

自 序

心理学是当代社会与行为科学中发展最迅速的学科之一。关于心理学的探索研究，目前可说是正处在“著人滋味，真个浓如酒”的时期。在心理学科的百花园中，实验心理学的发展尤为突出。实验法是心理学研究的主要方法，它不仅促使心理学成为一门独立科学，而且其日新月异的演进，也推动着心理学各领域的迅猛发展。诚然，由美国心理学家斯珀灵（George Sperling）1960年首次公布的部分报告法，证实了感觉记忆的存在。此外诸如信号检测论的应用，使对内部心理的分析更上层楼，反应时新法使认知心理学大展宏图，间接测量法导致了内隐记忆的新发现。由此等实验心理学的发展，在在有力地说明了实验方法是揭露心理和行为的规律性的重要途径和手段。一位心理学家可以对心理学任一领域任一分支感兴趣，可以专门从事工业心理、医学心理、教育心理或知觉心理、记忆心理、思维心理，以至社会心理的研究，但是他们必定有一个共同的特点，即是确切地掌握了实验心理学的研究方法，了解应当如何科学地考察心理和行为的规律。这就足以证明实验心理学的重要性。

本人有幸受邀参加《世纪心理学丛书》中《实验心理学》一书的编写工作，获得了一次珍贵的学习锻炼机会。本书是专为大学和师范院校学生撰写的一本教科书，一般读过《普通心理学》和《心理统计学》者，都可读懂此书。本书采用教科书的形式，宗旨是将实验心理学中已经确立起来的最基本的、最可靠的科学方法介绍给同学和读者，使同学们初步掌握实验心理学的重要理论，学会基本的实验设计，并具有相当的实验技能。

本书的成功出版，首先要归功于张春兴教授和卓鑫森先生。台湾师范大学张春兴教授费五年多心力，邀集两岸心理学者合作撰写《世纪心理学丛书》，对未来中国心理科学的发展，意义极为重大。台湾东华书局负责人卓鑫森先生，鼎力资助由繁简两种字体在两岸发行此《世纪心理学丛书》，对两岸文化学术交流，贡献卓著。本书的编写过程中，又得到主编张春兴教授的关怀和悉心协助，以及东华书局编辑们在体例、图表乃至内容上的诸多帮助。没有这些鼓励和帮助是无法写出本书的。还要说明的是，本书在写作过程中参考了国内外有关心理学专著、论文和文献资料，吸取了许多学者的实验成果也引用了我本人过去编写的著作和与人合著的著作，在此一并向原作者致深深的谢意。

最后要说明的是，由于本书编者自身水平和时间的限制，疏漏和缺失在所难免，恳切地期望广大读者批评指正。

杨 治 良 谨识
一九九七年一月于华东师范大学

第一章

绪 论

本章内容细目

第一节 实验心理学的性质

- 一、实验心理学的含义 3
- 二、实验中的几种变量 4
- 三、实验中主试者与被试者的关系 8
 - (一) 按实验程序进行的主-被试者的相互作用
 - (二) 干扰实验程序的主-被试者的相互作用
- 四、额外变量的控制 11
- 五、实验效度 13
 - (一) 影响实验内部效度的因素
 - (二) 影响实验外部效度的因素

第二节 心理实验的程序

- 一、课题的确定 17
 - (一) 课题的来源
 - (二) 实验类型的确定
 - (三) 问题的陈述
- 二、被试者的选择 19
 - (一) 依问题的性质选择被试者
 - (二) 依研究结果的概括程度来选择被试者
- 三、实验因素的控制 20
 - (一) 自变量的控制
 - (二) 反应的控制
 - (三) 反应指标的选择
 - (四) 仪器和材料的使用
- 四、实验资料的收集与分析 and 综合 24
 - (一) 心理学研究的资料类别
 - (二) 研究资料的整理与解释
- 五、撰写实验报告 26

第三节 实验心理学简史

- 一、心理实验发展的第一阶段 0
- 二、心理实验发展的第二阶段 1

本章实验

- 一、自变量和因变量的确定 6

本章摘要

建议参考资料

实验心理学一词，是 1862 年由德国著名心理学家、实验心理学的先驱之一冯特（Wilhelm Wundt，1832～1920）在他的《感官知觉理论贡献》论文集的导言里最早提出来的。冯特提出“实验心理学”是作为他创建的新心理学（new psychology）的代名词。为什么会出现创建新心理学的愿望呢？这是因为心理学在成为一门独立的科学之前始终作为哲学的一个附属部分，是由哲学家们用思辨的方法来进行研究的。冯特所说的实验心理学是指用实验的方法来进行研究的心理学，亦即科学的新心理学，它是为了与以前的心理学相区别而被提出的。但这时他所说的实验心理学只是指用实验的方法来研究心理学中的感知觉问题，而对于思维、想象、情感等高级的心理过程，他认为不能用实验的方法来进行研究的，他称用非实验方法研究的这部分心理学为民族心理学（folk psychology）。尽管冯特的观点有一定的局限性，但在当时乃至今日仍不失为心理学中比较好的研究方法，为科学心理学的创建立下了功劳。我们把 1879 年冯特在德国莱比锡大学建立第一个研究心理学的实验室这一事件作为心理学从哲学中分化出来，成为一门独立学科的标志，并称他为实验心理学的创始人。

虽然心理学的实验方法都是在加以控制的条件下进行研究的，但由于不同的心理学流派所主张的心理学研究的对象不同，在控制条件下观察什么则不尽相同。例如，构造主义（或结构主义）（structuralism）者就主张用内省法（introspection）观察自己的直接经验，并企图把意识分为最简单、最基本的心理元素；行为主义（behaviorism）者则主张在控制条件下观察人和动物的行为；等等。

在 20 世纪 50 年代，随着电子学和工程技术的发展，心理学的实验装置和测量仪器日益精密，从而使心理学的实验研究工作在客观性和准确性方面都得到了提高，并逐渐摆脱了不同心理学流派对实验方法的束缚和影响。心理学发展到今天，争论的问题已不再是实验方法对心理学的研究是否适用，而是如何使心理学的实验方法更加完善、更加自动化，以及如何用在实验室中发现的心理学规律来解决实际问题。

第一章的绪论，是为了使读者在进入以后各章的专题讨论前，先对实验心理学有一个概括性的认识。因此，本章的撰写旨在回答以下六个问题：

1. 实验心理学是什么样的学科。什么叫实验。
2. 什么叫自变量？什么叫因变量？并举例说明。
3. 什么叫额外变量？对额外变量的控制通常采用哪些方法。
4. 心理学研究的课题主要来自哪些方面。
5. 一个完整的实验报告应包括哪几项内容。
6. 从实验心理学简史的学习中，获得了哪些启示。

第一节 实验心理学的性质

一、实验心理学的含义

实验心理学 (experimental psychology) 就是在实验控制条件下对心理和行为进行研究工作心理学。这一心理学分支是以研究方法来定义的。根据这个定义, 只要是用实验法 (experimental method) 来研究的心理学问题, 都可以包括在实验心理学的范围之内。不过, 实验心理学所讲的实验法一般是指在严密控制条件下的实验室实验。实验心理学的内容通常包括两个部分, 一部分是阐述实验方法和实验设计, 另一部分则是阐述实验法在一些专门领域中的应用, 如心理物理学、知觉、学习、思维、情绪等专题。前者论述心理学实验的一般原理, 后者论述在具体研究领域如何应用实验方法等问题。

实验法不等于自然观察法 (naturalistic observation)。在使用自然观察法时, 研究者只能被动地仔细观察和记录研究对象在自然状态下所发生的情况, 而不能有任何干预。虽然长期的、系统的观察也可以发现事物之间的规律, 例如天文学中的规律大部分是这样发现的。但这种观察只能等待所要观察的事物出现时才能进行, 受自然条件的限制。而实验 (experiment) 则是人为地去干预、控制所研究的对象。实验者可以创造条件, 引发所需要的事件来观察其变化; 为了验证, 可以创造同样的条件进行重复观察。与自然观察法相比较, 实验法是探寻事物之间的规律的一种更加有效的办法。

现代心理学是一个非常庞大的学科体系, 包含有许多心理学分支, 例如普通心理学 (general psychology)、生理心理学 (physiological psychology)、学习心理学 (psychology of learning)、认知心理学 (cognitive psychology)、发展心理学 (developmental psychology)、社会心理学 (social psychology)、人格心理学 (personality psychology)、工业心理学 (industrial psychology)、教育心理学 (educational psychology)、管理心理学 (managerial psychology) 等等。在现代心理学的各分支学科中几乎都在应用实验法, 并且应用实验法的范围正在日益扩大; 即使在过去曾被认为难以进行实验研究的人格心理学和社会心理学也都在进行实验研究。心理学家把那些用实验法所得到的科学资料概括为实验心理学的综合成果, 从而发展了实验心理学; 同时, 实验心理学也成了现代心理学各分支学科的发展基础。实验心理学在现代心理学中的地位及其与其他心理学分支学科的关系如图 1-1 所示。

二、实验中的几种变量

从以上叙述中, 我们可以知道实验是一种控制的观察。与自然的或偶然的观察不同, 实验是实验者人为地使现象发生, 对产生现象的情景或影响现象的条件加以操纵、变化以及控制的观察。然而, 自然的或偶然的观察是研究者在自然的情境中任现象自然的发生, 对现象及其发生的情境不加入人为的干预而进行的。与其他研究方法相比, 实验法的主要特点是能较好地控制额外变量。伍德沃斯和施洛斯贝格 (Woodworth & Schlosberg, 1954) 指出了实验具有三个优点: (1) 在实验的方法中, 实验者可以在他愿意时, 使事件产生, 可以充分地进行精密的观察。(2) 实验在同样条件下是可以重复的, 别人可以验证它。(3) 系统地变化条件, 可以追究与此相随的事件的变化。在这三个优点中, 第 2 个优点是特别重要的。只有按照同样的手续, 能被别

人重复验证的事实，才能构成科学的知识。当然，每个实验的复杂程度是不同的。在简单、单一的实验里，只让一个条件变化而控制其余的一切条件。在妥善安排的复杂设计里，可以让两个或两个以上的条件同时变化，分析每个条件的单独影响以及它们之间可能的相互作用。在我们明确了“实验”的含义后，我们就可以来讨论什么是变量？变量主要有几种？

变量（或变项）（variable）是指在数量上或质量上可变的事物的属性。例如，光的强度可以由弱变强，呈现时间可以由短变长，智力的 IQ 可以由小到大，这些都属于量的变量。又如，人的性别有男有女，人的宗教信仰有佛教、道教、基督教、天主教和伊斯兰教等等，这些是质的变量。质的变量有时也可以用数字代替类别，以便于统计分析。

在自然科学中，常用数学方程式来描述一些现象。用实验法研究心理学问题时，也用数学方程式来探明变量与变量的关系。在实验中实验者所操纵的、对被试者的反应产生影响的变量称为自变量（或自变项）（independent variable，简称 I.V.）；由操纵自变量而引起的被试者的某种特定反应称为因变量（或依变项）（dependent variable，简称 D.V.）。例如，要研究灯光亮度对阅读速度的影响时，实验者所操纵的灯光亮度（变化烛光数）就是自变量，而阅读速度（以每分钟多少字来表示）就是因变量。因此，自变量和因变量是相互依存的，没有自变量就无所谓因变量，没有因变量也无所谓自变量。

除了自变量之外，还有其他许多因素都会影响因变量的变化。上例中，文字的形状、大小、排列方式、清晰度、熟悉度以及材料的呈现位置和呈现先后等都会影响阅读速度。此外，像大气压、被试者的高矮、胖瘦和头发长短等因素对阅读速度就没有什么影响。凡是对因变量产生影响的实验条件都称为相关变量（relevant variable，简称 R.V.），而对因变量不产生影响的实验条件称为无关变量（或无关变项）（irrelevant variable，简称 I.V.）。在相关变量中，实验者用以研究的变量称为自变量，实验者不用于研究的那些相关变量称为额外相关变量（extraneous relevant variable），或简称为额外变量（或外扰变项）（extraneous variable）。在实验中，额外变量是必须加以控制的。如果不控制额外变量，就会弄不清因变量的变化是由自变量的影响引起的，还是由因变量的变化引起的。因而就无法得出明确的结论。由于在实验中额外变量是必须加以控制的，所以额外变量也被称为控制变量（controlled variable）。评价一项实验设计的好坏的一个重要依据就是看研究者能否成功地控制那些额外变量。

自变量的种类很多，大致可以分为三类：

1. 作业 作业（task）是指实验中要求被试者作出特定反应的某种呈现刺激，例如，字母串、配对联想词表、Müller-Lyer 错觉图、字谜等等。如果把这些作业的任何特性作为自变量来操纵，则这种自变量即为一种作业变量（task variable）。

2. 环境 当实验呈现某种作业时，如果改变实验环境（environment）的任何特性，则改变了的环境特性即为环境自变量（environmental independent variable）。例如，我们可以改变实验室内的亮度或噪声（噪音）强度，也可以改变呈现刺激的时间间隔等等。

3. 被试者 被试者的特性因素如年龄、性别、健康状况、智力、教育水平、人格特性、动机、态度、内驱力等都可能影响对某种刺激的反应，这些因素

统称为被试者变量（或受试者变项）（subject variable）。在这些被试者变量中，有的是实验者可以主动操纵加以改变的，例如内驱力强度可以用禁食或禁饮的时间来加以操纵，而有的则是不能主动操纵的，只能进行测量，例如智力、教育水平、自我强度等。被试者本身固有的、实验者不能加以操纵使其改变的特性称为本性变量（或属性变项）（attribute variable）。

在某些情况下，研究者把几个不同的自变量当作一个复合自变量（complex independent variable）来操纵，以确定它们的综合效应。例如，有些学校为提高学生的学习成绩进行实验，采取了一系列的教学改革措施，如使用新的教材、加强课堂管理、奖励成就等。假定这些措施确实提高了学生的学习成绩，那我们是不大可能鉴别出哪一个自变量在起什么作用的，或许其中只有一种措施才真正起作用，但设计这种实验的目的并非要鉴别出某个变量的作用，而是考察其综合效应。因此，只要研究者不对复合自变量作出分析性结论，这类研究无可厚非。由于复合自变量更接近日常生活实际，并能解决某些实际问题，因而常被研究者所采用。

因为自变量的变化而产生的现象变化、或结果，称为因变量，也就是反应变量（response variable）、或反应测量（response measurement）。对于被试者的反应可以从下列几方面来测量：（1）反应速度，例如，简单反应时间、或潜伏期，走完一个迷津所需的时间，在一定时间内完成某项作业的数量等。（2）反应的正确性，例如，选择反应正确的次数、走迷宫入盲巷的次数、跟追踪盘离靶的次数或距离等。（3）反应的难度，有些作业可以定出一个难易等级，看被试者或动物能达到什么水平，如斯金纳箱就有三个难度等级。（4）反应的次数或机率，是指在一定时间内被试者能做出某种反应的次数，例如在心理物理学实验中，根据机率来规定阈限。（5）反应的强度，例如，情绪实验中的出汗量，皮肤电反射电阻变化的大小，等等。

除了上述反应指标外，被试者的口语报告内容（即口语记录）也是一项重要的反应变量。口语记录（protocol）是指被试者在实验时对自己心理活动进程所作叙述的记录，或在实验之后，他对主试者提出问题所作回答的记录。在心理学实验中，口语记录是很重要的参考资料，有助于我们分析被试者的内部心理活动。目前口语记录分析已被用于问题解决的实验中，借以了解被试者解决问题时所使用的思维策略。

