

效率还是价值: 高校教师评价制度的逻辑困境及其张力调适^{*}

苏 强¹ 蔡晓卫²

(1. 杭州电子科技大学 浙江高等教育研究院, 杭州 310018;
2. 浙江大学 马克思主义学院, 浙江大学马克思主义理论创新与传播研究中心, 杭州 310058)

摘 要: 近些年来盛行的国内高校教师评价活动, 科研导向、定量方式、绩效至上始终占据着主导地位, 反映出只重事实层面的合理性而轻价值层面正当性的现象。高校教师评价取向与标准主要非来自理论逻辑的自治, 而是现实中外来力量对高校学术生态的强力介入及其与高校博弈的结果, 进而形塑了教师的工作生活方式以及学术信念与行动偏好。因此, 从价值层面对时下评价制度的结果逻辑、计数逻辑与绩效逻辑予以全面检视, 继而调适教师评价目的性与工具性价值间的张力尤为必要。

关键词: 高校教师; 教师评价; 评价制度; 组织变革; 教师行为

近些年来, 我国高校纷纷启动并推进的教师评价制度改革, 不仅引起高校内部组织与治理结构的变迁, 也形塑了教师的学术选择偏好、职业生活形态乃至集体行动理性。然而, 在追求绩效量化指标精细化以增进科研产出效率与数量, 应对外部绩效问责与系统内部实力比拼的同时, 相当程度上弱化了对既有教师评价活动本然价值的反思, 进而在高等教育系统内部产生种种异化现象, 引发了政府、社会与学界广泛关注的“五唯”问题。习近平总书记在 2020 年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”(习近平, 2020, 第 348 页)。“五唯”关涉各级各类教育的功能与定位、学术与教育生态、评价泛化及错位、政策的科研偏好等系列问题, 试图对其进行全景式展现与系统分析非一篇短文所能驾驭, 也非本文立意所在。本文主要聚焦于建制性评价制度环境中高校教师所面临的学术工作与生活境况, 探究评价问题背后的深层机制与可能的行动路径, 以期政策转化落实及其相关问题的讨论提供一种经验性参考。

一、结果逻辑: 高度竞争性资源分配与评价标准的偏颇

我国政府作为高校的举办者具有指导权与问责权, 而高校对办学资源则存在高度依赖性。20 世纪 80 年代以来, 政府开启并推动了赋权高校的进程, 高校在形式上获得更多自主权的同时, 由政府主导和市场参与的多元评估, 如政府委托的学科评估、民间组织的大学与学科排行、ESI 学科水平评估等, 这些评估不无将科研成果和学术表现作为重要指标, 且或隐或显地影响政府政策决策与相关制度安排。而且政府资源配置特别倚重的各种工程或项目, 偏好选择那些高显示度、可计量的科研指标作为门槛, 促使高校群体在相互竞争中实现资源获取。现实中少有高校能置身于外, 因为其直接关系到

^{*} 基金项目: 2019 年国家社科基金专项“习近平新时代中国特色社会主义思想进教材进课堂进头脑传播效果研究”(19VSZ010)。

自身的绩效问责、学术地位与社会声誉,政府透过放权反而加强了对高校的约制。以若干所“双一流”建设高校建设方案文本为例,众高校对建设基础、目标与内容进行细致的指标设定,如在世界或国内排行榜中学校位次的提升,进入ESI前1%或1‰的数目等,将之视为跻身一流大学行列的重要指标,以作为学校整体建设成效的显性证据。

高校在将自上而下的政府政策目标转置为本校发展规划的过程中,出于绩效目标与任务的达成及确立自身比较优势,处在传导链底端的教师群体自然被视为核心的科研力量并委以重望重担。高校把组织目标与任务细化分解,具体化为教师绩效尤其是科研任务要求,将之与年度考核、职称晋升等制度相捆绑,纳入量化指标控制的人事管理系统,推动了高校内部组织管理与评价制度的全方位变革。

需要指出的是,这种高校内部压力的层层传递,不仅来自政府的非均衡资源配置、行政问责的压力,还有来自整个高等教育系统内部的高度竞争。由此,高校内部制度与行动逻辑的变迁,带有明显的利益驱动与零和博弈特点。高等教育系统是一个层次和类型多样化的场域,存在着高校层级、办学声誉、研究实力的结构性差异。出于知名大学的竞争优势、市场认同和办学收益所产生的溢出效应,其他高校在评价内容选择与方案制订上趋于模仿,如重视教师的科研业绩、突出科研权重,将科研产出作为提升追赶效率和实力比拼的利器,以降低外部压力与制度自主探索风险。从制度理论的观点来看,制度结构成为行动主体表现各种行为的具体场景,那些竞争能力强的组织往往试图把其自身的目标和程序作为一种制度规则,吸引其他行动主体进入该制度框架,以实现跟随性发展和获得自身利益的最大化(操太圣,2019)。其实还不止于此,处在系统中心外围的高校为了寻求地位上升,在以强校的评价制度为基本参照的同时,形成了更为严苛且具体的科研绩效细分机制。如有的高校在教师科研绩效考核基准分上,外加同等分值的所谓“资源分”——对教师使用的办公设施、设备与用品等加以折算——计入科研绩效总分等,以商业化操作迫使教师投入更多时间精力用于科研生产。

