

项目与政策评估 ——方法与应用

(第二版)

Evaluation in Practice
A Methodological Approach

[美] 理查德·D·宾厄姆

Richard D. Bingham

[美] 克莱尔·L·菲尔宾格

Claire L. Felbinger

朱春奎 杨国庆 等/译



Fudan Translation Series

复旦译丛 · 公 / 共 / 管 / 理 / 系 / 列 /

项目与政策评估

——方法与应用

(第二版)

Evaluation in Practice
A Methodological Approach

[美] 理查德·D·宾厄姆

Richard D. Bingham

[美] 克莱尔·L·菲尔宾格

Claire L. Felbinger

朱春奎 杨国庆 等/译

F062.4
32

复旦大学出版社

图书在版编目(CIP)数据

项目与政策评估——方法与应用(第二版)/[美]宾厄姆,[美]菲尔宾格著;朱春奎,杨国庆等译. —上海:复旦大学出版社,2008.4

(复旦译丛·公共管理系列)

书名原文:Evaluation in Practice: A Methodological Approach(Second Edition)

ISBN 978-7-309-05939-7

I. 项… II. ①宾…②菲…③朱…④杨… III. ①公共管理:项目管理-项目评价-研究②政策学-研究 IV. F062.4 D0

中国版本图书馆 CIP 数据核字(2008)第 021636 号

项目与政策评估——方法与应用(第二版)

[美]宾厄姆 [美]菲尔宾格 著 朱春奎 杨国庆 等译

出版发行 复旦大学出版社 上海市国权路 579 号 邮编 200433
86-21-65642857(门市零售)
86-21-65100562(团体订购) 86-21-65109143(外埠邮购)
fupnet@fudanpress.com http://www.fudanpress.com

责任编辑 张 炼
总 编 辑 高若海
出 品 人 贺圣遂

印 刷 浙江省临安市曙光印务有限公司
开 本 787×960 1/16
印 张 32.25
字 数 433 千
版 次 2008 年 4 月第一版第一次印刷
印 数 1—5 100

书 号 ISBN 978-7-309-05939-7/F·1353
定 价 65.00 元

如有印装质量问题,请向复旦大学出版社发行部调换。

版权所有 侵权必究

内 容 提 要

项目与政策评估是公共管理与政策研究的一个重要领域，也是公共管理与政策实践中的一个难点。《项目与政策评估——方法与应用》一书由导论、实验设计、准实验设计、反身设计、成本—收益和成本—效益分析、其他设计、自我演练、评估困境等八个部分组成。书中的论文为读者提供了多学科背景的大量项目与政策评估案例。书中部分章节运用了统计技术，但读者不必精通统计学。本书为高等学校教师提供了一本具有教材性质的评估方法专著，相关教师可以借助该书向那些需要参与评估实务的学生传授必须掌握的评估技术知识。同时，本书也是一本向初学者提供各种技术性参考资料的教材，可以直接为初学者学习与掌握评估实务的理论和方法提供指导与帮助。

译丛序

复旦版公共管理译丛在某种程度上是复旦版MPA系列教材的延伸,这是我们计划的复旦版公共管理系列丛书的一部分。

本系列译丛首批共有8本,它们是盖依·彼得斯的《美国的公共政策——承诺与执行》,小威廉·T·格姆雷和斯蒂芬·J·巴拉的《官僚机构与民主——责任与绩效》,理查德·D·宾厄姆和克莱尔·L·菲尔宾格的《项目与政策评估——方法与应用》,理查德·雷恩的《政府与企业——比较视角下的美国政治经济体制》,菲利普·库珀的《合同制治理——公共管理者面临的挑战与机遇》,科尼利厄斯·M·克温的《规则制定——政府部门如何制定法规与政策》,查尔斯·T·葛德塞尔的《为官僚制正名——一场公共行政的辩论》,以及多丽斯·A·格拉伯的《沟通的力量——公共组织信息管理》。

