理学小组和美国心理学会会议上分发的调查问卷。

此外，如果你觉得概率大约为0.48，那么你属于少数派。我们的84名被试中只有9名的答案介于0.40和0.60之间。然而，0.48恰好是比0.85更合理的估计。[\(26\)](#)

显然，大多数心理学家都夸大了成功复制已获得的发现可能性的信念。这种信念的来源及其对科学探究的影响是本章的主要内容。我们的论点是人们对随机抽样有很强的直觉；这些直觉从根本上说是错误的；这些直觉是无经验的被试和训练有素的科学家所共有的，并且在科学探究过程中它们的应用导致了不好的后果。

我们认为，人们觉得从群体中随机抽取的样本具有高代表性，即在所有基本特征上与总体相似。因此，他们期望由一个特定的总体中抽取的任意两个样本比抽样理论预测的更相似、更接近总体，至少对于小样本来说是这样。

将样本视为代表的倾向在各种各样的情况中都很明显。当被试被要求随机抛出一枚均匀的硬币时，他们所得到的结果是，在任何一个短时段中正面出现的比例都比机会法则预测的结果更接近0.50（Tune, 1964）。因此，响应序列的每个部分都高度代表着硬币的“公平性”。当被试连续预测随机生成序列的事件时，研究人员也观察到了类似的结

果，如概率学习实验（Estes, 1964）或其他连续的机会游戏。被试的行为就像随机序列的每个部分必须反映真实比例那样：如果序列偏离了总体比例，那么预期会出现另一个方向的矫正偏差。这种现象被称为赌徒谬误。

赌徒谬误的核心是对机会定律的公平性的误解。赌徒认为硬币的公平性使他有权利期望一个方向的任何偏差很快就会被另一个方向的相应偏差所抵消。然而，即使是最公平的掷硬币游戏，由于记忆和道德感的限制，也无法如赌徒所期望的那样公平。这种谬误并不是赌徒独有的。请思考以下例子：

已知一座城市八年级学生的平均智商为100。你选取了50个学生为随机样本来研究教育成果。第一个接受检验的学生的智商为150。你认为整个样本的平均智商是多少？

正确答案是101。令人惊讶的是，大多数人认为样本的预期智商仍为100。只有相信随机过程是可以自我矫正的，才能证明这种期望是合理的。诸如“错误相互抵消”之类的习语反映了积极的自我矫正过程。自然界中一些熟悉的过程遵循着这样的规律：偏离平衡会产生恢复平衡的力。相比之下，机会定律并不是这样的：随着取样的进行，偏差不会消失，只是被稀释了。

到目前为止，我们已经尝试着描述了关于机会的两个相关直觉。我们提出了一个代表性假设，根据这个假设，人们认为样本彼此间非常相似，而且样本与抽取样本的总体非常相似。我们还认为人们相信抽样是一个自我矫正的过程。这两种信念导致了同样的结果。两者都产生了对样本特征的预期，这些预期的可变性小于真实可变性，至少对于小样本来说是这样。

大数定律确保了非常大的样本确实能很好地代表其所在的总体。此外，如果自我矫正倾向在起作用，那么小样本也应该具有高代表性，且彼此相似。人们对随机抽样的直觉似乎符合小数定律，它断言大数定律也适用于小数。

假设有一位科学家，他的生活遵循着小数定律。他的信念将如何影响其科学工作？假设我们的科学家所研究的现象的大小相对于不受控制的可变性较小，也就是说，他从自然界接收到的信息的信噪比较低。这位科学家可能是气象学家、药理学家，也可能是心理学家。

如果科学家相信小数定律，那么他就会对基于小样本得出的结论的

效度过度自信。举例来说，假设他正在研究两个婴儿更喜欢玩哪个玩具。在研究的前5个婴儿中，有4个表现出对同一玩具的偏好。许多心理学家在这一点上会有一些信念，认为无偏好的零假设是错误的。幸好，这种信念并不是期刊出版的充分条件，尽管对一本书来说可能是这样。通过快速计算，我们的心理学家会发现，在零假设下得到一个极端结果的概率高达3/8。

诚然，统计假设检验在科学推理中的应用存在着巨大的困难。然而，显著性水平（或似然比，贝叶斯估计可能更好）的计算迫使科学家根据抽样方差的有效估计而不是根据自己的主观偏差估计来评估所获得的效果。因此，统计检验可以通过监管许多其想要遵守小数定律的成员，来阻止科学界过于仓促地拒绝零假设（I型错误）。此外，对于未能确认有效研究假设（II型错误）的风险，却没有类似的保障措施。

想象一位研究成就需求与成绩之间的相关性的心理学家。在决定样本量时，他可能会做出如下推断：“我期望什么样的相关性？ $r=0.35$ 。我需要 N 是多少才能使结果具有显著性？（查看表格） $N=33$ 。很好，这是我的样本量。”这个推理的唯一缺陷是我们的心理学家已经忘记了抽样变异的存在，这可能是因为他认为任何样本都必须高度代表其总体。但是，如果他对总体相关性的猜测是正确的，那么样本中的相关性可能低于或高于0.35。因此，对于 $N=33$ ，获得显著性结果（即检验功效）的概率大约是0.50。

在对统计能力的详细研究中，科恩对大、中、小效应给出了合理的定义，并为各种统计检验的效应估计提供了一套广泛的计算辅助工具（Cohen, 1962, 1969）。例如，根据建议的定义，在针对两种方法之间的差异的标准检验中， 0.25σ 的差异很小， 0.50σ 的差异中等， 1σ 的差异很大。文书工作人员和半熟练工人之间的平均智商差异是中等效应。在一项设计巧妙的科研实践中，科恩回顾了《变态和社会心理学杂志》（*Journal of Abnormal and Social Psychology*）一卷中发表的所有统计分析，并计算了检测3种效应大小的可能性（Cohen, 1962）。小效应检测的平均效应度为0.18，中等效应检测的平均效应度为0.48，大效应检测的平均效应度为0.83。如果心理学家通常期望中等效应并像上述例子中一样选择样本大小，他们的研究效应度确实应该约为0.50。

科恩的分析表明，许多心理学研究的统计效应度低得离谱。这是一种弄巧成拙的做法：它使得科学家们受挫，研究效率低下。研究人员如果检验了一个有效的假设，但没有得到显著性结果，他们就会认为大自然

是不可信的，甚至是充满敌意的。此外，正如J. E.奥弗奥尔（J. E. Overall）所表明的那样，缺乏统计效度的研究的盛行不仅是种浪费而且实际上是有害的：它导致在已发表的结果中存在着大量的无效的拒绝零假设的结果（Overall, 1969）。

由于对统计效度的考虑在设计重复研究中特别重要，所以我们在调查问卷中探讨了关于重复的态度。

假设你的一位博士生完成了针对 40 只动物的艰难且耗时的实验。他对大量变量进行了评分和分析。他的结果通常是不确定的，但前后比较得到了一个显著性结果， $t = 2.70$ ，这是令人惊讶的，且可能具有重要的理论意义。

考虑到结果的重要性、其惊人的价值及该学生所进行的数据分析，你是否会建议他在发表之前重复该研究？如果你建议重复，会建议他在重复实验中测试多少只动物？

我们向一些心理学家提出了这些问题，大多数人给出的答案倾向于重复：75名被试中有66人建议重复，这可能是因为他们怀疑这个唯一的显著性结果是偶然出现的。中间意见是建议该博士生在重复研究中测试20只动物。这是因为他们认为这一建议可能带来的结果具有启发意义。如果第二个样本中的均值和方差实际上与第一个样本中的相同，那么得到的 t 值将是1.88。根据本章开头脚注中的推理，学生在重复实验中获得显著结果的机会仅略高于1/2（ $p=0.05$ ，单侧检验）。由于预计20个重复实验的样本对于被试来说似乎是合理的，我们添加了以下问题：

假设你的学生实际上已经用另外 20 只动物重复了首次研究，并且在同一方向上获得了非显著性结果， $t = 1.24$ 。你现在的建议是什么？请从下列选项中选择一个（括号中的数字指的是选择每个选项的被试人数）：

- A. 他应汇总结果并将其结论作为事实发表（0）
- B. 他应将结果作为暂定结果来发表（26）
- C. 他应该对另一组动物进行实验（中位数是 20）（21）
- D. 他应该尝试找出两组之间差异性的成因（30）

值得注意的是，无论人们对最初的发现是否有信心，它的可信度肯定会因为重复而得到提高。不仅两个样本的实验效果方向相同，而且重复效果的大小完全是初始研究的2/3。鉴于我们的被试给出了建议的样本量（20），重复研究应该会与人们期望的一样成功。然而，选择结果的分布显示，在进行推荐的重复研究后，被试对学生的发现仍持怀疑态度。这种不好的状况是统计效度不足的典型后果。

与某些方面合理的回答B和C相比，最普遍的回答D是站不住脚的。我们怀疑，如果被试意识到这两项研究之间的差异甚至不具有显著性，是否会得到相同的答案（如果两个样本的方差相等， t 为0.53）。在没有统计检验的情况下，我们的被试遵循了代表性假设：由于两个样本之间的差异大于他们的预期，因此他们认为应该得到相应的解释。然而，尝试“找出两组之间差异性的成因”很可能是在解释噪声。

总的来说，我们的被试对重复进行了相当严格的评估。这遵循了代表性假设：如果我们期望所有样本彼此间非常相似，那么几乎所有有效假设的复制版都应该具有统计学上的显著性。成功复制的苛刻标准体现在对下列问题的回答中：

一名研究人员公布了你认为难以置信的结果。他的实验样本为 15，并公布了一个显著值， $t = 2.46$ 。另一名研究人员试图复制他的步骤，被试数量相同但得到的 t 值并不显著。两组数据的方向相同。

你正在审阅相关文献。被你描述为复制失败的第二组数据中 t 的最大值是多少？

我们的大多数受访者认为 $t=1.70$ 是无法复制的。如果将这两个研究的数据（ $t=2.46$ 和 $t=1.70$ ）合并，那么合并后的数据的 t 值约为3.00（假设方差相等）。因此，我们面临着一种似是而非的情况，在这种情况下，同样的数据如果被视为原始研究的一部分，会增加我们对这一发现的信心，但如果被视为一项独立研究，就会动摇我们的信心。这种双重标准尤其令人不安，由于存在许多原因，重复实验通常被认为是独立的研究，而且通常会通过列出确认和否定报告来评估假设。

与普遍的看法相反，当重复实验的样本通常大于原始样本时，重复实验可以进行。对已获得的发现进行复制的决定，通常反映出人们对该发现的极大喜爱，以及希望看到它被持怀疑态度的群体所接受的态度。由于该群体不合理地要求复制具有独立的显著性，或者至少近似具有显著性，因此必须使用大样本进行重复实验。为了进行说明，假设之前讨论过的那位不幸的博士生的初始结果具有效度（ $t=2.70$ ， $N=40$ ），并且如果他愿意接受 t 低于1.70的风险（只有0.10），那么他应该在重复研究中使用大约50只动物作为样本。如果初始结果较弱（ $t=2.20$ ， $N=40$ ），获得相同效度所需的重复实验的样本量应上升为约75只。

到目前为止我们所讨论的效应不仅限于关于均值和方差的假设，被试针对以下问题的回答证明了这一点：

你进行了一项相关研究，是对 100 名被试进行 20 个变量的评分。190 个相关系数中有 27 个在 0.05 水平上显著；其中 9 个的显著性超出 0.01 水平。显著相关系数的平均绝对水平为 0.31，从理论上讲，这一结果的模式非常合理。在这项研究的精确复制中，在 $N = 40$ 时，你预计 27 个显著相关性中有多少会再次显著？

当 $N=40$ 时，0.05水平上的显著性需要约0.31的相关性。这是原始研究中显著相关的平均值。因此，只有大约一半的原始显著相关（即13或14）在 $N=40$ 时仍然是显著的。此外，当然复制过程中的相关性必然与原始研究中的不同。因此，通过回归效应，原始显著系数最有可能降低。因此，与原始的27相比，8到10个重复实验的显著相关可能是一个人预期的慷慨估计。我们的受访者的估计值中位数为18。这比从100名被试中随机抽取40名被试重新计算相关性时发现的重复实验显著相关数还要多！显然，人们期望的不仅仅是复制样本中的原始统计数据；他们

希望复制结果的显著性，且不用考虑样本量。这种预期需要对代表性假设进行扩展；即使是小数定律也无法产生这样的结果。

虽然在实践中备受谴责，但人们总是期望抽样数据结果在总体中都是可复制的，进而相信这是普适性的原理。计算出3个焦虑指数和3个依赖指数之间所有的相关关系的研究者通常会非常自信地报告和解释所得到的单一显著相关性。他对这一不可靠的发现充满了信心，这是因为他相信所得到的相关矩阵具有高代表性，并且易于复制。

在回顾中，我们已经看到支持小数定律的人以下述方法进行科学实践：

- 他在小样本上研究假设的博弈，却没有意识到不利的可能性很高。他过高地估计了效度。
- 他对早期趋势（如前几名被试的数据）和观察模式的稳定性（如重要结果的数量和一致性）过度自信。他高估了显著性。
- 在评估自己或他人的重复实验时，他对显著性结果的可复制性有着不合理的高度期待。他低估了置信区间的宽度。
- 他很少将结果与期望的偏差归因于抽样变异性，因为他找到了任意差异的因果“解释”。他几乎没有机会发现行动中的抽样变异。因此，他对小数定律的信念将永远不会改变。

我们的问卷调查得到了相当多的证据，证明了相信小数定律是普遍存在的。[\(27\)](#)我们的典型受访者就是小数定律的拥护者，不管他属于哪个群体。在数学心理学会议和美国心理学协会大会的一般会议上，观众的反应中位数实际上没有差别，尽管我们没有说明这两种样本的代表性。显然，熟悉形式逻辑和概率论并不能消除错误的直觉。那么，我们能做些什么呢？对小数定律的信奉能否被废除，或者至少能得到有效控制？

研究经验不太可能有很大帮助，因为抽样变异太容易得到“解释”了。矫正经验既不能为我们提供动机也不提供虚假解释的机会。因此，统计学课程中的学生可以从群体中抽取给定大小的重复实验的样本，并且从个人观察中了解样本量对抽样变异性的影响。然而，我们还不确定是否可以通过这种方式矫正期望，因为相关的偏见，比如赌徒谬

论，在很多相互矛盾的证据面前仍然存在。

即使偏见无法消除，学生也能学会认识到它的存在，并采取必要的预防措施。由于统计学的教学并不缺乏告诫，对有偏见的统计直觉的警告可能也不能被视为不合时宜。明显的预防措施是计算。相信小数定律的人对于显著性水平、效度和置信区间有着不正确的直觉。显著性水平通常会被计算和报告，但效度和置信限度不会。也许它们也应该如此。

相对于一些合理的假设，例如科恩的小、大、中等效应，明确的效应度计算在任何研究之前都必须进行（Cohen, 1962, 1969）。这样的计算通常会让人认识到，除非将样本量乘以4，否则研究根本没有意义。我们不相信一个严谨的研究人员会在知情的情况下接受0.50的风险，即无法证实一个有效的研究假设。此外，效度的计算对解释无法拒绝零假设的负面结果非常重要。由于读者对效度的直觉估计很可能是错误的，因此公布计算值似乎并不是在浪费读者的时间和期刊版面。

在早期心理学文献中，普遍采用的报告方法是样本均值为 $M \pm PE$ ，其中 PE 为可能误差（即均值附近50%的置信区间）。这个惯例后来被抛弃了，取而代之的是假设检验公式。然而，置信区间实际上为我们提供了一个有用的抽样变异指数，而人们往往会低估这种变异性。强调显著性水平往往会模糊效应的大小与其统计显著性之间的根本区别。不管样本量是多少，一项研究中效应的大小是对重复实验中效应大小的合理估计。相反，重复实验中估计的显著性水平主要取决于样本量。如果可以明确大小和显著性之间的区别，并且将观察到的效应的计算大小作为一个需要发表的常规性指标，则可以矫正关于显著性水平的可复制性的不切实际的前景。从这个角度来看，至少接受假设检验模型对心理学来说，并不完全是一件好事。

真正相信小数定律的人承认他的许多行为都违背了善意的统计推断的逻辑。代表性假设描述了认知或感知偏见，这种偏见与动机因素无关。因此，虽然仓促拒绝零假设令人满意，但拒绝一个含有希望的假设则令人恼火，而真正相信的人会同时受到这两种情况的影响。他的直觉期望取决于对世界的一贯误解，而不是机会主义的一厢情愿的想法。鉴于一些社会评论的影响，他可能愿意以适当的怀疑态度来看待自己的统计直觉，并在可能的情况下用计算来取代印象形成。

参考文献

Cohen, J. (1962). The statistical power of abnormal-social psychological research. *Journal of Abnormal and Social Psychology*, 65, 145–153.

Cohen, J. (1969). *Statistical power analysis in the behavioral sciences*. New York, NY: Academic Press.

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment* (pp. 17–52). New York, NY: Wiley.

Estes, W. K. (1964). Probability learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 89–128). New York, NY: Academic Press.

Overall, J. E. (1969). Classical statistical hypothesis testing within the context of Bayesian theory. *Psychological Bulletin*, 71, 285–292.

Tune, G. S. (1964). Response preferences: A review of some relevant literature. *Psychological Bulletin*, 61, 286–302.

第11章 证据的权重与自信程度的决定因素

戴尔·格里芬
阿莫斯·特沃斯基

证据的权重和信念的形成是人类思想的基本要素。哲学家和统计学家从规范的角度解决了如何评估证据和信心的问题；心理学家和决策研究人员也对其进行了实验研究。这些研究的一个主要发现是，事实证明人们对自己的判断往往过度自信了。过度自信不仅限于非专业判断或实验室实验中。众所周知的观察结果表明，超过2/3的小企业会在4年内失败（Dun & Bradstreet, 1967），这表明许多企业家高估了自己成功的概率（Cooper, Woo & Dunkelberg, 1988）。除了一些明显的例外，如接受即时且频繁的反馈并对降水进行实际预测的天气预报员（Murphy & Winkler, 1977）、医生（Lusted, 1977）、临床心理学家（Oskamp, 1965）、律师（Wagenaar & Keren, 1986）、谈判家（Neale & Bazerman, 1990）、工程师（Kidd, 1970）以及安全分析师（Staël von Holstein, 1972）的判断都存在过度自信的现象。正如一位评论家所描述的那样，专家预测“经常是错误的但很少有人怀疑”。

过度自信很常见，但并不普遍。校准研究发现，在一些非常简单的项目中过度自信被消除了，并经常观察到不自信的情况（Lichtenstein, Fischhoff & Phillips, 1982）。此外，序列更新的研究表明后验概率估计通常表现为保守或不自信（Edwards, 1968）。在本文中，我们研究了证据的权重，并提出了一个可用来解释文献中观察到的过度自信和不自信模式的说明(28)。

信心的决定因素

针对给定假设的信心或信任程度的评估通常需要对不同类型的证据进行整合。在许多问题中，有可能需要区分证据的强度或极端程度，以及其权重或预测的效度。当评估由以前的老师写的研究生推荐信时，我们可能希望考虑证据的两个不同方面：第一，这封信有多积极或热情？第二，作者的可信度和知识水平如何？第一个问题指的是证据的强度或

极端程度，而第二个问题指的是证据的权重或可信度。同样，假设我们希望评估“硬币偏向于正面而不是反面”这个假设的证据。在这种情况下，样本中正面的比例反映了所述假设的证据强度，样本量反映了这些数据的可信度。证据强度与其权重之间的区别和效果大小（如两种均值之间的差异）与其可靠性（如差异的标准误差）之间的区别密切相关。虽然并不总是能够将证据的影响分解为强度和权重各自的贡献，但是在许多情况下，它们可以独立地变化。强或弱的建议可能来自可靠或不可靠的来源，并且可以在小样本或大样本中观察到相同的正面比例。

统计理论和机会演算为强度和权重的结合制定了规则。例如，概率论规定了样本比例和样本大小应如何组合以确定后验概率。大量的关于在不确定状况下进行判断的实验文献表明，人们没有根据概率和统计规则将强度和权重结合起来。相反，直觉判断会受到现有证据能代表所述假设的程度的过度影响（Dawes, 1988; Kahneman, Slovic & Tversky, 1982; Nisbett & Ross, 1980）。如果人们单独依赖代表性而不考虑控制预测效度的其他因素，他们的判断（如被采访的人将成为一名成功的管理者）将仅取决于他们的印象的强度（如问题中的人与成功管理者的相似程度）。然而，在许多情况下，人们似乎并没有完全忽视这些因素。相反，我们认为人们会关注自己感知到的证据的强度，然后根据其权重做出一些调整。

在评估推荐信时，我们认为人们应该首先关注推荐信的真诚度，然后再考虑作者的知识有限性。类似地，当判断硬币是偏向正面还是偏向反面时，人们会关注样本中正面的比例，然后根据投掷次数调整其判断。由于这种调整通常是不充分的，与恰当的统计模型相比，证据的强度倾向于支配其权重（Slovic & Lichtenstein, 1971; Tversky & Kahneman, 1974）。此外，人们倾向于关注证据的强度会导致他们没有充分利用控制预测性效度的其他变量，如基础率和可辨性。这种处理结合了完全基于印象强度的代表性判断，以及考虑到了证据权重的锚定和调整过程，尽管不够充分。夸特隆研究了锚定在印象形成中的作用（Quattrone, 1982）。

这一假设暗示了一种独特的过度自信和不自信模式。如果人们对极端证据的变化高度敏感，而对其可信度或预测性效度的变化不够敏感，那么当强度高且权重低时，判断将表现出过度自信，并且当权重高和强度低时他们将不自信。如下所述，该假设有助于组织和总结大量关于不确定性判断的实验证据。

思考一下根据推荐信预测研究生申请成功的可能性的例子。如果人们主要关注推荐信的真诚度而不充分考虑作者的可信度或预测者和标准之间的相关性，当遇到基于偶然接触的炽热信件时，就会过度自信；当他们遇到适度积极但知识渊博的信件时，就会不自信。同样，如果人们对硬币偏差的判断主要取决于样本中正面和反面的比例而不充分考虑样本量，那么当他们在小样本中观察到极端比例时就会过度自信，而当他们在大样本中观察到适度的比例时就会不自信。

在本文中，我们检验了当强度高且权重低时过度自信，并且当权重高且强度低时不自信的假设。前3个实验关注统计假设的评估，其中证据强度由样本比例决定。在本章的第二部分，我们将这一假设扩展到了更复杂的证据问题中，并研究了其对信心判断的影响。

评估统计假设

研究1：样本量

我们首先研究样本比例（强度）和样本量（权重）在涉及后验概率评估的实验任务中的相对影响。我们对35名学生进行了如下指导：

假设你在旋转一枚硬币，并记录硬币正面朝上和反面朝上的出现次数。与抛掷（平均而言）产生相同数量的正面和反面不同，由于硬币边缘有轻微瑕疵（以及质量的不均匀分布），旋转硬币会导致偏向于一侧或另一侧。现在假设你知道这个偏差是 3/5，即 5 次中有 3 次会落在一侧，但你不知道这种偏差是偏向正面还是偏向反面。

然后我们向被试提供不同的证据样本，样本量不同（3~33），正面的数量也不同（2~19）。所有样本大多数都是正面，并且我们要求被试估计偏差偏向正面（*H*）而不是反面（*T*）的概率（0.5~1）。被试收到了表11-1所示的样本比例和样本量的所有12种组合。那些给出了最接近正确值的判断的人，会获得20美元奖金。

表11-1 研究1的刺激物和反馈数据

正面的数量 (h)	反面的数量 (t)	样本量 (n)	后验概率 $P(H D)$	信心中位数 (%)
2	1	3	0.60	63.0
3	0	3	0.77	85.0
3	2	5	0.60	60.0
4	1	5	0.77	80.0
5	0	5	0.88	92.5
5	4	9	0.60	55.0
6	3	9	0.77	66.9
7	2	9	0.88	77.0
9	8	17	0.60	54.5
10	7	17	0.77	59.5
11	6	17	0.88	64.5
19	14	33	0.88	60.0

表11-1还显示，对于每个数据样本（ D ），根据贝叶斯规则计算假设 H 的后验概率（偏向正面的概率为3:2）。假设先验概率相等，根据贝叶斯规则可得

$$\log\left(\frac{P(H | D)}{P(T | D)}\right) = n\left(\frac{h-t}{n}\right)\log\left(\frac{0.6}{0.4}\right)$$

其中 h 和 t 分别代表正面和反面的数量， $n=h+t$ 为样本量。右边的第一项 n 代表证据的权重。第二项是样本中正面和反面的比例之间的差异，代表支持 H 反对 T 的证据强度。第三项在本研究中保持不变，表示的是两个假设的可辨性，对应符号检验理论中的 d' 。在对数坐标中绘制强度和权重的等支撑线，得到一系列斜率为-1的平行直线，如图11-1中的虚线所示。为了便于解释，我们将强度维度定义为 h/n ，与 $(h-t)/n$ 线性相关。每一条线连接所有为假设 H 提供相同支持的数据集。例如，样本量为9（6个正面和3个反面）和样本量为17（10个正面和7个反面）的两个样本，在 H 和 T 上产生了相同的后验概率（0.77）。因此，两个点（9, 6/9）和（17, 10/17）都在上方的线上。类似地，下方的线连接的数据集产生的后验概率为0.60，偏向于 H （如表11-1所示）。

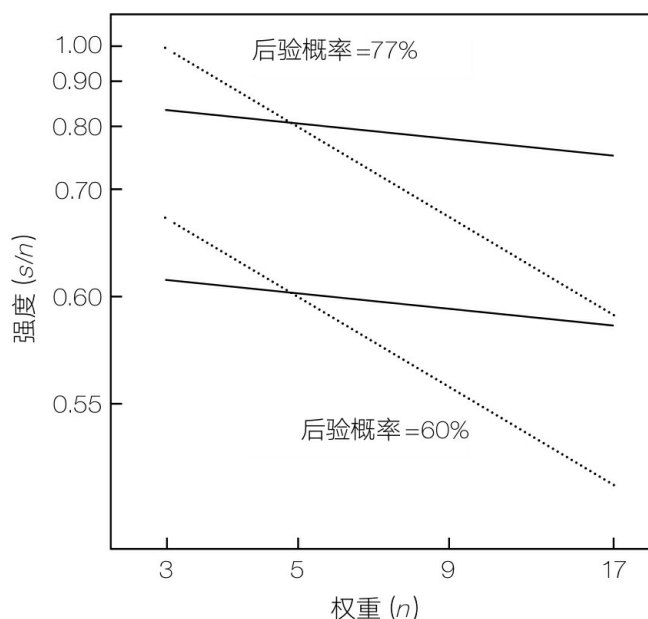


图11-1 强度和样本量的等支撑线

为了将观察到的判断与贝叶斯规则进行比较，我们首先将每个概率判断转换为对数概率，然后针对每名被试和中位数，分别将每名被试的这些值的对数与强度的对数 $(h-t)/n$ ，以及权重 n 的对数进行回归。回归结果与数据非常吻合：中位数的相关系数为0.95，中位被试的相关系数为0.82。根据贝叶斯规则，该指标中的强度和权重的回归权重相等（如图11-1所示）。相比之下，35名被试中有30名的强度回归系数大于权重回归系数（符号检验， $p < 0.001$ ）。在被试中，这些系数的中位数比率为2.2:1，有利于增加强度。[\(29\)](#)对于中位数数据，观察到的强度的回归权重（0.81）几乎是权重的3倍（0.31）。

回归分析得到的等支撑线为实线（如图11-1所示）。两组线条的对比高度揭示了两个值得注意的观察结果。首先，直觉线条比贝叶斯线条浅得多，这表明证据强度支配了它的权重。其次，对于给定的支持水平（如60%或77%），贝叶斯线与直觉线交叉表明强度高且权重低时会过度自信，以及强度低且权重高时会不自信。如后面所述，交叉点主要由竞争假设（ d' ）的可区辨性决定。

图11-2显示了作为两个不同样本量的（贝叶斯）后验概率的函数的给定证据样的中位自信程度。最佳拟合线是用对数概率尺度计算的。如果被试相信贝叶斯，那么实线将与虚线重合。相反，基于小样本（ $N=5$ ）的直觉判断表现出过度自信，而基于较大样本（ $N=17$ ）的判断

则表现出不自信。

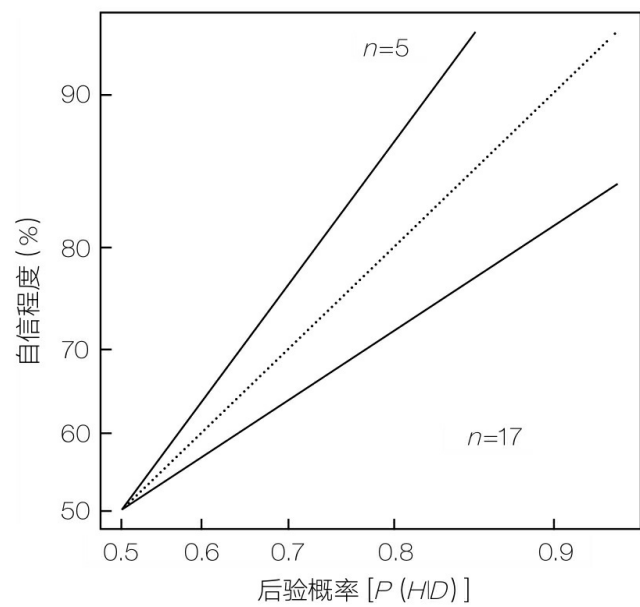


图11-2 样本量和自信程度

表11-1中描述的结果与之前证实的直觉判断是非规范性的结果基本一致（Kahneman, Slovic & Tversky, 1982; von Winterfeldt & Edwards, 1986）。此外，它们有助于调和明显不一致的发现。爱德华兹和他的同事（Edwards, 1968）使用了顺序更新范式，认为人们是保守的，因为他们没有从样本数据中提取足够的信息。此外，特沃斯基和卡尼曼研究了样本量在研究人员对其结果的可复制性的信心中的作用，得出的结论是，人们（即使是那些受过统计训练的人）也会根据小样本做出激进的推断（Tversky & Kahneman, 1971）。图11-1和图11-2表明了以样本比例而非样本量为主导的实验是如何产生两种结果的。在爱德华兹进行的一些升级实验中，研究人员向被试提供了中等强度的大量数据样本。这就是我们预期的不自信或保守主义的背景。此外，特沃斯基和卡尼曼研究的情况涉及基于相当小的样本的中等强度效应。在这种背景下，过度自信很可能会占上风。因此，保守和过度自信都可以通过证据权重的共同偏差产生，即强度重于权重。

如上文所述，关注证据强度的倾向使得人们忽视或降低了其他变量的权重，例如所讨论的假设的先验概率或竞争假设的可辨性。以下两项研究证实了这些影响。本节中的3项研究都采用了被试内设计，其中证据的强度和缓解变量（如样本量）在被试中各不相同。在这个过程中我

们可能低估了强度的主导作用，因为人们倾向于对研究中的任意控制变量做出反应，无论这样做是否符合规范（Fischhoff & Bar-Hillel, 1984）。实际上，在被试间的比较中，对样本量和基础率信息的忽略最为明显（Kahneman & Tversky, 1972）。

研究2：基础率

大量研究表明，在存在特定证据的情况下，人们往往会忽视背景数据，如基础率（Kahneman, Slovic & Tversky, 1982; Bar-Hillel, 1983）。这种疏忽可能导致不自信或过度自信。我们让40名学生想象他们有3种不同的外币，每种硬币的已知偏差是3:2。如研究1中那样，被试不知道每枚硬币偏向正面（*H*）还是偏向反面（*T*）。被试的两个假设（*H*和*T*）的先验概率是不同的。对于一半的被试，第一种类型硬币*H*的概率为0.50，第二种类型硬币*H*的概率为0.67，第三种类型硬币*H*的概率为0.90。对于另一半被试，这3种类型的硬币的*H*的先验概率为0.50、0.33和0.10。研究人员向被试呈现了一个样本量为10的样本，其中包括5~9个正面。然后要求他们给出自己对正在考虑的硬币偏向正面的信心（%）。同样，给出的判断最接近正确值的人可获得20美元奖金。表11-2总结了样本数据，每个样本的后验概率及被试信心判断的中位数。很明显，我们的被试高估了证据的强度，而低估了先验概率。

表11-2 研究2的刺激物和反馈数据

每掷 10 次硬币 正面朝上的次数	先验概率（基础率）	后验概率 $P(H D)$	自信程度中位数 (%)
5	9 : 1	0.90	60.0
6	9 : 1	0.95	70.0
7	9 : 1	0.98	85.0
8	9 : 1	0.99	92.5
9	9 : 1	0.996	98.5
5	2 : 1	0.67	55.0
6	2 : 1	0.82	65.0
7	2 : 1	0.91	71.0
8	2 : 1	0.96	82.5
9	2 : 1	0.98	90.0
5	1 : 1	0.50	50.0
6	1 : 1	0.69	60.0
7	1 : 1	0.84	70.0
8	1 : 1	0.92	80.0
9	1 : 1	0.96	90.0
5	1 : 2	0.33	33.0
6	1 : 2	0.53	50.0
7	1 : 2	0.72	57.0
8	1 : 2	0.85	77.0
9	1 : 2	0.93	90.0
5	1 : 9	0.10	22.5
6	1 : 9	0.20	45.0
7	1 : 9	0.36	60.0
8	1 : 9	0.55	80.0
9	1 : 9	0.74	85.0

图11-3显示的是（贝叶斯）后验概率函数 H 对高先验概率（0.90）和低先验概率（0.10）的自信程度中位数的判断。该图还显示了每种情况的最佳拟合线。从图中我们可以明显看出，被试在低基础率的条件下过度自信，在高基础率的条件下不自信。

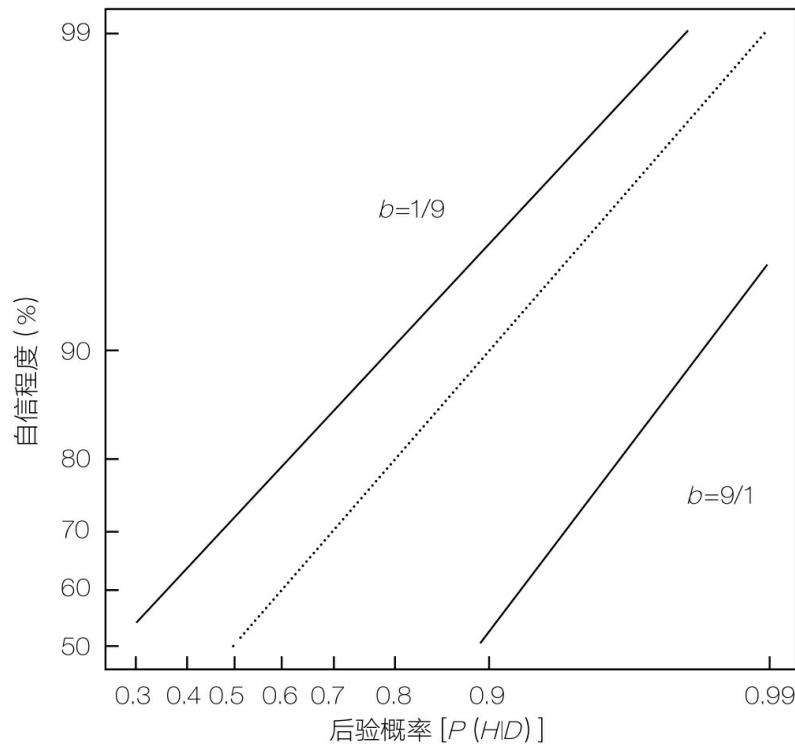


图11-3 基础率和自信程度

这些结果与D. M.格雷瑟（D. M. Grether）关于代表性启发式在后验概率判断中的作用的研究结果一致（Grether, 1980, 1990）。本研究的先验概率和数据均以数值的形式呈现，与本研究不同的是，格雷瑟的程序涉及从摇奖机中随机抽取有编号的球。他发现正如代表性所暗示的那样，被试高估了相对于先验概率的似然比，并且金钱激励的减少并没有消除忽视基础率的现象。格雷瑟的研究结果与C.卡默勒（C. Camerer）在其对市场交易的广泛研究中发现的结果一样，但与G.吉格瑞泽（G. Gigerenzer）、W.赫尔（W. Hell）和H.布兰克（H. Blank）的主张相矛盾，即明确的随机抽样消除了忽略基础率的现象（Camerer, 1990; Gigerenzer, Hell & Blank, 1988）。戴尔·格里芬发表了单独使用明确的随机抽样不会降低忽略基础率的现象的证据（Griffin, 1991）。

我们的分析表明，当基础率较低时人们容易过度自信，而当基础率

较高时，则会不自信。D.邓宁（D. Dunning）、格里芬、J.米洛伊科维奇（J. Milojkovic）和L.罗斯（L. Ross）在社会预测研究中观察到了这种模式（Dunning, Griffin, Milojkovic & Ross, 1990）。在他们的研究中，每名被试在对目标人物的偏好和行为做出预测之前采访了目标人物，采访内容为：“如果可以免费订阅，他会选择哪本杂志，《花花公子》（*Playboy*）还是《纽约书评》（*The New York Review of Books*）？”研究人员向每名被试提供了基于经验得出的关于问题的回答的预测基础率；如68%的受访者更喜欢《花花公子》。为了研究经验基础率的影响，邓宁等人分别分析了与基础率一致的预测（即“高”准确率预测）和与基础率相反的预测（即“低”基础率预测）。当基础率较低（自信程度=72%，准确度=49%）时，过度自信的情况比基础率较高时（自信程度=79%，准确度=75%）明显得多。此外，对于基础率超过75%的项目，被试的预测实际上是不自信的。这正是这一假设所暗含的模式，即被试根据对目标人物的印象来评估其更喜欢《花花公子》而不是《纽约书评》的概率，很少或根本没有考虑经验基础率，即这两本杂志在目标人群中的相对受欢迎程度。

研究3：可辨性

当我们考虑两个假设中哪个是正确的这一问题时，自信程度应取决于数据在多大程度上更符合一个假设。然而，人们似乎只关注给定假设的证据强度，而忽视了同样的证据与另一个假设的匹配程度。巴纳姆效应（Barnum effect）就是一个很好的例子。构建一个令许多人印象深刻的人格素描来非常准确地描述个人特征很容易，因为人们会通过与其个性相符的程度来评估描述，而不关心它是否也适用于其他人（Forer, 1949）。为了在机会设置中探究这种效应，我们向50名学生提供了与两种外国硬币相关的证据。在每种类型的硬币中，证据强度（样本比例）从7/12的正面到10/12的正面不等。这两种硬币的特征不同。被试被要求：

假设你在旋转一枚名为昆塔（quinta）的外国硬币。假设一半昆塔（X型）对正面有0.6的偏好（也就是说，X-昆塔在旋转时正面朝上的概率为60%）和一半的昆塔（Y型）对反面有0.75的偏好（也就是说，Y-昆塔在旋转时反面朝上的概率为75%）。你的任务是确定这是X-昆塔还是Y-昆塔。

然后他们收到了表11-3中所示的证据样本。在给出每个样本是来自X-昆塔或Y-昆塔的信心后，被试被要求对A-libnars（对正面有0.6的偏好）和B-libnars（正面出现的概率为0.5）做出同样的判断。硬币的呈现顺序是平衡的。

表11-3 研究3的刺激物和反馈数据

正面数量 (12个中)	析取假设 (d')	后验概率 $P(H D)$	自信程度中位数 (%)
7	0.6 vs 0.5	0.54	55.0
8	0.6 vs 0.5	0.64	66.0
9	0.6 vs 0.5	0.72	75.0
10	0.6 vs 0.5	0.80	85.0
7	0.6 vs 0.25	0.95	65.0
8	0.6 vs 0.25	0.99	70.0
9	0.6 vs 0.25	0.998	80.0
10	0.6 vs 0.25	0.999	90.0

样本数据的结果如表11-3所示，其中包括每个样本的后验概率和被试自信程度判断的中位数。将自信程度判断与贝叶斯后验概率进行的比较表明，我们的被试主要关注数据与偏好假设的匹配程度，而没有充分考虑数据与备选假设的匹配程度（Fischhoff & Beyth-Marom, 1983）。被试对贝叶斯后验概率的中位数的自信程度判断如图11-4所示，可用于比较高、低可辨性。当假设之间的可辨性较低时（当硬币的偏好为0.6或0.5时），被试略显过度自信。当假设之间的可辨性较高时（当偏好为0.6或0.25时），被试非常不自信。

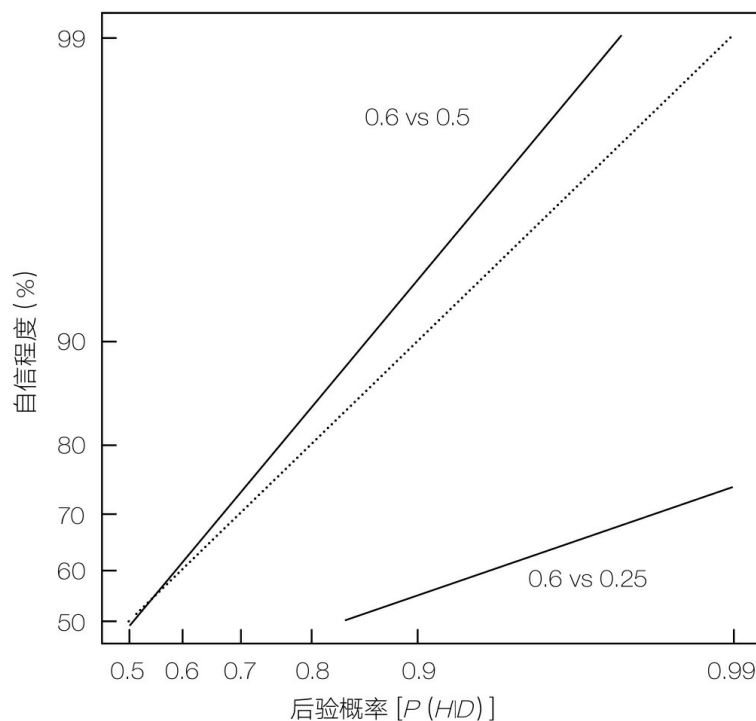


图11-4 可辨性和自信程度

在早期关于后验概率判断的文献中，大多数实验研究了高度可辨性的对称假设（如3:2对2:3）并且发现了一致的不自信（Peterson, Schneider & Miller, 1965）。然而，与我们的假设相一致的是，在对低可辨性假设的研究中发现了过度自信。例如，C. R.彼得森（C. R. Peterson）和A. J.米勒（A. J. Miller）发现，当各自的比率分别为3:2和3:4时，在后验概率判断中存在过度自信，L. D.菲利普斯（L. D. Phillips）和爱德华兹在比率为11:9和9:11时观察到了不自信现象（Peterson & Miller, 1965; Phillips & Edwards, 1966）。

对知识的信心

上文表明，人们对证据的强度比对其权重更敏感。因此，当强度高且权重低时，人们会过度自信，而当强度低且权重高时，人们会不自信。我们认为，这个结论不仅适用于随机过程（如硬币旋转）的判断，还适用于对不确定性事件的判断，如谁将在即将到来的选举中获胜，或某本书是否会进入畅销书排行榜。当人们评估此类事件的可能性时，我们认为他们评估的是自己对候选人或书籍的印象。这些印象可能基于随机观察或对选民和读者偏好的广泛了解。与机会设置类似，可以将印象的极端性与样本比例进行比较，而印象的可信度可以与样本量或者竞争

假设的可辨性相对应。如果人们只关注印象的强度而对权重认识不足，那么在机会过程的评估中观察到的过度自信和不自信的模式也应该存在于非统计证据的评估中。

在本节中，我们将这一假设扩展到了复杂的证据问题上。在这些问题上，强度和权重无法轻易被定义。我们首先比较了自己和他人的预测。之后，我们论述了目前的叙述是如何产生“难度效应”的。最后，我们探讨了常识问题中自信程度的决定因素，并将自信程度差异与效度错觉联系了起来。

研究4：自我与他人

在这项研究中，我们要求人们预测自己的行为（对此他们可能了解很多），以及他人的行为（他们对此知之甚少）。如果人们的自信程度主要取决于自己的印象强度而没有充分考虑其权重，那么我们预计对他人的预测比对自我的预测过度自信的情况会更加严重。

在一项涉及风险的任务中，14对彼此不认识的同性学生被要求对彼此的行为进行预测。研究人员让他们先进行5分钟的面谈，然后坐在各自的电脑前，在这个名为“公司丛林”的囚徒困境游戏中预测自己和同伴的行为。在每次实验中，被试都可以选择将他们的公司与同伴的公司“合并”（即合作），或“接管”其同伴的公司（即竞争）。如果一方试图合并，另一方试图接管，那么合作并购就会遭受严重损失，而企业掠夺者则会获得可观的收益。但是，如果两个人都试图在同一个实验中接管，那么他们都会遭受损失。本次实验中有20个收益矩阵，其中有一些旨在鼓励合作，另一些则旨在鼓励竞争。

被试被要求预测自己在10个收益矩阵中的行为，以及预测他们采访的人在另外10个收益矩阵中的行为。这两个任务的顺序是平衡的，每个收益矩阵在每个任务中出现的次数相同。除了预测每个矩阵的合作或竞争外，被试还要表明他们对每个预测的自信程度（50%~100%）。在完成预测任务后不久，被试与对手进行了20次实验但不进行反馈，并根据20次实验的结果获得报酬。

该分析基于完成整个任务的25名被试。总体而言，被试对自我的预测（ $M=84\%$ ）和对他人的预测（ $M=83\%$ ）几乎同样有信心，但他们对自我的行为的预测（ $M=81\%$ ）比对他人的行为的预测（ $M=68\%$ ）更准

确。因此，人们对他人的预测表现出了很强的过度自信，但在预测自己时却相对更为精准（如图11-5所示）。

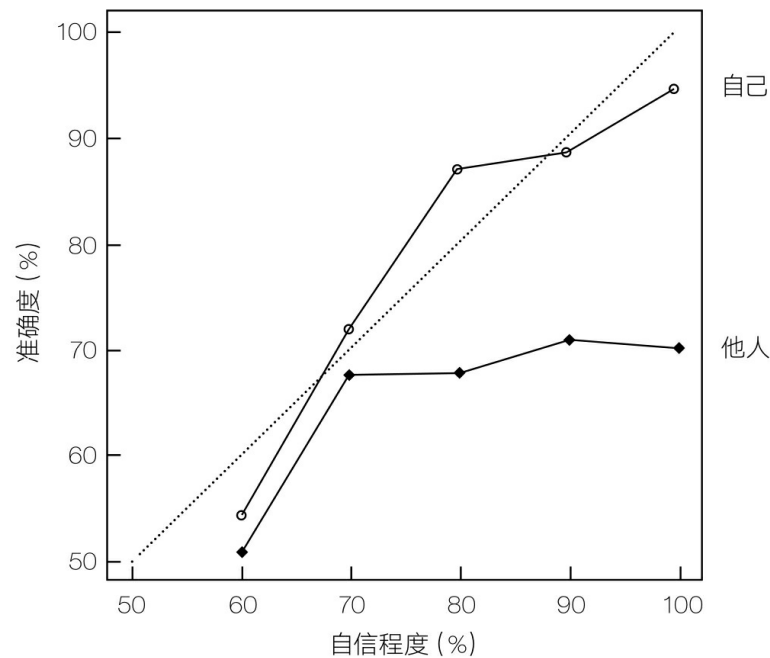


图11-5 对自己和他人的预测

在某些情况下，如果证据强度不极端，那么人们在对自己的行为进行预测时可能会不自信。例如，在选择工作的情况中，如果一个人有充分的理由去接受工作A并且也有充分的理由接受工作B，那么可能会存在不自信，但是他没有意识到，即使工作A只比工作B有一点小优势，通常也会导致其选择工作A。如果选择A而不是B的信心反映了两个工作的论证的平衡，那么2比1的平衡将产生大约2/3的信心，虽然选择A而不是B的概率可能更高（Koriat, Lichtenstein & Fischhoff, 1980）。在过去几年中，我们谨慎地与面临工作机会选择的同事接洽，并让他们估计自己选择一份工作而不是另一份的概率。对预测结果的平均自信程度为66%，但24名受访者中只有1人选择了最初给出较低概率的选择，总体准确率为96%。值得注意的是，在某些情况下，人们甚至在预测自己的行为时也表现出了过度自信（Vallone, Griffin, Lin & Ross, 1990）。因此，关键变量不是预测的目标（自己或他人），而是现有证据的强度和权重之间的关系。

对他人行为的预测充满信心，而对自己行为的预测却缺乏信心，这一倾向对决策分析具有耐人寻味的意义。决策分析师通常会区分受决策

者控制的决策变量和不受决策者控制的状态变量。通过确定决策变量的值（即决定你想要什么的价值）并将概率分配给状态变量（如他人的行为）来进行分析。一些决策分析师已经指出，他们的客户通常希望走相反的道路：在确定情况下，确定或预测他人的行为并为自己的选择分配概率。毕竟，他人的行为应该是可以根据他们的特点、需求和兴趣信息来预测的，而我们自己的行为是高度灵活的，且取决于不断变化的环境（Jones & Nisbett, 1972）。

难度效应

前面的分析表明，人们评估自己对两个相互竞争的假设中的一个的信心是基于他们对支持和反对这个假设的论据的平衡，而没有充分考虑数据的质量。当人们基于有限的知识而形成强烈的印象时，这种判断模式会导致过度自信；而当人们基于大量的数据形成适度的印象时，这种判断模式会导致不自信。

由于研究1至研究3中的强度和权重不能通过实验来控制，因此将该分析应用于常识问题是很复杂的。然而，与随机设置相似，我们假设对给定知识问题参数的平衡可以用样本中红球和白球的比例来表示。问题的难度可以通过两个假设的可辨性来表示，即在两个竞争假设中获得红球的概率之间的差异。当然，差异越大，任务越容易，即基于任意给定样本的更可能假设的后验概率越高。假设信心来自参数的平衡，即样本中红球的比例。这个模型预测的结果模式是什么？

3对假设的预测结果（样本量为10，如图11-6所示）定义了3个级别的任务难度：一项“容易”的任务，即在竞争性假设下获得红球的概率分别为0.50和0.40；一项“困难”的任务，概率为0.50和0.45；以及一个“不可能完成”的任务，即在两个假设下抽出红球的概率为0.5。我们为示例选择了非对称假设，以允许在校准数据中存在经常会观察到的初始偏差。

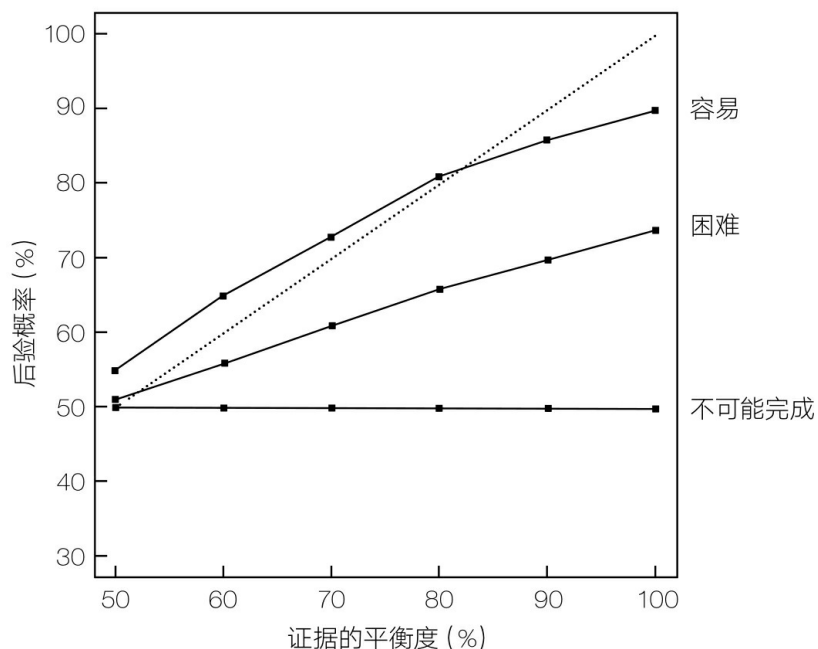


图11-6 项目难度的预测校准

将该模型的预测结果与利希滕斯坦和巴鲁赫·菲施霍夫（Baruch Fischhoff）针对任务难度效应的研究结果进行比较是具有指导意义的（Lichtenstein & Fischhoff, 1977；如图11-7所示）。他们的“简单”项目（准确率=85%）在大部分的置信范围内产生了不自信，而“困难”项目（准确度=61%）在大部分的置信范围内产生了过度自信，以及“不可能”项目（区分欧洲人和美国人的笔迹，准确率=51%）在整个范围内显示出了明显的过度自信。

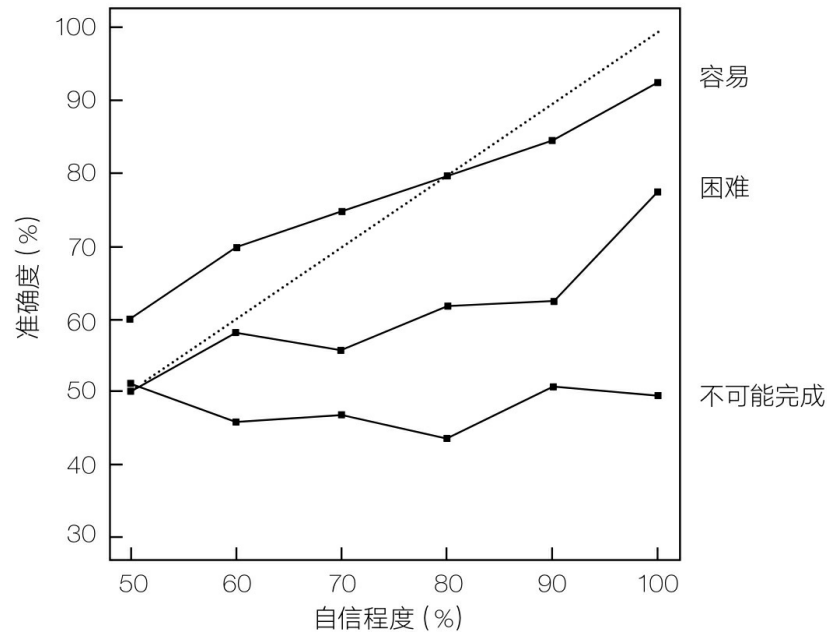


图11-7 项目难度校准

对比图11-6和图11-7可以发现，我们的简单机会模型再现了利希滕斯坦和菲施霍夫观察到的结果模式：人们对非常简单的项目有轻微的不自信，对困难的项目有持续的过度自信，对“不可能”的项目有明显的过度自信（Lichtenstein & Fischhoff, 1977）。这种模式源于这样一个假设，即对判断的自信程度是由竞争性假设的论据平衡所控制的。因此，这一说法可以解释所观察到的任务难度与过度自信之间的关系（Ferrell & McGoey, 1980）。

难度效应是校准文献中最一致的发现之一（Lichtenstein & Fischhoff, 1977; Lichtenstein, Fischhoff & Phillips, 1982; Yates, 1990）。它不仅存在于常识问题中，也存在于临床诊断（Oskamp, 1962）、未来事件的预测（Fischhoff & MacGregor, 1982; Wright & Wisudha, 1982）和文字识别（Keren, 1988）中。此外，难度效应可能有助于以不同方式解释其他发现。例如，G.克伦（G. Keren）的研究表明，世界级桥牌玩家的预测很准确，而业余玩家则过度自信（Keren, 1987）。克伦将这一发现解释为业余玩家的乐观偏好。此外，在预测桥牌结果方面，专业人士比业余爱好者更准确，而难度的差异可能是导致过度自信的原因之一。

难度效应也可以解释吉格瑞泽、U.霍夫雷奇（U. Hoffrage）和H.克莱因博尔汀（H. Kleinbolting）研究的主要发现（Gigerenzer,

Hoffrage, Kleinbolting, 1991)。在这项研究中，一组被试看到了成对的城市，研究人员要求他们选出人口较多的城市，并说明自己对每个答案的自信程度。这些城市是从所有德国西部的大城市名单中随机选择的。第二组被试被要求回答一些常识问题，如拉链是在1920年之前还是之后发明的？他们还被要求选择正确的答案，并评估自己对答案的自信程度。被试对城市人口的判断相当准确，但对常识问题的回答显示出了过度自信。然而，这两项任务的难度并不相同：城市人口判断的平均准确率为72%，常识问题判断的平均准确率仅为53%。因此，根据利希滕斯坦和菲施霍夫的记录，在后者中存在过度自信而在前者中不存在，可能完全是受到了难度的影响（Lichtenstein & Fischhoff, 1977）。事实上，当吉格瑞泽等人选择了一组难度可与常识问题匹配的城市问题时，被试在这两个领域中产生了同种程度的过度自信（Gigerenzer et al., 1991）。作者没有承认他们的研究混淆了项目生成（代表性与选择性）和任务难度（容易与困难）因素。相反，他们把自己的数据解释为对他们提出的理论的支撑证据，即对个人判断的过度自信是项目选择的结果，而当项目是从一些自然环境中随机抽样得到的时，这种自信就会消失。该预测在以下研究中得到了检验。

研究5：关于效度的错觉

在这个实验中，研究对象比较了1990年《世界年鉴》（*World Almanac*）中报告的美国各州的几个属性。为了保证抽样具有代表性，我们从所有可能的州中随机抽取了30对美国的州。研究人员向被试展示了成对的州（如亚拉巴马州、俄勒冈州），要求他们选择某一特定属性得分较高的州，并评估自己的答案正确的概率。根据吉格瑞泽等人的说法，这些判断中应该不存在过度自信，因为这些被选州是从自然参考类别中随机选择的（Gigerenzer et al., 1991）。相反，我们的研究表明，过度自信的程度取决于证据的强度和权重之间的关系。更具体地说，当证据权重低且证据强度高时，过度自信最为明显。这可能出现在人们可以很容易形成强烈印象的领域，即使这些印象具有较低的预测性效度。例如，面试官可以形成对未来研究生的心理素质的强烈印象，即使这些印象不能用于预测候选人的表现（Dawes, 1979）。

自然刺激物的使用可以排除强度和权重的直接操控。相应地，我们使用了3个属性，这些属性在被试可能形成的印象强度和可能具有的知识量方面有所不同。这3个属性分别是每个州的人数（人口）、高中毕业率（教育），以及每个州最后两次总统选举之间的投票率差异（投

票)。我们假设这3个属性会产生不同的自信程度和准确度。首先，我们预计人们对人口的了解比对教育或投票的了解更多。其次，我们预计对教育预测的自信程度大于对投票预测的自信程度，因为人们对各州的形象或刻板印象与前者的联系比后者更紧密。例如，如果一个州有更多的名牌大学，或者与更多的文化活动有关，人们可能会认为这个州的人比另一个州的人受教育程度更高。然而，由于这些线索与高中毕业率之间的相关性非常低，所以我们预计对教育问题的过度自信会比人口或投票问题更严重。因此，我们预测：被试对人口问题的预测的准确度和自信程度都高，对投票问题的预测的准确度和自信程度都低，对教育问题的预测准确度低但自信程度高。

为了检验这些假设，我们组织了298名被试各自评估成对州中一半的州（15个）的一个属性。在被试给出自己对15个问题的自信程度之后，他们被要求估计15个问题中有多少个回答是正确的。我们提醒被试：仅凭偶然性预估正确答案的值为7.5。

被试对3个属性中的每个属性的答案是正确的自信程度、准确度和估计频率的平均判断如表11-4所示，他们对所有3个属性的自信程度判断都表现出了显著的过度自信（ $p<0.01$ ），这与“如果从自然环境中随机抽取一组常识性任务，过度自信将为零”的说法相矛盾（Gigerenzer et al., 1991）。显然，过度自信的原因远不止带有偏好的选择。

表11-4 研究6中的自信程度和准确度数据

	人口 N = 93	投票 N = 77	教育 N = 118
自信程度	74.7	59.7	65.6
准确度	68.2	51.2	49.8
自信程度 - 准确度的差异	6.5	8.5	15.8
频率	51.3	36.1	41.2

我们在实验中观察到的自信程度和准确度模式与先前的假设一致，如图11-8所示。图中显示了3个属性的平均准确度及所有被试和项目的平均自信程度。在人口属性中，人们表现出相当高的准确度和适度的过度自信。对于投票属性来说，准确度是偶然的，但过度自信也是适度的。对于教育属性来说，准确度也是偶然的，但过度自信很严重。

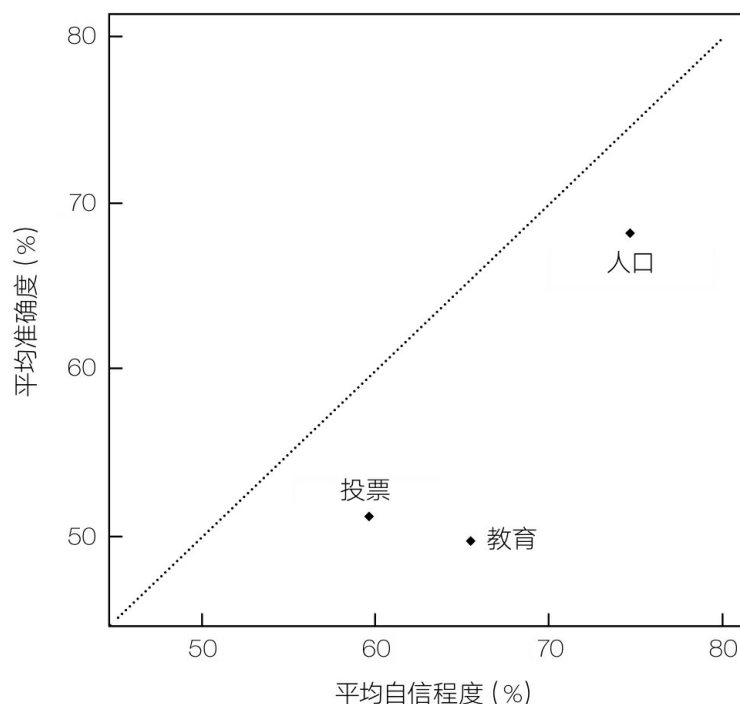


图11-8 3个属性的自信程度和准确度

目前的结果表明，过度自信无法通过难度效应得以充分解释。尽管准确度的差异很大（68.2对51.2， $p < 0.001$ ），但在人口和投票问题中产生了类似的过度自信的水平（6.5对8.5， $t < 1$ ， p 不显著）。此外，教育判断中的过度自信远远超过投票判断（15.8对8.5， $p < 0.01$ ），尽管他们的准确度几乎相同（49.8对51.2， $t < 1$ ， p 不显著）。

这种分析可以揭示过度自信与专业知识之间的关系。当可预测性非常高时，专家通常比非专业人员更准确。对比赛赔率的制订者

（Griffin, 1949; Hausch, Ziemba & Rubinstein, 1981; McGlothlin, 1956）和专业桥牌玩家（Keren, 1987）的研究与这一结论是一致的。然而，当可预测性非常低时，专家可能比新手更容易过度自信。如果无法通过现有数据预测精神病患者、俄罗斯经济或股票市场的未来状态，那么拥有丰富的系统模型的专家比那些对其了解非常有限的非专业人士更容易表现出对这些系统的过度自信。针对临床心理学家（Oskamp, 1965）和股票市场分析师（Yates, 1990）的研究结果与该假设一致。

频率与自信程度

现在我们来看看人们对个人答案效度的自信程度与他们对总体命中

率的估计之间的关系。例如，体育广播员可以被要求评估他对每场比赛的预测的自信程度，以及他希望正确预测的比赛数量。根据目前的说法，这些判断预计是不一致的，因为它们是根据不同的证据做出的。我们认为，对特定情况的自信程度的判断主要取决于支持和反对特定假设的论据之间的平衡，如两个对立团队的相对强弱。此外，对正确预测的频率进行估计可能要基于对任务难度、判断者的知识或在类似问题中的经验的一般评估。因此，在自信程度的平均判断中观察到的过度自信不一定适用于对预期准确度的整体判断。实际上，表11-4显示，估计的频率基本上低于正确预测的实际频率。事实上，后一种估计低于3种属性中的两种。(30)其他研究者也观察到了类似的结果（Gigerenzer et al., 1991; May, 1986; Sniezek & Switzer, 1989）。显然，即使知道自己的总体命中率不是很高，人们也可以对特定答案的效度保持高度自信。(31)这种现象被称为“效度错觉”：人们经常根据不可靠的数据（如个人访谈或投射测验）来对个案做出有信心的预测，即使他们知道这些数据的预测性效度较低（Kahneman & Tversky, 1973; Dawes, Faust & Meehl, 1989）。

频率估计与自信程度判断之间的差异是一个有趣的发现，但它并没有削弱人们在个别项目中过度自信的显著性。后一种现象很重要，因为人们的决定通常是基于他们对个别事件评估的信心，而不是对总体命中率的估计。例如，对新生企业家进行的广泛调查显示，平均而言，企业家对其特定的新企业的成功非常乐观（即过度自信），即使他们很清楚这类企业失败的一般规律（Cooper, Woo & Dunkelberg, 1988）。我们认为，进行新投资的决定主要是基于对个别事件的信念，而不是基于整体的基础率。倾向于选择个人观点或“内部”观点，而不是具有统计学意义的或“外部”观点，这是直觉判断与规范理论的主要背离之一（Kahneman & Lovallo, 1991; Kahneman & Tversky, 1982）。

还要注意人们在频率任务上的表现还有很多不足之处。平均而言，频率判断中的低估程度与个体概率判断的过度自信程度相当（表11-4）。此外，对于所有3个属性，被试的估计频率和实际频率之间的相关性可忽略不计（人口为+0.10，投票为-0.10，教育为+0.15）。这些观察结果不支持人们能正确估计其命中率的观点，并且自信程度-频率的差异仅仅是他们无法评估独特事件概率的表现。关于过度自信的研究一直受到一些作者的批评，因为它将一个频率标准（正确预测的比率）应用于非常见或主观的概率概念。然而，这个反对意见忽略了贝叶斯模式有可能会被校准的事实，因此，主观概率理论允许人们对自信程度和准确

度进行比较（Dawid，1982）。

结束语

前面的研究表明，在校准实验中观察到的过度自信并不是项目选择的产物，也不是测试难度的副产品。此外，过度自信并不局限于对离散事件的预测，它也经常出现在不确定数量的评估中（Alpert & Raiffa，1982）。

过度自信对处理人际事务非常重要。虽然过度自信并不普遍，但它很流行，往往是大规模的，且很难消除（Fischhoff，1982）。这一现象之所以重要，不仅因为它证明了直觉判断与机会定律之间的差异，而且主要是因为自信程度会控制行动（Heath & Tversky，1991）。有人认为，过度自信像乐观主义一样是适应性的，因为它使人们感觉良好，并促使他们去做一些原本不会做的事情（Taylor & Brown，1988）。然而，这些益处可能是以高代价换来的。在为患者做出诊断时，过度自信可能导致不恰当的医疗措施，对实验结果的过度自信可能导致糟糕的法律建议，对预期利润的过度自信可能导致失败的金融投资。可以说，如果人们对自己成功的机会会有一个更现实的评估，他们参与军事、法律和其他代价高昂的战斗的意愿就会降低。我们对过度自信的益处是否大于代价表示怀疑。

致谢

感谢罗宾·道斯、巴鲁赫·菲施霍夫和丹尼尔·卡尼曼的帮助。

参考文献

Alpert, M., & Raiffa, H. (1982). A progress report on the training of probability assessors. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 294–305). Cambridge, England: Cambridge University Press.

Bar-Hillel, M. (1983). The base rate fallacy controversy. In R. W. Scholz (Ed.), *Decision making under uncertainty* (pp. 39–61). Amsterdam, the Netherlands: North-Holland.

Camerer, C. (1990). Do markets correct biases in probability judgment? Evidence from market experiments. In L. Green & J. H. Kagel (Eds.), *Advances in behavioral economics* (Vol. 2, pp. 126–172). Santa Barbara, CA: Praeger.

Cooper, A. C., Woo, C. Y., & Dunkelberg, W. C. (1988). Entrepreneurs' perceived chances for success. *Journal of Business Venturing*, 3, 97–108.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, 34, 571–582.

Dawes, R. M. (1988). *Rational choice in an uncertain world*. New York, NY: Harcourt Brace Jovanovich.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668–1674.

Dawid, A. P. (1982). The well-calibrated Bayesian. *Journal of the American Statistical Association*, 77, 605–613.

Dun & Bradstreet. (1967). *Patterns of success in managing a business*. New York, NY: Dun & Bradstreet.

Dunning, D., Griffin, D. W., Milojkovic, J., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology*, 58, 568–581.

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment* (pp. 17–52). New York, NY: Wiley.

Ferrell, W. R., & McGoey, P. J. (1980). A model of calibration for subjective probabilities. *Organizational Behavior and Human Performance*, 26, 32–53.

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422–444). New York, NY: Cambridge University Press.

Fischhoff, B., & Bar-Hillel, M. (1984). Focusing techniques: A shortcut to improving probability judgments? *Organizational Behavior and Human Performance*, 34, 175–194.

Fischhoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review*, 90, 239–260.

Fischhoff, B., & MacGregor, D. (1982). Subjective confidence in forecasts. *Journal of Forecasting*, 1, 155–172.

Forer, B. (1949). The fallacy of personal validation: A classroom demonstration of gullibility. *Journal of Abnormal and Social Psychology*, 44, 118–123.

Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 513–525.

Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506–528.

Grether, D. M. (1980). Bayes' rule as a descriptive model: The representativeness heuristic. *Quarterly Journal of Economics*, 95, 537–557.

Grether, D. M. (1990). *Testing Bayes' rule and the representativeness heuristic: Some experimental evidence* (Social Science Working Paper 724). Pasadena, CA: Division of the Humanities and Social Sciences, California Institute of Technology.

Griffin, D. W. (1991). *On the use and neglect of base rates*. Unpublished manuscript, Department of Psychology, University of Waterloo.

Griffith, R. M. (1949). Odds adjustments by American horse-race bettors. *American Journal of Psychology*, 62, 290–294.

Hausch, D. B., Ziemba, W. T., & Rubinstein, M. (1981). Efficiency of

the market for racetrack betting. *Management Science*, 27, 1435–1452.

Heath, F., & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4, 5–28.

Jones, E. E., & Nisbett, R. E. (1972). *The actor and the observer: Divergent perceptions of the causes of behavior*. Morristown, NJ: General Learning Press.

Kahneman, D., & Lovallo, D. (1994). Timid decisions and bold forecasting: A cognitive perspective on risk taking. In R. Rumelt, P. Schendel, & D. Teece (Eds.), *Fundamental issues in strategy* (pp. 71–96). Cambridge, MA: Harvard Business School Press.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430–454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237–251.

Kahneman, D., & Tversky, A. (1982). Intuitive prediction: Biases and corrective procedures. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 414–421). Cambridge, England: Cambridge University Press.

Keren, G. (1987). Facing uncertainty in the game of bridge: A calibration study. *Organizational Behavior and Human Decision Processes*, 39, 98–114.

Keren, G. (1988). On the ability of monitoring non-veridical perceptions and uncertain knowledge: Some calibration studies. *Acta Psychologica*, 67, 95–119.

- Kidd, J. B. (1970). The utilization of subjective probabilities in production planning. *Acta Psychologica*, 34, 338–347.
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology. Human Learning and Memory*, 6, 107–118.
- Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? The calibration of probability judgments. *Organizational Behavior and Human Performance*, 20, 159–183.
- Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). Cambridge, England: Cambridge University Press.
- Lusted, L. B. (1977). *A study of the efficacy of diagnostic radiologic procedures: Final report on diagnostic efficacy*. Chicago, IL: Efficacy Study Committee of the American College of Radiology.
- May, R. S. (1986). Inferences, subjective probability and frequency of correct answers: A cognitive approach to the overconfidence phenomenon. In B. Brehmer, H. Jungermann, P. Lourens, & G. Sevo'n (Eds.), *New directions in research on decision making* (pp. 175–189). Amsterdam, the Netherlands: North-Holland.
- McGlothlin, W. H. (1956). Stability of choices among uncertain alternatives. *American Journal of Psychology*, 69, 604–615.
- Murphy, A. H., & Winkler, R. L. (1977). Can weather forecasters formulate reliable probability forecasts of precipitation and temperature? *National Weather Digest*, 2, 2–9.
- Neale, M. A., & Bazerman, M. H. (1990, forthcoming). *Cognition and rationality in negotiation*. New York, NY: The Free Press.
- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and*

shortcomings of human judgment. Englewood Cliffs, NJ: Prentice-Hall.

Oskamp, S. (1962). The relationship of clinical experience and training methods to several criteria of clinical prediction. *Psychological Monographs*, 76(28), 1–28.

Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology*, 29, 261–265.

Peterson, C. R., & Miller, A. J. (1965). Sensitivity of subjective probability revision. *Journal of Experimental Psychology*, 70, 117–121.

Peterson, C. R., Schneider, R. J., & Miller, A. J. (1965). Sample size and the revision of subjective probabilities. *Journal of Experimental Psychology*, 69, 522–527.

Phillips, L. D., & Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, 72, 346–354.

Quattrone, G. A. (1982). Overattribution and unit formation: When behavior engulfs the person. *Journal of Personality and Social Psychology*, 42, 593–607.

Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance*, 6, 649–744.

Snizek, J. A., & Switzer, F. S. (1989). *The over-underconfidence paradox: High Pi's but poor unlucky me*. Paper presented at the Judgment and Decision Making Society annual meeting in Atlanta, Georgia.

Staël von Holstein, C.-A. S. (1972). Probabilistic forecasting: An experiment related to the stock market. *Organizational Behavior and Human Performance*, 8, 139–158.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103, 193–210.

Tversky, A., & Kahneman, D. (1971). The belief in the law of small numbers. *Psychological Bulletin*, 76, 105–110.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.

Vallone, R. P., Griffin, D. W., Lin, S., & Ross, L. (1990). The overconfident prediction of future actions and outcomes by self and others. *Journal of Personality and Social Psychology*, 58, 582–592.

von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research*. New York, NY: Cambridge University Press.

Wagenaar, W. A., & Keren, G. (1986). Does the expert know? The reliability of predictions and confidence ratings of experts. In E. Hollnagel, G. Maneini, & D. Woods (Eds.), *Intelligent decision support in process environments* (pp. 87–107). Berlin: Springer.

Wright, G., & Wisudha, A. (1982). Distribution of probability assessments for almanac and future event questions. *Scandinavian Journal of Psychology*, 23, 219–224.

Yates, J. F. (1990). *Judgment and decision making*. Englewood Cliffs, NJ: Prentice-Hall.

第12章 模糊厌恶和相对无知

克雷格·R.福克斯
阿莫斯·特沃斯基

引言

现代决策理论的一个基本问题是对无知或模糊情况下的决策进行分析，在这种情况下，潜在结果的概率既不是预先指定的，也不能根据现有的证据轻易地进行评估。弗兰克·奈特（Frank H. Knight）解决了这个问题，他对可测量的不确定性和不可测量的不确定性进行了区分，前者可以用精确的概率表示，而后者则不能（Knight, 1921）。此外，他还认为企业家应该因承担不可衡量的不确定性得到补偿，而不是因为风险。同时，经济学家约翰·梅纳德·凯恩斯区分了概率和证据的权重，前者代表支持某一特定命题的证据的平衡，而后者代表支持该平衡的证据的数量（Keynes, 1921）。他接着问道：“如果两种概率在程度上是相等的，那么我们在选择行动方针时，是否应该选择基于更广泛的知识概率呢？”明确的概率和模糊的概率之间的区别已经被主观主义学派的支持者所否定。虽然萨维奇承认主观概率通常是模糊的，但他认为模糊在理性选择理论中没有作用（Savage, 1954）。

19世纪60年代初在该期刊上发表的一系列论文和评论重新引起了人们对无知决策问题的兴趣。埃尔斯伯格在其撰写的最有影响力的一篇文章中提出了一些引人注目的观点，即人们更喜欢为已知的而不是未知的事物投注（Ellsberg, 1961; Fellner, 1961）。埃尔斯伯格提出了最简单的被称为“双色”问题的例子，它用到了两个罐子，每个罐子中都有红球和黑球。1号罐子中有50个红球和50个黑球，而2号罐子中含有100个未知比例的红球和黑球。假设从一个罐子中随机抽出一个球，根据结果，人们有可能获得100美元或什么也得不到。对于任何一个罐子，大多数人似乎对投注于红球或黑球没有特别关注，但他们更愿意投注红黑球比例相同的罐子，而不是投注红黑球比例未知的罐子。这种偏好模式与期望效用理论不一致，因为它表明黑球和红球的主观概率在比例相同的罐子中，比在比例未知的罐子中更大，因此不能对两个罐子求和。

大约40年前凯恩斯就讨论过同样的问题：“在第一种情况下，我们知道罐子内包含相同比例的黑白球；在第二种情况下，每种颜色球的比例是未知的，并且里面的球可能是白色的也可能是黑色的。显然，在任何一种情况下拿出白球的概率是1/2，但是参数的权重在第一种情况中更大（Keynes, 1921）。”埃尔斯伯格在奈特和凯恩斯的影响下提出，人们在不确定的情况下采取行动的意愿不仅取决于事件的感知概率，还取决于其模糊性。埃尔斯伯格将模糊性描述为“一种取决于信息的数量、类型和‘一致性’的特征，并在估计相对概率时产生了一定程度的‘置信度’”（Ellsberg, 1961）。

使用埃尔斯伯格初始问题的几个变体的许多实验已经证实，相对于在模糊性上投注，人们更倾向于在确定性上投注（Camerer & Weber, 1992）。如上所述，这些观察结果为期望效用理论的描述效度提供了证据。此外，尽管雷法认为在通过抛硬币来决定是猜红色还是猜黑色的问题中，模糊性可以降低风险，但许多作者仍试图从规范的角度为风险偏好辩护（Raiffa, 1961）。

模糊厌恶引起了人们的广泛关注，因为除了显著的机会游戏外，决策者通常不知道潜在结果的准确概率。关于创业、上法庭或接受治疗的决定，通常是在对这些行动成功的机会缺乏明确认识的情况下做出的。那么在埃尔斯伯格实验中观察到的模糊厌恶是否适用于这样的决策？换句话说，相对于模糊的概率，人们更喜欢清晰的概率这一点是只局限于机会领域，还是可以扩展到基于世界知识的不确定信念？

为了回答这个问题，希思和特沃斯基进行了一系列实验，比较了人们对于投注模糊性的意愿与对于投注明确的随机事件的意愿（Heath & Tversky, 1991）。与模糊厌恶相反，他们发现人们更倾向于在感到自己特别有能力或知识渊博的情况下对模糊性信念进行投注，尽管他们更愿意在确定时投注。在一项研究中，被试被要求根据3种不确定性来源进行投注选择：1988年各州的总统选举结果、各项职业足球比赛的结果及来自已知成分的罐子的随机抽签结果。因富有政治知识和缺乏足球知识而被预选的被试更倾向于投注政治事件，而不是他们认为可能性相同的随机事件。同时，这些被试又更倾向于投注随机事件，而不是他们认为可能性相同的体育赛事。类似地，因富有足球知识和缺乏政治知识而被预选的被试表现出了相反的模式，他们喜欢足球胜过政治事件和随机事件。另一个与希思和特沃斯基的能力假说一致但与模糊厌恶不一致的发现是，人们倾向于投注他们的身体技能（如投掷飞镖）而不是相对应

的随机事件，尽管事实上对于技能而言成功的概率是模糊的，而对于随机事件而言是清晰的（Cohen & Hansel, 1959; Howell, 1971）。

如果模糊厌恶是由无知的感觉驱使的，那么正如前面讨论的那样，问题的关键在于这种心态是在什么条件下产生的。我们认为，当人们将对于一个事件的有限知识与对于另一个事件的丰富知识进行对比，或者将自己与知识渊博的人进行比较时，人们的自信心就会受到损害。此外，我们认为知识状态之间的这种对比是模糊厌恶的主要来源。当单独评估不确定性事件时，人们试图评估其概率，因为在一个好的贝叶斯公式中，人们相对很少关注二阶特征，如模糊性或证据权重。然而，当人们比较具有不同知识水平的两个事件时，对比度使得较不熟悉的投注不那么有吸引力而使得更熟悉的投注更具吸引力。这种被称为比较性无知假设的主要含义是当被试综合评估清晰和模糊的预期时会出现模糊厌恶，但当他们单独地评估每个预期时，模糊厌恶将大大减少或消失。

对实验文献的回顾，让我们发现了一个值得注意的事实：几乎所有对模糊厌恶的检验都采用了被试内设计，其中被试比较了明确的和模糊的投注，而不是不同被试评估每次投注的被试间设计。因此，该文献没有回答在对清晰和模糊赌注之间进行对比时是否存在模糊厌恶的问题。在下面的一系列研究中，我们检验了一个假设，即模糊厌恶在比较情境（或组内设计）中存在，而在非比较情境（或组间设计）中减少或消失了。

实验

研究1

我们通过问卷的形式向斯坦福大学的141名本科生提出了下述假设性问题。该问卷由几个不相关的项目组成，被试填写问卷即可获得学分。

想象一下，桌子上有一个袋子（袋子 A），里面装满了 50 个红色筹码和 50 个黑色筹码，第二个袋子（袋子 B）里装满了 100 个红色和黑色筹码，但你不知道他们的相对比例。假设你获得了如下游戏权利：首先，你要先猜测颜色（红色或黑色）。接下来，你从其中一个袋子中抽出一个筹码。如果抽出的颜色与预测的颜色相同，那么你将赢得 100 美元；否则你什么都赢不到。为玩这样一个游戏你最多愿意付多少钱？（0 ~ 100 美元）

袋子 A	袋子 B
50 个红色筹码	? 个红色筹码
50 个黑色筹码	? 个黑色筹码
共 100 个筹码	共 100 个筹码
我愿意为袋子 A（50 红色；50 黑色）最多支付：_____	
我愿意为袋子 B（? 红色；? 黑色）最多支付：_____	

大约一半的被试完成了上述比较性任务；两次投注的顺序是平衡的。剩下的被试执行的是非比较性任务：其中大约一半的被试单独评估明确性投注，剩下的被试单独评估模糊性投注。

被试对于每项投注的平均支付意愿如表12-1所示。与所有后续表格一样，标准误差（在括号中）和样本大小（ N ）列在均值下方。数据支持了我们的假设。在比较性条件下存在模糊厌恶的强有力证据：被试愿意为明确性投注比为模糊性投注平均多支付9.51美元 [$t(66) = 6.00$, $p < 0.001$]。然而，在非比较性条件下，没有模糊厌恶的痕迹，因为被试为明确性投注支付的费用略低于模糊性投注 [$t(72) = -0.12$, n.s.]。这种相互作用是显著的 ($z = 2.42$, $p < 0.01$)。

表12-1 研究1的结果数据

	明确性投注		模糊性投注	
比较性	24.34 美元		14.85 美元	
	(2.21)	$N = 67$	(1.80)	$N = 67$
非比较性	17.94 美元		18.42 美元	
	(2.50)	$N = 35$	(2.87)	$N = 39$

研究2

我们在这项研究中检验了比较性无知假设与真实货币。被试是通过在斯坦福大学心理学大楼张贴海报的方式招募而来的，他们被告知有机会参加一项简短的研究，会赢得高达20美元的奖金。我们招募了110名学生、教师和员工。由于反应不一致，其中6名被试被排除在外。

被试单独进行实验。在比较性条件下的被试为明确性投注和模糊性投注进行定价。在非比较性条件下的一半被试为明确性投注单独定价；另一半被试为模糊性投注单独定价。明确性投注包括从一个装有1个红色乒乓球和1个绿色乒乓球的袋子中抽取一个球。模糊性投注包括从一个装有2个乒乓球的袋子中抽取一个球，球可以是红色，也可以是绿色。首先我们要求被试猜测取出的球的颜色。接下来，要求他们做出一系列选择，如果他们的猜测是正确的，那么就可以获得20美元（否则什么也得不到），或者可以肯定地得到X美元。被试在一份回复表上标明他们的选择，该回复表以50美分的幅度从19.50美元到0.50美元的降序列出了各种确定金额（X美元）。我们还告知被试，他们中的某些人将被随机选中以获得真实的奖金。这些被试将随机选择一个选项，接受X美元或进行投注，这取决于他们的偏好。这个程序是激励相容的，因为被试只能通过歪曲他们的偏好来使自己遭受损失。

现金等价物是通过优先于不确定性投注的最低金额与首选投注的最高金额之间的中间值估算的。平均现金等价物如表12-2所示。本研究中引入的程序变化（实际投注、金钱激励、个人管理）并未影响结果的模式。在比较性条件下，被试平均投注1.21美元，比模糊性投注更高（ $t(51) = 2.70, p < 0.01$ ）。然而，在非比较性条件下，被试给模糊性投注的定价略高于明确性投注（ $t(50) = -0.61, p$ 不显著）。同样，相互作用是显著的（ $z = 1.90, p < 0.05$ ）。

表12-2 研究2的结果数据

	明确性投注		模糊性投注	
比较性	9.74 美元		8.53 美元	
	(0.49)	$N = 52$	(0.58)	$N = 52$
非比较性	7.58 美元		8.04 美元	
	(0.62)	$N = 26$	(0.43)	$N = 26$

关于研究1和研究2的解释有两条意见。首先，在比较性条件和非比较性条件下，被试都清楚地意识到自己不知道在比例模糊的袋子中，两种颜色的球各有多少个。然而，只有在比较性任务中，这一事实才会影响其定价。因此，模糊厌恶似乎需要在明确性和模糊性投注之间进行直接比较；仅意识到缺失信息是不够的（Frisch & Baron, 1988）。其次，值得注意的是，在研究1和研究2中，比较性情境增强了明确性投注的吸引力，而不是降低了模糊性投注的吸引力。然而，比较性无知假设并没有预测这些影响的相对大小。

研究3

除了上文中描述的双色球问题，埃尔斯伯格还引入了一个3色球问题（Ellsberg, 1961），如表12-3所示。这里有一个装有10个白球及20个未知比例的红球和蓝球的罐子。在决定1中，被试被要求做出选择，选择 f_1 ；抽得白球为赢（ $p=1/3$ ）；或选择 g_1 ，抽得红球为赢（ $0 \leq p \leq 2/3$ ）。在决定2中，被试被要求做出选择，选择 f_2 ，抽得白球或蓝球为赢（ $1/3 \leq p \leq 1$ ）；或选择 g_2 ，抽得红球或蓝球为赢（ $p=2/3$ ）。正如埃尔斯伯格所认为的那样，在决定1中人们通常喜欢 f_1 而不是 g_1 ，在决定2中，人们更喜欢 g_2 而不是 f_2 ，这与期望效用理论的独立性定理相反。

表12-3 埃尔斯伯格的3色球问题数据

		10 个球	20 个球	
	投注	白球	红球	蓝球
决定 1	f_1	50 美元	0	0
	g_1	0	50 美元	0
决定 2	f_2	50 美元	0	50 美元
	g_2	0	50 美元	50 美元

从比较性无知假设的角度来看，这个问题与双色球问题不同，因为这里对投注（特别是 f_2 ）的描述涉及了明确性和模糊性概率。因此，我们预测即使在非比较性情境中也存在模糊厌恶，其中每名被试仅评估一次投注。然而，我们预测在比较性情境中，模糊厌恶产生了更强的影响，其中每名被试都评估了明确性和模糊性投注。本研究对这些预测进行了检验。

被试是美国威拉姆特大学法学院的162名一年级学生，他们在课堂上完成了一份简短的问卷调查。3名拒绝支配地位的被试被排除在分析之外。被试被告知有人会根据他们的选择随机进行支付。这些说明包括与激励相容的支付方案的简要描述（Becker, DeGroot & Marschak, 1964）。被试被要求阐述表13-3中显示的投注的最低定价。在比较性条件下，被试对所有4个投注进行了定价。在非比较性条件下，大约一半的被试对两个互补的明确性投注（ f_1 和 g_2 ）进行了定价，剩下的被试对两个互补的模糊性投注（ f_2 和 g_1 ）进行定价。投注的顺序是平衡的。

设 $c(f)$ 为投注 f 的规定价格。正如预期的那样，大多数处于比较性条件下的被试给明确性投注的定价都高于模糊性投注。特别是，我们观察到28名被试的 $c(f_1) > c(g_1)$ ，17名被试的 $c(f_1) = c(g_1)$ ，以及8名被试的 $c(f_1) < c(g_1)$ ， $p < 0.01$ 。同样，我们观察到36名被试的 $c(g_2) > c(f_2)$ ，12名被试的 $c(g_2) = c(f_2)$ ，以及5名被试的 $c(g_2) < c(f_2)$ ， $p < 0.001$ 。此外，62%的被试表现出模糊厌恶，即 $c(f_1) \geq c(g_1)$ 和 $c(f_2) \leq c(g_2)$ 所暗示的模式，其中至少有一个不等式是严格的。

为了对比比较性和非比较性条件，我们为每名被试增加了两个互补的明确性投注（即 $c(f_1) + c(g_2)$ ）的定价和两个互补的模糊性投注（即 $c(g_1) + c(f_2)$ ）的定价。显然，对于非比较性条件下的主体，我们只能计算其中一个的总和。这些总和可以用来衡量投注在明确性和模糊性两侧的吸引力。这些金额的平均值如表12-4所示。结果符合预期，在比较性条件下，被试的平均定价明显高于模糊性投注，差异为10.68美元（ $t(52) = 6.23, p < 0.001$ ）。然而，在非比较性条件下，差异仅为3.85美元（ $t(104) = 0.82, p$ 不显著）。这种相互作用是非常显著的（ $z = 1.37, p < 0.10$ ）。

表12-4 研究3的结果数据

	明确性投注		模糊性投注	
比较性	55.60 美元		44.92 美元	
	(2.66)	$N = 53$	(3.27)	$N = 53$
非比较性	51.69 美元		47.85 美元	
	(2.94)	$N = 54$	(3.65)	$N = 52$

对个人投注的检验表明，对于更可能的投注 f_2 和 g_2 ，在比较性条件下，人们表现出了对明确性的强烈偏好（ $c(g_2) = 33.75$ 美元， $c(f_2) = 24.66$ 美元， $t(52) = 5.85$ ， $p < 0.001$ ），并且在非比较性条件下表现出了对于模糊性的中等偏好（ $c(g_2) = 31.67$ 美元， $c(f_2) = 26.71$ 美元， $t(104) = 2.05$ ， $p < 0.05$ ）。但是，对于不太可能的投注 f_1 和 g_1 ，我们发现在比较性条件和非比较性条件下，被试给出的明确性和模糊性投注之间的定价没有显著性差异，分别为（ $c(g_1) = 20.26$ 美元， $c(f_1) = 21.85$ 美元， $t(52) = 1.05$ ， p 不显著）及（ $c(g_1) = 21.13$ 美元， $c(f_1) = 20.02$ 美元， $t(104) = 0.43$ ， p 不显著）。因此，表12-4中显示的聚合模式主要由可能性更高的投注驱动。

研究4

在前3个研究中，不确定性是由机会模式产生的（即从具有比例已知或比例未知的罐子中抽取球）。我们接下来的研究将通过自然事件来检验比较性无知假设。具体来说，我们要求被试根据熟悉的城市（圣弗朗西斯科）和一个气候与其相似的不熟悉的城市（伊斯坦布尔）的未来温度来进行假设性投注。模糊厌恶表明，我们的被试（居住在圣弗朗西斯科附近）应该更倾向于投注他们非常熟悉的圣弗朗西斯科的温度，而不是投注伊斯坦布尔的温度。

如果给定城市的温度高于或低于某个特定值，那么我们会询问被试愿意为提供固定奖金的命题的每一方投入多少钱。具体描述如下。

想象一下，从今天起一周内，如果（圣弗朗西斯科 / 伊斯坦布尔）的下午高温至少为 60 °F（约为 16°C），那么你将获得 100 美元。你愿意为这个投注最多支付多少钱？

我愿意支付最多_____美元。

想象一下，从今天起一周内，如果（圣弗朗西斯科 / 伊斯坦布尔）的下午高温低于 60 °F，那么你将获得 100 美元。你愿意为这个投注最多支付多少钱？

我愿意最多支付_____美元。

在非比较性条件下，一组被试为上述两个与圣弗朗西斯科相关的投

注定价，第二组被试为相同的两个与伊斯坦布尔相关的投注定价。在比较性条件下，被试需要执行两项任务，为所有4个投注进行定价。事件（小于60°F/至少为60°F）和城市的顺序是平衡的。为了最大限度地减少顺序效应，在回答问题之前，我们要求所有被试思考要投注的城市或城市的下午高温的最佳猜测。

被试是加州大学伯克利分校校园里的189名行人，他们完成了一项时长为5分钟的调查（其中包括一些不相关的项目）以换取加州彩票。违背支配优势的10名被试被排除在分析之外。这里没有显著的顺序效应。设 $c(SF \geq 60)$ 表示愿意为预期“如果圣弗朗西斯科从今天起一周内的高温至少为60°F，那么赢100美元”支付的价格。如研究3所示，我们为每名被试添加了其愿意支付的投注的互补条件。特别是，我们计算了圣弗朗西斯科投注的 $c(SF \geq 60) + c(SF < 60)$ 和伊斯坦布尔投注的 $c(Ist \geq 60) + c(Ist < 60)$ 。表12-5中列出了这些总和的平均值。结果再次支持了我们的假设。在比较性条件下，与不熟悉的伊斯坦布尔温度相比，被试愿意在熟悉的圣弗朗西斯科温度投注上平均多付15.84美元（ $t(89) = 5.05, p < 0.001$ ）。然而，在非比较性条件下，被试愿意为圣弗朗西斯科投注平均多支付1.52美元（ $t(87) = 0.19, p < 0.05$ ）。这种相互作用是显著的（ $z = 1.68, p < 0.05$ ）。

表12-5 研究4的结果数据

	投注圣弗朗西斯科		投注伊斯坦布尔	
比较性	40.53 美元		24.69 美元	
	(4.27)	$N = 90$	(3.09)	$N = 90$
非比较性	39.89 美元		38.37 美元	
	(5.06)	$N = 44$	(6.10)	$N = 45$

同样的模式也适用于个人投注。在比较性条件下， $c(SF \geq 60) = 22.74$ 美元， $c(Ist \geq 60) = 15.21$ 美元， $t(89) = 3.13, p < 0.01$ 。类似地， $c(SF < 60) = 17.79$ 美元和 $c(Ist < 60) = 9.49$ 美元， $t(89) = 4.25, p < 0.001$ 。在非比较性条件下， $c(SF \geq 60) = 21.95$ 美元， $c(Ist \geq 60) = 21.07$ 美元（ $t(87) = 0.17, p$ 不显著）。类似地， $c(SF < 60) = 17.94$ 美元， $c(Ist < 60) = 17.29$ 美元（ $t(87) = 0.13, p$ 不显著）。因此，处于比较性条件下的被试愿意为圣弗朗西斯科投注支付的费用远远超过他们愿意为伊斯坦布尔投注支付的费用。然而，在非比较性条件下没有明显的这种形式的差异。注意，与研究1和研究2中观

察到的效果不同，当前效果是通过降低不太熟悉的投注的吸引力产生的。

研究5

我们用比较性无知的观点来解释前面的研究结果。或者，可以认为这些结果至少可以部分地用一个更普遍的假设来解释，即单独评估的预期现金等价物之间的差异，将通过它们之间的直接比较得到增强。无论所讨论的前景是否涉及不同的不确定性来源（这些来源因熟悉程度或模糊性而有所不同），这种增强都将适用。

为了检验这一假设，我们招募了129名斯坦福大学本科生为被试来回答一页长的问卷。如果帕洛阿尔托（斯坦福所在地）某一天白天的高温落在指定范围内，被试就可以获得100美元，研究人员要求被试说出其愿意为这个赌注支付的最高金额是多少。对于这两个赌注的描述如下所示。

[A] 想象一下，如果两周内帕洛阿尔托的下午高温超过 70 °F，那么你将获得 100 美元。你愿意为这个赌注最多支付多少钱？
我愿意最多支付_____美元。

[B] 想象一下，如果从今天起 3 周内帕洛阿尔托的下午高温低于 65 °F，那么你将获得 100 美元。你愿意为这个赌注最多支付多少钱？
我愿意最多支付_____美元。

处于比较性条件下的被试同时评价 [A] 和 [B]（顺序是平衡的）。处于非比较性条件下的被试中的约一半人单独评价 [A]，其余的被试单独评价 [B]。

因为帕洛阿尔托在春天（进行研究时）的温度更可能高于70°F而不是低于65°F，我们预计投注 [A] 通常比投注 [B] 更有吸引力。因此，增强假设意味着 $c(A)$ 和 $c(B)$ 之间的差异在比较性条件下比非比较性条件下更大。 $c(A)$ 和 $c(B)$ 的平均值如表12-6所示。结果不支持增强假设。在该研究中， $c(A)$ 大于 $c(B)$ 。然而，它们之间的差异 $c(A) - c(B)$ 在两个条件下大致相同（相互作用 $z=0.32$ ， p 不显

著)。事实上，任何一个前景的现金等价物的比较性和非比较性条件之间都没有显著性差异（对于A， $t(87)=0.53$ ， p 不显著；对于B， $t(85)=0.48$ ； p 不显著）。这种模式与前面研究的结果（特别是表12-5）形成了鲜明对比，这表明比较性条件下的所述价格与非比较性条件下的所述价格之间存在显著性差异。我们得出结论：在研究1至研究4中观察到的比较性无知效应不能用更一般的增强假设来解释。

表12-6 研究5的结果数据

	投注 A		投注 B	
比较性	25.77 美元		6.42 美元	
	(3.68)	$N = 47$	(1.84)	$N = 47$
非比较性	23.07 美元		5.32 美元	
	(3.42)	$N = 42$	(1.27)	$N = 40$

研究6

比较性无知假设将模糊厌恶归因于知识状态之间的对比。在前4项研究中，我们为被试提供了更熟悉和更不熟悉事件之间的比较。在最后一项研究中，我们为被试提供了他们自己和更专业的人之间的比较。

被试是美国圣何塞州立大学的本科生。我们在一份调查问卷中设计了以下假设性问题，其中包含几个不相关的项目。被试完成这些项目就可以获得课程学分。

考夫曼广阔家园（KBH）是美国最大的房屋销售商之一。他们的股票在纽约证券交易所上市交易。

- [1] 你认为 KBH 的股票周一的价格会比它昨天的收盘价更高还是更低？
- KBH 的股价将更高。

● KBH 的股价不变或更低。
- [2] 你更喜欢下列哪一个？
- 肯定会收到 50 美元。

● 如果对 KBH 的预测是正确的，那么收到 150 美元。

非比较性条件下的被试（ $N=31$ ）回答了上述问题。比较性条件下的被试（ $N=32$ ）回答了相同的问题，但在问题1和问题2之间插入了以下附加项目。

我们将向美国圣何塞州立大学的本科生、斯坦福大学的经济学研究生，以及专业的股票分析师展示这项调查。

然后我们要求被试以0~10的等级评定他们关于该项目的知识水平。

现在的解释表明，与更专业的人（即经济学研究生和专业的股票分析师）进行比较的建议将削弱被试的能力感，因此会降低他们为自己的判断投注的意愿。实验结果支持这一预测。在非比较性条件下，与赢得50美元的确定性前景相比，68%的被试更倾向于可以赢得150美元的不确定性前景；而在比较性条件下，这一比例只有41%（ $\chi^2(1) = 4.66, p < 0.05$ ）。

我们通过不同的被试群体（参加过心理学入门课程的斯坦福大学本科生）和不同的不确定性事件来复制这种效应。我们在一份调查问卷中设置了下述假设性问题，其中包含为获得课堂学分而需要完成的几个不相关的项目。

[1] 你认为荷兰过去 12 个月的通胀率是高于还是低于 3.0%？

- 低于 3.0%
- 至少 3.0%

[2] 你更喜欢下列哪个选项？

- 肯定会得到 50 美元
- 如果预测的通货膨胀率是正确的，那么会得到 150 美元

如前所述，处于非比较性条件下的被试（ $N=39$ ）评价上述项目，而处于比较性条件下的被试（ $N=37$ ）在回答相同问题的同时，在问题[1]和[2]之间插入了下述附加项目。

我们向心理学专业的本科生、经济学专业的研究生和专业的商业预测人员展示了这项调查。

然后我们要求被试以0~10的等级评定他们在该问题中的知识水平。

在非比较性条件下，38%的被试肯定支付不确定性预期，而在比较性条件下仅有11%的被试肯定支付（ $\chi^2(1) = 7.74, p < 0.01$ ）。因此，将其与更专业的人的意见进行比较，减少了被试对不确定性事件投注的倾向。请注意，这项研究的结果仅仅提到一个更专业的人群，应该与柯利、耶茨和艾布拉姆斯的发现区分开来，当人们预估到自己的决定将被同龄人评估时，模糊厌恶会得到加强（Curley, Yates & Abrams, 1986）。

市场实验

在我们将注意力转向目前调查结果的含义之前，问题出现了，当决策者有机会在提供激励和即时反馈的市场环境中做出多项决策时，模糊性和比较性无知的影响是否仍然存在？R.萨里（R. Sarin）和马丁·韦伯（Martin Weber）给出了这个问题的肯定答案，他们使用密封投标和双重口头拍卖比较了几个实验中被试对明确性和模糊性投注的定价（Sarin & Weber, 1993）。在一系列涉及科隆大学工商管理学研究生的研究中，如果从一个装有10个黄色网球和10个白色网球的不透明罐子中抽出一个黄色网球，就需要为这一明确性投注支付100德国马克。除了被试不知道黄色网球和白色网球的比例（这是从均匀分布中抽取的样本）外，这个模糊性投注的定义与明确性投注相似。在一些研究中，被试在每个市场上都进行了明确性和模糊性投注。在其他研究中，被试在一些市场上进行了明确性投注，在其他市场上进行了模糊性投注。因此，所有的被试都评估了明确性和模糊性投注。根据比较性无知假设预测：第一，明确性投注的定价一般高于模糊性投注；第二，当对明确性和模糊性投注进行联合交易时，价格之间的差异将比单独交易时更明显。数据支持这两种预测。在两种拍卖类型（实验11至实验14的最后一个交易期）中，明确性投注和模糊性投注的平均市场价格之差在联合市场中超过20马克，在单独市场中低于5马克。这种影响在两个口头拍卖中特别明显，在这两个拍卖中，单独市场中的明确性投注和模糊性投注的市场价格没有差别，而在联合市场中则有很大的差别（18.5马克）。显然，市场设置不足以消除模糊性和比较性无知的影响。

讨论

前面的研究为比较性无知假设提供了支持，根据这种假设，模糊厌恶主要是由事件或个体之间的比较驱动的，在没有这种比较的情况下，模糊厌恶会大大减少或消失。我们必须补充一点，比较性和非比较性评估之间的区别是指决策者的心态，我们试图通过实验环境来对其进行控制。当然，不能保证在比较性条件下的被试确实进行了建议的比较，或者在非比较性条件下的被试没有独立产生比较。例如，在埃尔斯伯格的双色球问题中，单独使用模糊罐子的人可能会自发地将比较调整到两种颜色的球各占50个的明确的罐子，特别是如果他们之前遇到过同样的问题的话。然而，在前面的研究中观察到的一致结果表明，实验操作成功地使被试在一种情况下进行比较而在另一种情况下不进行比较。

比较性无知假设表明，当人们单独地对一个不确定性前景进行定价时（如果伊斯坦布尔在未来一周内温度超过60°F，将得到100美元），他们很少或根本不关注自己给出的评估概率的质量或精确度。然而，当人们被要求在另一个情境下对这个前景进行定价时（如果圣弗朗西斯科在未来一周内温度超过60°F，那么得到100美元），他们会对这两个事件的知识对比变得敏感，不那么熟悉或不确定性前景的价格会低于更熟悉或更明确的前景（Heath & Tversky, 1991; Keppe & Weber, 1995）。同样，当人们意识到同样的前景也将由更专业的人评估时，不确定性前景就变得不那么有吸引力了。因此，模糊厌恶代表着不愿意对简单知识采取行动的态度，而这种自卑感只有在与其他领域或其他人的高级知识进行比较时才会出现。

理论意义

比较性无知效应违反了程序不变性原则，根据该原则，策略上同等的启发式过程应该产生相同的偏好顺序（Tversky, Sattath & Slovic, 1988）。在之前的研究中，不确定性和明确性投注在单独定价时价值相等，而当对两种投注同时定价时，后者的价值严格高于前者。和其他偏好逆转的例子一样，一个特定的属性（在本例中是关于概率的知识）在比较性评估中比在非比较性评估中更为显著（Tversky & Thaler, 1990）。然而，最值得注意的发现不是对一种新的偏好逆转进行的说明，而是埃尔斯伯格现象是一种内在比较效应这一结论。

比较性评估和非比较性评估之间的这种差异引出了哪种偏好更合理的问题。一方面，我们可以认为比较性判断情景反映了人们的“真实”偏好，而在没有比较的情况下，人们没有低估自己的无知。另一方面，可

能有人认为非比较性判断更为理性，而被试只是在与高级知识的比较中受到了刺激。正如我们所看到的那样，没有令人信服的论据支持另一种解释。理性选择理论（或更具体地说是程序不变性原理）认为比较性和非比较性评估是一致的，但该理论并没有提出一种协调不一致偏好的方法。

目前的研究结果对分析个人决策有什么影响？要回答这个问题，重要的是要区分在不确定状况下决策描述的研究中出现的两种现象：来源偏好和来源敏感度（Tversky & Fox, 1995; Tversky & Wakker, 1995）。来源偏好指的是这样一种观察结果，即前景之间的选择不仅取决于不确定性的程度，还取决于不确定性的来源（如圣弗朗西斯科的温度与伊斯坦布尔的温度）。来源偏好是通过表明一个人倾向于对从一个来源提取的命题而不是从另一个来源提取的命题进行投注来得以证明的，且还倾向于对第一个命题投注而不是对第二个命题投注（ $c(SF \geq 60) > c(Ist \geq 60)$ ， $c(SF < 60) > c(Ist < 60)$ ；参见研究4）。我们将模糊厌恶解释为来源偏好的一个特例，其中风险优于不确定性，如埃尔斯伯格给出的例子所示。[\(32\)](#)

来源敏感度是指决策权重的不可加性。特别是，对风险决策的描述性分析表明，当一个不可能性事件变成可能性事件或可能性事件变为确定性事件时，特定事件对前景价值的影响大于仅仅使一个不确定性事件的概率更大或更小（Kahneman & Tversky, 1979）。例如，将固定奖金从0增加到0.1或从0.9增加到1.0的概率比从0.3增加到0.4的概率有更大的影响力，特沃斯基和福克斯进一步证明了在这种模式中，不确定性比偶然性（即对于模糊性而非明确性概率）更明显，这被称为有界次可加性（Tversky & Fox, 1995）。换句话说，人们对随机的不确定性不那么敏感，尽管他们更喜欢不确定性而不是随机。因此，来源偏好和来源敏感度在逻辑上是独立的。

目前的实验表明，与来源敏感度不同，来源偏好是一种内在的比较现象，它不会出现在对不确定性前景的独立评估中。这表明基于决策权重或非加性概率的模型能够适应来源敏感度，但不能提供一个满意的来源偏好解释，因为它们不能区分比较性和非比较性评估（Quiggin, 1982; Gilboa, 1987; Schmeidler, 1989; Tversky & Wakker, 1995）。人们可以尝试使用偶然性加权方法来建立比较性无知效应的模型，其中与事件相关的权重取决于它是在比较性还是在非比较性情境中被评估的（Tversky, Sattath & Slovic, 1988）。这方面的主要困难，或任何其他

模拟比较性无知效应的尝试的主要困难在于，它需要事先说明决策者对所涉事件的感知能力，以及其他知识状态的显著性。虽然这些变量可以通过实验控制，就像我们在前面的研究中所做的那样，但它们不容易得到测量并纳入正式的模型中。

尽管在模拟相对无知方面存在困难，但相对无知可能具有重大的经济影响力。例如，一个对计算机行业很了解但对能源行业一无所知的人，在选择投资高科技初创企业或石油勘探时，可能会表现出模糊厌恶，但当每一项投资都是独立评估时，情况就不同了。此外，本解释指出，应考虑到这两项投资的顺序可能会影响它们的估值。特别是，如果在进行比较熟悉的投资之前进行考虑，不太熟悉的投资的价值就可能会更高。(33)根据目前的分析，最近针对金融市场中的模糊厌恶建模的尝试可能是不完整的，因为他们没有区分比较性评估和非比较性评估（Dow & Werlang, 1991; Epstein & Wang, 1994）。特别是，这种模型很可能高估了不确定性前景在单独评估的情境中的模糊厌恶程度（Sarin & Weber, 1993）。相对无知在经济交易中的作用有待进一步的实证研究。

致谢

感谢马丁·韦伯的帮助。

参考文献

Becker, G., DeGroot, M., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9, 226–232.

Camerer, C., & Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, 5, 325–370.

Cohen, J., & Hansel, M. (1959). Preferences for different combinations of chance and skill in gambling. *Nature*, 183, 841–843.

Curley, S. P., Yates, J. F., & Abrams, R. A. (1986). Psychological sources of ambiguity avoidance. *Organizational Behavior and Human Decision Processes*, 38, 230–256.

Dow, J., & Ribeiro da Costa Werlang, S. (1991). Uncertainty aversion, risk aversion, and the optimal choice of portfolio. *Econometrica*, 60, 197–204.

Einhorn, H. J., & Hogarth, R. M. (1985). Ambiguity and uncertainty in probabilistic inference. *Psychological Review*, 93, 433–461.

Ellsberg, D. (1961). Risk, ambiguity and the savage axioms. *Quarterly Journal of Economics*, 75, 643–669.

Epstein, L. G., & Wang, T. (1994). Intertemporal asset pricing under Knightian uncertainty. *Econometrica*, 62, 283–322.

Fellner, W. (1961). Distortion of subjective probabilities as a reaction to uncertainty. *Quarterly Journal of Economics*, 75, 670–689.

Frisch, D., & Baron, J. (1988). Ambiguity and rationality. *Journal of Behavioral Decision Making*, 1, 149–157.

Gilboa, I. (1987). Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics*, 16, 65–88.

Heath, C., & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4, 5–28.

Howell, W. (1971). Uncertainty from internal and external sources: A clear case of overconfidence. *Journal of Experimental Psychology*, 81, 240–243.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.

Keppe, H.-J., & Weber, M. (1995). Judged knowledge and ambiguity aversion. *Theory and Decision*, 39, 51–77.

Keynes, J. M. (1921). *A treatise on probability*. London, England: Macmillan.

Knight, F. H. (1921). *Risk, uncertainty, and profit*. Boston, MA: Houghton Mifflin.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3, 323–343.

Raiffa, H. (1961). Risk ambiguity and the savage axioms: Comment. *Quarterly Journal of Economics*, 75, 690–694.

Sarin, R. K., & Weber, M. (1993). Effects of ambiguity in market experiments. *Management Science*, 39, 602–615.

Savage, L. J. (1954). *The foundation of statistics*. New York, NY: John Wiley & Sons.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, 57, 571–587.

Tversky, A., & Fox, C. R. (1995). Weighing risk and uncertainty. *Psychological Review*, 102, 269–283.

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, 95, 371–384.

Tversky, A., & Thaler, R. (1990). Preference reversals. *Journal of Economic Perspectives*, 4, 201–211.

Tversky, A., & Wakker, P. (1995). Risk attitudes and decision weights. *Econometrica*, 63, 1255–1280.

第13章 支持理论：主观概率的非外延性表征

阿莫斯·特沃斯基
德里克·J.凯勒

外行和专家经常会被要求对如审判结果、手术成果、商业投资是否成功或足球比赛的赢家这样的不确定性事件的概率进行评估。这些评估在决定案件的判决、是否接受手术、是否投资企业、是否投注本地球队等问题中都发挥着重要的作用。不确定性通常是以口头的形式得以表达的（如不可能的或可能的），而对其进行数值估计也很常见。例如，气象预报员经常会报告下雨的概率（Murphy, 1985），而经济学家有时需要估计经济衰退的可能性（Zarnowitz, 1985）。主观概率的理论和实践意义激发了心理学家、哲学家和统计学家从描述性和规范性的角度来对其进行研究的兴趣。

事实上，信念的程度是否可以或者应该用概率的计算来表征，这个问题一直以来都存在着激烈的争论。与贝叶斯学派用递加概率测量来表征信念的程度不同，有许多怀疑论者对主观不确定性进行量化的可能性和智慧提出了质疑，不愿将概率定律应用于信念的分析。除了贝叶斯和怀疑论者，关于主观概率修正模型的文献也越来越多。这些包括D-S证据理论（Dempster-Shafer theory of belief; Dempster, 1967; Shafer, 1976）、可能性理论（Zadeh, 1978），以及各种类型的上、下概率（Suppes, 1974; Walley, 1991）。D.杜布瓦（D. Dubois）和H.普拉德（H. Prade）、I.基利波（I. Gilboa）和D.施迈德勒（D. Schmeidler）和P.蒙吉（P. Mongin）对最近的发展进行了回顾（Dubois & Prade, 1988; Gilboa & Schmeidler, 1994; Mongin, 1994）。就像贝叶斯主义者一样，修正主义者也支持对信念进行量化，他们在投注时会使用直接判断或偏好，但他们发现，相对于这一目的来说，对随机事件进行计算的局限性太大了。因此，它们用满足较弱要求的非递加性集合函数取代了经典理论中使用的递加测量。

贝叶斯和修正主义模型都有一个基本假设——外延性原则，即具有相同外延的事件具有相同的概率。然而，外延性假设在描述上是无效的，因为对同一事件的不同描述常常会产生具有系统性差异的判断。下

面的3个例子说明了这一现象，并推动了一种不受外延性假设影响的描述性信念理论的发展。

第一，巴鲁赫·菲施霍夫、P.斯洛维克和S.利希滕斯坦要求汽车机械师和外行评估汽车无法启动的不同原因的概率（Fischhoff, Slovic & Lichtenstein, 1978）。他们发现——“故障的原因不是电池、燃料系统或引擎”——当假设被分解成更具体的原因（如启动系统、点火系统）时，残差假设的平均概率从0.22增至0.44。虽然平均有15年经验的汽车机械师肯定知道这些可能性，但他们并不相信那些没有明确提到的假设。

第二，特沃斯基和卡尼曼构造了许多问题，其中概率和频率判断都与集合包含不一致（Tversky & Kahneman, 1983）。例如，一组被试被要求估计4页长的小说中以ing结尾的7字母单词的数量。第二组被试被要求估计以_n_结尾的7字母单词的数量。第一个问题的估计中位数（13.4）比第二个问题的估计中位数（4.7）高出近3倍，可能是因为以ing结尾的7字母单词比以_n_结尾的7字母单词更容易被想到。但是，似乎大多数评估第二类单词数量的人都没有意识到第一类单词已经被包括在其中了。

第三，违反外延性原则的例子并不局限于概率判断；在评价不确定性前景时，我们也可以观察到这些现象。例如，E. J.约翰逊（E. J. Johnson）、J.赫希（J. Hershey）、J.梅扎罗斯（J. Meszaros）和H.昆鲁斯（H. Kunreuther）发现，与覆盖所有住院类型的保险相比，那些（假设的）覆盖所有因疾病或事故住院的保险会让被试愿意支付更高的保费（Johnson, Hershey, Meszaros & Kunreuther, 1993）。显然，明确提及疾病和事故增加了人们认为的住院治疗的可能性，因此也增加了保险的吸引力。

这些结果与本文后面描述的许多其他观察结果一样，都与外延性原则不一致。我们区分了非外延性的两个来源。首先，外延性可能会因为记忆的限制而失效。如第2个例子所示，我们不能指望评判者能回忆起某一类别案件的所有实例，即使他能准确无误地识别出这些案例。一个明确的描述可以促使人们回想起相关的例子，否则他们可能会忘记。其次，外延性可能会失效，是因为对同一事件的不同描述可能会引起人们对结果的不同方面的关注，从而影响它们的相对显著性。这些效应可以影响概率判断，即使它们不会让人想起新的例子或证据。

我们认为，外延性常常失效代表了人类判断的一个基本特征，而不仅仅是一系列孤立的例子。它们表明，概率判断并不取决于事件本身，而是取决于对事件的描述。在这里，我们提出了一个理论，根据这种理论，对一个事件的概率的判断取决于对它进行的描述的确切性。这种被称为支持理论（support theory）的处理方法侧重于概率的直接判断，但也适用于不确定状况下的决策。我们将在下一节中介绍和描述这一基本理论。实验证据也将在下一节中得到介绍。在最后一部分中，我们将该理论扩展到了顺序判断中，讨论了信念系统的上下边界的指标，并阐述了当前发展的描述性和规范性影响。

支持理论

设 T 为一个包含至少两个元素的有限集合，我们可以将其理解为世界的状态。我们假设只有一个状态是存在的，但评判者一般不知道。 T 的子集被称为事件。我们对事件和对事件的描述进行的区分，即假设。设 H 是一组描述 T 中事件的假设。因此，我们假定每一个假设 $A \in H$ 对应

$$A' \subset T$$

于一个独特的事件。这是一个多对一的映射，因为不同的假设，如 A 和 B ，可能有相同的外延

（即 $A'=B'$ ）。例如，假设一个人掷了一对骰子，假设两个骰子的结果“和为3”和“积为2”是对同一事件的不同描述；也就是说，一个骰子显示1，另一个显示2。我们假设 H 是有限的，并且对于每个事件它至少包含一个假设。如果 $A' \in T$ ， A 是基本事件。如果 $A'=\emptyset$ ， A 就是空集。如果 $A' \cap B'=\emptyset$ ，那么 A 和 B 是互斥的。如果 A 和 B 在 H 中，并且它们是互斥的，那么它们的明确的析取也在 H 中，写作 $A \vee B$ 。因此， H 在互斥析取中是闭合的。我们假设 $\vee$ 是组合的和可交换的，那么 $(A \vee B)'=A' \cup B'$ 。

本公式的一个关键特征是显性和隐性析取的区别。如果 A 既不是基本的，也不是空集，并且不是显性析取，那么它就是隐性析取，或者简单地说是隐性假设（在 H 中存在不互斥的非空集 B 和 C ，使得 $A=B \vee C$ ）。例如，假设 A 是“安主修自然科学”， B 是“安主修生物学”， C 是“安主修物理学”。显性析取 $B \vee C$ （安的专业是生物学或物理学）与 A 有着同样的外延，即 $A'=(B \vee C)'=B' \cup C'$ ，但 A 是一个隐性假设，因为它不是显性析取。请注意，显性析取 $B \vee C$ 是由任一互斥集 $B, C \in H$ 来定义的，而一个可以共存的隐形析取是不可能存在的，因为有些事件不

能在未列出其成分的情况下对其进行自然描述。

评价框架 (A, B) 由一对排他性假设组成：第一个元素 A 是评价的焦点假设，第二个元素 B 是备选假设。为了简化问题，我们假设 A 和 B 是互斥的，评判者认为它们是具有排他性的，但我们不假设评判者可以列出隐性析取的所有组成部分。根据上面的例子，我们假设评判者知道遗传学属于生物学、天文学属于物理学，而生物学和物理学是互斥的。然而，我们并不认为评判者可以列出生物学或物理学的所有分支学科。因此，我们假设人们可以对包含关系进行识别，但无法进行完美的回忆。

我们把一个人的概率判断描述成一个从评价框架到单位区间的映射 P 。为了简化问题，我们假设当且仅当 A 为空集时， $P(A, B) = 0$ ，当且仅当 B 为空集时， $P(A, B) = 1$ ；我们假设 A 和 B 不可能同时为空。因此， $P(A, B)$ 是 A 而不是 B 的判断概率，假设其中一个且只有一个有效。显然， A 和 B 可以分别表示显性或隐性析取。标准理论中与 $P(A, B)$ 的外延对应的是条件概率 $P(A' | A' \cup B')$ 。目前的处理是非外延性的，因为它假设概率判断依赖于对 A 和 B 的描述，而不仅仅依赖于事件 A' 和 B' 。我们要强调，目前的理论适用于评判者提出的假设，但这些假设并不总是符合给出的口头描述。然而，呈现隐性析取的评判者可以将其视为显性析取，反之亦然。

支持理论假定有一个比例尺度 s （可以理解为支持度），它为 H 中的每个假设分配了一个非负实数，那么对于任何一对互斥的假设 $A, B \in H$

$$P(A, B) = \frac{s(A)}{s(A) + s(B)} \quad (1)$$

如果 B 和 C 是互斥的， A 是隐性的，且 $A' = (B \vee C)'$ ，那么

$$s(A) \leq s(B \vee C) = s(B) + s(C) \quad (2)$$

公式1为我们提供了一个主观概率的表征，它表示焦点假设和备选假设的支持度。公式2表明，隐性析取 A 的支持度小于或等于共存的显性析取 $B \vee C$ 的支持度，其等于它的各个元素的支持度之和。因此，支持度对显性析取来说是可加的，对隐性析取是次可加的。

我们认为，次可加性假设代表了人类判断的一个基本原则。当人们评估自己对隐性析取的信念的程度时，他们通常不会按照外延性的要

求，将假设分解成互斥的各个组成部分，并为它们添加支持度。相反，他们倾向于形成一个主要基于最具代表性或可用案例的整体印象。因为这种判断方式是有选择性的，而不是详尽无遗的，因此拆分会增加支持度。换句话说，我们认为隐性假设的概括性表征的支持度通常小于其互斥成分的支持度之和。记忆和注意力都可能产生这种效果。将一个类别（如非自然原因导致的死亡）分解成其组成部分（如致命的车祸、溺水）可能会让人们想到一些原本不会被考虑到的可能性。此外，明确提及一项成果往往会加强其重要性，从而加强对其的支持。虽然这种假设在某些情况下可能会失败，但是我们在下一节中描述的关于次可加性的强烈证据表明，这些失败代表的是例外，而不是规则。

与给定假设相关的支持度被解释为针对支持该假设的证据的强度的一种尺度，该假设对评判者来说是可用的。支持度可能是基于客观数据（如相关人群中的失业频率）或由判断启发式（如代表性、可得性或锚定和调整）介导的主观印象（Kahneman, Slovic & Tversky, 1982）。例如，“比尔是一名会计师”的假设可以通过比尔的个性与会计师刻板印象的匹配程度进行验证，“在明年年底之前东海岸会发生石油泄漏”的预测可能是通过评估类似事故得到的。支持度也可以反映评判者为支持所述的假设而采用的理由或论据（如果被告有罪的话，他就不会报案了）。由于基于印象和推理的判断通常是非外延性的，因此对于集合包含的支持函数是非单调的。因此， $s(B)$ 甚至可以超过 $s(A)$ ，尽管

$$A' \supset B'$$

。但是要注意的是， $s(B)$ 是不能超过 $s(B \vee C)$ 的。例如，如果一个类别的支持度由其实例的可得性决定，那么随机选择的以ing结尾的单词假设的支持度可以超过以_n_结尾的单词假设的支持度。一旦类别之间的包含关系变得透明，以_n_结尾的单词假设就被“以ing或任何其他_n_形式”结尾的单词假设所取代，它的支持度超过了以ing结尾的单词假设。

目前的理论从相对支持的角度对主观概率进行了解释。这一解释表明，在某些情况下，概率判断可以通过对支持度的独立评估来进行预测。稍后我们将探讨这种可能性。由下文中的讨论可知，根据目前的理论，支持度可以从概率判断中得到，正如效用可以从选项之间的偏好中得到。

结果

支持理论是由支持函数 s 表示的，它是不可直接观测的。接下来，我们用观察到的指数 P 来描述这个理论。我们首先列出了这一理论的4个结果，然后证明它们包含公式1和公式2。该理论的直接结果是二元互补

$$P(A, B) + P(B, A) = 1 \quad (3)$$

第二个结果是相称性

$$\frac{P(A, B)}{P(B, A)} = \frac{P(A, B \vee C)}{P(B, A \vee C)} \quad (4)$$

假设 A 、 B 和 C 互斥且 B 不为空。那么， A 相对于 B 的概率是独立于附加假设 C 的。

为了表述下一个条件，我们引入概率比 $R(A, B) = P(A, B) / P(B, A)$ ，这是 A 相对于 B 的概率。由等式1得出乘积法则

$$R(A, B) R(C, D) = R(A, D) R(C, B) \quad (5)$$

假设 A 、 B 、 C 和 D 不为空，且公式5中的4对假设互斥。那么， A 相对于 B 的概率和 C 相对于 D 的概率等于 A 相对于 D 的概率和 C 相对于 B 的概率。为了弄清生成规则的必要性，我们需要注意，根据公式1，公式5等号两边等于 $s(A) s(C) / s(B) s(D)$ 。偏好树理论也使用了在本质上相同的条件（Tversky & Sattath, 1979）。

公式1和公式2共同表明了分解原理。假设 B 、 C 、 D 是互斥的， A 是隐性的， $A' = (B \vee C)'$ ，那么

$$P(A, D) \leq P(B \vee C, D) = P(B, C \vee D) + P(C, B \vee D) \quad (6)$$

s 的特性包含了 P 的相应特性：判断概率对显性析取是可加性的，对隐性析取是次可加性的。换句话说，分解一个隐性析取可能会增加而不是减少它的判断概率。与概率标准理论中的公式3至公式5不同，分解原理（公式6）概括了经典模型。请注意，这一假设与较低概率模型，包括沙菲尔的模型存在差异，沙菲尔的模型假定了外延性和超可加性（即

如果 $A' \cap B' = \emptyset$ ，那么 $P(A' \cup B') \geq P(A') + P(B')$ ；Shafer, 1976)。

产生非加性概率的两个直觉是相互矛盾的。支持理论捕捉到的第一个直觉是：对隐性析取的分解增强了其组成部分的显著性，从而增加了支持度。沙菲尔理论捕捉到的第二个直觉是：在部分事实模糊的情况下，评判者持有某种程度的“储备”信念的尺度，而不是像贝叶斯模型所要求的那样将其分布在所有基本假设中（Shafer, 1976）。虽然沙菲尔的理论是建立在对信念的逻辑分析而不是心理分析的基础上的，但它也被几个作者视为一个描述性模型进行了解释。因此，它为我们提供了与现有理论进行比较的一种自然选择。

虽然之前的比例性（公式4）和乘积法则（公式5）没有得到系统性的检验，但一些研究人员观察到了二元互补（公式3）和分解原理（公式6）的一些方面。我们将在下一节中对这些数据和一些新的研究进行综述。下面的定理表明，上述条件不仅是支持理论的必要条件，而且是充分条件。相关证明见附录。

定理 1：假设 $P(A, B)$ 被定义为所有互斥的 $A, B \in H$ ，而且当且仅当 A 为空时才会消失。当且仅当 H 上存在一个满足公式 1 和公式 2 的非负比例 s 时，公式 3 至公式 6 成立。

该定理表明，如果概率判断满足所要求的条件，就有可能在不假设具有相同外延的假设也具有相同支持度的情况下，衡量与每个假设相关的证据的支持度或强度。我们在最后一节中给出了该理论的延伸，其中 P 被视为序数而不是基数尺度。在本节接下来的部分，我们将介绍次可加性的表征和处理方法。

次可加性

我们通过对次可加性更详细的表征来扩展该理论。假设 A 是一个隐性假设，其外延与基本假设 $A_1 \dots A_n$ 的显性析取相同；也就是说， $A' = (A_1 \vee \dots \vee A_n)'$ 。假定任意两个具有相同外延的基本假设 B 和 C 也具有相同的支持度；也就是， $B', C' \in T$ ，而 $B' = C'$ 意味着 $s(B) = s(C)$ 。因此，在这个假设下，我们可以这样写

$$s(A) = w_{1A}s(A_1) + \dots + w_{nA}s(A_n), 0 \leq w_{iA} \leq 1, i=1, \dots, n \quad (7)$$

在这种表征方式中，每个基本假设的支持度都由其各自的权重“折现”，这反映了评判者对该假设的重视程度。如果所有 i 的 $w_{iA}=1$ ，那么 $s(A)$ 是其基本假设的支持度之和，正如显性析取中那样。此外，对于某些 j ， $w_{jA}=0$ 表示 A_j 实际上被忽略了。最后，如果权重增加，那么 $s(A)$ 是一种加权平均 (A_i) ， $1 \leq i \leq n$ 。我们必须说明：公式7不应被理解为一个深思熟虑的折现过程——评判者通过对相应显性析取的支持度打折的方法来评估隐性析取支持度的过程。相反，权重表示的是评估过程的结果，在此过程中，评判者对 A 进行评估，但没有明确地将其分解为它的组成部分。我们还应该注意，基本假设是由相对于一个已知的样本空间定义的。通过细化描述的层次可以进一步分解这些假设。

请注意，虽然支持函数是唯一的，但是除了尺度单位外，“焦点”权重 w_{iA} 并不仅仅是由观察到的概率判断决定的。然而，这些数据决定了 w_A 所定义的“整体”权重

$$s(A)=w_A[s(A_1)+\cdots+s(A_n)], \quad 0 \leq w_A \leq 1 \quad (8)$$

整体权重 w_A ，也就是对应的隐性 (A) 和显性 $(A_1 \vee \dots \vee A_n)$ 析取的比例，为我们提供了一个测量由 A 产生的次可加性程度的便捷方法。我们认为，次可加性程度受到几个因素的影响，其中一个因素是对概率尺度的解释。具体来说，当概率被理解为个例的倾向，而不是等于或被评估为相对频率时，我们预计次可加性会更加显著。卡尼曼和特沃斯基分别将这些判断模式称为单数模式（singular）和分布模式

（distributional），并认为后者通常比前者更准确（Kahneman & Tversky 1979, 1982; Reeves & Lockhart, 1993）。(34)虽然许多有趣的事件不能用频率的术语来解释，但有些问题可以用分布模式或单数模式来进行描述。例如，人们可能会被要求评估从总体中随机选出的一个人死于意外事故的概率。或者，人们可能被要求评估因事故死亡的人口百分比（或相对频率）。我们认为，当评判者考虑到整个群体而不是一个人的时候，隐性析取“事故”更容易被分解成其组成部分（如车祸、飞机失事、火灾、溺水和中毒）。死亡的各种原因都在人口死亡率的统计中得到了反映，但在单个人的死亡中没有。更为普遍的情况是，我们认为在分布模式下分解隐性析取的倾向比在单数模式下更强。因此，我们估计一个频率公式会比一个表征单个案例的公式产生更少的折扣（即更高的 w_s ）。

条件

回想一下，在 $P(A, B)$ 被理解为给定 A 或 B 的情况下， A 的条件概率。为了获得一般性调节处理，我们加强了假设，假设在 H 集合中，如果 A 和 B 是 H 中不同的元素，那么他们的合取可表示为 AB ，也在 H 集中。自然地，我们假设这种合取是相关的和可交换的，即 $(AB)' = A' \cap B'$ 。我们还假设了分布性，即 $A(B \vee C) = AB \vee AC$ 。令 $P(A, B|D)$ 是在给定数据 D 中判断 A 而不是 B 的概率。一般来说，新的证据（如不同的信息状态）会产生一个新的支持函数 s_D ，它描述的是从 D 来看 s 的修正。在数据可以被描述为 H 的一个元素的特殊情况下，它仅仅限制了被考虑的假设，我们可以将条件概率表示为：

$$P(A, B|D) = \frac{s(AD)}{s(AD) + s(BD)} \quad (9)$$

假设 A 和 B 是互斥的，但 $A \vee B$ 和 D 不是。

以下是对这一形式的一些评论。首先，大家要注意，如果 s 是可加性的，那么公式9就可以简化为条件概率的标准定义。如果 s 是次可加性的，就像我们一直假设的那样，判断概率不仅取决于对焦点和备选假设的描述，还取决于对证据 D 的描述。假设 $D' = (D_1 \vee D_2)'$ ， D_1 和 D_2 是互斥的， D 是隐性的。那么

$$P(A, B|D_1 \vee D_2) = \frac{s(AD_1 \vee AD_2)}{s(AD_1 \vee AD_2) + s(BD_1 \vee BD_2)}$$

但根据次可加性，由于 $s(AD) \leq s(AD_1 \vee AD_2)$ 且 $s(BD) \leq s(BD_1 \vee BD_2)$ ， D 的分解可能更倾向于一种假说而不是另一种假说。例如，考虑到一位女性是一名大学教授，当“大学”被拆分成“法学院、商学院、医学院或其他任何一所学院”时，她的收入非常高的判断概率可能会增加，因为其中明确提及了高薪的职位。因此，公式9扩展了次可加性在证据表征中的应用。正如我们稍后说明的那样，它还允许我们对不同证据的影响进行比较，前提是它们可以被描述为 H 集。

想一下，一个 $n \geq 3$ 同时互斥且详尽（非空）的假设集合， $A_1 \dots A_n$ ，设 $\bar{A}_i$ 表示负性 A_i ，即对应于一个隐性析取的假设。考虑两项证据， B, C

$\in H$ ，假设每个 A_i 与 B 相对于 C 更加兼容， $s(BA_i) \geq s(CA_i)$ ， $1 \leq i \leq n$ 。我们认为 B 比 C 诱发了更多的次可加性， $s(B\bar{A}_i)$ 比 $(C\bar{A}_i)$ 被低估的程度

$$w_{B\bar{A}_i} \leq w_{C\bar{A}_i}$$

更大（如；参见公式7）。这个假设被称为强化，表明 $P(A_i, \bar{A}_i | B)$ 的评估通常会高于 $P(A_i, \bar{A}_i | C)$ 。更具体地说，我们认为每个由不同的评判者评估的(35) $A_i \dots A_n$ 的概率之和，在 B 下并不比在 C 下小。也就是说

$$\sum_{i=1}^n P(A_i, \bar{A}_i | B) \geq \sum_{i=1}^n P(A_i, \bar{A}_i | C) \quad (10)$$

次可加性意味着两个和都大于或等于1。前面的不等式表明，与所研究的假设更相符的证据增加了概率总和。值得注意的是，强化表明人们对假设的先验概率没有进行恰当的反应。下文中的示例说明了强化的含义，并将其与其他模型进行了比较。

假设几个嫌疑人中有一个（而且只有一个）犯了谋杀罪。在没有任何具体证据的情况下，假定所有嫌疑人都被认为具有相同的犯罪的可能性。进一步假设，初步调查发现了一系列的证据（如动机和机会），这些证据在一定程度上与每一个嫌疑人都相关。根据贝叶斯模型，所有嫌疑人的概率会保持不变，因为新的证据是非诊断性的。但在沙菲尔的信念函数理论中，凶杀案是由一名嫌疑人而不是另一名嫌疑人所为的判断概率一般会随着证据的增加而增加；因此，调查后的判断概率应该高于调查前。强化产生了一个不同的模式：二进制概率，如其中一个嫌疑人相对于另一个嫌疑人的犯罪概率；根据贝叶斯模型，在调查前和调查后，二进制的概率应该接近1/2。然而，凶杀案由某一特定嫌疑人（而不是任何其他嫌疑人）所为的判断概率应该随着证据数量的增加而增加。在下文中，我们将描述强化的实验检验。

数据

在这一节中，我们将讨论支持理论的实验证据。结果表明，用正态次可加性支持函数来对判断概率进行理解，可以统一解释文献中报道的几种现象；它还产生了新的预测，目前还没有经过检验。本节由4部分

组成。首先，我们研究了分解的影响，并考察了影响次可加性程度的因素。其次，我们将判断概率与证据强度的直接评级联系了起来。再者，我们研究了强化效应，并比较了不同的信念模型。最后，我们还讨论了不确定性条件下的合取效应、假设生成和决策。

分解研究

还记得分解原理（公式6）吗？它由两部分组成：显性析取的可加性和隐性析取的次可加性，它们共同导致了非外延性。二元互补（公式3）是可加性的一个特例。由于每个部分都有不同的解释，因此同时检验可加性和次可加性非常重要。根据这个原因，我们首先介绍了几个新的研究，它们在同一个实验中检验了分解原理的两个部分；然后我们回顾了以前的研究，这些研究为目前的理论提供了新的动力。

研究1：死亡原因。我们的第一个研究遵循了菲施霍夫等人关于故障树（fault tree）的开创性研究，使用的任务与J.E.拉索和K.J.科尔佐的研究相似（Fischhoff et al., 1978; Russo & Kolzow, 1992）。我们让斯坦福大学的本科生（ $N=120$ ）评估了各种死亡原因的可能性。我们告知被试，美国每年约有200万人（占总人口的近1%）因不同的原因死亡，并要求他们估计各种原因导致的死亡概率。一半的被试被要求只考虑一个最近去世的人，并评估他在一份特定死亡原因列表中死于其中的各种原因的概率。他们需要假设这个问题中的人是从前一年去世的人中随机挑选的。另一半被试则接受频率判断任务，他们评估了前一年死亡的200万人中因各种原因死亡的百分比。在每一组中，一半的被试会收到承诺：准确率最高的5名被试每人将获得20美元的奖励。

每名被试评估两种不同的原因列表中的一种，他要么评估一种隐性假设（如自然原因导致的死亡），要么评估一种共延的显性析取（如由心脏病、癌症或其他的一些自然原因导致的死亡），但不用两者都评估。表13-1列出了可以考虑的全部原因。死亡原因分为自然和非自然两种。每种类型都有3个组成部分，其中一个部分被进一步划分为7个子部分。为了避免出现非常小的概率，我们将这7个子部分限定在相应的死亡类型上（即自然的或非自然的）。为了给被试提供一些锚点，我们告知他们死于呼吸系统疾病的概率或频率约为7.5%，而死于自杀的概率或频率约为1.5%。

表13-1 研究1中死亡原因的平均概率和频率估计及对显性析取和共延性隐性析取的比较评估

假设	概率	平均估计值（%）	
		频率	真实值
3- 因素			
<i>P</i> （心脏病）	22	18	34.1
<i>P</i> （癌症）	18	20	23.1
<i>P</i> （其他自然原因）	33	29	35.2

续表

假设	概率	平均估计值 (%)	
		频率	真实值
Σ (自然原因)	73	67	92.4
P (自然原因)	58	56	
Σ/P	1.26	1.20	
P (事故)	32	30	4.4
P (谋杀)	10	11	1.1
P (其他非自然原因)	11	12	2.1
Σ (非自然原因)	53	53	7.6
P (非自然原因)	32	39	
Σ/P	1.66	1.36	
7- 因素			
P (呼吸道癌 自然原因)	12	11	7.1
P (食道癌 自然原因)	8	7	5.9
P (泌尿系统肿瘤 自然原因)	5	3	2.7
P (乳腺癌 自然原因)	13	9	2.2
P (膀胱癌 自然原因)	7	3	1.0
P (白血病 自然原因)	8	6	1.0
P (其他癌症 自然原因)	17	10	5.1
Σ (癌症 自然原因)	70	49	25.0
P (癌症 自然原因)	32	24	
Σ/P	2.19	2.04	
P (车祸 非自然原因)	33	33	30.3
P (枪械事故 非自然原因)	7	12	1.3
P (意外坠落 非自然原因)	6	4	7.9
P (火灾 非自然原因)	4	5	2.6
P (溺水 非自然原因)	5	4	2.6
P (意外中毒 非自然原因)	4	3	3.9
P (其他事故 非自然原因)	24	17	9.2

续表

假设	概率	平均估计值 (%)	
		频率	真实值
Σ (事故 非自然原因)	83	78	57.9
P (事故 非自然原因)	45	48	
Σ / P	1.84	1.62	

注：真实百分比取自1990年的《美国统计摘要》（U. S. Statistical Abstract）。

Σ =平均估计值之和。

表13-1显示，对于概率和频率的判断，隐性析取（如自然原因导致的死亡）的平均估计值，表示为 Σ （自然原因），小于各组成部分（心脏病、癌症或其他自然原因）的平均估计值之和。具体来说，前者等于58%，而后者等于22%+18%+33%=73%。表13-1中的8项比较采用的是曼-惠特尼检验，结果均有统计学意义（ $p < 0.05$ ）。这里需要说明一点：我们使用的是非参数检验，因为在比较单个测量变量和一组测量变量时，所涉及的方差不相等。

在本文中，我们将分配给共延性显性和隐性假设的概率之比作为次可加性的尺度。上述示例中的比率是1.26。这个指标被称为拆分因子，可以直接从概率判断中计算出来，而不像 w 是由支持函数来定义的。次可加性由一个大于1的拆分因子和一个小于1的 w 值表示。值得注意的是，次可加性本身并不意味着相对于适当的客观标准显性假设被高估了或隐性假设被低估了，它仅仅表明前者比后者更有可能被判断。

在本研究中，3-因素假设的平均拆分因子为1.37，7-因素假设的平均拆分因子为1.92，这表明次可加性程度随显性析取因素数量的增加而增加。对中位数（而不是平均值）的分析显示出了类似的趋势，结合版本和拆分版本之间的差异稍微小一些。概率和频率任务的比较显示，正如预期的那样，被试在判断概率时给出的次可加性估计比判断频率时给出的次可加性估计要高（ $F(12, 101) = 2.03, p < 0.05$ ）。概率判断和频率判断的平均拆分因子分别为1.74和1.56。

相对于1990年的《美国统计摘要》中的真实数据，这些判断通常是高估的。唯一明显的例外是心脏病，该疾病的实际死亡概率为34%，但被试判断的平均值为20%。由于被试对概率的判断高于对频率的判断，因此在概率判断中对实际值的高估更多，但就估计值与实际值之间的相

关性（针对每名被试进行的单独计算）而言，两项任务之间没有差异。金钱的奖励也没有提高人们判断的准确度。

下面的设计是对支持理论的更严格的检验，我们将其与其他信念模型进行了比较。假设 A_1 、 A_2 和 B 是互斥且全面的， $A' = (A_1 \vee A_2)'$ ； A 是隐性的； $\bar{A}$ 是 A 的负值。考虑以下可观测的值：

$$\alpha = P(A, B)$$

$$\beta = P(A_1 \vee A_2, B)$$

$$\gamma_1 = P(A_1, A_2 \vee B), \gamma_2 = P(A_2, A_1 \vee B), \gamma = \gamma_1 + \gamma_2$$

以及

$$\delta_1 = P(A_1, \bar{A}_1), \delta_2 = (A_2, \bar{A}_2), \delta = \delta_1 + \delta_2$$

不同的信念模型意味着这些值的不同组合方式：

支持理论， $\alpha \leq \beta = \gamma \leq \delta$ ；

贝叶斯模型， $\alpha = \beta = \gamma = \delta$ ；

信念函数， $\alpha = \beta \geq \gamma = \delta$ ；

回归模型， $\alpha = \beta \leq \gamma = \delta$ 。

根据焦点和残差假设的拆分，支持理论预测 $\alpha \leq \beta$ 和 $\gamma \leq \delta$ ；它还根据可加的显性析取预测 $\beta = \gamma$ ；由外延性贝叶斯模型可得 $\alpha = \beta$ 和 $\gamma = \delta$ ；由可加性可得 $\beta = \gamma$ 。沙菲尔的信念函数理论还假设了外延性，但根据超可加性，它预测 $\beta \geq \gamma$ 。以上数据及大量研究综述表明， $\alpha < \delta$ ，这与支持理论一致，但不符合贝叶斯模型和沙菲尔的理论。

根据观察， $\alpha < \delta$ 也可以由回归模型来解释，该模型假设概率判断满足外延性但偏向于0.5（Erev, Wallsten & Budescu, 1994）。例如，评判者可能从一个为0.5的先验概率开始，这个概率没有根据证据得以充分的修正。随机误差也可能产生回归估计。如果每个人的判断都偏向0.5，那么由一个单一判断组成的 β 将小于 γ ，而 γ 为两个判断的总和。此

外，这个模型预测 α 和 β 之间没有差异性，每个都由一个单一判断组成，或在 γ 和 δ 之间，每一个都包含两个判断。因此，支持理论和回归模型对 α 和 δ 之间的差异的来源做出了不同的预测。支持理论预测了隐性析取的次可加性（如 $\alpha \leq \beta$ 和 $\gamma \leq \delta$ ）和显性析取的可加性（即相加性， $\beta = \gamma$ ），而回归模型假定了外延性（ $\alpha = \beta$ 和 $\gamma = \delta$ ）和显性析取的次可加性（ $\beta \leq \gamma$ ）。

为了对比这些预测，我们让不同的小组（每组25~30人）评估了各种非自然死亡原因的概率。所有被试都被告知，评估对象是从前一年死于非自然原因的人群中随机挑选的。表13-2总结了研究中的假设和相应的概率判断。例如，第一行表示死亡是由意外事故或自杀引起而不是由其他非自然原因引起的判断概率 β 。与支持理论一致， $\delta = \delta_1 + \delta_2$ 明显大于 $\gamma = \gamma_1 + \gamma_2$ （ $p < 0.05$ ，曼-惠特尼检验，但 γ 没有明显大于 β ，这与回归模型的预测不一致。然而，我们不排除这种可能性，即向0.5回归可能会使 $\beta < \gamma$ ，这将导致 α 和 $d\delta$ 之间的差异。在最后一节中，我们考虑了适用于这种模式的支持理论的概括。

表13-2 对各种死因的平均概率和中位概率估计值

概率判断	均值	中位数
$\beta = P$ （事故或自杀，OUC）	64	70
$\gamma_1 = P$ （事故，自杀或 OUC）	53	60
$\gamma_2 = P$ （自杀，事故或 OUC）	16	10
$\gamma = \gamma_1 + \gamma_2$	69	70
$\delta_1 = P$ （事故，OUC）	56	65
$\delta_2 = P$ （自杀，OUC）	24	18
$\delta = \delta_1 + \delta_2$	80	83

注：OUC=其他非自然原因。

研究2：暗示性和次可加性。在转向针对拆分的其他演示之前，我们先讨论一些引出概率判断的方法论问题。可以这样说，让被试评估一个特定的假设传达了一个微妙的（或不那么微妙的）暗示，即这个假设是很有可能。因此，被试可能会将引起他们注意的假设作为关于判断其概率的信息。为了解决这个问题，我们设计了一个任务，其中的假设不包含任何信息，因此任何观察到的次可加性都不能归因于实验建议。

我们请斯坦福大学的本科生（N=196）对拥有某一数量孩子的美国

已婚夫妇的比例进行估计。被试被要求写下他们的电话号码的最后一位数字，然后评估与该数字相同数量的孩子的已婚夫妇的比例。我们向被试承诺，回答最准确的3个人将每人获得10美元。正如预测的那样，数字0到9的总百分比（在不同被试中相加时）远远超过了1。各组分配的均值为1.99，中值之和为1.80。因此，即使在选择焦点假设中几乎没有得到相关信息的条件下，次可加性也是非常明显的。被试高估了除无子女夫妇外的所有类型的夫妇的百分比，对于有两个孩子的夫妇的百分比的估计与实际的百分比之间的差异性最大。此外，0、1、2和3个孩子的概率之和为1.45，每种情况的概率都超过了0.25。因此，观测到的次可加性不能仅仅用高估小概率的倾向来解释。

我们请另一组被试（ $N=139$ ）对有“少于3个”“3个或3个以上”“少于5个”或“5个或5个以上”孩子的美国已婚夫妇的比例进行估计。每名被试都只需要回答4个问题中的一个。第一种情况的估计值为97.5%，第二种情况的估计值为96.3%。与早些时候观察到的次可加性形成了鲜明的对比，正如支持理论所表明的那样，对互补事件的估计基本上是可加性的。二元互补的发现具有特殊的意义，因为它排除了对次可加性的另一种解释，根据这种解释，对证据的评价偏向于支持焦点假设。

专家判断中的次可加性。次可加性是仅限于新手，还是也适用于专家？D.雷德尔迈耶（D. Redelmeier）、德里克·J.凯勒（Derek J. Koehler）、V.利伯曼（V. Liberman）和特沃斯基以医学判断为背景探讨了这个问题（Redelmeier, Koehler, Liberman & Tversky, 1993）。他们向斯坦福大学的医生（ $N=59$ ）展示了一个关于一名女性因腹痛到急诊室就诊的详细情景。一半的被试被要求将概率分配给两个特定的诊断（胃肠炎和异位妊娠）和一个剩余的选项（不是上述中的任何一种）；另一半被试要将概率分配给5种特定的诊断（包括前一种情况下的两种）和一种剩余类别（不是上述的任何一种）。被试被要求给出的概率总和为1，因为所考虑的可能性是互斥的。如果医生的判断符合经典理论，那么两种诊断结果条件下分配给剩余类别的概率应等于5种诊断结果条件下分配给未拆分结果类型的概率之和。然而，与支持理论的预测相一致，在两种诊断结果的条件中，剩余类型的判断概率（均值=0.50）显著低于5种诊断结果条件下未拆分结果类型的判断概率（均值=0.69； $p<0.005$ ；曼-惠特尼检验）。

在第二项研究中，特拉维夫大学的医生（ $N=52$ ）被要求阅读一段医疗情景，包括患者的年龄、性别、病史、症状和由检验结果组成的医

疗场景的描述。例如，有一种情况是，一名67岁的男子来到急诊室，他几小时前心脏病发作了。每位医生被要求评估以下4种假设中的一种的概率：患者在入院期间死亡（A）；患者活着出院，但在1年内死亡（B）；患者剩余寿命为1~10年（C）；患者的剩余寿命超过10年（D）。在本文中，我们将这些称为基本判断，因为它们将基本假设与它的补充假设进行了比较，这些补充假设是所有其他基本假设的隐性析取。在评估了这4个假设中的一个之后，所有的被试都评估了 $P(A, B)$ 、 $P(B, C)$ 和 $P(C, D)$ 或互补集。我们将这些称为二元判断，因为它们只涉及两个基本假设的比较。

正如预测的那样，基本判断基本上是次可加性的。上述示例中的4种假设的均值：A组为14%，B组为26%，C组为55%，D组为69%，均超过了医学文献中的真实值。在此类问题中，当每个单独的因素都在与残差项进行对比的条件下被评估时，拆分因子的分母就取1；因此，拆分因子只是分配给不同因素的总概率（对不同被试组进行求和）。在本示例中，拆分因子为1.64。与之形成鲜明对比的是，由两组不同的医生做出的二元判断显示出了近乎完美的可加性，互补对的平均概率总和为100.5%。

福克斯、W. H.罗杰斯（W. H. Rogers）和特沃斯基对太平洋证券交易所的32名专业期权交易员进行了调查，为专家判断的次可加性提供了进一步的证据。这些交易员对微软股票在未来某一特定时期的收盘价做出了概率判断，如该收盘价将低于每股88美元（Fox, Rogers & Tversky, 1994）。微软股票在太平洋证券交易所交易，交易者通常关心的是对其未来价值的预测。然而，他们的判断显示出了次可加性和二元互补的预测模式。四分分区的平均拆分因子为1.47，互补二元事件的平均加和为0.98。以汽车修理工为研究对象的菲施霍夫等人

（Fischhoff, Slovic, Lichtenstein, 1978）和以餐馆经理为研究对象的L. 杜布-里乌（L. Dube-Rioux）和J.E.拉索（J.E. Russo）在其他领域也记录了专家判断的次可加性（Dube-Rioux & Russo, 1988）。

回顾以前的研究。接下来我们对其他检验了支持理论的研究进行回顾。特沃斯基和福克斯要求被试将概率分配给数量不确定的不同区间，如即将进行的超级碗比赛的胜率或下周的道琼斯指数的变化（Tversky & Fox, 1994）。当一个给定事件（如“水牛队打败华盛顿队”）被拆分成单独的子事件（如“水牛队胜华盛顿队不到7分”和“水牛队胜华盛顿队至少7分”）时，被试的判断基本上是次可加性的。本研究得到的拆分

因子以显性析取中分量假设数的函数的形式呈现，如图13-1所示。其中共有对5种不同类型的事件的判断：未来的圣弗朗西斯科气温（SFO）、未来的东京气温（TKY）、美国国家橄榄球联盟（NFL）超级碗比赛的结果、美国国家篮球协会（NBA）季后赛的结果以及道琼斯指数的每周变化。回想一下，拆分因子大于1（即落在图中虚线以上）表示次可加性。图13-1所示的结果显示了所有来源的一致次可加性，其随着显性析取中子事件的数量增加而增加。

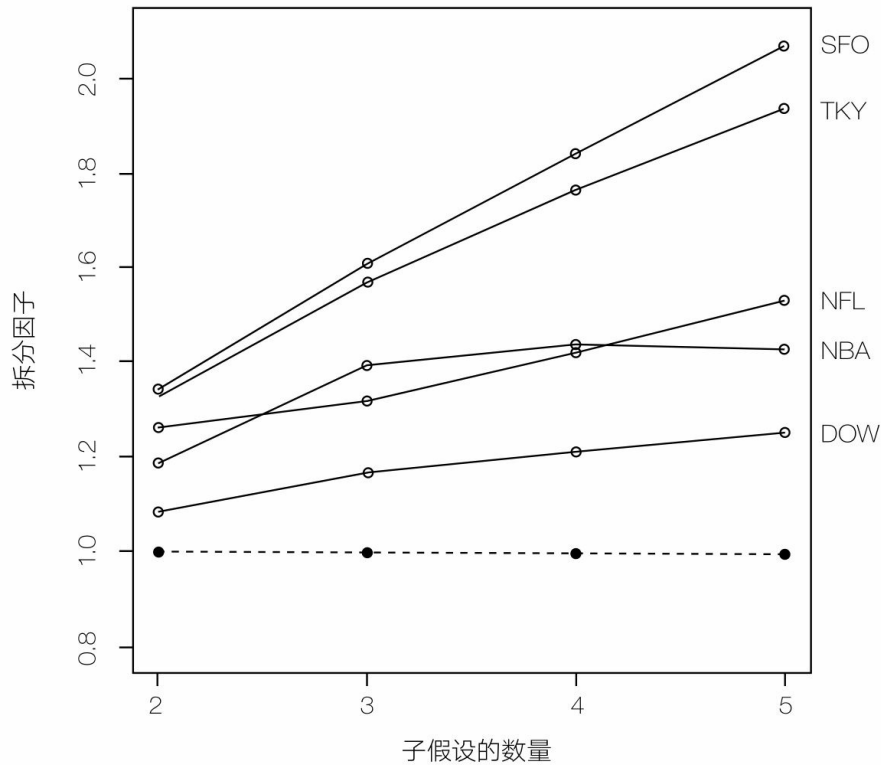


图13-1 特沃斯基和福克斯数据的拆分因子

注：SFO=圣弗朗西斯科气温；TKY=东京气温；NFL=1991年美国国家橄榄球联盟超级碗比赛；NBA=国家篮球协会季后赛；DOW=道琼斯指数周变化。

分配给互补假设的概率中位数如图13-2所示。每个假设在图中表示两次，一次作为焦点假设，一次作为补充假设。根据支持理论的预测，互补假设对的区间判断在本质上是可加性的，没有明显的次可加性或超可加性倾向。

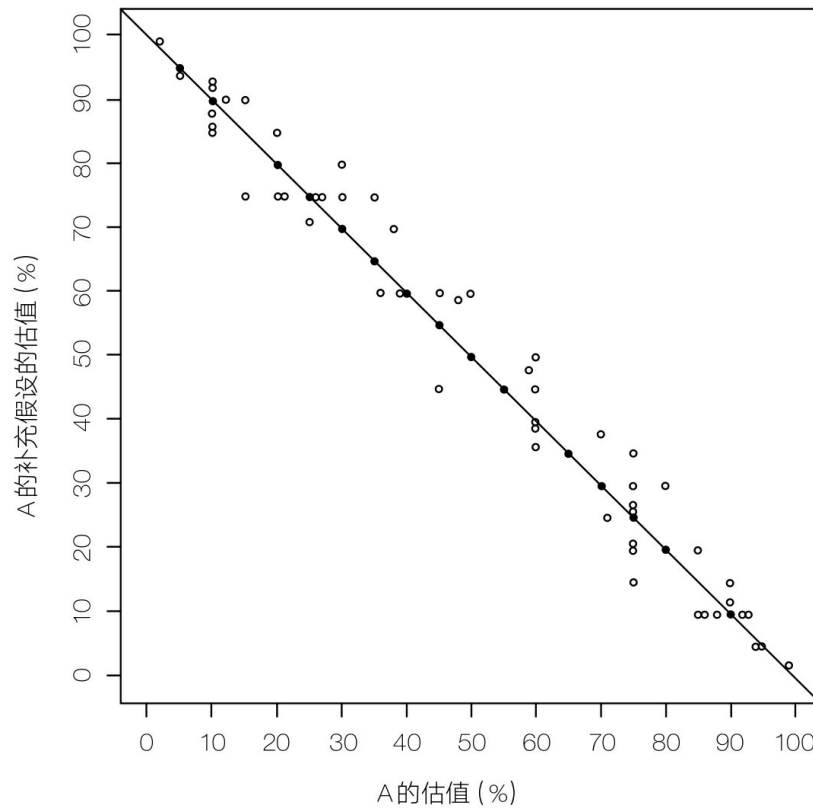


图13-2 基于特沃斯基和福克斯的二元互补检验

二元互补的更多证据来自瓦尔斯滕、D. V.布德斯库和R.兹维克（R. Zwick）的一项研究，他们向被试提供了300个关于世界历史和地理的命题，如“门罗主义是在共和党成立之前宣布的”，并要求他们估计每一个命题为正确的概率（Wallsten, Budescu, Zwick, 1992）。(36)每个命题的真假（互补）版本的不同之处在于日期。图13-3使用图13-2的格式显示了被试认为每个命题为真和为假的平均概率。同样，判断在整个范围内是可加性的（均值=1.02）。

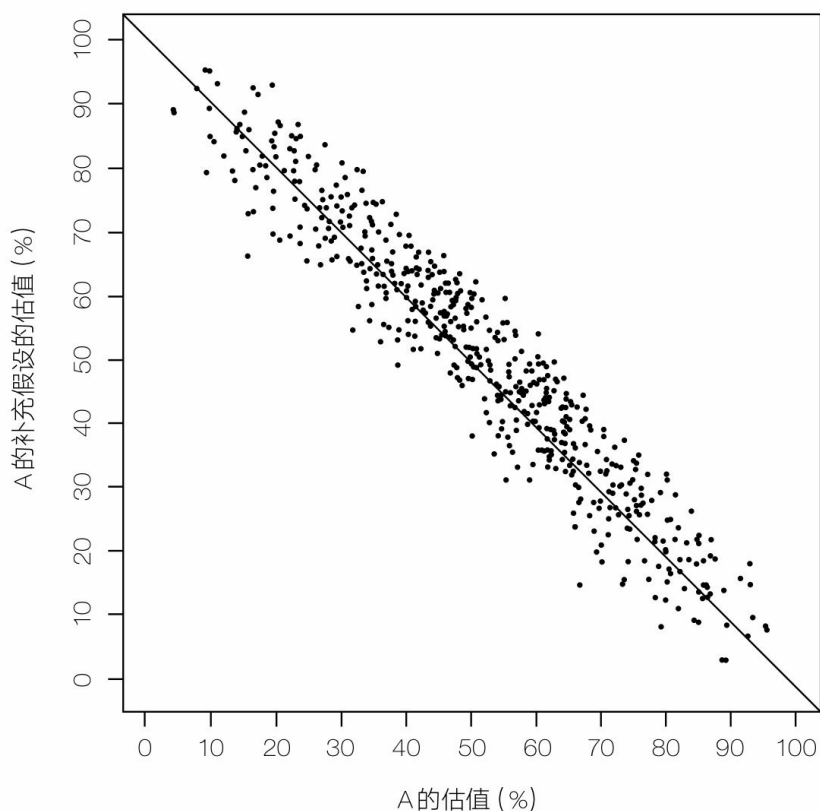


图13-3 基于瓦尔斯滕、布德斯库和兹维克的二元互补检验示意

接下来，我们将对主要发现进行简要总结，并列出支持每个结论的研究。

次可加性：将一个隐含的假设拆分为它的子假设增加了其总概率，从而产生了次可加性判断。表13-3和表13-4列出了提供拆分条件的检验的研究。每个实验中隐性假设的分配概率和显性析取中各组成部分的总概率都在列表中显示，同时还列出了由此产生的拆分因子。所有列出的研究都使用了一种实验设计，其中隐性析取和显性析取的组成部分分别由不同的被试组或相同的被试组进行评估，但存在大量的干预性判断。在显性析取中概率以子集数量的函数的形式呈现，并在所有其他自变量上折叠。表13-3列出了被试评估定性假设的概率，如比尔主修心理学的概率；表13-4列出了被试评估的定量假设的研究数据，如随机选择的成年男性身高为1.82~1.87米的概率。

表13-3 使用定性假设的实验结果：分配给共延隐性和显性析取的平均概率，以及测量次可加性的拆分因子

研究和主题	<i>n</i>	显性 <i>P</i>	隐性 <i>P</i>	拆分因子
Fischhoff, Slovic & Lichtenstein (1978)				
汽车失灵, 实验 1	4	0.54	0.18	3.00
汽车失灵, 实验 5	2	0.27	0.20	1.35
汽车失灵, 实验 6 (专家)	4	0.44	0.22	2.00
Mehle, Gettys, Manning, Baca & Fisher (1981): 大学专业				
	6	0.27	0.18	1.50
Russo & Kolzow (1992)				
死亡原因	4	0.55	0.45	1.22
汽车失灵	4	0.55	0.27	2.04
Koehler & Tversky (1993)				
大学专业	4	1.54	1.00 ^a	1.54
大学专业	5	2.51	1.00 ^a	2.51
研究 1: 死亡原因	3	0.61	0.46	1.33
	7	0.70	0.37	1.86
研究 4: 犯罪故事	4	1.71	1.00 ^a	1.71
研究 5: 大学专业	4	1.76	1.00 ^a	1.76

注：显性析取中的组分数量用*n*表示。无引文的编号研究请参考本文。

a: 由于组分分割了空间，因此假定一个会被分配到隐性析取的概率为1.00。

表13-4 定量假设的实验结果：分配给共延隐性析取和显性析取的概率，以及测量次可加性程度的拆分因子

研究和主题	n	显性 P	隐性 P	拆分因子
Teigen (1974b)				
实验 1: 二项式结果	2	0.66	0.38	1.73
	3	0.84	0.38	2.21
	5	1.62	1.00 ^a	1.62
	9	2.25	1.00 ^a	2.25
Teigen (1974b)				
实验 2: 学生的身高	2	0.58	0.36	1.61
	4	1.99	0.76	2.62
	5	2.31	0.75	3.07
	6	2.55	1.00a	2.55
Teigen (1974a)				
实验 2: 二项式结果	11	4.25	1.00a	4.25
Olson (1976)				
实验 1: 性别分布	2	0.13	0.10	1.30
	3	0.36	0.21	1.71
	5	0.68	0.40	1.70
	9	0.97	0.38	2.55
Peterson & Pitz (1988)				
实验 3: 棒球比赛胜利	3	1.58	1.00 ^a	1.58
Tversky & Fox (1994): 不确定性数量	2	0.77	0.62	1.27
	3	1.02	0.72	1.46
	4	1.21	0.79	1.58
	5	1.40	0.84	1.27
研究 2: 孩子的数量	10	1.99	1.00 ^a	1.99

注: 显性析取中的组分数量用 n 表示。无引文的编号研究请参考本文。

a: 由于组分分割了空间, 因此假定一个会被分配到隐性析取的概率为1.00。

从表中我们可以看出, 所观察到的拆分因子无一例外地都大于1,

这表明了一致的次可加性。我们在定性和定量假设中都观察到了次可加性，这一发现具有指导意义。在定性假设的评估中，当所讨论的事件以隐性的形式被描述时，次可加性至少可以部分地被解释为没有考虑到一个或多个组分假设。然而，在定量假设的判断中，观察到的次可加性不能被解释为这种提取的失败。例如，泰甘发现，在判断大学生身高落在给定区间的任务中，将区间划分为可单独评判的更小区间时，这种判断的百分比会增加（Teigen, 1974b, 实验2）。对于评估隐性析取事件（如更大的区间）的被试，我们认为，他们没有忽略给定区间包括几个更小区间的事实；相反，拆分操作增强了这些区间的显著性，从而提高了它们的概率判断。因此，即使在没有记忆限制的情况下，也可以观察到次可加性。

组分数目：在显性析取中，次可加性的程度随着组分数目的增加而增加。这很容易从支持理论中得到证实：将隐性假设拆分为互斥的组分会增加其总的判断概率，并且每个组分的额外拆分将进一步增加对初始假设的总概率的判断。正如所料，表13-3和表13-4显示了拆分因子通常随着组分数目的增加而增加（如图13-1所示）。

二元互补：互补假设的判断概率加和应该为1。表13-5列出了检验这一假设的研究。我们只列出了假设及其补充假设被独立评估时的研究，其中要么是由不同的被试进行评估，要么由相同的被试进行评估，但有大量的判断干预。我们列出了使用后一种实验设计的结果。表13-5显示，这些判断通常是可以加和为1的。二元互补表示人们会根据给定假设的补充假设对其进行评价。此外，它排除了对次可加性的另一种解释，即暗示效应或确认性偏差（confirmation bias）。与实验的结果相反，这些解释暗示了一种偏向于产生 $P(A, B) + P(B, A) > 1$ 的焦点假设。另一种可能是，人们可能倾向于将概率判断中观察到的次可加性归因于被试对概率理论的可加性原理缺乏了解。然而，这不能解释所观察到的频率判断中的次可加性（该概率判断中有明显的可加性）和二元互补中的次可加性（始终满足可加性）。

表13-5 检验二元互补的实验结果：分配给互补假设对的平均总概率、被试之间的标准差，以及实验中被试的人数

研究和主题	平均总和 P	标准差	被试人数
Wallsten, Budescu & Zwick (1992): 一般性知识	1.02	0.06	23
Tversky & Fox (1994)			
NBA 季后赛	1.00	0.07	27
超级碗	1.02	0.07	40
道琼斯指数	1.00	0.10	40
圣弗朗西斯科气温	1.02	0.13	72
东京气温	0.99	0.14	45
大学专业	1.00		170
研究 2: 孩子的数量 ^a	0.97		139
研究 4: 犯罪故事 ^a	1.03		60
研究 5: 大学专业 ^a	1.05		115

注：无引文的编号研究请参考本文。

^a表示一名被试只能评价事件或评价它的互补事件，但不能同时评价两个事件。

支持理论所暗示的二元互补与具有次可加性的基础判断的结合，既不符合贝叶斯模型，也不符合修正主义模型。根据贝叶斯模型，拆分因子应该等于1，因为拆分和结合假设具有相同的外延。沙菲尔的信念函数理论和其他低概率模型需要一个小于1的拆分因子，因为它们假设不相交事件并集的主观概率（或信念）通常大于其互斥成分的概率之和。此外，这些数据不能用信念函数的二元性（被称为似然函数）来解释，或者更一般的说法就是，不能用较高的概率来解释，因为这个模型要求对互补事件的判断概率之和大于1，这与事实正好相反（Dempster, 1967）。事实上，如果 $P(A, B) + P(B, A) = 1$ （如表13-5所示），那么概率上限和下限均降为标准可加性模型。当然，实验结果并没有否定使用作为表示不确定性的正式系统的概率上限、下限或信念函数。然而，本节回顾的证据表明，这些模型与引导直觉概率判断的原则不一致。

概率与频率：在表13-3和表13-4中列出了先前讨论的概率与频率之间的关系的研究，其中一些研究使用频率判断而另一些研究则使用了概率判断（Fischhoff et al., 1978; Teigen, 1974a, 1974b）。两种任务的比较的总结如表13-6所示，它证实了预测的模式：概率判断和频率判断中都存在次可加性，且前者比后者的次可加性更高。

表13-6 比较概率判断和频率判断的实验结果：由分配给共延显性和隐性析取的平均概率计算的拆分因子

研究和主题	<i>n</i>	拆分因子	
		概率	频率
实验 1：二项式结果	2	1.73	1.26
	5	2.21	1.09
	9	2.25	1.24
实验 2：学生的身高	6	2.55	1.68
大学专业	4	1.72	1.37
研究 1：死亡原因	3	1.44	1.28
	7	2.00	1.84

注：显性析取中的组分数量用*n*表示。

测量支持度

在上一节的形式理论中，支持函数是由概率判断得到的。我们是否有可能逆转这一过程，并通过证据强度的直接评估来预测概率判断？设 $\hat{s}(A)$ 作为对支持假设 A 的证据强度的评估。那么，这样的评级与源自概率判断的支持度之间有什么关系？也许最自然的假设是这两个维度是单调相关的；当且仅当 $s(A) \geq s(B)$ ， $\hat{s}(A) \geq \hat{s}(B)$ 。这个假设意味着，当且仅当 $\hat{s}(A) \geq \hat{s}(B)$ ，则 $P(A, B) \geq 1/2$ 。但它没有确定 $\hat{s}$ 和 s 的函数形式。为了进一步明确这些测量之间的关系，我们可以合理地假设支持度比率也是单调相关的。也就是说，当且仅当 $s(A)/s(B) \geq s(C)/s(D)$ ， $\hat{s}(A)/\hat{s}(B) \geq \hat{s}(C)/\hat{s}(D)$ 。

我们可以看出，如果满足两个单调条件，并且在单位区间上定义两个测量尺度，那么存在一个常数 $k > 0$ ，使得由概率判断得到的支持函数和直接评估的支持函数以 $s = \hat{s}^k$ 的幂转换形式相关。这就产生了幂模型

$$R(A, B) = P(A, B) / P(B, A) = [\hat{s}(A) / \hat{s}(B)]^k$$

表明

$$\log R(A, B) = k \log [\hat{s}(A) / \hat{s}(B)]$$

接下来，我们将使用该模型预测两个研究中获得的证据强度的独立

评估的判断概率。

研究3：篮球比赛。被试（ $N=88$ ）都是在电脑上订阅了新闻组的NBA球迷。我们在这个新闻组中发布了一份问卷，要求读者在一周内完成并以电子邮件的方式回复。在问卷调查中，被试评估了主队在接下来的20场比赛中获胜的概率。这20个结果构成了NBA太平洋赛区5支球队（菲尼克斯队、波特兰队、洛杉矶湖人队、金州勇士队和萨克拉门托队）之间所有可能的比赛，这样一来，对于每对球队都有两场比赛需要被评估（根据每个可能的比赛地点依次评估）。使用这一“专家”群体可以产生高度可靠的判断数据；事实表明，单名被试的评分与一组被试的平均判断之间的相关性中位数为0.93。

在做出概率判断后，被试对5支队伍的实力进行了打分。被试被告知：

首先，从5支球队中选择你认为最强的那支，然后把这支球队的实力设为100分。根据最强球队的实力，给剩下的球队打分。例如，如果你认为某支球队的实力只有最强球队（你给了100分的那支球队）的一半，那就给这支球队打50分。

我们将这些评级解释为对支持度的直接评估。

由于实力评级没有考虑到主场效应，我们将比赛中两个可能的比赛地点的概率判断进行了折叠。从 $\log [\hat{s}(A) / \hat{s}(B)]$ 预测 $\log R(A, B)$ 的回归线斜率为每名被试都提供了 k 的估计值。 k 的中位数估计值为1.8，均值为2.2；这个分析的中位数 R^2 是0.87。对于整合后的数据， k 是1.9，由此产生的 R^2 是0.97。图13-4中的散点图显示了基于球队实力的平均预测与平均判断概率之间的良好对应关系。这一结果表明，幂模型可用来在不参考偶然性或不确定性的强度评估中预测概率。它也加强了 s 作为证据强度衡量标准的心理学解释。

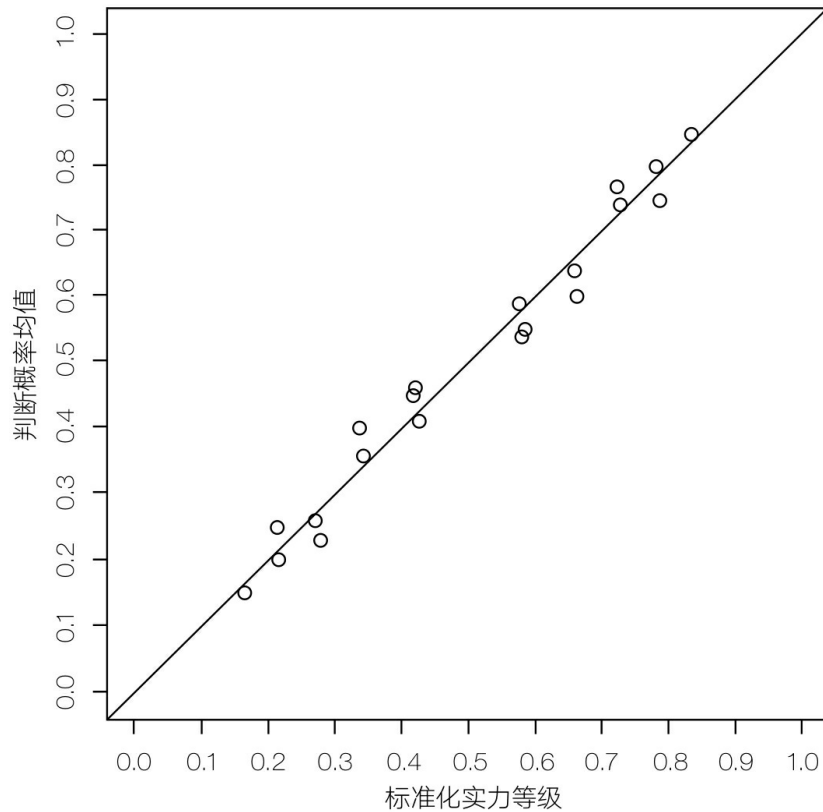


图13-4 作为标准化等级函数的篮球比赛的概率判断

研究4：犯罪故事。本研究旨在考察判断概率与不同情境中的支持度评估之间的关系，而且我们会在下一部分中探讨强化效应。为此，我们采用了泰甘及L. B.罗宾逊（L. B. Robinson）和R.黑斯蒂（R. Hastie）提出的一项任务，向被试展示了两个刑事案件（Teigen, 1983; Robinson & Hastie, 1985）。第一起是发生在计算机零部件制造公司的盗用案件，涉及4名嫌疑人（1名经理、1名买方、1名会计和1名销售员）。第二个案件是一起谋杀案，涉及4名嫌疑人（1名活动家、1名艺术家、1名科学家和1名作家）。在这两种情况下，被试都被告知只有1名嫌疑人有罪。在信息量较少的情况下，我们对每个案件的4名嫌疑人都进行了介绍，并简要描述了他们的身份和可能的动机。在信息量较大的情况下，犯罪嫌疑人的作案动机得到了强化。我们以一种类似于典型的推理小说的方式构建了每一个案例，这样一来随着更多的证据被揭露，所有的嫌疑人似乎都变得更加可疑了。

被试分别在阅读信息含量较少的材料后和阅读信息含量较多的材料后对嫌疑人进行评估。一些被试（N=60）需要判断嫌疑人有罪的概率。每名被试在每个案例中都做出了2个基本判断（一个特定的嫌疑人

有罪)和3个二元判断(嫌疑人A而不是嫌疑人B有罪)。其他被试($N=55$)对特定的嫌疑人的可疑程度进行评分,作为支持度的直接评估。这些被试通过给出一个介于0(表示该嫌疑人“一点儿也不可疑”)和100(表示该嫌疑人“最大程度上可疑”)之间的数字(表示该嫌疑人的可疑程度)来对每个案件中的2名嫌疑人进行评级。

与之前的研究一样,我们假设了二元互补关系,通过 $R(A, B)$ 与怀疑率的对数回归估计 k 。对于这些数据, k 的估计值是0.84, R^2 的估计值是0.65。因此,对怀疑度的评级为有罪判断提供了一个合理的预测。然而,篮球赛事研究中判断概率与评估支持度之间的关系强于犯罪案件研究。此外, k 的估计值在后者中比在前者中小得多。在上述的篮球研究中,一支实力是另一支球队实力两倍的球队,其获胜的概率也是后者的两倍。然而,在犯罪的故事中,一个可疑性比另一个可疑性多一倍的嫌疑人被判有罪的概率不足后者的两倍。这种差异可能是因为对球队实力的判断是基于更可靠的数据,而不是基于对可疑性的评级。

在前两项研究中,我们要求被试在所有现有证据的基础上评估每个假设的总体支持度。L. K.布里格斯(L. K. Briggs)和克兰茨采用了另一种不同的评估方式(Briggs & Krantz, 1992; Krantz, Ray & Briggs, 1990)。他们的研究表明,在一定条件下,被试可以评估一个孤立的证据在多大程度上可以支持所考虑的每一个假设。他们还提出了若干条将独立证据项目整合起来的规则,但没有将评估的支持度与判断概率联系起来。

强化效应

回想一下,支持度评估是非补偿性的,也就是说增加一个支持假设的证据并不一定会减少竞争假设的支持度。事实上,新的证据有可能增加对所有基本假设的支持。我们认为,这样的证据将增强可加性。在本节中,我们描述了有关强化的几个研究,并将支持理论与贝叶斯模型和沙菲尔理论进行了比较。

我们从前面讨论过的一个例子开始。在这个例子中,几个嫌疑人中的一人犯了谋杀罪。为了简化问题,假设有4名嫌疑人,在缺乏具体证据(低信息量较)的情况下,4名嫌疑人被认为有罪的可能性一样。假设进一步的证据(高信息量较)被引入,使每一个嫌疑人的可疑程度大致相同,因此他们的可疑程度还是相同的。设 L 和 H 分别为信息量较少

条件下和信息量较多条件下可用的证据。让 $\bar{A}$ 表示 A 的负值，也就是说它代表“嫌疑人 A 是无罪的”。那么，根据贝叶斯模型， $P(A, B|H) = P(A, B|L) = 1/2$ ， $P(A, \bar{A}|H) = P(A, \bar{A}|L) = 1/4$ ，依此类推。

与此相反，沙菲尔的信念函数采用的方法要求分配给每个嫌疑人的概率相加起来小于1，并表明在有直接证据的情况下（即信息量较多的条件下），总和将会比没有直接证据的情况下更高（Shafir, 1976）。因此， $1/2 \geq P(A, B|H) \geq P(A, B|L)$ ， $1/4 \geq P(A, \bar{A}|H) \geq P(A, \bar{A}|L)$ ，依此类推。换句话说，随着更多证据的出现，二元判断和基本判断都有望增加。在限制性情况下，当没有信念被保留时，二元判断接近1/2，基本判断接近1/4。

强化假设提出了一个不同的模式，即 $P(A, B|H) = P(A, B|L) = 1/2$ ， $P(A, \bar{A}|H) \geq P(A, \bar{A}|L) \geq 1/4$ ，依此类推。就像贝叶斯法则模型那样，二元判断是1/2；然而，与该模型相比，在信息量较多的条件下的基本判断预计将超过1/4，并且比在信息量较少的条件下的判断更大。虽然支持理论和信念函数方法在高信息量条件下都比在低信息量条件下产生了更大的基本判断，但是支持理论预测在两种条件下它们都将超过1/4，而沙菲尔理论要求这些概率小于或等于1/4。

假设同样可疑的嫌疑人对分析来说并非必要的。假设一开始嫌疑人的犯罪概率不相等，但是新证据并没有改变二元概率。因此，在这里，贝叶斯模型也需要在低信息量和高信息量的条件之间不加区分的可加性判断；信念函数方法需要超可加性判断，当信息量更多时，判断的超可加性就会降低；而且，强化假设预测，随着（相容的）证据的增加次可加性判断会变得更具有次可加性。

评估嫌疑人。带着这些预测，我们转向研究4中的犯罪故事。表13-7显示了在低信息量和高信息量的条件下，两起案件中每个嫌疑人的平均可疑度评分和基本概率判断。根据表可知，在所有情况下，概率判断和可疑程度评级的总和都大于1。显然，次可加性不仅适用于概率判断，而且适用于证据强度或信念程度（如某一嫌疑人是有罪的）的评级。对可疑程度的进一步调查发现，随着提供的信息越来越多，除了一名嫌疑人以外，所有人的可疑程度都有所增加。与我们的预测一致，这些嫌疑人的判断概率也随着信息量的增加而增加了，这表明次可加性增强了（见公式10）。唯一的例外是谋杀案中的艺术家，他在高信息量的情况下得到了不在场证明，正如人们所预料的那样，他后来的可疑性和评估都降低了。总体而言，在高信息量的条件下的可疑性评分和概率判

断都显著高于低信息量条件下的可疑性评分和概率判断（通过t检验，两种情况下的 $p<0.001$ ）。

表13-7 高 / 低信息量条件下每个嫌疑人的可疑程度和判断概率均值

案件和嫌疑人	可疑程度		判断概率	
	低信息量	高信息量	低信息量	高信息量
案件 1：盗用案				
会计	41	53	40	45
买方	50	58	42	48
经理	47	51	48	59
销售员	32	48	37	42
总计	170	210	167	194
案件 2：谋杀案				
社会活动家	32	57	39	57
艺术家	27	23	37	30
科学家	24	43	34	40
作家	38	60	33	54
总计	122	184	143	181

从规范性的角度来看，所有嫌疑人的支持度（如可疑性）都可以随着新信息的增加而增加，但是增加一个疑点的概率应该通过减少其他疑点的概率来补偿。罗宾逊和黑斯蒂之前发现，新的证据会增加对所有嫌疑人犯罪概率的评估（Robinson & Hastie, 1985; van Wallendael & Hastie, 1990）。他们的方法与我们的不同之处在于，他们的每名被试都评估了所有嫌疑人的概率，但这种方法也产生了大量的次可加性，典型的拆分因子系数约为2。他们拒绝将贝叶斯模型作为一种描述性解释，并提出沙菲尔的理论是一种可行的选择。然而，正如前面所指出的那样，观测到的次可加性与沙菲尔理论及贝叶斯模型不一致，但与目前的解释是一致的。

在犯罪故事中，增加的证据通常与正在考虑的所有假设相符。然而，彼得森和G.F.皮茨（G.F. Pitz）通过混合证据观察到了类似的效

果，因此支持了一些假设也反驳了另一些假设（Peterson & Pitz, 1988, 实验3）。他们要求被试根据1个、2个或3个线索（球队击球率、投手责任失分和该赛季的全垒打总数）来评估一个棒球队在某个时间区间内获胜次数下降的概率。在被试不知道的情况下要求他们在大量的问题中，将概率分配到这3个因素上，如少于80胜、介于80胜和88胜之间、超88场胜。平均来看，随着信息量的增加，被试将更大的概率分配给分区中的所有3个区间，从而显示出强化了次可加性。这些数据的拆分因子分别为1.26、1.61和1.86。这些结果证实了强化效应的稳定性，即使新增证据支持一些（但不是全部）研究的假设，也能观察到强化效应。

研究5：大学专业。在本研究中，我们通过替换证据而不是如之前研究中的添加证据来检验强化效应。我们沿用了T.梅尔（T. Mehle）、C.F.格蒂斯（C.F. Gettys）、C.曼宁（C. Manning）、S.巴卡（S. Baca）和S.菲舍尔（S. Fisher）的方法，要求被试（ $N=115$ ）评估某所未指明的美国中西部大学中某个社会科学专业学生主修某一特定专业的概率（Mehle, Gettys, Manning, Baca & Fisher, 1981）。被试被告知，在这所大学里，每名社会科学专业的学生都有且只有以下4个专业的选择：经济学、政治学、心理学和社会学。

被试根据学生在二年级修过的4门课程中的一门来估计他们主修某一专业的概率。其中两门课程（统计学和西方文明）是社会科学专业的典型课程；另外两门课（法国文学和物理）通常不是社会科学专业学生会选的课。一组独立的被试（ $N=36$ ）评估了一个社会科学专业的学生选修4门课程中的每一门的概率。根据强化效应，典型课程比不那么典型的课程将产生更多的次可加性，因为它们对4个专业中的每一个都具有更大的支持度。

每名被试都做出了基本判断和二元判断。在以往的研究中，基本判断会表现出明显的次可加性（平均拆分因子=1.76），而二元判断本质上是可加性的（平均拆分因子=1.05）。在前面的分析中，我们使用拆分因子作为一组互斥假设的次可加性的总体尺度。我们还通过该实验评估了 w （见公式8），这为次可加性提供了一个更为精确的尺度，因为它分别对研究中的每个隐性假设进行了估计。对于每门课程，我们首先从二元判断中估计各专业的支持度，然后利用公式从基本判断中估计各专业的 w ，公式如下

$$P(A, \bar{A}) = \frac{s(A)}{s(A) + w_{\bar{A}}[s(B) + s(C) + s(D)]}$$

其中A、B、C、D分别代表4个专业。

这种分析是针对每名被试分别进行的。各门课程和专业的w的平均值为0.46，这表明当一个专业被隐性地包含在残差中时，它得到的支持度还不到其显性支持度的一半。图13-5显示了每个专业的w（除以被试人数）的中位数，分别为4门课程的中位数。与强化效应一致，该图显示，典型的课程（统计学和西方文明）比不太典型的课程（物理学和法国文学）诱导了更多的次可加性（即w更低）。然而，对于任何一门具体的课程，各个专业的w值大致相同。当然，双向方差分析得出了非常显著的结果（ $F(3, 112) = 31.4, p < 0.001$ ），但是专业方面并不显著（ $F(3, 112) < 1$ ）。

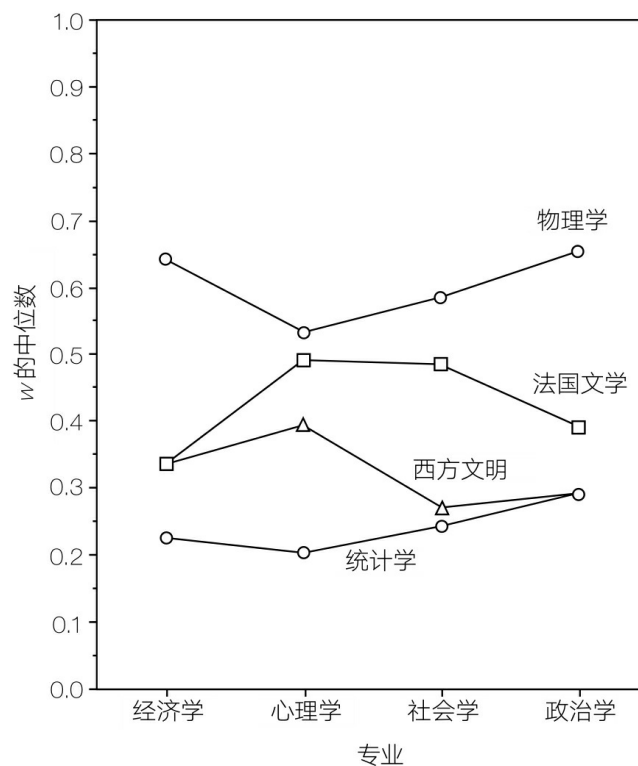


图13-5 大学专业预测的中位数w相对于每门课程的散点

意义

关于这一点，我们关注的是支持理论的直接后果。最后，我们从支

持理论的角度讨论了不确定状况下的合取效应、假设生成和决策。

合取效应。大量的研究都发现了合取效应， AB 合取事件的判断概率可能会超过它的子事件 A 的概率。当一个起初似乎是不太可能的事件（如一场巨大的洪灾在北美造成1 000多人死亡），得到一些看似合理的原因和说明的补充后（如在加州地震造成的洪水淹死了1 000多人），产生了一个被认为比最初不太可能发生的事件更有可能的合取事件时，这个效应是最强的，此时最初的事件是合取事件的子集

（Tversky & Kahneman, 1983）。支持理论表明，隐性假设 A 没有被拆分为共延的析取事件 $AB \vee \bar{A}B$ ，合取也是其中的一个组成部分。结果是支持 AB 的证据并没有被用来支持 A 。例如，在洪水问题上，地震引起洪水的可能性也许不会马上被想到；因此，除非明确提到，否则它不会对（隐性的）洪水假设提供任何支持。如果隐性析取在评估之前就进行了拆分的话（如果提醒被试洪水可能是由过多的降雨造成的，或由于地震、工程错误和破坏行为等造成了储水层的结构性破坏），根据支持理论合取效应应在这些问题中会消失。

频率评估中对拆分焦点假设或剩余假设的倾向可能有助于解释为什么当被试估计频率而不是概率时，合取效应会减弱但不会消失。例如，相对于频率判断，被试在概率判断中判断“ X 是55岁以上，至少有一次心脏病发作”的合取事件，比组成它的子事件“ X 至少有一次心脏病发作”更有可能（Tversky & Kahneman, 1983）。

区分两种不同的拆分操作可能具有指导意义。在合取事件的拆分中，一个（隐性的）假设（如护士）被拆分为互斥的合取事件（如男护士和女护士）。大多数但不是所有的合取效应的最初表现是基于合取拆分的。在种类拆分中，一个更高级的范畴（如非自然的死亡）会被拆分为“自然”的组成部分（如车祸、溺水和杀人）。本文中报告的大多数案例都是基于种类拆分的。M.巴尔-希勒尔（M. Bar-Hillel）和E.耐特尔（E. Neter）描述了使用种类拆分的合取效应，他们发现在许多情况下，一个陈述（如“丹妮拉的专业是文学”）被认为比一个包含更多内容的隐性析取（如“丹妮拉的专业是人文学科”）更有可能发生（Bar-Hillel & Neter, 1993）。这些结果既适用于被试对概率的直接估计，也适用于他们对相关事件投注的意愿。

假设生成。到目前为止，我们回顾的所有研究都要求被试评估他们进行判断的假设的概率。然而，在许多情况下，评判者必须提出假设并评估其可能性。在目前的处理中，替代假设的生成需要对残差假设进行

一些拆分，因此它相对于焦点假设的相对支持度也应该会增加。在缺乏生成替代假设的明确指示的情况下，人们不太可能分解残差假设，因此相对于那些未指定的假设，人们往往会高估被指定的假设。

这一影响已经被格蒂斯和他的同事所证实（Gettys, Mehle & Fisher, 1986; Mehle et al., 1981）。他们发现，与真实值相比，人们通常倾向于高估他们进行评估的特定假设的概率。事实上，对自己的判断过度自信有时可能出现，因为焦点假设是特定的，而它的替代性假设往往不是（Lichtenstein, Fischhoff & Phillips, 1982）。梅尔等人使用了两种操作来鼓励人们拆分残差假设：一组被试看到的是残差假设的例子，另一组被试被要求生成自己的示例（Mehle et al., 1981）。这两种操作都通过降低分配给指定的替代性方案的概率和增加分配给残差方案的概率来提高表现。结果表明，假设生成的效果是由人们想到的额外假设引起的，因为简单地提供假设给被试能产生相同的效果。杜布-里乌和拉索使用类似的操作发现，替代假设的产生增加了相对于特定类别的残差假设的概率判断，并减小了忽略类别产生的影响（Dube-Rioux & Russo, 1988）。对由被试自发产生的示例数量的分析发现，当他们产生足够多的示例时，忽略类别的影响可以完全被消除。

现在，假设有一个任务，在评估其概率之前，我们要求被试生成一个假设，如猜测哪部电影将在下一届奥斯卡颁奖典礼上获得最佳影片奖。我们让被试生成最有可能的假设，实际上这可能是在让他们选择自己喜欢的几个候选影片之一。这个过程等于一个拆分残差假设的过程，这降低了焦点假设的判断概率。与这一预测相一致的是，最近的一项研究发现，被要求生成自己的假设的被试比其他看到相同假设并进行评估的被试给出了更低的判断概率（Koehler, 1994）。对这些结果的解释——假设的生成使替代性假设更加显著——通过两个进一步的操作得到了验证。首先，我们提供一组指定的消除了生成和评估条件之间的差异的备选方案。在这两种情况下，残差应以同样的方式进行表示。其次，在生成假设和概率评估之间插入一个干扰性任务，足以降低生成任务所带来的备选方案的显著性，从而增加了对焦点假设的判断概率。

不确定状况下的决策。本文主要关注概率的数值判断。然而，在决策理论中，主观概率通常是从不确定前景之间的偏好中推断出来的，而不是直接评估出来的。因此，人们很自然地会问，拆分是否会影响人们的决策，以及影响他们对数值的判断。有相当多的证据表明确实如此。例如，约翰逊等人观察到，相对于没有列出具体事件的更为包容的保

险，如任何原因导致的死亡，被试愿意为明确列出保单涵盖事件的保险支付更高的费用（如恐怖主义事件或机械事故造成的死亡；Johnson et al., 1993）。

拆分可以从两个方面影响决策。首先，正如前面已经指出的那样，拆分往往会增加不确定性事件的判断概率。其次，即使事件的概率已知，拆分也会增加事件对决策的影响。例如，特沃斯基和卡尼曼要求被试在两种彩票中做出选择，这两种彩票根据从盒子中取出的弹珠的不同颜色支付不同金额（作为谨慎考虑选项的一个诱因，被试被告知，随机选择的参与者中有1/10的人会真正参与他们选择的投注；Tversky & Kahneman, 1986）。他们使用的两个问题只是在对结果的描述上有所不同。完全拆分的版本1如下所示：

盒子A:	90% 白色	6% 红色	1% 绿色	1% 蓝色	2% 黄色
	0 美元	赢 45 美元	赢 30 美元	输 15 美元	输 15 美元
盒子B:	90% 白色	6% 红色	1% 绿色	1% 蓝色	2% 黄色
	0 美元	赢 45 美元	赢 30 美元	输 15 美元	输 15 美元

不难看出，盒子B优于盒子A；事实上，这个版本中所有的被试确实都选择了盒子B。版本2是由将盒子A中两种导致损失15美元的结果（即蓝色和黄色）与盒子B中可以带来45美元收益的情况（即红色和绿色）相融合得到的：

盒子A:	90% 白色	6% 红色	1% 绿色	3% 黄色 / 蓝色
	0 美元	赢 45 美元	赢 30 美元	输 15 美元
盒子B:	90% 白色	7% 红色 / 绿色	1% 蓝色	2% 黄色
	0 美元	赢 45 美元	赢 45 美元	输 15 美元

与次可加性一致，产生相同结果的事件组合使盒子A更有吸引力，因为它将两个损失组合成了一个，而盒子B不再那么有吸引力，因为它将两个收益组合成了一个。事实上，在版本2中，58%的被试选择了盒子A，尽管它不如盒子B。C.斯塔摩（C. Starmer）和萨格登进一步研究了已知概率的拆分事件的影响（他们称其为事件的分裂效应，event-splitting effect），而且他们发现当一个积极的结果被拆分为两个组成部

分时，这一前景通常都会变得更加有吸引力（Starmer & Sugden, 1993）。这些结果表明，即使在明确了说明概率的情况下，拆分也会影响决策。

雷德尔迈耶等人进一步说明了拆分在选择中的作用（Redelmeier et al., 1995）。多伦多大学医学专业的毕业生（ $N=149$ ）面对一个医学情景：一名中年男子患有严重的呼吸短促。他们给一半的被试提供了详细的描述，指出：“显然，许多诊断是可能的……包括肺炎。”另外一半被试则得到了一份被拆分的描述，其中提到了肺炎以外的其他可能的诊断，如肺栓塞、心力衰竭、哮喘和肺癌。被试被问及是否会在这种情况下开抗生素（这种治疗方法对肺炎有效，但对拆分描述中提到的其他诊断无效）。他们预测拆分操作会降低肺炎的感知概率，因此被试开抗生素的倾向也会降低。事实上，在拆分描述的条件下，绝大多数的被试（64%）选择了不开具抗生素处方，而未拆分组被试选择开具抗生素处方和不开具抗生素处方的比例几乎相等（47%）。把肺炎单独列出来，增加了选择对肺炎有效的治疗方法的倾向，尽管描述中的症状明显符合一些众所周知的其他诊断。显然，拆分不仅会影响概率评估，还会影响决策。

尽管拆分在概率判断中起着重要的作用，但其背后的认知机制却更为普遍。因此，即使在不涉及不确定性事件的任务中，人们也会期望拆分效果。例如，范德·普利格特（van der Pligt）、J. R.艾泽（J. R. Eiser）和R.斯皮尔斯（R. Spears）让被试评估了5种能源（核电、煤炭、石油、水电、太阳能 / 风能 / 波）当前的占比和他们认为的理想分配，结果发现某种能源被单独评估时，会比它与其他能源一起被评估时获得更多的比例分配（van der Pligt, Eiser & Spears, 1987; Fiedler & Armbruster, 1994; Pelham, Sumarta & Myaskovsky, 1994）。这些结果都表明，拆分效应反映了人类判断的普遍特征。

外延

我们提出了一个有关信念的非外延性理论，在该理论中判断概率是根据对各个焦点假设和替代性假设的相对支持度或证据强度得出的。在该理论中，互斥假设的显性析取的支持度是可加性的，而隐性析取的支持度是次可加性的。实证研究证实了支持该理论的主要预测：第一，通过拆分焦点假设，概率判断增大了；通过拆分替代性假设，概率判断降低了；第二，主观概率在二元情况下是互补的，在一般情况下是次可加

性的；第三，相对于频率判断，概率的次可加性更明显，并且会通过相容的证据得到增强。支持理论还为从证据强度的独立评估中预测判断概率提供了一种方法。因此，它以单一的结构解释了广泛的实证结果。

在本节中，我们将探讨支持理论的一些外延和内含。首先，我们考虑这一理论的序数版本，并引入一个简单的参数表示法。其次，我们通过支持度的上下限来描述概率的上下限，以解决模糊或不精确的问题。最后，我们讨论了目前的研究发现了针对决策分析和知识工程的启发式程序设计的意义。

序数分析

在这一章中，我们把概率判断视为可表征信念程度的量化尺度。这一做法通常被解释为参考机会过程（reference chance process）。例如，假定某一候选人当选为总统的概率为 $2/3$ ，这就意味着评判者会认为，这一假设就像从一个装有 $2/3$ 的球为红球的罐子中取出一个红球的事件一样。因此，概率判断可以被看作是一个思维实验的结果，在这个实验中评判者将信念的程度与一个标准的机会过程相匹配（Shafer & Tversky, 1985）。当然，这种解释既不能保证一致性，也不能保证校准。

虽然概率判断似乎传递了量化信息，但将这些判断作为序数而不是基数来进行分析可能具有建设性意义。这一理解引出了支持理论的序数概括。假设对于 H 中的所有 A 和 B ， H 有一个非负尺度 s 和一个严格递增的函数

$$P(A, B) = F\left(\frac{s(A)}{s(A) + s(B)}\right) \quad (11)$$

其中，当 A 和 B 为互斥时， $s(C) \leq s(A \vee B) = s(A) + s(B)$ ； C 是隐性的，且 $C' = (A \vee B)'$ 。

序数模型的公理化超出了本文的范围。但是值得注意的是，为了在这种情况下获得一个本质上唯一的支持函数，我们必须做出额外的假设，比如下面的可解性条件：如果 $P(A, B) \geq z \geq P(A, D)$ ，那么存在 $C \in H$ ，使得 $P(A, C) = z$ （Debreu, 1958）。这种理想化模式在随机的情况下是可以接受的，如具有可连续调整扇形区的机会轮。下述定理

表明，假设序数模型和可解性条件、二元互补和乘积规则会产生一种特别简单的参数形式，这种形式与上文中用于关联评估和产生的支持模型相一致。证据见附录。

定理 2：假定序数模型（公式 11）和可解性条件。当且仅当存在一个常数 $k \geq 0$ 时，二元互补（公式 3）和乘积法则（公式 5）成立，则

$$P(A, B) = \frac{s(A)^k}{s(A)^k + s(B)^k} \quad (12)$$

如果 $k=1$ ，这种所谓的幂模型可以简化为基本模型。在该模型中，判断概率可能比各自的相对支持度更极端，也可能相反，这取决于 k 是大于还是小于 1。回想一下上文中的实验数据，它们为不等式 $\alpha < \delta$ 提供了强有力的证据。也就是说，只要 A_1 、 A_2 和 B 是互斥的，则 $P(A, B) \leq P(A_1, B) + P(A_2, B)$ ； A 是隐性的；而且 $A' = (A_1 \vee A_2)'$ 。我们还发现了等式 $\beta = \gamma$ 的证据（如表 13-2 所示），即 $P(A_1 \vee A_2, B) = P(A_1, A_2 \vee B) + P(A_2, A_1 \vee B)$ ，但是这个特性还没有经过广泛的验证。例如，由向 0.5 水平回归而产生的可加性的偏离可以由 $k < 1$ 的幂模型来表征，这意味着 $\alpha < \beta < \gamma < \delta$ 。请注意，对于互斥假设的显性析取，基本模型（公式 1 和公式 2）、序数模型（公式 11）和幂模型（公式 12）均假设了可加性支持，但只有基本模型才包含可加性概率。

上下指标

概率判断通常是模糊的和不够精确的。因此，要解释和正确使用这些判断，我们需要了解它们的不确定性范围。事实上，很多关于非标准概率的工作都与提供指标上、下限的置信程度的模型有关。然而，对这些指标的提取和解释，存在着理论和实践上的问题。如果人们很难评估某一事件的概率的确定值，那么他们可能更难评估两个确定值的上、下限，或者生成一个概率的二阶分布。评判者也许就能够就其评估的模糊性提供一些指示，但我们认为这种判断最好用定性而不是定量的术语来进行解释。

为此，我们设计了一种启发式程序。在该程序中，对概率的上、下限的判断是通过口头定义的，而不是用数字定义的。我们将这个程序称

为“阶梯法”（staircase method，如图13-6所示）。评判者面对一个不确定性事件（如下一个NBA总冠军将由一支东部球队而不是一支西部球队获得），并被要求从5个类别中选出一个来检验每个概率值。不“明显过低”的最低值（0.45）和不“明显过高”的最高值（0.80）分别由 P_* 和 P^* 来表示，且可作为指标的上、下限。当然，涉及不同的数字类别、不同措辞和不同幅度的情况可能会产生不同的指标。我们假设类别的标记是围绕中间类别对称的。我们可以将阶梯法看作对二阶概率分布或模糊隶属函数的定性模拟。

可能性 (%) :	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
明显过高																		x	x	x	x
稍微过高																x	x				
适当															x						
稍微过低										x	x	x	x	x							
明显过低	x	x	x	x	x	x	x	x	x												

图13-6 用来引出较高和较低概率的阶梯法示意

我们对 P_* 和 P^* 进行支持函数建模，分别表示为 s_* 和 s^* 。我们将这些尺度解释为 s 的高估计值和低估计值，并假设对于任意 A , $s_*(A) \leq s(A) \leq s^*(A)$ 。此外，我们假设 P_* 和 P^* 可以表示为

$$P_*(A, B) = \frac{s_*(A)}{s_*(A) + s^*(B)}$$

和

$$P^*(A, B) = \frac{s^*(A)}{s^*(A) + s_*(B)}$$

根据该模型，指标的上下限是通过证对证据的有倾向性的解读得到的； $P^*(A, B)$ 可以被理解为一个偏好 A 而非 B 的概率判断，而 $P_*(A, B)$ 正好相反。这种偏向性的大小反映了与基本判断相关的模糊性，也反映了启发式程序的特点。然而，在给定的过程中，我们可以将区间 (P_*, P^*) 作为不精确性的比较指标。因此，我们可以得出结论，如果与第一次评估有关的区间被包含在与第二次评估相关的区间内，那么一个判断就没有另一个判断那么模糊了。因为高估值和低估值不太可能比

评判者的最佳估计更加精确或可靠。我们认为 P_* 和 P^* 是 P 的补充，而非替代物。

为了检验标准理论所提出的上下限表示法，我们调查了人们对1992—1993年NFL季后赛结果的预测（Dempster, 1967; Good, 1962）。这项研究是在两场冠军赛开始的前一周进行的，在这两场比赛中，水牛城队将与迈阿密队争夺美国足球协会（AFC）的冠军，达拉斯队将与圣弗朗西斯科队争夺全美足球协会（NFC）的冠军。这两场比赛的获胜者将在两周后的超级碗比赛中互相比量。被试是135名斯坦福大学的学生，他们自愿参与了一项关于橄榄预测的研究以换取一张加州彩票。一半的被试评估了超级碗的赢家来自迈阿密和水牛城队这两支球队的概率。另一半被试评估了超级碗冠军将在达拉斯队和圣弗朗西斯科队之间产生的概率。所有被试都评估了两场比赛的概率。这些比赛的焦点假设和替代性假设的呈现顺序是平衡的。因此，每名被试使用图13-6所示的阶梯法进行了5次概率评估。

被试的最佳估计值呈现出了先前研究中观察到的次可加性和二元互补模式。4支球队赢得超级碗的平均概率之和为1.71；AFC和NFC的拆分因子分别为1.92和1.48。相对应的，每个事件和它的补集的平均概率之和为1.03。关于上、下限评估的分析，请注意本模型暗示的 $P_*(A, B) + P^*(B, A) = 1$ ，符合上、下限评估的标准理论。数据表明，该条件具有非常接近的近似值，平均和为1.02。

然而，目前的模型通常不符合上、下限概率的标准理论。为了说明这种差异性，假设 A 和 B 是互斥的， $C' = (A \vee B)'$ ，标准理论要求

$P_*(A, \bar{A}) + P_*(B, \bar{B}) \leq P_*(C, \bar{C})$ ，而之前的方法表明，当 C 为隐性时，反向不等式成立。这些数据显然违反了标准理论：迈阿密队和水牛城队赢得超级碗的平均较低概率分别为0.21和0.21，但它们之间的隐性析取（如一个AFC球队）仅为0.24。同样，达拉斯队和圣弗朗西斯科队赢得超级碗的平均概率也较低，分别为0.25和0.41，但一个NFC球队获胜的概率只有0.45。假设 s_* 具有次可加性，这些数据和目前的模型相符，但不符合低概率的标准理论。

规定性的影响

主观概率或信念程度模型有两个功能：描述性和规定性。关于非标

准概率模型的文献主要是规定性的。这些模型被视为评估证据和表达信念的正式语言。与之相反，支持理论试图描述人们做出概率判断的方式，而不是规定人们应该如何做出这些判断。例如，判断概率的论述通过拆分焦点假设而得到增强，通过拆分替代性假设而被削弱，这代表了一个不被规范性理论（加性或非加性）认可的一般性描述原则。

尽管支持理论具有描述性功能，但它也具有规定性的意义。它可以帮助设计启发式的程序并协调不一致的评估（Lindley, Tversky & Brown, 1979）。我们可以用一个感性的类比来对这一作用进行说明。假设一个测量员必须根据一个容易犯错的观察者对地标之间的距离的判断来绘制一个公园的地图。了解观察者可能存在的偏见可以帮助测量员绘制出更准确的地图。例如，由于观察者通常低估了隐藏区域的距离，测量员可能会放弃采用这些评估，使用平面几何定律来计算和评估它们的距离。此外，测量员可能希望通过应用适当的校正因子来对隐藏区域的估计值进行校准以减少偏差。同样的逻辑也适用于概率的引出。相关证据表明，人们往往低估了隐性析取的概率，尤其是对基本假设的否定。这种偏差可以通过要求评判者对比具有可比性的假设来得以减少，而不是评估特定假设相对于其补充假设的概率。

本研究的主要结论是，主观概率或信念程度在其他的空间分区中会带来不同的判断，在这个意义上它是非外延性的，是不可测量的。就像海岸线的长度测量会随着地图的详细程度的增加而增加，随着描述变得更具体，事件的感知概率也会增加。这并不意味着判断概率是没有价值的，它只是表明了这一概念比现有的形式理论所认为的更脆弱。本文所演示的外延性的失败说明了什么才是概率评估的基本问题，那就是需要考虑不可得的可能性。在需要生成新假设或构建新情景的任务中，这个问题尤其严重。我们认为，外延性原则在正常情况下是无懈可击的，但实际上它是无法实现的，因为我们不能指望评判者能完全拆分开任何隐性的析取。我们可以鼓励人们将项目拆分为它的组成部分，但不能期望他们考虑所有相关合取事件的拆分或生成所有相关的情景。在这方面，可加性概率分布的评估可能是一项不可能完成的任务。当然，评判者可以确保任何给定判断集的可加性，但这不能确保通过进一步细化后仍可以保持其可加性。

本研究和其他研究报告的证据表明，对不确定性的定性和定量评估都无法以逻辑上一致的方式进行，人们可能会得出这样的结论，即根本不应该进行这些评估。但是，这不是一个可行的办法，因为一般来说，

没有其他办法来评估不确定性。在测量距离的过程中，人类容易犯错误的判断方法可以被适当的物理测量所取代，而与此不同的是，没有客观的程序来评估事件发生的概率，例如被告的罪行、企业的成功或战争的爆发。因此，对不确定性的直觉判断必然在人们的思考和决策中发挥着重要作用。如何通过设计有效的启发式和矫正程序来提高思考和决策的质量，对于理论工作者和实践者来说都是一个重大的挑战。

附录

定理1: 假设 $P(A, B)$ 被定义为所有不相交的 $A, B \in H$ ，当且仅当 $A' = \emptyset$ 时会消失。当且仅当存在 H 满足公式1和公式2的非负比例尺度 s 时公式3至公式6成立。

证明：在当下，为了建立充分性，我们将 s 定义如下。设 $E = \{A \in H: A' \in T\}$ 为一组基本假设。选择一些 $D \in E$ 并设定集合 $s(D) = 1$ 。其他基本假设 $C \in E$ ， $C' \neq D'$ ，定义 $s(C) = P(C, D) / P(D, C)$ 。对于任何假设 $A \in H$ 都有 $A' \neq T$ ， $\emptyset$ ，选择一些 $C \in E$ 使得 $A' \cap C' = \emptyset$ ， $s(A)$ 定义如下

$$\frac{s(A)}{s(C)} = \frac{P(A, C)}{P(C, A)}$$

也就是说

$$s(A) = \frac{P(A, C)P(C, D)}{P(C, A)P(D, C)}$$

为了证明 $s(A)$ 是被单独定义的，假设 $B \in E$ 且 $A' \cap B' = \emptyset$ 。我们想证明这一点

$$\frac{P(A, C)P(C, D)}{P(C, A)P(D, C)} = \frac{P(A, B)P(B, D)}{P(B, A)P(D, B)}$$

根据比例关系（公式4），左边的比值等于：

$$\frac{P(A, C \vee B)P(C, D \vee B)}{P(C, A \vee B)P(D, C \vee B)}$$

右边的比值等于

$$\frac{P(A, B \vee C)P(B, D \vee C)}{P(B, A \vee C)P(D, B \vee C)}$$

消去一般项，很容易看出这两个比值是相等的，当且仅当：

$$\frac{P(C, D \vee B)}{P(B, D \vee C)} = \frac{P(C, A \vee B)}{P(B, A \vee C)}$$

这是成立的，因为这两个比值都等于 $P(C, B) / P(B, C)$ 。

为了完善 s 的定义，我们设当 $A' = \emptyset$ 时， $s(A) = 0$ 。对于 $A' = T$ ，我们区分了两种情况。如果 A 是显性的，也就是说对于一些互斥的 $B, C \in H$ 来说， $A = B \vee C$ ，设 $s(A) = s(B) + s(C)$ 。如果 A 是隐性的，那么 $s(A)$ 是 s 在所有显性 T 描述中的最小值。

为了建立所需的表征，我们首先显示所有互斥的 $A, B \in H$ ，这样 $A', B' \neq T, \emptyset$ ， $s(A) / s(B) = P(A, B) / P(B, A)$ 。回想一下， T 至少包含两个元素。我们必须考虑两种情况。

首先，假设一个 $A' \cup B' \neq T$ ；那么，存在一个基本假设 C 使得 $A' \cap C' = B' \cap C' = \emptyset$ 。在这种情况下

$$\frac{s(A)}{s(B)} = \frac{P(A, C) / P(C, A)}{P(B, C) / P(C, B)} = \frac{P(A, C \vee B) / P(C, A \vee B)}{P(B, C \vee A) / P(C, B \vee A)} = \frac{P(A, B)}{P(B, A)}$$

通过反复应用互补性可得。

其次，假设一个 $A' \cup B' = T$ 。在这种情况下，不存在 $C' \in T$ ，也就是不包含在 A' 或 B' 中，因此不能应用前面的参数。为了证明 $s(A) / s(B) = P(A, B) / P(B, A)$ ，假设 $C, D \in E$ 且 $A' \cap C' = B' \cap D' = \emptyset$ 。

因此

$$\begin{aligned}
\frac{s(A)}{s(B)} &= \frac{s(A)s(C)s(D)}{s(C)s(D)s(B)} \\
&= \frac{P(A,C)P(C,D)P(D,B)}{P(C,A)P(D,C)P(B,D)} \\
&= R(A,C)R(C,D)R(D,B) \\
&= R(A,B) && \text{(公式 5)} \\
&= P(A,B) / P(B,A) && \text{(根据需要)}
\end{aligned}$$

因此，对于任意一对互斥的假设，我们通过二元互补得到 $P(A, B) / P(B, A) = s(A) / s(B)$ ，且 $P(A, B) + P(B, A) = 1$ 。因此， $P(A, B) = s(A) / [s(A) + s(B)]$ 且 s 是唯一可供选择的单位，这是由 $s(D)$ 的值决定的。

为了设定 s 的特性，请回顾拆分（公式6）得到 $P(D, C) \leq P(A \vee B, C) = P(A, B \vee C) + P(B, A \vee C)$ ，当 $D' = A' \cup B'$ ，且 A 和 B 互斥时， D 是隐性的。左边的不等式意味着

$$\frac{s(D)}{s(D) + s(C)} \leq \frac{s(A \vee B)}{s(A \vee B) + s(C)}$$

因此， $s(D) \leq s(A \vee B)$ 。右边的等式意味着

$$\frac{s(A \vee B)}{s(A \vee B) + s(C)} = \frac{s(A)}{s(A) + s(B \vee C)} + \frac{s(B)}{s(B) + s(A \vee C)}$$

为了证明 P 的可加性意味着 s 的可加性，假设 A 、 B 和 C 非零且互斥（如果 $A' \cup B' = T$ ，那么结果是直接的）。因此，通过互补可得

$$\frac{s(A)}{s(B)} = \frac{P(A, B)}{P(B, A)} = \frac{P(A, B \vee C)}{P(B, A \vee C)} = \frac{s(A) / [s(A) + s(B \vee C)]}{s(B) / [s(B) + s(A \vee C)]}$$

因此， $s(A) + s(B \vee C) = s(B) + s(A \vee C) = s(C) + s(A \vee B)$ 。将这些关系代入由 P 的可加性暗示的公式中，可得到 $s(A \vee B) = s(A) + s(B)$ ，这意味着完成了定理1的证明。

定理2：假设序数模型（公式11）和可解性条件。如果存在一个常数 $k \geq 0$ ，那么二元互补（公式3）和乘积法则（公式5）成立，使得

$$P(A, B) = \frac{s(A)^k}{s(A)^k + s(B)^k}$$

证明：公式3和公式5由幂模型（公式12）表示，这很容易验证。为了得到这种表示，假设序数模型和可解性条件都满足。存在一个定义在 H 上的非负尺度 s ，以及一个从单位区间到自身的严格递增的函数 F ，对于所有 $A, B \in H$

$$P(A, B) = F \left[\frac{s(A)}{s(A) + s(B)} \right]$$

通过二元互补， $P(A, B) = 1 - P(B, A)$ ；因此， $F(z) = 1 - F(1 - z)$ ， $0 \leq z \leq 1$ 。函数 G 由下式定义

$$R(A, B) = \frac{P(A, B)}{P(B, A)} = \frac{F\{s(A)/[s(A) + s(B)]\}}{F\{s(B)/[s(B) + s(A)]\}} = G[s(A)/s(B)], B' \neq \emptyset$$

应用乘法法则， $s(C) = s(D)$ ，得到 $G[s(A)/s(B)] = G[s(A)/s(C)] G[s(C)/s(B)]$ ；因此， $G(xy) = G(x) G(y)$ ， $x, y \geq 0$ 。这是柯西公式的一种形式，它的解是 $G(x) = x^k$ （Aczel, 1966）。因此， $R(A, B) = s(A)^k / s(B)^k$ ，通过二元互补

$$P(A, B) = \frac{s(A)^k}{s(A)^k + s(B)^k}, \quad k \geq 0 \quad (\text{根据需要})$$

致谢

我们感谢玛雅·巴尔·希勒尔（Maya Bar-Hillel）、托德·戴维斯（Todd Davies）、丹尼尔·卡尼曼、大卫·克兰茨、格伦·沙费尔、艾尔达·沙菲尔和彼得·韦克尔给予的帮助。

参考文献

Aczel, J. (1966). *Lectures on functional equations and their applications*. San Diego, CA: Academic Press.

Bar-Hillel, M., & Neter, E. (1993). How alike is it versus how likely is it: A disjunction fallacy in stereotype judgments. *Journal of Personality and Social Psychology*, 65, 1119–1131.

Briggs, L. K., & Krantz, D. H. (1992). Judging the strength of designated evidence. *Journal of Behavioral Decision Making*, 5, 77–106.

Debreu, G. (1958). Stochastic choice and cardinal utility. *Econometrica*, 26, 440–444.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38, 325–339.

Dube-Rioux, L., & Russo, J. E. (1988). An availability bias in professional judgment. *Journal of Behavioral Decision Making*, 1, 223–237.

Dubois, D., & Prade, H. (1988). Modelling uncertainty and inductive inference: A survey of recent nonadditive probability systems. *Acta Psychologica*, 68, 53–78.

Erev, I., Wallsten, T. S., & Budescu, D. V. (1994). Simultaneous over- and underconfidence: The role of error in judgment processes. *Psychological Review*, 101, 519–527.

Fiedler, K., & Armbruster, T. (1994). Two halves may be more than one whole. *Journal of Personality and Social Psychology*, 66, 633–645.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 330–344.

Fox, C. R., Rogers, B., & Tversky, A. (1994). Decision weights for options traders. Unpublished manuscript, Stanford University, Stanford, CA.

Gettys, C. F., Mehle, T., & Fisher, S. (1986). Plausibility assessments in

hypothesis generation. *Organizational Behavior and Human Decision Processes*, 37, 14–33.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond “heuristics and biases.” In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 2, pp. 83–115). New York, NY: Wiley.

Gilboa, I., & Schmeidler, D. (1994). Additive representations of nonadditive measures and the Choquet integral. *Annals of Operations Research*, 52, 43–65.

Good, I. J. (1962). Subjective probability as the measure of a nonmeasurable set. In E. Nagel, P. Suppes, & A. Tarski (Eds.), *Logic, methodology, and philosophy of sciences* (pp. 319–329). Stanford, CA: Stanford University Press.

Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, 7, 35–51.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.

Kahneman, D., & Tversky, A. (1979). Intuitive prediction: Biases and corrective procedures. *TIMS Studies in Management Science*, 12, 313–327.

Kahneman, D., & Tversky, A. (1982). Variants of uncertainty. *Cognition*, 11, 143–157.

Koehler, D. J. (1994). Hypothesis generation and confidence in judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 461–469.

Koehler, D. J., & Tversky, A. (1993). The enhancement effect in probability judgment. Unpublished manuscript, Stanford University,

Stanford, CA.

Krantz, D. H., Ray, B., & Briggs, L. K. (1990). Foundations of the theory of evidence: The role of sche-mata. Unpublished manuscript, Columbia University, New York, NY.

Lichtenstein, S., Fischhoff, B., & Phillips, L. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). Cambridge, England: Cambridge University Press.

Lindley, D. V., Tversky, A., & Brown, R. V. (1979). On the reconciliation of probability assessments. *Journal of the Royal Statistical Society. Series A (General)*, 142, 146–180.

Mehle, T., Gettys, C. F., Manning, C., Baca, S., & Fisher, S. (1981). The availability explanation of excessive plausibility assessment. *Acta Psychologica*, 49, 127–140.

Mongin, P. (1994). Some connections between epistemic logic and the theory of nonadditive probability. In P. W. Humphreys (Ed.), *Patrick Suppes: Scientific philosopher* (pp. 135–172). Dordrecht, the Netherlands: Kluwer.

Murphy, A. H. (1985). Probabilistic weather forecasting. In A. H. Murphy & R. W. Katz(Eds.), *Probability, statistics, and decision making in the atmospheric sciences*(pp. 337–377). Boulder, CO: Westview Press.

Olson, C. L. (1976). Some apparent violations of the representativeness heuristic in human judgment. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 599–608.

Pelham, B. W., Sumarta, T. T., & Myaskovsky, L. (1994). The easy path from many to much: The numerosity heuristic. *Cognitive Psychology*, 26, 103–133.

Peterson, D. K., & Pitz, G. F. (1988). Confidence, uncertainty, and the use of information. *Journal of Experimental Psychology: Learning*,

Memory, and Cognition, 14, 85–92.

Redelmeier, D., Koehler, D. J., Liberman, V., & Tversky, A. (1995). Probability judgment in medicine: Discounting unspecified alternatives. *Medical Decision Making*, 15, 227–230.

Reeves, T., & Lockhart, R. S. (1993). Distributional Vs. singular approaches to probability and errors in probabilistic reasoning. *Journal of Experimental Psychology. General*, 122, 207–226.

Robinson, L. B., & Hastie, R. (1985). Revision of beliefs when a hypothesis is eliminated from consideration. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 443–456.

Russo, J. E., & Kolzow, K. J. (1992). Where is the fault in fault trees? Unpublished manuscript, Cornell University, Ithaca, NY.

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton, NJ: Princeton University Press.

Shafer, G., & Tversky, A. (1985). Languages and designs for probability judgment. *Cognitive Science*, 9, 309–339.

Starmer, C., & Sugden, R. (1993). Testing for juxtaposition and event-splitting effects. *Journal of Risk and Uncertainty*, 6, 235–254.

Statistical abstract of the United States. (1990). Washington, DC: U. S. Department of Commerce, Bureau of the Census.

Suppes, P. (1974). The measurement of belief. *Journal of the Royal Statistical Society. Series B. Methodological*, 36, 160–191.

Teigen, K. H. (1974a). Overestimation of subjective probabilities. *Scandinavian Journal of Psychology*, 15, 56–62.

Teigen, K. H. (1974b). Subjective sampling distributions and the additivity of estimates. *Scandinavian Journal of Psychology*, 15, 50–55.

Teigen, K. H. (1983). Studies in subjective probability III: The

unimportance of alternatives. *Scandinavian Journal of Psychology*, 24, 97–105.

Tversky, A., & Fox, C. (1994). Weighing risk and uncertainty. Unpublished manuscript, Stanford University, Stanford, CA.

Tversky, A., & Kahneman, D. (1983). Extensional Vs. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 91, 293–315.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions, Part 2. *Journal of Business*, 59, 251–278.

Tversky, A., & Sattath, S. (1979). Preference trees. *Psychological Review*, 86, 542–573.

van der Pligt, J., Eiser, J. R., & Spears, R. (1987). Comparative judgments and preferences: The influence of the number of response alternatives. *British Journal of Social Psychology*, 26, 269–280.

van Wallendaal, L. R., & Hastie, R. (1990). Tracing the footsteps of Sherlock Holmes: Cognitive representations of hypothesis testing. *Memory & Cognition*, 18, 240–250.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. London, England: Chapman & Hall.

Wallsten, T. S., Budescu, D. V., & Zwick, R. (1992). Comparing the calibration and coherence of numerical and verbal probability judgments. *Management Science*, 39, 176–190.

Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1, 3–28.

Zarnowitz, V. (1985). Rational expectations and macroeconomic forecasts. *Journal of Business & Economic Statistics*, 3, 293–311.

