

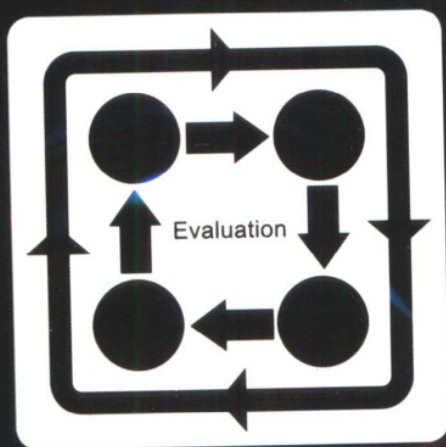
智勝
BEST-WISE

公共管理评估丛书

评估模型

Evaluation Models: [Second Edition]
Viewpoints on Educational and
Human Services Evaluation

Daniel L. Stufflebeam, George F. Madaus and Thomas Kellaghan
译者 苏锦丽等 审校 邓国胜等



北京大学出版社
PEKING UNIVERSITY PRESS

评估模型

Evaluation Models: [Second Edition]
Viewpoints on Educational and
Human Services Evaluation

在发达国家，“评估学”已经成为了一门专门的学科，并迅速发展成为超学科的显学。本书由评估大师Dr. Daniel L. Stufflebeam率同25位作者，合撰21世纪适用的方案评估的主要模型。书中首先介绍了发达国家计划或项目评估发展的历史，随后系统、详细地介绍了问题/方法取向、改善/问责取向、社会议题取向等不同类型的22种评估模式。最后，作者还推荐了21世纪最适用的几种评估模型，包括当前国际上最为流行的问责性评估模型、以实用为导向的评估模型等。本书不仅适用于公共管理学科、工商管理学科和教育学、心理学、社会工作、社会学的师生，也适用于市场调研、评估领域的实践工作者。

ISBN 978-7-301-11473-5



定价：59.00元



公共管理评估丛书

评估模型

Evaluation Models: Viewpoints on Educational and Human Services Evaluation

[Second Edition]

Daniel L. Stufflebeam, George F. Madaus and Thomas Kellaghan

译者

苏锦丽 王丽云 郭昭佑 王振世 张德胜 叶忠达 吴政达
刘淑滢 施香如 陈金盛 詹惠雪 苏永明 黄曙东 李安明
颜国梁 陈美如 林志成 王保进 游家政 曾祥榕

简体版审校

邓国胜 滕 瑶 张 弛 甄佩宏



北京大学出版社
PEKING UNIVERSITY PRESS

著作权合同登记:图字 01-2006-5735

图书在版编目(CIP)数据

评估模型/(美)斯塔弗尔比姆(Stufflebeam, D. L.)等著;苏锦丽等译. —北京:北京大学出版社, 2007. 1

(公共管理评估丛书)

ISBN 978-7-301-11473-5

I. 评… II. ①斯… ②苏… III. 评估—方法 IV. C93-03

中国版本图书馆 CIP 数据核字(2006)第 160406 号

簡體中文版由高等教育文化事業有限公司(Taiwan)授權出版發行

評估模型. Stufflebeam, D. L. 等著, 蘇錦麗 等譯.

2005 年, 第一版, ISBN 957-814-601-9

书 名: 评估模型

著作责任者: Daniel L. Stufflebeam 等著 苏锦丽 等译

责任编辑: 丁莉华

标准书号: ISBN 978-7-301-11473-5/G·2030

出版发行: 北京大学出版社

地 址: 北京市海淀区成府路 205 号 100871

网 址: <http://www.pup.cn>

电 话: 邮购部 62752015 发行部 62750672 编辑部 62117788

出版部 62754962

电子邮箱: law@pup.pku.edu.cn

印刷者: 北京飞达印刷有限责任公司

经销者: 新华书店

730mm×980mm 16 开本 39 印张 540 千字

2007 年 1 月第 1 版 2007 年 1 月第 1 次印刷

定 价: 59.00 元

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有, 侵权必究

举报电话: 010-62752024 电子邮箱: fd@pup.pku.edu.cn

译者、审校者简介

苏锦丽

现 职：台湾新竹教育大学教育学系教授兼
教务长

研究专长：教育行政、高等教育评鉴、学校校
务评鉴

最高学历：美国夏威夷大学教育行政博士

教授课程：教育行政学专题研究、引导研究

邓国胜

现 职：清华大学公共管理学院 NGO 研究
所所长、副教授

研究方向：公共组织问责、绩效评估与能力建
设、中国 NGO 研究

主要著作：《非营利组织评估》、《公益项目评
估》、《群众评议政府绩效》（合著）

公共管理评估丛书

- 《群众评议政府绩效》
- 《评估模型》
- 《政策执行与评估》
- 《第三部门问责与评估》
- 《民间组织评估体系》
- 《事业单位治理结构与绩效评估》
- 《香港社会服务评估与审核》

