

英国科研评价模式:由专业评价人员进行评价

British Model of Research Evaluation: Evaluating by Professionals

刘作仪

(国家自然科学基金委员会政策局, 博士 北京 100085)

近几年来, 科研评价成为了英国学术界和政策制定者的一个热点领域。由于英国的科研评价工作明显有别于欧共体盛行的以小组形式开展评价^[1], 因此, 学者们称其为“英国模式”。

所谓“英国模式”, 指的是由专业评价人员实施评价工作, 即在评价政府科研资助政策以及研究与发展的实施情况、产生的成果和影响时, 评价发起者只参与评价方案设计工作, 具体评价业务, 包括综合评价结论和建议等等, 一般完全委托给专业评价人员。专业评价人员通常具有很深的资历、丰富的评价经验、一丝不苟的学术态度, 他们提供的结论通常能获得行政部门和学术界的认可, 在政府制定政策过程中发挥了重要作用。

英国的专业评价人员指的是: (1) 政府部门内部的评价小组; (2) 科学政策研究机构, 例如 PREST 和 SPRU 中的有关人员^[2], 它们除开展评价理论与方法论研究之外, 还从事实际评价工作; (3) 一些专门为公共部门服务的私人咨询公司。需要指出的是, 参与评价的私人咨询公司的数量非常有限, 其原因在于科研评价工作不仅是一项知识性很强的活动, 需要各方面的知识投入, 而且需要保持绝对中立, 深受评价发起者和被评价方的信赖。这两个因素阻碍了私人咨询机构大量进入评价领域。

一、英国开展评价工作的政策背景

英国政府重视科研评价工作的主要原因在于: (1) 政府削减了研究投入; (2) 财政部 (Treasury) 和内阁办公室 (Cabinet Office) 要求所有公共管理领域证明公共资金的价值和交账 (accountability); (3) 政府科研资助部门必须对前景光明的新兴科技领域作出迅速反应; (4) 政府比以往更强调科学的实用性、技术转移和交叉学科研究, 并为此尝试利用新模式组织科学研究, 诸如资助机构联合启动研究计划、支持产学研合作研究和参与国际合作研究计划, 因此需要把评价纳入政策制定过程, 通过评价来阐明复杂、疑难问题, 例如不同资助模式的恰当程度、资助

产生的经济效益。面对上述 4 个方面的压力, 许多政府科研管理部门和机构建立了内部评估小组, 小组一般由 2~10 名全时的工作人员组成, 本文将他们称之为内部专业评价人员。小组在一位负有政策制定职责的高级公务员领导下开展工作, 这样做有利于把科研评价与政策制定结合起来。

尽管政府 R&D 管理部门面临相同的压力, 但是它们并没有采用统一的评价框架, 而是根据各自的需求和特征组织自己的评价工作^[3]。在这种情况下, 1986 年内阁办公室成立了组织协调各部门评价工作的评估办公室^[4]。评估办公室既不从事实际评价工作, 也不规定应该如何开展评价工作, 仅仅制定了一份用户友好的评价指南, 用于促进内部评价专家之间的思想交流, 传播政府评价工作理念^[5]。

评估办公室认为科研评价工作应该为科研管理与决策提供信息, 是政策制定过程中不可缺少的组成部分。这就要求内部评价小组不仅要组织评价工作, 还要研究、构建科研绩效、影响和投入等定量指标。例如, 研究理事会的内部评价小组特别重视参与研究绩效指标的拟订和内部管理信息的采集与利用, 将它们应用于研究计划的制定、预算分配、研究管理以及事后评价工作。再如, 贸工部 (DTI) 的评价办公室不仅仅采集和综合定量指标, 而且更强调主动参与 R&D 计划的制定及其详细评价。

二、内部评价小组在科研评价与管理中的作用

1. 在应用研究的评价与管理中的作用

1980 年, DTI 在其研究与技术政策司成立了研究与技术评估小组, 主要从事应用研究项目的评价。1984 年, 评估活动延伸到科研政策制定和实施的全过程, 即从制定研究计划到监控、评价研究计划实施情况与产生的影响。目前, 由于 DTI 不再支持与市场联系紧密的活动, 而是直接支持竞争前研究、支持大学与工业界的合作和技术转移以及国际合作计划, 使 DTI 必须改进其研究资助计划与管理

系统,评估新的资助模式产生的影响,为相关问题提供政策建议。为此,DTI制定了 ROAME 文件,以它为框架开展管理与评价工作^[6]。ROAME 由计划中各个要素的首字母构成:基本理由(Rationale)、目标(Objectives)、鉴定(Appraisal)、监控(Monitoring)和评价(Evaluation)。它是提交研究计划申请和评价受资助计划的标准格式,既有利于申请人从上述5个方面考虑如何提出申请计划,亦有利于评价人员对照这5个方面的内容评价计划实施情况。

