

· 基金纵横 ·

科技计划(基金)绩效评估的国际实践

——参加八国集团研究评估 2008 年会议的报告^{*}

吴建南

(西安交通大学, 西安 710049)

受国家自然基金委员会委派, 应美国国家科学基金会邀请, 在西安交通大学科学与技术研究院资助下, 我本人于 2008 年 11 月 3—4 日在美国华盛顿特区参加了八国集团研究评估工作会。该会议主题为“科学政策的科学”, 来自美国国家科学基金会、英国科学研究中心、德国科学基金会和马普学会、法国教育部、日本文部科学省、日本科技振兴机构、日本学术振兴会、以及加拿大、意大利、欧盟的科技计划(基金)绩效评估管理者就有关绩效评估的方法、数据来源与结果使用, 以及未来评估完善空间等话题进行了交流与讨论。我本人做了题为“中国国家自然科学基金管理科学部项目绩效后评估”的报告, 介绍了我国国家自然科学基金委员会项目绩效评估的研究与实践状况。根据会议发言与交流情况, 总结如下。

1 各国开展科学基金与科技计划绩效评估的情况

1.1 绩效评估的基础与方法

(1) 相关立法保障和政策支持在推动发达国家全面开展科技计划(基金)绩效评估的过程中发挥了重要作用。目前, 美国国家科学基金会开展的绩效评估, 是应历届国会的法案尤其是 1993 年《政府绩效与结果法案》(GPRA)的要求, 在“总统行政管理议程”(PMA)和预算与管理办公室(OMB)执行的“项目评级工具”(PART)的推动下进行的。

按照上述法案和计划要求, 美国科学基金采用年度财务报告、年度绩效报告、关于绩效和财务重点的公民报告, 以及在下一年度递交国会预算申请中包含的前一年度绩效报告和下一年度绩效计划 4 种途径展示其绩效状况。绩效评估紧密围绕上述报告

展开, 分为定性评估、定量评估和综合评估 3 类。其中, 外部专家委员会和顾问委员会的专家负责定性评估, 每年评审约三分之一的资助项目, 将三年定位一个评估周期; 科学基金内部的数据库和记录主要完成定量评估; 综合评估则由绩效评估顾问委员会等外部专家完成。

日本 1995 年颁布了《科学技术基本法》, 明确了科技评估的地位。近期, 科技计划绩效评估已经成为国家研究与发展评估的重点。日本于 2001 年成立了国家科学与技术政策委员会, 由日本首相担任主席, 各个中央政府部门和重大项目绩效评估的开展得到了进一步推动。

(2) 发达国家科技计划(基金)绩效评估开始关注计划特征和评估方式的合理配置。大型研究计划和涉及跨国合作的项目评估采用短期和长期兼顾的绩效评估方式。以欧盟 2006 年开展的“第七次研究框架项目”(简称框架项目)为例, 该项目是欧盟层面专门服务于跨年跨国科学项目评估的大型评估活动。从 1995 年起, 框架项目把 5 年 1 次评审和每年定期监测的途径相结合, 在 2006 年框架项目中, 将 5 年 1 次的评审改为项目结题后两年执行后评估, 将原本的外部专家年度监测改为评估组织机构进行成果监测。关注项目产出和效果的绩效评估则更多采用前评估和后评估相结合的方式。

以日本科学技术振兴机构的项目成果评估为例, 该机构在 8 年中分 3 次, 每次资助 5 到 10 个为期五年半的项目, 在项目实施前半年和完成后分别进行前评估和后评估, 后评估过程采用专家委员会的方式进行。加拿大科技产出评估也采用了和日本类似的后评估模式。

(3) 发达国家在绩效评估过程中注重将评估组

^{*} 国家自然科学基金资助项目和教育部新世纪优秀人才资助项目。
本文于 2009 年 5 月 10 日收到。

织机构和计划、基金的相关者结合,建立绩效评估工作网络。例如,美国国家科学基金会近期开展的“科学与创新政策”评估采用新理念,关注人力资本、社会活动和知识流动的情况。评估引入了研究专家和受到科学与创新政策影响的国内外企业和政府部门共同参与,形成了一张跨部门、跨行业、跨学科的工作网络。

德国科学研究基金近期推动了“德国卓越计划”的实施,根据大学的具体特征制定了具有针对性的绩效改进目标,致力于让高水平研究型大学提高其国际竞争力,而让一般水平的大学提供高水平的研究培训,让大学中的研究团队成为优秀的研究中心,得到了相关参与计划大学的认可。

