



# 科研评价改革与发展40年

——以基金委同行评议和中科院研究所综合评价为例

徐芳<sup>1,2</sup> 龚旭<sup>3</sup> 李晓轩<sup>1,2</sup>

(1. 中国科学院 科技战略咨询研究院, 北京 100080; 2. 中国科学院大学 公共政策与管理学院, 北京 100080; 3. 国家自然科学基金委员会 政策局, 北京 100085)

**摘要:**“科学的春天”至今已40年,科研评价也经历了改革与发展的40年。在回顾我国科研评价发展的三个阶段基础上,分别阐述了基金委项目评审的同行评议和中科院研究所综合评价这两个具有代表性和时代性特征的科研评价实践,进而总结我国科研评价改革与发展40年来取得的成绩,以及在新时代面临的挑战。

**关键词:**科研评价;改革开放;科学基金;同行评议;中科院研究所综合评价

**中图分类号:**G301 **文献标识码:**A **文章编号:**1002-0241(2018)12-0017-11

## 0 引言

1978年3月18日,被誉为给中国带来了“科学的春天”的全国科学大会在北京召开,这次大会澄清了长期束缚科学技术发展的重大问题,极大提振了广大知识分子的创新热情。大会还通过了《1978—1985年全国科学技术发展规划》,为中国科学事业的全面复兴奠定了良好开端。“科学的春天”40年来,作为科技管理重要工具的科研评价,它的改革与发展是中国科技事业发展和深化科技体制改革的缩影。本文选取科研评价改革与发展40年这一历史切面,旨在通过回顾我国科研评价发展与演进的三个阶段的基础上,以基金委项目评审的同行评议和中科院研究所综合评价这两个具有代表性和时代性特征的科研评价实践开展回溯分析,总结我国科研评价改革与发展40年来取得的成绩,以及在新时代面临的挑战,以期为科技界和管理者提供知识积累与启示。

## 1 科研评价政策与实践的发展与演进

恢复与重建(1978—1984):以人才和成果评价为主。1978年3月18日,党中央召开了具有深远历史意义的全国科学大会。邓小平同志提出“科学技术是生产力”、“知识分子是工人阶级一部分”等重要论断,扫清了“十年动乱”后科学技术发展的政治障碍,吹响了向科学技术现代化进军的号角,带来了“科学的春天”。在大会上,国家对7000多项科技成果进行了奖励,标志着我国中断了十多年的科技奖励制度得到恢复。国家开始倡导“尊重知识、尊重人才”,启动职称制度的恢复与重建。

1978年国务院率先恢复了《关于高等学校教师职务名称及其确定与提升办法的暂行规定》,并指出“原已确定提升的各等级职务一律有效,恢复名称,不须重新办理报批手续”。随后,国务院又正式批准开展教师职称评定工作,不久又在保留原有5个学术、技术职称的同时,增加了业务职

收稿日期:2017-12-25

基金项目:国家自然科学基金面上项目(71771205);中国科学院青年促进创新研究会项目(Y8X0101501);中国科学院战略研究与决策支撑系统建设专项基金项目(GHJ-ZLZX-2018-32-1)

第一作者简介:徐芳(1984—),江苏常州人,博士,中国科学院科技战略咨询研究院副研究员,研究方向:科技政策、科技管理与评价。

通信作者:李晓轩, xiaoxuan@casisd.cn

称。至此,我国创建了专业技术职称评定制。职称制度和科技奖励制度等的恢复,标志着我国科研评价进入恢复期。然而,职称制度恢复后,全国上下很快形成了职称评定热潮,产生了一些负面结果,使职称评定工作很快陷入混乱局面。为此,1983—1985 年三年间,国家暂停了职称评定,进行职称工作整顿。同一时期的科技奖励评价也出现了失真失控的局面。由于当时我国整体科技水平相对较低,以及科研评价中人情因素的影响等原因,评价远没有做到客观公正,更谈不上规范的同行评议。在成果评审时,许多成果不是“国际先进”就是“国内领先”,各种奖励“遍地开花”,科技奖励产生了较大负面影响<sup>[1]</sup>。随后,国家开始研究建立以适应当时经济、科技、教育体制改革需要的职称制度,科技界也在探索如何开展科研评价。

探索与实践(1985—2003):以机构实践为主。1985 年,中共中央发布《关于科学技术体制改革的决定》,改革的指导思想是“科学技术要面向经济建设,经济建设要依靠科学技术”,重大政策之一是改革拨款制度。对于主要从事基础研究的机构,在保证基本事业费的同时,试行科学基金制,通过同行评议、择优支持发放基金。《决定》同时指出设立国家自然科学基金会,标志着我国科研经费从“计划分配”向“竞争择优”过渡。科学基金制的实施和国家自然科学基金委员会(以下简称“基金委”)的成立,使得科研项目从任务分配变为自主申请,选题上充分尊重科研人员的学术意愿;项目遴选采用同行评议,尊重学术共同体意见,减少行政干预;项目实施过程中,科研人员有权自行组织研究,发挥科研人员的自主性。自此,科学共同体自主开展科研评价的模式逐渐开始盛行。

作为重要的科技管理工具,科研评价的运用也越来越广泛。1987 年起,原国家科委试行将科学计量学用于科学研究成果与水平的评价,并委托中国科技信息研究所根据三种有国际影响的大型检索

