

新时代我国科研评价体系重构中必须处理好八大关系*

邱均平 张裕晨 周子番

摘 要 没有科学的评价就没有科学的管理。目前,我国科研评价体系一定程度上存在着“SCI至上化”、量化考核泛化、同行评审形式化、评价导向行政化、学术研究功利化等评价失衡问题。在破除“SCI至上”等科研评价顽疾背景下,科研评价体系重构是适应新时代我国科研管理和科技创新发展的迫切需要。重构科研评价体系必须正确处理好八大辩证关系:管理评价与学术评价的关系、定量评价与定性评价的关系、工具性与目的性的关系、分类评价与统一评价的关系、评价理论与评价实践的关系、数量与质量的关系、论文评价与评价的关系、元评价与评价的关系。这些思考有助于更加理性、辩证地认识科研评价存在的问题,从而更全面地把握科研评价的科学内涵、要素、原则、作用,以便更好地指导科研评价实践活动。建议运用评价学理论指导科研评价实践,区分管理评价与学术评价,坚持分类导向、综合导向、质量导向,围绕评价目的合理使用论文指标以开展科学评价,重构科学合理的、适应新时代要求的科研评价体系。图3。表2。参考文献31。

关键词 科研评价体系 评价理论 评价指标 评价要素

分类号 G250.2

Eight Relationships in the Reconstruction of Scientific Research Evaluation System in China in the New Era

QIU Junping, ZHANG Yuchen & ZHOU Zifan

ABSTRACT

Without scientific evaluation, there is no scientific management. At present, our country's scientific research evaluation system has various issues such as "SCI First", peer review formalization, simplification of quantitative assessment, evaluation-oriented administration, and academic research utilitarianization. In the background of eliminating the stubborn problems of scientific research evaluation such as "SCI First", the reconstruction of the scientific research evaluation system is an urgent need to be adapted in order to meet

* 本文系国家社会科学基金重大项目“基于大数据的科教评价信息云平台构建和智能服务研究”(编号:19ZDA348)和教育部人文社会科学项目“高考综合改革试点政策评估与完善措施研究——基于后实证主义方法论视角”(编号:19YJC880138)的研究成果之一。(This article is an outcome of the major project “The Research on the Construction of a Cloud Platform for Science and Education Evaluation and Intelligent Service Based on Big Data” (No. 19ZDA348) supported by National Social Science Foundation of China and the project “Study on the Policy Evaluation and Improvement Measures of the Comprehensive Reform Pilot College Entrance Examination: Based on the Perspective of Post-positivism Methodology” (No. 19YJC880138) supported by Humanities and Social Science Foundation, Ministry of Education of the People's Republic of China.)

通信作者:张裕晨,Email: zyc@hdu.edu.cn, ORCID: 0000-0002-7380-2915 (Correspondence should be addressed to ZHANG Yuchen, Email: zyc@hdu.edu.cn, ORCID: 0000-0002-7380-2915)

the development of scientific research management and technological innovation in China in this new era. To reconstruct the scientific research evaluation system, eight dialectical relationships must be correctly handled: 1) the relationship between management evaluation and academic evaluation, 2) the relationship between quantitative evaluation and qualitative evaluation, 3) the relationship between instrumentality and purpose, 4) the relationship between classified evaluation and unified evaluation, 5) the relationship between evaluation theory and evaluation practice, 6) the relationship between quantity and quality, 7) the relationship between paper evaluation and evaluation, 8) the relationship between meta-evaluation and evaluation. These rational thinking can promote a more rational and dialectical understanding of the current problems existing in scientific research evaluation, which can lead to more fully grasp the scientific connotation, elements, principles, and functions of scientific research evaluation, so as to better guide the practical activities of scientific research evaluation. We suggest using evaluation theory to guide the practice of scientific research evaluation, distinguishing management evaluation from academic evaluation, insisting on classification-oriented, comprehensive-oriented, and quality-oriented evaluation, with the purpose of evaluation using paper indicators reasonably to implement scientific evaluation, reconstructing the scientific research evaluation system which is not only scientific and reasonable but also meets the new era's requirements. 3 figs. 2 tabs. 31 refs.

KEY WORDS

Scientific research evaluation system. Evaluation theory. Evaluation index. Evaluation element.

0 引言

有人类活动的地方总是伴随着评价活动的发生,科研评价对科研活动具有判断、预测、选择、导向、诊断、激励和合理配置资源等作用^[1]。自美国 1920 年开始科研评价以来,许多国家也陆续开展了系统的科研评价工作,并建立了较为完整的科研评价体系^[2]。其他西方国家大多效仿美国模式,成立专门评估机构。伴随着竞争性科技项目的设立,我国科研评价体制于 1985 年正式启动,并得到迅速发展。从 1993 年开始,中央及相关部委陆续发布了 20 余项涉及科研评价的文件(见表 1),有针对性地提出了科研评价的原则、目标和要求,对促进科技资源优化配置、提高科技创新能力、提升科技管理水平、保障我国科研事业健康发展起到了积极作用^[3]。

与此同时,我国科研评价体系也暴露出一些弊端,甚至产生评价失衡的系列问题,引起了

科技界广泛关注。科研评价失衡问题概括起来主要有四个方面:一是“SCI 至上化”,评价时误用 SCI、SSCI 的论文指标,过分看重论文的数量、期刊影响因子,而忽视了论文本身的质量;二是量化考核泛化,定量评价强调以各种数量化、绝对化的数字指标为核心,替代学术评价机制,导致学术研究急功近利;三是同行评审形式化,程序失范、人情关系、形式主义、本位主义等因素造成同行评审流于形式;四是评价导向行政化,以行政为主导的科研评价与学术评价产生错位,导致评价方法、评价指标的不合理使用。“四化”相叠最终导致了学术浮躁和学术研究功利化。国内有很多学者对科研评价问题产生的原因和对策进行了研究,比较有代表性的观点认为:现行的科研评价体系问题主要集中于评价的不公平、不公正、不公开、不准确四个方面^[4];政府在科技管理中的缺位和越位,以及科学共同体由于自身缺陷而未能有效发挥作用是造成当前科研评价问题的根本原因^[5-6];改革不合理的科研体制是终止学术之乱和评价之乱的

