

注意的新理论——特征整合论

王 魁

(北京大学心理学系)

一、概 述

自认知心理学兴起以后,曾经从不同的角度提出过几个注意的模型或理论,都在当时产生过一定的影响。然而目前最引人注目的一个新理论是 Anne Treisman(1977,1980)提出的特征整合论(feature integration theory)。它的核心是将客体知觉过程分成两个阶段:早期的前注意阶段和特征整合阶段。前者对特征进行自动的平行加工,无需注意,后者则通过集中注意将诸特征整合为客体,其加工方式是系列的。这个理论着眼于注意的整合功能,将注意与知觉的内部过程紧密地结合起来,是很有特色的。在此后的10多年间,Treisman 针对这个理论进行了大量的系统的实验研究,得到许多有价值的成果,进一步充实了它,并做了一些理论修正。现在,这一新的理论在心理学中占有突出的地位,也引起人工智能学界的重视。

特征整合论是在以前有关注意研究的背景中提出来的。本世纪60--70年代,注意的研究主要集中于注意过滤器在信息加工系统中的具体位置问题。以 Broadbent(1958)和 Treisman(1960)为代表的知觉选择模型或早期选择模型认为,注意过滤器处在知觉阶段,它在识别之前对信息进行选择,只让一部分感觉输入通过而被识别。以 Deutsch 等(1963)为代表的反应选择模型或晚期选择

模型与之相反,认为注意过滤器处于反应阶段,所有的感觉输入均可被觉察和识别,但后来只有一部分信息通过注意过滤器而得到反应。这两个对立的注意理论各有其实验证据,然而有些实验结果竟然可被这两个对立的理论所解释。后来侧重注意过滤器位置的观点有所减弱。有人认为,注意在不同条件下或为知觉选择,或为反应选择,甚至同时包含二者。此外,Kahneman(1973)提出资源分配理论,从心理资源分配的角度来阐述注意,避开注意机制在信息加工系统中的位置问题。值得特别强调的是著名心理学家 Neisser(1967)的观点,他从另一个角度解决了这个问题。他最先提出了前注意加工和集中注意加工的区别:前注意加工是自动的,可能是平行的加工;集中注意加工则是系列的加工。Neisser 认为知觉是主动的、灵活的,注意在概念驱动和数据驱动的共同作用下引导知觉,因而将注意的选择作用定位于靠近刺激还是靠近反应都是没有意义的。可是这种淡化注意位置的思想当时没有引起足够的重视。然而在 70 年代前后,大量的视觉搜索实验的结果,特别是有关早期的视觉研究,在总体上是支持 Neisser 的两种加工区分的。与此相近的还有 Shiffrin 和 Schneider(1977)区分的自动加工和控制加工,等等。不难看出,那些着眼于注意位置的理论 with Neisser 的两种加工观点相比,恰巧忽略了注意与知觉操作的联系。Treisman 提出的特征整合论无疑受到 Neisser 的观点影响,而且正是力图将注意与知觉的内部过程更加紧密地结合起来。

Treisman(1982)认为,知觉是构成特定时间和地点的特定的客体、状态和事件的暂时表征。它通常要包含语义记忆中贮存的信息与感觉输入相匹配,但不应将知觉与语义记忆结点的暂时激活相等同。知觉带有情景结构的性质,是将一些独立编码特征加以整合而构成的。特征整合论的出发点是知觉的特征分析说,它区分特征和客体:特征是维量的值,如三角形是图形维量的一个值;客体则是诸维量值的整合,如红色三角形。照它看来,对特征和客体的

加工是在知觉过程的不同阶段实现的。

依照这个理论,早期的前注意阶段以自动的平行的方式对一些基本特征,如颜色、线段朝向、明暗、大小等,以及某些复杂的特征如封闭性等进行加工,迅速而无需注意。这是一种自下而上加工。在这个阶段上,所知觉的特征处于自由飘移(free floating)状态,即特征不受其所属的客体的约束,如红色苹果的红色和形状不受这一苹果的约束,其位置在主观上是不确定的。知觉系统对各个维量的特征进行独立编码,如颜色、形状和线段朝向等是分别编码的,并且一个维量内的各个特征也是如此,如红色、黄色等都是分别编码的。这种加工由各特征模块实现,各特征编码构成相应的特征地图(feature map),如颜色地图、朝向地图等。这也就是前注意的表征。但是一个客体的知觉还必须对其各个特征加以联合,而特征的联合又要求对特征进行定位。这就需要第二个加工阶段即特征整合阶段。在这个阶段上,位置地图(map of locations)或位置总图(master map of locations)起着重要的中介作用。知觉系统中的位置地图表明全部特征的边界位置在哪里,或者说哪里有东西,但并未说明什么特征在什么位置。换句话说,位置地图只表明哪些位置是空的,哪些位置是实的。所谓实的位置意味着那里有某种或某些还未知道的特征。在前注意阶段,位置信息虽得到一定编码,但不能获取这些位置信息,特征加工可不管其位置,然而客体加工却离不开位置并需要集中注意。集中注意可在位置地图上作出选择,就像聚光灯那样进行扫描,扫描到哪里,那里的任何特征就通过该位置与它们在特征地图中的相应位置的联系而被提取并被整合,集中注意提供“胶水”(glue),把先前独立编码的一些特征结合为客体。特征地图中的位置只是通过它与注意所选择的位置地图中的位置的联系才能得到进一步的加工并进入意识。如果注意分散或过载,特征将不在正确的位置上结合起来,导致错觉性结合,一个客体的特征将与另一个客体结合起来。特征整合阶段的加工

方式是系列的,集中注意一次只扫描位置地图中的一个位置,被扫描位置上的特征整合而成暂时客体表征(temporary object representation)或文件(file)。它包含客体作用的时间、地点以及特征、联系和名称等。这个暂时客体表征具有情景结构的性质。它与识别网络(recognition network)互相联系,由此实现客体知觉或识别。客体知觉既包含自下而上加工,又包含自上而下加工。上述的理论模型可见图 1。

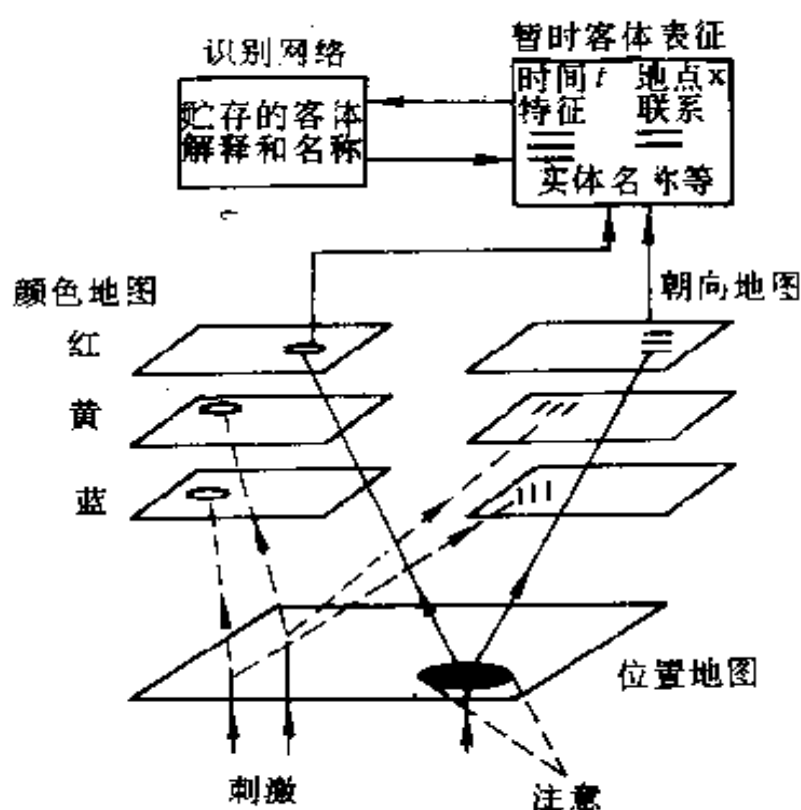


图 1 特征整合模型

上述模型是 Treisman (1985, 1988) 提出来的两个模型之一, 她提出的另一个模型 (Treisman, 1986) 的基本原则与此相同, 但也有一个区别: 上述模型将特征地图放在位置地图之上; 而另一个则相反, 特征地图放在位置地图之下, 这反映出 Treisman 本人对位置地图乃至整个模型的某些方面的看法还有犹豫。如 Treisman 所说的, 将特征地图放在位置地图之下, 意味着位置地图接收特征

地图或模块的输出；而特征地图放在位置地图之上则意味着，外部刺激最初是作为一个总体进入知觉系统的，各维量在位置地图上联合为一个表征，然后才分别进行各个维量的分析。她本人后来倾向于将特征地图放在位置地图之上的模型。我们也认为，特征地图在位置地图之上的模型似乎更适合特征整合论，否则将难于说明位置地图的作用。假如某个刺激一进入知觉系统就被分解为各个单独的维量，那么这些单独的维量后来怎样会聚在位置地图中的一个位置上呢？当然这不是没有办法解释的，不过需要增加其他额外的因素，使模型变得更加复杂起来，而将特征地图放在位置地图之上就显得简单多了。这确实是应当考虑的一个问题。

综上所述，我们可以看出有关特征整合论的实质的两个方面。第一，这个理论着眼于注意的整合功能，将注意与知觉的内部过程结合起来。这固然是该理论的一个特点，但这种整合功能是附随注意的选择功能的，或者说这两种功能连在一起，使注意所扫描的诸特征得以整合为知觉客体，因此这个理论本质上仍是知觉选择性质的一个模型，甚至可以说它是原先的知觉选择模型的深化或变式。第二，从知觉加工过程来看，这个理论区分两个阶段，先在前注意阶段进行特征分析，形成各种特征地图，然后再通过集中注意对位置地图的扫描，把各种特征加以整合，第一个阶段加工完成后再进行第二个阶段加工，因此这个理论模型似应属于信息加工的离散阶段模型的范畴。以上两点是特征整合论的核心问题。后来就这个理论出现的争议主要也是围绕这两点。Treisman(1990)本人曾对她的这个理论进行了修正，主要是有关第二点。这些在后面还要谈及。

无论怎样，目前形式的特征整合论比现有的其他注意模型乃至 Neisser 的观点都前进了一步。它对客体的知觉过程给出了一个完整的说明，包括注意在其中的功能。这样它就涉及到当前知觉研究的一些重大问题，如前注意的信息表征、知觉的基本单元等。

可以说特征整合论也是关于知觉的新理论。Treisman 在 10 多年间做了大量的实验来验证她的理论,其中有关错觉性结合(illusory conjunction)和非对称性搜索(search asymmetry)的实验尤为引人注目。后面将择要分别介绍。最后应当指出,这些实验都是以视觉为研究对象的,并且特征整合论在目前也是主要针对视觉道的,还没有涉及其他感觉道。

二、错觉性结合实验

错觉性结合是指不同客体的特征在注意分散或过载时发生的彼此交换的现象,如呈现的是绿 X 和红 O,被试报告的却是绿 O 和红 X。错觉性结合实验通常利用双作业任务,即安排两个作业,第一作业是集中注意作业,即将注意吸引过来;第二作业是非注意作业,这是真正的实验作业,考察它在缺乏注意条件下的进行过程和特点。注意过载在实验中也可起到注意分散的同样作用。但注意过载不需要安排双作业任务,而是通过缩短刺激呈现时间来实现,即非常迅速地呈现刺激。其实双作业任务也需要迅速呈现刺激。这些实验有力地支持特征整合论。它们说明,前注意阶段对特征进行独立编码,而特征整合为客体则需集中注意。下面介绍 Treisman(1982,1984,1986,1988)做的 3 类错觉性结合实验。

1. 字母错觉实验

这个实验应用双作业任务。第一作业的刺激为两个数字,第二作业的刺激为 3 个颜色不同的字母,涉及的字母有 5 个(T, S, O, N, X),颜色也有 5 种(粉红、黄、绿、蓝、褐)。第二作业的刺激分成两类:一类为 3 个不同颜色的不同字母;另一类也是 3 个彩色字母,但其中或有两个字母相同或有两种颜色相同。两个作业的刺激均印在一张卡片上,3 个彩色字母居中横排,其两侧各有一个数

字,可见图 2。实验时只要求被试注意两侧的数字。刺激卡用速示

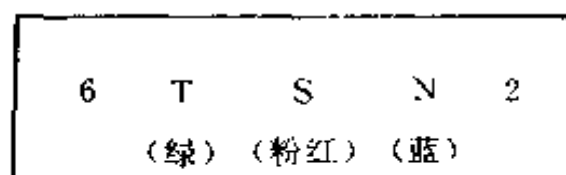


图 2 字母错觉实验刺激卡

器呈现。先呈现带注视点的空白卡 1 秒钟;然后呈现刺激卡,平均时间为 120 毫秒;再呈现掩蔽刺激,时间为 200 毫秒。要求被试先报告看到的数字(如 62),再报告看到的字母及其位置(左、中、右)和颜色,如左面的绿 T,中间的粉红 S 等。结果发现,第一作业(报告数字)的正确率高达 90%以上,而第二作业(报告字母)的正确率仅为 52%。令人感兴趣的是第二作业中出现两类错误:一类是特征错误,即被试报告出刺激卡所没有的字母或颜色,以图 2 的刺激卡为例,被试报告出“左面的黄 T”或“右面的蓝 O”等。另一类是结合错误,即被试所报告的字母或颜色确为刺激卡所有,但不在正确位置上,仍以图 2 的刺激卡为例,被试报告出“左面的粉红 T”或“中间的粉红 N”等。这就是说发生了不同位置上的特征交换,也即出现错觉性结合。这种性质的错误表明,注意分散确可导致错觉性结合。由于错觉性结合的前提是觉察特征,因此它们还表明,前注意阶段可对单个特征进行独立编码,特征是处于自由飘移状态的。

这个实验的结果有 3 点值得注意:①错觉性结合没有距离效应(distance effect),相邻字母与相隔字母间交换特征的可能性是相同的。②错觉性结合没有维量效应(dimension effect),颜色和字母这两个维量出现的错觉性结合在数量上没有显著差异。③错觉性结合没有相似性限制,特征是否交换与交换前两个位置的其他特征是否相似无关。这几点也反映出错觉性结合的特性。它们进一步说明,错觉性结合依赖于内部的信息加工而不是外部刺激

的物理特点。

然而,这个实验也还存在一个重大问题。这就是它还未能证实错觉性结合是知觉加工的结果,不受其他心理因素的影响。这些因素包括:①记忆因素,这个实验采用自由报告法,被试的报告离不开记忆信息提取,所以错觉性结合也可能发生在回忆过程中;②言语编码,由于这个实验要求被试作口头报告,所以错觉性结合也许是言语码而不是知觉码的交换;③反应倾向,实验发现被试有一种将特征结合起来报告的倾向。这些因素对错觉性结合的出现是否有影响是需要弄清楚的,否则就不能肯定错觉性结合的知觉性质。

针对上述因素,Treisman 做了进一步实验。首先做了字母继时匹配实验。该实验的刺激卡与前述实验类似,只是去掉了有重复特征的刺激卡。实验程序基本同前,但有一个重要变动,即在呈现刺激卡之前先呈现一个探测字母。探测字母分作 3 类:①同一性的,即探测字母(具有一定颜色)在随后的刺激卡中会出现;②特征性的,即探测字母仅在颜色或字母上与随后呈现的刺激卡上的某个颜色或字母相同;③结合性的,即探测字母与刺激卡上的字母和颜色的错觉性结合相同。要求被试报告刺激卡中是否有与探测字母相同的字母。结果发现,对这 3 类探测字母回答“有”的百分数如下:同一性的为 62%;特征性的为 23%;结合性的为 36%。这最后一类探测字母的实验数据实际上表明错觉性结合的数量。由于采用不同类型的探测字母和匹配任务,在这个实验中可较好地排除反应倾向因素的作用,也可减弱记忆因素和言语编码的作用,但还不能排除这后两个因素的作用。为此,Treisman 又做了字母同时匹配实验。

字母同时匹配实验的程序与以前的实验基本相同。实验所用的刺激卡见图 3。图中每张卡片的左右两侧各有一个数字,但未画出。实验时要求被试先报告两侧的数字,然后再报告刺激卡中是否有两个字母和颜色完全相同的一对字母。图 3 中各卡片左侧列出

对该刺激卡报告“有”的百分数,括号内是标准差。现在来看这些实验结果。图 3 表明,图中 a 和 d 两刺激卡都含有两个相同的彩色字母,所以报告“有”的百分数较高;并且 a 中只有一对相同的彩色字母,d 中却有两对,因此 d 的报告率高于 a。令人感兴趣的是 b、c 和 e 3 张刺激卡,它们本身并不含有两个相同的彩色字母,但被试却报告“有”,这个百分数实际上表明错觉性结合的数量,因为它们只有通过交换字母特征才能得到一对完全相同的彩色字母。其“有”的百分数低于 a 和 d 也就可以理解了。进一步来看,在 c 中要得到一对相同的彩色字母,存在着两种特征交换途径,而 b 和 e 只有一种交换途径,因此 e 的报告“有”的百分数高于 b 和 c。这些实验结果再次证明错觉性结合的存在。然而更为重要的,是通过这种同时匹配作业可以更好地排除记忆和言语编码等非知觉因素的作用。应当说,这些结果的确揭示了错觉性结合的知觉性质。

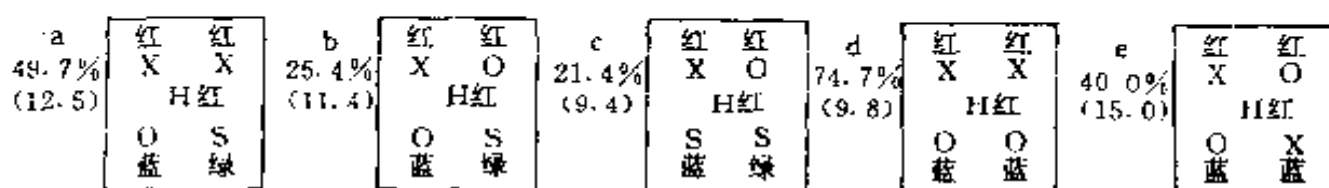


图 3 字母同时匹配实验刺激卡

2. 图形错觉实验

图形错觉实验也引起广泛的兴趣,因为图形具有更明显的知觉特性。在这类实验中大多不采用双作业任务,但实验设计思想和字母错觉实验是一样的。只是采用快速呈现刺激来造成注意过载,使被试不能将注意集中于所呈现的刺激。Treisman 曾经做过几个著名的实验,如三角形、箭头、美元符号 \$ 和“胡萝卜”实验等。

三角形错觉实验所用的刺激图见图 4。实验用计算机呈现刺激图。实验程序与以前实验基本相同。先呈现一个注视点,再迅速呈现一张刺激图,然后呈现掩蔽刺激。要求被试报告是否看到了三

角形,按相应的键作出“是”或“否”的反应。结果发现,当呈现刺激

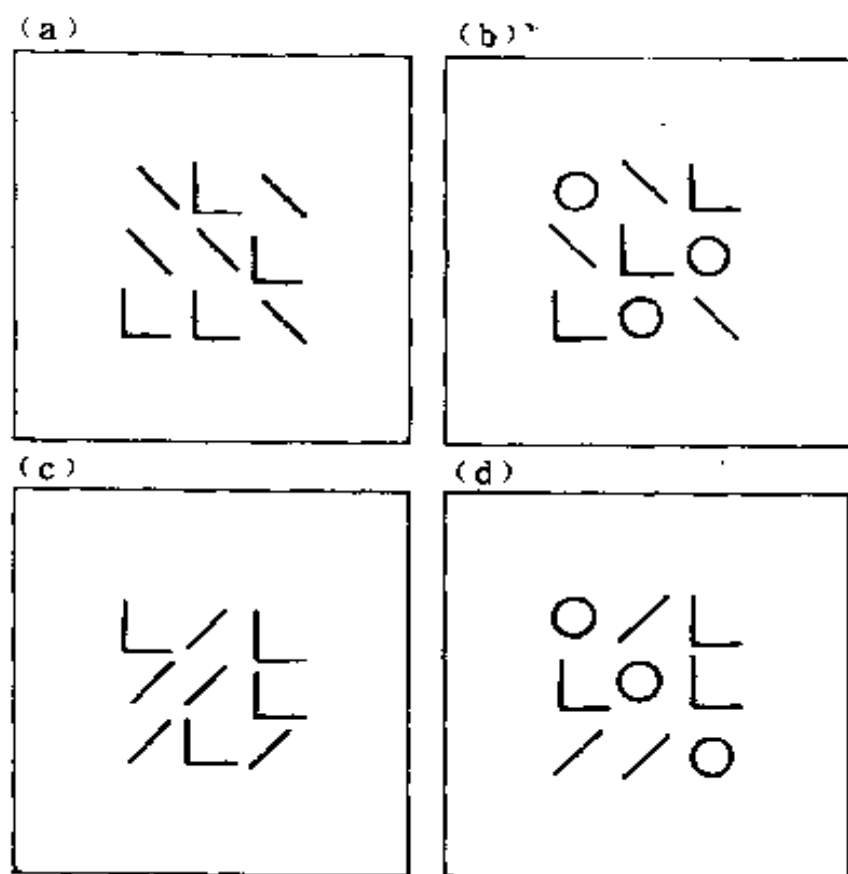


图 4 三角形错觉实验图形

图形 a 时,被试偶而报告看到三角形;而当呈现 b 时,被试报告看到三角形的次数就多了;但呈现 d 和 c 时,则极少看到三角形。将图中 a 与 c,d 相比,可以看出:在 a 中的斜线的朝向是适宜与给出的角构成三角形的;而 c 和 d 中的斜线的朝向却不适宜构成三角形。因此在 a 中可形成错觉三角形,即构成错觉性结合。这个结果也说明,线段朝向是前注意阶段加工的特征。有意思的是图中 b 的角和线段朝向与 a 相同,但 b 中有封闭的圆形,并且在其中看到的三角形显著地多于 a,这提示封闭性有利于形成错觉三角形,它似乎也是知觉早期加工的一个特征。但 Julesz(1981,1984)对封闭性

作为早期加工的特征有不同看法。这个实验证明,在注意过载时确可出现错觉性结合。

美元符号\$错觉也是一种图形错觉。这种错觉实验的程序与三角形的实验相同。它应用的刺激图见图5。在图中a和b都没有现成的美元符号\$,但有构成这个符号的成分,即S和有适宜朝向的斜线。然而斜线在a中是单独出现的,在b中是出现在一个封闭的三角形里。当呈现一张刺激图后,要求被试报告是否看到了美元符号\$。结果是被试常报告在a和b中都看到了美元符号\$,而且两者的报告率相近。这个结果提示,斜线不论是单独出现还是在三角形里出现,都可以在前注意阶段得到单独编码,当注意过载或分散时,便与S构成错觉性结合。与上述三角形错觉实验相比,我们看到三角形在前注意阶段似乎有两种表征:一种是包含封闭性的整体表征,另一种是作为3条不同朝向的线段的分别表征。这是一个涉及图形知觉的基本单元的问题,目前存在着争论。

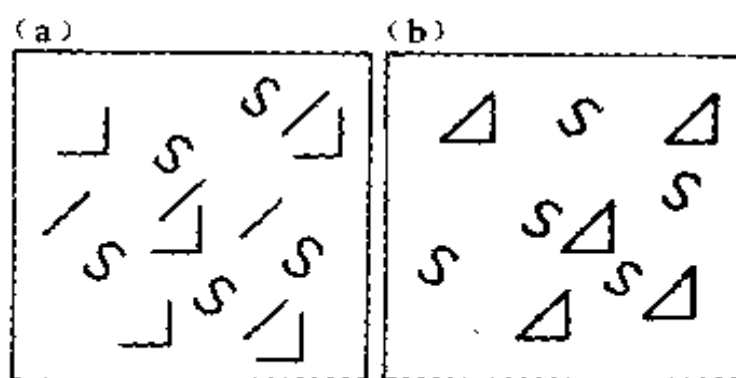


图5 美元符号错觉实验图形

“胡萝卜”错觉实验涉及其他有意义的问题。这个实验应用双作业任务,其实验程序与其他双作业实验基本相同。它采用的刺激卡见图6:其中a是测试卡,两侧各有一个数字,中间为3个不同的彩色图形,即橙色三角形(“胡萝卜”)、蓝色椭圆和黑色圆环;b为掩蔽刺激卡,其上方有一指示符号。在实验中,继刺激卡之后呈

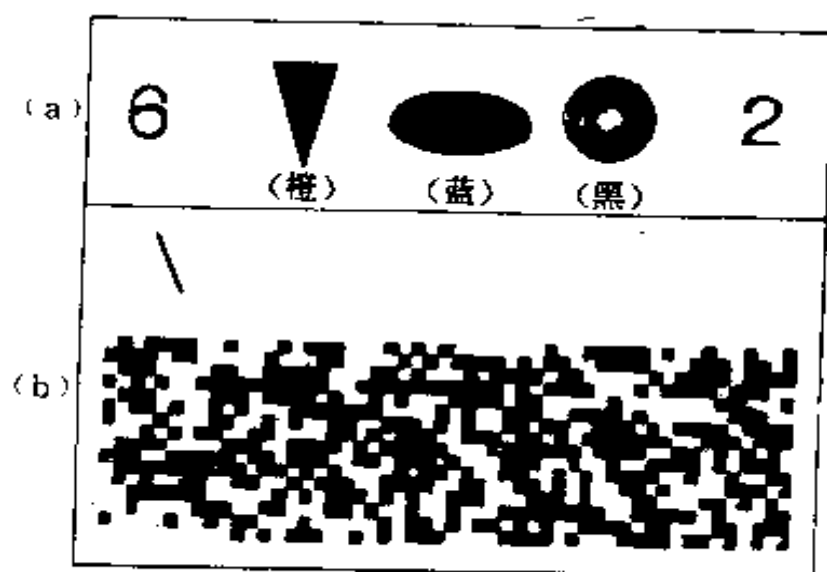


图6 “胡萝卜”错觉实验刺激卡

现掩蔽刺激卡,要求被试先报告刺激卡上的数字,再报告指示符号所指刺激卡中位置上是什么颜色的什么图形。每次试验的指示符号所指位置可不同。实验分3种条件:在第一种和第二种条件下即呈现刺激卡a,并且在第一种条件下事先如实地告诉被试将要呈现的这3个彩色图形;但在第二种条件下却事先告诉被试将要呈现橙色胡萝卜、蓝色湖泊和黑色轮胎;第三种条件下,应用的刺激卡分别在3个图形和3种颜色上与刺激卡a相同,但在图形与颜色的配对上与a不同,如蓝色三角形、橙色椭圆等。但是,在第三种条件下事先给被试的指示语与第二种条件的相同。因此第三种条件的实验与其他两种条件有一个重大区别,即呈现的彩色图形与告诉被试的不一致,如呈现的是蓝色三角形,而告诉的却是橙色胡萝卜。现在来看实验结果:在第一种条件下出现29%的错觉性结合,在第二种条件下只出现18%的错觉性结合。第二种条件下的错觉性结合少于第一种条件。这表明被试得知实物名称后,可以利用过去的有关知识,建立了期望,从而易化对特征的整合。然而现在还不清楚,错觉性结合是发生在利用过去知识之前的阶段还是

之后的阶段。这一点在第三种条件的实验中得到一定的说明。由于在该条件下呈现的图形的实际颜色与告诉被试的不同,即实际色和期望色不一致,如果错觉性结合发生在利用过去知识之后,那么当刺激卡上其他某个位置有期望色如橙色时,被试会比刺激卡中未包含这个期望色时报告更多的橙色“胡萝卜”,即出现更多的错觉性结合,以使所知觉的实际色与期望色相一致。然而结果却不是这样,被试在刺激卡包含或不包含期望色(橙色)两种情况下出现的“胡萝卜”错觉在数量上没有显著差异。这说明错觉性结合不受已贮存的过去知识的影响。据此,Treisman 认为,错觉性结合出自前注意阶段的感觉材料,是自下而上加工的结果,不受自上而下加工的约束。换句话说,错觉性结合发生在利用过去知识之前的阶段,而过去知识则有助于更有效地利用注意去整合各种特征。

“胡萝卜”错觉实验及从中得出的结论对特征整合论无疑有重要意义。但是,这个实验本身和结论都还存在问题。例如,在试验前告诉被试的是实物名称,而呈现的却是几何图形。这种不一致表明自下而上加工与自上而下加工不具有共同的情景,它可能使自上而下加工未能充分发挥作用。这就是说,错觉性结合似乎并非不受自上而下加工的影响,也许是实验本身的条件减弱了这种影响。此外,将自上而下加工限定在知觉过程的某一阶段(较晚阶段)的看法也是可以怀疑的。现在一般认为,自下而上加工和自上而下加工在整个知觉过程中都相互联系着,自上而下加工从一开始就对知觉过程起作用。这样看来,“胡萝卜”错觉实验涉及的问题还需要进一步研究。

3. 字词错觉实验

这个实验所用的一部分刺激卡见图7。其中a和c都是英文单词,b和d则是非字词即无意义音节。实验时先由实验者以拼音方式说出一个英文单词或非词,即为靶子。如靶子为car,即拼作c、

a, r; 然后用速示器迅速呈现一张刺激卡, 要求被试回答该刺激卡中是否有靶子, 若“有”则需指出其在卡片上的位置。针对 a 和 b 所用的靶子是单词 day, 对 c 和 d 用的靶子是非词 sen。可以看出, 在上述 4 种刺激卡中都没有靶子。但是结果发现, 被试对这几种刺激卡都报告有靶子, 即出现错觉性结合; 字词刺激卡和非字词刺激卡、字词靶子和非字词靶子在错觉性结合的数量上都没有显著差异。但合起来看, 当靶子是字词而刺激卡是非字词时, 错觉性结合稍多些。然而后来 Treisman 将实验方法加以改变, 改用自由回忆法, 这一点也不复存在了。这个实验同样表明, 自上而下加工对错觉性结合没有多大作用。看起来, 这个实验结果与“胡萝卜”错觉实验在趋势上是一致的, 但它们之间仍然存在着一个颇为重要的区别, 字词错觉实验没有完全否定词汇的自上而下的约束作用。结合其他的字词错觉实验, Treisman (1986) 承认, 关于词汇的知识所起的自上而下的约束作用, 在某种程度上可以调节错觉性结合的频数, 但它不是决定因素。这个看法克服了以前的偏颇, 显得较接近实际情况, 当然这个问题仍需进一步研究。这个实验也引出一个重要问题, 就是能否将字母看作知觉基本单元的问题, 这同样是需要加以研究的课题。

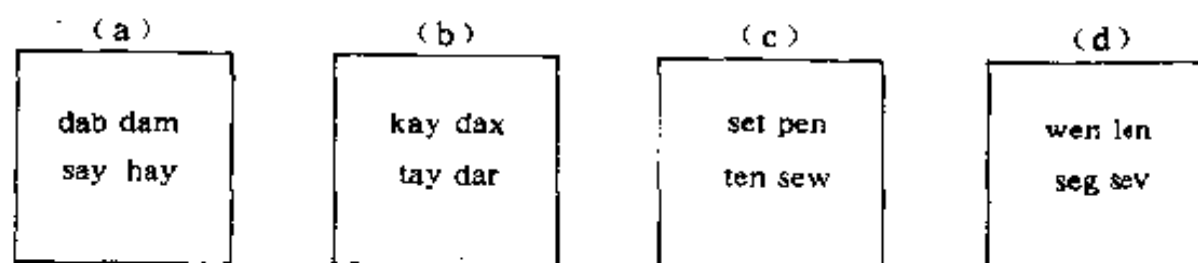


图 7 字词错觉实验刺激卡

总的来看, 几种错觉性结合实验得到的结果是支持特征整合论的。它们表明, 在前注意阶段, 特征是独立编码的, 处于自由飘移

状态,当注意分散或超载时就会出现错觉性结合,而特征整合为客体则需集中注意。可以说,这些实验结果主要支持了特征整合论的框架。这正是错觉性结合实验的重要意义的所在。

三、非对称性搜索实验

搜索的非对称性是视觉搜索中的一个有趣的现象。在提出特征整合论的初期,Treisman(1977,1980)曾经做过一些视觉搜索实验。其典型的结果是,一个单独特征的搜索是快速的和平行的,而特征结合(客体)的搜索则是缓慢的和系列的,例如,在若干个绿X和褐T(干扰项)中搜索一个蓝色或一个S字母(靶子),这种搜索完成得快,且不管有多少绿X和褐T,所需时间几乎不变;但是,如果要在同样的若干个绿X和褐T中搜索一个绿T,那么搜索将是缓慢的,而且所需时间随绿X和褐T的数目增加而呈线性增长。Treisman认为,识别特征可以不管其位置,属前注意阶段加工;而识别客体则需将注意集中于相应位置。这一类实验也支持特征整合论的框架。后来Treisman逐渐转向非对称性搜索实验,所谓非对称性搜索是指,在若干个甲类项目(干扰项)中找到一个乙类项目(靶子),与从同样的若干个乙类项目(干扰项)中找到一个甲类项目(靶子),两者的搜索速度有显著差异,出现非对称现象。换句话说,甲乙两类项目互易其靶子或干扰项的角色时,搜索所需时间不同。例如,从若干短的直线中搜索一个长的直线所需时间少,而从同样的若干长的直线中搜索一个短的直线所需时间多,前者快于后者。这种非对称性现象首先是Neisser(1963)发现的,后来Back(1973,1982)以及Julesz(1981)等都曾涉及过。Treisman(1985,1986,1988)做过许多非对称性搜索实验,除进一步论证其理论框架外,主要是研究特征地图的工作原则,这类实验以反应时为指标。下面择要加以介绍。

1. 不连续维量的有无性特征的非对称性

这类非对称性搜索的一个实验所用刺激卡见图 8。刺激卡分成 a、b 两类。在 a 中靶子是○，而干扰项是⊙；在 b 中这两类项目的角色倒过来了，靶子是⊙，而干扰项是○。这两类项目本身的

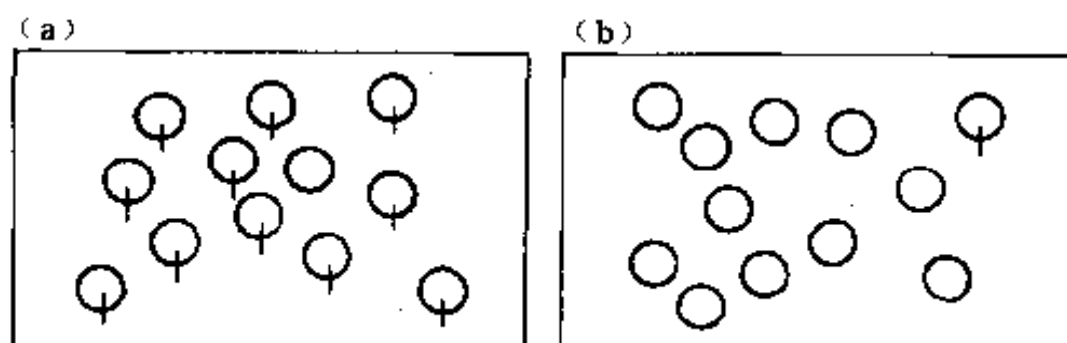


图 8 有无性特征的搜索实验刺激卡

区别即在于一个有与圆相交的竖直线，另一个则没有。无论在哪一种刺激卡中，靶子只有一个，而干扰项则按实验要求可有不同的数目。靶子与干扰项之和称为画面大小(display size)，这也是一个实验变量。靶子和干扰项在画面上的位置每次也是不同的。实验既用含有靶子的刺激卡，也用不含有靶子的刺激卡。每次试验之前将靶子告诉被试，然后用速示器呈现一张刺激卡，要求被试搜索靶子，按键作出“有”或“无”的反应，记录反应时。结果发现，在 b 卡中搜索⊙要显著地快于在 a 卡中搜索○。从反应时的最优拟合直线来看，⊙靶子的“有”反应的斜率为 4.0 毫秒/项目，“无”反应的斜率为 2.9 毫秒/项目，二者的差别很小，同时画面大小的效应也不大，即反应时未随画面中的项目增多而相应增加；○靶子的“有”反应的斜率为 19.7 毫秒/项目，“无”反应的斜率为 38.9 毫秒/项目，二者相差几近一倍，即“有”反应与“无”反应的斜率之比接近 1:2，而且画面大小的效应是很大的。

Treisman 对这些实验结果提出自己的解释。她认为，在 b 卡

中搜索(),只需判断在画面中有一无一条竖直线就可作出反应,而不必考虑竖线在哪个位置或与哪个圆相交。这种搜索属前注意加工的快速过程,是以平行方式实现的。然而,在a卡中搜索(),则需要对画面上的每个项目依次进行扫描,以判断哪个圆不与竖直线相交,即需要将注意依次集中于有关的位置。这种搜索属于集中注意阶段的慢速加工,是以系列方式实现的。同时,从这个结果还可以看到,既然两个靶子的差别仅在于有无一条竖直线,并且有竖直线的靶子的搜索快于无竖直线的,那么就可以说,前注意阶段中的特征地图的活动原则是表征“特征有”,而不表征“特征无”。这是一个很有趣味的看法,对进一步研究前注意阶段加工是有方法学意义的。依照这个活动原则,通过两个靶子的比较,也就是通过靶子和干扰项的比较,就可以查明究竟哪些特征可得到前注意加工,从而揭示前注意加工阶段的特点。其他种类的非对称性搜索也是这样。

2. 连续维量的数量多少性特征的非对称性

这一类非对称性搜索的一个很好的例子,就是前面提到的从短的直线中搜索长的直线快于从长的直线中搜索短的直线。现在可举另一个实验为例加以说明。该实验应用的刺激卡也分a,b两类,见图9。图中a卡以单线为靶子,双线为干扰项,b卡则相反。实

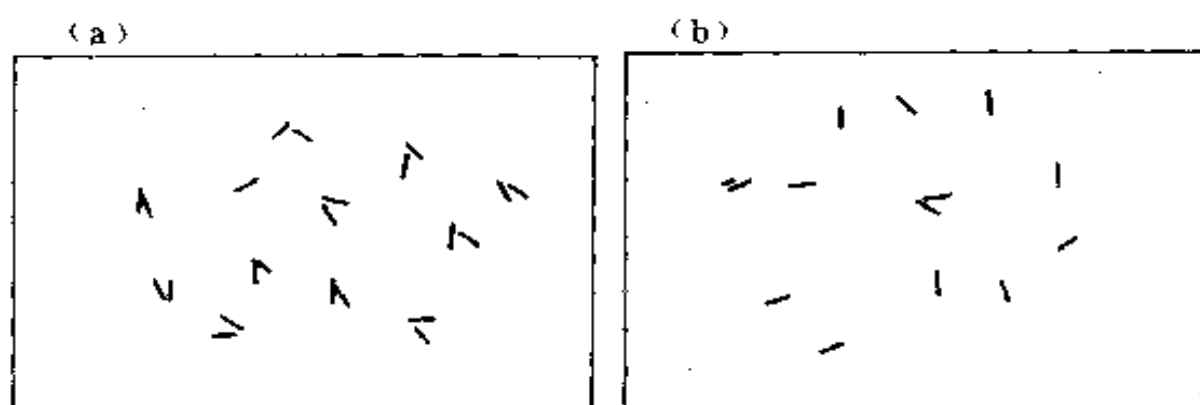


图9 数量多少性特征的搜索实验刺激卡

类,见图9。图中a卡以单线为靶子,双线为干扰项,b卡则相反。实

验程序与以前相同。得到的结果与长短线的实验也相同,从单线中搜索双线靶子要比从双线中搜索单线靶子需时较少。Treisman 认为,从数值小的干扰项中搜索数值大的靶子类似于前面提到的判断“特征有”;而相反的情况则类似于判断“特征无”。这样看来,特征的“有”和“无”可看作连续维量连续体的两个端点,而“多”或“少”则是这一连续体的中间点。由此可以引出特征地图的又一活动原则,即特征地图对连续维量是表征“特征多”,而不表征“特征少”。

3. 标准特征与偏离特征的非对称性

这里所说的标准特征是指原型,而偏离原型的则是偏离特征。如将直线看作原型,则曲线是偏离特征;同理,竖线作为原型,斜线则为偏离特征,等等。实验确定,偏离特征的搜索要快于标准特征,出现搜索的非对称性。图 10 是以竖线和斜线制作的刺激卡,a 卡中以竖线为靶子,斜线为干扰项,b 卡与此相反。结果表明,斜线靶子的搜索比直线靶子要快。

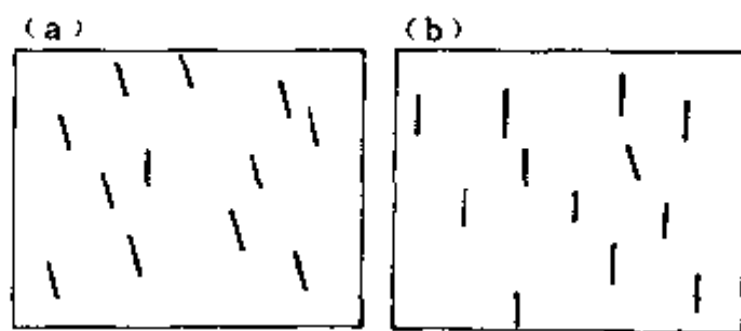


图 10 标准特征与偏离特征的搜索实验刺激卡

对于这种标准特征与偏离特征的搜索非对称性,Treisman 利用觉察器的激活分布予以解释,提出了一个可能的模型,见图 11。在这个模型中,对标准特征和偏离特征有相应的标准觉察器和偏离觉察器;标准特征刺激在标准觉察器中有最大的激活,而在偏离

觉察器中的激活几乎为零；偏离特征刺激则在偏离觉察器中有最大的激活，并在标准觉察器中也有较高的激活，尽管这后一点低于标准特征在标准觉察器中的激活，但却高于标准特征在偏离觉察器中的激活。所以，当搜索偏离特征时，由于标准特征（干扰项）在偏离觉察器中的激活很低，偏离特征极易从中“凸现”（pop-out）出来，因而搜索得快。但是，当搜索标准特征时，由于偏离特征（干扰项）在标准觉察器中仍处于较高的激活水平，与标准特征的激活水平相差不大，标准特征不易“凸现”出来，故搜索得慢。这个模型提出的解释还是假设性的，有待于进一步研究。从这类非对称性搜索实验结果却可引出特征地图的另一活动原则，即表征“偏离特征”，而不表征“标准特征”。

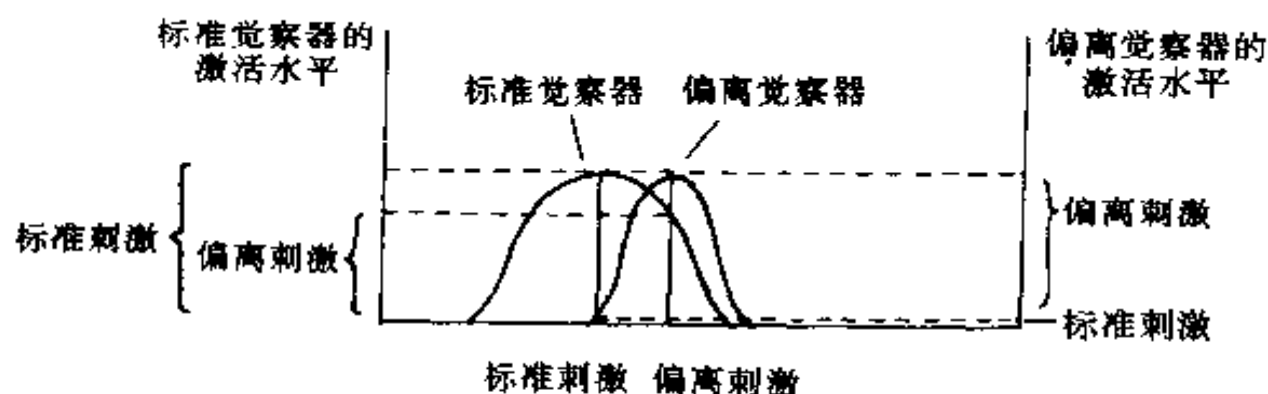


图 11 标准觉察器和偏离觉察器的特征激活分布模型

4. 拓扑特征与非拓扑特征的非对称性

这个领域的实验结果显得较为复杂，而且结论也不确定。Treisman 曾经用封闭的三角形和角做靶子分别进行搜索实验，所用的刺激卡见图 12。其中 a 卡的靶子为三角形，干扰项为角及直线；b 卡中的靶子为角，干扰项为三角形。实验结果表明，三角形靶子的搜索快于角。这提示封闭性这一拓扑特征是前注意加工的特征。这个结论与前面谈过的三角形错觉实验也是吻合的。但是关

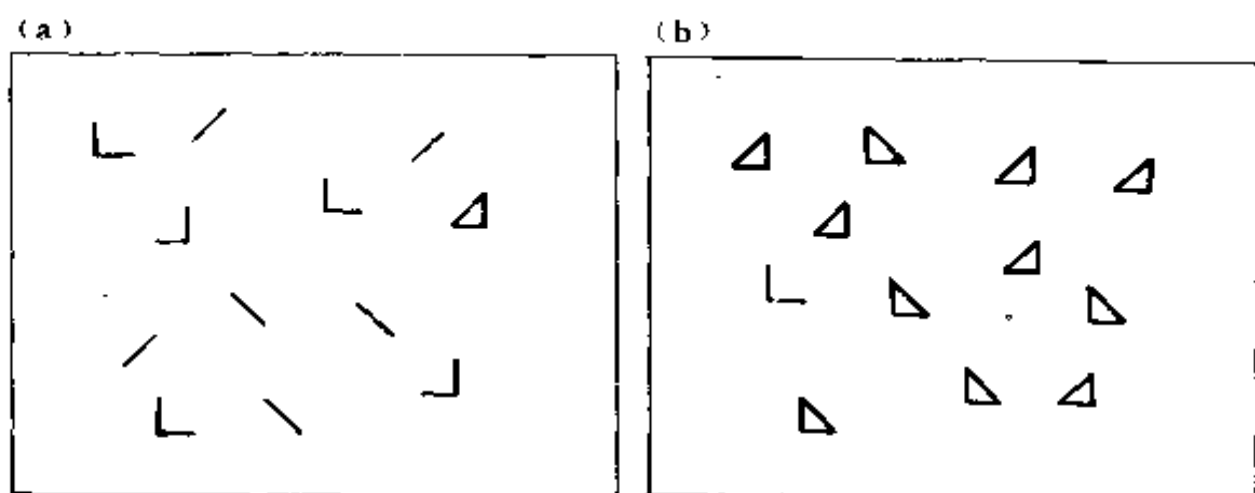


图 12 三角形和角的非对称性搜索实验刺激卡

于封闭性是否为前注意加工的特征,早已存在争论。Julesz(1981)根据他的质地分离实验结果否定封闭性是前注意加工的特征或知觉的基本单元。他认为自由线段的终端或终端子(terminater)是前注意加工的特征,它们在质地分离和搜索中有重要作用。终端子后来被称作质地子(texton)。以图 12 为例,直线和角都有两个终端子,三角形则没有终端子。如果要用终端子来说明上述三角形和角的搜索非对称性,那就意味着,从有终端子的干扰项中搜索无终端子的靶子要快于相反的条件。这种说法还缺乏充分的理由,而且它与特征地图在前注意阶段表征“特征有”的原则也是矛盾的。

然而,对估价封闭性特征来说,最重要的莫过于存在着与上述三角形实验相反的实验结果。Treisman 曾经应用封闭圆和开口圆做靶子分别进行搜索实验,所用刺激卡见图 13 的上半部。从图中可以看出,这些刺激卡是经过精心设计的。将封闭圆和开口圆作为靶子,就可以研究封闭性和线段终端两类不同性质的特征。并且开口又分 3 种大小,图中标出的 $1/2$, $1/4$, 和 $1/8$ 是指开口占圆周长的比例,目的是研究开口的不同大小对搜索的影响。这个实验的程序等都与以前实验相同。

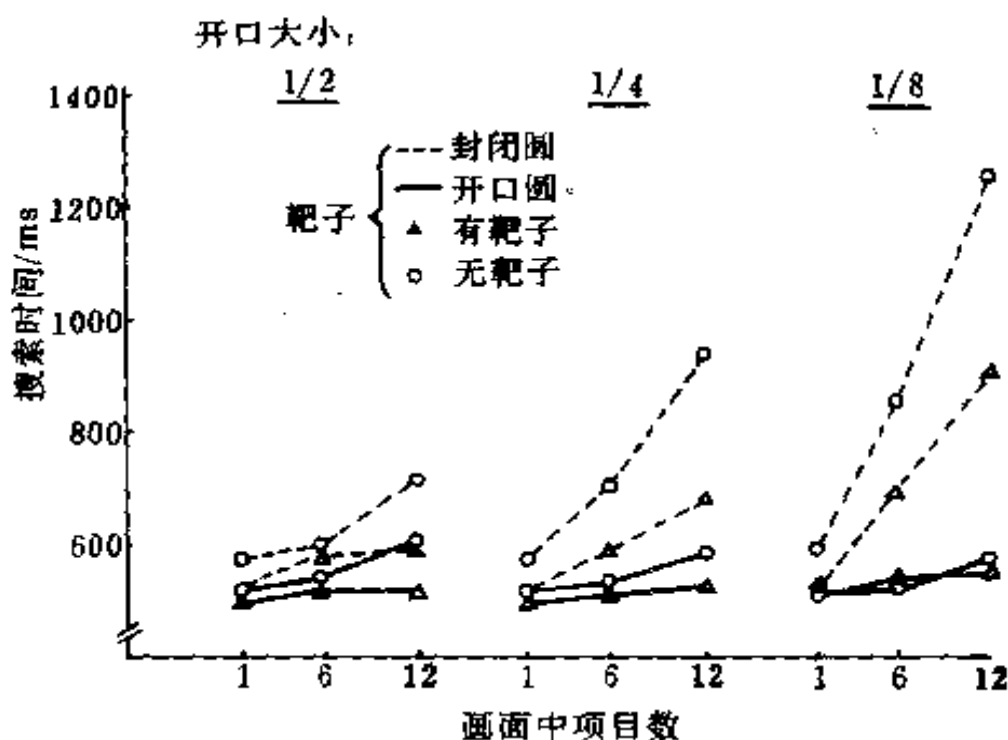
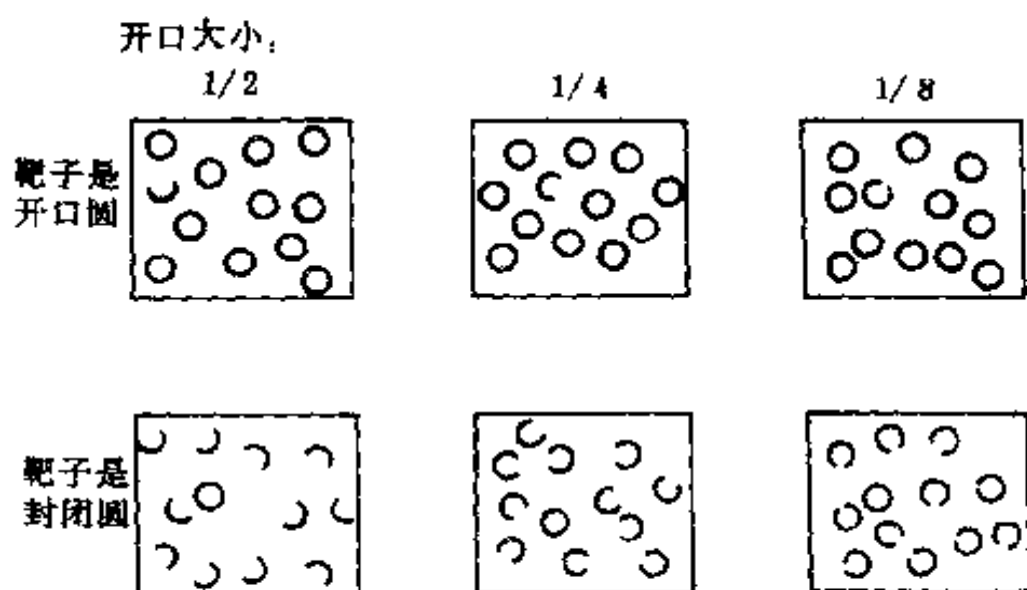


图 13 封闭圆和开口圆的非对称性
搜索实验的刺激卡和结果

实验结果见图 13 下半部。它们清楚地表明,对这两类靶子的搜索存在着强烈的非对称性。开口圆的搜索是快速的,不依赖开口

大小,反应时受画面大小的影响也很小,有靶子和无靶子的斜率相近,并且开口大小与画面大小没有相互作用。但是,封闭圆的搜索却是较慢的、系列的,其搜索速率受干扰项的开口大小的强烈影响,在开口较小($1/8$)和中等($1/4$)时,搜索时间与画面大小有明显的线性关系,且有靶子的斜率与无靶子的斜率之比接近 $1:2$ 。在开口较大($1/2$)时,封闭圆的实验结果不太明朗,有靶子的曲线在画面上的 6 个项目和 12 个项目之间是平坦的。尽管如此,总的看来,开口圆的搜索要快于封闭圆的搜索。

这个实验结果与三角形实验相反。将这两个实验的结果进行比较,就会出现十分复杂的局面。三角形搜索快于角,这提示封闭性是前注意加工的特征,但封闭圆搜索慢于开放圆又否定封闭性是前注意加工的特征。若用没有终端子来说明三角形搜索得快,但这又不能说明同样不具有终端子的封闭圆,而具有两个终端子的开口圆反而可以更快地从无终端子的封闭圆中搜索出来。这使人感到困惑。Treisman 也尝试从其他方面来分析这些矛盾的结果。她设想三角形可能在某个其他的简单特征上有别于角或线段。针对开口圆和封闭圆的实验结果,她提出这样的看法:开口圆具有的线段终端可在前注意阶段被觉察,因此开口圆可被快速搜索,不依赖画面中项目数和开口大小;封闭性可看作封闭程度的连续体,可在不同程度上被封闭圆和开口圆共有,当二者差别大时(干扰项开口的比例为 $1/2$),封闭圆较易搜索,而开口小时搜索就慢了。如果撇开三角形实验,这种看法不失为一种假设,Treisman 还做过关于包容性(在内、在外)的实验,这涉及另一个拓扑特征。总的来说,这个领域的研究还很不足,目前难以获得确定的结论。

以上扼要地介绍了 Treisman 所做的几种非对称性搜索实验。Treisman 在这个领域的实验研究中还发现了其他一些有意义的现象。她发现搜索的非对称可在一定条件下消失,例如,前面提到过的长直线靶子和短直线靶子的搜索非对称性可在一定条件下消

失。Treisman 做了一个实验,在一个画面中用作靶子的直线长 5mm,在另一个画面中的靶子长 10mm,但在两个画面中用作干扰项的直线都是长 7.5mm。这样一来,在前一个画面中是从长直线干扰项中搜索短直线,在后一个画面中则相反。结果发现,在这种条件下两个靶子的搜索时间没有显著差异,未出现搜索非对称性,这使人感到惊奇。Treisman 将韦伯定律引入数据的分析。如要计算刚才说的两个画面中的长直线、短直线靶子分别与其干扰项的差值的韦伯分数,那么就可以看到这两个画面的韦伯分数是一样的:

$$\frac{7.5-5}{7.5} = \frac{10-7.5}{7.5}$$

二者都是 2.5/7.5 即 1/3,这一点跟以前的实验是不同的。在以前出现搜索非对称性实验中,两个靶子互为对方的干扰项,两个画面的靶子和干扰项的长度差的绝对值相同,但韦伯分数不同。Treisman 由此认为,搜索的非对称性不仅取决于靶子与干扰项的差别的方向,如多对少等,还取决于这个差别与干扰项背景的关系。将韦伯定律引入搜索非对称性的研究无疑是有益的尝试,其意义和应用范围等问题都需更进一步研究。

Treisman 还发现,替代性特征搜索没有非对称性。所谓的替代性特征是指某个维量的一个特征未出现时可被该维量的另一特征的出现所取代,如颜色维量的各个特征(值),即红、黄、绿、蓝等各种颜色就是替代性特征。实验表明,搜索一个原来的特征与搜索一个取代它的特征都是快速的,例如,在一个研究中应用的刺激是红色的、绿色的和黑色的 O 和 X,将实验分成两组:一组要求被试在红色或黑色刺激背景中搜索绿色,另一组要求被试在绿色刺激背景中搜索非绿色即红色或黑色。结果两种条件下的搜索都是快速而平行的,没有非对称性。这种替代性特征在该维度上的表征是等值的,没有多与少、有与无或原型与偏离的差别。Treisman 还注

意到,这两组实验中的指示语--为正性的(搜索绿色)、一为负性的(搜索非绿色),结果未出现非对称性,这说明言语码的非对称性不会导致搜索的非对称性。

从以上的介绍中可以看出,非对称性搜索的实验研究是有理论意义的。它们首先证实了前注意加工阶段的存在,以及哪些基本特征可在前注意阶段得到加工,特别是确定了前注意阶段的特征地图的几个活动原则。其次,它们表明干扰项的特点对某些靶子的搜索速率有影响,这提示搜索不是孤立地寻找靶子,而是在靶子与干扰项的相互关系中进行的。这一点涉及的问题也是很多的。总起来看,关于搜索非对称性的实验研究加深了对前注意阶段加工的理解,并把两个加工阶段的相互关系突出出来了。

除视觉搜索实验等以外,Treisman 还做过一些质地分离实验。这种实验和视觉搜索实验相近。它的通常做法是在刺激画面上的左右两部分安排不同的项目。这两个部分的项目可以仅在某一个简单特征上不同,如在颜色、形状、线段朝向或大小上不同。这时两部分的边界显得“凸现”出来,可迅速得到觉察。这说明那些简单的特征可在早期的前注意阶段得到加工,但是,特征结合的边界却不会“凸现”出来。假如画面的一部分是红 V 和蓝 O,另一部分是红 O 和蓝 V,那么觉察边界需时较多。这说明特征结合不是早期阶段可以加工的。这些结果和视觉搜索实验也是一致的。

总之,前面介绍的错觉性结合实验、非对称性搜索实验等都一致地支持特征整合论。它们不仅为这个理论的框架或原则提供了实验依据,而且逐步深入到各个阶段的内部加工机制,其研究构思富有特色,实验方法也经过精心设计。这些也都吸引许多心理学家,使特征整合论更令人瞩目。

四、理论的修正和争议

特征整合论提出以后即以其思想新颖而受到关注。到了 80 年代中期,它的影响逐渐扩散开来,引起一些心理学家从事相应的研究。在最近的 10 年间,Treisman 做了大量的实验来支持其理论,而其他一些心理学家也得到了不同于 Treisman 的实验结果,呈现出研究工作比较活跃的局面。它既推动研究工作进一步深入,也促使 Treisman 对她的理论作了一些修正,并引发了理论上的论争。

80 年代后期,出现了一系列与特征整合论不相吻合或矛盾的实验结果。问题的焦点在于,依照特征整合论,单个特征是快速、平行搜索的,属前注意加工,而客体或结合特征是慢速、系列搜索的,属集中注意阶段加工;但是,其他人的一系列实验表明,结合特征也有快速、平行的搜索。例如,Nakayama 和 Silverman(1986)发现,双眼视差和颜色的结合或双眼视差与运动的结合都有平行搜索。McLeod 等(1988)发现,形状和运动方向的结合几乎是以平行方式搜索的。Steinman(1987)发现,双眼视差和朝向的结合也是这样。Wolfe 等(1989)发现,形状和颜色的结合以及其他的特征结合均有平行搜索。

在这方面,还应提到 Pashler(1987)的实验。他要求被试在绿 O 和红 T 中搜索绿 T,画面大小分 2,4,8,16 和 24 个项目共 5 种。结果发现,在画面为 2—8 个项目时,有靶子和无靶子的反应时曲线的斜率分别为 13.3 和 16.4,二者的比率为 1:1.24。这提示搜索几乎是平行的。但当画面大小为 8—24 个项目时,有靶子和无靶子的反应时曲线的斜率分别为 11.7 和 24.3,二者比率为 1:2.07,这说明对靶子的搜索是系列的,严格地说,是自我停止的系列搜索。这也与特征整合论的预测相一致。问题在于为什么在 2—8 个项目中会出现结合靶子的快速平行搜索。Pashler 认为,这是

由于被试在项目少和项目多两种情况下采取了不同的加工策略。他提出了分组搜索的概念,认为被试在搜索时会按 8 个项目来将全部项目分组,一组组地依次进行搜索,而在组内则是平行搜索的。虽然这个观点还需要得到进一步的实验验证和支持,但这种实验结果却不是特征整合论可以解释的。

Egeth 等(1984)做过一个很有意思的实验。他们要求被试在红色 N 和黑色 O 中搜索红色 O。呈现的项目分 5,15 和 25 个 3 种。实验分两种条件:第一种是在所有呈现项目中始终保持红色数字为 3 个,即在 5,15 和 25 个项目中分别有 2,12 和 22 个黑 O;第二种即为通常的结合靶子实验,即两种干扰项(黑 O 和红 N)的数目保持相等。结果发现,第一种条件下的有靶子和无靶子的斜率之比接近 1:1,而第二种条件下的相应的比率则接近 1:2。之所以出现这种结果,很可能是由于被试在第一种条件下只需在红色 N 中进行搜索,而可完全撇开黑色 O,出现选择性搜索或范畴效应,表现为快速的平行搜索。但在第二种条件下却不可能这样,出现通常的系列搜索。应当指出,特征整合论也可以解释第一种条件下的快速的平行搜索。在红 N 中搜索红 O,由于靶子和干扰项有共同颜色,搜索红 O 就变成搜索 O 了,单个特征的搜索是前注意阶段可以完成的,自然搜索就很迅速了。不过,这样的解释是以撇开另外一类干扰项为前提的,而搜索时要实现这一点却需要相应的加工,甚至需应用策略,这就涉及到自下而上加工和自上而下加工的联系了。

针对结合性靶子可有平行加工的实验结果,Wolfe 等(Wolfe 等,1989;Cave and Wolfe,1990)吸收了特征整合论的思想,提出了自己的模型,称作引导搜索模型(guided search model),见图 14。这也是一个包含早期的平行加工阶段和随后的系列加工阶段的两阶段模型。这个模型吸收了特征整合论的特征地图概念,也认为刺激的各种简单特征在平行阶段可分别表征于相应的特征地图

中。但这些特征在地图的位置上有不同的激活值,其大小受该刺激特征与靶子的相应维量特征的相似程度及与其他刺激的不同程度所影响。也就是说,在每个特征维量上,平行阶段确认那些与靶子的相应维量值最接近,且与画面中其他刺激不相同的刺激。这样,每个刺激的各个特征维量的信息或激活值再汇总到一起,就产生一个激活地图(activation map)。它表明各个位置上的刺激的激活值,反映着靶子在该位置的可能性大小。激活地图可引导注意,也即引导系列阶段对靶子的搜索,犹如一个注意地图。在系列阶段,注意从激活地图中有最大激活值的位置开始,按照激活值大小依次对这些位置作出核查,来搜索靶子,并将信息再反馈给激活地图,调整其激活值。如果某个刺激被确定不是靶子,这个刺激以后将不再加以考虑。此外,激活值也会受到噪音的干扰而减少或发生波动。由于激活地图的作用,这个模型就可以说明结合性靶子的快速搜索了。Wolfe 等曾经举过一个例子来说明这个过程,见图 15。如图所示,靶子是红色垂直线,干扰项为红色水平线和绿色垂直线;靶子在刺激画面的中心,其余为干扰项。与靶子的红色和垂直两个特征维量值最接近的诸干扰项的特征,都分别在颜色地图和朝向地图中得到表征(分别在相应位置用线标出来)。假设红色和垂直两个特征有相同的激活值,那么在汇总这些激活值的激活地图或注意地图上,靶子所在位置的激活是红色和垂直两个特征激活值的叠加,而干扰项所在的每一位置上却只有一个红色或垂直特征的激活值。在整个激活地图中,以处于中心位置的靶子有最大的激活值,它引导注意首先将靶子搜索出来,表现为对结合性靶子的快速平行搜索。这个模型在两个重要方面有别于特征整合论。其一,它将激活概念引入前注意的平行阶段,认为各个刺激不是等值的。其二,它强调前注意阶段对注意的引导。这两点后来对 Treisman 修正其特征整合论产生积极的影响。

面对许多业已确定的结合性靶子快速搜索的实验事实和已提

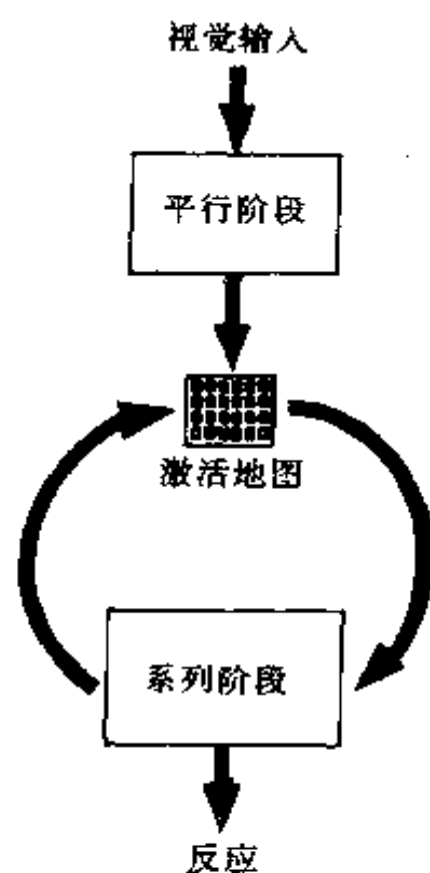


图 14 引导搜索模型

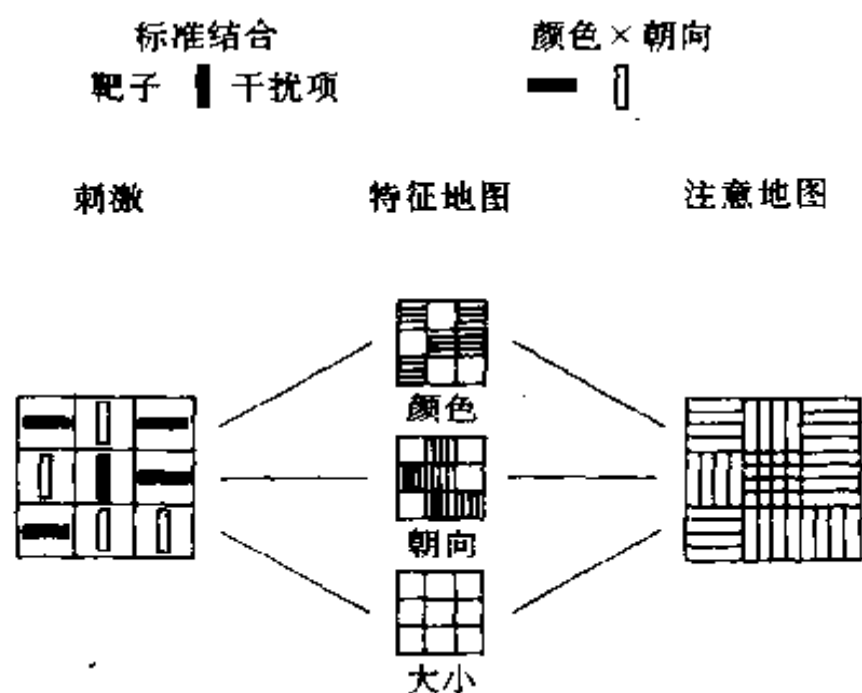


图 15 结合性靶子的快速搜索

出的一些新模型,Treisman(1988,1990)开始对其理论进行修正。这种修正主要是针对前注意阶段和集中注意阶段的关系。所涉及的具体问题也是关于前注意阶段的表征和它对注意阶段的作用。在这一点上,Treisman 显然受到 Wolfe 等人的影响,但着眼点却完全不同。Wolfe 等人在引导搜索模型中着眼于靶子的激活,然而 Treisman 则着眼于干扰项的抑制。她在原来理论的框架内提出特征抑制功能,以此为核心来修正某些内容。原来的理论认为,客体知觉或结合特征加工是通过集中注意对位置地图的系列扫描而实现的,扫描到哪个位置,与它有联系的特征就被激活或提取并被整合(参看图 1)。在这里,特征地图似乎只是一个数据库,为系列阶段的注意加工提供必要的数据库。现在 Treisman 认为,对位置地图上的某个位置的选择也可通过另一条途径实现控制。这就是特征地图对位置地图上的一些位置产生抑制,减弱其活动。进一步说,当干扰项的各特征地图对位置地图上的干扰项位置产生抑制,那么在位置地图上靶子所在的那个位置就会有高过其他位置的激活,干扰项位置所产生的干涉作用将减弱。这样,注意扫描的速度将加快,且不依赖干扰项数目多少,就会出现结合性靶子的快速平行搜索。Treisman 强调指出,她所说的这种特征抑制是一种用来易化选择性注意的策略,而不是两个相邻的相同特征之间自动发生的互相压制。特征抑制可有不同的范围,而且它也是可以逆转的,当注意的目标改变时,就可以解除先前的特征抑制。Treisman 把她关于特征抑制的这些看法称为特征抑制假设,并将原先的模型(图 1)加以修正,在模型中增加了从特征地图对位置地图起抑制作用的联系,见图 16。这个图没有画出全部联系,只画出一个颜色地图的特征抑制作用。图中假设的靶子是左右运动的绿色线条,干扰项是上下运动的绿色线条和左右运动的红色线条。将图 16 和图 1 加以比较,就可以看出特征整合模型发生的变化。

Treisman 的特征抑制假设与 Wolfe 等人的引导搜索模型使

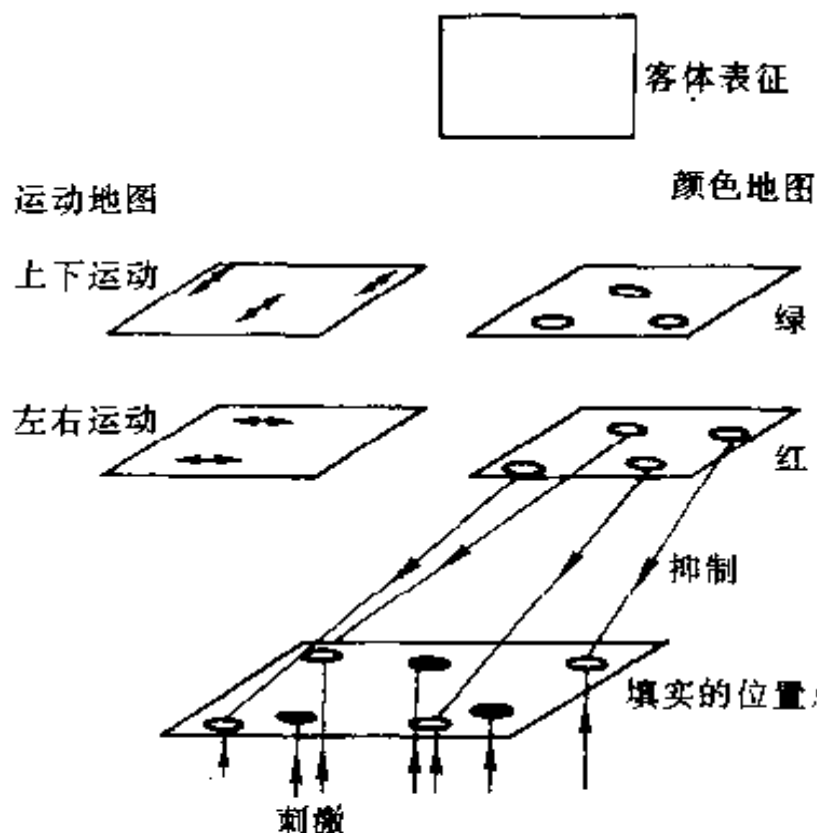


图 16 特征抑制示意图

人有殊途同归之感。它们甚至使人觉得,这两个模型在说明结合性靶子的快速平行搜索上只是表述的方式不同而已。但是,Treisman(1990)认为她的观点与 Wolfe 等人有原则性的区别。她做了一个实验,来考验这两个模型。这个实验还是视觉搜索实验,方法和程序与以前相同,但在实验条件上做了细致安排。实验要求被试搜索的靶子是倾斜 27° 的绿色线条。实验条件分两种:一种有两类干扰项,即倾斜 63° 的绿色线条和倾斜 27° 的灰色线条;另一种有四类干扰项,除前述两类以外,还有倾斜 27° 的粉红色线条与 90° 的绿色线条。呈现项目分 4 个,9 个和 16 个 3 种。Treisman 分别按上述两个模型对实验的结果作出预测:从引导搜索模型来看,第二种条件下的搜索应与第一种条件一样快,因为第二种条件下的额外两类干扰项比原来的两类(第一种条件)更少与靶子相似,不

会影响靶子特征的激活;从特征抑制假设来看,第一种条件下的搜索应比第二种条件要快,因为第二种条件多了两类干扰项,不同的特征愈多,抑制就愈难进行(特征抑制不是自动进行的)。实际得到的实验结果见图 17,第一种条件下的靶子搜索比第二种条件要快,符合特征抑制假设。

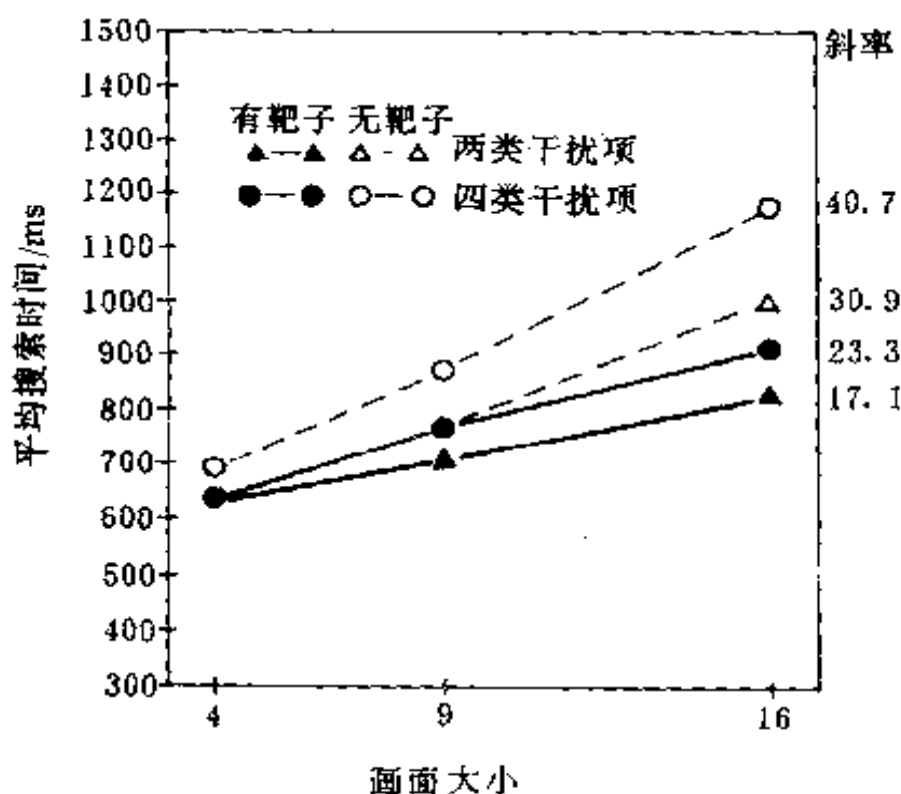


图 17 两类和四类干扰项的比较

Treisman 将特征抑制假设引进特征整合论,从而使它也可以说明结合性靶子的快速平行搜索。这给特征整合论带来不小的变化。首先,前注意阶段被赋予新的抑制功能,给模型增添了抑制过程。特征抑制反映了前注意阶段也存在着一种选择作用,Treisman 称之为前选择(preselection)。其次,与此有关,前注意阶段与集中注意阶段之间有了双向联系。以前的模型(图 1)只有从位置地图到特征地图的连线,表明特征提取;现在的模型(图 17)则还

有特征地图到位置地图的连线,表明特征抑制。这两点是 Treisman 对特征整合论作出的主要修正。它们表明,经过修正之后,前注意阶段与集中注意的特征整合阶段不再是截然分开的了。正如 Treisman(1988)所说的那样,以前的前注意和集中注意的两分法被另一种看法取代了,即认为注意是一个从广泛的分散注意到狭窄的集中注意的连续体。这对于一个注意理论来讲,不能不说是一个重大的修正。但是,特征抑制假设尚未说明特征地图和位置地图之间信息传递的动力过程,现在还不能说,经过修正的特征整合论已不再具有离散阶段模型的性质。公平地说,特征抑制假设引入特征整合论显得比较自然,它既保留了原来的理论框架,又使它变得更加灵活。当然,留下的问题还是很多的,特别是关于两个阶段的动力关系以及由此而来的每个阶段的内部关系问题。

特征整合论也受到批评,引起争议。在特征抑制假设提出来以前,对它的批评和争议主要是围绕结合性靶子的快速搜索。随着特征抑制假设的提出,一些争论的问题,其中特别是有关前注意阶段和集中注意阶段划分的争论(Tsal,1989)得到一定程度的澄清。其他的批评和争论仍然值得重视。Briand 和 Klein(1989)指出,特征整合论缺乏对特征的严格定义。这方面的论述往往忽视神经心理学和心理物理学已获得的关于早期视觉本质的事实。Tsal(1989)提出,特征整合论没有说明注意怎样将特征整合起来。这确实是一个重大问题。注意的整合功能在特征整合论中占有重要地位,并使它具有特色。但是 Treisman 强调的是注意对位置地图的扫描,扫描到哪里,那个位置上的诸特征被提取而得到整合,好像整合是自动的。Treisman 承认注意既受到自下而上加工又受到自上而下加工的影响。这种影响似不应只限于注意扫描,而也要包括整合在内。整合过程若得不到说明,特征整合论的立足就不会稳。Navon(1990)则认为,注意不是 Treisman 所说的“胶水”,而是一个“试管”(test tube),它划出一个工作空间的界限,排除其他认知活动

的干扰。

在所有对特征整合论的批评中,以 Navon(1990)的批评最直截了当,也最严厉。Navon 认为特征整合论是不能成立的。他对这个理论进行了逻辑分析并假设了其他模型与之比较。他提出的一个论点值得注意。他指出,依照特征整合论,特征是明显的即可直接加以利用,而位置则是潜在的,需要通过与之有关的信息才能加以利用。在搜索颜色和形状的结合靶子时,颜色和形状的信息可直接运用,但需要在位置地图上依次寻找该颜色的位置和该形状的位置,直到二者匹配为止。这种搜索应是在刺激项目 N^2 的方阵中进行,因此搜索时间将是刺激项目 N 的二阶函数,然而 Treisman 得到的实验结果却是一阶函数,所以特征整合论不能成立。Treisman(1990)对 Navon 提出反批评。她认为,一旦集中注意扫描到位置地图上一个位置,那里的特征信息就可通过与特征地图的联系而取得,因此搜索一个结合性靶子只需在位置地图上扫描 N 个位置,而不是 N^2 个位置。这与实验结果也是吻合的。Treisman 还指出,她所说的“胶水”与 Navon 的“试管”只是字面上的差别,二者的实质是一样的,都是排除其他项目干扰。

从以上所述可以看到,Treisman 的特征整合论自提出以来,在十年内既得到实验结果的充实,也受到与之不相吻合的实验结果的挑战,而且存在着有时十分激烈的争论。在这些因素的推动下,Treisman 提出特征抑制假设,对它进行了修正。这个理论对注意和知觉研究作出的贡献是无可怀疑的。可以预料,特征整合论还将进一步发展。它必将继续推动注意和知觉研究的发展。这也正是一个理论所起的重要作用。

(张奇萍协助整理)

主要参考文献

- [1] Briand, K. A., & Klein, R. M. Has feature integration theory come

unglued? A reply to Tsal. *Journal of experimental psychology: Human perception and performance*, 1989, **15**, 394—400.