高等教育系统内部利益博弈与外部压力相叠加,相当程度上强化了系统内部本就存在的办学模式和风格趋同机制。一方面,不同高校的教师评价标准与内容高度重合,淡化依据自身办学定位和职能特点建构评价制度。而事实上,不同类型或层次的高校,由于历史传统、学科布局与办学定位的差异,有着不同的办学模式和不同层次的特色要求。同质化评价标准不仅有益于特色办学文化的沉淀,也容易脱离本校教师的发展需求、职责侧重与专业能力的特点。另一方面,同一所高校内部漠视不同教师的岗位主责主业差异,采用统一指标与权重对其学术业绩进行衡量;这种不考虑学科特色、功能与定位的评价虽然容易操作,但通常会抑制学科发展的活力并加剧不同学科群落间的不当竞争。正如劳德尔等人所言,每个学科领域用来与学术社群沟通的形式不同,但是评价往往只用一套统一的标准与指标,如果继续采用单一的标准化方法,则科学社群彼此间的沟通将产生紧张关系,影响学科领域的发展,其中社会科学、艺术与人文学科受害最大(Laudel, Glaser, 2006)。为扭转不科学的教育评价导向,2020年国务院颁布了《深化新时代教育评价改革总体方案》纲领性文件,教育部、科技部随之出台相应文件对“五唯”问题展开专项行动,但受制于路径依赖的惯性与自我强化,高等教育系统中偏好高显示度科研结果的评价取向与逻辑难以短时改变,且以一种更为隐匿的机制促使教师科研产出最大化。

科研产出偏好是“五唯”系统中的制度源头。客观而言,它对组织目标在工具理性意义上的达成,以及对教师科研能力的基本识别有着效率与简明优势。但问题在于,这种制度框架所建构的竞争机制,深层上是把提升学校绩效的组织目标作为教师评价制度设计的圭臬。如高校在人事制度改革过程中引入“非升即走”的预长聘制,主要目的在于产出高显示度科研成果以应对外部绩效问责与系统内部竞争,并非为教师的安身立命、职业发展而设计。它无需通过行政强制,只要渲染一种以稀缺资源为导向的分配氛围即可将其纳入竞争轨道。众多青年教师为获得数量极其有限的教职,尽管集中心力在聘期内产出的科研成果颇丰,但富有想象力和创新性的研究严重缺乏,甚至换来的是恶性竞争、学术不端与矛盾激化的学术生态,近期中国论文被频繁撤回的报道、某些教师为获取教职以命相搏的事

件是典例。针对重科研产出的学术生态迷局,阿玛兰蒂认为,尽管教师绩效的形成性评价非常重要,但工具理性仍然是首选,这更多地与行政目的有关,而不是与提高教学质量相关(Amaranti, 2017)。在组织目标僭越个人发展的取向下,过分强调科研职能用于外部竞争的外适性,使得高校管理与学术研究的功利主义愈发盛行。它更倾向于把人的行动看作是基于个人成本与收益计算的理性人行为,也结构化为日益严苛的学术晋升与聘任考核条件、刚性的科研项目经费与成果指标。组织内部所建构的制度环境,相当程度上形塑了教师的学术取向与行为选择,成为一种纵贯教师学术生涯始终的外部驱动机制。其结果是,当评价标准聚焦于学术产出,教师为了自我生存与发展,按照评价标准及其指标来框定学术活动方向与内容,忽视那些无直接相关的社会性、情感性与道德性的发展活动。教学评价则因时效的滞后性、复杂性而多采用基本工作量考核。而更具高显示度的研究成果,自然就成为教师学术资本与政治权力获取的重要依据。正如布迪厄所言,大学教师的文化资本虽有相对自主性,但依旧处在经济资本之下,加上其居于的社会位置,实难摆脱经济与政治权力掌控者的支配。于是大学教师的处境渐趋向科层制企业中的受薪阶级,无法完全掌控其知识生产工具(Bourdieu, 1984, p. 152)。

需要追问的是,就教师评价制度的价值预期来说,结果逻辑是否符合当下“双一流”建设的归旨,是否符合学术职业的内在特质与教师生涯不同阶段的差异性需求?过度注重科研指标是否会遮蔽教师工作职责与任务的多样性,是否会挫伤教师的学术信念和工作激情,进而降低其职业归属感、成就感乃至安全感?希尔曾批评大学教师评价过度重视学术成果的现象,认为学术组织所发展的教师专业发展方案,似乎是不切实际的时尚精品,要想真正落实到教师本身,就应该推出量贩式如沃尔玛的差异化商品,即符合教师需求并将教师表现给予多样化呈现,才能满足其实际评价与教师发展的需要(Hill, 2009)。在科研绩效的压力和驱动下,外控的制度设计凸显的是新自由主义所强调的效率、绩效与市场价值,并借助多种信息技术全程监控教师的科研绩效,经由教师的自我审核以维持自己的学术表现与生产力,在工作中逐渐内化成为行动者的自我要求。

