

第一章 引 论

在实验心理学建立之前，心理学附属于哲学，还不是一门独立的学问。实验心理学的建立使心理学成了一门独立的科学。实验方法在心理学中的应用为心理学提供了收集材料的新手段，迅速改变了心理学的面貌，使它的发展大为加快。从实验心理学的建立到现在仅百余年，心理学的发展超过了以往许多世纪。现在，从感觉、知觉、注意、表象、学习、记忆、思维、言语到情绪、人格、意识各个专题，无一不应用实验方法进行着卓有成效的研究。在教育心理学、临床心理学、管理心理学中广泛应用的智力测验、能力倾向测验、成就测验、态度测验等，都应用了心理物理学的量表法。世界上第一本工程心理学的书就以《应用实验心理学》命名，《航空航天实验心理学》更是实验心理学在最复杂的劳动（操纵飞行器）中的具体应用。总之，可以说实验心理学乃是各门心理学的基础。

20 世纪 50 年代中期到 80 年代中期，认知心理学或信息加工心理学成为心理学中的主流。认知心理学把心理过程理解为一系列的信息加工过程，从机能上把人脑与计算机进行类比，因为计算机正是对信息作出一系列处理的加工系统。认知心理学强调，心理学的研究对象就是人的内部心理过程。但是，人们不能直接观察内部的心理过程，只能依靠输入（刺激）和输出（反应）的东西推测内部发生的过程。因此，认知心理学特别依赖于实验方法，只有正确而巧妙的实验设计才能保证推测到的内部心理过程——信息加工过程的可靠性。

许多心理学家认为，20 世纪 80 年代中期至今，认知神经科学是心理学最有希望的发展方向。认知神经科学把心理与大脑的统

一理解作为自己的目标，强调心理活动是大脑的功能这个简单的真理。它要求在具体的实验中，把心理过程与实时相应的神经过程联系起来。为满足这样的要求，研究者必须首先依靠实验设计来确认某种心理过程，只有在这样的前提下，通过仪器记录到的脑内变化过程才是有意义的，才是具有某种明确功能（心理活动）的脑活动。

综上所述我们可以说，心理现象是用自然科学的方法——实验的方法来研究的，也因为如此，心理学才能够不断地迅速发展。而且，使用实验心理学的实验方法获得的经验数据，对于评估关于人的智能的计算机模型，对于理解大脑永远是不可缺少的。

一、实验与观察

自然观察是许多科学分支的主要研究方法。天文学的发现是靠观察得到的，达尔文的进化论建立在对自然界周密的观察之上。儿童心理学、动物心理学也都依靠观察获得了重要的发现。在“文化大革命”期间，著名的心理学家荆其诚在干校养猪时，曾观察到非常有趣的母猪哺育幼仔的行为模式：母猪先后分娩仔猪以后，每一仔猪立即找到母猪身上的一个合适的乳头，而且从此以后始终从这一乳头吃奶。更有趣的是，母猪在喂奶时会发出一种特殊的呼噜鼾声，一旦它在喂奶中途入睡，停止发出呼噜声，其乳腺也会停止分泌乳汁；而这时正在吃奶的小猪中也必会有一头仔猪从侧卧的母猪背后绕到母猪的头部，用鼻子拱动母猪的鼻子，似乎在通知母猪继续放奶；于是母猪醒来继续发出呼噜声并分泌乳汁。这些行为模式在每一窝中都以完全相同的方式出现，以达到母猪哺育后代的目的（荆其诚，1990）。

实验方法是心理学研究的主要方法。自然观察只能等待所要观察的事物出现时才能进行，或只能对已有的事物进行观察；而实

验方法则是研究者主动控制条件下对事物的观察，它能对所观察的现象作出因果性的说明。

实验方法有如下的一些特点或优点：

①实验者总是带着特定的目的去进行实验。这样，他至少知道他将要观察行为的哪些方面，什么时候去观察它们。也就是说，实验者规定了他将要研究的事物。

②实验者设置的实验条件为他的观察创造了最好的条件，他可以在做好去测量和记录的充分准备时开始实验。这样，通过控制某事件的发生，他就可以使它重复产生，以便确信某种现象是不是前后一致。

