

国家中文核心期刊 中国人文社会科学核心期刊  
中国科技论文统计源期刊 中文社会科学引文索引期刊  
国家自然科学基金委员会管理科学部认定的管理科学 A 级重要期刊



ISSN 1003-2053

# 科学学研究

## STUDIES IN SCIENCE OF SCIENCE

- 技术科学发展模式初探
- 科学是客观的还是社会建构的?
- 专利技术产业化机理、关键问题与驱动策略
- 国家工程研究中心政策的历史演进及协调状况研究
- 创业企业架构能力、元件能力与绩效关系实证研究
- 企业自主创新的可拓创新模型构建与应用研究

中国科学学与科技政策研究会  
中国科学院科技政策与管理科学研究所  
清华大学科学技术与社会研究中心

合办

5  
2010  
Vol.28 No.5

目次

科技论坛

技术科学发展模式初探 ..... 刘启华 马万超 (641)

科学学理论与方法

科学是客观的还是社会建构的? ..... 王娜 (650)

从“统治”到“治理”——疯牛病危机与英国生物技术政策范式的演变 ..... 高璐 李正风 (655)

美国科学不端行为调查程序的历史演进 ..... 王阳 程晖 (662)

引文分析学的知识流动理论探析 ..... 梁永霞 刘则渊 杨中楷 (668)

非专利引文衍生的科学期刊共被引网络分析 ..... 高霞 官建成 (675)

科技发展战略与政策

专利技术产业化机理、关键问题与驱动策略 ..... 于晓宇 蔡莉 谢富纪 (681)

标准联盟的兴起及诱发技术垄断的机制研究 ..... 张米尔 冯永琴 (690)

国家工程研究中心政策的历史演进及协调状况研究 ..... 汪樟发 汪涛 王毅 (697)

科学基金国际评估的框架、内容与方法 ..... 吴建南 马亮 郑永和 (704)

中国高技术产业劳动生产率非均衡特征分析 ..... 吕韬 曹有挥 吴建楠 (713)

产业集聚的经济效应区域化还是城市化? ..... 徐盈之 朱依曦 (719)

科技管理与知识管理

基于 Wiki 的知识共享系统研究 ..... 张红兵 和金生 (727)

跨国公司战略管理知识转移机制的案例研究 ..... 胡玲 金占明 (732)

技术获取方式与企业绩效的关系研究 ..... 赵文红 梁巧转 (741)

创业企业架构能力、元件能力与绩效关系实证研究 ..... 徐中 姜彦福 谢伟等 (747)

资源整合能力、创业导向与创业绩效的关系研究 ..... 易朝辉 (757)

企业家精神对外部知识能力及网络能力的作用 ..... 李伟 聂鸣 李顺才 (763)

技术创新与制度创新

企业自主创新的可拓创新模型构建与应用研究 ..... 周志丹 李兴森 (769)

外部社会联系能否提升企业自主创新? ..... 谢言 高山行 江旭 (777)

高科技企业创新生态系统治理机制研究 ..... 张运生 邹思明 (785)

城市体系中创新扩散的空间特征研究 ..... 程开明 (793)

文章编号: 1003 - 2053 (2010) 05 - 0704 - 09

# 科学基金国际评估的框架、内容与方法 ——基于多案例的跨国比较研究

吴建南<sup>1</sup>, 马 亮<sup>2</sup>, 郑永和<sup>3</sup>

(1. 西安交通大学公共政策与管理学院, 陕西西安 710049; 2. 西安交通大学  
管理学院, 陕西西安 710049; 3. 国家自然科学基金委员会计划局, 北京 100085)

**摘 要:** 从国际评估的委托方、评估目的、被评估方、评估方、评估议题、评估方法与信息来源、评估程序与周期、评估报告、评估效果与影响等维度, 对包括德意志研究联合会、日本学术振兴会、欧盟第六框架计划、挪威研究理事会、奥地利科学基金和国家自然科学基金中德科学中心等在内的 6 个科学资助机构的国际评估案例进行了系统分析, 总结了科学基金国际评估的共同特点与主要特色, 为设计国家自然科学基金资助与管理绩效国际评估的方案设计与组织实施提供了借鉴与启示。

**关键词:** 国家自然科学基金; 国际评估; 绩效评估; 公共部门; 案例研究; 跨国比较

**中图分类号:** N945.16

**文献标识码:** A

伴随着财政支出绩效评估与公共问责的强化, 基于战略导向的绩效评估和绩效管理日益成为世界主要国家科学研究资助机构的发展趋势<sup>[1-3]</sup>。美国国家科学基金会 (NSF) 一直走在各国科学资助机构管理的前沿, 在绩效评估方面也不例外<sup>[4]</sup>。在 1993 年颁布实施的《政府绩效与结果法案》(GPRA) 的规范下, NSF 通过战略规划、年度绩效计划和年度绩效报告等环节实现了组织战略与绩效管理的对接, 为其实现卓越管理和支持基础研究提供了基础<sup>[5][6]</sup>。其他国家的科学基金也在效仿 NSF 的做法, 并纷纷开展主要由国际评估专家主持或参与的国际评估。2000 年, 德意志研究联合会 (DFG) 和德国马普学会组织了一次国际评估, 对其发展产生了巨大的积极影响<sup>[2][7]</sup>。2002 年, 日本学术振兴会 (JSPS) 也实施了国际评估, 为其实现组织转型和持续发展奠定了基础<sup>[1][2]</sup>。其他国家和地区的科学研究资助机构也都实施了国际评估并取得了广泛影响<sup>[8-10]</sup>, 呈现出科技政策与管理领域的一大亮点。

