

 新课程与教育评价改革译丛

译丛主编 董奇

# 教学中的测验与评价

Measurement and Assessment in Teaching  
(8th Edition)

【美】 Robert L. Linn & Norman E. Gronlund 著

国家基础教育课程改革  
“促进教师发展与学生成长的评价研究”项目组 译



中国轻工业出版社

# 教学中的测验与评价

Measurement and Assessment in Teaching  
J. H. Kortholt

Measurement and Assessment in Teaching  
J. H. Kortholt

新课程与教育评价改革译丛

译丛主编 董奇

Measurement and Assessment in Teaching (8th Edition)

# 教学中的测验与评价

【美】 Robert L. Linn & Norman E. Gronlund 著

国家基础教育课程改革  
“促进教师发展与学生成长的评价研究”项目组 译



中国轻工业出版社

## 图书在版编目(CIP)数据

教学中的测验与评价 / (美)林 (Linn, R.L.), (美)格朗伦德 (Gronlund, N.E.) 著; 国家基础教育课程改革“促进教师发展与学生成长的评价研究”项目组译.  
—北京: 中国轻工业出版社, 2003.1  
(新课程与教育评价改革译丛)  
ISBN 7-5019-3856-3

I. 教... II. ①林... ②格... ③国... III. ①教育  
测验—中小学 ②教学评议—中小学 IV. G632.0

中国版本图书馆CIP数据核字(2002)第092461号

## 版权声明

Authorized translation from the English language edition, entitled MEASUREMENT AND ASSESSMENT IN TEACHING, 8th Edition by Linn, Robert L.; Gronlund, Norman E., published by Pearson Education, Inc, publishing as Prentice-Hall, Copyright © 2000 by Prentice-Hall, Inc.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc. CHINESE SIMPLIFIED language edition published by CHINA LIGHT INDUSTRY PRESS, Copyright © 2002

本书译自美国培生教育有限公司出版的 *Measurement and Assessment in Teaching*, 8th Edition by Linn, Robert L.; Gronlund, Norman E.

本书中文简体版由美国培生教育有限公司授权中国轻工业出版社出版。

本书的所有版权归美国培生教育有限公司所有。未经美国培生教育有限公司授权许可, 不得以任何形式或方法复制或传播本书的任何内容。

策 划: 石 铁 卫 云

责任编辑: 朱 玲 卫 云 责任终审: 杜文勇

版式设计: 刘智颖 责任监印: 吴维斌

\*

出 版 人: 赵济清 (北京东长安街6号, 邮编: 100740)

网 址: <http://www.chlip.com.cn>

电子信箱: [bjwqtw@sina.com](mailto:bjwqtw@sina.com)

电 话: (010) 65262933

印 刷: 北京天竺颖华印刷厂

经 销: 各地新华书店

版 次: 2003年1月第1版 2003年1月第1次印刷

开 本: 787 × 1092 1/16 印张: 19.25

字 数: 250千字

书 号: ISBN 7-5019-3856-3/G · 367

定 价: 35.00元

著作权合同登记 图字: 01-2002-1225

· 如发现图书残缺请直接与我社发行部联系调换 ·

中国轻工业出版社读者服务部电话: 010-65121390

**国家基础教育课程改革重点项目**  
**“促进教师发展与学生成长的评价研究”**

# 新课程与教育评价改革译丛

## 编委会

主任：董奇

委员：（以下人名按姓氏笔画顺序排列）

王耘	王昌海	王文静	王斌华
刘永胜	刘志军	许明	许燕
孙凯基	芦咏莉	杨文钢	李灵
李烈	李家永	邹泓	范春林
庞丽娟	孟庆茂	赵希斌	赵学勤
赵德成	莫景祺	徐芬	徐玉珍
郭德俊	康长运	梁威	曾琦



# 总 序

全面推进素质教育是我国教育的根本任务。在素质教育实施过程中,构建符合素质教育要求的新的基础教育课程体系,已成为素质教育实施的核心。其中,如何适应新课程改革发展的需要,建立符合素质教育要求,促进学生成长、教师发展、学校教学质量提高的教育评价体系,已成为新课程改革中的一项重要任务。《基础教育课程改革纲要(试行)》明确提出“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”,给新课程的教育评价改革指明了方向。广大教育理论与实践工作者对构建新型教育评价体系的热情参与和期待,教育评价理论与实践领域自身发展与完善的需求,使教育评价改革成为新课程实施中备受关注的焦点。

