

21

世纪高等院校教材·师范类

# 化学教育测量与评价

周青 主编  
陆国志 杨辉祥 副主编

 科学出版社  
www.sciencep.com

21 世纪高等院校教材·师范类

# 化学教育测量与评价

周 青 主 编

陆国志 杨辉祥 副主编

科学出版社

北 京

## 内 容 简 介

教育测量与评价是当今世界教育科学研究的三大领域之一。化学教育测量评价是根据化学教育的目的和标准,通过系统的调查、比较和收集资料,对化学教育活动的过程、结果和有关因素进行描述或者价值判断的过程。本书根据目前我国教师教育中存在评价理念落后、评价能力弱和评价素质差,无法适应基础教育发展需要的现实,以新的教育理论与新课程理念为指导,系统介绍化学教育测量与评价的基本原理。全书共十章,分别介绍化学教育测量的基本原理和方法,新的课程评价观、评价方法、评价模式、教育统计基础、化学测量工具的编制与评价、专题评价(包括新课程理念下的化学课堂教学评价、化学教师评价、化学实验评价)。

本书适合从事中学化学教学的教师与教研人员,大学相关专业高年级学生、教师及研究生使用。

### 图书在版编目(CIP)数据

化学教育测量与评价/周青主编.—北京:科学出版社,2006

(21世纪高等院校教材·师范类)

ISBN 7-03-016842-9

I. 化… II. 周… III. ①中学-化学课-教育评估-高等学校-教材②中学-化学课-教育测验-高等学校-教材 IV. G633.82

中国版本图书馆CIP数据核字(2006)第007818号

责任编辑:王志欣 丁里 宛楠/责任校对:朱光光

责任印制:张克忠/封面设计:耕者工作室

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

新蕾印刷厂 印刷

科学出版社发行 各地新华书店经销

\*

2006年5月第一版 开本:BS(720×1000)

2006年5月第一次印刷 印张:19 1/2

印数:1—3 000 字数:365 000

定价:26.00元

(如有印装质量问题,我社负责调换〈环伟〉)

## 前 言

新的基础教育课程改革要求考试评价不仅要“关注学生的学业成绩,而且要发现和发展学生多方面的潜能,了解学生发展中的需要,帮助学生认识自我,建立自信”,以及要“改变课程评价过分强调甄别与选拔的功能,发挥评价促进学生发展、教师提高和改进教学实践的功能”。教育测量与评价是当今世界教育科学研究的三大领域之一。它不仅在教育科学体系中占有显赫的地位,而且在教育教学过程中具有重要而广泛的应用价值;在实践中具有强烈的导向作用,常常成为教育改革的突破口。教育测量与评价是研究如何测量学生的学业成绩、学习能力、兴趣爱好、思想品德以及教育上许多措施的绩效等问题的一门学科,它是改进教学、实行科学化教育管理和教育研究的重要工具。

把教育测量、教育评价的一般原理和方法应用于化学教育领域,逐步形成了化学教育测量与化学教育评价。化学教育测量与评价是根据一定的教育目的和标准,通过系统的调查、比较和收集资料,对化学教育活动的过程、结果和有关因素进行描述或者价值判断,并且通过协调取得有关方面对评价对象的共识,为提高化学教育活动的质量、效率等提供依据的过程。这个定义说明了化学教育评价的根据、方法、过程、性质和终结目的。学习和应用化学教育测量与评价,有助于教师深入地了解学生的发展变化,恰当地做出教学决策,改进教学,提高教学质量;有助于化学教育工作者认识、发现化学教与学的规律,使化学教育及其研究、管理工作实现从量性到质性的跨越,从经验到理论的升华。

