

Doi:10.11835/j.issn.1008-5831.pj.2021.04.001

欢迎按以下格式引用:陈敏,王轶.破“五唯”政策视角下的学术成果评价研究[J].重庆大学学报(社会科学版),2021(4):60-70. Doi:10.11835/j.issn.1008-5831.pj.2021.04.001.



Citation Format: CHEN Min, WANG Yi. Research on the assessment of academic achievements from the perspective of the policy of removing the “five standards for profession elevation” [J]. Journal of Chongqing University (Social Science Edition), 2021(4):60-70. Doi: 10.11835/j.issn.1008-5831.pj.2021.04.001.

破“五唯”政策视角下的 学术成果评价研究

陈敏¹,王轶²

(1. 国家开放大学 改革发展部,北京 100039;2. 北京工商大学 经济学院 100048)

摘要:最近,教育部、科技部等部门出台了多个破除“五唯”不良倾向的政策文件,特别强调要破除学术成果评价中的“五唯”导向。那么,我国学术成果评价究竟存在哪些问题、原因何在、如何解决?梳理已有文献发现:目前我国学术成果评价容易受到作者个人和单位级别的影响;学术成果评价与个人利益、单位利益相联系;单位的学术成果评价制度受到各级政府部门出台的政策文件影响;交叉学科学术成果评价存在盲区。分析认为:全球各评估机构发布学术评估指标的导向作用,国家、省部级及各级教学科研单位制定的学术成果评价制度,按照学科专业细分学术成果的相关制度等,均是造成不良评价倾向的诱因。为此,笔者借鉴发达国家学术成果评价的成功经验,结合我国学术成果评价的目的,建议重新审视我国学术成果评价制度,不唯论文不是不要论文,进而完善我国学术成果的评价方法,鼓励创作出有价值、高水平的原创学术成果。

关键词:学术成果;学术评价;破“五唯”;唯论文;评价制度;评价方法

中图分类号:G322 **文献标志码:**A **文章编号:**1008-5831(2021)04-0060-11

2020年2月17日科技部印发《关于破除科技评价中“唯论文”不良导向的若干措施(试行)的通知》,要求科研单位在科研成果评价中,破除重论文数量,轻论文质量,重论文影响因子,轻论文社会贡献的不良倾向。此后,教育部、科技部又联合印发了《关于规范高等学校SCI论文相关指标使用 树立正确评价导向的若干意见》,要求各科研单位建立科学评价体系,注重学术成果的实效。根据《2020年中国科技论文统计结果》发布的数据,2010—2019年,我国发表SCI论文301.91万篇,排名世界第二,仅2019年发表SCI论文49.59万篇,比2010年的12.15万篇高出

基金项目:北京市教育科学“十三五”规划2019年度一般课题“开放大学促进首都终身教育的路径与策略研究”(BADB19052);中国成人教育协会“十三五”规划2017年度项目“国家开放大学体系科研质量保障研究”(2017-141Y)

作者简介:陈敏,国家开放大学改革发展部,Email:chenmin@ouchn.edu.cn;王轶,北京工商大学经济学院。

37.44 万篇,近三年我国发表 SCI 论文数量正在快速增长。又据世界知识产权组织和美国康奈尔大学等机构联合发布的近三年《全球创新指数》统计,2016—2019 年间我国创新指数全球排名分别为第 25 位、22 位、17 位和 14 位,与之相应的我国全球创业指数也快速上升。中国电子学会 2020 年 12 月份数据显示,我国科技成果的转化率仅为 6%,而西方发达国家为 50%左右。由此,引发我们思考:我国学术成果到底该怎么评价?现存学术成果评价体系存在哪些问题?存在问题的根源在哪里?有无可借鉴的相关经验?换言之,如何破旧立新重构我国的学术成果评价指标体系,实现教育部、科技部提出的破“四唯”“五唯”目的。

一、学术成果评价及其文献梳理

近年来,关于学术成果评价研究的文献主要集中在两个方面:一是学术成果的评价方法,二是学术成果的评价制度。

(一) 学术成果评价方法研究

1. 从学术成果被引数量,判断学术成果的社会影响

目前,主张以成果的被引数量判断学术成果价值的代表性文献有:Aversa^[1]从学术论文的视角分析学术论文的被引数量与论文老化之间的关系,认为低被引论文的老化速度明显快于高被引论文;Smith & Presser^[2]首次采用 H 指数方法测评学者与学术成果质量之间的关系,认为学者的 H 指数越高,其学术成果质量越高;Bornmann^[3]采用引用行为研究方法分析学术成果的引用行为,发现对学术成果的引用行为会受到外在因素的影响,比如学术成果展现平台、作者所在单位和作者自身影响力;Egghe^[4]对 H 指数方法进行改进,增加了高被引学术成果的被引频次,认为该方法更能够反映学术成果的创新程度;张凌云等^[5]以旅游学科学术成果为例,从数量型和 H 指数方法对学术期刊与学术论文质量之间关系进行评价。

