

科研成果评价的常用方法与分析

徐文焕 罗长坤 黄国琼

【摘要】 系统查阅、收集、归纳、总结现有的基础科研成果评价方法,并加以归纳总结,主要包括定性评价即同行评议,定量评价即文献计量法,半定量评价即回溯法和案例研究法、定性定量相结合的评价方法。通过分析各种评价方法的特点和差异,以期为从事基础科研成果评价的部门及研究者提供有益的参考。

【关键词】 基础科研成果; 评价方法; 比较分析

Commonly used methods for evaluating scientific achievements and an analysis of them XU Wen-huan, LUO Chang-kun, HUANG Guo-qiong. *The Third Military Medical University, Chongqing 400038, China*

【Abstract】 Current methods for evaluating basic scientific achievements were retrieved, collected, and summarized, including qualitative evaluation, viz. peer review, quantitative evaluation, viz. bibliometrics, semi-qualitative evaluation, viz. backtracking methods and case studies, and the integration of qualitative and quantitative evaluation. A comparative analysis was made of the features of and differences between the various evaluation methods so as to give researchers and departments engaged in evaluating basic scientific achievements some beneficial insights.

【Key Words】 Basic scientific achievement; Evaluation method; Comparative analysis

20 世纪 90 年代以来,随着世界各国对基础研究投入逐年增加,基础科研成果评价问题已成为各国政府、科研管理部门及科技政策研究人员关注的热点。基础科研成果能否得到科学合理、客观公正的评价,对国家资源的优化配置、调动研究人员的积极性、不断提高基础研究成果的创新水平和影响等方面都有重要意义。其评价方法的形成是一个不断发展和完善的过程。目前,常用方法大致可分为 4 类,即定性评价、定量评价、半定量评价及定性定量相结合的评价方法。我们通过分析上述评价方法的特点和差异,以期为从事基础科研成果评价的部门及研究者提供有益的参考。

一、定性评价

广义来说,定性评价是对评价对象做出非量化的结论。定性评价的主要实践形式即同行评议(peer review)。同行评议是同行咨询(peer advice)、同行评估(peer evaluation)、同行审查(peer censorship)、专家鉴定(refereeing)等同义语的通用

术语^[1]。同行评议作为科研工作的评价方法已有数百年历史,它的广泛应用是从 20 世纪 50 年代开始的。美国国会技术评估办公室对同行评议下的定义是:“同行评议是用于评价科学工作的一种组织方法。这种方法常常被科学界用来判断工作程序的正确性,确认结果的可靠性以及对有限资源的分配,即诸如杂志版面、研究资助经费及公认性和特殊荣誉”。英国同行评议调查组对同行评议的定义是“有从事该领域或接近该领域的专家来评定一项研究工作的学术水平或重要性的一种方法”^[2]。

目前,国内外科研管理部门在评价基础科研成果时,普遍应用了同行评议的方法。同行评议一般包括通信同行评议、专家组或委员会同行评议、现场同行评议和内部同行评议几种^[3]。每位同行评议专家的意见通常通过专家填写的评审表表达。评审表基本由两部分组成,即专家根据评价指标的内涵按评定程度赋分,然后标明授奖等级。

同行评议的优势是,专家代替了科学外行,拥有了对学术问题的决策能力,操作得当的同行评议能够对科学研究发挥巨大的激励和支持作用。但由于

作者单位:400038 重庆,第三军医大学

专家评议制度尚不够完善,缺乏对评价专家的有效监督制约机制,最容易出现的问题是“马太效应”,可能因学派或利益、熟人等关系而削弱评议结果的权威性和客观公正性。据社会科学学家的研究报告,科学界的人才分布是呈宝塔形的,一个传统学科形成之后,在这个学科领域通常是由几个最具有权威的科学家控制的,当这种权威结合在一起时,便形成一种叫做“科学贵族”的现象。如果这些“科学贵族”控制了同行评议,对基础科研成果评价的结果影响会更大。这些人士如果是有见地的科学家,会对基础科学的发展起到极大的推动作用;如果这些人士固执己见,排斥异己,则对基础科学的发展将起到阻碍作用。

