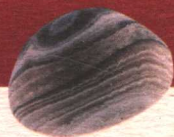


A

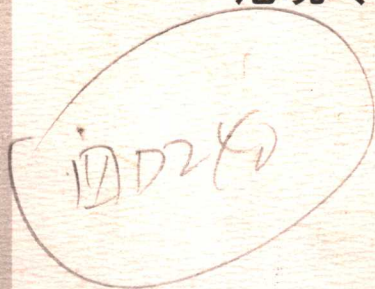
SERIES OF TEACHING THEORY

教学论丛书



教学评价论

范晓玲 著



湖南教育出版社

A

SERIES OF TEACHING THEORY

教学论丛书

教学评价论

范晓玲 著

湖南教育出版社

教学论丛书

教学评价论

范晓玲 著

责任编辑：祁光禄

湖南教育出版社出版发行

湖南省新华书店经销 长沙国防科学技术大学印刷厂印刷

850×1168 32开 印张：7.625 字数：190000

1999年10月第1版 2000年12月第2次印刷

印数：1501—4500

ISBN7-5355-3048-6/G·3043

定价：12.30元

本书若有印刷、装订错误，可向承印厂调换

目 录

第一章 评价——现代教学的价值评判手段	(1)
第一节 评价的意义	(1)
一、评价中有关概念的辨析.....	(1)
二、评价与教学活动的关系.....	(8)
第二节 教学评价方法	(15)
一、教学评价方法的有关问题.....	(15)
二、教学评价方法的类型.....	(20)
三、选择评价方法应遵循的一般原则.....	(28)
第三节 评价与教育教学的可持续发展	(32)
一、可持续发展战略的形成与拓展.....	(33)
二、可持续发展教育教学的特性和态势.....	(34)
三、教育评价的可持续发展.....	(39)
第二章 教学评价——现代教学的决策手段	(45)
第一节 教学评价的研究对象、学科性质及评价要素 与功能.....	(45)
一、教学评价的研究对象与学科性质.....	(45)
二、教学评价的特质要素与内容构架.....	(48)
三、教学评价的功能.....	(57)
第二节 教学评价模型的构建方法	(66)
一、指标构建的方法.....	(67)
二、权重的构建.....	(70)
三、数学模型的择取.....	(72)
四、教学评价模型构建的整体化构架.....	(74)

第三章 成就测验——学生成就商数的评价	(78)
第一节 成就测验	(78)
一、学业成就测验的意义.....	(78)
二、学业成就测验在教学中的基本类型.....	(79)
第二节 测验理论及其发展	(85)
一、经典测验理论.....	(86)
二、项目反应理论.....	(93)
三、概化理论.....	(95)
四、自适应测验.....	(95)
第三节 准标准化测验的编制	(96)
一、确定测验目的、制定测验大纲.....	(96)
二、命题.....	(99)
三、测题的质量分析.....	(107)
四、制卷与评卷.....	(110)
五、建立试卷库和题库.....	(112)
第四章 分数解释——学业成就测验结果的分析	(114)
第一节 测验分数	(114)
一、分数的本质.....	(114)
二、原始分数在评定学生学业成就中的缺陷.....	(118)
三、以原始分数报告学生成绩的基本指标.....	(120)
第二节 常模参照性测验的分析	(121)
一、百分等级.....	(121)
二、标准分数.....	(123)
三、正态化的标准分数.....	(125)
四、常模表.....	(127)
五、成就商数.....	(129)
第三节 目标参照性测验的分析	(131)
一、学生成绩的一般分析指标.....	(131)