2. 运用替代变量评价方法,分析学术成果的综合质量

随着网络化、数字化的快速发展,可以采用学术成果的下载量、浏览量、收藏量、粉丝量、推送量、点赞量、评论量、置顶量、转载量等替代变量来评价学术论文的质量。代表性研究有:Priem & Hemminger^[6]采用 Web2.0 服务,在学术论坛平台检查学术成果的使用和引用情况,如书签、微博等,试图通过网络社交平台来评价学术论文的质量;Piwowar^[7]认为,随着网络社交平台的发展,可以通过学术成果在社交平台上的下载量、浏览量、书签量、讨论量等数据来评价学术成果的质量;王轶^[8]认为,学术成果的学术影响力和社会影响力可以通过学术成果被四大文摘(《新华文摘》《高等学校文科学术文摘》《中国社会科学文摘》和《中国人大报刊复印资料》)的转载次数和微信、微博等浏览次数来衡量。

3. 评价学术成果内容,判断成果等级

代表性文献有:Yogatama^[9]以学术论文为例,研究学术论文作者、研究主题和发表期刊等指标来预测学术论文的引用,进而分析学术论文的质量;Livne^[10]采用机器学习技术分析数百万篇学术论文的质量,识别的特征包括:期刊社会声誉、论文作者、作者机构、引文特征和论文内容等指标,根据这些指标建立模型分析学术论文的未来被引数量,以此评价学术论文质量;索传军等^[11]基于认知计算的学术成果评价方法,从该领域的学术语义内容的视角对学术成果的内容进行评价;赵镇等^[12]基于 SCI 学术期刊的学科分类和分区,通过最高系数、基准系数计算他们的 p 值系数,这样的计算方法有助于消除不同学科学术期刊之间影响因子的差异,为客观评价学术成果价值提供参考。

4. 采用学术成果综合评价方法,评价学术成果质量

对学术成果质量采用定量与定性相结合的评价方法。代表性文献有:Van Raan^[13]采用论文被引数量和同行评价的方法来评价学术论文的质量;贺颖、邱均平^[14]认为,同行评议与科学计量相结合,建立遴选系统,是评价学术论文质量的另一种有效方法;张积玉^[15]以S大学为例,分析该大学在教师学术评价方面的探索,包括评价制度的改革、对不同教师群体采用不同的评价方法,在评价方法选择方面,要主观与客观相结合、学术水平评价与学术贡献评价相结合等;刘俏^[16]将Altmetrics指标与传统计量指标(即论文的被引频次)结合起来,计算学术论文的社会影响力和学术影响力。

5. 运用社会网络分析方法和同行评价方法,评价学术成果质量

部分学者采用社会网络分析的结构指标,通过逐级提取方法,提取学术成果中心度指标,据此对学术成果的质量进行评价,评价的方法有:网络节点、学派论等(Mcnamara^[17]; Upham^[18];高锡荣,杨娜^[19])。还有部分学者采用由同一领域专家对该领域内的学术成果进行评价的方法(贺颖等^[20]),这种评价方法在学术论文评价中应用广泛。

(二) 学术成果评价制度研究

对学术成果评价制度的研究,已有成果相对较少,主要原因在于学术成果评价制度往往会受到科研单位管理体制的影响。代表性文献有:Smith^[21]认为,目前高校存在很多校区,不同校区发展定位存在差异,那么学术评价就应该采取差异化评价方法,这样既有利于学校的发展,也有利于教师的成长;马永霞、仇筋熙^[22]以国家社科基金成果鉴定为例,强调学术成果的阶段性评价与同行评价等两项评价制度在提升学术成果质量中的关键作用;李法惠^[23]认为,在我国学术论文评价制度中,存在过度引用与抄袭之间的关系、过度引用与高质量论文之间的关系;刘明^[24]认为,学术论文不能“以刊评文”,而是强调学术论文的质量,建议完善我国学术评价制度,构建以学术界和读者为共同基础的评价机制;叶青、彭辉^[25]认为,目前学术成果评价制度存在不完善之处,建议在定量和定性评价基础上,采用代表性学术成果作为评价学术论文(或著作)的一项指标;Lamont & Mallard^[26]通过比较美国、英国和法国社会科学与人文学科的学术评价方法,认为人文社科学术成果的评价不能注重形式,需要学术评价委员会设立不同的专家委员会,加强同行评价及其加强对已有成果社会影响力和学术影响力的评价等;白如江等^[27]基于学术论文的外在因素和内在因素,分析了学术论文评价的未来发展趋势;李立国^[28]认为,学术评价制度就是牛鼻子,什么样的学术评价制度就会产生什么样的学术评价结果,因此评价制度具有极端重要性。

通过文献梳理发现,已有学者主要关注学术成果的评价方法,包括H指数法、社会网络法、同行评价法等;在评价制度方面,国外文献偏重于在学术自由、学术自治背景下探讨学术成果的评价制度,国内文献侧重于在理论层面探索学术成果的评价制度。上述研究成果为本研究提供了有益借鉴,但也存在不足需要回答,主要体现在:我国学术成果评价存在哪些不良倾向?其根源在哪里?结合科技部、教育部破“五唯”“四唯”文件精神,我国高校和科研机构在学术成果评价方面应该怎么做?回答和解决这些问题正是本研究的目的所在。

