

· 科学论坛 ·

美国科学基金绩效评估中的若干关键问题

陈建俞*

(上海交通大学医学院科技发展处, 上海 200025)

[摘要] 美国科学基金绩效评估是在 GPRA 框架下, 结合基础研究的特殊性, 针对战略目标开展的评估活动。美国科学基金采用定量与定性相结合的评估方法, 外部专家分别从项目层次、计划层次、科学部层次和部门层次对战略结果目标进行“自下而上”的评估, 而管理人员利用内部数据库对管理目标进行评估, 并由外部第三方来证实。美国科学基金绩效评估目的是改进其管理工作的效率, 使有限的科研资源得到优化配置, 其核心目标是将评估的最终结果作为制定部门预算的重要依据。

[关键词] GPRA; NSF; 绩效评估; 战略目标; 外部专家评审法

党的十八大以来, 一系列有关科研项目经费管理的办法相继出台。如何开展科研绩效评估, 充分发挥绩效支出的激励作用, 已成为目前科研管理部门亟待解决的问题。自 20 世纪中期开始, 西方各国就将绩效管理理论应用到科研管理工作中。具有里程碑意义的是 1993 年美国出台的《政府绩效与结果法案》(Government Performance and Results Act, 下文简称 GPRA), 这是世界上第一个通过立法形式建立的绩效评价制度。美国国家科学基金会(National Science Foundation, 下文简称 NSF)作为联邦政府资助基础研究的主要机构, 经过十多年的探索和实践, 已形成在 GPRA 框架下较为健全的绩效评估体系, 其绩效评估的理论与实践具有很好的参考价值。本文试图对美国 NSF 科研绩效评估中的若干关键问题进行探讨, 以期为当前科研管理部门开展科研绩效评估提供一点参考。

1 基础研究是否需要 GPRA 绩效评估?

美国 GPRA 绩效评估体系框架要求所有的联邦机构都要制订一个至少包括未来 5 年工作目标的战略规划(Strategic Plan); 要求每年提供将战略规划分解成量化实施目标的年度绩效计划(Annual Performance Plan); 要求对照年度绩效计划中的定

量目标检查其完成情况, 形成年度绩效评估报告(Annual Performance Report)。

GPRA 这一绩效评估理念, 对于 NSF 这种从事资助基础研究的政府机构来说, 无疑提出了巨大的挑战。许多科学家及科研管理人员纷纷提出, 基础研究领域不适合也不需要参与 GPRA 绩效评估, 其理由主要集中在以下几个方面:

(1) 基础研究是一个长期持续的过程, 很难将其划分为年度绩效目标, 也就很难通过定量的方式, 将目标与执行结果客观地比较并确定年度计划执行情况。而且基础研究具有很强的探索性和风险性, 研究成果的产生时间和发展方向是无法预测的, 试图通过实施“规划”来促使科学突破的产生, 尤其是为科学突破产生的过程设置“年度进展的里程碑(Annual Milestones)”, 不仅是徒劳的也将是有害的。这样不仅会使科学家更多地关注短期结果而趋于保守, 而且不再锐意创新。

(2) 基础研究大多具有回溯性, 评估成果往往涉及到多年前的投入, 科学研究的成果很难评估。另外基础研究不直接产生经济效益, 研究成果所产生的影响往往是隐藏的, 有些成果已经做出, 但是它的应用领域还没有完全揭示, 那些影响深远的项目也许在近期的产出效应却不一定很高, 或者某一个领

收稿日期: 2016-01-04; 修回日期: 2016-04-18

* 通信作者, Email: chenjianyu@shsmu.edu.cn

域的新发现也可在并不相关的领域中产生新的成果。

(3) 科学家的研究成果往往不是在某一个资助机构的单独支持下完成的,而是与其他经费来源共同资助的结果。把这种多渠道资助的成果仅作为某一个资助机构的成果指标,显然是不科学的。

然而,国会和白宫坚持认为,基础研究的资助机构实施 GPRA 是必要的,不过考虑到基础研究的特殊性,允许具体实施 GPRA 法案时可以有一定的灵活性,资助基础研究机构的年度绩效计划和绩效评估报告在定量指标方面也可以适当调整^{①②}。GPRA 实际上为科学界和支持科学研究的联邦机构提供一个机会,使之能够确保美国的研究资源在满足国家需求方面的有效使用,能够向决策者和公众阐明开展科学研究的理由与结果。