二、学生问题分析表·····	(135)
第五章 心力评价——学生心理素质的测评·····	(144)
第一节 心力的结构与评价·····	(144)
一、心力评价的建构·····	(144)
二、心力评价的意义·····	(147)
第二节 学力的测评·····	(149)
一、关于学力的界定·····	(149)
二、心理能力的划分方法·····	(153)
三、学力的评价·····	(153)
第三节 情力、意力和整合力的测评·····	(156)
一、情力与意力·····	(156)
二、整合力的测评·····	(166)
第六章 评定量表——教师教学的评价·····	(168)
第一节 评定量表及其构成要素·····	(168)
一、教学评定量表·····	(168)
二、教学评定量表的构成要素·····	(169)
第二节 教学评价指标体系和标准体系的设计·····	(171)
一、教学评价的指标体系与标准体系·····	(171)
二、教学评价指标体系和标准体系设计的原则·····	(177)
三、教学评价指标与标准体系设计的一般程序·····	(182)
第三节 教学评定量表的设计·····	(186)
一、教学评定量表的形式·····	(186)
二、教学评定量表的编制·····	(187)
第七章 综合评判——教学效果的主成分分析·····	(194)
第一节 主成分分析的意义·····	(194)
一、各种综合评判方法的比较·····	(194)
二、主成分分析的功能·····	(196)
第二节 主成分分析的原理·····	(197)

一、主成分分析的基本思想·····	(197)
二、主成分分析的任务·····	(198)
三、关于载荷的概念·····	(200)
第三节 主成分分析的应用·····	(200)
一、主成分分析的步骤·····	(201)
二、主成分分析在教师课堂教学评定量表构建中的应用 ·····	(202)
三、教师课堂教学评定量表及其数学模型的建立·····	(209)
第八章 元评价 ·····	(215)
第一节 元评价的基本思想·····	(215)
一、元评价的意义·····	(215)
二、教学元评价的内容构建·····	(216)
三、元评价的基本思想·····	(221)
四、元评价的方法·····	(223)
第二节 元分析·····	(226)
一、元分析的基本思想·····	(227)
二、元分析的基本原理·····	(229)
第三节 评价质量的分析·····	(231)
一、评价的效度·····	(232)
二、评价的信度·····	(236)
三、评价的效用性·····	(237)

第一章 评价——现代教学的价值评判手段

第一节 评价的意义

一、评价中有关概念的辨析

（一）测量、测验与考试（Measurement, Test and Examinations）

谈及“测”字，与此有关的概念主要有三个，即测量、测验和考试。在英文中，这三个概念分别用 Measurement、Test 和 Examinations 三个不同的词来表示。通常，人们总是在同一语义上使用着这三个概念。有些测量学者却从更为严格的意义上对它们进行了区分，认为它们是有所区别的三个概念。这些测量学者认为，测量是指人们对事物的属性或特征所进行的某种数量化的确定，测验和考试则是进行测量的工具。这是因为，测验和考试都只能够测量被试做测验或参加考试时的某些行为，而不可能测量到所欲测的所有行为领域。此外，他们还认为，考试与测验也是有区别的，其区别就在于有没有标准化，或标准化程度的高低。根据这一标准，他们认为，考试是没有经过标准化的或标准

化程度较低的一种测量工具，而测验则是经过了标准化的或标准化程度较高的一种测量工具。

上述这些区分主要是从狭义的测量与测验的关系出发的。若从广义的关系来看，测验主要是指对一组标准化的问题所作回答的描述，即由一个人对这样一组标准化的、系列的问题的回答，我们就可以获得对这个人有关特质的某种测量，而这种测量都是数字的。但是，测量的概念则不同，它除了有测验的含义以外，还含有更宽泛的意义。也就是说，我们除了可以用测验法来测量人的特质以外，还可以用各种其他的方法来测量人的特质，譬如用实验法、观察法、等级评定法，或是允许我们以量的形式获得信息的任何方法都属于测量。除此之外，测量不仅涉及到分数的获得，而且也涉及到应用的过程。