表 1 科研评价相关文件

年度	发布机构	发布内容
1993	全国人民代表大会常务委员会	中华人民共和国科学技术进步法
1996	全国人民代表大会常务委员会	中华人民共和国促进科技成果转化法
2000	科技部	科技评估管理暂行办法
2001	中共中央办公厅、国务院办公厅	关于加强专业技术人才队伍建设的若干意见
2003	科技部	科学技术评价办法(试行)
2006	国务院	实施《国家中长期科学和技术发展规划纲要(2006—2020年)》的若干配套政策
2009	科技部	科技成果评价试点暂行办法
2011	科技部、教育部、人社部等	国家中长期科技人才发展规划(2010—2020年)
2012	中共中央、国务院	关于深化科技体制改革加快国家创新体系建设意见
2013	教育部	教育部关于深化高等学校科技评价改革的意见
2014	教育部	高等学校科技分类评价指标体系及评价要点
2015	中共中央办公厅、国务院办公厅	深化科技体制改革实施方案
2016	中共中央	关于深化人才发展体制机制改革的意见
2017	中共中央办公厅、国务院办公厅	关于深化职称制度改革的意见
2018	中共中央办公厅、国务院办公厅	关于深化项目评审、人才评价、机构评估改革的意见
2018	国务院	关于优化科研管理提升科研绩效若干措施的通知
2018	科技部、教育部、人社部、中科院、工程院	关于开展清理“唯论文、唯职称、唯学历、唯奖项”专项行动的通知
2018	教育部	关于开展清理“唯论文、唯帽子、唯职称、唯学历、唯奖项”专项行动的通知
2019	中共中央办公厅、国务院办公厅	关于进一步弘扬科学家精神加强作风和学风建设的意见
2020	科技部	关于破除科技评价中“唯论文”不良导向的若干措施(试行)
2020	教育部、科技部	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

根本途径^[7]。综合来看,现行科研评价体系的失衡有主客观两方面的原因,具体来看则存在着若干对关系的失衡,包括管理评价与学术评价的关系、定量评价与定性评价的关系、工具性与目的性的关系、分类评价与统一评价的关系、评价理论与评价实践的关系、数量与质量的关系、论文评价与期刊评价的关系、元评价与评价的关系等的“失衡”,这些矛盾最终导致了科研评价“失衡”和评价功能“异化”。重构科研评价

体系必须正确处理好这些辩证关系。

表面上看,科研评价只是科技体制中的一个局部问题,但实际上对科技创新和科技发展具有全局性的影响。因为科研评价不仅是一个社会科学价值观念的直接体现,而且是科技管理的重要手段,评价的过程和结果直接影响科技资源的分配,评价的标准更是引导科学家科研行为的指挥棒。科研评价问题既是学术界讨论和研究的热点问题,也是科技管理部门一直

想解决而未能彻底解决的关键问题。习近平总书记指出“人才评价制度不合理,唯论文、唯职称、唯学历的现象仍然严重”,强调“要改革科研评价制度,建立以科技创新质量、贡献、绩效为导向的分类评价体系”。2020年2月,科技部、教育部连续出台《关于破除科研评价中“唯论文”不良导向的若干措施(试行)》(以下简称2020年《若干措施》),《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》(以下简称2020年《意见》)两个文件,直指科研评价中“唯论文”“SCI至上”的不良导向,在科技界引起了热烈讨论。

学术界对破除现有评价顽疾达成了广泛共识,但对如何构建一个科学合理、符合新时代要求的科研评价体系,仍需要不断探索。科研评价是一个复杂的科学系统,在破除“SCI至上”等科研评价顽疾的背景下,需要从科学评价的视角,建立起适合新时代我国科技发展体制的评价体系。本文从新时代特征和要求出发,运用评价学原理,对重构我国科研评价体系必须处理好的若干关系进行探讨,并给出了相应对策,希望能引发同行的全面思考,共同推动我国科研评价事业的健康发展。

1 正确处理管理评价与学术评价的关系,倡导区别对待原则

管理评价与学术评价是评价领域两个不同类别的评价,需要区别对待。学术评价是以科研活动或科研成果为对象的评价,包括对论文、项目、奖励等的评审^[8]。管理评价属于工作评价,是以学术评价结果为指标的评价,包括管理部门以一定的价值取向对人员、机构进行的绩效评价。两者在评价主体、评价对象、评价标准、价值取向向上都有显著不同。在当前科研管理中,“学术评价”和“管理评价”混淆不清,是导致量化考核泛滥、量化评价被误用的主要原因。近年来各科研评价领域“SCI至上”与这样的量化管理有着直接的联系。从管理的角度看,定量指标确实显

得更加“客观公正”,但有时却很难完全反映事实和价值。原本只具有相对参考意义、各种数量化的宏观学术评价指标被绝对化,进而又层层下放给机构、个人,在各种学术评估、评比、鉴定、考核、申报中,统统都要求提供一定的数量指标作为门槛^[9]。数量的堆砌却没有带来质量的提升,反而会导致“学术GDP”盛行,这样的评价方式最终也就异化成了功利主义。

正确处理管理评价与学术评价之间的关系,对两者的评价对象、评价标准、评价方法等作正确区分,是很有必要的。学术评价从根本上说,是学术性质的;而管理评价从根本上说,是一种管理上的常规激励手段。二者虽有联系但区别明显。把管理考核那一套搬来,用到学术评价上去,多半是不大合适的。在管理评价中,上层管理机构应以宏观管理为主,做好学术规范、学术制度建设;基层管理部门应做到服务与管理并重;对管理部门的考核也应该以此为重点。对科学研究人员和研究机构绩效、实力与水平的评价,应注重学术性,即以学术成果质量、标志性的成果产出为主要评价依据。在学术评价中,鼓励运用同行评议、文献计量等定性定量方法,对学术成果等进行综合评价,注重内容和效果评价。在评价实践中,评价主体需要提高对评价本质、特性和规律的认识,正确区分管理评价与学术评价,自觉地强调学术评价的“学术性”,把学术创新、研究成果质量摆在首要位置,避免不同评价方法的盲目“移植”和误用。

