



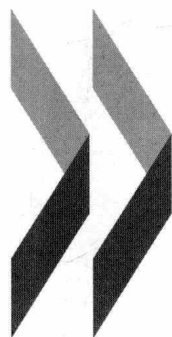
为了更好的学习

教育评价的国际新视野

经济合作与发展组织 编
窦卫霖 等译

Synergies for Better Learning
AN INTERNATIONAL PERSPECTIVE ON EVALUATION AND ASSESSMENT





为了更好的学习

教育评价的国际新视野

经济合作与发展组织 编

窦卫霖 等译

Synergies for Better Learning

AN INTERNATIONAL PERSPECTIVE ON EVALUATION AND ASSESSMENT

本书由经济合作与发展组织(OECD)秘书处负责出版,所表达的观点和采用的论据不代表 OECD 或其成员国政府的官方看法。

本书及所包含的任何地图均无意侵犯任何领土的地位及主权,不影响任何国际边界的划分,也不影响任何地域、城市或地区的名称。

有关以色列的统计数据由以色列有关当局提供并负责。OECD 使用这些数据无意损害国际法条款对戈兰高地、东耶路撒冷、约旦河西岸犹太人定居点的界定。

上海市版权局著作权合同登记号 图字:09-2019-325号

图书在版编目(CIP)数据

为了更好的学习:教育评价的国际新视野/经济合作与发展组织编;窦卫霖等译.-上海:上海教育出版社,2019.5

ISBN 978-7-5444-9092-4

I. ①为… II. ①经…②窦… III. ①教育评估-研究报告-世界 IV. ①G40-058.1

中国版本图书馆 CIP 数据核字(2019)第 085283 号

Originally published by the OECD in English under the title: *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education © 2013 OECD

© 2019 Shanghai Educational Publishing House for this Chinese edition

原书由 OECD 以英文出版,标题为: *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education © 2013 OECD

简体中文版由上海教育出版社出版。版权所有,违者必究。



责任编辑 廖承琳 谢冬华 周 晟

封面设计 郑 艺

为了更好的学习:教育评价的国际新视野

经济合作与发展组织 编

窦卫霖 等译

出版发行 上海教育出版社有限公司

官 网 www.seph.com.cn

地 址 上海永福路 123 号

邮 编 200031

印 刷 上海中华印刷有限公司

开 本 889×1194 1/16 印张 39.5 插页 1

字 数 980 千字

版 次 2019 年 5 月第 1 版

印 次 2019 年 5 月第 1 次印刷

书 号 ISBN 978-7-5444-9092-4/G·7511

定 价 298.00 元

如发现质量问题,读者可向本社调换 电话:021-64377165

翻译团队成员

(按照参与翻译的章节次序排序)

窦卫霖 张璐 杜海紫

马翔 宋晓凤 林江珊

陈钰童 隆思燮 陈圆

经济合作与发展组织

经济合作与发展组织 (Organisation for Economic Co-operation and Development, 简称 OECD) 是各国政府合作商讨应对经济、社会和环境全球化挑战的特别组织。OECD 还是促进各国相互理解, 帮助各国政府应对新形势、新问题的前线力量, 例如企业管治、信息经济以及人口老龄化的挑战。该组织提供了一个平台, 让各国政府能够相互比较政策经验, 寻求共同问题的答案, 共同确定良好实践, 以及致力于协调国内和国际政策。

OECD 成员国包括澳大利亚、奥地利、比利时、加拿大、智利、捷克共和国、丹麦、爱沙尼亚、芬兰、法国、德国、希腊、匈牙利、冰岛、爱尔兰、以色列、意大利、日本、韩国、卢森堡、墨西哥、荷兰、新西兰、挪威、波兰、葡萄牙、斯洛伐克共和国、斯洛文尼亚、西班牙、瑞典、瑞士、土耳其、英国和美国。欧盟委员会参与 OECD 的工作。

OECD 的出版物广泛传播该组织收集的统计数据, 内容涉及关于经济、社会和环境问题的研究成果, 以及其成员国达成一致的协议、指导方针和标准。

This translation is published by arrangement with the OECD. It is not an official OECD translation. The quality of the translation and its coherence with the original language text of the work are the sole responsibility of the author(s) of the translation. In the event of any discrepancy between the original work and the translation, only the text of original work shall be considered valid.