另外，在实际操作中，区分测量与测验或考试的不同是比较容易的，但是要区分测验与考试，并将二者截然分开却是相当困难的。这是因为，我们经常在不同的场合以同一语义使用着测验与考试这两个概念。有时我们所谈及的测验可能是没有经过标准化的或标准化程度很低的，同样有时我们所说的考试却又是经过了标准化的。譬如我们经常使用的单元小测验、学科测验等等，虽然名为测验，实际上并没有经过标准化；而高考、“托福”考试等等，虽然名为考试，但却经过了标准化。可见，不能以简单的方式来判断谁是标准化的，谁不是标准化的，而要依照某种工具编制的实际情况来进行判断。

、(二) 评价、评定与评估 (Evaluation, Assessment)

在评价中，我们还常常用到的另一组术语就是评价、评定与评估。与测量、测验及考试这几个概念一样，人们一方面在同一语义上使用着这三个词，另一方面又对它们作出区分。这些术语，不仅在汉语上用词不同，而且在英文中用词也不完全相同。譬如，英文中评价是用单词 Evaluation 来表示的，而评定和评估

则用单词 Assessment 来表示。

一般来说,评价意味着对某一事物的价值给予一般的衡量,评定与评估则用于对人的测定,它包括着对人进行分等、考试,证明其合格等等,如评定学生的学业成绩、评定教师的教学能力、评定一个人对某一特定职务或职称的申请等。评价与评估的区别在于,评价是在抽象领域中使用的一种术语,它既指人也指物,而且还需要对人或物作出价值的判断,它是从事研究和开发的教育研究人员的重要活动,如教育评价就是如此;评定或评估则是在具体领域中使用的一种术语,它是教育工作者在日常的教学工作中的重要活动。

不过,不同的国家对这些术语的使用也不完全相同。在英国,评价多指对效果的鉴定,以及对学习活动(如一堂课)、一门课及一整套课程的效果的判断。在美国,评价一词也可用来指评定,如评定学生的学习成绩,分析教材与观察教学活动,直接向教师、行政人员、学生本人、家长以及社区的其他人士了解情况等。①

评价的定义方法是多种多样的,如斯塔弗尔比姆(Stufflebeam)等人把评价表达成“是为了进行问题选择判断而描述获得和提供有用信息的过程”。在这一意义上,评价既包含了测验的意思,又包含了测量的意义。关于评价的第二种通俗概念的解释是学业成绩与目标一致性的测定。其他更简单的定义,则把评价看成是专业判断或是一种允许人们对有价值的事物所作判断的过程,它不是评价定性资料或数据,就是评价定量资料或数据。总之,无论是在英语、日语中,还是在汉语中,评价都包含有“价值与判断”的意思。价值是客体满足主体需要之间的一种特

① (英)德里克·郎特里著,赵宝恒、汪莲如、潘祖培译:《英汉双解教育辞典》,教育科学出版社,1992年版,第137页。

定关系，它通过人们的社会实践予以实现。价值判断则是价值关系在人脑中的反映，它是以一定的事实判断为依据，对事物的价值程度进行判断的思维过程。因此，作为完整的评价活动来说，不仅要根据评价目标收集定量与定性方面的资料进行量与质的分析，而且还要根据这些资料对评价对象作出最终的价值判断。

关于评定的定义方法也有许多种，有时它与评价一样，有时它又更多地用在个人问题的调查分析或判断上。但在大多数的情况下，人们更广泛地使用着评定一词。

在实际使用中，要截然区分评价、评估与评定是很困难的，因为它们的区别是相对的，其间存在着密不可分的关系。而从全面、整体的高度去看待评价时，也不必区分得太清楚。

（三）测量与评价的关系

作为一种通用的术语，评价是我们日常生活中使用频率较多的一个词汇。但是，人们对评价含义的认识却总是有些混乱，特别是用在教育中的时候，更是如此。这是因为在某些情况下，评价总是作为测量的同义词来使用的，于是在进行成就测验的时候，教师不是把它称为成就“测量”，就是称为成就“评价”，在这里很少有人去认真考查和区分这两个概念的特殊含义。然而在另外一些情况下，“评价”又用作与“测量”无关的评价方法的集合概念，这时评价与测量的概念是不同的，它们在使用上的区别为：评价是用于对学生的行为进行定性的描述，譬如行为的轶事记录法；测量则相反，它属于对学生行为进行定量分析的方法，譬如测验分数。