作为我国基础研究的主要资助机构之一, 国家自然科学基金委员会 (NSFC) 也在 2008 年提出国际

评估的设想, 预计在 2010 年正式实施<sup>[11]</sup>。然而, 基础研究绩效评估一直是科技政策与管理领域的一个恒久难题, 对于以资助基础研究为核心职责的 NSFC 而言更是如此<sup>[3][5][12]</sup>。绩效评估的特点之一即是其较强的经验性, 它是一门经验科学而非实验科学; 我们无法假设或假想, 但却可以学习他人的经验, 通过别人少走弯路, 通过“试验 积累经验<sup>[6]</sup>”。此前不久, NSFC 与 DFG 对其合作共建单位——中德科学中心共同开展了国际评估, 为 NSFC 国际评估提供了实践经验<sup>[13]</sup>。但鉴于国际评估的规模性、复杂性和专业性, 极有必要进一步汲取和总结国内外的已有实践经验, 为 NSFC 国际评估少走弯路和顺利实施提供充分依据<sup>[2]</sup>。虽然由于国情上的差异或管理体制的不同, 我们无法照搬其他组织的评估套路, 但却可以提供基于证据的学习和借鉴, 特别是在国际评估的一般做法和主要经验等方面。

NSFC 的工作人员和一些学者已经对科学资助机构的绩效评估问题进行了理论探讨, 认为 NSFC 绩效评估极有必要且具备了一定的可行性<sup>[3][12][14-16]</sup>。一些研究人员对国外一些科学资助

收稿日期: 2009 - 07 - 22; 修回日期: 2009 - 09 - 21

基金项目: 国家自然科学基金委主任基金项目 (J0910011); 国家自然科学基金项目 (J0910012, 70873092); 西安交通大学腾飞人才特聘教授项目

作者简介: 吴建南 (1970 - ), 男, 陕西西安人, 教授, 博士生导师, 博士, 研究方向为政府创新与绩效管理。

马 亮 (1984 - ), 男, 江苏铜山人, 博士生, 研究方向为绩效管理、组织创新与科技评估。

郑永和 (1964 - ), 男, 辽宁丹东人, 副教授, 硕士, 研究方向为科技政策与管理。

机构的绩效评估实践进行了考察和研究,如 NSF<sup>[4]</sup>, JSPS<sup>[1][2]</sup>, DFG<sup>[2][7]</sup>, 加拿大自然科学与工程研究理事会<sup>[17]</sup>, 等等。然而, 上述研究缺少一个统一的理论分析框架, 且主要是个案分析, 缺少多个案例的深入考察和系统比较, 在代表性和可推广性方面尚有待于进一步改善。本文期望在已有研究的基础上, 通过跨国多案例比较研究, 识别主要国家科学资助机构国际评估的共同属性和主要特征, 为设计 NSFC 国际评估总体方案提供经验借鉴。

具体而言, 结合已有公共部门绩效评估的理论和方法, 本文从科学基金国际评估的委托方、评估目的、被评估方、评估方、评估议题、评估方法与信息来源、评估程序与周期、评估报告、评估效果与影响等维度, 对包括德意志研究联合会 (Deutsche Forschungsgemeinschaft, DFG)、日本学术振兴会 (Japan Society for the Promotion of Science, JSPS)、欧盟第六框架计划 (the Sixth Framework Programmes, FP6)、挪威研究理事会 (Research Council of Norway, RCN)、奥地利科学基金 (Austrian Science Fund, FWF) 和国家自然科学基金 (NSFC) 中德科学中心 (Sino - German Center for Research Promotion, SGCRP) 等在内的 6 个科学资助机构的国际评估案例进行了系统分析。基于绩效评估的理论和方法, 结合国外实践经验和国家自然科学基金的实际情况, 本文对评估方案的设计思路和绩效报告的框架结构进行了阐述。

## 1 研究设计

### 1.1 案例选取与资料来源

尽管在社会科学研究中得到了广泛而持久的应用, 但案例研究却往往为人诟病<sup>[18]</sup>。其中的主要原因是案例选取的代表性不足, 难以反映总体的基本情况, 缺乏外推性。即案例虽可以点带面, 但也可能以偏概全。为了尽可能选取有代表性的国际评估案例进行研究, 我们采取多种方法和渠道获得案例来源, 确保案例之间的可比性和可鉴性。

通过咨询国内外基础研究绩效评估的有关专家, 结合已有的研究文献和笔者的网络检索结果, 本文选取以下 6 个科学资助机构的国际评估实践, 作为案例研究和比较分析的主要对象。它们是: 2000 年德意志研究联合会的国际评估; 2002 年日本学术振兴会的国际评估; 2001 年挪威研究理事会的国际评估; 2003 年奥地利科学基金的国际评估; 2005 年

中德科学中心的国际评估; 2009 年欧盟第六框架计划的后评估。上述案例具有较高的典型性和代表性, 从中分析的结果有助于 NSFC 国际评估的开展。而且这些案例的相关资料较为翔实, 在数据可及性较有保障。

本文的资料来源包括 NSFC 出国考察报告和上述案例产生的评估报告。具体来说, 我们的资料来源包括: 日本学术振兴会<sup>[1][2][19]</sup>; 德意志研究联合会<sup>[2][7]</sup>; 国家自然科学基金中德科学中心<sup>[13][20]</sup>; 挪威研究理事会<sup>[9]</sup>; 奥地利科学基金<sup>[8]</sup>; 欧盟第六框架计划<sup>[10][21]</sup>。上述文本资料为本文的案例研究提供了充分的信息和数据, 可以满足案例研究的一般数据需求<sup>[18]</sup>。此外, 研究中需要的其他资料则来源于上述科学资助机构的官方网站、年度报告和研究文献。

