

System Reform of Educational Evaluation

教育评估的 体制创新

袁益民 著

凤凰出版传媒集团
江苏科学技术出版社



System Reform of Educational Evaluation

教育评估的 体制创新

唐国兴 著

南京师范大学教育科学学院

南京师范大学教育评估研究所



System Reform of Educational Evaluation

教育评估的 体制创新

袁益民 著

凤凰出版传媒集团
江苏科学技术出版社

图书在版编目(CIP)数据

教育评估的体制创新/袁益民著. —南京:江苏科学技术出版社,2007.7

ISBN 978—7—5345—5488—9

I. 教... II. 袁... III. 教育评估—体制改革—研究—中国 IV. G449

中国版本图书馆 CIP 数据核字(2007)第 100524 号

教育评估的体制创新

著 者 袁益民
责任编辑 丁 鹏
特约编辑 钱 亮
责任校对 苏 科
责任监制 曹叶平

出版发行 江苏科学技术出版社(南京市湖南路 47 号,邮编:210009)
网 址 <http://www.pspress.cn>
集团地址 凤凰出版传媒集团(南京市中央路 165 号,邮编:210009)
集团网址 凤凰出版传媒网 <http://www.ppm.cn>
经 销 江苏省新华发行集团有限公司
照 排 南京奥能制版有限公司
印 刷 扬州鑫华印刷有限公司

开 本 850 mm×1 168 mm 1/32
印 张 8.25
字 数 180 000
版 次 2007 年 7 月第 1 版
印 次 2007 年 7 月第 1 次印刷

标准书号 ISBN 978—7—5345—5488—9
定 价 28.00 元

图书如有印装质量问题,可随时向我社出版科调换。

序

近几十年来,我国在普及基础教育和扫除文盲等方面取得了举世瞩目的成就,全民教育水平得到了极大的提高。这些进展不仅在我国历史上是空前的,而且在世界上也是罕见的。教育事业出现了可喜的局面,人们对它的可持续发展也充满了期待。只要坚持改革开放和开拓创新,未来完全可以打造出一个发展更加快速、全面、均衡、协调及和谐的教育。现在的当务之急是要处理好教育发展中还面临的一些挑战,如优质教育资源相对不足与人民群众接受优质教育要求之间的矛盾以及与此相关的教育机会不够均等和教育质量不够理想等问题。

随着教育规模的扩大和普及水平的提高,教育质量、效率和公平的问题越来越受到人们的关注。对于增加入学机会和满足基本的学习需求必须与有质量的教育齐头并进,各国已有较多的共识,对于教育是一项人权、教育公平是一项重要的质量指标的观点也有广泛认同。如果不能保证教育质量,并且是保证建立在平等基础上的教育质量,那么更大规模、更高水平的教育终将难以得到保障。而保障教育质量的一个关键措施是强化教育评估,可以说,质量评估是保障教育质量、效率和公平的基石。所以,关注质量首先就要关注教育评估。

教育质量是个国际通用概念,教育质量保障也是世界各国共同的现实命题,而教育质量评估则被全球各地越来越多的教育同行看作是教育改革与发展的一个优先领域。然而,对于如何开展教育评估,又存在各种各

样的理论、模式和方法上的不同。随着经济全球化和教育国际化程度的提高,国际性和地区性的教育评估活动越来越活跃,各国教育评估系统的相互借鉴、磨合和部分趋同的现象也越来越明显。在此情况下,借鉴基础上的自主创新显得尤为重要。

我国的现代教育评估起步较晚,理论研究和实践活动方面均相对比较滞后,与国外发达国家仍有实际差距。虽然也已经初步建立起教育评估的一些系统架构,有些评估系统的制度化程度也已经比较高,但是整体来说,教育评估体系仍然很不完善,亟待通过体制创新,构建与国际先进的理念、世界一流的模式、全球可比的方法相接轨的并且又具中国特色的现代教育评估体系。

袁益民先生在教育评估体制创新方面所做的这一研究是一项有益的尝试。他从我国教育评估的历史经验和现实问题出发,以比较开阔的视野,分析描述了教育评估体制的一种新的愿景。他的论述中既有一般性的介绍,也有深入的分析;既有心得体会,也有理论反思;既有现实感言,也有真知灼见;既有问题的探讨,也有具体的建议。相信该项研究将能引发更多对于教育评估方面的思考。

