

Evaluation

# 评估学

【德】赖因哈德·施托克曼 沃尔夫冈·梅耶 著

唐以志 译



人民教育出版社

# 评估学

主编：陈德胜、曹晓波 副主编：陈永成、陈超 曹  
晓波、曹

中国社会科学出版社



# 评估学

【德】赖因哈德·施托克曼 沃尔夫冈·梅耶 著

唐以志 译

 人 民 出 版 社

责任编辑：刘可扬  
封面设计：徐 晖  
版式设计：东昌文化  
责任校对：周 昕

### 图书在版编目（CIP）数据

评估学 / [德] 赖因哈德·施托克曼, [德] 沃尔夫冈·梅耶 著;  
唐以志 译. —北京: 人民出版社, 2012.11

ISBN 978-7-01-010543-7

I. ①评… II. ①施… ②梅… ③唐… III. ①评估方法  
IV. ①C931.

中国版本图书馆 CIP 数据核字 (2011) 第 276338 号

## 评 估 学

PINGGUXUE

[德] 赖因哈德·施托克曼 [德] 沃尔夫冈·梅耶 著  
唐以志 译

人民出版社 出版发行  
(100706 北京市东城区隆福寺街 99 号)

环球印刷(北京)有限公司印刷 新华书店经销

2012 年 11 月第 1 版 2012 年 11 月北京第 1 次印刷

开本: 880 毫米 × 1230 毫米 1/32 印张: 13.25

字数: 290 千字 印数: 0,001 - 3,000 册

ISBN 978-7-01-010543-7 定价: 39.00 元

邮购地址 100706 北京市东城区隆福寺街 99 号

人民东方图书销售中心 电话 (010) 65250042 65289539

版权所有·侵权必究

凡购买本社图书, 如有印制质量问题, 我社负责调换。

服务电话: (010) 65250042

# 前 言

赖因哈德·施托克曼

( Reinhard Stockmann )

这是一本介绍评估的书。在过去一段时间里，评估经历了一个快速成长的时期。最迟从“国际学生评估项目（PISA）”研究开始，在德语词汇中已经可以发现评估的概念了。即使在日常生活中，评估也占有一席之地，比如在《明星》周刊系列文章《爱的世界》中就写道：“女人总会精准地评估与谁交往。”（《明星》2007年第32期，第102页）

评估的概念不仅完全被滥用，而且在很多情况下使用评估概念的时候至少也没有探究其科学性，因为不是每一种评价都是评估。

当一些来自于日常生活的概念，譬如社会群体、能力、资质、机构、体系被有意识地引入科学语言，而且必须从日常生活语言的理解中加以提炼并用科学的含义加以精确解释时，评估的概念却走向了相反的道路。在某种条件下，对任何一个对象使用生活语言都会造成概念科学性和精确度的降低，因此，

## 2 评估学

现在对评估概念的范围进行重新界定就非常必要。

这不仅适用于评估的概念，而且同样也适用于那些借助评估概念来表述的评估方案。借助评估的概念，这些评估方案就与其他一些诸如鉴定、测验、成果监控或者各种形式的研究这些通用的方法严格区分开来。由于现在评估的概念很新潮，听起来似乎也很现代化、很科学，因而它被广泛地应用于行动策略中，而这些行动策略在设计的时候绝对不会朝着评估的方向来考虑。因此，人们需要很好地咨询和准确地检验，他那里是否产生并需要进行评估。

当然，评估的作用不仅被夸大，而且在不同情形下它总是被赋予相应的科学和政治意义。即便到今天（正如本书第一章将要阐述的），在德国几乎还没有人谈到评估文化，因此，评估还没有成为社会普遍接受的、在政治进程中系统使用并定期运用于理性决策的工作方式，很显然，它还只是处于起步阶段。在此期间，评估只是作为一种重要的调控工具：例如政府部门用于项目效果的检查、行政当局用于绩效考核、审计署用于支出项目的监督、政府和非政府组织用于项目的调控、企业用于质量管理的支持，等等。在一些政治领域，如教育和科研领域有一些永久建立的政府研究机构和下属机关，如劳动力市场与职业研究所或者联邦职业教育研究所，它们在很大程度上承担了这些评估任务。在更深层次的政治领域，如发展合作，由政府主管部门内部的评估人员或受托人负责对实施机构及其他援助机构进行评估，其中大部分评估是由自己内部实施的，还有一些是委托鉴定、咨询和研究机构来进行。在其他一些政治领域，比如外交、安全和文化政策，内务、农业和环境政