系统《科学引文索引》(SCI)、《科学评论索引》(ISR)和《科技会议录索引》(ISTP),对 1983—1986 年间中国科技论文被收录、索引的情况做统计分析,排出中国各高等学校和科研机构的“学术榜”<sup>[2]</sup>。自此,中国科学界开始采用国际通用的科学计量指标评估高等院校和科研院所的研究实绩。南大物理系主任最早将国际学术论文(SCI 论文)的概念引入科研评价和职称评审。在当时的背景下,在科研评价中引入 SCI 等国际评价指标,不仅显示了机构自身的学术水平和办学实力,而且确立了一种勇于争先、敢于到国际学术界去竞争、比试的气势,对中国科研提升质量和国际影响力都起到了助推作用<sup>[3]</sup>。正是由于南京大学的“一马当先”,才带动了大学和科研院所的“万马奔腾”,促使中国科学研究出现了全新的格局,科研评价也进入蓬勃发展的兴盛期,机构层面组织的各类评价工作也呈现百花盛开的局面。如,中科院在 1993 年启动研究所评价,基金委组织开展同行评议方法论研究并改进评议系统,1999 年国务院设立五大国家科学技术奖,科技部启动国家重点实验室评估,2002 年教育部启动首轮学科评估等。

制度建设与规范(2003—至今):政府介入管理。进入 21 世纪,“科研评价热”愈演愈烈,SCI 论文指标在科研评价用途上不断拓展,科技界对 SCI 论文量化评价(也包括项目、经费、专利、人才、奖励等的量化评价)越来越依赖,以 SCI 论文量化评价导向为主的问题也越来越凸显,政府开始介入到科研评价的管理中。2003 年,科技部、教育部、中科院、工程院、基金委联合发布《关于改进科学技术评价工作的决定》,提出实行分类评价,“倡导质量第一”,建立“公平、公正、公开”的评价原则等指导性意见。随后,科技部印发了我国关于科研评价管理的第一份系统性的文件——《科学技术评价办法》(试行),对科技计划及项目、研发机构、研发人员、科技成果等评价活动分别提出规范性要

求。此后,我国科研评价进入制度建设与规范期,按照政府要求开展的系列科研评价活动逐渐增加,如2010年中科院根据国家要求完成了知识创新工程的自评<sup>[4]</sup>。

十八大以来,科技创新发展迎来新的春天,科研评价改革成为政府关注的重点,并且重视在评价中发挥引导作用。在2016年召开的全国科技创新大会上,习近平总书记强调要改革科技评价制度,建立以科技创新质量、贡献、绩效为导向的分类评价体系;李克强总理要求改进对科研活动的评价机制。《国家创新驱动发展战略纲要》、《关于深化人才发展体制机制改革的意见》、《深化科技体制改革实施方案》等一系列文件对深入推进科研评价改革做出了系统部署。近期,中办和国办联合发布《关于深化项目评审、人才评价、机构评估改革的意见》,要求深化“三评”改革,构建科学、规范、高效、诚信的科技评价体系,推进分类评价制度建设,发挥好评价指挥棒和风向标作用。作为“三评”改革的具体措施,科技部五部委联合下发《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》,这些文件直面当前我国科研评价面临的突出问题,标志着我国科研评价进入深化改革期。

## 2 作为科研评价基石的同行评议制度建立与发展:以NSFC为例

同行评议是科研活动中最常用的评价手段,更是目的各异、对象多样的科研评价的基石。但同行评议并不意味着将专家组织起来开展科研评价,就能够自动地保障学术质量。同行评议若要行之有效,必须建立一套有关评价程序和评价主体的规则与规范,即形成同行评议系统;该系统只有在一定的运行条件与制度环境下,才能确保评议的公正性和有效性。在过去40年我国科研评价改革与发展历程中,基金委不仅因其规范高效的同行评议制度在国内外科学界赢得了广泛声誉<sup>[5]</sup>,

而且对我国同行评议系统的建立与发展起到引导和示范作用,通过建立与完善保障同行评议系统良好运行的制度环境及条件,以及组织开展同行评议研究与实践探索,加深了科学界对同行评议作为基础研究评价及管理机制重要性的认识和理解。

### 2.1 同行评议系统的建立与完善

“二战”之后,同行评议作为科研资助经费的核心分配机制在发达国家陆续得到普遍应用<sup>[6]</sup>,形成了由评议主体及其行为规范、评议程序、评议准则等所构成的相对成熟的同行评议系统,其中各组成部分对应着解决“谁来评”、“如何评”和“评什么”等关键问题,从而实现控制科研质量的目标。尽管各资助机构因职责和传统的差异在评议程序和评议准则上有所不同,但对于评议主体(即评审专家)在行为上的规范却基本相同。在改革开放中成立的基金委,为了改革计划体制下科研资源分配的弊病,在构建同行评议系统的过程中,充分借鉴了国外的做法和经验,特别是美国科学基金会(NSF)的评议程序和评议规范。

基金委成立于1986年,其前身是1982年设立的中科院科学基金局。科学基金制在中科院试行之初,就以“专家评审,择优支持”为特征<sup>[7]</sup>,即明确了“谁来评”和“评什么”的问题。但由于同行评议处于先行先试阶段,关于“如何评”的工作规范较为简单。例如,《中国科学院科学基金试行条例》仅对专家评审工作做出初步规范,规定同行评议在评审程序中的地位、参加评议的专家资格与人数,并对专家评议提出保密等行为要求<sup>[8]</sup>。直到基金委成立之后,我国才真正建立起相对规范和完备的同行评议系统。

