



新世纪教师教育丛书·修订版

袁振国 主编

JIAOYU PINGJIA YU CELIANG

教育评价与测量

金 娣 王 钢 编著



教育科学出版社

Educational Science Publishing House



新世纪教师教育丛书·修订版
袁振国 主编

教育评价与测量

金 娣 王 钢 编著

教育科学出版社
· 北 京 ·

责任编辑 杨晓琳
版式设计 贾艳凤
责任校对 贾静芳
责任印制 曲凤玲

图书在版编目(CIP)数据

教育评价与测量/金娣,王钢编著. —2版. —北京:教育科学出版社,2007.12(2008.5重印)
(新世纪教师教育丛书/袁振国主编)
ISBN 978-7-5041-3997-9

I. 教… II. ①金…②王… III. ①教育评估②教育-测验 IV. G449

中国版本图书馆CIP数据核字(2007)第148639号

出版发行	教育科学出版社		
社址	北京·朝阳区安慧北里安园甲9号	市场部电话	010-64989009
邮编	100101	编辑部电话	010-64989593
传真	010-64891796	网 址	http://www.esph.com.cn
经 销	各地新华书店		
制 作	北京金奥都图文制作中心	版 次	2002年12月第1版 2007年12月第2版
印 刷	北京人卫印刷厂	印 次	2008年5月第2次印刷
开 本	787毫米×1092毫米 1/16	印 数	3 001—8 000册
印 张	20.75	定 价	34.00元
字 数	280千		

如有印装质量问题,请到所购图书销售部门联系调换。

《新世纪教师教育丛书》修订版前言

振兴民族的希望在教育，振兴教育的希望在教师。

教师是一种专门化的职业，它有自己的理想追求、有自己的理论指导、有自觉的职业规范和成熟的技能技巧，具有不可替代的独立特性。教师不仅是知识的传递者，而且是道德的引导者，是思想的启迪者，是心灵世界的开拓者，是情感、意志、信念的塑造师；教师不仅需要知道传授什么知识，而且需要知道怎样传授知识，知道针对不同的学生采取不同的教学策略。教师职业的专门化既是一种认识，更是一个奋斗过程，既是一种职业资格的认定，更是一个终身学习、不断更新的自觉追求。中国教师队伍的培养和培训正在发生着历史性的变革，正在从发展数量向提高质量转变，提高质量将成为新世纪教师队伍建设的主旋律。在这种转变的过程中，无论是职前培养还是职后培训，无论是教育机构还是教师个人，都需要以一种新的姿态迎接这一转变。

从我们对广大中小学的调查中了解到，面对全面推进素质教育的新形势，当今教师迫切需要不断更新教育理念，提高将知识转化为智慧、将理论转化为方法的能力，提高将学科知识、教育理论和现代信息技术有机整合的能力，增强理解学生和促进学生道德、学识和个性全面发展的自觉性。为了响应这种挑战，广大的师范院校和教师培训机构都在积极探索教师教育的新内容和新方法。以华东师范大学为例，1996年起，就有组织地开发了现代教育理论与教育实践紧密结合的新课程系统和教

学模式，这些课程包括：教育新理念、课程理论与课程创新、现代教育技术、教育评价与测量、当代教学理论、教学策略、心理健康的指导和研究、网络教学、课件制作、教会学生思维、师生沟通的艺术、优秀班主任研究、中小学教学与管理案例分析、教育研究方法、基础教育改革的理论与实践等。参加课程开发的教师 60% 具有教授、副教授职称，80% 具有硕士、博士学位。这一项目列入了教育部师范司“面向 21 世纪高师教学与课程改革计划”重点项目。我主持了这一项目的研究和实践。根据边实践、边研究、边总结、边改进的方针，经过几轮教学，逐渐形成了一批相对成熟的教材，在反复教学的基础上，经过精选整合、修改补充，于 2001 年由教育科学出版社出版。由于这套丛书理念新、注重理论联系实际、强调可操作性，出版以后受到了读者极大欢迎，数次甚至数十次重印，为满足教师教育的新形势、新要求尽了绵薄之力。

