

Doi: 10.11835/j.issn.1008-5831.pj.2022.02.004

欢迎按以下格式引用:时艳芳.高等教育评价结果:困境、反思与改进[J].重庆大学学报(社会科学版),2022(2):108-120.

Doi: 10.11835/j.issn.1008-5831.pj.2022.012.004.

Citation Format: SHI Yanfang. Results of higher education evaluation: Dilemma, reflection and direction[J]. Journal of Chongqing University (Social Science Edition), 2022(2): 108-120. Doi: 10.11835/j.issn.1008-5831.pj.2022.02.004.



高等教育评价结果： 困境、反思与改进

时艳芳^{1,2}

(1. 四川师范大学 教育科学学院, 四川 成都 620021; 2. 张家口学院 教育学院, 河北 张家口 075000)

摘要: 教育评价结果是评价活动和实践改进的桥梁, 只有完善评价结果应用, 才能形成教育质量闭环的提升过程, 凸显评价的发展价值。但在具体实施过程中评价结果要么不被重视, 要么呈现为竞争性、等级性等不利于评价持久发展的现实样态, 尚未充分发挥以评估结果落实“强硬度”“长牙齿”、推进教育治理效能提升的目标。当前学界在谈及评价时通常会涉及评价结果相关问题, 但缺少专门研究。实践和理论的双重不足为开展评价结果研究提供了空间。2020年10月, 中国第一个关于教育评价系统性改革的文件——《深化新时代教育评价改革总体方案》(以下简称《方案》)提出扭转不科学的教育评价导向, 改进结果评价, 在对当前流行的结果评价的负面影响认识基础上进行反思, 改革阻碍高质量教育发展的“五唯”倾向, 促进“双一流”大学建设, 进而以评价赋能育人。该文件的出台既体现了国家对教育评价问题的重视, 也突显了解决评价中所出现问题的紧迫性。基于此, 以“改进结果评价, 强化结果使用”为价值追求, 从评价结果的呈现形式、反馈方式及相关主体的秉持态度、解读路径和运用取向等方面全方位审视高等教育评价结果。发现当前教育结果评价普遍存在“五轻五重”现象, 即呈现形式上重量化轻质性、反馈方式上重告知轻协商、对结果的态度重接受轻质疑、对结果的解读重实体轻关系、对结果的运用重功利轻价值。发现问题是为了解决问题, 而解决问题的前提条件是能追溯造成问题的根源, 只有如此, 才能做到有的放矢。遵循“理念—制度—操作”的思维逻辑, 深入系统地反思问题产生的根源, 发现其既有教育评价理念的偏差, 也有评价制度的不科学, 还有评价结果运用的复杂性。完善结果评价运用, 落实《方案》的评价理念, 实现“破五唯”的价值追求, 是研究的归属和落脚点。鉴于评价落地的关键在于人, 而人的任何行动都离不开制度的支持, 所以, 选取“人—制度”的分析框架, 提出详实可行的意见, 为发挥评价的发展性功能, 建立健全立德树人机制提供参考, 推进十九届六中全会提到的推动人的全面发展、建设社会主义现代化强国。具体包括提高行政人员的评价能力、增强实施人员的专业性、调动落实人员的积极性; 健全评价过程的监督机制、优化评价结果的使用机制、建立评价改进的追踪

基金项目: 国家社会科学基金“十三五”规划教育学一般课题“‘双一流’大学建设背景下地方本科大学国际化问题研究”(BIA170240)

作者简介: 时艳芳, 四川师范大学教育科学学院博士研究生, 张家口学院讲师, Email: yanfang023@126.com。

机制等建议。

关键词:高等教育评价;评价结果;破“五唯”;功利性;行政化;“双一流”建设;一流高校;一流学科;学科建设

中图分类号:G40-058

文献标志码:A

文章编号:1008-5831(2022)02-0108-13

一、问题提出与学术回应

高等教育评价是20世纪80年代在政府简政放权赋予高校更多自主权的背景下,作为政府对高等教育问责的一种手段而出现的。经过持续探索,高等教育评价的相关政策法规不断完善、组织机构不断健全、评估项目不断丰富、理论探讨不断深入,但在具体实施过程中出现了偏离以评促改、以评促建初衷的非理性现象。“评价最重要的意图不是为了证明,而是为了改进”^[1]。“教育评价的本质应当是推动教育活动的改进,评价能够完成的最大贡献是使得被评价者能更清晰地识别问题并找到自己的努力方向,这应该是教育评价最大的价值所在”^[2]。可见,评价不仅是对过往成绩的评判,更应为未来提供改进和努力的方向,所以,“突出以评估结果运用落实‘强硬度’‘长牙齿’,综合运用评估、督导、通报、挂钩和问责等举措,切实推进教育管理和教育治理效能提升”^[3]。无论是“双一流”大学遴选还是学生学习测评,任何评价都涉及结果应用问题,只是应用的主体、目的、方式、程度不同而已。一个完整的评估过程必然包括对评估结果的质量进行检验、对评估结果的意义做出解释、对评估结果进行反馈和有效利用等。

教育评价是一项费时费力的长期活动,本该守护教育质量的评价,在实践中反倒成了“评价性负担”^[4],评价对象疲于应付、怨声载道、消磨了热情、淡化了情感、降低了认同。为了评价而评价的异化现象,扭曲了评价的本真意蕴,消解了评价的实践指导价值。“评价数据常常处于静止而非动态、孤立而非关联的状态,整个评价链条出现一种‘虎头蛇尾’的表象”^[5],而且以“竞争性、选拔性、等级性的结果评价为主,集中表现为追求分数、升学率等”^[6],这不利于良性教育评价生态系统的形成,也不利于高质量教育的均衡发展。评估结束意味着万事大吉的线性化流程,严重限制了评价效用的发挥,即使“基于研究视角的学校评价指标、评价标准、结果分析等方面的探讨比较深入,但对教育管理者和实践工作者而言,如何应用学校评价结果更为重要”^[7]。