第14章 基于推理的选择

埃尔德·沙菲尔
伊塔马尔·西蒙森
阿莫斯·特沃斯基

其结果是一种被称为犹豫不决的内在不安感。幸运的是，它太熟悉了，不需要对其进行描述，因为描述它是不可能的。只要它持续存在，在注意之前与各种对象一起，我们就可以说是深思熟虑；当最初的建议最终被采纳并得以实施，或者被它的对手彻底扼杀时，我们就可以决定……赞成其中一方或另一方。与此同时，对想法的强化和抑制被称为决定产生的原因或动机。

——威廉·詹姆斯

我的方法是把半张纸用一条线分成两栏；然后，在三四天的考虑中，我在不同的标题下写下不同动机的简短提示，这些提示是在不同的时间想到的支持或反对的想法。当我把它们都放在一起时，我就会努力估计它们各自的权重……最后找到平衡点……而且，虽然理性的权重不能以代数的精确性来衡量，但是当把每一个想法分别加以考虑，或以比较的方式加以考虑时，整个问题就摆在我面前了。我想我可以做出更好的判断，不再那么容易鲁莽行事；事实上，我发现这类公式有很大的优势，即所谓的公正代数法（moral or prudential algebra），它有很大的用处。

——本杰明·富兰克林

引言

由于不确定性和冲突，无论大小决策通常都很困难。我们通常不确定自己的行为的确切后果，这可能取决于天气或经济状况，而且我们经常会在一个属性（如存钱）与另一个属性（如休息）之间的取舍上产生冲突。为了解释人们是如何解决这类冲突的，学习决策的学生要么使用传统的形式模型，要么使用基于推理的分析。在经济学、管理学和决策研究常用的形式化建模方法中，人们通常会将数值与每个选项关联起来，并将选择描述为价值的最大化。这种基于价值的方法包括规范性模型，如期望效用理论（von Neumann & Morgenstern, 1947），以及描述性模型，如前景理论（Kahneman & Tversky, 1979）。另一种研究决策的传统方法具有历史和法律方面的学术特点，它使用典型的政治和商业话语，采用非正式的基于推理的分析。这种方法确定了各种各样的理由和论据，这些理由和论据被认为可以进入和影响决策，并根据支持和反对各种选项的平衡来对选择进行解释。基于推理的分析可以在对美国历史决策的研究中找到例子，例如在古巴导弹危机期间所做的决策（Allison, 1971）、戴维营协议（Telhami, 1990）或越南战争期间所

做的决策（Berman, 1982; Betts & Gelb, 1979）。此外，基于推理的分析在商学院和法学院中通常被用来进行“案例研究”。尽管研究人员援引的理由可能并不总是与实际决策者的动机相对应，但人们普遍认为，根据推理进行的分析可能有助于解释决策，特别是在难以应用价值模型的情况下。

这两种传统之间几乎没有什么联系，它们通常被应用于不同的领域。基于推理的分析主要被用于解释非实验数据，特别是独特的历史、法律和政治决策。相反，基于价值的方法在偏好的实验研究和标准经济分析中发挥了核心作用。当然，这两种方法并不是不兼容的：基于推理的方法通常可以转换为正式的模型，而正式的分析通常可以被解释为基于推理的方法。在缺乏全面的决策理论的情况下，形式模型和基于推理的分析都有助于理解决策。

这两种方法都有明显的优点和局限性。形式化的、基于价值的模型具有严谨性的优点，这便于派生出可检验的应用。然而，基于价值的模型很难应用于复杂的、现实世界中的决策，而且它们往往无法捕捉到人们思考的重要方面。此外，对于基于推理进行选择的解释本质上是定性的，通常也是模糊的。而且，几乎任何事情都可以算作一种“原因”，这样每一个决定都可以在事后被合理化。为了克服这个困难，我们可以让人们说明自己做出决定的原因。可惜，指导决策的实际原因可能与被试报告的原因相一致，也可能不一致。正如有大量文献记载的那样，有时被试并不知道主导他们选择的确切因素，当被要求解释自己的决定时，他们会做出虚假的解释（Nisbett & Wilson, 1977）。事实上，对内省报告效度的怀疑已经导致许多学习决策的学生只关注观察到的选择。虽然口头报告和内省报告可以为我们提供有价值的信息，但我们在本文中会使用“推理”来描述影响决策的因素或动机，无论这些因素或动机能否被决策者阐明或识别。

尽管有其局限性，基于推理的决策概念仍有几个有吸引力的特点。首先，对原因的关注似乎更接近于我们通常思考和谈论选择的方式。当面临一个艰难的选择时（在学校之间或者工作之间做选择），我们总是试着找出支持和反对每一个选择的理由——我们通常不会尝试着去估计它们的总体价值。其次，选择以理性为指导为我们提供了一种自然的可用来理解决策过程中所特有的冲突的方式。从基于推理的选择的角度来看，当决策者有充分的理由支持和反对每一个选项时，或者有充分的理由支持相互竞争的选项时，就会产生冲突。与易于比较的数值不同，造

成冲突的原因可能很难调和。基于推理的分析也可以容纳框架效应（Tversky & Kahneman, 1986）和诱发效应（Tversky, Sattath & Slovic, 1988），这表明偏好很容易受到选项的描述方式的影响（如收益或损失），也很容易受到引发偏好的方法的影响（如定价和选择）。这些发现从价值最大化的角度来看是令人困惑的，如果我们假设不同的框架和启发式过程强调了选项的不同方面，从而引出了指导决策的不同原因，那么这些发现就更容易得到解释。最后，基于推理的选择概念可能包含比较性的考虑因素（如相对优势或预期后悔），这些因素通常不在价值最大化的范围内。

在本章中，我们探讨了基于推理进行选择的逻辑，并检验了一些有关推理在决策中的作用的具体假设。本章的内容如下：第1节探讨了推理在选择吸引力相同的选项时所起的作用。第2节探讨了选择和拒绝某选项的不同依据。第3节考察了强冲突和弱冲突之间的相互作用，以及人们寻求其他选择的倾向。第4节探讨了冲突与附加性选择集中的选项的关系。第5节对比了选择的具体原因与原因的析取的影响。第6节探讨了不相关的原因在决策过程中所起的作用。第7节为结束语。

在吸引力等同的选项之间做出选择

当决策者面临两个有相同吸引力的选项时，他们会如何解决这种冲突？为了研究这个问题，斯洛维克首先让被试将两个选项等同起来，然后让他们在每对同等价值的选项中做出选择（Slovic, 1975）。例如，一对由现金和优惠券组合而成的礼盒。如下表所示，每一对备选项都缺失一个组成部分，被试被要求填写表格中缺失的金额，使两个备选项具有同等的吸引力（在下面的例子中，被试自愿提供的值可能是10美元）。

	礼盒 A	礼盒 B
现金	—	20 美元
优惠券	32 美元	18 美元

一周后，被试被要求在两种相等的选项中做出选择。他们每个人都会被问到：哪个更重要？是现金还是优惠券？根据价值理论，这两个选项的价值明显是一样的，有同样的概率被选择。但事实相反，在上述礼盒的选择中，虽然88%的被试认为这两个选项的价值是相同的，但是他

们会选择在自己更重视的维度上价值更高的选项。

正如斯洛维克所指出的那样，人们似乎遵循着一种易于解释和证明的选择机制：根据更重要的维度进行选择，这比随机选择或选择右边的选项的选择理由更好（Slovic, 1975, 1990）。斯洛维克在许多领域复制了上述的结果，包括大学申请、汽车轮胎、棒球运动员和工作路径之间的选择（Slovic, 1975；关于更多的数据和引发程序的讨论，请参见特沃斯基等人进行的研究；Tversky, Sattath, Slovic, 1988）。所有的结果都是一致的——人们不会在相等的选项中进行随机选择。相反，他们通过选择在更重要维度上更优的选项来解决冲突，这似乎为选择提供了一个令人信服的理由。

赞成和反对的理由

考虑一下，我们必须从两个选项中选择一个，或者必须拒绝两个选项中的一个。在选择的标准分析下，这两个任务是可以互换的。在二元选择的情况下，人们被问到更喜欢哪个选项还是会拒绝哪个选项，应该是无关紧要的。因为如果人们更喜欢前者，他们就会拒绝后者，反之亦然。

正如富兰克林在本章的开篇所建议的那样，我们的决策一部分取决于分配给选项的优点和缺点的权重。我们认为，选项的正面特征（优点）将在选择过程中被放大，而选项的负面特征（缺点）的权重则在拒绝时更大。选择一个选项是因为它的正面特征，而拒绝一个选项是因为它的负面特征，这是很自然的。如果人们的决定是基于支持或反对选项的理由，那么在决定选择哪个选项时，他们可能会关注选择选项的理由，而在决定拒绝哪个选项时，他们可能就会关注拒绝某个选项的理由。这个假设引出了一个直接的预测：假设有两个选项，一个是更丰富的选项，具有更多的正面和负面特征，一个是更单薄的选项，具有更少的正面和负面特征。如果正面特征在选择时的权重大于拒绝时的权重，负面特征在拒绝时的权重大于选择时的权重，那么与负面特征相比，更丰富的选项可以同时被选择和拒绝。设 P_c 和 P_r 分别表示选择和拒绝某一选项的被试百分比。如果选择和拒绝是互补的，那么 $P_c + P_r$ 之和应该等于100。此外，根据上述假设，更丰富的选项的 $P_c + P_r$ 应该大于100，更单薄的选项的这一值则应低于100。沙菲尔观察到了这种模式（Shafir, 1993）。我们将下述问题以两个版本呈现给被试，只是括号内的问题有所不同。一半的被试看到的是一个版本，另一半被试看到另一个版本。

尽管呈现给被试的顺序是平衡的，但更丰富的选项在最后出现。

问题 1 ($N = 170$):

想象一下，在一场混乱的离婚案件之后，你在一个独生子女监护权案件中担任陪审员。由于在经济、社会和情感方面的模糊性，案件的事实变得复杂起来，你决定完全根据以下几个观察结果做出决定。(你会把孩子的监护权判给哪位家长？你会拒绝哪位家长获得对孩子的唯一监护权？)

	支持	反对
家长 A: 收入水平一般	36%	45%
健康水平一般		
工作时间为平均水平		
与孩子保持了良好的互动		
稳定的社会生活		
家长 B: 收入高于平均水平	64%	55%
与孩子的关系非常密切		
非常活跃的社会生活		
大量的出差		
较小的健康问题		

家长A是相对单薄的选项，没有显著的正面或负面的特征，也没有特别令人信服的理由可用来作为对孩子监护权的裁定依据。另一位家长B是相对丰富的选项，有充分的理由获得监护权（与孩子关系密切且收入丰厚），但也有充分的理由被拒绝授予单独监护权（健康问题和由于出差而大量缺席孩子的生活）。选项的右侧是选择判决同意和拒绝给予双方家长监护权的被试的百分比。家长B是大多数人的选择，既有大多数被试认为该家长应该获得孩子的监护权，也有55%的人认为应该剥夺其监护权。正如预测的那样，家长B的 $Pc+Pr$ ($64+55=119$) 显著大于100，选择和拒绝的前景值是互补的 ($z=2.48$, $p<0.02$)。这种模式可以通过观察结果进行解释，更丰富的选项（家长B）为决策者提供了更有说服力的理由来支持或反对其对孩子的监护权。

在货币赌博、大学课程和政治候选人的研究中，重现了上述结果（Shafir, 1993）。这里还有一个例子，考虑下面的问题，研究人员给

一半的被试呈现了“偏好”（prefer）选项，给另一半被试呈现了“取消”（cancel）选项。

问题 2 ($N = 172$):

偏好:

想象一下，你计划去一个温暖的地方度过一周的假期。目前你有两种价格合理的选择。旅游手册只提供了关于这两种选择的有限信息。根据现有的信息，你更喜欢哪个度假地点？

取消:

想象一下，你计划去一个温暖的地方度过一周的假期。你目前有两个价格合理的选项，但你不能继续保留关于两个选项的预订。旅游手册只提供了关于这两种选择的有限信息。根据现有信息，你决定取消哪项预订？

		偏好	取消
地点 A:	气候一般	33%	52%
	海滩一般		
	中等水平的酒店		
	中等水平的水温		
	夜生活一般		
地点 B:	阳光充足	67%	48%
	极美的沙滩和岩石		
	超现代的酒店		
	非常低的水温		
	非常大的风		
	没有夜生活		

这两个景点的信息是我们在决定下一个假期去哪里度假时可以得到的典型信息。由于很难估计每个景点的总价值，我们很可能会寻找做出决定的理由。地点A是相对单薄的选项，似乎并不引人注目，但在所有方面都是没有争议的。此外，选择地点B也有明显的理由：美丽的海滩、充足的阳光和超现代的酒店；当然，也有令人信服的拒绝它的理由：水冷、风大和缺乏夜生活。我们认为，美丽的海滩很可能在我们选

择的时候比我们拒绝的时候更有说服力，而缺乏夜生活可能在我们拒绝的时候比我们选择的时候扮演了更重要的角色。的确，地点B被优先选择和拒绝的比例超过了地点A（ $P_c + P_r = 67 + 48 = 115$ ， $p < 0.05$ ）。这些结果表明，选项并不是简单地按照价值排序的，选择的吸引力越大，拒绝的吸引力就越小。相反，选项的优缺点的相对重要性似乎因任务的性质而异。因此，当我们问自己更喜欢哪一个而不是考虑取消哪一个时，我们更有可能最终选择地点B（67%对52%， $z = 2.83$ ， $p < 0.001$ ）。

理性选择理论最基本的假设之一是过程不变原则，这需要策略上等价的方法来产生相同的偏好（Tversky et al., 1988）。选择-拒绝差异表明过程不变性的可预测性是无法满足的。这种现象与价值最大化是不一致的，但是从基于推理的选择的角度来看待问题就很容易理解：当我们选择的时候，选择的理由比我们拒绝的时候更有说服力；当我们拒绝的时候，拒绝的理由比我们在选择时更重要。

冲突中的选择：寻求选项

选择的需要常常会引发冲突：我们不确定如何权衡一个属性与另一个属性之间的关系，也不确定哪个属性对于我们来说最重要。我们经常试图通过为选择一方而不选择另一方寻找理由，以此来解决这种冲突，这是司空见惯的。有时，现有选项之间的冲突很难解决，可能会导致我们寻求更多的选项或是维持现状。在其他时候的情境是这样的，在选项之间进行的比较产生了选择一个选项而不是另一个的充分理由。使用推理来解决冲突存在着一些不明显的含义，下面我们将对这些含义进行讨论。本节集中讨论人们寻求其他选项的决定；在下一节中，我们将探讨向当前选项集中添加选项产生的一些效果。

在许多情况下，我们需要决定是选择当前的某个选项，还是再去搜索其他的选项。因此，一个想买二手车的人可能会满足于当下可以买到的汽车，也可能会继续寻找附加的选项。寻找新的选项通常需要额外的时间和精力，并且可能会有失去之前可得选项的风险。冲突在经典的决策理论中没有任何作用。在这个理论中，每个选项 x 都有一个值 $v(x)$ ，对于任何备选集合，决策者都会选择值最高的选项。特别是，只有当搜索附加选项的前景值超过当前可得的最佳选项的前景值时，人们才会去搜索其他选项。此外，对推理的依赖意味着，当我们有充分的理由做出选择时，应该更有可能选择一个可得的选项；当我们目前没有明确的做出选择的理由时，应该更有可能继续寻找其他的选项。

特沃斯基和沙菲尔验证了这一假说，他们给被试提供了几对选项，如不同概率和回报的投注或学生公寓不同的月租金标准和距离校园的远近；研究人员让被试从两个选项中选择一个，或者被试可以要求一个附加选项，但需要为此付出一些代价（Tversky & Shafir, 1992b）。被试首先会看到12个选项（投注或公寓），以便熟悉这些可得的选项。在研究投注之间的选择时，一些被试还会被要求回答以下问题。

冲突：

假设你要对以下两种情况做出投注选择：

(x) 有 65% 的机会赢 15 美元。

(y) 有 30% 的机会赢 35 美元。

你可以选择其中一种投注方式，或者可以通过支付 1 美元在当前的选项集中添加另外的一个选项。所添加的投注方式将从你查看过的列表中随机选取。

其他被试也会看到类似的问题，只是选项y被选项x'所替代，以便被试在以下选项中做出选择。

优势：

(x) 有 65% 的机会赢得 15 美元。

(x') 有 65% 的机会赢得 14 美元。

被试被要求表明他们是要增加另一种投注，还是在目前的选项集中进行选择。然后他们从选项集中选择自己喜欢的投注方式（添加或不添加选项）。被试被告知，他们选择的投注将会结束，其回报将与减去他们为附加选项支付的费用成正比。

研究人员在选择学生公寓的问题上，设计了一个类似的情境。一些被试被要求回答以下问题。

冲突：

假设你要在两套公寓中做出选择，这两套公寓的特点如下所述：

(x) 每月租金为 290 美元，距离学校 25 分钟车程。

(y) 每月租金为 350 美元，距离学校 7 分钟车程。

两套公寓都有一间卧室和一间小厨房。你现在可以在这两套公寓中进行选择，也可以继续搜索公寓（在你看过的清单中随机选择）。若继续搜索其他公寓，你可能会失去已经找到的一套或两套公寓。

其他的被试面对的问题与上述问题类似，只是选项y被选项x'所替代，以便他们在以下选项中做出选择。

优势：

(x) 每月租金为 290 美元，距离学校 25 分钟车程。

(x') 每月租金为 330 美元，距离学校 25 分钟车程。

注意，在这两对问题中，x和y之间的选择（冲突情境）并不是无关紧要的，因为x在一个维度上更好，而y在另一个维度上更好。相反，在x和x'之间的选择（优势情境）不涉及冲突，因为前者完全优于后者。因此，虽然在冲突情境中没有明显的理由选择一方而不选择另一方，但在优势情境中，有决定性的理由选择出其中的一个。

平均而言，在冲突情境下有64%的被试要求研究人员提供附加选项，而在优势条件下只有40%的被试要求附加选项（ $p < 0.05$ ）。换句话说，当可供选择的方案难以合理化时，被试更倾向于寻找附加选项，而不是当有一个令人信服的选择理由且决策相对容易时。

这些数据都不符合价值最大化原则。根据价值最大化原则，当且仅当附加选项的（主观）预期价值超过当前可用的最佳选择时，主体才应该寻找附加选项。因为在优势情境下给出的最佳选项在冲突情境下也是可用的，价值最大化意味着在冲突情境中寻求附加选项的被试百分比不应该比优势情境中的更大，而这与观察到的数据相反。

看来，寻找其他选项不仅取决于价值最大化所指的可用的最佳选项

的价值，而且也取决于在当下选项集中做出选择的困难程度。例如，在优势情境中，有明确且无可争辩的理由选择一个而不是另一个选项（如“这套公寓一样远，但我省了40美元！”）。一个令人信服的选择其中一个选项而不是其他选项的理由，可以降低寻找其他选项的诱惑。此外，当选择涉及冲突时，选择其中任何一个选项的理由都不那么充分，而且这个决定的合理性也更难以得到证明（如“我应该每月节省60美元，还是应该在离校园更近的地方居住？”）。在缺乏令人信服的选择理由的情况下，人们更倾向于寻找其他的选择。

冲突下的选择：添加选项

对原因的分析可以帮助我们解释所观察到的对无关选项独立原则的违反，根据这一原则，两个选项之间的优先顺序不应因引入其他选项而改变。这一原则遵循价值最大化的标准假设，并在分析消费者选择时经常被假设。尽管它具有直观的吸引力，但越来越多的证据表明，人们的偏好取决于选择的情境，即考虑当下的选项集。特别是，从给定的集合中添加和删除选项会影响人们对一直处于选项集当中的选项的偏好。在上一节中，我们考虑了人们在给定的一组选项的情境中寻求附加选项的倾向，而在这一节中，我们将说明因为添加选项而出现的现象，并从选择的理由的角度来解释它们。

价值最大化的一个可以验证的主要意义就是，一个原本不被喜欢的选项不会在加入新的选项后变成被喜欢的选项。具体来说，一个倾向于选项y而不是倾向于推迟选择的决策者，不会在y和x都可以选择的情况下倾向于推迟选择。选项的“市场份额”不能通过扩大给定的选项集来增加，这就是所谓的正则条件（**regularity condition**；Tversky & Simonson, 1993）。与正则条件相反，大量的实验结果表明，随着选项的增加，人们推迟选择的倾向会增强。例如，考虑一下人们面对两个相互竞争的选项和面对一个有吸引力的选项（他更愿意推迟选择）的冲突程度。从两种相互竞争的选项中选择一个可能是困难的：一个选项具有吸引力这一事实本身并不能为选择提供充分的理由，因为另一个选项可能同样具有吸引力。因此，增加一种选择可能会使决策更难被证明是正确的，并会增加人们推迟决策的倾向。

托马斯·谢林恰当地描述了一个相关现象。在书店里，他看到了两本很有吸引力的百科全书，他发现自己很难在这两本书之间做出选择，最后一本也没买。如果只有一本百科全书，他就会买。更普遍地说，在

某些情况下，人们更喜欢可用的替代性方案而不是现状，但没有令人信服的可以让其在这些选项中做出选择的理由，因此他们可能会无限期地推迟决定。

特沃斯基和沙菲尔在接下来的两组问题中证明了谢林所描述的现象，这两组问题分别由两组学生完成（*N*分别为124和121；Tversky & Shafir, 1992b）。

强冲突：

假设你正在考虑购买一个唱片播放器，但还没有决定要买什么型号。你路过一家正在进行为期 1 天的清仓大甩卖的商店，它们分别以 99 美元和 169 美元的价格推出了广受欢迎的索尼（SONY）播放器和顶级的爱华（AIWA）播放器，这两款播放器的售价都远低于标价。你会：

- | | |
|--------------------|-----|
| (x) 购买爱华播放器？ | 27% |
| (y) 购买索尼播放器？ | 27% |
| (z) 等了解了各种产品之后再做选择 | 46% |

弱冲突：

假设你正在考虑购买一台唱片播放器，但还没有决定要买什么型号。你路过一家正在进行为期 1 天的清仓大甩卖的商店，它们向你提供了一款受欢迎的索尼播放器，售价仅为 99 美元，远低于标价。你会：

- | | |
|--------------------|-----|
| (y) 购买索尼播放器 | 66% |
| (z) 等了解了各种产品之后再做选择 | 34% |

结果表明，与强冲突的情况相比，人们在弱冲突的情况下更倾向于购买唱片播放器（ $p<0.05$ ）。两款播放器看起来都很有吸引力，价格都很好，而且都在打折。决策者需要决定她是买一个更便宜、更受欢迎的型号，还是买一个更贵、更复杂的型号。这种冲突显然不容易得到解决，它会迫使许多被试推迟购买，直到他们了解了更多的其他选项。此外，当索尼播放器被单独给出时，就有了足够的理由购买：它是一款受欢迎的播放器，价格非常好，而且只销售一天。在这种情况下，决策者有充分的理由选择给定的选项，大多数被试决定选择购买唱片播放器而不是延迟。

在前面的例子中增加了一个竞争的备选方案后，推迟决策的倾向得到了增强。显然，冲突的程度及其得到解决的难易程度不仅取决于现有选项的数目，而且取决于对这些选项进行比较的方式。例如在下面的问题中，原来的爱华播放器被一个较差的型号所替代（ $N=62$ ）。

优势：

假设你正在考虑购买一台唱片播放器，但还没有决定要买什么型号。你路过一家正在进行为期 1 天的清仓大甩卖的商店，它们提供了一款受欢迎的索尼播放器，售价仅为 99 美元，远低于上市价格；还有一款相对较差的爱华播放器，售价仅为 105 美元。你会：

(x') 购买爱华播放器	3%
(y) 购买索尼播放器	73%
(z) 等了解了各种产品之后再作选择	24%

在这个版本中，与之前的强冲突版本相反，爱华播放器明显比索尼播放器差：它的质量较差，价格更高。因此，爱华的出现并没有减少购买索尼播放器的理由，它实际上是对这些理由的补充：索尼播放器的价格很好，只销售 1 天，而且明显优于竞争对手。因此，选择索尼播放器的比例比加入更差型号的爱华播放器前更大一些。当一个优势不对称或者相对较差的选项被添加到一个集合中时，优势选项的吸引力和选择概率得到增加的现象被称为“不对称优势效应”（Huber, Payne & Puto, 1982）。注意，在强冲突和优势情境的问题中，被试都有两个唱片播放器和一个延迟选择的选项。然而，当被试缺乏购买任何一款播放器的明确理由时，延迟购买的倾向要比他们有充分理由只购买一款而不购买另一款时大得多（ $p<0.005$ ）。

上面的模式违反了正则条件，只要添加的选项没有附带新的和相关的信息，正则条件就会成立。在上面的场景中，我们可以认为添加的选项（一种情况下是优势项，另一种情况下是劣势项）传达了消费者在寻找更好的交易机会的信息。回想一下，关于信息方面的考虑并不能解释上一节中的搜索实验，因为其中的被试已经看到了所有可能的选项。然而，为了进一步检验这一解释，特沃斯基和沙菲尔设计了一个类似的问题，它涉及实际收益，并且没有延迟选项（Tversky & Shafir, 1992b）。被试（ $N=80$ ）填写一份简短的问卷可获得 1.5 美元的报酬。在

问卷调查完成后，一半的被试有机会将1.5美元（默认值）的报酬兑换为两种奖品中的一种：金属斑马笔（以下简称“斑马笔”），或者塑料的百乐笔（以下简称“百乐笔”）。另一半被试只有用1.5美元换斑马笔的机会。研究人员将这些奖品展示给被试并告知他们，每个奖品的价格通常略高于2美元。在表明了自己的偏好后，被试会得到他们的选项。结果显示，当斑马笔是唯一的选择时，75%的被试选择了斑马笔而不是报酬，但在有两种笔的情况下，只有47%的人选择了斑马笔或百乐笔（ $p < 0.05$ ）。面对一个有吸引力的选项，被试有一个令人信服的理由放弃现金报酬：大多数人利用这个机会获得了更有吸引力、价值更高的奖品。此外，具有可比价值的竞争选项的可得性并不是选择其中一个选项的直接理由，这增加了被试保持默认方案的倾向。雷德尔迈耶和沙菲尔在由专业内科医生参与的虚拟医疗决策的实验中得出了类似的结果（Redelmeier & Shafir, 1993）。

在以上的研究中，添加一个竞争性选项可以增加默认选项的受欢迎程度。回想一下，一个选项的受欢迎程度也可以通过添加一个较差的选项来提高。因此，根据不对称优势效应，可以通过添加第3个选项z来增加x相对于y的偏好程度，选项z明显比x差但不比y差（图14-1）。J.休伯（J. Huber）、J. W.佩恩（J. W. Payne）和C.普托（C. Puto）在假设选项之间的选择中首次证明了不对称优势现象（Huber, Payne & Puto, 1982）。D. H.韦德尔（D. H. Wedell）使用货币投注研究报告了类似的发现（Wedell, 1991）。下面这个涉及实际选择的例子引自伊塔马尔·西蒙森（Itamar Simonson）和特沃斯基的相关研究（Simonson & Tversky, 1992）。在这项研究中，其中一组被试（ $N=106$ ）被要求在6美元和一支高级的高仕钢笔之间做出选择。36%的被试选择了钢笔，剩下的64%的被试选择了现金。第二组被试（ $N=115$ ）被要求在3个选项中做出选择：6美元现金、同一支钢笔，以及明显不那么吸引人的第二支钢笔。只有2%的被试选择了不那么吸引人的那支笔，但它的存在使选择高仕钢笔的被试比例从36%上升到了46%（ $p < 0.10$ ）。这种模式再次违反了前面讨论的正则条件。在其他消费品的选择中也发现了类似的行为模式。在另一项研究中，被试从“最佳”目录中看到了微波炉的描述和图片。其中一组被试（ $N=60$ ）被要求在价格为110美元的爱默生和180美元的松下电器之间做出选择。两件商品都在打折，比原价便宜1/3。在这里，57%的人选择了爱默生，43%的人选择了松下电器。第二组的被试（ $N=60$ ）则在加入了第3个选项的选择中进行决策：以10%的折扣获得价格为200美元的松下电器。只有13%的被试选择了较贵的松下电器，但是松下电器的出现使被试选择较便宜的松下电器的比例从43%上

升到了60% ($p < 0.05$)。 (37)

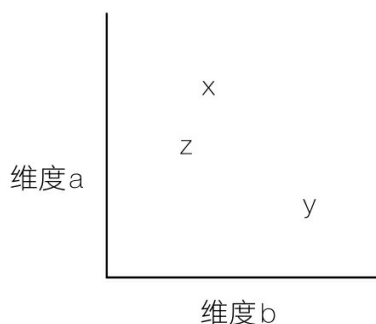


图14-1 不对称优势示意

注：x优于y的倾向可以通过添加一个选项z得到增强，z显然比x差，但不比y差。

西蒙森和特沃斯基将这些观察结果解释为“权衡对比”（tradeoff contrast; Simonson & Tversky, 1992）。他们认为，对某一选项的偏好程度的增强或者减弱，取决于所考虑的那一组选项集中对这一选项的权衡是有利的还是不利的。第二类效应称为极值厌恶（extremeness aversion），指的是，在给定的集合中，极值选项与中间选项相比吸引力较小（Simonson, 1989）。例如，考虑二维选项x、y和z，使y位于x和z之间（如图14-2所示）。根据价值最大化，中间选项y在三选一中比在二选一（y相比x或y相比z）中更不受欢迎。此外，极值厌恶表明了与预测相反的结果，因为y相对于x和z有很小的优势和劣势，而x和z都有更极端的优点和缺点。这种模式在几个实验中都被观察到了。例如，研究人员向被试展示了5架35毫米相机，它们的质量和价格各不相同。其中一组被试（ $N=106$ ）被要求在两款相机中做出选择：售价为170美元的美能达X-370和售价为240美元的美能达3000i。第二组被试（ $N=115$ ）得到了一个附加选项：定价为470美元的美能达7000i。第一组被试中选择两个选项的百分比相等，而第二组中57%的被试选择了中间组，即美能达3000i，其余的被试在两组极端方案中平分。因此，引入极端选项降低了另一个极端选项的市场份额，但没有降低中间选项的市场份额。请注意，这一影响不能归因于所提供的资料所传达的信息，因为被试在做出选择之前已经考虑了所有相关的选项。

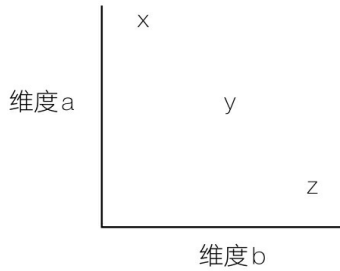


图14-2 极值规避厌恶示意

注：当x和z都可得时，与分别与x和z进行的二选一相比，选项y在三选一中相对更受欢迎。

我们认为，权衡对比和极值厌恶都可以通过推理进行理解。假设决策者面临两个选择，x和y；同时假设x质量更高，而y价格更低。如果决策者发现很难确定质量差异是否大于价格差异，那么会产生冲突。现在假设选择集还包括第3个选项z，它显然不如y但优于x，我们认为z的存在为选择y而不是x提供了一个理由。在某种程度上，在x和y之间做出选择是困难的，z的存在可以帮助决策者打破僵局。例如，在钢笔问题的研究中添加相对缺乏吸引力的钢笔，虽然其货币价值尚不清楚，但与高级的钢笔相比，其劣势是显而易见的，这为选择高级钢笔而不是现金提供了一个理由。同样，在相关维度上存在极值选项的情况下，中间选项可以被看作一种折中选择，比两个极端选项都更容易得到捍卫。事实上，口头报告表明，被试在做出这些选择时所产生的描述涉及了对不对称优势和折中的考虑；此外，当被试认为必须向他人证明自己的决定时，不对称优势会得到增强（Simonson, 1989）。值得注意的是，导致权衡对比和极值厌恶的争论本质上是相对的；它们建立于选项集中的选项的位置之上，因此不能轻易地将其转换为与单个选项相关联的值。

特沃斯基和西蒙森提出了一个形式模型，用类似锦标赛的过程来解释上述发现（Tversky & Simonson, 1993）。在这个过程中，每个选项都被拿来与其他可用选项进行比较，比较它们的相对优势和劣势。基于对选择的推理，该模型可以被看作前面定性解释的形式化类比。形式的或定性的分析哪种更有用，很可能取决于问题的性质和调查的目的。

明确的与析取的理由

人们有时会遇到不确定的情况，在这种情况下，他们最终选择了相同的行动方案，但原因非常不同，这取决于不确定性的解决方式。因此，参加过考试的学生可能会决定休假，或者是在考试及格时奖励自己，或者是在考试不及格时安慰自己。然而，如下图所示，学生可能不

愿意在考试结果悬而未决的时候去度假。特沃斯基和沙菲尔向66名本科生提出了下述问题（Tversky & Shafir, 1992a）。

析取版本

想象一下，你刚刚参加了一场艰难的资格考试。现在是秋季学期末，你感到疲惫不堪，且不确定自己是否通过了考试。如果这次考试不及格，你必须在几个月后再考一次——在圣诞假期之后。现在你有机会以极低的价格购买一个非常有吸引力的夏威夷 5 天圣诞假期套餐。特惠明天到期，考试成绩后天才能出来。你会：

(a) 购买假期套餐	32%
(b) 不购买假期套餐	7%
(c) 支付不可退款的 5 美元的定金可将优惠价格的资	61%

格保留到明天，即保留到你知悉资格考试结果之后。

选择每个选项的被试百分比如右侧数据所示。另外两个涉及及格和不及格的版本被呈现给两组不同的被试，每组为67名学生。这两个版本只在括号中的表达内容上有所不同。

及格 / 不及格版本

想象一下，你参加了一场艰难的资格考试。到了秋季学期末，你感到疲惫不堪，并且已经知晓了资格考试的成绩（通过 / 未通过）。未通过考试意味着几个月后——过了圣诞节还得再考一次。现在你有机会以极低的价格买到一个非常有吸引力的夏威夷 5 天圣诞假期套餐。特惠明天到期。你会：

	及格	不及格
(a) 购买假期套餐	54%	57%
(b) 不购买假期套餐	16%	12%
(c) 支付不可退款的 5 美元定金将优惠	30%	31%

价格的资格保留到明天

数据显示，超过一半的学生在知道自己通过了考试的时候选择了度假套餐；而在知道自己考试不及格的时候，选择度假套餐的比例甚至更高。然而，当他们不知道自己是否通过考试时，只有不到1/3的学生选

择了购买度假套餐，61%的学生愿意支付5美元，将优惠资格推迟到第二天，因为那时考试结果将会揭晓。[\(38\)](#)一旦知道了考试结果，这个学生就有了很好的去旅行的理由，尽管原因不同：如果通过了考试，假期被认为是一个辛苦但成功的学期后的奖励；如果考试不及格，假期将成为一种安慰和休整的时间，然后再重新考试。然而，由于不知道考试结果，学生缺乏去夏威夷的明确理由。请注意，考试的结果将在假期开始前很长一段时间内公布。因此，不确定性体现在做出决定的那个实际时刻，而不是最终假期。

去夏威夷的理由的不确定性使许多学生不愿花钱度假，即使这两种结果——通过或不通过考试——最终都支持了这一做法。特沃斯基和沙菲尔将上述的决策模式称为析取效应（Tversky & Shafir, 1992a）。显然，不同理由的析取（成功时的奖励或失败时的安慰）往往没有单独的明确的理由那么令人信服。实际上，上述学生中有相当一部分人愿意为那些最终不会影响他们的决定的信息付费——无论哪种情况，他们都会选择去夏威夷——但这给了他们一个更明确的理由来做出这个选择。为非工具性信息付费的意愿与经典模型不同，在经典模型中，信息的价值仅由其影响决策的潜力决定。

在无法选择推迟做出决定的情况下，人们倾向于明确的理由而不是析取的理由，这一点具有重要的影响。以下是特沃斯基和沙菲尔向98名学生提出的一系列问题（Tversky & Shafir, 1992a）。

赢 / 输版本

假设你刚刚玩了一个概率游戏，你有 50% 的机会赢 200 美元，50% 的机会输 100 美元。抛硬币，你会（赢了 200 美元 / 输了 100 美元）。你现在有第二个相同的投注机会：50% 的机会赢 200 美元和 50% 的机会输 100 美元。你会：

	赢	输
(a) 接受第二次机会	69%	59%
(b) 拒绝第二次机会	31%	41%

学生们首先看到了上述问题的获胜版本，一周后看到失败版本，10天后再看到前两个版本的析取版本。这些问题被嵌入了其他类似的问题中，所以不同版本之间的关系是不透明的。被试被要求分别对这些问题进行决策。

析取版本

假设你刚刚玩了一个机会游戏，你有 50% 的机会赢 200 美元，50% 的机会输 100 美元。假设这枚硬币已经被抛出去了，但是在决定进行第二次相同的投注之前，不知道自己是赢了 200 美元还是输了 100 美元：赢 200 美元的概率是 50%，输 100 美元的概率是 50%。你会：

- | | |
|-------------|-----|
| (a) 接受第二次机会 | 36% |
| (b) 拒绝第二次机会 | 64% |

数据显示，在赢了第一局后大多数被试接受了第二次投注，在输了第一局后大多数被试也接受了第二次投注。然而，当第一次投注的结果未知时，大多数被试拒绝了第二次投注。对个人选择的调查显示，大约 40% 的被试在无论第一次投注是获利和还是亏损后都接受了第二次投注。然而，其中 65% 的人在不知道第一次投注结果的析取条件下拒绝了第二次投注。事实上，这种反应模式（在两种情况下都接受，但在析取条件下选择拒绝）是最常见的一种模式，在所有被试中有 27% 的人表现出了这种模式。这种模式违背了萨维奇的保险原则，不能被归因于不可靠性（Savage, 1954; Tversky & Shafir, 1992a）。

研究人员请上述学生进行了一个具有正前景值的投注游戏，并且有一个损失不小的均等机会。根据第一次投注的结果，接受第二次投注的理由可能会有所不同。在赢的情况下，决策者已经获得了 200 美元，所以即使在第二次投注中输了，他总体上还是领先的，这使得这个选择非常有吸引力。此外，在输的情况下，决策者损失了 100 美元。第二次投注为他们提供了一个“扭亏为盈”的机会，对许多人来说，这比接受确定的 100 美元的亏损更具吸引力。然而，在析取条件下，决策者不知道自己是多了 200 美元还是损失了 100 美元；换句话说，她不知道自己进行第二次投注的原因是不赔钱还是因为它为自己提供了一个避免确定会损失的机会。在没有明确的理由的情况下，接受第二次投注的被试较少。

对上述问题的修订进一步支持了这一解释，即第一次投注的两种可能的结果都使决策者多了 400 美元，这样她在任何一种情况下都不会有损失。

“假设你刚刚玩了一个机会游戏，你有 50% 的机会赢 600 美元，50% 的机会赢 300 美元。假设这枚硬币已经被抛出去了，但是在决定进行第二次投注之前，你不知道自己是赢了 600 美元还是 300 美元。在第二次投注中，你有 50% 的机会赢 200 美元，50% 的机会输 100 美元。”

共有171名被试回答了这个问题，他们被平均分为3组。第一组被试被告知他们在第一次投注中赢了300美元，第二组被试被告知他们在第一次投注中赢了600美元，第三组被试被告知第一次投注的结果——300美元或600美元——是未知的（析取式）。在所有的情况下，被试必须决定是接受还是拒绝第二次投注，与之前的问题一样，第二次投注有赢200美元或输100美元的均等机会。在300美元、600美元和析取问题中，接受第二次投注的被试的比例分别为69%、75%和73%（回想一下，初始问题的对应数字分别是59%、69%和36%；在被试之间进行的有关这个问题的重复实验中得到了基本相同的数据），与初始问题的结果相反，这个修订后的问题中的第二个投注在析取式和非析取式中同样受欢迎。在初始情景中，第二次投注等于无损失或避免一定损失的机会，而在修订后情景中，无论第一次投注的结果如何，第二次投注都是无损失的。第二次投注的受欢迎程度的增加说明并不是析取情境本身阻碍了人们接受第二次投注。相反，似乎是缺乏一个特定的理由才导致了这种结果：当同样的理由适用于任何一种结果时，这种析取就不会再降低人们接受投注的倾向。

如上所述，决策情境的变化可能会改变被试所能想到的理由，从而改变他们的选择。在其他地方，我们描述了在一个一次性囚徒困境游戏中的析取效应，这个游戏是在计算机上进行的，被试可以获得真实的收益。在该研究中，被试（ $N=80$ ）玩了一系列囚徒困境游戏，且没有反馈，每个游戏都针对不同的未知玩家（Shafir & Tversky, 1992）。在这个设置中，当被试知道另一名玩家已经选择竞争时，他们的合作率为3%，当他们知道另一名玩家已经选择合作时，他们的合作率为16%。然而，当被试不知道其他玩家是选择了合作还是选择了竞争时（囚徒困境游戏的标准版本），合作的比例上升到37%。因此，许多被试在知道了对方的选择（合作或竞争）时就会选择竞争，而在不知道对方的选择的情况下就会选择合作。沙菲尔和特沃斯基将这种模式归因于不同的视角，这些视角构成了被试在不确定状况下的行为，而不是在不确定性被消除时的行为（Shafir & Tversky, 1992）。我们认为当其他玩家的决定

被知晓时且回报只取决于被试时，选择竞争的理由更有说服力，当另一方选择的策略不确定且博弈的结果取决于双方的选择时则不是这样。

上述“析取”的操作——从价值最大化的角度看没有直接关系——似乎影响了人们头脑中的做出决定的理由。下面描述的另一操作似乎可以改变人们的理由，但不直接影响选项的价值。

无价值特性

选择或拒绝的理由通常与所考虑选项的具体特性相关。选项的积极特性通常能为决策者提供选择该选项的理由，而其消极特性通常提供的是拒绝它的理由。当我们添加既不吸引人也不令人反感的特性时会发生什么？选择是否会受到价值很小或没有价值的特性的影响？

西蒙森和他的同事对无价值特性的影响进行了一系列的研究，并验证了这一假设，即人们不愿意选择那些由缺乏吸引力的理由所支持的选项。例如，在一项研究中，西蒙森等人预测：如果其他人选择了某一个选项，且如果做出这个选择的理由并不适用于我们自己，那么我们不太可能会做出同样的选择（Simonson et al., 1993）。加州大学伯克利分校商学院的学生（ $N=113$ ）被告知，由于预算削减，为了节省纸张和复印成本，他们将收到一份供两名被试共同使用的问卷。因此，当被试必须输入一个选项时，他们可以看到前一名“被试”做出的选择及给出的理由。前一名被试的选择和理由是被系统操纵的。例如这样一个问题：他们需要在美国西北大学和加州大学洛杉矶分校攻读MBA之间做出选择。在问卷的一个版本中，前一名被试选择了西北大学，并给出了（手写的）理由：“我在芝加哥地区有很多亲戚。”因为这个理由并不适用于大多数的被试，所以预计会降低他们选择西北大学的可能性。在第二个版本中没有给出前一名被试选择西北大学的理由。正如所料，那些接触到不相关理由的人选择西北大学的可能性比那些看到其他被试的选择但没看到他们给出的不适用于自己的理由而选择西北大学的人要小（比例分别为23%与43%， $p<0.05$ ）。值得注意的是，西北大学和加州大学洛杉矶分校在大多数学科上都非常有名（西北大学目前拥有排名最高的MBA课程；加州大学洛杉矶分校的课程排名很高，它与加州大学伯克利分校属于同一系统）。因此，被试不太可能根据另一名被试选择西北大学是因为他在芝加哥有亲戚的事实来推断西北大学的教学质量。

在一个相关研究中，西蒙森等人发现，赋予选项一个旨在加分的功

能实际上对决策者来说没有价值，这种情况下会降低人们选择这一选项的可能性，即使被试已经意识到自己无须为这个额外的功能或特性付出代价（Simonson et al., 1994）。例如，如果一个人买了某一个品牌的蛋糕粉，这时候告诉他可以买某个收藏盘（大多数人都不想要），那么他选择这一个品牌的蛋糕粉的概率会低于他选择另一个差不多品牌的蛋糕粉的概率（从31%降至14%， $p < 0.05$ ）。对于那些提供了无意义的赠品的品牌（在一项相关研究中），人们更难以证明选择它们的合理性，这种行为也更容易受到批判。对口头回应的分析表明，大多数没有选择添加功能选项的人都明确提到不需要添加该功能。应该强调的是，如上面提到的收藏盘子的促销方法，目前被广泛的公司采用，而且没有证据表明这些促销会导致人们对促销产品的质量做出任何推断（Blattberg & Neslin, 1990）。

上述操作都增加了“积极”的功能，尽管功能较弱或无关紧要，但不应降低选项的价值；然而，它们却显然给出了一个反对选择该选项的理由，尤其是在其他选项同样具有吸引力的情况下。显然，添加一个有潜在吸引力的功能（但被证明是无用的），实际上是提供了一个拒绝该选项的理由，而人们更倾向于选择那个没有“浪费”功能的竞争性选项。

结束语

人们的选择有时可能源于情感判断，而情感判断妨碍了对选项的全面评估（Zajonc, 1980）。在这种情况下，对做选择的理由进行分析可能被证明是没有根据的，并且决策者的努力实际上可能会产生一个不同的，甚至是更差的决定（Wilson & Schooler, 1991）。此外，其他选择可能会遵循标准的操作程序，不需要太多的内省反思。然而，许多决定都是经过仔细评估后做出的，人们试图做出他们认为最好的选择。当人们放弃了不那么吸引人的选项和面临着一个困难的选择时，他们往往会寻找一个令人信服的选择一个选项而不是另一个选项的理由。在这一章中，我们分析了理由在决策中的作用，并考虑了基于推理的分析是如何推进和完善基于价值最大化的标准的定量方法的。从这一角度出发，我们在实验环境中研究了一系列的假设。

做出决定的理由可能是复杂多样的。在前面的几节中，我们试图确定一些指导理性决策过程的一般性原则，因此理性思考的一些基本方法可能有助于我们理解决策过程。第1节显示了对更重要维度（那些可能提供了更有说服力的选择理由的维度）的依赖可以预测人们在原本吸引

力等同的选项之间的偏好。兼容性和显著性的概念在第2节中被提出，以解释在选择和拒绝任务中不同理由的不同权重。理由似乎倾向于某些框架操纵，我们很难从价值最大化的角度对这些操纵进行解释。然后，在第3节中我们证明了，对相互竞争的选项之间的精确关系进行操纵能加强或减少冲突，产生更容易或更难以合理化和得到证明的决定。提供可以说明选择一个选项的令人信服的理由的情境，显然会增加人们选择该选项的倾向，而那些妨碍上述推理过程的其他选项，往往会增加人们维持现状或寻找其他选项的倾向。在第4节中，我们进一步讨论了产生推理的决策环境是如何影响人们的选择的，选项的增减可以被理解为基于相对优势和折中的比较性考虑而形成的选择的理由。我们在第5节中讨论了析取理由的相对弱点。在这一部分中，一系列的研究对比了人们基于一个明确的理由做出决定的意愿和在不确定的情况下根据勉强的理由做出决定的意愿。我们在第6节中简要回顾了一种决策情境，在这种情境下，增加一个本意为增加选择倾向却不被被试接受的理由，其实会降低他们选择该选项的倾向，尽管该选项的价值并没有减少。

指导决策的推理的本质及其相互作用的方式，有待进一步的研究。有证据表明，各种各样的因素都在决策过程中发挥着作用。无论是出于人际的目的，还是出于个人内在的动机，我们经常会为自己所做的决定寻找一个可以令人信服的理由；这样，我们就可以向他人解释我们做出决定的原因，从而可以自信地做出“正确”的选择。对于风险和损失的态度有时会在神话和陈词滥调的基础上被合理化，有时我们的选择是基于计算具体的成本效益的道德或谨慎原则（Prelec & Herrnstein, 1991）。此外，决策的形式规则有时也可以作为人们参与讨论的论据。因此，当我们在x和z之间进行选择时，可以意识到，在之前的某个时候，我们认为x优于y而y优于z，因此通过传递性，我们现在应该选择x而不是z。蒙哥马利认为，人们会在决策问题中寻找优势结构，是因为他们要为选择寻求一个值得信服的理由（Montgomery, 1983）。类似地，特沃斯基和沙菲尔已经证明，确定原则在决策情景中的适用性会导致人们按照该原则的基本原理行事（Tversky & Shafir, 1992a）。事实上，我们已经多次观察到，在非透明的情况下，人们经常会违反理性选择的定理，而该定理通常只有在透明的情况下才会得以适用和满足（Tversky & Kahneman, 1986）。这些结果表明了理性选择的定理在某些特定的、适用性的决策中是一种有力的论据或理由，但它不是约束人们选择的普遍规律。

与假设稳定的价值和偏好的经典理论相反，人们往往没有确定的价

价值观，偏好实际上也是在诱导过程中形成的，而不仅仅是显露出来的（Payne, Bettman & Johnson, 1992）。基于推理的方法很适合这样的建设性解释。根据这种分析，决策往往是通过集中注意力证明一个选项比另一个选项更合理而做出的。不同的框架、情境和启发式强调了选项的不同方面，并提出了影响决策的不同理由和考虑因素。

依赖于推理去解释实验结果一直是社会心理学分析的特点。例如，对认知失调（Wicklund & Brehm, 1976）和自我知觉（Bem, 1972）的解释集中于人们试图解释他们的反态度行为时聚集起来的理由上。同样，归因理论也围绕着人们对他人行为进行归因的问题展开（Heider, 1980）。然而，这些研究主要集中于决策后的合理化，而不是决策前的冲突。虽然这两个过程密切相关，但仍有一些重要的差异。社会心理学的很多研究都是关于人们的决策是如何影响他们思考的方式的。相比之下，本文考虑的是人们思考问题时的推理是如何影响他们的决策的。最近，一些研究人员开始探索相关的问题。例如，M.比利希（M. Billig）采用了一种修辞方法来理解社会心理问题，根据这种方法“我们的内部审判是自我执行的无声宣言”（Billig, 1987）。N.彭宁顿（N. Pennington）和R.黑斯蒂应用基于解释的决策模型来解释司法判决，而菲利普·泰特洛克（Philip E. Tetlock）则阐述了社会责任在选择中的重要性（Pennington & Hastie, 1988, 1992; Tetlock, 1992）。在哲学方面，F.希克（F. Schick）的一篇文章从实践理性的角度对各种决策进行了分析（Schick, 1991）。早期有影响力的研究是S.图尔明（S. Toulmin）对伦理推理中论证的作用的研究（Toulmin, 1950）。

在这一章中，我们尝试探索了一些影响人们决策方式的原因和争论。基于推理的分析可能更接近于构成决策基础的心理学部分，从而有助于从经典理论的角度揭示一些违反直觉的现象。值得注意的是，本章所描述的许多实验研究的动机都是基于推理的定性分析所产生的直觉，而不是基于价值最大化的观点，即使后来可以用这种方式对其进行解释。我们不是要用推理分析的方法取代基于价值的选择模型，而是认为对推理的分析可以阐明反思性选择的某些方面，并为进一步的研究提供新的假设。

致谢

本章部分内容是由第一作者参加行为科学高级研究中心以谈判和争议解决为主题的暑期研讨班时编写的。本章第二作者来自加州大学伯克

利分校。我们感谢罗宾·道斯在形成初稿过程中给予的帮助。

参考文献

Allison, G. T. (1971). *Essence of decision: Explaining the Cuban missile crisis*. Boston, MA: Little, Brown.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). New York, NY: Academic Press.

Berman, L. (1982). *Planning a tragedy*. New York, NY: Norton.

Betts, R., & Gelb, L. (1979). *The irony of Vietnam: The system worked*. Washington, DC: Brookings Institution.

Bigelow, J. (Ed.). (1887). *The complete works of Benjamin Franklin* (Vol. 4). New York, NY: Putnam.

Billig, M. (1987). *Arguing and thinking: A rhetorical approach to social psychology*. New York, NY: Cambridge University Press.

Blattberg, R. C., & Neslin, S. A. (1990). *Sales promotion: Concepts, methods, and strategies*. Englewood Cliffs, NJ: Prentice-Hall.

Heider, F. (1980). *The psychology of interpersonal relations*. New York, NY: Wiley.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, 9, 90–98.

James, W. (1981). *The principles of psychology* (Vol. 2). Cambridge, MA: Harvard University Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.

Montgomery, H. (1983). Decision rules and the search for a dominance structure: Towards a process model of decision making. In P. Humphreys, O. Svenson, & A. Vari (Eds.), *Analyzing and aiding decision processes* (pp. 343–369). Amsterdam, the Netherlands: North-Holland.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive process perspective. *Annual Review of Psychology*, 43, 87–131.

Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 521–533.

Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the story model for juror decision making. *Journal of Personality and Social Psychology*, 62, 189–206.

Prelec, D., & Herrnstein, R. J. (1991). Preferences or principles: Alternative guidelines for choice. In R. J. Zeckhauser (Ed.), *Strategy and choice* (pp. 319–340). Cambridge, MA: MIT Press.

Redelmeier, D., & Shafir, E. (1993). Medical decisions over multiple alternatives. Working paper, University of Toronto, Toronto, Ontario, Canada.

Savage, L. J. (1954). *The foundations of statistics*. New York, NY: Wiley.

Schick, F. (1991). *Understanding action: An essay on reasons*. New York, NY: Cambridge University Press.

Shafer, G. (1986). Savage revisited. *Statistical Science*, 1, 463–485.

Shafir, E. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory & Cognition*, 21, 546–556.

Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive Psychology*, 24, 449–474.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, 16, 158–174.

Simonson, I., Carmon, Z., & O'Curry, S. (1994). Experimental evidence on the negative effect of unique product features and sales promotions on brand choice. *Marketing Science*, 13, 23–40.

Simonson, I., Nowlis, S., & Simonson, Y. (1993). The effect of irrelevant preference arguments on consumer choice. *Journal of Consumer Psychology*, 2, 287–306.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *JMR, Journal of Marketing Research*, 29, 281–295.

Slovic, P. (1975). Choice between equally valued alternatives. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 280–287.

Slovic, P. (1990). Choice. In D. Osherson & E. Smith (Eds.), *An invitation to cognitive science* (Vol. 3, pp. 89–116). Cambridge, MA: MIT Press.

Telhami, S. (1990). *Power and leadership in international bargaining: The path to the Camp David accords*. New York, NY: Columbia University Press.

Tetlock, P. E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 331–376). New York, NY: Academic Press.

Toulmin, S. (1950). *The place of reason in ethics*. New York, NY: Cambridge University Press.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, 59, 251–278.

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, 95, 371–384.

Tversky, A., & Shafir, E. (1992a). The disjunction effect in choice under uncertainty. *Psychological Science*, 3, 305–309.

Tversky, A., & Shafir, E. (1992b). Choice under conflict: The dynamics of deferred decision. *Psychological Science*, 3, 358–361.

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management Science*, 39, 1179–1189.

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Wedell, D. H. (1991). Distinguishing among models of contextually induced preference reversals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 767–778.

Wicklund, R. A., & Brehm, J. W. (1976). *Perspectives on cognitive dissonance*. Hillsdale, NJ: Erlbaum.

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology*, 60, 181–192.

Zajonc, B. (1980). Preferences without inferences. *American Psychologist*, 35, 151–175.

后记 特沃斯基是天生的决策理论家

丹尼尔·卡尼曼

本书所收录的文章展现了阿莫斯·特沃斯基这个传奇人物一生中的诸多闪光点。这些文章虽然写于几十年前，但它们至今仍散发着光芒。这些研究长盛不衰的一个原因是其论证非常清晰且有权威性。阿莫斯的许多作品给读者留下的印象是：他所做的研究非常权威，无法再改进。当然，阿莫斯并不总是正确的，但他从来不会在理由不充分时发表观点，而且其他人也无法对他提出的观点做出更好或更有说服力的阐述。

20世纪60年代早期，阿莫斯攻读研究生时，将数学思想引入社会科学领域的思想运动正处于鼎盛时期。他是受该运动启蒙的第一代学生，他与一些发起人都有着广泛的合作。这些早期经历为他使用形式语言打下了基础，也成为他之后思考的工具。形式语言本身能够保证精确性，但其他方面不是特别好。它完全有可能会变得非常无聊、完全不相关，甚至是完全错误的。阿莫斯使用精确的数学推算来避免出现近似正确的情况，他想要完全正确。他使用理论语言来帮助自己和读者清楚地思考重要的事情，他也因此成为心理学家中取得了特别成功的那一个。

阿莫斯的作品不会过时的另一个因素是他的写作质量。有一个著名的例子，心理学家就《心理学评论》在120年间发表的最佳段落达成了共识。获胜者是阿莫斯撰写的“相似性特征”一文（第3章）中的最后一段。许多心理学家在就读研究生时都很熟悉这篇文章的最后4句话。它们是这样写的：

比如“一篇文章就像一条鱼”这个比喻，一开始这句话令人费解。一篇文章不会具有腥、滑或湿这样的特点。但当我们想到一篇文章（像鱼一样的）有头有尾，并且偶尔也会在结尾处来一记翻尾时，那么这个困惑就解除了。

可供自我参照的结束语是阿莫斯思想的精髓，它精妙、清晰且有深度。本书的每一章中都有令人难忘的只属于特沃斯基的句子。在阿莫斯自己的用词中，“干脆”和“优雅”用于表达高度赞扬，而这两个词正是对其写作水平的完美概括。

本书的读者可以欣赏阿莫斯作品独特的精确性和优雅感，与他合作的人也有同感。多年来，我有机会观察到，阿莫斯写出令人难忘的句子的过程和很多诗人创作诗歌的过程一样。当一个想法在头脑中出现后，他通常会用铅笔将这个想法变成第一份初稿，为的是方便修改。完美来自耐心的打磨：首先要删除每一个多余的单词，然后思考留下的单词的替代词。接着重新开始这个过程，直到满意为止。当最初的想法很好时（阿莫斯的想法向来很好），最终也会得到一些无可挑剔的东西。

完美和完美主义有多种表现形式，焦虑、折磨和尽责也是其中的几种。阿莫斯的完美主义是以快乐的形式体现的。他喜欢思考，喜欢寻找最恰当的用词。他非常喜欢将原石打磨成闪闪发光的宝石这个缓慢的过程。他在工作中找到的快乐给了他无尽的耐心，他乐于向合作者分享。在我们一起工作的几年中，他经常说的一句话是“让我们把它做好！”，说这话的时候他总是很高兴。阿莫斯给了我们一些忠告，他承认在充满惊喜的旅程中会有一段漫长且颇具挑战性的路要走，但他目标明确，且坚信目标终将达成。分享这些旅程对他的合作者来说很有意义，阿莫斯的很多文章都是他独自写就的，但本书只收录了其中一篇。

有一些章节是在阿莫斯和我合作关系非常密切的12年间完成的。我们共同完成的判断和决策方面的作品比我们想象中更有影响力。它的影响部分源于成功地综合了各种方法，而对研究方法偶然天成的选择和阐述的格式可能是成功更为重要的因素。

阿莫斯是天生的理论家，也是经过训练的决策理论家。对他来说，将人们的判断和选择与概率和决策近似等同起来的逻辑是不完美的尝试，产生这种想法是很自然的事。我的研究领域是视觉感知，理解关于人们表征世界时是如何思考和选择的，对我来说很自然。事实证明，这些完全不同的方法的融合是富有成效的。

在解释我们的共同研究的影响时，研究风格带来的影响可能超过了研究本身。我和阿莫斯很少进行复杂的实验。相反，我们构建了一些例子，其中对单个问题（或一对相互匹配的问题）的简明回答就可以阐明一个特定的理论观点。我们发表的论文中包含了这些问题，并用特定格式将它们与其他文本分开。因此，读者必然能感受到我们研究中的参与者的体验。他们通常会发现自己身上也会出现我们报告中描述和解释的所有问题（包括预测错误）。个人经历的认知错觉和不连贯性选择对读者产生了强烈的影响。很明显，我们的研究对其他领域产生不寻常影响的主要原因是，我们很幸运：在认知或社会心理学等其他领域，没有研

究人员能如此自然地与读者进行过私人对话。

演示对我们的跨学科作品的传播至关重要。律师、哲学家、政治学家、医生和一些经济学家对他们的经历感到惊讶且印象深刻，并热衷于将他们学到的知识运用到自己感兴趣的领域。判断中的错觉和选择的不连贯性在许多应用领域中都被记录了下来，其中也包括核反应堆安全、复杂谈判的进程、足球运动员的选择以及退休存款的推进。

阿莫斯很高兴看到这项工作在其他领域产生的影响，但他参与最多的是“大型理性辩论”，即与决策理论和经济学方面的学者对话，这些学者希望将理性模型用作说明人类如何思考和选择的描述性理论。尽管在早期研究中我们没有明确地讨论理性，但我们公布的观察结果很快被认为是对于理性行动者模型（**rational agent model**）的重大挑战，并且引发了各种防御性反应。对于试图改变理性公理的内涵，将明显违背理性的例证重新归为理性的尝试，阿莫斯感到非常不耻和愤怒。他认为这些对于人类理性的防御充满了诡辩味道，毫无优雅可言，他迫切希望揭示它们的缺陷。在介绍选择和框架效应的那篇文章中（第5章），阿莫斯几乎做了全新的研究，这篇文章重新应用了早期的实验研究结果，得出了一套全新的理论观点。

从职业生涯之初，阿莫斯就努力且准确地描述了人类思维会以何种方式违背理性规则。一个反复出现的主题是外延逻辑和内涵直觉之间的对比，这是本书中几篇文章的核心思想。现实与描述之间的区别是阿莫斯思想的另一个关注点。他与其他合作者的很多后续工作都关注了规范性和描述性规则之间的这些对比。

很少有学者的作品可以在其去世几十年后还未成为历史文献，而是作为经久不衰的知识贡献重新出版。阿莫斯就是其中一个。

未来，属于终身学习者

我这辈子遇到的聪明人（来自各行各业的聪明人）没有不每天阅读的——没有，一个都没有。巴菲特读书之多，我读书之多，可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式；指数型技术在迅速颠覆着现有的商业世界；人工智能已经开始抢占人类的工作岗位……

未来，到底需要什么样的人才？

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才，而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力，获得异于众人的思考和整合能力。未来，将属于终身学习者！而阅读必定和终身学习形影不离。

很多人读书，追求的是干货，寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代，答案不会简单地出现在书里，因为生活根本就没有标准确切的答案，你也不能期望过去的经验能解决未来的问题。

而真正的阅读，应该在书中与智者同行思考，借他们的视角看到世界的多元性，提出比答案更重要的好问题，在不确定的时代中领先起跑。

湛庐阅读 App：与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上，把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本

怎么读是读者面临的最大的阅读障碍

“读书破万卷”不仅仅在“万”，更重要的是在“破”！

现在，我们构建了全新的“湛庐阅读”App。它将成为你“破万卷”的新居所。在这里：

- 不用考虑读什么，你可以便捷找到纸书、电子书、有声书和各种声音产品；
- 你可以学会怎么读，你将发现集泛读、通读、精读于一体的阅读解决方案；
- 你会与作者、译者、专家、推荐人和阅读教练相遇，他们是优质思想的发源地；
- 你会与优秀的读者和终身学习者伍，他们对阅读和学习有着持久的热情和源源不绝的内驱力。



下载湛庐阅读 App，
坚持亲自阅读，
有声书、电子书、阅读服务，
一站获得。

CHEERS

本书阅读资料包

给你便捷、高效、全面的阅读体验

本书参考资料

湛庐独家策划

- ✓ 参考文献
为了环保、节约纸张，部分图书的参考文献以电子版方式提供
- ✓ 主题书单
编辑精心推荐的延伸阅读书单，助你开启主题式阅读
- ✓ 图片资料
提供部分图片的高清彩色原版大图，方便保存和分享

相关阅读服务

终身学习者必备

- ✓ 电子书
便捷、高效，方便检索，易于携带，随时更新
- ✓ 有声书
保护视力，随时随地，有温度、有情感地听本书
- ✓ 精读班
2~4周，最懂这本书的人带你读完、读懂、读透这本好书
- ✓ 课程
课程权威专家给你开书单，带你快速浏览一个领域的知识概貌
- ✓ 讲书
30分钟，大咖给你讲本书，让你挑书不费劲



湛庐编辑为你独家呈现
助你更好获得书里和书外的思想和智慧，请扫码查收！

(阅读资料包的内容因书而异，最终以湛庐阅读App页面为准)

(1)即《选择、价值与决策》(*Choices, Values and Frames*)。卡尼曼在行为决策领域的研究过程中,发现了人类判断中极易被忽视的影响因素——噪声,并在《噪声》一书中详细介绍了其成因和应对方法。《噪声》中文简体字版已由湛庐引进、浙江教育出版社于2021年出版。——编者注

(2)1磅≈0.45千克;1英里≈1.6千米。为了确保实验结论的准确性,本书中的英制单位不改为国际制单位。——编者注

(3)感谢哈佛大学医学院的芭芭拉·麦克尼尔(Barbara J. McNeil)、美国塔夫茨大学医学院的斯蒂芬·波克尔(Stephen G. Pauker)、斯坦福大学医学院的爱德华·贝尔(Edward Baer),在设计临床问题和收集资料方面提供的帮助。

(4)对于斯坦福大学一组修过一门或多门统计学课程的高年级本科生($N=62$)来说,合取谬误的发生率要低得多(28%)。

(5)本篇论文于1977年在著名期刊《心理学评论》(*Psychological Review*)上发表,当时苏联尚未解体。——编者注

(6)为了从定性假设中推导出特征可加性,我们必须假设广泛结构的公理,以及广义和联合尺度的相容性(Krantz et al., 1971; 第10.7节)。

(7)所有实验的被试都是年龄为18~28岁的以色列大学生。该材料以小册子的形式呈现,并在小组环境中进行管理。

(8)本文中涉及的所有问题与实验中呈现给参与者的内容一样。

(9)如果 $p+q=1$ 且 $x>y>0$ 或 $x<y<0$ ，那么前景的值由 $v(y) + \pi(p)$
 $[v(x) - v(y)]$ 给定，因此决策权重不适用于确定性结果。

(10)为什么你总是对那些重要信息视而不见？你是否遗漏了信息背后的信息？马克斯·巴泽曼在《信息背后的信息》中介绍了5种突破信息获取屏障的方法。该书中文简体字版由湛庐引进、浙江人民出版社于2019年出版。——编者注

(11)将本分析扩展到具有许多（非零）结果的前景的过程涉及两个附加步骤。首先，我们假设在框架阶段连续（或多值）分布近似于离散分布，且结果的数量相对较少。例如，区间 $(0, 90)$ 上的均匀分布可以由离散前景 $(0, 0.1; 10, 0.1; \dots; 90, 0.1)$ 表示。其次，在多结果情况下，加权函数 $\pi_p(p_i)$ 必须依赖于概率向量 p ，而不仅仅取决于分量 p_i , $i=1, \dots, n$ 。例如，J.奎金(J. Quiggin)使用函数 $\pi_p(p_i) = \pi(p_i) / [\pi(p_1) + \dots + \pi(p_n)]$ (Quiggin, 1982)。如在存在两个结果的情况下，假设加权函数满足次确定性， $\pi_p(p_1) + \dots + \pi_p(p_n) \leq 1$ ，以及次比例。

(12)在修订版本问题9'至问题11'中，通过从包含100个顺序编号的票证包中抽取数字来生成获胜概率。在问题10'中，与赢得45美元（抽取1~20的数字）相关事件被包括在与赢得30美元相关事件中（抽取1~25的数字）。问题11的顺序设置被两个同时运行的随机设备取代：掷骰子（其结果决定游戏是否开启）然后从包中抽取编号的票据。现在这3个问题

都存在后悔的可能性，问题10'和问题11'在这方面不再有所不同，因为决策者总是知道替代选择的结果。因此，后悔理论无法解释在修订问题中观察到的确定性效应（问题9'相对于问题10'）或假确定性效应（问题10'相对于问题11'）。

(13)值得注意的是，问题11和问题12中使用的条件框架（案例3）在消除对阿莱斯悖论的共同反应方面比萨维奇引入的分区框架更有效（Slovic & Tversky, 1974）。这可能是因为在消除了其结果不依赖于一个人的选择的状态之后，条件框架清楚地表明了关键选项是相同的（即在问题11中达到第二阶段，问题12中的无法治愈的肿瘤，案例3）。

(14)由于本文主要关注具有已知概率的前景，因此我们不讨论由于含糊不清而导致消除性的重要违规行为（Ellsberg, 1961）。

(15)结果主义的概念在哲学和决策理论文献中有许多不同的意义（Hammond, 1988; Levi, 1991; Bacharach & Hurley, 1991）。

(16)认知能力会影响人们的决策。在《表象与本质》一书中，侯世达揭示了人类认知中隐藏的核心机制。该书中文简体字版已由湛庐引进、浙江人民出版社于2018年出版。——编者注

(17)在《直觉泵和其他思考工具》一书中，丹尼尔·丹尼特展示了数十种思考工具，教读者做出独立且清晰的思考。该书中文简体字版已由湛庐引进、浙江教育出版社于2018年出版。此外，丹尼尔·丹尼特在《丹尼尔·丹尼特讲心智》一书中探讨了心智的本质。该书中文简体字版已由湛庐引进、天津科学技术出版社于2021年出版。——编者注

(18)S.赫尔利 (S. Hurley) 提出了对只有一次机会的囚徒困境游戏中合作行为的“集体行动”解释 (Hurley, 1989, 1991)。她将这种行为解释为“非常理性”，因为据她所说，该行为的动机是“参与、尽自己的一份力量、参与.....一种有价值的集体代理形式” (1989)。然而，正如前面提到的道德论点一样，这种解释意味着当对方合作时，被试当然应该倾向于合作，这与目前的调查结果相反。

(19)“如果被试在囚徒困境实验中产生至少两种析取效应，那么在盒子B中放250分；否则，不在盒子B中放任何东西”这一非常简单的程序令70%的拿了一个盒子的人和29%的拿了两个盒子的人在盒子B中获得了250分。更复杂的规则可能更接近麻省理工学院程序所显示的表现。

(20)人们可以区分未来事件结果的不确定性和已经发生的事件结果的不确定性。虽然目前的研究没有系统地区分这两者，但S.B.格林 (S.B. Greene) 和D.J.约尔斯 (D.J. Yollles) 提供的数据有理由预言前者比后者会有更多的非结果主义推理 (Greene & Yollles, 1990)。

(21)在几十项使用基本形式的选择任务（带有“抽象”材料）的研究中，初始选择的成功率通常介于0和略高于20%之间 (Evans, 1989; Gilhooly, 1988)。

(22)有些人认为，在美国总统大选之前，美联储实际上参与了防止美元暴跌的行动。

(23)我们使用的“能力”这一术语，广义上包括技能、知识或理解力。

(24)在本次及之后的所有图中，我们绘制了 C 在 P 上的保序回归，即用最小二乘法做最佳拟合单调函数（Barlow et al, 1972）。

(25)感谢帕雷尔和卡尼曼为我们提供了这些数据。

(26)所需的估计可以用几种方式解释。一种可能的方法是遵循常见的研究实践，即用一项研究中获得的值来定义一个零假设的合理替代方案。然后，问题中要求的概率可以被解释为第二次测试的效度（即在第二组样本中获得显著性结果的概率）相对于第一个样本的结果所定义的备选假设。在方差已知的均值检验的特殊情况下，我们可以根据总体均值等于第一个样本均值的假设来计算检验的效度。由于第二组样本量是第一组样本量的一半，因此获得 $z \geq 1.645$ 的计算概率仅为0.473。从理论上讲，更合理的方法是在贝叶斯框架内解释期望概率，并根据适当选择的先验分布来计算它。假设先验均匀，则前景的后验概率为0.478。显然，如果先验分布有利于零假设，通常情况下，后验概率很可能会更小。

(27)W.爱德华兹认为人们无法从概率数据中提取足够的信息或确定性；他认为这种失败是保守主义的（W. Edwards, 1968）。我们的受访者很难说是保守派。相反，根据代表性假设，他们倾向于从数据中提取更多的确定性，而不是数据实际上包含的确定性。

(28)如果一个人过高地估计了他的偏好假设的概率，就会表现出过度自信。适当的概率估计可以凭经验确定（如通过人的命中率）或从适当的模型导出。

(29)为了探讨强度和权重之间的相关性的影响，我们用另一组刺激物进行重复实验，被选择的刺激物在两个独立变量之间具有较小的相关性（与 $r=-0.64$ 相比， $r=-0.27$ ）。这组刺激物的结果与文中公布的结果非常相似，即中位数数据的回归权重产生了有利于强度的接近2:1的比率。

(30)对于这个令人费解的观察结果，有一个可能的解释是，被试公布了他们确切知道的项目数量，而无须矫正猜测。

(31)这是矛盾性陈述的统计版本，即“我相信我的所有信念，但我也相信我的一些信念是错误的”。

(32)一些作者将这一发现解释为模糊厌恶，即人们更喜欢投注对给定概率 p 的更可靠的估计而不是不太可靠的估计（Einhorn & Hogarth, 1985）。然而，这种演示并没有建立来源偏好，因为它也没有考虑相关事件的补充。因此，上述发现可归因于这样一个事实，即与 p 的较不可靠估计相关的主观概率没有与 p 的较可靠估计相关的主观概率极端，即更接近0.5（Heath & Tversky, 1991）。更一般地说，经常被引用的结论是：人们对大概率的厌恶和对小概率的模糊性寻求是有疑问的，因为它所基于的证明不能适当地控制主观概率的变化。

(33)福克斯和韦伯收集的未公开的数据显示，在接触熟悉的前景之后评估不熟悉的前景时给予的投注价格更低。

(34)格尔德·吉仁泽（Gerd Gigerenzer）进一步论证：在单个事件概率判断中观察到的偏差，在频率的判断任务中会消失，但这里的数据回顾和其他地方的回归存在不一致的说法（Gigerenzer, 1991）。

(35)当一个人同时评估这些概率时，强化效应，如次可加性，可能不成立，因为在这个任务中引入了额外的约束条件。

(36)我们感谢作者向我们提供了相关数据。

(37)情境对选择的影响自然可以用在销售策略中。例如，一家位于圣弗朗西斯科的邮购公司，曾经出售价格为279美元的烤面包机。后来，他们又增加了第二种与第一种相似烤面包机，但稍大一些，价格为429美元，比最初的面包机的价格高出50%以上。该公司没有卖出很多新产品是必然的。然而，较便宜的电器的销量几乎翻了一番。（据我们所知，该公司没有预料到这种效果。）

(38)另外一组被试（ $N=123$ ）同时看到了不及格和及格两种版本，并被询问是否会在特定情况下购买度假套餐。在这两种情况下， $2/3$ 的被试做出了相同的选择，这表明析取版本的数据不能用假设来解释，即那些通过考试后选择度假的人在考试不及格时不会选择度假，反之亦然。请注意，虽然只有 $1/3$ 的被试根据考试结果做出了不同的决定，但超过60%的被试在不知道结果的情况下选择等待。

CHEERS
湛庐



NOISE

噪声

人类判断的缺陷
A Flaw in Human Judgment

诺贝尔经济学奖得主
《思考，快与慢》作者
丹尼尔·卡尼曼 全新力作

[以色列] 丹尼尔·卡尼曼
[法] 奥利维耶·西博尼
[美] 卡斯·R. 桑斯坦

Daniel Kahneman
Olivier Sibony
Cass R. Sunstein

著

李纾 汪祚军 魏子晗 等 译

 浙江教育出版社
ZHEJIANG EDUCATION PUBLISHING HOUSE

版权信息

本书纸版由浙江教育出版社于2021年9月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

书名：噪声

著者：丹尼尔·卡尼曼 奥利维耶·西博尼 卡斯·R. 桑斯坦

电子书定价：125.99元

Noise: A Flaw in Human Judgment.

Copyright © 2021 by Daniel Kahneman, Olivier Sibony, and Cass R.

Sunstein.

All rights reserved.

献给诺加、奥里和吉利

——丹尼尔·卡尼曼

献给范丁和莱莉亚

——奥利维耶·西博尼

献给萨曼莎

——卡斯·R.桑斯坦

重磅赞誉 Recommendation

从“偏差”到“噪声”，作为心理学家的卡尼曼，挑战的是经济学的“理性人”假设，并因开启了行为经济学的大门而获得诺贝尔经济学奖。他一直关注的是人类在决策中是如何犯错的，在这本书中，他深入分析了噪声的影响，但也乐观地预期，可以通过发掘埋没在大数据中的信息来减少决策中的噪声，这可以说是大数据渗透到行为经济学领域的新趋势。

巴曙松

北京大学汇丰金融研究院执行院长，中国银行业协会首席经济学家

继《思考，快与慢》之后，湛庐又推出了卡尼曼（与西博尼和桑斯坦合著）的新作《噪声》。本书开篇就对噪声（noise）与偏差（bias）做了区分。简单来说，噪声是无规律的错误，偏差是系统性的错误。你一旦听到这个区别，立刻就能知道它十分重要。《噪声》通过司法判决、医学诊断等多个领域的实际案例，探讨了噪声的成因以及减少噪声的策略。值得一提的是，三位作者最后还讨论了克服噪声的努力也可能带来负面后果。又一本读着不累且能时时受益的好书。

陈嘉映

首都师范大学哲学系教授

阅读卡尼曼等人的新作让我兴致盎然、心潮澎湃。可以肯定地说，通过理解并实践卡尼曼等人的思想与建议，我们能知道如何减少可能会困扰组织和个人的系统噪声，知道“决策信息的简约化”居然也如此意义重大！并且，更为重要的是，我们能够学习在噪声中运用科学知识、深刻的见解和坚实的原则，做出更为人道、更为公平的决策，减少时间、金钱和人才的浪费，拯救生命，传递善意，分享意义。而这些，都是一

个真实而积极的生命所追求的高级境界。

彭凯平

清华大学社会科学学院院长

人类判断的错误源于偏差和噪声。当人们表现出不同的偏差时，噪声就产生了。卡尼曼因在心理偏差研究方面做出的突破性贡献而获得诺贝尔奖，如今又开始对噪声进行探索。本书对于噪声的分析充满超凡的见解，并更新了我们对已有众多决策研究发现的认知。《噪声》一书中关于减少噪声的方法，既有助于让人变得聪慧，又有助于那些自命不凡的“专家”纠正无知。

张志学

北京大学光华管理学院教授

在判断与决策问题上，管理者和专业人士往往都非常自信。卡尼曼与两位合作者的这部新书《噪声》再次告诉我们：人类远远不如我们想象的那么理性，人类的判断也远远不如我们想象的那么完美。对于决策者来说，“认识你自己”是一个永恒的主题；警惕认知中的缺陷，比什么都重要。如果2021年只能推荐一本书，那我推荐《噪声》。

宫玉振

北京大学国家发展研究院管理学教授、BiMBA商学院副院长兼EMBA学术主任

《噪声》提供了一个正确的认知体系，对于日益富足且对未来颇有期待的大众和企业家而言，在后金融危机时代以及后疫情时代的大变局下，在全球政治经济形势高速变化的新环境面前，在决策过程所面临的噪声越来越多的现实之中，一个新的认知体系显得尤为重要。

朱 宁

上海交通大学上海高级金融学院金融学教授，副院长

现代经济学中的一个重要假设是“人是理性的，可以有效地判断信息”。但丹尼尔·卡尼曼和他的合作者指出，这个假设在很多情况下和事实并不相符，因此众多的理论结果都需要重新审视。这是对社会科学的重大贡献。人的判断错误包括两个部分：偏差和噪声。卡尼曼对偏差的分析已经广为人知，而本书对噪声的分析完成了对错误的全方位阐述。任何希望获得更高程度理性的个人或机构，都应该仔细阅读卡尼曼的这本新书。

刘 劲
长江商学院副院长

人生经历就是所有抉择的总和，我们无时无刻不在抉择，工作与生活在一次次的抉择中潜移默化地发生了变化。没有人可以永远做出正确的抉择，总有噪声在影响判断力。

噪声是不可消除的，但我们必须尽量将噪声的影响降至最低。《噪声》一书所提供的就是一套对于人生抉择大有裨益的减少噪声的方法论。

我要将《噪声》推荐给热爱生活的朋友。

毛大庆
博士，优客工场创始人

决策无处不在，而“噪声”与“信号”在其中穿插交错，互相影响。《噪声》一书道出了长久以来为人们所忽视的主观判断缺陷，人工智能虽然成了当今业界对高效率解决方案的新共识，但人工智能背后的“人”却依旧是无法回避的最大“噪声场景”。我们需要将个体复杂性充分融入组织决策考量之中，去探索商业与人的更合理关系。我们都期待

一个噪声更少的清朗世界。

吴 声
场景实验室创始人

自然界存在着尚未被感知的法则，而它们无形中影响着人们的决策。《噪声》这本书将“人类大脑”与“测量工具”同类对比，指出噪声的存在。这一重大发现将为企业的正确决策提供新的方法论，也将为企业家带来更深层次的思考。企业长期“与噪声共舞”，因此我们对本书中三位作者提出的新概念感到惊喜，并为作者对“噪声的形成”和“决策卫生”的一系列科学分析所折服。反观漫步者走过的25载，我们自上而下秉持着“敬畏、尊重、谦卑和良善”的管理理念，居然无形中用到了《噪声》提及的方法论，因而在极大程度上降低了噪声带来的内耗，为企业的持续发展赋能。相信这本由诺贝尔经济学奖得主创作的、直指“人类判断缺陷”的科学论著是普遍适用的，也必然能启迪你的智慧，为“人与世界的交互”注入新的创造力。

张文东
漫步者董事长、总裁

知识精英往往能在智力选拔中建立起优越感，斡旋于名利场中而获得成就感，但他们能否因此逼近理性呢？我们先贤提出的“君子时中”，可能是对信息决策中的偏差和噪声的一种敬畏吧。

归 江
信璞投资合伙人

很多大型企业和大型机构都强调，各个基层单位、各个员工的判断和决策，都要跟核心管理层制定的原则和准则保持高度一致。它们以为这样就可以做出好的判断和决策，但往往事与愿违。这就如同一支演奏水平低劣的交响乐团，面对指挥，各个乐手演奏出来的声音杂乱无章，

很不一致，很不和谐。对这种很不一致的错杂，应如何定义、如何测量、如何系统性地减少？卡尼曼的新书《噪声》第一次清晰系统地论述了这些问题。深入领会，学习应用，则基层与高层能更为一致。这样做会大大提高企业乃至整个社会的公平和效率。对于企业管理层而言，此书必备，此书必读，此书必用。

刘建位

央视十集《学习巴菲特》节目主讲人，《巴菲特选股10招》等书作者

人们担忧“算法统治世界”，其实是担忧暗藏在庞杂算法背后的“噪声污染”，但这种想法其实是倒因为果，因为这些噪声恰恰源于人的认知局限和判断缺陷。如何创造一个噪声更少的世界，是卡尼曼这本书关注的焦点。《噪声》是一本值得从头细读到尾的著作，对那些想提升认知水平、决策品质的组织与个人来说，“决策卫生”这个词一定会击中你的心灵。

段永朝

财讯传媒集团首席战略官，苇草智酷创始合伙人

面对环境、信息等多重不确定条件，如何审慎分析判断并做出恰当决策？卡尼曼教授的新作《噪声》提供了全新的方法论，让我们在警惕主观偏差的同时，学会发现并剔除系统噪声。

何 刚

《财经》杂志主编，《哈佛商业评论》中文版主编
《巴伦周刊》中文版联合创始人

人生是由大大小小的判断组成的。判断即人生。判断不仅会受偏差的影响，还会受噪声的影响。如果说偏差造成了“决策的不全面性”，那么噪声则造成了“决策的不一致性”。噪声和人的个性、偏好以及外部环境的变化都有关系，它往往随机出现，但影响巨大。《噪声》的启示

是，我们会和噪声相伴，但也有方法重塑决策框架，让自己做出更好的决策。

秦 朔
中国商业文明研究中心发起人

《噪声》是一本很特别的书，作者从心理学的角度出发，通过各种生动的案例和实验，把“噪声”这一不同于“偏差”的概念完整清晰地呈现给我们。三位作者抽丝剥茧地分析了当我们在各个领域做出判断时噪声可能带来的影响，以及如何减少其中的不利影响。希望自己做出正确决策的人都应该读一读这本书。

张 延
FT中文网出版人

《噪声》是一本将心理学研究应用于提升科学决策的跨界之书。噪声之所以存在，是因为人是情绪动物：决策者对洞见未来过度自信，不同人对同一情况的感知大相径庭，外部环境对心境的影响也会给决策者带来潜移默化的作用。关于如何减少噪声，书中给出了一系列操作性很强的决策卫生策略，同时敦促决策者保持开放的心态。当然，噪声也不是一无是处，在一些决策中，留下一定的自由裁量空间既能鼓励人们的创造性，也能使其更好地应对涌现出的新情况。

吴 晨
《经济学人·商论》执行总编辑

《噪声》的影响是震撼性的，因为它探究了人类判断中的一种具有根本性影响但被严重低估的危险。《噪声》是一本必读书，它提供了减少决策错误的可行方法。

罗伯特·西奥迪尼（**Robert Cialdini**）

畅销书《影响力》作者

地球已经被充分地探索过，科学家们不可能再发现一种以前未知的与大象体型相当的哺乳动物。而在决策领域，卡尼曼、西博尼和桑斯坦却发现了一个和“大象”一样大的问题：噪声。在这本重要的书中，他们向我们展示了为什么噪声很重要，为什么噪声比我们意识到的要多得多，以及如何减少噪声。遵照他们的建议将给我们带来能获得更多利润的企业、更健康的公民、更公平的法律体系和更幸福的生活。

乔纳森·海特（**Jonathan Haidt**）

坦普尔顿积极心理学奖获得者
积极心理学先锋派领袖，畅销书《正义之心》作者

在《噪声》一书中，三位作者巧妙地将他们关于人类判断缺陷的新颖见解，运用到人类在各个领域的判断问题中。做出判断和决策的人范围甚广，从点球教练到央行行长、军事指挥官，乃至国家元首，均在此列。《噪声》是心理学领域的一项了不起的成就，也是一座里程碑。

菲利普·E.泰特洛克（**Philip E. Tetlock**）

畅销书《超预测》作者

《思考，快与慢》《助推》《噪声》可以看成是“决策三部曲”。它们共同强调了所有领导者需要知道的东西，以帮助改善他们的决策，更重要的是改善整个组织的决策。《噪声》揭示了一种改善决策的关键杠杆，这在许多现有的行为经济学图书与文献中都未被提及。在噪声破坏你组织中的更多决策之前，我建议你尽快阅读《噪声》一书。

马克斯·巴泽曼（**Max Bazerman**）

畅销书《信息背后的信息》作者

杰作！卡尼曼、西博尼和桑斯坦将《噪声》引入生活，令人信服地

解释了为什么我们应该像对待偏差一样对待人类判断中的随机变化，并为减少判断中的噪声（和偏差）提供了实用的解决方案。

安妮·杜克（**Annie Duke**）

畅销书《对赌》作者

决策很关键。可惜，人们做出的许多决策都因噪声的存在而存在根本缺陷。因此，本书探讨的主题绝对是引人注目且必要的。本书以深入研究为基础，深思熟虑，而且易于理解。我带着好奇打开这本书，看完的时候觉得非常庆幸。因为这本书可以帮助我们在商业、政治和个人生活中做出更好的选择，照亮我们前行的道路。

丽塔·麦克格兰斯（**Rita McGrath**）

畅销书《瞬时竞争力》作者

《噪声》可能是十多年来我读过的最重要的一本书。它代表了一个真正的新想法，它很重要，而且你应该立即付诸实践。真是一本杰作！

安杰拉·达克沃思（**Angela Duckworth**）

麦克阿瑟天才奖获得者，畅销书《坚毅》作者

行为科学书籍的四个黄金标准是：见解新颖、证据严谨、引人入胜、能实际应用。很少有一本书能同时达到其中两个标准，但《噪声》却集这四个标准于一身。准备好接受世界上最伟大的思想家们带来的震撼吧，他们将帮助你重新思考如何评价他人、如何做决定，以及如何解决问题。

亚当·格兰特（**Adam Grant**）

《重新思考》作者

如何做一个聪明的决策者？



想知道你的判断
是怎么掉入噪声陷阱的吗？

扫码测一测，
立即获取答案及解析，
看看你的“降噪等级”。

1. 卡尼曼是世界上第一个凭借心理学研究获得诺贝尔经济学奖的人。
这个说法对吗？
 - A. 对
 - B. 不对

2. 卡尼曼指出：人类判断出错的原因有两种。一种是偏差，另一种是什么呢？
 - A. 误差
 - B. 噪声
 - C. 系统认知

3. 在做一个判断时，以下哪种方法出错的可能性最小？

- A. 使用算法模型
 - B. 进行头脑风暴
 - C. 追随明智判断者
4. 假设一位法官对于大部分案件都判得比较轻，但是对于某类案件，比如诈骗案，却总是判得很重。那么，下次他审理诈骗类案件的时候，你能准确预测判决结果吗？
- A. 能
 - B. 不能
5. 假设你看好一只股票，但是不确定现在的股价是不是太高了，想要知道在哪个价位买入更好。那么，以下哪种做法更有可能提高你的收益？
- A. 自己先做一个判断，然后假设这个判断是错的，试着考虑更多因素，重新判断，然后取两次的平均值
 - B. 问几个你信任的专业投资人，对他们的独立判断取平均值
 - C. 自己先做一个判断，过3个星期之后再判断一次，然后取两次的平均值
6. 企业在做员工绩效考核的时候，以下哪种做法不仅不能产生好效果，反而会引起巨大的问题？
- A. 强制性末位淘汰
 - B. 360度反馈制度
 - C. 结构化指标考核

在无法回避噪声的世界， 更好地追求高级境界与极致效益

彭凯平 清华大学社会科学学院院长

每年七八月份，我都会参加清华大学的本科生招生工作，由此而目睹了一个问题：高考揭榜之后，高中毕业生及其家长应如何做好志愿填报？学校和专业的选择非常重要，很有可能会决定一个学生一辈子的命运，其中牵涉的问题就与我们心理学中一个重要的研究领域紧密相关，那就是决策与判断心理学。而这一领域的集大成者，就是我要推荐的这本书的作者之一、著名心理学家和诺贝尔经济学奖得主丹尼尔·卡尼曼。

卡尼曼教授曾经是我在加州大学伯克利分校心理学系任教时的同事，他与斯坦福大学心理学系教授阿莫斯·特沃斯基（Amos Tversky）因“发现人在判断和决策上的系统性误差”而获得2002年诺贝尔经济学奖。1961年，卡尼曼教授毕业于美国加州大学伯克利分校，获得心理学博士学位。他曾任教于希伯来大学、不列颠哥伦比亚大学和加州大学伯克利分校，退休后去了普林斯顿大学，荣获伍德罗·威尔逊公众及国际事务学院荣誉教授称号。他的合作伙伴特沃斯基于1965年获得美国密歇根大学心理学博士学位，后在斯坦福大学任心理学教授。从20世纪60年代末开始，他们开创了一个心理学与经济学结合的领域——行为经济学，促使大批心理学家和行为经济学家（包括本书的另外两位作者西博尼和桑斯坦）开始研究人类判断出现错误的原因和补救措施，包括为什么人会过度自信，为什么人会有成见和确认性偏差（为什么我们总是寻

求、记住和过分强调支持我们信念的信息，而不接受反对我们但实则正确的信息）。

人类决策的心理动因一直是心理学界最为关注且开展过最多研究的核心领域之一。理性与非理性、主观意愿与客观影响，哪些才是影响决策品质的关键因素，这是心理学家们颇感兴趣的话题。从习惯上说，特别是在亚当·斯密的“看不见的手”这一理论大行其道的情况下，人们更愿意接受决策的理性效能，即“理性人假设”这个观念。然而，以卡尼曼为代表的行为经济学派的大量研究直接挑战了“人类决策是理性的”这种“正统经济学”的观念。2011年，卡尼曼教授出版了《思考，快与慢》，总结了这方面的大量卓有成效的研究成果。《思考，快与慢》一经上市，立即引起了全球广大读者的阅读热潮，也奠定了他在人类决策与判断领域以及行为经济学领域的开创者地位。

而我要推荐的这本新书《噪声》，与我刚开始提到的高考之后的志愿填报问题密切相关。学生们在选择学校和专业时，面对的是越来越多的可选项与决策建议。诚然，每一个提供建议的人的出发点都是好的，也都是本着负责任的态度，这无可厚非；但问题是，当把这些建议放在一起的时候，你会发现，这些建议本身就不一致，真应了那句老百姓常说的“萝卜白菜，各有所爱”。比如，张老师建议你学计算机科学，李阿姨建议你学财经，王老师建议你学心理学。假设你最终听从了王老师的建议，选择了学习心理学，但你还是会不停地想其他人的建议是不是更好，并因此感到困扰。

无须多言，我们在日常生活中也经常会遇到类似的情况：选哪只股票，打不打疫苗，选择哪一位追求者作为人生伴侣……在选择过程中，你会面对很多你并不想要或者令你判断不清的信息、建议、想法、观点等，现在我们知道了，所有那些不同的信息、建议、想法和观点，都有可能成为噪声。噪声极大地影响了我们决策的品质，也打破了我们内心的平静。它无时无刻不存在于我们身边，充满诱惑，甚至造成致命危

害。

很多时候，这些噪声看上去像是无关紧要的背景信息，于是被我们自己，乃至大量专业学者所忽视。但它没有逃过心理学家丹尼尔·卡尼曼、前麦肯锡合伙人和管理学教授奥利维耶·西博尼，以及法学家和行为经济学家卡斯·R.桑斯坦的眼睛。他们决定将噪声作为影响人类判断的一个重大问题来认真讨论。于是，我们有幸看到他们的讨论成果，即《噪声》一书的问世。这本书就是对噪声这一人类判断陷阱进行的一次系统而深入的剖析。在这本书中，作者融合了故事、研究和统计数据，提出了一个令人信服的观点，即噪声也会造成与偏差同等严重的损害，例如，破坏公平和正义，浪费时间和金钱，损害身心健康，造成生命价值与意义等方面的缺陷等。

在《噪声》这本新书中，作者探讨的是一个不同于偏差的判断错误问题。他们首先提出，错误是一种系统误差，通常出现在个人判断中。通过比较一个人的反应与理性标准的差异，我们可以发现他的判断中的错误有多严重。但是，我们不能以特定的判断来识别噪声。噪声是随机分布的，我们必须在研究一系列判断后通过统计的方法才能确定。

卡尼曼在书中介绍了商业、医学和司法领域中的几个令人信服的案例，其中的指纹鉴定案例让我印象深刻。同一个鉴定人员在不同时间居然对同一枚指纹给出了不同的鉴定结果。如果鉴定人员只有指纹可以查看，没有关于该案例的其他信息，那他每一次判断基本是一致的。但研究发现，如果不仅有指纹，还有其他信息，那么每一次鉴定结果就可能有所不同。这个案例的启发是：我们在制定决策时，维护至关重要的信息、排除非关键信息非常必要。

同时，卡尼曼的书还讨论了许多不同类型的噪声，不过最重要的讨论还是关于系统噪声的。卡尼曼认为：在很多情况下，在一个系统里，意见的多样性非常有必要，但噪声是你不希望存在的不一致和变异。卡尼曼和他的同事展示了评估和选择中不必要的分歧是如何形成“充满噪

声的系统”的。它们困扰着包括刑事法官、理赔员、司法鉴定专家、未来学家和医生在内的专业人士，他们经常对同一案例做出截然不同的判断和决定。

系统产生噪声的部分原因是，不同专业人员应用的标准不同。例如，有令人不安的证据表明，当多个医生评估同一位患者患心脏病、结核病、子宫内膜异位症、皮肤癌和乳腺癌的风险时，他们的诊断的不一致程度很高。可见，在如此嘈杂的系统中，错误只会累加而不会相互抵消。正如书中所说：“如果两个罪犯都应被判处5年监禁，但一个被判3年，另一个被判7年，虽然平均刑期是5年，但正义并没有得到伸张。”

系统之所以产生噪声，还因为随着时间的推移，相同的专业人员在不同的时间有可能应用不一致的标准。为了说明这一点，一项针对22名医生的研究发现，他们与自己原有意见不一致的可能性为63%~92%。他们每人在相隔几个月的时间里检查了同样的13张血管造影图。为了解释这种波动，作者使用“情境噪声”（occasion noise）来阐述研究结果，比如，天气与大学的录取决定理应是无关的，但是研究发现，情况并非如此。有心理学家研究了大学招生办公室的工作人员的682例录取决定，就发现：在阴天，他们更看重申请人的学术潜能；在阳光灿烂的日子，他们更看重申请人的非学术潜能。

《噪声》一书还进一步挖掘了噪声的其他特性，包括其原因和成分、如何测量，以及它和偏差之间的相互作用。作者解释了为什么讨论小组（与单一决策者相比）可能放大噪声，以及指南、规则和算法如何减少噪声。他们提供了一套工具，以帮助决策者识别和减少系统噪声。3位作者用28章的篇幅来阐述关于噪声的研究成果，外加3个附录，帮助读者进行噪声审查并通过制定清单来改进小组决策和预测。特别令我高兴的是，书中谈到了那些能够抑制系统噪声的人具有的特征：他们往往思维更加积极开放，会不断寻找和学习新的信息，并不断更新自己的信念和看法。这恰恰就是我经常提到的具有积极开放性思维的人。

在《噪声》一书中，三位作者还特别建议，为了减少决策过程中的噪声，最好先要求多人做出独立判断，然后再把他们聚集在一起解决分歧。另外，他们还解释了对直觉和特殊偏好进行限制的指导准则，并强调，长期以来人们都知道如何用这些准则减少偏差，因此也足够睿智，能够知道如何用这些准则减少噪声。为此，他们提供了一个关于招聘的有趣案例，来敦促组织使用结构化而不是非结构化的方式来面试员工。事实上，大多数面试官都喜欢随意地向求职者提出他们本身最喜欢的问题。但强有力的证据表明，当多个面试官每人按相同的顺序提出相同的问题时，他们所选定的应聘者的工作绩效会更好。

综合各项研究，卡尼曼在书中建议：组织可以通过任命一名“决策观察者”来解决噪声和偏差问题。这个观察员的职责是负责跟踪和指导互动的领导者或专家。但卡尼曼不建议决策观察者用冗长的问题清单的方式不厌其烦地进行问卷调查，而建议以观察与记录为主，因为调查问卷并不足以帮助这位决策观察者或其他任何人诊断出组织是否因噪声而破坏了自身的决策。引入决策观察者这种解决方案在认知失调的群体中是行不通的，因为在那里，直言不讳是不安全的；但它可以帮助那些决心做出正确判断的健康团队做出更加正确的决策。

由此看来，知道得越多，未必越好；想法越多，未必越好；选择越多，也未必越好。因为，在噪声的影响下，它们反而可能成为人们做出正确判断和决策的障碍。所以“听人劝，吃饱饭”有可能是错的！

总之，阅读卡尼曼等人的新作让我兴致盎然、心潮澎湃。可以肯定地说，通过理解并实践卡尼曼等人的思想与建议，我们能知道如何减少可能会困扰组织和个人的系统噪声，知道“决策信息的简约化”居然也如此意义重大！并且，更为重要的是，我们能够学习在噪声中运用科学知识、深刻的见解和坚实的原则，做出更为人道、更为公平的决策，减少时间、金钱和人才的浪费，拯救生命，传递善意，分享意义。而这些，都是一个真实而积极的生命所追求的高级境界。我也期待着《噪声》这

本行为决策领域的新作能够像《思考，快与慢》一样让你爱不释手，并且持续为你的美好生活提供力量。

穿越噪声的决策

朱宁 上海交通大学上海高级金融学院金融学教授，副院长

我抱着浓厚的兴趣读完了诺贝尔经济学奖得主丹尼尔·卡尼曼教授等的新作《噪声》一书，这本书汇集了卡尼曼教授近十年的最新发现。对于我所从事的行为金融学研究领域而言，《噪声》算得上是一本“专业书”，但是读完之后，我认为它很可能是卡尼曼教授继《思考，快与慢》后，又一次成功将大量、全新、重大的研究成果，以平实通俗的语言向广大读者分享。在此，我要向大家强烈推荐《噪声》一书。

噪声，“地下室漏水”的大问题

说来有趣，在我所研究的金融领域，也曾有一篇非常著名的论文《噪声》。这一论文的作者，是著名金融经济学家费希尔·布莱克（Fisher Black）。布莱克教授的论文讨论的是，股票市场中大量的个人投资者，也就是国内统称的“散户”，没有意识到自己的错误决定可能会对金融市场产生的影响。这篇论文对金融学研究，特别是行为金融学的发展做出了重要贡献。

无独有偶，我本人也曾发表过一篇名为《系统性的噪声》的论文，该论文讨论了个人在投资中所犯的错误，虽然这些错误数量众多、类型各异，但是如果进行系统性的分析，还是能发现其中的规律。

其实，无论是布莱克教授的《噪声》，还是我的《系统性的噪声》，这两篇论文想传递的信息，都与卡尼曼在《噪声》一书中想要传

递的信息高度一致：希望不管是投资界还是管理界的决策者，都能时刻记住决策过程是复杂和困难的，做决策的人应该紧绷这根弦。但卡尼曼教授的贡献在于，他将“噪声”这一重大发现系统性、结构化地呈现了出来，让这个影响决策的“地下室漏水”问题真正暴露在公众的视野中。正如他在书中所说：“噪声就像地下室漏水，它之所以能被容忍，不是因为人们认为它是可接受的，而是因为它一直未被发现。”在行文中，我们能感受到卡尼曼希望帮助投资者和管理者清醒地意识到，决策过程中存在着很多由不确定性和信息不完备导致的噪声，而这些噪声有可能引发致命的错误。而《噪声》一书的独特贡献，可能恰恰就在于——它不仅提出了关于人类判断错误的一个堪比“黑洞”的大问题“噪声”，而且系统性地研究了企业管理、医疗卫生、教育文化、刑事司法等领域，得出了噪声无处不在且影响巨大的重要结论。当然，这个问题并不是无解的，人们可以通过有意识地关注这些噪声背后的规律，以及调整自身的决策行为，有效地改善决策品质和提升管理水平。

噪声既然会产生系统问题，就需要系统解决

很多人读过卡尼曼的前作《思考，快与慢》，我也是他的忠实读者。时隔10年，卡尼曼再出新作。通篇阅读下来，我发现《噪声》这部作品不仅是前一本书的延续，而且实践性更强。这其实也体现了卡尼曼对于解决决策问题的不懈探寻以及雄心壮志。

《思考，快与慢》一书汇集了卡尼曼教授之前三四十年行为经济学研究之大成，堪称行为决策领域的百科全书。与此不同的是，《噪声》一书是卡尼曼教授与另外两位分别来自管理学、法学领域的教授通力合作的成果。他们三位共同探讨了行为科学将如何帮助决策者意识到并尽可能规避决策过程中的随机因素（噪声），也就是决策中不必要的变异，最终做出更为准确和科学的决策。

《思考，快与慢》立足于心理学和行为科学的学术研究，更多关注

的是人类在决策过程中普遍存在的一些行为倾向，以及这些倾向背后深层次的生理和心理因素。而《噪声》一书，则更多立足于管理科学和组织行为学，针对管理者在管理决策，特别是企业管理决策过程中，面对信息不完备和环境不确定的情况，在决策中可能产生的偏差和噪声。

从某种意义上说，《噪声》一书比《思考，快与慢》更加雄心勃勃。如果说《思考，快与慢》的主要目标是展现和揭示人类行为决策中普遍存在的大量非理性行为及其产生原因的话，那么《噪声》一书则是希望帮助读者意识、识别、分析和减少噪声，从而削弱噪声对完整系统的影响。

穿越噪声，让决策更快击中“靶心”

《噪声》这本书的核心内容，首先是希望帮助读者对噪声或者说决策过程中的随机因素，有一个清晰的认识和完整的了解。其次，它分析了这种决策过程中噪声的来源，及其可能对决策过程和品质产生的影响。但如果要说这本书最核心的价值，我觉得还是它帮助大家在意识到噪声存在的基础上，在熟悉和了解噪声的成因之后，能够有意识地克服和纠正自己在决策过程中会犯的错误，以改善和提升决策品质。

与《思考，快与慢》以及其他众多讨论管理决策过程中的问题的作品不同，《噪声》关注的是在环境不确定和信息不完备的情况下，所有人，特别是管理者在做决策时可能犯的错误和这些错误的来源。书中特别把“模式噪声”分解为“稳定的模式噪声”以及“情境噪声”这两种成分，并且讨论在面临诱发噪声的不同情境时，人怎样才能更好地避免在决策过程中犯各种各样的错误。卡尼曼指出：有一部分行为的错误，可能是由内在人格所驱动的，但有一部分错误，很可能是受外部环境影响所致。因而，决策者必须清醒地意识到自己可能会受到哪些方面因素的影响，并且有针对性地调整和纠正，力图达到减少噪声的目标。

《噪声》的三位作者希望能够帮助人们更好地超越人在认知过程中所犯的错误，能够更加准确地了解事实、了解这个世界，继而做出更加科学和准确的抉择。

在上述意义上，《噪声》看似和一些著作殊途同归，但是从另一个层面上讲，《噪声》又是一种超越，因为《噪声》提供了一个正确的认知体系。对于日益富足且对未来颇有期待的大众和企业家而言，在后金融危机时代以及后疫情时代的大变局下，在全球政治经济形势高速变化的新环境面前，在决策过程所面临的噪声越来越多的现实之中，一个新的认知体系显得尤为重要。

这，可能也是《噪声》一书最大的价值所在。

是为序。

哪里有判断，哪里就有噪声

我们非常荣幸，也非常感激，能够在中国出版这本书。《噪声》重点关注人类的判断，关注它是如何出错的，以及如何让它变得更好。无论是医学、法律、公共政策、商业，还是日常生活领域中，我们做出判断的过程中存在的问题，是偏好根据经验做判断的人类的共性问题。

数十年来，偏差问题备受瞩目。在《思考，快与慢》一书中，本书的作者之一卡尼曼，讲述了人类思维所能取得的惊人成就，同时也讲述了思维有时会产生错误——系统性误差，它被称为偏差。如果人们过度乐观，他们的判断会有偏差（也就是乐观偏差）。同理，如果他们过度关注短期而忽视长期目标，他们的判断也会有偏差（也就是即刻偏差）。许多人都会表现出乐观偏差和即刻偏差，尽管这些偏差可能是有益的（如果你很乐观，你可能更愿意努力尝试），但它们也可能造成严重的错误，进而导致严重的问题。

我们在本书中讲述了许多有关偏差以及如何减少偏差的内容，但本书的主要议题是另一种错误成分：噪声。我们将噪声定义为：判断中不必要存在的变异。偏差是平均的、共有的误差，噪声则是这些误差的变异。如果一位医生说某位病人有心脏病，而另一位医生说该病人只是压力过大，那么噪声就产生了。如果一位安检员说某一车间是安全的，而另一位安检员说该车间有危险，那么噪声也会出现。如果一位法官说被告应被判2周监禁，而另一位法官认为被告应被判2年监禁，那么噪声就非常显而易见了。

当然，噪声的存在本身并不令人惊讶。如果我们说某个问题涉及判断，就意味着我们允许存在分歧，甚至期望分歧的存在。与事实性问题

不同，对于需要做出判断的问题，理性的人可以存在不一致。激励我们写下本书的原因并不是噪声的存在令人惊讶，而是噪声之大令人惊讶。无论考察哪个领域——司法、商业、医学、法医学、公共政策等，我们都发现，噪声比我们想象的要大。更重要的是，噪声比这些领域的决策者所认为的还要大，而且往往大到超出他们所能容忍的范围。

噪声无处不在，这是我们写下本书的原因之一。一位心理学家、一位商学院教授和一位法学学者为一个图书项目在一起合作数年之久，这是很少见的。但是，完成如此庞大而又包罗万象的图书项目所需要的综合技能，超出了我们任何一个人的能力。我们的互补，为《噪声》的故事提供了广阔的视角，我们希望你能从中受益。

在许多方面，噪声仍是一个“未知的国度”，一个“有待探索的世界”。如果未受到重视，它将产生严重的问题，会导致严重的不公平。在某些情境中，它甚至应该被视作一种暴行或丑闻。而在大多数情境中，噪声会造成高昂的代价，损害企业、员工、消费者、投资者和许多其他人的利益。

不过也有好消息。一旦意识到了噪声的存在，你就可以减少噪声，并减小其危害。此外，你为“降噪”所做的工作往往能同时从根源上消除一些偏差问题。我们呈现了一系列案例研究的结果，展示了如何通过实施一系列技术性措施来测量和减少噪声，我们称这些措施为决策卫生策略。我们还重点介绍了“降噪”中你可能面临的挑战，以及有时需要做出的利弊权衡。

噪声是一个普遍性问题，任何做专业判断的人都应该关注它。我们撰写的《噪声》这本书，是为了献给所有参与决策并希望决策变得更好、更公平、更准确的人，包括企业领导者、政策制定者、法官、医生、人力资源主管、咨询顾问、研究人员和各个领域的专业人士。判断是人生阅历的核心组成部分，也是所有组织运作的中枢要点。

就全世界范围而言，在通过减少噪声降低成本、提高公平性方面，

我们可以做很多工作。无论你身在何处，无论你使用何种语言，情况皆是如此。

目 录

[重磅赞誉](#)

[测一测 如何做一个聪明的决策者？](#)

[推荐序1 在无法回避噪声的世界，更好地追求高级境界与极致效益](#)

彭凯平 清华大学社会科学学院院长

[推荐序2 穿越噪声的决策](#)

朱宁 上海交通大学上海高级金融学院金融学教授，副院长

[中文版序 哪里有判断，哪里就有噪声](#)

[引言 偏差与噪声，人类判断的两类错误](#)

[第一部分 寻找噪声](#)

[第1章 犯罪和充满噪声的判罚](#)

[第2章 系统噪声，给人达成一致的错觉](#)

[第3章 单一决策，仅发生一次的重复决策](#)

[第二部分 你的大脑是一种测量工具](#)

[第4章 什么是判断](#)

[第5章 测量误差，噪声与偏差的代价一样大](#)

[第6章 噪声分析：所有判断都存在3类噪声](#)

[第7章 情境噪声，无时无刻不在影响着我们的判断](#)

[第8章 群体是如何放大噪声的](#)

[第三部分 预测性判断中的噪声](#)

[第9章 判断与模型，简单的模型普遍优于人类判断](#)

[第10章 无噪声的规则](#)

[第11章 哪里有预测，哪里就有客观无知](#)

[第12章 常态谷：事情虽无法预测，但可以被理解](#)

[第四部分 噪声是如何产生的](#)

[第13章 启发式、偏差与噪声](#)

[第14章 匹配，找到与你的预测最精准匹配的共识](#)

[第15章 选取精确的量表，并多用相对判断](#)

[第16章 模式噪声的构成](#)

[第17章 噪声源，偏差是引人注目的图形，而噪声是不受我们关注的背景](#)

[第五部分 决策卫生，提升五大人类判断力](#)

[第18章 卓越的判断者，卓越的判断力](#)

[第19章 消除偏差与决策卫生](#)

[第20章 司法科学，信息排序是最大的噪声](#)

[第21章 甄选与汇总，超级预测的两大策略](#)

[第22章 医疗决策，用科学的诊断指南减少噪声](#)

[第23章 绩效评估，用基于外部视角的共识框架做出量化判断](#)

[第24章 人员招聘，以结构化指标衡量人才](#)

[第25章 中介评估法，做出明智决策的核心方法](#)

[第六部分 最佳的噪声水平](#)

[第26章 减少噪声的成本](#)

[第27章 尊严，人之为人的重要价值观](#)

[第28章 规则还是标准](#)

[回顾与总结 正视噪声问题](#)

[结语 一个噪声很少的世界](#)

[附录1 如何进行噪声审查](#)

[附录2 决策观察者检查清单](#)

[附录3 对抗噪声，修正预测](#)

[致谢](#)

[译者后记](#)

[注释](#)

偏差与噪声，人类判断的两类错误

试想一下，你的朋友组成了A、B、C、D共4支队伍，来到一个射击场。每队中有5个人，他们共用一支来复枪，且每人只开了一枪。图0-1显示了他们的射击结果。

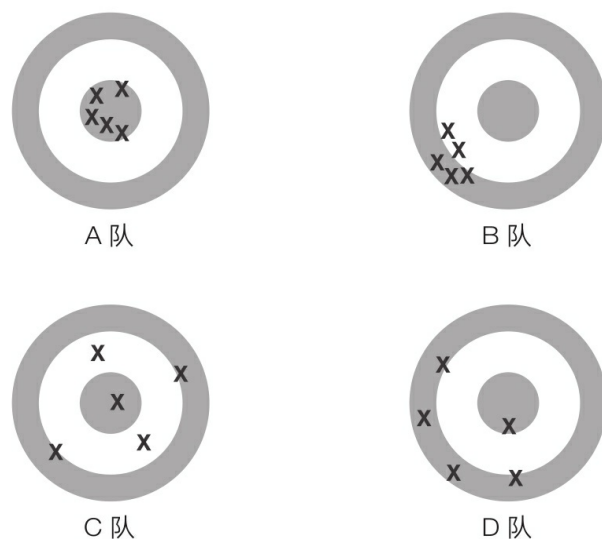


图0-1 4支队伍的射击结果

理想情况是，每一枪都能正中靶心。

A队几乎达到了理想情况，他们的每一发子弹都紧紧围绕着靶心，接近完美模式。

B队的每一发子弹都偏离了靶心，我们可以称其为偏差队。在图0-1中，我们可以根据B队偏差的一致性进行这样的预测：如果该队中的某位成员再开一枪，我们敢说子弹的落点也会与前5次落点的区域相同。偏差的一致性也许有原因可循：B队使用的来复枪的瞄准器歪了。

C队的子弹落点很分散，我们可以称其为噪声队。子弹的落点大致都在靶心四周，因而没有明显的偏差。如果该队的某位成员再开一枪，我们很难准确预测他可能击中的位置，而且我们也无法从C队的结果想出任何有趣的假设。我们只知道C队的成员不太擅长射击，但确实不知道为什么子弹的落点如此分散，会充斥着如此多的噪声。

D队是偏差与噪声共存队。与B队类似，D队的落点基本上都偏离了靶心；与C队的相似之处在于，D队的落点也很分散。

当然，这不是一本教射击的书，我们的主题是人类判断的错误。偏差和噪声，即系统性偏差和随机分散，是错误的不同组成成分，举射击的例子，只是为了阐明两者的差异。

[射击场只是一种隐喻](#)，用来说明人们在做判断时会出现什么样的错误，尤其是在代表组织成员做出各种判断时。在这些情境中，我们会发现图0-1展示的两类错误。有些判断存在偏差，它们整体偏离了目标；有些判断存在噪声，我们期望人们就某个目标达成一致，到头来他们却产生了严重的分歧。很遗憾，很多组织同时受到了偏差和噪声的困扰。

图0-2说明的是偏差和噪声之间的重要区别。如果隐去靶子看各队的子弹落点情况，那么最后每个队的射击结果就会如图0-2所示。

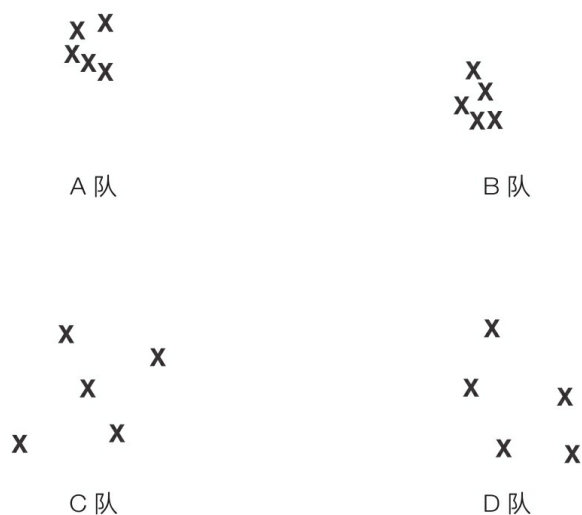


图0-2 隐去靶子后的子弹落点情况

此时，你无法分辨是A队还是B队的子弹落点更接近靶心，但你能一眼看出，相比于这两队，C队和D队存在更多的噪声。事实上，你对各队射击结果分散程度的了解与你观察图0-1时一样多。噪声的一个普遍特性是：你可以在对目标或偏差一无所知的情况下，识别噪声，并对它进行测量。

方才提到的噪声的普遍特性，对本书意义重大，因为在本书中，我们的很多结论，也都是通过研究人们在“真实答案”未知甚至永远无法获知的情况下做出的判断而得出的。当不同医生为同一位患者做出了不同的诊断时，我们可以仅研究他们的分歧点，而无须知道患者的病情；当电影公司主管评估一部电影的市场效益时，我们可以仅研究他们意见的差异性，而不必知道这部电影最终的票房，甚至不必知道它有没有拍摄完。也就是说，在衡量这些差异时，我们无须知晓谁对谁错。要测量噪声，我们只需看“靶子的背面”。

要想理解判断中的错误，我们必须同时理解偏差和噪声。有时候，噪声甚至是更为重要的问题。然而，在有关人类错误的公开讨论，以及世界各地的组织中，很少有人认识到噪声的重要性。偏差是台上的主角，而噪声只是幕后的一个不起眼的参与者。已经有成千上万篇科研文章和数十本畅销书讨论过偏差的话题，但讨论噪声问题的作品却少得可怜。本书的写作目的就是修正这一失衡的状态。

在现实的决策中，噪声的数量往往令人触目惊心。下面几个例子展示了在准确性非常重要的场景中，噪声存在的数量有多大：

- 医学诊断中存在噪声。不同医生对同一患者是否患有皮肤癌、乳腺癌、心脏病、肺结核、肺炎、抑郁症等疾病，会做出不同的判断。精神科诊断中的噪声尤其多，显然是因为精神科医生的主观判断对诊断结果起决定性作用。然而，在一些并不应该

存在噪声的领域，例如在对X线片报告的解读中，也存在着大量噪声。

- [儿童监护权判定中存在噪声](#)。儿童保护机构中的案件负责人需要评估儿童是否存在受虐待的风险，如果存在，则需要进一步评估是否需要将他们送去寄养。鉴于有些案例的负责人比其他负责人更有可能做出将儿童送去寄养的决策，所以该系统存在噪声。多年后，被某些过于严格的负责人送去寄养的不幸儿童，境遇大多很糟糕：犯罪率更高，青少年时期怀孕率更高，收入更低。
- [预测中存在噪声](#)。专业的预测人员对新产品的可能销量、失业率的可能增长、经营不善的公司破产的可能性，以及其他各类问题，都会做出分歧非常大的预测。他们不仅意见不一，而且各自的预测也前后矛盾。例如，当同一批软件开发人员被要求在不同的两天中分别评估完成同一任务所需的时间时，他们前后两次估计出的时间平均相差71%。
- [庇护权决策](#)中存在噪声。寻求庇护者能否被允许进入某个国家，就和买彩票差不多。一项研究发现，在将庇护申请随机指派给不同的法官后，一位法官批准了5%的申请，而另一位法官却批准了88%的申请。该研究的标题说明了一切：《难民轮盘赌》（*Refugee Roulette*）。（在后文中，我们将会看到许多“轮盘赌”。）
- 人事决策中存在噪声。不同面试官对相同应聘者的评估常常大相径庭。类似地，对相同员工的绩效评估，管理者之间也存在着很大差异。结果更多取决于评估者不同，而非被评估者的表现。
- 保释决策中存在噪声。被告是获准保释，还是移送监狱候审，部分取决于审理该案件的法官。有些法官更为宽容，有些则更

为严格。法官们对于哪些被告最可能逃逸或再犯的评估，也存在着显著的差异。

- 司法鉴定科学中存在噪声。在我们的认知范围内，指纹鉴定是绝对可靠的。然而，在判定犯罪现场的指纹与犯罪嫌疑人的指纹是否匹配时，指纹鉴定师们的意见有时也会不一致。不仅不同专家之间会产生分歧，即使是相同的专家，在不同场合看到相同的指纹时，也可能做出不同的判断。类似的差异性，在其他司法鉴定领域，甚至是DNA分析中也同样存在。
- [专利权授予决策](#)中存在噪声。关于专利申请的一篇重要研究论文的作者强调了专利申请涉及的噪声：“专利局是授予还是拒绝一项专利申请，很大程度上取决于该项专利申请被分配给了哪一位审查员。”从公平的角度来看，这种差异性显然是有问题的。

上述所有存在噪声的例子不过是冰山一角。无论你观察哪个领域中人类的判断情况，都有可能发现噪声的身影。为了提高判断品质，我们不仅需要克服决策中的偏差，还必须克服噪声。

本书共分为6个部分。在第一部分中，我们探讨了噪声和偏差的区别，并展示了无论是在公共部门，还是在私人机构，噪声都可能存在，其数量甚至达到大得惊人的地步。为了理解这个问题，我们从两个领域的判断着手，第一个领域涉及刑事判决（公共部门），第二个领域涉及保险（私人机构）。乍一看，这两个领域差异巨大，但它们在噪声方面却有很多共同点。为了证明这一点，我们引入了“噪声审查”（noise audit）的概念，目的是衡量一个组织中不同的专业人员对相同的事件在处理意见上存在多大分歧。

在第二部分中，我们研究了人类判断的本质，并探索了如何衡量其准确性和错误数量。判断容易受偏差和噪声的影响。我们将介绍，这两

类错误在造成的影响方面具有惊人的一致性。“情境噪声”是指同一个人或同一群体在不同场合中对同一事件做出判断而产生的差异性。我们还发现，群体讨论中存在很多看似无关的因素，比如发言次序，它们却导致了数量极其庞大的噪声。

在第三部分中，我们将深入探讨一种已得到广泛研究的判断类型——预测性判断（predictive judgment）。我们探索了规则、公式和算法在人类做出预测时的关键优势：与大众看法不同的是，与其说规则具有卓越的洞察力，还不如说规则是没有噪声的。我们讨论了影响预测性判断品质的终极局限，即对未来的“客观无知”（objective ignorance），以及这种客观无知如何与噪声一起对预测的品质产生了影响。最后，我们会探讨一个你肯定会问自己的问题：如果噪声无处不在，那为什么我之前没有注意到呢？

在第四部分中，我们转向人类心理学领域。我们解释了噪声产生的根本原因。这些原因既包括由各种因素引起的不同个体之间的差异，也包括个性和认知风格的差异、对不同因素进行加权时思考角度的差异，以及人们在使用相同的量表时出现的理解差异。我们探究了为什么人们会对噪声视而不见，以及他们为什么对不可能预测到的事件常常并不感到意外。

第五部分探讨了如何改进判断并防止出错这一实际问题。如果你主要对如何在实践中减少噪声感兴趣，那么你可以跳过第三部分和第四部分，直接阅读第五部分。我们调查了人们为降低医学、商业、教育、政府治理和其他领域中的噪声所做出的努力。我们介绍了利用“决策卫生”（decision hygiene）减少噪声的几种做法。我们展示了来自不同领域的5个案例研究。在这些领域中，关于噪声的记录已经有很多，人们为了减少噪声也做出了不懈的努力，并且取得了一定的成效。这些案例研究包括不太可靠的医学诊断、绩效评估、司法鉴定、招聘决策以及一般性预测。最后，我们介绍了中介评估法（mediating assessments

protocol, MAP) ——一种用于评估选项的通用方法，该方法融合了决策卫生在实践中的几种关键做法，目的是让噪声更少，从而得到更可靠的判断。

正常的噪声水平应该是多少呢？第六部分会回答这一问题。正常的噪声水平并不是零，这或许有点违反直觉。在一些领域中，噪声不可能被完全消除；在另外一些领域中，消除噪声的成本太高；还有一些领域中，消除噪声的努力会损害重要的竞争价值。例如，消除噪声的努力可能会影响士气，会让人觉得自己很像机器中的齿轮。当试图用算法解决问题时，人们会提出各种各样的反对意见，我们会对其中一些反对意见做出回应。尽管如此，当前的噪声水平还是令人难以接受。我们敦促私人机构和公共组织进行噪声审查，并以前所未有的严肃态度，付出更多的努力来尽力减少噪声。这样做可以减少很多组织中普遍存在的不公平现象，并降低诸多成本。

怀揣着这一愿景，我们在每一章的结尾都会提出一些简短的建议。你可以采用这些建议，也可以根据你的实际情况加以调整，无论这些问题是涉及健康、安全、教育、金钱、就业、娱乐，还是其他方面。了解因噪声而存在的问题，并试图解决这些问题，是一项需要持续去做的工作，也是一项需要我们共同努力的工作。我们所有人都有机会为这项工作做出贡献。这也正是我们写这本书的初衷。

NOISE

第一部分

寻找噪声

· 发现噪声

哪里有判断，哪里就有噪声

“同罪不同罚”令人难以接受。同样的罪行，有人被判处5年有期徒刑，有人却被判处缓刑。在许多地方，类似的事情正在发生。可以肯定的是，刑事司法制度中弥漫着偏差，但本书第1章关注的重点是噪声。一位著名的法官注意到了噪声的存在，他发现这种噪声是不道德的，并由此发起了一场在某种意义上改变世界（但还不够彻底）的运动。我们接下来要讲的故事发生在美国，但我们相信，类似的故事在其他国家也存在，甚至情况更为严重，并且它将来依然会出现。我们使用刑事判决的例子，也是想说明噪声会导致极大的不公平。

刑事判决过程充满了戏剧性，但我们也关注私人机构，因为私人机构也可能因噪声的存在而背负很大的隐形风险。为了阐明这一点，我们在第2章介绍了一家大型保险公司。在该公司，核保员的任务是为客户确定保费，而理赔员必须判定理赔的额度。你可能会猜想，这些任务简单而又机械，不同的专业人员会得出大致相同的数额。为了对这一猜想进行验证，我们精心设计了一个关于噪声审查的实验，实验结果不仅令我们惊讶，也让该公司的领导层惊讶和沮丧。据我们了解，数量众多的噪声使该公司损失了大量资金。我们使用这一案例是为了说明，噪声会造成巨大的经济损失。

这两个例子所涉及的研究都是大样本研究，也就是说，在此过程中，有很多人做出了大量判断。但是，许多重要的判断是单一的而非重复的，比方说：如何处理一个看上去独一无二的商机；是否发布一款全新产品；如何应对一场流行病；是否雇佣一个不太符合标准的人……这

类独特情境中的决策是否存在噪声？我们很容易认为，此类特殊情境中不存在噪声。毕竟，噪声是不必要的变异，单一决策怎么会有变异呢？在第3章中，我们会回答这个问题。我们做出的判断，即使在看似独一无二的情境中，也充满了各种各样的可能性，也存在着大量噪声。

这三章的主题可以用一句话来概括，这就是：哪里有判断，哪里就有噪声，而且它比你想象的还要多。一起来看看噪声有多少吧。

第1章

犯罪和充满噪声的判罚

假设某人被指控犯了罪，例如到商店行窃、私藏海洛因、袭击他人或持枪抢劫，该案件可能的判决结果是什么？

答案不应取决于该案件恰巧被指派给哪位法官、天气是冷是热，以及当地球队在前一天是输是赢。如果3个背景类似的人被指控犯有同一罪行，最后却得到截然不同的处罚，例如第一个人被判缓刑，第二个人被判2年有期徒刑，最后一个人被判10年有期徒刑，这样的结果显然会引起公愤。然而，在很多国家，这种不合理的情况时有发生——不仅在过去发生过，现在也依然可见。

长久以来，世界各地的法官对于案件的判决大都拥有自由裁量权。在许多国家，专家们对这种自由裁量权表示赞赏，认为它既公正又人道。他们坚持认为：刑事判决应基于多种因素，不仅要考虑罪行本身，还要考虑被告的性格及其所处的环境，“刑罚个别化”已成为当下主流；如果法官受到规章制度的约束，罪犯就可能受到不人道的对待，就不会被视为一个独特的个体，特殊情况也就无法被充分考量。在许多人看来，正当法律程序的理念似乎要求开放的司法自由裁量权。

20世纪70年代，大众对司法自由裁量权的热情开始瓦解，原因很简单：大量的证据表明，噪声无处不在。1973年，著名法官马文·弗兰克尔（Marvin Frankel）引发了公众对这一问题的关注。在成为法官之前，弗兰克尔不仅是言论自由的捍卫者，还是热情的人权倡导者，他帮助创立了人权律师委员会——一个号称“人权至上”（Human Rights First）的组织。

在人权捍卫方面，弗兰克尔有些激进，也因此他对刑事司法制度中

的噪声深感愤怒。以下是他对自己[帮助创立该组织的初衷](#)的描述：

如果美国联邦银行抢劫案的一个被告被定罪，最高会被判处25年有期徒刑，这意味着刑期是0至25年不等。我很快意识到，这个数字与其说取决于案件或被告本人，不如说取决于法官，即受到不同法官的观点、偏好和偏差的影响。因此，同一起案件、同一个被告，可能会因为审理法官的不同而导致判决结果截然不同。

弗兰克尔没有提供任何统计分析来支持他的论点，但他提供了一系列强有力的事例，来证明处境相似的人受到的对待存在严重不合理的差异。比如，两名男子均无犯罪记录，都因兑现假支票触犯法律，他们兑现的金额分别为58.4美元和35.2美元，第一个人被判处15年有期徒刑，而第二个人仅被判处30天监禁。对于两起类似的挪用公款案件，一名被告被判处117天监禁，而另一名被告被判处20年有期徒刑。因此，弗兰克尔对美国联邦法官“[几乎完全不受制衡的权力](#)”表示谴责，并认为这导致美国[每天都在发生着残酷专断的行为](#)。他认为，在[法治而非人治的社会](#)存在这种现象是令人难以接受的。

弗兰克尔呼吁美国国会结束这种残酷专断的“歧视”。他所说的歧视主要所指就是噪声，即量刑中存在的难以解释的差异性；同时，他也关注由种族和社会经济地位的差异所导致的偏差。为了消除噪声和偏差，弗兰克尔敦促改革，主张不应允许存在对刑事被告人量刑上的差异，除非这种差异可以“通过足够客观的测试来说明其合理性，以确保该结果不是特定官员、法官或其他人[独断专行的产物](#)”。不仅如此，弗兰克尔还主张通过制定“详细的影响因素清单”来减少噪声，这些清单“应尽可能包括[某种形式的数字，或其他客观的评分](#)”。

弗兰克尔在20世纪70年代初就写下了这样的话，因此不是在提倡“用机器取代人”这个主张。但当时他已经很接近这一目标了，这真令

人难以置信。他认为“法治需要一套在多种情况下都适用的客观的规则，从而可以对法官和其他人进行约束”。他明确主张[使用计算机作为量刑时有序思考的辅助工具](#)，此外，他还主张[成立一个量刑委员会](#)。

弗兰克尔的著作成为刑事司法史上颇具影响力的著作之一——不仅在美国，在全世界范围内皆是如此。他的著作有一个缺点：存在一定程度的非正式性，不够严谨。但他指出的事实令人错愕，也令人印象深刻。为了验证刑事判决中是否确实存在噪声，一些研究者紧随其后，展开了进一步研究。

1974年，弗兰克尔主导了一项早期的大规模研究。研究人员要求来自美国不同地区的50名法官给一系列虚构案件中的被告量刑，这些虚构的案件汇总于完全相同的待判决报告中。这项研究的基本发现是：法官们“[缺少共识是常态](#)”，[各种量刑之间的差异“令人震惊”](#)。根据法官的不同，一个[贩卖海洛因的毒贩](#)可能被判1~10年有期徒刑，[银行抢劫犯](#)可能被判5~18年有期徒刑。在[敲诈勒索案](#)中，量刑从最高的20年有期徒刑外加65 000美元罚款，到仅3年有期徒刑且无罚款不等。最令人吃惊的是，对于20起案件中的16起案件，法官们竟然对被告应判多少年无法达成一致意见。

在这项研究之后，人们还进行了一系列其他研究，所有研究都发现了类似的、令人震惊的“噪声”。例如，1977年，威廉·奥斯汀（William Austin）和托马斯·威廉姆斯（Thomas Williams）[对47名法官进行了一项调查](#)，要求他们对同样的5起案件做出判决，每起案件均涉及轻罪。对案件的所有描述资料包括法官在实际判决中使用的信息列表，如指控、证词、既往犯罪记录、社会背景和与性格相关的证据。该研究的重要发现是，法官的判决之间存在“大量差异”。例如，在一起涉及入室盗窃的案件中，建议的刑期从5年有期徒刑到仅监禁30天并处以100美元罚款不等；在一起涉及非法持有大麻的案件中，一些法官建议判处被告监禁，而一些法官则建议判处缓刑。

在1981年进行的[一项更大规模的研究](#)中，被试为208名美国联邦法官，这些法官要对16起完全相同的虚构案件进行判决。该研究的结果同样令人震惊：

在这16起案件中，法官们只对3起案件一致同意判处监禁，但即使在大多数法官同意判处监禁的情况下，他们建议的刑期长短也存在很大差异：在一起诈骗案中，法官们建议的平均刑期为8.5年，而最长的刑期是终身监禁；在另一起案件中，法官们建议的平均刑期为1.1年，而最长的刑期为15年。

尽管上述研究颇具启发性，但它们毕竟都进行了严格的实验控制，所以几乎可以肯定，它们还是低估了现实世界刑事司法中的噪声数量。在现实生活中，法官接触到的信息远远多于被试在这些精心设计的实验中得到的信息。诚然，一些额外的信息与案件是相关的，但也有充分的证据表明，一些微小或看似随机出现的无关信息反而会导致判决结果出现巨大差异。例如，相比于较疲劳时，在一天的开始时段或饭后休息充分时，法官更有可能批准假释；[如果处于饥饿状态](#)，他们则会更加严格。

一项针对几千个[青少年法庭判决](#)案例的研究发现：[当本地足球队在周末输掉比赛后，法官在接下来的星期一会做出更严厉的判决](#)，在本星期其他时间的判决则更为宽容。黑人被告首当其冲地受到这种更严厉判决的影响。另一项研究调查了过去30年中的150万个司法判决案例，也发现了类似的结果，即相比于赢球后的情况，如果当地球队在开庭前一天输球，法官会做出更加严厉的判决。反之，法官就会比较宽容。

研究人员对法国法官在过去12年内做出的600万个判决进行了研究，发现如果判决日恰好是被告的生日⁽¹⁾，那么法官往往会较为宽容。[甚至，像室外温度这种无关紧要的因素也会影响法官的决策](#)，通过对过去4年移民法院做出的207 000份判决进行分析发现，气温的变化会对判

决结果产生很大影响：当室外温度很高时，申请人获得庇护的可能性比较小。如果你希望获得庇护，那么你应该祈祷你的听证会那一天正好天气凉爽。

减少判决中的噪声

20世纪70年代，弗兰克尔法官的论点以及支持这些论点的实证结果引起了爱德华·肯尼迪（Edward Kennedy）的注意，爱德华·肯尼迪是美国前总统约翰·肯尼迪的弟弟，也是美国参议院最有影响力的成员之一。早在1975年，爱德华·肯尼迪就提议推行量刑改革法案，但遗憾的是并无任何进展。爱德华·肯尼迪不屈不挠，他年复一年地敦促美国国会实施该法案，并最终在1984年取得了成功。面对这些不合理的、具有差异性的量刑案件的证据，美国国会终于颁布了《1984年量刑改革法案》（*Sentencing Reform Act of 1984*）。

该法案希望，通过减少“法律赋予负责量刑、执行判决的法官和假释机构的[不受约束的自由裁量权](#)”来降低系统中的噪声。尤其是，国会议员在提到[“过于悬殊”的量刑差异](#)时，特别引用了纽约地区的调查发现：对相同真实案件的量刑可能是监禁3年，也可能是监禁20年。就像弗兰克尔法官建议的那样，国会立法允许设立美国量刑委员会（US Sentencing Commission），其主要职责很明确：发布强制性的量刑指南，并确立其限制范围。

1985年，量刑委员会通过分析1万起真实案件中类似罪行的平均刑期，制定了量刑指南。参与这一过程的[美国最高法院法官斯蒂芬·布雷耶（Stephen Breyer）](#)试图通过指出委员会内部存在的棘手分歧为过去的[做法辩护](#)：“为什么委员们不坐下来，真正使这件事合理化，而不仅仅是拿过去说事？原因很简单，我们不能这样做。我们之所以不能这样做，是因为有大量充分的证据表明我们应该采取截然相反的做法。你可以试图按照等级顺序列出所有罪行，考察其应该受到什么样的惩罚，然

后收集你的朋友给出的结论，看看结论是否一致。我可以告诉你的是，不可能一致。”

根据量刑指南，法官必须在量刑时考虑两个因素：被告所犯罪行和被告的犯罪记录。根据罪行的严重程度，量刑指南将罪行分为43个等级。被告的犯罪记录则主要指被告以前被定罪的次数和判决的严重程度。一旦将罪行和犯罪记录结合起来，法官就能根据量刑指南确定一个相对具体的量刑范围，其最高刑期与最低刑期之间的差距不得超过6个月或25%。法官也可以偏离该范围来加重或减轻处罚，但[必须向法院说明理由](#)。

虽然量刑指南是强制性的，但是它也并非完全没有调整的余地。它并没有达到弗兰克尔法官所希望的程度，也为法官提供了较大的裁量空间。最后，使用了各种不同方法并关注一系列不同历史时期的几项研究都得出了相同的结论：该指南可以减少噪声。更准确地说，[它减少了由于量刑法官身份的偶然性而导致的判决中出现的净差异](#)。

在那之后，[美国量刑委员会对量刑指南的效果进行了详尽的研究](#)。他们比较了1985年（该指南生效前）与1989年1月19日至1990年9月30日期间（该指南生效后）银行抢劫、贩卖可卡因或海洛因，以及挪用银行公款案件中的判决。结果发现，当法官根据量刑指南中的相关因素对罪犯进行判决时，量刑结果更为一致。《1984年量刑改革法案》实施之后，法官对每个等级的罪行的判决差异都比原先少很多。

[另一项研究表明，1986—1987年，法官之间在刑期长短上的差异为4.9个月；而1988—1993年，这一数字下降至3.9个月。一项涵盖了不同时期数据的独立研究](#)发现，量刑指南在减少法官之间的量刑差异方面取得了类似的成效。量刑差异是指审理过相似数量案件的法官在判决的平均刑期上的差异。

尽管有了这些发现，量刑指南还是遭到了猛烈的抨击。一些法官认为，有些判决过于严厉。事实上，这一点涉及的是偏差，而不是噪声。

就我们的目的而言，一个更为有趣的反对意见是：许多法官认为该指南极度不公平，因为它会妨碍法官充分考虑案件的具体情况。减少噪声的代价就是使决策变得更机械化，这让人难以接受。耶鲁大学法学教授凯特·斯蒂斯（Kate Stith）和美国联邦法官乔斯·卡布拉内斯（José Cabranes）写道：“[我们不能对案件的细节视而不见。我们要有洞察力](#)，注重公平性，只有在判断过程中充分考虑特定案件的复杂性，才能做到这一点。”

这一反对意见给量刑指南带来了严峻挑战，有些是法律层面的，有些则是政策层面的，但都未能动摇该指南的效用。[直到2005年，美国最高法院才取消了该指南](#)，主要是由于技术原因，而与此处涉及的争论完全无关。根据美国最高法院的裁决，该指南之后仅作为参考建议。在美国最高法院做出这一裁决后，大多数法官感到更开心了。因为[75%的法官更喜欢建议性制度](#)，只有3%的法官认为强制性制度更好。

将量刑指南从强制性改为建议性，会带来什么影响？哈佛大学法学教授克丽丝特尔·杨（Crystal Yang）没有采用实验或调查的方法对这个问题进行研究，而是利用了庞大、真实的判决案例集。该案例集涉及近40万名刑事被告。通过多种测量方法，[她的核心发现是：法官之间的量刑差异在2005年后明显增加](#)。当该指南是强制性准则时，相比于严厉程度一般的法官，相对严厉的法官判处的刑期仅会多2.8个月；当该指南仅作为建议时，这种差异增加了一倍。像40年前的弗兰克尔法官一样，克丽丝特尔·杨写道：“我的发现引起了人们对公平的广泛关注，因为在很大程度上，量刑法官身份的不同造成了‘同罪不同罚’这一不公正现象。”

在量刑指南成为建议性准则之后，法官更有可能根据自己的价值观做出判决。可见：强制性准则既能减少偏差，也能减少噪声。在美国最高法院做出将量刑指南作为参考建议的裁决后，美国非裔被告与犯有相同罪行的白人被告之间的量刑差距明显增加。同时，女性法官比男性

法官更有可能行使她们日益增加的自由裁量权，从而对犯人做出宽大处理。

弗兰克尔在2002年去世，3年后，强制性的量刑指南被取消，美国人又开始重新经历他所说的噩梦：无秩序的法律。

弗兰克尔法官为争取量刑指南而战的故事，能够让我们一窥本书将要论述的几个关键点。

第一，世界充满复杂性与不确定性，因此判断是一件困难的事情。这种复杂性在司法领域显而易见，在其他大多数需要专业判断的情形中也是如此。从广义上讲，这些情形包括由医生、护士、律师、工程师、教师、建筑师、招聘委员会成员、图书出版商、各类企业高管以及球队经理人等所做的判断。凡是判断，就不可避免地存在分歧。

第二，分歧的程度比我们预想的大得多。尽管很少有人会反对司法自由裁量权原则，但几乎每个人都会反对它所产生的差异。在理想情况下，判断应该完全相同，不应该存在差异。因为差异会导致层出不穷的不公平现象、高昂的经济成本和多种类型的错误。

第三，噪声是可以减少的。由弗兰克尔倡导并由美国量刑委员会实施的方法——颁布强制性量刑指南，是成功减少噪声的几种方法之一。另外，还有一些方法更适用于其他类型的判断。一些减少噪声的方法同时也可以减少偏差。

第四，减少噪声的努力往往会招致反对，让人举步维艰。这些问题必须得到解决，否则对抗噪声的斗争最终还是会失败。

- 消除噪声

判决中的噪声

- 实验表明，不同法官对相同案件的判决存在巨大差异。这种差异是不公平的。被告的判决结果不应取决于该案件恰好被指派给了哪位法官。
- 刑事判决不应取决于法官在听证会上的心情或室外的温度。
- 强制性的量刑指南是解决噪声问题的一种方法。许多人不喜欢这种做法，因为这会限制法官的司法自由裁量权，他们认为这种自由裁量权可能是确保公平和准确性所必需的。毕竟，每个案例都是独一无二的。

第2章

系统噪声，给人达成一致的错觉

我们最初接触噪声，并对这个主题产生兴趣，并不是因为遇到了像刑事司法案件那般富有戏剧性的案例。实际上，这次邂逅纯属偶然，缘于一家保险公司，这家公司与我们中的两人所属的咨询公司有合作。我们的研究揭示了营利性组织中存在噪声问题的严重性——组织会因为决策中的噪声损失惨重。我们在保险公司的经历有助于解释为什么噪声问题常常被忽视，以及要想解决这个问题，我们能做些什么。

保险公司的高管们正在权衡减少噪声这项工作是否有必要，即是否需要提升为公司做出重大财务决策的人员之间判断的一致性。虽然每个人都认为需要保持一致，但大家都明白这些判断不可能完全一致，因为它们具有一定的主观性。总之，有些噪声不可避免。

在谈到噪声数量时，大家出现了分歧。高管们怀疑，噪声对他们的公司而言是否真的是一个重大问题。非常值得称赞的是，他们同意通过一个简单的实验来解决这个问题。[我们称这个实验为噪声审查](#)。结果让他们感到惊讶。这个案例也有力地证明了噪声问题确实存在。

抽签会引发噪声

大型保险公司中的许多专业人员都有权做出对公司具有约束力的判断。例如，上述这家保险公司雇佣了许多核保员，他们会根据财务风险为投保人提供保费报价，比如为银行提供针对因欺诈或流氓交易造成损失的保险。公司还雇佣了许多理赔员来预测未来会产生索赔费用，并在发生纠纷时与索赔人进行谈判。

保险公司的每个大型分支机构中，都会有数位具备资质的核保员，当有人想知道报价时，任何有空的核保员都可能被指派去做这项工作。实际上，核保员是随机指定的，而报价数额却会对公司产生重大影响。如果保费高，而报价被客户接受，那么对保险公司是有利的，却存在将业务输给竞争对手的风险。低额保费更容易被客户接受，但对保险公司而言却是不利的。也就是说，对于任何风险，都有一个恰到好处的“金发姑娘价格”（Goldilocks Price）⁽²⁾——既不会太高，也不会太低。大多数专业人士判断出的平均价格很可能与这个“金发姑娘价格”相差不大，高于或低于这个价格，保险公司都会付出很大的代价。

理赔员的工作也会影响到保险公司的财务状况。例如，假设有人提交了一份索赔申请，而索赔人是一名在工业事故中失去右手的工人。就像核保员的分配一样，某位理赔员被指派受理该索赔案是因为他恰好有空。该理赔员的职责是收集案件的实情，并向公司提供最终成本的估价。他随后负责与索赔人的代表进行谈判，以确保索赔人获得保单中承诺的利益，同时保险公司也不用支付过多费用。

早期的估价很重要，因为它为理赔员与索赔人在未来的谈判中设定了一个隐性的目标。法律规定，保险公司有义务为每项索赔预留足够的资金。从保险公司的角度来看，这里的“金发姑娘价格”很重要，报价过高或过低都无法确保很好地解决问题：一方面，如果报价过低，索赔人的律师可能会选择诉诸法律；另一方面，过于慷慨可能会使理赔员因自由度过高而同意索赔人的无理要求。理赔员的判断对公司很重要，对索赔人而言也同样重要。

我们用“抽签”这个词来强调在选择核保员或理赔员时的随机性。正常情况下，保险公司只会为每一起案件分配一位专业人员，我们无法得知如果换成另外一位专业人员，情况会有什么样的不同。

当然，抽签也有好处，不是完全不公平的。进行“好的”（如一些大学的课程）或“坏的”（如美国的征兵工作）资源分配时，人们都能够接

受抽签的形式。抽签有它存在的意义，但我们所谈论的判断中的抽签并不分配任何东西，采用这种比喻是想要强调：它们只会产生不确定性。想象一下，假如一家保险公司的核保员是不存在噪声问题的，他们能够设定最佳保费，但一个“抽签装置”介入并修改了实际报价，显然这样的抽签是不合理的。在一个系统中，如果结果取决于随机选择的专业判断者的身份，这非常不合理。

噪声审查，系统噪声无处不在

以类似抽签的方式指派一位法官来审理刑事案件或选择一位射击手代表一支队伍出战，都会产生差异性，但这种差异性是不可见的。就像审查美国联邦法官在量刑方面的差异性那样，噪声审查是揭示噪声的一种方法，在这样的审查中，多人对同一起案件进行评估，于是他们评估中的差异性就可以显现出来。

核保员和理赔员的判断尤其适用于这种审查，因为他们都是根据书面材料做判断的。在噪声审查开始之前，保险公司高管向每个小组（核保员和理赔员）详细描述了5个代表性案例，随后要求每一组的成员对两三个案例进行独立评估。为了检查他们判断之间的差异性，研究人员事先没有告知这些员工该研究的目的。

在继续阅读之前，不妨想想，你会如何回答以下问题：在一家经营状况良好的保险公司中，如果你随机选择了两位有资质的核保员或理赔员，你预测他们对同一案件的估价会有多大差异？具体而言，两个估价之间的差异占他们估价平均值的百分比是多少？

在随后的几年中，我们让保险公司的众多高管们也回答这个问题，并获得了各界人士对此的估值。令人惊讶的是：有一个答案显然比其他所有答案都有人气，大多数高管猜测的差异值为10%或更少。我们调查了828位来自不同行业的CEO和高管，询问他们希望在类似的专家判断

中发现多少差异，最常见答案的中位数也是10%（排名第二常见的答案是15%）。10%的差异意味着，如果两位核保员中的一位将保费设定为9500美元，那么另一位可能报出10 500美元的保费。虽然这一差异不算微不足道，但它仍在组织能够容忍的差异范围内。

然而，我们的噪声审查发现的差异更大。根据我们的计算，核保员的中位数差异为55%，大约是大多数人（包括公司高管）估计值的5倍多。这一结果意味着，当一位核保员将保费定为9500美元时，另一位核保员很可能不是将保费定为10 500美元，而是定为16 700美元。另外，理赔员的中位数差异为43%。需要注意的是，我们强调的这些结果是中位数，事实上，在一半的案件中，两个判断之间的差异甚至更大。

当我们将噪声审查结果提交给高管们时，他们很快意识到巨大的噪声会让公司付出高昂的代价。一位高管估计，保险公司每年在承保方面的噪声成本是数亿美元，包括报价过高造成的业务损失和低价合同造成的收益损失。

没有人能准确地说出一项决策中到底存在多少错误或多少偏差，因为没有人能确切地知道每一种情况下的“金发姑娘价格”。但是，要想测量射击点的分散程度，并意识到差异性是一个问题，我们并不需要知道靶心在哪里。数据显示，要求客户支付的价格在一定程度上取决于被抽中的是哪位核保员。至少可以说，如果未经客户同意就做出这样的抽签式指派，客户也不会满意。更为普遍的是，与组织打交道的人希望系统能够提供可靠的一致性判断。总之，他们不希望自己的系统中存在噪声。

不必要的变异性与有利的多样性

系统噪声的一个特点是，它是不必要的，而我们想要强调的是：判断中的变异性并非总是不必要的。

让我们想想人们的偏好或品位问题。如果有10位影评人观看同一部电影，或有10位品酒师评价同一款酒，又或有10位读者阅读同一部小说，我们并不期望他们给出完全相同的评价。毕竟，人各有所好，这完全在意料之中，因为没有人愿意生活在一个所有人的好恶都完全相同的世界中。然而，如果我们的个人品位被误认为是专业判断，那么品位多样性就可能导致错误。如果一位电影制片人决定推进一个“小众”的项目（比如，转盘电话的兴衰），其原因仅仅是他个人喜欢这个剧本，而不管其他对这个项目的看法，那么就可能会犯下大错。

在最优判断会得到奖励的竞争环境下，人们也会期待和欢迎判断中的变异性。当几家公司或同一组织中的几个团队竞相为同一个客户问题制订创新性解决方案时，我们就不希望他们采用相同的方法；当多个研究团队试图攻克同一个科学问题时，情况也是如此，例如在研制疫苗时，我们希望各研究团队能从不同的角度去看待问题。即使是预测者，有时也会表现得像个竞争者一样。如果预测者能从别人没有想到的视角，正确地预测到经济衰退的可能性，那么他肯定会一举成名，而墨守成规、从不偏离一致性的人则将碌碌无为。在这种情况下，想法和判断的变异性同样有存在的必要，因为变异是第一步。在随后的阶段，这些判断的结果会相互竞争，最优判断会获得胜利。市场犹如自然界，没有变异，选择就不会起作用。

品位和竞争引发了有趣的判断问题，但我们的讨论重点是判断中存在的 unnecessary 的变异性。系统噪声是系统层面的问题，系统是组织，不是市场。当交易员对同一只股票的价值做出不同的评估时，其中一些人会赚钱，另一些人则不会，可以说正是分歧造就了市场。但是，如果随机选择其中一名交易员来代表他任职的公司进行评估，我们还发现他的同事做出了非常不同的评估，那么该公司就存在系统噪声，这就是一个大问题。

当我们向一家资产管理公司的高级经理提出我们的发现，并敦促他

们进行噪声审查时，这个问题就被很好地说明了。他们要求公司中42名经验丰富的投资经理估计一只股票的公允价格（当股价达到该价格时，投资经理不会进行买卖交易）。他们只是根据一页纸的业务描述进行分析，其中的数据包括过去3年的简化损益表、资产负债表、现金流量表以及对未来2年的预测。用我们在保险公司案例中使用过的噪声审查方法测量噪声水平，得到的噪声中位数为41%。同一家公司的投资经理在采用相同估值方法的情况下出现如此大的差异，这可不是一个好消息。

正如资产管理公司、刑事司法制度以及前面讨论的保险公司的案例一样：只要判断者是从一群具有同等资历的人中随机挑选出来的，噪声就是一个必然存在的问题。系统噪声困扰着许多组织和机构：比如，你去医院接受哪位医生的治疗、哪位法官在法庭上审理你的案件、哪位专利审查员审查你的申请、哪位客户服务代表听取你的投诉等。在这些由不同人做出的判断中，不必要的变异性会引发严重问题，包括金钱损失和广泛存在的不公平现象。

关于判断中不必要的变异性，一个常见的误区是认为它无关紧要，因为通常来说，随机错误是可以相互抵消的。当然，在关于同一案件的判断中，正负误差会趋于相互抵消，我们将详细讨论如何利用这一特性来减少噪声。但存在噪声的系统并非对同一案例做出了多次判断，而是对不同的案例做出了有“噪声”的判断。如果保险公司对一份保单的理赔金额估价过高，而对另一份保单估价过低，从平均值而言，两次估价看起来可能是适当的，但实际上保险公司却犯下了两次代价高昂的错误。如果两名罪犯都应该被判处5年有期徒刑，却分别被判处了3年和7年有期徒刑，那么尽管平均值是5年，但事实上正义并没有得到伸张。因此可以看出，在充满噪声的系统中，错误不会相互抵消，只会累加。

达成一致的错觉

几十年前已有大量文献记录了专业判断中的噪声，因为我们了解这

些文献，所以对保险公司的噪声审查结果并不惊讶。然而，令我们惊讶的是高管们在听到报告结果时的反应：公司中没有人预料到我们会观察到那么多的噪声数量，没有人质疑审查的有效性，也没有人声称观察到的噪声数量是可接受的。噪声问题及其导致的巨大成本对该组织来说似乎是一个新问题。噪声就像地下室漏水，它之所以能被容忍，不是因为人们认为它是可接受的，而是因为它一直未被发现。

这怎么可能呢？同一职位、同一办公室的专业人员之间怎么会有如此大的差异而不被察觉？高管们知道这对公司的业绩和声誉是一个重大威胁，他们怎么会没有注意到这一点呢？我们发现，组织经常无法识别系统噪声问题，组织对噪声的普遍无视与它的普遍存在一样有趣。噪声审查表明，受人尊敬的专业人员以及雇用他们的组织存在一种一致性错觉，事实上，他们在日常的专业判断中常常存在分歧。

想要理解一致性错觉是如何产生的，你可以把自己想象成一位核保员：你有5年以上工作经验，你知道自己在同事中享有较高声誉，并且你也尊重和喜欢你的同事。你对自己的工作能力很有信心。在仔细分析了一家金融公司所面临的复杂风险后，你给出了结论，认为将保费设定为200 000美元比较合适。这个问题虽然很复杂，但与你每天要解决的其他问题没有太大的区别。

现在想象一下，你被告知你的同事也拿到了同样的资料，并针对同样的风险进行了估价。然而，他们中至少有一半人设定的保费要么高于255 000美元，要么低于145 000美元，你会相信吗？这种想法的确很难让人接受。我们怀疑，那些听说过噪声审查并接受其有效性的核保员很难真正相信这一结论适用于他们自己。

多数时候，我们大多数人都有一个根深蒂固的信念，即这个世界就是它看起来的样子。当然，也很容易相信：“其他人对世界的看法与我差不多”。“其他人对世界的看法与我差不多”这样的信念也被称为“[天真的现实主义](#)”，它对于我们与他人共有的现实感而言至关重要，而且我

们很少质疑这些信念。我们在任何时候都对周围的世界持有一种单一的解释，通常而言，我们很少会投入精力来寻找其他可能的解释。在我们看来，一种解释就足够了，我们将其视为真实的经验。人们通常不会想到用其他方式来看待自己所看到的东西。

就专业判断而言，相信他人也像我们一样看待世界的信念每天都在以多种方式被强化。首先，我们与同事使用同一种语言，遵循着同一套规则来考虑决策中的重要因素。我们也有一些可靠的经验，可以就违反这些规则的荒谬判断与他人达成一致。我们把与同事之间偶尔出现的分歧看作对方的判断失误，却很难意识到，我们一致认定的规则是含糊不清的。这些规则虽然足以消除某些可能性，但并未明确规定可对特定案例做出的积极回应。我们可以与同事愉快地相处，根本不会注意到他们实际上并不像我们那样看待世界。

我们采访的一位核保员描述了他成为自己所在部门“老手”的经历：“当我还是新人时，75%的案件都是与我的主管一起讨论的，没过几年，我就不需要再这样做了。我现在被视为专家，久而久之，我对自己的判断越来越有信心。”像大多数人一样，这位核保员主要通过不断实践来建立对自己判断的信心。

这个过程的心理机制已被大众所熟知。信心来自主观的判断经验。个体做出这些判断越来越熟练、越来越轻松，这其中的部分原因在于，这些判断与过去经历过的判断类似。随着时间的推移，这位核保员学会了与过去的自己达成一致，他对自己的判断的信心也随之增强。他的叙述并没有表明，在最初的学徒阶段之后，他已经学会了与他人达成一致，已经知道了自己的观点与别人的观点能达到多大程度的一致，或是曾经尝试过防止他的做法偏离同事的做法。

对于保险公司来说，一致性错觉只有通过噪声审查才能被戳破。领导层为何一直没有意识到噪声问题？这里有几个可能的因素，但在许多情况下起最大作用的是对分歧的不适感。大多数组织喜欢共识与和谐，

不喜欢异议和冲突。现有的程序似乎是特意设计的，用来将分歧出现的频率最小化，以及在发生分歧时予以消除。

我们可以用明尼苏达大学心理学教授、杰出的成绩表现预测研究者内森·昆塞尔（Nathan Kuncel）给我们分享的故事来说明这个问题。昆塞尔正在帮助一所学校的招生办公室评估其决策过程。首先，一位评阅人阅读了一份申请文件并进行评分，然后将其连同评分一起交给了第二位评阅人，后者也对其进行评分。昆塞尔建议最好隐去第一位评阅人的评分，以免影响第二位评阅人的判断。学校的答复是：“我们以前就是这样做的，却导致了太多分歧，所以我们才改用了现在的方式。”要知道，这所学校并不是唯一一家认为避免冲突与做出正确决定至少同等重要的机构。

考虑一下许多公司采用的另一种机制：对失败的判断进行事后总结。作为一种学习机制，事后总结是有用的，但是如果真的犯了错误，比如某项判断严重偏离了专业规范，那么对该错误进行讨论就没有意义了。专家们很容易得出这样的结论：失败的判断与共识相去甚远。他们也可能将其作为一种罕见的例外情况记录下来。错误的判断比正确的判断更容易被识别。在做出广泛可接受的判断时，指出不合格同事犯的严重错误并排挤他们，对专业人士了解自己与他人的分歧程度并没有什么帮助。相反，对错误判断的简单共识甚至可能会强化一致性错觉。这样，人们永远也不会认识到“系统噪声无处不在”的事实。

我们希望你能认同我们的观点，认识到系统噪声是一个严重的问题，它是普遍存在的。判断天生是非制式的，所以噪声的出现也是自然而然的。然而，正如我们将在整本书中看到的，当组织认真审视和看待噪声时，观察到的噪声数量几乎总是令人震惊。我们的结论很简单：哪里有判断，哪里就有噪声，而且它比你想象的还要多。

- 消除噪声

保险公司中的系统噪声

- 我们非常依赖核保员、理赔员和其他人员的专业判断的品质。但我们在把案件指派给每一位专家时，常常错误地假定其他专家也会对该案件做出类似的判断。
- 系统噪声比我们想象中的大数倍，甚至超出我们可以承受的范围。如果没有噪声审查，我们将永远不会意识到这一点。噪声审查打破了一致性错觉。
- 系统噪声是一个严重的问题：它给我们造成了难以估计的损失。
- 哪里有判断，哪里就有噪声，而且其数量之大超出我们的想象。

第3章

单一决策，仅发生一次的重复决策

到目前为止，我们所讨论的案例研究都与重复决策相关。比如，对盗窃犯的恰当判决是什么？对某一特定风险该收取多少保费？虽然每个具体的案例在某种意义上都是独特的，但像这样的判断属于重复决策。医生诊断患者，法官审理假释案件，招生人员审查入学申请书，会计师准备税单等，这些都是重复决策。

正如前一章所介绍的，重复决策中的噪声可以通过噪声审查识别出来。当无本质差异的专业人员在相似的情况下做决策时，不必要的差异性很容易被定义和测量。然而，要将噪声这一概念应用于我们称为单一决策的判断中，似乎要困难得多，甚至是不可能的。

例如，让我们回想一下2014年世界面临的危机。在西非，无数人死于埃博拉病毒感染。世界是紧密相连的，大量预测表明，病毒将会迅速蔓延到世界各地，并对欧洲和北美造成尤为严重的影响。在美国，有人坚持要求暂停来自受感染地区的航班，并采取积极措施关闭边境交通线路，很多知名人士和有识之士也都赞成采取这些措施。

这是美国前总统奥巴马在他任期内面临的最困难的决策之一，这也是他以前从未遇到过，以后也不太可能再遇到的决策。最终，他没有选择关闭边界，相反，他派遣了3000人到西非，其中包括医务工作者和军人。并且，奥巴马领导了一个由国际组织组成的多元化国际联盟，并利用其资源和专业知识从根源上解决问题。

单一决策**vs**重复决策

像美国前总统奥巴马应对埃博拉疫情这样只需做一次的决策就是单一决策，因为它们不是个人或团队必须常常做出的决策，一般缺乏预先准备好的应对措施，并且具有真正独有的特征。在应对埃博拉疫情时，奥巴马和他的团队没有真正的先例可供借鉴。重要的政治决策通常是单一决策，军事指挥官的重大抉择也属于这一范畴。

对个人来说，你在找工作、买房子或求婚时做出的决策也有类似的特点。即使这可能不是你的第一份工作、第一套房子或第一段婚姻，尽管有无数人也面临过这些决策，但这个决策对你而言依然是独一无二的。在商业领域，公司负责人经常要做出对他们来说似乎是独一无二的决策，比如：是否推出一个可能改变“游戏规则”的创新计划，在流行病暴发期间应该关闭多少店铺，是否要在国外开设办事处等。

可以说，单一决策和重复决策之间的差异在于连续性而非类别。比如，核保员可能会处理一些非同寻常的案例。如果你已经是第四次买房，那么你可能已经开始将买房视为一个重复决策了。极端化的例子能够清晰地表明，对单一决策和重复决策进行区分是有意义的，比如，是否发动战争是一回事，进行年度预算审查则是另一回事。

单一决策中的噪声

通常，单一决策被视为与重复决策截然不同的类型。大型公司中由无本质差异的雇员所做的常规决策就是重复决策。社会科学家已经对重复决策进行了大量的研究，而高风险的单一决策则一直是历史学家和管理大师们的研究对象。这两类决策的研究方法有很大的不同。对重复决策的分析往往采用统计方法，社会科学家会通过对多个类似的决策进行评估，识别其中的模式，确定其规律并测量其准确性。相反，对单一决策的讨论通常从因果关系视角进行事后总结，集中探讨的是事件发生的原因。历史分析，比如对成功与失败的管理案例进行分析，是想了解某个独一无二的判断是如何做出的。

单一决策的特性对研究噪声提出了挑战。我们将噪声定义为对相同问题进行判断的过程中产生的不必要的变异。单一决策无法被重复，因而这个定义对它并不适用。毕竟，历史只发生一次，你永远无法将奥巴马在2014年向西非派遣医务工作者和军人的决策，与其他美国总统在特定时间处理特定问题的决策进行比较（尽管你可以进行推测）。你可以将你决定嫁给心仪之人的决策同其他与你相似的人的决策做比较，但这种比较显然不同于我们对同一案件中不同核保员提出的报价所进行的比较。也就是说，对于单一决策，我们没有直接的方法来考察是否存在噪声。

然而，单一决策并非不会受到那些在重复决策中产生噪声的因素的影响。在射击场上，C队（即图0-1中的噪声队）的射击手可能曾朝不同的方向调整来复枪的瞄准器，也可能只是他们的手不稳。如果只观察C队的第一名射击手，我们无法得知该队的噪声水平，但事实上噪声源是一直存在的。在做出单一决策时，你必须想象另一个决策者，即使他和你能力相当、有着相同的目标和价值观，他也会从相同的事实中得出不同的结论。作为决策者，你应该认识到，如果情境中的无关变量或决策过程有所不同，那么你就可能会做出不同的决策。

换句话说，我们无法测量单一决策中的噪声，但如果我们进行“反事实思考”（counterfactual thinking），则可以肯定噪声是存在的。就像射击手的手不稳将意味着单次射击可能落在靶心以外的地方一样，决策者内部以及决策过程中存在的噪声，也意味着单一决策可能会有所不同。

接下来，我们思考一下影响单一决策的所有可能因素。如果负责分析埃博拉疫情的威胁和制订应对计划的专家是不同的人——他们拥有不同的背景和生活经历，那么他们向奥巴马提出的建议会相同吗？如果以稍微不同的方式呈现相同的事实，那么讨论还会以相同的方式展开吗？如果关键人物当时的心情有所不同，或在暴风雪中开会，最终的决策会

不同吗？从这个角度看，单一决策似乎就有了可变的空間。可見，决策可能会受到很多我们未知的因素的影响，并最终变得不同。

关于另一个反事实思考的训练是，我们可以考虑不同国家和地区如何应对新型冠状病毒肺炎疫情。即使病毒在大致相同的时间内，以类似的方式侵袭世界各国，不同国家做出的应对也存在很大差异。这种差异为不同国家决策中的噪声问题提供了确切证据。但是，如果该病毒只侵袭了一个国家呢？在这种情况下，虽然我们不会观察到任何差异性，但决策中的噪声并不会减少。

控制单一决策中的噪声

对这一问题的理论探讨非常重要。如果单一决策与重复决策一样存在噪声，那么用于减少重复决策中噪声的策略应该也可以用于提高单一决策的品质。

有人建议，当你做出独特的决策时，你的直觉要将其视为某一类决策中的一种。这一建议似乎与我们的直觉相悖。有人甚至声称，概率思维的规则与不确定性情境下做出的单一决策完全不相关，而且单一决策需要使用一种截然不同的方法。

我们的意见恰恰相反。从减少噪声的角度来看，单一决策是仅发生一次的重复决策。无论只做一次决策还是做一百次决策，你的目标都应该是减少偏差和噪声，而且减少错误的实践方法在单一决策和重复决策中同样有效。

- 消除噪声

清楚认知单一决策

- 当你遇到一个不同寻常的机会时，你的应对方式就可能产生噪声。
- 请记住：单一决策是仅发生一次的重复决策。
- 那些造就你的个人经历，与你目前要做的决策实际上没多大关系。

NOISE

第二部分

你的大脑是一种
测量工具

- 发现噪声

判断永远不可能完美，但可以尽量准确

测量，在日常生活中的含义与其作为科学术语时的含义是一样的，就是使用测量工具给某个对象或事件在一定标尺上赋值。比如，你可以使用卷尺测量地毯的长度，也可以使用温度计来测量温度——无论是华氏还是摄氏度。

判断与此类似。法官为罪犯量刑就是依据罪行程度为刑期赋值，核保员会为每一项风险设定一个值（用金额表示），医生做出的诊断也是如此。需要指出的是：判断结果并不一定是数字，“排除合理怀疑”“黑色素瘤晚期”“建议手术”等也是判断。

因此，我们可以将“判断”描述为使用人类大脑作为工具的一种测量。测量这一概念隐含着准确性的目标——逼近真实值和使错误最小化。做出判断不是为了让人印象深刻，不是想表明立场，也不是为了说服他人。需要注意的是：此处的“判断”源自专业心理学文献，比日常语境中该词的含义要窄得多。判断不等同于思考，“做出准确判断”也不等同于“拥有良好的判断力”。

根据我们的定义，判断是可以用一个单词或一个短语总结出的结论。如果情报分析师写了一篇很长的报告，得出某国政权不稳定的结论，那么只有这个结论才是判断。判断就是测量，既与做出判断的心理活动相关，也与这项心理活动的结果相关。在英语中，“judge”除了表示“判断”，也指代做出判断的人，而这个人并不一定是法官。

虽然判断力求准确，但即使是在科学测量中，也不可能完美地达到这一目标，更别提判断了。在判断中总会存在一些误差，其中一些是偏

差，一些是噪声。

为了让你了解噪声和偏差是如何导致误差的，我们邀请你花不到一分钟的时间来玩一个游戏。如果你的智能手机上有带计时功能的秒表，那么利用它，你可以在不看显示器且无须暂停的情况下，测量连续的间隔时间。你的目标是在不看手机的情况下，让秒表连续记录5段10秒的时间。在开始之前，你最好先感受几次10秒有多长。

现在看一下你的手机上记录的每一段间隔时间。当然，手机难免也会存在一些误差，但这种误差很小。你会发现，手机上记录的每一段时间并非都是10秒，而且差距相当大。你试图精确地重现相同的时间，但根本无法做到。这种你无法控制的差异就是一个关于噪声的例子。

这一点都不奇怪，因为就人的生理和心理而言，噪声无处不在。毋庸置疑，不同的个体在生理上是存在差异的，这就像同一个豆荚中没有两颗完全相同的豌豆。在个体内部同样存在差异，比如，你的心跳不可能是完全规则的，你不可能完全准确地重复同一个姿势。如果有一个听力学家帮你检查听力，可能一些声音太小，你完全听不见；一些声音很大，你总能听见；然而，总有一些声音，你有时能听见，有时又听不见。

现在看看你手机上的5个数字。看出来一些规律了吗？例如，是每一段时间都短于10秒吗？如果是，那么这意味着你内心的钟表跑得太快了。在这个简单的任务中，偏差指的是你的平均时间与10秒之差（无论是正的还是负的）。噪声导致了结果的差异，其外在表现类似于我们之前看到的散点图。在统计学中，[标准差是测量差异的最常见指标](#)，我们将用标准差来测量判断中的噪声。

你可以把大部分判断，尤其是预测性判断与你刚刚做出的判断看成一回事，当做出预测时，我们在试图接近真实值。例如，经济分析师的目标是尽可能接近下一年国内生产总值增长的真实值，医生的目标是试图做出正确的诊断。（需要注意的是，本书中的“预测”一词并不是指对

未来进行预测，在我们看来，对现有病情的诊断也是一种预测。）

我们会充分利用“判断”与“测量”之间的类比来帮你理解判断中的噪声。进行预测性判断的人就像瞄准靶心的射击手或试图测量粒子真实重量的物理学家，判断中的噪声则意味着误差。简而言之，当判断的目的在于获得真实值，两个不同的判断就不可能都对。就像测量一样，在某些特定任务中，有些人可能会由于技能水平不高或训练不足，比其他人产生更多的误差。当然，做出判断的人也同样远非完美。因此，我们需要了解和测量他们的误差。

当然，大部分专业判断比测量间隔时间要复杂得多。在第4章，我们会对不同类型的专业判断做出界定，并探究它们的目标。在第5章，我们将讨论如何测量误差，以及如何量化系统噪声对误差的影响。在第6章，我们会进一步分析系统噪声并分析其组成部分，即不同类型的噪声。在第7章，我们将探讨其中一种噪声：情境噪声。在第8章，我们将展示群体如何常常在判断中放大噪声。

通过这几章，我们会得出一个简单的结论：同所有测量工具一样，人类的大脑并不完美，其中同时存在偏差和噪声。为什么会这样？会有多少偏差和噪声？让我们一起来揭开谜底吧。

第4章

什么是判断

本书论述的是一般意义上的专业判断，我们假设如果做判断的人有足够的力量，就能做出准确的判断。然而，判断这一概念本身包含着一个你不得不承认的事实：你永远无法确定一个判断是不是准确。

请思考“判断问题”和“主观判断”这两个短语的差别。我们认为“太阳明天会升起”或“NaCl是氯化钠的化学式”这样的陈述并非判断，因为任何理性的人都会完全同意上述观点。判断问题在于其结果存在不确定性，我们承认理性或有力量的人在某一问题上是不可能存在分歧的。

然而，分歧需要有一定的限度。事实上，“判断”这个词主要用于人们能够达成一致意见的场合。判断问题不同于品位或意见，在后两种情况下，彼此之间的差异是完全能被接受的。对于那些对噪声审查感到震惊的保险公司高管来说，理赔员在“披头士乐队和滚石乐队哪个更好”，或是“鲑鱼和金枪鱼哪个更好”之类问题上存在的分歧，并不会让人觉得不妥。

包括专业判断在内的判断介于两者之间：一端是事实或计算问题，另一端是品位或意见问题。因此，不同主体的判断必然存在一定程度的分歧。

到底多大的分歧是可以被接受的？这本身就是一个判断问题，而且取决于问题的难度。人们很容易对一个荒诞不经的判断达成一致。比如，在一起普通诈骗案中，法官们会一致认同：判罚款1美元或判无期徒刑都是不合理的；[在葡萄酒比赛中，评委们对哪种葡萄酒应该获奖可能会分歧很大](#)，但对于哪些葡萄酒应该被排除在获奖的门槛之外却往往能达成一致。

判断的经验

在进一步讨论判断经验之前，我们邀请你自己先做一个判断，完成这项练习有利于你理解本章的内容。

设想你是一个团队里的成员，负责评估一家中等规模的地区性金融公司的CEO候选人，该公司正面临着日益激烈的竞争。你需要评估以下候选人在工作两年后获得“成功”的可能性。成功的定义很简单，就是候选人在两年任期结束时，仍能保住CEO的职位。用0（不可能）到100（必然）范围内的数字来表示概率。

迈克尔·甘巴迪（Michael Gambardi）今年37岁。12年前从哈佛商学院毕业以来，已经任职过几家公司。早期，他是两家初创公司的创始人和投资人，但由于没有得到足够的资金支持而宣告失败。随后，他加入一家大型保险公司，并迅速晋升为欧洲地区的首席运营官。在该职位上，他发起并实施了一项有助于及时解决索赔问题的重要改进措施。同事和下属都认为他有效率，但也霸道、粗暴，在他的任期内，高管的离职率很高。同事和下属也认可他的正直以及愿意为失败承担责任。在最近的两年里，他一直担任一家中型金融公司的CEO，这家公司最初面临倒闭的风险，但他成功使公司转危为安。人们认为他是成功的，尽管与他个人合作很难。目前，他表示有兴趣换个工作。几年前面试过他的人力资源专家在创造力和精力两个方面给了他高分，但也表示他很傲慢，有时甚至是专横的。

让我们简单回顾一下已知信息，甘巴迪是一家地区性金融公司CEO的候选人，该公司已经取得了一定的成就，但也面临着日益激烈的竞争。你需要回答：如果甘巴迪被录用，两年后仍在职的可能性有多大？请在继续阅读后面的内容之前，给出

一个0~100的具体数字。如有必要，你也可以重新阅读上文的描述。

如果你认真思考这个题目，可能会发现自己很难做出判断。这里有大量的信息，但很多信息看起来并不一致，你必须努力构建一个帮助自己做决策的前后一致的印象。在建构这种印象时，你会专注于一些看起来非常重要的细节，但同时也可能忽略了其他细节。如果要求你解释下自己给出的数字，你可能会提到一些主要的事实，但这些又不足以全面地解释你的判断。

上述思维过程揭示了我们在“判断”过程中的几个心理特征。

- 首先，在前文所提供的线索（这些可能只是你需要了解的一部分信息）中，你对一些信息的关注度可能会高于对其他信息的关注度，并且你可能完全没有意识到这一点。你有没有意识到“甘巴迪”是一个意大利姓氏？你记得他是从哪所学校毕业的？这些设置是为了让你信息过载，从而让你无法回忆起该案例的所有细节。而且，你回忆起来的内容很有可能和其他读者能够回忆起来的内容不同。“选择性注意”（selective attention）和“选择性回忆”（selective recall）是人与人之间判断差异性的源头之一。
- 其次，你随意整合了所有线索，形成了一个关于甘巴迪的整体印象。这里的关键词是“随意”，也就是“非正式”，你并没有制订一个详细的计划来回答这一问题。在没有完全意识到的情况下，你的大脑建构了一个有关甘巴迪的优点、弱点以及他有可能面临的挑战的整体印象。这一“非正式”的思考能够让你快速做出判断，但它也有可能产生差异性：一个正式的过程，例如增加一系列数字，能够确保获得相对一致的结果，但在非正式的心理运作过程中，噪声难以避免。

- 最后，你将整体印象转换成一个用于衡量成功概率的数字。将整体印象与0~100中的某个具体数字对应起来是一个了不起的过程，我们在第14章会对此再次进行讨论。同样，你并不能确切地知道这么做的原因。比如，为什么你选择65，而不是61或69？很有可能的是，它只是在某个时间点，你脑海中忽然就跳出来的一个数字。你确认一下这个数字是否合适。如果你觉得不合适，大脑中则会跳出另一个数字。这一过程也会导致人与人之间的差异性。

既然复杂判断过程中的这三个步骤都会产生差异性，那么在回答有关甘巴迪的问题时存在噪声也就不奇怪了。如果你找几个朋友来阅读上面的案例，你会发现，你们对他成功的概率的评估也会有很大差异。我们让115名MBA学生来对之做判断，他们对甘巴迪能成功的概率的评估在10~95之间，这意味着非常大的噪声。

你可能会发现，秒表实验和甘巴迪问题体现的是两类噪声：秒表实验中产生的差异性单一判断者的噪声；而甘巴迪案例中，判断的差异性是不同判断者之间的噪声。从测量学的角度而言，第一个问题反映的是“个体内的信度”（within-person reliability），第二个问题反映的是“个体间的信度”（between-person reliability）。

判断的目的：只是响应“做出判断”的内部信号

你对甘巴迪问题的回答是一种预测性判断——关于这一概念，我们在前文中已经定义过，然而，它在一些很重要的方面不同于其他预测性判断，如曼谷明天的最高气温、今晚足球赛的结果或下届总统大选的结果。如果在这些问题上你不同意朋友的判断，你会在某个时间点知道谁是对的。但是，如果在甘巴迪日后表现的问题上你和朋友之间存在分歧，时间也无法告诉你谁是对的，原因很简单：甘巴迪根本不存在。

即使要评估的是一个真实的人，并且我们知道结果，我们也无法证实或证伪一个单一的概率判断（除非概率是0或1）。结果并不能表明“事前概率”（*ex ante probability*）是多少。如果一个被认为有90%的可能性会发生的事件并未真正发生，也并不能说明概率判断是不好的，毕竟，即使某个结果只有10%的可能性会发生，它也有可能真正发生。基于如下两个不同的原因，甘巴迪案例就是一个无法验证的预测性判断：甘巴迪是一个虚构的人物；关于该人物的判断结果是概率性的。

很多专业判断都是无法验证的。除非出现严重的错误，核保员可能永远无法知道某一特定保单的报价是过高还是过低。其他一些预测因为是有条件的，所以也无法验证。“如果我们开战，我们将被击溃”是一个重要预测，但它很可能不能验证。还有一种情况是：预测的时间跨度太大，以至于预测者无法验证结果，例如关于21世纪末全球平均气温的预测。

上述甘巴迪问题无法被验证的特征会影响你的判断吗？例如，你是否质疑甘巴迪是真实人物？你是否会在意下文能否告知你答案？你是否想过，即使告知你答案，你也无法解决当前所面临的问题？你可能没有想过这些问题，因为当你回答问题的时候，上述这些因素似乎无关紧要。

能否验证并不会改变你的判断经验。在一定程度上，当一个问题答案很快就被揭晓时，你可能会更加积极地思考，因为害怕出错，所以你的注意力更加集中。相反，你可能会拒绝对一个荒谬的假设性问题做太多思考，比如，如果甘巴迪有三条腿、能飞，他会是一个更好的CEO吗？总体而言，你会用解决现实问题的方式去解决一个假设性问题。这种相似性对心理学研究很重要，因为很多研究使用的都是这种假设性问题。

既然没有结果，你也不会去问是否有结果，那么按理说，你不会尽可能去减少判断与结果之间的错误，但你仍试图去做出正确的判断，给

出一个你有足够信心的数字。当然，你对自己的答案并没有十足的把握。你能意识到一些不确定性，并且正如我们所看到的那样，有可能存在一些超出你原本认知范畴的不确定性。但在某个时间点上，你决定不再进一步思考，而是给出最终答案。

是什么让你觉得自己做出了正确的判断？我们将这种感觉称为“做出判断的内部信号”（internal signal of judgment completion），这种内部信号与任何外部信息都无关。

如果你的答案与外部证据相契合，那么你就会觉得自己是对的。0或100这两个数字无法给你契合感：这两个数字所隐含的信心与现有的混乱、模棱两可、相互矛盾的证据是不匹配的。但是，不管你给出的是哪个数字，它都能够带给你所需要的一致感。正如你之前的经历告诉你的那样，判断的目标就是达到一致性的解决方案。

内部信号的本质特征在于：一致感是判断经验的一部分。它并不完全取决于真正的结果。因此，内部信号对于不可验证的判断与对于真实的、可验证的判断一样可用。这样就可以解释，为什么我们对于像甘巴迪这样的虚假人物做出判断的经验，就像是在真实世界中做判断。

评估判断的两种方法：结果和过程

能否被验证并不会改变判断的过程，但会改变事后对判断的评估。

结果可验证的判断可以由一个客观的观察者，根据一个简单的误差测量方式，即通过判断与结果之间的差异进行评定。如果天气预报员说今天的最高气温会达到21摄氏度，而事实是18摄氏度，那么我们可以说预报员犯了一个正3度的错误。显然，这种方法不适用于类似甘巴迪问题中的非验证性判断，因为这种判断没有真正的结果。那么，我们如何去界定一个判断的好坏呢？

答案是还有另一种评估判断的方法，这种方法既适用于可验证的判

断，也适用于不可验证的判断，而且其核心在于评估判断的过程。当我们在说好或不好的判断时，我们要么指的是判断的输出，比如甘巴迪的案例中你给出的数值；要么指的是判断的过程，也就是你是如何得出这个数值的。

评估判断过程的一种方法是，观察这一过程在应用于大量案例时的表现。以一个政治分析师为例，他在地方选举中预测了大量候选人的获胜可能性。他宣称，这些候选人中的100人有70%的可能性获胜。如果这100人中最终有70人当选，我们就会认为该政治分析师的预测很准确。虽然单个概率判断无法用正确或错误来评价，但其整体是可以被验证的。同样，通过审查大量案例的统计结果，我们可以更好地确定是否存在针对某一特定群体的偏见。

关于判断过程的另一个问题是，它是否符合逻辑或概率理论。大量关于判断中认知偏差的研究都在分析这个问题。

关注判断过程而不是结果，我们就可以对无法验证的判断进行评估，例如对虚构问题的判断或时间跨度很长问题的预测。虽然无法将其与已知的结果进行比较，但我们仍然可以评估这些判断是否正确。当我们将重心放在提升判断品质，而不仅仅是评估判断的好坏时，我们也会将注意力放在判断过程上。我们在本书中推荐的所有减少偏差和噪声的方法，都是为了实现这样的判断过程：能够在同类案例中从整体上最大限度地减少误差。

我们比较了两种评估判断品质的方法：一种是比较判断的结果，另一种是比较判断过程的品质。需要注意的是，对于可验证的判断，如果在单个案例中使用这两种评估方法，可能会得出不同的结论。一名熟练、谨慎的预测者在使用最好的工具和技巧对季度通胀进行预测时，也经常会出错。而在关于单个季度的预测中，即使是让黑猩猩掷骰子，也有可能给出准确的“答案”。

研究决策的学者为解决这一问题提出了清晰的建议：关注过程，而

不是单个案例的结果。然而，这并非现实生活中的惯常做法。专家们也经常评估自己的判断与可验证结果之间的一致程度，如果你问他们，他们判断的目标是什么，他们会说：尽可能地与结果一致。

总而言之，在可验证的判断中，人们通常声称，他们判断的目标是尽可能使自己的预判与结果一致。而事实上，无论是可验证的判断还是不可验证的判断，他们实际想要获得的是做出判断的内部信号，这种信号源自事实与判断之间的一致性。其实，他们应该追求的目标是：努力实现能够对一系列类似案件做出最佳判断的过程。

评估性判断vs预测性判断

到目前为止，本章主要集中讨论的是预测性判断的工作，随后讨论的也主要是这类工作。然而，在第1章中讨论的弗兰克尔法官以及美国联邦法官量刑中的噪声，考察的则是另一种类型的判断。给罪犯判重罪不是预测，而是“评估性判断”（evaluative judgment），目的在于使犯罪严重程度与刑期之间相匹配。葡萄酒比赛的评委和餐厅评论家做出的是评估性判断。对论文进行评分的教授、滑冰比赛的评委以及对科研项目进行评估以决定是否给予资助的委员会做出的判断也都是评估性判断。

在多项决策中对不同选项进行权衡也是一种评估性判断。类似的例子包括：经理们在一系列候选人中进行选择，管理团队在不同策略之间进行选择，以及总统们选择如何应对非洲的埃博拉疫情。可以确定的是，所有这些决策有赖于预测性判断来提供信息。例如，某位候选人第一年的表现如何；股票市场对某项战略举措的反响如何；如果放任不管，传染病会以多快的速度传播。然而，最终的决策过程需要在每个选项的优势和劣势之间进行权衡，而[这种权衡是通过评估性判断实现的](#)。

就像预测性判断，评估性判断也会出现一定范围内的不一致性。没有一位称职的美国联邦法官会说：“这是我最喜欢的判决，我根本不在

乎我的同事是否有其他看法。”从一系列策略中做出选择的决策者会有这样的预期：如果同事或其他人获得了相同的信息且具有相同目标，他们就会赞同自己的选择，至少不会偏离太远。评估性判断在一定程度上依赖于判断者的价值观和偏好，而不仅仅是个人品位或意见的问题。

可见，预测性判断和评估性判断的边界比较模糊，做出判断的人往往没有意识到二者间的差异。做出判决的法官和给论文评分的教授会努力思考，并力图找到“正确”答案。他们对自己的判断和做出判断的理由很有信心，在做出预测性判断（如：这款新产品销量如何）和评估性判断（如：我的助手今年表现如何）时，专家们感受相同、行为相同，当然也会以相同的方式阐述自己做出判断的理由。

“任意残酷行为”，噪声的最大问题

在预测性判断中存在噪声，则意味着哪里出错了。举个例子，如果两名医生在诊断上有分歧，或两名预测员对下一季度的销售额持不同意见，那么他们之中至少有一个人是错的。原因可能是其中一个人缺乏技能，也可能是存在其他噪声。不管是什么原因，错误的判断可能会给依赖诊断与预测行事的人带来严重的后果。

在评估性判断中存在噪声也会有问题，但原因不同。在司法系统中，如果法官是可以互换或随机分配的，关于同一案件的巨大分歧会违背人们对司法公正性和一致性的期望。如果对同一被告的判决差异很大，那么我们就是在做弗兰克尔法官所谴责的那种残酷专横之事。甚至那些信奉刑罚个别化的法官以及对抢劫犯的判决存有异议的法官也会认可：如果不同判决之间的差异太大，大到就像抽签一样，那就是有问题的。这一问题在其他场景中也存在，只是戏剧性没那么强：不同的教授对同一篇论文给出的评分差异巨大；不同的机构对同一家餐馆给出了不同的食品安全评分；不同的评委对同一位滑冰选手打出不同的分数。再比如，某人因患有抑郁症而获得了残障人士享有的社会保障，而另一

个状况相同的人则什么都没有得到。

即使不公平不是一个特别需要关注的问题，系统噪声也会带来另一个问题。受评估性判断影响的人期待这些判断反映的是系统的价值观，而不是个别法官的价值观。设想一下：一个客户抱怨笔记本电脑有缺陷，并得到了全额退款，而另一个客户仅仅收到道歉；一位在公司工作了5年的员工要求升职并获批准，而另一位绩效相同的员工则被婉拒.....这些都是很严重的问题。这样看来，系统噪声就是不一致，而不一致会损害系统的可信度。

噪声是可以测量的

只需对同一问题进行多次判断，我们就可以测量噪声，而且我们并不需要知道这一问题的真实值。正如引言中提到的射击的故事，当我们隐去靶子时，我们是看不见靶心的，但是可以看到子弹落点的分布情况。只要我们知道所有的射击手都是瞄准靶心的，那我们就可以测量噪声，也就是进行噪声审查。如果我们要求所有的预测者预测下个季度的销售额，预测的分散程度就是噪声。

偏差和噪声之间的这一区别对于改善判断品质非常关键。在我们无法验证判断是否正确的前提下，宣称可以改进判断听起来是自相矛盾的，但我们的确可以，因为只需从测量噪声开始。

无论判断的目标是获得精确的结果，还是在不同价值之间进行复杂的权衡，我们都不希望噪声存在。不过，我们通常可以对它进行测量。正如我们将会第五部分中讨论的那样，只要可以测量噪声，我们就可以减少它。

- 消除噪声

专业判断的本质

- 关于判断问题，你不要指望所有人完全一致。
- 是的，这是个判断问题，但有些判断太离谱了，所以是错的。
- 你对候选人的选择体现的只是你的个人喜好，这算不上严格意义上的判断。
- 决策既需要预测性判断，也需要评估性判断。

第5章

测量误差，噪声与偏差的代价一样大

显而易见的是，一致性的偏差会引发代价高昂的错误。如果体重秤在你每次称体重时都自动加上一定的重量，如果一位乐观的经理总是预测项目只需花费实际所需时间的一半，如果一位谨小慎微的经理总是年复一年地低估未来的销售额，那么后果都将会非常严重。

我们已经知道，噪声会引发代价高昂的错误。如果一位经理大多数时候预测的项目所需时间只是实际所需时间的一半，而在少数预测中又将前者估计成后者的两倍，那么我们是否可以说在平均水平上这位经理的判断是对的呢？答案是否定的。这些不同的误差是累加的，而不会互相抵消。

因此，我们想到了一个很重要的问题：偏差和噪声是如何引起误差的？多少偏差和噪声会引起误差？本章试图回答这些问题。这一章所要呈现的信息非常明确：在各种专业判断中，当以准确性为目标时，偏差和噪声在计算总体误差时扮演着相同的角色。在有些案例中，偏差是误差的主导因素；在其他案例中，噪声则是主导因素，而且这类案例比我们预想的更为常见。但在每一个案例中，每减少一个单位的噪声对总体误差的影响和每减少一个单位的偏差对总体误差的影响是一样的。因此，测量和减少噪声应该与测量和减少偏差同等重要。

这一结论是依据一种特定的误差测量方法得出的，这种测量方法由来已久，并且被科学界和统计学领域广泛接受。我们将在本章对其历史进行回顾，并简单介绍它的原理。

GoodSell应该减少噪声吗

假设有一家名为GoodSell的大型零售公司，这家公司雇用了一些销售预测师，他们的工作就是预测GoodSell在各地区的市场份额。可能是由于读过关于噪声的图书，预测部门的主管艾米·西姆金（Amy Simkin）进行了噪声审查。所有预测师都对同一个地区的市场份额进行了独立评估。

图5-1显示了噪声审查的结果（平滑得难以置信）。西姆金可以看到，这些预测分布在常见的钟形曲线（即正态分布，或称高斯分布）中。

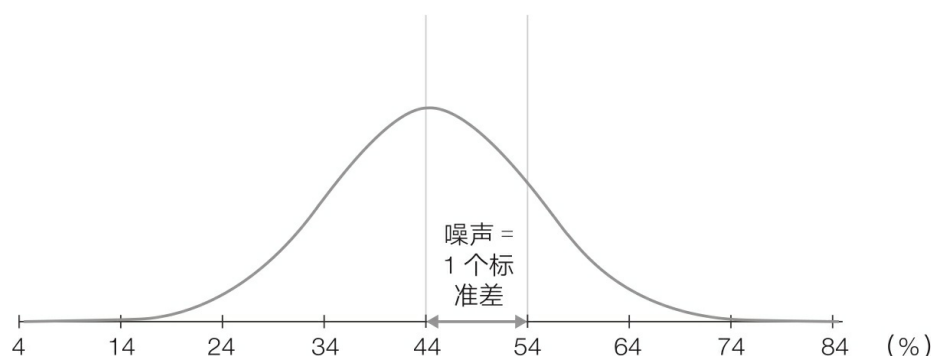


图5-1 GoodSell公司在某一地区市场份额的预测分布

频率最高的预测由钟形曲线的顶点所代表，市场份额为44%。西姆金可以看到，公司的预测系统具有很高的噪声：如果所有预测都是正确的，那么这些预测应该相同。但事实上，这些预测值的分布很分散。

我们可以给GoodSell预测系统中的噪声赋予一个数值。就像用秒表来测量间隔时间一样，我们可以计算这些预测的标准差。顾名思义，标准差表示一组数值偏离平均值的程度，在本例中为10%。对于每一个正态分布来说，大约2/3的预测值都落在偏离平均值正负一个标准差的范围内。本例指的是，约2/3的市场份额落在34%~54%这一范围内。西姆金现在获得了一个关于市场份额噪声的评估数据。需要提醒读者的是，为了获得更加稳健的评估，更好的噪声审查可以使用几个预测问题，但此处有一个问题就够了。

在第2章中，关于真实保险公司高管的案例也是如此。西姆金感到非常震惊，并决定采取措施。令人难以接受的噪声程度表明，预测师没有遵循他们应该遵循的程序。西姆金要求公司领导雇用一名噪声顾问来监督预测师的工作，从而提升他们的一致性。可惜，她的要求并未得到批准。领导的回复似乎也很合理：“当我们不知道自己的预测是对是错的时候，如何能够减少错误？如果平均起来错误很大（即存在一个较大的偏差），那它是一个需要尽快解决的问题。”最后，领导得出结论：GoodSell先需要确认这些预测师的预测是对是错，然后再谈改善预测品质的措施。

在那次噪声审查的一年后，预测师们所预测内容的真实结果出来了。目标地区的市场份额实际为34%。现在我们知道了每一位预测师的误差程度，即他们的预测值与实际结果之间的差异。如果预测值是34%，那么误差是0；如果预测值是44%，那么误差是10%；而对于很低的预测值24%，误差为-10%。

图5-2显示的是误差分布情况。它和图5-1中的预测曲线相同，不同的只是它的每一个数据点表示了将图5-1中相应的预测值减去真实值（34%）后的差。分布的形态没有改变，而且误差分布的标准差（我们对预测误差本身的噪声的测量结果）仍然是10%。

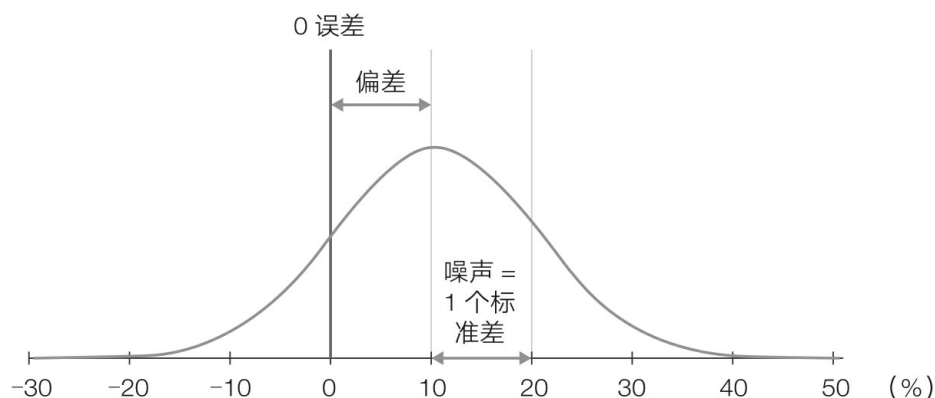


图5-2 GoodSell对某一地区市场份额预测中的误差分布

观察图5-1和图5-2之间的差异，就像图0-1和图0-2中从正面和隐去靶子观察射击规律之间的差异。在考察射击故事中的噪声时，我们无须知道靶子的位置，同样，知道真实结果对于预测噪声也无任何助益。

西姆金和她的领导现在知道了一些他们以前不知道的信息：预测的偏差数量。偏差是误差的平均值，在本例中是10%。因此，偏差和噪声在此案例的数据集中是一个相同的数字。（需要说明的是，噪声和偏差相等不是普遍的情况，但采用一个噪声和偏差相等的案例有助于读者理解它们各自的作用。）我们可以看到，大部分预测师都犯了过度乐观的错误，他们高估了公司可能获得的市场份额，大部分人的预测都落在了0误差线的右侧。事实上，结合该正态分布曲线的属性来看，大约84%的预测都高估了实际的市场份额。

面对上述结果，西姆金的领导自鸣得意地认为自己是对的，这些预测中存在着大量偏差！减少偏差固然是一件好事，但西姆金还是想知道，一年前的提议是明智的吗？就目前而言，再次提出减少噪声仍然明智吗？相比于减少偏差，减少噪声的意义是什么呢？

均方，衡量整体误差的精确规则

为了回答西姆金的问题，我们需要一个关于误差的评分规则，即对个体误差赋予不同权重并将其整合成测量总体误差的一个指标。幸运的是，这种工具确实存在，那就是[“最小平方方法”（method of least squares，也叫“最小二乘法”）](#)，它是由德国数学家高斯于1795年发明的。高斯是举世闻名的数学天才，生于1777年。他在十几岁时就做出了多项重大贡献。

高斯提出了一种方法，用于评估单个误差对总体误差的影响。他在测量总体误差时，使用的是“均方误差”（Mean Squared Error, MSE），即个体误差平方的平均值。

详细探讨高斯对总体误差的测量超出了本书的讨论范畴，而且他的解决方案并非直观易懂。为什么用均方呢？这是一个听起来有些奇怪的概念，但它建立在我们几乎所有的直觉上。

为了弄明白其中的原因，让我们来看一个看上去似乎完全不同但其实本质相同的问题。想象一下，你用一把尺子来测量一条线段的长度，要求精确到毫米，并且你可以测量5次。测量结果分别用图5-3中指向下方的三角箭头表示。

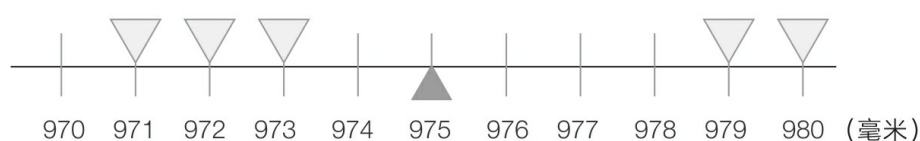


图5-3 对同一线段长度的5次测量结果

正如你看到的，5次测量结果都在971~980毫米这一范围内。那么，哪一个是对这条线段的最精确测量呢？一个可能是中位数，即两个最短的测量结果和两个最长的测量结果之间的那个测量结果，在本例中为973毫米。另一种可能是算术平均值，通常被称为平均值，在本例中为975毫米（图5-3中指向上方的三角箭头）。你可能倾向于认为平均值更为精确，你的直觉是对的。平均值包含更多信息，它受测量次数的影响，而中位数只受顺序的影响。

我们对实现最佳评估这一问题有很清晰的直觉，它与我们关注的对总体误差的测量关系紧密。它们实际上是同一个问题的两面，因为最佳评估能够使当前测量的总体误差最小化。因此，如果你凭直觉认为平均值是最佳测量结果，并且你的直觉没有错的话，那么用于测量总体误差的公式应该能够产生算术平均值，因为算术平均值的误差是最小的。

均方误差就具有这样的特征，它是唯一能够对总体误差进行测量的概念。在图5-4中，我们假定线段的真实长度有10个可能的整数值，进而计算5次测量的均方误差。例如，如果真实值是971毫米，那么5次测

量的误差就分别是0、1、2、8和9。这些误差的平方之和为150，均方为30。这是一个很大的数值，表明一些测量值离真实值很远。你会看到，当测量值接近975（平均值）时，均方误差会下降。随着测量值远离平均值，均方误差又会逐渐增加。可见，平均值是最佳评估结果，因为在这种情况下总体误差最小。

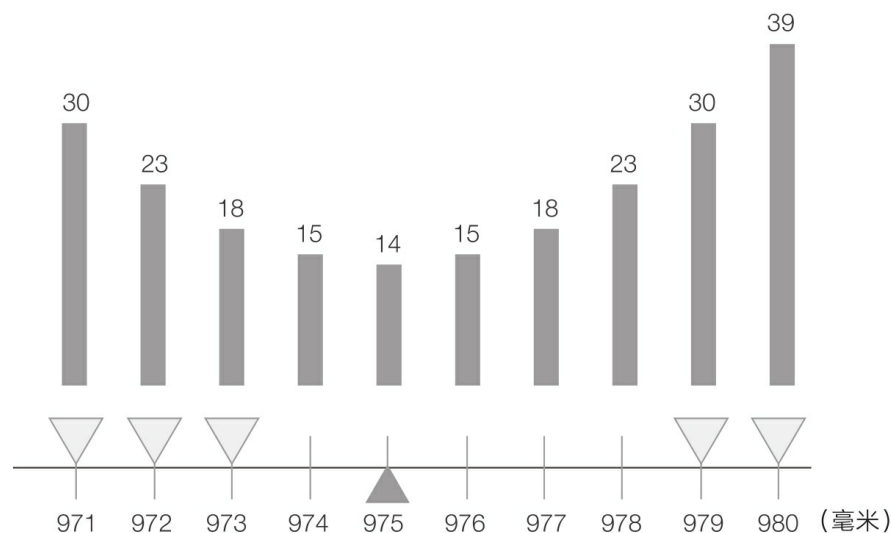


图5-4 10个可能的真实值对应的均方误差

你还会发现，当你的评估值远离平均值时，总体误差会迅速增加。例如，当你的评估值仅仅增加3毫米，如从976变化到979，均方误差就会翻倍。这是均方误差的关键特征：相比于小的误差，平方给大的误差赋予了更大的权重。

现在你应该明白了为什么高斯用于测量总体误差的公式被称为均方误差计算公式，以及这种评估方法为什么被称为最小平方法。对误差进行平方是其核心思想，其他任何公式都无法与我们的直觉——平均数是最佳评估值如此契合。

高斯的方法很快得到其他数学家的认可。作为其众多伟大成就之一，高斯用他的均方误差方法及其他数学创新解决了一大难题——重新发现谷神星（Ceres）。在此之前，这颗小行星只在1801年被短暂地追

踪到，之后便消失于太阳眩光中。高斯对这一问题的解决方案优于当时欧洲最好的天文学家。这些天文学家一直想方设法估算谷神星的轨道，然而他们测量望远镜误差的方法是错误的，这颗行星根本没有在他们预测的任何地点附近出现。高斯用最小平方方法重新进行了计算，当天文学家用望远镜对准高斯所预测的地点时，他们发现了谷神星！

不同领域的科学家很快就开始普遍采用最小平方方法来评估误差。两个多世纪过去了，当想要达到准确测量的目标时，最小平方方法仍然是评估误差的标准方法。用平方来赋予误差权重是统计学的核心。在绝大部分科学领域的应用中，均方误差方法处于绝对优势地位。正如我们将看到的，这一方法具有极高的应用价值。

单次测量中的误差

=

偏差

+

噪声误差

Error in a single measurement

=

Bias

+

Noisy Error

误差方程：无论偏差大小如何，减少噪声都有益处

偏差和噪声在误差中的作用很容易概括为两个表达式，我们将其称为误差方程。第一个误差方程将单次测量中的误差分解为你现在熟悉的两个部分：偏差（平均误差）和残留的“噪声误差”。

如果误差比偏差大，那么噪声误差是正的，反之则为负。噪声误差的平均数为0。这个误差方程并未提供什么新的信息。

第二个误差方程是对均方误差，即我们之前介绍的对总体误差测量的分解。均方误差可以简单表示为偏差和噪声的平方和。（回想一下，噪声是测量的标准差，它与噪声误差的标准差相同。）

下面的方程（两个平方之和）可能会让你想起一个学生时代常用的定理——勾股定理。或许你还记得，在一个直角三角形中，两条直角边长的平方和等于斜边长的平方。因此，误差方程式就更直观了，其中均方误差、偏差的平方和噪声的平方类似于直角三角形的三条边各自的平方。图5-5表明均方误差（黑色方块区域）的面积等于另外两个方块区域的面积之和。左图中噪声多于偏差，右图中偏差多于噪声。然而，两种情况的均方误差是相同的，均方误差的分解方程在这两种情况下都成立。

总体误差（均方误差）

=

偏差²

+

噪声²

Overall Error (MSE)

=

Bias²

+

Noise²



图5-5 均方误差的两种分解情形

正如上面的数学公式以及图5-5所示，偏差和噪声在误差方程中扮演了类似的角色，它们虽彼此独立，但被赋予了相同的权重。需要注意的是，在随后的章节中，我们在分析噪声的成分时也会用类似的平方和分解的方式。

误差方程解答了西姆金提出的问题。同等程度地减少噪声和偏差，对总体误差会产生什么影响？答案很明显：在误差方程中，偏差和噪声可以互换，因此无论是减少噪声还是减少偏差，对减少总体误差而言意义是一样的。在图5-2中，偏差和噪声刚好相等（都是10%），因而它们对总体误差的影响是等同的。

误差方程表明，西姆金最初想减少噪声的想法是正确的。无论何时发现噪声，你都需要想尽办法减少它。这一方程表明，西姆金的领导

所认为的“GoodSell应该在预测中的偏差的测量结果出来之后，再去减少噪声”的观点是错误的。对于总体误差而言，噪声和偏差是独立的：无论偏差的大小如何，减少噪声都有益处。

这个结论虽然很违反直觉，但非常重要。为了说明这一点，图5-6表明了减少相同数量的噪声和偏差所产生的效果。为了帮助你理解下图中左右两图的内容，最初的误差分布（来自图5-2）用虚线表示。

图 A：噪声不变，偏差减半

图 B：偏差不变，噪声减半

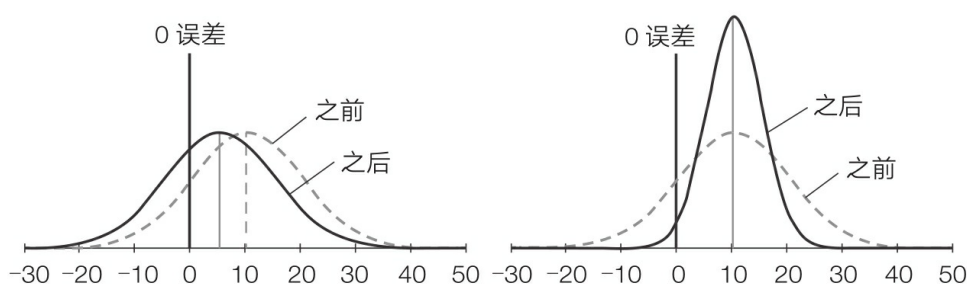


图5-6 偏差减半（A）与噪声减半（B）时的误差分布情况

在图5-6的图A中，我们假设西姆金的领导决定采用自己的方式：他发现了偏差，随后决定将其减半，如通过向过于乐观的预测师提供反馈的方式。他未对噪声采取任何措施。这种改进的效果是显而易见的：预测的总体分布更接近真实值了。

在图5-6的图B中，我们可以看到，如果西姆金的提议获得了领导的批准，其结果将会是：偏差没有改变，噪声减半。看似矛盾的是，噪声的减少似乎使问题变得更严重了——预测更加集中了（更少的噪声），而不是更准确了（并未减少偏差）。84%的预测落入真实值的一侧，几乎所有（98%）的预测都错误地高估了真实值。减少噪声似乎使预测更加不准确了——这肯定不是西姆金所希望的！

尽管看上去如此，但图5-6的图B中的总体误差和图5-6的图A中的总体误差减少的数量是一样的。图5-6的图B中情况变得更糟的错觉源自对偏差的错误直觉。测量偏差的目的并不是测量正误差和负误差之间的不

平衡，而是测量平均误差，即钟形曲线的顶点与真实值之间的距离。在图5-6的图B中，这一平均误差与原始情境相比并无差异——它仍然很高，占10%，但并没有更糟糕。的确，偏差变得更加显著，因为它占了总体误差中更大的部分——80%而不是50%，但这是因为噪声减小了。相反，在图5-6的图A中，偏差减少了而噪声没有。最终的结果是，图5-6的图A和图5-6的图B中的均方误差相同，也就是说，减少噪声和减少相同数量的偏差对均方误差的影响是相同的。

正如本案例所示，均方误差与我们对预测性判断进行评分的一般直觉相冲突。为了最小化均方误差，你需要尽可能避免大的误差。例如，如果你在测量长度，那么将误差从11厘米减少到10厘米的效果是将误差从1厘米减少至完全消失的效果的21倍。可惜，[关于这一点人们的直觉恰恰相反](#)：人们非常渴望一次性把问题全部解决，对小的误差高度敏感，但对两个大的误差之间的差异不敏感。即使你真心相信你的目标在于获得准确的判断，但你对结果的直觉反应与基于科学计算的准确性并不完全匹配。

当然，最佳的解决办法是既减少噪声，也减少偏差。既然偏差和噪声是彼此独立的，那就没有必要在西姆金和其领导的方案之间二选一。因此，如果GoodSell决定减少噪声，而减少噪声又可以使偏差更加清晰明了，那么这种选择就是正确的。也就是说，这确实是一件好事。减少噪声可以帮助公司进一步减少偏差。

然而，如果偏差远远大于噪声，那么减少噪声就不再是首要问题。GoodSell的例子给了我们另一个值得重视的教训。在上述简化的模型中，我们假定噪声和偏差是等同的。从误差方程来看，它们对总体误差的影响也是等同的：偏差和噪声各贡献了总体误差的50%。然而，正如我们所注意到的，84%的分析师会在同一个方向上犯错。如此之大的偏差（7个人中约有6个人朝同一个方向犯错）才产生了与噪声一样大的效果，因此在一些噪声比偏差更多的情境中，我们发现更大的误差就不足

为奇了。

我们在上文中用单个案例展示了误差方程的应用，这个案例就是预测GoodSell在某一地区的市场份额。当然，人们总是希望在多个案例中进行一次性噪声审查，方法是相同的：用误差方程计算各个案例的均方误差，然后对它们取平均值。均方误差就是偏差平方与噪声平方之和。对于西姆金而言，如果能得到多个地区的多个预测数据就更好了，无论它们是相同还是不同的预测师做出的预测。这些平均值能够让她对GoodSell的预测系统偏差和噪声有一个更清晰的认识。

噪声的代价

误差方程是本书的思想基础，它为减少预测性判断中的系统噪声提供了理论依据。原则上，减少预测性判断中的系统噪声这一目标与减少统计偏差同样重要。需要强调的是，统计偏差不是社会歧视的代名词，它只是一组判断中的平均误差。

误差方程和我们从中获得的结论均有赖于用均方误差来测量总体误差。这一规则适用于纯粹的预测性判断，包括预测和评估，它们都力求以最大的精度（最小的偏差）和最高的准确性（最小的噪声）来接近真实值。

然而，误差方程不适用于评估性判断，因为误差取决于真实值的存在，故而很难应用于评估性判断。此外，即使我们可以确定评估性判断中的误差，其代价也不太可能与它们的平方成正比。

例如，对于一家制造电梯的公司而言，评估一架电梯的最大负载的误差显然是不对称的：虽然低估意味着一定的代价，但高估可能会引发灾难。误差平方也不能用于评估“什么时候出发去赶火车”这样的决定，因为晚1分钟和晚5分钟的后果是一样的。在第2章中，当保险公司在为其保单估价或估计理赔额时，两个方向的错误都需要付出代价，但同样

没有理由假定这两个的代价是对等的。

上述这些例子充分表明：需要明确预测性判断和评估性判断在决策中所起的作用。关于良好决策，一个得到普遍认可的准则是：不应混淆自己的价值判断和事实。决策需要根据客观、精确的预测性判断做出，这些判断不应受到你的希望与恐惧、偏好与价值取向的影响。对于电梯公司而言，第一步是利用不同的技术解决方案对电梯的最大技术负载进行客观计算。安全性仅仅在第二步才需要被重点关注，即根据不同的安全边际来设定最大负载时。可以肯定的是，该选择在很大程度上取决于事实判断，诸如设置不同安全边际的代价与收益。同样，决定何时出发去火车站的第一步应该由客观的旅行时间决定，至于错过火车的代价和提前到火车站所浪费的时间，只有在你需要决定甘愿冒多大风险时才成为你需要考虑的因素。

同样的逻辑适用于会产生更严重后果的决策。在决定是否发动军事进攻时，指挥官需要权衡一系列因素，但他依赖的最重要信息是预测性判断。政府官员在对公共健康危机做出回应时，需要权衡不同选项的优势和劣势，如果缺少对每一选项的可能后果的准确预测，这种评估就无法完成。

在所有这些案例中，最终的决策都需要进行评估性判断。决策者需要考虑多种选项，并根据其价值做出最佳选择。但决策取决于潜在的预测，而这种预测应该是价值中立的。他们的目标是精确性（尽可能击中靶心），均方误差是测量误差的恰当手段。我们可以通过减少噪声来改善预测性判断，如果这样做不会大幅增加偏差的话。

- 消除噪声

误差方程的启示

- 减少偏差和减少噪声对准确性的影响是一样的。
- 不管偏差如何，在预测性判断中减少噪声总是有用。
- 如果判断值中有84%比真实值高，有16%比真实值低，就表明存在较大的偏差——此时，偏差和噪声数量相等。
- 预测性判断是每个决策的重要组成部分，精确度是其目标。要将个人的价值取向和事实分开。

第6章

噪声分析：所有判断都存在3类噪声

在上一章，我们讨论了单个案例中测量或判断的变异性。在单个案例中，判断的所有变异性都是误差，而误差由偏差和噪声组成。我们考察的判断系统，包括法院和保险公司的判断系统，它们的目的在于处理不同案例，并对这些案例进行区分。如果美国的联邦法官和保险理赔员对他们手头的所有案例都做出相同的判断，那他们就没有什么存在的价值了。对不同案例做出有差异的判断，在很大程度上是有意为之。

然而，人们并不希望在同一案例上存在判断的差异——系统噪声。你将看到，对同一批人在多个案例上做出的判断进行噪声审查，有利于更详细地分析系统噪声。

判决中的噪声审查

为了说明多起案件的噪声分析结果，我们对美国联邦法官的判决进行了[非常详细的噪声审查](#)。这一分析发表于1981年，也是我们在第1章介绍过的量刑改革运动的一部分。这一研究局限于判决，但我们能从中获得的经验是通用的，也适用于其他专业判断。进行噪声审查的目的不是检查弗兰克尔法官等人收集的逸闻趣事中的噪声，而是更系统地“确定量刑差异的程度”。

研究人员设计了16起虚构案件。在这些案件中，被告都犯下了罪行，即将接受审判。这些案件主要是关于抢劫犯或诈骗犯的，但他们在其他6个维度上存在差异，如被告是主犯还是从犯、是否有犯罪记录、是否使用武器（对抢劫犯而言）等。研究人员精心组织了面向美国全国范围内208名现任法官的结构性访谈。在90分钟的访谈中，[研究人员向](#)

[这些法官呈现16起案件的详细文件，并要求法官们做出判决。](#)

图表可能更有助于我们对研究的理解。请想象有一张大表格，表格中有16列，对应从A到P的16起案件；有208行，对应208名法官。从A1到P208的每一个单元格对应一位法官对每起案件的量刑。图6-1展示了这3328个单元格的表格样貌。为了研究噪声，我们需要将注意力集中在这16列上，其中每一列对应一个独立的噪声审查。

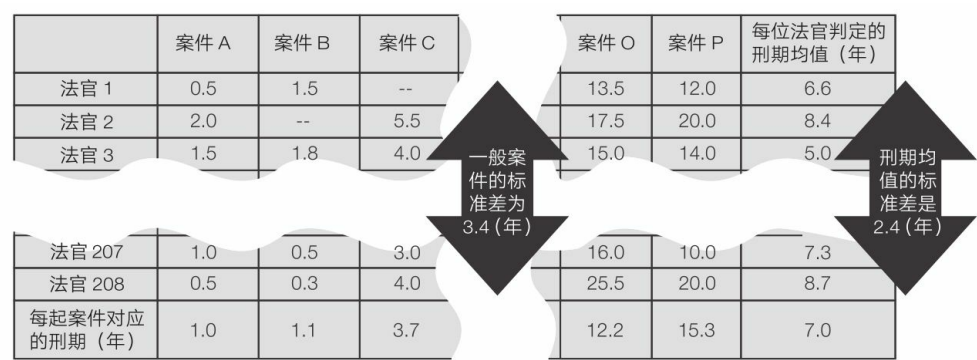


图6-1 判决研究示例

平均刑期，每一起案件的判决都存在大量噪声

由于没有办法客观地确定某一起案件量刑的“真实值”，所以接下来，我们将208名法官对每起案件判定刑期的均值作为对该案件的“公正”判决。正如我们在第1章中指出的，美国量刑委员会在将以往审判中的平均值作为制定量刑指南的基础时，也以同样的假设为前提。刑期均值假定，对每起案件对应的刑期取平均值，那么这起案件中的偏差为0。

我们很清楚，这一假设在现实生活中是错误的：一些案件对应的平均刑期与其他高度相似的案件的平均刑期相比，可能存在较大偏差，比如也许是种族歧视等因素所致。不同案件中存在的偏差也不尽相同，有些是正偏差，有些是负偏差，它们是造成误差和不公平的主要成因。令人困惑的是，[这种差异经常被称为“偏差”](#)。本章以及本书分析的焦点都

是噪声，它是误差的另一个成因。弗兰克尔法官强调了噪声导致的不公正，同时也引起了人们对偏差（包括种族歧视）的关注。同样，我们强调噪声的影响，并不代表我们认为测量和消除偏差不重要。

为了方便起见，我们在表格的最后一行给出每起案件对应的刑期均值。这些案件按照严重程度从低到高排列：案件A对应的平均刑期为1年；案件P对应的平均刑期为15.3年。所有16起案件对应的[平均刑期为7年](#)。

想象在一个完美的世界中，所有法官都是完美无瑕的司法公正的代言人，他们的量刑中不存在任何噪声。那么在这样一个世界中，图6-1看起来会是怎样的呢？显然，对于案件A来说，所有单元格中的刑期应该完全相同，因为每位法官都会判处被告1年刑期。其他案件的情况也是如此。当然，每一行中的数字仍然存在差异，毕竟案件不同，但每一列中的数字应该是相同的。不同案件对应的刑期之间的差异应该是导致该表格中出现变异性的唯一因素。

可惜，美国联邦法院并不存在于完美的世界中。法官们的量刑并不相同，每一列的数字之间的差异巨大，这意味着每一起案件的判决中都存在噪声。量刑中存在不应该存在的巨大变异性，而该研究的目的就在于对这种变异性进行分析。

判刑就像抽签

假设我们确实生活在上述完美世界中，每位法官对每起案件都会做出相同的判决，那么每列中的208个数字都应该相同。现在，为每一列从上到下增加噪声，并通过不时地在平均刑期上增加或减少刑期来改变一些数字。因为你的修改并不总是相同的，它们在每一列中都制造了变异。这种变异就是噪声。

本项研究的重要发现是，每起案件的判决中都存在大量噪声。我们

可以通过计算每起案件对应刑期的标准差来测量噪声。这些案件对应的平均刑期为7年，标准差为3.4年。

或许你对标准差非常熟悉，但采用具体的描述或许更有助于理解。假设你随机选择两位法官，并计算他们在对同一案件的判决中存在的差异。重复这个过程，计算所有案件中任意两位法官的判决差异，取平均值。这种测量方法——计算平均绝对离差，会让你觉得被告就是在法庭上“抽签”。假想这些判决呈正态分布，平均绝对离差就是标准差的1.128倍，这表明，两位随机选择的法官对同一案件量刑的平均绝对离差为3.8年。在第3章中我们谈到，保险公司的客户在寻找专业核保员时，就像在抽签。只不过在刑事判决中，被告所面临的类似抽签的状况，其后果更加严重。

在平均刑期为7年的情况下，不同法官之间的平均绝对离差达到3.8年，这令人不安，也是难以接受的。然而，我们有充足的理由相信，现实的执法中甚至存在更严重的噪声。首先，噪声审查中的被试处理的是虚构的案件，这些案件同时放在一起，很容易比较。现实世界中几乎不可能具备如此多的保持判断一致性的因素。其次，在法庭上，法官拥有的信息要比这项研究的信息多得多。除非新的信息是决定性的信息，否则它们更有可能导致法官们的分歧。因此我们认为，在实际的庭审中，被告面临的噪声要比我们在这里看到的多得多。

水平噪声

在接下来的分析中，我们会将噪声分解成不同成分。对于噪声，你脑海中的第一个念头可能是，噪声是由法官们在量刑时的严厉程度不一所致，就像弗兰克尔法官认为的那样。正如一些辩护律师告诉你的：法官们各有特点。有的人是“铁面判官”，他们比一般法官更严厉；有的人是“柔情法官”，他们比一般法官更仁慈。我们将这些差异称为“水平误差”（level errors）。（再次提醒：这里的误差是指判决结果与平均值之

间的差异；如果判决结果的平均值是错的，一个误差可能反倒纠正了这种不公正。）

任何判断都会存在不同程度的水平误差。例如，在进行绩效评估时，一些主管比其他人更宽容；在预测市场份额时，一些预测者比其他人更乐观；在是否需要实施背部手术时，一些骨科医生比其他医生的态度更积极。

在图6-1的表格中，每一行显示的是同一位法官对不同案件的判决结果。每一位法官判定的平均刑期展示在表的最右侧一列，这一列的数据反映了每一位法官的严厉程度。正如该表格所示，在这一维度上，法官之间的差异很大。表格最右侧这一列数字的标准差为2.4年。这种差异与公正无关。事实上，正如你所怀疑的那样，每位法官在平均刑期上的差异反映了不同法官在其他特征上的差异，例如他们的背景、生活阅历、政治倾向、偏见等。研究人员还考察了每位法官对于量刑的根本态度，比如，法官认为量刑定罪的主要目的究竟是让罪犯无法再危害社会（与社会隔离），还是对罪犯进行改造，抑或是对犯罪起威慑作用。研究人员发现，与持后两种观点的法官相比，持第一种观点的法官判处的刑期更短，监视居住的时间要长。此外，美国南部的法官比美国其他地区的法官判定的刑期明显要长。这说明，刑罚的严厉程度与保守的意识形态有关。

研究人员从中得出的一般性结论是，量刑的平均水平就像人格特征。你可以根据这项研究结果，将法官按“非常严厉”到“非常宽容”进行排序，就像用人格测试来衡量人的外向或平易近人程度。像其他人格特征一样，我们推测，量刑的严重程度与基因、生活阅历以及人格的其他方面相关，但这些因素与法官们审判的案件或被告无任何关系。我们用“水平噪声”（level noise）一词来代表每位法官所判处的平均刑期之间的变异性，这与水平误差的变异性是等同的。

模式噪声

如图6-1中黑色箭头所示，该案例的水平噪声是2.4年，[系统噪声是3.4年](#)。这两者之间的差异表明，除了法官们在平均刑期上的差异外，系统噪声还包括了其他成分，我们将这一成分称为“模式噪声”（pattern noise）。

要想理解模式噪声，我们再看一看图6-1，随机对一个单元格进行仔细观察，比如选择单元格C3。案件C对应的平均刑期在这一列的底部，你可以看到它是3.7年。现在，看一下最右侧的一列，找到3号法官对所有案件判定的刑期均值：5年。这个刑期均值比总体刑期均值少2年。如果法官严厉程度的变异是第3列中所有噪声的成因，你可能会预测，C3单元格中的刑期是 $3.7 - 2 = 1.7$ 年。然而，实际上C3中的数字是4年，这表明3号法官在这一案件上表现得尤为严厉。

使用同样[简单的相加逻辑](#)，你可以推测表格中每一列的刑期。但事实上你会发现，大部分单元格中的数字与推测的数字之间存在差异。观察每一行，你会发现，法官们在自己审理的所有案件中并非表现得同样严格：在有些案件上，他们比自己量刑的平均水平严格；但在其他案件上，他们则表现得要宽容。我们将这种残存的变异称为“模式误差”（pattern errors）。如果你在表格的每一个单元格中写下这些模式误差，就会发现：对于每一位法官（行）而言，这些模式误差之和为0；对每一起案件（列）而言，它们的和也是0。但模式误差对噪声的影响不能相互抵消，因为在计算噪声时，使用的是每个单元格中数值的平方。

有一个简单方法可以证明，刑期的简单相加模型并不适用于计算噪声。你可以看到，表格的每一列底部的刑期均值从左到右依次增加，但每一行的情况却并非如此。例如，208号法官对案件O中的被告判定的刑期比案件P中的更长。如果根据每位法官判定刑期的长短排序，那么

在他们的排序中，每起案件所在的位置会产生不同。

我们用“模式噪声”一词来表示上文提到的变异性，因为这种变异性反映了法官们对具体案件的态度复杂模式。例如，一位法官总体来说比其他法官更严格，但对于白领罪犯更宽容。另外一位法官可能总体上倾向于从轻处罚，但对于惯犯更严格。还有一位法官的宽严程度跟一般法官差不多，但如果被告只是从犯，他会表现出更多的同情，而如果受害者是老年人，这位法官会更严厉。我们用“模式噪声”这个说法是为了增加可读性。模式噪声的统计术语是“法官×案件的交互作用”——可以读作“法官乘以案件”。我们要对受过统计学训练的人说声抱歉，因为对他们而言无须进行这些解读。

在刑事司法的背景下，对案件的一些独特反应可能体现了法官的个人量刑哲学。其他的反应可能来自法官本人没有意识到的联想，例如，被告让他想起一个特别可恨的罪犯，或是被告长得像他的孩子。无论原因是什么，这些模式并不是偶然出现的。我们可以推测，如果重审同样的案件，他们还会做出相同的反应。然而，模式噪声在实践中很难预测，这导致本已像抽签一样的审判更加充满不确定性。这项研究的作者指出：“法官受犯罪情节或被告特征的影响而导致[审判中出现模式化的差别](#)”，是“同罪不同罚的另一种形式”。

你可能已经注意到，将系统噪声分解成水平噪声和模式噪声，与第5章提到的误差方程（将误差分解为噪声和偏差）遵循的是同样的逻辑。对噪声方程的具体表述请见下面。

系统噪声²

=

水平噪声²

+

模式噪声²

System Noise²

=

Level Noise²

+

Pattern Noise²

这个方程也可以采用误差方程那样的视觉表现形式（见图6-2）。我们将三角形的两条直角边表现为相等的长度。这是因为，在审判研究中，[模式噪声和水平噪声对系统噪声的贡献几乎相同](#)。

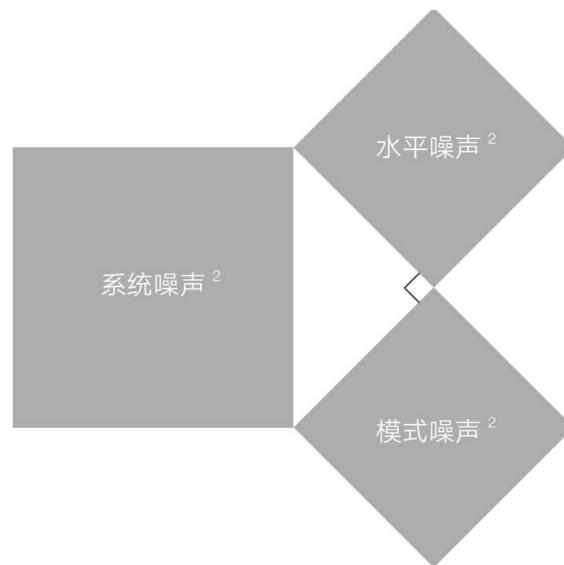


图6-2 分解系统噪声

模式噪声无处不在。医生决定患者是否需要住院，公司决定雇用谁，律师决定受理哪些案件，好莱坞的高管决定制作哪些电影.....所有

这些案例中都会存在模式噪声，不同的判断者会做出不同的判断。

噪声的成分

我们对模式噪声的处理掩盖了一个重要的复杂性：随机误差可能对模式噪声产生影响。

回想一下秒表实验：你连续测量10秒的时间长度，但每一次的测量结果与前一次都可能存在差别，此时你表现出了个体内的差异。同样，如果要求法官们在其他场景中再次为16起案件量刑，他们做出的量刑结果也不可能和第一次完全相同。事实上，正如你所见，如果该实验分别在同一个星期内的不同两天开展，他们做出的量刑结果也可能不同。如果法官因为种种原因（比如他的孩子有好事发生，他钟爱的球队昨晚赢得了比赛，或是天气很好）而心情愉快，那么他的判决会比其他时候更宽容。个体内的这种变异性与我们已经讨论过的个体间的稳定变异不同，但我们很难将这两种变异的成因区分开来。我们将这种转瞬即逝的因素所导致的变异称为情境噪声。

我们有意忽视了本研究中的情境噪声，而选择将噪声审查中法官独特的审判模式解释为稳定的态度。这一假设无疑过于乐观，但我们有理由相信，情境噪声在本研究中并未产生很大影响。参与这项研究的都是经验丰富的法官，他们对被告及其所犯罪行各项特征的意义已经具有一些固定的看法。在下一章，我们将更详细地讨论情境噪声，并展示如何将其与稳定的模式噪声进行区分。

总而言之，我们讨论了几种类型的噪声。系统噪声是由多个个体对同一案例进行判断时产生的不必要的变异。我们发现了噪声的两个主要成分。当同一批人评估多个案例时，我们就可以对这两种成分进行区分。

- 水平噪声是指不同法官做出的判决与平均值之间的变异。
- 模式噪声是指法官对特定案件做出的反应的变异。

在本研究中，水平噪声和模式噪声的数量几乎相等。然而，模式噪声中肯定会包含一些情境噪声，后者可以被看作随机误差。

我们借由在司法系统中进行的噪声审查对此进行了说明，同样的分析也适用于其他所有领域，比如商业、医疗、政府管理等。水平噪声和模式噪声（包括情境噪声）都会导致系统噪声，我们随后将多次介绍它们。

- 消除噪声

在现实世界中，噪声无处不在

- 法官之间表现出的不同严厉程度会产生水平噪声。当法官们对被告应该受到更严厉还是更宽松的惩罚意见不一时，就产生了模式噪声。模式噪声中有一部分是情境噪声——法官们自身的行为出现了不一致。
- 在一个完美的世界中，被告面临的将是公正的审判；但在现实世界中，他们面对的是有噪声的系统。

第7章

情境噪声，无时无刻不在影响着我们的判断

想象一名职业篮球运动员正在准备罚球的场景：他站在罚球线上，全神贯注地准备投篮。这是他练习过无数次的一系列动作，他能投中吗？我们无法预知结果。

在NBA的比赛中，球员们通常能够4罚3中。显然，一些球员比其他球员表现得更好，但没有人能做到百发百中^③。虽然篮筐一直都是距离地面3.05米，与罚球线的垂直距离一直是4.6米，篮球也一直都是重624克，精确地重复能够得分的一系列肢体动作却并非易事。变异不仅存在于球员之间，也存在于球员自身，这与我们推测的情况一致。罚球就像抽签，虽然库里比奥尼尔投中的概率更高，但依然像是在抽签。

这种变异来自何处？我们知道，很多因素都会影响球员的罚球命中率，比如长时间比赛的疲劳、比分接近时的心理压力、主场的欢呼声、对方球迷的嘘声。如果像库里或纳什这样的球员没投中，那么我们会做出如上的解释，但事实上，我们不太可能知道这些因素到底起了什么作用。因此可以说，球员表现中的变异性就是一种噪声。

第二次抽签

在罚球或人类的其他生理过程中存在变异性，这不足为奇。我们习惯了身体里的变异性，比如我们的心率、血压、下意识的反应、音调、手的颤抖等在不同时间都是不同的。无论我们多么努力地想要写下相同的签名，但每一张支票上的签名仍略有不同。

想要观察人类思维上的变异性尤其不易。当然，即使没有新的信息，我们也都有改变主意的经历，比如，前一天晚上令我们捧腹大笑的电影，此时看起来就可能平庸无奇，很容易淡忘；前一天受到严厉审判的那个人，现在看来应该得到宽恕；一个我们曾经不喜欢或不理解的观点，现在觉得很有道理。但是，正如这些例子所表明的，我们通常会将这种改变与相对没那么重要且非常主观的事情联系起来。

在现实世界中，我们也会在缺少有说服力的理由时改变主意，即使是对判断持高度审慎态度的专业人士也是如此。例如，同一位医生先后两次对同一病例做出的诊断常常会有显著的不同（见第22章）。在美国的一个大型葡萄酒比赛中，[专家先后两次品尝了同一种葡萄酒](#)，然而他们给出相同评分的可能性也只有18%，而且通常是对最差的酒打出的。仅仅相隔几个星期，指纹鉴定专家先后两次检查同一枚指纹时，也常常会得出不同的结论（见第20章）。[经验丰富的软件顾问](#)会在两种不同的情况下对同一任务的完成时间做出明显不同的估计。简而言之，篮球运动员不可能在两次投篮时采用完全相同的姿势，我们也并不总能在两种情境中对同一事实做出相同的判断。

至此，我们描述了像抽签一样“选择”核保员、法官或医生时所产生的系统噪声。情境噪声则类似于第二次抽签，这次抽签“选择”了专家做出判断的时机、他们的心情、他们以前是否遇到过类似案件，以及情境中的无数其他因素。第二次抽签往往比第一次抽签更抽象。如果在第一次抽签中“选择”了不同的核保员，我们是容易看到“选择”的结果，但要了解该核保员“实际反应”的其他可能性，则有赖于抽象的反事实思考。我们只知道这一判断是从一系列可能的判断中挑选出来的，而情境噪声是这一系列不可见的可能性产生的变异。

测量情境噪声

测量情境噪声并不容易，因此一旦检测到它的存在，人们会感到惊

讶不已。一旦人们形成了一个经过审慎思考的专业意见，他们就会找理由支持该意见。如果有人要求他们对自己的判断做出解释，他们就会以自认为具有说服力的论据来捍卫这一判断。如果他们再次遇到同样的问题，并且确定自己以前遇到过，他们就会给出与第一次相同的意见，以此来减少认知负担并维持一致性。设想一个关于教师的例子：如果一位教师给一篇作文评了非常高的分数，一个星期之后他又读了这篇作文并看到了第一次的评分，那么他这一次给出的评分不太可能与上一次差别很大。

因此，只要案例容易记住，直接测量情境噪声就很难。比如，你给一名核保员或法官展示一个他们以前处理过的案件，他们可能会回想起这起案件，从而重复自己以前的判断。一篇关于专业判断中存在变异性⁽⁴⁾的评论文章中提到了一些研究。这些研究发现，专家在同一环节做出了两次相同的判断，毫无疑问，[他们倾向于与自己保持一致](#)。

我们上文提到的实验使用了专家们不能再次辨认出的实验材料，从而绕开了这一问题。葡萄酒评委进行的是盲测；给指纹鉴定专家呈现的是他们已经鉴定过的一对指纹；给软件专家呈现的也是他们已经完成的项目。第二次任务是在几个星期或几个月之后进行的，并且实验人员没有透露这是专家们曾经鉴定过的案例。

还有另一个比较间接的方法可以确认是否存在情境噪声：借助大数据或计量经济学。当可以获得以往大量专业性决策的案例时，分析师们有时就能检验这些决策是否会受到特定情境或无关因素，比如一天的某个时间段或室外温度等的影响，如果这些无关因素产生了统计上的显著效应，则说明存在情境噪声。实事求是地说，不要指望发现可能导致情境噪声的所有外部因素，那些我们已经发现的因素就足以说明存在大量多种多样的因素。如果要掌控情境噪声，我们就必须了解它的产生机制。

群体智慧效应：一个是一群

思考下面这个问题：美国的机场数量占全世界机场总数的百分比是多少？当你思考的时候，一个答案可能会跳进你的脑海，但这个答案产生的方式与你回忆自己的年龄或电话号码时不同。你会注意到，你脑海中出现的数字是一个估计值，当然，它不是一个随机数字——1%或99%显然不是正确的答案。你给出的数字是一系列不可能被排除的可能数字中的一个，如果有人将你的答案加上或减去1%，你可能不会觉得这个答案比你给出的答案的准确性差。[正确答案是32%](#)。

[爱德华·沃尔（Edward Vul）和哈罗德·帕什勒（Harold Pashler）](#)这两位研究人员想到，可以要求人们回答这一问题以及其他类似的问题两次，而不是一次，并确保被试事前并不知晓他们需要第二次作答。沃尔和帕什勒假设，对两次答案取平均数会比单次答案（无论是第一次还是第二次的回答）的准确性更好。

结果证明他们是对的。总体而言，[第一次的答案比第二次的答案更接近真实值](#)，而最准确的估计值是两次推测的平均数。

沃尔和帕什勒的灵感来自一个非常著名的现象——群体智慧效应，它是指对人们的独立判断进行平均会提高判断的准确性。1906年，达尔文的表弟、著名博学大师弗朗西斯·高尔顿（Francis Galton）在一次乡村集市上目睹了787名村民估计一头获奖的公牛的重量。没有一位村民准确猜到公牛的实际重量——大概是1198磅^⑤，但他们猜测的平均值是1197磅，与实际值仅相差1磅。村民们估值的中位数（1207）与实际值也非常接近。虽然村民个体的估值充满噪声，但全部估值的平均值偏差极小，就这一点而言，村民们是智慧的群体。这一结果让高尔顿备感吃惊，因此，虽然他对普通人的判断并不看好，但他还是极力主张“民主判断的结果，比想象中更值得信赖”。

类似的结果是从成百上千个情境中获得的。当然，如果问题很难，

只有专家才能够获得相对准确的答案，那么群体的答案就不会更准确。但是，当让人们猜测透明罐子里夹心糖的数量、预测一个星期后当地的气温，或是估计美国一个州内两个城市之间的距离时，一大群人的平均答案可能更接近事实。出现这种情况的原因是一个基本的统计事实：对多个独立判断或测量进行平均会产生一个新的判断，虽然这一判断不一定会比个体判断产生更少的偏差，[但它会产生更少的噪声](#)。

沃尔和帕什勒想要看看同样的结果是否也适用于情境噪声：将同一个体的两次猜测进行平均，是否就像将不同人的猜测进行平均一样更接近真实值？沃尔和帕什勒的发现表明，答案是肯定的，他们给这个发现起了一个令人印象深刻的名字：内部群体（the crowd within）。

对同一个体的两次猜测进行平均，对于决策品质的提升效果并没有对两个独立个体的意见进行平均时更好。正如沃尔和帕什勒所指出的：“[你对同一个问题做出两次回答带来的好处，是向另一独立个体寻求建议时所获好处的1/10](#)。”显然，这种改进并不大，但如果你能等待一段时间再去做第二次猜测，你所获得的好处会大得多。沃尔和帕什勒在3个星期之后再问被试同样的问题，此时这种做法带来的好处上升到向另一个人寻求意见带来的好处的2/3。在无须额外信息以及外在帮助的情况下，这种方法已经非常不错了。这一结果明显为一种古老的决策智慧提供了理论依据，那就是：“睡一觉吧，明天再说。”

德国研究人员[斯蒂芬·赫佐格（Stefan Herzog）](#)和[拉尔夫·赫维格（Ralph Hertwig）](#)差不多与沃尔和帕什勒同时提出了另一种基于相同原则的方法。赫佐格和赫维格不是要求人们做出第二次评估，而是鼓励人们尽可能想出其他可能，从而做出与第一次评估不同的评估。这样一来，被试就会积极思考他们在第一次评估时没有考虑到的相关信息。具体做法如下：

首先，假设你的第一次评估是不正确的。其次，思考导致

第一次评估出错的几个可能原因——哪些假设或考虑的问题错了？再次，这些需重新考虑的问题意味着什么？第一次评估是过高还是过低？最后，根据这些新的观点做出不同于第一次评估的第二次评估。

与沃尔和帕什勒一样，赫佐格和赫维格会对前后两次评估进行平均。赫佐格和赫维格将这一技术命名为“自我重复抽样法”（dialectical bootstrapping）。相比于只让被试在第一次评估之后立刻进行第二次评估，这种方法大大提升了判断的准确率。因为被试迫使自己从一个新的视角考虑问题，他们创建了另一个样本，选取了另一个版本的自己，进而构成了两个互相分离的“内部群体”成员。因此，二者的平均产生了一个更准确的评估。两个即刻、连续进行的辩证性评估在判断准确度上的获益，是听取他人意见时的一半。

正如赫佐格和赫维格总结的那样，决策者最终要做的是在不同方法之间做出一个简单的选择：如果你能从别人那里得到独立的意见，那么就去做吧，因为这种真正的群体智慧很可能会提升你的判断水平；如果不能，你可以再次做出判断，以此创造一个“内部群体”。针对后者，你可以采用两种方式：要么隔一段时间再做出第二次判断，要么质疑自己的第一次判断，从另一个角度来看待问题。此外，不管是哪一种类型的“群体”，除非你有充足的证据表明需要对其中一次评估赋予更高的权重，否则对两次判断进行平均后的判断就是最佳判断。

除了实用的建议之外，这一系列研究还确认了一个关于判断的本质观点。正如沃尔和帕什勒指出的那样：“[被试的反应是从一个内部的概率分布中抽取的](#)，并不是根据被试掌握的全部知识来确定的。”这一发现与你回答美国机场问题时的过程相一致：你的第一个答案并不是根据你的全部知识甚至那些关键的知识得出的，你给出的答案只是你头脑中可能产生的一系列答案中的一个。我们从同一个人对同一个问题的判断中观察到的变异性，并不是一些高度专业的问题中存在的偶然现

象，事实上，情境噪声无时无刻不在影响着我们的判断。

情绪，情境噪声的源头

我们至少能够注意到一种情境噪声的来源：情绪。每个人都有过情绪影响判断的经历，而且你肯定意识到了，他人的判断也会受到情绪的影响。

情绪对判断的影响已经成为大量心理学研究的关注点。我们很容易通过一些策略来让人们暂时开心或悲伤，然后测量相应的情绪被诱发之后他们的判断和决策的变异性，比如让人们写下一段开心或悲伤的记忆，或只是简单地让他们看一段有趣的或催人泪下的电影片段。

一些心理学家花费了数十年时间来研究操纵情绪可能产生的效果。其中最高产的也许是澳大利亚心理学家[约瑟夫·福加斯（Joseph Forgas）](#)，他发表了大约100篇关于情绪的科学论文。

福加斯的一些研究验证了你的思考：一般来说，具有良好情绪的人会更加积极。相比于悲伤的记忆，他们更容易回忆起愉快的记忆，而且他们更倾向于赞同他人，更慷慨，也更乐于助人。消极情绪的作用则刚好相反。正如福加斯所说：“[对于同一个微笑，拥有积极情绪的人看到友好](#)，而拥有消极情绪的人可能感到不适。一个人情绪良好的时候，讨论天气会兴致盎然，而当他情绪不好的时候，该话题则显得枯燥乏味。”

换句话说，情绪会对你的思维产生可预见的影响，它关系到你会注意环境中的哪些事物、从记忆中提取哪些信息，以及如何解读这些信息。情绪还会产生另外一个更令人吃惊的效应——它会改变你的思维方式，这种效应甚至超乎你的想象。好的情绪是一把双刃剑，不好的情绪也可能成为困境中的一线希望，不同情绪的利与弊取决于具体的情境。

例如，在谈判中，好的情绪大有裨益。具有良好情绪的人会更富有合作精神、更容易促成互利互惠的局面。最终，他们往往能比情绪不好的谈判者获得更好的谈判结果。当然，成功的谈判也能让人们更开心，但在这些研究中，情绪并不是由谈判过程产生的，而是在谈判之前被诱发的。此外，[谈判过程中从情绪良好转向愤怒的谈判者也可能会获得更好的结果](#)——当你遇到一个顽固的对手时要记住这一点！

良好的情绪还会让人更容易接受第一印象。在福加斯的一项研究中，[实验人员要求参与者阅读一篇简短的哲学论文](#)，这篇论文随附了作者的照片。一些读者看到了一位典型的哲学教授——男性、中年、戴着眼镜；另一些读者看到的却是一位年轻的女士。和你猜测的一样，这项研究考察的是被试是否容易受刻板印象的影响：相比于年轻女士所写的论文，人们更偏爱由中年男性所写的论文吗？结果表明，的确如此。但更重要的是，在情绪良好的情境中，这种差异更大，拥有良好情绪的人更容易受到偏差的影响。

还有一些研究考察了情绪与上当受骗之间的关系。心理学家[戈登·彭尼库克（Gordon Pennycook）及其同事开展了一系列研究，来考察人们对毫无意义、看似深奥实则虚假的陈述的反应](#)。这些虚假的陈述是这样产生的：从一些广为流传的名言中随机挑选出名词和动词，然后组合成语法正确的句子，如“完整性平息了无穷的现象”或“隐藏的意义使无与伦比的抽象美变化无穷”。同意这些陈述的倾向是一种被称为“废话接受力”（bullshit receptivity）的特征。[自哈里·法兰克福（Harry Frankfurt）之后，废话已经成为一个术语](#)。法兰克福是普林斯顿大学的哲学家，他写了一本见解独到的书《论扯淡》（*On Bullshit*）。在这本书中，他将废话与其他类型的虚假陈述进行了区分。

确实，有一些人比其他人更倾向于接受废话。[他们可能会被看似令人印象深刻的断言所打动](#)，这些断言虽然看起来真实而有意义，但实际上空洞无物。这里的容易上当受骗并非是一种永久性的、无法改变的倾

向。[诱发良好的情绪会让人们更容易接受废话，以及更容易上当受骗。](#)他们不太容易觉察到具有欺骗性或误导性的信息。相反，[处于不良情绪中的目击者在看到这些具有误导性的信息时，更有可能忽视它们，从而避免做出虚假指证。](#)

即使是道德判断也会受到情绪的强烈影响。在一项研究中，研究人员让被试回答“[天桥难题](#)”（footbridge problem），这是道德哲学中的一个经典问题。在一个假想的情境中，5个人即将被一辆失控的电车撞死。研究人员要求被试想象自己就站在天桥上，电车很快就会从桥下经过。被试需要决定是否要把旁边一个大个子男人从天桥上推下去，使其落到铁轨上，这样他的身体就能使电车停下来。如果被试这样做了，那么大个子男人会死，但那5个人会因此而得救。

天桥难题体现了道德推理思维之间的冲突。英国哲学家杰里米·边沁（Jeremy Bentham）提出的功利主义认为，失去1条生命比失去5条生命要好。而康德（Immanuel Kant）的道德义务论则认为，不能以拯救更多人的名义杀害某个人。天桥难题明显包含个人的情绪因素：把一个人从桥上推下去落到电车轨道上是一种特别令人厌恶的行为。要从功利主义的角度把这个人从桥上推下去，需要人们克服对伤害陌生人身体这种行为的厌恶。只有少数人表示他们会这么做，在这项研究中，这部分人占比不到1/10。

然而，我们可以通过让被试观看5分钟视频短片的方式来诱发他们良好的情绪，这时他们打算将人推下天桥的可能性增加了3倍。无论我们是把“不可杀人”作为一个绝对的原则，还是愿意杀死一个陌生人来救活5个人，这都应该体现我们最深层的价值观。然而，我们的选择似乎只是取决于刚刚观看的视频短片。

我们详细描述了这些关于情绪的研究，因为我们需要强调一个很重要的事实：你并非在所有时刻都一样。随着情绪的变化（有时候你会意识到），你的认知机制也会改变（你可能根本意识不到）。如果你

面临一个复杂的判断问题，当前的情绪会影响你对这个问题的思考以及得出的结论，即便你认为你的判断没有受到情绪的影响，并且能很自信地阐明自己给出最终答案的理由。简而言之，你的判断充满噪声。

其他许多偶然因素也会诱发判断中的情境噪声。在不应该影响但实际上影响了专业判断的诸多外在因素中，压力和疲劳是两个主要因素。一项针对近70万次基层保健门诊就诊情况的研究表明，[医生在漫长的一天结束时开阿片类药物的可能性显著增加](#)，然而，没有理由认为预约下午4点就诊的患者比早上9点就诊的患者更痛苦。医生落后的看诊进度也不应该影响处方的决策。事实上，其他治疗疼痛的处方，如非甾体类消炎止痛药以及转介物理治疗，并没有显示出类似的模式。当医生面临着时间压力时，他们显然更倾向于选择快速的解决方案，虽然这样可能会产生一些严重的问题。其他研究也显示，[在一天将要结束时，医生开抗生素处方的可能性更大](#)，而[开流感疫苗处方的可能性较小](#)。

甚至天气也会对专业判断产生显著的影响。天气的影响可能会通过情绪产生作用，即天气虽不会直接影响决策，但会影响决策者的情绪，进而影响决策。[不好的天气与记忆力的增强有一定的相关性](#)；当天气炎热时，司法审判会更严厉；阳光明媚的天气会影响股市的走向。而在其他一些案例中，天气的影响则不那么明显。心理学家尤里·西蒙松（Uri Simonsohn）发现，高校招生人员在阴天时更关注候选人的学术表现，而在晴天时对非学术表现更敏感。他报告这些发现时采用的文章标题足以令人难忘——“[云让书呆子看起来不错](#)”。

判断会随机变异的另一个案例是不同案件所处的考察顺序。当一个人在思考一起案件的时候，此前一起案件的决策就成了一个潜在的参照点。包括法官、贷款专员以及棒球裁判员在内的做出一系列决策的专业人士，倾向于恢复某种形式的平衡：在连续朝着同一方向做出一系列决策之后，他们更有可能朝着相反的方向做出决策。严格意义上说，后面

这些决策是不合理的，因为这样难免会出现误差和不公正。例如，[在美国，当前面两个庇护申请获得庇护法官的批准时，下一个庇护申请获得批准的可能性会降低19%](#)。在申请贷款时，如果前面两个申请都被拒绝了，那么接下来的申请很有可能得到批准；但如果前面两个人的申请获得了批准，那么相同的申请此时则更有可能遭到拒绝。这种行为反映了一种被称为“[赌徒谬误](#)”（gamblers fallacy）的认知偏差，它指的是人们会认为随机序列中一个事件发生的概率与此前发生的事件有关，从而低估了“坏运气”再一次发生的可能性。

衡量情境噪声

相对于整体的系统噪声，情境噪声有多大呢？虽然没有一个原则适用于所有情境，但存在一个一般性的原则：就大小而言，本章所涉及的这些情境噪声要小于个体在水平噪声和模式噪声方面的稳定差异。

例如，正如上文所提及的，如果同一位法官连续两次批准了庇护申请，那么接下来的申请获得批准的可能性会下降19%。这种变异性肯定是有问题的，但如果与不同法官之间的变异性相比就是小巫见大巫了：天普大学法学院教授杰亚·拉姆基-诺加莱斯（Jaya Ramji-Nogales）及其合作者发现，在迈阿密的一家法院，一位法官批准了88%的庇护申请，而另一位法官只批准了5%的庇护申请。这是真实的数据，不是进行噪声审查的实验数据，因此申请者是不同的，但这些申请者的申请由哪位法官受理可以说是随机的。拉姆基-诺加莱斯及其合作者发现，申请人来源国的差异并不能解释这种差异。如果差异如此巨大，那么申请获得批准的概率下降19%好像也没有什么大不了。

同样，指纹鉴定人员和医生有时候也会做出自相矛盾的判断，但这种情况远远少于与其他人的判断不一致的情况。在我们回顾的每一个案例中，情境噪声在系统噪声中所占的比例都是可以测量的，情境噪声所占的比例皆小于个体间的差异。因此可以这样说：你不可能永远是同一

个人，随着时间的流逝，你远没有你所想象的那么前后一致；但令人欣慰的是，与其他人相比，今天的你更像昨天的自己。

造成情境噪声的内部原因

情绪、疲劳、天气、顺序效应等许多因素都可能导致同一个人在对同一案件做出判断时，产生我们不愿看到的变异。我们可能希望构建一个环境，在这个环境中，所有与决策有关的外部因素均是已知、可控的，至少从理论上来说，这样的情境可以减少情境噪声。但即使是这样的情境，也不足以彻底地消除情境噪声。

宾夕法尼亚大学心理学家[迈克尔·卡哈纳（Michael Kahana）及其同事研究了记忆的表现](#)。根据我们的定义，记忆不是一项判断任务，而是一项认知任务，可以对其条件进行严格控制，也很容易测量其表现。在其中一项研究中，他们对79名被试的记忆表现进行了非常深入的分析。被试要在不同的日子里参加23场测试会，在此期间他们需要对每组包含24个单词的24组单词进行回忆。他们所能回忆起来的单词数量可以作为记忆表现的指标。

卡哈纳及其同事感兴趣的不是被试之间的差异，而是哪些因素能够预测不同被试之间的差异。不同被试的警觉程度与前一天晚上的睡眠情况是否会影响记忆表现？他们的记忆表现是否会因在一次次的测试会中反复练习而得到提升？在每一次测试会中，记忆表现是否会由于感到疲劳或无聊而降低？是否有一些列表中的单词比其他列表中的单词更容易被记住？

上述所有问题的回答都是肯定的，但对结果的影响并不大。将所有这些因素纳入一个模型中，也只能解释某个特定被试记忆表现差异的11%。正如研究人员所指出的那样：“剔除这些可预测的因素的影响后，还剩下如此大的变异，这着实令我们吃惊不已。”即使在这种严格

控制的情境中，到底是哪些因素引发了情境噪声，仍是未知的。

在研究者分析的所有因素中，强烈影响被试对某组单词的记忆表现的一个预测因素，并非外部因素。被试记忆某组单词的表现在很大程度上受他的前一组单词记忆表现的影响。如果他在记忆前一组单词时表现良好，那么他接下来的表现也会很好；如果他在记忆前一组单词时表现一般，那么他接下来的表现也会一般。被试在记忆每一组单词时表现出的差异并不是随机的：在每一次测试会中，它会随着时间的推移而有所起伏，并且没有受明显的外部因素的影响。

用卡哈纳及其同事的话说，这些结果表明，记忆表现在很大程度上受“控制记忆功能的内源性神经过程效率”的影响。换句话说，大脑效率的即时变化并不源于天气或分心等外部因素的影响，而是一种大脑自身运作的特征。

大脑功能的内部变异很有可能以一种我们完全无法控制的方式来影响我们的判断，大脑功能的这种变异性应该可以打消人们认为情境噪声可以被彻底消除的想法。用篮球运动员的罚球进行类比并不像最初看起来那样简单：正如运动员的肌肉从来不会执行完全相同的动作一样，我们的神经元也从来不会以完全相同的方式运作。如果我们的大脑是一种测量工具，那么它永远是不完美的。

然而，我们仍然可以努力去控制这些不应出现但可控的影响，当判断是由群体做出时尤其应该如此。关于这一点，我们将在第8章中进行讨论。

- 消除噪声

情境噪声不是系统噪声的最主要成因

- 判断就像罚球：我们无论多么希望能够精确地重复，都永远无法做到完全相同。
- 你的判断依赖于你的情绪、你刚刚讨论过的案件甚至天气。在不同的时间里，你不可能一直是同一个自己。
- 虽然你做不到与上个星期的自己完全一致，但与今天的他人相比，今天的你更像上个星期的自己。也就是说，情境噪声不是系统噪声的最主要成因。

第8章

群体是如何放大噪声的

个体判断中存在噪声的结果已经很糟糕了，但群体决策中的噪声危害更甚。群体决策可能会由于一些无关因素而朝任何一个方向改变。谁先发言、谁后发言，谁说话更自信，谁穿着黑色衣服，谁和谁挨着坐，谁在某个时刻笑了 / 皱眉了 / 呈现出其他身体姿势.....所有这些因素都会影响结果。相似的群体每一天都会做出各种不同的决策，如雇用、晋升、破产、沟通策略、环境保护条例、国家安全、大学录取或新产品发布等方面的决策。

我们在前文曾提到，对多个个体的判断进行汇总可以减少噪声，而这里又强调群体会放大噪声，似乎显得很奇怪。然而，受群体动态过程的影响，群体也会放大噪声。有做出的判断与正确答案接近的明智的群体，但也有追随暴君的群体、增加市场泡沫的群体、相信奇迹或受共同幻想支配的群体。微小的差别可能导致一个群体坚定地说“是”，而本质上相同的另一个群体却坚定地说“否”。群体成员之间的互动会导致数量巨大的噪声，这也正是我们此处强调的重点。无论是对于相似群体之间的噪声，还是对于单个群体中的噪声，上述假设都是成立的。因此，无论是单个群体还是多个群体，都是如此。我们应该把这些群体对某个重要问题的判断视作一系列可能性中的一个。

音乐中的噪声

为了寻找证据，我们从一个看似不太可能的地方开始：普林斯顿大学社会学教授[马修·萨尔加尼克（Matthew Salganik）和他的合作者开展了一项关于音乐下载的大型研究](#)。实验人员创建了一个由几千人组成的

控制组（某流行音乐网站的访客）。控制组成员可以试听并下载72首新歌中的1首或多首。这些歌曲的名字都很生动：《深陷橘子皮》《啃》《眼罩》《棒球术士v1》《粉红侵略》等。还有一些歌曲名字跟我们的问题看起来好像十分相关：《最好的错误》《我是个错误》《信念高于答案》《生活的神秘》《祝我好运》《走出困境》等。

在控制组中，被试未被告知其他人说了什么以及做了什么等额外的信息，这样一来，他们就可以独立判断自己喜欢哪一首歌或希望下载哪一首歌。但萨尔加尼克及其同事还创建了其他8个组，对应8种群体情境，并将成千上万的被试随机分配到这些情境中。这8组被试获知的所有其他信息都是相同的，但有一处不同：人们可以看到同组中的其他人先前下载过哪些歌曲。例如，如果《最好的错误》是该组中深受喜爱的歌曲，那么所有成员都可以看到；同样，如果一首歌完全没有人下载，他们也可以看到。

因为不同的群体在任何重要的维度上均无差别，这项研究看起来就像是重复了8次。你可能会推测，好听歌曲的排名会上升，而不好听歌曲的排名则会下降，如果是这样，这些不同群体中的歌曲排名应该相同，或至少相似，即不同群体之间没有噪声。事实上，这也是萨尔加尼克及其同事试图去探讨的问题，他们考察的是一种特定的噪声源：社会影响。

该研究的核心发现是：不同群体中的歌曲排名差异巨大，也就是说，在不同群体之间存在大量噪声。在某个群体中，《最好的错误》可能非常成功，而《我是个错误》则非常失败；在另一个群体中，《我是个错误》极其成功，但《最好的错误》的表现则一塌糊涂。如果一首歌一开始就备受欢迎，它随后也一定会表现得更好，而如果它一开始没有获得这种优势，那么结果就难说了。

可以肯定的是，最差的歌曲（在控制组中表现最差）排名不可能靠前，最好的歌也不太可能垫底，而对于其他歌曲而言，任何事情都有可

能发生。正如作者所强调的那样：“相比于独立判断，人们在有社会影响的条件下，更难预测哪些歌曲会成功。”简而言之，社会影响在不同群体之间产生了明显的噪声。如果你仔细思考，你就会知道，单个群体内部也会存在噪声，因为他们很容易就喜欢一首歌或不喜欢一首歌，这取决于这首歌一开始是否受欢迎。

正如萨尔加尼克及其同事随后所展示的，群体的结果很容易被操纵，因为[流行程度会自我强化](#)。在后续实验中，他们动了点小心思，对控制组中的歌曲排名进行了反转。换句话说，他们谎报了这些歌曲的受欢迎程度，人们看到的最好的音乐其实是最差的音乐，反之亦然。研究人员随后观察了访客们的反应，结果是，最不受欢迎的歌曲深受喜爱，而原来最受欢迎的歌曲则表现非常差。即使研究人员误导了人们哪些歌曲是受欢迎的，但在人数非常大的群体中，受欢迎和不受欢迎程度受排名的影响是相同的。唯一的例外是，随着时间的推移，控制组中最好听的歌曲会逐渐变得更流行，这意味着反向排名也没有让它垫底。但是，对于绝大多数歌曲而言，反向排名决定了它们的最终排名。

我们很容易看出这项研究与一般性的群体判断的关系。假设有一个包含10名成员的小群体，他们要决定是否采用某项大胆的新举措。如果一两个支持者先发言，他们很容易使整个团队转向他们偏好的方向。如果最先发言的是持怀疑态度的人，情况也是如此，至少当人们能够互相影响时是如此。事实上，群体中的成员常常会互相影响，因此，仅仅是因为先发言的人不同，或者一开始下载某首歌的人更多，类似的群体会做出非常不同的判断。《最好的错误》和《我是个错误》的流行现象在各种专业判断中也存在。如果群体没有收到类似歌曲排名的信息，比如对某一大胆举措的热烈支持，该举措可能仅由于其支持者未发言而无法推进下去。

不仅仅是音乐下载，其他领域也一样

如果你是一个多疑的人，你可能会认为音乐下载只是一个特例，或者至少与其他的群体判断不同，然而，[在其他领域也出现了类似的结果](#)。我们来看一下在英国的公民投票（简称公投）中，人们对不同提案的支持情况。在公投中决定是否要投支持票时，人们自然要判断这不是不是一个好的主意。这种模式类似于萨尔加尼克及其同事的研究：最初涌现的流行度会自我强化，如果某项提案在第一天没有受到关注，那么它很快就会沉寂。在政治领域就像在音乐实验中一样，支持与反对在很大程度上依赖于社会影响，具体而言，依赖于人们是否能看到其他人投的是支持票还是反对票。

美国康奈尔大学[社会学家迈克尔·梅西（Michael Macy）](#)及其合作者在音乐下载实验的基础上构建了另外一个实验，目的是弄清楚：他人的观点是否会影响人们的判断，使得相应的政治观点受到民主党人的欢迎，而遭到共和党人的反对（或者相反）。答案简单明了：是的。在网络群体中，如果民主党人看到某一观点一开始就受到其他民主党人的支持，那么他们就会采纳这一观点，并最终导致大部分民主党人支持这一观点。但是，如果另一个网络群体中的民主党人看到，某一观点一开始就受到共和党人的支持，那么这些民主党人就会拒绝接受这一观点，并最终导致大部分民主党人都拒绝接受该观点。简而言之，政治观点同歌曲一样，最终的命运取决于最初的受欢迎程度。正如梅西等研究人员指出的那样：“少数先行者的随机差异”会对整个群体产生颠覆性的影响——让共和党人和民主党人都欣然接受一系列与彼此的立场毫不相关的观点。

我们还可以思考一个一般性的群体决策问题：[人们在网上如何对各种评论做出判断](#)。耶路撒冷希伯来大学教授列夫·穆奇尼克（Lev Muchnik）及其同事在一个网站上开展了一项实验，他们向人们呈现不同的故事，并允许人们发表评论，以及对这些评论投赞成票或反对票。研究人员可以人为地、自动化地给一些评论投出第一张赞成票。你可能会想，在成百上千名访客中，使某条评论多出一张初始赞成票根本无足

轻重，这个想法合情合理，却是错的。在看到第一张赞成票之后（别忘了这完全是人为操作的），下一个访客对该评论投赞成票的可能性增加了32%。

令人惊讶的是，这一效应随着时间的推移一直在持续。5个月后，开始时人为投出的那张赞成票，使得该评论的平均赞成票得票率增加了25%。最初的一张赞成票竟然产生了如此大的影响，这表明噪声确实存在。不管最初那一票是为何而投，它都使整体的受欢迎程度发生了巨大的改变。

这项研究为群体态度的转变以及群体内为何存在噪声提供了一条线索：相似的群体会做出非常不同的判断，而同一群体做出的判断也仅仅是一系列可能性中的一种。群体成员表达的赞成、中立、反对意见，其作用也类似于一开始投赞成票或反对票。如果群体中的一个成员立即表示赞同，那么其他成员也就有理由这么做。毫无疑问，当群体赞同某些产品、人、活动或思想时，可能并不是因为它们的内在优点，而是因为“提前投票”发挥了作用。当然，穆奇尼克的研究针对的是大规模群体，但同样的结果也会出现在小规模群体中，甚至更加富有戏剧性，因为最开始投下的赞同某个计划、产品或判决的赞成票经常会对他人产生更大的影响。

这里有一个相关的观点。我们曾经指出群体智慧效应指的是，如果你召集一大群人，问他们一个问题，他们的答案的平均值更有可能接近真实答案。对判断进行汇总是一种减少噪声，进而减少误差的非常好的方法，但是如果人们互相交流，那情况又会如何呢？你可能认为这样做是有好处的。毕竟人们可以互相学习，从而找出正确答案。在一些非常有利的条件下，互相分享知识、深思熟虑的群体确实会做得很好。然而，独立做出判断是发挥群体智慧的前提条件，如果人们不是自己做出判断，而是依赖于其他人，那么群体并不会更明智。

[有些研究已经表明了这一点](#)。在简单的评估任务——评估城市里的

犯罪数量、一段时期内增长的人口、不同国家国界线的长度等任务中，只要群体成员独自做出判断，群体会更明智；如果他们知道了其他人的评估，比如一个12人小组的平均估计值，那么群体比个体的表现还要糟糕。正如研究者指出的那样，社会影响是有问题的，因为它们降低了群体多样性，但并没有减少群体的误差。具有讽刺意味的是，即便一点点社会影响都会降低群体智慧，但对多个独立判断进行适当的汇总则可以产生令人难以置信的准确结果。

信息级联，极易放大群体判断的噪声

我们描述的一些研究中包括“信息级联”（information cascades）。信息级联很常见，它可以解释为什么一些相似的商业群体、政治群体以及其他群体会做出完全不同的决策，以及为什么一些小的变化会产生如此不同的结果乃至噪声。只有历史真实发生了，我们才能看到，而对于许多群体以及群体决策而言，存在着各种各样的可能性，而最终得以实现的只是其中的一种。

要想理解信息级联是如何发挥作用的，我们可以想象在一间大办公室中有10个人，他们在决定要雇用谁来担任一个重要职位。候选人有3位，分别是托马斯、山姆和朱莉。假设群体成员是按顺序发表自己的观点的，每一个人都会认真聆听其他人的判断。阿瑟第一个发言，他认为托马斯是最佳人选。芭芭拉现在知道了阿瑟的判断，如果她也认为托马斯是最佳人选，她肯定会认同阿瑟的意见。假如她不确定谁是最佳人选而她信任阿瑟，她可能也会认同托马斯是最佳人选。因为她足够信任阿瑟，所以她支持了阿瑟的判断。

现在轮到查尔斯发言。阿瑟和芭芭拉已经表明了他们想雇用托马斯，但查尔斯有自己的想法。基于他自己掌握的有限信息（他非常清楚自己的信息很有限），他认为最佳人选不是托马斯，而是朱莉。虽然查尔斯有自己的想法，但他也有可能忽视自己已知的信息，而只是附和

阿瑟和芭芭拉。如果此事发生，那并不是因为查尔斯懦弱，而是因为他是一个尊重他人的倾听者。他可能只是认为：阿瑟和芭芭拉都选托马斯，他们肯定有自己的理由。

第四位发言人是戴维，除非戴维认为他自己掌握的信息确实比前几人更有说服力，否则他也会附和前几个人的意见。如果戴维也这样了，那么戴维就处在一个“信息级联”中。事实上，如果戴维有非常充分的理由认为阿瑟、芭芭拉和查尔斯的选择是错误的，那么戴维可能会表示反对。但如果他缺乏充分的反对理由，那么他就会做出和前几个人同样的选择。

重要的是，查尔斯和戴维可能了解托马斯或其他候选人的一些信息，并且有自己独到的看法，而阿瑟和芭芭拉并不知道这些信息和独到的看法。如果这些信息得以分享，那么这些非公开的信息可能会改变阿瑟和芭芭拉的意见。如果查尔斯和戴维先发言，他们不仅能表达关于候选人的意见，而且可能提供对其他决策者产生影响的信息。但由于查尔斯和戴维是后发言的，所以他们的非公开信息就只有自己知道。