### 1.2 研究维度与分析策略

正如其他类型的绩效评估一样, 国际评估是一个综合集成系统, 包含许多可资分解和深入分析的评估要素。已有研究总结了公共部门绩效评估的要素框架, 包括评估目的、评估对象、评估内容、评估组织者、评估主体、绩效信息收集方式、评估结果使用方式等<sup>[6]</sup>。因其国际性, 国际评估有有别于其他绩效评估活动的特征, 需要在国际评估的案例分析中予以重视。

国际评估是相对于国内评估而言的, 但却并非单纯邀请国际专家“充场面”、“装门面”, 提高评估的档次或突出评估的国际化。许多科学资助机构之所以选取国际评估的形式, 主要是为了规避国内评估的利益相关性和评估“虚化”, 凸显国际评估的公正性、客观性和国际视野。“科学无国界”, 科学研究资助与评估也应该“无国界”。也因此, 国际评估越来越成为科学研究领域绩效评估的一个重要发展方向, 受到国内专家的注意<sup>[3]</sup>。

通过考察上述科学资助机构国际评估的“方法论”(methodology)部分, 结合公共部门绩效评估的通用要素<sup>[6]</sup>, 我们识别了国际评估的主要构成要素, 而它们也是未来国家自然科学基金国际评估设计与实施的重点内容。我们认为, 科学资助机构国际评估的分析维度包括如下方面: 评估目的、委托方、被评估方、评估方、评估议题、评估方法与信息来源、评估程序与周期、评估报告、评估结果与影响。具体来说, 上述维度包括以下内容。

(1) 委托方是国际评估的委托人或组织者, 他们不具体参与国际评估, 但却规定了国际评估的目

的、任务和要求,并为国际评估提供经费和资料等方面的支持。对于科学资助机构国际评估来说,委托方通常根据法律依据而提出,或者是其上级主管部门,有时也可能是科学资助机构自身。

(2)评估目的是国际评估期望达到的效果或实现的目标,如回顾总结、宣传激励、绩效改进等。评估目的可以是显在的,也可以是潜在的<sup>[6]</sup>,本文的分析主要是基于评估报告的显在目标,但也有结合考察报告对潜在目标的推测。

(3)被评估方即科学资助机构,它们在设立目标、主要职责、组织规模、成立时间等方面都存在许多差异,并由此影响国际评估的设计与实施。我们关注的是基础研究资助机构,它们的主要区别在于体量的大小和成立时间的早晚。

(4)评估方指具体实施国际评估工作的主体,通常包括国际评估专家、国内评估专家,以及科学资助机构的工作人员等。他们承担不同的评估职责,通过分工协调,共同完成国际评估的预定任务。

(5)评估议题即评估的内容领域和具体指标,它们是科学资助机构或评估委托方最为关心的关键议题,也是未来评估结论的主要指向。评估议题的数量是否合适、相互关系如何等是考量国际评估成败的重要方面之一。

(6)评估方法与信息来源。评估方法一般分为定性评估、定量评估,以及定性与定量相结合的评估方法<sup>[3]</sup>。国际评估的信息来源是决定评估能否进行并获得可靠有效的评估结论的前提条件,通常包

括科学资助机构运行管理自动积累的数据和国际评估专门收集的数据。

(7)评估程序与周期。国际评估通常由一系列活动环节构成,它们构成了一个有机链条,按部就班、有条不紊,使国际评估顺利实施。上述评估环节所耗费的时间加总构成评估周期,也是衡量国际评估是否可行或成本高低的重要指标之一。

(8)评估报告是国际评估的核心产品,它可能是一份独立报告,也可能是一系列评估报告。一般而言,鉴于国际评估的大规模和复杂程度,通常首先由科学资助机构完成自评估报告或绩效报告,或者邀请独立专家形成独立评估报告,然后在此基础上邀请国际评估专家作出判断并提出建议。

(9)评估效果与影响。国际评估实施以后,可能会产生最初设计时的预期效果和影响,但也可能引发意料之外的影响。这些影响可能是正面的,也可能是负面的,需要在设计国际评估方案时预先估计并规避潜在风险。

上述内容相互匹配,有机结合,共同组成了科学资助机构国际评估的分析框架(如图 1 所示)。国际评估分为三个部分:参与主体部分包括委托方、评估方和被评估方,它们之间存在委托—代理关系;评估方案部分包括评估目的、评估议题、评估方法与信息来源、评估程序与评估周期等,其中评估目的是核心,它决定了评估议题及后续要素的内容;评估结果部分包括评估报告、评估效果与影响,它们是评估方案的产物,并会对参与主体产生影响。