对于人的心理活动，我们无法直接观察到。但他的行为反应是受心理活动支配的，是我们可以观察到的。心理学家之所以对反应指标的测量感到兴趣，是因为他们相信从由刺激引起的反应指标可以推知某些更基本的心理过程。例如我们让被试者解一道数学题： $6+7+8+9=?$ 要求他们在解题时说出头脑中所进行的一切活动，并记录其得出的答案和使用了多少时间。结果发现，虽然被试者的答案都是对的，但解题所用的时间却有显著差异，有的解得快；有的解得慢。进一步分析他们的口语记录，发现解得慢的被试者是用累加法，解得快的被试者是用另一种策略，先做 $7+8=15$ ， $6+9=15$ ，再做 $15+15=30$ 。这就是说，在这个实验中，一个变量（一道数学题）导致另一个变量（解题所用的时间多少）的变化是由某个中介变量（intervening variable）（解题所用的策略）所引起的。现代实验心理学的目的就是要说明和解释人在完成某种活动时的心理活动是如何进行的，即通过刺激和反应（即 S—R）之间的关系来推断心理活动的方式。

然而，在没有控制的条件下，一个刺激可能会引起多种心理和行为的反

应；相反，一种反应也可能是由多种刺激复合作用的结果。因此，如果要建立 S—R 的对应关系，就必须发展一系列的有效措施，进行严密的实验设计。这是心理学实验不同于物理、化学等实验的第一个特点。

三、实验中主试者与被试者的关系

心理实验的对象统称为被试者 (subjects)，包括人和动物。在实验中，特别是以人为对象的实验中，如何处理好实验者（即主试者）和被试者的关系是实验取得成功的一个重要条件。在心理实验中主试者 (experimenter) 和被试者的关系如图 1-2 所示，包括两类性质不同的相互作用：(1) 按实验程序进行的相互作用，(2) 干扰实验程序的相互作用。

(一) 按实验程序进行的主-被试者的相互作用

在心理学实验中涉及到的人就是主试者和被试者。主试者对被试者的干预及被试者对主试者的实验态度都会对实验结果产生影响。由于心理实验都是通过被试者完成任务的方式进行的，所以主试者对被试者最直接的干预是向被试者交代任务。主试者为交代任务向被试者所讲的话，在心理实验中就称为指示语 (instruction)。

在以人作被试者时，指示语在实验中不仅是对被试者说明实验，更重要的是给被试者设定课题，这也是控制被试者这一有机体变量的一种手段。指示语不同，所得的结果也不相同。因此，主试者在给出指示语时应注意以下几点：

1. 要严格确定给被试者什么样的指示语 不同的实验会有不同的要求，有的要求被试者尽量做得准确，有的要求尽量做得快，还有的要求又准又快。此外，是让被试者按特殊的方式完成某任务，还是让他随使用什么方式去完成任务。类似这样的问题，主试者都要事先确定，写到指示语中去。

2. 在指示语中，要把被试者应当知道的事交代完全 主试者要求被试者所做的事，可能是他从未做过的，要说明将要给他呈现什么，要他怎样做等等。

3. 要保证被试者确实懂得了指示语 指示语要写得简单明确。切忌模棱两可，也不要专门术语。为防止被试者误解指示语，可以让被试者用自己的话重述让他做什么、怎么做等等。

4. 指示语要标准化 事先要把指示语写下来念给被试者听。应做到所给的指示语前后一致。不要任意改变同一指示语中的有关词句。对有些实验最好能使用录音机给出指示语。

在指示语不能充分控制反应时，就要很好地考虑刺激条件和实验装置，使刺激条件、实验装置与指示语配合起来，使被试者只能做出主试者所要求的反应。

(二) 干扰实验程序的主-被试者的相互作用

在心理实验中，除了主试者给出指示语及被试者按指示语完成任务的相互作用外，他们之间还可能存在着某种干扰实验、使实验结果发生混淆的相互作用。例如，主试者在实验中可能以某种方式（如他的表情、手势、语气等）有意无意地影响被试者，使他们的反应附和实验者的期望。这种现象称为实验者效应 (experimenter effect)。实验者效应往往会以一种颇为微妙的方式在起作用。例如，当主试者了解到他的少数几位被试者有某种反应倾向时，他随后观察到的被试者的资料也往往有某种反应倾向。这里，除主试者无意识地以某种方式影响被试者的反应外，还可能故意地对被试者暗示、

提醒或鼓动，或不能耐心地等待被试者的真实反应的出现就进行记录，或仅仅记录自己所期望的行为反应。这种现象甚至在用动物做实验时也可能出现。在罗森塔尔 (Rosenthal, 1966) 的一个研究中，他让选修心理实验课的学生做白鼠走迷津实验，一组主试者用来做实验的白鼠的笼子上贴有“走迷津伶俐”的标签，另一组主试者用来做实验的鼠笼子上贴有“走迷津呆笨”的标签。其实，这些白鼠是随机地被分到这些笼子里的。结果，标记为“走迷津伶俐”的白鼠比标记为“走迷津呆笨”的白鼠学习得快些。实验者的期望莫名其妙地影响了白鼠的行为，动物的行为居然附和于错误的标签。

人类被试者参加实验并不是消极被动的。他们总是以某种动机、态度来对待实验的。因而实验结果在很大程度上也依存于被试者对待实验的态度，我们必须高度警惕实验中的霍桑效应和安慰剂效应 (placebo effect)。所谓霍桑效应 (Hawthorne effect) 是指 1924 年美国芝加哥西部电力公司霍桑工厂所进行的一项实验。研究者系统地改变照明强度，目的是确定工厂的最佳照明条件，借以提高工作效率。结果，不论照明增加或是减少，工人的工作效率都在逐渐提高。后来才发现，这是由于工人们觉得参加实验是厂里在关心他们，从而提高了工作效率。这里，很明显被试者的态度成了自变量，从而影响了反应变量。与霍桑效应相类似的是安慰剂效应。有时，医生开给病人的“药物”实际上并非是药物 (如维生素片)，但当病人相信那是有效的药物，服用后也产生效果。这是病人心理作用的结果。可以这样说，凡是在以人作被试者的任何研究中都可能出现霍桑效应或安慰剂效应现象。被试者对指示语的理解、参与实验的动机、焦虑水平、有关经验、以及当时的心理、生理状态等，都会影响他们完成任务的质量和数量。而被试者的反应成绩又会影响主试者的行为。这种相互作用有的是不知不觉地在进行，主试者往往没有察觉到。因此，主试者给予被试者的某种处理，所获得的不一定就完全是此处理所引起的反应，自变量也不一定只是主试者加以操纵的那个自变量。

总之，在心理学实验中主试者和被试者 (人类被试者) 都是具有主观能动性的。主试者用指示语规定被试者的反应，试图控制额外变量，使因变量的变化成为自变量的一种效应；但与此同时，主试者与被试者又可能以某些干扰实验的方式不知不觉地相互作用，使额外变量成了实验中的自变量，从而混淆了实验结果。这是心理学实验不同于物理、化学实验的另一个特点。

四、额外变量的控制

额外变量是使实验结果发生混淆的主要根源。要提高研究的科学水平，就要采取一定的方法来控制额外变量。对额外变量的控制，通常采用以下几种方法：

1. 排除法 排除法 (elimination method) 是把额外变量从实验中排除出去。如果外界的噪音和光线影响实验，最好的办法是进入隔音室或暗室，这样可把它们排除掉。霍桑效应和实验者效应会影响实验结果，最佳的办法是采用双盲实验 (double blind experiment)。从控制变量的观点来看，排除法确实有效。但用排除法所得到的研究结果却缺乏推论的普遍性。例如，如果顾虑主试者与被试者的彼此接触会影响实验结果，而采用自动呈现刺激及自动记录实验结果的方法，则所得结果便不能对人们日常生活中的同类行为作出推论和解释。

2. 恒定法 恒定法 (constant method) 就是使额外变量在实验的过程中保

持固定不变。如果消除额外变量有困难，就可以采用恒定法。不同的实验场所、不同的实验者、不同的实验时间都是额外变量。有效的控制方法是在同一实验室、由同一实验者、在同一个时间对实验组和控制组使用同样的实验程序进行实验。如果实验时强度变化的噪音无法消除，则可以用噪音发生器发生恒定的噪音来加以掩蔽。除上述实验条件保持恒定外，实验者和控制组被试者的特性（如年龄、性别、自我强度、成就、动机等）也是实验结果发生混淆的主要根源，也应保持恒定。只有这样，两个组在作业上的差异才可归于自变量的效果。用恒定法控制额外变量也有缺点：（1）实验结果不能推广到额外变量的其他水平上去。例如，如果只用男性成人作为被试者进行实验，其结果不能推广到女性成人。（2）操纵的自变量和保持恒定的额外变量可能产生交互作用。例如，如果被试者是男性，实验者是富有魅力的女性，实验时，实验者可能使被试者分心。这是交互作用产生的额外变量。

3. 匹配法匹配法（matching method）是使实验组和控制组中的被试者的特点相等的一种方法。使用匹配法时，先要测量所有被试者和实验中要完成的作业具有高相关的特点；然后根据测得的结果把实验组和控制组的被试者的特点匹配成相等的。若要做“练习对学习射击成功影响”的实验，先预测一下被试者打靶的成绩，然后把两个预测成绩相等（击中环数相等）的被试者分别分到实验组和控制组，匹配成条件相等的两组被试者参加实验。这种方法在理论上虽然可取，但在实际上很难行得通。因为，如果超过一个特性（或因素）以上时，实验者常感到顾此失彼，甚至无法匹配。例如，实验者要同时考虑年龄、性别、起始成绩、智力等因素，力图使所有因素均匹配成相等而编为两组就很困难了。即使能解决此困难，也将使很多被试者不能参加这个实验。更何况，属于中介变量的诸因素，如动机、态度等，更是无法找到可靠的依据进行匹配。因此，匹配法在实际上并不常用。

4. 随机化法随机化（randomization）是根据概率理论，把被试者随机地分派到各处理组中。从界定的被试者总体中用抽签法或随机数字法抽取被试者样本，由于随机取样使总体中的各个成员有同等机会被抽取，因而有相当大的可能性使样本保持与总体有相同的结构。随机取样后，再把随机抽出的被试者样本随机地分到各种处理中去。例如，有三种处理组：实验一组、实验二组、控制组。给每一处理组指定一个数字，如 0、1、2，并定好先分给样本 A，再分给样本 B、C。如果在随机表上遇到“2”，就把样本组 A 定为控制组；再遇到“0”，就把样本组 B 定为实验一，依此类推。从理论上讲，随机法是控制额外变量的最佳方法，因为根据概率理论，各组被试者所具备的各种条件和机会是均等的，不会导致系统性偏差。它不仅能克服匹配法顾此失彼的缺点，还能控制难以观察的中介变量（如动机、情感、疲劳、注意等）。随机法不仅能应用于被试者，也能应用于呈现刺激的安排。例如，如果有许多处理施加于被试者，为了消除系列效应（即前面的处理对后面的处理的影响），可以用随机法安排各种处理出现的顺序。

5. 抵消平衡法抵消平衡法（counterbalancing method）是通过采用某些综合平衡的方式使额外变量的效果互相抵消以达到控制额外变量的目的的方法。这种方法的主要作用是控制序列效应（sequence effect）。如果给被试者施加一系列以固定顺序出现的不同处理，被试者的反应将会受到时序先后的影响。如果先后两种处理在性质上无关，就会产生疲劳的影响。这两种影响都可以使实验发生混淆，因而要加以抵消。如果只有 A、B 两种处理，最常

用的抵消序列效应的方法是用 ABBA 的安排。即对同一组被试者先给予 A 处理，再给予 B 处理；然后倒过来，先给予 B 处理，再给予 A 处理。如果对几组被试者给予两种以上的处理，为了抵消序列效应则可采用拉丁方实验 (Latin square experiment)。

6. 统计控制法上述各种方法都是在实验设计时可以采用的。这些方法统称为实验控制 (experimental control)。但有时候由于条件限制，上述的各种方法不能使用，明知有因素将会影响实验结果，却无法在实验中加以排除或控制。在这种情形下，只有做完实验后采用协方差分析 (或共变数分析) (analysis of covariance)，把影响结果的因素分析出来，以达到对额外变量的控制。这种事后用统计技术来达到控制额外变量的方法，称为统计控制 (statistical control)。例如，在对两班学生进行实验以比较两种教学方法的好坏时，虽然实验者事先知道此两班学生的智力不等，但限于条件，实验前却无法对智力因素加以控制使两班学生的智力水平相当。显然智力是影响实验结果的重要因素。实验后，使用协方差分析将智力因素所产生的影响排除后，就可以比较两种教学方法的优劣了。统计控制法除协方差分析外，还可用偏相关等方法。

五、实验效度

实验效度 (experimental validity) 是指实验方法能达到实验目的的程度。实验目的是验证假设，验证自变量和因变量之间的关系，使实验结果的推论可用以解释和预测其他同类现象。由于不同的实验者在设计上和在对额外变量的控制程度上极不相同，实验的效度也会有很大的不同。此外，每种实验都有几个不同的组成部分，其中每一部分也会影响整个实验的效度。了解影响实验效度的诸因素，将有助于我们评价实验设计的质量，提高实验设计的科学性。实验效度主要包括内部效度和外部效度。下面让我们先讨论影响实验内部效度 (internal validity) 的诸因素，然后再讨论影响实验外部效度 (external validity) 的诸因素。

(一) 影响实验内部效度的因素

实验的内部效度是指实验中的自变量与因变量之间的因果关系的明确程度。一项实验的内部效度高，就意味着因变量的变化确系由特定的自变量引起的。由于除了自变量以外，任何额外变量都可能对因变量产生影响，导致实验结果的混淆。这样我们就难以判定实验中自变量与因变量之间的关系的确定性。因此，要使实验具有较高的内部效度，就必须控制各种额外变量。在设计实验时，如果能考虑到以下六个方面的因素，将有助于提高实验的内部效度。

1. 生长和成熟 除了实验中的自变量可能使个体行为发生变化外，个体本身的生长和成熟也是使其行为变化的重要因素。特别是在以幼小的儿童为被试者而又采用单组前测后测实验的情况下，生长和成熟因素的影响就更大。单组前测后测实验通常是实验处理之前先对被试者的某种行为作一次测量，实验处理后再以同样方法测量一次，两次测量之差即表示实验变量 (即自变量) 产生的效果。很明显，这种设计忽略了前后两次测量之间被试者的生长和成熟因素，其实验效果易受生长和成熟因素的混淆，从而降低了内部效度。解决的主要办法是增设同样条件的控制组进行比较。

2. 前测的影响 在一般正常情况下，前后两次测量的结果会有一些的差异，后测的分数将比前测的高。这中间包括练习因素、临场经验、以及对实

验目的的敏感程度，从而提高了后测的成绩。特别是前后两次测量时间较近，这一因素的影响就更显著。

3. 被试者的选择偏性 在对被试者进行分组时，如果没有用随机取样和随机分配的方法，在实验处理之前，他们在各方面并不相等或有偏性，从而造成实验结果的混淆，降低了内部效度。

4. 被试者的缺失 如果是一项长期的实验，要保持原实验被试者的人数不变是相当困难的。即使开始参加实验的被试者样本是经过随机取样和随机分配的，但由于被试者的中途缺失，常常使缺失后的被试者样本难以代表原来的样本。这就降低了内部效度。

5. 实验程序的不一致 在实验过程中，实验仪器、控制方式的不一致，测量程度的变化，实验处理的扩散和交流等都可能混淆实验变量（即自变量）的效果。实验者知道实验目的所产生的“实验者效应”以及被试者知道实验目的或其自己正被研究所产生的“霍桑效应”和“安慰剂效应”等都将混淆实验变量（即自变量）的效果，从而降低了内部效度。

