

科技部科普专项经费资助课题

科普评估理论初探 与案例指南

KEPUPINGGULILUNCHUTAN YU ANLIZHINAN

〔张义芳 主编〕



科学技术文献出版社

科技部科普专项经费资助课题

科普评估理论初探 与案例指南

张义芳 主 编

科学技术文献出版社

Scientific and Technical Documents Publishing House

北 京

图书在版编目(CIP)数据

科普评估理论初探与案例指南/张义芳主编.-北京:科学技术文献出版社,2004.9

ISBN 7-5023-4762-3

I. 科… II. 张… III. 科学普及-评估-指南 IV. N4-62

中国版本图书馆 CIP 数据核字(2004)第 085234 号

出 版 者 科学技术文献出版社
地 址 北京市复兴路 15 号(中央电视台西侧)/100038
图书编务部电话 (010)68514027,(010)68537104(传真)
图书发行部电话 (010)68514035(传真),(010)68514009
邮 购 部 电 话 (010)68515381,(010)58882952
网 址 <http://www.stdph.com>
E-mail: stdph@istic.ac.cn
策 划 编 辑 科 文
责 任 编 辑 金 平
责 任 校 对 赵文珍
责 任 出 版 王芳妮
发 行 者 科学技术文献出版社发行 全国各地新华书店经销
印 刷 者 富华印刷包装有限公司
版 (印) 次 2004 年 9 月第 1 版第 1 次印刷
开 本 850×1168 32 开
字 数 256 千
印 张 10.5
印 数 1~2000 册
定 价 20.00 元

© 版权所有 违法必究

购买本社图书,凡字迹不清、缺页、倒页、脱页者,本社发行部负责调换。

(京)新登字 130 号

内 容 简 介

科普评估是保证科普事业持续健康发展的一项基本制度。本书在系统研究国外科普评估实践经验的基础上,研究借鉴教育和科技领域的评估理论与方法,系统解答了有关科普评估中的基本问题,主要包括以下方面的内容:科普评估的概念与作用;科普评估的结构框架与理论模式;各类科普评估的指标框架;评估程序与方法以及对我国未来开展科普评估工作的相关认识与建议。书中附录部分编译了国外有代表性的各类科普评估报告,相信对我国今后的科普评估实践具有很好的借鉴价值。

本书的最大特色在于内容的系统性和实用性,希望科普管理者、科普研究者以及从事科普工作的一线人员均能从中发现所感兴趣的内容,也希望本书能通过推动科普评估实践为提高中国科普工作的成效发挥一定的作用。本书可以作为科普评估培训教材和各类科普评估人员的参考书。

科学技术文献出版社是国家科学技术部系统唯一一家中央级综合性科技出版机构,我们所有的努力都是为了让您增长知识和才干。

目 录

序言	(1)
第一章 科普评估:概念、类型及其在国外的发展	(5)
第一节 科普评估的概念与作用	(5)
第二节 科普评估的类型	(9)
第三节 科普评估在国外的兴起	(12)
第二章 科普评估的结构框架与评估模式	(15)
第一节 科普评估的结构框架	(15)
第二节 科普评估的基本模式	(22)
第三章 科普评估的指标框架	(33)
第一节 建立评估指标框架的原则	(33)
第二节 建立评估指标框架的学理基础	(34)
第三节 战略规划或项目资助计划的评估	(38)
第四节 科普项目的评估	(40)
第五节 科普管理部门/执行机构能力与绩效的评估	(58)
第四章 评估程序与评估方法	(63)
第一节 评估程序	(63)
第二节 评估调查方法	(75)

第三节	评估分析方法	(90)
第四节	评估调查分析中应注意的问题	(97)
主要结论与建议		(99)
参考文献		(105)
附录 国外科普评估报告实例		(109)
附件一	美国国家科学基金会非正规科学教育计划 评估报告(1998)	(111)
附件二	美国国家科学基金会非正规科学教育计划 访评报告(2001)	(142)
附件三	“比尔·奈科学人”电视系列片评估报告(1996)	(188)
附件四	美国创新技术博物馆展厅评估报告(2000)	(256)
附件五	美国创新技术博物馆评估报告(2000)	(282)
附件六	英国“地铁中的科学”招贴画项目评估报告 (1998)	(314)