2 正确处理统一与分类的关系,倡导分类评价原则

分类评价是评价相对性与统一性的集中体现。狭义的分类评价,主要是指对评价对象进行分类,以此为基础开展评价工作。广义的分类评价,是按评价活动所涉及的各要素进行分类,包括按评价方法、评价主体、评价对象、评价时间、评价目的、评价标准、评价保密性、评价结果的表现形式分类(见图1)。科学区分评价的

不同类型,对于落实科研活动“分类评价”原则具有重要的意义。评价对象指的是被评价的客体。按学术评价所涉及的具体对象,可以将科研评价分为成果评价、人员评价、机构评价、项目评价、计划评价、学科研究评价、科学政策评价等七类;根据评价对象所涉及的范围大小,可以分为宏观、中观与微观学术评价^[9];根据学术研究所涉及的学科范围,可分为自然科学研究评价、人文社科研究评价两类;根据所涉及研究工作的层次,可以分为基础研究评价、应用研究

评价等。随着我国评价事业的不断发展,分类评价原则在各领域得到了不同程度的贯彻,如学科评估、双一流建设成效监测、大学评价、职称评聘等。但分类评价仍存在着分类标准模糊、分类评价失实等问题。如学术期刊的评价,基础研究领域影响因子会高于应用研究领域。不同学科的期刊也会存在较大差异,如生物化学类期刊最高影响因子会显著高于数学类期刊。所以,简单用影响因子这一标准去衡量不同领域的期刊,将难以得到合理的评价结果。

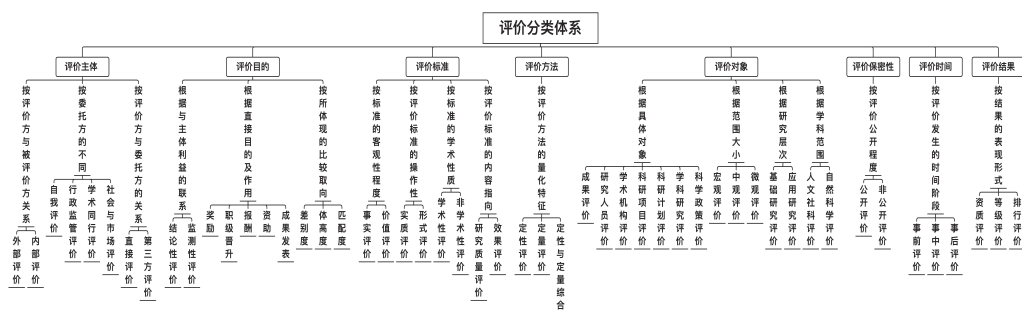


图1 评价分类体系

分类评价原则一直是我国科研评价体制改革的重点,在中央及相关部委发布的20余项涉及评价的文件中(见表1),提到“分类评价”的就有16项,占比超过76%。这些文件对分类评价的对象分类、具体岗位、评价标准都给出了具体的规定,但在具体的科研评价活动中,距离文件要求仍有很大的距离^[4]。分类评价是科研评价的重要原则,也是开展科研评价的基础。构建科学合理、适应新时代要求的科研评价体系必须正确处理统一评价与分类评价的关系。在最新的两份文件中也能体现这一原则,如2020年《若干措施》指出:“要建立健全分类评价体系,对不同类型的科研工作应分别建立各有侧重的评价路径。”2020年《意见》也提出:“要强化分类考核评价导向,实施分类考核评价,注重标志性成果的质量、贡献和影响。”分类评价仍将是新时代我国科研评价体系重构必须一以贯之的重要原则。

3 正确处理工具与目的的关系,倡导目标引领原则

评价本身具有较强的价值取向,这种价值取向可以通过评价目的和评价手段来体现。目的价值取向侧重于引导科学研究方向,促进学科发展;工具价值取向则倾向于将科研质量评价作为奖优罚劣的手段或工具^[10]。目的价值与工具价值统一于科研评价中,两者不可分割。但在实际工作中,人们往往重视科研评价的工具性,而忽视其目的性,导致“目的”与“手段”常常颠倒、混淆^[11]。科研评价的目的在于促进管理,引领科技活动的发展方向,引导科研人员投身科技创新,产出优秀成果。但是从现实中的实际效果来看,用以评价学术成果、学术绩效的“工具”,反倒变成了人们竞相追逐的“目的”,如此容易导致“唯论文”“唯数量”“急功近利”等

弊病^[12]。根据评价目的与主体利益的联系,可以将评价分为结论性评价和监测性评价两类(见图1)。在机构评价中,结论性评价基于历史统计的现实影响对绩效作出评判,并与其他类似的单位作比较,评价结果反映研究水平高低,并成为拨款的依据;监测性评价是面向未来的潜力前瞻性评价,只是为研究机构获取更大回报提供信息^[13]。如普赖斯通过统计SCI作者发文量和被引总次数,为预测诺贝尔奖获得者提供了许多有价值的参考信息,属于监测性评价范畴;瑞典皇家科学院组织的评审最终决定这些科学家能否获得诺贝尔奖,属于结论性评价的范畴。

科研评价中“工具至上”主义之所以盛行,一个重要原因在于对学术评价的本质认识不足,对评价的要素分析不够,尤其是对目的价值认识不明确,以致于造成对“工具”的误用。科研评价涉及许多方面,至少包括了十大要素:评价主体、评价目的、评价对象、评价标准、评价内容、评价过程、评价方法、评价指标、评价数据、评价结果。其相互关系(见图2)如下:评价主体与评价对象是评价过程中的实体性要素;评价目的是主导因素,主要是由评价主体中委托方的需求决定;评价对象本身的客观规律性,也在一定程度上决定评价目的的可能指向;评价标准与指标是核心,是学术目的的具体化;评价方法是手段,由评价方根据实际情况选定;对评价主体的认识会影响对评价目的、评价需求的认识;对评价目的、需求认识不清,显然会使学术评价的结果“答非所问”,无法达到预期的目标。任何工具如果不服务于一定的主体和目的,就失去了意义;如果工具服务于错误的主体和目的,就会造成损害。事实上,如果单单为科研行政管理部门考虑其管理上的方便性,而采取某些评价标准、方法、指标和程序,就很容易偏离本来目标,使科研评价丧失其“科学性”,沦为简单的行政管理考核。因此,构建科研评价体系需要正确处理目的与工具的关系,以目的引领评价导向。只有正确认识了学术评价的真正目

