

科研量化评估的历史建构 及其对大学教师学术工作的影响

沈文钦 毛 丹 蔺亚琼*

【摘 要】 20 世纪 90 年代中后期以来,随着重点大学建设工程的推进以及大学教师的专业化转型,量化评估成为评价大学教师的一种重要手段。基于历史的梳理,本文将表明量化评估如何被历史地建构成为一种合法化的评价手段。同时,基于对 8 所学校 36 位教师和行政管理者的访谈,我们分析了量化评估对大学教师学术生活的影响。文章指出,就积极意义而言,量化评估技术将贤能主义引入大学教师评估领域,推动了大学学术工作的专业化水平,并使得教师的职称评审等避免人情文化的影响,但这一评估技术也带来了一些意图之外的负面影响,如盲目追求数量而非质量、形式化的知识生产、教师科研的自主性受到侵害、应用导向而非发表导向的学科受到挤压,等等。

【关键词】 大学教师;量化评估;同行评议;学术工作

一、 导论

我们能观察到,量化在全世界范围内成为一种管理工具。政府部门很善于利用各种统计和数字来进行管理。^①在高等教育内部,对学者及其科学研究活动进行量化管理在全世界都有上升的趋势。^②在英国,大学教师越来越成为绩效考核(performance appraisal)的对象,尽管绩效评估与大学内在的制度逻辑是相冲突的。^③在澳大利亚,对大学教师的科研绩效评价已成为一项常规性制度,某大学根据申请外部资金情况、指导博士生情况和论文发表情况三项指标对教师进行考核,^④某历史系教师在申

*沈文钦,教育学博士,北京大学教育学院副教授,100871;毛丹,教育学博士,上海师范大学教育学院讲师,200000;蔺亚琼,教育学博士,华中科技大学教育科学研究院讲师,430074。

①C. Shore and S. Wright, "Governing by numbers: Audit culture, rankings and the new world order", *Social Anthropology*, vol.23, 2015, pp.22—28 以及 J. Z. Muller, *The Tyranny of Metrics*, Princeton: Princeton University Press, 2018.

②D. Brenneis, C. Shore and S. Wright, "Getting the measure of academia: Universities and the politics of accountability", *Anthropology in Action*, vol.12, 2005, pp.1—10.

③B. Townley, "The institutional logic of performance appraisal", *Organization Studies*, vol.18, 1997, pp.261—285.

④R. Whitley, J. Gläser and L. Engwall, *Reconfiguring Knowledge Production: Changing Authority Relationships on the Sciences and Their Consequences for Intellectual Innovation*, Oxford: Oxford University Press, 2010, p.170.

请晋升时,甚至需报告自己著作的被引情况。^①在欧洲大陆,对大学教师科研工作的量化管理也在加强,教授自主权一向很大的德国也未能幸免^②,从法国的情况来看,对教师的评价因高校类型的差异也存在不同,但同样也存在通过期刊分级、文章引用次数来评价教师科研的情况。^③挪威大学教师的科研产出呈上升趋势,而这一趋势与大学的激励政策有关。^④

相对而言,美国研究型大学对教师的科研考核较少采用量化模式。顶尖的研究型大学如伯克利和麻省理工学院在教师晋升和招聘新教师时都没有明确的量化标准和文章数量要求,评估报告中学者的被引数据仅作为参考。^⑤当然,即便在美国,学术晋升也越来越注重文章发表的数量。很多年轻教师被同事告知,在职称晋升时,最重要的因素是文章发表的数量。^⑥对美国社会学学者学术产出的分析显示,社会学学者的学术产出呈现激增趋势,其背后的一个原因是院校鼓励学者发表,并且多发表可以使工资收入获得增加。^⑦

最近20多年来,中国学者的科研论文呈急剧增长的趋势^⑧,但科研产出表现出数量优于质量的现象。即便是和香港、台湾地区的顶尖大学相比,中国大陆研究型大学的质量也相对更低。^⑨本文认为,中国学者论文发表数量的剧增,其原因之一是为了应对院校对其科研工作的量化考核。量化评估或审计文化(audit culture)对中国大学教师的影响,已经引起一些学者的关注^⑩,有学者分析了全球大学排名或国家科研评估制度对大学教师的影响^⑪,另有学者分析了中国大学教师对新管理主义改革的回应。^⑫

但总的来说,已有研究并未专门针对量化评估及其影响进行深入分析,存在几个有待填补的空白。第一,已有研究更多关注全球和国家层面的政策影响^⑬,对中国不同高校之间的差异关注不够;第二,已有研究注意到了科研评估制度对中国大学教师的影响,但并未特别关注一种特殊的评估技术即定量评估的影响。有鉴于此,本文通过对8所大学36名大学教师的访谈,结合相关政策文本的分析,对中国教师科研评估中的量化管理现象及其对大学教师学术工作的影响进行探讨。

为相对平衡地反映量化评估对不同学科教师的影响,我们分别访谈了工程科学、物理学、化学、中文、历史、社会学和历史学这几个学科的学者。访谈材料涵盖2004年、2007年、2011年、2015年等四

①J. Gläser and L. Laudel, "The social construction of bibliometric evaluations", in R. Whitley and J. Gläser (eds.), *The Changing Governance of the Sciences*, Dordrecht: Springer, 2007, pp.101—123.