西方公共管理的著作已有不少在我国翻译出版。本译丛选择的首批书籍大致可以分三类。一是偏重理论方面的,想介绍一些重要问题上的不同观点。比如在对官僚制的批评铺天盖地席卷而来之时,《为官僚制正名》为我们展示了一种针锋相对的观点。二是偏重操作和技术方面的,比如《沟通的力量》、《项目与政策评估》、《规则制定》。正如我们所知,公共管理是一门实践性很强的学科。三是涉及政府改革和治理的,比如《合同制治理》、《政府与企业》。此外,重量级学者盖依·彼得斯的《美国的公共政策》无疑为我们理解美国的公共政策提供了非常有价值的资料和观点。当然,对本译丛8本书的选择并不表明我们都赞同作者的观点,我们只是想给读者提供一些对于公共管理

教学、研究和实践可能有参考价值的东西。

翻译通常被认为是件费力不讨好的事。但事实上,谁都知道自近代以来西方思想对中国社会发展的影响之深,无论是哪一方面。因为在现代化的进程中,中国在很多方面只是个后来者。今天,当我们以开放的心态面对这个世界的时候,吸取前人的智慧,分享人类社会的文明成果,从而推进自身的发展和进步,构成了我们一个合乎逻辑的选择。以公共管理这一学科而言,西方学者的贡献对于推进这一学科在中国的发展和进步所起的作用是有目共睹的。今天,我们还在相当程度上受惠于他们的贡献,感受着他们的智慧给我们的启迪。因此,将他们的一些优秀的成果介绍到国内,其意义是不言而喻的。正是在这个意义上,我们应该对译者表示感谢,他们的努力应该是中国公共管理学科发展史上的重要一章。

本译丛首批书籍的翻译由我组织。在整个过程中,得到了美国加州大学(洛杉矶校区)马国泉教授的一如既往的支持。参加本书翻译的是我院的一些年轻教师和博士、硕士研究生。这些才华横溢的年轻人在翻译过程中体现出来的对两种文字的理解和把握以及他们本身具有的专业素质已经达到了相当的水准。此外,我们的一些博士生和硕士生也在迅速成长。相信不用很久,他们在翻译方面就可以独当一面。译界后继有人,这也是令人非常高兴的。在这里,作为组织者,我要对马国泉教授和本译丛的译者表示感谢。同时还要感谢邬红伟编辑,他的敬业和尽责促成了本译丛首批书籍的出版。

竺乾威

于复旦大学国际关系与公共事务学院

2007年6月8日

致 谢

我们感谢以下个人、组织和出版商慷慨授权本书使用以下材料：

“Improving Cognitive Ability in Chronically Deprived Children,” by Harrison McKay, Leonardo Sinisterra, Arline McKay, Hernando Gomez, Pascuala Lloreda, reprinted with permission from *Science* 200, No. 4339 (21 April 1978): 270 - 278. Copyright 1978 by American Association for the Advancement of Science.

“Evaluation of a Multimethod Undergraduate Management Skills Development Program,” by Marian M. Extejt et al., in *Journal of Education for Business* (March/April 1996): 223 - 231. Reprinted with permission of the Helen Dwight Reid Educational Foundation. Published by Heldref Publications, 1319 18th St. N. W., Washington, D. C. 20036 - 1802. Copyright 1996.

“Community Posthospital Follow-up Services,” by Ann Solberg, in *Evaluation Review: A Journal of Applied Social Research* 7, No. 1 (February 1983): 96 - 109, copyright © 1983 by Sage Publications, Inc. Reprinted by permission of Sage Publications, Inc.

“Conservation Program Evaluations: The Control of Self-Selection Bias,” by Tim M. Newcomb, in *Evaluation Review: A Journal of Applied Social Research* 8, No. 3 (June 1984): 425 - 440, copyright © 1984 by Sage Publications, Inc. Reprinted by permission of Sage Publications, Inc.