策，更确切地说，评估是隐秘进行的。

在此背景下，不仅厘清评估概念所要论述的内容以及如何将评估概念与其他概念加以区分很重要，而且弄清楚评估在社会中所扮演的角色也很重要。本书将不仅关注这些主题，而且还会阐述评估的效能，同时还会对评估的专业化实施和评估运用的科学依据进行论述。

因此，本书首先将对评估在现代社会中的角色进行分析（第一章）。正如前面已经简要介绍过的，评估的运用五花八门，因此有必要根据评估在社会中的不同作用将其进行系统分类。在接下来的内容中将对评估的历史发展、评估在各种实践中的特征、评估在未来面临的挑战等进行简要论述。与此同时，评估还面临着供给与需求之间的矛盾：一方面，评估的含义扩展了，对评估的需求也日益增长；另一方面由于在德国、欧洲，特别是大多数非洲、亚洲和拉丁美洲还只有很少的评估培训和继续教育机会，能提供的专业化评估仍然很有限。世界经济危机又使评估变成热门，因为当公共财政变得捉襟见肘的时候，提高项目和措施的有效性、效果和效率，为理性决策提供依据就显得十分重要，只有借助这些标准才能制定可行的优先选择及淘汰程序。而要解决这些问题，评估可以发挥很大的作用。

因此，必需尽可能地重视评估的专业化标准和科学基础，而这一点并不是很容易就可以做到的，因为评估不仅仅是经验社会学的一部分因而要按照经验社会学的多数规范来进行，而且它还是深受政治影响的职业范畴。首先，评估是一种委托研究，在多数情况下，谁付费，谁就会有决定权。评估本身不一

定会研究评估的结果，但评估最少会研究问题提出的方式。评估不仅需要通过它的科学性而且还要通过其有用性来加以定义，因此，必须随时找到两者之间的平衡点。本书的第二章将探讨评估与政治之间、应用科学与基础科学之间的对立关系，并介绍一种能满足不同需求条件下进行评估的研究模型，其中主要是对施托克曼确立的“评估中心（CEval）”模式进行阐述。这一模式是建立在不同的科学理论假说基础之上，原则上讲，它可以满足在各种政治领域和项目阶段开展广泛并有理论指导的评估。这一模式首先是从实践出发来阐述评估的科学依据，在实践中每一项评估都必须回答评估什么（评估有哪些对象）、为什么评估（评估有些哪些目的）、依据哪些评价标准、由谁进行评估（内部评估还是外部评估）、如何评估（用什么方法评估）这些问题。

评估模式看起来似乎给出了评估者所关注的一切。但理论模型和评估模式之间还存在着很多不明确的问题，其中至少还存在着一个带有普遍性问题，即它还没有涉及理论的严谨使用。第三章尝试对各种概念进行分类。在提出我们自己的分类模式之前介绍了各种系统的分类。评估有各种结构形态，如按其历史发展，古巴（Guba）和林肯（Lincoln）对评估进行了“时代划分”；或者是阿尔金（Alkin）的树状形态，在这种形态中，一些特有的模式来源于共同的“树干”，但分裂为不断精细分权（分支）的强有力的“树枝”。还有另外一种完全不同的由菲茨帕特里克（Fitzpatrick）、桑德斯（Sanders）和沃尔森（Worthen）进行的系统化分类，他们按照评估的运用进行分类。各种系统化的分类都有其优缺点，没有哪一种被证明



是最好的。因此在第三章我们进行了系统化分类的尝试，在这一系统化分类中运用了第二章提出的评估的四种不同功能。

尽可能广泛地了解各种不同的评估模式有利于从各种复杂的问题中甄别出恰当的模式，也就是说，没有必要为每一次的评估都开发一种新的模式来丰富评估方案的多样性，而是要尽最大可能改编已有的模式使其能更有效、更经济地使用。

通过这一工作步骤人们就可以发现他已经处于评估的过程规划中了。第四章将以流程图的形式来描述评估的规划、实施以及评估结果的应用。本章还将对评估的目的、评估方案、评估手段的开发、数据的收集和分析、报告撰写及评估结果陈述等整个评估周期的行动步骤进行论述，而本书作者长期以来开展各种评估所积累的经验也将体现在相应的内容中。为了确保评估的质量、有用性、可实施性和公正的操作程序，正如德国评估学会（DeGEval）和其他一些协会组织在它们的标准中所要求的，让与评估有关的人员主动参与进来是有很多好处的，因此，第四章介绍了评估中心（CEval）的参与式评估模式，这一模式将科学本位的评估理念（第二章第三节）与参与式方法很好地结合起来。这一模式的特别之处就在于，它努力实现参与者的最佳组合但又不损害科学的标准。为此，评估者和参与者（包括被评估者）被当作合作伙伴在评估过程中承担互补性的任务，这样就可以将专家的知识与专业性和具体的情境知识（Situationswissen）结合起来。