序 言

过去半个世纪以来,科普事业对提高我国公民的科学文化素养做出了巨大贡献。特别是近年来,中国各级政府对科普给予了高度重视和大力支持,对科普事业的投入显著提高,《中华人民共和国科学技术普及法》的实施更为科普事业的进一步发展提供了法律保证。然而,不能否认的是,由于多方面的原因,我国的科普工作长期存在着总体效率不高、责任不强、创新不昌、实效不明等问题,这些问题直接影响着我国科普事业未来的健康发展。面对科教兴国战略的深入实施、国内科教管理改革以及先进国家科普管理制度等的影响,健全我国的科普管理机制、通过实施科普评估制度提高科普工作成效已成为一个十分现实而迫切需要解决的问题。

我国科普工作的确有建立评估制度的必要性和紧迫性。在我国以往的科普工作中,科普管理部门与执行机构着重做什么事,开展什么活动,却没能顾及一个很重要的问题,即行动的效果或结果是什么,也就是做得怎么样的问题。比如,我国在科普事业上的持续投资到底收到了多大的价值?重要的科普活动或项目的执行效果到底如何?有什么值得借鉴的经验教训?如果继续开展,有哪些地方需要改进?科技馆等是否切实履行了其科普使命而物有所值?科教频道和众多的科普栏目的传播效果怎样?这些为科普管理所必不可少的信息,只有通过评估才能获得。而且评估本身也发挥着棒的无形管理作用,它能促使科普机构提高其工作的责任、

效果、效率和社会影响力,有利于实现资源的优化配置。可以说,缺少评估,任何管理都是不健全的。

评估是保证科普事业持续健康发展的一项基本制度。要建立这一制度,需要政府制订相关的政策法规去推动,而在此之前,则需要研究人员首先解决科普评估模式、框架、指标、方法等技术工具问题。2003年中国科技信息研究所受科技部政策体改司的委托进行的两年期的科普评估理论与实践研究课题正是基于这一需要而确立的一项重要研究任务。

然而,做好这一课题面临很大的复杂性和挑战,因为科普评估研究在世界上还属于一个尚在探索的领域。在课题的开始阶段,我们花费了大量的时间,从期刊和网络上查找国外资料,但所得甚少。我们发现,美英等国家虽然开展科普评估实践已有多年,但发表的研究文献却寥若晨星。在这种情况下,我们一方面采取了与国外的评估机构直接联系,获取它们的科普项目评估案例,另一方面,我们转向范围更广的教育评估,查找了大量教育评估文献,从中研究分析可供科普评估借鉴的内容。事实证明,我们采取的策略是有效的,因为从我们掌握的国外科普评估案例来看,它们的科普评估中确实借鉴了教育评估的很多理论和模式。

本课题的最大难点在于确定科普评估的指标框架。科普现象的复杂性、科普情景的多样性以及科普效果的难测性,使得指标框架的确立相当棘手,不可能建立一个放之四海而皆准的、适合各种情况的统一的指标框架。为此,我们采取了分类评估的方法,针对不同类型的科普评估建立不同的指标框架,比如,对科普项目评估,我们分科技展览项目、媒体项目和科技活动项目三个类型,分别确立了各自的指标框架。在这方面,我们没有现成的东西可以照搬,因为我们所看到的国外评估手册都不涉及指标的内容,我们所掌握的国外科普评估报告也都没有评估指标的明确说明,只有靠我们自己去演绎归纳。另外,对科普管理部门、执行机构的评

估,由于国内外基本上都未曾开展这方面的评估,也完全靠我们自己探索研究,这些都是我们格外花心思的地方。总之,对这一课题,我们投入了两年的时间,进行了大量认真细致的研究,但由于研究是探索性的,还需要接受未来实践的检验和修正。