二、学术成果评价的不良倾向分析

(一) 学术成果评价与作者个人社会影响和个人利益挂钩

1. 学术成果评价容易受作者学术身份的影响

在学术研究与评价实践中,学术成果评价往往受到学术成果作者学术身份的影响。主要存在

下面两类影响:一是已经在业界公认、有很多知名头衔的学者,其署名的学术成果一般很容易被认定为高质量的学术成果,主要原因在于这类人群已经在学术界树立了相对权威。另外,评价该类学术成果的专家都有可能与知名学者有着密切关联,如师生关系、同一师门关系等连带关系,所以知名学者的学术成果即使没有太多新意,其学术成果也很容易被认定为高质量学术成果。对于是否真正属于高质量学术成果,目前国内的评价指标体系很难认定。当然,这里并不能否认上述知名学者曾经发表过、出版过有思想、有创新性的高质量学术成果。二是熟悉高水平期刊写作套路的高产作者。有这么一批熟知国内学术评价中被认定为高水平学术期刊写作范式的作者,他们熟知这些期刊的偏好和范式,且与这些期刊形成了良好默契,其学术成果很容易在这些认定的高水平学术期刊上发表,发表后的论文在国内很自然地被界定为高水平的学术成果。如国内一些学者熟知 SCI 期刊的写作套路,且与 SCI 期刊形成了良好的默契关系,他们的文章很容易在 SCI 期刊上发表,发表后的 SCI 论文很容易在国内被认定为高质量学术成果。针对上面两类学者的学术成果,该如何评价,值得高校和科研单位深思。

2. 过于强调学术成果评价与个人利益之间的关系

作为科研激励的一种手段,学者发表的学术成果级别越高,获得的科研奖励越多,这是可以理解的,但是如果学者把学术成果的评价与个人利益完全挂钩,那么学术成果的评价就容易失去其本质的公正性,就容易功利化。过于功利化的学术成果评价不利于科学研究,特别是基础科学。现在国内各高校及科研单位,普遍将学者的学术成果等级直接与其职称、奖励、职务晋升、住房、人才引进、学术头衔等个人利益捆绑在一起,导致的结果是容易滋生学术腐败,学术成果商业化,学术创作短视化,直接影响了学术成果的创作环境,影响了学术成果评价的公正性。

(二) 学术成果评价与作者所在单位级别、单位利益挂钩

1. 学术成果评价与作者所在单位级别挂钩

目前,国内学术成果评价机构在评价学术成果的质量时,如果学术成果的质量差异不大,作者所在单位的行政级别越高,其学术成果被评定的级别就越高。主要表现在两个方面:一是学术论文发表方面。在国内高水平学术期刊上(如中国社会科学院中国社会科学评价中心发布的顶级权威期刊),一般都刊发“211”“985”“双一流”高校或者省部级科研院所科研人员的文章,非上述科研单位人员很难在上述顶级权威期刊上发表文章;甚至有部分期刊明确规定,不发表非“211”“985”“双一流”高校及地级市科研单位作者的论文,无论这些学术论文质量高低。二是学术成果评奖方面。上报省部级或国家级奖项的学术成果可能都在同一出版社出版,或者是在同一学术期刊上发表,就因为作者所在单位的行政级别不同,获奖的级别和难易程度存在明显差异。单位级别越高,越容易获得高级别奖项。

2. 学术成果评价与作者单位利益关联

目前社会上的各种高校综合排名、学科排名都与高校、科研单位产生的学术成果级别挂钩,挂钩的结果就是,高校和科研单位尽可能鼓励本单位人员多产出高质量学术成果,出重金招聘能产出高质量学术成果的学者。同时,学术成果挂靠单位也会动用单位资源去影响学术成果的评价,推动本单位多产出高质量学术成果。一般而言,被评价学术成果的级别越高,各单位动用本单位资源的可能性就越大。学术成果评价的结果不再仅仅取决于学术成果质量的高低,而是部分地取决于学术成果作者所在单位干预的程度和单位社会影响力的大小。学术成果的评价在某种程度上变成了

单位之间的利益之争,这种竞争必将导致学术成果评价的不公平、不公正。

(三) 学术成果评价容易受到各级政府政策文件的影响

1. 制度层面规定以学术成果的级别代替学术成果质量的评价

(1) 以学术期刊的等级代替学术论文质量的评价。目前我国高校和科研院所大都用学术期刊的级别来代替学术成果质量的评价。高校及科研机构要么根据南京大学中国社会科学研究评价中心发布的 CSSCI 来源期刊目录排名,要么根据北京大学图书馆发布的《中文核心期刊要目总览》中的期刊排名,要么根据中国社科院中国社会科学院评价中心发布的《中国人文社会科学期刊评价报告》将期刊分成不同等级,依据期刊等级来评价学术论文质量。上述以学术期刊等级代替学术论文评价的结果是,少量发表在最高级别期刊的短评文章,可能只有 1—2 页,也被认定为高质量的学术论文。还有一些发表在最高级别期刊的论文,发表以后却是“零”下载量,这就意味着该学术论文没有被所有学者浏览,这类学术论文认定为高质量学术论文值得商榷。