②J. Frost and J. Brockmann, "When qualitative productivity is equated with quantitative productivity: Scholars caught in a performance paradox", *Zeitschrift für Erziehungswissenschaft*, vol.17, 2014, pp.25—45.

③C. Paradeise and J. C. Thoenig, "Evaluating Faculty Contributions", in C. Paradeise and J. Thoenig (eds.), *In Search of Academic Quality*, London: Palgrave Macmillan, 2015, pp.83—107.

④S. Kyvik and D. W. Aksnes, "Explaining the increase in publication productivity among academic staff: A generational perspective", *Studies in Higher Education*, vol. 40, 2015, pp.1438—1453.

⑤J. C. Thoenig and C. Paradeise, "Organizational governance and the production of academic quality: Lessons from two top US research universities", *Minerva*, vol.52, 2014, pp.381—417.

⑥M. S. Anderson, E. A. Ronning and R. D. Vries, et al., "Extending the Mertonian norms: Scientists' subscription to norms of research", *The Journal of Higher Education*, vol.81, 2010, pp. 366—393.

⑦J. C. Hermanowicz, "The proliferation of publishing: Economic rationality and ritualized productivity in a neoliberal era", *The American Sociologist*, vol.47, 2016, pp.174—191.

⑧W. Liu, G. Hu, L. Tang and Y. Wang, "China's global growth in social science research: Uncovering evidence from bibliometric analyses of SSCI publications (1978—2013)", *Journal of Informetrics*, vol. 9, 2015, pp. 555—569.

⑨F. Li, Y. Yi and X. Guo, et al., "Performance evaluation of research universities in Mainland China, Hong Kong and Taiwan: Based on a two-dimensional approach", *Scientometrics*, vol.90, 2012, pp.531—542.

⑩L. Yi, "Auditing Chinese higher education? The perspectives of returnee scholars in an elite university", *International Journal of Educational Development*, vol.31, 2011, pp.505—514.

⑪J. Li, "The global ranking regime and the reconfiguration of higher education: Comparative case studies on research assessment exercises in China, Hong Kong, and Japan", *Higher Education Policy*, vol. 29, 2016, pp.473—493.

⑫Y. Huang, S. K. Pang and S. Yu, "Academic identities and university faculty responses to new managerialist reforms: Experiences from China", *Studies in Higher Education*, vol.43, 2018, pp. 154—172.

⑬J. Li, "The global ranking regime and the reconfiguration of higher education: Comparative case studies on research assessment exercises in China, Hong Kong, and Japan".

个时间段,使得我们可以考察历史性的持续和变化。此外,本文还收集了10所研究型大学有关招聘、科研奖励、职称晋升、工作量考核相关的政策文本,以及75所教育部直属高校的“十二五发展规划”和“十三五发展规划”文本,以了解科研量化评估在各高校政策文本中的体现。

本文结构如下:首先,根据对相关文本的梳理,我们将展示量化评估作为一种管理技术是如何历史性地进入高等教育场域并获得合法性的。其次,通过对各个高校政策文本的分析,我们将展示量化指标评估是如何体现在高校的不同政策之中的。最后,我们将分析国家以及院校的量化评估政策是如何影响高校教师的学术生活尤其是他们的知识生产行为的。

二、赶超发展遇上大学排名:科研量化评估的历史建构

20世纪80年代,中国大学的教师评价并不盛行量化评估的方法。当时的副教授、教授职称评审,一方面看资历,另一方面也主要看学者在学科领域的声誉和影响。换言之,此时同行评议发挥比较重要的作用。例如,在20世纪80年代,北京大学的教授评审需要在学校的学术委员会进行汇报、答辩(对北京大学某学校学术委员会成员的访谈,2012年)。

从20世纪80年代末开始,量化评估的方法首先在自然科学领域被逐渐引入进来。1987年,根据国家科委新技术局的要求,中国科技情报研究所采用文献计量学的方法对美国的SCI、ISR以及ISTP三种检索系统1983—1986年间收录的中国科技论文做了统计分析。1988年,中国科技情报研究所的尚忆初在《我国在世界所处的学术地位》一文正式发布了我国高等学校1983—1986年间发表论文前10位的情况。^①从1988年开始,教育部科技司开始定期出版《高等学校科技统计资料汇编》,该出版物统计了各高校的研发人员、研究经费和论文发表数量情况,为各高校之间科技论文数量的比较提供了参照。当时教师评价并未采用量化评估手段,大学教师的学术产出也相对较低,当年北京大学统计的科研活动人员是2412人,论文发表总数是1299篇,人均0.54篇,清华大学科技活动人员5768人,发表论文1293篇,人均0.22篇。^②

