

马普学会研究所评价对我国研究所评价工作的启示

〔关键词〕 科技评价 研究所评价 〔中图分类号〕 G3

廖方宇¹ 邓心安²

(中国科学技术大学管理学院¹ 北京 100080; 中国农业大学科技管理系² 北京 100094)

科技评价是把科技活动作为测定对象,对科技活动过程及其效果进行判定的行为过程。它起源于20世纪初的美国,随后世界许多国家相继开展了不同层次、不同类型的科技评价活动。我国最早开展的比较正式的科技评价是1984年国家重点实验室评估,90年代中期以来先后开展的有关科技机构评价有:国家工程技术研究中心验收评估、中国科学院研究机构年度绩效评价等。中国科学院在进入知识创新工程试点以后,开展了以研究所评价为核心的试点单位评价,标志我国国家研究所评价进入了一个新的阶段。

国家研究所评价是科技机构评价的主要内容和类型之一。国家研究所尤其是国家基础类研究所在促进国家目标(提升工业竞争力、健康医疗、环境保护、国家安全、促进就业等)的实现方面存在许多共性、特点,因而具有较大可比性和相互借鉴意义。本文在系统考察德国马普学会(The Max Planck Society)研究所评价的基础上,通过对中国科学院(以下简称“中科院”)基础类研究所评价活动的分析,就中国基础类研究所评价方法及存在问题进行了讨论。

一、国家研究所的定位及其评价功能

对研究所进行评价离不开其功能定位。所谓功能定位(Role),是指在一个国家或地区的研究与开发体系中研究所应当具有的功能及承担的使命。科技大国的研究与开发体系,一般都包含国家研究所(有的称为国立研究所或国家实验室)、企业研发机构、大学研究机构以及非营利科技机构等4个层次不同但相互联系的系统。不同层次或系统的研究所,其功能和使命均有所不同。

1. 马普学会的功能定位

马普学会设化学、物理与技术学部,生物医学学部,以及人文学部等3个学部,有80个研究所和科研机构(Institutes and research facilities),人员约11 200名,其中科学家3 100名;另有研究生、博士后、客座科学家等约7 600名。其功能定位是:“在自然科学、生物科学、人文和社会科学方面从事国际最高水平的基础研究”^[2],亦即开展需要大型设备、更多资金,以补充大学和大研究中心的具有特别创新性的前沿研究。

马普学会1992~2001年在国际顶尖杂志上发表论文数占全德的比例为:《自然》,32.14%;《科学》,35.96%;《细胞》,35.48%;《物理学评论快报》,28.38%。^[2]其基础研究水平和地位可见一斑。

2. 中科院功能定位

中科院的学科领域涉及数理、化学、生物、地学、技术科学等,实体组织体现在由基础研究所等5个业务局以及政策局管理或联系的80个研究所,人员约55 000名,其中科技人员36 000名;另有研究生、博士后、客座科学家等约18 000名。其功能定位是:面向国家战略需求,面向世界科学前沿,加强原始科学创新,加强关键技术创新与集成,攀登世界科技高峰,为我国经济建设、国家安全和社会可持续发展不断做出基础性、战略性、前瞻性的重大创新贡献。

3. 国家研究所的评价功能

研究所评价的功能主要有:第一,监督检查,即对公共科技资源的投入,予以公正、客观的评价,并向社会公众交待;第二,决策服务,即为提高科技活动的效率,运用科学而规范的方法、程序对科技活动及预期成果进行评价,以利于科学化决策。

研究所功能定位决定了研究所研究性质、经费以至任务的来源,进而决定了其评价功能。国家研究所的共性特点是,由政府财政拨款资助,从事体现国家整体利益的国家所需的基础性、战略性、前瞻性研究。国家研究所属于公共科研机构,而公共科研机构的评价,对于政府调整科技政策和资助重点、向社会公众“交账”等具有特别重要的意义,因而世界各国无不对其给予足够的重视。

