

哲学社会科学评价中数量与质量关系的思考

温芳芳

(武汉大学 科学评价研究中心,湖北 武汉 430072)

摘要:近年来,在哲学社会科学领域,量化评价引起太多的关注和思考,其中,既有合理的反思与批判,也不乏无端的非议与指责。文章对哲学社会科学评价中数量和质量的关系进行了梳理和解析,认为数量与质量相辅相成,构成了科研成果不可分割的两个方面;质量的评价必须要以一定的数量为基础;质量的提升源自于数量的积累;数量和质量并不冲突,二者融合于量化评价之中;量化评价方法本身没有问题,只是利益因素诱导出对该方法的误用和滥用,才导致了一系列的极端行为和失范现象。笔者呼吁能够辩证地看待数量与质量的关系,能够客观地评论量化评价方法,能够更加冷静地寻找哲学社会科学评价的改革之道。

关键词:哲学社会科学;科研评价;量化评价;数量与质量

中图分类号:G644 **文献标志码:**A **文章编号:**1008-5831(2012)02-0112-05

一、引言

2005年12月《人民日报》发表了《北大试行论文代表作制度》一文,报道北京大学为克服追求论文发表数量带来的弊端,正在试点推行论文代表作制度。北大中文系在教师聘岗时实施代表作制度,即教师在聘期内只要发表了一篇有较大影响力的论文,经院系学术委员会认定,在相关的学科领域里属于领先地位,对该学者的科研工作就不再量化考核。这一消息迅速被各媒体反复转载,引起教育界乃至社会各界人士的高度关注和震动。社会各界将舆论批评的矛头对准以数量为主要考核标准的学术评价体系,同时为北大推行论文代表作制度叫好^[1]。北京大学中文系主任温儒敏表示,在北大中文系并不是现在才开始试点论文代表作制度,而是一直将论文质量放在相当重要的位置。

一直以来,学界对哲学社会科学的量化评价颇有微词。近年来,关于成果数量与质量的讨论比比皆是,关于重数量轻质量评价方式的批判此起彼伏,关于改革哲学社会科学评价制度的呼声也越来越高。北大所推行的论文代表作制度在一定程度上推动了科研评价由重数量向重质量转化,可谓是哲学社会科学评价改革历程中的标志性事件。而且,北大此举再次引发了学界关于量化评价的批判与反思。尽管量化评价的确有其先天不足,在后天的应用中也出现了一系列的极端行为和异化现象,这些批判和反思在很大程度上真实地反映出了量化评价的种种弊端。但笔者认为,学界对量化评价的认识有一些不理智

收稿日期:2012-02-12

基金项目:国家社会科学基金青年项目“人文社会科学成果评价体系设计与实证分析”(09CTQ014)

作者简介:温芳芳(1984-),女,河南平顶山人,武汉大学科学评价研究中心博士研究生,主要从事信息计量与科学评价研究。

致谢:感谢中国社会科学院任全娥博士对本文所作的重要贡献。

和不客观的成分。那么,在哲学社会科学评价中数量与质量究竟是一种什么样的关系?是否重数量的评价方式真的是阻碍科学发展和人才成长的桎梏?是否重质量重创新就意味着要完全抛弃数量?带着这些问题,笔者对哲学社会科学评价中数量和质量的关系进行了梳理和解析,呼吁大家更为客观、冷静地看待量化评价方法。

二、对于哲学社会科学量化评价的批判与反思

20世纪80年代,同行评议机制被正式引入到国内,同行评议方法被大量应用于科学评价实践中。同行评议具有一定的优点,如同行专家对学科发展、研究力量等方面进行比较分析,能够做出比较准确的判断,但由于同行评议在评价过程中缺乏公认的评价标准,加上受诸多学术和非学术因素影响,而使同行评议流于形式,致使同行评议的科学性和公正性受到越来越多的质疑与批判^[2]。为了减少人为因素的影响,使成果评价尽可能客观,哲学社会科学评价引入了量化评价方法。20世纪80年代以来,随着科学计量学的兴起和科研计量评价共识的增加,越来越多的科研机构和管理部门将科学计量学和文献计量学方法引入科研绩效评价和科学家绩效评价,并作为科研管理与科学决策的重要依据^[3]。20世纪90年代中后期,许多管理部门和研究人员开始探索借助科学计量分析的方法评价社会科学成果。借鉴国外哲学社会科学的三大引文索引《社会科学引文索引》(SSCI)、《艺术与人文科学引文索引》(A&HCI)、《国际学术会议社会科学引文索引》(IS-STP),中国人文社科成果定量评价的兴起始于南京大学研制成功的《中文社会科学引文索引》(CSSCI)的应用,目前该数据库已经成为中国主要的社会科学评价工具之一。