二、计数逻辑:理性看待量化评价的价值与限度

评判教师对于既定评价标准及其指标的达成度,需要选择一定的评价方式与工具。由于教师个体学术能力或潜力的非直观性,通过外化成果予以甄别已成一种惯例。以至于目前各种以数字指标为核心的评价制度,将教师的教学、研究与社会服务等业绩表现量化赋分,以衡量教师学术产出价值和社会效用。这种制度化评价强调价值中立原则以及外在的可视化、可量化,建构起一种刚性的数字量化指标作为教师学术实力的认定依据。由于量化从外在形式上保证了评价的客观性和透明度,具有操作简便、经济、高效率的优势;定性评价则因目标模糊化、耗时性和复杂性等特征,为其落地实施带来不可控的风险。尽管众高校都声言采用定量与定性综合评价,但实践中出于便捷及规避矛盾纠纷的考量,量化成为一种标识意义上的程序性操作是常态。如高校行政部门建立的系列围绕期刊分区分级、高引用指数等指标建立的量化体系,各种数据始终作为衡量教师科研产出价值的重要依据,并覆盖至教师年度考核、职称晋升、学术奖励、岗位竞聘、薪酬分配等诸多环节。的确,量化评价对提高科研成果数量和质量以及国际学术交流能力助益良多;然而,对定量评估的日益依赖涌现出各种异化问题。

其一,量化评价容易诱致评价活动的价值取向偏离。在追逐“学术GDP”的高等教育场域中,因为评价的是数字化指标,科研成果的显示度、量化值往往成为评价活动与学术价值本身。在特定的制度安排下,面对如强调量化评价、过于频繁的考核、数量化考核结果与晋升及待遇挂钩、项目研究周期过短等,迫于压力也是出于风险规避的考虑,教师往往会理性地选择风险小、周期短但可能是低层次甚至重复性的研究,以提高产出的效率(阎光才, 2018)。量化评价体现了程序正义,但未必符合结果正义;简单量化并不能完整体现科研成果的内在价值,因为使用指标的优点是易于量化,不在于评估业绩表现的准确性,容易引发教师的争议以及或显或隐的抗拒。

其二,违背科研活动规律不利于学术创新。高校是以学术活动为中心的社会组织。学术研究是一

种自由探索未知领域的知识创造与理论创新活动,具有目标相对模糊性以及指向内容不确定性的特点,而且不同学科的属性、知识生产方法与路径也各异,单一的定量方式并不能准确呈现知识探索和创新的价值。采用一统的量化方式来管理所有学科,如通过发表数量、影响因子、引用频次等表面客观的指标,把不同学科纳入统一评价标准框架衡量科研成果价值,不仅违背了不同学科的知识生产规律,还轻忽了不同学科发表形式的差异,无法完整呈现受评者的学术成果价值。美国著名社会学家默顿指出,历史上有不少学术发表并非以文字进行,而是以演讲形式在教室、研讨会等场合中进行知识交流与传递。经济学家阿尔弗雷德·马歇尔、亚当·斯密,哲学家维特根斯坦等人,都不很热衷于文字的发表,但他们的影响力却无人忽视(赖鼎铭,2005)。

其三,削弱本土意识的独特性与主体性。目前高校量化评价尤为倚重的SCI数据库,它除了重视基础学科的学术问题,相关指标并不能全面反映科技创新水平和社会贡献;同时,SCI数据库存在广受诟病的语言、政治与地域等问题,如期刊分布地区的不均衡性,尤其侧重收录美国以及英语为母语地域的英文期刊。同样,21世纪初引入中国、近年来备受推崇的SSCI数据库也存在同样问题。诚然,SCI、SSCI对国内研究打开一种国际性视野,逐渐摆脱以往的封闭状态、融入国际学术话语体系的作用自不待言,但过分倚赖其作为重要的数字指标,形塑了国内高校教师的学术发表偏好。以SSCI为例,对它的偏好其来有自。由于研究生扩招带来学术群体的急速扩增,国内CSSCI来源期刊出于提高引用率等方面考量大幅降低刊文量,两者间此消彼长,加之论资排辈、利益纠葛等因素,发表难度逐年增大。而SSCI相对客观公正的评价标准与操作流程,相当程度上规避了学术评价中非学术性因素的介入;况且,在职称评审的科研权重与学术奖励额度上,也远高于国内绝大多数期刊论文。然而让人忧心的是,不论是问题意识、研究主题、使用语言还是研究范式都明显西化,淡化了学术研究对本土重要议题的理论解释力与情境关怀,容易诱发学术活动的依附现象乃至学术话语权的丧失。