如果缺乏灵活的调节机制,政府拨款资助的研究体制一般会给研究所机制和活力带来“先天性”负面影响。因而,为建立合适的体制与运行机制,就有必要对国家研究所进行评价,并据以进行调节。只有对研究所进行评价并作相应调整,才能保证其完成国家目标所赋予的使命,有效地利用社会公共资源,不断地保持灵活的机制与研究活力。

二、马普学会和中科院研究所的评价方法

1. 马普学会研究所评价

德国科技评价组织体系大体分为3个层次。第一,联邦和州政府,主要评估机构是科学委员会。第二,科研教育资助组织,如德意志研究联合会、马普学会、弗朗霍夫学会、洪堡基金会和大学交流基金会等,它们分别对本组织资助的项目以及本系统的研究单位进行评估。第三,大学和研究院所,各大学和研究中心对各自的研究所和课题组进行评估。马普学会研究所评价属于其中第二层次。其评价方式有以下3种。

(1) 预先评价 (Ex- ante evaluation) 其基本程序是: 通过对立项程序 (Appointment procedures) 进行预先评价, 明确评价概念与项目 (Concepts and Projects), 再建立研究所。由于组建研究所涉及长期基础建设、耗费大量资金, 因而对其计划进行评价非常慎重。对于候选组建的研究所, 需要进行两个阶段的评估。第一阶段, 先成立一个项目组。项目组的科学概念、人员的可行性将被进行正反两方面的预先评价。在马普学会科技人员的支持下, 有外部专家参加的专家委员会着手评价项目组基本的科学概念和项目组合适的候选人。第二阶段, 专家收集来自国际知名科学家的补充报告。在从项目到研究所的 5 年试运行期期末, 由马普学会成员和外部专家组成的新的评价委员会对科学概念的可行性、项目组的工作以及项目负责人的绩效实行后评价 (Ex- post evaluation)。只有在第二阶段通过的前提下, 才可能向马普学会委员会提出建议, 考虑将项目组升格为研究所。

(2) 日常评价 (Regular evaluation) 预先评价后, 对研究所绩效进行日常评价就显得重要了。对每个研究机构进行日常评价是资金得以高效分配的保证, 通常每两年一次, 并根据评价结果核定以后两年的研究经费额度。日常评价由各研究所科学顾问委员会开展。科学顾问委员会的任务是: 在当前研究所开展的活动范围内就马普学会研究方向进行咨询, 审慎地评估研究所存在价值, 以保证研究工作维持在国际性最高水准; 公开评价各个研究所取得的科学成就, 并对研究所以及马普学会今后的研究方向提出意见。

马普学会对研究所日常评价已形成一套系统方法。该方法由外部专家评价 (External expertise)、自评 (Self- evaluation)、主观评价 (Subjective appraisals) 和指标考核 (Performance indicators) 共同组成。

①外部专家评价。指研究所日常评价应主要由研究所以外的专家执行, 并强调要选择在相关学科领域中国际公认的代表组成科学顾问委员会。科学顾问委员会成员由马普学会主席任命, 在任命时要考虑来自研究所的意见。由于这些成员被公认为是相关学科领域有能力的代表, 一般为马普学会科学家信任和接受, 这对日后他们的建议被采纳和研究工作改进非常重要。

在科学顾问委员会中融入外部专家是保证评价程序质量的重要因素。2002 年, 科学顾问委员会共有 600 名成员, 有 97% 的科学家来自其他研究组织, 70% 来自国外, 反映了马普学会评价的国际化形象和客观公正性。

②自评。指在日常评价前期由研究所科学家准备自评报告 (Self- evaluative status report), 汇报当前研究情况和新的研究计划。

自评报告提供了当前论著发表情况和有关人力资源、预算及第三方资金 (Third- party funding) 状况的信息资料。内容包括: 组织与结构, 研究计划, 人员结构, 预算, 物质资源、设备及其空间布局, 青年科学家和客座研究人员, 与国内、国际研究机构的联系, 知识与技术转移, 与商界、政界和社会的联系, 论坛与国际会议, 论著发表情况, 任职、科学奖励和荣誉称号等。报告中对发表的论著进行评价,