量化评价引入哲学社会科学领域之后,以其操作简单、客观公正等优势受到科研管理部门的欢迎,并在科研管理中得到了广泛应用。但是哲学社会科学是以人自身以及人类社会作为研究对象,具有不同于自然科学的研究特点与发展规律,又深受本国政治、经济、文化环境的影响和制约。因此,当量化评价这种源自西方自然科学领域的方法被简单地引入中国的哲学社会科学评价之后,既不符合哲学社会科学的学术发展规律和特点,也不符合中国学术界的现实情况。在全面的、完善的、公正的、科学的人文社会科学学术成果评价体系还没有建立的情况下,各高等院校和科研机构纷纷出现学术评价与数量挂钩的做法。职称评定、岗位津贴、奖金发放等都有类似的要求:收入国外权威引文索引多少篇、中文核心期刊上发表论文多少篇、获国家级或省部级优秀成果多少项、出版专著多少本,等等^[4]。把学者的个人利益与学术成果的数量如此紧密地捆绑在一起,一定程度上引起了学风浮躁,滋长了弄虚作假

假之风,造成了学术腐败现象的扩大。从科研管理的角度看,忽略人文社会科学的内在特殊性与多样性,只用简单量化方法,会使管理的目的大打折扣,甚至走向其负面^[5]。

量化评价方法的缺陷与优点从一开始就相伴而生,对其过度应用也遭到越来越多来自学界及社会各界的非议与责难,称定量评价“违背学术规律,漠视个性与人性”^[6],认为“学术评价量化背离学术评价宗旨”,“学术评价要靠同行评议”,认为是学术评价量化导致了种种学术不端行为,甚至被称为“学术评价的鸦片和罂粟花”^[7]。一种本来可以行之有效的评价方法被过度地、超范围地、不按规律地加以滥用,导致社会矛盾和学界争议日益加深,显然有悖于初衷。与此同时,学界对于量化评价重数量轻质量的做法以及由此所引发的一系列负面影响进行了反思,呼吁改革现有的评价体系。

有学者指出,目前的量化考核评价制度,已经严重妨碍中国高等教育和学术研究的健康发展,已经与建立创新型国家和进一步繁荣发展哲学社会科学的要求不相适应,它抑制了学术创新,妨碍了学术事业的健康发展,必须尽快进行改革和完善。学术机构应克服急功近利与攀比心理,对学术人员的激励、考核、评价不能简单地看成果的数量,而应更多地看成果的质量。在学术竞争上应以质量取胜,学术管理上应坚持质量导向,学术考核上鼓励原始创新,克服急功近利与过度考核现象。应以学术的标准评价学术,以创新程度作为评价的主要标准。要改变“唯量”的论文考核方式,采取国际通用的同行评议、匿名审稿、论文代表作制度等方式,实现定量与定性考核相结合。同时,应为科研人员营造宽松的学术氛围,让他们潜心研究,减少由考核压力导致的急功近利行为。教育部李卫红副部长特别提出要“改进评价方法,强化质量标准,形成正确激励导向”。她主张“要鼓励运用现代先进的科学研究方法和手段,把创新程度作为衡量科研成果质量高低的核心要素。在专业技术职务晋升和岗位聘任过程中,强调精品力作的作用,让确有真才实学的人得到充分认可,让真正有利于个人成才的激励机制落到实处”。

三、哲学社会科学评价中数量与质量的相关性分析

针对哲学社会科学评价领域存在的关于“数量与质量相冲突”、“数量增长影响质量提升”等言论,笔者将采用一种实证分析的方法对科研成果的数量与质量之间的关系进行辨析。选用中国引文数据库,分别从个人和机构两个角度进行计量和分析。取中国引文数据库中高被引作者前50名和高