6. 统计回归 将实验结果进行统计回归（statistical regression）后，可能使实验者对实验变量（即自变量）的效果产生误解。如果选择具有极端特性的个体作为被试者（如高焦虑组和低焦虑组）进行实验时，尤需注意这种现象。统计回归现象是，第一次测量平均值偏高者，第二次测量平均值有趋低的倾向（向常态分布的平均数回归）；第一次测量平均值偏低者，第二次测量平均值有趋高的倾向（也向常态分布的平均数回归）。因此第二次测量虽在实验处理之后，其升高或降低只是受统计回归的影响，可能并非是实验变量（即自变量）所产生的效果。

（二）影响实验外部效度的因素

实验的外部效度是指实验结果能够普遍推论到样本的总体和其他同类现象中去程度，即实验结果的普遍代表性和适用性。以人的行为为对象所获得的实验结果，其推论法往往有相当的局限性。实验的外部效度主要受下列三方面的影响：

1. 实验环境的人为性 实验是在控制条件下进行的，实验环境的人为性可能使某些实验结果难以用来解释日常生活中的行为现象。实验室中的仪器设备会影响被试者的典型行为。被试者参与实验的动机也会影响其行为表现。而在实验室之外的日常生活中，就不会有这些因素的影响。因此，实验结果还不能完全等同于实验室之外的日常行为现象。

2. 被试者样本缺乏代表性 从理论上讲，从事于实验的被试者必须具有代表性、必须从将来预期推论、解释同类行为现象的总体中进行随机取样。但实际上这是很难做到的。因为，如果总体很大，即使能够随机取样，但心理学实验的被试者通常是自愿的，所以也很难把被随机选上的人全都请来实验。如果总体是无限的（例如，“七岁儿童”就是一个无限的总体，其包括过去的、现在的、将来的所有七岁儿童），随机取样实际上是行不通的。这样的实验结果自然会降低其外部效度。

3. 测量工具的局限性 实验者对实验变量（即自变量）和反应变量的操作性定义往往以所使用的测量工具的测量结果来加以考虑的。例如，把成就动机作为一个因变量，实验者常以某种成就动机量表所测得的分数来界定并评定其强度。但成就动机的测量工具有各种不同的形式，所测量出的分数并不代表同一种成就动机及其强度。如果在实验时采用的是某一种成就动机的

量表，那么所得出的实验结果便不能推论到采用其他成就动机的量表的情况中去。

实验的内部效度和外部效度是相互联系、相互影响的。提高实验内部效度的措施可能会降低其外部效度，而提高实验外部效度的措施又可能会降低其内部效度。这两种效度的相对重要性，主要取决于实验的目的和实验的要求。一般而言，在实验中控制额外变量的程度越大，则对因果关系的测量就越有效。因此，可以在保证实验内部效度的前提下，采取适当措施以提高外部效度。

第二节 心理实验的程序

实验的程序 (experimental procedure) 就是实验的进程。它是指实验在各个阶段应做的事。怎样合理地安排实验程序是心理学实验的重要内容。心理学实验研究是一种创造性活动,它没有一成不变的法则,不能把科学研究看成是科学家简单地在奉行某种常规活动。尽管如此,从大量的心理学文献中,我们仍然能够分析出心理学实验通常要遵循的基本程序。这基本程序是:(1)课题的确定;(2)被试者的选择;(3)实验因素的控制;(4)实验资料的收集、分析和综合;(5)撰写实验报告。

本节将概要地阐明上述五个步骤中所遇到的问题和各种可能的选择。

一、课题的确定

(一) 课题的来源

一项科学研究总是从发现问题开始的。选择课题是科学研究的第一步。学生和有经验的研究者一样,可以从不同的来源得到启发,提出研究课题。那么,课题从何而来呢?研究课题的来源通常有以下四个方面:

1. 实际需要 实际的工作中存在着许多问题需要实验研究来解决。例如,从事数学教学改革时,发现小学生学四则运算题有困难,就提出了小学生能不能学代数?小学生学代数是否比学四则题容易?学四则题对儿童的思维发展有什么影响?诸如此类的实际问题,都可以通过实验研究获得明确的回答。

2. 理论需要 从理论或学说中推演出的某个假设是否符合实际,这就需要实验来检验。例如,学习以后不再复习,遗忘量通常随时间的增加而增加。为了说明遗忘的原因曾提出过一种干扰说,根据这个理论来推论,在学习以后,回忆以前,如其他条件相同,则插入学习的材料越多,对原来学习材料的记忆影响就越大。根据这个推论就可设计《学习额外材料的数量对原来材料记忆影响的实验研究》等课题。

3. 个人经验 在学习、工作和日常生活中,经常会遇到一些心理学问题,例如,为什么“人芝兰之室,久而不闻其香”?50赫兹的电灯光为什么看不见其闪烁?高压水银灯下的红花为什么看上去是黑色的?为什么人能看到秒针走,却看不出分针在移动?“江山易改,禀性难移”的说法对吗?千千万万的人有千千万万的心理现象,所以人的心理现象就像万花筒一样,纷繁复杂而又绚丽多彩。针对这些实际问题就能设计出种种实验。

4. 前人的研究与文献资料 实验研究虽然不一定要从文献出发,但是,在确立课题前,系统地查阅有关文献是必要的。阅读文献可以发现什么问题已经解决、什么问题尚待研究。例如,心理学家(许淑莲、杨治良等,1984)看到国内三例有关先天性无痛症的报导,于是就检索国内外有关研究和文献,并未发现心理学上有这项研究,这样,想到了从心理学的角度研究无痛儿。其中有一位女孩,当时她只有十岁,是个三年级的学生。她看上去模样端正、活泼可爱。对她进行智力测验,结果属于正常偏低。对她情绪特点进行调查,发现她对一般恐惧的反应与正常的儿童一样,如怕看电影里的紧张恐怖镜头,看见小狗、小毛虫也会害怕。但对与身体损伤有关的恐惧与正常儿童不同。这些研究进一步揭示了先天性无痛儿的心理特点。

(二) 实验类型的确定

课题虽然有不同来源，但是，它们都是从提出问题开始的。对于“为什么”科学的探索，大致可分为两阶段或两个类型。第一阶段是探明规定某个行为的条件，第二个阶段是探明哪些条件与行为之间的函数关系。与这两个阶段相对应，可以把实验分为两种类型。

第一种类型是因素型实验（factorial type experiment），即探求规定行为的条件“是什么”的“什么型实验”，或是探明行为的规定要因的实验。在因素型实验里，逐个地除去、破坏或变化被看作是行为规定要因的几个条件，根据有无相应的行为变化，探明它是否是行为的规定要因。毫无疑问，这时候，对于被操作的条件之外的条件，都应当进行严密的控制。武德沃斯和施洛斯贝格根据他们的分类方法，把因素型实验看作是定性实验。

第二种类型是函数型实验（functional type experiment），即探求各种条件是“怎样”规定行为的“怎样型实验”，或是探明条件和行为之间的函数关系的实验。在函数型实验里，根据因素型实验的结果，系统地、分阶段地变化规定要因的条件，以进行确定条件和行为之间的函数关系的函数型实验，以找出行为的法则。武德沃斯和施洛斯贝格分类中的定量实验就相当于函数型实验。

如果将因素型实验和函数型实验做一比较，可以认为，因素型实验是函数型实验的前一阶段，具有函数型实验的预备实验的性质。当然，在不少实际的研究中，是将因素型实验和函数型实验作为一个实验来进行研究的。一般地说，根据以前的研究，在规定要因已被探明时，多半是直接进行函数型实验，这也是许多实验的基本过程。

由此可见，实验在提出课题的同时，还应明确所探求的问题是属于这两个阶段中的哪个阶段，即是属于因素型实验呢，还是属于函数型实验。

（三）问题的陈述

在课题确定及其所属的实验类型均明确之后，若能以假设的形式提出，那就更符合科学原则。假设是关于条件和行为之间的关系陈述。如果把对条件的叙述记为 a，把对行为的叙述记为 b，一般取“如果 a，那么 b”这样的形式。一切科学定律、法则虽然表面上不一定都符合这个形式，但实际上却包含先行条件（自变量）和后继条件（因变量）这样的逻辑关系。例如，对于缪勒-莱尔错觉（Müller-Lyer illusion）来说，“若变化夹角的大小，则视错觉就有变化”这一假设，是确定夹角是否是视错觉的要因这一因素型实验的假设。这是假设陈述的第一种方式。

假设的另一种陈述方式是用函数关系来表示。它用方程式 $b=f(a)$ 来表示自变量 a 与因变量 b 共变的函数关系，这个方程式读作 b 为 a 的函数，或 b 数量地依存于 a。应用这个模型，就可将上述关系改为：“视错觉的量与夹角的余弦成正比”。这就是函数型的假设。

二、被试者的选择

心理学实验的被试者是多种多样的，有正常人，也有疾病患者，有各种年龄段的人，也有动物，如鼠、狗、猴等。不同的实验对象常常要求不同的实验操作。因此，在制定研究计划之前，应确定研究中所用的被试者。涉及被试者选择的问题有：（1）是使用人类被试者还是使用非人类被试者？（2）被试者应具备哪些机体特征？（3）用哪一种取样方法才能使被试者样本代表总体？这些问题的解答主要取决于二个因素：即课题的性质及研究结果的概括程度。

（一）依问题的性质选择被试者

选用人类被试者还是非人类被试者，依据课题的性质而定。许多心理学研究选用人类被试者，因为它关心的是人类的心理和行为。有损被试者身心健康的实验就不能选用人类被试者，而应当考虑选用适当的非人类被试者。例如，拥挤度对本能行为的影响实验，这样的实验就不能选用人类被试者，而只能选用动物被试者。有的心理学实验使用非人类被试者是因为考虑到要严密地控制无关变量。例如，用同胎的幼小个体或限制它们的活动以削弱个体差异，然后进行实验研究。在人类被试者身上就不可能做到这样的控制，也有不少研究是先进行动物实验，然后在此基础上再进行人类实验。

（二）依研究结果的概括程度来选择被试者

在从事一项研究时必须依据研究结果的概括程度来选择被试者。心理学研究的群体可能是一个小群体，或仅具有某种特性的成员。例如，所有的生物、世上所有的人、某个国家的所有的人，某个民族的所有的人、所有的大学生、某个年龄组的所有成员、某个学校的学生……都可以作为心理学研究的总体。因此，选定什么样的被试者样本，要依研究的问题和据此而推论的全体而定。如果被试者的选择出现偏差，就会影响实验效度，用什么方法能减少这种偏差呢？简言之，可用如下二种方法：（1）随机抽样法。这是最基本的方法，实验用的被试者是随机抽选出来的。每个个体从总体中被选中的机会是均等的，任何个体的选择与其他个体的选择没有牵连，彼此之间的选择都是独立的。（2）分层随机取样法。当总体有不同大小的小组和层次组成时，分层方法最适用。例如，研究中国成人的听力与外国成人的听力是否有所差异，就必须用一定数量的被试者，他们最好是来自成人中不同的年龄段、不同的性别、不同的职业、来自国内不同的地区、甚至不同的民族。这样得到的结果才能代表中国成人的听力。

三、实验因素的控制

在实验过程中，对自变量、因变量和控制变量都要进行控制。有关控制变量的控制，已在第一节中作了讨论，这里讨论对自变量和因变量的控制。

（一）自变量的控制

在实验中对自变量的操纵、变化称为自变量的控制。对自变量控制的好坏，直接影响实验的成败。

对自变量的控制，首先要对自变量进行严格的规定，对心理学中一些含混不清的变量必须使之操作定义化，只有这样才能进行实验。那么什么叫操作定义呢？操作定义（operational definition）是由美国物理学家布里奇曼（Bridgman, 1972）提出的，他主张一个概念应由测定它的程序来下定义。操作定义的提出受到心理学界的欢迎。在心理学上，对一个心理现象根据测定它们的程序下定义就叫操作定义。例如，把“刚刚感受到”定义为“50%次感受到”，就可测定感觉阈限了。又如，疲倦（fatigue）没有一个共同的起点和尺度，怎么测量呢？如果定义为“工作效率的下降”，那么就可以进行测量和比较了。实验者可据此操纵这个变量了。因此，对一些含混不清的变量，一定要有操作定义。

其次，对于在刺激维度上连续变化的自变量，要做好三项工作：（1）要选一定数量的检查点，以找出自变量和因变量的函数关系。如果两者是线性关系，一般三至五个点就可以了。如果函数关系比较复杂，则至少要选五个检查点。（2）要确定好自变量的范围，对自变量范围的确定，有时前人的研

究可以提供线索，如在暗适应的研究中，一般过程在 0~60 分钟范围内。再如在两点阈的研究中，别人已对人体各部位的阈值做过测定，可作借鉴。若在无前人的经验时，就要通过预备实验来确定。(3) 要确定好各检查点之间的间距。间距的大小虽然和自变量的范围和检查点的数目有关，但还得根据实际情况而定。如果自变量和因变量的关系是接近于对数函数，则间距应按对数单位变化。

(二) 反应的控制

在实验目的明确、被试者已选定、刺激变量及操作也规定好了之后，如果把控制变量的控制方法也确定了，那么接下来的问题就是反应的控制了。可以设想，对于一个刺激，被试者个体所进行的、或能形成的反应种类是无限的。例如，被试者在注视一个视错觉图形时，有时把这一刺激作为整体而观察，也有时只注视特定的一部分。显然，这就会导致不同的实验结果。于是，把实验中的被试者的反应控制在主试者所设想的方向上，这就是反应的控制问题。

以人作被试者，往往用指示语来控制被试者的反应，指示语乃是心理实验中主试者给被试者交代任务时说的话。使用指导语时，应注意在允许的范围内做到引起动机，激发兴趣。被试者来到实验室时，不一定对参加实验感兴趣。因此主试者必须利用言词来引起他们的兴趣。在可能的范围内，告诉他们实验目的与应用价值，使他们认识到参与和合作的意义。

总之，指示语的内容与语调都可能影响实验结果。近年来，指示语已成为一个重要的自变量，广泛地得到心理学家们的重视。内隐学习的创始人雷伯(Reber, 1976)及其同事们进行坚持不懈的研究，积累了丰富的经验，获得了许多成果。他们的做法是：在内隐学习实验的学习阶段，向被试者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顺序的一种人工语法。给被试者的指示语，或是“学习并记住这些字串”(内隐学习)，或是“试发现这些字串的内在规则，以便更好地记住这些字串”(外显学习)。换句话说，记忆指示语产生内隐学习条件，而规则发现指示语产生外显学习条件。在实验的测试阶段，向被试者呈现一些新的字母串，让被试者判断有哪些符合学习阶段所呈现的刺激的规则，又有哪些不符合。这种用指示语的新方法导致了记忆研究的新发现，即除有意识的外显记忆(ex-plicit memory)之外，还存在着另一个相对独立的记忆系统——内隐记忆(implicit memory)。

(三) 反应指标的选择

在心理学实验中，一般常用的指标有：绝对阈限、差别阈限(以上见第四章第一节)、反应时(见第三章第一节)、反应持续时间、反应程度、完成量、错误率(上见第三章第一节)、完成一定的作业所需要的时间、达到一定基准所需要的次数以及口头报告等等。这么多常用指标，如何作出选择是由选择指标的条件决定的。选择指标的条件很多，主要有：

1. 有效性 即是指标充分代表当时的现象或过程的程度，也称为效度(validity)。哪一个指标最能充分代表当时的现象或过程，那么这个指标的有效性最高。选择任何指标首先要考虑其有效性，如果效度不高，指标就无用。反应指标的效度直接关系到实验的效度。为了使所用的指标具有较高的效度，应了解指标本身的意义是什么、此指标的变化意味着什么、利用此指标对所研究的现象最多能了解到什么程度、有何局限性、如何补救。只有