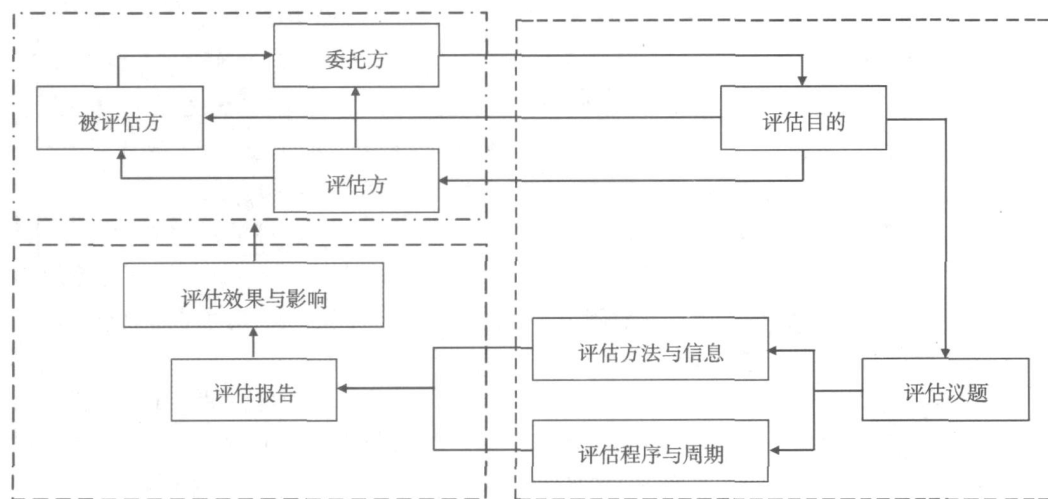


图 1 科学资助机构国际评估的分析框架

案例研究属于质的研究方法,即通过“解剖麻雀”,对案例本身的内涵进行深入挖掘,揭示以案例

为代表的总体的特征。但由于其不可重复性及研究人员可能存在的主观偏见,常常引发信度危机。多

案例研究可以较好地解决这个问题,特别是基于内容分析的案例研究,能够重复进行和相互校验,确保研究结论经得起考验<sup>[18]</sup>。遵照案例研究和内容分析的方法和程序,我们对上述案例的评估要素进行了深入分析。

## 2 结果与讨论

(1)委托方。案例研究发现,国际评估的委托方可以分为三类:FP6、RCN、FWF的委托方是所在国家或地区的科技管理部门,依照国家评估法律或科学资助机构法规组织国际评估;S-GCRP、DFG的委托方是所在国家的科技管理部门或上级管理部门,但没有明确的法律依据;JSPS的委托方是其自身,即科学资助机构自组织评估,也没有明确的法律依据。由此可见,国际评估一般是由科学资助机构的上级部门发起的,但法律依据并非关键。

(2)评估目的。科学资助机构国际评估的评估目的可以分为两大类:他组织评估,指上级主管部门为了解科学资助机构既定目标的达成情况,为下一

步资助安排和管理提供建议,包括FP6、RCN、FWF、DFG、S-GCRP;自组织评估,如JSPS,是科学资助机构自我加压,谋求组织转型与发展而开展的国际评估。显然,他组织评估居多,说明外部压力是评估的主要起因。

(3)被评估方。我们从科学资助机构的体量(以评估启动时的财政预算为指标)和时间跨度(以评估所涵盖的年限为指标)两个方面考察特征及对国际评估的影响。通常来说,体量越大意味着评估的内容范围越广,时间跨度越长通常需要更多的成本投入,二者共同影响评估的可行性和成本。

如图2所示,纵坐标表示科学资助机构的体量大小,横坐标指国际评估的时间跨度。我们对两个指标的中值做了分界线,以便进行比较和定位。分析发现,体量最小的S-GCRP的当年预算不足1亿美元;因其属多国资助,FP6的体量最高,远远超过其他科学资助机构。NSFC的体量适中,当年预算为10亿美元左右,因此其评估难度并不比其他国家高。

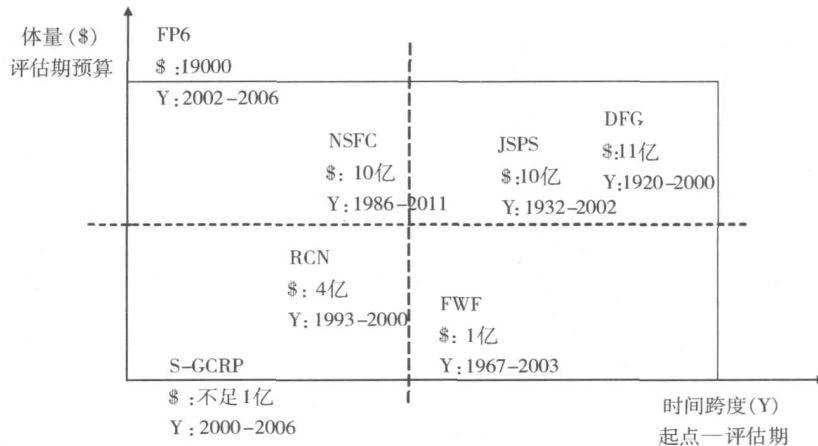


图 2 科学资助机构国际评估的体量与时间跨度

就时间跨度而言,所有案例的国际评估的时间跨度都是自其成立之日起至今。其中历史最久的是DFG,评估时间长达80年;时间跨度最短的是FP6,仅有5年。NSFC的评估时间跨度为25年,少于JSPS和FWF。尽管尚无完全对应的先例可循,但我们通过这种比较可以明确NSFC的评估定位。综合

而言NSFC的国际评估将不会远远难于上述案例,因此需要对NSFC的国际评估充满信心。

(4)评估方。利益相关性和知识拥有度是选取评估专家的两个关键因素<sup>[6]</sup>,前者指评估方与被评估方之间的利益共享或冲突程度,后者指评估方对被评估方绩效信息的掌握程度和处理能力(如图所

显然,被评估方的组织结构和资助的计划类型也是一个重要因素,后者决定了国际评估的复杂性,但由于各个科学资助机构在这方面的相似性较强,我们没有开展专门研究。

示)。评估方的理想状况是兼具低利益相关性和高知识拥有度,但现实情况却难以达到:国际评估专家与科学资助机构的利益相关性较低,但由于国别差异,他们对被评估方的了解和认识也较少;国内评估专家恰好相反,他们与科学资助机构关系密切,甚至