假设现在大家也想听听后续参与投票的人——埃丽卡、弗兰克和乔治的观点。如果阿瑟、芭芭拉、查尔斯和戴维都认为托马斯是最佳人选，即使埃丽卡等人有理由认为其他人选可能更合适，阿瑟等人也还是会做出相同的选择。当然，如果答案明显是错的，埃丽卡等人会反对这种越来越趋于一致的意见，但如果错误没有那么明显呢？这个例子的吊诡之处在于，阿瑟最初的判断启动了一个过程，其他人被引导进了信息级联中，即便有些支持托马斯的人实际上根本没有任何看法，甚至有人认为托马斯根本不是最佳人选，但最终结果依然是所有人都选择了托马斯。

当然，这个例子是人为设定的，然而在各种群体中，类似的事情经常发生。人们倾向于向他人学习，如果先发言的人似乎喜欢某个事物或者想去做某件事，人们会表示认同。如果人们不怀疑这些先发言的人，

或缺少一个明确的理由认为后者是错误的，那么至少在这些情况下，人们会选择认同。

我们想重点强调的是：信息级联会导致噪声可能出现在多个群体之间，有时出现的可能性甚至非常大。在上面的例子中，是阿瑟先发言，并且他看好托马斯。假设是芭芭拉先发言，而她更看好山姆，或是假设阿瑟的感觉稍微有点不同——他更喜欢朱莉，那么，一个可能的结果是，群体最终会倾向于选择山姆或朱莉，但并不是因为他们更好，而是因为信息级联。这也是音乐下载实验及同类实验中的核心发现。

需要注意的是，人们进入信息级联并不意味着他们是非理性的。如果人们不确定要雇用谁，追随他人未必不是明智之举。随着持同一观点的人越来越多，认同他们的选择仍然是明智的。然而，这里有两个问题：首先，人们往往会忽视一种可能性，即大部分人也跟他们一样处于信息级联中，因此他们也没有做出自己独立的判断。当看到3个、10个、20个人都欣然接受某种结论，我们可能会低估他们受前面的人影响的程度。即使他们的一致性反映的只是最初几个人的观点，但我们可能会认为这种一致性反映了某种群体智慧。其次，信息级联可能会导致整体朝着完全错误的方向前进，毕竟，阿瑟对托马斯的判断可能是错的。

当然，信息不是导致群体成员互相影响的唯一原因，社会压力也是很重要的因素。在公司或政府机构中，人们可能会通过保持沉默来避免自己显得不友好、爱争吵、迟钝或愚蠢。人们希望成为团队中的一员，这就是为什么人们通常会追随他人的观点和行为。人们可能认为自己知道什么是对的或什么有可能是对的，但他们表面上仍然倾向于与群体或少数优先发言者保持一致，从而在团队中保持良好的风度。

刚刚讲到的招聘故事也如出一辙，人们选择托马斯并不是因为他们通过彼此分享的信息了解到了托马斯的优点和长处，而是因为他们不希望被看成愚蠢或不合群的人。阿瑟支持托马斯的这一最初判断可能会引发一种从众效应，最终对埃丽卡、弗兰克或乔治施加了强大的社会压力

——仅仅因为其他人都喜欢托马斯，所以埃丽卡等人也选择了托马斯。就像信息级联一样，社会压力也会形成“信息级联”：人们可能放大了先发言者所持有的信念。如果人们支持托马斯，那么他们这样做可能不是因为他们真的喜欢托马斯，而是因为一个优先发言的人或一个有权势的人支持托马斯。这样一来，群体成员的一致性进一步增加，社会压力水平也增加了。这是一个在公司或政府机构中非常常见的现象，它可能会增加人们对错误判断的信心，并导致人们一致支持这个错误的判断。

社会压力会导致不同群体之间产生噪声。如果在公司中，某个人组织召开了一次会议，希望对公司发展方向做出重大改变。会议的发起者可能最先发表一番言论，进而导致人们一致支持这种改变。他们的一致性可能是社会压力的产物，而并不是自己的观点。同样，如果另一个人在会议一开始就表明了不同的观点，或者最初的发言者保持沉默，讨论可能会朝着一个不同的方向发展。总之，非常相似的群体可能会由于社会压力的影响而到达不同的终点。

群体极化，讨论往往会滋生更极端的结论

在美国和其他一些国家，刑事案件和一些民事案件通常是由陪审团参与审判的。人们希望陪审团成员经过互相商议，做出比个体更明智的决策。然而，针对陪审团的研究揭示了一种会产生噪声的社会影响：群体极化（group polarization）。这一概念指的是，人们在互相交流时，往往会提出比原有倾向更极端的观点。例如，在一个7人群体中，如果大部人都认为在巴黎设立一个新的办事处是一个好主意，讨论之后，群体的决定可能会变成：在巴黎设立一个新的办事处是一个极好的主意。内部讨论常常会导致群体更自信、更团结、更极端，三者通常以更大的热情展现出来。群体极化不仅发生在陪审团中，也发生在要做出专业判断的团队之中。

我们通过一系列实验研究了陪审团在“产品责任案”中做出的惩罚性

损害赔偿的决策。每位陪审员的决策对应一笔赔偿金额，目的是惩罚公司的不合规行为，并对其他公司起到威慑作用。我们会在第15章更加详细地讨论这项研究。为了说明群体极化问题，[我们来看一个实验，该实验比较了现实世界中的陪审团和“统计中的陪审团”](#)。首先，我们向899名被试呈现案件情境，并要求他们独立做出判断：用具有7个等级的量表来表达他们的愤怒程度、惩罚倾向，以及给出相应的赔偿金额。随后，基于这些被试的反应，我们利用计算机模拟出数百万个“统计陪审团”，即随机匹配的虚拟的6人群体。在每一个统计陪审团中，我们采用6人的中位数作为最终的裁决结果。

我们发现，这些统计陪审团的裁决非常一致，也就是说，噪声大大减少了。噪声水平的降低是因为对裁决结果进行了机械性的汇总，即对个体的独立判断进行平均会减少噪声。

然而，现实世界的陪审团不是“统计陪审团”，陪审员们会针对一起案件交流各自的观点。你有理由怀疑这些经过深思熟虑的陪审员是否真的会倾向于做出与评级为中位数的成员一致的判决。为了探明这一点，我们紧接着做了第二项研究。这项研究召集了3000多名有陪审员资格的人，[由他们组成500多个6人一组的陪审团](#)。

答案简单明了：总是互相商议的陪审团比统计陪审团具有更多的噪声。这清楚地反映了由于社会影响带来的噪声，互相商议增加了噪声。

这项研究还有一个有趣的发现。如果6人中评级为中位数的成员只有中等程度的愤怒，并且倾向于对相关人员从轻处罚，那么陪审团商议后的判决通常会更宽容；相反，如果各项选择均为中位数的成员非常愤怒，并且倾向于进行严厉惩罚，那么经过交流之后，陪审团会更愤怒，他们做出的判决也更严厉。当用赔偿金额来表达这种愤怒时，陪审团商议后的赔偿金额要比金额的中位数高。实际上，27%的陪审团选择的赔偿金额通常与陪审员选择的最高赔偿金额相等，甚至会比后者更高。能够相互交流的陪审团的噪声不仅比“统计陪审团”更高，而且加重了其成

员原有的倾向。

回想一下关于群体极化的基本发现：人们彼此交流之后，明显变得比原来更加极端了，我们的实验证明了这一现象。陪审团成员在商议后要么变得更加宽容（当评级为中位数的成员倾向于宽容时），要么变得更加严厉（当评级为中位数的成员倾向于严厉时）。同样，倾向于实施金钱惩罚的陪审团在商议之后也会变得比评级为中位数的成员更加严厉。

对群体极化的解释类似于对信息级联的解释：信息发挥着重要作用。如果大部分人倾向于进行严厉惩罚，那么群体成员会听到很多认为有必要进行严厉惩罚的观点——反方的观点更少了。如果群体成员能够互相交流，那么他们会朝向处于主导地位的观点转变，导致群体更加团结一致、更加自信、更加极端。而且，如果人们在意自己在群体中的声誉，他们也会朝着占主导地位的观点转变，这样也会导致群体极化。

当然，群体极化会产生误差，并且经常如此，但我们的主要关注点在于变异性。正如我们所看到的那样，对判断进行汇总会减少噪声，也正因为如此，判断数量越多，判断的品质越好，这也是为什么“统计陪审团”比单个陪审团噪声更少。同时，我们发现能够相互交流的陪审团会比“统计陪审团”产生更多噪声。当处于相似情境中的群体最后表现出巨大的差异时，其原因往往在于群体极化，其结果是产生巨大的噪声。

在商业、政府以及其他机构中，信息级联和群体极化都会导致群体在应对同一问题时产生巨大差异，最终的判断结果取决于少数人——那些率先发言的人或有巨大影响力的人，这是一个值得注意的问题，因为个人的决策有很大的噪声。我们已经看到，水平噪声和模式噪声会使得群体成员的观点产生不应有的差异，而且该差异比我们预期的更大。我们已经看到疲劳、情绪、可以比较等情境噪声会影响率先发言的那个人的判断，群体互动则会放大这种噪声。结果，经过商议的群体会比仅仅对个体判断进行平均的统计群体产生更大的噪声。

由于企业或政府部门的一些重大决策都是在商议之后做出的，我们尤其要对这种风险保持警觉。组织及其领导应该采取一些方法来控制其成员在判断中的噪声，比如对群体商议进行管理，从而减少噪声而不是增加噪声，我们提出的减少噪声的策略，其目的就在于此。

- 消除噪声

群体决策中的噪声

- 似乎任何事情都取决于它最初的受欢迎程度，因此，我们最好想尽一切办法来让自己发行的产品在第一个星期就获得巨大成功。
- 正如我一直怀疑的那样，政治或经济理念就像电影明星。如果人们认为其他人喜欢，那么这种理念就会大受欢迎。
- 我一直很担心，当我的团队聚在一起时，我们会更自信、更团结、对我们所选择的行动方针更加坚定。我认为，我们的内部流程可能存在一些问题。

NOISE

第三部分

预测性判断中的噪声

- 发现噪声

哪里有预测，哪里就有客观无知

很多判断都是预测性的，而预测可以被验证和评估，所以我们可以通过考察预测性判断来更好地了解噪声和偏差。在这一部分中，我们主要关注预测性判断。

在第9章，我们比较了专业人士、机器和简单规则预测的准确性，结论是：专业人士预测的准确性是三者中最差的——你可能并不会对此感到惊讶。在第10章，我们探究了上述现象背后的成因，并发现噪声是导致人类判断力降低的主要因素。

为了得出这些结论，我们需要评估预测的品质，这就需要一种测量预测准确性的方法，从而回答“预测与结果的协变关系”（covariation）这样的问题。例如，假设人事部门定期评估新员工的潜力，那么我们就可以在几年后考察这些员工的绩效表现，并对比其潜力评估与绩效评估之间的相似度。如果某位员工在入职时被评价为高潜力员工，并且他在随后的工作中也获得了较高的评价，那就说明针对他的潜力评估在一定程度上是准确的。

符合这一直觉的测量指标叫作“[一致性比率](#)”（[Percent Concordant, PC](#)）。它回答了一个更具体的问题：假设你随机挑选了两名新员工，那么在潜力评估中得分较高的新员工，在随后的工作中也表现较好的概率是多少？如果早期评估完全准确，则PC应为100%，即对两名新员工的潜力评估完美地预测到了两人的绩效排名。如果早期预测完全没用，那么一致性只会随机发生，所谓的“高潜力”员工表现好和表现差的概率相当，即PC为50%。我们将在第9章继续讨论这个已被广泛研究的有关

招聘的例子。再举个简单点的例子，[成年男性脚的尺码与身高的PC值为71%](#)，也就是说，如果对两个人“评头论足”，你会发现长得高的那个人同时脚也大的可能性是71%。

PC是一个衡量协变关系的直观指标，这是它的优点，但它并非社会科学家所使用的标准度量指标。标准度量指标是相关系数（**correlation coefficient**, r ），当两个变量正相关时，其值在0~1的范围内变化。在前面的例子中，身高和脚的尺码之间的相关系数约为0.6。

我们可以使用很多方法来审视相关系数。这里有一种很直观的方法：两个变量之间的相关系数就是指决定它们的因素中共有成分所占的百分比。例如，如果某个特征完全由遗传所决定，那么我们可以推测亲兄弟姐妹在该特征上的相关系数为0.5，因为他们拥有50%的共同基因，而堂兄弟姐妹之间的相关系数为0.25，因为他们拥有25%的共同基因。对于身高和脚的尺码之间的相关系数为0.6，我们也可以理解为，决定身高的因素中有60%也同时决定了脚的尺码。

以上两种协变关系的测量指标是直接相关的。表1列出了一系列相关系数的PC值。在本书的其余部分，当讨论人和模型的预测表现时，我们通常会同时应用这两个指标。

表1 [相关系数和一致性比率（PC）的对应关系](#)

相关系数	PC
0	50%
0.1	53%
0.2	56%
0.3	60%
0.4	63%
0.6	71%
0.8	79%
1	100%

在第11章，我们会讨论预测准确性的一个重要局限：影响未来的很多事件是无法预知的，因而大多数判断都是在我们所谓的客观无知的状态下做出的。然而令人惊讶的是，在大多数情况下人们往往会忽视这一局限，并满怀信心或过度自信地进行预测。最后，在第12章中，我们发现客观无知不仅会影响我们对事件的预测能力，甚至会影响我们对事件的理解能力，这也是“为什么噪声会隐而不见”谜题的一个重要答案。

第9章

判断与模型，简单的模型普遍优于人类判断

很多人都对预测未来的工作绩效感兴趣，不只是自己的，还有别人的。因此，绩效预测是用来考察预测性判断的实用例子。例如，一家大公司在招聘高管时，聘请了一家专业咨询公司对两名候选人莫妮卡和娜塔莉进行评估，并以取值为1~10分的量表对两人的领导力、沟通能力、人际交往能力、职业技能、应聘动机等维度进行打分（见表9-1）。你的任务是：预测她们在两年后的工作绩效，并用1~10分来评分。

表9-1 两名高管候选人的能力评估得分

	领导力	沟通能力	人际交往能力	职业技能	应聘动机	你的预测
莫妮卡	4	6	4	8	8	
娜塔莉	8	10	6	7	6	

大多数人在面对此类问题时，只是简单地盯着每一行数字并心算出平均分，然后快速做出判断。如果你也是这样，那么你可能会得出这一结论：娜塔莉是更优人选，因为莫妮卡的平均分比她差一两分。

判断还是公式

针对此问题，你采取的这种方法被称为“诊断性判断”（clinical judgment）。在此过程中，你会考虑相关信息，或许再快速计算一下，

然后利用直觉做出判断。事实上，诊断性判断就是我们在本书中简单描述的判断过程。

现在假设你以实验参与者的身份完成了上述预测工作。莫妮卡和娜塔莉的数据来自一个信息数据库，其中记录了此前聘用的数百名经理的信息，以及这些经理在5个维度上的得分。你可以使用那些评分来预测两人的工作绩效，而现在你还获得了两人的实际工作绩效数据。那么，想一想你对这两人的诊断性判断有多接近她们的实际绩效呢？

这个例子大致来源于[一项关于绩效预测的真实研究](#)。如果你曾参加过该项研究，你可能会对自己的预测结果非常不满意。一家国际咨询公司聘请了拥有博士学位的心理学家来做预测，结果发现，预测与绩效评估的相关系数仅为0.15（PC=55%）。也就是说，当他们评估一名候选人优于另一名候选人时，他们所偏爱的候选人最终获得更高绩效的可能性仅为55%，比随机选择的结果高不了多少。显然，这不是一个令人满意的结果。

也许你会认为，预测准确性之所以这么差，是因为评分信息对预测没有用。因此，我们不禁要问：对候选人的评分到底包含了多少有用的预测信息？如何将它们进行整合才能获得与实际表现相关性最高的预测分数？

有一种标准的统计方法可以回答上述问题。在上述研究中使用这种方法，可以使相关系数达到0.3（PC=60%）。结果虽然仍不尽如人意，但至少优于诊断性预测。

这种方法叫作“多元回归”（multiple regression），[它是对各种预测因素的平均值进行加权后获得预测分数的方法](#)。多元回归可以找到一组最佳权重，使整合后的预测分数与目标变量之间的相关性最大。最佳权重可以使预测的均方误差最小——这就是最小平方方法在统计学中举足轻重的有力证明。[你可能认为，与目标变量相关性越密切的预测因素，其权重也应该越大](#)；而无用的预测因素，其权重应该为0。然而事实上，

权重也可能是负数，例如候选人乘公交的逃票次数在预测其工作绩效上的权重就可能是负的。

多元回归是一个“机械性预测”（mechanical prediction）的例子。机械性预测种类繁多，从简单规则（如雇用完成高中学业的人）到复杂的人工智能模型不等。“线性回归”（linear regression）模型是最为常见的一种，因此该模型也被称为“[判断和决策研究的主力军](#)”。为方便起见，我们将线性回归模型称为“简单模型”（simple models）。

上文提到的莫妮卡和娜塔莉的案例，可以帮助我们对诊断性预测和机械性预测进行比较。[二者都具有如下一些简单的结构](#)：

- 用一组预测因素（如案例中对候选人的评分）来预测目标结果（如候选人的工作绩效）。
- 利用人类的判断做出诊断性预测。
- 基于某项规则（如多元回归），使用同一组预测因素来生成机械性预测的结果。
- 比较诊断性预测与机械性预测的整体准确性。

梅尔：最优模型击败了你

在了解诊断性预测和机械性预测之后，人们往往想知道两者之间的区别，即相比于公式，人类的判断会更优吗？

这个问题早已有人提出过，但是直到1954年，明尼苏达大学心理学教授[保罗·梅尔（Paul Meehl）](#)出版了《临床与统计预测：理论分析和证据综述》（*Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence*）一书后，该问题才引起了人们的广泛关注。梅尔回顾了20项研究，并根据一系列研究结果（如学业成就和精神

病预后效果）分析了诊断性判断和机械性判断的优劣。他得出了一个强有力的结论：简单的机械性规则普遍优于人类的判断。梅尔发现，临床医生和其他专业人员在整合信息的能力上表现极差，虽然他们自认为在这方面具有优势。

为了更好地理解上述发现为何如此令人惊讶，以及它与噪声的关系，我们需要明白简单的机械性预测模型是如何工作的。机械性预测最关键的特征是：它的预测规则适用于所有情况。每个预测因素都有特定的权重，这个权重不会因个案的不同而发生变化。你可能会认为，这种严格的约束会使模型比不上人类的判断，比如，在上文的例子里，也许你会认为莫妮卡的应聘动机和职业技能相结合是一项重要优势，能弥补她在其他方面的不足；也许你还认为，考虑到娜塔莉的其他长处，她在这两方面的劣势没什么大不了。也就是说，你会不由自主地设想她们两人不同的成功途径。凭借这些看似合理的诊断性推测，你针对两人的情况，对不同的预测因素赋予了不同的权重，而简单模型不存在这样的问题。

简单模型的另一个限制是，预测因素每增加1个单位，总是会产生相同大小的效果，即如果增加2个单位，那么所产生的效果是前者的2倍，而诊断性直觉经常与这一原则相悖。例如，娜塔莉的沟通能力是满分10分，如果你对此印象深刻，认为此分数值得你提高对其沟通能力的预测权重，那么你所做的就是简单模型所不能做到的。在加权平均公式中，得分10和9之间的差异与得分7和6之间的差异是相同的，但诊断性判断往往不遵循这一原则，相反，它反映了一种普遍性的直觉，即相同的差异在一种情况下可能无关紧要，在另一种情况下却可能非常重要。因此，我们认为没有哪个简单模型可以完整地描述你对莫妮卡和娜塔莉所做出的判断。

本文使用的例子就是梅尔模式的一个典型案例。正如我们所指出的那样，诊断性预测与工作绩效之间的相关系数仅为0.15（PC=55%），

而机械性预测获得的相关系数为0.3（PC=60%）。再回忆一下你在莫妮卡和娜塔莉的例子中对她们的优点所持有的信心。梅尔的结果强烈表明，你对自己判断品质的满意感只是一种错觉，即“效度错觉”（illusion of validity）。

人们做出预测性判断时总会出现效度错觉，因为我们无法区分预测性任务的两个不同阶段：对当前证据的评估阶段和对实际结果的预测阶段。如果要评估两名候选人哪个看起来更好，你通常会充满信心，但是这跟猜测哪名候选人实际上更好完全是两码事。比如，你可以胸有成竹地说“娜塔莉看起来是比莫妮卡更优秀的候选人”，但是，如果要断言娜塔莉将成为一位比莫妮卡更成功的经理，则要冒很大的风险，原因很明显：评估两名候选人所需要的大部分信息你都是知道的，但要想预测未来，就存在很大的不确定性。

然而，我们的思维对上述差别的感受是模糊不清的，事实上，几乎每个人都对二者的差别感到困惑。但是，如果你做出预测时表现出的自信与你进行案例评估时一样，那么你就产生了效度错觉。

即使是临床医生也无法避免效度错觉。梅尔的研究发现，最简单的公式，只要持续应用，竟然可以胜过医学专家的临床判断。你肯定能想象到临床心理学家对此做何反应，他们会感到震惊、怀疑，甚至会鄙视这种对神奇的临床直觉假装进行的肤浅研究。这种反应很容易理解，梅尔的发现与人类判断的主观经验相矛盾，大多数人都更相信自己的经验而非学者的主张。

梅尔本人对自己的发现也持模棱两可的态度，因为一提到他的名字，我们会想起“统计优于诊断性判断”这一论断，我们可能把他想象成人类洞察力的无情批判者，或者是“量化分析之父”，但事实并非如此。[梅尔不仅是一位学术研究人员，还是一位有着丰富临床经验的精神分析师](#)，他的办公室里挂着心理学家[弗洛伊德的照片](#)。同时，他也是一个多才多艺的人，不仅教授心理学课程，还教授哲学和法律学课程，他

还撰写了一些有关形而上学、宗教学、政治学甚至超心理学（parapsychology）的文章。这些特征都不符合一个刻薄的数据狂人的形象。梅尔对临床医生并没有恶意，但是正如他所说，存在“大量且一致”的证据表明：采用机械性的方法来整合意见会更具优势。

[“大量且一致”是个不偏不倚的表述。一篇发表于2000年的论文对136项研究进行了回顾](#)，清晰地表明机械性整合确实优于诊断性判断。这篇论文涵盖的研究主题广泛，包括对黄疸的诊断、军人的身体素质测评和婚姻满意度调查等，其中，63项研究表明机械性预测更准确；65项研究表明两者难分伯仲；8项研究表明诊断性预测更好。以上结果可能还低估了机械性预测的优势，因为机械性预测比诊断性预测速度更快、成本更低。此外，在许多此类研究中，[人类在判断时还具有不对等的优势，因为他们可以获取未提供给计算机模型的“私人”信息](#)。这些发现都支持了一个显而易见的结论：简单模型的决策优于人类判断。

戈德堡：你的判断模型击败了你

梅尔的发现引出了一些重要的问题：公式到底为什么会更优？模型在哪些方面可以做得更好？事实上，一个更好的问题是：为什么人类做出的判断很差？答案是：人类在许多方面都不如统计模型，其中一个主要弱点在于人类的判断过程存在噪声。

为了支持这一结论，我们来看另一项关于简单模型的研究，该研究始于美国俄勒冈州的小城市尤金（Eugene）。保罗·霍夫曼（Paul Hoffman）是一位富有且有远见的心理学家，他对当时的学术环境颇为不满，因此，他成立了一家研究所，招募了一批非常得力的研究人员，这使尤金市成为著名的人类判断行为研究重镇。

其中有一位名叫刘易斯·戈德堡（Lewis Goldberg）的研究人员，他因在“大五人格模型”的基础上发展出了“领导力角色模型”而闻名于世。

[在20世纪60年代后期，戈德堡基于霍夫曼的早期工作，开始研究用于描述个体判断行为的统计模型。](#)

建立这样一个判断模型和建立一个“现实模型”（model of reality）一样简单，因为两者所使用的预测因素完全相同。与我们最初的例子一样，预测因素是高管在工作绩效的5个维度上的得分，使用的工具也是多元回归。不同的是，该公式并非用于预测候选人的实际绩效，而是用于预测人的判断，比如你对莫妮卡、娜塔莉和其他高管候选人的判断。

用加权平均的方式对你的判断进行建模，可能看起来有些奇怪，因为你的判断并不是这样形成的。当你评价莫妮卡和娜塔莉的工作绩效时，你并没有采用这种规则，事实上，你可能没有采用任何规则。总之，判断模型并非描述实际判断过程的模型。

然而，即使你在实际判断过程中并未基于线性公式去运算，你的判断结果仍可能像是使用了线性公式一般。比如，[台球专家们在描述某一杆如何进球时，表现得就好像他们解开了复杂的方程一样，然而实际上他们并未真的那样做](#)。同理，你做出的预测就好像使用了简单公式一样，然而实际上你所做的要复杂得多。对于一个假设模型来说，即使它对过程的描述存在明显的错误，但只要该模型可以合理准确地预测人们的行为，那么它也是很有用的。简单模型就是这样的假设模型。一项针对判断研究的报告全面回顾了237项研究，发现判断模型和诊断性判断的平均相关系数为0.8（PC=79%），[尽管不是完全相关，但这种相关性已经足以支持所谓的“假设”理论了](#)。

戈德堡的研究想要解决的问题是：简单的判断模型预测实际结果的效果究竟如何？由于该模型只是对判断者的一个粗略的模拟，因此我们可以合理地假定它的预测效果不佳。那么，用模型替代判断者时，会损失多少准确性呢？答案可能会让你大吃一惊。当我们依据模型做出预测时，预测的准确性并没有降低，相反，在大多数情况下，判断模型反而表现更优，该模型甚至优于专业人士的预测。我们或许可以这样来理

解：替代品竟然比真品更好用。

这一结论已被许多领域的研究所证实。[早期一项关于预测学生毕业成绩的研究证实了戈德堡的结论](#)。研究人员要求98名参与者基于10条线索预测90名学生的GPA。研究人员根据这些预测，为每名参与者做出的判断建立了一个线性模型，并比较了参与者本人和模型预测的准确性。结果发现，对于这98名参与者来说，模型都比他们本人的预测更准确！几十年后，[一项对近50年研究成果的综述性研究也得出了同样的结论](#)：判断模型的表现一如既往地胜过判断者本人。

我们不知道这些研究中的参与者是否收到了有关个人表现的反馈，但是，如果有人告诉你，对你的判断进行粗略建模后的模型实际上比你本人预测得更准确（这极具讽刺性），想必你会感到非常沮丧。对于大多数人来说，判断活动是复杂、丰富且有趣的，这也恰恰是因为它不符合简单规则。当我们发明并应用一些复杂规则来做判断或对某些案例有了不同于其他案例的见解时，即当我们做出了无法用简单的加权求和模型去简化的判断时，我们会自我感觉更加良好，对自己的判断能力更加信心十足。但关于判断模型的研究进一步证实了梅尔的结论——很多细节都是无用的，复杂性和丰富性并不会使预测更准确。

为什么会这样呢？要了解戈德堡的发现，我们需要了解是什么导致你的实际判断与预测这些判断的简单模型之间有了差异。

基于你的判断建立起来的统计模型，不可能将所有用于判断的信息都纳入其中，模型能做的只是抽象和简化。尤其是，你的简单模型不会将你一直遵循的任何复杂规则表征出来。比如，你可能会认为沟通能力评分为10分和9分之间的差别要比7分和6分之间的差别更大，或认为在所有维度上得分均为7分的候选人比平均分相同但优势和劣势都更加明显的候选人更优秀，然而你的模型并不会表征这些复杂规则，即使你经常使用这些规则。

如果你的复杂规则行之有效，那么简单模型会因为不能重复你的规

则而导致自身的预测力下降。例如，假设你必须从一个人的技能和动机两个方面来预测他成功完成一项困难任务的可能性，那么加权平均并非好方法，因为动机再强，也无法弥补能力的不足，反之亦然。如果你使用复杂的预测规则，那么你的预测准确性将比无法获取复杂规则的简单模型更高。但复杂规则通常只会给你带来效度错觉，这实际上会降低你的判断品质。也就是说，少数复杂规则是有效的，但大多数是无效的。

此外，你的简单模型并不会表征你在判断中的噪声，它不能重现你在特定案例中由于随机反应而产生的正误差或负误差。同理，你在做出特定判断时会受到当时的环境和心理状态的影响，而模型并不会。这些判断的噪声带来的误差很可能与任何事物都不相关，这意味着在大多数情况下，我们可以将其视为随机误差。

[从你的判断中消除噪声通常会提高你的预测准确性](#)。例如，假设你的预测与结果的相关系数是0.5（PC=67%），此时你的判断中包含了50%由噪声导致的变异，而如果你的判断没有噪声，那么它们与结果的相关系数将提升至0.71（PC=75%）。由此可见，用机器减少噪声可以提高预测判断的有效性。

简而言之，用模型代替人类的判断意味着两件事：消除了人类的复杂规则，消除了噪声。判断模型比判断更有效这一强有力的发现说明：从人类判断的复杂规则中获得的好处（如果存在的话）不足以补偿噪声所带来的损失。你可能会认为自己比一般人更擅长思考、更有洞察力，但实际上只是你的噪声更多而已。

为什么我们以为复杂的规则更有效，实际上它们却损害了判断的准确性呢？一方面，人们发明的许多复杂规则并不正确；另一方面，即使复杂规则在原则上是有效的，它们也不可避免地仅适用于少数能被观察到的情况。例如，假设你已经得出结论：对于一个独创性极高的候选人，即使他在其他方面得分一般，也值得被雇用。可问题在于，从定义上看，具有独创性的候选人总是很稀缺。既然对独创性的评估可能不可

靠，在这一指标上得了高分的人就可能是侥幸，而真正具有独创性的人才往往无法被发现。即使在绩效评估中具有较高独创性的候选人最终真的表现得特别优秀，绩效评估本身也存在很多问题。两端的测量误差会不可避免地削弱预测的有效性，一些小概率事件尤其可能被忽略，复杂模型的优势很快就会被测量误差所掩盖。

[马丁·于（Martin Yu）和内森·昆塞尔（Nathan Kuncel）报告了一项比戈德堡所做的更激进的研究](#)。该研究基于莫妮卡和娜塔莉的案例，使用了一家跨国咨询公司的数据，这家跨国咨询公司聘请专家评估了3个独立样本中共847名高管职位的候选人。专家们在7个不同的评估维度上对这些候选人进行了评分，并使用他们的诊断性判断为每位候选人生成了一个预测总分，然而结果令人大吃一惊。

马丁·于和昆塞尔将判断结果与随机线性模型进行比较，而非与他们的最佳简单模型进行比较。他们为7个预测因素生成了10 000套随机权重，并应用了这10 000个[随机公式](#)来预测工作绩效。他们吃惊地发现，用任何线性模型来对所有案例进行预测，其结果均优于人类基于相同信息所做出的判断。在其中一个样本中，10 000个随机加权线性模型中有77%优于人类专家；在另外两个样本中，随机模型100%胜过人类专家。换句话说，该研究表明，所有简单模型的表现都比人类专家好。

这项研究得出的结论比我们从戈德堡的判断模型中得出的结论更有力。事实上，这是个非常极端的例子。在这种情况下，人类的判断确实非常糟糕，这就解释了为什么即使是不尽如人意的线性模型，其表现也超越了人类判断。当然，我们并不能因此下结论说机器绝对比人强，尽管如此，机械地遵守简单规则（马丁·于和昆塞尔称其为“无意识的一致性”（mindless consistency））可以显著提高针对困难问题所做判断的品质，这一事实说明了噪声对诊断性预测的巨大影响。

本章简要地说明了噪声对诊断性判断造成的负面影响。在预测性判断中，人类专家很容易被简单的公式所击败，其中包括真实模型、判断

模型甚至随机生成的模型。这一发现支持我们使用无噪声的方法——规则和算法，这也是下一章的主题。

- 消除噪声

判断中存在噪声，但判断模型中没有

- 人们往往相信自己的判断能更好地考虑问题的复杂性和微妙的细节，但复杂性和微妙的细节基本上没什么用，因为它们并不会提升简单模型的准确性。
- 在保罗·梅尔的书出版60多年后的今天，机械性预测优于人类的判断这一观点仍然令人震惊。
- 判断中有很多噪声，因此无噪声的判断模型会做出更准确的预测。

第10章

无噪声的规则

近年来，人工智能（Artificial Intelligence）特别是机器学习技术让机器能够执行许多以前只有人类才能执行的任务。机器学习算法可以承担人脸识别、语言翻译、分析医学影像等任务，并且可以以惊人的速度和准确性来处理计算问题，例如为成千上万名驾驶员迅速规划行车路线。它们还可以执行困难的预测任务：预测美国最高法院的判决；识别哪些嫌疑人更可能在保释期逃脱；评估儿童保护部门接到的哪些电话更紧急，并需要工作人员上门访视。

尽管如今我们一听到“算法”一词，首先想到的是上面这些应用，但这个词还有更广泛的含义。在词典中，算法的定义是：在解决计算或其他问题时（尤其是借助计算机）所遵循的步骤或规则。根据这一定义，我们在上一章中所描述的简单模型和其他形式的机械性判断也属于算法。

事实上，从简单的规则到最复杂且难以理解的机器学习算法，许多机械性方法都可以胜过人类的判断。机械性方法之所以有这种出色表现，一个关键原因可能是所有机械性方法均无噪声，尽管这不是唯一的原因。

为了研究不同类型的基于规则的方法，并了解每种方法为何以及在何种条件下更有价值，我们从第9章的基于多元回归的简单模型（即线性回归模型）开始我们的旅程。由此出发，我们将在复杂性频谱上朝着两个相反的方向前进，首先从极端简捷的一端开始，然后朝着逐渐复杂的方向前进（见图10-1）。



图10-1 4类规则和算法的相对复杂性

简捷：稳定之美

罗宾·道斯（Robyn Dawes）是20世纪六七十年代美国俄勒冈州尤金市研究人类判断行为的团队中的另一位明星成员。1974年，道斯在简化预测任务方面取得了突破。他的研究思路令人惊讶：他建议不要使用多元回归模型来确定每个预测因素的精确权重，而应给所有预测因素分配均等的权重。

道斯将均等权重的公式定义为“非最适线性模型（improper linear model）”。他出人意料地发现，[这些均等权重模型（equal-weight models）的准确性与合适的回归模型差不多，且远胜于诊断性判断](#)。

连“并非最合适的模型”的支持者也承认，这种说法是不可信的，并且[与统计直觉相悖](#)。的确，道斯及其助手伯纳德·科里根（Bernard Corrigan）最初曾努力将论文发表在学术期刊上，但是编辑们根本不认同。如果回顾一下上一章中的莫妮卡和娜塔莉的例子，你就会相信某些预测因素比其他预测因素更重要。例如，相比于职业技能，大多数人会给予领导力更高的权重。因此，简单的未加权平均值怎么可能比精细加权的平均值或专家判断更好地预测一个人的表现呢？

在道斯取得研究突破多年后的今天，人们已经很熟悉这种令其同时代人惊讶的统计现象。正如前文所解释的那样，多元回归模型计算出了最佳权重，从而使均方误差最小化，而多元回归使原始数据中的误差最小化，因此，公式会自行调整以便预测数据中的每个偶然因素。例如，如果样本中包含一些具有较高职业技能但是由于不相关原因而表现异常

出色的经理，该模型就将增加职业技能的权重。

这其中的挑战是：当将公式应用于样本之外时，也就是用它预测不同数据集的结果时，这些权重将不再是最优的。原始样本中的偶然因素不再存在，因为它们是“偶然因素”。在新样本中，具有较高职业技能的经理并不会都表现出色，而且新样本中具有原公式无法预测的新因素。要衡量模型预测的准确性，正确的做法是观察它在新样本中的表现，也就是观察它的“交叉验证相关性”（cross-validated correlation）。事实上，回归模型在原始样本上过于出色，因此交叉验证相关性的表现几乎总是比它在原始样本中的表现差。道斯和科里根在几种情况下对均等权重模型和多元回归模型（交叉验证后）进行了比较。他们采用的一个案例就是预测伊利诺伊大学90名心理学研究生第一年的GPA，使用的是与学业成就相关的10个变量，如能力测试分数、大学成绩、各种同龄人评分（peer ratings，如外向性）以及各种自评（如责任心）等。标准多元回归模型的预测相关系数为0.69，经过交叉验证后降至0.57（PC=69%）；均等权重模型与第一年GPA预测的相关系数与此大致相同，为0.6（PC=71%）。[许多其他研究也得到了相似的结果。](#)

当原始样本较小时，经过交叉验证后，准确性会降低更多，因为小样本的偶然性多，变异性较大。道斯指出，社会科学研究中通常使用小样本，以致所谓的最佳权重的优势消失殆尽。正如统计学家霍华德·怀纳（Howard Wainer）在一篇研究最适当权重估值的学术论文中所使用的副标题：[它并不重要](#)。用道斯的话说，“[我们不需要比我们的测量更精确的模型](#)”。均等权重模型之所以表现出色，是因为它不容易受样本中偶然因素的影响。

道斯的研究的直接理论成果值得广为人知：即使你缺少有关结果先前的数据，你也可以进行有效的统计预测，只需收集一些你认为与预测结果相关的预测因素即可。

假设你必须对已经在多个维度上获得评分的高管的绩效做出预测，

如第9章中高管的例子所示。你相信这些评分有很强的预测力，但是你没有每个评分预测的准确性数据。你也不可能花费几年的时间来追踪大量管理人员的绩效情况，但是，你可以基于这7个评分的均等权重模型来做预测。那么，这个均等权重模型的预测效果如何呢？[它与结果的相关系数将为0.25（PC=58%）](#)，远优于诊断性预测——相关系数为0.15（PC=55%），并且肯定与交叉验证后的回归模型非常相似，也不需要任何你没有的数据或任何复杂的计算。

用道斯的话说，均等权重模型具有“[稳定之美](#)”。研究判断的学生已对这句话达成共识。介绍这一观点的开创性文章的最后一句话给出了另一个精妙的总结：“[应用均等权重模型所需的全部技巧是决定要关注哪些变量，并知道如何将这些变量进行叠加。](#)”

更简捷：简约模型

另一种简化的方式是采用简约模型（frugal models）或简单规则。简约模型是对现实进行极端简化并无须复杂计算的模型，但在某些情况下，它们可以产生令人惊叹的预测效果。

令很多人感到惊讶的是，这些模型是基于多元回归的一个特征建立的。假设你使用了两个准确性很高的预测因素，它们与结果的相关系数分别为0.6（PC=71%）和0.55（PC=69%），且这两个预测因素彼此相关，相关系数为0.5。当将这两个预测因素进行最佳组合时，预测的准确性会有多好呢？答案令人失望，相关系数是0.67（PC=73%），这个结果比之前好，但并没有好太多。

该示例说明了一条一般性规则：将两个或多个相关预测因素组合后，预测效果相比于单个预测因素并不会好多少。因为在现实生活中，预测因素几乎总是相关的，所以这一统计事实支持使用包含少量预测因素的简约模型进行预测。与使用很多预测因素的模型相比，简单规则只

需少量计算或根本无须计算，就能在某些情况下达到令人吃惊的预测效果。

[一个研究团队于2020年发表了一项研究成果。他们将简约模型应用于一系列现实问题](#)，内容包括在案件待审期间法官是否该批准被告的保释申请。这项决策隐含着对被告行为的预测，如果错误地拒绝保释，被告将被不必要地拘押，从而对个人和社会造成巨大损失；如果错误地批准保释，则被告可能在受审前逃脱，甚至犯下其他罪行。

研究人员仅使用两个可高度预测被告在保释期逃脱可能性的已知变量来建立模型：被告的年龄（年龄越大，逃脱风险较低）和未按时出庭受审的次数（有未按时出庭受审记录的人，更可能逃脱）。该模型将这两个变量转换为一系列分数，并针对风险进行评分，在计算被告保释期逃脱的风险时无须使用计算机，甚至不需要计算器。

当用真实数据来测试时，该简约模型的表现与那些使用众多变量的统计模型一样好，而在预测逃脱风险方面，简约模型比几乎所有法官的判断都要好。同样的简约模型采用少数几个整数（-3~+3）对最多5个特征进行评分，并以此来对各种任务进行预测，如基于乳房X线片判断肿瘤的严重程度、诊断心脏病、预测信用风险等。在所有这些任务中，简约模型的表现都与复杂回归模型一样好，只不过它通常不如机器学习模型的表现好。

另一项研究也证明了简约模型的有效性。[另外一个研究小组研究了一个与上述案例相似但有所不同的司法问题](#)：预测惯犯。研究人员在评估被告再次犯罪的风险时，[使用的模型只有两个输入变量，但该模型的预测效果与使用137个变量的模型相同](#)。毫无疑问，这两个预测因素（年龄和先前被定罪的次数）与保释模型中使用的两个因素密切相关，[而大量证据表明，它们与犯罪行为也是紧密相关的](#)。

简约模型的吸引力在于其透明性和易用性，而且相比于其他复杂模型，它只需略微牺牲一点准确性就能获得这些优势。

更复杂：机器学习

在旅程的第二部分，让我们在复杂性频谱上朝相反的方向前进。如果我们可以使用更多预测因素，收集更多数据，发现前人未发现的关系模式，并对这些模式进行建模以实现更好的预测效果，那会如何呢？从本质上讲，这就是人工智能的目的。

[海量数据集](#)对复杂分析至关重要。而获得此类数据集越来越容易，是近年来人工智能快速发展的主要原因之一。例如，大型数据集可以机械地处理“断腿的例外”（broken-leg exceptions）这种情况。这个有点神秘的短语可以追溯到前文中梅尔假想的一个示例：设想有这样一个模型，它可以预测人们今晚去看电影的可能性，无论你对该模型有多大信心，如果你碰巧知道某人刚摔断了腿，你都可能会比模型更准确地预测他今晚是否会去看电影。

在使用简单模型时，“断腿原则”给决策者提供了重要启示：它告诉人们何时需推翻模型，何时则不需要这样做。如果你掌握了模型未考虑的如“断腿”这样的决定性信息，你就应该推翻模型的建议。此外，即使你缺少此类信息，有时你也不会同意模型的建议。在这种情况下，你试图推翻模型的行为，反映了你对相同预测因素做出反应的个人模式。这种个人模式很可能是无效的，你的干预可能会降低预测的准确性，因此你应该避免推翻模型。

机器学习模型之所以能够在预测方面表现出色，其中一个原因就是，它们能够发现人类所无法想象的各种“断腿”情况。在具有大量案例、海量数据的条件下，追踪观影行为的模型真的会学习，例如在固定观影日去了医院的人当晚不太可能去看电影。可以说，以这种方式改进对不常见事件的预测，可减少对人工监督的需求。

人工智能不是魔法，也不需要理解什么，它仅仅是在识别模式。虽然我们必须佩服机器学习的力量，但我们也要明白，人工智能可能要花

很长时间才能理解为什么断腿之人会错过电影之夜。

更明智的保释决策

在前面提到的研究团队将简单规则应用于保释决策问题的同时，[由塞德希尔·穆来纳森（Sendhil Mullainathan）^⑥](#)领导的另一个团队训练了复杂的人工智能模型来执行相同的任务。研究团队获得了更大的数据集——包含758 027个保释裁定的案例库。对于每种情况，研究团队可以获得和法官一样的信息：被告的罪行、犯罪记录、未按时出庭受审的次数等。除年龄外，参与训练的算法没有其他任何人口统计学信息适合使用。对于每一起案件，研究人员还知道关于被告是否被释放，以及他如果被释放，之后是否会按时出庭或被重新逮捕（被告中有74%的人获得保释，其中15%的人在那之后没有按时出庭，26%的人则被重新逮捕）的信息。[研究人员利用这一数据来训练一个机器学习算法](#)，并评估了该算法的表现。该模型是通过机器学习构建的，因此并不限于线性组合。如果它在数据中检测到更复杂的规律，它就会使用此模式来改进预测。

该模型用于预测嫌疑人在保释期逃脱的风险，因此将风险量化为数字，而非只产生是否准予保释的决定。这种方法确定了最大可接受风险的阈值，即如果风险高于该阈值，就应该拒绝保释。然而，研究人员发现，无论如何设置风险阈值，使用该模型的预测得分都高于法官的预测。穆来纳森的团队计算得出，如果将风险阈值设置为一个值，使模型预测的拒绝保释人数与法官判决的拒绝保释人数相同，则犯罪率最多可降低24%，个中原因在于，被关押的人最有可能再次犯罪。相反，如果将风险阈值设置为使该模型在不提高犯罪率的情况下，尽可能减少被拒绝保释的人数，则研究人员计算得出，被羁押的人数最多可再减少42%。换句话说，机器学习模型在预测哪些被告属于犯罪高风险人群方面，表现要比法官好得多。

利用机器学习建立的模型，也比使用相同信息的线性模型成功得

多，原因很有趣：[机器学习算法在变量组合中发现了一些会被线性模型遗漏的重要信息](#)。算法能对风险最高的被告进行归类，就证明它有能力找到很容易被其他模型忽略的模式。换句话说，数据中的某些模式尽管很少见，却非常准确地预测出了高风险人群。利用算法找到罕见但具有决定性作用的模式，让我们想起了“断腿”的概念。

研究人员还使用该算法为每位法官构建了模型，类似于我们在第9章中描述的判断模型（但不限于简单线性组合）。他们将这些模型应用于整个数据集，使团队能够模拟法官在遇到相同案件时可能做出的判决，并比较这些判决。结果表明，保释裁定中存在相当大的系统噪声，其中一些是水平噪声：根据宽容程度对法官进行分类时，20%最宽容的法官（即保释率最高的前20%的法官）准予保释的概率为83%，而20%最严厉的法官准予保释的概率为61%。法官对于哪些被告具有较高逃脱风险的判断方式也大不相同，被一位法官视为具有低逃脱风险的被告，可能被另一位更严厉的法官视为具有高逃脱风险。这些结果为模式噪声提供了清晰的证据。更详细的分析表明，案例之间的变异占总变异的67%，系统噪声占33%。[系统噪声包括一些水平噪声，即平均严厉程度之间的差异，但其中大多数（79%）是模式噪声](#)。

幸好，机器学习程序的高准确性并不以牺牲法官追求的其他目标，如种族平等为代价。从理论上讲，尽管该算法不使用种族相关数据，但它也可能会无意间加剧种族歧视。如果模型使用与种族信息高度相关的预测因素（如邮政编码），或是用于算法训练的数据源暗含偏见，则可能会出现种族歧视。例如，如果将过去的逮捕次数作为预测因素，而过去的逮捕次数受到种族歧视的影响，那么得到的算法也会存在歧视问题。

尽管从原则上讲，这种歧视无疑是一种风险，但在一些重要层面，该算法所做出的决策比法官群体中存在的种族歧视要轻微。例如，如果通过设置风险阈值使犯罪率与法官判决的犯罪率相同，则该算法可将有

色人种被判入狱的概率减少41%。在其他情况下，算法也得出了类似的结果，即提高准确性不必以加剧种族歧视为代价。正如研究小组所指出的：通过训练，该算法很容易用于减少种族歧视。

另一项不同领域的研究阐述了算法如何在提高准确性的同时减少歧视。[哥伦比亚商学院教授博·考吉尔（Bo Cowgill）考察了一家大型科技公司招聘软件工程师的情况](#)。考吉尔并未使用人工筛选简历的方式来筛选可进入面试流程的人，而是基于该公司收到并评估过的超过30万份简历，来训练机器学习算法进行筛选。该算法选出的候选人比人工筛选的候选人被录取的可能性要高14%。当候选人收到录取通知后，算法组筛选出来的候选人，比人工组筛选出的候选人接受工作机会的可能性要高18%。该算法还根据种族、性别和其他指标选择了一组更加多样化的候选人，而它更有可能选择“非传统”候选人，例如非名校毕业生、缺乏相关工作经验以及没有推荐信的候选人。在筛选软件工程师的简历时，人们通常倾向于选择符合这一群体所有典型特征的人，而该算法则为每个相关预测因素赋予了适当的权重。

需要明确的是，这些例子并不能证明算法始终是公平、无偏见和非歧视的。大家比较熟悉的一个例子是：一个用于预测求职者能否通过面试的算法，实际上是根据过去的晋升决策数据训练出来的，因此，这种算法必然会重蹈过去晋升决策中人类所有偏差的覆辙。

构建一个使种族或性别不平等持续存在的算法，不仅是可能的，而且十分容易做到。许多算法已做到了这一点。这些例子表明，人们越来越关注算法决策中的偏见，但是，在得出关于算法的一般性结论之前，我们应当记住：某些算法不仅比人类判断更准确，而且也更公平。

为什么我们不更多地利用规则

通过这一简短的机械性决策之旅，我们总结出，各种规则之所以会

超越人类判断，有两个原因。首先，如第9章所述，不仅仅是最新的和更为复杂的技术，所有机械性预测技术都能显著改善人类的判断。个性化的模式和情境噪声的结合会极大地影响人类判断的品质，因为简单的规则和无噪声是提高决策品质的关键。明智的简单规则比人类的判断要好很多。

其次，当数据足够丰富时，我们可以用复杂的人工智能技术找出有效的模式，并使其预测力远超简单模型。这些模型相对于人类判断的优势在于，它们不仅没有噪声，而且还具有利用更多信息的能力。

既然算法具有如此多的优点，得到大量证据的支持，那么为什么我们在本书中讨论的各种类型的专业判断，没有广泛地使用算法呢？尽管对算法和机器学习的讨论很热烈，但人们对它们的使用仍然很有限（一些特定领域除外）。许多专家不关心诊断性判断与机械性判断孰优孰劣，而是相信自己的判断。他们对自己的直觉充满信心，并对机器能否做得更好持怀疑的态度。他们将算法决策视为不人道的，认为使用算法是一种放弃责任的表现。

尽管算法决策已取得了令人瞩目的进步，但是在诸如医学诊断等领域，使用算法仍然不是惯常的做法，也很少有企业在招聘和晋升决策中使用算法。好莱坞电影制作公司的高管们是根据自己的经验判断而非公式来选择拍摄哪部电影的；图书出版商也在做同样的事情。而且，正如迈克尔·刘易斯（Michael Lewis）的畅销书《点球成金》（*Moneyball*）所讲述的那样，人们之所以对痴迷于统计的奥克兰田径队的故事印象深刻，恰恰是因为算法在运动团体中的运用是一种例外而非常规。即使在今天，教练、经理人以及与他们共事的其他人通常也更相信自己的直觉，并坚持认为统计分析不可能取代人类自身良好的判断力。

梅尔和他的合著者在[1996年的一篇论文](#)中，列举了精神科医生、医师、法官和其他专业人士反对机械性判断的至少17种理由，并对这些理由进行了驳斥。他们得出的结论是，需要结合社会心理因素来解释临床

医生对这类判断的排斥，这些因素包括“对技术性失业的恐惧”“了解不足”和“对计算机的普遍厌恶”。

从那时起，研究人员已经确定了导致这种排斥的其他因素。我们打算在这里对该研究进行完整的回顾，本书的目标是为改善人类判断提供建议，而不是像弗兰克尔法官那样，主张“用机器取代人类”。

但是，关于哪些因素会导致人类抵触机械性预测，其中的一些发现与我们对人类判断的讨论有关。最近的一项研究得出了一个重要观点：人们对算法并非全盘否定。例如，[当从人类的建议和算法的建议之间进行选择时，人们通常会选择后者](#)。对算法的抵制或厌恶并不代表一味地拒绝采用新的决策支持工具，人们愿意给算法机会，[而一旦发现它会犯错误，就不会再信任它](#)。

这种反应似乎是明智的：为什么要在你不信任的算法上浪费精力呢？作为人类，我们敏锐地意识到自己会犯错误，但这是我们不准备分享的特权，[我们希望机器是完美的，如果机器不完美，那就丢弃它](#)。

由于存在这种直觉性的期望，人们仍可能不信任算法，而继续相信自己的判断力，即使自己的判断明显不尽如人意。这种态度是根深蒂固的，除非算法能够达到近乎完美的预测准确性，否则这是不可能改变的。

幸好，可改进规则和算法的相关因素同样可用于改善人类的判断。我们不能奢望能够像人工智能模型一样有效地利用信息，但是至少可以努力模仿简单模型的简单性和无噪声性。在一定程度上，我们可以采用减少系统噪声的方法来提高预测判断的品质。如何改善我们的判断力是本书第五部分的主题。

- 消除噪声

没有噪声的规则和算法

- 当有大量数据时，机器学习算法比人和简单模型的预测要好。相比于人类的判断，即使是最简单的规则和算法也具有很大的优势，因为它们没有噪声，也不会尝试将复杂而无效的因素用于做预测。
- 既然缺少预测结果所需的数据，那么为什么不使用均等权重模型？它几乎和最合适的回归模型一样好，且比人类视情况而定的判断更胜一筹。
- 你不认同该模型的预测？是由于这里有“断腿”的例外情况，还是单纯因为不喜欢这个预测？
- 算法当然也会犯错，但是如果人类犯的错更多，那么我们应该相信谁？

第11章

哪里有预测，哪里就有客观无知

我们经常同一些公司高管分享第9章和第10章中的研究内容，并通过发人深省地发现让他们意识到人类的判断是有限的。这些发现已经存在了半个多世纪，应该很少有人能避开它们，但人们仍然对这些发现很抵触。

在我们的听众里，有一些高管会很自豪地告诉我们，相比于定量分析，他们更相信自己的直觉；其他人虽然没明说，但他们也有同样的看法。[有关管理决策的研究](#)表明，高管通常会凭直觉、感受或简单的判断（此处的判断与本书中“判断”一词的含义不同）来行事，尤其是级别较高的、经验丰富的高管。

简而言之，决策者喜欢听从自己的直觉，且多数人对直觉传递出的信息颇为满意。这里存在一个问题：这些极其自信的权威人士究竟从他们的直觉中“听”到了什么呢？

一篇有关[管理决策中的直觉的综述性文章，将直觉定义为对给定行动方案的一种判断](#)。这种判断是一种预感或关于正确性 / 合理性的信念，但这种判断没有明确的理由或依据，从本质上说，就是“知道是什么，却不知道为什么”。我们认为，这种“知道是什么，却不知道为什么”的直觉，实际上就是我们在第4章中提到的做出判断时产生的内部信号。

内部信号是一种自我管理的奖励，是一个人努力（有时或许没那么努力）做出判断并最终完成判断后的奖励。它是一种令人满意的情感体验，也是一种令人愉悦的一致感，它使我们感觉我们所考虑的证据和做出的判断是正确的，就如同玩拼图游戏时把所有的碎片都拼对了一样。

稍后我们将看到，隐藏或忽略那些与判断不匹配的证据还可以增强这种一致感。

内部信号很重要，但具有误导性，因为内部信号往往会被理解为一种信念而不是感觉。这种“感觉正确”的情感经验伪装成了我们对判断有效性所持的信心——我知道该判断是正确的，即使我不知道为什么。

然而，对判断所持有的信心并不能保证判断的准确性，许多充满信心的预测都是错的。尽管偏差和噪声都会造成预测误差，但此类误差最重要的来源并非受限于预测性判断实际有多好，而是受限于预测性判断应该有多好。我们将这一局限性称为“客观无知”（objective ignorance），这也是本章的重点。

客观无知

如果你经常进行预测性判断，那么你可以问自己一个问题，这个问题可能适用于很多工作，如选择股票或预测职业运动员的表现。为了方便起见，我们使用在第9章中出现过的筛选应聘者的例子。假设近年来你已经评估了100名应聘者，现在，通过将自己的评估与应聘者入职后的实际表现进行对比，你就能够对自己的招聘决策的品质进行评估。如果你随机选出两名你评估过的应聘者，你认为你的事前判断和事后结果达成一致的可能性有多大？换句话说，比较任意两名应聘者，你认为更有潜力的应聘者实际上表现也更好的概率有多大？

我们经常就此问题对高管进行非正式调查，高管们的答案大部分为75%~85%。我们怀疑，他们可能由于谦虚或不想被认为是自吹自擂才做出了这样的回答，否则他们给出的比例会更高。研究人员在与高管们私下一对一的交谈中发现，高管们实际的自信程度往往更高。

既然你已经熟悉一致性比率（PC），那么你可以轻易地看出此类评估的问题。PC为80%大致对应的相关系数为0.8，在现实中人们的预

测能力几乎无法达到这么高的水平。[近期一篇有关人才选拔的报告](#)指出，人类判断的预测相关系数远低于0.8，平均而言，人类判断的预测相关系数为0.28（PC=59%）。

鉴于人才选拔的难度很大，人类判断的表现糟糕也在情理之中。年轻人刚从事一份新工作时会面临许多机遇和挑战，这些机遇和挑战会以各种方式改变他的生活。他可能很幸运，遇到一位信任他、给他机会、在工作上提携他、帮他建立信心、能调动他的工作积极性的主管；他也可能没那么幸运，尽管自己没犯什么错，职业生涯却充满挫折。同样，他在私人生活中可能遇到的很多事也会影响他的工作表现，而这些事都是人们在当下无法预测的，就算是世界上最好的预测模型也无法预测，这种难以解决的不确定性包括所有与你的预测有关但当下无法得知的信息。

此外，与应聘者有关的很多事情从原则上说应该是可以知道的，但你在做判断时并不知道。对我们的目的而言，不管这种知识上的差距是源于缺少足够的预测性测试，还是由于获取更多信息的成本过高，抑或是由于你自己的调查疏忽，都无关紧要，不管是哪个原因，你都处于信息不完备的状态中。

难以琢磨的不确定性（未知之事）和不完备的信息（可知但不知之事）都将使完美预测变得不可能。这些未知信息并非源于判断中的偏差或噪声，而是源于任务本身的客观特征。这种由于重要信息缺失而产生的客观无知严重限制了人们判断的准确性，为避免用词太过专业，我们用“无知”来指代这种不确定性。这样可以避免混淆“不确定性”和“噪声”。不确定性是关于世界和未来的，噪声是本应相同的判断中出现的变异。

在某些情况下，我们能获得的信息比另外一些情况下更多，而且客观无知更少。此时，大多数专业判断的表现都是非常不错的。例如，医生可以对很多疾病做出准确的诊断，律师在面对法律纠纷时可以准确地

判断法官的判决结果。

但是，一般而言，你仍然可以想象到，从事预测性工作的人还是会低估自身的客观无知。[过度自信是已经被大量研究证明了的一种认知偏差](#)。具体而言，人们往往会过度高估自己做出准确预测的能力，即使是在信息有限的条件下。我们所讲的预测性判断中的噪声也可以被称为客观无知，哪里有预测，哪里就有客观无知，而且客观无知比你想象的要严重得多。

异常自信的权威：准确性和黑猩猩扔飞镖差不多

心理学家菲利普·泰特洛克是我们的好朋友，他是一个坚持真理且充满幽默感的人。2005年，他出版了《专家的政治判断》（*Expert Political Judgment*）一书。书名听起来是中性的，但实际上这本书对专家预测政治事件的能力进行了猛烈抨击。

泰特洛克研究了近300位专家的预测，包括著名的记者、受人尊敬的学者以及国家领导人的高级智囊团等。他想验证这些人的政治、经济和社会性预测是否正确，这项研究持续了20年之久。可见，想要验证长期性预测是否正确，你必须有足够的耐心。

泰特洛克的主要发现是：这些所谓的专家在对重大政治事件进行预测时表现得非常糟糕。书中有句玩笑话很有名：“整体上，普通专家预测的准确性和黑猩猩扔飞镖差不多。”更精确地说，那本书的核心内容是：那些以“[对政治和经济趋势发表评论或提供建议](#)”谋生的专家，他们“在‘展望’新趋势时，做得并不比《纽约时报》的记者或细心的读者好”。可以肯定的是，专家们讲故事的能力很强，他们可以分析形势，并用令人信服的方式来描绘事态的发展趋势，并满怀信心地在演播室里反驳那些提出反对意见的人，但是他们真的知道会发生什么事吗？事实上，他们可能并不知道。

泰特洛克撕开了专家们的面具，并得出了上述结论。对于每个预测性问题，他都要求专家给出三种结果（维持现状、很可能发生或不大可能发生）的对应概率。在理想情况下，即使让一只黑猩猩通过扔飞镖的方式进行选择，它都会以相同的概率（1/3）“选中”三个结果中的任意一个。泰特洛克发现，专家们预测的准确率并不比这一最低标准好多少。平均而言，他们评估那些未来真正发生了的事件时，给出的概率只比那些最终没有发生的事件稍微高一点，但他们常常表现得异常自信。那些对世界该如何运转拥有一套清晰理论的权威人士是最自信的，也是最不可靠的。

泰特洛克的发现表明，对具体事件进行详细的、长期性的预测根本不可能。这个世界是混乱的，一些微不足道的小事都可能引发严重的后果。例如，在受孕的瞬间，历史上的每个重要人物以及无关紧要的人物都有50%的可能性会以另一种性别出生。那样一来，注定会发生不可预见的事件，而且这些不可预见的事件的后果也是不可预见的。因此，你对未来的展望越远，客观无知就积累得越多。专家们在政治判断上的局限性并非源于预测者的认知局限，而是由他们对未来的客观无知所决定的。因此，我们的结论是：不应该将专家失败的预测归咎于专家本人。但是，他们确实应该受到批评，因为他们在尝试完成一项不可能完成的任务，却相信自己可以做到。

泰特洛克还有一项令人震惊的发现：长期预测毫无用处。几年后，他与妻子芭芭拉·梅勒斯（Barbara Mellers）合作，研究了人们在相对较短的时间内（通常不到一年）对事件进行预测的情况。他们发现：短期预测是困难的，但并非不可能，而且有些人始终比大多数人（包括情报界的专业人士）预测得好。泰特洛克和梅勒斯将这些人称为“超级预测者”（superforecasters）。在我们看来，客观无知会随着我们对未来展望的深入而增多，他们的新发现恰恰与这一观点相符。我们将在第21章中继续讨论超级预测者。

人的判断很糟糕，但模型也不尽如人意

泰特洛克的早期研究表明，对于时间跨度较大的政治预测，人们往往是无能为力的。要想证明一项任务是不可能完成的，只有在很多可靠的参与者尝试了该任务并且都失败了的情况下才能做到，时间跨度较大的政治预测便是如此。我们已经给大家展示过，对信息进行机械性汇总的结果通常比人类的判断更优。由于在预测方面的准确性，规则和算法能够更好地验证某些结果能否真正被预测。

前面的章节可能会让你形成一种印象，即算法在进行预测性判断时具有压倒性的优势，但是你有这种印象可能是被误导了。模型确实比人表现得更好，但并没有好很多。没有证据表明，在依据相同的信息进行预测时，人类表现得非常差而模型却表现得非常好。

在第9章，我们提到有一篇报告[对136项研究进行了回顾](#)，这些研究表明机械性的整合优于诊断性判断。尽管这种优势的确是“大规模且一致的”，但两者的表现差距并不大。该报告中有93项研究关注的是二选一的决策问题，它们衡量了临床医生和公式的“命中率”。总体来说，临床医生有68%的预测是正确的，而公式有73%的预测是正确的。报告中另有35项研究用相关系数来衡量预测的准确性。这些研究发现，临床医生的判断与真实结果的平均相关系数为0.3（PC=60%），而公式预测的相关系数是0.56（PC=69%）。在这两个指标上，公式总是比临床医生预测得好，但是机械性预测的有效性依然有限，使用模型并不能改变相当低的预测性的上限。

人工智能的预测性如何呢？正如前文所述，人工智能通常要比简单模型表现得好，但在大多数情况下，它的表现远称不上完美。例如，请回想一下我们在第10章中讨论过的保释预测算法，我们发现，被拒绝保释的人数如果保持不变，该算法可以将犯罪率降低24%，相比于法官所做的预测，这是一个较大幅度的改进。但是，如果该算法可以准确地预

测哪些被告会再次犯罪的话，它就能更大幅度地降低犯罪率。对未来将发生的犯罪行为进行预测是一种超自然的能力，它只存在于《少数派报告》（*Minority Report*）这种科幻小说中，因为对人类行为进行预测要面临大量的客观无知。

[由塞德希尔·穆来纳森和齐亚德·欧博迈亚（Ziad Obermeyer）完成的另一项研究对心脏病诊断进行了建模](#)。当患者有心脏病突发的迹象时，急诊医生必须决定是否需要进行检查。原则上，仅当患者心脏病突发的风险足够高时才应进行额外的检查——这些检查不仅昂贵，而且有一定的风险性和侵害性。低风险患者无须进行额外的检查。因此，医生在决定是否开检查单时需要先评估患者心脏病突发的风险。研究者建立了一个人工智能模型来完成这一评估。该模型基于大样本数据（160万名患者的440万次医保就诊记录），并且使用了2400多个变量，有如此大的数据量，该模型应该可以突破客观无知的限制。

不出所料，该人工智能模型的准确性明显超过临床医生。若想进一步评估该模型的预测性能，请你想一下，如果对那些被模型判定为具有最高发病风险（前10%）的患者进行检查，并发现其中有30%的人确实会突发心脏病，而那些被模型判定为具有中等发病风险的患者中，只有9.3%的人突发了心脏病，我们就可以据此推论，医生的表现受限于客观无知的程度，至少与其受限于判断力不足的程度相差无几。

否认无知是无知的另一种诱导

完美预测是不可能实现的，这似乎是显而易见的事。当然，断言未来是不可预测的也算不上什么具有突破性的见解。然而，众多研究证据都表明，人们做预测时会过分自信，这说明很多时候这一显而易见的事实被我们忽视了。

过分自信的普遍性让我们对非正式调查中那些相信直觉的决策者有

了新看法。我们发现，人们经常错误地将自信水平这一主观指标当成预测有效性的指标。例如，在第9章中，看完有关娜塔莉和莫妮卡的信息后，你做出了与信息相一致的判断，这时内部信号就会使你确信娜塔莉是更优的候选人。如果你对你的预测充满信心，你就已经陷入了效度错觉中：仅通过你所获取的信息进行预测，其准确性必然非常低。

那些对自己的判断极度自信的人否认自己的判断中存在噪声和偏差。他们不仅认为自己优于常人，甚至认为自己可以对一些实际上不可预测的事件进行预测。也就是说，他们从根本上否认现实的不确定性。用我们的术语来说，这种态度就是“否认无知”（denial of ignorance）。

梅尔及其追随者一直很困惑：为什么他们的研究结论未受到重视？为什么决策者总是依靠直觉来做决策？否认无知或许可以为此提供一个新的解释。当决策者聆听自己的直觉时，他们会听到内部信号，并感受到它带来的情感奖励。这种内部信号提示决策者已经做出了好的判断，它给了人们自信，让他们相信“知道是什么，却不知道为什么”的合理性。但是，对证据的真实预测效果进行的客观评估，很少能证明这样的自信是合理的。

放弃直觉确定性带来的情感奖励并非易事。因此，领导者们说，在高度不确定的情况下，他们更有可能依赖于直觉进行决策。如果现实无法给予他们把控感和信心，他们便会在直觉中寻求这些感觉。[在充满无知的情况下，否认无知就显得更加诱人。](#)

否认无知还解释了另一个谜团。面对我们之前给出的研究证据，许多高管给出了一个矛盾的结论。他们辩解道，尽管基于直觉的决策并不完美，但是，如果系统性的替代方案依然不完美，那么这样的方案就不值得采用。例如，人类判断的预测得分与员工绩效之间的平均相关系数为0.28（PC=59%），相应地，机械性预测可能做得更好，但并没有好很多，其预测的相关系数为0.44（PC=65%）。那么高管们可能会问，为什么我们还要采用机械性预测呢？

答案是，在诸如人事选拔这类重要的决策上，这种预测有效性的提高是很有意义的。某些高管会不断地在工作中做出重大改变，但得到的回报并不多。理性地讲，他们明白没有100%的成功，因此会努力做好各种决策以提高成功的可能性。他们也明白概率的含义：如果能够以相同的价格购买有65%中奖概率的彩票，没有人会去购买只有59%中奖概率的彩票。

问题在于，在很多情况下两者的“收益”是不一样的，直觉判断往往伴随着奖励，即内部信号。当算法能够获得更高的准确性时，人们更愿意相信算法，因为它的确定性带来的“收益”可以与内部信号相匹敌，甚至超越它。然而，如果替代方案是某种不一定会产生更高的预测准确性的机械性过程，人们就不愿意放弃内部信号带来的情感奖励了。

这一发现对于提高决策的品质具有重要意义。尽管所有证据都表明，机械性的预测方法和算法确实可以在一定程度上提高预测的准确性，但很多决策者仍会拒绝采用这些不让他们依据直觉进行决策的方法。只要算法还不够完美，人类的判断就不会被取代。而且在许多领域中，客观无知决定了算法不可能达到完美。这就是必须对人类判断进行改进的原因所在。

- 消除噪声

客观无知的“代价”

- 只要有预测，就会有客观无知，而且客观无知比我们想象的严重得多。我们是否想过，我们信任的专家预测的准确性真的比黑猩猩扔飞镖强很多吗？
- 如果你相信自己的直觉是基于内部信号，而不是基于你真正知道的事情，那么你就在否认自己的客观无知。
- 模型比人做得好，但没有好很多。我们发现，在大多数情况下，人类的判断能力处于中等水平，而模型略胜一筹，虽然只好了一点点，但也聊胜于无。
- 我们可能永远无法习惯于使用模型来做决策，我们需要内部信号来给自己信心。因此，请确保我们拥有最佳的决策流程。

第12章

常态谷：事情虽无法预测，但可以被理解

现在我们来考虑一个更大的问题：在这个世界上，有些问题容易解决，有些问题却充满了客观无知，那我们应如何自处呢？毕竟，在存在很严重的客观无知的情况下，我们很快就能意识到，用水晶球来预测未来是徒劳的，但我们通常并不会对这个世界产生这种体验。相反，正如上一章所提到的，我们一直在基于少量的有用信息来对未来做出大胆预测。在本章中，我们将讨论一种普遍存在的错误信念：事件虽然无法预测，却可以被理解。

这个信念意味着什么呢？我们从两个方面来讨论这一问题：一个是社会科学研究，另一个是日常生活中的经验。

预测生活轨迹

2020年，由普林斯顿大学社会学教授萨拉·麦克拉纳汉（Sara McLanahan）和马修·萨尔加尼克领衔的112名研究人员在《美国国家科学院院刊》（*Proceedings of the National Academy of Science*）上发表了一篇非同寻常的论文。他们想弄清楚，社会科学家对“社会脆弱家庭”（socially fragile families）的生活轨迹究竟有多了解。根据掌握的信息，社会科学家预测某个家庭将会发生的事件时，准确性如何呢？更具体地说，使用社会科学家通常在研究中收集的信息，来预测人们在日常生活中可能发生的事件，专家们的准确性究竟有多高？在我们看来，这项研究的目的是测量在社会科学家完成数据收集工作后，对脆弱家庭生活中可能发生的事件进行预测时还存在多少客观无知。

研究人员的资料来源于“脆弱家庭和儿童福利研究”（*Fragile*

Families and Child Wellbeing Study)。该研究是一项大规模的长期研究，追踪调查了儿童从出生到15岁的生活轨迹。这个庞大的数据库包含了5000多名儿童的家庭情况，每个家庭都有数千条信息。这些儿童大多是由美国大城市的未婚父母生下的。收集的数据信息包括孩子祖父母的教育程度和就业情况、所有家庭成员的详细信息，如健康状况、社会经济地位指数、各种问卷信息、认知能力和性格特征等。这项研究提供了一个宝贵的数据库，社会科学家已经在很充分地利用它来进行研究：根据这些关于脆弱家庭的研究数据，他们撰写了750余篇学术论文，其中很多都使用了儿童及其家庭背景的数据来解释这些孩子日后的生活表现，如高中成绩、犯罪记录等。

普林斯顿大学的研究团队所做的研究是，基于儿童早期的数据来预测他们15岁时可能出现的6种结果，包括是否流离失所、GPA以及家境的评估等。研究者采用了所谓的“共同任务方法”(common task method)，他们邀请了众多研究团队进行比赛，让这些研究团队基于“脆弱家庭”研究中收集的大量数据来预测6种结果，看谁预测得更准确。这种比赛在社会科学领域是一件新鲜事，在计算机科学领域却很常见。计算机科学领域经常会有组织请一些团队来参加比赛，完成诸如用机器翻译一段标准化的文字或从大量照片中识别动物图片的任务。获胜团队在比赛中所取得的成就定义了当时的技术水平，而且通常会在下一次比赛时被其他团队超越。社会科学领域的预测任务依靠的并不是技术的迅速发展，而是使用比赛中最准确的预测作为衡量结果可预测性的指标。这种做法是合理的。换句话说，无法预测的部分就是客观无知的残留量。

这一挑战引起了诸多研究人员的兴趣，报名的参赛者来自各个国家。最终的报告选取了160个高质量研究团队的结果。入选的参赛者大多自称数据科学家，并且在预测时使用了机器学习算法。

在比赛的第一阶段，参赛团队可以使用一半样本对应的所有数据，

[其中也包括6种结果](#)。他们会用这一“训练数据”训练预测算法，然后将算法应用于另一半样本数据。研究人员使用均方误差来衡量预测准确性，即每个案例的预测误差是实际结果与算法预测结果之差的平方。

最终获胜的模型准确度如何呢？基于海量数据集训练得到的复杂机器学习算法确实比简单线性模型的预测好，并且超过人类的判断。但是人工智能模型不比非常简单的模型好多少，预测的准确性仍然很低。[在预测“流离失所”这一事件的概率时，最佳模型的相关系数仅为0.22（PC=57%）](#)，对其他事件的预测也获得了类似的结果。比如，对抚养人是否失业或是否接受工作培训的预测，以及对孩子在“毅力”这一维度上自我评分的预测，与实际结果的相关系数为0.17~0.24（PC为55%~58%）。

在这6种结果变量中，2个是整合变量，它们的可预测性会更高一些：对儿童的GPA进行预测，其相关系数为0.44（PC=65%）；对儿童近12个月的物质贫困指数进行预测，其相关系数为0.48（PC=66%）。后一个指标是由11个问题的答案整合而来，包括“你曾经挨过饿吗”“你家的电话停机了吗”等。众所周知，整合指标一般比单一指标更具预测力，也更具可预测性。该项挑战的主要结论是：海量预测信息不足以预测人们生活中可能发生的单一事件，即使是整合变量，其预测力也非常有限。

这项研究中所报告的结果非常典型，社会科学家报告的大部分相关系数也都在这个范围内。[一项社会心理学方面的回顾性研究](#)涵盖了100多年来的25 000项研究，涉及800万名被试，该研究得出的结论是：“社会心理学效应所产生的相关系数（ γ 值）通常等于0.21。”在人体测量中常会出现更高的相关系数，比如我们之前提到的，成人身高与脚的尺码之间的相关系数为0.6，这种高相关性在社会科学领域很少见。[对行为和认知科学的708项研究进行的一项回顾性研究发现，只有3%的研究报告的相关系数大于0.5。](#)

如果你经常阅读那些“统计上显著”或“高度显著”的研究报告，那么如此低的相关系数可能会让你大吃一惊。统计学术语通常会误导读者，“显著”则是最具误导性的说法之一。当一项发现被描述为“显著”时，我们不应该下结论说这一结果的效应很强，它仅仅说明这项发现不大可能只是随机的结果。当样本量足够大时，相关性可能非常“显著”，但仍微不足道。

在关于这项挑战赛的研究中，对单一结果进行预测的有限性传达出一条令人不安的信息，那就是理解和预测之间存在差别。脆弱家庭的研究资料被认为是社会科学的宝库，确实如此，我们已经看到这些数据被广泛地应用于科学研究中，进行这些研究的学者们坚信，他们的工作将促进人们对脆弱家庭生活的理解。然而，这种促进与详尽地预测个人生活中可能发生的事件的能力无法相提并论。“脆弱家庭”挑战赛的发起者在论文摘要中郑重地告诫读者：[“研究人员必须认识到，虽然他们了解脆弱家庭的生活轨迹，但每一项预测都不够准确。”](#)

客观无知，理解和预测的上限

我们需要再次阐明这一悲观结论背后的逻辑。当“脆弱家庭”挑战赛的研究者们将“理解”等同于“预测”，或是将“预测的缺失”等同于“理解的缺失”时，他们所谓的理解就是具有特定含义的。然而，这个词还有其他的含义：当你说你理解某一数学概念或理解“爱是什么”时，你想说的可能不是你有预测能力这件事。

在社会科学研究以及大多数日常对话中，如果有人声称理解某事，他指的是理解导致该事件发生的原因。“脆弱家庭”研究中，社会科学家们收集并分析了数千个变量，现在他们想要找出观察到的结果的成因。如果医生了解病人生的是什么病，他们就会认为自己诊断出的病理就是所观察到的症状的成因。[理解就是描述因果关系，而预测能力就是衡量这一因果关系是否成立的指标](#)。相关系数这一被用于衡量预测准确性的

指标，衡量的是有多少因果关系是我们可以解释的。

如果你接触过基础统计学，并熟知“相关性并不代表因果性”这一被反复提及的警告语的话，你可能不会对前一段的最后一句话感到惊讶。举个例子，请你想一想儿童鞋子的尺码和儿童数学能力之间的相关性：很显然，其中的一个变量与另一个变量没有因果性。这种相关性源于：鞋的尺码和数学能力都会随着儿童年龄的增长而增加。这个相关性是真实存在的，它支持这样的预测：如果知道一个孩子的脚比较大，那么你可以预测他比那些脚较小的孩子的数学能力更强，但你不应该从这种相关性中推导出因果性。

我们也要知道，尽管相关性并不代表因果性，因果性却意味着存在相关性，哪里有因果性，哪里就有相关性。如果你发现成年人的年龄与脚的尺码之间没有相关性，那么你可以放心大胆地得出结论：青春期过后，年龄的增长不会使脚变大，你应该寻找其他可能导致人们脚码差异的原因。

简而言之，存在因果性就意味着存在相关性。当存在因果性时，我们应该能做出预测，并且这一相关性（即预测的准确性）可以衡量我们究竟在多大程度上理解了这一因果性。普林斯顿大学研究人员的结论是：社会科学家对诸如“流离失所”这类单一事件的预测力的相关系数为0.22，这表明了他们对这些家庭的生活轨迹的理解程度。客观无知不仅为我们的预测力设定了上限，也限制了我们的理解力。

那么，专业人士满怀信心地说理解自己所在的领域，这意味着什么？他们如何阐明所观察到的现象背后的原因，又是如何做出胸有成竹的预测的？为什么专业人士甚至所有人，似乎都会低估我们关于世界的客观无知？

因果思维，一旦发生就会被解释

如果在阅读本章的第一节时你就想知道“是什么原因导致了脆弱家庭中的孩子流离失所或其他一些结果”，那么你与研究人员的想法是一样的。你使用了统计思维（statistical thinking）：你关注了一些总体变量，如脆弱家庭的总量等；你也关注了描述总体变量的统计指标，如平均值、方差、相关性等。也就是说，你并没有专注于个别的案例。

[有一种思维模式会自发地出现在我们的脑海里](#)，它就是因果思维（casual thinking）。因果思维会创造出故事，故事中特定的人、事、物之间会相互影响。要想体验因果思维，你可以将自己想象成一名社工，你追踪调查了很多贫困家庭。你刚刚听说其中一个家庭的人目前无家可归，比如琼斯一家。你对这一消息的反应取决于你对琼斯一家的了解有多少。事情可能是这样的：这个家庭的经济支柱洁西卡·琼斯（Jessica Jones）几个月前失业了，而且找不到工作，从那时起，她就只能付一部分房租，她多次向大楼管理员求情，甚至请你帮忙求情（虽然你出面了，但大楼管理员无动于衷）。在这种情况下，琼斯一家的遭遇虽然很可怜，但我们并不会感到意外。事实上，这就好像是一连串事件的必然结果一样，一场“难以避免”的悲剧终会发生。

如果我们认同这是“难以避免”的事情，就忽略了这件事本可以轻易被改变，忽略了在人生的每一个岔路口，命运都有可能走向不同的方向。琼斯说不定可以保住工作；她也有可能很快找到另一份工作；或许亲戚帮助了她；你身为社工，或许为他们一家提供了更多帮助；大楼管理员或许更体谅他们一家，并多宽限了几周时间，以便琼斯早日找到工作并交上房租。

如果已知了结局，这些可能发生的结果将和实际结果一样，都不会令人感到惊讶。无论结果如何（琼斯一家是否被驱赶），一旦已经发生了，因果思维就会让我们觉得它完全可以解释，甚至可被预测。

理解常态谷

前文所述的这种现象有心理学依据。一些事会令人感到惊讶，比如一场致命的流行病暴发、一只声誉卓著的对冲基金被证明是骗局等。我们每个人在生活中，都会遇到一些令人意想不到的事情，比如与刚认识不久的人相恋、年轻的兄弟姐妹突然去世、继承了一笔意想不到的遗产等。而另一些事则是你完全可以预料到的，比如二年级学生会在固定的时间放学回家。

大多数人的经验都介于这两个极端之间，我们有时完全可以预料到某件事的发生，有时则完全无法预料。大多数事情发生在一个宽阔的“常态谷”（valley of the normal）中，那里的事情既不是完全可以预料的，也不是完全预料不到的。比如，此刻你无法预料下一段的内容，如果我们突然写了一段土耳其文，可能会让你大吃一惊。但是我们所讨论的内容如果没有偏离主题，你就不会感到惊讶。

在常态谷中，事情的发展和琼斯一家遭到驱赶一样，事后看似乎都是很平常的，尽管我们没有预料到它们会发生，也无法进行预测。[这是因为理解现实的过程是回溯性的](#)，当那些我们完全未预料到的事件（如琼斯一家遭到驱赶）发生时，我们会自动从记忆中搜索一个可能的原因，如萧条的就业市场、大楼管理员的不近人情等。找到合适的理由后，搜索就会停止。如果相反的结果发生了，这一搜索过程将同样找到令人信服的原因，如琼斯的坚强、管理员的通情达理等。

上述例子表明，通常情况下，许多事件从表面上看都是不言而喻的。你可能已经注意到，上述这两个版本中，大楼管理员似乎并非同一个人：第一个是无情的，第二个是友善的。无论如何，了解这个经理的唯一线索就是看他的所作所为。基于我们当下对他的了解，他的举止看起来是合乎逻辑的。正是因为事件已经发生了，你才能明白它发生的原因。

当你以这种方式解释那些“意料之外，但情理之中”的结果时，你最终会得到一个有意义的解释。这就是我们所谓的“理解”了故事，这也解

释了为什么现实在事后看起来是可预测的，因为该事件在发生时便解释了它发生的原因，于是我们产生了一种错觉，认为它是可以预测的。

我们对世界的理解，取决于我们编造故事来解释我们所观察到的事件的能力。并且，我们几乎总是能成功地找到原因，因为我们可以从无数的事实和信念中搜寻原因。例如，常听晚间新闻的人都知道，很少有无法解释的股市大波动。相同的新闻事件既可以用于解释股票指数的下跌（投资人因这一波动消息而担忧），也可以用于解释其上涨（投资人依然持乐观态度）。

如果找不到一个显而易见的原因，我们会在第一时间创造一个解释，以便将我们构建的世界模型中的空白填补上。这就是我们推断出未知事实的方式，比如在琼斯的例子中推断大楼管理员是一个友善的人。只有当我们的世界模型无法调整，无法对某一结果做出解释时，我们才会将这一结果归类为“令人惊讶的”，并开始为它寻求更为复杂的解释。只有当常用的后见之明不起作用时，我们才会真正感到惊讶。

这种持续对现实世界进行的因果解释就是我们“理解”世界的方式，我们对人生的理解，由对常态谷中不断产生的事情的事后解释所组成。这种感觉从根本上来讲是有因果性的：新事件一旦被人们所知，其他可能性就会被消除——编故事的过程几乎消除了所有的不确定性。正如我们在那些[“后见之明”的经典研究](#)中了解到的：即使主观上的不确定性存在过一段时间，当不确定性消失后，对不确定性的记忆也将消失不见。

内部视角和外部视角

我们来对比一下统计模式和因果模式这两种思考事件的方式。因果模式通过将事件实时划分为正常事件或异常事件，为我们节省了很多精力。在异常事件的迅速动员下，我们会从环境和记忆中努力搜索相关信

息。主动的期待，即专心地等待某一事件发生也需要耗费心力。相反，常态谷中的事件发展过程，几乎不需要我们耗费心理成本。你在路上遇见邻居时，他可能会向你微笑，或者看起来心事重重，因此只是冲你礼貌地点点头，如果这两种情况过去经常发生，那么它们就不会引起你的关注。如果他的笑容异常灿烂或点头过于敷衍，你很可能就会从自己的记忆中寻找可能的原因。因果思维避免耗费不必要的心力，同时保留了发现异常事件时所需的警惕性。

与因果思维不同，统计思维通常是费力的，它需要的注意力资源只有系统2思维（缓慢而审慎的思维模式）发挥作用时才能满足。除了基础水平外，统计思维还需要经过专门的培训才能掌握。这种思维方式基于整体的信息，将个别案例视为更大类别中的一个实例。它不会把琼斯一家的遭遇看作由一系列特定事件导致的必然结果，而是观察与琼斯一家具有相同预测性特征的先例，再判断在统计上这是不是有可能发生的结果。

这两种观点之间的区别是本书反复出现的议题。依赖单一案例进行因果思维是预测误差的重要来源，采用统计思维，也称为外部视角（outside view），是避免这些误差的方法之一。

我们需要强调一点：因果模式对我们来说自然得多，即使是那些原本基于统计思维的解释，也很容易转换成基于因果思维的解释。想一想“他们之所以失败是因为缺乏经验”或“他们之所以成功是因为有一位出色的领导者”这类表述，你很容易就能想到反例：有些经验不足的团队成功了，而拥有杰出领导者的团队却失败了。经验和才华与成功之间的相关性至多是中等水平，甚至可能更低，然而我们很容易做出这种因果性归因。只要因果性是合理的，尽管相关性很小，我们的思维也很容易将相关性转化为因果性。比如，杰出的领导者被视作成功的一个原因，而经验不足则是失败的一个原因，这样的解释令人满意。

如果不想放弃对世界的理解，我们或许要不可避免地依赖有瑕疵的

解释。然而因果思维和自以为了解过去的错觉，都会导致我们在预测未来时过分自信。我们将看到，偏好因果思维会使人们忽略噪声这一误差来源，因为噪声从根本上说是一个统计学概念。

尽管整个世界的可预测性比我们认为的要低，但因果思维帮助我们了解了一个比我们想象中还难以预测的世界，这也解释了为什么我们会认为世界比实际情况更具可预测性。在常态谷中，没有惊喜，也没有不一致，未来似乎与过去一样是可预测的。人们既“听”不到噪声，也“看”不到噪声。

- 消除噪声

被限制的理解

- 在日常事件中，事件之间的相关系数普遍为0.2（PC=56%）左右。
- 相关性并不代表因果性，但因果性意味着存在相关性。
- 大多数常规事件既不在我们的预料之中，也不会令我们感到惊讶。
这样的事不需要解释。
- 在常态谷中，一些事件既不在我们的预料之中，也不令我们感到惊讶，它们是不言而喻的。
- 我们认为自己理解正在发生的事情，但是我们料到了它会发生吗？

NOISE

第四部分

噪声是如何产生的

- 发现噪声

为什么噪声无处不在却很少被重视

噪声和偏差源自哪里？什么样的心理机制使我们的判断有了差异，又是什么样的心理机制使我们的判断出现了同样的错误？简言之，我们对噪声的心理机制了解多少？这是我们接下来将要讨论的问题。

首先，我们将介绍快速的系统1思维的一些运作方式，并说明它们是如何导致一些判断误差的。在第13章，我们将介绍系统1思维广泛依赖的3种重要的判断启发式（heuristics）。我们会解释这些启发式是如何导致可预测的、具有方向性的误差（统计偏差）以及噪声的。

第14章着重讨论匹配系统，这是系统1的一种特殊的思维方式，并且我们会讨论它可能产生的误差。

在第15章，我们将讨论所有判断过程中不可或缺的成分：做判断时所使用的量表。我们将为大家展示，选择合适的量表是做出良好判断的先决条件，无法清楚地定义量表或选择了不合适的量表是噪声的重要来源。

第16章将探讨噪声的一种心理来源——它可能是噪声的来源中最为有趣的一种，面对不同情境时，不同的人具有不同的反应模式。就像每个人的个性一样，这些反应模式也不是随机的，随着时间的推移，这些反应模式大多是稳定的，但是它们产生的影响是不容易预测的。

最后，第17章会总结我们当前对噪声及其成分的了解。这样一来，我们可以解答之前提出的谜题：为什么噪声无处不在却很少被重视？

第13章

启发式、偏差与噪声

本书是对横跨半个世纪的关于人类直觉判断研究的延续，也就是所谓的“启发式和偏差研究项目”（heuristics and biases program）。[《思考，快与慢》](#)一书对该研究项目前40年的研究内容进行了回顾，探讨了能够解释“直觉思维的奇妙与缺陷”的心理机制。该项目的核心思想是，当回答一个难题时，人们会使用简化的思维操作系统——启发式。通常来说，由快速的直觉思维（系统1思维）产生的启发式是非常有用的，并且能够得出适当的答案，但是有时候它们也会产生偏差，我们称之为系统性的、可预测的判断错误。

启发式和偏差研究项目关注的是人们的共性，而非差异。研究结果表明，导致产生判断误差的心理过程是人们共有的。因此，熟悉心理偏差（psychological bias）概念的人往往会认为心理偏差总是会导致统计偏差（statistical bias）。本书所使用的“统计偏差”这一术语，指的是大部分朝同一方向偏离真实值的测量或判断。毕竟，只有当判断者具有相同的心理偏差时，所导致的才是统计偏差。而若判断者以不同的方式在不同程度上出现心理偏差，就会导致系统噪声。当然，无论心理偏差引发的是统计偏差还是噪声，都一定会导致误差。

诊断偏差

我们通常是通过参考真实值来确定判断偏差的。在做预测性判断时，如果误差主要偏向同一个方向而非其他方向，偏差就出现了。例如，人们在预测完成一个项目需要多长时间时，他们估计的平均值通常比他们实际需要的时间少得多。这种常见的心理偏差就是计划谬误

（planning fallacy）。

不过，在做判断时我们往往没有真实值可以拿来比较。我们曾强调过，只有知道真实值时才能检测到统计偏差，那么，在真实值未知的情况下应该如何研究心理偏差？答案是，研究者如果观察到本不应该影响判断的因素对判断产生了统计效应，或是本应该影响判断的因素没有对判断产生统计效应，那么就可以证实判断过程存在心理偏差。

让我们再次用射击的例子阐明上述方法。想象一下，A、B两支队伍完成了射击，而这一次我们请你隐去靶子（见图13-1）。在这个例子中，你并不知道靶心究竟在哪里（即真实值未知），因此，你不知道这两支队伍的子弹落点偏离靶心的程度分别有多大。但是你被告知，在左图所示的情况下，两支队伍瞄准的是同一个靶心；而在右图所示的情况下，A、B两队瞄准的是不同的靶心。



图13-1 隐去靶子后观看射击结果以验证偏差

尽管靶心未知，但两种射击结果都证明了系统性偏差的存在。在左图所示的情况下，两支队伍的子弹落点在不同的区域，尽管它们本应该在相同的区域。这种模式就类似于你在实验中，观察两组投资者阅读内容完全相同的商业计划书，只是打印计划书所用的字体和纸张不同，如果这些无关紧要的细节影响了投资者的判断，就表明存在心理偏差。我们不清楚究竟是那些被圆滑的字体、光洁的纸打动的投资者判断得过于积极，还是那些阅读纸张较粗糙的计划书的投资者判断得过于消极，我们只知道他们的判断是不同的，尽管他们不应该这样。

右图所示的情况说明了相反的现象。由于两个队伍瞄准的是不同的靶心，因此两个队伍射中的区域应该是不同的，但落点却集中在了同一区域。例如，假设向两组人提出第4章中关于迈克尔·甘巴迪的那个问题，但问题有所不同：其中一组像你一样，被要求估计甘巴迪两年内仍能留任的可能性；另一组则被要求估计他三年内仍能留任的可能性。按理说两组应该得出不同的结论，因为三年内不能留任的可能性显然比两年内的更高，但是[证据表明，两组人所估计的概率相差很小，小到可以忽略不计](#)。答案本应该是有显著差异的，但事实并非如此，这意味着本应该影响判断的因素被忽略了。这种心理偏差被称为范围不敏感（scope insensitivity）。

判断的系统性错误已在许多领域中出现，“偏差”一词现在已被广泛用于多个领域，包括商业、政治、政策制定和司法。由于该词的普遍使用，它的含义也变得很广泛。我们在此处使用的是这个词在认知层面的定义，指的是一种心理机制和这种机制通常会产生的误差。除此之外，这个词还经常用于暗示某人对某个群体有偏见，如性别偏见或种族偏见。它也可能意味着某人支持某个特定的结论，例如，我们有时候会读到：某人因利益冲突或政治见解而产生偏见。所有这些心理偏差都会引起统计偏差和噪声，因此我们在讨论心理学的判断误差时，会将上述这些类型的偏差、偏见都包含在内。

我们强烈反对这个词的另一种用法——将代价高昂的失败归因于无特定所指的“偏差”，而在承认错误的同时，承诺努力消除我们的决策中的偏差。这些表述无非是说“错误已经铸成”“我们将努力做得更好”，再没有其他意义了。可以肯定的是，某些失败确实是由与特定心理偏差相关的可预见性误差引起的，我们也相信采取某些干预措施来减少判断和决策中的偏差和噪声是可行的，但是，将所有不好的结果都归咎于偏差，是毫无价值的解释。我们建议，只有在描述具体、可以识别的误差以及产生这些误差的机制时，才使用“偏差”一词。

替代偏差，我们常常用一个问题代替另一个问题

要体验启发式的过程，请你尝试回答如下问题，这个问题阐明了启发式和偏差的几个基本的主题。像之前一样，如果你能给出自己的答案，你将从中收获更多。

比尔今年33岁。他很聪明，但缺乏想象力，且大多数时候比较无趣。学生时代，他数学学得很好，但是社会研究和人文科目学得很差。

下表列出了8种可能发生的“比尔的现状”。浏览以下各项，然后从中选出你认为最有可能发生的两项。

- ☐ 比尔是内科医生，爱好玩扑克牌。
- ☐ 比尔是建筑师。
- ☐ 比尔是会计师。
- ☐ 比尔爱好演奏爵士乐。
- ☐ 比尔爱好冲浪。
- ☐ 比尔是记者。
- ☐ 比尔是会计，爱好演奏爵士乐。
- ☐ 比尔爱好爬山。

现在，请你再浏览一遍，然后选择两个与比尔最为相似的描述。你的选择可以与上次的相同，也可以不同。

几乎可以肯定的是，你所选择的最可能的类别与最相似的类别是完全相同的。我们如此肯定的原因是，[已经有许多实验表明：人们在这两](#)

[类问题上会给出相同的答案](#)。但是，相似性和可能性实际上是完全不同的。例如，你可以自问，如下哪种说法更合乎情理呢？

- 比尔符合我心目中爵士乐爱好者的形象。
- 比尔符合我心目中爱好演奏爵士乐的会计师的形象。

这两种陈述都不完全符合，但其中一个显然看起来没那么糟糕。相比爵士乐爱好者，比尔与“爱好演奏爵士乐的会计师”有更多的共同点。现在考虑一下：以下哪个陈述更有可能发生？

- 比尔爱好演奏爵士乐。
- 比尔是会计，且爱好演奏爵士乐。

你可能会选择第二个答案，但是这样做是不合逻辑的。比尔爱好演奏爵士乐的可能性必然高于他是一个“爱好演奏爵士乐的会计师”的可能性。回想一下文氏图（Venn diagram）[\[7\]](#)，如果比尔是爱好演奏爵士乐的会计师，那么他必然是个爵士乐演奏者。在描述中添加更多的细节只能使所描述内容发生的可能性降低，尽管添加细节能够使它变得更具代表性，而且更符合情况，就像当前的案例一样。

根据判断启发式理论，有时候人们会用一个简单问题的答案来回答一个较难的问题。因此，哪个问题更容易回答呢？是“比尔与一个典型的爵士乐乐手有多么相似”，还是“比尔有多大可能性成为一名业余爵士乐乐手”？众所周知，相似性问题更容易，因此人们在评估概率时，更可能给出与之相关的相似性问题的答案。

现在你已经体验到了启发式和偏差计划的核心思想：回答一个难题的启发式是去寻找一个简单问题的答案。用一个问题代替另一个问题会导致源于可预见性心理假设的误差，我们称之为心理偏差。

在比尔的例子中，心理偏差很明显。当相似性判断代替概率判断时，误差必定会发生，因为概率判断受限于特定的逻辑。也就是说，文氏图适用于概率，却不适用于相似性，正因如此，许多人都会犯源于可预见性心理假设的逻辑错误。

再说一个忽略统计属性的例子。请回想一下你在第4章中对甘巴迪是否成功问题的看法。大多数人对甘巴迪成功可能性的评估完全基于例子中给出的关于他的信息，并会尝试将关于他的描述与成功CEO的形象进行匹配。

你是否考虑过一个随机选择的CEO有多大可能在两年后仍然留任呢？你很可能不会考虑这个问题。你可以将这种基准概率信息（base-rate information）视为衡量一个CEO留任难度的方法。如果这种方法看起来很奇怪，那么请试想一下你会如何估计某个学生通过考试的概率。考试不及格的学生比例当然是很有价值的一个指标，因为它可以反映这场考试的难度。同样的道理，CEO的留任率对甘巴迪问题也是至关重要的。这两个问题都是采用我们所谓的外部视角的例子：当你采用外部视角时，你会把该学生或甘巴迪想象成一群相似案例中的一员。你会用统计思维来看待这个群体，而不是用因果思维去分析具体的案例。

采取外部视角可以扭转乾坤，防止出现重大错误。花几分钟进行一些调研你就会发现，[美国企业的CEO离职率大约为每年15%](#)。这一统计数据表明，一位新上任的CEO平均约有72%的概率在两年后仍然处在该职位上。当然，这个数字只是一个方面，在甘巴迪的案例中，那些具体信息都会影响你的最终估计。但是，如果你只关注被告知的与甘巴迪有关的信息，那么你会忽略一项关键信息。我们举甘巴迪的案例，原本是为了说明判断充满噪声，可我们花了几个星期才意识到，这个案例本身就很好地体现了我们现在介绍的这种偏差，我们称之为忽略基准概率偏差（base-rate neglect）。可见，即使是本书的作者也不能比其他人更自然地想到基准概率这件事。

用一个问题替代另一个问题不仅限于相似性和可能性之间的替代，另一个例子是将概率判断替换成在脑海中获取案例难易程度的判断。例如，人们对飞机失事或发生飓风的风险认知会在此类事件被媒体广泛报道后短暂上升。从理论上讲，风险判断应基于长期平均值，但实际上，近期发生的事件会被赋予更高的权重，因为它们更容易被人想起。用例证在脑海中出现的难易程度来替代频率判断，被称为“可得性启发式”（availability heuristic）。

用简单的判断替代困难的判断并不仅限于这些例子，实际上，替代是很常见的。我们常常会把回答一个较简单的问题视作一个“通用程序”，并以此来应付一个可能难住我们的问题。请思考一下，我们在回答下列问题时，往往会用什么比较容易的问题替代（例子中每个问题下方都提供了一个可能的选项）？

- 我相信气候变化吗？
我相信那些说它存在的人吗？
- 我是否认为这位外科医生有能力？
这个人说话充满信心和威严吗？
- 项目会如期完成吗？
现在项目已经按计划在执行吗？
- 核能是必要的吗？
我会闻核色变吗？
- 我对自己的生活感到满意吗？
我现在的心情如何？

不管问题是什么，将一个问题替代为另一个问题，都会导致你在给出答案时无法对证据的不同方面赋予恰当的权重，而错误地对证据赋予权重会不可避免地产生错误。例如，要想得到你关于生活满意度问题的确切答案，需要询问的显然不仅仅是你当前的情绪，现在也已经有相关研究证据表明，情绪实际上是被过分重视了。

同样，用相似度代替概率会导致忽略基准概率偏差，而基准概率在判断相似度时是完全不相关的因素。比如，在评估一家公司的价值时，应当很少或完全不考虑诸如公司计划书的美观度这类不相关的因素。这类因素对判断产生的任何影响都可能反映出你对证据赋予权重的不当之处，并导致错误的产生。

结论偏差，我们常常依结论寻找证据

[在写《星球大战》（*Star Wars*）第三部的电影剧本《绝地归来》（*Return of the Jedi*）的关键时期，该系列电影的制作人乔治·卢卡斯（George Lucas）与他出色的合作者劳伦斯·卡斯丹（Lawrence Kasdan）展开了激烈辩论。](#)卡斯丹强烈建议卢卡斯：“我认为你应该‘杀死’卢克·天行者并让利娅公主接手。”卢卡斯当即拒绝了这个建议。卡斯丹又建议，如果让卢克活着，那么另一个主角就应该死掉。卢卡斯再次拒绝，并补充道：“你不要光想着‘杀人’。”于是卡斯丹真心诚意地说出了自己对电影本质的看法。他向卢卡斯解释道：“如果你喜欢的人在旅途中遭遇了不幸，这段旅途才显得更有吸引力。这将使这部电影的情感分量大大增加。”卢卡斯的回答迅速而明确：“我不喜欢那样，也不相信那一套。”

这里的思维过程与你在考虑比尔是不是爱好演奏爵士乐的会计师时所经历的思维过程完全不同。再看一下卢卡斯的回答：先是“不喜欢”，然后是“不相信”。卢卡斯对卡斯丹的提议自动地做出了回应，这种自动化的回应有助于激发他的判断力，事实证明，他是正确的。

[这个例子说明了另一种类型的偏差，我们称之为结论偏差或者预判](#)。像卢卡斯一样，我们经常在判断过程初期就对某个特定结论有了一种倾向性。当我们这样做时，我们让自己头脑中的快速、依赖直觉的系统1思维给出结论，然后，我们要么绕过收集和整合信息的过程直接得出结论，要么转而动员系统2思维进行审慎性思考，找到支持我们预判结论的论据。[在这种情况下，证据就是有选择性且失真的](#)：由于证实性偏差和期望偏差，我们倾向于有选择地收集和解释证据，以支持那个我们已经相信或希望成真的结论。

人们通常会为自己的判断找到一些看似合理的解释，并且真的认为这是他们的信念的根据。验证结论偏差的一个很好的方法是，想象一下如果那些看起来支持我们信念的论点突然被证明是不成立的，情况会如何。回到前面的例子，卢卡斯向卡斯丹明确指出的“你不要光想着‘杀人’”并不是什么令人信服的论点，《罗密欧与朱丽叶》（*Romeo and Juliet*）的编剧应该不会同意卢卡斯的观点，如果《黑道家族》（*The Sopranos*）与《权力的游戏》（*Game of Thrones*）的编剧决定不再“杀人”，这两个剧集很可能播完第一季就会被下架。然而，我们敢打赌，这些强硬的反驳观点都不会改变卢卡斯的想法，相反，他很可能会提出其他的论点来支持他的判断，比如“《星球大战》是与众不同的”。

无论我们从哪个方面看，预判都是显而易见的。就像卢卡斯的反应一样，预判通常带有感情色彩。心理学家保罗·斯洛维奇（Paul Slovic）称其为情感启发式（affect heuristic）：人们通过询问自己的感受如何来决定自己如何思考。对于我们喜欢的政治家，我们会爱屋及乌，而对于不喜欢的政治家，我们连他们的外表和声音也都不喜欢，这就是那些聪明的公司要付出那么多努力使其品牌与积极情感产生联结的原因之一。教授们经常注意到，如果某一年他们获得了较高的教学评分，学生也会对课程所用的教材给予好评；而如果某一年学生不太喜欢这个教授，他们会给相同的教材差评。即使这一过程不涉及情感成分，相同的机制也会起作用：无论你的信念的真正来源是什么，[你总会倾向于接受任何看](#)

[起来支持该信念的论点，即使推理是错误的。](#)

结论偏差的一个更精妙的例子是锚定效应（anchoring effect），它是指一个任意数字对人们的定量判断产生的影响。[在一个典型的演示实验中，你可能会看到许多不容易猜出价格的物品](#)，例如一瓶陌生的葡萄酒，实验人员会要求你写下自己社保号码的最后两位数字，并问你是否愿意为那瓶酒支付同等金额的费用。最后，你需要说出你愿意为葡萄酒支付的最高费用。结果表明，利用社保号码所设定的锚点会影响你最终的购买价格。一项相关研究发现，那些因社保号码产生较高锚点的人们愿意支付的价格（超过80美元）比那些锚点较低的人们愿意支付的价格（不到20美元）要高出3倍之多。

显然，你的社保号码本不应该对判断葡萄酒价格产生影响，但事实就是如此。[锚定效应是一种非常强大的效应，谈判中经常被刻意用到](#)。无论你是在集市上讨价还价还是进行复杂的商业谈判，先下手为强总是理智的选择，因为接收锚定信息的人会不由自主地去思考你的报价在哪些方面可能是合理的。人们总是会试图理解他们所听到的，当人们遇到难以置信的数字时，他们会本能地想要去尽力消除这种难以置信的感觉。

过度一致性偏差

有一个实验可以帮助你体验第三种偏差。你将读到关于一位高管候选人的描述，该描述由4个形容词组成，分别被写在一张卡片上，并且这叠卡片刚刚做过洗牌处理。前两张卡片的描述如下：

聪明，执着。

按理说，你应该等到获得完整信息后再进行判断，但事实并非如此：此时你已经对候选人进行了评估，并做出了判断，但你无法控制这

一过程，你的判断过程根本无法暂停。接下来，翻开后两张卡片，现在你获得了关于这位候选人的完整的描述：

聪明，执着，狡猾，没有原则。

这时，你的评价可能就没有那么积极了，但是变化应该不会太大。作为比较，请你再考虑如下的描述，这可能是洗牌后的另一种情形：

没有原则，狡猾，执着，聪明。

第二种描述是由相同的形容词组成的，但是由于形容词出现的顺序不同，对我们来说，第二种描述的吸引力显然比第一种描述小得多。“狡猾”一词在“聪明”和“执着”之后出现时只有轻微的贬义，因为我们仍然（毫无道理地）相信这位高管总体上是好的。然而，对于第二种情况，由于出现在“没有原则”之后，“聪明”和“执着”也不再是积极因素了，反而可能会使坏人显得更加危险。

[这个实验说明了过度一致性（excessive coherence）偏差](#)：我们能迅速形成一致性印象，但要想改变一致性印象，过程却很缓慢。在上述例子中，我们仅依据很少的信息，就立即对候选人建立了积极的印象。证实偏差指的是当我们有预判时，我们会完全忽略那些与我们的预判相冲突的信息，这会使得我们对后续信息的重视程度过低。描述这种现象的另一个术语是光环效应（halo effect）——候选人是在积极的第一印象的“光环”笼罩下被评估的。在本书第24章中我们将看到，在招聘决策中，光环效应是一个非常严重的问题。

再来看另一个例子。在美国，政府会要求连锁餐厅标明食物的热量，以确保消费者能看到他们想要购买的芝士汉堡、牛肉汉堡和沙拉等食物的热量。然而，看到这些标签后，消费者真的会改变他们的选择吗？从现有证据来看，这一话题是存在争议的，但[有一项研究能给我们](#)

[一些启示](#)，该研究发现如果将标签放置在食品的左侧而不是右侧，消费者就更有可能受到热量标签的影响。当热量标签在左侧时，消费者会先接收该信息，显然他们会在看到食物之前就先想到“热量很高”或“热量不算太高”。他们最初的积极或消极反应将极大地影响他们的选择。相比之下，如果人们先看到食品，显然他们就会在看到热量标签之前先想到“美味”或“不太好吃”，同样，他们最初的反应将极大地影响他们的选择。针对希伯来语使用者的研究也支持了这一假设。希伯来语使用者是从右向左阅读的，热量标签在右侧会比在左侧时对希伯来语使用者的影响更大。

也就是说，我们总是会直接下结论，然后坚持己见。我们认为我们的观点是有依据的，但是我们所认为的依据和我们对它的解释很可能是扭曲的，至少在某种程度上是为了和我们的草率判断相匹配。这使得我们保持了出现在脑海中的故事的一致性。当然，如果结论是正确的，这个过程倒也无妨，但是，如果最初的评估是错误的，那么面对矛盾的证据也依然坚持己见就很可能放大错误。而且这种影响很难控制，因为我们无法忽略自己听到或看到的信息，很多时候甚至很难忘记。在法庭上，法官有时会要求陪审团忽视他们所听到的那些未经许可就呈现的证据，但这种要求是不切实际的。不过，这种要求可能有助于陪审团的审议，因为其明确了基于这种证据提出的论点可能会被驳回。

心理偏差引发噪声

到目前为止，我们简要地介绍了三种偏差，它们的运作方式是不同的：替代偏差会导致我们对证据不正确地赋权；结论偏差会导致我们要么绕开证据，要么以曲解的方式考虑它；而过度一致性偏差则会放大初始印象的效果并减少矛盾信息的影响。当然，以上三种类型的偏差都会产生统计偏差，也都会产生噪声。

我们先从替代偏差说起。如之前的例子，大多数人都会判断比尔的

形象与典型会计的形象有多相似，以此来判断比尔有多大可能性是会计师，在这个实验中，偏差是共性的。如果每一位受访者都犯同样的错误，就不会有噪声，但是，替代偏差并不总是会产生这种一致性。当把“气候变化是真的吗”这一问题替换为“我要相信那些说气候变化为真的人吗”，很容易看出答案是因人而异的，它取决于这个人的社交圈、偏爱的信息来源、政治面貌等因素。这样一来，同一种心理偏差就产生了不同的判断和个体间的噪声。

替代偏差也可能是情境噪声的来源。如果根据自己当下的心情来回答有关生活满意度的问题，那么同一个人的答案也必定因时而异。比如，一个人在早晨可能是快乐的，而在下午可能是痛苦的，情绪会随时间变化，这可能导致受访者反馈的生活满意度随调查者询问的时间不同而发生变化。在第7章，我们已经回顾了源于心理偏差的情境噪声的例子。

预判也会产生偏差和噪声。回想一下我们在引言中提到的那个例子：法官们批准庇护申请的比例差异惊人。如果一位法官只批准了5%的申请人，而同一法院的另一位法官的批准率却高达88%，那么我们可以肯定地说，他们在不同方向上存在偏差。从更宽泛的角度来看，偏差的个体差异会导致大量的系统性噪声。当然，系统也可以使大多数甚至所有判断者都产生相似的偏差。

最后，过度一致性也会产生偏差或噪声，这取决于所有（或大多数）判断者获取信息的顺序以及对信息赋予的含义是否相同。例如，一个长得漂亮、有吸引力的候选人会给大多数招聘者留下良好的第一印象。然而，如果应聘者所应聘的职位与外貌无关，那么积极的光环效应就会导致共同的错误，也就是偏差。

另外，做复杂的决策通常需要收集信息，而人们获取信息的顺序实际上是随机的。回想一下第2章中的索赔调解人，对于不同的调解人来说，不同的案件、索赔数据的获取顺序都是随机变化的，这会导致他们

对不同案件的初始印象随机变化。过度一致性偏差意味着初始印象随机变化将使最终判断随机扭曲，这就是系统噪声所带来的影响。

总而言之，心理偏差是一种普遍存在的心理机制，它们经常导致共性的错误。但是，当偏差的个体差异较大（不同的预判）或偏差的效应取决于情境（不同的触发因素）时，噪声就会产生。

偏差和噪声都会导致错误，这意味着减少心理偏差能够提高判断力。我们将在本书的第五部分讨论如何消除偏差这一主题。在那之前，我们还要继续探索判断的过程。

- 消除噪声

当偏差彼此不同，噪声便应运而生

- 即便知道了我们有心理偏差，我们还是应该克制将每个错误都归咎于所谓的“偏差”的冲动。
- 当我们用更简单的问题替代我们应该回答的问题时，一定会出现错误。例如，当我们利用相似性来预测概率时，我们往往会忽略基准概率偏差。
- 结论偏差会导致人们歪曲证据，从而使人们坚信其最初的立场。
- 我们会迅速形成一种印象，并始终维护这一印象，即使有矛盾的信息出现也依然如此。这种倾向被称为过度一致性。
- 如果许多人都有相同的偏差，那么心理偏差就会导致统计偏差。但是，通常人们的偏差彼此不同，这时，心理偏差就会导致系统噪声。

第14章

匹配，找到与你的预测最精准匹配的共识

现在请看向天空，你认为两小时内下雨的可能性有多大？你可能很容易就答出了这个问题。你做出判断时毫不费力。你可能在不知不觉中就把对天空黑暗程度的评估转换成了概率判断。

你刚刚做的事就是一种典型的“匹配”（matching）。我们将判断描述为一种思维操作，该思维操作会为主观印象或印象的某个方面在量表上找到对应的值。匹配是该思维操作的重要组成部分。当你回答“你的心情有多好？从1分到10分，请给出你的评分”或“请对你今天早上的购物体验评分，从1颗星到5颗星”这类问题时，你就是在进行匹配，你的任务是在判断量表上找到与你的心情或经历相匹配的数值。

匹配与一致性

回想一下上一章提到的比尔：比尔今年33岁，他很聪明，但缺乏想象力，而且大多数时候比较无趣。他上学的时候，数学学得很好，但人文社科类的课程的成绩却很差。我们要求你估计比尔从事各种职业和拥有各种爱好的可能性，然后我们看到你回答这个问题时用相似性判断替代了概率判断。你并没有真的去思考比尔成为会计师的可能性有多大，而是在思考他与该职业的典型形象有多相似。现在，我们转向另一个尚未解决的问题：你是如何做出这个判断的？

评估比尔的特征在多大程度上符合某个职业或某类爱好者的典型特征并不难。相比于会计师，比尔显然不像爵士乐爱好者，更不像喜欢冲浪的人。这个例子说明了匹配的非凡的通用性，在对人进行判断时，匹配的作用尤为突出。你可以回答无数有关比尔的问题。例如，如果你和

他一起被困荒岛，你会有什么样的感受？仅凭得到的那么一点信息，你可能立即就会得出一个直观的答案。不过，我们还为你带来了新的信息：据我们所知，比尔恰巧是一个有超强生存技能的经验丰富的探险家。如果这个消息使你感到惊讶（很可能会如此），这就说明你刚刚经历了一致性失调。

你之所以会非常惊讶，是因为新信息与你之前所构建的比尔的形象不匹配。现在想象一下，如果最初的说明中包含关于比尔的英勇和超强生存技能的信息，那么你将对该男子形成一个完全不同的整体印象——他也许是一个只有在户外才会变得活泼的人。因为你对比尔的总体印象本来就缺乏一致性，所以将该总体印象与职业或爱好类别进行匹配就会比较困难。不过，与之前你所经历的失调相比，你现在所经历的失调要小得多。

相互矛盾的线索使达成一致变得困难，也会使在判断中找到令人满意的匹配变得困难。复杂判断的特征就是存在相互矛盾的线索，可以预料的是，我们会在其中发现很多噪声。迈克尔·甘巴迪问题就属于这一类判断，该问题中有些线索是正面的而其他线索是负面的。在第16章，我们将再次讨论复杂判断问题。在本章中，我们先将重点放在相对简单的判断，特别是那些在“强度量表”（intensity scales）上做出的判断上。

强度的匹配

在有些判断中，我们所用的量表是定性的，例如职业、爱好和医学诊断。你可以通过看它们的测量值有没有大小、多少、强弱等之分，来判断它是不是定性的，比如“红色既不比蓝色多，也不比蓝色少”。

同时，也有许多判断的量表是定量的。比如，尺寸、重量、亮度、温度或响度等物理量的测量，成本或价值的测量，概率或频率的判断等都属于定量测量。还有一些抽象的判断也是定量测量，如信心、力量、

吸引力、愤怒、恐惧、不道德、惩罚的严厉程度等。

这些可以量化的维度有一个共同特征：对于在同一维度上的任意两个数值，都可以比较“哪个更多”。你知道鞭答是比打手心更严厉的惩罚，你也知道自己更喜欢《哈姆雷特》而不是《等待戈多》，就像你知道太阳比月亮更明亮，大象比仓鼠更重，美国迈阿密的平均气温比加拿大多伦多的更高是一样的。

人类具有非凡的直觉力，可以对比两个强度不同的量表，从而[将两个完全无关维度的强度相匹配](#)。你可以将自己对不同歌手的感情强度与城市建筑物的高度相匹配。例如，如果你特别喜欢鲍勃·迪伦（Bob Dylan），那么你可以将自己对他的热情与你所在城市的建筑物的高度进行匹配。你还可以将你所在国家或地区当下的政局紧张程度与夏季的气温相匹配。如果要求你用一本小说的长度来度量一家餐厅的美味程度，而不是让你给出常见的1~5颗星的评分，这种要求会让你觉得很奇怪，但并非完全不可行。在这种情况下，你依然能很清楚地表达你的意思。

在日常交谈中，量表的取值范围通常依赖于情境。比如对于“他攒了很多钱”这句话，如果你是在一位成功银行家的退休宴会上敬酒时说出的，那么它表达的含义与你在祝贺一位正在兼职做保姆的年轻人时说出的大不相同。“大”和“小”这类形容词的含义完全取决于参照系，“大老鼠从小象的鼻子上跑了下来”这句话能让我们深刻体会到这一点。

匹配性预测的偏差

接下来的这个难题说明了匹配的力量以及与之相关的[系统性判断错误](#)。

朱莉是一所大学的应届毕业生。请阅读下列有关她的信息，然

后猜测她的GPA是多少（标准分数为0~4.0分）：

朱莉4岁时能流利阅读。

她的GPA是多少？

如果你熟悉美国的成绩系统，那么你脑子里很快就会蹦出一个数字，这个数字很可能为3.7或3.8。你是如何快速得出这一结果的？这一猜测过程正好说明了我们刚刚描述的匹配系统。

第一步，你评估了朱莉在阅读上的早慧程度。这很简单，因为朱莉在很小的时候就能流利阅读了，这种早慧使朱莉被划分在了某个量表的某个类别之中。如果必须描述你所使用的这个量表，你可能会说它的最高类别是“极其早慧”，并且你也知道朱莉不属于这个最高类别（有些孩子在两岁之前就会阅读了），朱莉很可能属于次一级的类别，在这个类别中，孩子们虽然非常早慧但不算极其早慧。

第二步，你要判断出一个GPA值，使它能与你对朱莉早慧程度的评估相匹配。尽管你是无意识的，但你一定会寻找一个GPA值，使得它看起来符合“非同寻常但不是极其特别”这个标签。当你听到朱莉的故事时，你自然而然地就在脑海里进行了这种匹配性预测。

在执行这些评估和匹配任务时，人们如果刻意地去运算，将会花费相当长的时间，而通过系统1思维的快速思考，则可以迅速而轻松地做出判断。我们在这里讲的有关猜测朱莉的GPA的故事涉及复杂的、多阶段的心理过程，无法直接观察。“匹配”这种特殊的心理机制在心理学中虽然不常见，但支持它的证据是毋庸置疑的。我们可以确定，就如下两个问题，不管你问谁，你得到的答案都会是完全相同的两个数字，因为我们已经做了许多类似的实验。

- 朱莉所在的班级里，比她更早开始阅读的人占了多大的比例？

- 朱莉所在的班级里，比她的GPA更高的人占了多大的比例？

第一个问题很好解答：它只要求你基于现有的信息对朱莉进行评估。第二个问题需要我们对未来进行预测，当然会更困难些，但是我们的直觉思维会使我们情不自禁地用第一个问题的答案来回答第二个问题。

关于朱莉的这两个问题类似于我们先前在讨论效度错觉时描述的那两个向来容易混淆的问题。在关于朱莉的问题中，第一个问题要求你基于现有的信息评估“强度”；第二个问题询问的则是预测的强度。我们猜测，两个问题很难被区分开来。

对朱莉GPA的直觉性预测可以看作我们在第13章中介绍的心理机制的一个例子：用一个简单问题替代一个困难问题。你的系统1思维回答了一个简单得多的问题，把一个困难的预测问题简化成了：朱莉在4岁时就能阅读了，这多么令人印象深刻啊！在此之后，要把阅读早慧程度直接与GPA相对应，还需要一个步骤——匹配。

当然，只有当获得的信息与问题有关时，才会发生替代。如果你只知道朱莉跑得很快或她跳舞水平一般，在评估GPA的问题上你就一点信息都没有了。但是，任何看起来像是可以测量才智的指标都可能是可用的替代品。

用一个问题替代另一个问题时，如果两个问题的正确答案不一样，错误就会不可避免地产生。用阅读早慧程度代替GPA，尽管看似合理，但其实质是非常荒谬的。要想知道为什么，就想一想朱莉4岁以后可能会发生的事情。她有可能会经历一场可怕的事故，她的父母有可能会发生一场极不愉快的离婚风波，她也可能会遇到经常鼓励她的好老师，从而对她的人生产生巨大影响。总之，太多的事件都有可能会影响她上大学时的学业成绩。

只有当阅读早慧程度和大学的GPA完全相关时，匹配性预测才可能

是正确的，然而，事实显然并非如此。另外，完全不考虑朱莉的阅读早慧程度这一信息的做法也是错误的，因为她的阅读早慧程度确实提供了一些有用信息。最佳预测应该介于拥有完整信息时的预测和完全没有相关信息时的预测这两个极端之间。

当你只知道案例所属的类别，而对其他具体情况一无所知时，你对该案例能了解些什么？答案就是我们所谓的案例外部视角。如果要我们预测朱莉的GPA但不提供有关她的信息，那么我们肯定会预测GPA的平均值——可能是3.2，这就是外部视角的预测。因此，朱莉GPA的最佳估计值应该为3.2~3.8分。估计值的具体位置取决于信息的预测价值：你越相信阅读年龄可以作为GPA的预测因素，你的估计值就越高。就这一问题而言，信息的预测价值肯定是很薄弱的，因此最合理的预测应该接近平均GPA。有一种非常简单的技术可以纠正匹配性预测的误差，我们将在附录3中详细介绍。

我们总是难以抗拒地做出与证据相匹配的预测，尽管这种预测从统计学上看很荒谬。比如，销售经理通常会认为，去年表现比其他人更优秀的销售人员会在今年继续表现突出。有时候，当高管遇到才华横溢的应聘者时，他就会设想这位新员工将如何从企业底层晋升到最高层。制片人通常会认为：那些上一部电影大获成功的导演，他们的下一部电影也会取得成功。

上述这些匹配性预测的例子很有可能会以失败告终。另外，当情况糟糕时，匹配性预测的结果可能会过度消极。无论是过度乐观还是过度悲观，与证据进行匹配的直觉性预测往往太过极端。描述这种预测误差的一个术语叫作“非回归性”（nonregressive），因为它们没有考虑到“均值回归”（regression to the mean）这一统计现象。

然而，需要注意的是，替代和匹配并不总会影响预测。从双系统的角度来说，直觉性的系统1思维为出现的问题提供了快速的解决方案，但是这些直觉性思维必须经过反思性的系统2思维的认可，才会被转换

成信念。有时候，人们会偏好复杂性思维的反馈，而拒绝匹配性的预测。例如，相比于积极的证据，人们不太愿意对消极的证据进行匹配性预测。我们猜测，如果朱莉开始阅读的时间较晚的话，你可能就不会那么毫不犹豫地做出匹配性预测，即认为她的大学成绩差了。当获得的信息增多时，这种积极预测和消极预测之间的不对称性将逐渐消失。

我们可以通过外部视角来纠正各种直观预测。例如，在前面关于迈克尔·甘巴迪的未来前景的讨论中，我们建议你将对甘巴迪的成功概率的判断锚定在基准概率（新任CEO任期持续两年的成功率）上。对于诸如朱莉的GPA的定量预测，采取外部视角意味着将你的预测锚定在平均结果上。只有在非常简单的问题中，在可以非常确信基于当前的信息能够完美做出预测时，我们才可以忽略外部视角，而当需要做出认真严肃的判断时，我们必须考虑外部视角。

匹配中的噪声：绝对判断的局限性

我们在强度量表上区分类别的能力有限，这限制了匹配的准确性。在大小或财富维度上，有很多数值都可以被划分到“大”或“富”这种类别中，这是重要的潜在噪声源。

用富有来描述退休的投资银行家是没错的，但他到底有多富有呢？有许多形容词可供我们选择：小康、经济宽裕、富裕、超级富裕等。如果你提供许多人的财富信息，并要求你用一个形容词来形容每个人，而且不能对每个人的财富进行比较和排序，你可以分出多少种不同的类别？

发表于1956年的一篇有历史意义的经典心理学论文[《神奇的数字7，加2或减2》](#)（*The Magical Number Seven, Plus or Minus Two*），其标题就指出了我们在强度量表上可以区分的类别数。如果类别数超出这一范围，人们就会开始犯错——例如，人们会认为A应当划分到比B更

高级的类别中，但也会在一对一的比较时给予B高于A的评分。

请想象4条长度不等的5~10厘米的线段，这4条线段的长度呈等差分布。你每次只能看到其中的一条线段，你需要用数字1—4来评估该线段的长度：1表示最短，4表示最长。这个任务很简单。假设现在有5条长度不等的线段，重复这个过程，用数字1—5来评估线段的长度。这个任务仍然很容易。随着线段的增多，你从什么时候开始会犯错误呢？大约是从有7条线段时开始，7真是个神奇的数字！令人惊讶的是，该数字几乎不受线段长度范围的影响：如果线段的长度为5~15厘米，而不是5~10厘米，你仍然会在有7条线段之后开始犯错。当你听到不同大小的声音或看到不同亮度的光时，情况也是类似的。人们在同一维度上对不同刺激赋予不同类别标签的能力有一个限度：大约是7个标签。

这种区分能力的限度很重要，因为我们在不同维度上对强度值进行匹配的能力，应该不会比直接在这些维度上指定数值的能力更强。虽然匹配操作是快速的系统1思维的一种通用工具，也是许多直觉性判断的核心，但它是很粗糙的。

不过，神奇的数字7并不是一个绝对的限度，人们可以通过训练层级分类来对标签做出更好的区分。例如，在千万富翁这一群体中，我们又可以区分出好几种不同富裕程度的类别，法官也可以在重型犯罪中区分出多种不同严重程度的类别。但是，要使这种优化过程起作用，类别必须是事先存在并且边界定义清晰的。在为一组线段赋予标签时，我们很难将较长的线段与较短的线段区分开，并将它们视为两个单独的类别，因为当我们处于快速的系统1思维模式时，分类并不受主观意志控制。

有一种方法可以克服形容词量表在分辨能力上的局限性：进行对比，而不是使用标签，因为我们比较案例的能力比评价案例的能力强得多。

假设让你用一个20分制的量表来评价许多餐厅或歌手，你会如何做

呢？5分制的量表很容易把握，但是20分制的量表将很难保证你的评价的可信度。比如，在5分制的量表下我们可以给Joe's Pizza餐厅打3分，但是用20分制的量表时，我们该给11分还是12分呢？解决这个问题方法很简单，但是你需要花费很多时间才能做出判断。

首先，你要用一个5分制的量表对餐厅或歌手进行评分，将其分为5类。然后，你要在每个类别中对不同的对象进行排序，通常只需做几次对比即可完成：你可能知道你更偏爱Joe's Pizza餐厅还是Fred's Burgers餐厅，以及更喜欢歌手泰勒·斯威夫特（Taylor Swift）还是鲍勃·迪伦，尽管你将它们（他们）划分在了同一种类别下。一个更简单的方法是，你可以将5个类别再分为4个等级。即使是对那些你最不喜欢的歌手，你可能也可以区分出几种不同等级的厌恶程度。

上述行为的心理机制很简单：对判断对象进行直接比较能够获得比逐一评估各个对象更精确的区分效果。判断线段长度也是同样的道理：我们比较不同线段长度的能力要比确认某个具体的线段长度的能力强得多，此外，我们在比较两条线段的长度时，准确性也会更高。

我们在比较判断方面的优势适用于许多领域。如果你对人们的财富水平只有粗略的认知，那么相比于逐一标记每个人的财富水平，在同一范围内对不同个体的财富水平进行比较会更为清晰、容易。如果你对文章进行评分，那么将文章从最优到最差进行排序会比对文章逐一阅读和评分更为精确。比较判断和相对判断要比分类判断和绝对判断更敏锐，不过，正如上述的例子所表明的那样，前者也更费时费力。

使用明确有可比性的量表对目标进行单独评估，能够保留比较判断的一些优势。在某些情况下，特别是在教育领域，涉及入选录用或升迁推荐时，推荐人通常需要将候选人限定为某一类人群（如“自己教过的学生”或“具有相同经验水平的教师”）的前5%或前20%。但是，推荐人给出的评估值可能没有太大的参考价值，因为我们无法保证推荐者始终规范地使用统一的量表。当然，在某些情况下这些评估值还是有用的，

比如，当经理对员工进行评级或分析师对投资进行评估时，可以先找到那些将90%的对象都划分到“前20%”这个类别中的人并纠正他们的错误。使用比较判断是减少噪声的方法之一，我们将在本书第五部分对此进行讨论。

许多判断任务都需要将考查对象与量表上的某个类别进行匹配，或者使用一组有序的形容词，比如用“不太可能”或“极不可能”来评定事件发生的概率。然而，这种匹配其实很粗糙，因此充满了噪声。即使人们对判断的实质达成了共识，每个人对标签的理解也可能有所不同。[即便如此，强制进行比较判断还是可能减少噪声](#)。在下一章，我们将进一步探讨错误地使用量表会如何增加噪声。

- 消除噪声

匹配

- 我们俩都说这部电影非常好，但是你的喜欢程度似乎比我的要少一些。我们使用了相同的形容词，但是我们使用了相同的量表吗？
- 我们认为这部电视剧的第2季将会像第1季一样精彩。这时候，我们做的是匹配性预测，这是错误的。
- 在评阅论文时，你很难保持评分的一致性，那么你是不是应该试着给这些论文排个序？

第15章

选取精确的量表，并多用相对判断

假设你是民事审判案件中的一名陪审员。你目前所掌握的证据大致如下，你要据此做出判断。

琼·格洛弗（Joan Glover）诉General Assistance公司案

一个名叫琼·格洛弗的6岁小女孩吞下了大量非处方抗过敏药AllerFree，需要住院治疗很长时间。大剂量的药物使她的呼吸系统变得很脆弱，因此她很容易患上哮喘、肺气肿一类呼吸系统疾病。AllerFree药瓶的儿童安全盖设计是有问题的。

AllerFree的生产商是General Assistance公司，这是一家年利润为1亿~2亿美元的大型制药公司，生产各类非处方药。美国的一项联邦法规规定，所有药瓶都必须使用儿童安全盖。General Assistance公司完全无视该法规的初衷，在生产过程中使用了一种故障率比其他产品高很多的儿童安全盖。该公司的一份内部文件提到“这种愚蠢、完全没必要的联邦法规就是在浪费纳税人的钱”，并说他们受到惩罚的风险很低。该文件还提到：“就算是因为违反了这个法规而受到惩罚，那也是极其轻微的，不过就是要求我们今后提高保护盖的安全性而已。”尽管美国食品药品监督管理局的一名官员曾就该公司的安全盖问题提出警告，但该公司仍决定不采取任何纠正措施。

接下来，请你认真思考，然后做出三个判断。

愤怒程度：

下列哪个选项最能表达你对被告行为的态度？（请圈选相应的数字。）

完全 能接受	反对		震惊		完全 无法容忍	
0	1	2	3	4	5	6

惩罚倾向：

除了支付补偿性损害赔偿外，你觉得被告应受到怎样的惩罚？（请圈选相应的数字。）

无须惩罚	中等惩罚		严重惩罚		极度严重 的惩罚	
0	1	2	3	4	5	6

赔偿金额：

除支付补偿性损害赔偿外，还应该要求被告支付多少惩罚性损害赔偿，以示惩戒，从而避免被告今后重蹈覆辙？（请在下面的空白处写下你的答案。）

_____美元

琼·格洛弗的故事是本书的两位作者卡尼曼、桑斯坦以及我们的朋友兼合作者戴维·施卡德（David Schkade）[在1998年报告过的一项研究](#)中使用过的例子，但略微做了简化。我们将在本章详细介绍这项研究，希望你能体验其中的一项任务，因为目前我们已将这项研究视为噪声审查的一个富有启发性的案例，而且它重现了本书的多个主题。

本章的重点是介绍反应量表（response scale）的作用，它是最为普遍的一种噪声来源。人们的判断有所不同，可能并不是因为他们的判断有实质性的差异，而是因为他们使用量表的方式不同。假设你对员工绩效进行评级，在0~6分的量表上，你评了4分，在你看来这是很不错的分数。但是，对于相同的量表，有人可能觉得3分就已经是不错的分数了。“量表”的描述模糊不清是一个普遍的问题。已有很多研究分析

了“[高度怀疑](#)”“清晰而有说服力的证据”“出色的表现”“[不太可能发生](#)”等模糊表达所引起的沟通障碍。用这样的短语表达的判断不可避免会产生噪声，因为说者和听者对它们的理解是不同的。

在琼·格洛弗案件的研究中，我们发现表述不够清晰的量表可能引发严重的后果。该研究的主题是经陪审团裁决的惩罚性损害赔偿中存在的噪声。从琼·格洛弗案的第三个问题可以看出，美国以及其他一些国家的法律允许陪审团在民事案件中要求行为特别恶劣的被告履行惩罚性损害赔偿。惩罚性损害赔偿是对补偿性损害赔偿的补充，目的是进一步补偿受害者。像琼·格洛弗案那样，当某种产品已经造成伤害且原告成功地起诉了该公司时，被告公司将给予原告经济补偿，用于支付原告的医疗费用、弥补原告的收入损失等。同时，原告也可能会获得惩罚性损害赔偿，用于警告被告以及与被告有相似行为的公司。在此案中，General Assistance公司的行为显然应该受到惩罚，陪审团要求该公司履行惩罚性损害赔偿属于合理的诉讼请求。

惩罚性损害赔偿制度的一个主要问题是它的不可预测性。同样的不当行为造成的损害程度不同，因此相应的惩罚也没有统一的标准。用本书的术语来说，这个系统是有噪声的。惩罚性损害赔偿的诉讼请求常常被驳回，即使得到批准，其赔偿金额也不会使总金额相较于补偿性损害赔偿的金额有太多增加。但也存在一些例外，有些案件中，陪审团所裁定的惩罚性损害赔偿金额高得离谱，令人瞠目结舌。一个耳熟能详的例子是，[一家汽车经销商被处以400万美元的惩罚性损害赔偿，理由是该公司未告知原告，他们售卖的新宝马车是重新喷过漆的](#)。

我们对惩罚性损害赔偿进行了研究，邀请了899名参与者对琼·格洛弗案以及其他9起类似案件进行了评估。在这些案件中，原告都受到了损害，并起诉了据称应负相应责任的公司。与你刚刚体验的过程不同，在本项研究中，参与者仅需针对这10起案件回答上述3个问题（愤怒程度、惩罚意向和赔偿金额）中的1个。然后，参与者被进一步划分为不

同的小组，并被告知每起案件的一个版本。在这些不同版本的案件中，原告遭受伤害的程度和被告公司的收入水平均有所不同，实验条件总共有28种。该研究的目标是验证与惩罚性损害赔偿有关的一个心理学理论，并考察金钱量表（在该实验中以美元为单位）在这套法律体系中，作为噪声的主要来源所起的作用。

愤怒让我们快速启动快思考系统

几个世纪以来，哲学家和法学家一直在争论如何确定公正的惩罚。但是，我们的假设是，这一困扰哲学家的问题对普通人来说很好解决，因为后者会用一个简单的问题代替困难的问题，从而简化任务。当有人问你，General Assistance公司应该被处以多少数额的罚款时，你会直接选择回答“我有多生气”这个简单的问题。因此，惩罚倾向的强度就与愤怒程度匹配了起来。

为了验证这种愤怒假设，我们要求不同小组的参与者要么回答惩罚倾向问题，要么回答愤怒程度问题。然后，我们比较了该研究中28种实验条件下参与者在这两个问题上的平均评分。正如假设所预期的，愤怒程度和惩罚倾向的平均评分之间的相关系数高达0.98（ $PC=94\%$ ）。[这种高相关性支持了愤怒假设，即愤怒情绪是惩罚倾向的主要决定因素。](#)

愤怒是惩罚倾向的主要驱动力，但并不是唯一的驱动力。你是否意识到，在琼·格洛弗的故事中，相比对愤怒程度进行评级，你对惩罚倾向进行评级时，有一个细节吸引了你的更多注意？如果你意识到了，我们猜这个细节应该就是琼受到的伤害。在不清楚后果的情况下，你就可以判定某种行为是否令人发指，例如本案中General Assistance公司的行为就属于此类。而判断惩罚倾向的直觉性思维往往带有报应的色彩，报应粗略的表现就是“以牙还牙”原则。对报应的这种强烈渴望解释了为什么法律和陪审团对谋杀未遂和谋杀已遂的态度会有如此大的差别。如果一个意图行凶之人“有幸”杀人未遂，他受到的惩罚将轻得多。

为了验证损害程度是否会影响惩罚倾向而不影响愤怒程度，我们给不同组的参与者展示了琼·格洛弗案及其他几起案件的“严重损害”版本和“轻度损害”版本。严重损害的版本就是本章开头介绍的那个。在轻度损害的版本中，琼“不得不住院数日，且各类药物给她造成了严重的精神创伤。当她的父母尝试让她服用维生素、感冒药等对身体恢复有益的药片时，她会失控大哭并说自己很害怕”。这个版本描述了这个孩子的创伤经历，但是其损害程度远低于第一个版本中描述的长期生理损伤。在严重损害版本和轻度损害版本中，参与者对愤怒程度的平均评分（前者为4.24，后者为4.19）几乎相同。这与我们的预期相符，因为只有被告的行为对愤怒情绪有影响，这种行为引发的后果对愤怒情绪则没有影响。相比之下，严重损害版本的惩罚倾向平均评分为4.93，轻度损害版本的惩罚倾向平均评分为4.65，虽然差异不大，但这在统计学上已经很显著了。严重损害版本中赔偿金的中位数为200万美元，轻度损害版本中赔偿金的中位数为100万美元。其他几个案例也得到了类似的结果。

这些发现体现了判断过程的关键性特征：不同的判断任务会对证据不同方面的权重产生微妙的影响。对于评定惩罚倾向和愤怒程度的参与者来说，他们自己都没有意识到自己对“法律制裁是否应基于报应原则”这一哲学问题所持有的立场。他们甚至都不知道如何为案件的各种特征分配权重。不过，他们在对愤怒程度评分时，为损害分配的权重几乎为0，而在确定惩罚倾向时给这一因素分配了相当大的权重。请回想一下，参与者只看到了该故事的一个版本。因而，他们在对更严重的损害给予更严厉的惩罚时，结果的产生并非是基于仔细比较，而是因为他们两个版本中进行了自动匹配。参与者的反应更多依赖于快思考系统而非慢思考系统。

充满噪声的量表

该研究的第二个目标是找出惩罚性损害赔偿为何会充满噪声。我们

的假设是，对于被告应该受到多严重的惩罚这一问题，陪审团的想法往往是一致的，但对于如何将惩罚意图转化到赔偿金额这一量表上，他们出现了很大的分歧。

该研究设计比较了在相同案件的判决中，三种量表中存在的噪声数量，这三种量表分别是愤怒程度、惩罚倾向和赔偿金额（美元）。为了测量噪声，我们采用了本书第6章分析美国联邦法官的噪声审查结果时使用的方法。在那项分析中，我们假设某一案件中所有人的判断均值可以被看作是无偏差、公正的值。这只是出于分析目的所做的假设。需要强调的是，这个假设可能是错的。在理想状况下，陪审员使用的量表对每起案件的判断都应该是一致的。任何偏离平均值的判断都被视为误差，这些误差是系统噪声的来源。

第6章提到，系统噪声可分为水平噪声和模式噪声。在这里，水平噪声是指不同陪审员通常的严厉程度的变异，而模式噪声是指某位陪审员对不同案件的反应相对于其自身平均值的变异。因此，我们可以将判断的总体变异分解为三个组成部分，公式见下面。

$$\begin{aligned} & \text{判断方差} \\ & = \\ & \text{公正惩罚的方差} \\ & + \\ & \text{水平噪声}^2 \\ & + \\ & \text{模式噪声}^2 \end{aligned}$$

Variance of Judgments

=

Variance of Just Punishments

+

(Level Noise)²

+

(Pattern Noise)²

这种分析方法将判断的方差分解为三个部分，分别对愤怒程度、惩罚倾向和赔偿金额这三种判断进行了分析。

[图15-1显示了分析结果](#)。噪声最少的量表是惩罚倾向，其中系统噪声占变异的51%，与公正惩罚的变异几乎一样多。愤怒程度比惩罚倾向的量表噪声量更大，达到71%。到目前为止，赔偿金额量表的噪声量是最大的，判断中有94%的变异为噪声！

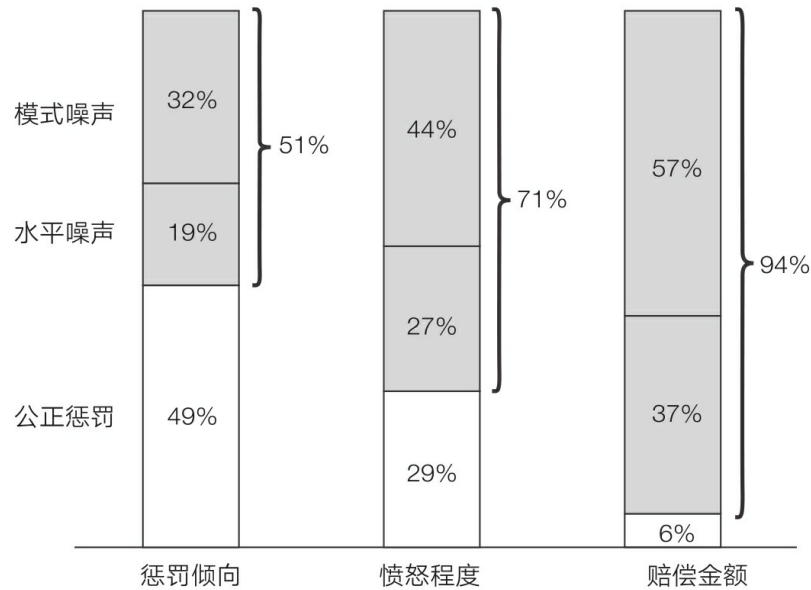


图15-1 判断变异的组成部分

三者的差异之所以令人惊讶，是因为这三种量表测量的内容几乎完全相同。我们在前面已经看到，正如愤怒假设所预测的那样，愤怒程度与惩罚倾向的值几乎完全相关。而惩罚倾向和赔偿金额恰恰回答的是同一个问题，即General Assistance公司应该接受多严重的惩罚，两者不过是单位不同罢了。那么，我们该如何解释图15-1中三种量表出现的巨大差异呢？

我们可能会认为，愤怒并不是一个非常精确的量表。没错，确实存在“完全可以接受”的行为，但是如果要有一个限度来界定你对General Assistance公司或其他被告有多愤怒时，那个限度是非常模糊的。“完全无法容忍”的行为指的是什么？量表缺乏清晰的上限，就会不可避免地产生噪声。

惩罚倾向要相对具体一些。“严厉的惩罚”比“完全无法容忍”更准确，因为法律已经规定了“极其严重的惩罚”的上限。你可能想朝被告扔鸡蛋，但是可能不会提出将General Assistance公司的CEO及其高管全部处死这样的建议。我们希望你不会。惩罚倾向量表的模糊性较低，因为它的上限更加明确。结果也正如我们所料，它的噪声更小。

在衡量愤怒程度和惩罚倾向时，我们使用了类似的评价量表，它们都是由相对明确的文字标签所定义的。但赔偿金额属于另一类量表，它的问题更大。

美元和锚定

我们关于这项研究的学术论文的标题“同样的愤怒，不同的惩罚：惩罚性损害赔偿的心理机制”（“Shared Outrage and Erratic Awards: The Psychology of Punitive Damages”）清楚地表达了其中心思想。在我们的实验中，陪审员们的惩罚倾向评分基本上是一致的，其评分大部分可以由其愤怒程度来解释。但是，赔偿金额能够更真实地模拟法庭情况，其噪声之大让人无法接受。

原因并不神秘。如果你在琼·格洛弗案件中提出了一个具体的损害赔偿金额，你肯定能真实地感受到自己在选择这个数字时实际上是比较随意的。感觉到随意性这件事本身传达了重要的信息：它能告诉你其他人也将随意地做出与你截然不同的决定，判断会充满噪声。事实证明，充满噪声确实是赔偿金额这类量表的一个突出特征。

哈佛大学著名心理学家斯坦利·史密斯·史蒂文斯（Stanley Smith Stevens）发现了一个令人惊讶的事实，那就是[人们对许多主观体验和态度的强度比例都有着强烈的直觉](#)。他们可以调整光线强度使其看起来是另一种光线亮度的“2倍”，人们会觉得10个月监禁的情感体验并不等于一个月监禁的10倍。史蒂文斯说，量表就是依据这类直觉等比绘制出来的。

你会发现，我们对金钱的直觉是通过理解“萨拉加薪了60%”或“我们的富豪邻居一夜之间损失了一半的财富”这类采用比例的表述而轻松实现的。惩罚性损害赔偿的赔偿金额量表是衡量惩罚倾向的等比量表。同其他等比量表一样，赔偿金额量表有个有意义的零点（0美元），而且

没有上限。

史蒂文斯发现，等比量表（如赔偿金额量表）可能会受单个中间锚定点（行业术语是“模数”）的约束。他在实验室中将观察者置于一定亮度的灯光下，并让观察者“将现在灯光的亮度称为10（或50、200），并相应地为其他亮度分配数字”。不出他所料，观察者分配给不同灯光的亮度数值与最初让他们采用的任意锚定点成正比。锚定点为200的观察者给出的判断值是锚定点为10的观察者的20倍。观察者判断的标准差与锚定点成正比。

在第13章中，我们介绍了一个有关锚定效应的有趣例子。在那个例子中，首先询问人们是否要支付与他们社保号后两位数字等金额的费用（以美元为单位）购买某件物品，这个问题极大地影响了他们之后购买该物品的意愿。更令人惊讶的是，最初的价格锚定点也会影响他们购买清单上其他物品的意愿。被说服愿意为无线鼠标支付较高费用的参与者也同意为无线键盘支付较高的费用。相比商品的绝对价值，人们似乎对可拿来比较的商品的相对价值更加敏感。该项研究的作者将单个锚定点的这种持续效应称为锚定的“[任意连贯性](#)”（coherent arbitrariness）。

来感受一下任意锚定点对琼·格洛弗案的影响。假设本章开头对该案件的介绍中包括以下信息：

在一个涉及另一家制药公司的类似案例中，受害者是一个小女孩，她遭受了中度的心理创伤（就像你先前读到的轻度损害版本一样）。惩罚性损害赔偿的数额是150万美元。

请注意，这时候如何处罚General Assistance公司这个问题突然就变得容易多了。确实，你可能已经想好了赔偿金额，将前面提到的赔偿金额乘以一个倍数（或比例）就可以了。这个倍数（或比例）对应琼受到的严重损害与另外那个小女孩受到的轻度损害之间的差异。此外，当你读到单一锚定点（150万美元）时，你的整个赔偿金额量表的取值范围

基本上就被确定下来了。现在，设置比上述两种情况更严重或更轻微的案件的赔偿金额对你来说也很容易了。

如果人们依据等比量表进行判断时需要锚定点，那么要是没有锚定点，将会发生什么？史蒂文斯给出了答案。在没有实验人员指导的情况下，人们在首次使用量表测量灯光亮度时，不得不随意地选择了一个值。他们将自己选的这个值作为锚定点，基于这个点做出一致性的判断。

你可能会发现，我们为琼·格洛弗案设定赔偿金额时所面临的任务，就是在没有锚定点的情况下确定量表的例子。就像史蒂文斯实验室中没有锚定点的观察者一样，你对General Assistance公司应该受到的惩罚的判断也具有很强的随意性。我们的研究的参与者也面临相同的问题：他们不得不对自己看到的第一起案件随意地做出决定。不过与你不同的是，他们不仅做出了一个随意的决定，而且还要为其他9起案件设定惩罚性损害赔偿。后面这9项判断就不再是随意的了，因为这些判断可以与最初判断的锚定点保持一致，因此它们彼此之间也更加一致。

史蒂文斯实验室的研究说明，个体生成的锚定点将对其随后做出的对赔偿金额绝对值的判断产生很大影响，但对这10起案件赔偿金额的相对值没有任何影响。如果最初判定的赔偿金额很高，那么后续判定的赔偿金额也会很高，但不影响它们的相对大小。这一推论可以得出一个令人惊讶的结论：尽管赔偿金额的判断看起来充满噪声，它却真实地反映了法官的惩罚倾向。为了测量这种惩罚倾向，我们只需用相对分数代替绝对的赔偿金额即可。

为了验证这一想法，我们将每个参与者给出的这10起案件的赔偿金额替换为按金额大小排序的分数，然后再次进行噪声分析。最高金额的排序分数为1，次高金额的排序分数为2，依此类推。这种将赔偿金额转换为排序分数的做法消除了所有陪审员的水平误差，因为除了偶尔会出现分数相同的情况外，1~10分的排序分布对每个人都是一样的。这里

要说明一点，该问卷有多个版本，每个人都评判了28个情境中的10个。我们将判断10个相同情境的参与者划分到同一组，然后对各组的反馈分别进行了分析，并且给出了平均值。

结果令人震惊：判断的噪声从94%下降至49%（见图15-2）。[将赔偿金额转换为排序](#)后我们发现，实际上，陪审员在不同案件中对什么是合适的惩罚基本达成了共识。事实上，赔偿金额排序中的噪声（如果有的话）还略低于最初的惩罚倾向评分中的噪声。

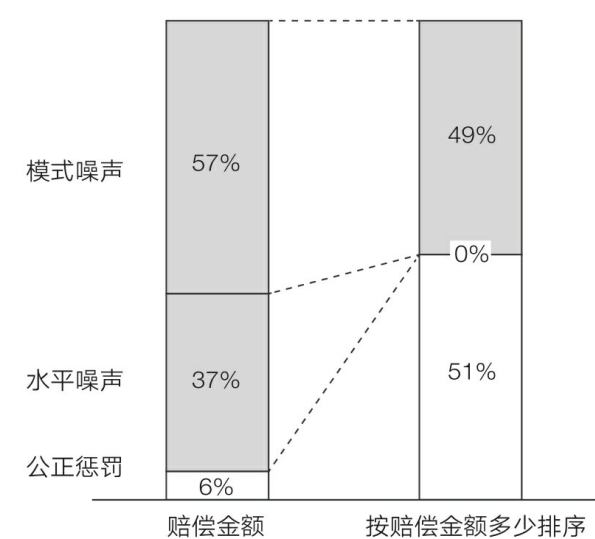


图15-2 数值中的噪声与排序中的噪声

一个不幸的结论

结果与我们提出的理论是一致的：所有案件的赔偿金额均锚定在每个陪审员在遇到的第一起案件中所选择的数字上。案件的相对排名反映了相对准确的态度，噪声水平相对较低；而赔偿金额的绝对数值基本上没有意义，因为它们取决于在第一起案件中选择中的任意数字。

颇具讽刺意味的是，陪审员在实际审判中所评估的案件往往是他们遇到的第一起也是唯一一起案件。美国的法律要求民事陪审团为一起案件设定一个赔偿金额，却不给出任何指导性的锚定点。美国的法律还明

确禁止向陪审团透露其他案件的赔偿金额。美国法律隐含的假设是，陪审员的正义感将直接引导他们根据对犯罪行为的周详思考，做出正确的惩罚判断。这个假设从心理学上说是不合理的，因为它假设了一种人类不具备的能力。司法机构应承认，司法人员也有局限性。

惩罚性损害赔偿是个极端的例子。专业判断很少用如此模糊的量表。但是，模糊的量表是很常见的，这意味着本书所介绍的惩罚性损害赔偿的研究对于商业、教育、体育、政府管理及其他领域都具有启示性。第一，量表的选取会对判断中产生的噪声量造成很大影响，因为模糊的量表是充满噪声的。第二，如果可以的话，用相对判断代替绝对判断可能会减少噪声。

- 消除噪声

重新理解量表

- 我们的判断有很多噪声，这是不是因为我们对量表的^{理解}不同？
- 对于作为量表参照点的锚定点，大家达成共识了吗？
- 为了减少噪声，我们应该用排序分数代替自己的判断吗？

第16章

模式噪声的构成

还记得朱莉吗？我们在第14章中提到的那个早慧的孩子。你曾尝试预测她的大学GPA。以下是有关朱莉的详细介绍。

朱莉是独生女。她的父亲是一位事业有成的律师，母亲是一位建筑师。朱莉大约3岁时，她的父亲患上了一种自体免疫性疾病，不得不居家办公。他花大量的时间陪伴朱莉，并耐心地教她读书识字。朱莉4岁时就能流畅地阅读。她的父亲也曾尝试教她算术，但朱莉觉得算术太难了。朱莉上小学时是个好学生，但她情感淡漠，不受同学欢迎。她经常独处，自从她和自己最喜欢的叔叔一起观察鸟类后，她就成了一名狂热的鸟类观察者。

朱莉的父母在她11岁时离婚了，朱莉因此非常伤心。她的成绩严重下滑，并且经常在学校大发脾气。上了高中以后，她在部分课程，比如生物课和创意写作课中表现出色。她的物理成绩非常出色，但是对其他大部分课程她都不太重视，因此她的毕业成绩只是B。

朱莉没有被她申请的名校录取，最终上了一所还不错的州立大学，主修环境学研究专业。大学的前两年，她经历了许多感情挫折，经常抽烟。大二下学期时，她产生了去医学院就读的强烈愿望，并开始认真地对待学习。

你预测朱莉毕业时的GPA是多少呢？

朱莉2.0问题：难与易

很显然，这个问题（我们叫它“朱莉2.0”）更加困难。对于“朱莉1.0”（第14章中提及的简单线索版本），你只知道她4岁时就能阅读。只有一条线索时，匹配就可以胜任该任务，你很快就能凭直觉估计出她的GPA。

如果存在多条线索，且指向同一个方向，匹配仍然可用。例如，当你读到“比尔是一位爱好演奏爵士乐的会计师”这样的描述时，你所拥有的信息（“缺乏想象力”“数学能力强”“社会科学课程成绩差”）描绘了一幅连贯、刻板的画面。同样，如果朱莉2.0中的大多数事件都与一个早慧、成就斐然的人物形象相吻合（也许在少数几个方面表现平平），你会发现这个问题并没有那么难。当手头的证据描绘出一幅连贯的画面时，快速的系统1思维就能毫无困难地做出判断。此类简单的判断问题很容易解决，而且大多数人都能给出相似的答案。

但朱莉2.0并非如此。这个问题有难度，因为存在多个相互冲突的线索。有一些线索表明朱莉有能力且积极性较高，但也有一些线索表明朱莉有性格缺陷、成绩普通。这样的故事俯拾即是，但这样的故事容易让人困惑，因为你无法对故事中的信息做出连贯的解读。当然，这种不连贯并不会使故事显得不切实际或不可信，因为人生总是比我们讲的故事复杂得多。

多个冲突的线索会导致模糊性，而模糊性决定了它是一个困难的判断问题。这也解释了为什么复杂问题比简单问题的噪声多。规则很简单：如果看待事物的方式可以有多种，那么人们就会用不同的方式看待它们。人们会选择不同的证据碎片来构建故事的核心框架，得出不同的结论。如果你发现很难构建一个故事使朱莉2.0的存在合情合理，那么其他读者肯定会构建出不同的故事来佐证他们的不同判断。正是这种差异制造了模式噪声。

你在什么时候对自己的判断充满信心呢？一般必须满足两个条件：你所相信的故事前后必须完全连贯，并且不存在其他有吸引力的合理解释。当你解释的所有细节都与故事完全吻合且相互强化时，这个故事就达成了完全连贯。当然，你也可以通过忽略不吻合的事件，或做出额外的解释来实现这种连贯性，尽管这样达成的连贯性可能不那么完美。对于其他合理的解释来说，情况也是如此。那些真正“解决”了判断问题的专家，不仅知道为什么自己所解释的故事是正确的，而且清楚为什么其他故事是错的。而一个人如果忽视其他合理的解释，或主动压制这些解释，他也能获得程度相当但品质较差的信心水平。

这意味着，一个人对其判断的主观信心程度并不能保证判断的准确性。不仅如此，压制其他可能存在的解释，或许会引起一致性错觉（见第2章），而[大量证据已经证明，确实存在这样一种知觉过程](#)。如果人们无法想象出与自己的结论不同的其他可能结论，他们会自然而然地认为，其他观察者也必然得出与自己相同的结论。当然，我们当中很少有人能对自己所有的判断都高度自信，所有人都经历过不确定的事件，也许就像你刚才阅读朱莉2.0的故事时一样。我们并非一直都高度自信，[但大多数时候我们的信心都高于应有的程度](#)。

模式噪声：稳定的还是变化的

我们将模式误差定义为：一个人判断一个案例的误差中，除去案例和判断者的影响后无法解释的部分。举两个明显的例子：一个平时很宽容的法官在判决某一特定类型的被告（比如违反交通法规的人）时变得异常严厉；平时很谨慎的投资者在看到令人兴奋的创业计划时，放松了警惕。当然，大多数模式误差并非如此极端。我们观察到，宽容的法官在面对惯犯时会严厉起来，在审理年轻女性的案件时则更加宽容，这些都是不极端的模式误差。

模式误差是由变化的因素和稳定的因素共同决定的。变化的因素包

括我们曾提到的情境噪声的来源，比如法官在审理案件中的某一时刻心情很好，或是法官刚好想到了近期发生的一些不幸的事。其他因素则更为稳定，例如，某个招聘者总是对来自某些大学的应聘者非常热情，或是某个医生总喜欢建议肺炎患者住院治疗。我们可以用一个简单的方程式来描述某一判断中的误差，方程式见下面。

模式误差

=

稳定的模式误差

+

变化的（情境）误差

Pattern Error

=

Stable Pattern Error

+

Transient（Occasion）Error

因为稳定的模式误差和变化的（情境）误差是独立且不相关的，所以我们可以扩展下面的方程式来进一步分析它们的方差。

模式噪声²

$$=$$

$$\text{稳定的模式噪声}^2$$

$$+$$

$$\text{情境噪声}^2$$

$$(\text{Pattern Noise})^2$$

$$=$$

$$(\text{Stable Pattern Noise})^2$$

$$+$$

$$(\text{Occasion Noise})^2$$

同分析误差和噪声的其他成分时一样，我们可以用直角三角形两条边长的平方和来表示扩展的方程式（见图16-1）：

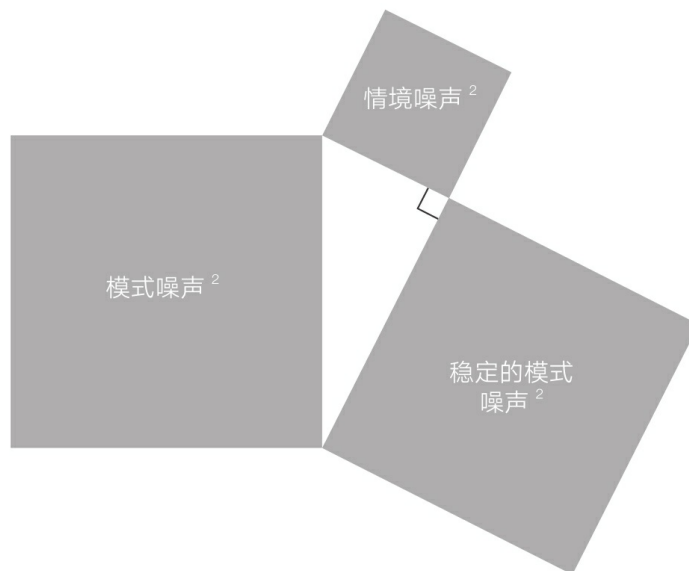


图16-1 模式噪声的成分

我们来举一个稳定的模式噪声的例子。你可以想象一下，一个招聘者要根据候选人在各个维度上的一系列得分，来预测候选人担任主管后的未来绩效。在第9章我们提到了判断模型。在该模型中，招聘者会给各个维度的评分赋予一个权重，权重与判断者对某一维度的重视程度成正比。不同的招聘者对各个维度的赋权各不相同：有的招聘者更看重领导力，而有的招聘者更看重沟通能力。这种差别导致候选人在不同招聘者心目中的排名不同。这就是稳定的模式噪声的一个例子。

对具体案例的个体化反应也会产生稳定的模式噪声，并且这种模式噪声还具有高度的特异性。想象一下，是什么导致你对朱莉的故事的某些细节更为关注，而对另一些细节不那么关注呢？可能是因为其中的某些细节与你的生活经历有许多相似之处。也许朱莉的某些事让你想起了自己的一位近亲：他离成功只有一步之遥，但最终功亏一篑——你认为这是他与生俱来的性格缺陷导致的。与此相反，朱莉的故事也可能会唤起你对某个挚友的回忆：他经历了艰难的青春后期，努力考入了医学院，现在已经成为一名成功的医学专家。对于不同的人来说，朱莉的故事所引发的联想是独特且不可预测的，但这些联想可能又是稳定的：你上个星期读朱莉的故事时联想到的人物，与现在读朱莉的故事时联想到的人物是相同的，并且你当时对朱莉的故事的个体化感受与现在的也是相同的。

判断品质上的个体差异是模式噪声的另一个来源。设想有一位拥有超级预测能力的预言家，包括他自己在内，没有人知道他有这种预测能力。他的准确性会使他偏离预测的平均值。在没有客观结果的条件下，这种偏离会被视为模式误差。当判断结果无法被验证时，超级准确的预测看起来就像模式噪声。

模式噪声还源于个体对案例的不同维度做出有效判断时其能力上的系统性差异。想象一下专业运动员团队的遴选过程。教练可能更关注与

比赛相关的各种运动技能，医生更关注运动员是否易受伤，心理学家则更关注运动员的动机和毅力。我们可以想到，当不同的专家评估相同的运动员时，大量的模式噪声将会出现。同样，全能型专家在进行判断时也会在某些方面更为擅长，在另一些方面不那么擅长。这种情况下，模式噪声更多地被描述为人们在知识、技能上的差异，而非误差。

专家们独自做决策时，技能的差异就是噪声。然而，当管理层有机会组建一个团队共同做判断时，技能的多样性将成为一种潜在优势，因为[不同的专业人士将会考虑不同方面，并相互补充](#)。我们将在第21章中讨论这种优势，以及如何获取这种优势。

在前面的章节中，我们讨论了在保险公司的客户被指派核保人时，或被告被指派法官时，所面临的两种类似抽签的情况。我们可以看到，第一次“抽签”，即从一群同行中选出一名专业人士这个过程，不仅选择了该专业人士判断的平均水平（水平误差），还选择了一个类似万花筒的组合：该专业人士独有的价值观、偏好、信念、记忆、经验和联想的组合。当你自己做判断时，你也会受这些因素的影响。你在工作中形成自己的思维习惯，并从导师那里获得智慧。获得成功帮你建立了信心，犯错让你避免重蹈覆辙。在你脑海中的某个地方，有你记住的正式规则，也有被你遗忘的规则，还有你知道可以忽略的规则。没有人能在所有方面都和你一模一样。你稳定的模式误差是独一无二的。

第二次“抽签”选择了你做判断的时机、你当时的心情以及其他本不应该影响判断却对判断产生了实际影响的看似无关的因素。这种抽签会产生情境噪声。设想一下，在你读朱莉的信息之前，你刚刚阅读了一篇有关大学生吸毒的报刊文章。这篇文章讲到了一名有天赋的学生下定决心要去攻读法学院，他为此努力学习但依然未能如愿，因为他在刚进入大学的时候曾由于吸毒落下了太多功课。因为这条信息刚刚进入你头脑中，所以这个故事会让你在评估朱莉的整体情况时更关注她的抽烟习惯。但是，如果你在几个星期后才评估朱莉的情况，你可能就不记得这

篇文章了。如果你昨天已经做了评估，你显然根本还没看到这篇文章。阅读报刊文章所产生的效应是短暂的，这就是情境噪声。

上面这个例子表明，稳定的模式噪声和不稳定的情境噪声之间并没有明显的分界线。对案例某些方面的独特敏感性是永久的还是暂时的，决定了两者的主要区别。当模式噪声的触发因素根植于我们的个人经验和价值观时，我们可以预期该模式噪声是稳定的，它反映了我们的独特性。

用人格进行类比

人们对某些特征或特征组合的反应是独特的，这一点不太直观，不太容易被理解。要想更好地理解它，我们可以思考另一种众所周知的复杂的特征组合：我们周围的人的人格。实际上，判断者对案例做出判断可以被看作人格研究领域中的一个更宽泛的研究主题下的特例，即在特定情境下人们会如何行事。对该宏大主题展开了数十年的深入研究后，我们了解到一些关于判断的知识。

心理学家长期以来一直试图理解和测量人格的个体差异。人们在许多方面彼此不同。[早期研究尝试在字典中搜寻能够描述一个人的词语](#)，并找出了18 000个单词。如今，占统治地位的人格模型——大五人格模型，将人的特质分为5个维度（外倾性、宜人性、责任心、开放性和神经质），大五模型的每个维度都包含一系列可辨识的特质。我们可以将人格特质理解为实际行为的一种预测指标。如果某人被描述为有责任心，我们就能够观察到相应的行为，如守时、信守承诺等。如果安德鲁在积极性上的得分高于布拉德，我们应该能观察到，在大多数情况下，安德鲁的表现都比布拉德更积极。然而，借助宽泛的人格特质来预测特定行为，其有效性实际上非常低。通常，相关系数为0.3（ $PC=60\%$ ）就[已经很高了](#)。

我们总是想当然地认为，行为可能是由人格驱动的，但它实际上也受情境的强烈影响。在某些情境中，没有人会表现出积极性，而在另一些情境中，每个人都会表现出积极性。安慰一位刚刚失去亲人的朋友时，安德鲁和布拉德都不会表现得很积极。但是，在足球比赛中，二人都会表现出一些拼劲。简而言之，行为必定是人格和情境的函数。

一个人与众不同且十分有趣的原因在于，人格与情境的结合并非是一个机械的相加函数。例如，情境引发的积极性强度对不同人而言是不同的。即使安德鲁和布拉德在积极性方面的平均得分相同，他们也不一定在每种情境中都表现出完全相同的积极性。也许安德鲁对同龄人比较强势，但对上级很服从；布拉德的强势表现不太受职位等级的影响，但当受到批评时，布拉德就特别容易产生攻击性，而[在身体受到威胁时](#)，他表现得非常克制。

人们在不同情境下的这种特有的反应模式，可能在时间上具有稳定性。尽管这种特有的反应模式不应被描述成广义的“特质”，但它们构成了我们所说的人格的大部分内容。虽然安德鲁和布拉德在积极性人格测试中的得分可能相同，但他们的积极性行为的触发因素或情境却不同。对于两个具有相同特质水平的人，比如同样固执或同样慷慨的两个人，我们可以这样理解：他们在该特质上的平均水平是相同的，但在不同的情境中，他们不一定具有完全相同的行为反应模式。

现在你可以看出有关人格特征的讨论与我们前面提到的判断模型之间的对应关系了。法官之间的水平差异与人格特质的个体差异相对应，代表了多种情境下行为的平均值。案件可以类比为情境。个体的平均判断只能在中等程度上预测他在某个具体问题上的判断，正如人格特质只能在中等程度上预测个体在某种特定情境下的行为一样。如果根据判断水平对判断者进行排名，那么他们在不同案例上的排名会存在很大差异，因为人们对不同案例的特征或特征组合的反应各不相同。个体判断与决策的鲜明特征就体现在他们对特征的独特敏感性，以及对具体案例

的独特反应模式中。

一般而言，人格具有独特性是好事，但本书关注的是专业判断，所以其中的变异性是一个问题，毕竟存在噪声就是存在误差。进行类比的目的是想揭示，即使模式噪声并不是随机的，我们也不可能对它进行解释。哪怕是做出不同判断的个体本身，也无法对它进行解释。

- 消除噪声

模式噪声从何而来

- 你似乎对自己的结论充满信心，但这不是一个容易解决的问题，因为线索指向了不同的方向。你是否忽略了证据的其他可能解释？
- 我和你面试了同一位候选人，我们俩平时都是很严格的面试官，但是这次我们做出了完全不同的判断。这种模式噪声从何而来？
- 人格的独特性使得人们具有创新性、创造力且有趣、生动，但是就判断而言，独特性并不是好东西。

第17章

噪声源，偏差是引人注目的图形，而噪声是不受我们关注的背景

哪里有判断，哪里就有噪声，希望你现在也认同这一观点。我们也希望，对你而言这已不再是一件超乎想象的事。这也正是我们着手开展研究的主要动力。经过多年的努力，我们对这个问题的思考已经逐渐深入和完善。现在，来回顾一下我们已经了解的噪声成分、这些成分对噪声的重要影响以及它们在判断中所起的作用。

噪声成分：影响判断的**3**种噪声

图17-1是我们在第5章、第6章和第16章中介绍的3个方程式的一个组合图，该图展示了对误差的3个连续的分解：

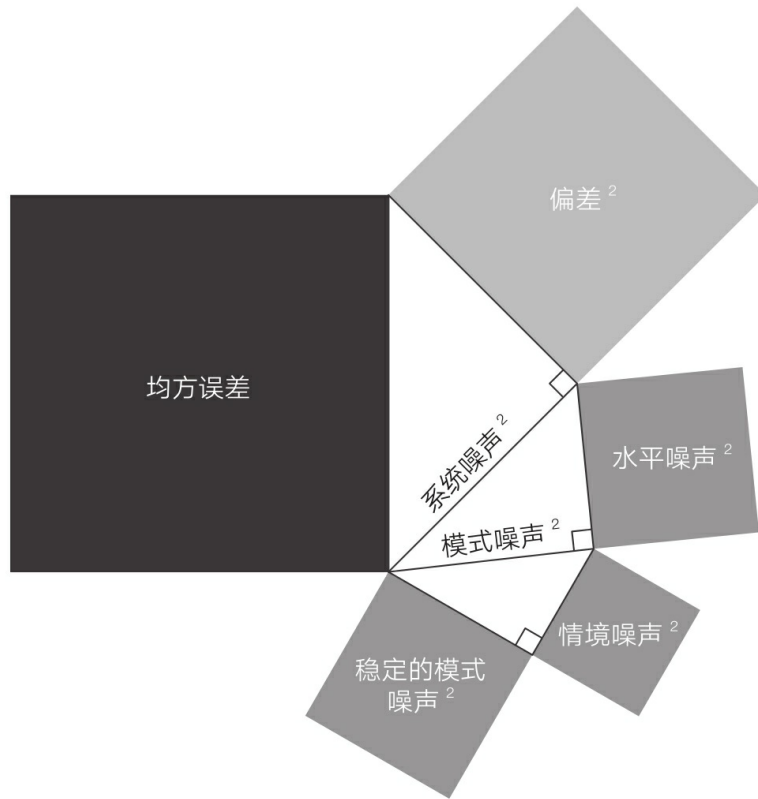


图17-1 误差、偏差和噪声的成分

- 将误差分解为偏差和系统噪声。
- 将系统噪声分解为水平噪声和模式噪声。
- 将模式噪声分解为稳定的模式噪声和情境噪声。

[现在，你可以看到均方误差如何被分解为偏差以及我们曾讨论过的3种噪声成分的平方。](#)

开始着手这一研究项目时，我们关注偏差和噪声在总误差中所占的相对权重。我们很快得出结论：相比于偏差，噪声通常是误差的主要成分，因此噪声非常值得我们进一步探索。

我们早期对噪声成分的思考基于复杂的噪声审查框架。在该框架中，多个人对多个案例做出判断。美国联邦法官的研究以及惩罚性损害

赔偿的研究，都是这样的例子。从这些研究中获得的数据能够对水平噪声进行可靠的估计。另外，虽然每个参与者都对每个案例进行了判断，但只进行了一次，因此无法判断残差是变化的还是稳定的。我们将残差称为模式误差。依据统计分析的保守性原则，残差通常被视为随机误差。换句话说，模式噪声被完全默认为情境噪声。

很长一段时间以来，这种将模式噪声解释为随机误差的传统做法限制了我们的思维。专注水平噪声，比如严厉和宽容的判断者之间、乐观和悲观的预测者之间的稳定性差异，似乎是自然而然的事。但是，有证据表明，那些无关的、变化的环境也会影响判断，从而产生情境噪声，这引起了我们的研究兴趣。

这些证据让我们逐渐认识到，人们做出的判断之所以充满噪声，很大程度上既不是因为个体的普遍性偏差，也并非变化的或随机的因素所致；特定个体对多重特征持续出现的个体化反应，决定了他们对具体案例的反应。我们最终得出结论，我们应该摒弃“模式噪声是随机变化的”这一观点。

虽然我们很想谨慎一点，尽量避免根据有限的案例做出过度概括的陈述，但是整合所做的研究后，我们发现，稳定的模式噪声实际上比系统噪声的其他成分更为重要。我们很少在同一研究中全面地探究误差的各个组成部分，因此需要进行精确的分析才能得出这样一个暂时性的结论。简而言之，以下就是我们已知和未知的内容。

成分的大小：稳定的模式噪声几乎是水平噪声的**4**倍之多

首先，我们对水平噪声和模式噪声的相对权重进行了几种估算。总体而言，模式噪声似乎比水平噪声对系统噪声的影响更大。例如，在第2章提到的保险公司的例子中，不同核保人在平均保费上的差异仅占总

体系统噪声的20%，剩余的80%都是模式噪声。在第6章提到的美国联邦法官的例子中，水平噪声（平均严厉程度的差异）在总体系统噪声中的占比略低于一半，而模式噪声则占一半以上。在惩罚性损害赔偿的实验中，系统噪声的总量在不同量表（惩罚倾向、愤怒程度或赔偿金额）上差别较大，但模式噪声的占比却基本恒定：三种量表下，模式噪声分别约占总系统噪声的63%、62%和61%。在第五部分，我们将介绍其他一些研究，特别是有关人事决策的研究，其研究的结果都与这一暂时性的结论一致。

在这些研究中，水平噪声通常不是系统噪声的最主要成分。这一事实已经传递了一则非常重要的信息，因为水平噪声是唯一一种（有时）无须组织进行噪声审查就可以监测到的噪声。当案例或多或少地被随机分配给各个专业人员时，他们的决策的平均差异就是水平噪声。例如，[针对专利局的研究发现](#)，审查员授予专利的平均倾向性存在很大差异，进而影响了有关这些专利的诉讼发生率。同样，[儿童保护部门的官员决定将儿童送到寄养家庭或寄养机构的倾向性也不同](#)，这会对儿童的命运产生长期影响。这些观测都仅仅基于水平噪声的评估。这些研究揭示的噪声问题令人震惊，但是，如果模式噪声多于水平噪声，那么它所反映的噪声问题的严重性至多只是实际情况的一半。这一暂时性结论也有例外。[法官在关于是否提供庇护的裁决中出现的令人震惊的巨大变异](#)，几乎可以肯定更多地源自水平噪声，而非模式噪声，但我们怀疑其中的模式噪声也很多。

下一步，我们通过对模式噪声的两种成分进行分解来对其进行分析。我们有充分的理由认为，模式噪声的主要成分是稳定的模式噪声而非情境噪声。对美国联邦法官在量刑中存在的噪声进行审查，展示了我们的推理过程。我们首先假设了一种极端情况，即所有的模式噪声完全是随机变化的。在这一假设前提下，法官们的量刑会随着时间的推移变得十分不稳定和不一致，乃至达到毫无道理可言的程度。因此，我们不得不做出这样的推测：[同一法官在不同情境中对同一案件量刑的平均差](#)

[异达到2.8年左右](#)。法官之间在平均量刑上的差异已经够令人震惊的了，要是同一名法官在不同情境下的量刑也存在这么大的差异，那实在是太荒唐了。因此，得出以下结论似乎更为合理：法官们会对不同的被告和不同的罪行做出有差异的反应，这种差异虽极为个体化，却是稳定的。

为了更加准确地量化模式噪声中有多少是稳定的模式噪声，又有多少是情境噪声，我们需要让同一名法官对每起案件进行两次独立评估。我们发现，在对判断的研究中通常不可能获得两个独立的判断，因为很难保证对案例的第二次判断完全独立于第一次判断。尤其是，当判断很复杂时，做出判断的人很可能会发现这是同一个问题，从而重复最初的判断。

[亚历山大·托多罗夫（Alexander Todorov）带领普林斯顿大学的一组研究人员设计了一个巧妙的实验范式](#)，攻克了这一难题。他们从“亚马逊土耳其机器人”（Amazon Mechanical Turk）^⑧网上招募了参与者。在他们的一项实验中，参与者观看了一些面孔图片（由计算机程序生成，但与真人面孔无异），并对这些面孔的各项属性进行了评分，比如亲和力、值得信赖程度等。一个星期后，实验人员用相同的面孔让相同的答卷人重复了该实验。

我们有理由认为，本实验中判断的一致性程度会低于专业人士的判断，如法官的量刑。虽然差不多每个人都会认同，有些面孔极具吸引力，有一些则令人反感，但在很大范围内，我们可以推测，对面孔的反应存在很大的个体差异。确实，观察者之间几乎没有共识。比如，在值得信赖程度的评分上，面孔图片之间的差异所导致的变异仅占判断变异的18%，剩余82%的变异都是噪声。

我们也有理由推测，这些判断的稳定性较差，因为实验参与者只是为了赚钱而参与在线答题，其判断品质会远低于专业判断。然而，实验结果发现，[噪声的最主要成分还是稳定的模式噪声](#)，其次才是水平噪

声，即不同的观察者对诸如面孔值得信赖的程度等属性的平均评分上的差异。情境噪声虽然也很大，但占比最小。

要求参与者做出其他判断，比如有关汽车或食物的偏好，或是回答一些接近专业判断的问题时，研究人员也得出了相同的结论。比如，在第15章中讨论的惩罚性损害赔偿的一项模拟研究中，参与者针对10个人身伤害案件评定了他们的惩罚倾向，两次评分之间间隔一个星期。同样，稳定的模式噪声是噪声最主要的成分。在所有这些研究中，判断者给出的结果彼此不一致，但他们自身的判断仍保持稳定。用研究人员的话说，这种“缺乏共识的稳定性”，为稳定的模式噪声提供了清晰的证据。

关于稳定的模式噪声所起作用的最有力证据，来自我们在第10章中提到的[对保释法官的大规模研究](#)。在这项杰出研究的一个部分中，研究者创建了一个统计模型，该模型可以模拟每位法官如何使用手头的线索来决定是否批准被告的保释申请。他们建立了173名法官的个性化模型，然后[运用这些模型来模拟法官对141 833个案例做出的判决](#)，每个案例产生了173个判决——判决总数超过2400万。应我们的要求，研究者慷慨地进行了特定分析，他们将判断的变异分为三部分：每个案例平均决策的“真实”变异、法官们在保释倾向上表现出的差异所导致的水平噪声，以及剩下的模式噪声。

这一分析的结果与我们的观点相符，因为该研究所测得的模式噪声完全是稳定的。这项研究是对预测法官决策的模型进行的分析，因而无法体现情境噪声的随机变异性。这些数据仅包含可验证的、稳定的个体预测规则。

结论非常明确：稳定的模式噪声几乎是水平噪声的4倍——稳定的模式噪声占总变异的26%，而水平噪声占总变异的7%。这种可识别、稳定、特异性的个体模式，要比不同法官之间审判的严格程度的差异大得多。

所有这些证据与我们在第7章中讨论的有关情境噪声的研究结果一致：尽管情境噪声的存在让人惊讶甚至不安，但没有迹象表明个体内部的变异大于个体间的变异。系统噪声中最主要的成分恰恰是最开始被我们忽略的成分——稳定的模式噪声，即法官们对特定案例做出的判决中的变异性。

考虑到相关研究较为匮乏，我们的结论只是暂时性的，但它们反映了我们对噪声的态度以及应对方式的改变。从原则上说，水平噪声——法官之间简单、全面的个体差异，应该是一个相对容易测量和解决的问题。如果有异常“严格”的评分者，或异常“谨慎”的儿童监护权法官，抑或十分厌恶承担风险的贷款人，雇用他们的机构可以试图使他们的判断维持在平均水平。例如，大学会要求教授们在评分时遵守事先确定好的成绩分布。

很可惜，我们发现，专注水平噪声会遗漏掉很大一部分个体差异。大多数噪声都不是水平差异的产物，而是交互的产物，如不同的法官如何对待特定的被告，不同的老师如何对待特定的学生，不同的社会工作者如何对待特定的家庭，不同的领导如何对待特定的公司愿景……噪声主要是我们的独特性或“判断人格”的副产品。降低水平噪声依旧是一个有价值的目标，但仅达到此目标仍然未能解决大部分的系统噪声问题。

对误差的解释

与噪声有关的很多问题都值得关注，但公众并没有意识到这些问题，有关判断与误差的讨论也几乎很少涉及噪声。尽管有证据表明噪声是存在的，并且产生噪声的机制很多，但人们很少将噪声作为影响判断的主要因素。为什么会这样呢？为什么我们从不用噪声来解释糟糕的判断结果，而是将其归咎于偏差呢？既然噪声无处不在，为什么人们很少将其作为判断误差的来源呢？

这个难题的关键在于：尽管误差的平均值（偏差）和变异性（噪声）在误差方程中起着同等作用，我们却是以截然不同的方式来看待它们的。我们对周遭世界进行合理化的常规方式让我们几乎无法识别噪声的作用。

在本书前面的章节中我们已经指出，人们在事前很难对事件进行预测，但对其进行事后解释却很容易。常态谷中的事件通常不足为奇，也容易解释。

判断也是如此。像其他事情一样，大多数判断和决策也发生在常态谷中，它们通常不会令人惊讶。如果对某件事的判断已经产生了令人满意的结果，那么该判断就是再正常不过的了，因而很少遭到质疑。当任意球的射手射门得分时，当心脏病手术成功时，当创业成功时，我们会认为决策者一定是根据正确的原因做出的选择。毕竟，这些原因已经被其结果证明是正确的。就像其他不足为奇的故事一样，一个成功的故事，一旦结果已知，就能自圆其说了。

然而，我们还是觉得需要对异常的结果进行解释，包括那些糟糕的结果，有时也包括好到出奇的结果，比如一个令人难以置信的商业冒险最终却获得了成功。用误差或特殊才能来解释这些结果是很普遍的做法，因为对于过去的重大冒险行为，一旦结果已知，人们就很容易将其解释为因为天才或愚蠢。心理学中著名的基本归因偏差说的就是这种倾向，人们通常会将责任或功劳归因于行动者。事实上，将这些结果解释为运气或客观环境所致或许更合适。另一个心理偏差，即后见之明偏差，则扭曲了人们的判断，从而导致那些无法预料的结果从事后看来是很容易预见到的。

对判断误差做出解释并非难事。为判断找原因比为结果找原因要容易得多。我们总是倾向于用动机来解释人们的判断。如果用动机不足以做出解释，我们还可以将误差归结为他们的无能。近几十年来，出现了另一种对糟糕的判断进行解释的常见心理机制：心理偏差。

心理学和行为经济学的大量研究发现了一长串的心理偏差，包括计划谬误、过分自信、损失规避、禀赋效应、现状偏差、未来过度折现（即刻偏差）等，当然也包括针对不同人群的歧视或偏好。人们已经很了解这些偏差在何种情况下会影响判断与决策，对于决策观察员怎样实时地对偏差进行识别也已知之甚多。

如果心理偏差可以被事先预测或被实时检测出来，那人们就可以对判断误差做出合理的因果解释。即使偏差只能被事后识别，只要它可以对未来做出预测，那也能够提供有意义的、暂时性的解释。例如，一名优秀的女性应聘者出人意料地应聘失败了，那么可以据此做出一个一般性的假设——存在性别歧视，可以通过对后续的招聘决策进行分析来证实或驳斥这一假设。相反，想象一个仅适用于某一具体事件的因果解释：在那个案例中，他们失败了，一定是因为他们过分自信了。这种说法没有任何实际意义，但它提供了一种让人感到满意的理解错觉。商学院教授[菲尔·罗森茨威格（Phil Rosenzweig）](#)坚定地认为，在讨论商业结果时，用偏见去进行空洞的解释是非常普遍的，这是因为人们迫切需要一个使经验合理化的故事。

从统计学上说，噪声无处不在

正如我们在第12章中指出的那样，我们惯常使用因果思维。我们会自然而然地去关注特定的事件，追踪并创建出因果连贯的故事。这其中，失败往往被归因于误差，而误差往往被归因于偏差。糟糕的判断结果很容易被解释，因而没有为用噪声解释误差留有余地。

因果思维的直接后果是我们看不见噪声。噪声本质上是一种统计概念：只有当我们采用统计思维对一组相似的判断进行思考时，噪声才可见。事实上，此时很难对噪声视而不见：它是对量刑判决和保费核准进行事后统计时呈现出的变异性；它是当你和其他人思考如何预测未来的结果时呈现出的各种可能性的范围；它是靶子上那些散布的弹孔。从因

果性上说，噪声并不存在；从统计学上说，它无处不在。

很遗憾，用统计思维看待噪声并非易事。我们毫不费力地就能为观察到的事件找出理由，但是若要对其进行统计学思考，必须学习并持续努力。因果思维是与生俱来的，而想要使用统计思维则困难重重。

这严重影响了我们是将偏差还是将噪声视为误差来源的选择。如果你接触过普通心理学，你可能还记得一幅插图：在模糊的背景中，一个突出而又充满细节的人物会脱颖而出。虽然人物很小，背景很大，但我们的注意力仍然会牢牢地集中在人物身上。图形 / 背景的例子可以与人类关于偏差和噪声的直觉进行很好的类比：偏差就是引人注目的图形，噪声则是不受我们关注的背景。这就是为什么我们在很大程度上意识不到自己的判断中存在重大缺陷。

- 消除噪声

噪声源

- 我们很容易看到判断在平均水平上存在的差异，但我们无法看到的模式噪声有多大呢？
- 你说这个判断是由偏差引起的，但如果结果不同，你还会这样说吗？你能确定其中是否有噪声吗？
- 我们专注于减少偏差，这是对的。同样，我们也应该关注如何减少噪声。

NOISE

第五部分

决策卫生，
提升五大人类判断力

- 发现噪声

对抗噪声，让你的决策更“卫生”

一个组织如何改进其专业人员所做的判断？尤其是，一个组织如何减少判断中的噪声？如果让你来解答这些问题，你会怎么做？

第一步是需要让组织认识到，专业判断中的噪声是一个值得关注的问题。为此，我们建议进行噪声审查（详细说明见附录1）。在噪声审查中，多个人会对同一个问题做出判断。噪声是这些判断之间的变异。在某些情况下，这种变异可以归结为能力不足：一些判断者业务熟练，另一些人则不然。当存在这样的技能差距时（无论是普遍情况，还是某些特定的情况），人们当然应该优先考虑改进技能上的欠缺。但是，正如我们所看到的，即使在能力出众、训练有素的专业人士的判断中，也存在大量噪声。

如果系统的噪声量值得关注，那么应该考虑用规则或算法代替判断，因为这样可以更好地消除噪声。但是规则有其自身的问题（我们将在第六部分中讲到），即使是人工智能最狂热的支持者也认同：在短时间内，算法不可能普遍取代人类判断。因此，提高判断力的任务紧迫性一如既往，这也是本书这一部分的主题。

当然，改善判断的一个明智做法是尽可能选择能力出众的判断者。在射击场，有些射击手打靶特别准。任何专业判断任务都是如此：精通的人既会产生更少的噪声，也会产生更少的偏差。关于如何找到最好的判断者，答案有时是显而易见的，如果你想解决一个象棋方面的问题，就去问一位象棋大师，而不是问写书的我们。但对于大多数问题，卓越判断者的特征很难辨别，这些特征是第18章的主题。

接下来，我们会讨论减少判断错误的方法。心理偏差和统计偏差都与噪声有关。我们将在第19章中看到，存在许多试图消除心理偏差的尝试，有一些非常成功，有一些则非常失败。我们简要回顾了消除偏差的策略，并提出了一种非常有前景但据我们所知尚未被系统性探讨过的策略。比如，让指定的决策观察者去搜寻一些诊断信号，这些诊断信号能够反映现实生活中的团队工作是否受到一种或几种常见偏差的影响。附录2提供了一个决策观察者可以使用的偏差检查清单。

然后就到了本书第五部分的重点：对抗噪声。我们会介绍决策卫生，即我们推荐的一种减少人类判断噪声的方法。我们呈现了五个不同领域的案例研究。在每一个领域中，我们会验证噪声的普遍性，以及由此造成的一些令人吃惊的故事。我们还将回顾减少噪声的一些成功或失败的尝试。当然，在每一个领域中人们都使用了多种方法，但为了便于阐述，每一章只会突出一种决策卫生策略。

第20章以司法科学的案例作为切入点，说明了排序信息的重要性。寻求一致性会导致人们根据当前的有限证据形成早期印象，然后去证实他们已经形成的预先判断。这就意味着，在判断过程的早期不要接触无关信息，这一点尤为重要。

在第21章，我们会谈到预测的案例，阐明了一个减少噪声的最重要策略的价值，这个策略汇总了多个独立判断。“群体智慧”原则是建立在平均多个独立判断的基础上的，这能确保减少噪声。除了直接进行平均之外，还有其他汇总判断的方法，我们也将结合预测举例说明。

在第22章，我们将审视医学诊断中的噪声以及减少噪声的尝试。它指出了一种减少噪声的策略的重要性和普遍适用性。这种策略我们已经在刑事判决的例子中阐述过，那就是判断指南（**judgment guidelines**）。判断指南可以成为一种强有力的减少噪声的机制，因为它们直接减少了法官之间在最终判决上存在的差异。

在第23章，我们将讨论商业活动中的一个常见挑战：绩效评估。在

这方面减少噪声的尝试证明了，使用一个基于外部视角的通用量表极其重要。这是一个重要的决策卫生策略，原因很简单，判断包含一个将印象转换成量表分数的过程，如果不同的法官使用不同的量表，就会产生噪声。

在第24章，我们探讨了人才选拔这一主题，该主题与绩效评估相关但又完全不同，在过去的100年中，人们已经对此进行了广泛的研究。它说明了一个必不可少的决策卫生策略的价值在于使复杂的判断结构化。结构化是指将一个判断分解为不同的组成部分，对数据收集过程进行管理，以确保输入的信息彼此独立，并且待所有信息都收集完毕后再进行整体讨论，从而做出最终的判断。

在第25章，我们在从人员招聘领域获得的经验基础上，提出了一种对选项进行评估的通用方法：中介评估法。中介评估法以“选项就像候选人”这一假设为前提，以图解的方式描述了如何将结构化的决策制定以及前面提到的其他决策卫生策略，引入重复决策和单一决策的典型决策过程中。

在开始之前，我们要说明一点：具体说明甚至量化每个决策卫生策略在各种情境中的价值非常重要，了解哪种策略最有益以及如何对它们进行比较也很重要。当信息流受到控制时，噪声能降低到什么程度？如果目标在于减少噪声，那么在实践中需要汇总多少判断？结构化判断是有价值的，但在不同的情境中，它的价值到底有多大？

由于噪声的话题很少引起人们的注意，这些问题仍然有待研究解决。在实际应用中，某一策略的优势取决于使用它的特定环境。想一想判断指南的应用：它们有时会产生巨大好处（我们将在一些医学诊断的案例中看到），然而，在其他情况下，采用该指南可能好处并不大。这或许是因为一开始就没有太多噪声，也可能是因为即使是最好的指南也不能减少太多噪声。在任何给定的情境中，决策者都应该力争更准确地了解每项决策卫生策略可能带来的好处及其相应的成本。我们将在第六

部分对其进行讨论。

第18章

卓越的判断者，卓越的判断力

至目前为止，我们主要谈论的是人类的判断，而没有对不同的判断者进行区分。显然，在任何一项任务中，有些人做出的判断要优于其他人。即使是由群体做出的判断，如果[该群体由高能力个体组成](#)，那么最终的判断也会更优。这就引出了一个重要的问题，即如何甄别出更好的判断者。

有三件事至关重要。如果做判断的人受过良好的训练、更睿智且拥有正确的认知风格，那么他的判断也会产生更少的噪声和偏差。换言之，好的判断取决于你的经验、思维能力，以及你的思考方式。好的判断者往往经验丰富且充满智慧，但他们也时刻保持着思维的开放性，愿意接纳新的信息。

专家和尊重型专家

毋庸赘言，判断者的技能影响其判断品质。例如，资深的放射科医生更有可能对肺炎做出准确的诊断；一些“超级预言家”对世界大事的预测，大大超越其他不那么厉害的同行；如果你召集一些在某个法律领域真正专业的律师，他们可能会对常见法律纠纷的判决结果做出大体一致的、准确的预测。专业水平高的人做出的判断，噪声更少，偏差也更少。

这些人是某个领域的真正专家。他们能做出比其他人更好的判断，这一点是可验证的，因为有结果的数据可以证明。至少在原则上，我们可以根据医生、预言家、律师以往判断的准确率来进行选择。出于显而易见的原因，这一方法在现实生活中可能遇到困难。我们并不建议你要

求你的家庭医生进行专业度测试。

我们已经指出，有一些判断的效果是无法验证的。在一定范围内，我们不能轻易得知或毫无争议地确定一些判断的真正价值，保险核保、刑事判决当属此类。此外，品酒、文章评分、书评和影评，以及其他数不胜数的判断也属此类。虽然这些领域的一些专业人士也被称为专家，但我们对这些专家所做判断的信心建立在同行对他的尊重的基础上。我们将这类专家称为“尊重型专家”（respect-expert）。

尊重型专家这一称谓本身并无冒犯之意，事实上，指出这些专家的判断的价值无法被准确评估这一点也并非一种批评，因为在很多领域中，情况本就如此。很多教授、学者和管理顾问也是尊重型专家。他们的可信度取决于学生、同行或客户对他们的尊重程度。在这些领域乃至其他更多领域中，某位专家的判断只能与他们同行的判断进行比较。

在缺少衡量对错的客观标准时，我们常常会看重尊重型专家的意见，虽然这些专家彼此之间也可能意见不一致。不妨想象这样一个画面：一群政治分析家对外交危机的起因以及可能的发展趋势有着截然不同的看法（这种不一致并非个别现象，如果他们全体意见一致才是怪事）。所有政治分析家都相信有一个正确的观点，而他们自己的观点是最接近正确的那一个。如果你仔细听，你会发现，几位分析家的观点都让人印象深刻，他们的论据也同样具有说服力，你现在不知道他们谁对谁错。如果他们的分析没有通过确定、可证实的预测来明确表达，那你以后也可能不会知道。你知道，至少有些分析家是错的，毕竟他们意见不一致，但你仍然尊重他们的专业性。

再来看一组根本不需要做出任何预测的专家。三位受过良好训练的道德哲学家齐聚一堂，其中一位是康德的追随者，第二位支持杰里米·边沁，第三位信奉亚里士多德。关于道德的要求，他们之间分歧很大。他们争论的内容包括撒谎是否合理、何时撒谎是合理的、动物的权益以及惩罚犯罪的目的。你认真聆听，可能会佩服他们的思维清晰、严谨，

也可能倾向于认同他们中的一位，但你对他们所有人都很尊重。

为什么会这样？说得更明确一些，为什么那些因自身判断的高质量而受到他人尊重的人，在没有任何客观证据证明专家专业性的情况下，会选择信任某些专家呢？是什么造就了尊重型专家？

一部分原因是他们遵循共同的准则或专业规范。专家往往需要从专门的组织获得专业认证，并在对应的机构中接受训练，受这些机构监管。完成住院实习的医生和向资深伙伴学习的年轻律师，不仅要学习使用各自行业的技术工具，也要接受特定方法的培训，以及遵守特定的规范。

有了共同的准则，专业人士就能知道在判断时应该考虑哪些因素、如何做出判断，以及如何确保最终的判断更加合理。例如，在保险公司，理赔员很容易就理赔金额达成一致，并阐述充分的理由，因为这些要考虑的因素已经包含在了理赔评估检查清单中。

当然，这种一致性并不能消除理赔员在理赔评估中存在的广泛差异，因为规范并没有对所有细节做出十分明确具体的规定，让理赔员照搬照做。事实上，规则留有解释的空间。专家仍然需要做出判断，而不是执行计算过程，这就是为什么噪声不可避免。即使受过相同训练的专业人士认可共同的准则，他们在实际应用该准则时，彼此之间也存在差异。

除了掌握共同的准则，专家还要具备丰富的经验。如果你的特长是下围棋、演奏钢琴，或是投掷标枪，那么你有可能成为天才少年，因为结果验证了你的水平。然而，核保员、指纹鉴定人员、法官则需要多年的经验才能获得认可，在核保行业，天才少年根本不存在。

尊重型专家的另一个特征是，他们能够满怀信心地做出判断并解释自己的判断。相比于那些自我怀疑的人，我们更加信任那些满怀自信的人。[自信启发式 \(confidence heuristic\)](#) 指出，在群体中，自信者具有更

高的影响力，即使他们的自信毫无理由。尊重型专家善于构建能自圆其说的理论，他们能够凭借经验识别事件的模式，与之前的案例进行类比推理，迅速形成并验证假设。他们很容易将所见的事实整合成一个连贯的故事，这种能力令他们信心十足。

智商高的人判断力更好

训练、经验以及自信是尊重型专家获得信任的决定性因素，但具备这些因素并不能确保他们做出的判断就是高品质的。那么，我们怎样才能知道哪些专家能做出更好的判断呢？

我们有充足的理由相信，智力水平与更好的判断相关。智力与几乎所有领域的良好表现都相关。在所有其他条件等同的情况下，智力不仅与更好的学业成就有关，[也与更好的工作绩效相关](#)。

对于如何测量智力或一般心智能力（General Mental Ability, GMA）^⑨，争议和误解一直存在。[人们对智力的本质也一直存在误解](#)。事实上，这些测试测量的是发展能力，这些能力一部分由遗传决定，一部分受环境（包括教育机会）影响。许多人也担心根据GMA测试做出的筛选会对一些社会群体产生不利影响，并质疑将GMA测试用于筛选是否合理。

我们需要将对这种测试的担忧及其实际的预测品质区分看待。自从一个世纪以前美国陆军采用心智能力测试以来，成千上万的研究考察了人们的认知测试分数与其随后表现之间的关系。通过大量研究，人们得出了非常清晰一致的结论。[就像一篇评论文章所指出的那样](#)：“GMA预测的是个体在其所选职业中能达到的专业水平及其表现。它在这两个方面的预测能力比预测任何其他能力、个性特征、性格特点的能力都要强，并且比根据工作经验进行预测更为准确。”当然，其他认知能力也很重要（本章稍后将详细讨论）。一些个性特征也很重要，其中包括[责](#)

[任心和毅力](#)——在追求长期目标时的意志和热情。当然，GMA未能将多种不同形式的智力纳入测量范畴，例如实践智力与创造力。心理学家和神经科学家还把智力分为晶体智力（crystallized intelligence）和[流体智力](#)（fluid intelligence），前者是指个体根据自身所掌握的关于世界的知识（包括数学运算）解决问题的能力，后者是指个体解决新问题的能力。

标准化的GMA测试测量个体在语言、计量和空间等方面的能力，尽管它还比较粗糙并存在一定的局限，但它仍然是迄今为止重要结果的最佳单个预测指标。正如那篇评论文章所指出的，[GMA的预测效力比心理学研究中的大部分测量方法都好](#)。随着工作内容复杂程度的增加，GMA与工作成就之间的相关性也会增加，这非常合乎逻辑——相比于从事简单工作的人，智力对火箭专家来说显然更为重要。对于高度复杂的工作，[标准化测验分数与工作绩效之间的相关系数可以达到0.5（PC=67%）](#)。我们在前面已经指出，[按照社会科学的标准，相关系数达到0.5代表非常强的预测力](#)。

在讨论高度专业性的判断时，人们经常会提到一个反对智力测量相关性的重要理由：所有做出这些判断的人很可能都拥有高GMA。相比于普通大众，医生、法官以及高级核保员通常都受过更好的教育，因而他们的认知能力测试得分往往会高出普通人很多。据此，你可能有理由相信，高GMA在他们之间并没有产生什么差别，它只是进入高成就群体的“入场券”，而不能解释这个群体中个体成就的差异。

这种观点虽然被普遍接受，但并不正确。毫无疑问，对于某一特定职业，相比于处于顶层的人群，底层人群的GMA分布范围更广：从事社会地位较低的职业的群体中有高GMA的个体，但[在律师、化学家或工程师这些职业中，几乎没有GMA低于平均水平的人](#)。从这个角度来看，高GMA是从事社会地位较高的职业的必要条件。

然而，这个测量指标并不能解释这些群体中的个体在成就上的差

异。[即使是认知能力的评测成绩位于前1%的群体（评测时的年龄为13岁），他们能获得的突出成就也与GMA高度相关](#)。在这前1%的群体中，那些处于前1/4的人获得博士学位、出版著作、获得专利的可能性比那些处于后1/4的人高出2~3倍。换言之，GMA差异的重要性，不仅存在于第99百分位数和第80或第50百分位数之间，甚至在第99.88百分位数和第99.13百分位数之间仍然存在！

[2013年的一项研究重点调查了《财富》500强企业的CEO和424位美国亿万富翁（财富排名前0.0001%的美国人）](#)，结果惊人地证明了能力与成就之间的关联。这项研究发现，这些超级精英群体是由智商最高的一群人组成的，这与预期一致。这项研究还发现，在这一群体内部，更高的受教育水平和能力水平，与更高的薪酬（对CEO而言）和净资产（对亿万富翁而言）相关。虽然偶有个例，比如史蒂夫·乔布斯、比尔·盖茨和马克·扎克伯格等是从著名大学辍学后成为亿万富翁的，但他们是“遮挡了森林的树木”——仅有约1/3的美国成年人获得了大学学位，但美国亿万富翁获得大学学位的比例高达88%。

结论显而易见。在需要做出判断的职业中，GMA会显著影响判断的品质，即使是在一群高能力的个体之中，情况亦是如此。有人认为存在一个界限，达到这个界限之后，GMA就不再起作用，但这种观点没有得到上述证据的支持。这一结论反而强有力地表明，如果专业判断是无法验证的且只能假定它更接近一个看不见的靶心，那么高能力个体的判断更有可能接近目标。如果你必须挑选一些人来做判断，那么你的最优选择是挑选那些高智商的人。

这一系列推理有一个严重的局限。由于你不能对每一个人进行标准化测试，你就需要去猜测哪些人属于高GMA人群。高GMA能显而易见地提高诸多方面的表现，包括有能力让别人相信你说的是对的。具有高心智能力的人，比其他人更能做出好的判断，也更有可能成为真正的专家；他们也同样能给同行留下深刻印象，获得后者的信任，在缺少现实

反馈的情况下成为尊重型专家。这样看来，中世纪的占星家就有可能是那个时代的高GMA人群。

信任那些举止言谈给人睿智之感或是能为他们自己的判断做出令人信服的解释的人，这种策略有一定的合理性，但并不完全理性，甚至可能适得其反。那么，是否有其他方法可以甄别出真正的专家呢？能够做出更好判断的人是否还具有其他明显的特征呢？

认知风格对判断的影响

不管心智能力如何，人们的认知风格或执行判断任务的方法都是不同的。研究人员已经研发了一些表现认知风格的工具。大部分测量方法都与GMA相关（或者彼此相关），但测量的重点各不相同。

其中一种测量是“认知反射测试”（Cognitive Reflection Test, CRT），该测试就是大众熟知的“球和球棒”问题：一个球和一个球棒的价格为1.1美元，球棒比球贵1美元，那么这个球值多少钱？[研究人员建议使用的其他测量问题包括](#)：在一场跑步比赛中，你超过了第二名，你现在是第几名？CRT旨在测量人们能够在多大程度上抑制闯进大脑中的第一个答案，也就是错误的答案：在球和球棒问题中回答“0.1美元”，在跑步比赛问题中回答“第一名”。[低CRT得分与现实生活中的一些判断和信念有关](#)，比如相信鬼魂、占星术、超感知觉（俗称“第六感”）。[CRT得分还可以预测人们是否会因为明显不准确的“假信息”而上当。该测试的得分甚至与人们使用智能手机的程度有关。](#)

许多人将CRT视为测量一种更宽泛的概念的工具，即测量[人们是否会习惯性地运用反射性或冲动性思维过程](#)。简言之，有些人喜欢深入思考，而另一些人在面对同样的问题时往往会相信自己一时兴起做出的判断。用我们的专业术语来说就是，CRT是一种测量人们倾向于使用缓慢的系统2思维还是快速的系统1思维的方法。

其他自我评估的方法也被用来测量这一倾向，当然，所有这些测试都是彼此相关的。例如，[认知需求量表考察](#)人们愿意在多大程度上仔细思考问题。要想在这项测试中得高分，你必须认同“我倾向于设定一些我要付出大量脑力劳动才能实现的目标”而不能认同“思考不是我的乐趣”。[认知需求高的人不太容易出现已知的认知偏差](#)。研究者还提及了一些匪夷所思的联系：“如果你不喜欢‘剧透’，你可能具有更高的认知需求；[那些在认知需求量表上得分低的人，更偏爱‘剧透’](#)。”

这是一个自我评估的量表，并且什么样的答案更容易得到大众认可，是显而易见的，因而会引发相当合理的质疑。那些希望给他人留下良好印象的人不太可能会认同“思考不是我的乐趣”这样的表述。因此，其他测验试图直接测量人们的技能，而不再采用自我描述的方式。

其中一个例子是“[成人决策能力量表](#)”（adult decision making competence scale）。这一量表测量人们在判断过程中犯一些典型错误的倾向性，比如风险知觉中的过度自信或不一致性。另一个例子是“[哈尔彭批判性思维测试](#)”（Halpern critical thinking assessment），这一工具主要评估批判性思维技能，包括理性思考的倾向和一套可习得的技能。在评估中，你需要回答类似下面的问题：想象一下，你的一位朋友不知道应该选择两个减肥项目中的哪一个，于是向你征求意见。其中一个项目称他们的客户平均减掉约11千克；另一个项目称他们的客户平均减掉约13.6千克。在进行选择之前，你认为需要先弄清楚哪些问题？如果你回答说，你想知道“有多少人减掉了这么多重量”“减肥效果能否维持一年以上”等问题，你就会因应用了批判性思维而获得相应的分数。在成人决策能力量表或哈尔彭批判性思维测试中，获得高分的人们在生活中似乎能做出更好的判断：他们较少承担由错误判断导致的不良后果，例如意外怀孕，或因忘记归还租来的影碟而不得不支付滞纳金。

认知风格和能力的测量以及其他测量都能对判断品质进行预测，这种看法似乎是合理的。然而，认知风格和能与判断品质之间的相关性

随任务的不同而产生差异。乌里尔·哈兰（Uriel Haran）、伊拉娜·里托夫（Ilana Ritov）和芭芭拉·梅勒斯在寻找[可以作为人的预测能力指标的认知风格](#)时发现，认知需求并不能预测谁会更努力地去获取信息。他们也并未发现认知需求与更高的绩效之间存在稳定的关系。

唯一能预测人们的预测品质的量表是由心理学教授乔纳森·伯龙（Jonathan Baron）发明的。该量表用于测量人们的“[积极开放性思维](#)”（actively open-minded thinking）。积极开放性思维是指个体愿意积极搜寻与自己先前的假设相矛盾的信息，这些信息包括其他人的不同意见以及与原来看法不一致的新证据。具有积极开放性思维的人会认同“允许自己被相反的意见说服是一种良好的品质”这类陈述，而不认同“改变想法是一种脆弱的表现”或“直觉是决策的最佳指南”这类观点。

换言之，虽然认知反射和认知需求的得分情况反映的是人们进行慢思考或审慎思考的倾向，但积极开放性思维更胜一筹。那些谦逊的人会一直提醒自己判断是一个不断发展的过程，并渴望被纠正。我们在第21章会看到，拥有这种思维模式的人是最好的预测者，他们会根据新的信息不断修正自己的思维和观念。好消息是，一些证据表明，[开放性思维是一种可习得的技能](#)。

这里，我们打算就如何在某些领域挑选出能做出良好判断的人给出一个硬性的结论，但从上面这个简短的描述中可以得出两个通用原则。首先，比较明智的做法是，认识到不同领域专家之间的区别：在一些领域，例如天气预报领域，专家的预测是可以被客观验证的，因此其专业水平可以分出优劣；而另一些领域的专家则是尊重型专家，比如政治分析家可能说得头头是道，令人信服，而象棋大师可能看起来谦卑内敛，也无法解释走出某些棋步的理由，然而，相比于后一种情况，你更需要对前一种专家的判断持怀疑态度。

其次，一些判断者有可能比其他具有同等资历和经验的人做得更好。如果有些人表现得更好，那么他们的判断将会出现更少的偏差或噪

声。在导致这种差异的诸多因素中，智商和认知风格是关键。尽管没有哪一项单独的测试或量表能够准确无误地预测人们的判断品质，但你可以试着去物色这样的人：他们愿意积极搜寻与自己的看法不一致的信息，并把这种信息整合进当前观点，且希望最终能改变自己的想法。

具有卓越判断能力的人，他们的个性可能不符合公众认可的行事果断的领导者形象。人们往往倾向于相信和喜欢这样的领导者：他们自信满满、口齿伶俐，似乎轻而易举或生来就知道什么是对的，而且这类领导者还能激发人们的信心。但证据表明，如果想要减少错误的判断，那么对于领导者或普通人而言，最好的做法是对反对意见持开放态度，并乐于接受“自己可能错了”的想法，如果依然想要坚持己见，那也是在听取了各方意见之后，而不是之前。

- 消除噪声

思维越开放，判断越好

- 你是一位专家，但你的判断可以被验证吗？你是一名尊重型专家吗？
- 如果需要在两种观点中选择一种，而我们对提出这些观点的人的专业性和过往经历一无所知，那么，我们可以选择听从更聪明的人的观点。
- 智商只起到一部分的决定作用，人们如何思考也同样重要。或许你可以选择那些最深思熟虑、思维最开放的人，而不是最聪明的那个。

第19章

消除偏差与决策卫生

许多研究人员和机构都力图减少判断偏差。在本章中，我们将分析[他们的核心发现](#)。我们将介绍消除偏差的各类干预措施之间的区别，并探讨其中一种值得进一步研究的干预措施。随后，我们将讨论如何减少噪声，并介绍决策卫生的概念。

干预：事后及事前消除偏差

减少偏差有两种主要方式，要描述它们的特点，一个好方法是回到测量类比。假设你家浴室里的体重秤在测体重时测量值平均会多出约0.23千克，那么你的体重秤是有偏差的，但这并不意味着它毫无用处。若要消除偏差，有两种可行的方法。第一种方法是，将它显示的数值减掉0.23千克，以此来纠正这种误差。当然，每次都这样做有点麻烦，而且有时你可能完全想不起来要这样做。第二种方法是调整刻度盘，一劳永逸地提高仪器的精度。

这两种消除测量偏差的方法，可以直接与消除判断偏差的两种干预方法进行类比：要么在做出判断之后进行纠正；要么在做出判断之前进行干预。

事后或纠正性地消除偏差往往是靠直觉进行的。试想你正在监管一个项目团队，团队成员估计他们可以在3个月内完工。你可以在团队成员判断的基础上加个缓冲期，将计划工期调整为4个多月或更长（计划谬误），以此来纠正当前估计可能存在的偏差。

此类纠正偏差的方法有时会更系统性地进行。英国财政部发布的

[《绿皮书》](#)（*The Green Book*）是一本介绍如何评估项目和方案的指南。该书鼓励做计划的人通过进行比例的调整来估算项目的花销和持续时间，以此来解决乐观偏差。在理想情况下，这些调整应参照某个组织的乐观偏差的历史水平。如果没有这样的历史数据，《绿皮书》建议对每个类型的项目应用通用的调整比例。

事前或预防性地消除偏差的干预措施又可以分为两大类。在最有前景的方法中，一些方法旨在改变做判断或决策的环境。这种改变也被称为[助推](#)（nudge），众所周知，它们的作用是减少偏差的影响，甚至利用偏差做出一个更好的决策。养老金计划的自动加入制度就是一个简单的例子。推行自动加入制度的意图是克服惰性、拖延和乐观偏差，确保员工为退休储蓄，除非他们主动选择退出该计划。事实证明，自动加入制度极其有效地增加了参与率。该计划有时还搭配“未来多储蓄”（Tomorrow Save More）计划，让员工将一定比例的未来工资增长用于储蓄。这种方式可以在许多地方应用，例如，自动加入绿色能源计划、贫困儿童的学校免费膳食计划，以及其他各种福利计划。

其他助推方式则在选择架构的不同方面发挥作用。它们可能会让正确的决策变得更容易，例如，通过减轻行政管理负担，提高人们获得心理健康护理服务的便利性。此外，它们可以将某一产品或某项活动的某些特征凸显出来，比如让原本的隐形费用更加清楚地显现。对百货商店和网站进行设计，可以轻松地助推人们克服偏差——如果将健康食品放在显眼的地方，则可能会有更多的人购买。

另一类事前消除偏差的方式旨在训练决策者识别并克服这些偏差，其中一些干预措施被称为[助力](#)（boosting）。它们通过让决策者学习统计学知识等方法来提升他们的决策能力。

[教育人们克服偏差是一项崇高的事业，而且很有用](#)，只是比看上去更具挑战性。例如，那些学了多年高级统计课程的人在统计推理中不太可能出错。不过，教人们完全避免偏差则很难。几十年的研究表明，那

些学会了在自己专业领域避免偏差的专业人士，往往难以将所学知识应用到其他领域。例如，天气预报员懂得不要对预报过分自信，当他们宣布有70%的可能性会下雨时，总体来说，遇到这样天气的日子里有70%最终都下了雨。然而，[当被问及常识性问题时，他们可能和其他人一样过分自信](#)。学习克服偏差的难处在于如何认识到：一个新问题与我们在别处所见的问题有相似之处，而我们曾在某处所见的偏差很可能也会在其他地方出现。

研究人员和教育工作者已经成功使用非传统的教学方式来增进这种认识。在一项研究中，波士顿大学的凯里·莫尔韦奇（Carey Morewedge）和他的同事使用了教学视频和“严肃游戏”，教参与者识别由证实性偏差、锚定效应和其他心理偏差引起的错误。每次游戏结束后，参与者都会收到自己所犯错误的反馈，并学习如何避免再次犯同样的错误。结果表明，无论是在完成学习后即刻测试，还是在完成学习8个星期后再进行测试，[这些游戏都使参与者在被问及类似问题时的犯错次数降低](#)。视频教学的效果要相对弱一些。在另一项独立研究中，安妮-劳蕾·塞利尔（Anne-Laure Sellier）和她的同事发现，那些在教学视频游戏中学会克服证实性偏差的MBA学生，可以在另一堂课上[应用学到的知识来解决商业问题](#)，即使没人告诉他们这两个练习之间有联系，他们也做到了。

消除偏差的局限

无论是事后纠正偏差，还是通过助推或助力来事前预防偏差，大多数消除偏差的方法都有一个共同点：它们针对的都是某种被人们假定存在的偏差。这种假设在通常情况下是合理的，但有时是错误的。

思考一下项目规划的例子。你可以合理地假定过分自信会对项目团队产生普遍影响，但你不能确定这是不是唯一的偏差，甚至不能确定它是不是主要的偏差。也许，由于在类似的项目中有过不好的经历，团队

负责人学会了在估计时特别保守，此时团队会表现出与你认为应纠正的误差正好相反的误差。又或许，该团队是通过类比其他相似的项目来做出预测的，并以该项目所花费的时间为锚点。还有可能，项目团队成员估计你会对他们的预估工期加一个缓冲期，便抢先对这个工期进行了调整，使得调整后的工期比他们的真实想法更乐观。

再举个例子，请试想一个投资决策。对投资前景的过分自信可能确实在起作用，但另一种强大的偏差——损失厌恶却会产生相反的效果，使决策者不愿冒可能赔本的风险。或者，试想一家公司要在多个项目之间分配资源，决策者可能既看好新项目的收益（过分自信），又不敢从现有项目中挪出一些资源。这是由现状偏差造成的问题，顾名思义，现状偏差是指人们倾向于保持现状。

上述例子表明，我们很难确切地知道是哪些心理偏差在影响判断。在复杂的情况下，多种不同的心理偏差可能会同时起作用，导致其在同一方向上共同叠加或相互抵消，进而产生不可预测的后果。

最后的结论是：事后或事前消除偏差，即纠正或预防特定的心理偏差，在某些情况下是有用的。当误差的大体方向已知并以明显的统计偏差的方式表现出来时，这些方法可以派上用场。那些注定存在严重偏差的决策类型，可能会因采用消除偏差的干预措施而受益。例如，计划谬误是一种足以消除过分自信的影响的可靠的偏差干预措施。

然而问题是，在许多情况下，误差的可能方向无法预知，比如因所有判断者之间的心理偏差不同且不可预测而出现系统噪声的情况。为了减少这类情况中的误差，我们需要广撒网，从而一次检测到多种心理偏差。

你需要一个决策观察者

我们建议，不要在决策前或决策后进行偏差检视，而要在决策过程

中即时进行。不过，人们很少会意识到自己正在被自己的偏差所误导，这种缺乏觉察本身就是一种已知的偏差——[偏差盲点](#)（bias blind spot）。相比于觉察到自己的偏差，人们通常更容易识别他人的偏差。我们认为，可以训练观察者去实时观测诊断信号，证实一种或几种常见的偏差正在影响他人的决策或建议。

为了阐明这个过程，请你想象一个群体要试着做一个复杂的、会产生重大结果的判断。判断可能是任意类型的，比如应对疫情或其他危机的政府决策、为一个病情复杂的患者寻求最佳治疗方案的会诊、要制定重大策略性行动的公司董事会议等。现在，想象一个决策观察者正对这一群体进行观察，并用一个检查清单来诊断是否存在导致该群体偏离最佳判断的偏差。

决策观察者不好当，并且毫无疑问，在一些组织中，安排决策观察者也是不现实的。如果最高决策者没有下定决心去纠正偏差，那么仅仅发现偏差并没有用。确实，决策者必须是发起决策观察并支持决策观察者的人。我们当然不建议你自己任命自己为决策观察者，这样你既不会赢得朋友，也不能影响他人。

然而，非正式的实验表明，用这种方式可以取得真正的成效。至少，如果使用得当，这种方式是很有用的，特别是当组织或团队的领导者真正愿意为此做出努力并且能够选出好的决策观察者时，因为好的决策观察者不容易受到自身偏差的严重影响。

这些情况中的决策观察者可以分为三类。第一类，在一些组织中，可以由监督员担任决策观察者。监督员不仅需要监管项目团队提案的主要内容，还要特别注意提案产生的过程以及团队的动态，这会使决策观察者对[可能影响提案产生过程的偏差](#)有所警觉。第二类，一些组织可以在每个工作团队中任命一个人担任团队的“偏差破坏者”，这位决策过程的“守门员”应实时提醒团队成员关注可能误导他们的偏差。这种方式的不足之处是，决策观察者被放在了团队中“唱黑脸”的位置上，并且可能

会很快耗尽其“政治资本”。第三类，一些组织可以依赖一位有中立视角优势的外部协调人员，但此做法的缺点是会泄露内部信息，并且需要一定的花销。

决策观察者要想发挥作用，就需要一些训练和工具。其中一个工具就是[一份关于他们试图去探测的偏差的检查清单](#)。使用偏差检查清单的好处很明显：[该清单对于提升高风险环境中的决策有悠久的历史](#)，而且尤其适用于防范过去犯过的错误。

举个例子。美国联邦机构在发布一系列实施成本高昂的规章之前，必须编写一份正式的监管影响分析报告，这些规章包括净化空气或水源、减少工作场所死亡事件、提高食品安全、应对公共卫生危机、降低温室气体排放、加强国土安全等。一份标题不讨人喜欢的技术文件，用近50页密密麻麻的文字来陈述要求，这些要求显然是为了抵消偏差而设计的。联邦机构必须解释为什么需要这些规章，同时考虑更严格和相对不那么严格的替代方案，考虑成本和收益，以无偏差的方式提供信息，并适当地给实施效果打个折扣。但在很多机构中，政府官员没有遵守那本文字密密麻麻的技术文件的要求，他们可能根本就没读过。作为回应，联邦政府的官员制作了一份简单的检查清单，篇幅只有一页半，以期降低机构忽视或没能注意到任何重要规定的风险。

为了说明偏差检查清单长什么样，[我们在附录2中展示了一个](#)。这个通用检查清单只是一个例子，决策观察者一定要制定一个符合其所在组织需求的偏差检查清单，以增强其相关性并[便于应用](#)。重要的是，一份清单并不会详尽地列出所有可能影响决策偏差的因素，它的目的是突出那些最常见以及最重要的偏差。

在决策观察中，使用合适的偏差检查清单有助于限制偏差的影响。尽管我们在一些非正式和小规模的实践中看到了令人振奋的结果，但由于还没看到有人对该方式的效果进行系统性的探究，我们也不知道各种可能的实施方式的利弊。我们希望能够激励实践者和研究人员开展更多

的实验，来考察决策观察者在真实环境中消除偏差的实际效果。

决策卫生，减少噪声的关键方法

偏差是我们经常看到甚至可被解释的错误。它是有方向的，这就是为什么助推可以限制偏差的有害影响，以及为什么助力判断可以对抗特定的偏差。偏差通常也是可见的，这就是为什么观察者可以期待在做出决策的过程中能够对偏差进行实时诊断。

然而，噪声是不可预测的误差，既不容易看到，也不容易解释，这就是噪声会造成严重损害却经常被忽视的原因。因此，减少噪声的策略对消除误差的作用，就好比预防保健措施对医疗的作用，这些策略的目的在于，在各种潜在的错误发生之前对其进行预防。

我们称这种减少噪声的方法为决策卫生。当你洗手的时候，你可能不知道自己到底在预防哪种细菌感染，你只知道洗手是预防各种细菌感染的好方法（不仅在疫情期间如此，平时也应该这样做）。同样，遵循决策卫生的原则意味着：即使你不知道想要规避什么样的错误，你也应该采用减少噪声的策略。

与洗手进行类比是我们有意为之。卫生措施可能很乏味，它们带来的益处并非显而易见，你可能永远不知道你预防了什么问题。而当问题真的出现时，你可能无法追溯到底是哪个卫生环节出了问题。因此，很难强制要求人们洗手，即便对方是对其重要性有着充分认识的医疗行业的专业人士，情况也是如此。

就像洗手和其他形式的预防性措施一样，决策卫生极其有用，但并不讨巧。纠正一个很容易识别的偏差至少可以让你获得一种实实在在的成就感，但是减少噪声的过程不会。从统计学上讲，减少噪声可以避免许多错误，但你可能永远也不会知道到底是避免了哪些错误。噪声是躲在暗处的敌人，即使躲开了敌人的暗箭，你可能也察觉不到。

考虑到噪声会造成的破坏，为了赢得这场无形的胜利，我们还是值得一战。后续章节介绍了在多个领域，包括司法、预测、医学和人力资源等领域中应用的几种决策卫生策略。在第25章，我们会综述这些策略，并展示如何将它们整合到一个减少噪声的方法中。

- 消除噪声

决策卫生，让你赢得隐形的胜利

- 你自己到底在对抗什么样的偏差吗？它会在什么方面影响结果？如果你不知道，那么可能有好几个偏差在同时起作用，而我们很难预测哪一个会占主导地位。
- 在开始讨论这个决策之前，先指定一个决策观察者。
- 如果在这个决策过程中，我们遵守了良好的决策卫生策略，那么很可能这个决策就是最好的。

第20章

司法科学，信息排序是最大的噪声

2004年3月，在西班牙首都马德里，一组被放置在通勤列车上的炸弹爆炸，造成192人死亡，2000多人受伤。人们在犯罪现场的塑料袋上发现了一枚指纹，并通过国际刑警组织将其传送到了世界各地的执法机构。几天后，美国联邦调查局（FBI）犯罪实验室最终确认这枚指纹属于一个居住在俄勒冈州的美国公民布兰登·梅菲尔德（Brandon Mayfield）。

梅菲尔德看起来有很大的犯罪嫌疑。他曾是美国陆军军官，娶了一名埃及女子为妻。后来，他曾作为一名律师，为一些被指控（后来被定罪）试图前往阿富汗加入塔利班的人辩护。因此他上了FBI的监视名单。

有关机构开始监视梅菲尔德，搜查了他的房子，窃听他的电话。在上述审查未能获得任何实质性证据的情况下，FBI逮捕了他。但他从未被正式指控过。梅菲尔德已有10年没出过国。在他被羁押期间，西班牙调查人员通知FBI，他们认为塑料袋上的指纹与梅菲尔德的指纹不符，而与另一名嫌疑人的指纹相符。

两个星期后，梅菲尔德获释。最终，美国政府向他道歉，支付了200万美元的和解金，并下令对发生这种错误的原因进行全面调查。调查的核心结论是：“[错误是人为所致](#)，而不是方法或技术上的问题所致。”

幸好，这种人为错误很少见。尽管如此，它们还是引人深思。美国经验最丰富的指纹专家怎么会误把一枚指纹认定为属于一个从未接近过犯罪现场的人呢？要找到答案，我们首先需要了解指纹鉴定的流程，以

及它与其他专业判断案例的关系。我们通常认为司法科学中的指纹鉴定是一门精确的科学，但实际上它会受鉴定人员心理偏差的影响。这些偏差会导致噪声，这些噪声又会产生大量错误，这样的情况之多远超我们的想象。我们来看看司法科学界是如何通过实施一项适用于所有环境的决策卫生策略来解决这一问题的，这个决策卫生策略就是：严格控制用于做出判断的信息流动。

指纹识别，身份识别中备受重视的技术

指纹是手指的脊线纹理在我们接触的物体表面留下的印记。在古代就已经有把指纹当作识别身份的显著特征的例子，而现代指纹技术则始于19世纪末。当时，苏格兰医生亨利·福尔兹（Henry Faulds）发表了一篇科学论文，首次提出将指纹用于身份识别的技术。

在随后的几十年里，指纹作为犯罪记录中的身份标识越来越受到重视，指纹技术逐渐取代了法国警官阿方斯·贝蒂伦（Alphonse Bertillon）发明的人体测量技术。贝蒂伦于1912年开发了一套对指纹进行比对的正式系统。群体智慧的发现者弗朗西斯·高尔顿爵士（Sir Francis Galton）在英国也开发了类似的体系。不过，这些奠基者并没有受到人们的称赞。高尔顿相信指纹是对个体进行种族划分的有效工具，而贝蒂伦可能是出于反犹太主义的倾向，于1894年至1899年在对阿尔弗雷德·德雷弗斯（Alfred Dreyfus）一案的审判中，做出了起决定作用但有误的专家证词。

警察很快发现，指纹的功用不只是可以作为惯犯的识别标志。1892年，阿根廷警官胡安·武塞蒂奇（Juan Vucetich）首次将隐藏在犯罪现场的指纹与嫌疑人的拇指指纹进行了比对。从那时起，收集隐藏指纹（在犯罪现场获得的指纹）并将其与样本指纹（在受控条件下从已知个体身上收集到的指纹）进行比对的做法，一直是指纹识别最关键的操作，并成为应用最广泛的司法鉴定的证据。

如果你见过电子指纹读取器，就像许多国家的移民局使用的那种，你可能会认为指纹比对是一项简单、机械、容易进行自动化作业的任务。但是，比起匹配两个清晰的指纹，把从犯罪现场收集到的隐藏指纹与样本指纹进行比对要烦琐得多。当你把手指紧紧地按在一个专门用来记录指纹印记的阅读器上时，会产生一个整洁、标准化的图像。相比之下，隐藏指纹往往只有一部分，或是不清楚，或是已经变形。它们提供的信息数量和品质与在精心准备的环境中收集到的指纹是没有可比性的。此外，隐藏指纹常常与其他指纹重叠，这些指纹可能属于同一个人，也可能属于不同的人。隐藏指纹的表面可能会出现污垢和其他瑕疵。确定它们是否与犯罪嫌疑人的指纹样本相符需要专业的判断，这就是人类指纹鉴定人员的工作。

拿到隐藏指纹后，指纹鉴定人员通常遵循一套称为ACE-V的过程，即分析（Analysis）、比较（Comparison）、评估（Evaluation）和核验（Verification）。他们必须先分析隐藏指纹，判断它是否有进行比对的价值，如果有，他们会把这枚指纹和样本指纹进行比对；比对后他们加以评估；评估后他们会得出相符（指纹是同一个人的）、排除（指纹不是同一个人的）或无法确认这三种结论。需要重新鉴定的决策会触发最后一步：由另一位指纹鉴定人员重新进行核验。

几十年来，对这一程序的可靠性一直没有人质疑。尽管目击者的证词被证明有时是非常不可靠的，甚至供词也可能是假的，但指纹一直被认为是最可信的证据——至少在DNA分析出现之前一直是这样。2002年以前，在美国的法庭上没有人质疑过指纹证据的可信度。例如，[当时的FBI网站就坚称](#)：“指纹鉴定是一种绝对不会出错的个人身份识别手段。”在极少数确实出现了错误的情况下，错误会被归结为鉴定人员能力不足或造假。

指纹证据在很长一段时间内都没有受到质疑的部分原因是，很难证明它是错的。一组指纹的“真实值”，也就是实际罪犯的真实身份往往是

未知的。对于梅菲尔德和少数类似案件来说，这个错误尤其严重。但一般来说，即便嫌疑人否认指纹鉴定人员的结论，人们也倾向于认为指纹证据更加可靠。

我们注意到，不知道真实值是普遍现象，但这并不妨碍我们对噪声进行测量。指纹分析中有多少噪声？或者更准确地说，假设指纹鉴定人员与法官或核保员不一样，指纹鉴定人员不是要给出一个数字，而是要做出一个明确的判断，那么他们有多大可能性会意见不一？为什么会这样？这些问题是伦敦大学认知神经科学研究员伊蒂尔·德鲁尔（Itiel Dror）最先着手研究的。他在一个被假定没有噪声问题的领域，进行了一系列类似于噪声审查的研究。

指纹分析中的情境噪声

对于一位认知科学家或心理学家来说，挑战指纹鉴定人员似乎有些奇怪。毕竟，正如你在《犯罪现场调查》（*CSI: Crime Scene Investigation*）系列节目中所看到的，这些都是需要戴乳胶手套、手持显微镜进行研究的硬科学。但德鲁尔意识到，指纹验证显然是一个判断问题。作为认知神经科学家，他断言：哪里有判断，哪里就有噪声。

为了验证这一假设，德鲁尔首先聚焦于情境噪声，即让同一位专家对同一证据进行两次验证，再观察他的前后两次判断之间的差异。[正如德鲁尔所说](#)：“如果专家不可靠到自相矛盾的地步，那么他们的判断和专业性的基础就值得怀疑。”

指纹分析为检查情境噪声提供了一个完美的测试平台，这是因为，不像医生或法官遇到的案件，成对的指纹不容易被记住。当然，必须留出适当的时间间隔，以确保鉴定人员不会记住指纹。在德鲁尔的研究中，一些勇敢、思想开放的专家同意，在未来5年中的任何时候，他们都愿意在不知情的情况下参与研究。此外，实验必须在专家的日常工作

过程中进行，这样他们就不会意识到有人在验证自己的技能。如果在这种情况下，鉴定人员在两次测试中的判断不一致，那就证明指纹鉴定中确实存在情境噪声。

缺乏独立性，司法科学的证实性偏差

以上述专家同意为基础，德鲁尔对此前的研究做了调整，又进行了两项研究，这一次，他引入了一个重要的变化。当第二次看到指纹时，一些检验人员会受到可能使这个案例出现更多偏差的信息的影响。例如，指纹鉴定人员在第一次验证时发现指纹是匹配的，但这次却得知“嫌疑人有不在场证明”或“枪支方面的证据表明他不是嫌疑人”。另外一些鉴定人员最初认定嫌疑人是无辜的或指纹无法确定，但在第二次验证中，他们被告知“侦探相信嫌疑人有罪”“目击者指认了他”“他供认了罪行”。德鲁尔称这项实验是对专家“可偏差性”的测试，因为实验人员所提供的背景信息激活了指纹鉴定人员在特定方向上的心理偏差（证实性偏差）。

事实上，鉴定人员很容易产生偏差。当同一批鉴定人员再次鉴定之前看到的相同指纹时，由于这次有了偏差信息，他们的判断也发生了改变。[在第一项研究中](#)，4/5的专家在面对强有力的背景信息时改变了他们先前做出的比对相符的决策。[在第二项研究中](#)，6位专家重新鉴定了4对指纹。在先前的24个决策中，偏差信息导致了4个决策的改变。可以肯定的是，他们的大多数决策并没有改变，但对于这类决策而言，1/6的改变可以算是很大的比例了。这些研究发现也得到了其他研究人员的证实。

正如我们所料，当决策一开始就很难做出，偏差信息又很强烈，而且对应的改变是从结论确凿的决策变为不确定的决策时，鉴定人员更有可能改变主意。然而，令人不安的是，[指纹鉴定专家往往是根据背景环境做出决策的](#)，而不是根据指纹中包含的实际信息来判断。

偏差信息影响的不仅仅是鉴定人员的结论（相符、无法确认或排除），它实际上还改变了鉴定人员对信息的感知，以及对这种感知的解释。[在另一项独立研究中](#)，德鲁尔和他的同事们发现，那些处于偏差环境中的鉴定人员，与那些没有接触偏差信息的鉴定人员看到的東西不同。当隐藏指纹与目标样本指纹同时出现时，鉴定人员观察到的细节（微小的细节）明显少于他们只看到隐藏指纹时所观察到的细节。[随后的一项独立研究](#)证实了这一结论，并且实验人员补充道：“我们并不清楚它是如何发生的。”

德鲁尔为偏差信息的影响创造了一个术语：司法科学证实性偏差。这一偏差后来为其他司法科学技术所引证，包括血型分析、纵火调查、骨骼遗骸分析和法医病理学。[即使是被普遍视为司法科学新黄金标准的DNA分析，也容易受到证实性偏差的影响](#)，至少在专家们评估复杂的DNA混合物时是如此。

司法科学专家易受证实性偏差的影响不仅是一个理论层面的问题，因为现实中尚没有系统性的预防措施来确保专家们不会接触偏差信息。[鉴定人员经常会在随证据一起提交给他们的传送信函中收到此类信息](#)，也经常会与警察、检察官和其他鉴定人员直接沟通。

证实性偏差还引发了另一个问题。为避免犯错，人们在ACE-V程序中加入了一项重要的安全措施，那就是在确认指纹的信息之前，由另一位专家独立进行一次验证。但大多数情况下，只有结果是“需要重新识别”时才需要独立验证。这很可能会导致证实性偏差，因为[执行核实工作的鉴定人员知道最初的结论](#)是“需要重新识别”。因此，验证的这一步并不像大家预想的那样，能够带来聚合多个独立判断的好处，因为验证实际上并不是独立的。

在梅菲尔德一案中，一系列的证实性偏差似乎起了作用。在这起案件中，不止两位，而是三位FBI专家为错误的身份识别“投下了赞成票”。后来[针对这一错误展开的调查](#)指出，第一位鉴定人员似乎非常相

信计算机系统自动从指纹数据库中搜索出的结果。虽然他明显没有接触到梅菲尔德的个人基本信息，但执行初步搜索的计算机系统提供的结果，加上处理一个极为引人注目的案件带来的心理压力，足以导致初步的证实性偏差。报告继续指出，一旦第一位鉴定人员做出错误的鉴定，随后的鉴定就被污染了。由于第一位鉴定人员是一位备受尊敬的监督员，“机构中的其他人难有异议”。最初的错误被复制和放大，导致人们几乎就认定了梅菲尔德有罪。到了这一步，[哪怕是一位备受尊敬的独立专家](#)接受法院指派代表梅菲尔德的辩护方审查证据，他也会和FBI的意见一致，确认指纹相符。

同样的现象也可能发生在其他司法鉴定领域。隐藏指纹识别被誉为各鉴定领域中最客观的分析手段，如果指纹鉴定人员在操作过程中都存在偏差，那么其他领域的专家也会如此。此外，如果一位枪支专家知道指纹是匹配的，这种信息也会影响他的判断。如果牙齿鉴定专家知道DNA分析已经确认了犯罪嫌疑人，那么这位专家就不太可能认为咬痕与犯罪嫌疑人不匹配。这些例子引发了人们对[一系列偏差](#)的恐惧：正如我们在第8章中描述的群体决策一样，由证实性偏差引发的初始错误变成了影响第二位专家的偏差信息，第二位专家的判断又会使第三位专家产生偏差，依此类推。

在明确了偏差信息会导致判断错误后，德鲁尔和他的同事们发现了更多的关于情境噪声的证据。即使指纹专家没有接触到偏差信息，[他们有时也会对自己先前见过的一组指纹改变看法](#)。在没有给出偏差信息的情况下，改变确实更少发生但无法杜绝，这些与我们预想的情况相符。[2012年，FBI委托进行的一项研究](#)更大规模地重现了这一发现，该研究要求72名鉴定人员再次查看他们在大约7个月前评估过的25对指纹。借助这一高质量鉴定人员的大样本，该研究表明，指纹专家有时容易受到情境噪声的影响，而且大约有1/10的决策改变了。大多数的改变要么从“相符”变成了“无法确认”，要么就是反过来，没有一个决策的改变是源于鉴定错误的。这项研究最令人不安的地方在于，它表明一些原本可

以定罪的指纹先前可能被判定为“无法确认”。当鉴定人员验证相同的指纹时，即使背景环境不是为了使他们产生偏差而是要尽可能让他们保持一致，他们的决策也存在不一致性。

少许噪声，但多少误差呢

这些研究结果反映出现实生活中有可能存在司法错误。我们不能忽视出庭作证专家的可信度问题：可信度是效度的条件，原因很简单，如果你自己的判断都会不一致，那么它也难以与真实情况一致。

究竟有多少错误是由司法鉴定中出现的错误造成的？美国一个致力于推翻错案的非营利性组织“[无辜者计划](#)”（Innocent Project）收集了350个改判无罪的案件。一份针对这些案件的回顾指出：在45%的案件中，司法鉴定科学的误用是促成错判的一个原因。这个统计数字听起来很恐怖，但法官和陪审员关注的重点有所不同：他们要知道包括指纹鉴定人员在内的司法鉴定专家犯相应错误的可能性有多大，才能知道应给予出庭作证的鉴定者们多少信任。

这个问题的最有力答案可以在总统科技顾问委员会（President’s Council of Advisors on Science and Technology, PCAST）的一份报告中找到。PCAST是一个由美国顶尖科学家和工程师组成的咨询小组，它在2016年[对刑事法庭中的司法鉴定进行了全面回顾](#)。其报告总结了指纹分析有效性的现有证据，特别是与错误身份识别（假阳性）的可能性相关的证据，例如关于梅菲尔德的指纹的鉴定。

这方面的证据少得可怜，就像PCAST指出的那样，收集这方面证据的工作直到最近才开展，这实在“令人痛心”。最可信的数据来自2011年FBI的科学家进行的唯一一项[大规模指纹识别准确性研究](#)。这项研究有169名鉴定人员参与，每个人都对比了大约100对隐藏指纹和样本指纹，并发现错误的鉴定很少——假阳性率约为0.17%。

0.17%的错误率的确很低，但报告指出：“鉴于一直以来在媒体宣传中指纹鉴定被赋予的高准确性，[这一比例要比普通公众乃至大部分陪审员认为的高很多](#)。”这项研究未包含任何有偏差的背景信息，而且鉴定人员知道他们在参加测试，这可能导致该项研究低估了实际案例中出现的错误。[在佛罗里达州进行的一项后续研究](#)发现的假阳性率要高得多。这些研究报告中的各种发现表明，我们需要更多的研究来确定指纹鉴定人员决策的准确性以及这些决策是如何做出的。

然而，所有研究得到的一致发现是，鉴定人员犯错的原因都是过于谨慎，这一发现令人欣慰。诚然他们无法做出百分之百准确的判断，但他们知道自己的判断可能会造成什么样的后果，并考虑到了潜在错误的代价。由于指纹识别的可信度非常高，错误的身份识别可能会导致灾难性的后果。其他类型的错误则不会有那么严重的后果。例如，FBI的专家观察发现：[在大多数案件中，“排除”与“无法确认”对案件本身产生的影响是一样的](#)，换句话说，在凶器上发现的指纹证据足以证明嫌疑人有罪，但没有指纹证据并不足以证明嫌疑人无罪。

不仅鉴定人员在做判断时会非常谨慎，证据还表明，专家在做出识别决策之前也会考虑再三。在FBI关于身份识别准确性的研究中，不到1/3的“配对”指纹对（隐藏指纹和样本指纹属于同一个人）被（准确地）判定为身份识别成功。与假阴性（排除）相比，[鉴定人员做出的假阳性判断（错误识别）也要少得多](#)。[他们容易受到偏差的影响](#)，但在这两个方向上的偏差并不等同。正如德鲁尔所指出的，“比起确定的‘相符’的结论，法医专家更倾向于做出‘不确定’的结论”。

鉴定人员接受过训练，因而会将错误的身份识别看成一种不惜一切代价也要避免的致命罪过。值得称道的是，他们遵循了这一原则，我们只能希望他们对身份识别错误的谨慎态度能使像梅菲尔德案这样的身份识别错误的热点案件少之又少。

倾听噪声，减少噪声的第一步

我们观察到在司法科学中存在噪声，这不应视为对法医学家的批评。这仅仅是我们反复观察的结果：哪里有判断，哪里就有噪声，而且比你想象的还要多。像指纹分析这样的任务似乎是十分客观的，以至于许多人根本不会将其视为判断的一种形式。尽管如此，不一致、分歧和偶尔出错在该领域也在所难免。无论指纹识别的错误率有多低，它都不是零，正如PCAST所指出的那样，陪审团应该意识到这一点。

当然，减少噪声的第一步必须是承认它可能存在。指纹识别领域的成员并没有自然而然地承认这一点，他们中的许多人最初对德鲁尔的噪声审查表示非常怀疑。“鉴定人员可能会在不经意间受到案件信息的影响”，这一说法激怒了很多专家。指纹学会（Fingerprint Society）主席就德鲁尔的研究进行了回复：“如果哪位指纹鉴定人员在决策过程中受到了影响……那么他就太不成熟了，[他应该去迪士尼工作](#)。”一家大型法医学实验室的负责人指出，接触那些可能使鉴定人员产生偏差的案件信息，“会令鉴定人员获得满足感，让他们享受工作，[而不会真正改变他们的判断](#)”。[就连FBI在梅菲尔德案的内部调查中都强调](#)，“隐藏指纹的鉴定人员通常会进行核查，他们知道之前的鉴定人员的结果，但这些结果并不会影响他们的结论”。这些言论基本等于否认了证实性偏差的存在。

即使意识到了偏差的风险，鉴定专家也无法避免偏差盲点，即他们倾向于承认他人存在偏差，但认为自己不会。[一项对21个国家400名鉴定专家展开的调查](#)显示，71%的人认同“认知偏差是整个司法鉴定科学中一个令人担忧的因素”，但只有26%的人认为“自己的判断受到了认知偏差的影响”。换句话说，大约一半的司法专业人士认为，他们的同事的判断有噪声，但他们自己的判断并不存在噪声。噪声可能是一个看不见的问题，甚至对那些工作职责就是“发现这种不可见”的人来说亦是如此。

对信息排序，一个好的决策者应该努力“保持怀疑”

多亏了德鲁尔及其同事的坚持，人们的态度正在慢慢改变，越来越多的法医实验室已经开始采取新的措施来减少他们分析中的误差。例如，PCAST在报告中赞扬了FBI实验室重新设计程序以尽量减少证实性偏差的风险这一举措。

必要的方法论步骤是简单明了的。它们阐明了一个适用于许多领域的决策卫生策略：通过对信息进行排序来限制过早地使用直觉。在所有判断中，有些信息是相关的，有些则不是，而且信息并非总是越多越好，在信息有可能诱导鉴定人员过早地根据直觉下判断并导致判断产生偏差时尤其如此。

本着这一精神，为确保鉴定人员判断的独立性，司法实验室采用的新程序只有在鉴定人员需要时，才会向他们提供所需的信息。换言之，实验室会逐步透露信息，尽可能地让他们“蒙在鼓里”。因此，德鲁尔及其同事设计的方法被称为“[线性序列揭露](#)”（linear sequential unmasking）。

德鲁尔的另一项建议也说明了相同的决策卫生策略：鉴定人员应记录他们在每一步做出的判断。他们应该在查看样本指纹之前记录对隐藏指纹的分析，再判断二者是否匹配。这一系列步骤能帮助专家避免只看到他们正在寻找的东西。他们应该在接触到有可能使自己产生偏差的背景信息之前，记录自己对证据的判断。如果他们在接触到背景信息后改变了主意，也应该记录这些改变及做出改变的理由。这项要求限制了先前的直觉使整个过程产生偏差的风险。

依据同样的逻辑，我们可以提出第三条建议，这也是决策卫生的重要组成部分，即当要求另一名鉴定人员核实第一名鉴定人员做出的身份识别时，第二个人不应知道第一个人的判断。

司法鉴定中存在噪声必然值得关注，因为事关生死，但这也说明了

一个问题。长期以来，我们完全没有意识到，指纹识别竟然也可能会出现错误，这说明我们对专家判断的信心有时被夸大了，噪声审查也会暴露出意想不到的噪声数量。简单地改变一下流程就可以改善这种不足，对所有关心提高决策品质的人来说，这一事实令人振奋。