2 NSF 绩效评估的战略目标是什么?

GPRA 绩效评估是“战略目标导向”的绩效评

估,NSF 根据其机构的宗旨与使命^③,制定了由上而下的战略目标。自 1997 年 9 月提交第一份 5 年战略规划以来,NSF 战略目标经历了 6 次较大的调整(表 1),战略目标的变化反映出 NSF 对其战略定位的重新认识,拓展了 NSF 绩效评估的外延和内涵。

在 1997—2003 年战略规划中,NSF 的绩效目标为“资助成果”、“投资过程”和“组织管理”;而在 2001—2006 年战略规划中,绩效目标被成果目标代替,包含“人员”、“创意”和“设施”等 3 个方面。2003 年 9 月,NSF 再次修订战略规划,将成果目标改为战略目标,以与 NSF 的使命更为吻合,并在原战略规划的基础上增加了专门针对 NSF 行政管理活动的“组织卓越”目标,以保证 NSF 是一个有能力负责责任的机构,能够支撑其他 3 个战略目标的完成。2006 年 9 月 NSF 在 2006—2011 年战略规划中将目

表 1 1997 年以来 NSF 战略规划目标的变化

1997 年 9 月	2000 年 10 月	2003 年 9 月	2006 年 9 月	2011 年 3 月	2014 年 3 月
FY1997—FY2003	FY2001—FY2006	FY2003—FY2008	FY2006—FY2011	FY2011—FY2016	FY2014—FY2018
1. 资助成果:注重所资助科学和工程领域教育与研究的长期效果;	1. 人员:由科学家、工程师、技术人员、有准备的公民构成多样化的、有竞争力的、面向全球的人力资源;	1. 人员:由科学家、工程师、技术人员和有准备的公民构成多样化的、有竞争力的、面向全球的人力资源;	1. 发现:鼓励研究推动前沿知识,着重发展具有巨大潜能和潜在利益的领域,使美国在基础研究和跨学科科学与工程方面处于领导地位;	1. 转变前沿:加强研究与教育的无缝整合,研究设施与发现的紧密耦合;	1. 转变科学与工程的前沿;
2. 投资过程:关注投资过程的绩效目标和用以实现成果目标有效资助模式;	2. 创意:与学习、创新和服务社会相关的跨学科科学技术前沿的发现;	2. 创意:与学习、创新和服务社会相关的跨学科科学技术前沿的发现;	2. 学习:培养具有世界一流水平的、有广泛基础的科学与工程人才,提高国民的科学文化素养;	2. 社会创新:指向基金项目与社会需求之间的紧密联系,突出新知识和创造力在经济繁荣和社会福利方面所起的作用;	2. 通过研究与教育鼓励创新满足社会需求;
3. 组织管理:强调管理活动的效率和效果以支持机构实现其使命。	3. 设施:共享、便利、卓越的研究和教育设施。	3. 设施:便利、共享、最先进的科学与工程设施设备,以及其他有利于发现、学习和创新的基础设施;	3. 研究设施:通过对先进仪器设备、网络设备以及实验设施的重点投入,建设国家的研究能力;	3. 争做一个模范组织:强调 NSF 在所有操作层面获得优秀和内涵的重要性。	3. 成为一个优秀的联邦科学机构。
		4. 组织卓越:一个灵巧敏捷有创新精神的机构,通过引领优秀的商务活动来完成其使命。	4. 管理:作为一个有能力负责的组织帮助科学工程领域的研究与教育做得更优秀。		

标结构修改为包括“发现”、“学习”、“研究设施”和“管理”四个方面的战略成果目标。2011年3月NSF重新构建了战略目标,在2011—2016年战略规划中战略目标改为“转变前沿”、“社会创新”和“争做一个模范组织”等3个方面,以期NSF的投资可以产生更长期的效果和更直接的影响,强调在全球范围内的领导力。2014年3月NSF在原有的战略规划基础上稍作了修改,突出了科学与工程领域。

