

教育心理学

与

学校心理学专题

王爱民 主编

*Selected Topics on
Educational Psychology and
School Psychology*



100



教育心理学与学校心理学专题

Selected Topics on Educational Psychology and School Psychology

主 编 王爱民

副主编 刘 文 Ray Witte 夏明珠

辽宁师范大学出版社

• 大连 •

©王爱民 2006

图书在版编目(CIP)数据

教育心理学与学校心理学专题/王爱民主编. —大连:
辽宁师范大学出版社, 2006. 7

ISBN 7-81103-404-2

I. 教... II. 王... III. ①教育心理学—文集②高等学校—教育心理学—文集
IV. G44-53

中国版本图书馆 CIP 数据核字(2006)第 079119 号

出 版 人:程培杰

责任编辑:王 丽 刘月娜

责任校对:孙晓艳

封面设计:方力颖

版式设计:孟 冀

出 版 者:辽宁师范大学出版社

地 址:大连市黄河路 850 号

邮 编:116029

营销电话:(0411)84206854 84215261 84259913(教材)

印 刷 者:大连海事大学印刷厂

发 行 者:辽宁师范大学出版社

幅面尺寸:178mm×230mm

印 张:16.5

字 数:330 千字

出版时间:2006 年 7 月第 1 版

印刷时间:2006 年 7 月第 1 次印刷

定 价:25.00 元

前 言

应用教育心理学是研究如何将心理学和教育心理学的原理及发现应用在教育领域中,学校心理学则将这些原理和发现与学校日常活动直接联系起来,利用这些原理与工具直接为学生及学校服务。这些学科涉及的领域包括如何帮助学生适应学校环境,增加学习兴趣,提高学习能力,增进心理健康以及对学生的学习、能力、个性、兴趣等方面的测量与评估,个体教育计划的制定以及对可能发生的问题的预防与干预。

2006年7月,由辽宁师范大学与美国迈阿密大学联合举办的首届中美应用教育心理学和学校心理学国际研讨会及培训班在依山傍海的北方明珠大连市召开。参加这次研讨班报告的是中美教育心理学和学校心理学领域的著名学者。为了保证研讨班的成员有一份可以阅读的材料,也为了使那些不能参加这次会议的学者、教师、研究生和本科生能够分享有关知识,我们决定将尚未发表的报告内容综合成一本书。本书涵盖了应用教育心理学和学校心理学的前沿内容。因为中国学者的报告内容已有相应的专著出版,本书只包括了美国学者的报告内容。

感谢本书作者们在自己繁忙的工作之余利用很短的时间完成了有关内容的写作。感谢译校者夏明珠、范姿和杨建湘等人以最快的速度 and 尽可能高的质量完成了大部分文章的翻译和校对工作。

俄亥俄州迈阿密大学 Susan Mosley-Howard 博士,美国纽约州立大学的 Shannon Gormley 博士,美国鲍灵格林大学 Eric Jones 博士以及辽宁师范大学邹晓燕博士对本书做了认真的审阅,并提出了宝贵的改进意见。同时感谢 Bonnie Thomas 女士对全书英文的校对工作。

感谢辽宁师范大学的领导和辽宁师范大学出版社的领导对出版本书的大力支持,对在较短时间内高质量完成本书排版、编校及印刷工作的辽宁师范大学出版社的工作人员致以诚挚的谢意。邹丽娜、张丽娜、李想、齐璐、彭菲菲等几位同学参与了本书的打印和校对工作,在此一并致谢。

由于本书的准备时间短暂,虽然我们已尽了最大努力,但翻译中不够完美、不恰切之处仍会存在。请读者对译文只作参考,如有疑问,请务必参考英文原文,并对本书的不足和错误不吝赐教。