本案例所说明的主要决策卫生策略（对信息进行排序）具有广泛的适用性，可作为预防情境噪声的措施。我们已经注意到，情境噪声是由各种各样的因素触发的，包括情绪，甚至温度。你不能指望控制住所有触发因素，但你可以尝试避免那些会对判断造成明显影响的因素。例如，你已经知道愤怒、恐惧或其他情绪可能会改变判断，于是你也意识到，如果可以的话，在不同时间重复审视自己的判断是一个不错的办法，因为情境噪声的触发因素在不同时间可能不同。

有一种不太显而易见的可能性，那就是你的判断在不知不觉中被另一种触发情境噪声的因素所改变，这种因素就是信息，甚至是准确的信息。就像指纹鉴定人员的例子一样，一旦你知道别人的想法，证实性偏差会导致你过早地形成一个整体印象，并忽略与其矛盾的信息。我们可以用阿尔弗雷德·希区柯克（Alfred Hitchcock）的两部电影的名称来概括这一点：一个好的决策者应该努力“保持怀疑”（*Shadow of a Doubt*，片名又译《辣手摧花》），而不要成为“已经知道太多的人”（*The Man Who Knew Too Much*，片名又译《擒凶记》）。

- 消除噪声

要对抗噪声，就要先承认它的存在

- 哪里有判断，哪里就有噪声——包括指纹识别。
- 我们掌握了案子的更多信息，但在专家们做出判断之前，不要把所知道的一切都告诉他们，以免他们的判断产生偏差。事实上，只告诉他们必要的信息即可。
- 如果给出意见的第二个人知道第一个人的意见是什么，那么第二个人的意见就不是独立的了，而第三个人的意见就更不可能是独立的了。如此一来，一连串的偏差便产生了。
- 若要对抗噪声，必须首先承认它的存在。

第21章

甄选与汇总，超级预测的两大策略

许多判断都涉及预测，比如，下个季度的失业率可能是多少？明年将售出多少辆电动汽车？2050年的气候变化会带来什么影响？盖完一幢新楼需要多长时间？某家公司的年收入是多少？新员工会有什么样的表现？新的空气污染管理制度的成本是多少？谁将赢得选举？这些问题的答案会产生重大影响，因为私人机构和公共机构的一些重要选择往往取决于这些答案。

预测分析员的工作职责是分析预测何时会出错以及为何会出错，他们对偏差和噪声（也被称为不一致性或不可靠性）进行了明确的区分。人们一致认为，在某些情况下，预测者的预测是有偏差的。例如，[官方机构在对预算进行预测时，会表现出不切实际的乐观](#)。平均来说，他们对经济增长的预测高得离谱，而对赤字的预测则低得不切实际。实际上，无论他们这种不切实际的乐观是出于认知偏差还是政治目的，都无关紧要。

此外，[预测者往往过于自信](#)：如果要求将他们的预测用置信区间而不是用单点估计值来描述，他们倾向于选择更窄的区间，而实际上这样做并不合理。例如，[一项正在进行的季度调查](#)要求美国一些公司的首席财务官估计下一年标准普尔500指数的年回报率。首席财务官们要提供两个数字：一个是最小值，即他们认为实际回报率有1/10的可能性低于该值；一个是最大值，即他们认为实际回报率有1/10的可能性高于该值。这两个数值的置信区间是80%。然而，事实上实际回报率落入这个区间的可能性只有36%，也就是说，首席财务官们对自己预测的准确性过于自信了。

预测者也充满噪声。J.斯科特·阿姆斯特朗（J. Scott Armstrong）在

《预测原理》（*Principles of Forecasting*）中指出，即使在专家中，“[不可靠性也是判断预测的误差来源之一](#)”。事实上，噪声是误差的主要来源。情境噪声很普遍：预测者自己的观点也并不总是前后一致。人与人之间的噪声也很普遍：预测者之间意见不一致，即便他们都是专家。如果[让法学教授预测最高法院的裁决](#)，你会发现大量的噪声。如果请相关专家预测实行[空气污染管理制度](#)的年度效益，你会发现巨大差异，例如从30亿美元到90亿美元不等。如果你让一群经济学家对失业率和经济增长情况做出预测，你也会发现他们的预测之间存在很大的差异。我们已经看到了[许多存在噪声的关于预测的例子](#)，而对预测的研究可以揭示更多问题。

改进预测的两种方法

研究也为减少噪声和偏差提供了建议。我们不会在此进行详尽描述，但会重点讨论两种广泛适用的减少噪声的策略。一种是应用我们在第18章提到的原则——选择更好的判断者，从而做出更好的判断；另一种是最普遍适用的决策卫生策略之一——汇总多项独立的评估结果。

对多个预测进行汇总的最简单方法是对它们取平均值。从数学的角度来看，平均值法能够保证减少噪声，具体地说，减少的量就是（1-判断总数平方根的倒数）。也就是说，如果你对100个判断取平均值，那么将减少90%的噪声，如果你对400个判断取平均值，则可以减少95%的噪声——基本上就消除噪声了。这一统计规律促使我们采用在第7章讨论过的群体智慧的方法。

由于平均值法不能减少偏差，对总体误差的影响取决于偏差和噪声的比例。这就是为什么当判断彼此独立时，群体智慧最有效，因为群体智慧中不太可能出现共同的偏差。大量实验证据表明，[对多次预测取平均值会大大提高预测的准确性](#)，例如在股票分析中，经济预测员的“共

识”性预测最准确。就销售预测、天气预测和经济预测而言，[一组预测者的未加权平均值优于大多数个体甚至所有个体的预测](#)。而且，通过不同方法获得的平均预测具有相同的效果：在一个涵盖30项各领域的实证比较分析中，研究人员发现[综合预测平均减少了12.5%的误差](#)。

直接取平均值并不是对预测进行汇总的唯一方法。群体选择策略与直接取平均值的方法一样有效，即[根据近期判断的准确性来选择最好的判断者](#)，然后对少数判断者的判断取平均值。对尊重专家意见的决策者而言，他们更容易理解并采取一种既依赖于汇总又依赖于群体选择的策略。

产生汇总预测的一种方法是利用“预测市场”（prediction markets），在预测市场中，个体就各种可能的结果下注，从而得到激励来做出正确的预测。如果预测市场的价格表明某些事件发生的概率大约是70%，那么它们大约有70%的可能性会发生。从这个意义上说，大多数时候，[预测市场的表现非常好](#)。来自各个行业的许多公司都[利用预测市场来汇总不同的观点](#)。

另一个汇总不同观点的正式程序是[德尔菲法](#)（Delphi Method）。该方法的经典范式是一个包括多轮信息反馈的过程，在此过程中，参与者向组织者提交评估（或投票），并且彼此保持匿名。在新的一轮中，参与者都会给出自己评估的理由，并对其他人给出的理由做出回应，这个过程仍然是匿名的。这一过程鼓励估计值趋同（有时要求新的判断值落在前一轮判断分布的特定范围内，从而迫使估计值趋同）。这种方法能够同时从汇总和社会学习中获益。

德尔菲法在很多情况下都很好用，但是[实施起来有一定的挑战性](#)。有一个更简单的版本可以在单个会议中实施，那就是“[迷你德尔菲法](#)”。这个方法也被称为“评估-讨论-评估法”（estimate-talk-estimate），它要求参与者首先给出独立的（未公开的）评估，然后进行解释，并说明理由，最后根据其他人的评估和解释做出新的评估。共识性判断是第二轮

中获得的个体估计的平均值。

良好判断计划的4个步骤

关于预测质量的一些最具创新性的研究超出了目前为止我们所探讨的内容。这些研究始于2011年，当时3位知名的行为科学家创立了“良好判断计划”（Good Judgment Project）。菲利普·泰特洛克（我们在第11章讨论过他对政治事件长期预测的评估）、他的妻子芭芭拉·梅勒斯以及唐·穆尔（Don Moore）联手提升我们对预测，特别是对为什么有些人擅长预测的理解。

良好判断计划首先招募了数以万计的志愿者，这些志愿者并非专家，而是来自各行各业的普通人。他们被要求回答数百个问题，例如：

- 印度或巴西是否会在未来两年内成为联合国安全理事会常任理事国？
- 在接下来的一年内会有国家脱欧吗？

从上述例子中我们可以看出，该计划主要关注国际大事。重要的是，回答这些问题引发了许多与日常生活更贴近的预测问题。比如一位律师被问到客户是否会胜诉，或是一个电视工作室被问到某个节目提案是否会大受欢迎，这些问题就涉及了预测技巧。泰特洛克和他的同事想知道是否有人尤其擅长预测，预测的能力能否习得或得到提高。

为了理解这些核心发现，我们需要解释一下泰特洛克和他的团队评估预测者时所采用方法的一些关键点。首先，他们使用的预测问题数量很大，而不只是一个或几个，因为那样的话，成功或失败可能全凭运气。如果你预测你最喜欢的队伍将赢得下一场比赛，而且它确实赢了，你也不一定是一个好的预测者。也许你总是预测你最喜欢的队伍会赢，

如果这是你的策略，而他们只赢了一半的比赛，那么你的预测能力就不能算是特别厉害。基于这方面的考虑，为了减少运气的作用，研究人员验证了参与者在大量预测中的平均表现。

其次，研究人员要求参与者对事件发生的概率进行预测，而不是仅仅给出“会发生”或“不会发生”这种非此即彼的答案。对许多人来说，预测就是做是非题——站这方或者站那方。然而，考虑到我们在客观上对未来事件是无法知晓的，预测它们发生的概率才是更明智的选择。如果有人2016年说“希拉里·克林顿有70%的可能性当选总统”，那他不一定是个糟糕的预测者。确切地说，有70%的概率会发生的事情仍然存在30%的不会发生的可能性。要知道预测者的水平高低，我们应该看他们估计的概率是否符合现实。假设一位名叫玛格丽特的预测者说，500个不同的事件发生的可能性为60%，结果其中300个真的发生了，那么我们就可以得出结论：玛格丽特的置信度被校准得很好。良好的校准是良好预测的前提之一。

再次，泰特洛克及其同事又对实验进行了改进，他们不仅要求预测者对一个事件是否会在12个月内发生做出一个概率估计，还让预测者根据新的信息不断修改自己的预测。假设你在2016年曾估计，英国在2019年底前脱欧的可能性只有30%，后来，新的民意调查结果显示，“脱欧”的选票持续增加，这时，你很可能会把你的预测值提高。公投结果公布后，我们仍然无法确定英国是否会在公布的时间内脱欧，但看起来可能性肯定会大得多。（事实上，英国已在2020年正式脱欧。）

泰特洛克和他的同事允许预测者根据新出现的信息更新预测，为了便于记分，他们把每次更新都视为一个新的预测。通过这种方式，良好判断计划的参与者被激励去密切关注新闻并不断更新他们的预测。这种方法反映出人们期望企业和政府预测人员也能够根据新的信息更新预测，尽管后者有可能因为改变主意而遭到批评。对这种批评有个非常常见的回应，有人认为它出自经济学家约翰·梅纳德·凯恩斯（John

Maynard Keynes）：“当事实改变时，我改变了主意。你又能怎样？”

最后，为了给预测者的表现打分，良好判断计划使用了格伦·W.布赖尔（Glenn W. Brier）于1950年研发的系统。该系统被称为“布赖尔分数”（Brier scores），可以测量人们的预测值和实际值之间的差距。

布赖尔分数是一种巧妙的方法，它可以绕过一个与概率预测相关的普遍存在的问题：预测者通过避免采取大胆的立场来对自己的预测做两手准备。再想想玛格丽特的例子，在我们的描述中她是一个校准效果良好的预测者，因为她将500个事件的发生率定为60%，其中300个事件确实发生了。这个结果可能没有看上去那么厉害。如果玛格丽特是一个天气预报员，她总是预测有60%的可能性下雨，而且500天中有300天下雨，那么玛格丽特的预测是很准的，但也是无用的，因为玛格丽特其实是在告诉你，你可能每天都要带把伞以防万一。拿她和尼古拉斯做个比较，尼古拉斯预测有300天下雨的概率是100%，有200天下雨的概率是0。尼古拉斯和玛格丽特有同样完美的校准：这两位预报员都预测有X%的日子会下雨，且实际也正是如此，但显然尼古拉斯的预测更有价值：他没有为自己的预测做两手准备，而是明确地告诉你是否应该带伞。从技术上讲，尼古拉斯的判断除了校准效果好外，还有很高的辨析度。

布赖尔分数会对准确的校准和准确的辨析度打出高分。为了得高分，你不仅要在平均水平上是正确的（即校准效果良好），而且要能够表明立场，区分不同的预测（即具有高分辨率）。布赖尔分数以均方误差的逻辑为基础，分数越低越好——0分就是完美。

我们已经了解了评分机制，那么良好判断计划的志愿者表现如何呢？其中一个主要的发现是，绝大多数志愿者的表现都很差，但是有2%的人表现突出。前面说过，泰特洛克称这些表现良好的人为超级预测者。他们几乎从不犯错，他们的预测明显高于随机水平。值得注意的是，一位政府官员表示，这些人的表现明显“[好于能够阅读情报和其他秘密数据的情报界分析师的平均水平](#)”。这个类比值得我们反思，情报

界的分析专家受过训练，能够做出准确的预测，而且，他们还可以了解机密信息，然而，他们的表现却比不上[超级预测者](#)。

永久测试版：一个特殊的思维循环

超级预测者为何如此优秀？与我们在第18章中的论述一致，我们可以合理地推测他们异常聪明。这种推测并没有错。在GMA测试中，超级预测者在良好判断计划中的表现比普通志愿者更好，而普通志愿者的成绩已经明显高于平均水平。但差别并不总是那么大，许多在智力测试中表现非常好的志愿者并没有成为超级预测者。除了常规智力，我们可以合理地预期超级预测者在数学方面的能力异常出色。他们的确如此。但他们真正的优势不是数学天赋，而是能够轻松自如地应用分析思维和概率思维。

想一想超级预测者对问题进行组织和分解的意愿和能力。他们不会对一个国家是否会退出欧盟、一场战争是否会在某地爆发这样的问题形成一个整体判断，而是将其分解为几个组成部分。他们会问：“什么情况下答案才是肯定的？什么情况下答案是否定的？”他们会问并试图回答一系列辅助问题，而不是给出一种直觉或整体的预感。

超级预测者也擅长从外部视角看问题，他们非常关心基准概率。正如第13章中对迈克尔·甘巴迪问题的阐述，你在关注甘巴迪的个人资料的细节之前，了解一下普通CEO在两年内被解雇或辞职的概率是很有帮助的。超级预测者会系统性地寻找基准概率，当被问及两个国家次年是否会因边境争端而发生武装冲突时，超级预测者们并不只关注或立即去关注这两个国家目前是否相处融洽。根据他们读到的新闻和分析，他们可能对此有一定的直观感受，但他们知道，对某件事情的直觉往往不可靠。相反，他们一开始会去寻找一个基准概率：他们会询问过去的边界争端升级为武装冲突的频率。如果这样的冲突很少发生，超级预测者将首先考虑这一事实，再去了解两国局势的详细信息。

简而言之，超级预测者的与众不同之处不在于他们智力过人，而在于他们明白如何运用智慧。他们运用智慧的技能反映了我们在第18章中描述的那种可能产生更好判断的认知风格，尤其是高水平的“积极开放性思维”。回想一下关于积极开放性思维的测试：它包括“人们应该考虑与他们的看法相悖的证据”和“关注与你意见不同的人比关注那些与你意见一致的人更有用”。显然，在这项测试中得分很高的人在新的信息出现时会大大方方地更新自己的判断，而不会反应过度。

为了描述超级预测者的思维方式，泰特洛克使用了“永久测试版”（perpetual beta）的说法。这是一个程序员常会使用的术语，指的是一个不打算在最终版本中发布，却被无休止地使用、分析和改进的程序。泰特洛克发现：跻身超级预测者行列的最有力的预测因素是“永久测试版”，即人们致力于更新看法和提高自我完善的程度。正如泰特洛克所说：“超级预测者之所以如此优秀，不在于他们是谁，而在于他们做了什么——艰苦的研究工作，仔细的思考和自我批判，对其他观点的收集和汇总，细微的判断和不懈的更新。”他们喜欢一个特殊的思维循环：[尝试，失败，分析](#)，调整，再试一次。

预测中的噪声和偏差

此时，你可能会想，人们可以被训练成超级预测者，或者至少变得更像一位超级预测者。事实上，泰特洛克及其合作者正在为此而努力。他们做的工作可以被视作了解“超级预测者为何表现如此出色，以及如何让他们更出色”的第二个阶段。

在一项重要的研究中，泰特洛克和他的团队将普通预测者随机分为3组，来测试不同干预措施对其后续判断品质的影响。这些干预措施恰好对应我们描述的改善判断的3种策略。

- 培训：一些预测者完成了一个概率推理的课程，以期提升他们

的预测能力。在这个课程中，预测者们了解了各种偏差（包括忽视基准概率、过分自信和证实性偏差）、对不同来源的多个预测取平均值的重要性，并将参照类别纳入考量。

- 团队合作（一种汇总判断的形式）：让一些预测者组成团队，这样一来，他们就可以看到彼此的预测，并能够进行讨论。团队合作可以通过鼓励预测者应对相反的论点、保持积极开放的思维来提高准确性。
- 甄选：对所有预测者的准确度进行评分，在年终时，排名前2%的人被指定为超级预测者，并有机会在接下来的一年里与精英团队合作。

事实证明，这3种干预措施都非常有效，从某种意义上说，它们提高了人们的布赖尔分数，而且我们发现训练是有效的，团队合作更有效，甄选的效果最优。

这一重要发现证实了汇总判断和选择优秀判断者的价值，但这并不是全部。有了关于每次干预效果的数据，泰特洛克和梅勒斯的合作者维莱·萨托帕（Ville Satopää）开发了[一种复杂的统计技术](#)，以梳理每一种干预措施究竟是如何改善预测的。他推断，从原则上讲，一些预测者的表现比其他人更好或更差的主要原因有3个：

- 他们可以更熟练地发现和分析环境中与自己要做出的预测相关的数据，这也说明了信息的重要性。
- 一些预测者可能总倾向于错误地高估或低估预测的真实值。如果在成百上千的预测中，你在整体上高估或低估了发生某种变化的可能性，那么可以说，你受到了某种形式的偏差的影响，要么是倾向于支持改变的偏差，要么是倾向于保持稳定的偏差。

- 一些预测者可能不太容易受到噪声或随机误差的影响。正如在任何判断中一样，当我们进行预测时，噪声可能有许多触发因素；预测者可能对某条新闻反应过度（这是我们称之为模式噪声的一个例子），可能会受到情境噪声的影响，还可能在使用概率量尺时产生噪声。所有这些误差（以及很多其他误差）的大小和方向都是不可预测的。

萨托帕、泰特洛克、梅勒斯及其同事马拉特·萨利霍夫（Marat Salikhov）称他们的模型为BIN预测模型，BIN是Bias（偏差）、Information（信息）、Noise（噪声）这三个单词首字母的合称。他们测量了这三种因素在多大程度上提升了三种干预措施效果。

他们的答案很简单：这三种干预措施主要是通过减少噪声起作用的。正如研究人员所说：“[干预措施提高准确性](#)的主要手段是抑制判断中的随机误差的出现，然而，训练干预的初衷是减少偏差。”

鉴于训练的目的在于减少偏差，一个不太出色的预测者会预测，减少偏差是培训的主要效果。然而，培训是通过减少噪声来起作用的。这种奇特的现象很容易解释。泰特洛克的训练的主要目的是对抗心理偏差。我们现在已经知道，心理偏差并不总是会造成统计偏差。当心理偏差以不同的方式影响不同个体的不同判断时，就会产生噪声。很明显，这里的情况就是这样的，因为所预测的事件是多种多样的。由于主题不同，同样的偏差会导致预测者反应过度或反应不足。我们不要认为他们会产生统计偏差，即预测者相信事件会发生或不会发生的普遍倾向。因此，培训预测者克服他们的心理偏差是有效的——通过减少噪声的方式。

团队合作对减少噪声有相当大的作用，同时也显著提高了团队提取信息的能力。这一结果与汇总的逻辑一致：几个人协同工作比一个人更善于发现重要信息。如果爱丽丝和布莱恩一起工作，爱丽丝发现了布莱恩错过的信号，那么他们进行联合预测会更好。在团队中工作时，超级

预测者似乎能够避免群体极化和信息级联的危险。相反，他们会汇总不同个体的数据和想法，并以积极开放的方式充分利用综合的信息。萨托帕和他的同事解释了这一优势：“[与培训方式不同的是，通过团队合作.....预测者可以利用这些信息。](#)”

甄选的总体效果最好，一些改进源于更好地利用了信息。超级预测者比其他人更善于发现相关信息，这可能是因为他们比一般参与者更聪明、更有动力、对做出此类预测更有经验。但甄选的主要作用还是减少噪声。超级预测者比普通人，甚至比受过训练的团队产生的噪声更少。这一发现也让萨托帕和其他研究人员大吃一惊：“[“超级预测者”的成功主要归功于他们在控制测量误差方面的出色能力](#)，而不是其他人无法复制的对新闻的透彻解读。

甄选与汇总的有效之处

超级预测项目的成功突出了两种决策卫生策略的价值：甄选（超级预测者都超级棒）和汇总（预测者组成团队进行合作时表现更好）。这两种策略可以被广泛地应用在许多判断过程中。只要有可能，你应该通过组建业务能力出众且能力互补的判断者团队（由预测者、投资专家、招聘人员组成）来实现战略性合作。

到目前为止，我们已经考虑了通过对多个独立判断取平均值的方式来提高精确度，就像在群体智慧实验中一样。对高准确性判断者的评估进行汇总将进一步提高判断的准确性。[通过汇总既独立又互补的判断，我们可以获得准确度上的进一步提高](#)。试想有4个人是一场犯罪行为的目击证人（确保他们不会相互影响至关重要），他们从4个不同的角度目击了犯罪行为，那么将他们提供的信息进行汇总后其质量会好很多。

组建一个专业团队来共同完成判断任务，类似于组建一套综合测试来预测候选人未来在学校或工作中的表现。该任务的标准工具是多元回

归（见第9章），它通过依次选择变量来执行。首先选择的是最能预测结果的第一个测试，然而，下一个测试不一定是第二有效的。相反，第二个测试提供了有效且与第一个测试无关的信息，为第一个测试增加了最大的预测力。例如，假设你有两个心理素质测验，其与未来的表现的相关系数分别为0.5和0.45，还有一个人格测验，其与未来表现的相关系数只有0.3，但与心理素质测验不相关。最佳的解决方案是首先选择更有效的那个心理素质测验，然后选择人格测验，因为这样会带来更多的新信息。

同样的道理，如果你要组建一个判断者团队，那么你当然应该先挑选最好的判断者。接下来，选择一个判断能力中等却能给团队带来一些新技能的人，而不是与第一个人高度相似且更善于判断的人，这样做可能效果更好。以这种方式组成的团队会更优秀，因为当不同个体的判断彼此不相关时，汇总判断的有效性会比判断存在冗余时提高得更快。在这样的团队中，模式噪声相对较高，因为个人对于每个案例的判断都会有所不同。矛盾的是，有噪声的群体的平均判断水平比高度一致的群体的平均判断水平更高。

我们需要注意的一个重点是：无论多样性如何，只有在判断真正彼此独立的情况下，对判断进行汇总才能减少噪声。正如我们在讨论群体中的噪声时所强调的，群体审议过程中增加的偏差往往比消除的噪声更多。一个组织想要利用多样性的力量，那它必须愉快地接受团队成员独立做出判断时产生的分歧。激发和汇总既独立又多样的判断往往是最简单、最便宜且适用范围最广的决策卫生策略。

- 消除噪声

像超级预测者一样思考

- 取4个独立判断的平均值，保证可以将噪声减少一半。
- 我们应该像超级预测者一样，努力使自身保持在“永久测试版”状态。
- 在讨论某个具体情况之前，我们需要先了解一下相关的基准概率是多少。
- 我们有一个很好的团队，但如何确保意见的多元性？

第22章

医疗决策，用科学的诊断指南减少噪声

几年前，我们的一位好朋友保罗被他的主治医师琼斯诊断为患有高血压。琼斯医生建议保罗尝试药物治疗，他给保罗开了利尿剂，但保罗的血压依旧很高，也就是说药物没起到效果。几个星期后，琼斯又开了第二种药物——钙通道阻滞剂，但效果依然不理想。

这些结果使琼斯困惑不已。虽然过去的3个月保罗每星期都会来复诊，但他的血压只是略有下降，仍然很高。对于下一步的治疗方案，琼斯医生感到很苦恼，保罗也很焦虑，毕竟他年纪不算大，身体也一向健康。随后，琼斯医生寻思着尝试第三种药物。

就在这时，保罗搬到了一座新城市，在那里他咨询了另一位名叫史密斯的医生。保罗向史密斯讲述了他这些日子与高血压抗争的经历。史密斯立即回答说：“买一台家用血压计，在家测测看血压值是多少。我认为你根本没有高血压。你可能只是得了‘白大褂综合征’——一进医生办公室，血压就升高！”

保罗照办了，果然，他在家时测得的血压是正常的，而且从那以后一直很正常。在史密斯医生向保罗说了“白大褂综合征”这件事的一个月后，即便是在医生的办公室，保罗的血压也没有再高过。

医生的主要任务就是做出诊断——判断病人是否患病了，如果患病了，还要鉴定出是哪种疾病。在许多情况下，诊断是常规化的，甚至是机械化的，并且人们还制定了规则和程序以减少噪声。医生通常很容易判断某个人的肩膀有没有脱臼或脚趾有没有骨折，而在一些对技术性要求更高的问题上，医生也能比较容易地做出判断。比如，[量化肌腱退化的程度时，医生的诊断产生的噪声就很小](#)。[当病理学家评估乳腺病灶的](#)

[穿刺活检结果时](#)，他们的评估是比较明确的，几乎没有噪声。

然而，有些诊断根本不涉及判断。医疗水平的进步往往是通过消除判断也就是将判断转变为计算来实现的。对于链球菌性咽喉炎的诊断，医生首先要对患者的咽拭子标本进行快速抗原检测，这种检测可以很快检测出患者是否出现链球菌感染。可以说，[即便有这种快速抗原检测结果](#)，链球菌性咽喉炎的诊断也存在噪声，但如果没有，那情况将更糟。如果你的空腹血糖水平为126mg/dL或更高，或是糖化血红蛋白（过去3个月内血糖的平均值）高于6.5HbA1c，你[就会被诊断为患有糖尿病](#)。在新型冠状病毒疫情出现的早期阶段，一些医生做初步诊断时是根据症状做出判断的，随着疫情的发展，检验变得越来越普遍，有了检验结果，判断就没有必要了。

当医生做判断时，他们的判断可能会有噪声，而且可能会出错，标准做法是建议患者再咨询另外一位医生，获取第二诊疗意见。[在有些医院里，第二诊疗意见是必须给出的](#)。当第二诊疗意见与首诊意见出现分歧时，噪声就出现了——当然，我们可能并不清楚哪位医生是正确的。一些患者（包括前文案例中的保罗）会惊讶地发现第二诊疗意见与首诊意见的分歧非常大。医疗领域存在噪声并不令人惊讶，真正令人惊讶的是其体量之庞大。

在本章中，我们要阐明医疗领域的噪声如何大到了令人惊讶的程度，并介绍医学界使用的一些减少噪声的方法。我们将着重讨论一个决策卫生策略：诊断指南。我们敏锐地意识到，与医学中的噪声以及医生、护士和医院采取的各种补救措施有关的内容中有非常多可写的，轻而易举便能写成一整本书。但值得注意的是，医学中的噪声并不局限于本章所关注的诊断性判断中的噪声，治疗过程也可能存在噪声，关于这个话题也有大量的文献可供参考。比如，对于心脏存在问题的患者，医生们对最佳治疗方案的判断存在惊人的差异，涉及的问题包括该用哪些药物、正确的手术方式是哪个，以及到底需不需要手术等。20多年来，

[达特茅斯·阿特拉斯项目](#)（The Dartmouth Atlas Project）一直致力于记录“美国在医疗资源分配和使用上存在的显著差异”，当然，[许多国家也存在医疗资源分配不均的情况](#)。对于本书来说，我们只需要简单地探讨诊断性判断中存在的噪声就可以了。

医疗诊断中的噪声

有大量文献介绍了医学界存在噪声，其中许多是实证性的且验证了噪声的存在，也有许多文献是指导性的。一些医学领域的论文一直在提出各种各样的减少噪声的策略，这些论文如同一座座金矿，蕴藏着众多极其珍贵的想法，值得许多领域的专家、学者参考借鉴。

若噪声存在，就证明其中一位医生明显是正确的，另一位明显是错误的并且可能受到某些偏差的影响，正如我们料想的那样，医生的技能在其中发挥了关键作用。例如，[一项针对放射科医生做出肺炎诊断的研究发现](#)，噪声中很大一部分源自医生技能上的差异。具体而言，“技能的差异可以解释诊断决策中44%的变异”，这表明“提升技能比使用统一的决策指南更有效”。训练和甄选对于减少误差、消除噪声和偏差至关重要，[在医疗领域也是如此](#)。

在医学的某些领域，[如放射学和病理学](#)，医生对噪声的存在一清二楚。例如，放射科医生将诊断差异称为“阿喀琉斯之踵”。目前尚不清楚到底为什么放射学和病理学领域的噪声如此受关注，可能是因为这些领域的噪声确实比其他领域多，也可能是因为这些领域的噪声更容易被记录下来，我们猜想易于记录可能是更重要的原因。在放射学领域，人们更容易实施一些简单的噪声测试以及误差测试，例如，你可以再看一遍影像片子来对之前的评估结果进行重新评估。

在医学领域，人与人之间的噪声，或者说评估者信度（interrater reliability），通常用[kappa统计量](#)来衡量。kappa值越高，噪声越小。

kappa值为1表示评估者们的评估完全一致；kappa值为0，则意味着评估者之间的一致性很差，就跟一群猴子往诊断列表上投飞镖差不多。根据该系数测量得到的信度，医学诊断的某些领域属于“轻度”或者“糟糕”，“糟糕”意味着该领域的噪声非常多。在大多数情况下，该系数测出的信度是“一般”，这当然有所改善，但依然表明其中存在着明显的噪声。随机选出100例药物，要求全科医生判断这些[药物之间的相互作用](#)是否具有临床意义，对于这一重要问题，医生们做出的判断“一致性很差”。在许多外行眼里，诊断肾脏疾病处于哪一阶段似乎比较简单，甚至很多医生也这样认为。但是，[肾病专家们在基于肾病患者的标准化检测结果进行诊断时](#)，仅勉强达到“中等程度的一致性”。

一项研究发现，关于[乳腺病变是否为癌变](#)这一问题，病理学家之间只能达成“一般”程度的一致性。[在诊断乳腺增生病变时](#)，一致性同样是“一般”。当医生根据核磁共振成像扫描结果[判断椎管的狭窄程度](#)时，一致性也是“一般”。这些发现很值得反思，因为我们说过，在医学的某些领域中噪声水平非常低，但在一些对技术性要求非常高的领域，医生们的诊断还远没有实现零噪声。关于是否会被确诊为患有癌症等严重疾病，患者有时就像在抽签，其结果取决于问诊的是哪位医生。

再来看看这些文献中的其他发现，它们源自那些噪声的严重程度尤其值得关注的领域。我们描述这些发现并不是为了对医务工作的现状做出权威性的论断（医务工作在不断地发展和改进，在某些情况下，发展和改进的速度很快），我们的目的是让大家认识到噪声是普遍存在的，不管是在不远的过去还是现在，噪声都屡见不鲜。

1. [在美国，心脏病是男性和女性的主要致死原因](#)。冠状动脉造影是检测心脏病的主要方法，可以评估急性和非急性情况下心脏动脉的阻塞程度。在非急性情况下，当患者反复出现胸痛，且有一条或几条动脉的阻塞程度超过70%时，医生通常会对患者实施支架置入术等治疗措施。然而，有资料显示，医生对血管

造影图的解释存在一定的变异，这可能引发不必要的手术。一项早期研究发现，评估血管造影图的医生对大血管阻塞是否超过70%持不同意见，[这种情况发生的可能性为31%](#)。心脏病专家普遍意识到，在解读血管造影图时存在潜在的变异，尽管他们不断努力并采取纠正措施，但这个问题仍未解决。

2. 子宫内膜异位症是本应该生长在子宫内的子宫内膜组织生长在子宫外所致，这种异位会令病人痛苦不堪，并且可能引发生育障碍。医生一般借助腹腔镜对这种疾病进行诊断，在诊断过程中，医生需要通过手术将腹腔镜的微型摄像头探入患者体内。研究人员让108名妇科医生观看了3名患者的腹腔镜检查视频并做出判断，其中2名患者患有不同程度的子宫内膜异位症，1名患者未患此病。然而，[这些医生在判断子宫内膜异位病灶的数量和位置时，产生了很大分歧](#)，并且这些判断在数量和位置上的相关性都很弱。
3. 肺结核是全球范围内传播最广、致死率最高的疾病之一。仅2016年，就有1000多万人感染肺结核，死亡人数接近200万。胸部X线检测是一种广泛使用的肺结核检测方法，它可以检测肺部是否出现了由结核杆菌引起的空洞。医生们在[肺结核诊断上存在的变异性](#)已有近75年的详尽记录了。尽管这种现象在过去几十年中已有所改善，但研究发现，肺结核诊断中依然存在显著的变异性，诊断一致性仅为“中等”或“一般”。此外，[不同国家的放射科医生在肺结核的诊断上也存在差异](#)。
4. 当病理学家分析黑色素瘤（最危险的皮肤癌）导致的皮肤病变时，只有“中等”程度的一致性。[研究人员要求8位病理学家对每个病例进行诊断](#)，所有病理学家全部达成一致或只有一位意见不一致的情况只占62%。另一项在肿瘤中心开展的研究发现，黑色素瘤的诊断准确率只有64%，这意味着每3个被诊断

为黑色素瘤的病灶中就有1个属于误诊。[还有一项研究发现](#)，纽约大学的皮肤科医生有36%的可能性无法根据皮肤活检样本诊断出黑色素瘤。[这项研究的作者总结道](#)：“医生如果未能在临床上正确诊断出黑色素瘤，就会放任这种潜在致命疾病对患者的生命构成严重威胁。”

5. 放射科医生通过乳房X线检测来诊断乳腺癌，这种诊断同样存在变异性。一项大型研究发现，放射科医生诊断结果的假阴性率为0%（每次的诊断都是正确的）~50%（甚至超过50%，也就是说在超过一半的情况中，放射科医生错误地将乳腺癌患者的X线片诊断为正常），而[假阳性率为1%~64%](#)（这意味着放射科医生将未患乳腺癌的人的X线片诊断为患有乳腺癌的可能性接近2/3）。同时，不同放射科医生诊断的假阴性率和假阳性率也不同，这证明确有噪声存在。

这些存在于不同人之间的噪声是当下的主要研究方向，但也有一些发现是关于情境噪声的。[有时候，放射科医生两次评估同一张影像片子时会给出不同的意见](#)，也就是说第二次给出的意见与第一次的意见不一致。但相比于与他人意见不一致的情况，与自己原有的意见不一致的情况较少出现。在重新评估[血管造影显示的血管阻塞程度](#)时，22名医生与自己原有意见不一致的可能性为63%~92%。我们可以看到，[在那些标准模糊和判断情境复杂的领域中](#)，评估者的信度可能非常差。

这些研究对这种情境噪声的成因没有明确的解释，而[另一项不涉及诊断的研究](#)发现了医学领域情境噪声的一个简单来源，这一发现值得患者和医生牢记。简而言之，医生在清晨比接近傍晚时更可能要求患者做癌症筛查。在一个大样本中，早上8点，乳腺和结肠筛查的预约率最高，为63.7%，到了上午11点，预约率下降至48.7%，中午时又上升至56.2%，下午5点则降至47.8%。因此，在一天中的晚些时候，医生建议就诊的患者做指南推荐的癌症筛查的可能性较小。

我们如何解释这些发现？有可能是因为医生们在遇到患有复杂疾病的患者后，他们的问诊进度会不可避免地落后，这些病情复杂的患者需要的诊断时间往往超过常规的20分钟。我们在前面提到过压力和疲劳等情境噪声的触发因素（见第7章），这样看来这些因素的确在起作用。一些医生为了赶上进度，跳过了对预防性健康措施的讨论。[还有一个例子也能说明临床医生会受到疲劳的影响](#)，那就是规范洗手的比率在轮班结束时会更低一点，可以说洗手这一过程也充满了噪声。

诊断指南的价值

在不同医学问题的情境中，对噪声的存在及其数量进行全面记录，不仅[对医学领域意义重大](#)，而且对人类的认知也有举足轻重的贡献。虽然我们不知道目前是否有此类记录，但希望时机成熟后有人能够完成这一壮举。不过，即使现在没有这类记录，现有的研究结果也提供了一些线索。

在极端情况下，一些问题和疾病的诊断基本上是机械性的工作，不需要做出任何判断。还有一些情况，诊断虽然不是机械性的工作，却很简单直接：任何受过医学训练的人都可能得出同样的结论。在某些场景中，噪声的数量也能得到控制，比如说肺癌专家在进行诊断时，他们的专业性虽不能完全消除噪声，但能使噪声数量最少。有些疾病的诊断给了医生很大的判断空间，并且诊断的相关标准过于宽泛，这会导致噪声数量巨大，且难以降低。我们将看到，大部分精神病学诊断就属于这种情况。

怎样才能降低医学领域的噪声呢？[我们在前面提到，培训可以提高医生的技能](#)，而高超的技能当然对减少噪声有帮助。[汇总多个专家的判断（如第二诊疗意见等）也能减少噪声](#)。算法是一个特别有发展前景的诊断手段，医生们现在正在使用深度学习算法和人工智能来减少噪声。例如，这类算法已经被用于检测乳腺癌患者的淋巴转移情况。目前[最好](#)

[的算法](#)的诊断准确性优于最好的病理学家。当然，算法是没有噪声的。[深度学习算法](#)在检测与糖尿病相关的眼部问题方面，已经取得了巨大的成功。人工智能在通过乳房X线检测诊断癌症方面的表现，[几乎和放射科医生一样出色](#)。如果人工智能进一步发展，它的表现可能会比人类更好。

未来医学界可能会越来越依赖算法。这个过程可以同时减少偏差和噪声，挽救生命并节省成本。但本章的重点是人类的判断指南，因为医学领域的情况能够充分说明，指南在某些实践中能够确保产生良好甚至极好的结果，而在其他应用实践中会产生更复杂的结果。

1952年由产科麻醉师弗吉尼亚·阿普加（Virginia Apgar）发明的阿普加评分（Apgar score），也许是诊断指南最知名的例子。过去，评估新生儿是否处于危险之中曾是医生和助产士需要做出的临床判断。阿普加评分给了他们一个标准化的指南，供他们评测婴儿的肤色、心率、反射动作、肌肉张力和呼吸状况。这些指标的首字母刚好组成Apgar：外貌（Appearance，指肤色）、脉搏（pulse，指心率）、痛苦反应（grimace，指反射动作）、活动（activity，指肌肉张力）和呼吸（respiration，指呼吸频率和力度）。在阿普加测试中，这5个指标分别对应3个分数：0、1和2。总分最高为10分。7分及以上就可被视为健康状况合格（见表22-1）。

[表22-1 阿普加评分指南](#)

类别	分数
外貌 (肤色)	0: 全身发紫或苍白 1: 身体红润, 但四肢发紫 2: 全身粉红或正常肤色
脉搏 (心率)	0: 没有心跳 1: 心跳小于 100 次 / 分钟 2: 心跳大于 100 次 / 分钟
痛苦反应 (反射动作)	0: 对呼吸道刺激无反应 1: 受到刺激时有痛苦的表情 2: 受到刺激时有痛苦的表情并伴有咳嗽或打喷嚏的反应
活动 (肌肉张力)	0: 四肢无力 1: 手臂和腿略有弯曲 2: 主动动作
呼吸 (呼吸频率和力度)	0: 无呼吸 1: 哭声微弱 (呜咽或咕噜) 2: 哭声响亮有力

要注意的是, 心率是评分中唯一一个可以直接评分的项目, 其他项目都需要先做一些判断。但是, 由于判断被分解为多个相互独立的元素, 每个元素都很容易评估, 即使是那些仅受过少量训练的医师, 在评估时也不太可能产生很大的分歧, 因此[阿普加评分产生的噪声很小](#)。

阿普加评分展示了指南是如何发挥作用的, 以及为什么指南能减少噪声。与规则或算法不同, 有了指南仍需判断——决策不是直接进行计算。医师在阿普加评分的每一个项目上都可能存在分歧, 因此在最终结论上也可能存在分歧。然而, 指南成功地减少了噪声, 因为它在预先定义好的维度上将一个复杂的决策分解成了许多简单的子判断。

如果用第9章讨论的简单预测模型来看待这个问题，我们就很容易看到指南的优势所在。临床医生是根据几条预测性线索来判断新生儿的健康状况的。情境噪声可能在其中发挥作用：就在某一天，或者就在某种情绪状态下，临床医生可能会关注一些相对不那么重要的预测因素，或忽略一些重要的预测因素。阿普加评分能够将医生的注意力集中在5个已经过实践验证的重要维度上。然后，评分标准清晰地描述了该如何评估每条线索，这大大简化了根据每条线索做出判断的过程，从而降低了噪声。最后，阿普加评分还规定了如何给各个预测因素赋权，并产生所需的总体判断，这个过程就是一项机械性的工作，而人类临床医生在给这些线索赋权时会存在差异。关注重要的预测因子，简化预测模型，以及进行机械性汇总——所有这些都降低了噪声。

类似的方法已被用于许多医学领域，比如用森特评分（Centor score）来指导链球菌性咽喉炎的诊断。病人如果出现以下症状或体征即得1分：咳嗽（Cough）、有渗出物（exudates，指喉咙后部有白色斑块）、颈部淋巴结肿痛（nodes）、体温高于38摄氏度（temperature）。这些症状的首字母组合刚好是Centor的前4个字母。这一指南由森特及其同事总结得出。根据病人的得分，医生可以决定是否推荐用咽拭子诊断链球菌性咽喉炎。[使用该量表进行评估和评分相对直接](#)，这样一来，有些人就不必接受链球菌性咽喉炎的测试和治疗了。

同样，乳腺癌诊断领域也制定了一套根据乳腺影像报告和数据系统（Breast Imaging Reporting and Data System, BI-RADS）进行诊断的指南，从而减少了对乳房X线片解读时出现的噪声。[一项研究发现，BI-RADS提升了乳房X线片的评估者之间的一致性](#)，这就表明，对于存在明显变异性的领域，指南可以有效地减少噪声。[在病理学领域](#)，关于人们使用指南来减少噪声，也有很多成功的例子。

伤脑筋的精神病学诊断

就噪声而言，精神病学是一个极端的例子。当使用相同的诊断标准诊断同一位患者时，精神科医生经常意见不一致。因此，[至少从20世纪40年代起，减少噪声就成为精神病学界的头等大事](#)。我们将看到，尽管指南在不断完善，但在减少噪声方面，诊断指南的作用非常有限。

[1964年，一项针对91名患者和10名有经验的精神科医生的研究](#)发现，两名医生意见达成一致的可能性只有57%。在另一项早期的研究中，两名州立医院的精神科医生单独对426名患者进行诊断，结果显示，他们在诊断精神疾病的类型时，诊断的一致性只有50%。另一项涉及153名门诊病人的早期研究发现，医生们在诊断方面的一致性只有54%。这些研究并没有明确噪声的来源。然而有趣的是，一些精神科医生倾向于将患者划分到特定的诊断类别中。例如，有些精神科医生更有可能诊断患者患有抑郁症，而另外一些则更有可能诊断患者患有焦虑症。

我们很快将看到，精神病学中的噪声水平仍然很高。为什么会这样？专家们给不出一个唯一且明确的答案，这意味着他们对噪声的解释本身就充满噪声。诊断类别过于宽泛无疑是其中一个因素。但是，[为了初步揭示其中的原因](#)，研究人员让一位精神科医生先面诊一位患者，在短暂休息后，再让另外一位精神科医生进行一次面诊。两位精神科医生随后会面，如果发现彼此意见不一致，再讨论具体的原因。

“医生之间的不一致”有很常见的原因：他们来自不同的学派，接受过不同的训练，拥有不同的临床经验，采用不同的面诊方式。“接受过发展心理学训练的临床医生，可能会将幻觉解释为受虐待所致的创伤经历的一部分”，但“[具有生物医学背景的另外一位临床医生](#)，可能会将幻觉解释为精神分裂过程的一部分”。这种差异就是模式噪声的例子。

然而，除了医生之间的不同，噪声的主要成因是“术语不够准确”。这种情况再加上专业人士对精神病学术语的普遍不满，促成了《精神疾病诊断准则手册（第3版）》（*Diagnostic and Statistical Manual of*

Mental Disorders, DSM III）于1980年改版完成。该手册首次提出了明确而详细的精神疾病诊断标准，朝着引入诊断指南迈出了第一步。

[DSM III促使大量的研究关注诊断是否存在噪声](#)。*DSM III*也被证明有助于减少噪声。[但这本手册远没有达到完美](#)。[即使在2000年对第4版——DSM IV（初版于1994年出版）进行了重大修订之后](#)，仍然有研究表明其中存在大量噪声。一方面，艾哈迈德·阿博拉亚（Ahmed Aboraya）及其同事得出结论称：“精神疾病诊断标准的使用已被证明能[提高精神疾病诊断的可靠性](#)。”另一方面，仍然[存在着一个严重的风险](#)，即对同一个病人要施以多种诊断，才能决定他是否应该入院治疗。

[该手册的第5版——DSM V](#)于2013年发布。[美国精神病学学会](#)曾希望*DSM V*能够减少噪声，因为其标准更客观、更清晰。但[精神科医生的诊断仍然表现出明显的噪声](#)。例如，塞缪尔·利布利奇（Samuel Lieblich）及其同事发现“[精神科医生就患者是否患有重度抑郁症很难达成一致](#)”。[DSM V的现场试验发现](#)其“一致性仅达到最低水平”，这“意味着，在实验条件下，训练有素的专业精神科医生对患者是否患有抑郁症达成一致的可能性只有4%~15%”。[另外一些现场试验表明](#)，*DSM V*实际上使情况变得更糟了，在所有主要领域中，一些诊断，比如混合性焦虑与抑郁障碍的噪声反而增加了。“*DSM V*如此不可靠，以至于它在临床实践中显得毫无用处。”

在精神病学中，[使用指南之所以很难取得成功](#)，主要原因似乎是“某些疾病的诊断标准仍然模糊，难以操作”。一些指南通过将判断分解成不同的标准来减少分歧，进而减少噪声，但如果这些标准是相对开放的，噪声仍然可能存在。考虑到这一点，我们呼吁制定更标准的诊断指南，具体的改进包括：（1）[明确诊断标准，舍弃模糊标准](#)；（2）给出症状及其严重程度的“参考定义”，其依据的理论是“当临床医生对症状是否存在能够达成一致时，他们更有可能在诊断上也达成一致”；

（3）除开放式谈话外，对患者进行结构化访谈，建议在访谈指南中纳

入24个筛选问题，如焦虑、抑郁和饮食失调等问题，以便医生做出更可靠的诊断。

这些改进建议听起来不错，但能够成功地减少多少噪声，其结果尚未可知。[一位观察者曾说](#)：“诊断过程依赖患者的主观症状、临床医生对症状的解读，而且缺乏客观的测量手段（如血液测试），这些因素都在精神疾病诊断中埋下了不可靠的种子。”从这个层面上说，要想减少精神病学中的噪声，困难重重。

在这一特定问题上，现在做出自信的预测还为时过早，但有一点很清楚，一般来说，医学诊断指南在减少偏差和噪声方面非常成功。它们帮助了医生、护士和病人，极大地改善了公共卫生。因此，[医学界需要更多的指南](#)。

- 消除噪声

消除医疗诊断中的噪声

- 医生之间的噪声水平远超我们想象。在诊断癌症和心脏病，甚至是看X线片时，专家们也会出现意见不一致的情况。这意味着病人得到的治疗可能是碰运气的结果。
- 医生往往会认为，无论是星期一还是星期五，是清晨还是傍晚，他们都会做出同样的决策。但事实证明，医生的言行可能完全取决于他们的疲劳程度。
- 医疗诊断指南可以使医生少犯错，从而维护病人的利益。这些指南对于整个医疗行业也有帮助，因为它们减少了变异性。

第23章

绩效评估，用基于外部视角的共识框架做出量化判断

让我们从一个练习开始：请选出3个你认识的人，他们可以是你的朋友或同事。请在友善、智慧和勤奋这3种特质上给他们进行1~5分的评分，其中1分是最低分，5分是最高分。现在再请另一个很了解他们的人（这个人可以是你的配偶、好友或者最亲密的同事）用同样的方式对这3个人进行评分。

在某些评估中，你和其他评分者很可能会给出不同的评分，那么出现这种差异的原因是什么呢？你可能会发现，答案在于你们是如何使用量表的，我们称这种量表为水平噪声。也许你认为，需要“特别出众”才能给5分，而另一位评分者则可能认为只需达到“还不错”的水平就能给5分。除此之外，你们之间的差异还有可能是由你们对被评估者的看法不同造成的，例如你们在理解被评估者是否友善以及如何确切地定义美德等问题上存在差异。

现在想象一下，如果要基于你们的评估来确定这3个人能否升职和加薪，并且假设你和另一位评估者是在同一家重视友善、智慧和勤奋的公司工作，你们之间的评分会有差异吗？这种差异会和前面的练习中出现的差异一样大吗？还是更大？不管你的答案是什么，公司的政策和量表上的差异都可能产生噪声，事实上，这种现象在不同组织进行绩效评估时普遍存在。

绩效评估本质上是一种判断工作

几乎所有大型组织都会定期对员工进行正式的绩效评估，而那些被

评估的人并不喜欢这种评估。[正如有家报纸的标题所示](#)，“研究发现，基本上每个人都讨厌绩效评估”。我们认为，每个人都知道绩效评估会受偏差和噪声的影响，但是大多数人并不知道其中到底存在多少噪声。

在一个理想化的世界里，评估绩效并不是一个判断工作，因为评估人借助客观事实就足以裁定员工工作表现的好坏。但问题在于，大多数现代组织与经济学鼻祖亚当·斯密（Adam Smith）提到的扣针厂没有什么共同之处。在亚当·斯密提到的扣针厂中，每个工人的产出都可以量化。但对一家公司的首席财务官或研发部主管来说，他们的产出是什么呢？当代知识工作者需要平衡多种工作目标，有时这些目标甚至是互相矛盾的。只关注其中一个目标，在做评估时就可能犯错，导致评估结果不准确，并带来负面的激励效果。例如，医生每天接诊的患者数量是医院生产力的一个重要衡量指标，但你并不希望医生一门心思只关注这一指标，更不会只根据接诊人数对他们进行评估和奖励。也就是说，即使是可量化的绩效指标，也需要放在一定的背景条件下进行评估，例如，评估销售人员的销售额或程序员编写的代码行数时，就要考虑具体情况，因为并非所有客户都同样“难伺候”，也并非所有软件开发项目都一模一样。鉴于这些挑战，我们无法完全根据客观的绩效指标对人们进行评估。因此，[以判断为基础的绩效评估无处不在](#)。

25%是信号，75%是噪声

有关绩效评估实践的研究文章已经发表了成千上万篇，[其中大多数人都发现这些评估充满了噪声](#)。这一发人深省的结论主要是通过对360度绩效评估的研究得出的。这类研究需要多个评估者共同参与，并且基于多个绩效维度对同一个人进行评估。在进行数据分析时我们会发现，这类研究的结果并不理想：真正的差异——个人绩效产生的差异通常不超过总差异的20%~30%，[其余70%~80%的差异是系统噪声](#)。

这些噪声是从哪里来的？得益于[多项关于工作绩效评估变异性的研](#)

[究](#)，我们现在已经知道，系统噪声的所有成分都在起作用。

在绩效评估的背景下，这些噪声成分是显而易见的。请想象两位评估者：琳恩和玛丽。如果琳恩的评估标准比较宽松，而玛丽的评估标准比较严格，换言之，琳恩对所有人的评分都比玛丽的更高，那么就出现了水平噪声。就像我们在讨论法官判决的案例时所看到的那样，这种噪声可能表示琳恩和玛丽真的对被评估者形成了不同的印象，也可能表示她们对被评估者的印象相同，但使用评估量表的方式存在差异。

现在，如果琳恩要对你进行评估，而且碰巧她对你和你的工作印象极差，她一向宽松的标准可能会与她对你的特殊（负面）的评价相抵消。这就是我们所说的稳定模式——一个特定评估者对特定被评估者的反应，这种模式是琳恩所独有的，这就是模式噪声的来源。

最后要讲的是情境噪声。玛丽可能在填写评估表之前发现有人把她停在停车场的车撞坏了，这让她在做评估时心情极差；琳恩可能刚刚收到了丰厚的奖金，这让她在做评估时心情极好。类似这样的事情当然可能会产生情境噪声。

关于如何将系统噪声分解为水平噪声、模式噪声和情境噪声这三种成分，不同的研究会给出不同的结论。我们肯定可以想明白，为什么不同组织的噪声成分会有所不同。当然所有形式的噪声都是我们不希望存在的。这些研究所传达的基本信息只有一个：大多数绩效评估结果与被评估者的真实绩效之间的关系远非我们想象中那么紧密。正如一篇综述文章中总结的那样：[员工真实的工作绩效和对工作绩效进行的评估之间的关系](#)可能很弱，或者说，两者之间的关系还不确定。

此外，组织中的绩效评估之所以无法反映[员工的真实绩效](#)，原因有很多。例如，评估者实际上可能不会尽可能准确地评估绩效，但可能会[“策略性地”对员工进行评估](#)。一些其他动机可能使评估者故意给员工打高分，例如为了避免煎熬的反馈性谈话⁽¹⁰⁾，[为了帮助一个一直在寻求晋升机会的人](#)，甚至只是为了摆脱一个表现不佳的团队成員——该成员

需要有良好的评价才可能被调到其他部门。

这些策略性的考虑当然会影响评估，但它们并不是唯一的噪声来源。我们之所以知道这些，全要归功于一种自然实验：仅以研发为目的的360度反馈系统。在这些系统中，受访者被告知他们的反馈并非用于绩效评估，如果评估者真的相信了这种说法，那么这就能防止他们夸大或降低评价。事实证明，研究性评估确实会对反馈产生影响，但系统噪声仍然很高，而且相比于被评估者的实际绩效系统，噪声能够解释更多的变异性。可以看出，即使是纯粹的研究性的反馈机制，评估中也仍然充满噪声。

一个早就发现却没有解决的问题

如果绩效评估系统如此不可靠，那么实施绩效评估的人就应该已经注意到并力图改进这个体系。事实上，在过去的几十年里，众多组织的确对这些评估制度进行了无数次的改革。这些改革也采用了一些我们提出的减少噪声的策略，但在我们看来，人们做得还远远不够。

几乎所有组织都使用汇总策略来减少噪声。人们通常将汇总评估与360度反馈系统相结合，这在20世纪90年代成为一些大公司的标准做法，《人力资源管理》（*Human Resources Management*）杂志在1993年推出了一期关于360度反馈系统的专刊。

虽然对所有评估者的评估进行平均有助于减少系统噪声，但值得注意的是，360度反馈系统并非是为了解决这个问题而开发的，它的主要目的是测量出一些领导看不到的东西。这个评估系统要求被评估者的同事和下属而不仅仅是其领导参与对他们的绩效评估，这样一来，评估的本质就已经改变了。从理论上讲，这种转变是一种改进，因为这使得人们的工作并不是为了取悦领导。360度反馈的盛行与非固定的、基于项目的组织形式是契合的。

一些证据表明，360度反馈系统是一个有用的工具，因为它可以对[客观、可量化的绩效做出预测](#)。然而，这种反馈系统自身也会制造一些问题。计算机的普及使得在反馈系统中添加更多问题变得轻而易举，组织的多重目标以及越来越多的限制也增加了岗位描述的维度，许多反馈问卷变得异常复杂，[过度设计的调查问卷](#)比比皆是。例如，有一份调查问卷要求每个评估者对被评估者在11个维度上进行46项评估。通常来说，人们是无法回忆和处理多个被评估者在多维度上准确且相关的表现的，因此，从某种角度来看，这种过于复杂的方法不仅无用，而且有害。我们知道，光环效应意味着看上去相互独立的维度实际上不会被独立对待，对最初几个问题给予一个强烈肯定或否定的评估后，后续问题的答案就会相应地偏向同一个方向。

更重要的是，360度反馈系统的发展成倍地增加了参与反馈的时间成本。中层管理者被要求填写针对各层级同事的几十份调查问卷，这种情况在很多组织中屡见不鲜，有时他们还需要填写针对其他组织中的员工的调查问卷，因为现在许多公司都要求客户、供应商和其他业务合作伙伴提供反馈。即便这样做的出发点是好的，但是当时间有限的评估者们遭受一系列问卷的“轰炸”时，我们就别指望他们能够提供高品质的信息。在这种情况下，减少噪声要付出的代价可能会非常大，这是我们将在第六部分讨论的问题。

最后，360度反馈系统也存在所有绩效评估体系都无法避免的一种通病——悄无声息地出现评估通胀。比如，一家大型工业公司曾观察到，公司[98%的管理者](#)被评为“完全符合预期”。当几乎所有人都获得了最高评价时，这些评估的真实价值就值得怀疑了。

相对判断的优点

从理论上讲，解决评估通胀问题的一个有效方法是在评估过程中引入一些规范，其中一个普遍的做法是[强制排名](#)（forced ranking）。在强

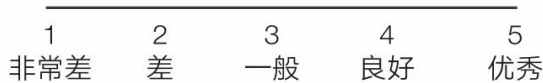
制排名体系中，评估者不仅不能给所有人最高的评分，而且评分还必须遵循预先设定的评分分布。杰克·韦尔奇（Jack Welch）在担任通用电气 CEO 时曾倡导强制排名这种做法，以此来防止评估通胀，以及确保绩效考核的“公正”。许多公司都采用了这个方法，但由于它会对员工士气和团队合作产生不良影响，人们渐渐就不再使用这种方法了。

无论排名有什么缺点，排名的噪声都比评分的要小。我们在惩罚性损害赔偿的例子中看到，相对判断中的噪声比绝对判断中的噪声要少得多，因此[这种关系也被证明适用于绩效评估](#)。

想要了解其中的原因，请看图23-1，它展示了两种评估员工表现的量表。A组对员工进行绝对评分，我们称之为匹配性操作——找到与员工“工作表现”印象最接近的分数。相比之下，B组要求在特定维度（如安全性）上将每个人与同组的其他人进行对比，要求主管使用百分位数来评定员工在特定人群中的排名。从图中我们可以看到，主管已在百分位数量表上找到了3名员工的位置。

A 组

员工 A 的工作表现：



B 组

请对你的下属在安全性这个维度上做出评估。安全性是指员工遵守相关规章制度、在工作中以安全的方式行事并展现出对安全工作实践的认识和理解的程度。

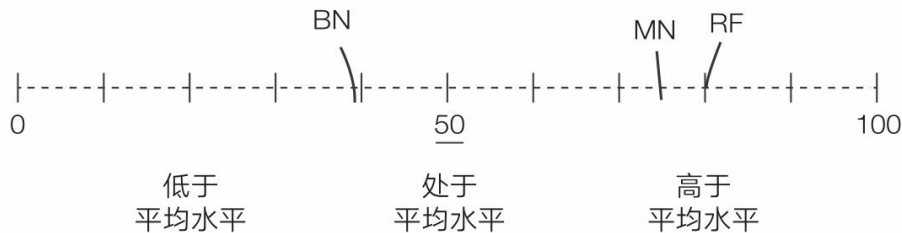


图23-1 绝对量表和相对量表的示例

B组的方法有两个优点。第一，每次只在一个维度上对所有员工进行评估（在这个例子中评估的是安全性），这是一种减少噪声的策略，即将复杂判断结构化地分解成多个维度的单一判断，我们将在下一章详细讨论这一策略。结构化是一种限制光环效应的手段，它通常可以将一个人在不同维度上的评估限制在一个小范围内。当然，只有在像本例这样对每个维度分别进行排名时，结构化才能起作用。如果在“工作表现”这种定义不清晰的总体判断上对员工进行排名，则不会减少光环效应。

第二，正如我们在第15章中讨论的那样，排序同时减少了模式噪声和水平噪声。当你将两个团队成员的绩效相互比较而不是单独给每个人打分时，出现评估不一致（产生模式噪声）的情况会少一些。更重要的是，排名自动消除了水平噪声。如果琳恩和玛丽对同一个包含20名员工的团队进行评估，尽管琳恩比玛丽的评估标准更宽松，她们给出的平均评分会有所不同，但她们给出的平均排名不会存在差异——一个标准宽松的评估者和一个标准严苛的评估者采用的是相同的排名方式。

事实上，强制排名的主要目的在于减少噪声，它确保了所有评估者都有相同的平均值和相同的评分分布。如果规定了评分的分布，排名就是强制的。例如，某一条规则可能会规定“被评为最优的人数不可以超过被评估者总人数的20%，垫底的人数也不可以少于总人数的15%”。

排名，但不强制

从原则上说，强制排名理应给评估带来巨大的改善，结果却往往适得其反。这里我们不打算就其可能带来的所有不良后果展开讨论，这往往与执行不力有关，而非原则本身的问题，但我们还是可以从强制排名体系的两方面问题中吸取一些具有普遍性的教训。

首先是绝对绩效和相对绩效之间的混淆。任何公司都不可能有98%的经理在绩效表现上跻身群体的前20%、前50%甚至前80%。但如果预期的绩效是事先定义好的绝对绩效，那么经理们都能“达到预期”也并非不可能。

许多高管并不认同几乎所有员工都能达到预期标准这一点，他们认为，如果是这样的话，肯定是期望标准设置得太低了，要不就是因为自鸣得意的组织文化。诚然，这种解释可能是正确的，但大多数员工也可能确实达到了高预期标准，而且在高绩效组织中，这种情况完全有可能出现。例如，当你听说所有成功执行太空任务的宇航员都完全达到了预期标准，你并不会嘲笑国家航天部门的绩效管理程序太过宽松。

重点是，只有当组织关心相对绩效时，依赖于相对评估的体系才是合适的。例如，无论人们的绝对绩效如何，都只有一定比例的人可以晋升，此时，相对评估才可能有意义——参考军队的晋升机制。但是，如果像很多公司那样，评估的目的只是衡量员工的绝对绩效水平，此时强制使用相对排名便不合逻辑了。强制规定将一定比例的员工评为不符合（绝对）预期，这不仅残忍而且荒谬——如果说在一支精锐部队中，有10%的人必须被评为“不满意”，那么这种做法就太愚蠢了。

其次，有些组织认为强制的评估分布反映了潜在的真实绩效分布（通常指接近正态分布），然而，即使群体中的绩效分布是已知的，同样的分布也可能不会出现在一个较小的群体中，如由一名评估者评估的小型员工群体。如果从几千人中随机挑选10人，你不能保证其中一定有两个人属于总体的前20%。“不能保证”是保守的说法，实际上，这种情况发生的概率只有30%。而在实践中，这个问题会更加严重，因为团队的构成不是随机的。一些团队成员可能全是高绩效人员，而另一些团队则可能完全由低绩效人员组成。

在这种情况下，强制排名的做法不可避免地会产生错误和不公正。例如，假设一个评估者所评估的团队由5个绩效表现相同的人组成，他

们的实际绩效表现相差无几，而其评分却要符合强制性、差别化的评分分布，这样的做法并不能减少错误，反而会增加错误。

批评强制排名方法的人经常把攻击的重点放在排名原则上，他们谴责这一原则的残忍、不人道，并认为采用这种方法会适得其反。无论你是否接受下述辩驳，强制排名的致命缺陷不是“排名”，而是“强制”。不管是因为将相对量表用于衡量绝对绩效，还是因为做判断者被迫要将无法区分的绩效区分开来，只要判断中强制使用了不合适的量表，不管你如何选择量表，都会自动增加噪声。

确保有一个达成共识的参考框架

对于各种组织为改进绩效评估的测量方式所做的努力，说其结果“令人失望”都算是保守的，这些努力的结果是：绩效评估的成本直线上升。2015年，世界四大会计师事务所之一[德勤（Deloitte）](#)经过计算发现，公司每年要花费200万工时用于评估6.5万名员工的绩效。绩效评估可以说是组织中最可怕的仪式之一，那些不得不进行评估的人和不得接受评估的人都痛恨绩效评估。[一项研究发现](#)，90%的管理者、员工和人力资源主管都认为，他们的绩效管理流程无法实现预期的效果，相关的研究证据也证实了大多数管理者的这种感受。这一发现令人震惊。虽然绩效反馈与员工发展计划相结合时可以带来一些改进，但绩效评估对员工工作积极性的打击同它带来的激励一样大。[正如一篇评论文章总结的那样](#)：几十年来，无论人们尝试了什么方法来改进绩效管理流程，绩效评估仍然会产生不准确的信息，并且对提升绩效几乎起不到任何作用。

由于对绩效管理流程感到绝望，越来越多的公司正在考虑做出一种极端的选择：彻底取消评估体系。这场“[绩效管理革命](#)”的支持者包括许多技术公司、一些专业服务机构和少数传统行业的公司，他们的目标是关注发展性的、面向未来的反馈，而不是回顾性的评估。有些公司甚至

采用非数字型的评估指标，这意味着他们放弃了传统的绩效评估方式。

对于绝大多数没有放弃绩效评估的公司来说，它们可以做些什么来对其进行改善呢？它们可以采用的减少噪声的策略就是选择合适的量表，以确保不同的判断依据的是相同的参照系。研究表明，改进评估方式并对评估者进行培训有助于提高评估者使用量表的一致性。

需要指出的是，绩效评估量表必须使用具体的描述，只有这些描述足够具体，人们对它们的理解才能够保持一致。许多组织都会使用“行为锚定评估量表”（behaviorally anchored rating scales），该量表上的每个评定等级都对应着具体的行为描述，图23-2的左侧给出了一个示例。

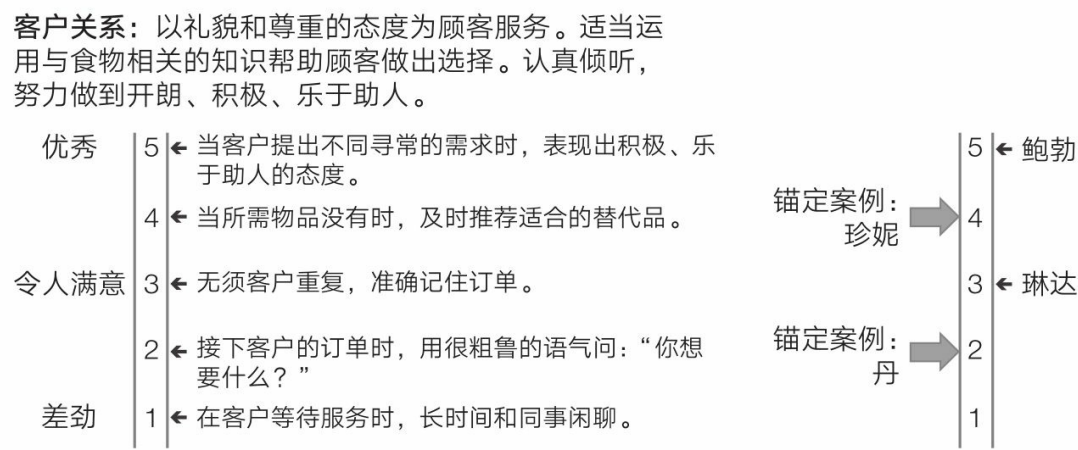


图23-2 行为锚定评估量表（左）和案例量表（右）的例子

然而，[有证据表明，行为锚定评估量表不足以消除噪声](#)。不过，进一步的参照框架培训（frame-of-reference training）则被证明有助于确保评估者之间的一致性。在这一步流程中，评估者将接受相关培训来对绩效的不同维度进行识别。他们[通过视频中的案例来练习进行绩效评估](#)，并通过与专家提供的“真实”评估做比较来了解自己的评估水平。视频中的案例可以作为参考案例，每一个案例都定义了绩效量表上的一个锚定点，这样就形成了一个案例量表，如图23-2的右侧所示。

使用案例量表对每个新成员进行评估时，我们都可以将评估对象与

锚定案例做比较，这样一来，评估就变成了一种相对判断。由于相对判断比评分更不容易受到干扰，案例量表比使用数字、形容词或行为描述的量表更可靠。

参照框架培训在几十年前就已经为人们所知，用它进行评估，结果显然更加准确，噪声数量也更小，然而，这种方法却难以得到普及。我们很容易猜到其中的原因，参照框架培训、案例量表以及其他力图达到相同目标的工具都很复杂，而且构建这样的工具需要花费大量时间。为了保证切实可用，它们常常需要根据公司的需求甚至为评估部门专门定制，并且要随着岗位要求的变化而不断更新。这些评估工具要求公司在已经投入巨大资金的绩效管理系统中追加新投入，这与现在的做法背道而驰。在第六部分，我们会介绍更多有关减少噪声成本的内容。

此外，对于任何组织来说，如果控制了评估者在评估过程中的噪声，也就限制了评估者按照自身意图来影响评估的空间。要求管理人员接受额外的评估者培训，在评估过程中投入更多精力并放弃对结果的一些控制，这些举措肯定会面对相当大的阻力。值得注意的是，迄今为止，[大多数关于参照框架培训的研究](#)使用的被试都是学生，而非实际管理者。

绩效评估这一大课题产生出了许多现实问题和哲学问题。例如，有些人会问，在当今的组织中，个人绩效的概念在多大程度上是有意义的，因为组织的成果往往取决于人们之间的互动。如果我们认为这个概念确实是有意义的，那么就需要想一想，在一个既定的组织中，个人绩效水平是如何分布的，例如，绩效是否服从正态分布，或者是否存在做出极大贡献的“[超级人才](#)”。如果你的目标是让人们表现出最好的一面，那么你自然会问，衡量个人绩效并用衡量结果来激发人们的恐惧与贪婪是不是最好的方法？这是不是有效的方法？

如果你正在设计或修改组织的绩效管理体系，那么你就需要回答以上问题乃至其他更多问题。我们的目的不是研究这些问题，而是提出一

个更为温和的建议：如果你确实测量了绩效，你的绩效评估过程很可能充满了系统噪声，因此这些评估过程可能基本上是无用的，甚至会适得其反。减少噪声是一项挑战，因为它无法通过简单的技术修复来解决。它要求评估者对自己做出的判断有清晰的思考。你很有可能会发现，可以对量表进行清晰的界定，可以训练人们始终按照相同的方法使用这一量表，以此来提升判断品质。这种降噪策略也适用于许多其他领域。

- 消除噪声

定义量表至关重要

- 我们在绩效评估上花了很多时间，结果却是，绩效评估结果只有1/4反映了绩效表现，而其他3/4都是系统噪声。
- 我们尝试了用360度反馈和强制排名来解决这个问题，但我们可能会让事情变得更糟。
- 如果出现了很多水平噪声，多半是由于不同的评估者对“好”或“优秀”的含义有完全不同的看法。只有当我们给他们具体的案例作为评估量表上的锚定点时，他们才容易达成一致意见。

第24章

人才招聘，以结构化指标衡量人才

如果你曾经找过工作，那么“招聘面试”这个词可能会唤起你的一些生动而紧张的回忆。工作面试，即求职者与未来的主管或人力专员会面，是进入许多组织的必经之路。

在大多数情况下，面试遵循着既定的流程。在寒暄之后，面试官会要求应聘者描述他们的个人经历，或者详细地描述他们经历过的某些特定情形。再问一些诸如个人成就与挑战、工作动机以及能给公司带来哪些改进之类的问题。面试官通常还会要求应聘者描述自己的个性，并解释为什么认为自己很适合所应聘的职位或公司文化，有时也会问应聘者的兴趣爱好。最后，应聘者通常有机会问面试官一些问题，面试官则会评估这些问题是否重要，以及应聘者是否有洞察力。

如果你的主要工作是负责招聘，那么你所用的甄选方法很可能就包含上述流程中的某些环节。一位组织心理学家指出：“[很少有人可以不经过面试就被录用](#)，这样的情况令人难以想象。”然而，几乎所有的专业人士在面试中做招聘决策时，都会[在某种程度上依赖直觉性判断](#)。

招聘面试的普遍性反映了人们的一种根深蒂固的观念，即在选择与我们共事之人时非常重视判断的价值。作为一项判断工作，人事选拔有个巨大的优势：由于它的广泛应用以及重要的作用，组织心理学家们对它进行了详尽的研究。1917年《应用心理学》（*Journal of Applied Psychology*）期刊的创刊号就将招聘确定为“[首要问题](#).....因为人的能力毕竟是一个国家最重要的资源”。一个世纪后，我们对各种人才选拔技术（包括标准面试）的有效性有了很多了解。没有任何一项复杂判断任务能得到如此之多的现场实验研究的关注。人事选择因此成为一个完美的测试案例，我们可以将从中获得的经验应用到包括多选项决策的很多

判断任务中去。

面试的风险

如果你对求职面试的研究不熟悉，接下来的内容可能会让你大吃一惊。从本质上说，如果你的目标是确定哪些应聘者在工作中会成功，哪些会失败，那么标准面试并不能提供非常有用的信息。更直截了当地说，它们往往毫无用处。标准面试也被称作非结构化面试，它与结构化面试有所不同。我们很快会介绍结构化面试。

为了得出这一结论，无数研究人员考察了评估者在面试中给应聘者的高评分与应聘者最终取得工作上的成功之间的相关性。如果面试中的高评分与成功之间的相关性很高，那么面试或其他以相同方式计算相关性的招聘技巧，都可以被认为是候选人绩效良好的预测指标。

这里需要注意的是，“如何定义成功”是一个非常重要的问题。通常情况下，绩效是基于上级评估的评分而来。有时，工作年限也是衡量成功的标准。当然，这些衡量标准都有一些问题，特别是绩效评估的有效性值得怀疑，这一点我们在上一章中已经指出。然而，为了评估雇主在选择员工时所做判断的品质，使用同一雇主在评估其员工时所做的判断似乎是合理的。

那么，这些分析得出了什么结论呢？在第11章中，我们提到了传统面试的评分与工作绩效评分之间的相关系数是0.28。[其他研究报告的相关系数为0.2~0.33](#)。正如我们所看到的，按社会科学的标准，这体现出了很好的相关性。但是，要据此做决策，这一相关性就不算好了。使用我们在第三部分介绍的一致性比率，可以计算出一个概率：根据前面提到的相关水平，如果你对两个候选人的了解只限于一个候选人在面试中比另一个表现得好，那么这个候选人确实比另一个绩效更好的概率仅为56%~61%。这当然比掷硬币好一些，但是据此做出重要决策并不是一

个稳妥的方法。

诚然，面试除了对候选人做出判断外，还有其他目的。尤其是，面试为公司提供了一个向候选人推销自己的机会，提供了一个与未来同事建立友好关系的机会。然而，从组织的视角来看，组织在人才选拔上投入了大量时间和精力，面试的最主要目的显然还是选拔。在这项任务上，面试做得并不成功。

面试中的噪声：面试官更倾向于与自己相似的求职者

我们很容易理解为什么传统面试在预测工作绩效时会产生误差。有些误差与我们所谓的[客观无知](#)（见第11章）有关。工作绩效取决于很多因素，包括你雇用的人适应新职位的速度，以及各种生活事件对其工作的影响。在招聘时，很多情况都是无法预测的。这种不确定性限制了面试的预测效度，事实上，也限制了任何其他人事选拔技巧的预测效度。

面试也是心理偏差的重灾区。近年来，人们已经意识到，面试官往往无意识地青睐那些与他们[有相似的文化背景或共同之处](#)的候选人，涉及的因素包括性别、种族和教育背景。许多公司现在已意识到偏差带来的风险，并试图通过对招聘专家和其他员工实施专项培训来应对这些风险。几十年来，其他一些偏差也逐渐为人们所知。例如，外貌在候选人评估中起着重要作用，即使是那些与外貌无关的职位也是如此。所有或者说绝大多数招聘人员都有这样的偏差，评定具体某个候选人时，如果外貌发挥了作用，招聘者们就会产生一个共同的误差，即在候选人评估中的正偏差或负偏差。

关于面试过程中也存在噪声这一事实，你应该不会感到惊讶：不同的面试官对同一个候选人的反应不同，得出的结论也不同。[两位面试官对同一位应聘者的评分的相关系数](#)为0.37~0.44（PC=62%~65%）。其中一个可能的原因是，应聘者在不同面试官面前的表现不一致。但即使

是多位面试官同时面试同一位应聘者，前者对后者的表现给出评分的相关性也很低。据一项元分析估计，面试官们的评分相关系数仅为0.74（PC=76%）。这意味着，你和另一位面试官在同一场面试中对同样的两位候选人进行面试，最后仍然有1/4的可能性无法就哪位候选人更优秀达成一致。

这种差异很大程度上是模式噪声的产物，即面试官们对某位应聘者的特殊反应是有差异的。大多数组织在预料到了这种差异后，要求几位面试官面试同一位候选人，并将面试结果以某种方式汇总。通常，汇总意见是通过讨论形成的，在讨论中必须达成共识——正如我们在前文中看到的那样，这一过程本身会产生问题。

一个更令人惊讶的发现是，在面试中存在很多情境噪声。例如，强有力的证据表明，应聘者能否获得面试官的聘用推荐，与面试过程中非正式关系的建立阶段所形成的印象有关。该阶段就是面试刚开始的2~3分钟里，面试官与应聘者寒暄，从而让应聘者放松的阶段。因此[第一印象非常重要](#)。

也许你认为，凭第一印象做出判断没什么问题。至少我们从第一印象中了解到的一些东西是有意义的。众所周知，在与新朋友交流的最初几秒钟里，我们确实能了解到一些东西。对于有经验的面试官来说尤其如此。但面试的前几秒钟恰恰反映了你从第一印象中获得的表面特质：早期的认知主要基于应聘者的外向性和语言能力。即使是[握手的感受](#)，也是能否获得面试官推荐的重要预测因素！我们可能都喜欢坚定的握手，但很少有招聘人员会有意识地将握手是否坚定作为关键的聘用标准。

面试官心理学：我们总是太相信“第一印象”

为什么第一印象会对面试的最终结果产生影响？其中一个原因是，

在传统面试中，面试官可以自由地将面试引向他们认为合适的方向。他们很可能会问一些能证实第一印象的问题。如果某位应聘者显得害羞和矜持，面试官可能会问一些尖锐的问题，比如询问应聘者过去在团队中工作的经历，但他们可能不会向那些看起来开朗、合群的人问同样的问题。他们收集的这两位候选人的相关证据将不尽相同。一项研究考察了面试官基于应聘者的简历和能力测试成绩，形成了对该应聘者积极或消极的第一印象之后的行为。结果发现，第一印象对面试过程有重要的影响。例如，面试官形成了[积极的第一印象](#)后，面试时提出的问题就较少，而且往往会把公司“推销”给应聘者。

第一印象的作用并非面试中唯一有问题的因素。另一个因素是，作为面试官，我们希望面前的候选人表现得合乎情理（在第13章中我们讨论过，这是我们过度寻求一致性的一种表现）。在[一个让人难以置信的实验](#)中，研究人员让学生扮演面试官或应聘者的角色，并告诉他们，面试中只能提问封闭式的是非题。然后，他们让一部分应聘者随机地回答问题。按照事先安排好的方法，应聘者会根据问题的首字母来回答“是”或“否”。研究人员有些讽刺地指出：“一些应聘者最初担心随机答题会被人发现他们在胡说八道，从而导致面试进行不下去。但他们的担心是多余的，面试顺利地进行了下去。”你想得没错：没有一位面试官意识到，候选人在随机地给出答案。更糟糕的是，当询问面试官是否能“根据与面试者在一起的这段时间，推断出很多这个人的相关信息”时，得到应聘者随机回答的面试官与得到应聘者如实回答的面试官，在认为自己能做到这一点的概率方面数值差不多。这就是我们创造连贯性的能力。就像我们经常能在随机数据中找到一个想象的模式，或从云的轮廓中想象出一个形状一样，我们能够在完全没有意义的答案中找到逻辑。

还有一些不那么极端的情况，请看下面这个例子：我们中的一人不得不面试一位应聘者，这位应聘者之前是一家中型公司的首席财务官。面试官注意到这个应聘者入职了几个月后就离职了，于是问他为什么这

样做。应聘者解释说，是因为“与CEO有战略分歧”。另一位同事也面试了这位应聘者，问了同样的问题，得到了同样的答案。然而，在随后的汇报中，两位面试官的观点却截然不同。其中一位之前就对该候选人形成了积极评价，因此认为候选人离开前公司的决定是正直勇敢的表现。而另一位面试官之前就形成了消极的第一印象，则把同样的事实解释为不灵活，甚至可能是不成熟的表现。这个故事说明，即使我们自认为对候选人的判断完全是以事实为依据的，我们对事实的解释都会受到先前态度的影响。

传统面试的局限性让我们严重怀疑能否从中得出有价值的结论。然而，面试中形成的印象是生动的，面试官通常对此充满信心。当把面试结论和候选人的其他线索结合起来时，我们往往会把面试看得太重，而把其他可能更具有预测性的数据，比如能力测试的成绩看得太轻。

面试的故事是情境噪声在实际生活中的一个例子。应聘教职的教授经常被要求为同行授课，以确保他们的教学技能达到学校的标准。这种授课当然不同于平常的正式授课。有人曾经目睹过一位候选人在这个试讲中给人留下了不好的印象，很明显是因为当时的压力过大，因为候选人的简历中提到了自己曾获得优秀的教学评价和几项优秀教学奖。然而，他在一个高度人为化的情境中的失败表现给人留下了深刻的印象，这种印象所占的决策权重，高于那些代表出色教学表现的抽象数据所占的决策权重。

还要记住，当面试不是唯一的候选人信息来源时，例如，还有能力测试、推荐制度或其他信息时，这些不同的信息必须汇总成一个整体性判断。这引发了一个问题，你现在应该能意识到这一问题：是应该使用判断（诊断性汇总）还是公式（机械性汇总）来汇总信息？正如我们在第9章中看到的，无论是预测一般性工作表现，还是预测具体情况中的工作表现，机械性方法都更好用。可惜，据调查显示，[绝大多数人力资源专业人士都更赞成使用诊断性汇总](#)。这种做法给已经充满噪声的面试

过程增加了另一个噪声源。

通过结构化流程提升人事筛选品质

如果传统面试和基于判断的招聘决策只会产生有限的预测效度，那我们该怎么办呢？幸好，研究还提出了一些关于如何改进人员选拔的建议，有一些公司正在关注这些建议。

谷歌就是这样的一个例子。它改进了人员选拔措施并报告了改进的结果。拉斯洛·博克（Laszlo Bock）曾任人力运营部高级副总裁，他在《重新定义团队》（*Work Rules!*）一书中讲述了这些事情。尽管谷歌专注于招聘最优秀的人才，并投入大量资源寻找合适的人选，但其仍举步维艰。对招聘面试的预测效度进行审查后发现“[相关度为0](#).....完全随机得一团糟”。谷歌为解决这一问题所做的变革，体现了近几十年来研究中提出的一些原则。这些变革措施也是决策卫生策略的例子。

其中一个策略是汇总，你应该已经很熟悉它了。人们在这种背景下使用汇总并不奇怪。几乎所有公司都会汇总多个面试官对同一个候选人的评价。谷歌也不甘落后，有时会让求职者参加25轮面试！博克得出的结论之一是面试次数应该减少到4次，因为他发现在4次面试的基础上再增加额外的面试几乎不会提升预测效度。不过为了确保预测效度，谷歌严格执行了一项规定：公司要确保面试官在相互交流之前对候选人进行独立打分。这条规定并不是所有公司都有。谷歌相信：汇总是有效的，前提是判断是独立的。

谷歌还采用了一种我们尚未详述的决策卫生策略：将复杂的判断结构化。结构化一词可以有很多意义。此处使用这个术语时，我们按照3个原则来定义结构化的复杂判断：分解、独立性和推迟整体性判断（delayed holistic judgment）。

第一个原则是分解，它将决策分解为多个组成部分，每个部分对

应一个中介评估法。这一步的目的与指南中确定子判断的目的相同：它能确保判断者将注意力集中在重要的线索上。分解就好像一个路线图，指明了需要什么样的数据，并且过滤掉不相关的信息。

以谷歌为例，人事决策可以分解成4个中介评估法：一般性认知能力、领导力、文化契合度（称为“谷歌特质”）和角色相关知识。某些评估会被分解成更小的成分。注意，应聘者的良好外表、流畅的表达、令人兴奋的爱好的，以及招聘人员在非结构化面试中可能会注意到的任何其他方面，无论是积极的还是消极的，都不在该列表上。

为招聘任务创建这种评估体系似乎是人们的共识。事实上，如果你正在招聘一名初级会计或行政助理，职位描述中就明确规定了所需的能力。然而，专业招聘人员都知道，对于非常规职位或高级职位来说，对核心评估维度进行定义会很困难，而且定义这一步骤经常被忽视。[一位知名的猎头](#)指出，以一种足够具体的方式定义所需能力是一项富有挑战且经常被忽视的任务。他强调了“在问题的定义方面有所投入”对决策者的重要性：与任何候选人会面之前，都有必要提前花一些时间，就那些清晰且详细的职位描述达成一致。这里的挑战是，许多面试官使用的是经过协商或妥协而产生的浮夸的职位描述。这些描述只是一份模糊的清单，列出了一位理想候选人应该具备的所有特征，没有提供如何对这些特征进行调整或在不同特征之间进行权衡和取舍的方法。

结构化判断的第二个原则是独立性，要求独立收集每个评估维度的信息。只列出职位描述的组成部分是不够的：大多数进行传统面试的招聘人员也知道他们要在应聘者身上寻找的四五种能力。问题在于，他们没有在面试过程中单独评估这些因素。每个评估维度都会互相影响，这使得每个评估都充满噪声。

为了克服这一问题，谷歌精心安排了各种方法，以确保评估以事实为基础且相互独立。最明显的举措或许就是引入了[结构化的行为面试](#)（structured behavioral interviews）。在这种面试中，面试官的任务不是

去决定他们是否喜欢某个候选人，而是收集评估结构中每个评估维度的相关数据，并在每个评估维度上为候选人打分。为此，面试官必须询问候选人在过去的某些情况下的行为，这些问题都是预先设定好的。面试官还必须记录答案，并参照一个预先制定好的评分量表，使用统一的评分标准进行评分。对于每一个问题，评分标准会就一般性答案、好答案或很好的答案给出具体的示例说明。这种标准统一的量表有助于减少判断中的噪声。我们在前一章中介绍的行为锚定评估量表也是它的一个例子。

这种方法听起来不同于传统的聊天式面试，事实也确实如此。这实际上更像是一场考试或审问，而不是一次商业会面。有证据表明，应聘者 and 面试官都不喜欢结构化面试，或者说，至少更喜欢非结构化面试。[究竟什么样的面试才算得上结构化面试](#)，针对这一问题的讨论仍在继续。尽管如此，有关面试的文献研究得出了一个一致的结论：[结构化面试比传统的非结构化面试更能预测应聘者未来的表现](#)，结构化面试与工作绩效的相关系数为0.44~0.57。用我们的PC指标来说，你通过结构化面试挑选出更优候选人的概率为65%~69%，明显高于非结构化面试56%~61%的概率。

谷歌在它所关心的一些维度上也使用了其他一些数据作为信息。为了测试与工作相关的知识，谷歌会依赖一部分[工作样本测试](#)（work sample tests），比如让应聘者编写几段代码。研究表明，工作样本测试是工作绩效的最佳预测指标之一。谷歌也使用“关系户推荐制”，但推荐人并不是由候选人本人指定，而是由与候选人有交集的谷歌员工指定。

结构化判断的第三个原则是推迟整体性判断，简单概括来说就是：不排除直觉，但推迟直觉。在谷歌，最终的招聘推荐是由招聘委员会共同做出的，该委员会审查候选人的完整资料，包括每次面试中每个评估维度上的评分，以及支持这些评估的其他相关信息。然后，委员会根据这些信息决定是否聘用候选人。

尽管这是一家以数据驱动文化著称的公司，尽管有证据表明机械性汇总数据比诊断性汇总数据更有效，但最终的招聘决策并不是机械性的，它仍然是一个判断。在这个过程中，委员会成员会考虑所有的证据，并对其进行整体权衡，以此来讨论这个人是否会在谷歌取得成功。这个决策不仅仅是计算得出的。

在下一章，我们会解释为什么我们认为用这种方法做出最终决策是明智的。但要注意，虽然决策不是机械性的，但谷歌最终会依据4位面试官给出的平均分做出招聘决策。他们也了解其他一些相关的证据。换句话说，只有在收集和分析了所有证据之后，谷歌才允许决策过程中存在判断和直觉。因此，每个面试官和招聘委员会成员快速形成直观印象并急于做出判断的倾向都得到了控制。

再重述一遍，分解、独立性和推迟整体性判断这三个原则，并不一定能为所有试图改进人事选拔过程的组织提供一个模板。但这些原则与组织心理学家多年来提出的建议基本一致。事实上，这些原则与本书作者卡尼曼于1956年在[以色列军队](#)中实施的一些选拔方法有相似之处，并且在《思考，快与慢》一书中有所描述。该选拔过程，就像谷歌公司使用的选拔流程一样，正式建立了一个评估结构——需要评估的性格和能力维度的列表。它要求面试官依次列出与每个维度相关的客观证据，并在评估下一个维度之前对该维度进行评分。并且，它允许招聘人员在完成结构化评估之后再运用判断和直觉做出最终决策。

有不可辩驳的证据表明，结构化的判断过程，包括结构化面试，在招聘中具有优越性。它能为采用这一方法的高管们提供[实用性的建议和指导](#)。正如谷歌的例子以及其他研究人员指出的那样，结构化的判断方法成本会更低，因为会面十分耗费时间。

尽管如此，大多数高管仍然相信非正式的、基于面试的方法具有不可替代的价值。值得注意的是，许多应聘者也相信只有通过面对面的面试，他们才能向未来的雇主展示自己真正的实力。研究人员称之为“[错](#)

[觉的持续](#)”。很显然，招聘人员和应聘者都严重低估了招聘判断中的噪声。

- 消除噪声

招聘中的结构化

- 在传统的非正式面试中，我们往往会有一种无法抗拒的直觉，感觉自己了解该应聘者，并且知道这个人是否符合要求。但是，我们必须学会不信任这种感觉。
- 传统面试问题很多，因为其中不仅有偏差，还有噪声。
- 我们必须让面试甚至整个人事选拔过程结构化。我们应该从一开始就更明确、更具体地定义要在候选人身上寻找什么特征，并确保自己在每个维度上独立地对候选人进行了评估。

第25章

中介评估法，做出明智决策的核心方法

不久前，我们两个人（卡尼曼和西博尼）与我们的朋友丹·罗瓦洛（Dan Lovallo）一起，介绍过组织中使用的一种决策方法。这一方法的首要目的在于减少噪声，我们称之为中介评估法。该方法涵盖了我们前几章所介绍的大多数决策卫生策略，并且它的应用范围很广泛，只要在计划或选项的评估过程中需要考虑和权衡多维度信息，我们就可以使用该方法。该方法能够以多种形式广泛应用于所有类型的组织中，包括不同的公司、医院、大学和政府机构等。

我们在这里用一个虚构的例子来说明这个方法，这个例子是根据几个真实案例改编而成的。我们虚构了一家名叫迈普科（Mapco）的公司，我们将跟随该公司的脚步，一起来研究一个重大的、具有变革性意义的企业收购项目。我们将重点介绍迈普科公司所采取的步骤与其他公司在类似情况下所采取的常规步骤有何不同。你将看到，两者的确有实质性差异，但这种差异很微妙，稍不留神就可能注意不到。

第一次会议：协议方法

收购竞争对手罗德科（Roadco）公司的想法由来已久，目前时机已经成熟，因此迈普科公司的领导人拟召开一次董事会议来讨论这一问题。迈普科的CEO琼·莫里森（Joan Morrison）召开了一次董事会战略委员会会议，初步讨论这次可能的收购案以及应采取哪些措施来改进董事会的收购决议。会议一开始，莫里森就提出了一个让委员们大吃一惊的建议：

我想提议董事会尝试一种新的决议流程来讨论罗德科收购案，新流程的名字不那么有魅力，它叫中介评估法，但其观点非常简单。它的灵感来自战略方案评估和求职者评估之间有相似性。

有研究表明结构化面试比非结构化面试效果更好，这一点你们想必已经很熟悉了，那么同理，让招聘决策结构化也可以改善决策品质。你们也都知道，我们的人力资源部在招聘决策中采用了这些原则。大量研究表明：结构化面试具有更高的准确性——这是我们过去使用的非结构化面试无法企及的。

我发现，对不同候选人的评价和在重大决策中对不同选项的评价有着明显的相似之处——不同的选项就像不同的候选人一样。这种相似性使我联想到，我们应该把评估候选人的那套方法应用于我们的工作中，也就是用在评估战略方案上。

委员会其他成员起初对这个类比感到困惑。他们争辩道：“原有的招聘流程就像是一台运转良好的机器，可以做出无数类似的决策，而且不会面临巨大的时间压力。相反，进行战略决策需要开展大量独特的工作，而且决策过程必须及时。一些委员向莫里森明确表示，他们将反对任何会导致决策延迟的提议，此外，他们还担心新的决议流程将给迈普科的研究人员增加不少职责内的负担。

莫里森对这些反对意见进行了直接回应。她向同事们保证：“结构化的流程不会拖延决策，它仅仅是为董事会讨论这项交易制定一个议程而已。”她进一步解释道：“我们应事先确定一个清单，定好从哪些方面对此次交易进行评估，就像面试官会列出岗位需求清单、确定候选人需要具备的特质或属性一样。我们将确保董事会对需要评估的各个方面进行逐一、独立的讨论，就像结构化面试中面试官们对候选人的不同维度依次进行评估一样。只有在这之后，我们才能开始讨论是接受还是拒绝

这项交易。这一流程将更有效地发挥董事会的集体智慧。

“当然，如果我们同意这种方法，这将对信息的呈现方式以及交易团队如何准备会议产生影响。因此，我现在想听听你们的想法。”

一位仍有疑问的委员问莫里森，这种结构化究竟给招聘决策的品质带来了什么好处，为什么她相信战略决策也能获得这样的好处。为了引导他理解其中的逻辑，莫里森解释道：“中介评估法能够将信息的价值最大化，因为它确保了各个维度在评估时是相互独立的。我们通常进行的董事会讨论看起来很像非结构化面试。我们总想着去达成最终的目标——做出决策，并根据该目标处理所有信息。我们从一开始就在想要如何做个了结，并且会尽可能快地达成目标。就像招聘人员在非结构化面试中的表现那样，我们有可能只是在用讨论来确认第一印象而已。使用结构化的方法将迫使我们等到完成所有评估之后再考虑达成决策这一目标。我们将独立的评估作为中程目标，这样一来，就能考虑到所有可用的信息，并确保我们对交易的某一方面得出的结论不会改变我们对交易的其他方面的解读。”

委员们同意试用这种方法，但是他们问道：“中介评估要评估什么呢？莫里森心中有没有预先定义好的列表？”“没有，”莫里森回答道，“如果我们用此方案来解决常规决策，那可能会有，但在当前这种情况下，我们需要自己定义中介评估的内容。这一点非常重要：我们需要自己来决定收购案的哪些重要方面需要评估。”战略委员会同意第二天再开会来完成这个任务。

第二次会议：定义中介评估法

莫里森继续说道：“我们要做的第一件事就是起草一份有关这项交易的完整独立评估清单。这些评估将由杰夫·施耐德（Jeff Schneider）的研究团队完成。我们今天的任务是编写评估清单。评估清单应该是全面

的，也就是说，你能够想到的任何与交易有关的事，都应该列入其中，并且它应该至少会影响清单中的一项评估内容。我所说的‘独立’是指相关事实最好只影响清单中的一项评估内容，尽量减少重复的情况。”

于是，整个团队开始忙碌起来，他们列出了一长串看起来与交易有关的事实和资料，然后把这些内容编排成一个评估项清单。参与者很快发现，难点在于如何让清单既简短又全面，且不同的评估项之间不会相互重叠，但这项任务依然是可以应付的。确实如此，团队最终列出的清单所包含的7项评估，与董事会期待在收购提案报告中看到的内容目录看起来很相似。除预期财务模型外，该清单还包括对目标公司管理团队的表现进行评估，对可能获得的协同效应进行评估等。

一些委员对会议未能就罗德科收购案产生新颖的见解而感到失望，莫里森解释说，那不是会议的目标，目前的目标是向负责研究此次收购的交易团队介绍情况。她说，每项评估都是交易团队报告中不同章节的主题，董事会将分别对其进行讨论。

在莫里森看来，交易团队的任务不是告诉董事会他们对交易的整体看法——至少现在还不是时候，而是对每项中介评估提供客观、独立的评价。莫里森解释道：“最终，交易团队报告中的每一章都应该以一个评分结尾，这个评分能够回答‘抛开最终决策中应给予这个主题的权重不谈，这项评估的证据在多大程度上支持或反对该交易’这个问题。”

引入外部团队

负责评估交易的团队负责人杰夫·施耐德当天下午召集了团队，并开展工作。与团队通常的工作方式相比，这次的变动不大，但他强调了这些变动的重要性。

施耐德解释道：“团队的分析师应该尽可能客观地做出分析。毫无疑问，评估应该建立在事实的基础上，但他们也应该尽可能使用外部视

角。”然而，团队成员并不确定他所谓的“外部视角”是什么意思，因此施耐德举了两个例子，用到了莫里森确认过的两个中介评估项。他说，要评估这项交易获得监管部门批准的可能性，他们首先需要找出基准概率，即那些与本次交易有可比性的其他交易最终获得批准的百分比。这项任务反过来要求他们定义一个与本次交易相关的参考类别，即一组被认为与本次交易具有足够可比性的交易。

施耐德接着解释了如何对目标公司产品研发部门的技术能力进行评估——这是莫里森列出的另一个重要评估。“仅仅以事实为基础来描述公司最近的业绩是‘好’或‘非常好’是不够的。我期待做出这样的判断：从最近的产品发布记录来看，这个产品研发部门在同类公司中处于前40%的位置。”他解释道，“总之，我们的目标是使评估尽可能有可比性，因为基于相对判断要比基于绝对判断好。”

施耐德还有一个要求，他说：“按照莫里森的指示，为降低评估之间相互影响的风险，评估应尽可能相互独立。”因此，他为不同的评估分配了不同的分析师，并要求他们独立工作。

一些分析师对此表示惊讶，反问道：“团队合作不是更好吗？如果你不想让我们交流，组建团队有什么意义？”

施耐德意识到，他需要解释为什么独立性是必须的。“你可能知道招聘中的光环效应，”他说，“当你对候选人的总体印象影响到你对他在某一特定维度上的技能评估时，就发生了光环效应。这是我们想避免的。”但一些分析师似乎认为光环效应不是一个很严重的问题，因此施耐德用了另一个比喻：“如果有4名犯罪目击者，你会让他们在作证前互相交谈吗？显然不会！你不想让证人之间彼此影响。”尽管分析师们没有觉得这种类比让他们茅塞顿开，但施耐德想，这起码传递了信息。

然而，施耐德没有足够的分析师来完成完全独立的评估。简是团队中经验丰富的成员，负责了两项评估。施耐德尽可能给她选择了完全不同的两项评估，他要求简在完成第一项评估并准备好报告之后再开始另

一项。另一个问题是对管理团队品质的评估，施耐德担心，分析师们很难将对管理团队内在品质的评估与对公司近期业绩的判断区分开来（团队当然会对这些业绩进行详细研究）。为了解决这个问题，施耐德外聘了一位人力资源专家就管理团队的表现发表看法。他想，这样一来就能获得更独立的意见。

交易团队发现，施耐德还有一个指示和以往不同：交易报告的每一章应聚焦一项评估，根据莫里森的要求，还要以评分的形式给出自己的结论。不过，施耐德也补充道，分析师们应该在每一章中都包含有关评估的所有事实性信息。“不要隐瞒任何事情，”他要求他们，“章节的总体基调必须与给出的评分保持一致，但如果有些信息与评分不一致甚至互相矛盾，也不要隐瞒。你的工作不是推销你的建议，而是代表客观事实。如果事情很复杂，那就接受吧——事情通常就是复杂的。”

本着这样的精神，施耐德鼓励分析师不要隐瞒自己对每项评估的信心。“董事会知道你没有完善的信息。如果你告诉他们，你真的毫无把握，这会对他们有帮助。如果你遇到了真正让你踌躇的事情，比如有可能会破坏这项交易的因素，你当然应该立即报告。”

交易团队按照指示开展工作。幸好，没有发现重大的破坏交易的因素。交易团队为莫里森和董事会编写了一份报告，涵盖了清单所列的所有评估。

决策会议：评估-讨论-评估法

为决策会议做准备时，莫里森阅读了交易团队的报告，她立刻注意到了一些重要的事情：虽然大多数评估都支持这项交易，但它们并没有绘制出一幅简单、乐观、全面的蓝图。一些评估非常支持这项交易，另一些则不然。她知道，这些差异是保持评估相互独立后的必然结果。当过度一致性受到控制时，现实并不会像大多数董事会报告所呈现的那样

具有一致性。“很好，”莫里森想，“不同评估之间的矛盾之处将引发问题和讨论，这正是我们需要在董事会上好好讨论的。当然，多样化的结果不会让决策变得更容易，但它会让决策变得更好。”

莫里森召开了一次董事会会议，目的是审查这份报告并做出收购决议。她向董事会成员解释了交易团队所用的分析方法，并要求董事会成员也遵循同样的原则。她说：“施耐德和他的团队一直努力使各项评估相互独立，我们现在也要独立地审查各项评估。这意味着我们将分别审查每一项评估，之后再讨论最终的决定。我们要把每一项评估视为一个独立的项目议程。”

董事会成员明白，遵循这种结构化的方法是有困难的。莫里森要求他们在讨论完所有评估之前不要对交易形成整体看法。但他们中许多人都是业内人士，对罗德科公司有自己的看法，不讨论这种整体看法感觉有点儿不自然。但他们知道莫里森想要什么，所以他们同意按她的规则行事，暂时不讨论自己的总体观点。

令他们惊讶的是，董事会成员发现这种做法非常有价值。在会议期间，他们中的一些人甚至改变了对这项交易的原有看法（尽管没有人知道，因为他们没有公开自己的观点）。莫里森主持会议的方式在很大程度上起了作用：她采用了[评估-讨论-评估法](#)，这种方法综合了深思熟虑和各种独立意见的优点。

莫里森是这样做的：在每次评估中，施耐德代表交易团队简要总结关键事实（董事会成员事先已经详细阅读了这些材料），然后，莫里森要求董事会成员在手机投票程序上对每项评估给出自己的评分。该评分与交易团队提议的评分可能相同，也可能不同。评分的分布情况会立即在屏幕上显示出来，但不显示评分者是谁。“这不是投票，”莫里森解释说，“我们只是想了解大家对每个话题的看法。”在讨论之前获得每个董事会成员的独立意见，这样做减少了社交影响和信息级联的危险。

董事会成员很快就某些评估项中达成了共识，但在讨论另一些评估

项的过程中出现了观点分歧。自然，莫里森引导大家花更多的时间讨论有分歧的项目。她要确保产生分歧的董事会成员们都发言，她鼓励他们的事实和论据表达自己的观点，但也要保持谦逊、接纳差异。有一次，当一位董事会成员在讨论中情绪激动、忘乎所以时，她提醒道：“我们都是理智的人，但我们还是会有分歧，这说明这本身就是一个连理智的人之间都会产生分歧的话题。”

当一项评估的讨论接近尾声时，莫里森要求董事会成员再次就评分进行投票。在大多数情况下，这一次投票结果比第一轮投票更一致。每一项评估都重复进行这样的流程——第一次评估、讨论、第二次评估。

最后，到了要对这项交易下结论的时候。为了便于讨论，施耐德在白板上写下了每项评估的列表，每项评估都有董事会给出的平均得分。面对这项交易的概况，董事会成员们应该如何决策？

一位董事会成员提出了一个简单的办法：直接使用平均分。也许他知道，正如我们在第9章所讨论的那样，机械性汇总比总体性的诊断性判断更优。但另一位成员立即反对，在他看来，对一些评估项赋予的权重应该远高于其他评估项。第三位成员则建议采用不同层次的评估。

莫里森打断了大家的讨论。“这不仅仅是简单地计算出各项评估分数的总分，”她说，“我们之前并没有靠直觉行事，但现在是时候运用它了。我们现在需要的是你们的判断。”

莫里森没有解释她的逻辑，她曾以惨痛的代价领悟到这个教训。她知道，尤其是在做重要决策时，人们会拒绝那些让他们束手束脚、禁锢他们判断的规则。她见识过当人们知道要用某个公式做决策时，他们是如何钻系统的漏洞的。他们会改变评分以得出期望的结论——这与整个工作的目的是相悖的。此外，尽管当下的情况并非如此，但她仍然担心，可能会出现一些在评估列表中未被包括但可能起到决定性作用的因素（如第10章讨论的断腿因素）。如果出现此类破坏交易或促进交易的意外因素，那么基于评估平均值的纯机械性的决策过程可能会导致严重

错误。

莫里森也知道，让董事会成员在这个阶段运用他们的直觉与让他们在一开始就运用直觉是非常不同的。既然这些评估就在案头，且所有人都知道这些评估，那么最终决策肯定会稳妥地基于这些以事实为依据、被全面讨论过的评分。当董事会成员盯着一份倾向于支持交易的中介评估清单时，如果有人要反对这项交易，他就需要拿出强有力的证据才行。按照这个逻辑，董事会对这项交易进行了讨论，并采用了与所有董事会大致相同的方式进行表决。

重复决策中的中介评估法

我们在一次性、单一决策的情境中介绍了中介评估法，实际上这个流程也适用于重复决策。想象一下，如果不是迈普科公司正面临一项收购案，而是它本身是一家风险投资基金，需要不断地对初创企业进行投资，那么，此方法同样适用，流程也大同小异，只有两个地方不太一样，这两个地方会使流程变得更简单。

首先，初始步骤，即确定中介评估清单只需做一次。基金有投资标准，适用于所有预投资项目——这些就是评估清单中的内容，没有必要每次都重新进行设计。

其次，如果该基金做许多同类型的决策，它可以利用自己的经验来校准自己的判断。例如，每个基金公司都需要做的一项评估是评估管理团队的表现。我们建议这类评估应该基于一个参考群体来进行。也许你同情迈普科的分析师，因为除了评估一个特定的目标之外，他们还要收集具有可比性的其他公司的数据，而这确实是一项挑战。

在重复性决策中，比较判断变得容易得多。如果你已经评估了数十家甚至数百家公司的管理团队，那么你可以将这些公司作为参考群体。一个很实用的方法是：创建一个由锚定案例所定义的案例量表。例如，

你可能会说，目标管理团队与我们收购的“ABC公司”的管理团队一样好，但不如“DEF公司”的管理团队好。当然，所有参与者都必须了解锚定案例并定期更新。定义锚定案例需要在前期投入一定的时间，这种方法的价值在于，比较判断（将该团队与ABC公司和DEF公司的团队进行比较）相比于由数字或形容词定义的绝对评分更可靠。

6个步骤，用中介评估法改善决策流程

为了便于参考，我们在表25-1中总结了中介评估法所做出的主要改变。

表25-1 中介评估法做出改变的主要步骤

步骤	方法
1.	在流程开始时，将决策结构化为中介评估（对于重复性判断，这个步骤只需执行一次）
2.	尽可能确保中介评估使用外部视角（对于重复性判断，使用相对判断，如果有可能，使用案例量表）
3.	在分析阶段，尽可能使评估相互独立
4.	在决策会议中，分别审核每项评估
5.	在每次评估中，确保参与者各自独立做出判断；然后使用评估 - 讨论 - 评估法
6.	在做最终决策时，要推迟使用直觉，但不要禁用直觉

到这里，你可能已经辨认出了我们在前几章中介绍的几种决策卫生技术：对信息进行排序；将决策过程结构化并进行独立评估；使用基于外部视角的共同参考框架；汇总多个独立判断。中介评估法旨在改变决策过程，通过应用这些技术尽可能多地引入决策卫生策略。

毫无疑问，这种强调过程而不是强调决策内容的做法可能会引起一些人的质疑和不满，我们前面所描述的研究团队成员和董事会成员的那些反应并不罕见。内容是具体的，而流程是通用的，运用直觉和判断往往生动有趣，遵循流程则显得沉闷乏味。传统观点认为，好的决策来自伟大领袖的洞察力和创造力，当我们是领导者时，我们尤其愿意相信这一点。而对于许多人来说，流程这个词常常意味着官僚主义、繁文缛节和拖延。

那些全面或部分实施了中介评估法的公司和政府机构的经验表明，这些担忧具有误导性。虽然让已经官僚化了的组织的决策过程复杂化并不能让事情变得更好，但是，决策卫生策略不会花更多的时间，当然也不需要官僚化。相反，它促使人们展开辩论和向权威提出挑战，而不是延续官僚机构那令人窒息的达成共识的一贯方式。

采用决策卫生策略的理由很简单。商界和公共部门的领导者们往往完全意识不到他们最大或最重要的决策中存在的噪声，因而也不会采取具体措施来减少噪声。在这方面，他们就像招聘人员继续依赖非结构化面试作为人才选拔的唯一手段那样，都忽视了自身判断中的噪声，过度信任这一方法的有效性，并且对改进它的步骤一无所知。

就像洗手并不能预防所有的疾病一样，决策卫生也不能防止所有的错误，也就是说，它不会使每一个决策都很明智。但就像洗手一样，决策卫生警示人们关注一个虽看不见却普遍存在且具有破坏性的问题——哪里有判断，哪里就有噪声。因此，我们建议将决策卫生作为减少噪声的工具。

- 消除噪声

评估-讨论-评估法

- 我们有一个结构化的招聘决策流程。为什么我们不为战略决策做一个结构化的流程呢？毕竟，不同的选项就像不同的候选人一样。
- 这是一个艰难的决策。我们应该根据哪些中介评估法做决策呢？
- 我们对这个计划的直觉和整体性判断是非常重要的，但我们先不讨论它。当我们获得了独立评估的信息之后，直觉将更好地为我们服务。

NOISE

第六部分

最佳的噪声水平

· 发现噪声

减少噪声不仅充满意义，而且迫在眉睫

1973年，马文·弗兰克尔法官呼吁要坚持不懈地减少刑事审判中的噪声，他这么做是对的。他那种不那么正式、依靠直觉的噪声审查以及随后更正式、系统的噪声分析手段，共同揭示了法官对犯下相似罪行的人进行判决时存在差异，而且差异之大令人震惊。

《噪声》这本书的大部分内容都可以理解为使弗兰克尔的主张成为常态所做的努力，以及对这些主张背后的心理学基础的阐释。在一些人看来，似乎只有刑事司法系统中的噪声令人无法忍受甚至汗颜，然而实际上，在其他很多情况下，噪声也不是完全可以接受的。据说，如果将私营企业和公共部门的人互换位置，他们会在工作中做出不同的判断。在保险行业、员工招聘及人力资源评估、医学、法医学、教育、商业以及政府部门中，个体之间存在的噪声是造成误差的主要根源。我们也看到，每个人都会受到情境噪声的影响，比如，一些看似无关紧要的因素可能会导致我们在不同的时间做出不同的判断。

但是，正如人们强烈反对量刑指南一样，减少噪声的工作常常遭到强烈甚至疯狂的反对。一些人认为，量刑指南死板、不人道，它本身就存在不公平。几乎每个人都有过这样的经验，他们向公司、雇主或政府提出合理要求，得到的回应却只是：“我们真的很愿意帮你，但我们也无能为力，因为我们这里有明文规定。”一些规则看起来很愚蠢，甚至无情，但它们的存在基于一个充分的理由：可以减少噪声，或许也能减少偏差。

即便如此，一些为减少噪声所做出的努力还是引起了人们的广泛关

注，虽然总体而言，它们很难甚至不太可能得到应有的重视。算法和机器学习的使用让那些反对声音有了新的理由，没有人愿意让“算法”来主导自己的生活。

来自耶鲁大学法学院的凯特·斯蒂思（Kate Stith）和美国联邦法院法官何塞·卡布拉内斯（José Cabranes）发出的批判声音具有较大的影响力，他们对量刑指南进行了猛烈抨击。虽然他们的反对意见主要集中于刑事量刑领域，但也可以看成是反对教育、商业、体育等任何其他领域中的减少噪声策略。斯蒂思和卡布拉内斯认为，量刑指南的推动因素是“对行使自由裁量权的恐惧，也就是对判断的恐惧，以及技术官僚对专家和中央计划的信仰”。他们认为，“对判断的恐惧”妨碍了审判者对“手头上每起案件细节”的考虑，[在他们看来，“没有任何机械性的解决方案可以满足司法要求”](#)。

这些反对意见值得深入研究。在涉及各种判断的环境中，人们常常认为，“追求正义”就是要禁止采用任何死板、机械性的解决方案，因而只能允许使用甚至不得不强制使用存在一定噪声的解决问题的过程和方法。许多人呼吁应注意“手头上每起案件的细节”，在大大小小的医院、学校和公司中，这种呼吁具有很强的感染力。我们已经看到，决策卫生策略包括多种减少噪声的策略，并且其中大多数策略不包含机械性的解决方案，当人们将一个问题分解成多个成分时，他们的判断就不再是机械性的了，然而，即便如此，还是有许多人不愿使用决策卫生策略。

我们已将噪声定义为不必要的变异，既然不必要，就应该被消除，然而，现实情况往往更复杂，也更有趣。在其他条件恒定不变时，噪声可能是不必要的，然而，情况并非总一成不变，消除噪声的成本可能会超过收益。有时，即使成本和收益分析表明不减少噪声造成的损失更大，对政府部门和私人机构而言，消除噪声也可能造成一系列不好的甚至难以接受的后果。

关于减少或消除噪声的反对意见主要表现在以下7个方面：

第一，减少噪声的代价很大，不值得。减少噪声的必要步骤可能非常烦琐，在某些情况下，甚至无法实现。

第二，为减少噪声而采用的某些策略可能会给系统本身引入新的错误，有时甚至会造成系统性偏差。如果政府部门的所有预测者都过分乐观，那么他们的预测结果或许就不只是有噪声这么简单了，有可能根本就是错的。如果医院里的所有医生针对每一种疾病开的药都是阿司匹林，那问题就不再是噪声而是大量误诊了。

我们将在第26章中探讨上述2种反对意见，在第27章中，我们将探讨另外5种反对意见。这些反对意见都很常见，未来你可能会在很多地方听到，尤其是在越来越依赖规则、算法和机器学习的场景中。

第三，如果希望让人们觉得自己受到尊重和有尊严，我们就必须容许一些噪声的存在。噪声可能是人们最终乐于接受的不完美过程的副产品，因为这个过程会让每个人（员工、顾客、求职者、学生、嫌犯等）有机会表达自己的观点，有机会影响“法官”行使自由裁量权，让每个人都有可能感受到被他人关注和倾听。

第四，噪声可能是容纳新的价值观乃至促进道德和政治的发展至关重要的事物。如果我们消除了噪声，当道德和政治承诺朝着新的、意想不到的方向发展时，我们做出回应的能力可能被削弱。无噪声的系统可能会导致价值观僵化。

第五，一些力图减少噪声的策略可能会促成投机取巧的行为，导致人们利用制度的漏洞，或规避禁令。容许一些噪声甚至很多噪声的存在，可能是防止恶行所必需的。

第六，存在噪声的系统可能有很好的威慑作用。如果人们知道自己受到的惩罚可能很严重也可能很轻微，那么他们有可能会尽量避免犯错，至少对于风险规避型的人来说是如此。所以说，系统应该保留一些噪声，让它们成为一种能产生额外威慑力的方式。

第七，人们不希望自己被当成一件物品或机器上的一个小齿轮。某些减少噪声的策略可能会限制人们的创造力，并影响团队士气。

虽然我们会尽最大努力本着同理心来回应以上反对意见，但这并不代表我们认可这些反对意见，至少我们不认可将其作为拒绝减少噪声这一总体性目标的理由。这里我们要提出一个随后会反复重申的观点：一个反对意见是否具有说服力，取决于具体的减少噪声策略的应用情境。例如，你可能会反对机械地遵循量刑指南，同时却认为汇总各种独立判断作为参考是个好主意。考虑到这些情况，我们得出的结论是：减少噪声不仅很有意义，而且迫在眉睫，哪怕存在一些反对意见。在第28章中，我们会对人们每天都需要面对但并非总能意识到的困境进行探讨，进而找出支持这一结论的证据。

第26章

减少噪声的成本

每当我们建议人们消除噪声时，他们可能会以成本过高为由进行反对，并认为在极端情况下，减少噪声根本不可能。我们已经在商业、教育、政府以及其他领域听到过这种反对意见，这种观点虽然有一定的合理性，但言过其实了，或者只不过是一个借口。

为了证明我们的这种反对意见更有说服力，举一个高中老师为文章评分的例子。有位高中老师每星期都要对学生撰写的25篇文章进行评分。如果这位老师在每篇文章上花费的时间不超过15分钟，那么评分会由于存在噪声而变得不准确和不公平。老师可能需要考虑使用一点决策卫生策略来减少噪声，比如可以邀请另一位老师给这些文章评分。这样一来，两位老师都需要把每篇论文阅读一遍。再比如，这位老师在阅读每篇文章时，可以多花一些时间，或者构建相对复杂的评估过程，也可以按不同的顺序多次阅读这些文章，这几种方法同样可以达到提升评分品质的效果。这位老师还可以制定一个详细的评分指南清单，这也能起到一定的作用。此外，每天在固定时间阅读学生的文章也可以减少情境噪声。

但是，如果这位老师以为自己的判断十分精准，不会受到任何噪声的干扰，那就完全没有必要做上述那些事了，因为根本没必要这么麻烦。这位老师可能会认为，没必要使用清单或请同事帮忙评阅。想要知道这是不是小题大做，就需要对以下问题进行严格的分析：教师评分的准确性会提高多少？更为准确的评分到底有多重要？减少噪声需要花费多少时间和金钱成本？我们很容易想象在减少噪声方面需要的投入的上限，我们也很清楚，如果这些文章是更低年级学生或者更高年级学生写的，那么投入的上限也应该有所不同。显而易见的是，离大学录取越

近，噪声带来的风险就越大。

各种私人机构和公共部门也会在各种复杂的情境中进行这种基础性分析，因此拒绝采取某些减少噪声的策略。对于某些疾病，医院和医生可能很难制定出可以消除变异性的简单指南。在医疗诊断意见存在分歧的情况下，尽可能减少噪声显得尤为重要，因为这样可以挽救生命。但是，减少噪声的可行性和成本也需要考虑。某些检查可以消除诊断中的噪声，但是如果该检查是侵入式的、危险的且费用高昂，而诊断中的变异性很小且只会导致轻微的不良后果，那就没必要让所有医生都要求病人去做这项检查。

对员工的评估很少涉及生死，但噪声会导致员工之间的不平等，并使公司付出高昂的代价。我们已经看到，需要考虑减少噪声的可行性。减少噪声的努力值得吗？明显误判的案件可能会引起人们的关注，并造成令人尴尬、不体面甚至糟糕的后果。一个机构可能会认为事无巨细地纠正每一个步骤并不划算，因此有时候其结论是短视的、自私的、错误的，甚至可能会带来灾难性的后果。这样看来，某种形式的决策卫生策略的确值得一试。不过，“减少噪声代价太大以至于不值得这么做”的看法并非总是错误的。

简而言之，我们需要权衡减少噪声的成本与益处，这也是进行噪声审查如此重要的原因之一。在许多情况下，通过噪声审查我们发现，噪声造成了极大的不公平或者高昂的代价，甚至两种状况同时存在，这时，减少噪声就势在必行了。

噪声越少，错误越多？

另外一种反对意见是：某些减少噪声的措施本身可能会引发大到令人难以忍受的错误。如果减少噪声的工具不太灵敏，那么上述反对意见就更有说服力了，确实，有些减少噪声的措施甚至会增加偏差。如果像

脸书（Facebook）或推特（Twitter）这样的社交媒体平台引入一种严格的指南，来删除所有包含粗俗言语的帖子，这样做确实减少了噪声，但也可能将大量原本并不违规的帖子一并删除了。这些假阳性的删除操作会导致有方向性的错误——偏差。

在现实生活中，有很多为了减少人们的自由裁量权和会产生噪声的做法而进行的改革。虽然其中一些改革举措出于良好的动机，但“治疗方法”却比“疾病”更糟糕。经济学家阿尔伯特·赫希曼（Albert Hirschman）在《反动的修辞》（*The Rhetoric of Reaction*）一书中指出了反对改革的三种常见观点。首先，在某种意义上说，[这样的改革可能适得其反](#)，使原本打算解决的问题变得更糟；其次，这样的改革可能是徒劳无功的，根本无济于事；最后，它们可能使其他重要的东西受到威胁。举例来说，有人认为保护工会的努力以及加入工会的权利妨碍了经济增长。适得其反、徒劳无功、构成威胁都可能成为减少噪声的反对意见，而在这三者之中，适得其反和构成威胁这两种说法可能是最有力的。有时候，这些反对意见只是一种说辞，其目的是破坏一项实际上会带来巨大好处的改革。但是，某些减少噪声的策略确实会损害一些重要的东西，也有一些策略可能存在适得其反的风险。

一些反对量刑指南的法官认可上述风险。他们很清楚马文·弗兰克尔法官所做的努力，也没有否认自由裁量权会产生噪声，但他们认为减少自由裁量权会引发更多错误，而不是减少错误。他们引用[瓦茨拉夫·哈维尔](#)（Vaclav Havel）的话并坚称：“我们必须摒弃一种自大的信念，即这个世界仅仅是一个有待解决的谜题、一部等待人们发明操作指令的机器，或是一堆等待输入计算机里的信息，因为这台计算机迟早会给出一个普适的解决方案。”拒绝普适性解决方案的一个原因在于，人们有这样一种坚定的信念——人类面对的状况千差万别，优秀的法官能够应对这些差异。这就意味着要忍受噪声，或至少拒绝某些减少噪声的策略。

在国际象棋游戏程序刚出现时，一家大型航空公司为国际航班乘客提供了国际象棋游戏程序，请乘客与计算机对战。这个游戏程序有几种不同的难度级别。在最低级别中，该游戏程序遵循一条简单的规则：只要有可能，就去将对手的军。这样的游戏程序自然不包含噪声，它每次都按相同的方式落子，并且始终遵循着这条简单的规则。但这个规则会导致大量错误。事实上，这个游戏程序的棋艺很糟，甚至没有太多经验的国际象棋新手也可以击败它。这当然也很重要：乘客能够获胜，会很愉快。

接下来，我们一起来看一下美国某些州所采用的刑事量刑政策，即“[三振出局](#)”（three strikes and you're out）。意思就是，如果一个人犯下了三重重罪，那么他一定会被判终身监禁。这一政策减少了由于随机分配法官所导致的变异性。这项政策的一些拥护者主要是担心水平噪声的影响，以及某些法官可能对惯犯太过仁慈。消除噪声正是“三振出局”立法的核心。

然而，即使“三振出局”政策成功降低了噪声，我们也可以提出合理的反对意见——这一成功背后的代价过高。有些犯下三重重罪的人不应该被判终身监禁：也许他们不是暴力犯罪，也许是他们悲惨的生活处境迫使他们犯罪，也许他们能够改过自新。许多人认为，无视特定情境的终身监禁不仅过于苛刻，还异常死板，因此，持反对意见的人认为采用“三振出局”政策来减少噪声的成本太高。

以“[伍德森诉北卡罗来纳州案](#)”为例。在该案件中，美国最高法院裁定，强制性死刑有违美国宪法，不是因为死刑太残忍，而是因为法律就是这样规定的。强制判处死刑纯粹是为了确保判决免受噪声干扰。也就是说，在某些特定情境中，杀人犯必须被处死。美国最高法院说，“刑罚个别化”之所以有存在的必要性，是因为“不考虑罪犯的过往经历和生活习惯，就对同一法律范畴内的每项罪行都处以相同的惩罚，这样的观念已经不再盛行了”。美国最高法院还认为，强制性判处死刑存在一个

严重的制度性缺陷：它“没有将被定罪的人视为独特的个体，而是将其视为无差别的群体中的一员，从而盲目地判处了死刑”。

当然，死刑涉及的风险极高，关于司法的分析可以应用于许多其他情境，尽管这些情境大多数与法律无关。评估学生的老师、评估患者的医生、评估雇员的雇主、确定保费的核估员、评估运动员的教练，如果这些人都采用过分刻板的减少噪声的规则，他们很可能会犯错。比如，如果雇主使用简单的规则来评估员工，对其做出提拔或解雇的决定，这些规则可能确实消除了噪声，却忽略了员工绩效的一些重要的考核内容。一个不能考虑各种重要因素、不存在噪声的系统做出的判断可能比依赖个体做出的有噪声的判断更糟糕。

在第27章，我们将以“每个人都是独特的个体，而不是毫无差别的群体中的一员”这一观点为例，来说明以上问题。目前，我们先重点探讨一个相对有些乏味的内容。一些减少噪声的策略会导致太多的错误，就像上文提到的愚蠢的国际象棋游戏程序。

尽管如此，这种反对意见似乎比实际更具说服力。如果一种减少噪声的策略容易导致错误，那么我们要做的不应该是轻易向高噪声水平妥协，而是尝试找到更好的减少噪声的策略。例如，我们可以采用汇总判断的方法，而不是采用愚蠢的法则，或者，我们可以制定明智的而不是愚蠢的指南或规则。从减少噪声的角度，一所大学应只录取考试分数最高的一批人。如果这一规则看起来过于简化，那么学校可以创造一个公式来综合考量高考成绩、在校成绩、年龄、运动表现、家庭背景等。复杂的规则可能更准确，因为它整合了所有的相关因素。同理，医生可以采用复杂的规则对某些疾病进行诊断。专业人士使用的指南和规则并不总是简单或粗糙的，其中一些有助于减少噪声，且不至于产生让人难以忍受的高昂代价（或偏差）。如果指南或规则行不通，也许我们可以引入其他适用于特定情境的决策卫生策略，比如汇总判断，或是像中介评估法那样的结构化的流程。

算法无噪声，但会导致偏差

人们通常在算法的背景下讨论减少噪声的潜在高额成本，因此人们越来越反对有偏差的算法。我们已经看到，算法消除了噪声，因此看起来很有吸引力。事实上，本书的大部分内容都可以算作支持使用算法的证据，因为算法没有噪声。但是，如果使用算法会增加种族或性别歧视，抑或不利于弱势群体，那么减少噪声将得不偿失。

人们普遍担心算法会导致歧视——这无疑是一个很严重的风险。数学家凯茜·奥尼尔（Cathy O’Neil）在《算法霸权》（*Weapons of Math Destruction*）一书中主张：[依靠大数据和借助算法来做决策可能会产生偏见](#)、加剧不平等，甚至威胁民主本身。另一种质疑的声音表示：“[存在潜在偏差的数学模型正在重塑我们的生活](#)，而负责开发它们的公司和政府都无意去解决这一问题。”独立新闻调查机构ProPublica称，“替代性制裁的惩罚性罪犯管理分析”（Correctional Offender Management Profiling for Alternative Sanctions）这种广泛用于对惯犯进行风险评估的算法，对少数族裔成员有强烈的偏见。

没有人会怀疑，我们有可能甚至很容易创建出一种无噪声但带有种族主义、性别歧视或其他偏差的算法。如果一种算法直接根据被告的肤色决定是否准许他获得保释，那么其中就存在歧视，这在许多国家是不合法的。将求职者是否可能怀孕纳入算法，则是在歧视女性。在这些案例以及其他案例中，算法能够消除判断中不必要的变异，但也会引发令人难以接受的偏差。

原则上，我们应该能够设计出一种不考虑种族或性别的算法，实际上这也完全可以做到。但有一个更具挑战性且备受关注的问题，那就是即使一个算法并不直接使用性别或种族作为预测因素，也可能产生歧视，进而导致偏差。

我们说过，算法存在偏差有两个主要原因。首先，无论我们是否有

意设计，算法都可能使用与种族或性别高度相关的预测因素。例如，身高和体重与性别相关，人们成长和居住的地区与种族相关。

其次，不同的数据来源可能导致歧视。如果一个算法是基于存在偏差的数据集训练所得，那么它就会产生偏差。例如，用来预测犯罪的“[预测性警务](#)”算法，其目的是改善警力的配置。如果现有的犯罪数据反映了某些区域的警力过度部署，或对某些类型犯罪的报案量相对较高，那么由此产生的算法将加剧这种歧视或使其永久化。只要原有的训练数据有偏差，就有可能有意或无意地设计出一种存在歧视的算法。因此，即使算法没有直接考虑种族或性别，它也可能像人类一样存在偏差。[事实上，就这一点而言，算法可能更糟](#)：由于消除了噪声，它们可能比人类的判断包含更稳定的偏差。

对于许多人来说，要考虑的一个关键的现实因素是：一种算法是否会对相同的群体产生不同的影响。[究竟如何测试](#)这种不同的影响，以及如何确定算法中的歧视、偏差或公平性具体包含哪些因素，这些都是异常复杂难解的问题，也超出了本书的范围。

然而，我们之所以提出这个问题，是因为相对于人类判断，算法还是具有独特的优势的。对于新手而言，我们建议仔细对算法进行评估，以确保其中没有输入不被法律许可的变量，并检测是否存在令人反感的歧视。要让人类个体接受同样的审查非常困难，因为他们的判断往往是不透明的。人们有时会无意中以一种外部观察者（包括法律制度）无法轻易看穿的方式表现出歧视。因此，在某些方面，算法比人类更透明。

毫无疑问，我们需要关注无噪声但有偏差的算法的成本，就像我们需要考虑无噪声但有偏差的规则的成本一样。关键问题是，我们是否可以设计一种在多种重要指标上的表现都优于真实世界中的人类判断的算法，它更准确、噪声更少、没有歧视、非常公正。大量证据表明，在人类选出的多个判定标准组合方面，算法的表现都可以比人类更好。请注意，我们说的是可以，而不是一定。例如，正如第10章所述，在保

释决策方面，算法可能比人类法官更准确，也更少产生种族歧视。同样，简历筛选算法可以比人类简历筛选者挑选出更好以及更多样化的人才。

基于这些例子和更多其他的例子，我们可以自然而然地得出一个结论：在充满不确定性的世界中，尽管预测性算法不太可能做到完美，但与通常充满噪声和偏差的人类判断相比，算法远没有那么不完美。算法的优势在于有效性（好的算法几乎总是可以做出更好的预测）和更少的歧视（好的算法比人类判断的偏差更小）。如果算法比人类专家更少犯错，而我们直觉上还是偏爱人类判断，这时候就应该仔细审视一下自己对直觉的执念了。

我们的总体性结论很简单，并且可以延伸至算法之外的主题。确实，减少噪声的策略可能对应高昂的成本，但是在很多情况下，高昂的成本只是一个借口，而不是忍受不公平和噪声的充分理由。当然，减少噪声的努力本身也可能产生误差，并以偏差的形式表现出来。如果是这样，我们就面临一个很严重的问题，但解决方案不应该是放弃减少噪声的努力，而应该是提出更好的方案。

- 消除噪声

要付出代价，但很值得

- 要消除教育中的噪声，必须花很多钱。当给学生评分时，教师本身存在噪声，但即便如此，我们也不可能让5位教师对同一篇论文进行评分。
- 如果不是依靠人类的判断而是交由社交网络来决定——“不管具体情境如何，任何人都不可以使用某些单词”，这样做确实可以消除噪声，但也会造成很多的错误。治疗方法可能比疾病本身更糟糕。
- 有些规则和算法的确存在偏差，但是人也有偏差。我们应该问的是：“我们可以设计出既无噪声，偏差又少的算法吗？”
- 消除噪声可能要付出高昂的代价，但这种代价是值得的。噪声可能会导致极端的不公平。如果减少噪声的方法过于粗糙，比如我们制定的指南或规则过于死板或在不经意间产生了偏差，我们也不应该放弃，而应该再试一次。

第27章

尊严，人之为人的重要价值观

设想如下三个场景：你申请房产抵押，但被拒绝，其原因并非有人实际考察了你的情况，而是因为银行有严格的规定，认为你这种信用等级的人不能申请房产抵押；你的条件很出色，而且一家公司的面试官也觉得你很不错，但你的求职申请被拒绝了，原因是15年前你曾被判刑，而该公司明令禁止录用任何有犯罪前科的人；你被判有罪，但不能被保释，这并非由于你在法官面前进行了个别化听证后被认为有问题，而是因为有一种算法认为具有你身上某种特征的人被保释后潜逃的风险超过了阈值。

在上述案例的情境中，许多人会提出抗议。他们希望被视为一个独特的个体，他们希望有一个真实的人可以看到自己的特殊处境。他们可能会也可能不会意识到，因人而异的处理方式会产生噪声。但是，如果这种方式要以噪声为代价，他们会坚持认为，这种代价是值得的。用美国最高法院法官的话说，只要法官“没有将被定罪的人视为独特的个体，而是将其视为无差别的群体中的一员，从而盲目地判处了死刑”，人们就会产生不满（见第26章）。

许多人坚持进行个别化听证，摆脱他们所谓的规则暴政，从而获得一种被区别对待和被尊重的感觉。正当的程序这个已经属于日常生活一部分的概念，似乎在要求与具有自由裁量权的人类审判者进行面对面的交流，这样他们才能考虑多种因素。

在许多文化中，这种对每个案子逐个进行审判的观点具有深厚的道德根基，这种道德根基可以在政治、法律、神学甚至文学作品中找到。很容易看出，莎士比亚的《威尼斯商人》（*Merchant of Venice*）反对无噪声的规则，书中呼吁法律和人类判断中通常应该存在慈悲。书中鲍西

娅（Portia）的结束语⁽¹¹⁾是：

慈悲不是出于勉强

它是像甘霖一样从天上降下尘世

它不但给幸福于受施的人，也同样给幸福于施与的人

.....

它深藏于帝王的内心

是一种属于上帝的德性

执法的人倘能把慈悲调剂着公道

人间的权利就和上帝的神力没有差别

慈悲因为不受规则的约束，所以包含噪声。尽管如此，在许多情况下，在无数组织中都会出现类似鲍西娅的请求，并能引起人们的共鸣。一个员工可能正在寻求晋升，一个买房人可能正在申请贷款，一个学生可能正在申请大学。在这些案例中，决策者可能会拒绝采用减少噪声的策略，因为他们与鲍西娅一样企盼拥有受社会赞许的仁慈。[他们可能知道，自己的做法充满噪声](#)，但如果这么做可以让人们感受到被尊重和被理解，无论如何他们都会欣然接受。

某些减少噪声的策略没有遇到此类反对意见。如果做出决策的是三个人，而不只是一个人，那么人们仍然感觉自己获得了个别化听证的机会。指南也可以给决策者很大的自由裁量空间。另一些减少噪声的措施，比如刻板的规则，确实会缩减自由裁量的空间，也因此可能招致人们的反对，人们会认为，这些措施冒犯了他们的尊严。

他们是对的吗？当然，人们常常很在乎自己是否有机会申诉。有机会被人倾听无疑具有重要的人文价值，但是，如果申诉的机会会造成更多不公正、更高代价，那就不值得了。我们已经强调过，在招聘、大学

录取和医疗等情境中，某些减少噪声的策略可能确实比较粗糙，却可以防止各种形式的区别对待，虽然存在噪声，但能从总体上减少误差。正如我们所主张的，如果减少噪声的策略比较粗糙，那么最好的办法就是尝试提出更好的策略——一种能更好地整合各种相关变量的策略。如果这种更好的策略能够消除噪声并减少误差，那么它显然比区别对待更有优势，即使它会减少或消除被倾听的机会。

我们并不是说区别对待无关紧要，只是，如果它会导致各种可怕的后果，例如明显的不公平，那么代价就太高了。

不断变化的新的价值观

想象一家公共机构成功地消除了噪声。比如，一所大学对“不当行为”进行了界定，每位教师和学生都知道它包含什么以及不包含什么。又比如，一家大公司精准地界定了什么是“腐败”，公司中的每个人都知道什么是被允许的，什么是被禁止的。再想象一下，一家私人机构可以通过明确指出只招收某些专业的毕业生等手段来降低噪声。然而，如果一个组织的价值观发生了变化，那会如何呢？某些减少噪声的策略似乎没有将这一因素考虑在内，因而它们的灵活性可能会成为一个大问题，而这一问题又与人们对区别对待和尊严的关注紧密相关。

[美国宪法中有一项以令人费解著称的裁决，可以帮我们理解这一点](#)。这项裁决发生于1974年，针对的是一项学校系统中的严格规定：学校要求怀孕的女教师在距预产期还有5个月的时候就开始无薪休假。教师乔·卡罗尔·拉弗勒（Jo Carol LaFleur）反对称，学校的规定具有歧视性，而且5个月无薪休假并不合理。这一反对意见获得美国最高法院认可，但是法院没有提到性别歧视，也没有说5个月时间不公平。相反，法院反对该规定的理由是，学校没有给予拉弗勒表明她不需要停止工作的机会。下面是法院的原话：

对于任何一位女教师是否有能力继续工作，校方没有根据医生或学校董事会的意见做出因人而异的决策。学校的规则包含一个不容反驳的“生理上无法胜任”的推定，而且即使医学证据表明有些女性的身体状况可能恰恰与其相反，该推定却仍被认为适用。

5个月的强制性休假看起来很荒谬，但法院并没有强调这一点。相反，法院控诉的是“不容反驳的推定”，以及缺乏“因人而异的决策”。如此说来，法院显然在用鲍西娅的方式争辩：慈悲不是出于勉强，应该指定专门的人来研究拉弗勒的特殊情况。

由于没有使用任何决策卫生策略，这样的做法会产生噪声。那么，谁有权判定拉弗勒的案件？许多其他处境相似的女性，她们得到的判定结果与拉弗勒的一致吗？在任何案件中，许多规则都是不容反驳的推定。某个限速规定能否被接受？有权投票或饮酒的最低年龄应该是多少？严格禁止酒后驾驶的规定又如何呢？考虑到这些例子，批评者们反对说：反对“不容反驳的推定”的论点过于牵强了，但他们反对的理由至少不是“这些推定的目的和实际效果是减少噪声”。

[当时有影响力的评论员为法院的裁决辩护](#)，他们强调说：人们的道德观念会随着时间的推移而发生变化，因此应该避免使用刻板的规则。他们认为，考虑到女性在当今社会所起的作用，相关的社会规范也应该顺应时代的变化。因人而异的决策在这种情况下特别合适，因为它能将不断变化的道德规范纳入其中。有规则约束的系统可能可以消除噪声，这很好，但如果固化了现有的规范和价值观，那就不太好了。

总而言之，有些人可能会坚持认为：有噪声的系统，其优势在于帮助人们适应新的价值观。例如，随着价值观的改变，如果允许法官行使自由裁量权，他们就可以根据当时的价值观，对贩卖大麻的罪行处以较轻的刑罚，而对强奸罪处以较重的刑罚。我们想要强调的是，如果有些

法官很宽容，而另一些法官并非如此，那么就会产生一定程度的不公平，即处境相似的人将受到不同的对待。但是，如果考虑新出现的社会价值观，那么不公平将更有可能被大众接受。

以上问题肯定不仅限于刑事司法制度或司法领域。一些公司可能允许判断和决策具有一定的灵活性，虽然这样会产生噪声，但可以确保公司政策能够随着新的观念和价值观的出现而做出相应改变。举一个我们亲身经历的例子：本书的一位合著者在几年前曾加入一家大型咨询公司，他所收到的不太具有时效性的“欢迎礼包”中，说明了他可以报销哪些旅行费用，包括安全抵达时给家人打电话的费用、熨西装的费用、服务生的小费等。这些规则是没有噪声的，但显然已经过时了，并且存在性别歧视。很快，这些过时的规则就被随着时代的发展而出现的新规则所取代。现在该公司对报销的要求是：必须“适当且合理”。

这一要求会产生噪声但依然得以采用的理由很简单，因为某些减少噪声的策略根本没有遇到反对意见。如果人们使用一个基于外部视角的通用量表，那么他们就可以随时对变化的价值观做出反应。在任何情况下，减少噪声既不需要也不应该是一劳永逸的。如果通过制定严格的规则来减少噪声，那么规则制定者应随时间的变化对规则做出改变。或许，他们可以每年对已有的规定进行一次重新审视。另外，随着新的价值观的出现，他们需要决定是否有必要制定新的规则。在刑事司法制度中，规则制定者需要降低对某些罪行的刑期要求，而增加对另一些罪行的刑期要求。他们甚至可能需要认定一些行为不再属于犯罪，而将一些在以往完全可被接受的行为认定为犯罪。

我们来一起回顾一下。有噪声的系统可以为新的道德观念留出余地，这可能是好事，但是在许多领域中，用这种观点来为高水平的噪声进行辩护就很荒谬。一些最重要的减少噪声的策略实际上已经融合了价值观的影响，比如聚合判断。而且，如果一家计算机公司对顾客投诉的笔记本故障问题予以差别对待，这种不一致性不太可能是新出现的价值

观所导致的。如果患有相同疾病的人得到了不同的医学诊断，也同样不可能是新的道德观念所致。总之，我们可以采取很多策略来减少噪声，甚至消除噪声，与此同时，我们也可以通过设计一些程序来适应价值观的变化。

利用制度的漏洞，逃避规则

在充满噪声的系统中，判断者们可以根据情境的要求做出调整，并对意料之外的事件做出反应。通过消除这种适应性，某些减少噪声的策略可能会产生意料之外的后果——激励人们去寻找可乘之机。容忍噪声的一个潜在理由是，噪声可能是私人机构和公共机构用来防止人们寻找可乘之机的副产品。

免税代码是我们很熟悉的一个例子。一方面，税收制度不应该包含噪声，它应该是清晰和可预测的，同一类的纳税人不应被区别对待。但是，如果我们消除税收制度中的噪声，“聪明”的纳税人就可能会找到逃避规则的方法。是制定明确的规则来消除噪声好，还是留有一定程度的模糊性好？税务专家对此一定争论不休，后一种情况具有不可预测性，但可以减少明确的法规可能导致的投机取巧和自利行为。

一些公司和大学禁止人们做“坏事”，却没有说明“坏事”具体指什么。这样一来，噪声将难以避免，但如果有一个具体的清单明确列出了什么是坏事，那么那些没有明确包含在清单中的可怕行为就会被容忍。

规则具有明确的边界，因此人们可能会绕过规则去实行一些严格意义上说不违反规则但会造成相同或类似危害的行为。青少年的父母都清楚这一点！如果不能制定出能够禁止所有不良行为的规则，我们就只能容忍噪声，因为不这样做会招致反对。

在某些情况下，制定明确、没有噪声的规则确实会导致有人寻找可乘之机的风险，而这种风险可能成为采用其他减少噪声策略的理由，如

汇总策略或容忍有噪声的方法等策略。关键在于：我们需要明确有多大的可乘之机，以及有多少噪声。如果可乘之机很小，但噪声很多，那么我们最好还是想办法减少噪声，第28章会继续讨论这个问题。

威慑与风险规避

假如我们的目标是对雇员、学生、普通市民的不良行为进行威慑，那么有一点不可预测性甚至很高的不可预测性，都可能并不会很糟。比如，雇主可能会想：“如果做出不良行为的后果可能是被罚款、被停职，甚至是被解雇，那么员工可能就不会这么做了。”那些刑事司法制度的管理者可能会想：“我们不用去管那些潜在的罪犯是否会猜测可能有什么样的惩罚。只要一想到存在各种不确定的、类似抽签的惩罚后果，人们就能够尽可能不越界，那么由此产生的噪声是可以容忍的。”

简而言之，这种论点难以反驳，但并不是很有说服力。乍一看，惩罚的预期值是最关键的。从预期值的角度来看，50%的概率受到5000美元罚款与肯定会受到2500美元罚款是等价的。当然，有些人可能会关注最坏的情况，风险规避者可能会被有50%的可能被处5000美元罚款所吓倒，但风险爱好者就没有那么容易被吓倒。要知道一个存在噪声的系统能否产生威慑力，我们首先需要知道，潜在的违法者是风险规避者还是风险爱好者。而且，如果我们想增加威慑力，最好的办法难道不是同时增加罚金和消除噪声吗？这样做还可以消除不公平现象。

创造力、士气和新颖的创意

减少噪声的策略会抑制人们做某件事的动机或降低其参与度吗？它会影响人们的创造性或妨碍人们取得重大突破吗？许多组织认为是的。在某些情况下，他们可能是对的。要知道是否确实如此，我们需要先搞清楚这些减少噪声的策略面向的对象是什么。

[回想一下，我们在前文中论述过的一些法官对量刑指南的强烈反对](#)，正如一位法官所说：“我们需要再一次学习如何信任法庭做出的判决。”通常，身居要职的人不喜欢别人剥夺他们的自由裁量权，因为他们会觉得受到了限制和束缚，甚至羞辱。当采取措施减少自由裁量权时，一些人会反抗，他们重视甚至珍视自己做出判断的机会。如果剥夺了他们的自由裁量权，让他们去做任何人都能做的事，他们就会感觉自己像是机器中的齿轮。

简而言之，有噪声的系统可能对保持士气有好处，但这并不是因为它包含噪声，而是因为它允许人们根据自己的意愿做出决定。如果允许员工以自己的方式回应客户的投诉，以他们认为最好的方式评估下属，或者确定他们认为合适的保费，那么他们可能会更喜欢自己的工作。如果公司采取措施来消除噪声，员工可能会认为这样会降低自己的主观能动性，因为这样一来，他们就只能按照规则行事，而不是发挥自己的创造性，他们的工作看起来会更机械，他们感觉自己甚至像个机器人。谁会愿意在一个限制自己独立做决策的地方工作呢？

组织之所以会照顾员工的这种感受，不仅是因为尊重员工，更是希望给员工留出能够提出新想法的空间。如果制定了规则，就可能会限制员工的灵感和创造力。

在组织中，很多人会持这样的观点，但并非所有人都是如此。需要对不同的任务进行不同的评估，例如，对脓毒性咽喉炎和高血压的诊断就不适合发挥创造性。但是，如果可以使员工更快乐和更有灵感，我们就更愿意忍受噪声，因为士气低落本身就是代价，而士气低落又会导致其他代价，例如糟糕的绩效。当然，我们应该能够在减少噪声的同时乐于接受新想法。某些减少噪声的策略，例如对重复判断进行建构，就可以做到这一点。如果我们想要在减少噪声的同时保持团队的士气，就可以选择具有这种效果的决策卫生策略。公司负责人应该明白，如果有严格的制度和规则，就应存在质疑和重新思考这些规则的程序，而不

是通过自由裁量权来破坏这些规则。

[在一系列颇具影响力的作品](#)中，杰出的律师和思想家菲利普·霍华德（Philip Howard）提出了类似的观点，他主张允许人们做出灵活的判断。霍华德希望不要将政策表述为法定规则的形式（这样可以消除噪声），而是将其表述为“要合理”“谨慎行事”“不会带来额外的风险”等一般性原则。

在霍华德看来，现代政府管理制度已经很不明智了，因为它过于呆板。教师、农民、开发人员、护士、医生以及其他专家都为规则所累，这些规则告诉他们要做什么以及如何具体去做。霍华德认为，应该允许人们利用自己的创造性来思考如何达成目标，无论这个目标是达成更好的教育成果、减少交通事故、获得更干净的水，还是促进患者的健康。

霍华德提出了一些很有吸引力的观点，但重要的是要知道他所提倡的方法会带来怎样的后果，包括是否会增加潜在的噪声和偏差。大多数人不喜欢抽象且严格的规定，但这可能是减少噪声和偏差的最佳方法。如果只是一般性原则，那么在解读和执行这些原则的过程中难免会产生噪声。这种噪声可能令人难以忍受，甚至会让我们颜面扫地。至少，需要仔细考虑噪声的代价——通常很少有人真的会去考虑。一旦我们意识到噪声会产生普遍的不公平以及很高的代价，我们就会得出结论：噪声是难以接受的，我们需要找到不损害重要价值观的减少噪声的策略。

- 消除噪声

人的尊严是无价的

- 人们很重视面对面互动，同时也很需要这种互动。他们希望有一个真实的人能够倾听他们的担忧和抱怨，并有能力让事情变得更好。这些互动不可避免地会产生噪声，然而人的尊严是无价的。
- 道德观念在不断发展。如果锁定所有内容，我们就没有给不断变化的价值观留出余地。一些减少噪声的措施太严格了，它们会阻碍道德观念的发展。
- 想要防止不当行为，就要容忍一些噪声的存在。如果学生们不确定抄袭会受到什么样的惩罚，那是好事，因为这样一来他们会避免抄袭。少量的不确定性（以噪声的形式表现出来）会增加威慑力。
- 如果消除噪声，我们就不得不提出明确的规则，这样违法者就有机可乘。如果噪声是防止投机取巧行为而不得不付出的代价，那么它是值得的。
- 有创意的人是需要空间的。人类不是机器人，不管你从事什么工作，都需要一定的回旋余地。如果被规则束缚住，你可能不会产生噪声，但你也不会有太多的乐趣，同时也无法提出有创造性的想法。
- 到头来，大多数为噪声辩护的理由都难以令人信服。我们可以尊重人的尊严，为道德发展留出足够的空间，并让人类的创造性得以展现，且无须忍受噪声导致的不公平及其高昂的代价。

第28章

规则还是标准

如果我们的目标是减少噪声和了解如何减少噪声（或在多大程度上减少噪声），那么我们就有必要区分两种约束行为的方式：规则 and 标准。很多组织通常会选择其中的一种，或将两者结合起来使用。

在商业领域，某家公司可能会有以下要求：员工必须在指定的时间内工作；每个人的休假时间不能超过两个星期；如果有人将公司的秘密泄露给媒体，那么他就会被解雇。我们也可以换一种表达：公司要求员工必须在“合理的工作日内”工作；能否休假取决于“公司的需要，并且视个人情况而定”；泄露机密“将会受到相应的惩罚”。

在法律上，一个规则的内容可能是：任何人不得超速；不得让工人接触致癌物；所有处方药必须贴有警示标识。相比之下，标准的内容可能是：人们必须“谨慎驾驶”；雇主必须“在可行的范围内”提供安全的工作场所；公司必须“合理地”决定是否要给处方药贴上警示标识。

上述示例说明了规则 and 标准之间的主要区别。规则旨在消除实施者的自由裁量权，而标准则会授予实施者一定的自由裁量权。凡是规则，都要严格减少噪声。解释规则的人需要先回答一个事实问题：驾驶员开车的速度是多少？工人是否接触了致癌物？药物是否贴上了必要的提醒标志？

在规则的约束下，进行事实调查的企业可能本身就会涉及判断，因此会产生噪声，或者受到偏差的影响。从我们遇到过的一些案例来看，人们会通过制定规则来降低这些风险，并且当规则包含某个数字时，噪声就会减少，比如“18岁之前，任何人没有投票权”或“限速105千米/小时”。规则具有一个重要的特征：它们减少了判断的作用，从这一点

来看，包括法官在内的所有应用规则的人，就都没有多少工作可做了，他们只需按规则办事。无论规则好坏，他们的回旋余地少了很多。

标准就完全不同了。当使用标准时，法官需要花大力气对一些开放性的法律条款进行解读，他们可能需要做出大量判断才能确定什么才算是“合理”和“可行”。除了查明事实以外，他们还需要对模糊的术语进行解读。也就是说，那些标准的制定者实际上将决策权转交给了他人，他们让渡了权力。

第22章中讨论的指南可能是规则，也可能是标准。如果是规则，它们会极大地限制判断，即使是标准，也不可能是完全开放的。阿普加分数是指南，而非规则，这样的指南并不禁止行使自由裁量权。当指南收紧到可以消除自由裁量权时，它们就会变成规则。算法是规则，而不是标准。

分裂与无知：标准引发噪声，而规则不会

我们一开始就需要明白，只要公司、组织、社会或群体内部存在巨大分歧，制定标准就比制定规则容易得多。在不了解某项禁令的具体含义时，公司领导大多会赞成“管理者不应滥用职权”。比如，在没有确定“调情”是否被允许的情况下，管理者会反对工作场所中的性骚扰；在没有明确“抄袭”一词的确切含义之前，大学会同意“禁止学生抄袭”；在没有明确是否应保护商业广告、威胁言论或猥亵言语的条件下，人们会同意“宪法应该保护言论自由”；在没有对“审慎”一词进行清楚界定之前，人们大多会赞成环境监管机构发布“审慎的规则”以减少温室气体排放。

制定标准但未明确细节就可能会产生噪声。这些噪声可以通过我们之前讨论过的一些策略进行控制，例如汇总判断和使用中介评估法。领导们可能想提出规则，但在实践过程中可能无法达成一致，于是就形成

了标准。宪法本身也包括许多标准，如“保护宗教信仰自由”。《世界人权宣言》也是如此规定：“人人生而自由，在尊严和权利上一律平等。”

让不同的人就减少噪声的规则达成共识非常困难，而这恰恰是需要制定标准而非规则的原因之一。公司的领导者可能无法就员工应如何与客户打交道的具体措辞达成一致，此时，制定标准可能是领导们最好的选择。在公共部门也有类似的例子，立法者可能在标准上达成妥协，并且如果容忍噪声是确保法律得以实施必须付出的代价，那么他们也能容忍由此导致的噪声。在医学上，医生可能会在疾病诊断标准上达成共识，而如果试图制定规则，则可能引起严重的分歧。

但是，社会和政治分歧并不是人们诉诸标准而非规则的唯一原因，有时，真正的问题在于人们缺乏能够制定合理规则的信息。比如，一所大学可能无法制定具体的规则来决定是否要提拔一名教师。雇主很难预见应该会导致留任或惩罚员工的所有可能情况。国家立法机关或许无法得知可悬浮颗粒物、臭氧、二氧化氮、铅等空气污染物的正常含量水平是多少。他们能做的就是发布某种标准，并依靠可信赖的专家来明确其具体含义，虽然这样做可能会产生噪声。

虽然规则可能导致多种偏差，但规则可以大大减少噪声，当然前提是每个人都遵守。如果某条规则规定年满21周岁的人可以购买酒精饮料，而不满21周岁则不能购买，那么只要人们遵守这一规则，几乎就不存在噪声。相反，标准却会引发噪声。

老板与被控制的下属

规则 and 标准之间的区别对于所有公共机构和包括各类企业在内的私人机构都非常重要，每当委托人试图控制代理人时，就需要在规则 and 标准之间做出选择。如第2章所述，保险公司核保员需要按“金发姑娘价格”收取不高也不低的保费，公司才能受益。老板会给这些核保员标准

或规则来指导他们吗？公司的任何领导给出的指导意见可能很具体，也可能很笼统，比如“这是常识”或“做出最佳判断”。此外，医生也会使用这样或那样的方式要求患者遵医嘱，“早晚各一颗”是规则，“每当你感觉需要时就服药”则是标准。

我们已经注意到，脸书等社交媒体公司必然会关注噪声以及如何减少噪声。如果某位用户上传的内容违反了明确的规则（如禁止裸露）时，该公司可能会要求员工将其删除。该公司也可能要求员工执行某项标准（例如，禁止发布带有霸凌或明显攻击性的内容）。脸书的“社区标准”首次发布于2018年，是规则和标准的有趣结合，社区标准发布后，用户的投诉不计其数，他们说该公司的标准产生了太多的噪声，可想而知，也造成了误差和不公平。一个受到普遍关注的问题是，成千上万的脸书审阅者要做出判断，而他们做出的决策千差万别。在决定是否删除帖子时，审阅者对允许和禁止的内容做出了不同的决策。要了解这种变异性为何不可避免，可以看看[脸书在2020年社区标准中规定的这些内容](#)：

仇恨言论指依据种族、民族、祖籍国、宗教信仰、性取向、种姓、性别、性别认同以及严重疾病或残障等应受保护的特征进行人身攻击。我们也为移民身份提供一些保护。我们对攻击的定义包括暴力、非人化的言论、贬抑的言论、宣扬排斥或孤立的言论。

在应用这类定义时，审阅者必然会产生噪声。究竟什么才算是“暴力或非人化的言论”？脸书也意识到了这一问题，并针对这一问题制定了明确的规则，从而有效地减少了噪声。这些规则被编入名为“实施标准”的非公开性文件中，这份文件约含12 000个单词，被[《纽约客》杂志](#)收录。在公开的社区标准中，管理图片内容的说明以这样一条标准开头：“我们会删除崇尚暴力的内容。”但确切地说，这是什么意思呢？相

比之下，实施标准则列出了具体图片，并明确告知内容版主应该如何处理这些图片，例如“烧焦或正在燃烧的人体”“肢解的人体部位”。说了这么多，我们来做个总结：社区标准看起来更像标准，而实施标准看起来更像规则。

同样，航空公司可能会要求飞行员遵守一系列规则或标准，如在停机坪上等候90分钟后是否要返回登机口，以及何时应打开安全带指示灯。航空公司可能喜欢规则，因为它们会通过限制飞行员的自由裁量权来减少失误，但在某些情况下，飞行员应该相信自己的判断。在这些情境中，即使标准会产生一些噪声，但它可能比规则更合适。

在上述案例以及更多其他案例中，在规则 and 标准之间做出决定的人必须考虑噪声问题、偏差问题或同时考虑这两个问题。对于企业来说，无论其规模大小，都经常需要做出这样的决策。有时，他们这样做完全是凭借直觉，而没有制定相应的框架。

标准有多种形式和尺度，但它们基本上不会包含这样的内容：在这种情况下，做出合适的行为。它们可以被表述得接近规则。例如，当明确界定了什么是“合适”以及限制了自由裁量权时，标准就类似于规则。规则 and 标准也可以相互配合使用，例如，人事部门可以在规则（“所有应聘者必须具有大学学历”）之后，附上标准（“在满足上述条件的情况下，选择最合适的人”）。

我们已经说过，规则应该减少甚至消除噪声，而标准通常会产生大量的噪声，除非采用某些减少噪声的策略。在私人组织和公共组织中，噪声通常是没有制定好规则的结果。如果噪声很大，比如每个人都可以看到处境相似的人没有得到同等的对待，此时就需要规则了。与刑事量刑一样，这一转变可能会引起强烈的抗议，而在强烈抗议之前通常先进行了噪声审查。

被“限制”的噪声

考虑一个重要的问题：什么样的人可以算作残疾人，有资格享受残疾人才能享受的经济福利？如果问题的描述是这样的，法官将做出充满噪声和不公平的临时决策。在美国，这种充满噪声且不公平的决策曾经是常态，其结果也是令人难以接受的。两个看似处于相同处境的人——都坐在轮椅上，或是同样遭受严重抑郁或慢性疼痛，却受到了不同的对待。作为回应，政府官员采用了更像规则的东西——残疾矩阵。该矩阵要求决策者根据受教育程度、地理位置、剩余的身体机能来做出相对机械性的判断，其目的是减少决策中的噪声。

在法学教授杰里·马肖（Jerry Mashaw）撰写的关于该问题的重要讨论文章中，他为这种消除噪声的判断起了一个名字：[官僚正义](#)（bureaucratic justice）。我们应该记住这个词。马肖非常支持创建这个矩阵，主要是因为这样有望消除噪声，实现最基本的公正。但在一些情况下，官僚正义的承诺可能难以实现。每当一个机构做出不受规则约束的决策时，噪声就有可能重新出现。

假如规则在特定情况下产生可怕的后果，法官可能会认为规则太苛刻，进而无视规则，因此，法官们可能会因为公民轻微的反抗而行使自由裁量权，这种情况很难被监督或发现。在私人企业，员工会忽略看似愚蠢的公司规则。同样，在法规过于严苛和过分形式化的情况下，负责保护公共安全和健康的行政机构也可能拒绝执行。在刑法中，陪审团否决权就是指，陪审团如果认为法律条文太过严格和苛刻，则可以拒绝遵守。

每当公共机构或私人机构试图通过制定严格的规则来控制噪声时，就需要随时警惕人们在背地里行使自由裁量权的可能。在实行“三振出局”政策时，为避免对已经有两项罪名成立的人进行重罪指控，检察官的惯常反应极难控制，甚至难以发现。

当这种情况发生时，就会有噪声，但是没人能够“听到”。我们需要监控规则的实施情况，以确保它们被按照预期的方式执行。如若不然，

噪声的存在就是一种提示，也说明规则需要修改了。

确保规则“足够”准确

在企业和政府机构中，人们通常凭直觉对规则 and 标准进行选择，但这个选择的过程可以更科学一些。简单来说，这个选择仅取决于两个因素：决策成本和误差成本。

对于各类决策者来说，使用标准进行决策的成本可能非常高，因为他们必须努力制定准则的具体内容。有时候，进行判断可能很麻烦，比如，如果要求医生做出最好的判断，他们可能不得不花时间考虑每一种情况，并且这些判断可能充满噪声。如果给医生明确的指南，使其据此去诊断患者是否患有脓毒性咽喉炎，那么他们的判断可能快速而直接。如果限速每小时105千米，交通警察就不必费力去思考应该允许人们开多快。但是，如果标准的内容是人们“不得以不合理的速度行驶”，那么交通警察就不得不考虑很多了，几乎可以肯定的是，执法过程一定会出现噪声。使用规则，决策的成本通常会低得多。

不过，情况还是很复杂。规则一旦确定了就可以直接应用，但在规则实施之前，需要有人确定规则具体是什么。制定规则很困难。有时，它的成本很高。因此，法律体系和私人公司经常使用“合理”“审慎”“可行”这样的词。这也是为什么此类术语在医学和工程学等领域中也起着同等重要的作用。

误差成本是指错误的数量和严重性。一个普遍的问题是代理人是否知识渊博且可靠，以及他们是否应用了决策卫生策略。若的确如此，那么一个标准就可以完美地解决问题了，而且噪声可能很少。委托人有理由不信任代理人，此时委托人必须制定规则。如果代理人不称职或有偏误，并且他们无法切实执行决策卫生策略，那么他们应该受到规则的约束。明智的组织深知，他们授予自由裁量权的程度与其对代理人的信任

程度密切相关。

当然，从完全信任到完全不信任是一个连续的过程。一个不可信的代理人可能会导致许多误差，但是如果产生的都是些小的误差，那么情况还是可以容忍的。一项规则可能只会导致很少的错误，但是如果这些错误是灾难性的，那么我们可能还需要一个标准。我们需要明白，没有普遍的理由认为，使用规则或标准会增加误差成本，当然，如果一条规则是完美的，它将不会产生任何误差，但是几乎没有哪条规则是完美的。

假设法律规定只有年满21岁的人才能购买酒精饮料是旨在保护年轻人免于承受与饮酒有关的各种风险，按照这种方式理解，法律将产生大量错误，比如，有些20岁、19岁、18岁，甚至17岁的人饮酒不会产生问题，而有些22岁、42岁或62岁的人却不能饮酒。一项标准可能产生更少的误差——只要我们能够找到合适的措辞，并且人们能够准确地使用这些措辞。当然，这很难做到，这就是为什么我们总能在酒精饮料的包装上，看到根据年龄制定的一些简单规则。

这个例子说明了一个更大的问题。每当需要做出许多决策时，很多噪声就可能产生，此时尤其需要制定明确的规则。当看到多位皮疹发痒和长痣的患者时，如果医生能依据明确的规则来做诊断，就很少会出现误诊，而如果没有这种规则，并采用开放式标准，决策成本往往会高得离谱。对于重复决策，机械性规则往往比临时判断更具有真正的优势。行使自由裁量权会产生很重的负担，而且由此产生的噪声成本或不公平可能令人难以接受。

精明的组织机构敏锐地注意到两种行为调节方式都存在弊端，他们会制定规则或接近规则的标准来减少噪声和偏差。为了最大限度地减少错误成本，他们愿意提前投入大量时间和精力来确保规则足够准确。

每一个领域都要“禁止噪声”？

在一些情况下，允许噪声的存在是不光彩的。人们有时会忍受噪声的存在，但不应该如此。一个简单的解决办法就是将开放式的自由裁量权或模糊的标准转变为规则或接近规则的东西。我们现在能意识到，有时候简单的做法就是正确的做法。然而，即使规则不可行或不太合适，我们也已经找出了各种能减少噪声的策略来解决噪声问题。

这些情况引出了一个大问题：法律系统应该宣布存在噪声是不合法的 吗？如果回答“是”，那答案就太过简单了，但是除了控制噪声，法律需要做的还有很多。这里有一种考虑问题的方法。德国社会学家马克斯·韦伯（Max Weber）控诉过“卡迪司法”（Kadi Justice），他将其理解为非正式的、临时的、不受一般规则约束的裁决。在韦伯看来，卡迪司法太过于就事论事了，它违反了法律规则。正如韦伯所说，法官“确实没有遵循正式的规则，也‘没有考虑人本身’就做出了判决，[恰恰相反，法官应是根据每个人的性格和具体情况，或根据公正性和具体结果的适当性来评判人的](#)”。

韦伯认为，这种方法“不是合理的决策规则”。我们很容易看出韦伯在控诉卡迪司法包含令人无法接受的噪声。韦伯认为，事前制定好的官僚主义的判断存在风险是好事。他将专门化的、专业的和有规则约束的方法看成是法律发展的最终阶段。但在韦伯提出以上论断之后的很长时间内，卡迪司法或类似的事物仍然无处不在。那么问题来了，我们应该如何处理这个问题呢？

我们并不是说减少噪声应该成为《世界人权宣言》的一部分，但是在某些情况下，噪声可以被视为侵犯人权的事物。从总体上说，全世界的法律体系都应该尽更大的努力来控制噪声，在刑事量刑、民事罚款，以及对庇护、教育援助、签证、建筑许可证或职业许可证的批准或拒绝中皆是如此。比如，一个大型政府机构正在招聘数百甚至数千人，但它的录用决定没有什么规律或理由可循，这其中就充斥着巨大的噪声。再比如，一个儿童监护机构是指派这名还是那名雇员来处理一起案件，可

能使不同幼童得到的待遇存在天壤之别。孩子的未来乃至一生就如同买彩票一般，这能让人接受吗？

在许多情况下，此类决策中的变异性是由偏差导致的，包括可识别的认知偏差以及某些形式的歧视。此时，人们往往会觉得这种情况让人无法容忍，于是引用法律作为一种纠正措施，要求采取新的、不同的做法来解决这些问题。全世界的组织都将偏见视为不好的东西，他们是对的；但他们没有将噪声视为不好的东西，这样是不对的。

在许多地区，噪声水平已经过高，并造成了高昂的成本以及严重的不公平。目前我们所说的这些只是冰山一角。为了降低成本，与不公平做斗争，法律应该做到更多。

- 消除噪声

规则还是标准，需视情况而定

- 规则简化了生活，减少了噪声，而标准允许人们视具体情况而定。
- 是选择规则还是标准？首先，弄清楚哪个会导致更多的错误；然后，弄清楚用哪个更简单、用哪个更麻烦。
- 当我们应该采用规则时，我们经常使用标准——仅仅因为我们没有注意到噪声的存在。
- 减少噪声不应该成为《世界人权宣言》的一部分，至少现在还不是时候，然而，噪声可能造成可怕的不公平。在世界各地，法律体系都应考虑采取强有力的措施来减少噪声。

正视噪声问题

噪声是我们在做判断时不希望存在的变异，然而噪声实在太多了，本书的核心目标就在于解释为何会如此，以及我们应该如何应对噪声。本书涵盖的内容十分丰富，我们在这里仅从一个宽泛的视角来对书中的核心观点进行简要的回顾和总结。

判断不是计算，也无须遵循精确的规则

需要注意的是，不要把本书中的“判断”一词与“思考”一词混淆。判断是一个相对狭义的概念，指的是以人的大脑为工具的测量。与其他测量方式一样，判断需要为一个对象赋值，不过这个值未必是数字。例如“玛丽的肿瘤可能是良性的”是判断，“国民经济不稳定”“弗雷德是我们新任经理的最佳人选”“这种程度的风险所对应的保费应该是12 000美元”等，这些也都是判断。判断是将各种信息非正式地整合到总体评估中，但判断不是计算，也无须遵循精确的规则。教师可以用判断来对论文进行评分，但不能用判断做多项选择题。

很多人以专业判断为生，每个人都可能在一些很重要的方面受到这些专业判断的影响，我们所说的专业判断者，包括足球教练、心脏病专家、律师、工程师、好莱坞高管和核保员等。专业判断是贯穿本书始终的重点，这不仅是因为它们已经被广泛研究，还因为这些判断的品质会对我们所有人产生巨大影响。此外，我们相信我们所学到的知识也同样适用于生活中其他方面的判断。

有些判断具有预测性，并且有些预测性判断是可验证的，即我们最

终会知道它们是否准确，短期预测大都属于此类，例如药物的疗效、流行病的传播过程以及选举的结果等。但是，有些判断是无法验证的，如长时间跨度的预测和对虚拟问题的回答，对于这些判断，我们只能通过评估产生这些判断的思维过程来评估其品质。此外，还有一些判断并非可预测的，而是可评估的，举例来说，法官的判决或一幅画在有奖竞赛中的排名，就不太容易与其客观的公正判决或这幅画的真实价值进行比较。

然而，令人惊讶的是，无论是否存在真实值，判断者都会表现得像存在真实值一样，他们的思考及行为就像是要瞄准一个看不见的靶心，并且希望不要偏离靶心太多。“见仁见智的判断”一词表明，判断可能存在分歧，但分歧是有限度的。判断的一个重要特征是存在有限的分歧。判断介于计算和独特的个人品位之间，计算意味着不允许存在任何分歧，而个人品位则意味着，除非处在极端情况中，人们完全不期望能够达成一致。

人类判断的错误：偏差和噪声

当一组判断中的大部分错误都指向同一个方向时，我们就认为这组判断出现了偏差，偏差即平均误差。例如，整队射击手连续命中靶子的左下方；公司高管年复一年地对销售额做出过高的估计；公司对本该撤销的失败项目持续进行投资。这些都是偏差。

消除一系列判断中的偏差并不能消除所有误差，消除偏差后仍然残留的误差缺少共性⁽¹²⁾。它们是我们做判断的过程中不希望存在的分歧，体现了我们将测量工具应用于实际时的不稳定性。这种变异就是噪声。噪声是本该相同的判断中出现的变异。我们用“系统噪声”这一术语来描述组织中具有同质性的专业人士，如急诊医生、量刑法官以及保险公司核保员在做决策时出现的噪声。本书的大部分内容都在讨论系统噪声的问题。

测量偏差和噪声

均方误差（MSE）是科学测量学中已经沿用了近200年的测量准确性的标准。均方误差的主要特征是：它将样本均值作为总体均值的无偏估计，同等对待正误差和负误差，并且不成比例地处理较大的误差，因此，均方误差不能反映判断误差的实际代价——误差的实际代价往往是非对称的。然而，专业决策往往需要做出准确的预测。对于一个即将面临飓风袭击的城市，低估和高估飓风威胁所需付出的代价显然是非对称的，但你不希望这些代价影响气象学家对风暴速度和轨迹的预测。对于此类以追求客观准确性为目的的预测性判断来说，均方误差是合适的判断标准。

偏差和噪声都可以用均方误差来独立测量，而它们都是误差的来源。显然，偏差通常是有害的，减小偏差总是能提高判断的准确性；而噪声同样有害，减少噪声也同样总是能提升判断力。但是从直觉上说，这一事实却不那么容易被大家接受，人们通常希望，即使判断中存在明显偏差，离散度也最好为0。当然，我们最终的目标是同时使偏差和噪声最小化。

一系列可验证的判断中的偏差指的是案例的平均判断与其真实值之间的差异。而对于无法验证的判断，则无法进行这种比较。例如，承保人为特定风险设定的保费，其真实值永远不可知。对于某项特定罪行来说，我们也无法轻易知道公正判决的真实值是什么。在缺乏真实值的情况下，一个最常用、最方便（虽然并非总是正确）的假设是：判断是不偏不倚的，多名法官的平均值就是对真实值的最佳估计。

我们可以通过噪声审查来评估系统中的噪声。在噪声审查过程中，几名专业人员对同一案例（真实的或虚构的）做出独立判断，我们可以在不知道真实值的情况下测量噪声，就像我们从靶子背面看到一堆弹孔的分布图一样。噪声审查可以衡量许多系统中判断的变异性，包括放射

科和刑事司法制度，有时还会让人们关注到技能或训练的不足。此外，噪声审查还可以对系统噪声进行量化，例如同一团队中的核保员对风险的评估有所不同的情形。

偏差和噪声哪个问题更大呢？这要依据实际情况而定，答案很可能是噪声。当误差的平均值（偏差）与误差的标准差（噪声）相等时，偏差和噪声对总体误差的贡献相同。如果判断呈正态分布，即标准钟形曲线，那么只有当84%的判断都高于或低于真实值时，偏差和噪声的影响才相等。这其实是很大的偏差，这么大的偏差很容易就能被专业人员检测到。而当偏差小于一个标准差时，噪声就是总误差中更大的错误来源了。

噪声是个问题，远超我们的想象

分歧在某些判断场景中是没有问题的，甚至是有必要的。观点的多样性对于激发创意和产生多种设想必不可少，标新立异的思维对于创新也至关重要。比如，电影评论家多元化的观点是优势而非劣势；交易员之间的分歧促进了市场的繁荣；相互竞争的创业公司之间的战略差异使市场能够优胜劣汰。但是，在我们所讨论的判断问题上，系统噪声永远是一个问题，如果两位医生给你做出了不同的诊断，那么其中至少有一种诊断是错误的。

系统噪声的绝对数量及其造成的破坏程度之大令人震惊，这也正是我们撰写本书的动力所在。两者都远远超出了我们的预期。我们在本书中提及了商业、医学、刑事司法、指纹分析、天气预报、绩效考核和政治等许多领域的案例，并从中得出了结论：哪里有判断，哪里就有噪声，而且其数量之大远超我们的想象。

噪声在误差中的重大作用与人们普遍拥有的一种观念相悖，即“随机误差不重要”，因为他们可以“相互抵消”。然而，这种观念是错误的

的，如果靶心周围散布着很多弹孔，那么说“平均下来射击手击中了靶心”是毫无意义的。如果一名应聘者的得分高于他的实际水平，而另一名应聘者的得分低于他的实际水平，则可能导致用人单位聘错人。如果一份保单的保费过高，而另一份保单的保费过低，这两种错误对保险公司而言都是代价高昂的：前一种情况可能会使公司丢了生意，后一种情况则会使公司赔钱。

简而言之，如果缺少正当的、导致判断差异性过大的理由，那么判断过程一定存在误差。而且，即使我们无法对判断进行验证以及对误差进行测量，噪声也是有害的。“同罪不同罚”是不公平的，在一个系统中，如果相关人员的专业判断缺乏一致性，那么这个系统就会失去公信力。

水平、模式、情境，噪声的3种类型

系统噪声可分为水平噪声和模式噪声。有些法官通常很严厉，而另一些法官则更宽容；一些股票预测者总是预测牛市，另一些则总是预测熊市；有些医生开的抗生素比其他医生多。水平噪声是不同个体平均判断上的变异性，判断量表的模糊性是水平噪声的来源之一。像“可能”这样的词或“0~6分量表中的4分”这样的数字对不同的人来说含义是不同的。水平噪声是判断系统中的误差的重要来源，也是减少噪声过程中的一个重要干预对象。

系统噪声还包含另一种成分，这种成分通常占比更大。无论判决的平均水平如何，不同的法官对于哪种罪行应受更严厉的刑罚的看法可能有所不同。法官们的不同判决会导致对不同案件的排序不同。我们称这种变异为模式噪声〔统计术语为“统计交互作用”（statistical interaction）〕。

模式噪声的主要来源是稳定的，如不同法官对同一案件所做出的个

体化、特异性的反应。其中一些差异反映了个体（有意识或无意识）遵循的原则或价值观。例如，一位法官对偷盗者可能特别严厉，而对违反交通法规的人则较宽容；另一位法官可能刚好相反。某些潜在的原则或价值观可能非常复杂，而判断者可能对此毫无意识。例如，某位法官可能对年龄较大的偷盗者比较宽容，他自己却完全没有意识到这一点。同时，对特定案例高度个体化的反应也可能是稳定的，比如，某位法官由于觉得被告长得像自己的孩子，从而对被告产生了怜悯之情，并对被告予以宽大处理。这位法官在不同时间里遇到这种情况，他都会如此。

这种稳定的模式噪声反映了法官的独特性：他们对案件的反应与他们独一无二的人格特征一样。人与人之间的细微差异通常很微妙也很有趣，但是，在需要一致性判断的系统中，这种由专业人员做出的判断间的差异是有问题的。在我们所考察的研究中，这种因个体差异而产生的稳定的模式噪声通常是系统噪声的最大来源。

尽管如此，法官对特定案件的不同态度也不完全是稳定的，也就是说模式噪声也包含一个可变成分，我们称之为情境噪声。如果放射科医生在不同的日子里对同一张影像片子做出了不同的诊断，或是指纹鉴定师有时认为两个指纹是匹配的，有时则认为是不匹配的，我们就能在其中检测到情境噪声。正如上述例子所示，如果判断者没能识别出某个案例是他以前处理过的案例，我们很容易在他做判断的过程中测量出情境噪声。另一种证明存在情境噪声的方式是发现与判断无关的背景因素对判断产生了影响。例如，当法官最喜欢的足球队获胜后，他们变得较宽容；医生在下午通常会开出更多的阿片类药物。

判断与噪声的心理机制

判断者的认知缺陷并非预测性判断中存在误差的唯一原因，客观无知往往起着更大的作用。实际上，有些事确实是不可知的，例如，新生的婴儿中，有多少人会在70年后抱上孙子；明年开奖的彩票，其中奖号

码是多少。还有一些信息也许是可知的，但判断者并不知道。人们往往会在做出预测性判断的过程中表现得过度自信，低估了他们的客观无知和偏差。

我们预测的准确性是有限的，而且这个限度通常很低，尽管如此，我们还是对自己的判断感到满意。这种令人满意的信心来自内部信号，即当事实和判断匹配成连贯的故事时自发产生的奖励，然而，我们对判断的主观信心未必与其客观准确度有关。

大多数人在得知自己的预测性判断准确度很低甚至不如公式时，都会很惊讶。然而事实上，即使是建立在有限数据上的简单线性模型或在信封背面就能写得下的简单规则，也常常胜过人类的判断，规则和模型的关键优势在于它们没有噪声。正如我们主观体验到的那样，判断是一个微妙而复杂的过程。这样说并非想要表明这种微妙之处大部分是噪声，而是说我们很难想象，盲目地遵守简单的规则都比人类自己做出判断更准确，但这的确是事实。

心理偏差显然是导致系统误差或统计偏差的来源之一，尽管没那么明显，心理偏差也是噪声的来源之一。如果并非所有判断者都共享同样的偏差，或他们产生偏差的程度各不相同，抑或偏差的影响取决于外部环境，那么此时心理偏差就会产生噪声。例如，如果一半的经理在做招聘决策时对女性存在歧视，而另一半经理又偏好录用女性，则整体上不会有偏差，但系统性噪声会导致许多招聘错误。另一个例子是第一印象效应，这是一种心理偏差，但如果证据的呈现顺序是随机的，这种偏差就可能产生情境噪声。

我们将判断过程描述为：非正式地对一系列线索进行整合，并在一个量尺上做出判断的过程。因此，消除系统性噪声需要判断者在线索的使用、给线索赋予权重以及使用量尺上保持一致。即使不考虑情境噪声的随机影响，这些要求也很难满足。

在单一维度的判断中，人们判断的一致性通常很高，比如两个候选

人中哪个更具魅力或哪个更勤奋，不同的招聘者往往能在这类评价性问题上达成共识。人们共有的对不同维度进行匹配的直觉过程让人们足够产生相似的判断，如人们将高GPA与超前的阅读能力相匹配。如果一个判断中只有少量线索，并且这些线索基本都指向同一个方向，那么人们也容易在这样的判断上达成一致。

当判断需要对多个彼此冲突的线索赋予一定的权重时，巨大的个体差异就会产生。对于同一位候选人，有些面试官会更重视才华或个人魅力，有些面试官则可能更看重勤奋或抗压能力。当线索不一致并且不能构成连贯的故事时，不同的人在重视哪些线索和忽略哪些线索上必然会表现不同，这就会产生模式噪声。

不易觉察的噪声

噪声不容易被人觉察，也很少被讨论，而且显然比偏差更少受到关注，因此你可能没怎么考虑过它。不过，噪声如此重要，却不容易被人觉察，这本身也是一种非常有趣的现象。

认知偏差以及其他情感或动机对思维的扭曲，常被用来解释糟糕的判断。分析人士用过度自信、锚定效应、损失厌恶、可得性偏差（availability bias）以及其他种类的偏差来解释那些糟糕的决策。基于偏差的解释常常令人满意，因为人类思维渴望因果性解释。每当出现问题时，我们都会去寻找原因，通常情况下我们也确实能够找到原因。在一些情况下，偏差看起来就像是不良决策的成因。

偏差在解释性方面颇具魅力，噪声却缺乏这种魅力。如果试图事后解释为什么某个特定的决定是错误的，那么我们可能很容易发现偏差，但不会发现任何噪声。只有用统计思维来看待世界，我们才会看到噪声，但这种统计思维并非与生俱来，因为我们往往更喜欢因果性的故事。我们的直觉系统缺乏统计思维，这就是噪声比偏差更少受到关注的

原因之一。

另一个原因在于，专业人士很少认为自己和同事的判断会有噪声。经过一段时间的培训后，他们通常会自行做出判断。指纹专家、经验丰富的核保员和资深专利审核员几乎不会去想其他同事会与自己的意见不一致，他们更不会去想，其他同事会如何质疑自己的意见。

大多数时候，专业人士会对自己的判断充满信心。他们总以为自己的同事会赞同自己的意见，却从来不会去弄清楚是否的确如此。在大多数领域，人们做出判断后永远不会将其与真实值进行比较，最多是让另一位尊重型专家对其判断进行审查。专业人士之间偶尔才会出现大到令人惊讶的分歧，而且一旦出现这种分歧，他们通常会找到理由来将其视为极端案例。组织程序也倾向于忽略或压制专家之间的分歧，这是可以理解的，毕竟站在组织的立场上，噪声是一种令人尴尬的存在。

决策卫生的6个原则

我们有理由相信，有些人的判断就是比其他人好。最佳判断者具有如下特征：拥有与任务相关的技能、智力以及特定的认知风格——用积极开放性思维来形容这种认知风格最为恰当。毫无疑问，优秀的判断者犯严重错误的情况很少。然而，鉴于导致个体差异的因素众多，即使是最好的判断者，我们也不应该奢望他们能够在复杂的判决任务中达成完美的一致性。背景、个性和经历的千差万别使得我们每个人都与众不同，这也是我们无法避免噪声的原因。

消除偏差是减少误差的一种策略。通常，人们要么在事前控制偏差的影响，要么在事后对偏差进行校正。我们提出第三种方案，它特别适用于团队决策：指定一位决策观察者来识别偏差出现的迹象，从而对偏差进行实时检测（参见附录2）。

对于降低判断中的噪声，我们的主要建议是采取决策卫生策略。我

们之所以选择“决策卫生”这个词，是因为减少噪声就像平常的卫生习惯一样，是为了预防未知的敌人。例如，洗手可防止未知病原体进入身体。同理，决策卫生可以防止未知的错误出现。决策卫生与它的名字一样乏味，远不如战胜可预测性偏差那样令人振奋，预防未知损害可能也得不到什么荣誉，但依然值得我们这么做。

组织中的减少噪声的工作也应该从噪声审查开始（参见附录1）。噪声审查的一项重要功能是使组织由衷地认真对待噪声。这将有利于对不同类型的噪声进行评估。

在前文中，我们描述了不同领域中减少噪声工作的成效与局限。现在，我们简单总结一下决策卫生策略的6项原则，描述它们如何解决引发噪声的心理机制并展示它们如何与特定的决策卫生策略之间建立联系。

原则1：判断的目的在于准确性，而不在于个性化表达。这是判断中必须遵循的首要决策卫生策略，它反映了本书对判断一词狭义、具体的定义。我们发现，稳定的模式噪声是系统噪声中很大的组成部分，它是个体差异以及“判断人格”的直接结果，判断人格会导致不同人在面对同一问题时持不同观点。基于这些发现，我们得出一个不受欢迎但又无法忽视的结论——做判断不是表达个性的场景。

需要明确的是，个人的价值观、个性和创造力在思考与决策的许多阶段都是重要的，甚至是必要的，包括目标选择、形成解决问题的新方案以及生成选项。但在对这些选项做出判断时，个性化的表达就成了一种噪声源。如果判断是为了达到准确性，且你希望其他人认同你的判断，那么你就需要换位思考：如果其他称职的判断者处于你的位置，他们会怎么想。

遵循这一原则的一种比较激进的方法是用规则或算法来代替判断。用算法做评估可以保证消除噪声——实际上，它是完全消除噪声的唯一方法。算法已在很多重要领域中得到应用，并且越来越受到重视，但

是，在重要决策的最终阶段，算法不太可能取代人类判断——我们认为这是件好事情。但是，适当地使用算法或至少让决策不那么依赖于专业人士的个人喜好，可以改善判断。例如，我们已经看到了，决策指南可以有效地限制法官的自由裁量权、提升医生诊断的一致性，进而减少噪声和提升决策品质。

原则2：使用统计思维，采用外部视角审视个案。当判断者将某个案例视为一系列相似案例中的一例而不是把它看作特例时，我们就说，这个判断者采用了外部视角。这种方法不同于常规思维模式，即只关注当前案例，并将其纳入因果性故事中。当人们基于自身独特的经历来形成对特定事件的独特看法时，其结果就是会产生模式噪声。外部视角可以解决这一问题：具有相同参照系的专业人员会有更少的噪声，此外，外部视角也有助于产生有价值的见解。

外部视角原则倾向于将预测锚定在一群类似案例的统计特征上。它也提示我们，预测应该是适中的——更技术性的术语是回归（参见附录3）。关注到以往结果的广泛分布及其有限的可预测性，有助于决策者调整自己的判断信心。人们不能因为在不可预测的事情上做出了错误的预测而遭受责备，但如果错误是由于他们在预测时过度自信，那么他们就应该受到责备。

原则3：对判断进行结构化，将其分解成几个独立的任务。这一“分而治之”的原则对于解决我们称之为过度一致性的心理机制所产生的影响是很有必要的。过度一致性会使人们曲解或忽略那些与已有结论或刚刚出现的故事不相符的信息。当同一案例的不同方面所产生的印象相互“污染”时，判断在总体上的准确性就会受到影响。比如，如果允许证人之间互相交流，你认为他们提供的证词还有多少价值呢？

人们可以通过将判断问题分解为一系列的小任务来减少过度一致性问题。这项技术与结构化访谈的原理类似。在结构化访谈中，访谈员一次只评估一个特征，即给前一个特征评完分之后，再对下一个特征进行

评分。结构化原则对一些医学诊断指南具有启发意义，例如针对新生儿的阿普加评分，该原则也是我们提过的中介评估法的核心。中介评估法将一个复杂的判断分解为多个基于事实的评估，其目的在于保障每个评估都是独立进行的。因此，尽可能将每个评估分配给不同的团队，并尽量减少他们之间的沟通，这样能确保每个评估的独立性。

原则4：抵制不成熟的直觉。我们介绍了完成判断时的内部信号，它赋予了决策者判断的信心。决策者不愿意放弃这种奖赏性的内部信号，这也是他们不愿意使用指南、算法或者其他束手束脚的规则的主要原因。显然，决策者需要对他们的最终决策感到满意，并从直觉的自信中获得回报感，但是他们不应该过早地给予自己这种回报感。通过对信息的权衡和仔细思考而得出的直觉决策远胜于快速判断，因此不必禁用直觉，但直觉应该建立在一定信息的基础上，且应该接受规则的约束与适当的延迟处理。

这一原则启发我们对信息进行排序：不应给做判断的专家呈现他们不需要的以及可能引发偏差的信息，即使这些信息是正确的。例如，在司法科学中，最好不要让检查员知道有关犯罪嫌疑人的其他信息。中介评估法的一个关键要素——控制讨论议程也是如此，一个有效的讨论议程将确保人们分别从不同的方面考虑问题，直到所有方面的评估都完成之后，人们再做出整体判断。

原则5：获取多位判断者的独立判断，再考虑汇总这些判断。很多组织的常规流程违反了独立性原则，尤其是一些会议流程会导致参与者的意见受到他人的影响。由于信息级联和群体极化效应，群体讨论通常会放大噪声。在讨论之前收集参与者的判断，这样一个简单的流程既可以揭示噪声的程度，又有助于以建设性的方式解决分歧。

对所有的独立判断求均值可确保减少系统噪声，但不能减少偏差。单一判断是从所有可能判断的总体中抽取的一个样本，增加样本量可以提高估算的准确性。当判断者具有多样化的技能以及互补的判断模式

时，求均值的优势会进一步增强。对一系列充满噪声的判断进行平均，其结果可能比全体一致的判断更为准确。

原则6：用相对判断和相对量表会更好。相对判断的噪声要比绝对判断的噪声小，因为我们在同一量表上对多个对象进行归类的能力是有限的，但是我们对众多对象进行两两比较的能力却更胜一筹。基于比较的判断量表要比需要绝对判断的量表噪声更少，例如，我们可以用每个人都熟悉的案例作为一把案例量表，判断者们只需要确定某起案件在量表上的相对位置即可。

我们刚刚列举的决策卫生原则不仅适用于重复性决策，也适用于单次的重大决策，即我们所谓的单一决策。单一决策中会存在噪声，这似乎与我们的直觉相悖：从定义上来看，如果你仅做一次决策，就没有可以测量的变异，也就是说不应存在噪声，但是噪声的确存在，而且会导致误差。如果我们只看到第一个射击手的射击情况，那么整个队伍的噪声是看不见的，但是当我们看到所有射击手的射击情况后，整个队伍的射击分布就变得清晰了。同样，考量单一决策的最佳方法是将其视为仅进行了一次的重复性决策。这就是决策卫生策略也能改善单一决策的原因。

强制采用决策卫生策略可能会费力不讨好。噪声是隐形的敌人，战胜隐形的敌人也只能取得隐形的胜利，但是，就像保持身体健康需要讲卫生一样，决策卫生至关重要。一位患者的手术成功后，你会相信这是外科医生的妙手回春挽救了患者的生命，但如果外科医生和手术室中的其他人员都没有洗手，这位患者可能已经一命呜呼了。做好杀菌消毒工作可能没有什么值得炫耀的，却会产生实实在在的好处。

噪声是隐形的敌人，我们应该重视这个敌人

当然，与噪声做斗争并不是决策者和组织要考虑的唯一因素。减少

噪声的成本可能太高了：一所高中可以要求更多的老师来阅读每篇作文，从而消除评分中的噪声，但是这种额外的工作量是没必要的。实践中不可避免地会有一些噪声产生，这是给予每一个案例个别化考虑的系统必然产生的副产品，它使得人们不像机器的齿轮，并确保了解决者的主观能动性。有些噪声甚至是人们所期待的，前提是它产生的变异有利于系统适应时代的变化，比如噪声反映了社会不断变化的价值观和目标，以及引发了能导致实践或司法变革的讨论。

也许其中最值得重视的是，降噪策略导致了令人无法接受的不良后果。许多有关算法的担忧实际上是被夸大了，但有一些担忧是合理的。即使算法可以成功地避免人类会犯的很多错误，它们也可能会犯人类永远不会犯的愚蠢错误，并因此失去可信度。算法也可能会因为设计不当或训练数据选取不当而产生偏差。另外，算法的模糊性也可能会引起人们的不信任。决策卫生在实践中也有其弊端：如果管理不善，它就会产生决策官僚化的风险，专业人士也可能因为自己的自主权受损而士气低落。

所有这些风险和局限性都应该得到充分考虑。不过，拒绝减少噪声的理由是否合理，取决于你所考虑的减少噪声策略具体是什么。拒绝汇总判断的理由，如成本太高或许不适用于拒绝使用指南。可以肯定的是，一旦减少噪声的成本超过其收益，就不应该再执着于减少噪声了。一旦进行成本收益分析，或许就能找到一个非零的最佳噪声水平。问题在于，在没有进行噪声审查的情况下，人们意识不到判断中包含了多少噪声，在这种情况下讨论减少噪声的难度，只不过是我不想测量噪声而找借口而已。

偏差会导致误差和不公，噪声也是如此，但我们在这方面所做的研究很少。相比于将判断误差归结为因果，当将判断误差归结为随机性因素时，人们对判断误差的容忍度更高，但这并不意味着随机性因素带来的判断错误的危害更小。若想在重大决策上做得更好，我们就应该认真

对待减少噪声问题。

一个噪声很少的世界

请想象一下，经过重新设计并减少了噪声的组织会是什么样子。医院、招聘委员会、经济预测机构、政府机构、保险公司、公共卫生机构、刑事司法系统、律师事务所和大学，都会对噪声问题保持警觉，并努力减少这些问题。噪声审查将会成为惯例，或许每年都能进行一次。

未来，组织的领导者将会在更多领域中使用算法来代替人类判断，或将其作为人类判断的补充。人们会将复杂的判断分解为更简单的中介评估。他们将了解决策卫生策略，并遵循它的6大原则。人们会做出独立的判断，并对这些判断进行汇总。会议看起来也会非常不同，讨论将会更加具有结构性。人们会更系统地将外部视角整合到决策过程中。分歧会出现得更加频繁，但也将会得到更具建设性的解决方案。

这将创造一个噪声更少的世界。我们可节省大量资金，改善公共安全和卫生等状况，令人们得到更加公平的对待，并防止许多可以避免的错误。我们撰写本书，就是希望你能够关注这样的机遇，我们也希望你能够抓住这一机遇。

如何进行噪声审查

附录1是一份关于如何进行噪声审查的实用指南。你可以以企业顾问的视角来阅读这部分内容，设想一下这家企业雇用你来对它的一个部门员工的专业决策进行噪声审查。

顾名思义，噪声审查的重点当然在于考察噪声的普遍性，然而，良好的噪声审查其实也能提供关于员工培训和工作监管中存在的偏差、盲点和缺陷的有价值信息。成功的审查有利于促进部门改革，包括改进专业人员的判断准则、他们接受的培训、支持他们判断的工具以及对工作的日常监管等。如果这项审查被认为是成功的，那么它就有可能被应用于组织中的其他部门。

噪声审查需要进行大量的工作，且要求高度关注细节，因为如果审查结果暴露出重大缺陷，审查的可信度肯定会遭到质疑。因此，对案件和程序的每一个细节都要深思熟虑，要经得起“充满敌意”的审查。我们所描述的这一过程，招募了很多可能对噪声审查提出严格批评的人来进行，以减少反对的声音。

顾问可以是企业外部人员，也可以是企业内部人员，除此之外，还包括以下相关主体：

- 项目团队。项目团队对研究的所有阶段负责，如果项目顾问来自企业内部，他们将构成项目团队的核心；如果顾问来自企业外部，那么企业内部的项目团队需要与他们密切配合，这样才能确保公司内部人员将噪声审查视为自己的项目，并将顾问视为他们的支持者。除了对数据收集过程进行管理、分析结果和

编写最终报告的顾问之外，项目团队还应包括相关领域的专家，并需要由他们来构建判断者要评估的案例。同时，项目团队的所有成员都应该有很好的职业信誉。

- 客户。噪声审查只有在引发重大改变时才有用，这就需要项目的“客户”——企业的领导一开始就要介入进来。可以预见的是，客户最初会对噪声的普遍性持怀疑态度。如果秉承开放的态度、怀有对审查结果的好奇，并且承诺当顾问的悲观预期被证实时愿意力挽狂澜，那么这种最初的怀疑实际上是一种优势。
- 判断者。客户可以指定一个或多个部门进行审查。选定的部门应包括大量的“判断者”，即为公司做出类似判断和决策的专业人士，而且判断者应该是可以有效替换的。例如，如果一个人不能处理一起案件，公司可以指派另一个人去处理，并认为他能够做出类似的判断。本书中介绍的这类例子包括美国联邦法官的量刑决策、保险公司对保费和理赔准备金的设定。对于噪声审查，最好选择类似这样的判断任务：可以以书面的形式完成，且可以以数字的形式，如货币、概率或评分来表示。
- 项目经理。应指定一位行政部门的高级经理担任项目经理，这项任务不需要相关人员具有特定的专业知识。在组织中，高级职位对于克服行政方面的障碍具有实际意义，同时也表明了公司对该项目的重视。项目经理的职责是对项目实施的各阶段提供行政支持，包括编写最终报告以及将结论传达给公司领导层。

构想案例材料

项目团队中的领域专家应具有公认的专业知识，例如，设定风险保

费或评估投资潜力的能力，他们将负责研发审查中使用的案例。有效模拟专业人士的判断是一项非常精细的任务，当审查暴露出严重问题并有可能受到详细审查时更是如此。因此，团队必须考虑以下问题：如果我们的模拟结果表明公司的决策存在很高水平的噪声，公司的相关人员会接受“实际判断中同样存在这样的噪声”的事实吗？只有在得到肯定答复的情况下，噪声审查才值得进行。

获得肯定答复的方法不止一种。本书第1章中介绍了对法官量刑的噪声审查，通过使用相关属性的简单列表来概括每个案例，并在90分钟内获得了对16个案例的评估。第2章介绍了保险公司的噪声审查，对复杂案例做了详细而真实的概括。这两种情况下的高水平噪声为噪声的存在提供了确凿的证据，因为，如果在简化的案例中都存在这么大的分歧，那么真实案例中的噪声只会更严重。

为了更好地理解每位判断者对案例做出判断的逻辑，我们可以为每个案例准备一份调查问卷，并要求判断者在对所有案例的判断都完成后才可以开始做问卷。问卷应包括：

- 一些开放性问题，以此考察导致参与者做出相应判断的核心因素。
- 一份与案例有关的事实列表，参与者可以在列表中对它们的重要性进行评估。
- 一些有利于促进“外部视角”的问题。例如，如果案例需要进行估价，那么参与者应该给出该案例与同类别案例的估价的对比，即结果是高于平均值还是低于平均值。

与高管们召开启动会

将用于审查的案例所需材料都准备好之后，应当安排一次会议，让

项目团队向公司领导层汇报审查内容。会议讨论的内容应当涉及研究的可能结果，当然，该可能结果中也应包括发现了令人难以接受的系统噪声。会议的目的是听取对即将实施的研究计划的反对意见，并获取领导层愿意接受任何结果的承诺，如果没有这样的承诺，进行下一步是毫无意义的。如果有人提出了强烈的反对意见，那么项目团队可能就需要完善案例材料并再次尝试。

当高管们接受噪声审查的设计时，项目团队应当让他们陈述对研究结果的期望。他们应讨论与下面这些问题类似的问题：

- 对于每一个案例而言，如果随机选择一组答案，你们预期的分歧程度有多大？
- 从商业的角度来看，可接受的最大程度的分歧是多少？
- 如果出现一定程度的错误高估或低估（比如15%），预计损失有多大？

应该详细记录上述问题的答案，以确保当审查的实际结果出来时大家仍记得且相信这些答案。

研究的管理

从一开始就应该告知被审查部门的经理，他们的部门被选择来做一项特殊的研究。有一点很重要，在描述项目时不能使用“噪声审查”这个词，尤其是在形容人的时候，应避免使用“噪声”和“充满噪声”这样的词，应使用“决策研究”这样的中性词。

部门经理要立即着手收集资料，并在项目经理和项目团队成员的参与下，向相关人员介绍任务，而且应当用概括性表述来介绍任务的目的是，比如“企业对决策者如何得出结论感兴趣”。

其中很重要的一点是，需要向参与研究的专业人士确认，包括项目团队在内的企业中的任何人，都不会知道他们个人的答案。如果有必要，可以聘请外部公司匿名处理数据。另外，强调审查不会对该部门产生任何特殊影响也很重要，让他们了解他们只是被选为部门代表，代表企业完成判断任务。为了确保结果的可靠性，部门内所有有资格参与的专业人士都应参与到研究中。此外，还可以专门为这个任务安排半个工作日的时间，这样有助于使参与者认识到任务的重要性。

需要注意的是，所有参与者都要同时完成这个任务，但是在任务进行的过程中，他们应该身处彼此独立的空间，并且不能相互交流。项目团队可以在研究过程中回答问题。

分析和结论

项目团队将负责对每一位参与者评估的多个案例进行数据分析，包括对整体噪声的数量、噪声的成分、水平噪声和模式噪声的数量等进行分析。如果案例材料允许，进行数据分析时还可以识别出反应中的统计偏差。项目团队还有另一个重要的任务，即通过分析被试在问卷中的反馈来确定判断中变异的来源。在这些问卷反馈中，被试解释了他们做出某种反应的理由，并表明了哪些事实对他们的决策起到了决定性作用。通过聚焦分布于两端的极端反应，团队可以探寻数据中的模式，寻找在员工培训、企业流程以及企业为员工提供的信息中可能存在的缺陷。

顾问和内部项目团队将共同开发工具和流程，应用决策卫生策略和消除偏差的原则，以改善部门内的决策与判断，这一过程可能会延续数月。与此同时，顾问和专业团队还将编写一份项目报告提交给该组织的领导。

此时，企业已经在它的一个部门中做了噪声审查的尝试。如果这个尝试被认为是成功的，接下来管理团队可能会对企业内的判断与决策开

展更广泛的评估以及品质方面的提升。

决策观察者检查清单

本附录给出了一份决策观察者（见第19章）检查清单的通用示例。这里提供的清单大致遵循做重要决策的时间顺序。

清单中每一项之后的参考问题都对相应的项目做了进一步说明。决策观察者在审视决策过程时应该问自己这些问题。

决策观察者不应照搬照抄这个清单。我们希望它发挥的作用是为决策观察者提供灵感和基准，并使决策观察者能够据此设计出适合自己的偏差观察清单。

偏差观察清单

1. 判断方法

a. 替代

- 团队对证据的选择和讨论的焦点是否表明，他们用一个简单问题替代了分配给他们的困难问题？
- 团队是否忽视了一个重要因素或过度重视了一个无关因素？

b. 内部视角

- 团队是否采用了外部视角作为其决策考量的一部分，并认真尝试运用相对判断而非绝对判断？

c. 观点的多样性

- 是否有理由怀疑团队成员的判断存在共同的偏差？正是这种共同的偏差导致他们出现共性的错误。那么，你能否想到了这个团队没有提及的一个相关观点或专业知识？

2. 预判和过早下结论

a. 启动预判断

- 决策者中是否有人会因为得出某个结论而比得出其他结论的人获利更多？
- 是否已经有人认定了某个结论？有没有理由怀疑该结论存在偏见？
- 持不同意见的人表达他们的观点了吗？
- 是否存在对已然失败的行动仍固执坚持的风险？

b. 过早下结论和过度一致性

- 在选择先前讨论过的考虑因素时，是否存在偶然出现的偏差？
- 是否充分考虑了其他可能性并积极寻找过支持其他可能性的证据？
- 令人不安的数据或观点是否遭到忽视或压制？

3. 信息处理

a. 可用性和显著性

- 参与者是否因事件是近期发生的、颇具戏剧性且与个人相关而夸大了该事件的相关性，哪怕该事件并不具有判断价值？

b. 对信息品质不重视

- 判断是否严重依赖于传闻、故事或类比？有数据可以证实这样的判断吗？

c. 锚定

- 准确度或相关性不确定的数字是否在最终判断中起到了重要作用？

d. 非回归性预测

- 参与者是否进行了非回归性推断、估计或预测？

4. 决策

a. 计划谬误

- 当使用预测时，人们是否质疑过其来源和有效性？是否使用过外部视角来质疑这些预测？
- 不确定的数字是否有置信区间？置信区间够宽吗？

b. 损失厌恶

- 决策者的风险偏好是否与组织一致？决策团队是否过于谨慎？

c. 即时倾向

- 采用的计算方式（包括所使用的折现率）是否反映了企业对短期目标和长期目标之间优先性的权衡？

对抗噪声，修正预测

匹配性预测是我们依赖直觉式匹配过程而产生的错误（参见第14章）。当我们基于现有信息做预测时，我们会进行匹配性预测，就好像我们能够依据这些信息对结果做出完美的或准确率非常高的预测一样。

让我们来回顾一下朱莉的例子，她4岁就能流利地阅读，那么她在大学时的GPA会是多少呢？如果你预测朱莉在大学时的GPA是3.8，这就意味着你从直觉上判断朱莉4岁时的阅读水平处于其年龄组的前10%（但不在前3%~5%）。然后，你心中有一个隐含的假定，即朱莉上大学时的GPA也应处于她所在班级的第90百分位数左右，那么与此相对应的GPA就是3.7或3.8——这是很常见的答案。

然而，这种推理在统计学上是不正确的，因为它严重夸大了关于朱莉的已知信息的判断价值。一个早慧的4岁孩子不一定永远都是“学霸”，当然了，小时候有阅读障碍的孩子也不一定一直在班里学习成绩垫底。

在大多数情况下，很多人的出色表现不可能一直延续下去；反之，糟糕的表现也往往会得到改善。尽管我们很容易就能想到一些社会学、心理学甚至政治学的原因来对这种现象做出解释，但这根本没必要，这纯粹是一个统计学现象。因为过去的表现与未来的表现并非完全相关，在一个方向或另一个方向上出现的极端观测值往往会在后续的观察中变得不那么极端，这种趋势被称为均值回归。因此，就匹配性预测而言，我们使用的术语是非回归性，因为它没有考虑到均值回归这一现象。

从定量的角度来看，如果能流利阅读的年龄是未来GPA的完美预测

指标，换句话说，如果这两个因素之间的相关系数为1，那么你对朱莉的判断就是正确的。然而，显而易见的是，现实情况并非如此。

有一种统计方法可以帮助人们做出更准确的判断，但即使对于受过一些统计学训练的人来说，这种方法也不是很直观，并且很难觉察。接下来，我们以朱莉的情况为例来介绍这种方法的实施流程（见图33-1）。

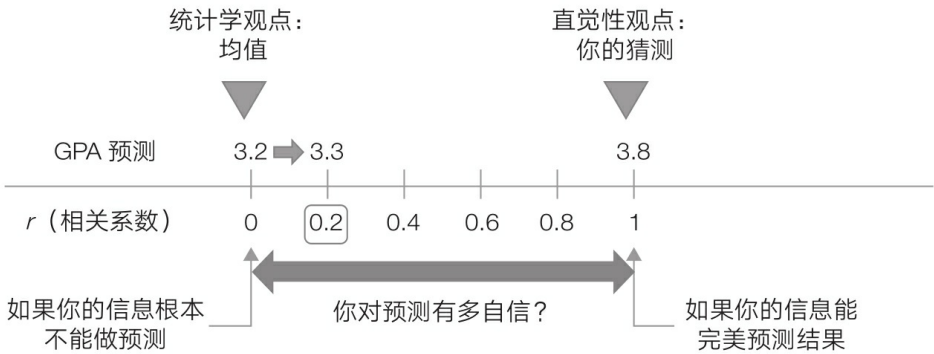


图33-1 以均值回归调整直觉性预测的图示

1. 用直觉去猜测

你关于朱莉的直觉或关于任何你有所了解的事件的直觉都并非毫无价值。你的快速的系统1思维可以轻松地将你所拥有的信息置于相应的预测量尺上，并为朱莉的GPA猜测一个数值。如果你拥有的信息具有完美的预测性，你就应该按照你的猜测来进行预测。猜完后请写下你的结论。

2. 寻找均值

现在，请你后退一步，暂时忘记关于朱莉的一切信息。如果你对朱莉一无所知，你会如何预测她的GPA呢？答案很简单：在没有任何信息

的情况下，你对朱莉的GPA的最佳估计应该是大学生群体的平均GPA——3.2左右。

以这种方式来看待朱莉问题遵循了我们在前文提到过的一种更广泛的原则，即外部视角。当我们采用外部视角时，我们将所考察的案例视为一类案例中的一则个案，然后从统计学角度去考察这一类案例。例如，请你回想一下，在迈克尔·甘巴迪问题中，外部视角是如何帮助我们思考新任CEO能否留任的基础概率的（参见第4章）。

3. 评估已有信息的诊断价值

这是很困难的一步，你需要问自己：“我目前拥有的信息的预测价值有多大？”这个问题很重要，而原因再清楚不过了。如果你只知道朱莉的鞋码，那么你会恰当地将此信息的权重设为0，并坚持平均GPA的预测；而如果你有朱莉各科成绩的列表，据此信息将能完美地预测她的GPA。这两个极端之间有许多灰色地带。如果你有朱莉高中时期优异的学业成绩的数据，那么这些信息将比她的阅读年龄更具诊断意义，但比起她的大学成绩，其诊断意义又会小得多。

你的任务是量化已知数据的诊断价值，你可以将这个诊断价值表示为已有数据与你所预测的结果之间的相关性。除极个别情况外，这一相关性肯定只是粗略估计。

想要合理地估计这一相关性，请你回忆我们在第12章中列举的一些例子。在社会科学领域，超过0.5的相关系数非常罕见。在多数情况下，相关系数达到0.2就已经十分难得了，而在朱莉的例子中，0.2可能就是相关系数的上限了。

4. 基于外部视角来调整你基于直觉的猜测，从而尽最大可能体现已有信息的诊断价值

最后一步是将已有的三个数字进行简单的算术整合，即对平均值进行调整，让它向基于猜测的值的方向偏移，偏移的程度取决于你所估计的相关系数的大小。

这一步只是我们对刚刚所做观测的进一步扩展：如果相关系数为0，那么你应该坚持平均值的估计；而如果相关系数是1，那么你应该忽略平均值并愉快地做出匹配性预测。因此，在朱莉的例子中，对其GPA的最佳预测应该是：由群体平均值向依据阅读年龄做出的直觉估计值的方向偏移，偏移的距离不超过平均值与直觉估计值距离的20%。据此，我们可以算出最佳预测值大约是3.3。

虽然这里我们只是举了朱莉的例子，但这种方法也可应用于本书讨论过的许多判断问题。例如，有一位销售副总裁需要招聘一名销售人员，他刚刚面试了一位十分优秀的候选人。根据对候选人的深刻印象，这位高管估计该候选人入职第一年的销售额能达到100万美元——相当于新入职员工第一年平均销售额的两倍。如何让副总裁的估计有回归性呢？需要注意的是，计算要取决于面试的诊断价值。在这个案例中，招聘面试能够在多大程度上预测出应聘者入职后的绩效呢？依据我们见过的研究证据，0.4的预测相关系数是一个相当高的估计值。因此，对新员工第一年销售额的回归性估计至多是： $50\text{万美元} + (100\text{万美元} - 50\text{万美元}) \times 0.4 = 70\text{万美元}$ 。

这一过程并不是很直观。如上例所示，值得注意的是，校正后的预测总是比直觉性预测更保守，它们永远不会像直觉性预测那样极端，而是会更接近于平均值，通常情况下会非常接近平均值。如果你校正了你的预测，对于那些已经赢了10次大满贯的网球冠军，你就不会赌他们还能再赢得10次大满贯。同样，你也不会预测一家价值10亿美元的成功的创业公司将会发展成价值数千亿美元的行业巨头。总之，校正预测不会在极端值上下注。

从事后看，这意味着校正预测将不可避免地造成明显的失误，然

而，预测并不是事后进行的。你应该记住，极端值是极其罕见的，与之相反的错误反而更为常见：我们往往以为自己预测的极端值在未来依然是极端值，但由于存在均值回归，情况往往会发生变化。这就是当目标是最大限度地提高预测准确性（即均方误差最小化）时，校正预测会优于直觉匹配性预测的原因所在。

致谢

我们要感谢很多人。林内亚·甘地（Linnea Gandhi）是我们的参谋长，她为我们提供了大量的指导和帮助，使我们的工作可以有条不紊地推进，让我们的工作充满欢声笑语。基本可以说，是她主持了本书的撰写。除此之外，她还为本书的初稿提出了许多宝贵建议，没有她，我们无法顺利完成本书。丹·罗瓦洛对成书也起到了非常重要的作用，本书的观点最初就源于他与人合著的一篇文章。我们的经纪人约翰·布罗克曼（John Brockman）在工作的每个阶段都热情洋溢、乐观、敏锐而又睿智，我们很感谢他。我们的主编，也是我们的引路人特蕾西·比哈（Tracy Behar）从诸多方面提升了本书的品质。阿拉贝拉·派克（Arabella Pike）和伊恩·斯特劳斯（Ian Straus）也从编辑角度提出了极好的建议。

同时也要特别感谢奥伦·巴-吉尔（Oren Bar-Gill）、马亚·巴-希勒尔（Maya Bar-Hillel）、马克斯·巴泽曼、汤姆·布莱泽（Tom Blaser）、戴维·布代斯库（David Budescu）、杰里米·克利夫顿（Jeremy Clifton）、安塞尔姆·丹内克尔（Anselm Dannecker）、维拉·德莱尼（Vera Delaney）、伊蒂尔·德鲁尔、安杰拉·达克沃思（Angela Duckworth）、安妮·杜克、丹·吉尔伯特（Dan Gilbert）、亚当·格兰特、阿努帕姆·杰纳（Anupam Jena）、路易斯·卡普洛（Louis Kaplow）、加里·克莱因（Gary Klein）、乔恩·克莱因伯格（Jon Kleinberg）、内森·昆塞尔、凯利·莱纳德（Kelly Leonard）、丹尼尔·莱文（Daniel Levin）、萨拉·麦克拉纳汉、芭芭拉·梅勒斯、乔西·米勒（Josh Miller）、塞德希尔·穆来纳森、斯科特·佩奇（Scott Page）、埃里克·波斯纳（Eric Posner）、露西亚·瑞奇（Lucia Reisch）、马修·萨尔加尼克、埃尔德·沙菲尔、塔利·沙罗特（Tali Sharot）、菲利普·泰特洛克、理查德·泰勒（Richard Thaler）、芭芭拉·特沃斯基（Barbara Tversky）、彼得·乌贝尔（Peter

Ubel）、克里斯特尔·王（Crystal Wang）、邓肯·瓦茨（Duncan Watts）和卡罗琳·韦布（Caroline Webb），他们对书稿的部分章节乃至全文进行了阅读和评论，感谢他们的慷慨帮助。

我们很幸运能够从这么多优秀研究者的建议中受益。朱利安·帕里斯（Julian Parris）在许多统计问题上提供了极其宝贵的帮助。没有塞德希尔·穆来纳森、乔恩·克莱因伯格、詹斯·路德维希（Jens Ludwig）、格雷戈里·斯托达德（Gregory Stoddard）和常惠（Hye Chang），我们就不可能完成介绍机器学习成就的那几章。我们对判断一致性问题的讨论很大程度上要归功于亚历克斯·托多罗夫（Alex Todorov）和他普林斯顿大学的同事乔尔·马丁内斯（Joel Martinez）、布兰登·拉伯瑞（Brandon Labree）、斯蒂芬·乌登伯格（Stefan Uddenberg）、斯科特·海浩斯（Scott Highhouse）以及艾莉森·布罗德富特（Alison Broadfoot）。这些了不起的研究者不仅热情分享了他们的见解，还非常友善地帮我们进行了一些特殊数据的分析。当然，如果书中出现了任何误解或错误，那一定是我们的责任。此外，我们还要感谢拉兹洛·博克、博·考吉尔、贾森·达纳（Jason Dana）、丹·戈德斯坦（Dan Goldstein）、哈罗德·戈德斯坦（Harold Goldstein）、布莱恩·霍夫曼（Brian Hoffman）、艾伦·克鲁格（Alan Krueger）、迈克尔·莫布森（Michael Mauboussin）、埃米丽·帕特南-霍斯坦（Emily Putnam-Horstein）、查尔斯·谢尔鲍姆（Charles Scherbaum）、安妮-劳尔·塞利耶（Anne-Laure Sellier）和昭田裕一（Yuichi Shoda）所提供的专业知识。

我们还要感谢一大批多年来给予我们支持的研究者们，包括莎瑞雅·巴德瓦杰（Shreya Bhardwaj）、乔西·菲舍尔（Josie Fisher）、罗希特·戈亚尔（Rohit Goyal）、妮科尔·格拉贝尔（Nicole Grabel）、安德鲁·海因里希（Andrew Heinrich）、梅根·约翰逊（Meghann Johnson）、索菲·梅塔（Sophie Mehta）、伊莱·纳克曼尼（Eli Nachmany）、威廉·瑞安（William Ryan）、伊夫琳·舒（Evelyn Shu）、马特·萨默斯（Matt Summers）和诺姆·齐夫-克里斯佩尔（Noam Ziv-Crispel）。在对许多我

们不太擅长的专业领域进行讨论时，正是由于他们的出色表现，才使得本书的偏差更小、噪声更少。

最后，我们这个由跨越两大洲的三位作者组成的合作团队在创作的最佳时机下，还是与挑战不期而遇，因为2020年是非常特殊的一年。如果没有Dropbox和Zoom的技术加持，我们肯定无法完成本书，因此，我们也感谢在这些优秀产品背后默默付出的工作人员。

译者后记

汪祚军

中国科学院心理研究所博士，宁波大学教授

感谢编辑部独具慧眼，能够邀请中国科学院心理研究所李纾研究员领衔本书的翻译工作。李纾老师是国内最早从事行为决策研究的专家，在行为决策领域深耕数十年，其提出的“齐当别”决策模型在心理学界具有重要的影响力。作为行为决策领域的一名研究者、李纾老师曾经的博士生，能够参与本书的翻译工作，我深感荣幸。

在拿到本书的英文原书后，我越读越爱不释手，深深被本书的独创性（阐述了一个非常重要但被普遍忽视的问题）、系统性（介绍了体系化而非碎片化的知识）和可读性所折服。我相信，三位重磅作者的这部诚意之作《噪声》，必将成为行为决策领域一部里程碑式的作品。

偏差和噪声是决策误差的两种核心成分，因此我们会很自然地将两者进行对比。以往在读到其他书中关于“心理偏差”的研究时，我都会产生惊奇和茅塞顿开之感（原来我们有这么多偏差！），但在将这些研究理论应用于实践时，又感觉很零散，不知如何下手。《噪声》这本书则不同，它很集中，整本书只讲了一个概念——噪声，并且借助司法、医疗、商业等十分重要的领域中的众多真实案例和研究成果，深入剖析了噪声的存在情况以及减少噪声的手段。

如果把解决各种心理偏差的方法类比成“工具箱”里的各种工具，那么《噪声》就是一把多功能的“锤子”。有时候工具太多，我们难以取舍，但有了《噪声》这把“锤子”，我们能做的事情就有了很多，至少，我们可以先敲敲看——进行噪声审查。

三位作者讲故事的能力也让我受益良多。用统计术语来说，噪声即

离散度或变异性，一个本来很抽象、晦涩甚至乏味的现象被三位作者用引人入胜的方式诠释出来，并用于解决现实生活中的各种重大问题，尤其是事关生命健康和社会公正的医疗和司法领域的问题。

作为这部皇皇巨著的主要译者之一，我一直承受着巨大的心理压力，很担心这样一部精品会毁在翻译这一步。我曾读过一些翻译较差的图书，并深受其苦。己所不欲，勿施于人。鉴于几次非常不愉快的阅读经历，我下定决心，宁愿多花数倍的时间来啃英文原著，也绝不交付语句不通的译文。我向翻译团队反复强调，我们的目标是“不要被人骂”。

本书由李纾老师精心组建的翻译团队共同完成。从第一部分到第六部分，主要译者分别为：黄元娜（博士，清华大学经济管理学院博士后）、汪祚军、刘洪志（博士，南开大学心理系讲师）、魏子晗（博士，天津师范大学心理学部讲师）、岳灵紫（硕士，友邦人寿销售经理）、郑党（博士，中国儿童中心助理研究员）。虽然本书的各个部分都有专人负责，但李纾老师和我对全书的译稿进行了细致的审校和修改。由于时间仓促，翻译工作中难免存在疏漏和不足，在此也敬请读者谅解。

多次修改和反复打磨，一直希望能达成“不要被人骂”的目标，然而，临近交稿，又觉得“有人骂才好”。鲁迅先生说的那种“连眼珠子也不转过去”的不骂，才是对作品的轻蔑和侮辱。只要有读者肯“转下眼珠”“找个茬”，便说明有人认真看了。若再有决策领域的同行能受本书启发，看出诸如“展望理论”（又译为“预期理论”）背后存在着的后续科学问题，那我们付出的时间和精力就是值得的。

引言 偏差与噪声，人类判断的两类错误

[射击场只是一种隐喻](#)：1778年，瑞士数学家丹尼尔·伯努利（Daniel Bernoulli）在一篇关于评估问题的论文中使用弓和箭进行了相同的类比。Bernoulli, “The Most Probable Choice Between Several Discrepant Observations and the Formation Therefrom of the Most Likely Induction,” *Biometrika* 48, no. 1–2 (June 1961) : 3–18。

[儿童监护权决策](#)：Joseph J. Doyle Jr., “Child Protection and Child Outcomes: Measuring the Effects of Foster Care,” *American Economic Review* 95, no. 5 (December 2007) : 1583–1610。

[预测中存在噪声](#)：Stein Grimstad, Magne Jørgensen, “Inconsistency of Expert Judgment-Based Estimates of Software Development Effort,” *Journal of Systems and Software* 80, no. 11 (2007) : 1770–1777。

[庇护权决策](#)：Andrew I. Schoenholtz, Jaya Ramji-Nogales, Philip G. Schrag, “Refugee Roulette: Disparities in Asylum Adjudication,” *Stanford Law Review* 60, no. 2 (2007) 。

[专利权授予决策](#)：Mark A. Lemley, Bhaven Sampat, “Examiner Characteristics and Patent Office Outcomes,” *Review of Economics and Statistics* 94, no. 3 (2012) : 817–827; Iain Cockburn, Samuel Kortum, Scott Stern, “Are All Patent Examiners Equal? The Impact of Examiner Characteristics,” working paper 8980, June 2002; Michael D. Frakes, Melissa F. Wasserman, “Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents? Evidence from Microlevel Application Data,” *Review of Economics and Statistics* 99, no. 3 (July 2017) : 550–563。

第一部分 寻找噪声

第1章 犯罪和充满噪声的判罚

帮助创立该组织的初衷：Marvin Frankel, *Criminal Sentences: Law Without Order*, 25 Inst. for Sci. Info. Current Contents / Soc. & Behavioral Scis.: This Week's Citation Classic 14, 2A-6 (June 23, 1986)。

几乎完全不受制衡的权力：Marvin Frankel, *Criminal Sentences: Law Without Order* (New York: Hill and Wang, 1973), 5。

每天都在发生残酷专断的行为：Frankel, *Criminal Sentences*, 103。

法治而非人治的社会：Frankel, *Criminal Sentence*, 5。

独断专行的产物：Frankel, *Criminal Sentences*, 11。

某种形式的数字，或其他客观的评分：Frankel, *Criminal Sentence*, 114。

使用计算机作为量刑中有序思考的辅助工具：Frankel, *Criminal Sentence*, 115。

成立一个量刑委员会：Frankel, *Criminal Sentence*, 119。

缺少共识是常态：Anthony Partridge, William B. Eldridge, *The Second Circuit Sentence Study: A Report to the Judges of the Second Circuit August 1974* (Washington, DC: Federal Judicial Center, August 1974) :9。

各种量刑之间的差异“令人震惊”：US Senate, “Comprehensive Crime Control Act of 1983: Report of the Committee on the Judiciary, United States Senate, on S. 1762, Together with Additional and Minority Views” (Washington, DC: US Government Printing Office, 1983) . Report No.98-225。

[贩卖海洛因的毒品贩子](#)：Anthony Partridge, Eldridge, *Second Circuit Sentence Study*, A-11。

[银行抢劫犯](#)：Partridge, Eldridge, *Second Circuit Sentence Study*, A-9。

[敲诈勒索案](#)：Partridge, Eldridge, A-5–A-7。

[对47名法官进行了一项调查](#)：William Austin, Thomas A. Williams III, “A Survey of Judges’ Responses to Simulated Legal Cases: Research Note on Sentencing Disparity,” *Journal of Criminal Law & Criminology* 68 (1977) : 306。

[一项更大规模的研究](#)：John Bartolomeo et al., “Sentence Decisionmaking: The Logic of Sentence Decisions and the Extent and Sources of Sentence Disparity,” *Journal of Criminal Law and Criminology* 72, no. 2 (1981)。完整讨论见第6章，也见Senate Report, 44。

[如果法官处于饥饿状态](#)：Shai Danziger, Jonathan Levav, Liora Avnaim-Pesso, “Extraneous Factors in Judicial Decisions,” *Proceedings of the National Academy of Sciences of the United States of America* 108, no. 17 (2011) : 6889–6892。

[青少年法庭判决](#)：Ozkan Eren, Naci Mocan, “Emotional Judges and Unlucky Juveniles,” *American Economic Journal: Applied Economics* 10, no. 3 (2018) : 171–205。

[当本地足球队在周末输掉比赛后，法官在接下来的星期一会做出更严厉的判决](#)：Daniel L. Chen, Markus Loecher, “Mood and the Malleability of Moral Reasoning: The Impact of Irrelevant Factors on Judicial Decisions,” *SSRN Electronic Journal* (September 21, 2019) : 1–70。

[在自己的生日当天可能会更宽容](#)：Daniel L. Chen, Arnaud Philippe, “Clash of Norms: Judicial Leniency on Defendant Birthdays,” (2020)。

[甚至像外界温度这种无关紧要的信息也会影响法官的决策](#)：Anthony Heyes, Soodeh Saberian, “Temperature and Decisions: Evidence from 207 000 Court Cases,” *American Economic Journal: Applied Economics* 11, no. 2 (2018) : 238–265。

[不受约束的自由裁量权](#)：Senate Report, 38。

[“过于悬殊”的量刑差异](#)：Senate Report, 38。

[美国最高法院法官斯蒂芬·布雷耶试图通过指出委员会内部存在的棘手分歧来为过去的做法辩护](#)：这句话出自杰弗里·罗森（Jeffrey Rosen）的作品，“Breyer Restraint,” *New Republic*, July 11, 1994, at 19, 25。

[必须向法院证明这样做的合理性](#)：United States Sentencing Commission, Guidelines Manual (2018) 。

[它减少了由于量刑法官身份的偶然性而导致的判决中出现的净差异](#)：

James M. Anderson, Jeffrey R. Kling, Kate Stith, “Measuring Interjudge Sentencing Disparity: Before and After the Federal Sentencing Guidelines,” *Journal of Law and Economics* 42, no. S1 (April 1999) : 271–308。

[美国量刑委员会对量刑指南的效果进行了详尽的研究](#)：US Sentencing Commission, *The Federal Sentencing Guidelines: A Report on the Operation of the Guidelines System and Short-Term Impacts on Disparity in Sentencing, Use of Incarceration, and Prosecutorial Discretion and Plea Bargaining*, vols. 1 & 2 (Washington, DC: US Sentencing Commission, 1991) 。

[另一项研究表明，1986—1987年，法官之间在刑期长短上的预期差异为4.9个月；而1988—1993年，这一数字下降至3.9个月](#)：Anderson, Kling, Stith, “Interjudge Sentencing Disparity”。

[一项涵盖了不同时期数据的独立研究](#)：Paul J. Hofer, Kevin R. Blackwell, R. Barry Ruback, “The Effect of the Federal Sentencing Guidelines on Inter-Judge Sentencing Disparity,” *Journal of Criminal Law and Criminology* 90 (1999) : 239, 241。

[我们不能对案件的细节视而不见，而是要有洞察力](#)：Kate Stith, José Cabranes, *Fear of Judging: Sentencing Guidelines in the Federal Courts*, Chicago: University of Chicago Press, 1998: 79。

[直到2005年，美国最高法院才取消了该指南](#)：543 U.S. 220 (2005)。

[75%的法官更喜欢建议性制度](#)：US Sentencing Commission, “Results of Survey of United States District Judges, January 2010 through March 2010” (June 2010) (question 19, table 19)。

[她的核心发现是，法官之间的差异在2005年后明显增加](#)：Crystal Yang, “Have Interjudge Sentencing Disparities Increased in an Advisory Guidelines Regime? Evidence from Booker,” *New York University Law Review* 89 (2014) : 1268–1342; p. 1278, 1334。

第2章 系统噪声，给人达成一致的错觉

[我们称这个实验为噪声审查](#)：该公司的高管们认真构建了几个具有代表性的案例描述，这些案例与员工每天都要处理的风险和索赔案例类似。他们将6个案例分别交给财产和意外伤害部门的理赔员处理，并将另外4个案例交给专门从事金融风险的核保员进行处理。在正常工作日中，员工获得了半天的时间来评估其中的两三个案例，为了检查他们判断中存在的变异性，研究人员事先并没有告知这些员工本研究的目的，并且每个人的评估工作都是独立进行的。我们一共得到了来自48位核保员的86个判断和来自68位理赔员的113个判断。

[天真的现实主义](#)：Dale W. Griffin, Lee Ross, “Subjective Construal, Social Inference, and Human Misunderstanding,” *Advances in Experimental Social Psychology* 24 (1991) : 319–359; Robert J. Robinson, Dacher Keltner, Andrew Ward, Lee Ross, “Actual Versus Assumed Differences in Construal: ‘Naive Realism’ in Intergroup Perception and Conflict,” *Journal of Personality and Social Psychology* 68, no. 3 (1995) : 404; Lee Ross, Andrew Ward, “Naive Realism in Everyday Life: Implications for Social Conflict and Misunderstanding,” *Values and Knowledge* (1997) 。

第二部分 你的大脑是一种测量工具

[标准差是测量差异的最常见指标](#)：一组数字的标准差是由另一个被称为“变异性”的统计量演化而来的。要想计算变异性，我们首先需要获得各个数字偏离平均数的值的分布情况，然后对这些偏离值进行取平方操作。变异性就是这一组数字偏离值的平方的平均数，而标准差是变异性的平方根。

第4章 什么是判断

[在葡萄酒比赛中，评委们对哪种葡萄酒应该获奖可能会分歧很大](#)：R. T. Hodgson, “An Examination of Judge Reliability at a Major U.S. Wine Competition,” *Journal of Wine Economics* 3, no. 2 (2008) : 105–113。

[这种权衡是通过评估性判断实现的](#)：决策领域的学者将决策定义为从不同选项中进行选择的过程，而将定量的判断作为决策的特例，因为此类判断中包含一系列连续可选的选项。从这一角度来看，判断是决策的特例。我们这里的看法有所不同：我们将从不同选项中做出选择的过程看作对每一个选项进行评估性判断的过程。换言之，我

们将决策作为判断的特例。

第5章 测量误差，噪声与偏差的代价一样大

[最小平方法是1795年由高斯发明的](#)：最小平方法常被称作“最小二乘法”，由阿德里恩·玛里·埃·勒让德尔在1805年首次提出。高斯宣称，他自己早在10年前就首次使用过这个概念，随后他基于这一概念提出误差理论和以他自己的名字命名的正态误差曲线。关于是谁最早提出最小平方法，学界已经有大量讨论，历史学家们倾向于相信高斯的说法 [Stephen M. Stigler, “Gauss and the Invention of Least Squares,” *Annals of Statistics* 9 (1981) : 465–474; Stephen M. Stigler, *The History of Statistics: The Measurement of Uncertainty Before 1900* (Cambridge, MA: Belknap Press of Harvard University Press, 1986)] 。

[用简单的数学公式表示](#)：我们将噪声定义为误差的标准差，因此噪声的平方就是误差的变异性。变异性的定义是“平方的平均数减去平均数的平方”。既然偏差是平均误差，“平均数的平方”就是偏差的平方。因此：噪声²=MSE -偏差²。

[关于这一点人们的直觉恰恰相反](#)：Berkeley J. Dietvorst, Soaham Bharti, “People Reject Algorithms in Uncertain Decision Domains Because They Have Diminishing Sensitivity to Forecasting Error,” *Psychological Science* 31, no. 10 (2020) : 1302–1314。

第6章 噪声分析：所有判断都存在3类噪声

[非常详细的噪声审查](#)：Kevin Clancy, John Bartolomeo, David Richardson, Charles Wellford, “Sentence Decisionmaking: The Logic of Sentence Decisions and the Extent and Sources of Sentence Disparity,” *Journal of Criminal Law and Criminology* 72, no. 2 (1981) : 524–554; INSLAW,

Inc. et al., “Federal Sentencing: Towards a More Explicit Policy of Criminal Sanctions III-4,” (1981)。

研究人员向这些法官呈现16起案件的详细文件，并要求法官们做出判决：判刑可以包括监狱服刑时间、监视居住时间和罚款等的任何一种组合形式。简而言之，我们这里主要关注量刑的一种主要组成部分——入狱服刑时间，而不考察另外两个组成部分。

这种差异经常被称为“偏差”：在多重案例、多个判断者的情境中，我们在第5章介绍的误差方程的扩展版中引入了一个概念，它反映了这种变异性。具体而言，如果我们将总体偏差（grand bias）定义为所有案例的平均误差，并且这种误差在不同案例中并不相同，则存在案例偏差的变异。这个方程就变成了： $MSE = \text{总体偏差}^2 + \text{案例偏差的变异性} + \text{系统噪声}^2$ 。

平均刑期为7年：本章中所提及的数字来自原始研究中的如下部分。首先，作者称，被告所犯罪行和被告的主要影响占总体变异性的45% [John Bartolomeo et al., “Sentence Decisionmaking: The Logic of Sentence Decisions and the Extent and Sources of Sentence Disparity,” *Journal of Criminal Law and Criminology* 72, no. 2 (1981), table 6]。然而，我们这里更宽泛地关注每个案例的影响，包括向法官呈现的案例的所有特征，例如被告是否有犯罪记录，在犯罪过程中是否使用武器。根据我们的定义，所有这些特征都是真实的案例变异，而不是噪声。相应地，我们将案例变异中不同案例特征之间的交互作用也进行了重新整合（这些解释了11%的总体变异性；见 Bartolomeo et al., table 10）。这样，案例变异占总体变异的56%，法官的主要影响（水平噪声）为21%，总体变异中的交互作用为23%。因此，总体变异中有44%是系统噪声。公正量刑的变异性可以通过表格（见Bartolomeo et al., 89）里每一个案例的平均刑期来计算：变异性为15。如果这占到总体变异的56%，那么总体变异为

26.79，系统噪声为11.79。变异的平方根即所呈现案例的标准差，为3.4年。法官的主效应或称水平噪声占总体变异的21%。这一变异的平方根是法官水平噪声的标准差，为2.4年。

系统噪声是3.4年：这一数值是所有16起案件变异性的平均值的平方根。对它的计算见前一个注释。

简单的相加逻辑：相加性假设认为，法官的严厉程度会在总体刑期上增加一个常数。这一假设不太可能是正确的：法官的严厉程度更有可能增加一个与平均刑期成比例的数值。这一问题在原始报告中被忽略了，因而无法评估它的重要性。

审判中出现模式化的差别：Bartolomeo et al., “Sentence Decision-making,” 23。

模式噪声和水平噪声的贡献几乎相同：以下方程也成立：系统噪声²=水平噪声²+模式噪声²。图6-1显示，系统噪声是3.4年，水平噪声是2.4年。这意味着模式噪声也约为2.4年。这一计算只是作为一个例子——四舍五入产生的误差导致实际的值会略有不同。

第7章 情境噪声，无时无刻不在影响着我们的判断

专家先后两次品尝了同一种葡萄酒：R. T. Hodgson, “An Examination of Judge Reliability at a Major U.S. Wine Competition,” *Journal of Wine Economics* 3, no. 2 (2008) : 105–113.

经验丰富的软件顾问：Stein Grimstad, Magne Jorgensen, “Inconsistency of Expert Judgment-Based Estimates of Software Development Effort,” *Journal of Systems and Software* 80, no. 11 (2007) : 1770–1777。

他们倾向于与自己保持一致：Robert H. Ashton, “A Review and Analysis of Research on the Test-Retest Reliability of Professional Judgment,” *Journal of Behavioral Decision Making* 294, no. 3 (2000) : 277–

294。作者偶然注意到，这41项研究中，没有一项是被设计用于检测情境噪声的，“在所有研究中，关于稳定性的测量都是其他研究目标的副产品”（Ashton, 279）。这句话表明，直到近期人们才开始关注情境噪声。

正确答案是约32%：Central Intelligence Agency, *The World Factbook*（Washington, DC: Central Intelligence Agency, 2020）。这一数值包括所有能从空中识别的机场、铺设或未铺设的跑道、关闭或废弃的设施。

爱德华·沃尔和哈罗德·帕什勒：Edward Vul, Harold Pashler, “Crowd Within: Probabilistic Representations Within Individuals”。

第一次的答案比第二次的答案更接近真实值：James Surowiecki, *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*（New York: Doubleday, 2004）。

但它会产生更少的噪声：平均判断的标准差（我们对噪声的测量）与判断者人数的平方根同比降低。

你对同一个问题做出两次回答带来的好处，是向另一独立个体寻求建议时所获好处的1/10：Vul and Pashler, “Crowd Within,” 646。

斯蒂芬·赫佐格和拉尔夫·赫维格：Stefan M. Herzog, Ralph Hertwig, “Think Twice and Then: Combining or Choosing in Dialectical Bootstrapping?,” *Journal of Experimental Psychology: Learning, Memory, and Cognition* 40, no. 1（2014）：218–232。

被试的反应是从一个内部的概率分布中抽取的：Vul, Pashler, “Measuring the Crowd Within,” 647。

约瑟夫·福加斯：Joseph P. Forgas, “Affective Influences on Interpersonal Behavior,” *Psychological Inquiry* 13, no. 1（2002）：1–28。

[对于同一个微笑，拥有积极情绪的人看到友好](#)：Forgas, “Affective Influences,” 10。

[谈判过程中从情绪良好转向愤怒的谈判者也可能会获得更好的结果](#)：A. Filipowicz, S. Barsade, S. Melwani, “Understanding Emotional Transitions: The Interpersonal Consequences of Changing Emotions in Negotiations,” *Journal of Personality and Social Psychology* 101, no. 3 (2011) : 541–556。

[实验人员要求参与者阅读一篇简短的哲学论文](#)：Joseph P. Forgas, “She Just Doesn’t Look like a Philosopher...? Affective Influences on the Halo Effect in Impression Formation,” *European Journal of Social Psychology* 41, no. 7 (2011) : 812–817。

[戈登·彭尼库克及其同事开展了一系列研究，来考察人们对毫无意义、看似深奥实则虚假的陈述的反应](#)：Gordon Pennycook, James Allan Cheyne, Nathaniel Barr, Derek J. Koehler, Jonathan A. Fugelsang, “On the Reception and Detection of Pseudo-Profound Bullshit,” *Judgment and Decision Making* 10, no. 6 (2015) : 549–563。

[自哈里·法兰克福之后，废话已经成为一个术语](#)：Harry Frankfurt, *On Bullshit* (Princeton, NJ: Princeton University Press, 2005) 。

[他们可能会被看似令人印象深刻的断言所打动](#)：Pennycook et al., “Pseudo-Profound Bullshit,” 549。

[诱发良好的情绪会让人们更容易接受废话，以及更容易上当受骗](#)：Joseph P. Forgas, “Happy Believers and Sad Skeptics? Affective Influences on Gullibility,” *Current Directions in Psychological Science* 28, no. 3 (2019) : 306–313。

[处于不良情绪中的目击者在看到这些具有误导性的信息时，更有可能忽视它们，从而避免做出虚假指证](#)：Joseph P. Forgas, “Mood Effects

on Eyewitness Memory: Affective Influences on Susceptibility to Misinformation,” *Journal of Experimental Social Psychology* 41, no. 6 (2005) : 574–588。

[天桥难题](#)：Piercarlo Valdesolo, David Desteno, “Manipulations of Emotional Context Shape Moral Judgment,” *Psychological Science* 17, no. 6 (2006) : 476–477。

[医生在漫长的一天结束时开阿片类药物的可能性显著增加](#)：Hannah T. Neprash, Michael L. Barnett, “Association of Primary Care Clinic Appointment Time with Opioid Prescribing,” *JAMA Network Open* 2, no. 8 (2019) ; Lindsey M. Philpot, Bushra A. Khokhar, Daniel L. Roellinger, Priya Ramar, Jon O. Ebbert, “Time of Day Is Associated with Opioid Prescribing for Low Back Pain in Primary Care,” *Journal of General Internal Medicine* 33 (2018) : 1828。

[在一天将要结束时，医生开抗生素的可能性更大](#)：Jeffrey A. Linder, Jason N. Doctor, Mark W. Friedberg, Harry Reyes Nieva, Caroline Birks, Daniella Meeker, Craig R. Fox, “Time of Day and the Decision to Prescribe Antibiotics,” *JAMA Internal Medicine* 174, no. 12 (2014) : 2029–2031。

[开流感疫苗的可能性较小](#)：Rebecca H. Kim, Susan C. Day, Dylan S. Small, Christopher K. Snider, Charles A. L. Rareshide, Mitesh S. Patel, “Variations in Influenza Vaccination by Clinic Appointment Time and an Active Choice Intervention in the Electronic Health Record to Increase Influenza Vaccination,” *JAMA Network Open* 1, no. 5 (2018) : 1–10。

[不好的天气与记忆力的增强有一定的相关性](#)：Joseph P. Forgas, Liz Goldenberg, Christian Unkelbach, “Can Bad Weather Improve Your Memory? An Unobtrusive Field Study of Natural Mood Effects on Real-

Life Memory,” *Journal of Experimental Social Psychology* 45, no. 1 (2008) : 254–257。阳光明媚的天气会影响股市走向：David Hirshleifer, Tyler Shumway, “Good Day Sunshine: Stock Returns and the Weather,” *Journal of Finance* 58, no. 3 (2003) : 1009–1032。

[云让书呆子看起来不错](#)：Uri Simonsohn, “Clouds Make Nerds Look Good: Field Evidence of the Impact of Incidental Factors on Decision Making,” *Journal of Behavioral Decision Making* 20, no. 2 (2007) : 143–152。

[赌徒谬误](#)：Daniel Chen et al., “Decision Making Under the Gambler’s Fallacy: Evidence from Asylum Judges, Loan Officers, and Baseball Umpires,” *Quarterly Journal of Economics* 131, no. 3 (2016) : 1181–1242。

[在美国，当前面两起案件获得庇护法官的批准时，下一个庇护申请获得批准的可能性会降低19%](#)：Jaya Ramji-Nogales, Andrew I. Schoenholtz, Philip Schrag, “Refugee Roulette: Disparities in Asylum Adjudication,” *Stanford Law Review* 60, no. 2 (2007) 。

[迈克尔·卡哈纳及其同事研究了记忆的表现](#)：Michael J. Kahana et al., “The Variability Puzzle in Human Memory,” *Journal of Experimental Psychology: Learning, Memory, and Cognition* 44, no. 12 (2018) : 1857–1863。

第8章 群体是如何放大噪声的

[马修·萨尔加尼克和他的合作者开展了一项大型音乐下载的研究](#)：

Matthew J. Salganik, Peter Sheridan Dodds, Duncan J. Watts, “Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market,” *Science* 311 (2006) : 854–856。另见：Matthew Salganik, Duncan Watts, “Leading the Herd Astray: An Experimental

Study of Self-Fulfilling Prophecies in an Artificial Cultural Market,” *Social Psychology Quarterly* 71 (2008) : 338–355; Matthew Salganik, Duncan Watts, “Web-Based Experiments for the Study of Collective Social Dynamics in Cultural Markets,” *Topics in Cognitive Science* 1 (2009) : 439–468。

[流行程度会自我强化](#)：Salganik, Watts, “Leading the Herd Astray”。

[在其他领域也出现了类似的结果](#)：Michael Macy et al., “Opinion Cascades and the Unpredictability of Partisan Polarization,” *Science Advances* (2019) : 1–8。另见：Helen Margetts et al., *Political Turbulence* (Princeton: Princeton University Press, 2015) 。

[社会学家迈克尔·梅西](#)：Michael Macy et al., “Opinion Cascades”。

[人们在网上如何对各种评论做出判断](#)：Lev Muchnik et al., “Social Influence Bias: A Randomized Experiment,” *Science* 341, no. 6146 (2013) : 647–651。

[有些研究已经表明了这一点](#)：Jan Lorenz et al., “How Social Influence Can Undermine the Wisdom of Crowd Effect,” *Proceedings of the National Academy of Sciences* 108, no. 22 (2011) : 9020–9025。

[我们来看一个实验，该实验比较了现实世界中的陪审团和“统计中的陪审团”](#)：Daniel Kahneman, David Schkade, Cass R. Sunstein, “Shared Outrage and Erratic Awards: The Psychology of Punitive Damages,” *Journal of Risk and Uncertainty* 16 (1998) : 49–86。

[由他们组成500多个6人一组的陪审团](#)：David Schkade, Cass R. Sunstein, Daniel Kahneman, “Deliberating about Dollars: The Severity Shift,” *Columbia Law Review* 100 (2000) : 1139–1175。

第三部分 预测性判断中的噪声

一致性比率（PC）：一致性比率与肯德尔和谐系数（Kendall's W）关联密切。

成年男性脚的尺码与身高的PC值为71%：Kanwal Kamboj et al., “A Study on the Correlation Between Foot Length and Height of an Individual and to Derive Regression Formulae to Estimate the Height from Foot Length of an Individual,” *International Journal of Research in Medical Sciences* 6, no. 2（2018）：528。

相关系数和一致性比率（PC）的对应关系：一致性比率的计算基于这样的假设：联合分布（joint distribution）是二元正态分布。表1中的数值就是基于该假设的近似值。感谢朱利安·帕里斯绘制表格。

第9章 判断与模型，简单的模型普遍优于人类判断

一项关于绩效预测的真实研究：Martin C. Yu, Nathan R. Kuncel, “Pushing the Limits for Judgmental Consistency: Comparing Random Weighting Schemes with Expert Judgments,” *Personnel Assessment and Decisions* 6, no. 2（2020）：1–10。本书所描述的专家们的相关系数0.15是指3个实验样本（总共847个样本）的相关系数的非加权平均值。真实实验中的相关系数与此处的这一简化的相关系数在某些方面有所不同。

它是对各种预测因素的平均值进行加权后获得预测分数的方法：对平均值进行加权计算的先决条件是所有预测因素都必须用可比较的测量单位测量。我们介绍的这个例子是满足这个先决条件的，所有的预测因素都是用0到10分的量表测量的。但在一些情况下，这一先决条件可能无法满足。比如，绩效表现的预测因素包括面试官给的0到10分的评分、相关工作经验的时长，以及工作技能的笔试成绩。多元回归分析程序在整合这些变量之前，会把这些变量先转换为标准分。标准分测量的是一个观测值与群体平均值之间的距离，它以

标准差为单位。例如，如果技能测试的平均分是55分，标准差是8分，那么标准分为+1.5就意味着技能测试为67分。需要注意的是，把每个个体数据标准化会消除所有平均值之间的误差痕迹，或者说会消除个体判断变异中的误差痕迹。

你可能认为，与目标变量相关性越密切的预测因素，其权重也应该越

大：多元回归最主要的特征是每个预测因素的最优权重都依赖于其他预测因素。如果一个预测因素与另一个预测因素高相关，两个预测因素的权重就不会都很大——这就跟变量被重复计算了一样。

判断和决策研究的主力军：Robin M. Hogarth, Natalia Karelaia, “Heuristic and Linear Models of Judgment: Matching Rules and Environments,” *Psychological Review* 114, no. 3 (2007) : 734。

二者都具有如下一些简单的结构：在这里的讨论主要基于透镜判断模型的研究框架，该模型被广泛用于人员测评的情境中。参见：Kenneth R. Hammond, “Probabilistic Functioning and the Clinical Method,” *Psychological Review* 62, no. 4 (1955) : 255–262; Natalia Karelaia, Robin M. Hogarth, “Determinants of Linear Judgment: A Meta-Analysis of Lens Model Studies,” *Psychological Bulletin* 134, no. 3 (2008) : 404–426。

保罗·梅尔：Paul E. Meehl, *Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence* (Minneapolis: University of Minnesota Press, 1954) 。

弗洛伊德的照片：Paul E. Meehl, *Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence* (Northvale, NJ: Aronson, 1996) 序言部分。

梅尔不仅是一位学术研究人员，还是一位有着丰富临床经验的精神分析学派的心理咨询师：“Paul E. Meehl,” in Ed Lindzey (Ed.) , A

History of Psychology in Autobiography, 1989。

“大量且一致”：“Paul E. Meehl,” in *A History of Psychology in Autobiography*, ed. Ed Lindzey (Washington, DC: American Psychological Association, 1989), 362。

一项发表于2000年的对136项研究的综述：William M. Grove et al., “Clinical Versus Mechanical Prediction: A Meta-Analysis,” *Psychological Assessment* 12, no. 1 (2000) : 19–30。

人类在判断时还具有不对等的优势，因为他们可以获取未提供给计算机模型的“私人”信息：William M. Grove, Paul E. Meehl, “Comparative Efficiency of Informal (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical-Statistical Controversy,” *Psychology, Public Policy, and Law* 2, no. 2 (1996) : 293–323。

20世纪60年代后期，戈德堡基于霍夫曼的早期工作，开始研究用于描述个体判断行为的统计模型：Lewis Goldberg, “Man Versus Model of Man: A Rationale, plus Some Evidence, for a Method of Improving on Clinical Inferences,” *Psychological Bulletin* 73, no. 6 (1970) : 422–432。

台球专家们在描述某一杆如何进球时，表现得就好像他们解开了复杂的方程一样，然而实际上他们并未真的这样做：Milton Friedman, Leonard J. Savage, “The Utility Analysis of Choices Involving Risk,” *Journal of Political Economy* 56, no. 4 (1948) : 279–304。

尽管不是完全相关，但这种相关性已经足以支持所谓的“假设”理论了：Karelaia, Hogarth, “Determinants of Linear Judgment,” 411, table 1。

早期一项关于预测学生毕业成绩的研究证实了戈德堡的结论：Nancy Wiggins, Eileen S. Kohen, “Man Versus Model of Man Revisited: The

Forecasting of Graduate School Success,” *Journal of Personality and Social Psychology* 19, no. 1 (1971) : 100–106。

[一项对近50年研究成果的综述性研究也得出同样的结论](#)：Karelaia, Hogarth, “Determinants of Linear Judgment”。

[从你的判断中消除噪声通常会提高你的预测准确性](#)：可以对相关系数进行校正以解决预测因素信度不高的问题，即，衰减校正。公式是：校正的 $r_{xy}=r_{xy}/\sqrt{r_{xx}}$ ，其中 r_{xx} 是信度系数（预测因素的真实变异占其观测变异的比例）。

[马丁·于和内森·昆塞尔报告了一项比戈德堡更激进的研究](#)：Yu, Kuncel, “Judgmental Consistency”。

[随机公式](#)：我们将在下一章更详细地讨论等权模型与随机权重模型。权重会被限定在一定的数字范围内，并且会被限定使用相应的符号。

第10章 无噪声的规则

[这些均等权重模型的准确性与最合适的回归模型差不多，且远胜于诊断性判断](#)：Robyn M. Dawes, Bernard Corrigan, “Linear Models in Decision Making,” *Psychological Bulletin* 81, no. 2 (1974) : 95–106。罗宾·道斯和伯纳德·科里根也提议使用随机权重。我们在第9章中介绍的管理绩效预测的研究也体现了这一思想。

[“与统计直觉相悖”](#)：Jason Dana, “What Makes Improper Linear Models Tick?,” in *Rationality and Social Responsibility: Essays in Honor of Robyn M. Dawes*, ed. Joachim I. Krueger, 71–89 (New York: Psychology Press, 2008) , 73。

[许多其他研究也得到了相似的结果](#)：Jason Dana, Robyn M. Dawes, “The Superiority of Simple Alternatives to Regression for Social Sciences Prediction,” *Journal of Educational and Behavior Statistics*

29（2004）：317–331; Dana, “What Makes Improper Linear Models Tick?”。

它并不重要：Howard Wainer, “Estimating Coefficients in Linear Models: It Don’t Make No Nevermind,” *Psychological Bulletin* 83, no. 2（1976）：213–217。

我们不需要比我们的测量更精确的模型：Dana, “What Makes Improper Linear Models Tick?,” 72。

它与结果的相关系数将为0.25（PC=58%）：Martin C. Yu, Nathan R. Kuncel, “Pushing the Limits for Judgmental Consistency: Comparing Random Weighting Schemes with Expert Judgments,” *Personnel Assessment and Decisions* 6, no. 2（2020）：1–10。与前一章一样，这里报告的相关系数是3个实验样本的相关系数的非加权平均值。在3个实验样本中，这一比较关系依然成立，临床专家判断的效度分别是0.17、0.16和0.13，而等权模型的效度分别是0.19、0.33和0.22。

稳定之美：Robyn M. Dawes, “The Robust Beauty of Improper Linear Models in Decision Making,” *American Psychologist* 34, no. 7（1979）：571–582。

应用均等权重模型所需的全部技巧是决定要关注哪些变量，并知道如何将
这些变量进行叠加：Dawes, Corrigan, “Linear Models in Decision Making,” 105。

一个研究团队于2020年发表了一项研究成果。他们将简约模型应用于一
系列现实问题：Jongbin Jung, Conner Concannon, Ravi Shroff, Sharad Goel, Daniel G. Goldstein, “Simple Rules to Guide Expert Classifications,” *Journal of the Royal Statistical Society, Statistics in Society*, no. 183（2020）：771–800。

[另外一个研究小组研究了一个与上述案例相似但有所不同的司法问题：](#)

Julia Dressel, Hany Farid, “The Accuracy, Fairness, and Limits of Predicting Recidivism,” *Science Advances* 4, no. 1 (2018) : 1–6。

[使用仅两个输入变量的模型就可以达到与使用137个变量的模型相同的](#)

[预测效度](#)：这两个例子都是基于极小的一组变量的线性模型（并且，在保释金的模型中，线性模型权重系数的近似值是基于舍入法求得的，它将模型转换成了一种粗略的运算）。另一个“并非最合适的模型”是单变量法则，它只考虑一个预测因素，而忽略所有其他预测因素。详见：Peter M. Todd, Gerd Gigerenzer, “Précis of Simple Heuristics That Make Us Smart,” *Behavioral and Brain Sciences* 23, no. 5 (2000) : 727–741。

[而大量证据表明，它们与犯罪行为也是紧密相关的](#)：P. Gendreau, T.

