



高等学校 绩效评价研究

张男星 等 著



科学出版社

高等学校绩效评价研究

张男星 等 著

科学出版社

北 京

内 容 简 介

本书较为系统、全面地论述了国内外绩效评价的理论基础、政策依据和研究现状,并以实证研究为特色,在持续研究的基础上,综合运用主成分分析、典型相关分析、DEA 分析、SFA 分析等多种主流量化评价分析方法,对我国 72 所教育部直属高校进行了投入与产出绩效评价和分析,同时基于资源优化配置的视角,从优化教师队伍、经费结构和管理创新等方面对高校绩效改进进行了探索。本书较好地体现出了理论研究和实证研究相结合、学术研究和工作的研究相结合、量化研究和质性分析相结合的特点。另外,本书对研究过程中的难点和关键问题进行了深入剖析,研究过程严谨,研究框架合理,研究结论真实可信,对引导政府改进资源配置,促使高校提高资源使用效率、提高绩效,具有重要作用。

本书可供高等学校和科研机构中的经济学、管理学、教育学等方面的研究人员、教师和学生,高校管理者及其他相关研究人员,以及政府相关管理人员参考,也可以供在职学习经济学、管理学、教育学的学生及攻读 MBA 和 MPA 的学生参阅。

图书在版编目(CIP)数据

高等学校绩效评价研究/张男星等著. —北京:科学出版社, 2018.3

ISBN 978-7-03-056206-7

I. ①高… II. ①张… III. ①高等学校—学校管理—研究—中国 IV. ①G647

中国版本图书馆 CIP 数据核字(2017)第 323132 号

特邀策划:来豫蓉

责任编辑:王丹妮/责任校对:王萌萌

责任印制:霍 兵/封面设计:无极书装

科 学 出 版 社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

艺堂印刷(天津)有限公司 印刷

科学出版社发行 各地新华书店经销

*

2018 年 3 月第 一 版 开本:720×1000 B5

2018 年 3 月第一次印刷 印张:9 3/4

字数:202 000

定价:82.00 元

(如有印装质量问题,我社负责调换)

前言

当今社会与世界的巨大变革已经对高等教育产生了深远的影响，而且这种影响还将持续下去。一方面，世界的复杂性、时代的不确定性和科技给经济社会带来的发展，导致社会普遍承认高等教育的重要性；另一方面，高等教育也正在成为社会各界批评的焦点。有人认为当今大学是一个巨大的、自私的甚至是贪婪的“怪物”，它向学生收取昂贵的学费，为科研向政府索要不合理的开支。而随着大学融入社会程度的提高，大学面临着新的有关其责任的争论，如公众对科学的合法性及旧知识生产者的自我合理性提出了挑战，审计社会的到来迫使大学实行新的管理体制并对社会负责。质量作为管理的核心与抓手，特别强调泰勒主义的任务划分、全面管理控制和成本核算。绩效评价作为高等教育质量管理的重要方式之一，就兴起于这一西方社会的发展情境之下，并通过全球化对质量关照的话语渗透，迅速超越了民族和国家的界限，体现出一种主要的政策借鉴和实践模仿的路径依赖。

近年来，我国在国家政策层面大力倡导开展绩效管理和评价，地方政府在积极推进高校^①绩效评价，教育部直属高校依然在高度关注高校绩效评价，学者也依然在持续研究与试验高校绩效评价。这说明在我国对高校开展绩效评价，已经在高等教育领域得到了更为广泛的认同，并且国家、地方、高校、学界等多方对此都达成了共识。

中国教育科学研究院绩效评价课题组（以下简称课题组）从 2008 年起一直致力于高校绩效评价的研究，一方面不断深化理论找寻高校绩效评价的本真本源，另一方面不断改进方法探索高校绩效评价的优化运行。依据持续研究在指标设计、方法模型、绩效呈现等方面的相对稳定性与连贯性，本书针对直属高校 2006—2015 年这 10 年的绩效状况展开研究，借此分析高校绩效及其与资源配置的关系问题。