NSF战略规划目标的历次调整表明,NSF的绩效评估越来越关注机构有没有“做正确的事情”,是否能够根据公众需求、外部环境以及机构资源条件与能力的变化不断修正或改进战略规划,从而使NSF能够在世界瞬息万变、科技日新月异的时代作为有能力、负责任的机构为美国教育和科学与工程领域的研究提供卓越的支持。

3 NSF 绩效评估的方法是什么?

NSF战略目标可分定性目标和定量目标,战略结果目标通常以定性的形式呈现,而大多数管理目标是定量的。NSF针对战略目标的不同性质,采取定量与定性相结合的评估方法,从不同层次、不同角度对NSF的绩效进行评估,并提交年度绩效报告(图1)。

战略结果目标的评估,是NSF根据GPRA的要求对教育和基础研究成果目标进行评估,主要采用外部专家评审法。NSF采用外部专家评审法已经有20多年,它是美国联邦机构中唯一邀请外部专家来评审的机构。专家评审的价值可从美国科学、工程和公共政策委员会(Committee on Science, Engineering and Public Policy)2001年的报告中得到认证。“因为我们不知道怎样估量正在产生的知识,当它的实际运用不能预测时,我们能做的最好是定期邀请这个领域的专家来评估研究进展情况,此过程称之为专家评审。这些专家,辅以定量方法,可以确定产生的知识是否是高质量的,是否直接导致资助机构使命的潜在重要性的主题,以及现有知识是否处于前沿,因此有可能加强对该领域的理解^⑧”。

对于管理目标的评估,NSF利用内部数据库系统由管理人员对绩效结果进行评估,同时通过外部第三方来审查所选择的目标,所选的结果仍由第三方来证实和验证。

2002年美国管理及预算办公室(Office of Man-

agement and Budget,简称OMB)提出了计划评估等级工具(Program Assessment Rating Tool,简称PART),作为评估联邦政府之间计划活动绩效的一种方法。PART通过计划评估的方式,起到对GPRA完善和补充的作用。PART评估主要关注“计划的目的和设计”、“战略规划”、“计划管理”和“计划结果”4个部分。每年NSF计划中大约有20%的计划参加PART评估。直至2008年,NSF所有计划都参加了PART评估。在联邦机构参加PART评估的1000多项计划中,18%的计划获得了最高等级“有效(Effective)”。而NSF所有参加PART评估的计划都获得了最高等级“有效(Effective)”。

2002年美国人事管理办公室(Office Of Personnel Management,简称OPM)开始对政府机构开展联邦人力资本调查(Federal Human Capital Survey,简称FHCS),其目的是为了测量员工对机构目前是否具备成功组织特征以及具备成功组织特征的程度的认知。此后联邦人力资本调查(FHCS)每2年举行一次。联邦人力资本调查(FHCS)主要围绕五项绩效管理指标,按照总统管理议程(President's Management Agenda)制定的三色等级评估体系来进行的。这五项绩效管理指标分别为“人力资本(Human Capital)”、“竞争力(Competitive Sourcing)”、“财务绩效(Financial Performance)”、“电子政务(E-Government)”和“绩效改进(Performance Improvement)”。2010年联邦人力资本调查(FHCS)被更名为联邦雇员观点调查(Federal Employee Viewpoint Survey,简称FEVS),原来每2年举行一次改为每年举行一次。联邦雇员观点调查(FEVS)主要围绕“员工敬业度(Employee Engagement Index)”、“总体工作满意度(Global Satisfaction Index)”和“工作环境包容性(The New Inclusion Quotient Index)”三大指标开展的。联邦雇员观点调查(FEVS)可为联邦政府如何运作其人力资本管理体系提供通用指标;可为OPM评估各政府机构及其人力资本战略管理的进展情况提供测量工具;也为高级管理人员回答“我能做什么使机构运行得更好”提供重要的情报。从2010—2015年联邦雇员观点调查(FEVS)的调查结果比较来看,NSF的得分逐年稳定上升,这说明基金委正朝着正确方向发展,已成功地完成了其使命。