- [2] Broadbent, D. E. *Perception and Communication*. London: Pergamon Press, 1958.
- [3] Cave, K. R. , & Wolfe, J. M. Modeling the role of parallel processing in visual search. *Cognitive psychology*, 1990, **22**, 225—271.
- [4] Egeth, H. E. , Virzi, R. A. , & Garbart, H. Search for conjunctively defined targets. *Journal of experimental psychology: Human perception and performance*, 1984, **10**, 32—39.
- [5] Julesz, B. Textons, the elements of texture perception and their interaction. *Nature*, 1981, **290**, 91—97.
- [6] Julesz, B. Textons, the fundamental elements in preattention, vision and perception of textures. *The Bell system technical journal*, 1983, **12**, 1619—1695.
- [7] Julesz, B. A brief outline of the textons theory of human vision. *Trends in Neuroscience*, 1984, **7**, 41—45.
- [8] McLeod , P. , Driver, J. , & Crisp, J. Visual search for a conjunction of movement and form is parallel. *Nature*, 1988, **332**, 154—155.
- [9] Nakayama, K. , & Silverman, G. H. Serial and parallel processing of visual feature conjunctions. *Nature*, 1986, **320**, 264—265.
- [10] Navon, D. Does attention serve to integrate features? *Psychological review*, 1990, **97**, 453—459.
- [11] Navon, D. Treisman's search model does not require feature integration; Rejoinder to Treisman. *Psychological review*, 1990, **97**, 464—465.
- [12] Neisser, U. Visual search. *Scientific American*, 1964, **210**, 94—102.
- [13] Neisser, U. *Cognitive psychology*. Englewood Cliffs, NJ: Prentice-Hall, 1967.
- [14] Pashler, H. Detecting conjunctions of color and form; Reassessing the serial search hypothesis. *Perception and psychophysics*, 1987, **41**, 191—201.

- [15] Schneider, W. ,& Shiffrin, R. M. Controlled and automatic human information processing: Detection, search and attention. *Psychological review*, 1977, **84**, 1—66.
- [16] Steinmen, S. B. Serial and parallel search in pattern vision. *Perception*, 1987, **16**, 389—399.
- [17] Treisman, A. Contextual cues in selective listening. *Quarterly journal of experimental psychology*, 1960, **12**, 242—248.
- [18] Treisman, A. Verbal cues, language and meaning in selective attention. *American journal of psychology*, 1964, **77**, 206—219.
- [19] Treisman, A. Strategies and models of selective attention. *Psychological review*, 1969, **76**, 282—299.
- [20] Treisman, A. , Sykes, M. ,& Gelade, G. Selective attention and stimulus integration. In *Attention and performance VI*, S. Donic (Ed.), 336—361. Hillsdale, NJ: Erlbaum, 1977.
- [21] Treisman, A. Perceptual grouping and attention in visual search for objects and feature. *Journal of experimental psychology: Human perception and performance*, 1982, **8**, 194—214.
- [22] Treisman, A. ,& Schmidt, H. Illusory conjunction in the perception of objects. *Cognitive psychology*, 1982, **14**, 107—141.
- [23] Treisman, A. ,& Paterson, R. Emergent features, attention, and object perception. *Journal of experimental psychology: Human perception and performance*, 1984, **10**, 12—31.
- [24] Treisman, A. Preattentive processing in vision. *Computer vision graphics and image processing*. 1985, **31**, 156—177.
- [25] Treisman, A. ,& Souther, J. Search asymmetry: A diagnostic for preattentive processing of seperable features. *Journal of experimental psychology: General*, 1985, **114**, 285—310.
- [26] Treisman, A. Features and objects in visual processing. *Scientific American*, 1986, **255**, 114B—125.
- [27] Treisman, A. Properties, parts and objects. In *Handbook of perception and human performance*, K. Boff, L. Kaufman, & J. Thomas (Eds.),

New York: Wiley, 1986.

- [28] Treisman, A. Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly journal of experimental psychology*, 1988, **40A**, 201-237.
- [29] Treisman, A. , & Gormican, S. Feature analysis in early vision; Evidence from search asymmetries. *Psychological review*, 1988, **95**, 15-48.
- [30] Treisman, A. Conjunction search revisited. *Journal of experimental psychology: Human perception and performance*, 1990, **16**, 459-478.
- [31] Treisman, A. Variations on the theme of feature integration; Reply to Navon. *Psychological review*, 1990, **97**, 460-463.
- [32] Tsai, Y. Do illusory conjunctions support the feature integration theory? A critical review of theory and findings. *Journal of experimental psychology: Human perception and performance*. 1989, **15**, 394-400.
- [33] Tsai, Y. Further comments on feature integration; A reply to Briand and Klein. *Journal of experimental psychology: Human perception and performance*, 1989, **15**, 407-410.
- [34] Wolfe, J. M. , Cave, K. R. , & Frazel, S. L. Guided search; An alternative to the FIM for visual search. *Journal of experimental psychology: Human perception and performance*, 1989, **15**, 419-433.

启动效应——无意识的记忆

朱 滢

(北京大学心理学系)

启动效应(priming effects)是一种无意识的记忆。启动效应的研究是整个 80 年代记忆研究的“热点”之一,预计在最近的将来,它仍然会吸引着大量的研究者。虽然启动效应研究的原动力不是来自艾宾浩斯,但是作为记忆的实验性研究之父,艾宾浩斯早在一百多年前已经指明了无意识记忆的存在,并给出了一种具体的实验方法。艾氏把记忆分为 3 类,头两类记忆都存在于意识之中。他说,“还有第三类……。已经消逝的心理状态,就是它们自己完全不回到意识中来,……。这些经验的大部分还是在意识外隐藏着,但产生很重要的效果,确证它们以前的存在”^[1]。艾宾浩斯还采用重学法(即节省法)来测量无意识的记忆。假设很久以前你学习过俄语,现在要你写出俄语单字,你是一个也写不出来了,换句话说,你不能有意识地回忆它们。但是,用重学法可以证明,你现在对那些单字仍然是有记忆的。不过,重学法测量的是无意识的记忆,并且它是间接地测量的,即从时间的节省而不是记住多少单字来测量。

现在我们要介绍的启动效应就是在意识外隐藏着,但产生很重要效果的一种记忆。今天测量启动效应的方法多种多样,但它们也同节省法一样,都是对记忆的间接测量。

下面我们分 6 个部分介绍启动效应的研究。

一、启动效应的概念与测量方法

启动效应指的是执行某一任务对后来执行同样的或类似的任务的促进作用。它与正迁移类似,但后者是在总体效果上进行测量的,启动效应的测量单位却十分具体。例如,学了“息、云、国……”等单字后,把它们与未学过的一些单字随机混合,在速示器上逐个以极短时间(5 毫秒以下)呈现,要求被试读出所见到的单字,典型的实验结果是,“息”、“云”、“国”等学过的字读出来的概率高于未学过的字,两者的概率之差就是对启动效应的测量,这就是测量启动效应的“知觉辨认”的方法(perceptual identification)。在这里,启动效应是由逐个对应于学习阶段的单字、图画等材料累计得来的。

启动效应又分直接启动(direct priming)与间接启动(indirect priming)。在直接启动中,测验项目与学习项目是相同的,例如在刚才的例子中,速示器上呈现的“息”字是被试学习过的;而在间接启动中,测验项目与学习项目有关,但并不相同。例如,学习了“医生”单词后,在测验中会使“护士”的反应时缩短。直接启动又称为重复启动(repetition priming),因为同一项目在学习和测验中都出现过。目前记忆研究中未作特别说明的启动效应就是指直接启动,我们以下讨论的启动效应也是指直接启动。

启动效应的测量方法除上面说到的知觉辨认外,还有词干补笔(word stem completion),补笔(word-fragment completion)等等。

下面以英文为例说明什么是词干补笔测验(参见表 1)。

被试学习一系列单字之后,主试给第 1 被试组提供单字的开头的 3 个字母,要求被试把心中首先想到的单字填出来。从表上我们知道,他们填成 shade, print 就算对;当然,第 2 被试组虽然没有

表 1 两组被试学习字单与词干补笔测验字单

学 习 字 单		测 验 字 单	
第 1 被试组	第 2 被试组	第 1 被试组	第 2 被试组
shade	chair	sha ____ (target)	cha ____ (target)
print	trade	pri ____ (target)	tra ____ (target)
		cha ____ (baseline for)	sha ____ (baseline for)
		tra ____ (baseline for)	pri ____ (baseline for)

学过 shade 和 print, 但给他们提供 sha __ 和 pri __ 的话(见测验字单第 2 被试组一行), 有时他们碰巧也能填对, 即填成 shade 和 print。这种碰巧填对的概率即机遇概率就作为测量启动效应的基础值。总体说来, 第 1 被试组填对 sha __ 与 pri __ 的概率要高于第 2 被试组, 第 2 被试组填对 cha __ 与 tra __ 的概率要高于第 1 被试组。学过字填对的概率减去未学过字填对的概率就叫作启动效应。

补笔测验与词干补笔测验类似, 但在测验中不是提供单字的头三个字母, 而是提供单字的缺笔字, 要求被试把他心中首先想到的单字填出来。例如, 学习 ASSASSIN, 给被试的缺笔字是 A __ A __ IN, 它是随机去掉一些字母构成的, 其余测验程序和启动效应的计算同于词干补笔。汉字的补笔测验与词干补笔测验请参考“加工水平、回想策略与不自觉记忆”一文^[27]。

现在人们把自由回忆和再认等传统的测验记忆的方法叫作自觉记忆测验(explicit memory tests), 其特点是这些测验要求被试有意识地努力去提取信息; 把知觉辨认、补笔和词干补笔等测验记忆的方法叫作不自觉记忆测验(implicit memory tests), 其特点是, 这些测验并不要求被试有意识地努力去提取信息, 只是要求他

们专注于完成眼前的作业就行,如在知觉辨认中,只要求被试读出短暂呈现的单字。

二、健忘症病人(amnesic patients)的实验性分离现象

人们把在同一自变量影响下两种测验产生相反结果的情形称为实验性分离(experimental dissociation)。E. Tulving 曾对实验性分离作如下的描述:“符合实验性分离逻辑的实验是这样的:控制单一的变量而比较在两种不同的测验任务中变量的效应……,如果变量影响被试在一种测验任务中的结果,但不影响另一种测验任务的结果,或者变量对两种测验任务的结果的影响有不同的方向,我们就说分离产生了”^[3]。

我们首先叙述健忘症病人的实验性分离现象。

记忆的神经心理学关于健忘症病人的研究是目前启动效应研究的原动力。健忘症病人由于某种形式的脑损伤而丧失记忆,不过,这些病人的知觉能力、语言和智力基本上保持完好。长期以来,人们相信健忘症病人不能把语词信息从短时记忆传递到长时记忆中去,即他们不能保存新鲜的经验。但是,1970年 Warrington 和 Weiskrantz 的实验改变了人们的看法^[4]。

他们让4位健忘症病人(3位 Korsakoff 综合症病人与一位颞叶切除术病人)与16位控制组病人(无脑损伤)学习字单,然后进行4种记忆测验:自由回忆与再认(现在称作自觉记忆测验一类即 explicit memory tests)、模糊字辨认(word fragment identification,测验时将字迹弄得模糊不清,要求被试说出是什么字)与词干补笔(现在称作不自觉记忆测验一类)。有些结果是预料之中的,健忘症病人在自由回忆与再认的测验中都比控制组差(见表2)。但使人惊讶的是,在其余两项测验中,健忘症病人与控制组的成绩一样好(词干补笔中控制组稍优于病人但差异不显著)。换句话说,健

忘症病人也能保持词语信息,但是这种保持要在特定的测验中才能表现出来。

表2 健忘症病人对语词信息的保持

组别 \ 任务	自觉记忆		不自觉记忆	
	自由回忆	再 认	模糊字辨认	词干补笔
控 制 组	0.55	0.75	0.45	0.69
健忘症病人	0.44	0.45	0.47	0.58

* 在自觉记忆测验中,病人差于控制组,但在不自觉记忆测验中,病人的成绩与控制组同样好。Warrington and Weiskrantz 1970,实验2。

1984年 P. Graf 等人在题为“健忘症病人不会遗忘的信息”的论文中^[52],报告了他们对3种类型的健忘症病人(柯萨柯夫健忘症、酒精中毒柯萨柯夫健忘症以及缺氧脑病病人)进行实验研究的结果。表3是3项实验中在意义加工的条件下(要求被试对所见到的单字作喜欢或不喜欢的5点量表评判),健忘症病人与控制组被试的

表3 3项实验的综合结果

组别 \ 任务	自由回忆	线索回忆	词干补笔
控制组	0.37	0.69	0.49
健忘病人	0.15	0.58	0.57

* P. Graf, L. R. Squire and G. Mandler 1984.

结果。从表上可以看到:①在自觉记忆测验中(自由回忆),病人显著地不如正常人,这表明病人的自觉记忆受到严重损伤;在另一项自觉记忆测验中(有线索回忆),结果是类似的。控制组69.0%与病人组58%在统计上有显著差异。②在不自觉记忆测验中(词干补笔),病人组57%,与正常组49%相比,表示病人的不自觉记忆仍然保留。③比较词干补笔与有线索回忆两种测验中两类被试的结果

可以看到,正常人从49%上升到69.0%,而病人从57%到58%几乎没有变化。这表明在有线索回忆时,当主试明确要求被试以缺笔字(如 sha __)作线索进行有意回想时,正常人大大地提高了成绩;但是病人由于脑损伤而不能回忆,所以主试的指示语并没有发生效力。

关于健忘症病人保留有记忆能力的研究,最初的兴趣只集中在对病人的诊断上,即这些病人是以其很差的自觉记忆与完好的不自觉记忆为特征的。随着研究的深入,人们对同一个人身上自觉记忆测验与不自觉记忆测验之间的分离现象赋予理论上的意义,认为研究这种分离现象是理解人类记忆本质的全新的途径。

三、正常人的实验性分离现象

1982年 Tulving 等人发表了题为“补笔的启动效应独立于再认”的论文^[6],他们用正常人作被试,也重现了在健忘症病人身上表现出来的不一致现象。Tulving 等人让正常人被试学习96个低频英语单字,学习后1小时进行再认测验与补笔测验,7天以后重复这两种测验,实验结果见图1。图1最显著的特点是,情节记忆(episodic memory)7天后遗忘很多(由再认结果显示出来),但由启动效应显示的不自觉记忆7天前后没有差别,即仿佛启动效应不会遗忘一样!另外,实验数据分析表明,测验中正确填对缺笔字的概率与被试自以为是学过的字或自以为是未学过的字无关,换句话说,补笔正确率是独立于再认的。

1983年 Jacoby 关于正常人的实验也十分典型^[7]。在实验的第一部分,他要求被试在3种条件下大声读出一系列视觉呈现的单词或心里想出来的单词。在“无上下关系”的条件下,呈现单词以前先呈现一排符号(××××),因此被试事先得不到关于该单词的信息;在“有上下关系”的条件下,呈现单词以前呈现它的反义词,这

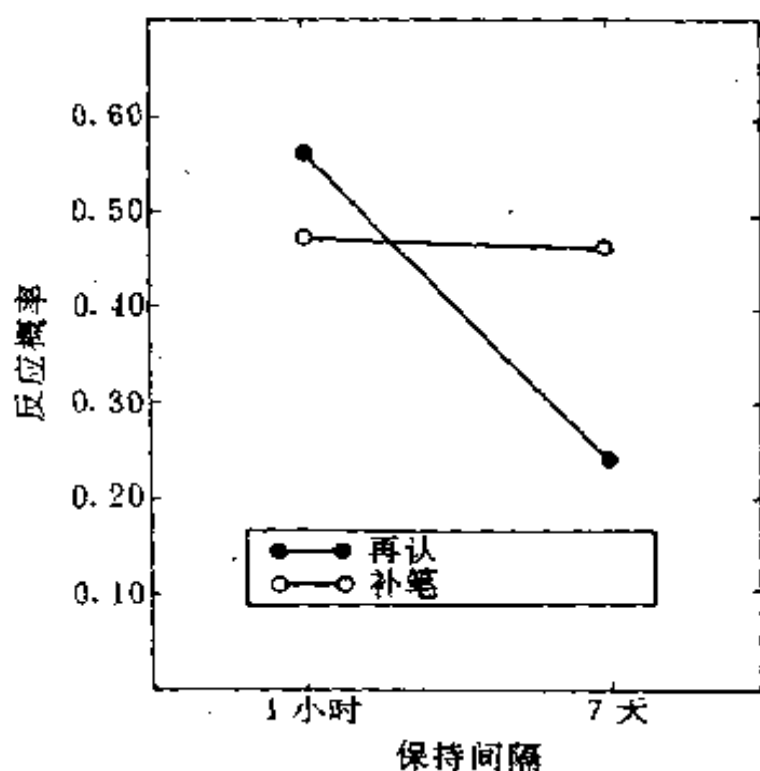


图1 7天前后再认与补笔的保持
(Tulving, Schacter, and Stark, 1982.)

样被试事先得到关于该单词的信息,从而可以预料它;在“想出”的条件下反义词首先呈现,但紧跟着出现一排问号?被试的任务是想出该单词来,而主试不呈现该单词(见表4)。通过这种程序, Jacoby 巧妙地改变着被试完成的对单词形状的感知觉加工和意义加工的数量。这就是,第一种条件下要求看清字形(感知觉的加工多),因为被试事先无法预料该单词,但同时也要求很少一点意义加工;第二种条件要求较少的感知觉加工,因为反义词首先呈现,这样被试事先就得到有关呈现的单词的信息。但相应地涉及较多的意义加工;最后,“想出”条件基本上不涉及感知觉加工,因为要求被试想出来的单词不在视觉上呈现出来,但是要求最大数量的意义加工。实验第一部分结束后,被试或者进行再认测验,或者进行知觉辨认测验,结果见表4。实验结果表明,再认成绩随实验条件从“无上下文关系”到“想出”一直上升,而知觉辨认却表现出明确的相反的结果。

果,出现了两种测验之间的分离现象。

表4 在3种实验条件下再认和知觉辨认的概率

	实 验 条 件		
	无上下关系	有上下关系	想出
测 验	××××	热的	热的
类 型	“冷的”	“冷的”	???
再 认	0.56	0.72	0.78
知觉辨认	0.82	0.76	0.67

在正常人身上也观察到自觉记忆测验与不自觉记忆测验之间的分离现象的意义重大。换句话说,没有脑损伤的正常人重复了脑损伤的健忘症病人身上的分离现象,这表明这种分离现象有着普遍的意义。

上面我们叙述了把模糊字辨认、词干补笔、补笔和知觉辨认等等不自觉记忆测验引入实验,首先在健忘症病人身上,而后在正常人身上都发现了实验性分离的现象,即在同一实验变量影响下,自觉记忆测验与不自觉记忆测验有着不同或相反的结果。

进一步要讨论的问题就是,怎样解释上述的实验性分离现象?不自觉记忆的性质是什么?自觉记忆与不自觉记忆之间的关系是怎样的?这些问题的提出与研究,标志着我们对记忆本质的认识正在深入。

四、关于启动效应的两种观点

由补笔、词干补笔等测验所显示的启动效应是人类记忆的无意识形式,这种看法是研究记忆的学者都赞同的。对健忘症病人来说,由于脑损伤而不能回忆(回忆总是有意识地进行的);对正常人来说,由于专注于眼前的作业(补笔等)而没有回忆。

但是,在回答(三)末尾提到的几个问题时,分歧就产生了。Tulving 等人提出了多重记忆系统的观点,认为记忆不是单一的系统,而是由几个系统组成的。启动效应代表着一种新的记忆系统,即知觉表征系统 PRS (perceptual representation system)^[8];而 H. L. Roediger 等人则提出传输适当认知程序的观点 (transfer-appropriate procedures approach)^[4],认为记忆测验是由各种认知过程构成的,测验之间的不一致反映了不同过程的作用。

1. 多重记忆系统的观点

(1) Tulving 的观点

Tulving 关于人的记忆是由多重系统构成的思想由来已久。1972年他把长时记忆区分为情节记忆 (episodic memory) 与语义记忆 (semantic memory),对记忆研究产生了重大影响。1982年在解释他的实验结果时(见图1),他认为启动效应反映了一种新的记忆系统的功能。他的推论是简单明了的:如果启动效应是由情节记忆系统来调整的话,我们应观察到“遗忘”,还应观察到启动效应与再认(情节记忆)的某种程度的相关,但实验结果表明,启动效应既没有遗忘也不存在启动效应与再认的相关。是不是启动效应反映着语义记忆呢?按照语义记忆最通常的观点,学习一系列单字后激活了它们的词汇节点,这种激活一直保持到补笔测验时缺笔字呈现的时候,新近激活的节点(这些节点与学过的字有关)比未受激活的节点(这些节点与未学过的字有关,但这些字与学过的字有部分相同的字母)更易接通,于是产生了启动效应。Tulving 认为,这种解释是不能成立的,由于启动效应7天前后无变化,人们难以想象激活能维持7天之久。于是他认为启动效应属于情节记忆、语义记忆之外的另一种系统。

1985年 Tulving 在“存在着多少记忆系统?”的论文中,进一步用随机独立性的概念 (stochastic independence) 来论证启动效应

代表着一种新的记忆系统^[9]。Tulving 1982 年实验(见图 1)结果的一部分列于表 5。

表 5 补笔的可能性

再 认	补 笔		总 计
	1	0	
1	0.512	0.271	0.783
0	0.136	0.081	0.217
总 计	0.648	0.352	

在表 5 中,1 代表再认正确或补笔正确;0 代表再认错误或补笔错误;0.512 表示既能再认也能补笔的单字比例;0.271 表示能再认但不能补笔的单字比例;0.136 表示能补笔不能再认的;0.081 表示既不能补笔也不能再认的; $0.512/0.648=0.790$,这表示在能补笔的单字比例中能再认的比例,它与整个能再认的比例 0.783 几乎相等。按照统计学的原理,这表明这种实验条件下再认这一事件与补笔这一事件是无关的,或者说彼此有随机独立性。把表 5 作图 2 中的 d 点(横坐标 0.783;纵坐标 0.790),a,b,c 点是实验中其他条件下的结果,这就是随机独立性的图解。对表 5 的数据作另一种处理,也得到相同的结果。概率论公式 $P(AB)=P(A) \cdot P(B)$ 表示:如果 A 事件的概率与 B 事件的概率的乘积等于 A,B 事件同时发生的概率,则 A 事件与 B 事件是无关的,或说相关为零。从表 5 可以看到,补笔的概率 0.648 乘以再认的概率 0.783 等于 0.507(预期值)。实际上既能补笔又能再认,即两事件同时发生的概率为 0.512(实验结果),0.507 与 0.512 几乎相等,因此可以认为在这种实验条件下补笔与再认的相关为零。

实验心理学家认为,“困难的情况是在概念上围绕着这样一个事例,即两种反应测量值都有高度可靠性(因此也有敏感性),但是彼此之间没有相关。这种情况在根本上强迫得出这样的结论:两种

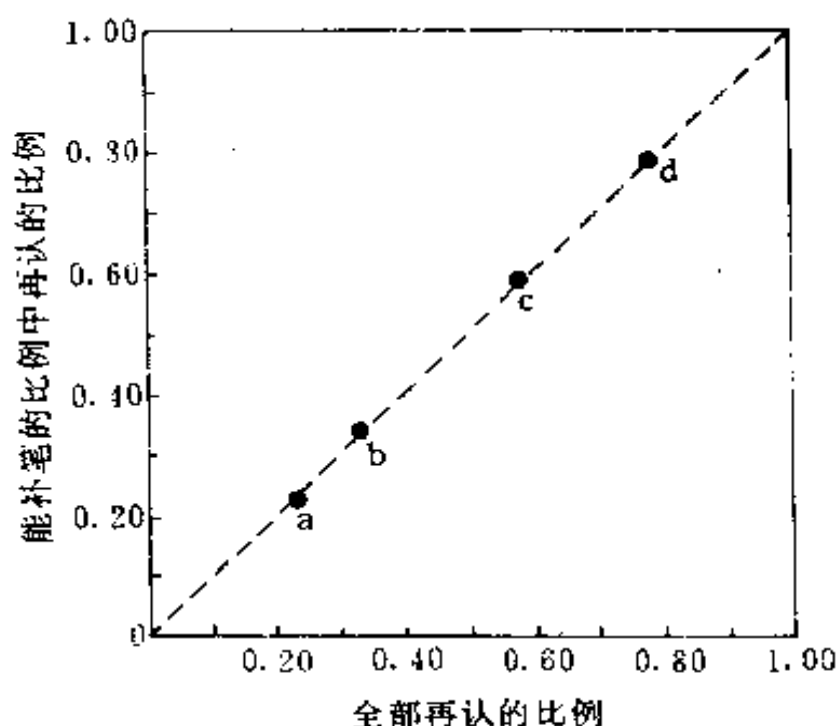


图2 随机独立性的图解

当某实验条件下能补笔的比例中,再认的比例等于全部再认的比例时,数据点就落在对角线上。Tulving 等人,1982。

反应测量标示两个相对独立的过程。如果自变量不是以同样方式影响两种反应测量,这个结论就特别肯定。在这点上我们只能作理论性的思考,因为资料不能提供给我们更多的东西了^[10]。补笔的启动效应与再认无关的实验结果进一步鼓舞着 Tulving 坚持他的多重记忆系统的思想。

1989年 D. Witherspoon 和 M. Moscovitch 用实验证明^[11],即使在两种不自觉地记忆测验中(例如补笔与知觉辨认),也可以发现随机独立性。如果启动效应代表着一种新的记忆系统,那么在测量启动效应的各种测验间,应该表现出随机相关性。实验结果与这种预期不符,给 Tulving 关于启动效应的解释打了一个问号。

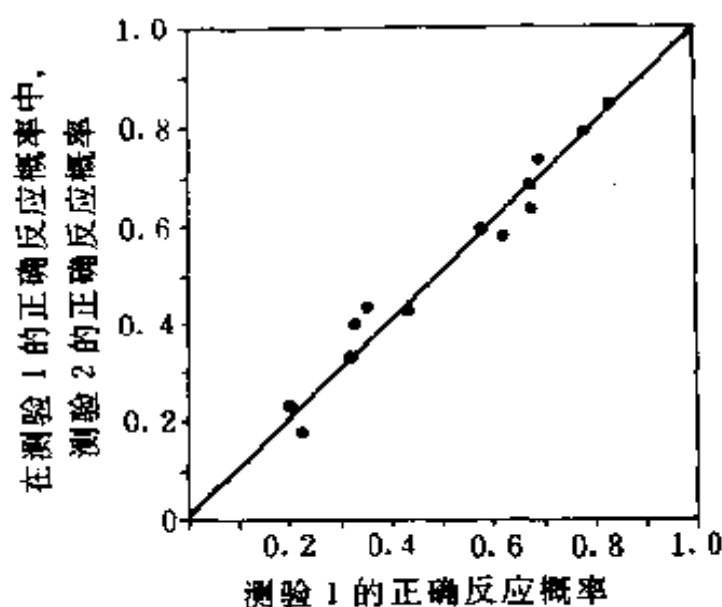


图3 连续两次测验获得的随机独立性

Tulving and Schacter 1990.

1990年 Tulving 继续进行关于启动效应代表着一种新的记忆系统的论证^[8]。他让被试学习一系列字单,例如包括 PYRAMID 和 MOSQUITO,然后针对这些单字连续进行两次补笔测验。有时候这两次补笔测验提供的缺笔字是相同的(如两次补笔测验都提供:“_Y _A _ID”作为线索),有时候这两次补笔测验提供的缺笔字有很少一点的重复(如“_O _Q _TO”和“_S _UI _O”分别在第一次测验与第二次测验)。实验结果表明,两次连续测验中,使用相同的缺笔字使两次测验结果高度相关;而在两次连续测验中,使用稍有重复字母的不同缺笔字使两次测验结果无关或相关为零。进一步的实验还使用如下方法:假设被试学习了 AARDVARK 和 UMBRELLA,然后进行两次连续的补笔测验,第一次测验的缺笔字包括3个字母,如“_A _D _R _”和“U _R _L _”;第二次测验的缺笔字包括5个字母(其中3个字母与第一次测验重复),如“_ARD _AR _”和“U _BR _LA”。实验表明,两次连续测验结果仍然是无关的或者说存在着随机独立性。图3代表着许多实验的结果,这些

实验的第一次测验是不自觉记忆测验,第二次测验或是自觉记忆测验,或是同样的不自觉记忆测验,但提供的线索不同。图上的每一个点代表一个实验或某实验中的一种实验条件。点的绘制方法同图2。由图3看到,连续两次测验都表现出随机独立性。

图3的结果虽然不能反驳 D. Witherspoon 和 M. Moscovitch 的挑战,但它仍然可以论证多重记忆系统的思想。Tulving 认为,从这些实验结果可以推测,头脑中接通有关信息从而产生启动效应的通路是不灵活的,非常特异化的(inflexible or hyperspecific)。换句话说,在一种场合下某一线索是否接通单字在头脑中的表征(a representation)与在另一场合靠另一线索接通同一单字在头脑中的表征无关。这种情形表明,产生启动效应的表征不是含抽象意义的有定位的痕迹(abstract focal traces)。如果产生启动效应的表征是这种痕迹的话,那么针对同一目标(单字)的两种不同线索应该唤起或刺激起同一表征,从而在实验结果上表现为至少是中等程度的相关。实际上,这种中等程度的相关,的确在自觉记忆测验中经常发生^[9]。由此看来,自觉记忆的表征与不自觉记忆的表征是不同的。这样,Tulving 推测启动效应代表着知觉表征系统(perceptual representation systems,简称 PRS),这种系统没有单字的固定的痕迹,也可以说是无痕迹的记忆系统(traceless memory system),但它可能包含单字的多重分布的表征,每一重表征的唤起是通过一个特殊的线索来实现的。从这一点看来,Tulving 关于启动效应的观点与 PDP 的观点有相似之处。

(2) Squire 的观点

Squire 关于记忆系统的划分代表着神经心理学家对记忆的看法及对启动效应的看法。他关于记忆系统的划分见图4^[12]。Squire 把记忆区分为两大类即陈述记忆(declarative)与程序记忆(procedural),其根据主要来自对健忘症病人的研究。在(二)中,我们叙述了有关健忘症病人的实验研究,这些研究最突出的结论是,病人

的自觉记忆受到严重损伤,但不自觉记忆(启动效应等)保持完好。换句话说,要求有意识努力的记忆很差,不需意识参与的记忆与正常人一样好。对许多各式各样的健忘症病人的研究还表明,他们仍保留有知觉、运动技巧、认知技巧、简单经典条件反射以及诸如习惯化、敏感化等等简单行为的可塑性。

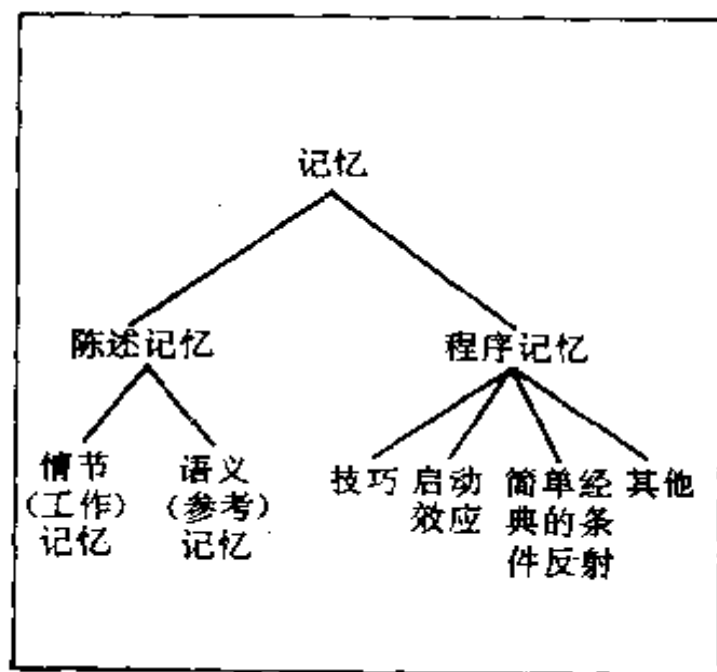


图4 Squire 关于记忆的划分

Squire 强调,区分陈述记忆与程序记忆是在探索生物系统实际上怎样工作的意义上提出来的。两大记忆系统存储的信息不同,这些信息被使用的方式不同,两种记忆需要的神经系统也不同。

陈述记忆是唯一能进入意识中的可以觉知的记忆,而程序记忆(各种技巧、启动效应等等)只能通过记忆系统的操作来实现(the memory can be expressed only in performance),因此这两类记忆反映的是不同的表征,这两种表征不能互相代替。Squire 强调,通过启动效应获得的信息不同于通过有意回忆获得的信息。例如,当1个刚才学过的单字与其余2个未学过的单字组成一小组,要求病人挑出刚才学过的单字时,病人做不到这一点,他的结果只在

机遇水平上;但是,当给予单字的头3个字母,要他填成有意义的单字时(词干补笔),学过单字填对的比例3倍于未学过的单字。这表明病人并未使用启动效应这种记忆去判断哪一个单字是学过的或未学过的。应该指出,Tulving 与 Squire 都一致认为,启动效应提取的信息不同于自觉记忆提取的信息,两种记忆有不同的表征。

要在一张图画中,迅速发现其中隐藏着的面孔是一种知觉技巧;但在发现它以后,能否再认这个面孔(判断见过还是未见过)则依赖于陈述记忆中的表征。因此陈述记忆存储的是发生在一定时间、地点的各种事件以及各种事实、映象和命题,它们可以被意识到;程序记忆存储的则是各式各样的程序或某种变化,这种变化可以促进某种认知操作的进行。

说到陈述记忆与程序记忆的神经机制,目前的研究还很少。一般说来,陈述记忆涉及脑边缘系统的神经结构,它必须依赖大脑皮质及一些特殊的脑区;而程序记忆中的启动效应,则可能是普遍存在于神经系统(包括脊髓)所有主要神经节中的一种特性。

多重记忆系统的观点能较好地解释健忘症病人的情形。由于各个记忆系统是独立的,当陈述记忆系统受损伤时,程序记忆系统仍保持完好,因而在健忘症病人身上出现了自觉记忆与不自觉记忆的分离或不一致。至于正常被试的分离现象可以理解为,一种变量影响陈述记忆但不影响程序记忆,或者发生了相反的情形,都会导致自觉记忆与不自觉记忆的分离。

2. 传输适当认知程序的观点

由 Roediger 等人提出的这种观点与多重记忆系统的观点不同,它不认为自觉记忆测验与不自觉记忆测验分别代表着不同的记忆系统,而主张记忆系统只有一个,自觉记忆与不自觉记忆测验之间的分离现象只是反映了两类测验所要求的认知程序(或过程)不同而已^[4]。这种观点包括4个基本假设:

①如果记忆测验所要求的认知过程与学习时所要求的认知过程相似或重叠,则测验的成绩好,否则就差。这一假设与“提取与加工的一致性保证好的记忆”的公认观点是类似的。②自觉记忆与不自觉记忆测验要求的提取过程不同,结果,这两类测验从学习时不同的加工过程获益也就不同。Roediger 进一步把这个假设分成两项。③学习时的意义加工、精细编码和心理映象等加工过程导致大多数自觉记忆测验(回忆和再认等)成绩良好。大量的实验表明,加工水平、句子的精细编码以及材料的有意义的组织等因素提高了回忆与再认这些自觉记忆测验的成绩。因此,可以认为自觉记忆测验是要求概念驱动过程的(conceptually driven processing)。④大多数不自觉记忆测验严重地依赖于学习时与测验时的知觉过程(perceptual processing)的匹配程度。许多不自觉记忆测验(例如知觉辨认、补笔与词干补笔等等)似乎是提取过去经验中的知觉成分。因此,可以认为不自觉记忆测验是要求材料驱动过程的(data-driven processing)。这样,概念加工中的各种变量对不自觉记忆测验就不会有什么影响,而表面特征方面的变量(variations in surface features,如感觉道)对不自觉记忆测验则有重大影响;另一方面,表面特征方面的变量对自觉记忆测验并没有什么影响。

传输适当认知程序的观点(简称程序观点)怎样解释两类测验间的分离现象呢?我们以(三)中 Jacoby 的实验为例(见表4)。Jacoby 的实验表明,学习时把单字写出来(“想出”的实验条件)比脱离上下文读它,在自觉记忆测验(再认)中有更好的成绩($0.78 > 0.56$);但在不自觉记忆测验(知觉辨认)中,读出单字比写出单字有更多的启动效应($0.82 > 0.67$)。Roediger 认为,这是因为阅读单字比写出单字更多地涉及材料驱动过程,因此对材料驱动测验(知觉辨认)能更好地传输信息,导致读出单字比写出单字有更多的启动效应;从有关的线索(热--?)写出单字比脱离上下文阅读($\times \times \times \times$ --冷)更多地涉及意义加工,因此对概念驱动测验(再

认)能更好地传输信息,导致写出单字比读出单字有更好的再认成绩。

许多其他两类测验间的分离现象也可以用程序观点来解释。

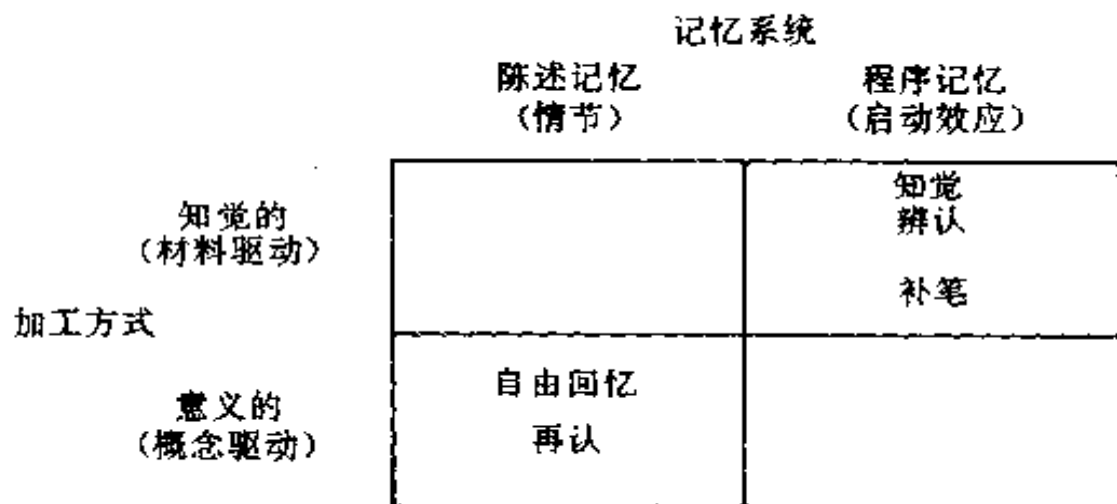


图5 区分加工方式与记忆系统作用的实验设计

但是,对于像 Jacoby 这样的实验结果,人们用多重记忆系统的观点也完全能够解释。在 Jacoby 的实验中,再认属于陈述记忆(情节记忆),而知觉辨认(启动效应)属于程序记忆,因此实验结果的不一致反映了不同记忆系统的功能。既然两种观点都能解释 Jacoby 的实验,那么反过来说,仅仅进行类似于 Jacoby 的实验是不能解决两种观点的争论的。我们必须寻找新的实验设计,其实验结果可以在肯定某一观点的同时否定另一观点。Roediger 提出了下面的实验设计(见图5)^[13]。他认为,如果我们同时比较4项记忆测验结果,我们就有可能判断,在同样实验条件下,实验结果是符合多重记忆系统观点的预测还是符合程序观点的预测。他还指出,不幸的是,以往绝大多数关于实验性分离现象的实验研究往往只用右上角的测验与左下角的测验作比较(例如,Jacoby 的实验就是再认与知觉辨认的比较),因而造成了自变量记忆系统与加工方式的混淆,无法区分两者各自的作用。例如,从纵向看即从多重记忆系统的观点来看,比较右上角测验与左下角测验的分离,可以证明多重

记忆系统的观点;但从横向看即从加工方式(程序观点)来看,同样的比较都可以证明程序观点。

1989年 Blaxton 进行了一项实验,同时比较4种测验结果,得出了有利于程序观点的结论(见表6)^[3]。

表6 Blaxton 的实验结果(1989,实验一)

任 务 组 别	材料驱动测验		概念驱动测验	
	字母线索回忆 (情节记忆)	模糊字补笔 (启动效应)	自由回忆 (情节记忆)	一般知识 (启动效应)
阅读组	0.45	0.75	0.19	0.34
产生组	0.34	0.46	0.31	0.50

实验中 Blaxton 首先要求被试在两种实验条件中的一种学习单字。例如,在无上下文关系中大声读出单字(×××××treason),或者在概念线索及头一个字母的帮助下,写出一个单字来(简称产生作用,例如,espionage—,t——,应该写出 treason)。然后,Blaxton 应用了4种测验。在字母线索回忆(graphemic cue-recall)中,主试提供 treasure(财富)给被试,要他以此为线索去回忆单字即 treason(叛逆),这样,线索在形状与声音上类似于目标词,但在意义上无联系。测验时告诉被试,要他们回忆在字形与发音上和线索词类似的词,所以测验属于情节记忆,但却依赖目标词与线索词在视觉与听觉上的类似性。在一般知识测验(general knowledge)中,要求被试回答一般性的知识问题,例如,“因为什么罪行,Rosenbergs 被处死?”学了 treason 这个词,被试就容易回答这个问题。一般知识测验虽然属于启动效应性质的测验,但它要求的认知过程却是概念加工。这样,Blaxton 周密的实验设计使人们能够检验实验性分离是发生在记忆系统之间还是发生在加工方式之间。

表6说明,虽然情节记忆与启动效应分属不同的记忆系统,但

如果进行的测验是材料驱动测验的话,阅读的作用总是大于产生作用(在情节记忆中 $0.45 > 0.34$;在启动效应中 $0.75 > 0.46$),原因是字母线索回忆与补笔都要求表面信息的加工;另一方面,虽然情节记忆与启动效应分属不同的记忆系统,但如果进行的测验是概念驱动测验的话,产生的作用总是大于阅读的作用(在情节记忆中 $0.31 > 0.19$;在启动效应中 $0.50 > 0.34$),原因是自由回忆与一般知识问题都要求概念加工。简言之,实验性分离现象有规律地依赖于加工方式。但是,表6同时也表明,实验性分离现象并不是有规律地依赖于记忆系统的。在情节记忆中,有时阅读作用大于产生作用($0.45 > 0.34$),有时阅读作用又小于产生作用($0.19 < 0.31$);在启动效应中,有时阅读作用大于产生作用($0.75 > 0.46$),有时阅读作用又小于产生作用($0.34 < 0.50$)。如果记忆系统的功能决定着分离现象的产生,那么在同一记忆系统内,应该出现一致的结果,分离只应该发生在系统之间,但实验结果与这个观点不符。总之,Blaxton 的实验是支持程序观点而否定多重记忆系统观点的。其他许多研究者也获得了类似的结果^[13]。

一般说来,程序观点在说明正常被试的两类测验间的实验性分离现象上是比较成功的。材料驱动测验与概念驱动测验各自要求的认知过程不同,若测验与学习的认知过程匹配,那么测验成绩就好;而测验时与学习时的认知过程失匹配,则导致实验性分离现象。但是,程序观点在解释健忘症病人的实验性分离时却遇到了麻烦。健忘症病人自觉记忆很差,而不自觉记忆与正常人一样好,程序观点解释为,他们在情节记忆中的概念过程有障碍,但在启动效应中的知觉过程保持完好。这种说明适用于部分健忘症病人的结果。但是,病人在那些要求概念驱动的不自觉记忆测验中也表现了与正常人一样的启动效应^[13],这就不能说他们的概念过程有障碍了。另外,不同类型的健忘症病人的情形更易于用多重记忆系统的观点来说明。

总之,程序观点能较好地解释正常人的启动效应,但多重记忆系统的观点能较好地解释健忘症病人的启动效应。

五、两种观点在方法学上的对立

面对实验性分离现象,程序观点与多重记忆系统观点各执一方。它们在观点上的对立,部分地根源于在方法学上的对立。持程序观点的多数认知心理学家认为,逻辑上我们不能从实验性分离现象推论出机能上独立的心理过程或系统;而主张多重记忆系统的神经心理学家及部分认知心理学家则相信,实验性分离证明着不同的心理过程或记忆系统。

1. 认知心理学家的看法

实验性分离包括几种类型的实验设计:单一的分離(a single dissociation);非交叉的双重分离(uncrossed double dissociation);交叉的双重分离(crossed double dissociation)和双向关联(reversed association)^[14]。

单一分离指的是,实验中一个自变量 V_1 影响测验任务 A,但不影响测验任务 B。在(二)中我们引述的 Warrington and Weiskrantz (1970) 的实验(表2)就属于单一的分离。从表2可以看到,被试变量影响回忆但不影响模糊字辨认;或者,被试变量影响再认但不影响词干补笔。人们认为,不能从单一分离推论出2个独立的心理过程,因为单一的心理过程完全可以解释单一分离的实验结果:某一心理过程只对测验任务 A 敏感,而对 B 完全不敏感。

非交叉的双重分离指的是,实验中变量 V_1 影响 A 任务但不影响 B 任务;而变量 V_2 影响 B 任务但不影响 A 任务。我们举 Craik 的实验来说明^[15]。他的实验是这样的:以视觉或听觉的形式呈现一系列低频英语单字,并且以2种不同水平的定向任务指导被试对

单字作不同层次的加工,即语音、字母形状知觉的(浅层)加工或意义方面的(深层)加工。学习单字完毕,先作再认,然后进行补笔。实验要点见表7,实验结果见图6与图7。图6表示,不同定向任务对再认有不同的作用(深层加工优于浅层加工,达到显著水平),但不同呈现方式(视觉或听觉呈现)对再认的作用是类似的(两种呈现方式的结果无显著差异)。图7表示,不同呈现方式对启动效应有不同

表7 克瑞克实验要点举例

定向任务的问题			目标词	补 笔
听觉(浅层)	视觉(浅层)	意义(深层)		
2 syllables?	contains E?	found in China?	pagoda	-ag-da

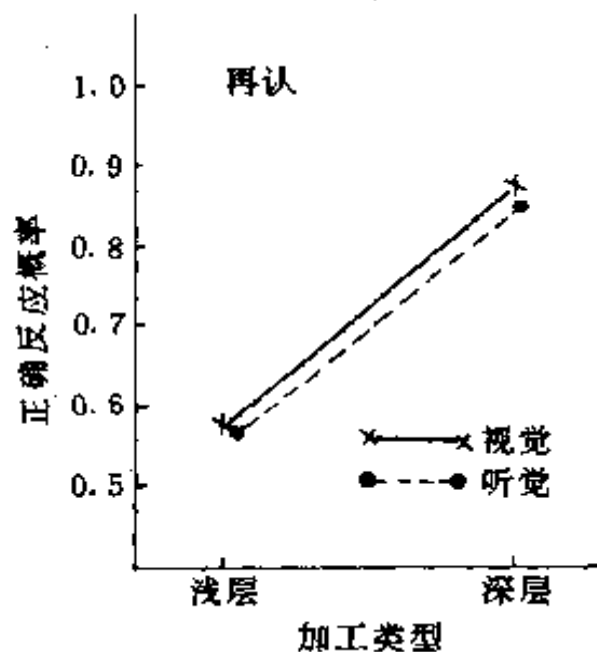


图6 不同感觉道不同加工类型对再认的影响

作用(视觉呈现的结果显著高于听觉呈现的),但不同定向任务对启动效应的作用是类似的(浅层与深层的加工结果大致相等)。这样,实验结果显示了两种测验任务之间(再认与补笔)完全的双重分离。

虽然非交叉的双重分离实验比单一分离的实验提供了更多的信息,但它也无法排除一种心理过程的可能性。从图6与图7中我们看到,浅加工与深加工效果都随着听(补笔)——视(补笔)——听(再认)——视(再认)逐步上升,这种情形表明,当以加工效率来衡量时,补笔与再认这两种测验任务在4种实验条件下都有类似的趋势,因此,用一种心理过程来解释就够了。Craik 自己也不认为他的实验结果支持多重记忆系统的观点。他认为,虽然他的实验结果表明,情节记忆系统(再认)对不同的感觉道呈现方式不敏感;但另外的许多实验证明,不同的感觉道的确对再认有重要影响。这样,不同的呈现通道不影响再认不是普遍的、绝对的。另一方面,虽然补笔显示的启动效应不受加工层次的影响(见图7),但也不能说深层的意义加工只影响再认而不影响启动效应。因为图7同时包含着这样的意思:听觉呈现——(视觉)补笔的结果显著优于未学过的

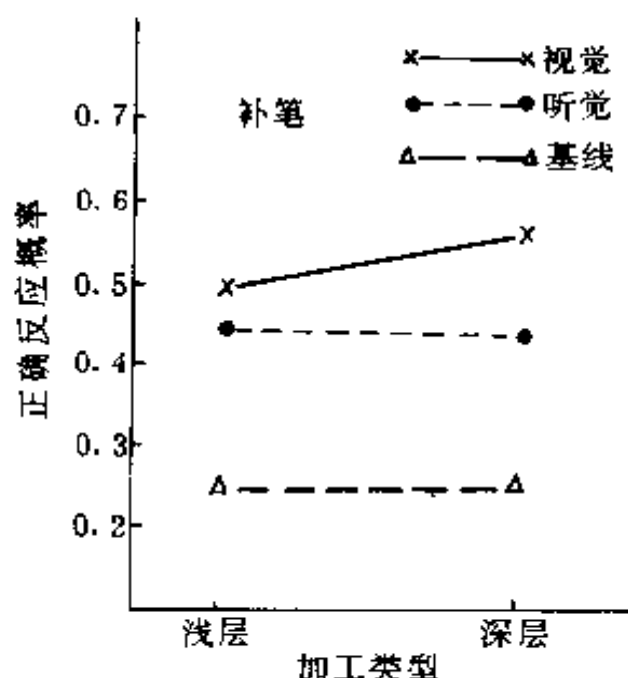


图7 不同感觉道不同加工类型对补笔的影响

单字的补笔(基线水平)。但是,补笔所需要的信息不能来自听觉呈

现(两种感觉道不一致),所以,补笔中的启动效应必须来自学习时对字词的抽象意义的加工,也可以说启动效应明显地受抽象的语义信息所影响。

交叉的双重分离指的是,在实验中,1个变量对2个任务有相反的作用。例如, a, b 是某变量的2个水平,则 $T_A(a) > T_A(b)$, 而 $T_B(a) < T_B(b)$ 。换句话说,当1个变量从水平 a 变到水平 b 时,实验结果在 A 任务变低而在 B 任务变高。上面所举的 Jacoby 的实验就属于交叉的双重分离。可以把这种情形看作是一种心理过程在2个任务中表现为负相关的结果,因此从交叉的双重分离不能推论出2种独立的心理过程。

现在让我们对上述的实验性分离的实验设计小结一下。

我们知道,一种心理过程的表现形式总是单向性的(monotonicity),即在变量影响下,2种测验任务的结果要么成正相关关系(两者都上升或都下降);要么成负相关关系(一个上升,一个下降)。单一分离、非交叉的双重分离和交叉的双重分离在逻辑上都不能推导出2个独立的心理过程或系统,原因就在于它们不能摆脱单一心理过程的单向性表现。如果我们能达到一种设计,其结果超出了单向性的表现,即在同一实验中,2个任务的结果既是正相关的,又是负相关的,那么,由实验推测2个独立的心理过程就是可能的了,这种设计叫作双向关联。下面是一项双向关联实验的结果,见表8。

表8是 Duchek and Neely 的实验结果^[16]。在再认与词汇判断两种记忆测验中,字词频率与加工水平的交互作用是不同的:再认中加工水平的作用使反应时在高频短于低频32毫秒,但在词汇判断中却使反应时在高频长过低频43毫秒。这样,在同一实验中出现了正相关(+114与-105, +71与-73)与负相关(-43与+32),这是单一心理过程解释不了的,因此,它可能是对两个独立的心理过程观点的有力支持。

表8 再认与词汇判断(lexical decision)中字词
频率×加工水平的分离(反应时:毫秒)

测 验 任 务	加 工 类 型		加 工 水 平 (语义·语音)
	语 义	语 音	
再 认			
高 频	1101	1219	+118
低 频	996	1146	+150
高低频之差	-105	-73	+32
词汇判断			
高 频	600	648	+48
低 频	714	719	+5
高低频之差	+114	+71	-43

但是,即使对于双向关联的实验设计,反对者也提出尖锐的批评。D. L. Hintzman 认为,“表面上双向关联似乎在理论上容许(人们)去发现不同测验任务背后的不同记忆系统,但实际上这是一个错觉”^[12]。他说,任何两种测验任务总是会有一定的差别的(不然就不成为两种测验任务),如果不同的任务涉及不同的心理过程,而不同的心理过程产生分离,那么我们可以说,任何两种测验任务的比较都会产生分离!事实上,在程序记忆系统内已发现了分离现象[见(四),D. Witherspoon and M. Moscovitch 的实验],而在陈述记忆系统内已知的分离就更多了。例如,在回忆中高频字比低频字成绩更好,而在再认中结果相反。程序观点认为,既然在同一系统内也出现了分离,那么用分离来推论不同的记忆系统在逻辑上就软弱无力了。Roediger 指出,为了弥补这个严重的缺陷,多重记忆系统观点的提出者不断地提出更多的记忆系统。例如,前几年把启动效应看作是一个独立的记忆系统,现在又在其中划分出字词系统(a word-form system)、结构描述系统(a structural-description system)和概念语义系统(a conceptual-semantic system)^[13]。

2. 神经心理学家的看法

神经心理学家从自己的科学实践出发,认为双重分离的实验设计是一种发展神经心理学的非常有效的方法,从双重分离获得的资料,可以推测存在着好几种潜在地独立而又交互作用的记忆系统。

首先应该明确,神经心理学家在研究人脑机能组织时,只是对脑局部损伤病人的行为变化作临床观察。因此,在神经心理学范围内原始资料的相互关系总是如此编排:脑的某一损伤部分导致某一测验任务的差的结果,相应地双重分离原则就是:损伤定位于 A 的一组病人比定位于 B 的一组病人在 P 测验任务中占优势而在 Q 测验任务中占劣势。如果出现这样的结果,我们就可以推论在这两组间有一种差异,即各组病人损伤部位的定位是明确地各不相同的。

著名的前苏联神经心理学家鲁利亚(Luria)在他的《神经心理学》一书第二章第四节“症候群的分析 and 心理过程的系统性组织”中,这样述评“双重分离”^[18]:“为了从脑的局部病理学资料中能够得出心理过程的有关结构,以及它们在大脑皮质中的定位的可靠结论,必须经过由单个症候的审定到整个症候群的描述,或者像通常所说的,转向对那些在脑的局部损伤的情况下所出现的行为改变的症候群的分析……事实证明:任何一个特定的皮质病灶确实破坏一些心理过程的进行,同时另外一些过程仍然保持完好(这个现象被美国神经心理学家 H. L. Teuber 称为机能的双重分离原则),包括在神经心理学研究的资料中的这种事实是很丰富的。例如,在左半球的顶-枕区(或下顶区)中的特定病灶,能导致知觉和运动的空间定向的破坏……但(这样的)病灶并不妨碍流畅的语言、对音乐旋律的再认或再现……但是,在皮质颞叶区(听觉区)局部损伤时,则观察到完全相反的情况……(这种损伤)导致听知

觉组织的破坏……流畅的、有选择性的言语受到严重破坏,但是,空间中的定向、运动的空间组织……照样保存着。

凡此种种十分清楚地证明,对症候群和那种在脑的局部损伤时所产生的“双重分离”进行仔细的神经心理学分析,使得有可能接近于对心理过程本身进行结构性的分析,并区分出那些属于这一组心理过程而不属于另一组心理过程的因素。这也使得有可能密切接近于解决关于心理过程的内部成分问题。这个问题在通常的心理学研究条件下始终是没有得到解决的。这样,大概可以把那些类似的心理过程加以区别开来,并且也可以将那些不同形式的心理活动接近起来”。

Luria 用具体的例子说明了相似心理过程的不同脑定位以及不同心理过程的相同的脑定位。例如,音乐和言语听觉似乎是同一个心理过程的两种变式,十分相像。但是,对一些具有脑的局部病灶损伤的病人的观察表明:左颞区某些部位的破坏会导致言语听觉的明显失调(完全不能区别近似的语音),但音乐听觉仍然保持着而未受损害。有一位优秀的作曲家在左颞叶脑溢血之后,再也不能区别语音和理解对他讲的话,但是他继续创作出了出色的音乐作品。另一方面,研究表明,像空间定向、计算和对复杂的逻辑语法结构的理解这些各不相同的心理过程,实际上都依赖于共同的定位:左半球顶-枕区(下顶部)。

另一位著名的英国神经心理学家 Weiskrantz 对于“双重分离”的述评提出3点主张^[19]:一,世界上没有“纯粹”的测验任务,即使最简单的任务也要求多种心理过程的参与;二,脑损伤(不管是结构的还是化学的)总是影响到多重神经通路,因此,神经心理学的测验结果也总是十分复杂的,即涉及不止一种缺陷;三,在上述复杂情形下,脑损伤造成的某种单一的缺陷,原则上仍然是可能测量出来的。从上述3点主张得出的一般结论是,神经心理学的分析必然要求多种的测验任务与多种不同定位的脑损伤来进行比较。

而双重分离正是一个可以满足这种要求的强有力工具。病变在许多情况下影响邻近部位的功能,重复运用两种或两种以上“双重分离”的方法,就能逐步使引起病变的定位越来越确切。换句话说,确定了一种分离,就提示了下一步应用的测验任务,最终确定病变的定位(在这里应该注意,当我们重复地应用双重分离来精确定位损伤部位时,我们同时也在逐步地分离出由准确定位的损伤部位造成的单一的缺陷来)。

根据双重分离原则获得的神经心理学证据,Weiskrantz 认为人类具有的记忆系统比目前我们已知的要多。他把记忆分为5个系统:短时记忆、知识与技巧、牢固的联想记忆、启动效应和事件记忆(event memory)。他还指出,这5个记忆系统的每一个还可以细分为若干个子系统(subsystems)或子类(subcategories)。

当 Tulving 等人应用分离原则把启动效应系统再细分为字词系统、结构描述系统和概念语义系统时,Roediger 不无嘲笑地评论说,按照已有的资料,我们将会有20—25种记忆系统!换句话说,多数认知心理学家否定分离原则(包括双重分离)在实验心理学、认知心理学范围内的应用。但是,对于神经心理学家来说,双重分离是一种十分有效的手段,自1955年 Teuber 提出,至今已广泛地为他们接受与应用。

认知心理学家与神经心理学家的矛盾在哪里呢?

我觉得,这也许是认知心理学家从心理水平来看待分离现象,而神经心理学家是从生理水平(或脑的结构机能的角度)来看待分离现象的缘故。因此,前者不满足于把分离现象归于不同的过程或系统就止步,而是不断地提问,每种分离现象的特点是什么,他们要求记忆理论能预测并解释各种变量对每种记忆测验的效果;而神经心理学家的目标不同,他们要寻找的正是心理过程的脑区部位(或生理机制),他们需要分离的手段不断地细分各个部位,愈来愈精确地定位它。

记忆心理学家 Craik 曾经说过：“在生理学或解剖学水平上，机制和结构的概念是必须的，但在心理学的水平上就不是必须的了”^[7]。我同意这种观点。由于在心理水平上不需要结构的概念，因此，心理学家拒绝双重分离的原则就是可以理解的了。

六、结 尾

系统地研究启动效应不过10来年的时间，但收获是巨大的。一百多年前，艾宾浩斯提出的无意识的记忆研究只是到了启动效应的研究才算真正开始。但也正因为其研究历史不长，人们预期今后也许会有更重要的发现^[13]。

现在，不论是心理学家还是神经心理学家都已认识到，我们测验记忆的各种任务在某种性质上是一个连续体而不是间断的。例如，Roediger 虽然提出材料驱动的记忆测验与概念驱动的记忆测验，但他同时表明，还有许多记忆测验是介于两者之间的^[4]；又例如，1983年 Craik 就提出^[7]，各式各样的记忆提取任务即记忆测验，按它获得的环境支持(即得到的线索)的多少以及意识参与的多少可以构成一个连续体(参见表9)，按表中的说明，自由回忆要

表9 不同的提取任务

测 验	环 境 支 持	自我发动的活动 (意识觉知?)
自由回忆		
线索回忆		
再 认		
再 学		
知觉学习		
程序学习		

求最多的意识参与而没有一点环境支持(因为在自由回忆时不提供任何线索)。相反,程序学习(包括现在我们知道的补笔、词干补笔等等)则获得最大的环境支持而无需意识参与。今天,既然我们已有的记忆测验任务是一个有最多意识参与到不需要意识参与的连续体,而我们的理论观点又是建立在记忆测验所获得的资料之上的,那么,建立一个包括意识记忆与无意识记忆的统一记忆理论就不仅是必要的也是可能的了。

其次,如果我们把有意识记忆与无意识记忆看作一对矛盾的话,我们要问,什么条件、什么机制使它们相互转化?这是一个很值得研究的问题。例如,要是我们提供测验材料 sha _ 给被试,当指示语是要被试回想刚才学过的单字(如 shade)时,这个测验就是有线索回忆,属于有意识记忆;当指示语要求被试迅速把他(她)心中首先想到的单字填出来时,这个测验就是词干补笔,属于无意识记忆。在这个例子中,看来指示语是一个促使记忆朝有意识方向或无意识方向转化的条件。但是,我们的被试是正常人,当指示语要求他们填空时,他们也许也进行回想,这是实验者无法控制的,换句话说,指示语作为一个转化条件并不可靠。

这些,就是我写完这篇文章以后得到的启示。

主要参考文献

- [1] H. 艾宾浩斯著,曹日昌译《记忆》,科学出版社,1965年,2页。
- [2] 朱滢等,“加工水平、回想策略与不自觉记忆”,《心理学报》1991,第三期,264--271。
- [3] E. Tulving (1985), Elements of episodic memory, clarendon Press. Oxford, pp. 73.
- [4] Roediger, H. L. Weldon M. S. & Challis B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A Processing account. In H. L. Roediger & F. I. M. Craik (Eds), Varieties of memory and Consciousness: Essays in honour of Endel Tulving (pp.

3—41). Hillsdale, NJ; Erlbaum.

- [5] Graf, P., Squire, L. R., & Mandler, G. (1984). The information that amnesic patients do not forget. *J. Exp. Psychol. Learn. Mem. Cognit.* **10**, 164—178.
- [6] Tulving, E., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word—fragment completion are independent of recognition memory. *J. Exp. Psychol. Learn. Mem. Cognit.* **8**, 336—342.
- [7] Craik, F. I. M., (1983). On the transfer of information from temporary to permanent memory. *Phil. Trans. R. Soc. Lond. B* **302**, 341—359.
- [8] Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science* **247**, 301—305.
- [9] Tulving, E. (1985). How many memory systems are there? *American Psychologist*, **40**, 385—398.
- [10] B. J. 安德伍德和 J. J. 肖内西著, 方俐洛等译, 《心理学实验方法》, 科学出版社, 1981年。
- [11] Witherspoon, D., & Moscovitch, M. (1989). Stochastic independence between two implicit memory tests. *J. Exp. Psychol. Learn. Mem. Cognit.* **15**, 22—30.
- [12] Squire, L. R. (1987). Memory and brain. New York: Oxford University Press.
- [13] Roediger, H. L. (1990). Implicit memory— Retention without Remembering, *American Psychologist*, 1043—1056.
- [14] Dunn, J. C., & Kirsner, K. (1988). Discovering Functionally independent Mental Processes; The Principle of Reversed Association. *Psychological Review*, Vol **95**, No. 1, 91—101.
- [15] Craik, F. I. M. (1987). an experiment on double dissociation. unpublished manuscript.
- [16] Neely, J. H. (1989). Experimental dissociations and the episodic / semantic memory distinction. In Roediger, H. L., & Craik, F. I. M. (Eds.), Varieties of memory and Consciousness; Essays in honour of Endel Tulving (pp. 245—246). Hillsdale, NJ; Erlbaum.

- [17] Hintzman, D. L. (1990), Human Learning and memory: Connections and dissociations, *Annu. Rev. Psychol.* 41, 121.
- [18] A. P. 鲁利亚著 汪青等译《神经心理学原理》, 科学出版社 1983年, 77—80页。
- [19] Weiskrantz, L. (1989). Remembering dissociations. In Roediger, H. L., & Craik, F. I. M. (Eds.), *Varieties of memory and Consciousness: Essays in honour of Endel Tulving* (pp. 101--118). Hillsdale, NJ; Erlbaum.

当代思维研究

杨 治 良

(华东师范大学)

思维,长期以来一直像磁石一样吸引着哲学家、心理学家、逻辑学家及其他许多科学家,也吸引着社会学家、教育学家的注意;它既引人入胜,又令人迷惑,实在是人们很愿意研究又很值得研究的一个课题。

思维一向被称作高级认知过程。说它“高级”,主要是因为传统上一直认为思维、推理和问题解决这些活动总是出现在信息加工过程的最后一环,这个观点到本世纪60年代更加牢固了;又因为知觉和记忆往往与神经过程有密切联系,而思维则似乎没有什么特定类型的神经过程与之对应,于是大家就确认思维高于其他认知过程了。看来,思维活动成为高级认知过程,得力于两个因素:第一,它包含着较低级的感觉、知觉、记忆、学习等过程;第二,它与低级认知过程又有质的不同,能从已知信息中推断出未知的信息。

对于思维这一高级认知过程的研究,也走过了长期艰难的历程,从早期的内省主义、行为主义、完形主义,发展到了今天的信息加工理论。

在心理学成为一门独立的科学之前,哲学家们早就注意上了人类的思维。17世纪兴起的联想主义心理学,继承了古希腊心理学的联想概念和联想规律,对思维问题提出了一些颇有见地的观点,对思维的概括性、间接性已有认识。

1879年,冯特建立了莱比锡实验室。他崇尚内省方法,而思维过程太复杂,飘忽不定,无法内省,他也就没有好好地研究。倒是他

的学生屈尔佩等人在符兹堡大学对思维进行了大量的实验研究，形成了历史上著名的符兹堡学派，提出思维不能归结为感觉或意象的论断，并发现了“决定倾向”现象。但由于内省法的缺陷和思想方法的简单，他们的理论今天已经不受重视了。

行为主义心理学派的出现，曾大大推进了思维研究。华生提出，思维是无声的言语，是一种行为；对思维的研究不能用内省，而要提倡实验。这一时期对思维的研究还以另一种形式曲折地表现出来，即研究动物和人的学习行为。桑代克就研究了老鼠关在笼子里的情况下所表现出的尝试错误行为。另一位叫如吉尔(Ruger, 1910)的心理学家则发现人类解决问题时也有类似的尝试错误模式。他们得出结论，认为人类思维类似于动物的尝试错误，只是比较隐蔽罢了。但他们对“顿悟”、“理解”这些概念仍不以为然。后来，为了应付格式塔学派的挑战，行为主义曾提出一些新的概念，如中介刺激、中介反应等，以弥补对有机体内部反应过程研究的不足，一时似有振作气象。但在认知的信息加工学派的研究大潮面前，新行为主义也黯然失色了。

格式塔学派对思维心理学作出了巨大贡献，尤其是在“顿悟”和“理解”方面。除了魏台默、苛勒的工作外，邓克尔对问题解决和功能固着的研究、陆钦斯对定势的研究都是我们所熟知的。

随着对思维研究的深入，当心理学家把注意力转移到对内在认知过程的剖析上时，信息加工理论应运而生了。

由于计算机的出现，心理学家开始把人比作一个计算机似的系统，他可以对信息加以编码、储存、提取和转换。计算机编程时常用的流程图也被用来描述人的认知活动。如果把流程图越来越具体化，就可以在计算机上越来越近似地模拟人的认知活动了。

对于人的计算机模拟还存在许多困难，也引起学者们的一些争议，但是它有一个重要作用却不容忽视，那就是它促进了对人的信息过程细节的探究。对人的信息加工容量的研究就是一例。同

时,计算机工作时是串行加工的,即同一时间只进行一种操作,这就使学者们想到人类信息加工(尤其是思维)是串行的还是并行的。许多研究人员认为是串行的,因为人在解决问题时很少同时考虑几条思路。但是,近来计算机科学界和心理学界对于并行处理的兴趣均日益浓厚。心理学界出了个叫联系(结)主义的小学派,认为信息加工是由大量密集排列的、相互影响(相互兴奋和抑制)的简单单元所构成的网络来承担的。尽管这种观点还有待大量确凿事实的支持,但看来其前途是很宽广的。

人工智能的研究对信息加工理论也有重大意义,它已经成为有关加工过程的学说、观点和术语的重要来源。例如,纽威尔和西蒙的通用问题解决者程序(1972)就是一个关于人类解决问题的方法的有价值的源泉;“产生式系统”在构筑认知的计算机模型中也起了很大的作用。

近来,一门新兴的学科——认知科学将思维研究引向新的天地。它关心的是实际生活中复杂问题的解决。其研究方法往往是详细研究少数被试在复杂任务上的表现,并常常要导出相应的计算机程序模型。这就要求研究者既有心理学知识,又精通计算机技术。

（总之,信息加工理论极大地促进了对认知过程的细微剖析。本文将从问题解决、概念形成、推理、决策、想象与创造等诸方面介绍这些成果。

一、问题解决

本世纪70—80年代以来,“思维”一词的定义渐渐有了变化。原来西方心理学界将思维看作是一种内在心理过程,它具有操纵人的内部符号,完成解决问题等功能。随着对人的心理模式研究的广泛和深入,学者们开始把思维看作是人们用来建立、运用和修正心

理模式的过程的集合。例如,在问题解决过程中,人就要探寻一种与问题相应的心理模式,以使用它确定一个最佳的(至少是满意的)行动方案。在这个新定义的指引下,心理学家们首先将兴趣转向了心理模型的建立上。他们向被试提出各种实验课题,记录被试的行为,藉此推测被试对课题的理解(模型)。他们设计或采用的实验课题往往带有人工性质,颇类似于人工概念的设计思想。典型的例子有移字码、富翁-强盗过河问题、河内塔、量水问题、密码算题(即 $DONALD + GERALD = ROBERT$ 之类的问题)。这些问题有几个特点:①新颖,相对于被试而言是第一次遇上的,过去经验的作用、机会是人人均等;②规模小,解答时无需渊博的知识,也无需成百上千个步骤,只消几个,最多20—30个步骤即可完成,且步骤多的问题往往有诀窍(例如河内塔问题);③明确具体,问题各要素都确定无疑,不会造成其他理解,属于那种 well-defined 问题。利用这些问题,研究人员探讨了两种类型的解题模式:状态动作模式和问题分解模式。

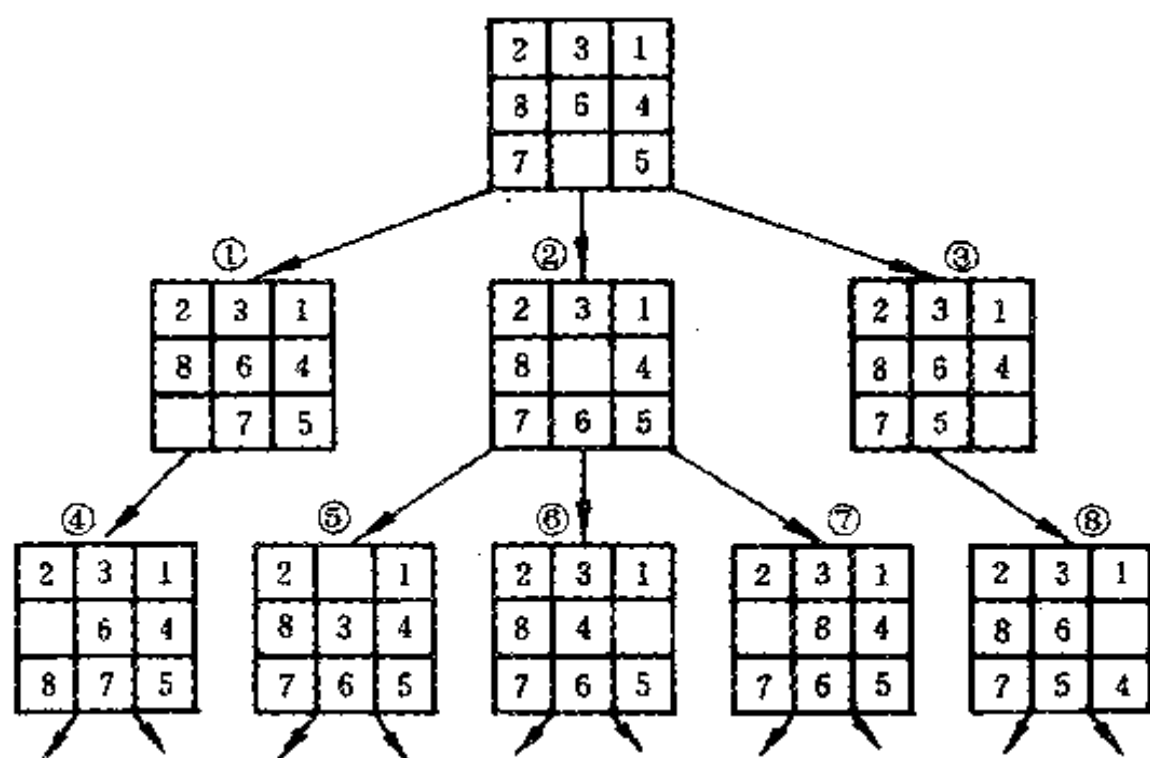


图1 移字码问题树形图

所谓状态动作模式,指的是在问题解决过程中,搜寻可以将初始状态转化为目标状态的一系列操作(动作)这样一种方式。例如对于移字码问题,被试就可能从下面树形图中搜寻到一系列合适的操作。当然,这样的搜寻往往是带有选择性的,被试往往优先考虑那些能使状态较快转化为目标状态的操作,这就是我们所熟悉的“登山法”。

问题分解模式则是将一个较复杂的问题分解为几个较简单的子问题。我们所熟知的“手段-目的分析”就是问题分解的一种方式,其要点是:①比较初始状态和目标状态,提出第一个子问题:如何缩小两者差距?②找出缩小差距的办法和操作;③如果提出的办法实施条件不够成熟,则提出第二个子问题:如何创造条件?④找出创造条件的办法和操作;⑤如果④中提出的办法实施条件也不成熟,则提出第三个子问题:如何创造条件?……如此循环进展,直至问题解决。问题分解模式比状态动作模式圆熟一些,当然它也不是万能的。研究人员分析了这两种模式的适用范围和工作特点。

托马斯于1974年率先采用富翁-强盗过河研究状态动作模式。实验在计算机上做,记录了被试完成各个操作的反应时和准确率。图2和图3分别说明问题的标准解法和测得的反应数据。在图2中,H代表富翁,Q代表强盗,B代表小船。每个状态图中都标出此岸H数和Q数,以及小船在此岸(1)还是在彼岸(0)。例如“320”说明此岸有3个H,2个Q,小船在彼岸。状态图之间连线上则标出操作内容。如1H1Q说明有1个H和1个Q坐上小船渡向对岸。

从图3的两项数据可以看出,321和110状态下被试颇感困难。在321状态下,可以有3个合法操作和2个不合法操作,而且可能陷入321-310-331-220-321循环。而110状态下的困难可能在于要作一暂时的倒退操作(110离000已经“不远”了,而221则似乎“远离”了000),被试们很难想到这一步,他们解题时用的大概就是“登山法”。