“Regulatory Strategies for Workplace Injury Reduction: A Program Evaluation,” by Garrett E. Moran, in *Evaluation Review: A Journal of Applied Social Research* 9, No. 1 (February 1985): 21 - 33, copyright © 1985 by Sage Publications, Inc. Reprinted by permission of Sage Publications, Inc.

“The Effects of Early Education on Children’s Competence in Elementary School,” by Martha B. Bronson, Donald E. Pierson, and Terrence Tivnan, in *Evaluation Review: A Journal of Applied Social Research* 8, No. 5 (October 1984): 615 - 629, copyright © 1984 by Sage Publications, Inc. Reprinted by permission of Sage Publications, Inc.

“Nutrition Behavior Change: Outcomes of an Educational Approach,” by Patricia K. Edwards, Alan C. Acock, and Robert L. Johnson, in *Evaluation Review: A Journal of Applied Social Research* 9, No. 4 (August 1985): 441 - 459, copyright © 1985 by Sage Publications, Inc. Reprinted by permission of Sage Publications, Inc.

“A Little Pregnant: The Impact of Rent Control in San Francisco,” by Edward G. Goetz, in *Urban Affairs* 30, No. 4 (March 1995): 604 - 612, copyright © 1995 by Sage Publications, Inc. Reprinted by permission of Sage Publications, Inc.

“The Costs and Benefits of Title XX and Title XIX Family Planning Services in Texas,” by David Malitz, in *Evaluation Review: A Journal of Applied Social Research* 8, No. 4 (August 1984): 519 - 536, copyright © 1984 by Sage Publications, Inc. Reprinted by permission of Sage Publications, Inc.

“A Cost-Effectiveness Analysis of Three Staffing Models for the Delivery of Low-Risk Prenatal Care,” by Elaine A. Graveley and John H. Littlefield, in *American Journal of Public Health* 82, No. 2

(February 1992): 180 - 184. Reprinted with permission from the American Public Health Association.

“Attitude Change and Mental Hospital Experience,” by Jack M. Hicks and Fred E. Spaner, in *Journal of Abnormal and Social Psychology* 65, No. 2(1962): 112 - 120.

“How Effective Is Drug Abuse Resistance Education? A Meta-Analysis of Project DARE Outcome Evaluations,” by Susan T. Ennett, Nancy S. Tobler, Christopher L. Ringwalt, and Robert Flewelling, in *American Journal of Public Health* 84, No. 9 (September 1994): 1394 - 1401. Reprinted with permission from the American Public Health Association.

“Are Training Subsidies for Firms Effective? The Michigan Experience,” by Harry J. Holzer, Richard N. Block, Marcus Cheatham, and Jack H. Knott, in *Industrial and Labor Relations Review* 46, No. 4(July 1993): 625 - 636. Reprinted with permission from Cornell University.

“Changing the Geography of Opportunity by Expanding Residential Choice: Lessons from the Gautreaux Program,” by James E. Rosenbaum, in *Housing Policy Debate* 6, No. 1 (1995): 231 - 269. Copyright 1995 by Fannie Mae Foundation. Reprinted with permission from the Fannie Mae Foundation.

“Changes in Alcohol Consumption Resulting from the Elimination of Retail Wine Monopolies,” by Alexander C. Wagenaar and Harold D. Holder, reprinted with permission from *Journal of Studies on Alcohol* 56, No. 5 (September 1995): 566 - 572. Copyright by Alcohol Research Documentation, Inc., Rutgers Center of Alcohol Studies, Piscataway, NJ08854.

“Fifth-Year Report: Milwaukee Parental Choice Program,” by John F. Witte, Troy D. Sterr, and Christopher A. Thorn, typescript

(December 1995). Reprinted with permission from John F. Witte.

“The Effectiveness of School Choice in Milwaukee,” by Jay P. Greene, Paul E. Peterson, and Jiangtao Du, in *Learning from School Choice*, edited by Paul E. Peterson and Bryan C. Hassel (Brookings Institution Press, 1998). Reprinted with permission from the Brookings Institution.