基金委发布的第一个政策文件即为《关于申请项目评审工作暂行办法》(1986年5月),对评审专家的遴选以及专家回避和保密等行为提出基本要求<sup>[9]</sup>。专家库是同行评议的重要基础设施,随着我



国基础研究队伍的不断壮大,基金委已逐步建立起全国学科面最广、数量最多、信息化水平最高的专家库,其人数从1986年的1万余人起步,目前已发展至近20万人,包括海外专家数千人。30多年来,基金委的专家遴选机制不断制度化,对专家评议行为的规范性不断强化,在最初关于专家回避与保密的基本要求之上,陆续进一步细化评审专家回避与保密的情形、规范评审专家行为、建立评议监督和问责机制等。

基金委成立之初就确定了由通讯评议和会议评审构成的两级项目评审程序<sup>[10]</sup>,两级评审的步骤以及通讯评议和会议评审的相关程序要求,不仅体现在基金委的管理办法中,而且在2007年国务院颁布施行的《国家自然科学基金条例》中得以确立<sup>[11]</sup>。自1986年起,基金委就针对陆续设立的各项项目类型的不同功能进行定位,对评审工作提出相应的要求,以引导专家评议和项目遴选。例如,对于青年科学基金项目,要求专家关注申请人的创新潜力,而非像评审面上项目那样要求考虑研究队伍构成<sup>[11]</sup>;对于面向国家经济社会和科技发展重大需求的重大项目,则要求专家考察研究内容与总体目标之间的相关性,还要评价研究队伍能否满足多学科交叉和综合性研究的开展<sup>[11]</sup>。这些类似的政策要求有效地引导了专家的评议工作,对专家开展同行评议也可视为一种有益的培训。

## 2.2 同行评议制度的建立与发展

同行评议制度涵盖了同行评议系统和保障同行评议运行的制度环境及配套条件,后者包括科学自主的理念、专业管理的队伍、监督制约的机制等。尽管各国科学基金组织的资助范围和政策导向不尽相同,但其评议工作的基本理念、管理人员的专业要求和评议的监督制约机制大体相似,这在很大程度上是为了保障同行评议公正有效地运行的缘故。我国同行评议制度的建立与发展伴随着科学基金制的演进历程,而且与科研管理学术

化和专业化的制度化进程密不可分。

基金委成立伊始,其领导层就认为基金委既是“学术机构”又是“管理机构”,而且“首先是一个学术机构”<sup>[12]</sup>;1986年提出了坚持至今的资助工作“十六字原则”,即“依靠专家,发扬民主,择优支持,公正合理”<sup>[13]</sup>。基金委突出学术和依靠专家的特点,不仅意味着改革开放带来的科研评价方式的变化,更标志着科研管理理念的转变。正如期刊论文录用的同行评议属于学术活动,资助机构项目遴选的同行评议同样是科研的重要组成部分<sup>[14]</sup>。基金委作为新型科研管理机构,之所以不同于传统的行政管理机构而将学术工作置于首位,也与同行评议在科学基金工作中的重要地位密不可分。

正因为同行评议的组织管理具有学术和管理的双重属性,科学资助机构的管理工作必须依靠大量业务水平高、学科知识丰富的专业人员<sup>[15]</sup>。基金委不仅在选聘资助管理人员中对专业经历和业务水平要求较高,而且通过各种途径不断保持和提升其专业能力,包括组织管理人员参与相关学科领域发展战略研究,使其掌握学术动态及发展趋势,熟悉专家队伍,“锻炼……组织和协调能力”<sup>[16]</sup>。

1988—1997年,基金委以其资助的56个学科为单元,组织开展大规模学科发展战略研究,负责各学科领域的管理人员深度参与其中。21世纪以来,学科发展战略研究成为基金委的常规工作,自2009年起还与中科院和工程院联合组织开展研究,提升了基金委管理人员在学术领域和同行评议组织管理方面的整体能力<sup>[17]</sup>。

为了保障同行评议公正性,上世纪90年代后期以来,基金委还陆续建立了向申请人全文反馈同行评议意见、监督委员会受理针对包括评议在内的资助工作各环节投诉举报、向评审会议派驻监督委员会等同行评议监督机制。这些监督机制既提升了同行评议的公正性,也为我国科研诚信建设发挥了示范作用<sup>[18]</sup>。

### 2.3 同行评议的研究与改进

如前所述,基金委在发展过程中,充分借鉴国外科学资助机构的管理制度及成功经验,其中包括国外同行评议制度与最佳实践,尤其是了解、研究与借鉴美国NSF的同行评议。早在中科院科学基金局时期,就组织翻译美国科学社会学家针对NSF同行评议的专题研究报告,为我国科学界了解国外同行评议工作、理解同行评议组织管理、建立我国同行评议体系等,提供了重要的参考借鉴。1986年基金委成立后,不仅继续、拓展与深化对国外同行评议的译介与研究,90年代后期起还与美国、德国、英国等国外资助机构联合举办同行评议双边研讨会,交流同行评议经验,考察同行评议制度,改进同行评议运行,促进我国同行评议研究水平提升<sup>[19]</sup>。

根据对1979—2017年知网(CNKI)期刊收录的以“同行评议”为主题的文献(共1781篇,其中研究论文占90%)进行统计分析发现,我国学术界的同行评议研究始于改革开放初期,上世纪90年代有所增长,21世纪以来发展显著(见图1)。这一趋势恰好与我国科学基金制的发展相一致,同基金委对同行评议研究的支持持续加大相符合。

