

中国教育管理全集

样长治校篇 (五)

王亮 主编

飞天电子音像出版社

目 录

教育实验法.....	1
行动研究法.....	53
文献法.....	110
内容分析法.....	145
调查访问法.....	162
校长的科研素质培训.....	190
科研型校长的基本要求.....	202

教育实验法

教育活动是人类的基本活动之一，教育实验的主要目的是研究如何提高“全人”发展的效率探索其客观规律。在教育实验中，主要在于研究、测定教育条件、环境、方法等与人的成长之间的关系。研究对象都为人类，因而具有很多特点，影响研究结果的因素亦很复杂。这一点与生理学、物理学及化学等实验明显不同。了解这些复杂的因素和特点十分必要，对于研究设计、解释和分析研究结果会有很大帮助。

1. 自变量及其控制

(1) 自变量 (Independent variable) 概念：又称独立变量，它是由研究者选定，并进行操纵、变化的能产生所欲研究的教育现象的因素，或因素的组合。作为研究者，操纵变化这些因素是想了解这些因素的不同变化，对人的教育或行为产生哪些影响。一般情况下，教育研究中的自变量常是不连续的，称为因素，同一因素的不同标准称为水平 (Levels)，又叫实验处理。自变量由研究者操纵变化，会引起受教育者（又称被试者）反应的变化，但需指出，在没有足够的证据之前，不能简单地认为自变量是因，受教育者的反应是果。它依据自变量的变化而变化，但不能简单地

认为它是自变量的果。

(2) 自变量的种类 1) 课题方面的自变量

这些自变量可为量的连续变化，也有属于质的不连续变化。例如，研究不同照度下学生视力的变化，这里的自变量就是不同照度（刺激光亮强度的一种表示单位），是量的变化。而不同学习方法和教学方法的学习、教学效果（常用一定的指标，如成绩、记忆量、学习兴趣、学习动机、自信心等表示）等研究中，自变量就属于质的变化，因不同的教学方法和学习方法难于用一个连续的量加以表示。

这些自变量可为简单的刺激，也可为各种条件组合的刺激，例如，讲课教师的教学经验、师生关系、教师的教态、年龄、相貌等，就与教学方法的操作混合在一起。

有些不同的刺激可作用于同一感觉道，或作用于不同的感觉道，这时其作用不同，因而作为自变量的性质也不同。⁵ 这些变量可在强度上和作用时间上不同，便构成不同的自变量及其水平。上述各种可以操纵的量和质的变化的刺激，一般称为课题变量。

2) 环境变量：不同的环境因素也可以作为自变量。例如，不同学校、不同地区、不同班级等环境不同的被研究对象，有些教育现象之间也可能不同，都

可作为研究内容。

3) 被试变量：即在外界条件一致的情况下，被试间不同程度的持续性特征，如年龄、性别、血型、不同文化背景、职业特点、健康状况、个体差异等都可作为自变量因素。但在大多数的研究中，这些被试变量常作为影响研究结果的无关因素（即控制变量）或称影响因素，加以控制。

4) 暂时的被试变量：即通过主试者或研究者的言语、态度以及用某些方法使被试的特性、机能状态等方面产生一时变化，这类自变量称之为暂时的被试变量。它们只在很少的情况下作为自变量因素加以操纵变化，而在更多的情况下作为无关变量因素（即控制变量）加以严格控制。被试不同的机能状态，构成对反应的准备状态不同，对研究结果的影响不同。这些因素几乎在所有教育研究中都会出现。它先于某种反应而变化并且影响所要研究的结果，因此，这类变量又称之为先行变量，有时写作 A 变量。教育研究中，常将反应函数写作 $R=f(O \times S)$ 或 $R=f(A \times S)$ 。这里的 O, A 即是被试变量，R 为反应。可见暂时的被试变量，即可作为自变量，又是影响实验结果的一个很特殊的变量，这是其他学科的研究中没有的一种变量。这种变量的存在，导致教育研究的复杂性，常会

造成研究中自变量的转移---或称实验污染。这类自变量还包括以下一些因素（赫尔的研究）：

习惯强度，指一定的反应和一定的刺激之间的联系强度，这种联系强度以从前的学习为基础，例同一个文字、词汇对不同的人联系程度是不同的；

内驱力，例如饥饿、本能要求等生理教育中所讨论的那些本体要求；

诱因，即预期可得到的奖励或惩罚，人类被试对实验结果的立时效应都属此类；5 抑制，使反应的当时准备状态削弱的一种因素或因素组合。如疲劳、满足、干扰、恐惧、谨慎等。上述的分析可见，自变量有些是有形的，有些是无形的，表面上看有些好像是刺激的缺乏，实际上刺激是存在的，例如实验教师的个人特征。尤其是被试变量及暂时的被试变量，在教育研究中，除个别研究作为自变量加以操作控制外，几乎所有的研究中，包括前述将其一小部分作为自变量的研究，都要对其进行严格的控制。否则会影响实验结果的准确程度，不利于探讨所要研究的问题。这一点要特别引起研究者的注意。