袁益民先生长期从事教育对外交流方面的工作,也多年参与一线教育评估工作,对国内外教育有比较多的了解,对教育质量保障有较深刻的理解,对教育评估也有其独到的见解。他从20世纪80年代初开始就与我国教科文组织的工作有很多的联系,参与过不少 APEID 联系中心和 UNESCO 协会方面的活动,还为全民教育九国(E-9)第六次高层会议和 APPEAL 等值教育地区研讨会等起草了国家报告。这些联系加深了他对教育评估标准、政策与体制等的思考。

我的理解是,该书的出版只是他在教育评估体制研究领域的初步成果,期望他今后在这方面开展更加深入

的研究。愿教育评估体制创新方面出现更多的理论与实践成果,愿我国的教育评估体制不断得到完善,为教育质量的提高,为现代学校制度的完善,为教育的现代化提供更加有效和有力的支撑。

中华人民共和国常驻联合国教科文组织代表团大使

张学忠

2007年5月28日

目录

引言	1
第一章 我国教育评估体制概述	6
一、现代教育评估在我国的发展	6
二、教育评估的现有体系	17
三、教育评估的管理体制	27
四、教育评估的现有政策概况	31
五、教育评估的影响分析	32
第二章 对教育评估的理性反思	37
一、当前教育评估存在的主要问题	37
二、教育评估实践的认识误区	43
三、教育评估的管理体制障碍	52
四、教育评估的政策保障问题	55
第三章 评估体制的社会文化及教育背景分析	57
一、教育评估与国情	57
二、教育评估与管理文化	59
三、教育评估与教育系统的特点	62
四、教育评估与教育目标	64
第四章 教育评估面临的新机遇、新挑战	68
一、教育改革与发展带来的机遇与挑战	68
二、经济社会发展提出的新要求	71
三、教育评估本身出现的新局面	73
四、全球化对我国教育评估的启示	75
第五章 国外教育评估体系简介	78
一、国外三类高等教育评估体制的异同	78
二、“E-9”七国基教评估体制简介	79

三、国际及地区性教育评估的进展	92
四、国外评估理论成果与未来趋势	93
第六章 用科学加人文的理念指导评估实践	96
一、树立正确的教育质量观	96
二、树立科学的教育价值观	106
三、树立人文的教育绩效观	111
四、树立有效的教育评估观	116
第七章 展望我国教育评估的创新体制	122
一、体制创新的指导思想	122
二、评估创新体制的基本框架	140
三、评估创新体制的法制基础	143
四、评估创新体制的功能定位	145
第八章 创新体制下教育评估制度的新视角	153
一、完善的评估管理体制	153
二、校本的院校认证制度	157
三、友好的学生评估制度	162
四、有效的考试评价制度	164
五、人文的教师评估制度	170
六、动态的课程评估机制	172
七、专业的管理信息系统	175
八、合理的区域教育评估制度	176
附录一	
与教育评估相关法律、法规文件及相关内容摘录	
一览表	179
附录二	
学校教育加强效能的条件:五项总结性研究的结	
果对照比较表	202
附录三	
2003 年国际学生评估项目(PISA)的研究主题及	
变量	204

附录四	
2005 年教育类卓越绩效准则的结构	206
附录五	
台北市优质学校教育指标	208
附录六	
欧洲学校教育质量指标	209
附录七	
欧洲高等教育内部与外部质量保证的标准与准则	210
附录八	
江苏省县(市、区)教育现代化建设主要指标	217
附录九	
江苏省县级人民政府教育工作督导评估和考核标准	219
附录十	
江苏省县级人民政府教育工作督导评估和考核 工作表	221
附录十一	
联合国教科文组织亚太地区全民教育中期评估 指标	227
附录十二	
国际学校理事会学校评估认证标准(第 7 版)	247
参考文献	252