当代教育的改革和发展已经开始聚焦在变革学校课程与教学的层面上,当前我国正在进行的基础教育课程改革就在相当程度上回应了我国教育所面临的挑战和现实问题。课程与教学的变革和发展需要一线教师的积极参与,教师自身的理论素养和实践能力是决定课程与教学改革成败的关键,教师发展是课程开发的中心,教师的评价素质是教师专业化的重要组成部分。新课程在课程理念、课程目标、课程功能、课程结构、课程内容、课程实施、课程评价及课程管理等方面都发生了重大变革,较原来的课程有了重大创新和突破。它既要求师范院校在现有的教师培养过程中融入新课程理念,使未来的教师能够了解新课程、理解新课程,又要求一线教师更新教育观念,改进原来习以为常的教学方法、教学行为和教学手段,重新确立和认识自己的角色,提高教师专业化水平。

我们编写的《化学教育测量与评价》一书在继承传统教育测量与评价精华的基础上,突破传统教育测量与评价的局限,体现当代基础教育课程改革指导纲要的基

本思想,以新的教育理念贯穿全书。本书共十章,分别介绍化学教育测量的基本原理和方法,新的课程评价观、评价方法、评价模式、教育统计基础、化学测量工具的编制与评价、专题评价(包括新课程理念下的化学课堂教学评价、化学教师评价、化学实验评价)。

我们在编写过程中试图使本书成为教与学的有效工具,力求使本书成为启迪思维、开阔视野、深入探究、培养方法的好材料。我们特意在每章中选择了适量的案例,便于读者理解、应用所学理论与方法。

本书由周青教授提出全书各章节框架及其编写体例要求并统稿、审定,陆国志副教授编写第一、二、四章和第六章,杨辉祥副编审编写第三、五章并参与审定工作,周青教授编写第七~十章。

在编写过程中我们参阅了国内外同行大量的文献,也引用了他们的一些研究成果与数据资料,在这里表示深深的谢意。

由于编者水平所限,再加上时间仓促,书中不妥或疏漏之处在所难免,敬请读者批评指正。

编 者

2005年9月于西安

# 目 录

## 前言

<b>第一章 化学教育测量与评价概述</b> .....	(1)
<b>第一节 教育测量与评价的基本问题</b> .....	(2)
一、教育测量与评价的发展历史.....	(2)
二、教育测量与评价的含义.....	(4)
三、教育评价的基本问题.....	(5)
四、教育评价相关概念辨析.....	(6)
<b>第二节 化学教育测量与评价的意义</b> .....	(11)
一、化学教育测量与评价的功能.....	(11)
二、化学教育测量与评价研究特点.....	(12)
三、化学教育评价的动态性.....	(16)
四、化学教育评价机制在化学教学管理中的应用.....	(25)
<b>第三节 美、英两国教育考试与评价现状</b> .....	(27)
一、美国教育考试与现状.....	(27)
二、英国教育考试与现状.....	(34)
<b>第四节 改革化学考试制度,培养创新人才</b> .....	(36)
一、现行化学考试存在的主要问题.....	(36)
二、现行考试对化学教学的影响.....	(36)
三、改革现行化学考试制度的对策.....	(36)
阅读材料:初中化学课程评价方案的研究.....	(38)
练习与思考题.....	(46)
参考文献.....	(46)
<b>第二章 课程评价</b> .....	(47)
<b>第一节 课程评价概述</b> .....	(47)
一、课程的概念.....	(47)
二、课程评价的定义.....	(49)
三、课程评价发展史.....	(49)
四、课程评价的目的.....	(54)
五、课程评价的内容.....	(55)
<b>第二节 课程评价的模式及其本质</b> .....	(55)