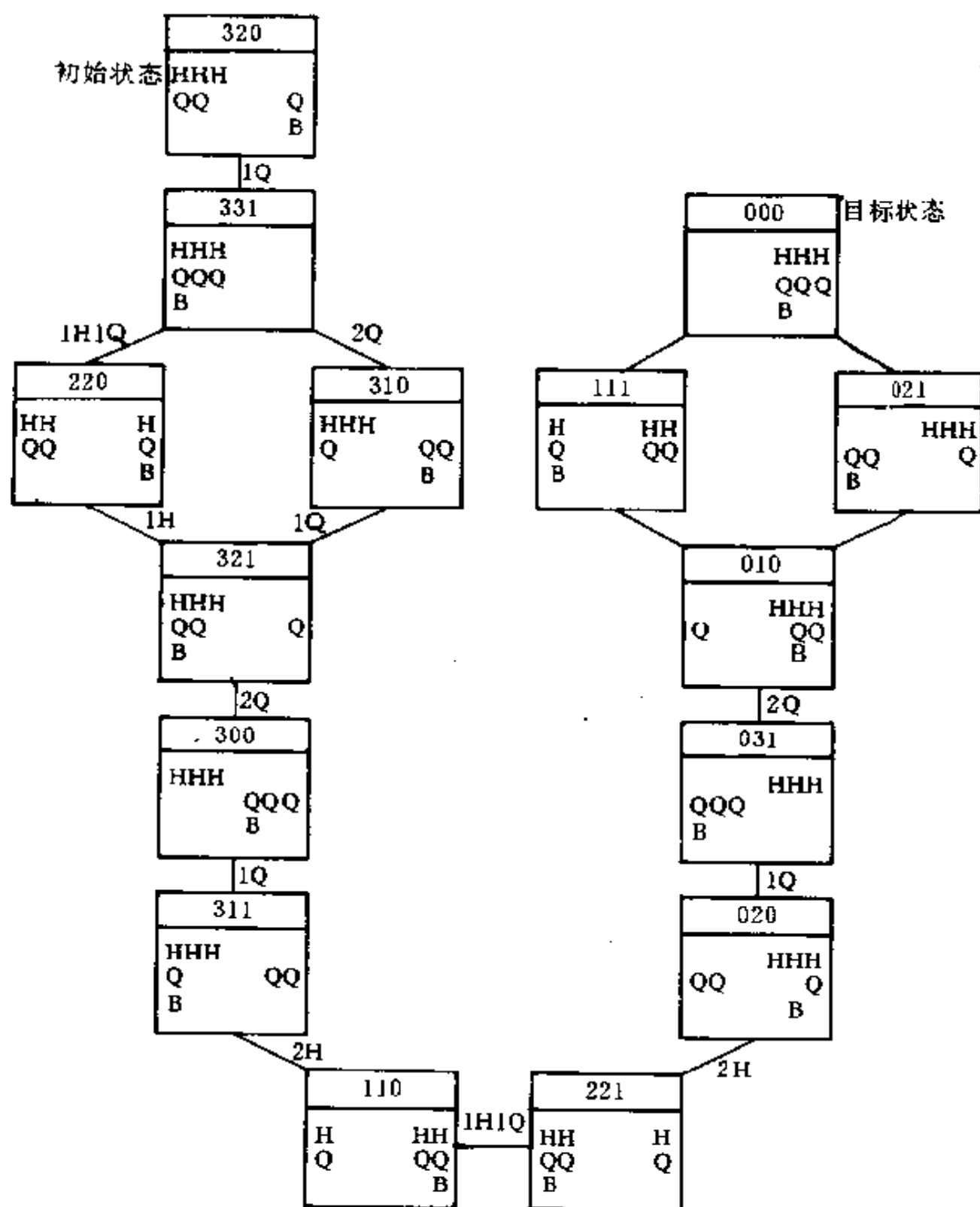


图2 富翁-强盗渡河问题答案

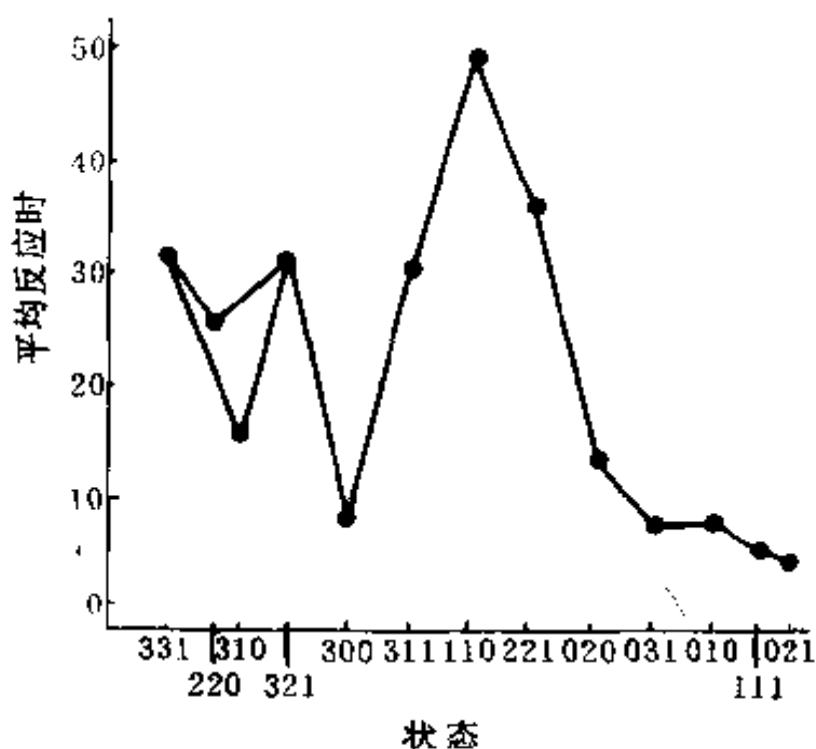


图3(1) 各状态下反应时

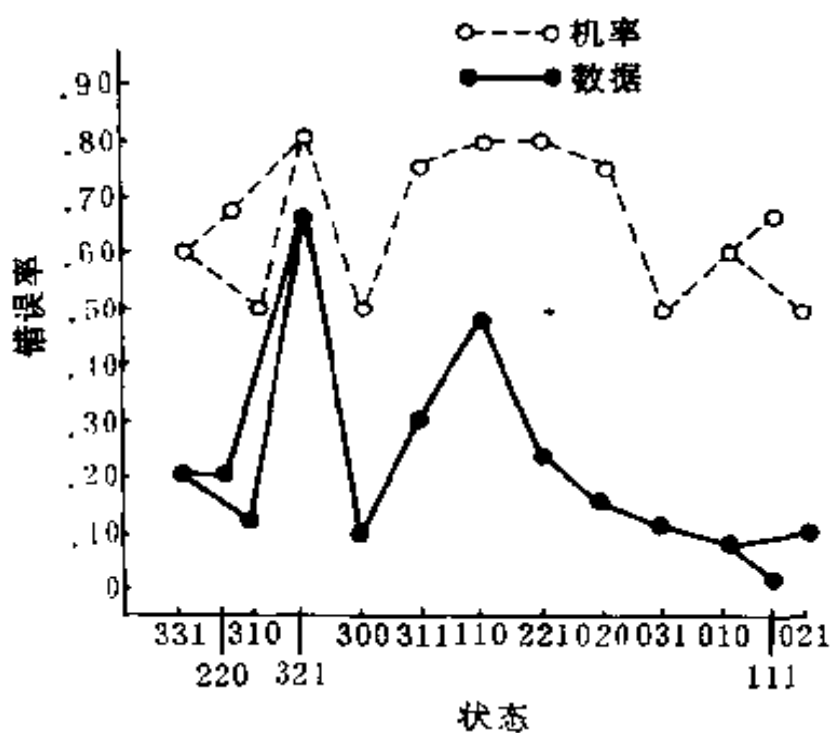


图3(2) 各状态下错误率

托马斯还报告了一个有趣的现象：有一组被试先从110状态做起，即先做本题后半部分，结果，再做前半部分时效率提高了（相对于从头到底做的控制组）；而控制组做后半部分时，似乎没从前半部分的操作中得到类似的好处，效率反而下降了。另外，先做过后半部分的被试在从头解题又做到后半部分时，反而比第一次做时效率降低。显然，前半部分的解题对后半部有消极影响。

1976年，西蒙和里德（Simon & Reed, 1976）将上述渡河问题略作改动：富翁和强盗由各3人改为各5人，小船可容人数由2人改为3人。大大提高了问题的难度。他们提出，被试解决问题时，策略上可能有变化。一开始，被试努力让两岸富翁和强盗的人数分别对等，即采用一种“平衡”策略，及至问题解决进展到某一环节时，策略发生变化，改为每走一步都力图使彼岸人数尽可能增加。另外，被试还要尽量避免出现前面曾经有过的状态，当然有时他们也随机地移动富翁和强盗。

为了检验上述假设，西蒙和里德做了两个实验。实验1被试分两组，控制组解决上述问题时不受任何提示，而实验组则被告知，在解题时会得出一个状态——彼岸有3个强盗，而小船却在此岸。用前面的数码表示，就是“521”。实验2的被试解2次本题，以观察训练效应。在所有实验条件下，都估价5个参数：①变换策略的概率；②检查有无陷入循环的概率；③随机移动的概率；④检查循环的概率有增加的比率；⑤随机移动的概率有增加的比率。结果发现，实验1中，实验组被试的策略由“平衡”法改为“手段目的”法的人数较控制组为多。这说明提示起了设立子目标的作用。实验2也表现出了预料中的训练效应。在第二次解题时，更多的被试出现了策略的变换。从上述两个实验来看，西蒙和里德的模型得到了证明。可以说，解决渡河问题时被试们确实无法预见许多步骤，只能“看一步，走一步”。这大概是短时记忆容量小造成的。

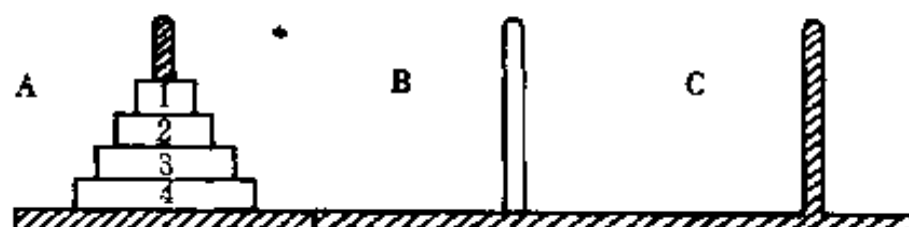
1977年，杰佛里斯等人提出了关于渡河问题解决的新模型。他

们认为,问题解决的每一步都涉及一种状态评价过程和记忆过程。被试评价每一步操作时可能遵循一定的规律。杰佛里斯提出这样一个“评价函数”:

$$e_i = aM_i + bC_i + cP_i$$

其中 M, C, P 分别是目标岸上的富翁人数、强盗人数和 $M-C$ 的配对数。例如目标岸上有2个富翁和2个强盗时,则 $M=2, C=2, P=2$ 。如果某一操作后 e_i 很大,则被试觉得这一操作很有价值。

记忆过程的特点则是这样的:工作记忆负责储存当前正在考虑中的少数几步操作;长时记忆负责储存已经经历过的各种状态。



4 个盘的河内塔问题

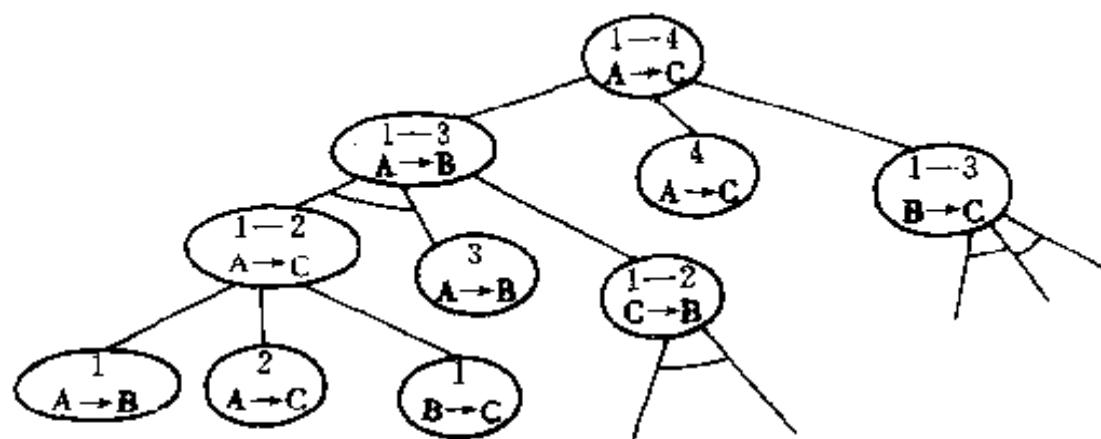


图4 河内塔问题分解简图

有些研究者还研究了类似于陆钦斯研究过的量水问题的解决过程,认为其模式颇类似于渡河问题。即多采取状态动作方式,一次考虑的步骤不多。

在河内塔问题和密码算题的解答中,研究者们看到了问题分解模式(图4)。例如,如吉尔(Luger, G. F., 1976)发现被试在解决问题时,往往连续的每一步都是为了接近某一个目标或子目标,而且一旦某一子目标达到了,达到子目标时用到的方法也往往会迁移到以后的操作中去。

西蒙对于河内塔问题解决也提出了几种策略。第一种是“目标递归”,实际上就是子目标策略。第二种是“知觉”策略。这种策略对记忆要求不高,只需按下列步骤行事:①找到尚未套入目标柱的最大盘(K);②找出阻碍K套入目标柱的最大盘;③如果没有其他盘阻碍,则将K移至目标柱;④如果有阻碍盘,则设立一个目标:将该盘移至其他柱子上;为了移动这个阻碍盘,还需重复②、③两步。这个策略只需记住一个目标,而不论盘数多少。而目标递归方式则需记住许多目标和子目标。第三种策略叫作“移动模式”策略,也可以说是一种口诀,共3句:

① 第奇数次移动时,移最小盘;

② 第偶数次移动时,移最上面的次小盘;

③ 盘数为奇数时,最小盘的移动顺序总是起始杆→目标杆→其他杆→起始杆……盘数为偶数时,顺序总是起始杆→其他杆→目标杆→起始杆……

当然,上述口诀并非人人能总结出来。

在河内塔问题的解决中,对问题的熟悉程度起了很大作用。安采和西蒙(1979)让一个被试做4次5个盘的河内塔问题,并要求他找出较好的解法。结果发现被试的策略不断进步。先是用状态动作模式,然后出现“知觉”策略,最后出现“目标递归”策略。这给研究者们一些启示和问题:也许问题分解模式是经验积累的结果?经验丰富的专门家与新入门的人思考问题有什么不同?由于人工问题意在平衡经验的影响,用它来研究专家思维已感吃力,于是新的研究模式出现了:结合专业能力和专业问题研究专家的思维。

知识和经验究竟是怎样影响思维的?某一领域的专家与刚入门的新手(以下简称“新手”)在解决问题上有什么不同?对于这类课题,恐怕只好用专业问题来研究了。目前国外除了用国际象棋、围棋、桥牌和扑克等比赛类专业问题来研究专家思维的特点之外,还广泛研究了数学、物理学、计算机科学、医学和政治学等多个领域的专家与新手思维的不同。

早在1965年,德格鲁特(De Groot)就从国标象棋高手和一般的好手的言语报告中总结出一个规律:他们在能预见的步数和搜寻走法的广度等方面均无差异,但高手却总能棋高一着。在进一步的研究中,德格鲁特分别给高手和新手看实际比赛棋局和随机摆放的棋子,然后让他们回忆(重新摆出)。结果,高手对比赛棋局的回忆成绩极好,而新手较差;但无论高手还是新手,对随机摆放的棋子的回忆成绩都较差。以后,在围棋、桥牌等项目的实验中也发现了类似的结果。

1973年,西蒙和吉尔马丁(Gilmartin)对上述结果提出了一种解释,其中运用了“组块”概念。他们认为,高水平的棋手在其长时记忆中储存着较多的模式或组块,在看到棋盘上有联系的子时,这些模式起着组织的作用,把许多棋子分成少数几个大组块,这样就出现了回忆成绩极好的现象。这就是“模式再认理论”。

但是,模式再认理论于80年代受到了一些批评。1981年,查内斯(Charness)研究了一些不同年龄、不同水平的棋手。结果发现,老年棋手和青年棋手水平即使相当,老棋手也在回忆棋子位置任务上成绩偏低。另外,棋手水平越高,其搜寻着法的深度、广度和速度也越大。这个发现与德格鲁特的结论也格格不入。

1982年,霍尔丁和雷纳兹(Holding & Reynolds)发现在记忆上没有差异的情况下棋艺高低能表现出来。他们让一些高水平的棋手和低水平的棋手同样完成两个任务:一是先看一些随机摆放的棋子,只看8秒钟,然后回忆;二是主试将刚才看的棋子重新摆

好,棋手决定走哪一步棋为最佳选择。结果是,棋手水平无论高低,回忆棋子位置的成绩都差不多;而在决定走哪一步棋时,高水平棋手的定棋质量显然较高。这说明,棋艺高低并没有表现在记忆的差别上。查内斯说明了记忆有差别的人的棋艺可以相当,霍氏和雷氏则发现记忆没有差别时棋艺也表现出高低。这两个结果是模式再认理论所无法解释的。

近年来,认知心理学家对学术领域的专家思维也进行了许多研究。

在物理学方面,蔡·格雷瑟和李兹(1982年)分析了物理学家和初学物理的人对20多个物理学习题的分类特点,发现初学者往往把表面上相似的习题分为一类,例如把附图中有斜面的习题都分为一类(“斜面问题”);而专家则把运用同一定理或解法相同的习题分为一类。这说明初学者易受问题的表层结构的迷惑,而专家则善于发现问题的深层结构——其内在含义。

在社会科学方面,沃斯等人(Voss et al., 1983)要求一组专门研究苏联问题的专家、一组正从事政治学研究的人和一组化学教授研究一个问题:怎样提高苏联农作物的产量?结果,苏联通中有许多人提出,问题讲得不明确,希望讲得再具体一些;而另两组被试则很少有人提出这个要求。从这个实验似乎可以看出专家思维的特点:在解答不太明确具体的问题时,他们倾向于将其分解成比较小的、容易把握的、明确的问题。

专家除知识经验高于新手以外,解决问题的策略是不是也与新手截然不同。一开始是有这种意见。例如拉金(Larkin et al., 1980)提出,新手常采用“向后法”(working backwards),而专家常采用从已知条件推向答案的“向前法”。但是吉克(Gick, 1986)提出,专家也有遇上难题的时候。这时他们往往时而用“向前法”,时而用“向后法”,甚至纯粹使用“向后法”,例如手段-目的分析法。心理学家以医学诊断课题来观察专家和新手解题策略时还发现,

专家和新手往往都采用“假设-演绎推理”的策略,即先根据症状提出1—2个假设,然后用演绎推理推测假设病症还应有什么症状,并检验有无这些症状。这里面既有“向前法”——提出假设,又有“向后法”——检验假设。在采用上述策略时,专家和新手在提出假设的多少以及提出假设前检验多少症状方面无显著差异,但是专家提出的假设质量高、具体明确、花费时间也较少。

据预测,未来对专家思维的研究可能集中到以下2个方面:①研究专家在解决难题时的行为;②尽量采用更精细的实验设计,研究多个技能水平的影响。

二、概念形成

概念形成的研究由来已久。自从1920年贺尔引入人工概念以后,这个领域的研究更趋活跃。海德布莱德、布鲁纳、古德瑞、奥斯汀、列温(M. Levine, 1975)运用人工概念所进行的研究也已为我们所熟悉。另外,对自然概念形成的研究也取得了很大进展。本文不打算罗列上述全部研究成果,而是仅就其中一些方面介绍一些深一步的研究。

首先是关于概念形成中被试策略的研究。我们都知道,布鲁纳、古德瑞、奥斯汀在他们于1956年发表的著名研究报告中认为,被试在形成人工概念时,采用了各种策略,即保守性聚焦、赌胜性聚焦、同时性扫描和继时性扫描。但是在以后的20多年时间里,对于他们的发现所作的解释是十分简单的。一般认为,被试把人工概念当作一个等待证实的假设;而他们在这个任务是很有策略的,策略的选择则取决于运用某一策略时的记忆负荷,被试总是偏爱负荷量小的策略。另外还有一种看法,运用不同的策略时,人们的心智操作活动也有明显的不同。

以上解释渐渐引起了学者们的怀疑。有些学者指出,从行为表

现主要将不同的策略分析出来是困难的。约翰逊(Johnson, 1978)举了一个例子:假如用某些问题来鉴别赌胜性聚焦策略,那么可能出现如下情形。一位被试在其首次尝试就猜对了,这时如果他接下去只改变1个维度,则似乎应该说他运用的是保守性聚焦策略;他也可以改变2个或3个维度,这时用的就是赌胜性聚焦策略了。但是,当被试用了一次赌胜性策略而且成功之后,就只剩下最多1个维度尚未检验了,而对剩下的这1个维度,他无法继续运用赌胜性策略了。这就是说,只根据一次试验就认为被试使用了某个策略似乎是不合理的。约翰逊进一步提出,如果解决某一概念问题时,策略甲确实高于策略乙,那么运用策略甲的被试成绩就会好一些,统计上就可以画出双峰曲线。

为了检验自己的观点,约翰逊设计了一种新的人工概念形成任务,叫作 zaps-duds(ZD)任务。在这个任务中,被试在看到由字母“X”和“O”组成的字符串(字母个数均为6)的基础上找出为什么有些字符串是“zap”,而有些是“dud”的规则。实验在计算机上进行,字符串和假设由被试键入,而计算机则给出“zap”或“dud”反馈,或告知被试假设是否正确。实验的统计量是形成概念所需的尝试次数。结果发现画出的曲线有4个众数(或峰)。约翰逊把这个结果解释为被试运用了4种策略。第一种叫作聚焦策略,约10次即可完成任务;第二种叫“双X型策略”,即每个字符中只输入2个“X”,约15-16次可完成任务;第三种策略是,被试输入几个字符串以后,转而只输入自己的假设,而不再输入字符串,这样做,需21次尝试可完成任务;第四种叫作“纯扫描策略”,被试从开始就不输入字符串,而只输入自己的假设,这种策略需30次尝试才能成功。约翰逊还发现,被试们各自使用的策略多数是不变的。

有些学者,例如迈丁和史密斯(Medin and Smith, 1981),提出这样一个论点:不同的策略并不影响认知操作的性质;被试选择策略仅仅对于需要记住多少信息有影响。而劳林等人(Laughlin,

Lange, Adamopoulos, 1982)的研究提出了相反的意见。劳林等人也让被试完成一个概念形成任务,这个任务要求被试找出一种颜色密码。劳林等人设想被试在完成任务过程中可能有2种策略,一是聚焦策略;二是一种类似于布鲁纳等人提出的扫描策略。结果发现,36%的被试运用了聚焦策略,他们平均作了5.27次猜测就完成了任务(理论上认为最少也要作4.5次猜测);另有31%的被试运用了扫描性策略,他们平均作了5.98次猜测(理论上的极限是3.02次)。可见尽管扫描性策略是一种有力的工具,但是被试运用起来却有困难。有意思的是,运用上述两种策略的被试比运用其他策略的被试做得出色一些,而当被试被引导着运用上述策略时,他们的成绩就上升。这些结果说明被试所选择的策略很可能对认知操作性质也产生了影响。

在布鲁纳等人看来,概念形成就是假设检验的过程。而这似乎也是很长一段时间里心理学界的共识。但到了80年代,这个见解也受到了挑战。挑战者的实验根据主要是假设记忆和内隐学习。

如果概念形成就是假设检验,那么被试在实验过程中对自己的假设应该是记得很清楚的,至少应该比对前面的刺激和反应记得清楚。但问题似乎并不是这样简单。柯洛格等人(Kellogg, et al., 1978, 1980, 1983)曾经用相当简单的概念问题研究被试在概念形成中的某些时刻,对前面一次尝试中的刺激、假设、反应和反馈的记忆情况。结果发现:①被试对前面一次尝试时的反应和反馈的再认成绩相当好;②对于刺激和假设的再认则较差,而对假设的再认为最差。更不同寻常的是,即使那个假设得到支持,其再认成绩也不比被否定的假设高。

柯洛格等人对上述现象作了如下解释:被试一开始注意的是概念正例中的特征;当要求他们提出假设时,他们通常选择最常出现的特征。而特征出现的频次是自动地、在无意识中累加的。所以学者们提出,概念形成的过程可能有两种类型:在某些情况下,人

们用假设检验的方法解决概念课题;但在另一些情况下,假设检验并不构成任务的中心成分,许多被试依赖于无意识中累加的频次来区分刺激的相关特征和无关特征。

尽管概念形成可能有以上两种途径,但是据多米诺斯基(1974)考察,如果坚持要求被试使用假设,被试们的成绩会提高。这说明,假设检验也许是概念形成的高效率手段。如果不是里伯和艾伦(Reber and Allen, 1978)以及布鲁克斯(Brooks)的一系列内隐学习研究提出了相反的证据,上面那个结论也许早就为人们接受了。里伯和艾伦精心设计了一种“人工语法”,用这种“语法”可以组成一个个字符串,里伯称这些字符串为“单词”。里伯在实验中把被试分成两组,并要求一组被试“努力记住字符串”;要求另一组被试“找出字符串排列的规则”,然后向被试一一呈现字符串,最后让被试们判别一些新的字符串是否符合语法。结果,那些仅仅要求记住字符串的被试成绩显著高于要求找出规则的被试。这说明,原先没有意识到字符串里头有什么规则的被试反而比较多地学到了里伯的人工语法。但是,这些被试却难以用语言准确地表述自己学到的东西。里伯等人还发现,这些被试学到的规则还不容易遗忘,并能长期运用。这种无意识中习得概念规则的学习,就叫作内隐学习。

里伯等人的研究结果不仅对假设检验理论是一个挑战,而且对有关概念学习的所有分析性理论都是一个挑战。他们的研究说明了一个问题:当刺激结构高度复杂时,相对被动的、非分析性的学习方式最为有效。里伯还认为,人类语言和另一些抽象概念的复杂结构就是在这种无意识的内隐学习中获得的。

目前,国内外对内隐学习的研究十分活跃。其研究的主要问题是:内隐学习得到的知识到底是什么?是人工语法本身?是一些简单的规则?还是许许多多的组块?这些知识又是怎样得到的?等等。

具体地说,内隐学习和外显学习一样,均属于复杂认知任务的

操作。外显学习过程类似于有意识的问题解决,它包括试图形成一个任务的心理表象,建立和检验任务操作的心理模型。这种过程是有意识的,需要作出努力的。在外显学习过程中,人们试图通过检验有意识的假设来发现规则。与外显学习过程的特点相反,内隐学习机制是自发地起作用的,不须任何有意识的智力来概括复杂的关系。

研究表明,使用无意识的学习机制发现的有意识思维更能检测微妙的和复杂的关系。而且,这种内隐学习过程是自动的和无意识的。雷伯(A. S. Reber)等人报告了内隐学习的一系列实验,从他们的研究中我们可概括出内隐过程的本质:①内隐知识能自动地产生,无需有意识地去发现任务操作内的外显规则(Reber, 1969, 1976);②内隐学习具有概括性,很容易概括到不同的符号集合(Brorks, 1978, 1988);③内隐学习具有无意识性,且内隐式获得的知识是不能用言语来表达的(Reber, 1976; Lewicki, 1986)。其实,内隐学习的特点,不尽这些。

国外的实验研究主要是以人工语法为对象进行的。内隐学习研究的创始人 Reber 及其同事们进行坚持不懈的研究,积累了丰富的成果。他们的做法是:在内隐学习实验的学习阶段,向被试呈现一系列由一套特殊规则产生的字母串,这些规则构成了指定各字母许可顺序的一种人工语法,给予被试指导语,或是“学习并记住这些串”(内隐学习),或是“发现隐藏的规则,以便更好地记住这些串”(外显学习)。换句话说,记忆指导语产生内隐学习条件,而规则发现指导语产生外显学习条件。在实验的测试阶段,向被试呈现一些新的字母串,让其判断哪个是符合由学习阶段刺激材料构成的规则,哪个不是。

我国心理学工作者杨治良于1991年在雷伯、马修斯(R. Mathews)研究的基础上,探讨了在无意义字母串学习中内隐学习的某些特点,兼用信号检测论分析内隐过程的情绪特点,同时也观察不

同条件下的启动效应(priming effect)。实验用无意义字串中安排“SCT”关键字,在每组的A,B两个系列,安排“SCT”的字串的比例是不同的。在第1,2组中,A系列中含“SCT”的字串较多,而B系列中含“SCT”的字串较少。在第3组中,A和B两个系列含“SCT”字串的比例恰巧相等。到了第4,5,6组,情况又反过来了,在A系列中含“SCT”的字串较少,而B系列中含“SCT”的字串较多。正是针对这一情况,测验中短暂呈现一个含有“SCT”字串或不含有“SCT”字串的刺激,要求被试确定它是“A”系列还是“B”系列。通过被试的选择情况,对各组进行比较。也就是说,通过被试对呈现无意义字串中“SCT”关键字的分辨能力的提高,来反映被试的启动效应。A和B两系列所含“SCT”字串频率之差异用对比度表示。对比度反映了学习次数的多少,这种学习效果在反应中显示出来。但是,他们(被试者)是在不知不觉中习得的,甚至自己也不知道学到了什么,只是从反应上看到了各实验组之间的明显差异。这种学习机制比发现的有意识思维的外显学习更能检测微妙而复杂的关系。这种学习过程比外显学习过程更有力,是自动的和无意识的。我们把这种学习过程称之为内隐学习过程。内隐记忆这一效果,也就是启动效应。启动效应是由于近期与某一刺激的接触而对这一刺激的加工得到易化。启动效应的测验,除了杨治良设计的这类知觉辨认方法外,还常用词语确定(lexical decision)、词的确 定(word identification)和词根或残词补全(word stem or fragment completion)等称之为补笔的方法。在补笔测验中,把学习的单词选出来加以残化,好比table--tab,让被试用第一个想到的词把它补全,如果被试表现出用词表上的词来补全残词的倾向,则表现出“启动效应”。正是由于这些测验都不要求被试直接回忆项目,因而它们常常作为内隐学习研究的特有的间接测量方法。

科洛格(1980)也发现了一个类似内隐学习的现象。他让一些被试做50道乘法题(心算),这可以转移被试的注意;又在被试算题

的同时呈现30张脸谱图。随后进行一项被试们意想不到的脸谱图形再认测验。结果,被试竟能相当不错地分辨出,哪些图是呈现过的,哪些是未呈现过的;成绩显著高于机遇水平。同时,被试对那些先前没有呈现过但又与呈现过的图形相象的脸谱图却束手无策,搞不清是否呈现过了。换句话说,被试们记住的不是一张张具体的图,而是掌握了那些脸谱图的某些特点。这个结果也说明了相当复杂的概念可以在无意识中偶然地获得。

在自然概念的研究中,值得一提的是一种新的研究材料。这种材料叫作假想动物,由一位叫卡明的人首创。他画出来的这些动物自然界并不存在,但动物身上的特征却是可以在自然界中找到的。用这种假想动物构造的“物种”或概念,既有人工的色彩,又有自然的特征,姑且称之为自然人工概念(或人工自然概念)。索卡尔(Sokal,1977)就运用这种材料进行过一个研究。他让3位有学问的人对一些假想动物进行分类。结果发现,尽管这3个人对各个动物应归入哪一类意见分歧不大,但是他们在根据哪些特征(或方面)来确定类别方面存在重大差别。

对于自然概念,研究得最深刻、名声也最大的当然是罗施(Rosch,1973年及以后)。她提出的概念形成的样例理论与假设检验理论并驾齐驱。

样例理论与假设检验理论是针锋相对的。在罗施看来,自然概念不像人工概念那么确定,人们头脑中的自然概念不是一个或几个关键特征,而是对概念样例的记忆。换句话说,自然概念的形成用不着假设检验参与。但是,近年来的研究发现上述说法并非完全正确。

马丁等人(Martin and Caramazza,1980)提出,成人在某些情况下会使用假设检验策略来形成自然概念。他们让被试对一些脸谱进行分类,并告诉这些被试(都是大学生),脸谱中没有哪一个或哪几个特征可以总是作为分类的决定性依据。尽管这样说了,被试

们却还是采用了假设检验策略,即系统地考察脸谱的各个具体特征并进行分类。

1981年,奥谢森和史密斯提出了一个新观点。他们认为,人们记忆中可能存在两类信息,一类是样例信息;一类是类别信息。当出现一个刺激,并且人要迅速判断这个刺激是否属于某个类别时,就要用到样例信息;而在严格地从逻辑上进行证明时,被试就要依靠类别信息——概念的定义、关键特征以及特征间的相互关系。拿这个观点去看前面马丁等人的研究结果就很容易理解了:正是自然概念的类别信息成了假设检验的基础。

自然概念有其类别信息,那么人工概念是不是也有样例信息?或者说,确定的、逻辑性强的概念是否也具有自然概念的一些特点?布尔尼(Bourne, 1982)的研究证明了这一点。他用完全精确定义的概念再现了原来认为只有在自然概念身上才有的现象,包括原型的选择,类别隶属程度的评价等。阿姆斯特朗等人(Armstrong, et al., 1983)甚至发现连数学上的奇数、偶数等完全精确定义的概念居然也有类别成员的隶属程度问题。他们让被试评价一些数“像”奇数或偶数的程度。照理说,奇数就是奇数,偶数就是偶数,哪有像不像的道理。可是结果确实说明,被试对不同的数字有不同的评价。阿姆斯特朗等人还让被试评价“母亲”、“主妇”、“公主”、“女招待”、“女警察”、“女喜剧演员”对于“女性”的隶属程度(或者说是例子的“好坏”),结果是:母亲是“女性”的最好例子,而女喜剧演员为最差。另外,被试在评价一些图形,如对于“平面图形”的隶属程度时也表现出差异(见表1)。阿姆斯特朗等人还发现,如果被试认为某个例子比较“好”,他反应得会快一点;如果他认为“不好”,反应得就会慢一点。上述结果说明,自然概念和人工概念之间没有不可逾越的鸿沟,它们可以得到同样的表征,并表现出同样的性质。这样,奥谢森和史密斯的两类信息的理论不仅在自然概念那里得到了证据,而且在人工概念和其他精确定义概念那里也

得到了支持。

表1 对一些样例的隶属程度评价

类 别							
偶 数		奇 数		女 性		平面几何图形	
样例	得分	样例	得分	样例	得分	样例	得分
4	1.1	3	1.6	母亲	1.7	正方形	1.3
8	1.5	7	1.9	主妇	2.4	三角形	1.5
10	1.7	23	2.4	公主	3.0	长方形	1.9
18	2.6	57	2.6	女招待	3.2	圆	2.1
34	3.4	501	3.5	女警察	3.9	不规则 四边形	3.1
106	3.9	447	3.7	喜剧女 演员	4.5	椭圆	3.4

近年来,许多学者积极投入对样例信息和类别信息的区别的研究。对于完整的概念形成理论究竟应该以样例信息为基础还是以类别信息或者以上述两者为基础,学者们各执一词,争论不休。不过,正如布尔尼(1986)所说的,一旦我们找到合适的角度,我们最终会发现,表面上水火不相容的理论实际上是相通的,这种情形历史上不乏先例。

三、推 理

人们的思维都是有逻辑的吗?显然不是。有位权威人士艾瑞克森(Erickson, 1978)甚至认为,事实上对多数三段论推理来说,人们的反应是错误的。

心理学家研究得比较集中的一种推理错误,就是“气氛效应”。这个效应其实早在1935年就由伍德沃斯和赛尔斯(Sells, S. B)提

出来了。他们认为,前提可能影响人们对结论的预期。例如,两个前提都是全称判断(即“所有的……”)时,人们就容易接受全称结论。所以,很多人会认为下面这个推理是正确的。

所有的 C 都是 M;
所有的 D 都是 M;
所以,所有的 C 都是 D。

后来,亨勒(Henle,1962)提出,人们推理上的错误可能源于对前提或任务要求的误解。但是她进一步又说,人们推理过程本身是合乎逻辑的。她这个观点是很难得到证实的,因为任何一种推理错误都可以归咎于某种误解,而不归咎于没有遵守逻辑规则。不过“误解”这个问题提得好,以后许多研究就是围绕着它进行的。

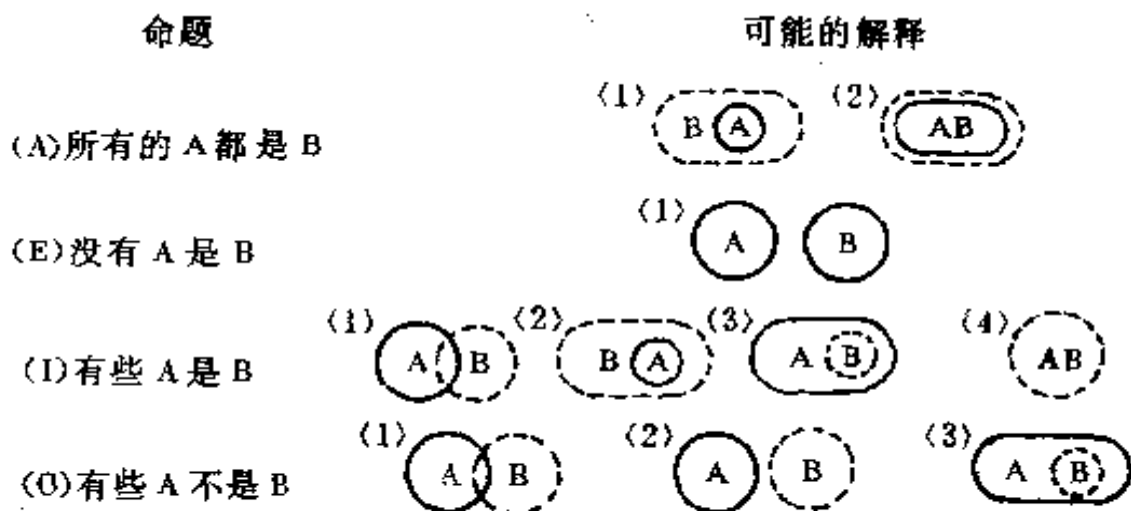


图5 各种命题可能的解释

1971年,塞拉索和普罗维特拉(Ceraso and Provitera)提出,三段论推理中前提的表达是不太明确的。对于全称肯定(A)可以有2种理解;对于特称肯定(I)可以有4种;对于特称否定(O)可以有3种;全称否定(E)最明确,只有1种理解(见图5)。这种不明确,使人们往往根据自己的解释来推理。所以,如果把前提讲得具体明确,推理时就可能少犯错误。例如,不说“有些 A 是 B”,而是说“有些

(但不是全部)A 是 B,而所有的 B 都是 A”,前提就明确了。塞拉索和普罗维特拉就根据这个思路设计了一个实验:让一组被试根据讲得十分明确的前提进行推理,而让另一组被试仍然根据用传统方式表述的前提进行推理。结果发现,前面一组在13个推理中,平均只错了2.5个,而后一组平均错误数多达5.5个。这个研究说明了对前提误解的结果。

艾瑞克森(1974,1978)进一步研究了对前提的理解问题,并提出了自己的理论模型。他首先让几所大学的学生用图来表示他们对“所有的 A 是 B”这个命题的理解,发现大约40%的反应是“A=B”,60%的反应是 B 包含 A(图6)。接着,艾瑞克森重点研究了两个前提都是全称肯定命题(AA)时的推理情况。

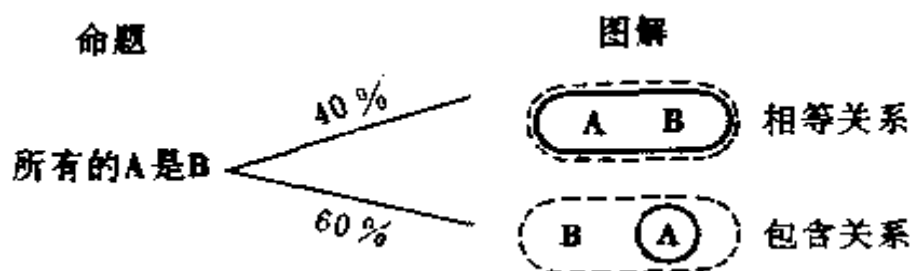


图6 对命题不同理解的百分比

首先,他预测了上述两种对全称肯定命题的理解在4种不同的逻辑格的情况下会导致什么样的结论,并提出了被试作出各种结论(A,E,I,O)的各个比例。随后,他用5个独立研究得出的数据来检验,发现这些数据与预测的比例十分符合。于是,艾瑞克森提出了推理的三阶段模式,即①理解前提;②将理解了的两个前提综合一下,得出一个结论;③根据结论和任务要求(例如是画图表示,是选择,还是口头回答等等)来作出反应。

显然,从亨勒到艾瑞克森,研究思路都是一脉相承的。而进入80年代以后,出现了一些新型的模式。

1981年,居约特和斯腾伯格(Guyote and Sternberg)提出了一

个比较复杂的模型。这个模型把人的理想推理能力和实际推理行为区分开来。这颇有点像心理语言学把人的语言能力和实际的言语表现区别开来。他们认为,被试看到前提和几个供选择的结论时的理想推理过程应该有以下两步:①对前提进行编码。②将前提信息加以综合。在编码时,被试应运用符号来表征各词项之间的关系,对于每个前提的每一种可能的解释都要考虑到,并用符号加以表征。接着,还要考虑两个前提的可能解释之间的所有组合,然后才能作出结论。这样,如果两个前提都是特称肯定命题,由于这种命题可以有4种解释(见图4),这样就可以得出16种组合。而实际上,很少有人能考虑这么多组合。所以,他们又提出推理的实际模型:①对前提进行编码;②对前提信息加以综合,这时被试实际上往往至少考虑1种组合,但至多也不超过4种,这是工作记忆容量所决定的;③选择一个标记来说明被综合的表征,这些标记就是前面提到过的A,E,I,O和“无结论”。这时被试往往倾向于选择较少可能解释的命题(例如“O”,“E”)和适合于前提“气氛”的命题;④进行反应。居约特和斯腾伯格用实验证明了自己的理论有一定的道理。但他们理论模型也受到了一些批评。

迄今为止,对推理问题解释得最具体、最具一般性的模式恐怕就是约翰逊-莱尔德(Johnson-Laird)等人的心理模型了。他们在实验中发现,对于“有些父母是科学家,所有的科学家都是司机”这两个前提,被试们倾向于得出“有些父母是司机”这个结论,而较少得出另一个同样正确的结论:“有些司机是父母”。其实,实验中凡是像上面那种逻辑格,即“A—B;B—C”(第四格)总是倾向于引出“A—C”(“C—A”也是对的)结论。另外,“有些科学家是父母,而所有司机都是科学家”(逻辑上是第一格:A—B,C—A)总是偏向于引出“有些司机是父母”(C—B)这一结论;而另一正确结论:“有些父母是司机”(B—C)却很少有被试发现。这种效应是过去不曾发现过的,故定名为“逻辑格偏向”。

约翰逊-莱尔德等人于 1978 年提出一种解释逻辑格偏向的原因,认为这是由于在短时记忆中组合前提表征所造成的,即对于 $A \rightarrow B, B \rightarrow C$ 前提,被试可以先对前一个前提进行编码,然后将后一个前提的表征直接联在后面,这样就得出 $A \rightarrow C$ 结论;而在 $B \rightarrow A, C \rightarrow B$ 情况下,由中项(B)不相邻,就必须先将第二个前提($C \rightarrow B$)储存在短时记忆中,而后再回忆第一个前提($B \rightarrow A$),得出 $C \rightarrow B, B \rightarrow A$ 这一组合(这与 $A \rightarrow B, B \rightarrow C$ 组合在形式上一致),从而得出 $C \rightarrow A$ 结论。

约翰逊-莱尔德和巴拉(Bara, 1984)还发现了一个新的效应。他们让被试看有关某个推理的结论,但结论只呈现 10 秒钟。结果,不但同样发现了明显的逻辑格偏向,而且还发现,被试反应“没有确定结论”的次数,在 $A \rightarrow B, B \rightarrow C$ 格情况下最少,而 $B \rightarrow A, B \rightarrow C$ 情况下最多。他们认为,这说明在某些逻辑格出现时,由于呈现时间短,被试难以组合前提,从而造成较高的错误率。从这个结果也可以看出,前提的整合过程确实是在工作记忆中进行的。

前面提到的都是演绎推理。归纳推理的研究在当代心理学界也受到了十分广泛的重视。其中对所谓的“4 卡片选择问题”的研究成了人们了解归纳推理过程和特点的一个重要窗口。

“4 卡片选择问题”是沃森(Wason, 1966, 1968)设计出来的。问题是这样的:在下面 4 张写有数字和字母的卡片上,都是一面是数字,另一面是字母;现在要验证下面这个命题——如果一面是元音字母,另一面就是偶数——请问你准备翻哪些卡片来验证?

[E] [K] [4] [7]

从逻辑上讲,要证明“如果 p , 那么 q ”这个命题的正确,只有从 $pq, p\bar{q}, \bar{p}q$ 和 $\bar{p}\bar{q}$ 这 4 种情况中找证据。其中只有 $\bar{p}\bar{q}$ 这种情况与原命题相冲突,而 pq 支持原命题, $\bar{p}q$ 和 $\bar{p}\bar{q}$ 与原命题无关。所以,上述问题的答案应该是翻“E”(p)和“7”(q)这两张卡片。但是,大多数被试却去翻“E”和“4”这两张卡片。对于这个结果,有一种解释

认为被试们有意去证实而不是推翻原命题,所以总是去翻可以支持假设的卡片(p 和 q),而忽视了有可能推翻原假设的卡片“7”($\bar{q}$)。如果拿比较具体的、贴近生活的假设,或者用大家都知道是错误的假设(例如“学哲学的人一定在牛津大学”)来给被试验证,被试的成绩就会好些了。

被试翻错卡片,也可能是因为理解上有问题。所以,布累丝韦尔(Bracewell, 1974)将问题阐述得更加清楚:“如果卡片的一面是‘J’,那么另一面是‘2’;但这并不是说,‘2’只能与‘J’写在一张卡片上。现在请你指出,为了检验上述假设,应该翻哪两张卡片?”这样一来,被试成绩好多了。

伊凡斯(Evans, 1984)等人则提出,被试翻错卡片是因为他们仅仅去翻那些写有命题中提到数字和字母的卡片,而没有进一步地加工信息。所以,如果假设是“一面是‘B’,另一面就不是‘3’”的时候,翻对卡片(应该翻‘B’和‘3’)的被试就多了。伊凡斯认为,被试翻卡片时的理由,他自己未必意识到。他在这个基础上提出两种不同的思维模式,一种叫“Ⅰ型”,一种叫“Ⅱ型”。Ⅰ型思维是非语词的、不能内省的(似乎也可以说是非逻辑的),而Ⅱ型思维则是语词性的、“合理”的思维。为了证明被试在完成4卡片选择问题时的思维模型是Ⅰ型,伊凡斯和沃森在被试解决完问题以后,把几个“标准答案”告诉被试,并请他们证明这些“标准”(其实有的是错的)答案是正确的。结果,被试们对所有“标准”答案(包括错误的答案)都很有把握地说出理由,证明其“正确”。这就说明,被试在解答问题时可能没有很清楚地意识到自己翻哪几张卡片的理由,但如果事后问及,他们会临时建构出理由来。

伊凡斯还进一步提出推理中的两种过程——“启发”过程和“分析”过程。启发过程只是注意到问题的表面特征,而分析过程则可以进行一些推论。对4卡片选择问题的研究说明,解决这个问题时被试用的多数是启发过程,即只是注意到问题中提到的内容。

心理学家还发现,在其他某些类型的任务中,被试有时也表现出上述现象。例如,沃森曾经做过这样的实验:给被试看3个数字(2,4,6),要求他们发现这些数字遵循一个什么样的规律;实验时让被试按照自己设想的规律说出后面应该是什么数字,每次连说3个,主试则对被试进行反馈。被试一旦十分确信自己的假设成立,就说出自己的假设。这个实验揭示出被试中一种压倒性的倾向:他们说出的数字系列总是和自己提出的假设相一致,直到有信心声明自己的假设正确为止;很少有人提出与自己的假设相违背的数字系列来检验假设或改变假设。结果,只有20%的被试发现了真正的规律——“后面的数大于前面的数”,而许多被试只是找到一个比较狭义的规律——“一个等差数列”。后来的实验也证明,如果不明确地告诉被试还可以采用另一种策略,他们就会一直沿用原来的方法。比如说,在指导语中提醒他们找出所有可能正确的规则,但没有效果。

四、决 策

决策过程和问题解决是有密切联系的。问题解决总要包括一定的决策,例如确定某种解法是否可行,或某一思路是否可以优先考虑,等等。而决策则可以看作是问题解决的一种特殊形式——从某些供选择的内容(如方案)中选择最好的一个。但是,尽管两者有密切联系,它们各自的侧重点却不同。问题解决侧重于产生想法,而决策侧重于从多种想法中加以选择。

决策可以根据不同的标准进行分类。根据决策的后果有无不确定性,可以分为风险性决策和非风险性决策;根据被选择的内容具有多少个特征,可以分为单因素决策和多因素决策;根据决策需要多少次选择,可以分为单阶段决策和多阶段决策。本节主要从风险性决策和非风险性决策来介绍一些研究成果。

研究决策过程的思路主要有两条。一条思路已经想了很久,就是研究一个理想的决策者应该怎样进行决策,并以此来劝戒决策者。这是一种“标准化”思路。在这个基础上,可以研究人们在实际决策中偏离那个“标准”多远。比较新的一条思路,就是研究决策背后的认知过程,研究其加工模式。

先谈谈对风险性决策的研究。人们很早就发现,决策者在面临风险时的决策与标准的理论计算不一致。对此,卡耐曼和特弗尔斯基(Kahneman and Tversky, 1984)提出了一个概念,叫作“损失厌恶”,也就是说,人对损失 1 元钱的厌恶要比对获得 1 元钱的欢乐强一些。下面这个简单的实验就能说明问题。他们问被试:如果一笔生意可以稳赚 800 美元,另一笔生意则有 0.85 的机会赚 1000 美元,但有 0.15 的机会不赚分文,那么你愿意做哪一笔生意。按照标准的决策理论应该这样计算:

$$\text{第一笔生意的期望值} = 1.00 \times 800 = 800$$

$$\text{第二笔生意的期望值} = 0.85 \times 1000 + 0.15 \times 0 = 850$$

从计算结果看,应该做第二笔生意。但是,大多数被试却选择了不承担任何风险的第一笔生意。卡耐曼和特弗尔斯基认为这就是人们厌恶风险的表现。经过细致的研究,他们制定了一条函数曲线来反映人们对于不同数量的金钱得失的评价(用“价值”表示评价的正负高低)。从图 7 可以看出,曲线在收入这一段上升速度随收入数量增加而减慢;在损失这一段,曲线却是随损失量的减少而加速上升。另外,曲线的斜率是损失这一段大,收入这一段小。这就是“损失厌恶”的具体表现。

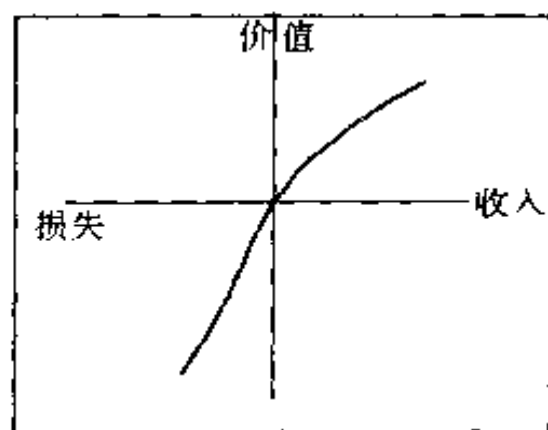


图 7 价值函数

卡耐曼和特弗尔斯基还指出,人们有时也会甘冒风险。如果把

刚才那个问题变成：第一笔生意有 0.85 的概率损失 1000 美元，而有 0.15 的概率不受任何损失；第二笔生意则无论如何总要损失 800 美元，问愿意做哪一笔生意。结果多数被试却选择了第一笔可能担多赔 50 美元风险的生意。看来有这么个规律：在有钱赚的时候，人们不愿冒险；而在要赔钱的时候，人们愿意冒险。

特弗尔斯基和卡耐曼(1981)还研究了被选择项目的表达方式对决策的影响。他们对被试提了这么一个问题：假定美国正在面临一种疾病的爆发流行，并已经估计到在这场流行病灾难中将有 600 人死亡。现在提出两种方案(A 和 B)，估计它们的效果是：①如果用方案 A，将有 200 人得救；②如果用方案 B，600 人全部获救的可能性是 $1/3$ ，而全部死亡的可能性是 $2/3$ 。请问应该采取哪一种方案？对于这个问题，约 72% 的人选择了方案 A。

后来，他们又对另一些被试提了上述问题，但这一次提出的是方案 C 和 D，并估计它们的效果是：①如果用方案 C，将有 400 人死亡；②如果用 D 方案，则没有人死亡的概率为 $1/3$ ，600 人全部都死亡的概率为 $2/3$ 。结果，这一次有 78% 的人选择了方案 D。但实际上，方案 A 和方案 C 是等值的，方案 B 和方案 D 也是等值的。可见方案的表达方式对决策有很大的影响。对此，特弗尔斯基和卡耐曼是这样解释的：在前一种情况下，被试考虑的是救人，所以他们不愿意采用方案 B，冒 600 人全部死亡的风险；而在后一种情况下，被试考虑的是会死多少人，于是就甘冒风险，以期 600 人全部救活。

对于比较复杂的推理，卡耐曼和特弗尔斯基(1984)也进行了一些研究。他们向 150 个被试提出下面这个问题(括号里是选择前面所讲的方案的人数百分比)：

设想你要同时作出下面两个决策，请先仔细考虑这两个决策，然后指出你愿意采用的方案。

决策(一)从方案 A 和 B 中选择；

A. 稳赚 240 美元 (84%)

B. 25% 的机会赚 1000 美元而
75% 的机会一分钱也赚不到 (16%)

决策(二)从方案 C 和 D 中选择:

C. 总要损失 750 美元 (13%)

D. 75% 的机会损失 1000 美元而
25% 的机会一分钱也不损失 (87%)

从以上结果可以看出选 A 的多于选 B 的,选 D 的多于选 C 的。另外,即选 A 又选 D 的占了 73%,而只有 3% 的人同时选了 B 和 C。但实际上,选 B,C 比选 A,D 强。因为选 A,D 的综合结果是 25% 的机会得 240 美元,75% 的机会失 760 美元;选 B,C 的综合结果则是 25% 的机会得 250 美元和 75% 的机会失 750 美元。接着,卡耐曼他们变换了一下问题的提法,原来决策(一)和决策(二)的方案都是分开写的,现在把它们组合起来,即从 A—C, A—D, B—C, B—D 中选择方案。结果被试们一致选择了正确的答案。

对于非风险性决策,学者们也作了很多研究。

卡耐曼和特弗尔斯基(1984)发现,人们在决策是否花力气去买廉价商品取决于原价与可以省的钱的多少。例如有个问题说:设想你去商店买一件 125 美元的衣服和 15 美元的计算器,后来听说离该店 20 多分钟路程(开车去)有个店计算器只卖 10 美元。问你愿意跑一趟吗?

对于上面这个问题,卡耐曼他们提出 3 种可能的策略。一是只看到能省钱就去跑一趟;二是考虑可以省多少钱,也考虑计算器的原价(15 美元);三是考虑省多少钱,并考虑个人所有的花费(包括衣服等等)。实验结果是,68% 的人愿意去跑一趟。但是如果问题中讲计算器原价是 125 美元(衣服是 15 美元),则只有 29% 的人愿意开 20 分钟车以省下 5 美元。可见只有第二种策略是合乎被试实际心理的。

如果决策所涉及的项目有多个因素(例如买一辆自行车,需要考虑到款式、价格、性能等多种因素),人们会怎样思考?按照正规的决策理论,应该:①列出有关因素(维度、特征);②确定各个因素的权重;③将各因素的价值作加权总和;④选择加权和最大的那个项目。但是在实际生活中,人们的决策并非如此。特弗尔斯基在1972年就提出,人们可能用逐级筛选的策略来决策。例如买一样商品时,先考虑价格,把那些价格高于预定数目的品种筛选掉;再考虑款式,把那些价格合适而款式不合心意的品种去掉;……最后找到合适的品种。西蒙(1978)则提出一种更简单的“满足”策略:被试面对动态的难以预测的情况,往往不奢望获得最满意的东西,而只是设置一个可以接受的最低标准,只要某项目满足这个标准就选中它。

为了弄清人们多因素决策的主要模式,佩尼(Payne, 1976)制定了一种实验技术。这一技术已经被证明是很有用的。佩尼先制作了大量卡片,这些卡片上都是说明一些公寓的价格、房间数等特征的。但是每张卡片上只写明其中某一所公寓的某一个特征(例如公寓A有2间房)。实验时,将这些卡片背面朝上,并按照“公寓×特征”矩阵排列。这样,被试就可以很方便地找到他所需要的卡片和信息,也可以很方便地找到某一公寓的所有方面的特征,或各个公寓在某一方面的特征。被试在实验中可以自由地翻看任意卡片。通过观察,佩尼将被试策略分成两大类:一类叫作“补偿性策略”;一类叫“非补偿性策略”。前面介绍的正规的决策方式属于补偿性策略,特弗尔斯基和西蒙提出的策略则属于非补偿性策略。佩尼还发现了其他一些补偿性的和非补偿性的策略。有趣的是,被试往往结合使用上述策略。例如,先运用非补偿性策略将选择范围缩小,然后利用补偿性策略作出最后决策。佩尼还发现,任务越复杂(公寓数目增加和特征数目增加都可以使任务复杂化),非补偿性策略用得就越多。佩尼的研究资料说明,人的信息加工容量与决策策略

有很大关系。决策时,人们既要找到最高价值的项目,又要最大限度地减轻认知负荷。当这二者冲突时,就会出现折衷的做法,即先用简单的策略缩小选择范围,然后再用复杂的方法对剩下的项目进行详尽的分析。

要进行有效的决策,人们常常要判断出现某些结果的概率。于是,这一领域引来了众多的研究。其中最有影响的两位学者还是特弗尔斯基和卡耐曼。他们在研究中发现,人们的概率判断结果与数学上的概率论计算所得出的结果大相径庭。概率判断往往受到“可提取性”和“代表性”的影响。前者使人们倾向于把容易想到的(可提取性强)事物当作比较容易出现的事件;后者使人们倾向于把符合某一类事物的典型特征的个体认作该类型的成员,而忽视该类型本身的基础比率。

对于不确定性结果的判断还有一种形式,就是对两个不确定事件之间的关系的估计。例如这样一个问题:某次医学试验发现,试验结果阳性的人中,有8个患病,4个未患病;结果为阴性的人中,有5个患病,10个未患病。请问试验结果与人是否患病的关系如何?在对这种关系的估价过程中,人们往往得不出正确的结论。这一方面是由于被试缺乏统计学知识;另一方面,这一类问题的记忆负荷较重。至于在日常生活中,人们往往得不到两个事物详尽的数量记录,当然就更难估计它们之间的关系了。不过杰宁斯等人(Jennings, et al., 1982)发现一个很有意思的现象:尽管人们有时没有注意到某些不太明显的联系,但是如果这些联系有所变化(相关高一些),他们却能觉察到。也就是说,人们不善于发现相关,却对相关的变化有些敏感。

目前,对决策过程的研究出现了一个新的方向。以往的研究多数探讨那些外部条件没有任何变化的决策活动。而许多决策活动往往是分步进行的,并且面临着变化中的世界;另外,前后的决策也会相互影响。这一切迫使心理学家去构造出一种动态的决策理

论框架。

五、想象与创造

想象是思维的一种形式。但是在本世纪 60 年代以前,想象并未受到心理学界的广泛重视。这一方面是因为想象活动较难研究,另一方面也是因为当时的传统思想认为,想象和幻想都是浪费人生,是愚蠢的。到了 60 年代,这方面的研究才发展起来。

辛格和麦克瑞文(Singer and McCraver, 1961)调查了 240 个美国成人,其中 96%的人说自己每天都有某种形式的想象活动。这个研究运用了“通用想象问卷”。该问卷是一个 6 点量表,涉及到约 100 种想象。

1972 年,辛格和安特罗布斯(Singer and Antrobus)发表了一个更细致的研究报告。他们先是编制了一个细致的问卷——“想象调查表”。这是一个 5 点量表,共有 400 个项目,分为 29 组(即 29 个分量表),分别测定想象的一些特征,例如想象出现的频率、对失败的恐惧、关于性和敌意的内容、罪错想象和想象中的视觉形象等等。利用这个调查表,他们对大批美国学生进行了问卷调查,然后进行因素分析,结果找到 3 个因素,即想象的 3 种风格。第一种风格叫“负疚—烦躁型”。有这种想象风格的人喜欢严格的自我解剖,成就动机很强,但想象内容太消极。第二种风格是“焦虑—错乱型”。具有这种风格的人往往很焦虑,对自己很怀疑,思想没有头绪,想象缺乏清晰性和完整性,而且不注重成就和毅力。第三种风格是“沉思型”。这种类型的人对想象和幻想有比较积极的态度,而且表现出很高水平的沉思默想。辛格还发现,在美国,具有不同文化背景的群体在想象的频度上有一定的差异。黑人频度最高,以下是意大利人、爱尔兰人、移居英国的日耳曼人。辛格认为,这些差异可能与各群体在社会上的地位有一定关系。

还有一个很有意思的研究,是克里普柯和桑耐先(Kripke and Sonnenshein,1978)完成的。他们要求被试在一个白天的时间内,每隔一段时间就报告他们的思想。结果发现,想象活动每隔90分钟左右达到顶峰,这时 α 波增强,眼动减少。

除了运用调查法、问卷法以外,心理学家还用实验方法来研究想象。他们发现,有些因素对想象的影响很大。

辛格等人曾运用信号检测任务来研究任务要求对想象的影响。信号检测任务有一个优点,就是对任务负荷可以很方便地进行各种变换。比如说,信号出现频率的高低,判断的复杂性(仅仅判断音高还是判断前后两个音的异同)等特征都很容易操纵。有一个研究(1966)就利用上述2个因素作为自变量,探讨它们对与任务无关的思想(也就是想象)的关系。结果发现,信号出现的频率越高,与任务无关的思想的次数就越下降;判断越复杂,这种思想的次数也越下降。但各个条件下反应正确率是比较稳定的。这说明任务的要求越高,需要的信息加工容量越大,剩下的加工容量越小,与任务无关的想象活动也就越少。这个研究还有一个发现,即任务对于被试的重要性也会影响他们的想象活动,任务越重要则越无心想象。

安特罗布斯和辛格等人(1966)还发现了一个“坏消息”效应:当一个人听到某个与自己有关的坏消息的时候,想象会骤然增加,而忽视当前的任务。他们在一个研究中让被试完成一些信号检测任务。但在开始工作之前,先对实验组被试播放一段自编自录的广播新闻,说美国在越南吃了败仗,要在全中国征兵了。然后比较实验组和控制组被试在完成信号检测任务过程中与任务无关的思维——想象的多少。结果发现,实验组被试在工作过程中出现的与任务无关的思维远比控制组多,而且工作还出了许多差错。后来,实验组被试在书面报告中告诉研究人员,听到那条新闻,就想到了自己的将来:会不会被征去当兵?能不能躲到加拿大去?越南战场怎

么样?……当然,实验结束时,安特罗布斯还得告诉被试,“征兵”一事纯属主试编造。

贝克尔(Becker, et al., 1973)等人研究了应激对想象活动的影响。他们把被试分成2组,一组被试看6分钟能增强应激的紧张电影,另一组则看6分钟平淡无奇的电影,然后两组被试都去完成信号检测任务。工作过程分3段,这样被试可以在工作间隙写下前一段工作过程中想象的内容。结果发现,看过紧张电影的被试在信号检测过程中,常常不由自主地反复回想起电影中的一些场面,有些念头想摆脱也摆脱不掉。但是他们与看平淡电影的被试相比,与任务无关的思想的多少并没有显著差异。

关于想象的功能。辛格(1978)在研究了大量文献的基础上提出了想象的3大功能:①预测和计划;②提醒人去做未完成的事情;③在沉闷的环境中保持人的警醒。

对于创造过程的研究,有两大方面。一是对创造性个人的研究;一是对创造活动本身的研究。

对于创造性个人的研究,用的主要是心理测量方法,其目的是揭示那些富有创造性的个体(科学家、作家等等)所特有的心理特征。

安·罗(Ann. Roe, 1952)研究了64位美国优秀科学家,包括物理学家、生物学家和社会科学家。安·罗与每位科学家都作了长时间的谈话,并让他们作了主题统觉测验、罗夏赫测验和一个智力测验。在搜集了这些资料的基础上,安·罗总结出这些优秀科学家的一些比较共同的特征:他往往是老大;出身于中产阶级新教家庭;父亲往往是一位专业人员;他小时候经常生病,要么就是失去父母之一;他的智商非常高,很小的时候就开始如饥似渴地读书;他感到孤独,与同学格格不入;他对姑娘不太感兴趣,结婚比较晚(平均27岁);他通常是在独立完成某个课题后决定以科学作为他的事业;他工作努力,坚韧不拔,节假日也常常不休息。这些描述最适合

于物理学家。相比之下,社会科学家稍微外向些,更关心人际关系;他们中有很人有一部分人有一位在家里说了算的母亲;他们的离婚率也比较高(41%);攻击性也比其他科学家略强一些。在其他方面,社会科学家与物理学家、生物学家相差不多。

卡特尔(Cattell, et al., 1955)等人也研究过科学家的人格特征。他们要求一些物理学家、生物学家和心理学家完成 16 人格因素测验。结果发现,这些科学家与普通人相比,往往比较内向、聪明、刚强、自律、多愁善感,勇于创新。另外,德莱夫达尔(Drevdahl)和卡特尔(1958)还发现,艺术家和作家的人格特征与科学家差不多,只是艺术家比科学家更多愁善感些,内心世界更紧张些。

另外还有些研究发现,科学家的思维倾向于收敛思维,而收敛思维测验成绩比较好的男孩子也倾向于成为科学家。这一点与我们平时的看法似乎不同,很多人认为发散思维是创造性思维的主要成分。

关于创造思维活动本身的研究,华莱士的四阶段论大概是最有名的了。这个理论是 1926 年提出来的,但当时没有什么实验论证,直到约 10 年后,帕特里克(Patrick, 1935, 1937)才为其提供了一些实验资料。但是对孕育期还是没有多少研究。后来穆瑞和丹尼(Murray and Denny, 1969)针对孕育期做了一个实验研究。他们先根据解决问题能力的强弱把被试分成两组,又把这两组再分为控制组和实验组,然后让被试解一道很复杂的题目。控制组要求在连续的 20 分钟内解决问题,而实验组要先做 5 分钟,然后干 5 分钟别的事,最后再用 15 分钟研究问题。也就是说,实验组有 5 分钟时间可以作为孕育期。从实验结果可以看出,孕育期的有无以及能力的强弱对这个问题的解决都没有显著影响。但这里有一个显著的相互作用,即能力较强的人受孕育机会的阻碍,而能力弱的人则受到这种机会的促进。表 2 中,“↓”表示受到阻碍,“↑”表示受到促进。这可能是因为能力弱的人准备期短,较早进入孕育期,5

分钟其他工作正好作为孕育期,从而提高了成绩。而能力强的人准备期长,5 分钟的孕育机会反而打乱了他的思路,成绩并没有什么提高。

表 2 孕育期的作用

	控制组	实验组
能力强	—	↓
能力弱	—	↑

当代心理学界对创造思维的理论研究也有丰富成果,这里仅谈一谈梅尼克(Mednick,1962)、柯斯勒(Kocstler,1964)、格鲁伯(Gruber,1980)和西蒙(1966)的理论。

梅尼克给创造性思维过程下了一个定义,说它是“联想的成分产生新的组合,这种组合满足一定的需要,具有一定的作用”。这个定义和创造过程的产品两个特征——新颖性和价值性十分吻合。他认为,任何两个概念之间都或多或少有些联系,平时联系得很紧密的两个概念(如“钟”和“表”)组合在一起没什么新鲜感,算不上创造性思维;如果平时联系很弱(如“软”和“表”),组合起来就新颖了。梅尼克还提出了产生新颖组合的一些方法。

柯斯勒提出的理论则更进一步。他认为,创造性思维是将两个先前没有什么联系的观点或理论联系起来。例如,阿基米德测量王冠体积的时候,就是把洗澡时出现什么现象(理论 1)和体积测量理论(理论 2)联系起来,从而解决了问题。

格鲁伯根据皮亚杰的理论分析了达尔文生物进化论的发展和成熟的过程。他认为,顿悟只是理论发展的长期过程的一个部分;如果过分重视顿悟,会把研究者的注意力从理论的长期发展上面引开。格鲁伯认为,达尔文理论的发展过程(也是创造性思维过

程)可以用皮亚杰的“同化”和“顺应”概念来解释。对于能够用原先的理论来解释的材料,同化之;对于不能解释的材料,顺应之。

西蒙认为,科学发现是一个非常缓慢的过程。要创造性地解决一个疑难问题,科学家要进行无数次的探索 and 实验。本来,人的长时记忆容量几乎是无限的,但长时记忆中的知识经验要进入短时记忆才能加入思维,而人的短时记忆容量太小了,所以要找到一个答案很不容易。这样,毅力和热情就成为有成就的科学家的“关键特征”。

那么孕育期和灵感期是怎么一回事呢?西蒙认为可以用“熟悉化”和“选择性遗忘”来解释。熟悉化指的是一个缓慢的、在长时记忆中建立问题表征和相关资料的过程。西蒙认为,搜集信息受短时记忆中目标系统的影响,信息进入长时记忆后,又反过来影响目标系统。这样,短时记忆中的目标信息和长时记忆中的问题信息就不断地相互作用着。当问题被搁在一边时,目标信息就遗忘了;当以后再研究这个问题时,短时记忆中就要组建一个新的目标系统,而这个新的目标系统就可能导致问题的迅速解决。

最后谈一谈思维心理学研究的趋势问题。

总的思维心理模型是怎样的?对这个问题,30年来已提出过多种模型[从布罗德本特(1958)到安德森(1983)]。其中某些论点得到了心理学界的广泛赞同。这些论点主要是:①思维是工作记忆内以及长时记忆与工作记忆之间发生的对符号的操纵;②对符号的操纵受长时记忆中储存着的规则制约,但从长时记忆中提取什么规则取决于当时工作记忆中的目标信息和内容信息;③在同一时间只能进行一步思维(系列加工)。英国心理学家吉路利(Gilhooly)于80年代将上述论点加以综合,使之成为一个理论模型,并借用记忆理论中的术语,称之为“模态模型”。从本文介绍的一些研究中可以看到,这个模型可以解释相当多的问题或现象。但是,近来有一些学派,例如联结主义学派所提出的模型中,符号操

纵观点和系列加工观点都被抛弃了,可以说完全否定了模态模型。还有一些中间派别,它们保留了符号操纵观点,而把系列加工观点抛弃了。

模态模型认为思维是对符号的操纵,这一点受到西尔勒(Searle, 1984)的批评。西尔勒认为,思想是有意义的,而意义不仅仅由符号来代表,思维也不能用简单的符号操纵来描述。后来,吉路利本人(1990)也接受了这个看法,把“操纵符号”改成了“操纵心理模型”。

模态模型并不考虑长时记忆中的信息加工活动,而许多学者认为在长时记忆中是进行着无意识的信息加工的,否则就不能解释创造性思维中的孕育期,也不能解释我们有时体验到的那种无法摆脱某一念头的情形。而如果承认长时记忆中也有信息加工,那么思维也许就是一种并行处理信息的活动了。可以预料,对于并行加工的研究将成为一种趋势;相应的理论模型也将越来越多。

思维研究的另一趋势,将是越来越重视研究人们对自然问题(相对于早期实验室研究中的人工问题、人工概念)的解决。人工问题对控制被试的先行经验等起到了很大的作用,但在研究人的先行经验对解决问题的影响方面就显得不够了。对自然问题解决的研究将成为对人工问题解决的研究的一个重要补充,两者相辅相成。

(邵志芳参加撰稿工作)

语 言 理 解

彭 聃 龄

(北京师范大学心理学系)

语言是人区别于其他动物的重要标志之一。人因为有语言,才能够彼此交际,交流思想,发展人际关系,传播人类文化,进行以语言为中介的思维活动,用语言计划行为,调节与控制行为。这一切都使人的行为和心理活动产生了一系列重要的特点。语言在人的行为和心理活动中的特殊作用,使它成为现代心理学的一个非常重要的研究领域。1921年,E. Esper 在“语言心理学”一文(见《Psychological Bulletin》)中,评述了1917—1921年间在语言领域内的心理学研究,当时引用的文献仅17篇。而在半个世纪以后,S. Fillenkaum (1971)在心理学年鉴上发表了“心理语言学”一文,评述乔姆斯基对语言心理学研究的巨大影响,其中引用的文摘就有1600篇。乔姆斯基曾说过:“语言研究……为人类心理过程的研究提供了一个非常有利的视角”(《语言与心理》1967)。可以毫不夸张地说,现代心理学关于人的心理的各方面的研究,都直接、间接地与研究语言有关。因此,只有了解语言的性质及其在人的行为和心理活动中的重要作用,才能真正理解和揭示人的心理的特点。本文只是从语言理解的角度,介绍有关语言的一些心理学研究。

一、语言理解的一般问题

1. 什么是语言理解

语言理解就是揭露语言的意义与作用。这是人脑对语言的一种主动、积极的加工过程。Moates(1980)指出,“理解是一个建构过程。在这个过程中,听话者使用对句子结构的知识、词汇的知识,对世界的知识以及对说话者怎样将他的愿望编码为主题结构的假定。听话者常常假定说话者试图提供某种意义,并根据说话者说出的话语构成一种可能的意义”。Caroll(1986)提出:语言理解是“从语言的表面结构提取深层的命题结构的一种积极的、推理的过程”。Taylor(1990)提出:“话语理解是指知觉和整合包含在话语中的各种信息(语言的或应用的信息;显明的或隐含的信息),从而产生在话语中所描述的事件的连贯而正确的图景”。Kintsch(1990)也指出,在语言理解中,“理解者的任务是建构由课文提供的,并且与读者的知识、信念、目标整合在一起的信息的一种心理表征”。

根据以上主张,我们认为,语言理解就是听者或读者借助于听觉的或视觉的语言材料,在头脑中建构意义的一种主动、积极的过程。它的特点主要有:①语言理解是一个建构意义的过程。所谓建构意义,就是从语言的表面结构提取深层的命题结构;②理解依赖于从外部输入的语言材料的性质和听者或读者头脑中已有的知识。建构意义的过程就是听者或读者运用已有的知识加工由外部输入的言语信息的过程;③在语言理解中,推理有重要的作用。推理可以补充语言信息的空缺,并预期将要出现的信息。

2. 语言理解的心理成分

语言理解包含一系列加工阶段,由一系列成分所组成。

Foss(1978)提出:句子的意义依赖于组成句子的词的意义。因此,理解过程的第一个成分是从句子中单词的语音表征获得词的意义表征。我们把它设想为词汇查寻(lexical lookup)的过程。……句子的意义还依赖于它的句法结构,因此,理解过程的另一成分就是从句子的语音和词汇表征来推论或建构出句子的句法结构。最后,就是将词汇和句法信息整合为句子的意义表征。

Anderson(1980)把语言理解分成3个阶段,即知觉(perception)、解析(parsing)和使用。知觉是对声音和书面信息的初步编码;解析是把信息转换成词的组合意义的心理表征;使用是对心理表征的实际应用。Anderson特别重视解析的作用。通过解析,人们将句子分解成短语或构成要素,并解释每一要素的意义,从而把词语表征转化为意义表征。

Granham(1985)把语言理解系统区分为以下几个子过程:①低水平的知觉加工;②单词识别;③解析,即从语法上分析词或句子;④语义解释;⑤模型建构(model construction),即建立某一特定情境的某种内部表征,或该情境的心理模型。

Taylor 和 Taylor(1990)列举了参与话语理解的各种心理成分:①知觉言语的声音;②认识单词;③理解每个句子;④把每个句子与前面的句子联系起来;⑤确认一篇说明文的题目;⑥确认一篇描述文中人物的目标;⑦确认每句话中第一个重复使用的单词的前行词(首语重复);⑧把重要的思想与不重要的思想分开,对重要的思想进行更多的加工;⑨作出推理与结论;⑩提取课文的主旨(gist)或故事的寓意(moral)。

总之,语言理解是一个过程,它经历一定的阶段,每个阶段又有相应的操作或处理。语言理解可以划分为3个水平:①词汇理解或词汇识别,即通过词形和词音达到单词的意义,或者说在语义网络中激活了与词对应的某个概念结点或命题;②句子理解,从句子的表面结构建构句子的深层结构;③话语或课文理解,建构句子之

间的联系,揭示课文或话语的宏观命题。

3. 口语与阅读理解

语言理解包括口语理解和阅读理解两方面。两者在加工的早期阶段是不同的。口语理解依赖于短暂存在的听觉意象和听觉的单词识别,而阅读依赖于稳定的视觉表征和视觉单词识别。但在句子和话语水平上,听觉理解和阅读理解是相似的。当单词识别技能得到发展后,听觉技能与阅读技能是可以互相预测的。研究表明,一到四年级学生两者的相关是不断发展的,以后到高中,稳定在0.6的水平上。(Sticht, Beck, Hanke, Kleiman 和 James, 1974)。对大学生进行的大样本的研究表明,阅读理解和口语理解的相关为0.80,因此,阅读理解能力与听觉理解能力是没有严格区别的(Palmer 等 1985)。

口语和阅读理解的相关很高,但不完全一样。在口语理解时,听者需要加工他所听到的言语,这受到说话人言语速度的限制。而在阅读理解时,读者可以按照自己的速度加工书面材料。Walker (1971)发现,用相同测验材料,读者的理解分数高于听者。阅读倾向于理解课文中明显的字面信息,而口语理解倾向于理解故事的要旨(Hildyard 和 Olson, 1978)。

二、词的识别

1. 词与词的识别

(1) 什么是词与词的识别

词是语言的建筑单位,是语言中“最小的自由形式”(布龙菲尔德《语言论》)或“最小的自由语”(董毓同《语言学大纲》)。高名凯教授(1965)也曾指出,词是“组成词汇系统的各个成员”,“一个词之

所以能成为一个词汇单位,必须具备 3 个特点:①在语音上具有某种统一性;②在语义上具有某种统一性;③在语法上具有某种统一性”。

我们不在这里讨论语言学家对词的不同定义,但在这些定义中有一点是共同的,即语言是由词构成的。在日常口语交际和书面语言中,人们自由使用的语言单位都是词。在口语中,词是音与义的结合体。它不仅传递声音与意义的信息,还传递构词法与句法的信息。而在书面语言中,词还具有图形的信息。因此,美国心理学家 J. E. Gibson 曾指出:词是特征的复合,即 5 种信息的复合物,包括图形信息、语音信息、构词法信息、语义信息和句法信息。因此,在书面语言中,词常常也称为形、音、义的统一体。

词的作用是标志事物或对象,或传递事物的信息。如果说,词形和词音是词存在的客观形式,那么,词义就是词所指称的事物或现象。我们听到“shū”或看到“书”,便知道它代表了一种记载人类文化成果的东西。人们通过听觉或视觉,接受输入的词形或词音信息,并在人脑中揭示出它所指称的事物或对象,即词义,这个过程就是词的识别过程。

词的识别不仅依赖于词音、词形的感觉输入,而且依赖于人脑中已有的关于词的各种知识。这些知识能对输入的感觉信息进行解释和预测。如果我们把人脑中的词汇比喻为一部词典(内部词典或心理词典),那么,词的识别也可以说是在内部词典或心理词典中进行查询和搜索的过程。当人们根据词形或词音在心理词典中找到了相应的词条,并获得对它的解释时,这个词也就被识别了。近年来,在国外阅读心理学和心理语言学著作中,流行着一个新的概念:词汇通达(lexical access),它的意思就是人们根据词形或词音在心理词典中找到了相应的条目和理解了它的意义。在这个意义上,词汇通达和词的识别有时可以互换使用。

(2) 单词识别的阶段

单词识别是一个过程,它需要经历一定的阶段,消耗一定的时间,并完成一定的心理操作。由于词的类型不同、所处的背景不同、读者的阅读熟练程度不同等,单词识别的时间长短是不一样的。一般来说,在阅读时识别一个单词的时间大约需要 70—200 毫秒 (Jackson 和 McClelland, 1975)。

J. I. Chumbley 和 D. A. Balota (1984) 把单词识别分为两个阶段:①词汇通达阶段,这是指将词的视觉特征与心理词典中的一个项目进行匹配,从而达到一个词汇项目,或者说在心理词典中找到了一个词的位置;②意义决策阶段 (meaning determination), 包括确定意义、发音和与单词有关的其他信息。只有在词汇通达之后,意义的确定才可以进行。

M. S. Seidenbeng (1985) 把单词识别分为 3 个阶段或 3 个过程:①词汇前加工 (prelexical processes), 分析书面语言或口语的输入,把它确定为一个特定的词;②词汇通达,激活与词典项目有联系的语义信息、语音信息、正字法信息和其他信息;③词汇后加工 (postlexical processes), 选择、精细加工和整合词汇的信息,以便理解句子和话语。在这种观点中,第一和第二阶段是分别与词汇通达阶段和决策阶段相对应的。

(3) 整词识别与特征分析

① 单词识别中的两种加工:在研究单词识别时,一直存在着两条不同的路线。一些人认为:单词识别是一种自下而上的加工,即从最小的、最低的单元达到最大的、最高的单元。根据这种看法,单词识别开始于对感觉资料的分析。这是一些细小的单元,由这些单元再联合成字母特征、字母和单词。另一些人认为:单词识别是一种自上而下的加工,它从目标词开始,这些词受到语境和读者知识的影响,然后检验输入的感觉资料是否与目标一致。前者强调单词的特征分析,后者强调整词的识别。现在越来越多的研究表明,词的识别是自下而上的加工与自上而下的加工交互作用的结果。

② 单词识别中特征分析的证据：在单词识别中存在着特征分析的阶段，这一看法得到了许多实验事实的支持。早期的实验如 Neisser(1963)的研究，Rumelhart(1971)的实验。在最后一实验中，Rumelhart 使用了一些特别准备的字母符号，如 A, B, C, D 等，每个符号都可以用 7 种不同的线段进行分析，如 $\neg$, $\subset$, 一 等。用速示器单个呈现这些符号，让被试设法加以确认。结果显示：符号的混淆非常有规律。一般来说，将一个符号误认为另一个符号的概率，依赖于被识别的符号包含着哪种直线线段。例如，D 常常与 $\neg$ 混淆。这说明人类可能存在特征觉察器一类的东西。

近年来，Just 和 Carpenter(1987)用眼动记录技术记录阅读，发现读者对课文中单词的注视时间随词的长度而上升。这一结果对高、中、低频率的词都一样。这说明长词比短词携带着更多的视觉信息，因此读者需要花费更多的时间来获得这些信息，并将它与一些相同长度的词区分开来。

③ 词优效应(word superiority effect, 简称 WSE): 1886 年美国著名心理学家 J. M. Cattell 最先发现了词优效应：在速示条件下，人们只能报告 3, 4 个无关的字母，但能报告两个在语义和句法上没有联系的单词。他的研究在方法上有两个缺陷：首先，他使用全部报告法，即在短暂呈现刺激后，要求被试报告(回忆出)所有看到的東西。在这种情况下，被试在感觉记忆中保存的信息可能比能够报告的信息多些。其次，没有排除猜测的作用。如果呈现的刺激是词，如 read，那么被试根据对词汇的知识和接受到的部分刺激，就可作出一定的猜测，而非词和假词没有猜测的基础。在这种情况下，知觉因素与反应因素是混在一起的。

这种方法学上的缺陷大约经过了 80 年才得到克服。1969 年，Reicher 的实验重新引起了人们对 WSE 的兴趣。Reicher 的实验控制了被试的反应因素，用速示器呈现刺激，包括词(word)，非词(owrd)和字母(d 或 k)。在刺激消失后，立即对某个字母进行再认

测验,如报告最后一个位置上的字母是 k 还是 d。结果发现,嵌在单词中的字母比嵌在非单词中的字母或单独呈现的字母,再认的正确率要高些。由于目标字母(d 或 k)都能与前 3 个字母组成一个单词(word 或 work),因此排除了被试猜测的作用。

继 Reicher 之后,许多心理学家研究了这个问题。例如 Baron 和 Thurston(1973)考察了正字法规则和语音学规则在词优效应中的作用,发现那些符合语音的拼写模式的假词(如 pake)和真词一样存在词优效应。Baron(1974)发现:化学家对化学公式存在词优效应;Henderson(1974)发现:有意义的缩写(如 FBI 美国联邦调查局)比同样长度的无意义字母串,在比较时存在优势。Aderman 和 Smith(1971)发现了期待在单词知觉中的作用。在出现词和非词的情况下,当被试期待出现词时,有 WSE;当期待出现非词时,没有 WSE。

对 WSE 的解释是各种各样的。我们在这里不作详细的介绍。但是,这一发现对特征分析模型提出了挑战。如果我们能更有效地加工在单词中的字母,这就意味着对词的加工有助于对字母的识别,而不是对字母识别有助于对词的识别。如此看来,对单词的识别就不是从字母到单词了。有人认为(Baron,1978),WSE 反映了我们的某种知觉策略,这种策略是以语言学知识为基础的。在一种情况下,我们可能按层次进行单词识别,即先确认特征,再识别字母和单词;而在另一种情况下,对单词特征的知识能帮助我们吧视觉特征整合成单词,而不是先确认它的个别字母。

(4) 研究单词识别的方法

研究单词识别的方法很多,如阈限法、反应时记录法、生理学方法等(K. Rayner,1989)。这里只介绍目前常用的几种方法。

① 词汇判断作业(lexical decision task,简称 LDT):在这种作业中,给被试呈现一串字母(如 DUT 或 DOG),有的是词,有的是假词或非词(ODT),要求他们尽快、尽量准确地决定它是不是

一个词。是词,按一个键;不是词,按另一个键。记录反应时和错误率,作为两个基本变量,以测定不同因素对单词识别的影响。这种作业要求被试加工词的视觉信息、语音信息,并且在心理词典中进行搜索,以确定输入的刺激是不是一个词,因而对探索词汇通达的问题有重要的作用。在研究词的视觉特征、语音特征和词频对单词识别的影响时,经常使用这种方法。但是,词汇判断作业不同于正常阅读,而且不能确定被试是否理解词义,因而有其局限性。

② 命名作业(naming task):即大声读出一个单词的声音,或命名一个单词,并记录命名的潜伏期(反应潜伏期)。这种作业也叫命名潜伏期作业(naming latency task,简称NL)。一般认为,被试在命名时,已经达到了词汇表征,并符合正常阅读的情况,因此在研究词频与语境的作用时,经常使用这种作业(包括研究重复效应)。但有人认为,命名发生在词汇通达之后,因此,应用这种方法不利于研究早期加工阶段的特点,特别是语音在单词识别中的作用。

③ 语义分类作业(semantic categorization task):这种作业要求被试作语义判断,即判断一个词是不是先前规定的语义范畴的一个例子,是作“Yes”反应,不是作“No”反应(按不同的键)。例如刺激词为语义范畴词“花”,目标词为“玫瑰”或“台灯”,对“玫瑰”作“Yes”反应,对“台灯”作“No”反应。这种作业要求被试经过词汇通达的全部过程,因而对研究后期的加工更有利。

2. 单词识别研究的类型

根据单词研究中使用的不同变量,我们可以把研究工作分为3类(Foss,1988):

(1) 经验性研究

经验性变量是指那些由语言经验得来的,对词汇项目的句法和语义描述不起作用的变量,如频度、熟悉性、新近性等。

研究词频对单词识别的作用,开始于本世纪 50 年代。1951 年 Howes 和 Solomon 发现,词频是影响单词再认阈限的一个重要因素。以后,Postman 和 Solomon(1952)通过研究,把词频的作用解释为新近性的影响。

1969 年 Foss 在一项著名的实验中,证明了频度在词汇通达中的作用。实验采用音素监测的方法。给被试呈现两个内容相同的句子,要求被试监测出其中的一个音素,如 |b| (目标音素)。在一个句子中,目标音素出现在高频词 travelling 之后;而在另一个句子中,目标音素出现在低频词 itinerate 的后面。结果,在低频词后的音素监测时间长于在高频词后的音素监测时间。这是因为对低频词的检索较困难,需要投入较多的能量,因而影响到对随后出现的音素的检测。

【附】 Foss 的实验句举例:

The travelling Basson player found himself without funds in a strange town.