被引院校前 50 名作为统计样本,分别统计其发文量、被引频次、下载频次、H 指数等指标数据(数据采集时间为 2011 年 6 月 17 日)。然后,分别将高

被引作者的指标数据和高被引院校的指标数据导入 SPSS13.0 进行 Spearman 相关性分析,结果如表 1 和表 2 所示。

表 1 高被引作者发文量与被引频次、下载频次和 H 指数的相关性分析

Spearman's rho		发文量	被引频次	下载频次	H 指数
发文量	Correlation Coefficient	1.000	0.505 **	0.635 **	0.512 **
	Sig. (2-tailed)		0.000	0.000	0.000
被引频次	Correlation Coefficient	0.505 **	1.000	0.826 **	0.926 **
	Sig. (2-tailed)	0.000		0.000	0.000
下载频次	Correlation Coefficient	0.635 **	0.826 **	1.000	0.863 **
	Sig. (2-tailed)	0.000	0.000		0.000
H 指数	Correlation Coefficient	0.512 **	0.926 **	0.863 **	1.000
	Sig. (2-tailed)	0.000	0.000	0.000	

** . Correlation is significant at the 0.01 level (2-tailed).

表 2 高被引院校发文量与被引频次、下载频次和 H 指数的相关性分析

Spearman's rho		发文量	被引频次	下载频次	H 指数
发文量	Correlation Coefficient	1.000	0.966 **	0.962 **	0.790 **
	Sig. (2-tailed)		0.000	0.000	0.000
被引频次	Correlation Coefficient	0.966 **	1.000	0.941 **	0.873 **
	Sig. (2-tailed)	0.000		0.000	0.000
下载频次	Correlation Coefficient	0.962 **	0.941 **	1.000	0.823 **
	Sig. (2-tailed)	0.000	0.000		0.000
H 指数	Correlation Coefficient	0.790 **	0.873 **	0.823 **	1.000
	Sig. (2-tailed)	0.000	0.000	0.000	

** . Correlation is significant at the 0.01 level (2-tailed).

一般来说,发文量在一定程度上代表着某一作者或者机构生产科学知识及文献的能力,而被引频次和下载频次则代表着某一作者或机构所生产的科研成果在知识交流中的质量和影响力。因此,在本次相关性分析中,将发文量看作是表征数量的指标,将被引频次与下载频次看作是表征质量的指标,而 H 指数则是数量和质量两方面的综合考察指标。从相关性分析结果看,无论是在个人层面还是机构层面,发文量与被引频次、下载频次和 H 指数都存在着显著的正相关关系。由此表明,数量与质量并不冲突,拥有高被引频次、高下载频次和高 H 指数的作者或者高校往往同时拥有较高的发文量。也就是说,无论是高被引作者还是高被引院校,其质量和影响力都基于一定的数量优势。

四、对于哲学社会科学评价的建议和思考

无论宏观评价还是微观评价,都要涉及到数量与质量的问题。数量易于统计,质量的鉴定则相对复

杂。但任何事物都有数量和质量两个方面的属性,要完整地界定一个对象既要定量又要定性,没有量也就无所谓质,没有质的界定,量也就没有意义。任何要把数量和质量割裂开的思想和行为都是不理智和不可行的。

(一)质量的评价必须要以一定的数量为基础

如果一篇论文都未发表,就无法认定某一国家地区或机构的科研绩效、学术水平和影响力。有一种观点认为,科研绩效评价应取消对论文数量的规定和要求。显然,这不符合实际,也是不负责任的。从 SCI 和 SSCI 等收录的论文数量看,科技发达国家的论文数量规模占明显优势,说明在遵循科研产出的客观规律和学术规范的前提下,论文发表的数量是科研实力、学术水平的一个客观评价指标。上文对科研成果数量指标与质量指标的相关性分析结果也表明,数量与质量具有较高的相关性,无论是个人还是高校,其科研产出的质量和影响力都是基于一定的数量优势。片

面追求论文数量,固然本末倒置,但把论文质量不高的原因归咎于论文数量太多同样不得要领^[8]。论文质量的控制要靠期刊编辑认真把关和同行严格评审,而实际的质量水平和学术影响力则要看其发表后被同行认可和接受的程度,即被引用的频次和影响力,以及成果转化应用所产生的社会和经济效益。

