

高等教育评价中的“数字陷阱”

刘少雪

(上海交通大学 高等教育研究院, 上海 200240)

摘要: 高等教育是育人的事业,其核心任务很难用具体数字来表达。数字化的描述和评价方法客观上为人们提供了更直接的考察高等教育及其价值体现的视角,但工具理性自身带有漠视人的情感和精神价值的弊端,也会将高等教育核心使命零碎化、浅表化,对人们全面认识和评价高等教育带来一定的负面影响。本文通过三个具体案例,分析了数字化形式表述高等教育所可能带来的问题,进而认为,数字化指标可以用来解释高等教育质量,但其使用需要设置若干前提条件,包括数字指标的解释力、指标间的逻辑关系、数据来源的真实性等;缺乏合理性和真实性的数据,对高等教育质量评价则是失真的、无效的,甚至是有害的。

关键词: 数字;指标;高等教育质量

作者简介: 刘少雪(1967—),女,山东莱阳人,博士,上海交通大学高等教育研究院教授,主要从事高等教育研究。

中图分类号: G640 文献标识码: A 文章编号: 2095-7068(2016)01-0028-08 收稿日期: 2015-12-22

高等教育是育人的事业,关系到人类的长远进步、现实发展和文明承续,因此,无论何时何地,培养人、探索未知和传承人类文明都是高等教育的根本任务。与可衡量的其他物质存在不同,高等教育的核心任务因为与“人”、“文明”和“未知”有关,因此很难用衡量其他物质存在通常用的具体数字表达,案例陈述往往成为人们描述和评价高等教育的主要方式。但随着现代工具理性在社会各方面的运用和渗透,效率和数字亦越来越成为人们衡量和描述高等教育的重要方式。数字化的描述和评价方法客观上能够为人们提供更直接的考量高等教育及其价值体现的视角,但不可避免地,由工具理性自身带来的漠视人的情感和精神价值的弊端,也体现在其可能导致将高等教育核心使命的描述零碎化、浅表化、庸俗化,对人们准确认识和评价高等教育带来负面影响。本文试图从一些具体的案例出发,阐释仅仅运用数字方式分析高等教育质量所可能存在的不足。

一、案例呈现与分析

当前,国内用数字方式来分析和表达高等教育发展的成果很多,本研究选取目前较为普遍的就业与专业设置、学生评教与教学质量、量化指标考核与科研发展等三个方面的一些典型案例进行关注和分析。

(一)就业与专业设置

进入新世纪后,我国高等教育快速从精英化阶段步入大众化阶段。2014年的高等教育毛入学率已经达到37.5%,普通高校在校生规模达到2500万人以上^[1],这也意味着每年会有数以百万计的高等学校毕业生进入劳动力市场。因此,每到毕业季,关于大学生就业的话题就会牵动社会各方面的注意力。为此,若干机构以不同形式发布了与大学生就业相关的数据和排名,也都会有不同的高校或专业成为舆论“讨伐”的中枪者。那么,这些数据排名是否科学合理呢?现以麦可思

公司发布的《中国大学生就业报告》(2013年和2014年)为例分析说明。

《中国大学生就业报告》显示,2013、2014年

我国高校本科毕业生半年内的平均就业率分别为91.8%和92.6%;从专业角度看,两个年份中排名最高和最低的十个本科专业分别可见表1和表2。

表1 2013年高校毕业生就业率最高和最低的10个本科专业^[3]

就业率最高的10个本科专业			就业率最低的10个本科专业		
序号	专业名称	就业率/%	序号	专业名称	就业率/%
1	建筑学	98.3	1	物理学	85.3
2	安全工程	96.9	2	生物科学与工程	85.7
3	地质工程	96.7	3	化学	86.2
4	机械电子工程	96.6	4	法学	86.3
5	护理学	96.1	5	生物技术	86.8
6	给水排水工程	95.9	6	生物工程	87.1
7	地理信息系统	95.6	7	音乐表演	87.6
8	工程管理	95.3	8	动画	87.6
9	矿物加工工程	95.0	9	应用心理学	87.9
10	信息管理与信息系统	94.9	10	应用物理学	88.0
10	电子商务	94.9	—	—	—

注:本年度全国本科专业毕业半年内的平均就业率为91.8%。数据整理自《2014年中国大学生就业报告》,www.eol.cn/html/c/2014jylps/index.shtml pcvVC2RyuJbvWu9tL-Xm3IfOBKUWzhFq1w-e8TDKe_。