定量评价之所以会引起人们如此强烈的反应,主要是因为就其开发、设计依据与现实效果而言,已经远远超出了作为一个工具本身所具有的外在功用价值,而是触及到人们对学术以及学术活动的内在本质的理解,对学术共同体内在精神气质、规范伦理和文化的体认(阎光才,岳英,2012)。存在缺陷绝不意味着我们不需要量化评价,关键在于如何对它进行价值功能定位与实践把握,是看作评价结果本身,还是仅仅作为一种评价工具。如果执持用它衡量学术成果价值并与资源分配紧密关联,极易形成趋利倾向的学术资本主义,进而对学术活动规律甚至整个学术环境造成严重侵害。若评价方式与资料搜集能做到多样化及针对性,将较容易获得教师的认同与信任,并可为其专业表现提供有效反馈及改进。因此,如何提升评价标准的适切性、加强评价过程的协商性、增进评价结果的信效度等,是教师评价制度所面临的根本性问题。

三、绩效逻辑:价值偏离引发组织与个人关系的内在紧张

评价结果的使用取决于评价目的与价值诉求。一般而言,教师评价具有绩效问责与专业支持两种目的。前者依循组织绩效为目标导引的管理逻辑,后者依循促进组织内部教师专业发展的学术逻辑。国内高校的教师评价制度,对绩效主导的问责取向尤为偏好。一项对国内18所高校的调查发现,受访者中的91.4%认为强调绩效问责的终结性评价占据主导地位,但教师最认同的形成性评价所占比重仅为4.3%(董彦邦,刘莉,2021)。绩效问责取向究其本质,是基于科学绩效管理,把质量与评价结果作为教师业绩认定、薪酬分配等方面的证据,实质上是反映了大学内部资源分配与奖惩制度,且尤为强调组织目标达成的路径与手段。需要进一步思考的是,无论是教学还是科研绩效评价,是否客观如实地反映教师的专业能力与潜质,是否符合不同阶段教师的学术成长规律,是否符合不同学科或岗位类型教师的任务特点?这尤其关涉绩效内涵的明确界定以及由之引发的公正性问题。更何况,绩效具有事实表现与价值判断属性,在评价的现实操作中却往往以可测化的显性指标为重,科研潜力与学术志趣等内隐性因素被忽视,引入教师专业发展的形成性评价严重不足。

退一步讲,即便绩效评价的信效度合乎逻辑,可以在某种程度上区分教师专业实力与水平,但绩效问责的意图与目的不在于发展教师,而在于把教师业绩表现给予分等以作为人事决策与管理的依据。它所呈现的主要是评价的甄别与奖惩功能,对借助评价来促进教师专业发展的考量还远远不够。这里不妨以当前高校广为采用的生评师为例。当前高校教师教学评价的普遍做法是生评师,这是一种被人们认为最为直接、重要与客观的教学评价方式。学生作为教学质量的第一感受者,所提供的教学反馈意见无疑具有第一手价值,国内外也因此把学生的教学反映作为评判教学质量的重要指标。然而,由于受到学生的评价能力、学习意图、个体特质等多重因素影响,容易造成评价结果的失真。如在一些思想政治理论课中,大多数学生更倾向于选择一些“好通过”且给分高的“水课”,学生也报以评教高分,两者间形成一种“教学相涨”现象;同时,人力与时间的高成本、人情关系等因素也使得督导评价、同行评价形同虚设。若把生评师结果作为教师履行教学水平高低的根本依据,难免让教师衍生教学活动的失落感、威胁感,而很难激起教师改进教学的强烈意愿和动机。

不可否认,高校作为一个理性的组织,在教师评价活动中自然要着重考量组织发展目标的达成,评判教师业绩之于学校发展的贡献度并给予一定奖惩,以提升学校办学绩效与核心竞争力。但在学术岗位与优质学术资源日益稀缺的境况下,置身其间的教师不论处在职业生涯历程中的哪个阶段,为谋求职业保障与发展空间都要持续地参与各种竞争。这种类似锦标赛竞争规则推崇的是实力至上通则,尽管与政府阶段性政策偏好存在关联,但高校内部对于科研产出的过度功利化追求极大加重了教师工作压力。有研究者对240名高校教师的调查表明,不同职级教师具有程度不同的教学科研和个人发展压力,压力均值为58.58分,达到中等偏上水平。研究还发现,教学年限越长,职业压力越大,这似乎超出人们的日常认知。其背后除了科研问责的压力,兼顾不同任务和社会责任的角色超载也是要因(Meng, Wang, 2018)。关于上述问题,人们基于学术职业与教师实践关怀立场的质疑声不断,但对于评什么、谁来评、如何评、怎么用,其本身的缺陷、评价的泛化与过频可能会产生哪些负面效应?对于这些根本性问题的反思有待深化。必须看到,这种围绕效率与绩效的制度建构,在激发教师科研潜能与产出数量的同时,又会引发其对结果预期不可控的担忧。评价结果一旦关乎教师的生存与发展乃至成为个人地位获致的学术资本与政治权力,过于激烈的竞争性评价不仅会减损教师的自我效能和职业角色认同,还可能影响其学术自信乃至职业生涯的发展方向,这种负面性往往会经由社交网络、信息平台与社会舆论的扩散发酵而形成放大效应。目前不在少数的优秀人才流向非学术领域,如目前掀起社会广泛关注与热议的名校教师或博士纷纷到街道办事处、中小学等机构就职等现象,背后折射出的是学术职业领域吸引力的急遽下滑、学人对学术职业认同感与归属感的危机。

