

通向结果之路： 有效发展评价的 设计与实施

THE ROAD TO RESULTS

Designing and Conducting Effective Development Evaluations

[美] 琳达·G·莫拉·伊玛斯

雷·C·瑞斯特 / 著

李扣庆 / 等译



经济科学出版社
Economic Science Press

图字：01—2011—6912 号

The Road to Results: Designing and Conducting Effective Development Evaluations Copyright
© 2009 by The international Bank for Reconstruction and Development/The World Bank

This work was originally published by The World Bank in English, as The Road to Results: Designing and Conducting Effective Development Evaluation in 2009. This Chinese translation was arranged by Economic Science Press. Economic Science Press is responsible for the quality of the translation. In case of any discrepancies, the original language will govern.

该作品中所有的调查结果、解释及结论均为作者本人观点，并不影响世界银行或其代表的政府的执行长官的观点。

世界银行不确保作品中数据的准确性。作品中所涉及的地图中的边界、颜色、统治等其他信息并不代表世界银行的立场及判断。

© 2011 中文简体字版专有出版权属经济科学出版社
版权所有 翻版必究

图书在版编目 (CIP) 数据

通向结果之路：有效发展评价的设计与实施 / (美)
伊玛斯, (美) 瑞斯特著; 李扣庆等译. —北京: 经济
科学出版社, 2011. 12
ISBN 978 - 7 - 5141 - 1271 - 9

I. ①通… II. ①伊… ②瑞… ③石… III. ①经济评
价 - 研究 IV. ①F014.9

中国版本图书馆 CIP 数据核字 (2011) 第 237645 号

责任编辑: 白留杰 李 剑
责任校对: 刘欣欣
版式设计: 代小卫
技术编辑: 李 鹏

通向结果之路：有效发展评价的设计与实施

[美] 琳达·G·莫拉·伊玛斯

[美] 雷·C·瑞斯特 著

李扣庆 等译

经济科学出版社出版、发行 新华书店经销

社址: 北京市海淀区阜成路甲 28 号 邮编: 100142

教材分社: 88191354 发行部电话: 88191540

网址: [www. esp. com. cn](http://www.esp.com.cn)

电子邮件: bailiujie518@126.com

北京中科印刷有限公司印装

787 × 1092 16 开 27.25 印张 640000 字

2011 年 12 月第 1 版 2011 年 12 月第 1 次印刷

ISBN 978 - 7 - 5141 - 1271 - 9 定价: 56.00 元

(图书出现印装问题, 本社负责调换)

(版权所有 翻印必究)

前 言

发展评价是整个评价体系的一个亮点。它研究的是发展领域影响深远、纷繁复杂而又迫切需要解决的问题。事实上，在我们看来，对扶贫、全球化及其对贫困人口的影响、全球变暖给弱小国家带来的影响、全球金融体系的结构不平等以及帮助冲突后国家的战略等领域的评价，只是发展评价帮助我们应对这些紧迫问题的几个例子。

全球范围内，政府和组织面临的问责压力越来越大，他们要向内外部的利益相关者展示绩效和更富有成效的发展。在这样的背景下，发展评价逐渐成为系统地探讨和回答“那又怎么样？”问题的主要方法。仅仅证明一个人很忙是不够的，现在还要证明一个人的工作是否有成效。

发展评价也是一个急需探索的领域。政府的数据系统可能很欠缺或者根本就没有，出于政治意图的信息腐败层出不穷，大量信息缺失，对所提供信息的可信度没有保障，在此情况下，寻找评价方法的工作是苍白无力的。面对这些挑战，发展评价表现出灵活性、创新性和创造性方法，探寻有效的方法向民众、政府官员、援助方、民间团体和媒体提供关于政府计划是否起到作用的信息。

本书力图通过强化发展评价这一工具，帮助从总体上改进治理体系，特别是帮助改进政策与计划的制订和实施。评价可以成为一个非常有力的公共管理工具，帮助政府和组织改进行为方式和结果。评价的贡献绝不只限于公共部门，评价也可以成为那些致力于支持贫困地区发展的民间团体、非政府组织以及援助方的有力工具。