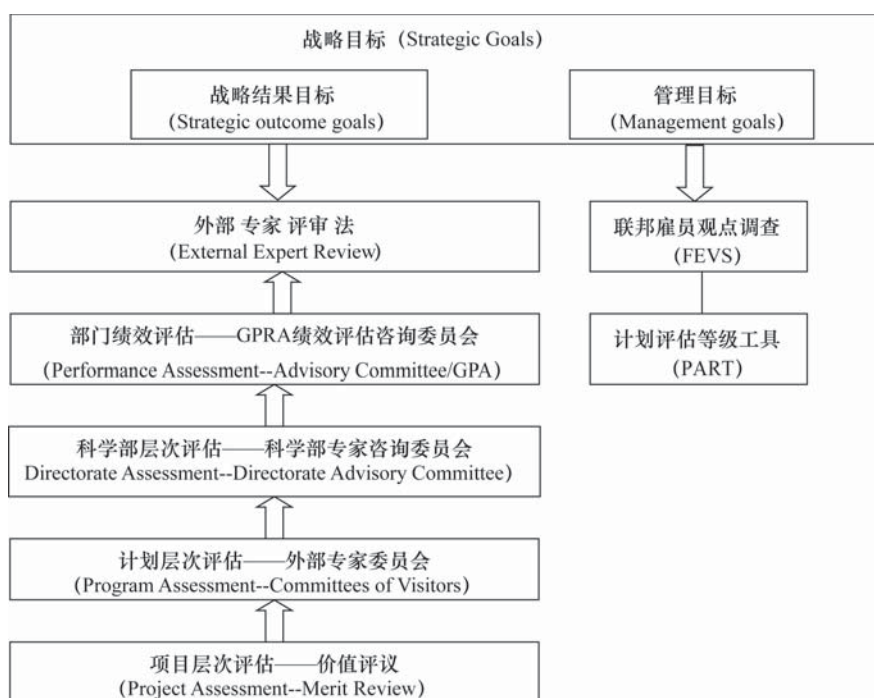


图1 NSF 绩效评估框架图(Performance Assessment Framework)

4 NSF 绩效评估的过程是怎样的？

NSF 的绩效评估过程是一个“自下而上”的评估过程,较微观的评估总是较宏观层次的评估的基础。根据不同层次,绩效评估对象可分为4类:项目层次评估、计划层次评估、科学部层次评估和部门层次评估(图1)。

项目层次评估是以两种不同的方式,主要针对具体项目的学术价值(Intellectual Merit)和广泛影响(Broader Impacts)进行评估。首先,项目申请者在提交申请书时必须提供前一个基金项目的结题信息。外部专家可获得这些信息,根据价值评议标准对申请书进行评审。计划主管在对项目提出给与资助或不予资助的意见时也会考虑这些信息。其次,项目负责人在基金资助期间要求每年提交进展报告。每年资金下拨前需要这些信息,以便决定是否继续资助。

计划层次评估是由外部专家委员会(COV)对项目价值评议过程的完整性和有效性进行评估。1999年随着GPRA的全面实施,NSF要求COV增加回顾GPRA评估环节(产出和成果)的职责。每一个外部专家委员会一般由5—20名外部专家组成,他们要在2—3天时间内评估1个或多个计划。为了保证评估的公平性和公正性,这些专家一般是来自学术界、工业、政府和社会公众多领域的代表。

所有外部专家委员会(COV)需要完成评估报告,说明计划对NSF战略目标的贡献程度。

科学部层次评估是由科学部专家咨询委员会(AC)对科学部的资助动议及预算进行评估,并对计划层次的评估报告和NSF资助计划的相应回应做出评估。每一个科学部都有一个专家咨询委员会,每年召开两次会议。科学部专家咨询委员会(AC)一般由18—25名在学术界、工业和政府有丰富经验的外部专家组成。每个专家咨询委员会都要遵从联邦咨询委员会法案(Federal Advisory Committee Act)的原则。

部门绩效评估是由GPRA绩效评估咨询委员会(AC/GPA)按照NSF战略规划中的战略结果目标对NSF整体绩效进行评估。绩效评估咨询委员会(AC/GPA)成立于2002年6月,成员大约有20人左右。这些人员来自学术界、产业界和政府机构,他们当中有企业的技术总监,首席信息官,军队的机械工程师,研究所的科研人员,大学的系主任,教务长,计算机与信息学,哲学与宗教学,环境学,人类学,天文学,植物学与微生物学领域的教授,副教授,助理教授等。GPRA绩效评估咨询委员会是基于委员会集体的经验和专业知识,对NSF是否作出重大成就进行评估。每年GPRA绩效评估咨询委员会针对每一个战略目标进行绩效评估的结果,形成报告后提交给基金委主任。