前言

我们在给各种公共行政专业的研究生班讲授项目与政策评估课程的过程中,发现学生们在对他人完成的项目评估作出评论时都存在着严重困难。很多学生难以区分各种设计类型。如果一个研究者设计的方案与教科书上的方案有所不同,很多学生便不能对其分类识别。

此外,学生通常认为发表在一个学术或专业期刊上的论文就是完美的。他们不能洞察作者在方法问题上的得失。

我们认为评估课程是一门方法课程,它旨在让学生掌握项目与政策评估技术,以及在评估研究中所遇到的方法问题。对他人成果的评估是学习过程的一个部分——不仅仅是挑别人工作成果的错误,更重要的是培养一种对评估难点较强的领悟力。彼特·罗西与凯瑟琳·莱尔(Peter Rossi and Katharine Lyall)清晰地指出了这一点:

做一名评论家很容易,因为所有经验研究的论文都或多或少存在着纰漏。在一个研究中有太多的地方会犯错误,或者由于未预测到的事件发生而使方案的设计和执行为难以完美。我们认为,评论家有责任对所评论的文章做两方面的判断:一是,如果可能的话,有必要对两种错误进行区分,即哪些是严重的错误,哪些是不太严重的错误。诚然,这是一个判断问题。但我们确实认为,负责的评论者应该对那些尚未很好理解大量通过实验产生的备忘录、工作文件、分析材料、最终报告主要内容的人提供指导。二是,只有把由计划和判断不当而造成的错误和在事件发生过程中不能提前预知而造成的错误区分开来,

才是公正的评论。这在区分项目是蹩脚的还是运气不好时至关重要。对前者的批评是正当的,对后者则需要善意的理解(1978,412—413)。

我们对罗西和莱尔的评论做进一步的分析。根据 Webster 词典的解释,所谓评论就是对一些问题表达自己的逻辑思考,它类似于一个艺术工作或行为过程,评论包括对价值、真理、正确性的评价,对美或技术的欣赏或解释。(Webster's, 1950,627)

因此,作为评论者,既要有正面的评论也要有负面的评论。需要指出的是,研究者需要创造有意义的方法去克服评估设计的难题。本书的第8章便是实例。梯姆·纽考姆(Tim Newcomb)在其关于西雅图能源保护项目评估中设计了一个独特的比较组。作为“评论者”,我们为作者这些创造性的想法而高兴。

至此,《项目与政策评估》这本书的宗旨便清晰可见了,即对在现实的评估研究中所遇到的不同的研究设计的技术、主要的设计问题(称内部效度问题)进行阐述。从第5章到第16章,每一章都会对一个评估设计做一个简短的介绍,并用一个评估案例进行具体讲解。每一章后面都附有一个对该评估工作的点评。

所受限制:这本书并不包括所有形式的评估监督、评估评价和过程评估等等(例外情况会在第1章中做简要讨论)。我们只关心那些有重大影响的评估。但这并不意味着其他的评估功能不重要,只是本书的目标并非包罗万象。

《项目与政策评估》是为研究生专业学位项目和社会科学的通用研究生课程所写。尽管我们也为公共行政项目的研究生授课,但这本书的受众不仅仅局限于上述两类读者。本书中的论文取自多个学科以阐释评估的共性。本书提供多学科背景的大量政策和项目评估案例。重要的是形成一套评估方法而不是论文中提出的实际项目的具体领域。各章正文中所提供的材料在某种程度上受到我们对不同设计实际运用偏好的制约。在导论之后,我们为对方方案设计更具体的解释感兴趣的读者提供了一份清单,其内容是对于各章节的补充阅读材料。本书同样适用于公共政策课程、公共行政、