引言

我国关于教育评估、测量、统计、考试、督导等相关研究成果已有很多,近年来,在基础教育和高等教育质量保障体系方面也取得了研究成果。但是,很少有从整个教育评估政策和体制层面切入的研究,本书试图从这一方面进行探讨。

教育评估中的许多问题最终都必须反映到政策与体制层面,才能得到根本的解决。创新评估体制不仅是发挥现代评估功能、落实评估政策、实现评估目的,及整体提高评估水平的一个有效的切入点,同时也是从根本上认识和解决评估乃至整个教育质量保障问题的必备条件。随着教育大众化、普及化程度的提高,教育质量越来越成为人们关注的议题。规模的扩大和数量的增加不仅不会自动带来质量的提高,而且相反,容易造成质量保障压力的增加,甚至引起质量的相对下滑。而质量方面的问题又会反过来制约就学机会、教育规模、性别平等和均衡发展等方面问题的解决。与此同时,在我国加入世界贸易组织后,通过改进教育质量来提高教育的国际竞争力的任务更为迫切,而随着教育资源的跨国配置和教育服务的跨国提供的增多以及学分跨国转移和学历国际互认需求的增长,教育质量评估标准与评估办法本身与国际接轨的压力也明显增加,教育评估体制及政策制约教育改革与发展的的问题日益凸现。

在这样的背景下,主动反思我国的教育评估的政策,创新和完善评估体制,不仅是保障教育质量的必须之举,改进教育质量的必然选择,而且也是面对经济全球化和教育国际化的必要应答和提升教育竞争力、谋求教育可持续发展的必由之路。作为今后一个时期教育

教育评估中的许多问题最终都必须反映到政策与体制层面,才能得到根本的解决。教育评估的相对落后首先是思想观念上的,同时也是技术方法上的,但是最终是体现在政策及体制上的。

规模的扩大和数量的增加不仅不会自动带来质量的提高,而且相反,容易造成质量保障压力的增加,甚至引起质量的相对下滑。

改革和发展的主题,我国必须积极面对教育质量和教育评估所面临的现实挑战,重新认识教育质量及其保障与改进途径,推动教育评估的体制创新,构建与国际接轨、有中国特色的评估体制,切实保障和改进教育质量。

对教育评估进行政策与体制层面的整体思考,必须对与教育质量、绩效、发展状况等相关的比较宽泛意义上的各类教育评估进行全面审视,包括教育督导、政府主导的管理性评估、社会中介性质的评估活动,教育考试、监测、鉴定,院校认证、行政性审批检查验收、立法机关的审查监督检查、研究性质的评估调查、社会评议与评论等。这些教育评估相关活动尽管在日常工作中并不都称为评估,而且在我国也是非常分散地进行的,相互之间缺乏沟通和联系,但是实质上均属于教育评估性质的活动,在国际上也均被归入教育评估范畴。本书试图聚焦在体制层面和创新角度,对与评估有关的教育系统内外的所有相关活动进行全面的反思,对教育评估体系进行系统的梳理。

高等教育与基础教育的现有评估工作的体制安排(包括项目管理、评估目标、政策依据、评估主体、组织方式和结果使用等)存在一些差别,但是本书仍将它们放在整个教育评估的大框架内进行讨论。这一安排的主要考虑包括:各层次教育评估的内在同一性和关联性(如基础教育阶段的学生综合评价,与高等院校入学前后的评估,甚至是正规教育以后的终生学习项目的评估都有着越来越紧密的联系)使这两个领域联系起来讨论具有内在的合理性;我国教育管理体制的集中化,和各类评估活动管理机关的一致性,使这样的讨论具有合法性;教育评估体制问题的独特性(它与国家教育管理的统一政策、行政架构和运作机制有很高的关联度,两个层面的体制安排问题本来就应该统一考虑)使其具有适切性;国外(如美国)同时涉及基础教育和高等教育的院

校认证的成功实践和已有经验,使其更具可接受性;本书试图探索整个教育评估体制(包括基础教育与高等教育等各个层次)的创新问题的研究定位也使其具有必要性;在统一的研究中,两者之间可以不排除必要的分别讨论的可行性,又使其保持更多的正当性。