表2 2014年就业率最高和最低的10个本科专业

就业率最高的10个本科专业			就业率最低的10个本科专业		
序号	专业名称	就业率/%	序号	专业名称	就业率/%
1	护理学	97.0	1	应用物理学	84.8
2	建筑环境与设备工程	96.8	2	表演	86.6
3	医学影像学	95.9	2	音乐表演	86.6
4	测绘工程	95.6	4	绘画	87.0
4	建筑学	95.6	5	法学	87.1
6	医学检验	95.5	6	冶金工程	87.3
7	工程管理	95.3	7	生物医学工程	87.9
8	安全工程	95.1	8	轻化工程	88.3
9	中医学	94.4	9	体育教育	88.4
9	数字媒体与艺术	94.4	10	生物技术	88.5
9	车辆工程	94.4	—	—	—
9	矿物加工工程	94.4	—	—	—
9	电气工程及其自动化	94.4	—	—	—

注:本年度全国本科专业毕业半年内的平均就业率为92.6%。数据整理自《2015高考志愿填报:就业率较低的10大本科专业》,http://edu.qq.com/a/20150619/019168.htm和《2015中国大学生就业报告发布》,http://learning.sohu.com/20150614/n414990956.shtml。

从表中可以看出:

第一,2013、2014年我国高等学校本科毕业生半年内的平均就业率为92%左右。这个数字是否属于国家、社会和高校可以接受的水平?从

我国高等教育发展历史可知,自新中国成立后到20世纪90年代中期,我国高等学校毕业生一直实行计划分配制度,高校毕业生毕业后都能够获得安排工作岗位(个别不服从分配的除外)。实行

毕业生自主进入劳动力市场、双向选择就业制度后,没有人能够保证所有的高等学校毕业生都能在短期内找到满意的工作岗位——这是市场运行的基本规律。但由于新中国成立后上大学—毕业—工作之间的密切联系,毕业生家长和毕业生本人对毕业后的就业状态抱有较高期望;另一方面我国人口基数和高等学校学生规模大,即使毕业生中出现小部分不能立即就业的群体时,其绝对数量也相当可观,因此易于激发社会多方面的反应;第三则是由于近年来的经济转型发展,劳动力市场的人才需求与高校毕业生的专业技能难以一一对应,使得社会对高校的人才培养工作颇有微词。可以说,在诸多因素的影响下,高校毕业生就业成为近年来的敏感话题。值得注意的是,即使在高等教育发达国家,高等学校毕业生的就业率也达不到100%,这其中既有劳动力市场的因素,也有毕业生自身的原因。例如,美国国家科学基金会2006年的毕业生调查显示,学士和硕士毕业生参加工作的比例为85%;劳工部委托人口调查局在2007年的调查中发现,在24岁及以下青年中,具有学士以上学位的就业率最高,男性为88.1%,女性为90%,具有副学士学位和接受一定高等教育的青年的就业率居次,男性为85.4%,女性为75.1%。^[2]英国高等教育统计局发布的《大学毕业生追踪调查》显示,2011年英国高等教育毕业生全职工作的比例也只有73.4%。^[3]如果仅从就业率的数据上看,我国高校毕业生的就业率水平明显高出美、英两国。那么,这是否意味着我国高校毕业生的就业率已经很高?我们是否可以对此表示满意?或者我们对此还有其他要求和期望?

第二,表1和表2显示的2013和2014年高校本科毕业生半年内的就业率最高和最低的差距分别是13和12.3个百分点。在平均就业率达到92%左右的水平上,最高和最低之间10%的差距,其实质性内涵应如何解释?特别是在2013和2014年就业率最高和最低的专业排名榜单中,就业率最高和最低的专业重合率都不高(50%),两年都列入最高前10位的专业有建筑学、安全工程、护理学、工程管理和矿物加工工程等5个专业;都列入最低名单的专业有法学、生物技术、音乐表演和应用物理学等4个专业。另外,教育部于2014年也公布了2012、2013年就业率较低的15个本科专业^[4],这15个专业与前面的两年的最低

