



教育类专业基础课系列教材

第二版

# 教育测量与评价

主编◎黄光扬



上海市  
著名商  
标

华东师范大学出版社

全国百佳图书出版单位

# 教育测量与评价

第 1 版



教育类专业基础课系列教材

第二版

# 教育测量与评价

主编◎黄光扬



华东师范大学出版社

## 图书在版编目(CIP)数据

教育测量与评价/黄光扬主编. —2版. —上海:华东师范大学出版社,2012.2

教育类专业基础课系列教材

ISBN 978-7-5617-9297-1

I. ①教… II. ①黄… III. ①教育测验—高等学校—教材 ②教育评估—高等学校—教材 IV. ①G449

中国版本图书馆 CIP 数据核字(2012)第 023091 号

教育类专业基础课系列教材

## 教育测量与评价(第二版)

主 编 黄光扬  
责任编辑 吴海红  
责任校对 王丽平  
装帧设计 卢晓红

出版发行 华东师范大学出版社  
社 址 上海市中山北路 3663 号 邮编 200062  
网 址 [www.ecnupress.com.cn](http://www.ecnupress.com.cn)  
电 话 021-60821666 行政传真 021-62572105  
客服电话 021-62865537 门市(邮购)电话 021-62869887  
地 址 上海市中山北路 3663 号华东师范大学校内先锋路口  
网 店 <http://hdsdcbs.tmall.com>

印 刷 者 浙江临安曙光印务有限公司  
开 本 787×1092 16 开  
印 张 21.5  
字 数 460 千字  
版 次 2012 年 5 月第一版  
印 次 2012 年 5 月第一次  
书 号 ISBN 978-7-5617-9297-1/G·5554  
定 价 41.00 元

出 版 人 朱杰人

(如发现本版图书有印订质量问题,请寄回本社客服中心调换或电话 021-62865537 联系)

## 第二版前言

教育测量与评价是所有成功教学的基础,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,也是学校和政府做出诸多教育决策的重要依据。由于教育测量与评价在实践中具有强烈的导向作用,因此,教育测量与评价常常成为教育改革的突破口和课程改革的支撑点之一。正因为如此,在教育领域乃至社会各界已有越来越多的人士关注着教育测量与评价的学科发展。

为适应 21 世纪高等师范院校“教育测量与评价”课程教学以及我国各级各类学校进行考试与评价改革的需要,我们在“21 世纪高等师范教育教材编写委员会”的授权下,组织编写了《教育测量与评价》一书,由华东师范大学出版社于 2002 年 8 月第一次出版发行。本教材自出版以来,被许多高等院校以及教师教育培训机构选做教材,个别省市把本教材选做自学考试有关开考专业的指定教材,许多教育工作者也把本书作为学习和研究的书籍,教育界许多人士和读者对本教材给予充分的肯定,使本教材在专业面比较窄的客观事实面前,保持了一定的读者群,并在短短几年中重印近 10 次。

根据社会发展以及读者的建议,我们对教材各章进行了与时俱进的修订和必要的校正,着重对教材内容与材料进行更新和补充,并对概念和问题作更加科学的阐述和说明。包括:教育部普通高校本科教学工作水平评估制度和评估方案;教育部以及国家体育总局最新颁布的《中国学生体质健康评价标准》;考试命题的教育测量法则和人文关怀;我国传统考试评价的文化特点反思;表现性测验评价原理与方法;教育测量与评价在国家教育质量监控以及教育管理体制改革中的应用;国际学生评价项目(PISA)新进展等许多方面。

本教材经过修订,进一步凸显了以下几个特点:一是体现时代精神,弘扬民族优秀传统文化思想;二是体现教育测量与评价学科内容整合的趋势;三是体现教育测量与评价学科内容的丰富性、发展性以及实际应用的广阔性和重要性;四是较好地处理了学科内容和方法的继承、发展与创新的关系;五是保