The itirerant basson player found himself without funds in a strange town.

以后,Rubenstein, Garfield 和 Milliken(1970)用 LDT 发现了类似的结果,即高频词的反应时比低频词短些。O'Connor 和 Forster(1981)用 LDT 作业研究了对非词的识别。发现对应于高频词的非词(拼错)比对应于低频词的非词,错误率要高。例如 MOHTER 是对应于高频词“mother”的非词,它可能引起高频词“mother”的激活,因而比 BOHTER 能引起更多的“是词”的反应。1983 年 Gordon 用两种方式呈现高频词与低频词,即分组呈现和混合呈现,比较它们对 LDT 的影响。结果发现,高频词在分组条件下比在混合条件下反应时快些;而低频词在两种条件下反应时没有区别。Gordon 认为:这是由于在高频词分组呈现时,被试可能降低了他们的判断阈限。

用命名作业也获得了一些有意义的结果。Geffen 和 Luszez (1983)发现,词频能影响命名前的停顿时间,即读低频词比读高频词前有更长的停顿时间,而对说出单词本身的持续时间没有影响。Rayner 和 Duffy(1986)也发现,在控制词的长度的情况下,被试阅读不常用词比阅读常用词,需要更长的注视时间。

近年来,心理学家们在研究词频和熟悉性的作用时,还发现了词的重复效应或重复启动效应(repetition priming effect):前面出现过的词对随后出现的相同单词有促进作用,即加快对后继单词的识别速度。1982年 E. Tulving 在定义启动效应时指出:“在后继过程中,如果遇到同一刺激或与之相关的刺激,则后继刺激将产生促进作用。其中同一刺激的启动称为直接启动,相关刺激的启动称为间接启动”。这里的直接启动就是重复启动。

从 70 年代以来,重复效应得到了广泛的研究。已有的发现可概括如下:

① 重复效应在各种作业中是普遍存在的,如 LDT、NL、再认、回忆、同义词搜索和语义搜索作业、字谜作业、词汇完成作业等。

② 词频和重复效应存在交互作用。重复对低频词的促进作用更大(Scarborough,1977)。

③ 词的重复效应可持续较长的时间:2 天(Scarborough)或 1 年(Kolers,1976)。

④ 非词也存在重复效应,但随着间隔时间的延长而迅速消失(Scarborough,1977)。

⑤ 重复频率影响重复效应的大小。

⑥ 刺激的具体特征(如大、小写)对重复效应可能有影响(Jacoby,1981)。

⑦ 当启动刺激未被意识到时,仍有重复效应,说明重复效应不是有意识的注意加工(Evett,1981)。

⑧ 向后掩蔽可能影响重复效应。

(2) 结构性研究

影响单词识别的结构性变量,是指词的结构成分(morphological constitution)和词汇的复杂性(lexical complexity),包括语义和句法的复杂性。

① 词汇通达中的词素结构(morphology),词由词素构成。词素可以定义为语言中最小的意义单位。例如 redistribution 一词,由 3 个词素(distribute, -ion 和 re-)组成。其中有些词素能独立存在,如 distribute,叫自由词素(free morphemes);有些词素不能独立存在,它必须依附于自由词素,如 -ion, re-,叫粘着词素(bound morphemes)。从另一个角度说,自由词素就是词根(root word),而粘着词素就是词缀。其中词根是携带主要词汇信息的词素,如 come,而词缀是加在词根或词基上(base),以构成新词的词素,如 comes, coming 中的“s”,“ing”等。在不同单词中,词缀是按相同语义反复使用的。

由于英语构词法的这一特点,使研究者产生了一种设想:词素信息和词根信息在心理词典中可能是分别存储和组织起来的(Mackay, 1978; Taft 和 Forster, 1975; Taft, 1981; Smith 等 1982)。根据这种看法,一个词如 decision 可以分解成词基-decide 和另一词素-ion。这两部分分别存储在心理词典中,提取时才将它们合在一起。这样存储符合认知经济原则,但增加了加工的难度或复杂性。

② 词汇分解(lexical decomposition);Mackay(1978)证明了词基和词缀是独立存储的。他用听觉方式给被试呈现一些动词,如 decide,让被试尽快说出有关的名词,如 decision。结果发现,被试作出反应所需的时间是随词的派生复杂性(derivational complexity)而变化的。从语言学上讲,后缀-ment 比-ence 简单,-ence 又比-ion 简单。这是因为带有后缀-ion 的词在由动词变名词时,元音要发生变化。如由 decide 变成 decision 时,元音[i]要由[ai]变成

[i]。-ence 又比 -ment 复杂,因为在由动词变成名词时,带有 -ence 的词,要出现音节的重新组成,如由 exist 变成 existence,其中辅音 t 在动词中属前一音节,而在名词中属后一音节;而带有 -ment 的词,在由动词变成名词时,音节的组合没有变化。由于词在语言学上的复杂性不同,因此命名这些单词的时间有明显差别。词在语言学上的复杂性越高,反应时就越长。

Taft 和 Forster(1975)使用 LDT 进行实验,发现将前缀与粘着词素(vade)联合起来构成的假词(如 revade)要比用前缀和非词(poch)合成的假词,词汇判断的时间要长些。这说明在心理词典中存在着粘着词素,如 vade,要确定前缀与粘着词素是否相容,是需要时间的。Taft 认为,一个由单一词素构成的词,可直接通达;而由不同词素组成的词,则需要先分解为它的词素成分,然后再到达它的词基;对带前缀的词,则需进行前缀脱落(prefix-stripping),然后再进行对基本词的搜索。如果搜索成功,就将前缀与基本词比较,看两者是否相容。因此,对词的识别是一个多阶段的搜索过程。近年来,合成词(compound word)的分解存储问题也引起了心理学家的极大关注(Andrews 1986, Sandra 1990)。

③ 韵律结构(rhythmic structure):Cutler 和 Clifton(1984)研究了韵律结构对词汇通达的影响。一些双音节的名词和动词具有不同的重音模式。名词是扬抑格和长短格,如 contrast;而动词是抑扬格和短长格,如 con-TRACT。研究发现,重音影响到词汇通达的速度,但对识别一些孤立的单词没有明显的影响。另一些研究者(Diehl, 1987; Grosjean 和 Gee, 1987)提出:发重音的音节是通向心理词典的一些重要的地点。

④ 词汇的两歧性(lexical ambiguity):歧义性是语言中普遍存在的特性。它是影响词汇通达的另一个结构变量。

Foss(1970)首先用音素监测技术研究了词汇的歧义性。发现当目标音素出现在歧义词后时,反应时较长,而当目标音素出现在

非歧义词后时,音素监测的反应时较短。以后,Cairns 和 Kammerman(1976);Swinney(1979),Bever,Garrett 和 Hurtig(1973)都研究了词的歧义问题,发现词的多重意义只能短暂保留;加工时间稍长,歧义性就将消失。

(3) 语境的研究

影响词汇通达的语境变量包括出现在单词前后的词、句子与课文。它们分别构成词语境、句子语境和课文语境。

最早研究句子语境对单词识别的影响,是 Tulving, Mandler 和 Bauml(1964)的一项实验。实验中先呈现语境句,接着呈现目标词。语境句如 Her closest relative was appointed as her legal ____, 目标词为 guardian。控制两个条件:①语境句中词的数量:8, 4, 2, 0;②目标词的呈现时间:0, 20, 40, 60, 80, 100, 120 或 140 毫秒。结果是,语境长度和目标词的持续时间对作业成绩都有影响。为了获得 50% 的正确率,当语境为 8 个单词时,目标词的呈现时间只需 40 毫秒。相反,没有语境时,为了获得同样的再认率,目标词的持续时间需 100 毫秒。

Ehrlich 和 Rayner(1981)用眼动记录技术发现,眼动是语境对个别单词限制程度的函数。他们编写了一些课文,这些课文对某一单词的限制程度分为高、低两种。在高限制语境下,被试容易预期某个单词,而在低限制程度下,被试难以作出预期。结果发现:①限制程度高的单词比限制程度低的单词,读者的注视时间较短;②当被试注视限制程度高的单词时,注视的持续时间较短;③高限制课文比低限制课文,读者发现单词拼写错误的可能性较低,说明当词可以较好预测时,被试对词的视觉特征不敏感。

语境的研究获得了一些重要的成果:

① 语境有促进和抑制两种作用。语境适当,有促进作用,使到达单词的时间缩短;语境不适当(语境与目标不一致),使到达单词的时间延长(Stanovich and West, 1983)。

② 词频与语境有交互作用,低频词从语境有效性中获得更大的好处(Stanovich,1981)。

③ 语境有助于迅速消除词的歧义。在语境条件下,人们倾向于用与语境一致的具体方式来解释一般性的名词,这个过程叫示例作用(instantiation),即用具体例子说明问题(Anderson and Ortony,1975;Bransford and McCarrell,1974)。

3. 单词识别的理论

(1) 单词产生器(logogen)模型

英国心理学家 Morton 于 1969 年提出了单词产生器(logogen)模型(见图 1)。这个模型的特点有:

① logogen 是一种内部词典装置,它像一种加法器一样,负责计算信息的总量。

② logogen 的信息输入有视觉的、听觉的即词音和词形的特征信息,也有由语境和上下文提供的语义信息。如看到、听到“书”(book),联想到笔(pen)等。这些信息都影响到 logogen 的计数值。

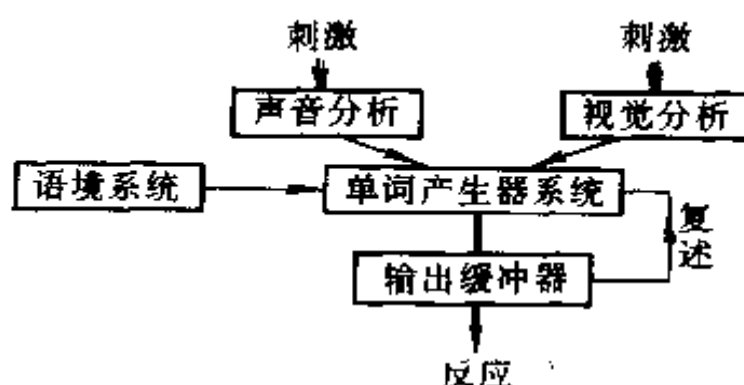


图 1 单词产生器模型流程图(1969)

词的物理特征和语义特征,是 2 个独立的信息来源。当输入某一刺激,或在语境的影响下,logogen 的计数值都会增加,并逐渐累积下来。当累积的信息超过了标准的阈限,便通过输出缓冲器作出一

个实际的反应。如果积累的信息没有达到阈限，它便经过复述而在系统内循环。图 2 说明了在各种条件下的单词产生器状态。

③ 词频的作用是影响阈限值的大小或判断的标准。高频词阈限低，容易得到激活；低频词阈限高、激活较困难。

④ 词频与语境的作用是相加的(additive)，而不是交互作用的，它们可能作用于信息加工的不同阶段上。

⑤ 单词产生器的活动具有持续性，一旦进入活动状态，就能持续一段时间。

以后，Morton 多次修改了这一模型(1982, 1984)。主要修改之处有：把单一的词产生器系统分为输入产生器和输出产生器，有些因素只影响输出系统，对输入没有影响；把单一的输入产生器分成听觉输入产生器和视觉输入产生器，这是因为视觉与听觉有时不存在跨通道的影响；把语境的影响归属到含义更广的“认知系统”中；在“认知系统”与听、视觉输入产生器之间存在着交互作用(图 3)。

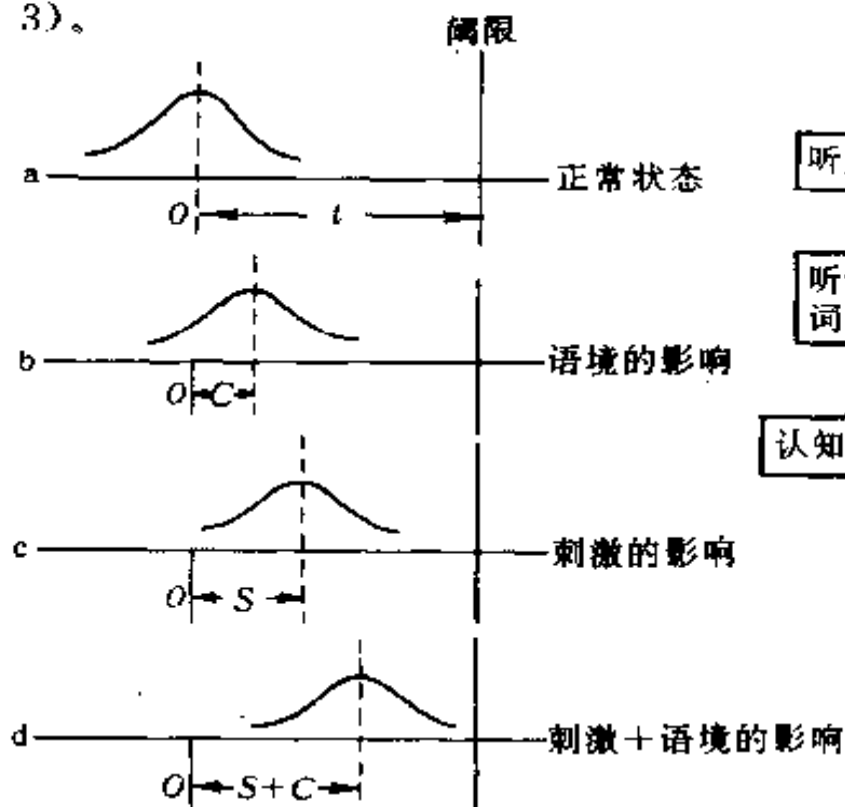


图 2 各种条件下的单词产生器的状态

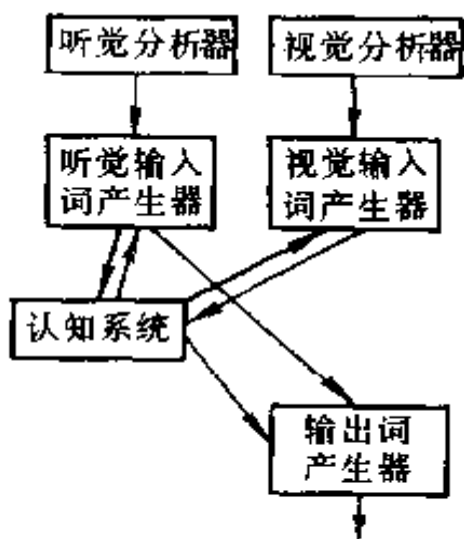


图 3 Morton 模型(1984)

Morton 模型是迄今为止在单词识别研究中影响最大的一个模型。它解释了语境的作用(Tulving 的实验),即有关语境增加了 logogen 的计数值,因而提高了它的平均激活水平;也解释了重复效应:前面出现的词降低了相应词条的激活阈限,因而对重复词有促进效应。

但是,模型也有局限性,它受到了后人的一些批评。例如,后人发现,词频与语境间存在交互作用,这是与 Morton 模型的预期相反的;又如,实验发现词的物理特征(如大、小写)有时对重复效应有作用,这用 Morton 模型难以进行解释;它对句子语境的作用也难以说明。

(2) 搜索模型(search model of lexical access)

词汇通达的搜索模型是由 K. L. Forster 及其同事(Taft 和 Forster, 1975, 1976; Forster, 1986)在 70 年代中期提出的。该模型认为,词汇是以系列方式进行扫描的。经过扫描最后在输入的感觉信息和一个词条(lexical entry)间找到一种适当的匹配。

词的感觉输入有两种基本形式,即视觉输入和听觉输入。为了解释感觉输入和词汇条目是怎样匹配的,搜索模型假定了两类词汇表征(lexical representation)。一类和词的正字法信息(如词的视觉形式)一致;另一类与语音信息(如词的听觉形式)一致。这两类表征分别叫作正字法通达文件(orthographic access file)和语音通达文件(phonological access file)。它们同时进入词典本身(lexicon proper)或“主文件”(master file)。在这里,一个词的全部信息(如意义、发音、拼写等)都得到利用。例如,当一个字母串以视觉方式呈现时,正字法通达文件得到序列的搜索。这时,如果在一个词汇表征和经过编码的感觉输入间找到适当的匹配,正字法通达文件就会直接到达相应的主文件项目。然后,再用最初的刺激(或这种刺激在工作记忆中的表征)来检验这个单词的全部拼法。如果两者匹配,就把这个字母串认成一个词;如果不匹配,就要在正字法

通达文件中重新进行搜索,以便找到适当的匹配。模型的框架见图4。

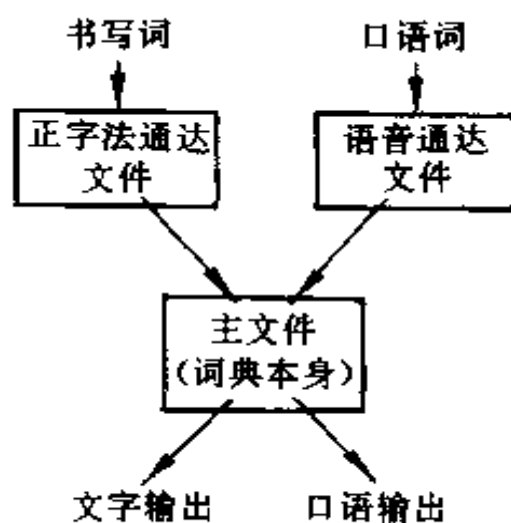


图4 词汇通达搜索模型示意图

Forster 提出,由于词典的容量很大,人们不可能每次对所有已知的单词进行搜索,因此每次搜索只局限于在一个特定的词条子集中进行。这种词条子集叫“bin”(箱子)。“bin”由感觉特征来定义,并且是在搜索程序开始前就被选定的。

Forster 强调,单词在“bin”中是依频度来存储的。因此,从心理词典中进行搜索也按词频的顺序来进行。高频词在低频词之前与输入的信息进行比较,因而在低频词之前受到通达,或者,对高频词的反应比对低频词的反应更快。

用搜索模型还可以解释词的重复效应。当一个单词经过一次通达后,它移到了搜索集的顶端,因而在第二次呈现时,就会很快到达它。由于低频词在搜索空间中的位置会因重复而得到更大的改善,因此,低频词的重复效应比高频词要大些。这一预测已经为实验所证实。

(3) 核证模型(verification model)

核证模型是由 Becker(1976)提出的。模型假定,单词识别依

赖于感觉信息存储和词典信息存储,它由特征提取和核证两个阶段组成。当一个词出现在感觉记忆中时,构成它的形状的一系列特征会引起视觉特征觉察器或视觉特征分析器的活动。然后,由特征分析器的输出进入词汇记忆(lexical memory),后者是长时记忆的一部分。在词汇记忆中,词是用一组特征(如线条,曲线等)来表征的,其中每个特征都有一个计数装置。当特征分析器觉察到单词中的一个特征时,在词汇记忆中具有这个特征的所有单词的特征计数值都上升。当计数值超过某个单词所需要的标准时,这个词就会成为一个潜在的知觉单元(percept)。核证模型认为,在同一时刻可能出现多个供选择的知觉单元。因此,加工系统必须从这些知觉单元中确定最合理的一个。

在上述情况下,单词识别进入核证阶段,所谓核证就是将已经激活的知觉单元与输入的刺激模式作比较。这种比较是按词汇记忆中词的排列顺序,依次进行的。每次只加工一个单词表征。当某个单词表征与输入的特征匹配时,与表征相联系的知觉单元就被选择出来,并据此作出反应;如果不匹配,加工系统就会选择下一个知觉单元,并按上述方式继续进行加工,直到找到正确匹配为止。例如,用视觉方式呈现单词 house。在特征提取阶段,与 house 类似的一些单词如 mouse, rouse 等,都被激活,并成为候选的知觉单元。以后在核证阶段,当加工系统把这些候选单元与原来呈现的字母串进行匹配时,只有 house 完全匹配,因而被选择出来。

模型既解释了在无语境条件下的单词识别,也解释了语境条件下的单词识别。如果在一个单词呈现以前,出现一个有语义联系的启动词(prime word),那么,对词的识别会加快。核证模型对这一现象的解释是:语义启动物使要核证的单词立即起作用,因而有效地绕开了对原始特征的计数作用,加快了整个反应的速度。

模型也解释了词频的作用。在一长列候选单位或候选表征中,熟悉的词排在前面,罕见的词排在后面。加工系统在核证时,先核

证高频词,后核证低频词,因而对高频词的识别比对低频词快一些。

核证模型强调词汇信息与感觉信息的匹配。因此,单词识别依赖于刺激项目在感觉记忆中的存储时间。只有当感觉信息保持时,核证才能进行。如果由于某种原因,如采用后向掩蔽的方法使感觉映象迅速消失,核证就无法完成。在这种情况下,人们怎样识别单词?这是用核证模型无法回答的。

(4) 相互作用激活模型(the interactive-actiration model)

近年来,在字词识别的模型中最有影响的是由 J. L. McClelland 等人(1981,1982,1986,1987)提出的相互作用激活模型。

模型假定的加工单元有些类似于词产生器。它们与语言的结构水平是对应的。由结构的下层往上走,可区分为视觉特征单元,字母单元和词单元等。当一个词用视觉方式呈现时,它会首先激活有关的视觉特征单元,然后激活适宜的字母单元,最后以类似的方式激活词单元。这个过程是对视觉信息的一种自下而上的加工。

相互作用激活模型假定,单词识别不仅依赖于自下而上的加工,而且依赖于自上而下的加工。也就是说,不仅低层单元的激活会影响到高层单元的加工,而且高层单元的激活也会反过来影响到低层单元的继续加工。它们的关系是相互作用的。

模型的另一个特点是:某个单元的激活还会抑制同一层次的所有其它单元的活动。例如,用视觉方式呈现字母串 HOUSE,这个字母串引起特征单元的激活,并在字母层次上激活 H,O,U,S,E 等字母,进而激活 HOUSE,HORSE,MOUSE,ROUSE 等词单元。这时,在词水平的激活又反过来影响字母水平,提高对字母单元 H,O,U,S,E 及 R 与 M 的激活作用。接着,信息继续进入特征水平,它使 H,O,U,S,E 的激活作用上升,而 R,M 开始失去激活。结果词单元 HOUSE 的激活水平明显地高出 HORSE 和 MOUSE 等词单元,并使 HOUSE 与其它词明显区分开来。在这种

情况下,单词 HOUSE 就被识别出来了。

图 5 说明了相互作用激活模型的一些特点。

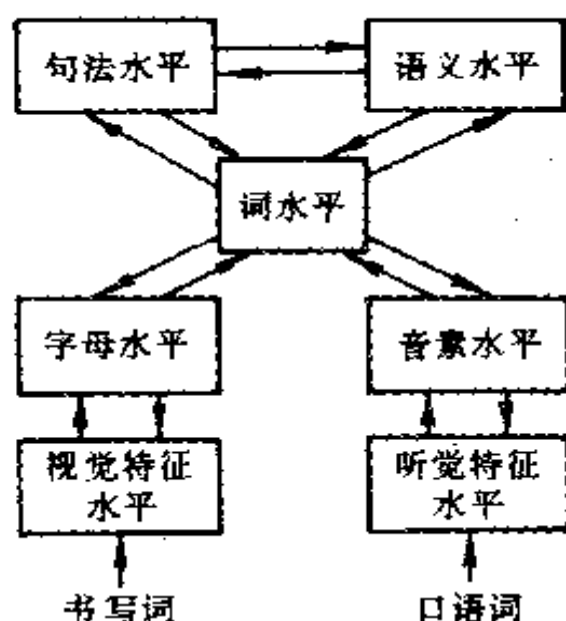


图 5 相互作用激活模型示意图

(5) 单词识别的分布表征模型(distributed representation model)

与每个单词相联系的信息是各种各样的,有字母、字母群、前缀、后缀、词的整体形状、第一个字母、语义特征、联想、句法关系、语音模式等。每种信息并不是与一个词单独发生联系,而是以一种分布表征的形式和其他词共同使用每一种信息(见图 6)。

在分布表征中,当一个词呈现出来时,具有相同信息的另外许多词也就不可避免地会激活起来。

在分布表征模型中,单词识别不需要从心理词典中搜索一个目标词。单词的任一属性的激活都可能增加一个熟悉的指标。当这个指标(index)达到阈限时,该字母串就会判断成一个单词。由于熟悉的单词比不熟悉的单词能更有力地唤起更多的属性,因此对熟悉词的词汇判断比不熟悉的单词快。同样,对词的判断快于非词,是因为前者唤起了许多有关的语义特征,而后者唤起的特征很

少(虽然不是全无)。

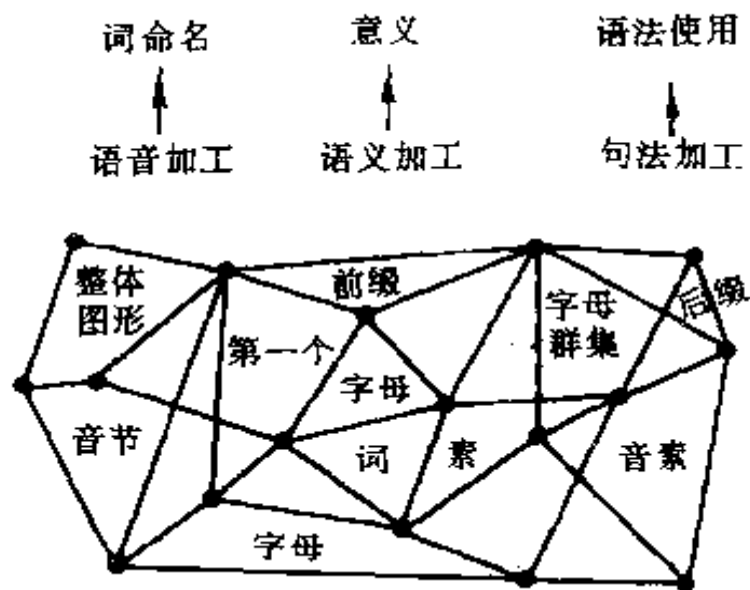


图6 关于词的各种信息的分布表征

在对单词命名的计算机模拟中,分布表征模型没有词典,它所模拟的许多效应与在人类被试中看到的效应是相似的,如词频效应、规则效应(regularity)(对符合字母-声音对应规则的词,如mint 比对不规则的词如pint 发音要快些)、词频与规则的交互作用(高频词规则效应小,低频词规则效应大,因为高频词是作为一个整体模式加工的,而低频词是按字母-声音转换规则加工的)等。

三、句子理解

1. 句法分析

(1) 句法分析与理解

句法分析也叫析句(parsing),它是指从语法上分析词或句子,将句子切分成不同的成分。确定哪个成分是主语、宾语、补语、修饰语以及它们与动词的关系。例如,对下面这个句子可以进行各

种不同的切分,结果有的容易理解,有的理解起来比较难,按照不同的成分进行切分,得到的意义也可能不同。

【例】落雨天留客天留我不留。

① (落雨天留)(客天留)(我不留)

② (落雨天)(留客天)(留我不)(留)

③ (落雨天留客)(天留)(我不留)

第一种切分,意义不清楚;第二、第三种切分,虽有意义,但释义完全不一样。可见,句法分析对句子理解有重要的作用。当一个句子受到不正确的切分时,理解将出现困难。在口语理解时,由于听话者不能反复感知所听到的言语刺激,正确地切分句子就显得更加重要。

人在理解语言时,是怎样切分句子的成分呢?

70年代初,W. Levelt 和 E. Martin 应用主观判断方法进行了句子切分的实验研究。实验者给被试呈现一个句子,要求被试判断句子中哪些单词联系最紧密。所用的实验句如:

The boy has lost a dollar.

从实验句中随便抽取 3 个单词,如 boy, has, a, 要求被试通过两两比较判断它们在句子中的关系:哪两个较近,哪两个较远。实验结果用聚类方法处理。结果发现,被试对句子切分的形式为:

((The boy)((has lost)(a dollar)))

可见,被试先把句子切分成一些短语,如 The boy, has lost, a dollar, 然后按(主语(动词,宾语))(S(VO))的形式切分句子。

Levin 和 Turner (1968)对阅读时眼音距(eyevoice span, EVS)的研究也发现,眼音距的大小与短语边界相吻合,在短语中没有说出的最后一个词是能够看到,并能被预测出来的。

由于短语常常由一些内容词和少数功能词组成,而研究发现,人在阅读时对功能词几乎不进行加工。因此,对短语的加工主要是对内容词的加工。

对一些比较复杂的句子来说,被试也往往按从句来切分句子。从句是比短语更高一层的加工单元。这里,我们可用 Fodor 和 Bever 的实验来说明。实验采用咔嗒声的听觉作业。给被试一耳呈现一个句子,另一耳呈现咔嗒声,要求被试写出这个句子,并标出咔嗒声所在的位置。句子是

((That (he (was happy))) (was (evident (from
1 2 3 4 5 6 7 8 9
(the way) (he smiled))))))

咔嗒声实际出现的位置如句中数字所示。其中数字“5”所在的位置是句子成分的边界。研究发现,当咔嗒声出现在位置“5”时,被试的判断最准确;当它出现在其它位置时,被试的误听较多,而且判断的位置转向“5”。这个事实说明被试倾向于按从句来切分句子。

(2) 句法分析的策略

析句是一种认知过程,在句子理解过程中占据工作记忆的部分容量。一个分析者的目的就是要正确而迅速地分析一个句子,而不让工作记忆负担太重。为了这一目的,析句者发展了某些策略。

① 标准句策略:NVN=SVO

词序是析句的一种重要线索。英语具有比较严格的词序。由名词—动词—名词(NVN)组成的最常见的结构,形成了标准的SVO句(主—动—宾)。因此,人们形成一种策略:把NVN解释为SVO,也叫标准句策略(canonical-sentence strategy)。在长句和复杂句中这种策略也很常见(Bever 1970)。

② 功能词策略(the function-word based parsing strategy)

功能词按其定义就是指起句法作用的词,因而它应该是句法分析的重要线索。

Kimball(1973)提出:“不论何时你发现一个功能词,那么就开始了一个新的句子成分”。

不同功能词有不同的策略。

* 在确认一个定冠词之后,寻找一个名词,它以一个名词短语结束:“the house”。

* 在确认一个介词后,寻找一个名词,它以一个介词短语结束:“of the house”。

* 在确认一个表示时态的助动词后,寻找一个主要动词,以一个动词短语结束:has gone。

问题是,人们是否使用功能词策略?如果功能词是标准功能词(prototypical),人们可能不用。因为标准功能词的语义内容很少,词很短,冗余性多,因而不为读者注意或注视,在口语中,标准功能词也不强调,几乎达到消失的程度。它们虽然存在于听觉信号中,但听话者知觉不到。既然人们觉察不到它们的存在,又怎样作为析句的线索?但大多数功能词(如 between, because)不是标准的,它们可能作为线索而起作用。

③ 晚终止(late-closure)策略和最简依附(minimal attachment)策略

在分析句子时,不要急于根据前面的材料对句子的结构作出判断。运用这种策略来分析语句,就叫 late-closure strategy,以下面两句为例(Frazier 和 Rayner, 1982)。

Since Jay always jogs a mile and a half this seems like a short distance to him.

Since Jay always jogs a mile and a half seems like a short distance to him.

第1句不是歧径句,第2句是歧径句(garden path sentence)借助于标准的 NVN=SVO 策略,名词短语 a mile and a half 开始可能被当成动词 jogs 的直接宾语,这种分析适合第1句,但不适合第2句。在第2句中,因为动词 seems 没有主语,读者到此要重新把 a mile and a half 分析为一个新从句的主语,因此阅读第2句的时间

长于第1句。

所谓最简依附策略就是在析句时,人们倾向于把动词后的名词或名词短语看成是它的直接宾语。

The city council argued the mayor's position forcefully.

(市议会激烈地争论市长的职位。)

The city council argued the mayor's position was incorrect.

(争论市长职位的市议会是不正确的。)

根据最简依附策略和标准句策略(NVN=SVO),第1句中的 mayor's position 容易当成 argued 的直接宾语。这种情况对第2句不适用。

(3) 句式与句子理解

人们日常听到、见到的句子大致有:简单、肯定、主动的陈述句,即 SAAD 句(simple-affirmative-active-declarative),被动句(P 句, passive)、否定句(N 句, negative)、被动否定句(PN)、命令句(imperative)、Wh-疑问句(WhQ)、Yes/No 疑问句(Yes/NoQ)、被动疑问句(PQ)、否定疑问式(NQ)、被动否定疑问句(PNQ)等。不同句式对理解有什么影响?

研究已经发现,在每种语言中,否定式是交往的一种必不可少的句式;对否定句的加工一般长于对肯定句的加工。其原因可能有:①否定句在句法上比肯定句复杂;②否定句的加工有两个阶段:先得到它的肯定意义,然后加以否定(Chase 和 Clark 1972);③它要求将它的内容与听者记忆中的东西作比较;④否定句有许多可能的意义,含义相对不够明确。

Coleman(1965),Horgan(1978)发现,对正常成人来说,被动句比主动句较难理解、产生和学会。原因可能是它与标准词序不一样,它没有主动句直接、生动。例如:如果让被试听一段课文,并进行复述,那么他们倾向于用肯定语气来回忆它,而不管它原来的语气怎样。

60年代初, Miller, McKeam 和 Slobin 等, 根据乔姆斯基的转换理论, 提出了复杂性派生理论, 用它来解释句式对理解的影响: 所有句式都由核心句通过转换规则派生而来; 转换的次数越多, 理解越困难, 但后来的一些研究, 发现这种理论不能预测对一些句子的理解时间。

70年代初, Goldman-Eisler 和 Cohen 提出了口语中句子的相对频度的问题。1983年, Taylor 等报告了书面语言中句子的相对频度。见表1。

表1 英文句子的相对使用频度

结 构	例 句	相对频度
SAAD 句	He read the book.	80.7%
(P.) 句	The book was read by him.	13.8
(N.) 句	He did not read the book.	2.2
PN 句	The book was not read by him.	1.7
命令句	Read the book.	1.1
Wh-疑问句	Why did he read the book?	0.6
Yes/No 疑问句	Did he read the book?	0
PQ 句	Was the book read by him?	0
NQ 句	Did he not read the book?	0
PNQ 句	Was not the book read by him?	0

结果发现, 句子结构的相对频率是与早期研究中发现的加工句子的相对难度大致平行的。

2. 理解中的语用与语义因素

在进行句法加工时, 人们常常利用语用和语义因素。语用因素

是指使用语境和有关世界的知识,语义因素是指使用句子中内容词的意义及其相互关系。

(1) 语用因素——合理性(plausibility)问题

句子理解的容易程度可能部分地依赖于它所描述的事件的合理性。合理的句子是指它描述了最可能发生的事件。在这些事件中,施事和受体(patient)对所给予的动作或动词,其角色是典型的;不合理的句子描述了高度不可能的事件;此外还有中性句。在这种句子所描述的事件中,施事和受体扮演的角色是可以互换的。如:

- ① The mother feeds her baby milk.
- ② The baby feeds its mother milk.
- ③ The baby smiles at the mother / The mother smiles at the baby.

句①是合理句(可能发生的)。对这种句子的理解只需要加工其中的内容词,如 mother, baby, feed, milk。运用日常的知识就能将这些内容词联系起来,几乎不必进行句法分析。而不合理句或中性句,就需要句法分析。当句法分析的技能降低时,如学前儿童或脑损伤病人,合理性这一因素将起重要的作用。

合理性因素的作用,还表现在对被动句的理解上。在一般情况下,被动句的加工比主动句难。但是,如果被动句中施事与受体不能互换位置(角色不能互换)时(不可逆被动句),这种被动句比可逆被动句容易理解。

- ① The patient was treated by the doctor.
- ② The boy was hit by the girl / The girl was hit by the boy.

不可逆句①只要求从句法上确证一下根据语义-语用内容已经得到理解的东西;而可逆句②要求句法分析。

Herriot (1969) 在一项实验中,给被试呈现一些句子:主动句或被动句,要求被试尽快说出句中哪个名词代表动作者,哪个名词

代表受动者。每种句子又有合理的如句①；不合理的，如句②两类：

① The patient was treated by the doctor.

② The doctor was treated by the patient.

结果，主动句与被动句的反应速度没有差异，而合理句与不合理句之间有区别。被试容易对合理句作出正确的判断。

当句子结构异常复杂时，正常成人可能依赖于合理性。

在下面的3个句子(中心嵌入句)中，①属于可能发生的句子(人抚摩、狗咬、猫咪咪叫)，因而容易加工；②属于中性句，要求细心地进行分析；③是不合理句。

① The cat that the dog that the man stroked bit miaowed.

② The boy whom the girl whom the man kissed saw left.

③ The man that the cat that the dog stroked bit miaowed.

Stolz(1967)发现，在中心嵌入句中，如果谓语在语用上与其主语一致；而不适合于其它从句的主语，那么这种嵌入句就能得到理解。Schlesinger(1968)发现，被试几乎总是采用与语义期待一致的解释，而不采用与句法分析一致的解释。

(2) 语义因素

语义因素指句子中个别内容词的意义及其相互关系。如果句子中的单词是不常用词，具有抽象的意义，特别是，这些词用一种抽象的、复杂的、新奇的方式联系起来，那么这种句子就是语义复杂的句子。我们可以比较下面两个句子。

① Energy equals mass times the speed of light squared
($E=mc^2$).

② The number 8 equals 2 times the number 2 squared ($8=2 \times 2^2$).

句①是爱因斯坦的著名的相对论公式，句②是在句法上相同的常识句。两个句子在理解上的差异主要是由于语义复杂性造成的。

French(1981)在一项研究中,比较了语义因素与句法因素的作用。在语义方面,他控制了子句中内容词可能整合的程度。

① 整合得很好的句子(SVO 形成了紧密的语义单元)。如:
The little baby drank the milk.

② 整合得不好的句子。如: The aunt saw the door and left.

③ 不合规则的或句法正确但无意义的句子。如: My tasty owner spilled the captain madly.

在句法复杂性方面也作了控制:

简单: The boy hit the ball.

复合: The boy hit the ball and run.

复杂: After hitting his sister, the brother cried.

混乱排列: the ate fat grass green cattle the.

实验中句子按一个词、一个词作快速视觉序列呈现(rapid serial visual presentation RSVP)。呈现速度有两种:①低速:每秒14个词(高于正常阅读速度);②高速:每秒20个词。结果发现在低速时,语义和句法因素都起作用;而在高速时,只有语义因素起作用。也就是说,按在快速呈现条件下报告的单词数计算,好整合句优于差整合句,差整合句又优于异常句。

Boothroyd & Nitttrouer(1988)让被试在噪音背景上听3种短句:正常的、异常的、乱排的。结果发现,影响句子识别的最重要的因素是语义因素。

3. 句子理解的模型

通过前面的分析,我们知道,在理解句子时,人们使用了各种各样的信息,包括语言的、非语言的信息。如语音信息、语义信息、句法信息、语用信息等。这些信息是彼此独立起作用,还是交互作用?

(1) 自主模块模型(autonomous modular model)

Forster (1979), Cairns (1984), Garrett (1982) 提出了自主模块模型。他们认为, 句子理解系统由3种加工模块组成, 即词汇加工、结构加工和解释加工。它们是按顺序发生的。

词汇加工器只根据语音信息从词典中提取单词; 而结构加工器 (structure processor) 只用句法信息产生句子的成分结构。结构加工器的产品成为解释加工器 (interpretive processor) 的输入, 由解释加工器使用有关世界的知识构成句子意义的表征, 它还综合若干个句子的信息, 必要时作出推理, 因而也可叫话语加工器 (discourse processor)。根据这种理论, 句法分析与语义分析是独立起作用的。句法分析系统对句子进行结构分析, 然后与语义系统匹配, 如果不成功, 则重新进行结构分析。

Rayner 等人 (1983) 的实验支持了自主理论。让被试阅读两个歧径句①与②, 并记录他们的眼动。

① The florist sent the flowers was very pleased.

② The performer sent the flowers was very pleased.

结果发现在句中单词 was 上, 停顿时间都较长, 并有回跳发生。

按自主理论, 在不同句子中, 被试都根据最简附着原则, 把 sent 分别确定为主句的动词, 它们的主语是 florist 和 performer。在这里, 句②中的语义信息未起作用, 因而出现了与句①同样的错误。相反, 如果语义起作用, 被试就应该根据 performer 的词义, 想到它是受花者, 而不是送花者。这样被试在 was 上的注视停顿时间应该有差别。实验结果表明, 被试在读到 was 时, 才将对句②中 sent 的错误理解改正过来。因而在 was 上的注视时间也较长, 并出现回跳。这说明语义分析并未参与对句子的理解, 语义分析与句法分析是独立起作用的。

Ferreira 和 Clifton (1986) 用类似的材料获得了相同的结论。他们的实验句是:

① The defendant examined by the lawyer turned out to be unreliable.

② The evidence examined by the lawyer turned out to be unreliable.

按最简附着原则,被试应该把 examined 看成是主要动词,它要求自己的施动者。如果被试对句子进行了语义分析,那么句①中的 defendant 是有生命的名词,它可以作为 examined 的主语(施事),而句②中的 evidence 是非生命的,它不能作为 examined 的施事。这种语义分析会帮助被试确定句②的主要动词不是 examined,因而加快对句②的阅读速度;而句①得不到这种帮助,因而阅读速度会受到歧义的影响而变慢。但实验结果发现,两个句子的语义区别对阅读速度没有影响。这一结果也有利于句法自主理论。

(2) 相互作用模型(the interactive model)

主张这种观点的心理学家有 Tyler 和 Marslen-Wilson (1977),Crain 和 Steedman(1985)等。这种模型主张,在语言理解的每一阶段都存在句法、语义和语用因素的相互作用。语义-语用信息可以限制、指导句法分析,得出正确的结构分析(强相互作用观点),也可以从对句法分析得到的多种结果中作出即时的选择,从而影响进一步的句法分析(弱相互作用观点)。

Tyler 和 Marslen-Wilson(1977)的实验支持了强相互作用的观点。他们以听觉方式给被试呈现句①与句②。在 landing planes 之后,立即以视觉方式呈现单词 is 或 are,要求被试尽快大声念出来。结果发现,对句①来说,被试读出 are 比读出 is 快,而在句②中,被试读出 is 比读出 are 快。

① If you walk too near the runway,landing planes...

② If you've been trained as a pilot,landing planes...

研究者认为,在句①背景(语境)的影响下,landing planes 被理解为“正在着陆的飞机”(planes that are landing),因而限制和

指导了句法分析,使被试较快地选择了 are。而在句②语境的影响下,landing-planes 被理解为飞机着陆的活动,因而使被试较快地选择了 is。这说明语境中包含的语义信息对 landing planes 作出了与语境一致的分析,对句法分析产生了直接的影响。

Crain 等用实验支持了弱相互作用的观点。他们用“快速系列视觉呈现”方法给被试呈现实验句,即一个词、一个词地呈现,要求被试作语法判断,即是否合于语法。实验句为:

① The teachers taught by the Berlitz method passed the test.

② The children taught by the Berlitz method passed the test.

结果,被试对句②的正确判断率显著高于句①。这是因为在被试看来,children 是受教育者,因而把 taught by 当成一个定语从句,taught 是从句中的动词,判断的正确率较高;而在句①中,教师是教育者,taught 可能成为 teachers 的谓语,因而导致了判断的错误。

四、话 语 理 解

话语(discourse)是指比句子更长的语言段落,如口语中的一段话;书面语言中的一段或一篇课文等。话语由前后连贯的句子构成,比句子更复杂。因此,对话语的理解,既要以对词和句子的理解为基础,又要经历更复杂的过程,执行更复杂的认知操作。本文限于篇幅,只介绍话语理解的两种模型。

1. 话语的微观结构模型

1978年 Kintsch 和 Dijk 在《心理学评论》上发表了“论课文理解和产生的一种模型”。1979年, Kintsh 在《教育心理学家》杂志

上；又发表了“论建立理解的模型”，提出了以分析话语微观结构为特征的理解模型。以后，这个模型经过一系列补充、修订，并不断趋于完善。

Kintsch 的基本假设是：

(1) 理解是指在读者头脑中建构由课文提供的信息的一种心理表征

理解一方面依赖于课文的性质与特点，即由课文输入的信息，另方面又依赖于读者的知识、信念和目标(见图7)。

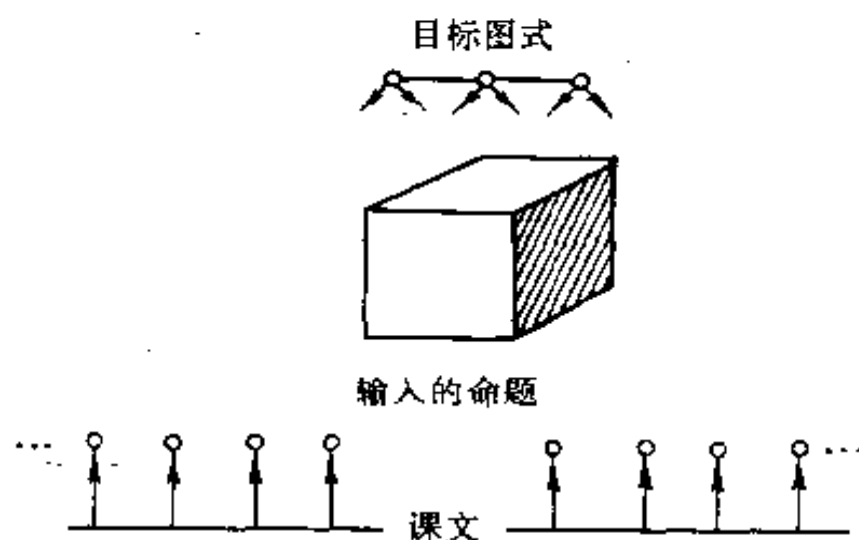


图7 课文理解的模型

(2) 话语的基本结构单元是命题

命题是一种概念单位，既可为真，也可为假。它不是一个句子，但可用句子表示。一个句子可以表示一个命题，也可以表示多个命题。

(3) 话语理解是由建构某种命题网络来完成的

这种网络表示了命题间的关系。如果两个命题具有一个共同的主题或逻辑中项(argument)，那么，这两个命题就彼此有关，或者说，是前后连贯的(coherence)。在话语理解中，命题间的连贯性有重要的作用，它可以用连贯图(coherence graph)来表示。

(4) 话语加工的周期性

由于工作记忆的容量有限,人们不可能把课文中所有以往的信息保持在工作记忆中,即让它处于激活的状态。因此,对话语的加工必须是周期性的或环状的。在各个加工周期或加工环之间,需要有一个记忆缓冲器起搭桥的作用。来自当前加工周期的少数命题,即一些最重要的命题,被保存在缓冲器中,从而把前面的课文与现在输入的信息联系起来。如果在缓冲器的内容与新的输入间没有联系,那么就需要读者进行额外的加工。它包括从长时记忆中搜索过去的课文,即复位搜索(reinstatement search)及以知识为基础进行的推理。通过复位搜索和推理构成必要的联系。

Kintsch(1979)在一项研究中,曾引用一段课文,说明上述模型的一些特点。课文的全文是:“Swazi 部落由于牲口的纠纷,与一个邻近的部落发生了战争。在这群武士中,有2个未婚的男子,Kakra 和他的弟弟 Gum。在一次战斗中,Kakra 被人杀死了。按照部落的风俗,Kakra 后来与妇女 Ami 结了婚”。

现在我们先分析第一个句子:“Swazi 部落由于牲口的纠纷,与一个邻近的部落发生了战争”。在这个简单句中,有5个独立的命题:①Swazi 是一个部落;②这个部落处于交战状态;③这场战争是与一个邻近的部落进行的;④战争有一个原因;⑤这个原因就是牲口的纠纷。由于命题①与命题②具有共同的主题或逻辑中项:部落,因此相互联系在一起。同样,命题②与命题③由于共有的主题——战争而联系起来。命题②与命题④也由于共有的主题——战争而联系起来;命题④与命题⑤由于“原因”这一共有的主题而联系在一起。5个命题的这些关系可以用连贯图表示出来(见图8)。从图上看到,命题②与大多数其它命题都有联系,因而是这组命题中最高级的命题,或主要思想。

现在我们再看第2个句子:在这群武士中,有2个未婚的男子 Kakra 和他的弟弟 Gum。这个句子包含的命题有:①有一群武士;

②有2个男子均为武士；③2个男子都没有结婚；④2个男子均有名字；⑤他们的名字叫 Kakra 和 Gum；⑥Gum 是 Kakra 的弟弟。我们同样可用一张连贯图表示这些命题的关系(见图9下方)。

连贯性分析——第一周期

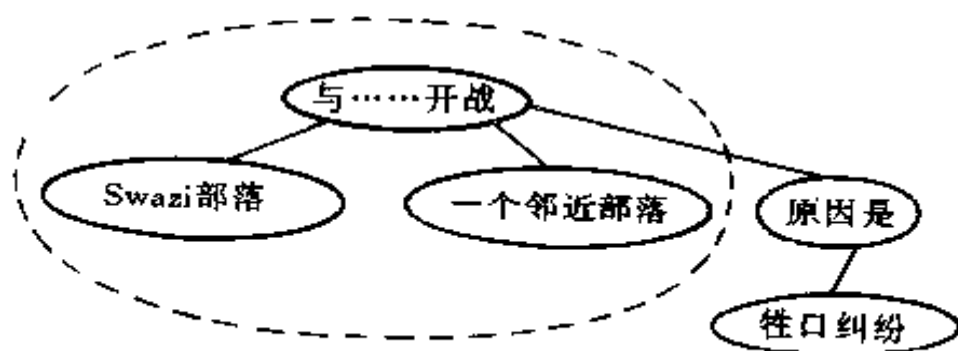


图8 第一个句子的连贯图的结构

连贯性分析——第二周期

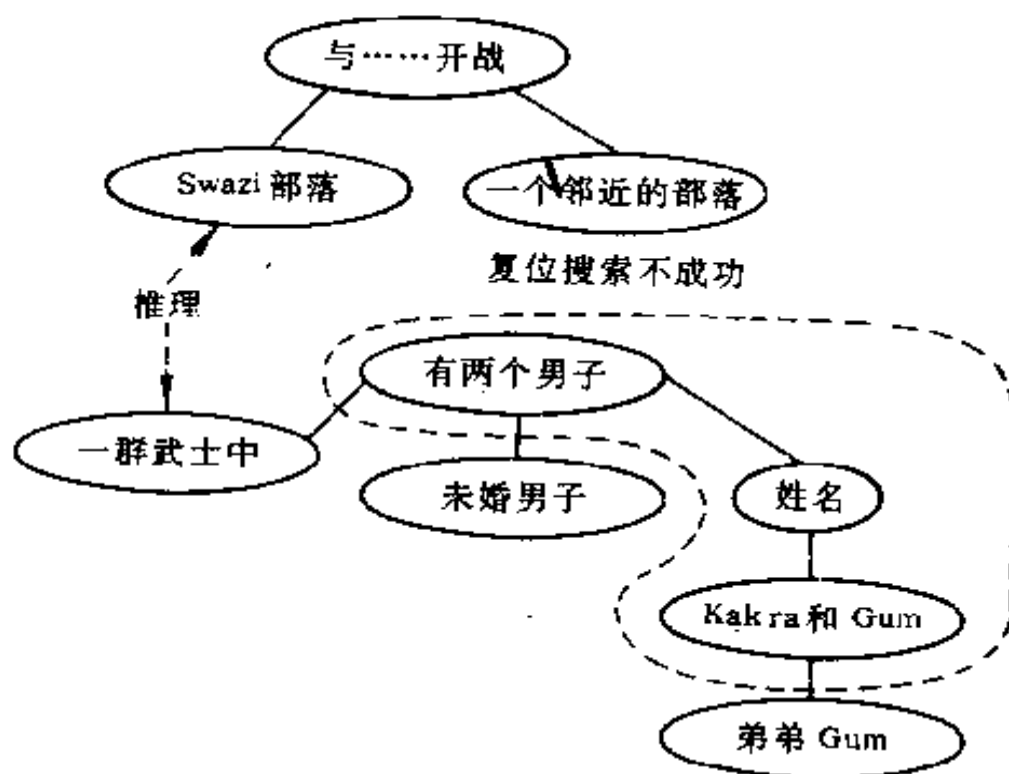


图9 第二个句子的连贯图的结构

为了建构整篇课文的心理表征,读者不仅要分别加工第一和第二个句子,而且必须把2个句子联系起来,即发现2个句子所包含的各个命题的关系。在进行第二个周期的加工时,由于工作记忆容量的限制,读者必须在命题和联系中进行选择,即在工作记忆中保存第一句中所包含的高级命题以及与它邻近的命题(如图9上方所示)。接着,他要进行复位搜索,即发现第一句的3个重要命题与第二句的6个新命题的关系。但是,从图上我们看到,在缓冲器中保留的命题和新输入的命题之间没有共同的主题或概念,因而无法通过复位搜索建立起它们之间的联系。在这种情况下,读者只有借助于推理来完成新旧命题之间的联系:这些武士是 Swazi 部落的成员,正是读者给课文补充的这个命题,把前后两个句子的命题联系起来了。

接着,进入第三个加工周期,即对课文的第三个句子(Kakra was killed in a battle)进行加工。这时,在短时记忆的缓冲器内保存着3个命题,由于这些命题与新输入的命题具有共同的概念:Kakra,因此,这两个句子很容易连贯在一起(见图10)。

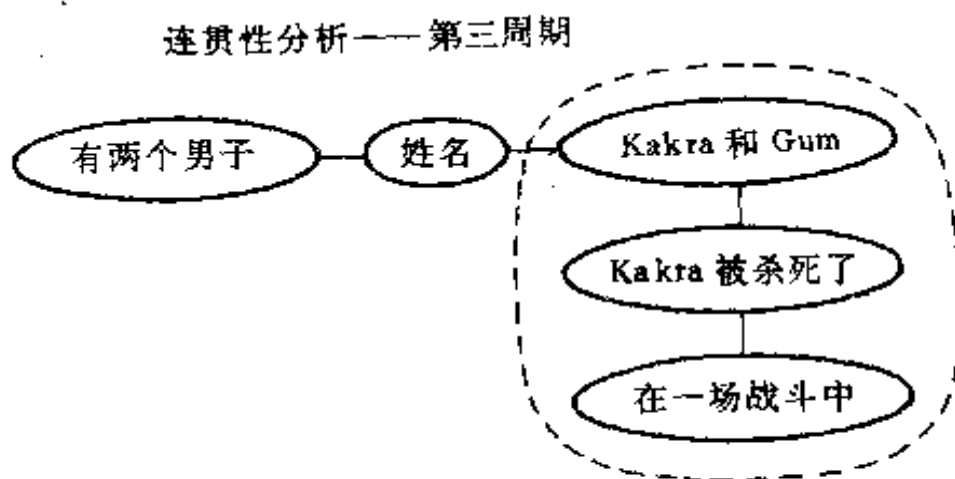


图10 第三个句子的连贯图的结构

依据同样的道理,我们还可以画出第四个句子的连贯图(见图11)。从图上看到,Kakra 不仅把第2、第3个句子联系起来,而且也

把第2和第4个句子联系起来。Kakra 死后,是按照部落的习俗和 Ami 成亲的。在这里,读者有关鬼婚的知识将对理解这段课文有重要作用。

连贯性分析——第四周期

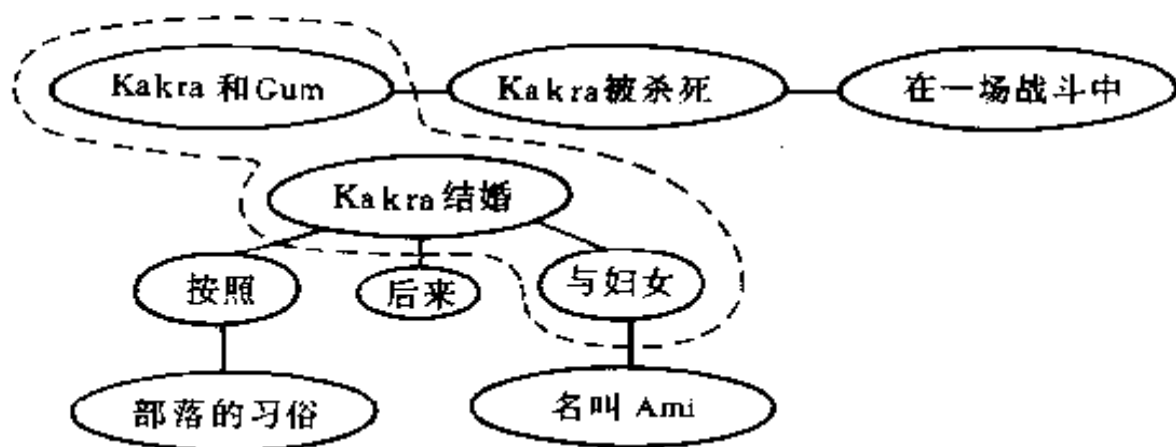


图11 第四个句子的连贯图的结构

在 Kintsch 的理解模型中,有3个重要的参数。一个是短时记忆缓冲器的容量,通常为4个命题;一个是每个周期的最大输入数,即每次可接受的新命题的数量。这个数量通常依赖于每个句子中命题的数量。最大数量为5—8个命题,平均为6.2个命题;一个是命题在长时记忆中保存的概率和以后被回忆出来的概率。短课文的平均学习概率相当高,等于0.64。Kintsch 用他的模型预测了课文的可懂度或难度。在他看来,在预测课文的难度时,词频和复位搜索的数量是2个最重要的预测指标。此外还有命题密度、推理的数量、加工周期的数量和命题中不同主题(或逻辑中项)的数量等。Kintsch 认为,可懂度不仅由课文决定,而且是特定读者与课文相互作用的结果。

2. 话语宏观结构的模型

话语的微观结构分析是指分析话语中个别句子的关系。在进行这种分析时,命题的概念具有核心的意义。而话语的宏观结构分

析是指分析话语的整体结构或意义。在进行这种分析时,图式的概念是特别重要的。

图式的概念起源于哲学家康德。以后,经过皮亚杰和 Bartlett 的努力,应用到现代心理学中来。图式是语义记忆的一种结构。它说明了一组信息的最一般的排列或可以预期的排列。例如,一种故事图式包含了我们在故事中期待发生的所有信息。故事图式包括探侦故事、神话故事、传奇故事的图式等多种。

下面我们用 Thorndyke 的模型来说明话语宏观结构模型的一些特点。

Thorndyke 描述了一系列规则,用来说明故事的良好结构。这些规则类似于乔姆斯基的短语结构规则(见表2)。

从表2我们看到,一篇故事是由场景(setting)、主题(theme)、情节(plot)和结局(resolution)组成的,而场景又由角色(人物)、地点和时间组成……在上表中,*号代表一个符号可以重复使用。例如,总的情节可以区分为一个或多个片段;而片段可以包含一种或多种尝试;一种尝试可以由一个或多个事件组成;而结果又可以包含一个或多个事件等。可见,故事语法的规则是一种一般性的知识结构,而不是一组固定的事实。在不同的故事中,它的具体表现形式可能是不同的。

表2 简单故事语法的规则(Thorndyke, 1977)

1. 故事 → 场景 + 主题 + 情节 + 结局
2. 场景 → 人物 + 地点 + 时间
3. 主题 → (事件)* + 目标
4. 情节 → 片断*
5. 片段 → 子目标 + 尝试* + 结果
6. 尝试 →

→ 事件*

→ 片段
7. 结果 →

→ 事件*

→ 状态
8. 结局 →

→ 事件

→ 状态
9. 子目标 → 所希望的状态
10.

人物
地点
时间

} → 状态

现在让我们用 Thorndyke 的故事语法来分析下面这篇故事。

① There was once an old farmer ② who owned a very stubborn donkey. ③ One evening, the farmer was trying to put his donkey into its shed. ④ First, the farmer pulled the donkey, ⑤ but the donkey wouldn't move. ⑥ Then the farmer pushed the donkey, ⑦ but still the donkey wouldn't move. ⑧ Finally, the farmer asked his dog ⑨ to bark loudly at the donkey ⑩ and then by frightening him into the shed. ⑪ But the dog refused. ⑫ So then, the farmer asked his cat ⑬ to scratch the dog. ⑭ So the dog would bark loudly ⑮ and would thereby frighten the donkey into the shed. ⑯ As soon as the cat scratched the dog, ⑰ the dog began to bark loudly. ⑱ The barking so frightened the donkey ⑲ that it jumped immediately into its shed.

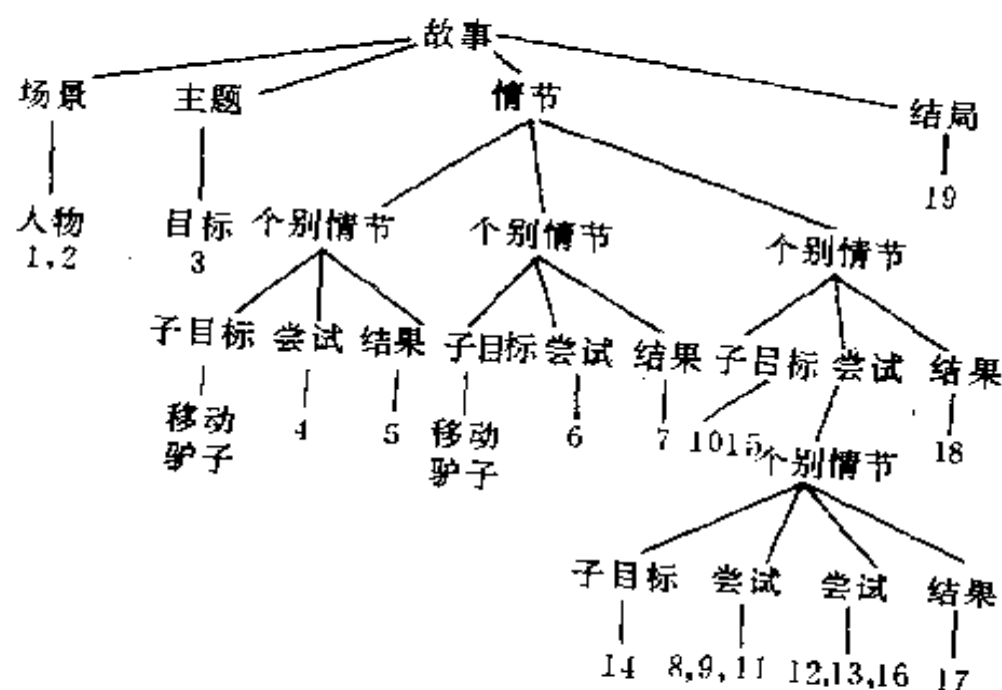


图12 对“农夫与驴子”一文的分析

(图中的数字与文中的数字对应)

研究故事的宏观结构,提供了故事的一般图式。这种图式说明

了故事中各种事件应有的顺序、具有层次性。图式还具有组织作用,一些与图式有密切关系的信息,能得到最好的回忆,而其它一些信息则往往被人们忽略;图式还提供了一种提取计划,它能唤起某些细节,而不是其它细节。图式具有预测作用,根据被激活的图式,人们能期待故事的发展和结局。图式指导着人们对话语的加工,当人们能对话语形成正确的图式时,理解较容易;相反,理解将出现困难。

主要参考文献

- [1] 彭聃龄、谭力海(主编)《语言心理学》,北京师范大学出版社,1991年。
- [2] Anderson, JR. ,Cognitive Psychology and Its Implication (2nd) San Francisco, Freeman 1985.
- [3] Carroll, D. W. ,Psychology of Language. Brooks/Cole Publishing Company 1986.
- [4] Chomsky, N. ,Language and Mind. Harcourt, Brace & World Inc. 1968.
Foss, D. J. And Hakes, D. T. Psycholinguistic and Introduction to the Psychology of Language. Prentice-Hall Inc. 1978.
- [5] Kintsch, W. and Van Dijk, T. A. ,Toward a model of Text Comprehension and Production. *Psychological Review*1978, **85**, 363-394.
- [6] Kintsch W. On Modeling Comprehension. *Educational Psychologists* 1979.
- [7] Rayner, K. and Pollatsek, A. The Psychology of Reading. Prentice-Hall Inc. 1989.
- [8] Taft, M. ,Reading and the Mental Lexicon. Lawrence Erlbaum Associates Ltd. 1991.
- [9] Taylor, I and Taylor, M. M. Psycholinguistics, Learning and using language. Prentice-Hall International Inc. 1990.

认知心理学与人工智能

陈 永 明

(中国科学院心理所)

认知心理学和人工智能是当代认知科学中的两个主要成员。这里,准备在认知科学的框架中,谈谈人工智能与认知心理学之间的相互关系和相互作用的问题。其目的,一是想强调心理学应该重视学科交叉点上问题的研究,在当代认知科学的发展中发挥它应有的作用,同时,需要关心和重视有关学科中的成就和方法;二是想说明人工智能的研究也需要以心理学的经验材料为基础,不管它是实验的还是直觉的;人工智能作为认知科学的一个成员,它的发展是与认知心理学研究联系在一起的。

当然,在认知科学的框架中,不同的研究者可以采取不同的方法或途径。认知心理学家提出某种假设或理论,然后用人当被试来做实验,对它们进行检验。人工智能研究者构建一个理论,然后把它转换或实现为一个计算机程序,对这个理论进行检验。可以说,两个学科均有其自身的长处和弱点。但是,如果它们彼此携起手来密切配合,那么,就会为人类认知的研究和机器智能的实现,提供良好的方法和环境。

一、人工智能需要人的智能作为模型

什么是人工智能?著名的人工智能专家明斯基(M. Minsky)认为,“人工智能是这样一门科学,它使机器去做那些如果由人来做的话,需要智能的工作”^[1]。另一位著名人工智能专家温斯顿(P.

H. Winston)在他所著的《人工智能》一书中说,“人工智能是研究如何使计算机去完成过去只有人才能做的智能工作”,“人工智能的主要目的,是要使计算机更加有用,并且要理解那些使智能成为可能的原理”^[2]。

既然人工智能是使机器去做那些需要人类智能才可做的工作,理解智能活动得以实现的原理,那么,人工智能研究的最佳途径应该是把人的智能活动作为模型,从宏观上对人类的智能活动进行模拟。纽厄尔(A. Newell)和西蒙(H. A. Simon)提出的物理符号系统的假设,就是计算机模拟的理论基础。西蒙曾指出,人的所有智能活动,都是可以用符号和计算来实现的^[3]。在研究人类解决问题的思维过程中,他们观察到,人解决问题时,并非纯粹地按逻辑规律办事,常常是根据经验和典型事例来解决问题。因此他们提出了启发式知识和启发式搜索的思想。这是一个重要的贡献。

当然,也有人认为,人们在研制飞机的时候,并没有把鸟的飞行方式作为模型,我们建造人工智能系统时,也不一定需要人的智能活动来作为模型。我们可以使机器采用一种完全与人不同的方法去完成同一件任务。这种类比意味着,人工智能研究不用考虑人的智能活动或心理学的证据,就可以研制智能系统,并成为认知科学的一个部分。这里,我们暂且不去讨论这种类比是否恰当,仅就这种主张本身来说,它是不可取的。这样做,会使人工智能脱离认知科学的轨道,而变成一门纯粹的工程学。

事实上,在人工智能研究者面前确实有两种可供选择的策略。

第一种策略,人工智能研究者可以放弃研究智能或思维活动的努力,专门去建造自动机或类似自动机的那些东西。显然,这种工作纯粹是工程学的事情。这对认知科学家来说也是有一定意义的。例如,它在说明完成某种任务所需要的最低限度的知识方面,是有用的;它可以提供一些必要的依据,用以解释人类的作业。但是,它本身已不属于认知科学,因为它的研究对象和目标脱离了人

类的认知。我们说它对认知科学家是有用的,只是像数学和统计学那样有用的意义上说的。

第二种策略,也是一种最现实的策略,即人工智能研究者把人的智能活动作为自己的模型,把人工智能的研究置于心理学的经验材料的基础上。当然,在这种情况下,我们还可以把“心理学的经验材料”这个概念的含义扩大一点,把个人对人的认知活动的直觉(或经验)也包括在内。有些人认为,没有必要根据心理学实验取得的经验材料来研制智能系统,但是,他们实际上却在依据自己关于人类智能活动的直觉或经验去构造计算机程序。我们认为,人工智能工作者应该意识到,与其说凭自己对心理事实的直觉去构造计算机程序,不如有意识地考虑利用比较客观的心理学方法收集起来的有关人类认知活动的材料。

人工智能的研究以人的认知活动为模型,有两方面的好处:

① 研究和了解人的认知活动,一方面能使我们知道计算机系统可以做些什么事;另方面也可以告诉我们,什么是完成某种任务的最佳方式。拿人的容错能力(识别面容大变的老朋友、根据不完整的照片作出辨认等)和遗忘机制来说,它们的结合,创造了世界上最灵活和最有效的认知系统。所以有理由认为,即使是人工智能领域的工程师,尽管他们并不想成为一个认知科学家,也应该利用人的认知方式来作为自己工作的模型。

② 把人的智能行为作为模型,模拟人的认知活动,可以使我们进一步揭示和了解人类认知的规律和机制。正如温斯顿所强调的,这也是人工智能研究的一个主要目的。

二、人工智能作为人类认知的理论模型

作为一个认知模拟的程序,应该在某种规定的层次上,尽可能地与人对同一任务的加工过程相类似。如果我们写了一个程序,这

个程序只是展示出被解释行为的片段,就不一定能对人的认知活动作出良好的解释。就是说,机器和人在外部行为上的类似性(简单的模仿),并不说明这个程序是人类认知活动的一个理论模型;只是研究计算系统及其输出,不可能对人的认知活动找到合适的解释。

1. 外部行为与人类似的 ELIZA 程序

ELIZA 程序是由魏瑞鲍姆(J. Weizenbaum)研制的^[4],在人工智能发展的早期阶段是相当出名的。它模仿一个心理治疗专家。它与要求心理咨询的人之间进行的对话,非常像是一位心理治疗的专家;它可以蒙骗许多人,使人误认为与他对话的是一位心理治疗家,而不是一架机器。例如,若患者说:“我感到孤独”,ELIZA 就会问:“你感到孤独有多久了?”;若患者说:“我哭了”,它就会问:“你为什么要哭?”。那么,这样一个程序是不是人类认知活动的一个良好的解释呢?事实上不是。尽管它与患者的对话具有惊人的真实性,但是,它内部的加工过程完全不同于人的内部心理活动过程。

这个程序的工作机制,采用了关键词识别和模板匹配的方法。它内部存有许多模块,每一模块包含3个部分:①关键词;②输入模板;③输出模板(见表1)。

在输入模板中的“ ϕ ”,可以匹配任意数量的词,包括这个位置上没有一个词;如果不是 ϕ ,而是1或2,那么,在这个位置上只能有1个或2个词。

我们还看到,在输出模板中有一个“4”,它表示,这个位置上应插入输入模板的第4个位置上所匹配的词。

假定有人给 ELIZA 输入这样一句话:

I remember the first time I went to the beach. ELIZA 首先对句中的个别词作预处理,例如,将 I 转换成 you。然后,它找出关键词 remember。在 remember 下面,有某个输入模板可用来与输入

表1 ELIZA 内部存储的工作模块

关键词: (remember

输入模板: ((ϕ you remember ϕ)

输出模板: (do you often think of 4)

(what else do you remember)

(why do you remember 4 just now)

(what is the connection between me and 4)))

句相匹配(见表1)。这时,它就根据输入模板把输入句分解成4个部分:第一部分为空,第二部分为 you,第三部分是 remember,第四部分包括 remember 后面的所有词串。该输入模板下面,有4个输出模板,如果它利用第一个输出模板,那么,它的工作过程就是把由输入模板分解出来的第四部分的词串插在输出模板中“4”所注明的位置上,并作为其输出。结果所产生的反应是:

“Do you often think of the first time you went to the beach?”