本译著中文版已取得 OECD 授权。本书非 OECD 官方译本, 翻译质量及其与原著的一致性均由译者负责。若出现翻译与原文(英文)不符的情况, 请以原文为准。

序言

教育评价的里程碑之作

袁振国

教育评价和测量是教育学中发展最早、最成熟,运用最广泛,影响最大的分支。美国的教育学硕士生、博士生所学专业差不多有一半与教育评价和测量有关。我国的高考,美国的 TOEFL、GRE 考试,英国的雅思考试,以及近几年被人们广为知晓的经济合作与发展组织(Organisation for Economic Co-operation and Development, 简称 OECD)开展的国际学生评估项目(Program for International Student Assessment, 简称 PISA),都是大型教育评价和测量项目(称“工程”或许更准确)的典型代表。这些项目参与人数之众,动员资源之巨,对社会和个人影响之大,简直无与伦比。然而,这些项目其实只是评价和测量理论及方法在教育中运用的一小部分。对学生成绩、教师绩效、学校办学水平、政府管理等的了解、判断、选择,无不与教育评价和测量有关。教育评价和测量不仅涉及方法和计算,而且涉及价值取向、理论框架、政策制定,更不要说教育行为本身。教育评价对教育实践具有巨大的导向作用,好的评价促进教育和个人的健康发展,不好的评价则会误导、阻碍教育和个人的发展。2018年9月,习近平总书记在全国教育大会上指出:“要深化教育体制改革,健全立德树人落实机制,扭转不科学的教育评价导向,坚决克服唯分数、唯升学、唯文凭、唯论文、唯帽子的顽瘴痼疾,从根本上解决教育评价指挥棒问题。”这深刻揭示了教育评价的重要性和改进教育评价的迫切性。

关于教育评价和测量的书可谓汗牛充栋,只要有教育著作的图书馆,就一定有教育评价和测量的书。关于教育评价和测量方方面面的理论及方法似乎都穷尽了,对于再获得一本与众不同、令人耳目一新的教育评价著作,似乎很难抱有希望。

令人惊喜的是,《为了更好的学习:教育评价的国际新视野》给我们带来了新的希望。这本由 OECD 教育与技能司组织动员 28 个国家(地区)的 120 多位专家学者精心打造的教育评价巨著,堪称教育评价的里程碑之作。

首先,这本著作是对全新社会发展环境下的教育的新审视。随着教育作用的不断凸显和对教育重要性认识的不断提高,各国不断增加教育经费,对有效利用公共资源的必要性的认识日益增强;全社会对教育的关注逐渐从数量发展转向了质量提高,转向了“全面素质教育”;教师的学历水平和基本素养不断提高,对教师教育专业化的要求越来越高;随着高等教育的普及,大众的受教育水平不断提高;新兴技术、高科技日新月异,给教育带来了极大的便利;作为一种服务的教育,商业潜力巨大;新媒体的兴起对教育问责的推动能量巨大,等等。社会各界对教育的评价和改进给予了

前所未有的关注。

其次，这本著作确立了以学生为中心的评价观，旨在通过教育评价和测量促进教育变革。教育评价是一把“双刃剑”，如何协调教育的标准化评价和多样性评价、教育的自评和他评（主动评价和被动评价）、单一性评价和综合性评价、群体评价与个体评价、结果评价和过程评价，等等，是教育评价的永恒话题。本书立足于为改进学生的学习统整各方面的工作，改进各项政策，同时为教师的发展、学校的领导和行政部门的管理带来具有人文精神的激励。

再次，这本著作首次清晰明确地将教育评价分为学生测评、教师考核、学校评估、学校领导考核、教育系统评估等五个不同层面，进行有针对性的评价，既突出各自的特点和要点，又努力兼顾相互之间的联系，具有广泛的适用性。在如此丰富的材料中，不同主体可以从中获得对自身最有参考价值的内容。