这样才能全面考虑，才能选择好指标。例如，大多数读者会注意到，地平线附近的月亮比天顶的月亮看上去大；落日时的太阳比正午时的太阳看上去大。其实，月亮和太阳的大小无论在何时都是一样的。因此，这一现象是一种视错觉，心理学上称之为月亮错觉(moon illusion)。雷曼(Reiman, 1920)发表了月亮错觉的实验结果。即落日时所看到太阳的大小是正午时看到太阳大小的3.32倍，而波林(Edwin Garrigues Boring, 1886~1968)和赫威(A.H. Holway, 1940)得出的错觉量为1.7~1.9倍，考夫曼和拉克(Kaufman & Rock, 1962)得出的错觉量为1.5倍。相比之下，雷曼实验的错觉量要大得多。雷曼测定月亮错觉的方法是，在正对太阳的方向上呈现一定大小的圆板，挪动圆板的距离，使之看上去与太阳的大小相等，分别求出落日时和正午时圆板的观察距离，根据这两个距离之比，算出落日时和正午时的太阳外观的大小之比。即，圆板与落日时的太阳的大小相等时，圆板的观察距离为16.6米；圆板与正午时的太阳的大小相等时，圆板的观察距离为5米，则正午时和落日时的太阳外观大小之比为1:3.32。但是，外观的大小并不是观察距离简单的一次函数，这从过去大小恒常性的许多研究结果来看是很明确的。因此，雷曼实验中所选择的指标的有效性不高。另外，如果用望远镜式的纸筒，挡去一切背景来看太阳，这时月亮错觉不存在了，相反会觉得落日时的太阳其大小看上去与当空时的太阳是一样的。这一例子说明了在选择反应指标时考虑其有效性的必要。

2. 客观性 是指此指标是客观存在的，是可以通过一定的方法观察到的。反应时、反应频率、完成量等都是客观存在的指标，是可以客观的方法测量和记录下来的。一个客观的指标一定能在一定的条件下重现。这样的指标能经得起检验，并能够重复进行实验，验证结果。

3. 数量化 指标能数量化，也就便于记录、便于统计，并且量化的指标就能进行比较。

除了上述三个条件之外，还有指标的可靠性和可辨别性等。但是，在这几个选择指标的条件中，有效性是最重要的。

在选择指标时，还要考虑技术设备的条件和可能性。如用脑电波来研究高级神经活动，这是有效的。但是，没有脑电波设备就无法以此作为指标。另外要注意：在测定、记录、观察反应的时候，常常会有仪器误差，操作者的记录误差等，这都要尽力防止。

(四) 仪器和材料的使用

心理学研究使用仪器和材料的主要目的是：呈现实验处理、控制环境、观察行为、及处理观察资料。可把心理学中的仪器材料分为以下四种类型：

1. 呈现刺激的仪器和材料 经常使用的有速示器，时距控制器，记忆鼓，色轮，闪光融合仪，棒框仪，镜画器，深度刺激仪，及用于各种测验的量表和罗夏墨迹测验图(以上见第十一章)等等。

2. 侦察或记录反应的仪器或材料 常用的有记数器，反应时仪，示波器，多道扫描仪，眼动仪(以上见第十一章)，以及各种形式的记录纸等。

3. 控制环境的装置 经常使用的如暗室、隔音室、屏蔽室、单向玻璃装置、闭路电视、眼罩等。

4. 电子计算机 其在心理学上是一种具有多种用途的仪器设备，既可作呈现刺激和侦察或记录反应的装置，也可进行模拟实验和统计处理等。

研究者可根据课题的需要和实际的可能选择使用仪器和材料。在选择使

用仪器和材料时，应以经济实用为原则，不要片面地追求高精尖。实际上，心理学实验，特别是演示实验，常常用简单的材料就可以进行。这就犹如医学上的对症下药一样。

心理学研究必须在良好的环境下进行。这就应当做到：实验室内空气新鲜，温度正常，光线自然而充足，闲人免入，实验时不得闲谈和吸烟等。这些都是最基本的条件。

四、实验资料的收集、分析和综合

实验开始后，研究者的任务是收集资料。由于心理现象的复杂性，在心理学研究中准确无误地收集资料往往是十分困难的。这就更需要研究者具备敏锐的洞察力，严谨的科学态度，对观察资料进行全面而细致的记录。

（一）心理学研究的资料类别

我国心理学家黄希庭等人将心理学研究所能收集到的资料大致分为如下四类：

1. 计数资料 (enumeration data) 就是按个体的某一属性或某一反应属性进行分类记数的资料。这种资料只反映个体间有质的不同，而没有量的差别。例如，被试者的男或女、成年或未成年；反应的有或无、对或错；等等。

2. 计量资料 (measurement data) 就是用测量所得到的数值的大小来表示的资料。例如，被试者的年龄(岁)、体重(斤)、脉搏(次/分)、反应频率(次/秒)、智商(IQ)，等等。

3. 等级资料 (ranked data) 介于计数资料和计量资料之间，可称为半计量资料。心理量表法所取得的资料都属于这类资料。例如，被试者的领导能力可划分为强、中、弱；某个团体对某项改革措施的态度可分为强烈赞成、赞成、无意见、反对、强烈反对，等等。

上述三种资料统称为数据。根据研究需要，上述三种资料在一定条件下是可以互相转化的。

4. 描述性资料 (descriptive data) 即非数量化的资料。在心理学研究中，数量化的资料固然重要，但描述性资料也同样重要。少数初学心理学实验的学生有一种片面的观点，他们只重视数据，实验时只记几个数目字，而对于实验的条件、过程、观察到的现象、及各种影响都不作记录。为了说明问题，不仅要有数量化资料，而且还要有描述性资料。描述性资料可以补充说明数据，使数据更有说服力。但描述性资料由于没有数量指标作为客观尺度，在对它进行解释时容易产生主观片面的错误。因此对它的解释务必更加小心谨慎。

（二）研究资料的整理与解释

实验和观察到的数据应当用统计方法加以整理并制成图表。在统计表中则要列出均值、极差、标准差和标准误等数值。如果研究是在实验组和对照组、实验组和实验组之间进行对比，则还需要注明其 t 值、F 值、和可信限的 P 值，等等。实验结果的分析，主要是应用统计法、通过计算和检验来求得问题的答案或求得对假设的证实或否定。在应用统计技术时，有两点要特别注意：

1. 选用某种统计方法都有其前提条件，不能不分青红皂白，随手拿来就用。例如，t 检验与 F 检验都假定资料的数值具有常态分布。因此，在选用这些方法之前应当先检查资料是否能满足其前提条件。如果不能满足，就应

改用其他合适的方法，或不勉强做定量分析。现代数学为我们提供了十分丰富的统计技术。

2. 根据实验所得到的资料，在验证实验的假设时，不外有两种结果：(1) 如果实验结果与假设相一致，则进一步演绎而找出更深的假设，推进实验；(2) 如果实验结果与假设不一致，则修正假设，反复实验；或放弃它；或建立新的假设，再进行实验。

统计分析之后的工作是根据结果的分析对问题做出结论，判断实验前作的假设能否成立。在说明结果时，可以将前人有关的研究与本实验的发现进行比较，以发现其异同。在作结论时，要实事求是，只能根据在本实验的条件下就所得到的资料和事实作出判断，不能做过分的推论。如果测验结果具有广泛的含义，要谨慎地指出。即使观察到两个变量有高度的相关性，也不能过于肯定它们之间有因果关系。因此，在得出结论时，必须时刻牢记要“敏于事而慎于言”，客观而真实地反映研究结果。

五、撰写实验报告

心理学实验完成之后，就必须把实验报告写好，实验报告是总结科研成果的一种形式。与其他学科一样，心理学的实验报告是对过去工作的总结，更重要的是能为进一步的研究提供线索。学习写好实验报告也是实验课的教学任务之一。写心理实验报告，一方面要完整全面地阐述实验进行的情况；另一方面又要写得简洁明了。

心理学实验课中所写的实验报告与科学研究的实验报告基本项目是相同的，但也有不同之处。科学研究的目的是要解决新问题，在科学研究的实验报告中必须提供新的研究成果，而在实验课中做实验是为了学习，常常是重复一个经典实验或验证某个已有定论的问题，因而在实验课的实验报告中往往只能提供前人研究成果的补充材料。但因为实验是在新情况下的重复，在结果中包含有新的因素，因此也需要整理、分析，写出高水平的实验报告。当然，实验研究的问题的范围很广，解决问题的方法各不相同。因此，在写实验报告时可能会有些差别，但基本的形式和要求是各类实验报告都必须具备的。

一个完整的实验报告，必须包括以下几项内容：摘要、题目、引言、方法、结果、讨论、结论、参考文献及附录。现具体分述如下：

1. 摘要 正式发表的科研报告，一般应写出论文摘要（abstract），把它放在正文的前面。论文提要应当以最概括、最简洁的语言写出，内容包括本课题所要解决的问题、方法、以及获得的结果和结论。如果论文在供国际交流的《心理学报》、《本土心理学研究》、《心理科学》等高水平的杂志上发表，则还应写出外文摘要附于论文的最后。近几年来，随着国际上联机检索文献的发展，目前一般在论文摘要的后面，还要列出论文中的关键词（key words），以便于学术交流。

2. 题目 题目（title）是说明要做的实验研究是属于哪方面的问题。一般要求在题目中既要指出自变量，也要指出因变量。例如，“照度对视觉敏度的影响”，在这一题目中，“照度”是自变量，“视觉敏度”是因变量。这样，只要一看题目就知道这个实验的总轮廓了。

3. 引言 在引言（foreword）中一般要求说明此实验的意义以及题目产生的过程，提出问题的背景材料或提出问题的假设，最好能引经据典，把这类实验的来龙去脉指出来。当然，语句要简练，一般不超过一千字。

一般来说，问题来源可从如下三个途径指明：

(1) 为了扩展以前的工作，或探讨过去尚未解决的问题，在引言中要把以前的工作简要地作介绍，以便与本实验进行衔接。

(2) 若题目来自对以某一理论为根据而提出的假设的论证，则在引言中对这一理论的内容和背景、及假设的由来要解释清楚。

(3) 若题目来自实际部门提出的问题，则在引言中就要对此实际问题进行介绍。

4. 方法 方法 (method) 主要在说明取得实验结果的实验设计，如实验设计有独到之处，或有创意，则应多花些笔墨，介绍这些对传统方法有突破的新方法。另外，在方法中对技术路线的介绍是必不可少的，其中包括：

(1) 被试者：要说明被试者选择的方式，被试者的年龄、性别，及其他有关方面的情况，被试者的数目、及如何进行分组的，等等。

(2) 仪器、材料：实验所用的仪器及材料的名称要一一写上，必要时注明仪器的型号。有时同类的仪器，型号不同，实验结果也可能不同。

(3) 实验程序：即具体说明实验是如何进行的，进行实验的原则、方法步骤、指示语是什么、要控制什么条件，等等。这部分要写得清楚、确切，以便他人可随时照此重复实验、验证这个实验。

5. 结果 结果 (results) 主要是指统计结果，也就是原始材料经过统计后，以图、表形式表示出此实验的结果。另外，把观察结果的记录及被试者的口头报告也列在这部分中。必须指出，在结果中所列的全部内容，必须都是来自本实验的，既不能任意修改或增减，也不要加入自己的主观见解，要使读者清楚地了解这个实验的客观结果。有不少实验报告撰写者把结果看作仅仅是图表的罗列，这也是片面的。对每一幅图、每一张表，一般都应配以文字解释。例如，说明曲线的走向，P 值的显著性水平。当然，大篇幅的分析和讨论应放在“讨论”这一栏目下。

6. 讨论 讨论 (discussion) 是根据实验结果对所要解决的问题给予回答，并指出假设是否可靠。如果结果不能充分地说明问题或者各部分有矛盾时，就要进行分析，找出原因。如果结果与别人的结果不一致时，可以进行讨论，提出自己的见解。当实验得到意外的结果时，也要进行分析，不能弃之不管。因为意外的结果有时会有意外的发现。在讨论中，还可以对本实验的程序、所用的仪器、及进一步的研究提出修改意见和建议。对于方法上有突破的实验，可以将方法单列一点，进行方法学上的专论。

7. 结论 结论 (summary) 是说明本实验证实了或否定了什么假设。结论一般以条文形式、用简短的语句表达出来。但是，结果必须恰如其分，不可夸大，也不必缩小，一定要以本实验所得的结果为依据，确切地反映整个实验的收获，切忌把话说过头。

8. 参考文献 要把参考文献 (references) 的题目、出处、作者、出版日期都写明，以便查找。文献的顺序一般在文章中出现的先后为序，也可按出版社的规定。

9. 附录 学生型的实验报告一般要把全部的原始资料都列入附录 (appendix)。因为同一结果，不同的人，在不同的时候都可以进行不同的分析和处理。另外，重要的实验材料、指示语等也应列入附录。

上述各项仅是为了便于阐述和讨论才把一篇科研报告解剖成几个部分。实际上，由于研究的题材和内容的不同，上述各项完全可以、而且应当根据

需要有所增删，也可以把其中的某些项合在一起写。但实验报告的基本思路 and 结构，一般都包括上述各项内容，同时也要按上述顺序行文。

第三节 实验心理学简史

心理学成为一门独立的科学不是一蹴而就的，它和物理学、生理学等学科一样，也有漫长而艰难的发展历程。心理学的形成和发展无不与实验心理学联系在一起。实验心理学在整个科学史上，尤其在心理学史上占有很重要的地位。它建立于十九世纪下半叶，开始于应用自然科学的方法研究心理学问题。它的诞生不仅使心理学获得了收集材料的新手段，而且使心理学建立在精确可靠的实验基础上。从实验心理学的建立到目前仅一百年的时间，它的发展是十分迅速的。

心理学是一门古老而又年轻的科学，也就是说，心理学自古有之，《论语》、《孟子》等古书中就曾论及心理学问题，荀子就是中国古代一位伟大的心理学家，他论及到身心关系、心物关系等一系列的心理问题。古希腊亚里斯多德对心理学贡献较大，他当时就提出诸如感觉、记忆、认识、情感、联想、统觉等问题。

实验心理学的发展，归纳起来大致经历了如下两个阶段：

一、心理实验发展的第一阶段

大约从 18 世纪中叶到 19 世纪中叶是心理实验发展的第一阶段，也可说是实验心理学的预备阶段。从 18 世纪中叶开始，生理学家、物理学家和天文学家研究了一些感觉和知觉范围内的问题。例如，反应时的测定、后象的延续时间、某些差别阈限的测定，等等。他们把自然科学的研究方法逐渐引入心理学的研究中，虽然当时应用的方法、技术比较简单，类似某些物理实验和生理实验。被试者的自我观察和陈述都带有十分初级的性质。例如，对某一声音听见或听不见，对某一强度的刺激觉察或不觉察，等等。尽管如此，长期被唯理论禁锢的心理学已被冲开了一道缝隙，心理学的研究活跃了许多，科学的心理学正被孕育着，这阶段前后大约经历了一个世纪。在这一段时间里，尽管收集了大量的在心理学范围内非常有趣的实验材料，但对大多数哲学家和自然科学家来说，人的心理是永远也不能用实验的方法来加以研究的。至 19 世纪中叶，随着社会实践、自然科学的迅速发展，与心理学关系密切的生理学有了一系列新的发展成果，例如，缪勒（Georg Elias Müller, 1850 ~ 1934）感官特殊能力的研究，赫尔姆霍茨（Hermann Helmholtz, 1821 ~ 1894）和韦伯等的感官生理的研究，以及人们对人类自身某些规律的认识又深入了一步。所有这一切都为心理学准备了科学的基础知识和研究方法。这时哲学也已为心理学积累了不少心理事实的材料，确定了研究对象。当时在正理学问题的研究上，莱比锡大学解剖学和生理学教授韦伯（Ernst Heinrich Weber, 1795 ~ 1878）首创了实验测量和用数学公式表示感觉的差别阈限（differential threshold），韦伯根据多年的研究结果，总结出关于人的感觉方面的某些规律性的知识，后人称之为韦伯定律（Weber's law）。莱比锡大学物理学教授费希纳（Gustav Theodor Fechner, 1801 ~ 1887）发展了韦伯的工作，他根据当时物理学和数学知识，对感觉阈限问题进行了深入的实验研究和精密的数学论证，提出一个感觉强度与刺激强度的所谓心物关系的对数定律，同时制定了心理物理学（psychophysics）的基本方法。费希纳的心理物理学实验具有两个明显的特点：（1）利用了专门为研究心理物理学制定的实验方法，即极限法、恒定刺激法、及平均差误法（见第四章第二节）。（2）对实验结果作数学处理。费希纳的心理物理学实验的这两个特