专业名单重合率也不高,在三个名单中都出现的专业只有1个(音乐表演),出现两次的专业有3个(生物科学、动画和表演)。依此来分析,我国高等学校设置的专业中,是否真正存在就业率低的专业?教育部对高等学校专业设置有要求,即需要“有稳定的社会人才需求”^[5],在巨大的毕业生就业压力下,要对毕业生就业率低的专业设置预警制度。如果用明确数字表达出来的就业率排名结果不一致,甚至差别明显,这样的统计或排名结果是否可信?高校是否可据此调整和设置专业?这里同样有2011年英国高等教育统计局发布的《高校毕业生就业去向调查》作对比,该报告显示,80.5%的英国高校学生毕业生从事的是与其专业相关的工作,66.3%的毕业生认为大学教育体现了其自身的价值。^[6]我国高校不同专业毕业生的就业率数字可否用来判断学生的就业去向,或者是否能够体现高等学校的教育价值?显然目前的数字很难达到这样的目的。

(二) A大学的本科课程教学学生评教情况

教学活动是“良心活”,教师的能力、技巧和投入对教学活动的成效影响很大。教师是影响教学活动成效的重要因素,但不是唯一因素,学生的投入和准备、教学内容与难度、培养体系的设计、教学制度安排等都会影响教学活动效果,因此,考察教学活动效果本身就是一件很复杂的事情。

近年来高等教育领域中的量化指标评估趋势,也影响到了教学领域,不少高校尝试对教师教学活动进行数字化评估,学生评教是其中的典型做法。学生评教起源于20世纪的美国。“我国高校学生参与教师教学评价,始于20世纪80年代中期。这段时间部分高校开始引入学生参与评价教学的评价方式,后逐渐为在多数高校得到认可和应用。”^{[6]13}本案例选取的A高校于本世纪初开始将学生评教作为监督和考核教师教学活动成效的一种方式。学生评教自被引入后,不断遭受教师的质疑,校方也曾多次改进评教体系和内容,但在教师晋升或奖励时,学生的评教结果还是会发挥一定作用。A大学2014—2015学年第一学期开设的3034门次本科生课程的学生评教结果呈现如下。除了选课人数少于10人,或参评学生少于50%的课程外,2823门次课程得到有效评

A大学本科课程教学学生评教情况的所有数据和问卷内容均来源于A大学教务处组织的2014—2015学年度的本科课程教学学生评教结果(供内部使用)。

价。在满分为100分的评价规则下,最高评教课程得分为98.2分,最低为38.4分;所有课程得分的平均数为85.2分,标准差为6.3;有2084门课程得分超过平均分(占73.8%);得分最集中的区间在85~95分之间,超过2/3的课程集中在这个区间(具体情况见表3)。

表3 A大学本科课程的学生评教结果分布表

分数段	课程门次	比例/%
95	163	5.8
90~95	1 003	35.5
85~90	938	33.2
80~85	438	15.5
75~80	173	6.1
70~75	52	1.8
65~70	35	1.2
60~65	9	0.3
<60	12	0.4
总计	2 823	100.0

从表3的数据中,会有这样的疑问产生:

第一,全校的本科生课程评教分数平均分超过85分,是否意味着A大学的本科生课程教学已经达到令人满意的程度?实际上,关于A大学本科课程教学的满意度曾经有一个参考数据(研究经过信度和效度检验合格,但因为非学术原因不能公开发表),即通过对该校某个本科2+2项目(在A校和其国外合作大学分别学习过2年)的学生调查结果表明,学生对A大学专业培养的满意度明显不如对国外合作大学的满意度高。调查学生对A大学的课程体系、教学内容、教材与参考书、教学方法、师资配备、教学效果、教学组织与管理等均有40%左右的“不满意”或“很不满意”;相反,对国外合作大学的上述方面仅在“教学方法”和“师资配备”两个方面分别有1.5%的“不满意”,没有“很不满意”的选项。两种来源不同的数据有如此明显的差距,作为学校的管理者和教师,更应该相信哪个数据?