20世纪70年代以来,美国大学已尝试建立与发展教师评价制度,但成功的经验并不多见。导致失败的原因包括行政组织对执行教师评价制度缺乏兴趣与信心、教师严重且非理性的抗拒、行政人员无法有效履行评价者角色、评价资料缺乏信度与效度、评价推动前未能有效界定教师职责与评价准则、评价资料未能有效涵盖学生学习成效,等等(苏锦丽,黄曙东,2009)。在英国,埃里克森等人对5888名教师的大规模调查发现,学术群体对大学治理与行政人员的满意度仅为10.54%,而支配和残酷的计量、过度的工作量是不满的首要驱动因素(Erickson, et al., 2021)。国内一项调查中的816名样本教师对评价制度也颇为不满,其中451人表示,受评价标准规制而迫于发表质量不高但有数量加持的论文,158人明确表示考核方式过于急功近利,71人指出考核次数过于频繁,普遍感到工作和考核压力大,希望组织营造宽松的评价氛围(李文平,2017)。可以说,高校教师学术工作受到行政权力的强力制约,评价的自上而下使得教师处于受动地位,组织内部的广泛民主参与和自主程度偏弱,使得教师成为整个评价活动的旁观者而非事实参与者。

大量研究结果表明:如果评价重视教师专业成长则具有正效应,教师的专业成就与职业发展诉求得以透过教师评价获得满足;反之,若只采用绩效问责的评价,将目标任务分解派发、以量化工分及按

分计酬等行为,会让教师产生被外部权力严重宰制之感,难以转化成内在的自觉自律。评价不是目的,而是一种手段。如何将评价取向在绩效问责与专业支持之间给予兼顾,考验着高校行政管理层的智慧与勇气。

四、行动逻辑:调适教师评价目的性与工具性价值间的张力

“破五唯”不只是政府政策的强制性驱动,更为关键的是评价治理体系的系统性变革,因为评价所关联的奖励制度、资源分配、学术资本等利益牵制着学术组织与个人的行为选择。高校是执行政府高等教育评价政策的主体,不同类型与层次高校宜根据指导性文件探索适合本校的评价方式和新标准制定任务,以政府政策变迁激活基层创新,反过来推动宏观政策的完善,这是改革得以成功的保障。

(一) 建立发展性的评价取向

高校内部教师评价制度改革,应确立“专业发展为主,绩效问责为辅”的价值取向,并将两种评价功能有机结合,取长补短,各尽其用,建构推动个人和组织相互促进、共同发展的长效机制。

第一,教师评价应服务于立德树人根本任务。在国家教育评价政策的引导下,高校教师评价应坚持全面考核与突出重点相结合,重视标志性成果的原创性、突破性和实际贡献,同时重点考核教学业绩。要依据教育部《新时代高校教师职业行为十项准则》等政策把师德师风纳入评价方案,形成合理化、可操作的教师学术行为规范细则,还要提升教学和人才培养业绩在职称评定标准中的占比,尤其要重视教师累计教育业绩和综合社会贡献,通过制度创新探索以证据为基础的、可比较的教学业绩评价方式(秦琳,2019)。至于社会服务是科研所带来的外溢效应,教师对服务的投入程度与学校政策有很大关联。如果学校重视服务职能,教师则将投入较多时间与精力,若过度强调,仍有可能因负担过重效果却适得其反(Terpstra, Honoree, 2009),并将重心放在同人才培养基本无关联的单纯的社会服务。因此,确定社会服务的标准、指标和权重须谨慎,尤其要凸显其人才培养功能属性。

第二,切实推进分类评价指标研制。教师群体内部存在结构分化与差异,要尊重教师个体性向、所在生涯阶段特征以及岗位主责主业的差别,不同系列岗位评聘与考核要有明显的区分度,将职级层次、学科属性、专业特长以及个人需求纳入整体考量,为不同岗位的教师提供清晰的职业发展预期,而非迫于外在政策驱动的点缀式或象征性安排。如思政教师职称分类晋升,尽管思政系列教师队伍建设得到前所未有的重视,但仍只停留在口头或文本上,特别是科研成果的单独认定,实际中并没有得到相关政策和职能部门的重视。因此,除了要注重思政教师的工作实绩,还应把在主流媒体、全媒体等平台发表的弘扬正能量、有创见性、点击率高的思想理论文章等作品,切实纳入评价范畴。