王爱民

2006年7月

目 录

第一部分 研究、测量与评估及其在学校、教室中的应用

- 3 心理与教育研究简介 王爱民撰
- 13 行动研究的方法及其在教育研究、教育实践中的应用
..... 王爱民、刘文、夏明珠、齐璐、杨丽珠撰
- 24 测量评估与考题编制 王爱民撰
- 35 学生课堂学习的测量评价及其对教师的意义
..... 雷·威特撰, 范姿译, 夏明珠、王爱民校
- 39 Assessment of Student Learning in the Classroom; Implications for Teachers in the Field
..... Ray Witte

第二部分 学校心理学简介

- 47 学校心理学: 一张培训与实践的蓝图 II
..... 吉姆·伊塞尔迪克等撰, 夏明珠译, 范姿、王爱民校
- 63 School Psychology: A Blueprint for Training and Practice II
..... Jim Ysseldyke et. al.
- 89 《学校心理学最佳实践》一书的进展及内容简介
..... 亚历克斯·托马斯撰, 杨建湘、夏明珠译, 杨建湘、王爱民校
- 90 Progress and Introduction of *Best Practices in School Psychology*
..... Alex Thomas

第三部分 心理健康及其管理与青少年自杀预防

- 109 心理疾病依然存在 雷·威特撰, 夏明珠译, 范姿、王爱民校
- 115 Psychopathology is Alive and Well Ray Witte
- 123 第一优先 雷·威特撰, 夏明珠译, 范姿、王爱民校
- 129 Number One Priority Ray Witte
- 138 寻求平衡 雷·威特撰, 夏明珠译, 范姿、王爱民校

- 146 In Search of Balance Ray Witte
158 制定计划 雷·威特撰,夏明珠译,范姿、王爱民校
169 Have a Plan Ray Witte
191 青少年自杀:症状的认识和预防
..... 莎拉·K. 富兰克林撰,夏明珠译,范姿、王爱民校
197 Adolescent Suicide: Recognition of Symptoms and Prevention
..... Sarah K. Franklin

第四部分 超常、阅读困难及其他特殊教育

- 207 促进超常儿童的学习进步:发现并培养早慧的才能
..... W. 托马斯·萨德撰,杨建湘译,王爱民校
213 Facilitating the Academic Advancement of Gifted Students: Identifying and
Nurturing Precocious Talent W. Thomas Southern
221 综合两家之长以培养阅读:组合策略
..... 凯西·M. 麦克麦亨-克罗斯特曼撰,夏明珠译,杨建湘、王爱民校
225 Developing Reading Using Some of This and Some of That: Combining Strategies
..... Kathy M. McMahon-Klosterman
231 教育自闭症儿童的一点经验
..... 凯特·A. 麦克麦亨-克罗斯特曼撰,夏明珠译,杨建湘、王爱民校
234 One Experience Teaching Young Children with Autism
..... Kate A. McMahon-Klosterman
238 Reading Fluency Intervention and Students with Learning Disabilities ...
..... William J. Therrien

256 作者简介

第一部分

**研究、测量与评估及其
在学校、教室中的应用**

心理与教育研究简介

王爱民 撰

美国俄亥俄州迈阿密大学

一、科学研究的目

科学研究的目的是为了发现新知识。新知识可以用来描述、解释和预期事物的发生,确定事物之间的因果关系,进而通过改变原因而改进其结果(Alexandera & Dochy, 1995)并以这些形式表现出来。描述主要指描述某种现象或定义一个概念。一个可以看见的现象不一定能被完全正确地理解。例如,在多动症这个概念被认真研究之前,很多教师可能会认为有多动症的学生淘气、坐不住、不听话,但他们不知道有多动症的学生在管自己的时候有多困难。对一个现象进行认真研究,并对这一现象作出尽可能全面正确的描述或定义,这个描述或定义就成为新知识。这样就产生了很多新知识,癌症也是一个例子。在癌症刚刚被发现的时候,没有人知道是什么原因造成的,也没有人知道究竟什么是癌症,什么不是癌症。经过长期的观察、描述、记录之后,才慢慢地发现癌症的共同规律及确定方法,从而可以对它作出正确的定义和区分。由此可见,描述一个现象或定义一个概念是新知识的一个方面。