③实验者设定了明确的实验条件，别人就可以重复实验，对他的结果做独立的检验。

④实验者可以控制一切条件，使之恒定，只改变某一条件，看实验结果是否就是这个条件引起的（Killing and Riggs, 1975）。

简要地说，实验方法可以‘产生’新的现象，这些现象是在自然状态下观察不到的。实验方法可以发现事物之间的因果关系。新的现象以及事物之间的因果关系都可以由不同的研究者各自做出检验。而且，实验是随时随地都可以进行的。

二、心理学实验和它的各种变量

一项心理学实验包含 3 种变量：自变量（independent variable）、因变量（dependent variable）和额外变量（extraneous variable）。那么，什么是心理学实验呢？心理学实验要求额外变量保持恒定，而仅仅操纵自变量去影响因变量。并且它还设定一个虚无假设（a null hypothesis）：因变量的平均值在不同的实验条件下没有显著差异，如果所获得的实验数据拒绝（或否定）虚无假设，那么实验者就得到了一个可靠的结论，即一点有用的信息——因变

量是明显地受自变量影响的 (Simon, 1974)。至今为止, 心理学知识的累积基本上都是在这个框架内实现的。这样的心理学实验框架包括两部分内容: 一是实验设计, 即怎样操纵自变量去影响因变量; 二是数据分析, 即对虚无假设进行显著性检验的问题。实验设计与数据分析是第一章要讨论的内容, 我们现在只对自变量、因变量与额外变量作一些介绍。

为了叙述方便, 我们先介绍两个实验心理学中常用的名词: 主试和被试。主试(experimenter)就是实验者, 即主持实验的人, 他发出刺激给被试, 通过实验收集心理学的资料。被试 (subject) 就是实验对象, 接受主试发出的刺激并作出反应。人与动物都可以作为被试。

下面我们介绍一个简单反应时间的实验。我们很熟悉这样的情景 在运动场上, 裁判员喊‘预备’口令不久, 枪响了, 百米赛跑的运动员一跃而起奔跑冲刺。从枪响的一刹那到运动员开始手离地面的一刹那这一段时间, 心理学上称之为反应时间, 简称反应时。一个人反应时间的长短可以用实验测量出来: 让一个被试安静地坐在桌子旁边, 主试说‘预备’后, 被试立即用右手轻轻压着一个反应键, 大约 2s 后主试给出声音, 被试一听到声音应尽快松开键。从声音出现到反应键松开的时间就是反应时间, 它可以由一个电钟记录下来。为了准确地测得某个人的反应时间, 测量应该进行几百次, 然后求得这些数据的平均值(毫秒)即为该人的反应时间。以上进行的是对声音的反应时间实验, 测得的是某人的听觉反应时间, 如果把声音刺激换成灯光刺激, 那么测得的反应时间就是视觉反应时间。

(一) 自变量

自变量即刺激变量, 它是由主试选择、控制的变量, 它决定着行为或心理的变化。在上面的实验中, 声音和灯光就是自变量, 它

决定着反应时的长短。主试选择自变量的目的是用自变量来改变行为。例如,主试如果用声音作刺激,就测得听觉反应时;如果用灯光作刺激,那么测得的就是视觉反应时。而视觉反应时总是比听觉反应时长,这就是说,由灯光引起的行为反应与由声音引起的反应快慢是不同的。还有,如果主试增加声音的强度,反应时间就会缩短,这也是行为的变化。强的声音和弱的声音都叫做声音的自变量,但它们处在不同的水平;强的灯光与弱的灯光引起的反应时也不同,这两种灯光刺激也处在不同的水平。当自变量的水平(数量)有了变化并导致行为的变化,我们就说行为是处在自变量的控制之下,或者说,自变量是有效的。

自变量可以分为 4 种:

1. 刺激特点自变量

刺激的不同特性会引起被试不同的反应。例如,灯光与声音引起的反应时不同,强度不同的声音引起的反应时也不同,我们把这类自变量称为刺激特点自变量。在记忆实验中,主试要求被试学习 50 个单词,这些单词也许是常见的,也许很少见到,那么单词在书刊报纸中出现的频率就是它的一个特点,我们可以研究单词的频率对再认的影响。在心理语言学实验中,句子的不同类型,如肯定句与否定句、主动句与被动句,就是刺激特点自变量,它可能会影响用句子匹配相应图画的快慢。