持体例创新、问题更新,增强了教材的可读性、启发性、教育性和科学性。

本教材由黄光扬任主编并主持修订。在编写和修订过程中,由黄光扬提出全书各章节框架结构、编写体例要求和修订思路,并邀请刘尧和龙文祥担任副主编,然后分工编写和修订,最后由黄光扬负责全书的修改、统稿和统校。本书各章节编写和修订分工如下(按章顺序):第一章、第六章、第七章、第八章、第九章由福建师范大学黄光扬执笔(其中第九章第五节由林云鹏执笔);第二章、第三章由江西师范大学董圣鸿执笔;第四章、第五章由安徽师范大学龙文祥执笔;第十章、第十三章由华南师范大学刘晓瑜执笔;第十一章由浙江师范大学刘尧和傅宝英执笔;第十二章由浙江师范大学刘尧和田里执笔。

在教材编写和修订过程中,我们参阅了国内外同行的最新资料,也引用了他们的一些新的研究成果与数据资料。对此,本书尽可能给予详细的注明,但由于受到条件限制事先未能一一致函征询,在这里表示歉意和深深的谢意。本书在修订过程中得到了华东师范大学出版社的大力支持,特别是赵建军和吴海红两位编辑为本教材的修订再版付出许多努力,在此表示感谢。由于我们水平所限,编写和修订的经验不足,书中可能还有不妥或错漏之处,敬请各位专家同行与读者批评指正。

编写和修订者

2012年2月

# 第一版前言

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。由于教育测量与评价在实践中具有强烈的导向作用,因此,教育测量与评价常常成为教育改革的突破口,也往往成为课程改革的支撑点之一。新世纪新理念,改革与创新是关键。为适应21世纪高等师范院校“教育测量与评价”课程教学以及我国各级各类学校进行考试与评价改革的需要,我们在“21世纪高等师范教育教材编写委员会”的授权下,组织编写了《教育测量与评价》一书。这本新教材,具有如下几个鲜明的特点:

## (一) 体现时代精神与要求

全面推进素质教育,尤其是国家新一轮基础教育课程改革,要求我们转变教育思想观念,树立新的课程观、教学观、质量观、人才观和考试评价观。本教材贯彻素质教育的思想,体现基础教育新课程改革的理念与精神,使考试评价不仅要“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”。

## (二) 体现学科内容整合的趋势

当前,教育测量和教育评价这两个学科内容具有相互整合的趋势。我们在教材的第一章中以翔实的资料,有力地说明了这一点。本教材用“教育测量与评价”作为书名,其用意也在于此。考虑到篇幅限制和教材内容的学术水准,本教材侧重于探讨教育测量与评价的基本原理和方法,以及把这些基本原理和方法应用于学生的课业发展、智能发展、创新能力发展、思想品德发展、体育发展等方面的测量与评价。本教材在整合学科内容、拓展研究广度的同时,把握理论与方法的深度;在重视阐述理论观点与原理的同时,注重它们的实际应用和指导,以期扩大本书的适用范围。

### (三) 较好地处理了继承、发展与创新的关系

近十几年来,国外教育测量与评价的新思想、新理念、新模式、新方法层出不穷,如表现性评价、质性评价、真实性评价、另类评价、实作评价、档案袋评价、动态评价、游戏化评价、多元智力评价以及成功智力评价等。这些新思想、新方法、新模式,有的比较成熟,具有可行性和推广使用价值;有的则很繁琐,缺乏可行性,不具普遍性;还有一些,花样虽新,但缺少内涵,被人喻为“杂耍式”评价方法,没有多少应用价值。为此,我们以科学、求实和创新的态度,对国外教育测量与评价的发展趋势,给予充分的关注和冷静的思考。在编写教材时,主要把握考试评价改革呈现多元化与人性化的趋势,重点突出教学与考试评价相融合、动态评价与静态评价相结合、质性评价与量化评价相结合、过程评价与结果评价相结合、纸笔测验与表现性测验评价相结合、主观题测验评价与客观题测验评价相结合等理念,较好地处理了继承、发展与创新的关系,使本书内容在理论与应用上更有生命活力。