的,才能做到有的放矢,使科研评价真正服务于科学研究事业的发展。

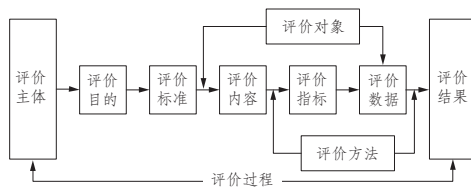


图2 评价要素关系

4 正确处理数量与质量的关系,倡导质量优先原则

数量与质量是事物的不同表现形式,是评价价值导向的两个方面。重数量轻质量一直是科研评价领域的突出问题。论文和经费数量经常被当作衡量科学家科研水平的唯一标尺^[14],职称评定和人才选拔也是按照一定标准的数量进行排队,许多大学排行榜更是演变成各种数量指标的盲目竞赛。本来量化考核作为一种日常管理手段并不是什么坏事,有助于奖勤罚懒,形成一定的激励机制,但一旦将其简单化、绝对化,变成数字的加减乘除,或者干脆用量化考核替代了学术界惯有的认同机制,就会严重影响对科研机构 and 科研人员的评价导向,必然会导致学术研究的急功近利。

构建科学合理的科研评价体系需要正确处理好质量与数量的关系,纠正以往过分看重数量指标、忽视学术研究质量的现象,树立质量优先的观念,先讲质量、再看数量。具体地说,就是在科研评价中,要提高质量指标的地位,加强质量指标的权重,而相对削弱对数量指标的要求。但这并不是否认数量指标的作用,质量是需要建立在一定数量基础之上的。数学中的大数定律就说明了“在随机事件的大量重复出现中,往往呈现几乎必然的规律”。进一步讲,某些数量指标本身就包括了质量属性。比如学术期刊论文本身就是要经过同行评审后才能发表的,是具有一定质量要求的。高水平期刊,其论

文录用率大概只有 5% 左右。为了扭转重数量轻质量的评价导向,近年来中央部委制定的文件中都对数量与质量关系作出了要求(见表 1):《关于深化项目评审、人才评价、机构评估改革的意见》提出要减少“三评”项目数量,改进评价机制,提高质量效率;《关于优化科研管理提升科研绩效若干措施的通知》提出要开展“唯论文、唯职称、唯学历”问题集中清理,建立以创新质量和贡献为导向的绩效评价体系。

代表作评价制度是不以学术成果的形式、级别和数量,而以科学研究机构、个人的代表性成果、成绩来衡量其工作业绩的方法^[15]。而评价“代表作”的标准,只能是学术成果的内容、研究质量和创新价值。代表作评价制度从程序上、根源上就是忽略数量指标的,所以对于矫正当前量化考核泛滥、学术浮躁、学术泡沫化的“症状”,的确是一剂“良药”。在国内诸多学者中,尤其是那些激烈反对量化考评体系的学者中,提倡“代表作”评价制度的呼声很高。宋丽萍^[16]认为,以质量为导向的、同行评议与文献计量相结合的方法将成为未来科研评价的主流。国际上也将代表作评价制度用于科研项目、科研机构等评价中,如加入了“内部控制质量”指标或“质量控制情况”指标的德国科研评价模式,采用具有质量控制功能指标的英国模式,强调质量导向评价维度的瑞典模式^[17]。

5 正确处理定性与定量关系,倡导综合评价原则

定性与定量是评价方法体系里的一种分类方式。评价方法是指处理客观评价对象及其有关数据资料的方式,总体上可分为:定性评价方法与定量评价方法,直接评估与间接评估,同行评议与引文计量、指标打分等。定性评价主要是指同行评议制,包括通讯评审法和现场评审法或会议评审法。从理论上说,双盲的通讯评审刻意回避了人情关系的影响,符合公正性的要求;会议评审实行民主投票,也不至于一言堂

或个别人说了算。理想的同行评议应以评价对象的能力或水平为依据,但在实践中,主观性等多种因素影响着同行评议的客观性。同行评审给人的感觉就是“走过场”和流于形式,其原因包括人情关系、本位主义、评审程序不合理、评审时间短、评审数量多等。定量评价方法主要有引文计量法和指标打分法。这些评价方法具有以下优点:一是客观性较强,尽量排除了主观因素的影响;二是具有形式上的定量性、精确性,便于不同层次、不同种类的对象之间进行相互比较。另外,由于大量地依据对象外在特征如课题来源、期刊级别、被引次数等进行计量,还具有便于公开、易于操作、成本较低等优点。在使用定量指标时,一定要对其有一个科学理性的认识:定量指标对定性评价有一定的支撑作用,但单一的定量指标用于科研评价有较大的不确定性与风险性。

定性评价与定量评价各有优势。在实际的评价实践活动中,应该将两者结合起来,互为补充,形成综合评价方法。定性是定量的前提,使得定量评价有据可循;定量反过来可以优化定性,使得定性评价更加客观、精准。在科研评价中应充分认识各类评价的目的和需求、评价对象的特征与规律等,选择合适的评价方法、评价程序、评价指标。对于国家竞争力评价、地区竞争力评价、大学评价、学科评估等宏观评价,可以采用定量评价方法;对于人才评聘、论文和项目评审等微观评价,采用定性评价为主、定量指标为辅助的综合评价方式更为妥当。关于定量与定性评价问题,国际上于 2014 年达成了共识,发布了“莱顿宣言”(The Leiden Manifesto for Research Metrics)^[18],对定量评价提出了十条原则(见表 2)。“莱顿宣言”之所以得到科研评价界专家学者的普遍认同,最大的原因在于它对“量化至上”的现象进行了纠正,如“量化评估应当支撑质化的专家评估,而不是取而代之”“对个人研究的评价应基于其综合作品的质性评价”,这两大原则对新时代我国科研评价体系重构仍有很强的现实指导意义。