一、课程评价模式的基本概念 .....	(55)
二、课程评价的本质 .....	(56)
三、课程评价的基本模式 .....	(58)
四、确定课程评价模式的评价指标的原则 .....	(61)
<b>第三节 课程评价的原则和方法</b> .....	(61)
一、课程评价的原则 .....	(61)
二、课程评价的类型 .....	(62)
<b>第四节 新的课程评价观</b> .....	(74)
一、当前我国课程评价的误区 .....	(74)
二、课程评价发展的特点 .....	(75)
三、发展性课程评价的特点 .....	(76)
阅读材料:课程评价的范式 .....	(87)
练习与思考题 .....	(88)
参考文献 .....	(89)
<b>第三章 化学教育测量工具</b> .....	(90)
<b>第一节 教育目标的分类研究</b> .....	(90)
一、布卢姆的教育目标分类理论 .....	(90)
二、加涅的教育目标分类理论 .....	(92)
三、梅里尔的教育目标分类理论 .....	(93)
<b>第二节 化学教育目标分类</b> .....	(94)
一、化学教育目标的概念 .....	(94)
二、化学教育目标设计的基本原则 .....	(95)
三、中学化学教育目标的学习水平分类 .....	(97)
四、化学教育目标的制订 .....	(99)
<b>第三节 教育测量的基本原理</b> .....	(102)
一、教育测量的特点 .....	(102)
二、教育测量的基本步骤及要求 .....	(104)
<b>第四节 命题计划的设计</b> .....	(107)
一、命题计划的意义与作用 .....	(107)
二、设计命题计划所要解决的主要问题 .....	(108)
三、命题计划的构成与使用 .....	(110)
<b>第五节 不同类型项目的编制</b> .....	(111)
一、项目的类型 .....	(111)
二、选择题、填空题及简答题的编制技巧 .....	(113)
<b>第六节 不同学习水平项目的编制</b> .....	(116)

一、记忆水平的项目 .....	(117)
二、理解水平的的项目 .....	(117)
三、简单应用水平的的项目 .....	(118)
四、综合应用水平的的项目 .....	(118)
<b>第七节 项目编排与试题的确定</b> .....	(119)
一、编排项目合成试题 .....	(119)
二、项目权重(分数)的确定 .....	(120)
三、完成试题双向细目表 .....	(121)
四、试题“等值复本”编制 .....	(122)
五、预试和试题的最后确定 .....	(123)
<b>第八节 问卷调查的测量学思考</b> .....	(123)
一、问题的提出 .....	(123)
二、问卷设计中的理论构思 .....	(124)
三、调查问卷的量化过程 .....	(126)
四、调查问卷信度和效度的考证 .....	(127)
练习与思考题 .....	(128)
参考文献 .....	(128)
<b>第四章 化学测量工具的实例</b> .....	(129)
<b>第一节 创新性化学习题的设计</b> .....	(129)
一、创新性化学习题的主要类型举例 .....	(129)
二、创新性习题设计遵循的原则 .....	(133)
<b>第二节 化学开放题及其编制方法</b> .....	(135)
一、化学开放题的有关理念 .....	(135)
二、化学开放题的编制方法 .....	(136)
<b>第三节 化学测量工具评价的案例</b> .....	(139)
一、命题指导思想 .....	(139)
二、试卷统计分析 .....	(139)
三、试题分析 .....	(140)
四、答卷中反映的问题 .....	(143)
五、对中学教学的启示 .....	(144)
练习与思考题 .....	(153)
参考文献 .....	(153)
<b>第五章 教育统计基础</b> .....	(154)
<b>第一节 数据的初步整理</b> .....	(154)
一、数据的概念 .....	(154)



二、数据的种类·····	(155)
三、统计图·····	(157)
第二节 统计相关概念辨析·····	(162)
一、算术平均数·····	(162)
二、方差与标准差·····	(163)
三、概率·····	(163)
四、正态分布·····	(165)
五、显著性水平·····	(166)
第三节 推断统计·····	(167)
一、统计检验程序·····	(167)
二、实验设计·····	(167)
第四节 SPSS 软件在教育统计中的应用·····	(169)
一、SPSS 简介·····	(169)
二、SPSS 在教育统计中的应用·····	(171)
三、教育统计常见问题类型及解决方案·····	(178)
阅读材料:高中化学新教材中科学方法教育的实施简述·····	(179)
练习与思考题·····	(181)
参考文献·····	(182)
<b>第六章 化学测量工具的评价·····</b>	<b>(183)</b>
第一节 试题质量的初步评价·····	(183)
一、测验分数的分布状态·····	(183)
二、平均分·····	(185)
三、标准差·····	(186)
第二节 试题的信度·····	(186)
一、相关系数·····	(186)
二、信度的概念·····	(187)
三、几种常用信度的意义及计算方法·····	(188)
第三节 试题的效度·····	(191)
一、效度的概念·····	(191)
二、几种常用的效度及估计方法·····	(191)
第四节 项目质量评价·····	(194)
一、项目的难度分析·····	(194)
二、项目的区分度分析·····	(197)
三、选择题的反应模式分析·····	(200)
第五节 调查问卷的生成及质量分析·····	(201)