一、高校绩效评价的说明

本次绩效评价（2016）继续沿用往年的投入、产出指标，包括人力、财力、物力投入指标和教学、科学研究、社会服务产出指标，对直属高校实行分类绩效

① 为论述方便，全书中的“高校”与“高等学校”不做统一。

评价（综合类高校 27 所、大理类高校 32 所、大文类高校 13 所）。

课题组通过实证分析表明：一是评价年限长短对绩效结果不造成影响，但评价年限延长便于对不同年份的绩效结果进行纵向比较，本次绩效评价将评价周期从 2006 年延长至 2015 年；二是指标无量纲化方法对绩效结果的影响甚微，但非线性无量纲化方法更能反映出指标的现实意义，本次绩效评价依然采用指数型功效系数对指标进行无量纲化处理；三是传统主成分法赋权与修正主成分法赋权没有统计学意义上的显著差异，本次绩效评价采用修正主成分法作为运行模型。

二、高校绩效评价的结果

本书运用主成分分析及 DEA^①模型、SFA^②模型等方法对 2006—2015 年高校的办学绩效进行评价，并采用位差法（产出排序与投入排序的差值）呈现出“绩效偏高”“绩效相当”“绩效偏低”3 类绩效表现结果。相关评价结果与具体分析如下。

第一，直属高校整体绩效状况相对稳定，近一半高校绩效偏高，大文类高校中绩效偏高的比例略高。72 所直属高校中，“绩效偏高”类别的高校共有 32 所，占总数的 44%，这些高校的投入获得了更多的产出，相对效率较好；“绩效相当”类别的高校共有 12 所，占总数的 17%，这些高校的投入获得了相应的产出，相对效率平稳；“绩效偏低”类别的高校共有 28 所，占总数的 39%，这些高校的投入未能获得相应的产出，相对效率较弱。

第二，4 成左右综合类和大理类高校、3 成左右大文类高校的教学绩效偏高。72 所直属高校中，教学“绩效偏高”类别的高校共有 30 所，占总数的 42%，其中包含 13 所综合类高校，12 所大理类高校，5 所大文类高校。教学“绩效相当”类别的高校共有 15 所，占总数的 21%，其中包含综合类高校 5 所，大理类高校 5 所，大文类高校 5 所；教学“绩效偏低”类别的高校共有 27 所，占总数的 37.5%，其中包含综合类高校 10 所，大理类高校 14 所，大文类高校 3 所。

第三，4 成左右综合类、大理类和大文类高校的科研绩效偏高。72 所直属高校中，科研“绩效偏高”类别的高校共有 29 所，占总数的 40%，其中包含 14 所综合类高校，10 所大理类高校，5 所大文类高校。科研“绩效相当”类别的高校共有 15 所，占总数的 21%，其中包含综合类高校 2 所，大理类高校 10 所，大文类高校 3 所；科研“绩效偏低”类别的高校共有 28 所，占总数的 39%，其中包含综合类高校 11 所，大理类高校 12 所，大文类高校 5 所。

第四，一半以上大文类高校社会服务绩效偏高，4 成左右综合类与大理类高

① DEA：数据包络分析(data envelopment analysis)。

② SFA：随机边界方法(stochastic frontier approach)。

校社会服务绩效偏高。72 所直属高校中, 社会服务“绩效偏高”类别的高校共有 33 所, 占总数的 46%, 其中包含 12 所综合类高校, 14 所大理类高校, 7 所大文类高校。社会服务“绩效相当”类别的高校共有 8 所, 占总数的 11%, 其中包含综合类高校 2 所, 大理类高校 4 所, 大文类高校 2 所; 社会服务“绩效偏低”类别的高校共有 31 所, 占总数的 43%, 其中包含综合类高校 13 所, 大理类高校 14 所, 大文类高校 4 所。

第五, 高校的教师投入结构和经费投入结构对高校绩效有显著的作用。在内部投入结构中, 学校的专任教师越多, 绩效水平越高, 但高校行政教师所占比例越高, 学校的绩效反而越低。在经费投入结构中, 财政拨款的比例和事业收入的比例对高校绩效的影响都是正向的且显著, 并且事业收入所占比例更大, 其他投入因素的作用不明显。

第六, 高校间绩效差异不在于区域, 而在于学校类型。尽管东、中、西部地区经济差异明显, 但区域间直属高校的绩效结果不存在显著的差异, 直属高校间绩效的差别主要体现在学校类别上。是否为“985 工程”高校对高校绩效有显著的正向影响, 但高校在办学历史上是否曾经历过合并并不影响高校的绩效水平。

三、提升高校绩效的建议

第一, 调整优化教师队伍结构。深入进行高校人事制度改革, 建设高质量、高水平的教师队伍。合理调整各类人员的比例结构, 提高科研和教学教师所占的比例, 适当降低和缩减一般行政教师所占的比例, 积极引进高水平的师资。高校教师队伍是实现高校绩效水平提升的重要因素, 优化教师队伍的结构不仅有利于提升高校的科研效率, 对提高教学效率和提升人才培养质量的效果也很明显。

