

# 信号检测论\*与古典的心理物理学方法

陈舒永 马谋超

北京大学心理学系 中国科学院心理研究所

从本世纪五十年代 Tanner 和 Swets 把信号检测论引入心理学的研究<sup>[1]</sup>以来,它已经成为研究心理物理学的一个有力工具。近年来利用信号检测论研究心理学问题的文章更是屡见不鲜。信号检测论对心理物理学的最大贡献是它可以把被试的辨别力和他的反应倾向性分开,而以古典心理物理学测得的感觉阈限则是二者不可分割的混合物。由于信号检测论的引入,确实解决了一些古典心理物理学方法所不能解决的问题。例如,关于精神病患者与正常人的大小常性,用古典心理物理学研究的结果很不一致。有的实验结果表明,精神分裂症患者和正常人比较,前者倾向于超常性;有的实验结果恰好与此相反;也有的实验结果证明,在大小常性的问题上二者并无差别。后来用信号检测实验得到的结果表明,在排除反应的倾向性以后,非妄想型精神分裂症患者比正常人大小常性的程度要低得多<sup>[2]</sup>。又如用古典心理物理学测定感觉阈限时,如果主试的指导语改变了,感觉阈限会随之发生变化。究竟指导语影响了被试的辨别力,还是影响了他的反应倾向呢?古典心理物理学就回答不了这个问题。用信号检测实验得到的结果表明,在不同指导语的影响下测得的感觉阈限所以不同,不是由于被试的辨别力发生了变化,而是由于改变了判断标准造成的结果<sup>[3]</sup>。另外还有些信号检测实验表明,用古典心理物理学测得的痛阈提高了,并不一定意味着痛觉感受性的降低,而常常是由于改变了报痛标准造成的<sup>[4]</sup>。

随着信号检测论的引入,把心理物理学的研究确实向前推进了一大步。正因为如此,也就提出了一个问题,那就是有了信号检测的方法以后,古典的心理物理学方法是否还有用处?是不是它已经完成了它的历史任务可以送进博物馆去了呢?下面的实验就是为了探讨这个问题而设计的。

## 实验方法

**实验一** 实验材料是两个高 50 mm、直径 40 mm 装有铁砂的圆柱体。它们的重量分别为 100 克和 108 克,后者作为信号,前者作为噪音。我们所以选择这两个重量,是因为在预备实验中被试有 60% 到 70% 的次数判断 108 克比 100 克重。

我们采用的是信号检测实验的有无法 (yes-no method)<sup>[5]</sup>。被试戴上眼罩,在排除视觉的条件下,以优势手(本实验的被试均为右手)提举重量。实验前让被试将两个重量比较 3—4 次,为的是让他练习一下提举重量的速度和高度,以便在实验中尽量保持一致;也可以使他对哪个轻哪个重有一个清晰的印象。实验时每次只呈现一个重量。让被试提举一

\* Signal detection theory, 简称 SDT, 有几种中文译名,如信号侦察论,信号觉察论,信号检测论等等,我们采用了最后一种译法。

次后,立刻对该重量是“轻的”(噪音)还是“重的”(信号)做出绝对判断。信号和噪音呈现的概率都是·50·每轮实验中信号和噪音各呈现 100 次。每对  $p(y|SN)$  和  $p(y|N)$  都是根据 200 次的实验结果计算出来的。每呈现 50 次休息三分钟。每 50 次实验中呈现信号和噪音各 25 次,呈现的顺序是随机排列的。每个被试共进行两轮实验,上下午各一轮。两轮实验都不是在同一天内进行的,一般相隔两三天,有个别被试两轮实验隔了一个月之久。实验前只告诉被试每次提举重量的速度和高度尽量保持前后一致,提举后立刻报告是轻的还是重的;不告诉他信号和噪音呈现的概率;也不对他的正确和错误的反应施加奖惩;使实验的情景尽量与古典心理物理实验一致。

被试共 14 人,男女各半,年龄 9—59 岁。其中 15 岁以下者 5 人,均系中小学生。成年被试中有工人、干部、保育员、实验员、资料员、教员和研究人员。在数量有限的被试中,年龄和职业的范围还是相当广泛的。