（3）自变量的控制：指在研究中对刺激变量的操纵变化。对自变量控制得如何，直接影响研究的成功与失败，控制包括以下几方面：1）给自变量规定

操作定义：实验中规定操作定义有以下三个好处：

可以使研究者的思考具体而清晰，例如“疲劳对识记的影响”研究课题中，“疲劳”一词定义很多，人人都有疲劳的体验，但要具体量化，却不是一件容易的事，只能规定操作定义，有了操作定义，才能使其具体操作，如规定“疲劳”是从事某种体力运动之时间数，则可用从事某体力运动一小时、二小时、三小时……作为疲劳的不同程度的指标，这一规定可使研究者对“疲劳”这一自变量的定义具体而清晰。“疲劳”这一自变量便可以操纵控制了。

操作定义可以增进科学研究者之间沟通的正确性。因为只有操作定义规定的概念或变量，才不易于产生误解。没有操作定义必使很多概念的控制含糊不清，这也必然要产生误解。

操作定义可减少一门学科所用概念或变量的数目。因为只有在操作定义明显不同时，才增加新概念或新变量。这里所讲的操作定义是指在定义一个概念时，不直接描述被定义事物的特征、性质，而是说明观察或测量被定义事物或变量所做的实际活动。但需指出，在对自变量规定使用操作定义时，要特别防止操作定义单一性偏差，影响研究的构想效度。例如，上例“疲劳”的操作定义是从事某体力活动的时间长

度，但是若研究的问题是大“疲劳”概念，这样疲劳的内涵，除体力疲劳外，还应包括脑力疲劳、心理疲劳，这种情况下，只从体力方面规定了操作定义，势必是不完全的。操作定义的片面性必然会带来研究效度降低的问题，这一点研究者必须注意。

2) 规定自变量的不同水平---检查点或实验处理：教育研究中的自变量，有些有质的区别，是连续的，例如，不同的教学方法，不同的学习方法，不同民族、性质等，这些不同质，便是不同的实验检查点，或称为实验处理。如果自变量是连续的量的变化，例如学生做习题的数量、补课的时间，每节课记忆的单词数等，则要求所选自变量的不同水平值---检查点，至少要三个，一般三至五个左右为宜，只有这样，才能足以找出自变量和因变量的函数关系，如果是两个检查点，则难以断定自变量与变量之间的关系是线性还是非线性，如果自变量与因变量之间的关系复杂，至少要选五个或五个以上的检查点。在规定检查点时，如何确定其范围？如何确定其间的间距亦是很重要的问题。确定自变量间距的原则是：两个不同的检查点能引起被试不同的反应。间距的大小可依具体的研究情况而定，一般要以已有的研究资料作参考，若无资料依据，则应通过预备试验来确定。自变量的间距可

以用算术单位，也可以用对数单位，这也要根据具体的研究需要确定。如果自变量范围不能事先确定，则要根据被试的反应，当其反应比较平稳时才能停止实验。

2. 因变量及其控制

(1) 因变量的种类：因变量又称之为反应变量（用符号 R 表示），因其主要是根据受测者在某种刺激作用下产生的反应并对其进行度量。反应变量是由自变量所引起的被测者某些教育行为的变化。这里需明确，因变量一定是由自变量引起，在教育研究中这点并不易分辨或觉察。很多情形下，某些教育现象或行为的变化，可能根本与自变量无任何关系，但由于时间上的巧合，或者是自变量引起某中间因素的变化，而间接地影响因变量的变化。例如，某年级某些学生的学习成绩，由于两位不同教师的讲课方法不同而不同。这一成绩，表面上看可以作为因变量，但实际上，两位教师的讲课方法并不必然是自变量，很可能是两位教师各自所用的不同讲课方法是讲两门不同的课程，或两位教师的仪态不同，或教学经验不同等等。不同的课程对学生来讲可能认识不同，对之学习兴趣可能也不同。实际上学生的学习成绩不同，可能是由于对不同课程的学习兴趣不一样所引起的，而

与教师的讲课方法无关。因此，学习成绩并不能作为两种讲课方法的因变量，如果误认为是因变量，并据此而进行推论，必然造成研究效度极低的情况。若两位教师所讲的课程相同，也可能是两位教师的仪态不同而引起学生的学习态度、兴趣等不同而造成学习成绩不同的。从这一点来看，判断被测者的反应是否由自变量引起，是一件并不轻松的工作。需要理论的支持、有关研究资料的佐证，以及一些统计方法的帮助。因变量的种类，对于教育实验研究来说，很多亦很复杂。一般来讲，作为反应变量的指标有：

