

 新课程与教育评价改革译丛

译丛主编 董奇


教育性评价

*Educative Assessment
Designing Assessments to Inform and
Improve Student Performance*

【美】Grant Wiggins 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

 新课程与教育评价改革译丛

译丛主编 董奇

Educative Assessment
Designing Assessments to Inform and Improve Student Performance

教育性评价

【美】Grant Wiggins 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

教育性评价 / (美) 威金斯 (Wiggins, G.) 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, 2005.2

(新课程与教育评价改革译丛)

ISBN 7-5019-4722-8

I. 教... II. ①威... ②国... III. 教育评估 IV. G449

中国版本图书馆 CIP 数据核字 (2004) 第 139360 号

版权声明

Copyright © 1998 by John Jossey-Bass Wiley & Sons, Inc.

All Rights Reserved.

Authorized translation from the English language edition published by John Wiley & Sons, Inc.

总策划: 石 铁

策划编辑: 赵 萍

责任编辑: 朱 玲 赵 萍 责任终审: 杜文勇

版式设计: 史春雨 责任监印: 刘智颖

出版发行: 中国轻工业出版社 (北京东长安街6号, 邮编: 100740)

印 刷: 北京天竺颖华印刷厂

经 销: 各地新华书店

版 次: 2005年2月第1版 2005年2月第1次印刷

开 本: 787 × 1000 1/16 印张: 20.75

字 数: 330千字

书 号: ISBN 7-5019-4722-8/G · 538 定价: 35.00元

著作权合同登记 图字: 01-2004-3666

咨询电话: 010-65595090, 65262933

发行电话: 010-65141375, 85119845

网 址: <http://www.chlip.com.cn>

E-mail: club@chlip.com.cn

如发现图书残缺请直接与我社读者服务部 (邮购) 联系调换



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、



教育性评价

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。

译丛新译作陆续问世,我们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“万千教育”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

于北京师范大学



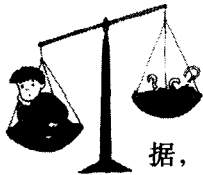
译者序

教育评价犹如水火，用之得当可以为利，用之不当可以为害。本书作者格兰特·威金斯（Grant Wiggins）就认为，美国大多数的中小學生是尚未充分发挥潜力的差生，其原因并非教师不尽心执教、学生不尽力求学，而在于流行于学校中的教学和评价体系尚未建立足以不断改进学生表现的内在机制。因此，为了匡正教育评价中的时弊，应该用有别于现行的教育评价的教育性评价（Educative Assessment）取而代之，并为此撰写了这本颇有见地的专著。

该书的突出特色是以高瞻远瞩的目光，按照教育评价的内在逻辑，用极其细腻的笔触，构建了别开生面的教育性评价体系，具体表现在以下几个方面：

其一，突出用理论指导实践。在理论层面，从教育性评价的根本目的或核心前提出发——即改进学生的表现——顺理成章地逐一推演出教育性评价的理念、原则、特征、要素，呈现出理论的严谨性和连贯性；实践层面，先是遵循有关的理念、原则，构建设计标准、评分规则，再以这些标准与规则作为“母体”，衍生出解决问题的实施策略和提供相应的实例，使人能站在理论的高度，从本质上把握教育评价实践，因此更便于人们在教育评价的实践中，按照正确的方向发挥自己的主体性和创造性，而不是一味地不明就里，对别人的东西照搬照用或亦步亦趋，也由此具有更合理的普适性和操作性。诚如作者本人所言：“本书置于理论和实践之间，其理论性表现在讨论了设计原则、评价标准，合理地说明了一种不同的评价学生的方法；其实践性表现在提供了设计模板、设计策略、解决设计问题的办法、大量的评价任务和评分规则的实例。”

其二，突出作为教育性评价质的规定性的两大基本特征，即真实性和反馈性。真实性主要指评价环境和评价任务要尽可能接近现实，以便能了解学生把所学的知识技能用于实际的真实表现；反馈性主要指在评价过程中要给学生提供获取和利用可了解自己表现的各种机会，从而能不断地根据评价目标对自我表现进行自我评价和自我调整，逐步完善自己的表现。由于真实性和反馈性是教育性评价的质的规定性，从而给人提供了判断某种教育评价是否是教育性评价的依