**实验二** 为了把我们的实验条件和反应倾向性变化比较明显的情况进行比较,我们用同样的实验材料,又在改变信号和噪音的先验概率的条件下进行了实验。

在这个实验中,我们用了两个相差较大的先验概率,即  $p(SN) = .2$  和  $p(SN) = .8$  的两个概率。每次实验前都把信号和噪音出现的概率,用通俗的语言向被试说清楚。例如,当使用  $p(SN) = .8$  的概率时,就告诉被试,在这 50 次实验中,重的将出现 40 次,轻的只出现 10 次,重的出现的次数多,轻的出现的次数少。每种概率共进行实验 100 次,分成两个 50 次进行。用两种先验概率所进行的 4 个 50 次实验,一半被试以 2、8、8、2 的顺序安排,另一半被试以 8、2、2、8 的顺序安排。其他实验程序和实验一相同。被试共 8 人,男女各半,年龄 24—59 岁。除没有用实验一里的儿童被试外,这里的被试绝大部分都和实验一的被试相同。

## 实验结果和讨论

根据实验一每人两轮实验的数据,我们按上下午分别计算了  $d'$  (辨别力或感受性)、 $\beta$  (判断标准)和  $DL$  (差别阈限)。 $DL$  是把  $p(y|SN)$  和  $p(y|N)$  转换成  $z$  分数以后,用画图法取 75% 点求得的。所计算出来的  $d'$ 、 $\beta$  和  $DL$  均列在表 1 中。

从表 1 所列的结果可以看到,在不同时间内测得的  $d'$ 、 $\beta$  和  $DL$ , 其各自的差别都很小,经考验均不显著。这说明在我们的实验条件下,不仅  $d'$  可以保持恒定, $\beta$  和  $DL$  也基

表 1 在上、下午两轮实验中  $d'$ 、 $\beta$  和  $DL$  的比较

实验时间	$d'$	$\beta$	$DL$
上午	1.80	0.75	8.70
下午	1.90	0.92	8.40
差异	0.10	0.17	0.30
$t$	0.05	0.95	2.00
$df$	13	13	13
$p >$	0.50	0.20	0.05

本上是稳定的,不受实验条件不同的影响。

$DL$ 、 $d'$  和  $\beta$  彼此之间的积差相关系数的计算结果见表 2。

表 2 在上、下午两轮实验中各参量的相关

参量	积差相关系数	$df$	$p$
$DL$ 和 $d'$	-.81	26	<.01
$DL$ 和 $\beta$	+.14	26	>.40
$\beta$ 和 $d'$	+.003	26	>.50

从表 2 所列结果可以看到,  $DL$  和  $d'$  为负相关, 相关量达到 .81, 十分显著; 而  $DL$  和  $\beta$  以及  $\beta$  和  $d'$  的相关量都很小且不显著。  $DL$  和  $d'$  有很高的负相关, 说明以恒定刺激法求得的感觉阈限和以信号检测法求得的感觉性之间有很高的一致性, 即差别阈愈大, 差别感受性愈小, 反之亦然。

由于  $d'$ 、 $\beta$  和  $DL$  在两轮实验中都保持恒定, 且  $DL$  和  $d'$  有很高的负相关, 可以看出, 在我们的实验条件下, 用古典的心理物理学法测定的感觉阈限和用信号检测法测定的  $d'$ , 可以同样好地反映被试的辨别力。

实验二的结果列在表 3 和表 4 中。

表 3 两种先验概率下  $d'$ 、 $\beta$  和  $DL$  的比较

$p(SN)$	$d'$	$\beta$	$DL$
.2	1.60	1.75	12.29
.8	1.59	0.54	7.04
差异	0.01	1.21	5.25
$t$	0.06	2.78	5.77
$df$	7	7	7
$p$	>.45	<.025	<.005

从表 3 的结果可以看出, 在两种不同的先验概率下,  $d'$  的差别很小, 很不显著。而  $\beta$  的差别和  $DL$  的差别却都是非常显著的。这说明  $d'$  确实不受先验概率的影响, 而  $\beta$  和  $DL$  都因先验概率不同而各异。

表 4 两种先验概率下各参量的相关

参量	积差相关系数	$df$	$p$
$DL$ 和 $\beta$	+.68	14	<.01
$DL$ 和 $d'$	-.52	14	<.05
$d'$ 和 $\beta$	-.02	14	>.05

从表 4 的结果可以看到,  $DL$  和  $\beta$  有相当高的正相关,  $DL$  和  $d'$  的负相关也是显著的, 而  $d'$  和  $\beta$  是不相关的。这说明  $d'$  和  $\beta$  是两个独立的参量, 前者代表被试的感受性, 后者代表被试反应的倾向性。  $DL$  和  $\beta$  和  $d'$  分别有显著的正相关和负相关, 但相关量都不