对研究所或其部门在国际上的交流联络进行详细而慎重的说明。青年科学家的职业培养是另外一项重要标准。自评报告篇幅依据研究所大小不同而异, 有 100~ 600 页, 最迟在考评研究所 6 周前呈报科学顾问委员会。

③科学顾问委员会综合评价。除了阅读自评报告外, 科学顾问委员会成员还到现场, 通过出席行政所长的报告自行判断研究所发展状况; 听取来自不同学部的报告, 与所长、独立青年研究小组组长以及青年科学家对话。由此, 科学顾问委员会成员才可能得出研究所工作的总体印象, 同时评价研究所的自评情况。

科学顾问委员会报告内容包括两大方面。一是研究所和单个部门的重要内容, 包括: 科学领域内联系, 国内外联系, 科学质量评价, 研究领域前景, 人员结构评价, 资金分配的合理性 (含第三方资金), 与其它马普学会研究所、德国国内和国外外部合作者合作情况, 对青年科学家的支持情况。二是进一步发展的建议方面, 包括: 改进或重组建议, 关闭一个部门或中断一个研究领域的考虑。

④公开性 (Transparency)。科学顾问委员会向马普学会主席提交包含其调查结果、评估及意见在内的详细书面报告; 马普学会主席将报告及其意见与建设措施转交研究所; 研究所对报告可持保留意见, 其意见可以在下一阶段被考虑。评价程序向涉及的每一人公开, 因此体现了评价的公开性。

科学顾问委员会成员既是批评者、有能力的顾问, 又是研究所科学家的密切合作者。这种做法有利有弊。因为, 任何给定的研究领域通常只有相对少量的国际著名专家, 而这些专家与高度专业化的马普学会研究所保持着学术联系, 渐渐地, 这种联系比客观评价所期望的更加密切, 此为不利的一面。但是, 这种现象对于评价工作的深度了解与实质性评价程序又有必要, 此为有利的一面。马普学会试图在深度了解与密切合作之间建立一种平衡。采取的办法是: 一方面, 保持科学顾问委员会成员与研究所业已存在的密切关系; 另一方面, 限制科学顾问委员会任期, 一般为 6 年, 最多为 12 年。

(3) 拓展评价 (Extended evaluation) 这是对研究领域进行评价和比较的一种方法, 目的是对研究所之间及其与国内、国际联系进行比较。该方法基于: 研究所各类相似研究题目结合构成了研究领域; 共性问题容易识别, 因而可以得出对该领域今后发展的总体建议。其步骤如下。

①确立评价标准。报告起草人对确认评价标准的一致性负有责任。每 6 年有 2 位在整个研究领域具有背景知识的外部国际调查人加入科学顾问委员会。他们参加所有研究所在给定研究领域的科学顾问委员会会议。报告起草人评价和比较每个研究所的自评报告、科学顾问委员会评价标准, 并保证评价标准前后一致。

②对一个研究领域的研究所评价之后, 各科学顾问委员会主任、2 个报告起草人、分管副主席以及学部主任聚在一起, 举行研究领域委员会会议。会议目的是对科学顾问委员会报告、报告起草人评价以及领域内可能的资源分配变化进行评议。

③最后, 研究计划高级委员会将对单个研究领

域委员会考察的结果进行讨论,中心议题是研究政策。

每年拓展评价的结果和研究领域委员会的陈述报告都将呈报给马普学会最高决策机构评议会。评议会由马普学会成员、政界和商界代表组成,因而投资方可以及时了解研究质量控制措施和资金分配的最新进展。