日本的国家大学科研组织评估体系中,包括现在的和未来的学 生和生源家庭、校友和他们的雇员、社会和产业等利益相关者都参与其中,成为绩效评估的主体。

1.2 绩效评估的数据来源与分析应用

(1) 在研究论证基础上,发达国家的绩效评估正在向两个“多”转变。首先是采用多种性质的绩效信息。以美国为例,通过对评估数据的专门研究,美国国家科学基金会指出“科学项目评估的最大障碍是缺少可靠数据”,近期国家科学基金在基金层面和项目层面的绩效评估在不断扩大信息来源,将定性和定量两种信息相结合。其基金层面评估包括对长期战略产出的定性描述,以及对年度目标执行情况的定性和定量评估;项目层面评估包括专家定性评估,以及对申请书和项目报告进行的定性和定量分析。同时,美国科学基金会还同多家教育科研机构联合开展了合作调查项目,采用新的步骤进行评估数据收集工具的完善,以保证评估信息的可靠性、真实性和全面性。加拿大科技产出评估采用多元混合方法,定性评估主要采用案例研究、座谈会等方式;定量评估则通过问卷调查、数据分析、计量模型、科学计量学研究等方式开展。

第二个“多”,是指评估指标体系覆盖了更多的绩效维度。绩效是一个由投入、过程、产出、结果、社会影响等方面组成的综合性概念。发达国家在选择绩效评估指标时关注对不同维度的综合考察和全面反映。在欧盟进行的第七次框架项目评估中,绩效指标涵盖了产出和结果两个方面,其中产出指标由科研管理系统生成的电子数据获得,结果指标由调查问卷、访谈和文献引用情况获得。意大利对本国研究与发展基础设施的评估也从投入、产出、活动、产出收益、质量和数量等多个维度入手设计评价指

标。关注项目产出的绩效评估不仅关注产出本身,也关注引起产出的过程,如日本科学技术振兴机构的评估不仅关注了结果指标,同时还采用研讨会、调查问卷等方式对过程指标进行评分。由英国科学理事会发起的经济影响评估也从支出、产出、经济影响等角度对不同绩效维度进行了考察。

(2) 在获得更为可信和全面的绩效评估结果的基础上,发达国家对评估结果的重视程度有所加强,使用范围有所扩大。在实践领域,绩效评估结果通常与部门预算、行动优先权安排挂钩。以日本为例,其国家层面的科学 研究与发展评估由科学和技术政策委员会主导。通过对部门政策进行前评估要求财政部回顾预算情况。由专家委员会进行大型项目评估用于制定国家研究与发展过程中资源配置优先权。部门层面的评估主要针对具体项目的执行情况进行,评估结果用于改进部门与拨款机构之间的沟通方式。在日本国立大学评估中,评估结果既向大学公布,用于修改第二年度的大学中期目标和执行计划,同时反馈给有关管理部门,用于对下一轮中期目标进行必要调整,以影响未来目标设置和资源分配等决策。欧盟框架项目的评估结果在影响欧盟成员国获得科技计划拨款、制定政策发展优先权、重构国内科研组织和增强国际合作、进行不同国家之间科研水平的标杆管理等方面,也产生了显著影响。

绩效评估的结果还用来推动科学学理论的新发展,以美国国家科学基金会为例,在完善科学与创新政策评估体系的过程中,研究人员将评估结果用于创造有关知识转移和生产的研究理论,并预测未来科技创新领域的可能增长点,进行社会网络分析,开展了大量有关人力资源和教育系统改进的国际比较实证研究,推动了科学学理论的创新。

1.3 对完善绩效评估的反思与未来实践构想

(1) 发达国家不同程度地认识到了绩效评估的实践困难。意大利国立评估委员会指出,科学项目评估经常采用自上而下推动的评估方式,导致信息不完备、不对称以及专家对于数据理解和解释的不一致,因此造成结果难以客观准确。同时,自上而下的评估模式较少关注被评估对象的参与空间,并不能充分调动其积极性。

在日本科技研究与发展评估的过程中,管理者认识到,制约评估发展的主要的挑战有两点:一是难以对研究与发展“情境”进行全面的分析和描述,以把握投入、活动、产出、顾客接受、短期效果、顾客效果和问题解决的效果等关键时点之间的关系。其次