当参与者能广泛地参与到规划和运用阶段时（取决于参与的程度），调查设计及调查工具的开发、数据调查和分析的实施则是由有资质的专家（也就是评估员）来负责的。第五章列

出了信息收集和评价的四个时间顺序阶段，首先是分析了相关的调查设计特别是根据影响评估进行调查设计的优缺点，重点关注了由评估所带来的一些特别的情形。在数据调查手段设计及应用的过程中，必须根据调查情形的变化灵活地进行调整，这一点通常在社会科学的基础研究中也是必要的。对调查数据的分析也需要毫无保留地遵循经验社会学研究的规则 and 标准。

对数据的解释必须特别注意评估的背景条件，因为在评估中进行评价必须严格地使用事先确定的标准、遵守客观的工作准则。尽可能不要让评估者、任务委托方、利益相关者的主观态度影响到评价。为了保证评估满足有用性和科学性这一双重要求，一方面（如第四章所描述的）有必要让利益相关者广泛地参加到评估过程中所谓的发现和应用过程；另一方面，在与研究有关的阶段，则特别要求评估者具有科学的专业经验。由于研究问题的复杂性，特别是评估所处背景的复杂性（如时间压力、资源不足、利益相关者的各种不同观点、兴趣和期望），因而要求每位评估者熟悉各种研究设计、质的和量的调查方法、统计评价方法，因为评估者是唯一一位能确保将研究问题尽可能与测量有机结合在一起的人。

在政治与科学对立的区域，要保证高度的科学性就特别要求评估者具备很强的与不同社会群体交往并了解他们利益的社会能力。在选择评估模式（第三章）、构建流程、进行研究和调查设计（第五章）时，必须对环境，也就是一项评估特殊的社会背景给予足够的重视。第六章从三个角度对评估的环境进行了探讨。首先是将评估的背景和社会制度作为研究的对象，正如在第一章论述过的，评估作为现代社会的一个重要组成部

分，有助于确保政策调控的合理化，有助于对社会进行宣传。考虑到社会宣传功能，要特别关注公共领域和公民社会的基础，既要有助于对政府行动的监督，也要有助于政治体系与公民之间利益的均衡。民主国家及其行政机关、媒体、公民社会成员之间的这种复杂关系经常会影响到评估的实施及其结果的使用。

参与评估的人不仅是自由作决策的个人，他们还是某些集体性组织的代表，他们的利益与社会某一阶层的利益紧密相关。一些组织在现代社会中扮演着重要的角色，他们借助于政党、社团和其他公民社会组织形式这样一些集体性组织来实现他们的利益诉求，他们首先会通过国家行政机关，但也会在一定条件下利用现代社会日益增长的复杂性以及基层的公民社会组织对政策调控施加越来越强的影响。这些组织委托进行的评估常常不是通过个人，而是通过其他机构来实施的，因此，必须特别关注有组织背景的评估。

这一点将会通过诸如组织的三个核心观点、个体与不同形式的成员构成的组织之间的关系、组织内部各调节功能之间的联系、组织形态结构内的执行单元等来说明，从而显示出组织的特征会对评估产生影响。要使组织内接受和支持评估，不仅参与者在组织中的责任很重要，而且组织内各部门之间的信息流、评估在各部门中所处的地位也很重要。

由于组织不仅要在固定的制度化的工作流程框架内遵循规则使命，而且还要实施有时间限制的措施、计划和项目等活动，因此，评估作为计划调控的重要组成部分起着重要的作用。为此，这种相互作用的关系被扩展到组织外部的群体，这

些群体或多或少参与了措施、计划或者是项目的实施，或者是从项目中受到影响。对这一部分多样化的社会群体，人们借用英语词汇“利益相关者（Stakeholder）”建立起了一个集合概念，比较贴切地表达了对评估产生影响的这些群体。

这一例子最终有助于评估者了解从事这一职业所要面对的评估的社会环境和所面临的各种挑战，有助于向他们说明从事这一职业所应具备的能力。评估绝非组织内有经验的实践者或经验社会学研究者希望通过这一委托研究获得额外收益的兼职工作。如果要扩大其社会功能，评估必须严谨和专业地进行。弄清楚评估的任务和特殊要求、介绍评估理念的现状及观点正是本书最主要的任务。