(2) 以学术成果的获奖级别代替学术成果质量的评价。目前,科技领域和人文社科领域设置了不同级别的奖项,包括国家级、省部级、市级、单位级等,还有一部分是行业、学会、协会组织的奖项。如果认真分析一下某一级别获奖者的名单就会发现,各个级别的奖项几乎总是被相同的一群人占有。如果我们再研究一下这些获奖题目还会发现,本次获奖题目与上一次获奖题目存在很高的关联性。那么,评价各级、各类学术成果奖项,是因人设奖还是因学术成果质量而设奖呢?显然,各类学术成果奖项不能完全反映该级别学术成果的真实学术质量。

2. 以科研项目级别代替学术成果质量的评价

目前,国内部分高校及科研机构科研人员即便产出了高质量学术成果,但成果折算成科研分数却并不高,而高级别科研项目的科研分数却可以折算成 10 余篇,甚至是 20 余篇高质量学术论文的科研分数。另外,国内大多数高校及科研单位明确规定,科研项目是科研人员晋升职称的硬性标准。在某个职称聘期内,科研人员如果没有国家级科研项目,即使有再多、级别再高的科研论文,也没有资格申报高一级的职称。部分高校及科研单位甚至将科研项目的等级视为学者学术成果的水平,包括现在高校和科研单位引进人才,一般也以学者所承担科研项目的级别来认定学者的学术水平。显然,以科研项目级别代替学术成果水平的做法值得思考和总结。

(四) 交叉学科学术成果评价存在盲区

随着社会融合发展,跨学科综合研究越来越多,我国原有的按照学科划分学术成果的模式已经难以适应社会发展的需要,有些交叉学科的成果很难划分到某些具体学科中。例如,人工智能与教育结合形成的人工智能教育学,神经学、会计学结合形成的神经会计学,互联网、人工智能、汽车制造业结合形成的基于 5G 网络人工智能的汽车驾驶技术等。交叉学科产生的学术成果往往也是全新的领域,或是多学科知识融合的结果。如果按照现有学术成果归属模式来划分成果的学科,那么新兴学科和交叉学科的原创学术成果很难归属于某个学科,这类交叉学科的学术成果很难被认定。成果不能得到认定或评价,意味着这类学术成果很难被认定为高质量的学术成果。即使是这类交叉学术成果去报奖、评奖,成果的学科归类、专家遴选等都会遇到同样的问题。

三、形成学术成果评价误区的原因诊断

(一) 全球各学科评估机构发布学科评估指标的导向作用

近年来,全球最大的学科评估机构——国际高等教育资讯机构 QS(Quacquarelli Symonds,简称

QS)每年发布全球学科评估结果,评估学科涉及社会科学、自然科学和工程技术等五大学科,评估结果作为全球学科评估、学科专业国际认证、留学生留学指南的重要依据,评估指标中最重要的指标之一就是该大学发表的SCI、SSCI 学术论文数量和引用次数。正是由于全球学科评估机构制定的学科评估指标存在较强的导向作用,导致了我国各高校及科研单位把SCI、SSCI 学术论文的数量作为本单位科研人员科研考核的主要指标。正是在各单位科研考核指标的指引下,科研人员把发表SCI、SSCI 学术论文作为其科研主攻方向。

(二) 各级政府制定的政策文件引导各高校、科研单位制定相应的学术成果评价制度

目前教育部制定的学科评估制度、双一流学科评选方法等,在评估指标中,明确列举了“中国版ESI 高被引论文”(自然科学学科)或“A 类期刊论文”(人文社会学科),并列举了一定数量的SCI、SSCI 学术期刊的数量作为评价的指标。在获奖方面,按照国家级和省部级等获奖数量来衡量各教学科研单位的科研水平,评估结果直接与高校、科研单位的发展密切联系。正是在上述政策文件的引导下,高校、科研单位制定了相应的学术成果评价标准来要求本单位科研人员执行,导致了各教学科研单位以学术成果奖励的级别、学术期刊的等级来认定学术成果水平。为了激励本单位科研人员产出不同级别政府部门所制定的评估指标中罗列的A类成果,各单位便在制度层面规定了A类成果的奖励机制,进而激励科研人员为了个人利益发表更多的A类学术成果。这些A类学术成果的学术贡献有多大,对社会的发展贡献有多大值得怀疑。而在评价指标中没有涉及的内容,自然也不是学者关心的内容。

(三) 学术成果按照学科专业细分制约了交叉学科学术成果的评价

按照学科专业设置科研项目、开展学术研究的确是国内外通行做法,然而社会发展推动了交叉学科的诞生,多学科交叉便产生交叉学术成果,交叉学术成果很难归属于某个学科专业,自然就很难在以学科专业来设置栏目的学术期刊上发表,不能发表意味着学术成果很难被认定和评价,其结果阻碍了交叉学科学术成果的评价和发展,进而可能影响学术成果的创新。上述分析不难发现,目前国内外高校、科研机构的学科专业设置在一定程度上制约了跨学科、交叉学科学术成果的产生与评价。