所以,对于“教育评估”的概念,这里采取比较宽泛的理解,是指与教育工作对象、机构、项目、状况、现象、事物、活动、政策等相关的,任何基于证据的事实和价值判断的过程,这种判断不论主体和客体的身份和地位的不同以及判断的对象和内容、基点和准则、手段和方法、结果和影响、时间和空间的不同。这一过程开始于对某变量进行测量,继而与前面定义过的参照点或基准相比较,根据测量结果与参数间的距离,作出是否和在多大程度上够标准或不够标准的事实和价值判断。只要符合这些活动和过程的核心要素就属于教育评估。

作为一个功能性概念,“教育评估体制”则是指评估教育相关方面或因素的制度化的成套活动的组织系统,模式化和程序化的管理体系,以及基于法律、法规或其他正式形式规定的,引导一个国家评估活动执行的,具有恒久性的方向定位的制度安排和运作机制的总和。这里侧重于教育评估的组织系统、管理体系以及政策规定的、制度安排的工作模式与运行机制等。英文中的“System”一词在翻译成汉语时可以指“系统”“体系”“制度”“体制”“秩序”“规律”和“方法”等,本书采用“体制”作为该词的在汉语中的基本对应概念。然而,在具体讨论过程中涉及到汉语中的约定俗成的提法时,将尽量尊重汉语中的习惯用法。但是,不论使用哪种汉语表达方式,本书在“系统”“体系”“制度”“体制”之间不作严格的区分。

为了避免理解的混淆,这里对本书中涉及的“评估”和“评价”两个概念的使用作简要说明。一是本书对“评

估”和“评价”概念不作定义上的区分。我国以往的教育评估实践中,一般基础教育阶段使用“评价”的居多(包括研究组织名称、政府文件和日常实际工作中),而高等教育领域更多地是使用“评估”,但其实两者没有实质意义上的区别。尽管注意到有一些学者对这两个概念作了区分,但是认为既然教育评估实践中是没有实质区分的,在外来文献翻译的过程中也是没有统一且明确的区分的,从理论上对其加以区分也没有实际意义,而且人为地或者牵强地给予不同的定义,可能会给理论与实践之间的对话带来困惑,给国内与国外之间的对话带来混淆。二是对“评估”和“评价”两词在不同语境出现的实际情形不作语义上的统一。既尊重工作实践中的实际使用惯例,也尊重不同语境和语言搭配关系中的习惯用法。本书把“评估”作为一个更加宽泛的概念,用它来等同和涵盖“评价”概念,并倾向于在涉及政策、体制及宏观管理时尽量使用“评估”概念,更多地对应于英文中的“Evaluation”,而在讨论手段、技术以及具体方法时,主张不排除使用“评价”概念,更多地对应于英文中的“Assessment”。涉及文件和机构名称等的直接引述完全尊重原文,间接引述也尽量尊重原文。

研究教育评估的整个体制如何创新的问题必然要首先基于对现有体制的认识,同时需要讨论与体制建立直接相关的政策、管理体系、系统目标等方面问题,还需要涉及现有的主要评估活动及其影响。为了使评估体制的创新建立在比较宽广的视野之上,还应该关注教育评估系统内外、教育系统内外和国内外的相关方面的发展状况和趋势。所以,本书以教育评估体制创新所涉及的基本问题为主线,力求从历史、现状和未来的不同维度,从理论、实践、政策的不同层面,从公共管理、教育和评估等不同的系统,从政府、社会和学校不同角度,从地方、国家和国际的不同视野出发,多方位地审视和宏

观分析教育评估问题,以便为教育评估体制的创新建立一个比较宽广而坚实的认识基础。本书的结构分为三大部分:现状的概述、反思和相关分析;新情况、新趋势和新理念;未来的方向、架构和制度。

大凡涉及体制创新,特别是作为一个十分复杂的社会科学研究领域和一个十分敏感的社会实践活动的教育评估体制的创新,都不会是简单或孤立的工作,而会涉及到许多方面的改革。同时,教育内外的环境变迁决定了教育评估的体制创新永远也不会有一个一劳永逸的解决方案,它必然是一个常创常新的课题。但是,如果能够摆脱“问题追赶型”研究模式,除了反思自己的不足,积极吸收别人的经验,适当超越时空限制,努力通过更宽的视野、未来的导向和长远的眼光,对教育评估体制进行根本性的重新反思和基础性的重新建构,那么,探索一套不会很快被“改革”的改革不是完全没有可能的。