最后，也是现有教育评价著作无可比拟、无法实现的，是这本著作运用了全世界 28 个国家（地区）的数据、案例、经验和政策，并对其成败得失进行了分析比较，为我们了解和理解广泛区域、不同文化的教育政策和实践，为开阔眼界、引为借鉴，提供了不可多得的资源和平台。

可以说，这是一本反映教育评价国际新视野的百科全书。无论是教育政策制定者、教育理论工作者，还是教育评价专业人员、教育一线工作者，都能从中得到启发和借鉴。

2019 年 5 月

前 言

真实、有效、可靠的评估(evaluation)和测评(assessment)是改进各级教育实践和提升学生学习能力的关键,是建立高绩效教育系统的核心。它也有助于发现和奖励教育实践者的工作,认证学生的学习情况。显然,促进评估和测评对学生及其家庭以及教育工作者和教育系统都大有裨益。因此,越来越多的国家开始实施雄心勃勃的学校改革方案,包括强有力的评估和测评。这可以巩固教育系统中的评估文化,并加强评估和测评框架在推动改革议程中的作用。

2009年底,经济合作与发展组织教育政策委员会(OECD Education Policy Committee)着手全面撰写关于评估和测评政策的国际评论,即《OECD关于改善学校成果的评估和测评框架的评论》(OECD Review on Evaluation and Assessment Framework for Improving School Outcomes,简称《OECD评论》)。它的目标是为各国提供分析和政策建议,说明如何将评估和测评纳入一个统一的框架,从而真正提升整个学校系统(school system)的绩效。此外,该评论还衍生了25个参与国的26份报告、15份由外部评审小组撰写的报告[以《OECD教育评估和测评评论》(OECD Reviews of Evaluation and Assessment in Education)系列出版物发布]和若干研究论文(以上所有报告均可在经济合作与发展组织官方网站 www.oecd.org/edu/evaluationpolicy 获得)。《OECD评论》在国际层面开创性地为评估和测评政策问题提供了全面分析。

OECD(经济合作与发展组织)的工作有助于各国相互学习,还可以让我们看到一些在国家辩论中难以见到的问题,并探索对策。这两个因素都清楚地支持了本报告*及其背后的工作。各国的直接参与对这一项目的开展至关重要。28个积极参与的国家投入了大量资源,并在外部评审和讨论中公布了其评估和测评政策。这种合作方式使各国能够更多地了解自己,并通过分享政策改革的影响和政策运行的最佳环境,拓展各国的知识储备。

《OECD评论》由OECD国家评估和测评专家小组(OECD Group of National Experts,简称GNE)监制,该专家小组是作为OECD教育政策委员会的附属机构建立的。OECD国家评估和测评专家小组由来自匈牙利罗兰大学(Eötvös Loránd University)的加博尔·豪拉斯(Gábor Halász)教授担任主席。爱尔兰教育和技能部(Department of Education and Skills,简称DES)助理总督导

* “本报告”即指本书。《OECD教育评估和测评评论》是从《OECD关于改善学校成果的评估和测评框架的评论》中衍生出的系列出版物,本书是其中的一本。——译者注

员埃蒙·默塔(Eamonn Murtagh)先生和韩国教育发展研究所研究员金王宏(Kwangho Kim)先生担任副主席。我们感谢他们富有洞察力、全力投入的指导。

4 《OECD评论》很大程度上得益于各组织的参与，它们分别代表了学生、家长、教师、学校领导、学校、教师教育者、研究人员和雇主。他们的代表服务于国家委员会，所做的工作有：准备书面意见，会见评论小组，参加 OECD 国家评估和测评专家小组的会议和见面会。该评论还得益于 OECD 工商业咨询委员会和 OECD 工会咨询委员会的贡献，以及其他对评估和测评政策感兴趣的国际组织，包括欧盟委员会、欧里狄克(Eurydice^{*})、教育监督机构国际常设会议(Standing International Conference of Inspectorates, 简称 SICI)、联合国教科文组织(United Nations Educational Scientific and Cultural Organization, 简称 UNESCO)和世界银行。