第三,要明确不同评价的属性与功能。当前,“破五唯”改革仍难以撼动短平快评价,而且评价泛化及对象错位仍是常态。高校要根据学术职业的特殊属性,实现新进教师、在岗教师以及从事专门研究等不同群体的考评办法和周期的差异化。对于资浅教师,必要的量化评价是其进步的基础和动力;对于成熟型教师,突出高质量成果与实际贡献应是重点。职称晋升与上岗竞聘等重在考察相匹配的业绩与实力;履约考评侧重双方约定任务的完成情况及其责权利,不应混为一谈。至于考核与评价周期,应关注教师专业成长和基于评估研究的未来发展前景,避免对教师进行过多过频的评价,激励其潜心致力于创新与更富挑战性的研究。如对从事长期探索、具有较强正外部性的基础研究,针对性开展长周期评价。此外,要弱化各种评价结果与资源分配的强关联,改变当下普遍存在的按分计酬与资源分配制度,消除教师学术活动中追求短平快的浮躁之风。

第四,评价标准应尊重学术职业品性与教师自身成长需求。受新管理主义思潮的影响,相当一部分高校管理者似乎存在一个误区:教师的压力越大、工作量越大,科研产出总量也就越高,以至于各种评价中科研业绩要求逐年高企,使得教师对“破五唯”政策效应缺乏一种真切实在的感受。因此,评价标准的确定应充分考虑学术职业与学术活动的特点,应是“跳一跳,够得到”的准线,不能为了适应外部竞争而无限拔高指标,导致教师迫于过高的、而自己又无力左右的科研任务,将大量本应属于自我

休闲、家庭与社会生活的时间投入工作,长期处在精神紧绷状态而疲于奔命。近年来,新闻媒体与网络上频仍出现的学人因过度劳累而倒下的事件,足以引起我们对教师评价制度异化状态的深刻反思。

(二) 注重评价过程的对话和协商

公共治理理论在教育领域的普遍应用,推动了全球高等教育管理的民主化进程。人们开始逐渐关注教师评价活动中不同利益相关者的价值诉求,重视评价者、受评者及学校行政领导者的共同合作,诉诸特定程序尊重受评者并赋予自由发表观点的机会。而“集体共识”与“专业支持”是评价活动能否顺利展开、赢得受评者信任的前置条件,因为不同关键利益相关者的立场与主张差异,容易引发价值观、利益冲突与制度障碍。为了解教师群体对“破五唯”政策的态度与体验,有调查结果显示,受访者为系列政策的认知与回应平平,无论是中立的适应者还是悲观的怀疑者,都对这一评价政策及其实施产生冷漠、困惑与不满等负面情绪。更令人诧异的是,他们从未在公众场合公开表达自己的不满(Lu, Zhang, 2021)。这固然体现了文化对个人情绪抑制和适应的影响,也反映出一些高校只是被动实施政策,基于管理层利益而忽视教师应有的学术权利赋予。克列诺夫斯基在分析制度化教师评价的负效应指出,如果教师缺乏参与制度设计的权利,更遑论其对评价的认同和接受。他强调,让教师参与是教师评价制度的重要环节,除了如符合受评者的互动原则,当教师参与整个评价程序时,对评价活动也会采取较为正面的态度(Klenowski, Carnell, 2006)。当然,自主、能动而有效的参与,首先在于教师是一个熟悉评价的目的与价值功能的主体,能够对自身的地位、自我价值和专业能力有所觉察与反思,进而与他在对话协商中就评价标准、结果应用等基本问题达成最大共识。

高校内部评价制度的出台与推动,应厘清评价的背景与目的,建立专业化双向对话平台,召集关键的利益相关者尤其是教师代表作为政策设计者和影响者,根本上改变把征求意见会作为例行程序而无建设性意义的官僚做派,为评价者与受评者双方创设贯穿评价全程的利益表达或申诉救济机制。同时,充分重视教师的学科定位与未来发展意向,尊重其教学与科研活动的相对自主性,让教师主动参与评价活动并利用结果反馈改进专业实践能力,以广泛争取教师对组织化评价的信任支持,缓解高校内部行政和学术之间的持续紧张与冲突。