本书的完成不仅仅只是作者为此付出了劳动。我们要感谢评估中心（CEval）的同事们，他们为本书的写作不断给予鼓励和建议，没有他们就不会有写作的源泉。要特别感谢辛勤参与文稿编排和图表制作的工作人员，特别要感谢米里阿姆·格拉普（Miriam Grapp）女士、安格莉卡·伦特维希（Angelika Nentwig）女士，没有她们的支持，这本书将不可能完成。最后，作者还要感谢芭芭拉·布德里希（Barbara Budrich）女士在本书的写作过程所表现出的责任心和耐心。

# 目 录

前 言 .....	1
<b>第一章 评估在社会中的作用</b> .....	1
第一节 评估——现代社会的一项发明创造 .....	1
一、目的 .....	3
二、评估的制度化 .....	8
第二节 评估的历史发展 .....	11
一、评估在美国的发展 .....	12
二、评估在欧洲的发展 .....	17
三、评估在德国的发展 .....	21
第三节 评估的现状 .....	23
一、国际 .....	23
二、德国 .....	33
第四节 评估目的的实现及面临的挑战 .....	39
结 论 .....	51

<b>第二章 科学本位的评估</b> .....	54
<b>第一节 政治与科学之间的评估</b> .....	54
一、评估与政治 .....	54
二、评估与科学 .....	57
三、评估的研究模式 .....	60
<b>第二节 依据</b> .....	66
一、定义 .....	66
二、评估什么? ——评估的对象.....	71
三、为什么评估? ——目的和任务.....	79
四、根据哪些评价标准进行评估? .....	85
五、谁评估? .....	89
六、怎样评估? .....	93
<b>第三节 施托克曼的 CEval 评估理论</b> .....	102
一、生命历程模型 .....	102
二、组织模型 .....	105
三、创新—扩散模型 .....	108
四、可持续性 .....	110
<b>第三章 评估模式及其理论基础</b> .....	116
引 言.....	116
<b>第一节 古巴和林肯的分类体系: 时代模式</b> .....	117
一、第一代评估——测量 .....	119
二、第二代评估——描述 .....	122
三、第三代评估——判断 .....	124

四、第四代评估——协商 .....	128
五、总结与评价 .....	132
第二节 阿尔金的系统分类：树状模式 .....	132
一、方法分支 (“Methods”) .....	134
二、运用分支 (“Use”) .....	140
三、价值判断分支 (“Valuing”) .....	146
四、总结和评价 .....	150
第三节 运用模式 .....	152
一、目标导向模式 .....	153
二、管理导向评估模式 .....	157
三、顾客导向评估模式 .....	160
四、专家导向的评估模式 .....	163
五、参与式评估模式 .....	170
六、总结和评价 .....	177
第四节 施托克曼和梅耶的分类：功能模式 .....	179
一、与人相关的分类 .....	180
二、与政治领域相关的分类 .....	185
三、与专业领域相关的分类 .....	191
四、总结与评价 .....	197
第四章 评估过程 .....	199
引 言 .....	199
第一节 规划阶段 .....	201
一、明确评估意图、界定评估范围 .....	201
二、评估方案与进度计划 .....	210

## 4 评估学

第二节 实施阶段.....	213
第三节 运用阶段.....	220
一、结果的描述和运用 .....	220
二、评估的效益 .....	228
三、评估的质量 .....	231
第四节 CEval 的参与式评估模式.....	237
<b>第五章 信息收集与评价.....</b>	<b>242</b>
引 言.....	242
第一节 数据调查准备：从评估设计到调查设计.....	244
一、实验方法 .....	248
二、准实验方法 .....	251
三、质的方法 .....	255
第二节 数据调查的实施——从方法使用的规划到 质量控制.....	261
一、数据调查方法的分类 .....	262
二、干扰因素和调查误差 .....	267
第三节 数据分析——从准备到运用.....	275
一、数据处理的过程 .....	278
二、统计分析方法 .....	281
三、显著性检验和代表性 .....	286
第四节 数据解释——从素材到评价.....	288
一、数据解释作为一种质的过程 .....	289
二、解释方法的种类 .....	293
三、分析性评价 .....	298



第五节 总结.....	301
<b>第六章 评估环境.....</b>	<b>303</b>
引 言.....	303
第一节 评估与公共领域.....	306
第二节 评估与组织.....	313
一、组织的成员与评估 .....	315
二、组织内的交往与评估 .....	320
三、评估在组织中的结构定位 .....	324
第三节 评估与“利益相关者”.....	326
第四节 结论与总结.....	330
参考文献.....	338
附 录.....	391
一 专业术语对照表.....	391
二 人名对照表.....	399
译 后 记.....	403