1989年,中国科技情报所受国家科委综合计划司委托,对我国1988年中文期刊上发表的科技论文进行统计与分析。此后,对三大国际检索论文的年度统计分析成为一项常规化的工作。1990年,中国政府对国家重点实验室进行评估,中国科学院紧接着对其下属的研究所开展了评估,当时的评估就很重视文章的数量。1992年,南京大学首次超过北京大学,成为国内发表SCI论文最多的高校。^③1998年10月23日,国家科技部举行新闻发布会,正式公布“1997年中国科技论文统计结果”,由于该排名结果由科技部发布,各高校领导层高度关注。在此之后,SCI的排名在国内高校影响越来越大,在重视SCI论文的政策导向下,一些传统的工科院校也开始重视SCI论文数量。

1998年,中央政府启动了985工程。2002年教育部正式启动了一级学科排名,学科排名的评价指标当中很多是量化指标。同年也就是2002年,浙江大学颁布《浙江大学研究生学位论文答辩有关规定》(浙大发研[2002]16号),规定博士生必须发表一定数量的论文才能参加论文答辩。此后,北京大学在2003年也采纳了这一政策,其他高校纷纷跟进。

2003年,上海交通大学向全世界公布了第一份世界大学排名,这个排名对高校也有很大影响。与此同时,一些民间的大学排名如武书连大学排名对高校也造成了越来越大的压力。在各种各样大学排

①尚忆初:《我国在世界所处的学术地位》,《中国科学基金》1988年第12期。

②中华人民共和国国家教育委员会科技司编:《1988年高等学校科技统计资料汇编》,武汉:武汉大学出版社,1989,第218—219页。

③中国科技信息研究所《中国科技论文统计与分析》课题组:《1993年中国科技论文统计分析》,《科学》1995年第2期。

名的推动下,各个高校越来越重视科技论文的数量。一位教师在2004年的访谈中指出,推动学校重视文章发表数量的一个背后推手是大学排名:

“这个不好弄,因为现在学校的排名都靠这个东西,这样的话,校领导们对这些都很紧张的,所以对学生们的要求也很高。……从学校而言,总是觉得排名的压力很大,跟北京上海的学校比起来,还是有差距的。所以学校要求要发文,这样一要求果然有效,排名就上去了。”(2004年访谈,某985大学教授)

随着世界一流大学建设工程的推进,SCI论文发表数成为学校领导者所追求的一个重要指标。一些高校在总结其办学成绩时,SCI论文和EI论文数在全国高校中的排名成为一个重要的指标。在此背景下,文章的数量、文章的期刊影响因子、文章的被引次数被广泛地应用于教师的评价(如年度业绩考核、职称晋升)当中。不仅如此,量化的指标甚至成为一些重要学术奖励的标准。很多学术奖励如教育部高校优秀成果奖、长江学者的评选均要求学者提供论文被引次数的数据。

三、高校量化评估制度与政策分析

量化评估越来越成为各个高校进行教师科研绩效管理的重要工具,量化考核的管理方式渗透到高校教师绩效评估的各个方面,包括学校发展规划、绩效考核、科研奖励、职称评定等,形成了一个以正式政策文本为基础的制度化体系。

1. 学校发展规划

从建国初期开始,我国高校就开始引入规划管理制度,发展规划既是国家宏观控制的工具,也是高校自主治理的重要方式^①,是一个学校未来若干年发展的指南针,对高校各项管理制度与政策都有着重要影响。笔者对教育部直属的75所高校的“十二五”和“十三五”学校发展规划文本进行分析发现,各校都将科研发表的数量、科研经费的数量纳入了“计划”范畴。某东部地区的211高校在其“十三五”发展规划中指出:“到‘十三五’末,力争具有高级职称的专任教师人均科研经费达到50万元/年,学校科研到位总经费达到30亿元,……力争实现国家自然科学奖的突破,新增教育部人文社科成果奖励2项。新增高水平(IF≥9)SCI论文40篇以上,累计发表SCI/EI论文7000篇、SSCI论文100篇、CSSCI论文1800篇。”

随着世界一流大学建设的推进,科研成果作为各项评估和排名的关键指标,成为提升高校排名和声誉的重要手段。一些高校在制定“十三五”规划时不仅对科研经费和科研发表有明确的量化指标,而且对指标的增长率提出了非常高的目标。如某财经类大学十二五期间SCI年均发表量为75篇,SSCI年均发表量为90篇,而“十三五”期间计划SCI和SSCI年均发表量各超过200篇,增幅超过100%。