第二, 完善高校经费的投入结构。高校不仅要有充足的经费来源, 还要能充分并有效使用这些资源。总体上, 财政拨款收入的比例为 62.5%, 事业收入的比例为 31.4%, 其他收入的比例为 6.1%。事业收入包括学费和科研事业收入, 主要由高校自主支配, 其所占比例对高校绩效的影响略大于财政拨款所占的比例。学费收入与招生规模有关, 此项收入基本固定。因此, 可以适当提高科研事业收入所占的比例, 保证财政拨款的比例, 提升高校的财务自主性水平, 提高经费使用效率。

第三, 提升高校管理创新水平。对于效率较低的高校在管理中遇到的复杂问题, 如创新效率、纯技术效率略低等问题, 要借鉴高效率学校的经验, 完善运行机制, 提升自身的管理水平。另外, 要明确责任主体, 强化对学校图书馆、教室及实验室的质量管理, 提高管理效率和办学效率。同时, 要完善高校内部管理体制, 优化管理流程, 提升技术水平, 实现技术进步和创新。

第四,给予高校充分的稳定发展的时间和空间。本书使用 10 年的数据发现,大文类院校的绩效表现相对较好,表明高校发展需要时间的积累,其产出具有明显的滞后性。因此,政府要给予高校充分的时间和空间,高校应当遵循教育规律,充分发掘自身特色和优势,通过管理、制度的改革追求稳步发展,而不能急功近利。

目 录

第一章 高校绩效评价发展的思考	1
第一节 高校绩效评价的理论追问	1
第二节 高校绩效评价的实施要素	12
第三节 高校绩效评价的影响作用	18
第四节 高校绩效评价的趋势	23
第二章 高校绩效评价的中外比较	27
第一节 高校绩效评价实施的异同	27
第二节 高校绩效评价指标的异同	41
第三章 高校绩效评价的实施及结果（2016）	51
第一节 高校绩效评价实施思想	51
第二节 高校绩效评价过程	56
第三节 高校绩效评价总体结果	78
第四节 高校绩效评价分项结果	80
第五节 总体绩效与分项绩效分析	84
第六节 小结	87
第四章 高校绩效数据包络分析	88
第一节 数据包络分析方法	88
第二节 相对效率不同时间点的评价结果	92
第三节 相对效率纵向动态的评价结果	101
第四节 小结	108
第五章 高校绩效随机前沿分析	109
第一节 随机前沿分析方法	109

第二节 高校绩效随机前沿分析	113
第三节 高校办学绩效分析结果及绩效变化趋势	117
第四节 小结	123
第六章 高校投入和产出趋势及绩效影响因素	125
第一节 不同类型高校投入变化趋势	125
第二节 不同类型高校产出变化趋势	131
第三节 高校绩效影响因素分析	138
第四节 小结	141
第七章 高等学校绩效评价的未来	143
后记	148

第一章 高校绩效评价发展的思考

对高校进行评价原本就是一项困难的活动,对高校进行绩效评价更是充满争议与风险,这也使得有关高校绩效评价的研究面临更大的挑战。近些年,我国国家层面在大力倡导开展绩效管理和评价^①,地方政府依然在积极推进高校绩效评价^②,教育部直属高校依然在高度关注高校绩效评价^③,学者也依然在持续研究与试验高校绩效评价^④。这说明对高校开展绩效评价已经在高等教育领域得到了较为广泛的认同,并且国家、地方、高校、学界等多方达成了共识。在这样一个新的背景下,高校绩效评价的研究也必然要进入新的历史阶段。那些依然存在的实践困惑,迫切需要在理论层面进行深入探讨。

第一节 高校绩效评价的理论追问

到目前为止,我们可以粗略地将我国高校绩效评价的研究分为三个阶段。第一阶段以单项研究为基本方式,以1994年《高校后勤管理工作绩效评价及作用》^⑤一文的公开发表为表征,持续到2008年,主要针对高校后勤、财政、教师或科研等某个领域进行单项绩效评价研究。第二阶段组织研究成为新生方式,以2009年中央教育科学研究所高等教育研究中心公开发表的《影响高校绩效的主因是资源有效利用》^⑥一文为表征,持续到2015年。这一阶段,单项研究依然存在,