Little, C. Goggin, “A Meta-Analysis of the Predictors of Adult Offender Recidivism: What Works!,” *Criminology* 34 (1996) 。

[海量数据集](#)：这里的“量”，应该理解为观测样本量与预测因素的比值。

要实现罗宾·道斯所说的“稳定之美”，意味着这一比例应该至少高于15:1，或者是20:1，才能使最优权重在交叉验证时表现得比单位权重好。通过对多个案例进行研究，罗宾·道斯和贾森·达纳认为，要“体现简单方案的优势”，这一比值应该要高于100:1。

[由塞德希尔·穆来纳森领导的另一个团队](#)：J. Kleinberg, H. Lakkaraju, J.

Leskovec, J. Ludwig, S. Mullainathan, “Human Decisions and Machine Predictions,” *Quarterly Journal of Economics* 133 (2018) : 237–293.

[研究人员利用这一数据来训练一个机器学习算法](#)：利用一部分训练数据

训练该算法，然后用该算法去预测一个随机选定的新数据集的结果，以评估该算法的性能。

[机器学习算法在变量组合中，发现了一些会被线性模型遗漏的重要信](#)

息：Kleinberg et al., “Human Decisions,” 16。

系统噪声包括一些水平噪声，即平均严厉程度之间的差异，但其中大多数（79%）是模式噪声：Gregory Stoddard、Jens Ludwig和Sendhil Mullainathan于2020年6月至7月与本书作者通过电子邮件进行过交流。

哥伦比亚商学院教授博·考吉尔考察了一家大型科技公司招聘软件工程师的情况：B. Cowgill, “Bias and Productivity in Humans and Algorithms: Theory and Evidence from Résumé Screening,”Smith Entrepreneurship Research Conference, College Park, MD, April 21, 2018。

1996年的一篇文章：William M. Grove, Paul E. Meehl, “Comparative Efficiency of Informal (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical-Statistical Controversy,” *Psychology, Public Policy, and Law* 2, no. 2 (1996) : 293–323。

当从人类的建议和算法的建议之间进行选择时，人们通常会选择后者：Jennifer M. Logg, Julia A. Minson, Don A. Moore, “Algorithm Appreciation: People Prefer Algorithmic to Human Judgment,”*Organizational Behavior and Human Decision Processes* 151 (April 2018) : 90–103。

而一旦发现它会犯错误，就不会再信任它：B. J. Dietvorst, J. P. Simmons, C. Massey, “Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err,” *Journal of Experimental Psychology General* 144 (2015) : 114–126。另见：A. Prah, L. Van Swol, “Understanding Algorithm Aversion: When Is Advice from Automation Discounted?,” *Journal of Forecasting* 36 (2017) : 691–702。

[我们希望机器是完美的，如果机器不完美，那就丢弃它](#)：M. T.

Dzindolet, L. G. Pierce, H.P. Beck, L. A. Dawe, “The Perceived Utility of Human and Automated Aids in a Visual Detection Task,” *Human Factors: The Journal of the Human Factors and Ergonomics Society* 44, no. 1 (2002) : 79–94; K. A. Hoff, M. Bashir, “Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust,” *Human Factors: The Journal of the Human Factors and Ergonomics Society* 57, no. 3 (2015) : 407–434; P. Madhavan, D. A. Wiegmann, “Similarities and Differences Between Human–Human and Human–Automation Trust: An Integrative Review,” *Theoretical Issues in Ergonomics Science* 8, no. 4 (2007) : 277–301。

第11章 哪里有预测，哪里就有客观无知

[有关管理决策的研究](#)：E. Dane, M. G. Pratt, “Exploring Intuition and Its Role in Managerial Decision Making,” *Academy of Management Review* 32, no. 1 (2007) : 33–54; Cinla Akinci, Eugene Sadler-Smith, “Intuition in Management Research: A Historical Review,” *International Journal of Management Reviews* 14 (2012) : 104–122; Gerard P. Hodgkinson et al., “Intuition in Organizations: Implications for Strategic Management,” *Long Range Planning* 42 (2009) : 277–297。

[管理决策中的直觉的综述性文章，将直觉定义为对给定行动方案的一种判断](#)：Hodgkinson et al., “Intuition in Organizations,” 279。

[近期一篇有关人才选拔的报告](#)：Nathan Kuncel et al., “Mechanical Versus Clinical Data Combination in Selection and Admissions Decisions: A Meta-Analysis,” *Journal of Applied Psychology* 98, no. 6 (2013) : 1060–1072。有关个人决策的更多内容，参见第24章。

[过分自信是已经被大量研究证明了的一种认知偏差](#)：Don A. Moore, *Perfectly Confident: How to Calibrate Your Decisions Wisely* (New York: HarperCollins, 2020) 。

[对政治和经济趋势发表评论或提供建议](#)：Philip E. Tetlock, *Expert Political Judgment: How Good Is It? How Can We Know?* (Princeton, NJ: Princeton University Press, 2005) , 239, 233。

[对136项研究进行了回顾](#)：William M. Grove et al., “Clinical Versus Mechanical Prediction: A Meta-Analysis,” *Psychological Assessment* 12, no. 1 (2000) : 19–30。

[由塞德希尔·穆来纳森和齐亚德·欧博迈亚完成的另一项研究对心脏病诊断进行了建模](#)：Sendhil Mullainathan, Ziad Obermeyer, “Who Is Tested for Heart Attack and Who Should Be: Predicting Patient Risk and Physician Error,” 2019. NBER Working Paper 26168, National Bureau of Economic Research。

[在充满无知的情况下，否认无知就显得更加诱人](#)：Weston Agor, “The Logic of Intuition: How Top Executives Make Important Decisions,” *Organizational Dynamics* 14, no. 3 (1986) : 5–18; Lisa A. Burke, Monica K. Miller, “Taking the Mystery Out of Intuitive Decision Making,” *Academy of Management Perspectives* 13, no. 4 (1999) : 91–99。

[当算法能够获得更高的准确性时，人们会更愿意相信算法](#)：Poornima Madhavan, Douglas A. Wiegmann, “Effects of Information Source, Pedigree, and Reliability on Operator Interaction with Decision Support Systems,” *Human Factors: The Journal of the Human Factors and Ergonomics Society* 49, no.5 (2007) 。

第12章 常态谷：事情虽无法预测，但可以被理解

一篇非同寻常的论文：Matthew J. Salganik et al., “Measuring the Predictability of Life Outcomes with a Scientific Mass Collaboration,” *Proceedings of the National Academy of Sciences* 117, no. 15 (2020) : 8398–8403。

在比赛的第一阶段，参赛团队可以使用一半样本对应的所有数据，其中也包括6个结果：这项研究只使用了4242个家庭的数据。由于隐私问题，有些脆弱家庭研究的家庭数据在分析时被删除了。

在预测“流离失所”这一事件时，最佳模型的相关系数仅为

0.22 (PC=57%)：为了确保分数的准确性，这场比赛的举办方用了均方误差 (MSE) 进行测量，这和我们在第一部分使用的指标是一样的。为了更好地进行对比，该研究还设定了一个基准模型。其他每个模型的MSE都可以与基准模型进行对比，基准模型是一种“无用的”预测策略：单一值拟合所有数据的预测策略，也就是说，每个样本的预测结果都是训练数据的平均数。为方便起见，我们将它的结果转换成了相关性。MSE和相关性之间的关系是： $r^2 = (\text{Var}(Y) - \text{MSE}) / \text{Var}(Y)$ ，其中， $\text{Var}(Y)$ 是结果变量的方差， $(\text{Var}(Y) - \text{MSE})$ 是可以预测的结果的方差。

一项社会心理学方面的回顾性研究：F. D. Richard et al., “One Hundred Years of Social Psychology Quantitatively Described,” *Review of General Psychology* 7, no. 4 (2003) : 331–663。

对行为和认知科学的708项研究进行的回顾性研究发现，只有3%的研究所报告的相关系数大于0.5：Gilles E. Gignac, Eva T. Szodorai, “Effect Size Guidelines for Individual Differences Researchers,” *Personality and Individual Differences* 102 (2016) : 74–78。

“研究人员必须认识到，虽然他们了解了脆弱家庭的生活轨迹，但每一项预测都不够准确”：注意，该研究刻意使用了一个已有的描述性大数据集，该数据集虽然很大，但它并不是为了预测特定结果而产

生的。这一点与泰特洛克的研究不同，泰特洛克的研究数据使用了所有看起来可能会有预测力的变量。它可能会找到一些不在已有数据集里但能想到的预测因素。因此，基于已有数据集，该研究无法证明无家可归在本质上是不可预测的，也无法证明其他结果是不可预测的。

理解就是描述因果关系，而预测能力就是衡量这一因果关系是否成立的指标：Jake M. Hofman et al., “Prediction and Explanation in Social Systems,” *Science* 355 (2017) : 486–488; Duncan J. Watts et al., “Explanation, Prediction, and Causality: Three Sides of the Same Coin?,” October 2018: 1–14。

有一种思维模式会自发地出现在我们的脑海里：与之对应的是外延性思维（extensional thinking）或非外延性思维，后者也叫意向性思维（intentional thinking）。Amos Tversky, Daniel Kahneman, “Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment,” *Psychological Review* 4 (1983) : 293–315。

这是因为理解现实的过程是回溯性的：Daniel Kahneman, Dale T. Miller, “Norm Theory: Comparing Reality to Its Alternatives,” *Psychological Review* 93, no. 2 (1986) : 136–153。

“后见之明”经典的研究：Baruch Fischhoff, “An Early History of Hindsight Research,” *Social Cognition* 25, no. 1 (2007) : 10–13; Baruch Fischhoff, “Hindsight Is Not Equal to Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty,” *Journal of Experimental Psychology: Human Perception and Performance* 1, no. 3 (1975) : 288。

与因果思维不同，统计思维通常是费力的，它需要的注意力资源只有当系统2思维（缓慢而审慎的思维模式）发挥作用时才能满足：Daniel Kahneman, *Thinking, Fast and Slow*. New York: Farrar, Straus and

Giroux, 2011。

第四部分 噪声是如何产生的

第13章 启发式、偏差与噪声

[《思考，快与慢》一书对该研究项目前40年的研究内容进行了回顾，探讨了能够解释“直觉思维的奇妙与缺陷”的心性机制](#)：Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011)。

[证据表明，两组人所估计的概率相差很小，小到可以忽略不计](#)：我们在这里提醒一下，研究判断偏差的心理学家所做的实验中，每个实验组不会只有5个被试（见图6-2），因为判断是有噪声的，不同实验组的结果并不都像图10-1那样聚集。人们对每种偏差所持的怀疑态度也是不同的，他们并不会彻底忽视掉重要的变量。比如，你可以确定，有很多被试会出现范围不敏感偏差，人们估计甘巴迪3年内仍能继续留任的可能性只比两年内仍能继续留任的可能性高一点点。将这一现象描述为范围不敏感是合适的，因为两者的差异远远小于它应有的水平。

[已经有许多实验表明，人们在这两类问题上会给出相同的答案](#)：Daniel Kahneman et al., eds., *Judgment Under Uncertainty: Heuristics and Biases* (New York: Cambridge University Press, 1982) 第6章；Daniel Kahneman, Amos Tversky, “On the Psychology of Prediction,” *Psychological Review* 80, no. 4 (1973) : 237–251。

[美国企业的CEO离职率大约为每年15%，详见](#)：Steven N. Kaplan, Bernadette A. Minton, “How Has CEO Turnover Changed?,” *International Review of Finance* 12, no. 1 (2012) : 57–87; Dirk Jenter, Katharina Lewellen, “Performance-Induced CEO Turnover,” Harvard

Law School Forum on Corporate Governance, September 2, 2020。

在写《星球大战》第三部的电影剧本《绝地归来》的关键时期，该系列电影的制作人乔治·卢卡斯与他出色的合作者劳伦斯·卡斯丹展开了激烈辩论： J. W. Rinzler, *The Making of Star Wars Return of the Jedi: The Definitive Story* (New York: Del Rey, 2013) , 64。

这个例子说明了另一种类型的偏差，我们称之为结论偏差或者预判：

Cass Sunstein, *The World According to Star Wars* (New York: HarperCollins, 2016) 。

在这种情况下，证据就是有选择性且失真的：我们这里所介绍的是一个简单的、有预判的例子，实际上，即使没有预判，随着证据的累积，也可能会在某一个结论上产生偏差，因为我们总是偏好简洁性和连贯性。当一个暂时性的结论生成后，证实偏差就会使我们更倾向于搜集或解释那些支持该结论的新证据。

你总会倾向于接受任何看起来支持该信念的论点，即使推理是错误的：

这一现象也被称作信念偏差。详见：J. St. B. T. Evans, Julie L. Barson, Paul Pollard, “On the Conflict between Logic and Belief in Syllogistic Reasoning,” *Memory & Cognition* 11, no. 3 (1983) : 295–306。

在一个典型的演示实验中，你可能会看到许多不容易猜出价格的物品：

Dan Ariely, George Loewenstein, Drazen Prelec, “‘Coherent Arbitrariness’: Stable Demand Curves Without Stable Preferences,” *Quarterly Journal of Economics* 118, no. 1 (2003) : 73–105。

锚定效应是一种非常强大的效应，谈判中会经常被刻意用到： Adam D. Galinsky, T. Mussweiler, “First Offers as Anchors: The Role of Perspective-Taking and Negotiator Focus,” *Journal of Personality and Social Psychology* 81, no. 4 (2001) : 657–669。

[这个实验说明了过度一致性偏差](#)：Solomon E. Asch, “Forming Impressions of Personality,” *Journal of Abnormal and Social Psychology* 41, no. 3 (1946) : 258–290, 这篇论文最早使用一系列排列顺序不同的形容词来说明这个现象。

[有一项研究能给我们一些启示](#)：Steven K. Dallas et al., “Don’t Count Calorie Labeling Out: Calorie Counts on the Left Side of Menu Items Lead to Lower Calorie Food Choices,” *Journal of Consumer Psychology* 29, no. 1 (2019) : 60–69。

第14章 匹配，找到与你的预测最精准匹配的共识

[将两个完全无关维度的强度相匹配](#)：S. S. Stevens, “On the Operation Known as Judgment,” *American Scientist* 54, no. 4 (December 1966) : 385–401。我们这里使用的匹配一词，其含义比史蒂文斯这篇研究中该词的含义更广泛，后者只是用来指代等比量表。我们将在15章讨论等比量表。

[系统性判断错误](#)：这个例子首次出现于丹尼尔·卡尼曼的《思考，快与慢》一书中。

[你得到的答案都会是完全相同的两个数字](#)：Daniel Kahneman, Amos Tversky, “On the Psychology of Prediction,” *Psychological Review* 80 (1973) : 237–251。

[《神奇的数字7，加2或减2》](#)：G. A. Miller, “The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information,” *Psychological Review* (1956) : 63–97。

[即便如此，强制进行比较判断还是可能减少噪声](#)：R. D. Goffin, J. M. Olson, “Is It All Relative? Comparative Judgments and the Possible Improvement of Self-Ratings and Ratings of Others,” *Perspectives on*

Psychological Science 6（2011）：48–60。

第15章 选取精确的量表，并多用相对判断

在1998年报告过的一项研究：Daniel Kahneman, David Schkade, Cass Sunstein, “Shared Outrage and Erratic Awards: The Psychology of Punitive Damages,” *Journal of Risk and Uncertainty* 16（1998）：49–86; Cass Sunstein, Daniel Kahneman, David Schkade, “Assessing Punitive Damages（with Notes on Cognition and Valuation in Law）,” *Yale Law Journal* 107, no. 7（May 1998）：2071–2153。该研究的费用由Exxon公司一次性承担。不过Exxon公司并未对研究者支付报酬，也未控制研究数据、研究结果或文章发表等流程。

高度怀疑：A. Keane, P. McKeown, *The Modern Law of Evidence*（New York: Oxford University Press, 2014）。

不太可能发生：Andrew Mauboussin, Michael J. Mauboussin, “If You Say Something Is ‘Likely,’ How Likely Do People Think It Is?,” *Harvard Business Review*, July 3, 2018。

一家汽车经销商被处以400万美元的惩罚性损害赔偿，理由是该公司未告知原告，他们售卖的新宝马车是重新喷过漆的：BMW vs. Gore, 517 U.S. 559（1996）。

这种高相关性支持了愤怒假设，即愤怒情绪是惩罚倾向的主要决定因素。要进一步了解情绪在道德判断中的作用，参见：J. Haidt, “The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment,” *Psychological Review* 108, no. 4（2001）：814–834; Joshua Greene, *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*（New York: Penguin Press, 2014）。

图15-1显示了分析结果：愤怒程度和惩罚倾向之间的高相关性（相关系

数为0.98)支持了愤怒假设,这可能会让你感到很困惑,因为评分中存在大量的噪声。不过,当你想起来相关系数计算的是判断的平均分时,你就不会那么困惑了。如果是100个人判断的平均值,其噪声(判断的标准差)就会变为原来的1/10。如果许多判断被整合在一起,噪声就不会产生很大的影响了,详见第21章。

人们对许多主观体验和态度的强度比例都有着强烈的直觉: S. S. Stevens, *Psychophysics: Introduction to Its Perceptual, Neural and Social Prospects* (New York: John Wiley & Sons, 1975)。

任意连贯性: Dan Ariely, George Loewenstein, Drazen Prelec, “‘Coherent Arbitrariness’: Stable Demand Curves Without Stable Preferences,” *Quarterly Journal of Economics* 118, no. 1 (2003): 73–106, 197。

将赔偿金额转换为排序: 转换成排序的同时必然会损失信息,因为判断间的距离信息并未被保留。假设只有3个案件,一位陪审员建议3个案件的损害赔偿金额分别为1000万美元、200万美元和100万美元。很显然,这位陪审员计划中对第一个案件和第二个案件的赔偿金额之间的差异要远远高于第二个和第三个案件赔偿金额之间的差异。但是一旦被转换成排名,上述差异就变得一样了——两者都只差一名。如果是将判断转换成标准分,就可以解决这一问题。

第16章 模式噪声的构成

大量证据已经证明,确实存在这样一种知觉过程: R. Blake, N. K. Logothetis, “Visual competition,” *Nature Reviews Neuroscience* 3 (2002): 13–21; M. A. Gernsbacher, M. E. Faust, “The Mechanism of Suppression: A Component of General Comprehension Skill,” *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17 (March 1991): 245–262; M. C. Stites, K. D. Federmeier, “Subsequent to Suppression: Downstream Comprehension

Consequences of Noun/Verb Ambiguity in Natural Reading,” *Journal of Experimental Psychology: Learning, Memory, and Cognition* 41 (September 2015) : 1497–1515。

但大多数时候我们的信心都高于应有的程度：D. A. Moore, D. Schatz, “The three faces of overconfidence,” *Social and Personality Psychology Compass* 11, no. 8 (2017) , e12331。

不同的专业人士将会考虑不同方面并相互补充：S. Highhouse, A. Broadfoot, J. E. Yugo, S. A. Devendorf, “Examining Corporate Reputation Judgments with Generalizability Theory,” *Journal of Applied Psychology* 94 (2009) : 782–789。我们要感谢斯科特·海浩斯和艾莉森·布罗德福特，他们为我们提供了原始数据，还要感谢朱利安·帕里斯帮我们做了进一步分析。

早期研究尝试在字典中搜寻能够描述一个人的词语：阿尔波特和奥德贝特（1963）有关人格特质的英文词典的研究，引用于Oliver P. John, Sanjay Strivastava, “The Big-Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives,” *Handbook of Personality: Theory and Research*, 2nd ed., ed. L. Pervin, Oliver P. John (New York: Guilford, 1999) 。

已经很高了：Ian W. Eisenberg, Patrick G. Bissett, A. Zeynep Enkavi et al., “Uncovering the structure of self-regulation through data-driven ontology discovery,” *Nature Communications* 10 (2019) : 2319。

在身体受到威胁时：Walter Mischel, “Toward an Integrative Science of the Person,” *Annual Review of Psychology* 55 (2004) : 1–22。

第17章 噪声源，偏差是引人注目的图形，而噪声是不受我们关注的背景

[现在，你可以看到均方误差如何被分解为偏差以及我们曾讨论过的3种噪声成分的平方](#)：如何分解偏差和噪声没有一个固定的原则，图17-1粗略地展示了某些我们讨论过的真实案例或虚拟案例的分解成分。具体而言，在该图中，偏差和噪声是等量的，就像GoodSell对销售量的预测那样。水平噪声的平方是系统噪声平方的37%，就像在惩罚性损害赔偿案例中那样。如图17-1所示，情境噪声的平方约为模式噪声平方的35%。

[针对专利局的研究发现](#)：参见引言中的参考文献，Mark A. Lemley, Bhaven Sampat, “Examiner Characteristics and Patent Office Outcomes,” *Review of Economics and Statistics* 94, no. 3 (2012) : 817–827。也可参见：Iain Cockburn, Samuel Kortum, Scott Stern, “Are All Patent Examiners Equal? The Impact of Examiner Characteristics,” Working paper 8980, June 2002; Michael D. Frakes, Melissa F. Wasserman, “Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents? Evidence from Microlevel Application Data,” *Review of Economics and Statistics* 99, no. 3 (July 2017) : 550–563。

[儿童保护部门的官员决定将儿童送到寄养家庭或寄养机构的倾向性也不同](#)：Joseph J. Doyle Jr., “Child Protection and Child Outcomes: Measuring the Effects of Foster Care,” *American Economic Review* 95, no. 5 (December 2007) : 1583–1610。

[法官在是否提供庇护的裁决中出现的令人震惊的巨大变异](#)：Andrew I. Schoenholtz, Jaya Ramji-Nogales, Philip G. Schrag, “Refugee Roulette: Disparities in Asylum Adjudication,” *Stanford Law Review* 60, no. 2 (2007) 。

[同一法官在不同情境中对同一案件量刑的平均差异达到2.8年左右](#)：在第6章中，我们估计的交互项的方差占总方差的23%，据此估计出

了这一判刑时间。因为，假设判刑时间服从正态分布，两个随机选择出来的观测值之间的绝对差异应该是1.128个标准差。

[亚历山大·托多罗夫带领普林斯顿大学的一组研究人员设计了一个巧妙的实验范式](#)：J. E. Martinez, B. Labbree, S. Uddenberg, A. Todorov, “Meaningful ‘Noise’: Comparative Judgments Contain Stable Idiosyncratic Contributions”（未发表的手稿）。

[噪声的最主要成分还是稳定的模式噪声](#)：Gregory Stoddard, Jens Ludwig, Sendhil Mullainathan, 他们在2020年6月至7月与本书作者有邮件往来。

[对保释法官的大规模研究](#)：J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, S. Mullainathan, “Human Decisions and Machine Predictions,” *Quarterly Journal of Economics* 133（2018）：237–293。

[运用这些模型来模拟法官对141 833个案例做出的判决](#)：该模型不仅生成了每一位法官对141 833个案件判决的排序，还生成了每一位法官是否判决保释的一个阈值。水平噪声就是这个阈值的变异性，而模式噪声就是案件排序的变异性。

[菲尔·罗森茨威格坚定地认为](#)：Phil Rosenzweig. *Left Brain, Right Stuff: How Leaders Make Winning Decisions*（New York: Public Affairs, 2014）。

第五部分 决策卫生，提升五大人类判断力

第18章 卓越的判断者，卓越的判断力

[该群体由高能力个体组成](#)：Albert E. Mannes et al., “The Wisdom of Select Crowds,” *Journal of Personality and Social Psychology* 107, no. 2（2014）：276–299; Jason Dana et al., “The Composition of Optimally

Wise Crowds,” *Decision Analysis* 12, no. 3 (2015) : 130–143。

[自信启发式](#)：Briony D. Pulford, Andrew M. Colmna, Eike K. Buabang, Eva M. Krockow, “The Persuasive Power of Knowledge: Testing the Confidence Heuristic,” *Journal of Experimental Psychology: General* 147, no. 10 (2018) : 1431–1444。

[也与更好的工作绩效相关](#)：Nathan R. Kuncel, Sarah A. Hezlett, “Fact and Fiction in Cognitive Ability Testing for Admissions and Hiring Decisions,” *Current Directions in Psychological Science* 19, no. 6 (2010) : 339–345。

[人们对智力的本质也一直存在误解](#)：Kuncel, Hezlett, “Fact and Fiction”。

[就像一篇评论文章所指出的](#)：Frank L. Schmidt, John Hunter, “General Mental Ability in the World of Work: Occupational Attainment and Job Performance,” *Journal of Personality and Social Psychology* 86, no. 1 (2004) : 162。

[责任心和毅力](#)：Angela L. Duckworth, David Weir, Eli Tsukayama, David Kwok, “Who Does Well in Life? Conscientious Adults Excel in Both Objective and Subjective Success,” *Frontiers in Psychology* 3 (September 2012)。关于毅力，请参阅：Angela L. Duckworth, Christopher Peterson, Michael D. Matthews, Dennis Kelly, “Grit: Perseverance and Passion for Long-Term Goals,” *Journal of Personality and Social Psychology* 92, no. 6 (2007) : 1087–1101。

[流体智力](#)：Richard E. Nisbett et al., “Intelligence: New Findings and Theoretical Developments,” *American Psychologist* 67, no. 2 (2012) : 130–159, 229。

[GMA的预测效力比心理学研究中的大部分测量方法都好](#)：Schmidt and Hunter, “Occupational Attainment”, 162。

[标准化测验分数与工作绩效之间的相关系数达到0.5（PC=67%）：](#)

Kuncel, Hezlett, “Fact and Fiction”。

[按照社会科学的标准，相关系数达到0.5代表非常强的预测力：](#)这些相关性来自元分析（meta-analyses），修正了在一般和受限情况下观测到的测量误差的相关性。对于这些修正是否夸大了GMA的预测价值，研究者之间存在一些争论。然而，由于这些方法论上的争论也适用于其他预测因素，因此专家们普遍同意GMA（连同工作样本测试，见第24章）是衡量工作成功与否的最佳预测指标。参见 Kuncel, Hezlett, “Fact and Fiction”。

[在律师、化学家或工程师这些职业中，几乎没有GMA低于平均水平的人：](#) Schmidt, Hunter, “Occupational Attainment,” 162。

[即使认知能力的测试成绩处于前1%的群体，他们能获得的突出成就也与GMA高度相关：](#) David Lubinski, “Exceptional Cognitive Ability: The Phenotype,” *Behavior Genetics* 39, no. 4（2009）：350–358。

[2013年的一项研究重点调查了《财富》500强企业的CEO：](#) Jonathan Wai, “Investigating America’s Elite: Cognitive Ability, Education, and Sex Differences,” *Intelligence* 41, no. 4（2013）：203–211。

[研究人员建议使用的其他测量问题包括：](#) Keela S. Thomson, Daniel M. Oppenheimer, “Investigating an Alternate Form of the Cognitive Reflection Test,” *Judgment and Decision Making* 11, no. 1（2016）：99–113。

[低CRT得分与现实生活中的一些判断和信念有关：](#) Gordon Pennycook et al., “Everyday Consequences of Analytic Thinking,” *Current Directions in Psychological Science* 24, no. 6（2015）：425–432。

[CRT得分还可以预测人们是否会因为明显不准确的“假信息”而上当：](#)

Gordon Pennycook, David G. Rand, “Lazy, Not Biased: Susceptibility to

Partisan Fake News Is Better Explained by Lack of Reasoning than by Motivated Reasoning,” *Cognition* 188 (June 2018) : 39–50。

[该测试的得分甚至与人们使用智能手机的程度有关](#)：Nathaniel Barr et al., “The Brain in Your Pocket: Evidence That Smartphones Are Used to Supplant Thinking,” *Computers in Human Behavior* 48 (2015) : 473–480。

[人们是否会习惯性地运用反射性或冲动性思维过程](#)：Niraj Patel, S. Glenn Baker, Laura D. Scherer, “Evaluating the Cognitive Reflection Test as a Measure of Intuition/Reflection, Numeracy, and Insight Problem Solving, and the Implications for Understanding Real-World Judgments and Beliefs,” *Journal of Experimental Psychology: General* 148, no. 12 (2019) : 2129–2153。

[认知需求量表](#)：John T. Cacioppo, Richard E. Petty, “The Need for Cognition,” *Journal of Personality and Social Psychology* 42, no. 1 (1982) : 116–131。

[认知需求高的人不太容易出现已知的认知偏差](#)：Stephen M. Smith, Irwin P. Levin, “Need for Cognition and Choice Framing Effects,” *Journal of Behavioral Decision Making* 9, no. 4 (1996) : 283–290。

[那些在认知需求量表上得分低的人，更偏爱剧透](#)：Judith E. Rosenbaum, Benjamin K. Johnson, “Who’s Afraid of Spoilers? Need for Cognition, Need for Affect, and Narrative Selection and Enjoyment,” *Psychology of Popular Media Culture* 5, no. 3 (2016) : 273–289。

[成人决策能力量表](#)：Wandi Bruine De Bruin et al., “Individual Differences in Adult Decision-Making Competence,” *Journal of Personality and Social Psychology* 92, no. 5 (2007) : 938–956。

[哈尔彭批判性思维测试](#)：Heather A. Butler, “Halpern Critical Thinking

Assessment Predicts Real-World Outcomes of Critical Thinking,” *Applied Cognitive Psychology* 26, no. 5 (2012) : 721–729。

[可以作为人的预测能力指标的认知风格](#)：Uriel Haran, Ilana Ritov, Barbara Mellers, “The Role of Actively Open-Minded Thinking in Information Acquisition, Accuracy, and Calibration,” *Judgment and Decision Making* 8, no. 3 (2013) : 188–201。

[积极开放性思维](#)：Haran, Ritov, Mellers, “Role of Actively Open-Minded Thinking”。

[开放性思维是一种可习得的技能](#)：J. Baron, “Why Teach Thinking? An Essay,” *Applied Psychology: An International Review* 42 (1993) : 191–214; J. Baron, *The Teaching of Thinking: Thinking and Deciding*, 2nd ed. (New York: Cambridge University Press, 1994) , 127–148。

第19章 消除偏差与决策卫生

[他们的核心发现](#)：Jack B. Soll et al., “A User’s Guide to Debiasing,” in *The Wiley Blackwell Handbook of Judgment and Decision Making*, ed. Gideon Keren, George Wu, vol. 2 (New York: John Wiley & Sons, 2015) , 684。

[《绿皮书》](#)：HM Treasury, *The Green Book: Central Government Guidance on Appraisal and Evaluation* (London: UK Crown, 2018) 。

[助推](#)：Richard H. Thaler, Cass R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (New Haven, CT: Yale University Press, 2008) 。

[助力](#)：Ralph Hertwig, Till Grüne-Yanoff, “Nudging and Boosting: Steering or Empowering Good Decisions,” *Perspectives on Psychological*

Science 12, no. 6 (2017)。

教育人们克服偏差是一项崇高的事业，而且很有用：Geoffrey T. Fong et al., “The Effects of Statistical Training on Thinking About Everyday Problems”, *Cognitive Psychology* 18, no. 3 (1986) : 253–292。

当被问及常识性问题时，他们可能和其他人一样过分自信：Willem A. Wagenaar and Gideon B. Keren, “Does the Expert Know? The Reliability of Predictions and Confidence Ratings of Experts,” *Intelligent Decision Support in Process Environments* (1986) : 87–103。

这些游戏都使得参与者被问及类似问题时的犯错次数降低：Carey K. Morewedge et al., “Debiasing Decisions: Improved Decision Making with a Single Training Intervention,” *Policy Insights from the Behavioral and Brain Sciences* 2, no. 1 (2015) : 129–140。

应用学到的知识来解决商业问题：Anne-Laure Sellier et al., “Debiasing Training Transfers to Improve Decision Making in the Field,” *Psychological Science* 30, no. 9 (2019) : 1371–1379。

偏差盲点：Emily Pronin et al., “The Bias Blind Spot: Perceptions of Bias in Self Versus Others,” *Personality and Social Psychology Bulletin* 28, no. 3 (2002) : 369–381。

可能影响提案产生过程的偏差：Daniel Kahneman, Dan Lovallo, Olivier Sibony, “Before You Make That Big Decision...,” *Harvard Business Review* 89, no. 6 (June 2011) : 50–60。

该清单对于提升高风险环境中的决策有悠久的历史：Atul Gawande, *Checklist Manifesto: How to Get Things Right* (New York: Metropolitan Books, 2010)。

一份简单的检查清单：Office of Information and Regulatory Affairs,

“Agency Checklist: Regulatory Impact Analysis,” no date。

我们在附录2中展示了一个偏差检查清单：该清单部分改编自丹尼尔·卡尼曼等人，“Before You Make That Big Decision”, *Harvard Business Review*。

便于应用：Gawande, *Checklist Manifesto*。

第20章 司法科学，信息排序是最大的噪声

错误是人为所致：R. Stacey, “A Report on the Erroneous Fingerprint Individualisation in the Madrid Train Bombing Case,” *Journal of Forensic Identification* 54 (2004) : 707–718。

当时的FBI网站就坚称：Michael Specter, “Do Fingerprints Lie?,” *The New Yorker*, May 27, 2002。

正如德鲁尔所说：I. E. Dror, R. Rosenthal, “Meta-analytically Quantifying the Reliability and Biasability of Forensic Experts,” *Journal of Forensic Science* 53 (2008) : 900–903。

在第一项研究中：I. E. Dror, D. Charlton, A. E. Péron, “Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications,” *Forensic Science International* 156 (2006) : 74–78。

在第二项研究中：I. E. Dror, D. Charlton, “Why Experts Make Errors,” *Journal of Forensic Identification* 56 (2006) : 600–616。

指纹鉴定专家往往是根据背景环境做出决策的：I. E. Dror, S. A. Cole, “The Vision in ‘Blind’ Justice: Expert Perception, Judgment, and Visual Cognition in Forensic Pattern Recognition,” *Psychonomic Bulletin and Review* 17 (2010) : 161–167, 165. See also I. E. Dror, “A Hierarchy of Expert Performance (HEP) ,” *Journal of Applied Research in Memory and Cognition* (2016) : 1–6。

在另一项独立研究中：I. E. Dror et al., “Cognitive Issues in Fingerprint Analysis: Inter-and Intra-Expert Consistency and the Effect of a ‘Target’ Comparison,” *Forensic Science International* 208 (2011) : 10–17。

随后的一项独立研究：B. T. Ulery, R. A. Hicklin, M. A. Roberts, J. A. Buscaglia, “Changes in Latent Fingerprint Examiners’ Markup Between Analysis and Comparison,” *Forensic Science International* 247 (2015) : 54–61。

即使是被普遍视为司法科学新黄金标准的DNA分析，也容易受到证实性偏差的影响：I. E. Dror, G. Hampikian, “Subjectivity and Bias in Forensic DNA Mixture Interpretation,” *Science and Justice* 51 (2011) : 204–208。

鉴定人员经常会在随证据一起提交给他们的传送信函中收到此类信息：

M. J. Saks, D. M. Risinger, R. Rosenthal, W. C. Thompson, “Context Effects in Forensic Science: A Review and Application of the Science of Science to Crime Laboratory Practice in the United States,” *Science Justice Journal of Forensic Science Society* 43 (2003) : 77–90。

执行核实工作的鉴定人员知道最初的结论：President’s Council of Advisors on Science and Technology (PCAST), *Report to the President: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* (Washington, DC: Executive Office of the President, PCAST, 2016) 。

针对这一错误展开的调查：Stacey, “Erroneous Fingerprint”。

即使是一位备受尊敬的独立专家：Dror, Cole, “Vision in ‘Blind’ Justice”。

一系列偏差：I. E. Dror, “Biases in Forensic Experts,” *Science*

360 (2018) : 243。

[他们有时也会对自己先前见过的一组指纹改变看法](#)：Dror, Charlton, “Why Experts Make Errors”。

[2012年，FBI委托进行的一项研究](#)：B. T. Ulery, R. A. Hicklin, J. A. Buscaglia, M. A. Roberts, “Repeatability and Reproducibility of Decisions by Latent Fingerprint Examiners,” *PLoS One* 7 (2012) 。

[无辜者计划](#)：Innocence Project, “Overturning Wrongful Convictions Involving Misapplied Forensics,” *Misapplication of Forensic Science* (2018) : 1–7。还可参见S. M. Kassin, I. E. Dror, J. Kukucka, L. Butt, “The Forensic Confirmation Bias: Problems, Perspectives, and Proposed Solutions,” *Journal of Applied Research in Memory and Cognition* 2 (2013) : 42–52。

[对刑事法庭中的司法鉴定进行了全面回顾](#)：PCAST, *Report to the President*。

[大规模指纹识别准确性研究](#)：B. T. Ulery, R. A. Hicklin, J. Buscaglia, M. A. Roberts, “Accuracy and Reliability of Forensic Latent Fingerprint Decisions,” *Proceedings of the National Academy of Sciences* 108 (2011) : 7733–7738。

[这一比例要比普通公众乃至大部分陪审员认为的高很多](#)：(PCAST), *Report to the President*, p. 95。

[在佛罗里达州进行的一项后续研究](#)：Igor Pacheco, Brian Cerchiai, Stephanie Stoiloff, “Miami-Dade Research Study for the Reliability of the ACE-V Process: Accuracy & Precision in Latent Fingerprint Examinations,” final report, Miami-Dade Police Department Forensic Services Bureau, 2014。

[在大多数案件中，“排除”与“无法确认”对案件本身产生的影响是一样](#)

[的](#)：B. T. Ulery, R.A. Hicklin, M. A. Roberts, J. A. Buscaglia, “Factors Associated with Latent Fingerprint Exclusion Determinations,” *Forensic Science International* 275 (2017) : 65–75。

[鉴定人员做出的假阳性判断（错误识别）也要少得多](#)：R. N. Haber, I. Haber, “Experimental Results of Fingerprint Comparison Validity and Reliability: A Review and Critical Analysis,” *Science & Justice* 54 (2014) : 375–389。

[他们容易受到偏差的影响](#)：Dror, “Hierarchy of Expert Performance,” 3。

[他应该去迪士尼工作](#)：M. Leadbetter, letter to the editor, *Fingerprint World* 33 (2007) : 231。

[而不会真正改变他们的判断](#)：L. Butt, “The Forensic Confirmation Bias: Problems, Perspectives and Proposed Solutions—Commentary by a Forensic Examiner,” *Journal of Applied Research in Memory and Cognition* 2 (2013) : 59–60。

[就连FBI在梅菲尔德案的内部调查中都强调](#)：Stacey, “Erroneous Fingerprint,” 713。

[一项对21个国家400名鉴定专家展开的调查](#)：J. Kukucka, S. M. Kassin, P. A. Zapf, I. E. Dror, “Cognitive Bias and Blindness: A Global Survey of Forensic Science Examiners,” *Journal of Applied Research in Memory and Cognition* 6 (2017) 。

[线性序列揭露](#)：I. E. Dror et al., letter to the editor: “Context Management Toolbox: A Linear Sequential Unmasking (LSU) Approach for Minimizing Cognitive Bias in Forensic Decision Making,” *Journal of Forensic Science* 60 (2015) : 1111–1112。

第21章 甄选与汇总，超级预测的两大策略

官方机构在对预算进行预测时，会表现出不切实际的乐观：Jeffrey A. Frankel, “Over-optimism in Forecasts by Official Budget Agencies and Its Implications,” working paper 17239, National Bureau of Economic Research, December 2011。

预测者往往过于自信：H. R. Arkes, “Overconfidence in Judgmental Forecasting,” in *Principles of Forecasting: A Handbook for Researchers and Practitioners*, ed. Jon Scott Armstrong, vol. 30, International Series in Operations Research & Management Science (Boston: Springer, 2001)。

一项正在进行的季度调查：Itzhak Ben-David, John Graham, Campbell Harvey, “Managerial Miscalibration,” *The Quarterly Journal of Economics* 128, no. 4 (November 2013) : 1547–1584。

不可靠性也是判断预测的误差来源之一：T. R. Stewart, “Improving Reliability of Judgmental Forecasts,” in *Principles of Forecasting: A Handbook for Researchers and Practitioners*, ed. Jon Scott Armstrong, vol. 30, International Series in Operations Research & Management Science (Boston: Springer, 2001) (以下简称*Principles of Forecasting*) , 82。

让法学教授预测最高法院的裁决：Theodore W. Ruger, Pauline T. Kim, Andrew D. Martin, Kevin M. Quinn, “The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decision-Making,” *Columbia Law Review* 104 (2004) : 1150–1209。

空气污染管理制度：Cass Sunstein, “Maximin,” *Yale Journal of Regulation* (草稿; May 3, 2020)。

许多存在噪声的关于预测的例子：Armstrong, *Principles of Forecasting*。

对多次预测取平均值会大大提高预测的准确性：Jon Scott Armstrong, “Combining Forecasts,” in *Principles of Forecasting*, 417–439。

一组预测者的未加权平均值优于大多数个体及至所有个体的预测：T. R. Stewart, “Improving Reliability of Judgmental Forecasts,” in *Principles of Forecasting*, 95。

综合预测平均减少了12.5%的误差：Armstrong, “Combining Forecasts”。

根据近期判断的准确性来选择最好的判断者：Albert E. Mannes et al., “The Wisdom of Select Crowds,” *Journal of Personality and Social Psychology* 107, no.2 (2014) : 276–299。

预测市场的表现非常好：Justin Wolfers, Eric Zitzewitz, “Prediction Markets,” *Journal of Economic Perspectives* 18 (2004) : 107–126。

利用预测市场来汇总不同的观点：Cass R. Sunstein, Reid Hastie, *Wiser: Getting Beyond Group-think to Make Groups Smarter* (Boston: Harvard Business Review Press, 2014) 。

德尔菲法：Gene Rowe, George Wright, “The Delphi Technique as a Forecasting Tool: Issues and Analysis,” *International Journal of Forecasting* 15 (1999) : 353–375。另参见Dan Bang, Chris D. Frith, “Making Better Decisions in Groups,” *Royal Society Open Science* 4, no. 8 (2017) 。

实施起来却有一定的挑战性：R. Hastie, “Review Essay: Experimental Evidence on Group Accuracy,” in B. Grofman & G. Guillermo, eds., *Information Pooling and Group Decision Making* (Greenwich, CT: JAI Press, 1986) , 129–157。

迷你德尔菲法：Andrew H. Van De Ven, André L. Delbecq, “The Effectiveness of Nominal, Delphi, and Interacting Group Decision Making Processes,” *Academy of Management Journal* 17, no.

4（2017）。

好于能够阅读情报和其他秘密数据的情报界分析师的平均水平：

Superforecasting, 95。

超级预测者： *Superforecasting*, 231。

尝试，失败，分析： *Superforecasting*, 273。

一种复杂的统计技术： Ville A. Satopää, Marat Salikhov, Philip E. Tetlock, Barb Mellers, “Bias, Information, Noise: The BIN Model of Forecasting,” February 19, 2020, 23。

干预措施提高准确性： Satopää et al., “Bias,Information,Noise,” 23.

与培训方式不同的是，通过团队合作.....预测者可以利用这些信息：

Satopää et al., 22。

“超级预测者”的成功主要归功于他们在控制测量误差方面出色的能力：

Satopää et al., 24。

通过汇总既独立又互补的判断，我们可以获得准确度上的进一步提高：

Clinton P. Davis-Stober, David V. Budescu, Stephen B. Broomell, Jason Dana.“The composition of optimally wise crowds.” *Decision Analysis* 12, no. 3（2015）：130–143。

第22章 医疗决策，用科学的诊断指南减少噪声

量化肌腱退化的程度时，医生的诊断产生的噪声就很小： Laura Horton

et al., “Development and Assessment of Inter- and Intra-Rater Reliability of a Novel Ultrasound Tool for Scoring Tendon and Sheath Disease: A Pilot Study,” *Ultrasound* 24, no. 3（2016）：134。

当病理学家评估乳腺病灶的穿刺活检结果时： Laura C. Collins et al.,

“Diagnostic Agreement in the Evaluation of Image-guided Breast Core

Needle Biopsies,” *American Journal of Surgical Pathology* 28 (2004) : 126。

[即便有快速抗原检测结果](#)：Julie L. Fierro et al., “Variability in the Diagnosis, Treatment of Group A Streptococcal Pharyngitis by Primary Care Pediatricians,” *Infection Control and Hospital Epidemiology* 35, no. S3 (2014) : S79。

[就会被诊断为患有糖尿病](#)：Diabetes Tests, Centers for Disease Control and Prevention。

[在有些医院里，第二诊疗意见是必须给出的](#)：Joseph D. Kronz et al., “Mandatory Second Opinion Surgical Pathology at a Large Referral Hospital,” *Cancer* 86 (1999) : 2426。

[达特茅斯·阿特拉斯项目](#)：大部分相关材料可以在网上找到。达特茅斯医学院有像一本书那样长的大纲，*The Quality of Medical Care in the United States: A Report on the Medicare Program; the Dartmouth Atlas of Health Care 1999* (American Hospital Publishers, 1999) 。

[许多国家也存在医疗资源分配不均的情况](#)：OECD, *Geographic Variations in Health Care: What Do We Know and What Can Be Done to Improve Health System Performance?* (Paris: OECD Publishing, 2014) , 137–169; Michael P. Hurley et al., “Geographic Variation in Surgical Outcomes and Cost Between the United States and Japan,” *American Journal of Managed Care* 22 (2016) : 600; John Appleby, Veena Raleigh, Francesca Frosini, Gwyn Bevan, Haiyan Gao, Tom Lyscom, *Variations in Health Care: The Good, the Bad and the Inexplicable* (London: The King’s Fund, 2011) 。

[一项针对放射科医生做出肺炎诊断的研究发现](#)：David C. Chan Jr. et al., “Selection with Variation in Diagnostic Skill: Evidence from

Radiologists,” National Bureau of Economic Research, NBER Working Paper No. 26467, November 2019。

[在医疗领域也是如此](#)：P. J. Robinson, “Radiology’s Achilles’ Heel: Error and Variation in the Interpretation of the Röntgen Image,” *British Journal of Radiology* 70 (1997) : 1085。与之相关的一个研究是 Yusuke Tsugawa et al., “Physician Age and Outcomes in Elderly Patients in Hospital in the US: Observational Study,” *BMJ* 357 (2017)。研究结果发现，医生们越不训练，表现就越糟糕，因此，对于医生来说，在积累实践经验和熟悉最新的证据和指南之间存在一种权衡。这项研究发现，结束住院医师训练后头几年的医生表现最好，因为他们还记得这些证据。

[如放射学和病理学](#)：Robinson, “Radiology’s Achilles’ Heel”。

[kappa统计量](#)：与相关系数类似，kappa值可以是负的，不过在实践中很少出现。以下是不同kappa统计量的一个特征：轻度吻合 ($\kappa=0\sim0.2$)，一般吻合 ($\kappa=0.21\sim0.4$)，中等吻合 ($\kappa=0.41\sim0.6$)，高度吻合 ($\kappa=0.61\sim0.8$)，近乎完全吻合 ($\kappa>0.8$)。参见Ron Wald, Chaim M. Bell, Rosane Nisenbaum, Samuel Perrone, Orfeas Liangos, Andreas Laupacis, Bertrand L. Jaber, “Interobserver Reliability of Urine Sediment Interpretation,” *Clinical Journal of the American Society of Nephrology* 4, no. 3 (March 2009) : 567–571。

[药物之间的相互作用](#)：Howard R. Strasberg et al., “Inter-Rater Agreement Among Physicians on the Clinical Significance of Drug-Drug Interactions,” *AMIA Annual Symposium Proceedings* (2013) : 1325。

[肾病专家们在基于肾病患者的标准化检测结果做出诊断时](#)：Wald et al., “Interobserver Reliability of Urine Sediment Interpretation”。

[乳腺病变是否为癌变](#)：Juan P. Palazzo et al., “Hyperplastic Ductal and

Lobular Lesions and Carcinomas in Situ of the Breast: Reproducibility of Current Diagnostic Criteria Among Community-and Academic-Based Pathologists,” *Breast Journal* 4（2003）：230。

在诊断乳腺增生病变时：Rohit K. Jain et al., “Atypical Ductal Hyperplasia: Interobserver and Intraobserver Variability,” *Modern Pathology* 24（2011）：917。

判断椎管的狭窄程度：Alex C. Speciale et al., “Observer Variability in Assessing Lumbar Spinal Stenosis Severity on Magnetic Resonance Imaging and Its Relation to Cross-Sectional Spinal Canal Area,” *Spine* 27（2002）：1082。

在美国，心脏病是男性和女性的主要致死原因：Centers for Disease Control and Prevention, “Heart Disease Facts,” accessed June 16, 2020。

这种情况发生的可能性为31%：Timothy A. DeRouen et al., “Variability in the Analysis of Coronary Arteriograms,” *Circulation* 55（1977）：324。

这些医生在判断子宫内膜异位病灶的数量和位置时，产生了很大分歧：Olaf Buchweltz et al., “Interobserver Variability in the Diagnosis of Minimal and Mild Endometriosis,” *European Journal of Obstetrics & Gynecology and Reproductive Biology* 122（2005）：213。

肺结核诊断中依然存在显著的变异性：Jean-Pierre Zellweger et al., “Intra-observer and Overall Agreement in the Radiological Assessment of Tuberculosis,” *International Journal of Tuberculosis & Lung Disease* 10（2006）：1123。关于“一般”诊断一致性的内容，参见Yanina Balabanova et al., “Variability in Interpretation of Chest Radiographs Among Russian Clinicians and Implications for Screening

Programmes:Observational Study,” *BMJ* 331（2005）：379。

不同国家的放射科医生在肺结核的诊断上也存在差异：Shinsaku

Sakurada et al., “Inter-Rater Agreement in the Assessment of Abnormal Chest X-Ray Findings for Tuberculosis Between Two Asian Countries,” *BMC Infectious Diseases* 12, article 31（2012）。

研究人员要求8位病理学家对每个病例进行诊断：Evan R. Farmer et al.,

“Discordance in the Histopathologic Diagnosis of Melanoma and Melanocytic Nevi Between Expert Pathologists,” *Human Pathology* 27（1996）：528。

还有一项研究发现：Alfred W. Kopf, M. Mintzis, and R. S.

Bart, “Diagnostic Accuracy in Malignant Melanoma,” *Archives of Dermatology* 111（1975）：1291。

这项研究的作者总结道：Maria Miller, A. Bernard Ackerman, “How

Accurate Are Dermatologists in the Diagnosis of Melanoma? Degree of Accuracy and Implications,” *Archives of Dermatology* 128（1992）：559。

假阳性率也为1%~64%：Craig A. Beam et al., “Variability in the

Interpretation of Screening Mammograms by US Radiologists,” *Archives of Internal Medicine* 156（1996）：209。

有时候，放射科医生两次评估同一张影像片子时会给出不同的意见：P.

J. Robinson et al., “Variation Between Experienced Observers in the Interpretation of Accident and Emergency Radiographs,” *British Journal of Radiology* 72（1999）：323。

血管造影显示的血管阻塞程度：Katherine M. Detre et al., “Observer

Agreement in Evaluating Coronary Angiograms,” *Circulation* 52（1975）：979。

在那些标准模糊和判断情境复杂的领域中：Horton et al., “Inter- and Intra-Rater Reliability”;and Megan Banky et al., “Inter- and Intra-Rater Variability of Testing Velocity When Assessing Lower Limb Spasticity,” *Journal of Rehabilitation Medicine* 51 (2019) 。

另一项不涉及诊断的研究：Esther Y. Hsiang et al., “Association of Primary Care Clinic Appointment Time with Clinician Ordering and Patient Completion of Breast and Colorectal Cancer Screening,” *JAMA Network Open* 51 (2019) 。

还有一个例子也能说明临床医生会受到疲劳的影响：Hengchen Dai et al., “The Impact of Time at Work and Time Off from Work on Rule Compliance: The Case of Hand Hygiene in Health Care,”*Journal of Applied Psychology* 100 (2015) : 846。

对医学领域意义重大：Ali S. Raja, “The HEART Score Has Substantial Interrater Reliability,”*NEJM J Watch*, December 5, 2018 [reviewing Colin A. Gershon et al., “Inter-rater Reliability of the HEART Score,” *Academic Emergency Medicine* 26 (2019) : 552]。

我们在前面提到，培训可以提高医生的技能：Jean-Pierre Zellweger et al., “Intra-observer and Overall Agreement in the Radiological Assessment of Tuberculosis,” *International Journal of Tuberculosis & Lung Disease* 10 (2006) : 1123; Ibrahim Abubakar et al., “Diagnostic Accuracy of Digital Chest Radiography for Pulmonary Tuberculosis in a UK Urban Population,” *European Respiratory Journal* 35 (2010) : 689。

汇总多个专家的判断也能减少噪声：Michael L. Barnett et al., “Comparative Accuracy of Diagnosis by Collective Intelligence of Multiple Physicians vs Individual Physicians,” *JAMA Network Open* 2 (2019) : e19009; Kimberly H. Allison et al., “Understanding

Diagnostic Variability in Breast Pathology: Lessons Learned from an Expert Consensus Review Panel,” *Histopathology* 65 (2014) : 240。

[最好的算法](#)：Babak Ehteshami Bejnordi et al., “Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women with Breast Cancer,” *JAMA* 318 (2017) : 2199。

[深度学习算法](#)：Varun Gulshan et al., “Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs,” *JAMA* 316 (2016) : 2402。

[几乎和放射科医生一样出色](#)：Mary Beth Massat, “A Promising Future for AI in Breast Cancer Screening”, *Applied Radiology* 47 (2018) : 22; Alejandro Rodriguez-Ruiz et al., “Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison with 101 Radiologists,” *Journal of the National Cancer Institute* 111 (2019) : 916。

[表22-1, 阿普加评分指南](#)：Medline Plus (last accessed February 4, 2020) 。

[阿普加评分产生的噪声很小](#)：L. R. Foster et al., “The Interrater Reliability of Apgar Scores at 1 and 5 Minutes,” *Journal of Investigative Medicine* 54, no. 1 (2006) : 293。

[使用该量表进行评估和评分相对直接](#)：Warren J. McIsaac et al., “Empirical Validation of Guidelines for the Management of Pharyngitis in Children and Adults,” *JAMA* 291 (2004) : 1587。

[一项研究发现, BI-RADS提升了乳房X线片的评估者之间的一致性](#)：

Emilie A. Ooms et al., “Mammography: Interobserver Variability in Breast Density Assessment,” *Breast* 16 (2007) : 568。

[在病理学领域](#)：Frances P. O’Malley et al., “Interobserver Reproducibility

in the Diagnosis of Flat Epithelial Atypia of the Breast,” *Modern Pathology* 19 (2006) : 172。

至少从20世纪40年代起，减少噪声就成为精神病学界的头等大事：

Ahmed Aboraya et al., “The Reliability of Psychiatric Diagnosis Revisited,” *Psychiatry (Edgmont)* 3 (2006) : 41。有关概述，详见 N. Kreitman, “The Reliability of Psychiatric Diagnosis,” *Journal of Mental Science* 107 (1961) : 876–886。

1964年，一项针对91名患者和10名有经验的精神科医生的研究：

Aboraya et al., “Reliability of Psychiatric Diagnosis Revisited,” 43。

为了初步揭示其中的原因： C. H. Ward et al., “The Psychiatric Nomenclature: Reasons for Diagnostic Disagreement,” *Archives of General Psychiatry* 7 (1962) : 198。

具有生物学背景的另外一位临床医生： Aboraya et al., “Reliability of Psychiatric Diagnosis Revisited”。

DSM III促使大量的研究关注诊断是否存在噪声： Samuel M. Lieblich, David J. Castle, Christos Pantelis, Malcolm Hopwood, Allan Hunter Young, and Ian P. Everall, “High Heterogeneity and Low Reliability in the Diagnosis of Major Depression Will Impair the Development of New Drugs,” *British Journal of Psychiatry Open* 1 (2015) : e5–e7。

但这本手册远没有达到完美： Lieblich et al., “High Heterogeneity”。

即使在2000年对第4版——DSM IV进行了重大修订之后： Elie Cheniaux et al., “The Diagnoses of Schizophrenia, Schizoaffective Disorder, Bipolar Disorder and Unipolar Depression: Interrater Reliability and Congruence Between DSM IV and ICD 10,” *Psychopathology* 42 (2009) : 296–298,特别是293页； and Michael Chmielewski et al., “Method Matters: Understanding Diagnostic Reliability in DSM IV and

DSM V,” *Journal of Abnormal Psychology* 124 (2015) : 764, 768–769。

[提高精神疾病诊断的可靠性](#)：Aboraya et al., “Reliability of Psychiatric Diagnosis Revisited,” 47。

[存在着一个严重的风险](#)：Aboraya et al., 47。

[该手册的第5版](#)：Chmielewski et al., “Method Matters”。

[美国精神病学学会](#)：Helena Chmura Kraemer et al., “DSM–5: How Reliable Is Reliable Enough?,” *American Journal of Psychiatry* 169 (2012) : 13–15。

[精神科医生的诊断仍然表现出明显的噪声](#)：Lieblich et al., “High Heterogeneity”。

[精神科医生就患者是否患有重度抑郁症很难达成一致](#)：Lieblich et al., “High Heterogeneity,” e–5。

[DSM V 的现场试验发现](#)：Lieblich et al., e–5。

[另外一些现场试验表明](#)：Lieblich et al., e–6。

[使用指南之所以很难取得成功](#)：Aboraya et al., “Reliability of Psychiatric Diagnosis Revisited,” 47。

[明确诊断标准，舍弃模糊标准](#)：Aboraya et al.

[一位观察者曾说](#)：Aboraya et al.

[医学界需要更多的指南](#)：一些有价值的注意事项参见Christopher Worsham, Anupam B. Jena, “The Art of Evidence-Based Medicine,” *Harvard Business Review*, January 30, 2019。

第23章 绩效评估，用基于外部视角的共识框架做出量化判断

[正如有家报纸的标题所示](#)：Jena McGregor, “Study Finds That Basically Every Single Person Hates Performance Reviews,” *Washington Post*, January 27, 2014。

[以判断为基础的绩效评估无处不在](#)：许多组织正在经历的数字化转型，在这个过程中，以判断为基础的绩效评估或许能够帮助创造新的可能性。从理论上讲，公司现在可以收集大量的关于每位员工绩效的详细的、实时的信息，基于这些数据，完全依据算法对一些职位进行绩效评估将成为可能。然而，我们此处关注的职位是那些不能完全从对其进行绩效评估的过程中排除判断的类型。E. D. Pulakos, R. Mueller-Hanson, S. Arad, “The Evolution of Performance Management: Searching for Value,” *Annual Review of Organizational Psychology and Organizational Behavior* 6（2018）：249–271。

[其中大多数人都发现这些评估充满了噪声](#)：S. E. Scullen, M. K. Mount, M. Goff, “Understanding the Latent Structure of Job Performance Ratings,” *Journal of Applied Psychology* 85（2000）：956–970。

[其余70%~80%的差异是系统噪声](#)：有一小部分（在有些研究中占总方差的10%）是来自研究者所说的评估者视角或者层级效应（level effect），指在组织中等级的意义上，而不是我们在这里定义的水平噪声。评估者视角反映了：在对同一个人进行评价时，老板和同事有系统性的差异，同事和下属也是如此。从善意的角度来解读360度评级系统的结果，人们可能会认为这不是噪声。如果组织中不同层级的人在看待同一个人的绩效表现时，普遍会看到不同的方面，他们对此人的判断应该有系统性的差异，而且他们的评估也应该表明这一点。

[多项关于工作绩效评估变异性的研究](#)：Scullen, Mount, Goff, “Latent Structure”；C. Viswesvaran, D. S. Ones, F. L. Schmidt, “Comparative Analysis of the Reliability of Job Performance Ratings,” *Journal of*

Applied Psychology 81 (1996) : 557–574; G. J. Greguras, C. Robie, “A New Look at Within-Source Interrater Reliability of 360-Degree Feedback Ratings,” *Journal of Applied Psychology* 83 (1998) : 960–968; G. J. Greguras, C. Robie, D. J. Schleicher, M. A. Goff, “A Field Study of the Effects of Rating Purpose on the Quality of Multisource Ratings,” *Personnel Psychology* 56 (2003) : 1–21; C. Viswesvaran, F. L. Schmidt, D. S. Ones, “Is There a General Factor in Ratings of Job Performance? A Meta-Analytic Framework for Disentangling Substantive and Error Influences,” *Journal of Applied Psychology* 90 (2005) : 108–131; B. Hoffman, C. E. Lance, B. Bynum, W. A. Gentry, “Rater Source Effects Are Alive and Well After All,” *Personnel Psychology* 63 (2010) : 119–151。

员工真实的工作绩效和对工作绩效进行的评估之间的关系：K. R. Murphy, “Explaining the Weak Relationship Between Job Performance and Ratings of Job Performance,” *Industrial and Organizational Psychology* 1 (2008) : 148–160尤其是151页。

员工的真实绩效：在讨论噪声的来源时，我们忽略了评价某些员工或员工类别的系统性偏差所引起案例噪声（case noise）的可能性。我们能够找到的关于绩效评估变异性的研究，没有一项与外部评估的“真实”绩效进行过比较。

“策略性地”对员工进行评估：E. D. Pulakos, R. S. O’Leary, “Why Is Performance Management Broken?,” *Industrial and Organizational Psychology* 4 (2011) : 146–164; M. M. Harris, “Rater Motivation in the Performance Appraisal Context: A Theoretical Framework,” *Journal of Management* 20 (1994) : 737–756; K. R. Murphy, J. N. Cleveland, *Understanding Performance Appraisal: Social, Organizational, and Goal-Based Perspectives* (Thousand Oaks, CA: Sage, 1995) 。

[为了帮助一个一直在寻求晋升机会的人](#)：Greguras et al., “Field Study”。

[它可以对客观、可量化的绩效做出预测](#)：P. W. Atkins, R. E. Wood, “Self-Versus Others’ Ratings as Predictors of Assessment Center Ratings:Validation Evidence for 360-Degree Feedback Programs,” *Personnel Psychology* (2002) 。

[过度设计的调查问卷](#)：Atkins, Wood, “Self-Versus Others’ Ratings”。

[98%的管理者](#)：Olson, Davis, cited in Peter G. Dominick, “Forced Ranking:Pros, Cons and Practices,”in *Performance Management:Putting Research into Action*, ed. James W. Smither, Manuel London (San Francisco: Jossey-Bass, 2009) , 411–443。

[强制排名](#)：Dominick, “Forced Ranking”。

[这种关系也被证明适用于绩效评估](#)：Barry R. Nathan, Ralph A. Alexander, “A Comparison of Criteria for Test Validation: A Meta-Analytic Investigation,” *Personnel Psychology* 41, no. 3 (1988) : 517–535。

[图23-1](#)：Adapted from Richard D. Goffin, James M. Olson, “Is It All Relative? Comparative Judgments and the Possible Improvement of Self-Ratings and Ratings of Others,” *Perspectives on Psychological Science* 6, no. 1 (2011) : 48–60。

[德勤经过计算发现](#)：M. Buckingham, A. Goodall, “Reinventing Performance Management,” *Harvard Business Review*, April 1, 2015, 1–16, doi: ISSN: 0017-8012。

[一项研究发现](#)：Corporate Leadership Council, cited in S. Adler et al., “Getting Rid of Performance Ratings: Genius or Folly? A Debate,” *Industrial and Organizational Psychology* 9 (2016) : 219–252。

[正如一篇评论文章中总结的那样](#)：Pulakos, Mueller-Hanson, and Arad, “Evolution of Performance Management”, 250。

[绩效管理革命](#)：A. Tavis, P. Cappelli, “The Performance Management Revolution,” *Harvard Business Review*, October 2016, 1–17。

[有证据表明，行为锚定评估量表不足以消除噪声](#)：Frank J. Landy, James L. Farr. “Performance Rating,” *Psychological Bulletin* 87, no. 1 (1980) : 72–107。

[通过视频中的案例来练习进行绩效评估](#)：D. J. Woehr and A.I. Huffcutt, “Rater Training for Performance Appraisal:A Quantitative Review,” *Journal of Occupational and Organizational Psychology* 67 (1994) : 189–205; S. G. Roch, D. J. Woehr,V. Mishra, U. Kieszczynska, “Rater Training Revisited:An Updated Meta-Analytic Review of Frame-of-Reference Training,” *Journal of Occupational and Organizational Psychology* 85 (2012) : 370–395; M. H. Tsai, S. Wee, B. Koh, “Restructured Frame-of-Reference Training Improves Rating Accuracy,” *Journal of Organizational Behavior* (2019) : 1–18。

[图23-2](#)：图左改编自Richard Goffin, James M. Olson, “Is It All Relative? Comparative Judgments and the Possible Improvement of Self-Ratings and Ratings of Others,” *Perspectives on Psychological Science* 6, no. 1 (2011) : 48–60。

[大多数关于参照框架培训的研究](#)：Roch et al., “Rater Training Revisited”。

[超级人才](#)：Ernest O’Boyle, Herman Aguinis, “The Best and the Rest:Revisiting the Norm of Normality of Individual Performance,” *Personnel Psychology* 65, no. 1 (2012) : 79–119; 以及Herman Aguinis, Ernest O’Boyle, “Star Performers in Twenty-First Century

Organizations,” *Personnel Psychology* 67, no.2 (2014) : 313–350。

第24章 人员招聘，以结构化指标衡量人才

很少有人可以不经过面试就被录用：A. I. Huffcutt, S. S. Culbertson, “Interviews,” in S. Zedeck, ed., *APA Handbook of Industrial and Organizational Psychology* (Washington, DC: American Psychological Association, 2010) , 185–203。

在某种程度上依赖直觉性判断：N. R. Kuncel, D. M. Klieger, D. S. Ones, “In Hiring, Algorithms Beat Instinct,” *Harvard Business Review* 92, no. 5 (2014) : 32。

首要问题：R. E. Ployhart, N. Schmitt, N. T. Tippins, “Solving the Supreme Problem: 100 Years of Selection and Recruitment at the *Journal of Applied Psychology*,” *Journal of Applied Psychology* 102 (2017) : 291–304。

其他研究报告的相关系数为0.2~0.33：M. McDaniel, D. Whetzel, F. L. Schmidt, S. Maurer, “Meta Analysis of the Validity of Employment Interviews,” *Journal of Applied Psychology* 79 (1994) : 599–616; A. Huffcutt, W. Arthur, “Hunter and Hunter (1984) Revisited: Interview Validity for Entry-Level Jobs,” *Journal of Applied Psychology* 79 (1994) : 2; F. L. Schmidt, J. E. Hunter, “The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings,” *Psychology Bulletin* 124 (1998) : 262–274; F. L. Schmidt, R. D. Zimmerman, “A Counterintuitive Hypothesis About Employment Interview Validity and Some Supporting Evidence,” *Journal of Applied Psychology* 89 (2004) : 553–561。请注意，当考虑某些研究的子集时有效性更高，如果研究使用专门为此目的创建的绩效评估，而不是现有的管

理评估，结论更是如此。

客观无知：S. Highhouse, “Stubborn Reliance on Intuition and Subjectivity in Employee Selection,” *Industrial and Organizational Psychology* 1 (2008) : 333–342; D. A. Moore, “How to Improve the Accuracy and Reduce the Cost of Personnel Selection,” *California Management Review* 60 (2017) : 8–17。

有相似的文化背景或共同之处：L. A. Rivera, “Hiring as Cultural Matching: The Case of Elite Professional Service Firms,” *American Sociology Review* 77 (2012) : 999–1022。

两位面试官对同一位应聘者的评分的相关系数：Schmidt, Zimmerman, “Counterintuitive Hypothesis”; Timothy A. Judge, Chad A. Higgins, Daniel M. Cable, “The Employment Interview: A Review of Recent Research and Recommendations for Future Research,” *Human Resource Management Review* 10 (2000) : 383–406; 以及A. I. Huffcutt, S. S. Culbertson, W. S. Weyhrauch, “Employment Interview Reliability: New Meta-Analytic Estimates by Structure and Format,” *International Journal of Selection and Assessment* 21 (2013) : 264–276。

第一印象非常重要：M. R. Barrick et al., “Candidate Characteristics Driving Initial Impressions During Rapport Building: Implications for Employment Interview Validity”, *Journal of Occupational and Organizational Psychology* 85 (2012) : 330–352; M. R. Barrick, B. W. Swider, G. L. Stewart, “Initial Evaluations in the Interview: Relationships with Subsequent Interviewer Evaluations and Employment Offers,” *Journal of Applied Psychology* 95 (2010) : 1163。

握手的感受：G. L. Stewart, S. L. Dustin, M. R. Barrick, T. C. Darnold, “Exploring the Handshake in Employment Interviews,” *Journal of*

Applied Psychology 93, (2008) : 1139–1146。

积极的第一印象：T. W. Dougherty, D. B. Turban, J. C. Callender, “Confirming First Impressions in the Employment Interview: A Field Study of Interviewer Behavior,” *Journal of Applied Psychology* 79 (1994) : 659–665。

一个让人难以置信的实验：J. Dana, R. Dawes, N. Peterson, “Belief in the Unstructured Interview: The Persistence of an Illusion,” *Judgment and Decision Making* 8 (2013) : 512–520。

绝大多数人力资源专业人士都更赞成使用诊断性汇总：Nathan R. Kuncel et al., “Mechanical versus Clinical Data Combination in Selection and Admissions Decisions:A Meta-Analysis,” *Journal of Applied Psychology* 98, no. 6 (2013) : 1060–1072。

相关度为0：Laszlo Bock, interview with Adam Bryant, *The New York Times*, June 19, 2013。另见Laszlo Bock, *Work Rules!: Insights from Inside Google That Will Transform How You Live and Lead* (New York: Hachette, 2015) 。

一位知名的猎头：C. Fernández-Aráoz, “Hiring Without Firing,” *Harvard Business Review*, July 1, 1999。

结构化的行为面试：有关结构化内面试的指南，参见Michael A. Campion, David K. Palmer, James E. Campion, “Structuring Employment Interviews to Improve Reliability, Validity and Users’Reactions,” *Current Directions in Psychological Science* 7, no. 3 (1998) : 77–82。

究竟什么样的面试才算得上是结构化面试：J. Levashina, C. J. Hartwell, F. P. Morgeson, M. A. Campion, “The Structured Employment Interview: Narrative and Quantitative Review of the Research

Literature,” *Personnel Psychology* 67 (2014) : 241–293。

结构化面试比传统的非结构化面试更能预测应聘者未来的表现：

McDaniel et al., “Meta Analysis”; Huffcutt, Arthur, “Hunter and Hunter (1984) Revisited”; Schmidt, Hunter, “Validity and Utility”; Schmidt, Zimmerman, “Counterintuitive Hypothesis”。

工作样本测试：Schmidt, Hunter, “Validity and Utility”。

以色列军队：Kahneman, *Thinking, Fast and Slow*, 229。

实用性的建议和指导：Kuncel, Klieger, Ones, “Algorithms Beat Instinct”。另见Campion, Palmer, and Campion, “Structuring Employment Interviews”。

错觉的持续：Dana, Dawes, Peterson, “Belief in the Unstructured Interview”。

第25章 中介评估法，做出明智决策的核心方法

中介评估法：Daniel Kahneman, Dan Lovallo, Olivier Sibony, “A Structured Approach to Strategic Decisions: Reducing Errors in Judgment Requires a Disciplined Process,” *MIT Sloan Management Review* 60 (2019) : 67–73。

评估-讨论-评估法：Andrew H. Van De Ven and André Delbecq, “The Effectiveness of Nominal, Delphi, and Interacting Group Decision Making Processes,” *Academy of Management Journal* 17, no. 4 (1974) : 605–621。另见本书第21章。

第六部分 最佳的噪声水平

在他们看来，没有任何机械性的解决方案可以满足司法要求：Kate

Stith, José A. Cabranes, *Fear of Judging: Sentencing Guidelines in the Federal Courts* (Chicago: University of Chicago Press, 1998) , 177。

第26章 减少噪声的成本

[这样的改革可能适得其反](#)：Albert O. Hirschman, *The Rhetoric of Reaction: Perversity, Futility, Jeopardy* (Cambridge, MA: Belknap Press, 1991) 。

[瓦茨拉夫·哈韦尔](#)：Stith, Cabranes, *Fear of Judging*。

[三振出局](#)：见斯坦福大学法学院三振出局的基本要素。

[伍德森诉北卡罗来纳州案](#)：428 U. S. 280 (1976) 。

[依靠大数据和借助算法来做决策可能会产生偏见](#)：Cathy O’Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Crown, 2016) 。

[存在潜在偏差的数学模型正在重塑我们的生活](#)：Will Knight, “Biased Algorithms Are Everywhere, and No One Seems to Care,” *MIT Technology Review*, July 12, 2017。

[ProPublica](#)：Jeff Larson, Surya Mattu, Lauren Kirchner, Julia Angwin, “How We Analyzed the COMPAS Recidivism Algorithm,” ProPublica, May 23, 2016。这个例子中的偏见主张是有争议的，对偏见的不同定义可能会导致相反的结论。有关此案例的观点以及与算法偏差的定义和测量相关的讨论，请参阅下文的注释“究竟如何测试”。

[预测性警务](#)：Aaron Shapiro, “Reform Predictive Policing,” *Nature* 541, no. 7638 (2017) : 458–460。

[事实上，就这一点而言，算法可能更糟](#)：虽然这种担忧重新引起人们的注意是在基于人工智能模型的领域中，但它并不限于人工智能领

域。早在1972年，保罗·斯洛维奇就指出对直觉进行建模会保留、加强甚至可能会放大现有的认知偏见。Paul Slovic, “Psychological Study of Human Judgment: Implications for Investment Decision Making,” *Journal of Finance* 27 (1972) : 779。

[究竟如何测试](#)：关于在COMPAS惯犯预测算法存在争议的背景下对这场辩论的介绍，参见Larson et al., “COMPAS Recidivism Algorithm”; William Dieterich et al., “COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity,” Northpointe, Inc., July 8, 2016; Julia Dressel, Hany Farid, “The Accuracy, Fairness, and Limits of Predicting Recidivism,” *Science Advances* 4, no. 1 (2018) : 1–6; Sam Corbett-Davies et al., “A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased Against Blacks. It’s Actually Not That Clear,” *Washington Post*, October 17, 2016; Alexandra Chouldechova, “Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments,” *Big Data* 153 (2017) : 5; Jon Kleinberg, Sendhil Mullainathan, Manish Raghavan, “Inherent Trade-Offs in the Fair Determination of Risk Scores,” Leibniz International Proceedings in Informatics, January 2017。

第27章 尊严，人之为人的重要价值观

[他们可能知道，自己的做法充满噪声](#)：Tom R. Tyler, *Why People Obey the Law*, 2nd ed. (New Haven, CT: Yale University Press, 2020) 。

[美国宪法中有一项以令人费解著称的裁决，可以帮我们理解这一点](#)： *Cleveland Bd. of Educ. v. LaFleur*, 414 U. S. 632 (1974) 。

[当时有影响力的评论员为法院的裁决辩护](#)：Laurence H. Tribe, “Structural Due Process,” *Harvard Civil Rights-Civil Liberties Law Review* 10, no. 2 (spring 1975) : 269。

[回想一下，我们在前文中论述过的一些法官对量刑指南的强烈反对：](#)

Stith, Cabranes, *Fear of Judging*, 177。

[一系列颇具影响力的作品](#)：请参阅Philip K. Howard, *The Death of Common Sense: How Law Is Suffocating America* (New York:Random House, 1995)；以及Philip K. Howard, *Try Common Sense: Replacing the Failed Ideologies of Right and Left* (New York: W. W. Norton & Company, 2019)。

第28章 规则还是标准

[脸书在2020年社区标准中规定的这些内容](#)：Hate Speech, Facebook: Community Standards。

[《纽约客》杂志](#)：Andrew Marantz, “Why Facebook Can’t Fix Itself,” *The New Yorker*, October 12, 2020。

[官僚正义](#)：Jerry L. Mashaw, *Bureaucratic Justice*, New Haven, CT: Yale University Press (1983)。

[恰恰相反，法官应是根据每个人的性格和具体情况，或根据公正性和具体结果的适当性来进行评判的](#)：David M. Trubek, “Max Weber on Law and the Rise of Capitalism,” *Wisconsin Law Review* 720 (1972) : 733, n. 22 [引自Max Weber, *The Religion of China* (1951) , 149]。

未来,属于终身学习者

我这辈子遇到的聪明人(来自各行各业的聪明人)没有不每天阅读的——没有,一个都没有。巴菲特读书之多,我读书之多,可能会让你感到吃惊。孩子们都笑话我。他们觉得我是一本长了两条腿的书。

——查理·芒格

互联网改变了信息连接的方式;指数型技术在迅速颠覆着现有的商业世界;人工智能已经开始抢占人类的工作岗位……

未来,到底需要什么样的人才?

改变命运唯一的策略是你要变成终身学习者。未来世界将不再需要单一的技能型人才,而是需要具备完善的知识结构、极强逻辑思考力和高感知力的复合型人才。优秀的人往往通过阅读建立足够强大的抽象思维能力,获得异于众人的思考和整合能力。未来,将属于终身学习者!而阅读必定和终身学习形影不离。

很多人读书,追求的是干货,寻求的是立刻行之有效的解决方案。其实这是一种留在舒适区的阅读方法。在这个充满不确定性的年代,答案不会简单地出现在书里,因为生活根本就没有标准确切的答案,你也不能期望过去的经验能解决未来的问题。

而真正的阅读,应该在书中与智者同行思考,借他们的视角看到世界的多元性,提出比答案更重要的好问题,在不确定的时代中领先起跑。

湛庐阅读 App : 与最聪明的人共同进化

有人常常把成本支出的焦点放在书价上,把读完一本书当作阅读的终结。其实不然。

时间是读者付出的最大阅读成本

怎么读是读者面临的最大阅读障碍

“读书破万卷”不仅仅在“万”,更重要的是在“破”!

现在,我们构建了全新的“湛庐阅读”App。它将成为你“破万卷”的新居所。在这里:

- 不用考虑读什么,你可以便捷找到纸书、电子书、有声书和各种声音产品;
- 你可以学会怎么读,你将发现集泛读、通读、精读于一体的阅读解决方案;
- 你会与作者、译者、专家、推荐人和阅读教练相遇,他们是优质思想的发源地;
- 你会与优秀的读者和终身学习者伍,他们对阅读和学习有着持久的热情和源源不断的内驱力。

从单一到复合,从知道到精通,从理解到创造,湛庐希望建立一个“与最聪明的人共同进化”的社区,成为人类先进思想交汇的聚集地,与你共同迎接未来。

与此同时,我们希望能够重新定义你的学习场景,让你随时随地收获有内容、有价值的思想,通过阅读实现终身学习。这是我们的使命和价值。

CHEERS

本书阅读资料包

给你便捷、高效、全面的阅读体验

本书参考资料

湛庐独家策划

- ✓ 参考文献
为了环保、节约纸张，部分图书的参考文献以电子版方式提供
- ✓ 主题书单
编辑精心推荐的延伸阅读书单，助你开启主题式阅读
- ✓ 图片资料
提供部分图片的高清彩色原版大图，方便保存和分享

相关阅读服务

终身学习者必备

- ✓ 电子书
便捷、高效，方便检索，易于携带，随时更新
- ✓ 有声书
保护视力，随时随地，有温度、有情感地听本书
- ✓ 精读班
2~4周，最懂这本书的人带你读完、读懂、读透这本好书
- ✓ 课 程
课程权威专家给你开书单，带你快速浏览一个领域的知识概貌
- ✓ 讲 书
30分钟，大咖给你讲本书，让你挑书不费劲



湛庐编辑为你独家呈现
助你更好获得书里和书外的思想和智慧，请扫码查收！

(阅读资料包的内容因书而异，最终以湛庐阅读 App 页面为准)

-
- (1) 我们怀疑，如果判决日恰好是法官的生日，他们也可能会更宽容，但这一假设尚未得到验证。
- (2) 这个典故出自英国作家罗伯特·骚塞（Robert Soutney）的童话故事《三只小熊》（*The Story of Three Bears*）。迷了路的金发姑娘误入了三只小熊的家，并且喝了小碗里的粥，坐了最小的椅子，睡在了最小的床上，因为那些是最适合她的。因此，“金发姑娘价格”是指最合适的价格。——编者注
- (3) NBA职业运动员罚球命中率最高数据为略高于90%。在撰写本书时，命中率前三名分别为斯蒂芬·库里（Stephen Curry）、史蒂夫·纳什（Steve Nash）和马克·普莱斯（Mark Price）。最低的命中率约为50%，美国著名篮球运动员沙奎尔·奥尼尔（Shaquille O’Neal）的命中率只有约53%。
- (4) 技术术语为“重测信度”（test-retest reliability），简称“信度”。
- (5) 1磅约等于453.59克。——编者注
- (6) 哈佛大学终身教授、“麦克阿瑟天才奖”获得者穆来纳森和普林斯顿大学心理学教授埃尔德·沙菲尔（Eldar Shafir）推出了行为经济学的重磅著作《稀缺》，首度提出“带宽=认知能力+执行控制力”，并指出处于稀缺状态中的人的大脑会被稀缺心态俘获，认知能力与执行控制力会变得低下。同时，书中还提出了一些改善稀缺心态的方法。该书中文简体字版由湛庐引进、浙江人民出版社2018年出版。——编者注
- (7) 文氏图是用于表示不同集合之间大致关系的一种“草图”。——编者注
- (8) 一个网站，人们可以通过在上面提供一些短期服务，如填答问卷，来获取报酬。
- (9) 目前，GMA比IQ更常用。
- (10) 这类似于教师给学生60以上的分数可以避免出题补考的麻烦。——译者注
- (11) 下面这段文字选自莎士比亚的《威尼斯商人》，朱生豪译，译林出版社2018年出版。——编者注
- (12) 此处的缺少共性意指残留的误差不像偏差那样指向同一方向。——译者注

湛庐文化
Essentialism

与思想有关
a mind about business

财富汇

精要主义

〔英〕格雷戈·麦吉沃恩 / 著



浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

〔英〕格雷戈·麦吉沃恩 / 著
Greg McKeown
邵信芳 / 译



精要主义

如何应对拥挤不堪的工作与生活
The Disciplined Pursuit of Less



浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

如果不能自己安排
生活的优先次序，
就只能
任由别人替你安排

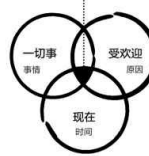
IF YOU DON'T
PRIORITIZE
YOUR LIFE,
SOMEONE ELSE
WILL.



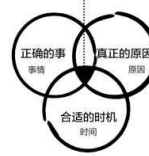
Explore

你付出的苦劳是个人努力峰值，
达成的切实结果才是
个人贡献峰值

个人努力峰值

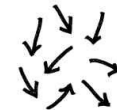


个人贡献峰值



Eliminate

如果答案不是
一个确定的 YES
就是一个肯定的 NO



很清楚目标



真的清楚目标

Execute

完成比完美更重要

一步一个台阶



版权信息

本书纸版由浙江人民出版社于 2016年4月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

书名：精要主义

著者：【英】格雷戈·麦吉沃恩

字数：210000

电子书定价：35.99美元

Greg McKeown. Essentialism: The Disciplined Pursuit of Less

Copyright 2014 by Greg McKeown

目 录

[版权信息](#)

[中文版序 对话麦吉沃恩，精要的力量](#)

[设计人生，以一种精要的方式](#)

[Part I 精要主义者的核心思维模式](#)

[导读 时间与精力只用于有意义的事](#)

[模式一选择，一种战无不胜的力量](#)

[模式二甄别，几乎一切都是噪声](#)

[模式三取舍，我要的究竟是什么](#)

[Part II 精要主义者养成的第一步：探索](#)

[导读 区分无意义的多数和有意义的少数](#)

[要素一抽离，为探索和思考留出空间](#)

[要素二审视，关注真正重要的事情](#)

[要素三游戏，拥抱内心孩童的智慧](#)

[要素四睡眠，保护你的身心资产](#)

[要素五精选，只接受前10% 的机会](#)

[Part III 精要主义者养成的第二步：排除](#)

[导读 摆脱无意义的多数](#)

[要素六澄清，做一个以一当千的决定](#)

[要素七勇气，优雅说“不”的力量](#)

[要素八放弃，止损就是最大的赢](#)

[要素九编辑，看不见的艺术](#)

[要素十边界，设定界线会带来自由](#)

[Part IV 精要主义者养成的第三步：执行](#)

[导读 让有意义的少数做起来毫不费力](#)

[要素十一缓冲，充足准备带来优势](#)

[要素十二扫除，锁定你的“最慢行进者”](#)

[要素十三进步，小胜的价值](#)

[要素十四心流，常规造就天才](#)

[要素十五专注，当下最重要的是什么](#)

[存在，设计你的精要人生](#)

[从精要主义者，到精要领导者](#)

[译者后记 不在琐碎中无足轻重](#)

对话麦吉沃恩，精要的力量

Q：《精要主义》即将在中国出版，能否与我们分享您此刻的感受？

A：中国对我来说很特别。每次来这里访问，我都会被中国人对自己前辈怀有的浓厚敬意深深打动。这种敬意很重要，它让我开始相信代际视角是辨认生命中最重要事务的最有效的方式。一个又一个瞬间倏忽而过，一封又一封邮件消耗着生命时光，人们很容易错误地认为刚收到的邮件非同小可，而事实上却不过是鸡毛蒜皮而已。一旦着眼于长远，胸怀一个十年目标，我们就能视野清明，眼光深邃。当我们在进行人生设计时，若能把祖父母和孙子女放在心里，就更能心明眼亮，洞悉事理。因而，中国人的代际文化遗产让精要主义具有了特别的适用性。

Q：何为精要主义？

A：我住在加利福尼亚的硅谷，我的公司THIS总部也在那里。也是在那里，我第一次发现了一个问题，这个问题在中国也同样存在：专注精要事务，可以通向成功；但成功会带来太多的选择和机会，其结果是最初通向成功的那个专注点土崩瓦解。成功变成了失败的催化剂，它会让人陷入吉姆·柯林斯（Jim Collins）所称的“盲目

地追求更多”的境地。要想走出这种困境，出路就是自律地追求“更少，但更好”。百折不挠、坚持不懈地追求精要事务，排除非精要的一切，建立一个让执行毫不费力的行为体系，这就是本书所要传达的内容。

Q: “精要主义者认为几乎一切都不重要”，有很多人质疑这个观点。您想通过它传达什么？

A: 很多成功者无法作出更多贡献，原因就在于他们坚持认为一切都是重要的。但是，精要主义者懂得如何将真正重要的事务和其他一切区分开来。

我们生活的世界中，几乎所有的一切都是毫无价值的，只有极少量事物具有非凡的价值，这是我们无法挣脱的现实。正如领导力与人际关系大师约翰·麦克斯韦尔（John Maxwell）写的那样：“几乎一切事情的次要性再怎么高估都不过分。”

Q: 这本书多次登上《纽约时报》的畅销书榜单，说明人们对这个话题反应热烈。本书所揭示的问题究竟有多大的普遍性？

A: 这个困扰我们的问题，其怪异之处在于，我们不仅知道它的存在，还为此津津乐道。最近就有人向我炫耀说，上两周她忙得焦头烂额，每晚平均只睡四个小时。她不仅不抱怨，还引以为豪。这种现象普遍存在。

为什么非常理性的人会做出如此不理性的行为？我觉得是因为我们被囚禁在了一个泡沫里，它大得无边，今天要在发达世界生存下去，必然会受到它的影响和感染。它是一切泡沫的泡沫：不仅映照出过去的泡沫（不管是郁金香、硅谷还是房地产泡沫），还加固了所有这些泡沫。我把它称为“升级版泡沫”。

泡沫的本质在于，某种资产的价值被荒谬地高估，一直到最终——泡沫破裂，我们不得不挠破头皮反思自己为什么一开始会激进得如此非理性。现在，被我们高估的资产就是那种什么都要做、什么都要有和什么都要实现的妄想。

Q：全新的精要主义观点对哪种传统的成功智慧构成了挑战？

A：成功意味着成为超人，能搞定一切，这种花言巧语欺骗了我们。当然，我们自己也在炫耀忙碌，以为忙就是成功和重要性的不二准则。精要主义就是要打破这种用忙碌衡量成功的浅见。

Q：您的公司**THIS**曾经与**Apple**、**Airbnb**、**Pixar**、**Yahoo**、**Google**、**Twitter**、**Facebook**、**Symantec**等公司进行过合作。您是否可以告诉我们，这些公司大体上追求的是什么？

A：因为考虑到能力的诅咒，这些公司对成功人士和团队的低效另有一番见解。它们聘用的都是精英型员工，能够胜任很多种工作。但让员工承担多种工作的风险在于，他们可能最终只能在无数个方向上取得一星半点的进步。人们觉得他们进展快速，但实际上不过是晕动症，而不是蓄势向前。因此，一般来说，这些公司希望自己的员工和团队专注于少量事务，从而创造更高的投资回报。

THIS公司已经开发出了一套系统，能帮助整个企业取消无意义的会议、浪费时间的电子邮件和无目的的项目，代之以在少量能带来改变的重要项目上的投资。事实证明，精要主义对企业以及在这些企业中的个体都大有益处。

Q：在您的企业或个人咨询中，有没有一些极具启发性的案例可以和大家分享？

A：就个人而言，我最近听一个人说，一场可怕的故事害他患上了偏

头痛，他每周只能工作15~20个小时。他说，《精要主义》这本书让他重新获得了生命，因为他认识到自己可以做到事半功倍。他真的做到了。和他的谈话很有意义。

就企业而言，我刚刚和一家大型企业的首席执行官通完电话，他已经运用精要原则让整个公司的面貌大为改观。这就是THIS做的事情：帮助每个人专注于最重要的事物，从而帮助企业实现突破升级。

Q：正规教育对于您目前所取得的成绩有多大的帮助？

A：我的本科专业是新闻学，它是仅有的几个能教你如何正确提问的专业之一——几乎所有的正规教育都在教你如何找到正确的答案。就其本身而言没什么问题，但是正确地提问是一种更高级、更有价值的技能。

Q：在您看过的电影中，您认为哪部最好地诠释了精要原则？为什么呢？

A：我喜欢电影《甘地》。甘地本人就是一个精要主义者，这部电影捕捉到了这一点。因为目的统一——为印度人民争取独立，他将其余一切都从人生中排除了。

近期我在南非演讲时，参观了位于凤凰城的甘地故居。我读到了甘地的诗，据说这是他写的唯一一首诗。在诗中，他写下了“把自己降为零”的诗句。这句诗可以作为他不断进行的精要实验的名称：他穿着家纺布（印度土布）衣服，并鼓励他的跟随者也这样穿；他三年不读报，因为他发现报上的内容只会给他的生活徒增无意义的混乱；他简餐缩食35年；他每周都有一天沉默不语。如果说他摒弃了消费主义，未免太过保守——离世的时候，他只有不到十样东

西。他有意不担任任何形式的政治职务，却成了印度国父。

甘地做出的贡献远不止于印度。阿尔伯特·爱因斯坦曾说：“在未来的时代，可能极少有人会相信，这样一个血肉之躯曾经在地球上匆匆走过。”我们无法否认，甘地的人生是一种真正有意义的人生，对精要事务的专注能力和对非精要事务的决绝摒弃是他成功的关键。

Q：哪本非商业类书让您学到了最有价值的商业经验？

A：我能不能以一篇文章而不是一本书来回答这个问题？这篇文章由田纳西·威廉斯（Tennessee Williams）所著，最初发表在《纽约时报》上，讲述了他在发表备受好评的剧作《玻璃动物园》（*The Glass Menagerie*）之后的经历。文章标题为《成功的灾难》（*The Catastrophe of Success*），该标题就清楚表明了他的论点。威廉斯在文中描述了剧本成功之后，他的生活中发生的种种变化，以及他如何偏离了那些最初使他成功的重要事情。

我被一个问题困扰了15年之久：“为什么已经很成功的人和团队不能突破进阶到更高的层次？”其原因正如威廉斯揭示的那样，就是成功。这显然是一个违反直觉的回答，并非显而易见。我们所有人，都需要学习如何在成功中变得成功。

Q：彼得·德鲁克说：“没有比高效率做无用功更无用的事了。”您怎么看这句话？

A：这显示了德鲁克的才华，我完全认同这个观点。绝大多数提高生产率的方法都不外乎增强适应性，提高效率。精要主义的本质在于排除最重要事务以外的一切，因为我们正在努力完成的事情中有太多是根本不应该做的。

Q: 很多人抱怨他们没有时间锻炼。精要主义如何能帮助他们获得锻炼时间？

A: 锻炼本身不是什么难事，每天决定是否锻炼才是真正让人头疼的。在任何一天，我们只有这么多可锻炼的项目。解决办法就是把这些项目发展为一种体系或常规，让未来每天的执行变得更容易。

Q: 您认为在未来，精要主义会变得更重要吗？

A: 我认为我们在未来面临的挑战还是和现在一样，但它将是一个被更加普遍接受的痛点。德鲁克说：“在几百年中，如果从一个更长远视角来书写我们这个时代的历史，史学家们可能会发现最重要的不是技术，不是因特网，也不是电子商务，而是人类的境况发生了史无前例的变化。这是历史上第一次，拥有选择的人数有了极大增长，而我们的社会却还完全没有准备好应对这一切。”

因此，对于首席执行官们来说，真正的挑战在于，当面临的选择如此之多时，如何授权整个企业专注于真正重要的事务上。最优秀的首席执行官能够在1米到1万米之间创造出清晰的目标和战略，并鼓励人们围绕焦点团结协作。听起来似乎很简单，但需要一个精要主义者授权企业所有成员都成为精要主义者，并付诸实践。

Q: 哪个问题您本来希望在采访中被问到，却没有问到？您将怎样回答？

A: 你决不会做的事情是什么？答案是：“听滚石乐队的歌。”因为我是披头士的粉丝！

Q: 最后您还有什么想说的？

A: 我想说，中文版《精要主义》的出版让我感到责任重大。当我每

天早上5点起床写这本书的时候，心中曾一次次地希望，这本书将给别人带来变化。现在，看到全世界有这么多人会受到此书长久的影响，这对我来说很有意义，我也对此心怀感激。



扫码下载“湛庐阅读”App，
搜索“精要主义”，
查看作者对本书的精彩解读视频。

前 言

The Essentialist

设计人生，以一种精要的方式

生之智慧，在于摒弃不必要之事。

——林语堂



非精要主义者



精要主义者

思考	所有事，所有人	更少，但更好
	“我必须做。”	“我选择做。”
	“一切都重要。”	“重要的只是少数。”
作法	怎样把一切都安排妥当？	怎样权衡取舍？
	盲目地追求更多	自律地追求更少
	对最紧迫的事情作出反应	停下来区分、辨别重要的事情
收获	不加思索地满口答应	非精要之事一律说“不”
	在最后一刻仓促执行	为方便执行清除障碍
	生活不如意	生活充满意义
	承担太多任务，工作成了折磨	精心挑选有价值的工作
	感到失控	一切尽在掌控中
	不确定该做的事情是否已经做了	完成该做的事情
	感到不知所措、精疲力竭	享受旅程的快乐

萨姆·埃利奥特（Sam Elliot）¹是美国硅谷一位颇有才干的经理人，但当他的公司被一家官僚习气浓重的大企业收购后，他发现自己精力涣散，不堪重负。

在新的岗位上，他认真负责，忠于职守，对很多要求甚至不假思索就一口答应。结果，他终日奔波于各种会议之间，企图让每个人都满意，把每件事情都做好。然而工作质量的下降，令他感到压力陡增。他仿佛成了处理琐事的专家，工作不仅不再给他带来满足感，而且还让他费力取悦的那些人也感到不满。

萨姆沮丧不已之时，公司为他安排了一项提前退休计划，但当时不过才五十出头的萨姆不想就这样彻底退休。于是，他萌生了创办一家咨询公司的想法，为先前从事过的行业提供咨询服务，甚至想到了要当现在公司的咨询顾问，赚雇主的钱。但是，这些想法都不能令他心动。因此，他向一位良师益友请教，得到了一个令人吃惊的建议：“留在现在的公司，但是只做咨询顾问的工作，其余一律不做。不要告诉任何人。”简单来说，这位导师建议他只做自己认为重要的事，忽略其他一切别人要求的事。

萨姆听从了这项建议！他制订了删减烦琐事务的日常计划，并开始说“不”。

开始时他还有点犹豫，试探着用一种较为保守的标准来衡量那些要求：“如果我有时间、有条件的话，能不能完成这项任务？”如果答案是不行，那就拒绝。令他又惊又喜的是，尽管人们被拒绝时会流露出失望的表情，但也会尊重他的诚实。

这些小小的胜利让他变得更加大胆。现在，如果有人提出要求，他会停下来，用一种更加严格的标准评估：“这是我现在必须用自己的时

间和资源来处理的最重要的事情吗？”

如果不能非常肯定地回答“是”，他就会拒绝这项要求。而令他高兴的是，虽然同事们刚被拒绝时有些失望，但很快就变得更加尊重他。

萨姆受到了鼓舞，开始将这种选择标准推行到所有事情上。过去，他总是自愿承担那些最后一刻才出现的工作，现在他想方设法不再主动揽活；过去，他常常是所有收件人中最早回复的那个，现在他不再争强好胜；如果兴趣不大，就不再参加电话会议；如果不需要信息，就不再去旁听每周的新情况通报会；如果不能有什么直接的贡献，就不再出席已经排期的会议。他解释说：“仅是一个邀请，并不足以让我参加。”

刚开始，他觉得这样做有些自我放任。但是通过选择性地接受任务，他为自己赢得了空间，进而找到了创造的自由。他可以集中精力攻克一个项目，作出周密计划，提前预测困难，努力清除障碍。由于不再开足马力试图搞定一切，萨姆得以专注于该做之事。只做真正重要的事情并忽略其他一切，这个新发现的专注式工作方法令他的工作恢复了高质量。他不再试图在多件事情上取得零星的进展，而是以排山倒海之势集中全力做成真正重要的事情。

坚持了几个月后，萨姆发现自己不仅白天的工作时间更加充裕，而且晚上下班后还有了更多时间享受家庭生活。他说：“我重新找回了自己的家庭生活！不用再加班至深更半夜才回家了。”现在，他下班后就关掉手机，不再做时刻听命于手机工作指令的奴隶。他还经常去健身房锻炼身体，与妻子一起外出共进晚餐。

令他始料未及的是，这些尝试并未产生负面的影响。主管没有为难他，同事们也没有怨恨他。恰恰相反，留给他的都是些对他而言意义重大、对公司来说也极具价值的项目，他的工作得到了比以前更多的尊重和重视。他在工作中获得了更多的满足感，业绩排名也随之上升，最终

赢得了职业生涯中最可观的奖励！

这个故事体现了精要的基本价值主张：若要最大限度成就真正重要之事，切忌贪多求全，事事应允。

你呢？你有多少次不假思索地答应别人的请求？你有多少次恨自己答应去做什么事，事后却疑惑不解地自问“我为什么要揽下这件事”？你是否经常为了取悦别人或者避免麻烦而点头答应？“是的”是否已经成了你的默认答复？

现在，请回答下列问题：

- 你是否曾感到心力交瘁？
- 你是否曾感到劳累过度却没有发挥所能？
- 你是否曾觉得自己只擅长做琐碎小事？
- 你是否曾感到自己忙忙碌碌却又效率低下？
- 你是否觉得自己一直在奔忙，却总是到不了任何地方？

如果你对上述任何一个问题的回答是肯定的，你就需要借助精要的力量。

更少，但更好

迪特·拉姆斯（Dieter Rams）曾任德国博朗公司（Braun）首席设计

师多年。他遵从“几乎一切东西都是噪声”的理念，认为只有极少的东西是重要的。他的任务是滤掉噪声，直抵精要。例如，24岁那年，公司安排他与人合作开发一种电唱机。当时的普遍做法是将转盘置于一个实木的盖子内，或者把唱机嵌入到一件起居室的家具中。然而，迪特和团队删繁就简，设计了一种透明塑料盖的唱机。由于这种设计在当时尚属首例，而且又显得过于惊世骇俗，人们担心它会把公司搞垮，因为没有人会买这样的唱机。但是，摒弃非精要之事从来都是需要勇气的。到了20世纪60年代，这种设计开始风靡，最终成了所有其他唱机竞相模仿的设计典范。

迪特的设计标准可以概括为一种典型的简洁原则，在德语中仅用三个单词即可体现：**Weniger aber besser**，意思是“更少，但更好”。这也是对精要最恰当的定义。

精要主义正是对“更少，但更好”的不懈追求。这意味着不是偶尔为之，而是把追求“更少，但更好”当成行事的铁律。

精要主义并不是下一个新年决心要对更多的事情说“不”，不是删减收件箱里的信件，也不是掌握新的时间管理策略，而是不断地停下来问问自己：“这些事情值得我投入时间和精力吗？”世间事务和机会之繁多，远非个人的时间和资源所能及。尽管其中不乏好事，有些甚至非常好，但绝大多数都是无意义的琐事，重要之事少之又少。学会区分差异——过滤筛选所有选项，并从中择其真正精要之事而为之，是精要之准则。

精要主义不是如何完成更多的事情，而是如何做好对的事情。它也不是提倡为了少做而少做，而是主张只做必做之事，尽可能做出最明智的时间和精力投资，从而达到个人贡献峰值。

精要主义者和非精要主义者的行为差异可见下页图片。在这两幅截

然不同的图像中，付出的努力是相同的。在左图中，精力被分散到了很多不同的事上，结果就是在每件事上都取得了零星进展，但都不能令人满意。在右图中，同样的精力被用到了少数几件事情上，由于精力集中从而在最重要的事情上取得了显著进展，获得了令人满意的结果。精要主义者拒绝接受我们能做所有事情的想法，而是要求我们进行真正的权衡取舍并作出艰难决定。在很多情况下，我们可以学着做一些一次性决定，它们相当于未来的一千次决定，这样就省得一次次用同样的问题来折腾自己。

精要主义，就是对自己的生活进行设计规划，而不是事事依循默认设置。精要主义者并不是应激性地作出选择，而是有意识地区别重要的少数和不重要的多数，精减非精要之事，然后清除障碍，从而为精要之事提供清晰顺畅的实现路径。换句话说，精要是一种自律的、系统的方法，用于决定个人贡献峰值的所在，然后轻松地执行。

精要主义让个人选择尽在掌控之中。它是一条能让人获得新的成就和意义的道路。在这条路上，我们尽享旅程的快乐，而不仅仅是到达目的地的欣喜。然而，尽管有种种益处，依然有太多的阻力干扰我们奉行“更少，但更好”这条铁律，这或许解释了为何有那么多人最后迷失在了非精要的道路上。



成功的悖论

在加利福尼亚州一个阳光明媚的冬日，我去医院看望妻子安娜。虽然身处医院，安娜看起来仍容光焕发。不过我知道，她其实非常疲惫，因为前一天刚刚诞下我们的大女儿，一个健康活泼、体重约3.2公斤的宝宝！

然而，在这个原本应该是我一生中最幸福、最安详的日子里却充满了紧张和不安。漂亮的新生儿正躺在疲惫母亲的臂弯里，我却还在接工作电话，回工作邮件，心中充满了会见客户前的那种压力和紧张。我的同事在邮件中写道：“星期五下午一点到两点之间最好别生孩子，因为我需要你过来参加会议，同客户见面。”现在正是星期五，尽管我非常肯定（或者说起码我希望）这个邮件不过是个玩笑而已，但依然感到了要去参加会议的压力。

我当然知道自己该干什么。这个时候我显然应该陪在妻子和宝宝身边。因此，当有人问我是否出席会议时，我鼓起勇气说……

“是。”

当妻子和刚出生的宝宝躺在医院里的时候，我去参加了那个会议，这真令我羞愧难当。事后，我同事说：“你决定来参加会议，这会赢得客户的尊重。”但是，我并没有在客户的脸上发现所谓的尊重。相反，他们都觉察出了我的不悦，这在他们脸上反映出来了。我到底干了什么？答应参加仅仅是为了不得罪人，而这样做却伤害了我的家庭，违背了我的原则，甚至还影响了客户关系。

结果，那场会面没有产生任何结果。不过就算有进展，这也是一个愚蠢的交易。为了让每个人都高兴，我牺牲了最重要的东西。

在对这件事情的反思中，我学到了重要的一课：

**If you don't
prioritize your
life, someone
else will.**

如果不能自己安排生活的优先次序，
就只能任由别人替你安排。

这段经历重新燃起了我对阅读的兴趣，我不知疲倦地探究为何聪明人总是用别的方式来作出个人生活和职业生涯中的各种决定？为什么我们的潜力要远远大于我们实际运用的？我们如何更深入地挖掘运用自己以及其他所有人的内在潜力？

为了找到这些问题的答案，我放弃了英格兰法学院的学习，前往加利福尼亚州斯坦福大学攻读研究生。我花了两年多时间与人合著了一本书，名叫《乘法器：最优秀的领导人如何让每个人变得更聪明》

（*Multipliers: How the Best Leaders Make Everyone Smarter*）。在这本书的启发下，我在硅谷创立了一家战略和领导力公司，它让我能够和全球

最有意思的一些公司中最能干的一部分人一起工作，帮助他们成为精要主义者。

在工作中，我见过被重重压力团团围住而喘不过气来的人，这样的人世界各地都有。我指导过的人中不乏“成功”人士，但他们却因为急于要把一切都做到尽善尽美而拼死拼活，承受着无以言表的折磨。我也见过有些人受制于控制欲极强的领导，意识不到自己并不是“必须”要去做所有这些别人要求的事，因为那纯粹是瞎忙，别人并不领情。我曾不知疲倦地研究过为什么那么多聪明、机灵、能干的人会深陷非精要主义的囹圄而无法自拔。

我的发现让我大为吃惊。

我曾和一位干劲十足的经理人一同工作过，他从小就从事并热爱技术工作。凭借自己的专业知识和工作激情，他赢得了越来越多的机会。因为急于建功立业，他尽其所能地广泛阅读，满怀激情地追求自己能做的一切。当我见到他的时候，他极度活跃，试图什么都学，什么都做。他好像每天都能找到一个新的爱好，有时候甚至是每个小时。在这个过程中，他没办法从不重要的多数中找到重要的少数，他觉得一切都重要。结果，他的精力变得越来越分散，只能在多件事情上取得零星的进步。劳累过度，却又未人尽其才。就在这个时候，我向他勾画了本书第8页图片中左边的那个图形。

他盯着那个图形看了很长时间，直到那种诡异的沉默被打破。随后，他情绪激动地说道：“这就是我的人生！”然后，我又画了右边的那个图形。“如果你能想明白，为了实现个人贡献峰值。自己最该做的一件事情是什么，那么将会发生什么呢？”我问道。他由衷地回答说：“就是这样。”

其实，很多聪明又有抱负的人都难以回答这个问题，他们都有极为

合理的理由。其中之一就是在社会中我们因为正确的行为（说“不”）而受到惩罚，因不正确的行为（说“好”）而受到奖赏。前者往往会令当时的气氛变得尴尬，而后者通常会得到赞美。这种情形导致的后果被我称为“成功的悖论”，主要可概括为以下4个可预见的阶段：

- 第一阶段：如果我们真的有清晰的目标，它有助于我们取得成功。
- 第二阶段：当我们成功时，就会被冠以“可以去找的人”的称号，成为“老好人”，只要你需要，就总能找到这个人。这样，我们就能够得到更多的选项和机会。
- 第三阶段：当我们拥有了更多的选项和机会，实际上也就被要求付出更多的时间和精力，这样便会导致精力的分散。我们被千头万绪的事情牵扯得越来越无法集中精力。
- 第四阶段：因为精力分散，我们无法实现原本可以实现的个人贡献峰值。成功所带来的后果最终瓦解了最初指引我们走向成功的那个真正清晰的目标。

奇怪的是，对成功的追求反而成为失败的催化剂，这一点需要反复强调以便我们能成功。换句话说，成功会让我们偏离最本质的事情，这些事情最初曾是成功的源泉。

这样的例子我们身边比比皆是。吉姆·柯林斯（Jim Collins）在其《再造卓越：巨人也会倒下，但有的公司从不放弃》（*How the Mighty Fall*）一书中探讨了那些曾是华尔街宠儿的公司为何最后会走向衰败。

他发现，很多公司最终失败的关键原因是“不假思索地贪多求全”。很多公司如此，很多在这些公司里工作的员工也是如此。但是这其中的原因何在？

非精要主义来袭

在几大趋势的共同作用下，非精要主义如风暴般强势袭来。

选择的数量呈指数级增长

我们都已经注意到了在过去十年中，人们面临的选择呈指数级增长。也许正因为如此，在这众多的选择中，我们反而看不到最重要的选择了。

正如彼得·德鲁克所说：“几百年后，当我们回看历史洪流中的当下这个时代，史学家们可能会发现，最重要的不是技术，不是因特网，也不是电子商务，而是人类的境况发生了史无前例的变化。这是历史上第一次，拥有选择的人的数量有了极大增长，而我们的社会却还未完全准备好应对这一切。”

我们没有准备好，在一定程度上是因为这是第一次，我们面临的选择之多大大超过了我们对此进行管理的能力。我们无法将重要之事从不重要之事中筛选出来。心理学家们称之为“决定疲劳”：要做的决定越多，决定的质量就越差。

高度互联与信息过载导致社会压力增大

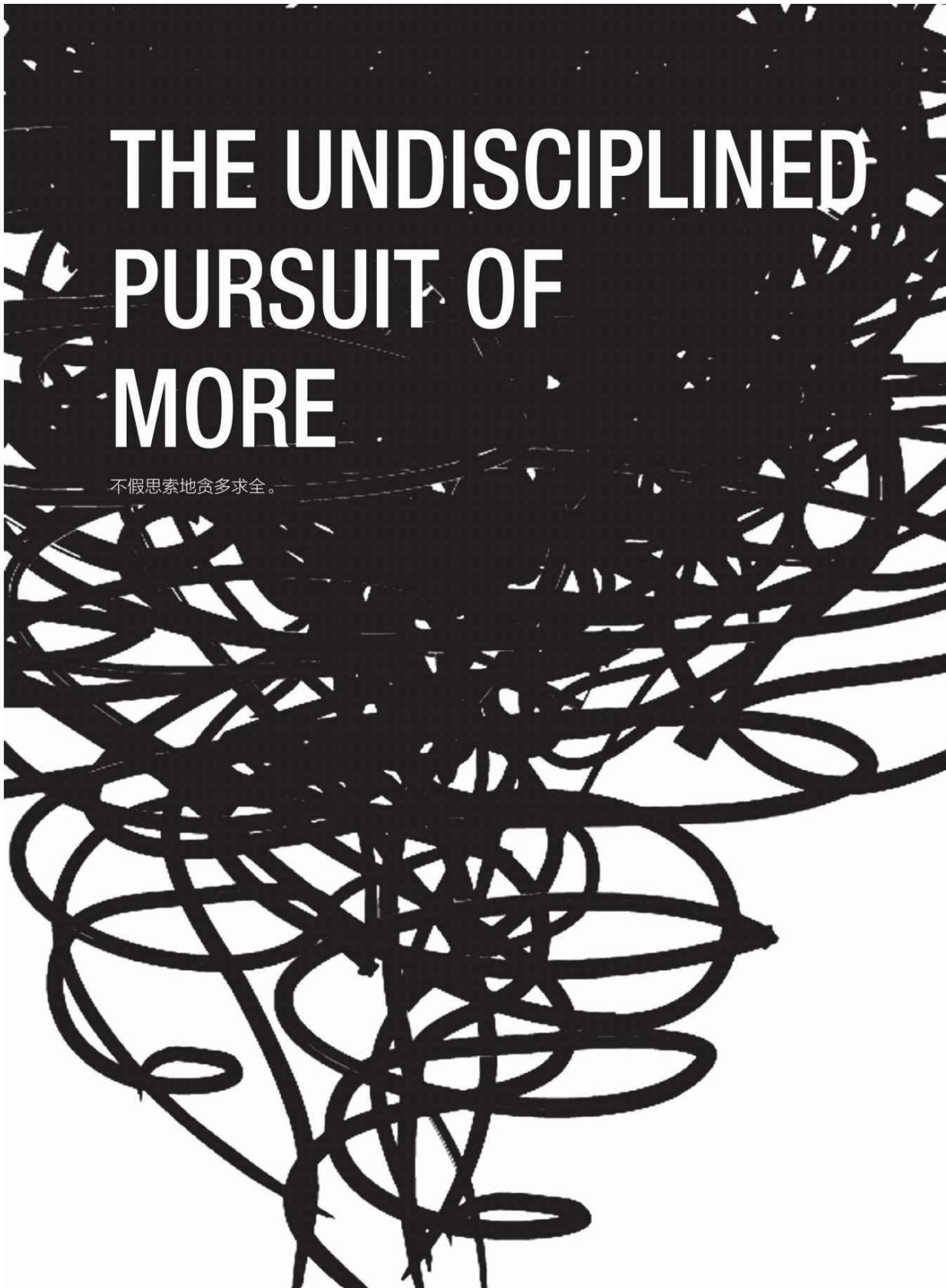
除了选择的数量呈指数级增长之外，干扰我们决定的那些外部影响因素的力量和数量也在增长。虽然现在有很多关于高度互联时代是什么样的，以及信息过载会如何分散注意力的论述，但是更大的问题是这种

连接如何导致了社会压力的增大。今天，技术已经为人们之间分享彼此的
关注点扫除了障碍，现在的问题不是信息过量，而是观点过多。

“你可以拥有一切”的错误引导

我可以拥有一切，做好一切，这种观点我们并不陌生。这种神话已经被兜售了很长时间，我相信今天所有活着的人都不能对此免疫。它在广告中被兜售，在公司里被追捧，还被植入到职位描述中，而这样的职位往往都会提出一大串的技术和经验要求。它也在需要参加过几十种课外活动的大学入学申请中出现过。

新的现象是，在如今选择和期望呈指数级增长的时代，这种神话将会产生多大的破坏力？它会令原本就已经忙碌过度的生活雪上加霜；它会造成一种嘴上宣称工作生活平衡，但实际上却希望员工的智能手机全年全天候开机待命的公司文化；它会让讨论10个“最优先项目”的员工会议开得不带任何讽刺色彩。



THE UNDISCIPLINED PURSUIT OF MORE

不假思索地贪多求全。

“priority”一词最早于15世纪进入英语。它是单数名词，意为“第一”或“优先的事物”。在接下来的5个世纪里，它一直保持着单数形式。直到20世纪，人们才创造了它的复数形式，并开始讨论“priorities”。我们为此找了一个不合逻辑的理由，就是通过改变这个词语来改变现实。现在，通过某种方式我们就可以拥有很多“第一”的事情，许多个人和公司就常常这么做。有个领导跟我说过他在一家公司的经历，那家公司有“优先1号，优先2号，优先3号，优先4号和优先5号”。乍一听给人一种有很多事情都是优先级的印象，但实际上什么都不是。

当我们努力要做到一切和拥有一切的时候，会发现自己正在牺牲一些收益作为交换条件，而这些收益是自己永远不会去主动追求的。当我们不去有目的地慎重选择该把精力和时间集中于何处时，别人——老板、同事、客户，甚至家庭，就会替我们作出选择。用不了多久，我们将对所有有意义的和重要的事情视而不见。要么自己慎重选择，要么受制于别人的计划，非此即彼。

曾经有一位名叫布鲁尼尔·韦尔（Bronnie Ware）的澳大利亚护士，他照顾过生命只剩最后3个月的病人，并记录下了这些人讨论最多的遗憾。在这些遗憾中，最大的遗憾是：“我希望有勇气真正按自己的想法生活，而不是按照别人的期望而活。”

为自己而活，不只是偶尔说说“不”，而是要有目的地、深思熟虑地、战略性地淘汰那些不重要的事情；不仅要摆脱那些明显浪费时间的事情，还包括放弃一些很好的机会；不要疲于应付那些让你在多条战线上作战的社会压力，要学会减少、简化，并通过淘汰其余一切来聚焦于绝对重要之事。

人生与事业的断舍离

这本书对你的人生和事业的意义，与一个专业的收纳达人将给你的衣柜带来的变化相似。设想一下，如果你从不整理自己的衣柜，它会变成什么样子？它会是整整齐齐，架子上只挂着很少的几件你爱穿的衣服吗？当然不是。如果你从不有意识地去整理它，它就会凌乱不堪，里面堆满了你很少穿的衣服。有时甚至变得实在无法使用，逼得你不得不去整理。所以，除非你有一套很规范的整理方法，否则，你要么继续留着那一堆很少穿的衣服，因为不知道哪些衣服该扔；要么因为不小心扔了可以穿而且想留着衣服而后悔不已；要么还留着一堆其实不想留但又没办法处理的衣服，因为你不确定它们被放在哪儿了，也不知该如何处理它们。

随着不穿的衣服越积越多，衣柜就会变得凌乱拥挤。同样，随着好心承诺和应承下来的事情越积越多，我们的人生也会变得杂乱无章。而且，绝大多数这样的努力根本没有尽头。除非我们找到系统的方法进行清理，否则这样的杂乱状态将会永远持续下去。

以下是精要主义者整理衣柜的方法。

1.探索

不要问：“这衣服我以后还会有机会穿？”而要问一些更加规范、更加尖锐的问题：“我喜欢这件衣服吗？”“我穿上这件衣服好看吗？”“我穿这衣服的次数多吗？”如果回答是否定的，那么你就知道该淘汰它了。

在你的个人生活和职业生涯中，问自己喜欢哪件衣服就如同问自己：“这件事情或这份努力有可能会对我实现目标做出最大的贡献吗？”本书第二部分将帮助你弄清这些事情是什么。

2.排除

假设你把衣服归为两类，一类是“必须保留”，另一类是“可能需要处理掉”。但是，你真的打算把“可能需要处理掉”的衣服塞进袋子里送走吗？归根结底，这里面有一种沉没成本：研究表明，对于自己已经拥有的东西，人们对其价值的评价要高于其实际价值，因此要放弃就更难。如果这还不能令你信服，那么就试试这个要命的问题：“如果我现在没有它，那么我会花多少钱来买它？”这么问往往能奏效。

换句话说，仅仅确定了哪些事情和努力不能帮你实现个人贡献峰值还不够，你还要积极地淘汰它们。本书第三部分将告诉你如何淘汰那些非精要的东西，不仅如此，它还要告诉你该怎么做才能帮你赢得同事、老板、客户和对手的尊重。

3. 执行

如果要使自己的衣柜保持整洁，你就要经常整理。用一个大袋子装好要扔掉的东西，把想留下来的东西堆成一小堆。你要了解弃物存放地和当地旧货商店的营业时间，并预先安排好去那儿的时间。

也就是说，一旦弄清了哪些活动和努力要保留——就是能让人实现个人贡献峰值的那些，你就需要一种方法来使执行变得毫不费力。本书第四部分将教会你创造一种能让重要的事情完成得毫不费力的方法。

当然，生活并不像衣柜里的衣服那般静止不变。衣服可以静静地待在衣柜里，早上我们离开时是什么样，晚上回去还是什么样（除非家里有小孩！）。但在人生的“衣柜”里，新衣服——占用我们时间的新需求，却在源源不断涌来。想象一下，如果你每次打开衣柜的门，都发现别人已经在里面塞了他们的衣服；如果你每天早上把衣柜收拾干净，到了中午却发现它已经被塞得满满当当，那是怎样一种感受。不幸的是，这是很多人生活的真实写照。有多少次你按照工作计划开始新一天的工作，但到了上午10点却已经在忙其他的事情，或是完全落后于工作进度

了？有多少次你在早上列了一个“待办事项”的清单，但到了下午5点却发现这个单子越拉越长？有多少次你期盼着能够待在家里与家人共度一个清静的周末，但直到周六早上你还淹没于各类琐事、各种约见和无法预见的倒霉事儿里？不过幸好，这里有条出路。

精要主义就是要为整理人生的“衣柜”创造一种方法。但是，它可不像你整理衣柜那样，只要一年整一次，一月整一次，或一周整一次就行。它是一种铁律，你必须在每次面临答应或婉拒的决定时执行。它是一种方法，帮你在很多好东西和一小部分真正重要的东西之间进行权衡取舍。它会告诉你如何做到更少，但更好，这样你就可以在人生的每个珍贵瞬间里尽可能地收获更多。

本书将告诉你，如何去过一种真正属于自己的生活，而不是依从他人的期望而活。它会给你指明一条路径，让你的人生和事业更有效率，更有成果，更令人印象深刻。它将告诉你一种系统的方法，让你弄清楚什么是重要的，淘汰掉不重要的，并轻松地执行那些精要之事。简而言之，它将教会你把“追求更少”运用到生活的方方面面。

从精要思维到精要行动

本书共分4个部分。第一部分勾勒了精要主义者的思维模式。接下来的3个部分将这种思维模式转化为一种系统化程序，用来执行“追求更少”这条铁律，你可以把它们用在你碰到的任何一种情况和做出的任何一种尝试中。

精要主义者的核心思维模式：时间与精力只用于有意义的事

这部分概述了我们试图躲避的三种现实，没有它们，精要思维就毫无意义或无从谈起。

- 选择：我们可以选择如何支配自己的精力和时间。若无选择，讨论权衡取舍就没有意义。
- 甄别：几乎一切都是噪声，只有少数事情价值非凡。正因为如此，我们应该费时费心地去辨别什么才是最重要的。有些事情比其他事重要得多，因此费力找到它们是磨刀不误砍柴工。
- 取舍：我们不能什么都有，也不能什么都做。如果可以的话，我们就大可不必评估和淘汰各种选项。一旦接受了取舍的现实，就不要再问：“我怎样才能完成这一切？”而是要问一个更加诚实的问题：“我要解决的问题是什么？”

只有理解了这些，我们才能像精要主义者那样思考。一旦彻底接受并理解了它们，那么本书中所提供的很多方法在我们的生活和工作中就会变得自然而然，并成为一种本能。践行这些方法要遵循以下3个简单的步骤。

探索：区分无意义的多数和有意义的少数

精要主义的一个悖论就是，相比非精要主义者，精要主义者实际上会探索更多的选项。尽管非精要主义者致力于什么都做或者几乎什么都做，但并未经过实际探索；而精要主义者在真正行动之前会系统地探索和评估多个选项，因为他们会在一两个大的设想或项目上大干一场，所以一开始就会慎重地探索更多的选项，来确保自己做出的选择是正确的。

通过运用更加严格的标准，我们能够挖掘出大脑中复杂的搜索引

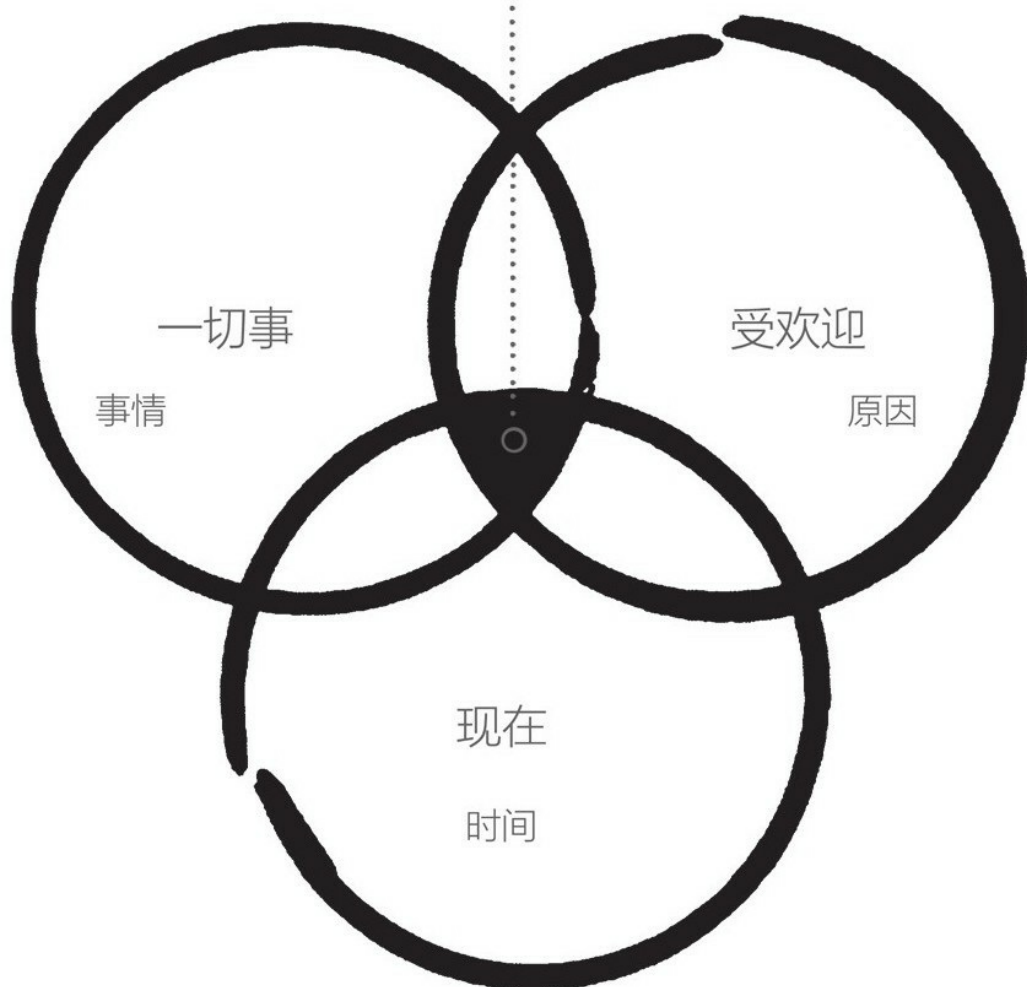
擎。如果我们搜索“一个好机会”，会找到很多页的搜索结果。但如果我们可以进行高级搜索并且问3个问题：令我深受启发的是什么？我特别擅长的是什么？在这世上能满足我重大需求的是什么？自然就不会有那么多搜索结果页需要查看了。但这正是练习的重点所在。我们不是要找过多的好事情去做，而是寻找个人贡献峰值的所在：重要的事情，恰当的方式，合适的时间。

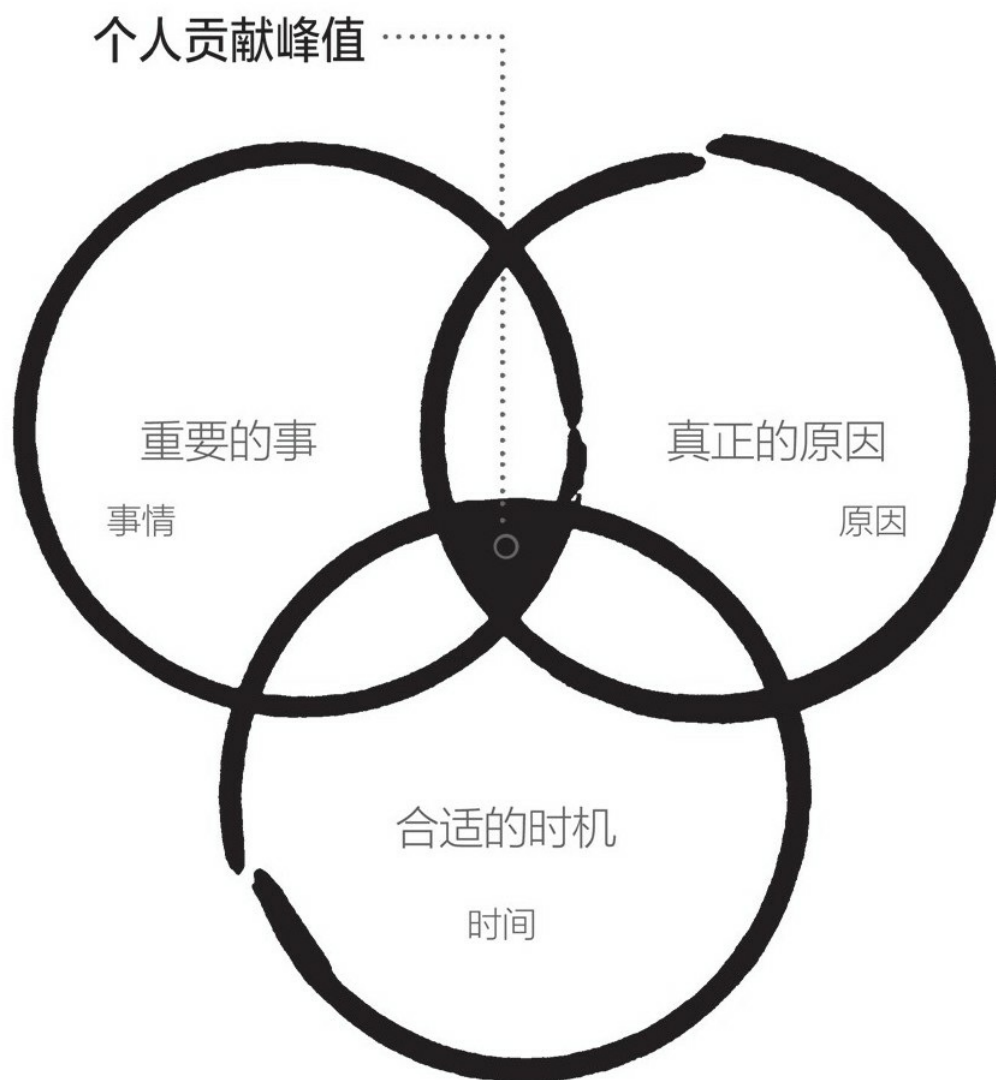
精要主义者在探索上面花了尽可能多的时间，倾听、辩论、质问、思考。但并不是探索完就结束了，探索的最终目的是从无意义的多数里识别出有意义的少数。

排除：摆脱无意义的多数

我们很多人答应一些事情，是因为要急于取悦别人或者让自己不同凡响。然而，说“不”才是实现个人贡献峰值的关键所在。彼得·德鲁克曾说：“说‘不’可以让人变得高效，因为说‘不’的人会说‘这不是我的事’。”

个人努力峰值





排除非精要之事意味着要对有些人说“不”，并且是经常性说“不”。这意味着要将那些社会期待拒之门外，要做到这一点极需勇气和同情。因此，淘汰非精要之事不仅是心智的训练，也是情感的训练，比如必要的时候对社会压力说“不”。

考虑到取舍的现实性，我们不能选择什么都做。真正的问题不是我们怎样才能做到这一切，而是谁来决定我们做什么和不做什么。记住，当我们丧失选择权的时候，别人会替我们做出选择。所以，要么慎重地选择有所不为，要么不由自主，任人摆布。

本书这一部分提供一种淘汰非精要之事的方法，让我们赢得时间去做真正重要的事。只有这样我们才能建立一种平台，尽可能地使执行变得毫不费力。

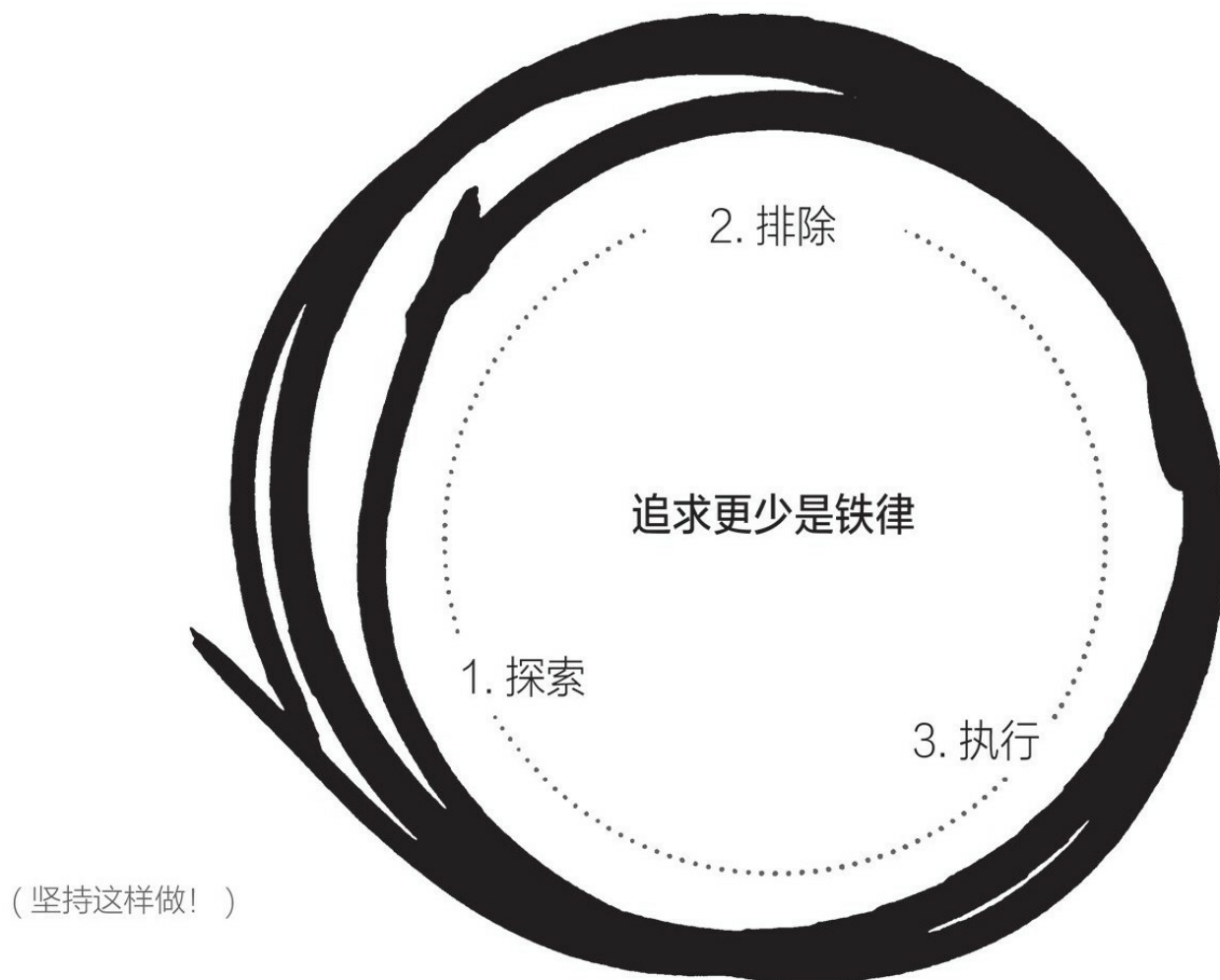
执行：让有意义的少数做起来毫不费力

不管是要完成一项工作任务，还是要迈向事业的新台阶，抑或是要为伴侣安排一个生日聚会，我们往往都会觉得执行的过程充满了困难和阻力，必须硬着头皮强迫自己去“完成”。精要方法却截然不同。不是逼着自己去执行，而是把时间省下来创造一种方法，用来扫除障碍，使任务执行起来变得尽可能轻松自如。

探索，排除，执行并不是呈周期性循环的独立事件，若我们能持续运用，就能获得越来越多的收益。

恰逢其时的一种理想

法国剧作家、小说家维克多·雨果曾说：“世上有一种东西比所有的军队都更为强大，那就是恰逢其时的一种理想。”“更少，但更好”正是一种恰逢其时的信条。



一旦我们允许自己在选择去做什么时变得更挑剔，一切都会改变。我们立刻就拥有了开启人生下一辉煌阶段的钥匙。我们拥有无尽的自由来学习排除非精要之事，不再受控于别人的安排，并自己做出选择。有了这种无敌的力量，我们就能找到自己对个人的人生和事业，乃至全世界的贡献峰值之所在。

假如学校取消课外作业，代之以能给社区带来变化的重要项目，那么结果会如何？假如所有的学生都能有时间考虑自己未来的个人贡献峰值，以免毕业时无路可走，那么结果又是如何？

假如企业不再开那些无意义的会议，而是让员工有思考空间并开展

最重要的项目，那么结果会如何？假如员工推掉一大堆浪费时间的电子邮件、目的不明确的项目、没有成效的会议，以便为公司和自己的职业生涯倾尽所能，那么结果又是如何？

假如社会不再告诉我们要拥有更多本事，而是允许我们去创造更多呼吸和思考的空间，结果会是如何？假如社会鼓励我们拒绝去做那些非己所愿的事情，拒绝即使缺钱也要买自己并不需要的东西，拒绝去取悦那些自己不喜欢的人，那么结果又如何？

假如不再有人向我们过分吹嘘更多的好处，也不再有人贬低更少的价值，那么结果会如何？

假如我们不再因忙碌而自认为自己很重要，并为此洋洋得意，结果会如何？假如我们在乎自己花了多少时间来悉心聆听，沉思默想，享受与生命中最重要的人在一起的时光，那么结果又会如何？

假如全世界人都不再盲目地追逐更多，而是自律地追求“更少，但更好”，那么结果又会如何？

我憧憬，所有人都有勇气为自己真正喜欢的而活，不再为取悦别人的期望所累。

我憧憬，每个人——儿童、学生、母亲、父亲、员工、经理、高管，乃至全世界的领导人，都能学着去挖掘出更多的智慧、才能和谋略，更积极地投入到一种有意义的生活中去。我憧憬，人们来到这个星球是为了要做些什么，并且能不改初衷，勇敢地去做。我憧憬，一场对话可以启发一项行动。

要驾驭这样的勇气，我们必须踏上正确的道路，因为生命何其短暂，而在这短短的余生里自己究竟要成就什么，是需要我们反思的。诗人玛丽·奥利弗（Mary Oliver）写过这样的诗句：“告诉我，你要干什么/

在这野蛮却宝贵的人生里？”

恕我冒犯，我要求你做更多的停歇，问问自己这个问题。

恕我冒犯，我要求你马上承诺，为自己留出更多空间享受本质和精华。你是否有那么一刻后悔做这个决定？有没有可能你某天醒来时会说：“我要是不这样忠实于自己，而是按别人的要求做完一切不重要的事就好了？”

恕我冒犯，我要帮助你建立一种方法，它将不公平地支持重要的少数多过不重要的多数。

恕我冒犯，我要求你在更大程度上成为一名精要主义者。本书不提倡回到从前那些简单的岁月，不是要你远离电子邮件，也不是要你不上网，更不是要你像一名隐士那样离群索居。那将是开时代的倒车。本书要提倡的是运用“更少，但更好”的原则，活在当下和未来。那将是一种创新。

因此，我要求你别像我女儿出生那天时的我那么傻。我有信心，这个决定会令你受益匪浅。可以想象，如果在这个星球上的每个人都能放弃一件虽不错但却不重要的事情，代之以真正重要的事情，那么世界将会变得多么美好。

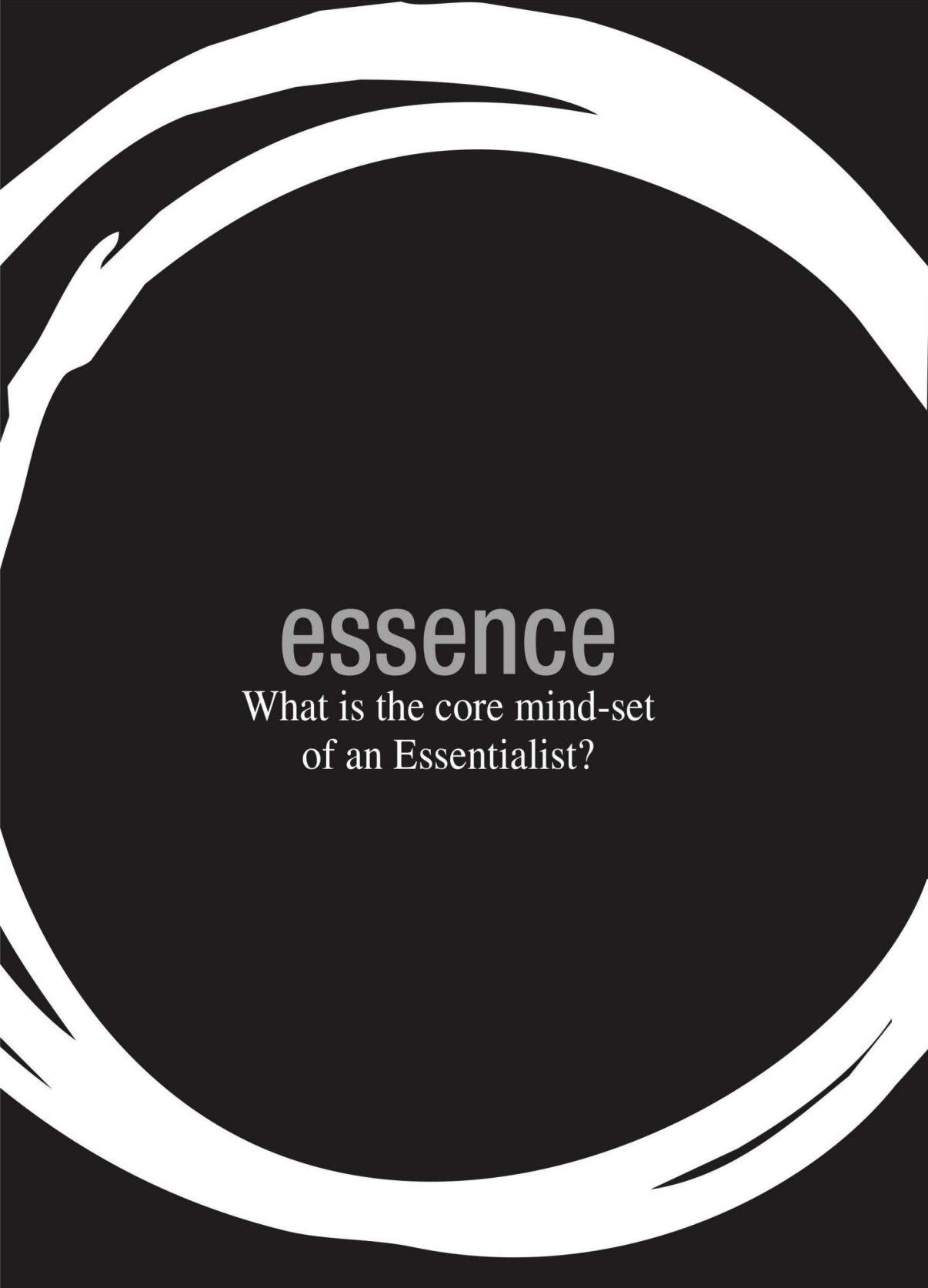
若干年（希望是很多很多年）后，当你走到人生的尽头，可能仍有很多憾事，但是，你不太可能会因为曾追求过精要主义而悔恨。回到此时此刻，为了这个机会——一个对自己真诚的机会，你愿意舍弃什么？到生命终结的那一天，你希望自己今天做出了怎样的决定？

如果你已经准备向自己的内心深处探求对此问题的回答，那么就已在精要主义的道路上准备启程了。让我们一起出发吧。

Part I

精要主义者的核心思维模式



An abstract graphic featuring a large, dark gray circle centered on a white background. This central circle is surrounded by several concentric, slightly irregular white rings of varying thicknesses, creating a layered, tunnel-like effect. The word "essence" is printed in a bold, lowercase, sans-serif font in a light gray color, centered within the dark circle.

essence

What is the core mind-set
of an Essentialist?

导读

时间与精力只用于有意义的事

精要主义不是一种教人再多做一件事情的方法，而是一种教人如何做事的不同方法。它是一种思维方式，但要把这种思维方式内化于心绝非易事。因为某些固有的想法总被人兜售推销，它们不断地把我们拉回到非精要主义的逻辑上来。该部分中每章都会呈现一个非精要主义的谬误，然后用精要主义的正确方法来改正它。

要像精要主义者那样行事，我们必须克服三种根深蒂固的想法：“我必须做”，“这些都重要”以及“我能二者兼顾”。这些想法仿佛神话中的海妖，危险可怕而又魅惑十足，会引诱我们坠落并淹没在浅水水域里。

要抓住精要主义的本质所在，我们应该摒弃这些错误的思维，拥抱三条核心真理：“我选择我要做的事”，“只有小部分事情是重要的”以及“我能做任何事但不是所有事”。这些简单的真理把我们从非精要主义的浑浑噩噩中惊醒，让我们去自由地追求真正重要的事情，并实现个人贡献峰值。

拂去非精要主义的重重迷雾，代之以精要主义的核心逻辑，我们会自然而本能地像精要主义者那样行事。

CHAPTER 01

CHOOSE

The Invincible Power of Choice

模式一

选择，一种战无不胜的力量

人之所以为人，就在于选择的能力。

——马德琳·恩格尔



非精要主义者

“我必须做。”

丧失选择权



精要主义者

“我选择做。”

运用选择能力



我坐在一幢摩天写字楼的大厅里，睁大眼睛盯着手中的一张纸。正是黄昏时分，楼里仅剩的几个人正在为晚上的活动涂脂抹粉，精心装扮。那张纸上涂鸦着的文字和箭头是一场长达20分钟的头脑风暴的产物，写的都是我此刻打算要干一辈子的事情。看着看着，我突然意识到缺了点儿东西——法学院未被列在上面。能注意到这点，是因为我已经在英格兰的法学院学习半年了。

我申请学习法律是因为总有人建议我要“让自己的选择具有开放性”。一旦学成，我就可以当律师，写法学著作，教授法律课程或者从事法律咨询工作。那样我就可以随心所欲地施展自己，或者起码可以这样推断。然而，我几乎从一开始学习法律，就没有在那些试图多面出击的追求中进行选择。我白天读法律书，晚上看那些重量级管理学家的著作，闲暇时间用来写作。这是一种典型的“横跨战略”，企图同一时刻面面出击。结果，我既没有在任何一项追求中彻底失败，也没有在任何一个方面大获全胜。很快，我就开始怀疑所有这些开放的选项究竟有什么重要的？

正当我困惑不解之时，有位在美国的朋友打来电话邀请我去参加他的婚礼。他已经买好机票并且寄了过来。我满怀感激地接受了他的邀请，离开英格兰，踏上了一次意想不到的冒险之旅。

在美国，我抓住一切机会与教师和作家们会面。有一次，我见到了一位非营利性教育集团的总裁。在我离开他的办公室时，他随口提了一句：“如果你决定留在美国，你应该过来加入我们的咨询委员会。”

这句题外话令人好奇不已。他不是问一个具体的问题，而是作了一种假设，那就是我有一种选择：“如果你决定留在……”他将之视为一个真正的选择。我因此陷入了思考。

我离开他的办公室，乘电梯下到大厅。我从别人桌上随手拿了一张纸，在大厅里找了个地方坐下，试图回答这个问题：“如果一生只能做一件事情，你会做什么？”

结果就是我之前提到的那样，我没有在那张纸上写下法学院。

在那一刻之前，我虽然在逻辑上一直都知道自己可以不学法律，但是在感情上，不学法律从来都没有成为我的选项。那一刻我意识到我浪费了自己的选择权。没有考虑法学院以外的其他学校就直接选择了法学院——不是因为我真的想去那儿上学，而是因为没有进行选择。我觉得，在那个时刻我第一次认识到了，一旦我们放弃自己的选择能力，别的力量或者别的人就会插手替我们作出选择。

几个星期以后，我正式从法学院退学，离开英格兰前往美国开始了作家和教师的职业生涯。大家能够读到这本书恰恰是因为那个选择。

然而，这个特殊的选择在我人生轨迹上所产生的全部影响中，我尤为看重的是它进一步改变了我对选择的看法。人们常把选择视为一种事物，但它不是事物。我们面临的选项也许可以算是事物，但选择是个行为动作。它不仅是我们拥有的东西，而且是我们做出的行动。这种经历令我茅塞顿开，我们也许无法完全左右自己会有哪些选项，但是可以永远掌控自己如何从中选择。

你是否曾因为觉得自己别无选择而纠结困惑？你是否因为陷于“这我不能做”和“这我必须做”的矛盾思维的夹缝中感到压力重重？你是否一点点放弃了自己选择的力量，直到盲目听命于他人为你设定路径？

如果你果真如此，那么你就跟很多人同病相怜了。

习得性无助与选择能力缺失

很久以来，人们过分强调了选择（我们面临的选项）的外部条件，而低估了选择（我们的行动）的内在能力。这不仅仅是一个语义学上的问题。试想一下，选项（事物）会被拿走，而进行选择的核心能力（自由意志）却不会。

我们是如何以及为何忘记了选择的能力，关于这个问题有一种重要的见解来自马丁·塞利格曼（Martin Seligman）²和史蒂夫·梅尔（Steve Maier）的经典作品。他们在对德国牧羊犬进行实验时偶然发现了一种现象，后来将其称为“习得性无助”。

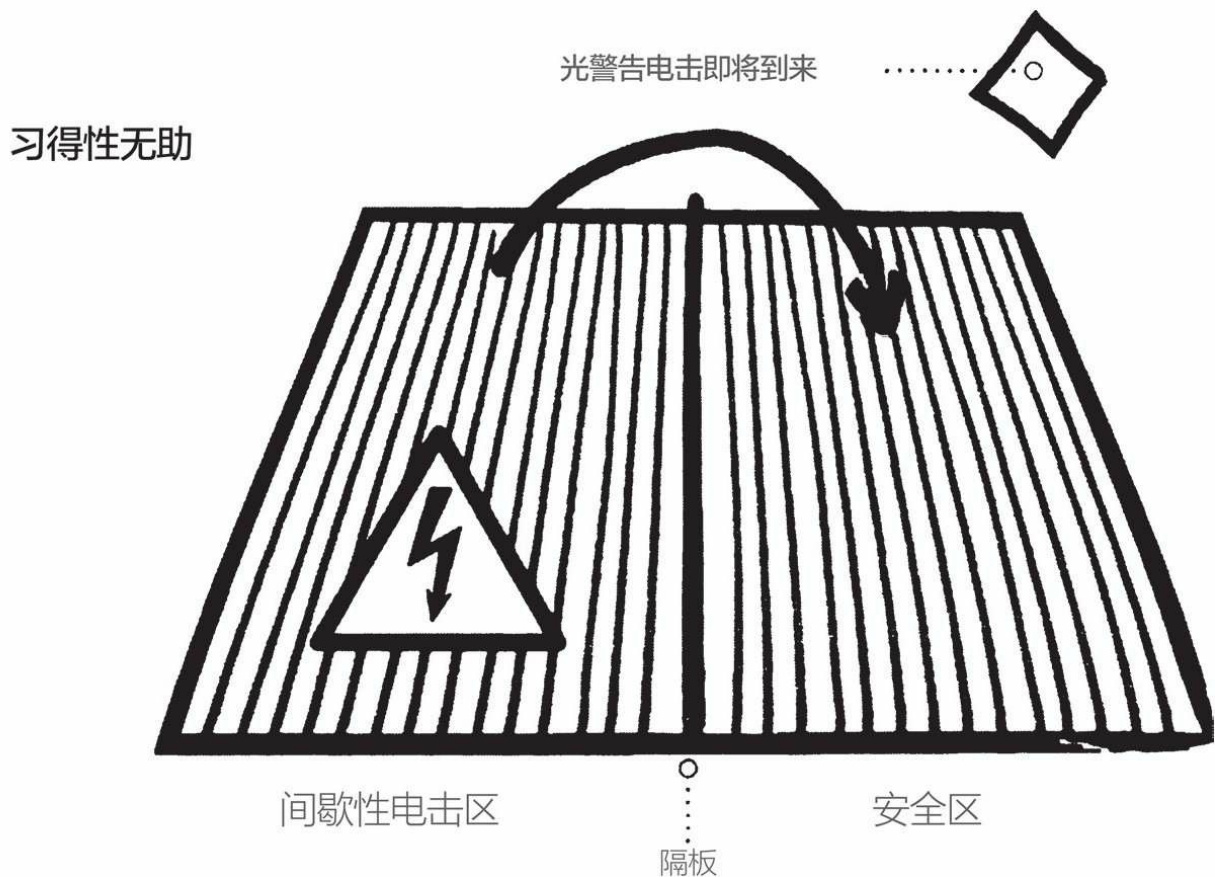
塞利格曼和梅尔将一群狗分成三组。第一组狗被拴上了系带，并被施以电击，但它们可以压下操作杆来停止电击。第二组狗被拴上了同样的系带，分到了同样的操作杆，面临同样的电击，不过有一个陷阱：操作杆坏了，致使狗无法对付电击。第三组狗仅仅被拴上了系带，不施加电击。

**The ability to
choose cannot
be taken away
or even given
away—it can
only be forgotten.**

选择能力不会被夺走，
不会被丢弃，
只会被遗忘。

然后，所有狗都被放到了一个笼子里，笼子中间放置有一块隔板，

隔板的高度是狗可以轻易跳过去的。隔板的一边有电击，另一边没有电击。有趣的现象出现了。那些在前期的实验中能够停止电击或者没有遭到电击的狗很快学会了跨过隔板，到达没有电击的那一边。但是无力应对电击的那些狗没有这么做，它们没有做出反应或调整，没有采取任何措施来避免电击。这是为什么呢？因为除了遭受电击，它们不知道还有其他的选择。它们有了习得性无助。



这一实验为人类以几乎相同的方式形成习得性无助提供了证据。我听说，有一个孩子一开始很勤奋地学习数学，他努力了一次又一次，但是成绩却总不见好，因此他最后放弃了。他认为自己所做的一切努力都是无用的。

我也发现，在我工作过的许多组织里也存在这样的习得性无助。当

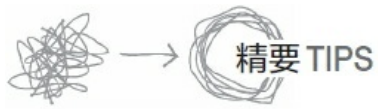
人们感到自己的努力无关紧要时，往往会采取两种方法中的一种来应对。有时他们会抽身退出不再尝试，就像那个数学怎么也学不好的孩子那样。另一种方法一开始并不引人注目。他们不是抽身退出，而是表现得极度活跃。他们会抓住每一个机会，积极参与每一项工作，热情万丈地解决每一个难题，竭尽全力做好一切。乍一看，这种行为未必像习得性无助。努力难道不是让人们相信自己重要和有价值的明证吗？但是进一步审视就会发现，这种强迫自己去做更多的方式不过是一种障眼法罢了。这样的人不相信自己可以选择机遇、任务和挑战，而是认为自己“必须什么都做”。

我首先要承认选择是困难的。根据定义，选择意味着拒绝一件或多件事情，这会带给人一种失落感。工作之外，选择会变得更加困难。每次我们去商店、饭店或者任何一个卖东西的地方，那里的一切都设计得精致美观，让人欲罢不能。不管是政治广告还是权威演说，其目的都是让听众无论如何都不要去给另一方投票。每当岳母（我岳母除外）来电让我们干这干那时，我们无论如何都不会觉得自己还有别的选择。如果以这样的眼光审视我们每天的生活，就会发现将自己的选择能力抛之脑后一点儿都不意外。

然而，选择却是成为精要主义者的核心所在。精要主义者必须高度重视选择能力，将之视为一种战无不胜的力量，它的存在独立于其他任何事物、任何人及任何力量。美国心理学之父威廉·詹姆斯（William James）曾这样写道：“我的第一个自由意志行为就是相信自由意志。”这就是为何在精要主义者的养成之旅中，你必须掌握的第一种并且是最重要的一种能力就是做出选择的能力，它涉及生活的方方面面。

忽略了选择的能力，就等于选择了无助。眼看着自己的能力被一点一滴蚕食，最后只得服务于他人的选择，或者服务于自己过去的选择。结果，我们将自己的选择能力拱手相让。这就是非精要主义者所为。

精要主义者不仅重视选择能力，而且还庆幸它的存在。他们深知，一旦将自己的选择权拱手相让，那么就不仅给予了他人这种力量，而且还明确地授权他人来替自己作出选择。



- 不加思考的努力等同于习得性无助。
- 无视拒绝带来的“失落感”。
- 高度重视选择能力，它的存在独立于其他任何事物、任何人及任何力量。

CHAPTER 02

DISCERN

The Unimportance of Practically Everything

模式二

甄别，几乎一切都是噪声

宇宙中存在的绝大多数事物——人类的行为和其他一切力量、资源、观念，都没有什么价值，也产生不了什么结果。只有少数事物效果非凡，影响巨大。

——理查德·科克



非精要主义者

认为几乎一切都重要

将所有机会等同视之



精要主义者

认为几乎一切都不重要

区分重要的少数和不重要的多数

在 乔治·奥威尔（George Orwell）的经典寓言小说《动物庄园》（*Animal Farm*）里，我们认识了一个虚构的主人公——一匹名叫“拳击手”（Boxer）的马，他被刻画得既忠诚又强壮。每当遭遇挫折和碰到问题时，他总说：“我会更努力。”哪怕是在最痛苦的境遇里，他都遵循着这样的信条，直到他耗尽了力气，伤病交加，被送到屠马场。他是一个悲剧性的角色：不管用意多么美好，他不断的努力实际上却加剧了庄园里的不平等，并导致了更多的问题。

我们是不是在某些情况下与“拳击手”有点相似？面对挫折，我们是不是常常只会让自己变得更加坚定，更努力而持久地工作？每当面临挑战，我们是否常作出“是的，我也能行”这样的回答？毕竟，从小受到的教育告诉我们，努力是结出硕果的关键，而且很多人为了外界强加给他们的任务和挑战拼尽全力，他们的业绩和能力确实给他们带来了丰厚的回报。然而，对于那些已经在努力工作的能人来说，努力工作的价值是否会有极限？是否存在一个临界点，在这个点上做得更多并不会带来更多？有没有一个点，在这个点上做得更少（却想得更多）反而会产生更好的效果？

做得更少，效果更好

记得小时候，我想自己挣点儿零花钱。那时在英国，当报童是一个12岁的孩子能找到的仅有的几份工作之一，一天的工钱是一英镑，工作时间大概是一小时。因此，在一段时间里，我每天早上上学前，都会拉着一个看上去比我还重的大袋子，挨家挨户地送一小时报纸（顺便说明一下，我们不能把报纸扔到人家门廊上就走，这和美国不一样。我们必须要把报纸拿起来用力整个儿塞进人家家门上的小信箱里）。这点零花钱挣得着实不易。

为了每天挣这一英镑，我历经千辛万苦，这永久而彻底地改变了我

对心仪事物所需成本的看法。从那时起，每当我想要买什么东西的时候，我都会把它换算成我需要送多少天报纸来衡量。一英镑的报酬等于一小时的奔波忙碌。我发现，按照这种比率计算，买一辆自己喜欢的玩具小汽车需要攒好一阵子的钱。

然后，我就开始琢磨如何才能加快攒钱的速度，我想到了自己周六上午可以不去送报纸而去给邻居家洗车。每辆车两英镑，我可以在一小时之内洗三辆车。突然，工作时间和英镑的比率由1：1变成了1：6。这给我上了重要的一课：某种工作会比其他工作带来更多的报酬。

几年后，我读大学，期间去一家培训公司工作。我在那儿的客户服务部工作，薪酬是每小时9美元。用时间和报酬之间的比率来衡量一份工作本非难事，但是我知道真正重要的是时间和结果之间的关系。

因此，我问自己：“在这份工作中能够获得的最有价值的东西是什么？”结果发现最有价值的莫过于留住原本要流失的客户。因此，我努力说服客户不要离开我们公司，很快便创造了客户流失率为零的成绩。每留住一个客户，我都能获得奖励，由此我学到了更多，挣到了更多，也贡献了更多。

努力工作当然重要，但并不是说更努力就一定可以带来更多收获。“更少，但更好”可以。

费兰·阿德里亚（Ferran Adrià）可以说是世界上最伟大的厨师，他主持的西班牙斗牛犬餐厅是全球最著名的餐厅。他起码在两方面集中体现了“更少，但更好”的原则。

- 第一，他擅长将传统食材打散成最基本的分子状态，然后再以人们从未想象过的方式创新重组。

- 第二，尽管斗牛犬餐厅每年晚餐预订量约为200万份，但每晚只为50名客人服务，每年只营业6个月。

事实上，在本书写作之际，费兰已经完全停止了供餐，而是将斗牛犬餐厅转变为专业的各类食品实验室，在那里，他唯一继续的是追求更高超的厨艺。

要习惯“更少，但更好”的想法也许不如听上去那么容易，尤其是当我们过去因付出了很多很多而获得了回报时。但是在一定程度上，更多的努力会导致发展进入不再变化或增长的稳定期，甚至停滞不前。诚然，结果和努力之间存在直接关联的理念是有吸引力的，它看上去很公平。然而，很多领域的研究结果却显示出了一幅迥然不同的图景。

无意义的多数与有意义的少数

绝大多数人听说过“帕累托原理”，该原理早在19世纪90年代由维尔弗雷多·帕累托（Vilfredo Pareto）提出，它指出80%的成果是由20%的努力产生的。很久以后，现代质量管理之父约瑟夫·摩西·朱兰（Joseph Moses Juran）于1951年在他的《质量控制手册》（*Quality Control Handbook*）中发展了此理论，并称之为“重要少数法则”（**the Law of the Vital Few**），即人们可以通过解决问题当中一个微小的部分来极大地提高产品的质量。他在日本找到了一个愿意进行试验的企业，当时该企业的品牌声誉很差，以生产成本低、产品质量差而出名。通过集中大量精力和注意力仅对少量真正重要的环节进行改进，他让“日本制造”这个词拥有了一个全新的含义。质量革命逐渐使日本崛起成为全球一大重要经济体。

对“不重要的多数”和“重要的少数”的区分可被用于人类各种大大小

小的努力中，理查德·科克（Richard Koch）对此进行了令人信服的实践。他写了好几本关于如何应用帕累托原理（80/20法则）的著作。的确，类似的例子比比皆是。

沃伦·巴菲特有句名言：“我们的投资策略接近于昏睡状态。”他的意思是说，他和他的公司只做相对较少的投资，并且会长时间地保持。在《巴菲特之道》（*The Tao of Warren Buffett*）一书中，玛丽·巴菲特（Mary Buffett）和戴维·克拉克（David Clark）解释道：“在从业早期，沃伦就断定自己不可能做出数百项正确的投资决策，所以他决定只投资自己有绝对把握的行业，然后在那上面押下重注。他90%的财富仅仅来自10项投资。有时候，你不做什么恰恰和你做什么一样重要。”简而言之，他在少数至关重要的投资机会上押重注，放弃那些仅仅只能算作好的机会。

有些人认为努力和收获之间的关系更加不是线性的，他们信奉科学家们所称的“幂次定律”。根据幂次定律，某种努力与其他努力相比，确实会使成果呈指数级增长。比如，微软公司前首席技术官纳森·米尔沃德（Nathan Myhrvold）曾表示（当面向我确认过）：“最优秀的软件工程师的开发效率比一般的软件工程师不是强**10**倍或**100**倍，甚至不是**1000**倍，而是**10000**倍。”这或许有夸张的成分，不过仍然说明了某种努力相比其他努力确实会大大改善效果。

一个不争的事实是：我们生活的世界中几乎一切都是无价值的，仅有极少的事物例外。正如约翰·麦克斯韦尔所写：“几乎一切事情的次要性再怎么高估都不过分。”

抛开1: 1的逻辑，我们就开始认识到了追求精要的价值。我们发现，很多我们追求的好机会远不如极少数真正的契机有价值。一旦认识到这一点，我们就开始扫描环境，追寻那些重要的少数，并且急切地淘汰那些不重要的多数。只有这样，我们才能拒绝好机会，拥抱真正的契

机。

这就解释了为何精要主义者要花费时间来探索自己的所有选项。这种额外的时间投资是值得的，因为有些东西实在是至关重要，它们会10倍地回报为找到它们所付出的努力。换句话说，精要主义者们用更多的时间来辨别，以便花更少的时间来做，正所谓磨刀不误砍柴工。

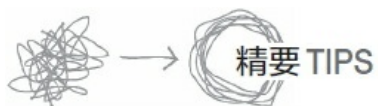
很多能干的人无法提升贡献度，是因为他们无法摆脱一切都重要的理念。但是精要主义者学会了辨别真正重要之事和其他事物之间的差别。为了将精要主义者的技能付诸实践，我们可以从一个简单的级别开始，一旦它成为左右每日决策的第二天性，我们就可以将它应用到更大、更广泛的个人及职业生活领域中去。要完全掌握这种技能需要思维方式的巨大转变。但这是可以做到的。

**A Non-
essentialist
thinks almost
everything is
essential.**

非精要主义者认为几乎一切都重要。

**An
Essentialist
thinks almost
everything is
nonessential.**

精要主义者认为几乎一切都不重要。



- 更多努力不一定能带来更多收获，“更少，但更好”可以。
- 几乎一切事情的次要性再怎么高估都不够。
- 扫描环境，寻找重要的少数，淘汰不重要的多数。
- 拒绝好机会，拥抱真正的契机。

CHAPTER 03

TRADE-OFF

Which Problem Do I Want?

模式三

取舍，我要的究竟是什么

所谓战略，就是如何作出选择及取舍。它意味着通过慎重选择来实现不同。

——迈克尔·波特



非精要主义者

想：“我两样都能做。”

问：“怎样才能什么都做？”



精要主义者

问：“我想作出什么样的取舍？”

问：“我要大干一场的是什
么？”

想象一下，你可以回到1972年，对每个标准普尔500指数的公司进行一美元的投资。到了2002年，哪家公司会给你带来最大的投资收益呢？是通用汽车？IBM公司？英特尔公司？根据《货币》杂志以及内德·戴维斯研究集团（Ned Davis Research）所做的分析，上述答案均不对。

正确答案是西南航空（Southwest Airlines）。这一答案实在令人大跌眼镜，因为航空业的利润是出了名的差。但是，赫布·凯莱赫（Herb Kelleher）领导的西南航空却年复一年地创造了靓丽的财务业绩。赫布的精要思维就是其商业成功的关键所在。

我曾参加过一次关于赫布商业战略的访谈活动。这项活动在很多方面亮点频出，但是当赫布开始谈及他在西南航空是多么谨慎地进行权衡取舍时，我不禁竖起了耳朵。他们公司不是所有的航线都飞，而是只提供点对点航线；不是抬高价格来覆盖餐食成本，而是决定不提供餐食；不提前分配座位，而是让乘客登机后挑选座位；不向客户推销炫目的头等舱服务，而是只提供普通舱服务。这些权衡取舍无一来自默认设置，而是通过精心设计做出的。每一项都是深思熟虑的商业战略中的一部分，用以降低成本。但是有些客户想要去往更多的目的地，想要购买昂贵的飞机餐等，公司这样做难道就不怕失掉这些客户吗？是的，但是赫布完全了解自己公司是一家什么样的公司——一家低成本的航空公司，也知道自己公司不是一家什么样的公司。这些都体现在他所做的权衡取舍中。

还有一个例子，也体现了赫布在工作中对精要思维的运用。他说：“你必须审视每一个机会，然后说，‘不，我很抱歉。我们不想做那么多件不同的事情，如果它们对我们努力要实现的那个最终结果没有太大贡献的话’。”

开始时，很多评论家、反对者和其他认为这种方法不可能成功的非精要主义者对西南航空进行了猛烈抨击。谁会在心智正常的情况下选择一家只飞往部分目的地且不提供餐食的航空公司，不管机票价格有多么便宜？然而，若干年后，西南航空的成绩就开始显山露水了。同业竞争者注意到了西南航空利润飙升并开始仿效他的成功做法。但是，他们并未全盘采用赫布的精要方法，而是采用了哈佛商学院教授迈克尔·波特（Michel Porter）所称的“横跨战略”。

用最简单的词来解释，横跨就是指继续原封不动地保留自己的现有策略，与此同时也试图采用竞争对手的策略。当时，这方面最明显的例子莫过于美国大陆航空公司（Continental Airlines）的做法。他们将自己新的点对点服务称为“大陆航空的小大陆”。

大陆航空采用了西南航空的一些做法，比如降低机票价格、取消飞机餐供应、停止头等舱服务、增加航班频率。但问题是，由于仍然保留着现有的商业模式（大陆航空的“小大陆”服务模式仅针对该公司的一小部分航班），所以运行效率并不高，不足以支撑他们在价格竞争上获得优势。因此，他们不得不在其他方面降低成本，最终导致了服务质量的下降。西南航空在关键性战略领域做出了理智、慎重的权衡取舍，而大陆航空只是迫于无奈才牺牲了一些收益，它们并不是公司连贯性战略的组成部分。根据波特的观点，“战略定位若要持续，必须同其他定位之间进行取舍”。企图同时运用两种水火不容的战略，最终削弱了大陆航空的竞争能力。

这种横跨战略让大陆航空付出了昂贵代价。由于航班延误，他们损失了数亿美元，而且按照波特的说法，“航班晚点和取消引发的投诉抱怨高达每日一千次”。公司总裁最后被炒了鱿鱼。无视取舍问题对于公司来说无异于做出了一种可怕的战略，对于个人来说也是如此。

有所取，必要有所舍

你有没有和这种人一起待过，他们总是试图要赶着再多做一件事情？这样的人明知道离会议开始只剩10分钟了，而去会场就需要10分钟，却还要坐下来回复几封邮件才出发；或者他们会答应在周五前准备好一份报告，即便在同一天还有一项重要工作必须按期完工；或者他们可能会答应周六晚上参加表亲的生日派对，即便在同一时间还有演出要去看。这种人的逻辑就是我可以两者兼顾，因此根本不考虑取舍。而更重要的问题是，这种逻辑是错误的。最终，他们要么开会迟到；要么在最后期限那天一项工作或者两项都无法交工（或者两项工作都完成得很不好）；要么参加不了表亲的生日派对，要么没去看演出。在现实中，接受任何一个机会都无疑需要你舍弃其他几个机会。

取舍是真实存在的，无论是在个人生活还是在职业生涯中都是如此。如果不接受这样的现实，我们就注定会像大陆航空公司一样，受困于“横跨战略”中，被迫毫无选择地牺牲一些收益，而这可能不是我们主动设计的。

雷曼兄弟前首席财务官艾林·卡兰（Erin Callan）在写给《纽约时报》的一篇很有深度的专栏文章中，讲述了她为不加选择地进行取舍而做出的牺牲。她这样写道：“我一开始时并不打算把自己的全部精力都投入到工作中去。但是，慢慢地就变成这样了。每过一年，都会有些略微的调整，这成了新的常态。开始时，我在周日会花半小时来整理电子邮件、工作事项清单和日程表，这样周一早上就会变得轻松一些。后来，我每周日要工作几个小时，再后来就变成了整天。工作与生活之间的界限渐渐消失了，最后只剩下了工作。”她的故事揭示了一个至关重要的真理：我们要么自己做出艰难的选择，要么允许其他人——不论是我们的同事、老板还是客户，替我们作决定。

我在工作中注意到公司高管是对取舍现实性接受程度最差的人。我

最近接触了硅谷一家市值400亿美元的公司的首席执行官。他跟我分享了他刚刚编写的，准备向全公司宣布的公司价值宣言：我们珍视激情、创新、执行和领导力。

这份宣言的问题之一就是，有谁会不珍视这些？另一个问题是，它并未告诉员工公司最珍视的是什么。它没有向员工指明，在这些价值观发生冲突的时候，他们应该选择什么。一些公司宣称自己的使命是平等地服务所有利益相关者——客户、员工、股东，这其中也存在类似问题。宣称对所有相关人士都同等重视就会给管理留下难题，因为没有明确指出若要对自身所服务的人群进行权衡取舍，我们该如何选择。

与之形成对照的是，强生公司成功地从1982年的一场灾难性氰化物谋杀丑闻中东山再起的故事。当时，强生公司占有37%的市场份额，泰勒诺是其最盈利的产品。那时，有报道称有7人服用泰勒诺后死亡。经调查发现，是装药的瓶子被动了手脚。强生公司该如何作出回应呢？

这个问题极其复杂。他们的首要工作是召回药店货架上的全部泰勒诺产品以保障客户安全呢？还是通过危机公关来控制损害并防止股东抛售公司股票？又或者是第一时间安慰并补偿受害人家庭？

幸运的是，他们有自己的信条：当时的公司董事长罗伯特·伍德·约翰逊（Robert Wood Johnson）曾于1943年写下一则宣言，被全文刻在了强生公司总部的一块石头上。与大多数公司使命宣言不同，这则信条对公司的组成人员进行了主次排序：客户第一，股东最后。

结果，强生公司迅速决定召回全部泰勒诺产品，即使这将会对公司造成巨大损失（据报道，涉及金额高达1亿美元）。是选择客户的安全还是1亿美元？这不是一个可以轻易做出的决定。但是，公司的信条让他们对“什么是最重要的？”这个问题没有任何疑问，并促成了一次艰难的取舍。

我曾与一个管理团队合作过，他们在安排事情的优先次序方面需要一些指导。当时，他们正纠结于确定5个最重要的项目，让他们的科技部门在下一个财年中完成，其中一名经理人对此尤其苦恼。她坚持要确定18个“最优先”项目，我坚持让她只选5个。她带着清单回到了自己的团队，两周后回来时带来了她努力精简后的版本——只减少了1个项目！（我一直很想知道那唯一一个被精简掉的项目为什么没有达到标准。）由于拒绝作出取舍，她最后把应该用于5个项目上的时间和精力摊到了17个项目上。不出所料，她没有取得理想的结果。她的逻辑是：我们可以搞定一切。很显然，事实不是这样的。

**We can try to
avoid the reality
of trade-offs,
but we can't
escape them.**

我们试图躲避
取舍现实性，
然而，
我们无处可逃。

要么创立伟大公司，要么创造伟大思想

要理解人们为何如此热衷于拒绝取舍，其实并不难。毕竟，从定义上就可以知道取舍会涉及我们想要的两件事物。你是想要更多收入还是更多假期？你想回完下一封邮件还是准时去参加会议？你看重工作速度还是质量？很显然，当我们必须在两个自己想要的事物中进行选择时，就会倾向于二者兼顾。然而，不管主观上多么想鱼和熊掌兼得，但实际上就是做不到。

每当面对取舍时，非精要主义者总会问：“我怎样才能二者兼顾？”而精要主义者却会问一个更尖锐但更能使人脱离困境的问题：“我要的是什么？”精要主义者会主动作出权衡取舍，成为自己行动的主宰，而不是坐等别人的驱使。正如经济学家托马斯·索厄尔（Thomas Sowell）所言：“所谓的解决方法是不存在的，有的只是取舍。”

商业巨著《从优秀到卓越》（*Good to Great*）一书的作者吉姆·柯林斯曾得到过彼得·德鲁克的忠告：你要么创立一家伟大的公司，要么创造伟大的思想，但是鱼与熊掌不能兼得。吉姆选择了创造思想。正是缘于这种取舍，尽管现在他的公司里只有3个全职雇员，但他的思想却已经随着他的著作走进了千万读者的心里。

尽管有时取舍会给人带来痛苦，但却代表着一个意义重大的机会。通过强迫自己权衡两者并战略性地选择那个对自己最有利的，就能极大地增加实现自己目标的机会。就像西南航空，因为选择始终如一，便可享受到成功的硕果。

最近在一次飞往波士顿的航班上，我就注意到了这样一个例子。当时我正与一对父母聊天，他们要去看望在哈佛读书的儿子。看得出来，他们为有一个能在哈佛读书的儿子感到骄傲。我很好奇这一家人采取了

什么样的策略，孩子才有机会进入哈佛读书。他们说：“我们也曾让他尝试做过很多不同的事情，但是一旦看清这件事不可能成为他的‘大事’，我们会商量不再让他继续做。”问题的关键不在于所有的父母都应该怀揣着让孩子进哈佛的梦想，而是在于那些运用精要思维的父母，能够自觉清醒地决定自己的目标就是让孩子进哈佛，并且懂得要取得成功必须做出战略性的取舍。

这种思维逻辑在个人生活中也是适用的。刚结婚的时候，我和安娜遇到了一个人，就我们所知，他婚姻美满，家庭幸福。我们就向他取经。他跟我们讲了很多，其中之一就是他们夫妇都决定不参加任何俱乐部。他没有参加当地的会所活动，他妻子没有加入读书俱乐部。他们这样做不是因为对这些事情没有兴趣，而是决定要把时间省下来陪在孩子们身边。多年以来，孩子成了他们最要好的朋友，为此他们也许错过了在高尔夫球场上或者在翻看破旧的《安娜·卡列尼娜》时结识朋友的机会，但都是值得的。

精要主义者将取舍视为生命中内在固有的组成部分，而不是先天消极的部分。他们不问：“我应该放弃什么？”而是问：“我要大干一场的是什么？”类似这样小小的改变，日积月累将会产生意义深远的影响。

在一篇发表于《纽约客》（*The New Yorker*）上的名为《笑吧，笑翠鸟》（*Laugh, Kookaburra*）的文章里，大卫·塞德瑞斯（David Sedaris）用幽默的笔法描述了他澳大利亚灌木丛的游览经历。在徒步旅行时，他的朋友兼当日导游跟他讲了一件她在一次管理课上不经意间听到的事情。“想象一种有四个燃烧器的炉子，一个燃烧器代表你的家庭，一个是你的朋友，一个是你的健康，一个是你的工作。为了成功你必须去掉一个燃烧器，为了真正成功你必须去掉两个。”

当然，这只是个玩笑罢了，我并不是建议大家为了像精要主义者那样生活，就必须在自己的家庭、健康和工作之间进行抉择。我要建议的

是，当有一个选项优先于家庭，另一个优先于朋友、健康或工作，而又必须从中进行选择时，我们就需要问一问自己：“我到底要的是什么？”

取舍既不能被忽视，也不能被诋毁。我们应该主动拥抱它，自觉地作出具有战略性和经过深思熟虑的取舍。



- 审视每一个机会，如果对努力要实现的最终结果没有太大贡献，就不做那么多不同的事情。
- 在关键性战略领域，做出理智、慎重的权衡取舍。
- 强迫自己权衡两者，并战略性地选择对自己最有利的。

Part II

精要主义者养成的第一步：探索





explore

How can we discern
the trivial many from
the vital few?

导读

区分无意义的多数和有意义的少数

精要主义的一大悖论就是相比非精要主义者，精要主义者实际上会探索更多的选择项。几乎所有的事情都能引起非精要主义者的兴趣，也会引起他们的反应。但是，正因为如此忙于追逐每一个机会和每一种想法，他们实际上在选项探索上花费的精力少之又少。相反，精要主义者的做法是在专注于某项事情之前，先广泛地探索和评估各种选项。因为精要主义者只在少数重要的想法和事情上集中精力“大干”一场，他们最开始时探索更多的选项，就能保证以后作出的选择是正确的。

在这一部分，我们将讨论探索精要之事的5个要素。非精要主义的巨大牵引力或许会让人们忍不住跳过这个要素或者仅仅是匆匆一瞥。但是，这个要素本身对于人们实现自律地追求更少具有极其重要的意义。要识别出什么是真正重要的，我们需要思考的空间，需要观察和倾听的时间，需要有享受娱乐的心态，需要聪明地让自己该睡就睡，需要在作出选择时自觉自律地运用严格的选择标准。

具有讽刺意味的是，在非精要主义者看来，空间、倾听、娱乐、睡眠和选择等都是毫无意义、会导致人们分心的事物。它们最多不过是被视为可有可无，有也不错，最次还可能被诋毁成软弱和浪费的明证。众所周知，野心勃勃或者工作高效的人是这样想的：“当然，我很想从工作日程表中留一些时间出来，专门用来思考，但这是一种奢侈品，我们现在用不起。”或者是：“玩？谁有时间玩？我们是来这里工作的！”或者，就像在一次员工入职过程中，一位领导跟我说的那样：“但愿你晚上能睡个好觉。在这里，这并不容易实现。”

如果你相信，过分忙碌和把精力用到极致是工作高效的证据，那么你也可能会认为探索、思索和反省所占用的时间应该被降到最低。然而，正是这些“可有可无”的活动为非精要主义者焦头烂额的忙碌提供了一种解毒剂。这些活动非但不是毫无意义、令人分心的，反而还是将真正的分心物和真正的要事区分开来的关键环节。

精要主义者将花尽可能多的时间用于探索、倾听、辩论、质问和思考。但是，探索本身并不是他们的目的，真正的目的在于将有意义的少数从无意义的多数中甄别出来。

CHAPTER 04

ESCAPE

The Perks of Being Unavailable

要素一

抽离，为探索和思考留出空间

没有伟大的孤独，一切严肃的事情都无法做成。

——巴勃罗·毕加索



非精要主义者

忙得无暇顾及人生



精要主义者

抽身而退，探索人生

弗兰克·奥布赖恩（Frank O'Brien）是Conversations公司的创始人，该公司是一家总部位于纽约的营销服务公司，已入选“美国成长最快的民营企业”500/5000榜单。面对现在疯狂的职场节奏，他采取了一种激进的做法。

每个月，他都会把自己这家50人的公司的每一名员工都集中到一个房间里，一待就是一整天。禁止使用手机，不准收发电子邮件，没有任何议项目。会议的目的仅仅是抽离出来，专心思考和讨论。请注意，每月中间的那个周五一般是工作效率低下的时候，人们在这一天不会在“真正的工作”上面有什么进展。但他不是在这一天召开这种会议的，而是在每月的第一个周一开这个会，一开就是一天。这种做法也不仅仅是一种内部纪律，因为连他的客户都知道，在这个“别打电话的周一”不要指望得到这家公司的任何答复。

他这样做，是因为觉得自己的员工如果随时待命的话，就会迷惘于何事才重要。他们需要空间来思考真正重要之事。他这样写道：“我认为，留出点时间来喘一口气，四下望一望，好好想一想，是必不可少的。你需要这样的安静清明来创新和成长。”此外，对他来说，这种会议类似于一种石蕊测试，提醒他是否有员工在非精要事务上花了太多时间：“如果有人因为太忙抽不出时间来参加这个会议，就说明我们要么工作低效，要么人手不够，需要补充。”如果员工忙得连思考的时间都没有，那么他们真的是太忙了，起码这段时间是这样的。³

我们需要逃离的空间，以便从不重要的多数中分辨出重要的少数。不幸的是，在这个时间宝贵的时代，默认情况下我们没有那样的空间，除非主动安排。跟我合作过的一个领导人承认说，在一个公司里待五年太长了。为什么？因为在公司上班时他忙得连考虑该不该留在那里的时间都没有。每日的工作任务和要求让他无暇抽身退一步看清局势。

无独有偶，全球一家重要科技企业的高级副总裁告诉我，他每周花在开会上的时间多达35个小时。这些会议消耗了他太多的时间，使他每月甚至连一个小时都抽不出来规划一下自己的事业，更不用说带领他的团队再上一个新台阶了。他没有给自己留出空间来讨论一下到底发生了什么事和究竟该干些什么，而是把时间消耗在了没完没了的业务报告和沉闷乏味的跨职能会议中，这样的活动往往什么问题都解决不了。

在评估事情的重要性之前，你需要探索一下自己的选项。非精要主义者往往对最新的想法做出反应，飞快地抓住刚刚出现的机会或是回复最新接收的一封邮件。精要主义者则会选择给自己创造探索和思考的空间。

留出空间，设计人生

我在斯坦福大学哈索普莱特纳设计学院（The Hasso Plattner Institute of Design at Stanford）工作期间，深刻领悟到了留出空间用以探索的价值。有一次，我走进一间教室，在那里我要讲授一门课程。我最先注意到的是那间教室里没有传统意义上的那种椅子，只有一些泡沫立方体充当座位，我很快就发现那些所谓的座椅坐上去很不舒服。就像设计学院里的一切东西一样，这也是一种设计。它的设计初衷是：因为这些立方体坐起来很不舒服，学生们坐几分钟后就宁愿站起来，四处走走，相互认识，而不仅仅只认识坐在自己两侧的同学。这就是设计的关键理念所在。学院就是利用物理空间来鼓励社交和思考方法创新的。

为了达到这样的目的，学院还建立了一个名为“黑色隔间”的隐藏空间。这是一间小屋子，故意设计得只能容下1~3人藏身。屋子没有窗户，完全隔音，而且被有意地去除了会分散注意力的因素。史考特·多利（Scott Doorley）与史考特·威硕夫（Scott Witthoft）合著的《制造空间》（*Make Space*）一书指出，这个“黑色隔间”“连低技术含量都不具

备，简直就是无技术含量”。它在一楼隐蔽地存在着。就像多利和威硕夫指出的那样，它不通往任何地方。思考，是人们去那里的唯一目的。有了这样一个思考和聚焦的空间，学生们得以后退一步，看得更清。

由于某种原因，人们对“焦点”一词有着错误的理解。就好像对“选择”一词的理解一样，人们习惯于将焦点视为一种事物。是的，焦点的确是我们的拥有物，但它（聚焦）同时也是我们的行为。

我说的焦点，并不是简单地意味着我们要挑出一个问题或一种可能性，然后很痴迷地思考它，而是要创造一个空间，来探索一百个问题和一百种可能性。精要主义者关注眼睛聚焦的方式，不是死死盯住某事，而是根据视野的变化进行不断地调整 and 适应。

In order to have focus we need to escape to focus.

为了找到焦点，
我们需要抽身而退，
全力聚焦。

在设计学院最近的一次会议上（在另一间没有座椅和桌子的房间

里，从地面到天花板都是白板，上面贴满了各种透明便利贴），我见到了杰里米·厄特利（Jeremy Utley）。他同我一起合作开发过一门新课程的雏形，有一次他突然灵光一现地称这门课为“设计你的精要人生”。

这门课的唯一目的是给学生们打造一个空间，用以设计他们的人生。每星期，它都会给他们的思考提供一个计划好的理由。这使他们被迫关掉笔记本电脑和智能手机，让自己的头脑开足马力地思考。练习经过深思熟虑后将最精要的少数从很多好东西中识别区分出来，这是他们的作业。其实，要培养这样的习惯未必一定要在设计学院，我们可以在自己的生活中创造更多的机会。

留出空间，专注探究

我认识一位经理人，他聪明机敏而又干劲十足，但美中不足的是，他经常分散注意力。在任何时候，他都离不开Twitter、Gmail、Facebook和其他各种即时聊天工具。为了创造一个没有任何分心物的空间，他曾一度让他的行政助理拔掉了他电脑上的网线。但即便如此，他依然可以找到很多其他的上网办法。因此，每当他要竭力完成一个特别重要的大项目时就不得不采取一个极端的方式——丢下手机，去一家没有上网条件的汽车旅馆办公。过了几乎与世隔绝的8个星期后，他才得以完成项目。

这个经理人居然被逼得要采用这样的方法，我对此感到有些遗憾。然而，他的手段固然有点极端，但目的却是毋庸置疑的。他深知，唯有思路不受打搅和干扰，才能达到个人贡献峰值，因此他努力创造这样的空间。

艾萨克·牛顿爵士的例子也许会给我们带来些启发。他在《自然哲学的数学原理》（*Philosophiae Naturalis Principia Mathematica*）一书上花了整整两年时间，阐述了他最著名的万有引力定律和力学三大定律。

这些定律是真正意义上的伟大突破，奠定了接下来3个世纪的科学思维，而这段几乎与世隔绝的研究岁月所具有的重要性也是不容置疑的。

理查德·韦斯特福尔（Richard S. Westfall）这样写道：“在牛顿名声大振的时候，有人问他是如何发现万有引力定律的。‘答案就是持续地思考……’他一旦要思考什么，就持续地思考，也就是排除一切杂念或是几乎排除一切杂念地思考。”换句话说，牛顿创造了进行高强度专注思考的空间，而这种不受打搅的思考空间使得他对宇宙间最基本构成要素的探索成为了可能。

深受牛顿的启发，我采取了一种相似的，但也许不那么极端的方法来写作这本书。每周5天，每天从早上5点到下午1点共8个小时，我把自己封闭起来专心写作。我给自己定下的基本规则是：直到下午1点前，不收发电子邮件，不接打电话，不安排约会，拒绝任何打搅。这样的状态也并不是总能实现，但是这种规则却产生了奇效。我设置了电子邮件自动回复，向对方解释在本书完工前，我将一直处于“修行模式”。这种方法给我带来的自由，怎么高估都不为过。通过创造空间来探索、思考和写作，我不仅提前完成了本书的写作，而且更好地支配了剩余的时间。

这看上去并不难理解，但是你能多久未从忙忙碌碌的一天里专门抽出时间，坐下来好好思考一下？我不是指你每天早晨上班路上抽出5分钟来列一个今日必做的事项清单，也不是指你在开会时溜会儿号琢磨琢磨手头的那个项目该怎么做。我指的是，有意识地留出没有分心物的时间和没有分心物的空间，除了思考，绝对不做任何别的事情。

今天的世界充斥着各种玩意儿，人们受着过度的刺激，因此，毫无疑问，如今要实现这样的专注思考已是难上加难。Twitter公司的一位领导人曾问过我：“你还记不记得无聊是什么感觉？这种感觉再也不会出现了。”他说的一点不错，也就是在几年前，如果你因为航班晚点而被

困于机场，或者在候诊室等待时，你可能会坐在那儿，发呆发愣，百无聊赖。然而现在，所有在机场候机的人和候诊室里等待的人都会选一样高科技工具，死死地盯着不放。当然，没有人会喜欢无聊，然而一旦所有无聊的机会都被剥夺殆尽，我们也同样丧失了可以用来思考的时间。

还有一种悖论需要人们去面对：工作节奏越快、越繁忙，就越需要我们在自己的工作日程中留出更多的时间用以思考。世界变得越喧闹，我们就越要开辟出一片可以安静反思的空间，真正实现全神贯注，心无旁骛。

不管你觉得自己多么忙，你依然可以从工作日中切出一块专门用以思考的时间和空间。比如，领英首席执行官杰夫·韦纳（Jeff Weiner），他每天都从自己的工作日程表中留出两个小时的空白，什么工作都不安排。因为接二连三的会议使他几乎没有时间来思考身边的世界到底发生了什么，他便养成了这样一个简单的惯例。刚开始时，这像是一种自我放任，而且还有浪费时间之嫌，但他最终发现这是最具价值的提升工作效率的方法，而且是独一无二的。他将之视为一种黄金法则，保证自己能够主宰自己的一天，而不是成为这一天的奴隶。

他向我解释说：“有那么一天我至今记忆犹新。由于情况特殊，我从早上5点到晚上9点一直都在不间断地开会，不是在参加电话会议，就是在各个会议室里。我清晰地记得，就在那一天结束的时候，当我想到自己对那一天的工作毫无发言权，而是任凭工作摆布时，自己非常沮丧。但是，那种沮丧的感觉很快就被一种感激之情取代了，因为自从我采用了现在这样的方法以来，这是我唯一记得的出现沮丧感的一天。”

这样的空间让他能够专注于思考最基本的问题：公司未来3~5年的状况如何？怎样才能最有效地完善一个已经很畅销的产品，或者处理客户未得到满足的需求？竞争优势怎样扩大，而竞争劣势又怎样消弭？与

此同时，他也利用这个空间来进行情感上的充电，让自己在问题解决模式和作为领导者而被期待的指导模式之间自如转换。

留出空间于杰夫而言，意义不止于一种惯例，而是一种更宏大的哲学的组成部分。他亲眼看见过盲目贪多求全给公司发展和高管生活带来的恶劣影响。因此，留出空间对他来说绝不是一种口号，也不是一个时髦词汇，而是一种哲学。

留出空间，阅读思考

比尔·盖茨的故事可以给大家带来更多的灵感。他经常性（而且是众所周知）地从日常工作中抽出一个礼拜时间，仅仅是为了思考和阅读。我曾参加过一次位于华盛顿州西雅图的比尔与梅琳达·盖茨基金会（**Bill&Melinda Gates Foundation**）总部举行的关于比尔的问答会。凑巧的是，他当时刚刚结束了最近的一次“思考周”。尽管我之前已经听说过他有这样的惯例，但是并不知道原来他从20世纪80年代就开始这样做了，并且一直坚持着，甚至在微软公司发展的鼎盛时期也不曾间断过。

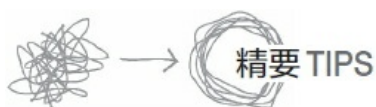
也就是说，在公司发展史上最繁忙、最迅猛的时期，他依然为自己创造时间和空间，每年安排两次什么也不干的一周，专门用来阅读文章（他的最高纪录是112篇）和书籍，学习新技术，勾画更远大的图景。现在，他仍然从自己日常的基金会工作中抽出时间来，不为别的，只是为了思考。

留出整整一周似乎有点奢侈和不可能，但是在每一天都享受一点点“思考周”，办法还是有的。我发现有一种做法很有用，也很简单，就是用每天的第一个20分钟来阅读经典文学（不是博客，不是报纸，也不是最新的海滩小说）。这种做法不仅打消了我以前一醒来就要去查收电子邮件的冲动，而且使这一天有了中心任务。这种阅读开拓了我的视野，让我回想起了很多重要的主题和理念，其中的精要和华彩足以抵御

时间的考验。

我偏爱励志文学，这种选择有点个性化，但要说到感兴趣的，则有以下这些：《禅宗，非理性的理性》《孔子的智慧》《摩西五经》《圣经》《道，知与不知》《古兰经的深意》《人如其所思》《甘地典藏》《瓦尔登湖》《摩门之书》《沉思录》《奥义书》。可读的书很多，只要保证选择的书写于我们这个过度互联的时代之前，并且具有永恒价值即可。这样的书籍能够激励我们去颠覆关于“何事最重要”的种种固有见解。

不管你是每天花2个小时，还是每年花2个星期，或者甚至只是每天早上花5分钟，这些都不重要，重要的是你要创造一个空间，让自己从忙碌的生活中抽身退出。



- 为自己创造探索和思考的空间。
- 适时抽身，退一步看清局势。
- 用每天的第一个20分钟来阅读经典。

CHAPTER 05

LOOK

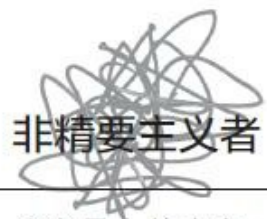
See What Really Matters

要素二

审视，关注真正重要的事情

知识，
那丢失在信息里的知识，
要到哪里去找回？

——T.S.艾略特



非精要主义者

注意最大的声音

听所有说出来的话

被海量信息压垮



精要主义者

关注声音所传递的信号

听没有说出来的话

审视查找信息的精华

已故作家诺拉·艾芙隆（Nora Ephron）最出名的作品包括《西尔克伍德》《西雅图未眠夜》《当哈利遇见莎莉》，每一部都获得了奥斯卡奖的提名。作为作家和编剧，艾芙隆的成功在很大程度上要归因于她捕捉故事精华的能力——这一技能是她早年当记者时练就的。但是，在高强度的新闻工作生涯中，对艾芙隆影响最深远的一课，还要追溯到她的高中时代。

查理·西姆斯（Charlie O.Simms）在比弗利山庄高中教一门新闻学课程。开课第一天，他的讲课和艾芙隆上过的任何一门新闻学课程并无二致，也是解释“导语”这个概念。他解释说，一个导语必须包含这篇新闻作品的“何时”“何地”“何事”“何因”“何人”这些要素，这些正是最重要、最精华的信息。然后，他给大家布置了第一项作业：为一个故事写一段导语。

西姆斯先提供了与这个故事有关的一些事实：“比弗利山庄高中的校长肯尼斯·彼得斯今日宣布，比弗利山庄高中全体教员将于下周赴萨克拉门托参加一场有关新型教学法的学术讨论会。人类学家玛格丽特·米德、大学校长罗伯特·梅纳德·哈钦斯，以及加州州长埃德蒙·杰·布朗将到会并发言。”

学生们噼里啪啦地敲打着键盘，努力要跟上老师的节奏。然后，他们纷纷交上了自己的“急就章”导语。这些导语每一篇都试图尽可能简明扼要地概括出“何时”“何地”“何事”“何因”“何人”这些要素：“玛格丽特·米德、梅纳德·哈钦斯及加州州长布朗将向全体教员发表讲话，主题是……”“下周四，高中全体教员将……”西姆斯检查了学生们写的导语，然后把它们放到了一边。

接下来，他告诉学生们，所有人写的都不对。他说，这个故事的导语正确的写法应该是“本周四不上课”。

“在那一瞬间，”艾芙隆回忆说，“我懂得了，新闻学不仅仅意味着要反刍事实，而且还要把握重点。弄清楚什么人、什么事、什么地方和为什么这些要素是不够的，还要理解这篇新闻到底意味着什么，以及它为什么重要。”艾芙隆还补充说：“他教我的某些东西不仅适用于新闻工作，也同样适用于生活中。”

在每一组事实中都隐藏着最重要的信息。一个优秀的新闻工作者知道，要找到最重要的信息，不仅需要研究这些信息，而且还要找出信息之间的关联（我的本科专业是新闻学，所以我对此高度重视）。这就意味着要使这些关系变得清晰明了，同时也意味着要在各部分综合的基础上构建整个故事，而且还要理解这些不同的信息碎片如何整合才能对每个读者具有意义。最优秀的新闻记者不只是简单地传播信息，他们的价值在于发现那些对人们真正重要的东西。

你是否也曾感到怅然若失，迷惘于究竟该关注什么？你是否也曾面对信息轰炸不知所措，纠结于这些信息该如何使用？你是否也曾对各种要求感到头晕目眩，不知道哪个重要，哪个不重要？你是否也曾在工作或家庭生活中未抓住要领，而等到发现错误时，已为时过晚？如果是，那么接下来的精要方法将会非常有价值。

把眼光放在大局上

1972年12月29日，美国东方航空公司（American Eastern Airlines）401号班机在佛罗里达埃佛格雷沼泽地（Everglades）坠毁，导致100余名乘客丧生。这是首次宽体客机失事意外，也是美国历史上最严重的空难之一。事故调查者们后来惊愕地发现，就所有重要的方面而言，这架飞机的操作状态完全正常。那么问题究竟出在哪里呢？

当时这架洛克希德飞机正准备降落，副驾驶艾伯特·斯托克斯蒂尔（Albert Stockstill）注意到了起落架状态指示器——一个小小的绿色指

指示灯没有亮起。但实际上起落架已经放下，有问题的是指示灯，而不是起落架功能。然而，由于机员们过度专注于起落架问题，导致他们未能注意到自动驾驶仪已经失效，到发现时已经为时过晚。也就是说，导致灾难的原因不是起落架，而是机员们忽视了更大的问题——飞行高度在下降。

如果你以新闻业为终身职业，那么不要再过多关注不重要的小细节，要把眼光放在大局上。不管你从事什么职业，甚至在你的个人生活中，你都可以把新闻记者的这种技能应用其中。通过训练自己寻找“导语”，你会突然发现自己能够找到已经忽略的东西。你不仅能看到每天的一个个点状事件，而且还能把它们联系起来看到各种趋势。你将不再对各种事实作出反应，相反，你会把注意力集中在真正重要的大事件上。

倾听话外之音，领悟言外之意

我们本能地知道，没人有能力研究自己在生活中接触到的每一条信息。要分辨出什么值得研究，我们必须自律地调查和筛选所有相互竞争而又相互冲突的事实、选项和观点，它们都很渴望得到我们的重视。

最近，我同《纽约时报》的专栏作家、普利策新闻奖获得者托马斯·弗里德曼（Thomas Friedman）一起讨论，如何从吵吵嚷嚷的无用信息中过滤出真正重要的信息。在见到我之前，他已经就自己正在撰写的专栏文章同几位消息人士进行了午餐会谈。其中一人起初以为，弗里德曼不会注意到餐桌上的各种插科打诨。但是，他其实是在听着的。午餐会上的所有交谈都逃不过他的耳朵，他只是在滤掉所有那些不能真正抓住他注意力的信息。然后，为了把分散的信息整合联系起来，他提了很多问题，但仅仅是针对那些激发他兴趣的信息而发问。

正如弗里德曼后来同我分享的一样，最优秀的新闻人，要听的是别

人听不到的东西。在午餐会上，他一直在听的不只是那些谈话中正题以外的东西，他听得更多的是那些没有说出来的话。

精要主义者都是出色的观察家和优秀的听众。他们深知，取舍的现实意味着自己不可能关注一切，因此他们会特别留意倾听话外之音，领悟言外之意。就像《哈利·波特》里著名的角色赫敏·格兰杰（Hermione Granger，她不可能是精要主义者，这一点我赞同，但是精要主义者在这方面都是一样的）所说：“实际上，我擅长逻辑思维，这有助于我通过无关细节抓住别人忽略的那些关键信息。”

非精要主义者也不是不听，但是他们听的时候，总想说点什么。他们会因为无关的噪声而分心，而且太注意那些无关紧要的细节。他们听到的，是最大的那个声音，但得到的，却是错误的信息。由于急于做出反应，他们抓不住要领。结果，套用刘易斯（C. S. Lewis）的一个隐喻，这些非精要主义者的所作所为就像在洪水泛滥的时候拿着灭火器跑来跑去，弄错了事情的关键所在。

在纷纷扰扰的现代职场，身边各种喧闹的声音企图把我们扯向四面八方，因此，现在比任何时候都要求人们要学会抵御海妖之歌的诱惑，集中精力关注新闻标题。运用以下几种方法，你也可以把自己培养成一个不对外挂牌的新闻记者。

培养“新闻记者”的四种方法

1. 写日记，成为自己人生的记者

显而易见，journal（日记）和journalist（记者）属于同一个词根。“记者”一词，按其字面的意思，指的是写日记的人。因此，要成为自己人生的记者，最明显而且很有效的方法就是写日记。

人是一种健忘的生物，这是一个悲哀的现实。我甚至还想说，这种健忘相当惊人。不信吗？

要是不信，你现在就可以测试这种说法的对错，只要回忆一下你自己上上个周四晚餐吃的是什么即可，再不行就问问自己三个星期前的那个周一参加了哪些会议。如果你不是太另类的话，那么你的脑中将会是一片空白。日记就像一个存储设备，当我们的硬盘驱动出错时，它将提供备份。有人曾对我说，好记性不如烂笔头。

在过去十年中，我一直保持着写日记的习惯，用一种违反直觉却很有效的方法：实际写下的要比想写的少。当人们开始写日记的时候，通常都会在第一天洋洋洒洒写下好几页。到了第二天，一想到要写那么多，就会气馁和却步，因此他们往往会拖延或者干脆就不写了。所以，在写日记时也应追求“更少，但更好”，在没有养成写日记习惯的时候，就不要写得太多。

我还想建议的是，每隔3个月，花1小时读一读这段时间里的日记条目。不要过于关注细节，比如3个星期前的预算会议或是上周四的意大利面食，而是要着眼于更宏大的图景或趋势。要捕捉标题，寻找自己每天、每周的生活中的导语。微小的、不断增加的变化一时很难发现，但是久而久之，就能产生累积效应。

2.实地探索，认清问题本质

简是一家设计学校“终端支付能力设计”课程的团队成员。这个课程给他们的挑战是设计一个婴儿保育箱，成本必须是一般保育箱成本（20000美元）的1%。简说，在发展中国家，“有400万低体重儿在出生后的前28天内死亡，原因是他们缺乏足够的脂肪来调节体温”。

如果他们将这个项目仅仅视为一个成本问题而草草上马的话，可能

已经制造出了一种便宜的塑料保温箱——看上去也不失为一种合理的解决办法，但结果是他们将无法从根本上解决问题。他们没有这样做，而是煞费苦心地寻找问题的关键所在。他们去尼泊尔目睹了问题的严重性，发现80%的婴儿不是在医院出生，而是在家里，那里还是缺电的农村地区。这样，团队真正的挑战就变得十分明确，那就是要创造某种根本无需用电的保温箱。有了这种关键性的认识，他们开始认真地解决手头的问题。最终，简和她的另外3名团队成员创立了一个名为“怀抱”的非营利公司，并创造了“怀抱巢”，它使用一种能在水里加热的类似蜡的物质，然后将这种物质放在一个类似豆荚的睡袋里，可以用于婴儿保温达6个小时或更长时间。通过深入实地探索，充分研究问题，他们得以进一步认清问题的本质，聚焦于最重要的细节，并最终使问题得到圆满的解决。

3.捕捉不正常与不寻常

玛丽亚姆·西曼（Mariam Semaan）是一位来自黎巴嫩的获奖记者。她最近完成了斯坦福大学的一项约翰·奈特新闻奖学金（John S. Knight Journalism Fellowship）研究项目，专门研究媒介创新和设计思维。我请她分享一下她的专业秘诀，即她多年来透过嘈杂表象直抵故事真正内核的经验所在。她给出了一个令人振奋的回答：寻找导语和捕获重要信息的能力都是可以后天习得的，你需要的是理解。要汲取故事的精华，需要深刻理解它的主题和语境，需要将它放在一个更大的图景中去俯仰观察，需要找出它和其他领域的关联。因此，她会读遍所有相关的新闻，努力找出其他报道忽略或关注不够的信息。“我的目标，”她说，“就是掌握故事的‘蜘蛛网’，因为它能让我捕捉到‘不正常’或‘不寻常’的细节和行为，它们与故事的自然发展进程格格不入。”

玛丽亚姆说，找到“一个给定故事的不同视角是非常重要的，它能够给主题带来新意，使之变得新鲜新颖、与众不同而又发人深省”。她

采用的技巧之一是角色扮演：让自己站在故事中所有角色的位置上，以便更好地理解他们的动机、理由和立场。

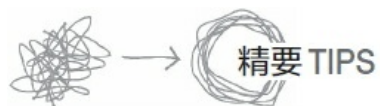
4.走出逃避推诿的怪圈

任何人，只要看过那些精通权术的政客如何接受采访，就都会知道这些人在答非所问方面，接受过多么登峰造极的培训。规避棘手的问题，对所有人都有诱惑力。通常，给一个模棱两可、漫无边际的回答，要比搜集必要的事实和信息来做一个考虑周全而有根据的回答要容易得多。但是，逃避推诿会导致一种不必要的恶性循环，那就是更严重的模棱两可和信息失真。要走出这个怪圈，办法之一是明确问题。

赛富时公司（Salesforce）的高级副总裁伊利·科恩（Elay Cohen），曾与其他5名队员组成小组，挤在一间闷热的宾馆房间里。在接下来的3个小时里，他们将和另外5个小组一起进行一场商业模拟竞赛。他们的任务是回答一系列关于如何处理假定的管理状况的问题。计时器已经开始计时了，但伊利的团队却还无法开始。因为每一个提出的答案都引发了更多的观点和评论，这样一来，原本应该是一个相当简单的解决问题的练习，却很快演变成了一个枝节横生、混乱无序的辩论。当时，我在观察和指导他们的队伍，这种情形出现15分钟后，我让他们停了下来，并问他们：“你们要回答的问题是什么？”每个人都很尴尬地停止了争论，但是没人回答我的问题。然后，有个人作了一个不相关的评论，随后全组又开始了离题万里的辩论。

我中途介入并再一次提出了我的问题，如此往复。最后，小组停止了辩论，并开始认真思考他们究竟想要实现哪些目标，以及要实现这些目标真正需要作出的决定有哪些。他们停止了讨论，费力地梳理了所有随意抛出来的想法和观点，倾听潜藏于其中的主题和能让这些主题产生有机联系的重大构想。然后，他们终于走出了“晕动症”的状态，获得了

一些进展。他们选定了一项行动计划，做出了必要的决定，并划分了职责。最后，伊利的团队大获全胜。



- 放眼全局，而不是纠缠细节。
- 去粗取精，自律地调查筛选。
- 写日记，成为自己人生的记录者。
- 实地探索，认清问题的本质。
- 捕捉不正常和不寻常的细节。
- 明确问题，走出逃避推诿的怪圈。

CHAPTER 06

PLAY

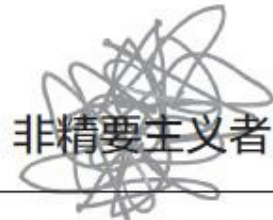
Embrace the Wisdom of Your Inner Child

要素三

游戏，拥抱内心孩童的智慧

最有智慧的人，会时不时听点儿废话。

——罗尔德·达尔



非精要主义者

对游戏不屑一顾

认为游戏徒劳无益、浪费时间



精要主义者

重视游戏

知道游戏能激发探索精神

在经典音乐剧《欢乐满人间》（*Mary Poppins*）的结尾处，性格暴躁而又沉闷的班克斯先生回到了家里，他已经被“解雇了，下课了，扔到大街上了”。但是，他看上去却绝对是一反常态的快乐，快乐到女仆认为他已经“不正常”，他儿子甚至认为“这听上去不像是自己的父亲了”。的确，当这位父亲给孩子们拿出那个修补好的风筝，并开始唱“让我们放风筝去”的时候，他几乎变了个人。离开了枯燥无味的银行工作，班克斯内心深处的那个孩童突然获得了重生。他的好心情效果非凡，大大提振了整个家庭的情绪，一扫先前的阴郁和忧愁，班克斯家成了欢乐、友爱和充满乐趣的居所。当然，这是一个虚构的故事，但它展现了让游戏重回日常生活，将会产生多么神奇的力量。

大多数人在孩提时代，都没有人正式地教过他们究竟该怎么玩。想象一下，当妈妈和新生儿玩躲猫猫时，小孩儿展现出的那种纯粹的快乐。想象一下，当一群孩子一起玩过家家时，那种无拘无束的想象力。再想象一下，当一个孩子用一堆旧纸盒搭建他自己的迷你王国时，他的那种被米哈里·希斯赞特米哈伊（Mihaly Csikszentmihalyi）⁴称作“心流”的状态。然而，随着年龄的增长，有人告诉我们，游戏不重要，它浪费时间，它毫无必要，它幼稚可笑。不幸的是，很多这样的负面信息恰恰是来自于那些富有想象力的游戏最应该得到鼓励而不是被抑制的地方。

school一词来自希腊语schole，意为“闲暇”。然而，诞生于工业革命时代的现代学校体系，已经扼杀了学习中的闲暇时光，乐趣也所剩无几。肯·罗宾逊爵士（Ken Robinson）⁵已经将研究学校教育中的创意学习当作自己毕生的工作，他注意到了学校实际上是在扼杀创意，而不是通过游戏来激发创意：“我们已经把自己卖给了一种快餐式的教育模式，它使我们的精神和活力日渐枯竭，就像快餐侵蚀我们的躯体一样。想象力是人类一切成就的源泉。我认为，我们现在教育孩子和我们自己

的方法正在系统性地破坏想象力。”在这个问题上，他无疑是正确的。

那种“游戏不重要”的想法跟随我们进入成年期，而且当我们步入职场后，它变得更加根深蒂固。遗憾的是，鼓励游戏的公司和机构实在太少太少，很多公司甚至在无意之中阻挠游戏。的确，有一些公司高管口头上承认游戏对于激发创造力功不可没，但他们中的绝大多数都没有创造一种能够真正激发探索精神的游戏文化。

对于这些，我们丝毫不感到意外。现代公司诞生于工业革命时代，那时它们存在的全部理由就是提高大规模生产产品的效率。而且，这些早期的管理人员还把眼光朝向了军队——一个更加没有游戏可言的组织，来寻求灵感（的确，军事化语言在今天的公司里还有很强大的影响力。我们现在仍然说员工们在前线，*company*一词本身就是表示军事单位“连”的术语）。尽管工业时代已经远去，但是那些习俗惯例、结构方式和制度体系在很多现代组织里依然随处可见。

游戏，我更愿意把它定义为任何一种我们仅仅为获得其中之乐而进行的活动，而不是一种达到目的的手段，不管是放风筝还是听音乐，抑或是到处乱扔棒球——也许它们看上去像是一种毫无意义的活动。游戏，通常被视为一种毫无意义的活动。但是，它实际上在很多方面都具有重要意义。美国国家游戏研究所（The National Institute for Play）创办人斯图亚特·布朗（Stuart Brown）通过研究大约6000个人的游戏史，得出了这样的结论：游戏，能够极为有效地改善包括个人健康、关系、教育、组织能力在内的一切。“游戏，”他说，“能使大脑具有可塑性、适应性和创新性。”同时，他简明扼要地指出：“在激发大脑功能方面，游戏的作用是无可比拟的。”⁶

大脑爱游戏

游戏在我们的生活中具有极其重要的价值。对动物王国的研究表

明，游戏对于关键性认知技能的发展极为重要，甚至还影响着物种的存活。研究员鲍勃·费根（Bob Fagan）对灰熊的行为进行了长达15年的研究，他发现最会玩儿的灰熊寿命最长。当被问及其中的原因时，他说：“在一个严峻挑战层出不穷而且无法预知的世界里，游戏能让这些熊更好地应对变化着的环境。”

雅克·潘克塞普（Jaak Panksepp）在《情感神经科学：人类和动物的感情基础》（*Affective Neuroscience: The Foundations of Human and Animal Emotions*）一书中得出了类似结论：“确定无疑的是，在游戏过程中，动物的行为往往变得更加灵活而且富于创造性。”

斯图亚特·布朗写道：“然而在所有的动物物种中，人类是最会玩儿的。我们生来会做游戏并且在游戏中生存发展。当我们玩儿的时候，人性得到了最纯粹的表达，个性得到了最真实的展现。游戏的那一刻，我们最深刻地感觉到自己活着，我们最美好的回忆也正是来源于此，这有什么好奇怪的吗？”

游戏鼓励我们进行新的探索，拓宽思维的疆域：创造新思想，对旧想法进行全新的审视。它让我们具有更强烈的好奇心，更能适应新奇的事物，更热衷于积极参与。游戏是成为精要主义者的根本条件，因为它起码以三种特殊的方式为人类的探索行为添砖加瓦。

首先，游戏拓展了人们选择的范围。它能帮助我们看到平常会忽略的种种可能性，并建立一些原本无法建立的联系；它能开阔视野，开放视角；它有助于我们挑战旧观念，更加乐于接受未经检验的新思想；它鼓励我们扩大自己的意识流，构想新故事。或者，就像爱因斯坦曾说的那样：“在我审视自身和自己的思考方式时，我得出的结论是，对我而言，幻想的天分比我积极吸收知识的才能更重要。”

其次，游戏是解压药。这种功用非常重要，因为压力不仅会导致生

产效率降低，而且还会抑制创造力，扼杀好奇心，损坏大脑的探索功能。这样的感觉你一定不陌生：工作压力让你透不过气来，突然间一切都不对劲儿了——你找不到钥匙，动不动就撞到东西，重要的报告落在厨房餐桌上了。最新研究成果表明，这种情况是由于压力导致大脑中情绪监测部位（杏仁核）活动增加，而同时大脑中负责认知功能的部位（海马体）活动减少——结果很简单，那就是我们无法清晰地思考。

我亲眼见过游戏在我孩子身上产生的与之相反的效应。当他们焦虑不安、无法控制事情时，我让他们画画，变化几乎立马就能出现。压力解除了，孩子们的探索能力失而复得。

最后，游戏可以对大脑的执行功能产生积极的影响。正如专门研究脑科学的精神病学家爱德华·哈洛韦尔（Edward M.Hallowell）²解释的那样：“大脑的执行功能，包括制订计划、区分主次、安排日程、预先准备、委派分配、作出决策和总结分析——简而言之，就是任何一个要在商界有出色表现的领导者所必须具备的技能。”

游戏能够刺激大脑中的相关区域，进行小心谨慎的逻辑推理和无忧无虑、不受束缚的探索。有鉴于此，人类思维领域的重大发现都出现在游戏时刻，也就不足为奇了。哈洛韦尔写道：“哥伦布正是在玩的时候发现了地球是圆的。牛顿正在放松自己的大脑时看到了苹果树，并突然发现了地心引力。沃森与克里克正在玩DNA分子各种可能的模型时，偶然发现了双螺旋模型。莎士比亚玩抑扬格五音步玩了一生。只要醒着，莫扎特就离不开玩。要说到大脑爱游戏，爱因斯坦的思维实验就是极好的例子。”

让游戏走进工作，让游戏走进生活

有些创意公司最终还是认识到了游戏所具有的重要价值。Twitter公司的首席执行官迪克·科斯特罗（Dick Costolo）鼓励员工玩喜剧表演，

他在公司发起了一种即兴表演课程。作为一个曾经的脱口秀谐星，他十分清楚即兴表演会迫使人们去努力拓宽思路，突破各种惯性思维的束缚，进行更加灵活、更有创造性的思考。

其他一些公司则鼓励在物理环境中开展游戏活动。IDEO公司在微型汽车里开会。在Google公司的大厅里，你很可能会碰巧发现一只巨大的、浑身覆盖着火烈鸟的恐龙（多个例子中的一个）。在皮克斯动画工作室里，艺术家“办公室”的布置各出奇招，从旧时代的西方小酒馆到小木屋，不一而足（在参观时，最令我惊艳的是那间从地面到天花板排列着数千个《星球大战》小雕像的房间）。

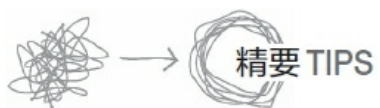
我曾经认识一位在出版公司工作的成功女士，她在自己的办公桌上放着一个史泰博公司（Staples）生产的“轻松按钮”。每个离开她办公室的人都会像孩子般激动地用手掌在那个大大的红色按钮上拍一下——录好的声音就会大声响起，向整个办公室宣布“小事一桩！”出了这位女士的办公室，沿着公司的走廊往后，另一位女士在她的办公室里放了一幅带框的装饰画，画中是儿童图书中的插图，这总能让她想起孩提时代阅读的乐趣。

可能对于有些人来说，办公桌玩具，覆盖着火烈鸟的恐龙，充满了动作片雕像的办公室，都是无价值的、令人分心的事物，但问题的关键却是它们的作用恰恰相反。以上这些人所做的努力“颠覆”着非精要主义者所谓的游戏不重要的逻辑，而且还把游戏视作创造性和探索活动的重要驱动力。

因此，我们该如何让游戏更多地走进我们的职场和生活？布朗在他的书中提供了一种入门方法来帮助读者重温游戏。他建议说，读者应从过去经历里开采出有关游戏的记忆。在你的孩提时代，什么游戏能让你激动不已？今天，你如何让这种激动重现？

**Play doesn't just
help us to explore
what is essential.
It is essential in
and of itself.**

游戏不只是有助于我们探索重要事物，
它本身就是重要事物。



- 让游戏重回日常生活。
- 用游戏激发创意。
- 从过去的经历中开发出有关游戏的记忆，让儿时的激动重现。

CHAPTER 07

SLEEP

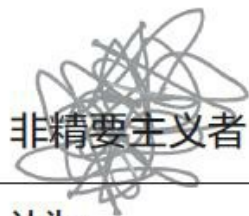
Protect the Asset

要素四

睡眠，保护你的身心资产

每晚，当我入睡时，我死去。次日，
当我醒来时，我重生。

——圣雄甘地



非精要主义者

认为：

少睡一小时，生产效率提高一小时。

睡眠是为失败者准备的。

睡眠是懒惰的温床。

睡眠会妨害“做成一切”。



精要主义者

懂得：

多睡一小时等于几个小时的更高效率。

睡眠是高成就者所必需的。

睡眠是创造力的孵化器。

睡眠有助于实现高水平的智力贡献。

杰夫直直地坐在床上，惊慌万分。他感到仿佛有一颗炸弹在脑中炸裂开来。他汗流不止，茫然失措。他竖起耳朵紧张地听着什么。到底发生了什么事？万籁俱寂。也许是他吃的东西引发了这样怪异的反应。他试图躺回去继续睡觉。

第二天晚上，同样的事情再度发生。过了几天，同样的事情发生在中午。他刚从印度回来，起初，他觉得这是药物产生的反应。为了在倒时差时改善睡眠，他服用了苯海拉明，同时还服用了疟疾药。但是，随着情况越来越糟，他发现自己的问题更加复杂了。折磨他的好像是焦虑症，但是他没有任何焦虑，只不过是身体上有些症状罢了。

杰夫的成绩特别优异，他很渴望能有所作为。交代一些故事背景，杰夫的祖父是美国和平队的早期官员。杰夫野心勃勃，干劲十足，并决意对世界有所贡献：他是Kiva的董事会成员；被提名为安永（Ernst&Young）年度企业家；被世界经济论坛选为全球青年领袖；与他人共同创立了一项成功的影响力投资基金；担任一家全球性微型贷款组织的首席执行官，该组织服务全球1200多万贫困家庭。他年仅36岁，却已成为所在行业的领军人物。

杰夫频繁出差，这令他时常睡眠困难。他的公司总部在西雅图，但在旧金山、印度和肯尼亚都有办公室。他需要例行性地飞赴伦敦参加会议，然后前往印度的5个不同的城市待6天，再去日内瓦同投资者们进行几小时的会晤，然后再返回西雅图待上一天半时间。在3年里，他有60%~70%的时间都在出差中，每晚平均睡眠时间只有4~6个小时。

在36岁这个成熟的年纪，这样的工作节奏开始给他的健康和工作能力带来威胁。伴随着每晚的睡眠障碍，他的状况越来越糟，身体器官开始一个个出现问题：心律不齐，站直的时候有疼痛感，消化障碍迫使他必须把食品混合起来吃，血压太低导致他快速起身时眼前一片黑暗。他

已经看过两次急诊。他不停地告诫自己，做完下一次生意后要放缓工作节奏，然后是再下一次，然后下下一次。但是，毫无疑问，他并没有这样做。他确信即使继续维持现在的状态，他也可以克服这些困难的。缩减工作是有成本的，需要作出一些取舍，这是他所不愿意的。但是，他很快便尝到了恶果：常常被迫在最后一分钟因为身体太虚弱而取消会议，或者本来要做一次演讲，却因为大脑如坠云里雾里，心有余而力不足。他开始怀疑，自己对公司造成的损失是不是大于贡献——这是肯定的。

最后，经过明确的诊断，医生给了他两个选择：长期服用药物进行对症治疗，或者彻底停止工作一到两年进行治疗和康复。刚开始时，杰夫拒绝作出这样的取舍。他曾是一个很有实力的铁人三项运动员，他认为自己可以运用和对付踝关节扭伤或者肩袖撕裂同样的逻辑来处理现在的状况。他自负地告诉医生，只需要休息几个月，就能全线回归正常的工作：“你瞧着吧！”

他休了两个月的假，但让他意想不到的的是，他彻底动弹不了了。他每晚睡眠时间长达14小时！然后一整天都在休息。有些天，他连床都起不了。整整6个星期，他什么都干不了，百无一用。他爬着回去见了医生，承认几个月是不够的，他需要休息更长的时间。

他言出即行，摆脱了给他生活带来压力的一切事情。他辞去了在几个董事会里的职务，甚至还决定离开自己的公司。他说：“下定决心脱离这些，是一件非常非常艰难的事情。当我离开董事会会议室的一刹那，眼中噙满了泪水，我告诉妻子，‘我不想以这种方式离开自己的孩子’！”

在执行治疗方案的过程中，他设计了一种新的生活，全身心地投入健康的恢复中。他改变了自己的饮食习惯，带着全家一起去法国南部生活了一年。治疗和气候及生活方式的变化产生了效用。他开始用一种新

的思维方式来理解自己在这个过程中学到的东西。

两年半后，杰夫在坦桑尼亚参加世界经济论坛的全球青年领袖评选活动。有一天晚上，在一个自由麦克风活动中，有些了解杰夫的人让他跟在场的200位成功人士分享心得体会。他情绪激动地告诉大家，他花了高昂的代价，学会了一个简单却极其重要的道理：“保护最珍贵的资产。”

投资自己的思想、身体和精神

就给世界作出贡献而言，最珍贵的资产莫过于我们自身。如果不能对自己进行足够的投资，我说的是自己的思想、身体和精神，那么我们就毁坏了用以达到个人贡献峰值的最佳工具。人们，尤其是充满野心的成功人士，用来破坏这种资产的最常见的方式之一就是睡眠缺乏。

如果让A型性格⁸占据上风，那我们就会像杰夫那样全然耗尽自己，生命将过快地燃烧殆尽。对待自身，我们应该像对待自己的职业和生意一样具有战略性。我们必须放稳自己的步调，滋养自己的生命，给自己加油添薪，用以探索、成长和实践。

在长时间的休息中，杰夫发现了自己对成功上瘾的一个有趣悖论：一个A型性格的人，让自己拼命工作并不难，让自己达到极限也是易如反掌！真正的难题是让一个通过应对挑战而创造成就的人去放弃拼命工作。他向那些成就突出的人解释说：“如果你自认为自己强大到可以做成一切，那么我有一个挑战给你。如果你真的很想做成某件事，那就和这个机会说‘不’，这样你就能让自己小睡片刻。”

我在21岁的时候也觉得，睡觉并不是什么非同小可的事情，少睡会儿也没有什么大不了的。对我而言，睡觉是一种无法摆脱的麻烦：它浪费时间，而这些时间本可以创造成效，只有弱者或者意志薄弱的人，才

会对睡眠欲罢不能。我曾经陶醉在成为超人的幻想中，每晚只睡几小时即可。为此，我甚至实验了几种非常激进、超常规的方法来减少睡眠时间。

我得知，有一项睡眠研究要求参与者在一整天之内只能每4个小时睡20分钟，我也试了试这种方法。时间不长的话，用这种方法还是可以忍受的。但是我很快就发现，虽然在技术上这种睡眠安排可以满足人的生存，但是它也存在着缺陷。例如，当我醒着的时候，大脑却几乎不怎么工作，更不用说作出计划，排出主次先后，或者是看到大局；要进行决策或选择也很困难；要去粗取精、区分重要和不重要则几乎更是不可可能。

这种方法很快就难以为继，但我依然固执地认为，睡得越少，干得越多。因此，我采用了一种新的方法，就是每周熬一个通宵开夜车。可是，这样做没什么效果。我妻子对这种做法颇有微词，她给我看了一篇文章，这篇文章彻底改变了我对睡眠的看法。文章否定了睡眠是效率之敌的观点，并令人信服地证明了睡眠实际上是最佳表现的驱动力。文章引述了一些顶级商业领袖的保证8小时睡眠的自豪言论。我还记得其中引用了比尔·盖茨的话，他人生中的每一次严重失误都是在睡眠不足的情况下发生的。自此以后，我努力保证每晚8小时睡眠。

你的情况如何？就想想上星期吧。你是不是有某晚睡眠不足7个小时？你是否一连好几晚都是睡眠不足7个小时？你是否发现自己曾经骄傲地说过或者自豪地想过：“不包括我。我不需要睡足8个小时。我只要睡4个或者5个小时就可以了（如果你此刻正一边读着，一边这样想着，那么你将从这个章节里获益匪浅）”。还有些人确实只需要较少的睡眠时间，不过据我了解，他们绝大多数人已经太习惯于疲惫状态，以至于都已经忘记了充分休息究竟是怎样一种感觉。

在非精要主义者看来，睡眠给他们那种早已精力过度分散、责任义

务超负荷、繁忙却并不总是有效率的生活带来了另一种负担。而精要主义者却认为，睡眠更多时候是实现个人贡献峰值的必要条件。因此，他们会在工作日程安排中系统性地、有意识地为睡眠留下一席之地，以便让自己干得更多、成就更大，拓展更多疆域。通过“保护自己的资产”，他们能够在每天的生活中保留一部分精力和创造力，以及解决问题的能力，以备不时之需——不像非精要主义者，他们永远无法得知自己会在何时何地为自己的疲劳虚弱绑架劫持。

精要主义者选择在当下做更少的事情，以便在未来做得更多。当然，这就是一种取舍。但是，日积月累，积少成多，小小的牺牲会带来巨大的收益。

睡眠属于傻瓜？

如果“保护资产”是如此的重要，那为什么人们会轻易地放弃宝贵的睡眠？就成就卓著者而言，这其中的原因之一，可能是他们就像我一样，错误地认为如果睡得更少就能创造更大的成就。然而，我们有足够的理由来颠覆这种观念，比如，越来越多的研究表明，一个晚上的好睡眠实际上能让我们更高效，而不是相反。

安德斯·爱瑞森（K.Anders Ericsson）曾对小提琴家做过一项著名的研究，马尔科姆·格拉德威尔（Malcolm T.Gladwell）将之通俗化为“1万小时法则”。安德斯发现，相比那些仅能称作好学生的人，最优秀的小提琴家在练习上花的时间更多。他的这一发现为非精要主义者提供了逻辑支持，即精通某项技能需要付出认真专注和苦心孤诣的努力。而且令人备受鼓舞的是，它让我们懂得，原来出类拔萃也是我们力所能及的，而不仅仅专属于那些最具自然禀赋的幸运儿。然而危险的是，它同时也可能会巩固非精要主义者“我必须什么都做”的思维定式，这种有害无益的谬见会误导人们去证明花更多更多的时间来工作是正确的，而这种工

作的收益实际上却在不断减少。

情况的确如此，但是我们应该看看这项研究中一个不那么为人所知的发现：睡眠是最优秀的小提琴家和优秀小提琴家之间的第二大重要分水岭。最优秀的小提琴家每天平均睡眠时间为8.6小时——这比美国人的平均睡眠时间大约多1小时。在一星期内，他们的平均午睡时间为2.8小时——比平均水平大约多2小时。研究者们得出结论，睡眠有助于这些顶级的演奏家补充体力和精力，从而更加专注地投入到练习之中。因此，当他们花了更多的时间来练习时，他们这几个小时的练习产生了更好的效果，因为他们已经得到了更充分的休息。

《哈佛商业评论》曾刊登过一篇名为《睡眠不足：业绩杀手》（*Sleep Deficit: The Performance Killer*）的文章，哈佛医学院的巴尔蒂诺睡眠医学教授查尔斯·切斯勒（Charles A. Czeisler）在文中解释了睡眠缺乏对优良业绩的破坏作用。他把缺少睡眠比作过量饮酒，其中的原理是熬一个通宵（也就是连续24小时不睡觉）或者一周内每晚只睡4~5个小时，“实际上会诱发相当于血液中酒精含量达到0.1%时的损害。试想一下：我们永远都不会说‘这个人是一名优秀员工！他喝下了所有时间！’然而，我们仍然在赞扬那些为工作牺牲了睡眠的人。”

睡眠通常被认为有助于身体的休息，但最近的研究表明，睡眠其实对大脑的作用更大。德国吕贝克大学一项研究提供的证据显示，一个完整的晚上睡眠能增强大脑功能，提升解决问题的能力。

《自然》杂志的报道称，在这项研究中，100多名志愿者接到了带着非常规迷阵的数字猜谜题，要求找到一个“隐藏密码”来揭开谜底。志愿者们被分成了两个小组，一组被准许8小时不间断睡眠，而另一组的睡眠是中断性的。科学家们随后观察哪些志愿者找到了隐藏的密码以及用了多长时间才找到。结果，在解决了问题的人中，来自连续睡眠8小时组的志愿者人数比来自睡眠不足组的志愿者人数多两倍。究竟是什么

么原因导致了这样的结果？研究者们解释说，睡觉时，我们的大脑会努力对信息进行编码和重构。因此，当醒来时，我们的大脑可能已经建立了新的神经细胞间的联结，因此能为解决问题打开更广阔的思路，提供更多的解决方案，而这几乎发生在一夜之间。

对于那些早起的人和熬夜的人，好消息还是有的：科学研究显示，哪怕是打个盹儿也能有助于提高创造性。仅举一个例子，发表于《美国国家科学院院刊》中的一篇报道显示，即使一个快速眼动睡眠期（Rapid Eye Movement, REM）也有助于提高不相关联的信息之间的整合。换句话说，一个短暂的深度睡眠有助于我们建立多种新的联系，去更好地探索世界。

简而言之，睡眠有助于我们达到个人贡献峰值，从而用更少的时间来创造更大的成就。尽管那种废寝工作的英雄主义文化仍将继续，但万幸的是，这种糟糕的情形正在日渐式微，这在一定程度上要感谢那些超高成就者——特别是在那些典型的倡导蜡烛两头烧、白天黑夜连轴转的行业里，公开炫耀自己拥有完整的8小时睡眠的人。在这些人中，很多都是精要主义者，他们明白健康的睡眠习惯带给了他们巨大的竞争优势，这是对的。

亚马逊网络购物中心的缔造者杰夫·贝佐斯（Jeff Bezos）就是其中一人。他说：“良好的睡眠习惯让我反应更灵敏，思维更清晰。如果有8小时睡眠，我一整天都感觉好多了。”网景公司联合创始人马克·安德森（Mark Andreessen）也是这样的人。他以前常常工作到凌晨，但是依然坚持在早上7点起床，现在，他已经改变了过去那种睡眠缺乏的状态。他说：“我会整天都盼望着能够回到家，再回到床上睡觉。”现在，按照他自己的话说，他的睡眠水平已经达到了“7个小时我开始变得机能下降，6个小时就不达标了，5个小时就会出大问题了，4个小时意味着我将无精打采”。周末他会睡12个小时以上。“这对我的能力发挥带来了

显著的影响。”他这样说道。

这些高管的言论被一篇题为《睡眠是成功企业家新的社会地位象征》（*Sleep is the New Status Symbol for Successful Entrepreneurs*）的文章所引述。《华尔街日报》的南希·杰弗里（Nancy Jeffrey）这样写道：“这是官方认定的。睡眠这种稀缺的日用品，在神经紧绷的美国，已然成为新的社会地位象征。曾经睡眠被嘲笑为懦弱和失败——那些曾喊过‘午餐属于失败者’的‘80后’高成就者也相信‘睡眠属于傻瓜’，现在睡眠已经被捧为创造性执行思维的复元伴侣。”对此，我们可以补充说，睡眠也是有辨识能力的精要思维的复元伴侣。

在《纽约时报》的一篇文章里，雷曼兄弟前首席财务官艾琳·卡兰（Erin Callan）讲述道：“在2005年的一次办公室聚会上，有位同事问我当时的丈夫我周末都干些什么。她也知道，我是一个能量十足而又精力充沛的人。‘她划皮划艇，攀岩，然后跑半程马拉松吗？’她开玩笑地问道。‘不’，他简洁地答道，‘她睡觉’。这一点儿不假。不赶工作的时候，我用周末时间为自己的身体充电，为即将到来的一周做准备。”

Google公司的小憩仓

如果在你的工作环境里，睡眠依然恶名不减，你可以考虑在工作中主动采取一些措施来明确地鼓励睡觉。如果那样做听上去有点激进，那么就考虑一下睡眠的诸多好处，比如提升创造力，提高生产效率，甚至减少医疗费用，这些都具有对最重要的因素产生直接影响的潜力。有了这样的认知，鼓励你的经理或者人力资源部制定一项书面政策就不难想象了（毕竟，很多公司都有针对饮酒问题的政策，而且，我们也已经认识到了饮酒和睡眠不足会对业绩产生类似的影响，这种情况是真实存在的）。例如，哈佛医学院的查尔斯·切斯勒提出了一项政策建议，规定员工在乘坐红眼航班后不得驱车赶去上班，其他公司准许头天晚上加班

工作的员工次日推迟上班。这样的公司和领导深知，“保护他们的资产”是一种受托责任。

得益于本书的写作，我最近去Google公司体验了一下他们著名的小憩舱。这是一个白色的宇宙飞船舱（就像你能想象到的某种能在20世纪70年代电视剧《默克与明蒂》[Mork and Mindy]中看到的東西），约2平方米大，足够让人躺下，但不能完全平躺。它有一个圆顶状的盖子，能够遮盖身体的大部分，但不能完全遮住，因此我起初感觉有点儿不自在，而且还担心自己能否睡着。30分钟后，小憩舱轻轻地颤动起来，提醒我时间到了，我根本无须担心。

小睡醒来后，我真切地感受到了自己是多么需要这样的休息。我感到更清醒、更敏锐，也更灵活了。

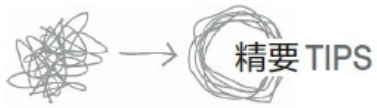
Google公司的小憩舱使用有一个签名日历表。我很想知道我在的那个星期有多少人使用了它？共有50人在小憩舱所在的楼层上班，我猜起码有10~20人使用了这个舱。但是我错了。根据签名日历表，只有一个人抓住了这个30分钟的午间睡眠机会，为自己的大脑和身体补充能量。不过，这个小憩舱的存在本身就具有一种重要的信号意义，告诉员工睡眠的重要性。

在本书的这一部分，我们讨论了如何探索和评估各种选项，以便从许多鸡毛蒜皮、平庸普通，甚至是一些还不错的事情中区分出最精要的少数。从定义上说，这是一个确定先后主次的过程。这其中的挑战包括过滤掉那些乍一看都很重要的选项。然而，精要逻辑告诉我们，现实中只有少数事情具有非凡的价值，其余绝大多数事情的重要性远远不及。睡眠不足会导致分辨能力下降，并进而损害我们确定主次先后的宝贵能力。

**Our highest
priority is
to protect
our ability to
prioritize.**

每个人的头等大事，
是确保自己有能力决断什么才是头等大事。

睡眠能让你更有能力探索世界，建立联系，在醒着的时候做得更少，但更好。



- 对自己的思想、身体和精神进行足够的投资。
- 粉碎睡眠的恶名，保护最重要的资产。
- 在工作中采取一些措施来明确地鼓励睡觉。

CHAPTER 08

SELECT

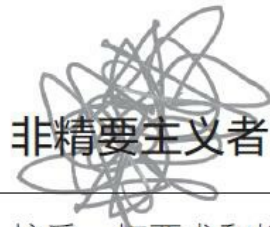
The Power of Extreme Criteria

要素五

精选，只接受前**10%** 的机会

内在的进程需要外在的标准。

——路德维希·维特根斯坦



非精要主义者

接受一切要求和机会

运用宽泛的、盲从的标准，比如，
“如果我认识的人在做，那么我也应该做。”



精要主义者

只接受前 10% 的机会

运用挑剔的、明确的标准，
比如，“这确实是我所追求的吗？”

在一篇名为《没有差不多。不是YES！就是No》（*No More Yes.It's Either HELL YEAH! Or NO*）的文章里，TED人气演说家德雷克·西弗斯（Derek Sivers）描述了一个简单的技巧，让人们在选择时更加精细。关键是要对这个决定进行一次极端测试：如果自己彻底地、无条件地确信要去做什么事，那么就答应，这就是德雷克法。任何事情，只要不是那么确信，就断然拒绝。或者就像Twitter公司的一位领导曾告诉我的：“如果答案不是一个确定的**YES**，那么就应该是一个肯定的**NO**。”这就是对精要原则的一个简明扼要的概括，它在探索的过程中具有举足轻重的地位。

德雷克自己就是这个原则的身体力行者。在他面试的求职者中，如果没有一个人能够打动他，他就会全部拒绝。最终，他准确地找到了合适的人选。当他发现自己对已报名要参加的好几个在世界各地举行的会议的兴趣并不大时，就决定待在家里，放弃所有这些会议，把省下来的12天时间用于效率更高的事情。为了决定在哪里居住，他否定了一些看起来非常好的地方（悉尼和温哥华），直到后来到了纽约，他马上知道这才是他理想的居住地。

再回想一下，当我们用一种宽泛的标准：这件衣服我未来哪天还有机会再穿，衣柜会遭遇到什么样的情况？它会被几乎不穿的衣服塞得满满当当、乱七八糟。但如果问问我真的喜欢它吗？那么我们就能狠下心来清除那些杂七杂八的衣服，腾出空间来放更好的衣服。我们可以举一反三地将这种标准运用到做其他选择的时候，不管事大事小，重要还是不重要，人生任何一个领域的选择都可以如此作出。

90%法则

最近，我和一位同事正在从大约100名申请者中为我们的“设计你的精要人生”课程挑选24个人。一开始，我们只设定了一套最低标准，比

如“每节课都能参加”。然后，我们设定了一套具有理想属性的标准，比如“是否准备好体验一次改变生命的历程”。运用这些标准，我们对每位申请人进行了从1到10的评分。对于得9分和10分的人，我们决定准予参加，低于7分的人则自动淘汰。我当时的任务是评估那些分数不高不低的人，即得7分和8分的人，这真是一项令人头疼的工作。在艰难地决定这些申请人是否足够好的过程中，我遵循了这样的思路：如果某物（在本事件中是某人）刚刚好或者差不多，也就是正好7分或者8分，那么结果就是淘汰。这大大简化了问题，令我如释重负。

这种方法可以被视为**90%法则**，在做每一项决定或每一次进退维谷的时候，你都可以运用它。评估每一个选项时，考虑一个最重要的决定标准，然后在**0~100**之间为这个选项打分。如果得分率低于**90%**，那就自动把评分降为**0**，并且断然淘汰它。这种方法可以避免在决策时举棋不定，或者更糟，不得不接受60多分或70多分的选项。试想如果某次考试你得了65分，你的感觉会是怎样的。那么，为什么要在作出人生的一项重要选择时有意去体会这样的感觉呢？

为了掌握精要技巧，在本环节中最重要，莫过于要注意承认取舍的存在。根据定义可知，运用具有高度选择性的标准本身就是一种取舍，有时候我们不得不拒绝一个看上去非常好的选项，心中怀着“完美的选项将会很快出现”的信念。可是，它有时会出现，有时又不会出现。关键在于运用高度选择性标准这个行为本身，将迫使你自己来选择你要等待的那个完美选项究竟是哪一个，而不是让别人，或让这茫茫宇宙来替你选择。如同任何一个精要技巧一样，它需要你通过设计来进行选择，而不是依从默认设置。

在做人生各个领域中的决定时，这种超级精挑细选的方法有着显而易见的好处：如果我们的选择标准过于宽泛，就会发现自己要面对的选项太多太多。更重要的是，对选项的价值进行评分，能让我们更加自

觉、更有逻辑、更加理性地作出选择，而不是听凭一时的心血来潮或受情绪支配。当然，要运用严苛的标准需要我们高度自律。但是，若不运用的话，代价会更高。

非精要主义者在个人生活和职业生涯中进行选择时，运用的是一种盲从的，无法明确表达的标准。比如，要决定承担什么样的工作项目时，非精要主义者运用的标准是盲从的，“如果经理让我做，那么我就做。”甚至更宽泛些：“如果有人让我做什么，那我就努力去做。”或者还有比这更宽泛的标准：“如果公司里有别的人在做这个，那么我也应该做。”在社交媒体高度发达的时代，我们对别人的所作所为更加了如指掌，这种宽泛标准会带来极其严重的负担，因为它会大大增加我们“应该”做的非精要事务。

我合作过的一个执行团队曾有一度确定了三种标准来筛选要承担的项目。但是，过了一段时间，他们就变得越来越不加选择了，最终这个公司似乎只剩下了一种通用标准，那就是客户让他们这么做的。结果，团队士气骤然大减，这不仅仅是因为承担了太多项目而让团队成员精疲力尽，不堪重负。还因为，没有一个项目被证明是值得做的，他们的工作缺乏明确的目标意识。更糟糕的是，公司的经营行为在市场上变得极其大众化，以至于失去了辨识度，而此前他们已占据了一个独一无二且利润可观的市场定位。

只有通过确定极端标准，他们才能摆脱那些耗尽他们时间和资源的70%和80%的工作，聚焦于最有意思的工作，从而最大限度地让自己在市场上被人一眼认出。而且，这种体系使员工能够自己选择工作项目，以便达到个人贡献峰值。曾经是听凭反复无常的管理层决定摆布，现在，他们也拥有了话语权。有一次，我看到这个团队里一个最安静、资历最浅的成员拒绝了最高管理者安排的工作。她的回答非常简单：“我们是不是应该考虑一下我们已经在执行的标准？”但是，在那种既有选

择性而又明确的标准制定之前，从来没有出现过这样的情况。

制定一种既有选择性而又明确的标准，能给我们提供一种系统的工具，用以辨别真正重要的事情，并过滤掉那些不重要的。

Make our criteria both selective and explicit.