为了完成“十三五”规划制定的科研绩效指标,一些学校开始采用关键绩效指标(Key Performance Indicator)对学校进行规划和管理,并将指标层层下放,落实到各个院系。某中部地区211高校在《关于拟定本单位2017年度目标任务工作的通知》中提到,科研工作的主要绩效指标主要包括:国家重大项目(如国家自然科学基金重大重点项目等)、政府重大奖励(如国家科技奖等)、高水平论文(如SCI、SSCI论文等),各院系在拟定年度目标时要注意“目标任务是可量化的,可考核的,见效果,有权重。”在不少高校,科研发表的数量也成为学校在年终考核各个院系绩效的核心指标。从各高校发展规划政策文本可以看出,学术论文的数量仍然是不少高校追逐的主要目标,而论文的质量则主要依靠论文所

^①齐明明、陈廷柱:《我国高校发展规划执行阻滞探源及其破解之道——基于合法性理论视角的分析》,《高等工程教育研究》2016年第4期。

发表的期刊是否收录于几大索引(如SCI、EI、SSCI、CSSCI)来衡量,简单用刊物影响力来衡量论文的影响力。

从上文可以看出,面对世界一流大学建设和各种排名评估的压力,许多高校都对科研发表提出了量化的目标,而这些攻关目标要落实到教师个人,则需要一系列微观的治理策略。这些策略既有职称晋升、资格控制、年度考核等制度上的规范性要求,又有针对科研发表和科研项目的物质奖励,在学校形成了一个细密的量化评估之网。下面本文将对这些微观的治理策略进行简要介绍。

2. 职称评定

对于高校教师来说,职称晋升是事关职业发展和切身利益的大事,也是各高校用来激励教师提高科研产出的有力途径。目前,大部分高校在职称评定上都规定了基本的量化科研成果指标,只有达到这个指标,才具有提交职称评定申请的资格。有些学校的职称评定规定中虽没有明确推行量化评估管理,但在实际操作中,发表文章数量更多的教师往往更占优势。

近年来,不少高校对于职称晋升制度进行了改革,对于职称评定中的科研发表要求做出了分类的规定。一是学科分类,理工科与人文社会科学的评价标准有所区分,但通常情况下高校的评价标准只细化到大类(如自然科学、工程与信息科学),还未细化到具体的每门学科。二是岗位分类,很多高校开始将教授、副教授岗位分为教学科研并重型、科研为主型和教学为主型。但是不管是哪一类别,都有着科研成果的量化硬指标。以某西部985高校为例,要申请正高职称,需满足如下要求:(1)主持国家级科研项目1项,或主持的科研项目累计到款200万(人文社科为40万)。(2)发表3篇SCI,SSCI论文(其中权威期刊论文2篇),5篇重要期刊论文。人文社科要求其中一篇为SCI,SSCI论文,或3篇权威期刊论文。(3)获得国家或省部级科研、教学成果奖。若无获奖,则需增加SCI,SSCI论文3篇(人文社科为SCI,SSCI论文1篇或权威期刊论文3篇)。

虽然在绝大部分学校的职称评审流程中,外审专家评审都是一个重要的环节,但是由于相关制度的不完善,外审专家的评议结果通常情况下都只是作为参考,科研成果的数量相比而言更为重要。从高校的角度来看,高校在职称评定中设置量化指标,一方面是因为其简单易行,另一方面也是为了杜绝职称评定中的人情关系因素,使这个过程更加公正客观,而参与表决的学术委员会成员也可以因此免责。

3. 年度考核

年度考核是激励教师提高科研产出的另一项策略。例如有些高校规定,教师每年要完成一定的“工分”,包括科研工分 and 教学工分。在科研工分的计算中,不同级别的刊物有不同的分数。如某985高校规定,每个教师得到的科研分数除以均值,如果小于0.5,则考核不合格,会影响到各项收入。而如果该值大于1,则为优秀。这个政策在院系的各个教师之间形成了一种激烈的竞争氛围,因为即使发表的绝对数量不少,但若低于全院的平均水平,仍有被评为不合格的风险。某985高校一位计算机学科的教师在接受访谈时向我们介绍,他们每年有2500分的科研工作量要求,每10万元课题经费可以折算成1000分(2010年访谈,计算机系博士生导师)。在有些高校,有些教师无法申请到课题,但为了完成科研经费的要求,不得不从校外找一个合作对象(如公司)来获得横向课题,实际上是自己支付课题经费(2011年访谈,某地方高校经济学教师)。

4. 科研奖励

为了提高教师的发表量,很多高校都制定了针对科研发表和科研项目的金钱奖励。一般来说,高校针对科研成果的奖励主要包括如下几个方面:(1)国家级或省部级科研获奖;(2)科研论文发表或专著;(3)国家或省部级科研课题。以西部某985大学为例,根据论文发表刊物的等级,每篇有0.3—30万元的奖励,在Science和Nature发表论文可获得高达30万元的奖励,发表1篇SSCI论文可获得3万