是其受资助者或评议人,但他们同时也对科学资助机构的运作与管理最为了解。因此,折衷的办法是综合国内与国外评估专家并构成一个评估组或评委会,取长补短,互有支撑。

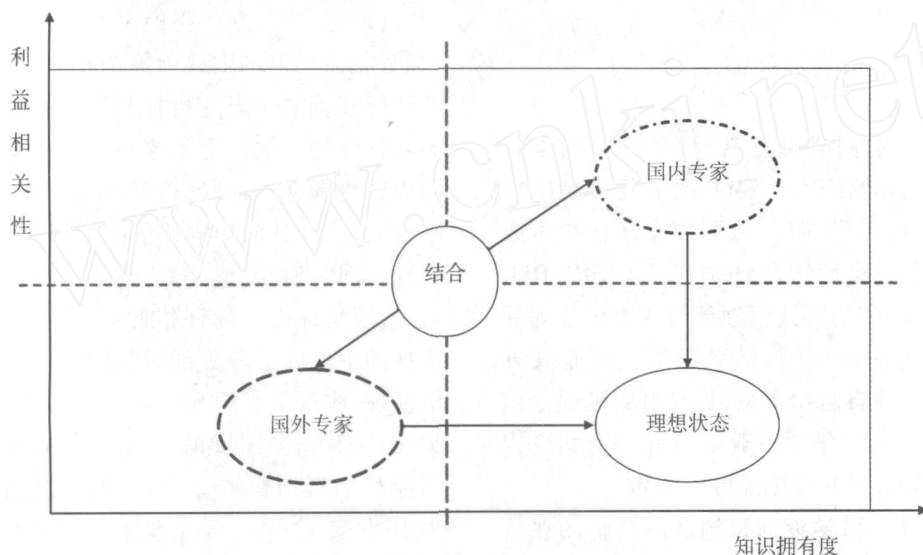


图 3 科学资助机构国际评估专家的选取

案例分析显示,科学资助机构的国际评估专家构成模式包括:完全受托评估,即科学资助机构发布国际评估招标项目,由国际专家组成的独立评估组经过竞标后完全代理评估任务,科学资助机构提供配合和支持,如 FP6、RCN、FWF 等即是如此。自评估+国际评估,指科学资助机构首先进行自我评估,然后邀请国际评估专家组对自评估报告进行判断和评审,如 JSPS、DFG 等的国际评估是这种模式。自评估+独立评估+国际评估,即科学资助机构首先开展自我评估,同时委托独立评估人员完成独立评估,最后邀请国际评估专家作出评判,如 S-GCRP 就采用这种模式。上述国际评估模式尽管有所不同,但其共同点都是采取双层或多级模式,兼顾评估的权威性、操作性和协调难度,最优化评估过程和效果。

就国际评估专家的构成来说,通常包括国外科技政策和科技评估专家、国外著名科学基金领导者和面向各个学科领域的国外著名科学家。国际评估专家的规模不等,最少的仅有 6 人(如 S-GCRP,国

内、国外各 3 人),最多的有 20 人(如 FWF,包括国外专家和 FWF 的专职人员)。其他案例的国际评估专家人数则相对集中:如 JSPS 共 7 人(国外 4 人、国内 3 人),DFG 共 10 人(国外 8 人、国内 2 人)、RCN 共 11 人(均为国际专家)、FP6 共 13 人(均为国际专家)。可见,国际评估中一般国外专家多于国内专家,规模在 10 人左右,并通常配有科学资助机构的工作人员参与其中。

(5)评估议题。科学资助机构国际评估的共同点包括:机构设立的必要性、设立目标定位及其达成情况(国际评估最为关键的评估内容)、资助方向及内容、效果与影响、问题与改进等。评估内容一般都按照逻辑模型的投入、活动与产出、结果与影响的因果链条表示,并重点采取目标实现程度的评估模式。就评估议题的特色而言,通常是由一系列的专题评估构成,多角度反映科学资助机构的运作与管理特点,具体体现为一系列支撑性报告或专题评估报告。此外,国际比较必不可少,尤其是发达国家之间科学研究水平和科学资助机构管理的标杆比对。最后,

完成该评估任务的 Technopolis 集团同时还对奥地利的另一个科学资助机构(奥地利产业促进会, FFF)进行了评估,因此其评估专家人数较多一些。



在国际评估的基础上解释评估结果的成因并寻求解决之道,如通过科学研究水平反映科学资助机构的绩效水平,根据科学资助机构的组织与管理状况解释绩效水平的高低<sup>[2]</sup>。

(6)评估方法与信息来源。国际评估的一个主要趋势即多种评估方法并举,采取三角测量(triangulation)的方式,定性评估与定量评估相匹配,多侧面、多角度地反映科学资助机构的绩效水平<sup>[9]</sup>。而且不同的评估方法对应于不同的评估议题,比如定性评估关注历史,而定量评估注重现状;呈现规模时采用定量评估,突出亮点时运用定性评估。

就国际评估的信息来源来讲,也是多渠道采集,实现多种证据的相互印证。首先是自有数据,即科学资助机构的历史资料和基础数据、所在国科技体制和法律制度信息等;其次是生成数据,即访谈、问卷调查、统计分析、基于证据的判断等国际评估专门收集的数据。

(7)评估程序与周期。就评估程序而论,国际评估通常包括以下几个环节:委托—启动—磋商—信息收集—评估—结束。评估之前各个关键利益相关者之间的沟通和磋商是评估的一个重要阶段,藉此明确评估目的、重点和要求等<sup>[8]</sup>。“信息收集—评估”阶段经常多轮反复,采取不断收集和评估的模式逼近评估结论。

国际评估的周期方面,一般评估周期为 1 年,如 JSPS、RCN、FWF、FP6;也有持续 2 年(如 S-GCRP)甚至长达 2 年以上的(DFG)。据此我们可以识别国际评估的工作量,它隐含在评估程序和周期中,需要进一步分析。

