

教育研究法

[英] J. D. 尼斯比特 N. J. 恩特威斯尔 著

张渭城 周南照 徐禾夫 译

教育科学出版社

图书馆

教育研究法

 J. D. 尼斯比特 N. J. 恩特威斯尔 著

张渭城 周南照 徐禾夫译

徐禾夫 校

教育科学出版社

一九八一年·北京

J.D.Nisbet N.J.Entwistle
Educational Research Methods
HODDER AND STOUGHTON
1977

本书根据英国霍德和斯托顿公司 1977 年英文版译出

教育研究法

〔英〕J.D.尼斯比特 N.J.恩特威斯尔著

张渭城 周南照 徐禾夫 译

徐禾夫 校

教育科学出版社出版 新华书店发行

北京顺义县燕华营印刷厂印刷

787×1092毫米 1/32 印张 6.5 字数137,000

1981年5月第1版 1981年5月第一次印刷

印数 1—10,000

书号 7232.29 定价0.65元

出版说明

本书作者尼斯比特是英国阿伯丁大学教育学教授，恩特威斯尔是英国兰开斯特大学教育研究教授。本书在1970年初版，至1977年，已再版四次。

本书以实际、直截的方式论述了教育研究法的一般原理，介绍了进行教育研究的一般步骤和具体方法，分析了这些方法各自的优缺点。本书可供我国教育科研工作者、教育行政人员及其他教育专业工作者参考。

本书前言和1—8、16章由周南照译，9—13章由张渭城译，14—15章及附录由徐禾夫译。

第四版前言

本书读者如想深入探讨有关问题，可参阅我们的姊妹作《教育研究的实施》一书（Educational Research in Action，1972年霍德和斯托顿公司出版）。该书选择了若干研究实例（这些实例不但本身有意义，而且能具体说明重要的研究方法），还收入了已发表的这些实例的研究报告的内容摘要，并作了相应的评论。这种编写方式使《教育研究的实施》成为内容更详尽、也许更有启发性的教本，它在许多问题上对《教育研究法》作了补充。

《教育研究的实施》第二章的论题是实验和调查的相对优点以及应用这两种方法所能得到的资料，内容与本书第二章“制订研究计划”相联系。此外，《教育研究的实施》还从实际的研究项目中援引具体实例，论述了抽样程序、家庭环境和师生态度的评定、考试评分、课堂观察、课程设计等问题，同时说明了《教育研究法》一书中提出的一般原理。

目 录

第四版前言

第 一 章	绪论：实验与调查·····	(1)
第 二 章	制订研究计划·····	(11)
第 三 章	抽样·····	(20)
第 四 章	谈话·····	(29)
第 五 章	问卷·····	(43)
第 六 章	资料处理·····	(55)
第 七 章	环境因素·····	(66)
第 八 章	能力和成绩·····	(79)
第 九 章	考试分数·····	(98)
第 十 章	创造力和分散思维·····	(115)
第十一章	个性评定·····	(125)
第十二章	态度测量·····	(138)
第十三章	观察和个案研究·····	(149)
第十四章	课程设计·····	(159)
第十五章	研究结果的阐释·····	(175)
第十六章	撰写报告·····	(187)
附 录	学校中的教育研究步骤·····	(197)

第一章 绪论：实验与调查

在教育研究领域，方法落后、技巧笨拙的现象非常普遍。许多人在这方面的知识很是可怜，因为他们从书报、杂志甚或广告中看到的常是质量粗劣的“情况调查表”一类东西；也许，他们自己就是这种“科研”项目的调查对象。于是，他们也就收集一大堆“问卷”或统计表格、量表，或找人谈话，或设想新的测验方法，然后把这些东西拼凑起来作为研究的论据。本书无意挫伤这种搜集论据的积极性，而只想提请人们注意改进近年来的教育研究方法。在实践的基础上，今天已有可能在教育科学研究的许多方面，主要是那些需要数量统计、科学调查的方面，提出一种带规范性的东西作为方法准则。当然，还有其它一些重要的研究方式（例如从历史上或哲学上进行探讨），但这方面的研究人员同样会发现那些材料与他们的工作有密切的关系。本书论述的教育研究方法，对于任何一个准备承担某种研究课题（不管这种课题初看起来是多么简单）的人来说，都是十分重要的。