附录 A 详细介绍了许多个人和组织，如各国协调员、OECD 国家评估和测评专家小组成员、各国评论小组成员和背景报告及委托研究论文的作者——总计超过 120 人。我们很幸运从他们的专长中受益，并且感谢他们对《OECD评论》的付出。15 次调查访问涉及 30 个外部评论员，以及一系列的研究和政策背景。总的来说，外部评论小组访问了大约 90 所学校，会见了 2 800 个人，这些成为各国评论调查结果的基础。感谢他们对建立这个集体知识库的宝贵贡献。

在 OECD 教育与技能司(OECD Directorate for Education and Skills),《OECD评论》由德博拉·罗斯维尔(Deborah Roseveare)领导的教育和培训政策处(Education and Training Policy Division)(2009 年 10 月至 2012 年 1 月)和迈克尔·戴维森(Michael Davidson)领导的幼儿教育和学校处(自 2012 年 2 月开始)负责执行。在 OECD 秘书处(OECD Secretariat),保罗·圣地亚哥(Paulo Santiago)(协调员)、德博拉·努舍(Deborah Nusche)和克莱尔·施博里奇(Claire Shewbridge)负责《OECD评论》并准备本报告。他们还负责个别国家的评论。海克·达妮埃拉·赫尔佐克(Heike Daniela Herzog)负责评论工作中的行政工作、会议组织和与各国的沟通。其他许多同事对本报告都有贡献(见后文的“致谢”)。OECD 的许多同事在《OECD评论》编写工作的关键阶段都提供了建议。尤其要指出的是，此项目带来了以下组织的合作与工作：国际学生评估项目(Programme for International Student Assessment, 简称 PISA),教与学国际调查(Teaching and Learning International Survey, 简称 TALIS),教育研究与创新中心的创新学习环境项目,用于收集和判断教育结构的政策和实践的系统层面描述信息的 INES 网络(NESLI),教育研究与创新中心关于纵向信息系统的工作(作为其教育和培训创新战略的一部分),教育研究与创新中心管理复杂教育系统的工作和有效学习环境中心评估教育设施质量的工作。

本报告于 2013 年 4 月 11 日在奥斯陆由挪威教育和研究部(Ministry of Education and

* Eurydice 是欧洲的一个国际组织,其主要任务是解释欧洲各教育系统是如何组织和工作的。——译者注

Research of Norway)、挪威教育和培训司(Norwegian Directorate for Education and Training)和OECD在其联合主办的国际会议上发布。

OECD将保持其在评估和测评政策中的领头作用,并在《OECD评论》和本报告的基础上继续努力。

致 谢

5

本报告的主要作者是 OECD 教育与技能司的克莱尔·施博里奇、德博拉·努舍、托马斯·拉丁格(Thomas Radinger)、保罗·圣地亚哥(协调员)。各章的主要作者为：

第一章	评估和测评焦点	保罗·圣地亚哥
第二章	评估和测评趋势	保罗·圣地亚哥
第三章	评估和测评框架：采用整体分析法	保罗·圣地亚哥
第四章	学生测评：以学生为中心	德博拉·努舍
第五章	教师考核：提高教师的专业素养	德博拉·努舍
第六章	学校评估：从合规到高质量	克莱尔·施博里奇
第七章	学校领导考核：培养学校中的教学型领导	托马斯·拉丁格
第八章	教育系统评估：为完善教育系统政策提供信息	克莱尔·施博里奇

海克·达妮埃拉·赫尔佐克负责本报告的制作和编排以及部分编辑和校对工作。OECD 秘书处成员、OECD 国家评估和测评专家小组成员、OECD 教育政策委员会成员、成员国国家协调员以及与该工作相关的研究人员和国际机构为各章的起草提供了宝贵的意见。作者团队要感谢为《OECD 教育评估和测评评论》做出具体贡献的各位专家，他们的专业知识和分析让本报告内容更加饱满可读(请见附录 A 表 A.3)。