在同行评议主题的研究论文中,1994年起出现受资助论文(且为国家自然科学基金资助),2002年以来受各项政府科研计划(包括国家自然

科学基金)资助的同行评议研究论文增长较快(见图2),主要刊发在科技政策研究的学术期刊上,基金委机关刊物《中国科学基金》刊发的同行评议研究论文亦不在少数。在同行评议研究论文作者中,基金委管理人员一直是一支独特的作者队伍,发文最多的前五位作者中有三位为科学基金管理人员<sup>[20]</sup>。

基金委对同行评议的研究既有组织科技政策专家共同开展的,也有管理人员结合个人资助工作进行。1991年委托中科院院士牵头组织开展的“同行评议方法论”研究,是我国第一次对同行评议制度进行的全面系统分析及考察,而且紧密结合科学基金资助工作,针对专家库建设、“非共识项目”、学科交叉研究等同行评议难点问题深入探讨,提出了具有前瞻性和建设性的政策建议<sup>[21]</sup>。基金委组织同行评议的管理人员利用工作便利,研究尝试新方法新机制,如计算机辅助指派专家<sup>[22]</sup>、网络预投票等<sup>[23]</sup>,研究成果丰富和改进了科学基金同行评议实践,促进我国同行评议制度持续完善,以及科研评价不断创新。

### 3 科研综合评价的发展与演进:以中科院研究所评价为例

中国科学院于1949年11月在北京成立,是国家自然科学最高学术机构、科学技术方面的最高咨询机构、自然科学与高技术综合研究发展中

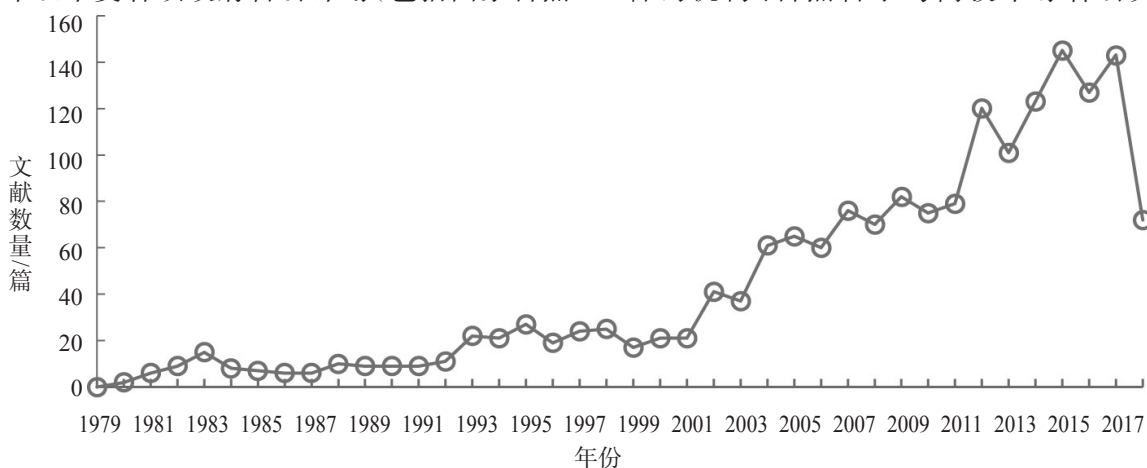


图1 1979—2017年同行评议主题文献数量变化

心。在科研评价工作方面,中科院走在全国前列,对我国科研评价工作发挥了引领和示范作用。1981—1984年,中国科学院学部对院所属的42个研究所进行了现场评估,开启了我国研究机构学术评价的先河。1982年,中科院设立主要面向全国的中国科学院科学基金,探索了以同行评议为基础的项目评审方法,为国家自然科学基金项目评审积累了实践经验<sup>①</sup>。1985年,中科院启动开放实验室建设,其后开展了对其的定期评估,对随后我国国家重点实验室评估起了示范作用。2004年,中科院主动要求国家对中国科学院一、二期知识创新工程试点工作进行评估<sup>[24]</sup>,国家科教领导小组委托科技部牵头组织实施,这是我国对大型科技计划和工程进行系统国家评估的范例。

伴随着我国科技事业的发展和科技体制改革的深入,中科院系统内部先后开展了研究所评价、项目评估、优秀人才评估、重点实验室评估等,其中,尤以自1993年启动的研究所综合评价涉及面最广,影响最大,其评价体系和方法为我国一些国立研究机构所借鉴。回顾中科院研究所评价历史,自上世纪90年代以来,结合不同时期的发展战略,中科院研究所综合评估经历了“蓝皮书评价体

系—二元评价体系—综合质量评估体系—重大产出导向评价体系”四阶段<sup>[25]</sup>,其评估方法经过了从以数量评价为主到以质量评价为主,从研究所大排名到分档分类,从单纯的外评估到自评估与交流评议相结合,从每年一评到五年一评年度监测与中期微调的发展与演进。

### 3.1 蓝皮书评价体系(1993—1998)

20世纪80年代中期,科研评价在我国迅速兴起。中国学术界开始对科研机构进行评价与排名,有些体系将中科院纳入被评对象之一。然而,由于外界缺乏对中科院的理解,外部评价不能恰当反映中科院完整准确的情况,反而导致公众对中科院产生误解。在此形势下,中科院于1993年启动了自己的研究所综合评价<sup>[26]</sup>。作为院对所宏观管理的重要工具,综合评价有利于院及时了解和把握研究所的现状和发展态势,回应国家和社会公众的关切。