正是由于这套丛书影响大、受欢迎程度高，所以更增强了我们的责任感。丛书出版的六年多来，教师教育的知识、观念不断更新，教师教育的实践不断发展，我们对教师教育课程的认识也不断深化，为此，根据教师教育的新形势和新要求，我们对《新世纪教师教育丛书》进行了修订。这次修订包括两方面，一是对第一版图书进行了较大修订，更新了内容，改善了结构，修饰了语言，修订了错误；二是丛书新增了若干选题，以反映教师教育的新要求。

祝愿丛书与我国一千多万中小学教师共同成长。

袁振国
2007 年 7 月

目 录

第 1 章 教育评价概述 / 1

第一节 教育评价的基本概念 / 2

第二节 教育评价的主要类型 / 8

第三节 教育评价的功能 / 18

第四节 教育评价模式 / 21

第 2 章 教育评价的历史和发展 / 33

第一节 教育评价的源流 / 34

第二节 西方教育评价发展的原因分析 / 35

第三节 我国教育评价的发展 / 41

第 3 章 教育评价的一般过程 / 58

第一节 教育评价的准备阶段 / 59

第二节 教育评价的实施阶段 / 68

第三节 教育评价结果的处理与反馈阶段 / 77

第4章 教育评价标准的编制 / 93

- 第一节 编制教育评价标准的依据 / 94
- 第二节 设计教育评价指标体系的方法 / 107
- 第三节 教育评价基准的制定 / 116

第5章 收集教育评价信息的方法 / 121

- 第一节 测验法 / 122
- 第二节 问卷法 / 133
- 第三节 访谈法 / 140
- 第四节 观察法 / 145
- 第五节 文献法 / 150

第6章 处理教育评价信息的方法 / 155

- 第一节 处理教育评价信息的方法概述 / 156
- 第二节 定性分析方法 / 159
- 第三节 定量分析方法 / 164
- 第四节 定性与定量相结合的处理方法 / 186

第7章 教育评价的再评价 / 195

- 第一节 教育评价再评价的标准 / 196
- 第二节 教育评价再评价的方法 / 202
- 第三节 教育评价再评价的实施 / 215

第8章 学力测验 / 228

- 第一节 学力的概念 / 229

第二节 学业成就测验 / 231

第三节 操作技能测验 / 245

第 9 章 学生评价 / 259

第一节 学生评价的概述 / 260

第二节 学生思想品德的评价 / 266

第三节 学生实践能力与创新精神的评价 / 274

第四节 学生评价的发展趋势 / 280

第 10 章 教师评价 / 290

第一节 21 世纪的教师评价观 / 291

第二节 课堂教学评价 / 294

第三节 班主任工作评价 / 312

后 记 / 322

1

教育评价概述

在当今世界教育领域中，教育评价、教育基础理论和教育发展被誉为三大研究课题。教育评价对于教育发展和改革，对于教育管理和决策，都有至关重要的作用，因而备受各国政府部门的重视。对于什么是评价、什么是教育评价、教育评价的类型有哪些、教育评价的功能是什么、教育评价模式有哪些等基本理论问题的探讨和研究，有助于对教育评价问题的深入研究。



第一节 教育评价的基本概念

一、教育评价的定义

在汉语中，评价是评定价值的意思。^①那么什么是教育评价呢？所谓教育评价，是指在系统地、科学地和全面地搜集、整理、处理和分析教育信息的基础上，对教育的价值做出判断的过程，目的在于促进教育改革，提高教育质量。

所谓教育评价，是指在系统地、科学地和全面地搜集、整理、处理和分析教育信息的基础上，对教育的价值做出判断的过程，目的在于促进教育改革，提高教育质量。

上述定义包含以下四个要点。

第一，教育评价的对象（即价值判断的对象）可以是教育领域中的任意元素，既可以是教育的参与者（人物），如教师、学生、教育管理人员等，也可以是教育现象和活动（事物），如教育方针、教育政策、教育活动、教育过程、教育效果等。

第二，教育评价的本质（即评价的主要性质）是对教育的价值做出判断，是评价者的主体需要与被评价对象的客体属性的一种特殊的效用关系运动。

第三，教育评价的手段（即评价的方法和技术）是运用科学的评价技术和方法。综合运用测量、统计、系统分析等手段进行综合分析判断，既有定量的，又有定性的。

^① 中国社会科学院语言研究所词典编辑室·现代汉语词典[M]·北京：商务印书馆，1978：882.