NSF 的年度 GPRA 报告是在 COV 报告和 AC/GPA 报告的内容分析基础上形成的,这是国会进行预算资源分配的重要依据。GPRA 报告主要由以下 3 个部分构成:(1) 战略规划(Strategic Plan); (2) 年度绩效计划(Annual Performance Plan); 自 2004 年起包含在提交给国会的年度预算报告中; (3) 年度绩效报告(Annual Performance Report)。自 2002 年起 NSF 将“GPRA 绩效报告(GPRA Performance Report)”和“责任报告(Accountability Report)”合并为“绩效与责任报告(Performance and Accountability Report)”,这是对《2000 年报告整合法案(Reports Consolidation Act of 2000)》的回应。自 2008 年起 NSF 又将名称改为年度绩效报告(Annual Performance Report),并连同机构财务报告(Agency Financial Report)、绩效财务亮点报告(Performance and Financial Highlights Report)一起提供财务管理和项目绩效信息。

作为一个联邦机构,NSF 在 GPRA 的框架下,制定战略规划,设立年度战略目标,每年针对战略目标的完成进展情况进行评估,并将评估结果向总统、国会和公众报告,以接受广泛的社会监督。当前我国科研绩效评估工作正处于白热化阶段,然而各级科研管理部门对科研绩效概念的理解却与西方发达国家有很大的不同,仍关注于各类产出指标的定量测评,此科研绩效非彼科研绩效。我国科研绩效评估工作任重而道远,首先理念上的改变尤为重要。

参 考 文 献

- [1] National science and Technology Council Committee on Fundamental Science. Assessing fundamental science. <http://www.nsf.gov/sbe/srs/ostp/assess>.
- [2] Committee on Science, Engineering, and Public Policy, National Academy of Sciences/National Academy of Engineering/Institute of Medicine. Evaluating Federal Research Programs: Government Performance and Results Act. Washington DC: National Academy Press, 1999.
- [3] The NSF statutory mission is “To promote the progress of science; to advance the national health, prosperity, and welfare; and to secure the national defense” —From the National Science Foundation Act of 1950 (P. L. 81-507).
- [4] National Science Foundation. National Science Foundation’s GPRA Strategic plan for fiscal years 1997—2003. 1997.
- [5] National Science Foundation. NSF GPRA Strategic Plan FY 2001—2006. 2000.
- [6] National Science Foundation. National Science Foundation Strategic plan for fiscal years FY 2003—2008. 2003.
- [7] National Science Foundation. Investing in America’s Future: NSF Strategic Plan FY 2006—2011. 2006.
- [8] National Science Foundation. Empowering the Nation Through Discovery and Innovation: NSF Strategic Plan FY 2011—2016. 2011.
- [9] National Science Foundation. Investing in Science, Engineering, and Education for the Nation’s Future NSF Strategic Plan for 2014—2018. 2014.
- [10] National Academy of Sciences, Committee on Science, Engineering, and Public Policy. Implementing the Government Performance and Results Act for Research: A Status Report, Washington, D. C., National Academy Press, 2001.
- [11] FY2016 Budget Request to Congress [EB/OL]. Available from: <http://www.nsf.gov/about/budget/fy2016/toc.jsp>.
- [12] 刘莹, 张大群, 李晓轩. 美国联邦科研机构的绩效评估制度及其启示. 中国科技论坛, 2007, (9): 140—144.
- [13] 阎波, 吴建南, 马亮. 科学基金绩效报告与绩效问责——美国 NSF 的叙事分析. 科学学研究, 2010, (11): 619—1628.
- [14] 李强, 李晓轩. 美国国家科学基金会的绩效管理 with 评估实践. 中国科技论坛, 2007, (6): 133—139.
- [15] 王文平, 刘云, 陈奎. 美国国家科学基金会绩效评估体系分析及启示. 第三届中国管理学年会——信息管理分会场论文集, 2008.
- [16] 张强, 韩莹莹. 当代美国联邦政府绩效评估的层级体系分析. 社会科学研究, 2006, (1): 94—99.
- [17] 王任飞, 郑永和, 刘云, 王文平. 美国国家科学基金的绩效评估体系. 中国财政, 2008, (10): 70—71.
- [18] 沈新尹. 关于对美国国家科学基金会基础研究绩效评价若干方法的思考. 中国科学基金, 2001, 15(5): 313—316.
- [19] 华薇娜, 闫苗苗, 刘芳. 美国国家科学基金会的绩效评估工具的方法和价值——以 2007 年评估结果为例. 甘肃社会科学, 2011, (4): 48—52.
- [20] 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. 中国基础科学, 2008, (2): 41—44.
- [21] 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. 科研管理, 2003, (2): 2—8.
- [22] 张强, 韩莹莹. 当代美国联邦政府绩效评估的层级体系分析. 社会科学研究, 2006(1): 94—99.
- [23] 何文盛, 曹洁, 张志栋. 美国政府绩效评价中项目评估分级工具——背景、内容与借鉴. 兰州大学学报(社会科学版), 2009(1): 92—96.
- [24] Cozzens, Susan E. Assessing federally-supported academic research in the United States. Research Evaluation, 2000, 12(1): 5—10.
- [25] Cozzens, Susan E. The knowledge pool: measurement challenges in evaluating fundamental research programs. Evaluation and program planning 20. 1 (1997): 77—89.
- [26] NSF Budget Request to Congress. <http://www.nsf.gov/about/performance/>.