① 2012年,党的十八大报告提出,要“创新行政管理方式,提高政府公信力和执行力,推进政府绩效管理”;2013年,党的十八届三中全会通过的《中共中央关于全面深化改革若干重大问题的决定》提出“严格绩效管理,突出责任落实,确保权责一致”。

② 2010年,上海市政府在出台的《上海市中长期教育改革和发展规划纲要(2010—2020年)》中明确指出“将财政投入与高等学校绩效考核衔接起来”,同时,2011年上海在全国率先开展了区域性高等学校整体办学绩效评价的实践探索。

③ 2014年,教育部直属高校工作咨询委员会的大会报告中明确提出“以科学评价为基础,通过绩效拨款,引导高校内涵发展、提高质量”。

④ 在中国知网以“高校绩效评价”作为关键词检索,2010—2013年共有近400篇有关高校绩效评价的学术论文,各年份分别为79篇、90篇、99篇和130篇。

⑤ 从中国知网检索的文献来看,这是国内第一篇研究“高校绩效评价”的文章。参见:冷寿中.1994.高校后勤管理工作绩效评价及作用.中国高校后勤研究,1:74.

⑥ 中央教育科学研究所高等教育研究中心.2009-12-09.影响高校绩效的主因是资源有效利用.中国教育报,第1版.

但把高校作为组织机构进行综合性绩效评价的研究则是重要的新生方向,并且地方政府或教育管理部门切实在实践中推行综合性的高校绩效评价,如上海市、重庆市、江苏省等。不少高校自己也在主动探索绩效评价的路径,如同济大学等。第三阶段是以多样研究为存在方式,以“教育部直属高校工作咨询委员会 2014 年咨询报告”为表征拉开序幕。该咨询报告意味着以中国高水平大学为代表的高校群体对绩效评价及绩效拨款的集体认可。这个阶段呈现出“集结式”特点:一是内容集结,即单项研究与组织研究等多种类型互补完善,呈现出点一面结合且更加深入、细化的呼应状态;二是方式集结,即理论研究与操作研究彼此支撑,呈现出理论深化与实践深化同步推进的融合状态;三是主体集结,即学界研究与业界研究协同,且随着政府、科研机构、高校及企业等不同主体的介入,呈现出学术主体、高校主体等多方主体共同研究的协同状态。三个阶段不是呈独立分割的条形式发展,而是呈包含、递进的螺旋式发展,后一个阶段是对前面阶段的覆盖与深化。

进入新阶段的高校绩效评价研究,必须要完成自身的理论建构与自证,必须主动思考与回应有关高校绩效评价的实践困惑,这些困惑已经成了推进高校绩效评价面临的重大实践和理论问题。

一、高校绩效评价的逻辑视角

是否可以对高校进行绩效评价,其实这个问题也可以变为“教育可评吗?”,其间隐藏的是对教育评价存在合理性的思考。尤其是当高校绩效评价以一种“投入—产出”的计量方式出现的时候,这种将企业使用的经济学概念绩效评价用于高校评价的做法是否合理,更加值得关注。新公共管理理念和方法下的绩效管理,宗旨就是实现既定的目标,而目标“绝大部分是用经济学的语言来界定的”^①。有论者提出“高等教育的产出(教学、科研、社会服务)均属于学术成果,不同于企业生产的产品和政府管理的行政绩效”^②。似乎高校作为培养人才的地方,开展的是人才培养活动,用企业使用的绩效管理方法来评价复杂的教育教学活动,是不合理、不科学的。

概括有关高校绩效评价合理性的疑惑,不外乎以下几种:一是评价对象高等学校不是物质生产部门,不能照搬企业评价的理论和方法。企业的特点决定了它以利润的获取为主要目的,产出绝大多数可以用物品或金钱来衡量。而高等学校的属性决定了它属于准公共生产部门,具有公益性,其产出更多指向人的发展状态,而非可以计件的物品。因此,用绩效评价来衡量高校的发展,有失公允。二

① Poltt C. 1990. Managerialism and the Public Service: The Anglo American Experience. Oxford: Basil Blackwell: 5.

② 宋燕. 2012. 我国高校绩效评价的反思及展望. 评价与管理, 6: 5-7.