(8)评估报告。科学资助机构国际评估的评估报告尽管内容和结构不同,特色各异,但一般都包括以下要素:前言、致辞、目录、缩略语表、摘要;评估背景、范围、目的、议题与方法;国家科技管理体制与基础研究政策环境;科学资助机构的历史沿革与发展态势;科学资助机构的目标定位及其实现程度评估;科学资助机构的资助情况;科学资助机构的管理状况;科学资助机构的资助成果与影响;评估结论与建议;附录及支撑性研究报告系列。上述内容有机组合,既为国际评估专家提供了判断的证据基础,也为科学资助机构改进绩效提供了清晰的思路和方向<sup>[2]</sup>。

(9)评估效果与影响。业界普遍认可的获得积极影响的国际评估案例包括 JSPS、DFG、S-GCRP,

它们的国际评估也成为其他国家科学资助机构学习的典范<sup>[2]</sup>。因为资料不足或时间较近而难以判断其评估影响的有 FP6、FWF、RCN。尽管没有明确结论,但其评估报告却表现出对积极影响的乐观态度<sup>[8-10][21]</sup>。总之,科学、合理的设计与实施是可以保证国际评估为科学资助机构及所在国家基础研究环境带来积极效果和正面影响。如果评估效果不佳,原因何在?我们认为,评估效果不佳的原因既可能源于单个评估要素的设计缺陷,也可能是因为评估要素之间不兼容<sup>[6]</sup>。

### 3 对国家自然科学基金资助与管理绩效国际评估的启示

#### 3.1 NSFC 国际评估方案设计

根据上述案例研究得出的研究发现和结论,我们对 NSFC 国际评估的设计思路和整体框架进行了研究,初步形成了相关建议。综合而言,NSFC 国际评估的设计需要基于以下三个方面:首先,学习借鉴国内外评估实践经验,包括案例选取与资料消化;其次,吸收公共绩效评估的理论基础与方法论,如评估要素分析框架和评估方法等;最后,深入理解评估对象(也即 NSFC)的组织管理及其运行环境,一方面认识基础研究资助与管理工作,另一方面分析科学基金制及其在中国的实践。

2008 年 9 月,在国外考察和国内调研的基础上,NSFC 计划局初步形成了《科学基金资助与管理绩效国际评估工作设想》。NSFC 的国际评估方案设计应以为基础,结合 2008 年 10 月召开的主题为“科学基金资助与管理绩效评估”的“双清论坛”以及后续一系列专题研讨会等达成共识的方案内容(如评估目的与对象、组织方式、时间跨度等)为主体,论证其合理性并进一步操作化。针对尚未取得共识的内容,如评估内容与议题、绩效报告内容,则需要提供备选方案并进行利弊分析,同时做好与有关方面的沟通。

通过了解得知,NSFC 绩效评估的法定依据是 2007 年初国务院颁布并实施的《国家自然科学基金条例》(以下简称《基金条例》),其中第五章第三十一条规定:“基金管理机构应当定期对基金资助工作进行评估,公布评估报告,并将评估报告作为制定基金发展规划和年度基金项目指南的依据。科学基金制是舶来品,在 1982 年引入中国并生根发芽,



而这也是 1993 年中央办公厅调研的主题:科学基金制的适应性、中国化及其扩散<sup>[22]</sup>。因此,NSFC 的评估目的是公共问责与绩效改进,一方面回应《基金条例》的评估要求和报告需求,另一方面总结科学基金制的成绩。

就 NSFC 的评估模式与参与主体来说,其评估组织机构的设置应依据 NSFC 的组织法《基金条例》,包括国家自然科学基金委员会及其上级部门,如国务院、财政部、科技部、国家审计总署等部门。鉴于国际评估的工作量较大、专业性很强,宜采取双层评估模式,即首先进行受托独立评估(entrusted independent evaluation),由邀请的独立评估专家组完成独立评估报告或绩效报告;然后开展国际专家评审(international expert review),由国际评审专家主持完成最终的评估报告。

NSFC 国际评估专家的来源可以包括 NSFC 顾问委员会委员、NSFC 2006 年举行 20 周年专题纪念邀请的国际专家、其他国外科学基金领导者和著名科学家,以及八国集团成员国科学资助机构推荐的代表等。尤为重要的是,需要明确国际评估专家的遴选时点、介入时点、作用力度、深入程度以及信息需求等,通过深入的沟通和磋商,最大化专家的作用。

NSFC 国际评估的对象是 NSFC 资助与管理活动的整体绩效,其评估内容与关键议题也应以《基金条例》为准,即国务院自然科学基金管理机构“负责管理国家自然科学基金,监督基金资助项目的实施”。案例研究表明,十分重要的是根据 NSFC 的战略定位(支持基础研究、坚持自由探索、发挥导向作用)和“十一五”期间实施的四项战略(源头创新战略、科技人才战略、创新环境战略、卓越管理战略)为核心设置评估议题,突出机构目标的实现程度。

目前 NSFC 国际评估的初步内容基本形成,尚需要与 NSFC 的领导层、国际评估专家等磋商并细化。NSFC 国际评估的三位一体模式包括战略与目标定位、管理与组织实施、结果与影响,三者都以研究类、人才类和环境类三大项目系列为贯穿线条。评估中尤其注重点面结合,选取若干重点专题进行评估,如同行评议、资助规模与强度、科学家参与管理、人力资源管理、信息与成果管理等科学基金制发挥作用的关键环节。最后,评估中需要兼顾历史追溯(科学基金的演变历程)和国际比较(国际同类机构比较),最终找准 NSFC 的坐标经纬度。

参照案例研究,NSFC 国际评估的程序包括以下环节:成立评估组,邀请国际评估专家,召开评估启动会;实施受托独立评估,形成绩效报告及基础性材料;国际专家评审,实地考察并按要求补充相关材料;完成国际评估报告,召开评估结题会。评估周期不应多于 2 年,最好以 1 年为宜。这就需要细化评估程序,并对信息收集和专家评审的难度做好预计。