新课程需要新的教育评价,其根本目的是促进学生的发展。新的教育评价观应该突显教育评价的发展性功能、强调评价主体的互动和参与、评价内容与方式的多元化以及评价过程的动态发展等等。但是,从新课程实施的现状来看,原有的教育评价体系与新课程的要求相差甚远,不能给新课程改革以强有力的支持,实践操作中存在着诸多问题,如过分地强调甄别与选拔,过多地关注对书本知识掌握的评价,过于重视终结性评价,过于注重纸笔测验与自上而下的评价等等,已影响了新课程改革工作的开展和学生的发展。

为了研究与解决当前教育评价理论与实践中的问题,更好地推动基础教育课程改革,教育部成立了国家基础教育课程改革“促进教师发展与学生成长的评价研究”项目组(以下简称“评价项目组”),由本人组织全国有关专家对教育评价改革中涉及到的诸多问题进行较为系统、深入的研究。两年多来,评价项目组的各位专家、成员努力工作,在积极参与新课程教育评价实践改革的同时,还立足于实践需求,进行了一系列有针对性的理论研究和国际比较研究。根据我国基础教育课程改革的实际需要,有选择地将国际教育评价研究与实践领域中新的研究成果和有效的评价方法等介绍给广大教师和教育理论工作者,是我们项目组的一项重要任务。

从国际视野看,基础教育课程改革近年来已成为国际教育改革的主旋律,而教育评价则是世界各国共同关注并致力研究的难题。目前教育科学、心理科学、认知科学、社会学、人类学、系统科学、脑科学等领域的最新研究成果为国际教育评价的发展提供了新的理论与科学基础。第四代教育评价理论的兴起,质性教育评价的



## 教学中的测验与评价

发展,表现性评价、真实性评价等多种教育评价方式的应用与推广,新兴教育评价理论对学生发展的促进、对课堂教学改革的推动,使西方一些国家的教育评价改革进入了一个新的发展阶段。许多国家在教育评价方面立足学校和课堂,不断进行探索和实践,已总结出了一些行之有效的教育评价改革方法,并形成了较为成熟的教育评价改革范式。为了满足新课程实施中教育评价改革的需要,开阔教育理论与实践工作者的视野,我们决定出版《新课程与教育评价改革译丛》,较系统地介绍国际上在该方面的新成果、新理论、新方法、新实践。该译丛将采取动态、开放、发展和面向实践需求的原则,不断选取当前国际教育评价领域中具有影响的研究成果进行翻译,并陆续出版。

译丛的出版旨在为我国的 教育评价研究搭建国际交流与互动的平台,为我国教育评价的理论与实践工作者、广大教师提供启发和借鉴。由于各国政治、经济、文化和教育等方面的巨大差异,对教育评价的看法也见仁见智,因此,广大教育工作者需要注意结合我国基础教育课程改革的特点和实际工作需求,批判地吸收译作中的观点和方法,努力将适合我国国情、各地区与各校教育实际的经验吸纳到我们的教育评价改革实践中来,以更好地推动新课程的实施。

在译丛的首批译作《有效的学生评价》、《促进教学的课堂评价》、《教学中的测验与评价》即将问世之际,我们首先要感谢教育部基础教育司领导对该方面工作的高度关注和大力支持,感谢译丛原作者的辛勤与智慧,感谢以不同形式参与本译丛工作的所有工作人员,感谢中国轻工业出版社的大力支持。我们希望,该译丛的陆续出版将会给我国的新课程改革注入新的活力,给我国广大教育工作者提供启发和借鉴,并不断丰富我国教育评价领域的研究。

教育评价改革是素质教育实施中的一项系统工程,更是一项关注学生未来发展的“希望工程”。我们希望社会各界都来参与、关心和支持我国的教育评价改革,愿广大教育工作者在教育评价方面的专业素质能得到不断提高,愿新课程背景下的教育评价改革能够促进所有的学生积极、主动地发展。