表2 莱顿宣言和旧金山宣言要点

莱顿宣言十原则 (Leiden Manifesto)	旧金山宣言 (DORA)	
(1)量化的评估应当支持而非取代质化的专家评审	基本原则	(1) 不使用基于期刊的指标去评价单篇文章的质量、单个科学家的贡献,或在招聘、晋升和资助的决策中使用。
	针对资助	(2) 明确用于评价资助申请人的科学生产力的标准,明确强调论文的科学内容远远比期刊的计量数据或发表期刊的地位更重要,尤其是对职业生涯早期的研究者而言。
(2)科研绩效的考量应基于机构、团队以及个人的科研使命		(3)为了达到评价目的,除了发表的正式研究成果,也要考虑所有研究成果(包括数据集、软件)的价值和影响力,并考虑更广泛的影响力评价方法,应该包括一些定量指标,比如评价其对政策、生产实践的影响力。
(3)保护卓越的本地化的相关研究	针对机构	(4) 明确用于达成聘用、任期和晋升决策的标准,明确强调论文的科学内容远远比期刊的计量数据或发表期刊的地位更重要,尤其是对职业生涯早期的研究者而言。
		(5) 同第3条。
(4)保持数据采集和分析过程的公开、透明和简单	针对出版商	(6) 极大地减少、最好是停止宣传期刊影响因子,或者将它与一系列的期刊评价指标(5年影响因子、特征因子、SCImago、h指数、编辑及出版时间等)一起提供一个更丰富的期刊表现视图。
(7) 提供一系列的论文级指标,以鼓励转向基于论文科学内容的评估,而不是基于期刊的评价指标。		
(5)允许被评估者验证数据和分析	针对提供指标的组	(8) 鼓励作者署名负责,并提供每一名作者的专业贡献。
(6)考虑发表和引用的学科差异		(9) 无论一个期刊是开放获取或收费订阅,移除所有对重新使用文献、列入参考文献表的限制,并使它们符合知识共享协议。
(7)对个人研究的评价应基于其综合作品的质性评价	针对研究人员	(10) 移除或减少对研究论文的参考文献数量的限制,并在适当的情况下强制引用原始文献,以支持评论,并给予首次发现成果的小组以认可。
		(11) 公开透明,提供所有的计算数据及方法。
(8)应避免评估指标的的不当的具体性和虚假的精确性		(12) 无限次地提供数据,并在可能的情况下提供对数据的计算访问。
(9)识别认清评价指标对科研系统的影响		(13) 明确不允许不适当的指标操作;明确什么构成不当操纵,以及将采取什么措施来打击这种行为。
		(14) 当使用、汇总或比较指标时,说明文章类型的变化(例如,评论与研究文章),以及不同主题领域的变化。
(10)定期审查评价指标并加以改进		(15) 当参与委员会制定有关资金、招聘、任期或晋升的决策时,应根据科学内容而非出版标准进行评估。
		(16) 在适当的情况下,引用原始文献而非综述,给予作者应有的认可。
		(17) 在个人/支持性陈述中可以使用一系列文章指标,以作为个人发表文章和其他研究成果影响的证据。
		(18) 挑战不适当地依赖期刊影响因素的研究评估实践,推广和教授注重于专业研究成果价值和影响的最佳实践。

6 正确处理论文与评价的关系,倡导论文适用性原则

近年来,我国科研评价中存在着过度看重论文数量及期刊影响因子、忽视科技成果质量的“唯论文”弊病。特别是 SCI 论文数量、被引次数、高被引论文、影响因子以及衍生出的 ESI 排名等相关指标,已经成为了科研评价以及职称评定、绩效考核、人才评价、学科评估、资源配置、学校排名等方面的核心指标,使得科技工作出现了过度追求 SCI 论文及相关指标,甚至以发表 SCI 论文数量、高影响因子论文、高被引论文为根本目标的异化现象。在这样的“SCI 至上”评价背景下,我国 SCI 论文数量一直快速增长。通过检索 Web of Science 数据库发现,2019 年我国发表 SCI 论文 54.8 万篇,约为美国(67.9 万篇)的 80.7%,连续多年位居世界第二。在一些世界著名大学排名指标体系中,也存在着“SCI 至上”的现象,如 ARWU 世界大学学术排名有近 40%的指标与论文数或论文折合数有关,THE 排名中论文指标也占到了 38.5%^[19]。高论文数带来了我国大学所谓的“高排名”,一些高校在世界大学排行榜中节节攀升,某些学科甚至排到了世界第一,比如清华大学计算机学科曾在近年 ARWU 世界大学学术排名中力压美国诸多顶尖高校排名第一。令人遗憾的是,排名的飙升、论文数量的井喷只带来了论文虚假繁荣,并没有带来我们国家“卡脖子”领域质的突破。

包括加菲尔德本人在内的科学计量学专家,对于如何运用 SCI 进行科研评价都会有适当的提醒:即使将 SCI 作为评价依据,论文本身质量、论文被引用率也应该是重要的参考因素。而国内的大部分情况却是,对科研人员进行评价、奖励,只看论文是否发表在 SCI 源刊上,是否被 SCI 索引,或者计算源刊的影响因子积点等,这种做法很明显是对 SCI 的误用。对于科研评价“唯论文”“SCI 至上”的现象,许多学者都强

烈反对。针对国内外出现的这种“以刊评文”现象,学者们于 2012 年 12 月在美国细胞生物学学会年会期间提出了著名的“旧金山宣言”(The San Francisco Declaration on Research Assessment, DORA)^[20],引发科学界的热烈讨论。“旧金山宣言”共列出了 18 条面向学者、机构、研究指标提供者、基金会的建议(见表 2),其中最重要的是“避免使用期刊影响因子去评价单篇论文或作者”。与“莱顿宣言”纠偏“重定量轻定性”不同,“旧金山宣言”针对“以刊评文”提出了一些建议,认为过度关注影响因子会导致科学家片面追逐热点,而非潜心科研成果创新。正如国外学者 Moed^[21]所言,研究成果及其影响是多维概念,当用论文、引用、期刊影响因子、社交媒体等指标来评估个人和团队时,必须受到严格限制,因为针对科研人员个人的评价指标已经被证明存在“虚假的精确性”,期刊影响因子的评价有效性尚缺乏足够的计量学证据,基于社交媒体的指标只能作为补充措施。我国学者韩毅^[22]也认为替代计量学(Altmetrics)的计量结果与传统以引文为基础的科研评价是相互补充关系,两者融合进行综合评价是必由之路。俞立平^[23]认为宏观评价要发挥 SCI 论文的主导作用,微观评价要兼顾代表作与 SCI 论文数量。