另外一个方面是解释。解释是对为什么会发生某种现象作出的说明。解释听起来与因果关系非常接近,因为在对某种现象的发生及发展变化作出解释的时候,人们会找出多种原因。比如说,在研究过程中,通过相关关系或其他关系找到可能有影响或有关的所有变量,从而对为什么发生某种现象作出解释。但是,这些解释不一定是正确的,这些关系也不一定是因果关系的。例如,教师发现学生上课越遵守纪律学习越好。这两者之间有一定的关系,一般来说,学生如果遵守纪律认真听讲,学习会好。但是如果只是遵守纪律,而且只是为了遵守纪律而遵守纪律,而不去认真地参与学习,这个学生的学习不一定就很好。因此在作出解释的时候,需要尽可能找到所有可能的关系和联系以便作出全面的解释。很显然,对某种现象或事物的发生、发展作出解释为我们带来新知识,但需要注意的是,这种解释在很多情况下不只是对因果关系的猜测,因此需要在以后的研究中对它进行验证或检验。

第三个是预测或预期。预测是指对某种现象会不会发生作出判断。例如,对一个小学生以后会不会考上重点高中作出预测;对一个中学生会不会考上重点大学或大学作出预测;对某个学生这次考试能不能考好作出预测。这些方面都是学生、教师和家长很关心的预测。其他方面的预测常常也很重要,有时也许更重要。例如,一名学生会不会产生心理问题;这个学生受到的压力会不会干扰他在考试中的发挥;这个学生对其他

事情的注意程度,对其他事情所投入的精力,以及教师、同学的影响会不会使他选择不同的职业方向。很多这样的预期是非常重要的,例如,预期学生能不能考上重点高中,就是看看哪些学生能考上重点高中,哪些学生有可能考不上重点高中。又如,哪些学生在得到什么样的帮助时成绩能够很快得到改进和提高。有了这个预期,只要给予足够的干预和注意,学校就知道在什么地方需要下工夫,以提高重点高中的升学率。可见,预期很有实用价值,只要预期是对的,我们就能知道应该对哪些人做哪些工作。

最后是因果关系。因果关系也是一个非常重要的知识。如果能够对因果关系进行检验或验证,就可以通过控制原因而改变结果。例如,预期某校 20% 的学生有可能考不上大学。假如这些学生有五种不同的情况。如果对这些学生采取同样的措施,这些措施不一定有针对性,效果也不一定好。但是如果根据所知的因果关系对不同的学生分别给予不同的处理、不同的引导方式和教育方式,就有可能使其改变薄弱的环节并发挥强项。通过改变这些直接针对他们特点的原因,就可以使他们在学业上有很大的进步,这样就可以通过已经知道的因果关系提高学生的学习成绩,提高学校的升学率。

理论是另一个有关的概念。简言之,理论就是概念以及概念之间的关系。概念是理论的重要部分也是理论的基石,理论需要有正确定义的概念作为基础。理论也包括概念之间的关系,例如,概念之间的相关关系、因果关系等。

我们可以将新知识与理论联系一下。新知识包括描述与定义概念。概念与概念之间的关系即解释、预期或者因果关系。由此可见,新知识应该且可以以理论的形式表现出来。因此科学研究的目的既可以说是为了发现新知识,也可以说是为了创建和验证理论。创建和验证理论是两个方面的,它们都是建立理论和发展理论时缺一不可的重要方面,这两个方面和研究方式有关系,需要通过研究方式作出进一步的解释。