董奇

2002年10月

于北京师范大学





## 译者序

每一名中小学教师都会在读书期间经历无数次的测验与考试,自己做了教师之后,也会在教学实践中把测验作为一种评价学生学业成就的重要方式,经常用它来检查学生的学习,督促学生的进步。但是,熟悉测验并不意味着每一位教师都能科学地使用测验。测验是一门科学,要用它公平、有效地评价学生的学业成就,考查预期教学目标的实现程度,有些重要的原则教师必须要遵守,有些基本的技术教师也要掌握。尤其是在新课程背景下,如何适应基础教育在培养目标、课程内容和教学方式等方面的变革,在日常教学过程中科学地开发和使用测验与评价,充分发挥评价促进学生成长和教师发展的重要作用,已成为摆在教师面前一个富有挑战性的课题。

《教学中的测验与评价》为我们开展相关研究与工作提供了很多有益的经验。这本书向中小学教师集中阐述了测验与评价工具的开发问题,比如,如何编制填空、判断、匹配和选择等各种客观性题目,如何使用解释性练习、论述题、表现性任务和成长记录袋等方式评价学生的复杂学业成就,如何通过观察、同伴评价和自我报告法评价学生的情感态度与价值观等。

尽管中美两国的教育存在很大的不同,但本书所介绍的一些理念与技术,确实能给我们以如下重要启发:

首先,本书明确指出,评价不等于测验,要考查学生的复杂学业成就、学科共通能力,以及情感态度与价值观,光靠测验还不行,还必须开发和使用表现性评价、成长记录袋、观察、自我报告等丰富多样的评价方式。我国的新课程要“改变课程过于注重知识传授的倾向,强调形成积极主动的学习态度,使获得基础知识与基本技能的过程同时成为学会学习和形成正确价值观的过程”。因此,教师在教学过程中对学生的评价不仅要用测验考查学生的知识与技能,还要高度关注过程与方法、情感态度与价值观等方面课程目标在学生身上的落实情况,要开发多样化的评价任务,综合使用多种评价技术,以全面评价预期教学目标的实现程度。

其次,本书十分注重测验、评价与教学的结合,把测验和评价看做是有效教学的有机组成部分。在传统的评价中,测验通常是在教学结束后对学习结果的检查,测验与评价在很大程度上是相互脱离的。但实际上,教师的教、学生的学与评价之间是紧密联系、相互促进的。为了实现教学与评价的结合,教师必须明确课堂教学



## 教学中的测验与评价

预期的学习成果,在提供恰当的教学的同时,使用简明有效的评价方法,考查学生学习的过程与结果,既关注终结性评价,又重视形成性评价,使评价成为改进教师课堂教学、促进教师专业成长的重要手段。

此外,本书还重视测验题目与评价任务的开发,为中小学提供十分具体的实践指导。评价对教学有导向作用,所以命题或开发评价任务的工作非常重要,它会直接影响评价能否有效地考查教师所关注的学习结果。但必须指出的是,我国的中小学教师在很大程度上缺乏相关经验,尤其是在开发表现性评价任务方面。因此,师资培训部门和其他有关机构要指导教师掌握纸笔测验中常用题型的编制技术与要求,并积极探索表现性评价任务的开发与使用,以帮助教师在教学过程中灵活地应用评价,为教学提供及时、准确的反馈,并进而不断改进课堂教学。如果经过长期的积累,能建立起一个全面而丰富的测验题目和评价任务库,将会在很大程度上方便教师的教学与评价。

本书由国家基础教育课程改革“促进教师发展与学生成长的评价研究”项目组组织翻译。本书的英文版中总共有19章内容,但考虑到我国中小学教师的知识结构基础和实际工作需求,我们没有翻译英文版中的第四至五章(关于测验与评价的信效度)和第十六至十九章(关于已出版测验的选择与使用),由其余的13章组成了新的中文版《教学中的测验与评价》。参加翻译的人员有赵希斌、王昌海、范春林、赵德成、王晓华、王明明、杨洁、曾细花和赵丽波等。杨静和程辰曦也参加了部分章节的初译。赵德成对整本书进行了校译和统稿。总项目组负责人董奇教授通读了全文,并最后定稿。“全力以赴,做到最好”是我们的工作信念,但毕竟我们的时间和精力水平有限,疏漏之处,在所难免,恳请各位读者提出批评意见。衷心地希望这本书能成为中小学教师在新课程实践中开展测验与评价工作的好帮手。