如果它使用第三个输出模板,那么,将产生下述反应:

“Why do you remember the first time you went to the beach just now?”

一般说来,若一个输入模板下面带有几个输出模板,那么,它就会轮流使用这些输出模板,从而产生不同的反应。

若有两个以上的关键词可与输入句相匹配,那怎么办呢?在 ELIZA 系统中,每个关键词都有个等级号,具有高等级者,优先进行处理。另外一种可能的情况,是没有任何一个具体的模板可以与当前输入句成功地匹配,怎么办?为了防止出洋相,所以在 ELIZA 内还设置了一块总模板,可与任何输入句相匹配。与它相联系的输出模板,均属于一般性的反应。如:“这很有意思,请继续说下去。”“请你

说详细一点好吗?”等等。

从上面的叙述中我们可以看到, ELIZA 所采用的技术是: ①识别在输入信息中出现的关键词; ②选择合适的输入模板(输入模板实际上是一条分解规则, 是由关键词触发的), 并实现对输入句的分解; ③选择某个输出模板(它实际上是一条重组规则, 由与它相联的输入模板触发的), 并实现输出反应的构建; ④有时亦需要识别与关键词相联的最小的上下文, 如关键词是 you, 它后面是否是 are 等; ⑤设置了一种实际上很简单的机制, 在找不到关键词的情况下, 使它能“机智地”作出反应。

尽管 ELIZA 的输出反应很像是一位心理治疗的医生, 但是, 这样一个程序, 实际上只是以一种不理解的方式来模仿人类某些外部行为的特征。即使它的输出很像人类的反应, 但从其内部加工的机制来看, 显然不足以成为人类认知行为的一种理论解释。人类的语言交际, 是在对句子理解的基础上进行的, 而不是依靠 ELIZA 这种方式来实现的。

从这里我们可以看出, 尽管机器和人两个系统表现出相类似的外部行为, 这也不能说前一系统是后一系统的模拟或解释。换言之, 尽管在某个描述层次上, 两个系统的行为是相同的, 但也不能得出结论说, 有关一个系统的行为结构的知识, 可以用来解释另一个系统的行为。这是因为, 两个系统之间行为的类似性, 可能不是发生在那种关键性的描述层次上。解决这个问题的唯一途径, 是进行心理学的科学实验, 获取人类认知活动的材料, 然后以这种材料为基础, 去构造可以解释人类行为的计算机程序。

2. 基于心理学经验材料的思维模拟程序 GPS

GPS (General Problem Solver) 程序是由纽厄尔和西蒙 (Newell, A. and Simon, H. A.) 研制的^[5]。它既是用来探讨人工智能的一个重要工具, 又是模拟人解决问题时的思维过程的一个著

名的实例。GPS 的建立,标志着对思维心理学的研究,从传统的实验方法跨入了信息加工的领域。

(1) GPS 所依据的心理学实验情境

实验任务是解初等符号逻辑题,被试是美国工科大学生,坐在黑板前,黑板上写着下述两个符号逻辑表达式:

$$(R \supset \sim P) \cdot (\sim R \supset Q) | \sim (\sim Q \cdot P)$$

同时,给被试12条规则(见表2),要求被试应用这12条规则,把黑板上左边的表达式转换成右边的表达式。其中所包含的一些符号的意思是:“点”(·)代表“与”;“楔形”(V)代表“或”;“马蹄形”(⊃)表示“蕴涵”;“否定号”(∼)代表“非”。

表2 符号逻辑作业的规则

要求用字母(P, Q, R, ……)和连接词·(点)、V(楔形)、⊃(马蹄形)和∼(否定号)组成表达式。例如, P, ∼Q, P V Q, ∼(R ⊃ S) · ∼P; 而 ∼∼P 则始终等于 P。

有12条用来转换这些表达式的规则(各条规则中的 A, B 和 C 可以是任何表达式或子表达式):

规则1, $A \cdot B \rightarrow B \cdot A$

$A V B \rightarrow B V A$

规则2, $A \supset B \rightarrow \sim B \supset \sim A$

规则3, $A \cdot A \leftrightarrow A$

$A V A \leftrightarrow A$

规则4, $A \cdot (B \cdot C) \leftrightarrow (A \cdot B) \cdot C$

$A V (B V C) \leftrightarrow (A V B) V C$

规则5, $A V B \leftrightarrow \sim(\sim A \cdot \sim B)$

规则6, $A \supset B \leftrightarrow \sim A V B$

规则7, $A \cdot (B V C) \leftrightarrow (A \cdot B) V (A \cdot C)$
 $A V (B \cdot C) \leftrightarrow (A V B) \cdot (A V C)$

规则8, $A \cdot B \rightarrow A$ 仅用于主表达式
 $A \cdot B \rightarrow B$

规则9, $A \rightarrow A V X$ 仅用于主表达式

规则10, $\left. \begin{matrix} A \\ B \end{matrix} \right\} \rightarrow A \cdot B$ A 和 B 是两个主表达式

规则11, $\left. \begin{matrix} A \\ A \supset B \end{matrix} \right\} \rightarrow B$ A 和 $A \supset B$ 是两个主表达式

规则12, $\left. \begin{matrix} A \supset B \\ B \supset C \end{matrix} \right\} \rightarrow A \supset C$ $A \supset B$ 和 $B \supset C$ 是两个主表达式

被试对这些规则的使用经过一定的练习以后,就开始解题实验。若被试提出,把某一条规则应用于一个已写在黑板左边的表达式,而且是合适的应用,实验者就在黑板左边一栏写出新的表达式,而在右边一栏写上所用的规则。表3给出了被试解题过程的实际例子。

表3 说明被试解题过程的例子

1. $(R \supset \sim P) \cdot (\sim R \supset Q)$	$\sim(\sim Q \cdot P)$
2. $(\sim R \vee \sim P) \cdot (R \vee Q)$	规则6用于表达式1的左右两部份
3. $(\sim R \vee \sim P) \cdot (\sim R \supset Q)$	规则6用于1的左部份
4. $R \supset \sim P$	规则8用于1
5. $\sim R \vee \sim P$	规则6用于4
6. $\sim R \supset Q$	规则8用于1
7. $R \vee Q$	规则6用于6
8. $(\sim R \vee \sim P) \cdot (R \vee Q)$	规则10用于5和7
9. $P \supset \sim R$	规则2用于4
10. $\sim Q \supset R$	规则2用于6
11. $P \supset Q$	规则12用于6和9
12. $\sim P \vee Q$	规则6用于11
13. $\sim(P \cdot \sim Q)$	规则5用于12
14. $\sim(\sim Q \cdot P)$	规则1用于13
	证讫

另外,还要求被试在解题过程中“出声”思维。通过录音,再整理成书面材料,作为被试解题过程的口语素材(见表4)。

表4 被试解题过程的部分口语素材

好吧,瞧方程式的左边,首先,我们希望利用规则8消去其中的一边,看来解这道题是太复杂了。现在——不对,——不对,我不能这样做,因为我将把整个表达式中的 Q 或 P 都消去了。我不能一开始就这样做,现在我得找一个办法,消去出现在方程式左边和右边的两个括弧内的马蹄形。我看不出这种办法,对了,如果把规则6应用于方程式的两边,那时,我就会看出我是否可以应用规则7。

实验者写:2, $(\sim RV \sim P) \cdot (RVQ)$

我差不多可以用规则7了,但是,有个 R 需要一个否定号。所以,我得寻找另外的规则。我要看一看,我是否可以把 R 改变为非 R。其实,我应该把规则6只用于方程式的左边。所以,用规则6,但只用于左边。

实验者写:3, $(\sim RV \sim P) \cdot (\sim R \supset Q)$

现在,我要照规则7所表达的来应用它,两者——请原谅,因为有个马蹄形,没办法用。所以,现在我正在查看——仔细看看这里的规则去找另外一条,看看我是否能够把第二个方程中的 R 改变为非 R。但是,我看不到有任何办法做到这一点。(叹气)。我简直没有办法找到第二条规则了。

(2) GPS 的作业环境、目标和方法

主试的任务,是从获得的实验材料中找出可以解释被试解题行为的人类解决问题的模型。然后,根据该模型编制程序,使这种程序能以被试所用的方式来处理符号;像被试那样考察各个符号逻辑表达式,估量着各条规则;它的输出与被试应用规则所得的结果相符。假如这样一套程序十分接近人类被试的实际解题过程,那么,它就可以成为一种关于人类解决问题的良好理论。GPS 就是在这种基础上产生的。

GPS 的作业环境与被试一样,它的操作对象是一些逻辑表达式;它用各种算子来转换这些表达式;算子就是被试使用的12条规则。GPS 的作业目标有3种类型:把对象 A 转换成对象 B,缩小对象 A 和对象 B 之间的差别,把算子 Q 应用于对象 A。

GPS 通过提出一些子目标及实现这些子目标的办法,一步步

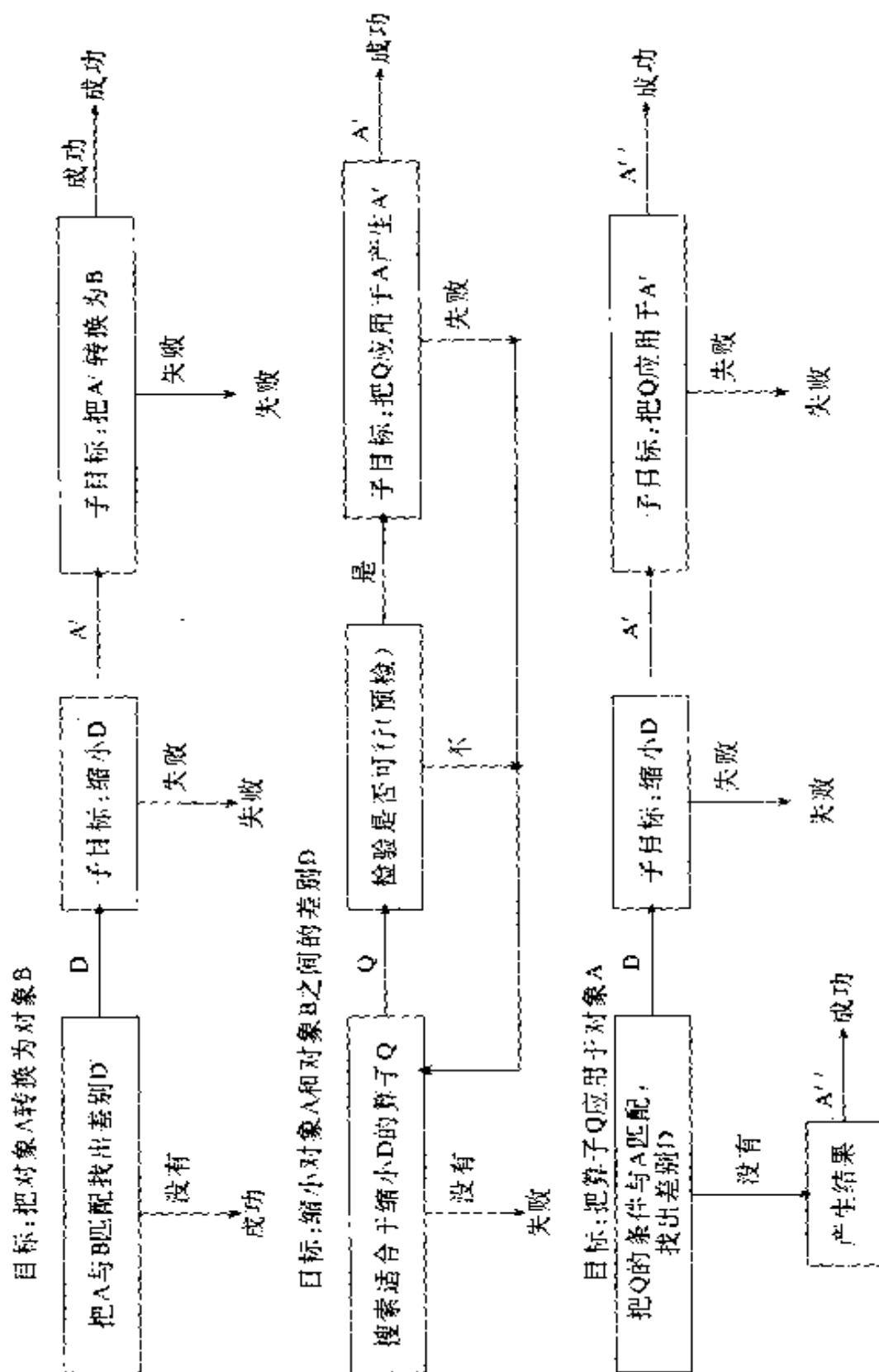


图1 GPS作业过程中采用的方法

地达到最终目标。为了做到这一点,它采用了3种主要方法,每一种方法都与上述的一类目标相联系(见图1)。具体地说,它们是:

① 若目标是把对象 A 转换成对象 B,那么,就把对象 A 和 B 进行比较,揭示出两个对象之间的差别 D。发现差别 D 后,就提出一个子目标即缩小差别 D。如果该子目标实现了,就产生一个新的对象 A',然后再建立一个子目标,把 A' 转换成 B。

② 若目标是缩小两个对象之间的差别,那么,就要找到一个合适的算子。GPS 可采用的、适合于缩小各种差别的算子,见表5。

③ 如果目标是把一个算子 Q 应用于对象 A,那么,第一步就要看看算子的条件是否得到满足。如果条件未满足,就要找出差别 D,并提出一个子目标,缩小这一差别;为了实现这个子目标,再去搜索合适的算子,并检查算子的条件与对象 A 是否匹配。

表5 适合于缩小各种差别的算子

加	项:R3, R7, R9, R10,R11,R12
删	项:R3, R7, R8,R11,R12
改变连接词:	R5, R6,R7
改变符号:	R5,
改变次级符号:	R2, R5,R6
改变组合:	R4, R7
改变位置:	R1, R2

实际上,这些方法构成了一个递归系统。GPS 在试图达到给定目标的过程中,生成一棵子目标树,通过这些子目标的实现,达到问题的最终解决。

(3) GPS 作业过程的跟踪及其与人的作业的等价性

表6给出了 GPS 对上述问题第一部分作业的跟踪。各行头部的空格,表示子目标与目标的关系。

表6 GPS 部分作业的跟踪

$L_0, \sim(\sim Q \cdot P)$	产生 $L_1: (\sim RV \sim P) \cdot (\sim R$
$L_1, (R \supset \sim P) \cdot (\sim R \supset Q)$	$\supset Q)$
目标1, 把 L_1 转换成 L_0	目标12, 把规则7用于 L_1
目标2, 从 L_1 中删去 R	目标13, 把 L_1 右部分内的连接词改为
目标3, 把规则8应用于 L_1	V
产生 $L_2: R \supset \sim P$	目标14, 把规则6用于 L_1 的右部
目标4, 把 L_2 转换成 L_0	分
目标5, 把 Q 加进 L_2	产生 $L_3: (\sim RV \sim P) \cdot$
否定	(RVQ)
目标2,	目标15, 把规则7用于 L_3
目标6, 把规则8用于 L_1	目标16, 改变 L_3 左右两部分的符
产生 $L_3: \sim R \supset Q$	号
目标7, 把 L_3 转换成 L_0	目标17, 把规则6用于 L_3 的
目标8, 把 P 加进 L_3	右部分
否定	产生 $L_4: (\sim RV \sim P) \cdot$
目标2,	$(\sim R \supset Q)$
目标9, 把规则7用于 L_1	目标18, 把规则7用于 L_4
目标10, 把 L_1 左部分内的连	目标19, 把 L_4 右部分内的连
接词改为 V	接词改为 V 否定
目标11, 把规则6用于	目标16,
L_1 的左部分	如前
	目标13,
	如前
	目标10,
	如前

这里对表6作一简要的说明。GPS 面临的问题是把 L_1 转换成 L_0 。为了解决这一问题, 它把两者进行比较, 揭示出 L_1 中有 R , 而

L_0 中没有 R 。这个发现导致形成一个“缩小差别”的子目标,即图中目标2,从 L_1 中删去 R 。为了实现这个子目标,就需搜索适当的规则,结果找到规则8。规则8有两种形式,而且都是可用的;GPS 选择了第一种形式,即 $A \cdot B \rightarrow A$ (表6中目标3)。结果产生一新的对象 $L_2: R \supset \sim P$,并形成一个新的子目标,即把 L_2 转换成 L_0 (表6中目标4)。由于规则8的应用,把表达式中的 Q 删去了,但 L_0 中是有 Q 的,所以又导致另一子目标,即把 Q 加进 L_2 中去(表6中目标5)。然而,GPS 经过估量,这个子目标被否定了,因为加一项的难度大于删一项。于是 GPS 返回到目标2,寻找另一能删项的规则;它找到规则8的第二种形式,并生成目标6。但是,经历了差不多同样的进程,最终还是放弃了,重新回到目标2。

为了解决删项的问题,GPS 准备应用规则7(表6中目标9)。它选择了 $(A \vee B) \cdot (A \vee C) \rightarrow A \vee (B \cdot C)$ 这一变式。但是它发现,子表达式中的连接词与 L_1 有差别,规则7不是直接可以应用的,必须首先要改变表达式中的连接词;于是,它找到并应用规则6,改变了 L_1 两个子表达式中的连接词(表6中目标10至目标14)。

当然,人们最关心的是,GPS 的作业过程是否基本上反映了被试作业的语言记录中的每个特征。这种语言记录是被试解题时内部思维过程的反映。作为人类解题时的一个思维过程的模拟程序,GPS 的作业过程必须反映被试语言记录的各个特征。下面看一下最初一段语言记录:

“好吧,瞧方程式的左边,首先,我们希望利用规则8消去其中的一边。”

“……不对,我不能这样做,因为我将把整个表达式中的 Q 或 P 都消去了。我不能一开始就这样做。”

这段语言记录说明,被试实际上应用了规则8,他把应用该规则而生成的表达式 L_2 与 L_0 作了比较,由于删去了最终的目标表达式中所需要的符号,因此,他又否定了规则8的应用。

“……我得找一个办法,消去出现在方程式左边和右边的2个括号内的马蹄形。我看不出这种办法,对了,如果把规则6应用于方程式的两边,那时,我就会看出我是否可以应用规则7。”

这段话与GPS作业跟踪(表6)中的目标9至目标14的作业进程是直接对应的。它清楚地表明,被试主要关心的是应用规则7,为了实现这一点,就需要改变两个子表达式中的连接词。于是,被试选择了规则6作为实现这一目标的手段。

分析这一段语言记录,只是想说明,GPS的作业进程基本上反映了被试解题过程的特征。GPS的实现,为我们建立关于人类思维的理论提供了新的有力的工具。正如西蒙指出的,这种模拟技术能清楚地向人们展示,人类的智能行为可以理解为一集规则的产物。

GPS(通用的问题解决者)的名称虽然有些夸张,但是,它借助于手段-目的分析等启发式策略,能成功地解决11种问题,除上述定理证明外,还有野人过河、河内塔等问题。它的成功在当时是非常令人振奋的。

总之,一个计算机程序,要成为人的某种认知活动的理论解释或模型,必须从内部心理过程这个层次上来进行模拟。单是程序的输出,不足以证明它具有解释人类认知活动的力量。一个人造的行为系统要用来解释人类的认知行为,那么,必须要依据人类的认知结构和认知过程来加以比较。认知心理学的任务,就是要发现那种人类外部行为之基础的认知结构和认知过程。心理学拥有的关于人的内部加工过程的材料,可以支持或说明一个程序是否是人类行为的真实的理论解释或模型。

三、心理学和人工智能在理论 发展中的相互作用

有些人认为,心理学家对设计和改善那些用来验证假设的方

法,作出了很大的努力和贡献,但是在提出假设或理论建构方面关注不够。这种评论是值得注意的。在心理学研究中,有时由于过分强调统计处理的结果及其作用,往往会忽视对过程本身作详细的说明。因此,许多心理学解释,从人工智能的角度来看,是过于简单的,往往是不可实现的。

人工智能工作者则倾向于强调一个理论应该是可实现的,就是可以用计算机程序的形式来表示的,而且,这种程序是可以成功地运行的。如果我们把人的认知系统看作一个计算系统,在物理符号系统的理论指导下研究人的认知过程,那么,对人类认知过程的任何一个好的理论解释,都应该是可以实现的。如果一个理论不能满足其实现的标准,那么,它很可能不是一种令人满意的解释。当然,要想使一种理论解释是可实现的,就需要对其过程作出非常详细的说明和规定。

但是,这里有一些具体情况。人工智能工作者可以撇开检验每一环节的要求,提出一种可实现的理论解释;其中有许多环节是为了达到程序能正常运行而暂时假定的,尚无经验材料作根据。对认知心理学来说,理论要以经验数据为基础;在缺乏经验材料的情况下,为满足可实现的要求而设想出某些环节,有时不仅不足以保证它是一种满意的认知解释,甚至可能会是错误的解释。所以,必须对这些假设的环节作进一步的经验检验。这样,才能使关于人类认知的理论解释得到进一步的改善,使模型既是可实现的,又是有经验材料为基础的。这一点,单靠人工智能是做不到的。

实验证据与认知理论的构建是联系在一起的,前者是后者的基础。我们认为,把心理学的科学实验工作与计算机实现相互结合起来,就会促进人类认知理论的发展。因此,在关于人类认知理论的构建和发展中,认知心理学和人工智能应该携起手来。

下面,我们看两个例子,以说明在理论的产生和发展过程中,人工智能和认知心理学是怎样相互交叉和相互作用的。

1. 图式和框架

第一个例子是关于图式(schema)的理论。这个理论起源于心理学,出现在英国心理学家巴特利特(F. C. Bartlett)的著作“Remembering”中^[6]。在20世纪30年代,巴特利特让被试读或听一些故事或关于事件的叙述,然后考察被试在回忆故事或事件时出现的误报和重构现象。

在他的一个最著名的研究中,他利用的一个故事,是北美印第安人的传说,叫“鬼魂的战争”(The war of the ghosts)。他让被试读这个故事,然后在各种不同的时间间隔里,检查被试对这个故事的回忆。被试在回忆这个故事时会出现什么样的系统性的记忆偏向,这是巴特利特关心的主要问题。结果发现,被试对故事内容作了“整理”,他们“忘记”的内容与已有的知识经验有关。例如,经过8天以后的回忆,出现下述情况:

① 故事简化了,一些专用名词(如“卡拉马河”)漏掉了;经验不熟悉的东西不易回忆起来。

② 表达方式成语化,如原文为“我的亲戚不知道我到哪儿去了”,在回忆中变成“家庭有牵挂”。

③ 用熟悉的词代替不熟悉的词,如用“钓鱼”代替“捕猎海豹”,用“船”代替“独木舟”等。

④ 故事的叙述更加连贯、合理,故事中的事件次序有了调整。

为了解释这类现象,巴特利特提出了“图式”这一重要概念。他认为,图式是对先前反应或先前经验的一种积极的组织,它是在任何适应良好的、有组织的反应中运转的;每当任何有序的、有规则的行为发生时,一个特定的反应之所以会出现,只因为它与其它有关的、已经系统地组织起来的反应联系在一起了;它们不是个别地而是作为一个整体运转的。也就是说,图式是过去获得的知识经验的一种抽象的表达,它可以对获得的材料进行重构和改造,并使人

的行为受先前经验的影响。

在巴特利特提出关于图式这个概念后的40多年时间里,人们对图式这个理论几乎没有什么研究。直到20世纪70年代,人们对图式的概念才开始注意和重视起来。为什么在70年代图式的理论在认知研究中突然受到人们的重视呢?当然,这一点很难确切地说清楚,但是,可以肯定,这与人工智能研究者对图式理论的兴趣有关。1975年,明斯基(Minsky, M.)写了一篇关于人工智能中用框架表达知识的文章^[7]。这篇文章很有影响。在这篇文章中,明斯基把知识的框架表达与巴特利特的图式理论联系起来了,文中多次摘引了巴特利特的论述。事实上,框架的思想与图式理论基本上是一回事,不过,巴特利特当初提出的图式概念尚较为抽象,而框架则把图式概念进一步具体化、形式化了。所以,明斯基亦明确地说,框架的思想不是新的,在巴特利特的图式理论中已经有了,但是,“框架系统”这个思想是新的。

知识是人类对客观事物及其规律的认识。客观事物可用其具有的属性或所处的状态以及它与其它事物间的关系来描述;而客观规律则反映了各种事物之间的动态转换、继承或因果联系等。这两者都可用框架来表示。具体地说,框架由3个部分组成,它们是:框架名、槽和槽的值(表7)。框架是一种知识结构,它用来表达某种固定的情境;当一个人碰到某种情境时,就从记忆中选取相应的框架,并依据框架所提供的知识来理解情境,作出某种预期或反应。

当然,巴特利特和明斯基探讨图式(或框架)理论的目的是不同的。巴特利特作为一个心理学家,他的目的是探求一种理论,用来解释人的记忆现象;而明斯基所关心的是:一个智能系统需要什么样的知识,有关的知识应该怎样相互联系并组织在一起,一个计算系统怎样才能有效地利用这些知识。

香克(R. C. Schank)和阿培森(R. P. Abelson),一位是计算机科学家,一位是心理学家,他们共同从事的研究和提出的“脚本”

表7 知识框架的基本形式

框架名:椅子		框架名:诊断规则		框架名:结论	
槽名	槽值	槽名	槽值	槽名	槽值
类别	家具	前提1	咳嗽	病名	感冒
腿	整数(缺省=4)	前提2	发烧	疗法	APC,1日3次
背的样式	软、硬	前提3	流涕	预后	良好
扶手数量	0,1或2	推论	结论		
		可信度	0.8		

(script)概念,可以看作是图式或框架理论的一个很有名的具体例证。人们即使为了理解一个很简单故事,也总是需要利用已有的知识去进行预期和推论,去填满故事叙述中缺省的或隐含的信息。比如说,我们读了“李四走进餐馆;服务员送来菜单”的叙述。为了理解这2个句子,我们就需要许多关于餐馆的知识。我们需要知道服务员所扮演的角色;菜单的作用是告诉顾客餐馆里已准备的饭菜,顾客可根据需要来购用等等。如果我们没有这些知识,那么,就不可能理解这种叙述。这种知识是有组织、有结构的。香克和阿培森把这种知识结构称之为“脚本”。从本质上看,脚本是先前形成的关于社会生活事件或活动的知识包,这些事件或活动都是定型化了的、有序的,而且是受文化约束的,如去医院看病,乘火车去旅行,到电影院去看电影等。

脚本是一种高级的知识结构;嵌入计算机程序中的脚本,在机器加工故事时,可以提供理解故事所需要的背景知识。脚本的结构是一张事件表(表8),所以,脚本的知识结构实质上是一集关于事件系列的预见,在故事加工中用于指导推理过程。香克和阿培森曾写了一本书,描述一个语言理解程序 SAM (Script Applier Mechanism)^[8]。这个程序就是应用脚本这一知识表征手段来使计算机理解故事的。程序根据输入故事的初步内容调用相应的脚本,并把故

事叙述的事件与脚本中包含的知识进行匹配,以实现故事内容的识别并推导出故事未明确叙述的有关事件。例如,输入下面一段故事:

“李四走进餐馆。他要了一份牛排。李四付了钱,离开了餐馆。”
然后问:“李四吃了什么?”

表8 餐馆脚本包含的知识

脚本:餐馆	
人物:顾客、服务员、出纳员……	
道具:餐馆、餐桌、菜单、食物、饭钱……	
事件:	
1. 顾客走进餐馆;	2. 顾客在餐桌前坐下;
3. 服务员拿来菜单;	4. 顾客点食物;
5. 服务员端来食物;	6. 顾客吃食物;
7. 顾客付饭钱;	8. 顾客离开餐馆。

在这段故事中,实际上并没有叙说李四吃了什么。回答这个问题,不仅需要具有在餐馆里进行通常活动的知识。而且还要具有把这些知识应用于特定情境的能力。脚本提供了这些东西,可用来进行常识推理。其推理机制大致是这样的:根据脚本中的事件1和事件4,程序认出了故事第一、二句话,并把李四指定为“顾客”,而把牛排指定为“食物”。然后,根据脚本中的事件7和事件8,程序认出了故事中的后一句话。这样,事件中的顾客就是李四,食物就是牛排。所以,程序就可以把脚本中的事件6具体化为“李四吃了牛排”,并作为对问题的回答。

当然,在应用脚本机制来理解故事时,首先要将输入句转换成概念依赖表达(conceptual dependency representation)。这里,限

于篇幅,我们就不去讨论 CD 表达本身了。但需要指出,香克等人的工作,在许多心理学家中引起很大兴趣。主要原因是:①这种加工过程的细节,为解释心理现象提供了某种可能性;②从人工智能角度提出的这一理论,为人们提供了一种考察图式的方式,也提供了一种如何从记忆中存取和利用这些图式的描述。这在当时是某种新的东西,它为心理学家研究人的知识结构提供了一种手段。

脚本的期待作用,会影响故事理解中所涉及到的某些过程。有些心理学家通过阅读时间的实验,对这种预测进行了检验。这种研究是以下述两个简单的加工假设为依据的:①被试对期待的句子的加工时间较少,而对未期待的句子的加工时间较多;或者,一个词出现在有关的脚本中,被试对它的再认较快,若出现在无关的脚本中,被试对它的再认则较慢。②阅读的时间随句子加工或词汇决策的难度呈函数式地增长。实验证实了这种假设。例如,一个句子,如果出现在有关脚本材料的后面,那么,就会比出现在其它课文后面时阅读的速度要快。

鲍尔(G. H. Bower)等人的研究还进一步表明^[9],当人们阅读一个去牙科医院的故事和一个去普通医院看病的故事后,在回忆或再认时,究竟哪个事件发生在哪个故事中,常常会发生混淆。他们发现,这种混淆现象往往发生在不同脚本中的类似情景(事件)之间。心理学家的研究和发现,又反馈到人工智能研究方面。鲍尔等人的发现,促使香克重新考虑关于脚本的概念,并导致这个理论的进一步完善化。从而,香克在他的著作《动态记忆》中,论述了“动态记忆”和“记忆组织包”的理论。

人类的记忆不是一种静态的结构,而是动态的。那么,什么是动态记忆呢?动态记忆是说,人的记忆是一种灵活的、可变的系统。它的特点表现为能够调整它对世界的原初的编码,反映出对世界的新的理解;它允许我们从自身的经验中学习,根据新经验的特点改变它自身的组织。比如,一个图书馆的记忆就不是动态的,它改

变起来很困难,而且,它的改变一定需要外界的干预。专家的记忆就不同了,当他的兴趣发生改变时,或者当他的科学知识发生变化时,他能够很容易地改变他自己的内部分类系统。一个专家是一个具有自我意识的人,他知道他知道什么,并能够观察他所知道的东西。因此,如果需要的话,他可以改变自己的知识的分类编码,也就是改变他自身的记忆结构的细节。所以说,他有一种动态的记忆。

按照香克原来的脚本理论,在人的记忆中有各种各样的脚本,它们的结构成分是固定的。拿付钱事件来说,一个餐馆脚本中有一个固定的付钱事件;商店购物的脚本中也有一个;修理汽车的脚本中又有一个。鲍尔等人的研究实际上对此提出了异议。因此,香克对自己的理论作了修改,提出在记忆中只有一个付钱的事件,所有这些脚本都共用这个付钱事件。在这个新的理论中,香克认为,我们记忆中所存储的不是一个个静态的脚本,而是许多一般性的事件的描述,它们可动态地聚合成一种高级的知识结构,叫作“记忆组织包”(memory organization package,简称 MOP)。一个 MOP 是由一集有序的事件描述组成的,这些事件描述指向于达到一个共同的目标(如“看病”)。但是,MOP 本身并不明确地包含某种实在东西的记忆,它只是提供了如何把有关的各事件的描述联系起来的索引。换言之,一个 MOP 只是包含有关的各事件描述的索引,一旦进入工作状态,它就根据索引把各有关的事件描述组织起来,形成一个类如表8所示那样的知识结构。所以,只要从记忆中提取到一个合适的 MOP,人们就可以对故事的发展作出预期。从这里我们可以看到,一个事件描述可以由多个 MOP 共享。

由此可见,调用一个 MOP 的最终产物,类似于某种他们以前称之为脚本的东西。但 MOP 具有与脚本概念不同的新的性质。这个新的理论,吸取了鲍尔等人的心理学研究中发现的关于记忆混淆现象所提出的思想。

可见,一个理论,它起源于心理学,经过多年的潜伏期,未被人

们注意和利用,然后,却由人工智能研究重新赋予了活力。人工智能的研究又推动了心理学家去作进一步的实验研究,其结果又反过来促使人工智能理论的进一步完善化。在这里,我们清楚地看到,一个科学理论在两个学科之间的相互影响和相互合作之下,逐步提高和完善化的过程。

2. 平行扩散激活网络

第二个具体的例子是平行扩散激活网络理论的形成。

(1) 层次网络模型

1969年,奎林(M. R. Quillian)研究出 TLC (Teachable Language Comprehender) 程序^[10]。奎林的工作是第一次尝试用语义网络来模拟人的长时记忆。这个模型的直接意图是想模拟人的自然语言的理解过程。这里,我们介绍一下该系统的主要成分,系统的记忆网络和提取信息的方式。

在 TLC 系统中,记忆网络中的信息是由3种类型的数据组成的。

① 一些代表现实世界中各种客体的概念,如鱼、鸟、动物等,并用节点来表示。

② 每个概念有它的属性,如鱼有鳍,鸟有羽毛,动物有皮肤等。

③ 各概念之间有联想链把它们彼此联结起来构成一个记忆网络,并使这种网络形成一种有层次的结构。例如,“动物”范畴,在它下面有若干条链,把它与“鸟”、“哺乳动物”等联结起来;同时,从“鸟”下面又伸出若干条链,把“鸟”与“麻雀”等联结起来。

为了使系统的存储空间节约一些,奎林采用了一条原则,即一个特性在网络中只存储一次。例如,“有皮肤”这个特性,只是在“动物”这个概念节点上存储一次。在这种情况下,系统确认在“动物”节点下属的各层次上的所有概念节点,都具有“皮肤”这个特性;同

样,“有翅膀”这个特性,也不是在网络中每一种鸟的概念节点上都存储的;这个特性仅仅存储在它们上属的“鸟”概念节点上,系统确认,每一下属的具体的鸟概念节点,均继承这一特性。这种下属概念继承上属概念特性的安排,称之为“强认知经济原则”。

从心理学的观点看,TLC模型的一个最重要的方面,是它对输入句进行加工的方式。如果问它“麻雀是动物吗?”这样一个问题,它将从“麻雀”与“动物”两个概念节点出发,开始双向搜索。如果两个搜索通道相交,那么,系统对该问题的回答,原则上是肯定的,否则,回答就是否定的。

奎林从人工智能研究的需要出发,提出了一个称之为语义网络的记忆模型。认知心理学家对这个模型非常感兴趣,因为它对人的句子加工提供了一些非常精确的预测。这些预测是以系统的两个基本参数为依据的:①搜索过程实现相交所需的时间;②既然系统的记忆网络是一种层次结构的组织形式,而且是按照认知经济的原则进行安排和运转的,所以,在两个概念节点之间,相隔的链愈多,搜索过程所需的时间愈长;相隔的链愈少,搜索过程所需的时间愈短。这样,人验证句子“麻雀有翅膀”就比验证“鸟有翅膀”需较长的时间。

柯林斯(A. M. Collins)和奎林同时还做了一个心理学实验^[11]。实验发现了层次结构和认知经济原则的证据,也就是说,实验的结果证实了模型的预测。实验数据表明,若两个概念节点之间相隔一条链,搜索所需的时间大约是75毫秒。一个从人工智能研究的需要提出的关于记忆的理论,具有很强的心理学现实性,这在当时是非常令人振奋的。但是,没过多久,其他的一些心理学实验发现,这个理论是有缺陷的。这里,我们简要地介绍一下有关的一些心理学实验研究。

威尔金斯(A. J. Wilkins)^[12]考察了概念范畴之间的联系,计算了句子(如“鸡是鸟”和“麻雀是鸟”)中主语和述语的结合频率。

所谓结合频率,是指两个词(如“鸡”和“鸟”)在书面语言中一起出现的频率。结果发现,结合频率可以更好地预测句子验证所需的时间。也就是说,句子中两个概念的结合频率高,搜索过程所需的时间较少,该句子的验证较快;句中两个概念的结合频率低,该句子的验证就较慢。接着,康拉德(C. Conrad)也发现^[13],被试对结合频率高的句子,验证的速度较快;对概念结合频率低的句子,验证的速度较慢。后来,李普斯(L. J. Rips)等人进一步发现^[14],句子验证的时间,在同一层次的概念范畴之间亦有很大差别。被试验证句子“猪是哺乳动物”所需的时间较长,而验证句子“母牛是哺乳动物”所需的时间较短。还有一些其他的例子,说明实验中所发现的现象与模型的预测是相反的。例如,人们验证某个东西是哺乳动物,比验证某个东西是动物往往需要较长的时间。但是,根据层次网络模型的预测,验证前一句要比验证后一句快,因为前一句两个概念在网络中相隔的距离较近。

(2) 扩散激活模型

层次网络模型受到严重的挑战。1975年,柯林斯和洛夫特斯(E. F. Loftus)在进一步的实验和对已有的实验结果作了全面考察的基础上,提出了人类记忆的扩散激活模型^[15]。这个模型保留了人类记忆中知识存储的方式是一种网络结构这一重要的观念,但是,对层次网络模型作了许多重要的改进。其中包括:

① 丢掉了呆板的层次结构。扩散激活模型是一种相互高度联结的网络结构;网络中的节点由各种链联结在一起,犹如许多公路、铁路和航线把世界上大小城市联结起来那样。扩散激活模型是根据语义类似性的线索组织起来的。两个概念共有的特性愈多,概念之间的联结也愈紧密。网络的结构不是根据概念之间的逻辑关系来安排的,因此,网络不再具有那种层次结构的形式。

② 修改了它的强认知经济原则。也就是说,在概念之间,尽管有些特性是继承的,但是,对某些概念来说,其特性在记忆网络中

确实是重复储存的。这一原则,称之为“弱的认知经济原则”。

③ 以扩散激活的机制代替了直接的搜索过程。当一个概念被加工时,相应于该概念的节点就进入激活状态,并且,激活沿着从该起始节点伸出的链向外扩散,去激活那些与它相联的节点;这些被起始节点激活的节点,又可平行地去激活那些与其相联的其它节点。这样,激活从起始节点向外扩散,形成一个激活圈。当遇到来自另一起始节点的激活圈时,就实现相交;当激活相交点上的总量达到某个阈限时,就会触发一个合理的反应。

④ 提出了链的强度假设。在扩散激活模型中,联结各节点之间的链的“可达性”是不同的。链的可达性越大,激活传送的速度越快,而且,在传送过程中激活减弱的速度就慢。这个关于链的强度假设,说明了概念的结合频率效应。也就是说,概念的结合频率愈高,联结它们的链的可达性就愈大;在结合频率较低的概念之间,链的可达性就较小。激活通过可达性大的链传送时,速度快,因而验证句子所需的时间就少。

从这个实际的例子中我们再一次看到,人工智能和心理学研究之间的相互交叉和相互作用,如何导致对心理现象作出越来越完善的、在计算上愈趋精确的解释。

这里,还有一点也是应该说一下的。我们知道,上述两个实际例子是从不同的侧面来描述人类记忆的。近年来,心理学家沙基(N. E. Sharkey)一直努力试图把这两种不同的途径综合起来,构成一个单一的、在心理学上看起来是合理的计算模型。这就是他提出的KAN(Parallel Spreading Activation Knowledge Access Network 的简称)模型^[16]。这个模型的主要特点是把“脚本”的思想与“扩散激活的联想网络”结合起来了。

KAN模型的核心是扩散激活网络,脚本被表征为该网络中一集相互联系的概念。或者说,一个脚本可以表征为扩散激活网络中的一个区域。在该区域内,有些概念节点(如“吃”、“食物”、“盘

子”、“桌子”等)是彼此直接相联的,而另一些则可能是通过一个中介节点而联接起来的。从本质上说,这个网络区域受一个脚本子节点所制导和支配。该子节点代表许多概念节点之间的一个联接站,通过它,这些概念节点集合在一起构成一个脚本(见图2)。按照KAN模型的假设,当来自故事的信息输入时,网络某个区域的有关节点就被激活,同时,激活开始向邻近的、与它们相联的节点传播。一个节点从它邻近节点接收的激活量,依赖于它们之间联结的强度。

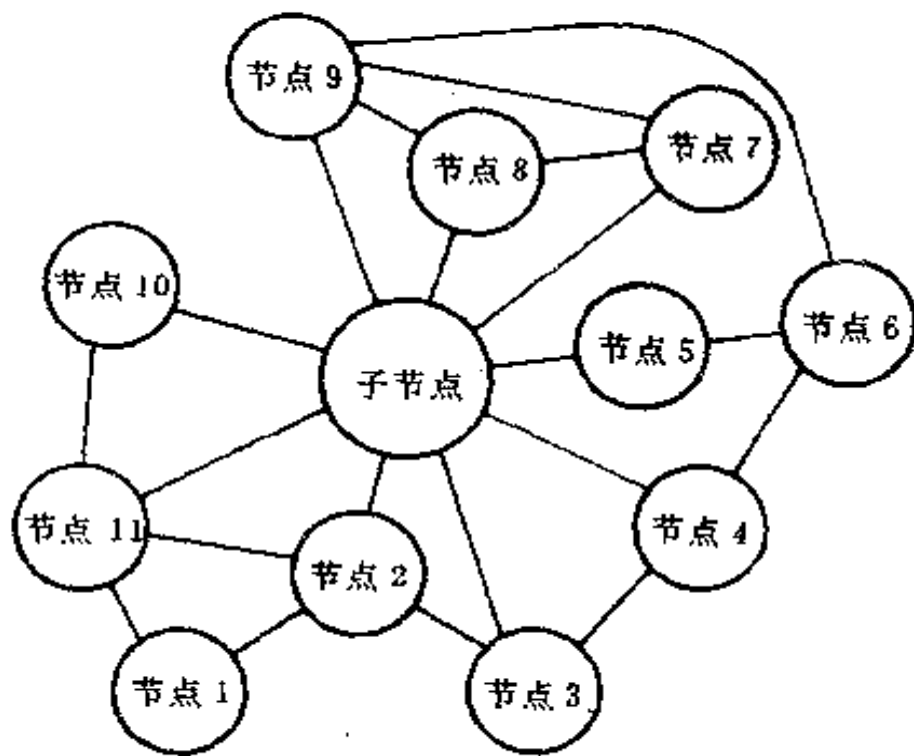


图2 相应于一个脚本的网络区域

当激活通过网络扩散时,激活就在有关的脚本子节点上累加起来,直至其中一个脚本子节点达到阈限。这时,受这个子节点所支配的网络区域(脚本)就处于活动的焦点中。系统的加工活动就集中在知识库的这个区域内进行。

KAN模型的实质,是要把“脚本”的思想与“扩散激活的联想网络”思想结合起来;它对人记忆中知识的结构及其怎样利用的问

题,为我们提供了一种在心理学上可能是合理的解释和预测,也为我们将来的经验性研究提供了有价值的参考。

总之,我们有理由认为,尽管心理学和人工智能在认知科学的发展中各自起着重要的作用,但是,它们之间的结合,将会给予人类认知的研究更大的推动。

一方面,人工智能研究者可以从心理学研究的经验材料中获取有益的启发,根据心理学研究提供的证据,来构造或检验自己的系统,从而使自己的系统更具有人类智能的特色。从长远来看,这会使人工智能研究得益非浅。实际上,人工智能要真正成为认知科学的一个部分,那么,以严格加以控制的心理学实验的经验材料作为自己的基础,这是必需的。

另一方面,心理学家借助于人工智能的方法和手段,提供的关于人类认知的模型或理论的形式,就不再限于一种言语描述。况且,这种言语描述,从人工智能的观点来看,是不怎么精确的。实质上,如果心理学家把他们的理论实现为一个计算机程序,那么,就会使他们的模型或理论更加确切和精细。

还有,认知心理学和人工智能研究者的携手合作,必然会推动关于人类认知理论的深入发展,不断地加深我们对人类认知活动的认识。

主要参考文献

- [1] Minsky, M., 《Semantic Information Processing》1968.
- [2] P. H. 温斯顿;《人工智能》科学出版社,1987年。
- [3] Simon, H. A., “认知科学的一些最新进展”,《心理学报》1991年第二期。
- [4] Weizenbaum, J., “ELIZA — A computer program for the study of natural communication between man and machine”, *Communication ACM*, 1965, 9, 36--45.
- [5] Newell, A. and Simon, H. A., “GPS, 一个模拟人类思维的程序”,《心理学动态》,1983年,第1期(陈永明编译)。

- [6] Bartlett, F. C. ,《Remembering》1932.
- [7] Minsky, M. ,“A framework for representing knowledge”, in《Cognitive Science——Perspective From Psychology and Artificial Intelligence》(edited by A. Collins and E. E. Smith, 1988).
- [8] Schank, R. C. and Abelson, R. P. ,《Scripts, Plans, Goals and Understanding》1977.
- [9] Bower, G. H. and et al. ,“Scripts in memory for text”, *Cognitive Psychology*, 1979, **11**, 177—220.
- [10] Quillian, M. R. ,“The teachable language comprehender”, *commun. ACM*, 1969, **12**, 458—476.
- [11] Collins, A. M. and Quillian, M. R. ,“Retrieval time from semantic memory”, *J. Verb. Learn. Verb. Behav.* , 1969, **8**, 240—247.
- [12] Willkins, A. J. ,“Conjoint frequency, category size and categorization time”, *J. Verb. Learn. Verb. Behav.* , 1971, **10**, 382—385.
- [13] Conrad, C. ,“Cognitive economy in semantic memory”, *J. Exp. Psychol.* , 1972, **92**, 149—154.
- [14] Rips, L. J. and et al. ,“Semantic distans and the verification of semantic relation”, *J. Verb. Learn. Verb. Behave.* , 1973, **12**, 1—20.
- [15] Collins, A. M. and Loftus, E. F. ,“A spreading activation theory of semantic processing”, *Psychological Review*, 1975, Vol. **82**
- [16] Sharkey, N. E. ,“A model of knowledge-based expectations in text comprehension”, in《Knowledge Structures》(Edited by J. A. Galambos et al. , 1986).

学习和记忆的神经基础

邵 郊

(北京大学心理学系)

一、学习和记忆的脑解剖观

1. 概言

人们都知道经验和阅历能改变人的行为和认识。如所谓“吃一堑，长一智”，或曰知过必改。然而如从生理心理学的角度来看，经验改变的是脑内神经元的生化和生理反应的特性，以及因之产生的它们之间的功能联系和解剖结构的生长变化。行为的改变乃是这些变化的结果。

人类有意识的学习就是主动地去追求经验，来促使自己的脑细胞发生这些变化。学习的成功需要记忆，记忆就是脑内神经元网络中那些被经验改变了的特性的保持。

心理学中所谓记忆痕迹(engram)，或记忆的贮存和提取等概念，只不过是一种比喻的说法。实际上在脑内是找不到任何事物的痕迹和存贮地点的。因为：①任何一件事情的经验都包含着多种心理属性的表征，如视觉的、听觉的、空间和时间知觉的、其他感觉的和感情的。这些心理表征成分所涉及的功能类型、数量、解剖部位和它们的相互联系都是极庞杂而且是广泛分布的；②个别的和成群的神元可能接受多种经验的影响，即参与多种事件的记忆；③在不同的经验影响下神元未必有事件特异性的物质改变。目前，与其把任何一件事情的记忆痕迹想象为某一处和某一

群神经元的具体的物质变化,不如设想为脑的许多部分受过同一种影响的神经元的动力模式的改变。这种动力模式,如在空间和时间上有特别组织方式的活动,一旦形成可以潜伏下来,当经验过的事物再现或遇到与之有联系的事件时即可恢复活动,即记忆的唤起。同时还须考虑到每次参与回忆的神经元未必都是完全相同的一群。因此每次回想一件事的完全性和确切性都不一样。然而,探索与各种学习任务有关的神经网络所涉及的脑区和神经元的一些基本的物质变化仍然是研究学习和记忆的重要课题。

2. 学习任务和记忆的多维属性及其神经概观

人们的任何一种经验或学习任务都包含着多种的心理内容。例如,我们在某时、某处、某种场合下认识了一个穿着考究、举止不凡的人,并且与他(或她)握过手,谈过话,留下了较深的印象。这一经验的内容有时间的和空间的成分、视觉和听觉的成分、好或恶的成分和动作的成分。R. P. Kesner 和 B. V. DiMatta (1987)^[1]把这些心理内容归纳为5种心理属性,即时间、空间、感知觉、感情和反应的属性。并提出有关这5种心理属性的记忆的神经结构,即所谓记忆的属性模型及其神经概观(Attribute model of memory and its neural aspects)。其观点首先肯定有两种记忆系统分别由不同的神经结构调处。一种是属于资料基础的记忆(data-based memory)系统。例如,对于每天处理的一些临时性的事物的记忆。在这一系统中,海马(hippocampus)(图1)调处关于外界事物之间的暂时关系的记忆;杏仁核(amygdala)调处内部事态的记忆;尾核(caudate nucleus)调处自我中心的空间定位的记忆;一级感觉皮质、运动皮质、脑干和小脑调处刺激和反应之间的联系的记忆(S-R 系统)。另一种是属于期望基础的记忆(expectancy-based memory)系统。例如关于习得的某种行为法则或认识的规律的记忆。在这种系统中,他提出前额皮质的背外侧或内侧部分(dorsolateral or

medial prefrontal cortex)调处时间和自我中心的空间认知图;眶额皮质(orbito-frontal cortex)调处感情和自主反应的认知图;二级感觉联络皮质(secondary sensory association cortex)调处感知觉认知图;前运动和补充运动皮质(premotor and supplementary motor cortex)调处躯体反应的认知图。

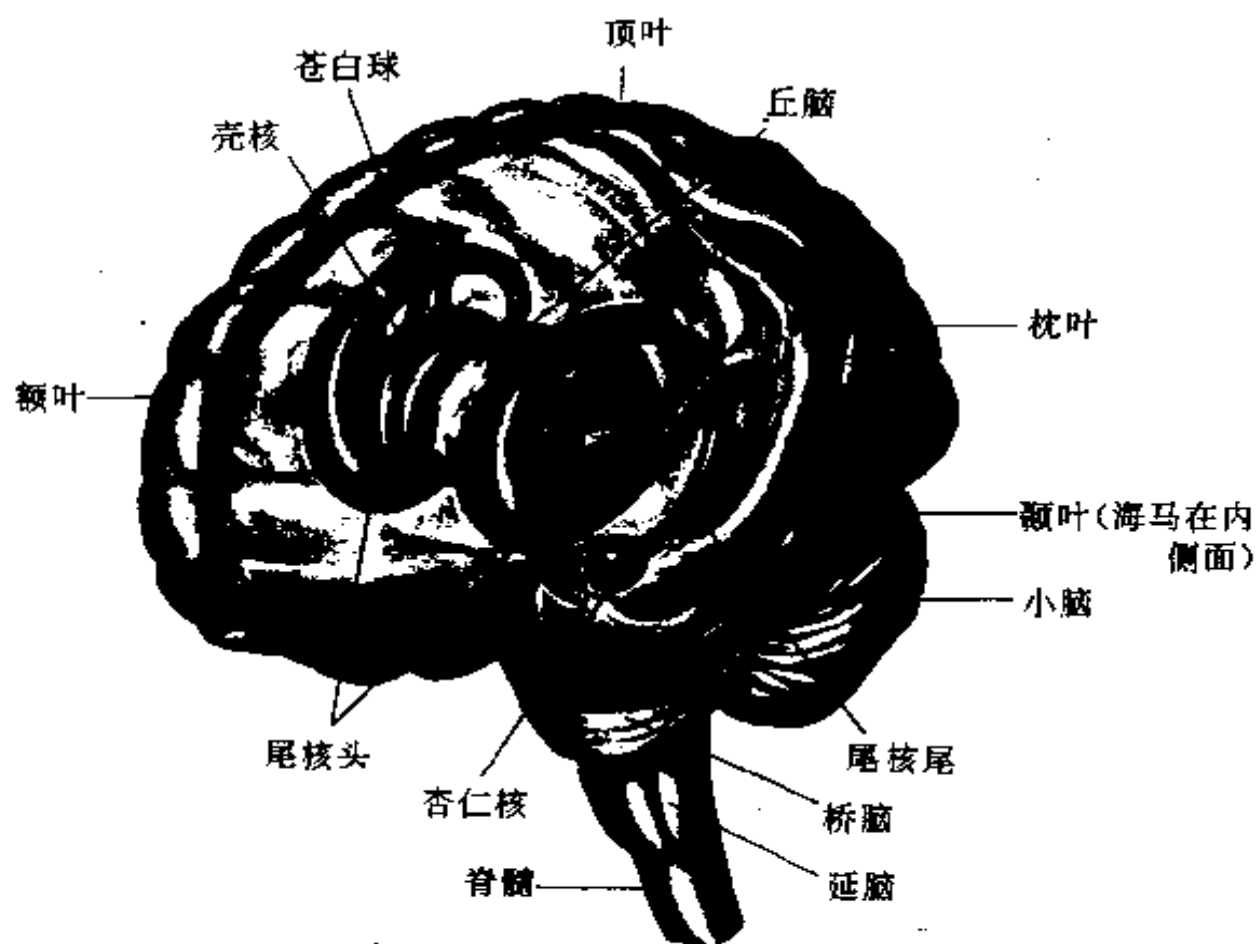


图1 大脑两半球透视观

此外,背外侧或内侧的前额皮质和后顶皮质 (posterior parietal cortex)的交互作用可认为是认知的脚本原稿的神经基质;后顶皮质和二级感觉联络皮质则可能起一种图解的作用;背外侧或内侧前额皮质和眶额皮质共同调处心情的状况(mood);而二级感觉联络皮质和前运动及补充运动皮质也是调处技巧行为的神

经基质。

最后还假定为各种心理属性服务的神经部位和形成认知图的神经部位之间虽然有许多的交互作用,但彼此都可以独立地工作。

下面分述有关两种记忆系统的各种属性记忆的神经基础的实验研究。

3. 资料基础的记忆系统及有关各种属性记忆的神经结构

(1) 海马和外界事物的记忆

按照记忆的属性模型的假设来说,海马编码和记取的主要是关于外部事物的情况的信息,特别是构成时间和空间属性的信息,这类信息也受感知觉和躯体反应属性的信息的某种调节性质的影响。此外,海马的工作完全属于资料基础的记忆系统。

海马编码空间属性的证据来源很多。对大脑右半球颞叶皮质和部分海马损毁的病人的测验,发现他们很难回忆刚刚看过的一个盘子上的玩具的位置。海马损伤的部分越大,此种缺陷就越严重(M. L. Smith and B. Milner, 1981)。在 Korsakoff 病的患者中也发现他们在记忆房间内的物体的空间位置上有类似的困难(W. Hirst and B. T. Volpe, 1984);然而,这种病人也有海马以外的脑损伤。

动物实验的证据有:切除了两侧海马的猴子对物体的空间位置的记忆也有障碍(J. K. Parkinson and M. Mishkin, 1982);海马损毁的猴子能学会从两个相同的物体中选择左边的或右边的一个,但不能改变已习得的位置选择的习惯,即不能完成位置的颠倒学习,然而能完成物体的颠倒学习(B. Jones and M. Mishkin, 1972)。

大鼠的实验可进一步说明海马在空间信息的记忆中所具有的重要作用。实验的第一步是训练大鼠在一放射形的8臂迷津(图2)中任意进入任何一臂得到食物(麦圈,强化),10秒钟之后将大鼠从迷津中取出,隔1分钟或2小时之后将其重新放入迷津中,测验它是

否还记得进入前次得到食物的那个臂。这种实验也叫作延迟的空间配样本任务,主要的是测验大鼠在一系列的空间位置中再认前次选择的某一位置的能力。所以在延迟的测验中正确的作业应当是再进入前次在其中得到食物的那一个臂,即正确的作业是有记忆的表现。经过长久训练的大鼠能在1分钟、甚至2小时的延迟之后仍记得进入前次得到食物的那个臂(在多次重复测验中只偶尔出错)。在此种测验完成之后,给受试大鼠做小范围的背侧海马损毁手术。手术后以1分钟和2小时的延迟时间分别地重新测验,结果发现背海马部分损毁的大鼠在1分钟的延迟之后其测验的成绩与手术前的一样;但在2小时的延迟之后测验的成绩大为下降,表现了记忆的缺陷(R. P. Kesner and B. V. DiMattia, 1984)。类似的实验证明较大范围的海马损毁,或与海马有双向联系的隔区损毁之后,即使在1分钟的延迟,甚至在没有延迟的测验中也表现出严重的空间位置记忆的缺陷(M. A. Hunt et al., 1986)。

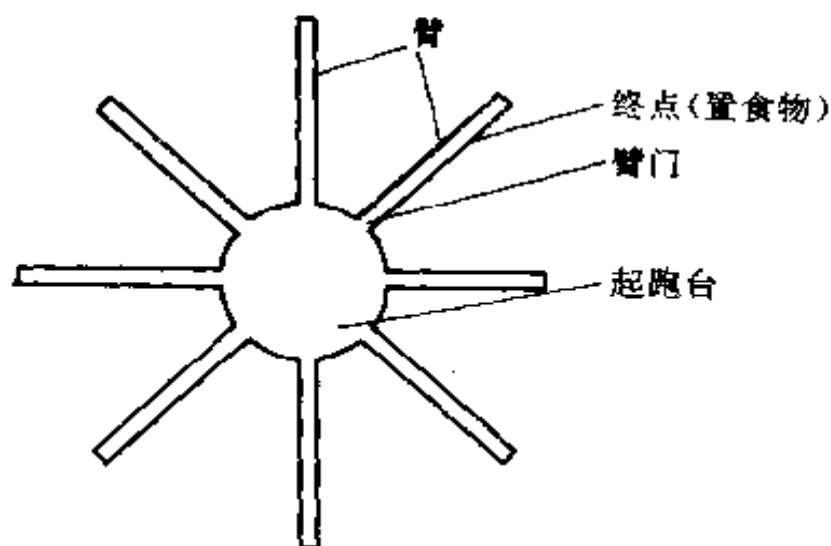


图2 放射型8臂迷津平面示意图

利用放射形8臂迷津测验大鼠对空间位置的记忆能力,并观察海马或与海马有密切的功能关系的隔区的损毁对此种属性的记忆能力的影响,也有多种设计各异的实验。结果都证明在需要记忆有

关地点或物体的空间位置的资料性质的信息时,海马和隔区,特别是内侧隔区(medial septum),起着极为重要的作用。

与此同时,神经生理学的实验也在大鼠和猴子的海马中找到了对地点-空间的特定位置有区别反应(如某些地点引起放电率增加)的神经元,称为地点细胞(place cells)(J. O'Keefe, 1979; E. T. Rolls et al., 1985)。

此外,在需要暂时记忆(资料基础的)曾接触过的事物的先后次序(时间顺序)时,海马也起重要作用,即所谓海马具有编记时间密码的功能。例如,克萨考夫(Korsakoff)病患者不能再记忆刚学习的一些词的先后次序(W. Hirst and B. T. Volpe, 1984),海马损毁的猴子不能学会按照规定的先后次序反应两个同时出现的刺激(D. P. Kimble and K. H. Pribram, 1963)。

用大鼠做的实验也得到了证明。实验的第一步是训练大鼠在 8 臂迷津的各个臂中取食。经过多次的训练之后,让每只大鼠在每天一次的练习中进入当天按随机规定的先后次序开门的臂中取食。当大鼠从 8 条臂的最后一条臂中取得食物之后立刻(约在 90 秒之内)开始下一步的测验。每天只测验一次。测验时,同时打开 2 个臂的门。随机地有时开当天练习时的第 1 和第 2 两门;有时开第 4 和第 5 两门,或第 7 和第 8 两门。在这种测验性的练习中,让大鼠学习必须进入按照当天练习时的次序先进去的一臂(例如,开第 1 和第 2 门时进第 1 门,第 4 和第 5 门时进第 4 门)才能得到食物(强化)。这种测验性质的学习结果证明所有受试的大鼠对第 1 和第 2 门,以及第 7 和第 8 门的先后次序都会有显著的记忆。但对排在中间的门(如第 3, 4, 5, 6 门)的先后次序的记忆却很难(R. P. Kesner, et al., 1984)。背海马做了小范围的损毁手术的大鼠则只能记忆最后两个门(即第 7 和第 8 门)的先后次序。而不能记住最先的两个门(即第 1 和第 2 门)的先后次序,如在第一步练习之后延迟 10 分钟再做这种测验,背海马损毁的大鼠都不能记得进入所

有各门的先后次序。表明它们对先后的时间信息的记忆能力有了严重的缺陷。另外的研究还发现因内侧隔区的损伤造成的背海马中小量乙酰胆碱的匮乏(乙酰酯酶-AchE 量减少)也产生类似的时间属性的记忆缺陷。背侧或腹侧海马的大范围的损伤,或大范围的隔区损伤造成的背海马内大量乙酰胆碱酯酶的减少对时间属性的记忆能力的影响更大。受此种损伤的大鼠即使在无延迟的测验中也不能记得任何两门的先后次序(R. P. Kesner, et al., 1986)。

上述的各种观察和实验的结果有力地证明海马,可能还有与之有密切联系的内侧隔区,调处资料基础的记忆系统内的空间和时间属性的信息。海马的损伤范围越大,可利用的空间和时间信息的保持时间越短。

(2) 海马损毁后未受影响的能力

人们还可以设想在某些学习任务中,空间和时间关系不是学习的主要线索,海马损毁对于完成这类学习任务应无影响,或影响甚微。

这类学习任务的一个典型例子是味觉厌恶的习得。在这种学习中,要编码和记住的主要线索是味觉和随之而来的感情和体内生理状况的变化。例如,大鼠在喝了有某种甜味的水(如糖精水)后接受 X 射线照射,并产生了病感或呕吐,此后即不再喝甜味的水(J. Garcia, et al., 1966)。在这种学习中,主要的线索是水的滋味和体内的不适感。实验证明海马损伤的大鼠在习得这种味觉厌恶的行为上无异于正常大鼠(P. J. Best and J. Jr. Orr, 1973)。又如,上面提到过的海马损毁后的猴子仍能学习物体的辨别,并且也能完成物体辨别的颠倒学习,即学习将以前选择的正刺激变为负刺激的任务(B. Jones 和 M. Mishkin, 1972)。

因为设想海马的工作属于资料基础的记忆系统,所以可以推测在利用期望基础的记忆系统的学习中,海马损毁也不会产生严重影响。这类的学习任务有技巧的学习、某种反应的学习、以自身

为中心的左右辨别的学习和其他事物的辨别学习。在这些学习任务中,空间和时间的信息都不是关键性的记忆线索。例如,海马损伤的病人能学习种种的运动和认知的技能,包括镜中描画、镜中阅读和河内塔(Tower of Hanoi)等技能(J. J. Cohen and L. R. Squir, 1981; N. Conen, 1984)。海马损伤的猴子能学会从一根棍棒上摘下糖圈的技巧(R. P. Kesner, et al., 1985)。海马损毁的大鼠也能学会在迷津的特定地点做急转弯的动作(R. P. Plunket, et al., 1973)。这种学习主要靠自身的运动觉的线索。

最应注意的是,在空间时间成为关键线索的学习任务中,如果空间时间的线索是以认知图的形式被利用的,即属于期望基础的记忆系统的性质,海马的损毁将不会产生有害影响。这种预见有许多重要的根据。例如,临床的研究曾发现,一位切除了两侧大部海马的病人(H. M.)很难按照新住处的平面地图从一个房间走到另一个房间。但他却能画出他手术后的早期曾住过的房子的准确的平面图,并能将所有房间都画在正确的位置上(N. Cohen, 1984)。

用大鼠做的实验结果也可作为这种预见的根据。在一个实验中,训练大鼠在一个 17 臂的放射形迷津中,只准在其中的 8 条臂中取食,其余 9 条臂中永无食物。经过训练的大鼠每次练习时学会了不重复进入已吃过其中的食物的臂,也不会进入不曾有过食物的 9 条臂中的任何一臂。在学习达到试者满意的标准之后,给大鼠做了切断繖-穹窿(Fimbria-fornix lesions)手术。手术后的大鼠在每次练习中仍然不进入任何一条不曾有过食物的臂,但会重复进入已经在其中吃过食物的臂(D. S. Olton and B. C. Papas, 1979)。这说明对于迷津空间的认知图的记忆(属期望基础的记忆系统)未受海马功能损伤的影响,而受影响的是临时的工作用的空间信息的记忆(属资料基础的记忆系统),即不能记住刚刚进入过的个别臂的位置。

试者在另一种比较简单的实验中,先训练大鼠从 3 个不同的

起始箱奔向一个特定的目的箱获得食物(强化);学会之后给大鼠做损毁纹-穹窿手术;手术后观察它们能否从一新的起始位置奔至原先的目的箱。这叫作迁移测验。设想大鼠在测验中,如果能从新位置奔至原先的目的箱,必然要利用在学习中获得的对整个迷津的认知图。测验的结果看到纹-穹窿(海马的出入纤维)损毁的大鼠和正常大鼠(对照组的)一样,都能从新起点奔至原先的目的箱。说明海马的损伤不影响空间的认知图,即不影响期望基础的记忆系统的空间属性的记忆。

总之,上述实验都可证明海马不属于保持认知图的系统,而是为资料基础的记忆系统的空间、时间的信息加工服务的。

(3) 杏仁核和体内状况的记忆

杏仁核参与情绪行为的机制已有多方面的证据:①杏仁核,特别是其中央核和下丘脑及其他自主神经系之间有许多交互的神经纤维的联系(B. S. Kapp, et al., 1984);②损毁杏仁核的病人和动物性情都变得柔顺,几乎完全没有攻击性了。对有害刺激的情绪反应明显地减少。例如杏仁核损毁的动物毫无惧怕地探索新异的和可能有危险的东西,以及降低了以惧怕为基础的学习能力(G. V. Goddard, 1964; B. S. Kapp et al., 1984; M. Narabayashi, 1972);③电刺激人的杏仁核引起多种情绪体验(如怕、怒或愉快),电刺激动物杏仁核引起攻击、逃走、心率快、呼吸粗、竖毛及瞳孔放大等生理反应(B. R. Kaada, 1972; V. H. Mark and F. R. Ervin, 1970);④在动物和人的杏仁核中埋植电极都可获得乐于自我刺激的行为(R. M. Wurtz and J. Olds, 1963; R. G. Heath, 1964);⑤人类的不愉快情绪经常伴有杏仁核的 EEG 变化(R. G. Heath, 1964)。这些事实说明杏仁核参与情绪体验和情绪行为的机制。

人们经常体会到曾经引起过强烈的情绪反应的事情最不容易忘记,所谓“入骨三分”。可以推测杏仁核可能是通过调理情绪体验的机制参与记忆系统。最早的设想是杏仁核记录学习中强化物的

性质(R. J. Douglas and K. H. Pribram, 1966)或中介刺激和强化的联系(B. Jones and M. Mishkin, 1972)。有许多动物实验的结果支持这种看法。例如,在大鼠学习躲避强的电击(2 毫安, 2 秒)脚掌的实验中,每次练习后,通过埋植的电极刺激杏仁核,可以破坏学习效果の长时保持。如果学习时用的是弱的电击(0.5 毫安, 0.5 秒),练习后杏仁核的电刺激反而有促进学习的影响(P. E. Gold 等, 1975)。

另有实验证明:大鼠在一放射形的 8 臂迷津中学习按照预示大奖(5 粒食物)或小奖(1 粒食物)的暗号再次进入前一次取得大奖或小奖的臂中的实验中,当给大鼠提供的暗号是预示大奖的时候,与前一次练习的间隔长达 24 小时仍有较好成绩,此种实验称为符号的延迟的空间配样任务。在这种实验中,如果在长延迟的配样测验中,当预示大奖的暗号出示的 10 秒期间给大鼠的杏仁核电刺激,就会妨害它的配样成绩;但是,在预示小奖的暗号出示期间给杏仁核电刺激则不影响其配样成绩(即在后一次的测验练习时仍然再进入前一次获小奖的臂中)(R. P. Kesner and R. G. Andrus, 1982)。这一实验似乎显示杏仁核的电刺激干扰了对大奖的记忆。如果认为大奖可能产生较强的情绪活动,那么杏仁核的电刺激影响对大奖的记忆可能是通过对记忆中起重要作用的感情成分的干扰。

此外,有的实验还证明杏仁核损毁的大鼠不会在偶然一次练习中删去奖励物之后增加跑的速度,即不发生挫折效应(R. G. Henke, 1977),杏仁核损伤的猴子也不能随奖励物的多寡改变它们的作业成绩(J. S. Schwartzbaum, 1960)。

根据上述的资料可以认为,杏仁核在由强烈的强化产生的感情性质的记忆中起一定的作用。但是关于杏仁核在记忆的时间和感情属性之间的交互作用中起什么作用的研究尚无直接的实验资料。

(4) 杏仁核损毁后不受影响的记忆力

根据假设的记忆的属性模型推测,有两类工作的记忆不会受杏仁核损毁的影响。一是虽然需要资料基础的记忆加工,但强化物却不是重要的记忆情节时,杏仁核损毁的动物会做出无异于正常动物的学习成绩。如猴子表现的以视觉辨别为主的再认记忆(M. Mishkin, 1982)。其次是,在利用期望基础的记忆系统时,即使是感情属性的记忆成分为主的学习任务,杏仁核损毁的动物仍能完成。例如,杏仁核损毁的猫虽然不能学会积极地躲避电击的反应,但是正常的猫在学会这种反应之后,杏仁核被损毁后也还能保持这种反应(J. V. Brady, et al., 1954)。

(5) 杏仁核与海马之间功能的可分离性

综上所述,杏仁核和海马对记忆的贡献是不同的。杏仁核在记忆感情成分为主的事件中起重要作用。海马在记忆以空间位置为主要线索的事物(例如陆标)中起重要作用。

除上面例举的实验证据之外,还可再举两例。实验是用电击脚的手段训练大鼠抑制一种行为(如不从台上跳下,或不进入暗室)以躲避电击。如果动物在一次练习之后用弱电流通过埋植的电极刺激杏仁核或海马,可产生遗忘效果,因为这种刺激可干扰杏仁核或海马的正常工作,致使动物不再出现抑制性的躲避行为。然而如果在第二次练习开始之前,先在另一笼内电击一下动物的脚掌,作为唤起记忆的感情成分的提醒刺激。在这种情况下,只有前次练习后海马受电刺激的动物恢复了记忆,而杏仁核受刺激的动物仍未能恢复记忆(L. J. Baker, et al., 1981)。这是由于练习后的杏仁核刺激已破坏了对电击脚掌产生的情绪后果的记忆(感情属性的记忆),因而以后的提醒刺激已不可能再起作用。

另一实验结果则与此相反。实验仍用电击脚掌的办法训练大鼠抑制与上述实验相同的行为以躲避电击。但在此实验中,电击脚掌是在一边箱壁发出荧光的情况下给的,这是为了加强学习任务

的外界空间线索的重要性。动物在练习之后也用与上一实验相同的方式和电流强度刺激杏仁核或海马。在 24 或 48 小时之后测验记忆。结果发现,只有在练习后海马受电刺激的大鼠表现有遗忘的效果;练习后杏仁核受刺激的大鼠则无遗忘表现(R. P. Kesner and J. D. Hardy, 1983)。这暗示在对环境事态的空间属性的记忆中,海马起重要作用,海马的功能受干扰则妨碍此种记忆;而杏仁核则与此种记忆无关。

应当说明,在一般的学习任务中,空间、时间的线索和事物的其他线索给予学习者的感情影响(如,好或恶)何者占主要地位,有时因个体而异,难于估计。但任何学习任务都或多或少地包含着多种可记忆的属性。因此,如果一个人的海马和杏仁核同时受损伤,其记忆力就显得更坏了。

(6) 尾核和自我中心的空间定位

在解剖关系上,尾核处于整个新皮质和运动系统的中间位置。从它和新皮质以及其他脑区的联系上看,它可以分为前背、后腹和尾部 3 个部分。其前背部分的主要功能是调处空间属性和反应属性的记忆以及在这两种属性交互作用的基础上实现的自我中心的(即以自身为坐标轴的)位置记忆。当然,自我中心的位置记忆也涉及到感知觉的属性(特别是本体感觉的)。

有许多研究的结果支持这种论断。兹举动物实验中的一例说明此问题:D. G. Cook 和 R. P. Kesner(1984)在一放射形 8 臂迷津中训练一组大鼠学习一种自我中心的定位任务(空间位置记忆)。方法是將大鼠放在每次练习时任意选择的一条臂的终端(起跑点)。让大鼠从此臂跑到毗邻的一臂中吃食(强化)。有的鼠学转入左边邻臂,有的学转入右边邻臂。在练习中,如果进入其他 5 条臂都得不到食物,经过多次练习,大鼠都能学会从起跑的一臂出来转入左边(或右边)的邻臂。向左或向右转的线索是由自身的运动感觉提供的,所以这是一种自我中心的定向任务。与此同时,这组大

鼠也在另一个 8 臂迷津中学习,在一次练习中,不重复进入任何一条已经进去吃了食物的臂。当这两种学习任务成功之后,给大鼠做尾核前背部损毁手术。两星期后测验它们对这两种学习任务的记忆,结果发现尾核前背部损毁的大鼠忘记了转入毗邻臂(左或右转)的习得行为,但仍能在另一迷津中不重复进入曾进入过的任何一臂。这一发现,可以说明学习进左边或进右边一臂,是一种自我中心的空间定位任务,要依赖尾核前背部的工作;而学习在 8 臂迷津中不重复进入任何一臂,是一种利用其他物标的(他物中心的)空间定位任务,似乎不需要尾核的工作。这一解释也符合在杭廷顿氏(Huntingtons)和帕金森氏(Parkinson's)病患者中发现的情况。这两种病都被诊断为尾核功能的障碍。

杭廷顿氏病人不能学会从镜中阅读的技能,因为这种技能要靠自我中心的定向能力(M. Martone, et al., 1984)。帕金森氏病人也难以完成需参照自身的空间关系来定向的任务。例如在暗室中让这种病人端正地坐着,将荧光屏上的一条光线调整到垂直于地面的方向。他们常常调不准。如果让他们倾斜着坐则更难调准。正常人一般都能完成这种任务。

(7) 尾核和海马两者的功能可分离性

尾核和海马在记忆功能上的区别和彼此独立工作的性质可从实验和临床研究的一些结果中看到。例如,尾核损伤的大鼠,如上所述,丧失了 8 臂迷津中习得的向左或向右转入邻近一臂的记忆,而仍然记得不重复进入任何一条曾进去过的臂;海马损伤的大鼠则与此相反。即仍能完成转入左或右臂的任务,而不能完成不重复进入任何一臂的任务(D. G. Cook and R. P. Kesner, 1984)。向左边或右边邻臂转的记忆是一种自我中心的空间位置属性的记忆,这是由尾核调处的。在 8 臂迷津中学习在一次练习中不重复进入曾进去过的臂,要靠对各条臂的特殊标记的临时记忆。这是属于资料基础的记忆系统中的他物中心的空间位置的记忆,由海马调

处的。

同理,尾核损伤的杭廷顿氏病患者不能学会镜中阅读技能,但有对测验的字的再认能力。而海马损伤的病人能学会镜中阅读技能,但丧失了对测验字的再认能力(M. Martone, et al., 1984)。

(8) 一级感觉皮质、运动皮质、脑干(Brain Stem)、小脑(Cerebellum)和刺激-反应(S-R)的联系

按照记忆的属性模型的设想,一级感觉皮质及其丘脑中转核、运动皮质及其脑干和小脑的线路都是为刺激反应(S-R)联系的信息编码和提取服务的。刺激反应联系的信息包含感知觉属性和躯体反应及自主系的反应的属性,为此服务的这一部分神经线路主要是在资料基础的记忆系统之内工作的。