准确性方面的指标，正确率或错误率。若相同条件下（除自变量之外），亦可用正确或错误的次数表示。

速度或敏捷度方面，一般来讲，内部的教育历程越短，速度越快。常有两种情况：工作量一定，完成工作所需要的时间；时间一定所完成的工作量。这两种指标都可作为速度或敏捷度指标。

刺激的强度水平不同，也可以作为反应的指标。例如，学习实验中，通过不同难度的智力测验题目。

5 概率或频率，某些反应出现的概率或频率，也常被用作反应变量的指标。⑤反应的强度或力量。⑥各种成就测验与心理测验的量表分数及评价者的评定分

数，这虽然是人为的主观评价，但若依据的评价标准可靠，这些评价分数是可靠的。⑦高次反应变量，即用一个图或表表示反应的多种情况，例如学习曲线，既可表示学习的正确或错误率，又能表示整个的学习进程情况。在一项研究中有时用一种反应变量的指标，有时用多种反应变量的指标，共同标明反应，还有时可用其再生指标——称为衍变量表示。究竟应该如何选用恰当的反应变量指标，这要根据研究的需要及有关的研究经验确定。

(2) 因变量的控制：对因变量进行控制，就是要达到反应变量确实能够反映受测者真实的身心反应。

为此，在选定反应变量指标时要做到以下几点：

1) 规定好反应的操作定义：作为教育现象或行为变化，往往都是含义不太清楚的问题。例如反应的快慢，或好与不好等等。为要使其能够具体度量，就要规定一定的标准，这样才能加以数量化。为使反应能达到具体度量，所规定的一系列标准或具体指标，就是反应的操作定义。反应的操作定义，要特别注意指标的单一偏差，因为这种单一方法的偏差，严重影响研究的构想效度，这意味着测量、记录方法单一或考虑不周全。也就是说，应该用两种指标度量反应的，

而研究中只采用了一种方法，这就势必造成单一方法的偏见。例如，常用学习成绩一种指标反映教育效果，就是一例，教育研究中反映教育效果的重要指标---学生学习兴趣、动机的提升，几乎被忽略。

2)因变量应具备的特点·反应指标要具有有效性，要能真实地反映反应的情况。即反应的指标能够度量某刺激所引起的真实反应，而不是其他别的什么，故这种特点又称为特异性。·反应指标要具有客观性，即不同的研究观察者或不同次的观察只要是反应相同，其指标也应该相同。当然允许观测的随机误差。一些主观性较大的度量指标，规定好反应标准以后，也可认为是一种客观性很好的度量指标。·指标要能够数量化，或用数字，或用次数，或用一定的反应不同程度的语义，例如好、很好、非常好之类，也可认为是具有数量化特征的指标。·反应指标应能准确、真实地度量反应的变化。在应用百分数表示正确或错误时，应特别注意天花板、地板效应。这一效应即指因反应指标的量程不够大，而造成的对接近两极端的反应不能区分的情况。

例如，同是 100% 的正确，或同是 100 分，各个受测者在此方面的反应并不尽相同，有的很吃力，有的很轻松，有的勉勉强强，有的绰绰有余，但作为记

录的反应只是一种，这就是天花板效应。地板效应也类似，不过是反应的另一端。

3) 反应指标的平衡：教育研究中的反应变量有上述多种。这些指标中的两种或多种，有时在一种反应中同时都出现。例如，反应速度要快，最易出现反应的正确率下降，错误率上升的问题，同样反应中若强调准确性，就可能出现反应速度减慢的现象，这一类问题怎样解决，怎样平衡这两种反应指标之间的关系，尚需进一步研究。尽管已有的研究结果揭示，时间（速度）指标，可能比错误率或正确率能更好地反映实际情况。这一点在因变量的控制中是很有意义又很困难的事情，故研究者根据自己的研究应花费一定的气力解决这一问题。因变量的选择与确定也会影响研究的构想效度，因此须要下大力气去斟酌。

3. 无关变量的控制

(1) 无关变量的含义：无关变量，又称控制变量、参变量、额外变量等，这是由于描述问题的角度不同引起的。一是说：除自变量之外，一切能够影响实验结果，而实验中需要加以控制的变量；另一是说：一切与所研究的条件和行为无关，但在实验中又是影响反应变量（行为，又称实验结果）的因素。可见二者所指的是同一个内容。但这两个名称又都不是那么