制定一种既有选择性而又明确的标准。

如果不是一个确定的**YES**

Vitsoe公司总经理马克·亚当斯（Mark Adams）花了27年时间，有意识地把选择性标准运用到工作中。

Vitsoe公司从事家具制造。这个行业因为快速大批量的生产而声名狼藉：每个季节都会提供大量颜色不同、款式新颖的产品。然而，Vitsoe公司却几十年只提供一种产品：606万用置物柜系统。为什么呢？原因非常简单，Vitsoe拥有极为独特的标准，而606万用置物柜系统是唯一达到这个标准的产品。

这个系统集中体现了本书开头述及的“更少，但更好”这种精要主义

标准的特点，迪特·拉姆斯对此十分推崇。也许是巧合之中的巧合，606万用置物柜系统也是由迪特设计的。但是，Vitsoe公司在人才选拔方面，采用的标准更具选择性。

他们最开始的时候，遵循了一个基本原则，那就是宁可人手不足，也不草草地招进不合适的人。按照这个标准，他们在物色新员工的时候，设立了严格而系统的选择程序。第一步，他们对求职者进行电话面试。这么做是故意的，因为他们想在形成第一印象时，排除视觉因素的影响。同样，他们也想听听未来员工在接电话时的表现，以及在那段分配给他的访谈时间里，这个员工的办事条理性是否强到可以找一处僻静的地方。通过这种高效的方法，他们淘汰了很多。

第二步，一名求职者要接受全公司很多人的面试。如果他能通过几轮面试，那就可以被邀请去和团队一起工作一天。然后，管理人员会给整个团队发一份调查问卷，了解他们对这个求职者的印象。但是，他们问的不是那种显而易见的问题，而是“他或她喜欢在这里工作吗？”，以及“愿意让他或她留下来和我们一起工作吗？”在这个阶段，公司不提供录用通知，求职者也不作出任何含蓄的承诺，目的是为了双方都能尽可能公正地看清对方。如果双方都感到合适的话，求职者将会继续接受最后的面试，并接到录用通知书。如果团队并不100%地确定对方是否合适，那么就会拒绝。

曾经有一名求职者申请一个在置物柜安装团队工作的职位。这是一个很重要的职位，安装人员代表了产品和公司的形象。这名求职者在安装置物柜系统时表现得很出色，但是在向马克汇报工作的时候，团队对他有了一丝担忧。在那一天结束时，他们要把工具收起来，这名求职者却只是把工具扔进了盒子，盖上了盖子。在你我看来，这看起来不过是个轻微的过失，几乎不值得一提，更不至于让这一天原本毫无瑕疵的工作蒙上阴影。但是这个团队却认为，这种行为暴露了疏忽和草率的问

题，同他们期待中的那种合适人选是相违背的。马克听后表示赞同，然后很礼貌地告诉求职者，他并不是适应Vitsoe公司文化的最佳人选。马克和他的团队认为：

**If it isn't a
clear yes, then
it's a clear *no*.**

如果不是一个确定的 Yes，
那就是一个肯定的 No。

但是，他们强化这种高选择性的筛选程序还不仅仅是一种直觉反应（尽管那也非常重要）。一个看起来反复无常的决定，实际上经过了一个严格而且连续的探索过程，用来研究什么有效，什么无效。例如，他们得知，某人在孩提时代玩乐高玩具的热情与其对Vitsoe公司文化的适应性之间存在较高的关联度。这种关联并不是他们凭空建立起来的。他们在几年里作了各种各样的尝试，有些坚持下来了，有些放弃了。

这个团队在进行评估的时候，采用的也是明确的标准。他们的首要标准是：“这个人会是一个绝对合适的人选吗？”这就解释了为什么在他们设计的选择程序中包含了多次面试，也解释了他们开发这种工作日试工法的原因，以及他们发放调查问卷的初衷。像任何一个真正的精要主

义者一样，他们努力收集有针对性的信息，以便作出一个知情的、缜密的、深思熟虑的决定。

Box公司的首席执行官阿伦·李维（Aaron Levie）采用的招聘标准与之类似。他在乎的仅仅是，他愿不愿意每天与这个人一起共事。“我们考虑这个问题的方法之一就是”，他说，“这个人原本会不会成为这个团队的创始成员之一？”如果答案是肯定的，他就知道自己找到了合适的人选。

三条最低标准+三条极端标准

在决定自己应寻找什么样的机会时精挑细选只是问题的一个方面，当机会自己找上门来的时候，问题就变得更复杂了。得到了一份工作，但却不是自己想要的；来了一个项目，不是自己的主要业务但很容易来钱；有人让我们帮忙解决困难，事情我们乐意做，但是没有报酬；一个熟人有个分时共享资源，不过地点不太理想，但是有折扣。我们该怎么办？

对错过好事儿的担心会产生很大的影响。我们怎么能说“不”呢？机会就在那里，等着我们去拿。我们还可能曾经追求过这样的机会，但是现在要得到它简直易如反掌，而我们却瞻前顾后。但如果只是因为它容易得到而接受它，我们就会冒未来不得不拒绝另一个更有意义的机会的风险。

在建立一家综合广告公司的时候，南希·杜阿尔特（Nancy Duarte）就曾面临过这样的处境。2000年，这家公司业务涉猎广泛，从企业形象建立到印刷及网站开发，再到展示设计（一种所有设计公司都讨厌做的工作），什么都做。但是，由于没有一项专长能够独树一帜，公司就与其他任何一家设计公司没有什么区别。

后来，南希读到了吉姆·柯林斯的《从优秀到卓越》。在书中，柯林斯主张如果你只对一件事情充满激情，而且你又最擅长它，那么你就应该只做这件事情。就在那个时候，南希认识到了让公司独树一帜的真正机会在于，要做其他公司不愿意做的业务：展示设计。

专注于做别人不做的事情，使他们能够创造知识、工具和专业技能，让自己成为全球会展业务的领头羊。但是，要做到这一点，他们不得不放弃其他一切机会，就算是在经济不景气的时候，就算面对有偿工作，也必须如此。要独一无二，这就是代价。换句话说，他们必须更加挑剔地选择自己要承担的业务，以便把所有的能量集聚到自己的专长领域，把它越做越好。

当机会向你走来时，不妨用以下这种简单、系统的流程来执行精挑细选的标准。第一步，把机会记下来。第二步，写下3条“最低标准”，如果这个机会要进入考虑范围的话，就必须达到这3条标准。第三步，写下3条理想的或“极端标准”，如果这个机会要被考虑的话，也必须满足这3条标准。按照这个流程，一旦这个机会不能满足第一组标准，那么答案很显然就是**No**。但是如果它不满足3条极端标准中的2条，答案依然是**No**。

机会			
你面临的机会是什么?			
最低标准			
选项要被考虑,			
最低标准是什么?			
极端标准			
选项要获得首肯,			
理想的标准是什么?			

布鲁克林最美味的比萨

在做人生重大决定的时候，更严苛的标准能让我们更好地激发大脑中复杂的搜索引擎功能。想象一下，在Google搜索中查找“纽约的好餐馆”和查找“布鲁克林市区最美味的比萨”，结果会有什么不同。如果我们搜索的是“一个好的职业机会”，大脑中就会出现很多很多页的结果，让你来浏览和选择。相反，为什么不进行高级搜索，问3个问题：“我深深热爱的是什么？”“能发挥我才干的是什么？”“能满足世界上一个重要需求的是什么？”很自然，搜索结果就不会有那么多页了，然而这正是练习的目的所在。我们要找的不是大量可做的好事情，而是能让我们绝对地实现个人贡献峰值的那一件对的事情。

恩里克·萨拉（Enric Sala）就是用这种方法，找到了生命的召唤。在职业生涯早期，恩里克是美国加利福尼亚州拉荷亚著名的斯克里普斯海洋研究所（Scripps Institution of Oceanography）的教授。但是，他始终有一种挥之不去的感觉，那就是自己当时的职业路径只不过是接近于他真正应该做的事情。因此，他离开了学术研究领域，投身到《国家地

理》杂志的工作。这个成功的转型，使他在华盛顿州获得了新的、有趣的机会，他感受到自己进一步靠近了正确的职业路径，但依然没有踏上那条对的路。就像很多干劲十足而又野心勃勃的人一样，恩里克早期的成功分散了他的精力，使他无法对目标有清晰的认识。后来，他看了雅克·库斯多（Jacques Cousteau）率领著名的“卡里普索号”（Calypso）科考船进行航海探险的故事，从那一刻起，在全球最美的海洋中潜水开始成为他的梦想。几年后，当一个绝佳的机会出现时，他再次掉转了自己职业车轮的方向，去征战能够真正实现个人贡献峰值的领地：成为《国家地理》杂志的驻站探险家。这样他就能花大量时间在最偏远的地方进行潜水探险，同时还可以运用自己在科学和通信方面的优势对全球性政策施加影响。要获得一份梦想中的工作，其代价就是放弃很多不错的工作，甚至是非常棒的工作，以及与这些好工作相媲美的各种机遇，去等待那个他能够热情十足地说出“**YES**”的那个机会。而这种等待是值得的。

只有很少的人，能够从事自己热爱的，能发挥自己才干的，以及服务世界需要的工作，恩里克是这少数人中的一个。他的主要目标是帮助建立类似于国家公园的机构，来保护海洋中最后的原始区域——一个真正伟大的贡献。

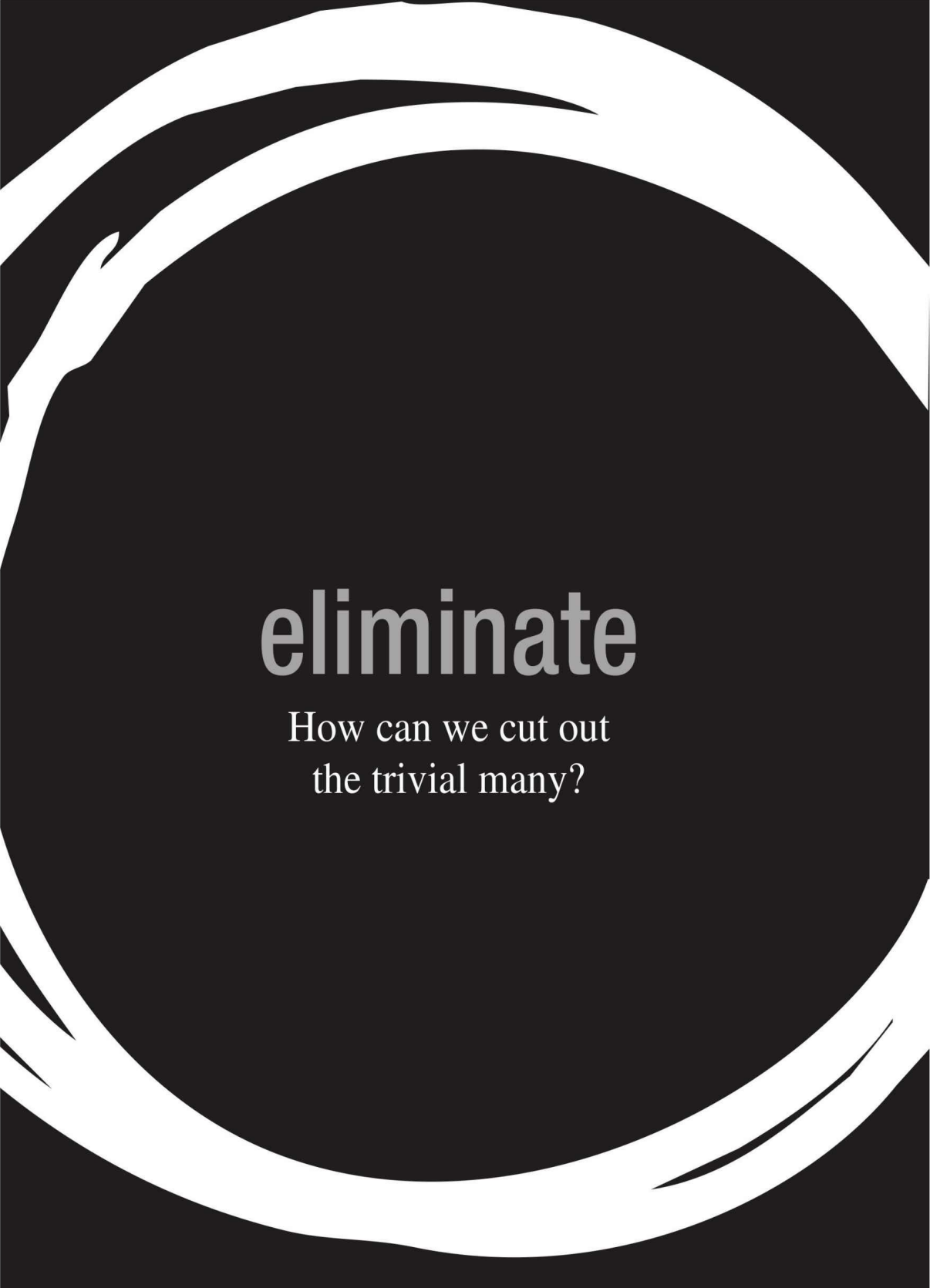


- 如果答案不是一个确定的Yes，就是一个肯定的No。
- 只接受前10%的机会。
- 制定一种既有选择性又明确的标准。
- 放弃“不错”，等待“最好”。

Part III

精要主义者养成的第二步：排除





eliminate

How can we cut out
the trivial many?

导读

摆脱无意义的多数

再回想一下本书前言部分谈及的那个关于衣柜的隐喻。到了这个时候，你已经仔细打量了挂在你衣柜里的所有东西，并且把衣服分成了“必须保留的”和“可能应该扔掉的”两堆。但是，你真的打算把那堆“可能应该扔掉的”衣服装进袋子里并送走吗？

换句话说，仅仅确定哪些行为和努力不可能带来个人贡献峰值是不够的，你还得积极地淘汰它们。该部分将告诉你，如何排除那些不重要的，这样你就能为真正重要的事情作出更大的贡献。不仅如此，你还将学到实现上述目标的方法，让你的同事、领导、客户和对手对你尊重有加。

扔掉旧衣服并非易事。毕竟，人们还有点恋恋不舍，也还会担心“假如”过了几年后悔不该扔掉那件带着大大的垫肩和花哨的条纹的上衣，那怎么办。这种感觉是正常的，研究表明，人们对于自己已经有的东西的估价，往往高于它们的实际价值，于是扔掉这些东西就更难。如果你还不舍得抛弃那件比喻中的上衣，就问问这个致命的问题：“如果没有它，我愿意花多少钱去买它？”同理，在人生中，在决定要淘汰哪些事情时，也问一个致命的问题：“如果没有这个机会，我愿意做什么来得到它？”

当然，要找到那条铁律，去拒绝那些出现在你工作和生活中的机会——往往还是非常好的机会，比扔掉衣柜里的旧衣服要困难得多。但是，你必须找到它，因为你要记住，任何时候你没有对不重要的事情说“NO”，就意味着你默认自己对它说了“Yes”。因此，一旦你已经对选

项进行了充分探索，你要问自己的问题就不是：“在那些具有竞争性的优先级事情中，我应该选择哪个？”而是要问这个本质性的问题：“我应该拒绝哪个？”这个问题才能揭示你真正应该优先做的事情，才能真正为你的团队打开最好的前行之路，才能显示你真实的意图并帮助你实现个人贡献峰值，不仅仅为你自己的目标，也为你团队的使命。这个问题，能够传递一种罕见而宝贵的明确性，而这种明确性是获得改变职业和生活中游戏规则突破点所不可或缺的。

CHAPTER 09

CLARIFY

One Decision That Makes a Thousand

要素六

澄清，做一个以一当千的决定

不休不止，向着一个目标前进，这便是成功的秘诀。

——安娜·巴甫洛娃



非精要主义者

愿景及使命陈述模糊、普通

拥有具体的季度目标，但不能激励和鼓舞成员为上新台阶而努力

设定了价值理念，但缺乏执行的指导原则



精要主义者

战略具体实在，鼓舞人心

目标意图具有重大意义且可以量化

做一个以一当千的决定

让我们先来做一个游戏。以下是3家公司的使命陈述。你试试能否为每家公司找到匹配的使命陈述。

公司	使命陈述
1. 爱科集团 农业设备的主要制造商和经销商，产品主要有：替换零件、拖拉机、牧草设备和各种机械等	A. 用一流的服务、持续的创新、优良的品质和诚信如一来创造利润
2. 多福集团 垃圾车等机械设备及喷墨式打印机、电路板组装等电子设备的制造商	B. 在服务客户和股东利益的每一个市场成为领头羊
3. 迪安食品公司 食品及饮料公司，主要生产牛奶、奶制品和豆制品	C. 公司首要目标是实现长期股东价值最大化，遵守经营所在地法律法规，始终秉持最高道德标准

你匹配对了吗？如果你觉得一头雾水、不知所措，那么你就和很多人同病相怜。面对千篇一律、毫无辨识度的陈述，大家都觉得这是不可能完成的任务。这类模糊、浮夸的任务陈述，现在依然在某些领域被视为“最佳作法”，但是很多情况下它们并不能像预期的那样奏效，它们无法激起员工清晰的目标意识。

本书的这个部分主要涉及如何排除非精要事务，以便把所有精力投入到最有意义的事情中去。我们要学会排除的第一种非精要事务的类型，就是任何与你的奋斗目标南辕北辙的事情。这听上去很简单，但是

要做到这点，你首先必须明确自己的目标是什么——这就是本章要解决的问题。

从“很清楚”到“真的清楚”

与我合作过的经理人，常常说自己公司的目标或战略“很清楚”，似乎这样就够了。然而，任何一个戴眼镜的人，都知道“很清楚”和“真的清楚”之间存在很大的不同。同样的问题也存在于个人的职业规划上。每当我问别人：“未来5年，你真正的职业目标是什么？”令我吃惊的是，直到现在，能回答这个问题的人依然寥寥无几。

要不是因为目标明确总是能够预示人们如何工作，这个问题也就不那么重要了。在同执行团队合作的过程中，我很讶异，如果这个团队仅仅是“有点明确”而不是“真的明确”该干什么，会发生什么。

首先，人的动力会大打折扣。事实证明，如果没有目标，团队的积极性和合作效果都会打折。你可以花大力气通过培训来改善管理者在人际沟通、团队合作和进行360度反馈报告方面的能力，但那是徒劳无功的。如果团队缺乏明确的目标及角色定位，问题就会越来越严重，越来越多。

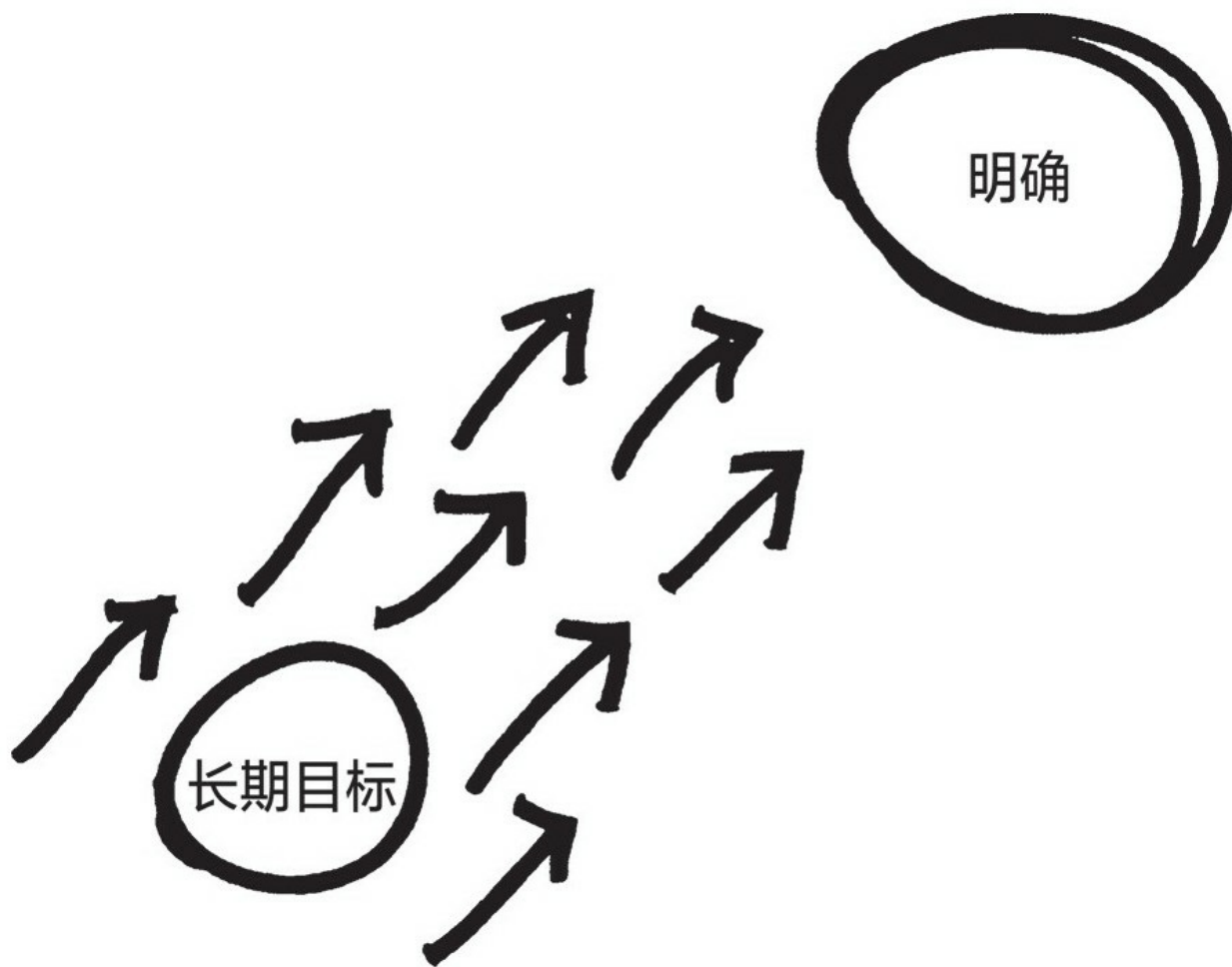
这不仅仅是我的理论，或者是我在某本商务类书籍中读到过的内容。通过收集500多人在100多个团队中的经历，我发现了一个不争的事实：人们一旦对团队目标及自身角色定位缺乏明确的认识，就会感到困惑不解、压力重重、灰心失意；相反，当他们有较高程度的明确性时，就会积极振作起来。

没有明确的目标，人们往往会在无意义的事情上虚掷光阴，耗费精力。当他们足够明确自己的目标时，就有能力在真正重要的领域实现更大的突破和创新，这种成就甚至连他们自己都不曾认识到。在我的工作

中，我发现当团队缺乏明确的目标时，会出现两种典型的模式。

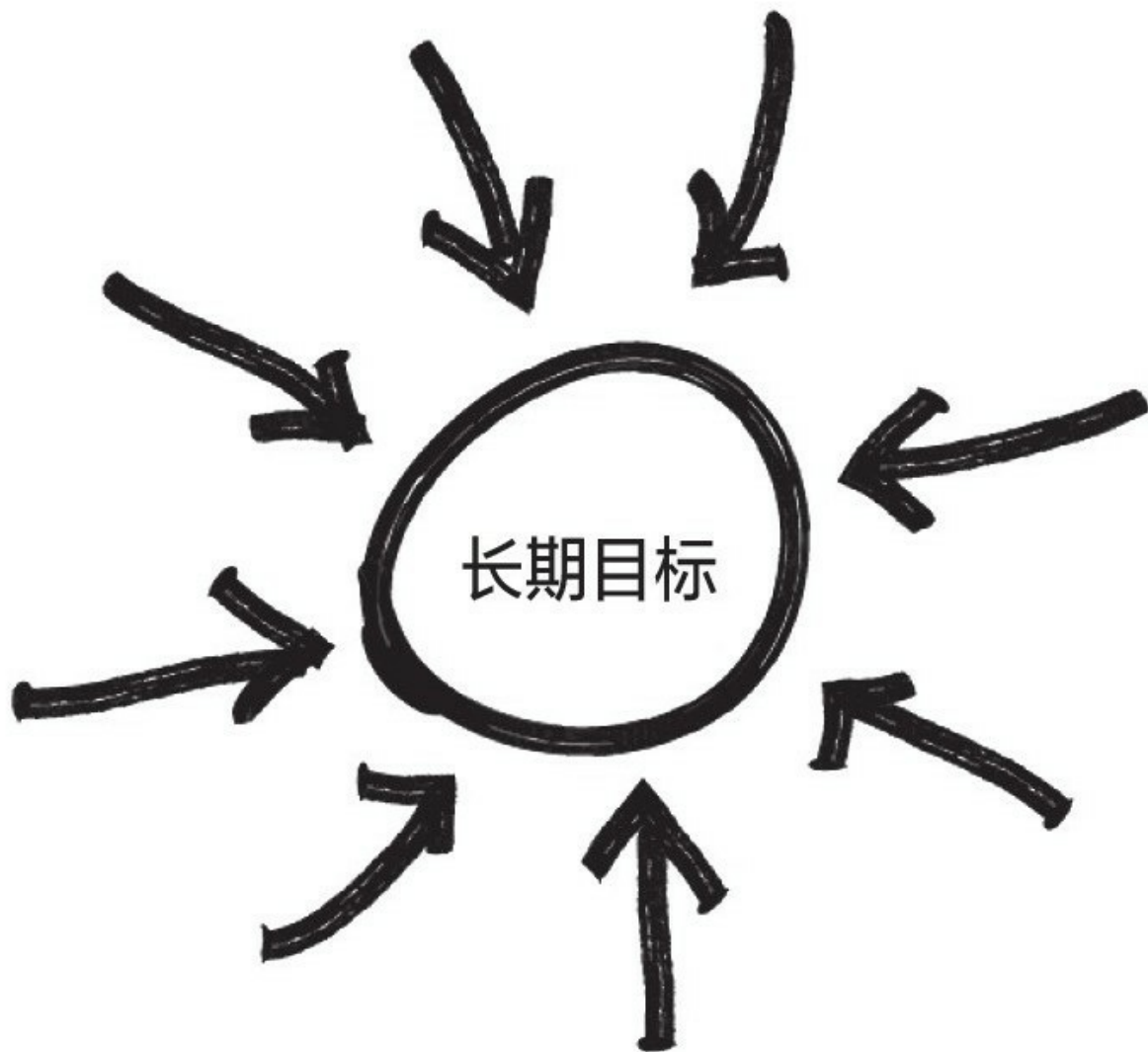
模式1：设计自己的游戏，制定自己的规则

在这种模式中，团队极度热衷于赢得老板的青睐。这种模式的问题在于：当人们不知道游戏最终会变成怎样时，就会迷惘于如何去赢得老板的青睐。结果，他们各自设计自己的游戏，制定自己的规则，竞相争宠。他们不再集中自己的时间和精力来作出更大的贡献，而是机关算尽施展计谋，比如处心积虑把竞争者比下去，妄自尊大，对老板所说的一切随声附和、言听计从。这些行为非但毫无意义，而且还具有破坏性，会产生适得其反的效果。

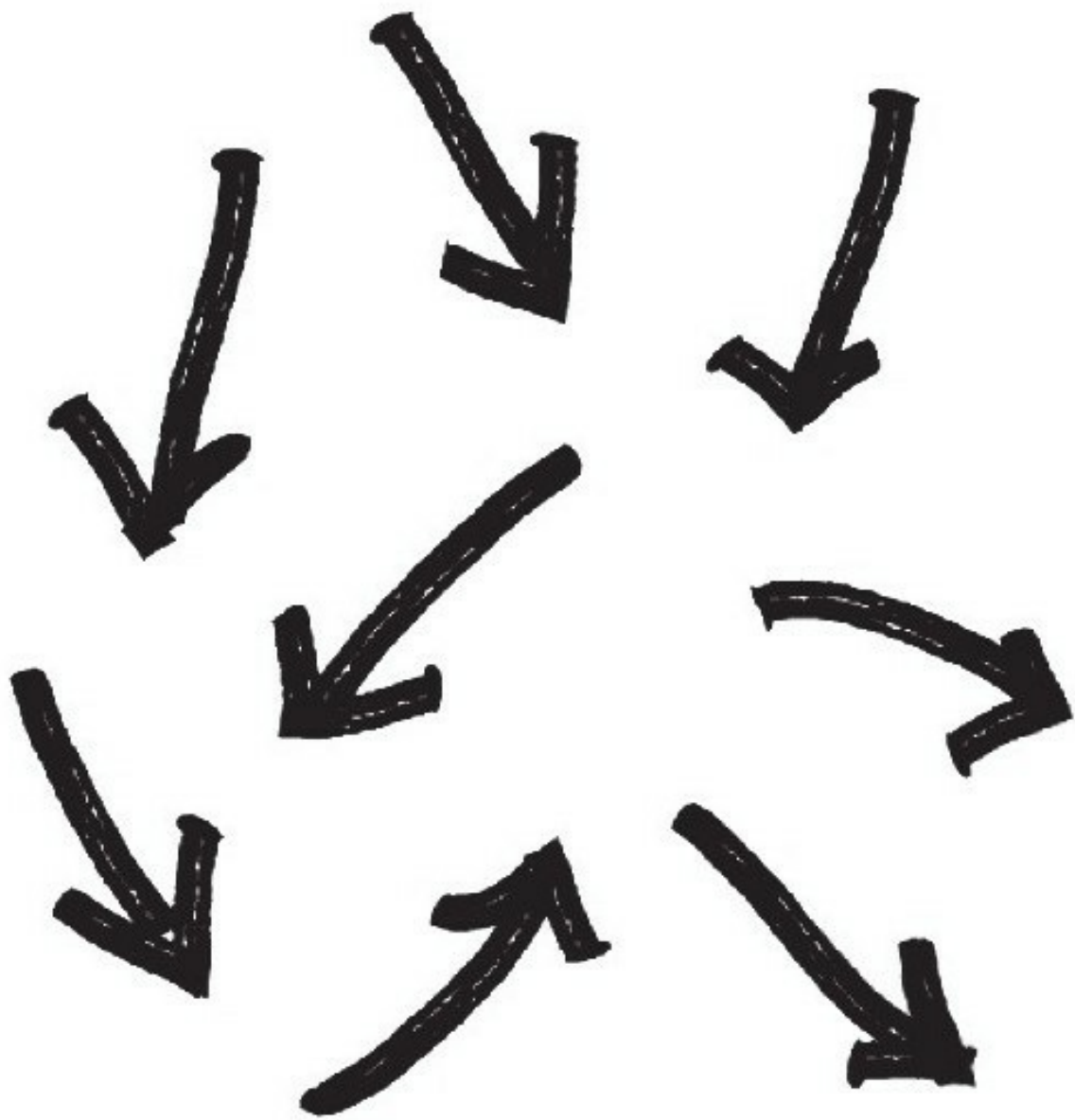


在个人生活中，我们也在做着类似的事情。一旦我们不清楚人生的目标究竟何在，换句话说，我们对于目标、追求和价值缺乏明确意识时，就会设计自己的社交游戏。我们浪费时间和精力来力图把别人比下去。我们高估了不重要的东西，比如把豪车美宅、Twitter上的粉丝数量或者自己上传到Facebook上的照片等虚无缥缈的东西看得过重。结果，忽略了真正重要的事情，比如花时间陪伴心爱的人，滋养自己的心灵，关心自己的健康，等等。

模式2：假装一切都好



在这种模式里，缺乏目标的团队变得群龙无首。因为方向不明，人们追求的净是些短期利益，丝毫不顾及自己的行为对于团队整体的长期目标是否有所贡献（或者在某些情况下，是否会产生破坏力）。有时候，这样的行为确实具有良好的动机，有些就个人而言，甚至还具有重要意义。但是，当人们在团队中工作时，多个不同任务产生的分歧和差异，却无益于团队整体目标的达成。这种团队每向前一步，就会后退五步。



与此类似，当一个人参与了太多不同的任务——哪怕是好任务时，也可能导致无法实现最重要的使命。其中的一个原因是，多种行为无法产生合力，因此不能对整体的意义有所增益。比如说，学习5种不同的专业，每一种都很好，但也不能等同于一个学位。同理，在5个不同行业里的5种不同的工作，并不能保证职业的向前发展。方向不明，缺乏目标，什么好就做什么，并不足以实现贡献峰值。正如拉尔夫·瓦尔多·

爱默生（Ralph Waldo Emerson）所说：“那败坏了人类和国家的罪恶就在于那种工作，它时不时地为了一个权宜的目的，而偏离了主要轨道。”

相反，如果团队有着真正明确的目标，每个成员了解自己的角色定位，团队的动力就会出现神奇的变化。动力增长，蓄势待发，团队整体的贡献就会日积月累，不断攀高。

那么，如何让团队和个人的努力具有明确的目标呢？方法之一，就是确定精要目标是什么。

精要目标

要了解什么是精要目标，最好的方法也许是先确定什么不是精要目标。我们可以运用一个2×2的矩阵来探究其中的含义。



在左上角象限里，我们可以归入“我们要改变世界”这样的愿景和使命，这种陈述听上去鼓舞人心，但实际上普通得让人留不下任何印象。在左下角象限里，我们可以找到一大串模糊、普通的价值理念，比如“创新”“领导力”“团队精神”，都是典型的平淡无奇、千篇一律，丝毫

不能引发任何激情。在右下角象限里，我们可以设定一些必须关注的短期季度目标，比如“利润在去年的基础上提升5%”，这些短期策略因为足够具体，就会引起关注，但往往缺乏鼓舞力量。

精要目标，却是既鼓舞人心又明确具体，既充满意义又容易衡量。如果操作正确，一个精要目标是一个可以以一当千的决定。就像你决定自己要当医生，而不是律师一样。一个战略选择可以排除其他所有选项，为你未来5年、10年甚至20年的生活规划路线。一旦重大决定已经做出，所有随之而来的其他决定就能更好地聚焦。

玛莎·莱恩·福克斯（Martha Lane Fox）受英国首相之命担任英国首位“数字化倡导者”（Digital Champion）⁹时，获得了机会对这个新角色进行描述。你可以想象玛莎可能会尝试的所有模糊不清、枯燥或者晦涩难懂的解释方法，如同呆伯特连载漫画（Dilbert）一般呼之欲出。

然而，玛莎和她的团队却提出了这样一个精要目标：“到2012年末，保证让每一个英国人都能上网。”多么简洁、具体、鼓舞人心而又易于衡量！它让每个团队成员都确切地知道自己要干的究竟是什么，以便能协调行动和精力，来精简掉其余所有的无关事务。这让每个团队成员，不管资历深浅，都有权推却并表示：“但这个新想法真的有助于实现我们的目标吗？”而且，它也让他们能够更好地赢得合作者的支持，从而大大加速目标实现的进程。纵然他们尚未实现自己的全部抱负，明确目标也让他们得以创造出比其他任何情况下都大得多的贡献。

这就是我们的公司、团队和职业生涯所需要的目标陈述。那么，我们怎样才能制定出一个既具体可行又鼓舞人心，既充满意义又容易衡量的目标陈述呢？

制定精要目标的两大要点

1.重在实质，而不是形式

每当制定目标陈述的时候，不管是为自己的公司、团队，或者是为自己，人们总是念念不忘那些无关痛痒的形式上的细枝末节，比如：“我们应该使用这个词，还是那个词？”于是，套用陈词滥调和时髦词汇就在所难免了，其结果就是我在本章开头时所引用的那些模糊不清、毫无意义的陈述。一个精要的陈述并不需要精雕细琢，它重在实质，而不是形式。问一个更加重要的问题，它将在你未来作出每一个决定时提醒你：“如果我们只能在一件事情上做到出类拔萃，那将会是一件什么事情？”

2.具体的目标才拥有鼓舞人心的力量

这就是说，要实现目标明确，启发鼓舞更加有用。一说到鼓舞和启发，我们常常会想到那些花言巧语和高谈阔论。但是，尽管这些虚华的辞藻的确有些鼓舞人心的作用，但我们需要记住的是，具体的目标不仅能够提振士气，还能提升境界。一个充满力量的精要目标能够鼓舞士气，这在一定程度上是因为它具体得足以回答“我们如何知道自己已经成功了”这个问题。

比尔·米汉教授（Bill Meehan）向我充分阐释了这其中的道理。比尔教授为麦肯锡公司的首席执行官和高管提供战略咨询服务长达30年，现任教于斯坦福商学院，讲授“非营利机构战略管理”课程。我在研究生期间参与了这门课程的学习，当时他给大家布置的作业之一，就是对一些非营利机构的愿景及使命陈述进行评价。

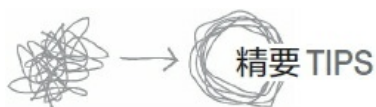
该课程学员们查看了100多个案例，发现其中最浮夸的那些陈述恰恰是最没有启发性和鼓舞性的。例如，其中一个组织的使命是“消灭世界上的饥饿”，但是考虑到这个组织只有5名成员，这个使命就是空话一

句。在这一大堆大而无当的理想主义表述丛中，有一个陈述立刻被我们所捕获，并使我们从中大受启发。它的来源略微出人意料：演员/社会企业家布拉德·皮特（Brad Pitt）惊讶于卡特里娜飓风后新奥尔良市重建工作进展之慢，创建了一个名为“让它重新归位”的机构，核心目标是“为居住在下九区的家庭建造150所价廉环保且能抵御暴风雨的房屋”。这个陈述摒弃了浮夸和空洞，目标明确性使它真实可信，而这种真实性又使它具有鼓舞人心的力量。“我们如何知道自己已经成功了？”答案就在这里。

真正明确，才能真正优秀

精要目标的适用性远远不局限于职位描述或者公司使命陈述，一个真正意义上的精要目标能够指导你获得更明确的目标意识，帮助你绘制生活的路径。例如，南非前总统纳尔逊·曼德拉（Nelson Mandela）在长达27年的监禁生活中成长为了一名精要主义者。1962年入狱时，他的一切几乎都被剥夺殆尽：家庭、名誉、骄傲，当然还有自由，均不复存在。他选择用这27年专注于真正重要之事，摒弃其余一切——包括他自己的愤恨在内。他把消除南非的种族歧视当做自己的重要使命，在这个过程中，他创造了一笔留存至今的遗产。

设定一个精要目标并非易事。它需要勇气、卓识和远见，来看清什么样的行为和努力能够提升个人贡献峰值。它需要我们面对尖锐的诘问，作出真正的取舍，并且执行严格的纪律约束，以割舍掉那些极具竞争力却会诱使我们偏离真正目标的重要事务。然而，为精要目标作出这些努力是值得的，因为只有当目标真正明确时，人员、团队和组织才能充分调动起来，并达到真正的优秀。



- “很清楚”不够，“真的清楚”才行。
- 精要目标既具体可行，又鼓舞人心；既充满意义，又容易衡量。
- 只有目标真正明确，才能达到真正优秀。

CHAPTER 10

DARE

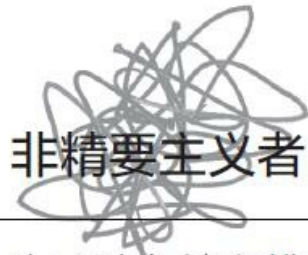
The Power of a Graceful “No”

要素七

勇气，优雅说“不”的力量

勇气，就是在压力下展现优雅。

——欧内斯特·海明威



非精要主义者

为了避免社交尴尬和压力而
不说“不”

应承一切事情



精要主义者

敢于坚定地，决绝地，优
雅地说“不”

只答应真正的要事

在合适的时机正确地说“不”能够改变历史的进程。

类似的例子不胜枚举，其中一个便是在深受种族隔离之苦的蒙哥马利市一辆公共巴士上，罗莎·帕克斯（Rosa Parks）沉静而又坚决地拒绝让出座位，这几乎在顷刻之间集结起了推进民权运动的力量。帕克斯回忆说：“当（公共汽车司机）看到我还坐在座位上时，他问我是否准备站起来，我回答说，‘不，我不准备让座’。”

与普通观念不同的是，她勇敢地说“不”，并不是出于特别的自信或一般意义上的个性。实际上，当她被选为全美有色人种进步协会蒙哥马利分会主席秘书时，她解释说：“在那里，我是唯一的女性，他们需要一个秘书，我胆小得没敢说‘不’。”

她在公共汽车上作出的决定，更多的是出于她对自己在此时此刻刻意要做出的选择充满了坚定的信念。当司机命令她离开座位时，她说：“我感到有一种决绝如同冬日晚上的被子一样，盖过了我的全身。”她没有意识到，自己的一个决定会激起一场波及全世界的运动。但是，她知道自己需要的是什麼。她知道，即便是在被捕的时候，“这是我最后一次在这样的屈辱中乘车”。为了告别屈辱，哪怕冒着监禁的风险，也在所不惜。是的，于她而言，这就是头等大事。

诚然，我们都希望自己不会面临罗莎·帕克斯那样的处境，但我们依然可以从中获得启发。需要勇气说“不”的时候，想一想罗莎·帕克斯。记住她那种信念的力量，面对社会压力，坚定自己的立场，不向无意义屈服。

你可曾感觉到，自己认为对的和别人迫使你去做，二者之间剑拔弩张？你可曾感觉到，你内在的信念和外在的行动势不两立？你是否曾违心地答应，仅仅是为了避免冲突和摩擦？你可曾感觉到，在拒绝老

板、同事、朋友、邻居或家庭成员的邀请或要求时，自己是多么的恐惧和胆怯，因为你怕他们会失望？如果你的答案是肯定的，那么你并不孤独。带着勇气和优雅来驾驭这样的时刻，是成为精要主义者最重要也是最难掌握的技能之一。

我并不打算专门就勇气写一个章节。但是，随着我对精要主义理解的愈发深入，我更加清楚地认识到了，勇气是排除冗余和无益之事的關鍵所在。没有勇气，“自律地追求更少”不过是挂在嘴上的空话，是茶余饭后的又一种谈资，其深度不会超过皮肤的厚度。任何人都可以夸夸其谈地说，专注于最要紧的事情是多么重要，而且很多人确实这样说过，但真正敢于这样做的人却寥寥无几。

我说这些并没有带任何批判色彩。害怕说“不”，我们对此有着充足的理由。我们担心自己会失去一个重大的机会。我们害怕破坏现状、制造事端、过河拆桥。一想到会让自己敬重、喜欢的人失望，我们就无法忍受。但这些都不足以让我们成为坏人。这是人性的一部分。然而，尽管拒绝别人是如此艰难，但如果不这样做，我们就会错过重要得多的东西。

有一位叫辛西娅的女士，给我讲述过一个她父亲曾经计划带她夜游旧金山的故事。当时12岁的辛西娅和她父亲对这个计划中的“约会”期盼了好几个月。他俩把整个旅程计划得细致入微：辛西娅将参加父亲最后一个小时的讲座，4点半在会议室后面集合，在所有人都涌上来和他说话之前快速离开。他们将搭上一辆有轨电车去唐人街，吃他们最喜欢的中国菜，买一个纪念品，在那里欣赏一会儿风景，然后像她父亲喜欢说的那样去“看一场电影”。接下来，他们将坐出租车回酒店，跳入泳池快速地游会儿泳（她父亲在泳池关闭后偷偷潜入的水平堪称一绝），从客房服务那里点一份淋着热浓糖浆的圣代，看一场午夜秀。他们把所有细节讨论了一遍又一遍。那种期待本身也是整个经历的一个组成部分。

一切都像计划的那样进展着，直到他父亲正要离开会议中心的时候，偶然遇见了一个老同事兼商业伙伴。他们两人已经数年未见，辛西娅看到了他们那么热情地拥抱对方。那位朋友说：“看到你正在和我们的公司合作，我真是太高兴了。当我和洛伊斯得知此事时，我们觉得这简直就是完美。我们想邀请你，当然还有辛西娅，去渔人码头好好享受一顿海鲜晚餐。”辛西娅的父亲回答说：“鲍勃，看到你真是太高兴了。去渔人码头吃晚餐太棒了！”

辛西娅却垂头丧气，灰心不已。她朝思暮想的观光游览和冰激凌圣代转眼就泡汤了。还有，她不喜欢吃海鲜，她都能想象得到整夜听着大人们之间不停说话会是多么无聊。但是，他父亲紧接着又说：“但是今晚不行。辛西娅和我有一个特别的约会计划，不是吗？”他朝辛西娅使了个眼色，并拉起她的手，然后一起出了门，去继续那个难忘的旧金山之夜。

无巧不成书，辛西娅的父亲就是管理思想家史蒂芬·柯维（Stephen R. Covey），就在辛西娅给我讲这个故事的前几个星期，他刚刚离世。因此，在回忆起那个旧金山之夜的时候，辛西娅感情深挚。他那个简单的决定“使得他和我永远骨肉相连，因为我知道对于他来说，最重要的就是我”！她这样说道。

作为同时代中最受欢迎、读者最多的商业思想家之一，史蒂芬·柯维是一名精要主义者。这不仅仅是由于他经常向全世界的重要领导人和国家元首传授精要主义原则——比如“头等大事就是要把头等重要事情当作头等大事”，他自己也是身体力行，亲身实践。而在那个时刻，当他把与女儿共度欢乐时光当作头等大事的时候，他留下的快乐回忆直到他死后还是令他女儿久久难以忘怀。从某种角度来看，他做出那样的决定似乎平淡无奇。但很多人若处在他当时的位置上时，可能会接受朋友的邀请，因为他们担心拒绝会显得鲁莽无礼和不知好歹，或者会错失一个与

老朋友共享晚餐的良机。那么，为何在当下，大胆选择重要的事情，放弃不重要的事情，是如此艰难？

一个简单的答案就是，我们分不清什么才是重要的事情。当事情发生时，我们变得毫无防御能力。而与之相反的是，当我们对什么是要事确定无疑时，我们就如同拥有了一个力场，保护自己免受来自四面八方的非精要事务的侵袭。

罗莎·帕克斯拥有深刻而清晰的道德感，这给了她非同寻常的坚定信念。史蒂芬·柯维拥有明确的期盼，这让他和自己心爱的女儿共度了难忘的快乐之夜。几乎在每一种境遇下，对何事重要的确定性都能为我们补给力量，来对非精要事务说“不”。

因为说“不”，所以高效

在当下选择重要之事为何如此艰难，这其中的第二个原因其实再简单不过，无非就是我们内心对社交尴尬畏首畏尾。实际上，作为人，我们都对与同类交往乐此不疲。毕竟，几千年前，当人类过着以采集、狩猎为主的群居生活时，就依赖于同别人的联结。然而，我们对群体对我们期望的服从，心理学家称之为规范性从众，虽然如今这种服从已经不再是生死攸关之事，但那种渴求却始终深深地根植于我们的头脑中。这就是为什么，不管是一个老朋友邀请你去吃饭，或是你的老板让你负责一项重要的、备受瞩目的项目，抑或是一个邻居拜托你帮忙参加家长会的糕点义卖活动时，哪怕是一个拒绝的念头都会给人带来身体上不舒服的感觉。我们觉得内疚，我们不想让别人失望，我们害怕破坏关系。但是，这些情绪会让明确性变得混乱不堪，会分散人的注意力，导致人们认识不到事实就是要么拒绝，然后遗憾几分钟；要么答应，换来几天、几周、几个月，甚至几年的后悔。

要摆脱这样的困境，唯一的出路就是学会说“不”，坚定地，决绝

地，却很优雅地说“不”。因为一旦我们这样做，就会发现自己对打击别人和惹怒别人的担心其实是被夸大的，而且还会发现人们实际上会更加尊重我们。自从成为精要主义者以来，我发现人们普遍都尊重和羡慕那些有着坚定信念勇于说“不”的人。

彼得·德鲁克是我眼中的现代管理学之父，他就是一位精通优雅拒绝这门艺术的大师。以研究“心流”著称的匈牙利教授米哈里·希斯赞特米哈伊，曾为了写作《创造力：心流与创新心理学》（*Creativity: Flow and the Psychology of Discovery and Invention*）¹⁰采访过一批创新达人。德鲁克的回答引起了希斯赞特米哈伊的浓厚兴趣，因此他进行了逐字引用：“您2月14日的来信让我深感荣幸并受宠若惊，因为多年来我一直钦佩您和您的工作，并从中学会了很多。但是，亲爱的希斯赞特米哈伊教授，我恐怕不得不让您失望了。我可能无法回答您的问题。您说我是富有创造力的人，但我不知道那是什么意思.....我只不过是坚持埋头苦干而已.....如果我说富有成效的秘密之一就是准备一个很大的废纸篓来装所有像您发来的那种邀请函，希望您不要觉得我很冒失或者无礼。根据我的经验，提高生产效率的形式是不要做任何帮助别人的工作，而是把自己全部的时间花在上帝安排我们去做的事情上，并且把它做好。”

How many times have you said no today?

今天，你说了几次“不”？

作为一个精要主义者，彼得·德鲁克深信“因为说不，所以高效”。

非精要主义者凡事都答应，是因为感觉到了那种社交尴尬和压力。他们常常为了追求取悦别人之后那种兴冲冲的感觉而自动说“Yes”，根本不假思索。但是，精要主义者知道，在这种兴冲冲的感觉过去之后，随即而来的就是懊悔的折磨。他们知道，如果这样做，自己很快就会感觉到受了欺负，因而愤恨不满——对别人，也对自己。最后，他们会清醒地认识到自己将面临一种不愉快的现实：为了应许新的承诺，自己现在不得不牺牲更重要的东西。当然，关键不在于对所有要求都全盘拒绝，而是在于拒绝不重要的事情，以便应承真正重要的事情。对一切说“不”，频繁地说，优雅地说，只有千真万确的要事除外。

优雅说“不”的六大原则

我们怎样才能学会优雅地说“不”呢？下面就是一些基本的原则和具体的方案，教你怎样来优雅地说“不”。

1.把决定和关系分开来

当别人让我们去做什么的时候，我们往往会把这个要求和自己同他们的关系搅和在一起。有时候，它们看上去是如此的相互关联，以至于让我们忘记了拒绝这个要求并不等同于否定这个人。只有当人们把决定和关系分开来的时候，才能做出一个明确的决定，然后另行找到勇气和同情来传达它。

2.优雅地说“不”并不代表必须使用“**No**”

精要主义者更多地选择“不”，而不是说“不”。可能在某个时候，说“不”最优雅的方式就是直言不讳地说“不”。但是，“您能想到我，真让我受宠若惊，但我恐怕爱莫能助”，或者是“我非常愿意，可是我已经力不从心了”，这样清楚并且礼貌地拒绝别人、却不使用“不”字的方法很多很多。在本章的后面部分，你将找到更多例子，帮助你优雅地说“不”。

3.关注权衡取舍

在答应的时候，对自己将要牺牲什么考虑得越多，拒绝就会变得越容易。不能清晰地认识到机会成本——也就是我们所放弃的东西的价值，那么我们就很容易跌进非精要主义的陷阱，自欺欺人地告诉自己我们可以搞定一切。其实，我们不能。一个优雅的“不”，来自于对权衡取舍一清二楚却又不会言明的算计。

4.自我提醒，每个人都是推销员

这并不是说你必须对别人疑心重重。我也不是要暗示大家，说人是不值得信任的。我只是说，每个人都在販售一些东西——一种想法，一

种观点，一种主张，来换取你的时间。清楚地知道被贩售的是什么，有助于我们更加谨慎地决定是否去购买它。

5.说“不”需要用人气换尊重

说“不”通常会在短时间内对关系产生影响。归根到底，当有人提出要求，却没有得到满足时，他或她的第一反应很可能就是烦恼、失望，甚至气愤。不利之处是一看便知的，而潜在的有利之处却不那么显而易见：当最开始时的烦恼或失望、气愤渐渐褪去时，尊重便会显露出来。一旦我们有效地推回别人的请求，就等于告诉他们，我们的时间是非常宝贵的。这便是行家里手和初出茅庐者之间的分水岭。

在这方面，一个恰当的例子就是平面设计师保罗·兰德（Paul Rand）有勇气对史蒂夫·乔布斯（Steve Jobs）说“不”。当时乔布斯正在为NeXT公司物色一个商标，他让曾为IBM、UPS、安然公司（Enron Corporation）、西屋电气公司（Westinghouse Electric Corporation）、美国广播公司（American Broadcasting Corporation, ABC）设计过商标的兰德提供一些设计方案。但是，兰德并不想提供“一些方案”。他只想作出一种设计。于是，他便说：“不，我只为你提供解决问题的方案，而且你得给我支付报酬。当然，你没有必要一定采用这种方案。如果你需要多种方案，那就另请高明吧。但是，我将会拿出我最好的设计方案来。用或者不用，都由你来决定。”结果毫无意外，兰德解决了问题，设计出了乔布斯想要的“宝石”标志，但是这个事件中真正值得关注的是兰德的“推脱”在乔布斯身上所起的作用。在后来谈及兰德时，乔布斯这样说道：“他全面透彻地考虑了客户和像他那样的专家之间的所有正式关系，从这个意义上说，他是我所合作过的人当中，最具专业水准的行家之一。”兰德拒绝乔布斯时，无疑是冒着风险的。他用短期内的人气损失换来了长期的尊重，这是物有所值的。

精要主义者接受自己不能始终都让所有人喜欢的现实。诚然，怀着敬意，合情合理并且优雅地说“不”，会在短期内带来社交成本。但是，精要人生的题中之义，就是认识到从长远来看，受人尊重比被人喜欢重要得多。

6. 一个明确的“不”比一个模糊的“是”更优雅

任何一个曾处于表态接收端的人都知道，相较于不答复，或是用一些诸如“我将努力解决问题”或明知自己不行却还说“我可能行”这样不作许诺的回答来蓄意误导别人，清楚地说“我不干”就有价值多了。模棱两可不等于优雅，延迟说那个最终还是要说的“不”，只会让它更难说出口——它会引起对方更深的怨恨。

有效说“不”的八种方法

请记住，精要主义者并不是偶尔才说“不”。这是他们的常备技巧之一。经常性地、优雅地说“不”，那么，面对别人的要求时，你就有各式各样的回应方法。以下八种应对方法，你可以放入自己的说“不”技巧大全中。

1. 尴尬的停顿

不要因为害怕尴尬的沉默而束手束脚，主动拥抱它。把它当做工具。当你面对要求的时候（很显然这只在面对面时有用），停顿一会儿。先数到三，然后再作出自己的判断。或者，如果你更勇敢一些的话，就等着别人来填补空白。

2. 委婉地说“不”或“不，但是”

我最近收到一封电子邮件邀请我去喝咖啡。我答复说：“我目前正忙于写书。但是，我想一写完书就和您聚一聚。请告诉我，在夏季快要结束的时候是否方便聚会。”电子邮件也是练习说“不，但是”的好办法，因为它能给你机会，来反复斟酌修改用词，使你尽可能优雅地说出那个“不”字。而且，很多人发现，电子邮件的空间距离能减轻人们对尴尬的担忧。

3.“我需要先看一下日程表，然后再答复您。”

我认识一个领导人，她觉得自己整天的时间都被别人操纵了。她是一个典型的非精要主义者，虽然聪明能干，但却不会说“不”，结果很快就成了一个“万事帮”类型的人。人们常来找她，问她：“你能不能帮我完成X项目？”为了当个好公民，她答应了。但是，这些五花八门的工作事项很快就让她感到不堪重负。后来，她学会了用一种新的措辞来答复请求：“我需要先看一下日程表，然后再答复您。”这才让她的境况得到了改观。这种措辞让她有时间先停顿一下，好好想一想，最后答复别人：很遗憾，我无法提供帮助。这种措辞也让她能够重新获得对自己决定的控制权，而不是别人一提出要求，就仓促答应。

4.使用电子邮件的回复功能

当收件人外出或者不在办公室的时候，发件人会很自然地希望获得邮件自动回复。实际上，这是接受度最高的一种说“不”的方式。使用这种自动回复的人，没有说他们不想回复你的电子邮件，而是说在这一段时间里，他们无法答复你。那么，为什么仅仅在节假日时才用这种功能呢？在我写这本书的时候，我就设定了主题为“处于修行模式中”的电子邮件自动回复。这封回复邮件这样写道：“亲爱的朋友，我目前正在写一本新书，这极大地占用了我的时间。很遗憾，我无法以我喜欢的方式来回复你。我为此深表歉意。——格雷戈。”你猜结果怎么样？大家似

乎适应了我暂时的不存在，也不介意我不回应的行为。

5.“好的。那么我要取消的是什么？”

在很多人看来，对单位的高层领导说“不”，这几乎是无法想象的，甚至是可笑的。然而，如果说“YES”，会导致你无法集中精力完成自己的工作，那么此时，说“不”也是你的职责所在。在这种情况下，说“不”不仅是合情合理的，也是至关重要的。为达到这个目的，一种有效的方法是提醒你的上级，如果你答应做这件事，那么你将无法做其他事，就让他们来解决这个取舍问题吧。

比如，经理来找你，问你能不能做X事项，你可以答复说：“好的，我很乐意优先完成这项工作。但是，为了集中精力在这个新项目上，我应该把哪些项目先放一放？”或者说：“我很想把这个工作做好，但是考虑到我还有其他的任务，如果我再承担这一项，我将无法把工作做得好到令自己满意。”

我认识一位领导，他的一名下属就是这样答复他的。他一点儿也不想干扰这位工作高效、做事有条理的员工，于是他收回了这项并不太重要的任务，派给了一个做事不那么有条理的员工！

6.带着幽默感说“不”

最近有一位朋友邀请我一起参加一项马拉松训练。我的回答很简单：“没门儿！”他笑了笑说：“啊，你可真是言出必行呀。”由此可见，精要主义者这个名声多么有用！

7.“欢迎你X。我愿意Y。”

比如，“欢迎你用我的车。我愿意保证你一定能拿到钥匙。”这样一

说，也委婉表达了“我不能开车带你”的意思。你要说的是你不想做什么，但你用来表达这个意图的语言却是你愿意做什么。这是一个特别好的方法，用以驾驭一个你愿意稍稍支持一下但又不能全力以赴的请求。

我非常喜欢这种结构的答复，因为它同时表达了对他人以及对你自己的选择能力的尊重。它让双方都注意到了自己可以做些什么选择。

8.“这个我做不了，但是X可能对此感兴趣。”

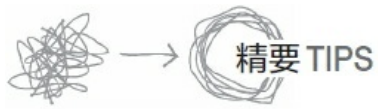
很多人都倾向于认为自己的帮助具有无可比拟的价值，但是人们在提出要求的时候，通常并不在乎帮助他们的究竟是不是我们，只要有人帮忙就行。

安·泰勒及洛夫特女装零售公司（Ann Taylor and LOFT）的首席执行官凯·克里尔（Kay Krill）过去常常对拒绝社交邀请感到头疼。结果，她最后还是参加了很多毫无兴趣的社交活动。很多办公室聚会，她发现自己一到那儿就后悔不该来。

后来有一天，她的一位良师益友告诉她，应该把那些无关紧要的人和事驱逐出自己的生活，这样就可以把所有的精力投入到真正于己有益的事情上去。这个建议让她重获新生。现在，她已经很善于挑选取舍了。熟能生巧，礼貌地拒绝别人于她来说已是易如反掌。凯解释说：“说‘不’对我来说一点都不难，因为我知道什么对我是重要的。我唯一希望的是，我要是能早点学会拒绝就好了。”

说“不”本身就是一种领导能力，而不仅仅是一种外在的技能。对于任何一种能力，我们起初都是经验不足。对于说“不”，我们也是新手。然后，我们学会了一些基本技巧，会出错，并从中学习，然后发展更多技能，并不断实践。不久，我们拥有了一个本领大全供自己使用，最后我们就掌握了一门社交艺术。我们可以优雅、体面地处理几乎来自所有

人的所有请求。海德思哲国际咨询公司（Heidrick&Struggles）首席执行官汤姆·弗里尔（Tom Friel）曾对我说：“我们需要学会慢点说‘好’，快点说‘不’。”



- 把决定和关系分开。
- 优雅地说“不”并不代表必须使用“No”。
- 关注权衡取舍。
- 自我提醒，每个人都是销售员。
- 说“不”是用人气换尊重。
- 一个明确的“不”比一个模糊的“是”更优雅。

CHAPTER 11

UNCOMMIT

Win Big by Cutting Your Losses

要素八

放弃，止损就是最大的赢

人生中一半的麻烦源于答应得太快，
拒绝得太慢。

——乔希·比林斯



非精要主义者

问：“为什么现在要停止这个项目，我已经在这上面投入了这么多？”

认为：“如果我坚持努力，就能取得成功。”

不愿承认错误



精要主义者

问：“如果不是之前已经有投资，我现在愿意在这个项目上投入多少？”

认为：“如果我现在就此罢休，这些时间和金钱可以用在何处？”

愿意脱手止损

不管如何评估，协和式超音速喷射客机都是航空工程的一大惊人创举。坐这种客机，从伦敦飞到纽约只要2小时52分钟59秒，比传统飞机的飞行时间短了一半，这使其成为世界上飞行速度最快的客机。

然而，令人遗憾的是，这种飞机在财务上却出奇地失败。当然，商业上失败的伟大思想、创新和产品比比皆是。但是，这个失败的独特之处在于，它持续亏损长达40余年。每当它超出预算的时候，英法两国政府就会砸入更多的资金。他们这样做，并不是不知道这些不断追加的投资收回的希望微乎其微，前期的投入就更是打水漂了。由于飞机上座位有限，预定很少，生产成本又高，因此即使是采用最夸张的估算方法，这个项目也是无利可图的。的确，当英国内阁文件按照国家档案30年后可供公众查阅的规则公之于众的时候，人们得知，当时英国政府的部长们早就知道这项投资“经济前景黯淡”。

为什么在这么长的时间里，这些有头脑、有才干的英法政府官员要一再投资这项一看便知的亏本买卖。原因之一就是常见的“沉没成本”这种心理现象。

沉没成本，该忽略就忽略

沉没成本指的是，由于已经投入其中的成本无法收回，而对某项已知的亏本生意继续投入时间、金钱或精力的倾向。毫无疑问，这很容易导致一种恶性循环：投入越多，就越想破釜沉舟坚持到底，直到投资获得回报。在某事上投入得越多，要从中抽离就越困难。

研发和制造这种协和式超音速喷射客机的沉没成本约10亿美元。英法两国政府在其中砸的钱越多，就越是欲罢不能。个人也同样很容易为沉没成本所俘虏。这就解释了，为什么我们能够忍受一场难看的电影而不中途退场，是因为我们已经为电影票花了钱。房子装修似乎永不见完

工，但我们却总是不断地往里面投钱，大概也是出于同样的原因。也正是受制于沉没成本，哪怕公交车或地铁迟迟不来，我们也会一直等着，而不去叫一辆出租车；明知一段关系有害无益，我们也会精心维护，哪怕自己的努力只会让情况更糟。诸如此类的例子俯拾即是。比如亨利·格瑞伯汉姆（Henry Gribbohm），他最近把自己毕生的全部积蓄，共计2600美元，都花在了一场狂欢节游戏上，结果却只赢得了一个大号香蕉娃娃。他花得越多，就越是铁了心要赢回来。亨利说：“‘我一定要把钱赢回来’的这个念头让人欲罢不能，但结果却适得其反。”随着对这个毫无意义的活动投入越来越多，他越无法从中脱身。

你是否也曾在无意义的事情上不断投入时间和努力，而不是斩仓止损？你是否也曾在成功无望的投资上越砸越多，而不是收手脱身？你是否也曾在没有前途的道路上一条道走到黑，因为不想承认从一开始就选错了方向？尽管已经深陷想补偿损失反而损失更多的怪圈，非精要主义者却依然不愿迷途知返，冲破藩篱，挣脱陷阱。而精要主义者却有勇气和自信承认自己的错误并脱身抽离，忽略已经沉没的成本。

沉没成本实在是太普遍，但它不是唯一需要我们小心提防的非精要主义陷阱。以下是另外几种常见的陷阱及相应贴士，帮助你礼貌、优雅地并以最少的代价解救自己。

六个常见陷阱与六条应对策略

陷阱1，禀赋效应

主人翁意识效力非凡。就像一句谚语所说的那样，世上从来就没有人会洗租赁的车！这是禀赋效应（**The Endowment Effect**）在作祟，也就是当一个人没有某件物品时，他就倾向于低估这件物品的价值，而当他一旦拥有该物品，就会倾向于高估它的价值。

在一项证实禀赋效应作用力的研究中，诺贝尔经济学奖获得者丹尼尔·卡尼曼（Daniel Kahneman）及其同事做了一项实验。他们把被试随机分成两组，并只给第一组每人发放了一个咖啡杯。然后问第一组人，这个咖啡杯他们想要卖多少钱。问第二组人，他们愿意花多少钱来买这个杯子。结果，“拥有”杯子的人表示低于5.25美元不卖，而没有杯子的人只愿花2.25~2.75美元来买这个杯子。换句话说，仅仅是因为拥有的缘故，就让这些杯子的主人们高估了杯子的价值，并且更加不愿舍弃它们。

在你的生活中，我相信你一定可以想到这样的东西，一旦你想要放弃它们的时候，它们就会显得更加有价值。想象一下，一本闲置在书架上多年未翻的书，一个仍未开封的厨房用具，或者一件从米尔德里德姨妈（Aunt Mildred）那里买的但从未穿过的运动衫。不管它们是否对你有用或者让你高兴过，它们属于你这个事实，都会让你觉得它们比不属于你的时候更有价值。

遗憾的是，对于不重要的事情以及不重要的物品，我们都有这样的偏见。如果自己是团队主管，一项毫无进展的工作项目就会显得格外要紧。如果自己是募捐活动的召集人，那么不参加志愿糕点售卖活动就会难上加难。当我们感到自己是某项活动的“主人”时，就倍加欲罢不能。但是，这里有一个有用的小窍门。

对策1，假装你还不曾拥有它

汤姆·斯塔福德（Tom Stafford）描述了一种简单的克服禀赋效应的方法。不要问：“这件物品我估值多少？”而是问：“如果手头没有，我愿意花多少钱来得到它？”对于机会和义务，也应该采取同样的方法。不要问：“如果错过了这个机会，我会有怎样的感受？”而是要问：“如果我没有这个机会，我愿意花多大的代价去得到它？”同样，我们也可

以问：“如果我不是已经参与了这个项目，我愿意费多大的劲儿来参与其中？”

陷阱2，对浪费的恐惧

俄亥俄州立大学心理学教授霍尔·阿克斯（Hal Arkes）在研究决策过程中的判断力问题时，有个难题让他困惑不已。为什么成年人比小孩更容易受到沉没成本的影响？他认为，这其中的原因是成年人终身都受控于“不要浪费”这条规则，因此当我们成年时，我们已经能训练有素地避免出现浪费，甚至对自己也是一样。“放弃一个你已在其中投入许多的项目，会让你感到一切都浪费了，而我们都被教导过浪费是应该避免的。”阿克斯说道。

为了说明这点，他向一群参与者演示了以下这个情境：假定你已经花了100美元买了一张去密歇根的周末滑雪旅游票。几个星期后，你又花50美元买了一张去威斯康星州的周末滑雪旅游票。你认为自己会更喜欢威斯康星州的那场滑雪旅游。当你把新买的威斯康星州滑雪票放入钱包时，发现这两张滑雪票都是同一个周末的。去卖掉或退掉其中任何一张都已经太晚了。你必须选定用哪张。阿克斯问大家：“你们将去哪里滑雪旅游？”超过一半的参与者都表示，他们愿意选择票价更贵的那场，尽管他们并不那么喜欢它。他们为此给出的理由是，使用那张便宜的票，会比使用那张贵的票，浪费更多的钱。很自然，人们不想放弃在一个错误的选择上已经浪费的东西，但一旦我们真的这样做了，就注定要浪费更多。

对策2，承认失败，让成功开始

我记得有一位友人，他从来都不会停下车去问路，因为他从不愿承认自己已经迷路了。我们总是浪费时间和精力开着车一圈圈地来回绕，

却哪儿都到不了——一个典型的非精要行为。

只有当我们承认自己犯了错，才能让这个错误留在过去。而如果我们总是一味地否认错误，就会始终漫无目的地原地打转，无法取得任何进展。承认错误，没有什么可羞耻的，因为我们真正承认的，是自己现在比过去更聪明了。

陷阱3，强己所难

在电影《窈窕淑男》（*Tootsie*）中，达斯汀·霍夫曼（Dustin Hoffman）扮演了一个拼命找工作的男演员。电影以一连串滑稽可笑的试镜失败开始。在其中一次试镜中，男主人公被告知“你太年轻了”。而在下一次试镜时，又被告知“我们要找的是一个年轻一点的人”。然后，再下一次就是“你的身高不符合要求”，他答道：“我可以更高。”但导演却回答说：“不，我们在找一个矮一些的人。”因为迫切想要得到这个角色，霍夫曼扮演的那个演员解释说：“我不这么高也是可以的，看见了吗，我是垫了增高鞋垫了。我可以变矮。”但是导演同样也很坚持：“这我知道，但我们要找的不是你这样的。”这个准演员还是不泄气，他再次回击说：“我可以变得不同。”在这个故事中，值得关注之处在于我们常常像霍夫曼所扮演的那个角色一样，太过努力地去追求不属于自己的东西。不管是在个人的生活中，还是在职场生涯里，人们都很容易去强己所难，强行匹配。那么，出路何在？

对策3，寻求中性的第二意见

每当我们因非要强行促成并不合适的事情而精神不安时，找一个参谋通常会有用处。由于事不关己，我们怎么选择都不会影响到他，于是他就可以让我们不再费力不讨好地去做一些根本无法做成的事情。

我曾经为了强行完成一个根本完不成的项目，浪费了好几个月的精力。回想起这件事，我发现投入越多，情况越糟。但是非理性的反应却是追加投入。我自认为我是可以搞定的！我不愿接受自己一直在浪费精力这个现实。最后，我向一个朋友倾诉了自己的挫败。他对这个项目不带任何感情色彩，没有沉没成本的负担，可以从不同的角度来评估我的决定。听完我的叙述后，他说：“你又不是和这个项目结婚了。”这句简单的话解放了我，让我停止了对这个项目没完没了的投资。

陷阱4，现状偏见

仅仅是由于自己一直在做某事而继续做下去，这种倾向有时被称作“现状偏见”（**Status Quo Bias**）。我曾工作过的一家公司，使用着一套在我看来早就严重过时的员工评价系统，我很想知道它究竟已经用了多久了。当我在公司里查找这个系统的发明者时，发现就连公司资深的人力资源总监也不知道这套系统的来历。更令人震惊的是，这位总监在公司工作的10年里，居然没人问过她有关系统的问题。仅仅因为它们早已存在，就听之任之、盲目接受，不再费心费力地询问质疑，这实在是太容易了。

对于这样的现状偏见，我们可以从会计学中借鉴一种治疗方法。

对策4，运用零基预算法

通常，会计在进行预算分配的时候，会以上一年度的预算作为下一年度预算规划的基底。但是，零基预算法（**Zero Based Budgeting**）却是以零为基底的预算方法。换句话说，预算提议中每一个项目支出的必要性都必须从零开始证明。这样做尽管比较费力，但是好处很多：它能根据实际需要而不是历史经验，对资源进行有效的分配，能够发现夸大其词的预算需求，引发对过时操作的关注，鼓励人们进一步明确目

标并提升支出与项目的契合度。

我们在工作中也可以运用零基预算法。不必以现有的任务为基础来进行自己的时间预算，而是要假定一切都难以预料，先前所有的承诺都已不复存在。然后，从零开始，问问自己愿意给今天加上什么任务。对于一切事情，不管是财务问题，还是你所负责的项目，甚至是你身处其中的各种关系，都可以这样操作。时间、精力和资源每使用一次，都必须重新证明使用的必要性。如果它不复适用，那就彻底消除。

陷阱5，随意承诺

有些人的时间被一些软性承诺塞得满满当当，这些都是他们不知在什么地方见到什么人时，因为随便发发评论或随意交谈了几句而不经意间答应下来的。你知道我所指的就是那种——你和邻居聊起了她在家长会的工作，和同事谈到了她现在负责的一个新项目，或者和朋友说起了一家他很想尝尝的新餐馆，而还没等你完全反应过来，你就已经做了某种许诺。

对策5，先停顿，再说话

这听起来似乎是显而易见的事情，但是在答应一件事之前先停顿5秒，能大大降低做出后悔承诺的可能性。先别让“这听起来很棒，我很愿意”这样的话脱口而出，而是问问自己：“这事儿重要吗？”如果你已经随随便便做出了一个让自己后悔不已的承诺，想一个好办法抽身而退吧。道个歉并且告诉对方，当时你作出这个承诺的时候，没有充分认识到它会带来什么。

陷阱6，害怕错过

本章已经提供了充分的例证，告诉大家绝大多数人天生就非常厌恶损失。因此，脱离现行轨道的障碍之一就是害怕自己会错过好东西。

对策6，逆向试行，战胜害怕

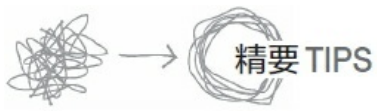
近年来，“原型法”理念（**Prototyping**）在商界很流行。构建一种原型或者大尺度模型，可以让公司试运行一种理念或产品，而不用预先进行大规模投资。同样的理念可以被反向运用，以一种相对而言低风险的方法排除不重要的东西，领英的董事丹尼尔·夏皮罗（**Daniel Shapero**）称这种方法为“逆向试行”（**Reverse Pilot**）。

在逆向试行中，你可以检验取消一项尝试或活动是否会产生负面效应。例如，我合作的一位领导者在公司里担任了一个新的高管职务，他接手了其前任费了很大力气实施的一项工作：每周为其他领导者写一份大篇幅、非常直观的涉及多个主题的报告。这极大地消耗了他团队的精力。他假设这个报告并未给公司带来太多价值。为了检验这个假设是否正确，他进行了一次逆向试行。他停止了报告的发布，等着看会出现什么样的反应。他发现，没有人对这个报告流露出想念的意思，几星期以后，甚至都没人提起这个报告。结果，他得出结论：此报告对公司并无重要意义，可以取消。

在我们的社交生活中，也可以进行类似的逆向试行。你是不是经常会对顾客、同事、朋友或者家庭成员做些承诺，你总是觉得这些能给他们带来重要影响，但实际上他们可能根本没有注意到？悄悄地取消或者至少把一个活动缩减几天或几周，你大概就能评估出它是否真的重要或者是否真的没人在乎。

就算用上了这些技巧，“不履行承诺”还是比从一开始就不承诺要难。要对已经承诺的事或人说“不”，是会让人深感愧疚，但我们也得直面它，没有人愿意违背自己说过的话。不过，学会这样做，以自己的勇

气、专注和自律赢得尊重，是成为精要主义者的关键所在。



- 小心禀赋效应，假装你还不曾拥有。
- 克服对浪费的恐惧，承认失败，让成功开始。
- 不要强己所难，寻求中性的第二意见。
- 小心现状偏见，运用零基预算法。
- 不要随意承诺，先停顿，再说话。
- 不要害怕错过，逆向试行，战胜害怕。

CHAPTER 12

EDIT

The Invisible Art

要素九

编辑，看不见的艺术

我在一块大理石中看到了一位天使，
我不停地雕刻，直到将它释放。

——米开朗琪罗



非精要主义者

认为要改善就意味着要增加
对每个词、形象或细节都恋
恋不舍



精要主义者

认为要改善就需要删减
删减分散注意力的词、形
象和细节

每年奥斯卡金像奖的重头戏是“最佳影片奖”。在颁奖典礼播出的前几周，媒体就对此奖项的归属进行预测，大多数观众还会熬夜观看此奖的颁发。而那天晚上还有一个没有被大肆炒作的奖项：最佳剪辑奖。看看现实情况吧：当最佳剪辑奖宣布的时候，大多数观众会转换频道，或者跑到厨房再满满地装上一大碗爆米花。然而大多数人不知道的是，其实这两个奖项是高度相关的：自1981年以来，每一部获得最佳影片奖的影片无一不是最起码获得了最佳剪辑奖的提名。实际上，在获得最佳剪辑奖提名的影片中，有大约2/3的影片最后都获得了最佳影片奖。

在奥斯卡金像奖的历史中，最受尊敬的（如果不是最著名的）电影剪辑师是迈克尔·卡恩（Michael Kahn），获得8次提名——超过了这个行业中的任何人，并3次获奖。虽然他的名字不是家喻户晓，但他剪辑的电影却是。经他剪辑的著名影片有《拯救大兵瑞恩》《夺宝奇兵》《辛德勒的名单》《林肯》。在37年时间里，他剪辑了几乎所有史蒂文·斯皮尔伯格（Steven Spielberg）的电影，成为斯皮尔伯格电影工作中不可或缺的帮手。然而，卡恩的名字却鲜有人知。因此，电影剪辑有时候被认为是“看不见的艺术”。

很明显，剪辑就是严格淘汰那些无意义的、不重要的或者不相关的内容，它是一种精要工艺。那么，优秀剪辑师是怎么炼成的？当美国电影艺术与科学学院（The Academy of Motion Picture Arts and Sciences）的电影剪辑行会成员坐下来挑选最佳剪辑提名者时，正如马克·哈里斯（Mark Harris）描写的那样，他们“非常努力地不去看那些他们应该去看的”。也就是说，一个优秀的电影剪辑师会让人很难不去看重要内容，因为他已经淘汰了除了绝对必须存在的元素之外的一切东西。

在本书第5章里，我们把探索过程比作成为新闻记者，它需要提问、倾听和联系要点，以便将有意义的少数和无意义的多数加以区别。

因此，在精要主义者养成过程的下一个阶段，淘汰这些无意义的事务，就意味着要充当自己人生和领导力的剪辑师。

杰克·多西（Jack Dorsey）是Twitter和移动支付公司Square的联合创始人兼首席执行官。他有着一种较为罕见的精要管理方法。在我最近参加过的一次晚宴上，他发言说，首席执行官就好比是公司的首席编辑。还有一次，他在斯坦福大学更深入地解释了这个问题：“关于编辑，我的意思是可做之事千千万万，但是重要之事只有一二。各种各样的想法.....来自工程师、支持者、设计师的各类信息如潮水般不断涌来，该做之事多得泛滥.....作为一名编辑，我不断地吸收这些输入的信息，决定哪一点或是哪一交汇部分，是值得去做的。”

删减冗余，删减冗余

编辑的工作并不仅仅是说“不”，这连三岁小孩儿都会。编辑也不仅仅是淘汰，实际上，编辑还以某种方式增加补充。我是说，一个优秀的编辑通过深思熟虑的删减，为思想、环境、情节和人物增添生命活力。

与此类似，在生活中，训练有素的编辑工作可以提升你的贡献水平。它能让你变得更加专注，为真正重要的事情投入能量精力。它能给最有意义的关系和行为以更大的空间来繁荣生长。

编辑，能让执行变得不费力气，因为它扫清了一切容易导致分心的、不必要的和棘手的事情。或者，就像一位图书编辑说得那样：“我的工作是为读者的生活变得尽可能轻松自如，帮助他们尽可能清晰地理解最重要的信息和道理，这就是我的目标。”

当然，编辑也需要进行权衡取舍。一名编辑不会试图安排每一个章节，每一个情节转变，每一个细节，而是会问：“这个章节、情节转变

或者细节，能让故事更出色吗？”而作者，不管是电影、图书还是新闻作品的作者，都很容易过分专注于某种观念或者作品主体，特别是那种呕心沥血之作。要删除作品中的几个段落、几页纸，甚至几个章节，简直如同割肉般疼痛难忍，因为它们最初是用了几个星期、几个月，甚至可能是好几年才写成的。但是，这种训练有素的删除却对这门手艺非常重要。你必须像史蒂芬·金（Stephen King）说的那样：“删减冗余，删减冗余，哪怕这会让你这个以自我为中心的小文人心碎，删减冗余。”

当然，编辑一部电影，或一本书，或任何其他创意作品，并不等同于编辑自己的人生。在生活中，我们不能奢望重新聆听刚刚进行的那次交谈，或再参加一次刚刚主持过的会议，或再做一遍刚刚做过的一个报告。然而，编辑工作所固有的四条简单原则，可以用来把无意义之事逐出我们的生活。

编辑自己人生的四条原则

1.删减，自律地抛弃那些好选项

显而易见的事情是，编辑必须把容易让读者犯晕并且遮蔽信息和故事的内容删除掉。编辑精良的电影和图书都是赏心悦目并且便于理解的，这都是有据可查的。

当我们作出决策的时候，决定删减选项是可怕的，但事实是，这正是决策过程的精华所在。实际上，decision一词的拉丁词根“cis”或者“cid”，字面意思即为“切”或“杀”。

在scissors（剪刀）、homicide（杀人犯），或者fratricide（杀害兄弟者）这些词中，也有这个词根。既然归根结底，选项更少能让决定“更加赏心悦目”，我们就必须严格自律地抛弃那些不错的，甚至非常好但却成为障碍的选项。诚然，要下决心删除好东西是痛苦的。但是，

当我们意识到自己赢得的每一个额外时刻都能用于更好的事情之上时，那么最终每一次删减都会带来快乐——也许不是那么立竿见影，但终能实现。这也许解释了为什么史蒂芬·金会写“写作是人做的，编辑是神做的”。

2.精缩，用有意义的活动取代无意义的活动

想必很多人都容易有这样的想法：“我必须道歉，如果我有更多时间的话，我就会写一封更短的信。”的确，少做可能更难，不论是在艺术创作中，还是在人生中，都是如此。每一个词，每一个场景，每一个活动都必须更有价值。为了让每一个词都有价值，编辑必须无情。你能不能只用一个句子，而不是两个，就把意思说清楚？现在用两个词的地方能不能只用一个词？艾伦·威廉姆斯（Alan D.Williams）在《编辑是什么？》（*What Is an Editor*）一文中指出，“编辑要问作者两个问题：你所说的正是你想说的吗？你的表达是不是尽可能地清晰、简洁？”精缩，意味着尽可能清晰而简洁地表达。

同样，在生活中，精缩让我们事半功倍，以一当十。比如，当格拉汉姆·希尔（Graham Hill）搬进位于纽约的40平方米的寓所时，他想知道自己可以把物品精简到什么程度。最终结果是一种被他称为“小珠宝盒”的设计。这个盒子很管用，因为每一件家具都是多功能的。例如，画左边的墙是一个大大的投影屏幕，可以用来看电影，其中还藏着两张客床，当有客人来住宿时可以拉出来用。右边的墙折向下，越过长沙发，就是一张双人床。每一件家具都具有双重或者三重功能，换句话说，所有物品都在为公寓生活作出更大的贡献。这种设计如此具有新意，因此希尔把它变成了一个企业，专注于让小空间发挥大功用的艺术。他还给它起了一个恰当的名字：LifeEdited.com（编辑后的生活）。

但必须明确的是，精缩并不代表要马上做更多，而是意味着要少浪费。它指的是，用更少的词表达同样的思想，用更小的空间获得同样的功用，或者用更少的努力取得同样的成就。要在生活中运用精缩原则，我们必须改变行为与意义之间的比率，淘汰许多无意义的活动，并用特别有意义的那一个来取代它们。比如，我合作过的一家公司里有一名员工（他在公司地位足够稳固，毫无失业之忧）经常不参加其他人都参加的每周例会，只是问下别人会上讲了什么。这样，他就把一个2小时的会议精缩成了10分钟，把节省下来的其余时间用于完成重要工作。

**The Latin root
of the word
decision——cis
or cid——literally
means “to cut”
or “to kill.”**

“decision” 一词的拉丁词根——“cis” 或 “cid”
字面意思为“切” 或 “杀”。

3.纠正，反观核心，调整方向

编辑的工作不仅包括删减和精缩，还包括纠正错误，也许只是一个微小的改动，比如纠正一个语法错误，但也可能是一个复杂问题，比如匡正论证中的各种谬误。要做好纠错工作，编辑必须对自己正在加工的这个作品的核心主旨了如指掌。迈克尔·卡恩解释说，他并不总是按斯皮尔伯格说的去做，而是按他自己认为斯皮尔伯格真正想要的去做。因为理解了最核心的意图，迈克尔·卡恩所做的纠正工作，甚至连斯皮尔伯格本人都不能用言语表达出来。

同理，在我们自己的职业生涯或个人生活中，我们可以经常性地反观自己的核心目标来调整航向。正如我们在本书第9章中讨论过的那样，具有一个支配一切的精要目标，能帮助我们进行自我审视，审视自己的活动和行为是否与自己的目标相符相配。如果行为有误，我们就要编辑修正。

4.少量编辑，知道在什么时候谨慎克制

这似乎有点违反常理，但最好的编辑并不认为需要改动一切。他们知道，有些情况下，必须让某些特定的内容保留原样，这是对编辑判断力的最佳运用。这也从另一种途径体现了编辑是一门看不见的艺术。最好的外科医生不是开刀最多的那个，同样，最好的编辑常常是对作品改动最少的，最谨慎克制的那一个。

成为自己人生的编辑，也必须知道什么时候应该谨慎克制。方法之一，就是编辑自己要插手干预的那种倾向。比如，当我们被拉进一个电子邮件话题时，要克制住常常会有的那种第一个回复全部的冲动。参加会议时，克制住想插话的冲动。我们可以等待，观察，看看情况怎么变

化。少做，不但是—种效力强大的精要策略，而且也是一种非常有用的编辑策略。

非精要主义者将编辑工作视为—种分散独立的任务，只在事务繁重、令人难以应付时才不得已而为之。但是拖延太久才去编辑，会让我们迫不得已地作出—些不是自己选择的重大舍弃。相反，持续性地编辑自己的时间和行动，能让我们—路上都可以作出—些更小但却是深思熟虑的调整和改变。要成为精要主义者，就要让删减、精缩和纠正成为我们日常工作的天然组成部分，要让编辑成为生活的自然节奏。



- 删减，自律地抛弃那些好选项。
- 精缩，用有意义的活动取代无意义的活动。
- 纠正，反观核心，调整方向。
- 少量编辑，知道在什么时候谨慎克制。

CHAPTER 13

LIMIT

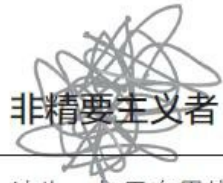
The Freedom of Setting Boundaries

要素十

边界，设定界线会带来自由

“不”，是一个完整的句子。

——安妮·拉莫特



非精要主义者

认为：如果有界线，自己将受限制

视界线为约束

努力尝试直接说“不”



精要主义者

懂得：如果有界线，自己将不受限制

视界线为自由

事先设定界线，避免真接说“不”

金永是韩国一家科技公司的员工，她发现自己的婚礼筹备正好与董事会会议的准备工作撞车，这个会议将在她大喜之日的前三周召开。金永的经理李孝利吩咐她把董事会会议上的报告所需的脚本和幻灯片都准备好。金永连着好几天每天都工作15个小时，很快完成了工作，这样就可以把董事会会议的准备时间用于筹备婚礼了。看到她提前完成了工作，她的经理很吃惊，也很高兴。金永现在有连续5天的时间可以全心投入自己的婚礼筹备活动。

可是后来，金永的经理紧急要求她在董事会会议之前完成一项额外的项目。在一起工作的几年时间里，金永从来没有跟李孝利说过“不”字，哪怕她说出的“**Yes**”会让她短时间内陷入一片混乱之中（这样的事情经常发生）。到目前为止，金永已经花了不知道多少时间，也不知道做出了多少牺牲，来执行每一个要求和任务，并利索、完整地交付了工作。然而这一次，她毫不犹豫地对她的主管说“不”。她没有道歉，也没有为自己的答复过多地辩护。她只是简单地说：“这次的婚假我已经计划好了，并为之努力工作，我理应享有它.....我没有过错。”

不可思议的事情发生了。团队中的其他所有人都说了“不”，经理李孝利只好自己完成这个任务。一开始，李孝利很生气。她花了整整一周才完成这个任务，因此很不高兴。但是，费劲地干了好几天后，她发现自己工作中存在各种各样的毛病。她很快就意识到，如果要使自己的管理更高效，她必须停下来整顿局面，让团队中的每一个成员明确期望、责任和结果。最后，她很感激金永让她看到了自己工作方式上的失误。通过设定界线，金永不仅使经理看到了团队动力中的不健康状态，同时也为改变创造了空间，她的所作所为为她赢得了持久的感激和尊重。

模糊的界限与被侵蚀的生活

界线的丧失是非精要时代的典型特征。当然，一方面，技术已经完

全模糊了工作和家庭之间的界线。现在，对于别人期望我们能够干活而言，已经不存在任何界线。我最近让一个行政助理给我提供与客户的会面时间，其中还包括了星期六早上。尽管那个会面并没有那么紧迫，安排在周六也没人认为有何不寻常之处。但是我很好奇，星期六已经成为新的星期五了吗？然而，大多数人没有意识到的是，问题不仅在于界线已经模糊了，还在于工作的界线已经恶意地侵蚀了家庭生活的领地。很难想象，大多数公司的领导者，如果看到员工周一上午带着孩子来上班会感到很舒服。可是如果看到员工周六或周日来办公室为一个项目加班，他们似乎不会有任何异样的感觉。

克莱顿·克里斯坦森（Clayton Christensen）是哈佛商学院教授，著有《创新者的窘境》（*The Innovator's Dilemma*）一书。¹¹他也曾被人要求做出这样的牺牲。当时，他正就职于一家管理咨询公司，一个合作伙伴过来要求他周六加班赶一个项目。克莱顿简要地答复说：“哦，对不起。我已经承诺过了，每个周六都要陪伴我的妻子和孩子。”

对方很不高兴，气冲冲地离开了，但是后来他又回来告诉克莱顿：“克莱，我已经同团队里的每一个人都谈过了，他们都答应了星期天来加班。所以我也希望你也能来。”克莱顿叹了口气，然后说道：“感谢你这样做。但是礼拜天也不行。我的礼拜天已经给了上帝，因此我不能来加班。”这个合作伙伴之前就已经很懊丧了，现在更是心灰意冷。

克莱顿并未因为坚持立场而丢了饭碗，而且，尽管在当时他的选择并不受待见，但最终他因此赢得了尊重。界线发挥了效用。

克莱顿回忆说：“这件事给我上了重要的一课。如果我当时破例一次，那么我大概已经破例好多次了。”界线就如同沙堡的围墙一般，一旦倒下，其余就会瞬间崩塌。

我不否认，设定界线绝非易事。仅仅因为金永和克莱顿从中受益，

并不足以说明界线永远都会起作用。金永可能会失去工作机会。克莱顿不愿意在周末加班，这可能会限制他的职业发展。界线的确会索要高昂的代价。

然而，不推脱的代价更高——你会失去选择人生头等大事的能力。对于金永和克莱顿来说，在职场中赢得尊重和拥有陪伴上帝及家人的时间是最重要的，因此他们慎重地并且战略性地将这些事列为头等大事。归根到底，如果你不设定界线，界线就无从谈起。或者更糟，界线是有，但它们顺从默认设置或者由他人说了算，而不是由你自己主动设计。

非精要主义者总觉得界线就是约束和限制，是自己超级高产的人生道路上的拦路虎。对于非精要主义者来说，设定界线无异于自证懦弱。他们认为，如果自己足够强大，界限就毫无必要。他们能够应付一切，搞定一切。但是由于没有限制，他们最终被拉扯得细若游丝，搞定一切终成幻影。

而与之相反，精要主义者视界线为一种授权。他们承认，界线保护了他们的时间，使之免受劫持，界线也让他们摆脱了对有利于实现他人目标而非自身目标之事说“不”的负担。他们知道，清晰的界线能让他们积极主动地淘汰各种要求和累赘之事，以免自己精力分散，无法专注于真正重要之事。

If you have limits you will become limitless.

设定界线，才能获得自由。

免受他人事务烦扰的四个方法

当然，设定界线的挑战远远超出了职场范围。在个人生活中，的确有些人在要求我们牺牲时间的时候，毫无界线意识。你是否也经常感觉到自己的周末被别人的事情所劫持？在你的生活中，是否也有这样的人，丝毫觉察不到他/她正在越过别人的界线？

在我们的生活中都会有这样一些人，和他们维护关系要比跟其他人费劲。这些人总是把他们自己的问题当成我们的问题。他们让我们偏离自己的目标，他们只关心自己的日程事项。如果我们听之任之，那么他们就会榨取我们的时间和精力，来为他们的重要事务而非我们自己的核心事务效力，这样我们就无法实现自己的目标。

因此，我们该如何从金永和克莱顿·克里斯坦森身上吸取经验，设定界线，让它们保护我们免受他人事务的烦扰呢？下面有几个方法供你参考。

1.别人的问题不是你的问题

我不是说我们不应该帮助别人。我们当然应该为别人提供服务，关爱他们，使他们的生活有所改变。但是，如果别人把他们的问题变成我们的问题，那么我们就不是在帮他们，而是在纵容他们。一旦替他们承接了问题，我们所做得一切就变成了剥夺他们解决问题的能力。

作家亨利·克劳德（Henry Cloud）在《界线》（*Boundaries*）一书中讲了一个故事，说的就是这种情况。有一次，一对父母过来看他，希望他能“治好”他们那位25岁的儿子。他就问：“为什么孩子没有一起来？”这对父母说：“他不觉得自己有问题。”在听了他们的讲述之后，克劳德得出结论，这结论让这对父母很吃惊：“我觉得你们的儿子是对的。他没有问题.....有问题的是你们.....你们付钱，你们焦急，你们担心，你们计划，你们费尽心力让他继续这样下去。那么，他就没有问题了，因为你们已经从他那里把问题接过去了。”

然后，克劳德给他们打了个比方。想象一个邻居，他从来不给他家的草坪浇水，但是只要你打开你的洒水器，水就洒到了他家的草坪上。你的草地渐渐变黄并行将枯死，但是邻居看了看他家的青青绿草，暗自思忖：“我家的院子不错呀。”于是，所有的人都是输家：你的努力白费了，邻居也从没有养成给草地洒水的习惯。怎么解决？正如克劳德说的：“你必须筑起篱笆，让他的问题离开你的院子，待在它们该待的地方。”

在职场中，有人总想使用你的洒水器来给他们自己的草浇水。也许是老板让你为他的一个宠物项目效力；也许是一个同事让你为一个她自己都无暇完善的报告、讲演或计划贡献智慧；也许是一个同事在走廊上拦住你，喋喋不休地跟你说话，而你正要去参加一个重要的会议，有一个要紧的电话要打，或者桌上还堆着重要的工作。

不管是谁想要榨取你的时间和精力，来为他自己的目的服务，唯一的解决办法就是筑起篱笆。不是在他们提出要求的时候，而是必须提前筑起篱笆，清楚地划分出哪些是超越了界线的，以便将浪费时间和侵犯界线之事阻挡在关隘之外。不要忘了，迫使这些人自己解决自己的问题，不仅于你有利，也于他们有益。

2.没有界线，就没有自由

有一个关于一所学校的故事，巧妙地阐述了这个道理。这所学校临近一条繁华的道路。最开始的时候，孩子们只能在操场上一小片划出来的区域里玩耍，这片区域临近一幢住宅楼，大人们可以在那里看着孩子们。但是后来，有人围着操场筑起了一道围墙。现在孩子们可以在操场上的任何一个地方玩耍了。他们的自由，实际上比之前增加了一倍还多。

同样，如果我们不在自己的生活中设定界线，那么就只能被囚于别人替我们设定的限制之中。相反，如果我们有了清楚的界线，我们就能自由地从整个领域中进行选择，或者从我们精心选择的要去探索开发的全部选项范围中进行挑选。

3.找到自己的坏事者

当我让领导者们确定自己的界线时，他们几乎都做不到。他们知道自己是有界线的，但就是不能用语言表达出来。一个简单的事实是：如果你不能和自己、和别人说清楚这些界线，那么你期待别人能够尊重或者找到你的界线，恐怕是不现实的。

想想那个频繁把你从精要道路上拉出去的人，列一个坏事者清单——这个人提出的各种要求和行为，你都拒绝，除非它们正好和你自己

的优先要事或者事项安排有交叉重叠之处。

找到你的坏事者的另一种快速测试方法就是，每一次你感到别人的要求冒犯了你或欺骗了你，你都要记下来。你无须通过某种极端的方法就能注意到它。即使是一个令你不快的小小“刺痛”（我觉得这种描述有助于形容你的界线所受到的那种微小侵犯），不管是来自一个你不想得到的邀请，还是一个不请自来的“机会”，抑或是一种帮个小忙的请求，都是你发现自己隐藏的界线的线索。

**Sets boundaries
in advance
that eliminate
the need for
the direct
“no”.**

事先设定界线，
避免直接
说“不”。

4.制定社会契约

我曾与一位同事搭档工作，他处理项目的方法与我截然相反。因此，有人预测我们之间会闹得不可开交。然而，我们的工作关系却十分融洽。这是为什么呢？因为从合作一开始，我就郑重其事地亮出了自己的优先事项，说明了在项目存续期间哪些额外的工作我乐意做，哪些不乐意做。“让我们说好我们想得到什么，”我开门见山说道，“以下这些事情对我来说很重要……”而且，我让他也这样做。

这样，我们就完成了一个“社会契约”，这与本章开篇故事中金永和她老板定的那个可不一样。就双方真正要着力实现的目标提前达成共识，互相了解对方的界线，这让我们避免了浪费彼此的时间，不去给对方提出徒增烦恼的要求，也不去分散对方的注意力，使其无法专注于要紧之事。结果，我们都在这个项目上实现了个人贡献峰值，而且尽管也有意见不同，但在整个工作过程中，我们相处得十分融洽。

通过不断练习，你会发现，设定界线会变得越来越容易。



- 不要把别人的问题，变成你的问题。
- 设定界线，才会拥有自由。
- 找到自己的坏事者。
- 制定社会契约。

Part IV

精要主义者养成的第三步：执行





execute

How can we make doing
the vital few things
almost effortless?

导读

让有意义的少数做起来毫不费力

关于执行，有两种思考方法。

非精要主义者通常会用蛮力执行，而精要主义者却通过淘汰无意义之事来节省时间，设计一套体系，让执行变得毫不费力。

在本书前言部分中，我们谈到人生就好比一个塞得过满的衣柜，精要主义者会将它整理得井然有序。我们谈到，如果想让自己的衣柜保持整洁，就需要进行一种常规性的整理。就需要一个很大的袋子，来装要淘汰的东西，而需要留下来的东西只有很小很小的一堆。你必须知道弃物存放地和当地旧货商店的营业时间，你还要预先安排好去那里的时间。

也就是说，一旦确定了哪些行为和努力要在人生中加入保留，你就必须设立一套体系来执行。不能等到衣柜满得要被撑破时，才去用超人般的努力来清理。你必须有一套体系，这样保持衣柜整洁不过是例行公事，毫不费力。

做容易做的事是人的天性。在本书这个部分，你将学会如何让正确的事情——重要的事情，执行起来尽可能轻松自如，顺利无阻。

CHAPTER 14

BUFFER

The Unfair Advantage

要素十一

缓冲，充足准备带来优势

给我6个小时砍倒一棵树，我会先花4个小时来磨快斧子。

——亚伯拉罕·林肯



非精要主义者

认为最好的情形会发生

不到最后一刻不行事



精要主义者

为意外情况设置缓冲

提前做好最充分的准备

在《圣经》中有一个关于约瑟夫的故事，是他把埃及从一场长达7年之久的可怕饥荒中拯救了出来。埃及法老做了一个梦，他自己无法解释，于是就让手下那些最聪明的智囊们给他好好解释一番。但他们也无能为力，幸好有人想起了关在狱中的约瑟夫善于解梦，于是约瑟夫便被召去帮法老解梦。

法老梦见自己驻足于一条河边，看见7头肥壮的母牛从河里上来，在岸边吃草。随后，又有7头干瘦的母牛从河里上来，它们吃掉了第一次上来的那7头肥壮的母牛。对此，约瑟夫解释说，这个梦意味着埃及将迎来7个丰收年，但之后会遭受7个荒年。因此，他建议法老委派一个“谨慎且智慧”之人将7个丰年中每一年的收成保留1/5，积蓄起来以备荒年之需。法老同意了 this 提议并委任约瑟夫为地位仅次于他的高官，掌管埃及事务。约瑟夫成功地推行了他的计划，当7个饥荒年到来之时，埃及和周边地区的所有人，包括约瑟夫一大家族的人都幸免于难。故事虽简单，但其中包含了一个很有用的经验，精要主义者们运用它来使执行变得毫不费力。

我们生活的世界无法预知，这是确定的现实。像饥荒这样的极端事件自不待说，意料之外的事情更是比比皆是。谁能知道交通状况是畅通无阻还是拥堵不堪？谁能预知自己要搭乘的航班将会晚点或是被取消？谁能料到自己明天会否在光滑的路面上滑倒摔伤手腕？同理，在公司里，我们无法预先知道供应商会迟到，同事会出错，或者客户会在最后一刻变卦等。无法预料却是我们唯一可以预料的（确定无疑）。既然如此，我们要么坐等这样的时刻到来然后对此作出反应；要么提前准备，设置缓冲，未雨绸缪。

“缓冲”字面上的解释是用以避免两物接触以免相互伤害的事物。例如，在受保护区的外缘设立的缓冲区，就是辟出一块区域作为额外空间，将保护区和任何可能侵入其中的潜在威胁隔离开来。

有一次，我试图向我的孩子解释缓冲的概念。当时，我们一起坐在车里，我试着用游戏的方法来解释它的意思。我让他们想象一下，我们要一口气开5公里直奔目的地，中途一次都不停。孩子们立刻感受到了挑战重重。我们无法预测前方会遭遇什么，周边会发生什么，不知道绿灯会亮多久，前面的车辆会不会突然转弯或者紧急刹车。要避免发生碰撞，唯一的方法就是与前车保持距离。这个车距就是缓冲。如果别的车突然移动，这个缓冲就让我们有时间作出反应和调整，从而避免猛然停止或发动引发摩擦。

同样，在工作和生活中只要留一个缓冲，我们就可以减少执行必要任务时的阻力。

在和孩子们玩汽车“游戏”时，他们注意到了当我因为说说笑笑而一心两用时，就会把保留缓冲区这回事抛到脑后，而与前车靠得太近。这样一来，我就不得不在最后一刻做出一些“不自然”的举动，比如突然转向或猛踩刹车来调整距离。在生活中，如果不重视和保留缓冲，那么同样的事情也会发生。我们忙忙碌碌，无法凝神聚力，不知不觉任务期限到了，该做重要报告了——不管我们在其中额外花了多少时间；结果，却只能被迫无奈在最后一刻“转向”或者“紧急刹车”。化学知识让我们懂得气体会扩散并充满其所在的整个空间，同样，我们都体会到了项目和任务也是扩散型的，哪怕我们尽了最大努力，它们也会用完所有分配到的时间。

想想这样的事情在你参加过的报告、会议和研讨会中是多么频繁。你是不是经常发现有人要在很短的时间内讲完很多张幻灯片？你参加过的会议里是不是有好多都让你感到主讲人突然中断了一个有意思的谈话，因为他必须把准备好的所有内容都讲完？这样的情形我见过太多了，它甚至俨然成为一种默认模式。因此，当我与一位有着不同理念的主讲人一起工作时，真的有一种挣脱牢笼的感觉。他当时正在策划一场

为时4个小时的研讨会。通常，会议的最后10分钟被用于提问和评论，但他却建议留出整整一个小时。对此，他的解释是：“我喜欢留出充足的时间，以防意外。”最开始时，他的意见没有被采纳，因为大家觉得这太没有必要了，并要求他按惯例讲。果不其然，讲课超出了预定的时间，这位主讲人只好仓促讲完剩余的内容。于是，后来的课程就按照他的意见调整为留出一个小时进行提问和评论。这下事情就进展得如他所愿了，因为这一次他设置了缓冲。现在，课程可以按时结束，主讲人可以专心讲课而不用紧赶慢赶了。

我认识一位母亲，她在准备与家人度假的过程中学到了相似的道理。以前，当他们要度假的时候，她总是直到临行前一晚才打包。这样一来，她就难免会熬夜，泄气，睡眠不足，甚至需要通宵达旦才能完成，而且还容易缺这少那，导致出发延迟，甚至必须通过匆忙赶路来弥补。但是这一次，这位母亲提前一周就开始打包。头一天晚上，她确认车上装满了东西，这样次日早上她只要把孩子们叫醒，招呼所有人上车就大功告成了。这个办法非常有效。他们晚上美美地睡了一觉，第二天早早就出发上路，东西都带齐了，就算遇上堵车也不用着急，因为他们已经为此留出了富余的时间。结果，他们不仅按时到达，而且享受了一个快乐的假期，没有遇上一点磕磕绊绊的情况。

基于最佳情况，考虑最坏可能

非精要主义者往往基于最佳情况来考虑问题。对于这样的人，我们并不陌生（我们中间的大多数人，包括我自己在内，都曾是这样的人），他们会习惯性地低估一件事情所需的时间：“这只要5分钟就行”，或者“周五前我就能完成这项工作”，或者“一年足以完成我的巨作”。可是，这些事情却难免需要更多的时间，不是有这样那样的意外情况发生，就是任务比预想的更复杂，或是从一开始就过于乐观地估计了形势。一旦出现这样的情况，他们就只好疲于应付，自食其果。要么

通宵熬夜加班加点；要么偷工减料走捷径，只交个半成品，或者更糟——根本就干不了；要么就干脆把烂摊子甩给其他同事。不管用哪种方法来对付，他们都无法以自己的最高水准来执行任务。

精要主义者的做法就迥然不同。他们高瞻远瞩，他们规划安排，他们筹备应急，他们预料意外，他们为无法预知之事置备缓冲，以便意外情况不可避免地发生时，自己能有回转腾挪的余地。

非精要主义者若得到一笔意外之财，往往会消耗挥霍，而不是为穷日子积蓄储备。这种情况在有些国家找到石油后采取的行动中可见一斑。1980年，英国发现了北海油田，政府税收突发横财，10年之内收入高达1660亿英镑（合2500亿美元）之多。围绕这笔横财该怎么花，人们进行了激烈的争论。然而，不管怎么争论，这笔钱到底还是花了。英国政府没有为意想不到的灾害（比如，事后才知一场大萧条正在逼近）设立基金，而是把这些钱另作他用。

与此不同的是，精要主义者倾向于未雨绸缪，以便有备无患。挪威也曾从石油税收中获利巨大，但不同于英国所为，挪威政府把这笔巨额收益中的大部分投资于一项基金。现在，这项基金的价值已经增长到了7200亿美元之巨，是全球最大的主权财富基金，能为将来未知的情势提供缓冲。

现如今，人们的生活节奏越来越快，快得就像以160公里时速行驶在距离前车仅2.5厘米之遥的马路上。倘若前车作出意想不到的举动，哪怕是最微小的，比如减一点点速，或者转一个极小的弯，都会造成严重的追尾，因此，不能有一丝一毫的闪失，而这样导致的结果就是，事情执行起来变得压力重重，令人沮丧万分。

设置缓冲区的三个窍门

下面是一些小窍门，告诉你如何通过设置缓冲来使工作步入正轨，保持头脑理智清醒。

1. 做完善的准备

我在斯坦福大学读研究生时，发现完善的准备是取得高分的关键。每次一拿到课程大纲，我就会复印几份，和整个学期的校历粘在一起。在还未开始上课前，我就知道了有哪些重要的研究课题，并立马开始动手准备。这样的准备减轻了整个学期的压力，因为知道自己有充足的时间来完成所有作业，就算工作量突然增加，或者家里有点儿急事而不得不缺几次课，或者有其他意外情况出现，也无甚大碍。

罗尔德·阿蒙森（Roald Amundsen）和罗伯特·法尔肯·斯科特（Robert Falcon Scott）二人比赛看谁能成为现代人类史上到达南极点的第一人的故事，显示了做最全面、最充分的准备是多么重要。两个人目标相同，但做法却迥异。阿蒙森为所有容易出错的环节做好了一切准备，但斯科特却对一切抱乐观的态度。斯科特为这次探险只带了一支温度计，所以当那支温度计打破时，他忍不住怒气冲天，而阿蒙森带了4只温度计上路。斯科特为他的17人团队储备了1吨食物，而阿蒙森准备的食物是他的3倍之多。斯科特把返程的补给物资贮藏在一个地方，上面只插了一面旗子作为标记，这意味着如果他哪怕有一丁点儿偏离线路，他的团队就可能找不到补给了。与他形成鲜明对照的是，阿蒙森做了20个标记，相距数公里远，以确保其团队一定能够看到。罗尔德·阿蒙森为这次行程做了煞费苦心的准备，并且查阅了大量资料，但罗伯特·法尔肯·斯科特的准备工作却少得可怜。

阿蒙森在制订计划时全面考虑了挫折和缓冲，但斯科特却只期盼理想的状况。当斯科特的队员们疲惫不堪、饥寒交迫之时，阿蒙森的队伍却很顺利地行进着。最后，阿蒙森成功地抵达了南极点，而斯科特和他

的队友们却惨死于途中。

完善准备的重要性同样也显现在商业活动中。实际上，这个例子已被吉姆·柯林斯和莫滕·汉森（**Morten Hansen**）用来说明为何有的公司能在极端困难的情况下逆风飞扬，而其他公司却不能。这两位作者从20400家公司里筛选出7家，他们发现那些最成功的公司在预测未来方面并不比那些不那么成功的公司高明多少。相反，它们承认自己无法预测意外情况，因此事先做好了更充分的准备。

2.比预计时间多留50%

我认识一个人，她总以为自己到商店只要5分钟就行，因为她有一回真的只用5分钟就到了那里。但实际上，没有10~15分钟根本就不够。这件事本身没什么大不了的，但不幸的是，她对于生活中的绝大多数事情都是这么估计的。结果，迟到便成了家常便饭。更糟糕的是，她总是处于压力之下，这让她感到内疚不已。在这样一种循环之中生活久了，她甚至都意识不到自己正处于一种常态化的压力之下。这种状况还影响了她的健康。然而，她依旧固执己见，认为只要5分钟就能到达商店，只要半小时就能完成电话会议，一星期就能写完重要报告，或者任何其他事情她都可以完成——这些她的确偶尔可以做到。但是这让她和她周围的人付出了很大代价。其实她只需要准备一些缓冲，就可以把那些仓促完成的事情做得更好。

你是否也曾低估过完成任务所需的时间？如果是，那么你就和很多人同病相怜了。这种现象被称作“规划谬误”（**Planning Fallacy**）。这一术语由丹尼尔·卡尼曼于**1979**年提出，它指的是人们倾向于低估任务完成所需的时间，即便是在以前已经承担过相同任务的情况下也是如此。在一项研究中，37名学生被问到他们认为自己需要多长时间才能完成毕业论文。当这些学生被要求在“一切尽可能顺利”的前提下估计论文

完成时间时，他们的平均估算值是27.4天。而当他们被问到如果“一切情况糟得不能再糟”，那么完成论文需要多长时间时，他们的平均估算值是48.6天。最后，这些学生实际完成毕业论文的时间平均为55.5天，只有30%的学生在他们估算的时间里完成了任务。令人不解的是，人们一方面承认自己有低估的倾向，但与此同时又总是相信自己当前所作的估计是准确的。

在关于人们为何会低估任务完成时间的诸多解释中，我认为“社会压力说”是最有意思的一种。一项研究发现，如果人们以匿名的方式估计任务完成所需的时间，那么他们就不会对规划谬误感到愧疚。这意味着，通常我们实际上知道自己无法在给定的期限内完成任务，但是我们不愿意向别人承认。

不管原因如何，结果都无外乎当我们做到时往往已经晚于自己预先所说的时间：开会迟到，工作迟交，账单迟结，等等。这样，本来可以顺利完成任务，执行起来却变得阻力重重，令人沮丧不已。

要避免出现这样的情况，方法很简单，只需在我们估算的任务或项目完成所需的时间上加50%的缓冲量即可（如果觉得50%的缓冲量太过奢侈，那就想想实际完成时间比预计时间多50%的情况是多么常见吧）。如果你为一个电话会议留了一小时的时间，那么就再额外增加30分钟吧。如果你估计带儿子去参加足球训练需要10分钟，那么就在练习开始前15分钟离开家吧。这样做不仅有助于缓解担心自己会迟到时感受到的压力（想象一下，如果我们时间充裕，那么堵车时就大可不必着急心慌），而且如果发现任务比预想的完成得更快、更轻松（不过对于绝大多数人来说这种事情发生的概率很小），那么这多出来的时间就让人感觉像是得了一笔奖金一样。

3.做好情景规划

沃顿商学院风险管理和决策流程中心（The Risk Management and Decision Processes Center at Wharton）主任艾万·米歇尔-克嘉（Erwann Michel Kerjan）提出，每一个人，上至国家元首，都应该建立风险管理策略。比如，他曾与世界银行合作寻找世界上抗冲击能力最弱的国家，结果发现在85个国家里排名第58位的摩洛哥建立了应对风险领域的行动计划。

在与一些国家政府合作建立风险管理策略的过程中，艾万建议他们从5个问题开始：

- ①我们面临的危险是什么，在哪里？
- ②哪类财产和群体将暴露于危险之中，暴露程度有多大？
- ③它们有多易受危险攻击？
- ④这些危险将给个人、企业和政府预算造成何种财务负担？
- ⑤如何做最优投资以降低危险并增强经济、社会应变能力？

我们也可以将上述5个问题运用到我们自己建立缓冲的尝试中。想好自己要在工作和家庭中完成的最重要的任务是什么，然后问下面这5个问题：

- ①在这个任务中你面临的危险是什么？

②最糟糕的情形是什么样的？

③它将产生什么样的社会影响？

④它将在财务方面造成什么样的影响？

⑤怎样做最优投资以降低风险并增强财务和社会应变能力？

你对关键性的第5个问题的回答将会引导你去建立缓冲，来保护自己免受未知事件的影响——或许是增加20%的项目预算，或许是找一名公关人员来应对可能出现的负面报道，或许是召开一次董事会会议应付股东期望。

我们永远无法预测所有情形和所有可能性，并做好万全准备，未来实在是太难预测。精要主义者对于这样的现实不是采取回避的态度，而是通过建立缓冲来减少意外事件带来的摩擦。



- 作完善的准备。
- 比预计时间多留50%。
- 做好情景规划。

CHAPTER 15

SUBTRACT

Bring Forth More by Removing Obstacles

要素十二

扫除，锁定你的“最慢行进者”

为学日益，为道日损。

——老子



非精要主义者

堆积权宜之计

付出更多



精要主义者

扫除进步的障碍

实现更多

在商业小说《目标》（*The Goal*）中，三个月内要把小工厂扭亏为盈的任务压得主人公艾利克斯·罗格（Alex Rogo）喘不过气来。刚开始的时候，他看不到一丝可能性。后来，一位教授指导他说，只要他能找到这家工厂的“瓶颈”在哪儿，就能在短时间内创造不可思议的进步。所谓的瓶颈，他知道就是那些阻碍整个体系运转前进的障碍。教授告诉他，就算他对厂里其他各个方面都做出改进，只要不能破除这些瓶颈，工厂就很难有实质性的起色。

在努力弄清楚这些话的含义的时候，他和他的儿子及一帮朋友进行了一次徒步旅行。作为童子军领袖，他负责在太阳下山之前，把所有男孩子都带到营地。但是，所有参加过这种旅行的人都知道，让一群小男孩都以同样的速度行进并不掉队，这实际上是很难的。艾利克斯很快就碰上了一个问题：有些童子军走得很快，而另一些却走得很慢。特别是一个叫赫比的小男孩，走得最慢。结果，走在队伍最前面的男孩和落伍的赫比之间，距离相差好几公里。

为了解决这个问题，起初艾利克斯用的办法是让走在前面的那些人停下来，等等落后的人。这办法让队伍在短时间内聚了起来，但是一旦又开始前行，同样的差距就再次形成。

因此，艾利克斯决定换一种方法。他把赫比放在这组人的前面，把其他孩子按速度进行排列：从最慢到最快。虽然，把走得最快的人放在队伍的最后面是违反常理的做法，但是这么做能让这群人开始以整齐划一的队列行进。每个男孩都可以跟上自己前面的那个男孩。这种方法的好处在于，他马上可以看到整个队伍的情况，男孩子们可以安全、同时到达营地。不利之处在于，这个队伍现在是以赫比的速度在前进，因此他们将会晚到。那么，艾利克斯该怎么办呢？

他找到的对策是，想尽一切办法为赫比减负。因为走得最慢的小孩

恰恰在队伍的最前面，如果赫比每小时都能走快一米，整个队伍就能以同样的增速更快地到达营地。

这个发现对艾利克斯来说意义重大。赫比的任何一点加速，哪怕极其微小，也会对整个队伍的行进速度产生立竿见影的效果。因此，他把赫比背包里的东西（他随身带的多余的食物和用品）都拿了出来，并分给了队里的其他人。这种方法马上奏效，队伍行进速度立刻加快。他们最后提前到达了营地。

仿佛从中顿悟，艾利克斯明白了如何运用这种方法来让自己的工厂扭亏为盈。他不再费力地去改进厂里的一切状况，而是找出其中的“赫比”：工厂运行程序中较之其他部分更慢的那个部分。为了找到这个“赫比”，他找出了材料堆积最多的那台机器，并想方设法提高它的效率。这还反过来促进了倒数第二个“最慢行进者”提高效率。依此类推，整个工厂的生产效率都得到了提高。

谁是你的“赫比”

你工作或生活中“最慢的行进者”是什么？阻碍你取得重要成就的障碍是什么？通过系统定位并移除这个“掣肘”，你将大大减小阻碍你执行重要事务的摩擦力。

但是它并不能随随便便实现。仅仅是漫无目的地找到一些需要改进之处，最多只能实现一些无关紧要的、短期的改善，最差的结果，还可能是让你在不重要事务上耗费时间和精力。但是，如果你真的想提升系统整体的运行效果，不管这个系统是生产程序，部门工作流程，或是日常生活中的常规性事务，你都需要找出这个“最慢行进者”。

非精要主义者在执行的时候，总是被动应战，率性而为。因为，他们总是在遭遇危机后才被动反应，而不是提前预测危机，所以就不得不

采取应急之计——相当于把手指插进堤坝的漏洞，幻想整个大坝不会坍塌。由于擅长使用锤子，非精要主义者认为一切都是钉子。这样，他们用的力气越来越大，但结果却导致了更大的阻力和更多的挫折。的确，有些情况下，你给人的压力越大，他或她的反抗力也就越大。

精要主义者不会默认权宜之计。显而易见的和眼前的障碍不是他们的目标，他们关心的是那些延缓前进步伐的阻力。他们会问：“重要事务实现道路上的拦路虎是什么？”非精要主义者埋头于与日俱增的压力和堆积如山的解决方案，而精要主义者却以一次性投资来解除障碍。这种方法的意义不仅仅在于解决问题，而且也有助于让你以更少的付出达到效益最大化的目的。

三步扫除“最慢行进者”

亚里士多德谈到过三种形式的工作，但在现代社会中，我们倾向于只强调两种。第一种是理论性工作，其终极目标是探求真理；第二种是实践性工作，其目的是行动。但还有第三种：创制性工作（Poietical Work）。哲学家马丁·海德格尔（Martin Heidegger）将创制描述为“展现”（Bringing forth）。这第三种类型的工作就是精要执行方式。

人们很少花时间认真考虑什么样的努力会产生成效，什么样的不会。但就算我们认真考虑了，也很容易认为执行意味着加法而不是减法。如果我们想要出售更多的产品，那么我们会增加销售人员。如果我们想要生产更多的产品，那么我们会扩大生产。这种方法可以找到清晰的证据支持。然而，要改善效果，还有其他思路可循。精要主义者不关注需要增加哪些精力和资源，而是关注哪些限制和障碍需要连根拔起。那么，他们是怎么做的呢？

1. 弄清核心目标

只有弄清了要达到的结果是什么，我们才能知道应该扫除的障碍是什么。如果不知道自己真正想要实现什么，所有的改变都是武断妄为。因此，不如问问自己：“如何知道自己已经成功了？”在本章，让我们假定你的目标是，在本周四下午2点之前，准备好一份15页的报告草案，并用电子邮件发给客户。注意：这结果非常精确，一点儿都不含糊。

2.锁定障碍

不要仓促上马，三思而后行。问问自己：“在我完成任务的道路上有哪些障碍？”“是什么阻碍着我完成任务？”把这些障碍列一个清单。它们可能包括：缺乏必要的信息、精力不足、追求完美。“哪一个障碍扫除之后，其余的障碍会立刻消失？”根据这个问题，来给所有的障碍排列优先顺序。

在锁定你的“最慢行进者”时，需要记住的关键一点是：即便是一些“有成效的”事情，比如为了获取信息而进行研究或者给别人发电子邮件，或是修改报告以求第一稿就达到完美，这些都可能是障碍。记住，要达成的目标是完成报告的草案。任何延缓此目标执行的活动都是值得怀疑的。

妨碍重要目标实现的障碍多种多样。然而，在任何一个特定的时间，有且只有一个头等要事，如果主要障碍岿然不动的话，盲目扫除障碍不会产生任何效果。以我们的任务为例，如果获取信息是你的主要障碍，你可以雇人替你做些研究，但前述报告的起草依然没有获得进展。因此，正如艾利克斯最先修理了效率最低的那台机器，紧接着修理效率倒数第二低的那台，然后依此类推一样，不要试图一下子修好所有的机器，我们也应该一个一个地消除障碍物。

3.扫除障碍

如果你的“最慢行进者”是你追求报告的完美性，你有很多很多的想法能让报告更完善，但是在现在的情况下，你的核心目标是交上初稿。因此，要扫除障碍，你必须考虑“完成比完美更重要”，忘记“这报告必须完美”。允许自己不用对第一稿进行润色。主要障碍一扫除，这项工作的任何其他方面都会变得更容易。

“最慢行进者”甚至还可能是一个人——不管是在一个项目上给你开绿灯的老板，还是不批准预算的财务部门，抑或是一个不愿在签字栏里签字的客户。为了减少与他人的摩擦，你要采用“用蜂蜜捕捉更多苍蝇”的办法。给他发一封电子邮件，但不要问他是否已经为你做了X事（很显然，他没有做），而是去看看他，问问他：“你在做X事的时候碰到了什么障碍和困难，我该怎么帮你扫除这些障碍？”不要纠缠别人，而是要真诚地提供支持。这样做，你会比只发一封邮件然后提一个要求，得到更加热情的回复。

**An Essentialist
produces more
——brings forth
more——
by removing
more instead
of doing more.**

精要主义者产出更多
——展现更多
——其方法是扫除更多，
而不是做得更多

我的孩子们还很小的时候，我正在读研究生，我妻子被每天从早到晚照看孩子的种种要求折磨得精疲力竭，她真的不知道该怎么办。当时，我正在读关于约束理论（The Theory of Constraint）的书，所以对此特别在意。当我们运用上述步骤时，发现制约我妻子为孩子成长作出最大贡献的主要障碍在于，她没有时间去计划、思考和锻炼。毕竟，要带三个很小的孩子，她几乎不可能有不被打搅的时间。因此，我们努力去终结这个障碍。我放弃了很多课外活动，晚上待在家里。我们还找了人，一周来照看孩子几个小时。这样一来，当我们和孩子们待在一起的时候，就能更加全身心地投入。换句话说，我们俩虽然做得少，但收效却很好。

扫除障碍并不一定困难重重，也不需要超人般的努力。相反，我们可以从小处开始。就像移走山顶上的一块卵石，只需轻轻一推，动量就自然产生了。



- 弄清核心目标，确定扫除的障碍是什么。
- 锁定障碍，但不要试图一下子扫除所有障碍。
- 扫除障碍，从小处开始。

CHAPTER 16

PROGRESS

The Power of Small Wins

要素十三

进步，小胜的价值

每天都做一点接近美好未来的事。

——道格·法尔博



非精要主义者

大处着手，收获甚微
追求华而不实的结果



精要主义者

小处着手，成果巨大
为小胜喝彩

想 想最近一次你在开车时，被警察示意靠边停车的事。你当时是否心里纳闷过：“这是一张好单子，还是一张坏单子？”显然不大可能。因为每个人都知道，通知单是坏的，不是吗？但是，加拿大里士满警局一个富有创见的警察认为，这种观念应该被颠覆。

打击犯罪有非常明确的办法：制定新的、更加严酷的法律，实施更重的量刑，采取零容忍手段。换句话说，就是做更多我们已经做的——仅仅是更加强有力而已。多年来，里士满警察局就是默守着这些在所有地方的警务管理系统中核心的、长期存在的成规，同时也承受着典型的后果：高达65%的累犯率和呈螺旋式上升的青少年犯罪率。这种情况始终存在着，直到沃德·克拉彭（Ward Clapham），一个年轻的、具有前瞻思维的新长官出现并挑战它们。他提出了质疑：为什么我们所有的警务工作都必须如此的被动、消极并且滞后于现实？如果我们不把精力仅仅集中在犯罪行为实施之后的犯人抓捕和进行更严酷的处罚上，而是用大量的资源和精力，将犯罪行为扼杀于其发生之前，那将产生什么样的结果？引用一句前英国首相托尼·布莱尔（Tony Blair）的话来说，如果我们不仅对罪行强硬，而且对犯罪的原因也持强硬态度，那会如何？

就在这些问题之中，孕育出了“发放肯定票”的新颖想法。运用这种方法，警务人员不是专注于抓捕实施犯罪行为的年轻人，而是寻找做好事的年轻人，这些好事可以小到把垃圾丢到垃圾桶里而不随地乱扔，骑车时佩戴头盔，在指定地点滑滑板，或者按时到校等。通过给这些年轻人发肯定票来表扬这些积极行为。这种票当然不像违章停车罚单那样是用来处罚的，而是可以换成一些小小的奖励，比如免费观看电影或免费参加当地青少年中心的活动，这些健康有益的活动同时也具有防止年轻人在街上无所事事从而滋生事端的积极效果。

那么，里士满的这种非常规的方法对重新规划警务工作的作用如何？结果证明，效果好得令人惊讶。虽然没有产生立竿见影的效果，但

是他们把这种方法当作长期战略来实施，十年以后，肯定票制度把累犯率从60%降到了8%。

人们可能想象不到，在警察局这样的地方，也可以找到精要主义的用武之地，但实际上沃德创立的肯定票制度为轻松执行提供了重要的借鉴。

非精要主义者采用的方法不外乎是把一切都做大：什么都要做，什么都要得到，什么都要搞定。非精要主义遵循着错误的逻辑：努力越多，成就越大，但现实是，天上的星星我们越是伸手去够，就越难让自己离开地面。

精要主义者则采用截然不同的方法。他们不会贪多求全，急于求成，平均用力，而是从小处着手，庆祝进步。他们摒弃好高骛远，鄙视华而不实，努力追求在重要领域中的每一个微小而平实的进步。

通过捕捉并奖励这些获得“小胜”的人们，沃德的方法让人们看到了鼓舞进步所具有的力量。这里曾经发生过一个感人的故事。有个男孩子救了一位险些被车撞到的姑娘，警官把他拉到路边，给了他一张肯定票，并对他说：“你今天干了一件伟大的事情，你将来一定会有所作为。”这个男孩回家后，把这张肯定票贴到了墙上。几个星期以后，他的养母问他是否打算把这张票兑现奖励。令她吃惊的是，那男孩说，他永远不会。因为一名成年人告诉过他，他可以出人头地，这比免费吃顿比萨和打场保龄球有价值多了。

这种积极互动以每年4万次的规模增长，并持续了10年，这样你就能认识到为什么它可以创造不同了。每一次，当年轻人因为做了好事而得到肯定和称赞时，他就会有更大的动力去坚持做好事。久而久之，做好事就习惯成自然，做起来毫不费力了。

斯坦福监狱实验与英雄想象工程

人们常常觉得，要有所改变，就得从某件宏伟的大事开始。就像我认识的一个经理人，他大张旗鼓地宣布自己要给女儿们建造一个精巧的玩具屋，可是后来，由于目标过于庞大和浮夸，他不得不放弃了这个不能承受的重负。在这个问题上，人们常常被一种逻辑所迷惑：要干大事，就必须从大处着手。然而，想想企业里所有那些“庞大的”、激动人心的壮举，最后还不是像那位经理人的玩具屋计划一样，没有任何成效。

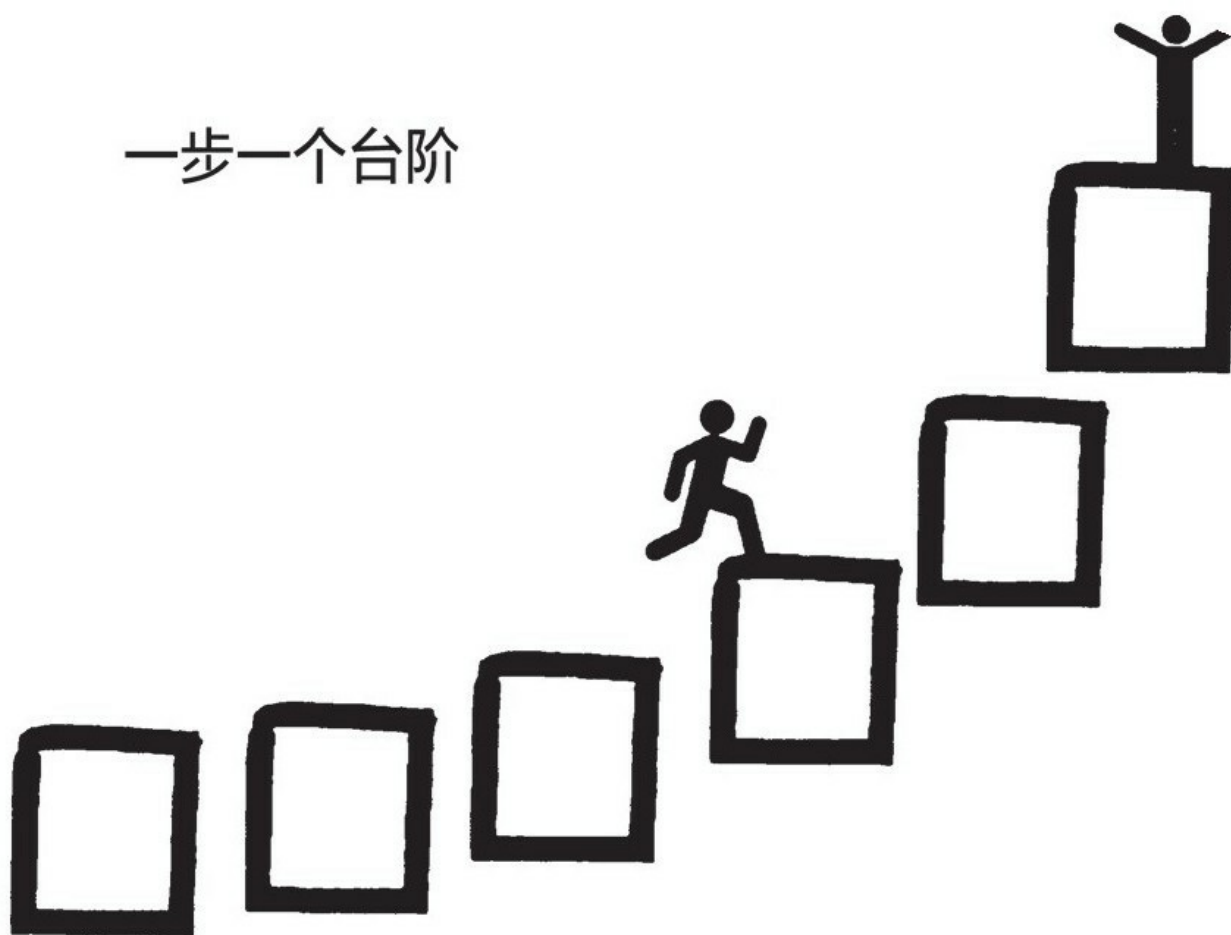


研究表明，在所有形式的人类动机中，最有效的莫过于进步。为什么？因为一个小小的但却实实在在的胜利能够创造推动力，并让人们对自己未来的成功持肯定的信念。1968年，弗雷德里克·赫茨伯格（Frederick Herzberg）在《哈佛商业评论》上发表了一篇题为《再论如何激励你的员工》（*One More Time: How Do You Motivate Employees?*）的文章，此文是《哈佛商业评论》史上最受欢迎的文

章。在文中，弗雷德里克·赫茨伯格写到，研究显示，成就和对成就的认可，是人类行为的两大主要内在驱动力。，特丽萨·阿马比尔

（Teresa Amabile）和史蒂文·克莱默（Steven Kramer）收集了数百人的数千条不具名日记条目，日记描述了这些人工作日的状态。根据这成千上万的回忆，阿马比尔和克莱默得出结论，“每天的进步——哪怕只是小小的一点进步”，都可以对人们的感受和表现产生巨大的影响。“在所有有助于提升工作日的情绪、动机和认知的激励物中，最重要的莫过于取得有意义的进步。”他们这样说道。

一步一个台阶



不要再从大处着手，然后除了浪费时间和精力，一无所获。要真正做好重要的事情，必须从小处着手，创造动力。然后，运用这种动力创造下一个胜利，再下一个，长此以往，直至取得重大突破。如果这

样做，我们就能毫无阻力、毫不费力地取得进步，突破性进展就仿佛在一夜之间降临。曾任斯坦福大学教授的教育学家亨利·艾林（Henry B. Eyring）这样写道：“经验告诉我，个人和组织是这样完善的，最佳着眼之处，莫过于那些微小的改变，它们就发生在我们常做的事情上。持之以恒和不断重复是力量之源。”

**To really get
essential things done
we need to start
small and build
momentum.**

要真正做好重要的事，
必须从小处着手，
创造动力。

有一回，在和美国心理协会前主席菲利普·津巴多博士（Dr. Philip Zimbardo）¹²共进午餐时，我了解到他是著名的斯坦福监狱实验的策划人。1971年夏，津巴多请了一群健康的斯坦福大学学生参加实验，给他们随机分配了“狱警”和“犯人”的角色，然后把他们关进在斯坦福大学地下室里临时搭建的“监狱”里。没过几天，这些“犯人”开始表现出抑郁和极度紧张的症状，而“狱警”则变得残酷无情和虐待成性（出于显而易见的原因，实验提前结束）。这其中的问题在于，被人当作犯人和狱警来对待，在短短几天就产生了一种动力，它激励着主体要表现得像犯人和狱警那样。

斯坦福监狱实验颇具传奇色彩，阐述其内涵的著述也颇丰。但令我感到好奇的是：如果仅仅以特定方式对待就能影响这些斯坦福学生，让他们逐步采取这些负面的行为，那么同样的情况是否也能为更加积极的行为所用？

的确，现在津巴多正在借鉴这种思路尝试一种大型的社会实验“英雄想象工程”。这项工程的运作逻辑是通过灌输英雄主义原则，增加人们做出勇敢行为的可能性。津巴多相信，通过鼓励和奖励英雄行为，我们可以有意识地精心创立一种制度，让英雄行为最终变得水到渠成、轻而易举。

鼓励小胜的三个技巧

我们可以选择。我们既可以下功夫建设一种制度，让执行精要之事不费力气，也可以听命于一种制度，哪怕它实际上只会让执行要事变得困难重重。沃德的肯定票制度属于前者，它取得了很好的效果。在为人生设计制度的时候，我们也可以将同样的原则运用到自己所面临的选择中去。

我和妻子安娜试着在我们的子女教育机制中运用这样的理念。我们曾一度很担心家里的屏幕时间究竟多到了什么样的程度。在对电视、电脑、平板电脑和智能手机的轮番使用中，孩子们的时间就这样轻易浪费在了意义不大的娱乐活动中。我们试图改变他们的习惯，但是正如你能想象到的那样，我们遇到了阻力。不管什么时候，我们要关电视或者想要限制他们的屏幕时间，孩子们总是抱怨不休。作为父母，我们必须有意识地管理这种状况。这让我们不得不放下重要的事情。

因此，我们采用了一种代币制度。每周一开始，孩子们可以领到10个代币。每个代币可以换30分钟的屏幕时间，或者在周末的时候换50美分，总共加起来是每周5美元或者5个小时的屏幕时间。如果哪个孩子读书30分钟，他或她就可以赚到一个额外的代币，也可以换成屏幕时间或者钱。结果简直不可思议：一夜之间，屏幕时间下降了90%，读书时间上升了同样的百分比，而且我们必须投入到管理中的努力越来越少。也就是说，不重要的活动显著减少，而有意义的活动显著增加。当初只不过投入了一点点努力来建设一种机制，而现在它顺利无阻地发挥着效用。

在家里和在工作中，我们都可以建立这样的机制。关键在于，要从小处着手，鼓励进步，庆祝小胜。以下是一些技巧。

1.关注最低可行性进步

在硅谷流行着“完成比完美更重要”的理念。这并不是让大家去制造垃圾。其中的思想，在我看来就是不要在不重要的事情上虚掷光阴，而是要把事情做成。在企业界，这种想法被表述为创造一种“最低可行性产品”（Minimal Viable Product）。它指的是：“对目标客户有用且有价值的尽可能简单的产品是什么？”

同样，我们也可以采取“最低可行性进步”（Minimal Viable

Progress) 的方法。问问自己：“对于我们努力要完成的重要任务，有用且有价值的最小进步是什么？”我在本书的写作中就采用了这种方法。比如，当我还在探索阶段的时候，在动笔（或敲打键盘）之前，我会在Twitter上面分享一个小小的想法（我的最低可行性产品）。如果能够在那上面引起共鸣，我就会给《哈佛商业评论》写一篇博客文章。通过这种只需要花费很少精力的迭代过程，我能够知道应该在什么地方搭建一座桥梁，让自己所思考的事情和同别人生活最相关的事情之间建立起联系。

皮克斯动画工作室把这个方法用在了他们的电影中。这些电影的创作不是开始于脚本，而是开始于情节串联图板，或者叫作连环漫画版电影。他们尝试各种想法，并观察哪些有效，在小范围内进行数百次实验。然后，做出一个影片放映给小规模人群看，以便提前获得反馈。这样，他们就可以用尽可能少的投入，获得尽可能快的长进。皮克斯动画工作室的首席创意官、迪士尼幻想工程的首席创意顾问约翰·拉塞特（John Lasseter）说：“实际上，我们不是在完成电影，而是在发行电影。”

2. 做最低可行性准备

应对一个重要目标或一个最后期限，有两种截然对立的方法。一种是早动手，小步前进；另一种是晚动手，一日千里。“晚而大”意味着要在最后一刻完成全部，通宵加班，把它做完。“早而小”意味着以尽可能少的时间投入，尽可能早地开始。

你只要在一个项目或任务到期前两周仅仅投入10分钟时间，就不用到最后时刻花大把时间焦头烂额地仓促赶工了。有了目标或者期限，就问问自己：“我此刻能够做的最低可行性准备是什么？”

有一位领导者，他非常善于发表激动人心的演说。他解释说自己的

秘诀就是在演说的6个月前就开始为这些鸿篇阔论做准备。不是准备整个讲稿，只是开始而已。如果未来几个星期或几个月内你有一个重大报告要做，那么你现在就打开文件，花4分钟时间，想到什么就写什么。然后，关闭文件。不要超过4分钟。开始就够了。

一位在纽约的同事有一种很简单的方法：任何时候她安排了会议或者有电话要打，就会很精确地花15秒钟备忘一下那个会议的主题。这样在开会那天早上，当她坐下来准备发言要点的时候，可以参考这个备忘录。她不需要规划整个会议的议程，只要花几秒钟提前准备，就会大有收益。

3.直观地嘉奖进步

你记不记得，在你5岁的时候，学校组织了一场募捐活动？你记得那个用来直观地展示学校向目标迈进的大温度计吗？你是否记得，看到温度计一天天上升时，心情多么激动，多么欢欣鼓舞？或者，你的父母给你准备了一个星形图，每次你吃了菠菜，或按时睡觉，或打扫了房间，你就能得到一颗星星，然后很快你就不用什么激励物也在做这些事情了。

直观地看到自己正在朝着目标进步，会产生一种巨大的力量。不管是在家里，还是在单位里，不要不屑于把这样的方法运用到你自己的重要目标上。

相较于设定庞大的、高远的和不现实的目标，从小处着手，奖励进步，才能成就更多。作为一种额外奖励，这种明确地强化成功的行为，能让我们在过程中收获更多的快乐和满足。



- 关注最低可行性进步。
- 做好最低可行性准备。
- 直观地嘉奖进步。

CHAPTER 17

FLOW

The Genius of Routine

要素十四

心流，常规造就天才

常规，对智者而言，与雄心壮志无异。

——W.H.奥登



非精要主义者

凭借蛮力执行重要事务

让非重要事务成为默认设置



精要主义者

设计一种处理重要事物的
常规，让执行毫不费力

让重要事物处于默认状态

奥运游泳名将迈克尔·菲尔普斯（Michael Phelps）在获得2008年北京奥运会金牌前的几年里，他在每一次比赛中都遵循着同样的常规。提前两小时到达比赛地点，始终按照一套精确的模式做伸展放松运动：800次混合式练习，50次自由式练习，600次扶板打腿练习，400次拉浮标练习，还有一些其他练习。做完热身运动后，他会把身体擦干，戴上耳机，在按摩台上坐下来，不躺下。从那一刻起，他与教练鲍勃·鲍曼（Bob Bowman）不再有任何交谈，直到比赛结束。

比赛开始前的45分钟，他会穿上比赛服。开赛前30分钟，他跳进热身池，游600~800米。只剩10分钟的时候，他会走进准备室，找一个单独的座位，不同任何人挨着。他喜欢让座位两边都空着，这样便可以放自己的东西：一边放护目镜，一边放毛巾。听到他的比赛广播后，他会走到比赛跳台边，一如既往地做两种伸展运动，先直腿，再屈膝，每次都是先活动左腿。然后摘下右边的耳塞。叫到他名字的时候，他会把左边的耳塞也摘下来，踏上跳台，每次都是从左边上去。他会用毛巾把跳台擦干，每次都是如此。然后，他会站着，用双手拍打后背的方式摆动双臂。

菲尔普斯解释说：“这不过是常规。我的常规。它贯穿我的整个人生。我不想改变它。”教练鲍勃为他设计了这种体能练习常规，但这并不是全部。教练还给菲尔普斯睡觉时和一醒来时想什么也设定了常规，称之为“看视频”。当然，并不是真的有录像带，而是对完美比赛进行一次想象。在精致的细节和缓慢的动作里，菲尔普斯想象着每一个瞬间——从跳台上的起始位置，到每一次划臂，一直到浮出水面，沐浴着胜利，任由水珠从脸上滑落。

这样的精神常规活动，菲尔普斯并不是偶尔才做，而是每天上床睡觉前、每天早上一醒来就做，几年来一贯如此。当鲍勃要在练习中给他增加难度的时候，他就会大喊：“把录像带放好！”然后，菲尔普斯就会

挑战自己的极限。最后，这种精神常规已深深植入了菲尔普斯心中，以至于鲍勃在比赛前几乎不需要小声提醒“准备好录像带”，菲尔普斯就已经随时候命“按下播放键”。

在谈及这种常规时，鲍勃说：“如果你要问菲尔普斯，在比赛前脑子里想的是什么，他会告诉你，他其实什么都没想。他只不过是在走这个程序。但这么说也不对，其实是他的习惯在主导着他。当比赛来临时，他的计划已经进行了一半多，每一步都是成功的。像计划的那样做完所有的伸展运动，像脑子里想的那样进行热身，头上戴的耳机里播放的正是他想听的。实际的比赛，不过是他那天早就开始的程序中的一个步骤，除了胜利，还是胜利。它是自然延伸出来的。”

众所周知，菲尔普斯在2008年北京奥运会上获得了创纪录的8枚金牌。在他创下这个辉煌战绩的几年后，我在北京访问时，禁不住好奇，为何菲尔普斯和其他奥运选手，都能看似毫不费力地创造出这些惊人的运动奇迹？当然，相比世界上其他所有的运动员，这些奥运选手练习时间更长，训练得更艰苦，但是当他们跳入泳池，踏上跑道，或者走进溜冰场时，看上去是多么轻松自如。这种自如还不仅仅是艰苦训练的自然延伸，它同时也证明了正确的常规会给人们带来多大的益处。

非精要主义者认为，重要的事情都是被逼着做成的。执行不过是一种很生硬的努力。为了把事情完成，人们埋头苦干，用的是九牛二虎之力。

精要主义者的做法则截然不同。他们设计一种常规，让自己所认定的精要事务沿着默认的路径实现。当然，在某些情况下，精要主义者仍然需要付出努力，但是由于有了正确的常规，每一分努力都会产生事半功倍的效果。

让事情变得看似简单

常规，是移除障碍最有力的工具之一。没有常规，人们就被非重要事务牵扯得没有余力。但是，如果我们能够创造一种处理重要事务的常规，执行就如同拥有了自动驾驶仪。这样，重要之事就不用我们有意求之，而是水到渠成，无须思虑。我们也就大可不必日日耗费宝贵的精力，安排何事为先，何事为后。但最开始时我们需要稍费些力气，创造一种常规，接下来就只需遵行即可。

大量的科学研究解释了常规使难事变易的运作机制。一种简单的解释是，当我们重复完成某项任务的时候，神经元，或者神经细胞，会通过名叫“突触”的通信网关建立新的联系。通过重复，联系得到了加强，而且大脑能够更加容易地激活它们。比如，你学一个新单词，就需要时不时地重复几次才能掌握它。以后，你要想起这个词，就必须激活这些突触，直到最终你不用有意识地去想也能知道这个词。

同样的过程也可以用来解释，如果我们每天开车从A点到B点，最后我们想都不用想就能开到；或者同样的菜烧过几次之后，就不用再看菜谱了；还有，为什么脑力工作会变得一次比一次容易。在重复中，我们掌握了常规，任务就变成了习惯。

执行重要事务的能力也是熟能生巧，这就像其他任何能力一样。试想一下，你第一次在工作中履行一种重要的职责。刚开始时，你感觉自己就是个新手，没有自信，笨手笨脚。你把所有的意志力都用在了集中精力上。紧接着，决策疲劳就出现了，你还可能动不动就分心。这些都是再正常不过的现象。然而，一旦你一遍遍地重复履行那种职责，你就会变得自信。你不再会偏离目标。你能更好更快地履行职责而不用那么专注、那么费力了。常规的作用力来自大脑掌控全局的能力，它不断生长，直至成为一个完全无意识的过程。

常规还有另外一种认知优势。一旦脑力活动转入基底神经节，精神空间就得以释放，可以专注于新事物。这让我们能够自动驾驭某项重要

事务的执行，并在同一时间积极从事另一项，而关注程度和贡献水平并不会降低。“实际上，大脑活动开始变得越来越少，”《习惯的力量》

（*The Power of Habit*）一书的作者查尔斯·杜希格（Charles Duhigg）这样说道，“大脑可以完全停工……这是一个真正的优势，因为它意味着你可以把全部精神活动投入到其他事情上去。”

有些人觉得，常规听起来像创造能力和创新能力的死亡——厌倦的终极表现。人们甚至把它当成乏味和平淡的同义词，就像“这于我而言，不过是例行公事”。当然，有的常规的确会成为例行公事——那些不正确的常规。但是，正确的常规实际上可以提升创造力和创新性，因为它减少了对精力的占用。我们无须再把有限的自律性用于一次又一次地做同样的决策，只要把决策嵌入到常规中，就可以把这种自律调拨到其他精要事务上去。

米哈里·希斯赞特米哈伊的创造性实验，证实了具有高超创造力的人，是如何运用严格的常规来释放自己头脑的。“绝大多数有创造力的人早就发现了自己的睡眠、饮食和工作的最佳节奏，并严格地遵守，哪怕在做别的很有吸引力的事情时也不例外。”米哈里说道，“他们穿衣注重舒适，与人相处讲究意气相投，只做自己认为重要的事情。当然，这些偏好并不讨那些他们不得不与之打交道的那些人喜欢……但是，个性化的行为模式有助于心灵免受那些需要耗费注意力的各种期待的干扰，并有助于集中精力处理重要事务。”

硅谷最有创造力的一家公司的首席执行官，有一种乍一看枯燥乏味、扼杀创造力的常规。他每周一早上9点会召开一个3小时的会议，雷打不动，也从不安排在其他时间。这一规定如此具有强制性，以至于在这家全球化的公司里，所有的高管都知道不能安排任何与这个会议相冲突的旅行。这是一条铁律。乍一看，这没什么特别之处。但是，其独特之处在于这个常规会议所产生的高品质的想法。因为，首席执行官把规

划会议和决定谁参加或谁不参加会议所需要的智力成本降到了零，与会者能够全身心投入到解决创造性问题中去。而且，他的团队也确实轻松地提出了创新、独特的主意和方案。

五招释放正确常规的力量

杜克大学的研究表明，在人们所做的选择中，有40%是深度无意识的，我们并不在一般意义上思考它们。这其中，危险和机遇并存。机遇在于，我们可以发展新的才能，它们最终会成为本能；危险之处在于，我们可能会建立一些常规，它们会产生适得其反的效果。如果缺乏充分的意识，我们会深陷于一些非精要的习惯之中——比如每天早上一起床就查看电子邮件，每天在下班回家的路上买个甜甜圈，或者浪费午餐时间在互联网上胡言乱语，而不是利用时间思考、反省、充电或与朋友及同事沟通联络。因此，我们该如何摒弃那些让人受困于非精要习惯的例行事务，代之以让重要事务执行起来毫不费力的常规性活动呢？

1.彻底检修你的触发器

几乎所有人都有自己想要改变的行为习惯，无论是少吃些垃圾食品，少浪费些光阴，还是少一些烦恼。但当我们努力做的时候，就会发现哪怕是改变一个最简单、最最微小的习惯，都困难得如此令人惊讶、令人不安。仿佛有一种引力无情地把我们拉回到炸薯条暖洋洋的怀抱中，那个有着傻瓜猫图片的网站上，或者对自己无力把控的事情不断升级的担忧里。怎么才能摆脱这些习惯的巨大牵引力？

在关于杜希格的著作《习惯的力量》的采访中，他表示：“在过去的15年中，我们认识到了习惯如何起作用，如何改变我们，科学家们解释了每一个习惯都是由刺激、常规和激励组成。这个刺激就是一个触发器，指示你的大脑进入一种自动模式，以及选择要使用的习惯。然后就

是常规活动，即行为本身，它可以是身体的、精神的，或者情绪的活动。最后是激励，它将帮助你的大脑弄清这个习惯是否值得记住。久而久之，这个循环——刺激、常规和激励，将会因为刺激和激励神经性地相互交织而变得更加无意识。”

这就意味着，如果想改变自己的常规，你并不需要真的去改变行为。你需要找到引发非精要活动和行为的刺激事件，并设法让这个刺激事件与确属精要的事情相关联。比如，如果下班回家路上的那家面包店是导致你买甜甜圈的触发器，下一次路过时，你就要让它提醒你去街对面的熟食店买一份沙拉。或者，如果闹钟一响你就会去查看电子邮件，那么就用它来刺激你起床读书。刚开始时，要抵制进面包店或查看邮箱的诱惑的确是困难的。但是，你每次执行新的行为，每次拿起沙拉，都在强化大脑中这种刺激与新行为之间的关联。很快，你将会下意识地、自动地执行新的常规。

2. 创造新的触发器

如果要想行为有新的变化，就不能局限于现有的刺激事件，我们可以创造全新的刺激事件来触发重要常规的执行。我用这种方法培养了写日记的常规习惯，它给我带来了神奇的功效。在很长一段时间里，我不过是三天打鱼，两天晒网地随便写写。一整天都拖着不写，然后到了晚上，我就会给自己找借口明早再写。但是，到了第二天早上我肯定又不会写，拖到晚上就有两天的东西要写，工作量太大了。因此，我再次拖延，就这样一天天地拖下去。后来，我听某人说，他养成了每天在固定时间写下寥寥数行的习惯。这听上去很容易，但我知道自己需要某种刺激物来提醒自己要在每天的特定时间内写作，否则我还是会像以前那样拖延。因此，我开始把日记本放在包里，紧挨着电话。这样一来，每天晚上当我把手机从包里拿出来充电（这个习惯我已经坚持得很好了）时，我就会看到日记本，它会提醒我在上面写东西。现在，写日记已经

成了我的一种本能，仿佛自然天生，我甚至还对它心怀期待。我已经坚持写了10年，几乎一天都未落过。

3.先啃最难啃的骨头

雷·津恩（Ray Zinn）是麦克雷尔公司（Micrel）的创始人兼首席执行官，该公司在硅谷经营半导体业务。雷·津恩在很多方面的做法与别人截然相反。他以75岁的高龄，在一个20岁大学辍学生当道的行业和城市里打拼。1978年，他和生意伙伴投资30万美元开办了这家公司，自成立起年年盈利（他们合并两家制造工厂的那一年除外）。上市以后，公司的股价从未跌破发行价。雷·津恩将成功归于他们对盈利高度自律的专注。作为首席执行官，他已领导这家公司达35年之久，在这漫长的35年里，雷·津恩始终如一地遵循着一种极其稳定的常规。他每天早上5：30醒来，周末亦然（他已经坚持了50多年）。然后，他锻炼一小时，7：30吃早饭，8：15到单位。下午6：30与家人共进晚餐，晚上10点睡觉。但是，真正让雷·津恩达到个人贡献峰值的是，他一整天的常规活动遵循一个基本原则：先攻克最难的事情。他告诉我：“我们要考虑的事情已经太多了，为什么不建立一种常规来减少一些呢？”

运用上述窍门建立一种常规，让攻克一天中最困难的任务成为每天早上的第一件事情。找到一种刺激物——不管它是出现在写字台上的第一杯橙汁，还是你在手机上定的闹铃，抑或是任何一件你已经习惯一大早就最先做的事情，来触发你坐下来，把精力集中到最难的事情上去。

4.建立混合型常规

在同一时间做同样的事情，而且天天如此，难免令人感到乏味。为了避免这种常规疲劳，你完全可以给一周的每一天都安排不同的常规。前文提到的Twitter和移动支付公司Square的联合创始人杰克·多西就把自

己每周的常规活动安排得有滋有味。他把一周分成几个主题。星期一的主题是开会以及处理公司经营方面的工作。星期二主攻产品开发。星期三致力于营销、传播和发展。星期四关注开发人员和合作伙伴。星期五的主题是公司文化建设。这种常规设计，有助于这家高速成长的新兴公司在面临千头万绪时能够冷静应对，镇定处事。同时，这也让他每天能够集中精力在一个主题上，免得四面开弓，左支右绌。他每周都坚持同样的常规，从不例外，这样时间一长，人们就了解了他的惯例，并能根据这种常规安排各种会议和活动。

5.陈规陋习逐个击破

如果过分着迷于常规的好处，以至于忍不住要一下子修检和改进多种常规，那将是一种灾难。我们已经从本书前一章学到了，要获得大成效，必须从小处着手。因此，先从对每天或者每周的常规作一个改变开始，从那里开启进步之旅。

我并不是在暗示任何一种小的改变都是轻而易举的。很多非精要的常规活动已经根深蒂固，而且带有情绪因素。它们是由一些强烈情绪熔铸而成的。如果有人认为只要动动手指头，就能改掉它们并代之以新的常规，那就太天真了。学习有意义的新技能，从来都不是一件容易的事情。但是，我们一旦掌握了这些新技能并让它们成为无意识行为，就是一次了不起的成功，因为这些技能会一直陪伴我们度过余生。常规也是如此，一旦建立起来，就会成为永不枯竭的馈赠。



- 彻底检修你的触发器。
- 创造新的触发器。

- 先啃最难啃的骨头。
- 建立混合型常规。
- 陈规陋习逐个击破。

CHAPTER 18

FOCUS

What's Important Now?

要素十五

专注，当下最重要的是什么

生活只在当下。如果放弃当下，你将无法细致深入地过好生活中的每一刻。

——行禅师



非精要主义者

满脑子充斥着过去或未来

考虑过去或未来的重要事务

为未来担心，为过去焦虑



精要主义者

专注于当下

专注于当下的重要事务

享受当下

拉里·吉尔威克斯（Larry Gelwix）曾是高地中学的橄榄球队教练，在他执教的36年里，这支球队获得了418次胜利，20次荣膺全国冠军，仅有10次失败。对于他的成功，他是这样描述的：“我们战无不胜。”有着高地中学这样的辉煌战绩，他说起这话来名正言顺，底气十足。但实际上，他真正所指的东西比这些赫赫战功更有意义。当他说“win”的时候，他同时也在用首字母缩写的形式表示一个问题，这个问题左右着他对球员们的期许：What's important now?（当下重要的是什么？）

通过让球员们充分重视此时此刻，专注于最重要的东西——不是下周的比赛，不是明天的练习，不是下一场球赛，而是此时此刻，拉里实现了让赢球几乎毫不费力。但他是如何做到的呢？

首先，当下重要的是什么？这是球员们在整场比赛中时刻不忘的问题。不再纠结于上一场比赛的失利，不再因为担心会不会输掉比赛而劳神费思，因为这样做没有任何益处，拉里鼓励球员们只专注于当下正在进行的比赛。

其次，当下重要的是什么？这个问题让他们专注于自己此刻的表现。拉里相信，取得胜利在很大程度上取决于球员们是否把重心放在自己的比赛或者对手的比赛上。如果球员们开始想着其他球队，那就丧失了焦点。不管是有意还是无意，如果他们开始想以其他球队的战术来进行比赛，那就分散了注意力，分化了战斗力。通过专注于此时此刻的比赛，他们可以统一战术，凝聚战斗力。这种高度团结可以让比赛战略执行起来顺利无阻。

可以肯定的是，拉里对胜负有着一套根本性的精要方略。他告诉球员们：“输掉比赛和被人打败是不一样的。被人打败，意味着别人比你们更优秀。他们更快，更强，更有天赋。”对于拉里而言，输掉比赛有

着不同的含义。它意味着你丧失了重心，意味着你不能在要事上专心致志。一个简单而又强大的理念告诉我们：要实现个人贡献峰值，你就必须全身心专注于此时此地的要事。

除了当下，别无其他

想想如何在你的生活中运用这个理念。你是否曾纠结于过去的错误.....一遍又一遍，就像被按下重播键的录像机一样，无休无止地重放着？你是否因为担心未来而忧虑重重，任由光阴飞逝？你是否花更多的时间来思量以己之力无法控制的事情，而不去考虑自己有用武之地的可控之事？你可曾发现自己脑子里总是忙于筹备下一个会议，下一个任务或者下一阶段的生活，而对现在这个心不在焉？对过去的错误不能释怀，或者对未来要面临的事情颇有忧虑，也是人之常情。然而，在担心过去或未来上花费的任何一秒钟，都在分散抽离着你对此时此地重要事务的关注。

古希腊人有两个词用来表示时间。第一个是柯罗诺斯（Chronos），第二个是凯洛斯（Kairos）。希腊神柯罗诺斯的形象是一位白发苍苍的老人，他的名字在字面上意味着钟表的滴答声，就是我们用来计量的那种物理时间（要争分夺秒高效使用）。凯洛斯就不一样了。尽管它很难准确翻译，但它指的是恰好的，正确的和不同的。柯罗诺斯代表着数量，而凯洛斯代表着质量、契机。后者只有当我们完全投入到此时此刻之中——活在当下，才能体会到。

我们实际上就只有当下，这想来令人费解。照字面上来理解，我们不能控制未来，只能控制现在。当然，我们可以借鉴过去，也可以畅想未来。然而，只有在此时此地，我们才能执行真正重要之事。

非精要主义者总是念念不忘过去的成功或失败，并沉浸在未来的挑战 and 机遇里，因此错过了当下和此刻。他们注意力涣散，没有焦点，心

不在焉。

精要主义者则聚焦于当下。他们体会着在凯洛斯时刻里的生命，专注于真正重要的事情——不在昨天，也不在明天，而是在当下。

最近，我和安娜在一个忙碌的工作日一起吃午饭。通常，在我们见面吃中饭的时候，都会忙不迭地了解对方早上忙了些什么，或者晚上要安排哪些活动，根本无暇享受此时此刻正在吃的这顿午饭。但这一次，上菜的时候，安娜提议做一个实验：我们全身心投入吃饭的这一刻。不再反刍早上的会议，不讨论孩子练完空手道谁去接，也不商量今晚做什么菜。我们有意放慢进餐速度，细嚼慢咽，充分享受此刻的午餐时光。我对此兴趣盎然。

正当我缓慢地吃下第一口食物时，变化出现了。我注意到了自己的呼吸。几乎是无意识地，我发现自己的呼吸在放缓。刹那间，就连时间也仿佛放慢了前进的脚步。我不再心不在焉，而是觉得自己的身心都完全融入了此处。

这种感觉陪伴我一直到了下午，那时我发现了另一个变化。没有什么能分散和打断我的思绪，我可以专心致志于自己的工作。我平静沉着，全心投入手头的任务之中，因此每一项都进展得很顺利。我不再像往常那样把自己的精力分散到很多很多相互抵触的主题上，而是专注于此刻最重要的那一个主题上。完成工作不仅变得很容易，而且还让我享受到了乐趣。这一刻，头脑变得灵光，灵魂也得到了滋养。

小野二郎是世界上最出色的寿司大师，也是大卫·贾柏（David Geld）执导的电影《寿司之神》（*Jiro Dreams of Sushi*）的主人公。在85岁那年，他已经做寿司做了几十年，对他来说，做寿司简直就是小菜一碟。但是，他的故事所要讲述的，绝不是一个熟能生巧、巧能生精的道理。看他做寿司，你会发现一个人完全地沉浸在了这一刻里。

精要主义者就是以这种方式度过整个人生的。正因为如此，他们会把全部精力投入手头的工作，心无旁骛，聚精会神。他们深知，事情的难易取决于努力与否，你若刻苦，它便容易；你若闲散，它便困难。

多任务执行与多点聚焦

从斯坦福大学毕业多年后，我偶然在学校碰到一位老同学。当时我正用学校一间办公室的电脑工作，他走过来跟我打了个招呼。短暂的寒暄之后，他告诉我，他正处于待业状态。他讲了讲自己想找的工作是什么样的，并问我能否帮忙。我于是问了他几个问题，看看怎样才能帮上忙。但是才聊了20秒，他就收到了一条手机短信。然后，他什么也没说，低头盯着手机看，并开始回复短信。面对这番情境，我就像往常一样，停下来等着。

10秒钟过去了。我就站在那里，他继续拼命地发着短信，什么也不说，也不搭理我。出于好奇，我等着看这种情景究竟还会持续多久。但是等了足足2分钟，这对于等待的人来说已经算很久了，我放弃了等待，回到办公桌，继续自己的工作。又过了5分钟，他再次出现，并打断我的工作。这一次，他是想继续刚才的谈话，再次要求我帮他找工作。我知道有个岗位在招聘人，本来打算推荐给他的，但是出了刚才那件事后，我开始犹豫是不是要推荐他去参加面试，因为他可能会突然缺席：身体或许在那里，但思绪已飘远了。

讲到这里，你可能会希望我讲讲同时执行多项任务的坏处——真正的精要主义者从不会试图同时做一件以上的东西。但是，实际上，同时做两件事情轻而易举：洗盘子的时候听收音机，边吃边聊，一面收拾书桌一面琢磨上哪儿吃中饭，一边发着短信一边看电视，等等。

我们没法做的是同时专注于两件事情。当我谈到当下时，我并不是在主张同一时间只能做一件事情。我说的是，同一时间只专注于一件事

情。同时执行多项任务与精要并不冲突，假装自己可以多点聚焦才是精要的大敌。

活在当下的三个技巧

为了充分投入当下的一切，我们该怎么做？以下是一些简单的技巧供参考。

1.想清楚当下最重要的是什么

最近，我在纽约向一个领导团队讲了整整一天的精要主义。我尽情享受了这一天，并自始至终感受着当下。但当我回到房间酒店的时候，我突然感到自己被拉向了四面八方。身边每一样东西都在提醒着我应该做的一切：查看电子邮件，听留言，读一本自己觉得有必要读的书，从现在开始花几个星期准备一个报告，把这一天的经历触发的有意思的想法记录下来，以及其他事情。让我觉得不堪重负的，倒不仅仅是事情的数量，而是在同一时刻很多事情都争着抢着要被优先完成的这种压力，对此我再熟悉不过了。焦虑和紧张让我停了下来。我跪下来，闭上眼睛，问自己：“当下重要的是什么？”沉思片刻后，我明白了，在知道当下重要的是什么之前，重要的是搞清楚“当下重要的是什么”这个理念！

**Multitasking itself
is not
the enemy of
Essentialism;
pretending we
can “multifocus” is.**

同时执行多项任务与精要并不矛盾，
假装可以多点聚焦，才是精要的大敌。

我站了起来，开始整理。我把所有散乱地堆在我身边的东西收了起来，放到该放的位置上去，免得它们分散我的注意力，也免得我每次经过时，它们迫使我按它们的吩咐去做。我关掉了手机。别人没办法给我发短信，这让我如释重负。我打开笔记本记录下这一天，这让我收起了心。我用铅笔给脑海中出现的所有事情列了一个清单。为了明确任务，我问自己：“你要做些什么才能平静地睡去？”我决定，第一重要的事是跟妻子和孩子们联系。接下来要做的就是能让次日早上的头几个小时尽可能轻松一些的事情：安排好叫醒服务和房内早餐，把幻灯片拷到电脑上，熨好衬衣。我划掉了当时并不重要的事情。

当你面临很多很多的任务，以至于不知该从何下手的时候，停下来，深呼吸。投入到当下，问问自己此时此刻最重要的是什么——不是明天最重要的是什么，甚至不是一小时后最重要的是什么。如果不确定，那就把所有争抢你注意力的事情都列出来，然后划掉所有此时此刻不重要的事情。

2. 专注当下，把未来清除出去

把未来清除出你的脑海，你将会更加专注于当下重要的。这个时候，我采取的下一步措施是坐下来，给那些可能重要的事情列一个清单——不是针对当下的。因此，我打开了日记本的另一页。这一次，我问自己：“因为今天发生的一些事，什么事是你某天想去做的？”这个清单不是用来记录确定承诺的，它只不过是用来把这些想法从脑子里清除出去，留在本子上而已。这样做有两个目的：第一，它保证了我不会忘了这些想法，它们也许将来会被证明有用；第二，它减轻了我必须现在就执行它们的那种压迫感和分心的危险。

3. 给每一个清单排列优先顺序

接下来，我给每一个清单都排列了先后顺序。然后，我开始做“当下重要的是什么”这个清单上所列的事情，一次一件。我冷静地完成着清单事项，每完成一件，就从清单上划掉。到了睡觉时间，我不但做完了所有需要在当时完成的事情，而且还因为认真专注而更出色、更高效地完成了任务。

凯洛斯时刻

杰夫瑞·罗杰斯（Jeffrey A. Rodgers）是Cornish&Carey Commerical Newmark Knight Frank的执行副总裁，曾有人向他传授了用停顿来放松的方法。那个时候，杰夫瑞意识到自己每天晚上下班回家时，思绪还牢牢地钉在跟工作有关的事情上。相信你我都有过这样的感觉：身体的确是离开了办公室，但是思绪却还在那里，因为脑子里总是无休无止地循环播放着白天的事情，并且担心着第二天必须要完成的各种各样的事情。

因此，现在当他到达家门口时，他就开始做被他自己称作“用来更新的停顿”。方法非常简单，他只是停一会儿，闭上眼睛，缓慢地深呼吸一次。当他呼气的时候，工作上的事情就烟消云散了。这样，他穿过前门，走近家人时，就已不再思虑重重。这正符合老子的主张：“事善能，动善时。”

越南籍佛教禅宗僧侣一行禅师被称作“世上最心平气和的人”，他花了一辈子的时间，探索如何活在凯洛斯时刻里，尽管叫法可能不同。他主张安住于正念或者保持“初心”。他曾写道：“正念帮助我们回到并安住于当下。每一次你回到当下，就会觉察到自己的幸福处境，幸福就来了。”

专注于当下，对他的一切行为都产生了影响。他每天都和其他僧侣们一道，用整整一个小时来享用一杯茶水。他解释说：“倘若你正要饮

一杯茶，当你举着杯时，你仿佛在吸气，要将思绪带回己身，这样你就完全回归了当下。当你真正存在于当下时，生命也在那里，这杯茶水便是写照。在当下，你是真实的，茶水亦是真实的。你将不会被过去、未来、你的工作或者各种思虑所吞没。你将脱离所有这些苦恼，在自由状态中享受着这杯茶。这便是幸福，这便是平和。”

注意每一天里属于你的凯洛斯时刻。把它们写在日记里。想一想，是什么触发了这样的时刻，又是什么让你从中抽离。既然你已经知道是什么触发了这一时刻，那就努力重新创造它。

适应凯洛斯时刻，不仅有助于实现个人贡献峰值，而且还能让你的生活更幸福。



- 想清楚当下最重要的是什么。
- 专注当下，把未来清除出去。
- 给每一个清单排列优先顺序。

结语

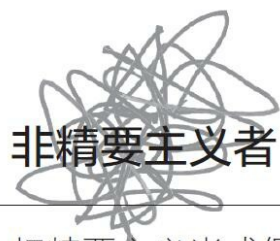
BE

The Essentialist Life

存在，设计你的精要人生

避免碌碌无为。

——苏格拉底



非精要主义者

把精要主义当成偶尔做的事情

带着遗憾的痛楚回顾自己的人生



精要主义者

把精要主义作为处理所有事情的生活方式

过一种有目的、有意义的生活

莫罕达斯·甘地（Mohandas K.Gandhi）早年留学英国，学习法律。他家庭殷实富裕，职业前景诱人，前途一片光明，每天醒来，都是信心满满。对于自己的主攻目标，他十分清楚：取得律师资格，然后过上舒适的生活。然而，正在他抓住时机进行环球旅行的时候，一切都变了。

甘地来到了南非，目睹了那里的压迫。突然间，他找到了一个更加崇高的目标：解放全世界被压迫者。

为了一心一意达到这个目标，他舍弃了生活中所有的其他事情。他把这个过程叫做“把自己降为零”。他穿着手织的土布衣服（印度土布），并鼓励追随者们也这样穿。他3年不看报纸，因为觉得报上的内容只会给他的生活带来不必要的混乱。他简餐缩食35年。他每周都有一天沉默不语。如果说他摒弃了消费主义，那就未免太过保守——离世的时候，他只有不到10样东西。

更重要的当然是他把毕生献给了帮助印度人民争取独立的伟大事业。他从未有意担任过任何形式的政治职务，但他被尊称为印度的“国父”。而他的贡献远不止于印度。甘地逝世的时候，美国国务卿乔治·马歇尔将军（General ; George C.Marshall）说：“圣雄甘地是人类良知的代言人，是他，用谦逊和真理的力量，战胜了帝国。”爱因斯坦评价他说：“在未来的时代，可能极少有人相信，这样一个血肉之躯曾经在地球上匆匆走过。”

甘地的人生是一种真正有意义的人生，这恐怕是毋庸置疑的。

甘地是充分而完全地按照精要方式生活的典范，但我们没有必要为了从榜样中受益而亦步亦趋地去模仿他。我们可以净化非精要生活，拥抱精要生活——用自己的方式，在自己的时代，以自己的能量为限。我

们要过的生活，不是简单足矣，而是要有巨大的贡献和伟大的意义。

精要生活

关于精要主义，有两种思维方式。第一种是把它当成偶尔做的事情；第二种是认为你就是它。在前一种思维中，精要主义是你早已过度填塞的生活中的又一种增加物。而在后一种思维中，精要主义则是一种不同的方式——一种用来处理所有事情的更简单的方式。它是一种生活方式，一种无所不包的生活和领导方法。它就是我们的本质所在。

精要主义在很多宗教传统中有着很深的渊源。释迦牟尼放弃了王族的生活，自愿选择了艰苦修行。他由此获得了顿悟，并创立了佛教。犹太教也有着相似的由来：它的创始人摩西也是放弃了他在埃及作为被收养王子的优裕生活，去往蛮荒之地过牧羊人的生活。就在那里，摩西亲眼看见耶和华的使者从燃烧着的荆棘丛里向他显现，知晓了自己最重要的使命就是带领以色列人脱离在埃及的奴隶生涯。先知穆罕默德过的也是一种精要生活：他缝补自己的鞋子和衣服，给自己的羊挤奶，还让自己的伊斯兰门徒们照做。施洗者约翰也是俭朴生活的典范：住在沙漠里，穿着粗陋的骆驼毛衣服，以蝗虫野蜜果腹。基督教贵格会信徒也是坚定地保持着信仰的精要元素：比如，他们实行“简单的明证”，承诺只过最基本的生活。当然还有耶稣，他曾当过木匠，后来在传道时也没有金银财宝，没有政治职务，没有财产。

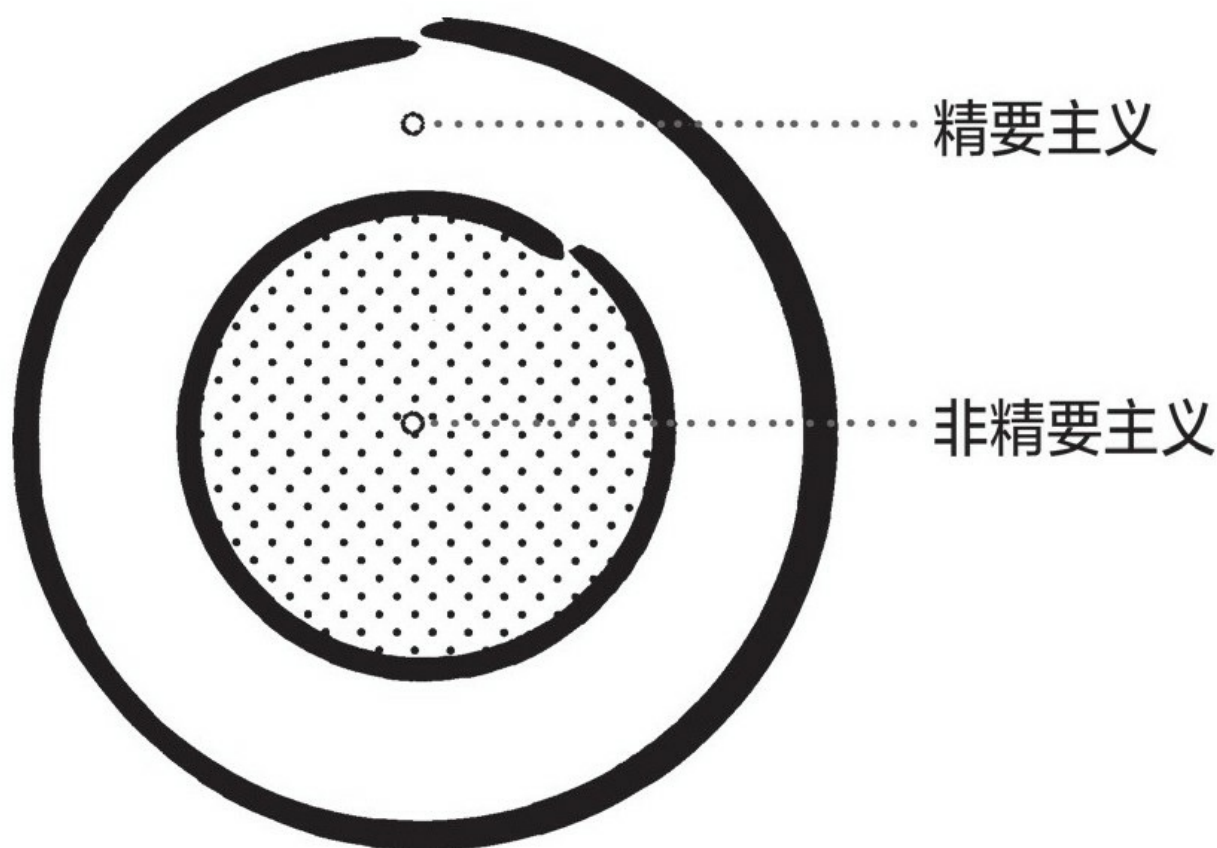
可以看到，历史上还有很多形形色色的著名人物，都在生命中奉行了“更少，但更好”的哲学。举几个例子来说，史蒂夫·乔布斯、列夫·托尔斯泰、迈克尔·乔丹、沃伦·巴菲特、特蕾莎修女、亨利·大卫·梭罗（他曾写道：“我确实相信简单。这是惊人的，也是悲伤的。多少琐碎事务，就连最聪明的人也认为他必须照料一天……所以，要简化生命的问题，区分必要的和真实的。”）。

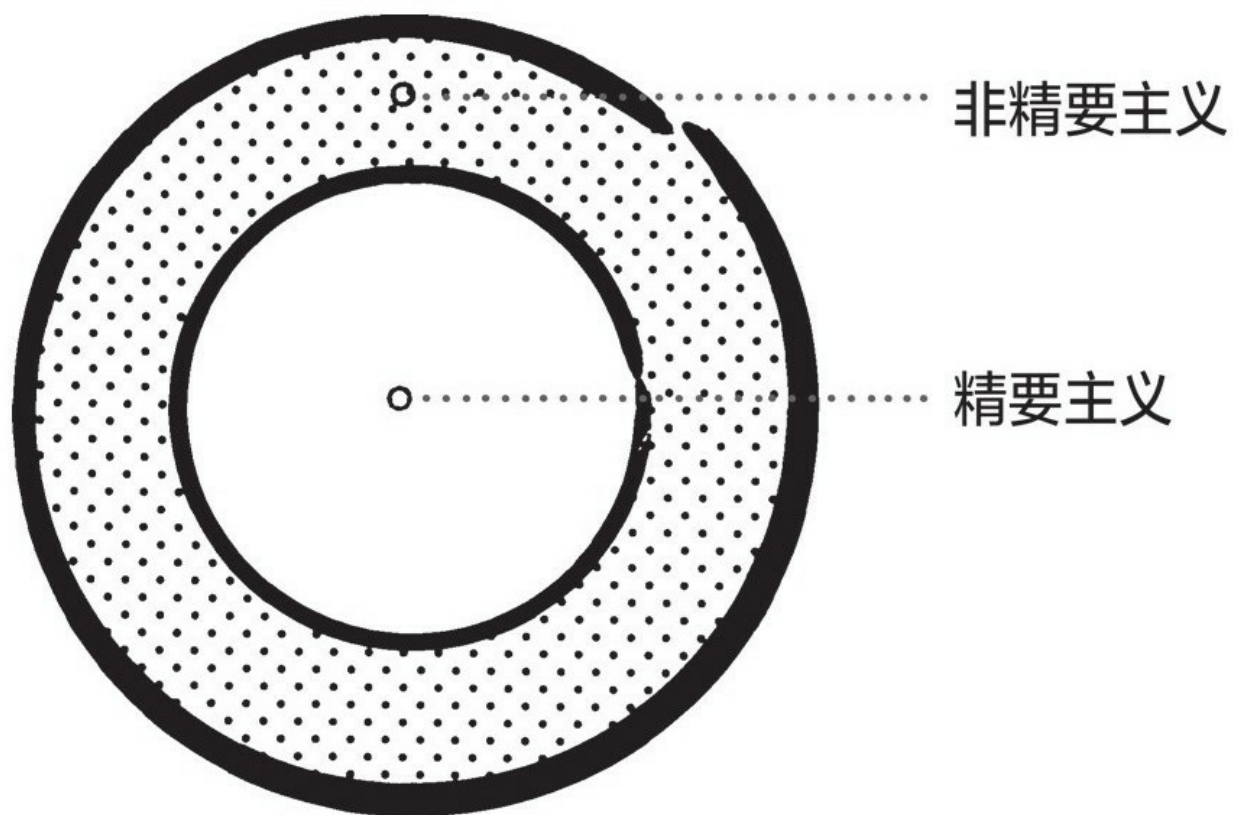
的确，在每一个人类活动领域的翘楚中，我们都能找到精要主义者。他们中，有宗教领袖、新闻记者、政治家、律师、医生、投资者、运动员、作家、艺术家，等等。这些人以各自不同的方式实现了个人贡献峰值。但是，他们共享着同一种品质：他们绝不只是在嘴上说说“更少，但更好”，而是深谋远虑地选择了全面拥抱精要主义。

不要在乎自己所在的职业、领域或者行业，我们也一样可以作出同样的选择。

彻底拥抱精要主义

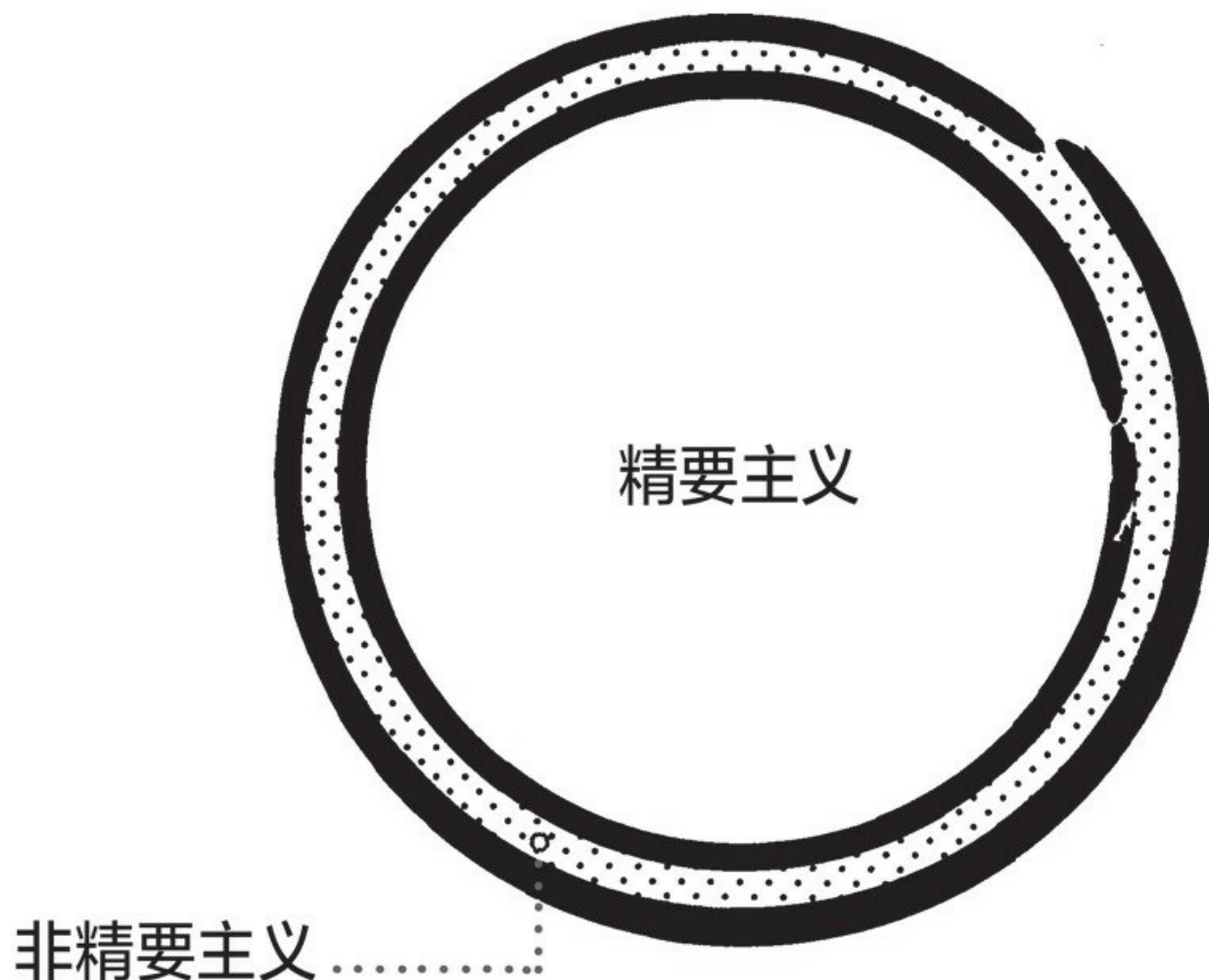
到了本书的这个部分，你已经学到并吸收了所有的精要原理和技能。在本章中，我们必须采取最后一个步骤，学会运用这些技巧，不仅仅为了偶尔践行精要主义，而是为了要成为一个真正的精要主义者。





一个碰巧采用精要行为的非精要主义者，和一个偶尔倒退回非精要行为中去的精要主义者之间，存在巨大的差别。问题在于你以哪个为主，哪个为次？绝大多数人身上，既有一点儿精要特征，又有一些非精要特质，但问题在于：你内在的核心究竟属于哪一类？

核心为精要主义的人，从投资中获得的收益远远大于那些表面精要的人。不错，收益是积少成多的。我们为追求精要方式和消除非精要方式所做的每一项选择都是精要主义自身发展的基础，它会变得越来越习惯成自然，直至成为实质上的习惯。假以时日，那个内在的核心就会向外扩张，直到它把我们仍深陷于非精要泥沼中的那部分几乎完全遮蔽。



人们很容易陷入到我们在本书前言部分讨论过的“成功的悖论”。我们有明确的目标，它能指引我们走向成功。但是，成功给我们带来了新的选项和机遇。这听起来很美妙，但是请记住，这些选项会在无意间分散我们的注意力，魅惑我们，引诱我们偏离正轨。目标的明确性被阴云遮蔽，很快我们会发现自己因为同时在做太多的事情而心力交瘁。现在，我们无法实现个人贡献峰值，只能在无数个方向上取得一星半点儿的进步。最终，我们的成功会沦落为失败的催化剂。而摆脱这个恶性循环的唯一方法，就是精要主义。

但是，精要主义并不仅仅关乎成功，它关照的是一个有意义、有目标的人生。当我们回顾自己的工作和生活时，更愿意看到一条长长的，

上面布满没有什么意义的“成就”的细目清单，还是仅有少许有着真正意义的主要成就的清单？

如果，你允许自己彻底拥抱精要主义，不管是在家里还是在单位，以那样的方式生活，用它来指导自己所做的每一件事情，那么，它便会成为你观察和理解世界的一种方式。深刻地改变自己的思维，直到我们已经讨论过的精要方法和你将开发的其他精要方法，变得自然而然，如同本能一般。

希腊语中有一个词叫“metanoia”，指的是心思的转变。人们常常认为，转变只发生在头脑中。但是正如一句谚语所说：“一个人在心里怎样思量，他就会如其所思。”一旦我们心中充满了精要的精髓，我们就是精要主义者。我们会变成一个不同的、更好的自己。

而一旦你成了精要主义者，就会发现你与其他任何一个人都不相同。当别人满口答应时，你发现自己说的是“No”；当别人在埋头苦干时，你发现自己正在用心思考；当别人侃侃而谈时，你在细心聆听；当别人在聚光灯下为名为利你争我斗时，你却甘坐冷板凳，等待发光的那一天；当别人热衷于美化自己的简历、更新招聘网站上的个人档案时，你却正忙于一砖一瓦地构建有意义的事业大厦；当别人喋喋不休地抱怨（吹嘘）自己有多忙时，你却一笑而过，满怀同情，却一点儿也不认同；当别人活在压力和忙乱中时，你却过着一种有效率、极具满足感的生活。在一个每时每刻都有太多事情的今日社会，像精要主义者那样活着，在很多方面都无异于一种平静的革命行为。

完全过上精要生活，并不总是一帆风顺的。在很多方面，我自己也依然在与之抗争。别人要我做什么事情时，我还是会本能地想要去取悦他们，哪怕我明知此事并不重要。碰到机会时——尤其是那些好机会，我仍会禁不住想“我可以都做”，而实际上根本做不到。我还要努力抵抗那种急着查看手机的强烈冲动，在最糟糕的时候，我甚至怀疑自己的墓

志铭会不会是“他查看了电子邮件”。我愿意第一个承认，转变并不会在一夜之间发生。

但是，时间长了，做起来自然就会越来越容易。拒绝变得不那么难受了。决定更清晰了。排除无关紧要的事情变得越来越自然和本能。我感觉自己对选择更加掌控自如，人生变得不同了。敞开心胸，全然拥抱精要主义，这些变化也会属于你。

现在，精要主义已不仅仅是我做的事情。我正在稳步成为一个精要主义者。刚开始时，它是一些有意为之的选择，然后就成了一种生活方式，接下来它将改造我，成为最核心的我。我几乎天天都会发现自己做的事情越来越少——为了贡献更多。

成为精要主义者于我的意义，完好地体现在一些小小的瞬间里。它意味着：

- 放弃社交活动，与孩子们在蹦床上摔跤嬉闹。
- 为了写作，拒绝了去年的国际客户。
- 每周留出一天不查看任何社交软件，让自己全身心沉浸在家里。
- 为完成这本书的写作，8个月里每天早上5点起床，一直写作到下午1点。
- 推迟工作截止日期，和孩子们一同去野营。
- 出差时，不看电视或电影，以便有时间思考和休息。
- 经常一整天都在做当天最重要的工作，哪怕这意味着做不了自己

任务。清单上的其他事情。

- 把自己正在读的小说搁一边，因为它不是今天的头等要事。
- 过去10年中，一直保持几乎每天写日记的习惯。
- 放弃一次演讲的机会，与妻子安娜共度约会之夜。
- 为和93岁的爷爷通一次定期的电话，舍弃在Facebook上的时间。
- 拒绝斯坦福大学最近提供的讲师工作，因为知道这意味着我将没有时间通过讲座来传播精要主义，以及和家人在一起。

这张清单越拉越长，但是我要在此强调的是，聚焦在重要事务上是一种选择。它是你的选择。它本身是一种令人难以置信的解放。

几年前，在我放弃了法学院的学习之后，我思考着自己下一步的职业规划。安娜是我的决策参谋，我尝试了几十个，可能是几百个不同的想法。有一天，我们开车回家的路上，我说道：“我去斯坦福完成研究生学业如何？”像这样“如果……怎么样”的问题实在太多了。通常情况下，这些想法都是昙花一现。但是这一次，我马上感到了一种明确性：就在那一瞬间，就在那些话刚刚脱口而出时，我知道，这就是我的精要道路。

因为当我想到其他规划时，这种明确性就消失了，这让我非常确信自己已经走上了正轨。有好几次，我开始投入其他规划，但总是没过几分钟就戛然而止，总是不对劲。因此，我把精力都集中在这一个规划上。在等待学校答复的过程中，出现了很多其他的机会，有一些相当有吸引力，但我全部拒绝了。尽管，对于学校是否会接纳我存在不确定

性，但我并不感到焦虑和紧张。相反，我平静，专注，并掌控着一切。

我只申请了斯坦福大学——两次都是。当我最终在第二次收到邀请时，我再清楚不过地知道了这就是我应该做下去的最重要的事情。这是合适的时间里的一条合适的路。这是对精要主义的一种冷静的自我确认。

如果当初没有选择精要主义，我可能从来没有追求过“斯坦福或失败”的策略。我也许从来没有给《哈佛商业评论》写过文章。而且，毫无疑问，我也许永远写不出你们现在正在阅读、吸收，或者更幸运一些，正在认真思考如何转化成自己生活的这些文字。

自律地追求更少

成为精要主义者长路漫漫，但是它带来的收益是无穷无尽的。运用以下这些方法自律地追求更少，将会使你的人生不断向好。

1.更大的明确性

还记得我们在本书前言部分讲到的那个比喻中的衣柜吗？当你继续清理人生的衣柜时，你将体会到自己正在重新整理真正重要的事务。生活将不再以高效地划掉任务清单上的事项或者仓促完成工作日程表上的一切为重，而是以改变最开始时把什么放入清单为重。每天，你将比前一天更清楚，最重要的事务何以能比仅次于它的重要事务重要得如此之多。这样一来，执行这些头等要紧的事务，就会变得越来越轻松。

2.更强的掌控力

你将对自己停顿、推却或者慎重行事的能力拥有更多信心。你将越发感到自己不再为人所用去完成别人的任务清单或工作安排。记住，如

果你不给自己的人生排列优先顺序，别人会来代劳。但是如果你决意要给自己的人生安排主次先后，你是可以做到的。决定权在于你。它是你的内在力量。

3.更多的途中之乐

专注于当下的要事，将会使你充分享受此时此刻的人生。于我而言，生活在当下的一大好处是我能够留下快乐的回忆，它们只在沉浸于当下时才有。我笑得更多，我崇尚简单，我拥有更大的喜悦。

过一种真正重要而有意义的生活

精要主义者的生活是一种有意义的生活，它拥有真正的重要性。

关于这一点，我想到了一个故事。故事主人公的一个三岁大的女儿死了。在悲痛中，他打算把女儿短短小生命中的点点滴滴整合成一个视频。但当他浏览所有的家庭视频时，发现少了点儿东西。他给每次的家庭外出郊游和旅行都拍摄了视频，他有很多镜头——这不是问题所在。他后来发现，虽然有足够的镜头记录下了他们到过的地方——一家人看过的风景、尝过的美食、游览过的地标，但却几乎没有留下他女儿的特写镜头。一心忙于记录外围的景致，却遗漏了最重要的东西。

我在写作本书的过程中，深深领悟到了这个故事蕴含的两大最个人化的真理。第一，家庭在我生命中占据极其重要的地位。相形之下，其他所有的一切都将退化为毫无意义的东西。第二，生命留给我们的时间少得可怜。于我而言，这并不令我沮丧，反倒还让我觉得振奋。它消除了选择错误时的那种恐惧，让勇气深入我的骨髓。它挑战我，让我在使用宝贵的时间时——也许“宝贵”一词太过寡淡，挑剔得更加不近情理。我知道有个人，他在旅行时参观过世界各地的公墓。我开始时觉得有点儿奇怪，但是现在我认识到，这个习惯能让他把自身的必死命运放置于

最重要的位置上。

精要人生，是一种了无遗憾的生活。如果你已经正确地识别出什么是真正的要务，并将自己的时间和精力投入其中，那么你很难会为自己作出的选择抱憾不已。相反，你将会为自己所选择的生活感到骄傲。

你愿意选择去过有目的、有意义的生活，还是愿意带着遗憾的痛楚回顾这只有一次的人生？如果你只能从本书中获得一点启发，我希望是：不管你在人生中面临什么样的决策、挑战，或是站在什么样的十字路口，简单地问问自己：“什么是重要的？”然后排除其余的一切。

如果你摩拳擦掌要向自己的内心探寻这个问题的答案，那么你就已经在精要道路上整装待发了。

**As these
ideas become
emotionally true,
they take on
the power to
change you.**

当你感受到这些想法的真实性时，
它们就拥有了改变你的力量。

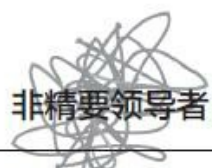
附录

Leadership Essentials

从精要主义者，到精要领导者

毋庸置疑，少数极富思想而又信念坚定的人可以改变这个世界。这也确实是唯一可以改变世界的方法。

——玛格丽特·米德



非精要领导者



精要领导者

思维模式	所有事，所有人	更少，但更好
人才	疯狂招聘员工，制造“傻瓜大爆炸”	对人才挑剔得近乎荒谬，淘汰拖团队后腿的人
战略	采取骑墙战略，把一切都视为优先事务	通过回答“如果只能做一件事，那是哪件？”来明确重要意图，淘汰无意义的分心物
授权	对于谁在干什么模棱两可。决策反复无常	高度关注团队每个成员的角色和贡献目标
沟通	讲话很隐晦	为弄清重要事情而仔细聆听
问责	检查大多内容或者一下子查清全部。扰乱团队目标，然后让该目标不了了之	与员工共同检查，逐步了解他或她是否可以排除障碍，实现一些小小的进展
结果	团队四分五裂，目标众多而散乱，进步微小而分散	团队步调一致，攻坚克难，迈上新台阶

领英的首席执行官杰夫·韦纳认为“更少，但更好”是领导力最强大的运行机制。当他执掌公司大权的时候，他可以很轻易地采用硅谷很多新创公司惯用的经营方法，让公司什么都做。但是，他却拒绝了很多的确很不错的机会，只为了能够抓住最好的机遇。他用“FCS”这个首字母缩写形式（来自FOCUS）来向员工传授自己的思想。这几个字母代表的是：

- Fewer things done better（更少，但更好）
- Communicating the right information to the right people at the right time（在合适的时机向适宜的对象传递恰当的信息）
- Speed and quality of decision making（决策的速度和质量）

事实上，这些正是精要领导的意义所在。

精要团队

作为思维和行事的方式，精要不仅事关公司和团队的领导方法，而且也关系到我们如何管理自己的人生。事实上，我在本书中与大家分享的很多理念，最初就是在我同一些领导团队进行合作的过程中逐渐清晰明朗起来的。

我至今已经收集到了500多个人在1000多个团队中的工作经历。我让他们回答了一系列问题，都是关于当他们在一个团结的团队里工作的时候，他们有什么样的感受，管理人起了什么样的作用，以及最后的结

果如何。然后，我让他们对比一下当他们在一个分裂的团队里工作的时候，情况又是如何，管理人起了什么样的作用，这对最终结果造成了什么样的影响。

调查结果让人大吃一惊：当目标高度明确的时候，团队及其成员能够以压倒性的优势茁壮成长。而当他们对团队的目标及其自身的定位和角色严重不明确的时候，就会出现混乱、紧张、失望的局面，并最终导致失败。一位高级副总裁在看到其团队的结果时，简洁地概括道：“明确性的意义堪比成功。”

这从一个方面解释了为什么“更少，但更好”原则，不仅对于建设有作为的团队大有裨益，而且还有助于人们过上真正有意义的生活。今天的团队讲求速度和机会。当团队团结统一时，丰富的机会可以成为好事。然而，当团队缺乏明确的目标时，要从这形形色色的机会里辨别出哪些是真正重要的就会变得很困难。意想不到的后果就是非精要管理者会让自己的团队追求太多的目标——他们自己也会做太多的事情，结果导致团队发展停滞不前。而一位精要领导者会做出一种截然不同的选择。因为目标明确，她能够把“更少，但更好”的原则应用到人才选拔、方向指导、角色定位、沟通交流及责任追究等所有环节中去。结果，她的团队就能步调一致，攻坚克难，迈上新台阶。

精要领导者的要素

到了本书的这一部分，你已经了解到了非精要思维的种种弊端，并且用精要主义的基本原则替代了错误的逻辑。但是，精要主义并不只是对个体有意义。如果你担任着某种形式的领导职务，不管是在一个只有2个同事的团队，还是一个有着500员工的部门，又或是学校的某个社团，那么你在精要道路上要采取的下一个步骤，就是将同样的技巧和思维方法应用到你的领导工作中去。

将精要方法应用到领导力的每一个重要方面，好处是显而易见的。尽管如此，还是让我们再花点时间，简要地阐述一下这些内容，以便我们能够更加清楚地认识到，究竟应该如何做一个精要主义者领导者。

1.对人才招聘挑剔得近乎荒谬

非精要主义者聘用员工的时候往往很疯狂、很冲动，然后就是为裁员或培训那些拖后腿的员工而忙得不可开交，无暇顾及正事儿。刚开始时，为了保持发展速度，这种聘用狂欢似乎也是情有可原。但在现实中，一次错误招聘的代价要远远大于一个人手短缺。招聘太多不合适的人（一次错误招聘往往会导致更多的错误招聘，因为一个不合适的人会吸引来更多不适合的人）的代价被盖伊·川崎（Guy Kawasaki）称作“傻瓜大爆炸”（Bozo explosion），他用这个术语来描述一个原本优秀的团队或公司陷入平庸时的境遇。

相反，一个精要领导者对人才极为挑剔。他会很自律地坚守完美的招聘标准不放松，不管要看多少份简历，参加多少次面试，开展多少次人才调查，他都会毫不犹豫地淘汰那些拖团队后腿的人。结果，这个队伍就成了一个全明星员工团队，他们共同打造了大于个体能力加总的集体合力。（要了解关于此主题的更多内容，请参考本书第8章）。

2.在真正明确核心意图之前，不要停止争论

目标尚未明确，非精要领导者就实施了骑墙战略：他们试图追求很多目标，做很多事情。结果，团队的精力被分散到无数个方向上，而在每一个方向上取得的进步几近于零。他们在不重要的事情上耗费时日，而对真正重要的事情却疏于照管（请参考本书第9章中关于目标及核心意图的重要性的内容）。现在，公司机构中关于“一致性”的讨论非常多，的确，一个团队目标越是一致，其贡献就越大。明确的目标有助于

实现一致性，而导向模糊则无一例外会导致偏差。

3.支持极端授权

非精要主义者不给员工授权，对于谁在干什么总是模糊不清。就保持队伍的灵活性而言，这样做常常是合乎情理的。但是，这样做导致的实际后果却是虚假的灵活性。当人们不知道自己究竟该对什么负责以及自己的业绩表现会被如何评价的时候，当决策是或者似乎是反复无常的时候，当角色定位不清晰的时候，他们很快就会放弃。或者更糟糕些，他们变得热衷于让自己显得很忙、很重要，而真正的工作却一点儿都不做。

精要主义者懂得，授权的关键在于明确性。他不会让角色定位过于原则性和模糊。他会保证让团队中的每个人都真正清楚自己该做出什么样的贡献，以及其他员工都在贡献什么。最近，一位首席执行官坦陈，他的执行团队存在不明确性，这阻碍了整个公司的发展。为了弥补损失，他大大精简了工作程序，直到只剩下4份直接报告，而每一份都对公司的发展负有明确的功能性责任。

网络支付公司Paypal的创始人彼得·泰尔（Peter Thiel）是一位标新立异的企业家及风险投资家，他将“更少，但更好”的原则发挥到了另类高度。他坚持让贝宝员工只选择一种优先级任务，并且只专注于这个任务。贝宝公司总经理基思·拉布伊斯（Keith Rabois）回忆说：“彼得要求每个人仅被分派一项优先处理事项。除了当前分派给你的首要任务之外，他拒绝同你讨论任何其他事项。甚至我们的2001年年度评审表也是要求每位员工只能确定一项个人对公司最有价值的贡献。”这样一来，员工们都获得了授权在明晰界定的任务范围内做任何事情，他们觉得这样可以为公司的共同使命作出更多贡献。

4.在合适的时机向适宜的对象传达恰当的信息

非精要领导者说话很隐晦，因此员工不确定他们所说的究竟意味着什么，他们的指令常常不是笼统得无法执行，就是转变太快，以至于令人猝不及防。而精要领导者则在合适的时机向适宜的对象传达恰当的信息。他们说话简洁明了，在沟通中选择克制，从而让团队保持专注；他们说话清楚易懂，避免使用毫无意义的术语行话，传达的信息始终如一，听起来似乎还有点儿乏味。但这样做，团队成员就能够透过这些琐碎的噪声获得重要的信息。

5.经常检查以保障有意义的进步

非精要领导者不重视责任追究。这其中主要的，也是显而易见的原因就是一个人追求的目标越多，就越难坚持完成所有的目标。实际上，非精要领导者可能会无意间教会员工根本不要期待善始善终。反过来，团队成员很快就会知道失败没有后果，他们会抄近路，或者优先做容易做的事情，而不是重要的事情。因为他们明白，领导宣布的每一个目标都不会被重视太长时间，它们很快就会被其他能引起瞬间兴趣的事情所取代。

精要领导者会花时间确定真正需要做的事情，从而让后续行动变得很容易，很顺利，而且真正做到善始善终。他们会经常与员工一起检查工作的进展情况，奖励他们的每一点进步，帮助他们扫除障碍，增强团队动力，保持团队专注性，让员工取得更加有意义的进步（请参考本书第16章的相关内容）。

按照“更少，但更好”的原则进行管理，有助于团队发挥集体优势，并成就真正卓越之事。

优秀的精要主义者、真正的先见型领袖埃拉·巴哈特（Ela Bhatt）女

士，一生建树累累，比如获得了富有声望的英迪拉·甘地和平奖，创建了数十所旨在改善印度妇女贫困状况的机构（印度自我就业妇女协会），她是希拉里·克林顿心目中的女英雄之一。她这样写道：

在所有的美德中，我最崇尚的是简单。因此，我常认为简单可以解决几乎所有的问题，不管是个人的，还是世界性的。如果人生简单地过，那么人们就不要再频繁地撒谎、争吵、偷盗、嫉妒、发怒、虐待和杀戮。人人自足充裕，世间便无须再有囤积、投机、赌博和憎恨的行为。个性美好，你便美好。这就是简单之美。

实际上，这也是精要领导力的美丽之处。

不在琐碎中无足轻重

《精要主义》是一本永不会过时的书，因为它稳稳地戳中了人们在生活和职业中如何保持重心以对抗在琐碎中无足轻重这样一个永恒的命题。这一命题不仅令前人困惑不解，也足以让今人和后人绞尽脑汁。它同时又是一本非常及时的书。这是一个比以往任何时候都过度互联的时代，智能终端的束缚捆绑和超量信息的入侵干扰已经把人们切分成了越来越细碎的存在，如何保持自我的完整性以抵御在机械、琐碎和无效的图囿中一步步滑向无足轻重的危机？以往面临危机时，人们往往都会自觉或不自觉地向古之智者先贤乞灵，但这一次也许不完全够，因为我们的时代给了这个难题鲜明的时代特征，它们构成了活生生的、无法挣脱的、不在其中不知甘苦的新的境遇，因此为难古人无异于作茧自缚。在这个时候，活跃在领导能力、战略设计、集体智能及人文系统研究一线的财经商业作家格雷戈·麦吉沃恩的《精要主义》一书，或许能给我们提供一种符合时代特质的解困视角。

人只要还有点儿向前的动力和扑腾的能力，就不会甘于成为庸碌之辈，因此很多人为了最大限度地实现价值而东奔西突。但是任何成功和对成功的追逐都不能脱离具体的、现实的社会条件，因此认清现实无疑是找到出路的首要条件。在林林总总的社会现实中，格雷戈·麦吉沃恩在书中精要地揭示了我们今天面临的一些突出现实。其一，技术的发展

和信息的互通等条件为人类的开疆拓土和外援求索提供了更多选择，但选择的质量却不尽如人意。现代管理学之父彼得·德鲁克曾说：“这是历史上第一次，拥有选择的人的数量有了极大增长。而我们的社会却还没有完全准备好应对这一切。”我们没有准备好，在一定程度上是因为这是第一次，我们面临的选择之多大大超过了我们对此进行管理的能力，无法将重要之事从不重要之事中筛选出来，这导致了选择质量的恶化。其二，除了选择数量大大增长之外，干扰我们作出明智决定和行为的外部影响因素的力量和数量也在增长。以手机24小时待命和信息实时回应等为代表的极端任务指令无时无刻不在宣告着外在世界对人的期待和要求正在呈指数级增长。因为无法在万事万物中择其精要而为之，还要对各种外部期待和要求俯首帖耳、唯命是从，因此满世界都是一个个方向不明、轻重不辨、气喘吁吁的累人或累人联合体——企业。盲目地追求更多，让本已经负重过度的生活和事业雪上加霜，进而与价值实现的目标南辕北辙。

新的现实需要新的做法和活法来成就有效率、有质量、有意义的事业和人生，而做法活法千千万，没有人能够施展其于万一，因此明智之举是寻得法中之法、路中之路。本书孜孜传达的精要主义在某种程度上就是对法中之法、路中之路积极探求所得的成果。用作者的话说，精要是一种自律、系统的方法，用于主动自觉地决定个人贡献峰值所在，然后斩除一切障碍和冗余，轻松自如地执行那些最重要的事。它的基本价值主张可以概括为：若要最大限度成就真正重要之事，切忌贪多求全，事事应揽。而其中所谓的精要，根据作者的逻辑，则不外乎精华、精髓、重要和本质之所在。

曾经成功过的人和企业需要精要主义，而正处在成功华冠上的人和企业似乎也不能例外。不少成功到头来都摆脱不了一种悲剧性的命运，因为它不过是一种瞬间状态而已，在某种条件下甚至还可能发展成一种自掘坟墓的矛盾性存在。目标明确、重点突出和在要事上心无旁骛，可

以帮助人们取得成功，成功会带来更多的机会和选择项，而更多的机会和选择项却导致了精力的分散和更多时间的占用，人们被千头万绪之事牵扯得越来越无法集中精力，因而也就无法实现原本可以实现的个人贡献峰值。成功所带来的后果最终瓦解了最初指引我们走向成功的那个真正清晰的目标，让我们偏离了最本质的事情，这些事情最初曾是成功的源泉。一旦偏离精华和本质，对成功的追求最终就成了失败的催化剂。这种悖论无时无刻不在提醒着我们要不断地回到那个有助于实现个人贡献峰值的精要目标上去，运用精要思维，像精要主义者那样思考、行动和生活。

从对精要主义的技术性剖解入手，本书逐层深入地揭示了精要主义在其作用力和影响力范畴内的三重境界：如何做好一件具体的事情，如何成就成功的人生和事业，如何满足一个世界性的需求。同时，作者牢牢地遵循着精要主义讲求本质和精髓的思路，在精要主义究竟是什么的多种可能性解答中，着力聚焦于精要主义的三种最基本的属性：一种做一切事情的不同方法，一种观察和理解世界的思维方式，以及一种本质的、核心的存在方式，也就是说你就是它，它就是最核心的你。旁征博引和破立并进是本书在写作中的两大突出特点。在本书提供的丰富例证中，我们遇见了无数践行精要主义的知名企业和商业精英，比如有勇气对史蒂夫·乔布斯说“不”的平面设计师保罗·兰德，抽出整整一周时间专门思考和阅读的比尔·盖茨，以及Google公司著名的小憩舱。也看清了无数和自己类似的普通人在偏离或者遵循精要主义时，人生和事业迥然不同的境遇。面对在看清后该如何行动的追问，作者在边破边立中规划了一条步骤清晰的精要人生实现路径。这条路，不仅仅通向成功，还关照着一个人有意义、有目标的人生，人类各个活动领域的翘楚，如圣雄甘地、佛祖释迦牟尼、先知穆罕默德等都汇聚于此，他们以“把自己降为零”的先驱者姿态昭示着简化生命的非凡意义。

法国作家维克多·雨果曾说：“世上有一种东西比所有的军队都更为

强大，那就是恰逢其时的一种理想。”在这过度互联的时代中，要找回淹没在海量信息中的知识，要在数不胜数的事务中集中攻克有意义的少数，要在生活和工作的多种过法和做法中心无旁骛地聚焦于最重要的那种，要在纷纷扰扰的逐利世界中旗帜鲜明地活出超越空洞追求的意义，精要主义正是恰逢其时的一种原理。

也许出身际遇天注定，祸福兴衰如天之风云般不可测，修成圣贤、臻于翘楚对于万众而言最终可能只是个终不可及的幻影，但在人力所能左右的疆域里，于抉择取舍之间时刻倾听内心志向之声，于千头万绪中明辨轻重缓急，于万事万物中斩断一切有碍进步之事，藏大局于胸，不为一时一事所缚，进退有度，涵养心智，活在当下，尽享途中之乐，这或许是人人皆可为之的。

翻译此书的日日夜夜，我于笔耕的辛苦中对精要主义产生了深深的敬意。我也是忙于不重要的多数人，幸福的是我比很多人提前获得了这种拨云见日般的启悟，它时常鞭策我去全神贯注地聚焦于真正该做的要事上，而我也真的这么做了。感谢湛庐文化把此书的翻译任务交给我。感谢张云美、邵永康、黄文忠、黄卓然、邵信定、何扣红、罗雪挥等亲朋好友的关心、帮助和支持。感谢在我成长路上以各种形式帮助过我或即将帮助我的每一个人。

湛庐，与思想有关.....

如何阅读商业图书

商业图书与其他类型的图书，由于阅读目的和方式的不同，因此有其特定的阅读原则和阅读方法，先从一本书开始尝试，再熟练应用。

阅读原则1 二八原则

对商业图书来说，80%的精华价值可能仅占20%的页码。要根据自己的阅读能力，进行阅读时间的分配。

阅读原则2 集中优势精力原则

在一个特定的时间段内，集中突破20%的精华内容。也可以在一个时间段内，集中攻克一个主题的阅读。

阅读原则3 递进原则

高效率的阅读并不一定要按照页码顺序展开，可以挑选自己感兴趣的部分阅读，再从兴趣点扩展到其他部分。阅读商业图书切忌贪多，从一个小主题开始，先培养自己的阅读能力，了解文字风格、观点阐述以及案例描述的方法，目的在于对方法的掌握，这才是最重要的。

阅读原则4 好为人师原则

在朋友圈中主导、控制话题，引导话题向自己设计的方向去发展，可以让读书收获更加扎实、实用、有效。

阅读方法与阅读习惯的养成

(1) 回想。阅读商业图书常常不会一口气读完，第二次拿起书时，至少用15分钟回想上次阅读的内容，不要翻看，实在想不起来再翻看。严格训练自己，一定要回想，坚持50次，会逐渐养成习惯。

(2) 做笔记。不要试图让笔记具有很强的逻辑性和系统性，不需要有深刻的见解和思想，只要是文字，就是对大脑的锻炼。在空白处多写多画，随笔、符号、涂色、书签、便签、折页，甚至拆书都可以。

(3) 读后感和PPT。坚持写读后感可以大幅度提高阅读能力，做PPT可以提高逻辑分析能力。从写读后感开始，写上5篇以后，再尝试做PPT。连续做上5个PPT，再重复写三次读后感。如此坚持，阅读能力将会大幅度提高。

(4) 思想的超越。要养成上述阅读习惯，通常需要6个月的严格训练，至少完成4本书的阅读。你会慢慢发现，自己的思想开始跳脱出来，开始有了超越作者的感觉。比拟作者、超越作者、试图凌驾于作者之上思考问题，是阅读能力提高的必然结果。

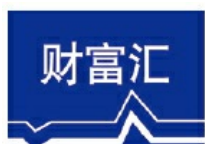
好的方法其实很简单，难就难在执行。需要毅力、执著、长期的坚持，从而养成习惯。用心学习，就会得到心的改变、思想的改变。阅读，与思想有关。

[特别感谢：营销及销售行为专家孙路弘智慧支持！]

✎ 我们出版的所有图书，封底和前勒口都有“湛庐文化”的标志

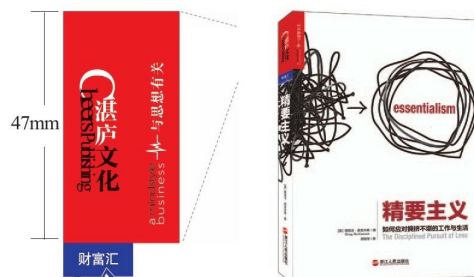


并归于两个品牌



✎ 找“小红帽”

为了便于读者在浩如烟海的书架陈列中清楚地找到湛庐，我们在每本图书的封面左上角，以及书脊上部47mm处，以红色作为标记——称之为“小红帽”。同时，封面左上角标记“湛庐文化 Slogan”，书脊上标记“湛庐文化 Logo”，且下方标注图书所属品牌。



湛庐文化主力打造两个品牌：财富汇，致力于为商界人士提供国内外优秀的经济管理类图书；心视界，旨在通过心理学大师、心灵导师的专业指导为读者提供改善生活和心境的通路。

✎ 阅读的最大成本

读者在选购图书的时候，往往把成本支出的焦点放在书价上，其实

不然。

时间才是读者付出的最大阅读成本。

阅读的时间成本=选择花费的时间+阅读花费的时间+误读浪费的时间

湛庐希望成为一个“与思想有关”的组织，成为中国与世界思想交汇的聚集地。通过我们的工作和努力，潜移默化地改变中国人、商业组织的思维方式，与世界先进的理念接轨，帮助国内的企业和经理人，融入世界，这是我们的使命和价值。

我们知道，这项工作就像跑马拉松，是极其漫长和艰苦的。但是我们有决心和毅力去不断推动，在朝着我们目标前进的道路上，所有人都是同行者和推动者。希望更多的专家、学者、读者一起来加入我们的队伍，在当下改变未来。

湛庐文化获奖书目

《大数据时代》

国家图书馆“第九届文津奖”十本获奖图书之一
CCTV“2013中国好书”25本获奖图书之一
《光明日报》2013年度《光明书榜》入选图书
《第一财经日报》2013年第一财经金融价值榜“推荐财经图书奖”
2013年度和讯华文财经图书大奖
2013亚马逊年度图书排行榜经济管理类图书榜首
《中国企业家》年度好书经管类TOP10
《创业家》“5年来最值得创业者读的10本书”
《商学院》“2013经理人阅读趣味年报·科技和社会发展趋势类最受关注图书”
《中国新闻出版报》2013年度好书20本之一
2013百道网·中国好书榜·财经类TOP100榜首
2013蓝狮子·腾讯文学十大最佳商业图书和最受欢迎的数字阅读出版物
2013京东经管图书年度畅销榜上榜图书，综合排名第一，经济类榜首

《牛奶可乐经济学》

国家图书馆“第四届文津奖”十本获奖图书之一
搜狐、《第一财经日报》2008年十本最佳商业图书

《影响力》（经典版）

《商学院》“2013经理人阅读趣味年报·心理学和行为科学类最受关注图书”
2013亚马逊年度图书分类榜心理励志图书第八名
《财富》鼎力推荐的75本商业必读书之一

《人人时代》（原名《未来是湿的》）

CCTV《子午书简》·《中国图书商报》2009年度最值得一读的30本好书之“年度最佳财经图书”
《第一财经周刊》·蓝狮子读书会·新浪网2009年度十佳商业图书TOP5

《认知盈余》

《商学院》“2013经理人阅读趣味年报·科技和社会发展趋势类最受关注图书”
2011年度和讯华文财经图书大奖

《大而不倒》

《金融时报》·高盛2010年度最佳商业图书入选作品
美国《外交政策》杂志评选的全球思想家正在阅读的20本书之一
蓝狮子·新浪2010年度十大最佳商业图书，《智囊悦读》2010年度十大最具价值经管图书

《第一大亨》

普利策传记奖，美国国家图书奖
2013中国好书榜·财经类TOP100

《真实的幸福》

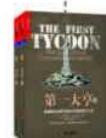
《第一财经周刊》2014年度商业图书TOP10
《职场》2010年度最具阅读价值的10本职场书籍

《星际穿越》

2015年全国优秀科普作品三等奖

《翻转课堂的可汗学院》

《中国教师报》2014年度“影响教师的100本书”TOP10
《第一财经周刊》2014年度商业图书TOP10



湛庐文化获奖书目

《爱哭鬼小牟》

国家图书馆“第九届文津奖”十本获奖图书之一
《新京报》2013年度童书
《中国教育报》2013年度教师推荐的10大童书
新阅读研究所“2013年度最佳童书”

《群体性孤独》

国家图书馆“第十届文津奖”十本获奖图书之一
2014“腾讯网·啖书局”TMT十大最佳图书

《用心教养》

国家新闻出版广电总局2014年度“大众喜爱的50种图书”生活与科普类TOP6

《正能量》

《新智囊》2012年经管类十大图书，京东2012好书榜年度新书

《正义之心》

《第一财经周刊》2014年度商业图书TOP10

《神话的力量》

《心理月刊》2011年度最佳图书奖

《当音乐停止之后》

《中欧商业评论》2014年度经管好书榜·经济金融类

《富足》

《哈佛商业评论》2015年最值得读的八本好书
2014“腾讯网·啖书局”TMT十大最佳图书

《稀缺》

《第一财经周刊》2014年度商业图书TOP10
《中欧商业评论》2014年度经管好书榜·企业管理类

《大爆炸式创新》

《中欧商业评论》2014年度经管好书榜·企业管理类

《技术的本质》

2014“腾讯网·啖书局”TMT十大最佳图书

《社交网络改变世界》

新华网、中国出版传媒2013年度中国影响力图书

《孵化Twitter》

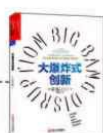
2013年11月亚马逊（美国）月度最佳图书
《第一财经周刊》2014年度商业图书TOP10

《谁是谷歌想要的人才？》

《出版商务周报》2013年度风云图书·励志类上榜书籍

《卡普新生儿安抚法》（最快乐的宝宝1·0~1岁）

2013新浪“养育有道”年度论坛养育类图书推荐奖



延伸阅读

《全新思维》

- ◎ 著名未来学家、趋势专家丹尼尔·平克畅销8年的《全新思维》再度登陆中国，全新译本、经典升级。
- ◎ 未来是那些跨领域型人才、故事型人才、创意型人才的天下。对于当今那些仍然奋斗在会计、律师、工程师岗位的专家型人才来说，是时候给自己的工作来点不同以往的色彩了。
- ◎ 著名管理学大师汤姆·彼得斯称赞该书是一大奇迹，把作者誉为新时代的“哥白尼”。



《绝佳提问》

- ◎ 未来最需要的是不是知道答案的人，而是能够提出美丽问题的人。掌握提问的无穷力量，改变传统的思维方式，真正实现从0到1。
- ◎ 《企业家》网站（Entrepreneur.com）评选的2014年度25本最令人惊叹的图书；《赫芬顿邮报》2014年度改变思维的10本必读书；福布斯网站创意领袖暑期必读书。



《真实的幸福》

- ◎ 积极心理学之父马丁·塞利格曼扛鼎之作，畅销全球20年，销量达到200万册！
- ◎ 提升幸福感不可不读的心理学经典。
- ◎ 清华大学心理学系主任、加州大学伯克利分校心理学系终身教授彭凯平倾力推荐。



《创造力：心流与创新心理学》

- ◎ “心流”之父米哈里·希斯赞特米哈伊经典力作。
- ◎ 深入访谈77位卓越创新者、14位诺贝尔奖得主，为心理学带来革命性影响。
- ◎ 运用“心流”理论，从91位创新者的工作方式、性格特征、思维模式中，寻找让每个人的生活丰富而充盈的创造力。



[1]此处使用化名。

[2]美国心理学家，“积极心理学之父”，认知疗法的主要倡导者之一。想了解马丁·塞利格曼的更多思想，请阅读由湛庐文化策划的幸福著作（套装共5册），包括《持续的幸福》《真实的幸福》《活出最乐观的自己》《认识自己，接纳自己》《教出乐观的孩子》。——编者注

[3]石蕊测试，比喻简单而具有决定性的试验。——编者注

[4]“心流”理论提出者、积极心理学奠基人、创造力大神。想了解他的观点，推荐阅读湛庐文化策划、浙江人民出版社出版的《创造力》。——编者注

[5]全球最具影响力的教育家，Thinkers 50顶尖思想家，创造力和创新领域的全球杰出思想家，排名第一的TED演讲人。想了解有关肯·罗宾逊的更多思想，请阅读由湛庐文化策划、浙江人民出版社出版的《让思维自由》《发现你的天赋》。——编者注

[6]想了解斯图亚特·布朗的更多观点，推荐阅读由湛庐文化策划，中国人民大学出版社出版的《玩出好人生》。——编者注

[7]著名儿童心理学家、曾在哈佛医学院任教22年，现任哈洛韦尔认知及情绪健康中心负责人。他也是注意力缺失领域的顶尖专家和作者，共出版了《分心的孩子这样教》《分心不是我的错》《分心也有好人生》等18部著作，其中多部中文简体字版已由湛庐文化策划出品。——编者注

[8]脾气较火爆，遇事易急躁，喜欢竞争，有闯劲，但缺乏耐性，总是试图做两件以上的的事情，无法处理休闲时光。——编者注

[9]该项目由英国数字化创新研究所与蒂赛德大学共同开展。申请成为“数字化倡导者”，将会接受数字技术方面和教学技能方面的培训，然后通过非正式授课或影片展映等形式，将其所了解的数字技术传授给所在社区的居民，减少数字隔离。——译者注

[10]本书中文简体字版已由湛庐文化策划、浙江人民出版社出版。——编者注

[11]克莱顿·克里斯坦森是当今最具影响力的创新大师，著有5本著作，其中《创新者的处方》和《创新者的课堂》中文简体字版已由湛庐文化策划，中国人民大学出版

社出版。——编者注

[12]当代著名心理学家，斯坦福大学心理学教授，《纽约时报》畅销书作家，其中由湛庐文化策划出品的《津巴多普通心理学》已成为最畅销的心理学教材。——编者注

Thinking

in Systems: A Primer



系统之美

决策者的系统思考

系统思考泰斗，《第五项修炼》作者彼得·圣吉的老师

德内拉·梅多斯

10年手稿首度出版

[美] 德内拉·梅多斯◎著
邱昭良◎译

Donella H. Meadows

 浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE



作者简介 德内拉·梅多斯

- 世界上最伟大的系统思考大师之一，师从系统动力学创始人杰伊·福瑞斯特，是知名的“世界模型Ⅲ”的主创人员，也是“学习型组织之父”，《第五项修炼》作者彼得·圣吉的老师。
- 1972年，作为畅销书《增长的极限》的第一作者，她引领了社会观念变革的先河，引发了全球对地球承载能力和人类选择的大辩论，另著有9本关于全球建模和可持续发展问题的著作，并于1996年创立了可持续发展协会。
- 她始终处于环境与社会分析研究领域的最前沿，1991年被授予环境保护领域“皮尤学者奖”，1994年又荣获麦克阿瑟基金会“天才奖”，曾获得普利策新闻奖提名。

译者简介 邱昭良

- 管理学博士，学习型组织研修中心·中国学习型组织网创始人，中国学习型组织促进联盟主席、首席专家，中国人民大学培训学院特聘教授，国际组织学习协会会员，美国项目管理协会会员、认证项目管理专业人士（PMP）。
- 著有《系统思考实践篇》《学习型组织新实践》《学习型组织新思维》，译著包括《系统思考》《情景规划》《欣赏式探询》等。



更多精彩书讯，请搜索关键词：
@湛庐文化 新浪微博



Thinking
in Systems: A Primer

系统之美

决策者的系统思考

[美] 德内拉·梅多斯 (Donella H. Meadows) 著
邱昭良 译

 浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

版权信息

本书纸版由浙江人民出版社于2012年8月出版

作者授权湛庐文化（Cheers Publishing）作中国大陆（地区）电子版发行（限简体中文）

版权所有·侵权必究

书名：系统之美

著者：（美）德内拉·梅多斯

字数：206000

电子书定价：28.99美元

各方赞誉

左延安 中国工业经济联合会主席团主席，安徽江淮汽车集团公司原董事长

系统思考是整体地、动态地、连续地思考问题的思维模式，是在复杂动态系统中的一种以简驭繁的智慧。决策者唯有深刻领悟系统思考的精髓，才能在正确的时间、正确的地点，做正确的事情。

李兰 国务院发展研究中心公共管理及人力资源研究所副所长，中国企业家调查系统秘书长

系统思考能力是企业家和公共事务管理者等决策者最亟需提升的一项核心能力，本书在此方面具有重要意义和实用价值。

孟庆俊（Stephen Meng）美国生成创造领导力研究发展中心

系统思考虽然是出自于系统动力学的思维工具，其哲学基础却源于东西方探询真善美深处的智慧。创造领导力的系统思考是了解当今世界各种复杂系统的问题与机遇、实现可持续发展的关键策略。邱昭良博士精译之作《系统之美》为重塑我们的创造性系统思维提供了“生生之美”（Creativity in Beauty）的哲理基础和实践工具，既能在求真的过程中体悟系统之美，又可基于系统观，反思真实光景、汲取系统智慧，防微杜渐，并提供尽善尽美的系统设计。借此跨学科的远见卓识，可以践行“生生之德”（Creative Creativity）的智慧，应对各种复杂性挑战，创造更广大、和谐与健康的社会生态经济（Socio-eco-nomic）的可持续发展之道，为人类谋福祉。

宋铠 美国凯斯西储大学系统工程博士，组织学习专家，中国台湾中央大学管理学院资深教授

今天人类面临的各种复杂问题，各个国家、种族、人群、团队之间的争斗，在很大程度上源自于工业时代根深蒂固的心智模式以及由此主导的教育体制——把活的系统当成机器，认为一切事物都是可以被拆解和还原的，可以用理性和逻辑去预测和控制。然而，这显然是错误的。

我们周围充满了各种活的系统，不仅需要逻辑分析，也离不开艺术、直觉、爱与慈悲等智慧。我们必须注重心与物的平衡、左脑和右脑的平衡，冷静洞悉系统的动态性、整体性和规律性。如果能把这种根植于东方智慧的全脑思维技能推广普及，对于社会的和谐、个人的发展，都是价值无限的。

邱昭良 管理学博士，学习型组织研修中心·中国学习型组织网创始人，《系统思考实践篇》作者，《系统思考》译者

“会干活”的前提是“会思考”。如果你想克服自己思维中固有的缺陷，学习和提升系统思考能力，应对复杂性挑战，学会与系统共舞，那么本书对你来说将是一部从入门到精通的智慧宝典。

奥兰·扬（Oran Young）美国加州大学圣巴巴拉分校教授，IHDP科学委员会主席

本书的出版具有划时代的意义.....它将在很长一段时期里重塑人类对社会生态系统的理解，其影响方式就如同20世纪六七十年代《寂静的春天》一书教会人类理解生态系统的本质一样。

莱斯特·布朗（Lester Brown）地球政策协会创始人兼主席

在阅读本书时，我始终感受得到大多数决策者和合格的系统思考者之间的巨大差距。如果本书有助于缩小这一差距，这将是丹娜最伟大的贡献。

亨特·洛文斯（Hunter Lovins）美国自然资本管理咨询公司创始人和主席

任何一位希望创办并成功管理一家公司、一个社区或国家的管理者，都需要阅读《系统之美》这本书。如今，学习如何在系统中思考，已经成为实现变革的基本素质之一。本书是系统思考领域最好的一本书。

比尔·麦吉本（Bill McKibben）Deep Economy作者

德内拉·梅多斯是我所认识的最睿智的人之一，她几乎对于任何问题都能一针见血地指出其潜在的本质，并发现新的问题。本书阐述了她

是如何进行思考的，因此对于我们这些经常困惑于她是如何做到这一点的人来说，无异于无价之宝。这确实是一部经典。

艾默里·洛文斯（Amory Lovins）落基山研究所主席、首席科学家

本书清晰易读，有助于各行各业的读者把握我们这个复杂世界到底是如何运作的，并善加利用。

贾尼丝·莫洛伊（Janice Molloy）Pegasus Communications公司内容总监，《系统思考者通讯》总编

梅多斯具有一种独特的能力，她引导人们超越复杂的概念，洞悉系统的神秘之美。在这本她去世后才出版的书中，她为人们揭示了各种不同的系统如何以相似的规律运作，并为我们提供了改变内在的系统结构以取得更好结果的具体方法。本书是过去十年间系统思考领域10部必读著作之首。

献词 献给世界的复杂性

纪念逝去的丹娜

1993年，德内拉·梅多斯（我们叫她丹娜）完成了本书的初稿。后来，这部手稿并未出版，但是在圈内已经非正式地流传了好多年。2001年，就在正式完成本书之前，丹娜因意外永远地离开了我们……自她离开后，虽然十几年过去了，但她的论述显然并未过时，反而很多得到了印证，并在持续地让广大读者从中受益。作为一位科学家和作家，丹娜是系统建模领域最好的传播者之一。

1972年，作为第一作者，丹娜和其他合著者出版了畅销书《增长的极限》，该书被译成多种文字在全球热销。在书中，她向人类提出了警告，今天看来，这些预见和警告是非常准确的——如果不加抑制，人类持续以以前那种不可持续的方式发展下去，将会对全球造成恶劣影响。在当时，这个话题成为了世界各国各大报纸杂志的头条新闻，人们开始认识到人口和消费的持续增长可能对生态系统和社会系统造成严重的破坏，而后者正是我们人类赖以生存的基础；同时，对经济增长不加节制的追求也可能导致很多本地、区域甚至是全球范围内的系统最终陷入瘫痪。因此，那是一本具有划时代意义的巨著，在该书出版后的数十年间，伴随着石油价格的狂飙、气候变化的严酷现实、全球人口迈向70亿大关，以及物质增长带来的诸多破坏性后果，该书中的发现以及其后的更新，一再成为世人关注的焦点，掀起了全球性的反思和行动。

简而言之，丹娜引领了社会观念变革的先河，促使人们觉醒，改变看待这个世界和系统的方式，以此改变当今社会的发展进程。如今，系统思考作为应对我们周围世界各种复杂性挑战的有力工具，已经得到普遍认可，在环境、政策、社会和经济等领域得到了广泛应用。不论是大系统，还是小系统，都有其内在的规律和原理。只有理解了系统的各种行为特征，我们才能顺势而为，正确地引领系统产生持续的变革，并将其应用于我们生活和工作的各个层面上。丹娜之所以写这本书，就是想将系统思考的观念与方法传递给更多的人。也正是基于这个原因，我和可持续发展协会的同事们决定在她去世后把这本书的手稿正式出版。我们认为，现在是时候了。

一本书真的可以帮助这个世界，帮助包括你在内的各位读者吗？我认为的。也许你在一家公司工作，或者你自己就是某家企业的老板，在当今复杂多变的商业环境中，如果具备系统思考的智慧，你就可以更好地理解你所在的企业或组织是如何与其相关的商业生态系统相互关联、和谐发展的。如果你是政策制定者，你就可以更加透彻地理解自己制定政策的本意、各种可能的阻力、“副作用”以及“上有政策、下有对策”的规则规避行为。如果你是一名管理者，无论是在公司还是社会，每天都要面对很多矛盾、挑战和各种重要的问题，如果不具备系统思考的智慧，不仅会做出很多糟糕的决策，疲于应付，而且有可能使局势越变越糟。在日常生活中，人们经常会谈论社会（或组织）应该如何如何变革，我们也经常会倡导或捍卫某一种价值观。但是，社会的变革是一个缓慢渐进的过程，哪怕是一些最细小的改变都需要历经数年时间。在全球化趋势日益深化的今天，作为一名普通公民，要想积极而持久地保持本土特色文化，我们能够做的，绝不仅仅是一味地抱怨或叹息。如果能够洞悉社会、文化系统的内在结构和运作机理，我们就可以有更大的作为。

如果你想将我上面提到的那些事情加以实践，那么本书将是你的行动宝典。虽然市面上关于系统建模、系统思考的书不在少数，但显然大家都需要一本深入浅出、能启迪人们思维的系统思考应用入门读物，为我们答疑解惑——我们经常遇到的系统性难题是什么、为什么，我们如何才能更好地应对，甚至重新设计系统。

在丹娜撰写本书时，她刚刚完成了《增长的极限》一书20周年修订版，取名为《超越极限》。她是环境保护领域的知名学者，曾为美国国家地理协会研究与勘探委员会工作。此外，她长期在达特茅斯学院教授系统学、环境学和伦理学。在所涉猎的各个领域，她都非常活跃，积极参与到各种活动和项目之中，她把这些活动当做是我们生活中经常会面临的各种复杂系统的外在行为。

尽管丹娜最初的手稿经过了多次编辑和结构上的重新调整，但你在本书中看到的很多例子仍然取自她1993年完成的初稿。可能有人会觉得这有些过时，但我们在编辑本书时，之所以选择保留这些案例，是因为它们给我们的启示至今仍然像当初一样深刻、有效，一点儿也没过时或失效，很多甚至已经被时间所验证。例如，20世纪90年代初期苏联的解体以及其他一些社会主义国家的巨变；北美自由贸易协议的签署；伊拉克入侵科威特，之后在撤军途中烧毁了很多油井；纳尔逊·曼德拉被从

狱中释放，南非的种族隔离法被废除；劳工领袖瓦文萨当选为波兰总统，诗人哈维尔当选为捷克总统；国际气候变化委员会发布了它们的第一份评估报告，指出：“人类活动的排放物，正在使大气中温室气体的浓度显著增加，这将加剧温室效应，并导致地球表面温度进一步升高。”为此，联合国在巴西里约热内卢召开了环境与发展大会。

在那段时间里，丹娜在世界各地旅行，参加各种行业会议和活动，她定期阅读《国际先驱导报》，在一个星期内就找到了很多系统性问题的案例，它们都需要更好的管理或彻底的重新设计。她之所以从报纸上选择案例，是因为这是我们每个人日常生活中真实存在的事件。一旦我们开始把各种相关的日常事件联系起来，并且透过事件察觉到背后的某个趋势，而这些趋势又是潜在的系统结构的外在征兆，我们就已经具备了系统思考的能力。你可以用这种新的方式去管理、决策，并在这个充满了各种复杂系统的世界里更好地生活。通过将丹娜的手稿正式出版，我希望大家帮助大家提高理解和分析你身边的系统的能力，更积极地投身于各种变化。

这是一本循序渐进的系统思考入门指南，也是认识动态复杂系统的有力工具。在当今这个复杂多变的世界里，系统思考的能力显得尤为重要而迫切。这本书讲的是世界的复杂性，也献给世界的复杂性。我希望本书可以帮助那些真心希望更好地塑造未来的人们。

戴安娜·莱特

导读 决策者的系统思考

如果你是一家企业的老板，不管是大企业，还是小企业，你都可能遇到下列难题：

- 想把业务做大做强，但驱动业务成长的关键因素究竟是什么？
- 企业做到了一定规模之后，似乎遇到了成长的“天花板”，徘徊不前，究竟应该如何突破、再创新高？
- 面对竞争加剧、价格大战，究竟应该跟进还是不跟进？
- 面对产品质量问题、销售不畅等困境，究竟应该采取怎样的对策？
- 面对工作与家庭的矛盾，究竟应该如何妥善处理？
- 面对工作应酬与身体健康的冲突，究竟应该如何兼顾？
- 管理团队中，每个人的看法都不相同、争吵不休，如何达成共识、形成合力？
- 管理团队能力不强，究竟如何提高他们的能力？
- 如果你的孩子不愿意“子承父业”，你是否考虑过企业的前途？

如果你是一位职业经理人，不管是企业高层，还是某个部门的负责人，你都可能遇到下列难题：

- 如何组织、协调相关部门完成老板交代的任务？
- 如何更好地与老板以及“创业元老”进行沟通与协作？
- 如何解决你看不惯的一大堆问题？
- 如何带领团队如期、高质量地完成任务？
- 面对项目延期，加班是否能够解决问题？
- 面对其他部门的指责，你应该如何反馈和应对？

如果你是一位家长，面对“调皮”的孩子，如何抓好他的学业？如何与家庭成员和睦相处？

如果你是一位政府公务员，面对一个突发事件，如何快速有效地处置？

.....