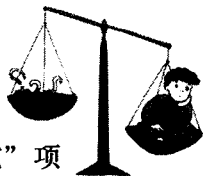
教育性评价

据，换言之，不论什么教育评价形式，只要符合这两大特征，就是教育性评价。因此，教育评价的改革不在于是否用某种新颖的教育评价形式替换某种传统的评价形式，而是要以这两大特征为指南，承前启后，推陈出新。比如，在当前美国教育评价改革中几成众矢之的的标准化考试，如果在实行这种考试时编制真实性考题，允许学生在考试中有得到和利用反馈的机会，调整只是在教学之后才举行，并只举行一次的这种常规的考试安排，消除考试的保密性，等等。那么，这种考试同样可以成为教育性评价。

其三，突出构建教育评价体系的内在逻辑。首先，本书的主要内容大致是按照什么是教育性评价、怎样设计教育性评价、如何实施教育性评价这个顺序依次进行的剥茧抽丝似的理论分析；其次，在各章的编写上，也基本是先确立所论主题的逻辑起点，后着重讨论与此有关的各因素的逻辑组合和关系。因此，该书与我国已翻译出版的其他一些以提供形形色色具体实例为特色的美国教育评价著作相比，更加体现出一种理论的深刻性和体系的严密性。

从以上特点可以看出，该书主要是一本以理论分析见长的专著，而不是一本教人如何逐步照做的“工作手册或辅导手册”（原书作者语），但长于理论并不必然就短于实践，正如该书作者所说：“教师总是过于急切地想要一种马上就可使用的便利工具，而不是努力掌握作为改革根基的策略、原则、逻辑。我积十年之久的工作经验，从中认识到，如果不从概念上搞清什么是良好的评价，那么，就决不会有讲究实效和自我矫正的学生评价。”教育性评价体系的目的一一不断改进学生的表现，与我国教育评价改革的目的——促进学生的成长与发展，不谋而合，它所欲矫正的美国学校教育评价的弊端——如过分倚重纸笔测验、编制不着边际的试题、只在教学之后才举行的一次性考试的安排、秘而不宣的考试内容和评价规则、禁止学生在接受评价中寻找和利用反馈等——在我国中小学校也不鲜见。因此，我国中小学校教师和教育评价设计者了解本书所讨论的理论——评价的目的、原则、逻辑、标准、性质等，有利于在理论的指导下从实质上认清我国教育评价偏差的根本原因，在知其然的基础上正本清源，弃旧图新，主动积极地寻找和开创标本兼治的各种措施、手段、途径，这或许比拘泥于某种具体的做法更具有实用价值。当然，这样说并不否定我们借鉴国外某些具体评价方法、形式的巨大作用和现实意义，而是想指出，若要继续健康地推进我国中小学校的教育评价改革，有所预见地避免南橘北枳的不良后果，现在更需要用理论来指导实践。就这一点而言，这本书来得正当其时。

译者序



本书是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”项目组组织翻译的。参加翻译的有华中师范大学的方形（前言与第一章）、吉艳艳（第二、三、八、九、十、十三章）、黄红霞（第四、五、六、七、十一、十二章）。译稿完成后，先由方形对全文进行了校改和统稿，后由项目组成员赵德成、边玉芳做了部分审稿工作，最后由姜英杰统校。总项目负责人董奇教授通读了全文，并最后定稿。囿于译者的水平，译文中当存不少谬误，恭请读者批评指正。

译者

2004年12月



前 言

本书阐明了在学校实施以学习为中心的评价的基本理由，论述了教育工作者在设计和使用考虑学生需要的评价方式时应该关注的工具、技术等问题。本书论证了我们需要采用与多数学校现在的学生评价完全不同的一种评价，解说了这种新的评价体系应该具有的形态和怎样用来满足学生的需要，而不是像现在的评价一样，为了追求效率和便利就牺牲学生的需求。本书还提供了我们可以用来判断评价体系的效用和价值的标准及尺度。

本书的所有建议都以一个简单明了的原则为出发点：评价应该专门用来改进和培育学生的表现，而不仅是像现在大多数学校的测验那样用来审计学生的表现。审计测验（不论是全国性的还是教师编制的，一般是间接的多项选择题测验或简答题测验）不管技术上多么无懈可击，都不能为评价的主要“用户”（即学生）提供满足其需要的服务，因为这些测验具有内在的缺陷，根本不能使学生得到对学习来说必不可少的榜样和反馈，也决不能使教师改进其工作表现。由于现在的每一个大型测验对学习结果都是一锤定音的，也由于这些测验的试题和评分规则都是秘而不宣的，所以这些传统的测验方法不能给我们的学校提供最需要的东西：一条有助于学生系统地自我矫正其表现的途径。