责任编辑：丁莉华

《公共管理评估丛书》

编委会

总主编 邓国胜（清华大学公共管理学院NGO研究所所长）

主 编 江明修（台湾政治大学公共行政系教授）

陈锦棠（香港理工大学应用社会科学系首席讲师）

《公共管理评估丛书》总序

公共组织不以私人利益为追求目标,因此公共组织往往容易缺乏内在发展的动力。另一方面,公共组织又具有相对的垄断性,缺乏市场竞争机制,因此组织自我改进与发展的外在动力也不足。公共组织的这种先天不足,会直接影响到公共组织服务质量与绩效的提高。因此,公共组织最迫切需要解决的问题就是动力源泉!

公共组织的动力有两方面的来源,第一是组织内部自身,即组织的使命与员工对使命的认同。但遗憾的是,公共组织经常会出现使命失灵的现象;第二是组织外部的压力,但这种外部的压力与市场竞争机制又有很大不同,它不是建立在优胜劣汰的基础上,而是建立在评估的基础上。也就是说,外部的评估是公共组织提升绩效与服务质量的重要动力来源。这也是为什么评估成为新公共管理的热点的原因。事实上,无论是发达国家,还是发展中国家,政府在推进行政体制改革或公共事业改革方面,都离不开评估这一管理的核心工具。

中国的改革更离不开评估。改革实际上是对利益进行重新调整与分配的过程。随着中国改革开放的不断深化,

改革的阻力与难度也越来越大。而从技术层面入手、从管理改革入手,就具有较强的正当性与可行性。它不仅符合世界的潮流,容易为利益各方所接受,而且也容易取得立竿见影的效果。而作为管理改革的核心工具,评估是最理想的技术手段之一。也就是说,从评估切入,通过评估来推动中国的改革是未来一个切实可行的路径与选择。

然而,评估又是一件非常困难的事情,甚至可以说是管理学的世界性难题。毕竟评估是一种基于价值的判断,具有较强的主观性。因此,没有绝对客观、公正与科学的评估。特别是对于公共组织而言,由于其目标的多元与弹性和指标的难以测量,评估工作更是难上加难。在这种情况下,公共组织的评估就更需要小心谨慎,更需要理论的指导。在西方国家,评估作为一门专业与学科,始于20世纪60年代后期。到20世纪80年代,这门学科已经相对成熟。然而,中国公共管理学界对评估的研究才刚刚起步。可以说,在很大程度上,中国的公共管理是实践先于理论。

近年来,由于实践的需要,国内各个地方开展了各式各样的评估活动,例如政策评估、政府绩效评估、项目评估等。然而,由于中国相关理论研究的滞后和评估专业人员的匮乏,在缺乏理论指导和缺乏专业评估人员的情况下,有的地方开展的评估不仅没有取得预期的效果,反而产生了一些负面的效应。因此,中国的公共管理学界理应积极回应社会的需求,加强评估领域的研究、评估知识的传播和评估专门人才的培训。

正是基于以上考虑,我们组织出版了这套《公共管理评估丛书》。本丛书具有以下特点:

第一,系统性。这套丛书不仅包括政府绩效评估、事业单位绩效评估、非营利机构绩效评估、社会福利机构评估方面的专著,也包括评估理论模型、政策评估、公共服务与事业评估等方面的著作。

第二,前沿性。本丛书抓住了当前评估领域的热点与前沿性问题。例如,在政府绩效评估方面,侧重于群众评议政府绩效的理论与方法。群众评议政府绩效是未来发展的趋势,有关这一问题的研究,即使在发达国家,也是一个相对较新的领域。而在国内,这可能也是第一部专门阐述群众评议政府绩效的著作。

第三,实用性。本丛书的另一个特点,是注重可操作性。有的著作不仅阐述了评估的思想、理论,也尽量给出具体的评估方法、解决评估难题的方案,有的还列出了一些具体的评估指标体系与评估问卷,有的还给出了实际的评估案例及案例分析。因此,这套丛书不仅适宜于公共管理领域的教师、MPA 学员使用,也特别适宜公共部门的管理者与实践者参考。