第二,评教课程的最高分与最低分之间的差距达60分,是哪些因素导致了如此大的差距?是授课教师,还是授课内容,抑或是授课方式导致学生对某门课程的评价如此低?经过仔细查看,发现最高得分(98.2分)的3门课程和最低得分(38.4分)的1门课程都是外语类课程,除其中

1门最高得分课程是英语课程外,其余3门都是法语课程。在全部有效评价的课程中,法语类课程共有37门,其中“法语写作”课程开设3门次,且评教得分均处于同类课程中的最低得分(分别为71、65.2和38.4分),其余法语类课程的得分均在85分以上。这里的疑问是,在法语类课程中,不同课程得分差别巨大,是否与课程内容及难度相关?但同为“法语写作”课程,不同班级的评教分数差别很大,那么根源在任课教师还是选课学生?另外,从目前的数据中可以看出,在3门次的“法语写作”课程中,最高得分的课程在全部课程排名中已经处在最后2%的水平上,更不用说最低得分的课程就是全部评教课程的最低分。那么从管理部门的角度,如何看待这门课程?从任课教师的角度,又如何评价这门课程?从课程设置本身,这门课程是否有存在的价值和必要?如何在师生间沟通处理好设与不设之间的平衡?……由此可以引出的问题很多,但仅从上面的数据中是看不出来的。

第三,评教分数呈现严重偏正态分布(见图1),这个结果是否意味着这次评教本身不够严谨有效?按照教育统计学规律,正常的评教成绩分布直方图,如基本呈正态分布状,会比较有效地反映学生对学校开设课程的评价情况,即评教效果是好的,基本上可以信任;现在的评教结果呈严重偏态,峰值处与得分最集中的课程数不能重合,说明评教的效果不理想。通过考察学生评教的问卷内容,其中“理论教学”类课程的评教内容主要关注三方面内容:课程内容、课程及任课教师等。由此提出的问题是,以这样三个问题来评价教师的教学工作,依据是什么?这三个指标之间的关联性如何?它们是否可以反映教师的工作投入和状态?问卷内容是否进行过可靠性检验?与学校的本科课程教学目的、人才培养目标是是否

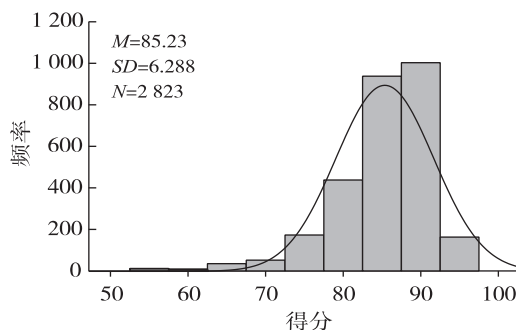


图1 A大学本科课程评教分数聚集图

匹配?评教指标体系是否能够规避评教过程中的干扰性因素?如果所有这些问题都不能在评教之前就得到解答,评教结果自身及其应用都是值得商榷的。

综合上面的分析,由于A大学没有提供学生评教的信效度检验,影响了评教结果的真实可靠性,难以反映A大学本科生对学校开设课程的态度和真实评价。因此,从严格意义上说,这样的评价是没有意义的。但对于管理者和教师来说,这些有先天缺陷的分数又是有意义的——因为它们可能会被运用到学校对学院和教师的考核评价

中,甚至可能在教师的聘任与职务晋升等关键性事项中发挥作用。这样的数字确实会在高等教育管理中产生“陷阱”作用。

(三)C大学在不同年代对教师的考核要求及科研成果比较

近年来,对教师进行以量化指标为核心的考核,成为国内高校的普遍做法。具体的量化指标在不同高校之间会有差异,但宗旨都是把高校教师的工作分成若干不同的方面并指标化。表4反映了C大学自20世纪80年代以来对教师考核变化要求。

表4 C大学在三个不同年份对教师的考核要求

年份	教学要求	科研要求
1984	7门次	560小时,完成年度研究计划所必需的调查研究、方案论证、各种设计、研究试验、上机计算、撰写论文、科技文件材料整理归档、成果鉴定、成果推广、技术后勤、课题组长组织工作等基本小时数
1992	7.5门次	300小时,(副教授以上)每两三年在国际或国内一级学术刊物或国际性、全国性的学术会议上发表(宣读)一篇与自己职称相称的论文;或取得一项科研成果(包括设计新实验,研制实验装置,编写新教材、教学参考书、专著及教学研究方面所取得的成果)
2013	4门次	聘期内(三年):学术论文不少于4篇;主持省部级基金项目1项,或独立出版个人专著1部,或以第一作者获得省部级及以上教学或科研成果奖1项,或以第一作者(SSCI通讯作者)发表B类及以上论文1篇(不与前面论文重复),或发表被SSCI引用频次较高的学术论文(包括国际会议论文),或海外出版社出版外文著作