2. 环境特点自变量

进行实验时环境的各种特点,如温度、是否有观众在场、是否有噪音、白天或夜晚等,都可以作为自变量。在记忆实验中,两组被试都在同一实验室学习,但在测验时,第一组被试在原来实验室进行,而第二组被试换一间实验室进行,研究者想要知道,不同的测验环境是否对记忆有影响。这可以说是典型的环境特点自变量。初学者常常忽视了时间这个环境特点自变量。在暗适应过程中,时间是一个最重要的自变量。正是随着时间的流逝,处在黑暗中的眼睛

的感受性才逐渐提高了。时间这个自变量在记忆研究中是如此重要和无处不在，你甚至可以说，几乎没有不用时间作自变量的记忆实验。

3. 被试特点自变量

一个人的各种特点，如年龄、性别、职业、文化程度、内外倾个性特征、左手或右手为利手、自我评价高或低等等，也都可以作为自变量。老年记忆的研究中常把老人的记忆与青年人的记忆作比较；在儿童心理学的研究中年龄、性别的选择十分重要。对于被试特点自变量，主试只作选择而不能改变它，这和主试可以任意调节声音刺激的强度是不同的。

4. 暂时造成的被试差别

当被试来到实验室时，他们在各方面都是大致相同的。但是当主试对被试进行分组时，一组被试与另一组被试的差别便产生了。例如，研究者对 3 种不同的学习方法是否对记忆有不同影响感到兴趣。第一组被试对每次呈现的 3 个单字机械复述即死记硬背；第二组用每次呈现的 3 个单字造一句子；第三组对 3 个单字所代表的实物形象进行想像。例如，如果 3 个单字是狗、水、月，那么被试就可以想像一只狗对月亮汪汪的时候掉进水坑了。三组的学习的时间是相等的。这样，三组被试由于使用的学习方法不同而产生了差别。这种差别可以产生不同的记忆效果。应该指出的是，被试的暂时差别通常都是由主试给予不同的指示语而造成的。

(二) 因变量

因变量即被试的反应变量，它是自变量造成的结果，是主试观察或测量的行为变量。在刚才叙述的记忆实验中，三组被试学习完毕即进行再现和再认测验，能够再现和再认的单字在全部学过的单字中的百分比就是反应变量。在对声音和灯光的反应时实验中，测得的反应时的长短就是反应变量。对因变量的测量与选择有下

面几个问题需要讨论。

1. 因变量的可靠性即信度

信度是指一致性，同一被试在相同的实验条件下应该得到相近的结果。如果同一被试在相同的实验条件下有时（结果）得分很高，有时得分很低，那么我们就可以说，这种因变量（或测量被试反应的方法）是不可靠的，它缺乏一致性。

2. 因变量的有效性即效度

当确是自变量而不是其他各种因素造成了因变量的变化时，我们就说这种因变量是有效的。例如，在问题解决的实验中，你规定的因变量（反应变量）是在一定时间内被试解决问题的多少。看来这是一个很有效的因变量。当要解决的问题很多而且是按困难程度越来越难排列时，解决问题的数目作为因变量是有效的；但是，如果问题很多但非常容易，那么解决问题的数目就不能说明或测量一个人解决问题的能力，而只不过是说明或测量他的阅读速度罢了。

如果因变量的变化不是由自变量造成的，而是由其他的因素造成的，我们就说这种因变量是无效的，或者说，产生了自变量的混淆（例如将阅读速度与解决问题的能力混淆了）。对这个问题下面我们还要专门讨论。

3. 因变量的敏感性

自变量发生变化可以引起相应的因变量的变化，这样的因变量是敏感的。如果自变量的变化不能引起相应的因变量的变化，我们就说这样的因变量是不敏感的。不敏感的因变量有两类典型的例子，一类叫高限效应(ceiling effect)。当要求被试完成的任务过于容易，所有不同水平(数量)的自变量都获得很好的结果，并且没有什么差别时，我们就说实验中出现了高限效应。例如，你要求被试学习 30 个单字，有些人读一遍，有些人读两遍，有些人读三遍。测验表明，3 种情况下被试都能再认 90%以上的单字。在这种情况