四、发达国家学术成果评价的经验借鉴

(一) 英国学术成果评价标准

英国科研机构特别是英国大学,关于学术成果评价的依据可以概括为四个方面:(1)学术成果能否促进学术交流,即学术成果被引用的情况,以及学术成果能否对政府的政策制定产生影响;(2)学术成果能否提高本领域的学术交流水平;(3)学术成果内容是否规范,主要包括学术论文的逻辑层次是否清晰,学术内容是否规范,专业术语是否规范等;(4)学术成果能否促进知识传承,即论文发表后对本学科的知识是否有承前启后作用,与已有学术研究是否存在连续性等^[29]。

根据上述四方面评价指标对学术成果质量进行评价,其结果一般分为五个等级:最高一级为四星级,即学术成果在世界范围内属于原创性,对本学科的研究有重要推动作用,学术成果严谨规范,比如学者的学术成果获得了诺贝尔奖;三星级指学术成果在世界范围内得到了认可,有一定的原创性,但还达不到国际上公认的原创性标准;二星级指学术成果的质量达到或基本达到国际公认的标准或水平,学术成果有一定原创性,学术成果较为严谨;一星级指学术成果在国内属于原创性,在本

国本学科得到认可,学术成果内容较为规范;无星级指学术成果原创性存在不足,尚未达到本国或本学科认定标准。通过对上述评价指标和评价方法的分析发现,英国科研机构对人文社科科研成果的评价虽然注重定量与定性结合,但显然更注重定性评价。

(二) 德国学术成果评价方法

德国科学委员会成立于1957年,20世纪80年代开始就着手完善学术成果的评价体系,并试图通过学术成果评估、学术成果排序和创新评估方法等进一步完善学术成果评价体系。20世纪90年代,德国科学委员会进一步完善学术成果评价程序,评价过程更加透明。据此,德国科学委员会发布了《科研成果评价与监督建议》。根据建议,德国学术成果评价方法及内容主要包括以下五方面:(1)同行评议。这是学术成果质量评价的必须环节,在学术成果质量评价、教师资格晋升、学术成果获奖、学术机构研究能力评价等方面都采用同行评价方法。(2)量化指标。经常使用的量化指标包括:学术成果有无项目资助、成果作者数量、成果参考文献数量、学术成果被引次数、H指数等。(3)排序。德国从20世纪80年代开始就采用不同形式的排序评价方法,专家根据学术成果的质量指标和数量指标标准进行排序。排名将充分考虑学术成果之间的差异,包括同类型学术成果纵向和横向的比较,这种比较主要是确定该学术成果的创新性在国内外同类研究中所处的位置,确定创新性的程度。(4)项目资助。评价学术成果质量一般会考虑学术成果是否获得项目资助、项目资助级别、学术成果与所资助项目之间的关系。(5)学术不端行为。主要检测该学术成果是否存在剽窃、学术造假、学术成果抄袭、研究方法不当、研究逻辑不严谨、不合法署名、学术不规范、一稿多投、一稿多发等,对学术不端行为的要求极为严格,基本实行一票否决制^[30]。

(三) 美国学术成果评价实践

美国早期关于学术论文的评价主要模仿英国,第二次世界大战以后,学术概念悄然兴起,科研激励机制逐步建立,学术出版物(包括学术论文)成了科研人员树立学术地位的重要抓手。特别是Boyer^[31]发表“Scholarship reconsidered: Priorities of the professoriate”以后,美国高校及科研机构对学术成果的评价日趋规范。目前,学术成果评价主要包括以下几方面内容:(1)同行评价。美国的同行评议比较严格,所以美国很多高校把期刊发表成果的数量作为评价学术成果价值的主要手段。在发表的成果数量与质量方面,专家主要查看美国科学情报所发布的SSCI数据库(社会科学引文索引)、ISSTP数据库(国际学术会议社会科学引文索引)和A&HCI数据库(艺术与人文科学引文索引)。(2)定量分析方法,即引用指数分析法。主要分析学术成果参考文献数量、交叉学科文献数量等,以此判断学术成果的创新程度。(3)学术影响指标。主要考察该学术成果的社会引用、学术成果获奖、学术成果在企业和社会中的应用情况。比如:以美国总统名义设立的学术成果奖励和各级学会设立的学术成果奖励等,这些都是学术成果学术影响力的重要指标^[32]。(4)学术成果违规惩戒制度。为了规范学术行为,美国专门颁布了《版权法》《商标与不公平竞争法》,其严厉的惩戒制度大大降低了学术抄袭行为,增强了研究人员的学术自律。

(四) 澳大利亚推出学术成果评价卓越研究项目

澳大利亚对学术成果质量的评价主要是基于该国的卓越研究项目,该项目采用较多的方法就是同行评议方法和引文分析方法。这两种方法在使用过程中相互补充:(1)同行评议方法。卓越研究项目中的同行评议同英国同行评议的划分等级类似,不过澳大利亚划分六级。其中五级是学术成果质量远超世界水平,四级是学术成果质量超过世界水平,三级是学术成果质量处于世界平均水