二、研究方式

研究方式主要有两种:探索性研究方式和验证性研究方式。探索性研究方式通常被称为质的研究,验证性研究方式通常被称为量的研究。也有人把探索性研究叫做主观性研究,把验证性研究叫做客观性研究。这两种研究方式还有很多其他的名字。主观性研究不是一个很准确的名字,但也揭示了探索性研究中有很多地方研究者可以有自己的意见、想法和主观性,不是所有时候都必须完全客观。因为有自己的主观解释和分析,就使得不同人由同样的研究内容中得到不同的研究结果。但是在进行验证性研究时,只要大家用一样的工具,用一样的数据,得出的结果就应该完全一样,因而相对更客观。我个人比较喜欢用探索性研究和验证性研究这两个词,尽管探索性研究与质的研究有很高的重合度,验证性研究与量的研究有很高的重合度,但它们并不是完全一样的。主要的不同是,探索性研究不一定是质的研究,对探索性研究同样可以进行量化,而且同样可以进行统计分析。然而无论用不用统计分析,探索性研究通常是用于创建理论的。在研究者创建理论时并不能保证创建的理论一定是正确的。验证性研究是

用来验证理论的,一个创建的理论只有经过验证之后才可以被认为是正确的,才可以被接受。在验证性研究时,如果一个理论没有被验证,那么这个理论就可能需要进行进一步的探索性研究或验证。

一方面,探索性研究和验证性研究是研究的两个不同方式。另一方面,探索性研究和验证性研究也是科学研究两个不同阶段。探索性研究是对所有可能有联系的变量以及它们之间可能的关系作出探索,作出尽量全面的发现。验证性研究是对那些在探索性研究中认为可能有的关系和联系进行验证,找到哪些是真正有关的变量和因素,并验证这些因素之间的关系(Hammersley, 1996)。

三、验证性研究从样本到总体的三项推论

进行验证性研究时,具体进行的研究和希望进行的研究是有区别的。例如,对于群体来说,我们希望研究一个总体。例如,所有的小学生、所有学物理的学生。但实际进行研究时,不可能对所有的学生都进行研究,第一是因为不需要,第二是因为花费的时间和精力太多。例如,我们要知道三年级学生的平均身高,一个方法就是把所有三年级的学生进行测量,算出平均身高;另一个方法就是从全国各地选取样本,如每个省选3000个,并对这些样本进行平均计算。只要样本足够大、有代表性,样本平均数和总体平均数即使不完全一致,也会非常地接近。因此不需要对所有的三年级学生进行测量,样本能很好地代表总体。其中最重要的一点是样本一定要具有代表性,一定要代表总体。从样本到总体,对于研究的被试来说是一个推论。只有样本能代表总体,且足够大,这个推论才能准确、可信。

另外一个推论就是定义一个自变量以后,对自变量的操纵或具体测量和自变量的理论定义之间的关系,与样本和总体之间的某些关系相似。因为定义自变量这种理论上的概念可以涵盖很多内容,但是在具体操作的时候不可能把所有的方面都做到,只能像从总体中选取样本一样,选取一个或几个有限测量或操作的办法,这就叫做操作性定义的概念。操作性定义的概念一定要对理论上定义的概念具有代表性,从而可以将结果推广到理论定义的概念。此外,因变量的理论定义和操作定义也具有同样的关系,操作定义必须代表理论定义,这样研究才有效(Ary, Jacobs & Razavieh, 2002)。

以因变量为例。例如,要测量小学生的速算能力,测量的方法有很多。一种操作性定义的办法就是给小学生100道速算题,并记录小学生完成速算题的时间,根据所花的时间评定他们的速算能力。另外,也可以同样给小学生100道速算题,但只给他们两分钟的计算时间,以在这两分钟之内做完多少道题指标,做对多少道题也可以是一个指标。用这些不同方法的结果就不完全一样,但是在一定程度上它们都可以代表小学生的速算能力。