因此，为了不断地改进并最终达到出色的表现，学生需要教育性评价。这种评价至少要有两大要素：其一，使用真实性任务，也就是能使学生了解并从事成人在实际工作中要处理的任務；其二，为教师和学生提供便于利用的反馈和机会，使他们可以改进在完成真实性任务或类似任务时的表现。

说实话，这个主张既不激进也不新颖。从某种意义上说，本书不过是提醒我们追忆布卢姆（Benjamin Bloom）及其同事多年前对知识的真正“应用”和综合的阐述：应用要求解决新问题，而综合则要求随机应变的表现和独特的成果。可是，值得注意的是，我们当前在遵循布卢姆《教育目标分类学》中的教诲方面却仍然如同40年前那样，几乎毫无进展。有鉴于此，不妨说，我们提倡的教育性评价是一种与时俱进的布卢姆分析的最新版本（当然，特别是在不能将学术性操



教育性评价

练与表现性理解予以区分的时候，我们对其分析重点做了适当的调整与修正)。本书说明，如果教育性评价成为学生学习和教师工作的中心环节，它对评价设计、教学、课程、教师问责制、等级评定与成绩报告都具有的深远意义。

我在本书中力图达到两个看似永远都不可调和的标准：说的话虽刺耳难听，但有利于行。毫无疑问，有些读者可能期望本书提供可循序照办的评价设计指南。凡是想得到这方面具体的辅导书、练习册、设计演练的读者，可以求助于我们的组织——学习、评价和学校结构中心（Center on Learning, Assessment, and School Structure, CLASS）——该中心提供关于评价设计的各类书籍、音像制品、软件，还主办训练班、研讨会等。本书的目的则是给读者提供可以作为任何评价设计训练基础的设计标准。从长远的观点看，这些标准远比训练本身重要，因为可以避免评价工作沦为死板僵硬的教条，或成为阻塞人们思路的设计。设计如同学生的表现一样，也是一种依据标准而不断自我矫正的活动。

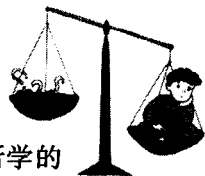
本书的读者

从事各级各类教育工作的人都可以了解和应用教育性评价，如中小学任课教师、教育局长、中小学校长、其他教育管理人员、课程与评价专家、大学教育评价与改革专业的教授和学生、从事师资培训的教师等。任何与教育学院学生直接打交道的人，任何在研究生教育类课程中从教或就读的人，都必须懂得如何评价和报告学习成果，因此了解本书所讨论的理念和实践必然有所裨益。任何负责专业或技术评价或测量的人也可从本书中获益。有人认为，当今的测验为了技术上的精确和成人的利益，而牺牲了改进作为用户的学生的表现和学生的需要，凡是持有这种观点的人也必将从本书找到许多可以用来匡正评价时弊的办法和手段。

本书的内容

在第一章我展示了对教育性评价的一种设想：教育性评价具有什么形态？当用真实性任务评价学生，学生又能得到并利用有用的反馈时，会发生什么事情？随后三章构成本书的第一编，阐述教育性评价的各个要素。

在第二章我说明并探讨了教育性评价的两大要素或两大原则之一：真实性任务。真实性任务模拟成人的任务，使学生从事有价值的工作，并产生有用的反



馈。于是，学生能够不断地改进自己的理解和表现，也就是能够应用自己所学的学科知识。

在第三章我考察了教育性评价的另一要素：反馈。说明反馈为何要贯穿整个评价过程，而不仅是评价后才给出富有意义并便于用户使用的反馈。有用的反馈如何不同于表扬和批评？为何与已知的目标挂钩的反馈是公正评价的有机组成部分？我们怎样使反馈促进学生达到出色的表现？这些问题是该章的重点。评价和反馈以前老是被认为是教学以后的事情，但只有利用真实性任务和反馈来不断地评价学生的理解力，才能帮助学生对自己表现进行自我评价和自我调整，也才能彻底地改变我们对教学本身的观念。我们需要一种教学理论，它承认在学习和评价中使学生对自己学习结果做出反应的重要性，因此学生在某种程度上可以进行有目的的试误和调整。