可以看出,自20世纪80年代中期以来,C大学对教师的年度任务考核内容已发生显著变化。20世纪80年代,学校对教师的教学工作量要求很明确,7门次的标准班级授课任务是教师工作量的主干;科研工作有时间要求,但工作内容比较灵活,除了常规的调查研究、撰写论文外,“科技文件材料整理归档、成果鉴定、成果推广、技术后勤、课题组长组织工作等”也可以纳入到教师的科研工作量计算。由此可以看出,这一时期学校对教师的科研工作要求比较宽松,基本尊重教师的自主选择。相比80年代,90年代的教学工作量要求依然繁重,平均增加0.5门次的授课任务;同时,科研工作要求也比80年代聚焦,最明显的是有了发表论文或其他成果形式的明确要求,但成果体现依然允许有多种选择。到21世纪,教师的教学任务量明显降低,但科研工作的要求更加集中明确,包括论文方面有了数量和发表刊物等级的要求,项目或奖励有来源等级,著作也有出版社等方面的明确要求或限制。由之可以清楚看出,C大学对教师的考核要求越来越明确、集中,

并向越来越数字化的方向转变。

学校的考核要求,一定会转化到教师的实际工作中。在不同时期的要求中,教学工作量都是“硬任务”,且教学工作量的安排,并不完全取决于教师自身,因此,一般在学期/年初,教师的教学工作会被首先落实。科研工作的表现、类型、成果等工作,与教师个体的能动水平有明显关系。在考核要求相对模糊阶段,教师的科研工作可能会表现为多种方式,“成果效率”可能不高;但在精准要求阶段,完成考核任务成为教师第一要务,效率可能会有明显提高。表5是对C大学在这三个时期教师科研产出的简单对比。

表5 C大学在三个年份的科研产出比较表

年份	师均论文发表数/篇	师均承担项目数/项	师均获奖成果数/项
1985	0.93	0.54	0.03
1993	0.91	0.84	0.02
2014	3.63	1.88	0.02

若无特别说明,表4、5、6中涉及C大学的数据,均整理自其公开出版物和部分内部资料。

从表5中看出,C大学在这三个阶段以论文和项目为指标的科研产出有明显进步。其中,20世纪90年代与80年代相比,80年代师均发表论文不到1篇,承担的科研项目数也很少(0.5项);到90年代初,师均论文数略有降低(0.02篇),但承担项目数有明显进步(1.56倍),再考虑到这一时期教师的教学工作量增加,可以说这一时期的教师科研产出还是有明显进步的。但更大的进步体现在2014年的数据上,其中,论文和项目数比前面两个年份都是呈倍数增长,如对1985年论文和项目数都超过2.5倍以上,对1993年则是论文数超过近3倍、项目数超过1.2倍。这充分证明了数字化的考核要求对教师的科研产出有相当明显的正面影响。

从表5的数据中,还可以发现反映科研质量的另外一个指标——获奖成果——在三个不同年份的变化。在三个年份中,1985年的师均科研成果获奖数最高,其后的两个年份没有明显变化。进一步查阅相关材料,发现1985—1990年间是C大学迄今为止获得科研成果奖励项目数量最多、等级最高的辉煌时期。相对来说,2014年C大学教师发表了更多学术论文,承担了更多项目,但教师的获奖量不仅没有提高,反而有所下降。由此产生的疑问是:论文发表和承担项目的数量和级别是否能够反映一所大学的科研能力和水平?如果能,那么为什么以反映质量为核心的师均成果奖励数没有随着教师科研产出数量的增加而提高?如果不能,为什么当前我国的大学普遍采用以论文和项目数来表示大学的科研能力?发表和承担项目的数量及级别在用于评价大学和教师的科研能力时应该发挥什么作用?比较C大学三个时期的考核要求,1984年对教师的科研考核指标最软、最难以考量和把握,教师发表的论文和承担的项目数也最少,但科研成果的师均获奖率最高,甚至后面还有一段时间成果获奖的辉煌期。这对C大学来说意味着什么?2014年时学校对教师的科研工作任务要求最明确,教师和管理者都很容易把握这些指标,其后的教师科研表现也基本实现了校方的预期,但成果的质量并没有同步得到社会和同行的认可。那么这是不是大学期望的结果?另外,大学教师除了可以罗列出教学时数和发表论文、承担项目外,大学内的很多其他诸如教学改革、调查研究、与外界沟通交流,特别是需要直指人心、打动学生的思想工作等,