关于特别事件的信息的获得需要特别的感觉系统的充分工作,包括初级感觉的中转系统对感觉输入的关键特性的有效的加工。因为在感觉系统内的信息加工是为各种属性的记忆活动提供重要资料的,所以在此加工的、属于感知觉属性的信息也就形成了任何资料基础的记忆的主要成分。就是在感觉系统内也可看到与学习和信息的贮存有关的重要变化。例如:研究者曾发现在用听觉刺激作为条件刺激的经典条件反应训练中,中脑下丘(inferior colliculus),内侧膝状核(medial geniculate nucleus)和听觉皮质(auditory cortex)中有些神经元的活动发生了与训练有关的变化——潜伏期短的反应增加(J. F. Disterhoft and D. K. Stuart, 1977)。应当注意的是,一级感觉皮质对于记忆任何感觉类型的各种特征来说并非完全不可缺少。例如,切除一级视觉皮质的猴子虽然不再能辨别物体,但仍保留着对物体位置的视知觉,即能指出测试物的位置(L. Weiskrantz, 1980)。这是由于视觉的不同特性有平行的加工系统。又如,大脑半球的新皮质全部切除的家兔仍能学习和保持瞬膜的条件反应,即由听觉刺激(CS)配合向角膜吹气(US)形成的瞬膜对声音的伸展反应(D. A. Oakley and I. S. Rus-

sell, 1977)。然而, 去掉新皮质的动物的反应显得更呆板和缺乏主动性(D. A. Okley, 1983)。

看来, 皮下的中转核, 包括脑干和小脑的神经网络能够调处简单的刺激反应之间的联系。现已发现小脑的齿状核和居间核(dentate and interpositus nuclei)属于中介瞬膜的简单条件反应的系统, 损毁这些核即不能获得和保持瞬膜的条件反应(D. A. McCormick and R. F. Thompson, 1984)。另外, 还发现这些核的神经元有与瞬膜条件反应的形成一致的放电活动的变化; 电刺激这些记录的地点也可引起类似条件作用的瞬膜反应。因此研究者们提出, 刺激反应之间的连接和条件作用的瞬膜反应的记忆应当是贮存在小脑的齿状核与居间核中的。其他系统主要是通过与这些核的连接对这种记忆起某种作用的。而其他的脑区也许在对其他种感觉刺激的条件反应的形成中有类似的重要性。

(9) 一级感觉皮质与海马之间功能区别和可分离性

一级感觉皮质和海马虽然都参与对外界事物的信息的加工和记忆过程, 但是两者的工作是可以分离开来的。

例如, 两侧视觉皮质损毁的猴子仍然能相当准确地察觉和摸到在他们的视野中短促呈现的物体, 然而却不能认识物体(即所谓心理盲)(L. Weiskrantz, et al., 1977)。一级视觉皮质损伤的病人也能指出一个物体的位置, 甚至辨别两种不同的物体, 但他们却说: “并不曾看见这个物体”(L. Weiskrantz, 1980)。海马损毁的动物和人则与此相反, 他们不能记忆物体的空间位置, 但能完成物体的辨别和再认的任务(R. P. Kesner and B. V. DiMattia, 1984; M. Mishkin, 1982)。这种功能上的分离可能是由于视觉信息的平行的加工线路所致。如视觉信息可通过上丘和外侧膝状核到达一级视觉皮质, 也可以经上丘和枕核(pulvinar nuclei)到顶叶皮质, 然后转到海马。

看来一级皮质更多地参与事物的认识机制, 而海马则更多地

参与对事物的空间关系的记忆机制。

(10) 小脑与海马之间的功能区别和可分离性

记忆的属性模型认为海马和小脑虽然都参与对外界事物的记忆,但记忆的属性有所不同。海马涉及的是外界事物的空间关系的记忆,小脑参与的是对外界事物的某种因果关系和自己行为的联系(即 S-R)的记忆。因此两者的功能是可以分离的。例如,海马损毁的家兔可以经过声音与角膜刺激(吹气)配合的条件作用形成瞬膜对声音的伸展反应(R. F. Thompson, et al., 1984)。海马损毁的病人也能用经典的条件作用的方法习得条件的眨眼反应(L. Weiskrantz and E. K. Werrington, 1979)。这类简单的条件反应的习得皆属小脑的工作。如上面提到的,小脑的齿状核和居间核损毁即不能习得这种条件反应。相反,海马损毁的动物(大鼠)却不能像正常动物一样地学会跑迷津(D. S. Olton, 1983),而小脑损毁的动物则仍能记忆学会跑迷津的方式(K. S. L shley and McCarthy, 1926)。这是因为跑迷津(特别是 8 臂迷津)需要记忆空间位置,空间属性的记忆属于海马的工作,所以海马损毁的影响严重。空间属性的记忆似与小脑无关,因此不受小脑损毁的影响。

4. 期望基础的记忆系统

(1) 背外侧或内侧前额皮质和时间认知图

前额皮质是大脑皮质额叶的一部分,包括 8, 9, 10, 11, 12, 44, 45, 46 和 47 区(图 3)。这些区域都接受发自丘脑内侧背核(mediodorsal thalamic nucleus-MDN)的神经纤维,故统称为 MDN 投射区。前额皮质可分为背外侧部(包括 9, 10, 44, 45 和 46 区)、眶额部(包括 11, 12 和 47 区)和额眼野部(frontal eye-field region, 包括 8 和 9 区)。前额皮质除了接受 MDN 的传入纤维之外,还接受来自视、听、躯体感觉联络皮质、嗅觉系统、下丘脑(hypothalamus)、海马、蓝斑(locus coeruleus)和中缝核(raphe

nuclei)的传入。此外前额皮质还有到下丘脑、中央灰质(central gray matter)、壳核(putamen)、尾核、运动皮质以及视、听和躯体感觉联络皮质去的传出纤维。从这些联系上看,前额皮质很可能成为一个高级整合中枢。

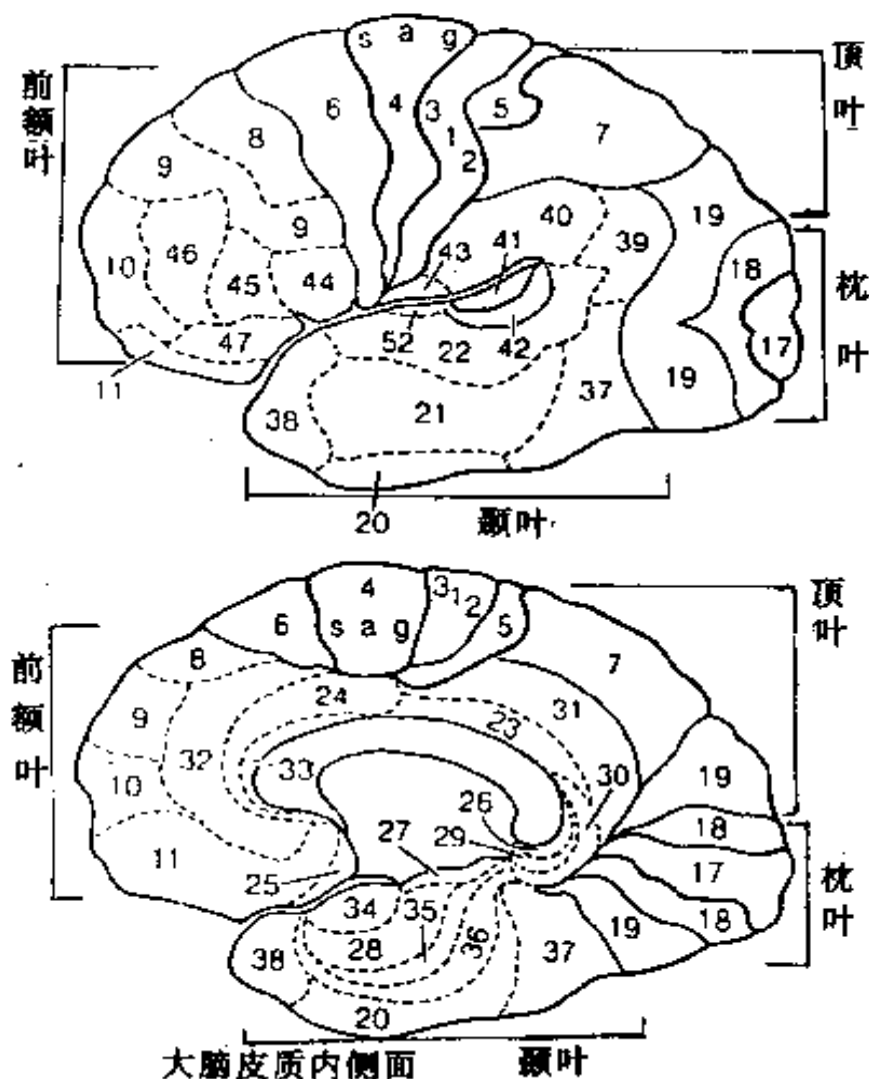


图3 人类大脑皮质主要区域示意图

前额皮质的结构和丘脑 MDN 的分区投射的模式在哺乳动物中(包括人类在内)都是非常相似的。在灵长目中,MDN 的外侧小细胞部分投射到背外侧前额皮质,MDN 的中央大细胞部分投射到眶额部皮质。在大鼠中,MDN 的外侧部分投射到内侧前额皮质。中央部分投射到嗅沟背岸的额皮质(眶-岛皮质)。由此来看,大

鼠的内侧前额皮质与猴或人类的背外侧前额皮质可能有相似的功能。

记忆的属性模型设想人类和猴子的背外侧前额皮质和大鼠的内侧前额皮质以及向这部分皮质投射的丘脑内侧背核贮存和提取期望基础的记忆系统中的时间信息——形成时间认知图。它们接受从二级感觉联络皮质调处的感知觉认知图中传来的重要信息，它们又影响由前运动皮质(premotor cortex)和补充运动皮质(supplementary motor cortex)调处的反应认知图系统。这一设想可以从前额皮质受伤的病人的症状中得到证明，这种病人缺乏主动性、不容易计划行动的程序、皮质的配套放电(corollary discharge)减少。此外，他们不能记忆所经历的事情的时间先后，不能计划复杂的动作或行为的先后次序。

最明显的是额叶皮质损伤的病人能记忆曾给他们看过的字或画，但不能记忆它们出现的先后次序(B. Milner, 1971)。这一事实说明额皮质损伤不影响对事件项目的记忆，但影响对次序的记忆，即不能贮存和提取时间信息。

在猴类中，背外侧前额皮质损毁也产生类似上述病人的症状。例如，额叶皮质损毁的猴子不能完成有 5 秒钟延迟的左右交替选择的任务，但能完成中间有 15 秒钟间隔的左右反应对的任务(如左选一次，右选一次为一对，中间停 15 秒，再做下一对)(K. H. Pribram and W. E. Tubbs, 1967)。这是因为延迟的交替反应突出了反应的时间次序特性(即如，先左后右或先右后左)。而在后一种任务中，每做完一次交替反应(对)，中间有较长的间隔，动物有可能将一次交替截作为一项动作的信息，这样就不需要记住左右反应的次序了，因此不受额皮质损毁的影响。

在大鼠中，内侧前额皮质的功能相当于灵长动物的背外侧前额皮质，因此损毁大鼠的内侧前额皮质后也出现对次序属性的记忆障碍。

一个最有力的证明来自 R. P. Kesner and B. V. DiMattia (1987) 的一个实验。实验是在一放射形的 8 臂迷津中训练大鼠。方法是,当大鼠学会进入每条臂中取食之后,每天在一次练习中先将 8 条臂的入口关闭,然后按随机选择的先后次序打开当天任意选择的 4 个入口。即强迫大鼠按照试者规定的先后次序进入 4 个臂取食。当大鼠从最后一个臂中吃了食物出来之后,立即开始测验它对次序的记忆。方法是:打开刚才训练时选择的 4 个臂中的任何两个,如,第 1,2 或第 3,4 或第 2,3,让大鼠进入刚才练习时在先的一个臂中才能得到食物(如,第 1,2 中选第 1 或第 2,3 中选第 2)。经过长时的练习,正常大鼠都可完成这种任务。但当学习完成之后,吸出大鼠的内侧前额皮质,等大鼠康复之后再做上述的测验练习,结果发现内侧前额皮质损毁的大鼠不再能完成这种任务。说明它们的次序记忆受到损害。但是在另一种项目属性的记忆测验中,大鼠的成绩却不受内侧前额皮质损毁的影响。测验是这样做的:当大鼠按规定的次序先后在 4 条臂中吃完食物之后。立刻打开其中的任何一个(第 1,2,3 或 4)入口和另外一个不曾进入过的路口。只有进入刚才进入过的臂才能得到食物。经过多次练习,正常大鼠也都能完成这种任务;内侧前额皮质损毁对此种任务的记忆无显著影响,特别是在第一段练习时每次选择的 4 个臂不变的条件下更不受影响。这说明大鼠的内侧前额皮质损毁不影响项目的记忆。

值得注意的是,电生理的研究发现,在猴子做延迟反应、延迟交替反应或延迟配样本反应时,背外侧前额皮质中有大量的神经元在延迟期间增加放电频率。而在这些神经元中,有许多是在动物学会必须记住指示正确反应的关键暗号时才增加放电频率。因此研究者曾提出前额皮质首先是加工短时记忆中的信息的时间顺序(J. M. Fuster, 1980),或控制工作记忆中注意的次序(K. H. Pribram, et al., 1964)。

(2) 背外侧或内侧前额皮质和空间认知图

背外侧前额皮质除调处时间先后的记忆之外,可能也编码、贮存和提取期望基础的记忆系统中的、有关自我中心的空间图的信息,它接受二级联络皮质调处的感知觉认知图的重要输入。这一假设得到许多实验和临床观察的证明。例如,有一种实验是训练猴子在两个相同的物体中选择左边的(或右边的)一个得到食物,这是一种以自身为坐标的,即自我中心的、空间关系的识别任务。正常的猴子很容易学会这种辨别;而且在学会之后还可以再学会把选择的方向颠倒过来,如学会选左边的物体后,再反过来学习选右边的,也能成功。而背外侧前额皮质损毁的猴子则很难完成后一种学习任务。然而,如果学习的任务不是选择左边或右边的物体,而是学习在两个相同的物体中选择其中靠近某种标志物的一个(即他物中心的空间认知图),外侧前额皮质损伤的猴子也能完成颠倒学习的任务(W. Pohl, 1973)。

电生理实验已发现,在左右辨别的延迟反应的练习中,前额皮质中许多神经元对指示左或右的暗号有区别性的放电频率。而另一些神经元则在正确的运动反应出现前放电(似乎是准备性质的, H. Niki, 1975)。

更有趣的一个实验结果是,当训练猴子学会数数时,要求它们记住前次在某一键上扣击的次数。正常猴子可以记住,背侧前额皮质损毁的猴子则很难完成这种任务(R. Passingham, 1978)。这暗示记住扣击的运动次数也需要记住扣击的空间位置。因此就需要有背侧前额皮质的工作。

临床观察也发现,额叶皮质损伤的病人在利用自我中心的空间定向的测验中也暴露了同样性质的缺陷。例如,他们不能按照人体的前或背面图上指定的位置触摸自己身上的相当部位。然而,他们却能按照房间的平面图上标出的路线在实际的地板上的许多点子中选择相应的连接路线来走(J. Semmes, et al., 1963)。前一种任务属于自我中心的空间定向,需要前额皮质的工作。病人故不能

完成,后一种任务属于他物中心的空间定向,则不受前额皮质损伤的影响。

(3) 背外侧或内侧前额皮质与海马之间的功能区别和可分离性

背外侧或内侧前额皮质都有处理时间属性记忆的功能。在猴和大鼠中也发现了海马与背外侧或内侧前额皮质之间的解剖联系(L. W. Swanson, 1982; R. S. Goldman-Rakic et al., 1984)。说明这两个脑区之间可能有重要的功能关系。

至少有两种可能:①内侧前额皮质和海马可能是平行工作的,每一脑区都全面地加工时间信息;②两个脑区可能是串联工作的,各自贡献不同的成分。记忆的属性模型暗示:两个脑区可以独立工作,但也有串联的关系。海马被认为属于资料基础的记忆系统的整合部分,它编录情节的、变化的时间信息;而背外侧或内侧前额皮质则被认为属于期望基础的记忆系统的整合部分,编录以时间策略和法则形式存在的时间认知图的信息。

根据上述假定,人们可以预见这两个脑区的分别损毁可以产生不同的记忆缺陷。如在前面的几节中提到的前额皮质损伤的病人丧失了对以前熟悉的事情的时间次序的记忆,而海马损伤的病人则无此种记忆缺陷,但却不易记住临时遇到的事件的先后次序。又如,额叶损伤的病人在解决汉诺(Hanoi)塔难题的测验上成绩都是很差的(T. Shallice, 1982)。这被认为是由于解决这种问题需要记住和计划运用时间和空间关系的策略,而这是属于额叶皮质的功能。相反的,海马损伤的、患有遗忘症的病人(如, H. M.)则能学会解决汉诺塔难题。其次,额叶皮质损伤的病人常常有一种类似补偿的机制,即将注意盯在一件事物上。如学会了按颜色来分卡片之后,不易改变为按图形分卡片。而海马受损伤的 H. M. 则能改变卡片分类的法则(B. Milner, 1964)。

内侧前额皮质损毁的大鼠能记忆要选择的一系列变化的空间

位置中的第一个位置,但不能记住其他的位置。这也被认为是由于注意专一的补偿机制所导致的后果(R. P. Kesner, 1983)。

(4) 背外侧或内侧前额皮质与尾核之间的功能区别和可分离性

背侧尾核前部与背外侧或内侧前额皮质之间在功能上有许多相似性。因为内侧前额皮质和背侧尾核前部有直接的神经联系(I. Divac, 1972)。记忆的属性模型假设背侧尾核前部的主要工作是在新的和变化的情境中选择适当的身体运动。这是属于资料基础的记忆系统的。而背外侧或内侧前额皮质则是调处自我中心的空间认知图的;这就是说,它是指导在熟悉和不变的情境中应做的适当的身体运动,这是属于期望基础的记忆系统的工作。但两个脑区处理的都是自我中心的空间关系。而者的区别可见于如下事实:

第一,损毁婴猴的背外侧前额皮质,在2岁以前,不出现做延迟的空间位置反应的缺陷(P. S. Goldman and H. E. Rosvold, 1972),即不妨碍资料基础的记忆。相反,损毁婴猴的尾核则立即破坏延迟反应的学习(A. Kling and T. J. Tucker, 1967),说明尾核的损毁影响资料基础的记忆。

第二,在猴子做延迟反应的延迟期间,只有在前几秒时电刺激背外侧额皮质才会破坏其作业,稍后刺激即无影响;而在延迟期间,无论什么时候刺激背侧尾核前部都会破坏其作业(J. S. Stann, 1969)。这也说明这两个脑区在工作上可能有所不同。

此外,在临床的实验中看到帕金森氏病人和额叶损伤病人都在奥伯特(Aubert)测验(调准垂直线)中表现有缺陷,但额叶损伤病人只是在坐着倾斜的椅子时才显露出困难,帕金森氏病人则在坐正时也难调准(H. L. Teuber, 1976)。这一事实说明前额叶皮质和尾核在处理自我中心的空间关系上确实有其不同之处。

(5) 眶额皮质和感情反应(自主系)的认知图

记忆的属性模型提示在人类、猴和大鼠等动物中眶额皮质及

其丘脑投射部分贮存和提取有关感情形成的认知图的信息,是属于期望基础的记忆系统的。对奖赏和厌恶性刺激的感觉传入影响感情认知图,它又借助自主系的反应图来选择适当的自主反应。

此观点也有各方面的根据:例如,伤及眶额皮质的额叶损伤病人的临床资料说明这类病人表现各种的个性改变,其特点有:好戏谑、快乐、易激怒、突然抑郁和缺乏社会行为的判断力。这类病人大多数缺乏对社会法则的重视,因此不能享受社会奖励的愉快或知识性的愉快,他们的感情似乎有了改变,对痛刺激的反应也降低。在猴子中,眶额皮质损毁后攻击行为减少,对厌恶的视觉刺激的情绪反应也消失或减弱(C. M. Butter, et al., 1970)。大鼠的眶额皮质损毁后攻击行为同样也减少(B. Kolb, 1974)。

眶额皮质损毁的动物习得的反应不再受到奖励(强化)时也不易消退。所以很难完成物体或空间位置辨别的颠倒学习(C. M. Butter, 1969; S. D. Iversen and M. Mishkin, 1970; B. Jones and M. Mishkin, 1983)。单位神经元放电的研究发现动物在期待奖励或惩罚时眶额皮质中有反应各异的神经元(S. I. Thorpe, et al., 1983)。

此外,电刺激猴、猫和狗的眶额皮质产生包括血压和呼吸变化的各种自主系的反应(B. R. Kaada, 1951)。眶额皮质与控制自主反应的下丘脑和杏仁核有神经联系。后两个脑区都参与情绪活动的机制(J. Aggleton, et al., 1980; L. J. Porrina, et al., 1981)。

(6) 眶额皮质与杏仁核之间的功能的可分离性

眶额皮质和杏仁核看来都参与感情经验的记忆机制,但两者似乎有一定区别。杏仁核被认为是贮存和提取变化的情节性的感情属性的信息,因此是属于资料基础的记忆系统的整合部分;眶额皮质则被认为是贮存和提取类似奖励或惩罚的规则和策略这样的、属于感情性认知图的信息,因此是期望基础的记忆系统的整合部分。根据这种设想,可以预见眶额皮质损毁的动物应在已习得的

以某种方法获得奖励物或躲避惩罚的行为方面显露缺陷,而杏仁核损毁的动物则不会有这方面的缺失。有两个实验与此预见一致:一个实验证明,当猫学会了主动地躲避电击之后,损毁了眶额皮质,它们就会丧失了躲避的记忆。但是杏仁核损毁之后则不会丧失这种记忆;相反的是,如果在学习躲避反应之前损毁了杏仁核,学习任务就不能完成了(J. V. Brady, et al., 1954)。另一实验发现,猴子的眶额皮质或杏仁核损毁之后都影响物体或地点辨别的颠倒学习,但表现的方式有所不同。眶额皮质损毁的猴子在颠倒辨别练习的早期阶段出现的错误最多,似乎表示它们不能抑制以前习得的对奖励的预期行为。杏仁核损毁的猴子发生的错误则在颠倒练习的整个阶段都比较多。这被认为是由于杏仁核损毁妨害形成新的刺激与奖励的联结(B. Jones and M. Mishkin, 1972)。

(7) 眶额皮质与背外侧或内侧前额皮质的功能区别和可分离性

记忆的属性模型假定眶额皮质是为感情性质的认知图服务的,而背外侧或内侧前额皮质则是为自我中心的空间和时间认知图服务的。这种区别可见证于眶额皮质损毁的猴子仍能学习延迟的位置辨别任务和延迟的交替反应任务,背外侧前额皮质损毁的猴子则不能(S. D. Iversen and M. Mishkin, 1970)。眶额皮质损毁的猴子在不给奖励的情况下,消退习得的反应有困难,背外侧损毁的猴子则无此种困难。

(8) 后顶皮质和他物中心的空间认知图

后顶皮质(PPC)是顶叶的一部分(图3),包括5,7(上顶叶),39(角回)和40(下顶叶)区,这些区域的传入来自外侧和后丘脑、枕核(pulvinar n.)、一级和二级感觉皮质。从这些区域传出的神经纤维投射到额和颞叶的联络皮质、外侧和后丘脑、壳核(putamen)、尾核、上丘(superior colliculus)、脑干和脊髓的神经核(A. Brodal, 1981)。

在猴类中,上顶叶标为5区,下顶叶标为7区。这两个区域与视觉皮质、躯体感觉皮质、额叶的大部分、上颞沟皮质、扣带回、对边半球的相应区、基底神经节、丘脑的枕核和外侧后核及腹外侧核、中脑上丘和桥核都有双向联系(J. Hyvarinen, 1982)。

大鼠中相当于5,39和40区的投射尚不十分清楚。已知丘脑的外侧部分、主核的后部和外侧部分有投射到这些皮质区的神经纤维(K. S. Lashley, 1941; W. F. McDaniel, et al., 1978)。大鼠的丘脑的后外侧部分被认为是枕核的先区(A. Brodal, 1981)。因而,在大鼠和猴子中可以找到相对应的丘脑投射区。从上述的神经解剖的关系上看,后顶皮质也被认为是一个高级的整合中枢。

设想后顶皮质和它的丘脑投射可能贮存和提取期望基础的记忆系统中的、他物中心的空间认知图的信息。它从二级感觉联络皮质中形成的感知觉认知图中接受主要的传入。它控制在前运动皮质和补充运动皮质中调处的反应认知图系统。这种设想也有临床资料和动物实验的根据。例如,给右顶皮质受伤的病人做知觉匹配测验,即判断两个卡片上的点子的位置、线条的方向或两线之间的距离是否相同(这是一种空间关系的辨别任务),他们会做出许多的错误;但颞或额皮质受伤的病人做此种测验的成绩无异于常人(E. K. Warrington and P. Rabin, 1970)。又如,测验证明后顶皮质损伤的病人在过9至16秒后就不能记忆3个方块曾摆过的位置(E. De Renzi, et al., 1977),过3至10秒之后就不能再认刚刚看过的无意义的图形(N. Butter, et al., 1972),也不能在自己住的城市的地图上找到自己住处的区域和画出自己的办公室或自己家中的空间布局(H. Hecaen and M. L. Albert, 1978)。

还有许多测验都证明后顶皮质损伤的病人在他物中心的空间认知图的信息贮存和提取方面有严重缺陷。两侧后顶皮质损伤,此种缺陷更为严重。

在动物中,猴子在后顶皮质损毁后所表现的记忆缺陷和在后

顶皮质受伤的病人中所观察到的情况十分相似。例如,后顶皮质损毁的猴子不能辨别地点标记物(简称路标)的距离(A. D. Milner, et al),但仍能辨别图形和物体(M. Petrides and S. D. Iversen, 1979);在根据物体与路标的距离做选择的学习任务(简称路标辨别)中和颠倒这种选择的学习中,后顶皮质损毁的猴子表现很坏。由此看来,猴子的后顶皮质和人类的后顶皮质参与相类似的空间辨别和他物中心的空间认知图的信息加工。

损毁后顶皮质的大鼠在学习迷津的时候也表现有严重的困难。例如,大鼠在放射型的8臂迷津中可以学会不进入4条永无食物的臂和在一次练习中不重复进入已在其中吃过食物的臂;但当大鼠学会之后,有的损毁后顶皮质、有的损毁视觉皮质、有的只做假手术(不损毁皮质),手术后测验发现这3组动物仍能不重进曾进入过的臂。视觉皮质损毁的大鼠和假手术大鼠也能不进入那4条永无食物的臂;但是后顶皮质损毁的大鼠却多次进入以前永无食物的臂。在每次练习中记住已经进去吃过食物的臂的位置是资料基础的记忆系统的工作,后顶皮质损伤的大鼠似乎未丧失这种记忆能力。记住以前永无食物的4条臂的位置是期望基础的记忆系统的工作,后顶皮质损毁的大鼠丧失了这种记忆。这一实验结果证明在大鼠中后顶皮质的损毁与在人和猴中一样,也影响已建立的空间认知图的利用。

(9) 后顶皮质和海马的功能的区别和可分离性

后顶皮质和海马都涉及他物中心的空间属性的记忆。这两个脑区必定有重要的功能联系。神经解剖方面的证据是:后顶皮质7区的神经元有些轴突投射到旁海马回(parahippocampal gyrus)和前下托(presubiculum),旁海马回通过内鼻区(entorhinal)皮质投射到海马,前下托直接投射到海马,海马又通过内鼻区皮质和旁海马回投射到7区(G. W. Van Hoesen, 1982)。

前面已提到,海马调处情节性的、变化的空间和时间属性的信

息记忆；后顶皮质调处他物中心的空间图的信息记忆，即与空间性质有关的规则或策略的记忆。由此可以预见，这两个脑区个别的损毁会产生不同的记忆缺陷。例如后顶皮质损毁后会丧失对以前熟知的地点位置的记忆，而海马损毁后则不会有此种记忆缺陷。

临床资料证明，顶皮质受伤的病人不能画出他们的旧居处的地图和所在的空间位置(F. DeRenzi, 1982)，但海马损伤的病人(如 H. M.)却还能记得手术前的住处所在的地方(空间位置)(N. Cohen, 1984)。

一个有趣的证明实验是这样安排的：让大鼠在一个 12 臂的放射形迷津中学习进入各个臂中吃食物。学会后开始进一步的训练，在每一次练习中只开 4, 6, 8 或 10 个臂的门，让大鼠进去吃食；在做完每次的练习后，过 15 分钟再将大鼠放入迷津中，但这次只开两个门，其中一个门是前一次练习时不曾开过的。大鼠进入此次新开的这个臂可以得到食物，进入前次开过的臂得不到食物，记作错误。经过多次训练之后，当在先的一次练习中开的臂少的时候，如只开了 4 个臂门，正常大鼠在 15 分钟后的两个门的选择中都能较多地选择新开门；但当先一次练习中开的门多时(如开 6 个门)，随后的选择的错误就会增加。有趣的是当先一次练习中开的门再增加时(如开 10 个门)，随后的选择误差反而减少了。这被认为是由于大鼠在需要记忆的项目少和项目多时，可能利用不同的策略。项目少(开的门少)时用回想的记忆——即回想曾进去过的臂的位置；项目多(开的门多)时用前瞻的记忆——即瞻望尚未进去过的臂。两种记忆实质上都属于资料基础的记忆系统。此实验发现，正如所预见的，海马损毁的大鼠在少开门和多开门的练习后的两门之间的选择测验中都是失败的；而后顶皮质损毁的大鼠则只在练习中开的门多时才会在随后的选择测验中出现较多的错误(R. P. Kesner and H. Kametani, 1986)。这一结果证明，设想为期望基础的记忆系统服务的后顶皮质的损毁不影响对资料性质的(臂门临

时的)空间位置的记忆;后顶皮质损毁的大鼠在项目多的练习后的选择测验中错误增加,可能是由于项目多增加了海马负担所致。

(10) 后顶皮质与背外侧或内侧前额皮质之间的功能区别和可分离性

记忆的属性模型设想后顶皮质调处的是他物中心的空间认知图的信息,背外侧或内侧前额皮质调处的是自我中心的空间认知图和时间认知图,因此他们对记忆有各自的特殊贡献。例如,额叶皮质受伤的病人不能以自身的左右来定向,但能按照画出的图样走过地板上的一系列点子(他物中心的空间定向);顶叶皮质受伤的病人与之相反,不能照上法走,但能按自身的左右来定向(J. Semmes, et al., 1963)。又如,背外侧额叶皮质损毁的猴子不能完成左右辨别的颠倒学习,但能完成路标辨别的颠倒学习;而顶皮质损毁的猴子则与此相反(W. Pohl, 1973)。

在时间认知方面,额皮质损毁的猴子不能学会按规定的顺序进行选择的任务,而顶皮质损毁后则无此缺陷(J. V. Brody and K. H. Pribram, 1978)。

(11) 二级感觉联络皮质和感知觉认知图

二级感觉联络皮质为高级组织的感知觉的信息编码,为记忆与感觉传入的交互作用提供一种背景,形成感知觉的认知图。这个系统的工作影响期望基础的记忆系统中的其他各种的认知图。各种感觉系统都可视为二级感觉联络皮质的部分。视觉系统的一部分在下颞皮质,关于这一部分的研究最多,可作为一个例子来说明。

下颞皮质接受的大部分是来自视觉皮质加工过的视觉信息,由下颞皮质再加工的信息传到杏仁核、海马和尾状核的尾部。下颞皮质在视觉的学习和认知中极为重要。下颞皮质损毁的猴子学习视觉的物体辨别任务非常困难,但不影响学习听觉、触觉或嗅觉的辨别任务。此外,下颞皮质损毁的猴子更不能学习配样本一类的任

务,但它们仍有正常的视敏度和完整的视野(P. Dean, 1976, 1974)。看来它们丧失的是视觉的高级功能。

右颞叶皮质受伤的病人也有类似的视觉障碍,如对见过的面孔的再认和回忆看过的几何图形都有困难(B. Milner, 1968)。

单位神经元放电的研究发现下颞皮质中有些神经元的特性与这部分皮质调处感知觉认知图的设想一致。例如,猴子的这些神经元的接受野很大,大多数神经元反应颜色、形状、质地或这些性质的联合刺激;另有一些神经元有选择地反应特别的物体,如猴子的脸和手(C. G. Gross, et al., 1984)。在猴子做延迟的配样本任务时,在下颞皮质中可以找到在延迟期间持续放电的细胞(J. M. Fuster and J. P. Jervey, 1982)。

两边的枕-颞区受伤的病人不认识熟人的面孔、称为面孔不识症(prosopagnosia)。这种病人可以辨别不认识的面孔的异同,而且能再认比面孔更为复杂的刺激物,但不能把面孔的视觉与有关熟人的特殊经验整合起来(A. R. Damasio, 1985)。

这些事实可说明下颞皮质似乎是调处视知觉的认知图的信息,属于期望基础的记忆系统的部分。

(12) 下颞皮质与一级视觉皮质的功能可分离性

一级视觉皮质和下颞皮质都负担视觉功能,这两个区域有神经解剖上的联系:一级视皮质被认为是编码情节性的、变化的视觉信息;下颞皮质调处视知觉认知图,可能是以模样或类别形式贮存和提取视觉信息。

两个区域的功能的这种区别有许多事实可证明:例如,一级视觉皮质损毁的猴子的视敏度受损害,但不影响视觉图形的辨别学习;相反,猴子的下颞皮质损毁不影响视敏度,但妨害图形辨别学习(A. Cowey and L. Weiskrantz, 1967; W. A. Wilson and M. Mishkin, 1959)。又如,枕-颞部受伤的病人不能认识熟人的面孔;但无视敏度的损失,因此能辨别不同的人面孔(A. R. Damasio,

1985)。

(13) 下颞皮质与后顶皮质之间的功能可分离性

已知下颞皮质是调处视知觉认知图的,而后顶皮质是调处他物中心的空间认知图的。两者对记忆的贡献是不同的,而且是可分别工作的。例如,下颞皮质损毁的猴子不能学习匹配样本或不配样本的任务,不能再认曾看过的物体;而顶皮质损毁的猴子则无此种缺陷。相反,后顶皮质损毁的猴子不能学习路标辨别(属他物中心的空间认知图),而下颞皮质损毁的猴子则无此种缺陷。

(14) 前运动和补充运动皮质与躯体反应认知图

在人类和猴子中,6区在一级运动皮质(PM-4区)的前方,它包括外侧面上的前运动皮质(PMC)和内侧面的补充运动皮质(SMC)。在大鼠中,与前运动皮质同性质的区域亦在4区前(图4),前运动皮质内有丰富的神经元发出轴突投射到4区和7区,但前运动皮质没有直接到脊髓中去的投射。然而有通过皮质网状路线间接下行的传出(H. Freund, 1984)。投入前运动区的神经纤维来自许多脑区,有小脑深部的核、腹外侧丘脑、额叶皮质、视区、听区和顶叶的7区。补充运动皮质的传出纤维加入皮质脊髓束。传入的纤维来自顶叶5区和1,2区的外侧部分。

两侧前运动皮质受伤的病人不能站立和迈步。补充运动皮质受伤的病人缺少自发的运动,说话也减少,两手的配合活动和模仿别人的动作皆受到障碍(H. Freund, 1984)。

设想前运动和补充运动皮质负责贮存和提取期望基础的记忆系统中的反应图。这一设想的根据来自临床资料、动物的切除部分脑区的实验和电生理的研究。例如在人类,前运动皮质受伤都会产生运用不能症(apraxic),其特点是病人虽然仍有足够的力量、协调作用、感觉和意志,但是却不能完成系列的或有次序的随意动作(R. K. Deuel, 1977; D. Kimura and Y. Archibald, 1974)。这表明前运动皮质可能在安排动作的次序和计算动作的时间上起重要的作

用。

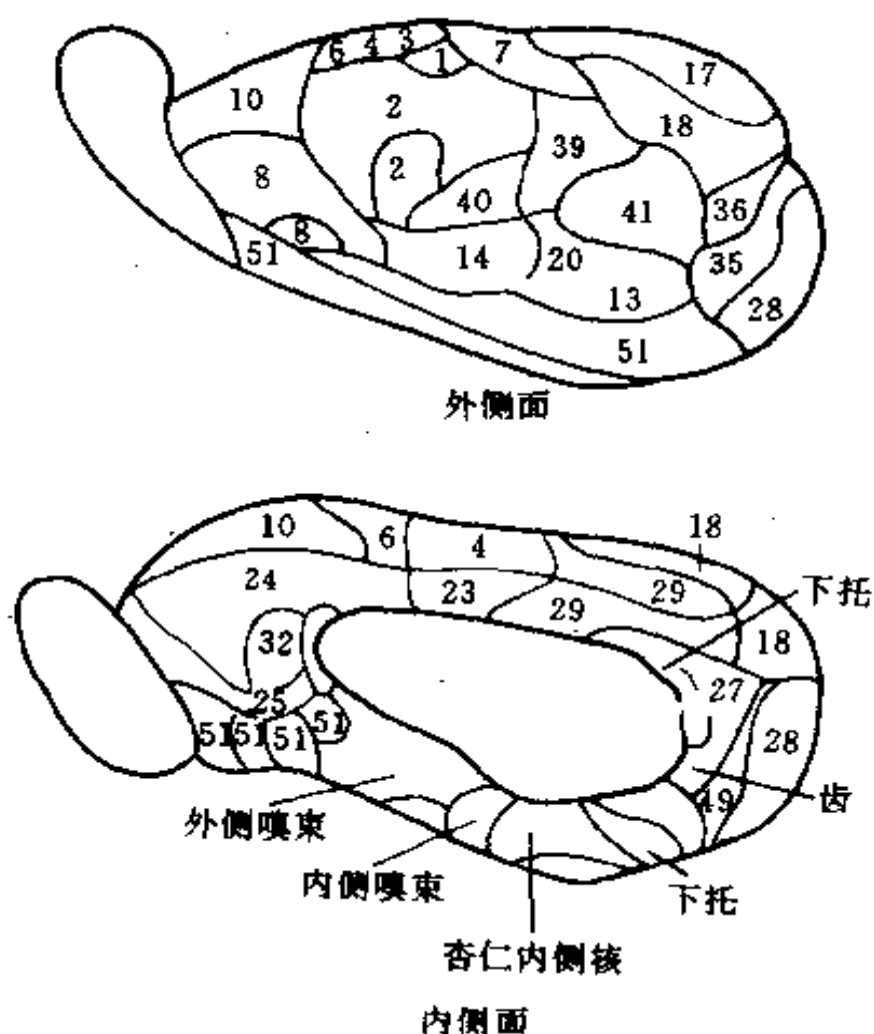


图4 大鼠大脑皮质区域示意图

补充运动皮质的损伤也产生运动的次序性的障碍(J. Pailard, 1982)。此外,补充运动皮质还涉及到计划和发动运动的功能。例如,当一个人去做或想象某种有次序的复杂的动作时,其补充运动皮质的局部血流急剧增加(这是用注射放射性标记的氩测定的)。前运动皮质和感觉运动皮质区的血流则是在执行这一动作的前后才增加的(P. E. Roland, et al., 1980)。更具特点的是,补充运动皮质在做复杂的动作时比做简单的动作时更为兴奋。

在猴子中,前运动皮质和补充运动皮质也有贮存和提取反应

图的功能,如前运动皮质和弓状沟部分损毁的猴子在做复杂的有顺序的动作时发生困难(R. K. Deuel, 1977)。如先训练猴子做一种复杂的开箱子的工作,其中包括按一定的顺序開箱子的门闩和取出箱中的食物;训练成功之后,损毁猴子的前运动皮质或后顶皮质。手术后,前运动皮质损毁的猴子尽管很快地恢复了运动能力,并能做开闩的每一个个别的动作,但却不能按一定的操作顺序打开箱闩;后顶皮质损毁的猴子能顺利完成打箱闩的任务,但却摸不到箱中的食物。又如,前运动皮质损毁的猴子可以拿到伸手就能够到的物体,但不能伸到玻璃挡板的后面去拿东西(L. Moll and H. G. J. M. Kuypers, 1977)。

大鼠的前运动皮质同型区损毁后能做简单的延迟的交替反应,但在学习按照门上的黑色推门和白色拉门的动作时发生了困难。单位神经元放电的研究发现,在猴子的前运动皮质和补充运动皮质中,有些神经元当要求做某种动作的信号出现时或在准备动作期间,其放电有持续性的变化的特点(S. P. Wise and E. J. Evarts, 1981)。这些事实都暗示前运动皮质和补充运动皮质是期望基础的记忆系统中的反应图的神经基质。

5. 各种属性记忆的工作动态

属性记忆的结构如何工作,可作如下的几个假定:

第一,存在的记忆结构一般的是处于不活动状态的(D. O. Hebb, 1949; D. J. Lewis, 1976)。在一些相当的属性刺激出现时,有关的神经区域就活动起来。这种活动状态称为短时记忆并且受多种过程的影响,其结果是增加了记忆结构组织中的持久性变化的可能性。

第二,每种属性的信息平行地激活资料基础和期望基础的相应的神经区,随后便是这两个系统之间的交互作用。这两个系统的神经活动的水平随情境的需要而变化。例如,就他物中心的空间属

性而言,当重要的空间信息能在期望基础的记忆系统的空间认知图中编码时,后顶皮质就会比海马更为活跃。相反,当重要的空间信息因为是新异的或变化的,而不编作认知图时,海马就会比后顶皮质更为活跃。有关其他属性的各脑区的活动水平亦同此理。

第三,期望和资料基础的记忆系统的工作都是短时的和能量有限的,这反映关键性的神经元系统的活动是短时的。

第四,选择性的注意过程不仅能影响哪种神经系统将要活动,而且影响活动的程度。此外,注意的控制水平也会改变,即从不需要费力的自动注意改变到需要很努力地控制的注意。自动的注意产生极小的神经活动,努力的注意(例如预演和思考帮助记忆的方法)能增加活动的持续时间和活动起来的神经线路的数量。例如,后顶皮质中的某些细胞只有当猴子集中注视视野中的物体时才活动起来(V. B. Mountcastle, et al., 1975),这些物体还必须是伸手够得到的和具有某种动机意义的。又如,猴子的下颞皮质中的神经元,当猴盯住或注意视觉辨别物时才增加放电率(C. G. Gross, et al., 1979)。还须注意的是,在脑损伤的病人中有各种的注意缺陷,因个别的脑区涉及的属性而异。例如,后顶皮质受伤的病人表现不能注意由他物指示的空间目标。此种缺陷在任务困难时表现更明显,因为它需要更高度的注意。例如,有一个这样的病人能和正常人一样地指出试者刚才敲击过的一些木块,但不能学习通过视觉的针迷津。此病人没有其他方面的遗忘症,其他的认知和口语测验的结果也都正常(B. Milner, 1965)。

细察后顶皮质损伤的病人常常只能一次注意一个物体或问题的一个片段,不能把一些个别的空间关系整合成一个整体。例如,可以描述一张画的个别的部分,而不能说明整个画的主题是什么。这种现象暗示简单的空间关系的加工可自动地进行,而对于外界环境的空间关系的高级的整合则需要包括后顶皮质在内的更复杂的加工。

有许多动物实验的结果可以得出同样的结论。看来,注意过程是个别的神经线路活动的水平和持续时间的主要决定因素。在相当于某些属性的脑区被激活之后,它们活动的时间长了就会触发巩固过程。巩固过程能将主要的信息保持在长时间的活动状态,其结果是增加了调处长时记忆的神经网络的组织发生结构变化的可能性。巩固过程可以发生在任何重要的脑区,这决定于特别属性和期望或资料基础的记忆的要求。然而巩固过程更有可能开始于资料基础的系统中,因为这个系统是由新异的和变化的情境信息的传入激活的,而新的情境更容易触发集中的注意。因为外界的事态变化常常是记忆的重要成分,所以巩固过程的激活更经常地开始于海马中。中介巩固过程的一种神经机制可能是长时的加强作用(long-term potentiation),这种现象已在海马中观察到了(T. J. Teyler and P. Discenna, 1984)。设想期望基础的记忆系统内的认知图就是信息贮存之处,所以认为巩固过程能够通过信息从海马或杏仁核的转移中改变新皮质内的神经元的活动,或直接改变新皮质的神经元的活动。应该说,短时记忆、注意和巩固过程是处理每种属性记忆的主要脑区的固有功能。

二、学习和记忆的神经元基础

学习和记忆的脑的解剖基础前节已作了说明。本节将深入一步讨论学习(或经验)引起的脑内的细胞水平的变化——即神经元的生理和结构方面的改变。

1. 短时记忆

在人们的经验中有许多事不需要永久的记着,有的是浮光掠影,留意一下即可忘掉。如查一个不是经常用的电话号码,查得之后只待拨完电话,片刻即忘,这谓之短时记忆。短时记忆被认为是

由刺激引起的神经元的较持久的活动来维持的 D. O. Hebb(1949)曾设想这是刺激引起的神经兴奋在一定的神经元环路中发生的类似振荡性质的作用;后来的实验发现,电刺激猴子的下颞皮质可以将它刚刚看过的图形的记忆抹去(R. Kovner and J. S. Stamm, 1972);更晚的研究者们又发现,在猴子做延迟的配样本实验时,当样本刺激出现后,其下颞皮质中有些神经元开始增加放电频率,在样本刺激消失后的整个延迟期间一直持续着(J. M. Fuster and J. P. Jervey, 1981)。这些事实支持了 D. O. Hebb 的设想。

2. 长时记忆

短时记忆可以消失,也可以转变为长时记忆。这包含着更复杂的神经元变化的过程。要分析这些过程,必须从最简单的学习过程开始。一个最简单的学习模型是经典的条件作用。例如,铃声最初不能引起狗的唾液腺分泌,但是铃响时有食物吃。多次之后,狗听到铃声就会流出唾液。这就是条件反应的学习过程。这种学习的效果可维持较长的时间。D. O. Hebb(1949)曾将这种学习模型作为长时记忆的基本的依据来研究。经典条件作用的成功,现在解释为异突触的易化作用(heterosynaptic facilitation)。意思是说,在一个神经元上的一种突触的活动可加强另一种突触的效力。例如,食物刺激(无条件刺激-US)的突触活动能加强铃声刺激(条件刺激-CS)的突触的效力。Hebb 设想,只有在后突触神经元放电的时候才会发生突触的加强作用。

异突触的易化可用图 5 说明。图中 CS 突触对后突触神经元 A 的作用本来是很弱的,不能使 A 放电。US 突触对 A 有固有的强效作用,能使 A 放电,引起唾液腺分泌。而只有在 A 兴奋时,CS 突触的活动触发 A 放电的效力才有可能被加强。至于异突触的易化如何发生和保持,可能包含有下述的一些变化:

① 后突触神经元反应的变化;

- ② 前后突触长大的变化；
- ③ 后突触膜的神经递质的受体的增加；
- ④ 神经元轴突上生出新的终扣以及与别的神经元建立全新的突触。

这些变化的任何一种都会影响神经元在网络中相互作用的方式，都可作为长时记忆的基础。

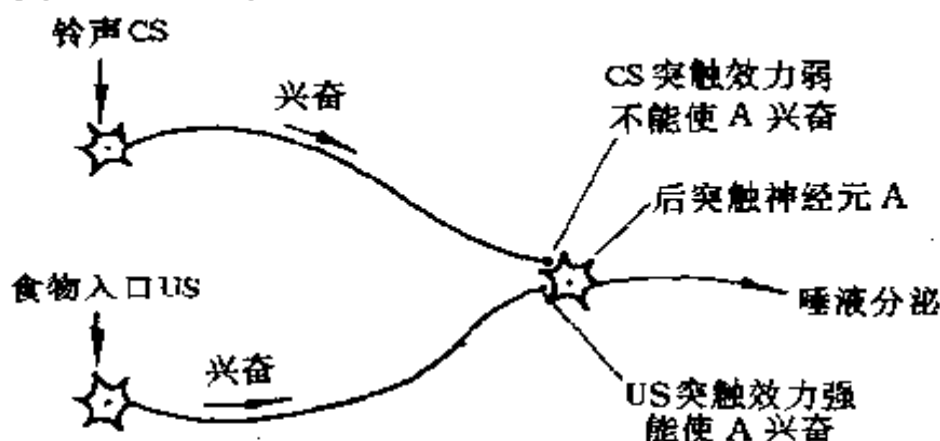


图 5 异突触易化机制的示意图

3. 长时的加强作用(long-term potentiation)

(1) 长时的加强现象

这是使后突触神经元的反应性改变的一种现象。一般的学习都需要感觉刺激来引起神经活动，这种活动的结果是神经元的特性的改变，这种改变无疑也有结构改变的因素。人们已认识到这种变化是很微妙的，不容易在整个脑内查出来，所以研究者们就转向用电流刺激神经系统的确定部位，期望能引起可测量到的物质变化。因为电刺激可能比自然的感觉刺激引起更强的神经活动，而且可以局限在一个特别的部位，这样就能确定要观察的准确地点，从而有可能发现这个地方的变化性质。长时的加强现象就是这样发现的。

1966 年 T. Lømo 看到，如果先在 7 至 10 秒之间，以上百个电脉冲刺激进入海马结构的轴突束，20 分钟后再以单个电脉冲刺激

这些轴突,从海马的齿状回中记录到的棘波电位就会比前面没有高频刺激时的棘波高出 250%。这种效果可维持很久,因此这一现象就叫作长时的加强作用(见图 6)。以后用微电极记录齿状回中的单位神经元的反应,确证这是由于增加了后突触神经元的兴奋性,也就是说,传入的轴突和齿状回的神经元之间的突触效力加强了。

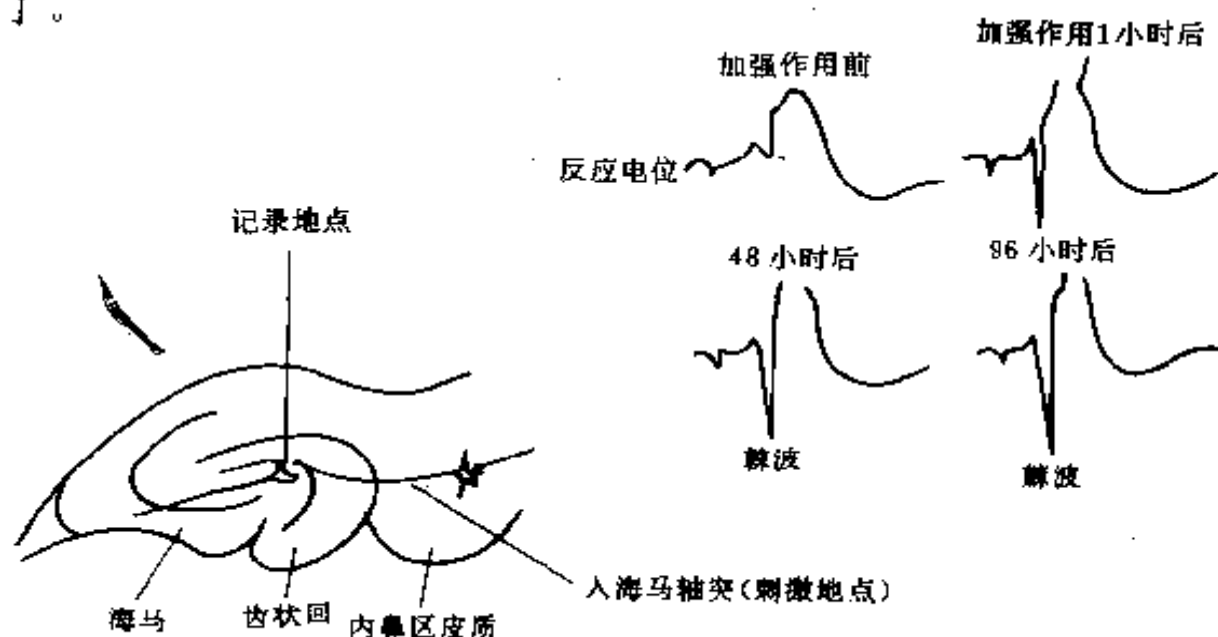


图 6 加强作用的刺激和记录齿状回中细胞群的反应电位变化的示意图

长时的加强作用可保持几个月 (T. V. P. Bliss and T. Lømo, 1973)。这种特性类似学习,即由短时爆发的神经活动引起了这个神经线路中的持久的变化。

长时的加强作用也能在海马的活切片中产生。在这种情况下,研究者能够更确切地看到加强作用发生的地点。此外还可以改变脑组织的培养液中的离子和化学成分以观察其影响。

(2) 长时加强的生理机制

有几位研究者曾用长时加强现象作为异突触易化的证明,以支持 Hebb 提出的长期记忆的细胞基础的假说。例如 W. B. Levy

和 O. Steward(1983)曾用经典条件作用的刺激模式来研究长时加强能否导致异突触的易化。已知,齿状回只接受很少的对边内鼻皮质的投射,所以如果只刺激一边的内鼻皮质不能使对边的齿状回产生足以导致长时加强的兴奋。两位研究者在观察一侧齿状回的兴奋后突触电位加强现象时,先用密集的电脉冲同时刺激两边的内鼻皮质;或先刺激对侧的内鼻皮质,20 毫秒再刺激同侧内鼻皮质。这两种方法都可促成在单脉冲刺激对侧内鼻皮质时其齿状回内的兴奋后突触电位的加强,暗示发生了异突触的易化。更有趣的是,如果先刺激同侧内鼻皮质,或在两侧内鼻皮质刺激过后 200 毫秒再刺激同侧内鼻皮质,都不会发生加强作用。这一结果证明一群神经元上的某些突触的活动可以增加另一些突触的效力。但原来效力强者必须与弱者同时,或在弱者稍后活动才有此结果。这与经典条件作用所要求的条件刺激和无条件刺激呈现的时间次序非常相似。

上述结果也暗示,异突触易化乃是后突触神经元中发生变化的结果。海马结构中的主要的兴奋性神经递质似乎是氨基酸物质——谷氨酸(glutamic acid),与长时的加强现象有关的后突触变化似乎是后突触膜上的谷氨酸受体数目的增加。在离体的海马切片上的研究证明,在刺激与海马的神经元有突触连接的轴突之后,出现了长时加强现象的切片中,其突触膜的谷氨酸受体数目比未经刺激的切片中的增加了 30%(G. Lynch, et al., 1982)。研究者在离体的海马切片中的进一步的研究发现,造成长时加强现象的重要环节是钙离子(Ca^{2+})进入细胞浆内(G. Lynch and M. Baudry, 1984),钙进入细胞引起谷氨酸受体增加。研究者设想钙使一种特别的蛋白酶活化,被钙激活的这种蛋白酶不可逆转地分裂氨基酸链,从而改变了蛋白质的结构。可能的情况是蛋白酶将后突触区域的神经膜里面的一层叫作 fodrin 的蛋白质撕裂成多股的、股间的裂隙,将以前被盖着不能活动的谷氨酸受体释放出来。

(3) 长时的加强与学习的关系

长时加强现象的生理机制如果是由于刺激引起的后突触膜的谷氨酸受体的增加,如果这也可能是某种学习的生理基础,那么长时的加强和学习应当是能互相影响的。有些研究的结果证明确实如此。例如,当动物学习条件反应时,海马内的神经元便活动起来。研究者们发现条件反应的训练也影响海马结构中的突触效力。如 D. J. Weisz, et al., (1982) 发现,在经典条件作用的瞬膜反应的训练期间,齿状回的神经元群的棘波幅度两倍于训练前的幅度。此外, R. F. Thompson 等 (1983) 还发现了从接受训练的家兔的海马中分离出的突触膜上的谷氨酸的结合增加。这可以说是学习引起了海马内神经元的结构和反应特性的改变。又如,已知动物对刚刚去过的地方的记忆依赖海马的工作,海马损伤后即失去这种记忆能力。人们由此可以设想,如果以有效的方法阻止海马内神经细胞的谷氨酸受体增加也将会影响动物的这种记忆。G. Lynch, M. Baudry (1984) 训练大鼠在 8 臂的放射形迷津中觅食,学习在一次练习中不重复进入已在其中吃过食物的臂。与此同时,通过植入大鼠脑中的微型导管给其侧脑室中注入一种化学物质——leupeptin,这种物质能阻止被钙激活的蛋白酶的活动,使之失去解放被盖着的谷氨酸受体的作用。实验中看到,接受了 leupeptin 注射的大鼠像海马损毁的大鼠的表现,常常重复进入已经进入过的臂。这一结果显示海马在学习中的作用可能包含着长时加强的机制。

4. 神经元的结构生长的可能性

长时记忆的神经元基础的变化可能还包含着神经元树突分枝的增加和树突刺数目的增加,当然也还有其他方面的生长变化,如轴突变粗和突触长大等,然而结构生长是一种比较缓慢的变化,在短期的实验中难以发现,而且在实验中给动物设计的学习任务往往涉及到脑的许多区域的神经元的不同程度的活动,不容易像直

接刺激进入海马的轴突那样,可以确切地知道发生变化的脑区。因此研究者绕过了这一困难,去观察有更多的学习机会的生活环境对脑细胞生长的影响。

M. R. Rosenzweig 及其同事们(1962),曾设计了一种实验,观察丰富和贫乏的生活环境对大鼠脑发育的影响。他们将刚断乳的同窝鼠分成 3 组:一组放在大的,有转轮、梯子、滑梯、各种玩具和同伴(或 8 只鼠)的笼中饲养——即在生活丰富的环境中成长;另一组单个的在小笼中饲养,笼内除食和水外别无他物——即在生活贫乏的环境中成长;第三组为对照组,在实验室中常规的饲养笼中成长。在这 3 种条件下生活的大鼠长到 80 天或更老些时,皆被杀死,取出脑进行解剖和生化的测量,发现了有趣的结果。

① 生活丰富鼠的枕叶皮质的厚度最厚,生活贫乏鼠的最薄。

② 枕叶皮质中的锥体细胞的树突分枝、树突刺和树突的大小,在生活丰富鼠中,都是最多和最大;相反,在生活贫乏鼠中则是最少和最小。

③ 大脑皮质中胶质细胞的数量在生活丰富鼠中为最多。

④ 大脑皮质中的胆碱脂酶(AchE)的含量也以生活丰富鼠的最多。

神经元的结构和生理的这些变化经过控制的条件更为严格的实验分析确信是由于生活丰富的环境提供了更多的学习机会造成的(E. L. Bennett, et al., 1979)。由此也可以认为长时记忆的基础也包含有神经元的结构生长的因素。

主要参考文献

- [1] Kesner, R. P. and DiMattia, B. V., Neurobiology of an Attribute Model of Memory, In Progress in psychobiology and Physiological Psychology, A. N. Epstein and A. R. Morrison eds., 1987, Vol. 12, 207—277.
- [2] Carlson, N. R., Foundations of Physiological Psychology, 1988, Allyn

and Bacon, Inc. Boston, 302—402.

- [3] Brodal, A. , *Neurological Anatomy*, 3rd Ed. , 1981, Oxford Univ. Press, London.
- [4] Rosenzweig, M. R. and Leiman, A. L. , *Physiological Psychology*, 1982, D. C. Heath and Company, Lexington, 593—636.
- [5] Kolb, B. and Whishaw, I. Q. , *Fundamentals of Human Neuropsychology*, 1980, W. H. Freeman and Company, San Francisco.

信息加工与脑能量函数

沈 政

(北京大学心理学系)

本世纪 40 年代信息论刚诞生不久,心理学家们就开始以信息加工的概念改造传统心理学,60 年代建立起当代心理学发展的主干——认知心理学。70—80 年代间认知心理学将人类认知活动中信息加工区分为两个性质不同的过程:自动过程(automatic processes)与控制过程(control processes)。前者又称为非随意的(involuntary)、不耗心神的(effortless)、心理资源无限的(unlimited mental resource);后者则称为随意的(voluntary)、耗费心神的(effortful)、心理资源有限的(limited mental resource)。这种分类原则对当代心理学的影响十分广泛,已渗透到工程心理学、认知心理学、情感心理学和心理生理学等许多研究领域。80 年代,在行为神经科学中出现了一个新的研究领域,即利用脱氧葡萄糖技术和正电子发射层描技术揭露了脑区域性葡萄糖代谢率与行为模式之间的关系,开展了脑能量代谢与心理活动的相关研究。这一时期脑模拟研究领域出现了人工神经网络的新模型,并概括出计算神经网络功能和结构特性的能量函数。因此,将心理学、行为神经科学和人工神经网络理论中的这些概念加以比较,找出其间的内在联系进行跨学科的研究是本文的基本出发点。

本文不但介绍上述 3 个领域中的研究进展,并讨论其间可能存在的关系,还试图简要地介绍作者实验室近年在这一研究中,所得到的一些初步结果和今后发展的理论设想。希望本文内容作为作者 4 年前起草的“多学科间发展的生理心理学”^[1]一文的续集,

增进读者对我们的了解,在促进我国心理学发展中起到抛砖引玉的作用。

一、信息加工的自动过程和控制过程

Posner 和 Snyder(1975)最先提出了关于信息加工中的自动激活概念,指出这一概念的 3 个操作标准^[2]:不是在意识控制下进行的,主体对此过程一无所知,该过程并不干扰其他心理活动。Posner(1978)进一步指出自动激活的信息加工过程是以往学习的结果,是内在编码及其联结在重复刺激作用下的激活^[3]。他还提出,与此相对应的是注意过程,其特点为容量有限性(limited capacity)。Schneider 和 Shiffrin(1977)明确而完整地提出了信息加工中控制过程和自动过程的概念^[4,5],用以对选择性注意,短时记忆搜索和视觉搜索等心理过程进行综合解释。Schneider 等(1984)进一步指出^[6],自动过程是一种快速的并行传入过程,也是一种不费心神、不受短时记忆容量限制、不受主体意识直接控制的加工过程。自动加工过程是主体对同一刺激多次重复应用而发展起来的,是熟练技巧行为的基础。与此相对应的控制加工,是一种缓慢的串行传入过程;也是耗费心神、容量有限的加工过程;又是主体对不断变化的刺激进行反应的过程。因此,控制过程由主体随意加以调节。Kahneman 和 Treisman(1984)明确总结出自动加工过程的标准^[7]:非随意性,不需意志参与即可自动开始,一旦开始也无法随意终止;自动加工过程不耗费精力,它既不受其它随意活动的干扰,也不干扰其它随意活动。几种自动加工过程可以同时并行性进行,彼此没有干扰,没容量限制,无意识地进行着。

我们首先概括性介绍自动加工和控制加工过程的这些标准和特点在注意、记忆、知觉理论研究中的反映,再讨论它们在心理生理学和工程心理学中的影响,着重阐明心理资源的含义及其在心

理学理论发展中的意义。

1. 认知活动的两种加工过程

首先,两种加工过程的概念在 70 年代中期带来了注意理论的突破性进展。70 年代中期以前,心理学中关于注意研究的理论,占主导地位的是瓶颈说(Bottleneck),利用滤波实验方案(filtering paradigm)或双重刺激方案(double-stimulation paradigm),通过视觉或听觉系统对被试呈现大量字、词,在大量记忆负荷的背景下,测定反应的正确率。这类实验所积累的事实表明,选择性注意妨碍或降低对未注意刺激的知觉加工过程。据此,把注意看成是大量刺激信息在单一通道竞争过程中,发挥选择作用的心理过程。这种选择作用发生在知觉信息加工的早期,故又称为注意的早期选择模型(early-selective model)。与这种模型相对应的是晚期选择模型(late-selective model),认为注意的选择作用发生在对刺激进行反应的决策过程,即在产生知觉以后和给出反应之前,注意发生重要作用。Kerr(1973)把这两种注意模型融合为一体,提出了单一有限容量的中枢处理器模型(a single limited-capacity central processor, LCCP),认为在同一瞬间,只能很好地进行一种心理活动,其它心理活动或次级作业任务不可能很好地完成^[8]。LCCP 可以从记忆中提取信息、传递刺激、反应决策或形成假设;但在某一瞬间只能完成一项,多项任务只能分时地串行进行。所以,这是一类有容量限制的有意识的注意(conscious attention)。Posner 和 Snyder(1975)提出,除了这类有意识注意过程之外,还存在一种无意识的自动激活过程。Kahneman 和 Treisman(1984)将这类过程按自动性程度分为 3 类:①强自动性知觉过程既不为集中注意所增强,也不为分散注意所妨碍;②部分自动性过程,虽然它能在分散注意时完成,但集中注意力可使该过程的作业成绩更好;③偶然的自动性过程,一般需要注意参与,但偶尔不注意时也能正常进行

下去。70 年代中期以来,在注意研究中更多采用选择集实验方案(selective-set paradigm),在被试记忆负荷较小的背景下,测定被试对少数刺激物或刺激物某一特征进行选择反应的反应时。这类研究积累的事实表明,选择性注意能显著加速对期望目标的反应。这类研究还发现了 3 个基本事实:非注意材料的语义加工、分类效应和阈下知觉。这 3 个事实对注意、知觉和记忆理论的发展提供新的线索。就注意理论而言,这些事实证明,在非注意通道中,阈下知觉刺激的加工对注意通道的知觉没有干扰作用,但在知觉形成中,非注意通道的加工往往补助注意通道的知觉加工,使客体的某些特征不致漏掉。利用 Stroop 效应的实验,有力地证明了这一点。语词材料的语义加工过程可以在非注意通道中进行。总之,选择注意的控制过程和非注意的自动过程,是并行互补的信息加工过程。自动加工的通道是多重并行,而且互不干扰,没有容量限度。同样,容量有限的选择注意通道也可以进行多重并行加工,只不过根据加工深度要求的不同,对有限注意资源(attentional resource)进行适当分配,形成了对多项任务的分配注意。由此可见,70 年代中期以前,注意理论主要是单一通道滤波理论或称瓶颈理论,而 70 年代中期,由于两种信息加工过程的概念使注意理论变迁为多通道理论,即包括有限注意资源分配的选择注意通道,又包括不耗费资源的自动加工通道。两类通道相互补充和转化的动态过程,就是注意变换的动力过程。

Hasher 和 Zacks(1979)首先提出将自动过程和耗费心神的过程(automatic and effortful processes in memory)作为研究记忆的理论框架^[9]。他们将那种从有限容量的注意机制中吸取较小能量(minimal energy)的心理操作称为自动过程。它的发生不干扰其他认知活动的进行。他们又区分出两类自动过程:一种制约于遗传性的自动过程,没有年龄、性别和个体差异,对信息流的基本特性、物体名称、空间、时间、发生频率等进行自动加工;在各种环境条件

下,总是以恒定水平进行着。另一类自动加工过程制约于学习和实践,随着学习和实践而不断提高的加工过程是熟练技能的重要基础。与两类自动加工过程不同,耗费心神的过程又可称为意识的控制过程,是练习和精细的记忆活动,它的发生常常干扰其他认知过程。Johnson 和 Hasher(1987)将自动过程又称之为非意识的、无策略的加工过程和无策略加工的记忆(memory without strategic processing)^[10]。有意的学习(intentional learning)与无意的学习(incidental learning),外显的或自觉的记忆测验(explicit)与内隐的或不自觉的记忆测验(implicit)之间的不一致,近年已成为记忆研究的焦点^[11]。通过重复启动效应,即对刺激的加工由于受到先前经历的同一刺激的影响,而得到促进的现象,揭露了内隐记忆的许多规律^[12]。Friedrich, Henik 和 Tzelgov(1991)研究了词汇存取中的自动过程。他们认为在语义启动机制中,至少有 3 个成分:①视觉的、语音的和语义表征间的编码存在着固有联结性,自动激活或扩散性激活这种联结性就会造成完全自动的启动效应;②主词相关的目标期望集(a set of expected targets related to the prime word)所引起注意过程参与下,可能实现一种非自动启动机制(nonautomatic priming mechanism);③词汇存取后的语义匹配过程(postlexical semantic-matching process)。这 3 种机制在语义启动效应中发生复杂的相互作用^[13]。扩散激活的完全自动的启动效应,对记忆过程来说是一种快速的易化机制;基于目标期望集的非自动启动机制需要注意分配变化的时间,所以是一种慢过程,既包括相关词提取的易化过程,也包括无关词提取的抑制过程。自动的和非自动的语义启动机制均发生在词汇存取(lexical access)过程之前,启动词和目标词之间的语义匹配,则是词汇决策之后发生作用的机制。由此可见,对于语词记忆启动效应的研究,引出自动加工过程和非自动加工过程间的复杂关系,是当前记忆研究中的核心问题。

两种加工过程的概念在知觉理论中的意义,可以通过 3 种知觉理论加以分析,即特征结合理论、RBC 理论和拓扑计算理论。特征结合理论由 Treisman 和 Gelade(1980)提出^[14],10 年后加以修正^[15],近年仍在深入研究继续完善之中^[16]。这一理论将知觉形成过程分为两个阶段:前注意(pre-attentive stage)和注意阶段(attentional stage)。前注意阶段对物体各种特征进行搜索,各种特征形成多维向量,如颜色维度、方向维度、位置维度等。这种搜索过程不需注意参与,因而是自动加工过程。注意参与下的串行加工过程,可将各维度上的特征加以结合,实现了特征结合的目标。注意集中参与的控制过程,才能很好地实现特征结合,否则注意分散将会造成知觉模糊或错觉(错觉性特征结合)。当某一维度上的目标或特征非常明显,则注意引导的搜索过程很快指向该特征。目标必须定位地与其他特征分离出来,才能保证不发生错误结合。否则在许多搜索任务中,特征结合可能是伪装的或错误的。由此可见,这种知觉理论强调资源有限的注意过程,在特征结合中的重要作用,而自动加工过程只为控制过程提供可选择性结合的素材。在 Treisman 特征结合理论的基础上,Duncan 和 Humphreys(1989)明确地将视觉搜索的加工过程分为 3 个阶段:不受资源限制的并行性自动加工过程、竞争性匹配过程和视觉短时记忆阶段。第二阶段是注意资源有限的过程。可见,后两个阶段都是控制加工过程。对于不受注意资源和记忆容量限制的自动激活的无意识知觉过程,又称为阈下知觉。它的研究大大地丰富了知觉理论^[18]。

RBC 理论(recognition-by-components, RBC)是 1987 年由 Biederman 提出的^[19,20],将物体识别或辨认分为 5 个阶段:边缘检测、特征分离、几何子(geons)及其关系的激活、模板激活匹配和物体确认。前 3 个阶段都是自动加工过程,后 2 个阶段是记忆和注意参与的控制加工过程。在传统拓扑心理学、现代认知心理学,神经网络的并行分布处理原理、能量函数概念和统计数学等多学科理

论指导下, McCafferty(1990)提出视觉拓扑计算理论^[21], 主张在视觉信息加工早期阶段中, 即有高层次知识的参与, 有助于初始态的确定; 视觉信息加工的全过程都贯穿着节能的原则, 视知觉形成过程就是能量的组合优化过程。他对 Triesman 特征结合理论的实验数据进行分析, 证明前注意并行加工阶段的客观性。

从上面讨论的注意、记忆和知觉理论中, 我们可以看到自动加工过程、控制加工过程、资源(resource)、容量(capacity)、能量(energy)、心神耗费(effort)、意识(consciousness)、知识(awareness)等名词概念虽各有不同, 但基本以资源是否有限为重要差别, 这种差异不但体现在认知活动, 也体现在作业的操作过程和意向活动之中, 这里更能体现出心理资源的含义。

2. 作业成绩与心理资源

心理资源、容量或能量等(resource, capacity, energy)概念最早出自于 James(1890)的著作《心理学原理》。在 50—60 年代中, 认知心理学和工程心理学均对这些概念赋予新的含义。如前所述, 在认知心理学中, 通常指注意和记忆的容量有限; 在工程心理学中, Knowles(1963)提出了人类作业的操作模型。他认为, 作为操作者的人好像是容量有限的资源库(limited capacity resources), 面对一些工作任务时, 这种资源可以分配地利用^[22]。70—80 年代, 工程心理学热衷于测量不同工作负荷时已经利用的和可以利用的资源以及这种测量的主观心理参数和客观生理指标。Kahneman(1973)明确指出, 注意的容量模型是指完成心理操作时可以利用的能量^[23]。Norman 和 Bobrow(1975)分析了双项任务作业中容量或资源分配问题, 提出了作业-资源函数中资源限定(limited-resource)的加工过程和资料限定(limited-data)的加工过程^[24]。Wickens(1984)在总结资源或容量概念发展的历史资料时, 概括出单源论(single-resource theory)和多源论(multiple-resource

theory)的理论观点,指出工作负荷测验必须综合地分析主任务、次任务参数(primary and secondary task parameters)、生理测量和主观心理测验等4项结果,并综述了许多致力于寻求心理容量、资源、心理能量与生理和物理能量间相互关系的研究报告^[25],他指出瞳孔、心率、区域性脑血流量、脑区域性葡萄糖代谢率和平均诱发电位等生理参数均可作为心理资源分配的重要生理学参数指标。同时他也指出没有一项生理指标可以完全代替心理资源的全面分析。Brown(1982)指出心理负荷(mental load)是多维度的,任务要求可能是感觉、知觉、注意、知觉运动等多方面的,随时变化的;心理负荷和心神耗费(mental effort)并非完全一致,前者常用任务要求的难易加以度量,后者用执行任务者主观努力程度加以度量。然而,面对同一难度的任务,不同人可能采用不同策略,耗费心神的程度也相差很大。除了个人能力、技能不同,还有动机因素。总之,心神耗费、任务要求和动机水平3者相互制约,决定了心神耗费的测量是十分困难的问题,必须同时进行任务分析、主观评定,还应对双重任务中心理资源的分配和适当的生理指标等加以综合分析。Eysenk(1982)认为,动机、唤醒水平、人格特质等多种变量,对于完成某项作业所需的心理资源或耗费心神的程度均有一定关系。对引诱性刺激作用的分析是研究这种复杂关系的较好途径。他概括了这类刺激对作业的8种影响:①它可以增强选择性注意这一控制加工过程;②加速内部信息加工过程和外部的行为反应率,但往往降低作业质量;③影响内部动机状态也常引起动机水平的波动或焦虑状态;④常常不利于作业的持久性改善,因为它能降低本能的动机水平;⑤较强的引诱性刺激降低并行共享的信息加工过程(自动加工过程);⑥增高唤醒水平;⑦与作业成绩的关系是曲线性的;⑧可能增加注意的涣散力。在这8项作用中,较为重要的是信息并行加工过程的降低,可导致信息的有效利用率降低,这是引起作业成绩变差的主要原因。刺激与作业成绩的曲线关

系,可能最初改善作业成绩是由于其增强选择性注意的信息控制加工过程;随后由于它引起自动加工过程的并行处理变差以及动机的波动或焦虑状态,因此又引起作业成绩的下降。

3. 两类加工过程与脑事件相关电位

80年代以来,由于平均诱发电位记录和分析技术的广泛应用,直接用于研究人类认知活动中神经信息加工过程的脑功能,并积累了丰富的科学资料。因此,认知心理生理学(Cognitive Psychophysiology)已成为心理学中较活跃的研究领域。对于认知活动中自动加工和控制加工过程的研究,以及心理资源分配问题的研究,脑事件相关电位(ERPs)分析是较有成效的研究领域。

Hink 等(1977)首先报道了支持资源有限论和资源分配理论的 ERPs 结果^[28]。他们发现 ERPs 中的 N_1 波的波幅可以作为心理资源分配的生理指标。在注意分配的条件下, N_1 波幅介于集中注意时波幅和不注意时波幅的中间值。Para-suraman(1978)也得到同样的实验数据^[29]。然而,80年代以来,更多的研究报告表明, P_3 波比 N_1 波对于心理资源分配具有更可靠的一致性变化。Näätänen(1988)指出 N_1 波对于注意的启动是一项较好的生理指标,当刺激呈现或长时间呈现后停止时, N_1 波幅明显增强^[30]。

在综述大量研究报告基础上,Näätänen(1988)提出,对刺激物理特性的自动加工过程总是伴随 ERPs 的不匹配负波幅值增大(mismatch-negativity,MMN),因此,他认为 MMN 波可以作为信息自动加工的指标,而选择性注意的控制加工过程更多地反映为加工负波(processing negativity)的变化。虽然 MMN 和加工负波的潜伏期较为接近,但后者伴有 P_3 波的同时性变化,表明加工负波反映了语义加工和选择性注意过程的脑功能。对于 ERPs 理论来说,MMN 是外源性 ERPs 成分,加工负波则是内源性 ERPs 成分。前者与自动加工过程有关,后者与控制加工过程有关。Coles

等(1986)的著作中,总结 ERPs 在心理生理学体系建立中的作用时指出,ERPs 研究不仅是认知活动脑功能之窗,也可以作为工程心理学测定工作负荷的有效工具^[31]。

我们已经从认知心理学、认知心理生理学和工程心理学等角度讨论了信息加工的自动过程和控制过程的概念,这一概念的应用,导致了心理资源有限论和分配论的新理论。既然心理资源和耗费心神是两类信息加工过程间的主要分界线,那么究竟心理资源的物质基础是什么呢?一些心理学家已明确使用心理资源分配,就是可以利用的能量(energy)变化^[32]。因此,我们从生理学或神经科学的角度探讨脑能量代谢、神经信息加工与行为模式之间的关系,可能有助于我们对心理资源和两种信息加工过程的理解。

二、神经信息与能量代谢

虽然从发展智能化计算机的目标出发,对大脑的模拟研究已有半个世纪的历史,近 10 年来却发生了巨大变化,以致近年来已形成了“计算神经科学”(Computational Neuroscience)的分支。计算神经科学的研究目标是,揭露脑的电信号和化学信号如何表达与加工神经信息,及其在智能活动中发生变化的规律^[33],要阐明人类和动物的神经系统是怎样使用它的微观组件及其相互作用来表征和处理信息的^[34]。因此,在这里我们首先弄清神经信息的电学表现形式、化学表现形式及其与相伴随的脑能量代谢之间的关系,然后再讨论它们与行为模式或认知活动的关系,以便加深对人类信息加工过程的理解。

1. 神经信息

神经信息既有电学表现形式,又有多种生物化学形态。在整体水平上,电学表现形式呈现出大脑多种形式的电活动,包括自发电

活动和诱发电活动、交流电变化、慢电变化和直流电变化^[35]；在细胞水平上，其电学表现形式可分为两类：一类是符合全或无规律的调频式神经脉冲，包括神经干上的动作电位(action potential)和神经细胞的单位发放(unit discharge)；另一类是级量反应(graded response)，包括感受器电位、突触后电位和后电位，其中突触后电位又包括兴奋性突触后电位(EPSP)和抑制性突触后电位(IPSP)；后电位包括后兴奋电位(AEP)和后超极化电位(AHP)。除了上述这些受到刺激后的即时性生物电反应外，最近十多年，神经生理学研究发现了多种长时程效应，如长时程增强效应(LTP)、长时程抑制效应(LTD)、输入输出转换性长时程增强(ITTO-LTP)和长时程抑制效应(ITTO-LTD)。由此可见，神经信息可根据出现的部位、性质、时程等标准，区别出多种电学表现形式：细胞体和轴突上、突触前和突触后成分、全或无和级量反应、后电位与突触后电位、即时效应和长时程效应等等不一而足。这足以说明神经信息的复杂性。然而，如此多种形式的生物电信号，却都产生于神经细胞的膜上，因此，80年代以来片膜-钳技术(patch clamp)在神经生理学界得到普遍重视，揭露了细胞膜上离子通道的功能特性与神经信息的关系。分子生物学发现^[36]细胞膜是具有流动性的双层脂质膜，跨膜内外有许多大蛋白分子在神经信息的传递和转换中发挥重要作用，其中离子通道蛋白分子和受体蛋白分子尤为重要。这些大蛋白分子在膜内外穿行几次，尚有游离端在膜内外浮动，每个分子都由许多亚单元组成。在神经信息传递中，突触前末梢释放神经递质，作用于突触后膜与其膜上的受体蛋白发生结合，引起受体蛋白分子发生变构作用，其中一些受体蛋白又具有离子通道的功能。另一些受体蛋白中介于鸟苷三磷酸(GTP)相依性蛋白(G-蛋白)的作用，引起一系列细胞信使的传递作用，最终也使离子通道蛋白分子变构。所以，神经信息的生物化学表现形态也是多种多样的，有几十种小、中分子量的神经递质、神经调质、多种受