世界充满了各种各样的问题。在我们的工作、生活、学习和成长过程中，每时每刻似乎都离不开问题的分析与解决，离不开决策的制定。

按理说，经过无数次的历练，我们每个人似乎都应该成为解决问题与制定决策的高手，但事实并非如此：

- 很多人只是机械地应对，专注于事件层面的响应，而并未触及问题产生的根源；
- 很多问题的成因似乎很复杂，各种因素相互影响，难以理出头绪；
- 很多对策并未抓住问题的本质，只是“治标不治本”的“症状缓解”，或者不全面地解决问题，导致“按下葫芦浮起瓢”，到处救火、应接不暇；
- 很多决策反而产生了一连串的副作用，使局势更加恶化。

事实上，著名系统思考专家、美国麻省理工学院教授约翰·斯特曼（John Sterman）的研究表明，人们用来指导自己决策的心智模式，在应对系统的动态行为方面具有天生的缺陷。人们通常持有一种基于事件层面、因果关系而非回路观点，而忽略了反馈的过程，意识不到行动与反应之间的时间延迟，在交流信息时也未理解存量和流量，并且对于在系统进化过程中可能改变反馈回路强度的非线性特征不敏感。因此，可能产生我称之为“系统思考缺乏症”的五种典型症状：

- 只见树木，不见森林；
- 只看眼前，不看长远；
- 只看现象，不见本质；
- 头痛医头，脚痛医脚；
- 本位主义，局限思考。

任何一个决策者，包括企业家、管理者、公务员，乃至每个人，要想避免上述“系统思考缺乏症”，都需要具备睿智地分析和解决问题、制定决策的能力。系统思考是你应对复杂性挑战、做出睿智决策、与系统共舞的核心技能。

自20世纪60年代以来，系统思考作为与还原论相对立的思维模式，采用了整体论的哲学思想和辩证唯物主义的分析方法，并整合了现代复杂性科学的最新研究成果，不仅具有完备的理论体系，而且拥有实用且

有效的方法与工具，已在欧美等发达地区和国家的企业经营管理、生态环境、交通与水利规划以及社会事务管理、公共政策制定等诸多方面取得了显著成效。

我在《系统思考实践篇》（中国人民大学出版社，2009）一书中提到，系统思考是“见树又见林”的艺术，它具有四个方面的重要价值：

- 解决复杂问题、制定睿智决策；
- 推动组织成长、提升组织能力；
- 激发集体智慧、加速组织学习；
- 建设学习型组织的核心技能。

总之，如果你想克服自己思维中固有的缺陷，学习和提升系统思考能力，本书对你来说将是一本从入门到精通的必读之作。

一部系统思考的经典之作

我之所以认为本书是系统思考从入门到精通的必读书，没有丝毫过誉之意，主要原因在于以下几点。

●高屋建瓴的系统指南

在本书的第一部分，作者简明扼要地讲解了系统动力学的基本概念（如存量、流量、反馈、回路等）和使用方法，并以动物园为隐喻，以每个人都可以理解的生动案例（如浴缸、温度调节器、人口和工业经济、库存、石油经济和渔业系统等），对常见的系统结构及行为进行了阐述。我相信，只要用心，每个读者都可以掌握基本的系统思考方法。

需要说明的是，作者从事系统思考研究、应用与教学工作数十年时间，是系统思考领域大名鼎鼎的“世界模型III”的主创人员，绝对是系统思考领域的大师级人物。虽然她长期从事教学工作，但她并不像有些学究们那样，把系统思考当做充满数学方程式和建模的“高深学问”，而是深入浅出，选取读者身边常见的、耳熟能详的事件作为案例，用系统思考的原理与方法对其进行分析，让读者可以轻松入门，并在潜移默化之中领悟系统思考的精髓、提升系统思考的能力。当然，这一切都基于作者对系统思考的深邃洞察和严谨应用，高屋建瓴，举重若轻，更显难能可贵。

●深入解析系统运作机理

在本书的第二部分，作者对系统为什么会运作得如此精妙、为什么总是让我们感到出乎意料，以及常见的系统结构（基模）进行了解读，可以让读者“知其然，并且知其所以然”。

在第3章中，作者详细阐述了系统的三个基本特性：适应力、自组织和层次性。正是这些特性的存在和相互作用，使得系统衍生出复杂多样的结构，继而呈现出缤纷多彩的行为特性。这是对系统运作内在机理的深刻见解。只有深刻理解这些原理，才能更好地理解和驾驭系统。

由于社会上主流的思维模式与系统思考是相悖的，人们的心智模式也不能很好地考虑到现实世界中的“并发症”或者系统的某个侧面，因此，动态系统的行为经常让我们感到出乎意料。在第4章中，作者给出了一个“警告清单”，包括6项注意事项：

- 为了在复杂的世界里自由遨游，你需要把注意力从短期的事件上移开，看到更长期的行为和系统内在结构，不要被表象所迷惑；

- 在非线性的世界里，不要用线性的思维模式；
- 你需要恰当地划定系统的边界（这一点通常并不容易）；
- 你需要考虑到多重限制性因素以及相对强弱；
- 你需要理解无所不在的时间延迟；
- 你需要清晰地意识到“有限理性”。

这既可以看成是作者对我们思维中常见误区点亮的“红灯”，也可以作为完善自己思维的行动指南。

此外，虽然我本人并不太主张初学者使用“系统基模”（因为初学者在使用“系统基模”时容易“削足适履”、“生搬硬套”），但“系统基模”作为对常见结构或者一些普遍存在的客观规律的总结，便于人们透过纷繁复杂的表象、快速把握问题的本质及其内在结构，对学习者的还是有一定指导和参考价值的。在第5章中，作者基于自己对系统的深刻理解和丰富经验，总结出了8种常见的“系统基模”（作者称之为常见的陷阱和机会）：政策阻力、公地悲剧、目标侵蚀、竞争升级、富者愈富、转嫁负担、规避规则和目标错位。我认为这些总结非常有见地、有价值，不可多得。

●与系统共舞的行动指南和生存法则

这是本书另外一项重要的价值体现，即使对于那些已经可以熟练应用系统思考的原理和方法的决策者来说，这也是一个需要常加温习、反复揣摩，以求心领神会、融会贯通的宝贵指南。

在本书第6章，作者论述了如何寻找系统中的“杠杆点”（即在系统中的某处施加一个小的变化，就能导致系统行为发生显著的转变），提供了12个可能的备选选项——而这正是我本人在应用系统思考方法分析和解决问题的过程中，经常遇到、曾苦苦思索的难题，相信也是众多系统思考学习者在实践中必不可少的指导准则。虽然作者谦称这一清单还不完善，还有很多可改进的空间，但毫无疑问，这是我所见过的最详细、最全面而深入的寻找“杠杆点”的实用指南。

在第7章中，作者则更进一步，对如何更好地与系统共舞、顺势而为提出了15项指导原则，堪称是在系统世界中的生存法则。这些都是作者数十年“功力”的积淀，是应对复杂性挑战、驾驭动态系统的“内功心法”。虽然看起来很简单，但要真正做到，实属不易。即使对于我这样一个研究、应用系统思考十余年的系统思考学习者来说，也弥足珍贵。我愿意把这本书放到案头、手边，经常对照研习、不断自省。

本书的阅读指南

根据本书的结构和内容，我认为，无论对于系统思考初学者，还是已有一定经验的进阶学习者，本书都是一本难得的学习参考资料。

如果你是系统思考初学者，还不了解系统的基本原理和方法，我建议你从第1章开始读下去，循序渐进，终会有成。同时，阅读之余，还可以根据作者的讲解和指导，拿起纸笔（或使用配套的系统动力学建模软件iThink、Stella、Vensim等），边学边练。因为系统思考是一种思维技能，只做到“知道”的程度是远远不够的，动手练习才是知行合一的关键，而反复练习则是技能水平提升不可或缺的重要途径。

如果在阅读和练习的过程中感觉有难度，不妨暂时停下来，把之前读过的内容再温习一遍，或者找对系统思考更有心得的朋友交流一下。等有了一定心得之后，再开始新的学习旅程。当然，更为快捷有效的方式是参加一次“系统思考应用实务”的专业培训，1~3天时间就可以快速入门。

如果你已经有了一定的系统思考训练和学习、应用心得，就可以略过第一部分，从第二部分（第3章或第4章）开始阅读。这样一方面可以联系自己的学习、应用心得以及遇到的实际问题，从书中找到一些解疑释惑的理论支撑；另一方面，也可以加深自己对系统内在运作机理和规律的认识，提升自己的系统思考能力。

如果你是系统思考的高手，本书对你来说也有价值。你可以从本书第5章或第三部分开始阅读。如上所述，我认为，第三部分一定是你需要反复揣摩、常加温习的“系统宝典”。

需要说明的是，由于作者长期从事环境、生态、社会 and 宏观经济与发展领域的研究与教学，她举的很多例子都是与此相关的，因此，本书特别适合从事社会事务管理的政府机关、事业单位公务员，以及环境、生态、社会、行业管理、经济发展等领域的从业人士阅读。当然，基于我自己十余年企业经营与管理的实际经验和系统思考应用心得，我认为它对于企业家、管理者来说，也是一本难能可贵的参考手册，不仅有利于企业家搞清楚自己企业的采购、库存、销售等运作系统，而且可以洞悉企业的成长引擎、找到睿智解决复杂性系统问题的“根本解”和“杠杆解”。

好了，希望大家就此开始提升自己系统思考和决策能力的学习之旅，以便更好地应对复杂性的挑战，与系统共舞！

导读人 邱昭良

- [前言](#)
- [引言 系统多棱镜](#)
 - [无处不在的系统](#)
 - [重塑系统，发现更大的世界](#)
 - [找到属于你的系统之美](#)
- [第一部分 系统的结构和行为](#)
 - [第1章 系统之基础](#)
 - [总体大于部分之和](#)
 - [从关注要素到透视游戏规则](#)
 - [理解系统行为的动态性](#)
 - [反馈：系统是如何运作的](#)
 - [自动洄游的鱼：调节回路](#)
 - [脱缰的野马：增强回路](#)
 - [第2章 系统大观园](#)
 - [单存量系统](#)
 - [双存量系统](#)
- [第二部分 系统思考与我们](#)
 - [第3章 系统之美：系统的3大特征](#)
 - [适应力](#)
 - [自组织](#)
 - [层次性](#)
 - [第4章 系统之奇：系统的6大障碍](#)
 - [别被表象所迷惑](#)
 - [在非线性的世界里，不要用线性的思维模式](#)
 - [恰当地划定边界](#)
 - [看清各种限制因素](#)
 - [无所不在的时间延迟](#)
 - [有限理性](#)
 - [第5章 系统之危与机：系统的8大陷阱与对策](#)
 - [政策阻力：治标不治本](#)
 - [公地悲剧](#)
 - [目标侵蚀](#)
 - [竞争升级](#)
 - [富者愈富：竞争排斥](#)
 - [转嫁负担：上瘾](#)
 - [规避规则](#)

- 目标错位
- 第三部分 改变系统
 - 第6章 系统之杠杆点：系统的12大变革方式
 - 12.数字：包括各种常数和参数
 - 11.缓冲器：比流量力量更大、更稳定的存量
 - 10.存量—流量结构：实体系统及其交叉节点
 - 9.时间延迟：系统对变化做出反应的速度
 - 8.调节回路：试图修正外界影响的反馈力量
 - 7.增强回路：驱动收益增长的反馈力量
 - 6.信息流：谁能获得信息的结构
 - 5.系统规则：激励、惩罚和限制条件
 - 4.自组织：系统结构增加、变化或进化的力量
 - 3.目标：系统的目的或功能
 - 2.社会范式：决定系统之所以为系统的心智模式
 - 1.超越范式
 - 进入谦卑的“空”
 - 第7章 与系统共舞：系统的15大生存法则
 - 跟上系统的节拍
 - 把你的心智模式展现在阳光下
 - 相信、尊重并分享信息
 - 谨慎地使用语言，并用系统的概念去丰富语言
 - 关注重要的，而不只是容易衡量的
 - 为反馈系统制定带有反馈功能的政策
 - 追求整体利益
 - 聆听系统的智慧
 - 界定系统的职责
 - 保持谦逊，做一名学习者
 - 庆祝复杂性
 - 扩展时间的范围
 - 打破各种清规戒律
 - 扩大关切的范围
 - 不要降低“善”的标准
- 附录
 - 系统术语表
 - 系统原理概要
 - 常见的系统陷阱与对策
 - 采取干预措施的杠杆点

- [系统世界的生存法则](#)
 - [模型公式](#)
- [译者后记](#)

前言

即使你的工厂被拆除了，只要它的精神还在，你就能很快重新建立起另一家。如果一场革命摧毁了旧政府，但新政府思想和行为的系统模式没有变化，它就仍然难逃再次被推翻的命运。关于系统，我们很多人经常挂在嘴上，但几乎没有多少人真正理解。

罗伯特·波西格，《禅与摩托车维修艺术》

本书是我个人在系统思考建模与教学方面30多年的经验积累，也凝聚了几十位智者的研究智慧，他们中的绝大多数都曾在麻省理工学院系统动力学小组从事过教学研究工作。这其中最重要的人是我的恩师、该小组的创始人杰伊·福瑞斯特（Jay Forrester）教授。除了福瑞斯特教授之外，我的其他老师还包括艾迪·罗伯特（Ed Roberts）、杰克·皮尤（Jack Pugh）、丹尼斯·梅多斯（Dennis Meadows）、哈玛特·博塞尔（Hartmut Bossel）、巴里·里士满（Barry Richmond）、彼得·圣吉（Peter Senge）、约翰·斯特曼以及彼得·艾伦（Peter Allen），他们中的一些人曾经是我的学生，后来则成为了我的老师。总之，我在本书中呈现的一些观点、案例、引用、书籍以及其他相关资料，都来自于一个更大的智慧群体。在此，我对他们表示由衷的敬仰和感激。

我也曾受教于一些其他学科的思想家们，包括格雷戈里·贝特森（Gregory Bateson）、肯尼思·博尔丁（Kenneth Boulding）、赫尔曼·戴利（Herman Daly）、爱因斯坦、加勒特·哈丁（Garrett Hardin）、瓦茨拉夫·哈维尔（Vaclav Havel）、路易斯·曼福德（Lewis Mumford）、贡纳尔·默达尔（Gunnar Myrdal）、E.F.舒马赫（E.F.Schumacher）等。就我所知，他们从来没有使用计算机进行过系统建模，但他们都是天生的系统思考者。此外，本书还包括一些公司的高层管理者以及很多世界各地古代先贤的智慧。虽然表现形式各异，但殊途同归，真正的系统思考从来都是超越学科和文化的，同时也可以跨越历史。

既然讲到跨越，我还要感谢派系之争。虽然系统分析者共同使用着一些概括性的概念，但他们也是常人，也受人性的左右，这意味着他们也会形成不同的流派。事实上，系统思考领域已经形成了很多不同的学派。在这里，我使用的是系统动力学的语言和符号，因为这是我所接受的教育。同时，我在本书中也只是呈现了系统理论的核心，而不是发展的前沿。此外，我也不会针对最抽象、最深邃的理论展开分析，除非我

认为适当的介绍和分析有助于大家理解和解决实际的问题。等到系统理论发展到了相对成熟的阶段，到时候才有必要写另外一本书详加阐述，我相信会有那么一天。

因此，我想告诫大家，本书和其他所有书籍一样，也存在偏见和不完整性。我在本书中阐述的内容可能只是系统思考领域的九牛一毛，如果你有兴趣去探索，你会发现一个更加广阔的世界，而远不止本书所展现的这个小世界。我的目的之一就是，让你对系统思考感兴趣；而我的另一个目的，也是最重要的目的就是，让你具备基本的理解和应对复杂系统的能力——不管你在阅读本书前后是否接受过正式的系统训练。

德内拉·梅多斯，1993

引言 系统多棱镜

拉塞尔·阿克夫

运筹学理论家

管理者所遇到的问题通常都不是彼此孤立的，而是相互影响、动态变化的，尤其是在由一系列复杂系统构成的动态情境之中。在这种情况下，管理者不能只是解决问题，而应善于管理混乱的局势。

无处不在的系统

在早期教授系统课程时，我经常拿出一个“机灵鬼”玩具——它由一根长长的、松松的弹簧制成，可以弹起来、落下去，或者在两只手之间倒过来、倒过去；如果把它放在楼梯上，它就会自动地“拾级而下”。

我把“机灵鬼”放到手里，手掌向上，用另外一只手从上面抓住它，接着我把底下的手拿开。“机灵鬼”的下端垂了下去，又再弹回来，在我的手指上下，不停地伸展再复原。

“是什么原因让‘机灵鬼’像这样上下弹跳？”我问学生们。

“是你的手。你把手拿开了。”大家回答。

于是，我拿起装“机灵鬼”的那个盒子，也把它同样托起来，用另外一只手从上面抓住它，然后像魔术师一样，吹了一口气，把底下的那只手移开。

什么也没有发生。盒子还是被我抓在手里，一动不动。

“你们再想一想，是什么让‘机灵鬼’这样上下弹跳的？”

答案显然是在“机灵鬼”身上。当人们在按压弹簧或移开手时，弹簧自身的结构会使它做出相应的举动，但人们往往只注意到了手的动作，而忽视了弹簧的内在结构。

这是系统理论的一个核心观点。

一旦我们看清了结构和行为之间的关系，我们便能开始了解系统如何运作，为什么会出现一些问题，以及如何让系统转向符合人们预期的行为模式。当今世界持续快速地变化发展，而且日益复杂，系统思考将有助于我们发现问题的根本原因，看到多种可能性，从而让我们更好地管理、适应复杂性挑战，把握新的机会。

那么，什么是系统呢？系统是一组相互连接的事物，在一定时间内，以特定的行为模式相互影响，例如人、细胞、分子等。系统可能受外力触发、驱动、冲击或限制，而系统对外力影响的反馈方式就是系统的特征。在真实的世界中，这些反馈往往是非常复杂的。

拿“机灵鬼”玩具的例子来讲，原理简单易懂；但如果系统是一个人、一家公司、一座城市或者一个经济体，就不那么简单了。在很大程度上，系统自身就能产生一系列相关的行为。对于一个系统来说，某个外部事件可能引发某些行为，而同一个事件之于另外一个系统的结果就可能迥然不同。

现在，想一想下列观点的含义：

- 经济的繁荣或衰退不是由某个政治领导人左右的——上下波动是市场经济内在的结构；
- 一家公司市场份额的丧失很少是由竞争对手造成的——虽然其对手可能在逐渐积累起竞争优势，但至少部分市场份额的丧失要归因于公司自身的业务政策；
- 石油出口国不是油价上涨的唯一罪魁祸首——如果石油进口国的经济不是建立在脆弱的石油供应上，其石油消费、定价和投资政策不那么容易受到石油供应的影响，石油出口国的行动就不会触发全球油价上涨，导致经济混乱；
- 流感病毒不会攻击你——相反，是你自己的身体状况正好适合流感病毒的生长；
- 吸毒上瘾不是因为吸毒者软弱，其实不管多么坚强或有多少人关爱，没有任何人可以治愈毒瘾，连上瘾者自己也不能——只有理解了上瘾是更大的社会性问题的一部分，会受其他多个因素的相互影响，人们才能开始应对这一问题。

类似这样的论述可能令很多人心神不宁，而在另外一些人看来则是纯粹的常识。我认为，这其实是两种截然不同的做法，一种是顺应系统，另外一种是与其对抗。它们源自人们不同的生活体验，大家对这两者并不陌生。

其实，在接受现代的理性分析教育之前，我们人类早已与各种各样的复杂系统打过交道，并且驾轻就熟。举例来说，人体自身就是一个典型的复杂系统——我们是由各种器官组成的一个整体，各种器官之间相互连接、协调运作、自我调节和成长。不只是我们自身，我们遇到的每个组织、每只动物、每个花园、树木、森林都是一个个复杂的系统。我们凭直觉建立起了关于系统如何运作、如何与它们和谐相处的认知，无须任何分析，也往往不需要语言的表述，这是非常自然的事情。

然而，现代系统理论却经常和计算机、方程式等联系在一起，搞出

一大堆高深莫测的术语、行话。但是，正所谓“大道至简”，系统理论所表达的一些真理，经常在某种程度上是为人所共知的常识。因此，我们完全可以摒弃繁杂的系统术语，直接回归传统智慧。

重塑系统，发现更大的世界

由于复杂系统中存在反馈延迟，经过一段时间之后，问题会变得更加严重，而且更加难以解决。

——一针不补，十针难缝。

——千里之堤，溃于蚁穴。

根据竞争排斥原理，如果一个增强回路使一位参与者取得了胜利，占据了竞争的优势地位，那么它将在未来的竞争中继续获胜，甚至将消灭几乎所有竞争者。

——盖有者，将予之；无者，并其所有亦夺之。（《圣经·马可福音》第4章第25节）

——富者愈富，贫者愈贫。

在受到外力冲击或影响时，通路和冗余众多的多样化系统，往往比几乎没有差异的单一化系统表现得更加稳健。

——不要把所有的鸡蛋都放到同一个篮子里。

自工业革命以来，西方社会的主流思维模式逐渐演变为科学、逻辑以及与直觉和“整体论”相对的“还原论”，并取得了巨大成就。无论是出于心理上还是政治上的需要，我们都更倾向于认为，导致问题的原因是“在那儿（外部）”，而不是“在这儿（内在）”。这种“归罪于外”的思维习惯，几乎不可避免地让我们去责备或怪罪他人，推脱自己的责任，并迷恋于寻找能够摆脱或解决问题的“控制按钮”、“药丸”、产品或技术支持。

诚然，通过专注于外部因素，一些严重的问题得以成功解决，比如预防天花、增加农作物产量、大件货物运输等。但是，由于这些问题都是一些更大的系统中不可分割的一部分，我们的一些“解决方案”已经产生了明显的副作用，并进一步激化了问题。

与此同时，一些问题深深嵌入在复杂系统的内在结构之中，无法被摆脱或转移。例如饥荒、贫困、环境退化、经济波动、失业、慢性病、药物成瘾以及战争等，尽管人们在消除这些问题方面做出了大量努力，不断改进分析方法，并取得了很多技术进步，也没有人故意制造出这些

问题、想让这些问题存在，但是它们依然存在。这是因为它们从本质上看都是系统性问题——系统的内在结构决定了我们所不愿意看到的行为特征。只有重新找回人们的直觉，停止相互指责和抱怨，看清系统的结构，认识到系统自身恰恰是问题的根源，找到重塑系统结构的勇气和智慧，这些问题才能真正得以解决。

可能有人会说，这些道理浅显易懂，甚至是老生常谈，但从某种角度看，它们也是全新的，甚至是有些颠覆性的。这些道理可能让你感到很舒服，因为这些想法其实就存在于我们的认知之中；这些道理也可能让你感到不安，因为我们必须以不同的方式去观察、思考和做事。

本书所讲的就是观察和思考世界的不同方式。有些人一看到“系统”二字，一提到系统分析领域，就不由自主地会有顾虑，即使他们一辈子都在进行系统思考。所以，为了便于各位读者更好地理解系统，我会尽量用“非技术化”的方式来论述，少用一些数学公式或计算机语言。

在本书中，我大量使用了各种图表，因为在讨论系统问题时，如果只用文字会产生问题。谁也无法否认，字、词和句子必须以线性的方式、按照逻辑顺序，一个一个地蹦出来；而系统则是一个整体，同时发生和并存，它们相互连接，不只是单向、线性的关系，而是同时存在多个方向上的多种连接。从某种意义上讲，为了更好地讨论系统，我们有必要使用一种“新语言”，以便完整地呈现系统的特性。

因此，图表胜于文字，因为你一眼便可以看到整个图表，包括它所有的构成部分以及它们是如何相互连接的。在本书中，我会从非常简单的图表开始，逐步构建系统框架。根据我的经验，我相信，读者朋友们会很容易理解这种图表化的语言。

我们首先从基础开始，包括系统的定义以及对其构成要素（部件）的解析。尽管这是一种还原论式的、非整体的方式，但它对于我们理解系统的原理必不可少。接着，我会把这些要素放回原位，好让大家明白它们之间如何相互关联，这也是在勾画系统的基本运作单元：反馈回路。

然后，我会带领大家走进“系统大观园”参观——其中放置了我收集的一些常见的、很有趣的系统。作为示范，我会选择其中几种系统详加解读，让大家看清楚它们是如何运作的，为什么会如此运作，以及我们能在哪里找到它们。事实上，你能认出它们，它们就在我们身边，甚至

就在你的体内！

以此为基础，我会带领你共同讨论这些系统如何以及为什么能够如此优雅地运作，为什么它们经常让我们感到出乎意料、到处碰壁。我也会解释系统的一些“怪异”特征的成因：

- 为什么系统中的每个要素或事物都在忠实、理性地运作，而所有这些善意的行动加起来却经常得到很差的结果？

- 为什么系统整体发展变化的速度要快于或慢于每一个成员的想象？

- 为什么过去一直奏效的一些做法，现在却突然失效了？

- 为什么系统会突然毫无任何征兆地，呈现一种你以前从来没有见过的行为？

这些讨论将引领我们审视一些在系统思考领域——企业、政府、经济和生态系统，生理学和心理学等——被反复提及的常见问题。例如，当我们发现一些社区在共享水资源，或者几所学校在共享财务资源时，就会说：“这是‘公地悲剧’的又一个典型案例。”又如，当我们看到公司现有的业务规则和奖励政策阻碍了新技术的开发时，就会说：“这完全就是‘目标侵蚀’。”再如，当我们研究决策与组织的关系时，经常能发现“政策阻力”的现象，无论是一个家庭、社区，还是一个国家，都是如此。同时，我们也经常能发现“上瘾”的情况，不管引起上瘾的媒介是咖啡因、酒精、尼古丁，还是麻醉剂。

系统思考研究者有时将这些常见的、会引发特定行为的系统结构称为“系统基模”（archetypes）；而我在准备撰写本书时，将其称为“系统的陷阱”，后来我又加上“对策”二字，称其为“系统的陷阱与对策”，因为这些基本模型既是一些顽固、棘手和潜在危害性极大的问题的根源，也是实现有效行为改变的“杠杆点”。只要懂得系统原理，就可以在恰当的地方施加干预措施，从而获得期望的转变。

了解了这一点之后，我将和你一起探讨如何改变我们生活中一些常见难题的系统结构，进而学习如何寻找到启动变革的“杠杆点”。

最后，我将总结一些宝贵经验，这些都是我所知道的大多数系统思考专家分享的智慧。对于希望继续探索系统思考的朋友，我在本书附录中提供了一些深入学习的资源和指南，包括系统术语表、系统原理概要、常见的系统陷阱以及本书第一部分所涉及的系统思考模型的方程式。

找到属于你的系统之美

多年之前，在我们的系统思考研究小组从麻省理工学院转移到达特茅斯学院后，一位工程系教授曾观摩了我们举办的一次研讨会。在观摩了一会儿之后，他拜访了我们的办公室。他说：“你们这些人十分与众不同，你们问不同的问题，看得见我所看不见的事物。总之，你们以不同的方式来看待这个世界。怎么才能做到这一点？为什么这样做呢？”

这也是我希望在本书中讲清楚的问题，尤其是最后的推论。我并不认为系统思考的观察方式比还原主义的观察方式更优秀，我认为二者是互补的，相互具有借鉴意义。就像有时候，你可以通过你的眼睛去观察某些事物，而有时又必须通过显微镜或者望远镜去观察另外一些事物。系统理论就是人类观察世界的一个透镜。通过不同的透镜，我们能看到不同的景象，它们都真真切切地存在于那里，而每一种观察方式都丰富了我们对这个世界的认知，使我们的认识更加全面。尤其是当我们面临混乱不堪、纷繁复杂且快速变化的局面时，观察的方式越多，效果就越好。

通过系统思考的多棱镜，我们可以重新找回对整个系统的直觉，并且训练我们理解系统各个构成部分的能力，看清楚系统各个构成要素之间的关系，分析未来可能的行为趋势，以更具创造性的方式重新设计系统，更有勇气地去面对系统性的挑战。这样，我们就能充分发挥我们的洞察力，去打造一个完全不同的自我和一个崭新的世界。

系统寓言

盲人摸象

在古尔（Ghor）城旁边，有一座城市，城里的居民有不少是盲人。有一次，国王及其随从、车队经过这里，在城市边上安营扎寨。据说，国王骑着一头大象，令国民十分敬畏。

所有人都期盼着能够看一看这头大象，就连城里的一些盲人也随着疯狂的人流涌向国王的营地，去看大象。

他们当然看不见大象长的什么样，只能在黑暗中摸索，把一片片凭着感觉得到的信息拼凑起来。

每一个盲人都认为自己是正确的，因为他们都对大象的某一部分有着真切的感觉……

摸到大象耳朵的那个盲人说：“大象是一个很大、很粗糙的东西，宽阔而平坦，就像一块厚地毯。”

摸到大象鼻子的那个人说：“你说的不对，我说的才是对的。它就像一个直直的、中空管子，很可怕，很有破坏性。”

另外一个盲人摸到了大象的腿和脚，他说：“大象是强壮有力的，很结实，就像一根粗柱子。”

每一个盲人都真切地感知到了大象身体的一个部分，但他们的理解都是片面的.....

这是一则古老的寓言故事，它告诉我们一个简单却经常被忽略的真理：不能只通过了解系统的各个构成部分来认识系统整体的行为。

第一部分 系统的结构和行为

第1章 系统之基础

波尔·安德森

美国科幻作家

无论一个问题多么复杂，如果能以正确的方法去看待，它都会变得简单起来。

总体大于部分之和

系统并不仅仅是一些事物的简单集合，而是一个由一组相互连接的要素构成的、能够实现某个目标的整体。从这一定义可见，任何一个系统都包括三种构成要件：要素、连接、功能或目标。

例如，你的消化系统包括牙齿、酶、胃、肠等要素，它们通过身体血液的流动和一系列化学反应产生相互的连接；消化系统的功能是将食物转化为人体所需的基本营养成分，并将这些营养成分输送到血流中（另一个系统），同时通过新陈代谢，排出各种废物。

再如，一支足球队是一个系统，它的要素包括球员、教练、场地和足球等；它们之间通过游戏规则、教练指导、球员技能、球员之间的交流以及物理法则等产生连接；而球队的目标是赢球、娱乐、锻炼或赚钱等。

同样，一所学校、一座城市、一家工厂、一个公司以及国家经济等，都是系统。动物是一个系统，树也是一个系统，而森林则是一个更大的系统，包含了很多树木和动物等这些子系统。地球是一个系统，太阳系、银河系都是系统，包含了地球和其他很多子系统。

因此，一个系统中可能包含很多子系统，而它也可以嵌入到其他更大的系统之中，成为那个更大的系统中的一个子系统。

那么，有什么事物不是一个系统吗？是的，没有任何内在连接或功能的随机组合体就不是一个系统。举例来说，随机散落在路上的一堆沙子，就其本身来说就不是一个系统，因为它们之间没有什么稳定的内在连接，也没有特定的功能。你可以任意添加或取走一些沙子，而它们仍旧只是路上的一堆沙子。

但是，对于系统来说，如果你更换了其中的要素，系统就被改变了。例如，如果你新引进或开除了某一个球员，或者将你消化系统中的

某些器官进行了一些调整，那么它们很快就不是原来的那个系统了。当一个生物死去，使其成为一个有机系统的多种连接不再产生作用时，它就丧失了作为一个系统的存在状态，尽管它仍是一个更大的食物链系统中的组成部分。类似地，在一座老城中，人们彼此熟识、经常交流，就会形成一个社会系统；而一个充满了各色陌生人等的新街区，就不是一个社会系统，直到这些人之间产生了一些新的连接关系，一个系统才会形成。

系统之思

对于一个系统来说，整体大于部分之和。任何一个系统都包括三种构成要件：要素、连接、功能或目标。它具有适应性、动态性、目的性，并可以自组织、自我保护与演进。

可见，系统既有外在的整体性，也有一套内在的机制来保持其整体性。系统会产生各种变化，对各种事件做出反应，对各种错误或不足进行修补、改善和调整，以实现其目标，并生机勃勃地生存下去，尽管很多系统本身可能是由各种无生命的要素构成的。系统可以自组织，并且常能通过局部的瓦解来进行自我修复；它们具有很强的适应性，很多系统还可以自我进化、演变，生成另外一些全新的系统。

从关注要素到透视游戏规则

苏菲教育故事

因为“一加一等于二”，所以你自认为只要知道“一”，就能知道“二”，但是你忘了，你还必须理解两个“一”之间的关系。

构成系统的要素是比较容易发现的，因为它们多数是可见、有形的事物。例如，树是由树根、树干、树枝、树叶这些要素构成的。如果更仔细地观察，你还会发现其中有一些更小、更具体的单元，如流动着液体的叶脉以及叶绿体等。一所大学也是一个系统，它由建筑物、学生、教师、管理人员、图书馆、图书、计算机等构成——诸如此类，不胜枚举。

当然，要素并不一定是有形的事物，一些无形的事物也可以是系统的要素。比如在一所大学中，学校的声誉和学术能力就是该系统中至关重要的两大要素。事实上，当你想罗列出一个系统中的所有要素时，你会发现那几乎是一项不可能完成的任务。你可以把一些大的要素分解为若干子要素，并进而细分为子子要素，但很快，你就会迷失在系统中，正如人们所说的“见树不见林”。

为避免这种情况，你应该从细究要素转向探寻系统内在的连接关系，即研究那些把要素整合在一起的关系。

系统多棱镜

请思考以下问题：

如何才能知道你观察的是一个系统，而不是一堆材料的集合？

- 1) 你能够识别出各个部分吗？
- 2) 这些部分相互之间有联系吗？
- 3) 这些部分单独作用时产生的影响和它们整合在一起时产生的影响有所不同吗？
- 4) 这些影响和长期的行为在各种环境中都是固定不变的吗？

在树木系统中，内在的连接关系是那些影响着树木新陈代谢过程的物质流动和化学反应，也就是让系统中某一部分对另外一些部分的状况做出反应的各种信号。

例如，晴天时，当树叶散失了水分，负责输送水分的叶脉中的压力就会减小，从而从树根那里汲取更多的水分；相反，如果树根察觉到土壤变得干燥、水分减少，叶脉中压力减小的信号就会让树叶关闭毛孔，以避免流失更多的水分。

在温带，随着白昼的逐渐缩短，落叶树木会释放出一些化学信号，使树叶中的养分向树干和树根传输，从而导致叶茎枯萎、树叶脱落。甚至当某个部分遭到害虫攻击时，树木也能发出一些信号，让自身产生驱虫的化学反应或形成更加坚固的细胞壁。

没有人了解是什么让树木做出这样的反应，但这也不足为奇，因为研究系统的要素毕竟要比研究其内在的关系简单得多。

在一个大学系统中，内在连接包括入学标准、学位要求、考试和分数、预算和现金流、人们的闲谈等。当然，最重要的是知识的交流，这或许才是整个系统的根本目的。

系统之思

系统中的很多连接是通过信息流进行运作的。信息使系统整合在一起，并对系统的运作产生重要影响。

系统中的某些连接是实实在在的物质流，例如树干中的水分，或者学生在大学中的改变；还有很多连接是信息流，也就是系统中影响决策和行动的各种信号。这类连接通常很难被发现，但只要你用心，就会看到它们。例如，学生们可能会通过一些非正式的信息渠道，去了解每门课获得高分的概率，从而决定选修哪些课；消费者可能会参考其收入、

储蓄、贷款额度、家中的存货量、商品价格和数量等信息，来做出是否购买的决策；政府在出台切合实际的治污法规之前，也需参考水污染的种类和数量等信息（当然，了解问题存在的相关信息是必要的，但还不足以采取行动——我们还需要了解资源、动机和结果等信息）。

如果说基于信息的连接都很难被发现的话，那么与系统的功能或目标有关的连接就更难被察觉了。只有通过分析系统的运作，我们才有可能明确地表述出系统的功能或目标。要想推断出系统的目标，最好的方法就是仔细地观察一段时间，看看系统有哪些行为。

如果一只青蛙向右转捉住了一只苍蝇，然后向左转又捉住了另一只苍蝇，接着又向后转捉住了第三只苍蝇，那么我们就可以判断出青蛙的目的并非是向左、向右或向后转身，而是为了捕捉苍蝇。如果一个政府宣称要保护环境，却只为此拨付了很少量的资金，投入很少的精力，那么我们就可以判断出政府的实际目的并非保护环境。因此，必须通过实际行为来推断系统的目标，而不能只看表面的言辞或其标榜的目标。

系统之思

总的来说，“功能”一词常用于非人类系统，而“目标”一词则用于人类系统。但它们之间的区分并不是绝对的，因为很多系统兼具人类和非人类要素。

温控系统的功能是让建筑物内的温度保持在一个设定的水平；植物的功能是结出果实、繁育更多的植物；国家经济的目标是保持增长。几乎每一个系统都有一个重要的目标，那就是确保自我永存。

系统的目标不一定符合人们的初衷，或系统中某个个体的意愿。在很多情况下，系统中各个要素的目标是不一致的，并都会或多或少地对系统整体行为产生影响。最后，系统所呈现出来的结果很有可能事与愿违，谁都不满意，谁都不愿意看到。就像我们现在的社会，吸毒和犯罪日益猖獗，但是没有人主观上想要这样。以下是我们所处的社会系统中一些角色的意愿，而这些意愿汇总起来，就有可能形成上述景象。

- 有些人想要尽快摆脱心灵上的伤痛；
- 农民、商人和银行家都想要赚钱；
- 贩毒人员对法律的约束无所畏惧，但是又害怕警察的打击；
- 政府颁布禁毒法令，并借助警力维护法律，打击毒品贩卖；
- 富人们居住在离穷人很近的地方；

●吸毒者更关心如何保护自己，而非戒除毒瘾。

在社会系统中，上述这些要素各自的目标看起来都是正当的，但它们组合成为一个系统，相互影响，就造成了吸毒和犯罪日益蔓延并很难被根除的恶果。由于系统中嵌套着系统，所以目标中还会有其他目标。

例如，如果说一所大学的目标是创造和保护知识，并将知识代代相传的话，在其中，学生们的目标就可能是取得好的分数，教授们的目标或许是保住饭碗，而管理人员的目标是预算平衡。这些个体的目标有可能与总目标冲突：学生们为了获得好的分数可能在考试中作弊，教授们可能会忽视教学而一心只顾着发表论文，管理者可能会解雇优秀的教授以实现预算平衡。

一个成功的系统，应该能够实现个体目标和系统总目标的一致性。我们会在后面讨论系统的层级时，再来深入探讨这一问题。

如果系统中的个体是一个接一个地发生变化，那么，我们就能够识别出系统中有哪些要素、它们之间的内在联系、系统的目标以及各种要素的相对重要性。但事实并非如此。虽然系统中的某些要素是很重要的，但一般说来，改变要素对系统的影响是最小的。即使更换了一支足球队中的所有队员，它仍是一支球队（当然，有可能表现得更好或者更糟）；一棵树的细胞、树叶年年都在不断地变化，但它仍是同一棵树；你的身体每隔几周就会更换掉大部分细胞，但那仍是你的身体；大学中的学生每年都在不断流动，教授和管理人员也会缓慢变化，但它仍是一所大学；即使更换了所有成员，通用汽车公司和美国国会也依然如故。总之，只要不触动系统的内在连接和总目标，即使替换掉所有的要素，系统也会保持不变，或者只是发生缓慢的变化。

系统之思

系统中最不明显的部分是它的功能或目标，而这常常是系统行为最关键的决定因素。

相反，如果内在连接改变了，系统就会发生巨大的变化。拿球队来说，如果球员之间的关系更亲密了、配合程度提高了，即使还是那些球员，整个球队也可能会变得耳目一新；如果改变足球或篮球比赛的规则，我们肯定会看到一种全新的比赛；如果我们改变了树木中的内在连接（例如，不再让它吸入二氧化碳、呼出氧气，而是相反），那它就不再是树了（或许会变成动物）；如果在一所大学中，不是让教授给学生来打分，而是让学生给教授打分，或者在争论时不是以理服人而是以暴取胜，那么我们就不能称其为大学了。这或许是个有趣的组织，但绝不是大学。总之，改变系统中的内在连接，会让系统发生显著的变化。

同样，功能或目标的改变也会对系统产生重大影响。如果仍旧保留

那些球员和规则，但改变比赛的目标——看谁输而不是谁赢；如果树木的生存目标不是为了繁衍后代，而是为了获取土壤中所有的营养成分，以无限成长；如果大学的目标不仅是传播知识，还要实现赚钱、教导民众、赢取球赛等目的，情况会怎么样？显然，目标的变化会极大地改变一个系统，即使其中的要素和内在连接都保持不变。

有人可能会问：要素、内在连接和目标对系统来说，哪个是最重要的呢？以系统的观点来看，这个问题本身就是个伪命题。

系统提示

系统三要件的关系

对一个系统来说，要素、内在连接和目标，所有这些都是必不可少的，它们之间相互联系，各司其职。一般来说，系统中最不明显的部分，即功能或目标，才是系统行为最关键的决定因素；内在连接也是至关重要的，因为改变了要素之间的连接，通常会改变系统的行为；尽管要素是我们最容易注意到的系统部分，但它对于定义系统的特点通常是最不重要的——除非是某个要素的改变也能导致连接或目标的变化。

例如，更换国家最高领导人。像前苏联领导人由勃列日涅夫换成戈尔巴乔夫，或者是美国总统由卡特换成里根，虽然这个国家的土地、工厂和亿万民众仍然保持不变，但仍有可能把整个国家引向全新的方向，因为新的领导人可以为土地、工厂和民众制定新的规则，或让国家系统的目标发生改变。当然，也有可能什么都不会发生，因为土地、工厂和民众是国家系统中长期存在的物质要素，变化相对缓慢，任何一个领导人对整个国家的改造程度都是有限的。

理解系统行为的动态性

西班牙生态学家 罗蒙·马格列夫

大自然中蕴藏的信息使我们能重现一部分历史。河流的改道、地壳的运动……所有这些世界变迁的真实痕迹，都像是遗传系统中的信息存储器——随着存储信息的不断增加，系统的结构也会日益复杂。

“存量”是所有系统的基础。所谓存量，是指在任何时刻都能观察、感知、计数和测量的系统要素。如其名称所示，在系统中，存量是储存量、数量或物料、信息在一段时间内的积累量。它有可能是浴缸中的水、人口数量、书店中的书、树木的体积、银行里的钱，等等。但是，存量不一定非得是物质的，你的自信、在朋友圈中的良好口碑，或者对世界的美好希冀等，都可以是存量。

系统之思

存量是对系统中变化量的一种历史记录。

存量会随着时间的变化而不断改变，使其发生变化的就是“流量”。所谓流量，是一段时间内改变的状况。例如浴缸中注入或流出的水量、出生或死亡的人数、买入或卖出的数量、成长或衰退、存入或取出、成功或失败等（如图1—1所示）。



图1—1 存量—流量图

系统提示

如何阅读存量—流量图

在本书中，存量用方框来表示，流量则用流入或流出存量的、带箭头的水管来表示。在每个流量上标有一个T型图案，代表“水龙头”，表示流量可以被调高或调低、打开或关闭；在流量的前端或后端，有时会画一个“云朵”图案，表示该流量的资源和消耗，也就是该流量从哪里来、到哪里去。虽然我们可以进一步明确这些来源和去处，但这样会把我们对当前系统的分析变得过于庞杂。因此，出于简化当前讨论的需要，我们可以把这些因素用“云”来概略地表示。

例如，地下的矿藏是一个存量，随着该矿藏被人们发现和开采，会产生一个矿藏开采的流量。由于矿藏的形成（流入量）可能是数百万年前各种复杂地质变化的综合作用，很难全面表述，因此这里用一个简单的存量-流量图进行描述（如图1—2所示，图中没有画出流入量）。



图1—2 矿藏储量慢慢被开采、消耗

大坝后面水库中的水也是一个存量，流入量有雨水和江河的来水，

流出量包括水的蒸发和堤坝放水（如图1—3所示）。

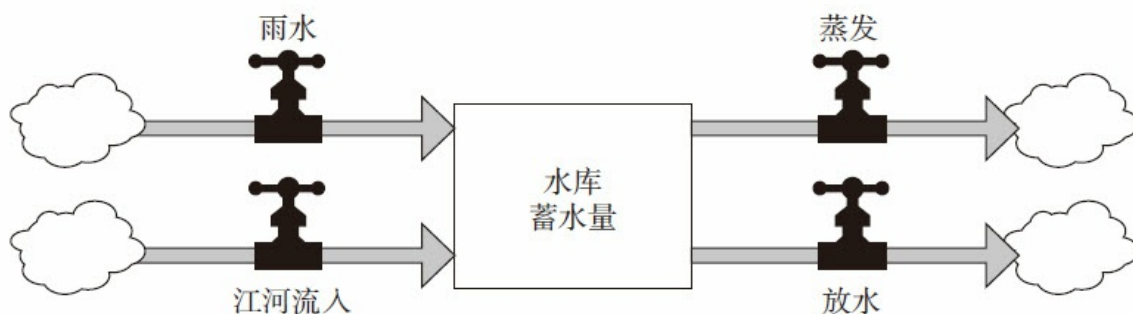


图1—3 水库中的存量和各种流入、流出量

一片森林中所有树木的蓄积量也是一个存量，流入量是树木的生长，流出量包括树木的自然死亡和伐木工的砍伐。被砍伐的木材累积起来，会形成另外一个存量，即伐木工厂里木材的存货量；而当木材出售给客户时，就会产生一个流出量，减少库存（如图1—4所示）。

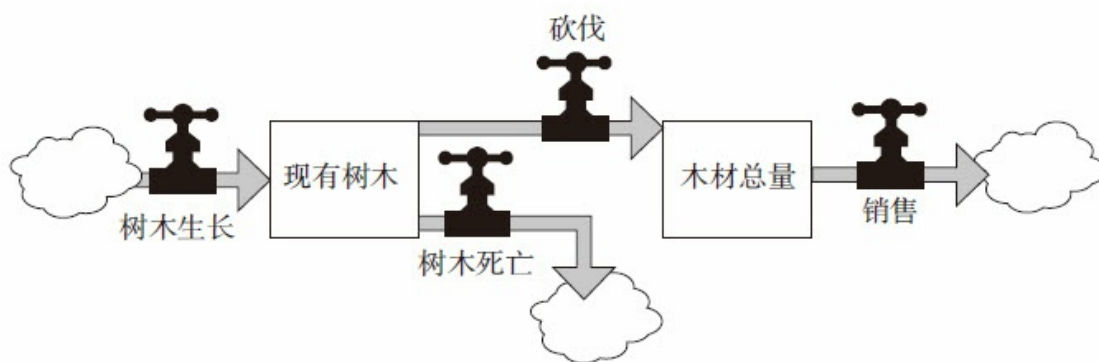


图1—4 一片森林中的存量和各种流量

如果你能够理解各种存量和流量的动态特性，也就是它们随时间流逝而产生的各种行为变化，你就能更好地理解复杂系统的行为。现在，让我们以大家都多少有些体验的“浴缸”为例，来理解存量和流量的动态特性（如图1—5所示）。

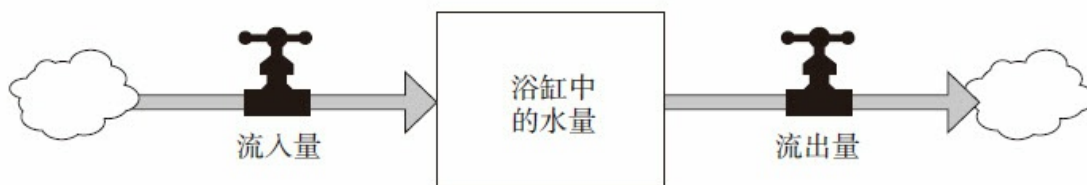


图1—5 浴缸系统的结构——一个存量、一个流入量和一个流出量

想象一下，当一个浴缸盛满了水，其排水口被塞住，自来水管也关着时，这是一个毫无变化、没有活力、乏味的系统。现在，如果我们拔掉塞子，水就会流出去，浴缸中的水位会不断下降，直至浴缸中的水完全流尽（如图1—6所示）。

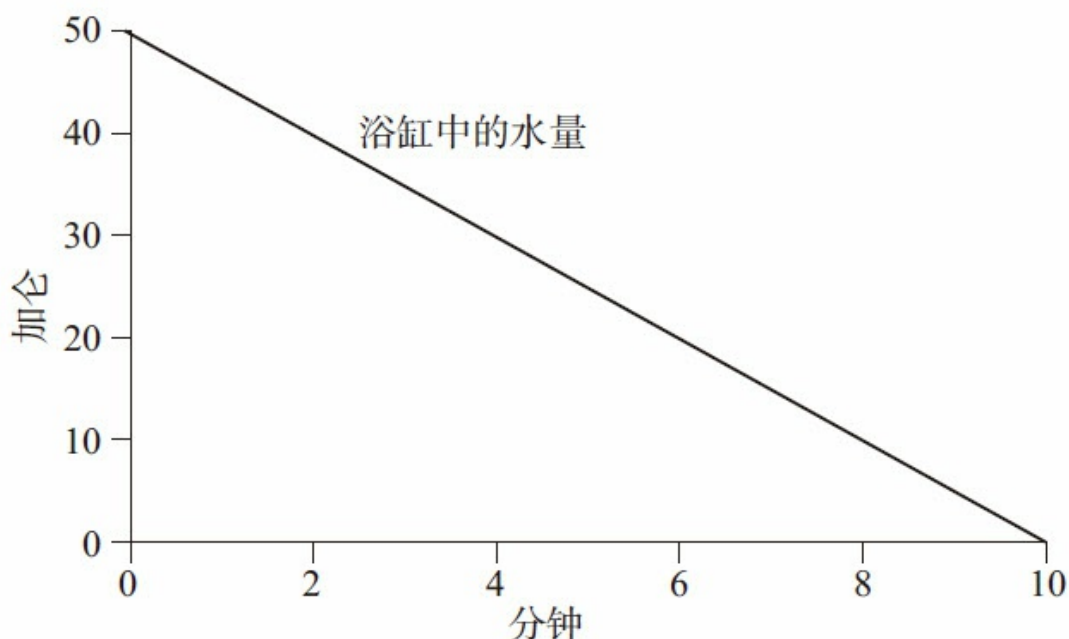


图1—6 当塞子被拔掉时，浴缸中的水位变化情况

系统提示

关于行为模式图的阅读提示

系统思考者经常会使用图表来辅助理解系统的动态变化，了解系统行为随时间而变化的趋势或模式，而不只是关注个别的事件。借助行为模式图，我们可以判断系统是否正在趋向某个目标或极限点变化，也可以了解其变化的速度。

图表中的变量可以是存量，也可以是流量。在阅读这类图表时，要重点关注其变化模式，即表述变量数值变化的线条的形状和方向，相对来说具体的数字并不重要。

横轴是时间轴，有助于你探究问题的来龙去脉，促使你关注所研究问题的时间范围。

现在，让我们再想一想那个盛满水的浴缸。我们再次拔出塞子，但这次，当浴缸还剩下一半水量时，我们拧开水龙头，让水流进浴缸，并使流入浴缸的水与流出的水保持同等速率。这时会发生什么呢？很简单，浴缸里的水位将保持不变，即处于一种动态平衡的状态。（如图1—7所示）

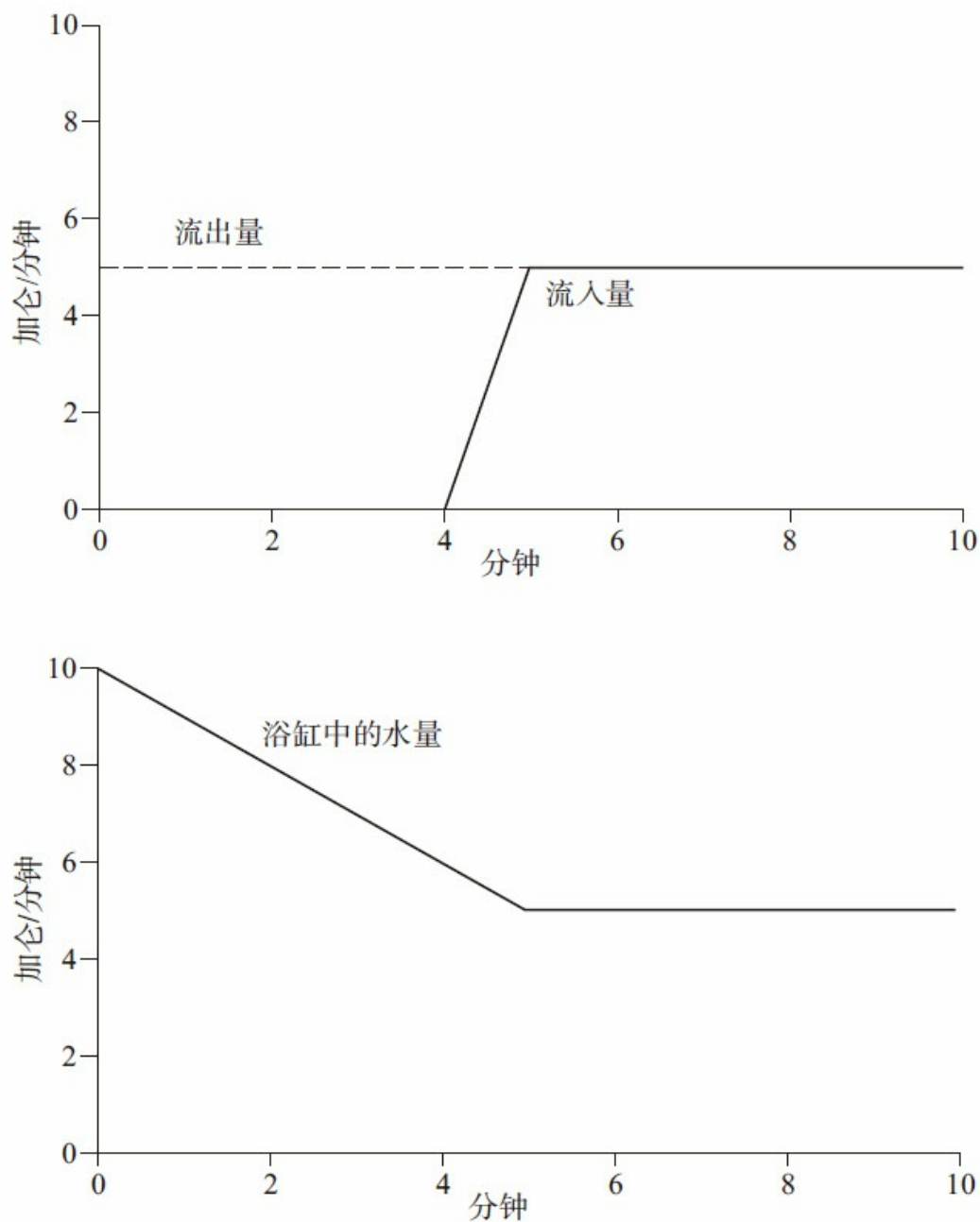


图1—7 水持续流入、流出5分钟的情况，以及浴缸中水量变化的结果

假如我们把水龙头开得再大一些，让水流入的速度稍大于流出的速度，那么浴缸中的水位将会缓慢上升。接着，我们再把水龙头调小，使水的流入和流出速度保持一致，那么浴缸中的水位又会停止上升。相反，拧小水龙头，水位则会缓慢下降。

上述的浴缸就是一个非常简单的系统，只有一个存量、一个流入量和一个流出量。由于在我们考察的时间范围内（数分钟），浴缸中水的

蒸发量微乎其微，所以我们未将这部分流出量计算在内。所有的模型，无论是心智模型还是数学模型，都是对现实世界的简化。借助这些模型，我们可以了解浴缸系统所有可能的动态变化。由此，你可以推断出几项重要的原则，它们同样适用于其他更为复杂的动态系统：

- 只要所有流入量的总和超过了流出量的总和，存量的水平就会上升。
- 只要所有流出量的总和超过了流入量的总和，存量的水平就会下降。
- 如果所有流出量的总和与流入量的总和相等，存量的水平将保持不变；事实上，无论在何种情况下，当系统的流入量和流出量相同时，系统就处于动态平衡的状态。

人类的大脑似乎更加容易关注存量，而不是流量。更进一步地说，当我们关注流量时，我们更容易倾向于关注流入量，而不是流出量。因此，我们有时候会忽视这样一个事实：如果要灌满浴缸，不能只是提高流入速率，还需要降低流出速率。每个人都能认识到，要想维系当今这个离不开石油的经济体系，我们可以加大勘探力度，不断发现新油田；但是，并不是每个人都可以很好地认识到，同样的结果也可以通过减少石油的消耗来实现。如果能在能源使用效率方面实现更大的突破，这和发现一个新油田、增加可用石油储量的效果是一样的。当然，从中获利的人是不同的。

系统之思

要想使存量增加，既可以通过提高流入速率来实现，也可以通过降低流出速率来实现。请注意，要灌满一个浴缸，不是只有一种方式。

类似地，可以通过招聘更多的人来扩大公司规模，也可以减小职工的离职率或辞退速度；这两种策略的成本可能差异很大。可以通过投资、建立更多的工厂和机器设备来增加一个国家的财富，也可以减少工厂和机器设备的磨损、故障或停工；一般来说，后者可能成本更低。

你可以突然调整浴缸的流量——完全打开排水管或关上水龙头阀门，但要想快速地改变存量（水位）就要困难得多。即使你把排水管完全打开，浴缸里的水也不可能一下子排空；同样，即使你把水龙头开到最大，浴缸也不可能马上被灌满。存量的变化需要时间，因为改变它的流量运作需要时间。这是一个关键点，是理解各种系统行为为什么如此运作的一把钥匙。譬如，因为存量的变化一般比较缓慢，它们可能表现为延迟、欠货、缓存、压舱物以及系统中动量的源泉等。存量，尤其是比较大的存量，在应对变化时，只能通过逐步的增加或释放来实现，即使对于突然的变化也是如此。

人们经常低估存量的内在动量。

例如，人口的增长或停止增长、森林中木材的蓄积、水库蓄水、矿藏的耗尽等，都需要花很长的时间。国家的富足，大量工厂和基础设施（如高速公路、发电厂等）的建立，都不是一朝一夕之功，即使你有大把大把的钞票也不可能一夜之间建成。一旦一个依赖石油的经济体被建立起来了，大量的熔炉、汽车发动机每天运作就都需要燃烧石油，即使石油价格暴涨，它们也不可能迅速改变，去消耗另外一种能源。对地球大气臭氧层造成破坏的污染物，是数十年甚至近百年来人类活动迅速增加且疏于污染治理的累积结果；而要想清除这些污染物，也需要花费数代人的心血和智慧。

系统之思

存量的变化一般比较缓慢，即使在流入量或流出量突然改变的情况下，也是如此。因此，存量可以在系统中起到延迟、缓存或减震器的作用。

因此，可以说，存量的改变设定了系统动态变化的速度。工业化的进展速度不能超过工厂和机器设备建设的速度，也不能超过培养出经营这些工厂、操控这些机器设备的合格劳动者的速度。森林不可能一夜之间长成。一旦污染物在地下水中沉积，就只能随着地下水更新的速度去慢慢消除，而这可能需要数十年甚至数百年。

在系统中，由于存量变化缓慢而产生的时间滞后可能会导致一些问题，与此同时，它们也是系统稳定性的根源所在。土壤是数个世纪沉积而形成的，它们不可能瞬间被冲刷、流失殆尽；人类社会数千年积淀下来、世代相传的知识和技能，也不可能转眼间被遗忘；虽然当代人抽取地下水的速度远快于其补充的速度，但由于地下水储量丰富，在相当长的时间里，即使地下水位一直缓慢下降，也还不至于达到无以为继的地步。因此，存量变化缓慢所产生的时间滞后，让人们有了一定的余地去调整、尝试一些做法，并根据反馈来修订那些不奏效的政策。

如果你对存量的变化速度有正确的认知，你就不会“拔苗助长”，期待事物变化的速度超出其特定规律；同时，你也不会过早地放弃，因为你知道一项措施要想见到成效，也需要时间；此外，你也可以更好地把握系统动量所展现的机会，“顺势而为”，就像一个高超的柔道选手善于利用对手的力量那样，聪明地实现自己的目标。

对于存量在系统中所起的作用，还有一个更为重要的原则，那就是：由于存量的存在，流入量和流出量可以相互独立，并在一定时期内不必保持平衡或一致。这一原则可以引导我们直接了解反馈的概念。

系统之思

由于存量的存在，流入量和流出量可以被分离开来，相互独立，并可以暂时地失衡。

如果没有这一原则，很多事情将难以想象。比如，要让炼油厂生产、加工汽油的速度必须与全社会消费石油的速度完全一致，几乎是不可能的事；要让木材砍伐的速度精确地等于树木生长的速度，也是不现实的。石油公司可以生产出汽油，储存在油罐中；森林中木材的储量，以及木材公司的库存等，这些都是存量。正是由于这些存量的存在，即使短期内某些流量的波动很大，人们的生活仍然可以保持一定的确定性、连续性和可预测性。

事实上，人类发明了成千上万种存量维持机制，以确保流入和流出量相互独立和稳定。

例如，为了使当地居民和下游的农民能更稳定地生活和工作，不必担心江河来水的波动性导致旱涝不均，人们在河流上游兴建了水库；为了让你能够支付各项生活开销，不必让自己花钱的速度完全等于你赚钱的速度，人们建立了银行；为了让生产能够顺畅地进行，不必受最终用户需求波动性的影响，人们建立了供应链体系，并在其各个环节（从生产商、分销商、批发商到零售商）都保留一定的库存量，这同时也可以让消费者在想要购买某种产品时能够及时买到，不必完全受工厂短期内生产波动的影响。

由于大量存量维持机制的存在，大多数个人和组织的决策也会受到存量水平的影响：如果库存过高，就会降低价格，或者增加广告或促销方面的预算，以增加销售量，削减库存。

如果家中厨房里食物的储量不多了，你就会去商店采购，以补给食物储备；如果地里种植的作物产量增加或减少了，农民就会想办法浇水、施肥或杀虫，谷物公司就会考虑要预定多少仓库或车皮去贮存和运输这些粮食，投机商就会预测未来的粮价走势，考虑买入或卖出，养殖户就会多养或少养一些牲畜。水库中的水位变化也会引发一连串的修补措施，避免其过高或过低。同样，你钱包中的钱，石油公司的石油库存量，造纸厂的原料储量以及湖中污染物的含量等存量过多或过少，都会引发各种各样的行动反馈。

人们不断地监控存量的变化，根据其状况和特定规则，制定决策并采取相应行动，以增加或降低存量水平，使其保持在可接受的范围内。这些决策累加起来，会对各种相关的系统造成复杂的影响，带来不同程度的起伏、涨落，也造成了各种问题或取得了成功。因此，从系统思考的角度来看，我们这个世界可以被视为各种各样存量的组合，围绕着这些存量，存在着各种不同的存量调节机制，而后者主要表现为各种各样的流量。

这意味着，系统思考者将世界视为各种“反馈过程”的组合。

反馈：系统是如何运作的

杰伊·福瑞斯特

信息反馈控制系统是所有生物和人类行为的基础，从缓慢的生物进化到最先进的卫星发射.....我们所做的任何一件事，无论是个人，还是某个行业或社会，都离不开信息反馈控制系统。

如果存量飞速地增长、急剧地下降，或被维持在某个特定的范围内，不管周围情况如何变化，我们可以很肯定地说，这个系统中存在一种控制机制，并正在发挥作用。换言之，如果你看到某一种行为持续了一段时间，就一定存在导致这种行为产生的作用机制。这种作用机制是通过反馈回路来运作的。因此，长期保持一致的行为模式是反馈回路存在的首要线索。

当某一个存量的变化影响到与其相关的流入量或流出量时，反馈回路就形成了。反馈回路可能非常简单而直接。想象一下你的储蓄存款账户，假设银行和你约定按照复利的方式向你支付利息（也就是俗话常说的“利滚利”），利息的多少取决于你账户中的余额和当前的利率。这样的话，你账户中的余额（存量）会影响到利息的多少，而利息作为一个流入量，也会使下一年的账户余额增加。按照这种算法，银行每年向你支付的利息并不是一个固定的数值，而会随着上一年账户余额的增减而变动。这样就形成了一个简单的反馈回路。

系统之思

一个反馈回路就是一条闭合的因果关系链，从某一个存量出发，并根据存量当时的状况，经过一系列决策、规则、物理法则或者行动，影响到与存量相关的流量，继而又反过来改变了存量。

当你每个月检查自己的活期存款账户对账单时，你也会发现另外一类反馈回路，也很简单、直接。假设现在你的活期存款账户中可用现金（存量）很少，你可能会因为压力而去做更多的工作、赚更多的钱。这些钱会作为流量，进入你的银行账户，这样就可以提高可用现金存量的水平，达到让你满意的程度。如果账户中可用现金很多，你可能愿意放松一下，不用再那么努力地工作，这样赚的钱（流入量）就减少了。这一类反馈回路能使你的可用现金水平（存量）保持在一个自己满意的范围内。当然，赚钱并不是唯一作用于现金存量的反馈回路，你还可以调节自己的消费支出（流出量）。显而易见，这是另外一个调节现金存量的反馈回路。

反馈回路可能导致存量水平维持在某一个范围内，也可能使存量增长或减少。在任何一种情况之下，只要存量本身的规模发生了改变，与

之相关的流入量或流出量也会随之而变。不管是谁或怎样监控存量的水平，一旦存量水平有变化，系统就会启动一个修正的过程，调节流入量或流出量的速度（也有可能同时调整二者），从而改变存量的水平。这又会产生一个反馈信号，再次启动一个控制行动，从而形成一系列连锁反应（如图1—8所示）。

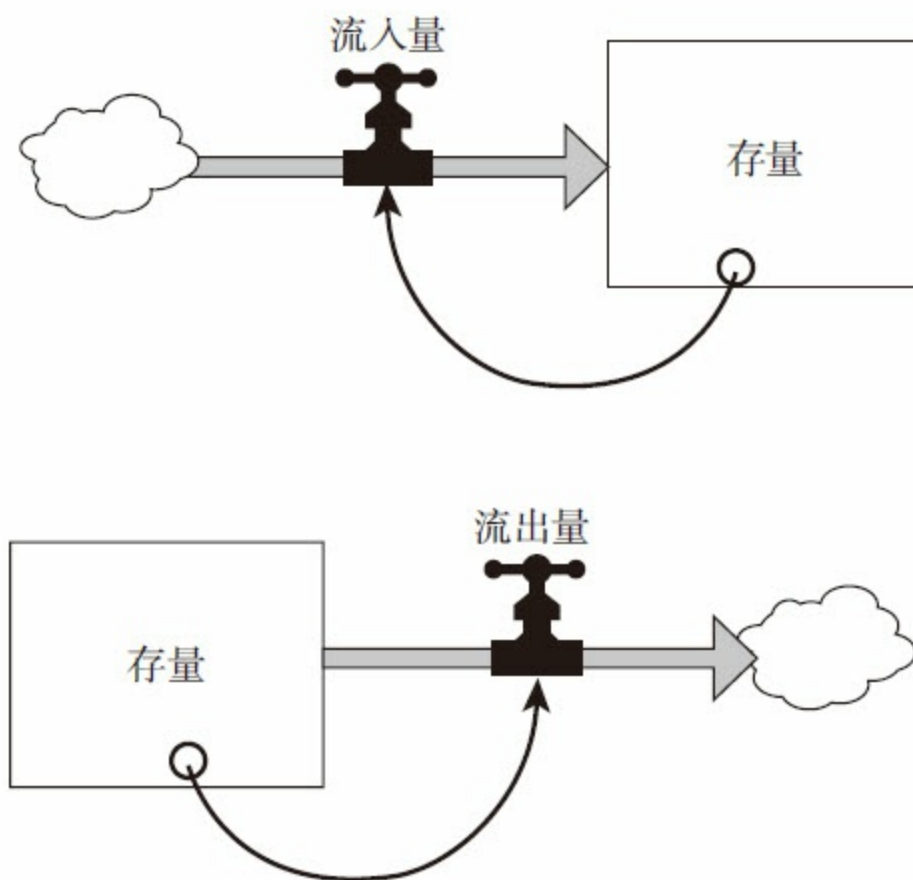


图1—8 带反馈回路的存量-流量图

每一个图都清晰地标注出了存量、影响和改变存量的流量，以及指出行动方向的信息连接（以一个带箭头的曲线来表示）。它强调的是，行动或改变通常是通过调整流量的方式进行的。

不是所有系统都有反馈回路。一些系统是相对简单的、由若干存量和流量构成的、两端开放的链条，它们可能会受到外部因素的影响，但是链条上存量的水平并不影响其流量。然而，更为常见的是包含反馈回路的系统，通常也更为简练且令人惊奇。我们稍后就会见识到。

自动洄游的鱼：调节回路

有一种常见的反馈回路，其作用是使存量的水平保持稳定，就像我们在前面提到的活期存款账户的例子一样。当然，我们所说的“使存量的水平保持稳定”，并不是说存量的水平要完全精确地保持在某一个固定的数值上，而是说保持在一个可接受的范围之内。接下来，我会列举一些为大家所熟知的例子，它们都包含一个能使存量保持相对稳定的反馈回路。这些例子可以让我们更为深入地了解反馈回路。

如果你习惯喝咖啡（或茶），当你感觉有些倦怠时，你可能会煮上一杯浓浓的黑咖啡，让自己重新振作起来。你，作为喝咖啡的人，在头脑中有一个期望的精神状态；当你察觉到实际精神状态与期望状态之间存在差异时，你会通过喝咖啡这一系统，摄入咖啡中的咖啡因，从而调整自身能量的新陈代谢，使自己的实际精神状态（存量）接近或达到期望的水平。当然，你喝咖啡可能还有其他目的，比如喜欢咖啡的味道或者是一项社交活动等，在此不展开讨论。

请大家注意图1—9中名词短语的标签，和本书中其他所有图表中的标签一样，它们都是无方向性的。比如说“身体内储存的能量”而不是“能量水平低”；“咖啡摄入量”而不是“喝更多的咖啡”。我这样做的原因是，反馈回路通常可以往两个方向运转。在上面这个案例中，该反馈回路既可以起到为你补充能量的作用，也可能导致能量供应过量。如果你喝了过多的咖啡，发现自己能量过剩，过于亢奋，你不得不活动一会儿，让过多的咖啡因代谢掉。与期望的能量水平比，过高的能量会产生一种差异感，告诉你的身体“太多了”，这样会让你减少咖啡摄入量，直到身体的能量水平降至合适的范围。因此，使用无方向的标签可以提示我们，影响你身体的能量水平（存量）的反馈回路可以往不同的方向运转。

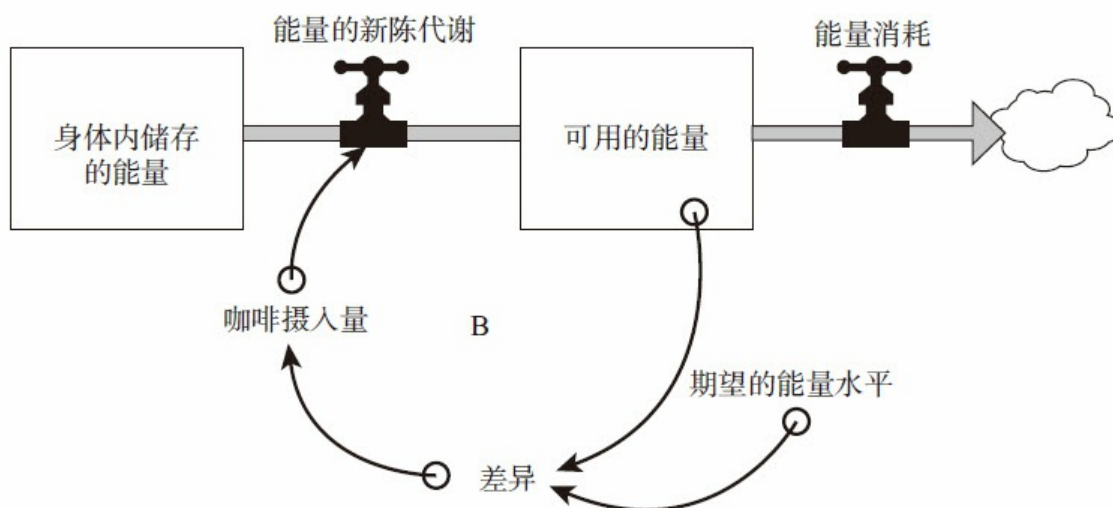


图1—9 喝咖啡的人的能量系统

在我们刚才的讨论中，我把身体能量的流入量来源做了简化处理，忽略不计。现在，让我们更详细地探究一下这个问题，把系统图变得稍微充实、复杂一些。但是，请记住：所有系统图，不管其繁简程度如何，都是对现实世界的简化；我们每个人都在以不同的复杂或简化程度来看待这个世界。

在本例中，我标出了另外一个存量——身体内储存的能量，它可以被咖啡因激活。我之所以这么做，是想告诉大家，在我们现在讨论的系统之外还有更多的东西，不只是一个回路这么简单。每一个喝咖啡的人都知道，咖啡因只是一种短期内起效的刺激物，它能让你的“马达”更高速地运转，但却不能为你的“油箱”补足燃料。过了一定时间，咖啡因的加速作用消失了，而你的身体则因为过快消耗了大量能量而比以往更为疲乏。这一落差可能再次激活一个反馈回路，让你再次跑到咖啡机前去煮咖啡（久而久之，这样就会产生一种“上瘾”结构，我们将在后面进行讨论）；或者可能激活其他一些反馈活动，比如吃一些食物、散散步、睡睡觉等，这比喝咖啡更为长期而健康一些。

这一类反馈回路具有保持存量稳定、趋向一个目标进行调节或校正的作用，我们称之为“调节回路”^[1]。在图中，我在该回路的内部标了一个字母“B”，以示区别。当系统中存在调节回路时，面对各种变化，它都会采取措施，消除这些变化对系统的影响，使存量保持在某一个目标值或可接受的范围之内，系统行为会因此表现出“寻的”或“动态平衡”的特征。无论你是想让存量水平升高或降低，调节回路都会想方设法，将其拉回到预期状态或设定的范围之内。

关于咖啡，我这里还有另外一个调节回路的例子，但它是通过物理法则来起作用的，而不是依靠人的决策。大家都知道，煮完咖啡以后，如果你没有及时把它喝掉，咖啡会逐渐冷却到室温状态，而它冷却的速度取决于咖啡的温度和室温之间的差距：二者的差距越大，咖啡凉得就越快。与上面的案例类似，这一回路的作用方向也有两种——另外一种情况是，如果你在夏天做了一杯冰咖啡，它将逐渐变热，直到达到室内温度为止。该系统的功能是缩小咖啡的温度和室温之间的差距，直至差距为零，不管二者的差距是正还是负（如图1—10所示）。

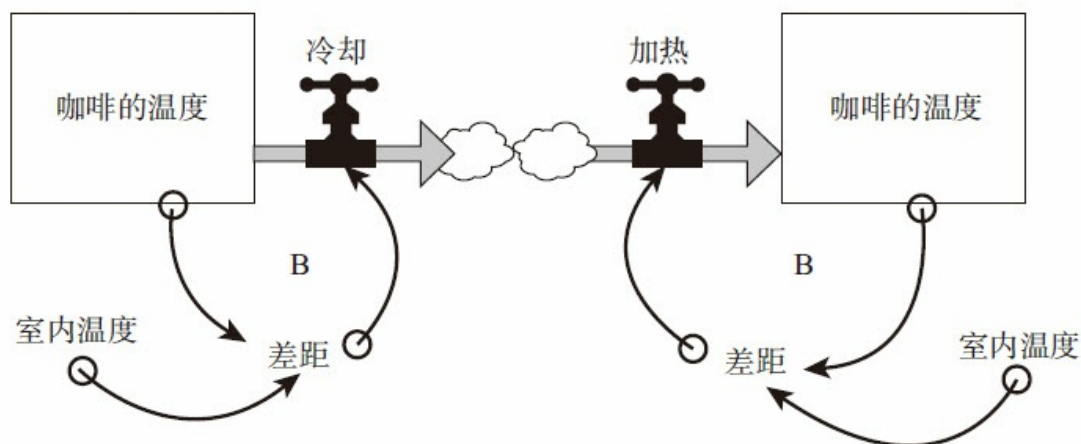


图1—10 一杯咖啡逐渐冷却（左图）或变热（右图）的回路

开始时，咖啡有不同的温度，可能是仅低于沸点（热咖啡）或者仅高于冰点（冰咖啡）一点点，如果你没有把它们喝掉，经过一段时间之后，它们的温度变化情况如图1—11所示。在这里，你可以看到调节回路“自动寻的（返航）”的行为特征。不管系统存量的初始值怎样，也不管它是高于或低于“目标”状态，调节回路都会将其引导至目标状态。一开始变化很快，后来逐渐趋缓，直到存量 and 目标之间的差距消失。

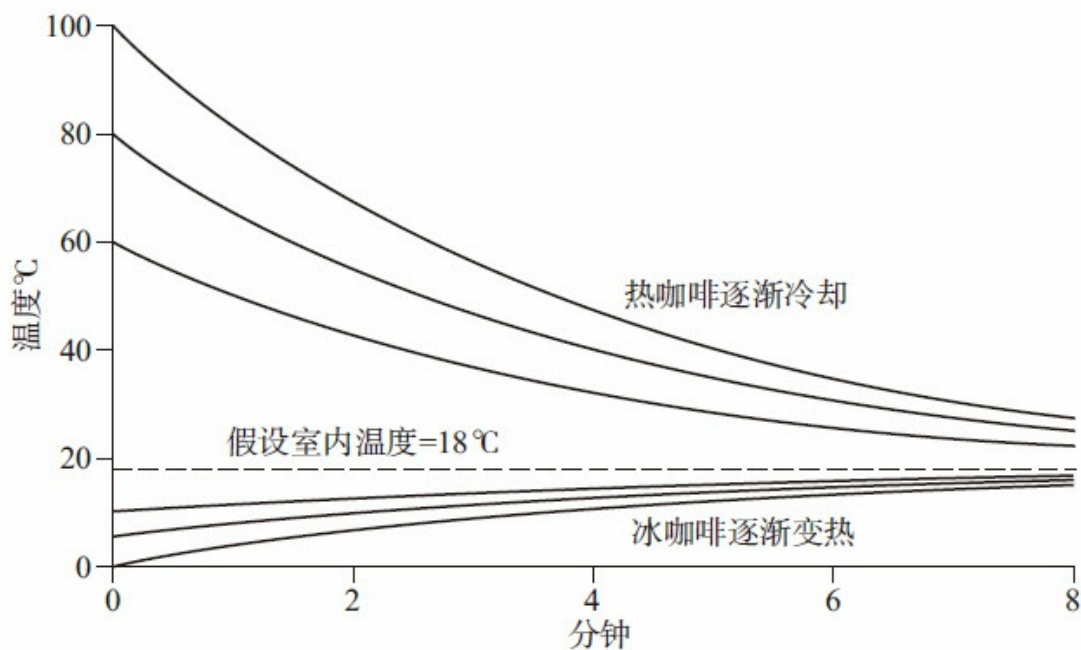


图1—11 咖啡的温度逐渐接近室内温度（假设室温为18℃）

调节回路的这一行为模式——逐渐接近系统设定的目标，在大自然中是很常见的。例如，放射性物质逐渐衰变、导弹的自动制导、固定资产的折旧、水库的蓄水或放水、你的身体对血糖浓度的调节，以及你在停车入位时，都会经历类似的行为模式。你还可以找出更多的例子，这个世界充满了自动复位的调节回路。

系统之思

在系统中，调节回路是保持平衡或达到特定目标的结构，也是稳定性和抵制变革的根源。

反馈机制的存在并不意味着它可以很好地起作用。有时候，相对于外部变化或影响而言，如果这些反馈机制不够强大，它就无法将存量水平引导至期望的状态，系统就会失效或被改变。反馈其实是系统中各种要素之间的相互联系，是构成系统的信息要件。因为种种原因，反馈有可能会失效。

例如，信息有可能到达得太晚，或者没有传送到合适的地方；信息有可能不清晰、不完整，或者难以被解读；反馈触发的行动可能力度太弱、太滞后、受到资源的限制，或者根本无效。

大千世界是纷繁复杂的，在实际中，一些调节回路的目标可能永远也无法达到。当然，我们上面所举的例子很简单，咖啡的温度一定会逐渐与室温相同。

脱缰的野马：增强回路

巴尔扎克

我需要休息，让我的大脑重新焕发活力，而旅行就可以让我得到休息。但是，要能去旅行，我必须得有錢。为了賺到錢，我必须去工作……我陷入了一个“恶性循环”之中，根本不可能逃出它的魔爪。

简·丁伯根

我们经常会遇到这样一种状况，看起来好像是循环迭代的：利润下降是因为投资不足，而投资不足是因为利润不佳。

第二类反馈回路的作用是不断放大、增强原有的发展态势，自我复制，像“滚雪球”一样。它们是一个良性循环或恶性循环，既可能导致系统不断成长，越来越好；也可能像脱缰的野马，导致局势越来越差，造成巨大的破坏甚至毁灭。我们将这一类回路称为“增强回路”^[2]。为了表示区别，我在这类回路内部标注了一个字母“R”。在这类回路的作用

下，系统的存量越大，存量的流入量也就越多，导致存量进一步变得更大；反之亦然。总之，增强回路会强化系统原有的变化态势。

举例来说：

- 两个小孩子发生了争执，一个孩子打了另一个孩子一拳，后者就会踢前者一脚，这样就导致前者更大力度地反击.....就这样，冲突不断升级。

- 物价升高，要想让人们维持原有的生活标准，就需要给工人涨工资；而工资越高，产品的价格就需要更高，以便企业能够维持获利.....而这意味着，物价水平将变得更高，同时又需要给工人涨工资，如此循环不已。

- 野兔数量越多，有生育能力的兔妈妈就越多，生下来的兔宝宝也就越多；而兔宝宝越多，等它们长大以后，兔妈妈就变得更多，又会产下更多的兔宝宝。

- 土壤流失越严重，植被就越稀少；而植被越稀少，巩固和维系土壤的根就越少，从而导致更大的土壤流失，植被更加稀少。

- 我练习弹钢琴的次数越多，技术水平越好，从琴声中体味到的乐趣就越多，从而让我更加愿意弹琴，进行更多练习。

增强回路也很常见，当你发现系统中某个要素具有自我复制或繁殖的能力，或者持续增长时，你就能找到推动其增长的增强回路。还记得我们在上文中提到的那个银行存款账户的例子吗？银行账户的余额越大，你所能获得的利息就越多，这使你的存款金额更大，你下一期获得的利息更多（如图1—12所示）。类似的例子还有很多，包括人口、经济体系等。

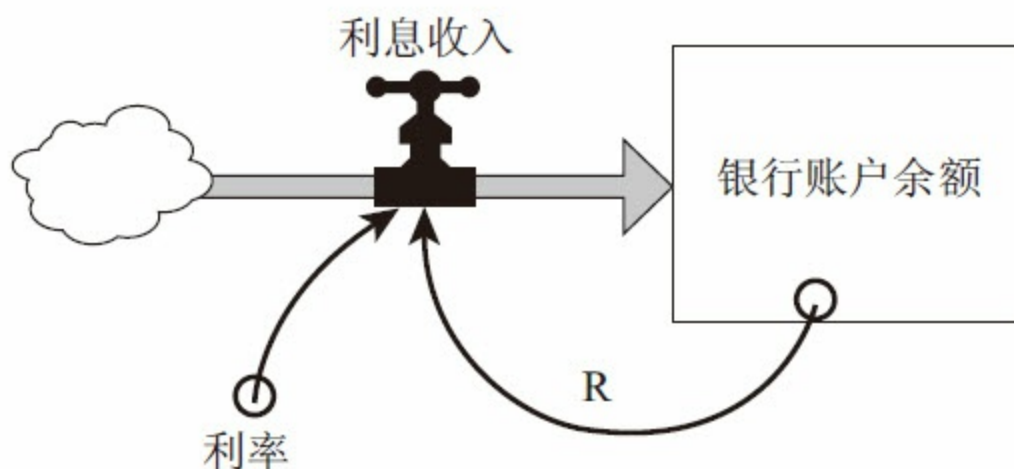


图1—12 银行存款账户增强回路

图1—13显示的是在这一增强回路的作用之下，你在银行中的钱

（初始值是100美元）是如何增长的。我们假设你在12年中没有对这一账户进行过其他存取款操作。图中有五条曲线，分别对应的是五种不同利率的情况，从年利率2%~10%。

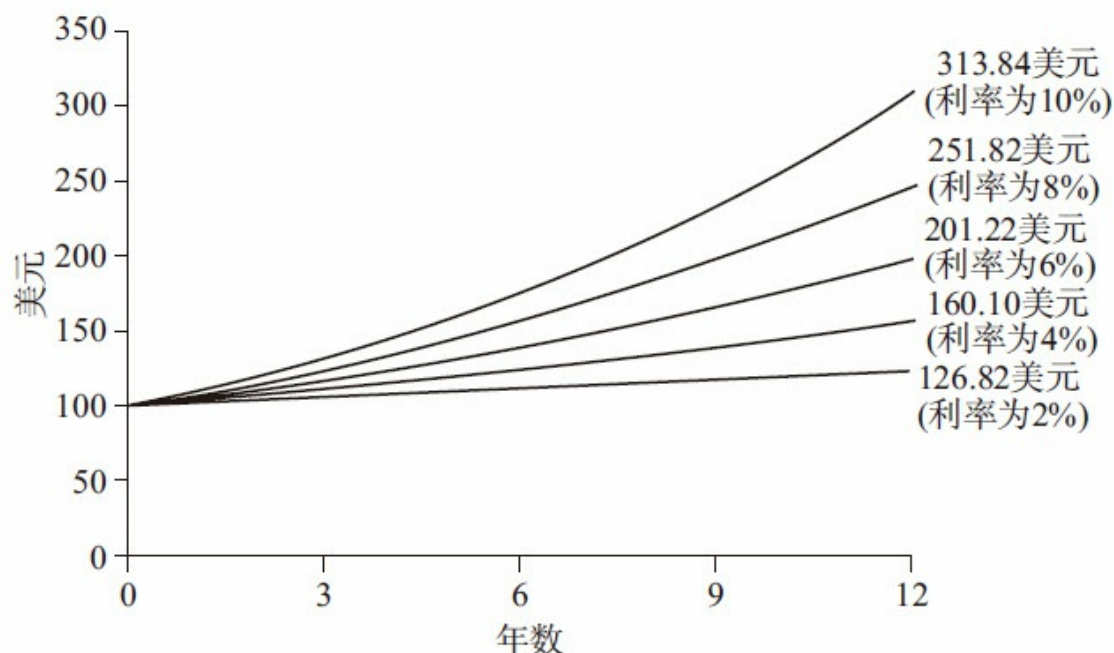


图1—13 在不同利率情况下，银行存款账户余额的变化状况

这不是简单的线性增长，每一年的变化不是固定的。虽然在利率较低的情况下，银行账户余额头几年的增长看起来像是线性的，但实际上，它的增长是越来越快的。余额越大，增加得越多，此类增长被称为“指数曲线”。当然，这是好消息还是坏消息，取决于到底是什么在增长——存在银行中的钱在增长，这就是好消息；你借了高利贷，要支付的利息在增长，这就是坏消息。当然，感染艾滋病的人数在增长、玉米地里的害虫在增长、国家的经济在增长，或者军备竞赛过程中武器装备在增长，不同的人对好坏的评判是有差异的。

系统之思

增强回路是自我强化的。随着时间的变化，增强回路会导致指数级增长或者加速崩溃。当系统中的存量具有自我强化或复制的能力时，你就能找到推动其增长的增强回路。

如图1—14所示，你拥有的工厂和机器设备（一般被称为“资本”）越多，能生产出的产品和服务（“产出”）就越多；这些产品和服务被销售出去以后，你就能有更多的钱，去投资建设更多的工厂和机器设备。

就这样，你做得越大，赚得越多，就能做得更大。这是一个增强回路，也是任何经济体系增长的核心引擎。

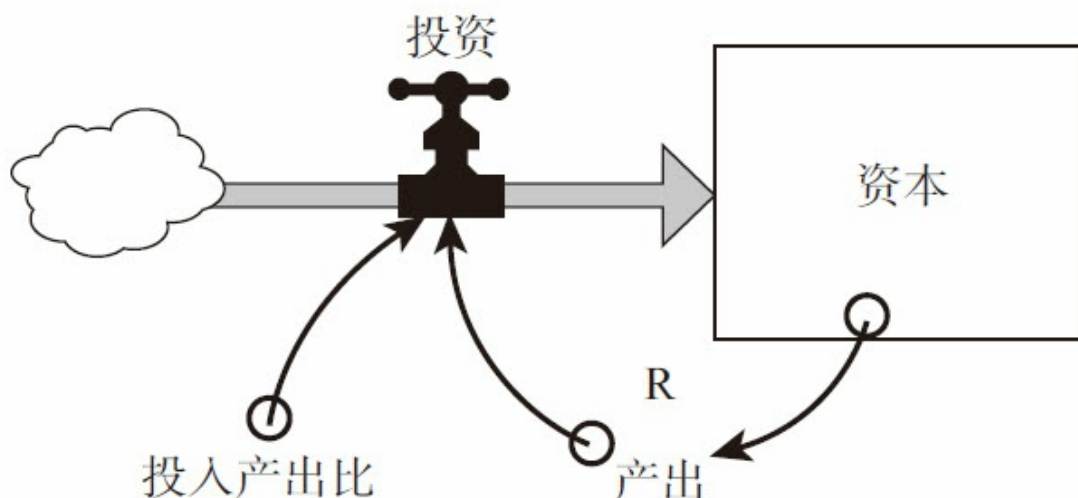


图1—14 资本再投资推动成长

到现在为止，你可能已经了解了调节回路和增强回路对于一个系统是多么重要，它们是系统的基础。有时候，我会让学生们试着去想象一下，如果没有反馈回路，我们人类在制定决策时会怎么样。也就是说，我们要在没有任何关于存量水平信息的情况下，做出与此相关的某项决策，结果会怎么样呢？请你也去思考一下吧。你对反馈回路了解得越多，就会发现它几乎无所不在。

关于“无反馈”型决策，我的学生们提到最多的就是恋爱和自杀。但是，我对此持保留意见。我也想把这一问题留给你自己去思考，看看要做出恋爱或自杀这样的决策是否真的可以不包括反馈回路在内。

系统之思

想一想：如果A能引起B，那么B是否也有可能引起A呢？

我想提醒大家注意的是，如果你发现反馈回路无处不在，那么你已经处于成为一名系统思考者的“危险处境”之中了。因为你不只是看到了A如何影响B，也会开始探究B是否会以某种方式影响到A，以及A是否会增强它自身，或者相反。当你看到电视台晚间新闻上说，美联储要出台某些措施以控制经济时，你也会了解到国家的经济体系肯定会有某些应对措施，反过来影响到美联储。当某人告诉你人口增长导致了贫困

时，你也可能会问自己，贫穷是否会以某种方式导致人口的增长。诸如此类，随着你对世界的认识日益加深，可能会使自己深陷其中、痛苦不堪，因为世界是如此复杂，想要看清楚非常困难，甚至是不可能的。

系统提示

对增强回路和时间翻倍的提示

因为我们经常会遇到增强回路，所以很容易知道这一速算诀窍：对于指数级增长来说，存量翻倍所花费的时间，约等于70除以增长率（以百分数来表示）。

举例来说，如果你把100美元存入银行，年利率是7%，那么10年后，你的钱会翻一倍（ $70/7=10$ ）；如果利率只有5%，那么这笔钱要翻倍就需要花14年时间。

这样，你看到的世界就不再是静态的，而是动态的。你将停止抱怨、指责他人，而是开始探寻“系统究竟是怎样的？”反馈的概念让我们看到，系统本身就可以产生其自身的行为，这是认识世界的一把新钥匙。

到现在为止，我们所探讨的案例，都是在一个例子中只包含一个或一类反馈回路。当然，在真实的系统中，根本不是这个样子。同一个系统中会存在很多不同类型的反馈回路，它们经常以异常复杂的方式相互联系在一起。即使某个单一的存量，也有可能同时受到好几个增强回路和调节回路的影响，它们的力度不同，作用方向迥异。某一个流量也可能受到3个、5个、10个、20个存量的影响。它们可能使某个存量增加，而使另外一些存量减少，或者又可能引发一系列决策，去调整另外一些存量。在一个系统中，有如此之多的反馈回路彼此联系在一起，相互影响：有的试图使存量增长，有的想使其消亡，或者努力让彼此保持平衡。正如你所见到的结果，复杂系统的行为是复杂多变、异彩纷呈的，难以预测和驾驭，绝不只是保持稳定或平滑地趋向一个目标、呈指数级增长或加速衰败这样简单。我们在后文中会见识到这一点。

[1] 原文为“balancing feedback loop”，亦可译为“保持平衡的回路”。——译者注

[2] 原文为“reinforcing feedback loop”，亦可译为“不断强化的回路”。——译者注

第2章 系统大观园

爱因斯坦

所有理论的目标都是将基本要素尽可能减少和简化，而不是考虑完整地呈现真实的体验。

学习一项新事物最好的方法之一，是通过具体的范例，而不是抽象的理论。所以，在本章中，我会给出几种常见、简单但很重要的系统范例，来帮助大家更好地理解系统，包括复杂系统的一些基本原则。

这就像我们去逛一个动物园，有利也有弊。好处是，我们可以在一个地方快速看到很多不同种类的动物，让自己对动物有一个总体的概念；但问题是，动物园中的动物远非动物的全部，我们只不过接触到部分代表，而且按科属进行了分类——这边是猴子，那边是熊。虽然你能通过与熊的对比，观察到猴子的行为特征，但这样的观察是有缺陷的：一是动物园为便于管理，将各种动物彼此分隔开；二是动物园无法真实地再现动物们的生存环境。而在大自然中，各种动物是混杂在一起的，相互影响，与生态环境密不可分。

因此，我们在这里所提到的几类系统，在真实的情境中通常也是相互关联、相互作用的；不仅如此，它们也会和我们没有提到的其他一些系统相互影响，共同组成了我们身处其中的各种复杂系统，嘈杂、喧闹、纷繁、多变。

好了，现在，让我们把生态系统放到一边，先走进“系统大观园”，一种动物接一种动物地看看吧。

单存量系统

系统1.1：一个存量、两个相互制衡的调节回路的系统

典型代表：温度调节器

在上一章所提到的咖啡冷却的例子中，你已经认识了调节回路逐渐趋向于某个目标（“寻的”）的行为特征，那么，如果有两个这样的回路，情况会怎么样？是否会牵引着一个存量朝向两个不同的目标变化？

这类系统的一个典型范例是居室的温度调节器装置，它控制着你房间的制热或制冷（如果是传统的火炉，只能制热；如果是现代的空调，就既能制热，也能制冷）。跟其他模型一样，图2—1中的温度调节器是

一个简化了的家用制热系统的工作原理。

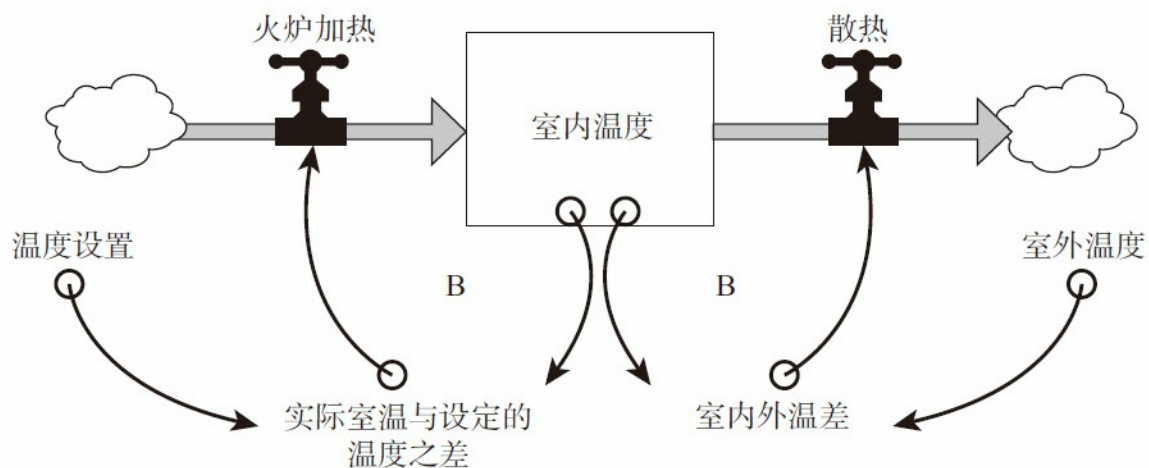


图2—1 受温度调节器和火炉控制的室内温度

这一系统的工作原理很简单：当室温低于设定的温度时，温度调节器探测到这一差异后会发出启动火炉加热的信号，从而提高室内温度；当室温升高，超过设定温度时，温度调节器则不再加热。这一调节回路由一个存量维持（如图2—1中左半部分所示）。如果系统中没有其他因素，并且你所设定的室温是 18°C ，那么该系统的运作情况将如图2—2所示。由于一开始室内温度很低，火炉会打开并开始工作，房间里的温度逐渐提高，当室温达到设定的温度时，火炉会关掉，房间会一直保持在你所设定的目标温度。

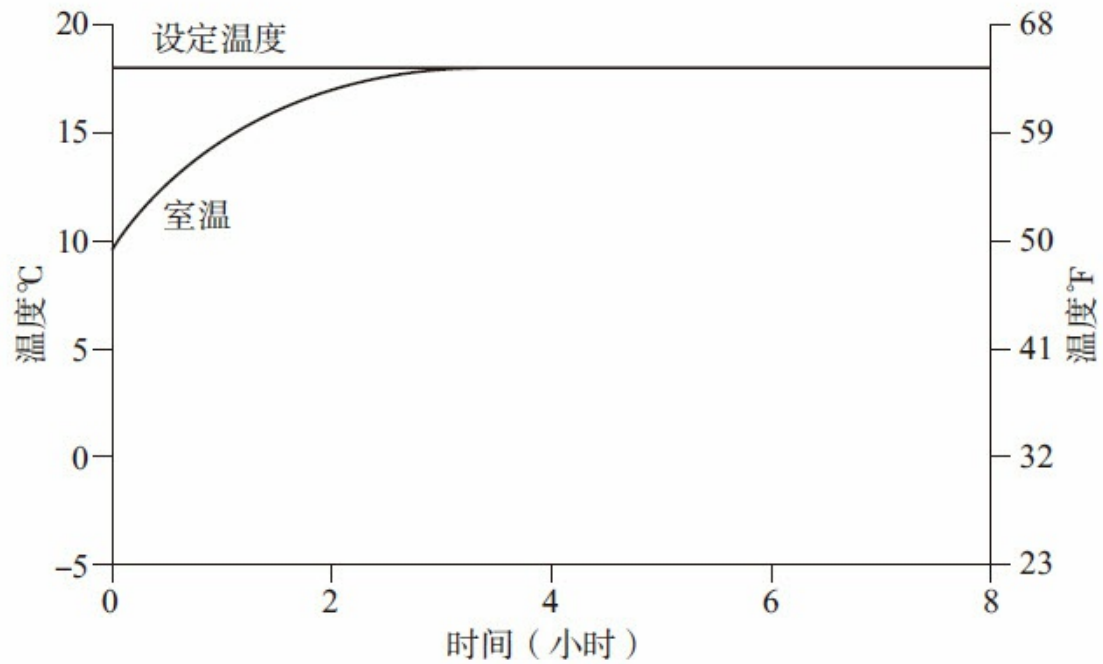


图2—2 室温快速升高到设定的温度

然而，事实并非如此，这也不是系统中唯一的回路，因为热量会散失到室外。热量流出的调节回路如图2—1中的右半部分所示。就像咖啡冷却的案例一样，这一回路的目标是使室内外温度一致。如果这是系统中唯一的回路的话（也就是说假设没有火炉），系统运作的情况就会如图2—3所示，一开始室内较温暖，然后热量逐渐散失，室内温度逐渐下降，最后和室外温度相差无几。

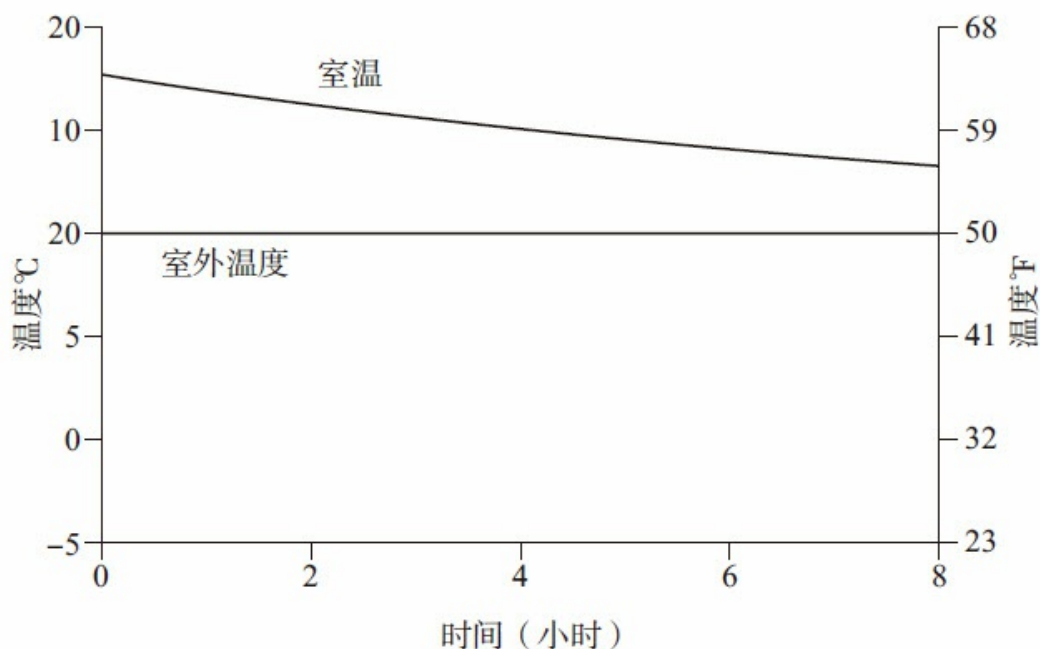


图2—3 室温慢慢下降接近室外温度

以上这种情况是假设房间的保温效果不是很好，由于存在室内外温差，一些热量会从室内散失到室外。房屋的保温效果越好，温度降低的速度就会越慢。

好了，现在让我们看一看，当以上两个回路同时运作时，情况会怎么样？假设房间的保温效果足够好，火炉的功率很充足，制热的回路将取代制冷的回路居于主导地位；结果是，房间里会很温暖，即使刚刚你还在冰天雪地的日子里处于一个冰冷的房间里（如图2—4所示）。

随着室温的升高，向外流出的热量也在增加，因为室内外温差加大，但是由于火炉持续加热，流入的热量超过了流出的热量，所以室温会逐渐达到目标温度。至此，火炉释放的热量与从室内流失的热量达到了均衡。

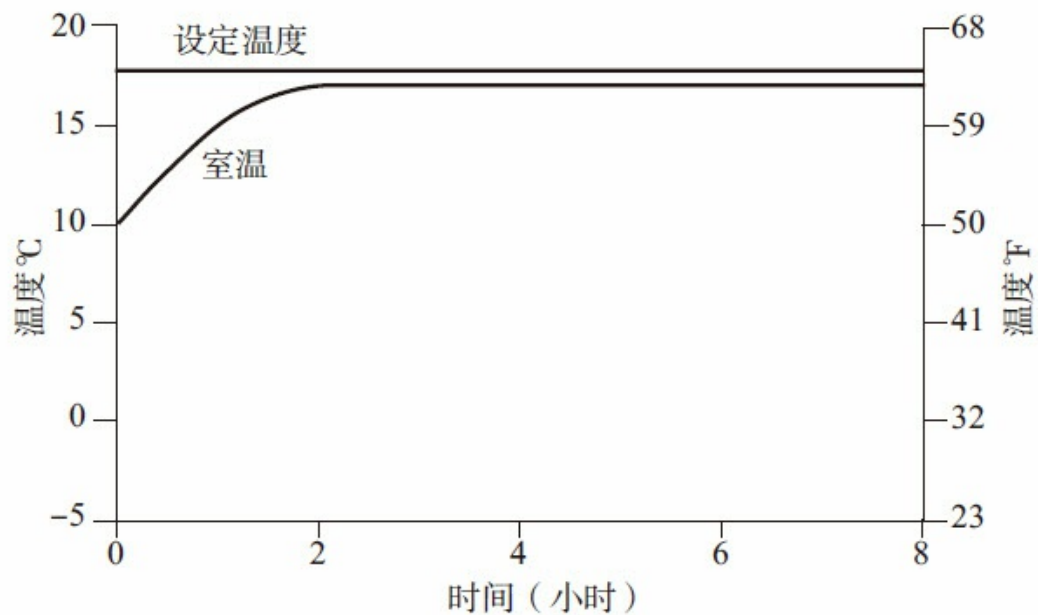


图2—4 制热回路取代制冷回路居于主导地位

在这个案例中，虽然温度被设定为18℃，但均衡的室温会略低于18℃。这是因为存在向外的热量散失——即使火炉因室温未达到设定目标值而一直加温，但同时仍有一些热量在源源不断地流失到室外。这是两个相互矛盾的调节回路，这样的系统会呈现类似的特征，有时会产生人们意想不到的结果。就像你试图让一个底下有漏洞的水桶装满水一样，一切尝试都是徒劳的。更糟糕的是，如果漏出的水受一个反馈回路的影响，水桶里的水越多，水桶底的水压越大，从破洞中流出的水流量也会增加。在上述案例中，如果我们试图让室内比室外更暖和一些，那么屋里越暖和，向外的热量散失也会越快。这就要让火炉花更多的时间去工作，以弥补更多的热量散失，与此同时便伴随有更多的热量在散失。因此，保温效果更好的房间，热量散失更慢，这样往往比一个保温效果差却装备着一个大火炉的房间更令人感到温暖舒适。

对于家用制热系统，人们已经知道需要将温度设定得比他们实际需要的温度稍高一点。当然，具体高多少，是一个相对棘手的问题，因为在越冷的日子里，热量向外散失的速率越高。对此，人们并不会太精确地计算或控制，只要大致地设定一个自己感到舒适或能够接受的温度即可。

但事实上，在现实生活中，对于其他与此有着同样结构（一个存量、两个相互制衡的调节回路）的系统来说，系统中的存量会持续地变

化，如果你试图去控制它，可能会产生一些问题。

例如，假设你试图将商店里的库存量维持在特定水平，由于从订购到进货存在一定的时间延迟，你不可能立即以新货物补足已售出的货物，也无法准确地预计在你等待订购的货物到来之前可能售出多少货物，你就可能面临断货的风险，库存量不可能一直充裕。同样，相似的状况还会出现在下列情况下：公司试图保持现金收支的平衡、水库试图保持蓄水量的稳定，或者化工厂试图在一套连续反应装置中保持化学药品的浓度一致。

系统之思

由反馈回路所传递的信息只能影响未来的行为。它不能足够快地发送一个信号，去修正由当前反馈所驱动的系统行为。哪怕是非物理性的信息，也需要时间反馈到系统之中。

从这里，我们可以得出一条很重要的系统基本原则：由反馈回路所传递的信息只能影响未来的行为，不能立即改变系统当前的行为。因为信息经由反馈回路的传递需要时间，如果你根据当前反馈做出了一项决策，它不能足够快地发送一个信号，去修正由当前反馈所驱动的系统行为，这期间必然有一定的延迟。所以，你的决策只能影响未来的行为，不能改变当前的系统行为。

系统之思

在一个由存量维持的调节回路中，设定目标时，必须适当考虑补偿对存量有重要影响的流入和流出过程。否则，反馈过程将超出或低于存量的目标值。

为什么说这条原则很重要呢？因为它意味着，在行为与结果响应之间经常会有时间延迟。也就是说，一个流量不能立即对其自身做出调整，它只能对存量的变化做出反应，而这必然是在一段时间的延迟之后，等待信息反馈达到一定程度。至于时间延迟的长短，取决于具体的系统情境。例如，对于一个浴缸，你可能只需要花很少时间，就可以估计出水的深度，从而决定调整水的流量。但对于一个复杂的经济系统，一个决策可能需要很长时间才能见到效果，而信息的反馈通常非常缓慢、微妙、杂乱，难以把握。因此，很多人在对经济学相关问题进行建模时，经常假设消费或生产会快速地对诸如价格等要素的变化做出反应，这肯定是错的。真实的经济系统肯定不是这样运作的。

从温度调节器这一简单系统中，我们还可以得出一条具体的原则：在类似的系统中，流量的散失和补充过程是持续的、动态变化的，不能静止地看。如果意识不到这一点，存量的目标水平就难以维持。

例如，假设你希望室内温度达到18℃，你必须考虑到热量的持续散失，从而将温度设定得略高于期望值；如果你想偿付自己的信用卡欠款，你必须考虑到利息支出和期间的开支，从而稍微提高偿还金额，以便补足还款期间所发生的费用；如果你希望增加员工人数，必须尽快地

招聘到位，以防在招聘期间再有员工离职。

换句话说，对于类似的系统，你必须考虑到所有重要的流量，否则系统的行为就可能让你大失所望。

在我们结束这一小节之前，让我们看一看当外部温度波动时，温度调节器系统是如何运作的。图2—5显示的是一个正常工作的温度调节器系统在24小时内的一般运作情况，同期室外温度在夜间降低到了冰点以下。由于火炉供给的热量很好地弥补了向室外散失的热量，在室内暖和起来之后温度几乎没有变化。

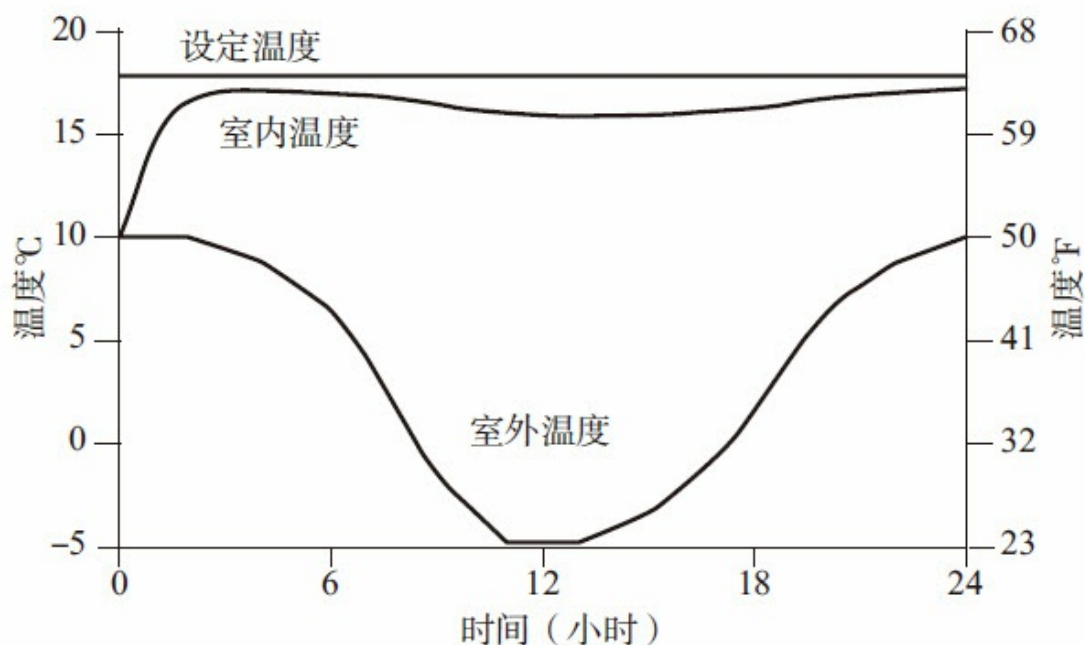


图2—5 当气温很低时，保温效果较好的房间室温变化情况

每一个调节回路都有它的转折点，此时其他回路会取代该回路而居于主导地位，使存量远离它的目标且无法自动回到动态平衡状态。这在温度调节器系统中也可出现。例如，你减弱火炉制热的功率（假设更换一个更小的火炉，或少放一些柴火），或者加大制冷回路的影响力（如室外温度更低、降低保温效果，或者打开门窗）。图2—6反映的就是室外温度与图2—5相同但热量散失更快的情况。在这种情况下，当室外温度较低时，火炉不能确保供给足够多的热量，在一段时间里，使室温降低并逐渐趋向室外温度的回路占据了主动，于是房间里的温度显著降低了。

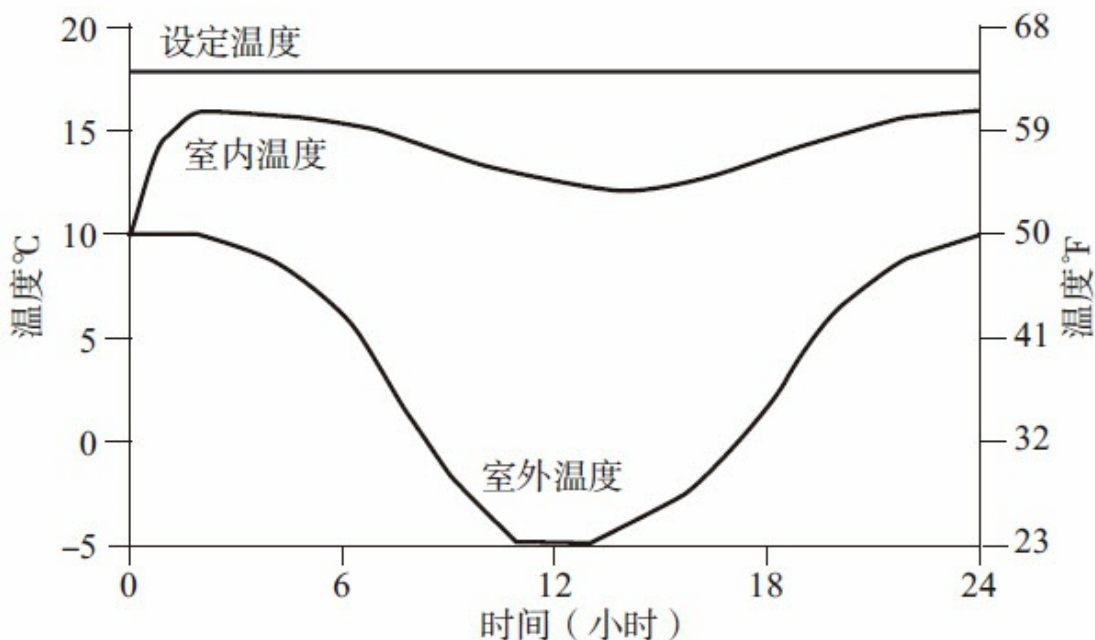


图2—6 当气温很低时，保温效果不好的房间室温变化情况

如果你能参照本书附录中的公式做出这个模型，你就可以看到，随着时间的推移，图2—6中的各种变量如何相互关联。一开始，室内外的温度都很低；由于火炉供给的热量超过了散失到屋外的热量，室内逐渐暖和了起来。经过一两个小时之后，由于室外温度逐渐升高，火炉源源不断补给的热量完全弥补了外流的热量，室温就能达到并保持在接近期望的温度。

但是，当室外气温开始下降时，外流的热量也加大了，虽然火炉还是全速地工作，但产生的热量不足以很快地弥补二者的缺口（即火炉产生的热量少于外流的热量），于是室温开始下降。最后，室外温度又开始回升，外流热量减少，火炉补充的热量终于又占了上风，室温开始回升。

按照我们在浴缸案例中提出的法则，当火炉补充的热量超过外流的热量，室温就会上升；反之，温度就会降低。如果你能认真研究一下这一系统行为图的变化，并将其与系统的反馈回路图结合起来，你就能很好地理解系统在结构上如何相互关联，以及结构如何影响系统的行为——该系统有两个反馈回路，它们之间相互影响，此消彼长，随着时间推移而动态变化。

系统1.2： 一个存量、一个增强回路以及一个调节回路的系统

典型代表：人口和工业经济

如果有一个增强回路、一个调节回路，同时作用于一个存量，情况会怎样？这其实是最常见、最重要的系统结构之一。在众多案例中，与我们每个人都息息相关的人口和工业经济体系就是这样的系统（如图2—7所示）。

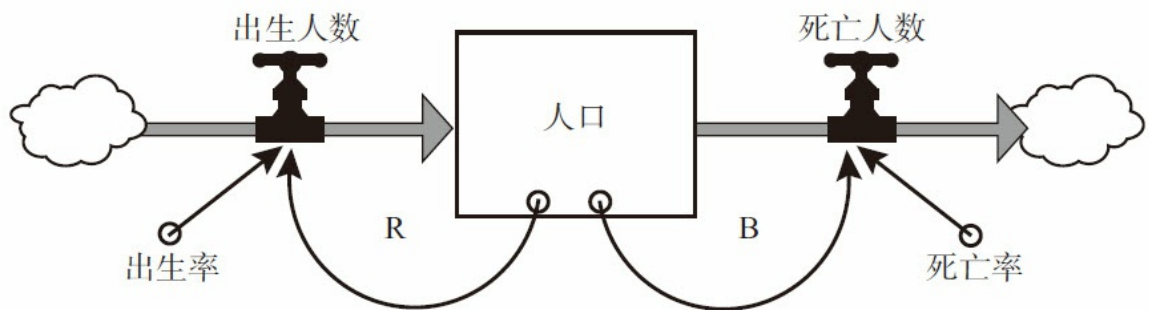


图2—7 人口受出生（增强回路）和死亡（调节回路）的影响

人口受到一个增强回路和一个调节回路的影响：增强回路决定新出生的人数，受出生率的影响，导致人口数量增长^[1]；调节回路影响当期死亡的人数，受死亡率的影响，导致人口数量减少^[2]。

如果出生率和死亡率是常数（在真实情况下很少如此），这一系统的行为就很简单：它将以指数方式增长或减少。至于变化的方向，取决于决定出生人数的增强回路和决定死亡人数的调节回路谁的效果更强。

例如，2007年全球人口总数为66亿，出生率约为21‰，死亡率是9‰。由于出生率高于死亡率，增强回路占了主导地位，人口总数以指数级方式增加。如果出生率和死亡率一直保持不变，那么，一个现在出生的孩子，到他或她60岁时，将看到世界人口已经翻了一番还多（如图2—8所示）。

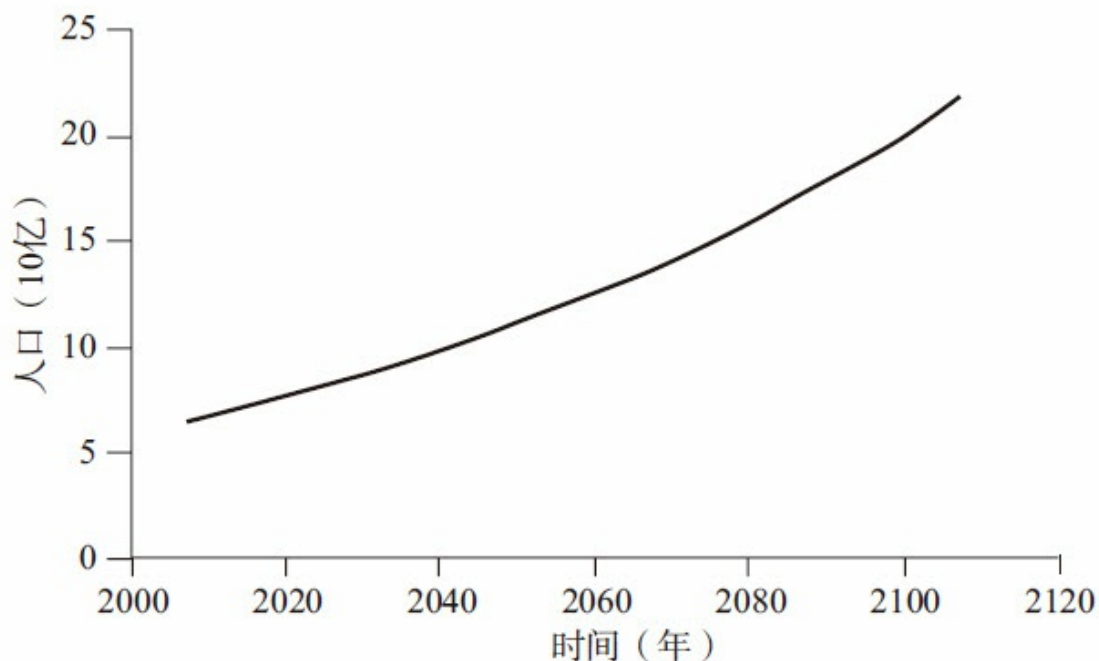


图2—8 世界人口增强回路

按照2007年的出生率和死亡率，世界人口将呈现指数级增长态势。

假设由于一次可怕的瘟疫，死亡率提高了，比如说达到了30%，而出生率仍保持在21%，那么决定死亡人数的调节回路将占据主导地位。由于每一年死亡的人数多于当年新出生的人数，人口总数就会逐渐下降（如图2—9所示）。

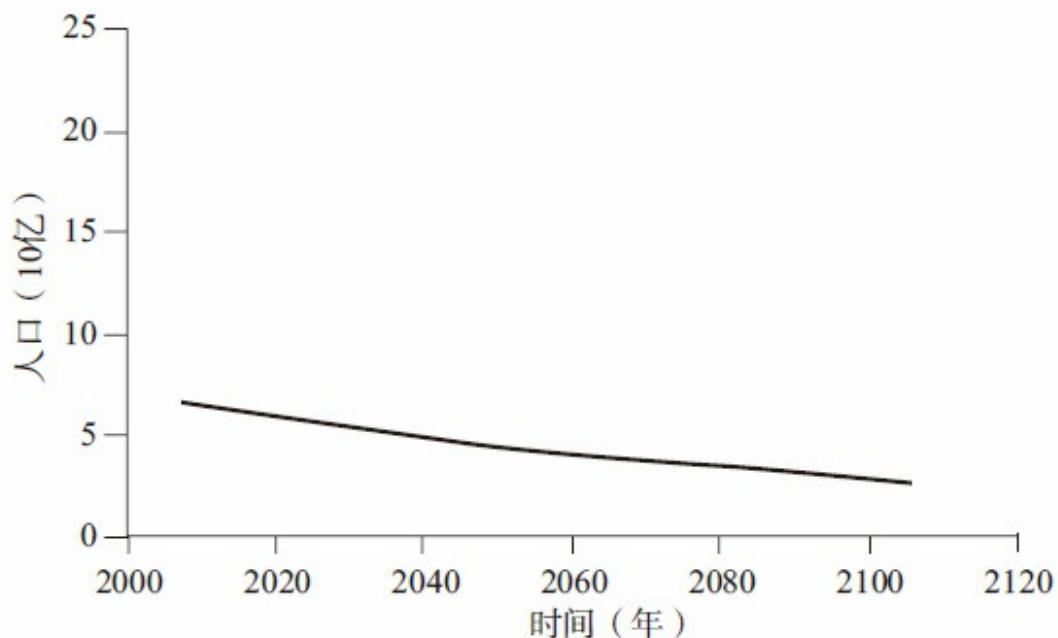


图2—9 世界人口调节回路

如果死亡率达到30‰，而出生率维持在2007年21‰的水平，世界人口将呈现指数级减少的态势。

如果出生率和死亡率不是常数，都会随着时间变化，事情就变得更有趣了。过去，联合国在进行长期人口预测时，通常会假设随着经济的发展，各国的平均出生率会下降；同样，一直到最近，人们通常会假设死亡率也会逐渐降低，但降低的速度会慢很多，因为世界上大部分国家的死亡率已经比较低了。然而，由于艾滋病的蔓延，联合国现在假设，在未来50年中，一些受到艾滋病影响的地区，人均寿命增长的趋势将会减缓。

受流量（新生和死亡）变动的影响，系统中存量（人口）的行为也会随时间而变化，但变化曲线是不规则的。例如，如果全球的出生率稳步回落，到2035年与死亡率持平，并且二者在此后均保持稳定，世界总人口变化将趋向平稳：每年新出生的人数刚好等于当年死亡的人数，达到动态平衡（如图2—10所示）。

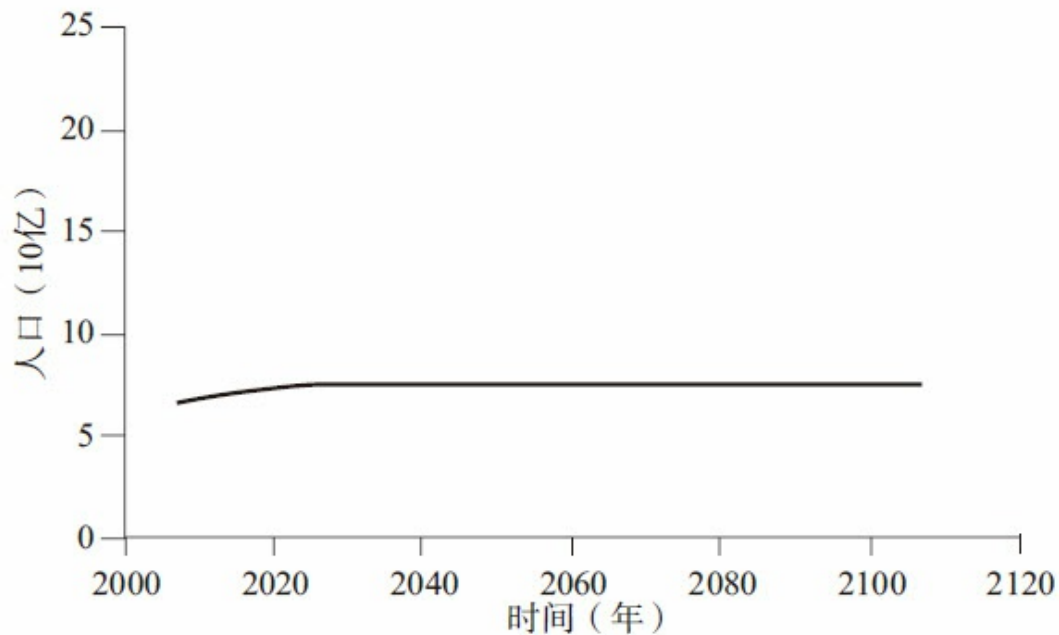


图2—10 世界人口动态平衡

若出生率等于死亡率，人口将保持稳定。

这一行为是反馈回路之间“主导地位转换”的例子。“主导地位”是系统思考中的一个重要概念，当一个回路相对于另外一些回路居于主导地位时，它对系统的行为就会产生更强的影响力。虽然系统中经常有好几

个相互矛盾的反馈回路同时在运作，但只有那些居于主导地位的回路才能决定系统的行为。

在人口系统中，一开始，由于出生率高于死亡率，推动人口增长的增强回路就居于主导地位，所以系统的行为是指数级增长；但是，随着出生率的降低，这一回路的影响力逐渐弱化；最后，它刚好等于与死亡率相关联的调节回路的力量。在这一点上，没有哪个回路占据主导地位，系统最终达到动态平衡状态。

在温度调节器系统中，当室外温度降低而热量从保温效果不好的房间里流失，超过了火炉补充的热量时，你也可以发现“主导地位转换”的现象。此时，冷却回路取代了加热回路而居于主导地位。

系统之思

当不同调节回路的相对优势发生改变时，系统常会出现一些复杂的行为，由一个回路主导的某种行为模式会变为另外一种。

人口系统只有少数几种行为模式，而这取决于出生率、死亡率等关键变量的状况。对于只有一个增强回路和一个调节回路的简单系统来说，可能性就那么几种。如果增强回路居于主导地位，与增强回路和调节回路相关联的存量就会呈现指数级增长；如果调节回路居于主导地位，存量就逐渐衰退（趋向于一个目标）；如果两个回路势均力敌，存量就会维持在一个特定水平上（如图2—11所示）。如果两个回路的相对优势随时间而变化，出现“轮流坐庄”的局面，系统就会波动（如图2—12所示）。

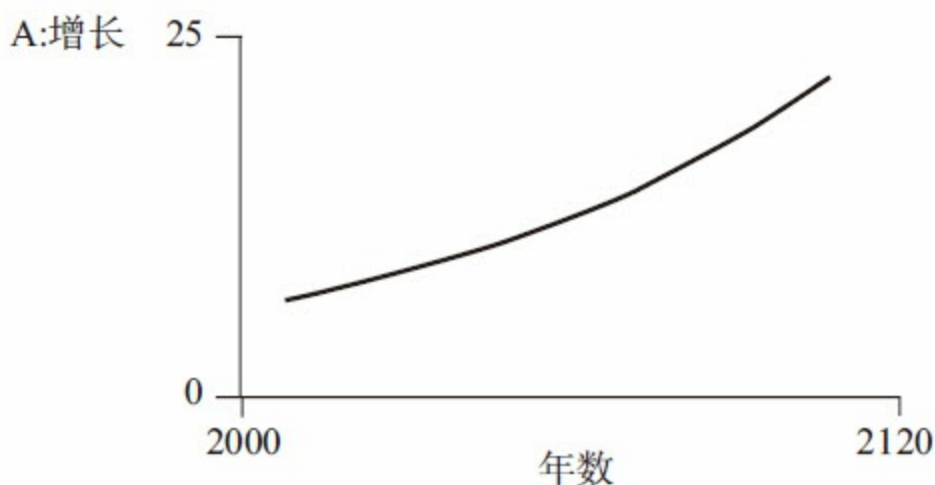


图2—11 人口系统的三种可能行为：增长、衰退和稳定

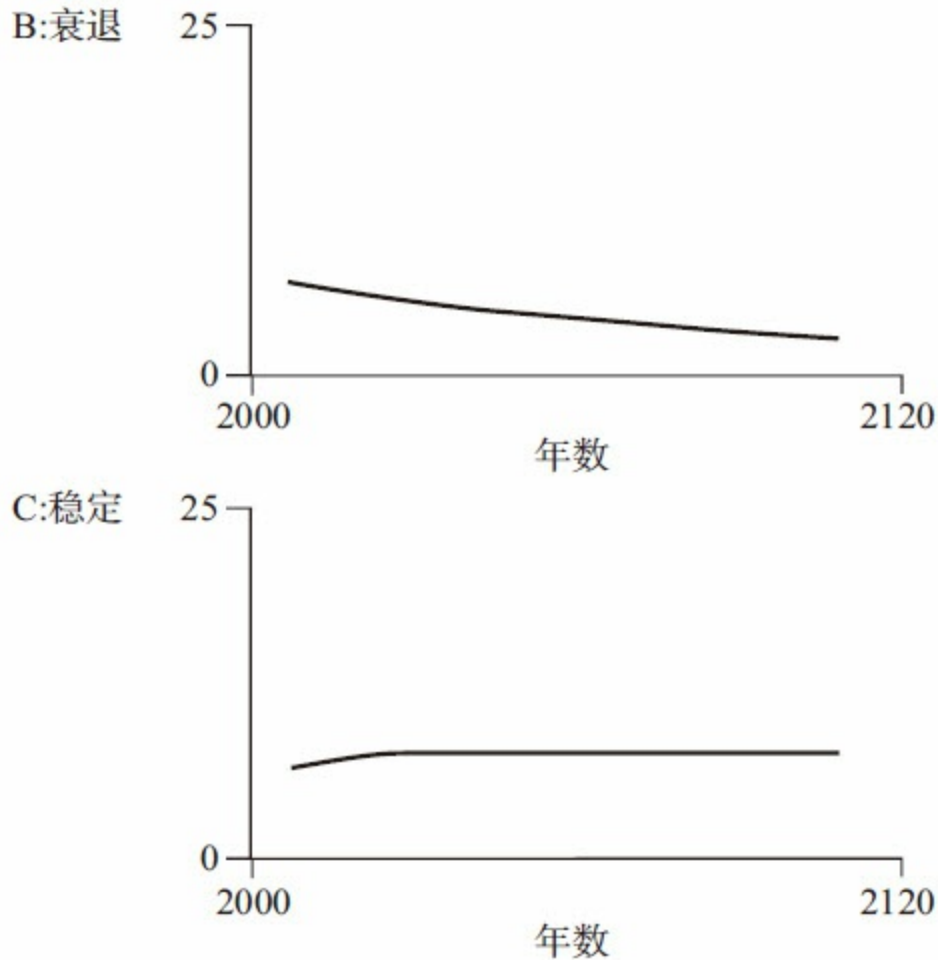


图2—11 人口系统的三种可能行为：增长、衰退和稳定（续）

这里，我选取了一些极端的人口情景，以展示模型中的一些要点以及它们所产生的新情景。当你面对某个情景时，例如听到经济预测、公司预算、天气预报、未来气候变化以及股票经纪人关于某只股票价格走势的预测等，你可以思考下列问题，以帮助你判断这些表述是否真实有效地反映了系统潜在的结构。

- 各种驱动因素会不会以这种方式发挥作用？例如，出生率和死亡率可能处于什么样的水平？
- 如果驱动因素这样发挥作用，系统将以何种方式应对？例如，出生率和死亡率真的会像我们想象的那样影响人口存量的变化吗？
- 影响各种驱动因素的又是什么？例如，什么会影响出生率，什么会影响死亡率？

实际上，第一个问题很难回答，因为这是对未来的猜测，而未来从

本质上讲是不确定的。虽然你可能认为自己对未来很有把握，但除非未来真正到来了，否则仍然无法验证你的观点是否正确。通过系统分析，人们可以进行一系列情景测试，以便观察当各种驱动因素处于不同状况时，系统状况如何。这通常是系统分析的一个目的。但是，你仍然需要拟订各种情景，并判断哪种或哪些情景可能发生，需要认真分析。

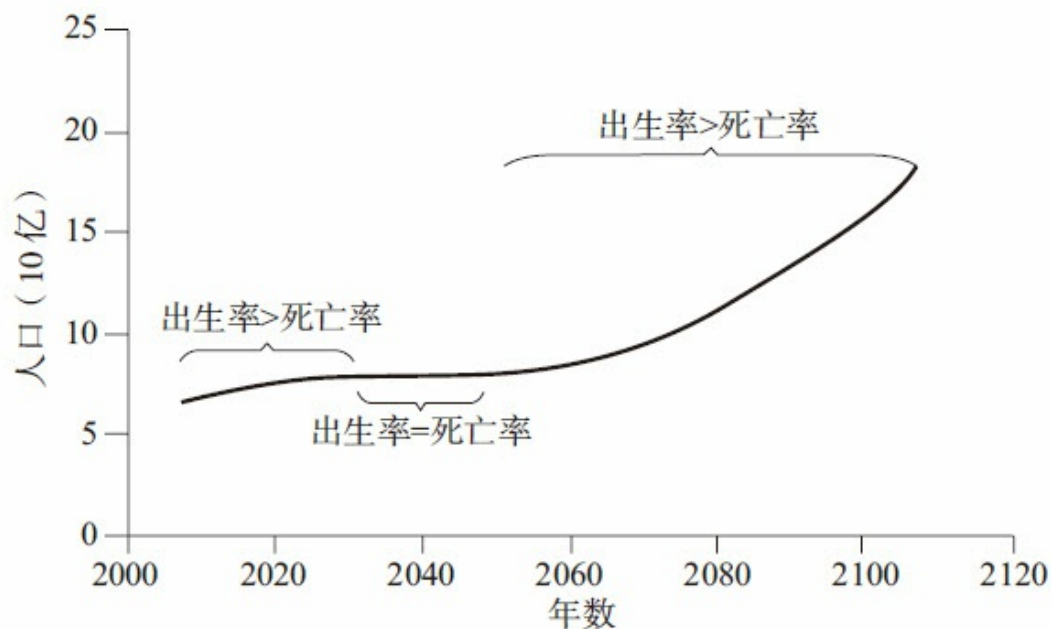


图2—12 驱动出生和死亡的反馈回路发生主导地位转换

动态系统分析的目的通常不是预测会发生什么情况，而是探究在各种驱动因素处于不同状况时，可能会发生什么。

系统提示

测试模型价值的问题

- 1) 各种驱动因素会不会以这种方式发挥作用？
- 2) 如果驱动因素这样发挥作用，系统将以何种方式应对？
- 3) 影响各种驱动因素的又是什么？

相对于第一个问题，第二个问题更为科学，而这取决于模型的质量。如果模型质量高，能够反映系统特有的动态性，我们就可以更好地解释，当某种或某些驱动因素以一种方式变动时，系统可能以何种方式应对。

系统之思

系统动力学模型可探究未来的多种可能性，让我们进行“如果……会怎么样”这样的思考。

如上所述，对于人口系统，不管你认为出生率与死亡率变化的可能性怎么样，对第二个问题的答案大致是肯定的，因为我们这里所使用的人口模型是很简单的，其行为变化模式只有那么几种可能性。当然，我们可以进一步使其精细化，例如对人群按年龄划分。但是，上面这个简单模型基本上就能反映真实的人口变动状况，具体数字可能不太准确，但基本的行为模式却是真实的。

系统之思

模型的价值不取决于它的驱动情景是否真实（其实，没有任何人能够对此给出肯定的答案），而取决于它是否能够反映真实的行为模式。

第三个问题“影响各种驱动因素的又是什么”，指的是什么东西会影响流入量和流出量。这是与系统的边界有关的问题，需要认真研究，看看哪些驱动因素是完全独立的，而哪些是系统内部的变量。

例如，人口规模是否会反过来影响到出生率和死亡率？是否还有其他因素，如经济、环境或者社会趋势等，会影响到出生率和死亡率？人口规模是否会影响到这些经济、环境和社会方面的因素？

当然，这些问题的答案都是肯定的。出生率和死亡率也是被一些反馈回路所左右的，其中一些反馈回路本身也受人口规模的影响。人口这一系统本身只是一个更大的系统中的一部分。

在这个更大的系统中，经济是一个很重要的影响人口的子系统。在经济系统的核心，也存在一个“增强回路+调节回路”的系统，其结构与人口系统类似，因此这两个子系统也具有类似的行为模式（如图2—13所示）。

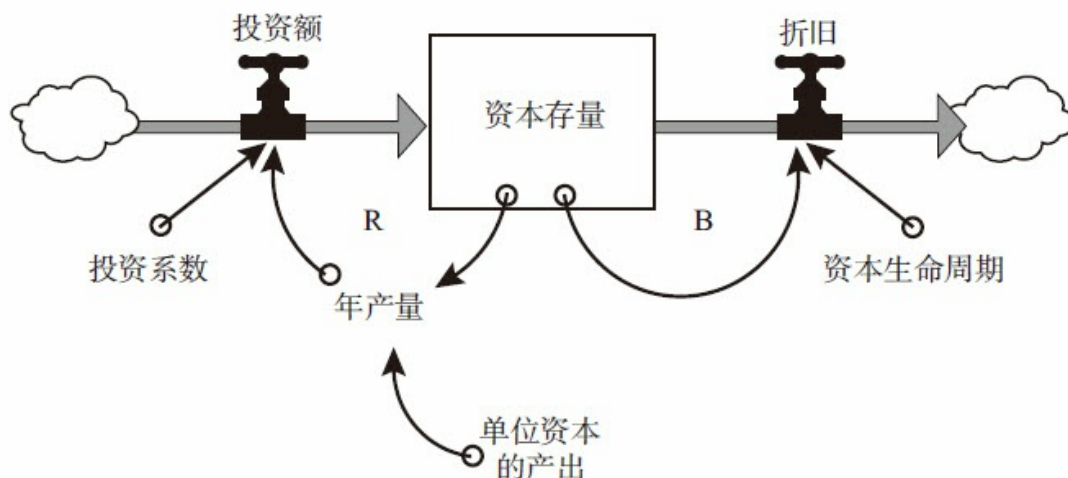


图2—13 经济系统行为变化模式

与人口系统类似，经济系统中也有一个驱动增长的增强回路（外部投资）和一个导致衰退的调节回路（折旧）。

在经济系统中，实体资本（如机器设备和工厂等）的存量越大、生产效率（即单位资本的产出）越高，产量（产品和服务）也就越大。而产量越大，就可以有更多的投资形成新的资本。这是一个增强回路，就像人口系统中新出生人数所在的那个回路。在这里，投资系数相当于出生率。投入产出比越高，资本存量的增长就越快。

另一方面，实体资本会由于折旧和磨损而逐渐消耗。这是一个调节回路，类似于人口系统中死亡人数所在的回路。资本的“死亡率”取决于资本的平均寿命：生命周期越长，每年资本淘汰或置换的比例就越小。

因为这个系统和人口系统结构相同，它们也具有相同的行为模式。和世界人口一样，近年来，世界资本存量也被增强回路所主导，从而呈现出指数级增长态势。至于未来它将继续增长还是保持不变，抑或衰退，取决于影响它的增强回路和调节回路二者孰强。这体现在：

- 投资系数——每年社会上有多少产出用于再投资，而不是被消耗掉；
- 资本的效率——要想获得特定的产出，需要消耗多少资本；
- 资本的平均生命周期。

如果用于再投资的投资系数是一个常数，资本的效率也相对稳定，资本存量的变化趋势将取决于资本的生命周期。在图2—14中，有三条曲线，显示的是三种资本平均生命周期的资本存量变化情况。当资本生

命周期相对短暂时，资本消耗速度快于更新速度，再投资量无法抵消折旧的影响，经济将慢慢陷入衰落；当折旧刚好与投资持平，经济就会处于动态平衡；假如资本的生命周期很长，资本存量将呈指数级增长。资本的生命周期越长，增长的速度越快。

这里体现了系统的一项基本原则：你可以通过减小流出量的速率或者加大流入量的速率，来使存量增长。我们之前已经遇到过，这又是一个典型范例。

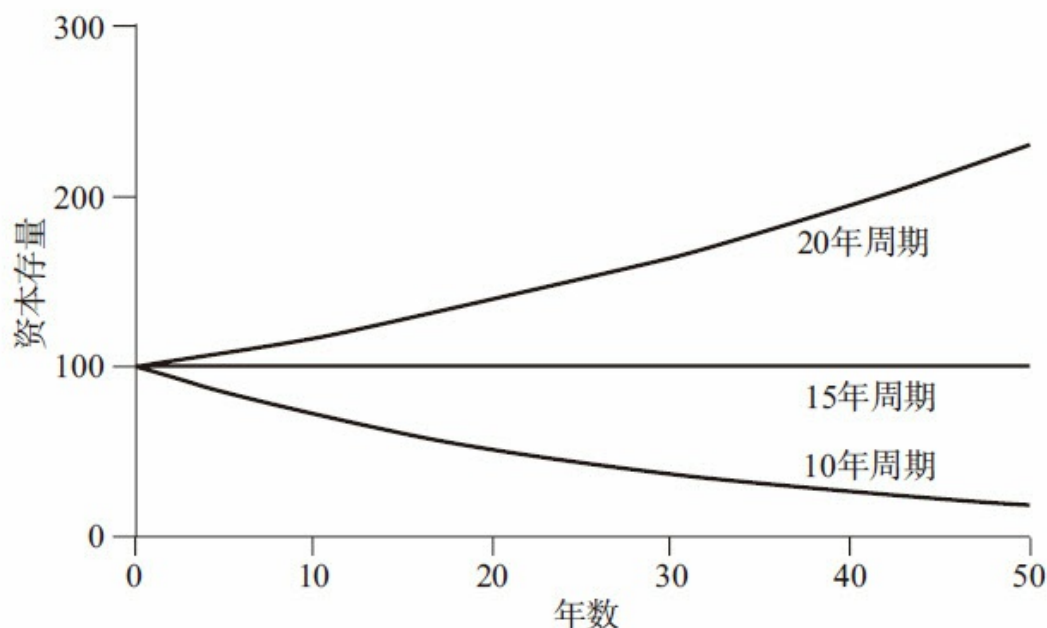


图2—14 资本存量的增长与资本生命周期之间的关系

假设单位资本的产出比为1: 3，投资系数是20%，在这种情况下，资本生命周期为15年时，折旧与资本更新速度持平；资本生命周期越短，资本存量将陷入衰减。

正如人口系统中出生率和死亡率受很多因素的影响一样，在经济系统中，也有很多因素影响产出率、投资系数和资本生命周期，例如利率、技术、税收政策、消费习惯以及价格等。人口系统也会影响投资，包括为产出提供劳动力、增加消费需求，并由此减小投资系数。经济系统的产出也会以多种方式反馈并影响到人口系统，例如经济越富裕的地区医疗保健条件通常较好，从而降低了死亡率；同时，经济富裕地区的出生率通常也较低。

事实上，在长期经济系统建模时，都会考虑经济系统和人口系统以及它们之间的联系，以反映二者如何相互影响。经济发展的核心问题是

如何防止资本积累的速度慢于人口增长速度。这是两个增强回路，前者受经济增长的回路所推动，若前者快于后者，就可以使人们越来越富裕，否则就将陷入越来越贫穷的泥沼之中。

系统之思

具有相似反馈结构的系统，也将产生相似的动态行为。

如果我将上面提到的经济系统与人口系统叫做同一类“动物”，不知道你是否会感到奇怪？一个是关于工厂设备、进货出货和经济往来的经济系统，另外一个是关于婴儿出生、长大、再生儿育女、逐渐变老并最终逝去的人口系统，二者在表象、运作等多个方面似乎都有很大差异。但是，从系统的角度看，它们之间有一个很重要的共同点：相同的反馈回路结构，即都有一个存量，受到一个推动成长的增强回路和一个导致消亡的调节回路的影响，也都有一个老化的过程，无论是用钢筋水泥建造的工厂，还是车床、汽轮机，都会像人一样慢慢变老，并最终寿终正寝。

我们之前讲过，我们所观察到的系统行为主要是由其自身所引起的。同样，系统理论的另外一项核心见解是，具有相似反馈结构的系统，也会产生相似的动态行为，即使这些系统的外部表现是完全不同的。

例如，虽然人口与工业经济系统表面上差异很大，但它们的行为模式却基本相似：可以自我更新，以指数级方式增长，都会逐渐老化和衰亡。之所以如此，是因为它们有相似的系统结构。同样，咖啡杯的冷却与房间室温的降低、放射性物质的衰变、人口或工业经济系统的老化和衰亡也基本相似，这些都是调节回路作用的结果。

系统1.3：含有时间延迟的系统

典型代表：库存

想象一下一家汽车经销商的仓库，它有一定的库存量，也有一个流入量（从各家工厂订货交付的汽车）和一个流出量（因销售给客户而被提走的汽车）。从结构上看，汽车库存量的行为变化模式会像一个浴缸的蓄水量。

现在，让我们看一下这一系统是如何运作的。假设要维持足够10天

销售的库存量（如图2—15所示）。汽车经销商需要保持一定的库存，因为每天到货交付的数量和销售量不可能完美地匹配，而且每天客户的购买量也很难预测。此外，经销商还需要多保持一些额外的库存作为缓冲，以防供应商偶尔出现交货延迟或其他意外情况。

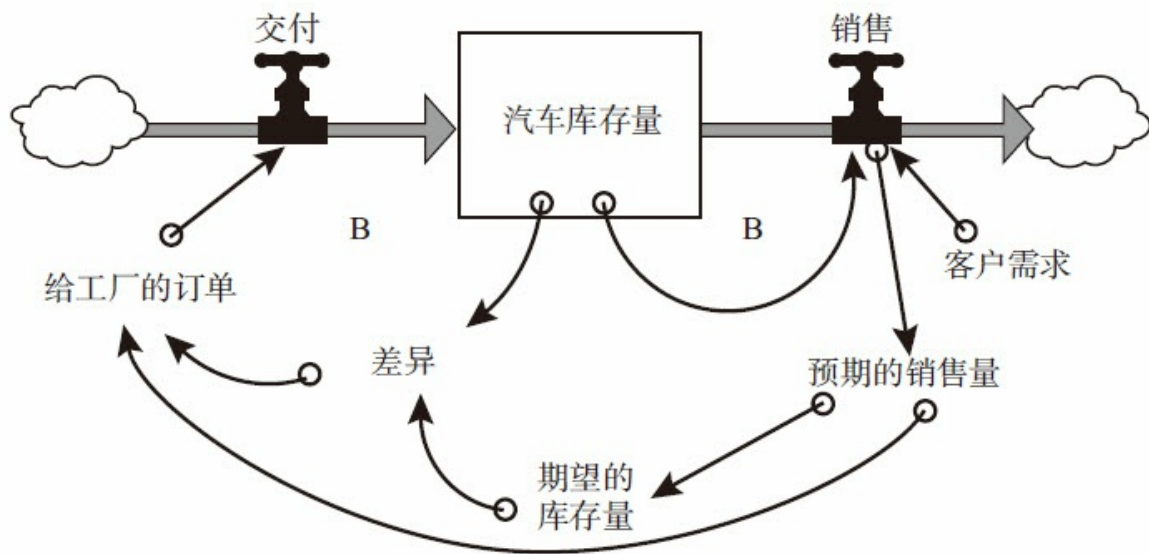


图2—15 汽车库存量的行为变化模式

一家汽车经销商的库存量由两个运转方向相反的调节回路来保持平衡，一个是因销售而产生的流出量，另一个是到货交付的流入量。

经销商会对销售进行监控、分析和预测，例如，如果他们发现销量有增长趋势，就会据此增加给工厂的订单，以便增加库存，满足未来可能加大的销量。因此，销售量变大，意味着未来预期的销售量也会变大，导致实际库存与期望库存之间的差距加大，这将导致向工厂下达的采购订单量增加。一段时间之后，到货量将增加，从而提高库存量，应对未来可能出现的更大销售量。

从结构上看，这一系统是温度调节器系统的翻版，都是一个存量受两个相互制衡的调节回路的影响。在本例中，一个调节回路是把车销售出去，从而导致库存减少，另一个是向供应商订购的车辆到货交付，从而导致库存增加（也可以看做是补足因销售而减少的库存）。图2—16显示的是顾客需求增加了10%的情况。波澜不惊，是吧？

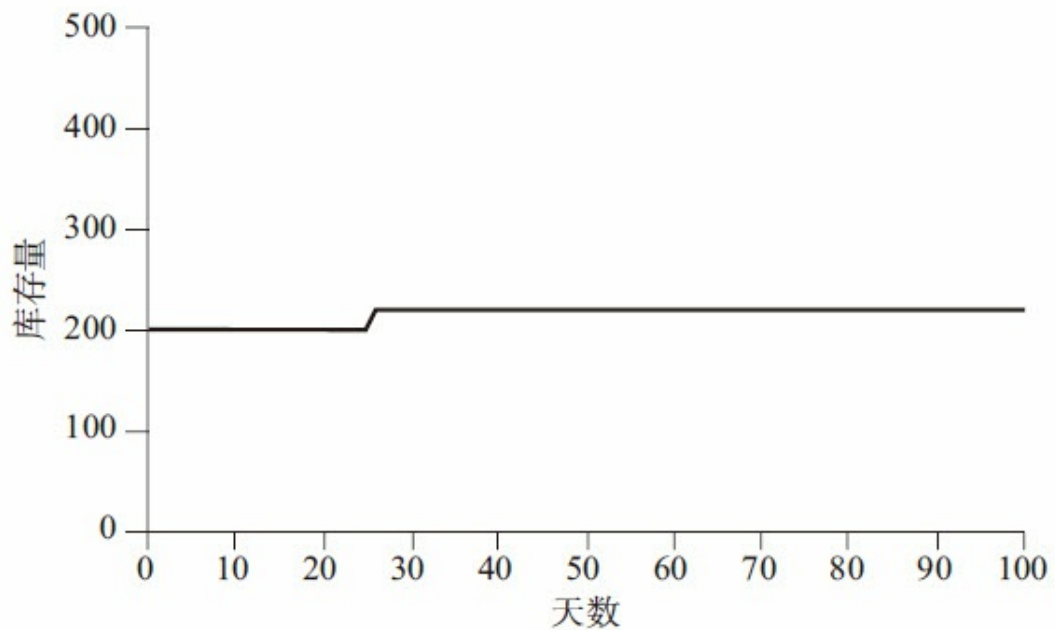


图2—16 库存量变化情况

从第25天开始，顾客需求永久性地增长了10%，相应地，经销商的库存量也增加了这么多。

在图2—17中，我们对这个简单的模型进行了一些微调，加入了三个时间延迟（感知延迟、反应延迟和交货延迟），这些都是我们在现实中经常会遇到的情况。

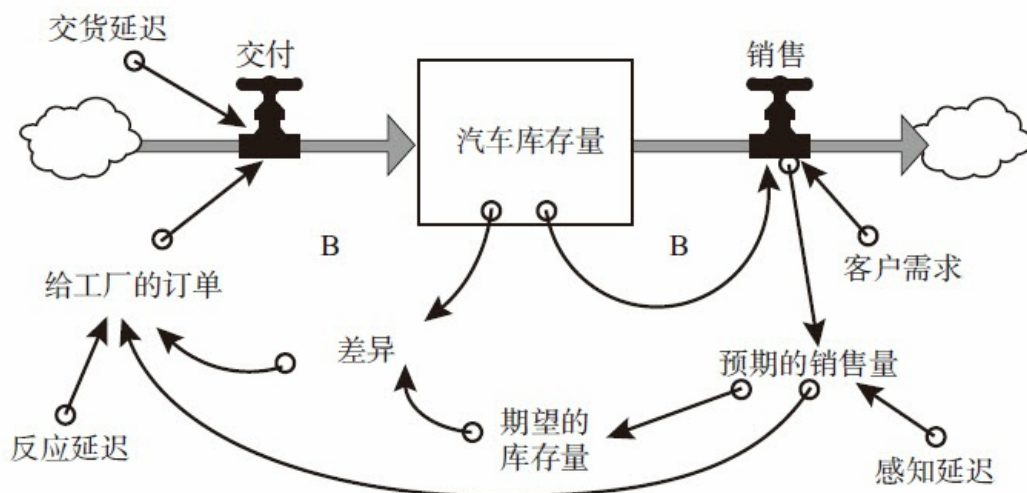


图2—17 有时间延迟的汽车库存量行为变化模式

首先是感知延迟，这与人们的主观认识有关。经销商不可能对销售量的任何变化都立即做出反应。在制定订货数量决策时，人们通常会将

过去一段时间内的销售量进行平均，以发现销售量的变化是真实的趋势，还是短期内的波动或异常。

其次是反应延迟，即当形势已经很明朗了、需要调整订单数量时，经销商也不会在某笔订单里将所有缺货一次性地调整到位。相反，他会在今后的每笔订单中多增加一部分。也就是说，即使当他相信销售量的变化趋势是真实的，他也只会部分地进行调整，以便在今后的几天内进一步确认这一趋势。

第三是交货延迟。从供应商的工厂收到订单、加工生产并发货到交付给经销商，要花5天时间。

虽然这一系统和简单的温度调节器系统一样，只是由两个调节回路构成的，但它的行为变化却和后者有着明显的差异。例如，想象一下，当客户需求 and 随后的销售量增长了10%之后，库存量会有怎样的变化？图2—18是库存量随时间变化的示意图。

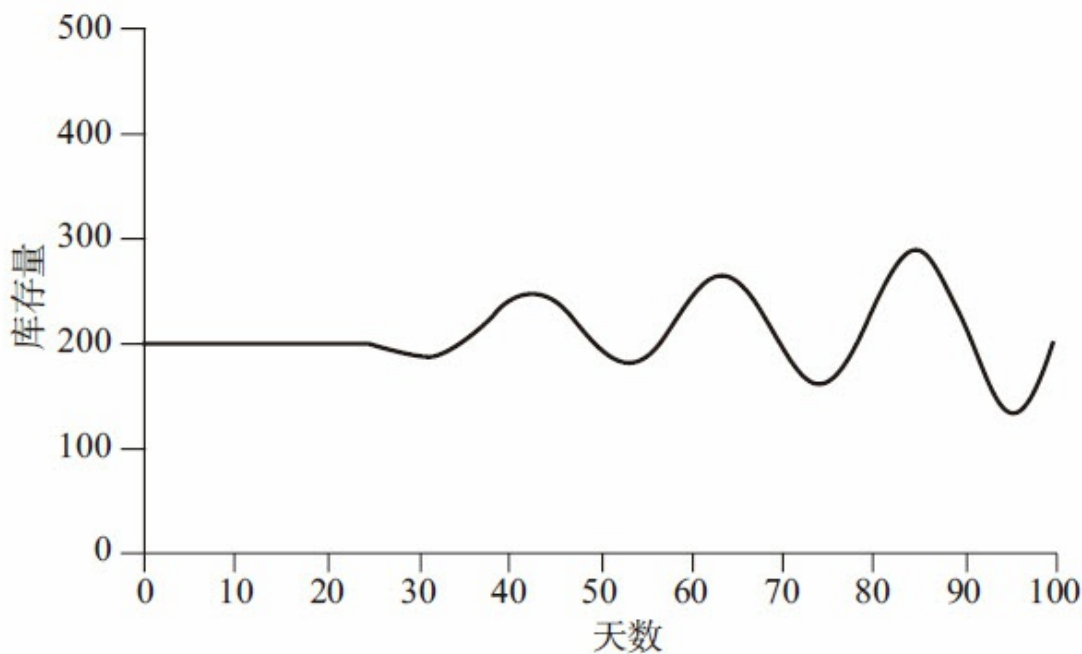


图2—18 库存量随时间变化情况（一）

当系统中存在时间延迟时，库存量会随销售额增加而上下振荡。

是的，剧烈振荡！一开始，销售量的微小增长，导致了库存的下降；经过几天的观察，经销商感知到销售量增长的趋势确实存在，而且将会持续，所以他们开始订购更多的车，不仅要满足增大的销量，而且

要补足之前因销售而导致的库存差距。但是，因为存在交货延迟——即在经销商发出采购订单到其拿到车之间，有一段时间的延迟或滞后，在这期间，库存量继续降低，经销商会进一步加大采购订单的数量。

最后，大量的订货终于源源不断地到货交付，补足了之前的差额，并持续地推高库存量，因为在前一段时间里经销商下了太多的订单。现在，他们意识到了自己的错误，并开始减少订单量。但是，之前所下的大额订单仍然不断地到货交付。因此，他们更大幅度地削减订单量。事实上，由于他们无法洞悉未来的走势，几乎不可避免地会将订单削减得过多。于是，库存会再次变得过低……如此循环往复，库存量会围绕新的预期库存量上下振荡。从图2—18、图2—19可以看出，少数几个时间延迟会造成系统行为多大的变化！

系统之思

调节回路上的时间延迟很可能导致系统的振荡。

稍后我们会探讨有哪些方法可以抑制这些振荡，但首先，很重要的一点是，我们必须理解为什么会产生这些振荡。这不是因为这家经销商的经理或决策人愚蠢，而是因为他们置身于一个缺乏及时的信息反馈的系统之中。由于存在物理的分割和时间延迟，他们很难甚至不可能实时地了解自己的行为对于库存量变化的影响。他们也不知道顾客下一步会做什么；当顾客的行为有了一些变化，他们也不能肯定这些变化是否会持续下去。当他们发出一个订单，他们也不能得到立竿见影的响应。在真实的商业环境中，信息不对称以及物理延迟的情况是非常普遍的。因此，在很多其他系统中，类似这样的振荡也是很常见的。想象一下，当你在酒店中享受淋浴时，如果冷热水混合器与喷头之间有很长的管道（会导致反馈延迟），那么你在调节水温时，很可能会直接地体验到水温振荡的“乐趣”。

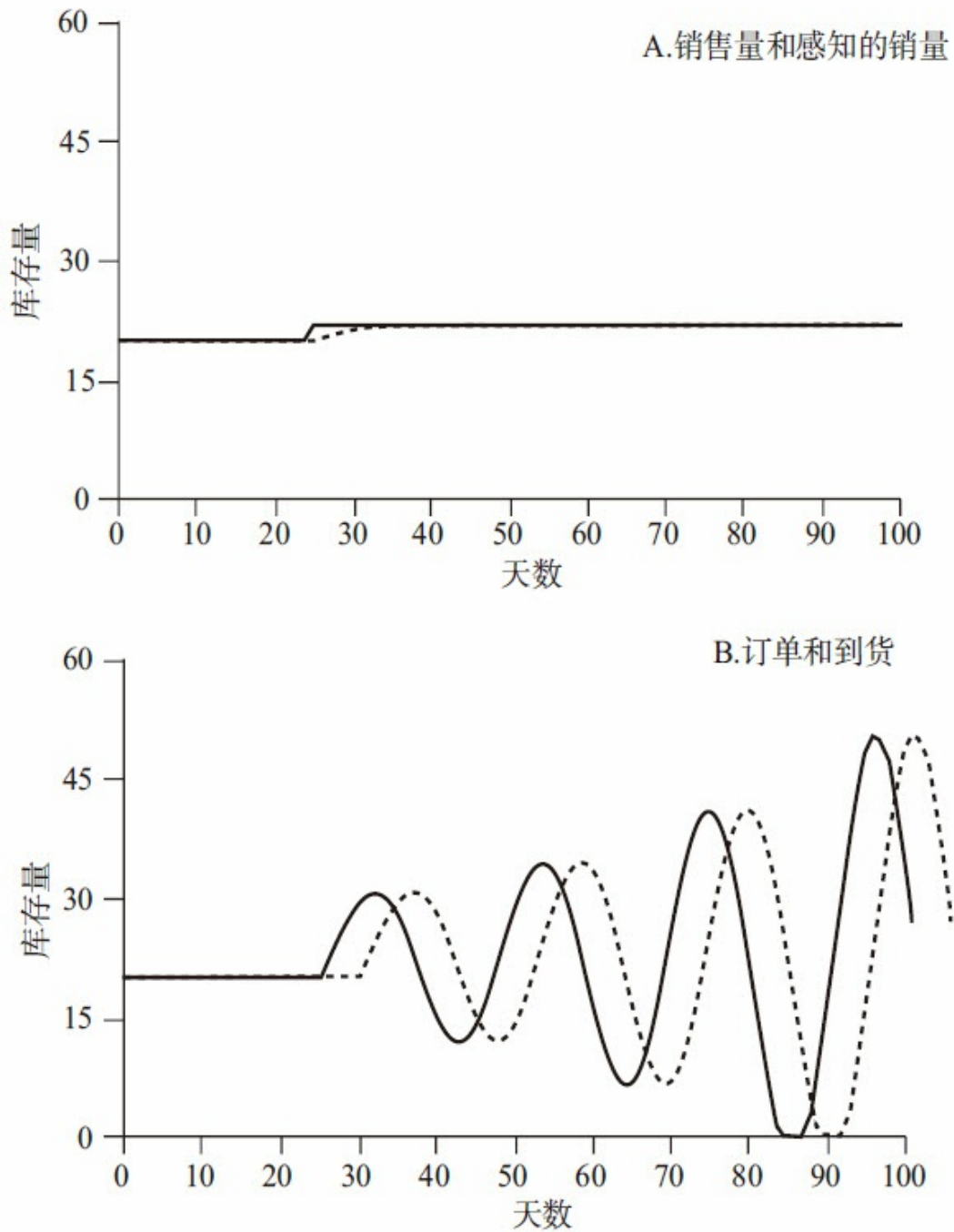


图2—19 订单和交付对需求增加的反应

A中的实线显示的是在第25天，销售量有一个很小但突然的阶梯式上升；虚线是经销商对前三天的销量进行平均，所感知到的销量的变化；B中实线和虚线分别是对应的订单和实际到货情况。

至于说一个延迟在什么情况下会引起怎样的振荡，这并不是一个简单的问题。我会用上述库存系统来向你解释具体的原因。

“这些振荡真是令人无法忍受，”经销商气愤地说，因为他们本身也是一个会学习的系统，所以决心改变一下库存系统的行为：

我要缩短时间延迟。虽然对于交货延迟，我左右不了，因为那主要取决于供应商，但我可以加快自己的反应。在确定订单数量时，我以前选取的是前5天的平均销量，现在我只选取前两天的平均数，这样就可以更快地响应销售量的变化了吧。

如果他们真的这样做了，图2—20就是调整之后库存量变化的示意图——从图上看，虽然供应商缩短了感知延迟时间，但基本上没有变化，库存量的振荡甚至有些恶化。

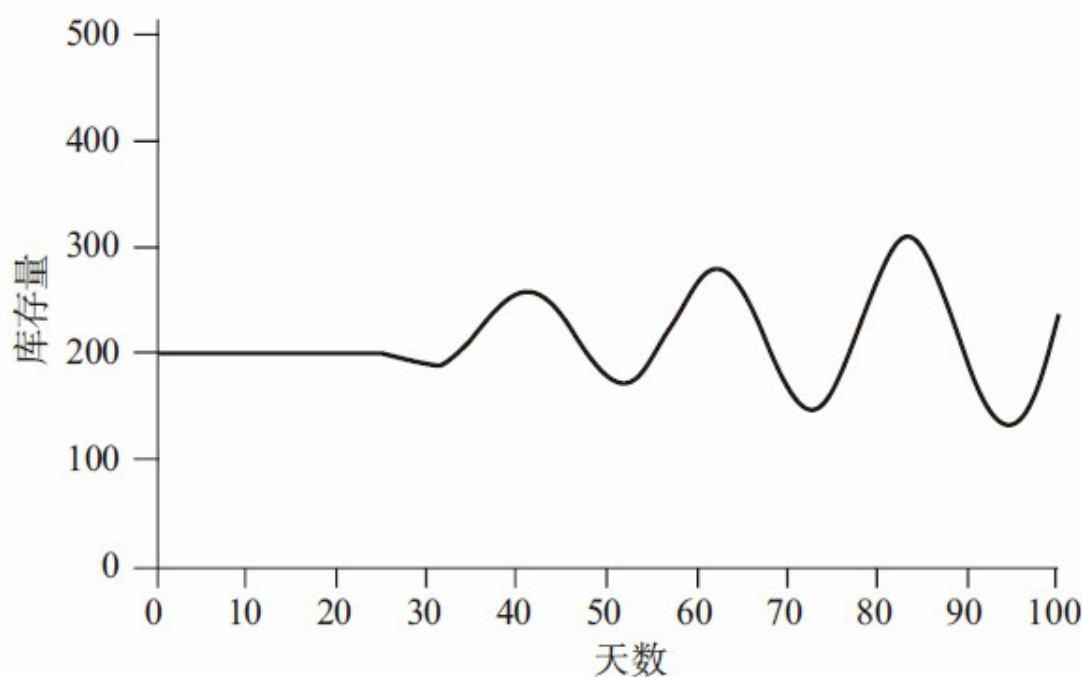


图2—20 库存量随时间变化情况（二）

即使缩短了感知延迟时间，库存量对于需求增加的反应模式依然如故。

如果这位经销商不是缩短感知延迟时间，而是缩短其反应时间，例如将感知到差距的时间从3天缩短到2天，情况会怎么样？

事实上，事情只能变得更糟，如图2—21所示。

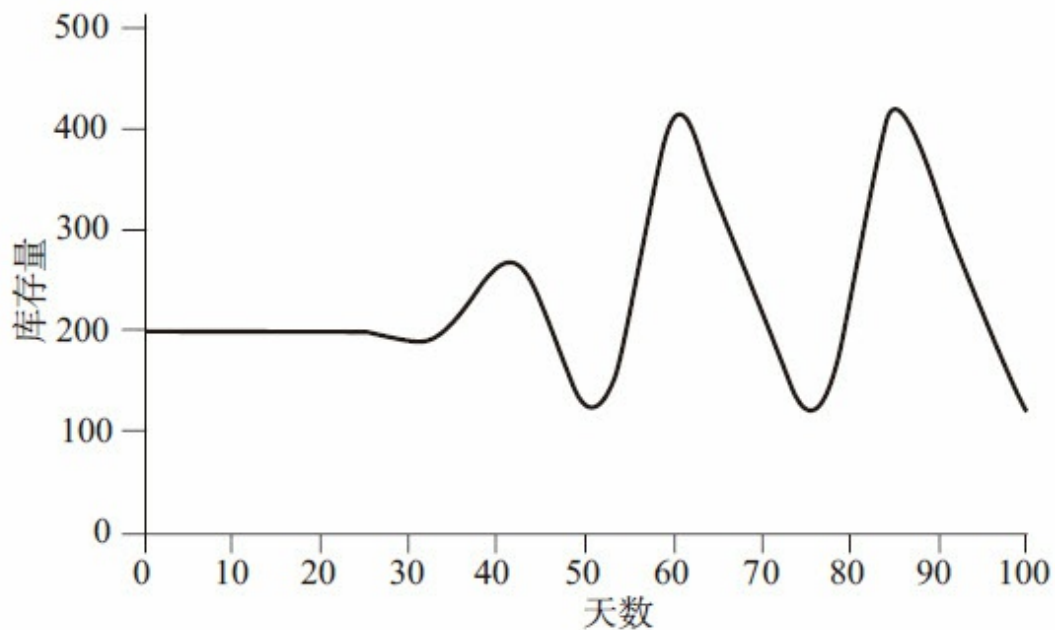


图2—21 库存量随时间变化情况（三）

如果缩短反应延迟时间，库存量的振荡将更加恶化。

作为一个善于学习、调整的人，经销商知道必须做些什么来改变这种状况。“高杠杆效应，但方向错误”，经销商这样对自己说。他已经注意到，销售量的微小变化会显著影响到订单和库存量的变化，而自己奉行的试图抑制振荡的政策是不奏效的。

系统提示

“变化”的可能后果

其实，这种“好心办坏事”或“越采取干预措施，问题越恶化”的情况很常见。人们通常出于好意，试图借助一些政策或干预措施来修补系统出现的问题，但结果往往事与愿违，甚至将系统推向错误的方向。同时，你的动作越大，对系统的影响就越强烈。当我们试图改变一个系统时，系统的行为往往违背我们的直觉，出乎我们的意料。

对于这个系统，造成问题的部分原因不是经销商的反应太慢了，而是太快了。在系统既定的状况下，经销商有些反应过度了。如果经销商不是将反应延迟从3天缩短到2天，而是延长到6天，事情就会好转许多，如图2—22所示。

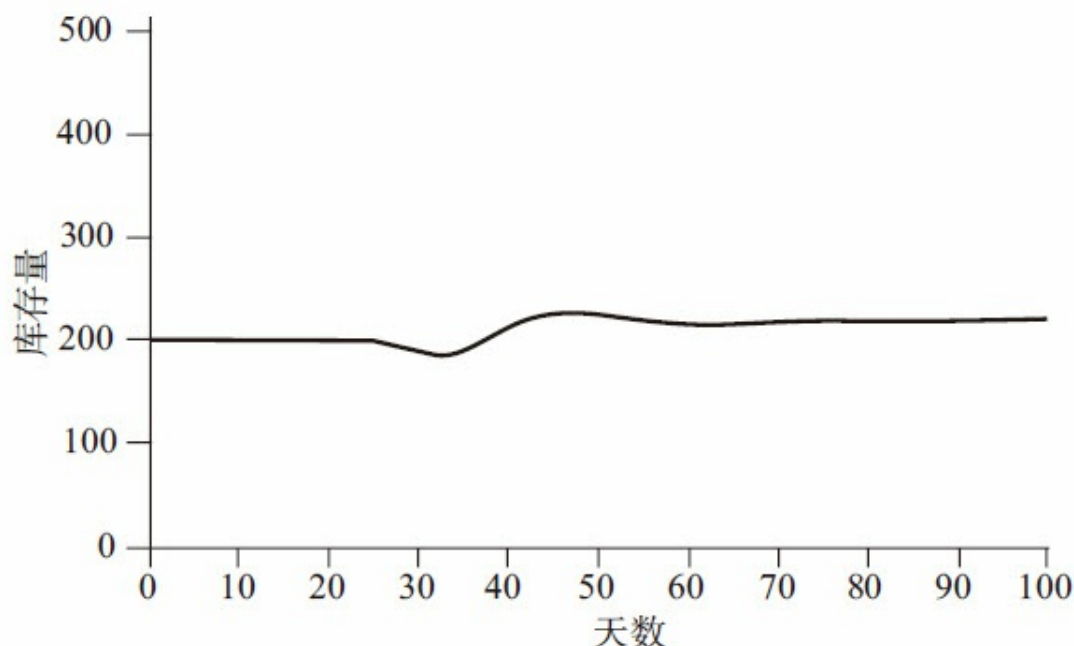


图2—22 库存量随时间变化情况（四）

在延长反应延迟时间的情况下，对于同样的需求变化，库存的反应很平缓。

如图所示，在同样销售变化的情况下，振荡明显地减缓了，而且系统会很快地找到新的均衡状态。

系统之思

在系统中，时间延迟是普遍存在的，而且它们对系统行为有很强影响。改变一个延迟的长短，可能会导致系统行为的很大变化，也可能不会，这取决于该延迟的类型以及与其他延迟相比的相对时间长短。

在该系统中，最重要的延迟是交货延迟，它不在经销商的直接控制之中。但是，即使没有改变这一延迟的能力，经销商也可以学会如何更好地管理库存。

改变系统中的延迟可能使系统更容易被管理，也可能完全相反。所以，有时候连一些系统思考专家对于时间延迟的问题也有些迷惑。为此，一看到系统中存在时间延迟，我们总是非常警觉，会认真分析它们存在于哪儿，时间多长，是信息流的延迟还是物理过程的延迟等。在某种意义上可以说，如果我们不知道延迟在哪儿、时间多长，我们就不可能真正理解系统的动态行为。从上文我们已经知道，一些延迟可能成为强有力的政策杠杆，延长或缩短它们可以使系统行为产生显著变化。

从一个更大的视角上看，一家经销商的库存问题可能是微不足道、可以被解决的，但想象一下，如果这是全美未售出的汽车库存，会怎么样？

- 订单的增减不仅会影响到整车组装厂和零部件供应商的产量，而且会影响到钢铁厂、橡胶厂、玻璃、纺织制品和发电厂等一系列相关企业。在这个庞杂的系统中，到处都有感知延迟、生产延迟、交货延迟和建设周期延迟。

- 让我们再考虑一下汽车生产和就业之间的关联。增加产量会提高就业人数，从而使更多的人可以去买车。这是一个增强回路，当然也可以往相反方向运转：产量减少，就业人数减少，购车数量降低……

- 还有另外一个投机性股票买卖的增强回路，即基于近期业绩，投资者买进或卖出汽车制造商和供应商的股票，产量的提升必将导致股价的上涨，反之亦然。

这是一个非常庞大的系统，不同工业部门之间相互联系，经由各种延迟相互影响，推动彼此的振荡，并被各种乘数效应和投机因素所放大，这是商业周期形成的主要原因。虽然总统或政治领袖会更容易抑制或强化经济回暖或衰退带来的乐观或悲观情绪，但这些周期并非源自总统。总之，经济是一个无比复杂的系统，充满了各种调节回路和时间延迟，它们本质上具有波动性。

双存量系统

系统2.1：一个可再生性存量受到另外一个不可再生性存量约束的系统

典型代表：石油经济

到现在为止，我们所讨论的系统都没有考虑外部因素的约束。例如，在工业经济系统中，资本存量未考虑原材料和产出的限制；人口系统中，未考虑食物的限制；在温度调节器系统中，也未考虑火炉可能缺油的限制。因为我们的目的是研究这些系统内在的动态性，所以在构建系统模型时进行了简化，没有考虑外部的约束条件。

但是，任何真实的实体系统都不是孤立存在的，其外部环境中都有各种相互关联的事物。

例如，一家公司离不开稳定的能源和原材料供应商，离不开员工、管理者和顾客；成长中的玉米离不开水分、养料，也少不了昆虫的袭扰；人口与食物、水和生存空间息息相关，也少不了就业、教育、健康医疗和很多其他因素。同时，任何使用能量和处理原材料的系统，都需要置放废弃物的场所，或者处置废弃物的过程。

因此，任何物理的、成长的系统，或早或晚都会受到某种形式的制约。这些限制因素通常以调节回路的形式存在，在某些条件下，这些调节回路会取代驱动成长的增强回路成为主导性回路，要么是提高流出量，要么是减少流入量，从而阻碍系统的进一步成长。

在现实环境中，受限制的成长是非常普遍的，以至于系统思考专家将其当成一种“基本模型”^[3]，命名为“成长上限”（limits-to-growth）。所谓“系统基模”，指的是一些常见的系统结构，可以导致人们熟悉的一些行为模式。在第5章中，我们将介绍更多的系统基模。事实上，以后每当我们看到一个成长的系统，不管是人口、一家公司，还是一个银行账户、一则谣言、一种流行病，或者新产品的销售，我们都可以找出驱动其增长的诸多增强回路，也必然能找到最终限制其增长的调节回路。即使尚未占据主导地位，似乎还看不到它们对系统行为的影响，但这些调节回路肯定存在，因为没有任何真实的物理系统可以永无止境地成长下去。

系统之思

在呈指数级成长的实体系统中，必然存在至少一个增强回路，正是它（或它们）驱动着系统的成长；同时，也必然存在至少一个调节回路，限制系统的成长，因为在有限的环境中，没有任何一个物理系统可以永远地增长下去。

例如，再热销的新产品，也总会有市场饱和的一天；核反应堆或原子弹中的链式裂变反应威力再强大，也终将耗尽核燃料；再凶猛的病毒，总有一天会无人可感染；再蓬勃发展的经济，也将受到实体资本或金融资本、劳动力、市场、管理、资源或污染等诸多条件的限制。

在我看来，资源与污染是一组相对的概念，正如资源是流入量，是供应源，有可再生性资源和不可再生性资源之分，污染的限制也有的是可修复的，有的是不可修复的。如果环境没有足够的能力吸收污染物，或使其无害化，污染就是不可修复的；反之，就是可修复的。当然，环境对污染物的消纳能力通常是有限且可变的。因此，在这里我们所说的受资源约束的系统，和受污染影响的系统具有相同的动态行为，只不过二者方向相反。

系统之思

限制性调节回路发源于可再生性资源或者不可再生性资源，是有所区别的。区别不在于成长能否永远持续，而在于成长以何种方式终止。

对成长的限制有可能是临时的，有可能是永久性的。有时候，系统可以找到其他途径，暂时或相对长期地“绕过”阻碍因素，再次实现成长。但最终肯定会产生某种类型的调适，要么是系统适应了限制因素，

要么是限制因素适应了系统，或者彼此适当调整。当这些调适出现时，可能会产生一些有趣的系统动态。

现在让我们看一个实例。假如有一家公司，通过提炼一种不可再生性资源（比如石油）来赚钱。它们刚刚发现了一个巨大的新油田，如图2—23所示。

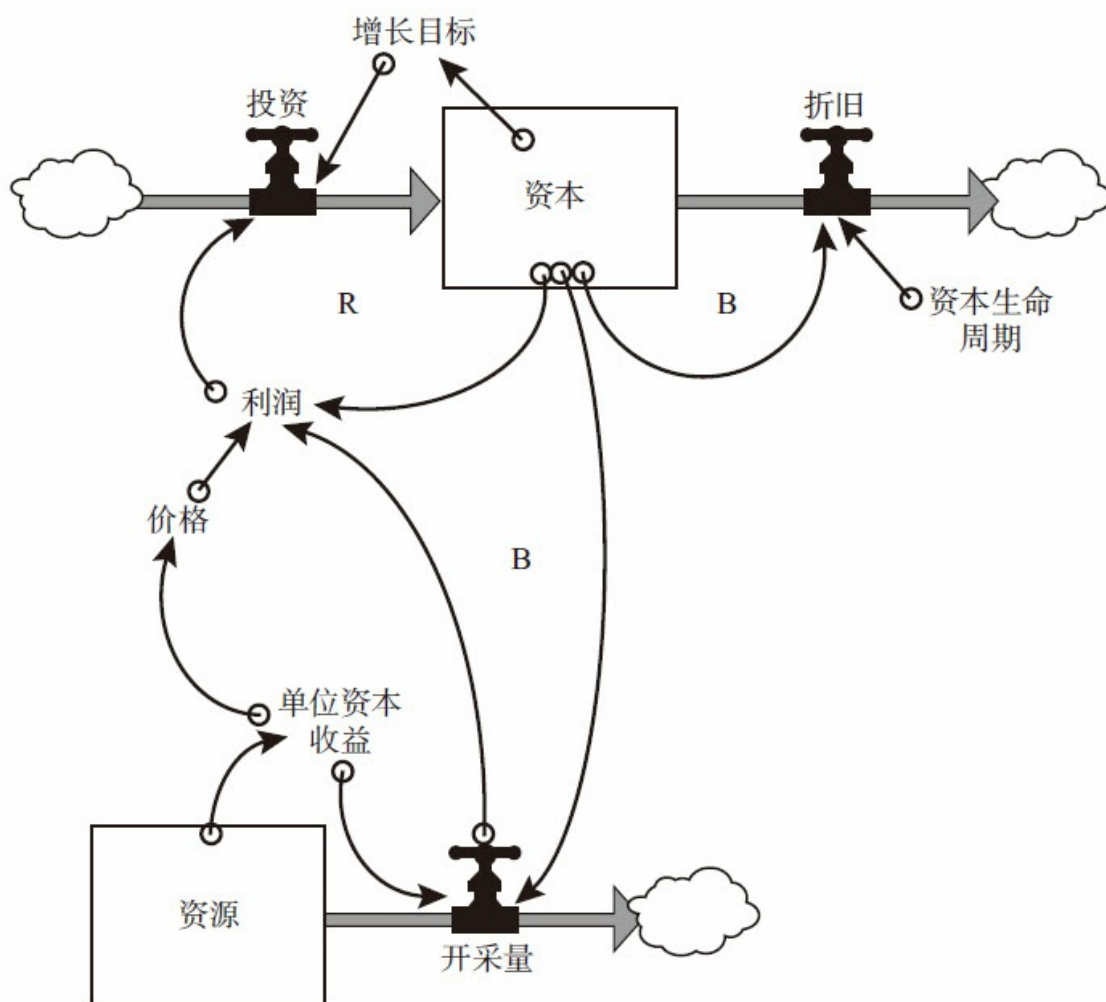


图2—23 有一个增强回路且受到不可再生资源限制的经济资本系统

图2—23看起来可能比较复杂，但它从本质上还是我们上面已经看到过的资本—增长系统，只不过用“利润”代替了“产出”。

驱动折旧的调节回路大家已经熟悉了：资本存量越多，机器设备的磨损和消耗越大，从而减少资本存量。在本例中，资本存量是开采和提炼石油的设备，按20年来提取折旧，这意味着每年计划提取5%的减值

损失费用。为了保持并扩大产能，石油公司将获得的部分利润用于再投资，来提高资本的存量。所以，这里有一个增强回路：更多的资本存量可获得更多的资源开采量，创造更多的利润用于再投资。在这里，我们假设公司每年预计有5%的资本增长。如果当年利润不能支持5%的增长，则将全部利润均用于再投资。

众所周知，利润等于收入减去成本。在本例中，我们将收入简化为石油价格乘以石油产量；成本等于资本总额乘以单位资本的运营成本（包括能源、劳动力、原材料等）。同时，为了简便起见，我们将价格和单位资本的运营成本都假定为常数。

在这里，我们未将单位资本的资源收益假定为常数，这是因为这些资源（石油）是不可再生的，随着石油开采量的增加，一口口油井终将枯竭，下一桶石油的获得将比上一桶更加困难。对于石油存量来说，开采量是流出量，但是没有流入量；剩余的资源要么埋藏得更深，要么浓度更低，或者开采难度更大。人们不得不更多地使用一些更加昂贵和技术复杂的措施去获取该资源。

这是一个新的调节回路，并最终会限制资本的增长：更多的资本导致更快的开采速度，从而更快地降低资源储量。资源储量越少，单位资本的资源收益就越低，利润就越少（假设价格是固定的），再投资比率就越小，资本的增长速度也将降低。我们假设资源枯竭可以通过运营成本和资本效率反映出来。在现实商业世界里，这两方面都是真实存在的：一些油田要么因运营成本过高，要么因资本效率过低而被废弃。无论哪种情况出现，后续的行为模式都是一致的。如图2—24所示，这是一个典型的衰竭行为。

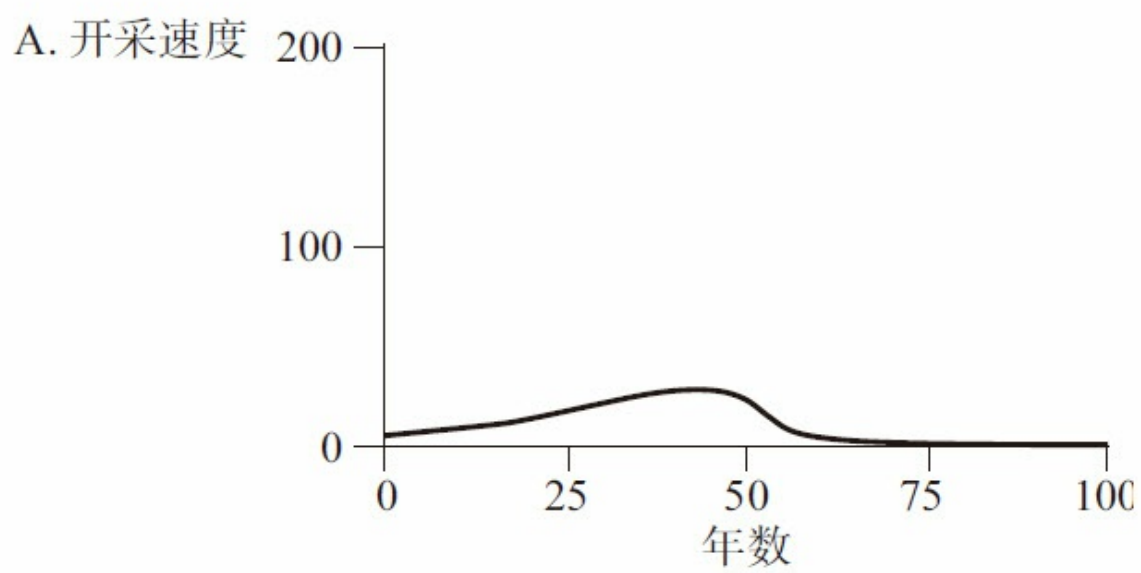


图2—24 资本积累的速度越快，资源枯竭得越快

开采（A）会带来利润，导致资本（B）的增长，并将逐渐耗尽不可再生性资源（C）。

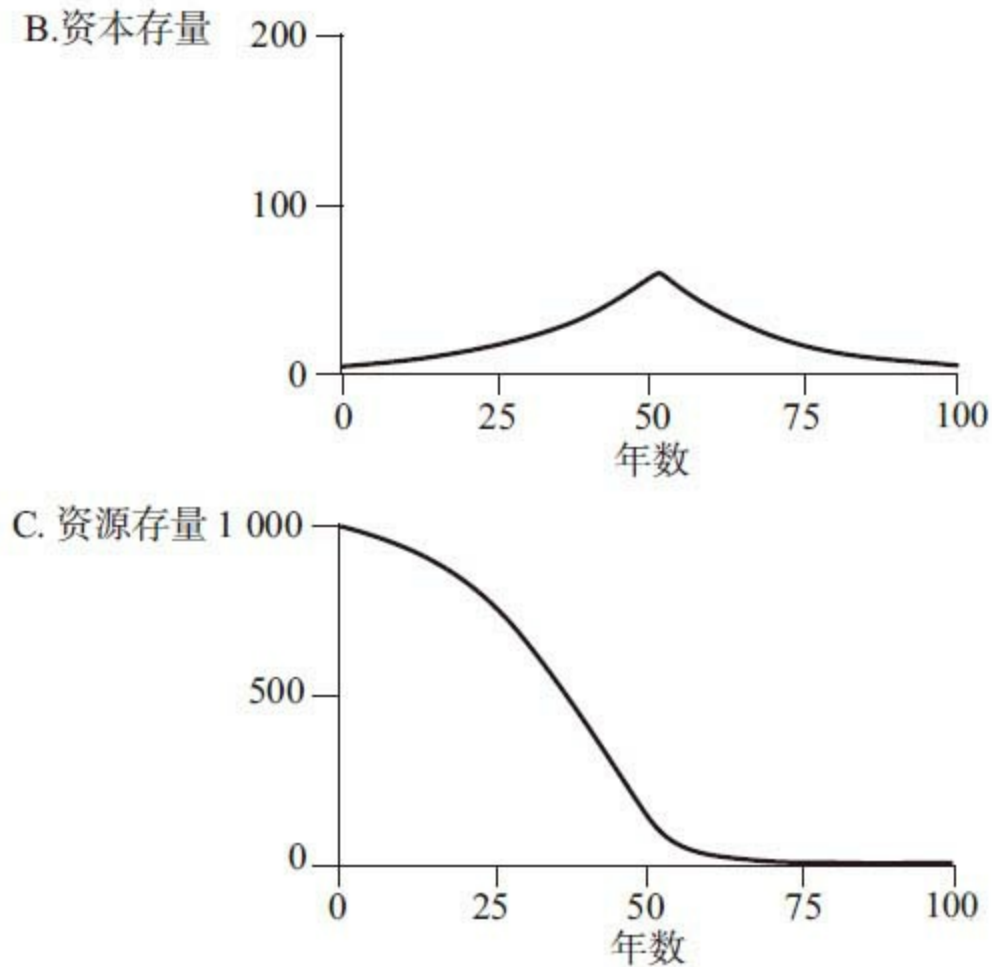


图2—24 资本积累的速度越快，资源枯竭得越快（续）

开采（A）会带来利润，导致资本（B）的增长，并将逐渐耗尽不可再生性资源（C）。

一开始，地下石油储量很充足，按最初的设计规模可开采200年。但是，由于增长的指数效应影响，实际开采量在第40年左右的时候达到了顶峰，之后迅速衰竭。按照每年10%的投资系数，资本存量和开采速度均每年增长5%，因此到第14年翻了一番。28年后，资本存量变成了原来的4倍，开采量却受单位资本产出率下降的拖累而未能同步增长。到第50年，资本存量的维持成本已经超过了从资源开采中获得的收入，因此利润不足以维持投资的增长，投资小于折旧。很快地，随着资本存量的衰减，运营也停止了。剩余的最昂贵的资源被长留在地下，因为从经济和商业的角度上看，把它们开采出来已经没有价值了（如图2—25所示）。

如果资源的实际储量被证明是最初估计的2倍或4倍，情况会怎么

样？当然，石油开采总量会有很大差异，但在每年再投资率为10%、资本增长率保持5%的情况下，资源每翻一番都会使开采高峰期延长14年左右，与石油开采相关的产业、社区、就业等的繁荣期也相应地延长了。

系统之思

当一个变量以指数级形式逼近一项约束或限制时，其接近限制的时间会出乎意料地短。

如果你构建的资本存量依赖于—项不可再生性资源，那么它增长得越高、越快，下跌得就可能越深、越快。在资源开采或应用以指数级增长时，若不可再生性资源储量翻番或扩大4倍，只能稍微延长一些开发的时间而已。

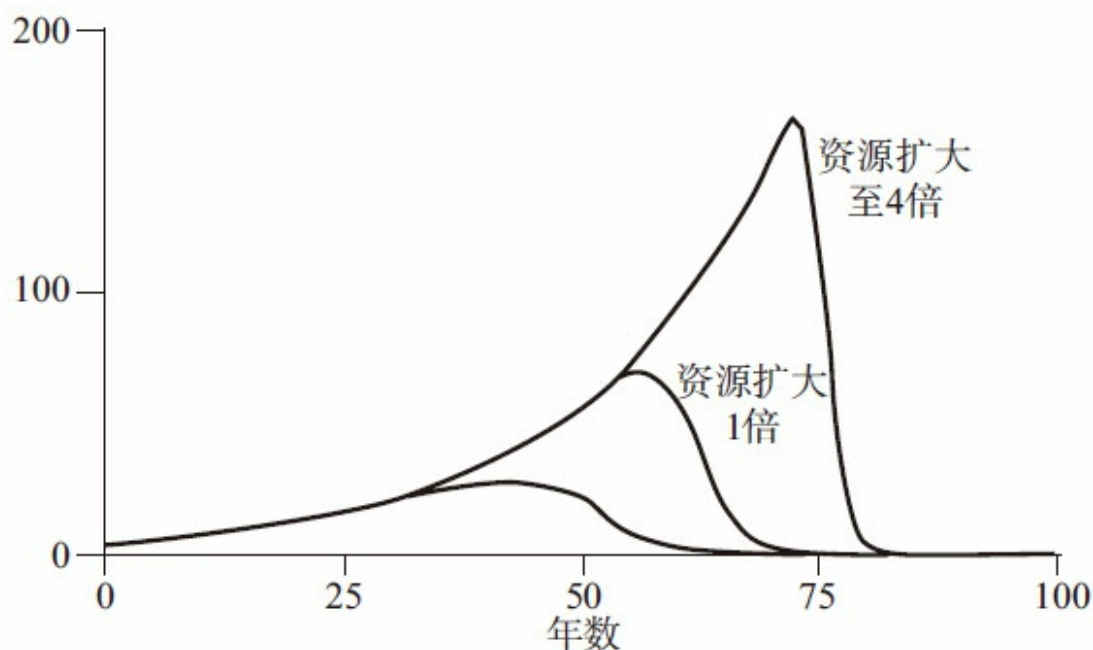


图2—25 资源储量每翻一番都只能使开采高峰期延长14年左右

如果你只关心最大限度地开采资源赚钱，那么该系统中最重要数字就是资源的最终规模。比如，你是一名油田或矿井上的工人，你关心自己能工作多少年以及社区是否稳定，那么你最需要关注的两组数字是：资源的规模以及预期的资本增长率。我们知道，反馈回路的目标对于系统行为是很关键的，这就是一个很好的例子。对于不可再生性资源的管理，真正需要做出抉择的是快速致富、“过把瘾就死”，还是无须那么富有，但可以持续的时间更长。

图2—26显示的是，假设扣除折旧后预期的年资本增长速度分别为1%、3%、5%和7%时，开采速度的变化情况。当增长速度为7%时，设计开发周期为200年的油田，将在40年内达到开发高峰。别忘了，你的决策不只会影响到公司的利润，也会对社会和地区自然环境造成影响。

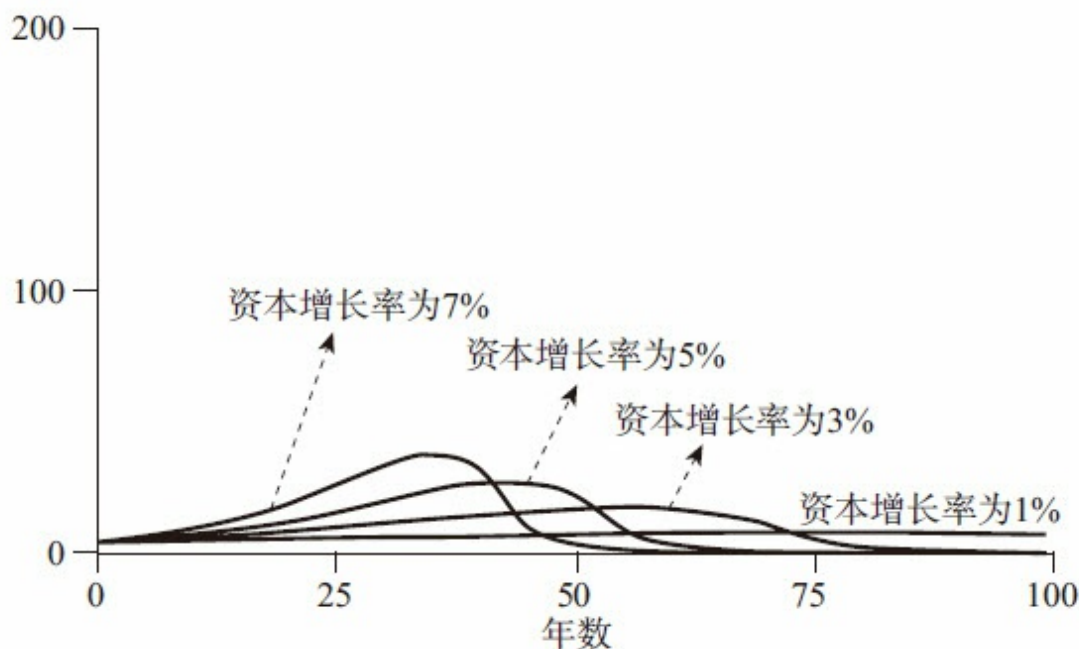


图2—26 资本增长率与开采高峰的关系

再投资比例越高，开采高峰期到来得越快。

之前我们曾讲过，为了简化，我们假设价格是常数。但是，如果价格是变动的，情况会怎么样呢？假设短期内资源对消费者很重要，价格过高将减少需求。在这种情况下，随着资源日渐稀缺，价格会快速升高，如图2—27所示。

价格越高，公司的利润就越高，所以投资增加，资本存量持续上升，而且更多的剩余资源可能被开采出来。把图2—27和图2—24做一下对比，可以发现，价格上涨的主要结果是让你积累更多的资本存量，但不能避免最终的崩溃。

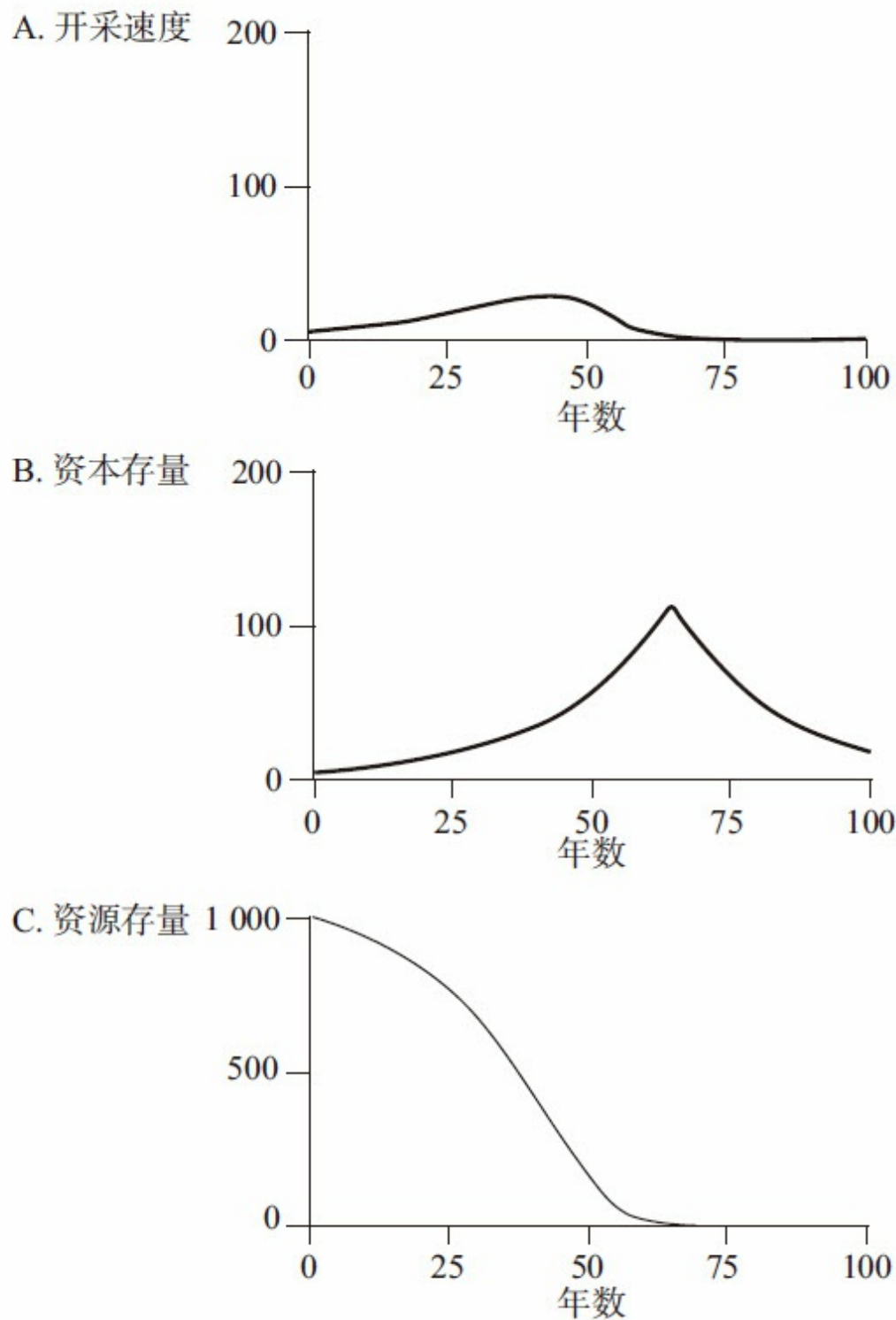


图2—27 价格上涨最终导致资源枯竭

价格上涨加大了资源的稀缺性，有更多的利润用于再投资，使得资本存量（B）变得更大，并使开采时间（A）更长，结果是资源存量（C）被快速耗尽。

顺便提一下，如果价格不上涨，而是技术进步导致运营成本降低，也会导致相同的结果。例如，人们发明了一种先进的油井回收技术，或者从接近枯竭的铁矿中提取低等级铁矿石的精选过程，或者一种氰化物萃取方法，可以从金或银的尾矿中提取贵金属等，就可能出现类似情况。

我们都知道，单个的矿山、化石燃料储藏和地下水都可能会枯竭。世界上有很多废弃的矿井、油田，都可以验证我们上面所探讨的情况。那些资源或能源企业也知道这些系统动态。所以，当它们在一个地方正常运营但发现资本效率开始下降时，就会转到其他地方投资，以发现和开发另外一处资源储藏。但是，如果存在地域限制，这些公司最终还会展开全球性运作吗？这个问题就留给各位读者去判别吧。

系统提示

依靠不可再生性资源的经济体系

根据资源耗尽的动态行为特性，初级资源存量越大，新的发现越多，驱动成长的增强回路相对于限制性的调节回路的影响力就越强，导致资本的存量越高，开发速率越快；然而，一旦生产高峰过去，经济衰退也开始得更早，速度更快，而且幅度越大。

也许，我们应该尝试建立一个完全依靠可再生性资源的经济体系了。

系统2.2：有两个可再生性存量的系统

典型代表：渔业经济

假设在上文所述的资本系统基础上，我们给资源存量增加一个流入量，使其成为可再生性资源。这样，我们就能得到我们即将开始探讨的渔业经济系统。在这里，可再生性资源是鱼，资本存量是渔船。其实，类似这样的系统还有很多，例如树木与伐木场、牧场与奶牛等。

- 有生命的可再生性资源，如鱼、树木、草等，可以经由一个增强回路实现自我再生。
- 无生命的可再生性资源，如阳光、风、河水等，虽然没有可再生的增强回路，但是不管存量的当前状态如何，它们都有稳定的补给输入来源。
- 感冒病毒感染、日用品销售等，也与“可再生性资源系统”具有相同结构。对于病毒感染来说，易感人群是可再生性资源存量；对于日用品销售，潜在消费者也是可再生的存量。就像虫害只能部分地损害农作物，而不能完全摧毁农作物一样。作物可以再生，而昆虫可以吃得更多。

在所有这些案例中，对于约束性资源存量，都有一个流入量，保持资源存量的补给（如图2—28所示）。

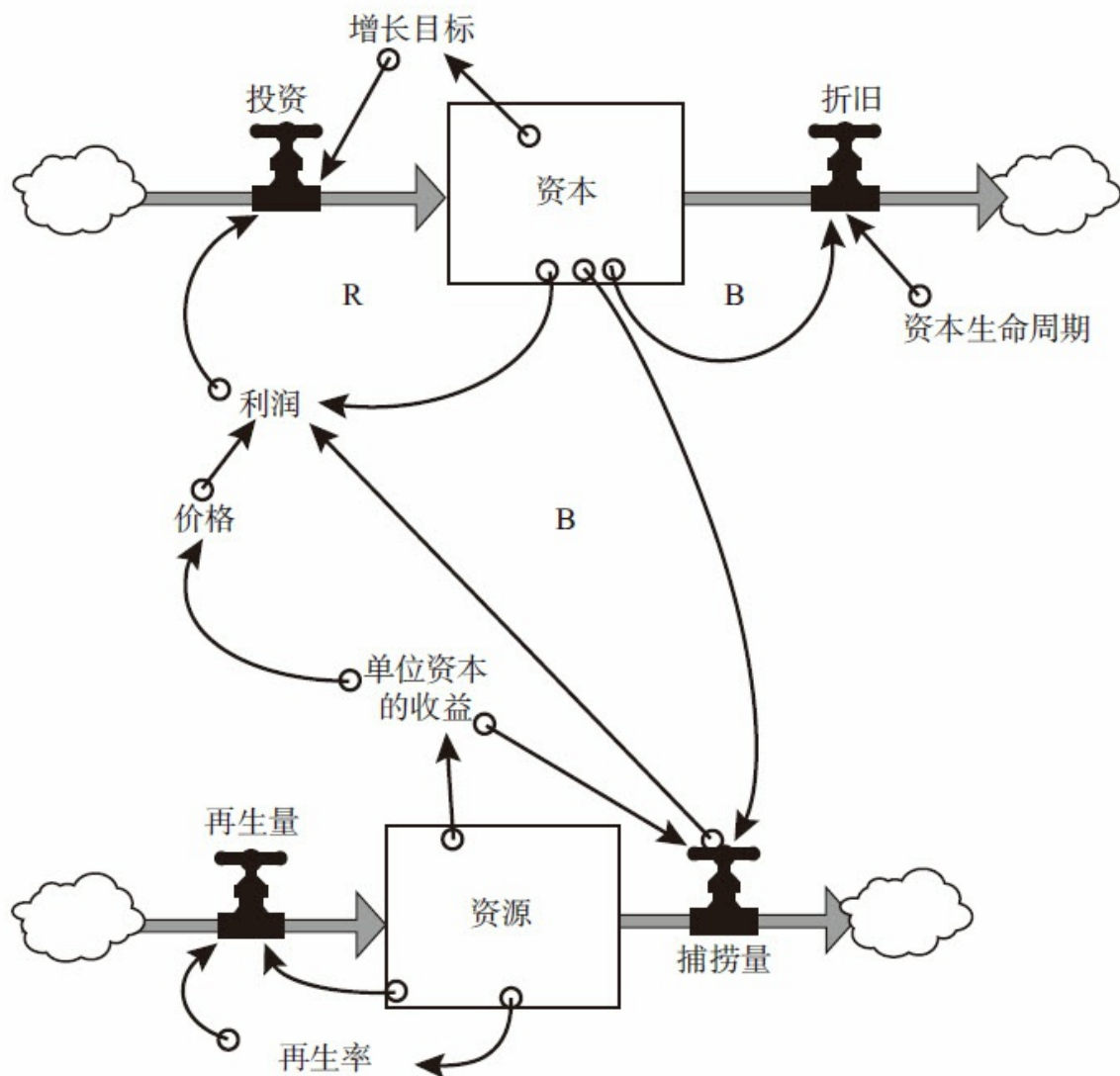


图2—28 经济资本受一个增强回路推动而成长，并受一个可再生性资源约束的情况

我们以渔业为例。首先，假设资本的生命周期为20年，每年产业增长率维持在5%。其次，和不可再生性资源一样，我们假设渔业资源由于日益稀缺而价格上涨，这将带动开发力度加大，并需要增加资本投入。对于渔业公司来说，吨位更大的渔船可以在海上航行更远的距离；而安装声纳设备，可以及时追踪鱼群的动向；装备数海里长的拖网或者船载冷冻系统，可以捕到更多的鱼，并将其从遥远的海上带回港口。所有这些都需要花费资本。

鱼的再生率不是固定的常数，而是依赖于同一海域中鱼的数量，也就是说鱼群的分布密度。如果鱼群密度很大，由于受到食物和栖息地的限制，再生率会接近于零；而随着密度的降低，鱼群再生速度会加快，因为有更多的食物或更大的空间。但是，到了某个点时，鱼群再生率会达到最高峰。超过这个点，如果鱼群数量继续减少，鱼群繁殖速度不是越来越快，而是越来越慢。这是因为每条鱼都很难再找到同类，或者是另外一个物种侵入了它们的领地。

这一简化的渔业经济模型受到三种非线性关系的影响：价格、再生率、单位资本的收益。价格取决于鱼的种类和数量——越稀少的鱼，价格越高；再生率取决于鱼群的密度——鱼群密度越低，鱼的繁殖率越低，但是鱼群密度变得过大之后，鱼的繁殖率也越低；单位资本的收益取决于捕鱼技术和方法的效率。

这一系统可以产生很多种不同的行为模式，图2—29是其中之一。

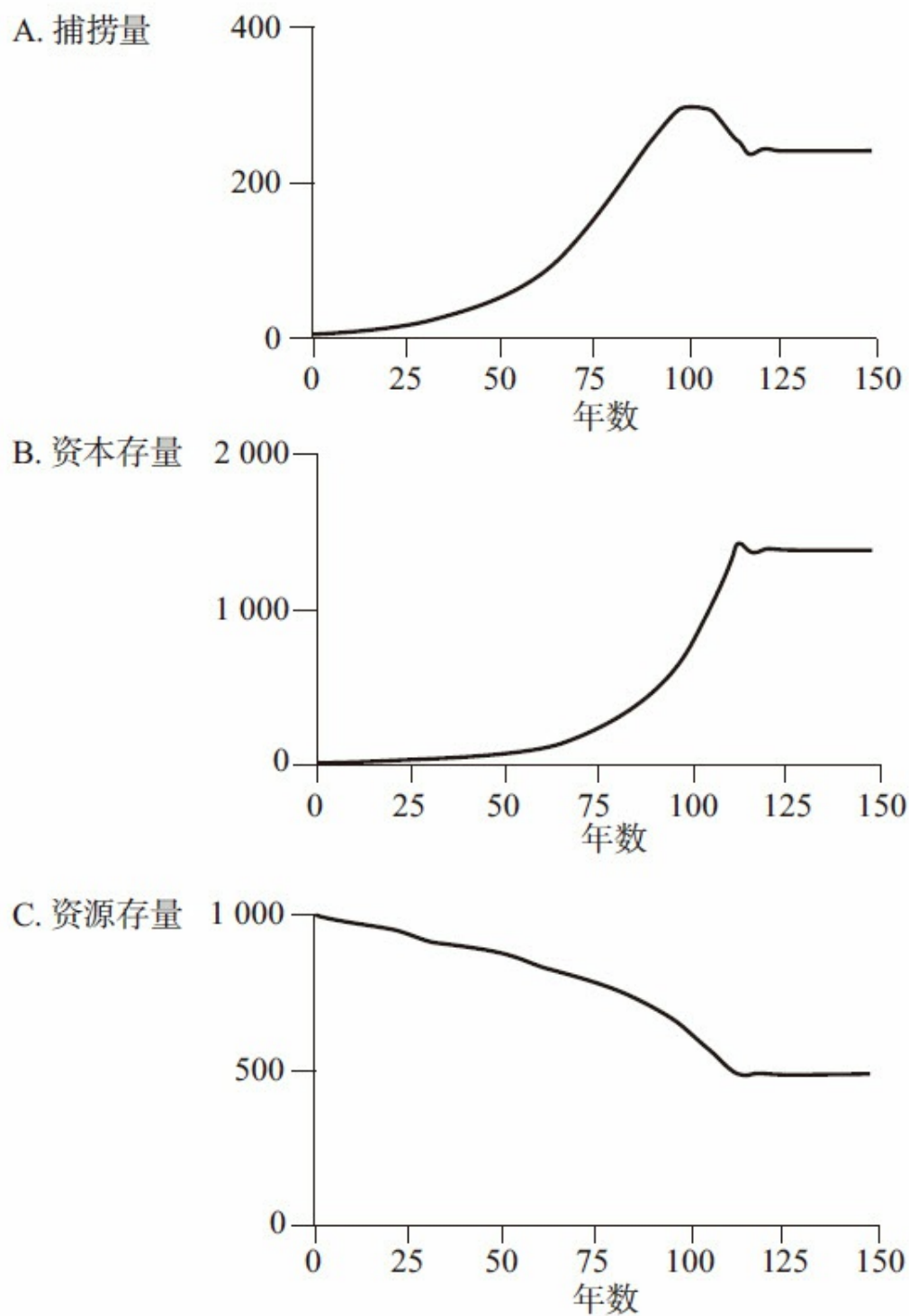


图2—29 捕捞量、资本存量及资源存量的关系

每年捕捞量 (A) 产生的利润，推动资本存量 (B) 的增长，在稍微过量增长之后，捕捞量趋于稳定，也导致资源存量 (C) 保持稳定。

从图2—29中可见，一开始资本和捕捞数量以指数级方式上升，鱼

群数量（资源存量）则快速下降，但这提高了鱼群的再生率。在其后的数十年中，资源再生数量可以弥补、应付捕捞增加的速度；但最后，由于捕捞量增长太快，鱼群数量下降到不能满足日益扩大的捕捞量需求，捕捞量降低，从而降低了船队的利润率和投资速度。这形成了一个调节回路，使得船队和鱼资源之间达到了平衡。船队达到一定规模后，不能永远扩大下去，但是可以一直保持较高且稳定的效率。

然而，只要对单位资本产出做很小的调整，改变其控制的调节回路的力量，就能产生显著的差异。假设人们发明了一种改进船只效率的技术，通过更好的声纳设备，可以发现更稀少的鱼群。这样，即使鱼群减少了，但每一艘船获得同样捕捞量的能力都稍微提高了一些（如图2—30所示）。

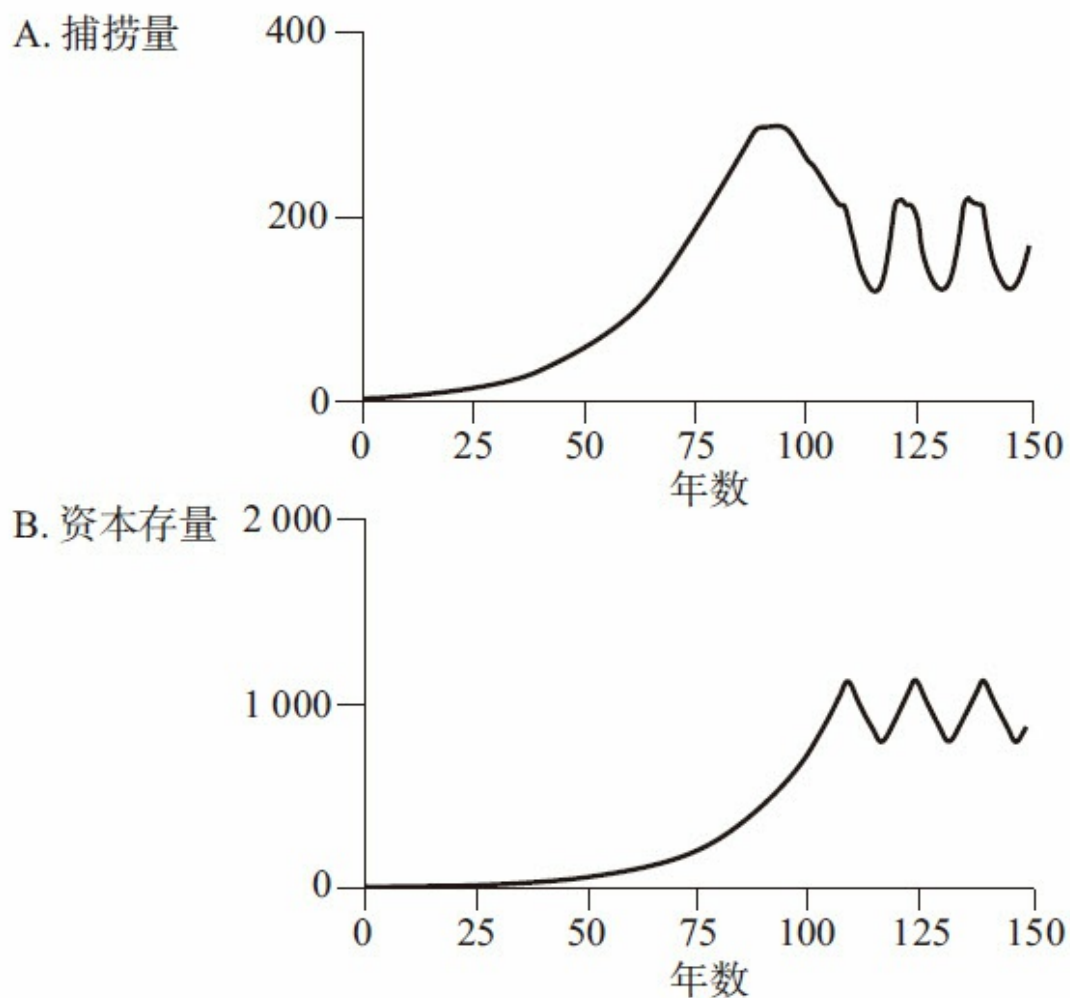


图2—30 技术改进带来系统振荡

单位资本的收益稍微升高一些，例如提高技术的效率，会造成围绕捕捞量（A）、资本存量（B）和资源存量（C）的稳定状态的过度开发和振荡行为。

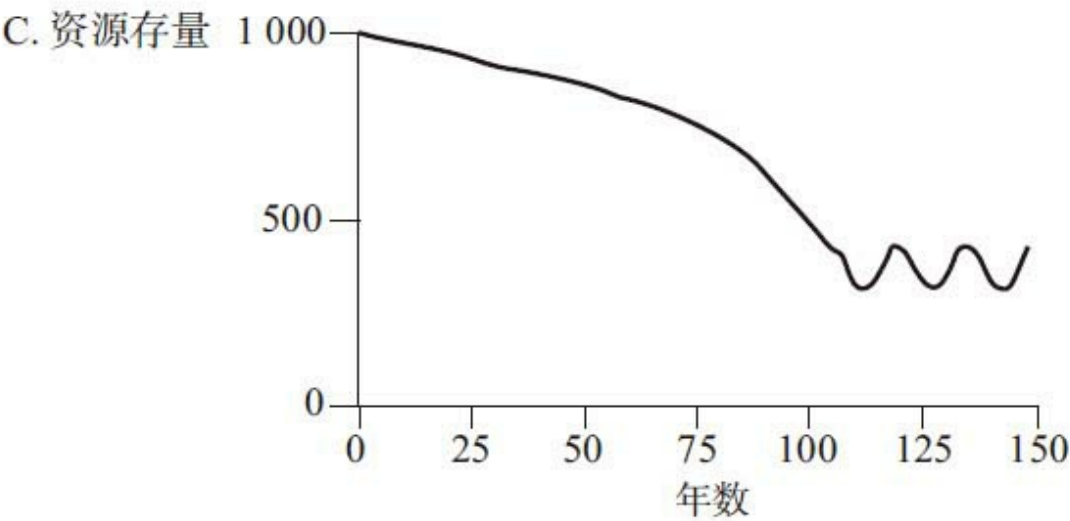


图2—30 技术改进带来系统振荡（续）

单位资本的收益稍微升高一些，例如提高技术的效率，会造成围绕捕捞量（A）、资本存量（B）和资源存量（C）的稳定状态的过度开发和振荡行为。

我们之前曾见过“高杠杆效应，但方向错误”的状况，图2—30又是另外一个案例。技术改进，增加了所有渔民的生产率，却导致系统陷入不稳定之中，振荡再次出现。

系统之思

不可再生性资源主要受限于存量。所有存量一次到位，然后被逐渐开发使用（流出量）。之所以不能一次性开发，主要是因为资本（和开发条件）的限制。由于存量是不可再生的，开采速度越快，资源的生命周期就越短。

如果捕鱼技术得到提高，船只可以在鱼群密度很低的情况下更为经济地维持运作，但结果只能是鱼和捕鱼业接近彻底消亡。海洋变成荒漠，鱼也逐渐变成了不可再生性资源。图2—31显示的就是这一情景。

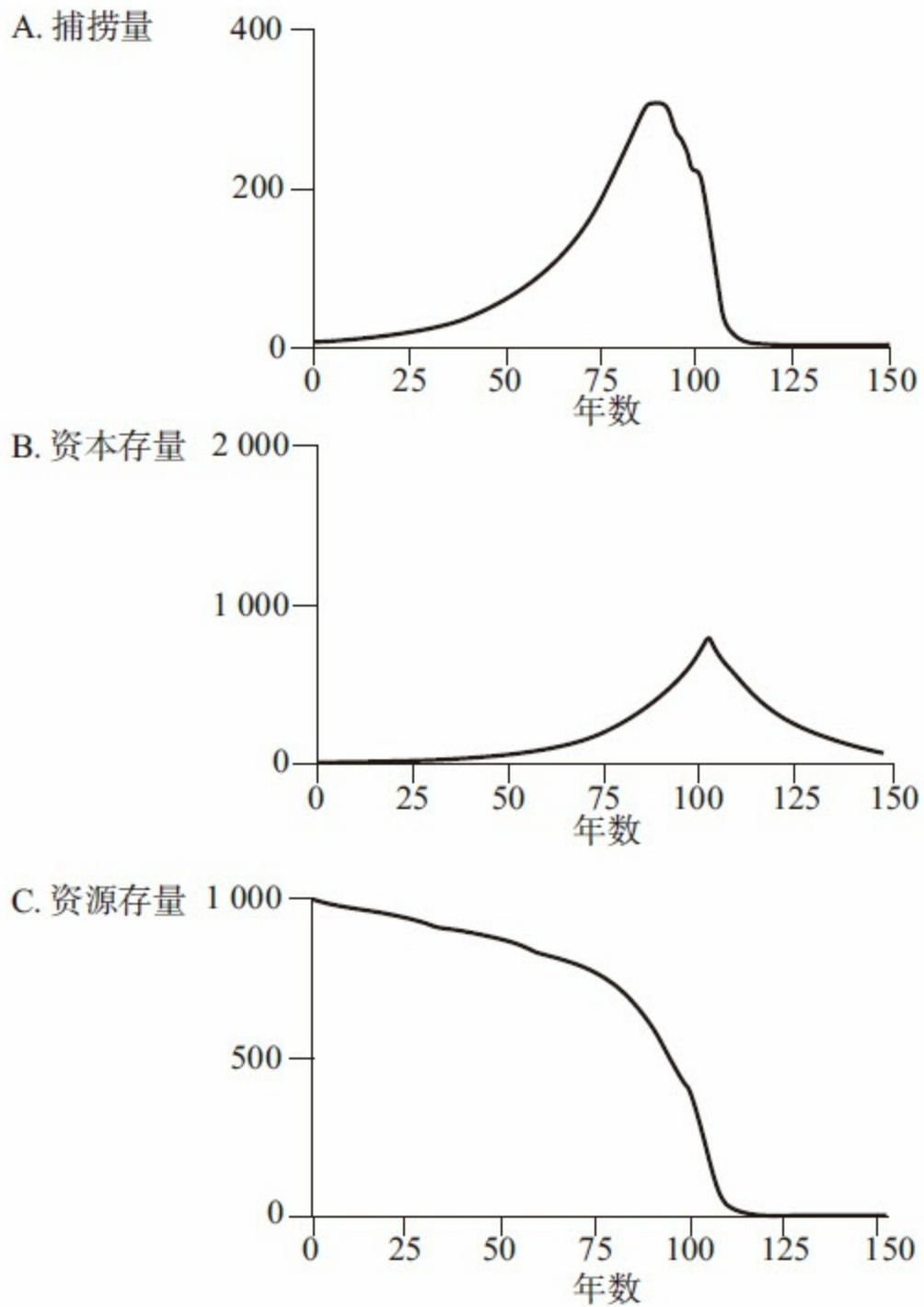


图2—31 技术进步最终导致崩溃

单位资本产出上升得越高，越会造成捕捞量（A）、经济资本（B）和资源存量（C）的过度开发和崩溃的行为模式。

与我们这里所讨论的简化模型相比，在很多真实的开发可再生性资

源的经济系统中，一旦资本和开发活动的增强回路得以启动，资源通常就会被加速攫取、耗竭，等开发者退出之后，只剩下极少量的“活口”，苟延残喘，或者侥幸得以休养生息而恢复。数十年之后，同样的故事又会再次上演。有些可再生性资源的循环周期很长，例如新英格兰的伐木工业，从成长、过度砍伐、崩溃，到资源的逐渐再生，然后再次砍伐.....现在已是第三个循环了。但是，并不是所有的可再生性资源都是如此。由于科技和开发效率的提升，越来越多的资源趋于枯竭，无法再生。

系统之思

可再生性资源主要受限于流量。只要开发（流出量）的速度等于资源再生（流入量）的速度，它们就可以被无限地开采或捕捞；如果开发的速度快于再生的速度，资源存量最终可能低于某个关键转折点，从而转变为“不可再生性资源”，逐渐耗尽。

到底可再生性资源能否在过度开发之后“劫后余生”，取决于在资源严重衰竭的那段时间里发生了什么情况。例如，当鱼群数量变得很少时，它们就会非常脆弱，一次污染、一场龙卷风，或者缺乏基因的多样性，都可能使鱼群彻底灭绝。再如，对于森林和草场资源，裸露的土壤也很容易被风化或流失。生态环境中其他竞争者也有可能“乘虚而入”，侵入、占领这一小生境。当然，在条件合适时，几近枯竭的资源也可以存活下来，并再次繁衍、生息。

在这里，我们可以看到可再生性资源系统的行为模式有三种可能性：

- 过度开发，然后逐渐适应、调整至相对稳定的平衡状态，并长期保持；
- 过度开发，超出了均衡状态，之后上下振荡；
- 过度开发，之后导致资源的枯竭，产业崩溃。

系统提示

两个影响因素

实际会出现哪种结果，取决于两方面：第一，关键转折点是否被突破。一旦关键转折点被突破，资源的种群数量实现再生的能力就会被破坏；第二，在资源逐渐衰减的过程中，抑制投资增长的调节回路的力度。如果该调节回路可以在关键转折点到来之前，快速起作用，控制资本的增长，那么整个系统就能平滑地达到均衡状态；如果该回路速度比较慢，不够有效，系统就会振荡；如果该回路非常弱，或者起作用的速度很慢，这样一来，即使资源已经降低到难以再生的水平，但资本仍在持续增长，最终的结果是，该资源和产业都将崩溃。

在成长上限结构中，不管是可再生性资源，还是不可再生性资源，

物质的存量都不可能永远增长，但是二者对于系统的限制，从系统行为的动态角度上讲是非常不同的。之所以会有差异，原因在于存量和流量的不同。

对于所有复杂的系统来说，判断系统未来行为走势的诀窍在于，了解什么样的系统结构包含哪些可能的行为，以及什么状况或条件可以触发这些行为。换句话说，如有可能，我们可以调整系统结构和相关条件，从而减少破坏性行为发生的概率，增加有利行为出现的概率。

[1] 在特定出生率的情况下，人口越多，当期新出生的人数也就越多，从而进一步增加了人口数量，这是一个增强回路。——译者注

[2] 在特定死亡率的情况下，人口越多，当期死亡的人数也就越多，从而导致人口总数的减少，这是一个调节回路。——译者注

[3] 也可称为“系统基模”，简称“基模”，下同。——译者注

第二部分 系统思考与我们

第3章 系统之美：系统的3大特征

奥尔多·利奥波德

生态学家

如果土壤有机质作为一个整体是良好的，那么每个部分都是好的，不管你是否理解；如果生物圈在演进的过程中，已经建立了我们喜欢但不了解的东西，那么只有傻瓜才会丢弃那些看似无用的部件。一个精明的钳工，首先想到的是确保每一个齿轮和车轴都完好无损。

我们在第2章中介绍了几种简单的系统，它们有着不同的结构，因而呈现出不同的行为模式。其中有一些系统是非常优雅的，即使受到各种限制，依然在这个充满斗争的世界中顽强地生存着，保持着它们的从容与淡定，执著地从事着它们的工作，包括保持室温的稳定、开采油田，或者在渔业资源和船队规模之间保持平衡等。

如果受到太大的冲击，系统可能会四分五裂，或表现出我们未曾见过的行为。但是，在很大程度上，它们都应对得很好。这就是系统之美：它们运作得如此精妙，各种机能和谐运行。想象一下，一个社区在应对风暴时的景象，人们争分夺秒地救助受难者，各种智慧和技能都涌现出来，如同一个上足了发条、高速运转的机器。等灾情过后，一切又恢复如常。

为什么系统会运作得如此精妙？请选定一个你所熟悉的高效运作的系统，比如一台机器、一个社区或者生态系统，并认真观察。幸运的话，你可能会看到以下三个特征中的一个或几个：适应力、自组织和层次性。

适应力

生态学家 霍林

如果系统被一大堆常数所禁锢，它就很难成长和进化。

“适应力”（Resilience）在工程学、生态学或系统科学领域有很多种

定义。基于人们的目的不同，常见的字面意思也有差异。

- 如果形容一个物体，适应力指的是其在被按压或拉伸之后，能够恢复到原有形状、位置的能力；
- 如果形容一个人，适应力指的是他快速恢复的能力，包括力量、精神、幽默感或其他方面；
- 如果形容一个系统，适应力指的是系统在多变的环境中保持自身的存在和运作的能力。与适应力相对的是脆弱性或刚性。

系统之思

系统之所以会有适应力，是因为系统内部结构存在很多相互影响的反馈回路，正是这些回路相互支撑，即使在系统遭受巨大的扰动时，仍然能够以多种不同的方式使系统恢复至原有状态。

我们知道，一个单一的调节回路会驱动系统存量到达预定的状态，对于存在多个类似回路的系统来说，就会显现出适应力，因为这些回路以不同的传导机制起作用，有不同的时间周期，也存在一定冗余——如果其中一种机制失效了，另外一种就可以补位。

如果有一组反馈回路，可以修复或重建反馈回路，系统的适应力就比较强，也可称为“元适应力”（**meta-resilience**）。由具有更高适应力的反馈回路组成的“元元适应力”（**metameta-resilience**），往往具有更加复杂的系统结构，有更强的复原能力，可以学习、创造、设计和进化。这类系统具有很强的自组织性，也是系统的基本特征之一。

系统之思

适应力总是有限度的。有适应力的系统可能是经常动态变化的。相反，一直保持恒定的系统恰恰是不具备适应力的。

人体就是一个令人称奇的、具有很强适应力的系统。它可以抵御成千上万种病毒、细菌等有害物质的入侵，可以适应各种不同的温度以及差异很大的食物，可以根据需要调整血液供应，可以修补、愈合创伤，可以加快或减慢新陈代谢速度，甚至可以在一些器官受损或缺失的情况下做出适当的调整或补偿。在自组织系统的基础上增加智能性，就可以实现学习、交际、设计等过程；再加上器官移植技术，我们就可以极大地提高人体的适应能力。当然，这不是无限的，因为至少从现在看来，无论是人体自身还是智慧，都无法长生不老，任何人或器官最终都难逃死亡的宿命。

生态系统也具有相当强的适应力，多个物种相互依存，在同一片蓝天下迁徙，随着天气的变化、食物的丰俭以及人类活动的影响而繁衍兴旺或衰败消亡。由于很多种群和整个生态系统具有令人难以置信的丰富基因及变异能力，它们也具备“学习”和进化能力。如果时间足够久，它们就可以塑造出一个全新的系统，以充分利用各种变动的机会，获得生存和发展。

适应力与一直保持静止或恒定是不同的。有适应力的系统可能是经常动态变化的。事实上，短期的振荡、阶段性的发作，或者周期性的兴衰、高潮与崩溃，都是正常状况，而适应力可以使其复原。

相反，一直保持恒定的系统恰恰是不具备适应力的。因此，区分静态的稳定和适应力非常重要。静态的稳定很容易被观察，它是以一定周期内系统状况的变动来衡量的；而适应力则很难被观察到。除非超出了限度、调节回路受到冲击或破坏，或者系统结构被分解，否则你很难了解适应力是如何产生和运作的。如果没有完整的系统视角，人们看到的就只是系统表面呈现出来的动态或静态，而不是适应力。实际上，人们经常为了稳定或者提高生产率等目的而牺牲系统的适应力，有时候也可能会为了其他一些更容易被识别的系统特性而破坏系统的适应力。

- 给奶牛注射转基因的延缓生长激素可增加牛奶的产量，却不会相应地增加奶牛的食物摄入量。该激素可以将奶牛其他一部分身体机能的新陈代谢能量转化为产奶。虽然这样做可以增产，但其代价是降低了奶牛的适应力，使奶牛的健康状况恶化，寿命缩短，更加依赖于人类的管理。

- 近年来一些企业推行出了“及时生产”（Just-in-time）模式，不管是部件到制造商，还是产品到零售商，都降低了库存的波动性，减少了成本。然而，这种模式也使生产系统更加脆弱，容易受到燃料供应、交通流量、计算机瘫痪、劳动力短缺或其他阻碍的影响。

- 在欧洲，数百年来对森林的严格管理，已经逐渐将自然的原生林替换为单一树龄、单一树种的人工林，有的甚至大多不是本地树种。这样的森林管理模式可以产出更多更适合制造纸浆的木材。但由于缺乏多个物种之间的相互作用、交替荣枯，容易导致土壤的贫瘠，并且更加容易受到病虫害的危害，从而使得森林更为脆弱。此外，它们似乎对一种新的侵害更为敏感，那就是：人类工业或生产活动造成的空气污染。

对于人类来说，很多慢性疾病，如癌症、心脏病等，都源自人体适应力机制的崩溃，这类机制可以修复DNA、保持血管的弹性，或者控制细胞分裂。对于生态系统来说，很多生态灾难的发生也是因为适应力的丧失，例如一些物种的消失、土壤生化机制的破坏或者毒素的积累等。同样，各类大型组织（如企业、政府等）适应力的丧失，也是因为其对环境感知和响应机制、反馈过程过于冗长、低效，要么存在很多层级，要么有很长的时间延迟或信息失真。稍后，我们会在讨论组织层级

时对此进行探讨。

系统之思

不能只是关注系统的生产率或稳定性，也要重视其适应力，即自我修复或复位的能力，战胜干扰、恢复机能的能力。

我认为，适应力是系统运作的一个基础，正是因为适应力的存在，系统才可以正常地发挥和维持各种功能。因此，一个有适应力的系统就是一个大平台，在该平台支撑起来的空间里，系统可以自由地驰骋，一旦接近危险的边缘，就会遇到一堵柔软的、有弹力的“墙”将其反弹回来。随着系统适应力的下降，支撑的平台就会变小，那道保护墙也会变矮、变硬，直到系统如同运行在刀尖上，只要有一点震动，就随时可能坠落。因为人们通常更加关注系统是如何运作的，而忽视了其运作的空间，所以，在一般人看来，适应力的丧失似乎是突然来临的，是一种意外。但是，在此之前，系统其实早已是千疮百孔了。

当你认识到了适应力的重要性，你会找到很多方法去保持或增强系统自身的康复能力。

例如，在自然的生态系统中就蕴藏着这种认识，所以捕食者可以有效地控制害虫的数量；在和谐的养生保健生活方式中也体现了这样的认识，所以人们不只是有了疾病再去治疗，而是会增强人体自身的免疫力；同样，在积极的救援项目中也包含了这样的认识，所以救援者不只是简单地提供食物或金钱，而是努力创造条件，让人们可以自给自足，正所谓“授之以鱼，不如授之以渔”。

自组织

生物学家 路德维格·冯·贝塔朗菲

进化看起来并不是只由环境变化所引起的、贯穿于物种诞生和整个生存历程的一系列意外事件，而是有一定规则的未来；揭示这些规则将是人类最重要的任务之一。

一些复杂系统最令人称奇的特征，就是它们具有学习、多元化、复杂化和进化的能力。依靠这种能力，单个受精卵经过不断生长、分化，最终演化成一只青蛙、兔子或者人，体现了令人难以置信的复杂性；依靠这种能力，大自然中的一捧泥土有机质也能生生不息，滋养着不计其数、多姿多彩的生命物种；依靠这种能力，人类社会从刀耕火种，进化到发明了蒸汽机、水泵、专业化分工、大规模流水线生产、摩天大楼，以及全球化的通信网络。

系统所具备的这种使其自身结构更为复杂化的能力，被称为“自组织”（**self-organization**）。无论是从一片雪花身上，还是窗户上的冰晶，或者是过饱和溶液的结晶体中，我们都可以看到简单的自组织的工作原理和机制；但是，大自然中还包含着一些更为复杂的自组织过程，比如一颗种子生根发芽，一个孩子学会讲话，或者一个社区里的居民自发地联合起来反对有害垃圾倾倒等。

对于一个有机系统而言，自组织是一个非常普遍的特性，以至于很多人认为这是理所当然的。否则，我们就有可能被周遭世界中纷繁复杂的系统搞得眼花缭乱。当然，如果我们能够对自组织特性引起重视，我们就将会更好地鼓励而非破坏系统的自组织能力。

与适应力相似，人们也经常会出于追求短期生产率和稳定性的目的而牺牲掉系统的自组织特性。例如，把人和其他有机系统当做机器和生产过程中的附属品；或者减少农作物的基因变异性；或者建立官僚政治或组织，将人等同于符号或数字。

自组织特性会产生出异质性和不可预测性：系统有可能演变成全新的结构，发展出全新的行为模式。它需要自由和试验，也需要一定程度的混乱。但是，这些状况可能令人恐慌，或者威胁到现有的权力结构。结果是，教育体系往往限制了儿童的创造力，而不是激发这种能力；经济政策往往倾向于支持现有的大企业，而不是鼓励创新型的创业企业；同时，很多政府倾向于管制人民，而不是允许人们自发地组织起来。

幸运的是，自组织作为有机系统的一个基本特性，对于大部分冲击力都有一定免疫力。尽管以法律和维持秩序的名义，自组织能被长期压制、残酷打压，但它不可能被彻底消灭，而会顽强地持续下去。

过去，一些系统理论研究者曾认为自组织是系统的一种复杂特性，不可能被完全理解，人们也曾运用计算机建模技术去仿真、模拟一些系统的行为，当然，这主要针对的是一些内在作用机理清晰、可以定量描述的系统，而不是一些可以进化的复杂动态系统，因为人们会主观地认为后者不易被理解。

然而，新的发现表明，仅用一些简单的组织原则，就可以引起非常多样化的自组织结构。

想象一下，一个等边三角形，在每一条边的中间增加另外一个等边三角形，其面积是前者的 $\frac{1}{3}$ ；依此类推，得到的图形被称为“科赫雪花”（Koch snowflake，如图3—1所示）。它的边

长很长，但可以被包围在一个圆形之中。这种结构只是分形几何学的一个简单范例，这是数学和艺术的一个交叉领域，它通常使用一些相对简单的规则来产生精美的形状。

类似地，基于一些简单的分形规则，使用计算机就可以生成各种精致、优美、复杂的类似蕨类植物的形状。同样，虽然单个受精卵细胞裂变、生长成人的过程非常复杂，但其中也可能包含一系列类似的、相对比较简单几何规则。例如，在人体的肺部，肺细胞按照分形几何学的规则排列，所以，在有限的空间里，可以更大程度地扩展与空气的接触面积。事实上，如果把所有肺细胞展开的话，其表面积足以覆盖整个网球场。

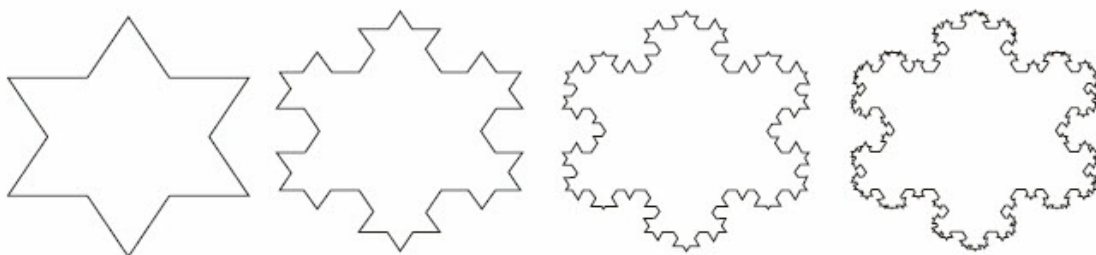


图3—1 “科赫雪花”演变图

即使像“科赫雪花”这样精美、复杂的图形，也可以从一组简单的组织原则或决策规则演化而成。

由此可见，一些简单的组织规则就可以产生复杂的自组织系统，以下是其他一些例证：

系统之思

系统通常具有自组织的特性，具有塑造自身结构、生成新结构、学习、多样化和复杂化的能力。即使是非常复杂的自组织形式，也有可能产生于相对简单的组织规则。

- 所有生命都是基于DNA、RNA和其他蛋白质分子等遗传机制中内含的基本组织规则繁衍生息的，从病毒到红树林，从变形虫到大象，均是如此；

- 农业的发展和相关机制都始于一个简单的创意，即人类可以在一个地方定居下来，拥有土地，并选择和培育作物；

- “上帝创造了万物，大地居于它的中心；城堡位于大地的中心；教堂位于城堡的中心”——这是中世纪欧洲人眼中的社会和物理结构的组织原则。

现代科学证明，自组织系统可以产生自一些简单的规则，衍生出多种多样的技术成果、物理结构、组织和文化。科学本身也是一种自组织系统，它倾向于认为，这个纷繁复杂的大千世界，往往生成自一些简单的规则。当然，究竟是否如此，科学到现在为止仍然未能给出答案。

层次性

18世纪英国诗人 乔纳森·斯威夫特

因此，自然主义者观察到，
一只跳蚤，会猎食比它更小的跳蚤；
更小的跳蚤仍然会猎食比它更小的东西；
如此反复，以至于无穷。

在新结构不断产生、复杂性逐渐增加的过程中，自组织系统经常生成一定的层级或层次性。

人们一般会认为，一个大的系统中包含很多子系统，一些子系统又可以分解成更多、更小的子系统。例如，你体内的细胞是某个器官的一个子系统，而那个器官又是你身体这一有机系统中的一个子系统；而你自身又是一个家庭、一支球队或一个组织的一个子系统；而它们又是一个城镇或城市、国家的一个子系统，依此类推。系统和子系统的这种包含和生成关系，被称为层次性。

很多事物，如公司、军队、生态系统、经济体系、有机体等，都具有层次性。这并不是偶然的。如果各个子系统基本上能够维系自身，发挥一定的功能，并服务于一个更大系统的需求，而更大的系统负责调节并强化各个子系统的运作，那么就可以产生并保持相对稳定的、有适应力和效率的结构。如果没有类似关系，很难想象系统最终会演变成什么样子。

系统寓言

为什么事物的组织呈现层次性

从前，有两个钟表匠，一个叫霍拉（Hora），一个叫坦帕斯（Tempus），他们都能制造精致的钟表，也各自有很多顾客。他们店里的顾客总是络绎不绝，电话响个不停，新订单源源不断。然而，多年以后，霍拉变得很富有，而坦帕斯却越来越穷。这主要是因为，霍拉发现了层次性原则。

霍拉和坦帕斯制造的手表都由近百个零件组成，坦帕斯依次组装这些零件，但是，在组装过程中，如果他不得不放下手头的活计去干其他事，比如去接个电话，半成品就会散成一堆零件；等他回来后，就只好从头开始组装。因此，顾客的电话越多，他就越难找出一整段不被打扰的时间，以完成一只手表的组装工作。

相反，霍拉制造的手表不像坦帕斯的那么复杂，他先把大约十个零件组装成一个稳定的部件，然后把十个部件组成一个更大的集合；最后只要把这些组件装在一起，一只手表就做好了。即使霍拉也和坦帕斯一样，不得不放下手头的活去接听顾客的电话，这也只会影响到他手头很小一部分工作。这样，他可以比坦帕斯更快、更有效率地制造手表。

只有存在稳定的媒介形式，一些简单的系统才能进化成复杂的系统。这样形成的复杂系统，天然地就具有层次性。这或许可以解释为什么在自然系统中，层次性比比皆是。相对于其

他各种可能的复杂形式，层次性结构是少数几种随时间而进化的结构之一。

系统之思

层次性是系统的伟大发明，不只是因为它们使系统更加稳定和有适应力，而且因为它们减少了信息量，使得系统各部分更容易记录和跟进。

在具有层次性的系统中，各个子系统内部的联系要多于并强于子系统之间的联系。虽然每件事物都和其他事物存在联系，但不同联系的强度并不一样。例如，在一所大学中，同一个院系或年级的人会更加熟悉，交流更多，与其他年级或院系的交流通常较少；组成肝脏的细胞，彼此之间存在更加密切的联系，而它们与组成心脏的细胞之间联系就较少。如果层级中每个层次内部和层次之间的信息连接设计合理的话，反馈延迟就会大大减小，没有哪个层次会产生信息过载。这样，系统的运作效率和适应力就得以提高。

在某种程度上，层级系统是可以被拆解的。由于各个子系统内部存在较为紧密的信息流，其自身也有一定功能，在被拆开之后，至少可以部分地像系统一样发挥作用。当层次被打破之后，子系统之间的边界通常被割裂开来。将系统不同层次拆分开，我们可以分别对其组成部分（例如细胞或器官）进行更为深入的研究，从而增加对系统的了解。因此，从这个意义上讲，基于还原论、解剖式的科学研究让我们学到了很多。但是，我们不能忽视各个子系统之间的重要联系，正是它们将各个子系统联系在一起，形成更高的层级，并可能在更高的层级上，生成让我们意想不到的行为。

例如，假设你得了肝病，医生通常会针对你的肝脏来进行治疗，而不会太关心你的心脏或扁桃体（因为它们处于同一个层级上，都是人体的一个器官），也不会考虑你的个性或者肝脏细胞核里面的DNA，因为它们分别位于更高或更低的层级上。但是，这当然有很多例外，也许的确需要上升到更高的层级去考虑整个层级结构，或许是你的工作使你长期接触某种化学物质，从而损害了肝脏的健康；或者需要深入到更低的层级去探究根源，或许你的肝病要归因于你的DNA功能障碍。

同时，你要认识到，随着时间的推移，自组织系统可以演化出新的层次，改变整合的程度，因此，你要考虑的因素也需随之调整。比如，过去美国的能源系统几乎是相互割裂的，但现在却完全不一样了。如果人们的思维没有随着能源经济的发展而进化，就会惊讶地发现他们已经变得如此依赖资源，而其决策已经与周围的世界背道而驰了。自组织系统可以形成层次，你可以在很多地方观察到类似过程。

例如，一个自由职业者如果工作量太大，就会雇用一些人作为帮手；一些小型、非正式的赢利性组织，吸引了很多会员，就会扩大预算，直到有一天，会员们会说：“嘿，我们需要有

人来组织一下”；一些单细胞的集合体会逐渐形成独特的功能，并产生一个分支循环系统以支持这些细胞，以及一个分支神经系统去进行协调。

一般来说，层次是从最底层开始向上进化的，从局部发展到整体，从细胞发展到器官和有机体，个人发展到团队，单人作业发展到生产管理。早期的一些农民逐渐聚居起来，形成了自治的村镇，而村镇之间产生了贸易行为。生命起源于单细胞细菌，而不是一头大象。层次性原本的目的是帮助各个子系统更好地做好其工作，不幸的是，系统的层次越高或越低，越容易忘记这一目的。因此，很多系统因为层次的功能失调，而不能实现并预定的目标。

如果团队成员过分追求个人荣誉，而忽视了团队整体的目标，团队就有可能失败；如果身体里一些细胞打破了该层次应有的功能，开始快速繁殖，就形成了人们常说的癌症；如果学生认为其目的是尽可能获得高分，而不是获取知识，就有可能出现考试作弊或其他不当行为；如果公司为了自己的利益而向政府部门行贿，市场竞争秩序和整个社会的公共福利就会受到损害。

当某个子系统的目标而非整个系统的目标占了上风，并牺牲整个系统的运作成本去实现某个子系统的目标，我们将这样行为的结果称为“次优化”。

当然，与次优化同样有害的问题是太多的中央控制。如果大脑直接控制身体的每一个细胞，导致细胞不能自我维持其功能，整个有机体就会死亡。如果大学领导直接决定每一位学生和老师的专业或研究方向，学生和老师就不能自由地探索自己感兴趣的知识领域，大学的使命也就不复存在；如果球队所有运动员完全听从教练的指挥，而不顾及场上的感觉，整个团队就不会有几分胜算。在历史上，当权者对经济的过度控制，无论是公司还是国家，都曾引发过一次次的大灾难，这样的例子比比皆是。

要想让系统高效地运作，层次结构必须很好地平衡整体系统和各个子系统的福利、自由与责任。这意味着，既要有足够的中央控制，以有效地协调整体系统目标的实现，又要让各个子系统有足够的自主权，以维持子系统的活力、功能和自组织。

系统之思

系统的层次性表明系统是自下而上进化的，上一层级的目的是服务于较低层级的目的的。

适应力、自组织和层次性是动态系统有效运作的三个原因。促进或精心管理系统的这三种特性，可以增强其长期保持有效运作的能力，保

持稳定。但是，别忘了，系统运作也可能充满了意外.....