评估方法方面,需要在正式国际评估之前进行试评估,探索和比对多种方法和信息源,并形成基本统一的评估模式。国际评估的信息来源与评估内容的各个维度的匹配关系,包括 NSFC 的已有信息(文件档案、资助信息、成果案例等)和国际评估的生成信息(文档分析、访谈、问卷调查、实地考察、案例研究等)两大类。

国际评估是一项庞大的系统工程,过程中尤其需要多方面的充分沟通与磋商,逐步明晰国际评估的以下内容:评估框架与内容、评估的关键议题;绩效信息的支撑性、可及性与收集成本;评估方法选择与匹配组合;评估的总体周期及各阶段的时间分布;评估的人力需求及专业技术;评估的结果呈现方式,也即“绩效故事”的讲述;评估的潜在风险、困难和负面影响,等等。

### 3.2 NSFC 绩效报告设计

在评估模式基本确定的情况下,为国际评估专家提供的 NSFC 整体绩效报告,是总体方案设计中尤为重要内容。显然,如果绩效报告无法满足国际评估专家的信息需求的话,无异于“返工”,必将耗费大量的人力财力物力,而最终的评估效果却未必尽如人意。因此,一方面需要完成自组织评估报告或独立评估报告,另一方面则要提供大量基础性的评估材料供国际评估专家决策。基于已有案例,必须结合 NSFC 的实际情况并与国际评估专家沟通,预测其绩效信息的需求,并组织绩效报告设计工作。

综合而言,绩效报告的一般评判标准包括:效度,即绩效信息的完备性;信度,指绩效信息的准确性,不同的评估专家可以据此得出大致相同的结论;可读性,指结构和表述符合阅读要求;质量(读者满意度),即读者需求的满足程度<sup>[23]</sup>。除此以外,国际评估专家的信息需求还受以下因素影响:国际评估专家对绩效的理解(什么是绩效?)可能不同,需要沟通并明确;国际评估专家的学科背景与分析视角不同,也需要引导和支持;最后是绩效信息的可获得

性,应向专家表明哪些信息可以获得,哪些信息难以获得<sup>[24]</sup>。

在绩效报告时有必要明确设计逻辑,首先确定 NSFC 绩效本身的内涵和表现形式,如 NSF 在开展绩效评估时指出,“研究和教育工作是实现 NSF 结果目标的先决条件,而这些工作主要是在机构外完成的。因此评价 NSF 绩效的挑战在于:“NSF 的工作人员并不从事研究与教育项目,而是支持他人申请开展活动,NSF 影响而不是控制成果和结果”。其次,有必要描述 NSFC 绩效的形成机理与演进机制,全面反映 NSFC“十岁”、“二十岁”、“二十五岁”的成长历程。NSFC 资助与管理绩效评估的关键特点是其面向历史的回顾性,特别是基于改革开放 30 余年科技政策与管理体制改革的变革背景,针对 NSFC 历史演变的刻画是绩效报告的一个关键和亮点。最后,需要识别 NSFC 绩效的关键影响因素,也即系统分析内部管理的关键因素,并综合考虑外部环境的不可控因素的潜在影响,为专家评判提供参考。

绩效报告的要素与结构可以采取独立报告或系列报告两种方式,但已有案例表明绩效报告一般都需要包括以下内容:前言、致辞、目录、缩略语表、摘要;评估背景、目的、议题与方法;科学基金的历史沿革与发展态势;中国科技管理体制与基础研究政策环境;科学基金的战略与目标定位;科学基金的管理与组织实施;科学基金的资助成果与影响;科学基金的关键议题与解决途径;评估结论与建议;支撑性研究报告系列。值得注意的是,针对历史沿革的分析需要结合档案和访谈,体现时间跨度;绩效报告的主体内容需要注意报告方式、信息可及性和专家的需求,尽可能满足上述因素;围绕一些重点内容,有必要部署和提供专题评估报告和案例研究报告,深化评估内涵。

## 4 结 论

在公共预算管理、绩效评估、公共问责越来越受到重视的情况下,国际评估日益成为各国科学研究资助机构的主要手段和发展趋势<sup>[1][3]</sup>。在此背景下,NSFC 也提出了国际评估的设想,并逐步开展国外考察和国内研究。从国际评估的委托方、评估目

的、被评估方、评估方、评估议题、评估方法与信息来源、评估程序与周期、评估报告、评估效果与影响等维度,本文对包括德意志研究联合会、日本学术振兴会、欧盟第六框架计划、挪威研究理事会、奥地利科学基金和 NSFC 中德科学中心等在内的 6 个科学资助机构的国际评估案例进行了系统分析,总结了科学基金国际评估的共同特点与主要特色,为 NSFC 资助与管理绩效国际评估的方案设计与组织实施提供了借鉴与启示。在此基础上,我们还对 NSFC 国际评估的方案设计和绩效报告设计的一些要点进行了初步研究,为未来国际评估的顺利开展提供了参考。

已有研究只对单个案例进行了分析,但没有引入跨国比较的视角,难以反映国际评估的全貌<sup>[1][4][7][13][17]</sup>。本文对有代表性的若干国际评估案例进行了比较分析,有助于我们全面认识国际评估的重要方面。显然,我们无法穷尽已有科学资助机构的所有国际评估,因此未来研究需要进一步了解国内外的的发展状况,丰富和完善本文提出的分析框架和研究维度。另一方面,本文的信息来源主要是国际评估的考察报告和评估报告,采取的分析策略是倒逼和回溯,但国际评估究竟是如何开展的,其过程和隐含于中的挑战和问题、经验和教训等都难以获知,由此也降低了本文进一步分析的可能。因此,未来研究有必要通过面向国际评估专家和实际参与者的访谈和座谈会等形式,深入了解和进一步理解国际评估的设计和 implementation 过程,为 NSFC 及国内其他科学资助机构的国际评估提供借鉴。