当前学界在谈及评价时一般会涉及评价结果相关问题,但缺少对评价结果的专门研究。“教育评估结果是指对评估对象所具有或所达到的预定目标价值程度做出的判断”^[8]。“如果不强调结果,会导致人才培养的弱化,过于重视结果评价所带来的功利性评价思维会压抑和异化师生的发展”^[9]。“评估人员、评估理论与方法技术等种种因素导致评估结果存在失真现象”^[8]。“任何以定量形式表现的教育评估结果都是相对的。(学科)评估结果总体上可信度比较高……问题出在教育评估的结果跟各种利益绑定得太厉害了。由于对评估结果不仅影响学校声誉而且有可能与资源分配挂钩的预测和恐慌,‘应付’好学科评估就成了高校中学科建设的头等大事。高校为了能评出好名次,绞尽脑汁、想尽办法,不惜对学科进行关停并转、整合取舍,甚至不择手段、削足适履,无所不用其极”^[10]。“评估结果有的给出明确结论,包括定性等级式结论和分数式等级制,有的只是一种事实、状态说明或表征。结果使用形态包括正式使用和非正式使用、实质使用和形式使用、直接使用和间接使用、平行使用与交叉

使用。使用中存在过度使用和不当使用问题”^[11]。王薇较为系统地研究了学校评价结果的解释,包括主体、类型、方法、标准、模型建构等。^[12-15]

2020年中共中央国务院印发的《深化新时代教育评价改革总体方案》,是新中国第一个关于教育评价系统性改革的文件,是“在我国教育评价政策体系中所处层级最高、系统性最强、调整范围最广”^[16]的文件。深化评价改革意在“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”,手段是“改进结果评价,强化过程评价,探索增值评价,健全综合评价”^[17]。其中,结果评价以育人为导向,以立德树人为根本标准,过程评价、增值评价和综合评价服务于育人目标的实现。改进结果评价是“在认识当前流行的结果评价负面影响的基础上进行反思”^[9],从而改革阻碍人才培养的“五唯”倾向,以评价赋能育人,促进“双一流”建设。“改进”一词说明结果评价有合理性的一面,切忌全盘否定,走向另一个极端。“‘五唯’的本质和要害是‘唯’,就是将其绝对化、片面化、形式化和一刀切。我们在反对一种错误倾向的时候,要防止可能掩盖的另一种错误倾向”^[10]。2022年,三部委《关于公布第二轮“双一流”建设高校及建设学科名单的通知》(以下简称《通知》)中,有7所大学首次入选“双一流”名单,29所大学的学科数量有所增加,更有15所大学被公开警示及撤销学科^[18],这种以绩效为杠杆的动态调整机制,既是评价改革的初步成效,也是影响面最大的评价结果。

总之,评价结果是评价活动得以发挥作用的关键环节,只有克服短视行为、功利化倾向,破除高校、教师和人才等评价中的“五唯”,才能发挥评价结果的导向、鉴定、诊断、调控和改进作用。但当前教育评价结果存在诸多现实困境,学界的研究又稍显不足,这一矛盾提供了继续探索的空间。以“改进结果评价,强化结果使用”为价值追求,从评价结果的呈现形式、反馈方式及相关主体所秉持态度、解读路径和运用取向等方面全面审视当前教育评价结果存在的问题。遵循“理念—制度—操作”的思维逻辑,系统深入地反思问题产生的根源。选取“人—制度”为分析框架,提出有针对性的优化策略。为落实《方案》的基本精神,发挥评价的发展性功能,建立健全立德树人机制提供参考。

二、教育评价结果的现实困境

全方位谛视当前高等教育评价结果发现,普遍存在“五轻五重”现象,这既是阻碍评价发展的原因,也是改进评价结果的动因。

(一)评价结果的形式:重量化轻质性

无论是宏观的“双一流”评选、本科教育教学评估,还是微观的教师评价、学生评价等,评价结果多采取或分数或等级或在分数基础上划分等级的量化评价形式,具体表现为评价对象在每一评价项目上的评定结果。如第四轮学科评估结果按照“精准计算、分档呈现”的原则,根据“学科整体水平得分”的位次百分位,将前70%的学科分为9档公布,从C-到A+^[19]。这些客观量化的数字和等级,是教育评价科学化探索,直观清晰、操作简单、公开透明,既能回应各种质量问责,也为选拔和奖惩提供依据,还可以发挥一定的诊断导向功能,有其存在的合理性。

但大学教育是以知识为原材料,指向生命的活动,“大学评估中可量化的事项极其有限,而且这些可测量的未必是大学至关重要的部分”^[20]。如果根据期刊级别设定论文分数,就难以判断知识的内在价值和社会贡献,如果将充满活力的育人过程变成干巴巴的数字,就丧失了教育活动的价值判断和人

文关怀。量化的泛滥使得“评价从一种需要专业知识与技能的实践行为转变为一种自动化的机械操作”^[20],降低了评价活动的专业性和评价结果的公信力。唯量化趋势带来的“数字陷阱”^[21],容易催生盲目攀比、弄虚作假、片面狭隘、短视功利等背离教育本真的行为。因为“离开了边界约束单纯地追求数量化的形式公平存在着巨大的风险”^[4]。忽视对数据含义的详细描述和深入挖掘,评价结果就容易蜕化为大量数据的堆砌,既缺乏深刻的思想提炼,也难以发挥回应问题、制定政策和促进改善的作用。现实中出现了数量繁荣与质量干瘪的悖论,层出不穷、日渐虚高的排名、分数终将失去灵魂的卓越。