90年代初,中国科技整体水平相对低下,资源有限,科研成果少,同行水平和质量无法保障,在这种情况下,中科院提出对研究所采用“蓝皮书评价体系”,即采用定量指标评估研究所的状态和绩效,同时采用国际通行的定量指标(如SCI论文数

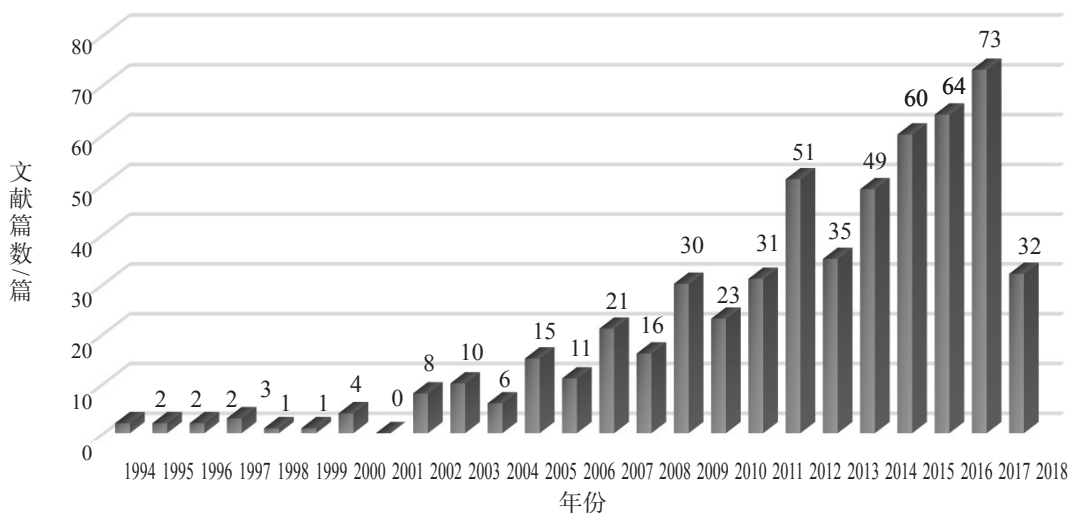


图2 1994年以来受资助同行评议论文发表数量变化

①[http://www.cas.cn/jzzky/ysss/bns/200909/t20090928\\_2529245.shtml](http://www.cas.cn/jzzky/ysss/bns/200909/t20090928_2529245.shtml).



等)评估基础类研究工作,以求在短时间内帮助研究所提高科研水平和成果产出质量。简言之,蓝皮书评价体系主要发挥对研究所产出的监测作用,其评价结果并没有在资源配置等管理决策中发挥作用,重在激励和促进研究所发展发挥积极的正向意义。1994年发布了第一次评价结果,因其发布报告封面为蓝色,得名“蓝皮书评价”。

### 3.2 二元评价体系(1999—2004)

1998年,国家决定在中科院实施知识创新工程。知识创新工程的总体任务是到2010年前将中科院建成具有强大和持续创新能力的国家自然科学和高技术创新中心<sup>[27]</sup>,为国家创新体系建设探索道路。为鼓励和引导研究所根据知识创新工程目标开展基础性、前瞻性和战略性(简称“三性”)的科研活动,中科院改革研究所综合评价体系为二元评价体系(即目标完成情况评价和“三性”贡献评价)<sup>[28]</sup>,其特点是定量评价与定性评价相结合,关注重大创新成果、目标完成度和政策引导情况。评价结果表达为研究所分类排名,并作为院对研究所调整资源配置的重要依据以及研究所所长年薪的依据。

与蓝皮书评价体系相比,二元评价体系本质上是绩效评价,在知识创新工程实施初期,大量资源投入到研究所,院对研究所的管理相应地也加强

了,表现之一是通过二元评价体系引导研究所研究方向,同时通过设定绩效目标以便阶段性地开展绩效评价。在评价方法上,除了定量指标评价外,对科技目标完成情况采用专家评议。定量指标也在蓝皮评价基础上做了较大调整,一是二元评价体系中的指标以“三性”贡献为基础重新构建,二是由强调研究活动的近期直接“产出(outputs)”,扩展为关注影响覆盖学术界和社会经济的中长期“成果(outcomes)”。三是评价标准进一步提升。以“高质量科学论文”为例,SCI论文中只有高端影响因子论文(约占SCI论文总数的10%)入选计分。

### 3.3 综合质量评估体系(2005—2010)

2005年中科院进入知识创新工程第三阶段,该阶段核心是提升研究所的创新能力。为此,中科院改革二元评价体系为综合质量评估体系,其特点是以研究所年度基础数据定量监测和自评交流为基础,系统评价研究所创新绩效、发展状态和规划布局<sup>[29]</sup>。该体系包括研究所自评估、外部专家同行评议、相同领域研究所交流评议、管理专家现场评估以及综合决策分档等多个环节(见图3),各环节侧重点不同。关于该体系中的定量评价工具,中科院在原有定量评价基础上,研究开发了科技创新能力指数工具<sup>[30]</sup>,与以往量化评价中归一化