是零散科研项目如何产生扩散效应的机制并没有得到充分认识,绩效评估往往追求“简便易行”,可能会以牺牲更为准确的发现和结论为代价。尤其是在引入竞争机制、关注拨款的经济性以及考虑不同政策关系的前提下,可能遭遇更多的挑战。

美国国家科学基金会进一步指出,以往经验让人们对未来绩效评估的开展前景难以充分乐观。这些经验包括:传统的评价工具都有致命缺点,导致缺陷严重的分析结果;联邦政府对把科学政策分析和评价联系起来的容纳力欠佳,直接原因是资源不足、时间不足和方法不足;由于对科学项目绩效决定性因素的使用、定义和理解不清晰,评估难以解决所有问题。此外,跨国项目评估中的成员国参与有限、评估中的协调机制运行不灵、难以充分评估对社会经济系统的影响等问题也成为制约科学项目评估完善的瓶颈。

(2)发达国家正在努力应对绩效评估中的现实挑战。主要做法包括积极开展对绩效评估的科学研究、增强绩效评估机构的独立性以及进一步关注有关绩效评估的国际交流与合作等。美国国家科学基金会和英国研究理事会是关注绩效评估研究的典型。在同高等学府、研究机构和学者合作进行的长期研究中,双方通过共享研究信息和知识产权,全面提高科学项目评估的科学性、信息有效性和使用价值。

日本和意大利较为关注评估机构运行的独立性。其中,日本在其中央政府和各个中央政府部门之下分别设立了负责科技项目评估的独立法人机构。上述机构既可以直接负责组织评估,必要时还可以将评估工作委托独立性更强的外部评估组织进行。意大利也在组建一个更具权威性的评估机构,以更好地对科技项目进行前评估。

在国际交流与合作方面,来自美国国家科学基金会和欧盟的实践专家指出,应该关注全球化背景下科学知识的形成过程,进一步共享有关评估的知识。通过此次会议,参会国家认识到本国的绩效评估体系应注意借鉴其他国家的经验和做法,取长补短。因此,搭建有关参会国家的交流平台也称为本次研讨会的一个共识。

2 推进我国科学计划(基金)绩效评估的建议

改革开放30年见证了中国科技计划(基金)的快速增长。近年来,为了规范和加强科技计划(基金)经费绩效考评的组织和管理,提高科技计划经费的使用效益,财政部出台了一系列意见和办法。此次八国集

团研讨会充分表明,财政部强化对科技计划(基金)经费的绩效考评管理工作,是在科学发展观引导下的重要举措,能够促进相关部门进一步树立问责意识。通过本次赴美参加研讨会,我们根据发达国家的实践情况,结合我国科学计划(基金)绩效评估的需求,对进一步做好评估工作提出如下建议。

2.1 进一步认识绩效评估的复杂性,谨慎而有效地推动实际工作

绩效评估如同一把“双刃剑”,在带来正面效果的同时,往往也伴随着不可预期和难以控制的负面效果。本次八国集团的有关发言提供了清晰的证据,即越是在科技计划(基金)绩效评估制度发展先进、应用广泛的国家,绩效评估暴露出的问题和挑战越多,管理者对这些挑战的理解也越深入、越全面。八国集团高度关注绩效评估并开展积极讨论的原因,不仅是出于对绩效评估正面效果的期待,同时也是为了解决现有绩效评估中的问题和困惑。

从这一意义上讲,为了更为有效的应对可能的复杂性,增强目前有关办法和规定的实践效果,应从两个方面入手。首先,进一步关注我国科技计划(基金)管理部门的绩效评估现状,对不同类型评估活动的开展情况进行摸底。根据主管部门的组织结构、管理模式、经费额度等分类梳理绩效评估实践中的问题。对计划、课题承担单位的需求进行充分的了解和把握。另一方面,考虑到美国具有稳固的法案基础和长期的实践经验,建议进一步分析美国1993年政府绩效与结果法案对科技计划经费绩效评估的影响,进一步提高有关政策对绩效评估的关注度和指引力。