是高等学校是一个复杂的、多样化的体系,难以准确把握各级各类高校间的评价定位及差异。高校由不同层级和类型的学校构成,有专科、本科、研究型等不同层级之分,有职业教育和普通教育等不同类型之别,甚至还有各种混合型、综合型的高校。要对这些高校一一进行标准的分类,本身就是一项复杂困难的工作,更何况是分类的评价。三是在评价高校的投入和产出的时候,是很难量化的,尤其是那些无形的投入与产出。例如,高校产出有多种形式,既有可以量化的产出,如学生培养人数、发表的论文、科技发明等,也有难以测评的产出,如教育公平、教育价值、校园文化等。相比之下,在高等学校,无形的产出的作用丝毫不亚于有形的产出。以量化的指标,很难对一所学校的办学绩效作出定论。基于类似因素,有论者认为“高校绩效评价要做到非常完美,是非常困难的,或者说是几乎不可能的”^①。

有关高校绩效评价合理性的疑虑,尤其是用定量的指标来评价复杂办学活动的应然性反思,反映了对高校办学进行绩效评价的困难和挑战,但我们也可以从三个方面来分析高校绩效评价存在的合理性及其发展逻辑。

1. 本体论视角

从教育学逻辑看,高等学校是追求高深学问的场所。无论大学的职能怎样演变和发展,高等学校始终从事着传承知识、生产知识、传播知识和转化知识的活动,并在这些活动过程中承续、生发、完善着其自身的职能。从这个视角看,绩效评价在高等学校中的实施可以看作是高等教育适应知识生产模式变化的要求。^②甚至有论者认为,“知识经济时代必须审视发展状态又了解经营环境,才被认为是具有适应性和反思性的机构,才能获得生存和发展”^③。既然知识是可以生产、保存、传播、转化运用的,那么在知识形态流变的过程中,就必然会产生出可以计量的要素,如生产了多少?什么时候生产的?多少人参与了生产?生产的知识被运用了多少?等等。这些源自于高校本体存在过程中的计量要素,就是对高校进行绩效评价的合理性原子。本体论回答了高校的活动是可以计量的问题,至于如何计量则是合理性以外的思考。但已有学者在对知识生产模式进行研究,如吉本斯等研究者围绕不同学科知识生产活动的历史演变及发展,将知识生产划分为生产模式1和生产模式2。^④知识生产模式1,即知识的契约式学科生产模式,表明知识生产是在传统的学科认知环境中进行的,知识的生产与积累主要是通过

① 周雪峰. 2011. 绩效评价很难完美,但定位要准. 大学(学术版), 10: 27-28.

② 郭芳芳, 张男星. 2012. 高深知识的生产变革与高等教育绩效评价. 复旦教育论坛, 6: 5-9.

③ Senge M P. 2010. The fifth discipline: The art and practice of the learning organization. Performance Improvement, 30(5): 37.

④ 迈克尔·吉本斯, 卡米耶·利摩日, 黑尔佳·诺沃提尼, 等. 2011. 知识生产的新模式. 陈洪捷, 沈文钦, 等译. 北京: 北京大学出版社: 21.

大学中被制度化了的学术职业来实现的。知识生产模式 2, 即知识的大众式跨学科生产模式, 表明知识是在一种跨学科的应用情境中生产的, 并以社会弥散体系为主要组织形式, 呈现出非制度化、非等级性、多样化的特点。对于知识生产如何被计量的问题, 可以通过知识生产模式的演变加以剖析。

2. 政治论视角

大学的发展, 逐渐延伸出科学研究、社会服务和文化传播等功能, 不仅有其探究高深学问的本体论存在缘由, 而且有为社会和国家服务的功能的凸显。正如布鲁贝克所说: “大学存在两种高等教育哲学, 一种哲学是以认识论为基础的, 另一种哲学则以政治论为基础。”^①高等教育的主要职能是保存、传授和发展高深学问, 但同时也担负着为公众服务的职能。随着高校的发展及其与政府之间关系的演变, 高校被普遍要求有效运行并被实施问责。过去 20 多年里, 西方发达国家对于高等教育公共财政的支出, 即“钱的价值”问题越来越关注。^②在美国大学于市场导向的策略下特别凸显“绩效责任”与“绩效管理”的同时, 欧洲各国政府也开始把绩效拨款方式引入大学发展中。^③如英国教育与科学部 1985 年在其绿皮书《20 世纪 90 年代高等教育的发展》中明确提出: “虽然衡量高等教育绩效具有很大的困难, 但是, 如果政府要实现控制公共支出和促进纳税人金钱使用达到效益最大化的目标, 如果高校和与高等教育相关的其他方面要在资源分配决策中处于有利形势, 就必须进行这样的努力。”^④许多国家的政府都出台了法律, 要求大学参与评价, 显示办学的效果和效率, 即办学绩效。