第四章探究了为何传统的测验不能合适地验证理解力，也不能改进理解力的问题。什么是理解力？为什么我们竭力教学之后还有那么多学生缺乏理解或仍存误解？为什么传统的测验难以恰当地评价理解力？该章指出了一种或许惹人恼火的理论，根据该理论，理解力由下面五个互相区别又互相依存的能力组成：精确细微的讲述能力、随机应变的应用能力、由表及里的观察能力、体谅他人的共感能力和恰如其分的自知能力。虽然这个理论对评价具有多种意义，但许多含义可以从前面各章所述的观点衍生而来：要看出学生的理解力，必须要使用真实性工作和回应反馈，必须要看学生是否能够合理地解释答案、进行自我评价，真正地把知识应用于各种情景。

在本书的第二编，我转向考察评价设计方面的问题，重点探讨按什么逻辑次序组合评价设计的各要素，探讨可以作为任何评价设计的评判标准的一套设计契约。第二编的内容不是告诉你如何设计一项具体的评价任务，而是使你知道如何了解你使用的评价任务的得失。为了使评价有效，设计者在教学前需要探询一系列问题，其基本问题如下：要断定学习是否成功，我们需要什么证据？要产生那样的证据，应该提出什么样的学习要求？

第五章讨论了如何利用第二章至第四章所述的要素（真实性任务、反馈和理解力）来设计真实性评价。地方的教师必须学会像评价者那样思考——一个戏剧性的角色转变——去问一问：什么是学习证据，我怎样通过教学去寻找这个证据？而不是像现在我们常做的那样，只是暗自思量：我要传授哪些课程内容，对此又要求学生做些什么？该章及随后两章详细述说了支撑一个教育性评价体系的



教育性评价

设计标准，由此可以澄清内容标准、表现标准、评价标准之间的区别，而所有这些标准都对改进我们学校起着举足轻重的作用。

第六章主要讨论评价任务和真实性问题。讨论的中心问题是：什么是一项真实性任务？为什么真实性任务不能和表现性评价混为一谈？真实性和效度之间具有什么样的关系？

第七章探讨了评分规则 and 标准，回答了经常被人问到的关于评分规则设计及使用的许多问题，列举了良好的或不良的评分规则的实例。该章和前面各章一样，旨在使学生充分地了解表现的预期影响——作文是否引人入胜？论据是否令人信服？解决方法是否奏效等等——而不是像我们现在常做的那样，只是要求学生把注意力集中在作文中的内容准确性、组织和写作常规等方面。从目的着眼来看待表现，意味着学生必须学会符合标准地完成任务——而不是学着怎样取悦老师或一味地循规蹈矩，照章行事。

本书第三编考察了设计工作的意义。

第八章建议把成长记录袋派上一个大有作为的新用场——学习的进程可以被看成是一个产生有效的、发人深省的学习成长记录袋的过程——因此可以通过成长记录袋这个透镜考虑评价的“宏观”问题。我们怎样设计学生成长记录袋来确保合理的评价？有什么种类的成长记录袋？不同的目的如何影响成长记录袋的设计？讨论的中心议题是成长记录袋的所有权问题：到底是谁的成长记录袋？是我们的还是学生的？对这些问题的不同回答又怎样影响到我们的工作？

第九章考虑了课程问题：以表现为基础的教育性评价对课程有什么意义？如何确保课程或教学计划促使学生的表现不断地靠近或达到高标准？这些问题本身就蕴涵着答案：课程不能脱离评价，也不能先于评价。我提出了对课程采用“反向”构建的思路：先要确定我们希望学生理解的东西和改进这种理解的评价任务，惟有到了此时，我们才能决定如何最好地开发出帮助学生掌握有关学识和能力的一堂课的教学活动。这种以表现为基础的课程观同样使学生（和教师）可以按事情的轻重缓急来安排教学内容和教学时间，并认识到课堂教学与活动的更为深远的意义。换言之，一个评价驱动的课程体系能使学生更明确学习的目的，一目了然地看到学习的意义。这种课程前后一致，首尾贯通，有方向引路，不断重复其中关键的概念、理论、难题和表现类型，从而使学生形成深刻的理解力，掌握核心知识。该章最后提供了课程设计模板与课程同行评议的范例。