《方案》提出“改进学科评估,强化人才培养中心地位,淡化论文收录数、引用率、奖项数等数量指标”,《“双一流”建设成效评价办法(试行)》(以下简称《办法》)提出“综合客观数据和主观评议,避免简单以条件、数量、排名变化作为评价指标”^[22],既说明了这一问题存在的普遍性,也是试图破除唯量化评价范式的努力。“任何教育评价最终都要归结到定量或定性评价…….破‘五唯’与二者方法的选择密切相关,是绕不过去的坎,改进完善定量定性评价成为当前评价改革的重要内容”^[23]。

(二) 评价结果的反馈:重告知轻协商

反馈本质上是信息或者数据的传递,“评价结果反馈是以某种方式向被评对象、有关部门或人员通报评估结果的一种形式”^[24]。只有建立起“评价—反馈—改进”的封闭循环系统,及时、全面、准确地将结果反馈给评价对象尤其是实践主体,他们才能明晰自身的优点和问题,进而明确改进方向。但现实中存在缺少反馈机制、无反馈、重告知轻协商等现象,违背了评价的内在规律与发展逻辑,弱化了评价指导力,限制了评价功能的发挥。

“当评价组织者对把握的价值体系、评价体系体现的评价初衷,没能负责任地在公布结果时说明……形成价值信息不对称,会形成不良的评价生态”^[23]。有的评价主体运用公告、通知等形式将评价结果告知评价对象,这凸显了评价主体的权威,但缺少沟通协商机制,使得即使评价结果有待商榷,也被作为最终结论。如针对“双一流”建设高校的评选结果,网友调侃到“华南农学哪家强,华南理工大”,这既降低了对评价结果的认同度和接受度,也影响了评价活动的公信力,甚至造成评价者与被评者之间的对立。有的评价主体反馈意识淡薄,仅仅将评价结果作为工作总结的支撑材料呈现在质量报告中,代表该项工作的完成,而真正的行为主体,往往不知道评价结果或者仅被告知一个评价等级。如教师教学评价活动,尽管拥有领导听课、同行评价、学生评教等多种形式,教学督导的教学质量分析也年年进行,但仍然存在大量教师不清楚教学评价具体内容、自身的不足与优势有哪些等情况,这种教师缺位、失语的教学评价显得尤为荒唐。

(三) 对评价结果的态度:重接受轻质疑

评价是一种价值判断活动,评价主体的价值取向、利益驱动、专业能力,评价工具的信效度、专业性、运用条件,评价对象的本质属性、具体情境、发展历史,评价所采取的形式、所依赖的数据信息等都会对评价结果产生影响。所以,对任何评价结果都应该不断追问谁评价的、如何评价的、依据什么标准评价的以及结果在多大程度上代表了评价对象等问题,而不是盲目的全盘接受。

教育活动的复杂性、教育效果的滞后性决定了评价的限度,“有些活动可测量也值得测量,但能被测量的并非总值得测量,值得测量的也并非总可以测量”^[25]。所以,只有当教育评价是对教育的评价,才能保证评价的效度。目前很多评价,“在很大程度上其实都是‘指标’的需要,未必是出于真正学术

的需要”^[10],如学科评估无法测量学科本身的构成要素、发展脉络以及学科间的关系等。评价对象为了维护良好形象、获得理想结果,“一些高校有意选取有利于本校的排名,奖项的选择也是避轻就重……部分高校为了追求排名制造数据,严重违背办学规范和学术道德”^[23]。造成“最终的评价结果不再是一份真实的体检表,而是成为被评价者的美颜照和艺术照”^[4]。如首轮26所985高校“双一流”建设成果自评报告被吐槽,同一所大学在不同的大学排行榜名次相差甚至高达百余名令人瞠目结舌。可见,只有对评价结果的信效度不断质疑和审视,才可能客观理性地对待之,全盘接受只能带来实践的混乱和无措,在与初衷相悖的道路上越行越远。

(四)对评价结果的解读:重实体轻关系

评价结果解读是评价结果和评价结果应用的中介,直接影响着教育评价对实践改进发挥作用的方式和方向。其实质是评价结果对评价者、评价对象意味着什么,意义何在等问题的追问。只有科学地看待评价结果、合理地解释评价结果、深入地挖掘评价结果的意义,才能有效利用评价结果,充分发挥评价促进发展的工具性价值。

评价是主体依据客体对自身需要满足程度进行价值判断的活动。所以,只有从关系的视角审视,才能理解评价结果的意蕴,克服绝对性地看待结果的弊端。首先,评价结果内隐着“对谁而言”的主体言说立场,多元主体基于自身的信念、价值、目的等,对同一结果会得出有云泥之别的结论,而我们在谈论评价结果时,往往忽略了对主体的关注,只是极端化、单一化、孤立化地看待评价结果。如对于当前备受关注的大学排行榜,存在忽略西方话语体系的影响、忽视每所大学的发展历史、脱离大学的现实定位、不顾及大学与区域发展的关系等问题。其次,对基于不同价值判断标准得出的评价结果,应有不用的解读路径,否则就背离了评价的目的。如始于2003年的第一轮教育部本科教学评估,属于合格评估,以改进工作、达到基准为目的,如果用区分和选拔的标准来看,就会出现很多问题。对于新一轮评估的解读,如果仍然停留在合格评估上,则难以达到促进发展的作用。

当然,任何一种解读都是主体基于自身利益的合理选择,这“不是一种失误,而是一种局限,一种评价主体存在方式及由此决定的评价思维的局限”^[26]。教育结果有认知和非认知之异、有可测与不可测之差、有长期和短期之别,影响机制存在一因多果和一果多因的复杂情况,不能以此代彼也不能顾此失彼,如此才可能科学地解读评价结果。