## 参考文献:

- [1] 龚旭. 绩效评估与组织变革——日本学术振兴会的国际评估与年度评估[J]. 中国科学基金, 2009, (1): 60 - 64.
- [2] 科学基金绩效评估考察团. 科学基金国际评估考察团出访日本和德国报告[R]. 国家自然科学基金委员会:北京, 2008.
- [3] 赵学文, 龚旭. 科学研究绩效评估的理论与实践[M]. 北京:高等教育出版社, 2007.
- [4] 龚旭, 夏文莉. 美国联邦政府开展的基础研究绩效评估及其启示[J]. 科研管理, 2003, 24 (2): 1 - 8.
- [5] 郑永和, 刘云, 何鸣鸿. 科学基金绩效评估方案设计

- 中的若干问题研究 [J]. 中国基础科学, 2008, (2): 41 - 44.
- [ 6 ] 吴建南. 公共部门绩效评估:理论与实践 [J]. 中国科学基金, 2009, (3): 149 - 154.
- [ 7 ] 房强. 德国科研系统未来发展纲要——国际评审委员会评估马普学会和德意志研究联合会 [J]. 科学对社会的影响, 2001, (4): 5 - 7.
- [ 8 ] A mold E. Evaluation of the Austrian Industrial Research Promotion Fund (FFF) and the Austrian Science Fund (FWF): Synthesis Report[R]. Technopolis, 2003.
- [ 9 ] A mold E, Kuhman S, Meulen Bvd A Singular Council - Evaluation of the Research Council of Norway [R]. Technopolis, 2001.
- [ 10 ] Rietschel E T, A mold E. Reports of the Expert Groups on the Ex Post Evaluation of the Sixth Framework Programmes [R]. Commission of the European Communities, 2009.
- [ 11 ] 钱炜, 罗晖. 科学基金:试水科研投入总体绩效评估——访国家自然科学基金委主任陈宜瑜代表 [N]. 科技日报, 2009 - 03 - 09.
- [ 12 ] 龚旭, 赵学文, 李晓轩. 关于国家自然科学基金绩效评估的思考 [J]. 科研管理, 2004, (4): 1 - 6.
- [ 13 ] 韩建国, 陈乐生, 朱东华. 科学基金国际评估的实践——中德科学中心评估工作 [J]. 中国科学基金, 2009, (3): 173 - 176.
- [ 14 ] 王辉. 谈谈科学基金绩效评估中的几个问题 [J]. 中国科学基金, 2002, (1): 47 - 49.
- [ 15 ] 王任飞, 郑永和, 刘云. 关于我国科学基金绩效评估的若干思考 [J]. 中国基础科学, 2007, (6): 52 - 55.
- [ 16 ] 孟激, 程卓蕾. 科学基金机构的绩效评价与管理 [J]. 科学学与科学技术管理, 2009, (4): 9 - 14.
- [ 17 ] 孟激, 杨慧娟. 战略导向下的绩效评估——加拿大自然科学与工程研究理事会的评估制度与实践 [J]. 2008, (5): 294 - 297.
- [ 18 ] 罗伯特·殷. 案例研究:设计与方法 [M]. 重庆:重庆大学出版社, 2004.
- [ 19 ] JSPS Japan Society for the Promotion of Science External Review Report[R]. Japan Society for the Promotion of Science, 2002.
- [ 20 ] 朱东华, Hombostel S. 中德科学中心评估小组报告 [R]. 北京:国家自然科学基金委员会国际合作局, 2007.
- [ 21 ] Commission of the European Communities. On the Response to the Reports of the Expert Groups on the Ex Post Evaluation of the Sixth Framework Programmes [R]. Commission of the European Communities, 2009.
- [ 22 ] 于维栋. 科学基金制——科学研究永葆活力的催化剂 [M]. 北京:科学技术文献出版社, 1994.
- [ 23 ] Fountain J R, Patton T K, Steinberg H I. Improving performance reporting for government: new guidance and resources [J]. Journal of Government Financial Management, 2004, 53 (1): 60 - 66.
- [ 24 ] 吴建南, 马亮, 郑永和. 科学基金国际评估如何报告绩效:关于日本学术振兴会绩效报告的叙事分析 [J]. 科学学与科学技术管理, 2009, (12): 55 - 59.

## The framework, issues and methods of international evaluation of national science funding institutions: a cross-national comparative multi-case study

WU Jian - nan<sup>1</sup>, MA Liang<sup>2</sup>, ZHENG Yong - he<sup>3</sup>

(1. School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an 710049, China;

2. School of Management, Xi'an Jiaotong University, Xi'an 710049, China;

3. Bureau of Planning, National Natural Science Foundation of China, Beijing 100085, China)

**Abstract:** The principals, agents, objects, purposes, issues, methods and sources of information, procedures and duration, results and reports, as well as effectiveness and impact of international evaluation of several national science foundations, including Japan Society for the Promotion of Science (JSPS), Austrian Science Fund (FWF), Research Council of Norway (RCN), Deutsche Forschungsgemeinschaft (DFG), the Sixth Framework Programmes (FP6), and Sino - German Center for Research Promotion (SGCRP), are comprehensively analyzed, which in turn results in a summary of common attributes and characteristics of international evaluation. Such analysis supplies suggestions and directions for the design and implementation of international funding and management performance evaluation of NSFC in the future.

**Key words:** national natural science foundation; international evaluation; performance evaluation; public sector; case study; cross-national comparison