(1) 理由部分必须阐明需要政府资助的原因,其中主要是经济方面的原因,即如果没有政府介入工业界将不会开展此项研究。

(2) 目标包括投入目标和技术目标,它们最好可以测度,同时还要列出难以量化和测度的目标,诸如计划可望产生的社会效益与经济效益。这样做有利于在计划执行期间或计划完成以后参照目标衡量计划执行的成功程度。

(3) 鉴定部分要详细说明单个项目组成计划的方式,列出计划涵盖的技术领域、遴选项目的标准和机制。

(4) 监控指检查受资助计划的进展所采用的系统和程序。必须追踪单个项目和整个计划的管理情况,如果出现了问题,或者计划的目标不恰当,都可以迅速作出调整。

(5) 评价部分不必列出详细评价计划,只需要针对相关政策信息的需求,拟订出开展评价工作的框架,并说明通过评价可以明确解决的问题。

从 DTI 评价工作的实践来看,ROAME 至少发挥了以下几方面的作用:(1) 为 DTI 的评估专家提供了一种评价思想;(2) 通过监控和事后评价活动,把有关信息反馈给管理人员,可以提高他们识别、理解问题的水平,从而促进了“不断学习和提高”这样一种文化氛围的形成;(3) 按 ROAME 格式书写的记录,是评估小组获取管理信息资源的一种重要渠道,这些信息既可以用于 DTI 制定科研资助政策,也可以用于设计评价方案;(4) 对事后评价工作而言,可以对照当初实施计划的理由和计划拟定的目标来考察计划实施情况。

2. 在基础研究评价与管理中的作用

农业与食品研究理事会(AFRC)、自然环境研究理事会(NERC)率先在内部成立了评价小组,由负有政策制定职责的官员监督小组的评价工作,还指令一名高级官员负责把评价与资助机构的其他活动整合起来。这种线性管理结构有助于确保评价获得理事会高层的认可。

为了实时监控和评价资助项目,内部评价小组制定了绩效指标,建立了相关数据库。例如,AFRC 于 1972 年率先建立了项目管理数据库。目前,该数据库除了收录经费与人员投入、研究计划的构成等

数据项之外,还包括定性和定量研究绩效指标,其中定性指标来源于项目立项时对项目和人员的评价,定量指标是出版物和引用情况。评价小组要定期向 AFRC 的有关管理部门通告这些数据,使它们能充分利用这些信息支持同行专家对 AFRC 的资助工作及所属研究机构进行评价。同样,NERC 的内部评估小组要提供财务和管理信息,开展一系列文献计量研究,参照当初设定的目标分析科学家和研究机构的产出。分析研究结果也用于支持同行专家的评价工作。

科学与工程研究理事会(SERC)的做法与上述两个研究理事会不同:没有自己的研究投入和绩效指标数据库;重视研究计划的评价,而不是项目评价;倾向于雇佣外部专家深入分析定性问题,而不是单纯依靠内部小组支持的同行评议^[7]。但是,目前 SERC 正同教育与科学部以及其它研究理事会合作建立一套测度产出的方法和指标,使之适用于整个科学(不包括国防科学)的评价。此外,SERC 还肩负培养科学家的重任,因而需要记录博士学位论文和博士后论文这类数据,用于评价人才培养活动的实施情况。

由于内部评估小组不仅积极主动地在理事会内部,而且还向广大共同体交流自己的工作,因此,它们提供的定量指标获得了广泛的认同。通过把特定评价业务、评价支持研究和对理事会工作人员进行定量评价方法的培训工作委托给从事科研评价理论研究的学术组织,评估小组与评价理论研究专家形成了密切的相互作用。这样,政策制定部门可以迅速了解、掌握评价方面的专门知识,在实际应用时按照每个组织的需求状况进一步发展和完善这些知识。各个研究理事会还精选了部分获资助的研究人员、同行评议专家和视察小组(visiting group)成员参与开发定量指标,例如向他们咨询衡量成功的标准和指标,确保任何新颖的方法都可以成为强有力的评价证据。尽管如此,内部评估小组提供的研究投入和绩效指标信息是在研究理事会内部发掘的,它们只可能适应于某类研究,因此,需要把它们整合到利用同行评议方式进行的评价活动中。

总之,在英国基础研究资助机构中,内部专家对评价工作和政策制定工作起支持作用,它们是英国专业科研评价队伍中不可缺少的组成部分。尽管同行评议仍是评估基础研究最令人满意的方式,不可能被内部专家评价方式替代,但可以肯定的是,评价小组能提供极有价值的政策信息,能为同行评议过程提供许多证据。