一、初步整理分析 .....	(202)
二、信度分析 .....	(202)
三、效度分析 .....	(203)
四、区分度分析 .....	(203)
五、对问卷进行总结,找出其不足,以待改进 .....	(204)
六、发展问卷的当地及全国常模 .....	(204)
七、化学测量工具评价的案例 .....	(204)
练习与思考题 .....	(212)
参考文献 .....	(213)
<b>第七章 化学测试结果的分析与评价</b> .....	(214)
<b>第一节 学生个体的测验成绩评价</b> .....	(214)
一、以常模参照评价学生的测验分数 .....	(214)
二、以目标参照评价学生的测验分数 .....	(217)
三、评价学生成绩时应注意的问题 .....	(219)
<b>第二节 学生集体测验分数的评价</b> .....	(220)
一、根据分数分布估计测验成绩 .....	(220)
二、根据平均分和标准差估计测验成绩 .....	(221)
三、不同学生集体的测验分数比较 .....	(222)
练习与思考题 .....	(223)
参考文献 .....	(223)
<b>第八章 促进学生发展的课堂教学评价</b> .....	(224)
<b>第一节 课堂教学评价概述</b> .....	(224)
一、对评价及课堂教学评价的认识 .....	(224)
二、促进学生发展的课堂教学评价的理论基础 .....	(225)
<b>第二节 促进学生发展的课堂教学评价</b> .....	(227)
一、新课程改革下,课堂教学的改革背景 .....	(227)
二、课堂教学的要素分析 .....	(228)
三、课堂教学评价主体的确定——注重多元化 .....	(229)
四、课堂教学评价方式的选择——体现多样性 .....	(229)
五、课堂教学评价过程的展开——强调人文性和发展性 .....	(230)
六、课堂教学评价标准的确立 .....	(231)
<b>第三节 中学化学课堂教学评价的主要功能</b> .....	(234)
一、全面落实化学教学目标 .....	(235)
二、全面评价教学水平,激发教学积极性 .....	(235)
三、帮助教师改进教学方法,提高教学水平 .....	(235)

第四节 中学化学课堂教学评价及案例分析·····	(236)
一、在确定中学化学课堂教学评价指标时遵循的主要原则·····	(236)
二、化学课堂教学质量评价量表的编制·····	(237)
三、课堂教学评价案例分析·····	(238)
练习与思考题·····	(243)
参考文献·····	(243)
<b>第九章 教师评价</b> ·····	(244)
第一节 教师评价概述·····	(244)
一、教师评价的概念·····	(244)
二、教师评价的历史发展·····	(246)
三、教师评价的历史经验与教训·····	(248)
第二节 教师评价的种类与模式·····	(249)
一、教师评价的种类·····	(249)
二、教师评价的模式·····	(250)
三、教师评价的原则·····	(251)
四、教师评价的目的·····	(252)
五、教师评价的功能·····	(253)
第三节 发展性教师评价:教师评价的新理念·····	(253)
一、发展性教师评价的特点·····	(254)
二、发展性教师评价的意义·····	(255)
三、发展性教师评价的原则·····	(256)
四、发展性教师评价的评价指标体系·····	(257)
五、发展性教师评价的方案·····	(259)
第四节 新课程观下教师评价的特征·····	(260)
一、评价目的:促进教师专业发展·····	(260)
二、评价功能:强调展示成就与改进激励·····	(261)
三、评价内容:突出综合素质、重视个体差异·····	(262)
四、评价方法:注重质性评价、主张自我反思·····	(263)
五、评价主体:以自评为主、各方协同参与·····	(264)
第五节 化学教师的评价方式·····	(265)
一、学习成绩的评定·····	(265)
二、自我教学评定·····	(266)
练习与思考题·····	(268)
参考文献·····	(268)
<b>第十章 化学实验教学目标及其评价</b> ·····	(269)