(五)对评价结果的运用:重功利轻价值

以评促建、以评促改、以评促管、以评促强,是教育评价的本体功能,完善评价结果运用,综合发挥导向、鉴定、诊断、调控和改进作用。运用评价结果在全面诊断现有教育发展水平,发现薄弱领域和环节,为下一步改进提供依据,并发挥对未来工作的激励以及可持续的改进效应。但现实中往往把评估结果与相关利益直接或间接挂钩,成为评价对象赢得声誉、获取资源、争取利益的手段,呈现出严重的功利取向,带来了诸多异化现象。

评价结果过度的利益捆绑,使评价对象不惜运用各种手段谋取理想的排名或等级,在一定程度上影响了评价结果的客观性。还造成评价对象大肆宣扬成绩、刻意回避短板的常态,形成注重横向的相互比拼和竞争,忽视纵向的自我发展和提升的畸形心态。“加剧了高校内部的无序竞争和生态失衡。一些高校借着‘资源整合’的名义,以‘砍杀’或拆并非重点、非特色、非优势学科为代价,来确保重点学

科的权威地位”^[10]。加速了“五唯”倾向,却忽视对“唯”背后所沉淀的意义的追问,忽视了教育的生命本质,漠视人的内心情感和精神价值的行为取向,丧失了对多元价值的追求。总之,评价结果的功利取向带来教育评价“应然价值和发展意义不断被消解,加重了教育的不公平,破坏了教育生态的平衡”^[27],呈现出越来越多的消极性和负面性。所以,必须克服过分注重教育评价结果的短视行为,扭转功利化的倾向,在相关政策、利益和关系上淡化身份,使其“基于纯粹理性的教育理解,它不是与相关利益主体相联系的工作理解,也不是基于管理立场的现实导向”^[26]。

三、对教育评价结果困境的归因

发现问题是为了解决问题,而解决问题的前提是追溯造成问题的根源,只有如此,才能做到有的放矢。造成教育评价结果问题产生的原因是复杂的,既有教育评价理念的偏差,也有评价制度的不科学导向,还有评价结果本身运用的复杂性。

(一) 教育评价理念的偏差

“现在所面临的问题主要不在评估技术本身(当然这方面也需要不断地研究和改进),而在于评估背后的理念和逻辑出了问题”^[10]。“大学确立它的地位的主要途径有两种,即存在着两种主要的高等教育哲学,一种哲学主要是以认识论为基础,另一种哲学则以政治论为基础”^[28]。所以,教育评价是指向基于高深学问的人的发展和社会的进步,本质上具有学术性、人文性和服务性。但受我国行政管理模式的影响,行政化思维倾向普遍存在,具体表现为,一是用管理代替服务,发挥行政权力,一是将复杂问题简单化,追求效率。

教育评价的根本旨趣是推动教育目的的实现和促进教育的发展,具有服务性和工具性。一方面为有关部门了解情况及进一步决策提供依据,另一方面为被评对象改进工作提供方向。但由于对评价性质的理解不到位,行政主体往往将评价作为一种行政管理,实施主体将评价作为一种权力资源。由行政部门发起、操纵的评估,采取至上而下的科层制方式推行,评价结果与资源配置紧密相联。这种“行政化的思维方式,以短平快的办法,用最熟悉的运动方式捧出个别的一流”^[10],是与教育规律相悖的。而高校对资源的严重依赖性又不得不全员投入到既伤身又伤神的评价运动中。评价主体凭借科学守门人的身份,在评价过程中确证了自己的优越性,还借助对评价结果的话语权影响评价对象。这种错误的评价权力观,既容易使评价主体框定自我、迷失自我,也容易使评价对象在资源的束缚下过度迎合评价。声望利益的捆绑、人情文化的影响等,容易造成评价中的马太效应,加剧教育的不公平,丧失评价过程的客观性,影响评价结果的可信度,丢失评价活动的本真含义。

教育评价是评价主体运用一定方法、依据一定标准对教育活动进行的价值判断。其中,评价方法和评价标准取决于评价主体的立场、教育活动的属性以及两者之间的关系。管理者、企事业单位、教师、学生等都对教育活动的某一方面具有发言权,多元评价主体是高等教育评价的内在意蕴。教育活动是面向人的,人的存在样态、发展方式等决定了教育活动的复杂性。“人的本质总是与其基本的生命活动状态关联在一起的,是通过自身对象性活动和历史实践活动所呈现出来的自由自觉的生命活动”^[29]。任何评价都不可能是绝对的客观白描,而是特定视角下特定价值观的体现,价值既是客体属性对主体需要的满足,更是“一种意义赋予,一种人的合目的性的主动建构”^[30]。可见评价主体的多元

性、教育活动的人文性、价值判断的主观性等决定了教育评价的复杂性。面对复杂的教育,行政化思维将“职称、论文、文凭、奖项、帽子”等转化为一个个数字,然后进行排序定级,成为快捷高效、统一采用的方式。这既解释了“五唯”这一顽瘴痼疾出现及存续的原因,又注定了其会招致诟病。

所以,评价理念的偏差是导致以唯量化形式表征评价结果,以实体思维解读评价结果,以告知形式反馈评价结果,以盲目态度接受评价结果,以功利性衡量评价结果的根源所在。

(二) 教育评价制度不够科学

“任何一种文化的兴盛与危机都与其‘制度-机构’状况密切相关”^[31]。分数、名次、等级“凭借其强大的生杀予夺威力裹挟着资金、政策和价值渗透深入到大学的各个层面。高度利益绑定的量化评价,以及量化发展模式的路径依赖成为大学、学院、学者个人输不起的数字游戏”^[4]。尽管学界对名目繁多、违背教育本质的评价颇有微词,但又不得不参与其中,如“第一轮学科评估,南京师大、东北师大没参加,第二轮西南大学没参加,第三轮差不多所有学科比较强的学校都参加,到第四轮基本上没有一所大学敢不重视”^[10]。之所以如此,就在于将评价结果与资源配置、奖优罚劣等制度导向相捆绑,这不但催生了“当前大学无所不在的急功近利的思想,也成了学术不端行为和学术腐败的重要诱因”^[4]。评价对象通过各种手段谋求“看上去很美”的数字,不但能够得到现实利益,而且一旦通过评价便成为“实力”的象征,进而带来无形的边际效应。这进一步刺激了大学对“数字”的追求,而“本源意义上教育的情感投入、默默奉献和脚踏实地反倒被挤压驱逐成了稀缺品”^[4],形成劣币驱逐良币的恶性循环。“又爱又恨”成为学界对教育评价最直观的感受。