需要指出的是,破除“唯论文”,并非不看论文。SCI 之所以成为定量评价的标准,可能只是因为其有一套选刊的标准程序,有一定的门槛,所以我们把它当作甄别工具。尽管存在着对 SCI 的种种批评,但不容否认,这种定量评价的形式仍然有许多优点,如形式上的公正,评审指标和评审程序很容易公开,便于公示和接受监督,操作方便,容易实现自动化处理,评审效率较高等。事实上,学术期刊在录用论文之前,都需要经过同行专家的评审,这在一定程度上保证了录用论文的质量,毕竟录用的比例是比较低的,尤其是高水平期刊,在这方面的评审更为严格。所以,完全放弃类似 SCI 的定量评价方法,目前不仅不现实,而且还可能是危险的。在

学术规范尚未得到普遍遵循、行政主导学术的局面尚未完全改观的情形下,贸然放弃具有公正形式的定量指标评价方法,可能导致学术评价中更广泛的不公正。SCI 论文被用于科研评价备受诟病之处在于“SCI 至上”,即过度依赖 SCI,而不在于发表 SCI 论文。因此,新时代要破除的是“唯论文”中的“唯”,而不是论文本身。

7 正确处理元评价与评价的关系,倡导科学评价原则

元评价是对评价的评价。在我国目前的科研评价体系中,还缺少一种元评价机制。我国评价事业蓬勃发展,但评价的效果如何?评价的程序是否规范?还有哪些值得改进之处?甚至具体到某些评价是否有必要?这些问题值得我们反思。元评价机制的主要目的是形成规范,对评价主体尤其对评价方形成制约,从而提高评价质量。从元评价的类型上看,可以从对同行评审专家的再评价、对第三方机构的再评价角度来区分^[24];也可以对评价主体、评价方法、评价程序、评价指标等评价要素进行元评价^[25]。元评价的方法主要有内容分析法、经验总结法、评价信度效度分析等^[26]。元评价的实施主体既可以是评价主体,也可以是评价委托方;既可以从内部进行评价,也可以从外部进行评价。元评价是提高评价质量的有效手段^[27]:对评价流程的元评价可以发现问题,改善评价;对评价结果进行元评价,有利于改善评价结果的公布方式,提高评价质量;对评价方法进行元评价,有利于改进评价技术和方法,淘汰不适宜的评价方法;对评价目的进行元评价,有利于评价组织者反思评价本身,并加强对科研评价自身的管理。评价的次数、周期等,都应该符合科技活动发展的规律,过多或过少都是不适宜的。

缺乏对评价方法和指标体系的元评价,是许多科研评价项目结果不准确的重要原因。传统的评估方法是分项打分、加权平均、求综合分

的综合指数法,又称指标打分法,这是目前学术评价领域应用最成熟、最普遍的一种方法。常用的指标打分法,是以评估指标之间的线性关系为假设前提来进行的,但在实际的评价实践工作中存在着非线性的性质,不适宜用统一的线性评价方法^[28]。况且,指标的完备性和互斥性(不相关性)是理论上建立指标的准则,这是由其线性的假设前提所规定的。然而在实践中,由于指标概念模糊不清,指标划分往往不具备完备性和互斥性。指标与指标之间存在函数关系或其他间接的相互依赖关系,这种“多重共线性”都是需要尽量避免的。当前学术评价中流行的指标打分法,看似客观、公允,但其内在的线性逻辑却存在着许多固有的缺陷。这并不等于说该方法不能用,而是需要对其指标体系进行元评价,注意其假设前提,不要陷入“工具法则”的陷阱之中。正确处理元评价与评价之间的关系,建立元评价机制,对这些评价程序、方法、指标形成稳定的指导性规范,将有助于解决当前科研评价中出现的科学性问题。

8 正确处理理论与实践的关系,倡导理论指导原则

评价理论是指人们由评价实践概括出来的系统性的概念、原理和结论^[29]。我国在科研评价实践活动中积累了丰富的经验,也取得了许多积极成果,但总体上评价理论基础薄弱,评价制度有待健全。虽然围绕科研评价方法、流程、指标体系、数据处理、信息集成、结果运用等内容,目前已形成了较为科学的理论体系,但是评价活动是一个复杂的过程,评价系统作为社会活动大系统中的一个子系统,涉及的理论问题众多,我国评价基础理论研究仍严重滞后于各类评价实践活动。长期缺乏系统的评价理论指导,是当今各领域评价问题未能有效解决的重要原因。我国科研评价基础理论研究总体滞后,评价领域研究的重大课题、重要专著等研究成果鲜有突破,而科研评价在我国起步相对较

晚,许多评价理论是“拿来主义”,与我国科研评价具体实践差距较远,难免会有一些水土不服。再加上理论偏向于实际操作性,导致指导性不足等问题凸显。

完整的科研评价理论与方法体系由评价方法论、评价理论基础和基本理论模型、各种具体的科研评价方法及应用模型、实施过程管理与控制的方法和规范四个部分组成^[30]。科研评价实践需要评价理论的正确指导:从评价问题的界定到评价方案的设计,再到评价信息的收集过程,并不是随意进行的,而是要依据一定的理论基础(见图3);在评价方法层面,从方法论指导到评价工具选择也必须有理论依据。应该说,国内外评价基础理论的研究,为促进我国科研评价实践发展起到了巨大的促进作用。例如,邱均平教授等著的《评价学:理论·方法·实践》(上、下册)^[31]对评价理论问题进行了全面、系统和富有创见的探讨,是我国首次以“评价学”命名的开山之作,并首次构建了该学科的理论、方法、应用三结合的内容体系,具有开创性和奠基性意义。但每一种理论的发展,都带有其所处的时代特征,尤其是信息技术、智能技术等新技术的发展,必将极大地丰富科研评价实践的内容,从而促进评价基础理论体系的完善。