在进行课题研究的过程中,我们越来越深刻地认识到,制度化的科普评估对确保我国科普事业的未来发展意义重大,应作为推动我国科普工作的一个重要组成部分,列入我国科普事业发展纲要。科普评估制度势在必行,如能真正落实,必将促进科普工作绩效的提升,加快科普组织的成长,开创我国科普事业的崭新局面。

第一章 科普评估:概念、类型 及其在国外的的发展

20 世纪 70 年代以来,评估开始受到国际社会的普遍关注,并广泛应用于投资、教育和科技等领域,国外科普界也日益重视评估的问题。然而在我国,科普评估尚未普及,科普评估研究才刚刚开始,整个科普界对评估的定义、评估的类型、评估的作用等基本问题尚很陌生,因此,本章首先介绍一些有关评估的基本概念。

第一节 科普评估的概念与作用

一、科普评估的定义

评估近年来得到广泛应用,尤其是在投资领域,已经成为政府、金融机构或建设单位等投资主体进行项目投资决策的基本依据,然而这种含义的“评估”与我们所关心的科普评估完全属于不同的学科体系。前者是应用经济学的一个分支,以成本-效益为核心,以增加利润为目的。而后者属于社会学研究的范畴。因为其一,科普工作总体上具有社会公共项目的共同特点:非营利性和追求社会效益。其二,科普属于教育的范畴,科普评估采用的是教育评估的研究方法。因此我们对科普评估的定义来源于教育评估的描述,即:科普评估是根据委托方的明确目的,由评估机构遵循

一定的原则、程序和指标,运用科学、公正和可行的方法,对科普组织、计划或活动进行价值评判,并提出改进的建议和措施的专业化活动,其目的在于提升科普教育的品质。

科普评估是评估在科普领域的应用。它具有一般意义的评估所具有的特点:

第一,时限性。科普评估是在一定的时间范围内完成(如一周或一个月),而不是贯穿于项目、计划或政策的始终。这也是评估区别于监测的主要特性。

第二,系统性、规范性和科学性。科普评估与以往工作总结最大的区别就在于,评估是有一套完整的理论和指标体系,规范、科学和有目的的评定活动。而中国传统的工作总结往往缺乏系统的定性和定量指标,随意性较大。

第三,学习性。科普评估不仅仅是对科普组织、计划或活动的相关性、绩效和成功度的简单评价,而是通过评价活动本身进行学习。甚至可以说,学习加责任就等于评估。事实上,通过评估过程进行学习的知识更具有权威性,特别是对一系列的项目、计划或政策进行的评估,从中所学习的经验更适合在更大的范围内推广。

二、科普评估与监测的区别

监测与评估是人们在实际工作或研究过程中经常混淆的两个概念。

监测是一种持续性的职能,它是指收集信息数据,了解工作是否按原计划进行,即检查工作是否按计划投入资源、开展活动、产出成果的过程。这种职能主要是为了给管理部门和主要利益相关群体及时提供项目、计划或政策的进展状况,如果发现项目的执行与原计划不符,就要采取措施纠正偏差,使其恢复到原计划设定好的轨道。其着眼点不在于项目、计划或政策的目标结果上。

评估与监测的主要区别在于:

第一,评估是一定时限内的工作;而监测是连续性的职能,它贯穿于计划、项目或政策过程的始终。

第二,评估是根据具体情况,有选择地进行。也就是说,并不是每一个项目、计划或政策都需要进行评估;而监测活动原则上需要对项目、计划或政策都执行。

第三,监测是针对单个项目进行;而评估可以针对一个或几个项目、计划或政策进行。如行业评估。

第四,监测只需要记录实际的数据,用于与计划数据对比;而评估不仅可以利用监测的数据,通常也需要进行一些项目、计划或政策以外的数据对比。例如,与行业标准规范对比,或与同类同规模的项目、计划或政策的数据对比。