本书的材料源自国际发展评价培训项目（IPDET）。IPDET 是世界银行和卡尔顿大学（Carleton University）每年夏天共同举办的一个发展评价培训项目。它将世界各地的学员汇聚在一起，用一个月的时间共同学习发展评价。本书是 IPDET 核心课程的精华，希望借此能与更多对发展评价感兴趣的人们分享 IPDET 课程。

作为本书作者，我们要感谢那些给我们提出咨询建议、阅读本书手稿、提出批评意见以及鼓励我们坚持完成本书的人。还有一些人是我们在这里就

要感谢的，他们是：迈克尔·巴顿（Mickeal Patton），帕特里瓦·格拉原（Patrick Grasso），马丁·艾布拉姆斯（Martin Abrams），米尔·达贝尔斯坦（Miel Dabelstein），格雷格·杰克逊（Gregg Jackson），基因·斯威梅尔（Gene Swimmer）和南希·博都斯（Nancy Porteous）。以上各位都阅读了本书的一些章节并提出了宝贵意见。我们非常感谢他们为此付出的心血。世界银行出版办公室的圣地亚哥·庞波（Santiago Pombo）给了我们强有力的支持。他始终坚定地鼓励我们完成本书。

我们还要感谢卡尔顿大学两位与我们共同合作 IPDET 的朋友：凯伦·金斯伯格（Karen Ginsberg）和芭芭拉·莱文（Barbara Levine）。没有 IPDET 也就不会有这本书了。他们是非常棒的合作伙伴，也是非常好的朋友。

最后，我们还要感谢黛安·诺瓦克（Diane Schulz Novak），写作和修改书稿的整个过程中，她都给了我们很大的帮助。她的敬业、细心和娴熟的技能都是不可或缺的，没有她，我们也就不会写这篇序，更不会有后面的书了。

2

我们俩是三十多年的同事和朋友。1981 年我们在美国政府责任办公室首次见面，并开始合作共事。其后我们又在世界银行做了十多年的同事。合作得越久，友谊也越来越深厚。随着我们走向事业的顶峰，是的，我们现在终于可以向评价界献上我们的合作果实了。

Ray R. & Diane M. Novak

中文版序

能为此书的中文版作序，我们甚为高兴。这是在中国出版的关于发展评价的首部教材，内容全面，而且是中文版本，因而极具里程碑意义。

本书付梓问世，犹如向一个强盛民族打开了通向评价这一公共管理工具的大门。利益相关者，不论是内部的或是外部的，均可以借助这个工具衡量政府和组织对战略、政策、计划和项目的执行方法与效率，以系统地提高对政府和组织的问责，因而意义可谓重大。我们不能再像过去那样只是对资金的消耗额度和去向进行简单的记录，而是要进一步回答几个关键问题：“这些资金是否用在了最迫切的干预活动上？资金的利用效率如何？干预活动的效果怎样？干预活动能否持续？有没有产生预期影响？”

发展评价面临诸多困难：如数据缺乏，信息因政治目的而被歪曲，信息不透明，通过其他方式获取的信息不一定可靠等。面对重重困难，发展评价要在收集和提供干预措施有效性信息的方法上有所创新。

在发展领域，人们强烈呼吁，要突破捐助机构采用的那种衡量其资助活动成效的评价体系，应该采用一种能适用于各种类型干预活动的评价工具，以系统地评价政府和组织如何实现预期结果以及最终的效果。本书正是在此大背景下产生的。2001年，世界银行和卡尔顿大学联合推出了国际发展评价培训项目，简称 IPDET（参见 www.ipdet.org）。IPDET 由两周的发展评价基础课程和两周的独立研讨会组成。研讨会可对具体的发展评价议题进行深入探讨。自 2001 年起，IPDET 每年夏天在加拿大渥太华的卡尔顿大学举办，来自全球各地的参与者汇聚一堂，共同学习发展评价的理论与实务。作为 IPDET 的联合主任，我们将 IPDET 的核心课程材料进行扩充和改编，形成了英文版的《通向结果之路：有效发展评价的设计与实施》（2009）。我们的目标是既让那些参加 IPDET 的学员能有一本发展评价的教科书，也要让无法到卡尔顿大学参加 IPDET 的人以一种易于理解和消化的方式分享有关内容。只是截至目前，读者仅限于懂些英语的人。