很高,恰好说明  $DL$  是  $\beta$  和  $d'$  的混合物。

我们知道,信号检测论是用  $d'$  和  $\beta$  这两种指标独立地测量被试的感受性和主观态度的;而古典的心理物理实验则是通过测量感觉阈限来测量感受性的。被试的主观态度属于非感觉方面的因素。由于古典的心理物理法不能把在实际作业中混在一起的这两类因素分别进行测量,通常是借助于实验程序的安排和指导语来控制被试的主观态度的,并假设在控制了主观因素的条件下,所测得的感觉阈限是可靠的。但这种假设的有效性究竟怎样呢?我们的实验结果为回答这个问题提供了论据。我们在实验一中所用的指导语是简单而明确的。我们只要求被试判断一个给定的重量是重的还是轻的。其他实验程序的安排,除了呈现的刺激数量较少(只有一个信号和一个噪音)外,和一般的做绝对判断的古典心理物理实验是十分类似的。在实验一中我们避免了人为地对被试施加影响以改变他的反应倾向的措施,也就是期望在这种条件下,被试的反应倾向能保持恒定,从而不致影响所测量的感觉阈限。另外,在古典的心理物理实验中,对于象进行实验的时间,被试的年龄和过去的经验等实验条件,本来也是尽量地加以控制的。而我们在实验一中都有意地把两轮实验安排在不同日的上午和下午,所用的被试从9岁的儿童到59岁的老人,从小学二年级的学生到大学的教师,他们在年龄和经历上的差异也是比较大的。我们所以做这样的安排,为的是要看一看,在这些条件变化比较大的情况下,被试的反应倾向会不会发生变化,从而影响到所测量的感觉阈限。从实验结果看来,即使在这些实验条件变化比较大的情况下,两轮实验的  $\beta$  值的差异和  $DL$  值的差异,仍然是不显著的;并且  $DL$  和  $\beta$  的相关也不显著。这说明即使在我们没有严格控制的实验条件下,那些非感觉方面的因素对感觉阈限的影响仍然很小,仍然可以忽略不计。而这里唯一显著的乃是  $DL$  和  $d'$  的负相关。根据信号检测论, $d'$  是不受反应倾向影响的参量。在实验一中出现了  $DL$  和  $d'$  很高的负相关,说明在上述的场合下,所测得的  $DL$  的确是可靠的。根据古典的阈限理论,感受性是以阈限的倒数来定义的,因此在这种条件下,两种方法对于感受性的测量是相通的。Treisman 和 Watts 在讨论信号检测论和测量感觉阈限的传统方法之间的关系时,也曾指出两种方法之间的一致性<sup>[6]</sup>。所不同的只不过是他是从用恒定刺激法做的实验结果来估计  $d'$  的,而我们则是从信号检测的实验结果来估计  $DL$  罢了。

可是在实验二中,由于先验概率对被试的反应倾向有影响,在不同的先验概率下, $\beta$  值的差异是很显著的, $DL$  值的差异也达到了非常显著的水平,并且  $DL$  和  $\beta$  是密切相关的。只有  $d'$  的差异很不显著,考验的结果, $p > .45$ ,并且  $d'$  和  $\beta$  的相关量只有 .02。这说明在实验二的场合下,被试的反应倾向确实影响了感觉阈限的测定,而  $d'$  则不受反应倾向的影响。我们的实验结果还从量的方面说明了它不受反应倾向影响的程度。正是在这里显示出了信号检测法比古典的心理物理法优胜的地方。

从上面对于两种实验结果的对照和分析可以看到,在通常的情况下,当被试的反应倾向变化不大,或者实验条件虽然发生变化但不致影响被试的反应倾向时,古典的心理物理法和信号检测法对感受性的测量是同样有效的。但当某些主观因素或人为地加在被试身上的措施足以影响他的反应倾向时,则用信号检测法比用古典的心理物理法测量感受性更为可靠。

总之,我们一方面要看到信号检测论对心理物理学研究的巨大贡献,要利用这个有

力的工具去研究解决一些古典心理物理学所不能解决的问题; 另一方面也要看到古典的心理物理学, 在一定条件下仍不失为测定感受性的一种可靠的方法, 并且用起来比信号检测法要简便些; 此外, 即使在进行信号检测实验以前确定信号和噪音的差别时, 也是离不开古典心理物理学方法的。

### 参 考 文 献

- [1] Tanner, W. P., Jr. and Swets, J. A.: A decisionmaking theory of visual detection, *Psych. Rev.*, 61, 401—409, 1954.
- [2] Price, R. H. and Eriksen, C. W.: Size constancy in schizophrenia: A reanalysis, *J. abnormal Psych.*, 71, 155—160, 1966.
- [3] Clark, W. C.: The psyche in psychophysics: a sensorydecision theory analysis of the effect of instruction on flicker sensitivity and response bias, *Psych. Bull.*, 65, 358—66, 1966
- [4] Clark, W. C. and Yang, J. C.: Acupunctural analgesia? Evaluation by signal detection theory, *Science*, 184:1096—1097, 1974.
- [5] Green, D. M. and Swets, J. A.: *Signal detection theory and psychophysics*, New York: Robert E. Krieger Publishing Co., 1974.
- [6] Treisman, M. and Watts, T. R.: Relation between signal detectability theory and the traditional procedures for measuring sensory thresholds: Estimating  $d'$  from results given by the method of constant stimuli, *Psych. Bull.*, 66, 438—454, 1966.