下,不同的学习遍数(即自变量的不同水平)并未造成再认百分比(因变量)的变化。这时我们就说,在这种情况下再认这一指标(因变量)是不敏感的。另一类不敏感的因变量例子是低限效应(floor effect),当要求被试完成的任务过于困难,所有不同水平的自变量都获得很差的结果,并且没有什么差别时,我们就说实验中出现了低限效应。在刚才的例子中,如果你要求被试学习 200 个单字而不是 30 个,测验又由再认改为再现,结果表明,3 种情况下被试的再现百分比都不超过 20%。

如果你在实验中选择的是不可靠的、无效的或不敏感的,那么你从实验结果中所得出的结论也就并不能证实你的假设,而只是反映了实验方法中的毛病而已。

(三)额外变量或自变量的混淆

当我们在一次实验中确定了自变量与因变量以后,就应该使实验的其他条件保持恒定,因为只有这样实验中的因果关系才能得到明确的说明。所以,额外变量就是在实验中应该保持恒定的变量。

例如,在听觉反应时的实验中,下列变量应保持恒定:声音的频率与强度 预备时间 反应方式(听到声音松开键,还是先松开键听到声音后去按压键)用利手反应等等。对任何一项实验来说,需要控制的变量都是极多的——它要比实际上能够控制的多得多,但是,只要我们把足以影响因变量的一些主要因素控制住,实验结果就是可靠的。我们怎样决定哪些因素必须控制住呢?这要参照别人有关的实验以及凭你自己的实验经验。

如果应该控制的变量没有控制好,那么它就会造成因变量的变化。在这种情况下,研究者选定的自变量与一些未控制好的因素共同造成了因变量的变化,这就叫自变量的混淆。因此我们也可以说,额外变量就是潜在的自变量。

例如,在对声音的反应时实验中,如果没有控制好预备时间,预备时间有时较长(如 4s),使被试的准备松弛了;有时较短(如 0.5s),又使被试来不及准备,这都会对反应时产生不利的影响。结果测出来的反应时的长短就不仅是由声音决定的,同时也是由预备时间决定的。这样,如果我们实验的目的是要考察声音强度对反应时的影响,但由于预备时间忽长忽短也造成了反应时的变化,那我们就不能得出关于声音强度与反应时关系的正确结论。

三、实验范式

实验方法在各种心理过程以及各个心理学分支的研究中,具体表现为各种不同的实验范式(the experimental paradigms)。例如,研究言语过程的一种实验方法——斯楚普测验(stroop test)就是一种实验范式。在斯楚普测验中使用一系列颜色词(红、绿、黄等),但词义与书写该词的颜色不匹配。例如“红”字用绿色写,“黄”字用红色写等。实验中当要求被试尽快说出字的颜色时,被试常常自动地首先把字读出来了,这就是颜色命名的过程与读字过程的竞争。由于阅读是一种自动化过程,因此人们倾向于先读字,这样,字义的自动加工过程就干扰了有意的颜色命名的过程。自从 1935 年斯楚普测验范式首次出现以来,在这几十年里与它有关的研究已经超过 700 篇文献,研究者通过将它稍作改动已将它广泛应用于个别差异、年龄差异、性别差异、大脑左右半球差异等领域(Macleod, 1991)。近年来,人们又将斯楚普范式应用于记忆、无意识知觉的研究中(耿海燕、朱滢,1998)。记忆研究中的“加工水平效应”(the level effect of processing)也是一个能被用来说明什么是实验范式以及它的作用的恰当例子。在非随意学习中(incidental learning),要求被试回答有关单字的各种问题,如这个单字是用大写字母写的吗?这个单字与某个字是否押韵?这个单字能否归为

动物一类？回答这些问题要求进行不同水平的加工，而后来的记忆成绩也就是不同水平加工的函数。深层的有关意义的加工（如归类）导致最好的保持，其次是语音的加工，最差的是字型的加工。应用加工水平的实验程序研究记忆曾风靡一时，至今，它仍然是各种记忆新概念的一块试金石。例如，到 1992 年为止已发表的关于内隐记忆（implicit memory）是否存在加工水平效应的实验研究就有将近 40 个（Challis & Broadbent, 1992）。

从上述两个例子可以看到，某种实验范式实际上就是相对固定的实验程序，它的设计一般有两种用途或目的。第一，为了使某种心理现象得到更清晰准确的描述和表达。第二，为了检验某种假设、新提出来的概念。有些实验范式只局限于某一领域，有些实验范式经改动可以适用于许多领域。实验范式还随着研究的扩展与深入不断地涌现出来。例如，为了研究启动效应（priming effect），Tulving 等人于 1982 年设计了补笔（word fragment completion）这一实验范式，如今这一范式已广泛运用于内隐记忆的研究中。另一方面，有些实验范式又随着研究内容的变化而逐渐被淘汰。例如，记忆研究中对偶联合范式（paired—associate procedures）从 60 年代以后已逐渐停止使用。