---

第一节 化学实验教学评价观的发展·····	(269)
一、化学实验教学评价观的演变·····	(269)
二、传统化学实验教学评价的主要弊端·····	(271)
三、化学实验教学评价的发展方向·····	(272)
第二节 化学实验教学目标·····	(272)
一、实验观察能力培养·····	(273)
二、实验操作技能培养·····	(274)
第三节 化学实验教学中的评价方法·····	(277)
一、分组实验教学评价方法·····	(277)
二、活动表现评价方法·····	(279)
三、化学实验教学评价的新发展——档案袋评价方法·····	(284)
第四节 化学实验操作技能测评·····	(287)
一、测评量表的设计·····	(288)
二、测试过程·····	(289)
三、测试结果分析与讨论·····	(289)
练习与思考题·····	(291)
参考文献·····	(291)
附表 1 正态分布表·····	(293)
附表 2 $t$ 值表·····	(297)

# 第一章 化学教育测量与评价概述

1996年,“国际21世纪教育委员会”向联合国教科文组织提交了一份报告。15名来自各国的权威专家经过3年多的研究,根据现代社会面对的矛盾和未来教育面临的挑战提出要适应未来社会的发展,教育必须围绕“四种”基本学习能力来重新设计、重新组织,即学会认知(learn to know)、学会做事(learn to do)、学会共同生活(learn to live together)、学会生存(learn to be)。国家新一轮基础教育课程改革,也要求我们转变教育思想,树立新的课程观、教学观、质量观和考试评价观。教