以上所说的都是验证性研究需要注意的方面。探索性研究一般不注重这些方面。例如,发现一个问题,要对一个问题进行探索。因为不是所有的人、所有的样本都会有

这样的问题,样本的选取根据研究的现象而定,而不是代表总体。对于操作定义和理论定义之间的关系也不是需要注重的方面。因为在进行探索性研究的时候,理论定义还没有形成。在没有理论定义的条件下就没有办法进行操作性定义。探索性研究的任务之一是对某个现象或某个概念作出解释和定义。所以探索性研究是验证性研究的基础,只有通过探索性研究,从理论上对需要研究的概念和方面作出定义之后,才有可能对理论定义进行操作性定义,从而对操作性定义的概念及概念之间的关系进行检验,从而验证探索性研究所发现的理论。

因为探索性研究有很多不同的方式,很多时候数据收集的方式就被认为是探索性研究的研究方式,所以要作全面地介绍需要很多篇幅,而验证性研究则是清楚明确、易于讲解的。

验证性研究包括三个主要方面。第一方面是样本的选择;第二方面是研究设计;第三方面是数据收集与分析。

四、样本选择

样本选择是总体与样本之间的关系,总体是被定义的,样本是被选择的。例如,要研究小学生的速算能力,所有小学生都属于这个总体,我们选择 100 个小学生去做速算题,那么这 100 个小学生就是样本。样本的选择有很多种方法,最关键的是样本要具有代表性,要能够代表总体。如果我们选了 100 个最好的小学生或者选了 100 个最不会做题的小学生,这样得出的结果都不能完全代表总体,都是有偏差的。所以在选择样本时需要注意样本必须代表总体。如果样本不能代表总体,比如我们只能在辽宁做,辽宁的小学生可以代表辽宁的小学生,但不一定能代表全国所有小学生,那么就需要对样本作出明确的描述,描述得越详细、越清楚越好。这样读者就能够根据对样本的描述了解到样本在多大程度上代表总体,代表什么样的总体,以及能不能代表读者所在地区的总体。具体进行样本选择时有很多方式,如随机选择、系统随机选择、分组选择等。此外,还可以进行有目的的选择,选择那些具有代表性的、最能代表某方面问题的被试。另外,还有滚雪球式的选择,发现一个人后由这个人介绍另一个人,不断增加被试(Stuart, 1984)。验证性研究一般需要随机选择,一定要让样本代表总体。探索性研究是要对某个问题进行深入研究,所以不一定要代表所有人但一定要代表有需要被研究的特性的人。例如,我们要研究某种教学法对学生有什么不同的影响,我们要随机选择在校的学生,一半放到新教学法组,另一半放到旧教学法组。但是如果我们要进行深入的探索性研究时,例如我们发现学校有三个学生有某种特殊问题,需要进行解决,这三个学生就应该作为被试,即作为有目的的选择被试,从而能对需要研究的方面进行深入研究。比如在几十年前,在我们都不知道多动症这个概念,或者对自闭症没有很好的了解的时候,我们只发现了一些现象,但不知道究竟是怎么回事。在这种情况下,就需要对特定的群体或个体进行深入的研究。对这两个研究方法进行比较就可以看到,在进行验证

性研究时,随机选择样本是特别重要的。如果样本代表总体,研究结果就可以推广到总体,也就是说从研究的一部分学生推广到所有相似的学生或者同类的学生。而有目的的样本可以使研究集中在有相应特征的样本里,这样推广性要相对小得多,但能使研究者深入地研究有关问题。

五、数据收集

数据收集的主要方法有考试法、问卷法、访谈法、观察法。考试法是指对学生的知识、能力进行检验,考的人知道正确答案,要看被考的人是不是知道正确答案。但是如果问一些孩子对学习是不是感兴趣,虽然我们知道对学习感兴趣是一件好事,但是对学习感不感兴趣这样的问题没有正确答案。因为孩子可能对有的方面感兴趣,对有的方面不感兴趣,有的孩子可能对所有的方面都感兴趣,有的孩子只对一两个问题感兴趣。在上学的过程中,我们希望孩子对某些东西很感兴趣,同时对其他方面也感兴趣,这样才能在考试中考好,不偏科,将来能够继续升学并受到很好的教育。这个要求是和我们考试、升学的一般要求相一致的。如果有的学生在一个方面表现突出但在其他方面相对弱一些,如这个学生在突出的那个方面很有成就,也可能被某个学校破格录取或被某单位破格录用。虽然这种事情发生的概率相对较小,我们也不鼓励这样,因为我们都希望学生能打好基础,但是在实际生活中,在某一方面很有特长,在其他方面相对较弱的学生很可能以后很有成就,只是这种方式与我们现在的考试系统不一致,这样的学生不容易通过考试。如果我们的考试系统采取推荐的方式,那么这个学生就很有可能突出出来,从而被推荐。