点，在心理实验的发展中标志着心理学的研究方法向前迈了重要的一步，这意味着心理学从利用物理学和生理学的实验方法过渡到利用自己的特殊的心理学方法的开始，即标志着心理实验的“心理学化”的开端。多次重复同一试验、并对所得结论作统计处理，意味着心理实验从比较草率的、不精确的实验逐渐向专门的实验心理学过渡。费希纳的《心理物理学纲要》的出版，对以后心理学实验方法的建立起了重要的作用。

心理实验在这第一阶段具有以下特点：（1）应用的方法类似于某些简单的物理学和生理学的实验方法，研究的问题只限于某些简单的心理现象的量的测量，如视觉敏度、正后象的延续时间、差别阈限的测定、反应时间的测量，等等。（2）实验技术简单，对条件控制、及实验结果的数据处理都还有欠缺。（3）被试者的自我观察与陈述都带有十分初级的性质，例如，听见或听不见某一声音、两个刺激的强度差别觉察或不觉察，等等。

二、心理实验发展的第二阶段

心理实验发展的第二阶段始于19世纪60年代，一直延续到今天。开始阶段是以冯特的工作为标志的，这一阶段也可称为实验心理学的创立、传播和发展阶段。1879年，冯特在莱比锡大学建立了第一个正式的心理实验室，还培养了一批各国的心理学家。这对实验心理学的传播起了很重要的作用。对于冯特来说，心理实验就是系统的自我观察，而一切实验手段只是自我观察的辅助手段。在冯特看来，心理学的研究对象是所谓纯粹的“直接经验”，冯特所采用的实验方法仅仅是进行“科学的内省”。冯特还认为，“经验”是由许多心理元素构成的，他希望通过“内省”把“经验”分解为简单的心理元素，诸如感觉和感情，犹如化学中把水分解为氢和氧一样。冯特认为实验心理学的主要任务是在严格控制的自我观察的帮助下精确地分析个体经验。这就意味着，心理功能只有分解成简单的感觉成分才可以放到实验室中去研究。所以，冯特的实验法又称作内省实验法（或内省分析法）（introspective experimental method）。冯特的学生铁钦纳（Edward Bradford Titchener, 1867~1927）发展了其师的“内省实验法”，在内省实验法上，铁钦纳为提高自我观察的效度设定了更多的限制，要求被试者接受更严格的训练，使研究的范围越来越窄。虽然铁钦纳的极端化倾向限制了其体系的发展，但他严格的科学态度对心理学基础研究所作的贡献已被载入史册。

这一阶段心理实验的特点是：在实验中提出了被试者做系统的自我观察计划，被试者起着“观察者”的作用。冯特的心理实验理论也决定了其实验方法在心理学中应用的范围。冯特的兴趣主要在感觉方面，对于思维等“高级的”心理过程，冯特认为是不能用实验方法进行精确研究的，对这些“高级”的心理过程的研究，他提出了特殊的非实验的方法，称之为“民族心理学”的方法。他的学生屈尔佩（Oswald Külpe, 1862~1915）着重进行高级心理研究，并提出了“无表象思维”的理论，1904年在符茨堡大学建立了符茨堡学派（Würzburg School）的实验室，符茨堡学派是以自我观察的实验方法研究思维过程。冯特对此表示坚决反对。

除冯特式心理实验之外还有不少别的种类的心理实验，因素型实验或函数型实验就是其中之一。这些实验的任务不在于精确地分析意识过程，而是企图找出一定现象产生的原因，或是阐明两个变数之间的函数关系。艾宾浩斯（Hermann Ebbinghaus, 1850~1909）所做的许多记忆实验就属于这一类。艾宾浩斯运用严格的实验方法研究记忆，这就突破了心理实验局限于感觉、

知觉的范围，为研究“高级心理过程”的基本实验方法与材料提供了基础。同时，艾宾浩斯的兴趣从主观行为转向客观行为，他不用记忆的主观经验，取而代之的是客观指标——回忆量。艾宾浩斯的记忆研究开拓了高级心理过程实验研究的新时代。

此外，在 19 世纪 90 年代兴起的“测验式实验”或“心理测验”以及“动物心理实验”、“儿童心理实验”等新类型的心理实验，都有别于冯特式的心理实验。这个变化过程有以下几个特点：（1）制定和应用了实验研究的一般方法，注重对实验条件的严格控制；（2）制定和应用了实验研究的特殊方法，如记忆研究法、情绪研究法，等等；（3）广泛应用最新科学技术成就和统计学方法，这类新类型的心理实验的目的是力图得到精确、可靠和客观的实验结果。

到了 20 世纪初，在美国兴起了一个新的心理学学派——行为主义（behaviorism），创始人是华生（John Broadus Watson, 1878 ~ 1958）。华生于 1913 年发表的《一个行为主义者看来的心理学》的文章，是这一学派诞生的标志。行为主义反对冯特把意识当作心理学研究的对象、把内省当作心理学研究的方法，主张心理学研究的对象为行为，认为心理学要走生物科学的道路。华生拒绝把自我观察作为心理学的一种完全的研究方法，因为自我观察至少不能用于研究婴儿和动物，婴儿和动物是不具备报告自我观察的能力的。行为主义认为，人和动物可观察的活动是行为，因而，心理学研究的对象只能是行为。对行为的研究包括刺激和反应两个方面，所以，华生的行为主义也称刺激-反应（S-R）心理学（参见图 1-3）。刺激（stimulus）是指外界环境和身体内部的变化，如光、声、渴、饥等。反应（reaction）是指有机体所做的任何外部动作（外部反应）和腺体分泌（内部反应）。反应有遗传反应和习得反应两种。复杂反应和动作技能是通过建立条件反射学会的。

刺激 反应

信号（输入） 信息加工（讯息处理） 行为（输出）

在中国，行为主义的代表人物是郭任远（1898 ~ 1970），他和华生都否认本能的存生，郭任远毕生从事动物个体行为发展的研究，1921 年在美国发表《取消心理学上的本能说》，批判麦独孤（William McDougall, 1871 ~ 1938）的本能理论，震惊了美国心理学界。他坚信动物的理论只能从解剖生理的现实生活环境和个体发育的历史研究中探索。他对鸡胚胎行为发展的研究和小鸡啄食行为发展的研究证明，小鸡啄食的动作是从鸡胚胎的心脏跳动之际开始发展起来的。他训练猫不吃老鼠的实验研究表明，通过发育时期生活条件的控制，可以使猫亲鼠、护鼠、甚至怕鼠。这项实验也受到国际心理学界的重视。郭任远是国际上有影响力的心理学家，他的著作有《人类的行为》（1923）、《行为学的基础》（1927）、《行为主义》（1934）。在 1946 ~ 1970 年期间，他定居香港，著有《行为发展的动力形成论》。

行为主义在美国统治了四、五十年，到了 50 年代后期，美国心理学中出现了一个新兴的理论方向和研究领域，到 70 年代已成为美国心理学的一个主要方向，这就是认知心理学（cognitive psychology）。认知心理学反对局限于研究孤立的、外观的、可观察的反应，而致力于了解心理活动的过程。这种活动就是传统心理学中称之为认识活动的东西，也就是全部的信息加工（或讯息处理）（information processing）活动。认知心理学把研究的重

点转到内部心理过程上去了。图 1-3 (参见上页) 用模型的方式表示行为主义和认知主义的区别。在 70 年代, 曾任美国心理学会主席的麦基奇 (W.J.McKeachie, 1921 ~) 在一篇报告中说, 心理学是什么, 这一概念已发生改变, 在詹姆斯的经典教科书里, 把心理学定义为“对意识状态的描述和解释”, 后来改为“研究行为的科学”, 今天我们的定义又改变了, 心理学又回到意识上来了。但是, 更多的教科书和心理学家把心理学定义为研究行为和心理过程的科学。这就既包括外观的、可观察的动作, 又包括内部的心理活动。在研究方法上, 行为主义强调严格的实验方法, 排斥一切主观经验的报告, 认知心理学则既注重实验室研究、又重视主观经验的报告。目前, 认知心理学已经成为心理学中的一个重要潮流。可以这样说, 心理实验方法发展的第二阶段, 前期主要受到行为主义的影响, 后期主要受到认知心理学的影响。

从以上心理学实验发展的历史我们可以看出, 由于在心理学中采用了实验法, 才有了对心理现象进行客观研究的手段, 才能从对心理现象的一般哲学推论, 进入到具体心理过程及其物质基础的分析研究, 从而能越来越深入地揭示各种心理活动的规律, 大大丰富了心理科学。从某种意义上说, 正是因为实验心理学和心理物理法的诞生和发展, 才使心理学完全从哲学中分离出来, 成为一门独立的科学。

实验方法的建立, 特别是从 20 世纪 50 年代以来, 心理学已广泛应用了现代科学和工程技术的最新成就。近年来, 随着科学技术的日益发展, 国际上的一些新技术、新概念和新方法已被引进心理学领域中, 如控制论、信息论、微电子技术、分子生物学、数理分析等, 使心理学的实验研究水平有了显著的提高。心理学家从信息加工的角度去看待知觉、学习、记忆及思维等心理过程。心理学与邻近学科有了越来越多的共同语言, 从而更有利于其发展。这使我们认识到, 心理学应注意从邻近学科中汲取有益成果, 这样才有助于逐步揭示心理现象的奥秘。

实验方法的应用, 使心理学的发展大为加快。从心理学实验方法的建立到现在才一百多年, 而心理学的发展却大大超过以往许多世纪。虽然科学的心理学不能归结为实验心理学, 实验方法也不是心理学研究的唯一方法。但是, 任何当代心理学的教科书都以大量的篇幅证明, 现代心理学中的大量事实大都来自先前的实验研究, 心理学实验的诞生和发展有力地说明了实验方法是揭露心理和行为的规律的重要途径和手段。就以武德沃斯 (Robert Sessions Woodworth, 1869 ~ 1962) 等所著的《实验心理学》一书中所列举的参考文献来说, 1938 年的第一版列举了 1770 项, 到 1955 年的第二版就列举了 2480 项, 其中有 50% 是旧版中未曾引用过的。1971 年的第三版的参考文献目录长达 100 页, 不下于 4000 项, 其中 80% 以上是旧版未曾引用过的。它所引用的资料主要来源于美国的材料, 主要也只是涉及心理学基础领域的实验研究。

一位心理学工作者可以对心理学的任一领域或任一分支进行特别研究, 可以专门从事儿童心理、教育心理、医学心理、或知觉心理、思维心理, 以至社会心理的研究。但是, 如果他想成为一个真正严格的、科学的心理工作者, 他就必须要很好地掌握实验心理学的研究内容和方法, 了解应当如何科学地考察心理和行为的规律。

实验研究和任何一种研究方法一样, 有它一定的局限性。这是因为实验

研究是在一定控制条件下进行的，实验结果有时会与现实生活中人们的心理活动不完全一样。这种现象的发生在科学研究中并不少见，更不是不能解决的。只要研究课题来自实际，研究结果又不断到社会实际中去检验；同时，在研究过程中又不仅仅注意心理活动的单因素的实验研究，还注意多因素的交互作用的研究，那么，其研究结果是会与现实生活相一致的。

综上所述，综观心理学研究的历史，可以这样说，实验法是心理学的主要研究方法，实验心理学的诞生意味着科学心理学的建立。它是使心理学成为一门独立科学的基石。当然，它也不是心理学的唯一的研究方法，它和观察法(observation method)、临床法(clinical method)、问卷法(questionnaire method) 等研究方法相辅相成，构成完整的心理学方法学。

本章实验

一、自变量和因变量的确定

(一) 目的：

- (1) 通过动作学习的过程了解心理实验中确定自变量和因变量的方法；
- (2) 学习使用触棒迷津。

(二) 材料：触棒迷津、小棒、遮眼罩、秒表、记录纸。

(三) 程序：

(1) 三人一组，被试者带上遮眼罩，用小棒走迷津（实验前被试者勿看迷津，也勿用手触摸迷津）。主试者对被试者的指示语必须这样说明：“在排除视觉的条件下，尽快地学会用小棒走迷津，中间不要停顿，要积极运用动觉、记忆和思维，争取早些学会”。主试者把小棒放在迷津的入口处，然后让被试者用优势手拿住小棒，手臂悬空。

(2) 被试者手执小棒静候。主试者在每次开始前 2 秒钟，先发出“预备”口令，主试者再说“开始”时，被试者才用小棒走动。在发出“开始”口令的同时，主试者开动秒手表。

(3) 被试者在走迷津的过程中，凡进入盲巷一次就算出错一次，主试者记下错误次数。

(4) 当被试者的小棒进入迷津终点，主试者立刻说“到了”，同时停秒表，记录走一遍迷津所用的时间（秒）。再做下一次的准备工作。

(5) 学习遍数因被试者而异，均以连续三次不出错为学会的标准。

(6) 若被试者在学习途中感到疲劳，可在某次走到终点后休息几分钟。

(四) 结果：将每次走迷津所用的时间和错误次数列成表格（见下页）。

本实验记录用纸

自 变 量 因 变 量	学 习 遍 数															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	...
所需时间 (秒)																
错误次数 (次)																

注：学习遍数因被试者而异，均以连续三次不出错为学会的标准，学习遍数超过 15 次者可自行加行。

(五) 讨论：

(1) 本实验中自变量是什么？为什么在实验前要对所用的自变量提出一个操作上的定义？

(2) 本实验用什么作因变量的指标？它的作用是什么？

(3) 本实验控制了那些变量？

二、自变量的范围和间距

(一) 目的：通过对皮肤两点阈的测定，学习确定心理实验中变量范围

和间距的方法。

(二) 材料：两点阈量规、遮眼罩、记录纸。

(三) 程序：

(1) 主试者选定被试者的左手背(A区)与前臂背面为测量区，前臂背面又分为二个区(B、C区)，本试验取A、C两个区。测量前，在各区内标上蓝印或划出圆圈，刺激在圈内进行。

(2) 在使用两点阈量规时，必须垂直接触皮肤，对两个尖点施力均匀，接触时间不超过2秒钟，先在自己手上练习几次，再在被试者的非测验区练习几次。

(3) 实验序列的长度和起点，可根据初步测验后确定，大致在7~15毫米的范围内取5个点，每步变化在1毫米左右。每种间距做20次，顺序随机安排。

本实验记录用纸一

刺激代号相对应的实际刺激值

代号	1	2	3	4	5	...
刺激值(m)						

(4) 每隔100次，休息5分钟。被试者有时产生一种持久的两点后象(即使只有一点或距离很近的两点刺激，也有感到两点的印象)，此时应给予休息，使其恢复两点的标准。每隔数十次，可插入一次一点刺激，测验被试者是否有这种后象。

(5) 主试者对被试者的指示语必须这样说明：“在排除视觉的条件下，当被试者感觉到是两点时说‘二’；当被试者感觉到一点时说‘一’；当被试者不能确定感觉到的是几点时说‘不’”。