根据教育的观点，人们又常常把评价定义成学生实现教育目标程度的系统测定过程。从这个定义可以看出，它既蕴含了评价的两个实质性内容，同时又表现出其不足之处。具体来说，一是注意了评价含有系统过程的意思，而忽视了非控制观察法对学生实现教学目标程度的作用；二是评价总是假设事先早已确定了教

学目标。如果事先没有确定目标，那么判断学生学习进步的程度显然是不可能的。

事实上，评价是比测量更加综合的概念，可以把评价定义为确定学生达成教育目标程度的系统过程。这个系统过程包括描述学生的行为变化和与希望的行为变化有关的价值判断的测量和非测量的技术。换言之，它既包含对学生行为的定性描述，又包含对学生行为的定量描述，还包含与行为描述有关的价值判断。所以说，评价是对学生实现教学目标程度的行为进行系统的定量与定性描述，最终作出价值判断的过程。测量则局限于对学生行为的定量描述，它既不包括定性描述，也不包含对有关行为的价值判断。也就是说，测量只是评价过程的一个组成部分，而评价可以不依赖测量，从这一意义上说评价远远超出了简单定量描述的层面。下图清楚、简洁地表明了评价与测量之间的联系与区别。

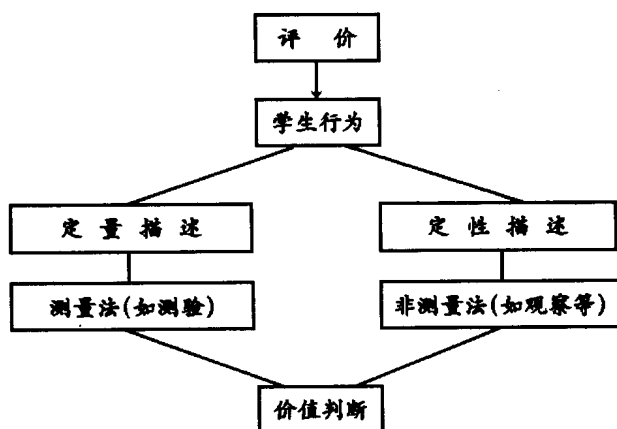


图 1-1 评价与测量的关系

评价的重点在于强调教育目标的实现程度。譬如，通过教学，学生能够准确、迅速做有关乘法题的程度如何？学生对数字系统的掌握程度有多大？在学生的学习中，若是没有日常的辅

导，学生也会进步吗？如果进步了，那么进步的幅度又有多大？在有效地利用时间方面学生有什么进步吗？如果有的话，是多少呢？学生的书法字迹写得更清楚了吗？如果是的话，又有多清楚呢？诸如此类的问题，都是我们必须要求每个学生进行自评和回答的一些典型问题。因此，在评价中，各种方法都是必要的，一种合情合理的评价方法既包括测量技术，又包含非测量技术，因为每一种方法都有其特殊的功能与用途。

虽然在测验、测量、评价和评定这四个术语中，测验所涉及的范围是最狭窄的，但是它的使用频率却是最高的。测量也好，评价、评定也好，都要用到测验，测验是它们所使用的共同工具。不过，一般的教育或心理测量则是指测验意义上的测量，所以本书我们所说的测量是特指测验的。

评价与测量有着密切的关系，它们之间的关系具体表现为价值判断与事实判断的关系，即测量是评价的基础之一，评价需要借助测量的结果进行判断。在使用中，人们很容易把测量等同于评价，这是非常错误的做法。可以说，目前我国所实施的教学评价尤其是对学生学习的评价，都是测量意义上的评价，而非真实意义上的评价。