(二)质量的提升源自于数量的积累

论文质量并非与数量成正比,但应认识到科学研究是一个长期积累的过程,无论对于一个国家或地区,还是一个机构或个人,科技论文的产生都必须有一个数量积累的过程,才能有整体质量的逐步提高。这一方面有赖于学术规范和监督机制的加强和完善,另一方面则应注重学术评价的导向作用,摒弃单纯以论文数量评定绩效的做法,提倡数量与质量相结合的评价方法。例如,对基础理论研究型论文,采取质量影响因子法评价,将发表论文的总被引频次除以论文总篇数作为评价指标,可促使科研人员慎重发表文章,因为每一篇文章如果不能对总被引频次有较大“贡献”,反而会导致发表文章数越多,个人质量影响因子越低。再如,可采用“代表作法”,即只推荐一定数量最有代表性的作品进行被引频次统计或由同行评定其质量高低,也可有效抑制片面追求论文数量的势头。那种将数量与质量对立,抛弃数量只追求质量提升的做法完全违背了科学发展和人才成长的规律。那种一生成果寥寥却对人类有巨大影响的科学精英固然伟大,但毕竟只是少数。对于大多数科研人员来说,还是要遵从“量变引起质变”的自然规律。尤其对于那些刚刚进入科研领域的年轻人,更是要经历数量的积累才能有质量上的突破。

(三)数量与质量并不冲突

在对待论文数量与质量的问题上应持辩证的观点。今天学术界存在的种种失范与不端问题,其根本的原因并不在于评价方法本身,更不能把所有的责任都归结为成果数量的急剧增长。数量与质量本身并不冲突,只是在引入和使用定量评价方法时,操作程序过于简单机械,又赋予了过高的利益诱惑,人为地把数量与质量对立,使定量评价走向了只重数量、忽视质量的极端,从而导致了种种异化现象和失范行为,其实质是利益因素诱导出对评价的误读与误用。这一切与量化评价方法本身没有必然的联系,那种把量化等同于“重数量轻质量”的观点是十分错误的。实际上,量化评价的核心思想就是提倡数量与质量相结合。发文量代表科研工作者的生产能力,期刊的级别和影响因子在一定程度上代表着论文的质量,而被引频次则反映出该项成果在知识交流系统的

贡献度和影响力。综合了发文量、影响因子、被引频次的量化评价方法能够很好地践行数量与质量相结合的评价思想。

(四)同行评议与定量评价孰优孰劣

科研成果的评价分为两个方面:一是从专业学术角度,对成果内容本身的创新性、科学性和应用性及其发表价值的评审,这主要靠同行评议;二是从科研管理角度对科研人员或机构的绩效考核,这主要是运用科学计量学的方法对已发表的论文及其学术影响力,尽可能做出符合实际情况的量化评价结果。两者的操作主体和功能目的不同,既相对独立又相互联系。前者是学术评价的基础,评价主体是同行专家,主要是评审论文成果的价值以决定是否能够发表及其发表的场合;后者主体是科研管理部门或机构以及科学计量分析人员,主要是对已发表论文成果数量和质量的综合评价,其目的是作为绩效考核、职务评聘、资源分配的依据。同行评议与文献计量评价各有利弊,都非完美无缺。之所以采用其中一种与不采用另一种方法,并非因为某一种方法优于另外一种,而是根据评价目的和评价对象性质的不同所做出的合理化选择,与评价方法本身的优劣没有关系。目前,中国一些学者对成果量化评价大加抨击,同时强调应重视同行评议,甚至提出要推翻定量评价回归同行评议,这其实是对评价方法存在误解。同行评议和定量评价都有其不可替代性,二者应该长期共存、互为补充。

(五)学术失范究竟谁之过

在对量化评价的种种批判与反思中,笔者也注意到了一种非常极端的言论,认为学术界一切失范问题的元凶就是量化评价,甚至认为量化评价已经成为阻碍中国哲学社会科学以及学术人才成长的桎梏。那么量化评价是否应该承担学术界所有的责难与批判?笔者认为,量化评价方法本身没有问题,只是人们对该方法的误用和滥用才导致了一系列的极端行为和失范现象。实际上,量化评价方法在最初被引入时,在科研管理领域发挥了积极有效的作用,并且短时期之内迅速推广应用,这就足以说明量化评价有其科学性和合理性的一面。如果推翻了量化评价,那是否就可以解决目前学术界所有的问题,又能否建立一种新的更为有效的评价制度?我们可以轻易地打破一个旧世界,但在这之前应该思考有没有能力去构建一个新世界。目前,在哲学社会科学评价领域,关于量化评价的批判很多,但是对新的评价方法和体系的研究却很少,也就是说大家都意识到了旧的评价体系的弊端,但是还没有能力构建一