3. 实践论视角

20 世纪 70 年代以来, 绩效评价被广泛用于高等教育领域, 绩效评价结构在一些国家甚至与经费等资源分配环节直接挂钩。^⑤我国从 20 世纪 80 年代末 90 年代初开始引入“绩效评价”的概念, 在人才管理、评价激励等方面进行了探索; 进入 21 世纪, 我国政府对绩效管理日益重视, 绩效评价已经在企业、政府部门、医疗卫生等各行各业广泛运用, 这也包括高等教育领域的教师、科研、后勤、财政的绩效评价在内。2010 年颁布的《国家中长期教育改革和发展规划纲要(2010—2020 年)》明确提出“引入竞争机制, 实施绩效评估, 进行动态管理”,

① 约翰·S. 布鲁贝克. 1998. 高等教育哲学. 王承绪, 郑继伟, 张维平, 等译. 杭州: 浙江教育出版社: 13

② Geuna A, Martin B R. 2003. University research evaluation and funding: An international comparison. *Minerva*, 41: 277-304.

③ Orr D. 2005. Can performance-based funding & quality assurance solve the state Vs. market conundrum? *Higher Education Policy*, 18: 31-50.

④ Britain G. 1985. *The Development of Higher Education into 1990s*. London: H. M. S. O: 49.

⑤ Gaither G, et al. 1995. *Measuring up: The promises and pitfalls of performance indicators in higher education*. ASHE-ERIC High Education Report, 66(Pt 5-6): 4.

2012年党的十八大提出“创新行政管理方式,提高政府公信力和执行力,推进政府绩效管理”,党的十八届三中全会通过的《中共中央关于全面深化改革若干重大问题的决定》也进一步提出要“严格绩效管理,突出责任落实,确保权责一致”,尤其是倡导市场在资源配置中发挥决定性的作用,更是极大地推动了高等教育领域内的绩效评价。这些政策和实际做法从实践层面增强了高校绩效评价的合理性。

同时,我们也必须认识到,绩效评价属于一种评价类型,每一种类型的评价都可以通过形成不同的评价指向、采用不同的方法论、运行不同的方法模型来实施。所以,不可以同一个批判层次混淆评价类型存在合理性与评价方式合理性这两种不同性质的问题。就实践效果来看,国外对高校绩效评价一直也存有争议,但尽管如此,绩效评价一直被政府和高校所重视并不断完善。以英国为例,尽管其高校绩效拨款模式遭到多方人士的怀疑、责难和批评而没有得到彻底推行,但它的提出在很大程度上冲击了英国高等教育的传统拨款理念,为此后的改革打下了必要的基础。^①我国对高校绩效评价的探索虽然时间短,但也得到了许多高校管理者和专家的认同。武汉大学前校长顾海良教授认为,“无论从哪个方面看,高校绩效评价确实是对中国高校评估模式的一种积极的、开创性的探索”^②。

二、高校绩效评价的量化方式

绩效评价能够通过量化方式反映出高校的办学质量吗?对这个问题产生疑问的人似乎认为,无论如何都难以用一些量化指标把高校的办学质量表征出来。其原因如下:一是绩效指标本身不可穷尽,高校的办学涉及方方面面,很难用十几个或几十个指标就能涵括;二是有些指标难以量化,如学校理念、精神、文化,就很难用具体指标衡量出来。因此,巴恩斯和巴尔在20世纪80年代就提出,“绩效指标几乎不能测量真正的绩效,它们测量的只是那些能够测量的”^③。

的确,基于高等教育的特殊性,高校在办学过程中并非所有活动及其结果都可以作为变量被观测到。而且,源于技术和数据的限制,即使被观测到的那些活动及其结果,也未必都能够被形成指标并予以量化。另外,囿于质量的内隐性及对质量标准认知的不统一,那些予以量化的活动及其结果又不一定能够完全代表质量,这是高校绩效评价的难题,相应地,疑惑也在所难免。但是,研究遇到难题并不等于所研究的问题不可能得到解决,即使此时不能完全解决,也可以部分

① Hannah S B. 1996. The Higher Education Act of 1992: Skills, constraints, and the politics of higher education. *Journal of Higher Education*, 67(5): 498-527.

② 顾海良. 2013. 高校绩效评价是对中国高校评估模式的积极探索——兼评《高等学校绩效评价论》. *教育研究*, 3: 148-149.