评价与测量既有着密切的联系，又有着本质的区别。其区别主要是：首先，从性质上看，测量可以说是一种纯客观的过程，客观性是测量质量的首要标准，客观的测量期望测量者尽可能地排除各种主观因素的影响。评价是在测量的基础上进行的活动，但是它却超出了描述层面，而试图去确定行为的价值。从本质上来说，评价并不是纯客观的，而是同时具有客观性和主观性两个基本特性，即其客观性表现在它是建立在对教育现象的客观测验、客观检验和客观描述的基础上的，其主观性则表现为它又是以主体的需要和目标为准绳进行评判的。其次，从复杂性上看，测量较为单一，它只重视在数量上对事物的特性进行测定，评价

则要复杂得多，它不仅重视数量分析，而且还重视性质分析。再次，从内容上看，测量着重对教育或心理现状的确定，评价则侧重于对教育或心理情况的解释与判断。最后，从所分析的资料上看，测量主要是通过测验所收集的量的资料进行分析，评价不仅包括了测验所得的量的资料，而且还包括用观察法、谈话法、逸闻记录、录音等技术所获得的非数量的、质的资料。

由此可见，测量与评价一方面是不完全相同的两个概念，另一方面又是互为关联的两个过程。这就好像两个学生可以获得相同的测量结果，但是我们却可以对这两种测量结果进行不同的评价一样。譬如，在五年级末有两名学生阅读能力的测验成绩都达到了五年级的水平，虽然如此，我们对这两名学生的评价却不完全相同。这是因为，在五年级初，一名学生的阅读能力只有三年级的水平，而另一名学生的阅读能力已达到了四年级四个月的水平了。因此，教师对这两个学生的成绩评价是不一样的，要具体说明其中一个学生的进步在平均水平以上，而另一名学生的进步却处于平均水平以下，即一名学生的进步速度快，一名学生的进步速度慢。在此基础上，进一步的评价还需要借助其他非测量的手段，分析这两名学生之所以有这种差别的原因，以及对两名学生今后的发展作出预测等等。

在这里需要特别强调指出的是，我们从不测量或评价泛指的人，我们所要测量或评价的是一个人的特质或特性，譬如一个人学习潜能、代数知识、诚实、毅力、教学能力等方面的特质。这样一来，大家就不至于把它们与人的价值的评价混为一谈。可是，在现实生活中，教师、家长和学生并不十分注意区别这些概念的不同。由于现代教学评价的思想传入我国只有 20 年左右的时间，其兴盛也只是近几年的事，但远还没有达到全民皆知的水准。在这样的历史背景下，如果说家长和学生不去区分这些概念还情有可原的话，但作为教师也不去区分这些概念却怎么说不可

过去，因为这毕竟是与教学息息相关的过程，它涉及到一名教师如何去准确、全面地评价学生的学习状况和自己的教学情况。所以，教师应当十分清楚地辨明评价与测量的区别，以免误用和滥用。

二、评价与教学活动的关系

（一）评价与教学过程

可以这样说，教学的最终目的就是期望改变学生各方面的行为。从这一点上来讲，评价是构成整个教学过程的有机组成部分。而“期望学生行为发生变化的方面”是由学校和教师设立的教学目标，评价只是确定这些目标达成程度的过程。当人们不断地讨论教师的教与学生的学的相互关系的时候，却忽略了它们与评价的相互关系。我们认为在整个教学过程中，弄清教师的教、学生的学与教学评价这三个方面的相互关系是十分必要的，它对教学有着非常重大的作用与意义。关于这一点，可在以下教学评价的具体进程中得到反映。

1. 识别和定义目标。根据学生行为的理想变化来识别和规定教学目标。无论是在教学方面，还是在评价方面，首要的工作就是根据课堂教学所规定的期望的学习结果来确定教学目标。譬如，学生最终希望得到的学习经验是什么？学生应当具备什么程度的知识和理解？学生能够表现出什么样的技能技巧？学生会发展什么样的兴趣和态度？学生的思维习惯、情感和行为会发生什么样的变化？我们要使哪些学生的哪些特定行为发生变化？当我们在这些行为变化方面获得成功的时候，学生的感受如何，他们愿意是否发生这些变化？对此，只有根据特殊行为的识别目标并清晰地阐明这些目标，我们才能看出教学过程的发展方向，才能建立评价学习结果的基本步骤。