元的奖励。国家级或省部级科技成果获奖有5—200万元的奖励。

对于论文的质量,大部分高校采用期刊的影响力和是否收录于权威论文索引库来进行等级划分,通常分为A类、B类和C类三个等级,不同等级的论文有不同的权重,比如有些高校规定一篇A类论文可以等同于三篇C类论文。

对于理工科论文来说,不少高校将《Science》、《Nature》作为A类,SCIE为B类(有的根据SCIE的分区又划分更多的等级),EI为C类。而对于社会科学而言,不少高校将SSCI收录期刊和学科最权威的1-2本期刊作为A类,学科重要期刊为B类,其他CSSCI论文为C类。刊物的级别决定了奖励的级别。此外,绝大部分高校都要求以上科研获奖或论文必须为第一作者或者独立作者成果,第二作者或者第二署名单位将不能被视为有效成果。

四、量化评估管理对大学教师的影响

如前所述,作为一种资源分配的制度化手段,科研的量化审计裹挟着资金、政策和价值,深入到大学的各个层面。无论是宏观层次的教育部等政府管理部门、中观层次的高校抑或是微观层次的院系和研究者,都成为这套治理体制的行动者。学术人员是其中最为关键的利益相关者,科研审计同他们的日常生活息息相关,其收入津贴的多与寡、学术荣誉的高与低、职称晋升的难与易,将取决于他们在论文数量与质量、经费等级等指标上的数据表现。科研审计与量化评估对中国大学教师的学术工作影响深远,引发了一系列或积极或负面的影响。科研的量化评估一方面促进了中国学术工作的专业化程度,但另一方面也引发了“形式化科研”、重数量轻质量等意外后果。

(一) 科研审计与学术工作的专业化

依托于量化评估强大的分配功能和动员能量,中国的学术研究——尤其是自然科学、部分工程科学和社会科学——加速了与国际学界接轨的步伐,并在整体上有效提升了学术职业的专业化程度(professionalization)。这里的专业化不是指从组织形式上确立学术职业的准入门槛(如博士学位)、成立专业性的学会,而是指研究范式、方法技术、理论概念、写作规范等知识创造的多个环节开始大范围且较深程度地受到国际学术界规范的影响。

科研量化指标的出现和广泛运用伴随着中国大学的重点建设,创建世界一流是其重要目标,作为国家意图落实的政策工具,科研评估有效地引导着学术研究与世界接轨。新中国建国后科研事业经历了误区和弯路,对改革开放之后我国科研人员从事研究的专业水准和伦理带来了负面影响。发端于上世纪80年代末的量化评估,在学术界引入了前所未有的学术竞争。出于发表国际论文的压力,以理工科为主的学者不得不有意识地提升研究的专业水平,以满足国际发表的要求。

学术工作专业化的一个重要方面是学术评价中的“贤能主义”(meritocracy)和“普遍主义”开始确立,在相当大的程度上改变了以往职称晋升、学术评比中的模糊性,由此研究者的权力关系和工作焦点也发生了相应的学术转变。科研审计引入了个人主义式的竞争,对于奉行“集体主义”的人事传统进行了改造与转换。此前,学术评价模糊,在同行评议孱弱的体制下,非学术标准极易凌驾于学术标准之上。派系、山头、资历、师承、人际等“特殊主义”因素弥漫于学者的职业生涯,学者需要投入精力来维系学界人情,以换取发展机会。与此相对,量化审计的指标清晰了然,尽管不可避免地存在着缺陷,但它将学术表现凸显到至关重要的位置,搭建了学术锦标赛的竞技场,学术表现成为学术工作的焦点。何人晋升,谁受褒奖,都确立了普遍遵循的明晰标准,尽可能排除人际干扰。至少,在制度运行的主导逻辑下,学界分层的规则开始朝向基于学术表现的方向移动。这种体制鼓励学术勤奋,

鼓励学者将学术目光投注于更大的学界,而不必陷入工作院系内的学术政治。在同行评议制度被扭曲的情况下,定量评价成为一种形式上公正的制度。一位985校长指出,在没有实行量化评价制度时,每年评职称的时总是有很多人打招呼,实行定量评价后,打招呼的人少了,所以他认为不能贸然取消量化评估制度(某985高校校长,2014年)。