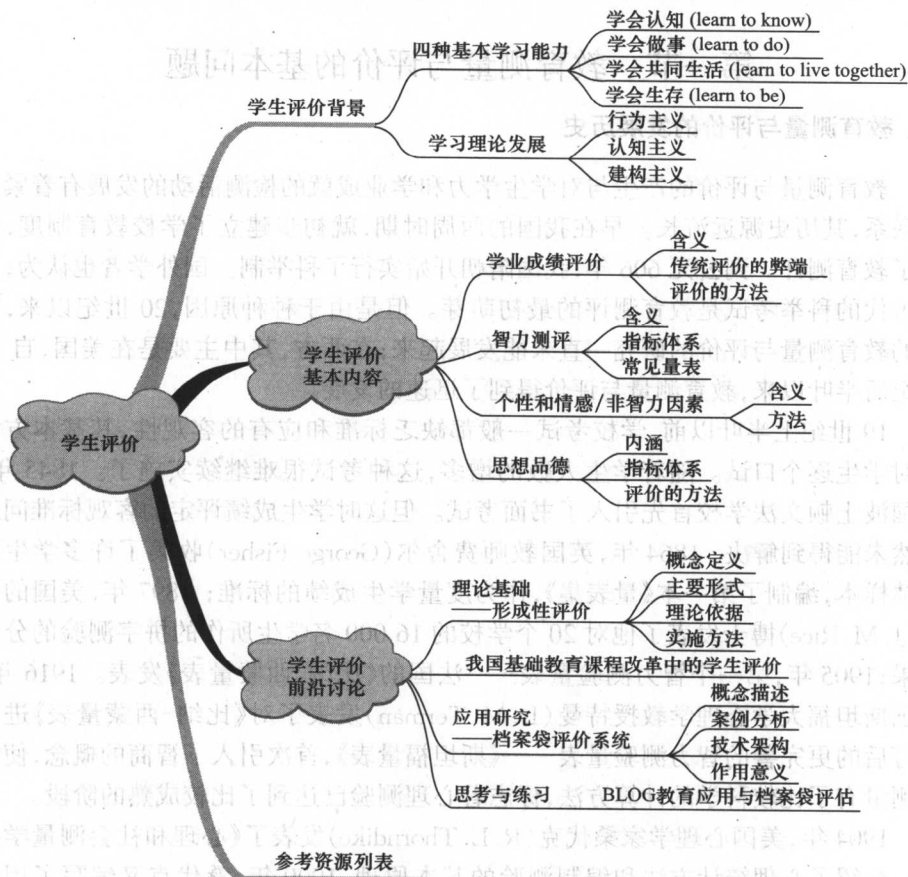


图 1-1 学生评价概念图

育测量与评价是当今世界教育科学研究的三大领域之一。它不仅在教育科学体系中占有显赫的地位,在教育教学过程中具有重要而广泛的应用价值,而且在实践中具有强烈的导向作用,常常成为教育改革的突破口。要实现上述“四种”基本的学习能力,全面推进素质教育,掌握和了解教育测量与评价的基本原理与方法是为关键的。教育测量与评价是我们实现教育目标的有力保证体系。图 1-1 是在新的教育理念指导下学生评价概念图。

可以看出,教育测量与评价涉及学生发展的很多方面,是一个非常复杂、重要的问题。教育要实现真正地促进学生的全面发展,尊重学生的个性,挖掘学生的潜力,必须重视教育测量与评价的理论与方法研究。那么,教育测量与评价的历史渊源怎样?什么是教育测量与评价?他们与教育测验、教育考试、教育评估等概念有什么联系和区别?在这一章中,我们将对这些基本问题做一简单探讨。

## 第一节 教育测量与评价的基本问题

### 一、教育测量与评价的发展历史

教育测量与评价的产生与对学生学力和学业成就的检测活动的发展有着紧密的联系,其历史源远流长。早在我国的西周时期,就初步建立了学校教育制度,开始了教育测评。到公元 606 年,我国隋朝开始实行了科举制。国外学者也认为,中国古代的科举考试是教育测评的最初萌芽。但是由于种种原因,20 世纪以来,我国的教育测量与评价的研究一直未能发展起来;在西方,其中主要是在美国,自 19 世纪后半叶以来,教育测量与评价得到了迅速的发展。

19 世纪上半叶以前,学校考试一般都缺乏标准和应有的客观性,其基本方法是对学生逐个口试。随着学生人数的增多,这种考试很难继续实施了。1845 年,美国波士顿文法学校首先引入了书面考试。但这时学生成绩评定的客观标准问题仍然未能得到解决。1864 年,英国教师费舍尔(George Fisher)收集了许多学生的成绩样本,编制了第一本《量表集》,作为度量学生成绩的标准;1897 年,美国的莱斯(J. M. Rice)博士发表了他对 20 个学校的 16 000 名学生所作的拼字测验的分析结果;1905 年,第一个智力测验量表——法国的《比纳-西蒙量表》发表。1916 年,美国斯坦福大学心理学教授特曼(L. M. Terman)发表了对《比纳-西蒙量表》进行修订后的更完善的智力测验量表——《斯坦福量表》,首次引入了智商的概念,使智力测量有了比较科学的计算方法,标志着心理测验已达到了比较成熟的阶段。

1904 年,美国心理学家桑代克(R. L. Thorndike)发表了《心理和社会测量学导论》,介绍了心理统计方法和编制测验的基本原理,1909 年,桑代克又编写了用于书法、拼字、作文、图画测验的标准测量工具,使教育测量开始走上了科学化的道路。因此,桑代克被称为教育测量学的鼻祖。1918 年以后,教育测量的使用范围

逐渐从小学发展到中等以上学校,许多大学也开始设立了教育测量学课程。至 20 世纪 20 年代末,教育测量已发展到全盛期,教育测量在学力检测与教育成就的量化、客观化、标准化方面取得了很大的成绩。

随着教育测量运动的发展,其弱点也逐步暴露出来。由于当时的测验都是围绕着教科书编制的,测验内容只要求学生记忆教材的知识内容,未能包括对学生的社会态度、兴趣、情感的检查,不能反映学生的全面发展和实际需要,因而它很快引起了人们的不满和批判。