目前，在教育研究法方面训练有素、经验丰富的人材十分缺乏。如果仅仅从现有的书籍考虑，一个初踏上教育研究岗位的新手也许会想，他需要掌握的无非就是统计学的知识，并以此自慰。诚然，我们不乏各种统计学的教科书，但是，统计方法只是研究方法的一个方面，尽管它是必不可少的一个方面。本书很少论及统计方法，因为读者或者已经熟

悉统计学，或者案头已有好的统计学书籍可供参考。本书着重阐述没有现成材料可作尺度的测量方法。现在已经有许多书籍详细说明通行的测验方法。我们在叙述测验方法时只是着重指明它们在教育研究方面的作用。强调这一点，可能有助于阐明处理各种科研题目所应用的方法。

本书前三章论述设计和制订教育研究计划的一般问题，第三章介绍取样的程序，其内容可与每一本统计学入门书的概论部分相连接。第四、五章详细说明正确使用问卷和进行研究性交谈的各个步骤。第六章用通俗的语言概括介绍目前广泛使用的各种资料处理方法；它们既适用于研究题目较小的个别学生，也适用于依靠电子计算机设备进行研究的专业小组。

第七章到第十二章分别论述社会交往、智力、态度这三个主要方面的测量方法——说明如何确定并使用对环境、家庭条件、社会等级、能力和学生成绩以及个性和态度的评价方法。第十三、十四两章着重讨论相关分析和课程设计研究方面的两种特殊方法。最后两章分别指出在阐明研究结果时容易犯的错误，指导撰写专题报告。书后有关于学校中进行教育研究的步骤的附录。

调 查

要把科学研究纳入某种死板划一的模式是很困难的：许多研究项目需要用不同的方法进行多方面的探讨。我们可以在调查（或描述性研究）与实验研究之间划一条大的杠杠。

“调查”一词包括范围很广的研究工作，这种工作主要是通

过考察现状而不是通过建立实验条件或分组实验来完成的。但是，这种调查不是单纯地记录有关事实，它还可以包括测验、研究各类分数的分布情况、用复杂的测量方法分析学校中的对照组等等。通过调查的设计，特别是通过样本的选择，研究人员就使调查带有某种实验的因素。有些问题是无法用对儿童进行实验的办法来解决的，但是通过创造性地运用调查方法却可以提供解决这类问题的依据。比如，通过对聋哑儿童或能讲两种语言的儿童的深入调查，结合使用相应的实验手段，就可以探明言语能力的发展过程。

调查工作要有价值，就离不开科学的比较，不是把某一较大抽样的全面数据进行比较，就是把两组可供对照用的调查结果进行比较。比如，对乡村儿童与城市儿童进行严格限制时间的语言推理测验，其结果即可提供这种比较。同样，家长和教师对“家长教师协会”一类组织所持的不同态度，也可以进行专题调查，加以对比。如果收集的资料可以分类，那么连一般性的调查也能提供比较的基础。这方面一个典型的实例就是1954年由英国教育部进行的一项中学生早期辍学情况的调查。该调查把早期辍学学生按能力水平和社会等级进行分类。调查表明了放弃中学课程的聪明学生的人数比例以及这种比率在不同社会阶层中的差异。