③ Barnes J, Barr N. 1988. *Strategies for Higher Education: The Alternative White Paper*. Edinburgh: Aberdeen University Press: 15.

解决;即使当下不能够解决到位,也可以逐步完善。事实上,国外同行在研究中也存在类似的困惑,但是他们仍然认为,“通过可以观测到的指标,是能够反映出一些关键问题的”^①。任何关于高校办学质量的判断,除了观念性感觉意见的表达,都必然是能够以数据或事实来说明事物物质性本真的思维活动,不存在无数量支撑的抽象式的质量呈现,所以数量或量化与质量之间并不是必然矛盾的。

首先,高校办学的质量是建立在数量化的活动积累之上的。当高校自身的发展在缺乏数量的时间积累并达不到一定数量规模的时候,很难奢谈质量。毕竟我们要评价的是现代大学,而不是古老的私塾。即使是实施精英高等教育的国外高水平小型大学,也是有上千人规模的支撑,而且办学历史都不短,比如,美国1891年建立的加州理工学院,2013—2014学年在校生也达到了2181人。^②因此,只有当事物活动中的各种要素在数量上不断积累、集聚、增长时,才能够有力量引发事物的质量得以生长与显现,一所大学的发展更是如此。高校办学的活动要素无非是人的要素、物的要素、财的要素,分别以教师和学生、基础设施设备、各种经费等为主要活动载体及条件。如果这些要素没有时间积累且达不到一定的数量并形成一定的规模,现代大学的办学质量就是空谈。正因为如此,可以通过对高校办学活动要素的量化,来反观高校的办学质量处于什么样的程度,同时反观高校办学不同方面的质量状况。因此,将数量与质量进行绝对的等同或者割裂,对于高校绩效评价而言都是有误的。

其次,高校办学的质量是可以通过结构化的“数量群集”来表达的。质量的表达可以是感知式的定性经验判断方式,这种方式主要是基于大量的现象观察形成高度的抽象概括,属于归纳式的方法论应用。但这种方法的使用存在很大的对象性局限,致使同一个评价主体很难在同一个时刻大量感知到众多对象的发展现状,也很难从中抽取特点进行概括性的判断与比较,而且这种感知式的判断往往只能得出模糊性的结论,很难达到较高的准确度。再者,不同的评价主体,由于个人经验的差异性,体现在事物判断的结果方面就很有可能完全不同,事物质量的评价认同度也难以实现统一。那么,如何把众多需要评价的高校放在相同时刻进行质量的比较呢?这个时候还需要运用质量表达的另一种方式,即结构式的量化判断方式。这种方式主要是基于对质量的观念性理解形成反映质量的指标群,将这些指标作为观察要素,去聚集观看所要评价的高校并进行比较,属于演绎式的方法论应用。在这种方法中,从某种程度上而言,结构化的数量指标就是质量的表达,可以说任何事物的评价都是基于结构化指标的量化而来的。所以,基于

① Horne J, Hu B. 2005. Estimation of cost efficiency of Australian Universities. *Mathematics & Computers in Simulation*, 78 (2-3): 266-275.

② 加州理工学院. 2015. 学校概览. <http://www.caltech.edu>[2015-01-19].

一定理念、指导思想、方法论所形成的高校绩效评价的指标维度与结构化的指标群，本身就是对高校办学质量的一种表达方式。至于对高校办学质量的理解是否认可，则另当别论。

在研究中运用数学工具进行分析，有很多优势，比如，严密、科学的逻辑推理和数理及计量分析，能够使得研究得出更有意义的科学结论。在今天万物都以数的方式存在的大数据时代，在进行高校绩效评价时，必须高度重视对数及数学方法的使用。对高校绩效评价的批判性认识，要消除量化手段与追求质量评价之间的对立，充分看到数量化绩效指标群中所暗含的的质量的内涵与所指。

三、高校绩效评价的维度指标

如果结构化的指标群可以或多或少地反映高校活动的质量，那么接下来的问题是：这些指标是否能够独立而完整地、从牵连复杂的高校整体办学活动中剥离出来？换言之，怎样看待类似于高校办学这类教育指标之间的“绝缘性”？