当然,教育评估体制问题本身涉及到对于许多根本性问题的回答,如它的功能定位、各类评估之间的关系、评估机构角色的定位与监管措施、如何保障评估本身的质量、教育评估的法制保障等,不可能在一本书中得到完整的回答。本书提出的一些革新思路充其量只能算是一种可供选择的建设性方案,如果能够给予那些对教育评估的政策与体制感兴趣的决策者、研究人员和其他教育工作者一点新的启示,就足可聊以自慰了。

评估是一个十分复杂的社会科学研究领域,一个极易引起争议的理论范畴和一个十分敏感的社会实践活动。

第一章 我国教育评估体制概述

教育评估体制的创新是建立在体制的现实基础之上的,也扎根于其历史背景和社会环境的土壤。所以,在讨论体制创新的思路和具体建议之前首先有必要对现代教育评估在我国的发展作一简要的回顾和大致的总结,对教育评估的现有体系、管理体制和政策作一介绍,并对现有评估的影响作一初步分析。

一、现代教育评估在我国的发展

(一) 世纪回顾

1. 总体印象

作为古代教育评估制度的封建科举制度在 20 世纪初废止后,由于众多原因,现代意义上的教育评估的理论研究与实践活动在我国一度没有得到很好的发展。直到 1977 年我国恢复高考制度以后,为了适应教育改革的发展的需要,迅速地提高学校教育的质量,保证教育政策的执行和目标的实现,教育评估研究和实践才逐步地得到了恢复和发展。

回顾近一个世纪我国教育评估走过的坎坷道路,可以比较安全地作出这样的判断:现代教育评估在我国的产生、生长、曲折和发展有其深刻的社会根源。从上个世纪初教育测量引入我国并得到一定的发展,到曲折徘徊后陷入停顿;从现代教育评估的空白期,到一边倒的单一苏联模式;从无所适从的政治迷茫期,到理论研究的间断;从恢复高考到最近二十多年来的逐步发展等。我们可以清楚地看到,政治社会因素是影响教育评估在

政治社会因素
是影响教育评估在
我国发展的主要
动因。

我国发展的主要动因。

我们从教育评估在教育内外的环境因素影响下自身发展的线索和规律这一个侧面也不难看出,在上个世纪大部分时间里,教育评估的发展显得没可能、没需求、没必要。如果民族的生存问题还没解决,动荡的社会里不能放下一张可以安静读书的书桌,那就谈不上教育的发展,更谈不上促进教育质量提高的评估的发展;同时,如果教育的生存问题还没解决,老百姓的基本学习需求远远不能得到满足,那就还不需要过早地考虑更高层次需求的满足,就没必要花时间和精力来评判是非曲直、善恶功过的问题,也没有评论好坏优劣、成败得失的正当前提,更没有评价公平效益、价值附加的“奢侈”需求;再者,如果连评估工作在教育发展中应该发挥什么作用的问题还没有解决,学校教育过多地受到外部力量的影响和左右,或者教育管理仍然实行政府大包大揽,教育的计划、实施和总结反馈“一竿子到底”,那也就没有开展独立的评估活动和独立的评估研究的必要。

从这个意义上来说,现代教育评估作为一种专门的社会活动,其发展所需的社会条件在我国过去一个世纪的大部分时间是不充分的,其作为一个相对独立的专业和完备的制度而存在的理由更是不存在的。这也从根本上部分地解释了为什么现代教育评估传入我国后,虽然在漫长岁月中也有总体的上升(如研究在深入,重视程度在加大,活动在增多,政策在完善等),但一直未能发展到理想的水平。

2. 主要特点

百年教育评估发展主要呈现出六个基本特点。

(1) 非连续性。封建科举制度与现代教育评估制度之间没有很好地过渡和衔接,在引入和开展现代意义上的教育评估的过程中的大部分时期,也没有出现较长时