当前周期性迎评备战现象普遍存在,评价对象常常在评价临近时,全员出动、全副武装甚至以牺牲日常的教育教学工作为代价,而评价工作一结束便万事大吉。之所以如此,和“一评定终身”的制度导向不无关系。当前很多评价既不考虑评价对象的现实基础,也较少追踪基于评价结果的改进后效,评价对象只要在某一轮、某一阶段评价所倡导的“唯”的指标上表现突出,即可获得相关利益,从此高枕无忧。这种评改建分离的政策导向忽略对评价中发现的问题的解决,忽视评价的螺旋上升效应,尚未为评价对象创造一个自由宽松、专注于发展的环境,难以发挥持续过程管理和动态效果反馈的发展性功能。相对固化的评价结果,耗费了评价主体大量的时间精力,消磨了他们的评价热情,打击了评价对象参与的积极性,降低了他们过程性投入的动力,使教育评价在某种程度上反倒成为提高教育质量和实现人才培养目标的障碍。可喜的是,这种现象正在逐步改进,如“‘双一流’建设高校动态调整机制的设计,为‘非双’高校明确了追赶方向,留足了追赶空间”^[32]。

(三) 教育评价结果运用的复杂性

发挥评价结果的实践改进功能,是一项系统性、综合性的工作,需要整体发力、上下衔接、相互呼应,只有人、财、物的协同配合,才能达到理想的效果。这既需要各主体的协同配合、共同努力,也需要他们各就其位、各司其职;既需要他们以“育人”这一价值统摄作为制高点,也需要彼此在价值优先考虑下做出利益让渡;既需要营造改革的环境氛围,也需要提供相应的政策支持;既需要有充足的资源保障,也需要有专业的人力支撑。现实样态是行政权力具有资源配置权,但行政主体对教育实践的内在规律与发展逻辑往往不甚了解,以致无法开展与教育本质相匹配的评价活动、不能科学解读评价结果的意蕴、忽略对评价结果价值的深入利用,经常出现越位、缺位和错位的乱象。评价实施主体一方面受

制于行政委托,难以保持独立性,一方面得益于评价权力,难以摆脱利益牵制,在双重处境中,影响评价的客观性。落实主体尤其是处于基层的师生常常以局外人的角色游离于评价之外,在不能及时得到评价结果或仅被告知评价等级的情况下,既无法明了问题所在,更无从着手改善实践。

教育效果的严重滞后性、行为惯习的较强稳定性决定了运用评价结果指导实践的长期性。任何改革既会受到已有习惯的抵制,还要承担可能出现的不良后果,这注定是一个困惑和痛苦的过程。人本身的“习惯仿佛像一根缆绳,我们每天给它缠上一股新索,要不了多久,它就会变得牢不可破”^[33]这一特性,也会起到延缓作用。所以,在改革动力不足、前景不明确、风险不可期的情况下,面对教育评价结果所反映出来的问题,评价对象要么视而不见,争取不求有功但求无过地维持现状;要么快速投机地追求数量上的跃进以获取资源。这些做法既放弃对教育生命本质、心灵对话的追求,也舍弃对教育评价发展效用的探索;既不能协同深入地推进教育过程的实质性变化,也形不成教育质量优化过程的闭环。

四、教育评价结果的改进路向

对照教育评价结果中的问题,落实《方案》的评价理念,实现“破五唯”的价值追求,构建符合中国实际、彰显中国特色、具有世界水平的教育评价体系,是研究的归属和落脚点。评价“落地的关键在于‘主体’,在于‘人’”^[34],所以,需要突出评价相关人员的重要性。而人的任何行动都离不开制度的支持,所以,还要彰显评价制度的发展性。

(一) 突出评价相关人员的重要性

教育评价人员涉及行政人员、评价人员和落实人员等。随着教育评价的发展,“特别需要一个评价专业指南,严格规定一切专业活动,规范专业人员的操守。从而在评价政策的制定,评价活动的开展和评价数据的使用等方面,建立起一套专业标准”^[35]。

1. 提高行政人员的评价能力

行政人员是评价的组织者和评价方案的制定者,“联合国教科文组织把各国教育行政人员有没有相当的教育评价能力作为评估一个国家教育发达程度和教育效能的依据”^[36]。可见,提升行政人员的评价能力非常重要,这里的评价能力主要指对评价的认识及指导能力。

首先,行政人员的角色定位应由管理者转变为服务者。教育评价的学术属性要求其从政府主导回归到同行评价,走向专业化。如果行政人员定位不合理,就会以行政代替学术,违背教育规律。“管理就是服务,当一种服务不被服务对象所认可,还让服务对象有排斥感,这种服务是不合格的”^[10]。所以,行政人员要由管理者变为真正的服务者,扮演好提供支持和搭建平台的角色。其次,行政人员的服务方式要由管理走向治理。在深化教育管理体制改革、促进管办评分离的背景下,行政人员要构建评价利益相关者间的新型关系,引导多元主体平等地参与评价过程,通过协商回应形成相互制衡的共治格局,以推进治理能力和治理体系的现代化。

2. 增强实施人员的专业性

实施人员是运用方案实际开展评价工作的人,他们对方案“是否具备足够的理解能力、运用能力,以及运用之后的反思能力与重建能力,至关重要”^[34]。他们是增强教育评价学理依据、优化教育评价过程、改进评价结果的关键因素。只有他们真正地理解评价的目的、科学地使用评价工具、理性地看待