1929 年,美国经历了一场严重的经济危机,教育的社会效果问题越来越多地受到人们的关注,这就直接推动了以教育的社会价值为依据的教育评价活动的发展。从 1934 年开始到 1942 年,美国俄亥俄州立大学教授泰勒(R. W. Tyler)受卡内基基金会的资助,进行了历时 8 年的课程与评价研究,这就是教育评价发展史上著名的“八年研究”。在对以往的课程和测验设计进行了尖锐批评的基础上,泰勒提出了一套以教育目标为核心的课程和测验编制原则,试图以此把社会的要求、学生的需要反映在课程与测验中。为了把这一思想与早期的测量区别开来,泰勒和他的同事正式提出了教育评价的概念:“在本质上,评价过程乃是一种测量课程和教学方案在多大程度上达到了教育目标的过程”,即认为教育评价就是衡量教育活动达到教育目标程度的一种活动,测量是它的手段。根据泰勒的理论进行的评价活动能够获得关于目标完成情况的信息,有助于发现存在的问题,改进教育工作。所以,这一理论受到了较为广泛的欢迎。

泰勒等的教育测量和评价观点以及相应的“目标评价模式”是教育测评历史上第一个比较完整的、影响较大的理论。自 20 世纪 30 年代以来的几十年中,这一理论一直占据着主导地位。为了发展泰勒评价模式的实施技术,美国教育心理学家布卢姆(B. S. Bloom)等提出了教育目标分类学理论,对完善教育评价理论起了重要作用。但随着测评实践和理论研究的发展,这一理论也逐渐暴露出一些根本性的缺陷。泰勒理论的核心是目标,但由于教育目标概念与教育结果的概念紧密相关,这就容易使人们将更多的注意力集中到对教育结果的评价上。因此,泰勒评价更多地被看成是一种对结果的总结性评价;事实上,形成性评价可能有着更重要的作用。此外,人们还认识到,目标本身的合理性也需要评价,对那些预期目标以外的教育效果也不能排除在评价之外。

在对泰勒评价观进行分析和批判的基础上,又有人提出了新的教育评价观。克龙巴赫(L. J. Cronbach)修改了泰勒的评价定义,将评价定义为:“为做出关于教育方案的决策,收集和使用信息。”他强调:“评价能完成的最大贡献是确定教程需要改进的方面。”斯塔佛比姆(D. L. Stufflebeam)给评价的定义是:“为决策提供有用信息的过程。”并强调:“评价最重要的意图不是为了证明(prove),而是为了改进(improve)。”在这一观点的基础上,他提出了著名的 CIPP(context-input-process-

product)模式,即把对背景、输入、过程、结果四个方面的评价结合起来的一种评价模式。受斯塔佛比姆等的影响,美国教育评价标准委员会给评价的定义是:“评价是对某些现象的价值和优缺点的系统调查。”

## 二、教育测量与评价的含义

### (一) 测量的含义

什么是测量? S. S. Stevens 认为,“广义而言,测量(measurement)就是根据法则赋予事物数量。”也就是说,按照一定的规则给事物的属性指派数字或符号的过程就是测量。这是迄今为止公认的测量定义。例如,要测量一下桌子的高度,我们可以拿尺子来量一量,看看它有多高。尺子是人们根据一定的法则制定的量具,利用它就可以把事物的属性,即桌子的高度用数字表示出来,如 0.75m。这种测量属于客观测量,因为它基本上不受观察者的主观判断的影响。

### (二) 教育测量的含义

人们不仅能对事物的许多物理、化学属性如长度、质量、距离、体积、温度、硬度、组成成分及运动速度等做出越来越准确的测量,也能够对人自身的许多生理及心理特性如身高、体重、血压、血型、脉搏、肺活量以及人的知识水平、能力、气质、性格、兴趣、态度等精神特性进行测量,这些精神特性方面的测量,不仅表现在学校教育领域,而且涉及社会许多部门。教育测量,就是针对学校教育影响下学生各方面的发展,侧重从量的规定性上予以确定和描述的过程。学校教育实践活动的客观需要,促进了教育测量科学研究及学科发展。教育测量学是一门发展较早、应用较多、内容较丰富的教育科学分支。