平,二级是学术成果质量低于世界水平,一级是学术成果质量远低于世界水平,0级是未达到参评标准,不做评价。(2)引文分析方法。该方法主要是分析期刊文章、文章文献与学术成果之间的关系^[33]。这种方法是从文献分析入手的共引分析(即两篇文献被后来发表的文献同时引用,那么这两篇文献之间是否存在关联),该方法是以知识结构、邻里联系法则、映射为方法论基础,来分析研究对象之间的逻辑关系,挖掘潜在的有用知识。这种分析方法又包括三种分析方式:一是引文描述性统计分析,主要分析文章引文数量与文献之间的逻辑关系;二是引文线性关联分析,主要分析文献之间的关联和影响;三是文献有向链状分析,主要是引用文献之间有向的引用关系。通过引文分析方法就可以揭示学术成果所在学科的发展脉络、发展历程以及研究热度。

目前,澳大利亚卓越研究项目主要从时间维度(作者、机构、关键词和参考文献)和地理维度(作者、机构和国家信息等)展开。引文分析方法虽然主要是通过客观性评价来评价学术成果的质量,但这种方法没有统一的标准来验证分析结论的有效性和正确性^[34],导致客观性评价结果也受到质疑。当然,澳大利亚卓越项目也关注学术成果有无基金项目、社会影响、社会应用等。

五、重构我国学术成果评价体系的政策建议

学术成果评价的目的决定了评价的方法和手段,那么如何认识我国学术成果评价的目的呢?笔者认为,我国学术成果评价的目的至少体现在以下两个方面:一是发现学术成果的内在价值。内在价值体现于学术成果的学术价值和社会价值,学术价值表现在通过评价发现该学术成果对某领域学术研究的贡献;社会价值表现在通过成果评价发现该学术成果对社会经济发展和科学技术进步的贡献。二是促进人才培养。通过学术成果评价,能够判断人才培养的标准与质量,能够判断人才发展的阶段和水平,进而激发学者从事学术研究的积极性,提升学术成果的质量。要实现我国学术成果评价的目的,就需要在破“四唯”“五唯”政策视角下对我国目前的学术成果评价体系进行改革,破旧立新,建立新的学术成果评价体系。

(一)重新审视我国学术成果评价制度

完善我国学术成果评价制度是实现我国学术成果正确评价的前提,为此,在完善学术成果评价制度方面,我们要坚定“四个自信”,即“中国特色社会主义道路自信、理论自信、制度自信、文化自信”。在借鉴他国学术成果评价成功经验的基础上,结合我国国情,完善我国学术成果评价制度,避免被西方国家制定的学科评价体系所牵制,避免否定国内学术成果展示的平台,避免国内的学术期刊盲从国外学术成果展示平台,破除“唯SCI、SSCI论文”至上的不良倾向。

1. 从国家层面完善符合我国国情的学术成果评价制度

科技部、教育部于2020年连续出台了多项文件,旨在改变目前高校、科研单位盲从SCI、SSCI等外文期刊的局面,当然仅仅停留于这几个文件还远不能改变目前学术成果评价的制度与方法。为此,建议科技部、教育部在联合出台学术成果评价文件的基础上,进一步完善涉及学术成果评价、学科评估等相关文件,发挥国家层面学术成果评价制度的引领作用。

2. 引导和完善教学科研单位完善学术成果评价制度

国家层面制定了学术成果评价制度,还需要引导各高校、科研单位完善科研人员科研考核和评价制度,让学术成果回归本源。从事学术研究的动力源自学者的学术兴趣,学者的学术成果应该对所研究领域的学术研究有推动作用,研究成果有助于推动社会经济的发展。引导教学、科研单位建

立合理的科研奖惩制度,科研成果奖励的确有必要,但不能将学者的科研成果评价与其个人利益全方位挂钩。

3. 完善论文成果质量的相关评价制度

国家提出破除“唯论文”不良导向,不代表学术论文不重要。我们既要扭转片面追求发表SCI论文、SSCI论文的不良局面,又要在国际学术期刊上发出中国的学术声音。既要鼓励在国际顶级学术期刊上发表高水平的学术成果,又要破除“唯论文”“唯SCI论文”“唯SSCI论文”至上的观点与做法。为此,建议完善学术成果评价制度,推行分类评价、定性与定量结合、综合评价制度,实施代表作制度。

4. 成立第三方学术成果评价机构

建议各省市成立第三方学术成果评价机构,这一评价机构最好为非官方机构,评价机构评估工作独立运行、独立完成,第三方评估机构主要负责学术成果的评价,评价结果主要用于教学科研单位的各项审核与评估,包括教学科研单位的学科水平评估、学科专业评价、办学水平评估、博士点申报评估、学术成果奖励等。第三方评估机构的评价结果接受社会监督,国家主管部门负责对第三方学术评估机构的资质进行审核。