都需要教师较多的精力、时间和精神投入。如果大学把教师的工作都以数字指标来标示的话,那么这些难以用数字精确衡量的精神和文化层面的工作是否还需要教师承担?如果大学和教师不能继续承担此类工作,大学与研究机构之间是否还有区别?大学及大学校园存在的价值意义,培养学生、学术研究及推动科学进步在当下的意义史无前例的重大,但若因此而忽略了其在传承和发扬人类文明、承担改良社会风气之先的重任,大学的意义会黯淡许多。但传承人类文明、改良社会风气的责任,过程比结果重要,且很难用结果替代过程,因此非常有必要让大学和教师能够感受到适度的宽松和信任,既能由此体现出大学对教师和学生、大学教师职业的信任和尊重,也是大学能够获得更多教育成果及存在价值的基本前提。

二、讨论

高等教育作为教育的一部分,树立百年育人、长期教育的观念和情怀至关重要。数字性指标能够在某个或某些方面帮助人们直观地考察高等教育质量的部分表现,但如果过分依赖数字性指标评价高等教育质量,则还有很多重要问题需要讨论清楚。

第一,评价中所选取的数字指标,是否有科学合理的依据?简单地说,不同的指标只能反映高等教育质量的某一或某几个方面,因此,不同的高等教育质量评价,只能根据目的需要选取不同的指标及指标体系。因此,服务于不同目的的指标选取,即使合理,也只能保证其合理性在评价某一或某几个方面时的有限范围。从这个意义来说,任何时候选取科学合理的指标都是合理使用数字性指标的第一步。但如何选择科学合理的指标?选择时是否有合理的规则和流程?如何评价考察选取的指标是否合理科学?这些则是指标使用前需要解决的关键问题。以上面提到的就业率为例。这个就业率所表述的是学生在毕业后半年内是否签订了工作协议,因此其核心内涵为是否有工作协议,但工作协议不反映毕业生正在或将要从事工作的性质、类型,以及毕业生对该工作的满意度等,也反映不出毕业前学生所学专业对其从事的工作是否有价值、价值有多大。用没有反映学生专业价值的工作协议来评价不同学科专业的就业率,没有逻辑意义,因此也难有衡量

价值。用文章数和项目数来考核教师的合理性同样值得怀疑。在近年来大学普遍采用量化考核指标后,我国大学的学术发表能力和国际影响力不断提高,但论文泡沫、论文枪手、学术造假等与考核指标密切相关的质疑声不断;大学教师作为职业化“写手”(论文和项目申请书)的素质和能力也有明显提高,但大学中另一核心群体——学生——被教师关注的越来越少,质疑高等教育质量的外部声音越来越多,这也是就业率、学生评教等数字化指标能够有市场需求的原因所在。借用克拉克·科尔所说的“高等教育越来越少只是为成人就业做准备,而越来越多成为一个人整个人生的一部分”^{[7]264}来评价毕业生就业率与高校专业设置的问题,高等学校不应该也不需要把就业率与专业设置联系在一起。

第二,数字本身是否真实可靠?无论是就业率统计,还是学生评教,都需要大范围收集有效数据。但目前的数据发布者都没有对数据来源的可靠性进行说明。在我国高等教育领域真实数据获取有很多现实障碍的情况下,采用不同方式获取大规模数据是不可避免的,但越是在这种情况下,越需要对所获得的数据进行可靠性分析。特别当这些数据指标有可能被应用到评价个人或群体时(如就业率可能会影响到高校的专业设置、学生和家长专业选择,学生评教结果可能会影响到教师考核及晋升等),真实可靠是影响数据意义的核心因素。高等教育质量是一个内涵极其丰富的概念,对其能够产生影响的因素也很多,人们对好的高等教育质量有若干共识,不理想的高等教育质量可能是受多种不同因素共同影响的结果,仅靠某些因素分析就对高等教育质量进行评价和判断的做法是危险的。富有大学管理经验的克拉克·科尔还说过:“即使意图再好,成绩难以界定,也不易估价。这是为什么在美国高等教育机构教师的晋升中至少有半数(也许多至四分之