尽管学界对量化审计颇多微词,但不少研究者意识到数量评估中承载的普遍性原则对研究者的保护效应。在学术流动初露端倪的中国学术界,这种基于学术的普遍主义有利于破除长期近亲繁殖和人际庇护对于学术研究造成的压制与不公。一位文科青年教师到一所较为封闭的院系任职,院里同事多为该校毕业生,他指出:“我是一个外来人,我埋头做好自己的研究,教好课,和同事维持和谐的关系,就足够了,不必要处心积虑地讨好(领导和同事)。分奖金、评职称、申帽子,需要成果过得去,否则就算有人背后力挺,他们也没有办法。一个外来人要扎根下来,就是要不断地发表。”另一位教师追忆起她毕业离校时,导师对她的教导:“你们那个地方(指工作地),运作的东西比较复杂,这种情况下你就埋头写文章,多发表,暂时摒弃完美标准,只要心中知道哪里有缺陷就够了。这是一种对你的保护……现在学术竞争加剧了,他们总需要推一些能干的人出来”。

不过,基于学术绩效的个人主义如何改造着学术职业的集体主义,普遍原则如何校正着特殊原则,还需要进行更复杂的理解,也需要更长的时段和整体的制度配套才能完成。在以团队的方式进行知识创造的领域,许多学者不能拥有自己独立的实验室,一个大的科研团队常常包含数位甚至十多位教师。他们生涯发展所需的资源、机会、平台首先是在团队内部进行分配,这关系着学者个体的学术绩效,他们需要平衡学术能力与人际间的权力关系。量化评估引入的个体竞争与传统的集体主义文化,二者并非简单的替代和此消彼长的关系。

随着学术竞争加剧与普遍主义原则的强化,科研活动占据了学者大量的时间和精力。在量化评估盛行之前,学术界全神贯注投身于科研者更多出于内在的激情和高度自律,尚未出现深度奖惩的制度将学者导向科研,对于大多数学者而言,学术压力的轻重,研究时间的短长,多为学者自我选择的结果。时至今日,情景大不相同。越来越多的学者在学术工作上投入愈来愈多的时间,尤其在研究型大学,越来越多的学者专注于学术工作,这在微观和日常的层面映射着研究专业化这一宏观趋势。“只要有时间,我就投入到工作中。外界对我们有很多误解,以为大学老师最轻松,一周随便上上课,其余都是休息时间,还有寒暑假,我都不晓得该怎么向外行解释。平时都不会休息太久,一有空就在琢磨研究。”上述受访者对工作日常的叙述在研究型大学中颇为普遍。

总体而言,在国家一整套大学重点建设的政策与制度中,针对学者学术产出的量化评估是最普遍而且渗透性最强的治理术,它深度重构了学者的学术工作,并推动了学术工作的专业化程度。

(二) 科研审计的意外风险:研究的形式化

在促进专业化的同时,科研的量化评估为研究者构筑起理性化的“铁笼”,这也引发了一系列负面的非意图后果,其中最为微妙但又最具风险的便是“研究的形式化”。概而言之,“研究的形式化”指科研的导向从认知转向指标,研究首先不是为了解决重要的理论或实践问题,而是为了发表论文、申请课题、获得体制的奖励——指标数据宰制了科研。许多研究者和体制的评价者很少关注研究内容本身,而更多关心研究能否带来数据的增长:论文的数量、课题的数量、获奖的数量,等等。科研沦为形式性的工作,其内容和实质意义被抽空,内核不再重要,研究从知识产生的情境中脱域。“为学术而学术”的科研观,在科研审计的奖惩下,转变为“为论文而论文”,“论文学者”、“课题学者”等戏称应运而生。

由于每个教师都面临着完成考核任务的要求,因此写作论文、申请课题甚至指导学生都被教师们

认为是“完成工分”：“有的老师凑公分所需要，于是必须抓住学生写一点东西，学生人头也算工分，所以我觉得总的来说就是导师为了自我的发展，为了自我的生存，甚至为了自我挣钱，来招学生”（T5，工程学）。而且，这种计件制的绩效文化在中国有深厚的文化基础，具有很高的合法性。^①

职称晋升是学者学术生涯当中最为重要的关键事件，因此很多教师为了实现晋升，不得不屈从于量化评估的体制。许多研究型大学的受访者已然观察到“大家都在拼命，而且都在赶工，不是说真正的做出质量来，（而是）我要把指标做上去。”

在职称晋升这一重要环节，“它（指评审制度与评审专家）看文章的数量，它不看到底解决什么问题，这件事就比较麻烦，这种评价导向影响是非常大的”。（2012年访谈，物理学）在一些领域，例如生物领域，“科研的泡沫特别大，就是发表了论文90%是没有任何科学意义的，只是为了发论文而发论文。现在的评价体系就导致人们为了发论文，可能就真心地去想解决一个科学问题的人太少了，不是没有，但是太少了。”