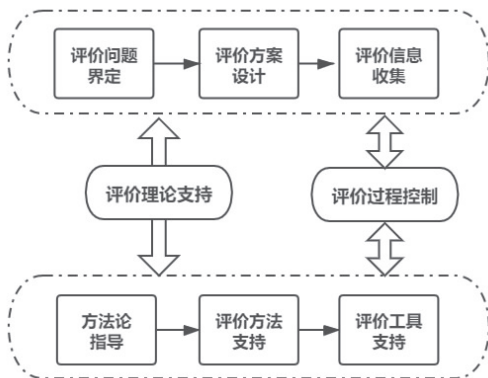


图3 基于评价活动的评价理论体系

9 结语

科研评价作为一项具有社会科学价值的科技管理实践活动,最终目的在于促进科技活动的健康发展。其作用就在于根据委托方明确的目的,通过科学、可行的方法,依据规定的原则、程序和标准,对科技活动各个方面的事项进行论证、评审、评议、评估和验收,确定科技活动成效状态和发展趋势。评价是个系统工程,涉及评价主体、评价目的、评价客体、评价方法等十个方面的要素。在破除“SCI至上”、重构科学的科研评价体系时,就要紧紧围绕这些要素和目标展开,以评价理论为指导,以适应新时代我国科技发展水平为重要任务,探讨现行科研评价体系、方法、原则及实践是否遵循了科技创新活动的规律,体现管理评价与学术评价的基本价值理念,催生出真正高水平的科技成果;是否发挥了“评价指挥棒”应发挥的引领作用,促进科学技术的发展,提升创新能力水平,促进经济社会发展。

没有科学的评价就没有科学的管理。科研评价作为科学管理的重要手段,首先需要深入分析其失衡的根源性问题,最重要的是要处理好八个辩证关系,倡导八大原则:正确处理管理评价与学术评价的关系,倡导区分对待原则;正确处理统一与分类的关系,倡导分类评价原则;正确处理工具与目的的关系,倡导目标引领原则;正确处理数量与质量的关系,倡导质量优先原则;正确处理定性与定量的关系,倡导综合评价原则;正确处理论文与评价的关系,倡导论文适用性原则;正确处理元评价与评价的关系,倡导科学评价原则;正确处理理论与实践的关系,倡导理论指导原则。当然,我们在审视现有评价体系各种不足的时候,还要正确对待其好的一面,避免矫枉过正。科研评价体系改革和重构应该是一个循序渐进、不断优化的过程。

参考文献

- [1] 汤建民,邱均平. 评价科学在中国的发展概观和推进策略[J]. 科学学研究,2017,35(12):1813-1820,1831. (Tang Jianmin, Qiu Junping. Development situation and prompting strategies of evaluation science in China [J]. Studies in Science of Science, 2017, 35(12): 1813-1820, 1831.)
- [2] 赵蓉英,王旭,王建品,等. 中国科技评价研究态势与推进策略[J]. 情报科学,2019,37(5):7-12. (Zhao Rongying, Wang Xu, Wang Jianpin, et al. Research situation and prompting strategies of S&T evaluation in China [J]. Information Science, 2019, 37(5): 7-12.)
- [3] 汪雪锋,张硕,刘玉琴,等. 中国科技评价研究40年:历史演进及主题演化[J]. 科学学与科学技术管理,2018,39(12):67-80. (Wang Xuefeng, Zhang Shuo, Liu Yuqin, et al. Forty years of research on science and technology evaluation in China: historical and theme evolution [J]. Science of Science and Management of S. & T., 2018, 39(12): 67-80.)
- [4] 王海燕,张昕妍. 我国科技评价体系改革的困境与对策[J]. 中国软科学,2018(4):10-17. (Wang Haiyan, Zhang Xinyan. The predicament and countermeasure of the reform of science and technology evaluation system in China [J]. China Soft Science, 2018(4): 10-17.)
- [5] 代涛,李晓轩. 我国科技评价的问题分析与改革思路[J]. 中国科学院院刊,2013,28(6):750-755. (Dai Tao, Li Xiaoxuan. Analysis on the problems and perspectives of China's S&T evaluation [J]. Bulletin of Chinese Academy of Sciences, 2013, 28(6): 750-755.)
- [6] 杜鹏,李凤. 是自上而下的管理还是科学共同体的自治——对我国科技评价问题的重新审视[J]. 科学学研究,2016,34(5):641-646,667. (Du Peng, Li Feng. Top-down management or autonomy of scientific community: rethinking on China's S&T evaluation [J]. Studies in Science of Science, 2016, 34(5): 641-646, 667.)
- [7] 朱剑. 科研体制与学术评价之关系——从“学术乱象”根源问题说起[J]. 清华大学学报(哲学社会科学版),2015,30(1):5-15,180. (Zhu Jian. Relations between research system and review system: causes to the lack of regulation in academic research [J]. Journal of Tsinghua University (Philosophy and Social Sciences), 2015, 30(1): 5-15, 180.)
- [8] 李楠. 关于深化高校科技评价改革的思考[J]. 中国高等教育,2014(9):18-20. (Li Nan. Thoughts on deepening the reform of science and technology evaluation in universities [J]. China Higher Education, 2014(9): 18-20.)
- [9] 邱均平,朱少强. 宏观与微观学术评价之关系探讨[J]. 图书馆论坛,2006(6):51-53. (Qiu Junping, Zhu Shaoqiang. On the relationship between macro and micro level evaluation of scientific research [J]. Library Tribune, 2006(6): 51-53.)
- [10] 陈晓清,詹启伟,姜田. 关于深化高校科技评价改革的思考与建议[J]. 中国高校科技,2014(1):23-25. (Chen Xiaqing, Zhan Qiwei, Jiang Tian. Thoughts and suggestions on deepening the reform of science and technology evaluation in universities [J]. Chinese University Science & Technology, 2014(1): 23-25.)
- [11] 刘在洲,张云婷. 高校科研质量评价中若干关系探讨[J]. 科技进步与对策,2018,35(17):146-150. (Liu Zaizhou, Zhang Yunting. The quality evaluation of the university's scientific research of several relations [J]. Sci-