在教育史上最早的一项客观调查是1897年由赖斯进行的“拼写练习调查”^①。他把美国小学每天花在拼写课上的平均时间以及他访问过的每所学校的拼写水平等资料统统收集起

^① J. M. 赖斯：《无益的拼写练习》，Forum, 23, 1897年。

来。奇怪的是，调查结果说明拼写成绩同所花的时间并无联系。从那时以来，调查研究这一科学方法在其复杂程度和量表使用方面都有了很大发展。与赖斯首创的调查研究截然不同，1936年和1947年的“苏格兰智力调查”，对这两年的所有十一岁儿童进行了同样的测验，并用测验结果来验证这样一种理论：文化越低的人生育率越高，而这又将导致民族智力的下降^①。从1948年至1964年，苏格兰每四年就进行一次全国阅读调查，方法是精心挑选有代表性的样本，以检查全国的阅读水准^②。为准备普洛登报告^③所作的调查报告非常详细，已编成上下两卷出版，上卷刊载报告本文，下卷收录各项调查结果。

以上实例表明了教育研究中调查方法适用的广阔课题范围和它的复杂性。这种研究方法的价值在很大程度上取决于调查所选择的样本有多大的代表性。局限于某一领域某所学校的小范围调查研究，能为有关的具体部门提供有价值的情况，但其结果则可能与其它团体进行的类似的研究相矛盾。因此，选择合适的样本具有举足轻重的意义。这个问题，我们将在第三章深入讨论。

① 苏格兰教育研究会：《苏格兰民族智力发展趋势》，伦敦，1949年版。

② 教育和科学部：《阅读的进步，1948—64年》，伦敦，1966年版。

③ 教育和科学部：《儿童及其小学》（普洛登报告），2卷本，伦敦，1967年版。

调查方法的一个很大的局限性在于，调查的结果不能证明某种因果关系。例如，我们或许会发现，在学生的个人爱好和学业成绩之间有一定的关系，喜欢打扮得花枝招展的学生考试成绩多数不好，但不能因此就认为是学生对色彩的个人爱好引起成绩不好；我们只能把后者看作前者的伴随物，而且，一种衡量尺度产生的某些数值总是和另一种衡量尺度产生的特定数值联系在一起，要证明因果关系是不可能的。但如果某种情况总是无例外地跟随着另一种情况发生，那我们就可以推导出某种因果关系。在调查中，通常是收集同一时间发生的各种事实、数据，这样就使人无法知道，哪些情况发生在先，哪些情况发生在后；甚至有些“跟踪”（随访）研究用几年时间来收集情况，也只能找到一些有限的因果证据。每个社会问题，总是有许多错综复杂的因素交织在一起；除非进行十分复杂的统计分析，调查方法对我们认识哪种基本因素在起作用帮助甚少；简单的分析实际上还可能使人导致错误的结论。卡方测验最主要的数值或相互关系并不一定就是主要的因素，这些结果只是表明某些测定数据是最可靠的，可是对最重要的因素也许测量的过于粗略，也许根本就没有列入调查项目内。

作者提出上述告诫，并不是说调查方法不适用于教育研究。实际上，这个方法比实验方法用得更广泛，甚至可以说已到了滥用的程度。所以，使用这种方法，必须特别注意正确地分析和阐释调查结果，尤其是在试图推导出某种因果关系时，更应谨慎从事。

实 验 设 计

在运用调查方法过程中所碰到的困难比实验设计中碰到的困难要小,但是会出现另一些困难。在最简单的实验计划中,如果一个变值——**自变量**——发生变化,也就会引起另一个变值——**因变量**——的变化。前边提及的赖斯“拼写调查”,如用实验方法进行比较研究,就要把儿童进行分组,以比较他们在拼写方面的不同进步。其中一组为**对照组**(控制组),用作衡量自变量(学习拼写所花时间)变化的标准,这一组的拼写教学时间仍与原来的一样。另一组——**实验组**——则减少拼写学习时间,过一阶段以后进行拼写测验,把两组儿童的成绩加以比较。即使这样,我们也不能十分肯定地说是这种变化引起了那种变化,因为拼写时间的减少也许意味着阅读时间的增加,或者拼写课上的更加专心。从教师方面来说,由于有研究人员在一旁观察、试验,可能对自己的教学工作采取更积极的态度(霍索恩效应),而这也许是儿童学习成绩变化的真实原因。