除了要感谢核心团队成员(海克·达妮埃拉·赫尔佐克、德博拉·努舍、保罗·圣地亚哥、克莱尔·施博里奇)外，我们还要向许多在不同阶段为《OECD 评论》的撰写工作付出辛苦劳动的人致敬：从挪威教育和研究部借调过来的莫滕·罗森奎斯特(Morten Rosenkvist)(2010年2月至7月)以及两位来自卡罗·施密德项目的研究员兼顾问斯蒂芬妮·迪福(Stefanie Dufaux)(2010年9月至2011年8月)和托马斯·拉丁格(2011年9月至2012年9月)。他们的分析工作极大地丰富了该评论的知识库。托马斯·拉丁格还承担了对学校领导考核的分析，并协调总结评估和测评框架特点表格的编制工作。斯蒂芬妮·迪福和托马斯·拉丁格负责该评论中的统计工作。维奥莱纳·福贝尔(Violaine Faubert)、马莱娜·伊索雷(Marlène Isoré)、艾利森·莫里斯(Allison Morris)和安·那哈尔·普彼司克(Anne Nayral de Puybusque)在 OECD 实习期间，作为研究助理，总结了评估和测评领域的关键文献。珍妮特·卢尼(Janet Looney)作为一名外部顾问，准备了研究论文，作

为《OECD 评论》的背景分析。OECD 秘书处的弗朗西斯科·贝纳维德(Francisco Benavides)和黛安娜·托莱多(Diana Toledo)作为专家,分别为 OECD 在智利和墨西哥的评估和测评报告做出了贡献。

- 6 萨布里纳·莱奥纳尔杜齐(Sabrina Leonarduzzi)(2009 年 6 月至 12 月)和萨拉-杰恩·莫斯(Sara-Jayne Moss)(2010 年 1 月至 6 月)为《OECD 评论》的撰写提供了行政支持。卡桑德拉·戴维斯(Cassandra Davis)(教育与技能司)和安-里斯·普里让(Anne-Lise Prigent)(Public Affairs and Communications Directorate,公共事务和通信司)提供了关于本报告制作和对外传播的建议。布莱恩·基利(Brian Keeley)(公共事务和通信司)准备了本报告的“纲要”部分。特蕾泽·沃尔什(Therese Walsh)和达米安·加尼斯(Damian Garnys)(均来自公共事务和通信司)在编辑方面提供了帮助。

我们还要感谢德国学术交流中心(German Academic Exchange Service)和德国国家学术基金会(German National Academic Foundation),它们对斯蒂芬妮·迪福(2010 年 9 月至 2011 年 3 月)和托马斯·拉丁格(2011 年 9 月至 2012 年 6 月)作为卡罗·施密德项目的研究员兼顾问赴 OECD 秘书处任职提供了赞助。还要感谢挪威教育和研究部对 OECD 秘书处借调莫滕·罗森奎斯特的支持,以及主持召开国际会议(2013 年 4 月 11—12 日)发布本报告。

目 录

纲要	1
引言	1
共同主题	1
挑战与方向	2
概述	5
评估和测评越来越受关注	5
评估和测评的主要趋势	5
主要政策挑战	7
主要政策方向	8
共同的政策主题	10
第一章 评估和测评焦点	15
评估和测评的重要性日益凸显	16
方法论	16
本报告的结构	18
第二章 评估和测评趋势	21
引言	22
评估和测评趋势	22
影响评估和测评的背景	29
参考文献	35
第三章 评估和测评框架：采用整体分析法	39
引言	40
定义和范围	40
教育情境	41
管理	43

总体参考：学生学习目标	52
设计	56
评估和测评能力	70
结果使用	75
评估和测评政策的实施	80
对未来政策发展的建议	87
参考文献	95
第四章 学生测评：以学生为中心	103
引言	104
分析方法	104
影响、驱动因素和情境发展	107
管理	115
程序	130
能力	147
结果的报告和使用	155
对未来政策发展的建议	164
参考文献	173
附录 4. A1 初中学生测评框架特点	185
附录 4. A2 小学和高中学生测评框架特点	213
第五章 教师考核：提高教师的专业素养	257
引言	258
分析方法	258
影响、驱动因素和情境发展	259
管理	265
程序	275
能力	292
结果使用	301
对未来政策发展的建议	311
参考文献	318
附录 5. A 教师考核框架的特点	326
第六章 学校评估：从合规到高质量	359
引言	360
分析方法	360
概念框架	361