(二) 进一步完善我国学术成果评价方法

1. 建立多元化的学术成果评价标准

因学科不同,学术成果的评价标准也存在显著差异。为此,需要对学术成果进行分类,确定该成果属于理工类还是人文类学术成果,确定大类后,根据学术成果所研究的领域划分为基础性和实证性学术成果。不同学科和不同性质的学术成果,要按照一定标准对其创新程度进行分级,基础性学术成果在质量评价时更多强调学术成果的创新性,并延长评价的时效性。对于应用性学术成果,应该更多考虑学术成果数据的有效性、代表性和可信度,成果服务于社会的实际价值。评价方法要灵活掌握,有些学术成果可能介于基础研究和应用研究之间,具体采用何种评价方法,需根据具体情况而定,某些学术成果很难划清是理论研究还是实用性研究,可能还需要将两类学术成果的评价方法综合使用。

值得注意的是,对于多学科交叉学术成果的评价需要采用更为创新的评价方法,包括专家评价、同行评价、社会评价等,甚至还可以采用追溯评价等方法,既要鼓励这类成果的产出,又要评价出交叉学术成果的学术价值和社会价值。

2. 推行“以刊评文”和“以文评刊”相结合的评价方法

“以刊评文”的侧重点在于高水平期刊发文量少、发文周期长,且资源稀缺、竞争激烈,因此更容易在大量文章中挑选出期刊认为高质量的学术论文。“以刊评文”强调的是期刊把握学术论文质量关。“以文评刊”侧重的是带着怀疑的眼光,通过同行评议、学术论文的引文数据等形式对学术论文的内容和社会关注度进行评价,通过上述评价可以侧面反映高水平期刊是否真的发表了高质量学术成果,成果质量是否与期刊相匹配。因此,“以刊评文”和“以文评刊”相结合是评价学术论文质量的一种有效方法,减少了目前科研机构中单纯“以刊评文”的不足。

3. 完善定量评价与定性评价相结合的综合评价方式

定量评价方法和定性评价方法各有侧重,评价的角度存在差异。定量评价方法侧重学术成果参考文献数量、被引数量、下载数量、浏览量等信息,往往忽略学术成果内容的评价,包括论文内容

是否存在创新、论文是否属于原创性等内容;定性评价侧重于主观评价,主要是对学术成果内容的评价,评价结果可能受到评价人的学术水平、评价人与被评价人之间关系的影响。因此,在评价学术成果过程中,根据评价的目的,应不断完善定量与定性评价相结合的综合评价方式,并确定两种评价方法的主次或侧重。

参考文献:

- [1] AVERSA E S. Citation patterns of highly cited papers and their relationship to literature aging: A study of the working literature [J]. *Scientometrics*, 1985, 7(3/4/5/6): 383-389.
- [2] SMITH N M, PRESSER P. Embed with the faculty: Legal information skills online [J]. *Journal of Academic Librarianship*, 2005, 31(3): 247-262.
- [3] BORNMANN L, DANIEL H D. What do citation counts measure? A review of studies on citing behavior [J]. *Journal of Documentation*, 2008, 64(1): 45-80.
- [4] EGGHE L. Theory and practise of the g-index [J]. *Scientometrics*, 2006, 69(1): 131-152.
- [5] 张凌云, 齐飞, 黄晓波, 等. 2003—2014 年我国旅游学术期刊和学术论文评价 [J]. *旅游学刊*, 2015 (12): 85-100.
- [6] PRIEM J, HEMMINGER B H. *Scientometrics 2.0: New metrics of scholarly impact on the social Web* [J]. *First Monday*, 2010, 15(7): 32-51.
- [7] PIWOWAR H. *Altmetrics: Value all research products* [J]. *Nature*, 2013, 493(7431): 159-185.
- [8] 王轶. 中国经济类高被引论文特征研究 [J]. *重庆大学学报(社会科学版)*, 2019(1): 118-128. Doi: 10.11835/j.issn.1008-5831. pj. 2018. 09. 001.
- [9] YOGATAMA D, HEILMAN M, O'CONNOR B, et al. Predicting a scientific community's response to an article [C]// *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics*, 2011: 594-604.
- [10] LIVNE A, ADAR E, TEEVAN J, et al. Predicting citation counts using text and graph mining [C]// *Proc. the iConference 2013 Workshop on Computational Scientometrics: Theory and Applications*, 2013: 16-31.
- [11] 索传军, 盖双双, 周志超. 认知计算: 单篇学术论文评价的新视角 [J]. *中国图书馆学报*, 2018(1): 50-61.
- [12] 赵镇, 陈金源, 曹蓓, 等. 基于美国《科学引文索引》分区的新系数算法对不同学科间学术论文评价方法探讨 [J]. *科技管理研究*, 2019(1): 61-64.
- [13] RAAN A F J. Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises [J]. *Scientometrics*, 1996, 36(3): 397-420.
- [14] 贺颖, 邱均平. 基于科学计量的同行评议专家遴选系统模型构建研究 [J]. *图书情报工作*, 2011(14): 28-31.
- [15] 张积玉. 以量化为基础以代表作为主的综合化学术评价制度构建: 基于 S 大学的经验 [J]. *重庆大学学报(社会科学版)*, 2019(6): 84-96. Doi: 10.11835/j.issn.1008-5831. pj. 2019. 05. 004.
- [16] 刘俏. *Altmetrics* 指标与传统计量指标对学术论文评价的关系研究 [J]. *情报理论与实践*, 2018(7): 60-64.
- [17] MCNAMARA D, WONG P, CHRISTEN P, et al. Predicting high impact academic papers using citation network features [C]// *Trends and Applications in Knowledge Discovery and Data Mining. Springer Berlin Heidelberg*, 2013: 14-25.
- [18] UPHAM S P, ROSENKOPF L, UNGAR L H. Positioning knowledge: Schools of thought and new knowledge creation [J]. *Scientometrics*, 2010, 83(2): 555-581.
- [19] 高锡荣, 杨娜. 基于社会网络分析方法的论文评价指标体系构建 [J]. *情报科学*, 2017(4): 97-102, 144.
- [20] 贺颖, 刘慧, 刘友存, 等. 基于可视化 ACA 技术的同行评议专家学术范式辨别研究 [J]. *图书情报工作*, 2010(14): 28-31.
- [21] SMITH D K. Multi-campus system approaches to academic program evaluation [J]. *New Directions for Institutional Research*, 1980, 27(4): 43-67.
- [22] 马永霞, 仇茹熙. “不唯”≠“不评”: 论人文社会科学成果评价方式的改进 [J]. *重庆大学学报(社会科学版)*, 2021(3): 54-66. Doi: 10.11835/j.issn.1008-5831. pj. 2020. 08. 004.
- [23] 李法惠. 学术论文评价中“过度引用”质疑 [J]. *出版发行研究*, 2009(12): 56-58.