第十章考察了因实施新型的评价和课程而产生的一个实际问题，也是许多人



关心的问题：要有与之配套的新型的等级评定和成绩报告制度。现行的成绩报告卡和等级评定制度不能本着公正和诚实的要求，让我们知悉我们本应该了解的复杂的表现：根据标准和合理的期望，学生现在是怎样做的。该章考察了新的成绩报告卡，描述和说明了等级评定的新方法，从而表明了我们可以从什么地方入手来更有效地和更有益地区分学业进步、学业成就和学习态度。

第十一章我指出教师应该欢迎问责制，而不应该抵制或半推半就地容忍问责制。如果反馈是所有学习的关键，那么也必然是教学的关键。问责制应该建立在反馈的基础之上，应该要求教师不断地寻求反馈和应答反馈（不管反馈是来自用户还是来自评价结果）。当要求我们把寻求和利用反馈作为我们日常实践的一部分时，我们就有了真正的，而不仅是政治上的问责制。教师和学生都应该认识到，寻求和接受关于他们表现的反馈完全符合自身利益，完全不同于他们以前所经历的所谓督导和评估。这些做法常常束缚了师生的手脚，或是毫无用处。该章结尾处描述了处于问责制核心地位的一种激励性和专业性的机制——同行评议程序。借助这种程序。可以对评价设计草案进行修订、纠错、完善。我们必须以身作则地实行我们提倡的东西。我们总是要求学生齐心协力，学会互相帮助，共同达到标准，这样的方法同样可以用于我们成人的设计工作。确实，如果没有旁人对我们的设计工作做出公正严格的评议，那就几乎不可能遵循前面有些部分谈到的设计逻辑。这样的评议是真正专业性的，又是特别激励人的，前提是我们公开地、诚实地、负责地实施这种评议。

最后两章讨论了可行性问题，提出了以后的措施。一个教师或一个学校系统应该从哪里起步？前进的道路上可能有什么绊脚石？预想的障碍是实在的，还是虚幻的？有哪些现存的理念和范例有助于我们对改革计划和政策做出必要的修正，以实现我们描画的远景？有哪些最可能导致改革成功的策略？又应该避免哪些易于常犯的错误？这些是该章予以回答的若干问题。

我本人讲究实际，不抱幻想。我知道落实本书所说的每一个建议或每一个想法都决非易事，我的建议也许在读者所处的环境下根本不可能实行。因此我写作本书面临的严峻挑战，就是能为读者指明前进道路的方向，提出跨越障碍而不是向障碍低头的指导原则。本书的所有理念都来自现实学校中的经验，虽然其中许多学校还真是缺乏怀疑者所“认定”的导致成功的因素：时间、金钱、才能。

本书各章的基本立论是，所有的学生，尤其是最差和最好的学生都能具有远远超出他们现有成就的表现。现在学生没有应有的表现，这倒不是因为教师和学生



教育性评价

生有什么毛病或不思进取，也不是因为有什么其他社会或学校问题干扰我们的工作，而是因为我们学校本身实施的是一种最不能促进表现的制度。

本书的新颖之处是毫不忌讳地坦言，我们学校中绝大多数都是尚未发挥潜力的差生，其根本原因在于现在学校暗地流行一种关于表现的错误理论：教学和勤奋就能导致学生娴熟出色的表现。可是娴熟出色的表现决不会来自按照某种逻辑顺序而凑集起来的不相联系的教与学。相反，要使学生有娴熟出色的表现，必须给学生清晰而有价值的表现目标、有用的反馈、因势利导的辅导以及提供让学生表现的各种机会，使学生不断地显现欲达到目的的表现，从而逐渐地、甚至飞跃地通向卓越的境界。显而易见，并非所有的学生都能达到我们所说的最高标准（一个自相矛盾的措辞），但我们仍然不能认清我们的教与学方法是多么的无效（因为是靠输入驱动的）。只要把评价仍然看作是教与学之后才做的事情，那么不管学生现在的学习有多好或多差，也不管学生现在的学习动机有多强或多弱，我们就不能极大地改进学生的表现 [本书对如何提供实现卓越表现所需的机会谈的多，而对如何给学生提供激励谈的少。如果读者想更多地了解有关激励，有关学生如何通过良好的评价增强内在和外在在学习动机的内容，那就请阅读我那本于1993年发行的书《评价学生表现》(Assessing Student Performance)]。

我对于评价的思想，也是根据来自读者、用户以及落实我的想法而产生的效果的反馈，不断地发展变化的。本书的大多数材料可以看作是全新的，虽然有一半章节的内容可以追溯到早期的论文（但可以说采用了与早期论文不同的角度进行了改写，凡是熟悉我早期论文的读者可以看出这一点）。例如，第十章的内容，五年内改写了三次。本书的用处，与其说在于它可能包含某些绝对真理，不如说在于它必然要引起的人们的疑问、讨论和研究。