### (四) 体例创新,增强了教材的可读性

本教材不仅在篇章结构的设计上有其独到之处,而且在编写体例上进行了一些创新。我们试图使新教材成为教与学的有效工具,力求使新教材成为启迪思维、开拓视野、深入探究、培养创新精神的好材料。所以,我们进行了教材编写体例的独特设计,讲究了撰写内容的行文风格。比如,每一章开头都列出学习目的,它告诉读者学习本章后应掌握哪些要点;每一章开头都列出编写者及任课教师建议的阅读文献,便于读者进一步学习和研究。此外,在每一章正文内容中都精心设计一个课堂讨论题、一个研究性学习专题、一个重要内容补充或案例。最后,每一章末尾都要求列出关键术语、内容提要与小结以及练习与思考题,想方设法帮助读者掌握重要的概念、观点和内容,加深理解和掌握各章重要内容。本教材由黄光扬任主编。在编写过程中,由黄光扬提出全书各章节框架结构及其编写体例要求,并邀请刘尧和龙文祥担任副主编,然后分工编写,最后由黄光扬负责全书的修改和统稿统校。本书各章节编写分工如下(按章顺序):第一章、第六章、第七章、第八章、第九章由福建师范大学黄光扬执笔(其中第九章第五节由林云鹏执笔);第二章、第三章由江西师范大学董圣鸿执笔(其中第三章中的第二节由董圣鸿和黄光扬共同执笔);第四章、第五章由安徽师范大学龙文祥执笔;第十章、第十三章由华南师范大学刘晓瑜执笔;第十一章由浙江师范大学刘尧和傅宝英执笔;第十二章由浙江师范大学刘尧和田里执笔(其中第一节由福建师范大学黄光扬执笔)。在教材编写过程



中,我们参阅了国内外同行大量的文献资料,也引用了他们的一些研究成果与数据资料。对此,本书尽可能给予详细的注明。由于受到条件限制事先未能一一函征询,在这里表示歉意和深深的谢意。由于编者水平所限,再加上时间仓促,书中不妥或错漏之处在所难免,敬请各位同行与读者批评指正。

本书编写者  
2002年3月

第二版前言 .....	( 1 )
第一版前言 .....	( 1 )
<b>第一章 教育测量与评价的学科发展 .....</b>	<b>( 1 )</b>
第一节 教育测量与评价的基本问题 .....	( 1 )
一、教育测量与评价的含义 .....	( 2 )
二、教育评价的基本问题 .....	( 5 )
三、教育评价相关概念辨析 .....	( 7 )
第二节 教育测量与评价的发展历史 .....	( 9 )
一、中国是考试制度的发源地 .....	( 10 )
二、中国科举制度的世界地位 .....	( 11 )
三、教育测量学科的诞生 .....	( 14 )
四、教育测量运动的蓬勃开展 .....	( 15 )
五、美国的“八年研究”是教育评价的 催生剂 .....	( 16 )
六、教育测量与评价理论的发展 .....	( 18 )
第三节 教育测量与评价的学科地位和 作用 .....	( 20 )
一、教育测量与评价是现代教育科学 研究的三大领域之一 .....	( 20 )
二、教育测量与评价在教育改革中 具有重要的作用 .....	( 25 )
三、教育测量与评价是教师的专业 素养和能力 .....	( 27 )
关键术语 .....	( 30 )
内容提要与小结 .....	( 30 )
练习与思考题 .....	( 32 )
<b>第二章 教育测量与评价的类型和功能 .....</b>	<b>( 33 )</b>
第一节 教育测量与评价的主要类型 .....	( 34 )

一、按测量与评价在教学中运用的时机	
分类	( 34 )
二、按解释测量结果或评价结果时的	
参照点分类	( 35 )
三、按测量与评价被试行为表现的性质	
分类	( 36 )
四、按测量与评价的内容分类	( 36 )
五、教育测量与评价的其他分类	( 37 )
第二节 教育测量与评价的主要功能	( 40 )
一、实现教育判断的功能	( 40 )
二、改进教师教学的功能	( 41 )
三、促进学生学习的功能	( 42 )
四、行使教育管理的功能	( 42 )
关键术语	( 43 )
内容提要与小结	( 44 )
练习与思考题	( 44 )