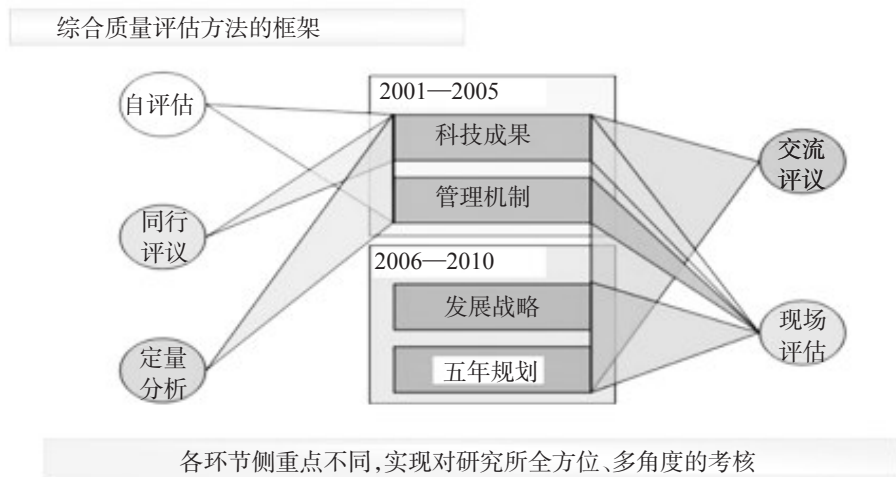


图3 中科院研究所综合质量评估体系框架图

处理方式不同,该指数根据指标权重进行赋分,可以实现研究所的横向可比和纵向可比。关于该体系中的定性评价方法,配合研究所年度自评交流,中科院开发了表征研究所科研活动8个关键方面的自评工具,指导研究所对照检查,并通过ARP平台,促进了研究所交流。科技创新能力指数和研究所自评交流工具为院所两级的态势分析提供了有力支撑。

综合质量评估体系实现了定量评价与定性评价相结合,同行专家评价与管理专家评价相结合,自评估与外部专家评估相结合,现有成绩评价与未来发展评价相结合,基于汇报、数据和材料的评估与现场实地评估相结合,是对研究所全方位、多角度、多信号的反馈。

### 3.4 重大成果产出导向评价体系(2012年至今)

2010年知识创新工程结束,中科院进入新的发展阶段,并公布了其2020年的发展战略(又称“创新2020”),旨在鼓励研究所在科技进步、经济社会发展 and 国家安全方面作出重大贡献<sup>[31]</sup>,同时要求研究所根据“创新2020”制定“一三五”规划,即一个定位、三个重大突破和五个重点培育方向。“一个定位”是指研究所确定的主要研究领域、核心竞争力、以及在国际上的预期地位,并避免与其他研究所同质化。“三大重大突破”是指研究所在今后5~10年内有可能产生的不超过三个的基础性、前瞻性和战略性的六大类重大成果产出。“五个重点培育方向”是指研究所在今后5~10年内不超过五个具有独特特征的优先研究事项、未来的竞争优势或潜在的突破。为推动“一三五”规划的实施,中科院调整综合质量评估体系为重大成果产出导向的研究所综合评价体系,同时构建了“两个环节一个基础”。两个环节包括“一三五”中期评估和“一三五”目标完成情况验收评估,一个基础是研究所关键指标年度监测。

“一三五”中期评估是重大成果产出导向评价

体系中的重要一环,其特点是借助国际专家力量帮助研究所判断能否产出重大成果。在评估程序安排上,借鉴国际上评估的一般做法,包括提前给专家提供材料,现场评估安排2~3天,主要程序包括预备会议、所长报告、重点学科领域报告、参观与座谈、专家组讨论、与院机关和所领导班子反馈6个环节,同时专家组现场评估后1个月内反馈专家组报告。之后专家意见会反馈到研究所,由研究所针对专家意见,提出解决方案。经过三年37个研究所国际专家诊断评估探索与试点,实践表明国际评估提升了研究所追求卓越的理念,促使其产生重大成果产出,带动了研究所科学文化建设和科学精神培养。“一三五”目标完成情况验收评估在规划结束时开展,并不针对研究所整体,而是针对重大突破和重点培育方向进行专项验收评估。

## 4 我国科研评价的挑战与机遇

过往40年,得益于政府、科学共同体和社会公众的共同努力,我国科研评价取得了有益的经验与可喜的成绩。整体上,科研评价在科技管理中的地位越来越显著,科研评价活动的规范管理与制度化建设呈现健康发展的趋势。从前文所述关于基金委项目评审的同行评议和中科院研究所综合评价实践,可以看到,同行评议在中国的实践日趋成熟与完善,应用也越来越普遍,无论是在方法或制度层面,同行评议都在不断改进与优化,也得到了政府和社会公众等各界越来越多的认可。其次,伴随着中国科技事业的上升和科技体制改革的深入,中科院研究所综合评价与中科院同期的发展战略紧密结合,有效发挥了科研评价作为宏观工具在科技管理中的导向作用<sup>[32]</sup>。再者,随着我国科技实力的提升和全球国际化大趋势的发展,作为国际科技界通行惯例的国际评估,也在科研评价活动中逐步探索引入,尽管在国际评估方法和程序上仍有改进的空间,但国际评估对提升中国科技实力和国际声誉的正面意义是值得肯定的。