第五,监测通常通过定期的报告、报表、考察等方式,向项目、计划或政策的执行人员了解投入、活动和产出情况;而评估主要通过阶段性的座谈、考察或问卷调查等方式向服务对象了解工作的效果、效率、影响和持续性。

三、科普评估的作用

评估作为一种管理工具,具有品评鉴定、激励、导向、诊断、改进发展等多种功能。

- 激励的功能:可促进机构努力向上,以求最佳绩效。
- 回馈的功能:即检讨反省。评估是科普过程的一系列反馈环。评估能促进组织“省思”和学习。
- 品评的功能:根据事实评判其绩效。
- 改进的功能:评估的主要目的在改进。通过评估,可以了解计划或项目的优点与缺失,进而形成改进的意见和建议,使未来计划或项目的实施更趋完善、更有成效。
- 品管的功能:评估可使产出维持一定水准,亦即对计划或项目的执行过程具有品质管制的作用。

- 诊断的功能:通过收集资料和分析,可指出计划、项目、组织中的问题和困难,作为改进的依据。

在科普领域开展评估目前有两大主要意图:一是“证明”(to prove);二是“改进”(to improve)。所谓“证明”,就是依据收集的证据,判断科普组织、计划或项目的绩效,也就是要证明组织运作、计划或项目实施的“有效”或“无效”。所谓“改进”,则强调在评估过程中,了解组织、计划或项目的优点与缺失,进而形成改进的意见和建议,使未来的组织工作、计划或项目的实施更趋完善、更有成效。我们认为,科普评估应特别突出改进发展的目的。正如美国著名的评估学者 Stufflebeam 所说:“评估的目的不在证明什么,而在力求改进。”目前的许多教育评估似乎多偏重“证明”,而非针对评估对象的短处,积极地辅助其改善,最终达到受评对象自我成长的理想目标。故评估的目的是什么?是一种品质保证、绩效责任、认可的制度、改良的依据?或作为机构发展的参考指标?这些是我们进行具体评估前所应明确的。

另一方面,要真正有效发挥评估的上述作用,评估工作本身首先必须是高质量的。为保证评估的质量,美国教育评估标准委员会(Joint Committee on Standards for Education Evaluation)制定了30项后评估标准,用以评估其评估工作。这套教育评估标准同样适用于科普评估,它共有30项,分为四个组群,分别是效用的标准、可行的标准、正当的标准及精确的标准。

- 效用的标准:即评估应当是有用的,应满足评估用户的信息需要(在科普领域,评估应能协助被评估对象确认其优点、存在的问题并提出改进的建议)。

- 可行的标准:即评估应当是可行的。评估应运用可操作的、严谨的评估程序,所需的资源、人力和时间必须符合成本效益。

- 正当的标准:即评估应当是合法和讲求伦理的。评估过程中应维护有关团体的权益,在出现利益冲突时能坦然处之,不妥

协,谨遵应有的道德规范。

- 精确的标准:即评估应当是精确的。评估应建立在可靠的信息收集与分析基础之上,应清楚准确地描述评估对象的价值,显示评估计划、程序、结果等,并提供有效的研究结论与具体建议。

第二节 科普评估的类型

根据不同的标准,可以将评估分为多种类型。按照时间顺序,评估可分为预评估、形成性评估和总结性评估;按照评估者的来源,可分为自我评估、外部专家评估和参与式评估;按照评估内容,则至少可分为科普项目评估、科普计划评估和科普管理部门/执行机构的评估。

一、预评估、形成性评估和总结性评估

预评估是指科普项目或计划开始实施之前所进行的评估,也称事前评估。由于是在实施之前进行,因此实际上是对项目和计划可行性分析的评估。预评估的结果一方面可以决定项目或计划是否实施,以避免对低效科普项目或计划的大规模投入。另一方面预评估取得的数据可以作为基准线,在项目或计划完成后进行前后对比。

形成性评估是指在项目或计划开始后到完成前之间的任何一个时段进行的评估,也称事中评估。它的目的在于检查项目或计划的设计,或者评估实施过程中的重大变更及其影响,或诊断实施过程中的困难、问题,从而为项目或计划执行者提供改进的方向。