马普学会研究所评价由委员会操作。同行评议是马普学会研究所评价的核心。它基于一种信誉,表明只有在一个学科领域的同行才有资格充分地评价其他同事的工作。其做法是,选择本学科最优秀的同行来评价该学科领域最佳科学家的科研成效。除采用同行评议方法外,其他考虑的有关科研绩效的因素包括发表论著数量、争取第三方资金情况、投入产出结果、学生培养情况、所长与博士研究生或博士后比率以及每位所长投入到每位学生的时间、奖励和其它科学荣誉称号。这些其它因素都可以用数据进行计量,它们综合在一起只是构成同行评议的补充,并且用于中期到长期的趋势分析。

2. 中科院基础类研究所评价

(1) 中科院研究所评价体系 中科院研究所评价的特点是,按行政管理并参照学科性质将研究所分为基础研究、高新技术研究与发展、社会可持续发展科技研究和产业化等4类,建立了统一的、互有侧重的评价指标体系(图1所示)。其中重大创新贡献评价每2年1次,只涉及部分研究所;基础指标测评为年度评价,1年一次,涉及全部研究所。

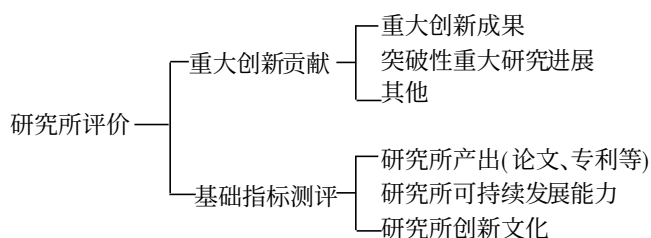


图1 中国科学院研究所评价体系

(2) 基础类研究所评价方法 基础类研究所评价采用以计量学评价为主,结合国内同行(通讯)评议、管理部门综合评价等综合评价方法。评价指标包括科技论文、国际学术会议报告、科技奖、研究项目、重大创新贡献。根据这些指标的量化值进行评分,并在基础类研究所进行内部排序,包括单项指标排序和综合排序。

(3) 基础类研究所评价主体 评价主体除包括基础研究局外,还包括涉及基础类研究所的生物局、资环局、高新技术研究与发展局、政策局;复评委员会;人事教育局、京区党委;中科院评估研究中心。它们分别负责:评价标准和指标制定;组织专家初评;复评;创新文化和所长管理行为评价;综合统一测评等各项评价工作。评价专家几乎全部来自国内,有相当比例属于管理部门的工作人员。如2002年中科院研究所评价,有676名院内外科技和管理专家应邀参加了同行评议。这些同行一般是临时被邀请参加研究所评价,没有形成专门委员会。

(4) 基础类研究所国际化评价探索 近年来,中科院一些以基础研究为主的科研院所如数学与

系统科学研究院、物理所等,为争创国际一流,作为一种探索,自发地组织了国际化评价。有关院所领导和专家认为:不同类型的研究所对此应有所不同,在一些与国际差距小的研究单位实行国际化评价方法比较可行,但在更多科研院所实行还值得探讨和研究。中科院人教局拟在“海外华人专家库”基础上,建立起基础研究国际同行专家库,为部分基础类研究所试行国际化评价作准备。

三、对我国基础类研究所评价的启示

1. 国家研究所评价功能与评价“外部化”程度

国家研究所存在“交账”心理。对其评价不仅应当实现外部评价的“决策咨询”功能,而且更应该实现“监督检查”功能。中科院研究所评价的目的主要是检查知识创新工程试点研究所的改革情况,评价基本实现了机构内部评估的“决策咨询”功能,但没有实现外部“监督检查”功能。马普学会较好地实现了“监督检查”和“决策咨询”这两个功能的统一,并要求对外界公开(Transparency)。其科学顾问委员会中有97%的科学家来自其他研究机构、70%来自国外。中科院研究所评价的同行评议专家大多来自国内或机构内部。其原因主要是研究所评价的宏观性与相对封闭;其次是因为科学共同体尚未得到充分发育,机构以外的专家在被邀从事评价工作时并未将学术评价活动视为一种荣誉。