影响、驱动因素和情境发展	361
管理	367
程序	387
能力	402
结果的报告和使用	414
对未来政策发展的建议	429
参考文献	436
第七章 学校领导考核：培养学校中的教学型领导	447
引言	448
分析方法	448
影响、驱动因素和情境发展	455
管理	465
程序	474
能力	492
结果使用	495
对未来政策发展的建议	500
参考文献	509
附录 7. A 学校领导考核框架的特点	517
第八章 教育系统评估：为完善教育系统政策提供信息	533
引言	534
分析方法	534
影响、驱动因素和情境发展	536
管理	539
程序	544
能力	564
结果的报告和使用	571
对未来政策发展的建议	583
参考文献	589
附录 A 《OECD 关于改善学校成果的评估和测评框架的评论》是如何形成的	595
附录 B 《OECD 关于改善学校成果的评估和测评框架的评论》中的政策方向	
概要	603
译后记	607

表 格 目 录

表 3.1	不同国家(地区)运用结果进行问责和促进发展的情况(2012 年)	45
表 3.2	不同国家(地区)评估和测评框架中发挥作用的专门的中间机构(2012 年)	48
表 3.3	不同国家(地区)评估和测评框架的全面性与系统性程度(2012 年)	64
表 3.4	评估和测评框架各组成部分内部的关键连接	67
表 3.5	评估和测评框架各组成部分之间的关键连接	69
表 4.1	初中阶段学生测评中使用的参考——ISCED 2 级(2012 年)	129
表 4.2	初中阶段中央考试的特点——ISCED 2 级(2011—2012 年)	132
表 5.1	决定教师考核程序的职责分配(2011—2012 年)	272
表 5.2	不同类型教师考核使用的参考文件(2011—2012 年)	277
表 5.3	参与教师考核过程的评估人员(2011—2012 年)	293
表 5.4	参与教师考核过程的评估人员(2011—2012 年)	293
表 5.5	教师考核对专业发展的影响(2011—2012 年)	302
表 5.6	教师考核对职业决策的影响(2011—2012 年)	305
表 5.7	与教师考核挂钩的奖励(2011—2012 年)	308
表 5.8	对教师表现不达标的应对措施(2011—2012 年)	309
表 6.1	学校评估的目的	368
表 6.2	校外评估的职责	373
表 6.3	校外评估的法律框架	375
表 6.4	学校自我评估的要求	378
表 6.5	学校自我评估的频率	381
表 6.6	校外评估与学校自我评估之间的联系(2009 年)	384
表 6.7	学校评估涉及的领域(中学,2009 年)	387
表 6.8	为合规报告收集信息	395
表 6.9	评估教学与学习质量的课堂观察指标	399

表 6.10	支持学校自我评估的国家举措	406
表 6.11	利用学校评估结果和学校绩效衡量标准进行问责(2009 年)	414
表 6.12	校外问责制评估结果的使用情况(2009 年)	422
表 7.1	各国公立学校校长考核方法分类	465
表 7.2	学校领导考核结果的使用情况	495
表 8.1	政策框架及教育系统评估的发展(2012 年)	540
表 8.2	OECD 成员国教育系统评估战略方法指标(2012 年)	541
表 8.3	OECD 成员国的国家测评(2012 年)	549
表 8.4	国家测评的测试题型(2012 年)	553
表 8.5	OECD 成员国参与学生成就国际研究情况(2012 年)	557
表 8.6	OECD 利益相关方调查的使用情况(2012 年)	559
表 8.7	教育系统评估的纵向信息应用情况(2012 年)	561
表 8.8	国家教育系统和地方教育系统评估机构(2012 年)	565
表 8.9	国家教育系统评估结果报告(2012 年)	572