评价结果、不断地积累实践经验,才能真正实现评价的专业化。

加强评价人员的培训,是提升其专业性的重要手段。培训内容可包括教育评价的本质与价值、评价人员的角色与职责、评价实施的理念和策略、评价工具的设计与运用、评价结果的解读与利用等专题。培训形式切忌采取单一的讲座式,应将工作坊、任务导向式、小组焦点研讨、案例分析等融为一体,以优化培训效果。培训时间采取常规和集中相结合的方式,常规培训侧重教育评价基本理论。集中培训可围绕某一次评价,促成教育评价领域的专家学者、评价活动的实施者和评价对象等围绕评价全过程进行协商对话,呈现多元主体的多重视角,以在交流讨论中达成共识。可以借助信息技术手段,丰富培训形式、扩大培训规模、加强过程管理、收集相关信息。充分发挥大数据优势,了解强弱项的诊断信息,把握纵向的发展趋势,进行横向的相对比较,为应用评价结果改进实践提供参考。

3. 调动落实人员的积极性

落实人员是指根据评价结果进行改进的评价对象,只有他们将评价结果转化为“治理能力、管理能力和教学能力的一部分”,评价“才真正实现了内化与转化:内化于教育主体的教育理念,转化为他们日常的教育行为”^[34]。要做到这一点,必须明确评价对象的意义追寻、满足评价对象的内在需求,做到评价过程和结果的公平、公正。《通知》中提到“北京大学、清华大学在第二轮‘双一流’建设中自主确定建设学科并自行公布”,就是进一步给与落实人员自主权、调动其积极性、体现其专业性的最好例证。

符合评价对象内在属性的评价才可能得到评价对象的认可,所以,评价工具的选择与确定应以评价对象的本质把握为基础,以与其职责、内涵和结构相适应为依据,以全面准确体现被评价对象的真实状态为本质追求,以促进其不断调整完善自身为目的。就评价理念、评价目的、评价过程等与评价对象及时地进行沟通协商,广泛吸纳利益相关者参与评价,倾听他们的意见、获得他们的理解、争取他们的参与。这种具有交互建构性的评价,才有可能使评价对象明晰评价结果的含义,合目的性和合价值性地解读评价结果,明确评价对自身的意义,从而参与后续的建设性行为。关注评价对象的差异性、发展性和个性化,实现对评价对象的联动、整体和动态评价,这种实质意义上的公平、公正评价,才能得出信效度较高的评价结果,才能得到评价对象的认可,也才能找出评价对象存在的不足,为其提供改进措施,进而促进评价对象的发展。

(二) 彰显评价制度的发展性

制度是调整活动主体之间社会关系的规范体系,是教育评价发展的风向标,优化制度设计是保障和促进高校教育评价发展科学性、客观性和持续性的重要手段。加强教育评价制度的顶层设计,转变教育管理方式,超越教育评价现状,引领教育评价健康发展。

1. 健全评价过程的监督机制

为了保证评价过程和结果的真实性和客观性,减少政府“守门员”和“裁判员”双重角色的影响,规避评价中的人情怪圈,回归健康的学术生态,有必要建立健全评价过程的监督机制。《办法》中提到“建设高校在动态监测、中期自评和周期评价中应确保材料和数据真实准确。凡发现造假作伪等情形的,建设主管部门将视情节予以严肃处理。情节严重的,减少支持直至调整出建设范围。”但如何发现造假行为、如何规避虚假数据是评估中面临的难题,需要不断进行实践摸索。

建立具有公信力的专业机构和社会组织等第三方评估机构,是提高评价可信度的有效路径。在我国缺少独立的第三方评估机构的现实境遇下,要厘清其与政府的边界,切断其与评价对象的利益链,必须使评估机构真正“由委托走向授权,评估机构不隶属于教育行政部门,评估的权利由法律赋予^[36]”。在这个过程中,政府需要加强对第三方评估的监管,通过政策引导,“制定教育评价专业指南,建立第三方教育评价机构的准入和监管机制,以及进行教育评价人员的专业资格认证^[35]”等路径,保证其规范性,提升其专业性。建立健全各方协商对话机制,增强评价各方的责任感,在沟通中加深对评价的认识。“评估者以及各方利益相关者应该多方共同协商一个评估标准,标准制定过程要公开、民主、广泛听取不同方面的意见^[10]”。《办法》中提到按照“建设高校自我评价、教育部委托相关机构提出初步评价结果、‘双一流’建设专家委员会形成综合评价意见、‘双一流’建设主管部门综合研究确定评价结果”等程序评价“双一流”高校周期建设成效,这体现了建立评价结果协商机制的初步尝试,唯有如此才能保证评价的可接受度和认可度。

2. 优化评价结果的使用机制

为了扭转当前评价中的“五唯”倾向,克服评价对象对量的过分追逐,发挥评价结果的激励与约束,形成闭环的良性系统,必须优化评价结果的使用机制。首先,健全评价结果的反馈机制。建立评价结果反馈通道,确保评价结果及评价标准及时、有效、全面地传递至评价对象,否则任何评价对评价对象的影响都是无力和无效的,也无法调动他们的参与意愿。丰富评价结果的反馈形式,实现评价结果多维多样化的呈现,量化和质性相结合,尤其侧重描述性反馈,形成有针对性的评价结果分析报告,使评价结果成为评价对象改进工作的切实指南。破除人才评价、学科评估、绩效考核等对量化指标的异化和扭曲使用倾向,“以量化指标为基础,找非量化点突破,以此逐步建立起适应新时代的教育评价体系^[16]”。其次,淡化评价结果的利益负载。改革教育资源的配置方式,不得将评价结果与学校工程项目、经费分配、评优评先等挂钩。“最大限度地降低评价的利益名利捆绑负载,这样才能有效克服评价水分,杜绝粗制滥造冲动^[4]”。不计算总分、不发布排名,如此评价对象才不会盲目追随评价甚至迎合评价,不受评价的排名、指标等制约和干扰,心无旁骛地做好教育,回归教育的本真,教育评价才能充分发挥工具价值。就如《方案》中提到的“使潜心育人的评价制度更加健全,促进学生全面发展的评价办法更加多元”。