三)的教师以资历取代成绩的一个理由。鉴于极少合适的绩效,又缺乏明显的高绩效的证据,至少全体教师的一半(多半在研究型大学和优秀的文理学院以外)只根据资历晋升。”^{[7]173}大学教师是一群以传承文化、探索未知为基本使命的知识精英,他们之间的共性要远远少于他们之间的个性差别,而且正是因为他们之间是不同的,他们才有可能共处一个校园。承认了他们之间的差别,就难以对不同的人、不同的工作进行精确评价和比较,以模糊代替精确是科尔管理大学的明智选择。

第三,精确的数字化模式是否适合高等教育质量管理?精确管理最先出现在企业管理中,它对提高企业的生产效率、降低成本等方面富有成效。但与企业相比,高等学校承担了育人、探索未知、服务社会、传承文明等多项使命,因此决定了高等学校的发展是多目标性的,其质量维度也是多向的;更由于高等教育是以人(包括教师和学生)、文明、科学等为核心要素,就决定了高等教育难以套用精确管理的模式去进行质量管理,简单地说,高等教育质量管理需要更加复杂的系统管理模式。但“随着系统复杂性的增加,我们做出关于系统行为的精确而有意义的陈述的能力将降低,越过一定阈值,精确性和有意义(或适用)几乎成为相互排斥的特性”^{[8]30},以简单的数字化指标来判断和指挥高等教育的做法从管理学的意义上说是不适宜的。

由于高等教育目标和使命的多元化,决定了高等教育质量的测量和评价是个复杂工程,用数字化指标解释高等教育质量,需要若干的前提条件,包括指标对高等教育质量的解释力、指标间的逻辑关系、数据来源的真实性等。缺乏合理性和真实性的数据,对高等教育质量评价则是失真的、无效的,甚至是有害的。因此,防止高等教育质量评价中的“数字陷阱”,对于高等教育研究者来说是非常急迫而且相当艰巨的任务。

参考文献

- [1] 中华人民共和国教育部. 2014年全国教育事业统计公报[EB/OL].(2015-07-30)[2015-10-10]. http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/201507/t20150730_196698.html.
- [2] 王晓阳,方军. 美国联邦政府对高校毕业生就业市场的培育与引导[J]. 世界教育信息, 2009, (1).
- [3] Chris Havergal. Graduate unemployment back to pre-recession levels[EB/OL].(2015-08-27)[2015-10-10]. <https://www.timeshighereducation.com/news/graduate-unemployment-back-pre-recession-levels>.
- [4] 晋浩天. 教育部公布近两年就业率较低的本科专业名单[EB/OL].(2014-10-16)[2015-10-10]. <http://www.>

jyb.cn/high/gdjyxw/201410/t20141016_601306.html.

[5]教育部.普通高等学校本科专业设置管理规定[EB/OL].(2012-09-14][2015-10-10]. http://www.moe.gov.cn/s78/A08/gjs_left/moe_1034/s3881/201305/t20130523_152287.html.

[6]王丽荣.中美高校学生评教指标体系的比较研究[D].济南:山东大学,2008.

[7]克拉克·科尔.高等教育不能回避历史——21世纪的问题[M].王承绪,译.杭州:浙江教育出版社,2001.

[8]马驥良.札德的科学思想和方法[M].西安:陕西师范大学出版社,1992.

[责任编辑:罗雯瑶]

Misleading Quantitative Evaluation in Higher Education

LIU Shao-xue

(Graduate School of Higher Education, Shanghai Jiaotong University, Shanghai 200240, China)

Abstract : The core missions of higher education as an undertaking of talent development cannot be simply evaluated in quantitative terms, though quantitative evaluation and description may sometimes offer straightforward insight into higher education and into its expression of values. However, due to the failure of instrumental reason to fully consider human emotional and spiritual values, the quantitative approach to higher education may lead to segmented and superficial misunderstandings of the core missions of higher education, and therefore inaccurate evaluation of higher education. Through an analysis of three cases, the paper discusses the possible problems caused by the quantitative description of higher education. It argues that though quantitative indicators may help interpret education quality, the use of these indicators needs a number of pre-conditions, including the insightful interpretation of the indicators, the relationship among different indicators, the authenticity of the data sources, etc. The illegitimate and invalid data may lead to misleading assessment of higher education.

Key words : data; indicator; quality of higher education