三、利用外部专业评价人员的方式

1. 实时评价大型 R & D 计划

英国科技管理部门认为,外部评价人员的评价比内部评价人员的评价更深入,更具有可靠性和公正性,因此,大型 R&D 计划的评价一般都委托给外部评价人员。例如,英国国家信息技术计划^[8](Alvey 计划)的实时评价就是由外部评价人员完成的。

在开始实施 Alvey 计划时,DTI 决定在计划运行过程中要开展实时评价,通过把评价结论立即反馈到计划运行过程中,提前发现问题并作出纠正。Alvey 计划的实时评价包括两个部分:由 SPRU 评价计划的适宜性和影响,由 PREST 评价计划的结构和组织。这两个机构的科技政策理论研究人员要把评价过程中开发的专业评价知识传授给 DTI 的评估小组和参与评价工作的 DTI 经济学家,使他们掌握最新评价专业知识和经验。

2. 制定评价手册

利用外部专业评价人员制定评价手册是一种一次性使用外部专家的方式,它不同于 DTI 和一些研究理事会采取的与外部评价专家之间需要长期维持联系的方式。

英国健康和执行部门(HSE)就采用这种评价方式。通过委托评价理论研究人员起草详细的评价手册,指导内部评价人员对研究活动进行事后评价^[9]。评价理论研究人员在起草评价手册时,不仅要充分反映 HSE 的研究人员、研究计划制定人员以及研究成果用户(包括 HSE 的其他部门、外部研究人员和工业界)的意见,而且还要提供评价案例,供内部评价人员参考。由于在起草评价手册时调查了受评价活动影响者的意见,HSE 的评价工作也获得了很高的可信度。

3. 外部专业评价人员和同行小组相结合的评价模式

经济和社会科学研究理事会(ESRC)采用外部专业评价人员和同行小组相结合的评价模式,对它管理的研究计划、研究中心和培训计划做非常详细的定性评价。其中,专业评价人员发挥两方面的作用:采集研究绩效方面的定量数据,指导评价。例如,ESRC 在评价地方研究实验室时,专业评价人员要与实验室的管理者共同商定评价的范围,然后向同行评议小组提供详细的评价证据。这些证据不仅涉及研究质量,还包含管理问题、相关性和影响。评价座谈会、同行小组可根据这些证据作出判断。从表面上看,这种评价模式和欧共体的小组评价模式之间没有明显差别,但是,这种评价模式中的评价专家在确定评价内容和评价实施方面比欧共体评价模式中的评价人员有更多的控制权。

四、结论

1. 充分发挥内部评估组织和外部评价专家的作用

使两者相互密切配合,对科研评价工作非常重要

内部评价专家的产生和发展来源于政府行政管理工作的需要,他们一般有制定研究计划、管理研究、为同行评议提供支持以及委托和实施评价的职责。由于他们熟知自己部门的使命、了解评价的需求和目标,因此在设计评价方案的过程中,他们的意见起了决定性作用。但是,评价是一项专业性和学术性极强的工作,不仅需要丰富的调查、分析和解决问题的能力,还要通晓管理学和经济学的有关专门知识、熟练掌握评价方法,而内部评价人员不可能花大量时间充分掌握这些知识。可见,如果没有外部评价专家为内部评价专家提供专业培训和咨询服务,政府机构要建立评价职能是非常困难的。

2. 以下几个因素可能影响评价的可接受程度

(1) 为评价活动提供有力的组织保障。资助机构要有高级公务员负责评价活动,只有这样评价活动才能受到高层的认可,且不受任何要挟和胁迫。

(2) 允许不同资助机构对自己的评价活动进行安排。不同资助机构都有自己特殊的评价需求,评价体系要反映这些需求。不能草率地对各个机构的评价工作强加一些硬性要求,否则,评价体系将缺乏可操作性。

(3) 把评价作为政府 R&D 资助与管理全过程的一部分。只有这样,才能重视评价工作并对其做周密部署,使政策制定者从评价中获得最大利益。

(4) 外部专家在学界的声望深刻地影响评价公正性和评价结论的可信度。评估专家可能被指责为得出了在政治上有利的结论,以获取更多政府委托评价任务;小组评价易受专家地位的影响,小组成员总是想要保护自己所属专业领域的声望;而声望卓著的专家易于被信任。

(5) 在评价时要广泛咨询广大科学共同体。只有充分调查被评价团体及其同行、科研管理人员乃至研究成果用户的意见,并且适时利用评价专业技能,科研人员才可能接受评价专业人员的结论。