实验范式的出现多带有一定的理论背景。例如， Craik 和 Lockhart (1972) 曾批评 Atkinson 和 Shiffrin (1968) 的如下主张：信息由短时记忆传递到长时记忆依赖于信息在短时记忆中维持的时间长短。Craik 等人认为，这种看法是片面的，短时记忆中复述的性质才是决定性的因素。机械复述（即死记硬背）对回忆帮助不大，只有涉及意义的复述才会导致良好的记忆。而加工水平范式正是上述思想的体现。因此，我们要学会把抽象的观念变成可以具体操作的实验，从实验范式中学习是极为重要的。

四、心理学规律的性质

R. S. 武德沃斯和 H. 施洛斯贝格(1965)曾指出,虽然实验心理学注重实验研究的定量工作,但有一些重要的变量,它们的性质是质量的而不是数量的。例如,“感觉道”的差异就是性质的不同,不同的感觉器官在感知周围环境方面各有不同的功能。R. S. 武德沃斯等人认为,过于偏向定量的工作,就会给研究工作设置许多障碍,从而把许多基本的科学问题都给掩盖起来了。

几十年以后,人工智能和认知心理学的创始人之一 Simon,从心理学规律的性质的角度讨论了心理学研究中的定性与定量的问题。他认为,心理学不应当向物理学看齐,企图用几个基本公式来概括所有的心理现象是不切实际的。实际上,其他科学领域的规律也不一定都是定量的规律。生物学、化学、分子生物学、地球物理学、甚至物理学的一些规律就属于定性结构的规律。例如,帕斯特的“疾病是由病菌造成的”这个规律并不是一个定量的规律,但它对人们寻找疾病产生的原因却起了指导作用,并推动了医学的发展,因此人们承认它是基本规律。又例如,现代分子遗传学大致源于如下的概括,即生物的遗传性状是由 4 种核苷酸排列成的长链决定的,这也属于定性结构的规律。心理学也有一些定量的规律,如短时记忆容量大约为 7。但心理学的规律现在和将来都是局限于一定范围内的,其普遍性是有限的,因而它主要是定性结构的规律。心理学的成功不是看它与物理学多么相像,而是看它描述和解释人的行为有多好。目前,心理学所提出的一些一般化的定理还不能给出精确的定性描述,还不属于定性结构的规律。例如,“人类在问题解决过程中受着短时记忆的限制,受到计算能力的限制”,这个理论并不那么确切,但它可以指导我们去分析问题,如它可以指导计算机的人-机界面设计(司马贺,1986;Simon, 1990)。

为什么心理学的发展不能以物理学为标准呢？这不是一个容易回答的问题。让我们举认知心理学为例来作一点说明。认知心理学作为心理学的主要潮流曾大大推动了心理学的前进，但是认知心理学仅强调用计算机程序来类比人的心理过程，即仅仅从软件层次上来进行类比，而不涉及硬件，不涉及产生心理过程的大脑。这表明它遵循的是信息论的思想，即只强调信息本身，而认为信息的载体是无关紧要的。这是一种只强调形式而不重视物质本身的倾向或方法，这种方法在物理学中已获得成功，它对物理学中的一系列基本问题都可以用简单而漂亮的形式系统来给予统一的描述和解释。但是，几十年来，这种方法已使人工智能不时陷入困境，它应用于心理学的局限性也愈来愈明显。在知觉、注意、表象、记忆等基本过程中常常出现对立的观点和流派而看不到解决的前景。也许，心理学应当从生物学中得到启发。生物学的知识依赖于对千百万种微生物、植物和动物的研究，每种生物生存的环境又千差万别。因此，仅仅在最抽象的和质的水平上才能谈论生物学的最一般的规律，许多生物学的规律都只适用于单一的物种。由于人的心理活动既受自然环境与社会环境的制约，又受人的大脑的约束，因而人的心理活动也就可以说是世界上最复杂的过程。也许，比较而言，心理学的规律更类似于生物学的规律而不似物理学的规律（陈霖、朱滢和陈永明，1996）。