其它因素的影响可以在一定程度上加以控制。实验方法的一个核心问题,就是如何控制正在调查研究的变量以外的各种变量,或者使它们随机化。扩大调查研究的范围就可使一个变量随机化。比如,如果你认为可能是学校的类型或教师的年龄在起作用,那就要把抽样面扩大到各类学校和各种年龄组的教师。由于没有哪类学校或哪种年龄占支配地位,因此它们的作用就不会影响实验的结果。如果调查研究范围较小,那就需要仔细选择合适的学校和教师作为对象,或者

确保这些变量的影响已经消除，以便控制变量。

即使在外变量的影响已经严格控制或者已经随机化的情况下，使实验组和对照组的条件保持相等，仍然是十分重要的问题。只有在这时，学生拼写成绩的变化才能真正归诸于教学方法的变化。还有，使实验组和对照组等价的方法，一是**作有意识的选配**，一是**随机配置**。如用后边这种方法，就要把有关学校的所有学生都编上号，而后用随机数表把学生分入两大组内，或者采用随机取样办法抽取学生名字。如果每组人数都相当多，那么两组在抽取机会上的重要差异就可能很小。另一种方法——**有意识的调整**——旨在确定可能影响学习成绩的重要变量，并保证两组成员在这些变量方面的条件相同。实验组的每个成员与对照组的一个成员进行比较，这两人在性别、年龄、智力、社会出身、拼写能力等方面都要具备同样的条件。为了编组，就需要把自己认为重要的全部变量加以测验，并且保证所选学生的学习成绩是根据相同的记分制得出的，然后再把确实可作比较的成对学生随机分为实验组与对照组。这种成对成对学生的比较，当然是编制可比组的最可靠的方法，但往往并不实际可行。如果没有条件采用这种方法，就可以稍作变动，使两组从总体上看，在学习成绩、年龄、性别、智力、社会出身等方面的学生比例上保持相等。这种总体上相等的分组办法较之分为可比的成对儿童的办法要容易得多。

除了可比性以外，两组儿童都要能代表作为研究对象的所有学生。编组时，许多儿童都可能排除在外。不能编组成对的学生不要编入组内。如果这种编组没有代表性，实验结果就不能作为一般结论。本书后面有关章节将专门论述选择

有代表性的样本，以及使用带有偏向的样本可能造成的全部问题。

这项简单的拼写实验包括的其它工作，可按下列步骤进行：把在性别、年龄、智力、社会出身、拼写能力等方面可作比较的儿童编成实验组和对照组，并对这两组进行预先测验（测验目的是测量因变量，在这个实例中，拼写测验内容不是先前已经学过的词，而且包括实验期要引入的新词）。对照组继续按原来方式、时间进行教学，这样就能使自变量保持不变。实验组减少教学时间，以观察预期的变化。实验期结束后，对两组进行实验后测验（或者是同样的拼写测验，或者是两种在内容和要求上相近的测验）。然后，把两组的成绩进行比较，并对两组平均分数的明显差异作统计学上的检验。

上面所举实例，当然过于简单。一堂拼读课很可能有一个最合适的教学时间长度，而两组儿童之间的比较并不能为此提供确切的证据。但是，如果编几个上课时间长短不同的实验组，而对照组根本不上拼读课，那么，把一系列平均分数进行比较，就很可能指明较为合适的上课时间长度。用来处理两个或两个以上相同小组的统计方法称作方差分析。简单的双向方差分析使我们能处理上述包括两个变量（拼写成绩与教学期限）的问题。然而，由于教育问题一般总是包含许多变量，所以，实验设计势必更加复杂。控制变量（对照变量）有一个严重的缺点，那就是我们事后无法知道这些变量本身是如何影响实验的。很清楚，智力会影响拼读成绩，但性别也会引起某种差异。复杂的实验设计应能使人对各个自变量的作用进行分别研究（拼写成绩与一节课的时间