(三) 促进评价方式的正当合理

现阶段,由于人们在认识论上的理解误区与评价能力不足,以及外部评价规则尚未得到根本性改变,高校行政化操控下的诸多评价举措存在形式大于实质问题,而一个成功的评价所需的基础条件为信度和效度。信度与效度的缺乏是教师评价体系自身固有的问题,是改进评价首要解决的方面。遗憾的是,现实中出于追求评价的效率与简便,它仍未受到应有重视和恰当使用。信度是指评价过程的稳定性和可靠性,不会因为评价者的不同而产生迥异的评价结果;提高信度的途径是综合考察教师的各种业绩表现,重视内容效度的分析与审查。因此,要对管理者、教师等不同评价主体进行系统培训,深入了解各种评价方式的优缺点及适用范围,透过评价理念的更新、评价技能和方法的应用、评价主体间的合作与交流等方面,提升评价专业素养和实践水平,最大限度地减小评价结果的偏误或差异,并增进不同评价者之间的一致性。

与此同时,要提升评价效度。效度是指评价结果为评价目的所能提供的最大贡献度,而采取综合评价覆盖学术成果的不同表现形式是关键。在评价活动中,要结合定性与定量方法进行分析,但实践中教师绩效评价仍注重所谓“客观依据”的关键指标。一份基于20所高校教师对绩效评价的感知与体验的调查指出,受访者中的81.8%表示定量评价占主导地位,但教师最认同的定性评价仅占18.8%(董彦邦,刘莉,2021)。由于教师工作职责涉及多个领域,如普罗杰斯将教学、研究、专业发展、学生指导、大学服务、社区服务、管理、院系服务与发表,纳入其建构的教师评价模型框架(Prodeggers, 1980)。40多年前的教师评价框架尚且如此,时至今日其内涵更是丰富繁杂,为组织化的评价活动带来严峻挑战。对此,可根据评价目的与对象,界定评价内容、评分标准与程序,从受评者、同事、管理

者与学生等多角度搜集教师业绩资料。如通过教师自评、同行评价、学生调查、同事会谈等多种途径,将丰富多样的陈述材料、典型案例以及定量分析数据作为证据,各种证据之间相互检证,发挥相辅补足的效用。当然,这并不是要把综合评价的所有要素一并纳入某种评价活动,而是根据评价目的和特定情境需求的有机组合与取舍。

综合评价的操作难度和成本远大于其理论图景展示,但可规避单一评价工具所造成简单、粗暴与武断的后果,更为贴近学术活动本真,有助于提升评价活动的实效和公信力。2021年国务院办公厅印发《关于完善科技成果评价机制的指导意见》,强调对成果创新水平与影响力的同行评议。但由于承担价值判断职责的专家在评价过程中,有意无意地使用量化指标替代学术成果的实质价值,以及学术权威、行政意志与人情关系在其间的介入,导致教师对现行的同行评议制度颇有质疑。为保证评价过程的客观公正,应设立国家层面的同行专家数据库,教育行政部门制定同行评议操作指南,各高校依据自身实际形成具体实施方案,建立起完善的评议程序与伦理规范,强化同行专家的责任和规则意识,从不同层面增强同行评议制度的规范性与权威性。

(四) 将后设评价作为重要机制

评价作为一个价值判断的过程,如果评价活动本身存在问题,那么其效果也将大打折扣,从而影响评价目标预期与应有功能发挥。充分发挥教师评价的作用,不仅要重视评价前的准备、评价中的执行,还应在评价结束后对评价本身加以反思改进,努力保障评价活动的科学、合理与公正。这种后设评价或者说元评价,是按照一定的理论和价值取向,对评价的方法与技术、结果与应用的整体效果作系统分析,采用数理统计等方法来预判其偏差对评价结果的可能影响,进而提出具体的改进建议,作为下一轮评价活动设计的循证依据。

评价过程中难免会受到评价主体、标准、方法与工具,以及评价者自身的专业程度、经验或认知偏好等因素影响,为评价的预期目标实现造成障碍。有研究者通过教师对评价政策认同程度的调查发现,理想的政策与现实情况之间存在较大落差,政策内容认同与执行认同差距的平均值为45.7%,主要体现在教师对评价程序、评价指标、评价标准、评价结果等方面(刘莉,等,2020)。在这个意义上,实践中并不存在一种最优的教师评价系统,但基于反思立场寻求次优方案是一务实选择,而尊重学术价值和教师的尊严则是反思的逻辑基点。

后设评价作为评价方案设计的参考依据,应纳入评价活动成为一个核心机制,充分发挥其规范、监督与改进功能。首先,在教师评价工作启动之前,评价组织者要根据后设评价的价值取向,规范评价项目的标准、内容与技术路径,在实施过程中根据设定的程序和方法有机推进。其次,通过对评价活动整体情形的调研与数据分析,将受评者反应、评价质量和最终效果等情况回馈组织者,以把握评价活动本身缺陷与背后机制。最后,透过评价标准与内容、方法与技术、管理与应用等方面的改进,不断提高评价活动的专业化水平。在一轮评价活动结束后实行后设评价,不仅可以对评价取向、技术与伦理等基础性问题给予重新审视,增进评价规划、设计和实施的系统化与可操作性,还可使评价结果与应用更具针对性与信服力,从而推动制度设计的不断成熟完善以及教师学术生涯的稳步递进。