译者

2002年10月



## 前 言

本书——教学中的测验与评价（第八版）——旨在向任课教师或者即将成为教师的读者介绍一些与改善教学有关的测验与评价知识。全书的主题有两个，一是评价在教学过程中扮演着重要的角色，二是它的有效性很大程度上取决于我们编制和选择测评工具的能力。

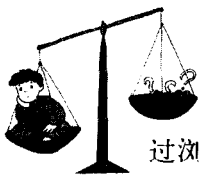
为了反映出近期在教育测验和评价方面发生的重大变化，我们对第七版内容进行了充分的修改，从而形成了新的第八版。比较大的变化包括：（1）增加了一章新内容（第十章），主要是出于教学、评价以及与家长等人交流的考虑，介绍如何创建和使用学生作品成长记录袋；（2）在几个章节中介绍了与内容、表现标准有关材料；（3）在第一章中对当前评价的现状进行了总结与回顾，并讨论了评价中存在的问题和今后发展的趋势；（4）增加了一系列关于表现性评价方法的内容；（5）更新一些测验资料，以反映评价的新版本和新方法。

本书可作为教育测验与评价领域的初级教程，无论是本科生还是研究生，不需要测量或统计方面的知识背景，都可以使用本书。本书的写作十分通俗和简洁，有大量的案例说明有关重要概念和方法。统计的概念只在需要理解和讨论某些问题时才会提出，而且其重点在于解释而不是计算。

和前几版一样，本书试图让读者通过学习达成下面的一系列目标：

1. 理解测验与评价在教学过程中扮演的角色。
2. 能科学界定教学的目的和目标，以便编制合适的测验与评价。
3. 编制课堂测验和评价工具，以测量各种各样的学习结果，无论是简单的还是复杂的。
4. 能通过课堂观察、同伴评价和自我报告获得评价信息。
5. 能以合适的方式实施测验和评价，并有效地使用这些结果。
6. 能正确解释测评结果，充分认识到结果的意义以及测验中存在的错误。
7. 能认识到在学校中使用的各种测验和评价工具的优势和不足。
8. 理解测验和评价如何能够促成有效的记分和报告系统并进而改善教学决策。
9. 能正确解释和使用已出版的测验，并能根据既定目的选择最合适的测验。
10. 理解使用外部统一测评的趋势和有关问题。

要达成上述学习目标，你还要在阅读本书的同时完成其它一些练习和任务。通



## 教学中的测验与评价

过浏览公司网站<http://www.prenhall.com/linn>, 你可以复习和应用各章节中的内容, 并对某些与测验和评价有关的问题进行深入探讨。此外, 你还可以一边阅读本书, 一边亲身参加一些实践活动, 比如编制测验和评价工具、评分、分析结果、选择测验和其它评价工具, 或解释测验和评价结果。不管怎样, 将书中所学应用于实践, 不仅有助于巩固你的学习, 而且能提高你使用测验技术的实际能力。

我们要特别感谢审阅了第七版书稿的所有人, 他们是: 杜肯斯大学 (Duquesne University) 的 S.M. 布鲁克哈特 (Susan M. Brookhart)、加利福尼亚州立大学 (California State University Hayward) 的 D.J 布朗 (Donald J. Brown)、西南路易斯安那大学 (University of Southwestern Louisiana) 的 J. 弗莱兹 (Jim Flaitz)、鲍尔州立大学 (Ball State University) 的 B.E. 格里德利 (Betty E. Gridley)、乔治亚大学 (University of Georgia) 的 C.J. 休伯特 (Carl J. Huberty)、休斯顿大学 (University of Houston) 的 R.M. 琼斯 (Robert M. Jones)、马里兰大学 (University of Maryland) 的 R.W. 里斯兹 (Robert W. Lissitz)、SUNY 的 R.F. 麦克莫里斯 (Robert F. McMorris)、爱达荷州立大学 (Idaho State University) 的 S. 萨尔兹曼 (Stephanie Salzman)、宾夕法尼亚州立大学 (Pennsylvania State University) 的 H. 休恩 (Hoi Suen), 以及北卡罗来纳州立大学 (University of North Carolina) 的 K. 怀特 (Kinnard White)。