尽管中国早已实施了很多评价活动，如政策分析、经济分析和管理研究、调查研究、项目竣工审核以及经验总结等，但相对来说，评价在中国还是个新现象。中国要加入世界贸易组织的愿望可能是促使中国引入评价的动力。中国于 2001 年 12 月成功入世，入世的要求之一是加强政府的透明度。2000 年，受中国科技部委托，中国国家科技评价中心对中国高科技研发计划（“863”计划）15 年来的活动进行了评估，并公布了评估结果。之后，中国对评价保持了浓厚的兴趣。2006 年，中国财政部国际司派代表参加了 IPDET 培训。中国国家科技评价中心继续开展许多重要的评价工作，与荷兰政策与业务评价局进行了首次科技项目的联合评价，并于 2004 年完成了项目评价报告。2005 年，我们

应邀对约 100 名来自中国国家科技评价中心和其他政府部门的官员进行发展评价培训。2006 年，中国第一次将系统性监测与评价内容列入其五年规划。2007 年，中国财政部与合作伙伴一起发起了一年两期的上海国际发展评价培训项目（简称 SHIPDET），对本国和本地区内的评价人员进行培训。这个至今仍在顺利开展的项目合作伙伴包括中国财政部、世界银行、亚洲开发银行和亚太财经与发展中心。在这些活动的推动下，评价已经在中国和整个亚太地区打下了坚实的基础。但是在此期间，中国一直缺少一本中文版本的评价教材。

中文版《通向结果之路》提供了一本发展干预评价的教材，可供政府官员、大学教师、培训师以及现在和将来的专业人士研读，因此，该书出版的意义不言自明。在亚太财经与发展中心副主任李扣庆先生的主持下，本书的翻译工作耗费了大量的心血，最终才得以问世。我们向李扣庆、吴宁沁、曹声容、赵敏等表示深深的谢意，感谢他们投入了大量的周末和假期时间来精心完成本书的翻译工作。中国财政部的同事也对此书出版做出贡献，我们在此一并致谢。

中国的评价能力提升之快，让人惊叹不已。我们很骄傲，也很荣幸，能为此贡献一份微薄之力，相信本书会进一步推动中国的评价事业不断前进。

引言

发展的分析框架、概念框架和政治框架日新月异。新的发展议程要求以更加开阔的视野解读部门和国家发展战略与政策。它强调在发展周期的所有阶段都要进行学习和持续反馈。

事实上，发展评价可以被看做是一种公共产品：

评价超越了单个组织的边界。一项好的评价研究可以给整个评价界带来积极的溢出效应。发展评价具有国际公共产品的特征。（Picciotto 和 Rist 1995，P. 23）

随着发展议程涉及面的不断扩大，复杂程度的不断增加，发展评价也在不断跟进以适应形势需要。随着人们对结果的重视和对千年发展目标的认同，评价人员正逐渐从传统的实施和产出导向的评价转向结果导向的评价。为全面应对各国所面临的挑战，发展领域关注的焦点已不再是项目，发展评价人员正在探索用于评估国家、部门、主题、政策甚至全球层面结果的方法。评价界认识到不仅要以全面而且要以合作的方式应对发展中国家挑战的重要性，开始重视合作伙伴关系，发展评价人员也越来越多地参与联合评价。这些联合评价有很多方面的优点，但也增加了发展评价的复杂度（OECD 2006）。而且，在这种情况下，评价人员发现越来越难以度量单个发展组织的绩效，界定单个组织的贡献也越来越难。