体蛋白分子、细胞内信使(第二、三、四信使)和离子通道蛋白分子。这样,神经信息就在电学表现形式和化学表达形态间发生顺序性和并行性的交替变换,实现着信息传递、放大和转换等加工过程。

2. 脑能量代谢

在神经信息加工过程中,到处都伴随着能量代谢的变化,离子泵逆离子浓度梯度将一些离子由低浓度送往高浓度的过程需要耗能,神经递质、调质、受体蛋白、离子通道蛋白、G 蛋白、细胞内信使等活性物质的生物合成都需要能量。哪里神经信息加工活跃,那里的能量代谢率就最高。因此,脑区域性能量代谢率也是神经信息加工和认知过程的灵敏尺度。对脑信息与脑能量关系的这种概念,是 80 年代以来在许多科学事实基础上形成的。

重量只占体重 2% 的脑,其耗能量却达全身的 20%,脑能量代谢率高居全身各器官之首。80 年代神经生物学采用脱氧葡萄糖组织化学法和正电子发射层描(PET)的活体测定,已积累相当多的科学事实^[37],说明认知过程中脑区域性糖代谢率发生动态变化。Sokoloff(1984)系统总结了脱氧葡萄糖组织化学法所发现的科学事实^[38],在视皮层中糖代谢率随光刺激强度呈对数关系而增加,与心理物理学中的费希纳定律十分吻合。双眼视觉剥夺,两半球视皮层糖代谢率对称性降低。单眼视觉剥夺后,由于在两侧皮层和外侧膝状体糖代谢率高低不同,呈黑白交替的条纹状组织图像。证明了糖代谢率的研究发现与 Hubel 等人电生理所发现的眼优势柱完全吻合。在人类被试中,注入¹⁸F-脱氧葡萄糖,利用 PET 测定技术,发现用简单黑白反转棋盘格作为刺激物时,17 区初级视皮层糖吸收率最高,18 区次级视皮层次之;改用复杂彩色风景图为刺激物时,则 18 区糖吸收率最高,而 17 区次之。当用语言刺激时,被试左半球糖吸收率高于右半球。音乐刺激时,对音乐行家来说,音乐知识必然参与其认知活动,引起左半球额-颞区糖吸收率高于右

半球,可能控制性信息加工过程正是在这里最活跃;对于外行,则引起右半球额-颞区糖吸收率的增高。由此可见,糖代谢率可以作为认知过程信息加工水平的灵敏指标,反映出认知过程的生理基础。

大脑的能量代谢,99%以血液供应的葡萄糖为燃料,经糖酵解、三羧酸循环和氧化磷酸化 3 个阶段,将氢和电子由葡萄糖传递给氧分子生成水,67%的能量由三磷酸腺苷(ATP)负载,33%的能量变为热噪声损耗。所以脑能量代谢的效率为 67%,是任何热力机所无法比拟的。这是由于在脑能量代谢过程中,葡萄糖的能量并不直接传到 ATP 分子中,二核苷酸类辅酶作为电子、氢传递体的同时,也发挥能量传递的作用。每摩尔葡萄糖分子经过糖酵解和三羧酸循环,将绝大多数有效的能量转移到所生成的 10 摩尔的还原型辅酶 I (NADH)和 2 摩尔的还原型黄素辅酶(FADH_2)分子中。10 摩尔 NADH 和 2 摩尔 FADH_2 在随后的氧化磷酸化过程中,依次将氢和电子传递给辅酶 Q,细胞色素 b, c_1 , c, $a_1 + a_3$ 等,最终达到氧分子。氧化磷酸化过程发生在细胞内的线粒体中,1 摩尔 NADH 被氧化释放的能量,形成 3 摩尔 ATP;1 摩尔 FADH_2 被氧化时,可形成 2 摩尔 ATP。所以,脑内的还原型二核苷酸辅酶的含量变化,可以作为脑能量代谢的有效指针。它有着天然的荧光特性,吸收 340nm 紫外光,可发射出蓝色荧光,既可用吸收光谱法测定,又可用荧光发射光谱测定。生物化学家们早就利用这种特点,作为测定脑内各种氧化还原酶活性的重要方法学基础。

3. 各种行为模式中的二核苷酸类辅酶

二核苷酸类辅酶包括烟酰胺腺苷二核苷酸(辅酶 I, NAD), 烟酰胺腺苷二核苷磷酸(辅酶 II, NADP)和烟酰胺黄素二核苷酸(黄素辅酶, FAD)。它们是生物能量代谢中的重要辅酶,这类酶的二态特性,即氧化型与还原型之间的变换,发挥氢、电子和能量传

递的重要作用。自 1984 年以来,我们实验室采用吸收光谱和发射光谱法,研究了跳跃行为、穿梭行为、操作行为和多行为中鼠脑各区二核苷酸类辅酶的含量变化,以此为方法学基础,研究区域性脑能量代谢与行为模式的关系,以及一些药物作用的行为药理学机制。

我们首先选择与能量代谢关系较为密切的跳跃行为^[39],分别将 0.1, 0.5, 2.5 $\mu\text{mol/ml}$ 的 NADH 注入鼠大脑皮层,结果发现 NADH 增强跳跃行为的剂量相关效应,随 NADH 剂量增大,跳跃的次数和高度增加。精神运动兴奋剂苯丙胺可以增加脑内 NADH 的含量,为脑代谢传递较多的能量,所以能显著增多大鼠的跳跃次数^[40]。咖啡因是一种常见的中枢兴奋剂,增加鼠脑内 NADH 含量的效应弱于苯丙胺,作用时间也较短,故对跳跃行为的影响不如苯丙胺强。苯丙胺在海马中引起 NADH 含量增高的同时,可促进大鼠在穿梭箱中的非条件性逃跑反应和条件反射性躲避反应,并缩短反应时^[41]。60mg/kg 和 120mg/kg 的咖啡因对鼠脑 NADH 和 FAD 的影响不同^[42]: 60mg/kg 的咖啡因使 6 个研究的脑区 NADH 含量普遍增加,在小脑和脑干内的 FAD 含量也显著增加; 120mg/kg 咖啡因不再引起 FAD 含量的增高,间脑内的 NADH 含量也不再增高^[43]。与此相应,120mg/kg 咖啡因对穿梭和操作行为均有抑制效应,30mg/kg 咖啡因可明显增强穿梭行为,60mg/kg 时可明显增强操作行为^[44]。总之,这 3 种行为模式中,无论是苯丙胺还是咖啡因引起脑内 NADH 含量增高与脑内注入 NADH 具有相似的行为效应,能源物质 NADH 含量增高伴有信息加工过程的增强,产生明显的行为效应。

除上述药物与行为实验之外,我们还利用光导纤维将弱激光导入鼠大脑皮层,观察激光诱导出的脑事件相关电位及脑内二核苷酸辅酶的变化^[45]。结果表明,与眼前闪光刺激相比,脑内弱的脉冲激光可使 ERPs 中 N_2 波和 P_3 波幅值增高,此时在大多数研究

的脑区内 NADH 和 FAD 含量均明显降低,说明激光普遍提高了脑能量代谢率,可能是引起 ERPs 幅值增高的分子生物学基础。由此可见,脑能量代谢与 ERPs 之间存在着密切关系。如前所述, N_2 波可能是 MMN 波的一种,也可能是加工负波,决定于是否伴有 P_2 波的同时变化。我们的这一研究,发现脑能量代谢率增高伴有 N_2 和 P_2 波的变化,说明这是耗能的控制性信息加工过程。

综上所述,神经信息的电学和化学表达形式总伴随着能量代谢的变化;反之,改变能量代谢或测定能量代谢的动态变化,总能发现其与行为或认知活动的相关变化。所以,脑能量代谢与脑神经信息和心理资源之间必然存在着内在联系,能量与信息加工及实现信息加工的神经网络的结构与功能之间的关系是认知心理学、神经科学和信息科学跨学科研究的重大课题。由于人类信息加工和脑能量代谢的关系较为复杂,如何着手这一重大课题,是近几年来引起我们反复思考与探求的问题。与人类信息加工和脑能量代谢相比,电子器件搭成的人工神经网络较为简单。正是在这类简单网络研究中,80 年代却发现了通过计算其能量函数对网络结构和信息加工能力,进行定性与定量的预测。因此,吸收人工神经网络的能量函数概念,对于加深理解人脑信息加工过程是有益无害的。

三、人工神经网络的能量函数

神经网络是类似生物脑或神经系统的网络模型。人工神经网络的研究从 40 年代起,半个世纪来经历了曲折的 4 个历史时期^[45],现已掀起第二个发展热潮。美国著名学者 Hopfield(1982)提出的神经网络模型及其能量函数^[46],拉开了第二个高潮的序幕。

1. 网络的数学模型

人工神经网络的数学模型包括两个方程:信息存贮功能的网络学习方程和信息提取功能的网络动力方程:

$$(1) \text{ 权矩阵方程 } \begin{cases} T_{ij} \approx \frac{1}{N} \sum_{s=1}^P g_i^{(s)} g_j^{(s)} & \text{当 } i \neq j \\ T_{ij} \approx 0 & \text{当 } i = j \end{cases}$$

$$(2) \text{ 动力方程 } h_i(t+1) = \text{sgn} \left[\sum_{j=1}^N T_{ij} h_j(t) \right]$$

如符号方程值大于 0, 则 $h_i(t+1)=1$, 表示神经发放提取信息; 符号方程小于 0, 则 $h_i(t+1)=-1$, 表示神经元不发放, 无法提取信息。在这一数学模型中, 网络各单元的关系类似于固体物理学的自旋玻璃模型 (spin glass), 是一些彼此吸引的吸引子, 遵从热力学基本规律。网络的各种构型均与其能量函数一个确定值所对应。所以, Hopfield 的能量函数和网络自由能概念, 是其理论的基石。

2. 能量函数及其在网络分析中的意义

从高能态达最小值时, 网络收敛得到稳定解, 是其神经网络理论的核心。用公式可表示为:

$$E = - \frac{1}{2} \sum_{ij} T_{ij} h_i h_j$$

式中负号表明, 网络的每次构型变换或每次迭代, 能量总是下降或保持不变 $E \leq 0$ 。网络的噪音或随机背景活动由网络间交叠而引起, 用温度 (T) 或其倒数 $\beta = 1/T$ 所表征。温度非零时, 即存在噪音条件下, 网络的自由能 F_β 是其成分间交叠 m 的函数:

$$F(\beta) = \frac{1}{2} m^2 - T \{ \ln [2 \cosh(\beta m \cdot g)] \}$$

式中 g 是存贮模式, h 是状态矢量, m 是网络的交叠, $m^S = g_i^{(S)} \cdot \bar{h}_i$ 。由此可知, 交叠是 i 位神经元存贮模式 $g_i^{(S)}$ 和其平均热力学状

态 $\bar{h}_i$ 之积。当 $T \geq 1$ 时, 上式无解; 只有 $0.461 < T < 1$ 时, 方程才有稳定解; $T < 0.461$, 网络处于亚稳定状态, 也没有真正的解。所以, 人工神经网络噪音过高和过低都不会有稳定解。

设 P 为网络存贮模式数, N 为网络的神经元数, α 为每个神经元的存贮效率, 则 $\alpha = P/N$ 。通过上述能量计算, 网络的 $P < N/2 \cdot \ln N$ 时, 网络的交叠性 m 与其存贮效率之间的关系为:

$$m = 1 - \sqrt{\frac{2\alpha}{\pi}} \exp\left(-\frac{1}{2\alpha}\right)$$

经模拟研究发现, 临界存贮率 $\alpha_c = 0.145 \pm 0.009$, 经公式计算 $\alpha_c = 0.144$, 两者十分吻合。将上面两个公式结合起来, 可以得到的结论是: 当噪音背景或温度 $T = 0.461$, 每个神经元存贮效率 $\alpha > 0.145$ 时, 网络的各成分间的交叠和存贮能力最佳。

3. 人工网络和生物脑的类比

我们引述的这些概念、公式表明, 在人工神经网络中, 通过能量计算, 可以对网络结构、功能特点进行设计和改善。那么, 在生物脑的研究中, 既然大量事实表明能量代谢也能很灵敏地反映出脑功能特点, 那么如何从人工神经网络研究中吸取计算神经科学的方法, 将生物脑的研究纳入智能化信息系统的研究轨道上来, 将是神经科学、心理科学和信息科学跨学科研究的结合点。如果把脑能量代谢和脑信息的神经科学知识与信息科学的知识加以比较, 我们可以发现两者之间存在着某些相通之处。神经信息中全或无的调频编码类似于二进制数字计算机, 神经细胞间这种调频式通讯联系类似立体声电台; 神经信息的级量反应类似普通调幅电台, 其编码方式类似模拟计算机; 脑内一些功能结构相当一些功能相似的人工神经网络, 神经元的关系相当网络单元间的交叠 m 。脑能量代谢中 33% 的热耗损相当人工网络的热噪声 T 。脑能量代谢的效率与网络效率之间, 也存在着某些相似之处; 然而, 人脑和电脑的

差异却远远大于其间的相似性。在 Hopfield 网络中,各单元之间是对称的、相似的,单元间连接在初始状态下是随机的;而人脑是亿万年生物进化和个体生活中数年、数十年发展的结晶,脑各区域的细胞并不是匀一的,其间的联结也不是随机的。脑功能的鲁棒性、容错性、并行分布性比人工神经网络及其硬件实现的神经计算机无比优越。脑能量代谢即使在安静时也不是热平衡的。在认知活动中,与其有关的脑结构总是代谢率最高,从系统外吸收的能源物质——葡萄糖最多,这些都与人工网络有着本质的区别。所以,在信息科学和神经科学跨学科研究中,既要寻求其结合点,又应认清两者研究对象的天壤之别。

四、面孔和说话人识别中的 两类信息加工过程

我们已经分别讨论了心理学、神经科学和人工神经网络这 3 个领域中,关于信息加工与能量的理论问题。那么,如何吸收这些知识进行多学科间研究呢?经过几年的探索和方法学准备,我们实验室自 1988 年以来,已选定面孔和说话人识别作为研究信息加工与能量问题的认知模式,通过整体、细胞和分子水平的神经生物学实验手段,取得脑信息和脑能量代谢的数据,再进行数学分析,力图阐明面孔和说话人识别这两种认知模式中存在的两类信息加工过程及其与脑网络和细胞能量函数的制约关系。面孔是一种特殊的图像模式,说话人噪声则是一类听觉信息模式。然而,这两种视觉、听觉信息模式,又常是不同层次语义码的信息源,包括人的姓名码、身份码、表情码、年龄和特殊经历的传记内容等。当人们看到一张照片或听到一个人的噪声时,除注意观察面孔的图形模式和倾听语声外,常联想起此人的年龄、性别、职业、个性及其特长、爱好、经历及与自己的关系等许多信息。总之,面孔和说话人识别,从

认知心理学的角度来说,包括注意、知觉、记忆等许多认知过程和多层次信息加工过程。因此,必然包含着自动加工和控制加工的两类耗能不同的信息加工过程。现将我们刚刚起步的一些研究工作与国外一些相应研究工作分别介绍如下:

1. 认知神经心理学研究(Cognitive Neuropsychology)

Rumelhart 和 McClelland(1986)在其影响深远的理论著作^[47]结束语中预言,人工神经网络理论与神经心理学研究相结合将相互受益取得新进展。脑障碍是大自然为人类带来的一种灾难,却也为脑功能的研究提供了难得的模型。令人惊异的是面孔和说话人噪声识别障碍中,竟存在某些有趣的共同特点:对熟悉人面孔和噪声的确认与对陌生人的鉴别由两类不同层次的脑网络完成。对一些陌生人面孔与噪声鉴别是否同一人的脑结构,是枕叶视皮层和颞叶听皮层;对熟悉人面孔和噪声的确认,分别由颞叶和外侧下顶叶完成。这说明,对熟悉人的视、听信息存贮从视、听皮层区发生一定距离的转移。因此,一些脑病人可能只丧失熟悉人确认的视、听信息加工能力,称之为亲人不识症,这类病人却保留对一些陌生人分辨为是否同一人的识别能力。相反的障碍也会单独发生,病人无法分辨陌生人,所有陌生人对他们来说都是一副面孔、一个声调;这类病人却能很好确认自己的亲人。对亲人不识症病人进行脑 ERPs 记录分析表明^[48],尽管病人认不出熟悉人的照片,但这些照片引出的 ERPs 中 P_3 波幅值仍高于病前不相识的陌生人照片所引出的 P_3 波,不仅 P_3 波幅值高,潜伏期也长。这些事实说明,病人对熟悉人面孔认知过程中,尽管不能在眼前照片与长时记忆中存贮的熟悉图像加以匹配,并将结果投射到意识中来,但以 P_3 波幅增高为指针的耗能的控制加工过程仍在脑内进行着,而且所需加工的时程较长。可见脑生理学上耗能的信息加工过程,与知觉心理学中的选择注意的串行加工过程发生了分离现象。这种疾病造成

的病理现象是值得深入研究的问题。除亲人不识症之外,利用全面性痴呆病人的实验研究,也证明熟悉人面孔确认是一种控制加工过程。Bäckman 和 Hellitz(1990)研究了正常老人和退行性痴呆老年病人对熟悉面孔照片确认能力的差异^[49]。他们从 40 年代社会名流,包括名运动员、演员、作家和政治家的照片 40 张,将照片投射在幕布上,每张呈现 7 秒钟,间隔 2 秒连续呈现 10 张照片后休息 15 分钟,再进行测验。请被试看 20 张照片,要求他们指出哪些是 15 分钟前曾见过的照片,并正确说出这些照片的名人身份、职业、姓名等。这一实验结果表明,痴呆老年病人不能像正常老年人那样,利用从前积累的知识对社会名流照片进行确认。这说明对熟悉人面孔的确认,包含有深层次的语义加工过程,是在知识的记忆和选择注意控制下的信息加工过程。与此不同,对陌生人照片的鉴别,主要是自动编码的加工过程。

Ozkes 和 Gilleard(1989)利用 54 张照片和 54 张画片作认知材料,对中青年组和老年组被试进行实验^[50],每张照片或画片呈现 10 秒,分别重复 1,3,5 次呈现,经过一段学习时期,一部分被试立即进入测验时相;另一部分被试经过 10 分钟后再进入测验时相。要求被试对每张照片所见过的次数和在视野中呈现的部位作出正确回答。结果发现,两个年龄组的成绩没有显著差异;对这些被试按文化程度、教育背景重新分组,也未发现对照片或画片学习和回忆过程的显著差异。这篇文章的作者认为,这一结果可以证明陌生人面孔照片和画片的认知学习,既不受年龄影响、也不受文化教育的影响,是一种相对稳定的自动编码过程。

由于认知神经心理学对面孔和说话人识别的研究能提供其他方法得不到的科学事实,1990 年 3 月起我们与海军总医院神经外科建立了协作关系,对脑手术病人进行神经心理测验、ERPs 分析和脑生化分析的综合研究。我们从国内外著名神经心理测验中,选择面孔识别等 5 项空间知觉信息加工的测验方法,也选择和制定

了说话人识别等 6 项听觉信息加工的神经心理学测验方法^[51]。我们把这 11 项测验得到的 12 项分数作为研究视觉、听觉信息加工的一个系统。经过一年多试用,将脑手术病人和胸外科手术病人术前、术后测验成绩及 2 组病人测验成绩进行统计分析。初步结果表明,每组病人术前、术后的测验成绩无显著差异,说明这套测验具有一定的信度和效度。脑手术和胸手术病人组之间,面孔识别测验、形板测验(记形、记位)和范畴测验成绩达显著差异,可见复杂视觉信息加工能力对脑损伤是十分灵敏的功能特性。对这套测验产生的差异进行主成分分析后发现:因素 1 负荷了组间方差的 40.9%,而形板测验一项在因素 1 中的因素负荷达 0.8;虽然因素 3 仅占组间差异的 12%,但面孔识别测验却占该因素负荷的 0.86;因素 4 仅占组间差异的 8.8%,但范畴测验一项的因素负荷为 0.969;虽然在多变量分析中,听觉信息加工的各项目在脑、胸手术病人间的测验成绩差异未达显著性水平,但在主成分分析中,所得因素 2 占组间差异的 19.7%,而说话人识别测验的因素负荷却达 0.94。这些初步结果表明,这套神经心理测验系统中,面孔识别等复杂空间与视觉信息加工和说话人识别的听觉信息加工,在多因素脑功能模型的建立中具有重要作用。这里我们采用面孔识别测验材料对被试来说都是陌生人面孔,可以说信息的自动加工过程占主导地位,在脑功能障碍中具有重要意义。应该坦率地承认,我们的这项研究作为时较短,数据也不够多,还需要进一步改进和完善。

2. 认知心理生理学研究(Cognitive Psychophysiology)

认知心理学创造了许多实验方法,研究正常人面孔认知的规律。左构脸与右构脸的研究发现了左侧脸负载着较多的信息;正位脸与倒立脸的研究发现了面孔认知的翻脸效应;面孔旋转的研究发现了心理旋转效应;正常脸与重组脸的研究发现了面孔认知的

拓扑编码规律;熟悉脸与陌生脸的研究发现了不同的编码过程和脑网络。这些研究表明,面孔认知过程至少包含7种编码:图形码、结构码、身份码、姓名码、表情码、面部言语码和视觉语义码;熟悉性判断、身份判断和姓名判断的反应时依次增长的事实,揭示3者是顺序进行的信息加工过程。熟悉人确认至少包含3种编码,即结构码、身份语义码和姓名码;陌生人识别则以图形码和视觉语义码的两种编码过程为主;在面孔识别中,最普遍而又共同的加工过程是并行处理,随加工深度要求不同则有顺序的串行加工过程;各种编码过程中,均可并行同时提取许多特征,实现底-顶的加工策略,也存在着顶-底的语义指导的加工策略^[52]。总之,认知心理学发现的这些规律,对于深入研究人类信息加工的自动过程和控制过程的关系,提供了良好的前提。

心理生理学以脑事件相关电位为基础,吸收了认知心理学对面孔认知研究的理论与方法,形成了认知心理生理学的新研究领域。文献中积累的事实表明,从简单素描的面孔图到真实面孔照片,随复杂性增加和要求记忆功能的参与,面孔刺激引出的 ERPs 中,较长潜伏期成分增多^[53],面孔与非面孔刺激的 ERPs 差异主要反映在潜伏期为250毫秒以前的成分^[53—55],大体在140—240毫秒之间。在熟悉人照片匹配实验中,不匹配时引起160毫秒以前的负波,以右半球为主^[56]。照片的身份、职业匹配实验中,不匹配时则引起两半球广泛性不匹配负波,潜伏期约450毫秒^[57]。

我们实验室自1988年以来,研究了正常被试面孔识别中的 ERPs,发现以双关图为认知材料时,将其认知为面孔时比认知为非面孔时, P_2 波的潜伏期加长,说明面孔认知比非面孔认知的加工过程复杂。以熟悉人和陌生人的正面脸照片为实验材料时,发现熟悉的正面脸较陌生脸引出较高幅值的 P_3 波;熟悉和陌生人左、右侧位面孔照片,对 ERPs 有相反效应,熟悉人左侧脸照片比陌生人的照片诱发高幅值 P_3 波;熟悉人右侧脸照片比陌生人诱发出低

幅值 P_2 波。这一结果提示,熟悉人面孔负载较多的信息,伴随更高的能量耗费的控制加工过程;熟悉人左侧脸负载的信息较右侧脸多,而陌生人右侧脸负载的信息多。在另一项面孔匹配的实验中发现,两张照片不匹配较匹配时,在左、右两侧颞、顶区诱发出幅值较高的 N_{400} 波,这与语义启动效应的 ERPs 有相似的现象。除正常人类被试的这些实验研究外,我们还以恒河猴为对象,研究了 6 种照片的 ERPs 诱发效应。结果表明,熟悉猴和熟悉人照片较球照片能诱发出高幅值的 P_2 波,熟悉人与熟悉猴照片比陌生人与陌生猴照片还引出更为明显的 N_{400} 波。

总结上述实验结果,我们得到这样的初步印象:随刺激面孔的复杂性和信息量增多,人类被试 ERPs 发生显著变化的 ERPs 潜伏期变长,从面孔与非面孔、熟悉人与陌生人到面孔的匹配性,发生显著差异的 ERPs 成分,依次为 P_2 、 P_3 和 N_{400} ,说明加工过程逐渐复杂。信息量多的刺激,引起幅值高的 ERPs 成分,表明更多耗能的控制加工过程参与。猴 ERPs 的变化除与人类被试的上述变化相似外,也表现出不同的规律。猴 ERPs 差异只发生在 300 毫秒以后的成分,200 毫秒以前的成分没有显著差异,可能这种识别过程对猴的难度比人类大的缘故。

说话人识别的研究,在认知心理生理学中尚属空白之地,没有国际文献可以借鉴。近几年来,我们实验室先后在人类被试和猴中,完成 4 篇研究报告^[58-61],所得到的初步印象可归纳为以下 4 点:①陌生者的噪声辨别与熟悉者的噪声识别,具有不同的神经回路:陌生者的辨别在颞区进行,熟悉者确认则由顶叶参与下完成,这一结果与神经心理学在病人中的发现^[62]是一致的。②对陌生人和陌生猴的噪声辨别 ERPs 主要是在颞区,潜伏期 100—200 毫秒成分发生变化,不伴有 P_2 波改变;说明此时主要是 ERPs 的外源性成分变化,是对声音物理特性的反应,信息的自动加工过程为主要加工方式。③陌生人噪声辨别的 ERPs 变化与噪声呈现概率有

关,小概率噪声引起的不匹配负波更明显,说明是非选择注意与自动加工过程为主;相反,熟悉人噪声确认所引起长潜伏期 P_3 波的变化与呈现概率的关系并不明显,可提示控制加工过程为主要加工方式。④猴对同类噪声识别比对人类噪声识别的神经过程较快,表明较多的并行性自动加工过程参与噪声识别。

我们所进行的认知心理生理学研究具有自己特点:对比人类被试和猴类被试在面孔和噪声识别中 ERPs 反应的异同。这一研究方案的目的,在于通过 ERPs 的桥梁将人类面孔和说话人识别过程的研究,与猴的认知模式中脑变化联结起来,以便将加工过程的脑机制研究,深入到脑细胞水平,寻求两类信息加工过程的神经细胞水平和分子水平的物质基础。

3. 认知生理心理学研究(Cognitive Physiopsychology)

生理心理学是心理学体系中历史悠久的基础理论学科,长期以来以低等脊椎动物,特别以大鼠为材料,研究简单声、光刺激引起运动反应的本能行为、习得行为或情绪性行为。1972 年 Gross 报道了关于猴下颞叶皮层具有复杂物体的认知功能,开辟了颞叶皮层认知功能的研究领域^[63]。80 年代初,英国牛津大学心理学实验室首先报道猴杏仁核中存在一些面孔认知单元^[64]。几年以后发现,对熟悉人与熟悉猴面孔识别发生特异反应的神经元主要分布在猴颞上沟上沿的颞叶皮层中。最令人惊奇的是,这些面孔认知单元大体可分为两类^[65]:一类是观察者为中心的细胞(viewer-centred cells),不论熟悉人还是陌生人,只要有面孔图像呈现,这类细胞就发生反应。根据观察者与被观察者相对位置关系,这类细胞又可分为 5 种,即正面脸、左侧面脸、右侧面脸、上仰 45°脸和下俯 45°脸。另一大类细胞称对象为中心的细胞(object-centred cells),不管是正位、侧位、还是仰、俯脸,只要是特定的熟悉人面孔出现,都发生同样的反应^[66]。前一类细胞似乎是并行的,以自动加工过程

为主,后一类细胞则是特异选择性控制加工过程的单元。最近 Roll 将面孔认知的细胞电生理研究数据,用人工神经网络的并行分布处理原理进行概括,提出对熟悉面孔存在一组为数不多的神经细胞,按疏编码规则,对一批熟人进行并行分布式群集编码。颞叶视觉信息加工后,输出到边缘系统的杏仁核,将视觉信息与味觉等多种信息聚合,并通过旁海马回、内嗅区皮层与海马的联系,构成自联想网络。这一网络的并行分布式加工,才是熟悉面孔认知的基本机制^[67]。

我们实验室已做好面孔识别的猴脑细胞生理学研究的准备工作。与其他国家现有的研究方法相比,我们已经花费多年时间摸索出一种独特的技术,即神经脉冲和激光-荧光脉冲的同时记录^[68]。石英光导纤维镀金属膜以满足细胞外电活动记录的要求,石英光导纤维将氮分子紫外激光导入脑内,脑细胞中的 NADH 分子受激发射出的蓝色荧光再由光导纤维导出。激光-荧光脉冲的幅值反映出脑内 NADH 含量的变化,从而可以计算出记录到神经脉冲的那些细胞的能量代谢水平。这样就可以直接研究神经脉冲发放率与能量代谢的关系,为神经信息加工过程与能量代谢的关系进行数学模拟。以脑 ERPs 为纽带把猴面孔和说话人识别细胞能量代谢及信息加工的数据,用以解释人、猴被试中,具有相同 ERPs 变化的机制。这样就可能对面孔和说话人模式识别中,两类信息加工过程的认识提供新的科学事实。对加深理解信息加工与能量代谢关系的命题,我们在大鼠行为药理学实验中所积累的行为和生化数据,都有一定的科学参数价值。这项跨学科多层次的研究工作正在逐渐展开,前途光明,任重而道远,希望它能为我国心理学事业增加些光彩。

主要参考文献

- [1] 北京大学心理学系编著《当代西方心理学评述》,辽宁人民出版社,

1991, 372—466。

- [2] Posner, M. I., Snyder, C. R. R., Facilitation and inhibition in the processing of signals. In P. M. A. Rabbitt, S. Domic (ed); *Attention and Performance*, Academic Press, London, 1975, **5**:669—682.
- [3] Posner, M. I., *Chronometric explorations of mind*, Erlbaum, Hillsdale, 1978.
- [4] Schneider, W., Shiffrin, R. M., Controlled and automatic human information processing: I. Detection, search and attention, *Psychol. Rev.*, 1977, **84**:1—6.
- [5] Shiffrin, R. M., Schneider, W., Controlled and automatic human information processing: I. Perceptual learning, automic attending, and a general theory, *Psychol. Rev.*, 1977, **84**:127—190.
- [6] Schneider, W., Durnais, S. T., Shiffrin, R. M., Automatic and control processing and attention. In R. Parasuraman, D. R. Davies (ed); *Varieties of Attention*, Academic Press Inc. Orlando, Florida, 1984, 1—27.
- [7] Kahneman, D., Treisman, A., Changing views of attention and automaticity, *Ibid.* 29--61.
- [8] Kerr, B., Processing demands during mental operations, *Memory and Cognition* 1973, **1**:401—412.
- [9] Hasher, L., Zacks, R. T., Automatic and effortful processes in memory, *J. of Experimental Psychology: General*, 1979, **108**:356—388.
- [10] Jchson, M. K., Hasher, L., Human learning and memory, *Ann Rev. Psychol.*, 1987, **38**:631—668.
- [11] 朱滢、王洪斌、范津、周爱宝,“加工水平、回忆与短时记忆”,《心理学报》1991(3):264—271。
- [12] 张蜀林、黄希庭,“关于内隐记忆的研究”,《心理学动态》1991(2):30—37。
- [13] Friedrich, F. J., Hlenik, A., Tzelgov, J., Automatic processes in lexical access and spreading activation, *J. of Experimental Psychology: Human Perception and Performance*, 1991, **17**:792—802.
- [14] Treisman, A., Gelade, G., A feature integration theory of attention,

Cognitive Psychology, 1980, **12**:97--136.

- [15] Treisman, A., Sato, S., Conjunction search revised, *J. of Experimental Psychology: Human Perception and Performance*, 1990, **16**:459—478.
- [16] Treisman, A., Search, Similarity, and integration of features between and within dimensions, *J. of Experimental Psychology: Human Perception and Performance*, 1991, **17**:652—667.
- [17] Duncan, J., Humphreys, G., A resemblance theory of visual search, *Psychological Review*, 1989, **96**:433—458.
- [18] 张积家、彭聃龄,“无意识知觉和知觉的信息加工理论”,《心理科学动态》1991(2):26—29。
- [19] Biederman, I., Recognition-by-components: a theory of human image understanding, *Psychological Review* 1987, **94**:115—147.
- [20] Biederman, I., High-level vision, In D. N. Osherson, S. M. Kosslyn, J. M. Hollerbach (ed): *Visual Cognition and Action*. The MIT Press, Cambridge Mass, 1990, 41—72.
- [21] McCafferty, J. D., *Human and Machine Vision*, Ellis Horwood, New York 1990, 1—8.
- [22] Knowles, W. B., Operator loading tasks, *Human Factors*, 1963, **5**:151—161.
- [23] Kahneman, D., *Attention and Effort*, Prentice-Hall, Englewood Cliffs, NJ, 1973.
- [24] Norman, D., Bobrow, D., On data limited and resource limited processing, *J. of Cognitive Psychology*, 1975, **7**:44—60.
- [25] Wickens, C. D., *Processing Resources in Attention*, Academic Press Inc, Orlando, 1984, 63—102.
- [26] Brown, I. D., Measurement of mental effort: Some theoretical and practical issues, In G. A. Harrison (ed): *Energy and Effort*, Taylor and Francis Ltd London 1982, 27--38.
- [27] Eysenck, M. W., *Attention and Arousal*, Springer—Verlag, Berlin 1982, 67--94.

- [28] Hink, R. F., Van Voorhis, S. T., Hillyard, S. A., Smith, T., The division of attention and the human auditory evoked potentials, *Neuropsychologia*, 1977, **15**:597—605.
- [29] Parasuraman, R., Auditory evoked potentials and divided attention, *Psychophysiology*, 1978, **15**:460—465.
- [30] Nāātānin, R., Implications of ERP data for psychological theories of attention, *Biological Psychology*, 1988, **26**:117—163.
- [31] Coles, M. G. H., Donchin, E., Porges, S. (ed): *Psychophysiology: Systems, Processes, and Application*, Till and Till, Middletown, NJ, 1986.
- [32] Posner, M. I., Rothbart, M. K., The concept of energy in psychological theory, In G. R. J. Hockey, A. W. K. Gaillard, M. G. H. Coles (ed): *Energetics and Human Information Processing*, Martinus Nijhoff Publishers, Dordrecht 1986, 23—40.
- [33] Sejnowski, T. J., Koch, C., Chwrecland, P. S., Computational Neuroscience, *Science*, 1988, **241**:1209—1306.
- [34] 郭爱克、潘弘、冯春华、王倩、木日先一,“计算神经科学的一个范例”,《生物物理学报》1990, **6**:135—146。
- [35] 沈政、林庶芝,《生理心理学》,华夏出版社,1989,79—93。
- [36] Smith, C. U. M., *Molecular Neurobiology*, John Willey and Sons, Chichester 1989.
- [37] 李小云、沈政,“正电子发射层描术及其在心理学研究中的应用”《心理科学通讯》1985, **3**:54—58。
- [38] Sokoloff, L., *Metabolic Probes of Central Nervous System Activity in Experimental Animals and Man*, Sinauer Associ. Inc. Sunderland, MA 1984.
- [39] ShenZheng, Lin Shu-Zhi Brain NADH and jumping behavior in the rat, *Life Science*, 1985, **37**:731—338.
- [40] Shen Zheng, Lin Shu-Zhi Effects of amphetamine and caffeine on jumping behavior and brain NADH in rats, *Acta Pharmacologia Sinica*, 1987, **8**:79—100.
- [41] ShenZheng, Wang Gang, Lin Shu-Zhi Two-way shuttlebox avoidance

- conditioning and brain NADH in rats. *Physiology and Behavior*, 1990, **48**:515-517.
- [42] Lin Shu-Zhi, Shen Zheng, Effects of amobarbital on pterins and dinucleotides contents in rat brain, *Acta Pharmacologia Sinica*, 1991, **12**: 12-16.
- [43] 林庶芝、沈政:“咖啡因对大鼠穿梭、操作行为和脑代谢的影响”《中国药理学报》1992, **13**:143-146.
- [44] 沈政、林庶芝、谢家农,“激光皮层平均诱发电位与脑辅酶”《生物化学与生物物理进展》1990, **17**:453-456.
- [45] 沈政、林庶芝,《脑模拟与神经计算机》第一章第一节,北京大学出版社,1992年。
- [46] Hopfield, J., Neural networks and physical systems with emergent collective computational ability. *Proc. Natl. Academy. Sci USA*, 1982, **79**:2554-2558.
- [47] Rumelhart, D., McClelland, J. (ed): Parallel Distributing Processing Explorations in the Microstructures of Cognition, Foundations MIT Press, Cambridge, MASS, 1986.
- [48] Renault, B. et al., Brain potentials reveal covert facial recognitions in prosopagnosia. *Neuropsychologia*, 1989, **27**:905-912.
- [49] Bäckman, L., Herlitz, A., The relationship between prior knowledge and face recognition memory in normal aging and Alzheimer's disease. *J. Gerontology*, 1990, **45**:94-100.
- [50] Ozekes, M., Gilleard, C., Remembering face and drawings: a test of Hasher and Zack's model of automatic processing in a turkish sample. *J. Gerontology*, 1989, **44**:122-123.
- [51] 沈政、林庶芝,《脑模拟和神经计算机》第五章第三节,北大出版社,1992年。
- [52] Bruce, V., Young, A., Understanding of face recognition. *British Journal of Psychology*, 1986, **77**:305-307.
- [53] Bötzel, K., Grüsser, O. J. Electric brain potentials evoked by pictures of faces and non-faces: a search for "face-specific" EEG potential. *Ex-*

perimental Brain Research, 1989, **77**:349—360.

- [54] Jeffreys, D. A., Mussel hitc, M. J., A face-responsive visual evoked potential in man, *J. Physiology*, 1987, **390**:26.
- [55] Jeffreys, D. A., A face-responsive potential recorded from the human scalp, *Experimental Brain Research*, 1989, **78**:193—202.
- [56] Barrett, S. E., Rugg, M. D., Perrett, D. J., Event-related potentials and matching of similiar and unfamiliar faces, *Neuropsychologia*, 1988, **26**:105—117.
- [57] Barrett, S. E., Rugg, M. D., Event related potentials and sementic matching of faces, *Neurepsychologia*, 1989, **27**:913—922.
- [58] 林庶芝、沈政,“噪声鉴别的脑事件相关电位研究”,《心理学报》1991, **4**:412—417。
- [59] 沈政、林庶芝、邬鹏飞,“猴对熟悉与陌生噪声识别的脑事件相关电位”,《中国神经网络第二届学术大会论文集》上册,1991,22—27。
- [60] 沈政、王秦湘、林庶芝,“猴对噪声识别及其听觉平均诱发电位”,《心理科学》1992, **2**:32—35。
- [61] 刘雅颖、沈政,“熟悉与陌生人噪声识别的脑事件相关电位研究”,《心理科学》1992(3):29—34。
- [62] Van Lanck, D. R., Cumings, J. L., Kreiman, J., Dobkin, B. H., Phonagnosia; a dissociation between familiar and unfamiliar voices, *Cortex* 1988, **24**:195—209.
- [63] 沈政,“颞叶皮层的认知功能”,《生理通讯》1991, 增 **1**:6—7。
- [64] Leonard, C. M., Rolls, E. T., Wilson, F. A. W., Baylis, G. C., Neurons in the amygdala of the monkey with responses selective for faces, *Behavioral Brain Research*, 1985, **15**:159—176.
- [65] Perrett, D. I., Mistlin, A. J., Chitty, A. J., Visual neurones responsive to faces, *TINS*, 1987, **10**:358—364.
- [66] Hasselmo, M. E., Rolls, E. T., Baylis, G. C., Nalwa, V., Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey, *Experimental Brain Research*, 1989, **75**:417—429.

- [67] Rolls, E. T., Treves, A., The relative advantages of sparse versus distributed encoding for associative neuronal networks in the brain, *Network*, 1990, 1:407—421.
- [68] 单嘉骥、沈政,“石英光纤金属镀膜微电极的问世”《中国生理心理学第六届学术会议论文集》1991,第78页。

当代情绪研究的理论来源 与方法学原则

孟 昭 兰

(北京大学心理学系)

一、理 论 来 源

尽管詹姆士·兰格学说被看作情绪心理学的第一个真正的理论,然而当我们追溯情绪研究的理论来源的时候,眼光却应该从更广的领域去认识情绪心理学的发展渊源。

纵观当代心理学发展史,情绪研究来自4条理论路线。首先是詹姆士(W. James, 1884)学说^[1],它在情绪研究未开垦的土地上,点燃了一盏灯,指明了一条路。许多学者及其研究沿着这条理论路线,作出了关于情绪研究的奠基性工作。第二为弗洛伊德(S. Freud, 1916)学说^[2],它耕耘了另一片不失润泽的土地。学者们在一片反对声中,艰辛而曲折地在它上面探寻着坦途。第三为达尔文(C. Darwin, 1872)进化论^[3]。它比上述两个学说出现得早,也曾是一盏明灯,但很久以来失去了光泽,直到近20多年来又被人点燃了,照亮着当代一些学者的研究道路。第四条理论路线就是行为主义(J. Watson, 1930)^[4]。它从20—30年代迅速兴起,在心理学研究中居统治地位。它的思想路线不能说不无缺陷,但是在方法学上,建立了客观化和操作化原则。因此,对于情绪研究来说,其成就仍是不可抹煞的。

上述这些思想路线,尽管有着截然不同、“互无”影响的意味,但是对后来情绪理论发展的影响,却是互相交织的。

1. 詹姆士-兰格学说是关于情绪的第一个理论

詹姆士心理学注重内省和人的本能,这意味着人的内在精神因素是詹姆士心理学的核心。他从内省体验的角度探索情绪的机制,提出了情绪是机体变化的反馈知觉的定义,为人们指出了几条重要的思考线索:首先,情绪被假定为可感受到的意识体验,它是一种主观的、可内省的、在个体观念中可存在的感受状态。这一观点在后人的研究中被苦苦思索,他们却难以从这种主观内省的迷雾中走出来。因此,对情绪体验的机制和研究方法虽初见端倪,需要走的路仍是漫长的。学者们始终在追求着去解决它。

经过几十年来的努力,情绪体验概念化的内涵逐渐丰富起来。人们认识到,从发生上说,情绪是有机体适应生存过程在脑中留下的痕迹。中枢神经系统,特别是新皮质发生之后,这种痕迹带有感觉(feelings)的性质。情绪体验在意识里是一个持续的过程,它对情绪自身起监视的作用,从而组织、协调和控制认知加工和机体行为。情绪体验具有特定的主观享乐色调,每种不同情绪的特定享乐色调是永不改变的,它与情绪的外显行为——表情,具有先天的一致性^[5]。在方法学方面,主观内省方法被视为必要的,但又不足以作为呈现情绪规律性的手段。近年来,由于对体验性质认识的加深,以及对情绪机制的了解,在情绪测量的多种方法已经出现的基础上,尝试采用多变量测量的综合方法已经逐渐被人所认识。

由此可见,由詹姆士所暗示的情绪体验的概念已经深化。情绪心理学家都知道,这是一个困难的、但又不能回避的问题,需要作艰苦的探索。诸如,体验作为情绪的心理实体,是为怎样的物质载体的过程所携带,情绪体验的信息操作过程的特点是什么,它与认识过程的信息操作有什么不同,它们是怎样相互影响的,等等问题。情绪心理学所给予我们的,不但是浮浅的,而且几乎是乏善可陈的。

詹姆士为人们提出来的第二个思考线索：情绪过程是刺激事件引起自主神经系统所支配的身体器官变化的知觉反馈。应当说，詹姆士的论断只是一个概念假设，然而与詹姆士同时提出“情绪是一种内脏反应”的兰格(C. Lange, 1885)^[6]，则是在他所作的一系列生理学实验的基础上提出来的。詹姆士-兰格学说对后人的研究工作所起的重要作用在于，它刺激了大量的关于情绪与有机体身体器官活动的联系的实验研究。还引起了由坎农(W. Cannon, 1927)^[7]所倡导的关于情绪脑定位的大量始终不衰的生理学研究。诸如本世纪 30 年代帕佩兹(J. Papez)^[8]提出的边缘系统概念，后来被称为“帕佩兹环路”。麦克林(P. Maclean)^[9]在 40 年代命名边缘系统为“内脏脑”和后来又有“情绪脑”之说的一系列研究，为情绪整合提供了神经学的基础性解释。在 50 年代，林斯里(D. Lindsley)^[10]关于网状结构的研究第一次提出了情绪激活的特殊神经传导通路。与此同时，欧德(J. Olds)和米纳尔(P. Milner)^[11]在下丘脑找到“愉快”和“痛苦”中枢，以及更多的关于下丘脑-脑垂体-肾上腺系统与情绪相联系的生物-化学机制。近年来，学者们发现海马更多地卷进认知、记忆工作，而杏仁核逐渐被认为是情绪发生的核心部位。还有许多关于大脑两半球情绪功能差异的研究。所有上述这些研究成果成为情绪心理学实验研究和理论发展的重要的自然科学基础。

詹姆士提出的第三个思考线索包含着起因的含义。詹姆士说，因为哭，所以愁；因为动手打，所以生气；因为发抖，所以怕。对于这些因果关系的陈述虽然在大多数情况下是被颠倒了，但是它向人们暗示了情绪体验和运动反应之间的联系。对于詹姆士这一含义的了解是经历了一个曲折过程的。人们起初反对的是，并非因为哭才忧愁，而是因为忧愁才导致哭泣。这在很大程度上是针对情绪体验与自主神经系统的关系而言的，也就是针对情绪体验与身体器官运动之间的关系而言的。在这个问题上，既从詹姆士得到了启

发,又对他提出了修正。然而体现詹姆士关于体验与运动反应之间的关系的后来研究中,更重要的是关于骨骼肌系统支配的表情研究。固然,人们首先认为这是与达尔文相联系的。其实,詹姆士理论的模式图是把自主性内脏系统与躯体骨骼肌系统的反馈作用并列的,只是由于在当时强调了兰格的关于血管系统活动对情绪的作用,人们的注意几乎全部集中在内脏反馈的研究上。近二、三十年来提出的面部反馈理论已经取代了内脏反馈理论,注重了面部肌肉运动模式与情绪体验之间的关系,但是对于这种见解的发生与发展,仍然不能不说与詹姆士有着一定的渊源。

2. 行为主义在研究方法上作出重要贡献

在进一步阐述弗洛伊德与达尔文之前,让我们先介绍行为主义路线和认知理论在情绪研究中的作用与缺陷。自本世纪20年代后的50年间,行为主义路线在心理学研究中所居的统治地位,不能不涉及对情绪的研究。行为主义以其自身的基本观点出发来解释情绪,引起了阵容庞大的研究。它提出了关于情绪的操作性定义,按照条件反射和操作条件反射原理,带出了观察情绪反应和解释情绪行为的许多方法,力图把对情绪的解释建立在客观的、可验证的实验过程之上。例如,华生认为情绪是一种反应模式,或模式化了的反应行为,它包括内脏、腺体及全身的强烈变化。由此可见华生接受了詹姆士的内脏机制假设。然而由于行为主义基本思想的限制,它摒弃了詹姆士所涉及的情绪的主观内省方面。华生为情绪确定了两个客观的指标:一为反应模式,另一为活动水平。这就确立了情绪的客观的可观察性和可测量性的实施原则。但是行为主义的基本理论模式限制了华生的思路,他把情绪归结为怒、怕、爱3种形式,而且它们是由遗传而来的。这些基本论断封闭了对人类复杂情绪的深化研究道路。

行为主义对情绪研究毕竟是有贡献的。华生所提出的情绪定

义和一般的行为主义研究方法为后人的研究展示了一条可行的途径,例如用斯金纳箱所做的动物因受电击而引起的按压杠杆率的下降实验,作为动物恐惧的指标;用剥夺水引起动物饥渴的方法作为引发情绪的刺激,以节律性咔嗒声作为电击的条件刺激所引起的反应比率的下降被称为条件性焦虑反应等,如此以水剥夺时间的长短、强化间隔率的不同,建立起恐惧或焦虑反应的动物模型^[11]。

由斯金纳创建的操作条件反射(1938,1953)应用很广泛。有的学者试图模拟人类生活情境,通过操作条件反射的方法,进行动物复杂行为的实验,以期解释某些人类情绪行为。例如阿姆赛尔(A. Amsel,1959)^[12]所提出的挫折效应,既可解释动物又可解释人类的情绪行为。他为挫折提出了一个操作性定义:“挫折是先前体验到奖赏后,又体验到无奖赏时的情况”。挫折的基本模型为:“在经过强化的行为的基础上,建立抑制性条件反射,暂时性地提高了反应率”。实验结果表明,在食物强化(奖赏)反应被消退后(惩罚),动物出现反应率暂时性提高;表明动物在产生了挫折情绪后表现了积极的行为反应,这就是挫折效应。这类实验的结果在一定条件下,完全可以解释人在受挫折后所产生的积极反应效果。例如,小学生在受到责备或惩罚后,反而能派生出积极的改正行为并能得到积极的效应。它反映了一种负性情绪可以在行为上起积极作用的可能性,并且可用重复的实验证明它。

运用操作条件反射技术进行条件性情绪反应的设计,更接近动物或人的生活情况,并可得到数量化的结果。例如,实验程序规定为,首先给动物建立稳定的积极操作程序(以食物强化的)或消极操作程序(以电击强化的),在此基础上,插入另一个积极的(食物)或消极的(电击)刺激,即可观察到动物的条件性情绪反应。在这种情况下,无论是积极操作基础或消极操作基础,附加积极刺激均产生欢快反应,附加消极刺激均产生恐惧反应。如果插入的刺激

是原操作的时间间歇或延缓,所产生的情绪反应就不同了。如积极操作背景上给以原食物强化的间歇或延缓,则出现愤怒反应;而在消极操作背景上(电击)出现的间歇或延缓,则发生宽慰反应。这一实验程序的设计符合人或动物生活中常见的情况。原初建立的积极或消极操作类似于通常情况下从事着某种具有积极或消极性质的活动,插入刺激意味着遇到了新的情况,这一新的情况可以加强或改变原来的操作,产生不同的情绪反应^[13]。

操作条件反射技术比经典条件反射更能反映动物或人的复杂的生活情境,使之更接近动物或人的生活实际。然而行为主义的基本思想和所建立的方法限制了探索心理的内在机制,排斥了心理的意识方面和情绪的体验方面,这就注定了它的研究在复杂的心理活动面前束手无策。从而被后来发展的其它学派所取代。

3. 以阿诺德为首创建的评价学说是关于情绪的第二个理论

本世纪 60 年代初阿诺德(M. Arnold, 1960)两卷集巨著《情绪与人格》^[14]的发表,标志着情绪研究进入了一个新的时期。由她所倡导的情绪的理论路线,包容着 3 大领域的工作。首先是她继承并改造了詹姆士的情绪学说,她以认知-评价论代替詹姆士的内脏反馈知觉的假设。这一点看来是阿诺德的主要功绩,在文献上多次被学者们所肯定。第二,至于她对行为主义的扬弃却很少人提及。事实上所有认知心理学家所探索的一直是刺激与反应之间的中间环节是什么。第三,她利用了当时的生理学中的全部最新成就。

随着认知心理学的崛起,许多情绪心理学者转向情绪的认知理论的研究,成为 60 年代以来情绪研究的一大流派。例如,沙赫特-辛格(S. Schachter and J. Singer, 1962)^[15]提出环境、认知和生理激活三因素论,明确指出单独环境因素或生理因素均不能引起情绪;单独环境与生理二因素结合也不能引起情绪;只有认知的参与、认知对环境因素以及生理唤醒的评价过程,才是情绪的机制。

三因素论在相当长的时期内被后来那些因袭行为主义的学者们所接受和引用。又如,以扬(P. Young, 1973)^[16]为代表的认知派学者提出,认知冲突与生理唤醒结合是情绪的原因。这一派学者实际上主张,在情绪过程中有两种冲突发生。一种是外在刺激作用与内部心理图式之间的冲突;一种是已有的生理模式与心理冲突所引起的生理变化之间的冲突。冲突论被许多认知心理学者认定为情绪的内在机制。但是,无论是三因素论或是冲突论都没有说明情绪的实体到底是什么。

与其他的认知派情绪学者不同,曼德勒(G. Mandler, 1975)^[17]努力于揭示情绪活动的深层结构。导源于格式塔心理学、信息科学和生理激活论,他提出了远比沙赫特-辛格更广的理论体系。他在信息加工的框架中加入情绪和动机概念,试图建立和解释包括情绪和动机在内的信息加工系统。

曼德勒以认知评价为思考情绪问题的核心,容纳自主系神经激活和意识体验作为情绪的整合成分,确定认知评价、自主系激活和意识体验为情绪产生的3个过程。他把认知评价解释为意义分析,认为意义分析的来源和结果取决于个体经验、期望、个性特征和信念等因素的相互作用。这些因素每次的不同结合都会引起在性质上不同的情绪。

曼德勒对自主系唤醒的内导功能,不仅把它归结为情绪的原因,而且提出了它使有机体对环境事件产生警觉。他认为,交感系统与副交感系统的交互作用除了维持有机体激活水平的平衡以外,还为保持有机体的警觉状态使之更加有效。由此可见,曼德勒没有离开詹姆士,而是修正和发展了他的学说。曼德勒发展詹姆士的另一方面是关于主观内省方面的分析。虽然曼德勒还说不清关于情绪的意识体验机制,但是他十分重视情绪结构中的意识体验成分。他提出了人的意识容量有限的看法,指出情绪的产生必然出现在心理的很小的意识窗口上,才能被主体所体验。他暗示,情绪

包含着对情境事件的意义分析和自主性唤醒这两类无意识的心理过程,它们整合于意识之中时就是情绪体验。曼德勒以焦虑为例,认为焦虑体验是由失助与干扰相结合而产生的。失助代表焦虑的认知成分;干扰导致自主性唤醒,代表焦虑的内脏活动成分。失助的认知分析与干扰的机体活动相伴随,就产生恐惧、紧张、神经质等主观体验。如果这些成分首先占据了有限的意识空间,主体就难以发现应付当前局面的有效方法,这时思维和行动上都变得狭窄而刻板,焦虑体验就难以被摆脱。如此看来,认知分析与生理唤醒在意识中的整合过程就是情绪的过程。情绪体验此时执行着控制有机体的作用,支配着有机体的行为倾向。这一观点与普里布拉姆(R. Pribram, 1970)提出的情绪的监测机制完全吻合^[18]。

以曼德勒为代表的认知结构派沿着信息加工理论建立了情绪发生的框架。这一派的主要思想强调认知分析是情绪的首要因素,它决定情绪的性质,自主性唤醒决定情绪的强度,二者在意识中整合为情绪体验。体验一旦出现就以其特有的适应性功能和动机作用对有机体施以影响。曼德勒观点的优胜之处在于提出了、并开始解释意识中的情绪体验成分,这对于克服下面我将分析的情绪副现象论十分有力。

4. 拉扎勒斯建立了完整的“纯”认知理论体系

很多认知派情绪心理学者并不是都与曼德勒一样,探索情绪的内在体验机制。与此相反,多数学者的注意力只是集中在刺激-反应的中间环节的认知加工上。最具有代表性的是以拉扎勒斯(R. Lazarus)^[19,20]为首的“纯”认知派。拉扎勒斯继承和修正了行为主义,丢弃了詹姆士学说,绝口不提情绪的主观方面。他也很少涉及生理学方面,只把生理反应作为情绪现象中伴随的必然反应。他集中精力研究人的环境关系与认知过程对情绪的决定作用。拉扎勒斯的理论体系比较完整,对情绪发生的认知过程有详尽的分

析,有很强的说服力。但人们仍不无遗憾地认为,它也有着严重的理论缺陷。

拉扎勒斯把情绪定义为“预料环境中对主体是好的或不好的信息的生理心理反应的组织或模式”。简单说,情绪是模式化了的生理心理反应。从这一点而言,他没有脱离行为主义,但是他比行为主义远为复杂。他反对把情绪看为单一变量,认为既不能把情绪归结为内驱力,也不能归结为生理激活,他把情绪看作为一种综合症,一种由4种因素构成的症候群,它们是环境因素、认知因素、生理因素和行为因素。

拉扎勒斯理论中有2个重要的概念,即评价与应付。

拉扎勒斯认为评价是一个过程,而且是分等级的。初级评价是对环境信息的性质的评价。与主体无关的刺激,在初级评价后,评价过程即终止。而对伤害性、威胁性和挑战性刺激则进行次级评价。这些紧张性刺激常常是指向未来的、形成威胁或挑战的预期,从而无论是现实的或预期的,紧张刺激常常是心理应激的来源。在次级评价中,对主体本身的评价时常成为中介因素影响第三级评价——再评价。再评价常发生在对紧张性刺激的评价过程中,在面对威胁或挑战中,主体需要估量刺激作用的后果,要寻找应付的策略和手段,要估量所采取应付手段可能发生的后果等等。情绪发生在评价之中,情绪也影响再评价,再评价又调整情绪。

拉扎勒斯重视人与环境的相互作用,他把人与环境相互作用的过程,从情绪的发生及调节来说,称为应付过程。他认为应付是一种人对环境信息的努力,这种努力当个体的需要超过他所能适应或承受的能力和范围时,为控制与调节这种需要或预期时发生。应付来自评价,是评价后采取的策略和手段,它的功能既为改变或调整行为,也为改变或调整情绪。因此,应付可以指向情境中的问题,也可以指向情绪。这样,当面对处于心理应激的患者时,心理医生视情况既可从分析与改变患者的遭遇入手,也可以从分析与改

变患者的情绪状态入手。

应付有多种多样的可能,拉扎勒斯对应付源也作了分析。他认为应付源可来自客观,也可来自个体主观方面。诸如身体健康与能量是应付的能量来源;个体的积极态度、意志或信念是应付的精神资源;个体的人际交往技能和问题解决技能是应付的重要手段。在客观上,社会支持及经济资源也是应付所不可缺少的,由此可见,个体采取应付策略是有局限的。个人方面的生理心理条件限制、环境过于苛刻、事件威胁性太强等等,都给人的应付过程增加困难。这常常是加剧应激的原因。

拉扎勒斯在分析人与环境的相互关系时,建立了交往(transaction)概念,在刺激与反应之间提出“关系”概念。人与环境以及人际之间的交往是发生在情境事件与认知评价之间、评价与应付之间、应付与再评价之间、还发生在认知、应付与情绪之间。在解剖人与环境的关系时,他指出,个体方面包括个人投入水平和信念水平;从环境方面说,包括事件的新颖性、不确定性、可预测性、紧迫性和持续时间等。

综览以上对拉扎勒斯观点的概略陈述,从他所建立的概念体系,对评价、应付过程的详尽分析、对人境关系和人际交往与评价、应付过程的复杂交融的解剖,早已超越了行为主义刺激反应的简单模式,形成了自成体系的、涉及情绪发生的认知系统。

拉扎勒斯从心理应激患者的实际出发,从诊断与治疗的过程中,总结出了极有特点的关于情绪的认知理论。他的理论的建立,深深地扎根于人的实际的关系与交往之中,所分析的是人的实实在在的认知活动和应付行为。他观察了情绪与认知的联系和互相影响,端正了传统上把情绪与理智对立起来的观念。他否认把情绪看为人的原始功能,而把它看作带有社会文化烙印的心理反应,把人的情绪、心理看作社会文化的一种现象。

然而,拉扎勒斯丢弃了詹姆士关于心理的主观内省的思想线

索。在这一点上,他俨然像一位行为主义者,否认情绪体验作为心理实体的存在。他始终把情绪看作一种反应;作为反应,情绪是认知的终端。因此他断言,情绪是认知的结果。这就十分不幸地陷入了副现象论。拉扎勒斯从不涉及情绪的个体及种族发生学,从而也不涉及具体情绪,这也不失为重大的理论缺陷。他还断言,情绪在心理学中存在的价值仅在于它对行为的分类和描述起作用,但对它的理论解释则是不必要的,从而把情绪排除在心理学的理论体系之外,这就取消了对情绪机制进一步研究的余地。

5. 弗洛伊德是情绪理论的另一个先驱

情绪研究发展的另一条理论来源是弗洛伊德学说。弗洛伊德在1916年发表了《精神分析概论》,建立了精神分析学说。弗氏理论的核心概念为本能内驱力和无意识,他对情绪的全部解释均由这两组概念出发。他认为本能内驱力的激活与释放同有机体生物学的需要与满足相联系,从而在内驱力得到释放时导致快乐情绪,而内驱力在机体内的激起与积累则产生紧张情绪。因此,他认定情绪是内驱力心理能量的释放,情绪是一个欲表露的源于本能的心理能量的释放过程。他解释内驱力的压抑是本我、自我和超我之间的冲突所引起,因此他在说明焦虑和抑郁时,认为它们都是由能量压抑所引起。

后来产生的新精神分析学派改造并发展了弗氏学说。以拉波特(D. Rapaport, 1953)^[2]为代表,他把情绪放在更大的心理环境来解释,把情绪同人格、动机联系起来。这一派学者普遍认为情绪参与在过程之中,或者认为情绪就是动机。关于情绪与人格的关系,他们的基本观点在于确定人的感情特征在整个相对稳定的心理特征群中占重要位置。这一观点为许多人格心理学家所接受。新精神分析学派把情绪的动机性和特质化诉诸为心理治疗的重要原则。他们认为儿童时期发生的重大情绪性事件或经历被压抑,其能

量会固着下来,永久性地成为人格构成的一部分,并可成为情绪或人格障碍的早期原因。

弗洛伊德理论为许多心理学者所反对,批评他所提出的基本原初概念是缺乏科学根据的,也是缺乏科学内涵的,虽然在心理治疗中它始终是有生命力的。但是,早期弗洛伊德学派早已解体。除了新精神分析主义者继承他的基本概念并加以改造以外,一些重要的思想还被后来崛起的一批情绪心理学家所吸收。这体现在汤姆金斯(S. Tomkins, 1970)和伊扎德(C. Izard, 1977)为代表的情绪理论中。

无论是汤姆金斯或伊扎德的著作,都很少谈到他们的理论与弗洛伊德学派的血脉联系。但是他们所强调的诸如“情绪在内驱力中起放大的作用”(汤姆金斯)^[22],“情绪是基本的动机系统”(伊扎德)^[23]等论断乃是他们的理论核心,这不能不追溯到弗洛伊德。

汤姆金斯断然否定了把动机归结为内驱力的看法,认为这是一个错误。其错误之处在于把内驱力本身的信号作用同它的放大器混淆了。他分析说,有机体的需要本身——内驱力提供机体需要的信号,例如饥渴的需要发出内驱力激活的信号,与此同时,伴随饥渴产生一种急切感,这是情绪。饥渴的急切感加强和促进有机体觅食的动机。因此,情绪起着放大或缩小内驱力的作用。而且,情绪作为心理适应的手段,远比生理适应更加灵活和及时。因为机体过程及内驱力活动严格按照生物节律起作用,例如人按时补充食物。然而情绪的放大或缩小生理需要的作用可以在一定程度上改变生物节律而使之适应人的社会环境。情绪的发生同身体器官的节律活动相比,可以超出时间、条件的限制。对于人的各种社会性需要来说,情绪对它们的调节的灵活性则更大。因此,感情活动由什么对象以及在什么程度上所引起的灵活性,服务于人类复杂的生活条件和人际关系条件。汤姆金斯的这一观点不能不说对弗洛伊德提出了质的修正。

伊扎德的动机论容纳了更复杂的内涵,他所建立的情绪-动机体系是从整个人格系统出发的。他所提出的人类4个动机系统——内驱力、情绪、情绪-认知相互作用和情绪-认知结构——是由整个人格的6个子系统——内稳态、内驱力、情绪、知觉、认知、动作——组合而成。按照伊扎德的看法,在人类庞大的动机系统中,情绪是核心。无论是与内驱力相联系的动机,或是同知觉、认知相联系的情绪,抑或是蕴含在人格结构中的情绪特质,都起重要的动机作用。这种观点早已远远地超越了弗洛伊德的内驱力学说,也启示了我们,为什么情绪经常处于人的心理活动的前沿,也为什么几乎在一切心理障碍中,均有情绪方面的反映;情绪既是心理障碍的原因,也是心理障碍的结果。

因此,可以把伊扎德和汤姆金斯因袭和改造弗洛伊德学说的主要方面归纳为3点:①沿用内驱力概念,认为内驱力是整合情绪活动的来源之一。②改造心理释放概念而成为动机概念,并描述了人类的动机系统。③重视情绪的内在、主观因素,探索情绪体验的内在机制。

6. 达尔文重新被人所重视

达尔文的一系列关于进化论的著作,包括《动物和人的表情》一书在一百多年前问世,对于全球范围的科学思维、学术思想和社会思想发生了重大的影响。可是在心理学界,一百年来,达尔文关于表情进化的理论几乎被遗忘殆尽。尽管在本世纪20年代以来有一些关于面部表情的研究,其结果对于推动情绪研究没有起多大的作用。只是到了60年代初期,当汤姆金斯和伊扎德为代表的一批学者在建立他们的理论时,重新审视了达尔文关于情绪的适应功能和表情的有用性的论断,提出了一条崭新的、与行为主义或认知派截然不同的理论路线。

按照达尔文的观点,有机体在进化过程中所产生的适应行为

方式,经过遗传变异和淘汰选择,在它们的结构和机能上留下痕迹。生物进化到一定阶段,由于神经系统的发生和发展,感觉功能成为比形态变化更有效的适应手段。随着动物感受性和感觉能力的提高,感情性功能——那些被称之为本能的争斗、逃避、趋近等行为成为重要的适应工具。人类的感情性反应能力则随着大脑皮质的发展而成为分化的情绪。

情绪心理学家们从达尔文以及其后一百多年来神经科学的发展所得到的启示,对近二、三十年来情绪理论的发展起着重要的作用。伊扎德指出,从进化的观点看,大脑新皮质体积的增长和功能的分化同面部骨骼肌系统的分化和情绪的分化是平行的。这意味着,①大脑两半球高级机体感觉区和运动区的形成和分化,以及通过横纹肌随意运动实现的外周运动,特别是面部运动模式的形成和分化与情绪的形成和分化均是同步的。②有机体在进化中获得的感情性反映日益同大脑皮层及皮层下部位的联系和分工相关联,并与运动器官相联系,这导致感情性适应行为随面部运动模式的形成在进化中储存于皮层下结构,成为分化了的情绪的内在体验和外显行为的机制。

在这一思想指导下,汤姆金斯和伊扎德对面部表情和情绪体验作了深入的探讨。从他们的许多论述中,我试图作一些概括。对面部表情和情绪体验归纳几点假设^[30]。

首先,人类婴儿的基本情绪是先天的;标志具体情绪的面部表情是先天预成的程序化模式。这个论断已从多方面的研究得到证明:①对先天盲婴的观察表明,他们具有与正常视觉婴儿同样的面部表情。②跨文化研究证明,前文化民族同文化民族的表情是一致的。③正常婴儿前语言发育阶段的基本情绪表情是不学而能的。随着婴儿年龄的增长,婴儿表情的显露从反射性向整合性、随意性转化,逐渐受情境和语言的制约。尤其当表情在人际间发挥着信息交流作用的时候,表情经常表现为被夸张的、被抑制的或被掩盖的。

然而表情的社会化和可被随意支配的特性,并不能抹煞它们的先天模式化的基本性质。

关于表情先天模式化的论断的争议,其实并不在于表情是否可受随意的支配,这本是不可、也无需否认的事实。提出表情先天性论断的含义在于,它是一种重要的成分,参与到情绪发生的整个机制之中。

其次,面部运动反馈是情绪的激活器。面部运动同身体其它运动器官一样,具有内导反馈的作用。模式化的面部表情的活动从皮层下某部位,如下丘脑与杏仁核,到外周面部肌肉系统有天然的神经通路。外界刺激事件引起感觉皮层和边缘系统的兴奋,激活在下丘脑或杏仁核内贮存的情绪模式,这时在面孔上显露为一种具体情绪的表情。这一表情活动本身的内导反馈再次引起皮层整合,产生情绪体验。因此,伊扎德曾强调,情绪是一个过程,它在意识里,以体验的形式持续存在,表情只是整合情绪的成分之一。

伊扎德和汤姆金斯深化了对情绪的内在体验的理解,综合多种理论来源考查体验的性质和机制。从詹姆士理论得出了情绪体验是大脑皮层对内脏活动的感知,即体验是一种特定的知觉;从弗洛伊德理论得到了情绪是无意识的,也可以是意识的,但体验必然存在于意识之中的概念。这两种理论线索在情绪心理学中已成为普遍的一般原则,然而伊扎德从更深广的方面探讨体验的性质,把体验纳入面部反馈假设之中。伊扎德对体验的理解可概括如下:

首先,从种族进化和个体发展的观点看,环境信息在脑内留下的痕迹的活动,也就是在高等动物和人类婴儿身上逐渐贮存的骨骼肌外显行为模式的活动,带有感觉的性质。它们是最初的感情性反映。这种反映为意识提供了一种粗糙的情绪觉知,因此,它既是个体最初的情绪体验,又是最初的意识。研究指出,婴儿在出生最初几周内,脑内接受的感觉材料多数来自体内器官和本体感受器的活动,它们是婴儿最初情绪体验的刺激来源,也是最初的意识感

受的刺激来源。人类婴儿的神经系统的发育水平,脑的构造与功能特性的高度分化与整合,神经冲动弥散性地扩散到脑的各级部位,是感情性反映在脑内持续存在的基础,也是意识以感受体验形式在脑内持续存在的基础。这种感情性反映随着婴儿的发育成长,从粗糙的、模糊的、时断时续的存在形式,向分化的、清晰的、持续的存在形式转化。

其次,情绪体验的意识特征具有一种监测功能。这是一个既新颖、给人以很大启发,又难以证实的论断。原来,从对边缘系统被称之为“内脏脑”以及后来的“情绪脑”的时候起,人们已经推测情绪体验与皮层下部位的整合活动的密切联系。然而,只有神经生理学家普拉布里姆(Pribram, R.)对此作了更详尽的神经学解释^[18]。普拉布里姆指出,脑核心部位的神经元聚合体是由具有短而细的神经纤维的神经元团块所组成。这些神经元具有多突触的性质,它们之间由多支纤维形成的神经网络进行平行传导,所产生的电位变化诱发感受状态并在脑内留下痕迹。这些神经元之间的信息传递对血液循环携带的化学物质特别敏感,使之成为一种特殊的感受器官。它们的敏感性使之成为持续监测它们自身活动状态的工具,并成为协调自身活动状态、维持或改变、加强或控制自身的器官。它们可以在低唤醒水平持续存在,不受自主神经系统的直接影响,而由脑干、皮下神经节和丘脑系统所直接调节。这些机构不但是心境持续存在的重要机制,而且为认知加工和行为提供动机背景。

再次,情绪是相对独立的心理过程。伊扎德把情绪定义为:“情绪是具有神经生理的、神经肌肉的和现象学的复杂过程”。从前述面部肌肉运动模式的先天预成性,以及情绪体验从进化中获得的可感受性,使人们推测情绪整合有其特定的相对独立于认知加工的过程。让我们试图作这样的理解:表情的先天预成模式的活动包含着感情性信息,对这些感情信息加工的过程是由特定的脑核心部位的神经过程所调解,它可以在没有基于学习和经验的认知的

参与下发生。研究证明,在人类婴儿还没有对外界物体在脑内进行匹配加工的能力之前,却发生了对活动着的人类面孔报以感情反应的能力。先天表情模式可被机体内外刺激所激活,它所卷进的感情性信息可以在没有认知活动的情况下进行整合,产生情绪体验。这些与表情活动相联系的先天预成感情性机构的存在,正是儿童感情与其后发展起来的认知开放系统相联系并发展起来的基础。在成人的研究中发现,对爱好的判断是直接的,它在认知判断发生之前就已经发生了^[24]。扎恩斯(R. Zajonc)据此推测:与其它表象一样,表情活动所激活的感情体验,在感情加工中可以被传递、储存和提取,因此,感情体验可以表象的形式起作用。这种感情表象由表情活动内导所激活,不存在特定的符号过程,不必以认知加工为中介。

这一论断是扎恩斯在与情绪的认知论代表拉扎勒斯的争论中作出的。如前所述,拉扎勒斯曾断言,情绪是认知的结果而陷入副现象论。新近的神经学研究利用高技术脑扫描器和计算机,正在系统地绘制情绪的脑通路图。最新研究证明,海马部位更多地卷入记忆及其它认知活动,而埋藏在颞叶深处的杏仁核与其它部位的一些联系,被认为是情绪中枢。不仅如此,拉道克斯(J. LeDoux)^[25]发现了在丘脑和杏仁核之间有一条携带情绪信息的直接通道,通过这条“捷径”到达杏仁核的信息传递比原来知道的通道——信息从丘脑上至皮层,再由皮层下至海马和杏仁核——要快2—3倍。而且,即使通过皮层的长通道被损伤,实验动物仍能对大声发出恐惧反应。拉道克斯的发现十分重要,让我以图1、图2说明这个问题。

图1是描述情绪“捷径”的模式图。整个图像是图中间小方框部分的放大的想像图式。全图按所标示的数字顺序表示情绪发生的神经信息通路。一个闪光刺激①的信息传至丘脑②和大脑皮层③,然后传送到边缘系统部位,比如杏仁核④,同时引起身体生理反应⑤。这是原来被人们所理解的情绪的神经通路。神经学家

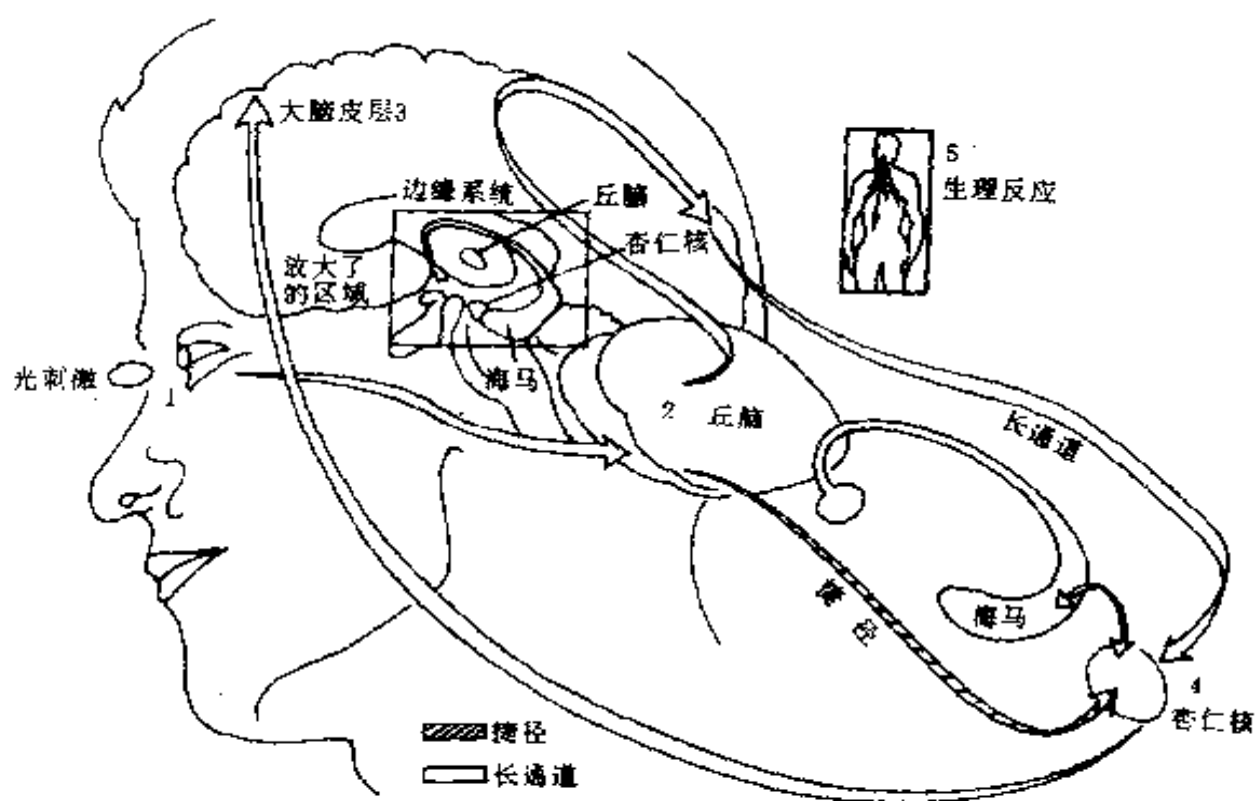


图 1 情绪脑环路“捷径”模式图

拉道克斯发现了另一条由丘脑至杏仁核的、完全不通过大脑皮层的直接通道。即在图上标示为“捷径”的通道。原始情绪，如恐惧，信息通过这条通道到达杏仁核，作出粗略的情绪判断，要比通过原来的长通道快 2—3 倍。而后，信息传至皮层，作出精确的记忆提取和思维分析。

图 2 展示情绪与认知是两个相对独立的系统，又表明二者之间的联系。图 2 左边(A)表明认知加工系统的相对独立模型及其引起情绪体验的过程；右边(B)表明情绪的相对独立的系统及其对认知加工的影响；其中(C)部分表明自主神经系统支持情绪活动的传导模型。