本丛书的另一特色在于它是两岸三地公共管理领域的学者共同努力的结果。其实,我个人一直都很看重香港和台湾地区公共管理的理论与实践经验。西方的一些公共管理理论虽然很前沿,但有些可能不适于我国,或我们暂时还做不到。而香港和台湾地区与祖国大陆的情况很相似。特别是台湾地区的学者较早到西方学习公共管理,20 世纪 80 年代后陆续回台湾,他们较早将西方公共管理理论应用于台湾地区的实践,并将其本土化。无论如何,香港和台湾地区有很多经验值得我们学习,更有许多教训值得我们吸取。因此,这套丛书主要由祖国大陆、香港和台湾地区从事公共组织评估研究的学者和实务工作者共同撰写而成。

最后,我要特别感谢台湾智胜文化事业有限公司副总经理陈春燕女士。几年来,春燕女士多次来清华大学,不仅为我带来了台湾地区公共管理学界的最新书籍,而且帮助我们与台湾地区公共管理学界牵线搭桥,对我们的研究起到了很大的帮助作用。本套丛书的出版也离不开她的大力支持。

在中国,公共管理领域的评估研究还是一个初生的婴儿,本套丛书也一定存在许多不足之处,还望广大读者不吝指正。

邓国胜

2006 年 9 月于清华大学

原著序言

当我们尝试进行正式的评估时,各种不同的概念会浮上心头,例如:价值、优点、值得、成长、准则、标准、目标、需求、规范、委托人、评估报告阅读者、效度、信度、客观性、实务的意义、绩效问责、改善、输入、过程、产出、形成性的、总结性的、成本、冲击、信息、可信度,以及评估本身一词。当评估者与其委托人及同事沟通时,他/她除了要清楚这些概念的意义外,也需将这些概念及其意义整合为一个具体架构,以便引导评估者进行评估。

评估的概念化不是仅实施一次就结束的活动,也不是静态的,而是与评估领域的理论与实务共同成长、齐头并进的。此外,任何特定研究的设计与实施皆会包含对正在进行工作的许多想法,因此我们有必要确认及定义评估报告阅读者与信息的需求、评估对象、评估目的、探究步骤、需检视的关心事项与议题、待评估的变项、解释发现(译者注:即评估结果)的基础以及评估工作品质的标准。

因此,尝试将评估概念化,在快速增多的文献中成为最有影响力论著的主题之一,已不足为奇。而这些作品的内容包含丰富的理论观点,并在许多方面呈现差异,因此评估

作品具有复杂的特性一点也不令人惊讶。因为实施评估时必须面对各种各样的情境与政治脉络,是服务导向的,以及每位作者都有不同的背景与信仰。评估概念化的方式将依据评估过程目标的角色、希望呈现收敛的或分歧的研究发现的程度、对建构主义者或客观主义者结果发现及阐释的偏好、使用或不用实验控制、运用理论来决定变项及检视变项间关系的程度,以及硬性数据或软性数据在获得结论时所扮演的角色,而有所差异。因此,评估者在某一评估上遵循某种研究途径,在其他情境下又应用其他的研究途径是可以理解的。因为评估会面对不同的情境与哲学观点,幸运的是,评估者可从文献中发现评估过程各种概念化的方法,并找出一种最适合特定情境的研究方法。

然而,这种多样化的评估概念途径,使得寻求能支持所有评估原则的共识慢慢浮现。这项共识透过教育评估标准联合委员会(the Joint Committee on Standards for Educational Evaluation)提出的标准而具体化。基本上,这些标准要求评估应具有效用性、可行性、伦理性及精确性。这些标准及定期检视与修正标准相关机制的出现,意味着评估已发展成一门专业。当这些标准应用在北美教育评估上的同时,教育以外的领域及世界上其他国家也应用或至少参考过这些标准。

本书是1983年选辑的修订版,文章的选录有两项重要的考量。首先,保留下来的篇章是因为它们在评估历史上有里程碑的地位,且能适当地描述评估的模式,而所删除的章节是因为它们呈现的信息已经过时。其次,修订版新增自1983年以后评估发展的章节,所增加的素材取自于教育以外的领域及自然主义(naturalistic)的评估。这些考量使本书维持原有的7章,修正了3章,以及新增了15章。

上述的编辑方针使本书及时反映评估观念的发展,特别是在方案评估上,并分成五篇。第一篇包括评估的历史、模式、隐喻、定义及可供选择或替代的途径(alternative approaches)。在第二、三、四篇的文章则呈现了当前评估的主要学术思想,这是由此领域的大师们所执笔。在第二篇,文章是以问题/方法取向加以分类,其中包含目标导向的评估、结果评估、测验在评估的角色、差异评估、实验设计、成本分析、澄清听证或司法评估、个案研究、批判