发展越来越复杂，对发展的要求越来越高，评价设计也越来越难。它需要运用多种错综复杂的技术和方法，并设定更高的标准以产生影响。

对新的评价方法和混合技巧的需求已经超出了经济学的范畴。迫在眉睫的问题，如气候变化，就需要能够评价可持续性的新方法。环境问题、多国影响、难以获得共有可比性的数据，以及无法预料事件的持续发生等，都需要复杂的、融合多种方法的评价方法。

大概没有哪个学科能够主导应对大多数人所遇到的多种挑战和满足大多数人希望的工作。既然没有这样的单一学科，就经常会出现跨越不同学科的现象，评价人员越来越敢于博采众长、兼收并蓄地使用各种社会科学工具（Picciotto 和 Rist 1995，第 169 页）。

评价能力建设——建立受过发展评价实践和方法训练的评价团队——是大多数发展中国家面临的挑战。发展中国家全国性评价协会以及区域性评价组织的兴起是发展评价界走向专业化所迈出的重要的第一步。发展评价人员找到了致力于满足其需求的国际专业组织——国际发展评价协会（IDEAS）。对发展评价能力建设同样起到积极推动作用的是大学中发展评价研究生课程的开设和地区性培训中心的设立。

本书旨在为发展评价能力建设提供一个工具，帮助发展评价人员思考和探索新的评价

体系，特别是设计和实施能够应对发展挑战的结果导向的评价。

国际发展评价培训项目（IPDET）由世界银行运营评价部（现独立评价局 IEG）于 2001 年创建。IEG 在世界银行学院的帮助下，与卡尔顿大学合作在加拿大渥太华举办了第一期培训。自 2003 年后，该项目已在十多个国家提供了为期一周或两周的 IPDET 定制课程。2007 年上海国际发展评价培训项目（SHIPDET）创建。

尽管 IPDET 持续不断地改进以适应不断变化的发展要求，但目标对象仍基本定位于从事或即将从事发展评价的人员。它努力为评价地方、全国、区域或全球层面的发展干预提供常用工具。它的目标群体主要是在双边或多边发展机构、发展中国家政府部门、非政府组织中从事评价工作的人员以及国会议员和私人顾问们。

IPDET 的总目标是提升参与者设计和实施有效的发展评价的知识、技巧和能力，以推动基于事实的决策制定。它通过十四个授课模块，概要地介绍了实施有效的发展干预评价的路径。

本书基于这些模块，又有所扩展。它全面讨论了发展评价人员面对的问题，并为他们完成发展评价提供指导。通过阅读本书，更多的从事发展评价工作的人将掌握能进行更具说服力的发展结果评价的基本工具。

目 录

引言	1
基本原理	1
第一章 发展评价简介	3
什么是评价?	3
评价学科的起源和历史	11
发展评价的沿革	16
发展评价的原则与标准	18
发展评价实例	23
小结	26
练习	26
名词解释	26
参考书目	27
网址	31
第二章 推动发展评价的主要因素	34
发达国家及发展中国家的评价活动概况	34
新出现的发展问题的意义	47
小结	61
名词解释	62
参考书目	62
网址	70
发展评价的准备与实施	71
第三章 建立结果导向的监测与评价体系	73
结果导向的监测与评价体系的重要性	73
什么是结果导向的监测与评价体系?	74

传统的监测与评价体系和结果导向的监测与评价体系	75
建立结果导向的监测与评价体系的十个步骤	78
小结	94
练习	95
名词解释	96
参考书目	96
网址	97
第四章 预期评价背景与项目变革理论	99
前期分析	99
确定主要客户和关键利益相关者	101
理解评价背景	104
发掘已有知识	105
变革理论的构建、使用和评估	106
小结	120
练习	120
名词解释	122
参考书目	122
网址	125
第五章 考虑评价方法	128
评价的一般方法	128
评价方法的优势与挑战	146
小结	150
练习	150
名词解释	150
参考书目	151
网址	155
设计与实施	159
第六章 评价问题及设计矩阵的开发	161
问题的来源	161
问题的类型	162
问题类型与变革理论的关系	167
识别和选择问题	167
开发合适的评价问题	169