Key issues concerning NSF performance assessment

Chen Jianyu

(Department of Scientific Technology, Shanghai Jiao Tong University School of Medicine, Shanghai 200025)

Abstract NSF performance assessment is ongoing under the GPRA framework, combined with the particularity of basic research. NSF employs a mix of both qualitative and quantitative methods to assess and evaluate strategic goals. Formalized examination takes place during merit review of project, COV and AC/GPA assessments, and GPRA reporting. Additionally, NSF administrative staff and managers make use of internal data system to assess and review progress in achieving the quantitative management goals, with selected results verified by external third parties. Its purpose is to improve the management efficiency and make the limited research resource optimize configuration. Its core goal is to make the assessment results as the important basis for planning department budget.

Key words GPRA; NSF; performance assessment; strategic goals, expert review

· 资料信息 ·

我国化工学者在乙炔乙烯分离方面取得重大突破

2016年5月19日, *Science* 杂志以 First Release 的方式在线发表了浙江大学邢华斌教授与利莫瑞克大学、德克萨斯大学圣安东尼奥分校等单位的合作研究成果“Pore chemistry and size control in hybrid porous materials for acetylene capture from ethylene”。该文提出了离子杂化多孔材料分离乙炔和乙烯的新方法,这对于气体分离技术的发展具有重要的科学意义和应用价值。(论文链接: <http://science.sciencemag.org/content/early/2016/05/18/science.aaf2458?rss=1>) 该项研究得到了国家自然科学基金优秀青年基金项目、重点项目和面上项目(项目批准号:21222601、21436010 和 21476192)的资助。

气体吸附分离过程中普遍存在分离的选择性和容量难以兼具的现象(trade-off 效应),导致设备投资和能耗居高不下。针对此问题,研究者在文章中首次提出了离子杂化多孔材料吸附分离乙炔和乙烯

的新方法。一方面,通过无机阴离子的强氢键作用实现乙炔分子的高度亲和识别,获得目前最高的乙炔乙烯分离选择性。另一方面,通过调控阴离子的几何分布和孔径大小,实现气体分子-气体分子或气体分子-多孔材料间的协同相互作用,获得迄今为止所报道的最大的吸附容量,从而解决传统气体吸附过程分离选择性和容量难以兼具的巨大挑战。在极低乙炔分压(0.025 bar)时吸附容量达 2.1 mmol/g,乙炔/乙烯(1/99, v/v)的 IAST 选择性达 39.7-44.8。该多孔材料具有很好的扩散传递性能,另外该吸附剂兼具高选择性和吸附容量的现象。

该研究成果不仅为乙烯和乙炔的高效分离与节能降耗提供了解决方法,而且也为其其它吸附分离材料的设计提供了新的途径。

(供稿:化学科学部 孙宏伟 张国俊)