量化审计当然也重视奖励研究的质量，不过质量这种兼具主观性的评价，也被简单地操作化为客观的指标——文章发表的期刊等级及其声誉。CNS(Cell、Nature、Science)级别的一篇文章，动辄奖励数十万，这背后实乃是“重赏之下必有勇夫”的物质激励逻辑。体制着重奖励的是在CNS、权威期刊发表论文的行为和结果，而非研究的学术价值与社会意义。一类体制宠儿便如此诞生：凭借CNS的加持，获得巨额的科研经费，购买昂贵的设备，聘请更多的人手，进而继续发表更多的CNS。研究者个体名利双收，在官僚体制的科研审计下如鱼得水，至于他们是否真正做出了极有价值的研究，是否收获了国际科学界最高级别的学术赞誉，人们并不关心。

科研审计不仅细致地制定了各项指标的兑换值（如奖金额度、晋升机会），同时严格限定着科研产出兑换利益的有效时间，其典型体现为评估的时间段：一年一小考的年度考、三年一大考的聘期考、各种表格中的“近三年”或“近五年”、多种基金评审和人才认定中的年龄限制，这些时间管理的技术营造出“争分夺秒”的紧迫与“时不我待”的恐慌。

面对职业生涯发展和收入津贴的压力，许多研究者不得不在“多的成果”与“好的研究”之间做出抉择，向前者妥协成为审计文化下的生存理性。“十年磨一剑”在当今的制度下意味着极大的职业风险，“不发表则出局”是所有研究机构处于学术职业生涯晋升期学者的命运。“人家是三年磨一剑，可能还没磨出什么东西，就让他走人了”，“（如果）你十年磨一剑，学生也不来了，没有学生愿意陪你磨。”（复旦大学某教授访谈，2011年）

相比此前学术评价的模糊化和弱利益，科研审计充分调动起研究者的工作投入，“越多越好”随之成为学者在激烈竞争下面对不确定性的行动策略。当科研评估简化为计数工作之后，对于多数人而言，“量多者胜”成为第一生存法则，第二法则才是“质优者胜”，量多是前提，质优是亮点，有量又有质，才可能在学术资源更为激烈的竞争中胜出。在科研审计政策的“引导”下，不少研究者在知识生产方面随之出现一系列相应的策略。访谈资料显示，一些研究者可能选择容易在短时期内出成果、发论文的题目，耗时长、沉潜性的研究不受青睐，或者暂不顾及。“本来他可以做很高水平的文章但是他做的不多，搞高水平的文章需要很大的精力，如果有的教员讲究实际，为了评职称，他就追求发大量的意义不大的文章。”

论文作者数量的增加也受到了科研量化审计的驱动，合作能够提升研究者的产出数量、影响力或者“工分”。一些研究者采取了“稀释”的策略，可以浓缩为一篇学术精品的论文拆分为数篇论文，“一

^①A. B. Kipnis, “Audit cultures: Neoliberal governmentality, socialist legacy, or technologies of governing?”, *American Ethnologist*, vol. 35, 2008, pp.275—289.

杯浓茶稀释为几杯淡茶,每篇也有创新或见解,数量就上去了”。

研究界是高度分化的世界。社会情境对于知识的实质内容施加着怎样的限制,存在着较强的学科差异,因而“科研形式化”在不同学科的语境下会有不同的表现。譬如,相比于更容易跨越地域差异而在国际范围内拥有一致目标的自然科学,工程研究的价值则相对更需要结合当地的经济、社会和文化进行评判。解决实践问题、促进产学研结合与转换,是许多工科领域重要的目标。目前科研量化审计的标准主要参考了部分自然科学的标准。对于其他领域,科研审计显示着行政系统“无知的傲慢”,指标和标准的确定忽略了某些领域或机构的特征。然而,职称晋升和项目申请过程中对CNS和权威期刊论文、影响因子、被引频次等指标的倚重,对理工科的应用研究形成了挤压之势。一位重要工科大学的教授谈到,当整个学校卯足了劲去比拼SCI论文数及被引指标之后,面临着“应用学科被搞掉”的风险。这所大学也素有承担军工项目的传统,军工项目因其特殊性和机密性,更多以委托的信任机制进行经费拨款,极少采用课题申报的竞争机制,也很少纳入“纵向课题”的统计口径。但目前“科研项目 and 经费”的评比中,来自国家的纵向经费才有份量,来自国家部委或其它部门的委托课题渐受冷落。随着科研量化审计的盛行,这所学校的军工课题也呈现出衰减之势,研究者都或主动或被迫地被吸入体制的选拔之中,那些难以被量化、可视化的研究面临着日渐边缘化的处境。

五、结论与讨论

如前所述,量化评估的技术是在20世纪80年代末之后被逐渐引入中国大学的。随着学科评估、世界大学排名等评估逐渐渗入高等教育场域并深入人心,以文章发表数量、论文被引次数为导向的量化评估技术也获得了巨大的合法性。通过对高校政策文本的分析可以发现,量化评估技术已经被广泛采用于教师招聘、教师晋升、奖励、教师考核等不同环节。