3. 建立评价改进的追踪机制

《方案》中提出既评估最终结果,也考核努力程度及进步发展,《办法》中提到设立常态化建设监测体系,注重考察期中期末建设目标达成度、高校及学科发展度,形成监测、改进与评价“三位一体”评价模式,综合评价结果作为下一轮建设范围动态调整的主要依据。体现了注重评价发展的过程性、持续性和改进性,所以,应建立评价改进的追踪机制,真正发挥评价的调控和改进作用,持续提高建设水平。

首先,建立持续性监督改进机制。将总结性评价与形成性评价有机地结合起来,形成闭环的评价过程,对改进效果进行再评价,真正推动教育实践的持续改进,实现评价结果的稳定性和动态性协调发展。可以运用“动态数据获取技术、大数据挖掘与分析技术、数据呈现技术等^[23]”,对教育发展状态进行常态监测和直观呈现。其次,建立评价结果的实践指导机制。评价主体与评价对象就评价结果开展

对话,客观理性地对待评价结果,探讨评价结果的参考价值,找到实践改进的突破点,不断修正问题,从而强化评价结果的使用。定期讨论基于评价结果的整改进度、人员参与度、实践问题、如何克服困难等,使基于评价结果的改进活动常态化。最后,建立有限的问责制。在提供各种改进保障的前提下,评价对象的工作仍然实施不力、进展缓慢、缺乏实效,应该给与一定的惩戒。

五、余论

改进结果评价是深化新时代教育评价改革的关键环节。立足实际,剖析评价结果在呈现形式、反馈方式及相关主体所持态度、解读路径和运用取向等方面所面临的困境。从理念、制度和操作层面,深入反思问题产生的根源。从人员和制度两方面着手,为《方案》精神的落实,提供参考。改进结果评价道阻且长,如何在“破五唯”的同时建立更好的评价机制?如何评价对象的成长性、特色性发展?如何在服务国家发展战略的同时淡化评价结果中国家政治平衡的色彩?评价的边界如何确定?哪些人负责提升评价相关人员的素质等,这些问题都值得我们不断去探索。

参考文献:

- [1] 韦永琼. 教育评价能否关照人的内在性? [J]. 山西大学学报(哲学社会科学版), 2019(3):100-106.
- [2] 杜瑛. 我国高等教育评价的范式转换及其协商机制研究[D]. 上海:华东师范大学, 2010.
- [3] 人民网. 教育部启动实施新一轮高校本科教育教学审核评估[EB/OL]. (2021-02-07). <http://sn.people.com.cn/n2/2021/0207/c393584-34569503.html>.
- [4] 高江勇. 大学教育评价中的过度量化:表现、困境及治理[J]. 中国高教研究, 2019(10):61-67.
- [5] 王薇, 张鑫. 学校评价结果解释模型的建构[J]. 中国考试, 2021(1):89-94.
- [6] 许海霞, 王蕊, 马陆亭. 教育评价改革的几个关键问题[J]. 中国考试, 2020(8):20-23.
- [7] 王薇. 奖励·问责·改进:让学校评价促进学校发展[J]. 中小学管理, 2017(3):51-53.
- [8] 潘杰宁, 黄海滨. 高等教育评估结果的失真分析[J]. 太原师范学院学报(社会科学版), 2007(1):145-146.
- [9] 周光礼, 袁晓萍. 聚焦“四个评价”深化教育评价机制改革[J]. 中国考试, 2020(8):1-5.
- [10] 陈学飞, 叶祝弟. 中国式学科评估:问题与出路[J]. 探索与争鸣, 2016(9):59-74.
- [11] 刘振天. 高等教育评估结果的使用及其规范探究[J]. 中国高等教育, 2013(21):22-25.
- [12] 王薇, 张鑫. 学校评价结果解释模型的建构[J]. 中国考试, 2021(1):89-94.
- [13] 王薇. 学校评价结果解释标准探析[J]. 上海教育科研, 2020(2):47-52.
- [14] 王薇. 学校评价结果解释的基本问题与前瞻研究[J]. 上海教育评估研究, 2019, 8(6):7-11.
- [15] 王薇. 学校评价结果解读:视角、方法与应用[J]. 上海教育科研, 2019(5):39-44.
- [16] 马陆亭, 王小梅, 刘复兴, 等. 深化新时代教育评价改革研究(笔谈)[J]. 中国高教研究, 2020(11):1-6.
- [17] 教育部. 中共中央国务院印发《深化新时代教育评价改革总体方案》[EB/OL]. (2021-10-13) [2021-12-09]. http://www.moe.gov.cn/jyb_xwfb/xw_zt/moe_357/jyzt_2020n/2020_zt21/.
- [18] 教育部 财政部 国家发展改革委 关于公布第二轮“双一流”建设高校及建设学科名单的通知[EB/OL]. (2022-02-09). http://www.moe.gov.cn/srcsite/A22/s7065/202202/t20220211_598710.html.
- [19] 教育部第四轮学科评估结果(完整版):权威高校学科排名[EB/OL]. (2020-07-01). https://kaoyan.eol.cn/nnews/201712/t20171229_1577397.shtml.
- [20] 王建华. 量化评估与大学发展[J]. 高等教育研究, 2020(11):33-41.