我们还要衷心感谢那些审阅本版书稿的所有人, 他们提供了许多宝贵建议, 他们是: 南缅因大学 (University of Southern Maine) 的 J. 博德里 (Jeff Beaudry)、宾州印地安那大学 (Indiana University of Pennsylvania) 的 R.L. 布里斯科 (Roger L. Briscoe)、杜肯斯大学 (Duquesne University) 的 S.M. 布鲁克哈特 (Susan M. Brookhart)、西南路易斯安那大学 (University of Southwestern Louisiana) 的 J. 弗莱兹 (Jim Flaitz)、马里兰大学 (University of Maryland) 的 R.W. 里斯兹 (Robert W. Lissitz)、印地安那大学 (Indiana University) 的 J.E. 麦克恩伊奈伊 (John E. McEneaney)、维吉尼亚大学 (University of Virginia) 的 T.R. 穆恩 (Tonya R. Moon)、威斯康星大学 (University of Wisconsin) 的 P. 尼科尔斯 (Paul Nichols)、内华达大学 (University of Nevada) 的 P. 帕金斯 (Peggy Perkins)、东密歇根大学 (Eastern Michigan University) 的 P. 波凯 (Patricia Pokay) 以及坦普尔大学 (Temple University) 的 G.E. 斯奈尔博克 (Glenn E. Snelbecker)。

感谢有关出版商慷慨允许我们在本书中使用了一些他们享有版权的材料; 我们还要感谢我们的妻子对我们的耐心和支持, 感谢梅里尔编辑室 (Merrill editorial staff) 的鼎力相助。

R. L. 林 (Robert L. Linn)

N. E. 格朗伦德 (Norman E. Gronlund)



# 目 录

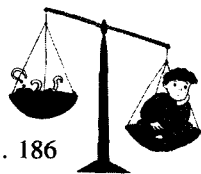
第一章	教育测验与评价的背景、问题和趋势 .....	1
	教育评价：温度计和改革杠杆 .....	2
	四十年来基于测验的教育改革 .....	2
	全国性评价重点 .....	8
	测评中的技术进步 .....	13
	公众对测评的关注 .....	15
	小结 .....	21
	练习题 .....	22
第二章	测验和评价在教学中的作用 .....	23
	教学决策需要评价信息 .....	24
	评价、测验和测量 .....	25
	评价的一般原则 .....	26
	评价与教学过程 .....	28
	评价方法的种类 .....	30
	评价种类小结 .....	36
	小结 .....	38
	练习题 .....	40
第三章	评价的基础：教学的目的和目标 .....	41
	作为学习成果的教学目标 .....	43
	从极度具体的目标到思考性课程 .....	45
	陈述教学目标的方法 .....	52
	小结 .....	59
	练习题 .....	60
第四章	课堂测验与评价的设计 .....	61
	课堂测验和评价的目的 .....	62
	设计测验和评价的操作细目表 .....	64
	选择合适的题目和评价任务类型 .....	71
	准备测验题目和评价任务时需要考虑的一些问题 .....	77



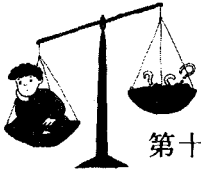
## 教学中的测验与评价

	小结 .....	85
	练习题 .....	87
第五章	简单形式客观测验题的编制 .....	89
	简答题 .....	89
	是非题或二选一题 .....	98
	匹配题 .....	105
	小结 .....	110
	练习题 .....	111
第六章	客观性试题的设计：选择题 .....	113
	选择题的特征 .....	113
	选择题的作用 .....	115
	选择题的优点与不足 .....	120
	设计选择题的建议 .....	123
	小结 .....	134
	练习题 .....	135
第七章	复杂成就的测量：解释性练习 .....	137
	解释性练习的本质 .....	138
	解释性练习的形式和使用 .....	138
	解释性练习的优点和不足 .....	145
	设计解释性练习的建议 .....	147
	小结 .....	152
	练习题 .....	153
第八章	复杂成就的测量：论述题 .....	155
	论述题的形式和使用 .....	155
	论述题的优点和局限 .....	159
	设计论述题的建议 .....	162
	评分标准 .....	166
	论述题评分的建议 .....	171
	小结 .....	175
	练习题 .....	175
第九章	复杂成就的测量：表现性评价 .....	177
	表现性评价的类型 .....	177
	表现性评价的优势和局限性 .....	183
	设计表现性任务的建议 .....	184