教育评价，包括绩效评价中的一些指标会存在不独立，甚至相互包含与交叉的现象，从而影响评价结果的有效性。这种担忧不难理解，首先，我们难以严格区分出归属不同高校职能的办学活动。高校的三大职能相互支撑、相互影响，具有高度的统一性，且对于三者的分化，也多是从理论研究的分析角度进行的，事实上，实践中很难严格区分出某种办学活动是纯粹归属于某种职能的活动，甚或根本就不存在绝对单一的职能活动。如同“产学研一体化”思路下的科学研究，其间同时也包含着人才培养和社会服务的活动。即使我们倡导要重视教学，但教学不也同样需要并包含科学研究吗？只不过是称之为教学研究而已。其次，即使在同一类职能的活动中，我们也很难准确分化出单纯指向某种产出的办学活动。高校办学活动所形成的影响或产出确实存在投入效应不清晰、不对应的现象。无论是培养人还是发表科研成果，都无法精准地将结果归因于某种特定投入，甚至还可能会出现同为投入与产出或者互为投入与产出的状况。例如，对于高校的自筹经费，既可以看作是高校办学的投入，以显示高校办学的总体投入规模；也可以看作是高校办学的产出，以显示高校通过自身努力所获得的竞争力。所以，在不同的价值取向的导引下，不同指向的绩效评价可以有不同的活动要素划分。也正因为如此，马丁等认为，“高校绩效指标既可以是序数的，也可是基数的；既可是绝对性的，也可是相对性的；既可是固定的、机械的，也可是一些非正式的如同行或声誉排行等”^①。

指标的独立性、不重复性，的确是关乎评价效果的重要标尺。但是，对于

^① Martin C, Stephen M, Mauric K. 1988. The Use of Performance Indicators in Higher Education: A Critical Analysis of Developing Practice. London: Jessica Kingsley Publishers: 22.

高校办学本身的特点,包括活动的职能归属难以严格区分,活动的因果关系难以一一对应,活动的集中指向难以准确细化,等等,确实存在似乎与统计科学指标要求不完全相符的现实特征。不过,教育类型的评价真的需要在评价指标上做如此精准的科学化分吗?如同教育实验,它始终是一种非自然科学实验,似乎根本不能算作科学实验,但依然为教育研究所需要,依然为社会科学研究所认可,因为任何一种社会科学研究都很难在完全脱离现实环境的条件下实现其研究的科学性。所以,放置于高等教育领域的绩效评价,与进入教育场域的科学实验活动一样,都带有某种非自然科学的人文社会科学色彩,既要借助自然科学的技术和方法,又不完全达致其精细的要求,从而形成了自己领域的使用特点。这是一切人文社会科学研究中使用自然科学研究方法与技术存在的普遍难题。

在实际操作中,对于高校绩效评价指标的选取、剥离及其指标之间彼此独立的问题,可以采取一些做法来逐渐予以解决。一是对所有指标的内涵进行充分界定。根据研究需要,分别将指标从高校办学活动中抽离出来,形成类型、层级、范畴、意义、作用、基本指向等方面的相对独立性区分,避免在高校办学活动大格局下的重复。二是对指标的筛选,综合采用统计学和教育学方法。评价实践中必然会涉及某些具有教育价值、导向价值的指标,会被基于数据事实的统计学方法判定为重复或不独立而需要删去。所以,属于教育评价的高校绩效评价必须将两类方法综合使用,实现数学意义和教育学意义的评价统合,避免因教育活动的独特性所带来的统计学方法方面的误差。三是对于一些确实不易界定或异常模糊且难以操作的指标,可以寻找替代指标。在调查研究中,类似的情况会不时遇到,选择相近的替代指标也能在一定程度上确保研究的科学性。

世界原本就处在一片混沌之中,万物都有联系,绝对意义上的独立是不存在的。硬性实现指标的绝对剥离,实质上是违背了自然存在的规律。毕竟,无论是在自然、社会还是心理层面,所有的独立其实都是相对的,并没有事物完全处在独立的“绝缘真空”状态。后现代主义者甚至质疑:“人类建立的智力结构之外是否存在一个确定的自然世界……科学只是由科学家创建的一套规则游戏。”^①为此,在推进高校绩效评价研究的过程中,不必将高校这一教育活动场所的世界本源特点悬置,而戚戚然执念于剥离出充分科学的、绝对精准的绩效指标,这似乎既不可能也无必要。只要基于一定的价值取向、范畴与分类来进行评价,相对独立的绩效指标还是可以抽离出来的,并且也可以在理论研究中不断得以完善。

四、高校绩效的比较角度

当绩效指标设定之后,便可依据指标采集数据并对高校实施绩效评价。对于

^① 亨利·N. 波拉克. 2005. 不确定的科学与不确定的世界. 李萍萍译. 上海: 上海科技教育出版社: 8.