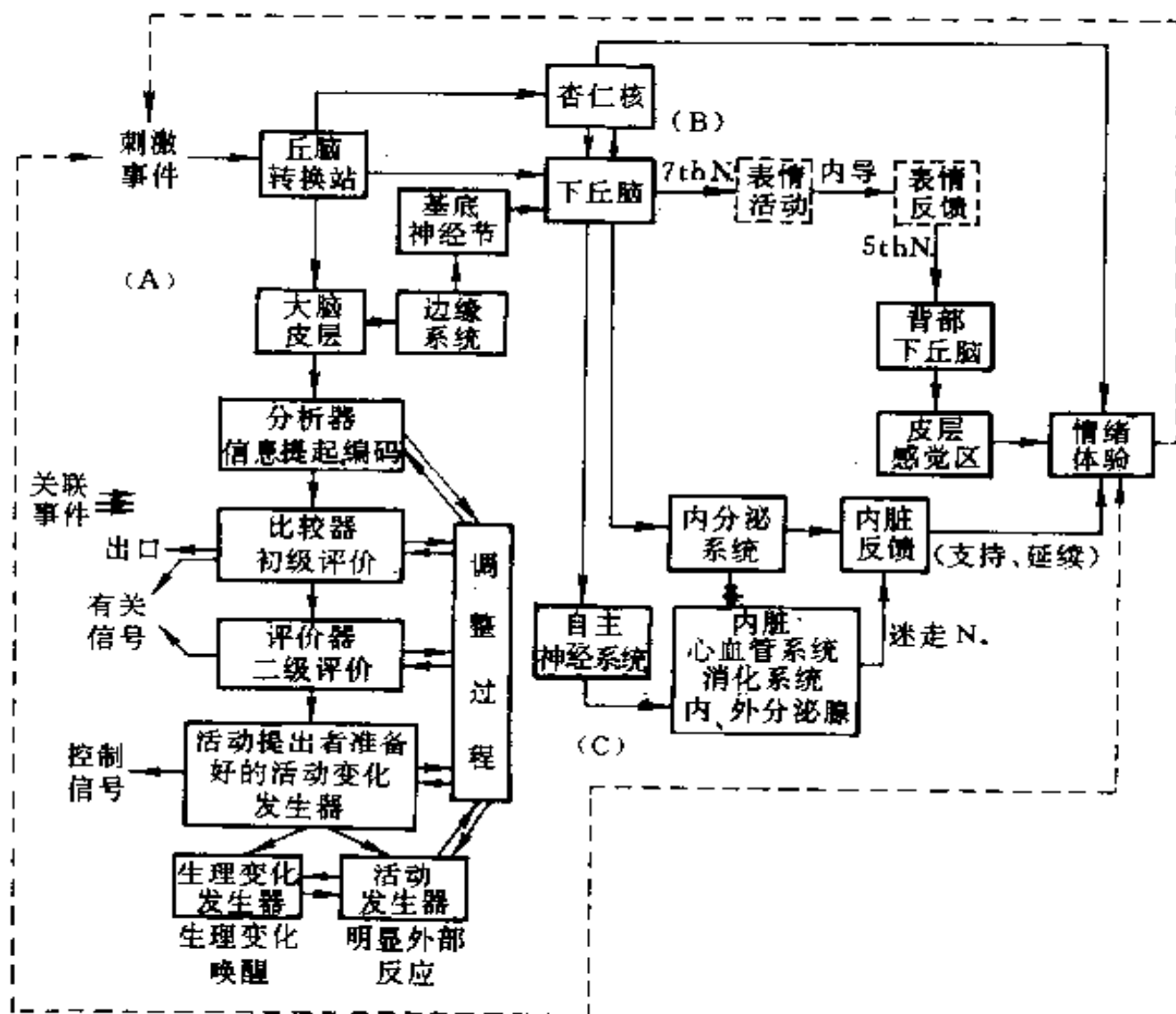


图2 情绪过程

(根据 Izard, C., Human Emotions Frijda, N., The Emotions. 孟昭兰修订)

拉道克斯的发现支持了扎恩斯、伊扎德等人的主张。它澄清了在情绪理论争论中的两大问题：①情绪过程可以在无意识觉知的情况下和无认知加工参与的情况下发生；②情绪与认知在脑中有不同的系统，尽管它们是相互联系着的。拉道克斯的发现为我们了解情绪的发生机制提供了神经解剖学证据。

二、方法学原则

情绪是一种多成分、多维量、多种类以及多水平整合的复杂心理过程。这一事实对情绪研究方法学提出了严重的挑战。然而,情绪理论的发展以及人们在理论上达到的共识日趋接近,已制订出来的研究方法的有效性日益显著。在过去,行为主义把情绪仅看为反应这个单一变量,条件反射和操作条件反射的研究方法相应而出现。内省学派把情绪归结为主观体验这个变量,根据口头报告的自陈法被广泛采用。当詹姆士把情绪确定为内脏反馈时,生理测量则成为情绪研究的唯一方法。近年来,各派学者则达到了这样的共识,即无论研究情绪的什么问题,都不能只从单一变量测量去解决。为了从方法学方面澄清这一共识,就需要把情绪具有普遍性的各个变量规定下来,并把它们转化为制订研究方法的原则。因此,在情绪研究中,既要坚持客观的原则,又要超出心理学通用的方法;要遵循它本身固有的特性,控制多变量因素,采用多项指标,作出综合分析。

近年来,一些带有普遍性的情绪特性已逐渐为多数情绪心理学家所重视和确认。因此把它们确定为情绪研究的普遍原则,将有助于研究方法的确定与课题的选择和研究质量的提高。下面仅将就已被大多数人公认的原则,结合实例试作方法学方面的阐述。

1. 情绪分化原则

承认情绪可分化为多种各不相同的具体情绪,一般来说,现在已不成问题,许多学者正在试图把情绪进行分类。但是把分化为具体的情绪确定为情绪变量而纳入研究设计的,却只是少数。这是因为从理论的建立到方法的制定还需要一个过程。如前所述,分化情绪理论把情绪的分化和进化放置在表情的分化和进化的基础上,

这就为研究具体情绪找到了一个可靠的方法。它就是面部表情测量。

近二、三十年来,对面部表情测量方法及其理论根据的专门研究已经使表情测量成为符合客观化原则的方法。它的理论根据在于^[30]:

(1) 面部表情是传递具体情绪信息的外显行为,是提供人们在感情上互相了解的鲜明标记

情绪过程既包括情绪体验,也包括情绪表现,而表情既是情绪的外部表现,也是情绪体验的发生机制;既是最敏锐的情绪发生器,也是最有效的情绪显示器。这就从机制上说明了以面部肌肉运动模式作为情绪标志的根据。

(2) 面部表情具有全人类普遍性

研究表明前文化民族的面部表情与文明社会民族的面部表情十分相似。从盖达塞克(C. Gajdusek)所拍摄的非洲新几内亚前文化民族影片中的表情,都是在文化民族中见过的,参照作出这些表情的人们的社会行为,能够了解这些表情的含义亦与文化人无异。诚如艾克曼(P. Ekman)所指出的,外国人的表情不同于外国语,不经学习就可作为有效的信息在人们之间交流^[26]。

(3) 新生婴儿具有不经学习所显露的基本情绪的面部表情模

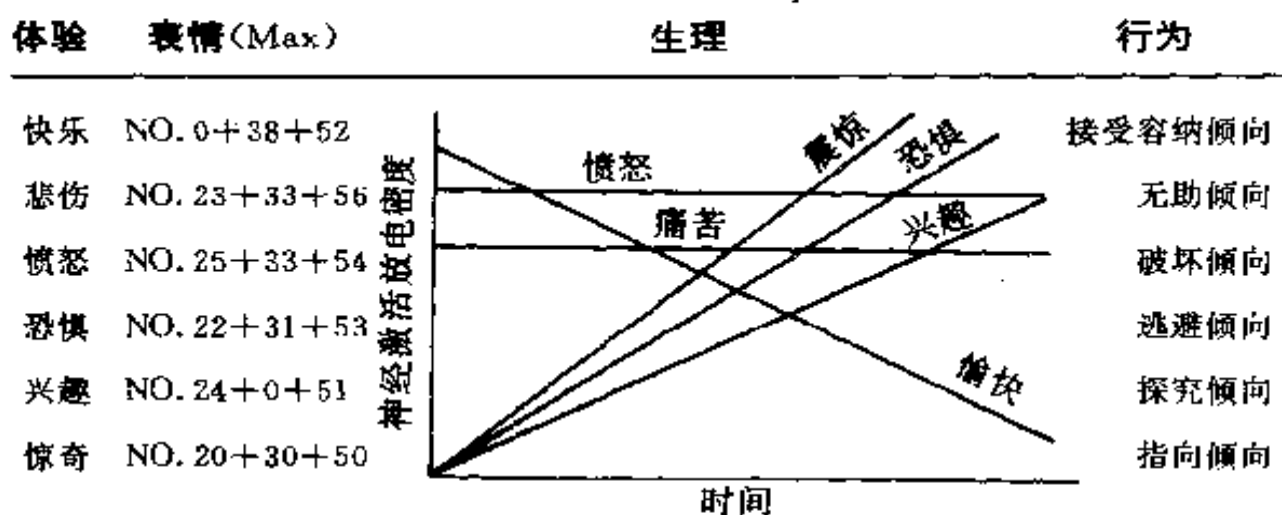
按面部不同部位的肌肉运动而模式化,面部表情就是面孔各部位神经肌肉运动的模式。毫无疑问,面部反应模式携带着心理学的意义,那就是或快乐、或悲伤等具体情绪。但是,对表情进行测量的原则在于:所要测量的是面孔各部位的肌肉运动本身,而不是面部所给予观察者的情绪信息。因此,测量的最终目的是揭示其显露的情绪信息,而测量本身必须是产生情绪信息的物质过程。为此,测量方法必须严格遵循解剖学基础。整个情绪表情产生的过程,要受严格的解剖生理过程所支配。它包括精细分化的肌肉组织系统的活动和灵敏传导的躯体神经系统活动。在测量中,使用肌肉刺激,把面孔上一块块肌肉的单一活动和互相有牵连的肌肉的组合活动通过肌电记录和照像记录保存下来,从而得到了面部各个分别独立的部位的变化模型和可观察到的、由这些肌肉活动所支配的面容。艾克曼(P. Ekman)的“面部肌肉活动编码系统”(FACS, 1976, 1978)^[27],伊扎德的“最大限度辨别面部肌肉运动编码系统”(Max, 1979)^[28]和“表情辨别整体判断系统”(Affex, 1980)^[29],是目前主要被采用的测量标准。

艾克曼的方法是把面部按活动单位进行测量的。他首先把面部分为额-眉区、眼-睑区和鼻颊-口唇区 3 个部位,各个部位均可呈现不同的活动模式。按照一块肌肉的活动或几块肌肉的组合活动分别记录为活动单位,并分别由面部 3 个部位的一个或几个不同的活动单位的活动,组成为显示某种情绪的面部表情。伊扎德的方法与艾克曼类似,只不过他制订了 Max 和 Affex 两种测量方法,互相印证来使用。对这些测量方法的制订和使用,在孟昭兰著《人类情绪》一书中有详细的记载^[30]。

艾克曼和伊扎德的面部表情测量方法已经成为情绪研究普遍使用的测量手段。本作者在对中国婴儿和成人表情模式研究与伊扎德的 Max 进行比较之后,曾经运用 Max 进行过多项关于婴儿具体情绪研究的测量^[31-33]。

2. 情绪各成分相互作用原则

情绪是由多种成分组合成的心理现象。近年来,已由一些情绪心理学家总结并公认的,情绪是由外显行为、内在体验和神经生理基础3部分所组成。其中,情绪的外显行为是指面部表情的外部表现和整体行为适应模式,主要为身体姿态表情和声调表情行为。众所周知,表情是由躯体神经系统的随意运动实现的,它们可以被人为地修饰、掩盖或夸大,并受社会、时代、文化习俗及教育教养的影响和个人境遇的制约。但是,它们在人的进化过程中所获得的原始形式仍经常可观察到。因此,情绪的成分可归纳为4个方面:面部表情和内在体验是情绪的主体成分,生理激活是情绪发生的基础和必然伴随物,模式化行为倾向是附加成分。把伊扎德的“三成分论”、汤姆金斯的“生理激活论”和普拉切克(R. Plutchik)的“系统化模式”结合起来,图示如下:



这里,标志各种情绪表情的3个相加的数目字即伊扎德的Max量表上表示面孔3部位的标号^[28]。其测量的具体方法不在此赘述。

关于主观体验测量,编制形容词检表则是测量各种具体情境下发生的情绪和测量心境的普遍使用的方法^[30]。这些检表既可用

来检测正常人,又可用来测量情绪异常的患者,而且十分便于应用。由于这方面的测量只能进行情绪的现象学描述,无法避免口头报告的主观性缺欠而不宜于单独使用。

关于生理测量,上列图示中的曲线图是汤姆金斯^[22]提出的,按情绪过程的神经激活密度(单位时间神经放电量)的变化趋势作为区分具体情绪的指标。但它只提供了一个一般的某种情绪发生过程的神经元放电变化,没有作出多项身体生理变化的区分指标。

确实,现有的生理测量方法还不能标定具体情绪的特异性。而且,在理论上,越来越明确并为人普遍接受的是,生理变化既不是具体情绪发生的直接原因,也就不像面部表情那样成为具体情绪的客观标志了。但是,关于生理活动变化与情绪现象多方面的联系,以及多年来大量研究结果的积累,许多资料提供了标定具体情绪的某一方面的生理指标。心血管变化、呼吸、汗腺、以及内分泌化学变化,都与情绪的发生和变化密切联系。例如,皮肤温度的差异可用来作为区分愤怒与恐惧的指标之一:愤怒,皮温上升;恐惧,皮温下降。血压变化可用来作为区分爆发怒与压抑怒的差异:爆发怒,舒张压上升;压抑怒,收缩压上升。尤其应当指出的是,生理指标在标定情绪的维度方面更为灵敏。例如,脉搏容积和皮肤电阻变化是情绪紧张维度的灵敏指标。有报道指出,在疼痛刺激引起的紧张情绪反应中,脉搏容积显著下降,皮肤电变化则因汗腺分泌而电阻下降,导电率上升^[34]。一般来说,心率及呼吸率是情绪激活维度和强度的灵敏指标,然而它们既在正性情绪快乐中发生变化,也在负性情绪愤怒或恐惧中发生变化。但由于正、负性情绪的紧张维度的根本差异——愉快情绪的紧张维很低,而怒、怕等负性情绪的紧张维则很高——可以在参照多种指标中加以区分。而且,呼吸率在呼气与吸气的比率上的差异也能标明情绪的差异。如在兴奋激活维增加的情况下,吸气与呼气之比大于1时(即吸气相对长于呼气时)表示情绪之紧张维上升;而吸气与呼气之比小于1时(即吸气

相对短于呼气时)表示情绪之愉快舒松维上升。

上述生理测量举例显示的复杂情况给我们一个启示,既然生理指标难以直接表示具体情绪,各生理指标对情绪维度又十分敏感,为什么不把以面部表情和主观体验所标定的具体情绪与各具体情绪的维度特征结合起来,进行具体情绪的综合测量呢?我们发现,伊扎德曾把情绪的主观体验测量和维度测量相结合,形成了相对应使用的“分化情绪量表”(DES)和“维度等级量表”(DRS)。DRS 是用来测量各具体情绪的维度模式的,它从享乐度、激动度、紧张度和确信度四维来判定具体情绪。DES 是用来直接检测具体情绪的。这样,作为研究方法,当我们在引发任何具体情绪时,以面部表情测量(Max)、主观体验报告(DES)和维度等级报告(DRS)2种测量方法综合使用,将会弥补单独使用它们之中任何一种的缺欠。

让我们以快乐与恐惧为例。快乐是最普遍的具有正性享乐色调的情绪,它在体验里呈现为满意和自信的意识状态,快乐使人处于轻松和活跃中,使人有超越的自由感和舒展的轻松感。快乐的面孔显示为额头平展,呼气时发出“呵呵”声,口部是张开的,从而面颊肌肉上推,双眼眯起,眼眉外角下垂,显示快乐体验中的舒展、轻松状态。笑是快乐的典型表现,而微笑则不发声、不张嘴,但保持了面部其它部位的肌肉运动模式,从而双嘴角上翘,面颊上推和双眼微眯的模式更明显。快乐体验使人具有亲切感,对人对事都易于接近,有容纳和接受的行为倾向。因此,快乐的享乐度和确信度较高,紧张度很低^{[28][5]}。

与此相反,恐惧是处于危险或受到威胁情境中产生的退缩性负情绪。它在体验中呈现不肯定感、不安全感和危机感。恐惧使人受到抑制、被惊呆,有逃避、退缩倾向。由于呼吸时吸气长、呼气短,故而呈现嘴唇向后平拉,口微张、眼睁大、双眉心和鼻根变窄的面容。恐惧使人处于高度紧张和激动而压抑之中,确信度和享乐度极

低,其维量特征和面孔特征与快乐相比,均截然不同^{[28][5]}(见图 3, 4)。

情绪各成分是“各司其职”,又互相联系的。表情(包括面部表情及姿态、声调表情)是通讯交流的信号,体验是承载情绪这一心理状态的主体,生理活动则是情绪的物质过程。它们是情绪过程的各个方面,是不可分割的。因此,对情绪各成分进行同步综合测量是保证测量的客观性、体现情绪过程多变量复杂性的较好方法。

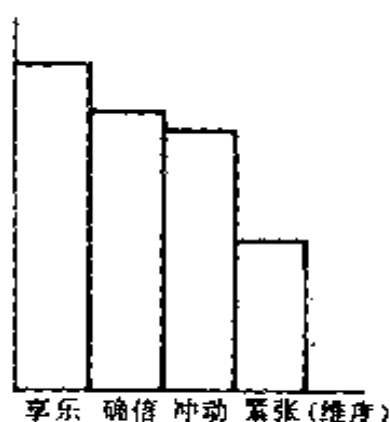


图 3 快乐的 DRS 测量

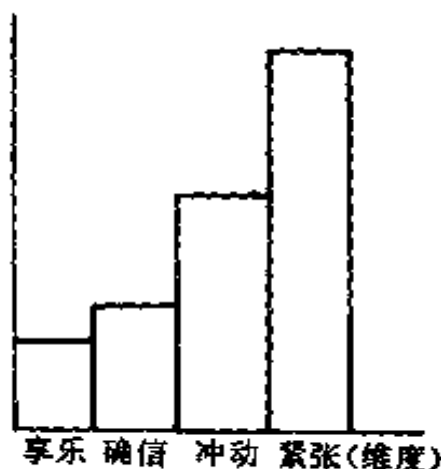


图 4 恐惧的 DRS 测量

3. 情绪的通讯和交流原则

情绪的外显成分作为一种标记,在人际间起着通讯和交流的作用。它携带并传递信息,向他人通报主体自身的体验感受。表情和体验是统一的情绪过程不可分割的两个方面。正是由于二者的一致性,表情表达着体验,表情传递的信息才可以为他人所理解。表情的符号信息不是人为地确定的,接受情绪信息的人只能通过自身已有过的体验经验才能识别他人的表情。这一特征为情绪带来一种特有的品性,那就是,表情不但传递信息,而且能在人际间蔓延和互相感染。就这一点而言,情绪的通讯职能有别于语言,它不但传递有关事实和事件的内容,而且可以在感情体验上使传递

者与被传递者双方互相沟通,引起移情或同情。这就比语言信息传递所达到的理解有更深的含义。

情绪的通讯和交流职能是情绪所具有的一种根本品质,在研究方法和实验设计中加以考虑和运用,会给工作带来很大的方便。比如,凡是试图引起被试者某种表情或体验而不使用语言词汇标示时,运用表情暗示或摹仿均能达到预期的效果。以孙绍邦的硕士论文实验中所使用的方法为例,他在进行“面部反馈假设”的检验研究中,为了检验被试者在发出某种表情时是否有相应的内在体验产生,首先需要被试者呈现出所要求的具体表情。但是这一要求不能用语言明示给被试者,因为语言提示就等于向被试者泄露了所要检验的体验是什么了。据此,作者采用向被试者呈现表情录像片的方法,通过这一媒介让被试者摹仿录像片所呈现的面部模式,然后报告他(她)所体验到的情绪。结果作者成功地引发了被试者的表情,得出了这种表情的内导反馈引起的相应体验。

这个实验所使用的方法成功地利用了情绪的通讯交流原则,在实验室工作中证实了面部表情内导反馈确实能引起相应体验的结论^[35]。这一结果,不但直接地检验了面部反馈假设,而且通过表情传递而产生相应体验的实验方法,又恰恰证实了情绪的信息传递功能。

情绪的通讯交流作用不但可用于研究设计,而且它本身就是情绪研究的重要课题。例如,婴儿在前语言阶段的学习是利用情绪信号进行的。这种利用情绪信号进行学习的能力称之为情绪的社会参照性。典型的实验是一岁婴儿面对视崖深边这一不确定情境时,就会抬头向站在视崖外对面的母亲面孔上寻找“指示”。当母亲报以肯定鼓励的微笑表情或威吓、阻止的恐惧表情时,这两种表情反应导致婴儿迥然不同的行为——跨过视崖的深端或不跨过去。王垒的实验设计利用情绪的信息通讯原则,在婴儿的视崖行为中加入一个体验变量,从而证明表情信号不单纯只是一种情境信息,

而是通过引起婴儿相应的情绪感受而调节其行为的。当那个插入的体验与外来的信息不一致时,内在体验常常成为婴儿行为的重要凭借^[36]。王垒的实验结果所得渡过视崖深端的人数差异如下所示。

在引发婴儿愉快体验背景下:

愉快情境(母亲表情)信息 8(样本数) 8(渡崖人数)

恐惧情境(母亲表情)信息 11(样本数) 2(渡崖人数)

在引发婴儿恐惧情绪背景下:

愉快情境(母亲表情)信息 13(样本数) 6(渡崖人数)

恐惧情境(母亲表情)信息 11(样本数) 1(渡崖人数)

按上述,把情绪的通讯交流原则从方法学的角度确定下来,能在很大程度上使研究的设计和方法的采用带有情绪的特色。这意味着,把内在体验与外显行为作为统一体来处理情绪的研究方法问题,揭示情绪在人际间和人群中交流和蔓延的现象,为开辟情绪社会化研究打开了一条通道。

4. 情绪与其它心理因素相互作用原则

情绪自身是一个完整的活动系统,又包容在整个的心理活动系统之中。因此,情绪以不同的方式与其它心理过程交互作用。从方法学的角度而论,在诱导情绪发生的过程中,所派生的认识活动成为一个重要变量而起作用。在实验设计中,正好可以利用认知这一中介变量作为所要测量和所要比较的指标。让我们以陈宏图的实验为例:在检测先前存在的情绪体验对后来诱导的情绪过程是否有影响的实验中,通过对小学生被试进行手工作业操作所给予肯定和否定的评价来诱导情绪体验。这时产生的体验作为“先前存在”的体验。肯定或否定的评价(以灯光显示)意味着被试操作的“成功”或“失败”。在此之后出现滑稽卡通片来引发愉快情绪反应。这时产生的情绪反应作为“后引”情绪。以“反应时”作为“先在”情

绪对“后引”情绪的影响的指标。正是在这个过程中,出现了一个认知中介变量,那就是被试的操作效果对他(她)所揭示的意义——成功或失败——的认知加工占据了儿童的思维空间。成功引起的喜悦体验,使被试直接面对滑稽卡通故事而发出笑的情绪反应;而失败引起另一些被试的沮丧体验,使他们的思维停留在对“失败”的思考而面对滑稽卡通故事却迟迟没有反应——反应时加长了。实验者从两组儿童反应时的差异来判断他们的成功和失败情绪对后来反应的影响^[37]。这里,认知因素是在情绪诱发过程中不可避免地发生的中介物,它对后来情绪反应可起促进的作用,也可以起干扰的作用。实验设计利用了这个中介物,就找到了“反应时”这个可用以进行测量的指标。

孟昭兰所作关于情绪对认知操作的影响的系列研究中,所提出的假设是不同种类情绪对认知操作的效果有不同的影响。实验过程揭示了当被试者(1岁4个月婴儿)在受到负性情绪干扰时,在问题解决的认知操作中产生了利用线索缓慢和动作刻板的加工现象,拖延了完成作业的时间,借此找到了以被试者完成手工操作的延续时间作为测量指标。这项研究不但课题本身是情绪与认知的关系问题;而且在方法学上,也利用了认知这一中介变量,以操作延续时间作为资料收集的客观指标^[31-33]。

在诱导情绪发生的过程中,所派生的认识活动这一中介物,并非每次都可以成为用以测量的可靠指标。与此相反,它有时可起到干扰情绪发生的不好作用。而诱发的情绪作为自变量是实验进行的前提条件,因为在情绪研究中,如果不能诱导出一种作为实验进行的基础的情绪,实验的下一步就无法进行。为了达到这个目的,避免或排除认知因素的影响和干扰是十分必要的。例如,当采用呈现电影片断的方法去引发被试者(成人)的情绪时,实验设计的目的本来是利用影片的内容情节作为情绪刺激,使被试者受到感染而发生某种体验。可是,被试者的思路却可以沿着影片的一些情

况和线索而发展着,诸如,影片的故事情节是否合理、影片片断的剪辑是否在故事情节上衔接、影片的艺术价值、视听效果是否够好、以及被试自身对影片的总体评价等等,这些不在设计之内出现的思维活动却在很大程度上干扰了引发某种情绪的预期。这一现象提醒我们,在情绪研究中,如不能排除和避免情绪与认知交互影响的这一客观事实,那么,在实验设计中必须加以严肃的考虑。

上述这一现象并非只发生在对成人情绪的研究上,即使在婴儿被试中也会如此。我们在一项尚未完成的实验中发现,对7—10个月的婴儿,在每间隔为4—5周的连续测验中,由于在同一实验室中进行,并从事相同的操作测量,还由于测验过程中有一个测验(陌生人接近被试3秒钟)会引起婴儿负性情绪反应,实验者发现,婴儿对实验情境的记忆使他(她)在以后的实验中拒绝进入实验室。类似这样的现象在情绪研究中经常发生,认知作为中介物对情绪引发产生干扰是情绪研究所遇到的最大困难之一。

至于情绪与人格特性之间的联系则更为密切。情绪是组成人格特性的重要方面,它既以情绪特质的形式体现,又以情绪-认知相结合所形成的特质,以及情绪-生理驱力相结合所形成的特质的形式体现。实验设计必须考虑到人格特质作为中介因素在情绪反应中发生影响,这样才不至于把情绪问题简单化。因此在情绪研究中应考虑首先把被试按气质类型(婴儿、儿童被试)和人格类型(成人被试)作出分类,排除人格特质变量这一中介物的干扰。刘小冬进行了一项关于婴儿气质类型特性之一的“情绪性”的分化实验。通过因素分析得出了4种基本情绪(愉快、兴趣、愤怒和恐惧)之间,在对某一情境的反应中不存在一致的相关关系^[38]。表明在儿童气质测量中采用分化了的具体情绪作为“情绪性”维度,比只采用笼统的“情绪性”维度将更为准确。这项研究十分深刻地表明,情绪系统在构成人格特质中的影响不是单一维度而是多维的。情绪与人格之间有着多维度的交织关系。

由此可见,情绪与其它心理因素十分密切的相互作用的特点,是情绪研究方法学的重要原则之一;也是情绪研究设计所面临的最复杂和难以处理的问题之一。作者在本文中提出这个问题,是希望能与对情绪研究有兴趣的同仁一道,共同努力,找出在实验设计中更有效的方法;去克服这一困难。

5. 情绪的自我调节原则

情绪过程的心理载体是体验,是由中枢神经系统各级水平所整合;它被表情所携带,而表情是由外周神经系统实现的随意运动所支配,再附加上内脏反应,构成情绪过程的心理-生理实体。然而由于表情为随意运动所支配,因而人的情绪的外显行为是可以进行自我控制和自我调节的。因此,在情绪研究中,情绪的可调节性和可控制性乃成为方法学中一条重要的、必须予以考虑的原则。

在过去,面部表情之所以长期被认为不能作为情绪研究的客观指标,就是由于它可以由被试主体所随意支配。尽管在大多数场合下,人们的表情是自然流露的,这种可控性有一定的限度,但鉴于社会情境和社会人际关系,特别是在实验的条件下,被试者随意修饰或掩盖真实表情,就导致得不到客观性测量数据的后果。因此,为了保证测量的客观性和科学性,表情测量经常只用于婴幼儿和精神异常患者,而很少用于正常成人。这是考虑到,婴幼儿尚未或很少受到社会关系的影响;精神病人在一定程度上失去主动的社会适应能力,他们的表情比较自然地流露于外,尽管它们是病态的。因此,对成人更多地采用量表法,用“情绪分化”和“维度水平”相结合,并以生理指标作为参照,所得到的综合指标将更为可靠。对婴幼儿,可选用面部表情测量,以行为倾向(姿态和声调表情)和生理指标为参照,所得到的综合资料也将是可靠的。

主要参考文献

- [1] James, W., Psychology. Cambridge: Harvard University Press. 1890.
- [2] Freud, S., Introduction to Psychoanalysis. 1916.
- [3] Darwin, C., The expression of emotion in man and animals. London: John Murray. 1872.
- [4] Watson, J., Behaviorism. University of Chicago Press. 1930.
- [5] Izard, C., Human Emotions. New York: Plenum Press. 1977.
- [6] Lange, C., The Emotions. Eng. transl. publ. 1922. Williams & Wilkins. Baltimore.
- [7] Cannon, W., The James-Lange theory of emotion; a critical examination and an alternative theory. *Am. J. Psycho.*, **39**. 1927.
- [8] Papez, J., A proposed mechanism of emotion. *Arch. neurol. Psychiatry*, **38**. 1937.
- [9] MacLean, P., Psychosomatic disease and the visceral brain! *Psychosomatic Medicine*. III. 1949.
- [10] Linsley, D., Emotion, in S. Stevens (Ed.) Handbook of Experimental Psychology, Wiley. New York. 1951.
- [11] Olds, J. and Milner, P., Positive reinforcement produced by electrical stimulation of the septal area and other regions of the rat brain. *J. Comp. Physiol.*, **47**. 1954.
- [12] Amsel, A., Frustrated factors in selective learning with reward and nonreward as discrimination. *J. Experi. Psycho.*, **57**. 1959.
- [13] Sidman, M., Normal sources of pathological behavior. *Science*, **132**. 1960.
- [14] Arnold, M., Emotion and Personality. Columbia University Press. 1960.
- [15] Schachter, S., and Singer, J., Cognitive, social and physiological determinants of emotional state. *Psychol. Review*, **69**. 1962.
- [16] Young, P., Feeling and Emotion. in Wolman, B. (Ed.) Handbook of

General Psychology, 1973.

- [17] Mandler, G., Mind and Emotion. New York; wiley. 1975.
- [18] Pribram, R., Feelings as Monitors. in M. Arnold (Ed.) Feelings and Emotions. New York; Academic Press. 1970.
- [19] Lazarus, R., Stress, Appraisal and coping. New York; Springer. 1984.
- [20] Lazarus R., Towards a cognitive theory of emotion. in M. Arnold (Ed.) Feelings and Emotions. New York; Academic Press. 1970.
- [21] Rapaport, D., On the psychoanalytic theory of affect. *International Journal of psychoanalysis*, **34**, 1953.
- [22] Tomkins, S., Affects as the primary motivational system. in M. Arnold (Ed.) Feelings and Emotions. New York; Academic Press. 1970.
- [23] Izard, C., Emotions as Motivations. in Nebraska Symposium on motivation. 163—200. 1978.
- [24] Zajonc, R., Feeling and Thing. *American Psychologist*, **35**, 2, 1980.
- [25] Goode, E., Where emotions come from. *U. S. News and World Report*, June **24**, 1991.
- [26] Ekman, P., Measuring facial movement. *Environmental Psychology and Nonverbal Behavior*, **1**(1) Fall, 1976.
- [27] Ekman, P., The facial action coding system (FACS). Palo Alto, California Consulting Psychologists Press. 1978.
- [28] Izard, C., The maximally discrimination facial movement Coding system (Max). Newark; Instructional Resources Center, University of Delaware. 1979.
- [29] Izard, C., A system for identifying affect expression by holistic judgments (Affex). Newark; Instructional Resources Center, University of Delaware. 1980.
- [30] 孟昭兰,《人类情绪》,上海人民出版社,1989年。
- [31] 孟昭兰等,“幼儿不同情绪状态对其智力操作的影响”,《心理学报》,1984年3期。

- [32] 孟昭兰,“情绪的组织功能”,《心理学报》,1988 年 2 期。
- [33] 孟昭兰等,“幼儿怒情绪对认知操作的影响”,刘范编《心理发展的近期研究》,北京师院出版社,1990 年。
- [34] 邵郊等,“以多种生理指标观测痛觉反应的初步报告”,《北京医学院针刺麻醉研究论文集》,1974 年。
- [35] 孙绍邦等,“面部反馈假设的检验研究”,《心理学报》,印刷中,1992 年。
- [36] 王垒,“婴儿情绪背景下母亲表情对婴儿行为的影响”,硕士毕业论文,未发表(1987 年)。
- [37] 陈宏图等,“关于先发情绪体验对后发情绪过程的影响的一项实验”,《心理学报》,1990 年 1 期。
- [38] 刘小冬等,“关于婴儿情绪性的分化实验研究——气质的情绪性特征在不同情绪中的表现是一致的吗?”《心理学报》,1990 年 4 期。

心理学研究方法

王 重 鸣

(杭州大学心理学系)

心理学的理论发展与实际应用,离不开行之有效的研究方法。当代心理学的每一项重要进展,都与研究思路的突破和研究方法的更新密切有关。心理学研究方法内容十分丰富,从研究的构思、设计和取样,到数据的采集、测量、统计分析和研究结果的报告。心理学研究方法领域的最新进展,集中反映在研究思路的综合化、测量方法与分析手段的精细化、理论建模的系统化等方面。本文将把心理学研究的理论构思、设计、测量与分析诸方面融为一体,介绍和讨论心理学研究方法的新进展和新成果。

一、心理学研究的理论构思与设计

在心理学研究中,理论构思是至关重要的环节。通过理论构思,使整个研究具有较高层次的理论出发点和清晰的分析路线。在此基础上,就能够作出比较周密的研究设计方案。

1. 心理学研究的理论框架与制定计划

(1) 心理学研究的层次与理论指导

与其他许多学科的研究一样,心理学研究具有不同的研究层次与分析水平。首先是描述水平,侧重于对研究对象的现状作出一般的描述和说明。例如,学生学习动机的强弱、职工工作满意感的高低、弱智儿童心理缺陷的轻重、或者人类记忆容量的大小,等等。

在这一层次上的研究,只是对当前的心理特征或状态作出评估与测定,并不关心心理特征或活动的前因后果。在分析的手段方面,这类研究主要依靠描述统计(descriptive statistics),因此,属于较低研究层次。许多探索性的心理学研究处于描述水平。其次是解释水平,重视寻找影响心理发展特点或心理活动的主要因素。比如,学习动机为什么有高低?工作满意感与哪些因素有关?弱智儿童的生理与家庭条件有何影响?儿童认知发展如何形成若干阶段?等等。围绕这类问题而开展的研究就处于较高一级层次,需要对多种因素的交互影响作比较透彻的分析,常常需要采用某些多元分析的方法,例如复相关、聚类分析或因素分析等。再高一层次则是预测水平的心理学研究,强调对所研究对象将来的发展进行预测。比如,根据学习动机与能力的高低预测学习的成绩,从工作条件与满意感预测进一步的工作实绩,以弱智儿童当前诊断得到的心理与生理指标预计行为矫正方案的效果,等等。进行预测水平的研究,关键是要构建好相应的因果关系模型,注重纵向的动态分析,在具体方法上应运用多元回归分析、通径分析(path analysis)或者结构方程建模技术(LISREL)等高级统计手段。

无论采用何种研究层次,都需要有更高层次的理论指导。辩证唯物论为心理学研究提供了一系列科学的指导原则。从方法论的角度来看,有两方面的原则是十分重要的,即关于心理“过程”的思想和有关因果关系验证的论述。恩格斯在《路德维希·费尔巴哈和德国古典哲学的终结》一书中精辟地指出了唯物辩证法的基本思想,“即认为世界不是一成不变的事物的集合体,而是过程的集合体”,强调了进行动态过程分析,这对于心理过程的研究具有深刻的含义。在《自然辩证法》一书中,恩格斯论述因果关系问题时指出,单凭观察所得的经验是决不能充分证明必然性的,因果关系的证明是在人类活动中、在实践中、在劳动中。他提出了以人类的活动与实验操纵作为心理活动因果关系验证的根本途径。

(2) 心理学研究的计划制定

心理学研究计划的制定,是一个理论构思和方法设计的过程,也是心理学研究取得成果的必要条件。完整的研究计划应包括 3 个部分:

① 研究目的与理论框架:研究计划中应首先确定研究的具体目的和问题,并提出研究的理论框架。研究的目的应明确具体,理论框架应紧密联系本研究的实际。例如,进行一项有关奖励与学习成绩关系的研究,需要首先确定理论框架。在奖励与成绩之间,显然存在学习动机这一中介变量。这样,研究实际上涉及“奖励—动机—成绩”这一关系,因此,可以选用有关的动机理论或奖励理论作为研究的理论出发点。动机理论又可分成侧重于需要与激励因素分类的“内容理论”和着力于动机发展变化的“过程理论”,在确定理论框架时需要明确本研究的理论层次,即是描述“激励因素”(内容理论定向)、还是从奖励特征分析对动机及作业成绩的影响(过程理论定向)。在此基础上,需要提出较具体的假设,指出有关理论对于本研究的意义。当然也可以根据前人的理论与研究结果,提出自己新的理论框架。

② 以往研究的考察:在研究计划中,必须全面回顾本课题有关的以往研究,以便提出新的具体课题或问题。这样,研究的起点就比较高。以往研究的考察并非罗列或堆积以往文献,而是要对以往研究作出理论上的总结和分析,提出新的问题。不少研究正是由于忽视了这一步而成为低层次研究或重复研究。

③ 研究的设计与具体方法:在确定了研究的理论框架和着重要解决的问题以后,下一步就是进行研究的具体设计和提出具体方法。研究设计的一个关键是构建主要变量的关系图,使研究图式具有明确的结构。然后,需要确定研究的总体、取样方法、测量指标与工具、统计方法等。心理学研究主要有两大范式:一种范式是选用较大样本作一、二次测量,称为大样本范式,适用于进行较大范

围的调查或测量;另一种范式是对较小样本作多次测定,叫作小样本范式,较适用于对少数被试作深入研究与分析。无论是大样本少测量还是小样本多测量,测量数据的样本都应达到一定数量,才能保证样本的代表性。为了进行有效的多元统计分析,样本一般应达到150的规模,以便获得比较稳定的统计分析结果。这些都是在制定研究计划时应加以重视的。

2. 心理学研究的效度分析

近年来心理学研究方法的重要发展之一,是日益重视研究的效度,并把研究的效度分析作为研究设计与评价的一项关键内容。心理学研究中应特别重视4个方面的效度:内部效度、统计结论效度、构思效度和外部效度,并强调从以下方面加强研究的效度。

(1) 明确研究中的变量关系

研究的内部效度是指研究中自变量与因变量之间关系的明确程度。一项心理学研究要达到足够的内部效度,必须明确规定其自变量和因变量,以便决定相应的因果关系。方法论的研究表明,自、因变量含糊或者根本没有考察的研究,多半是探索性的低层次研究,谈不上具有内部效度。研究的内部效度取决于对研究课题的周密构思(包括自、因变量的选择与关系图的构建)和实验控制的程度。研究中需要控制的因素比较多。例如,被试心理特征的差异、取样的代表性、测量的准确程度、实验信息的扩散等。

(2) 加强多重检验与代表性

研究的外部效度是指研究结果对于其他总体、时间和背景条件的普遍意义。通常,现场研究由于是在现实背景下对真实被试进行的,因而具有较高外部效度,而实验室实验则因为其实验情景与任务等方面带有很大的人为性,其外部效度就可能比较低。因此,要提高外部效度,需要加强研究取样、情景和测量指标等的代表性,并通过多重测量与检验,使研究的结果更具有普遍意义。

(3) 注重数据质量与统计分析假设

研究的统计结论效度是有关研究统计分析程序有效性的检验,这与研究数据的质量和统计假设密切相关。心理学实验数据常常难以达到许多统计分析方法所要求的正态分布,数据测量包含的误差也会在很大程度上影响统计分析结果及其解释。这方面有一些常见错误,例如,对小样本数据进行多元统计分析,把被试内设计(within-subjects design)的重复测量数据当作被试间设计(between-subjects design)的数据加以统计分析,对统计分析中显示出包含高残余变异(residuals)的变量作过分的解释,等等。这些都会直接影响研究的统计结论效度。

(4) 强调构思层次与网络化

研究的构思效度涉及到研究设计和测量的理论构思及其操作化的有效性。这也是评价研究质量的一项重要标准。例如,为了研究某项奖励制度的效果,要是只简单地比较该奖励制度实行前后的成绩或者满意感,在研究的构思上就很不够要求,在理论上没有明确、系统的构思。为了在研究中加强构思效度,应强调研究构思的层次与概念间网络关系,并确定相应的操作指标。就奖励制度来说,需要分析奖励制度的心理学特征及其影响因素,进而确定测量的指标。研究方法的一个新的趋势是日益重视用多种方法和多维操作对理论构思进行研究,改进研究的构思效度。

3. 心理学研究的实验设计

(1) 不等同对照组设计

大量的心理学实验都采用了实验组与对照组相比较的方式考察实验效应。可是,在通常条件下,实验所要求的被试随机化难以真正实现,因此,对照组在被试条件不等同的情况下与实施实验处理实验组进行比较,属于一种准实验设计(quasi-experimental design),又叫作不等同对照组设计。由于被试条件不等同,对于这类

实验设计结果的统计处理,既不能简单地采用 t 检验法对变量的事前、事后测量值一一作出差异检验,又不应使用一般的方差分析法作出分析。比较有效的统计方法是对实验数据进行协方差分析(ANCOVA),本文第三部分对这种方法进行讨论。

(2) 纵向研究设计

心理学研究中,日益重视采取纵向分析的研究图式,即在研究设计中强调时间维度、从发展的观点研究心理过程与心理特征。纵向研究往往能提供比较充分的实际依据,支持所揭示的因果关系,因而,纵向研究设计在儿童心理学(如独生子女、超常儿童和弱智儿童等)、教育心理学、管理心理学、社会心理学、临床心理学等许多领域,都是极为有力的研究设计图式。纵向研究设计的关键是在一定时间间隔上,对相同被试进行多次测量,考察对象心理发展的特点或者实验处理前后一段时间内的心理变化。常用的纵向研究是间歇时间序列设计,这种设计要求包含几十次观察与测定值,以便从整体发展趋势上分析心理特点与过程。纵向研究中还应特别避免周期性因素干扰(季节、态度、情绪等因素的变化),并使被试保持稳定、防止更换或散失的情况。此外,多次测量的程序和指标应尽可能保持一致和可靠。

纵向研究图式的一种变式称为交叉滞后组相关设计。这种设计采用平行测量指标,在不同时间点上对自变量与因变量作出相应的测量,并运用相关分析方法(见图 1),求出交叉相关系数(r_1, r_2)、同步相关系数(r_3, r_4)以及稳定性系数(r_5, r_6)。研究表明,当稳定性系数和同步系数比较一致的情况下,可以通过交叉相关系数之间的差异,推论 A, B 变量之间的因果关系方向。这时,如果 $r_1 > r_2$,就认为变量 A 引起 B ;反之, B 引起 A 。由于相关系数本身并不能证明因果关系,因而需要把这种设计与实验操作结合在一起,加强研究的功效。

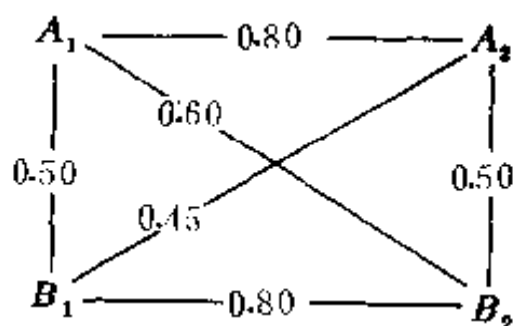


图 1 交叉滞后组相关设计

(3) 因子式实验设计

心理学家越来越认识到,心理活动与特征是由个体的、社会的、内部的、外部的多种复杂因素共同作用和交互影响的结果,因此,心理学研究就需要采用适应于这种复杂交互关系的研究设计图式。因子式实验设计(factorial experimental design)就是这种研究设计之一,它具有较高的生态效度(ecological validity)。虽然因子式实验设计已是人们十分熟悉的设计图式,但以往研究中常常把因子看成相互独立的因素,只关心其主效应(main effect)。近年来,研究方法中日益重视运用因子式实验设计来分析多种因素之间的交互作用,并认为交互作用在一定意义上比主效应更有理论价值。这是因为在现实生活中,各种因素总是相互影响与作用的。在因子式实验设计时,应重视 4 个方面的问题。①因子的构思和各因子水平的确定应注重层次性,实际上,因子构建的过程可以使研究思路更为清晰、结构更分明。②各区组的被试必须真正是从总体随机抽取而又在各组随机分配的,随机化是因子式实验设计的重要条件。③在统计分析时,应区分“被试内设计”与“被试间设计”,大部分因子式实验设计为了提高实验效率,多采用被试内设计,让每一被试在不同实验条件下与自己比较,因而接受了多重实验处理与测量(repeated measures),对于这类设计的实验结果,需要运

用不同于通常方差分析的方法,在许多统计软件包中,都有适合于多重测量实验数据的程序。④因子式实验设计中应特别重视因子间的交互作用(interaction),在出现交互作用时,需要进行单效应分析(simple effect analysis),从而对交互作用的性质与意义作出分析与解释,更系统地了解因子间的复杂关系。

二、心理学研究的测量

1. 心理测量理论

心理测量是心理学实验科学的重要组成部分,这不仅需要以系统的心理测量理论作为基础,而且要求设计可靠、有效的测量量表。心理测量理论主要有两大类:经典测验理论和现代测验理论。

(1) 经典测验理论

经典测验理论起源于早期有关智力和智力测验的研究,以Spearman 的分数模型为基础:

$$X=T+e$$

X 表示测验原分数, T 表示心理特征的真分数, e 表示由随机因素造成的误差分数。认为真分数呈正态分布、且与误差可加及相互独立的。目前流行的一些心理测验及其信度、效度的定义与估算方法都是以这个分数模型为依据的。经典心理测验理论对心理学研究起了重要的指导作用,但由于它是从测验水平上对测验结果进行分析,而忽视了相似总分背后所蕴含的不同的反应模式,加上其测验参数在很大程度上依赖于所测样本的特点,因而存在一定的局限。

(2) 现代测验理论

现代测验理论是近一、二十年发展起来的,主要有潜特征理论和项目反应理论。这种理论的特点是注重从项目水平上对心理特

征进行测量与分析,并以项目特征曲线和项目信息函数表示项目的测量参数与特征(难度、辨别力、信度和测量标准误等)。项目特征曲线是由从具有特定潜特征的总体随机选取的受测者作出特定反应的条件概率。现代测验理论为心理测量标准化和电脑化提供了有力的理论依据,也为设计适应性心理测验(adaptive testing)程序和心理测验项目库提供了具体的手段。适应性心理测验程序能够运用电脑,按被试的能力水平,实时调用适当层次的心理测验项目,进行有效的实时心理测量、诊断和训练。目前,现代测验理论及方法已广泛用于心理测量的各个方面,例如测验项目分析、测评量表设计和适应性测验的设计与实施。

根据现代测验理论,对心理测量的参数作了如下规定:①以项目特征曲线的斜率表示项目辨别参数(a_i),斜率越大,辨别力越高,这样,能力量表上较小的差异就会在项目通过率(P)上表现出较大的变化;②以某一项目正确反应的概率 $P(\theta) = 0.50$ 位置所对应的能力量表值表示项目的难度参数(b_i);③以能力量表值为零值时(毫无能力)的“正确反应”概率值作为项目的猜测参数(c_i),这一参数对于多重选择项目题尤其重要。图 2 表示出典型的项目

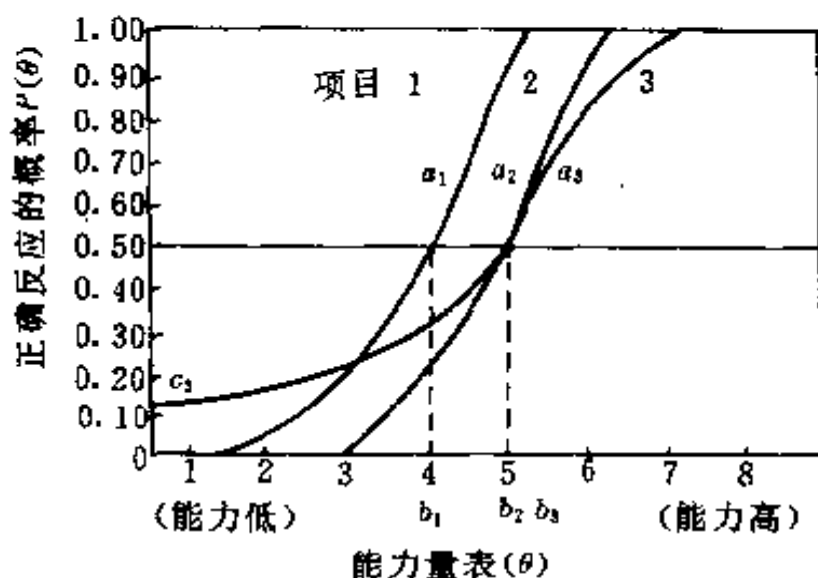


图 2 项目特征曲线

特征曲线及其测量参数。由于这些参数是以正确反应的概率为基础的,因而比较稳定。

(3) 心理测量的信度与效度

目前比较流行的信度与效度估算方法是以经典心理测验理论为依据的。由于 $X=T+e$, 而且 T, e 是相互独立的, 因此, 心理测量分数的方差(或变异度)为真分数与差误分数方差之和, 即

$$S_X^2 = S_T^2 + S_e^2$$

测量的信度则是测量分数反映出被测特征真实变异的程度:

$$r_{XX} = 1 - (S_e^2 / S_X^2)$$

所以, 差误变异(S_e^2)越小, 测量信度越高。具体估算信度的方法比较多, 例如, 重测信度(test-retest reliability)、折半信度(split-half reliability)和内部一致信度(internal consistent reliability)等。重测信度通过对测验作间隔 1 个月左右的事前、事后两次测量结果的相关系数作为指标, 简便易行。可是近年来的研究表明, 重测信度可能导致严重的误估, 特别是难以监测测量中常常出现的反应定势与偏向以及项目结构低劣等问题。因而, 在心理学研究中更主张采用内部一致信度——克劳伯克(Cronbach, L.)的 α 系数, 其公式为:

$$r_\alpha = \frac{K}{K-1} \left[1 - \frac{\sum S_i^2}{S^2} \right]$$

式中, K 为测验项目数, S_i^2 为第 i 个项目的方差, S^2 为总得分的方差。当测验项目为是否题或对错题等两分项目时, 用各项目的通过比例(P, q)表示方差, 以 Kuder-Richardson 公式(KR20)估算测量信度:

$$r_{KR20} = \frac{K}{K-1} \left[1 - \frac{\sum Pq}{S^2} \right]$$

式中, P 为某项目的通过率, q 为没有通过该项目的比例($q=1-P$)

P), S^2 同上式。内部一致系数计算方便,目前许多统计软件包都有相应的统计程序。运用内部一致系数作为信度指标的关键,是在测量项目设计时应使项目具有明确的内部结构,在对多种概念或变量进行测量时,可设计多个子量表。

测量效度的估算比较复杂。如果说信度是测验自身两次测量或多组子量表之间的关系,那么效度则是测验与某种外部标准(效标)之间的关系。因此,测量的效度不仅与测验本身的特点有关,而且与效标的质量密切有关。这一点在以往研究中往往被忽视。测量的效度也有几种类型,例如内容效度、预测效度、同时效度和构思效度等,各适合于不同的研究目的。在心理学研究中,最经常采用的也许要算预测效度(predictive validity)和构思效度(construct validity)了。预测效度是指测量结果与将来行为表现或业绩之间关系的密切程度,比如智力得分与以后学习成绩之间的相关程度、培训成绩与以后工作成绩之间的相关等等。这里,智力得分和培训成绩叫作预测指标,学习成绩与工作成绩则称为效标。由于预测效度涉及对实际学习、工作与生活中行为表现的预测,因而被认为是最严格的效度类型。

构思效度与测验的理论构思有关,对于测验结果的理论分析来说,这是最重要的效度指标。通常采用因素分析和聚类分析等方法对测量效度作出评定。假如属于同一构思的测验项目在相同基本因素上有高荷重或聚类在一起,就认为测量的构思效度比较高。近年来,心理学家日益重视对测量结果进行多重效度验证(multiple validation),认为单一的方法(如因素分析)不足以充分验证测量的效度,因而主张用聚合效度(convergent validity)和辨别效度(discriminant validity)作为指标,以多种测量方法对相同心理特征评估的相似结果和同一方法对不同心理特征测量结果的差异程度,对构思效度进行估算与评定。Campbell 和 Fiske 为此提出“多特征-多方法矩阵法”(multi-trait multi-method matrix),考察“异

特征-同方法”、“异特征-异方法”测量结果的相关系数是否有相似模式以及“异方法-同特征”相关系数(效度系数)的显著水平,以此评估测量的构思效度,见表1。

这种多特征-多方法矩阵法已成为评估构思效度的重要新途径。

表1 多特征-多方法矩阵

	特征	方法1			方法2			方法3		
		A ₁	B ₁	C ₁	A ₂	B ₂	C ₂	A ₃	B ₃	C ₃
方法1	A ₁	(.89)								
	B ₁	.51	(.89)							
	C ₁	.38	.37	(.76)						
方法2	A ₂	(.57)	.22	.09	(.93)					
	B ₂	.22	(.57)	.10	.38	(.94)				
	C ₂	.11	.11	(.46)	.59	.58	(.84)			
方法3	A ₃	(.50)	.22	.11	(.67)	.42	.33	(.94)		
	B ₃	.25	(.58)	.12	.43	(.66)	.34	.67	(.92)	
	C ₃	.11	.11	(.45)	.34	.32	(.58)	.58	.60	(.85)

注:():信度系数; (○):效度系数;

△ 不同特征-同一方法的相关系数(异特征-同方法);

△ 不同特征-不同方法的相关系数(异特征-异方法)。

2. 心理测验效度验证的若干问题

近年来,我国心理测验与测量的研究在测验修订、移植、常模制定以及自编设计等方面取得了引人注目的新进展。但是,在研究方法特别是测验效度验证(validation)上也存在一些值得重视的问题。在许多心理测验的修订、移植或设计编制过程中,比较重视采集大样本数据来制定“常模”,却忽视了严格的效度分析与验证,例如用很小的样本来估算测验的信度与效度。许多研究表明,效度验证工作是一项极为重要而复杂的研究任务,其意义也许远远超过了制定常模,是在心理测验研究中尤其需要重视的环节。效度分

析中的另一常见的问题,是用新编测验与已有较完善测验得分之间的相关来证明新测验的效度,即以后者作为效标。比如,用某一新测验与韦克斯勒成人智力量表得分之间的高相关作为新测验有效性的依据。心理测验专家 Anastasi(1990)认为,当新测验只是已有测验的简化版本或者两者只是个人与团体测验版本不同时,可以用后者作为效标,以两者之间的相关来检验新测验。否则,就不宜用已有的较完善的测验作为效标。正如前面所指出的,在测验效度验证时,应慎重地选择效标,常用的指标有学习成绩、工作绩效、行为表现等实际效标,也可采用某种理论构思作为效标。

在估算出预测效度或者其它效度系数以后,需要对其含义作出解释。效度系数多大才算达到可接受的水平呢?首先要求效度系数具有统计意义,其次才能对其作出进一步的解释。效度系数的解释还需要运用“估算的标准误”(standard error of estimate, SE_{est})作为依据。 SE_{est} 与测量的标准误相似,它表示出效标估算值所允许的差误范围,即:

$$SE_{est} = SD_y \sqrt{1 - r_{xy}^2}$$

式中, r_{xy}^2 是效度系数的平方值, SD_y 是效标分数的标准差。从公式可知,当 $r_{xy} = 1.00$ 时,“估算的标准误”为零,而当效度系数很小时,则 SE_{est} 接近于效标的标准差 SD_y , 这时预测力就很低。因此, $\sqrt{1 - r_{xy}^2}$ 表示出相对于随机猜测结果所产生的差误大小。例如,效度为 0.80 时, $\sqrt{1 - r_{xy}^2} = 0.60$, 差误量就是随机猜测所得结果的 60%, 换言之,运用该测验进行预测的话,可以使差误比随时猜测结果少 40%。可见,即使是 0.80 这么高的效度系数(通常效度系数数值在 0.30—0.60 之间),仍然有较大幅度的预测差误。一般情况下,测验多用于筛选人员或普通诊断,这时,效度系数达到 0.30—0.40 就认为可以接受了,因而并不对 SE_{est} 作严格的检验。需要强调的是,即使证明具有高效度的测验,如果在使用时操作程

序不当,或者条件控制得不严格,也无法取得可靠、有效的心理测量数据。因此,许多心理学家认为,测量参数的验证是一个从测验设计、量表制作、取样、数据采集、统计分析等诸方面相互影响的研究工作程序,其每一个环节上的失误,都会给测量参数带来损失。这是在心理测量中尤其要注意的。

3. 心理测量量表的设计

在心理学研究中,除了使用标准化的心理测验以外,常常需要自行设计和编制量表。心理测量量表是心理学研究的特殊而有效的工具,依照研究的目的不同,常用的有 5 种量表形式:总加量表、累积量表、一致定位量表、语义区分量表和行为定位量表。下面分别介绍和讨论各种量表的设计要领。

(1) 总加量表

总加量表又称为 Likert 量表,是最常用的心理测量量表。它以量表各个项目测量的总得分作为测量的指标,因此得名。这种量表通常采用 5 点或 7 点等距量表的形式,在设计时,要求在试测的基础上对各项目得分与总分之间的相关系数进行检验。由于总分包含着各项目得分,因而这一“项目-总分”相关系数往往高估,需要用以下公式加以校正:

$$r = \frac{r_{Ti}\sigma_T - \sigma_T}{\sqrt{\sigma_T^2 + \sigma_i^2 - 2r_{Ti}\sigma_T\sigma_i}}$$

式中, r_{Ti} 为第 i 项目得分与总分之间的相关系数, σ_T 、 σ_i 分别为总分和项目得分的标准差。在以上校正的基础上,应保留校正后高相关的那些项目,删除低相关项目,再进行信度、效度的检验和进一步的项目分析(难度或辨别力等),从而制成总加量表。

(2) 累积量表

累积量表强调项目的单维性(unidimensional),这与现代测验理论的假设十分相似。它以一组形成难度阶梯或按其他标准形成

阶梯的项目制成量表,测量同一维度特征,从而以阶梯项目累积得分来确定特定的心理特征水平。当需要测量多个维度的特征时,可以设计多组阶梯项目,构成累积量表。在对量表进行项目检验时,需要分析其同质性,假定在高一阶梯项目通过时,各个低阶梯项目也应通过,否则就称作“反转反应”。“反转反应”过多,则表示出量表的单维性不符合要求。累积量表以“复制指数”(R)计算反转反应的程度,其公式为:

$$R=1-\frac{e}{nK}$$

式中, e 为反转反应的数目, n 是被试数目, K 为项目数目。一般以 $R=0.85-0.90$ 作为可接受的边界。 R 低于这一范围则需要项目修订。

(3) 一致定位量表

一致定位量表是以心理物理测量作为基础的,其设计方法较为严密,因而量表质量比较高,特别适合于进行态度测量。设计一致定位量表时,先写出各个项目的陈述语句,可多写一些项目供筛选用。然后,由 50—80 位评判人以 11 点量表(最赞成—最不赞成)对每一项目打分,并以评级得分的中数作为该项目的量表值,以分数分布的第 75 百分位与 25 百分位之差作为项目的离散度。量表一般由 20—25 个项目组成,应使所选项目的量表值尽可能均匀分布于 1—11 分的范围里,在遇到相似量表值的项目时,应选择离散度较低的项目。在确定项目以后,采用“同意”和“不同意”作为反应的类别。在用一致定位量表进行测试时,以反应为“同意”的那些项目量表值的中数作为量表的得分。

(4) 行为定位量表

心理测量中常常会遇到这样的问题:由于量表项目的内容比较抽象,使测量结果很大程度上受到主观因素的影响,出现了偏差。为了避免这类问题,可以设计和采用行为定位量表,以具体的

行为表现(例如关键事件)对量表值作出定义,这种量表称为行为定位量表。在设计量表时,需要挑选一些对所测行为表现比较熟悉的人,把他们随机分成4个评定小组,按照以下5个步骤进行:①由第一评定组讨论所测行为的关键特征并确定行为表现的重要维度(方面),从而作出测量构思;②由第二评定组就第一组提出的重要行为维度,分别提出好、中、差3类关键行为事例;③由第三评定组进行“重译”,即让他们把随机排列的好、中、差的行为事例重新归类到各自的行为维度,并删除无法归回原类的事例;④第四评定组采用一致定位量表设计的方法,对经过“重译”而保留的项目确定其量表值;⑤对项目进行试测与评定,检验各项目的评级者间信度(inter-rater reliability)。如果信度符合要求,则以1—9点或1—11点量表轴标出行为事例相对应的量表值位置,从而制成量表,图3是行为定位量表的例子。研究表明,行为定位量表由于以具体的关键行为事例进行量表值定位,定义明确,测量误差较小,效度也比较高。

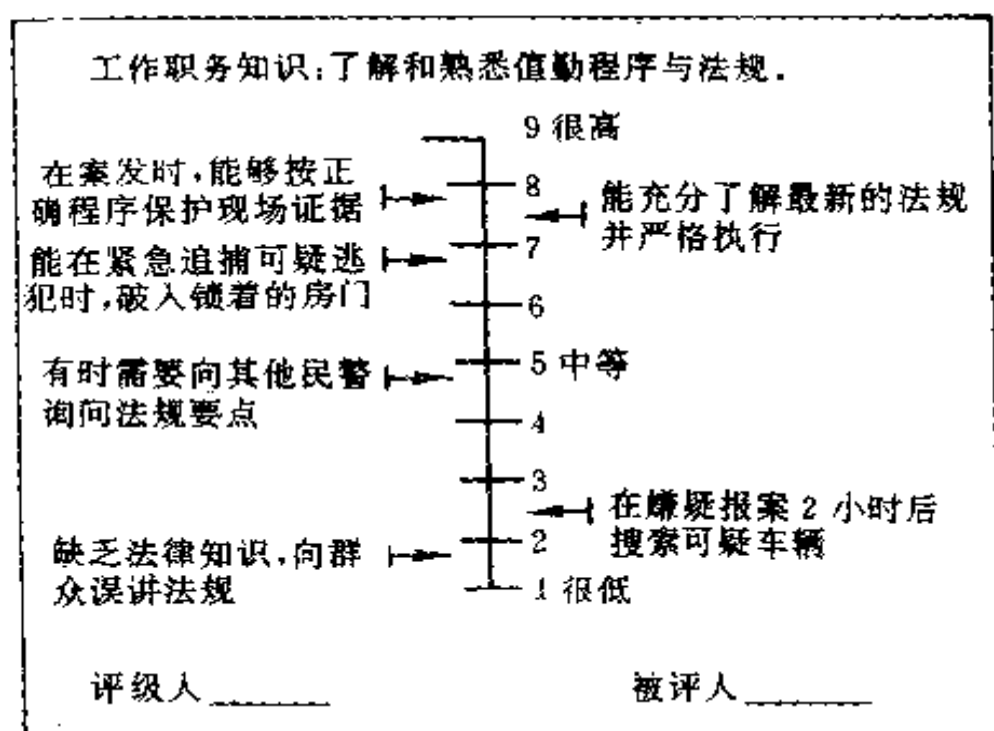


图3 民警值勤工作的行为定位量表(部分)

(5) 语义区分量表

语义区分量表采用了三维语义空间(评价、潜能和活动三维度)对关键概念的心理结构作出测量,从而评定有关的心理特征(如,态度、认知结构等)。量表的每一维度由 3—15 组双极形容词组成测量指标。这种量表在知识结构研究、心理语言学、社会心理学和管理心理学等许多领域得到广泛应用。管理心理学中著名的

为同类概念,这种方法又称为“距离-聚类分析”。还可以进行“被试间”聚类分析,找出相似认知或态度结构的群体。

心理量表的设计是一项复杂而严肃的工作,需要在理论构思、量表形式和项目结构等诸方面都作出细致、周密的考虑与安排。在设计涉及个性、认知、心理诊断和态度测量的量表中,应特别注重项目语句撰写的质量,避免专业术语、暗示、概念笼统、多重含义以及易受社会规范直接影响的语句,从而减少量表使用中可能引起的失真和偏差。同时,应特别重视测验技术的培训、心理测验的严格保管,避免心理测验的误用或滥用,并尽可能把心理测量与心理咨询、教育、培训、矫治等工作紧密结合在一起。

4. 知识获取与表征的测量方法

随着认知科学的新发展,许多新的心理测量方法得到广泛的应用,其中特别是一些与知识结构表征及知识获取有关的测量方法,在心理学各个领域的研究中得到完善与发展。

(1) 元素分类法

元素分类法(element sorting method)通过对概念元素的分类,对知识结构关系作出测量。具体做法是,先把与所测量的知识范畴有关的一系列概念写在卡片上,然后把不同概念的卡片充分混合,要求被试按照一定的标准确定概念之间的相似程度,以此分类堆放,并给每堆概念所属的类别命名。这样,根据各概念之间的相似程度和分类方式测量出被试在特定知识或态度范畴的心理结构。

(2) 全方格法

全方格法(repertory grid technique)起源于心理学家 G. Kelly 的人格构思心理学(personal construct psychology)。Kelly 认为,每个人用生活中形成的人格表征系统指导随后的行为。他为了验证和应用其理论,建立了全方格法,用以测量人格构思结构或知

识与态度结构。这种方法要求被试按照自己确定的维度标准,对认知元素(例如概念)作出分类,一般的做法是每次从备样元素中任取 3 个元素(概念),从中区分出两个比较相似而第三个相对来说不同的元素,并说明分类的理由(即维度标准);然后,利用方格距离确定这些认知元素在相应维度空间的位置,并对维度作出命名;以这样的方式,3 个概念成一组地做下去,可以测量出所有元素概念在多维空间的结构关系。由全方格法测得的数据,可用层次聚类分析法进行统计分析,也可以运用多维量表法(multi-dimensional scaling)分析基本维度,并确定其性质。使用多维量表法的条件是,概念组基本上属于同一知识范畴。研究表明,在多维量表法分析中采用三维解法时,比较容易获得理想的、语义明确的知识结构。

(3) 网络通径搜寻法

网络通径搜寻法(network pathfinder method)是由认知心理学家 Schvaneveldt, R. 等开发的知识工程测量与分析方法。这种方法也是以知识元素(概念)之间的相似性或接近性作为基本数据,用所测概念作为节点(nodes),以概念之间接近程度的最短加权连接线构建知识网络通径结构。这种知识结构有两个重要的参数: L 参数和 R 参数。 L 参数决定了概念节点的最大数目; R 参数则是多维空间通径距离的测度。通常,以 R 参数无穷大时不同 L 参数值来构建知识网络结构。我们在有关决策知识获取与测量的研究中运用了网络通径搜寻法。

(4) 过程追踪测量法

心理学研究中日益重视信息加工过程的分析与考察,在研究方法上逐步形成了适合于心理过程研究的测量手段——过程追踪测量法(process tracing technique),主要包括口语报告法(verbal protocols)和程序信息搜索法。口语报告法是从认知心理学有关人类解决问题的研究中发展起来的。这种方法由任务设计、口语报告、记录与转译、数据编码和结果统计分析等 5 个部分组成。其中,

应特别重视实验任务结构设计和数据编码系统设计这两个环节。为了保证测量的信度,对于口语报告记录的编码应由两人以上独立编码,以便提高编码质量,并以“评级者间信度”加以检验。程序信息搜索法注重于对信息加工过程中信息搜索路线、顺序和时间数据的测量。这种测量可以通过眼动监测仪测量被试注视与搜寻信息的轨迹,也可以对被试在实验中检索、搜寻信息项目的按键反应作实时追踪测定。

以上各种量表与测量方法各有特点,应根据研究的具体目的和构思,选用适当的量表形式和测量。

三、心理学研究的统计分析

有关心理学统计分析方法的著述已经很多,这里不准备重复介绍一般的统计计算方法,而是结合研究设计类型、测量数据特点和研究目的要求,讨论统计分析中比较新的思路、方法及其使用中可能遇到的问题。主要介绍复相关、协方差分析、因素分析和结构方程分析(LISREL)。

1. 复相关分析与心理预测

心理学研究中越来越重视多种因素的复杂相互作用和采用多种指标(变量)预测心理特征或行为表现的发展与变化。例如,用个性与能力变量的共同作用来预测学习成绩的变化,或运用数个认知变量共同预测心理发展的趋势。这种课题的统计分析,可以运用复相关方法(multiple correlation)。复相关分析可以处理多个变量与一个变量的关系,通常用于多个自变量(预测指标)和一个因变量(效标)之间关系的分析。复相关系数的公式为:

$$R_{1 \cdot 23}^2 = \frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{23}^2}$$

式中, $R_{1 \cdot 23}$ 表示变量 X_2, X_3 联合起来与变量 X_1 之间的相关, 一般用 $R_{1 \cdot 23}^2$ 表示所能解释的变异量(方差), r_{12}, r_{13}, r_{23} 分别是 X_1, X_2, X_3 之间的两两 相关系数(积矩相关)。由于心理活动是由多种因素决定的, 因此, 运用多个指标一般能够提高预测能力。可是, 在实际应用中, 并不是预测指标(变量)数目越多, 复相关系数就会越高, 这是因为随着指标数量增多, 测量误差的影响也相应增大了, 加上各个预测指标之间还会相互影响。因此, 预测指标的数量应该适宜, 不能过多。那么, 选用什么样的预测指标才能获得较好的预测效果呢? 心理学研究表明, 当选择 2 个互不相关的预测指标时, 可以获得较高的复相关系数。但是, 当 2 个预测指标之间具有高相关, 而其中之一与效标没有相关, 这一预测指标就可以解释另一指标的一部分残余变异(residuals), 即指标中不与效标关联的那部分变异, 起到抑制误差变异量、改善整个预测效度的作用, 我们把那个与效标没有直接关系的预测指标, 称为抑制变量。图 5 表示出上述变量(指标)之间的关系。图中预测指标 P_1 与 P_2 之间呈高相关, 而 P_2 与效标 C 没有什么关系, P_1 则与效标 C 有较高相关。这时, 把 P_1, P_2 结合起来预测 C , 有利于抑制 P_1 的误差变异而增强复相关和预测能力。

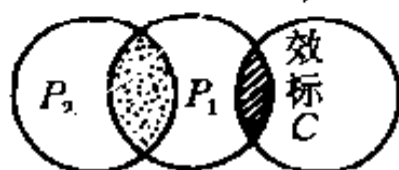


图 5 运用抑制变量提高效率

假定图中 $r_{P_1C} = 0.30, r_{P_2C} = 0.00, r_{P_1P_2} = 0.80$ 。这时, 虽然 P_2 与 C 没有直接关系, 但是 P_2 有可能产生抑制作用, 仍然能提高预测效应。从复相关公式可知:

$$\begin{aligned}
 R_{C \cdot P_1 P_2}^2 &= \frac{r_{CP_1}^2 + r_{CP_2}^2 - 2r_{CP_1}r_{CP_2}r_{P_1 P_2}}{1 - r_{P_1 P_2}^2} \\
 &= \frac{0.30^2 + 0.00 - 2 \times 0.30 \times 0.00 \times 0.80}{1 - 0.80^2} \\
 &= 0.25 \\
 R_{C \cdot P_1 P_2} &= 0.50
 \end{aligned}$$

可见,结合运用了 P_1, P_2 后,使预测效度系数从 0.30 提高到 0.50,显著加强了预测能力。复相关的这种特点使之在教育心理、儿童心理、临床心理、管理心理和社会心理等领域都有广泛的应用价值。在运用复相关方法时,应注意选用较大的样本(150—200 人),并对所涉及的变量(指标)进行仔细筛选。

2. 协方差分析与不等同对照组实验设计

许多心理学实验都采用了实验组与对照组进行事前-事后测量比较,以此推论实验处理的效应,并用 t 检验考察统计差异。可是,研究中难以真正做到两组统计比较所需要的随机化取样,因而属于“不等同”对照组实验设计。这种设计下很容易造成系统偏差,即两个组在与因变量有关的某些关键变量上出现系统的差异,使实验结果的统计分析失误。协方差分析(ANCOVA)就是一种能以很大程度上消除系统偏差和组内误差变异,它实际上是把方差分析与回归分析结合在一起的统计方法。在基本上做到随机分组的实验中,协方差分析的主要目的是减少误差变异,但如果每组被试人数小于 10 时,仍需要运用这种方法校正事后测量可能引起的偏差。在非实验研究中,协方差分析则主要用来校正事前、事后测试中的系统偏差。例如,某项研究比较了 4 种应激情景对于被试血压的影响,发现第三种情景比其他情景更容易引起心理紧张,但第三组被试即使在最低应激条件下,其血压仍比较高。这就需要运用协方差分析方法对 4 个组被试的血压进行均值校正。校正时,需要先确定关键变量作为协变量(co-variate),比如上例中被试进入应

激情景以前的血压就可以作为协变量。其它常用的协变量还有被试的智力(IQ)、个性特征、成就动机、态度以及先前的成绩等。有的情况下,可以对多个协变量同时作出校正。具体运算可以采用 SPSS 或 SAS 软件包中的相应软件进行,限于篇幅,不在这里详细讨论。对于协方差分析的结果,需要首先检验协变量的选择是否得当,一是检查协变量与因变量之间是否有线性关系;二是看各回归线的斜率是否基本相等。如果这两点都不符合要求,那么协变量选择就不适宜,还有一种可能是变量的测量指标有误。

3. 因素分析与因果模型构建

心理学研究统计分析的新趋势是日益重视多元分析和结构模型的构建。可是,统计分析方法的复杂化是以研究理论构思的精细化为条件的。因此,在运用多元分析方法时,应特别注重理论构思。多元统计分析的另一条件是合理选取样本,不但要求有一定的数量($N=150-200$),而且强调样本的代表性和数据的常态分布特征。此外,多变量的测量容易产生测量误差,这就对测量的信度和效度有更高的要求。

(1) 经典因素分析方法

经典因素分析方法又称为探索性因素分析,主要是从斯皮尔曼早期有关智力结构的研究发展起来的,对心理学各领域的理论和实践起着重要的作用。经典因素分析目前较多运用于变量分类、关系简化和量表项目结构分析。现今流行的各类统计软件包均有因素分析的程序。在因素分析的使用与运算中,应注意以下问题。

① 因素数目的确定:因素分析中需要确定基本因素的数目,通常以 Kaiser 的特征值等于 1 的标准或 Cattell 的陡阶检验法,即以各因素特征值连线陡降呈台阶后的那个因素为基准,确定因素数目。在确定因素数目时,还要结合理论构思上的某些考虑。研究表明,Kaiser 标准特别适合于主成分分析(principal components

analysis),并且在变量数目为20—50之间时最可靠。在许多统计软件包中,这种标准是“默认”(default)的因素抽取方法。

② 因素参照轴的旋转:在因素荷重计算的基础上,需要对因素的参照轴作出旋转,以便获得最明确、最有意义的因素模式。因素参照轴的旋转分为直转和斜转两种。直转法假定因素之间呈正交关系、相互独立,常用的是Kaiser提出的“方差最大”旋转法(VARIMAX),认为以方差最大为标准旋转时,可以对因素结构作出最清晰的解释。斜转法则假设所分析的因素参照轴相交成一斜角,相互间有一定的相关,常用的斜转方法是从直转法演变而来的,其中较流行的有“斜交因子最小”旋转法(OBLIMIN)。由于许多心理因素都是相互关联与影响的,因此,斜转法往往能获得比较符合实际的结果。在这种情况下,因素分析得到互有相关的一组基本因素,可以进一步作二阶因素分析(second-order factor analysis),寻求基本因素背后存在的深层次因素。

③ 基本因素的解释:在获得比较明确的因素荷重模式以后,应对基本因素的含义作出解释,以基本因素上具有较高荷重的变量的性质,对因素命名。因素解释中的一个常见问题,是对每一因素下“可解释方差量”的百分数作绝对的解释,简单地认为百分数最大的因素“最重要”,而忽视每一基本因素所包含的相对变量数目以及参照轴旋转方法的影响。另一个问题是对包含变量较少的基本因素作过分的解释,实际上,基本因素如果包含太少的变量,常常意味着因素结构不良或者分析方法不当。这些都是因素分析中应当注意的。

(2) 验证型因素分析

因素分析最新的发展是验证型因素分析的方法。这种方法通过结构方程模型法(structural-equation model),运用多重指标来表达不可观察的心理学潜变量(如焦虑、动机、智力等),从而以实际数据“验证”因素的结构关系。它对于心理学研究中的理论建模

具有显著的意义,已成为心理学各领域强有力的研究手段。验证型因素分析中最流行的软件是 LISREL 统计软件包。LISREL 以最大似然统计理论作为基础,包括测量模型和结构模型。前者表明外显可测指标与潜变量之间的关系,可用米检验测量的质量;后者则代表了所研究的潜变量中外源变量与内源变量之间的关系,说明某种理论构想或因果关系结构。在研究中,只有当测量数据与模型相拟合时,才能进一步分析潜变量之间的因果关系结构。结构模型的构建可以借用多元分析中的因径分析(path analysis)方法;所不同的是,因径分析采用可测变量直接建模,而验证型因素分析则要求运用潜变量,并以可测指标的因素荷重表示它们之间的关系。在 LISREL 结构模型中,以 ξ 表示外源潜变量(可看作自变量), η 为内源潜变量(可看作因变量),用 X, Y 分别表示外源与内源潜变量的可测指标(又称测量变量),以 λ 表示潜变量与测量变量间的关系(即因素荷重);用 γ 表示内源与外源变量之间的因果效应,以 ϵ, δ 表示可测指标的测量误差。心理学研究表明,验证型因素分析是一种严密的验证因果关系特征的方法,它所构建的模型具有参数不变的优点,并且较好地解决了测量误差、公共因素方差估算和参照轴旋转不变性等问题。由于验证型因素分析不像经典因素分析方法那样对所有因素荷重作出估算,而是只分析与理论假设有关的那些因素荷重,其余可测指标的荷重则设定为零值,因此,理论模型及其假设就显得至关重要。这是在使用这一方法时尤其要重视的。

(3) 因果模型的构建

心理学研究中日益强调理论模型的构建和理论的发展,特别是理论构思在多元统计分析中的重要性。因果模型的构建需要满足一些基本的条件,主要有 6 个方面:①模型的理论陈述应尽可能规范化,使之包括主要的原因与结果变量以及与之有关的中介变量,常用的办法是作出初步的结构关系图;②对变量之间的因果关