本实验记录用纸二

被试者反应记录表

	A区						C区					
	1	2	3	4	5	...	1	2	3	4	5	...
被 试 者 报 告 「 两 点 」 或 「 一 点 」												

(6) 主试者在每次刺激前先发出预备信号：“注意”。在被试者作出反应后，要记录下来。

(7) 在全部测量完毕后，在记录中挑出有“不”的序列，再重复那些序列，获得确定的判断。

(四) 结果：分别求出左手A区和C区的两点阈限，并对两均数的差别进行考验。

(五) 讨论：

(1) 说明身体部位不同，触觉有差异，两点阈值从一个侧面揭示了触觉感受性的大小。

(2) 如果要测定皮肤某部位的两点阈限，自变量的范围和间距应如何安排？

本章摘要

1. 实验心理学就是在实验控制条件下对心理和行为进行研究工作心理学。这一心理学分支是以研究方法定义的。也就是说，只要是用实验法来研究的心理学问题，都可以包括在实验心理学的范围内。

2. 实验是一种控制的观察，与自然的或偶然的观察不同，实验是实验者人为地使现象发生，对产生现象的情景或影响现象的条件加以操纵、变化与控制的观察。

3. 实验具有三个明显的特征：实验者掌握有主动权；实验具有可核对性、或验证性；实验者可以系统地变化条件，观察因这些条件的变化而引起的现象上的变化。

4. 变量是指在数量上或者质量上可以改变的事物的属性。在心理学的实验中，研究者必须考虑三类变量：(1) 已知的、对有机体反应发生影响的变量，叫作自变量，它是由实验者操作、掌握的变量；(2) 由操纵自变量而引起被试者的某种特定的反应，叫因变量，它是由实验者观察或记录的变量；(3) 在实验中应保持恒定的变量，叫控制变量。

5. 心理实验的对象统称为被试者，包括人和动物。在实验中，特别是在以人为被试者的实验中，处理好实验者（即主试者）与被试者的关系是实验取得成功的一个重要条件。

6. 额外变量（即控制变量）是使实验结果发生混淆的主要根源。对额外变量的控制常用的方法有：排除法、恒定法、匹配法、随机化法、抵消平衡法和统计控制法。

7. 实验效度是指实验方法达到实验目的的程度。换言之，实验效度就是实验结果符合客观实际的程度。实验效度主要包括内部效度和外部效度。

8. 实验的程序是实验的进程，它是指实验在各个阶段应做的事。心理学实验通常遵循的基本程序是：(1) 课题的确定；(2) 被试者的选择；(3) 实验因素的控制；(4) 实验资料的收集、分析和综合；(5) 撰写实验报告。

9. 课题的确定是科学研究的第一步，研究课题的来源通常有四个方面：实际的需要，理论，个人的经验，过去的研究与文献。

10. 心理学实验的被试者是多种多样的。有正常人，也有疾病患者；有各种年龄阶段的人，也有动物。不同的实验对象，通常要求有不同的实验操作。

11. 操作定义的提出，受到心理学家们的欢迎。在心理学上，对一个心理现象根据测定其程序下的定义称之为操作定义。

12. 一个完整的实验报告必须包括以下几项内容：摘要、引言、方法、结果、讨论、参考文献及附录。

13. 实验心理学的发展，大约经历了两个阶段。从 18 世纪中叶到 19 世纪中叶是心理实验的第一阶段，也可说是实验心理学的准备阶段。在这一阶段具有以下特点：应用的方法类似于某些简单的物理学和生理学的实验法；研究的问题只限于某些简单心理现象的量的测量；被试者的陈述带有十分初级的性质。

14. 心理实验发展的第二阶段始于 19 世纪 60 年代，一直延续到今天。开始阶段是与冯特的工作相联系的，这一阶段也可称为实验心理学的诞生、传播和发展阶段。心理实验发展的第二阶段，前期主要受到行为主义的影响；后期主要受到认知心理学的影响。

15. 综观心理学研究的历史，可以这样说，实验法是心理学的主要研究方法。实验心理学的诞生意味着科学心理学的建立，它是使心理学成为一门独立科学的基石。当然，它也不是心理学的唯一的研究方法，它和观察法、临床法、问卷等方法相辅相成，构成完整的心理学方法学。

建议参考资料

1. 王甦等 (1993) : 当代心理学研究。北京市：北京大学出版社。
2. 杨治良 (1988) : 心理物理学。兰州市：甘肃人民出版社。
3. 张春兴 (1991) : 现代心理学。台北市：东华书局 (繁体字版)。上海市：上海人民出社 (1994) (简体字版)。
4. 波林 (高觉敷译, 1981) : 实验心理学史。上海市：商务印书馆。
5. 黄希庭 (主编) (1988) : 心理实验指导。北京市：人民教育出版社。
6. Atkinson, R.C. et al. (1988). Stevens' handbook of experimental psychology (2nd ed). New York: A Wiley-Interscience Publication.
7. Kantowitz, B.H., & Roediger, H.L. (1984). Experimental psychology (2nd ed.). New York: West Publishing Co.

第二章 实验设计

本章内容细目

第一节 实验设计的基本类型

- 一、被试者内设计 46
 - (一) 实验前后设计
 - (二) 定时系列设计
 - (三) 抵消实验条件的设计
- 二、被试者间设计 51
 - (一) 随机组设计
 - (二) 配对组设计
- 三、混合设计 56

第二节 多变量实验技术

- 一、多自变量实验的优点 58
- 二、多因素实验设计 63
- 三、拉丁方设计 65

第三节 实验数据的统计分析

- 一、统计表和统计图 66
 - (一) 统计表
 - (二) 统计图
- 二、实验数据的初步整理 73
 - (一) 偶然误差与系统误差
 - (二) 集中量
 - (三) 差异量
- 三、显著性检验 82
 - (一) 显著性检验的含义
 - (二) t 检验
 - (三) F 检验
 - (四) X^2 检验

本章摘要建议 参考资料

前面谈到，心理学是一门实验科学。为了完成好一个实验，并使其在方法上能站得住脚，必须首先做好实验设计，这是实验成功的关键。一个好的实验设计能够用最少的人力、物力来获得最多和最有效的实验数据以达到最佳的实验目的。否则，如果实验设计不当、对非实验条件控制不严、实验方法和程序安排不合理，即使实验的规模较大、实验次数较多，但是所得到的实验数据却不能为最初所提出的问题提供合理的答案，达不到原来的实验目的。这样的实验设计不是好的设计。

一个良好的实验设计是实验研究能够作出成果的重要因素，实验设计是

整个实验研究的重要组成部分。有人以为只有实验終了得出数据之后再考虑统计处理，这样想法是片面的。事实上，只有在实验开始之前就有精心的实验设计，才能事半功倍。心理是一种复杂的现象，它跟自然现象和社会现象一样，也有着自身本身的规律性。实验就是揭示这些规律性。但心理科学的实验设计要比其他科学更为困难，这是因为无论是实验对象，还是环境条件，反应测定都要受到很大的限制。因此，掌握好实验设计技术也就更加重要。科学的探索要求不断地发展自己的手段和方法，也就是要不断地创造出新的“工具”。这些新的方法和“工具”，理所当然地能反映在实验设计之中。历史事实表明，方法上的创新能使我们产生意想不到的效果。

顾名思义，实验设计就是对实验的设计。在长期的科学实验中，科学工作者已经总结出一些在实验设计时所必须遵守的共同原则、具体的实验方法和技术上应注意的细节。本章内容主要讨论下面六个问题：

1. 何谓实验设计，它有哪些基本类型。
2. 多自变量实验设计有哪些优点。
3. 统计表有何功能。
4. 统计图有何功能。
5. 标准差的主要用途有哪些。
6. 什么叫显著性检验？在使用显著性检验时应注意什么。

第一节 实验设计的基本类型

实验设计 (experimental design) 乃是进行科学实验前做的具体计划。它主要是控制实验条件和安排实验程序的计划。它的目的在于找出实验条件和实验结果之间的关系, 做出正确的结论, 来检验解决问题的假设。心理实验设计的内容包括: 刺激变量 (或刺激变项) (stimulus variable) 的确定及其呈现的方式, 反应变量 (或反应变项) (response variable) 的指标及其测量方法, 对一切有关变量 (或变项) (variable) 的控制措施, 确定被试者人数和选择被试者的方法, 拟定主试在实验开始前对被试者要说的指示语, 规定实验次数, 安排实验程序, 规定使用仪器的型号, 处理实验数据的方法等等。实验设计要对实验结果有预见性, 要保证严格按照实验设计进行才能取得有效的实验数据。

每一个实验设计都必须回答三个基本问题: (1) 实验采用多少自变量? 例如在一个阅读速度的研究中取“照明强度”为自变量。(2) 各自变量内又采用多少处理水平? 例如照明强度又分为强、中、弱等处理水平。(3) 在各自变量和各处理水平中用相同的被试者, 还是用不同的被试者? 根据这三个条件的组合, 就可构成许多不同类型的实验设计。

一般根据对上述三个问题的回答, 就可把实验设计的类型大体上分为三种: 被试者内设计、被试者间设计以及同时包括被试者内与被试者间的混合设计 (mixed design)。

被试者如果在自变量发生变化的所有情况下接受实验, 则是被试者内设计。例如, 自变量是照明强度, 假如有强、中、弱等三种照度供操纵; 因变量是观测一个在标准条件下视角为一分的 C 型视标的距离。采用被试者内设计时, 同一被试者需要在三种照明条件下都接受测试。

如被试者只接受多个自变量情况中的一个, 即不同的被试者接受不同自变量的处理, 这类设计便是被试者间设计。在上例中, 如实验者令一组被试者在强照度条件下测试, 令另一组被试者接受中等照度条件下的测试, 令第三组被试者接受弱照度条件下的测试, 如此类推, 直到各组都分别接受某一种照度条件下的测试。这种实验设计就是被试者间设计。

为什么在被试者内设计和被试者间设计中, 实验都可以同时操作两个或两个以上的自变量呢? 乍看之下, 这种处理似乎违背了实验法“单一变量”的原则。其实, 由于实验者采用了统计方法, 可以通过平衡多变量的关系, 通过控制其他一些自变量, 使某一阶段变化特性显露出来, 这一问题就迎刃而解了。因此, 可以说统计技术是实验设计的基础。

所谓混合设计, 是指在一项实验中, 有些自变量是被试者内的, 而有些自变量是被试者间的。例如, 一个被试者接受甲变量的每一种情况, 但只接受乙变量的一种情况。根据被试者内设计和被试者间设计的特征, 这时的甲变量是被试者内自变量, 乙变量则是被试者间自变量了。这类实验设计即是混合设计。

一、被试者内设计

被试者内设计 (或受试者内设计、单组实验设计) (within-subjects design) 是每个被试者须接受自变量的所有情况的处理。其基本原理是: 每个被试者参与所有的实验处理, 然后比较相同被试者在不同处理下的行为变化。这种实验设计下的同一被试者既为实验组提供数据, 也为控制

组提供数据。因此，被试者内设计无需另找控制组的被试者。

在实验研究中，如果实验者主要想研究每一个被试者对实验处理所引起的行为上的变化，一般可考虑采用被试者内设计。被试者内设计又可分为三种子类型：

(一) 实验前后设计

实验前后设计 (experimental before-after design) 是指在实验条件处理前对被试者进行观测的结果与实验条件处理后所做的同样观测的结果加以对比的设计。也就是说，这种设计类型是实验 (处理) 前后的比较设计。我们可用表 2-1 来表示这种设计的模式。

实验 (处理) 前、后设计，具有二个优点：(1) 能较明显地检查出实验处理的效果如何。因为在这种实验中，前、后被试者是同一的，如果能够控制好无关变量的话，那么实验处理前后的差异，就是实验处理的结果。(2)

表 2-1 实验前与实验后设计模式

组数	前观测	处理	后观测
1	Y_1	X	Y_2

说明： Y_1 表示实验处理前对被试者观测所得值

Y_2 表示实验处理后对被试者观测所得值

X 表示实验条件处理

对被试者的需要量较少，一组被试者当二组被试者用，无须再增设被试者控制组。不仅效率提高了，而且被试者变量也得到了较好的控制。

然而，我们应该看到不足之处。各类被试者内设计的共同缺点是，这类设计需要每个被试者在实验中花费很多时间。通常一个被试者被要求执行几项任务，这势必会产生疲劳，影响实验结果。而前、后设计还有二个不足之处：(1) 由于前、后两次观测之间存在时间间隔，这就会带来外来影响。一般来说，前、后间隔时间愈长，则影响就愈大。反之，如果前、后两次观测时间很接近，倒可以认为这是实验处理引起的行为差异。(2) 容易产生顺序误差。即前面的观测影响后面观测的结果，从而影响实验结果。例如实验中第一次观测会产生学习、疲劳、情绪等效应，从而影响第二次观测的结果。

(二) 定时系列设计

定时系列设计 (time serial design) 是指实验处理前对一组被试者作一系列的定时重复观测，然后实施实验处理。再对被试者作一系列的定时重复观测，分析自变量 (实验处理) 对因变量的关系。定时系列设计可以看作是前、后设计的扩展形式。定时系列设计，一般先分别求出实验处理前和实验处理后的平均数，也有分别取众数或中数的。我们用表 2-2 来表示这种设计的模式。

表 2-2 定时系列设计模式

组数	前观测	处理	后观测
1	Y_{1a} 、 Y_{1b} 、 Y_{1c}	X	Y_{2a} 、 Y_{2b} 、 Y_{2c}

说明： Y_{1a} 、 Y_{1b} 、 Y_{1c} 分别表示实验处理前之观测值

Y_{2a} 、 Y_{2b} 、 Y_{2c} 分别表示实验处理后之观测值

X 表示实验条件处理

定时系列设计除了具有上述前、后设计的二个优点之外，还具有以下二个优点：（1）降低由于一次观测而得到被试者不正常行为的机率。实验处理前和实验处理后分别对每一位被试者作一系列测定，这就能降低只作一次观测而得到不正确结果的机率。（2）提供测量过程中的信息。无论是在实验处理前，或是在实验处理后，均对被试者进行一系列的观测，这样就可以使我们看出发展趋向。

但是，问题还具有两面性，定时系列设计的这些优点，也加重了前、后设计所存在的弱势：（1）由于更多次的观测，势必延长实验时间，从而会有更多的外来影响。（2）也正是更多次的观测，更易引起顺序误差，更易导致练习、疲劳、紧张或厌烦等效应，影响实验结果。

（三）抵消实验条件的设计

抵消实验条件的设计（reversal experimental condition design）是指抵消实验过程中无关变量的一种设计。前面讲到，有些无关变量在某些实验情况下既不能被消除，又不能保持恒定。例如，单组实验往往由于前一处理影响后一处理的效果，产生顺序误差。为了抵消顺序误差，最简单的方法就是用 ABBA 的排列顺序来安排实验顺序。其模式可用表 2-3 表示。

从表 2-3 中可见，首先给被试者作第一种处理（ X_A ）和第一种处理后的观测（ Y_{2a} ），其次作第二种处理（ X_B ）和处理后的第二次观测（ Y_{2b} ），再重复第二种、最后重复第一种的处理和观测。整个实验程序的安排，就是抵消实验条件的设计的基本形式。

这里不妨举一个例子。有甲乙两种操作方式，我们要比较这两种操作方式哪一种更正确，更快些。我们知道，以人作为被试者的实验，在不同时间所得结果必然有所不同，甚至同一人以同一操作方式作一系列实验，要使其其中两次结果完全相同也是一件不容易的事。定时系列规则是解决此类问题的方法之一。但是定时系列规则不能消除和顺序有关的实验误差。