### 第三章 教育测量与评价的质量特性 ( 45 )

#### 第一节 教育测量与评价的信度 ( 46 )

一、信度的意义 ( 46 )

二、信度的统计定义 ( 46 )

三、信度的估计方法 ( 47 )

四、标准参照测验的信度分析 ( 53 )

五、测量标准误与测验信度的关系 ( 54 )

六、评分者信度 ( 55 )

#### 第二节 教育测量与评价的效度 ( 56 )

一、效度的意义 ( 56 )

二、效度的统计定义 ( 58 )

三、内容效度及其分析方法 ( 59 )

四、结构效度及其验证方法 ( 60 )

五、效标关联效度及其估计方法 ( 62 )

六、结果效度的概念与分析思想 ( 64 )

#### 第三节 教育测量与评价中题目(项目)的

    难度 ( 64 )

    一、难度系数的计算方法 ( 65 )

    二、难度指标的等距变换 ( 66 )

    三、测验题目的恰当难度和恰当难度分布 ( 68 )

第四节 教育测量与评价中题目(项目)的 区分度 .....	( 68 )
一、区分度的意义 .....	( 68 )
二、区分度的计算 .....	( 69 )
第五节 教育测量与评价方案的可用性 .....	( 73 )
一、科学性、公平性和可行性 .....	( 73 )
二、针对性、区分性和简洁性 .....	( 76 )
关键术语 .....	( 77 )
内容提要与小结 .....	( 77 )
练习与思考题 .....	( 78 )
<b>第四章 编制教育测验的一般原理与方法 .....</b>	<b>( 82 )</b>
第一节 测验题目类型与测量功能 .....	( 82 )
一、客观性试题的特点及其编写技巧 .....	( 83 )
二、主观性试题的类型及其编写要领 .....	( 91 )
第二节 测验目标与教育目标研究 .....	( 96 )
一、教育目标与测验目标 .....	( 96 )
二、教育目标分类研究 .....	( 97 )
三、确定测验目标应注意的问题 .....	( 102 )
第三节 测验蓝图设计与测验编制要领 .....	( 103 )
一、设计测验的基本考虑 .....	( 103 )
二、设计测验蓝图 .....	( 105 )
三、测验的编制与组织 .....	( 106 )
关键术语 .....	( 110 )
内容提要与小结 .....	( 110 )
练习与思考题 .....	( 111 )
<b>第五章 制订教育评价表的一般方法和步骤 ..</b>	<b>( 112 )</b>
第一节 教育评价表的基本类型与制订原则 ..	( 113 )
一、教育评价表的基本类型 .....	( 113 )
二、制订教育评价表的基本原则 .....	( 114 )
第二节 教育评价表的构成及编制方法 .....	( 115 )
一、教育评价表的构成要素 .....	( 115 )
二、制订教育评价表的方法 .....	( 117 )
关键术语 .....	( 123 )
内容提要与小结 .....	( 123 )
练习与思考题 .....	( 123 )