在我们学习实验心理学、想要使心理学有一番作为时，应当好好思考心理学规律的性质。

问 题

1. 怎样理解实验心理学是各门心理学的基础？
2. 实验与观察有什么不同之处？
3. 什么叫做自变量的混淆？怎样避免自变量的混淆？
4. 举出几个实验范式的例子，并说明它们的应用，提出它们的理论设想。

5. 怎样理解心理学的规律更类似于生物学的规律而不似物理学的规律？

参 考 文 献

- 司马贺. 1986, 人类的认知. 科学出版社, 北京. 3~4
- 陈霖, 朱滢, 陈永明. 1996, 心理学和认知科学. 见: 21 世纪初科学发展趋势课题组著. 21 世纪初科学发展趋势. 科学出版社, 北京. 100~110
- 武德沃斯 R S, 施洛斯贝格 H. 1965, 实验心理学. 曹日昌等译. 科学出版社, 北京. 2~3
- 荆其诚. 1990, 现代心理学发展趋势. 人民出版社, 北京. 6
- 耿海燕, 朱滢. 1998, 关于无意识知觉的研究. 心理学动态, 6(1), 14~19
- Atkinson R C & Shiffrin R M. 1968, Human memory: A proposed system and its control processes. In: K W Spence & J T Spence (Eds.). The Psychology of Learning and Motivation: Advances in Research and Theory, Vol. 2. Academic Press, New York
- Craik F I M & Lockhart R S. 1972, Levels of processing: A framework for memory research. Journal of Verbal learning and Verbal Behavior, 11: 671~684
- Challis B & Broadbent D R. 1992, Level of processing affects priming in word fragment completion. JEP: Learning, Memory, and Cognition, 18: 595~607
- Kling J W & Riggs L A. 1975, Woodworth & Schlosbergs experimental psychology. Holt, Rinehart and Winston Inc, 1~9
- Macleod C M. 1991, Half a century of research on the stroop effect: an integrative review. psychological Bulletin, 109(2): 163~203
- Simon H A. 1974, How big is a chunk? Science, 183, 482~487
- Simon H A. 1990, Invariants of human behavior. Annu. Rev. Psychol, 41: 1~19

第二章 实验设计与数据统计分析

实验研究的目的是要检验理论假设的正确性或探索理论意图实现的可能性。所谓实验设计是指研究者针对需要验证的实验假设，为有计划地搜集观察资料而预先建立和依据的设计模式。

从理论上讲，心理学的实验设计是一门心理学与统计学相结合的实验技术科学，它隶属于心理科学方法论的范畴。广义地说，实验设计包括以下几方面：形成统计假设，并为检验假设、搜集和分析数据制定有效的计划；阐明检验统计假设所遵循的决策；按计划搜集资料；按计划分析资料；对统计假设的真伪作出归纳性推断。实验设计是研究工作的重点。良好的实验设计不仅是实验过程的依据和处理结果的先决条件，也是使科学研究获得预期结果的一个重要保证。

根据实验控制条件严密程度的不同，也就是根据对内在无效来源和外在无关因素的控制程度、能否随机选择和分配被试以及能否主动操纵实验变量，又可将实验设计分为真实验设计、准实验设计和非实验设计。

一、实验设计中的基本问题

研究者在对心理学研究中提出的问题进行实际实验时，往往希望能用较简单的实验设计来达到实验所需要的精确度。要想达到这一目标，研究者就必须考虑选择一个良好的实验设计，并要对设计中的各种变量进行控制，以及采取使误差减低到最低限度的手段——而这也正是我们在本节中所要讲述的问题。

（一）变量的选择与控制

1. 变量的选择

在科学研究中，研究者总是通过各种实验来探索原因和结果之间的因果关系。在心理学研究中，实验研究的目的在于探查行为（反应）与实验中所施加的刺激之间的关系或说明关系的性质。而这种行为（反应）与刺激的关系则正是用变量间的关系来表示和实施的。在研究中所讨论的变量包括自变量和因变量、调节变量、控制变量、中间变量、无关变量。这些内容已在第一章中讲述过，这里不再重述。