第四,教育评价的目的(即评价活动所要达到的境地)是为了促进教育改革,提高教育质量。为了达到这一目的,要为被评价者(或被评价单位)诊断各种教育问题,探索改进措施,选择行为决策。

为了深入理解教育评价的定义,从价值论、认识论和实践论的角度,剖析评价和价值、评价与认识及评价和实践的关系。

(一) 评价和价值

由教育评价的定义可知,教育评价的本质是对教育的价值做出判断,而进行价值判断离不开一定的教育价值观。因此,没有正确的价值观,就谈不上客观、公正、有效的评价。

1. 价值

什么是价值?目前学术界仍是众说纷纭,还未形成一致的看法。在日常生活中,“价值”泛指事物或人的功能和作用。在哲学上,有关价值的定义主要有:意义说、满足需要说、兴趣说、情感说、欲望说、先验性质说、情境说、功能说、有用说和结果内在说等。^①马克思曾指出:“‘价值’这个普遍的观念是从人们对待满足他们需要的外界物的关系中产生的。”^②他还说,价值“是人们利用的并表现了对人的需要的关系的物的属性”,^③“表示物的对人有用或使人愉快等等的属性”,“实际上是表示物为人而存在”。^④根据马克思的科学论断,我国哲学界一般认为,价值从本质上属于一种关系范畴,即是通过主体和客体的相互关系而体现的,这种关系的联结涉及主体对客体的需要和客体客观属性。只有当主体具有某种需要,而同时客体本身也具有满足主体需要的客观属性,才能体现出价值。缺少主体的需要,或者主体有需要,但客

① 王玉梁.论价值本质与价值标准[J].学术研究,2002(10).

② 马克思,恩格斯.马克思恩格斯全集:第19卷[M].北京:人民出版社,1965:406.

③ 马克思,恩格斯.马克思恩格斯全集:第26卷[M].北京:人民出版社,1965:139.

④ 马克思,恩格斯.马克思恩格斯全集:第26卷[M].北京:人民出版社,1965:326.



体本身没有满足这种需要的客观属性，那么主、客体就没有形成关系的可能性，也就无法谈论价值。

2. 教育价值

所谓教育价值，就是指教育能够满足人和社会需要的程度。具体说来，教育价值主要体现在两个方面：教育对人发展的价值和社会的价值。教育对人发展的价值主要是指教育对人的精神需要、物质需要的满足，也就是对人的知识增长、能力的提高、个性的发展、心理上的满足以及身体的发育等需要的满足，以提高人的价值实现能力和身体素质，增强人的创造自觉性，以及求得人的全面的、自由的发展，从而为每个受教育者最大可能地实现其个人价值和个体价值社会化准备条件。教育的社会价值主要是指主体对教育在社会内容方面的价值，它大致包括教育的政治价值、经济价值和文化价值。教育的政治价值主要体现在教育对维护和巩固政治制度的作用方面；教育的经济价值主要表现为提高人的劳动技能，促进生产力的发展方面；教育的文化价值主要表现在传递和继承文化、发展和创造文化方面。

所谓教育价值，就是指教育能够满足人和社会需要的程度。

由于教育价值的主要体现是对人的发展价值和社会价值，因此，受人的身心发展需要和社会的政治、经济、文化、人才需要的影响极大。这就形成了教育价值观的多样性，如对于一心要让子女上大学的家长来说，他们评价一所中学教育价值的标准，就是子女能否考上大学；对于一个立志要成为艺术家的人，评价教育价值的标准是，教育为其成为艺术家所提供条件的程度；社会各界、各行各业评价教育的价值，同样是以教育培养的人满足他们需要的程度为标准。因此，对于同一教育现象，具有不同教育价值观的人、对教育有不同要求的人，会得出完全不同的评价结论。而社会发展和人的自我完善对教育的要求，又是通过社