3. 评价方法的日益复杂化要求评价人员具备一定水准的专业知识。内部评价人员和外部评价人员相结合组成评价共同体,既有利于内部评价人员迅速掌握最新的评价专业知识,又有利于外部评价人员发现政府机构的需求和评价中存在的问题,从而开展研究、推动评价理论的进步。

4. 尽管评价工作的伦理准则至今还未见有机构明确提出,但是要加以深入研究,以确保和维护评价工作应有的独立性和责任感。

参考文献与注释

[1] 在欧共体的小组评价方式中,小组成员一般由著名科学家和工业界的专家组成,成员的选择依赖于评价背

以信息技术手段促进管理体制创新

To Promote Innovation of Management System by Means of Information Technology

刘燕鹏¹ 方克定² 齐文虎¹ 姚予龙¹
(中科院地理科学与资源研究所¹ 北京 100080;国家行政学院³)

本文是“关于深化资源环境管理体制研究”项目的分报告,本项目由中央机构编制委员会办公室发布。课题组先后对国土资源部、农业部、国家发展计划委员会、水利部、国家海洋局、国家测绘局、国家地震局、国家林业局、国家环保总局等委部局进行了调研,并于2001年初专门赴深圳对深圳规划国土局、运输局、环保局等部门进行了调研。1991年我们就开展了同样名称的项目,对国务院农口部门的资源环境管理信息系统进行了调研。综合两次调研的成果,提出以下基本结论:管理信息系统做为一种有效的技术支撑手段,在一定程度上可以促进管理体制创新、理顺管理机制,利用管理信息技术深化政府机构改革是可行的;如果下届政府机构不做大的调整,应在以消化1998年机构改革成果为主的前提下,大力推进政府信息化,藉以缓解现存老大难问题、减少机构改革的成本、避免事后机构反弹。建议加强信息管理工作,以“登记、上网、专管、共用”为基本方针。

一、国务院农口部门资源环境管理信息系统的基本情况

景和目标。在整个评价工作过程中,科研管理部门的内部专业评价人员要为小组提供有关背景信息、事实和证据,他们仅仅起辅助开展评价工作的作用,并不指导评价工作,也不影响评价的结论。

[2] PREST: 英国 Manchester 大学的工程、科学与技术政策研究所; SPRU: 英国 Sussex 大学的科学政策研究中心。它们都是世界知名的科技政策研究机构。

[3] CABINET OFFICE,《Annual review of Government Funded Research and Development》, HMSO, London, 1990

[4] 目前, 由于各个部门充分认识到了评价工作的重要性, 不再需要一个专门机构推动此项工作, 因而取消了该办公室。

[5] CABINET OFFICE, 《Research and Development Assessment: a Guide for Customers and Managers of R & D》,

各单位管理信息系统建设情况一览表		
单位名称	上级主管	信息系统名称
国家信息中心	国家发展计划委员会	国家经济信息系统
国土资源部信息中心	国土资源部	数字国土工程
水利部信息中心	水利部	全国实时水情计算机广域网系统
农业部信息中心	农业部	农牧渔业管理信息系统
国家林业局信息中心	国家林业局	森林资源清查
国家海洋信息中心	国家海洋局	国家海洋信息系统
国家基础地理信息中心	国家测绘局	国家基础地理信息系统
国家地震局信息中心	国家地震局	全国地震通信网络的建设与完善、地震应急响应信息系统工程
国家环保局信息中心	国家环保局	尚未建立系统

根据1998年的三定方案,各委部局均成立了信息中心,其基本职能规定为面向管理、服务机关、服务社会,基本达到了所规定的职能。各单位均成立了信息化领导小组。与1991年时的情况相比,信息中心的建设发生了巨大的变化,与同期信息技术发展水平保持了一定程度的均衡。根据不同行业特点,信息中心建设各有特色。如国家测绘局建成了“全国基础地理信息中心”,负责国家基础地理信息系统的建设、管理和数据分发服务等;国家海洋信息中心从1993年开始建设和开发国家经济信息系统的分系统——国家海洋信息系统(其它见上表)。

HMSO, London, 1989

[6] L. Georgiou & E. Davis (eds), ‘Evaluation of R & D - A Policymaker’s Perspective’, Proceedings of the International Workshop on Assessment and Evaluation, HMSO, London, 1988

[7] SERC 肩负与其他部门共同资助英国的战略研究、联合资助学合作计划、资助英国参与国际科学合作。很明显,评价这些资助工作非常复杂,不能单纯依靠同行评议方法。

[8] 该计划由3个政府机构共同资助(DTI、国防部和SERC),涉及公司之间在竞争前的合作和公司与合作。

[9] PREST, ‘Protocol for the Evaluation of HSE Research’, Manchester, 1988 (责任编辑 王宏章)