种新的行之有效的评价制度。由此,中国哲学社会科学评价进入一种尴尬的境地,一方面猛烈地批判量化评价,另一方面又不得不继续沿用这一评价方法。笔者认为,任何一种评价方法都不是万能的,与其批判量化方法本身的种种不是,倒不如更加冷静地思考如何趋利避害,如何更加合理地利用这种方法。

五、结语

现代科研工作存在着巨大的压力,而且科研制度本身对成就和成功都起着激励的作用。科学家同其他社会劳动者没有什么不同,并没有摆脱其他行业的人们所具有的感情、奢望和人性弱点。现代科研毕竟是一项职业,要想获得成功,一个研究人员必须尽可能拿出自己的成果,特别是有分量的成果,唯有如此,才能得到同行的认可和社会的承认,才能确保获得政府的资助,建立自己的实验室,创造条件招收研究生,争取当选科学院院士,并企盼有朝一日被邀请到斯德哥尔摩领取诺贝尔奖^[9]。这种夹杂着强烈愿望和复杂情感的心态不可避免地会渗透到科学研究及成果分配的过程中。在这样的科研制度和环境中,确实很容易滋生浮躁的学术风气和种种不端的行为,这种情况与科学评价方法本身并没有直接的因果关系。

任何一种评价方法,一旦与利益紧密结合,又没有相应的监督约束机制做保障,不可避免地会产生一些异化,各个国家、各个领域、各个行业概莫能外。只不过当这种问题发生在以严谨认真而著称的学术界、发生在以追求科学真理为己任的科研人员身上,就显得尤为恶劣和不能容忍。近年来,在哲学社会科学领域,量化评价承担了太多的评论与质疑,其中,既有合理的反思与批判,也不乏无端的非议与指责。我们

支持以质量和创新为导向的哲学社会科学评价改革,也为北大试行的“代表作制度”拍手叫好。与此同时,笔者也呼吁大家更加理性地辨析数量与质量的关系,能够客观辩证地看待量化评价。希望大家能够结合哲学社会科学自身发展的规律,以及中国学术界的现实情况,更多地从制度、环境、文化等方面探讨如何规范和优化科研评价,而不要过多地纠缠于评价方法本身的优劣和不足。

参考文献:

- [1]曹建文.北大中文系论文代表作制度探路学术评价体系改革[N].光明日报,2006-01-04.
- [2]刘文达.关于基础研究同行评议的约束机制构建问题[J].科技管理研究,1999(6):13-18.
- [3]邱均平.文献计量学[M].北京:科学技术文献出版社,1988:189.
- [4]任全斌.我国人文社会科学研究成果评价体系研究[D].武汉:武汉大学博士论文,2008.
- [5]上官子木.学术管理的有效性:量化管理方式并不适用于学术管理[N].南方周末,2002-09-19.
- [6]黄安年.违背学术规律 漠视个性与人性——评目前高校流行的“量化”管理制[EB/OL].(2008-4-5).http://www.acriticism.com.
- [7]叶再林,聂聃.索引崇拜——科研评价的罂粟花[J].社会科学论坛,2007(2):101-107.
- [8]朱大明.学术评价中论文的数量与质量[EB/OL].(2007-09-12).http://www.edu.cn.
- [9]朱进宁.背叛真理的人们——科学界的弄虚作假[M].方玉珍,译.北京:科学出版社,1988.

Contemplation on the Relationship between Quantity and Quality of Humanity and Social Science Evaluation

WEN Fang-fang

(Research Center of Chinese Science Evaluation, Wuhan University, Wuhan 430072, P. R. China)

Abstract: During recent years quantitative evaluation has attracted a variety of concerns in the field of humanity and social science, including both reasonable introspection and unreasonable critics. In fact, quantity and quality are two sides of research achievement and one can't be separated from the other. The qualitative evaluation must be based on quantity and the promotion of quality also comes from the accumulation of quantity. The two elements are integrated in the method of quantitative evaluation. This paper analyzes the relationship between quantity and quality, in the hope that people can have a dialectical attitude towards quantitative evaluation methods and can find a proper way to realize the innovation and reform of humanity and social science research evaluation.

Key words: humanity and social science; scientific research evaluation; quantitative evaluation; quantity vs quality

(责任编辑 彭建国)