总结性评估是在项目或计划结束后进行的全面、系统评估,也称事后评估。总结性评估的目的是通过分析评估找出项目或计划的优点与缺失,并通过及时有效的信息反馈,为未来新项目提供经验教训。总结性评估一般更容易为科普机构理解和接受,也更受到重视。

预评估、形成性评估和总结性评估在评估时点、评估目的、评估主体上互有差异。在理想状况下,一次全面的项目或计划评估应当包括预评估、形成性评估和总结性评估三个阶段,但在实际评估工作中,由于经费和时间的限制以及评估目的的不同,经常有选择地进行。

二、自我评估、外部专家评估和参与式评估

根据评估者的来源,评估又可以分为自我评估、外部专家评估和参与式评估。自我评估即项目或计划的实施者进行的内部自我评估。专家评估是指项目或计划的实施机构聘请外部专家进行的评估。参与式评估是吸收受评对象(如科普项目执行者)和项目受益者(科普对象)参与评估实施过程的评估。在实际工作中,选择何种评估类型,将取决于评估的目的和具体的情况。

自我评估的优点首先在于评估者是项目或计划实施机构的成员,对组织内部结构与运行机制较为了解,对项目 and 计划实施的整个过程较为熟悉,对科普社区的社会经济背景较为了解;其次,自我评估的成本较低,在项目或计划实施的每一个阶段都可以进行评估,有利于及时发现问题、解决问题;最后,自我评估的结果或建议容易在执行过程中得以实现。这是因为一方面评估者本人就是项目或计划的执行人;另一方面他们对于评估过程中发现的问题有切身的体会。然而,自我评估也存在一些局限。由于评估者是实施机构的内部成员,因此评估结果往往缺乏客观公正性。而且,即使自我评估很客观,评估结果也不容易为有关方面完全相信。另外,由于自我评估者的日常工作较多,难以集中精力进行评估,因此自我评估较为粗略。

而专家评估则可以弥补自我评估的不足。专家评估往往较为客观公正,评估结果易于被有关方面所接受;专家的专业评估知识也可以增进评估的科学性与规范性;而且专家评估往往是集中一

段时间进行评估,效率较高。但是,专家评估也有其缺陷。首先,专家评估的成本较高,因而也不可能在项目或计划执行的每一阶段进行,更不可能对每一个项目都进行专家评估;其次,外部专家对项目或计划实施的整个过程、运行机制和背景并不十分清楚,而专家往往又需要在较短的时间内完成评估工作,因此有人戏称专家评估为“旅游式评估”;最后,专家评估的结果或建议并不容易在项目或计划的执行过程中得以真正落实。

可见,自我评估的长处正是专家评估的短处,而自我评估的短处正是专家评估的长处。有时,在条件允许的情况下,将两种评估方式结合起来,评估的效果可能更好。参与式评估正是在此基础上演绎发展出的新兴评估方式。

参与式评估是指外部评估者吸收受评对象和科普目标群体的成员共同组成评估小组,对科普项目、计划等进行评估。

参与式评估要求受评对象和科普目标群体人员参与评估的领域、评估指标的选择、收集和整理数据、分析结果、将评估的信息用于目标的实现等全过程。而不是传统的由专家或项目管理机构得出评估结论。需要说明的是,参与式评估绝不是为了简单地增加评估的人员,而是通过受评对象和目标群体成员的直接参与提高评估的质量。

参与式评估的最大特色是打破了传统的思想认识和评估方式,即“自上而下”,迷信上级和专家权威的评估。而是鼓励受评对象和目标群体成员参与评估过程,充分调动利害关系人的积极性。

一般而言,参与式评估不是一种自发的行为。因此,为了进行参与式评估,必须运用积极的、操作性强的方法保证利害关系人的参与。如果可能,还需要对他们进行参与式评估的培训。

三、项目评估、计划评估和管理部门/执行机构的评估

如前所述,科普评估是对科普组织、计划或活动,通过科学系