<b>第六章 教育测验的常模及其建立方法</b> .....	(124)
<b>第一节 教育测验常模的意义与类型</b> .....	(124)
一、测验常模的意义 .....	(125)
二、测验常模的主要类型 .....	(127)
<b>第二节 百分等级常模及建立方法</b> .....	(131)
一、百分等级常模的意义与应用 .....	(131)
二、百分等级常模的建立方法 .....	(132)
<b>第三节 标准分数常模及建立方法</b> .....	(136)
一、标准分数的基本定义 .....	(136)
二、标准分数常模的建立方法 .....	(137)
三、正确理解与使用标准分数 .....	(138)
<b>第四节 标准分数在我国高考中的试点应用</b> .....	(141)
一、建立标准分数制度是高考标准化的重要环节 .....	(141)
二、在高考中使用原始分数的局限性 .....	(142)
三、高考标准分数制度的内容 .....	(143)
四、高考标准分数转换方法 .....	(144)
<b>关键词语</b> .....	(146)
<b>内容提要与小结</b> .....	(146)
<b>练习与思考题</b> .....	(147)
<b>第七章 教育测验的标准及其建立方法</b> .....	(148)
<b>第一节 教育测验标准的含义</b> .....	(148)
一、对标准参照测验及其标准的再认识 .....	(149)
二、认识几个与标准参照测验相关的概念 .....	(150)
<b>第二节 标准参照测验分数的解释</b> .....	(152)
一、正确百分数 .....	(152)
二、掌握分数 .....	(153)
三、内容等级分数 .....	(154)
<b>第三节 标准参照测验的定性标准——测验内容领域规范</b> .....	(155)
一、汉布莱顿模式 .....	(155)
二、美国 IOE 模式 .....	(156)
三、英国 GCSE 模式 .....	(157)
四、我国高中毕业会考标准的建立模式 .....	(157)

五、我国高等教育自学考试标准的建立 模式 .....	(158)
第四节 标准参照测验的定量标准——合格 分数线 .....	(160)
一、基于题目分类下的经验判断方法 .....	(160)
二、基于题目双基度指标下的经验评 判法 .....	(161)
三、基于题目“目标—难度”层次评估法 .....	(163)
四、基于已知“掌握—未掌握”效标组的 统计分析法 .....	(164)
关键术语 .....	(165)
内容提要与小结 .....	(165)
练习与思考题 .....	(166)
第八章 学生课业发展的测量与评价 .....	(167)
第一节 课业考评改革的基本认识与目标 ..	(168)
一、课业考评的主要作用 .....	(168)
二、我国学校传统课业考评的文化特点 及目前存在的问题 .....	(169)
三、国外课业考评改革的主要经验 .....	(171)
四、课业考评改革的主要目标 .....	(174)
第二节 学生课业发展的主要内容及参照 点体系 .....	(178)
一、学生课业发展的主要内容 .....	(178)
二、评价学生课业发展的参照点体系 .....	(182)
第三节 评价学生课业发展进步的主要方法 ..	(185)
一、客观题评价法与主观题评价法 .....	(185)
二、表现性测验评价法 .....	(187)
三、评定量表评价法 .....	(193)
四、档案袋评价法 .....	(196)
五、动态评价法 .....	(198)
第四节 成绩评定等级制的原理及应用 .....	(202)
一、绝对评价等级制及应用 .....	(202)
二、相对评价等级制及应用 .....	(204)
关键术语 .....	(206)
内容提要与小结 .....	(206)
练习与思考题 .....	(208)

<b>第九章 学生智能发展的测量与评价</b> .....	(209)
<b>第一节 学生认知过程基本能力发展的测量与评价</b> .....	(210)
一、观察能力发展的测量与评价 .....	(210)
二、记忆能力发展的测量与评价 .....	(212)
三、思维能力发展的测量与评价 .....	(215)
四、想象能力发展的测量与评价 .....	(216)
<b>第二节 学生智力发展的测量与评价</b> .....	(219)
一、关于智力的经典定义 .....	(219)
二、智力结构的理论及发展 .....	(221)
三、经典智力发展的评价方法 .....	(226)
四、智力测验历史及我们的态度 .....	(233)
<b>第三节 学生创新能力的测量与评价</b> .....	(235)
一、创新的内涵及相关概念 .....	(235)
二、创新能力的评估方法 .....	(237)
三、创造能力的测量与评价 .....	(238)
<b>第四节 学生能力倾向发展的测量与评价</b> .....	(240)
一、学习能力倾向测验 .....	(240)
二、职业能力倾向测验 .....	(241)
<b>第五节 学生多元智力发展的测验与评估</b> .....	(242)
一、多元智力理论的基本观点 .....	(242)
二、发展多元智力的学校教学改革 .....	(244)
三、多元智力评估的理念与方法 .....	(246)
<b>关键术语</b> .....	(250)
<b>内容提要与小结</b> .....	(251)
<b>练习与思考题</b> .....	(252)
<b>第十章 学生人格心理发展的测量与评价</b> .....	(253)
<b>第一节 人格心理测量与评价的意义</b> .....	(253)
一、人格的不同定义 .....	(254)
二、人格的基本内容 .....	(254)
三、健康人格的特征 .....	(256)
<b>第二节 人格心理测量与评价的方法</b> .....	(256)
一、人格心理测量与评价的方法简介 .....	(257)
二、国内外知名人格测验及应用简介 .....	(259)
<b>关键术语</b> .....	(261)
<b>内容提要与小结</b> .....	(262)
<b>练习与思考题</b> .....	(262)