## A NOTE ON SDT AND CLASSICAL PSYCHOPHYSICAL METHODS

Chen Shu-yong

(*Department of Psychology, Peking University*)

Ma Mou-chao

(*Institute of Psychology, Academia Sinica*)

In order to compare the merits of SDT and classical psychophysical methods two experiments on weight discrimination were carried out on subjects of different ages, sexes and occupations. Both SDT and the method of constant stimuli were used. It has been shown that when the subject's response bias is not influenced by the change of experimental conditions, the measurement of sensitivity by the method of constant stimuli is just as efficient as that by SDT. But when some non-sensory factors have an effect on the subject's response bias, the measurement of sensitivity by SDT is more reliable. The authors hold that although SDT has been considered a powerful tool to measure sensitivity, the classical psychophysical methods are still useful under certain appropriate conditions. They are indispensable methods in sensitivity studies, and cannot be replaced by SDT.

# 记忆与学习的脑化学研究

## II. 海马内注射胰蛋白酶与胰凝乳蛋白酶对大白鼠暗箱迴避模式记忆的影响

罗胜德 李德明 孙丽华

中国科学院心理研究所

在实验探索“记忆之迷”的进程中,近十余年来,有关脑内蛋白质与肽类基质的重要职能,已越来越为广大学者所重视。以弗莱克斯纳 (J. B. Flexner 等, 1963)<sup>[4]</sup>、格拉斯曼 (E. Glassman 等, 1964)<sup>[5]</sup>、海登 (H. Hyden 等, 1966)<sup>[6]</sup>、与昂格 (G. Ungar 等, 1971)<sup>[9]</sup>为代表,借助蛋白质类代谢抑制剂、同位素示踪、生化分析与生物测定等手段,曾经取得一些实验资料,表明脑内(特别是海马部位)的蛋白质,是痕迹固着必不可少的环节。然而,具有不同结构组成的种种蛋白质,它们在不同水平的种种模式记忆中,究竟起什么作用?这一方面,与其说至今缺乏雄辩的科学假设,还不如说是缺乏令人信服的实验证据。

酶的主要特征之一,就是对其底物作用的高度专一性<sup>[3]</sup>。借助作用机制明确的种种酶,有可能揭示其底物的功能,因此,我们提出了工具酶药物 (instrumental enzyme drugs) 研究途径的思路<sup>[1]</sup>。前一研究首先选择了两种工具酶:胰核糖核酸酶 (RNase) 与胰蛋白酶 (Trypsin), 它们的作用底物是极其不同的核糖核酸 (RNA) 与蛋白质类生物大分子;通过海马内微量注射对暗箱迴避模式记忆巩固组与不巩固组的比较研究,结果发现了胰蛋白酶/不巩固组 (T/NC) 的遗忘效应<sup>[1]</sup>。可是,是不是作用底物为蛋白质与肽的种种酶制剂的作用皆同?本实验进而引用了另一种工具酶——胰凝乳蛋白酶 (Chymotrypsin), 旨在比较研究同属水解蛋白质与肽类底物的两种内切酶(胰蛋白酶与胰凝乳蛋白酶<sup>[2,7,10]</sup>) 对于同一种模式记忆的巩固保持,影响究竟是否相同。

### 实验方法

基于前一研究的结果与讨论<sup>[2]</sup>,本实验免设录用。被试 26 只大鼠达到标准后给药,分别于 72 小时后(第 8、9 实验日)、半日后(第 22、23 实验日)与一月后(第 36、37 实验日)进行记忆保持测验。最后进行统计分析的三项指标同前,即分别代表条件反射的质、量与速度的条件反射连续出现的次数 ( $SN_{CR}$ )、条件反射的百分率 ( $CR\%$ ) 以及条件反射的平均反应时 ( $RT_{CR}$ )。

本研究所用实验仪器同前<sup>[2]</sup>。

### 实验结果

海马内注射生理盐水、胰蛋白酶与胰凝乳蛋白酶,对暗箱迴避条件反射各项指标的影