刚才所举的例子是围绕问卷法展开的,问卷法是指当我们不知道问题是怎么回事的时候,我们问被试,让他们回答这些问题。比如对哪门课感兴趣?每天花多长的时间学习?对哪些教学方法或方式感兴趣?教师、家长、其他同学对学生有什么影响等都可以作为问卷的问题。问卷法主要有两种方式,一种是开放式问卷,一种是封闭式问卷。开放式问卷是让被试回答问题,封闭式问卷是给出选择,比如说对某句话同意的程度或者不同意的程度。例如1是完全不同意,5是完全同意,让学生选择一个数字来进行回答。对封闭式问卷得到的结果进行数据分析相对容易,而开放式的问卷就像探索性研究一样需要把它进行编码整理,然后才可以进行数据分析(Fowler, 1995)。访谈法和问卷法几乎相同,问卷法是在纸上回答问题,访谈法是面对面地回答问题。访谈法要比问卷法多花很多时间,但是访谈法可以有很多灵活性。结构式访谈和问卷法几乎没有区别,就是规定好问哪些问题,规定好如何回答,但是如果被试不能回答问卷,比如有残疾、不能写等各种原因,访谈法就可以通过问被试问题并让他回答。但是半结构式的访谈或者没有结构的访谈则与问卷法有很大的区别,它可以使研究者在访谈的过程中发现问题,并使研究者进行更深入地调查,这些是问卷法不可能做到的。问卷是事先印好的,不可能根据被试的回答相应改变问题。虽然可以用“根据回答通过计算机再给出问

题”这种方式使问卷有变化,但是这与一个有经验的访谈者在访谈的过程中,根据访谈的结果提出新的问题仍有一定的差距。访谈法在探索性研究中是非常常用的,在验证性研究也可以应用,但是用得较少。

另一种数据收集方式是观察法。观察法是对某种现象进行观察、记录,然后对记录的结果进行分析。观察法与访谈法一样,如果是有结构的,要非常明确观察什么,知道怎么观察,观察哪些方面,如何记录,它和问卷法是完全平行的。对于任何无关事件都不观察,只观察事先确定的所要观察的方面,只记录事先确定好的观察内容。这适用于验证性研究,只需要观察已经定义的变量,从而可以对这些变量的因果关系进行计算和检验。但是如果不是有结构的观察,知道大概要观察什么,然后根据观察的情况作相应的调整,或者记录观察到的所有情况,就需要采用开放式观察法。这种开放式的观察一般用于探索性研究。研究者可以在观察的过程中发现问题,了解相关的方面,对需要研究的问题进行全面深入的观察、描述和定义。