在基础研究领域,科学本身没有国界。考虑到评价成本等情况,同行评议的专家通常有一定比例来自于机构内部。但由于基础研究的前沿性和国际性,为了保证评价公开、公正和高水平,应从制度上保证有相当比例的专家来自本国大学、政府部门、产业界以及海外,此即提高评价“外部化程度”。比例多少合适,是基于对语言、评价成本和国内研究水平等现实情况的综合考虑。“国情”虽然可以作为过渡条件,但不应成为偏执于“内部评价”的“托辞”。正如著名数学家丘成桐所言,科学领域是有所谓“中国国情”或“非中国国情”的。他认为推进中国科学评价制度是一个非常要紧的问题,而不愿意邀请外国专家评议显然是一个很大的损失(转引自“六外籍院士为中国科学把脉”,《望》,2000)。

为此,建议提高我国研究所评价的外部化程度。对于基础类研究所而言,评价专家“外部化”主要是指“国际化”,包括评价专家国际化和评价标准参照的国际化。国际化的评价对提升研究水平将会起到极大的促进作用。马普学会认为,只有建立在国际化、学科交叉、灵活机制的基础上,从事新的和富有前景的前沿研究题目的使命才可望成功。^[2]

2. 评价指标的规范性与评价周期

中科院研究所评价指标包括研究所科研工作绩效、人才队伍建设、科技项目对外争取情况等,同时也包括创新文化建设和所长管理行为、重大创新贡献。这些指标在评价内涵上存在互相重复的成分,如创新文化建设评价和所长管理行为评价,研究所绩效中自然离不开创新文化和所长管理行为,创新文化和所长管理行为的功效最终体现在绩效中。人为分割,无异于重复评价。马普学会研究所评价并未将后者单独分离出来。

在中科院研究所评价指标体系中,产出指标为

上年数据;重大贡献也只考虑前2年。这对于以基础研究为主的研究所来讲,容易产生负面效应。基础研究须经较长时期的探索才可望产生出原创性成果;而在这种年年评的制度中,谁还敢于碰那些创新性强、难度大的课题?!这种制度无意识地助长了好大喜功、急功近利的倾向。

建议:规范评价指标体系,避免重复设置;简化计量指标,并把它们作为同行评议的补充;适当延长基础类研究所的评价周期,从目前的1年1次延长到2年乃至4年1次。

3. 分类导向评价与评价标准的一致性

马普学会十分重视评价标准的一致性,如在其研究所拓展评价中,需要报告起草人(Rapporteur)评价和比较每个研究所的自评报告、科学顾问委员会评价标准,并保证评价标准前后一致。

中科院的研究所,除产业化类外,其余3类都涉及到基础类研究所,学科领域与主管局相互穿插。为此,研究所评价采用了“分类导向指标”,这些指标由院评估研究中心与有关业务局协商后确定,其中基础指标的评价标准由4个业务局根据各自联系或管辖的研究所特点分别制定。由此容易造成政出多门、各类研究所评价标准不一的问题。由于过于强调“分类”和所谓特色,容易忽视普遍性规律。例如,本来基础研究中涉及生物,生物研究中不乏基础研究,如果说由于历史原因而形成了基础局、生物局、资环局并存的局面尚属于操作层面的话,那么,在评价同一生物学领域时,分别由不同的管辖局制定标准、组织初评和复评,难免造成标准多样化。虽然这种做法使得评价工作分工和目标明确,但分类评价的标准缺乏国内外甚至机构内的可比性。而且,这些标准也不能相对稳定,如2002年标准与2001年的标准就不一致。

对基础类研究所评价,不论其属于哪一个部门或系统管理,都应采用一致的评价标准,并保持其前后的相对稳定性,以便进行同一领域时间纵向与研究所之间横向的对比。只有加强研究所评价的可比性,才能走向国际化,增强评价的说服力,才能更好地制定资助政策。马普学会用拓展评价来评价相近领域研究所之间以及国内外差异,有利于从更广泛的视角进行比较。