2. 设计和说明学习经验。设计和说明的学习经验要与所描述的教学目标相一致，这是把课程内容和教学方法与所设计的学习经验结合起来的关键所在。我们希望学生的行为朝着理想的方面发生变化，而评价这种变化的重点应该放在学生的学习过程，而不是学生的学习结果。尽管我们的评价常常针对的是学生的学习结果，对学习结果进行着直接评价，对学习过程只是间接地进行评价，但是评价人员不能因此而忽略了对过程的评价，应当意识到对学生学习过程的评价是重点，学习结果的评价是为学习过程的评价服务的。当然，学习过程的评价相对于学习结果的评价来说，难度要大得多，所花费的时间要多得多，其计划和准备的过程也是相当复杂的。这是因为学习过程的评价所强调的大多都是隐性的东西，除非对学习经验的设计和构建是由清晰阐释的目标作为基础的，而且由学生进步的评价作为估计，否则我们是无法对学生的学习过程进行合理评价的。

3. 确定学生的进步状况。对照所规定的教学目标来确定学生学习的进步程度 评价与教学过程的第一步——规定教学目标——不仅回答了评什么的问题，而且也回答了怎样评的问题。譬如，我们如何选择理想的行为变化的情形呢？我们如何选择最恰当的评价方法呢？我们如何构建或选择特定的评价技术呢？我们如何解释和应用评价结果呢？诸如此类的问题都与评价学生的进步程度有着直接的关系。为此，我们必须对照教育目标进行仔细的、认真的评价。

4. 利用评价结果来改进学习和教学。评价学生的学习进步程度不仅对教师和管理者来说是不可缺少的，而且对学生自身来说也是必不可少的。这是因为，对教师来说，评价可以提供反馈信息，可以帮助我们澄清希望学生学习的东西，帮助教师改进教学。对学生来说，评价的优点更是数不胜数，因为它可以为学生提供有关其进步的信息，帮助学生认识到学习的难点所在。对学

生和教师双方来说，评价象征着未来学习活动的准备状态，有助于学生和教师对未来学习经验的选择。对于管理者来说，评价可以提供行之有效的管理方法，譬如对特殊班级的安置等，因为它们起着导向作用。为此，我们要十分慎重地选择和构建评价的方案或方法。

当然，评价还有许多其他方面的用途，譬如向父母汇报学生的学习进步的情况等。但是人们对评价的其他间接用途很少充分了解。由于慎重开发评价技术可能给教师提供使用课程内容和教学方法效果方面的线索，故这种信息还有助于预先确定课程目标的适合性与可行性。因此，评价既有助于教学过程，又有助于直接改进学生的学习。

（二）评价与课程发展

在评价中还应特别强调的是课程建设和发展的问题，因为评价与课程的发展有着直接的关系。如何将评价与课程建设和发展联系起来呢？其基本方法就是编制课程手册，即针对每一主要的课程内容领域任命一个教师委员会，然后让教师委员会在主要的课程内容领域里开发其工作单元。Lindvall 等人总结了作为教师委员会在开发课程中应遵循的原则，大意如下：^①

1. 在行为领域里，确定理想的结果或特别的目标。在特殊的教学单元里，教师委员会的首要任务是正确地确定学生在掌握某一单元之后，期望他们还能做什么。这里的重点在于要根据所确定的学生行为来解释这些目标，而不能根据教师将要去做的工作去解释这些行为，因为教师打算做的事并不能描述学生的学习活动。每一种事实的报告，都应描述学生在经过学习和练习以后将能做什么。

^① C. M. Lindvall, Stella Nardoza and Margaret Felton, *The Importance of Specific Objectives in Curriculum Development*, University of Pittsburgh Press, 1964.