在看到科研评价经验与成绩的同时需要意识到,我国科研评价仍然面临不容忽视的挑战。首先,从评价内容看,科研评价不应单纯追求数量,而要更加注重科技对经济社会的贡献,即科学的影响力,这成为各国政府关注的热点和重点问题。国际上已有相关探索与实践,但科学影响力的边界难以清晰界定,以及如何设计最合适的影响力分类方式和评估指标体系,是政府和科学界的难题。对于中国而言,在建设世界科技强国的征途上,如何探索中国特色科学影响力的评估制度和方法,使之与中国学科体系和发展阶段,以及国家社会发展需求相适应,尚需要长期的理论研究和试点探索。其次,从评价相关利益方来看,政府作为科学研究的资助者、监督者、协调者和消费者,在很多科研评价实践活动中发挥了较强的主体作用。科学共同体作为科学研究的生产者、组织者和评价者,在科研评价活动中的话语权还较弱。其原因之一在于目前我国政府和科学共同体在科研评价中的关系还没有从根本上理顺,我国科研管理行政色彩过重,在科研评价上政府的诱导力量强大,科学共同体的自主性得不到充分发挥。科研评价最终目的是为促进科学繁荣,这需要政府在评价中少些干预,进一步深化“放管服”改革,给予科研人员和机构更多自主权,让科学共同体发挥在科研评价中的“决定性作用”,同时科学共同体也要尽快成熟与成展起来,在科研评价中有所担当与作为。

尽管现阶段我国科研评价存在很多问题,科研评价改革的探索也已在中国陆续开展起来。一批处于优势地位的科研机构或高校是科研评价改革探索的核心力量。如,中科院于2012年即提出“重大成果产出导向”研究所评价体系并通过引入国际专家开展评估工作,清华、上交大和南京大学等高校针对整个创新链中不同阶段和不同类型的科技活动<sup>[33-35]</sup>,开展了各具特色的分类评价尝试。这

些探索发挥了重要的排头兵作用。同时,中央政府部门最近也颁布了一些新政策来规范各种评价活动。目前,改革的两个主要动力源——政府和科学共同体有必要加强合作互动,以寻找最有效和可操作的方式,使这项改革能真正有效地解决存在的问题。

中国需要探索建立与国家和科技自身发展相适应的新型科研评价体系,并在国际科研评价制度改革中发挥积极作用。从国际上看,科研评价也是很多国家多年来面临的巨大挑战。过于注重论文数量、期刊的影响因子等量化指标而非科学研究本身,这一趋势已成为国际上的普遍问题<sup>[36]</sup>。各国政府也注意到这一问题,并意识到其对满足经济发展和社会需求的负面影响,但均没有找到有效的解决方案。中国有可能在解决这个世界性难题上先行一步,找到一条在评价导向上既注重科学质量与卓越又注重国家经济社会影响的新型道路,为促进全球科学的健康发展贡献一份中国力量<sup>[37]</sup>。中国在解决这个问题上的优势主要体现在这几个方面:一是自2003年以来,我们一直关注这个问题,并且已经对问题的根源和可能的解决方案有了丰富的理解,二是中国科技界已经就解决该问题形成高度共识,三是中国政府对这个问题的高度重视从而形成了改革评价制度的良好政治基础。在2018年的两院院士会上,习近平总书记再次郑重发出了改革科研评价的号召。科研评价改革是一件极具挑战的大事,不能孤军奋战。中国需要和全球科学界同仁一道建立起全球化的科研评价治理体系,相互合作,分享经验,共同应对挑战。

### 参考文献

- [1] 李真真. 怎样评价基础科学研究?[J]. 中国高校技术市场, 2001(9):20-21.
- [2] 徐芳,李晓轩. 跨越科技评价的“马拉河”[J]. 中国科学院院刊,2017,32(8):879-886.
- [3] 龚放,曲铭峰. 南京大学个案:SCI引入评价体系对中国大