除了外部评估的推动之外,学术共同体发育不良、同行评议文化没有彻底制度化也是量化评估得以肆意扩张的重要原因。中国在期刊论文发表、课题评审、学术荣誉评审、教师职称评定等方面,已比较普遍使用同行评议的手段。但在很多情况下,同行评议的原则没有得到彻底的贯彻。在期刊论文发表方面,很多期刊仍然没有采取严格的同行评议制度,论文发表中的人情因素仍然存在。在课题评审和学术荣誉评定中,拉关系、托人情的现象普遍存在。在同行评议文化缺失的情况下,大学乃至学者均通过文章发表的数量来提升显示度和声誉,科研计数成为普遍执行的法则。

对于科研的量化评估,之前学术界几乎是一边倒地加以批判。和这些观点不同,本文认为,由于中国高等教育在20世纪70年代末之后才开始重新起步,整个大学教师群体面临着走向专业化的历史要求,量化评估的技术在其中发挥了一定的积极作用。科研的量化审计犹如一把双刃剑,一方面,在学术指标式治理的驱动下,中国的学术工作发生了积极的变化,专业化水平在提升,科研在学术工作中日益占据重要的位置;但另一方面,在学术自主性较弱和同行评议发育不良的学术空间内,量化评估也引发了以“科研形式化”为典型的系统风险。学术工作中多样化选择(如“十年沉潜”)的空间被日渐压缩,大学教师的科研自主空间受到侵害。过于强调论文发表的数量和效率,使得一些基础性、研究周期较长、风险较高的研究可能得不到应有的重视。而且,我国高校中教师科研的量化评估是一种典型的“指令性绩效管理”(directive performance management),根据已有研究,这一管理方式会降低教师的士气,给他们带来挫折感而非积极的效果。^①

^①M. Franco-Santos and N. Doherty, “Performance management and well-being: A close look at the changing nature of the UK higher education workplace”, *The International Journal of Human Resource Management*, vol.28, 2017, pp.2319—2350.

量化评估技术所存在的问题,已经被越来越多的学者所意识到^①,一些高校也出台了改革措施,如吉林大学在其“十三五”规划中提到:“在哲学社会科学领域探索实施量化考核与非量化考核相结合的评价办法,引入‘代表作评价制度’。对于经过长期潜心研究产生的重大成果给予追溯性、补偿性的奖励和评价。”复旦大学等学校也在几年前开始实行“代表作评价制度”。但是这些改革尝试仍然面临着很大的困难,最主要的困难是非量化评价所依赖的同行评议制度还存在诸多问题。同行评议程序复杂、成本高昂,而且需要学界同仁的大力付出。此外,由于学术系统内部的等级制,使得少数精英主宰了学术资源和资源分配的过程,政治、关系资源或社会资本常常在其中起作用,伤害了同行评议的公正性^②,而同行评议机制自身存在的自主性弱、容易受行政力量干扰、资本不足等问题也使得同行评价制度存在重重困境。^③

总的来说,由于目前高校的评价仍然为外界所主导,而外界的评价更多是基于量化的指标之上。尤其是在新的双一流大学建设的浪潮之下,量化指标存在着愈演愈烈的风险。中国学者如何面对这一量化评估体制并做出回应,还需要进一步观察。

(责任编辑:蒋永华)

Development of Quantitative Evaluation of Scientific Research and Its Influence on the Academic Work of University Teachers

SHEN Wen-qin, MAO Dan, LIN Ya-qiong

Abstracts: Since the 1990s, along with the development of different projects aiming to build key universities (e.g. 211 project and 985 project) and the professional transformation of university teachers, quantitative evaluation with the aid of bibliometric techniques has become an important means to evaluate university teachers' research performance in China. Based on an analysis of the key policy texts of different periods, this paper will show how bibliometric evaluation has historically been constructed as a legitimate evaluation method. At the same time, based on interviews with more than 20 teachers and administrators from three universities in China, we will analyze the impact of quantitative assessment on the academic life of university teachers. In a positive sense, the quantitative assessment has introduced meritocracy into the field of evaluating the performance of university teachers, which has promoted the professionalism in academic work and made the academic promotion avoid the influence of *Guanxi* culture. However, this assessment technique has also brought about some unexpected negative effects such as the blind pursuit of quantity rather than quality, ritualism in knowledge production, violation of the academic autonomy of the faculty, and discouragement of application-oriented disciplines.

Key words: university teachers; quantitative evaluation; peer review; academic work

①陈洪捷、沈文钦:《学术评价:超越量化模式》,《光明日报》2012年12月18日。

②阎光才:《学术共同体内外的权力博弈与同行评议制度》,《北京大学教育评论》2009年第7卷第1期。

③周玉容、沈红:《大学教学同行评价:优势、困境与出路》,《复旦教育论坛》2015年第13卷第3期;蒋凯:《教育学术共同体建设中的同行评议制度》,《北京大学学报》(哲学社会科学版)2012年第49卷第2期。