会各阶层具有不同教育价值观的人们对教育的不同要求反映出来的。这就要求我们在实施教育评价时，必须采取实事求是的态度，在正确价值观的作用下，如实地反映被评价对象的真实价值。

（二）评价和认识

评价从本质上说是一种认识活动。马克思主义认识论认为，认识的本质是能动的反映，而评价尽管过程是复杂的，但它首先是客观社会存在的反映。可见，二者都是主观形态的意识活动，反映的对象都是客观存在的。但是，我们也不能不看到作为特殊认识的评价与人的认识活动之间的区别。

1. 对象的区别

我们通常所说的认识，是指对客体本身某方面的本质或规律的认识，因此其对象是客体本身。而评价的对象不是客体本身，即不是客体的实体性属性，而是客体的社会属性。

2. 主体性的区别

诚然，进行认识和评价的主体都是人，但认识的目的在于揭示事物的本质联系，而评价则不同，主体在评价事物时，总是把客观事物的属性同自身的需要紧密联系起来，主体的需要不能排除在评价内容之外，为此，评价较认识具有更强的主体性。

3. 反映形式的区别

评价、认识的本质是反映，然而，两者的反映形式是有区别的。认识常以理性的、抽象的思维形式来反映客体的本质和规律，而评价是只有在主体的需要和兴趣的关系中才能得到实现的特殊反映。每一种评价不仅是行为主体生活状况的反映，而且是评价主体世界观的体现。因此，评价作为价值判断，常常以理性和抽象思维之外的形式来反映客体与主体需要的关系。



明确两者之间的区别，便于我们更深入地了解两者之间的关系。认识是评价的基础和前提，只有在对事物有了一定认识之后，才能评价事物，同时，评价又为进一步的认识提供指导。

（三）评价和实践

教育评价活动本身就是一种实践活动，评价正确与否是要通过实践检验的，这里包含两层意思。

第一，把评价理论和方法运用于实践，接受实践的检验。

在教育评价的实践中，人们对事实判断的客观性深信不疑，对价值判断的客观性却一直持怀疑态度。究其原因，是因为事实判断和价值判断是不同的。事实判断是关于客体本身是什么的判断，而价值判断是关于客体对主体的意义是什么，对主体意味着什么的判断。两者的本质区别在于，在价值判断中多了一种对于价值而言是决定性的因素——人的需要。而人的需要是复杂的，“人”和“需要”这两个概念的内涵是复杂的。但这种复杂性并不能使我们放弃对教育评价客观性的追求，反而坚定了对教育评价客观性的追求——那就是把评价的现有成果，即评价理论和方法运用于实践，接受实践的检验。对于实践证明是科学的理论和方法，应该保留，并继续运用，对于实践证明是不科学的理论和方法，我们要加以修正或摒弃。

第二，评价的结果要接受实践的检验。

教育评价既基于对教育客观规律本身的认识，又基于对满足人和社会需要的价值关系的认识。教育本身的规律以及教育对人和社会的价值，就构成了教育评价活动的两个尺度，其一称之为合规律，其二称之为合目的。所谓“合规律”指评价要合事实、合逻辑、合规范，“合目的”是指教育评价目的和教育评价依据的合理性。而要衡量评价结果是否“合规律”和“合目的”，则是需要接受实践检验的。这种检验主要从三方面来观察：其一是价值客体的发展水平是否与评价结果相一致；其二是价值主体需要的满足程度；其三是评价主体的主观愿望是否实现。

国外学者曾用一个简单的公式对评价做了形象的说明：

评价 = 定量描述（或定性描述）+ 价值判断

从这个意义上说，评价就是在定量描述或定性描述的基础上进行价值判断的活动。

二、教育评价与相关概念的区别

（一）教育评价与教育测量

所谓教育测量，就是依据一定的法则（标准）用数值来描述教育领域内事物的属性，是事实判断的过程。要对教育的价值做出判断，必须取得有关教育活动大量的、系统的信息。要获取这些信息，教育测量是有效手段，它是教育评价的基础。