- 陆大学基础研究的影响[J]. 高等理科教育, 2010, 1(3): 4-17.
- [4] 知识创新工程评估工作组. 中国科学院开展知识创新工程(1998—2010年)评估工作[J]. 中国科学院院刊, 2011, 26(2): 127-133.
- [5] 国际评估专家委员会. 科学基金资助与管理绩效国际评估报告[R]. 北京: 国家科技评估中心, 2011.
- [6] 罗伯特·弗洛德曼, J·布瑞特·霍尔布鲁克, 卡尔·米切姆, 著. 夏国军, 朱勤, 译. 同行评议、研究诚信与科学治理: 实践、理论与当代议题[C]. 北京: 人民出版社, 2012.
- [7] 中国科学院科学基金委员会办公室. 改革科研管理的有效尝试: 科学基金工作情况汇报[J]. 科研管理, 1984(3): 34-40.
- [8] 胡维佳. 中国科技政策资料选辑(上中下)[M]. 济南: 山东教育出版社, 2006.
- [9] 国家自然科学基金委员会. 国家自然科学基金委员会关于申请项目评审工作暂行办法[EB/OL]. [2018-9-10]. [http://www.110.com/fagui/law\\_145850.html](http://www.110.com/fagui/law_145850.html).
- [10] 唐教庆. 完善科学基金制促进科学研究为四化建设服务[J]. 中国科学基金, 1987(1): 26-40.
- [11] 国家自然科学基金委员会. 国家自然科学基金规章制度汇编[M]. 北京: 法律出版社, 2016.
- [12] 师昌绪. 我的感言[J]. 中国科学基金, 2010, 24(1): 3-4.
- [13] 唐教庆. 完善科学基金制促进科学研究为四化建设服务[J]. 中国科学基金, 1987(1): 26-40.
- [14] Baldwin M. In referees we trust?[J]. Physics Today, 2017, 70(2): 44-49.
- [15] 吴京生. 关于中国的科技体制改革问题[J]. 科技导报, 1986, 4(2): 62-64.
- [16] 师昌绪. 在人生道路上: 师昌绪自传[M]. 北京: 科学出版社, 2011.
- [17] 龚旭. 科研资助管理与学科发展战略: 国家自然科学基金委员会的学科发展战略研究考察[J]. 中国科学基金, 2016(5): 410-416.
- [18] 王新, 张蓁, 唐靖. 追求卓越三十年: 国家自然科学基金委员会发展历程回顾[J]. 中国科学基金, 2016(5): 386-394.
- [19] 龚旭. 同行评议与科学基金政策研究[J]. 中国科学基金, 2007(2): 91-94.
- [20] 李芬, 朱紫阳, 丁枝秀. 关于中国同行评议研究状况的分析报告: 基于CNKI核心期刊的文献研究[J]. 中国科学基金, 2009(3): 177-182.
- [21] 吴述尧. 同行评议方法论[M]. 北京: 科学出版社, 1996.
- [22] 冷疏影, 俞肇元, 胡勇, 等. 地理学学科语义网及其在科学基金智能辅助指派中的应用[J]. 中国科学基金, 2015(1): 48-53.
- [23] 温明章, 乔建军, 刘文娟, 等. 国家自然科学基金评审会前网络投票的尝试与分析[J]. 中国科学基金, 2015(3): 196-200.
- [24] 中国科学院. 中国特色国家创新体系建设的成功实践: 知识创新工程(1998—2010年)评估报告[R]. 北京: 科学出版社, 2012.
- [25] 李晓轩, 杨国梁. 中国科学院研究所评价的逻辑模型研究[J]. 科学学与科学技术管理, 2009, 30(4): 5-8.
- [26] 李晓轩, 石兵. 中国科学院研究所评价浅议[J]. 中国科学院院刊, 2003, 18(2): 131-134.
- [27] 知识创新工程评估工作组. 中国科学院开展知识创新工程(1998—2010年)评估工作[J]. 中国科学院院刊, 2011, 26(2): 127-133.
- [28] 中国科学院知识创新工程评估方法研究课题组. 中国科学院知识创新工程评估方法[J]. 中国科学院院刊, 2012, 27(4): 485-492.
- [29] 中国科学院科技评价研究组. 关于我院科技评价工作的若干思考[J]. 中国科学院院刊, 2007, 22(2): 104-114.
- [30] 柳卸林, 支婷婷. 中国科学院的知识创新工程与能力提升[J]. 中国软科学, 2009(1): 56-65.
- [31] 白春礼. 以重大成果产出为导向改革科技评价[J]. 中国科学院院刊, 2012, 27(4): 407-410.
- [32] Xu F, Li X. The changing role of metrics in research institute evaluations undertaken by the Chinese Academy of Sciences (CAS) [J]. Palgrave Communications, 2016(10): 1-6.
- [33] 清华大学物理系. Tenure-track 制度激活教师队伍[EB/OL]. [2017-08-11]. [http://www.tsinghua.edu.cn/publish/thunews/9914/2011/20111111084759369801776/20111111084759369801776\\_.html](http://www.tsinghua.edu.cn/publish/thunews/9914/2011/20111111084759369801776/20111111084759369801776_.html).
- [34] 南京大学. 陈骏校长在全国“两会”期间畅谈高教改革和创新发展[EB/OL]. [2017-08-11]. <https://xgc.nju.edu.cn/65/bd/c1519a91581/page.htm>.

- [35] 张杰. 院系中长期国际评估的思考[EB/OL]. [2017-08-11]. 533-534.  
<http://plan.sjtu.edu.cn/info/1004/1004.htm>.
- [36] Fathi S. San Francisco declaration on research assessment[J]. Journal of Experimental Biology, 2013,216(12): 605-605.
- [37] Cheng J P, Li X X, Xu F. Science-evaluation reform on the road in China[J]. National Science Review, 2018(5): 605-605.

## 40 Years Reform and Development of Research Evaluation in China: Based on Case Studies on NSFC Peer Review and CAS Research Institutes' Comprehensive Evaluation

XU Fang<sup>1,2</sup>, GONG Xu<sup>3</sup>, LI Xiaoxuan<sup>1,2</sup>

(1. Institutes of Science and Development, Chinese Academy of Sciences, Beijing 100080, China; 2. School of Public Policy and Management, University of Chinese Academy of Sciences, Beijing 100080, China; 3. Policy Bureau, National Natural Science Foundation of China, Beijing 100085, China)

**Abstract:** Since 'the Spring of Science' in 1978, scientific research evaluation has undergone 40 years of reform and development in China. With an overview of the three stages of the development of scientific research evaluation in China, this paper expounds the two representative and contemporary scientific research evaluation practices, namely, peer review by NSFC and Comprehensive Evaluation of CAS research institutes, and then summarizes the achievements of the reform and development of scientific research evaluation in the past 40 years, as well as the challenges faced in the new era.

**Key words:** research evaluation; reform and open-up in China; peer review; NSFC; comprehensive evaluation of CAS research institutes