评价设计	170
小结	176
练习	177
名词解释	177
参考书目	178
网址	178
第七章 因果性问题、描述性问题和规范性问题的设计选择	179
将问题与设计相关联	179
因果性问题设计	182
描述性问题设计	196
规范性问题设计	198
更严格的评价设计需求	199
练习	202
附表 不同设计类型的关键术语和特性概要	203
名词解释	204
参考书目	205
网址	209
第八章 选择和构建数据收集工具	210
数据收集策略	210
合理指标的特点	213
定量数据和定性数据	214
收集数据的工具	217
小结	253
练习	253
名词解释	254
参考书目	255
网址	259
第九章 抽样策略选择	260
抽样简介	260
样本类型：随机和非随机	260
确定样本大小	266
小结	268
练习	268
名词解释	269
参考书目	270

网址.....	271
第十章 计划和实施数据分析.....	272
数据分析策略.....	272
分析定性数据.....	273
定量数据分析.....	283
定性数据与定量数据的关联.....	292
小结.....	294
练习.....	295
名词解释.....	296
参考书目.....	297
网址.....	300
迎接挑战	303
第十一章 复杂干预评价	305
发展评价大局观.....	305
联合评价.....	306
国家项目评价.....	309
部门项目评价.....	311
主题评价.....	313
全球和地区合作项目评价.....	314
小结.....	317
名词解释.....	317
参考书目.....	318
网址.....	321
引领	323
第十二章 评价活动的管理	325
设计矩阵的管理.....	325
评价的承包.....	326
不同参与人员的角色和职责.....	328
人员、任务和预算的管理.....	332
小结.....	339
练习.....	339
名词解释.....	340

参考书目	341
网址	344
第十三章 结果的展示	346
精心设计沟通策略	346
撰写评价报告	348
形象地展示信息	353
口头演讲	362
小结	364
练习	365
名词解释	365
参考书目	365
网址	367
第十四章 指导评价人员：评价的道德、政治、标准和指导原则	368
道德行为	368
政治与评价	371
评价标准与指导原则	376
小结	379
练习	379
参考书目	380
网址	381
第十五章 展望未来	384
从过去到现在	384
展望未来	385
练习	392
参考书目	392
网址	394
附录一 OECD DAC 发展评价网络评估发展合作关键规范和标准概述	395
附录二 Outer Baldonia 中学职业培训项目的设计矩阵	409
附录三 缩略语表	415
后记	420

基本原理

“真正的天分表现为评估不确定的、危险的和相互矛盾信息的能力。”

——温斯顿·丘吉尔

第一章 发展评价简介

- 什么是评价？
- 评价学科的起源和历史
- 发展评价的沿革
- 发展评价的原则与标准
- 发展评价实例

第二章 推动发展评价的主要因素

- 发达国家及发展中国家的评价活动概况
- 新出现的发展问题的意义

第一章 发展评价简介

本章介绍了项目、计划和政策评价的定义和一般概念，进而讨论了发展干预的评价，即通常所称的发展评价。

本章包括五个主要部分：

- 什么是评价？
- 评价学科的起源和历史
- 发展评价的沿革
- 发展评价的原则与标准
- 发展评价实例

什么是评价？

理解评价的含义、目的及其用途，对理解发展评价是十分重要的。本章的这一部分将

- 对评价进行定义
- 确定评价的目的
- 列出评价的好处
- 指出评价人员评价什么
- 分析监测与评价的关系
- 明确职业评价人员的作用和活动

评价的定义

评价有多种定义。牛津英语词典将其定义为：

1. 一种鉴定或估价（物品等）的活动；价值的计算或声明；2. 估计或确定（数学表达式、物理量等的）数值，或是估计概率、证据等作用的行为。

在评价学科内，评价这一术语有多种含义。定义的差异反映了评价目的侧重点的不同，如问责与学习；也反映了与计划、项目或政策进展相关联的评价时间侧重点的不同。