- [21] 刘少雪. 高等教育评价中的“数字陷阱”[J]. 苏州大学学报(教育科学版), 2016(1):28-35.
- [22] 财政部. 关于印发《“双一流”建设成效评价办法(试行)》的通知[EB/OL]. (2020-12-15)[2022-01-06]. http://www.mof.gov.cn/zhengwuxinxi/caizhengxinwen/202103/t20210324_3675365.htm.
- [23] 瞿振元, 张炜, 陈骏, 等. 深化新时代教育评价改革研究(笔谈)[J]. 中国高教研究, 2020(12):7-14.
- [24] 汪建华, 徐颖婕. 教育评估结果反馈探究[J]. 上海教育评估研究, 2020(2):65-69, 86.
- [25] 杰瑞·穆勒. 指标陷阱:过度量化如何威胁当今的商业、社会和生活[M]. 阎佳, 译. 上海:东方出版中心, 2020:3-4.
- [26] 刘庆昌. 一种弱功利的教育评价哲学[J]. 教育发展研究, 2018(12):1-11.
- [27] 刘志军, 徐彬. 教育评价:应然性与实然性的博弈及超越[J]. 教育研究, 2019(5):10-17.
- [28] 约翰·S. 布鲁贝克. 高等教育哲学[M]. 王承绪, 等译. 杭州:浙江教育出版社, 2001:13.
- [29] 韩庆祥, 邹诗鹏. 人学:人的问题的当代阐释[M]. 昆明:云南人民出版社, 2002:191.
- [30] 李铭磊, 韩延伦. 回归生命:教师教育本体价值刍论[J]. 当代教育论坛, 2018(5):48-54.
- [31] 刘冲, 巴登尼玛. 文化、社会与教育的关系[J]. 当代教育与文化, 2019(6):14-22.
- [32] 段从宇, 伊继东. 高校教师流动的本质内涵及合理性判别:兼论“双一流”建设背景下的高校引才[J]. 高校教育管理, 2019(3):89-96.
- [33] 赵振杰. 论教育哲学视野中的习惯养成[J]. 辽宁师范大学学报, 2005(1):60-64.
- [34] 李政涛. 把新时代教育评价改革深化到“评价能力”的提升那里去[J]. 中国教育学刊, 2020(12):8.
- [35] 张娜. 从对教育的评价到促进教育的评价:教育评价国际研究进展综述[J]. 基础教育, 2017(4):81-88.
- [36] 王红, 吴强. 高等教育评估的马克思主义哲学审视[J]. 国家教育行政学院学报, 2015(12):23-28.

Results of higher education evaluation: Dilemma, reflection and direction

SHI Yanfang^{1,2}

(1. School of Educational Sciences, Sichuan Normal University, Chengdu 620021, P. R. China;

2. College of Education, Zhangjiakou University, Zhangjiakou 075000, P. R. China)

Abstract: Evaluation results are the bridge between evaluation activities and practice improvement. Only by perfecting the application of evaluation results can we form a closed-loop process of improving education quality and highlight the development value of evaluation. However, in the process of concrete implementation, the evaluation results are either ignored, or presented as competitive, hierarchical, etc., which are not conducive to the sustainable development of evaluation. The evaluation results have not been fully utilized to implement the goal of “strong hardness” and “long teeth” and promote the improvement of educational governance efficiency. At present, when academic circles talk about evaluation, they usually involve the related problems of evaluation results, but there is a lack of specialized research. The double deficiency of practice and theory provides space for further research on evaluation results. In October, 2020, the first document on systematic reform of educational evaluation in China—“The General Plan for Deepening the Reform of Educational Evaluation in the New Era” (hereinafter referred to as the “plan”), proposed to reverse the unscientific orientation of educational evaluation and improve the evaluation of results, aiming at reflecting on the negative impact of the current popular evaluation of results, thus reforming the “five-dimensional” tendency that hinders the development of high-quality

education, promoting the construction of double first-class universities, and further educating people with evaluation empowerment. The publication of this document not only reflects the importance attached by the state to the evaluation, but also highlights the urgency of solving the problems in the evaluation. Based on this, with the value pursuit of “improving the evaluation of results and strengthening the use of results”, this paper comprehensively looks at the evaluation results of higher education from the aspects of presentation form, feedback mode, attitude of relevant subjects, interpretation path and application orientation. It is found that there is a common phenomenon of “five-light and five-heavy” in the current evaluation of educational results, that is, it presents the dilemmas of emphasizing quantity over quality in form, emphasizing notification over negotiation in feedback mode, emphasizing acceptance over questioning in attitude towards results, emphasizing entity over relationship in the interpretation of results, emphasizing utility over value in the application of results, etc. Finding the problem is to solve the problem, and the premise of solving the problem is to trace the root cause of the problem. Only in this way can we aim at the target. Following the thinking logic of “idea-system-operation”, we should deeply and systematically reflect on the root causes of problems, which are caused by the deviation of educational evaluation ideas, the unscientific evaluation system and the complexity of application of evaluation results. Perfecting the application of results evaluation, implementing the evaluation concept of the “Plan” and realizing the value pursuit of “breaking through the five only” are the ownership and foothold of the research. In view of the fact that the key of evaluation lies in people, and any action of people can not be separated from the support of the system, this paper selects the analysis framework of “people-system” and puts forward detailed and feasible opinions, which can provide reference for giving full play to the development function of evaluation, establishing and perfecting the mechanism of cultivating people by virtue, and promoting the all-round development of people and building a socialist modernization power mentioned in the Sixth Plenary Session of the 19th CPC Central Committee. Suggestions are put forward including improving the evaluation ability of administrative personnel, enhancing the professionalism of implementers and mobilizing the enthusiasm of implementers, perfecting the supervision mechanism of evaluation process, optimizing the use mechanism of evaluation results, and establishing the tracking mechanism of evaluation improvement, etc.

Key words: evaluation of higher education; evaluation results; break the “Five Only”; utilitarianism; administration; “double first-class” construction; first-class universities; first-class discipline; discipline construction

(责任编辑 王淼卉)