长短 $\times$ 智力 $\times$ 性别对比)。此外，在这些因素中还可能有这样的情况：某种课时长度可能对聪明的女孩和迟钝的男孩更有效，而另一种课时长度可能对迟钝的女孩和聪明的男孩更适合。由方差分析所测出的自变量相互作用的影响可能相当复杂。比如，某种课时长度可能只适用于某种教学方法，如果改用另一种教学方法，采用另一种课时长度就可能更合适（即时间与方法的交互作用）。或者也可能出现这样的情况：某种课时长度对年龄较小的男孩和年龄稍大的女孩较为合适，而另一种课时长度则对年龄较大的男孩和年龄较小的女孩更加合适（时间 $\times$ 性别 $\times$ 年龄的交互作用）。

方差分析的统计方法对这些复杂的实验设计仍然是适合的。这方面有许多专著可供读者参考，其中之一就是D·G·刘易斯的《教育实验设计》（伦敦，1968年版）。使用实验设计法进行的另一种研究课题实例，将在第十四章论述。课程研究是精心设计的实验项目能对教育研究作出巨大贡献的一个重要领域。对教育研究领域两种主要方法——调查和实验——的简要介绍，表明它们各有长处，也各有缺点。我们把教育研究方法作这种大体的划分，读者决不可因此把它们看作是截然分家的。有时往往很难确定某项课题研究到底采用的是哪种方法，而且也不是所有的研究课题都可按这两种方法进行归类。这里没有提到医学研究中所用的临床法或者**个案研究法**，即对个人或对小集体进行详细的观察或测验的方法，这种方法经常作为进行更大范围研究的一种准备。进行这种详细观察的常用测量方法将在第十三章中介绍。但不管采用哪种研究方式，认真制订计划都是必不可少的第一步。调查和实验，对于了解和认识教育问题，都不失为系统的科

学方法，它们都需要研究人员在动手收集资料之前作出深思熟虑的计划。第二章就将论述制订研究计划的步骤。

小 结

大部分教育研究项目都是以调查或实验的方式作规划的。调查用来描述现实状况，变量是在正常的教育环境中测量的。采用这种方法的重要一环是认真选择样本，但是，即使分组具有代表性，仍然不能测出它们的因果关系，不能误认为某种现象一定伴随另一种现象产生。调查使我们能测定现实环境中重要的教育变量和社会变量之间的关系。更认真的话，还可以抽取子样本(二次抽样)——例如聋哑儿童和听力好的学生——从而使调查带有实验方法的某种因素。

在实验中，要控制自变量的值，以便显示出实验组与对照组的真实差异。分组时必须通过匹配测验法或随机分配(概率配置法)使各组保持条件相等而都具有代表性。通常的方法是用实验前测验和实验后测验来了解因变量的变化，然后再用统计学的方差分析法来比较各组之间的实验后测验结果的差异情况。

第二章 制订研究计划

人们惯于把研究者看成是幽居密室、多少与世隔绝的学究。如果把这种旧框框用在教育研究上，那就会把人引向歧途。要知道，集体的努力，实际的经验，与教师和其他教育专业人员的密切合作，以及对儿童的深切了解——所有这些，正是现代教育研究的特点。然而，人们的传统观念也有一点道理，这就是：科学研究本质上是一种思想的谋划，而不只是行动的纲领。本书论述的研究方法只在帮助人们对教育问题进行更敏锐的思考。科学研究的原始材料不是支离破碎的见闻，而是通过收集论据而得到检验和修正、从而证实其正确性或暴露其不足处的思想和观点。

一个研究项目所包含的工作可以分成以下七步：

1. 确定并准确地表述所要研究的问题；
2. 阅读有关的研究成果；
3. 确定收集和分析资料所要用的方法；
4. 选择和确定所要研究的样本；
5. 收集资料；
6. 加工、分析并阐释研究的结果；
7. 撰写研究报告。