为提高同行评议质量,一方面要选择合适的专家参加评议。评价基础科研成果的专家应当从研究经验丰富、学术眼光敏锐、战略意识强和知识面宽的专家中遴选产生,并注重吸纳在一线从事高水平基础研究的专家参加评议。另一方面要根据专家的权威性和可信度对不同专家的评议赋予适当的权重系数,建立高质量的同行专家库。

二、定量评价

定量评价是依据某种模型对科学研究活动与结果的相关数据进行计算,得出定量的结论。如文献计量学、经济计量学,如成本效益分析等。其中文献计量学方法尤为重要,特别是在评价基础科研成果时,可以作为定性同行评议的补充。

文献计量法用于分析科学研究活动,始于 19 世纪末和 20 世纪初,直至 20 世纪 30~40 年代,有关文献特征的几个规律发现,为文献计量学的发展奠定了坚实的基础。20 世纪 60 年代末,由美国学者德里克·普赖斯(Derck J de Solla Price)首次提出了文献计量学(Bibliometrics)这一术语。《科学引文索引(Science Citation Index,缩写为 SCI)》于 1955 年由美国学者 K·加菲尔德提出,它的编辑出版为利用文献计量手段分析科学研究成果特征提供了一个有利的工具。文献计量学利用的文献特征包括:科技论文(专著)的数量、科技论文(专著)的引证、引用率、平均被引次数等。科技论文数量可以反映一个国家或某个研究队伍的规模和科研能力,但反映不出论文的质量;引文指标可以反映论文的影响、时效性和效用等。从某种程度上说,引文指标反映的是论文质量。但是,文献引用的动机比较复杂,按照

温斯托(M. Weistock)的研究,可以将文献引用的动机分为 15 种。一般来说可以归纳成 3 种类型:一是指以前研究成果的创新;二是对研究工作者的感谢;三是批评和否定过去的研究工作。研究表明,批评与否定的引文占的比例较小,约 5%,大多数引用都是正面引证^[4]。

中国科技信息研究所从 1987 年起对中国科技论文所作的年度分析以 SCI 为主要统计源,并以此统计结果,即论文收录数和引用数来排序,以衡量各科研单位或学校在基础研究方面的总体水平和实力^[5]。现在对基础研究成果的评价仍应采用这种统计标准,主要是通过出版的论文、引用次数统计对基础科研成果进行评价。

文献计量法的优点是,具有较强的科学性,不受个人主观因素和其他非科学因素的影响。文献计量法在用于基础科研成果评价过程中也暴露了一些缺陷。比如数据来源问题。现在对基础科研成果评价常用的 SCI 标准,其扩展版 SCIE 只收录了全球科技期刊中的 6000 余种,该数据库只有相对的统计学意义;其次,还存在着严重的语言偏见问题。如德国和法国期刊发表了大量使用本国语言撰写的论文,因无法被 SCI 收录,导致 SCI 得出德语论文在减少、德国影响力在下降的结论,而事实上德文文献却上升了 15%。同样,语言偏见也出现在对中文期刊上。所以,建议对从事基础科研成果评价的有关部门和专家应正确使用 SCI 标准,既要重视它的优势所在,又要克服它存在的问题,充分结合定性评价的方法,正确评价基础科研成果的价值。

三、半定量评价

半定量评价法是指介于定量评价和定性评价之间的第 3 类评价基础科研成果的方法,包括案例研究法和回溯法。

1. 案例研究法:它是最先用于分析某项具体技术的发展轨迹。例如, Jewkes、Sayre 和 Stiller 利用这种方法追溯了创新的演进历程,即从某项创新的出现追溯到对创新起支持作用的基础科学发现,考察了 R&D 和经济的关系。20 世纪 60 年代至 70 年代,美国政府机构为理解科学研究和经济增长的关系,开展了大量案例研究。例如,美国国防部的“研究项目的事后认识”(project hindsight),国家科学基金会(NSF)的“技术回溯和科学项目中的重大事件”(TRACES)。这些研究表明,案例研究法可以