贴切，在没有更好的名称之前，姑且用“无关变量”。

(2) 无关变量的鉴别：在一个教育实验中，确定哪些因素属于无关变量，最好是通过查阅文献或用因素型实验来确定，但是这很难做到。怎么办呢？只能根据已有的理论、知识和经验，来分析、确定一个教育实验中该控制的无关变量是哪些，其中有哪些必须控制，又有哪些没有必要控制。例如，在知觉实验中，被试的社会关系，实验前一天吃的什么，有些什么活动等就很难进行控制，实际上也没有必要进行控制，因而它就不能称作无关变量。在一个实验中，根据实验者的分析所确定的无关变量，有些是科学的、正确的，有些是非科学的（不正确的）。那么，实验中对无关变量的控制是否成功，无关变量的确立是否正确，能不能鉴别呢？回答是肯定的。一般可以由实验结果误差分散的大小来鉴别。实验误差分散小时，实验者的分析正确，可以认为实验被充分地控制了。相反，虽然充分地控制了自变量和反应，而且反应和度量也认为是妥当而灵敏的，但结果的误差分散较大，就有理由认为，有些无关变量没有得到很好的控制。遇到这种情况，就要认真分析无关的因素是什么。如果经过认真分析，有些无关变量还是难以作为记述因素抓住时，就中止实验，从别的角度设计实验，或

等到找出控制这些变量的新方法和技术时再实验。

(3) 关于无关变量的控制：对无关变量控制得如何，直接影响研究的效度，因此，在研究设计中必须认真考虑。

1) 对部分被试变量及暂时被试变量的控制方法：被试变量是指外界条件一致的情况下，被试间不同程度的持续性特征，例如，年龄、性别、民族、文化、职业以及其他较稳定的个体差异。暂时的被试变量是指非持续性的被试机能状态，例如，疲劳、兴奋水平、诱因、抑制等。这些变量不作为自变量时，也会影响被试的反应，应作为无关变量加以控制，具体方法如下：用指导语控制。指导语在实验中的作用很大，它使被试者明确如何进行反应，因此，指导语的明白程度，指导语中是否含有暗示的成分等，是会直接影响反应的。为使被试的反应一致，必须对指导语标准化。这里所说的标准化是指所用词语，主试者的表情、语气、声调等方面都要进行标准化，即对所有的被试，指导语都应一样。主试者对待被试者的态度应予规范化。即应热情，有礼貌，尊重人，但这方面又须有度，以避免产生新的影响实验结果的变异。双盲实验法，即在教育实验中，令主试者也不知道试验的内容，对实验内容来讲主试是一位盲者，而教育实验中的被

试，一般情况下对实验内容总是一位盲者，故称为双盲实验。

在一些具有明显主试者效应的教育实验中，主试者会影响被试者的反应，故应采用双盲法控制被试。控制被试者的个体差异。在选择被试样本时，要按被试个体差异的不同层次，对被试进行划分，然后进行随机取样，以保证被试样本的代表性。同时在研究条件允许的情况下，尽量加大样本容量，也是控制被试变量的一个方面。实验组、控制组法。为控制被试个体方面的无关变量因素的影响，使两组被试者个体差异在各方面相等，一般常采用随机取样、随机分组的方式进行取样与分组，这样可保证实验组与控制组被试尽量相等。

2) 对环境变量及部分被试变量等的控制方法：主要包括对不作为自变量的环境方面的因素、实验执行中所产生的无关变异因素，以及部分被试变量或暂时被试变量因素的控制。方法主要有以下几个方面：无关变量的消除：对无关变量的控制，最简单的方法是消除无关变量，使实验在“单纯的”条件下进行。教育实验在微格教室、隔音室里进行，就是为了消除作为无关变量的视觉刺激和听觉刺激。消除无关变量的方法，一般情况下较简单易行。但有些问题须

注意：一方面，在消除无关变量的时候，实验条件开始失去"真实性"，离日常状况较远，而变得"抽象"了。另一方面，在消除无关变量的时候，反而会导入难以控制的新的变异产生：期望、紧张、恐惧等被试变量，招来教育上更复杂的情况。因此，使用此种方法要用得适当。对于教育实验中不能消除的变量，如年龄、性别、身高、体重、遗传、性格、智能等变量，更不能采用消除的方法来控制。

无关变量保持恒定：无关变量的保持恒定，也是对无关变量进行控制的最基本方法，应用的范围很广。不能采用消除法的有机体变量，如年龄、性别、身高、体重、遗传等，形状、大小、呈现时间等刺激变量，都要使其保持恒定，使其效果固定。另外，实验常常要在同一时间、同一房间里进行，通常要在同一的条件下使用同一部仪器（或同一型号，效度较一致的几台仪器），主试的态度保持一定，温度、湿度也保持一定，实验其他情境保持恒定等等。为了控制练习和疲劳等效果，一般将明显表现这些效果的最初数次的尝试除外，而用练习曲线、疲劳曲线平坦时那一段实验次数，作为正式实验。这称之为练习法。练习法的意图是将练习和疲劳等效果保持恒定而控制最终的实验效果。

无关变量效果的平衡：用无关变量的消除法或恒定法控制无关变量有