教育测量是为了了解学生的发展,尤其是为评定学习成绩而进行的测量活动。这是教育测量活动最原始的动机,也是教育测量学科发展最早的立足点。在学校教育背景下,通过课程、教师和学生的相互作用进行教学,无论是教育者还是受教育者,都需要了解学生学到些什么,掌握程度如何,学习者有哪些变化,学生的成绩和发展进步用什么方式加以考核记载,所有这些基本的、原始的教育工作要借助于教育测量活动加以实现。

教育测量关注学校的教学效果,而教学效果是教与学双方共同作用的结果。因此,教育测量的结果在用于评定学生学习效果的同时,还被用于了解教师的课堂教学效果。一句话,教育测量反馈的是关于课堂教与学两方面的信息。

教育测量关注学生的发展,而学生的发展是多方面的。除了掌握学科的知识与技能外,还要发展学生的道德、情感、态度、价值观、兴趣、思维能力、实践能力、创造能力等方面。一句话,教育测量涉及学生在德育、智育、体育、美育、劳动技能以及个性、心理素质等许多方面。然而,由于许多身心特性存在着复杂性和模糊性,



使得教育测量比物理测量有更多的困难,教育测量的结果也不像物理测量的结果那么容易理解。

### 三、教育评价的基本问题

#### (一) 评价的含义

在我们的日常生活和文化活动中,经常用到“评价”(evaluation)这个词语。从评价某人的烹调手艺,到评价某人的为人处世;从评价某学生的科技作品,到评价某作家的文学新作;从评价某高校的办学指导思想是否明确以及学校定位是否合理,到评价一所高校的整体办学水平等。可以说,无处不在使用着“评价”这个词语,我们每个人自觉不自觉地参与评价活动或处于被评价的位置上,学校教育工作者尤其如此。

那么,何谓评价?广义地讲,评价泛指衡量、判断人物或事物的价值。评价活动的过程是对人物或事物的价值进行分析、衡量和判断的过程。在评价过程中,无论是事实判断还是价值判断,都需要以事实为依据,通过收集多方面的资料证据,对人类社会活动的效果、物质产品和精神产品的质量及价值等做出判断。

#### (二) 教育评价的概念

当把“评价”一词特别地用于学校教育领域或课堂教学情境时,在一些情况下,“评价”就是“教育评价”一词的简称;在另一些情况下,它指的是教育目标类中最高层次的认知能力水平——评价。下面是国内外一些学者对“评价”在教育教学活动情境下的若干典型的解释:

(1) 格兰朗德(N. E. Greenland)认为,评价是为了确定学生达到教学目标、程度、收集、分析和解释信息的(课堂)系统过程;评价包括对学生的定量描述(测量)和定性描述(非测量)两方面。根据格兰朗德的观点,评价总是包括对测量结果需求程度的价值判断(例如,玛丽学习数学取得了长足的进步)。一个完整评价计划将包括测量和非测量两种方法,用公式加以形象地表达,即:评价 = 测量(定量描述) + 非测量(定性描述) + 价值判断。

(2) 斯塔佛比姆等认为:“评价是一种划定、获得和提供叙述性和判断性信息的过程。”这些信息涉及研究对象的目标、设计、实际影响的价值及优缺点,以便指导如何决策、满足教学效能核定的需要,并增强对研究对象的了解。斯塔佛比姆还说,“评价最重要的意图不是为了证明,而是为了改进”。

(3) 美国教育评价标准委员会曾对“评价”给出一个简明的定义,即“评价是对某些现象的价值如优缺点的系统调查,为教育决策提供依据的过程”。

(4) 泰勒指出:“评价过程在本质上是确定课程和教学大纲在实际上实现教育目标的程度的过程。”