技术及理论为基础的评估。第三篇是阐述改进/绩效问责取向的研究途径,包含了消费者导向评估、决策导向评估及绩效责任评估。进入第四篇是有关社会议题导向(倡议)评估模式,包含了响应式评估、建构主义者评估、增能评估及民主审议评估。最后一篇,将描述三个跨越模式的主题,包括以实际利用为焦点的评估、评估的标准及后设评估的方法论。

就某种意义来说,本书所描述的评估模式并非测试既有理论的数学模式,而是记述每一模式的作者对于评估主要观念及结构的理想化看法,这些看法形成了作者达成辩护性的描述与判断的指导方针的基础。我们知道此领域的一些作者反对评估模式有一对应的替选观点,然而我们认为诸如信念或信仰的其他名称,可能只会使读者更为困惑。我们很乐意在此描述所谓的概念化,不是发生了什么模式,而是将模式视为依照不同作者对于评估的信仰所进行的研究。在此观念下,模式是有关评估实施中如何挑出及陈述问题的理想化或示范式的观点。

我们希望强调的是,模式不应被视为分离的选项,虽然各个模式在主要的观点上可能有很大的差异,诸如对目标的处理及实验控制的使用,但它们仍有重叠之处,例如,所有的模式都要求结果的检视,而且大部分的模式也包括过程的检视。在 Scriven、Stake 及 Stufflebeam 所提出的模式中,我们可以看到这些互有重叠的明显例子。它们都强调对适当的准则,进行综合评估的重要性,以及提出对方案与其他对象的优点判断。然而,这些模式在一些显著的方面上是有差异的,例如,改进取向与总结性判断就有相对重要性的不同,这些重叠模式的实务意义是:当使用者设计特定的评估时,他们可以结合不同模式的要素。

我们对本书各章的作者怀有无尽的感激。首先要感谢授权我们重印其著作的作者,以及特别为本书撰文的作者。我们也很感谢 Kluwer-Nijhoff 出版公司的 Zachary Rolnik 和 Michael Williams 对我们的全力支持。接下来要感谢 Seamus ó hUallacháin、Brian Carnell、Marguerite Clarke、John Coyle、Ida Holmstedt、Catherine Horn、Diane Joyce、Amandine Passot、Sally Veeder、Hilary Walshe 及 Lori Wingate 在编辑上、技术上及行政上卓越的能力,以有效协助出版的工作。

我们深信本书的出版将会吸引而且有助于实际从事评估的工作者,其中包括委托评估研究并应用成果的委托者、评估者以及评估方案的行政主管及工作人员。我们同样深信本书可以作为方案评估及教学课程的教科书。而且,对于实际参与评估工作的人而言,他们将会证明本书是极为宝贵的参考资料。我们希望本书能明显地增进实际参与评估的工作者所能了解评估的复杂性,欣赏评估的替选观点,增强从文献中撷取理论建议的能力,具有对于各种评估途径批判性评价的能力,以及遵守此领域的专业标准,最后,能够改进评估的品质及效用。

目录

▶	第一篇·方案评估:概论	1
	第一章 方案评估:历史的回顾	3
	第二章 评估的模式、隐喻及定义	23
	第三章 21 世纪方案评估的基本模式	39
▶	第二篇·问题/方法取向的评估模式	99
	第四章 方案评估的理论基础	101
	第五章 结果评估	113
	第六章 测验在评估中的角色	133
	第七章 差异评估模式	149
	第八章 学校实务评估中田野试验的角色:一种特殊的设计	169
	第九章 改善教育决策与评估的成本分析	187
	第十章 澄清听证会:对其程序的观点	199
	第十一章 个案评估研究:进步的十年?	215

第十二章	教育批判:一种质性研究方法	227
第十三章	方案理论:重在如何执行而非是否可行	243
▶	第三篇·改进/绩效问责取向的评估模式	271
第十四章	以评估改进课程	273
第十五章	评估的各种意识形态	289
第十六章	CIPP 评估模式	321
第十七章	绩效问责:对州与地方决策者的政策意涵	369
▶	第四篇·社会议题导向(倡议)的模式	397
第十八章	强调响应式的方案评估	399
第十九章	自然主义探究法的知识论和方法论基础	423
第二十章	发展评估的论述	445
第二十一章	赋权评估的步骤:从加州到凯普镇	461
第二十二章	审议的民主评估实务	477
▶	第五篇·跨模式的主题	495
第二十三章	以实用为导向的评估	497
第二十四章	评估专业标准与原则	515
第二十五章	后设评估方法论	533
	参考文献	551
	索引	601

第一篇

方案评估：概论

- ▶ 第一章 方案评估：历史的回顾
- 第二章 评估的模式、隐喻及定义
- 第三章 21世纪方案评估的基本模式