<b>第十一章 学生思想品德发展的测量与评价</b> …	(263)
<b>第一节 学生思想品德发展概述</b> ……………	(263)
一、学生思想品德的概念与结构 ……………	(264)
二、学生思想品德的形成与发展 ……………	(266)
<b>第二节 学生思想品德测量与评价的意义</b> …	(271)
一、学生思想品德测量与评价的可能性 ……	(271)
二、学生思想品德测量与评价的意义 ……	(272)
三、学生思想品德测量与评价的原则 ……	(274)
<b>第三节 学生思想品德测量与评价的内容     及方法</b> ……………	(276)
一、学生思想品德测量与评价的一般 内容 ……………	(276)
二、学生思想品德测量与评价的方法 ……	(279)
<b>关键术语</b> ……………	(285)
<b>内容提要与小结</b> ……………	(285)
<b>练习与思考题</b> ……………	(285)
<b>第十二章 学生体育发展的测量与评价</b> ……	(286)
<b>第一节 学生体育发展的目标及评价概述</b> …	(286)
一、学生体育发展的一般目标 ……………	(287)
二、基础教育《体育与健康》课程目标及 评价 ……………	(289)
<b>第二节 学生体能发展的意义与内容</b> ……	(291)
一、学生体能发展概述 ……………	(292)
二、体育锻炼对体能发展的影响 ……	(294)
<b>第三节 学生体能发展的测量与评价</b> ……	(296)
一、学生体能发展的测量与评价原则 ……	(296)
二、学生体能发展的测量与评价方法 ……	(297)
三、中国学生体质健康评价方法和 标准 ……………	(304)
<b>关键术语</b> ……………	(306)
<b>内容提要与小结</b> ……………	(306)
<b>练习与思考题</b> ……………	(306)
<b>第十三章 现代教育测量与评价的发展趋势</b> …	(307)
<b>第一节 经典测验理论的特色</b> ……………	(307)
一、经典测验理论的主要优点 ……………	(308)
二、经典测验理论的主要局限性 ……	(308)



第二节 现代测验理论的优势 .....	(309)
一、题目反应理论的主要优点 .....	(309)
二、题目反应理论的主要局限性 .....	(310)
三、经典测验理论与现代测验理论的 比较 .....	(310)
第三节 现代教育测量与评价的发展趋势 .....	(312)
一、教育测量与评价的政治性得到加强 .....	(312)
二、教育测量与评价的教育功能得到 强化 .....	(312)
三、教育测量与评价的理论研究得到 突破性进展 .....	(313)
四、计算机技术的发展提高了测验的 效率 .....	(314)
五、测量理论与认知心理学结合更加 紧密 .....	(314)
关键术语 .....	(315)
内容提要与小结 .....	(315)
练习与思考题 .....	(315)
附录 .....	(316)
附表1 $W$ ·显著性检验时 $S$ 的临界值表 .....	(316)
附表2 正态分布表 .....	(317)
附表3 积差相关系数( $r$ )显著性临界值表 .....	(321)
附表4 $t$ 值表 .....	(322)
附表5 高考标准分与百分等级对照简表 .....	(324)
主要参考文献 .....	(325)