表 2-3 抵消实验条件设计的模式

处 理	X_A	X_B	X_B	X_A
后观测	Y_{2a}	Y_{2b}	Y_{2c}	Y_{2d}

为了消除这类影响，方法之一就是使用抵消实验条件的设计。本例就可用表 2-4 表示其模式。

表 2-4 抵消实验条件设计的实例

处理	操作方式甲	操作方式乙	操作方式乙	操作方式甲
结果	试验一之观测值	试验二之观测值	试验三之观测值	试验四之观测值

这种实验设计在处理 X_A 的情况下得到两个数值，在处理 X_B 的情况下也得到两个数值。统计时就通过分别求平均数后，加以比较。这样的实验设计，既安排了二种操作方式在程序上的均等，同时又提供了估计随机误差的可能

性。

抵消实验条件的设计具有三个明显的优点：(1)能较好地控制被试者变量，这是单组实验设计的共同优点。单组实验设计不但比两组实验设计用的被试者数量少，而且由于两种实验条件使用同一被试者，从而较好地控制了实验变量。(2)能较好地控制顺序误差。在心理实验中有许多变量与时间有关，例如学习、练习、迁移、挫折、疲劳等等。这些变量特别需要用这类实验设计来消除顺序误差。如果一半被试者所作的顺序为 ABBA，另一半被试者所作的顺序为 BAAB，那就抵消得更彻底。(3)时间上比较经济。这类实验一般观测次数不多，上面操作方式甲乙的比较，也只要通过四次观测就能完成某一被试者的实验。因此，被试者较易配合，以顺利完成实验，同时也排除了某一些实验误差。

抵消实验条件设计虽有上述优点，但也有其局限性。主要有两点：(1)反应变量在时间维度(轴)上的关系是线性时才能使用。抵消实验条件的设计的前提条件是反应变量在时间维度(轴)上的关系是线性关系。例如单纯的前影响作用，无论是从试验一到试验二，还是从试验三到试验四，都是后者受前者的影响。但是，当这些变量的实际效果与行为效果的关系是非线性时，则采用这类设计也不能达到抵消实验的顺序误差的目的。(2)对有些实验不适用。例如用两种学习方法学习同一实验材料就不适用。若一定要用这种设计，就必须改用同一被试用两种学习方法，学习两种难度相等的不同材料，才能消除练习误差，同时还应用相同的计量单位，否则无法进行比较。以上我们介绍了三类单组实验设计，清楚了这些设计所具有的优点。在心理实验中确实有不少运用这种设计方法达到较好效果的例子。例如：杨治良等(1979)曾用单组实验前后设计，对痛的成分进行心理学研究。采用这种设计的出发点是愿接受痛刺激的被试者不可多得，为了弥补单组实验和被试者数量较少的不足，实验采用单组实验前后设计。在这样的设计中，被试者变量能得到较好的控制，因为能消除由被试差异引起的误差，从而对自变量的效果作出更精确的估计。同时，为尽可能减少单组实验中的被试者在实验中花费很多时间，实验采用了心理量表法记分，通过被试者一次反应，记录较多的反应指标，而这些指标之间又存在着很大的一致性(见表2-5)。从表2-5可见，众多的生理和心理指标增加了对照程度，在一定程度上弥补了被试者较少的缺陷。

表 2-5 痛反应量表

级 别		一	二	三	四	五	六		
刺 激 量	t 值	1	2	5.6	8.3	13.9	19.5	2'	
	mA	0.37*	0.74	2.06	3.06	5.13	7.22	1C	
动 作 单 位	兴奋神经 纤维类型	无或少量 小幅度 A A	增大的 A	接近最大的 A	最大 A 小幅度传 导快的 A	最大 A 增大的 A	最大的 A 较大幅度传 导慢的 A		
	传导速度范围 (35)(m/s)	42.8 ~ 55.9	40.0 ~ 55.9	41.5 ~ 56.8	34 ~ 62.4 20 ~ 33.4	35 ~ 56.8 21 ~ 32.9	34.5 ~ 59.4 13 ~ 38.8		
感觉成分(主诉)		触	轻麻	重麻	轻痛	中痛	重痛	极	
情 绪 成 分	主 诉	无所谓	无所谓	无所为	稍有不适	难过、害怕	非常难过 非常害怕	恐 拒	
	外 表 行 为	语言	自如	自如	自如	语调稍变	轻声 叫痛	叫“喔唷”	连 大
		动作	协调	协调	欲动 能控制	头动、 手动	上身动、握 拳、转身	屈肌反射 全身动	
		表情	安详	安详	注意会神	轻皱眉 眨眼	争眉 紧闭目	咬牙 重皱眉	
出汗状况		无	无	无	身体发热	略有	手汗	大	

相当于各级刺激的强度为被试的平均值(采自杨治良, 1979)

应看到单组实验设计有其局限性,同时,正如前面所述,对有些实验不适用。单组实验设计在不能避免某种系统误差时,就要用被试者间设计或混合设计来完善实验设计。

二、被试者间设计

被试者内设计要求每一个被试者接受所有自变量处理,关键是如何安排各处理的先后次序。被试者间设计(或组间法)(between - subjects design)是要求每个被试者(组)只接受一个自变量处理,对另一被试者(组)进行另一种处理,故又称这种设计为独立组设计(independent - groups design)。这种设计的主要问题是决定哪一个被试者(组)接受哪一个实验处理。若有两种以上的处理,有多少种处理就采用多少个被试者(组)。被试者间设计或独立组设计有两类设计技术:随机组设计和配对组设计。

(一)随机组设计

随机组设计(或随机分组设计)(random - groups design)是将被试者随机地分配在不同的组内接受不同的自变量处理。这一假设是将被试者随机分配到不同的组,若对各组用一样的课题,在相同的条件下进行测量,其结果就成为相等组(或等组)(equivalent groups),则它们的成绩在统计上应是相等的。换句话说,假设各组在与实验课题有关的特性上(如年龄、智力、性格特征等)没有差别(在统计允许的限度以内),而实验结果却出现了差别,这差别就是由于处理的不同而引起的。

从统计理论来讲,特别是从抽样理论来说,此设计的各被试者样本(随机组)在未受不同的实验处理前,他们的作业平均数在统计上是没有显著差

异的，如果有，也仅仅是抽样变动。各自样本的平均数 \bar{X} 都是总体平均数 μ 的无偏估计值。换言之，各随机组在未经受不同处理之前是相等的（在统计允许的限度以内）。若这些随机组经受不同的实验处理后，经 t 检验和 F 检验后发现作业的平均数有显著性差异，那么这些差别是实验处理的不同引起的。各实验处理的平均数不是来自总体的平均数 μ ，而是来自各实验处理总体 $\mu_1, \mu_2, \dots, \mu_k$ 。这就是根据随机理论所作的逻辑推理。

怎样才能做到随机分组呢？常用的方法有两种：同时分配法和次第分配法。这两种方法各有其特点，可择宜采用。

1. 同时分配法 同时分配的条件是，被试者同时等候，而实验者可随意调派其中任何一个被试者。同时分配法通常有三种技术。

(1) 抽签法：先将所有被试者编号，记入纸片，每一纸片号码代表一个被试者，然后将纸片放入容器内搅匀，按组抽取。若要将 40 个被试者分为四组时，第一次抽 10 片，代表第一组，第二次也抽 10 片，代表第二组，依次类推。假定 40 个被试者中有 16 个女生，则每组应各有 4 名女生。所以可以在 16 名女生中先随机选 4 名，再在 24 名男生中随机选 6 名，归入第一组，依次类推。

(2) 笔划法：若要将 40 个被试者分为四组时，首先将被试者依其姓氏笔划数进行次序排列，再查随机数表每一数列的第一位数，只取第一个数为 1、2、3、和 4 的数字，分别归属四个组，各查 10 个，共查满 40 个以后，按姓氏笔划先后对入 1、2、3、4 所表示的组别。用此法时只要注意在查第一位数为 1、2、3、4 的随机数时各查 10 个即可。

(3) 报数法：若要将 40 个被试者分为四组时，采用类似于体育课上的报数，假定被试者都坐在教室内，实验者令其从第一排报数，报 1 的被试者都被分在第一组、报 2 的被试者都分在第二组，依此类推，只是要注意原有顺序的影响。若每排报数的方向随机改变，例如用 1234, 4321, 2341 等不同的顺序报数，则随机分组的效果将更好。

2. 次第分配法 次第分配法的条件是，由于实验持续时间较长或其他原因，实验者知道有一群被试者，但不知道究竟那位被试者什么时间会来，只能根据预先拟好的原则进行分派，而且当实验结束时，各组要符合随机组的要求。这里介绍两种技术。

(1) 简便法：按被试者出现在实验的先后分派，第一名属第一组，第二名属第二组，第三名属第三组，依此类推等。这有点像同时分配法中的第三例。一般而言，使用本法能满足随机的条件，然而它取决于被试者报到的次序是否符合随机原则。

(2) 区内随机法：为了避免被试者非随机出现的可能性，可按照被试者来到实验室的先后，使用区内随机次序分派被试者归属各组。例如，可根据随机数表来分配被试者。

随机组设计有优点也有缺点。其优点是：(1) 用随机分配被试者的方法可控制两组被试者变量的差异，分组方法简单可行。(2) 由于对每一被试者只作一次观测，可消除某些实验误差，如消除学习误差的影响。这种设计的缺点是：(1) 分成等组的方法仍欠精密。(2) 若两组在不同时期观测，就有可能插入实验以外的偶发事件，影响因变量的观测结果。

(二) 配对组设计

配对组设计（或对等组设计）(matched - groups design) 是随机组设

计的一种逻辑扩展。配对组设计的目的是使各组的特性更加相同。这种设计可控制组内变异和组间变异。在心理学研究中常常会遇到某些变量，特别是机体内部的变量，例如智力、态度……。它们对实验变量（自变量），例如对教材、教法等来说是额外的、无关的，而对因变量来说却是有关的。因此，它们是共变量。共变量若不受到实验控制，则进行 F 检验时作为实验误差的组内误差往往就不合理地扩大了。配对组设计就能解决此类问题。这一设计要求把共变量分成几个等级，经过测量，然后把具有同一等级特征的 K 个被试者加以配对，此时，每一个配对组便是一个层级。配对组的 K 个被试者每人只接受一种实验处理，至于谁应接受哪一种实验处理，则用随机分派的方法来决定。表 2 - 6 乃是配对组设计的基本模式。

表 2 - 6 配对组设计的基本模式

层级	实验变量					层级平均数
	A ₁	A ₂	A ₃	A _K	
1	X ₁₁	X ₁₂	X ₁₃	X _{1K}	$\overline{X_1}$
2	X ₂₁	X ₂₂	X ₂₃	X _{2K}	$\overline{X_2}$
3	X ₃₁	X ₃₂	X ₃₃	X _{3K}	$\overline{X_3}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	X _{n1}	X _{n2}	X _{n3}	X _{nK}	$\overline{X_n}$
实验处理 平均数	$\overline{X_{\cdot 1}}$	$\overline{X_{\cdot 2}}$	$\overline{X_{\cdot 3}}$	$\overline{X_{\cdot K}}$	$\overline{X_{\cdot \cdot}}$

层级	实验变量					层级平均数
	A ₁	A ₂	A ₃	A _K	
1	X ₁₁	X ₁₂	X ₁₃	X _{1K}	$\overline{X_1}$
2	X ₂₁	X ₂₂	X ₂₃	X _{2K}	$\overline{X_2}$
3	X ₃₁	X ₃₂	X ₃₃	X _{3K}	$\overline{X_3}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	X _{n1}	X _{n2}	X _{n3}	X _{nK}	$\overline{X_n}$
实验处理 平均数	$\overline{X_{\cdot 1}}$	$\overline{X_{\cdot 2}}$	$\overline{X_{\cdot 3}}$	$\overline{X_{\cdot K}}$	$\overline{X_{\cdot \cdot}}$

（采自杨治良、乐竟泓，1990）

在表 2 - 6 上，每一等级的每一处理只是观察一个被试者的做法，有时所观察的基本单位不是一个被试者，而是一个团体或一个子集合。

配对法常有两个步骤：第一，令所有被试者做“共同作业”，即接受预备测验，获得作业分数；第二，根据作业分数形成配对组。

第一步：共同作业

共同作业（common task）亦称为先检验（或前测）（pretest）。配对组设计的优劣完全依赖实验作业是否与共同作业有高度的相关。先检验与实验作业的相关越高，组间差异越小，则接受不同处理后的实验结果越能反映实验处理的差异。先检验作业有两种：一种是和实验作业有高度相关的其他作业；另一种是利用被试实验作业的初期表现。通常，同一作业两个阶段的行为表现是相关的。很多学习实验以智力测验为配对分组的依据，但在不能

找出和实验作业相关的作业时，对有些无练习效应的实验可使用实验练习的成绩作配对依据。

第二步：配对分组

配对分组 (matched groups) 乃是取得先检验测试分数后，再用两种方法形成配对组。

方法一：将被试者按先检验作业分数的高低排列 (见表 2-7)。现设有三种自变量处理 (情况)，将前三个被试者依区内随机方式派到 A、B、C 三组，接着将后三个被试者如法分配，直到分配完毕。从表 2-8 可以看到，

表 2-7 30 名被试者先检验分数的高低排列表

被试者	分数	被试者	分数	被试者	分数
1	75	11	40	21	27
2	65	12	39	22	25
3	64	13	37	23	21
4	63	14	37	24	19
5	60	15	34	25	19
6	45	16	32	26	16
7	45	17	31	27	15
8	42	18	30	28	13
9	41	19	30	29	10
10	41	20	30	30	2

(采自杨治良、乐竟泓，1990)

按这种方法分配的各组在平均数和标准差上都很接近。

方法二：取先检验分数中相同或非常接近的，以三人为一单位，然后按区内随机法逐次分派到 A、B、C 各组。每三个被试者的分数接近的程度完全是任意的，如表 2-9 的分派限定三个被试者分数差落入 4 的范围内。因为有了这样的限制，超过标准的被试者被剔除。如表中每组被试者只有 8 人，各组分数相差甚微。

表 2-8 30 名被试者按方法一配对分组的结果

被测者 分组	1	2	3	4	5	6	7	8	9	10	\bar{X}	S
A	65	63	42	41	34	31	30	19	19	2	34.6	18.5
B	64	60	41	39	37	32	27	25	15	11	35.0	16.5
C	75	45	45	40	37	30	30	21	16	13	35.2	17.1

说明： \bar{X} 为平均数，S为标准差 (采自杨治良、乐竟泓，1990)

表 2-9 30 名被试者按方法二配对分组的结果

被试者 分组	1	2	3	4	5	6	7	8			\bar{X}	S
A	63	45	41	37	34	30	19	13			35.6	15.7
B	64	45	41	39	31	30	21	15			35.8	15.2
C	65	42	40	37	32	30	19	16			35.1	15.3

(采自杨治良、乐竟泓，1990)

二种方法相比，方法二中各被试者的平均值比方法一更相近，但由于有随机分配的程序因素参与其中，只比较平均数还不能说明问题，应再比较所分各组的标准差。若被试者在方法二中的标准差比在方法一中的更接近，这就有力地说明了方法二的各组较为接近，而根据方法二的处理，各组的标准差总会小一些。

配对组设计的作用在于控制组内变异与组间变异。它的优点是在实验处理之前，就把组间变异缩到最小和要求两组组内变异比单独的随机分配更接近相等。因此，这种设计能对被试者个别差异给予很多的控制，小型实验用配对设计，其效果比用随机分组的效果更为显著。但这种设计的缺点是，实验者因分配被试者而大大增加其工作量。

三、混合设计

混合设计 (mixed design) 乃是在一个实验中同时采用两种基本设计的实验设计。它要求一个自变量用一种设计处理，如被试者内设计处理，而另一个自变量用不同种类的设计处理，如被试者间设计处理。这时，实际上是在进行两个实验。如果某个实验要求处理三个自变量，可按同样的原理对三个自变量采用不同的设计。注意，这里谈及的三个自变量不同于一个自变量的同类的三种水平。这里举一个例子，若将实验限定在两个自变量的范围以内，在最简单的双变量实验设计 (2×2) 中，每个自变量在二种水平 (A 变量有 A_1 、 A_2 ，B 变量有 B_1 、 B_2) 上的混合设计有如下几种可能：