	表现标准 .....	186
	评分细则和等级量表 .....	186
	核查表 .....	196
	学生参与到评定中 .....	199
	小结 .....	200
	练习题 .....	201
<b>第十章</b>	<b>成长记录袋及其应用 .....</b>	<b>203</b>
	学生成长记录袋有哪些要求 .....	204
	成长记录袋的优势与不足 .....	204
	成长记录袋的使用目的 .....	206
	选择成长记录袋内容的指导 .....	210
	学生在选择成长记录袋内容和自我评价中的作用 .....	217
	评价标准 .....	219
	在教学和交流中运用成长记录袋 .....	222
	小结 .....	222
	练习题 .....	223
<b>第十一章</b>	<b>观察、同伴评价和自我报告的使用 .....</b>	<b>225</b>
	轶事记录法 .....	226
	学生评价与报告 .....	232
	同伴评价 .....	232
	自我报告技术 .....	241
	态度测量 .....	242
	兴趣调查 .....	245
	人格测量 .....	248
	小结 .....	250
	练习题 .....	251
<b>第十二章</b>	<b>试题编排、测评实施与评估 .....</b>	<b>253</b>
	课堂测验的编排 .....	253
	课堂测评的实施与评分 .....	263
	对课堂测验与评价的评估 .....	266
	建立一个有效的题目和任务库 .....	275
	小结 .....	277
	练习题 .....	278



## 教学中的测验与评价

第十三章 评分和报告 .....	279
评分和报告系统的功能 .....	280
等级和报告系统的类型 .....	282
多样化的评分和报告系统 .....	285
召开家长会 .....	288
向家长报告标准化测验的结果 .....	290
小结 .....	292
练习题 .....	293







## 第一章

# 教育测验与评价的背景、问题和趋势

问责的要求、州、国家和国际评价项目，全国性的内容和表现标准，以及全球竞争等，都促成了人们对测验和评价需求的增加。这些因素既导致了教育测量新趋势的形成，同时也反映了这种新趋势。作为教育改革工具的测验和评价，随着人们对它们依赖的增加，产生了有关使用和解释测评的公平性问题。

提起教育测验，大家想到的大多是标准化的纸笔测验，通常，这是课堂外的某些机构所要求的测验。自从外部统一的测评在美国中小学教育中被广泛采用之后，教育测评已为公众所熟悉。这些测评服务的目的各有不同、形式多样。它们包括机器记分的多项选择测验题、要求做简短或扩展回答的测验题，或者某种类型的动手做表现性评价任务（如，一个科学实验或创作一件艺术作品）。

大部分的州和学区都规定，在某些特定的年级里要有测验或评价。这些测验或评价包括从出版商处购买的成就测验，以及专为该州或该学区设计的评价。外部统一的测评可以用于不同的目的，如监控学生的进步、证明学校或系统的教育质量、向家长汇报，或对某个学生做出重要决定（如能否升级或中学毕业）。测验和评价也可以作为国家或国际调查的一部分（如全国教育进步评价或第三次国际数学和科学研究），或用于联邦教育方案评估（如1994年改善美国学校教育法案第一条）。当然，对于本书的许多读者来说，出现在脑海中的另一种标准化测验，也许就是刚刚参加过的大学入学考试（SAT）或研究生入学考试（ACT）。

大多数学生，尤其是那些在美国接受教育的学生，在进入大学前对各种标准化测验都有着相当多的经验。在讨论美国的教育状况及教育革新时，标准化测验和评价起了十分突出的作用。由于诸多原因，标准化测验也成了激烈争论的焦点，如有人抱怨标准化测验歧视少数民族和女性，有人认为标准化测验过分强调客观性知识和低级技能，从而误导了教育。考虑到标准化测验对教育决策的重要性以及相关的问题和争议，教师要理解标准化测验，并且对它的正确使用和滥用能准确区分。

尽管本书的大部分内容会集中在课堂测评工具的编制和使用上，但我们有必要首先讨论一下外部统一测评的发展趋势与问题。外部统一测评为课堂评价的某些方面提供了背景支持。在许多情况下，这种测评有助于明确课程实施的重点，使教学