- ence & Technology Progress and Policy, 2018, 35(17): 146–150.)
- [12] 汪虹, 梅伏生, 吴绮慧, 等. 高等学校科研浮躁情绪产生的原因及对策[J]. 科技进步与对策, 2002(9): 175–176. (Wang Hong, Mei Fusheng, Wu Qihui, et al. Reasons and countermeasures of blundering emotion caused from scientific research in universities[J]. Science & Technology Progress and Policy, 2002(9): 175–176.)
- [13] 刘恩允. 高校科研评价的问题与对策[J]. 高等工程教育研究, 2004(1): 39–42. (Liu Enyun. Problems and countermeasures of scientific research evaluation in universities[J]. Research in Higher Education of Engineering, 2004(1): 39–42.)
- [14] Geuna A, Martin B R. University research evaluation and funding; an international comparison[J]. Minerva: A Review of Science, Learning & Policy, 2003, 41(4): 277–304.
- [15] 刘明. 学术评价体制检讨[J]. 浙江社会科学, 2005(5): 210–214. (Liu Ming. Review of academic evaluation system[J]. Zhejiang Social Sciences, 2005(5): 210–214.)
- [16] 宋丽萍. REF与科研评价趋向[J]. 图书情报工作, 2011, 55(22): 60–63, 100. (Song Liping. Interpretation of research evaluation trends based on REF[J]. Library and Information Service, 2011, 55(22): 60–63, 100.)
- [17] 马恩斯, 周静, 王峻岭. 比较研究下高校科研评价指标的优化[J]. 中国高校科技, 2019(10): 47–50. (Ma Ensi, Zhou Jing, Wang Junling. The optimization of scientific research evaluation indexes in universities under comparative research[J]. Chinese University Science & Technology, 2019(10): 47–50.)
- [18] Hicks D, Wouters P, Waltman L, et al. The Leiden Manifesto for research metrics[J]. Nature, 2015, 520, 429–431.
- [19] 夏国萍, 管恩浩. 建设世界一流大学亟待处理好四重关系——基于三大世界大学排行榜的视角[J]. 高校教育管理, 2018, 12(4): 27–35. (Xia Guoping, Guan Enhao. The construction of a world-class university in China must deal with four relations; based on world university rankings of ARWU, QS and THE[J]. Journal of Higher Education Management, 2018, 12(4): 27–35.)
- [20] 自然科研签署《旧金山宣言》, 倡导科研评估不要再过度依赖基于期刊的指标[EB/OL]. [2020-06-24]. http://www.sohu.com/a/136891901_465226. (Natural Science signed the *San Francisco Declaration*, advocating that scientific research evaluation should no longer rely on journal-based indicators too much [EB/OL]. [2020-06-24]. http://www.sohu.com/a/136891901_465226.)
- [21] Moed H F. Appropriate use of metrics in research assessment of autonomous academic institutions[J]. Scholarly Assessment Reports, 2020, 2(1): 1.
- [22] 韩毅. 非正式交流回归语境下科技评价的融合路径取向[J]. 中国图书馆学报, 2016, 42(4): 64–74. (Han Yi. The aggregation path orientation of scientific and technological evaluation under the revival context of informal communication[J]. Journal of Library Science in China, 2016, 42(4): 64–74.)
- [23] 俞立平. 规范使用SCI下科技评价的问题与解决路径研究[J]. 情报资料工作, 2020, 41(2): 64–69. (Yu Liping. A probe into the problems and solutions of scientific and technological evaluation under standard use of SCI[J]. Information and Documentation Services, 2020, 41(2): 64–69.)

- [24] 朱少强,唐林,柯青. 学术评价的元评价机制[J]. 重庆大学学报(社会科学版),2010(3):81-85. (Zhu Shaoqiang,Tang Lin,Ke Qing. Meta-evaluation mechanism of academic evaluation[J]. Journal of Chongqing University(Social Science Edition),2010(3):81-85.)
- [25] 马宁锋,李卫东,薛艳泽. 基础教育学业成绩元评价的探析[J]. 中国电力教育,2009(8):29-30. (Ma Ningfeng,Li Weidong,Bi Yanze. Meta-assessment of basic education academic achievement [J]. China Electric Power Education,2009(8):29-30.)
- [26] 王敏. 对大学课程评价之“元评价”的思考[J]. 云梦学刊,2008(5):120-121. (Wang Min. The thought of “meta-evaluation” of the evaluation of university curriculum[J]. Journal of Yunmeng,2008(5):120-121.)
- [27] 俞立平,潘云涛,武夷山. 元评价对提高科技评价质量的影响研究[J]. 科学学研究,2012(4):512-516. (Yu Liping,Pan Yuntao,Wu Yishan. A effective method to enhance science and technology evaluation quality: meta-evaluation[J]. Studies in Science of Science,2012(4):512-516.)
- [28] 蔡国梁,黄斌,李玉秀,等. 科技进步状况的多指标可拓综合评价[J]. 数学的实践与认识,2005(1):59-65. (Cai Guoliang,Huang Bin,Li Yuxiu,et al. Extension synthesize evaluation of science-technology condition [J]. Mathematics in Practice and Theory,2005(1):59-65.)
- [29] 叶继元. 人文社会科学评价体系探讨[J]. 南京大学学报(哲学·人文科学·社会科学版),2010,47(1):97-110,160. (Ye Jiyuan. Approaching evaluation system in humanities and social sciences[J]. Journal of Nanjing University(Philosophy,Humanities and Social Sciences),2010,47(1):97-110,160.)
- [30] 全允桓. 科技评价理论与方法的体系结构[J]. 科技成果纵横,2003(5):18-20. (Tong Yunhuan. System structure of science and technology evaluation theory and method[J]. Perspectives of Scientific and Technological Achievements,2003(5):18-20.)
- [31] 邱均平,文庭孝,等. 评价学:理论·方法·实践[M]. 北京:科学出版社,2010. (Qiu Junping,Wen Tingxiao, et al. Evaluation science:theory method practice[M]. Beijing:Science Press,2010.)

邱均平 杭州电子科技大学中国科教评价研究院资深教授,院长。浙江 杭州 310018。

张裕晨 杭州电子科技大学中国科教评价研究院讲师,博士研究生。浙江 杭州 310018。

周子番 杭州电子科技大学管理学院博士研究生。浙江 杭州 310018。

(收稿日期:2020-08-07;修回日期:2020-10-15)