提供历史进程方面的知识,为基础科研成果的评价提供一种思路。

2. 回溯法:回溯法研究与案例研究方法是相关的,不同的是回溯研究通常与多方面科学或技术创新有关而非单一方面。早期这类研究的动机是根据研究开发的有效应用,证实对其投资的正确性。较近期的研究,如 Comrore 和 Dripps 为 NIH 做的研究,涉及了基础研究成果对科学或临床进展的贡献^[6]。

案例研究与回溯研究方法无疑给基础科研成果评价开辟了一种新的思路,案例研究与回溯研究通常与多方面的科学或技术创新有关,其目标是要认证创新和特定先前事件,比如资金资助等之间的关联,其他评价方法是做不到这一点的。但遗憾的是案例研究与回溯研究相对其他方法费时又费钱,如美国学者 Judson 就分子生物学进展所做的研究,历时 10 年,足迹踏遍欧美。Wade 仅局限于下丘体的发现,所做调查研究也花去许多年,这也是妨碍案例研究与回溯研究评价方法广泛应用的原因。

四、定性与定量相结合的评价

由于定性评价中存在太多的主观性,使得评价结果偏离公正性。定量评价中存在一些不合理的问题,使得评价结果带有一定的不科学性,如现在对基础科研成果评价特别强调 SCI 数量,已经对科学评价带来了负面影响。所以,近年来国内外不少学者为确保基础科研成果评价的科学性和准确性,对基础科研成果评价方法进行了更深层次的探索研究,提出定性评价与定量评价相结合、内部评价与外部评价相结合,运用多种方法对基础科研成果进行综合评价。从具体评价方法来看,我国目前常用的是同行评议、文献计量法、聚类分析法等。国外在具体评价过程中运用的方法则较多,如同行评议、文献计量法、案例法、回溯法、经济计量法等。相比较而言,国外对基础科研成果的评价更注重实证研究。

最近,美国国家标准与技术研究所的战略规划与经济小组对基础研究成果产生的经济效益运用了经济回报率的测度进行了测算。一些学者试图运用更合适的方法和手段来研究与测度公共基础研究与

特定预期效果之间的定量关系。比如运用求偏指标弹性(Partial—indicator elasticity)的方法对基础研究的投入回报进行研究^[7],从理论方法角度看,这是一个创新。

总之,几十年来,国内外许多学者从不同角度、运用不同方法就基础科研成果评估的理论与方法进行了有益探索和尝试,也取得了较丰富的实证经验,对基础性科研成果的评估逐渐形成了比较成熟的理论与方法。在评价中运用多种评价方法相结合的综合评价,定性评价中参照一定的定量指标,既可以避免定性评价的主观性和不确定性,又可避免定量评价疏于质量指标的缺陷。

为提高基础科研成果评价的质量,增强国家的科技竞争实力,建议有关科研主管部门继续加大对基础科研成果评价理论与方法的研究,突出评价的导向性,从系统整体的观点出发,科学合理地对基础科研成果进行评价,不仅能对基础科研成果的学术价值给予客观公正的评价,同时还会对基础研究方向选择、资源配置提供经验、参考和借鉴,使有限的资金能更准确地投资到国家所需要的、可行的重点研究领域,使基础科研成果评价工作真正起到竞争、激励、约束的作用,促进国家科技事业的发展,促进我国原始性创新成果的产出。

参 考 文 献

- 1 黄亚明,何钦成. 科技成果评估中的常用方法. 中华医学科研管理杂志, 2004, 17: 12.
- 2 郭碧坚,韩宇. 同行评议制-方法、理论、功能、指标. 科学学研究, 1994, (3): 63-66.
- 3 吴述尧,主编. 同行评议方法论. 北京:科学出版社, 1996. 3-7.
- 4 国家自然科学基金绩效评估研究课题组. 国家自然科学基金绩效评估指南研究(内部资料), 1998. 5.
- 5 沈新尹. 引文计量与基础研究成果评价. 科学学与科学技术管理, 1996, 17(1): 6.
- 6 阮祖启,译. 吴述光,韩宇,审校. 科学质量的评估. 美国国家科学基金会, 1994. 6-22.
- 7 葛朝阳,郑刚,陈劲. 基础研究的经济回报率测度与评价:国外研究述评. 科研管理, 2003, 4(1): 46.

(收稿日期:2006-04-18)