2. 实验设计中的控制

实验研究的设计与实验控制的应用有着紧密的联系。尽管良好设计的实验和其他形式的实验研究可追溯到几百年前，但由科学工作者用控制的方法来进行系统的实验研究也仅是近 100 年的事情。研究实验控制的问题包括以下几个方面：

（1）控制的概念

控制是指对实验条件的限制。在心理测量学的研究中，需要施行控制的研究是很多的。例如，在对各类考试、测验进行命题时，既要考虑控制试卷中各个题目的难度，也要考虑控制整个试卷的平均难度。在实验心理学的研究中，所实施控制的实例就更多了，这里就不一一列举了。

（2）控制的应用

①对变量的控制。在实验研究中，随机这一概念仅作为控制变量的技术是一个相当重要的发明。利用随机化原则对自变量进行控制包含两种含义：一是从限定的总体中随机抽取被试作为实验的样本；二是将抽取的被试随机分配到相应的实验处理中去。从限定的总体中随机抽取样本的方法有：简单随机抽样和分层随机抽样等方法

②使用控制组。在实验中,采用控制组对实验结果的评价是非常有利的。米尔(J. S. Mill)的研究进一步指出了“控制”这一词的含义,他的第一个原则是一致原则。此原则表明 如果 a 在 A 后面发生,那么 A 有可能是 a 的原因。对于“有可能”要认真考虑,因为显然有这样的情况,即使 a 总在 A 后面发生, A 也可能不是 a 的原因。在米尔的第二个原则中,他主张 如果 a 总是跟在 A 后面发生,并且如果 A 不存在,则 a 也不存在,那么就可以断定 A 是 a 发生的原因。这后一原则恰好是第一原则的拓广,它包括可组成 A 不存在时的情况,这恰恰是控制的引入,是控制组引入实验的基础,它描述了一种非常普遍的科学研究的方**法**。

3. 实验误差

在实验设计的研究中,研究者面临的一个重要问题就是要估计实验误差。没有对实验误差的估计,研究的结果就不易解释清楚,并可能产生错误的解释。所以有必要了解实验误差的定义以及它的主要来源。

实验误差是存在于实验单元内作同样处理所得观测数据间的变差的度量。从理论上来讲,实验误差有 3 个主要来源:首先是有**一个**内在的变差存在于要施加处理的实验被试中;其次是由于在实验的环境条件和操作过程中因缺乏一致性的结果而产生的变差;再次是重复实验的变差。林德奎斯特(E. F. Lindquist, 1953)提出了相类似的分类原则,只是他的分类原则更针对于具体的实验。他按误差是否与被试、组别、重复相联系而将其分为三类:即:S 类型误差(subjects),G-类型误差(groups)和 R-类型误差(replications)。研究者在进行实验设计时,应力图使上述三类误差的来源减少到最小值。为达此目的,可采取如下措施:

①调整实验中的被试,以减少其内在的变差效应。

②在可能的情况下,增加实验重复的次数,对实验结果进行综合评价。

(二)实验中的效度

在心理学研究中,常有一些无关因素与实验变量相混淆,影响实验的结果。这就是影响实验设计的内部效度和外部效度。

1. 影响内部效度的因素

实验研究的内部效度是指实验变量(处理)能被精确估计的程度。排除对实验结果产生干扰的无关因素,使研究者相信实验结果确实是由实验变量引起的,这时则说这个研究具有内部效度。反之,如果对与实验无关的因素控制不充分,一些无关的因素可能与实验变量产生混淆,使实验结果受到影响,则说这个实验研究缺乏内部效度。一般来说,下面的这些因素常会与实验变量相混淆,成为影响内部效度的因素。

(1)历史(经历)

历史指在实验过程中,与实验变量同时发生,并对实验结果产生影响的特定事件。当出现这种情况时,研究者往往无法判断实验结果是由处理(自变量)引起的,还是由特定事件引出的。

(2)成熟或自然发展的影响

指在实验过程中随着时间的延续,被试身心发生变化,如变得较为成熟,变得疲倦、对实验失去兴趣或饥渴等。

(3)选择

在实验过程中,由于没有采用随机化的方法来选择被试和分配被试,因而造成在实验处理前被试的组与组之间在很多方面并不相等或有偏向性。

(4)测验

在许多研究中,研究者为了取得实验前被试的初始状态,常对被试实施前测验,而这种前测验可能会积极或消极地影响实验处理实施后所进行的测验。

(5)被试的亡失