所以，本书置于理论和实践之间。其理论性表现在讨论了设计原则、评价标准，合理地说明了一种不同的评价学生的方法（熟悉《评价学生表现》的读者，不妨把本书看作是前书的姊妹篇）；其实践性表现在提供了设计模板、设计策略、解决设计问题的办法、大量的评价任务和评分规则的实例。对于本书采用的材料，根据来自用户的反馈，以及做咨询工作时将这些材料用于学校、学区、州教育厅的经验，我和同事一直不断地予以增补和修订。读者也可以利用本书提供的设计标准和模板来草拟或不断修订自己的评价方案。但是，本书不是一本工作手册或辅导手册。

“告诉我，周一我可以做什么？”全国各地都可以听到教师发出这样的呼声，



尽管我们最终必须做到这一点，但这种要求立竿见影的直接关联性的呼声有其害处。教师总是过于急切地想要一种马上就可使用的便利工具，而不是努力掌握作为改革根基的策略、原则、逻辑。我积十年之久的工作经验，从中认识到，如果不从概念上搞清什么是良好的评价，那么就决不会有讲究实效和自我矫正的学生与学校表现的评价。我们要掌握评价设计，就像掌握评价本身一样，我们只有通过反馈来建立理想评价（体现在某个标准或远景上）的清晰概念，才能做到真正的掌握。没有什么包打天下的技术，也没有什么包治百病的药方。如果我们仅凭习惯和猜想工作，我们就难以改进工作。只有根据标准和反馈来判断我们的工作，才能改善设计，提高效益。这就是我想对读者说的最有用的观点。

最后，我要指出的是，如果读者想自修评价设计，CLASS 可以提供这方面的设计工具、全套工作表格、培训材料、评价任务和评分规则的样本手册。欲知关于 CLASS 的材料、培训、项目的详情，请按下列方式和我们联系：

发信函到：CLAAS, 56 South Main Street, Pennington, NJ, 08534

打电话到：(609) 730 - 1199

发电邮到：info@classnj.org

上因特网：www.classnj.org

格兰特·威金斯 (Grant Wiggins)

新泽西州彭宁顿

1997 年 12 月



作者简介

格兰特·威金斯是设在新泽西州彭宁顿的非营利性教育组织 CLASS 的主席和项目主任。CLASS 就各项改革事宜为学校、学区和州教育部提供咨询；主办全国性会议和研讨会；开发评价与课程变革方面的音像制品、软件和印刷材料。威金斯和他在 CLASS 同事曾为美国一些最具影响的评价改革计划（包括肯塔基州的表现本位评价体系和佛蒙特州的成长记录袋评价体系）担任了顾问。他们还成立了两个专门从事评价改革的全州性联合会，最近还为北卡罗来纳州标准和问责委员会设计一个表现本位和教师管理的成长记录袋评价体系。

威金斯于 1972 年从安纳波利斯的圣约翰学院获得文学士学位，于 1987 年从哈佛大学获得教育博士学位。他是《评价学生表现：探讨测验的目的和局限》一书的作者，该书也是由 Josser - Bass 出版社出版发行的。他的许多关于课程和评价改革的论文发表在《教育领导》、《大学生联谊会学刊》等学术期刊上。他有 14 年中学教学和辅导的经验作为其工作的基础。



目 录

第一编 评价的基本元素

第一章	教育性评价	3
	一种面向未来的设想	3
	课堂所闻	3
	课堂所见	5
	对教育管理和教育政策的影响	7
	核心前提:评价目的	8
	从不同的角度看待当前的测试	10
	教育性评价的界说	12
	激励学习,激励教学变革	16
	结 语	17
第二章	确保真实性表现	19
	真实性任务	19
	影响和背景:不仅仅是真诚的努力	25
	评价任务:并非是教学活动	27
	作为智力试误的表现	31
	处理后果:通过评价的现实疗法	35
第三章	提供持续反馈	39
	什么是反馈	41
	用评价导致卓越:写作即修改	50
	对自我评价、自我调整的评价和传授	53
	使用带样例的反馈	57
	教育性评分规则和标准:旨在预期效果	59
	循环任务和纵向评价	60
第四章	增进学生的理解力	63
	知识与理解力	64