建议根据科技发展趋势和国家战略需求来调节学科领域评价指标,改变目前按基础类研究所分类、按部门导向的评价为按学科领域的评价。相似领域采用同一标准,才便于进行国内外研究所之间的比较,增强研究所竞争能力。

4. 评价的独立性与交互性

科技评价一般应当遵从独立、客观、公正、回避、外部评估、科学规范、可行性等原则。只有加强评价的独立性,才能维护客观、公正、公平的原则。

马普学会研究所评价的主体是学部立项委员会和科学顾问委员会,相对独立。中科院的评价主体包括多个部门,最高决策机构是秘书长办公会议,既是评价工作的委托者,又是评价工作的实际操作者,其客观性受到一定的质疑。由此产生的不利影响从评估报告可以反映出来。中科院评估报告较多地关注排序,主要内容是每个研究所评估总分及其排序,各单项指标得分及其排序;评价报告

展现出一定程度的自我乐观,正面的评价结论很多,针对具体研究所提出问题和措施建议的很少,只是在报告中笼统地提到,似有“为评价而评价”之嫌。马普学会的评价报告注重单个研究所的研究质量、发展潜力与对外联系等。

评价过程的交互使得评价意见易于采纳,从而使评价易于实现“决策咨询”功能。马普学会日常评价对研究所评价的交互性发挥了重要作用。

5. 同行评议与计量评价

同行评议一般是指某一领域的同行专家采用一致的评价标准,共同对涉及领域的一项事物进行评价的活动。计量评价指将文献等评价指标量化,然后加权打分。

马普学会研究所评价方法以国际同行评议为主,即以同行专家到被评估的研究所实地考察后的直接、感性评价为主;辅以计量指标。而且,马普学会只是在分析中期到长期趋势时才采用文献计量指标,将文献计量“作为同行评议系统主观评价的补充”^[2]。中科院基础类研究所评价以指标量化为主,其中文献计量占有核心地位。

建议:以同行评议为主,在有必要进行指标量化评价时(如缺乏足够的同行评议专家),也要充分认识到文献计量指标的使用具有局限性和时间性,以及文献计量中的“量”与“质”的关系。文献计量指标可作为同行评议的辅助手段或补充,但不应成为主要导向。如果将其作为主要导向,与资源分配直接挂钩,各研究所则可能被引导到对文献数量的机械追求而忽视对研究本质的关注,从而导致一系列弊端。也就是说,直接强调文献计量指标,容易导致“重视量而忽视质”,割裂研究内涵。同行评议由于注重研究内涵,才不至于因为短期计量指标问题而中断对真正创新研究的持续支持。

6. 国家研究所与其他研究机构的联系

近现代基础科学的发展表明,基础研究成果促进新技术发展和新型产业生长的进程常常比预料的要快得多;新技术的产生越来越依赖于基础科学理论突破;从基础理论到新技术产生,再到形成新型产业并由此促进国家经济建设的发展周期越来越短。为便于找准目标,并及时有效地利用国家创新体系中的其他要素资源,就需要加强基础类研究所与其他机构的联系。

马普学会很注重评价研究所与国内工业界和大学的联系。我国基础类研究所这种联系较为松散,在评价时尚没有形成类似的概念性指标。“产学研工程”实施多年,但在对其活动和效果进行判定,以便对建立相互之间的联系与合作予以指导方面尚显不足。为此建议,我国在评价基础类研究所时,应充分考虑基础研究尤其是应用基础研究与工业界和大学的联系。

参考文献

- [1] 中国科学院评估研究中心, 2003、2002 中国科学院知识创新工程试点单位评价报告
- [2] The Max Planck Society, 2002. Evaluation
- [3] Masayuki Kondo, December, 1998. R&D Evaluation in Japanese MITI. Proceedings of the APEC Symposium on the Evaluation of S&T Programmes among APEC Member Economies. Wellington, New Zealand

(责任编辑 王宏章)