2.2 关注科学计划(基金)绩效评估机制的设计和完善,强化绩效评估的管理效果

本次八国会议表明,发达国家在实施科技计划(基金)绩效评估之前,会对绩效评估的运行机制进行充分论证,并开展不定期的回顾和完善。从绩效评估的定位和目的出发,在凝练清晰评估定位的基础之上,系统设计与之相适应的绩效评估指标体系、信息收集机制和结果使用方式,已经成为美国等发达国家的基本经验。其中,不同利益相关者的参与、多来源信息和多维度指标的采用,根本目标是还原科技计划(基金)绩效的形成过程。对于我国而言,在提高资金使用效益的目标定位下,目前应该进一步完善绩效评估机制,考虑科技计划(基金)运行中的多主体参与、绩效产生存在时间滞后等特征,使之适应科技计划(基金)绩效改进的战略需求,更好地

服务于创新型国家的建设。

2.3 进一步加强国际交流和合作,促进绩效评估知识的共享

此次中国首次应邀参加八国集团的讨论,是我国进一步与发达国家交流的开始。尽管我国的绩效评估同其他参会国家还存在着内容和形式上的不同,但在对绩效评估的理解、现实绩效评估的困难和复杂性、未来完善绩效评估的意愿等方面,中国和其他参会国家都有相似的需求和共同的愿景。

通过此次会议,参会国家认识到受到全球化和愈发深度的国际合作等因素的影响,本国科技计划(基金)绩效评估不仅需要在学习其他国家经验的基础上完善,同时也是产生全球化的科技管理知识的重要途径。因此,相互之间的取长补短具有多方面的积极意义。我国应该主动增进发达国家对我国绩效评估的认识,并在未来推动可能的国际合作,更好地服务于人类社会文明和科学技术的发展。

INTERNATIONAL PRACTICE OF PERFORMANCE EVALUATION IN S&T PROGRAM (FOUNDATION): A REPORT FROM G-8 RESEARCH ASSESSMENT 2008 WORKSHOP

Wu Jiannan

(School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an 710049)

· 资料·信息 ·

北京大学城市与环境学院在中国陆地碳循环研究方面取得重要进展

2009年4月23日出版的*Nature*杂志刊发了北京大学城市与环境学院朴世龙研究员与方精云教授发表的论文“*The carbon balance of terrestrial eco systems in China*”,报道了他们在中国陆地碳汇研究方面取得的新进展。核心内容之一,不仅是因为它的微小变化就能导致大气CO₂浓度的明显波动从而进一步影响全球气候的稳定,而且由于它与“京都议定书”(限制一个国家化石燃料使用的国际公约)有着紧密联系。我国是当今最大的工业源CO₂排放国家之一,因此阐明中国陆地生态系统在全球碳循环中起碳汇还是碳源的作用是中国科学家乃至国际社会普遍关注的重大环境问题。北京大学城市与环境学院全球气候变化研究小组,利用已有的土地利用和资源清查数据、大气CO₂浓度观测数据、遥感数据、以及气象数据,借助遥感、GIS等新技术的支持,并结合大气反演模型(atmospheric inverse model)和基于过程的生态系统碳循环模型,综合研究中国陆地碳汇/源的时空格局及其机制。研究结果表明,上世纪80年代和90年代,中国陆地生态系统碳储量平均每年增加0.19—0.26PgC(1PgC=1015gC)。中国陆地生态系统碳汇大小相当于此间中国工业源CO₂总排放量的28%—37%,显著地高于欧洲(7%—12%),跟美国相近(20%—40%)。中国陆地生态系统的碳汇主要与我

国人工林的增加、区域气候变化、CO₂浓度施肥促进植被生长、以及植被恢复尤其是灌丛的恢复有关。此外,增加秸秆还田等农业管理措施也增加了我国农田生态系统土壤碳储量的积累。

这项研究不仅首次采用自上而下的大气反演模型和自下而上的过程模型及地面资料有机结合的途径,系统地分析了我国陆地生态系统碳汇大小及其机制,提高了对陆地生态系统在全球碳循环中作用的认识,而且阐明了中国陆地生态系统净吸收的二氧化碳量可以部分抵消其工业源排放量,为制定CO₂的排放策略提供依据,并增加我国在联合国气候变化框架协议谈判中的砝码。*Nature*杂志在同一篇专门发表来自著名碳循环专家Dr. Gurney的评述,报道了该文章的重要性及其意义。*Nature*杂志编辑认为“这份中国陆地碳循环的综合评估论文的发表填补了全球碳平衡中一个重要的空白地区”,而著名的全球变化专家Dr. Marland在*Nature*杂志的This week's news栏目中评论“这是一篇令人‘impressive’的文章”。

该成果受到国家自然科学基金项目的资助。

(北京大学科学部供稿)