所谓教育测量，就是依据一定的法则（标准）用数值来描述教育领域内事物的属性，是事实判断的过程。

但是，教育测量与教育评价有着根本区别。教育测量本质上是一个事实判断过程，而教育评价实质上是一种价值判断过程。它们的区别主要表现在以下两个方面。

第一，由于教育测量是对事实作判断，在判断的法则（标准）确定后，如果排除测量误差的影响，则不同的人进行测量应能得到相同的结果，即教育测量具有较强的客观性；而教育评价是对教育活动的价值做出判断，由于评价主体的价值观念和标准有所不同，因此判断的结果可能是不相同的。

第二，教育测量是在事实判断基础上，进行赋值的过程，因此它注重量化，但教育评价既有定量的评价，也有定性的评价，就是说测量的结果是评价的主要依据之一，评价的价值判断标准是多方面的。如某学生英语期末考试获得 84 分，全年级 60% 的人低于他。这对于教育测量



来说,任务已完成。因为分数已描述了该生当前的英语水平。但是对于教育评价者来说,还应考查该生以往的成绩,即与以往成绩相比是进步了还是倒退了。若是退步了就要找出他退步的原因,分析他哪部分知识、技能有缺陷,应如何补救。可见,教育评价不仅要当前结果做出描述,还要考查其发展过程,诊断其症结所在,提出补救措施。

教育评价和教育测量的关系可表述如下:

$$\text{教育评价} = \begin{cases} \text{测量 (量的记述) + 价值判断} \\ \text{非测量 (质的记述) + 价值判断} \end{cases}$$

(二) 教育评价和教育评估

教育评价和教育评估是非常接近的两个概念,在许多场合下是可以通用的。事实上,从认识论的角度分析,评价和评估的本质是精神对物质、意识对存在的一种反映。在教育领域中,二者均从一定社会的角度出发,按照一定的教育价值观来考查和评定被评价对象的社会价值,判断其好与坏、功与过、正确与不正确及其表现的程度。

在实际中具体使用时,不同的范围和场合又有不同的习惯用法,如高等教育领域多称教育评估,在督导部门也称督导评估,而在普通教育领域则多称为教育评价。在我国大陆教育理论界把“evaluation”多译作“评价”,有时也译作“评估”,似已成习惯,而在我国的台湾省,常译作“评鉴”,也有译作“评量”或“评价”的。在我国正式公布的文件中有时用“评价”一词,有时用“评估”一词。可见,教育评价和教育评估在使用中并无严格的界限。

第二节 教育评价的主要类型

教育评价涉及的范围广、内容多,门类繁杂,为便于研究和应用,可依据不同的标准做出分类。

一、按评价范围分类

(一) 宏观的教育评价

宏观的教育评价是以教育的全领域或涉及宏观决策方面的教育现象、措施为对象的教育评价。如对教育目标、教育结构、教育制度、教育内容、教育方法、教育行政管理、教育社会效益等方面的评价。它是总体的、全局性的、战略的、宏观的、高层次的。

(二) 中观的教育评价

中观的教育评价是以学校内部各方面工作为对象的教育评价。评价的内容包括学校的办学条件、办学水平、领导班子、教师队伍、思想政治教育工作、教学工作、体育卫生工作、总务工作、团队工作、家长工作、学校社会效益等方面。

(三) 微观的教育评价

微观的教育评价是以学生的发展变化为对象的教育评价。评价的内容包括对学生的思想品德、知识技能、健康状况、审美情操、劳动技能等方面。

二、按评价的内容分类

(一) 条件评价

条件评价是对教育方案可行性的评价，也就是对达到目标所需条件的评价。为了识别各备选方案的优劣，需要对实现目标所需成本费用、可利用的人力和物力资源，解决问题的策略和相应的程序设计等进行调查研究。为了设计实现目标的最佳方案，可以把两个或多个竞争性策略中的最好方面结合起来，增强其可操作性。