(苏强工作邮箱:xsdsq@163.com; 本文通信作者为蔡晓卫, 邮箱:1364152794@qq.com)

参考文献

- 操太圣. (2019). “五唯”问题: 高校教师评价的后果、根源及解困路向. *大学教育科学*, (1), 27—32.
- 董彦邦, 刘莉. (2021). 高校医学院本部教师最认同什么样的科研评价制度?——基于对我国18所高校医学院本部教师的调查. *复旦教育论坛*, 19(5), 105—112.
- 赖鼎铭. (2005). 量化指标并非学术评鉴的万灵丹: 以国外几种代表性的学术评鉴为例. *图书馆学与资讯科学*, 31(1), 14—33.
- 李文平. (2017). 大学教师对教师评价制度的满意度调查分析. *高校教育管理*, (3), 95—104.
- 刘莉, 朱莉, 刘念才. (2020). 目标群体视角下高校教师科研评价政策认同研究——基于20所“双一流”建设高校的问卷调查. *清华大学教育研究*, 41(1), 100—110.

- 育研究, 41(4), 73—82.
- 秦琳. (2019-10-24). 如何破解高校“唯论文”顽疾. 光明日报, (14).
- 苏锦丽, 黄曙东. (2009). 美国大学教师评鉴制度之探讨. 教育政策论坛, (2), 1—44.
- 习近平. (2020). 习近平谈治国理政. 北京: 外文出版社.
- 阎光才, 岳英. (2012). 高校学术评价过程中的认可机制及其合理性——以经济学领域为个案的实证研究. 教育研究, (10), 75—83.
- 阎光才. (2018). 高水平大学教师本科教学投入及其影响因素分析. 中国高教研究, (11), 22—27.
- Amaranti, M. (2017). Uso de Resultados de la Evaluación Docente para Mejorar la Calidad de la Docencia Universitaria. *Congreso Universidad*, 6(1), 91—114.
- Bourdieu, P. (1984). *Distinction: A Social Critique of the Judgement of Taste*. Cambridge: Harvard University Press.
- Erickson, M., Hanna, P., & Walker, C. (2021). The UK Higher Education Senior Management Survey: A Statactivist Response to Managerialist Governance. *Studies in Higher Education*, 46(11), 2134—2151.
- Hill, H. C. (2009). Fixing Teacher Professional Development. *Phi Delta Kappan*, 90(7), 470—477.
- Klenowski, V., Askew, S., & Carnell, E. (2006). Portfolios for Learning, Assessment and Professional Development in Higher Education. *Assessment and Evaluation in Higher Education*, 31(3), 267—286.
- Laudel, G., & Glaser, J. (2006). Tensions Between Evaluations and Communications Practices. *Journal of Higher Education Policy and Management*, 28(3), 289—295.
- Lu, H., & Zhang, X. R. (2021). Emotional Responses Toward a New Research Policy Among Academics in a Chinese University. *Front Psychol*, 12(11).
- Meng, Q., & Wang, G. (2018). A Research on Sources of University Faculty Occupational Stress: A Chinese Case Study. *Psychology Research and Behaviour Management Journal*, 11, 597—605.
- Prodegers, S. B. (1980). Toward Systematic Faculty Evaluation. *Regional Spotlight*, 13(2), 1—19.
- Terpstra, D. E., & Honoree, A. L. (2009). The Effect of Different Teaching, Research, and Service Emphases on Individual and Organizational Outcomes in Higher Education Institutions. *Journal of Education for Business*, 84(3), 169—176.

(责任编辑 范笑仙)

Efficiency or Value: Logic Dilemmas and Tension Adjustments of University Faculty Evaluation System

Su Qiang¹ Cai Xiaowei²

(1. Zhejiang Institute of Higher Education, Hangzhou Dianzi University, Hangzhou 310018, China;

2. School of Marxism Zhejiang University, Research Center for Marxism Theory Innovation and Communication of Zhejiang University, Hangzhou 310058, China)

Abstract: In recent years, the prevalent evaluation activities in university faculties have always been dominated by scientific research orientation, quantitative method, and performance orientation, reflecting the phenomenon that only emphasizes the rationalities at the factual level and ignores the legitimacy at the value level. The evaluation orientation and standards of university faculties are not mainly derived from the self-consistent theoretical logic, but from the strong interventions of external forces in the academic ecology of universities and the results of competitions among universities in reality, which in turn shapes faculties work life styles, academic beliefs, and action preferences. Therefore, it is necessary to comprehensively reflect on the result logic, counting logic, and performance logic of the current evaluation system at the value level, so as to adjust the tension between the purpose value and instrumental value of faculty evaluation.

Keywords: university faculties; faculty evaluation; evaluation system; organizational reform; teacher behavior