第4章 系统之奇：系统的6大障碍

温德尔·贝里

美国肯塔基州农民、作家

麻烦的是，我们几乎完全愚昧无知，即使最博学的人也是无知的……知识获取从本质上看永远都是在破除无知，获得一些启示。我们关于世界的认识，首先告诉我们的就是，世界远远大于我们对它的认识。

我们在第2章中所讨论的一些简单系统的行为，其实已经令很多人感到困惑了，但是，以我多年从事系统思考教学和研究的经验来看，动态系统的行为变化，远远超出了我们的想象，并且，它们会持续地让我们感到惊奇。我自认为了解动态系统，但一旦遇到现实世界中的实际问题，我也经常感到尴尬和束手无策。

从中，我得到以下三项启示：

- 我们认为自己所知道的关于这个世界的任何东西都只是一个模型。每一种语言、每一个字，都是一个模型；所有的地图、统计数据、图书、数据库、方程式和计算机程序，也都是模型；包括我们头脑中认知和描述世界的方式，即心智模式，也是模型。所有这些都不是真实的世界，永远也不可能是。

- 我们的模型通常是与现实世界高度一致的。这就是我们为什么会成为这个星球上最为成功的一个物种的原因。尤其是我们经由对周围世界的直接观察、深入体验所建立起来的心智模式，包括对自然、人和组织的认知，是非常复杂而精密的。

- 然而，与第二点相反的是，我们的模型仍远远达不到能完整地描绘世界的程度。这就是我们为什么经常会犯错误、会感到出乎意料的原因。在同一时间里，我们的大脑只能跟踪少数几个变量。哪怕基于正确的假设，我们也经常会得出不合逻辑的结论；或者依据错误的假设，得出看似符合逻辑的结论。例如，对于一个指数级增长过程所引发的增长量的变化，我们大多数人都会感到惊讶；而对于复杂系统中的振荡，只有很少的人能够凭直觉从容应对。

系统之思

我们认为自己所知道的关于这个世界的任何东西都只是一个模型。虽然我们的模型确实与现实世界高度一致，但远未达到能完整地代表真实世界的程度。

简而言之，我们需要保持多方面的均衡：一方面，对于世界是如何运作的，我们确实有很多了解，但另一方面，这还远远不够；一方面，我们的知识储备令人惊异，但另一方面，我们也是如此无知；一方面，我们可以提高自己的认知，但另一方面，我们又不可能做到尽善尽美。这看似矛盾的几对关系，我都相信它们的存在，因为这就是我在多年对系统的研究中所学习到的。

在本章中，我们将讨论为什么动态系统的行为经常出人意料。换句话说，我们将从侧面解释，为什么心智模式不能很好地考虑到现实世界中的“并发症”，以及从系统的角度可以发现哪些问题。这将是一个警告清单，显示出哪些地方隐藏着障碍物。在这样一个相互联系、相互影响的世界中，你不能对这些障碍物掉以轻心。为了在复杂的世界里自由遨游，你需要把注意力从短期事件上移开，看到更长期的行为，看到系统内在的结构；你需要清晰地界定系统的边界，以及“有限理性”；你需要考虑到各种限制性因素、非线性关系以及时间延迟。如果不能很好地兼顾系统的适应力、自组织和层次性等特征，你很可能会误读系统、不当干预或者错误设计。

如果你真的理解了所有这些系统特征，那么你可能不会经常感到惊讶，但无论如何，你还是会感到意外。这对你来讲，到底是好消息还是坏消息，取决于你是否想要控制世界，并愿意接纳系统带给你的惊喜。

别被表象所迷惑

经济学家 肯尼思·博尔丁

系统是一个大大的黑箱子，

我们无法打开它的锁；

我们所能看到的，

只是什么东西进去了，什么东西出来了。

理解输入与输出的搭配，

再与其他参数联系起来，

这样，我们就可以发现输入、输出和状态之间的关联。

如果它们之间的关联清晰而稳定，

我们就可以预测出它可能的态势；

如果我们的预测不灵——但愿不要如此，

我们只能被迫打破它的盖子！

系统会呈现出一系列事件，迷惑我们，或者说，是我们自己在观察世界时被自己愚弄了。每天的新闻会告诉我们关于选举、战争、政治辩论、灾难以及股市涨跌的消息；我们大部分日常谈话也是关于什么时

间、什么地点发生了什么事，例如某支球队获胜了，某条河水泛滥了，道琼斯工业平均指数冲破10000点大关了；哪儿发现了一个大油田，哪儿的森林被砍伐了……如果把系统视为一个黑箱子，那么事件就是这个黑箱子时时刻刻的产出。

一些事件可能是很壮观的，例如崩溃、暗杀、巨大的胜利、可怕的悲剧，它们会引发人们的各种情绪。虽然我们已经多次从电视屏幕、报纸头条或网站首页上见过类似事件，但每一次事件都与上次不同，从而不断吸引着我们，正如每天的天气预报都不相同，我们却从未对它失去兴趣一样。就这样，各种各样的事件源源不断地吸引着我们的注意力，也时常让我们感到惊讶，因为我们看待世界的方式几乎没有预见性，也不能揭示其内在的原因。就像冰山浮在水面之上的部分一样，事件只是一个更为巨大的复杂系统中为人可见的一小部分，但往往不是最主要的。

如果我们可以看到相关的事件是如何积累形成动态行为模式的，我们就不会感到太惊讶了。

例如，一支球队处于连胜状态，大有夺冠的趋势；河水的消涨变化在增强，雨季水位更高，而旱季水流更少；道琼斯指数已经连续两年呈现上涨的趋势；发现新油田的频次越来越少；森林的砍伐与日俱增等。

系统的行为就是它的表现或绩效水平随时间变化的趋势，有可能增长、停滞、衰退、振荡、随机或进化。如果新闻报道有深度的话，它会把当前的事件与历史背景联系起来，这样我们就能够更好地理解行为层面的变化趋势，而不只是停留在较浅的事件层面。当遇到一个问题时，善于进行系统思考的人要做的第一件事，是寻找数据，了解系统的历史情况以及行为随时间变化的趋势图。这是因为，系统行为的长期趋势为我们理解潜在的系统结构提供了线索，而系统结构又是理解系统会发生什么以及为什么发生这些事情的关键，让我们不仅“知其然”，而且“知其所以然”。

系统结构是各种存量、流量和反馈回路的相互关联与作用。我们常用包含各种变量和箭头的图表（如存量—流量图等）来描述系统结构。结构决定了系统可能存在哪些行为。例如，一个调节回路会呈现出达成目标的行为，也就是说系统会接近或保持动态平衡状态；而增强回路则会引发指数级增长。二者联系在一起，则可能呈现出增长、衰退或均衡三种行为模式。如果其中包括时间延迟，则可能产生振荡。如果它们是周期性波动或快速爆发，系统则可能产生更多令人惊讶的行为。

系统之思

系统结构是行为的根源，而系统行为体现为随时间而发生的一系列事件。

系统思考需要反复审视结构和行为，善于系统思考的人会将二者联系起来，理解事件、行为以及结构之间的关系。例如，当你的手从“机灵鬼”玩具上移开时（事件），它就会弹起来并来回振荡（行为），而这一行为是因为“机灵鬼”自身的弹簧（结构）所引起的。

在“机灵鬼”这个简单例子里，事件、行为和结构之间的区别是很明显的，但对于其他系统，可能就不那么明显了。事实上，现实世界里很多分析都只是停留在事件层面上。如果你炒股，你就会发现，每天晚上的股评，很多都是就事论事的分析：因为美元下跌，所以股市上涨了；或者利率上升了，民主党获胜了，两个国家发生了冲突等。

这样的分析不能使你预测明天会发生什么，也不能让你改变系统的行为，比如使股票市场更少波动，或者使公司更健康，更好地鼓励投资等。

相对于事件层面的分析，大多数经济分析都会更进一步，到达行为层面。例如，一些计量经济学模型往往会以复杂的方程式，来发掘和表示收入、储蓄、投资、政府开支、利率、产出以及其他变量的历史趋势，以及它们之间的统计关系。

这些基于行为层面的模型比事件层面的分析更有价值，但它们仍然有一些根本性的问题。首先，它们普遍过分强调了系统流量，而对存量关注不足。经济学家们喜欢追踪流量的行为，因为那是最有趣的变量，也是系统最快表现出来的变化。

例如，经济新闻报告中经常会关注国民生产总值GNP（流量），而不是实体资本总量（存量）。前者是一个国家或地区的国民经济在一定时期（一般为一年）内，以货币形式表现的全部最终产品和服务价值的总和；而后者是一个国家或地区所有的工厂、机器设备、农场和工商企业等资本的总和，正是它们才能生产出各种产品和服务。但是，如果不了解这些存量如何经由各种反馈过程对相关的流量产生影响，你就不能很好地理解经济系统的动态，或者这些行为产生的原因。

其次，计量经济学家们试图发现各种流量之间在统计上的关系，但这往往是徒劳的，他们不过是在寻找一些不存在的東西。其实，在任何一個流量与其他流量之间都没有稳定的关系。流量有大有小、有开有关，存在各种组合，受存量而非其他流量的影响。

让我们用一个简单的系统来做一下解释。假设你对温度调节器一无所知，但是你有一组数据，记录了过去一段时间内房间里热量流入、流出的量。通过比较这些流量在过去这段时间里的变化，你可能会发现它们之间的关联，因为在正常情况下，它们都受同一个存量（室温）的控制，它们之间确实存在联动性。

你可以用公式去预测明天的室温，但是有一个前提条件，那就是：系统没有发生变化或故障。只要系统的结构发生了一些变化，你所发现的两个流量之间的关系很可能就会改变。例如有人打开了窗户，改善了保温效果，调整了火炉开关，或者忘了加油。如果房间的主人请你把屋子弄暖和一些，或者室内温度突然降低，需要你去想个办法，或者你希望用更少的燃料去维持同样的温度，那么在行为层面的分析就很难奏效了。你不得不再深入发掘系统的结构。

因此，基于行为的计量经济学模型在预测短期经济走势时很有效，但是在做长期预测时却表现很差；同样，它们在帮助人们找到如何改善经济的对策方面也无能为力。

同时，这也是各类系统让我们感到惊讶的原因之一。我们太沉迷于系统产生出来的事件，却很少关注系统行为的历史，也不善于从后者中发现线索，去揭示潜在的系统结构。要知道，系统结构才是系统行为与事件产生和演进的根源所在。

在非线性的世界里，不要用线性的思维模式

《混沌》作者 詹姆斯·格雷克

线性关系很容易理解：越多越好。线性方程组是可解的，因此广泛存在于各种教科书中。线性系统有一个重要的模块化属性，即你可以把它们拆分成一个个部件，然后重新组装起来。

非线性系统通常不可解，不能被拆分和拼装。非线性关系意味着身处其中的参与者可以随时改变游戏规则。变化的不确定性使得非线性关系难以被计算，但它也可以比线性系统产生出更为丰富多彩的行为。

我们通常并不特别擅长理解关系的本质。在系统中，如果两个要素之间的关系是线性的，就可以用一条直线来表述，它们之间有着固定的比例。例如，我在地里施了10磅肥料，就可以多收2斗谷物；如果施20磅肥料，收成将增加4斗，依此类推。

非线性关系是因与果之间不存在固定的比例关系；二者的关系只能

用曲线或不规则的线来表示，不能用直线。例如，如果我在地里施了100磅肥料，收成可增加10斗；如果施200磅肥料，收成也不会再增加了；如果施300磅，甚至出现减产。为什么呢？因为土壤的有机质被破坏了，庄稼被“烧”死了。

世界上到处都是非线性关系。因此，如果我们以线性思维来观察这个世界，就会经常感到惊喜。很多人都知道，一分耕耘，一分收获；付出两分耕耘，就可能有两分收获。但是，对于非线性系统就并非如此。两分耕耘，可能只能得到1/6的收获，也可能得到四倍的收获，或者根本没有收获。

以下是一些非线性关系的范例：

- 随着高速公路车流量的增加，车辆密度（每公里内汽车的数量）从零增长到一定限度时，汽车行驶速度会受到轻微的影响；然而，在超过这个限度之后，只要汽车密度再稍微增加一点儿，汽车行驶速度就会显著降低。之后，就会形成交通拥堵，汽车行驶速度会降到零。

- 在相当长的时间里，土壤流失都不会对收成造成多大影响，但是一旦表层土壤被冲蚀掉，植物的根系裸露出来，就会造成作物收成的大幅下降。

- 少量一些格调高雅的广告可以唤起人们对某款产品的兴趣；但是一大堆庸俗喧嚣的广告则可能招致人们的反感。

通过上面这些例子，你可以看出为什么非线性关系会让人感到惊讶——它们不符合看似合理的预期和推理，并不是少许努力就有少许回报，更多努力就会有更多回报；或者少数破坏性行为只产生可以容忍的少量伤害，更多这样的行为产生更大的伤害。如果以类似这样的预期来看待非线性系统，将会不可避免地碰壁或犯错。

理解非线性是非常重要的，不仅因为它们有悖于我们对行动与结果之间关系的正常预期，更重要的是，它们改变了反馈回路的相对力量对比，有可能使系统从一种行为模式跳转到另外一种。

对于我们在第3章中提到的几类系统而言，非线性关系是发生“主导地位转换”的主要原因，例如从指数级增长模式突然转变为衰退，就是因为非线性关系，导致系统由增强回路占主导转变为调节回路占主导。

为了让大家更形象地理解非线性关系的效果，我们来看一下云杉色卷蛾入侵北美森林造成破坏性伤害的案例。

云杉色卷蛾、冷杉和杀虫剂

树木的年轮记录显示，在过去至少400多年时间里，每隔一段时期北美地区云杉色卷蛾就会泛滥成灾，杀死大量云杉和冷杉树。但是，直到本世纪，几乎也没人对此特别留意。这是因为，在林业行业，有价值的是美国五针松，而冷杉和云杉一向被归入“杂木”类。然而，由于原生松树植株的消失，林业企业终于开始关注起云杉和冷杉树了。这也使得云杉色卷蛾突然被视为一种危害性极大的昆虫。

所以，从20世纪50年代开始，北美北部林区开始每年喷洒DDT农药，以控制云杉色卷蛾的泛滥。尽管每年都喷药，但色卷蛾虫害每年都会再度爆发。DDT的使用从50年代一直延续到70年代，直到DDT被禁止使用为止。在此之后，人们开始喷洒杀螟硫磷、高灭磷、西维因以及甲氧氯等农药。

尽管后来人们意识到杀虫剂不能彻底解决虫害问题，但它们仍然被认为是不可缺少的。正如一位护林人所说：“杀虫剂为我们赢得了时间。所有森林管理员都希望赢得时间，让树木生长，直到成材。”

到20世纪80年代，杀虫成本已经非常高昂了，加拿大新布伦瑞克省（New Brunswick）当年用于防虫的费用竟高达1250万美元。一些市民开始抗议，反对滥用有毒杀虫剂。尽管如此，云杉色卷蛾每年仍要毁掉多达2000万公顷（约5000万英亩）的森林树木。

不列颠哥伦比亚大学的霍林（Holling）和新布伦瑞克大学的戈登·巴斯克维尔（Gordon Baskerville）两人使用计算机建模技术，对虫害问题进行了全方位的系统分析。他们发现，在开始使用农药之前，在大多数年份里，云杉色卷蛾的数量都很少；它们受到很多捕食者的控制，包括一些鸟类、蜘蛛、寄生蜂，以及其他疾病的影响。然而，每隔一二十年，就会有一次色卷蛾虫害大爆发，持续约6~10年时间。接下来，害虫数量会减少，直到下一次爆发。

色卷蛾倾向于先侵害香脂冷杉，其次是云杉。在北部林区，香脂冷杉是最有优势的树种。依靠自身的特点，香脂冷杉可以排挤云杉和桦树，使它们无立足之地，从而使得整片森林逐渐成为香脂冷杉——“人”的天下。然而，每次色卷蛾虫害的爆发都会大大减少冷杉的数量，使得云杉和桦树有了生长的空间。然后，冷杉数量会慢慢恢复，并再次称雄。

随着冷杉数量的增加，色卷蛾虫害大爆发的可能性也迅速增加。但二者的关系是非线性的。色卷蛾的繁殖能力很强，数量增长远快过相应的食物供给的增长。最后，只要有两三个温暖、干燥、非常适合色卷蛾幼虫生长的春天，就可能爆发大规模的虫害。如果只在事件层面上分析，可能还会有人把虫害的爆发归咎于某个春季的天气呢。

色卷蛾数量增长太快，超过了它们的天敌能够控制的限度，二者的关系也是非线性的。在大多数情况下，大量色卷蛾也将导致捕食者数量的快速增加。此时，这是一个同向变化关系——色卷蛾越多，捕食者数量增长越快。但是，超过某个点之后，捕食者就不能再加速繁殖了。此时，二者关系转变为非线性关系，色卷蛾更多，但捕食者不能更快地繁殖。最后，大规模虫害爆发了，一发不可收拾。

现在，只有一件事可以阻止虫害的爆发，那就是：色卷蛾吃掉了大片大片的冷杉，导致自己的食物减少。到了这个时候，色卷蛾数量会迅速减少——这也是非线性的。原先，色卷蛾自我繁殖的增强回路（色卷蛾数量越多，新出生的色卷蛾就越多）占据主导地位；现在，大量色卷蛾被饿死的调节回路占了上风。之后，在冷杉死掉的地方，云杉和桦树又生长起来……如此周而复始。

数十年来，色卷蛾—冷杉—云杉的系统周期性振荡，但是总体来讲，仍在一定限度内保持着生态的稳定性。它们可以一直这样延续下去。色卷蛾的主要作用是不让冷杉一树独大，保持树种和环境之间的动态平衡。但是，生态上的稳定并不等于经济上的稳定。在加拿大东部，当

地经济几乎完全依靠伐木，而这主要依靠的就是冷杉和云杉的稳定供给。

当人类开始使用杀虫剂之后，整个系统的平衡就被打破了，系统内部各种非线性关系之间微妙的平衡点也很难达到了。因为杀虫剂杀死的不只是害虫，还有害虫的各种天敌，因此削弱了大自然中存在的控制害虫的反馈回路。这样一来，冷杉的密度一直保持在较高的水平上，使得色卷蛾的数量难以降低到数量开始快速降低的那个点，因而虫害可以每年爆发一次。

这样的森林管理措施使色卷蛾始终保持着“亚爆发状态”，也使森林管理当局坐在了一个随时可能爆发的“火山口”上，深陷其中，欲罢不能。继续用药的话，只能勉强维持，疲于应付；而如果放弃用药，色卷蛾的反扑力度又将空前强大，甚至可能彻底毁掉整个森林。

系统之思

系统中的很多关系是非线性的，它们的相对优势变化与存量的变化是不成比例的。反馈系统中的非线性关系导致不同回路之间主导地位的转换，也相应地引起系统行为的复杂变化。

恰当地划定边界

生态经济学家 加勒特·哈丁

当我们进行系统思考时，一个基本的误解是容易把它和另外一个流行词——“副作用”混淆。“副作用”这个词意味着，存在一些“我没有预见到或者想到的”结果。其实，副作用只不过是“主要结果”的副产品而已。进行系统思考很难，我们很容易曲解自己的语言，为自己开脱，让自己相信没有必要这么做。

还记得我们在第1章和第2章中画的一些结构图中存在一些云状物吗？千万不要小看这些云！它们是系统之所以让我们感到意外和惊奇的主要来源之一。

云表示的是流量的源头和终点。它们是一些存量，是流量的来源和去处，但是为了便于讨论，我们暂时对其进行了简化和忽略。它们标记着系统图的边界，但很少是系统真正的边界，因为系统很少有真正的边界。从某种意义上讲，每一种事物都与其他事物存在着联系，或多或少，或深或浅。例如，海洋与陆地之间并不存在泾渭分明、清晰明确的界线，社会学与人类学也很难截然分开，汽车的排气管和你的鼻子之间也存在着联系。所谓的边界，只是人为的区分，是人们出于观察、思考、理解、表达、交流等方面的需要，而在心理上设定的或社会上一般公认的虚拟边界。

系统最大的复杂性也确实出现于边界上。例如，在德国和捷克的边境线两侧，两国居民并非水火不容，而是你中有我，我中有你；在森林与草原之间，物种也在相互渗透，森林里的动物会溜达到草原上，草原上的动物也会在一定程度上渗透进森林中。因此可以说，恰恰是边界上

的无序、混杂，成为了多样化和创造力的根源所在。

在我们上一章提到的汽车库存系统里，导致经销商库存增加的流入量（即汽车到货量）的源头以云来表示。当然，汽车并不是从云彩里凭空掉下来的，它们需要经历一系列复杂的生产、运输过程，离不开各种原材料，以及资本、劳动力、能源、技术、管理等要素。类似地，库存的流出量（即汽车销售量）的去向也可以用云来表示，并不意味着这些汽车化为云朵飘到天上去了，而是销售给了千家万户使用。

因此，我们在分析问题时，究竟要不要深刻地考虑制造商的原材料、成品库存等因素，以及消费者对汽车的使用、维修、更换等问题，还是将其简化为一朵云，主要取决于我们要研究的是哪个系统，这些存量是否对于我们所关注的系统行为变化有显著影响。如果原材料的供应充足而稳定，而且顾客对汽车的需求是持续的，那么用云来表示就是有效的。相反，如果原材料出现短缺，或者产品供过于求，抑或是顾客的需求出现了重大变化，而我们在考虑问题时没有把这些存量划入系统的边界进行考虑，那么未来系统将出现令我们意想不到的事件。在这种情况下，就不能再将其化简为云、忽略不计了。

但是，在图4—1中，依然存在云，系统的边界还可以继续扩展：制造汽车的原材料来自化工厂、钢铁厂或炼油厂，而它们的输入都是来自于地球。同时，加工过程不只是制造出了产品，带动了就业，也创造了福利和利润，产生了污染。用户用旧了的汽车报废之后，要么被丢进了垃圾填埋场、焚化炉，要么开进了拆车厂、循环再生中心。在那些地方，它们继续对社会和环境施加着影响：垃圾场的废物、废气、废液泄漏，有可能造成水、空气污染；焚化炉产生浓烟和灰烬；循环再生中心则将一些材料回收利用。

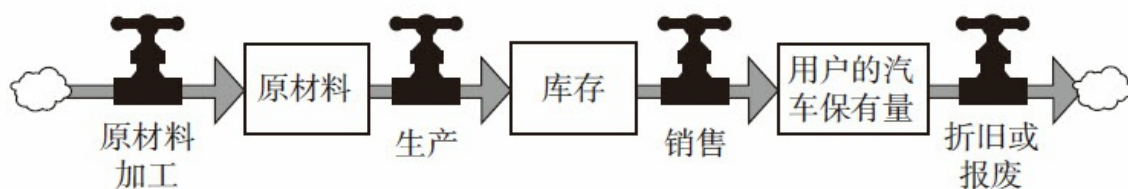


图4—1 揭开“云”背后的一些存量

究竟要不要考虑各种流量的所有环节，从矿山到垃圾填埋场，或者叫“从摇篮到坟墓”，这主要取决于谁想知道什么，出于何种目的，以及

在多长的时间范围考虑问题。从长期来看，完整地考虑流量的各个环节是很重要的；同时，随着实体经济的增长，社会的“生态足迹”不断扩大，“长期”也正在不断向我们靠近，成为“短期”。

似乎在一夜之间，垃圾填埋场已经被装满了，这让很多人感到诧异和不可理解，因为在他们看来，垃圾丢出去就应该是“消失”了。同样，很多原材料也不是取之不尽、用之不竭的，包括煤炭、金属、石油、天然气和水源等，都在以惊人的速度枯竭。

从足够长的时间维度上看，即使是矿山和垃圾填埋场也不是故事的全部。整个地球系统更大的物质循环驱动着各种物质的转化，沧海桑田不断变迁。亿万年之后，今天在垃圾填埋场里的所有东西都可能位于高山之巅，或者深海之下，形成新的金属矿藏或石油煤炭。对于我们居住的这个星球而言，没有什么东西是“云”，也没有最终的边界。即使空中飘浮着的真正的云，也是水循环的一部分。每一件事物都来自他处，每一件事物都走向他处，所有的事物都处在不断的运动、变化之中。

当然，这并不是说，每一个模型都必须包含所有的关联，直到囊括了整个星球。在建模时，云是模型必不可少的一部分，因为建模本身就是一个形而上的过程，必须合理地划分边界。有些东西确实是凭空生出来的，例如我们经常会说：“怒从心头起”，或者“爱意油然而生”，愤怒、爱、仇恨、自尊，等等，从字面意义上讲都是从“云”里来的，很难讲清楚来由。因此，如果我们试着去理解某一件事，就必须将其简化，这就意味着设定边界。而且，这样做通常是安全的。例如，我们把人口的出生和死亡看做是来自白云和化为云，并没有什么问题。

图4—2显示的就是“从摇篮到坟墓”过程的实际边界。但是，对于解释其他一些问题，例如出现大量的移民或人口的剧烈波动，抑或是探讨产房床位、墓地空间是否充足等，这些边界就显得过于宽泛且无用了。



图4—2 更大的云

对于系统思考者来说，认识和把握系统的边界是一个很难的课题，因为在系统中，并不存在一个明确、清晰划定的边界，而是要我们根据自己的需求和实际情况去划定。边界划定不当，很可能会带来一些问

题。

如果边界划定得太窄，一些对系统行为有显著影响的因素未被认真分析和对待，系统就可能产生令你意想不到的行为。

例如，你想解决城市交通拥堵问题，而未考虑到人们的居住、生活模式，这一问题将很难得到有效解决，一些对策甚至可能会引发更大的问题。比如，你建起一些快速路，试图缓解交通压力，但这吸引了一些房地产公司在快速路沿线开发地产项目；相应地，这些地产项目又增加了更多的交通流量，导致高速路也开始拥堵起来。

系统之思

世界是普遍联系的，不存在孤立的系统。如何划定系统的边界，取决于你的分析目的，也就是我们想问的问题。

如果你想使流经你所居住城镇的一条河的河水变得清澈，你就不能把边界只划定在你所居住的城镇，必须把整条河都考虑进来，因为如果上游的污染源得不到有效治理，你的努力就将是徒劳的。同时，你要考虑的还可能还包括河流两岸的土壤和地下水。当然，你可能无需考虑另外一个流域的问题，更不用考虑地球水循环问题。

在规划国家公园时，人们以前只是划定了一个物理的边界，认为把这里面保护起来，问题就能解决了。然而，在保护区之外，野生动物会定期迁徙、四处游荡，河流水系（包括地下水）也会流入、流出，人员往来穿梭，因此界线之外的人类活动、经济发展也会影响到保护区。近年来，这样的例子屡见不鲜，包括酸雨、温室效应引发的全球变暖、气候变化等。即使没有气候变化，想有效管理国家公园，你也必须把边界设定得比法定边界更宽。

与此相反，在进行系统分析时，人们经常陷入另外一个陷阱：把系统边界设定得太宽。在分析一个问题时，往往要画好几页的图表，密密麻麻，用很多箭头把各种事情都联系起来——在他们看来，这样才是一个系统嘛！唯恐如果你少考虑了一件事，就显得你不那么系统。

其实，“我的模型比你的大”只是一个浅薄的想法，其结果是分析过于庞杂，堆积了大量的信息，反而遮掩了问题的真实答案。例如，在模拟全球气候问题时，堆砌了所有的细节，列举了各种各样的原因，虽然很充实、有趣，但反而更容易让人迷失，无法发现如何减少二氧化碳排放、控制气候变化的关键。

在思考一个问题时，正确地设定边界很重要，但这和学术研究上的

边界以及政治上的边界不是一个概念。河流可能是两个国家之间天然的界线，但对于水源的数量和质量而言，如果也以这个作为界线，那结果就可能不尽如人意了。同样，大气治理也不能仅限于政治上的边界。对于温室效应、臭氧空洞或海洋污染来说，国家之间的边界没有任何意义。

理想情况下，对于每一个新问题，我们都应该在头脑中为其划定一个合适的边界。但是，我们往往很难保持这样的灵活性。一旦我们在头脑中划定了一些边界，它们就会逐渐变得根深蒂固，甚至理所当然。为此，很多争斗都是与边界有关的，例如国家之间的战争、贸易壁垒、人种差异、文化隔阂、公共部门与私营企业之间的责任划分、贫富差距、污染与被污染、现代人与子孙后代，等等。对于经济和政府、艺术与艺术史、文学与文学评论之间的边界问题，在学术上可以争论很多年。事实上，大学本身就是边界僵化的一个典例。图4—3展示了更多关于边界划定的示例。

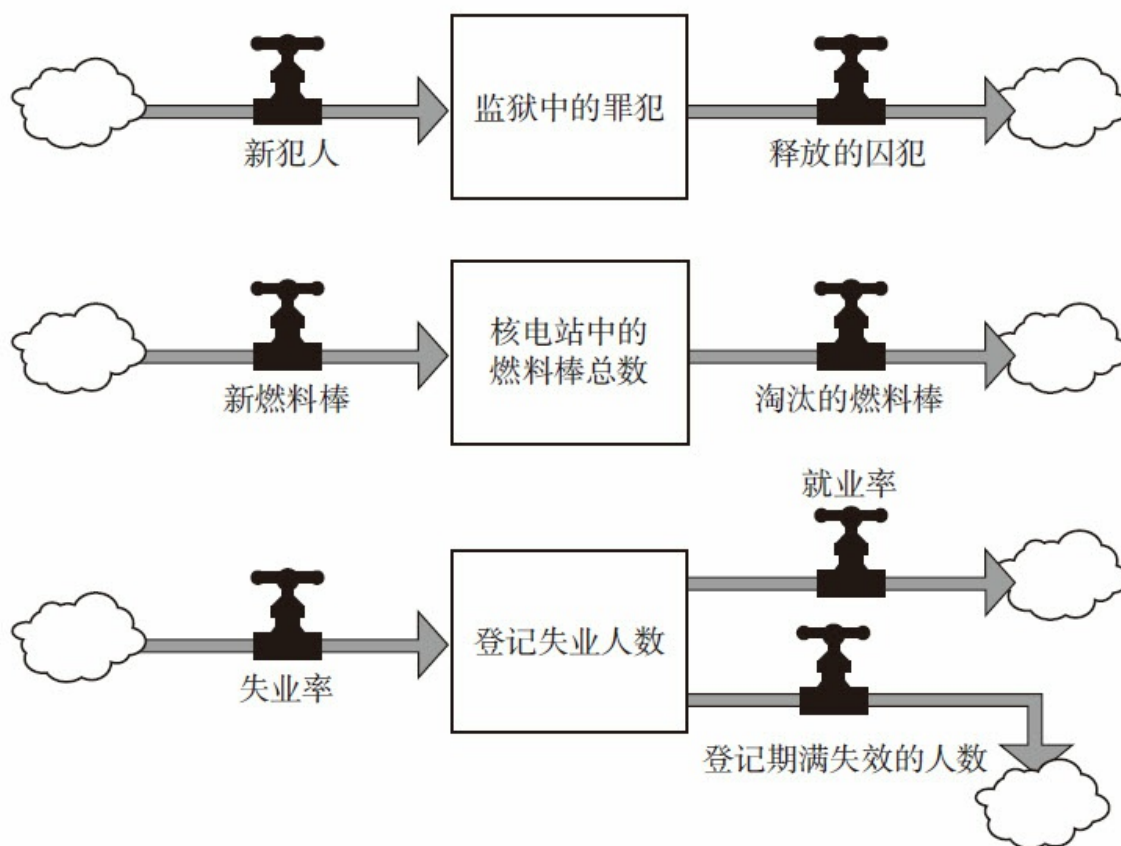


图4—3 更多关于“云”的示例

这些都是一些复杂的系统，你可以以此作为起点，但不能局限于既定的边界进行分析，不能让“云”阻止了

你的思考。根据你所研究的问题，你可能需要超越这些边界进行思考。例如，什么因素导致新犯人的增加？燃料棒淘汰之后去了哪里？失业人员在其登记期满失效以后怎么办？等等。

系统提示

正确划定系统边界

从某种程度上看，划定系统的边界需要较高的艺术性。请记住：边界是我们自己划定的，对于新的问题，出于不同的目的，可能需要并且应该对其进行重新考虑。事实上，在面对每一个新任务时，都应该忘掉在上一次任务中行之有效的边界划分；而针对当前问题的需要，应创造性地划定最合适的边界。对很多人来说，这是一个不小的挑战。但是，为了有效地解决问题，这也是十分必要的。

看清各种限制因素

系统之所以让我们感到惊奇，是因为在我们的思维中，倾向于认为单一的原因只会引发单一的结果。同一时间内，我们往往只能考虑一件或至多几件事情。而且，我们不喜欢考虑限制因素或范围，尤其是当我们在制订自己的计划或力求实现既定目标的情况下。

但是，我们所生活的这个世界并非如此简单。通常情况下，很多原因一起起作用，会产生很多种结果；多个输入产生多个输出；而且，几乎所有输入和输出，都会受到各种限制。例如，工业生产过程就需要：

- 资本
- 劳动力
- 能源
- 原材料
- 土地
- 水
- 技术
- 信用
- 保险
- 客户
- 管理
- 公共基础设施以及政府服务（如治安、消防、教育、医疗卫生等）

- 员工和客户的家庭、社区
- 稳定、健康的生态系统，供应或者支持各种原材料、生产要素，吸纳或处理生产中的废弃物等

在一小片土地上种植谷物，则需要：

- 阳光
- 空气
- 水
- 氮
- 磷
- 钾
- 各种微生物
- 松软肥沃的土壤以及有机质
- 控制野草和昆虫的一些系统或机制
- 保护其免受工业生产废弃物伤害的措施

根据德国化学家尤斯蒂斯·李比希（Justus von Liebig）提出的著名的“最小因子定律”，我们在上面列出的只是种植谷物必不可少的一些要素。在李比希看来，如果缺少其中任一项要素（如磷元素），即使有再多的其他要素都不管用；同样，如果问题是缺少钾元素，一股脑地施加再多的磷元素也毫无意义。

正如俗话说：巧妇难为无米之炊。如果没有酵母，即使有再多的面粉，也做不出面包来。小孩子不管吃了多少碳水化合物，如果缺少蛋白质，身体也不会健康。同样，不管有多少客户，没有资本，公司也无法正常运营；相反，资本再多，如果没有客户，公司的运营也难以为继。

这就是“限制因素”的概念，虽然很简单，但是很多人却对此存在误解。

例如，农学家认为，通过检测土壤中各种营养成分的含量，可以确定最佳的施肥量。但是，他们是否把作物生长所需要的各种必不可少的限制因素都考虑到了？人工制造的化肥会对土壤中各种微生物和有机质造成什么样的影响？化肥是否会影响或限制土壤其他功能的发挥？化肥的生产又受到哪些限制？这些问题是否已经得到周全的考虑，并可以清楚地回答？

一些发达国家将资本或技术转移到欠发达国家或地区，但是效果不

佳，受援助国的经济并未有大的起色，很多人对此大感困惑。其实，经济发展也有很多限制因素，不只是资本和技术，甚至这些东西根本不是受援助国经济发展最主要的限制因素。

在人类经济发展史上，确实有一段时期，制约生产力提升的最大因素是资本和劳动力。解决了这两个问题，就能推动经济发展。因此，大多数经济发展措施也只重点关注这两个因素（有时候也会考虑技术因素）。然而，随着资本和劳动力“瓶颈”的突破，经济逐步发展，日渐与生态系统息息相关，限制因素也开始转变为清洁的水源、空气、垃圾填埋场、能源以及原材料供应等。在这种情况下，如果还是按照传统只关注资本和劳动力，就事倍功半了。

在系统动力学领域，一个经典模型是由系统动力学奠基人、麻省理工学院教授杰伊·福瑞斯特开发的公司成长模型。在该模型中，一家初创企业非常成功，快速成长，而问题的关键是认识并处理各种限制因素，这些因素会随着公司自身的成长而不断变化。

例如，如果公司雇用了更多的销售人员，订单就会大量增加，从而超出了工厂的产能，导致交货延迟、部分客户流失。此时，产能成为最主要的限制因素。于是，管理者决定投资建厂，以缓解欠货压力。新厂建设过程中，需要雇用大量的新员工，并对其进行培训，而这通常慢于厂房和机器设备安装、到位的速度，并不受太多的重视，因而效果欠佳。于是，开始出现一些产品质量问题，再次导致客户流失。此时，员工的技能成为最重要的限制因素。所以，管理者又开始加强对员工的培训。逐渐地，质量得以改善，新的订单又大量增加，但是这时候，订单执行和跟踪系统又出现了拥堵……如此，不一而足。

系统之思

在给定的一段时间内，对于系统来说，最重要的一项输入是限制或约束力度最大的那个因素。

事实上，对于每一个工厂、儿童、流行病、新产品、技术、公司、城市、经济和人口等，都存在多重限制。你不仅需要知道有哪些限制因素，它们中的哪个或哪些起着主要作用，而且需要认识到，如果增长变缓或者约束变强，就说明起主导作用的限制因素正在发生改变。植物的生长与土壤的相互作用，公司扩张和市场的博弈，经济发展与资源之间的关系都是动态变化的。当一种因素的制约被解除了，成长就开始启动，而成长本身会改变各种限制因素之间的强弱对比，因此，相对最为稀缺的一种因素开始逐渐发挥作用。这时候，成长就会减缓或陷入停滞。要将注意力从相对丰富的资源（制约力没有那么强）转移到潜在的下一个最主要的限制因素，需要对成长过程的真正理解，并具备有效把控的能力。

任何一个有着多重输入和输出的物质实体，包括人口、作物生长、生产过程、经济发展等，都受到多重限制因素的制约。随着系统的发展，系统自身会影响和改变各种限制因素，也在这一过程中实时地受到各种限制因素的影响。系统与其限制环境之间构成了一个相互进化的动态系统。

系统之思

任何成长都存在限制，有些限制是自发的；而有些则是系统施加的。从根本上讲，关键不是追求持续成长，而是选择在哪些因素的限制之下维持生存。

然而，理解系统受到多重因素限制，对即将到来的下一个限制因素保持警惕，并不是实现持续成长的秘方。在限定的环境里，任何物质实体要想永远保持持续成长是不可能的。因此，从根本上讲，关键不是追求持续成长，而是选择在哪些因素的限制之下维持生存。如果你的公司能以可承受的价格生产某些精美的产品或服务，订单就会蜂拥而至，直到你无法应付，出现某些瑕疵或限制，降低了产品的完美程度，或者提高了它的价格；如果某一个城市居民的生活环境明显优于其他城市，人们就会络绎不绝地聚集到这个城市，直到超过该城市可以容纳的限度，导致生活环境变差。

成长总是不可避免地会受到限制，其中有些限制是自发的，而有些则是系统施加的。没有任何物质实体可以永远成长。如果公司管理者、政府当局、社会人口不能认识到其成长过程所面临的限制，并自觉地对其成长过程加以选择和控制在，那么环境也会做出选择，并施加限制。

无所不在的时间延迟

捷克共和国第一任总统、剧作家 瓦茨拉夫·哈维尔

我惊恐地意识到，我急迫地希望重建民主，有些做法却几乎已经和理想主义者一样了。我一直希望更快地推动历史的发展，却犯了“拔苗助长”的错误。

我意识到，当我们试着创造一个新事物时，我们必须学会等待。我们必须充满耐心地播种，精心浇灌土地，让种子自己发芽、生长，它需要时间。你不可能愚弄植物，你更不可能愚弄历史。

无论是植物的成长、森林的荫蔽，还是民主的发展，都需要时间。同样，你投入邮箱里的信到达目的地，需要时间；顾客注意到价格的变化，并据此调整他们的购买行为，需要时间；建造一座核电站，需要时间；机器的磨损，需要时间；新技术的普及，也需要时间。

我们经常对事物发展需要多少时间感到出乎意料。对此，福瑞斯特告诉我们，当我们在搭建模型或应对延迟时，可以先去问一问系统里的每一个人，他们认为时间延迟有多长，对其进行分析、做出最精准的预测，然后再扩大三倍。我也发现，对估计写一本书需要多长时间，这一修正方式也同样适用。

在系统中，时间延迟比比皆是。每一个存量都是一个延迟，大部分流量也有延迟，包括运输延迟、感知延迟、处理延迟、成熟延迟等。以下是我们在建模过程中发现的一些很重要的延迟范例：

- 在病毒感染和症状发作去就医之间存在延迟（有时候也称为“潜伏期”）——有可能是几天，有可能是几年；
- 在污染发生与污染物在生态系统中扩散或积累、造成危害之间，存在延迟；
- 能繁家禽和农作物从出生到成熟之间存在延迟，使得这些大宗商品的价格具有周期性振荡的特性：猪的周期为4年，奶牛的周期为7年，可可树的周期为11年；
- 改变人们认为合适的家庭规模的观念至少需要一代人的时间；
- 生产线更换工装以及资本存量的周转也有延迟。设计一款新车并将其投放市场需要3~8年的时间，它在市场上作为新车销售的生命周期约为5年，而它可以在路上平均行驶10~15年。

正确地划定系统的边界取决于讨论的目的，对于重要的延迟也是这样。如果你所关心的波动持续数周，你大可以不必考虑数分钟或数年的延迟；如果你关心的是长期的人口规划、经济预测，时间跨度可能持续数十年，你通常可以忽略为期数周的振荡。因此，一项延迟是否显著取决于你试图理解的频率处于哪一档。

在第2章“系统大观园”中，我们已经用实例展示了一些重要的反馈回路是如何影响系统行为的。改变延迟的长短可以彻底地改变系统行为，同时，延迟也常常作为敏感的政策杠杆点。这是因为，如果系统中的一个决策点对某一些信息存在反馈延迟，使反馈延迟更长或更短一些，决策就可能偏离目标，从而导致人们为达到目标而采取了更多或更少的行动。同时，如果采取行动太快，则可能因过度反应而放大短期的波动，产生不必要的振荡。时间延迟决定了系统的反应速度有多快，达到目标的准确性，以及系统中信息传递的及时性。矫枉过正、振荡和崩溃也经常是由时间延迟所引起的。

系统之思

当在反馈回路中存在较长的时间延迟时，具备一定的预见性是必不可少的。如果缺乏预见性，等到一个问题已经很明显了才采取行动，将会错过解决问题的重要时机。

认识时间延迟，有助于我们理解为什么戈尔巴乔夫可以在几乎一夜之间改变了苏联的信息系统，却不能改变实体的经济——这需要数十年时间。我们也可以更好地理解为什么东西德国合并，社会所经历的痛苦期要远远长于政治制度的变革。因为建设新的发电厂需要很长时间，电力工业经常被周期性的波动所困扰，要么是供过于求、产能闲置，要么是供不应求，导致拉闸限电。因为海洋对气候变化的反应时间延迟长达数十年，导致人类燃烧化石燃料排放的二氧化碳已经对气候变化造成了严重影响，一两代人都难以修复。

有限理性

亚当·斯密

只要每一个人都竭尽所能地管理好自己的资本，使其价值最大化，就能支持本国的产业……从主观上讲，他的确并不想提高公共利益，也不知道自己对公共利益有多大帮助……他的目的仅仅是促进个人的利益所得……他在一只“看不见的手”的引导下，在实现自己利益的过程中，无形之中增进了公共利益。通过追逐个人利益，他可以更多、更有效地促进社会利益，甚至比他真心希望促进社会利益时还要有效。

如果市场这只“看不见的手”真的能够引导个体在追逐私利的同时也增进集体的福利，那确实是太棒了。那样的话，不只是物质上的自私将成为社会美德，对于经济的数量模拟也将变得容易得多。无需考虑他人的利益，或者复杂的反馈系统的运作。难怪亚当·斯密的模型在200多年的时间里都一直有着如此强大的吸引力。

不幸的是，事实并非如此。这个世界呈现给我们更多的现实是，人们倾向于理性地从自己短期的最大利益出发，但每个人的行为汇集起来的结果却是所有人都不愿意看到的。

例如，世界各地的游客蜂拥至夏威夷的威基基海滩或瑞士瓦莱斯州的采尔马特冰川，却开始抱怨这些地方已被众多旅游者破坏了；农民生产出了过多的小麦、黄油或奶酪，然后价格突然暴跌；渔民们过度捕鱼，却最终导致自己没了饭碗；众多公司看到有利可图就加大投资，最终导致供大于求和商业的周期性波动；穷人本来就不富裕，却比富人生育和抚养着更多的孩子。

为什么？

这就是世界银行经济学家赫尔曼·戴利所说的“看不见的脚”（**invisible foot**），或者是诺贝尔经济学奖获得者赫伯特·西蒙（Herbert Simon）所说的“有限理性”（**bounded rationality**）。

系统之思

有限理性意味着，人们会基于其掌握的信息制定理性的决策，但是由于人们掌握的信息通常是有限的、不完整的，尤其是对于系统中相隔较远或不熟悉的部分，由此导致他们的决策往往并非整体最优。

例如，渔夫并不知道哪里有多少鱼，也不了解同一天里其他渔夫多捕或少捕了多少鱼。同样，一位商人也无法确切地知道其他商人正在计划哪些方面的投资，有哪些顾客愿意购买，或者各种产品之间如何相互竞争。他们不知道市场的规模，也不知道自己当前所占的市场份额。他们有关这些方面的信息都是不完整的，存在时间延迟，而且他们自身的反馈也存在延迟。因此，经常出现系统性地投资过度或者产能不足。

正如西蒙所说，我们并非无所不知、理性的乐观主义者；相反，我们是浮躁的“自足自乐者”，在做下一个决策之前，总是试图最大程度地满足当前的需求。我们会以理性的方式尽力维护和扩大自身的利益，但是却只能基于自己所知道的信息进行思考。除非他人有所行动，否则我们不会知道他们计划做什么。我们也极少看到自己面前存在的所有可能性，也通常不会预见到自己的行动对于整个系统的影响，甚至有可能选择性地忽略这种影响。因此，我们只能在自己有限的视野范围内，从当前几种很明显的选择中进行抉择，并坚持自己的看法，不会考虑整体的长期最优方案。只有在被迫的情况下，我们才会改变自己的行动。

一些行为科学家认为，我们甚至不能很好地解读自己所掌握的那些有限的信息。

- 我们会对风险做出错误的估计，将其中一些事情的危害程度估计得过高，或者轻视其他一些事情的危险性。

- 我们也容易过度夸大当下状况的重要性，对眼前的经验非常重视，而未对过去给予足够的重视。

- 我们会更加关注当前的事件，而对一些长期的行为不那么关心。对于未来的价值，我们会按照自己的价值判断，打一些折扣，这些价值判断会受到经济或生态等方面的影响。

- 对于所有输入的信号，我们不能正确地评估它们的重要性。我们不会全盘接受自己不喜欢的，或者不符合我们心智模式的所有信息。

这就是说，即使为了最优化自己个体的利益，我们有时也不能做出完全正确的决策，更别提系统整体的利益了。

有限理性理论对主流经济学提出了挑战，后者是建立在200多年以前亚当·斯密的政治经济学基础之上的。如果你对经济学有一定了解，就很容易看到二者之间的差别有多大，这也导致了长期激烈的论战。发源于亚当·斯密的经济理论首先假设，每个市场主体都是基于完备的

信息、完全理性地做出行动的“经济人”（homo economicus）；其次，当各个行为主体按照这些规则行动时，他们的行动累加起来，就会产生对每个人来说都是最优的结果。

从长期证据来看，这两个假设都无法成立。在下一章中，我们会深入探讨其中一些最常见的结构，它们可能导致有限理性，并最终酿成灾难。它们是一些常见的现象，包括上瘾、政策阻力、军备竞赛、目标侵蚀以及公地悲剧等。总之，如果不能理解有限理性，你将会备感意外。

假设你因为某种原因，离开了自己所熟悉的生活环境，来到一个完全陌生的社会之中，对于那里人们的行为，你很难理解。

或者，之前你是一个坚定的反对派，而现在突然成了政府的一员。或者，你之前是一名普通员工，现在突然变成了管理层^[4]。或者，你之前是一名环保主义者，一直批评某家大企业的新建开发项目，而现在，一夜之间，你成了该企业的员工，需要推进这一项目。

如果发生这类巨变，各个方面都突然不同了，你将如何应对？频繁地经历这样的转变，能帮助人们拓宽视野吗？

在你的新位置上，你会经历新的信息流、激励和限制因素、目标、差异以及压力——只有发生这种彻底的转变，你才能如此真切地经历和感受到有限理性。也许你会回想起以前从另外一个角度是如何看问题的，而且在你转变了系统视角之后，也会激发一些创新，但这无疑是不可能发生的。如果你变成了管理者，你可能不再会把员工看做是生产线上有功劳的伙伴，而是将其视为需要缩减的成本；如果你变成了投资人，你也可能和其他投资人一样，在繁荣期过度投资，在衰退期投资不足；如果你变得一贫如洗，可能也会对每天的每一项开支斤斤计较、精打细算，并期盼着哪一天能有发达的机会，或者打算多生几个小孩；如果你是一名渔夫，渔船是抵押贷款购置的，要供养家庭，又对鱼群数量的状况也不了解，在这种情况下，很可能也会过度捕捞。

系统之思

要想改变行为，首先要跳出你所在系统中固有的位置，抛弃当时观察到的有限的信息，力求看到系统整体的状况。从一个更广阔的视角来看，可以重构信息流、目标、激励或限制因素，从而使分割的、有限的、理性的行动累加起来，产生每个人都期盼的结果。

在教学中，我们通过游戏模拟演练也能发现类似结果。在这个过程中，学生们置身于模拟的情境，扮演各种不同的角色，对真实、不完整的信息流做出反馈。作为渔夫，他们会过度捕捞；作为发展中国家的最

高领导人，他们会优先考虑核心产业的需要，而不是人民的需求；作为上流社会成员，他们会精心装饰自己的爱巢；作为社会的底层，他们会变得冷漠或逆反。如果换成是你，也会如此。在美国心理学家菲利普·津巴多（Philip Zimbardo）实施的著名斯坦福监狱试验中，令人感到震惊的是，在很短的时间内，扮演者的态度和行为就能和真正的看守与囚犯十分相似。

有时候，在信息有限的情况下，个人也能做出理性的决策，但有限理性并不能成为人们目光短浅的借口，而是为我们提供了理解为什么会产生这些行为的机会。处于系统的特定位置上，他的所见、所知都是有限的，而其行为是合理的。在有限理性的情况下，如果换成另外一个人，结果仍是相同的。因此，只是责备个人，并不能有助于产生更加符合人们期望的结果。

令人感到惊奇的是，只要稍微强化一下有限理性，更好、更及时地提供更多、更完备的信息，行为的转变其实可以是很快、很容易的。

系统寓言

荷兰房屋的电表

在阿姆斯特丹郊区的一个地方，有一些同一时期建造的独栋别墅，样子几乎一模一样。由于某种未知的原因，其中一些房屋的电表被安装在了地下室，而另外一些则被安装在前厅里。

这些电表都有一个透明玻璃罩，里面有一个小的水平金属圆盘。家庭用电越多，圆盘就转得越快，而电表的刻度盘上显示着累积的用电度数。

在20世纪70年代早期石油禁运和能源危机时期，荷兰政府开始重视能源的使用。统计发现，在这个地区，有些家庭的用电量比其他家庭少1/3。对此，没有人可以给出合理的解释，因为所有家庭都是类似的，用电价格也一致。那么，为什么会有这样的差别呢？

调查结果表明，差别取决于电表的安装位置。用电多的家庭，电表都是安装在地下室，人们很少能看得见电表；而用电少的家庭，电表则是安装在前厅中，每当人们走过，都能看到电表的小圆盘在转动，提醒人们本月的电费在不断增加^[2]。

尽管存在有限理性，只要系统的结构设计得很精致，仍然可以在合适的时间、合适的地点做出合适的反馈，维持着适当的功能。例如，在一般情况下，你的肝脏只会得到必要的信息，去完成其需要执行的任务。在未受扰动的生态系统和传统文化中，每个人、物种或种群，都以自己的方式服从并服务于系统整体，保持着整体的稳定性。尽管每个个体都有自己的策略，但这些系统和其他很多系统总体上都是可以自我调节的。正常情况下，它们不会产生问题，也无须设置治理机构或者制定这样那样无用的政策。

系统之思

系统中每个角色的有限理性可能无法产生促进系统整体福利的决策。

从亚当·斯密之后，人们一直相信，自由竞争市场是一种设计得当、可以自我调节的系统。在某种程度上，它确实如此；然而，在另外一些情况下，事实就并非这样了，而且大家都很容易发现类似的证据。自由市场经济体系的确允许生产商和消费者，对生产机会和消费选择拥有最好的信息，以做出公平的、不受限制的、理性的决策。但是，这些决策本身并不能纠正整体系统内生的垄断倾向，以及一些不利的副作用（外部性），比如对穷人的歧视，或者产能的波动。

让我们借用一句常用的祈祷词，并稍加改造来祈祷吧：

让神赐给我们一颗平静的心，在结构精致的系统中自由地使用我们的有限理性；让神赐给我们勇气，去重塑结构不良的系统；让神赐给我们智慧，去理解其中的差别。

受到信息、动机、抑制因素、目标、压力以及对其他角色的限制等因素影响，系统中的每一个角色都存在有限理性，这可能会产生促进系统整体福利的决策。如果不能，即使在同一个系统中放进新的角色，也不会改善系统的表现。要想有所变化，就必须对系统的结构进行重新设计，改进信息、动机、抑制因素、目标、压力以及对某些特定角色的限制等。

[1] 在西方，通常劳资双方、员工和管理层之间存在难以调和的利益冲突或矛盾。——译者注

[2] 这个故事是我1973年在丹麦科勒克勒（Kollekølle）参加一次会议时听说的。——作者注

第5章 系统之危与机：系统的8大陷阱与对策

约翰·罗尔斯顿·索尔

政治学家

理性的精英们相信，通过科学技术的发展，可以对世界上每一件事做到一览无余，但是他们都缺乏一个更为广阔的视野。从马克思主义者到耶稣的信徒，从哈佛MBA到部队军官，都是如此……虽然主张各异，但他们都有一个共同的担忧：如何使特定的系统有效运作；同时，社会文明也日益变得没有目标和难以理解。

对于时间延迟、非线性、模糊的边界，以及其他一些令我们出乎意外的特性，在任何系统中都可找到。一般来说，它们是系统固有的特性，不可以被改变。这个世界是非线性的。如果为了管理的方便，非要用数学或机械式方法使其线性化，即使可行，也是不明智的。事实上，这几乎是不可能的。同样，边界也与你所要研究或应对的问题相关联，虽然有必要对其进行组织和澄清，但它本身也是非常模糊、短暂且易变的。要想使复杂的系统不再那么让我们出乎意外，最主要的途径就是加强学习，提高对复杂性挑战的理解、尊重和利用能力。

但是，有些系统非常难以理解和驾驭，不只是出乎我们的意料，它们甚至是违反常理的，其结构方式注定会产生一些问题，让我们陷入巨大的麻烦之中。系统性问题的表现形式很多，有些是独特的，但有些却非常常见。我们把产生常见问题行为模式的系统结构称为“基模”（**archetypes**），诸如公地悲剧、目标侵蚀和竞争升级等。这些基模是如此常见，以至于我只用了不到一周的时间，就在《国际先驱导报》上找到了足够多的例子，以支持本章对各种基模的论述。

仅仅理解问题产生的基模结构是不够的，试着去容忍它们也不足够，我们需要去改变它们。对于它们造成的破坏，人们常常指责其中的某些参与者，或归咎于某些事件，但实际上，这不过是系统结构使然。用指责、惩罚、建立或调整某些政策这样的“标准”应对方式，很难修正结构性问题，就像试图让编排好的剧情有一个好的结果，或者对破镜子修修补补一样，都难以奏效。这就是我把这些基模称为“陷阱”的原因。

当然，系统陷阱也是可以避开的，但前提是要预先识别出陷阱，不去触发它们或者改变其结构，比如重设其目标，增强、减弱或改变反馈回路，增加新的反馈回路等。因此，除了把它们叫做“陷阱”之外，我也

把这些基模称为“机会”。

政策阻力：治标不治本

小约翰·库斯曼，《国际先驱导报》，1992

邓白氏公司首席经济师约瑟夫·邓肯曾说过：“从历史上看，我相信投资税贷款是刺激经济的一个有效措施。”

但是，反对者也大有人在。他们认为，无人可以证明经济发展得益于投资贷款。在过去30年中，这一政策一再被批准、调整和废除。

就像我们在第2章中所讲到的，调节回路结构的主要表现就是会消除外部力量对系统的影响，使系统特定的行为模式保持相对稳定，没有太多变化。对于我们每个人来说，这是一个伟大的结构，因为它可以使我们的体温保持在摄氏37℃（或华氏98.6°F）。但是，其他一些长期持续的行为模式，可能并不符合人们的预期，往往被视为一个问题。尽管人们发明了各种技术、采取了多项政策措施，试图去“修复”它们，但系统好像很顽固，每年都产生相同的行为。这是一种常见的系统陷阱，人们习惯称之为“治标不治本”^[1]或“政策阻力”^[2]。

在我们的日常工作与生活中，“治标不治本”的例子比比皆是：

- 对于农产品，人们年复一年地采取各种措施，试图减轻供大于求的情况，但产品过剩的问题仍然存在；
- 对于毒品泛滥，社会采取了各种禁止、打击措施，但毒品泛滥依旧；
- 当市场不景气、从本质上看并不利于投资时，政府依然会出台投资税收贷款或其他刺激投资的政策，但事实上几乎没有什么效果；
- 在美国，不要指望任何一项单一的政策可以降低医疗成本；
- 在几十年时间里，美国政府一直在“创造就业”，但失业率长期居高不下。

我相信你也可以举出一大堆类似的例子，人们前仆后继地努力，但均于事无补、徒劳无功。

“政策阻力”来自于系统中各个参与者的有限理性，每一个参与者都有自己的目标，都会对系统进行监控，观察一些重要变量的变化态势，如收入、价格、房屋供给、毒品交易或投资等，并将其与自己的预期或目标进行对比。如果存在差异，每一个参与者都会采取某些措施，试图扭转当前的局势，使其符合自己的预期或目标。一般来说，目标与实际

状况之间的差异越大，行动的压力或强度就越大。

当各个子系统的目标不同或不一致时，就会产生变革的阻力。想象一下我们在第2章提到的单存量系统，例如城市里的毒品供应，不同的参与者对同一个存量有不同的期望，希望将其拉向不同方向：吸毒者希望毒品供应充足；执法部门希望减少乃至杜绝毒品；贩毒组织则希望毒品的供应量既不太多也不太少，以保持价格和收入相对稳定；普通居民真正想要的是社会治安稳定，减少吸毒者抢劫的风险。每一个参与者都尽力采取措施，以实现自己的目标。

如果某一个参与者占据了优势地位，使得系统存量朝一个方向运动，那么，其他一些参与者将会付出加倍的努力，把系统存量往相反方向拉：执法部门设法切断了毒品走私的渠道，导致毒品供应量减少；市面上毒品的价格暴涨，吸毒者不得不实施更多犯罪（如抢劫），以筹集到更多的钱去购买毒品；而价格暴涨给贩毒者带去了更多的利润，使得他们可以更大力度地加强毒品的走私（如购买飞机或轮船，逃避边境检查）。结果是，一方的努力成果会被抵消掉。事实上，某一方的成果越大，反方向的抵消力量往往也会越大，导致存量与以前的状况没有太大的差别，而这同时也是每一个人所不希望看到的。

在一个具有“政策阻力”的系统中，多个参与者有不同的目标。如果任何一方的态度有所让步或放松，其他各方就会把系统往更靠近自己目标的方向拉，导致系统更加远离让步一方的目标。因此，每一方都不得不付出巨大的努力，以使系统保持在谁也不希望看到的状态。事实上，这一类系统结构以类似“棘轮”的模式在运作：任何一方增强的努力，将导致其他所有人的努力也得到加强。这种僵持不下并不断强化的模式很难缓解，有人可能会说：好吧，为什么大家不能都退后一步呢？要做到这一点很难，需要大量的信任和沟通，而有些沟通是很难做到的（如上面所讲的贩毒的例子）。

“政策阻力”的结局可能是悲剧。

1967年，当时的罗马尼亚政府认为需要增加本国的人口，遂做出决定：45岁以下妇女的流产行为是非法的。并开始一刀切地禁止各种流产行为。很快，出生率就增加了两倍。接下来，罗马尼亚的人口就遭遇到了“政策阻力”的报复。

虽然避孕和流产仍然是非法的，出生率却缓慢回落到接近政策出台之前的水平，而育龄妇女死亡率较之前增加了两倍。这主要是由于存在大量危险的、非法的流产。此外，这项政策也造成孤儿大量增加，因为一些孩子虽然被生了出来，但其父母却并不想要他们，又不能流产，于是被抛弃成为孤儿。一些贫困家庭无力抚养多个孩子，深知无法给他们良好的教育，于是对

政府增加人口的政策进行抵制，因为这不仅会增加他们自己的生活成本，而且也不利于孩子长大成人以后的生活。

应对“政策阻力”的一种方式，是努力压制它。如果你拥有足够大的权力，你可以行使权力去压制它，但相应的代价可能是招致怨恨（因为可能不符合人们的期望），而一旦权力有所放松，则可能带来爆炸式的反弹。这就是罗马尼亚人口政策制定者、独裁者齐奥塞斯库（Nicolae Ceausescu）所面临的状况。他尽力维持自己的权力，以压制对其政策的反抗。当他的政府倒台后，他和家人都被处以死刑。新政府颁布的第一部法律，就是废除对流产和避孕的禁令。

相对于压制，应对“政策阻力”的另一种方式是，放弃、废止无效的政策，将资源和能量应用于增强和坚持更具建设性的目标。这是违反人们的直觉的，因而几乎是不可想象的。在系统中，你可能得不到这种指示，但是如果你走错了方向，你也不会走很远，因为你要花费很大的精力去采取纠正措施。这时候，如果你安静下来，那些抵抗你的人也会安静下来。这种情况发生在1933年，美国终止了禁酒令，由此导致的混乱也基本上停止了。

安静下来，可以为人们提供更深入地审视系统内部反馈的机会，让我们理解人们行为背后的有限理性，发现更加符合系统各种参与者目标的方式，使系统的状态逐渐迈向更好的方向。

例如，一个国家要提高人口出生率，可以先去了解为什么父母不愿意多生孩子，然后再根据具体原因采取顺应人们行为模式与意愿的政策。父母少生孩子的原因，可能是他们没有相应的资源、生活空间或时间，也可能是因为他们对未来缺乏安全感。在罗马尼亚禁止流产的同一时期，匈牙利也对低出生率感到忧虑，因为这可能会使未来缺乏足够的劳动力，从而导致经济衰退。匈牙利政府发现，住房是影响家庭生育的一个原因。于是，政府出台了相应的奖励计划：人口较多的家庭可以拥有更大的房屋。这个政策起到了效果，但效果也很有限，因为住房只是影响生育的因素之一。但是，相对于罗马尼亚政府采取的政策，匈牙利的政策效果很明显，而且避免了灾难性的后果。

应对“政策阻力”最有效的方式是，设法将各个子系统的目标协调一致，通常是设立一个更大的总体目标，让所有参与者突破各自的有限理性。如果每个人都能为了同一个目标而和谐地相处，其结果将令人惊奇。对此，人们最为熟悉的例子就是战时的经济动员，或者战后或灾后的重建。

关于瑞典的人口政策，在20世纪30年代，瑞典的人口出生率陡然下跌，而瑞典政府和罗马尼亚与匈牙利一样，也对此颇感忧虑。不同于罗马尼亚和匈牙利，瑞典政府评估了自己的目标以及国民的目标，认为双方有一个基本的共识，那就是关键不在于家庭人口的数量，而在于育儿的质量。每个孩子都应该是被渴望、珍视的，都应该有条件接受优良的教育和健康保健，没

有孩子只存在物质需求。这些就是政府和国民共同的目标，可以把各方协调起来。

最后，瑞典政府出台了一系列政策，而它们看起来与当时的低出生率是格格不入的，因为它包括人们可以自由地避孕和流产——这符合“每个孩子都应该是被渴望、珍视的”原则。这一政策也包括广泛地进行性教育、不过分限制离婚、免费的产科护理、对有需要的家庭给予支持，以及大大增加教育与保健投资等。自从该政策出台之后，瑞典的人口出生率上升和下降数次，都没有给国民和政府造成大的恐慌，因为政府一直关注的是比人口数量更为重要的一个目标。

有时候，并不能在系统中找到一个和谐的总体目标，但这是值得人们努力尝试的一个方向。只有放弃一些狭隘的目标，考虑整个系统更为长期的福利，才有可能找到这一目标。

系统多棱镜

陷阱1：政策阻力

当系统中多个参与者有不同的目标，从而将系统存量往不同方向拉时，结果就是“政策阻力”。任何新政策，尤其是当它恰好管用时，都会让存量远离其他参与者的目标，因而会产生额外的抵抗，其结果是大家都不愿意看到的，但每个人都要付出相当的努力去维持它。

对策

放弃压制或实现单方面的目标。化阻力为动力，将所有参与者召集起来，用先前用于维持“政策阻力”的精力，去寻找如何实现所有人的目标，实现“皆大欢喜”，或者重新定义一个更大、更重要的总体目标，让大家愿意齐心协力去实现它。

公地悲剧

《国际先驱导报》，1992

经过数月的争吵，科尔总理领导的基督教民主联盟上周与反对派社会民主党达成了协议，收紧对避难申请的审批，以阻止大量经济移民的涌入。

对于人们共同分享的、有限的资源，很容易出现开发（或消耗）逐步升级或增长的态势。这时，就容易陷入“公地悲剧”陷阱。

1968年，生态学家加勒特·哈丁发表了一篇堪称经典的论文，对这一常见的系统结构进行了论述。哈丁在文章开头以一个普通的草场为例，指出：

试想一下，有这么一块草场，对所有牧民免费开放。显然，每一个牧民都会尽力扩大自己的畜牧规模……自觉或不自觉地，明确或含蓄地，他们都会问：“如果我再增多一头牲畜，效益如何？”

由于牧民全部享有额外增加的牲畜的收入，正面的效益接近+1。然而，如果每个人都这么

做，将导致过度放牧效应，但这个结果会由大家共同来承担。因此，对于单个牧民而言，负面影响只是若干分之一。

由此，每位理性的牧民都能得出结论，个人能采取的唯一明智的行动就是不断扩大牧群的规模。但是，这是每个人都能想到的，大家都有这样的共识。因此，悲剧就此酿成……每一个牧民都被锁定在系统中，迫使他们无节制地增加牲畜数量。然而，资源毕竟是有限的。每个人拼命努力追求自己的最大利益，最终的命运就是集体毁灭。

在有限的环境中，有限理性的结果就是这样。

在任何一个系统中，通常都有一些共享的资源（如草场）。对于那些容易出现“公地悲剧”的系统来说，共享的资源不仅是有限的，而且在过度使用时会出现严重的侵蚀和衰竭。也就是说，超过一定限度之后，资源越少，其自我再生的能力也越差，或者更可能的情况是被彻底破坏。以草场为例，草越少，牲畜就越容易把草连根吃掉；由于没了草根，土壤就容易被雨水冲刷走，变得更加贫瘠，草也就更难生长了。如此往复，这是另外一个恶性循环的增强回路。

任何一个系统也离不开资源的使用者（如奶牛和牧民），他们有很强的增长动力，且增长速度不被系统的状况所影响。拿单个牧民来说，没有理由、动机和强烈的反馈，使其不再扩大自己的牧群规模，从而防止过度放牧。相反，他们会尽最大努力去争取自己的利益。

对于那些充满期待的准移民来说，除了一心指望着从德国政府宽松的移民政策中受益，他们也没有理由去考虑这一现实，即太多移民不可避免地会迫使德国政府收紧这一法律。事实上，当得知德国政府正在讨论或有可能出台更为严格的移民法律的消息时，更多的人行动了起来，争取赶上“末班车”。

使用者越多，就会使用更多资源，导致每一个使用者可用的资源减少。如果每个使用者都遵循有限理性原则（“没有理由只让我一个人不扩大牧群规模”），就没有理由让任何人减少对资源的使用。就这样，对资源的使用速度将超过资源能够承载的限度。因为使用者得不到这方面的反馈，过度使用资源的状况还在持续，资源也日渐枯竭。最后，加速枯竭的回路开始启动，资源被破坏殆尽，而所有的使用者都将颗粒无收。

你可能会想，不会有人这么短视，以至于大家最后同归于尽吧？但是，不幸的是，这样的状况在历史上比比皆是，以下只是几个很常见的范例：

- 对某个景色秀丽的国家公园来说，如果对游客数量不加限制，将很快人满为患，自然的美景被破坏殆尽；

- 使用不可再生的化石燃料对每个人来说都有直接的好处，尽管由此会增加二氧化碳排放，产生“温室效应”，引发全球气候变化，但人们仍一如既往；

- 如果每个家庭想要几个孩子就要几个，却由整个社会负担所有儿童的教育、医疗保健和环境保护费用，那么新生儿的数量将很快超过社会能够负担的水平（正是这个例子使哈丁写了那篇著名的论文）。

这些例子都离不开对可再生性资源的过度使用，这是你已经在“系统大观园”一章中见过的结构。“公地悲剧”不仅可能出现在对公共资源的使用之中，也可能包括公用设施、公共排污场地等。如果一个家庭、公司或国家可以让整个社区来消化或处理其废物，例如将污染物未加处理就排放到河流的下游或下风向，那它们的成本就可以降低，利润得以增加，或增长得更快。这样做收益很大，而代价却很小，充其量只占污染结果的一部分，甚至完全没有代价。在这种情况下，就没有充分的理由让污染者停止排放污染。在这些案例中，限制公共资源使用速度的反馈回路力量都很弱。

系统之思

“公地悲剧”之所以产生，一个重要原因是资源的消耗与资源的使用者数量增长之间的反馈缺失了，或者时间延迟太长。

如果你认为上述解释很难理解，那么就问一问自己：你是否愿意为了减少空气污染而少开一天车？你是否愿意自己动手处理自己生产的垃圾？这一系统的结构，使得对整个社区和未来更负责任的行为对自己好处不大且成本高昂，反而不如自私行为更为便捷且有利。

防止“公地悲剧”有以下三种方式：

方式一：教育、劝诫。帮助人们看到无节制地使用公共资源的后果，号召并激发人们的美德品行。劝说人们有所节制，以社会舆论谴责或严厉惩罚来威慑违规者。

方式二：将公共资源私有化。将公共资源分割给个人，每个人都要对自己行为的结果负责。如果某些人缺乏自控力，对资源的使用超出了其所拥有的资源的承载能力，他们也只能自食其果，伤害不到其他人。

方式三：对公共资源进行管制。哈丁将这种选择称为“达成共识，强制执行”。管制可以采取很多种形式，从对某些行为的严格禁止，到

配额制、许可制、税收调控以及鼓励措施等。要想奏效，管制必须有强制性的监管和惩罚措施。

上面所讲的第一种解决方式试图通过道德压力，使参与者对共同资源的使用低于承载限度，从而保护资源免遭耗竭之虞；第二种方式（私有化）则旨在在资源的状况及其使用之间建立起直接的反馈联系，让同一个行动者对其行动的后果承担责任，不管是赢利，还是亏损，资源的占有者仍然可以过度使用资源，但他将为自己的无知或非理性而埋单；第三种方式（管制）通过监管者及使用者的互动，在行动和资源的状况之间建立了间接的联系，为了使这个反馈起作用，监管者必须有能力实施监管，并可以准确地把握公共资源的状况，同时也要有有效的威慑措施，并真心愿意维护整个社区的福利。

在一些“原始”文化中，人们通过教育和劝诫，世代有效地管理着公共资源。然而，哈丁并不相信这种方法总是有效。如果公共资源只靠传统或“信用”系统来保护，反而有可能让那些不尊重传统、不按常理出牌或不讲信用的人钻了空子，占了便宜，从而遭到破坏。

相对于教育与劝诫，私有化的方案更为有效。但是，很多资源不能简单地被分割、私有，例如大气和海洋，这只剩下一选择——“达成共识，强制执行”。

其实，在生活中到处都存在“达成共识，强制执行”的安排，大多数规则是如此常见，以至于人们往往习以为常、视而不见。每一种这样的安排既保护人们使用资源的自由，又同时限制人们无节制地使用公共资源。例如：

- 在繁忙的十字路口，通过交通信号灯对公共道路空间的使用进行管制。每个人都要遵守交通规则，不能随心所欲地想什么时候通过就什么时候过。然而，当绿灯亮起，轮到你通过时，你可以更安全、更顺畅地通行。如果没有信号灯，每个人都随心所欲，反而会乱成一锅粥。

- 在闹市区，使用计时收费设备来限制大家对于公共停车位的使用，不仅要收费，而且会限制你占用停车位的时间。你不能想停哪儿就停哪儿、想停多久就停多久，在这些方面你没有自由，但是，正是由于有了计时收费设备，当你有需要时，才有更高的概率来找到停车位。

- 对于银行里的钱，你不能随心所欲地“自助”服务，不管这对你来说多么划算。必要的保安设备和机制，例如保险柜、安全门、保安等，能够防止你把银行当做公共资源。相应地，你自己存放在银行里的钱也受到了保护。

- 你不能随意占用某些特定波长的波段，因为它们被用于电台和电视台的广播。要想建立电台进行广播，你必须得到相关监管部门的许可。如果不对此类行为加以限制，电波信号将彼

此覆盖，变得混乱不堪。

●很多城市里的垃圾处理系统运行成本高昂，千家万户不得不按照其生产的垃圾数量缴纳相应的垃圾处理费。这是把以前免费使用的公共资源转变为按使用量收费的监管系统。

你可能已经注意到了，在上述这些案例中，“达成共识，强制执行”的机制有很多种不同的表现形式：交通信号灯按照轮流放行的规则，让大家依次使用公共资源；车位计时收费设备采取的是按使用时间收费的规则；银行使用的则是物理的屏障加上严厉的惩罚；无线电广播频率的使用则需要向政府部门申请、得到审批；而垃圾处理收费的方法则修复了单个家庭对公共资源的使用及其影响后果之间的联系，让每个家庭直接从经济利益上得到收获。

大多数人在大多数时候都会遵守这些管制系统，只要大家能够达成共识，并且理解其目的。但是，所有管制系统都必须使用强制监管与治安权，对不合作者进行惩罚。

系统多棱镜

陷阱2：公地悲剧

当存在一种公共资源时，每个使用者都可以从这种资源的使用中直接获利，用得越多，收益也越大，但是过度使用的成本却需由所有人来分担。因此，资源的整体状况和单个参与者对资源的使用之间的反馈关联非常弱，结果导致资源的过度使用及耗竭，最终每个人都没有资源可用。

对策

对使用者进行教育和劝诫，让他们理解滥用资源的后果。同时，也可以恢复或增强资源的状况及其使用之间的弱反馈连接，有两类做法：一是将资源私有化，让每个使用者都可以直接感受到对自己那一份资源滥用的后果；二是对于那些无法分割和私有化的资源，则要对所有使用者进行监管。

目标侵蚀

埃里克·伊普森，《国际先驱导报》，1992

在本轮经济衰退中，英国人发现，经济就像一台装上了下行指令的发动机。这一全国性的灾难，现在更被视作进一步衰退的预兆。星期日的《独立报》（*Independent*）在头版发表了一篇文章称，“我们有一种不祥的预感，温莎失火就是整个国家的写照，它的根源在于国家的新政策失当……”

工党的贸易与工业发言人罗德·派斯顿（Lord Peston）坚称，“我们知道我们应该做什么，但由于某些原因，我们现在无法去做。”

政治家、企业家和经济学家们将原因归咎于，政府未能给青年人提供合格的教育，从而导

致劳动者和经理人的技能普遍不足，投资萎缩，经济政策不当。

一些系统不只是对试图改变它的政策措施具有阻力，竭力维持在一个大家谁都不愿意看到的不良状态，更为糟糕的是，它们还在持续恶化。这是一种被称为“目标侵蚀”^[3]的基模。这类例子包括企业的市场份额逐渐下滑，医院的服务品质不断下降，河水或空气质量持续恶化，虽然定期节食但体重依然在增长，美国公立学校的状况日渐衰落，包括我自己一度坚持的慢跑也越来越懈怠了。

通常，反馈回路中的主体（如上述例子中提到的英国政府、企业、医院、体重超标的人、校长以及慢跑者等）会有一个绩效目标，或期望的系统状态。如果系统的实际状态与目标或期望相比存在差距，主体就会采取行动。因此，这是一个常见的调节回路，应该会使绩效表现达到并保持在期望的水平上。

但是，在这些系统中，主体感知到的系统状态与系统的实际状态并不相同。一般而言，主体对坏消息更加敏感，倾向于更加关注并相信坏消息，而非好消息。当实际绩效有变化时，最好的结果会被当做偏差忽略掉，而最坏的结果却被记住了。这样，主体感知到的状况会比实际状况更为糟糕一些。

在这个系统基模中，还有一个重要特点：期望的系统状态会受到感知到的状态的影响。也就是说，标准也不是绝对的。当感知到的绩效水平下滑了，目标也可以相应地下调。因此，你可能经常在生活中听到类似这样的话：“好吧，看来我们只能如此了”、“唉，我们至少不能比去年做得更差吧”、“唉，看看周围，每个人都麻烦不断”。

按照调节回路的作用，应该使系统状态维持在一个可接受的水平上，但是它却被一个具有向下趋势的增强回路所掩盖。也就是说，感知到的系统状态越差，期望就越低；期望越低，与现状的差距就越小，从而采取更少的修正行为；而修正行为越少，系统的状态也就越差。如果任由这一回路运转下去，将导致系统的绩效不断降低。

这一系统性陷阱也被称为“温水煮青蛙效应”。它源自一个古老的寓言故事，说的是如果把一只青蛙突然放到热水中，它会立刻跳出来，但是，如果把它放到凉水中，但逐渐加温，青蛙就会舒服地待在那儿不动，直到被煮死。也许青蛙会想：“这里好像变热了一些。但是与刚才相比，也不是那么热，没有什么大不了的吧。”

由此可知，目标侵蚀是一个渐进的过程。如果系统状态快速改变，通常会引发明显的修正过程。但是，如果系统状态是逐步下降的，变化速度非常缓慢，不容易引起人们的注意，或使参与者产生了麻痹心理，忘记了系统之前一直保持的良好状态，每个人都似乎不知不觉、顺理成章地将期望值越降越低，努力越来越少，实际的绩效表现也就越来越差。

对于目标侵蚀，有两个对策：一是不管绩效如何，都要保持一个绝对的标准；二是不断地将目标与过去的最佳标准相对照，而不是和最差的表现相比。如果感知到的绩效比较乐观，而不是相对悲观，或者将最佳结果作为标准，而视最差结果为临时性的挫折，那么同样的系统结构就能把系统状态拉向越来越好的绩效。下行的增强回路，即“破罐子破摔”的恶性循环，将逐渐变成一个向上的增强回路，“事情越好，我就越努力工作，把事情做得更好”。

如果我能把这一教训应用于我的慢跑，说不定到现在我都可以跑馬拉松了。

系统多棱镜

陷阱3：目标侵蚀

绩效标准受过去绩效的影响，尤其是当人们对过去的绩效评价偏负面，也就是过于关注坏消息时，将启动一个恶性循环，使得目标和系统的绩效水平不断下滑。

对策

保持一个绝对的绩效标准。更好的状况是，将绩效标准设定为过去的最佳水平，从而不断提高自己的目标，并以此激励自己，追求更高的绩效。系统结构没有变化，但由于运转方向不同，便能成为一个良性循环，做得越来越好。

竞争升级

克莱德·哈伯曼，《国际先驱导报》，1992

星期天，伊斯兰激进分子绑架了一名以色列士兵，并威胁杀死这名士兵，除非以色列军队迅速释放被囚禁的加沙地区主要穆斯林派别创始人。这次绑架，引发了一波激烈的暴力冲突……三名巴勒斯坦人受枪伤，一名以色列士兵在巡逻途中被一辆过往的车辆袭击致死。除此之外，加沙地区摩擦不断，示威者向以色列军警抛掷石块，而以色列士兵则向示威者开枪，发射橡皮子弹，至少导致20人受伤。

在本书前面的部分，我已经提到过一个竞争升级^[4]的例子，就是小孩子打架的系统。你打我一拳，我就踢你一脚；你用了一份劲儿，我就

还你两份……很快，两个人就会打得不可开交。

“以眼还眼，以牙还牙”是导致竞争升级局面出现的决策规则。竞争升级源自于一个增强回路，相互竞争的参与者都试图超越对方，占据上风。系统中参与各方的目标都不是绝对的（把室内的温度设定为18℃）而是相对的，取决于系统中其他参与者的状况。也就是说，不管室内温度是多少，我都要比另外一个房间的温度高一度。

就像其他系统陷阱一样，竞争升级也不一定就是一件坏事。如果大家竞相争夺的是一些符合人们预期的目标，例如研发出速度更快的计算机、治疗艾滋病的方法等，它就能加速整个系统的进步。但是，如果它驱动的是恶性对抗、暴力、争吵、噪音或愤怒，它就真的是一个很危险的陷阱了。竞争升级最常见而可怕的例子是军备竞赛、地区或种族之间不可调和的矛盾与冲突，经常酿成暴力事件。

每一个参与者期望的系统状态都是相对于其他参与者而言的，并试图超越对方，领先一步，连并驾齐驱都不行；而且每个参与者都有高估对方的敌意、夸大对方实力的倾向。

在冷战时期，美国和苏联竞相夸大对方的武器装备实力，使自己不断发展更多的军备。一方每增加一件武器，都会导致另一方的紧张，并试图超过对方。虽然每一方都指责对方搞军备竞赛，但从系统的角度看，正是他们自己使得对抗不断升级。自己生产的武器触发了一个“多米诺骨牌效应”，使得自己未来需要生产出更多的武器。这一系统耗费了数十亿美元，拖累了两个超级大国的经济，生产了令人难以想象的大规模杀伤性武器，对世界安全构成了巨大威胁。

竞争升级另一个常见的例子是，竞选中的负面宣传。为了取胜，一名候选人诽谤对手，对手就会反击，“以眼还眼，以牙还牙”，如此发展下去，每个竞选人都被搞得几乎一无是处、体无完肤，让选民无所适从，也玷污了整个民主进程。

此外，常见的例子还包括价格战，一方为了战胜竞争者报出了低价，另外一方就会出更低的价格，引发连锁反应，如此发展下去，每一方都会遭受损失，但没有任何一方愿意认输、让步。最终的结果，可能导致一方的破产或两败俱伤。

为了吸引消费者的注意力，相互竞争的两家公司在打广告时也会竞相出招，导致竞争升级。

例如，一家公司的招牌做得亮一些，推销声音大一些、更吸引人；另外一家则会把招牌做得更大、更亮，推销声音更响、方式更炫。第一家公司只好想办法再度超过对手。如此发展下

去，广告就变得怪招迭出，无孔不入，过分炫耀、吵闹，让消费者的感觉麻木，对每一则广告信息都没留下什么印象。

在鸡尾酒会上，一个人大声说话，就会导致其他人必须以更大的声音交谈，反过来又使得那个人必须以更响的声音说话，如此循环，使得整个聚会都吵闹不堪；同样，相互竞争导致豪华车的车身不断加长，摇滚乐队的恶俗程度不断增加……诸如此类，不胜枚举。

尽管有时候是破坏性的，竞争升级也可能是平静、礼貌的，涉及效率、精妙、质量等。但是，即使竞争升级是朝着好的方向发展，也可能有问题，因为它不容易停下来。

例如，在医疗设备方面的不断追求，使每一家医院都竞相引进更先进、功能更强大、昂贵的诊断设备，试图胜过其他医院，而最后的结果却使得医疗保健成本大大超出了人们的承受限度，一些设备的功能也过于复杂、没有实际价值；对道德品格的不断追求，也可能导致装腔作势或假仁假义的伪善；对艺术的不断追求，导致建筑风格从精美的巴洛克式发展到过分装饰的洛可可式，再发展下去就成了“画蛇添足”的低劣、媚俗之作；对环境友好生活方式的不断追求，可能发展成为刻板、极端的清教徒主义。

由于竞争升级的系统结构是一个增强回路，它是以指数级方式发展起来的，一旦超过某个限度，其使竞争激化的速度会超出绝大多数人的想象。如果不加制止或打破这一回路，这一过程通常的结果是竞争的一方被打到，甚至是两败俱伤。

应对竞争升级陷阱的一种方式，是单方面裁军，也就是一方主动地让步，降低其系统状态，从而引导竞争对手的状态也随之下降。但是，按照系统通常的逻辑或规则，这一选择几乎是不可想象的。但实际上，它确实奏效，只要一方真有这样的决心，并且他没有被竞争对手因其单方面的让步而击倒在地。

应对竞争升级系统的另一种更为优雅的方式是，谈判达成裁军协定。这是一种结构性的改变，是关系到系统设计的活动。它将引进一组新的平衡控制回路，使竞争能够被控制在一定程度内。例如，父母介入调解来制止孩子们的争吵；针对广告的大小、呈现方式等出台一些规定；在武装冲突地区部署维和部队等。对于竞争升级的系统，停战协议通常不容易达成，对于参与各方来说，也从来不会是皆大欢喜，但是相对于深陷其中而言，停战是更好的选择。

系统多棱镜

陷阱4：竞争升级

当系统中一个存量的状态是取决于另外一个存量的状态，并试图超过对方时，就构成了一个增强回路，使得系统陷入竞争升级的陷阱，表现为军备竞赛、财富攀比、口水仗、声音或暴力升级等现象。由于竞争升级以指数级形式变化，它能以非常令人惊异的速度导致竞争激化。

如果什么也不做，这一循环也不可能一直发展下去，最后的结果将是一方被击倒或两败俱伤。

对策

应对这一陷阱的最佳方式是避免陷入这一结构之中。如果已经深陷其中，一方可以选择单方面让步，从而切断增强回路的运作；或者双方进行协商，引入一些调节回路，对竞争进行一些限制。

富者愈富：竞争排斥

塞尔维亚·纳萨尔，《国际先驱导报》，1992

最富有的那些人，在所有纳税者中只占1%的比例，在避税方面具有相当大的灵活性.....他们可能现在已经拿到了奖金，而不计入下一年的收入（那时税率水平会提高），将股票期权变现.....并且想方设法将各种收入提前。

利用积累起来的财富、权力、特殊渠道或内部信息，可以创造出更多的财富、权力、渠道以及信息。这些都是另外一个被称为“富者愈富”的基模的例子。俗话说：“成王败寇。”在现实世界中，这一系统陷阱也比比皆是，竞争的赢家获得了有利条件，从而可以在未来获得更大的发展。从结构上看，这是一个增强回路，系统中的参与者会迅速被分化为两类：赢家和输家。前者的发展越来越好；后者则越来越差。

玩过“大富翁”游戏的人都知道“富者愈富”的系统。游戏一开始时，所有参与者都处于平等地位，当有人首先获得了财产，并在自家地盘上盖起了“旅馆”之后，就能从其他参与者身上获取“租金”，这样他们就可以有钱买更多的旅馆。就这样，你有更多的旅馆，就可以得到更多的旅馆。最后，当一个参与者买了所有的东西，或者其他参与者受挫而退出时，游戏就结束了。

在圣诞节期间，每家都会在家门外挂上精美靓丽的灯饰。如果有一年这个街区搞了一次评比，某一家因为其灯饰最美丽而获得了100美元的奖励，第二年这一家很可能在灯饰方面多支出100美元，从而使他家的灯饰更为漂亮；当这一家连续三年赢得了胜利，他们家的灯饰每年都比上一年更为靓丽时，这一比赛很可能就会被终止。

正如《圣经》所讲：“凡有的，还要给他，使他富足。”赢家赢得的越多，他所拥有的就越多，未来就有越大的可能获胜。如果这种竞赛发生在一个有限的环境中，比如赢家得到的都是从输家那里攫取的，那么输家逐渐就会破产、被迫离开或饿死。

在生态学领域，富者愈富是一个普遍存在、广为人知的概念，也被称为“竞争排斥法则”。按照这一法则，争夺完全相同的资源的两个不同

物种，不能共生于同一个生态小生境之中。因为这两个物种是不同的，为了生存，每一种都需要繁殖得更快，或者比对方更有效地使用资源。如果能做到这一点，一方将占有更多的资源，而这将使其繁殖得更多，并持续保持优势地位。逐渐地，它将彻底主导这一小生境，并将落败的一方完全清剿出境——这通常不是通过直接的对抗来完成的，而是通过占用了所有的资源，一点儿也没留给居于劣势的竞争对手。

对于这一陷阱，卡尔·马克思也曾有所涉及，并以此对资本主义经济进行了抨击。两家公司在同一个市场上竞争，与我们上面刚刚提到的两个物种在同一个生态小生境中竞争的行为模式将完全一致。如果某一家公司通过各种手段，例如提高效率、恰当的投资或者更好的技术、甚至是行贿等，获得了些许的优势，它就会获得更多的收入，从而有更强的实力用于扩大再生产、新技术研发、做广告或者更多地行贿等，这将启动一个资本积累的增强回路，从而使其生产能力与收入规模越来越大，与其他竞争对手的差距也越来越大。如果市场是有限的，也没有反垄断法律阻止其不断扩大，这家公司将占领全部市场。

一些人认为，苏联的解体动摇了马克思的理论，但是他对市场竞争的独特分析并未被否定。在任何存在同质化竞争的市场中，都存在或曾经存在这样的现象，即从系统上看，市场竞争的本质是消除市场竞争。正是由于存在富者愈富的增强回路，美国众多的汽车公司逐渐减少到三大巨头（由于存在反垄断法，所以不是一家）；在美国大多数主要城市，每个城市现在只剩下一家新闻类的报纸；在每一个市场化的经济体中，我们也可以发现，从长期来看，公司数量存在下降的趋势，而公司的规模有扩大的趋势。

富者愈富这一陷阱会持续加大贫富差距，从而使其具有很大的危害性。不仅富人会比穷人有更多的避税方式，而且：

- 在大多数社会中，贫穷家庭的儿童要么失学，无法接受教育，要么勉强到最差的学校接受最差的教育。由于缺乏必要的劳动技能，他们只能从事一些简单劳动，从而获得很低的报酬，使他们难以脱贫；

- 低收入人群和贫困家庭缺乏可抵押的资产，使其无法从大多数银行贷到款。因此，他们要么没有资本去投资、“以钱生钱”，要么只能无奈之下通过非法渠道借高利贷，即使贷款利息是合理的，也是穷人支付利息，富人从众多穷人那里敛走了利息；

- 在世界上很多国家或地区，土地都不是平均分配和拥有的。大多数农场主都是租用他人土地的佃户。他们必须把部分收成交给土地的拥有者，才能获得在那片土地上劳作的许可，并且永远买不起土地，而地主则可以用从佃农那里收取的租金来购买更多的土地。

这些只是众多反馈机制中的少数几种，它们使得收入、资产、教育和机会等分配持续失衡且不断加剧。穷人们只能买得起很少一部分资源（如食物、燃料、种子、肥料等），而且还得付出最高的价格，因为他们经常是无组织、缺乏话语权的，政府也只用很少的精力去关注他们的需求。新的创意和技术往往最后才光顾他们，而疾病和污染却最先影响他们；他们要去做各种危险、低报酬的工作，除此之外别无选择；他们的孩子生活在拥挤不堪、疾病横行的恶劣环境中，却无法接种疫苗、接受教育。

那么，我们怎样才能规避富者愈富的陷阱呢？

第一种方式是，多元化。一些物种和公司有时会采用这一策略，以规避互斥性竞争。一个物种通过学习或进化，去开发新的资源；一家公司也可能创造出一种新的产品或服务，不与现有竞争对手直接抗衡。尽管生态小生境和市场会呈现垄断的趋势，但他们也能衍生出多样化的分支，产生新的物种、新的市场等。一段时间之后，新的小生境或市场又会吸引来竞争者，再次使系统迈向互斥性竞争的态势。

因此，多样化也不能保证可以规避富者愈富的陷阱，尤其是当占据垄断地位的公司（或物种）有力量摧毁所有的衍生物，或者把他们买下来，或者剥夺他们赖以维持生存的资源。所以，多样化不能作为穷人的一个策略。

另一种方式是，通过植入一个反馈回路，避免任何一个竞争者完全控制，使富者愈富的反馈回路处于可控的状态。这就是“反垄断法”理论上应起的作用，当然在实践中有些也能部分地起到类似作用。然而，一些非常庞大的公司可以动用各种资源，去减弱反垄断法的管制。

规避“富者愈富”基模最明显的方式是，通过设定游戏规则或机制，定期“校正”竞技场，防止一家独大。在传统社会和游戏中，设计者们本能地在系统中设置了一些平衡或牵制各自优势的方式，使得游戏可以保持公平和有趣。例如，“大富翁”游戏一开始，每个参与者都是公平的，这样使大家都有机会赢，不管你之前的战绩如何；很多运动对于弱势参与者提供了一些额外支持；很多传统社会中也有一些本土化的仪式，类似于北美地区印第安人的“冬节^[5]”，富人拿出他们的很大一部分财产分送给穷人。

有很多机制可以打破富者愈富、贫者愈贫的循环，例如，在税法中

规定对富人苛以更多的税、社会慈善、公共福利、工会、全民基础性义务教育和医疗保健、征收遗产税等。在大多数工业社会中，都有一系列此类措施的组合，应对富者愈富的陷阱，以保证社会公平；而在一些传统社会中，通过类似“冬节”或其他施舍、馈赠的方式，实现社会财富的再分配，从而巩固馈赠者的社会地位。

这些均衡机制可能源自基本的道德，也可能来自实践中对富者愈富循环后果的理解，比如失败者可能因为破产、彻底失望而“揭竿而起”，造成破坏性的结局。

系统多棱镜

陷阱5：富者愈富

如果在系统中，竞争中的赢家会持续地强化其进一步获胜的手段，这就形成了一个增强回路。如果这一回路不受限制地运转下去，赢家最终会通吃，输家则被消灭。

对策

多元化，即允许在竞争中落败的一方可以退出，开启另外一场新的博弈；反垄断法，即严格限制赢家所占有的最大份额比例；修正竞赛规则，限制最强的一些参与者的优势，或对处于劣势的参与者给予一些特别关照，增强他们的竞争力（例如施舍、馈赠、税赋调节、转移支付等）；对获胜者给予多样化的奖励，避免他们在下一轮竞争中争夺同一有限的资源，或产生偏差。

转嫁负担：上瘾

比尔·克林顿，《国际先驱导报》，1992

你们可能也感觉到了，我们现在正处在一个惊人的螺旋式下降的通道之中。因为很多成本被不断地转嫁到工商企业身上，他们的负担不断加剧，使得更多的工商企业主无力给其雇员缴纳保险.....现在每个月，全美大约有10万人正在失去健康保险。

他们之中，很多人都符合联邦医疗补助计划的受益条件。但是，由于联邦政府的赤字，他们都被排除在政府救济之外，要么教育水平下降，要么无力支付抚养孩子的开支、从事其他投资，或者没有余钱去缴纳税款。

基思·里奇伯格，《国际先驱导报》，1992

据说，如果你想使一位索马里人发怒，只要拿走他的阿拉伯茶就行了.....阿拉伯茶是由一种常绿灌木新鲜的嫩枝和叶芽制成的，在药理学上，它和安非他命有关。

22岁的法拉·穆罕默德·艾哈迈迪（Abdukadr Mahmoud Farah），说他从15岁时就开始饮茶.....“原因是它能让我不去想当前的状况。当我喝茶时，我很快乐。我好像能做任何事情，我也不会感到疲倦。”

大多数人都能理解个人上瘾的状况，包括酒精、尼古丁、咖啡因、糖和海洛因等，但并非每个人都知道，在一些大型系统中也存在上瘾症状，并有各种不同的表现形式。例如，一些企业或行业对政府补贴的依赖，农民种地对肥料的依赖，西方国家经济体系对廉价石油的依赖，或者军火制造商对政府合同的依赖等。

这一陷阱有很多名称，包括上瘾、依赖性、将负担转嫁给干预者^[6]等。该系统的结构包括一个存量以及相关的流入量和流出量。这一存量可能是物理存在的，如农作物产量，也有可能是形而上的，如幸福的感受、自信等。同时，这一存量由一个参与者进行调节。该参与者有一个预设的目标，通过将其感知到的该存量的实际状态与预设目标相对照，决定要采取的行动。这本质上是一个调节回路，要么改变流入量，要么调节流出量。

举例来说，你是一个青年，生活在一个饥荒与战争连绵的国度，你的目标是提高自己的幸福感，以使你感到高兴、精力充沛、无所畏惧。与现实相比，你期望的状况和实际状况存在巨大差距，同时你也几乎没有什么办法可以缩小这一差距。但是，你可以做的一件事是：吸毒。毒品对于改变你的真实状况没有任何帮助——事实上，它很可能使你的状况变得更为糟糕。但是，毒品有助于改变你对现状的感知，而且速度很快，它会麻痹你的感官，使你产生幻觉，让自己感觉亢奋和勇敢。在这种情况下，你会怎么办？

类似地，假如你经营着一家业绩不佳的公司，要想改变公司的经营状况，费时费力而且实属不易；但是，你有办法获得政府的补贴，这样你就可以使公司运营下去，并维持自己体面的社会地位。或者，你是一位农夫，想提高农作物的产量，但土地很贫瘠，而要想改善土壤的品质，需要花费很长时间；但是，你可以大量施化肥，使庄稼长得更壮，当季的收成会不错。你会怎么办？

显然，很多人都会选择那些简便、容易、可以快速见效的干预措施。但是，这样做的后果也是可以预料的——你通过类似的干预措施创造的系统状态不会持久：幻象终会消失，你的生活环境依然严酷；政府补贴也会用完，你的公司依旧没有市场竞争力；化肥会被作物吸收或耗损，而土地依然贫瘠（甚至更加贫瘠）。

我们周围类似上瘾、依赖性或转嫁负担的范例还包括：

- 以前养老是由家庭来承担的，对于一些家庭来说，负担很重。于是，人们开始建设社会保障体系、老人社区和养老院。现在，大多数家庭不再有时间、空间、技巧或意愿去赡养老人。

- 过去，长途运输主要通过铁路，短途通勤依靠地铁和公共汽车；后来，政府通过建设高速公路，改变了这一状况。

- 以前，小孩子们都能进行心算或借助纸和笔进行计算，但随着计算器的普及，孩子们的计算能力大大下降。

- 人体能够通过自身的免疫系统，对一些疾病（如天花、肺结核和疟疾等）产生抗体，但是随着预防接种和抗生素等药物的使用，人体自身的免疫能力不断下降。

- 由于现代医疗技术和药物的发展，基本上改变了人们的主动保健意识和对健康生活方式的重视，而是把这种责任转嫁给医生和药物。

将负担转嫁给干预者有可能是一件好事情。有时候，人们也会有意而为之，因为会提高系统保持在期望状态的能力。例如，通过接种天花疫苗，可以完全实现对天花病毒的免疫；如果不这样做，很可能造成天花病毒的感染和流行。虽然这样做，部分人也能凭借自身免疫能力而存活，但毫无疑问，前者好于后者。因此，有些系统可能真的需要一个干预者。

但是，干预也可能变成一个系统陷阱。如果系统内部的自我修正反馈机制不足以维持系统的状态，此时，一个善意的、有能力的外部人员，看到系统中的纠结，挺身而出，承担了部分工作量。很快地，他就能帮助系统恢复到大家都希望的状态。于是，大家都来感谢或祝贺，干预者自己通常也会沾沾自喜。

紧接着，原来的问题又出现了，因为干预者只是承担了部分工作，而并未采取任何措施去消除问题的产生根源。于是，当干预者再次采用类似方案去掩饰问题时，期望的系统表象再次出现了，但这只是一个假象。在表象之下，问题的根源仍未被触及。这使得干预者需要更多地应用最初“介入”的方案。

如果这种干预削弱了系统原本维持其自身状态的能力，不论干预者是由于主动破坏造成的，还是无意忽略，陷阱就形成了。一旦系统的自我调节能力萎缩了，就需要更多的外部干预措施，才能达到期望的效果。这进一步削弱了系统自身的能力，如此循环.....

为什么每个人都会跌入陷阱？首先，干预者可能无法预见最初看起来迫切需要的少量介入措施将启动一连串连锁反应，使自己越陷越深。

随着系统对干预者的依赖日益增加，最终会耗尽干预者的能力。美国医疗保健系统正是在经历了这样一系列事件之后，苦苦挣扎，急需改革。

其次，善意提供帮助的个人或团队（干预者）可能未想到长期失控的可能性，以及系统能力将随着将负担转嫁给强有力的干预者而日益弱化。

如果干预措施是药物，就将变成药物上瘾。一旦陷入某项上瘾行为，你陷得越深，就会陷得更深。1934年创立于美国俄亥俄州的“嗜酒者互诫协会”对“上瘾”的定义之一，就是一再重复同一种愚蠢的行为，试图得到某些不同的结果。

在解决问题时，也容易产生“上瘾”的状况。这通常指的是针对问题的症状，寻找一个快速见效但不太彻底的解决方案，这妨碍或转移、推迟了人们花费更大的精力、采取更长期的措施去解决真正的问题。具有上瘾性的政策是阴险的，因为它们很好“兜售”，也很简单，容易让人上当。

举例来说，虫害威胁到了农作物的生长，怎么办？与其系统地审视并控制耕种方式、单一作物栽培模式、自然生态系统破坏等导致虫害爆发的诸多因素，不如喷洒杀虫剂简单快捷。虽然这样能快速消灭害虫，但也导致生态系统的更大破坏。这将导致未来昆虫更大规模的爆发，不得不需要使用更多的杀虫剂。

再比如，油价上涨了，怎么办？与其承认石油这样的不可再生性资源最终将枯竭、耗尽，想方设法提高燃油的利用率，或转向其他能源，不如控制价格。在上个世纪70年代世界石油危机爆发时，面对油价的暴涨，美国和苏联的第一反应都是试图控制石油价格。在这种情况下，我们可以假装什么事情也没有发生，继续使用石油，而这只能使资源枯竭的问题更加严重。当这一政策不再奏效时，就会爆发石油战争，或者去发现更多的石油。就像一个喝醉酒的人在屋子里到处搜寻，试图再找到另外一瓶酒一样，我们也在到处污染海洋、侵入最后的未开发的处女地，只为寻找到另外一块大油田。

系统之思

戒除上瘾，回到非上瘾的状态固然是值得肯定的，但更理想的状况是：提前预防，防止上瘾。

打破上瘾结构是痛苦的。这可能是戒除海洛因所产生的肉体上的痛苦，也可能是因减少石油消费导致物价上涨所引发的经济上的阵痛，或者昆虫侵害的后果以及由此带来的天然捕食者数量的恢复。戒除意味着拒绝采用那些快速见效但有明显副作用的措施，从根本上面对系统的真实状态（通常是逐渐恶化的），并采取那些能彻底解决问题的根本措施[7]。有时候，如果一步到位的话，可能产生剧烈的阵痛（因为对于人们来说，问题的症状本身也是棘手的）。因此，在有条件的情况下，戒除

可以逐步完成，首先采用一种非上瘾性政策，以最小的扰动使系统得以休养生息，恢复原有的自我调适能力。例如，针对药物上瘾者，提供团体支持，进行家庭隔离，以恢复其自我认知；提高汽车的单位使用成本，以降低对石油的消耗；采用混养或农作物轮种的方式，以减少农作物对昆虫攻击的脆弱性。然而，有时候可能没有逐步戒除上瘾的办法，只能采取“休克疗法”，快速戒掉坏习惯，并承受相应的痛苦。

要想防止陷入上瘾的状态，需要以适当的方式干预，即增强系统自身应对负担的能力。“帮助系统改进自身”这一观点，可能比接管并运作系统成本更低，也更容易，但一些自由派政治家却似乎无法理解。秘诀在于，不要以英雄式的接管开始，而是从提出一系列问题开始。这些问题包括：

- 为什么自然的纠正机制不奏效呢？
- 如何移除影响成功的各种障碍？
- 如何让推动成功的各种机制更为有效？

系统多棱镜

陷阱6：转嫁负担

当面对一个系统性问题时，如果采用的解决方案根本无助于解决潜在的根本问题，只是缓解（或掩饰）了问题的症状时，就会产生转嫁负担、依赖性和上瘾的状况。不管是麻痹个人感官的物质，还是把潜在麻烦隐藏起来的政策，人们选择的干预行动都不能解决真正的问题。

如果选择并实施的干预措施，导致系统原本的自我调适能力萎缩或受到侵蚀，就会引发一个破坏性的增强回路。系统自我调适能力越差，就需要越多的干预措施；而这会使得系统的自我调适变得更差，不得不更多地依赖外部干预者。

对策

应对这一陷阱最好的办法是提前预防，防止跌入陷阱。一定要意识到只是缓解症状或掩饰信号的政策或做法，都不能真正地解决问题。因此，要将关注点从短期的救济转移到长期的结构性重建上来。

如果你是干预者，要想办法恢复或增强系统自身解决问题的能力，然后自己择机抽身退出。在需要干预时，以这样的方式实施干预，是最睿智的。

如果你是患有依赖症的人，在摆脱干预措施之前，要以建立自己系统内在的能力作为基础。如果已经处于上瘾结构之中，要马上行动。你拖得越久，戒除的过程就越长、越艰难。

规避规则

《国际先驱导报》，1992

卡尔文：霍布斯，我有了一个计划。

霍布斯：是吗？

卡尔文：如果我从现在到圣诞节，每天自发地做十件好事，圣诞老人将会仁慈地对待我。我的人生将从此揭开新的一页。

霍布斯：好的，这是你的机会。苏茜正朝这里走过来。

卡尔文：或许我可以从明天开始，每天做20件。

只要哪里存在规则，哪里就存在“规避规则”（rule beating）的可能。规避规则意味着，采取一些迂回或变通措施，虽然在名义上遵守或不违反规则的条文要求，但在本质上规避了系统规则的原本意图。如果规避规则的行为导致系统产生严重的扭曲或不自然的行为，就是一个需要警惕的问题：一旦失去控制，系统将会具有强大的破坏性。

一些扭曲了自然界、经济体系、组织和人性的规避规则行为，可能是毁灭性的。以下是一些案例，有些很严重，有些稍微好一些：

- 一些政府部门、大学和公司经常在财政年度末突击消费，产生一些无意义的支出。这是因为，如果它们今年不把预算内的钱花完，明年就有可能被砍掉一些预算；

- 在20世纪70年代，美国佛蒙特州出台了一项土地利用法案，被称为《250法案》（Act 250）。该法案对面积为10英亩（约4万平方米）及以下的住宅开发设置了复杂的审批流程。现在，佛蒙特州有非常多仅比10英亩大一点儿的住宅项目；

- 为了减少粮食进口、扶持本地种粮农民，一些欧洲国家在20世纪60年代对饲料进口设置了限制。在制定进口限制措施的过程中，没有人想到木薯——一种富含淀粉的根状作物，也是一种很好的动物饲料。因此，木薯并未被列入限制目录。所以，进口商以前是从美国进口玉米，而自那之后变成了从亚洲进口木薯。

- 按照《美国濒危物种法案》（The U.S. Endangered Species Act）规定，禁止对濒危物种栖息地进行商业开发。于是，一些土地拥有者在发现自己的土地上出现了濒危物种时，就有意地对其进行猎杀或投毒，以便自己的土地可以被开发。

请注意，“规避规则”并非违背规则，但只是遵守了规则的“表象”。当附近有一辆警车时，司机会遵守限速规定；因为木薯没有被列入谷物饲料的范畴，所以进口木薯不算是进口饲料；当领地内有濒危物种时，可以在把它们杀死之后再开发……因此，这只是遵守了“法律上的条文”，而不符合立法的本义或实质。规避规则的行为提示我们，需要从

整个系统的角度，重新修订法律，弥补自身存在的一些纰漏。

规避规则通常是下级对上级制定的刻板的、有害的、不切实际或不明确的规则的反应。对于规避规则的行动，人们一般有两种应对方式：

第一种方式是，通过强化规则及其实施力度，试图扑灭、镇压规避规则的行动。但是，这通常会激起系统更大的变形。以这种方式应对，将使人们在陷阱中越陷越深。

另一种方式是，把规避规则看做有用的反馈，对规则进行修订、改善、废除，或给予更好的解释。这将是系统的机会之所在。更好地设计规则，意味着要尽可能预见到规则对各个子系统的影响，包括可能出现的各种规避规则行为，并调整系统结构，充分发挥系统自组织的能力，将其引导到符合系统整体福利的方向上来。

系统多棱镜

陷阱7：规避规则

“上有政策，下有对策”。任何规则都可能会有漏洞或例外情况，因而也存在规避规则的机会。也就是说，虽然一些行为在表面上遵守或未违背规则，但实质上却不符合规则的本意，甚至扭曲了系统。

对策

设计或重新设计规则，从规避规则的行为中获得创造性反馈，使其发挥积极的作用，实现规则的本来目的。

目标错位

《国际先驱导报》，1992

政府本周五正式承认了经济学家们几个月来一直在谈论的一则消息：日本未实现一年前设定的3.5%的经济增长目标。

1993年GNP增长率为3.5%，1990年是5.5%。自从本财政年度开始，经济始终处于停滞或萎缩.....

现在，经济增长预测已被大幅拉低，来自政界和工商界的压力很可能会迫使财政部采取经济刺激措施。

我在本书第1章中曾说过，影响系统行为最有力的方式之一就是，调整它的目的或目标。这是因为，目标设定了系统的方向，定义了需要采取校正措施的差距，并指示着调节回路运作的预期状态以及成败。如

果目标定义不当，不能测量应该被测量的东西，不能真实地反映系统的状态，那么系统就不可能产出期望的结果。就像传统神话故事中可以许的三个愿望一样，系统也有一个可怕的趋势，即只会产出你要求它们产出的结果。所以，一定要慎重地对待系统的目标。

如果期望的系统状态是国家安全，并将其定义为军费开支数额，那么系统的行为就是军备竞赛。这可能带来国家安全，也可能不能。事实上，如果军备开支过高，超出了合理、必要的限度，或者挤占了经济建设投资，国家安全形势甚至可能被削弱。

如果期望的系统状态是良好的教育，并以平均分摊到每位学生身上的教育经费为衡量标准，就能确保教育经费能用到每一位学生身上。如果以标准化的考试成绩作为教育质量的衡量标准，就会产生提高考试成绩的行为。当然，这些标准能否与良好的教育相互关联，是值得考虑的。

在印度推行计划生育的早期，该项目标准被定义为宫内避孕器放置的数量。于是，医生们为了完成任务，往往未经病人同意就进行放环。

在这些案例中，结果往往与人们的初衷相悖，其中最常见的一个错误就在于系统设计中定错了目标。也许在此类错误中，最为严重的就是以国民生产总值（GNP）作为衡量国家经济成就的指标。GNP反映的是一国经济当期生产的最终产品和服务的货币价值。从这一指标推出的那一刻起，对它的批评之声就几乎没有中断过：

国民生产总值不能反映我们子女的健康、他们的教育质量以及幸福感。它也不包括诗歌艺术之美、婚姻关系的牢固、公开辩论的智慧，或者政府公务员的廉正。它既不能测量我们的聪明才智，也无法测量我们的勇气、学习、同情心以及爱国之心。它只衡量当期的产出，而这与人生的意义毫无关系。

国民经济核算系统与国民经济的真实状况相去甚远，因为它不记录我们的家庭生活，只是表面的消费。

此外，GNP也将好的和坏的混在一起。如果有更多的交通事故，就会产生更多的医疗账单和维修费用，推高GNP。而且，它只计算被交易的产品和服务。如果每个父母都雇人来抚养其子女，GNP就会提高。它不反映分配的公平性。富人购置昂贵的第二套住房，对GNP的贡献要大于贫困家庭购买的廉价保障性住房。它衡量的是成果而不是成就感，衡量的是产品和消费的总额而不是效率。一盏新的节能灯，同样的亮度只需耗费1/8的电量，使用寿命也比传统灯泡长10倍，而推广它却会拉低GNP。

GNP衡量的是“吞吐量”，也就是每一年制造和购买的材料流量，而不是资本的存量，而房

屋、汽车、计算机、立体声音响等才是真正的财富和喜悦的来源。可能有人会争辩说，在最好的社会状态下，资本的存量会通过最低的流量来加以维持和利用，而不是最高的。

每个人都希望经济繁荣昌盛，但并不是每个人都希望GNP不断提高。然而，世界各国政府都将萎靡不振的GNP视为经济不振的信号，并会采取各种行动以推动GNP的增长。其中很多行动纯属浪费，比如刺激一些无效率的、无人想要的产品和服务的生产；另外一些甚至可能适得其反，带来一些副作用，比如短期内为了刺激经济发展而加速砍伐森林，往往会威胁到经济、社会及环境的长期发展。

如果用GNP来定义社会的目标，社会就会尽全力去创造GNP。除非你在目标中明确设定并定期衡量、报告有关福利、权益、正义或效率等指标，否则就无法产出它们。如果各个国家设定了不同的目标，如致力于以最低的生产量去实现最高的人均财富，或最低的婴儿死亡率、最大的政治自由、最清洁的环境、最小的贫富差距等，而不是尽力实现最高的人均GNP指标，那么，世界将会有很大的不同。

与规避规则恰恰相反，追求错误的目标、陶醉于错误的指标是系统扭曲的另一个特征。在规避规则的情况下，系统试图逃避某个不受欢迎或设计不当的规则，尽管在表面上仍遵守这一规则；而对于追求错误的目标，系统忠实地遵守着规则，并产出了特定的结果，但实际上这并非是每个人都真正需要的。如果你在某个地方发现，“因为这是规则”而发生了一些愚蠢的事情，你就面临着追求错误目标的问题；如果你发现，“因为有办法绕过规则”而发生了一些愚蠢的事情，你就面临着规避规则的问题。围绕同一个规则，这两种系统扭曲现象可以同时存在或发生。

系统多棱镜

陷阱8：目标错位

系统行为对于反馈回路的目标特别敏感。如果目标定义不准确或不完整，即使系统忠实地执行了所有运作规则，其产出的结果却不一定是人们真正想要的。

对策

恰当地设定目标及指标，以反映系统真正的福利。一定要特别小心，不要将努力与结果混淆，否则系统将只产出特定的努力，而不是你期望的结果。

系统寓言

帆船设计的目标

过去，人们进行帆船比赛，不是为了赢取百万美元大奖或国家荣誉，只是为了享受这项运动的快乐。

他们用自己已有的船只来进行比赛，这些船都有其日常的用途，比如钓鱼、货物运输或周末环游。

但人们发现，如果比赛船只能在速度和机动性上基本一致的话，比赛会更加有趣。于是，规则改进了，人们根据长度、受风面积以及其他一些参数，将船只划分为不同的等级，规定选手只能在同一级别上进行比赛。

很快，专门为比赛而设计不同类别的帆船就出现了。这些船尽可能地提高了单位帆面积的推进速度，减轻了标准尺寸船舵的重量。这些船往往奇形怪状、很难操控，根本不像是你想去海钓或周末出游时驾驶的那一类帆船。随着比赛变得更加严肃，规则也愈发苛刻，船只设计也更加怪异。

现在，比赛的帆船都非常快速、高度灵活，但是几乎不适于航海。它们需要专业的运动员和船员来管理。除了去比赛，没人会觉得那些美洲杯帆船赛上的赛艇有任何其他用处。围绕现有规则，那些船只只是被极度优化了，以至于失去了所有适应力。只要规则发生任何一点变化，它们都将变成一堆废物。

[1] 原文为“fixes that fail”，在《第五项修炼》一书中被译为“饮鸩止渴”，我们在本书中译为“治标不治本”，因为一些应对措施只是不能从根本上解决问题，未必使问题恶化。——译者注

[2] 原文为“policy resistance”，直译为对政策的阻力、抵抗或反弹，我们在本书中译为“政策阻力”，意味着系统自身结构具有内在的惯性，对外界施加的变革力量产生一定的阻力或对抗。这与我们常说的“上有政策，下有对策”或医学上的“耐药性”有相似之处。——译者注

[3] 原文是“drift to low performance”，也被称为“目标侵蚀”（eroding goals），本意是绩效表现节节下滑，日渐恶化，如同走下坡路，越来越差。因为在《第五项修炼》等书中普遍采用“目标侵蚀”的译法，为避免读者产生困惑，我们在这里也译为“目标侵蚀”。——译者注

[4] 原文为“escalation”，原意是逐步升级、扩大或加剧，也被译为“恶性竞争”。但是，正如作者所讲，这一基模并非总是“恶性”的或破坏性的，如果竞争的双方争夺的是一些符合人们预期的目标，它也可能引发好的结果。所以，我们在本书中将其译为“竞争升级”。——译者注

[5] “冬节（potlatch）”是北美地区印第安人冬季的一个节日，富人会拿出自己的财产或礼物馈赠给他人，也被称为“散财宴”。在中国某些地区，也有类似的施舍、馈赠仪式。——译者注

[6] 原文为“Shifting the Burden to the Intervenor”，直译为“将负担转嫁给干预者”，对于干预者来说，本来可能是好心，想帮忙，但是却“惹火烧身”，越陷越深，难以自拔；对于被帮助者来说，产生依赖性，越来越需要干预者的帮助或干预。在《第五项修炼》等书中，这一基模被译为“舍本逐末”或“饮鸩止渴”（二者在本质上是一致的）。——译者注

[7] 通常这些针对问题症状、可以快速见效但“治标不治本”的解决问题的方案，被称为“症状解”；能从根本上彻底解决问题的方案，被称为“根本解”。——译者注

第三部分 改变系统

第6章 系统之杠杆点：系统的12大变革方式

劳伦斯·马尔金

《国际先驱导报》，1992

IBM宣布增加裁员25000人，并大幅压缩研发开支，下一年的研发费用将减少10亿美元……。董事会主席约翰·阿克斯（John Akers）说，IBM在研发方面仍然是世界和行业的领导者，但还有改进的空间，尤其是“转向增长的领域”即服务业，这需要较少的资本，但从长期看利润回报也不高。

在了解了系统的上述机理之后，你可能会问：我们如何改变系统的结构，以便更多地产生我们所期望的结果，并尽量减少不期望的情况呢？对此，麻省理工学院的福瑞斯特教授基于多年来对企业中系统性问题的深入研究告诉我们：一般管理者都能很准确地定义当前的问题，识别这些问题产生的系统结构，并基本能够猜中到哪里去寻找“杠杆点”——也就是说，在系统中的某处施加一个小的变化，就能导致系统行为发生显著的转变。

“杠杆点”这一概念并非系统分析的独创，而是根植于人们的日常生活与很多神话故事之中，如“银弹”、“命门”、特效疗法、秘密通道、魔法口诀、扭转历史潮流的孤胆英雄、几乎不费力气地穿越或跨过巨大障碍的方法，等等。人们不仅愿意相信杠杆点的存在，也很想知道它们到底在哪里，以及如何运用它们。杠杆点就是权柄。

但是，福瑞斯特继续指出，虽然身处系统之中的人们通常凭直觉判断到哪里去寻找杠杆点，但多半是往错误的方向推动系统的变化。

在这方面最典型的一个例子，发生在我自己做世界模型系统分析时。应罗马俱乐部^[1]的邀请，福瑞斯特教授组织了一个研究小组，目的是探究一些全球性的重大问题，如贫穷和饥荒、环境破坏、资源枯竭、都市退化以及失业等，是如何相互关联以及发展演进的。利用系统动力学的方法，福瑞斯特教授开发了一个计算机模型，并找出了一个明确的杠杆点：增长。不只人口在增长，经济也在增长。在带来好处的同时，增长也有成本，特别是一些我们无法计算的成本，包括贫穷、饥荒、环

境破坏等。可以说，我们研究的清单上的所有问题，都与增长有关。而我們所需要的，是一种更慢一些的增长，是一种完全不同的增长模式，在有些情况下甚至是不增长或者负增长。

世界各国的领导人也都将经济增长作为解决各种问题的答案，但是他们尽全力推动的，却都是错误的方向。

福瑞斯特另外一部经典著作是于1969年出版的《城市动力学》（*Urban Dynamics*）。在该书中，福瑞斯特指出，低收入家庭住房补贴是一个杠杆点。补贴越少，城市状况越好——对于城市中的低收入人群也是如此。这一模型推出时，正值美国大规模推行低收入家庭住房补贴政策。福雷斯特的观点在当时受到了很多人的批评和嘲笑。但从那之后，很多推行了此政策的城市一个接一个地没落了。

拿福雷斯特的话来说，复杂系统的特征之一是“违反直觉的”，而寻找和撬动杠杆点也通常不能靠直觉。即便是直觉的，我们也经常选错方向，结果将试图解决的问题弄得更糟。

在我看来，对于动态、复杂的系统而言，要想找到杠杆点，没有快速或简单的公式，如果给我几个月或几年的时间，我可以发现杠杆点，但因为它们是违反直觉的，即使我真的发现了系统的杠杆点，也几乎没有人会相信我。这让人很受打击，也很痛苦。尤其是对于那些不只是希望了解复杂系统，还想做些事情让我们这个世界变得更美好的人来说，更是如此。

正是经历了多次这样的挫折，在一次有关全球贸易体制的会议上，我举出了几个可以干预系统的杠杆点。尽管还不完善，但我愿意在下面讨论这个清单，目的是抛砖引玉，希望能够帮助大家更好地找到系统的杠杆点。当然，这个清单还有很大的改进空间，你也可以进一步完善它。其实，那天我在脑海中浮现出来的几个泡泡，并非空穴来风，而是在我几十年对很多不同种类的系统进行严谨分析的基础上提炼出来的，其中也包含了很多其他智者的成果。但是，复杂的系统就是复杂的系统，对其进行一般化的归纳是十分危险的。所以，我想提醒读者朋友们注意：你下面所读到的内容不是“成品”或定论，而仍然是“在制品”；它们也不是寻找杠杆点的“处方”，相反，你可以把它们看做是一个邀请，我为你提供了一些参考或指引，并希望你能以更广阔的视角来思考系统的变化。

随着系统复杂度的增加，它们的行为也变得更加令人惊奇。想象一下你的活期存款账户，你可以存款和开支票，如果账户有足够大的余额，就会有一些利息流入，并且也将有一些银行费用流出；如果你的账户里没有钱，就会产生一笔债务，并会逐渐累积起来（债务也有利息）。现在，让你的账户与数千个账户有关联，并且允许银行将所有这些账户的结余存款作为贷款贷出去；更进一步，上千家这样的银行连接起来，形成了美国联邦储备系统。你将会发现，一些简单的存量和流量，相互连接起来，将创造出一个非常庞大、复杂且动态变化的系统，其复杂程度超乎想象。

这就是为什么杠杆点经常是违反直觉的。此外，也有很多的系统理论支撑并进一步深化这些可能的杠杆点。

12.数字：包括各种常数和参数

想象一下我们在第1章中提到的浴缸，那是一个基本的存量—流量系统。在那个系统中，存量和流量的大小及其变化快慢都可以用数值来表示：如果水龙头的阀门转动不灵活，得花很长一段时间才能把浴缸注满或关掉水龙头；如果下水道堵塞了，流出量的数值就很小。类似这一类参数，其中一些是固定不变或者被物理锁定的，但很多是可变的。这样一来，就有很多可调控的点。

再考虑一下国家的债务体系。它看起来像是一个很奇怪的存量，是一个漏钱的“大窟窿”。这个“大窟窿”的深度被称为年度财政赤字。政府的税收收入会填补、缩小这个窟窿，而政府的各种开支、花费又会加深这个窟窿。国会和总统将其大部分时间花在讨论如何增加（政府开支）或减小（税收）这个窟窿的深度或尺寸的很多参数上。由于这些流量也与选民相关，因此这些参数也会成为政治的筹码。但是，不管这些争论有多激烈，也不管是哪个党派在执政，近年来，这个漏钱的“大窟窿”都一直在扩大，区别只是速度快慢而已。

为了控制我们每天呼吸的空氣的清洁度，政府设定了一些被称为“空气质量”的标准参数；为了保护现有的森林面积，政府设定了每年允许砍伐量等指标；为了维持利润水平，公司会调整诸如工资率或产品价格等参数。

每一年允许砍伐的森林面积、最低工资、政府在艾滋病研究或者隐

形轰炸机上投入的经费、银行从你的账户上收取的服务费等，所有这些都是参数，是对水龙头阀门的调节。此外，单位解雇或招聘新人，包括政府公务员，也是在调整参数。另一方面，也可以通过改变阀门转动的速度，对系统进行调节。但是，如果还是使用原有的旧阀门、往原有系统中注水，并根据同样的旧信息、目标和规则进行转动，系统行为将不会有大的变化。把票投给比尔·克林顿还是老布什，肯定有很多不同，但如果每一任总统都置身于同样一个政治系统之中，那么从本质上看并没有什么不同。在国家财政系统中，改变系统中金钱流动的方向，或许可以产生显著的变化，但这几乎难以实现，因为无论换成你还是我，都会把自己放到首先获益的位置上去。

在我看来，通过数值（尤其是流量的大小）来调节系统是效力最低的一种方式，无法改变系统基本的结构。只是调整一些细枝末节，就像对泰坦尼克号甲板上的座椅进行重新布置一样。然而，大多数人会把90%的注意力都集中在参数上，尽管这样并不会起到太大的杠杆效应。

这并不是说参数不重要。在短期内，或者对于那些与流量直接相关的个体而言，它们可能是很重要的。例如，对于受税收和最低工资等政策变量直接影响的人群，或者直接从事与此相关的工作的人，都是如此。但是，改变这些变量极少能改变国家经济系统的行为。如果系统是慢性滞涨，参数变化了几乎也不可能一下子暴涨；如果系统是剧烈波动的，参数变化一般也不会使其稳定下来；如果是增长失控，参数也无法使其减缓。

不管我们对竞选捐款做何种限制，都无法消除政治因素。美国联邦政府无论如何调整利率，也无法改变经济周期（在经济向好时，我们经常会忘掉它；而在经济陷入衰退时，我们又会一再感到震惊）。在实行了数十年世界上最严格的空气污染标准之后，洛杉矶的空气质量虽有所改善，但远非洁净。扩大警察开支，也没有消除犯罪现象。

由于下面我将会展示一些数字作为杠杆点的案例，在这里我想先提出一个郑重声明：只有当我们实在找不到其他11个因素时，才能把参数当做杠杆点。例如，通过利率或出生率，来控制增强回路的增长。如果参数是系统目标，那就另当别论了。这类关键数字绝非平凡之辈，大多数系统都会受到这些关键参数的影响。但不幸的是，很多系统都已经偏离了最初设计的状态。因此，关注这些数字并没有太大的价值。

以下是一个朋友通过网络和我分享的一个故事，能很好地说明这一点：

当我想把自己的房子租出去时，我花了很多的时间和精力，试图算出“公平”的租金到底是多少。

我试着考虑所有的变量，包括房客的相对收入、我自己的收入和现金流需求，日常维护费用以及资本开支、权益与抵押贷款的利息率，我为房屋所付出的劳动价值，等等。

结果，我彻底迷失了，根本找不到结果。于是，我找了一个理财顾问去咨询。她说：“在你看来，好像有一个精确的合理租金水平，在这个水平之上，房客就被剥削了，而在这个水平之下，你又被压榨了。但事实上，并不存在这样一条明确的线。在你和房客之间，是一个较大的灰色区域——在这个区域之内，你们都可以达成合理的交易。所以，别再自寻烦恼了，照常过你的日子去吧。”

11.缓冲器：比流量力量更大、更稳定的存量

想象一下，有一个巨大的浴缸，流入和流出速率较缓慢；同时，有一个小号的浴缸，但流入量、流出量很大。这就是湖泊和河流的区别。相对于湖泊而言，河流更容易出现泛滥成灾的情况，因为与河流相比，湖泊的存量相对于其流量更大，也更加稳定。在化学和其他一些学科领域，一个较大、较稳定的存量被称为“缓冲器”（buffer）。

缓冲器具有保持稳定的力量。因此，你会把钱存入银行，而不是把它放到自己钱包中（流量是日常生活费用开支）；同理，商家会保持一定的库存量，而不是等顾客买走了存货之后马上订购新的货；为了保护濒危物种，我们需要使该种群的数量大于最低限度的繁殖数量；由于美国东部没有更大的钙存量去中和酸性物质，所以东部的土壤相对于西部的土壤对酸雨更加敏感。

通过提高缓冲器的容量，我们通常可以使系统稳定下来。但是，如果缓冲器过大，系统也将变得缺乏弹性，它对于变化的反应速度将过于缓慢。同时，要建立、扩大或维护某些缓冲器的容量，也需花费巨大的时间和资金，例如建设水库或仓库等。为此，一些企业发明了“零库存”的“及时生产”模式。在这些企业看来，与耗费巨资维持固定的库存相比，偶尔的波动或缺货造成的损失并不是很大。此外，较小的库存也具有更大的灵活性，可以更快地响应需求的变化。

所以，有时候很神奇，改变缓冲器的大小会成为杠杆点。但是，缓冲器通常是物理实体，不太容易改变。例如，美国东部土壤对酸性物质的吸收能力就不是缓解酸雨危害的杠杆点；水库大坝的蓄水能力也是固定不变的。因此，在这个清单中，我并没有把缓冲器排得太靠前。

10.存量—流量结构：实体系统及其交叉节点

具有若干个存量和流量，并物理地连接起来的系统，其结构对于系统如何运作具有巨大的影响。以前，匈牙利的道路系统是以首都布达佩斯为中心，从该国一端到另一端的所有车辆都必须穿过布达佩斯市中心，这导致了严重的空气污染和交通拥堵。该国曾经采取过多项控制污染和治理拥堵的措施，包括安装并改造了大量的交通信号灯、出台限速政策等，但都没有什么效果。

造成上述系统行为的根本原因在于道路系统结构设计不良，因此，修正该系统的唯一方式就是重建道路系统。洛基山研究所的艾默利·洛温斯（Amory Lovins）带领他的团队，通过把弯曲的管道取直、把过窄的管道扩大等简单的措施，在节约能源方面取得了令人惊奇的效果。如果我们对美国所有建筑都做类似的节能改造，节约下来的能源足以让很多发电厂关门。

但是，实体系统的重建通常是最慢、也是最昂贵的改变系统的办法，有的存量—流量结构甚至可能是不可改变的。

美国人口“生育高峰”一开始给小学教育造成了压力，接着是中学、大学，然后是就业和住房；现在，社会又不得不考虑如何支撑大量人口的退休养老需求。对此，我们基本上束手无策，因为等5岁的人变成了6岁，64岁的人就变成了65岁，这是顺理成章且不可阻止的。同样，这也可用来解释，对臭氧层有破坏作用的氟利昂分子的生命周期，因为低效率的汽车大军需要耗费10~20年才能淘汰，远远超出了污染物被清除的速度。

物理结构对系统是至关重要的，但它们很少是杠杆点，因为改变物理结构通常不太容易而且见效慢。恰当的杠杆点，需要从一开始就被设计好。一旦实体的结构建立起来了，要想找到杠杆点，就需要理解系统的限制和瓶颈，在尽可能发挥它们的最大效率的同时，避免出现较大的波动或扩张，超出其承受能力。

9.时间延迟：系统对变化做出反应的速度

反馈回路中的时间延迟对系统行为有着显著影响，它们通常是造成振荡的主因。如果你试图去调节一个存量，使其达到你的目标，但是你接收到存量状态的信息存在延迟，你将会面临矫枉过正或未达目标的情况；如果信息是及时的，但你的反应速度存在延迟，情况依然如此。

例如，建设一座发电厂需要耗费数年时间，而其运行寿命可能会超过30年。这些时间延迟使得发电厂的建设数量不能很好地匹配快速增长的用电需求。即使有非常好的预测能力，也难免用电紧张（供不应求）和供过于求的反复振荡，几乎世界上所有国家的电力工业都曾经历过类似状况。

当存在长期的时间延迟时，系统将不能对短期的变化做出反应。这就是大规模的中央计划经济体系必然运作不佳的原因，比如苏联或通用汽车公司。

因为我们知道时间延迟很重要，所以只要愿意去找，我们就能发现它们。

从污染物被倾倒在地上，到其渗透、污染地下水之间，存在时间延迟；从某一个儿童出生，到其长大成人、准备生儿育女之间，存在时间延迟；从某项新技术首次实验成功，到其被市场接纳、广泛采用之间，存在时间延迟；从产品价格调整，到其供需不平衡之间，也存在时间延迟。

在居于主导地位的反馈回路中，若反馈过程中存在时间延迟，将对存量变化的速度产生重要影响。如果延迟时间太短，容易导致反应过度，风声鹤唳，草木皆兵，并因过分敏感而导致振荡被放大；相反，如果延迟时间太长，将导致反应迟钝，使振荡得以衰减或突然爆发，这取决于延迟的时间到底有多长。对于存在某个临界值或危险水平的系统来说，一旦超过了一定限度，过长的延迟将造成不可逆转的伤害，从而导致矫枉过正并崩溃。

系统之思

既然时间延迟无法消除，那么，放慢系统的增长速度，使得技术和价格可以与增长保持一致，将具有更大的杠杆作用。

我将时间延迟的长度列为一个高杠杆点，但事实上，时间延迟通常不是很容易改变的。很多事物的发展有其内在规律，该花多长时间就得花多少时间。你不可能一夜之间积累起一大笔资本，孩子也不可能在一夜之间长大，拔苗助长也无法加快庄稼的生长。但是，减缓变化的速度一般并不难做到，这样的话，虽然反馈中的延迟不可避免，也不至于造成很大的麻烦。这就是在我排列的杠杆点清单中，成长速度位于时间延迟之前的原因。

同样，这也可以解释为什么在福瑞斯特的世界模型中，放缓经济的增长速度与加速技术开发或放开市场价格相比，是一个更大的杠杆点。它们都是试图加快调整速度的措施，但是，现实世界的实体资本存量（例如工厂及设备、应用技术等）变化不会那么快，即使面对新的价格

或创意也是如此——其实，价格和创意本身也不会马上变化，更不要说全球文化。

当然，如果在系统中存在可变的时间延迟，那么，改变延迟时间就能取得显著效果。但是，你必须非常小心，确保自己正在往正确的方向改变。例如，在金融市场上，用力减少信息和资金转移的延迟时间，只能加剧金融风暴的出现。

8. 调节回路：试图修正外界影响的反馈力量

现在，我们开始把关注的焦点从系统的实体部分转移到信息和控制部分，在这些地方存在更多的杠杆点。

调节回路在系统中是普遍存在的。大自然具有极强的平衡和自我进化能力，人体也能精妙地调控自身，把一些重要的存量维持在安全边界内。上文所述的温度调节器系统就是一个典型的案例。它的目标是将存量（室内温度）维持在期望的水平上。任何一个调节回路都有如下三项要素：一个预定的目标（如设定的室温）、一套检测达到或偏离目标的监测设备（如温度计），以及一个反应机制（如加热器、空调器、风扇、水泵、水管、燃料等）。

一个复杂的系统内部通常都有不计其数的调节回路，因此具有较强的自我纠正能力，可以适应不同的状况和影响。在某一段时间里，其中一些调节回路可能是不起作用的，例如核电厂的应急冷却系统，或者人体通过流汗、打寒战来维持体温的机制等，但这些回路的存在对于系统的长期价值而言是至关重要的。

我们常犯的一个大错误是试图扔掉这些“应急”反应机制，因为它们不是经常使用，而且看起来成本不菲。短期来看，这样做似乎没有什么影响；但从长期来看，这样做会大大降低系统对不同状况的适应力。在这方面，最令人心痛的一个例子就是，我们人类对濒危物种栖息地的侵略；另一个例子是，我们很多人将更多时间用于工作、赚钱，而忽视了个人休息、娱乐、社交和思考的时间。

调节回路具有将相关联的存量保持在预定目标值附近的能力。这主要取决于该回路上所有参数和连接的组合状况，它们决定了调节回路监测的准确性与速度、反应的灵敏度和力度、校正流量的直接性以及规

模。有时候，这些地方就存在着杠杆点。

正如很多经济学家相信，市场是一种调节型的反馈系统。它们的确具有神奇的自我调节能力，例如价格会随供求状况而波动，并对供求状况进行调节，使其保持相对的平衡。对于生产商和消费者来说，价格是市场的“晴雨表”，是核心的信息枢纽。价格机制越清晰、透明、及时和准确，市场的运作就越顺畅、有效。如果价格反映了全部的成本信息，消费者就能知道自己实际可以支付多少，而生产商也知道自己以一定的效率来生产可以赚到多少。毫无疑问，公司和政府都注意到了市场价格的杠杆效应，并且试图通过补贴、税赋以及其他方式对其进行调节，但是实际效果往往南辕北辙。

实际上，这些调节扭曲了市场信息，从而弱化了市场信号的反馈力度。在这里，真正的杠杆其实是顺其自然，不要干预，包括反垄断法、广告法、防治污染法、反不当补贴法，以及其他各种干预市场自由竞争的措施。

如今，一些强化和澄清市场信号的方法（如全成本核算等）不会走得太远，因为另外一些相关的调节回路被大大削弱了，包括民主政治体制。从它的设计原理看，已经在人民和政府之间设置了自我纠正的反馈机制。人民享有选举其代表的权力，并通过信息公开，了解他们选出的代表的所作所为，从而据此做出投票，决定让其继续履职或下台。这一过程的有效运作，有赖于选民和民选代表之间的信息流是自由、完整、准确的。为此，国家每年都要耗费数十亿美元，以防止信息的偏差。如果让那些试图扭曲市场价格信号的人有权力影响政府领导，允许其散播对其合作伙伴有利的信息，这些必不可少的调节回路将难以有效运作。市场和民主都将受到侵蚀。

系统之思

调节回路的力量需要与其预定要校正的影响大小相对应，这一点至关重要。可能的影响力量越大，调节回路的实力也需越强，否则就有可能无法发挥校正的作用。

在寒冷的冬日，人们通常都会紧闭门窗。只有这样，温度调节器系统才能正常运作。如果你把所有门窗都打开，与这种突如其来的变化相比，该系统的调节能力就不匹配了，室温会迅速下降，温度调节器系统就不起作用了。

再如，以前人们依靠传统的捕鱼设备，海洋渔业生态具有良好的调节能力，但是后来，人们发明了声呐定位、大功率拖船以及其他一些先进的设备，使得捕捞能力迅速增加。与这种变化相比，海洋系统的调节能力难以应对，海洋渔业资源迅速退化。

类似地，随着一些大企业的崛起，主导产业的控制力越来越强，需要政府加大检查和控制

的力度；经济全球化也使得跨国监管和全球治理显得越来越重要。

为了改善系统的自我校正能力，我们需要增强调节回路的力量，以下是一些实例：

- 通过预防医学、锻炼和良好的营养补给，来提高人体战胜疾病的能力；
- 推行综合治理，鼓励增加自然的捕食者来控制害虫；
- 通过信息公开法案，来限制政府的舞弊；
- 建立环境破坏监督报告系统；
- 保护举报者；
- 对私营企业征收排污费、污染税、履约保证金等。

7.增强回路：驱动收益增长的反馈力量

调节回路是自我修正的，而增强回路是自我强化的。它每运作一次，就能获得更大的强化自身运转方向的力量，使系统的行为朝原有方向越转越快。

例如，越多的人患上流感，就有更多的人被传染；新生儿越多，他们长大之后就会生育更多的孩子；你在银行里的存款越多，你得到的利息就越多，从而让你的钱更多；土壤流失越严重，地上的植被就越稀少，土壤就越容易板结并被雨水冲走，从而导致更多的流失；临界物质中的高能中子越多，它们与原子核的碰撞就越多，从而产生更多的高能中子，导致核爆炸或熔毁。

增强回路是系统中出现增长、爆发、腐蚀和崩溃的根源。如果系统中存在一个不受抑制的增强回路，该系统最终会被其摧毁。这就是它们比较罕见的缘故。通常，或早或迟都会激活一些调节回路，来抵消其影响。

对于流感来说，易感人群都感染了，或者人们采取了更有效的预防措施，流感的增强回路就受到了抑制；对于人口增长来说，随着人口增加，死亡率也会上升，与出生率持平，或者人们看到了不加限制地生育的后果而采取各种节育措施，使出生率降低，这样一来，人口的增强回路就会受到限制；对于土壤流失来说，所有土壤都被冲刷走了，露出了岩层，而数百万年之后，岩石也会粉碎变成新的土壤；或者人们停止了过度放牧，兴修水利、植树造林，土壤流失也会慢慢减少、停止。

在上述所有例子中，最初的结果是增强回路若不受约束可能出现的状况，后者则是在增强回路的基础上，增加了一些调节回路，对其进行干预，以减少增强回路自我放大的力量。

系统之思

一般来说，与增大调节回路的力量相比，减少增强回路的产出成果，也就是说放缓增长的速度，可能是更有力的杠杆点，而且其结果比听任增强回路不受约束地运作更好。

在世界模型中，人口和经济增长率都是杠杆点，因为只有减缓人口和经济的增长，才能让很多调节回路有时间发挥作用，包括技术、市场和其他调节措施，所有这些机制都有一定的局限性和时间延迟。同样，当你开快车时，降低行驶速度，是预防交通事故更为有效的杠杆点，而不能光指望刹车的响应速度或者其他先进的驾驶技术。

我们在前面讲过，在社会中，有很多增强回路使得资源向竞争中的赢家倾斜或集中，使得他们在其后的竞争中变得更为强大，这就是我们前文所说的“富者愈富”的陷阱。富人因有余款放贷而获得利息，穷人却要因匮乏资金而贷款并支付利息；富人可以雇用理财师和会计师，找到政治上的靠山，在扩大收入的同时减少税赋支出，而穷人却做不到；富人可以让其子女接受良好的教育，并给其丰厚的遗产，而穷人也做不到。与“富者愈富”强大的增强回路力量比起来，“反贫穷项目”的调节回路是软弱无力的，难以奏效。相反，如果能够弱化增强回路的力量，就可能有效得多。一些值得探讨的做法包括征收累进式收入税、遗产税，以及提高社会公共义务教育水平等。如果富人可以影响政府削弱而非强化这些措施，那么政府本身就从一个平衡调节性结构转变为一种强化“富者愈富”系统的力量，后果不堪设想！

当围绕出生率、利率、（土壤）流失率以及与“富者愈富”等增强回路有关的地方去寻找杠杆点时，越深入，通常收获也会越大。

6.信息流：谁能获得信息的结构

在第4章中，我们探讨了荷兰在住房开发过程中安装电表的故事，一些电表被安装在地下室，另外一些则被安装在前厅。在房间没有其他区别的情况下，后者家庭的电力消耗比前者低30%，差别仅仅在于电表是否被安装在更容易被人们看到的位置。

我很喜欢这个故事，因为它说明了系统的信息结构是一个高杠杆点。它不涉及参数的调整，也不是对现有反馈回路的强化或弱化。它是一个新的回路，让人们在之前得不到信息的地方可以获得反馈。

信息流的缺失是系统功能不良最常见的原因之一。增加或恢复信息可能是一个强有力的干预方法，而且通常比重建系统的物理基础设施更为容易、成本低廉。举例来说，摧毁全球渔业的“公地悲剧”之所以发生，就是因为在渔业资源状态（即海洋里还有多少鱼）和捕捞能力投资决策（即是否要投资建设大功率的渔船）之间，几乎没有反馈回路。与主流的经济观点不同，我并不认同鱼的价格可以提供上述反馈信息。相反，当鱼越稀缺时，鱼的价格就越高，也就让捕鱼人觉得更为有利可图，从而不辞辛苦地出海，把仅剩的一点儿鱼也捕捞殆尽。这是一个不恰当的反馈，是一个最终导致崩溃的增强回路。因此，我们真正需要的不是价格信息，而是剩余鱼群数量的信息。

系统之思

信息流的缺失是系统功能不良最常见的原因之一。增加或恢复信息可能是一个强有力的干预方法，而且通常比重建系统的物理基础设施更为容易、成本低廉。

在恰当的地点、以有效的方式恢复缺失的反馈是非常重要的。以另外一种“公地悲剧”抽采地下水为例，仅仅告诉每一个用水者“地下水位正在下降”的信息是不够的，甚至可能导致人们竞相把井打得更深。相反，把用水价格与地下水的蓄积量挂钩，也就是说当抽采速度超过地下水的补充速度时，迅速提高水价，可能更为有效。

其他一些强制性反馈的例子并不难找。想象一下，纳税人可以得到政府的详细信息公开，及时了解到自己缴纳的税款的用途；每一个城市或公司的取水管都放到同一条河的下游，让它们马上能够了解到自己排放出去的污水的情况；任何一个打算投资建设核电厂的政府公务员或企业老板，都得在自家的草坪上堆放核电厂排放的废物；每一位鼓吹宣战的政客都得自己上前线.....

从人性的角度上看，存在一种系统性的倾向，即人们避免对自己的决策承担责任。这就是为什么有如此多的反馈回路缺失了，也能说明为什么这类杠杆点受一般大众的欢迎，而在那些有权势的人那里不受欢迎。同时，如果你有了权力，可以这么做，我相信也是有效的。

5.系统规则：激励、惩罚和限制条件

规则定义了系统的范围、边界和自由度。例如，《圣经》上说“不可杀人”；宪法规定“每个人都有言论自由的权利”；总统的任期为4年，并且连任不超过两届；商务活动的基本规则是“合同具有法律效力”；棒球的规则是每队9个人，你需要触到每一个垒，三次击球失败就出局；如果你抢劫银行并被抓住，就要去坐牢。

宪法是社会规则最典型的例子。物理定律（如热力学第二定律）是一些绝对规则，也就是说不管我们是否理解或喜欢它们，它们都起作用。法律、处罚、激励以及非正式的社会约定都是一些人为规则，强度依次降低。

为了展示规则的力量，我会在课堂上让我的学生想象一下，如果改变大学的规则，会是一幅什么景象？假设让学生来评价老师，或者相互评价；假设没有学位，当你想学习一些新东西时，你就来上大学，学到了就离开；假设根据解决现实世界问题的能力来授予教授职位，而不是发表的学术论文数量；假设以小组为单位来评分，而不是个人.....规则一变，情况肯定大有不同。

当我们努力重塑规则，并且明白这些规则的变化如何影响人们的行为时，我们才能真正懂得规则的力量。这就是为什么当国会制订新的法律时，说客会云集于首都；为什么负责解释和制定宪法的特别法院，甚至比国会的权力还要大（宪法是制定规则的规则）。如果你想了解系统功能失调最深刻的原因，那就特别留意这些规则以及什么人拥有控制规则的权力。

当别人向我介绍了新的世界贸易体系之后，我的直觉系统使我警觉了起来，也正是基于这个原因。这一体系的规则由大公司制定，运作也由大公司把持，服务于大公司的利益。它的所有规则中几乎排除了来自社会其他方面的任何反馈。它的大部分会议也是封闭的，甚至对于媒体也是不公开的，没有信息流，没有反馈。它迫使各个国家进入一个追逐自身利益最大化的增强回路之中，相互竞争，并且弱化环境与社会的安全保障，以便能够吸引到公司的投资。这些“处方”将触发“富者愈富”的回路，产生巨大的权力集中和中央计划经济体系，并最终自己把自己摧毁。

4.自组织：系统结构增加、变化或进化的力量

有机系统和一些社会系统具有的最神奇的功能，是它们能够通过创造全新的结构和行为，彻底改变自身。在生物系统中，这种能力被称为“进化”；在人类社会、经济领域，则被称为技术进步或社会革命。用系统的语言来讲，这被称为“自组织”。

系统之思

自组织是系统具有最高适应力的表现形式。一个能够自我进化的系统，可以通过改变自身，来适应各种变化，以维持生存。

“自组织”意味着系统可以改变自身的任何一个方面，包括增加全新的物理结构（如大脑、翅膀或计算机等）、增加新的调节回路或增强回路，或者新的规则。

人类的免疫系统有能力对之前没有遇到过的各种伤害做出应对、人类的大脑可以捕获新的信息并突然冒出一些全新的想法，都是这个原理。

自组织的力量看起来非常神奇，以至于我们倾向于将其视为神秘的、不可思议或神化了的事物。比如在经济学领域，经济学家经常把技术当做魔法来对待，它们不知道是从哪儿冒出来的，也没有什么成本，但是每年可以稳定地对生产力的提高起一些作用。几个世纪以来，人们一直对大自然中令人叹为观止的多样性保持着同样的敬畏之心，也有很多人相信，只有一个神奇的造物主才能做到这一点。

对自组织系统进一步的研究表明，如果真的存在一个神奇的造物主，他也不能创造出进化的奇迹。不管是他、她或者是它，只不过是设定了非凡的自组织规则。这些规则在全局上左右着在什么情况下，系统应该在哪儿、如何、做哪些增减。

复杂的科学研究表明，通过计算机模拟，只是根据一组很简单的规则，就能够衍生出非常复杂、令人惊讶的模式。DNA携带的遗传代码是所有生物进化的基础，它也只是由4种不同的字母组成，每3个字母组合成不同的单词。这种模式及其复制和重组的规则，对于一些物种来说，保持了30亿年未变，而在这个漫长的演化过程中，涌现出了难以想象的自我进化的生物变种，有的很成功，有的则失败了。

要想实现进化，自组织机制需要一些原材料，以及一些实验、变异和测试，以便能够从海量、多变的信息存量中选择出可能的模式，并对各种新模式进行检测。对于生物进化来说，原材料是DNA，基因突变是多样性的来源之一，测试机制是环境变化、适者生存。对于技术演进来说，原材料是各种研究机构和科研人员，通过多种机制，例如图书馆、论文以及人的大脑等，创造、继承和累积、存储了大量的科学知识，多样化的来源是人的创造力，而测试机制可能是市场竞争，也可能是政府和基金会组织的赞助，或者符合人们的需求等。

当你理解了系统自组织的力量，你就更容易理解，为什么与经济学家推崇的技术相比，生物学家更加推崇生物多样性。经过亿万年的进化和积累，DNA的存量非常巨大，具有难以想象的广泛的多样性，这正是无限进化潜能的来源。相比之下，虽然科学图书馆、实验室、大学以及

研究人员等也有很多，一代代研究者不断被培养、从事研究，构成了人类技术潜能的来源，但这显然是无法和生物遗传基因的多样性相比的。当今时代，大量物种濒临灭绝，如果任由这种情况发展下去，从系统的角度看，这对于生物进化来说就是在犯罪。这就像把某个科学分支或领域的所有书籍、杂志全部销毁，或者让某一类科学家全部消失一样。

同样的状况也出现在人类文化上。经过数千年的积淀，人类文化也是多种多样的，储藏着各种行为指令的集合或保留剧目，它们也是人类社会演化的宝贵财富和源泉。不幸的是，人们并不那么珍惜或欣赏社会文化中所蕴含的宝贵的进化潜能。相比而言，人们对此的重视度甚至还比不上对地松鼠某一项基因变异的关注。在我看来，部分原因在于几乎每种社会文化都有一种信念，即认为只有我们的文化才是最优的。

强调单一的文化认同关闭了学习的大门，也削弱了人类社会的适应力。任何系统，包括生物、经济或社会，如果变得如此单一，也就会变得僵化、难以自我进化；如果某一个系统，在结构上鄙视、限制实验，不允许差异和多样性的存在，并消除了这一创新的原材料，从长期来看，将注定灭亡。要记住，我们生活的这个星球是高度多元化的。

在这里，介入点是很明显的，但通常并不受欢迎，那就是：鼓励多样性和实验。在很多人看来，多样性意味着“失控”——让一千种花朵自由绽放，任何事情都有可能发生！谁喜欢这样呢？还是让我们更稳妥一些吧，把杂草除掉，种上自己喜欢的花朵，这样会显得更为美观、整洁。但是，消除了多样性，无论对于生物系统，还是对社会、文化，或者市场而言，都将是一场灾难。

3.目标：系统的目的或功能

为什么系统的目标是一个杠杆点，而且要比系统的自组织能力更为优先考虑？这是因为，推动人们试图控制、消除多样性的动因，正是系统的目标。如果目标是将全世界都置于一个中央计划体系的控制之下，就像成吉思汗、基督教会、迪士尼或沃尔玛等构建的“帝国”一样，那么我们上面提到的所有可能的选择，例如物理存量和流量、反馈回路、信息流，甚至是自组织行为，都可以转化为实现这个目标上。

这就是我不参与关于“基因工程是好事，还是坏事”的争辩的原因。就像所有技术一样，是好是坏取决于谁来使用它以及要实现什么目标。

唯一可说的是，如果公司将其用于生产市场化的产品，那将是一个非常不同的目标，有不同的选择机制以及进化的方向，与我们这个地球上目前所见的所有事物都不同。

就像前文我们提到的单一循环的例子显示的那样，系统内大多数调节回路都有自己的目标——保持浴缸里的水位、保持适宜的室内温度、保持充足的库存量、保持足够的蓄水量。对于系统的组成部分而言，这些目标都是重要的杠杆点。大多数人其实都意识到了这一点，例如，如果你想让屋里更暖和一些，你会去调节空调的旋钮，将其设定到合适的温度。但是，对于那些更为复杂的系统整体而言，也存在更大的、但更不明显的高杠杆性目标。

即使系统里的人也经常认识不到他们所在的系统整体的目标。“为了赢利”，大多数公司都可能这么说，但那只是一个规则，是持续经营的必要条件。那么，这场游戏的目标到底是什么？为了成长、增加市场份额，还是将世界（包括顾客、供应商和管理者等）更大程度地置于自己的控制之下？如果是这样的话，公司的运作就需要逐渐扩大掌控范围，以规避不确定性。经济学家约翰·加尔布雷思（John Galbraith）很早之前就认识到了公司的目标是侵吞一切。其实，这也正是癌症的目标。实际上，这是每一个有机群体的目标。但是，因为总有一些更高层次的调节回路，去限制某个局部获得主导权，使之无法极力成长并最终控制整个系统，所以只有当这些回路无法奏效时，才会出现这样最差的情况。

例如，在市场体系中，维持市场公平竞争的目标，会超过单个公司消灭所有竞争对手的目标；在生态系统中，维持种群平衡和进化的目标，会超过单个个体无限制繁殖的目标。

我在前面说过，只要系统结构不变，改变系统中的参与者只是一个低层次的干预方式，除非那个参与者有权力制订规则，并可改变系统的目标。这种情况非常罕见，但效果却往往令人惊讶。组织中一位新领导者上台了，设定了一个新的目标，就把成百上千甚至数百万聪明而理智的人带往了新的方向，从达特茅斯学院到纳粹德国，都是如此。

这也是美国前总统罗纳德·里根所做的，我就曾亲眼目睹。在他入主白宫之前，有位著名的总统曾说过：“不要问政府能为你做什么，而是要问你能为政府做什么。”大家都能肃然接受这样的观念。然而，里根一再表示，目标不应该是让人民去帮助政府，也不是让政府去帮助人民，而是让政府减少干预。可能有人会说，为了更大的系统变革和公司权力的增长，他不得不出此下策。我也曾这么想过，但不可否认的是，自从里根上台之后，美国的公共政策乃至全球舆论，都被完全改变了。

所以，系统中的某个参与者可以清晰地设定、阐述、重复、支持并坚持新的目标，从而引导了系统的变革，里根就是这样一个例证。

2.社会范式：决定系统之所以为系统的心智模式

福瑞斯特关于系统的另一个著名论断是这样说的：一个国家的税法是如何制定的并不重要。这里面有一个社会公认的观念，即税赋的公平分配问题。不管税法上怎么说，公平或不公正，复杂或简单，关于欺骗、豁免或抵扣、起征点、实际的税款支付方式等，都会逐渐符合社会公认的“公平”观念。

这些社会公认的观念，一些潜在的基本假设以及关于社会现实本质的普遍看法，构成了社会的范式（paradigm），或者是一整套世界观，它们是人们普遍相信的、关于世界是如何运作的一系列基本假设、规则或信念。这些信念都是隐含的，因为在一个社会中，几乎每一个人都已经知道它们，因而不须特别申明。

例如，人们普遍认为可以用金钱来衡量某件事物的真实价值，金钱也有其真实的含义。因此，如果人们都不愿意多付钱，那就意味着这件东西不值钱。再如，人们普遍认为成长是件好事情；大自然是富含资源的宝藏，可以为人类服务；自然进化的进程被“智人”的出现终止了；个人可以“拥有”土地，等等。这些只是我们当前文化中很少一些有关社会范式的基本假设，而这些假设在另外一些文化中很可能是匪夷所思或惊世骇俗的。

对此，美国作家爱默生这样写道：

每一个国家、每一个人都会本能地以自我为中心，对其周边的物质存在做出反应，这准确无误地反映了他们的思想状态。为了了解这一点，你可以观察一下自己周围的每一个真理和错误、每一个人内心的想法，包括与你息息相关的社会、住房、城市、语言、各种各样的仪式、礼节、报纸杂志等；观察一下今天的各种想法；观察一下原木、砖块、石灰和石头如何被组合成方便人们居住的形状，以符合大多数人头脑中的主流想法……当然，它符合以下规律，即思想的轻微改变将被迅速放大，导致外部事物的显著变化。

系统之思

范式是系统之所以成为系统的根源。根植于这些范式，产生了系统的目标和信息流、反馈、各种存量和流量，以及系统中的所有东西。系统的目标、结构、规则、时间延迟和各种参数，都受范式的直接影响。

古埃及人建造了金字塔，因为他们相信在人死后还有来世。我们建造了摩天大楼，因为我们相信市中心的土地很值钱，寸土寸金。不管是哥白尼和开普勒发现地球并非宇宙的中心，还是爱因斯坦论证了物质和能量是可转换的，或者是亚当·斯密提出了“市场中单个主体的自私行为

汇集起来，却能令人惊奇地产生出集体的福利”，凡是在范式层面上采取干预措施或推动变革的人，都产生了巨大的杠杆效应，并从根本上改变了系统。

你可能会说，范式比系统中的其他东西都更难改变，因此这一项应该放到寻找杠杆点的清单中的最后一位，而不是倒数第二位。但是，在范式变革过程中，没有任何物理的变化，也无需昂贵的代价，甚至也没有时间延迟或是缓慢的过程。对于个人来说，这种转变可以在毫秒之间发生，所需要做的只是头脑中的一闪念、眼睛一眨或者采用一种新的观察方式。当然，对于整个社会来说，这将是另外一码事——人们抵制范式的转变，远甚于抵制任何其他变化。

所以，我们怎样才能改变范式呢？研究科学范式变革的著名学者托马斯·库恩（Thomas Kuhn）对此进行了深入的研究，他指出：

- 你需要持续地留意，在旧的范式中有哪些异常和失效之处；对于新的范式，你需要不停地宣讲和行动，并持之以恒。
- 在可能的情况下，加入新范式已被接纳并占主导地位的社会环境之中，切身体验新的范式。
- 不要与反对改革的人接触，浪费时间；相反，你要与积极的变革代理人合作，置身于心胸开阔、愿意接纳新事物的社会人群之中。

系统结构师告诉我们，可以通过建构系统的模型来改变范式。因为建模的过程会让我们跳出系统之外，使我们把系统视为一个整体。我个人曾经通过这种方式，改变过自己的范式。所以，我想告诉你，这种方法确实有效。

1.超越范式

与改变范式相比，在更高的层次上，还有另外一个杠杆点，那就是使自己摆脱任何范式的控制，保持灵活性，意识到没有范式才是“真实”的。每一个人在认识世界方面都有巨大的局限性，这个世界太大、太复杂，远远超出了我们人类的理解力和认知范围。即使那些能够持续不断地塑造和调整自己世界观的人，也无法对这个世界拥有完整的认识。

我们需要在自己的内心认识到各种范式的存在，并将这一点也视为一种范式，以赤子之心对待整体现实。这样才能进入一种“空”（not-knowing）的状态，进入佛教所讲的“开悟”境界。

我们每个人几乎都会受范式的控制，从而对每一种事物的多种可能性视而不见，偏执一端，想当然地认为只有自己的看法才是对的，其他看法都是错误或荒谬的，并迅速地反对其他看法，在自己认准的道路上越走越远。然而，在“空”的境界里，就没有权力与控制，没有决断，甚至没有存在的缘由，更没有无谓的作为，内心深处不执著于某一个信念。虽然这种境界看似玄妙，但事实上，每一个接纳该观点的人，不管是暂时的，还是长期的，都会发现这将是彻底放下的基础。如果没有任何范式或世界观是正确的，为了实现你的目标，你可以选择任何合适的手段。如果你对于如何达成目标没有任何想法，你可以仔细观察大自然的运作，倾听宇宙之声。

正是在这个超越范式的自主空间里，人们可以抛弃一切贪念，放弃控制、执著、封闭，打破一切统治、禁锢的枷锁，活在永恒的愉悦之中，即使被烧死在火刑柱上、被钉死十字架上，或者遭受其他迫害，但是其影响力仍然会持续上千年。

进入谦卑的“空”

由于系统是高度复杂、特异且动态变化的，要想有效地干预它，其实并没有一定之规。我在这里给出的清单也只是一家之言。它们是一些可能的选择，顺序也不是固定的。在每一项中，都有很多的例外，你也可以根据实际情况灵活调整它们的顺序。虽然这些想法多年来一直存在于我的潜意识里，但我也没有因此变成“超人”。杠杆点所处的层次越高，杠杆作用就越大，系统抵制变化的力量也就越强——这就是为什么社会总是会排斥或者消灭那些真正开悟的人。

但是，即使我们知道杠杆点在哪儿，以及可以往哪个方向去推动它，却也往往很难理解它的神奇效力，更难掌握和使用这些杠杆点。你必须努力思索，审慎地分析系统，并抛弃自己的范式，进入谦卑的“空”的境界。最后，看似无为，却可能是最为根本性的、战略性的、有效的杠杆点；看似疯狂，放下一切，却能优雅地与系统共舞。

[1] 罗马俱乐部是一家由企业家、政治家和科学家组成的国际团体，在20世纪70年代曾资助麻省理工学院系统动力学小组，使用系统动力学方法和计算机模拟技术，对世界人口和实体经济增长的原因及后果进行研究。该项研究成果——《增长的极限》一书出版后，受到全球广泛关注。——译者注

第7章 与系统共舞：系统的15大生存法则

G.K.切斯特顿

作家

在这个世界中，真正的难题不是判断世界本身是理性还是非理性。最常困扰我们的往往是，世界是基本理性的，但并非完全理性。生活是理性的，但如果完全按照逻辑来推论，却可能到处是陷阱。它看起来有一定量化关系，却不可能完全精确地用数学公式来度量；它看起来有一定规律，却又随时充满了惊喜。

在工业社会长大的人，若热衷于系统思考，很可能会犯一个严重的错误。他们可能会假定，通过系统分析，可以认清系统中的相互联系以及复杂纠葛，借助计算机的威力，最后找到预测和控制系统的钥匙。不幸的是，这是错误的观念，其根源在于工业时代根深蒂固的心智模式，即相信存在一把预测和控制的钥匙。

一开始，我也是这么认为的。我们所有这些在麻省理工学院就读系统专业的学生都是这么假定的。我们或多或少都天真地认为，通过学习，就可以练就一双洞察复杂世界的“慧眼”。大家都为此而着迷，努力钻研，如同许多前人所做的一样。但是，我们高估了自己的发现。这么做并不是想存心欺骗他人，只是表达了个人的意愿或期望。对我们来说，系统思考不仅是微妙、复杂的“头脑体操”，更是为了让系统高效地运作。

就像一位即将开始印度探秘之旅的探险者却误打误撞闯进了西半球一样，我们确实发现了一些东西，但它们根本就不是我们自认为的那样。系统思考与我们过去习以为常的观察、分析、认知世界的方式是迥然不同的。后来，随着对系统思考的理解日渐深入，我们慢慢发现，系统思考的价值比我们之前想象的还要大，只不过要做到这一点，必须使用与以往不同的方式。

我们认识到的第一个问题是，理解了如何修补一个系统和实际动手去修补它，完全是两码事。之前，我们曾就“实施”这一话题进行过很多次热烈的讨论，当时，我们的真实想法只不过是“如何让管理者、市长和相关机构的负责人接受我们的建议”。

然而实际情况是，就连我们自己也没有采纳自己的建议。我们理解了上瘾的结构，也多次给别人讲解，但我们自己却不能一天不喝咖啡；

我们都知道目标侵蚀的动态特性，但我们自己的慢跑又能坚持多久呢？我们一再警告别人不要陷入竞争升级和转嫁负担的陷阱，但我们在对待自己的家人时却又一再犯类似错误。

社会系统是人类文化思考模式的外在体现，也是深层次的需求、情绪、优势和劣势的反映。改变它们绝非易事，不是简单说一句“我们现在正面临严峻的挑战”，人们就能改变，也不是因为我们知道了改变的好处就能改变。

我们认识到的第二个问题是，系统洞察力会让你产生更多的问题。虽然对系统的洞察力让我们理解了很多之前所不曾理解的事情，但它们不能帮我们理解所有的事情。事实上，你所理解的事情越多，新出现的问题也就越多。就像人类所发明的其他透镜一样，系统思考也能让我们透视微观世界和宏观宇宙，发现很多神奇的事物，在过去这些可能都是一些难以理解的奇迹，但同时它也能让我们发现很多新的奇迹。借助这一项新工具，我们能够揭开很多根植于人类思维、心灵深处的神秘事物，更好地理解系统如何运作。下面，我列举了我们发现的几个问题：

系统行为

在某个特定系统的某个具体点上，增加一个新的信息反馈回路能使系统更好地运作。但是，决策者往往抵制他们所需要的信息：他们要么漠不关心，要么拒绝相信，或者根本不知道如何解释这些信息。

如果某个反馈回路只能围绕并导向某个特定的价值观，系统的结果将能让所有接受那个价值观的人感到满意，而让其他人不满意。我们不必改变任何人的价值观，但必须让系统围绕真正的价值观运作。

如果某个系统在各个参与方看来都不合常理，它将导致各种低效、丑恶现象、环境退化以及人类的痛楚。但是如果我们将其消灭，我们就将失去整个系统。旧的已破，新的难立。这将是最可怕的事。

系统中的人们会容忍一些有害的行动，因为他们害怕改变。他们不相信可以建立一个更好的系统，感觉自己无力去提出或实现任何系统化的改进。

系统行为引发的问题

为什么人们以自己的方式主动地对信息加以整理和过滤？他们是如何决定选择并处理哪些信息，而摒弃或贬低另外一些信息的？对于同样的信息，不同的人是如何做出不同的解读，并得出不同的结论的？

什么是价值观？它们来自何处？它们是通用的，还是由文化决定的，各有差异？是什么原因让一个人或社会放弃追求“真正的价值”，而满足于廉价的替代品？如何确定一个反馈回路是否为关键回路，尤其是如何从无法衡量的质量上来判断，而不只是数量？

为什么人们在最小的结构化和最大的自由度的情况下，创造力会如此惊人？一个人观察世界的方式如何能够被广泛共享，使得机构、技术、生产系统、建筑、城市等都建构于这种世界观之上？系统如何创造文化？文化如何创建系统？一旦文化和系统都被发现存在不足，它们是否能够通过瓦解和混沌而得以变革？

为什么人们会如此轻易地相信自己无能为力？他们为什么会变得这么愤世嫉俗，丧失实现自己愿景的能力？他们为什么更加愿意听别人告诉自己无法改变现状，而不相信他人告诉自己能做些什么呢？

系统思考者并非首批或唯一提出类似这些问题的人。当我们开始探究这些问题时，我们会发现各个学科领域、图书典籍、历史记录，也都曾问过同样的问题，并在某种程度上给出过答案。我们的研究的特别之处，并不在于我们的答案，甚至不是我们的问题，而是系统思考工具本身，它发端于工程学和数学，应用了计算机技术，受机械论思维模式的影响，追求对系统的预测和控制，这让它的实践者会面临人类最深刻的未解之谜。我坚信，即使对于最坚定的技术统治论者，系统思考也会让

我们发现，要想应对这个充满各种复杂系统的世界，需要的不只是技术统治。

自组织、非线性、反馈系统从本质上是不可预测和被控制的，因此，我们只能以最为一般的方式去理解它们。想要准确地预见未来并完美地提前做好准备，是不现实的；想要让复杂的系统只做符合我们期望的事情，也是不现实的。即使在最理想的情况下，也只能是暂时地实现。我们永远无法完整地理解这个世界，无法像还原论者所期望的那样彻底解构这个世界。科学本身，从量子理论到模糊数学，都会引导我们走入无法回避的不确定性之中。只要不是最为琐碎、具体的目标，我们就无法将其优化；我们甚至都不知道到底要优化什么；我们也不能跟踪每一件事情的发展、变化。虽然我们人类常把自己当做无所不知、无所不能的征服者，但无论是对于大自然，还是人与人之间，抑或是我们自己创造的各种组织，我们都尚未建立起一种恰当的、可持续的关系。

对于那些坚信自己是宇宙主宰的人，系统思考所揭示的不确定性是令他们难以接受的。如果你不能理解、预测和控制系统，那该怎么办呢？

系统之思

未来是不可预测的，但它可以被想象，并在人们的脑海中栩栩如生、呼之欲出；系统不可以被控制，但它们可以被设计和重构。

然而，只要我们认识到并愿意放弃控制的错觉，稍加等待，系统思考就能得出另外一种结论，如此鲜明耀眼，那就是：我们可以有很大的作为空间，但要换一种截然不同的方式。我们不能让风起云涌、变化万千的大千世界变得四平八稳、毫无意外，一切尽在掌握，但我们可以预料到各种意外，从中学习，甚至能从中获益；我们不能把自己的意志强加于系统之上，但我们可以聆听系统的声音，听它告诉了我们什么，并发现如何顺应系统的特性，使我们的价值观更好地与之匹配，从而创造出另外一些更好的事情来，而这都是无法只靠我们的意志来实现的。

我们无法控制系统，或将其搞清楚，但我们可以与系统共舞！

我已经知道这一点可以做到。通过漂流、园艺、演奏音乐和滑雪，我懂得了如何与大自然伟大的力量共舞。这些都需要一个人保持完全清醒、高度关注、竭尽全力，并对各种反馈做出快速响应。对于我来说，这些感受都是独一无二的，而在我从事研究、管理、政府公务和与人相

处时，从来没有出现过类似的情况。

但是，这并不意味着它们不需要同样的投入，从我们开发的每一个电脑模型中涌现出来的信息，也需要与系统共舞。要想成功地在这个系统的世界里生存，需要我们付出更多，不只是计算能力。我们需要奉献出全部的人性，包括理性分析、识别真理和谬误的能力、直觉、同理心、对未来的期许以及道德的力量，等等。

在本章中，我试图总结出一些最为通用的“系统智慧”，作为本书的结尾。这些都是我个人通过对复杂系统的建模以及与众建模者共同探讨而学习到的，是我亲身收获的一些经验教训，是植根于系统的基本原理之上提炼出来的一些概念和实践做法，虽然不尽完美，但大家可以在自己的工作和生活中应用它们、感受它们。它们是基于反馈、非线性和自组织等世界观之上的一些行为。当达特茅斯大学那位工程学教授注意到我们这些搞系统的家伙是“另类的”，并且疑惑为什么我们的看法与他有差异时，我想，这些可能就是他发现的不同之处。

我在这里列出的清单可能是不完整的，因为我依然是系统思考学派中的一名学生。同时，这也不是学习和应用系统思考的“独门秘籍”，有很多种方法可以学会与系统共舞。但是，大家可以把它作为学习与系统共舞的起点。我注意到，当我的同事们遇到一些新的系统时，他们会自然而然地这么做。

跟上系统的节拍

在你想以任何方式去干预系统之前，首先要观察它是如何运作的。如果它是一首乐曲、一段激流或者是某种大宗商品价格的波动，就要研究它的节拍；如果它是社会系统，就要观察它是如何运作的。研究它的历史，询问那些曾长期关注它的人们的意见，让他们告诉你曾经发生过什么。如果可能的话，去寻找或者制作一张图表，显示系统实际数据的时间变化态势——人们的记忆不是一直很可靠的，尤其是对于相对久远的事件。

这一方针看起来简单，实则不然。除非你能养成这样做的习惯，否则就会多走很多的弯路。从系统的行为开始，强迫你自己关注于事实，而不是各种理论。同时，这也有助于防止你快速陷进自己的信念、误解或其他类似误区之中。

在我们周围，其实有太多的误解，几乎到处都是。

比如说，很多人坚信年降水量在减少，但是当你查看年降水量的数据之后，你会发现，真实情况是降水量的波动性增加了，并非降水量绝对值的变化——干旱程度是加剧了，但洪涝灾害也更大了。再如，当一些权威专家宣称牛奶价格要走高之后，真实情况是牛奶价格持续走低；当专家告诉我汇率会下降时，其后的情况却是汇率一路走高；虽然人们一直预期财政赤字会下降，但赤字占GNP的比例却比以往任何时候都要大。

观察系统中的各种变量如何一起变动或者不一致地变动，是非常有趣的。但是，你要去观察真实发生的状况，而不是听人们对于发生的状况的解释，这样可以规避许多有意或无意的因果假设。例如，新罕布什尔州每一位行政委员看起来都支持减税以支持城镇的发展，因为很多人想当然地认为税率和增长率是有关联的，但是如果你把增长率和税率做成一张图，你会发现它们之间的分布是随机的，如同新罕布什尔州冬日的夜空繁星，完全没有内在的关联。

从系统的行为开始，也能把个人的思想引导到动态的分析上，而不是静态的研究——不只是问“哪儿出了问题”，也包括“到底是怎么弄成这样的？”、“是否有可能是另外的行为模式？”、“如果我们不改变方向，继续发展下去，事情最后会变成什么样？”同时，注意了解系统的力量，你可以问：“在系统里，什么运作得很顺利？”从几个变量的历史数据开始，以散点图的方式去揭示它们之间的关系，不仅可以发现系统中存在哪些因素，也可以了解它们之间是如何相互连接的。

最后，从历史数据开始，也能让人们根据系统的真实行为去定义问题，而不是用通用的或假想的趋势来干扰或分散注意力，掩饰因为“没有我们偏好的解决方案”而产生的不安感。倾听任何讨论，不管是在家庭中的对话，还是在公司会议上的意见分歧，或者是媒体上专家们的辩论，并且留意观察人们是如何得出解决方案的——在大多数情况下，都是“预测、控制”的模式，或者是把自己的意志强加于人，没有关注系统的状况及其原因。

把你的心智模式展现在阳光下

当我们画出系统结构图，并接着写出来各种变量之间关系的方程式，这样就强迫我们把自己内心隐藏的各种假设投射出来，并精准地表述它们。因为我们的模型需要保持完整、符合逻辑，并且前后一致，所以我们不得不把自己关于系统的每一个假设都摆出来，让其他人（也包括我们自己）能够看到它们。虽然心智模式是非常微妙的，但一旦要放

到模型中，我们的假设就不能再摇摆不定，在某一次讨论中这么说，到下次讨论时又那么说，前后矛盾，显然是行不通的。

系统之思

心智上的灵活性，是你在充满了各种灵活的系统中生存的必要条件；这包括愿意重新划定系统的边界、注意到系统转换到一种新模式，以及知道如何重新设计系统结构的能力等。

你不必非得以系统结构图或方程式的方式来表露自己的心智模式，尽管这样做是非常好的方法。你也可以通过语言、列表或者图片、箭头等方式，来表达你的想法，什么东西与另外一些东西存在关联，这些都包含着你的心智模式。不管以什么方式，你这样做得越多，你的想法就会变得越清晰，你承认各种不确定性并修正自己错误的速度就越快。这样，你就能学到更多，变得更为灵活。

请始终牢记，你所知道的每一件事，以及任何人知道的任何事，都只是一个模型。把你的模型拿出来，放到人们看得见的地方，再邀请其他人来挑战你的假设，并补充说明他们自己的假设。这样做不是为了比较、选出哪种假设、解释哪个模型是最好的，而是要尽可能多地收集各种可能的解释，并把它们都当做是合理的，除非你发现了一些证据，让你可以排除其中的一种或几种解释。在这个过程中，你也要保持警惕，因为其中可能包含着你自己的情绪或偏好，让你选择性地看到支持你的观点的证据，或者剔除掉了你自己不认同的一些假设。

把模型也拿出来，放到阳光之下，让它们尽可能地精确，用各种证据对其进行检验，如果没有得到证据的支撑，也应该勇于舍弃。这就是科学的态度和方法。但是，即使在科学研究领域，这样的做法也不常见，更不用说在社会科学、管理学、政府管理或日常生活中，就更为罕见了。

相信、尊重并分享信息

到现在，你应该已经了解了信息如何把系统结合到一起，并且知道信息的延迟、偏差、分散或缺失如何使得反馈回路的功能失调。很明显，如果决策者缺少信息，他们就无法做出应对；而如果信息不正确，反应也不可能正确；如果信息是滞后的，更不可能及时做出反应。在我看来，系统中的大多数错误，都是由于信息的偏差、延迟或缺失所致。

如果可以的话，我想向各位再多提醒一点：你不能歪曲、延迟或隐瞒信息。如果扰乱了系统的信息流，系统的运作就会陷入混乱或疯狂。相反，如果能使信息更及时、准确、完整，系统就会运做得更好，轻松而自在。

例如，1986年，美国联邦政府出台了新的《有害气体排放法案》（*The Toxic Release Inventory*），要求美国所有公司每年向政府报告它们所有工厂对外排放的有害气体总量。通过《信息自由法案》（*The Freedom of Information Act*），公众可以获得这些信息。从系统的观点看，后者堪称是整个美国最重要的一部法律。1988年7月，第一批有关化学排放的数据被公开。虽然报告了的排放就不算是非法的，但是把它们发表在企业年报上也不怎么好看，尤其是那些被列入“本地区十大排放单位”之列的排放大户压力更大。因此，虽然没有法律诉讼，没有强制性的减排指标，没有罚金，没有惩罚措施，但情况就发生了明显改善。在两年之内，全美国的化学排放量减少了40%（这只是从报告的数字来计算的，我们假定它也是符合事实的）。一些公司已经出台了相关的政策，计划将其排放量减少90%，这仅仅是因为之前被企业内部控制的信息被公开了。

从某种意义上讲，信息就是权力。任何对权力感兴趣的人都会很快认同这一观点。媒体、公众人物、政治家、广告发布者等对公众信息传播有一定影响力的机构，都有很大的权力，甚至超出了大多数人的想象。在很多时候，它们会为了自己的利益，在短期内对信息进行过滤，有选择性地发布信息，并对信息流进行引导。这可能就是我们所在的社会系统经常会变得失控的重要原因之一。

谨慎地使用语言，并用系统的概念去丰富语言

我们的信息流主要是由语言来组成的，而人们的心智模式也大多是通过词语来表达的。因此，尊重信息首先意味着避免语言污染——尽可能清晰、准确地使用语言；其次，要想办法扩展我们的语言，以便能够更有效地谈论复杂性。

弗雷德·考夫曼（Fred kofman）在一篇系统学期刊的文章上写道：

语言可以作为一种媒介，通过它，我们可以创造出新的理解和新的现实。事实上，我们不是在讨论我们所见的东西，我们只能看到我们能够讨论的东西。我们对世界的看法取决于我们的神经系统和语言的交互作用，这两者都是“过滤器”，影响着我们能看到的事物……语言和组织的信息系统都不是客观地描述外部存在的方式——它们从根本上塑造了其成员的感知和行动。要重塑（社会）系统的测量和沟通系统，就要在最根本的层面上重塑所有潜在的交互作用。与战略、组织结构或文化等比起来，语言作为现实的表述方式，是更为根本的。

如果在一个社会中，人们不停地在谈论“生产力”，而很少使用或很难理解“适应力”这个词，那么整个社会就将变得更有“生产力”，但“适应力”会降低。同样，如果人们不能理解或使用“承载能力”这个词，整

个社会将很快超过其承载能力；如果人们谈起“创造就业”时，就意味着必须要靠公司去做一些什么事，那么社会中的大多数人就没有紧迫感和动力，去为自己和他人创造出更多的工作机会。如果公司的使命是“创造利润”，就可能会轻视员工在价值创造过程中的角色。同样的情况还有“和平卫士”，这可能意味着必须拥有导弹，或者不可避免的“附带损害”，以及可能需要采取的“最终解决方案”或“种族清洗”，包括温德尔·贝里（Wendell Berry）所谈论的“暴政统治”。

让我感触很深的是，我们已经观察到，在过去的大约150年里，语言在日渐退化，要么是变得没有意义，要么是对其原本含义的破坏。同时，我也相信，伴随着语言的日渐退化，个人和社区的衰变、瓦解也会逐渐加剧.....

他接着说道：

随着表述的退化，语言已基本上丧失了标示作用，因为它未被认真地使用，已没有了特定的所指。人们的注意力被各种百分比、分类或抽象的函数瓜分.....用户很可能不再需要依赖或指望语言，因为它不能定义任何个人的立场或行动基础。它唯一实用的价值就是支持“专家意见”，大量不带个人色彩的技术行动已经开始了.....这将是一种残暴的语言：暴政统治。

系统之思

尊重语言的第一步是尽可能地保持它的真实、具体和有意义——这和保持信息流的清晰、准确是相同的。第二步是扩展语言，使其能反映我们对系统的理解。

为什么爱斯基摩人有那么多词语来形容雪，那是因为他们曾经深入研究和学习如何充分地利用雪。他们已经将雪变成了一种资源，当成了他们可与之共舞的系统。而我们所处的工业社会，只是刚刚开始拥有和使用系统的词汇，因为我们刚刚开始关注和利用复杂性。类似“承载能力”、“结构”、“多样性”这样的词汇，甚至是“系统”这个词，都是旧的词汇，但现在，它们的含义正变得越来越丰富，意义越来越精准。同时，我们也必须学会发明一些新的词汇。

电脑中的文字处理程序有拼写检查能力，这让我必须手工增加一些词语，因为它们原来并不包含在电脑的词典中，包括写作本书过程中常用的下列词语：反馈、吞吐量、次优化、自组织和可持续性。

关注重要的，而不只是容易衡量的

在我们的文化中，数字总是令人着迷的。这让人们自然地产生了一种想法，即那些可以测量的东西要比不能测量的更为重要。你若是不信，可以先想一分钟。这可能意味着，我们认为数量比质量更重要。如

果数量是某一个反馈回路的目标，那么它将成为我们关注的焦点，是我们的语言和体系的中心。无论是激励、评估，还是奖励，都离不开数量这个指标。这样的结果就是数量更为优先受到重视。你可以观察和考虑一下，在你周围的世界中，在自己的脑海里，到底是数量还是质量更为突出和重要？

作为系统建模者，我们曾经不止一次地被那些从事科学研究的同事们所嘲笑，因为我们无法将一些重要但无法衡量的变量纳入模型之中，诸如“偏见”、“自尊”或者“生活质量”等。由于计算机模拟需要数值，我们不得不编造出一些定量化的方法，以衡量那些定性的概念。比如说，我们假设偏见的取值范围位于-10到+10之间，0分意味着毫无偏见，-10分意味着完全负面的偏见，而+10分意味着完全正面的偏见，即你不可能做错事。现在，假设某人对你的偏见度是-2，或者+5，-8，这将对你的工作表现有什么影响？

实际上，在我们的工作中，真的有一次必须把偏见和绩效之间的关系放到模型之中。^[1]该项研究是一家公司想知道如何更好地对待少数民族裔员工，以及如何帮助这些员工走上公司的领导梯队。我们面试过的每一个人都认为，在偏见和绩效之间真的存在着某种联系。虽然采用哪种衡量方式是武断的，可以是1到5，也可以是0到100，但是如果不把“偏见”放到研究之中，肯定是不科学的。因此，我们试着把“偏见”放到模型之中，并努力去测量它。当公司里的员工被问及偏见和绩效之间的关系，他们提出来的几乎都是一种非线性关系，在此之前我从未在模型中见过此类关系。

如果某件事物难以量化，我们往往对其视而不见或者忽略它，这会导致模型不完善。在前面，你已经知道了系统常见的陷阱，其中包括围绕那么易于测量的东西来设定目标，而不是根据那些真正重要的东西。所以，不要再掉进此类陷阱。上帝赋予人类的不只是数数的能力，也包括评估品质的能力。希望你能成为一名质量检测员，到处走动，随时随地检查、确认质量是否达标。

举例来说，如果某件东西是丑陋的、俗气的、不相称的、比例失调、不可持续的，或者道德格调低下、危害环境、有损人格的，就不要让其通过。不要再让“如果你不能定义和测量它，你就不必关注它”这类的说教限制住你的手脚。没有任何一个人可以定义或测量正义、民主、安全、自由、真理，或者爱，也没有任何一个人能够定义或测量任何价值观。但是，人们却都离不开这些东西。如果我们不去谈论它们，不去检查它们是否存在或达标，在设计系统时也不去设法实现这些目标，它们将不复存在。这是不可想象的。

为反馈系统制定带有反馈功能的政策

吉米·卡特总统有一种非凡的能力，他可以用反馈的概念来思考系统性问题，并制定反馈的政策。不幸的是，因为国会和公众并不理解反馈，所以他在面对如何解释自己的想法、以说服国会和公众时，一度感到很困难。

例如，在石油进口量飙升的时期，他建议对美国国内消费的汽油征税。如果石油进口持续增加，汽油税也将进一步提高，直到能够抑制住对石油的需求，并迫使人们寻找石油的替代品，并减少进口。如果石油进口量降至零，汽油税也将降为零。

这一建议没有获得通过。

卡特也曾试图处理从墨西哥偷渡美国的非法移民潮。他认为，只要美国和墨西哥之间在发展机会和生活标准方面存在巨大的差距，就什么措施也阻止不了非法移民潮。与其在边境检查和保安方面投入大量的人力、物力、财力，不如把这笔钱作为给墨西哥的投资，帮助墨西哥发展经济。只要能够坚持下去，移民潮早晚会停止。

但是，这一建议也没有得到采纳。

很显然，对于动态的、自我调节的反馈系统，不能用静止的、刚性的政策来进行管制。虽然根据系统当下的状态设计出一项政策相对容易、快捷，代价通常也不大，但这很难行得通。相反，好的政策必须能够根据系统状态的变化及时地灵活调整。尤其是在面对复杂的系统、存在多重不确定性的情况下，最好的政策不仅要包括反馈回路，也要包括一种机制，对其中的各种反馈回路进行调整（可称之为“调整回路的回路”），适时地进行改变、纠正或扩大。从本质上看，这是把“学习”功能融入管理过程之中，使得政策具备了灵活性，从而能够更好地与系统共舞。

历史上签署的《蒙特利尔议定书》（*Montreal Protocol*），目的是保护地球大气层中的臭氧层。当1987年签署该议定书时，人们对于臭氧层破坏的速度、危害性以及各种化学物质对臭氧层的破坏作用等，都不太确定。因此，议定书设定的目标是，要减少对臭氧层最具破坏力的化学物质的生产和排放速度。但是，它也规定了一种机制，通过持续地监测臭氧层的变化状况，并根据臭氧层被破坏的实际状况与预期的差距，重新召开国际会议，调整逐步淘汰危险化学物质排放的时间表。结果三年后，也就是在1990年，该议定书被重新调整，加快了减排和淘汰的速度，并增加了更多的化学物质，因为臭氧层被破坏的速度以及危害性远大于人们在1987年预计的状况。

这就是一个带有反馈功能的政策的例子，它的结构之中具有学习的功能。我们希望它能及时地发挥作用。

追求整体利益

请记住，层级组织存在的目的是服务于最底层，而非最顶层。千万不能放大系统的某个部分或某个子系统的重要性，使其凌驾于系统整体之上，反而忘记了系统整体的存在。这是犯了典型的“一叶障目，不见泰山”的错误。同时，正如肯尼思·博尔丁所讲的，不要因为优化某件根本没必要做的事而招来更大的麻烦。因此，要瞄准那些能增强系统整体性能的要害，包括增长、稳定性、多样性、适应力以及可持续性，而不必在意它们是否容易衡量。

聆听系统的智慧

帮助并鼓励那些有助于系统自我运行的力量和结构。请留意，在这些力量和结构中，有多少是位于层级的底部的。不要成为一个粗鲁、莽撞、没脑子的干预者，破坏了系统内在的自我调节能力。在你介入之前，关注一下那些已经存在的价值是什么。

我的朋友内森·格雷（Nathan Gray）曾在危地马拉担任过援助人员。他告诉我，他对那些援助机构很失望，因为它们的意图是“创造就业”、“增加创业能力”以及“吸引外部投资者”。但实际运行起来时，它们却对当地市场视而不见——在那里，各行各业，从篮子制造、蔬菜种植、牲畜屠宰到糖果销售，大量中小企业和工商业者都正干得如火如荼，他们为自己创造了“就业岗位”，并在这些工作中展现出了高效的创业能力。内森花了很多时间与当地市场的工商业者交流，询问他们的生活和企业经营，从中学习如何才能帮助这些中小企业扩展业务规模、增加收入。最后他得出结论，他们真正需要的不是外部投资者，而是内部的融资和支持。合理的利率、易于获得的小额贷款，以及文化教育、会计服务等，将对整个社区产生更为长期的价值，这比从外面引进一家工厂或者生产线意义更为重大。

界定系统的职责

这是系统分析与设计的一项指导原则。对于系统分析，这意味着要搞清楚系统是怎样产生出它的各种行为的，包括有哪些触发事件和外部影响，引发了系统的哪些行为，经过了多少环节，谁在这些环节中起着什么样的作用，等等。有时候，这些外部事件是可控的（如减少饮用水中的病原体，以降低传染性疾病的发作概率），但有些是不可控的。如果只是责怪或试图控制外部影响，将容易使人们忽视系统内部的职责。

事实上，在系统内部，总有一些较简单的任务，可以增强自身的职责，从而更好地应对外部影响。

增强系统的“内在责任”，意味着在设计系统时，要在决策及其结果之间建立起反馈回路，让决策者直接、快速、强制性地看到其行为的后果。就像飞行员位于飞机的前部，面对着所有的仪表，可以直接地了解到自己每一个决策的后果。这样，飞行员就是负起了“内在责任”。

以前，达特茅斯学院每个办公室和教室都安装有单独的温度调节器装置；后来，学院拆除了单独的空调，改造、安装了中央空调设备，由一台电脑集中控制整栋建筑的温度。这样做据说是为了节能，但根据我作为一名基层教员的观察，我认为这样降低了系统的“内在责任”，导致的后果是室温大幅振荡。当我的办公室温度过高时，因为没有办法关闭空调，我只能给温控中心打电话。以前学院曾有巡查人员，隔一段时间（数小时或数日）做一些修正，但经常是矫枉过正，所以后来设立了一部客服电话。但是，这样也不能及时解决问题。在我看来，另外一种解决问题的方法是让系统自己承担起更多的责任，即让每一位教授都能自己控制室内的温度，然后直接根据他们所使用的能量进行收费，这样就可以把“公共物品”（集中温控）变成私人物品，从而提高系统的效率。

再如，为了让系统承担起“内在责任”，应当要求各个城镇或公司都把自己的取水口设在河流的下游，而把排水口设在上游，这样就意味着，如果你自己排放的废水不达标，你马上就可以看到，而且必须承担相应的后果。对于因吸烟而造成的疾病，或因骑摩托不戴头盔、开车不扣安全带等个人疏忽而造成的事故所产生的医疗费用，保险公司或社会保障基金都予以赔偿或支付。国会在立法时，不能包含自身免责的豁免条款（现在有很多这样的条款，比如因积极的行为所产生的招聘需求不受限制，以及无须发布环境影响报告等）。当统治者可以公开宣战但无须自己带着队伍去征战，很大一部分责任就失去了。现在，发动战争可能变得更加不负责任，因为只需要按下一个电钮，就能在很远的地方造成巨大的伤亡，而且按下电钮的人根本看不见任何伤害。

哈丁曾经建议过，如果有人想阻止他人堕胎，就是在违反“内在责任”，除非他们自己真的愿意把孩子抚养成人。

以上几个案例只是“抛砖引玉”，希望能引发读者朋友的思考，在我们当今的文化中，几乎很少在系统中寻找相应的责任，而这是有效行动的根源；同时，在设计系统时，我们也很少考虑让参与各方去承担他们所应承担的责任，去体验他们自己的行动所产生的后果。

保持谦逊，做一名学习者

经过多年的系统思考研究和实践，我已经学会了更多地相信自己的直觉，无需太多的理性思考，我也尽可能地保持这两方面的平衡，但是，我仍然随时准备着应对各种意外的出现。数十年来，我一直在与系统打交道，无论是电脑建模、大自然、人群组织，还是企业系统，我始终提醒自己：每个人的心智模式都是不完整的，而世界是如此复杂，因

此我还有很多很多不知道的东西。

当你发现自己不知道时，真正要做的不是虚张声势、自欺欺人，也不是回避或畏缩不前，而是学习！而学习的方式就是试验，或者就像是巴克敏斯特·富勒（Buckminster Fuller）所说，通过“试错”来学习。在复杂系统的世界中，认准一个方向，一声令下，勇往直前地冲锋陷阵，不懂变通，并不是睿智的做法。即使你确信自己的方向正确，“坚持到底”也可能只是一个一厢情愿的想法。你必须随机应变、相机而动。当你没有把握时，假装“一切尽在掌握”，不仅是人们在犯错时经常产生的一种反应，也无法让人们从失败中学习。应该怎么办呢？那就是采取幅度小而稳妥的措施，持续地监控，认真地观察系统的方向，并且愿意顺势而为，改变自己的路线。

这其实是很困难的，意味着犯错误并需要承认错误，这更不容易。心理学家唐·迈克尔（Don Michael）称之为“拥抱失误”。但是，拥抱你的失误需要很大的勇气。

不管是我们自己，还是我们的同事，或者任何一个“局中人”，都很难知道究竟发生了什么，并且有可能照旧前行，就像我们真的掌握了事实、对所有问题都了如指掌一样，我们也会很肯定、相信自己知道事情的后果，期望自己能够获得最好的结果。此外，当我们处理复杂的社会性问题时，我们会假装自己知道要干的事，这只会降低我们的可信度.....不信任制度和权威的人数正在增加.....认识到不确定性的行为，可以在很大程度上帮助我们改变这一恶化的趋势。

“拥抱失误”就是学习的条件。它意味着搜索、使用和分享“我们到底在什么地方失误了”的信息，了解哪些与我们的期望或希望的状况不符。“拥抱失误”和承受高度不确定性，都会强化我们个人和社会的脆弱性。通常，我们会隐藏自己的脆弱性，无论是对自己还是对他人，都是如此。但是，对于那些真正接受其责任的人，需要的知识远多于社会上一般的人，也更能深入地发现自我。

庆祝复杂性

让我们面对现实吧，大千世界是混乱不堪的。它是非线性的，狂躁不安，又动态变化；在某一个瞬间，它是一种状态，但到了下一个时刻，它又是另一种状态。谁也不知道它要到什么地方去，根本无法精确地测量，也算不出平衡点。它是自组织的，始终处于进化之中。它同时演化出了多样化和统一性。正是由于这些原因，我们所处的大千世界才如此变化万千，异彩纷呈。

在我们人类的大脑中，吸引我们关注的，往往是直线而非曲线，是

整数而非分数，是整齐划一而非参差不齐，是确定无疑而非神秘莫测。但是，我们也存在完全相反的另外一面，因为我们本身也是从动态复杂系统中进化而来，被复杂性所塑造，并且处于复杂系统的结构之中。只有一部分人，喜欢把建筑设计得方方正正、有棱有角，像一个盒子；在人类的历史上，这只是近年来才出现的一种趋势。另外一部分人本能地承认，大自然的设计是分形的，无论是在显微镜下，还是在放大镜中，每一个切面上都有着迷人的、无穷变化的细节。这部分人建造了哥特式的大教堂和波斯人的地毯，创作出了交响曲和小说、设定了盛装狂欢节以及设计了人工智能程序。所有这些复杂性的展现，在我们周围的世界中比比皆是。

我们可以欢庆并鼓励自组织、无序、变异和多样性，至少部分人或你的一部分可以。其中一些人甚至将其作为道德准则，就像奥尔多·利奥波德（Aldo Leopold）的“土地伦理”（Land ethic）^[2]所阐述的那样：“当某件事倾向于保护生物群落的一致性、稳定性和自然之美时，它就是对的，否则就是错的。”

扩展时间的范围

人类最糟糕的发明之一就是利息，并由此产生了诸如回收期 and 贴现率等点子，所有这些都为人们忽略长期利益提供了借口。

在工业社会，无论是企业还是政府考虑的时间范围都是有限的，要么是当前投资的回收期，要么是当期的任期，远远小于大多数家庭的时间范围——后者会持续到孩子长大或孙子辈。而在很多美洲原住民的文化中，他们的决策所谈及和考虑的时间范围长达未来七代。考虑的时间范围越长，生存的机会就会越好。就像肯尼思·博尔丁所言：

大量历史证据表明，如果一个社会失去了后代身份的一致性，失去了对于未来的积极期望，也将失去处理现在问题的能力，并将很快崩溃……除了一些匪夷所思的观点，如“我们应该像鸟儿那样筑巢而居”之外，其实我们的后代在某些方面与鸟儿也是类似的。所以，我们是不是可以把某个地方糟蹋了，然后再移居到另外一个地方去呢？然而，出于为将来着想的考虑，我完全不能接受这样的解决方案……

按照系统理论，严格来讲，没有长期和短期的区分。在不同时间范围内发生的各种现象彼此都是相互嵌套和依托的。此刻你所采取的一些行动，除了即刻就会有一些效果之外，也可能在未来很长时间之后仍然残留着一些影响。其实，此时此刻，我们每个人都正在经历着某个人或

某些人一段时间之前的一些行动的影响，这些行动可能发生在昨天、去年，也有可能是几十年前乃至数个世纪之前的事情。非常快速的过程和非常缓慢的过程之间的连接，有时候很强，有时候很弱。当那些非常缓慢的过程居于主导地位时，事情看起来似乎没有什么变化；而当非常快速的过程占据主导时，事情变化发展的速度可能让人目不暇接。在一般情况下，系统中大的和小的、快的和慢的，都是不断结合、分解、再组合的。

当你沿着一条崎岖不平、障碍遍布的未知小路散步时，如果低着头、只盯着你脚下的那一两步，无疑是很不明智的。当然，只盯着自己前面的伙伴，从不注意自己的脚下，也很愚蠢。这时候，你需要做的就是，既要关注长期（前方的路况），又要留意短期（脚下的状况）。

打破各种清规戒律

不管你是学什么专业的，教科书上是怎么说的，或者某个专家是如何认为的，你都不要盲从，放弃所有的规则，只要紧紧地遵从着系统的指引就可以了，无论它去向何方。可以肯定的是，这样毫无疑问会打破很多传统的清规戒律。为了理解系统，你必须学会向生态学家、化学家、心理学家和神学家们学习，当然也不能局限在这个范围内。你必须深入了解他们的术语行话，以及他们从自己独特的角度得出的结论，取其精华，去其糟粕，并将它们整合起来，融会贯通。当然，这并不容易。

把系统视为一个整体，要求人们的思考必须“跨领域”，也就是把不同学科或领域的人放到一块儿，让他们相互交流、切磋、研究。要想让“跨领域”沟通真正奏效，必须有一个待解决的真实问题，而且来自各个领域的代表必须真心愿意参与解决这一问题，而不只是学术研究。大家必须都进入到真正的学习模式，愿意承认自己的无知，愿意接受新的知识，不只是相互学习，而且要向系统学习。

这完全可以做到。一旦做到，这将非常令人振奋。

扩大关切的范围

要想在一个充满了各种复杂系统的世界中生存，你不仅需要扩展时

间范围，也要拓宽思考范围，也就是说扩大你的关切范围。当然，这样做也符合很多文化中的道德规范。我们姑且不论道德上的争论，但从系统思考的角度看，这样做也是必不可少的。因为真正的系统是相互连接的，我们人类的任何一部分都与他人以及整个地球生态系统不可分割。在这个事物存在普遍联系的世界里，任何一项事物，如果离开了其他相关联的事物，都很难单独存在。

例如，如果你的肺不工作了，心脏也不可能像往常一样继续跳动；如果员工都不干了，你的公司也无法顺畅运作；如果洛杉矶的所有穷人都没有了，富人也将不复存在；如果非洲不存在了，欧洲也不会稳定如初；如果全球环境被破坏了，全球经济也将难以为继。

其实，除了系统之外，大多数人都已经知道了“事物是相互联系的”这个道理。在这一点上，社会公认的道德规范和系统思考的实践原则是一致的。我们只不过是让人们真正做到他们已经知道的。

不要降低“善”的标准

如前所述，系统基模之一是“目标侵蚀”，在当今时代，这一基模最危险的一个实例是：现代工业文明正慢慢侵蚀着人们的美德表现。这一陷阱的作用已经很明显了，后果也很可怕。

现在，有一些不良行为被揭发出来，通过媒体的宣传被放大，但很多人认为这并不值得大惊小怪，甚至觉得稀疏平常。你可能也会这么想：毕竟，我们只是凡人。有太多人性向善的例子没有被注意或提起，因为它们“不是新闻”。而且，凡事都有例外，有坏人，也一定会有圣人存在。不能指望所有人都按照同一个标准来作为。

但是，如果这样的话，人们的期望就会被降低。期望的行为与实际行为之间的差距就会缩小。继而，被确认并逐渐灌输为理想的行为就更少。同时，公众舆论也会充满了愤世嫉俗、玩世不恭的态度。一些公众领袖显然是在“说一套，做一套”，不讲道德和是非，甚至有些人道德败坏，他们也不会对此有所作为。相反，太讲理想会遭到嘲笑。人们关于道德信念的陈述都被认为是可疑的。在公共场合，谈论起恨比谈论爱更为容易。对此，文学批评家和自然主义者约瑟夫·伍德·克鲁奇（Joseph Wood Krutch）是这样说的：

虽然人们之前从未对自己所拥有的感到过满足，或者对实现自己内心渴望的能力有丝毫怀疑，但与此同时，他们也从未认可过对自我过低的评价。他们相信，科学的方法可以帮助人们创造财富、释放潜能，也能从生物学和心理学上解释人们这么做的原因……但是，他们真正关

心的，其实是财富和权力，而不是精神上的富足。

从上文我们已经知道该怎样应对“目标侵蚀”了，因此，不要过度关注坏消息，而对好消息不闻不问，一定要保持客观的标准，不能降低。

系统思考只能告诉我们该做什么，但它本身不会去做。让我们回到知与行之间的鸿沟上。知易行难，虽然系统思考不能填补这一鸿沟，但它可以引领我们来到鸿沟的边缘，让我们更好地进行分析，并继而找到突破点——从人类精神的角度，告诉我们能做什么，以及必须做什么。

[1] 这一故事是Pugh-Roberts联合公司的艾迪·罗伯特告诉我的。——作者注

[2] 奥尔多·利奥波德在其著作《沙乡年鉴》中首次倡导的一种环境伦理。——译者注

附录

系统术语表

英 文	中 文	释 义
Archetype	系统基模	是一些常见的系统结构，能够产生特定的行为模式。
Balancing feedback loop	调节回路	是一种逐渐趋于稳定或特定目标（寻的）的、调节性的反馈回路，也被称为“负反馈回路”，因为它能消除或平衡施加给系统的变化力量。
Bounded rationality	有限理性	从系统的一个部分来看，做出决策或行动的逻辑是合乎情理的，但从系统整体或更大的系统层面上看，这些逻辑就不合情理了。
Dynamic equilibrium	动态平衡	在这种状况下，如果不看流入量和流出量，存量的状态是稳定、不变的。只有当所有流入量和流出量完全相等时，才可能出现这种情况。
Dynamics	动态	一个系统或其中任何一个组成部分随着时间的推移而展现出的行为变化。
Feedback loop	反馈回路	是通过影响与同一个存量相关的流入量或流出量，而使该存量发生改变的机制（规则、信息流或信号）。这是从存量开始的一个闭合的因果关系链，根据存量的水平，通过一系列相关的决策和行动，影响与存量相关的流量，反过来又会通过流量来改变存量。
Flow	流量	是一段时间内，进入或离开某一个存量的物质或信息。

续前表

英 文	中 文	释 义
Hierarchy	层次性	按照不同层次来组织起来、构成一个更大的系统。系统中包含一些子系统。
Limiting factor	限制因素	是系统必不可少的一种输入，在某个特定时刻，它会限制系统的活动。
Linear relationship	线性关系	在系统中，两个要素之间存在因果关系，而且因与果的变化有着固定的比例，可以把它们在图表中以一条直线来表示。其效果是累积的。
Nonlinear relationship	非线性关系	系统中，两个要素之间存在因果关系，但因与果的变化并不存在固定的比例。
Reinforcing feedback loop	增强回路	是一种不断放大或强化的反馈回路，也被称为“正反馈回路”，因为它会强化既有的变化方向。它们既可能是良性循环，也可能是恶性循环。
Resilience	适应力	是系统从扰动中恢复到动态稳定状态的能力，在受到外力影响之后，系统产生了暂时的改变，然后还原或修补、恢复正常。
Self-organization	自组织	是系统构建自身、产生新的结构、学习或多元化的能力。
Shifting dominance	主导地位转换	随着时间的推移，相互制衡的反馈回路之间的相对力量强弱发生了改变。
Stock	存量	在系统中，物质或信息随着时间推移而逐渐积累。
Suboptimization	次优化	某个子系统的目标取代系统整体的目标居于主导地位，导致的一种不合理行为——系统以全部的成本去实现某个子系统的目标。
System	系统	是一系列相互联系和有序组织在一起的要素或部分，它们会产生特定的行为模式，通常被称为其“功能”或“目标”。

系统原理概要

系统

- 总体大于部分之和。
- 系统中的很多相互连接是通过信息流进行运作的。
- 系统中最不明显的部分是它的功能或目标，而这常常是系统行为最为关键的决定因素。
- 系统结构决定系统行为（结构影响行为）。系统行为是系统随着时间流逝展现出来的一系列事件。

存量、流量和动态平衡

- 存量是对系统中变化量的一种历史记录。
- 只要所有流入量的总和超过了流出量的总和，存量的水平就会上升。
- 只要所有流出量的总和超过了流入量的总和，存量的水平就会下降。
- 如果所有流出量的总和与流入量的总和相等，存量的水平将保持不变，即系统将保持动态平衡。
- 要想使存量增加，既可以通过增加流入速率来实现，也可以通过减小流出速率来实现。
- 存量可以在系统中起到延迟、缓存或减震器的作用。
- 由于存量的存在，流入量和流出量可以被分离开来，相互独立。

反馈回路

- 一个反馈回路就是一条闭合的因果关系链，从某一个存量出发，并根据存量当时的状况，经过一系列决策、规则、物理法则或者行动，

影响到与存量相关的流量，又返回来改变了存量。

- 在系统中，调节回路是保持平衡或达到特定目标的结构，也是稳定性和抵制变革的根源。

- 增强回路是自我强化的，随着时间的变化，增强回路会导致指数级增长或者加速崩溃。

- 由反馈回路所传递的信息，哪怕是非物理性的信息，只能影响未来的行为，它不能足够快地发送一个信号，去修正由当前反馈所驱动的行为。

- 在一个由存量维持的调节回路中，设定目标时，必须适当考虑补偿影响存量的重要的流入和流出过程。否则，反馈过程将超出或低于存量的目标值。

- 具有相似反馈结构的系统，也将产生相似的动态行为。

主导地位转换、延迟和振荡

- 当不同调节回路的相对优势发生改变时，系统通常会出现一些复杂的行为，由一个回路主导的某种行为模式变为另外一种。

- 调节回路上的时间延迟很可能会导致系统的振荡。

- 改变一个延迟的长短，可能会导致系统行为的巨大变化。

情景和测试模型

- 系统动力学模型可探究未来的多种可能性，让我们了解“如果这样.....会怎样”之类的问题。

- 模型的价值不取决于它的驱动情景是否真实（其实，对此没有任何人能够给出肯定的答案），而取决于它是否能反映真实的行为模式。

系统受到的约束

- 在呈指数级成长的系统中，必然存在至少一个增强回路，正是它（或它们）驱动着成长；同时，也必然存在至少一个调节回路，限制系

统的成长，因为在有限的环境中，没有任何一个物理系统可以永远地增长下去。

- 不可再生资源受限于存量。

- 可再生资源受限于流量。

适应力、自组织和层次性

- 适应力总是有限度的。

- 不能只是关注系统的生产率或稳定性，也要重视其适应力。

- 系统通常具有自组织的特性，具有塑造自身结构、生成新的结构、学习、多样化和复杂化的能力。

- 层次自下而上地进化；上一层级的目的是服务于较低层级的目的。

系统之奇的根源

- 系统中的很多关系都是非线性的。

- 世界是普遍联系的，不存在孤立的系统；如何确定系统的边界，取决于你的分析目的。

- 在任何给定的时间，对于系统来说，最重要的一项输入是限制最大的那个因素。

- 任何有着多重输入和输出的物质实体，都受到多重限制因素的制约。

- 任何成长都存在限制。

- 当一个变量以指数级形式逼近一项约束或限制时，其接近限制的时间会出乎意料地短。

- 当在反馈回路中存在较长的时间延迟时，具备一定的预见性是必不可少的。

- 系统中每个角色的有限理性可能无法产生促进系统整体福利的决策。

常见的系统陷阱与对策

政策阻力（Policy Resistance）

陷阱：当系统中多个参与者有不同的目标，从而将系统存量往不同方向拉时，结果就是政策阻力。任何新政策，尤其是当它恰好管用时，都会让存量远离其他参与者的目标，因而会产生额外的抵抗，其结果是大家都不愿意看到的，但每个人都要付出相当的努力去维持它。

对策：放弃压制或实现单方面的目标。化阻力为动力，将所有参与者召集起来，用先前用于维持政策刚性的精力，去寻找如何实现所有人的目标，实现“皆大欢喜”，或者重新定义一个更大的、更重要的总体目标，大家愿意齐心协力去实现它。

公地悲剧（The Tragedy of the Commons）

陷阱：当存在一种公共资源时，每个使用者都可以从这种资源的使用中直接获利，用得越多，收益也越大，但是过度使用的成本却需由所有人来分担。因此，资源的整体状况和单个参与者对资源的使用之间的反馈关联非常弱，结果导致资源的过度使用及耗竭，最终每个人都没有资源可用。

对策：对使用者进行教育和劝诫，让他们理解滥用资源的后果。同时，也可以恢复或增强资源的状况及其使用之间的弱反馈连接，有两类做法：一是将资源私有化，让每个使用者都可以直接感受到对自己那一份资源滥用的后果；二是对于那些无法分割和私有化的资源，则要对所有使用者进行监管。

目标侵蚀（Drift to Low Performance）

陷阱：绩效标准受过去绩效的影响，尤其是当人们对过去的绩效评价偏负面，也就是过于关注坏消息时，将启动一个恶性循环，使得目标和系统的绩效水平不断下滑。

对策：保持一个绝对的绩效标准。更好的状况是，将绩效标准设定为过去的最佳水平，从而不断提高自己的目标，并以此激励自己，追求更高的绩效。系统结构没有变化，但由于运转方向不同，便能成为一个

良性循环，做得越来越好。

竞争升级（Escalation）

陷阱：当系统中一个存量的状态是取决于另外一个存量的状态，并试图超过对方时，就构成了一个增强回路，使得系统陷入对抗升级的陷阱，表现为军备竞赛、财富攀比、口水仗、声音或暴力升级等现象。由于对抗升级以指数形式变化，它能以非常令人惊异的速度导致竞争激化。如果什么也不做，这一循环也不可能一直发展下去，最后的结果将是一方被击倒或两败俱伤。

对策：应对这一陷阱的最佳方式是避免陷入这一结构之中。如果已经深陷其中，一方可以选择单方面让步，从而切断增强回路；或者双方进行协商，引入一些调节回路，对竞争进行一些限制。

富者愈富（Success to the Successful）

陷阱：如果在系统中，竞争中的赢家会持续地强化其进一步获胜的手段，这就形成了一个增强回路。如果这一回路不受限制地运转下去，赢家最终会通吃，输家则被消灭。

对策：多元化，即允许在竞争中落败的一方可以退出，开启另外一场新的博弈；反垄断法，即严格限制赢家所占有的最大份额比例；修正竞赛规则，限制最强的一些参与者的优势，或对处于劣势的参与者给予一些特别关照，增强他们的竞争力（例如施舍馈赠、税赋调节、转移支付等）；对获胜者给予多样化的奖励，避免他们在下一轮竞争中争夺同一有限的资源，或产生偏差。

转嫁负担（Shifting the Burden to the Intervenor）

陷阱：当面对一个系统性问题时，如果采用的解决方案根本无助于解决潜在的根本问题，只是缓解（或掩饰）了问题的症状时，就会产生转嫁负担、依赖性和上瘾的状况。不管是麻痹个人感官的物质，还是把潜在麻烦隐藏起来的政策，人们选择的干预行动都不能解决真正的问题。

如果选择并实施的干预措施，导致系统原本的自我调适能力萎缩或受到侵蚀，就会引发一个破坏性的增强回路。系统自我调适能力越差，

就需要越多的干预措施；而这会使得系统的自我调适变得更差，不得不更多地依赖外部干预者。

对策：应对这一陷阱最好的办法是提前预防，防止跌入陷阱。一定要意识到，只缓解症状或掩饰信号的政策或做法，都不能真正地解决问题。因此，要将关注点从短期的救济转移到长期的结构性重建上来。

规避规则（Rule Beating）

陷阱：“上有政策，下有对策”。任何规则都可能会有“漏洞”或“例外情况”，因而也会存在规避规则的行为。也就是说，虽然一些行为在表面上遵守或未违背规则，但实质上却不符合规则的本意，甚至扭曲了系统。

对策：设计或重新设计规则，从规避规则的行为中获得创造性反馈，使其发挥积极的作用，实现规则的本来目的。

目标错位（Seeking the Wrong Goal）

陷阱：系统行为对于反馈回路的目标特别敏感。如果目标定义不准确或不完整，即使系统忠实地执行了所有运作规则，其产出的结果却不一定是人们真正想要的。

对策：恰当地设定目标及指标，以反映系统的真正福利。一定要特别小心，不要将努力与结果混淆，否则系统将只产出特定的努力，而不是你期望的结果。

采取干预措施的杠杆点

- 12.数字：包括各种常数和参数
- 11.缓冲器：比流量力量更大、更稳定的存量
- 10.存量—流量结构：实体系统及其交叉节点
- 9.时间延迟：系统对变化做出反应的速度
- 8.调节回路：试图修正外界影响的反馈力量
- 7.增强回路：驱动收益增长的反馈力量
- 6.信息流：谁能获得信息的结构
- 5.系统规则：激励、惩罚和限制条件
- 4.自组织：系统结构增加、变化或进化的力量
- 3.目标：系统的目的或功能
- 2.社会范式：决定系统之所以为系统的心智模式
- 1.超越范式

系统世界的生存法则

1. 跟上系统的节拍。
2. 把你的心智模式展现在阳光下。
3. 相信、尊重并分享信息。
4. 谨慎地使用语言，并用系统的概念去丰富语言。
5. 关注重要的，而不只是容易衡量的。
6. 为反馈系统制定带有反馈功能的政策。
7. 追求整体利益。
8. 聆听系统的智慧。
9. 界定系统的职责。
10. 保持谦逊，做一名学习者。
11. 庆祝复杂性。
12. 扩展时间的范围。
13. 打破各种清规戒律。
14. 扩大关切的范围。
15. 不要降低“善”的标准。

模型公式

即使不用计算机，我们也一样能从系统中学到很多东西。但是，哪怕是一个非常简单的模型，一旦你开始探索，就会惊喜地发现自己是多么希望更深入地学习，并创建出自己正式的系统数学模型。

本书中的模型都是用STELLA建模软件创建的，该软件由isee systems（原名为High Performance Systems）公司出品。我在下面给出的公式可以很容易地写入各种建模软件中，例如Vensim软件（Ventana Systems Inc.出品）、Stella和iThink软件（isee systems Inc.出品）等。

下列这些公式是用于第1章和第2章中讨论的那9个动态模型的。“转化器”（Converters）可以是常数，或是基于系统模型中的其他要素得出的计算结果。（t）是时间的缩写，（dt）代表时间间隔，即从本次计算到下一次计算之间的时间长度。

浴缸——图1—5、图1—6、图1—7

存量：

浴缸中的水量（t）=浴缸中的水量（t-dt）+（流入量-流出量）×dt

初始存量值：浴缸中的水量=50加仑

t=分钟数

dt=1分钟

运行时长=10分钟

流入：流入量=时间从0～5分钟，流入速度是0加仑/分钟；时间从6～10分钟，流入速度是5加仑/分钟

流出：流出量=（流出速度是）5加仑/分钟

咖啡杯的冷热实验——图1—10、图1—11

●冷却

存量：咖啡的温度（t）=咖啡的温度（t-dt）-（冷却×dt）

初始存量值：

咖啡的温度=100℃，80℃，60℃……分别对应三个对比模型

t=分钟数

dt=1分钟

运行时长=8分钟

流出：冷却=差距×10%

转化器：差距=咖啡的温度-室内温度

室温=18℃

●加热

存量：咖啡的温度（t）=咖啡的温度（t-dt）+（加热×dt）

初始存量值：咖啡的温度=0℃，5℃，10℃……分别对应三个对比模型

t=分钟

dt=1分钟

运行时长=8分钟

流入：加热=差距×10%

转化器：差距=室内温度-咖啡的温度

室内温度=18℃

银行账户——图1—12、图1—13

●存量:

银行账户余额 (t) = 银行账户余额 (t-dt) + (利息收入×dt)

初始存量值: 银行账户余额=\$100

t=年

dt=1年

运行时长=12年

流入: 利息收入 (\$/年) = 银行账户余额×利率

型 转化器: 利率=2%, 4%, 6%, 8%, 10%.....分别对应五个对比模

温度调节器——图2—1至图2—6

存量:

室内温度 (t) = 室内温度 (t-dt) + (加热-散热) ×dt

冷 初始存量值: 室内温度=10℃——冷房间制热; 18℃——热房间制

t=小时

dt=1小时

运行时长=8小时和24小时

流入: 加热=5或实际室温与设定的温度之差的最小值。

流出: 散热=室内外温差×10%.....对应“正常的”房子; 室内外温差×30%.....对应“保温效果不好的”房子。

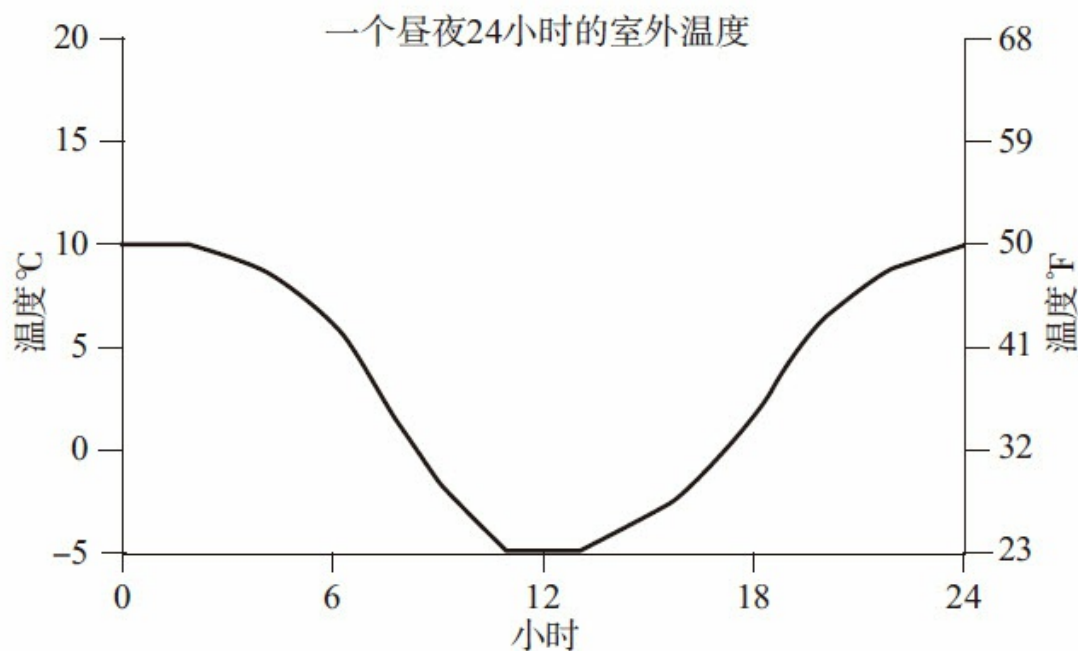
转化器: 温度设置=18℃

实际室温与设定的温度之差=0或（温度设置-室内温度）之差的
最大值。

室内外温差=室内温度-10℃.....对应持续测量的室外温度（图2—2
至图2—4）；室内温度-24小时的室外温度.....对应于一天一夜（图2—
5至图2—6）。

24小时的室外温度变化，从白天10℃（50°F）到夜
间-5℃（23°F），

如图A—1所示：



图A—1

人口——图2—7至图2—12

存量：人口（t）=人口（t-dt）+（出生人数-死亡人数）×dt

初始存量值：人口=66亿

t=年

dt=1年

运行时长=100年

流入：出生人数=人口×出生率

流出：死亡人数=人口×死亡率

转化器：

图2—8：

死亡率=0.009.....或每1000人中死亡9人

出生率=0.021.....或每1000人中生育21人

图2—9：

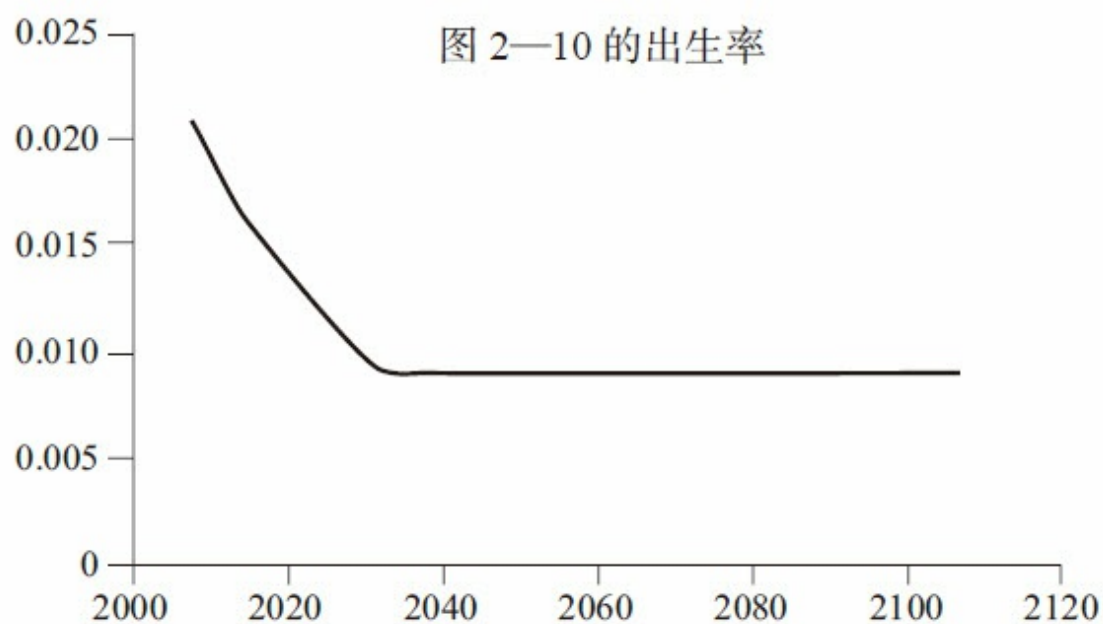
死亡率=0.030

出生率=0.021

图2—10：

死亡率=0.009

出生率从0.021经过一段时间后下降到0.009，如图A—2所示：

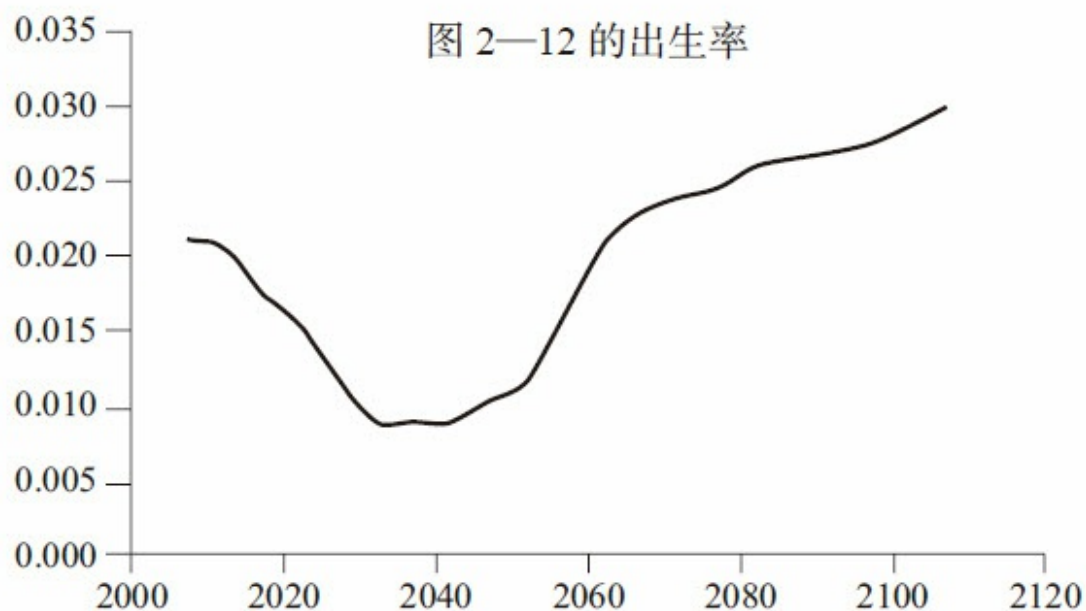


图A—2

图2—12:

死亡率=0.009

出生率从0.021经过一段时间下降至0.009，又上升至0.030，如图A—3所示:



图A—3

资本——图2—13和图2—14

存量：资本存量（t）=资本存量（t-dt）+（投资额-折旧）×dt

初始存量值：资本存量=100

t=年

dt=1年

运行时长=50年

流入：投资额=年产量×投资系数

流出：折旧=资本存量/资本生命周期

转化器：年产量=资本存量×单位资本的产出

资本生命周期=10年，15年，20年.....分别对应三个对比模型

投资系数=20%

单位资本的产出=1/3

库存——图2—15至图2—22

存量：汽车库存量（t）=汽车库存量（t-dt）+（交付-销售）×dt

初始存量值：汽车库存量=200辆

t=天

dt=1天

运行时长=100天

流入：交付=20.....时间0至5；给工厂的订单（t-交货延迟）.....时间6至100）

流出：销售=汽车库存量或客户需求的最小值

转化器：客户需求=每天20辆汽车.....时间0至25；每天22辆汽车.....时间26至100

预期的销售量=感知延迟下的平均销量

期望的库存量=预期的销售量×10

差异=期望的库存量-汽车库存量

给工厂的订单=最大值（预期的销售量+差异）或0.....图2—18；

=最大值（预期的销售量+差异/反应延迟）或0.....

图2—20至图2—22

图2—16，延迟：

感知延迟=0

反应延迟=0

交货延迟=0

图2—18，延迟：

感知延迟=5天

反应延迟=3天

交货延迟=5天

图2—20，延迟：

感知延迟=2天

反应延迟=3天

交货延迟=5天

图2—21，延迟：

感知延迟=5天

反应延迟=2天

交货延迟=5天

图2—22，延迟：

感知延迟=5天

反应延迟=6天

交货延迟=5天

受到不可再生性资源限制的资本存量——图2—23和图2—27

存量：资源（t）=资源（t-dt）-（开采量×dt）

初始存量值：资源=1000.....图2—24、图2—26、图2—27； 1000，
2000，4000.....对应图2—25的三个对比模型

流出：开采量=资本×单位资本收益

t=年

dt=1年

运行时长=100年

存量：资本(t)=资本(t-dt) + (投资额-折旧) ×dt

初始存量值：资本=5

流入：投资额=利润或增长目标的最小值

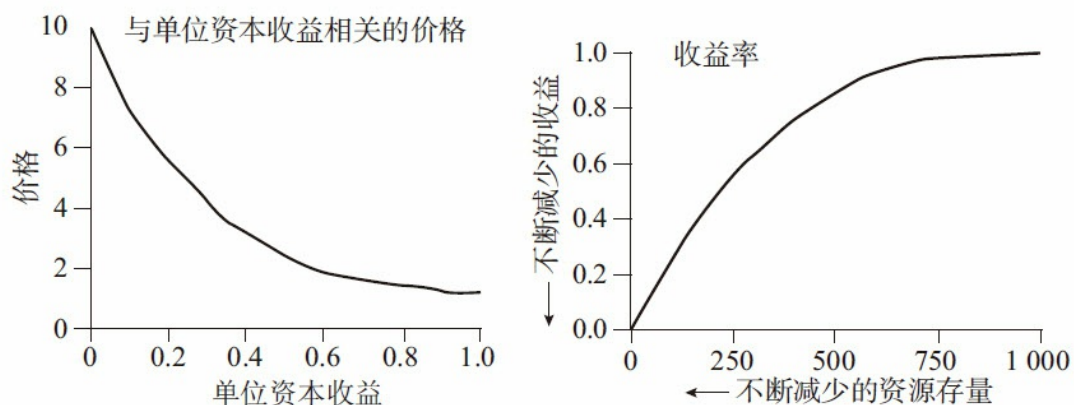
流出：折旧=资本/资本生命周期

转化器：资本生命周期=20年

利润=（价格×开采量）-（资本×10%）

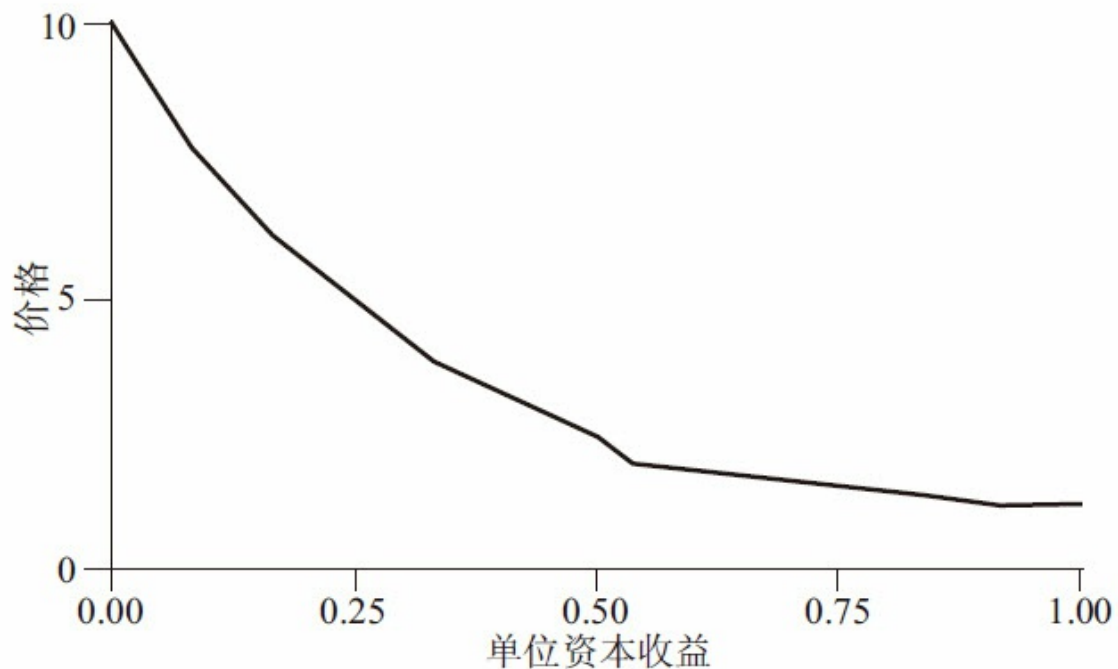
增长目标=资本×10%.....图2—16至图2—26；资本×6%，8%，10%，12%.....对应于图2—26的四个对比模型

价格=3.....图2—24，图2—25，图2—26；图2—27，价格从1.2开始，此时单位资本收益很高；随着单位资本收益下降，价格上升到10，如图A—4所示：



图A—4

单位资本收益从1开始，此时资源存量很高，随着资源存量的下滑，单位资本收益变为0，如图A—5所示。



图A—5

受到可再生性资源限制的资本存量——图2—28至图2—31

存量：资源（t）=资源（t-dt）+（再生量-捕捞量）×dt

初始存量值：资源=1000

流入：再生量=资源×再生率

流出：捕捞量=资本×单位资本收益

t=年

dt=1年

运行时长=100年

存量：资本（t）=资本（t-dt）+（投资额-折旧）×dt

初始存量值：资本=5

流入：投资额=利润或增长目标的最小值

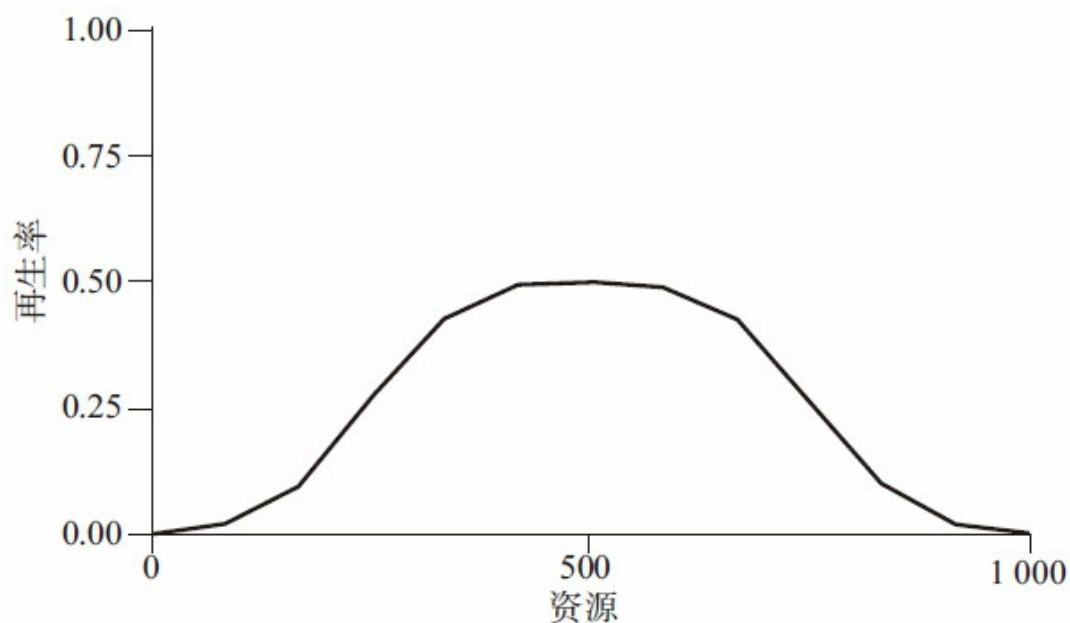
流出：折旧=资本/资本生命周期

转化器：资本生命周期=20年

增长目标=资本 $\times$ 10%

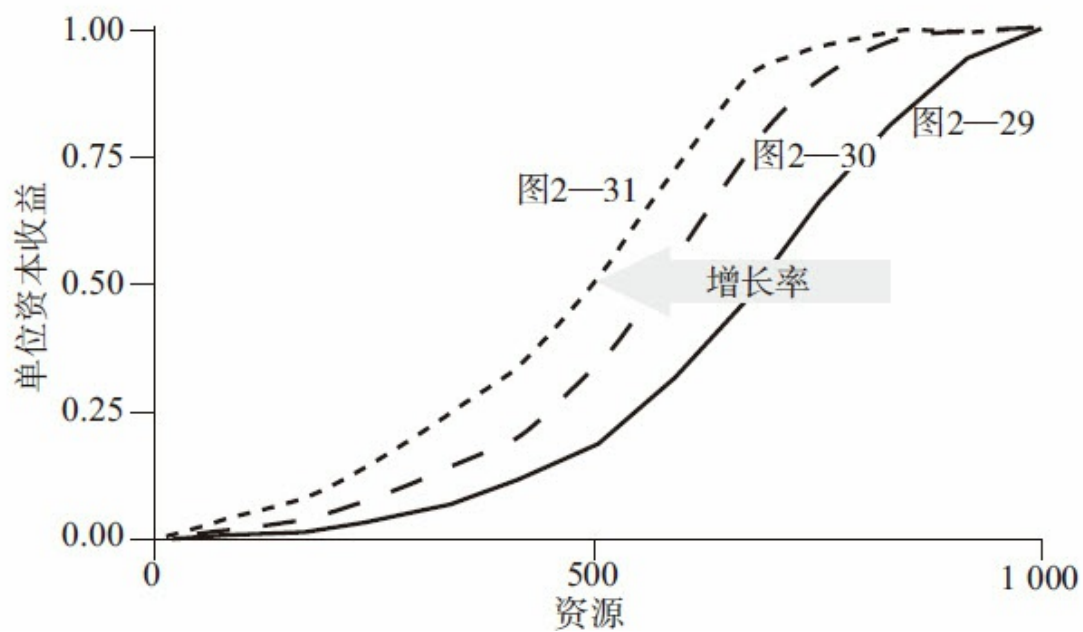
利润=（价格 $\times$ 捕捞量）-资本

价格从1.2开始，此时单位资本收益很高，随着单位资本收益下降，价格上升到10。与前面的模型一样，价格和收益之间也是一种非线性的关系（如图A—6所示）。



图A—6

再生率在资源储备很充足或彻底耗尽时是0。在资源变化范围的中间段，再生率达到最高点0.5左右（如图A—7所示）。



图A—7

单位资本收益从1开始，此时资源储备充足，但随着资源存量的减少，单位资本收益下降（非线性）。单位资本收益的增长从总体上看，图2—29增长率最低，图2—30增长率稍微高些，图2—31增长率最高。

译者后记

九个月，全部利用业余时间，我的第7部译著《系统之美》正式审校完成。又到了最后要写些什么的时候，心中确实是感慨万千。因为每一次翻译都是一次独特的“痛并快乐着”的过程，一方面要靠使命感支撑着自己度过一段苦行僧般的生活，另一方面也是自己在高强度地深入学习、思索、领悟和提升的过程。这本书也不例外。

虽然我自己学习和应用系统思考已经有16年的时间了，但仍感这一体系的博大精深，既有诸多大师级人物基于控制论、信息论、系统论等科学创立的完备体系，又有实用的方法和工具。这着实需要持续不断地操练，才能逐渐提升自己的思维能力，达到“全面思考”、“深入思考”、“动态思考”、“整体思考”的境界，透过事物的表象看到内在的本质，超越一个个孤立的事件看到系统整体的发展变化、来龙去脉，超越本位或局部看到相互关联的整体，对事物运作的内在机理、关键要素以及可能的干预措施等建立较为清晰的认识。当然，在学习的过程中，自己认真学习、用心琢磨是一回事；如果有高手指点，进步就会更快。

回想我自己的学习历程也是如此。我最早在1996年开始接触彼得·圣吉的《第五项修炼》，但当时对如何应用其中的“第五项修炼”（即系统思考）并没有什么感觉。之后，我开始套用一些“基模”，有了一些模模糊糊的感觉。直到2003年，我开始翻译丹尼斯·舍伍德的《系统思考》一书，并有幸结识了孟庆俊、刘兆岩、宋铠教授等“高手”，才算是真正入了门，学习并应用这一实用的方法。之后，我先后给数百家企事业单位进行过“系统思考应用实务”的培训，教学相长，同时自己也一直不停地在学习、实践，于2009年出版《系统思考实践篇》，算是有了“小成”。但是，凭心而论，我自己清醒地认识到，相对于纷繁复杂的大千世界，我自己的“系统思考”能力仍是微不足道的。因此，当看到这本书时，我怦然心动，因为本书作者绝对堪称系统思考的大师，师从系统动力学创始人杰伊·福瑞斯特教授，是知名的“世界模型III”的主创人员、《增长的极限》第一作者，长期从事系统思考研究、实践和教学工作。在本书中，作者不仅深入浅出地阐述了系统思考的基本原理与方法，而且给出了应用系统思考的“内功心法”，让人受益匪浅。事实上，本书初稿写成于1993年，在其后多年的教学中不断得到了验证和锤炼，也在系统思考领域有着良好的口碑，曾广为流传。因此，我本人不仅很

喜欢这本书，也非常荣幸能将其翻译成中文。

本书由我进行翻译，北京学而管理咨询有限公司的崔玲女士协助翻译了第1章和附录部分，并对全书进行审核、校对，最后由我进行全文审校、定稿。

基于十余年的亲身经验，我深感系统思考的价值之重，希望本书能让更多渴望学习和提升自己系统思考能力的读者朋友们有所受益。

邱昭良

2012年7月于北京中关村



Cheers Publishing

出品 Published by: 湛庐文化 Cheers Publishing

总经理 General Manager: 陈晓晖 Chen Xiaohui

出品人 Publisher: 韩焱 Han Yan

法律顾问 Legal Consultant:

北京诚英律师事务所 吴京菁律师

北京市证信律师事务所 李云翔律师

湛庐文化 豆瓣小站

<http://site.douban.com/109787>

高级副总裁 Senior Vice President: 张晓卿 Anna Zhang

夏红卫 Xia Hongwei

总编室主任 Editor-in-chief Assistant: 李阳 Lydia Lee

对外合作部主任 Director of Copyright Dept.: 赵丽琴 Lina Zhao

质量总监 Quality Director: 周裕 Zhou Yu

营销中心 Marketing Dept.

总经理 Marketing General Manager: 陈漪 Tanya Chen

财富汇 Finance, Economics & Management

副总编 Deputy Editor-in-chief: 董寰 Dong Huan

心视界 Psychology & Positive Psychology

副总编 Deputy Editor-in-chief: 季阳 Christine G

特约编辑 Executive Editor: 李琳 Li Lin

封面设计 Designer:  梁庆博 Arthur Liang

采购热线: 010-64970851/64979054

传真 FAX: 010-64979547-8032

投稿请至: service@cheerspublishing.com





Thinking in Systems: A Primer

一部引领个人、组织与社会变革的经典之作

一部应对复杂性挑战、与系统共舞的智慧宝典

- 3大系统特征
- 8大系统陷阱与对策
- 12大系统变革方式
- 15大系统生存法则

左延安 中国工业经济联合会主席团主席，安徽江淮汽车集团公司原董事长

系统思考是整体地、动态地、连续地思考问题的思维模式，是在复杂动态系统中的一种以简驭繁的智慧。决策者唯有深刻领悟系统思考的精髓，才能在正确的时间、正确的地点，做正确的事情。

李兰 国务院发展研究中心公共管理及人力资源研究所副所长，中国企业家调查系统秘书长

系统思考能力是企业家和公共事务管理者等决策者最亟需提升的一项核心能力，本书在此方面具有重要意义和实用价值。

孟庆俊 美国生成创造领导力研究发展中心

邱昭良博士精译之作《系统之美》为重塑我们的创造性系统思维提供了“生生之美”（Creativity in Beauty）的哲学基础和实践工具，既能在求真的过程中体悟系统之美，又可基于系统观反思现实问题，汲取系统智慧，防微杜渐，并提供尽善尽美的系统设计。借此跨学科的远见卓识，可以践行“生生之德”（Creative Creativity）的智慧，应对各种复杂性挑战，创造更广大、和谐与健康的社会—生态—经济（Socio-eco-nomic）的可持续发展之道，为人类谋福利。

宋铠 美国凯斯西储大学系统工程博士，组织学习专家，台湾中央大学管理学院资深教授

我们周围充满了各种活的系统，不仅需要逻辑分析，也离不开艺术、直觉、爱与慈悲等智慧。我们必须注重重心与物的平衡，左脑和右脑的平衡，冷静洞悉系统的动态性、整体性和规律性。如果能把这种根植于东方智慧的全脑思维技能推广普及，对于社会的和谐、个人的发展，都是价值无限的。

奥兰·扬 美国加州大学圣巴巴拉分校教授，IHDP科学委员会主席

本书的出版具有划时代的意义，它重塑了人类对于社会生态系统的理解，并会在未来很长时间内产生深远的影响。正如20世纪六七十年代《寂静的春天》一样，将引领我们更好地理解生态系统的本质。

湛庐文化
Cheers Publishing

上架指导：管理 / 组织管理

ISBN 978-7-213-05011-4



9 787213 050114 >

定价：56.90元