1. 表 2 - 10 中的 1、6、11、16 组是非混合设计，而 2、5 组代表同种混合设计，3、9 组也代表同种混合设计，等等。真正的混合设计有十二组。其实只要用对角线的一半编号组就可以说明问题，真正的混合组有二个六组，要么是 2、3、4、7、8、12，要么是 5、9、10、13、14、15。这两个六组只在接受处理的方向上不同，而接受处理的内容基本一致。

2. 现再以 2 组为例，展开随机组和配对组 2×2 混合设计。步骤一：将全部被试者分成二个随机组，一组接受变量 A_1 ，另一组接受 A_2 ；步骤二：将 A_1 组被试者分解为二个配对组 (G_1 、 G_2)，将 A_2 组也分解为配对

表 2 - 10 混合设计一例

		自变量 A			
		随机组	匹配组	不完全的被试者内设计	完全的被试者内设计
自变量 B	随机组	1	2	3*	4
	匹配组	5	6	7	8
	不完全的被试者内设计	9*	10	11	12
	完全的被试者内设计	13	14	15	16

说明：表中数字编号是任意的，和*代表同种混合。

(采自杨治良、乐竟泓，1990)

组 (G'_1 、 G'_2)；步骤三： A_1 组中的两个配对组接受变量 B_1 情况；另一组接受 B_2 情况，在 A_2 组中的二个配对小组也分别接受变量的 B_1 、 B_2 情况。上述情况可归纳为表 2 - 11。

3.再以 9 组为例，第一部是构成两个随机组，一组是接受变量 A_1 情况，另一组接受 A_2 情况。另一变项 B 的二个情况 B_1 、 B_2 以任何方式给 A 变项的二组中所有的被试者，构成被试者间和被试者内混合设计，从而把全部被试者的材料总合起来的时候，渐近误差将被抵消。

2×2 实验仅是混合设计的简要形式，若有需要，完全可以实现 2×4 、 3×3 、 3×6 等等的设计。例如，作 2×4 的随机组与被试者内混合设计可用表 2 - 12 的矩阵来表示，表中的数据是假想观察值。在决定一个实验是否用混合设计时，我们有必要在两种水平考虑：首先看看所涉及的有效变量是

表 2 - 11 2×2 混合设计

		自变量 B		
			B_1	B_2
自变量 A	A_1	G_1	接受	
		G_2		接受
	A_2	G_1	接受	
		G_2		接受

(采自杨治良、乐竟泓，1990)

表 2-12 2×4 混合设计

			自变量 B				X
			B ₁	B ₂	B ₃	B ₄	
自 变 量 A	A ₁	S ₁	0	0	5	3	8
		S ₂	3	1	5	4	13
		S ₃	4	3	6	2	15
	A ₂	S ₄	4	2	7	8	21
		S ₅	5	4	6	6	21
		S ₆	7	5	8	9	29
X		6/S	23	15	37	32	X=107

(采自杨治良，乐竟泓，1990)

否要求特殊设计，如指导语变量要求被试者间设计；在这一点上没问题时，就得从方便、经济和注意统计正确性的角度作进一步的考虑。

第二节 多变量实验技术

在上一节中，我们讨论了实验设计的基本类型，在以人或动物为研究对象的心理学实验中，首先要解决的问题是：实验处理中用了相同的被试者还是用了不同的被试者。根据这一原则，上一节中我们讨论了被试者内设计、被试者间设计和混合设计。尽管这样的分类也穷尽了各种实验，但在心理学实验中，大多数比较重要的实验都是多自变量实验。所以在这一节里，我们对多自变量的实验设计作进一步介绍。

一、多自变量实验的优点

多自变量（或多自变项）（multiple independent variable）是指在一个实验中包含有两个或两个以上的自变量。它不是指同一自变量的多个水平。在心理学的刊物上较少发现只用一个自变量的实验研究报告。一个复杂的心理现象很少是单一刺激因素决定的，往往是多因素的结果。例如，想研究不同的教学效果，可能有教学方法的影响，也可能有教师水平的影响，还有学生自身水平的影响等。这种教学效果，实际上是受多因素交互影响的结果，要使实验更接近实际，就应采取多因素，即多自变量的实验设计。典型的实验是同时操纵两到四个自变量。

那么为什么在心理学实验中在较多的情况下要采用多自变量实验呢？这是因为多自变量实验具有三个明显的优点：

1. 效率高 有两个自变量的实验要比分别做两个只有单一自变量的实验效率要高，也就是说事半功倍，花同样的时间，做了一倍，甚至二倍、三倍的工作。

2. 实验控制较好 做一个实验时某些控制变量比进行两个实验时更易于控制和恒定。如被试者条件是一样的，同一个被试者，同样的身心状况。再如时间条件也是一样的，日期相同、时间相同。其他条件就更不用说了，如相同的外界环境、相同的温度、相同的湿度等等。这就在很大程度上排除了许多实验误差，减少了实验污染。

3. 实验结果更有价值 有多种自变量的实验所得的结果，由于在多种情况下都证明是确实的，这样就比多个单独实验所概括的结果更有价值。例如，我们想要知道两种奖励办法中哪一种促进了中学生学习体育。第一种奖励办法是对正确完成规定的运动项目的学生给予物质奖励；第二种是给予提前下课，即每正确完成一个规定的运动项目允许学生提前下课，哪种更好些需要用实验证实。而且，在把它作为一项中学的规定之前，我们还应把这两种办法放到语文、数学等课程中去实验。在这里，不同课程的被试者是第二个自变量。显然，把奖励办法与不同的课程这两个自变量结合在一项实验中，要比进行两个连续的单自变量的实验为好。特别是，当一个自变量产生的效果在第二个自变量的每一水平上不一样时，交互作用（interaction）就发生了。在有交互作用的情况下，分别讨论每一自变量的效应就不够了。因为一个自变量的效应依赖于另一个自变量的水平。这里，我们举个例子。杨治良等人（1981）做了一个实验，目的是想了解年龄对再认能力的影响。第一个自变量是年龄，这里我们选取初中生年龄组和大学生年龄组。第二个自变量是实验材料，这里我们选取具体实物图形组和词组两个组。因变量取再认能力 d' 作指标（信号检测论方法用于再认实验，采用 d' 作为再认能力的指标）这个实验的方法是采用再认法，把被试者识记过的材料和没有识记过的材料混在

一起，要求被试者把两种材料区分开来。实验结果如图 2-1 所示（见下页）。

从图 2-1 可见：第一，在哪个实验条件下，初中生的再认能力较强；第二，对具体图画的材料，再认能力较强。然而，此图是将两个变量（年龄大小和材料性质）分开绘制的。在左图，初中生的再认能力指标中，既包括了对具体图画的材料，又包括了对词的材料。大学生的情况也是一样。

如果把同样一个实验结果，换另一种方式绘图（如图 2-2），我们就可以看到，实际情况并非这样简单。图 2-2 将两个变量画在一起，显示了交互作用的发生。它告诉我们：第一，在此实验条件下，再认能力较低的情况仅仅是在以具体图画作材料的大学生组发生；第二，无论是对具体图画材料，还是词材料，初中学生都表现出较高的再认能力，比较图 2-1 和图 2-2，我们看到了区别。具体地说，大学生对词材料的再认能力，在二张图上有不同的分析。图 2-2 能明确指出一个自变量的各水平受到另一个自变量水平的不同影响的交互作用就代表了这种不同的影响。

为进一步分析多自变量的优越性，我们也可以设想，两个自变量没有显出交互作用。图 2-3 就是虚构的情况。从图 2-3 上可以看到，任何一个自变量的效果对于另一个自变量的两个水平来说都是相同的。也就是说，第一，初中学生总是表现出较高的再认能力，不论实验材料是图画还是词。第二，对于图画材料，不管是初中学生还是大学生都表现出较高的再认能力。这就是图 2-3 说明的问题。平行线意味着没有交互作用的发生。

实验的实际情况并不是这样。实际情况如图 2-2 所示，再认能力的高低既依赖于年龄大小，也依赖于实验材料；一个自变量的效应依赖于另一个自变量水平。

上述三种情况的图解分析揭示了多自变量实验能揭示变量间是否存在交互作用。这是单自变量实验所不能完成的，我们想象这个实验是分两部分进行的。第一部分实验中，只有年龄大小被当作唯一的自变量。那么，实验材料就成了控制变量。我们知道，控制变量在实验过程保持恒定。如果选择的是图画作实验材料，实验结果将表明，初中学生比大学生表现出较高的再认能力。但是，实验者将不会知道，如果用词作材料会产生什么结果？在第二部分实验中，只有实验材料被当作唯一的自变量。年龄大小就成了控制变量。若用初中生做被试者，将揭示出不同的结果。我们可以看到，实验虽然做了两个，但得到的信息反而比一个双自变量实验少。二个单自变量实验所获得的信息反而比一个双自变量实验少。

图 2-2 是实际的实验结果，它表明了交互作用。一个自变量的效果依赖于另一个自变量的水平。具体地说：第一，若用词作材料，初中生和大学生的年龄因素对再认能力的影响较少；第二，若用图画作材料，初中生和大学生的年龄因素对再认能力的影响较大；第三，若用初中生作被试者，实验材料的因素关系不大；第四，若用大学生作被试者，实验材料的因素关系就很大。这里我们不难发现，一个双自变量实验实际上等于四个单自变量的实验，而且还要看单自变量实验是否做得好。

综上所述，当一个自变量的水平受到另一个自变量水平的不同影响时，交互作用就发生了。在有交互作用的情况下，分别讨论每一个自变量的效应就不够了，因为一个自变量的效应依赖于另一个自变量的水平，因而在实验结果的分析讨论中，除了对单独因素的分析之外，还必须分析讨论出现交互作用的原因和后果。

多自变量的实验是效率高、价值高的实验。在这些实验中，每一种因变量的测量至少能提供两个变量信息，从而有可能研究变量之间所有可能发生的交互影响。在实验中自变量数目增加时，能发生交互影响的数目也迅速增加。两个自变量只有一个交互作用影响的可能。用三个自变量（自变量 A、B、C）就有四种交互作用影响的可能，如 AB、AC、BC、ABC 四种。假定每一自变量有两个水平，三种自变量的交互影响由四条线生动地表现出来；两条表示变量 B、两条表示变量 C，与横坐标变量 A 起交互作用，见图 2-4，它们的交互作用由四条非平行线表示出来。

二、多因素实验设计

多因素实验设计 (multifactors experimental design) 是指在同一个人实验里可以同时观测两个或两个以上的自变量的影响，以及自变量与自变量交互作用的效果的实验设计。它与上述所讨论的只用一个自变量的实验设计是不同的。包括 X_a 和 X_b 两个自变量的设计，叫做双向（或二向）析因设计，简称为 $A \times B$ 因素设计。包括 X_a 、 X_b 和 X_c 三个自变量的设计叫作三向析因设计，简称为 $A \times B \times C$ 因素设计。在因素设计中，每个因素（自变量）又可以包括几个水平。若两个自变量 X_a 和 X_b 各有两个水平，则可称为 2×2 设计；若自变量 X_a 有两个水平， X_b 有三个水平，则称为 2×3 设计。例如，有人想研究大学生对红、绿、黄三种灯光的反应是否与灯光的强度有关，在这项实验中，A 代表灯光强度，是一种因素（自变量 X_a ），包括强和弱两种水平；B 代表灯光频率，是另一个因素（自变量 X_b ），包括红、绿、黄三种水平。这种二向析因设计就是 2×3 因素设计。若它还想同时研究反应时是否有性别差别，则 C 代表性别（自变量 X_c ），分男女两个水平。这种三向析因设计就是 $2 \times 3 \times 2$ 因素设计。当然也可以有多于三个因素的设计。不过，因素多，解释结果的困难就大，交互作用也就多，因而研究者一般是将研究限于 2 至 4 个变量。在析因设计中，研究者要操作的实验处理的个数就是各自变量的水平个数的乘积。例如，在一项二因素实验设计中，设 A 因素有 p 个水平，B 因素有 q 个水平，则研究者就应操作 $p \times q$ 个实验处理。在确定了实验中所要操作的实验处理后，就要决定怎样分配被试者了。下面我们举一个研究实例。

杨治良等人（1993）在研究汉字内隐记忆的实验研究中，通过四个 $2 \times 2 \times 2$ 方差设计实验，试图寻找汉字认知范畴里内隐记忆存在的条件。此方差设计的三个因素为：（1）A 因素，即注意程度因素，分目标字和非目标字两个水平；（2）B 因素，分直接测量和间接测量两个水平，直接测量为再认判断任务，间接测量则是汉字偏好判断任务；（3）C 因素，即时程因素，分单元一和单元二两个水平，参见图 2-5。结果发现了内隐记忆存在的三个必要条件是：非目标汉字、汉字的整体加工和偏好判断任务。此例说明了，正是这些多因素实验设计，才发现了汉字内隐记忆存在的条件。

多因素实验设计可以概括成某一个基本形式。假设实验因子有 A、B 这两个，A 因子有三个水平，B 因子有四个水平，那么就有 $3 \times 4 = 12$ 种组合，每一个组合称作一种处理，现随机抽取被试者 1、2、3、……24，共 24 个分别安排在各种处理中接受实验，实验设计列表如表 2-13。

表 2-13 多因子设计表

实验因子		B 因子			
		B ₁	B ₂	B ₃	B ₄
A 因 子	A ₁	1,2	3,4	5,6	7,8
	A ₂	1,10	11,12	13,14	15,16
	A ₃	17,18	19,20	21,22	23,24

(采自杨治良、乐竟泓, 1990)

实验所得的数据,可表明哪种处理(即水平的结合)效果比较好。但到底多种处理的效果是否有显著差别,还要作进一步的统计检验。

三、拉丁方设计

拉丁方设计(或拉丁方格设计)(Latin - square design)是多变量实验设计中一种较为常用的设计方案。心理实验中采用循环法平衡实验顺序对实验结果的影响,就使实验顺序、被试者差异都作为一个自变量来处理。只要是实验中自变量的个数(因素)与实验处理水平数相同,而且这些自变量之间没有交互作用的存在时,都可采用拉丁方设计方案。这里对于这些因素之间没有交互作用的假设是很重要的。否则,若按没有交互作用的统计方法处理实验结果,只能是准实验设计。此设计的基本模式如表 2-14。

表 2-14 拉丁方设计

被试者	实验顺序		
	B ₁	B ₂	B ₃
C ₁	A ₁	A ₂	A ₃
C ₂	A ₂	A ₃	A ₁
C ₃	A ₃	A ₁	A ₂

(采自金志成, 1991)

表 2-14 中 A₁、A₂、A₃ 为实验处理的三个水平。C₁、C₂、C₃ 是被试者的三种不同类型,存在个体差异,被试者可为一人,也可为多人。对于这种拉丁方设计的实验结果,可用独立样本拉丁方变异数分析法,可分析出被试者间、实验顺序间、实验处理间的差异情形。就是说实验顺序、被试者差异对实验结果的影响都可分析出来。

假如,被试者是随机抽取的,且随机分派,并保证各组被试者相等。这时的结果分析可用相关样本拉丁方分析法来处理。可分析出被试者内的顺序(B因素)、处理(A因素)及部分 A×B 的交互作用。

拉丁方设计能够抵消实验中因实验顺序、被试者差异等造成的无关变量效果,因此在心理实验中经常被应用。随着数理统计的发展,多因素设计的方案越来越多,多因素的实验设计执行起来很费劲,被试者的数量要多,结果的统计处理也较复杂,因此,选用哪一种实验设计,要因地制宜,以做到恰到好处。