要作好这些数据的收集还有很多需要注意的方面。其中最需要注意的就是测量工具的效度与信度。效度主要是指测量工具和数据收集方法是不是真正测量了所需要测量的方面。例如内容效度是指所测量的内容是否符合数据收集方法在设计的时候所计划测量的方面。比如测量小学二年级学生的乘法运算,我们可以出几道乘法题让他们做,并根据测验结果确定这些学生是否掌握了乘法运算。但是如果我们出几道应用题,这些应用题不是小学生能读懂的,虽然他们会进行具体的运算,但是因为他们读不懂题,结果造成不能做对题,这样的题实际测量的不是学生的数学能力,而是他们的阅读理解能力。这样的题目虽然涉及的是数学乘法,但实际测量的是他们的语文阅读能力,如果把这些题目当作测量小学二年级数学乘法能力的测题,那么它的内容效度就不高,因为它没有测到所要了解的内容。还有一个很重要的效度是预测效度。预测效度有两种,一种是预测将来会发生的事情,一种是预测现在正在发生的事情。预测将来事情的需要很容易理解,比如我们预测小学二年级的学生中哪些学生可以考重点高中,有些预测可以非常准确而有些预测会有很大的误差。比如用智力测验的成绩预测学生学习成绩常常很准确,因为智力测验的设计就是为了预测学生的学习成绩。这样通过智力测验就可以预测哪些学生学习很好,哪些学生学习困难,哪些学生学习一般。对那些学习困难的学生可以进行特殊干预从而使他们成绩提高;而那些学习成绩很好的学生可能一般的教育已经不能满足他们的需要,这就要对他们进行特殊教育,提供特殊环境或升级等,这样才能满足他们的需要(Messick, 1995)。

同时性预测效度的一个主要应用是筛选测验。筛选测验是用一个小的测验对总体进行筛选,然后对那些被筛选出来的样本再进行详细的测验。比如要找出哪些学生是超常学生,根据正态分布,超常学生一般定义在2%左右,如果用智力测验评定,即智商在130以上。判断一个学生是不是超常需要很多测验和观察。如果对所有学生作这种

详细、认真的测查会花费我们很多精力和资源,但如果我们先作一个十分钟的测验,这个测验筛选出最好的 5%,这 5%里的 40%就是需要选出的整个总体的 2%。这种详细的测验只是对这 5%的人,从而为我们省下很多宝贵的时间和资源。同时性预测效度的另一个重要应用是间接测量通过直接测量不能获得的信息。比如在美国很多学生用毒品,但这是被禁止的,所以当涉及这类问题时,学生不会真实地回答这些问题。假定在过去的研究中发现学生吸毒与他有几个朋友吸毒有很大关系。假定研究结果发现如果一个学生有 6 个以上的朋友吸毒,那么这个学生吸毒的可能性就在 90%以上。我们就可以利用这个知识,不直接问他是否吸毒,而问在他的朋友里有多少人吸毒。这不是一个直接问被试是否吸毒的问题,但是通过对这个问题的回答以及以前研究所了解的这两个因素之间的关系就可以作出比较准确的推测,从而发现直接测量得不到的答案。应用这样的方法有道德方面的约束,不可以随便乱用。如果用得合适的话就可以得到很多新的信息,滥用则会造成不良影响。因为这叫隐瞒被试,如果经常作这类隐瞒被试的研究,对研究者和研究都会产生很多负面的影响。

评定数据收集好坏的另外一个指标就是信度。主要的信度指标是再测信度。再测信度就是用同样的工具对同样的被试在其没有变化的情况下进行再测。观察再测的结果与初测的结果是否一致,如果测量的结果一致就可以被相信。反之,则结果不可相信,因为其不稳定,有很多变化。比如我们先进行一次考试,10 天以后进行再测,在此期间不教学生考试中的任何内容,如果两次考试结果非常一致,那么这个考试就是一个很可信的考试。不管是现在考,还是将来过一段时间再考,只要学生没有继续学习这方面的内容,学生在这方面的成就和能力就可以被很可信地测量出来的。但如果同样的考试,今天考、明天考、过一段时间考的结果不一样,这就提示这个考试存在很多误差,不能很一致地测量学生的成绩,因而结果不可信。同样,如果问卷、访谈、观察的问题给得好,每次回答都会很一致,如果问题给得不好,被试只是进行随机的回答,结果就会不一致。如果一个数据收集工具所得到结果的信度很低,那么通过这个工具所收集的数据就没有什么价值,因为它随时会变,根据其进行任何分析、任何研究都不会有多少可信度,因为再测一次数据就变了,结果也就变了。