- [24] 刘明. 学术论文评价亟待制度创新[J]. 自然辩证法通讯, 2001(1):7-9.
- [25] 叶青, 彭辉. 人文社科领域学术成果认定与评价方法的研究进展[J]. 社会科学, 2013(3):98-105.
- [26] LAMONT M, MALLARD G. Peer evaluation in the social sciences and the humanities compared: The United States, the United Kingdom, and France[R]. Report for the Social Sciences and Humanities Research Council of Canada (Ottawa: SSHRC), 2005.
- [27] 白如江, 杨京, 王效岳. 单篇学术论文评价研究现状与发展趋势[J]. 情报理论与实践, 2015(11):11-17.
- [28] 李立国. 学术评价与教师治理[J]. 国家教育行政学院学报, 2020(1):9-10.
- [29] PHILLIMORE A J. University research performance indicators in practice: The University Grants Committee's evaluation of British universities, 1985—1986[J]. Research Policy, 1989, 18(5):255-271.
- [30] 王飞. 德国科学委员会对科研成果评价及科研诚信关系的分析与建议[J]. 科学与社会, 2017(1):31-40.
- [31] POSTON L, BOYER E L. Scholarship reconsidered: Priorities of the professoriate[J]. Academe, 1992, 78(4):43-61.
- [32] 邱均平, 任全斌. 国内外人文社会科学科研成果评价比较研究[J]. 国外社会科学, 2007(3):58-66.
- [33] Australian Research Council: ERA 2012 Evaluation Handbook[R]. Canberra: Australian Government, 2012.
- [34] 鲁超, 刘清. 引文分析可视化现状[J]. 情报杂志, 2010(11):48-52.

Research on the assessment of academic achievements from the perspective of the policy of removing the “five standards for profession elevation”

CHEN Min¹, WANG Yi²

(1. Department of Reform and Development, The Open University of China, Beijing 100039, P. R. China; 2. School of Economics, Beijing Technology and Business University, 100048, P. R. China)

Abstract: Recently, the Ministry of Education, the Ministry of Science and Technology and other departments have issued policy documents to abolish the malpractice of “five standards for profession elevation”, with special emphasis on removing the “five standards” orientation in academic assessment. So, what are the problems in the academic assessment system of China, what are the reasons, and how to solve them? After combing the existing literature, we find that: at present, the evaluation of academic achievements in China is susceptible to the rank of authors or units. Personal interests and unit interests are related to the assessments. The units' academic assessment system is affected by policy documents issued by government departments at all levels, and there are blind areas in the evaluation of interdisciplinary academic achievements. After analyzing, the orientation of academic assessment indicators issued by evaluation institutions around the world, the academic assessment systems established at all levels of teaching and research institutions, and the relevant systems for subdividing academic achievements according to disciplines and majors, are causes for the ill practices mentioned above. Drawing on the successful experience of academic assessment in developed countries and taking the purpose of China's academic assessment into consideration, we suggest re-examining China's academic assessment system, so as to improve it and encourage scholars to have more meaningful and high-level academic achievements, where removing the standards of essays only doesn't mean absolutely abandoning essays as an indicator in the assessment.

Key words: academic achievements; academic assessment; removing the “five standards for profession elevation”; essays only; assessment system; evaluation method

(责任编辑 彭建国)