其他的信度还有 AB 卷信度,如果考试用不同形式的考卷,这些考卷之间是否一致是一个很重要的问题,如果 B 卷比 A 卷的得分平均少 10 分,那么谁答 A 卷谁就获益,考试成绩就相对高。这样 A 卷的结果就不可以与 B 卷的结果一起比较。如果 AB 卷之间的信度很高,不管学生用哪一种试卷考试所得的成绩都会基本一样。这种考试方式在很多时候是很有用的,例子之一是 AB 卷交叉使用可以防止学生作弊。

还有一个信度是内部信度。内部信度是所有题目之间是否一致。如果所有题目都一致就表明所有的题目都是测量特定的某个方面。比如要测量学生的自我约束能力,假设有 10 个题目,原则上如果一个学生在 1 个题目上得分高,那么在其他 9 个题目上

应该相对都高,反之亦然。虽然对某一个学生来讲可能某一方面强,某一方面差,但所有的学生一起回答这个问题时,从总体来讲应该是一致的,都是测量他们的自律。如果10个题目里其他题目得分高,只有一个题目得分低或不随其他题目的变化而变化,那么这个题目显然不是测量自律的,我们就应该对这个题目进行删除或者改写。

评分者信度是另一个信度指标。它是指不同教师批改同样的问题、阅读同样的作文或者批改同样的应用题时,是不是对同样的答案给出同样的成绩。如果评估者信度很高,也就是教师们都能遵循事先规定的标准,那么给出的成绩就会很一致。如果评估者信度很低,教师们想怎么批改就怎么批改,学生的成绩就不完全是自己学习的成果,在很大程度上决定于哪个教师给的成绩。如果有的教师给的高,有的教师给的低,而学习不好的学生碰到给成绩高的教师,他的成绩可能会比碰到给成绩低的教师学习好的学生高。这样对学生是不公平的,所以评估者信度高是很重要的(Feldt & Brennan, 1989)。

六、研究设计

探索性研究和验证性研究的研究设计有许多不同。验证性研究是为了验证理论,探索性研究是为了发现和创造理论。验证性研究一个常用的方法叫做因果关系比较法(Causal Comparative),是指对已经存在的不同类别的事物进行比较。比如比较男生和女生在学习上有什么不同的特点,男生和女生就是已经存在的类别。又如比较重点学校和非重点学校的学生在学习态度和学习能力上有什么不同;或者比较重点学校和非重点学校的教师在教学投入、教学法、知识水平上有什么不同,这些重点学校和非重点学校都是已经存在的组和类别。对这些存在的类别进行比较从而找到与之有关的、可能的因果关系的推论就叫做因果关系比较法。这不是对因果关系的直接研究,因为比如说男生和女生之间学习特点的不同,不一定是因为他们是男生所以才这样或者是女生才那样的,很可能有许多其他的因素影响了男女学生的学习特点。所以用因果关系比较法进行总结的时候要十分小心,不能过分肯定这二者之间的因果关系,这只是对因果关系进行某种推测或进行第一步研究。知道了这两者之间有不同之后,要继续深入研究为什么会有这些不同,对这些新知识进行利用,以便找到真正的因果关系,从而改变现状。

还有一种常用的方法叫做相关法。相关法就是测量需要研究的各个方面,从而考察各个方面之间有没有联系。比如说学生对学习越有兴趣学习可能就越好,在学习上花的时间越多学得就越好,都是数量之间关系的例子。让我们对花费时间越多学得就越好这个关系作一个简单的解释。在进行简单相关研究的时候,相关假设是一个线性关系。这个线性关系就是在测量的一个方面学习时间多少和另一个方面学习成绩之间,如果二者之间有正相关的关系,那么学习时间越多学生学习成绩就会越好。如果我们把它当作因果关系,那么就可以通过增加学习时间,让学生学习越好。这显然是不对