公共策划、城市研究、社会学、政治科学、刑事司法、教育学、社会福利、医疗等相关领域。

本书并没有假定读者的学历背景都在本科以上。书中部分章节运用了统计技术,但学生不必精通统计学。除非特殊说明,学生们可以假定书中运用的统计数据都是合理的。在《项目与政策评估》这本书中,重要的是设计方法,而不是统计。

本书的组织

本书由八个部分组成:导论、实验设计、准实验设计、反身设计、成本-收益和成本-效益分析、其他设计、自我演练、评估困境。第一部分包括四章。第1章导论讲述了过程评估、输出评估、结果评估。第1章假设:评估是一个连续的过程,特定项目中不同类型的评估适用于不同的时期、不同的对象。第2章介绍了四个基本的评估设计,并讨论了设计对效度的威胁。第3章阐述基本的测量方法并讨论了测量的信度和效度。第一部分的最后一章关注于两种当今十分流行的特定的测量范式,即绩效测量和标杆设定。

第二部分包括三章,讨论并阐述三个实验评估设计。这些评估设计的共同特点都是把实验对象随机分配到了实验组与控制组。这三章包括前-后测控制组设计、所罗门四组设计和单后测控制组设计。

第三部分阐释了准实验评估设计,由四章组成。这四章分别讨论了前-后测比较组(不是控制组)设计、回归-间断设计、间断的时间序列比较组设计和单后测比较组设计。

第四部分关注准实验评估设计,但又不同于第三部分。第四部分中各组仅仅相互比较。在这部分中阐述单前-后测设计和简单的时间序列设计两种评估设计。

第五部分包括其他设计的两个变动或者说是扩展,也正是这种变动和扩展的独特性才被写入本部分。它们是成本-收益分析和成本-效益分析。

第六部分同样包括了在基本设计上的独特变化和改进,即补丁设计和后设评鉴设计。

第七部分要求学生自我演练。该部分中的三章提供给学生三个有趣的政策评估项目的案例,但没有给出任何解释和评论,目的是让学生自己对论文进行评判并学以致用。

第八部分总结中包括三篇论文。该部分阐释评估的困境。其中的两篇论文在对父母教育选择计划进行评估时,所用的数据是一样的,但却各自得出了不同的结论。我们的同事切尔·陈·鲍文(Chieh Chen Bowen)就此发表了评论。

最后,我们要向给本书提供过帮助的人表示感谢。首先感谢克利夫兰州立大学的城市研究中心对我们提供的机构支持。特别感谢我们的一些同事给予的评论和指导,他们是南加利福尼亚大学的乔·胡利(Joe Wholey)、明湖休斯敦大学的罗格·杜兰(Roger Durand)、西南得克萨斯州立大学的帕崔克·谢尔德(Patricia Shields)、南部伊利诺伊州立大学的保罗·卡汉(Paul Culhane)的建议对本书的第二版起到了至关重要的作用。同时感谢美利坚大学的琼·参(Jane Chan)和罗宾·格拉克(Robin Gluck)两位同学帮助我们所做的大量手稿校对工作。最后,我们感谢恺瑟姆出版社的戴维·埃斯特宾(David Estrin)、鲍勃·格姆利(Bob Gormley)、凯瑟琳·米勒(Katharine Miller)和特德·鲍伦(Ted Bolen),感谢他们为我们提供的编辑和出版服务。当然最终的文责还是我们自己承担。

我们对这本书的第二版寄予厚望。我们坚信如果学生真正理解了书中的阅读材料和讨论,他们将能够对在他们今后的职业生涯中很可能遇到的大部分项目、政策评估或评估设计进行理性的评价。

参考文献

1. Rossi, Peter H., and Katharine C. Lyall. 1978. An Overview Evaluation of the NIT Experiment. In *Evaluation Studies Review Annual*,

Vol. 3 Beverly Hills, Calif. : Sage Publications, 1950.

2. *Webster's New International Dictionary*. 2d ed. Springfield, Mass. :
Merriam.

目 录

第一部分 导论 / 1

第 1 章 评估过程 / 3

第 2 章 评估设计 / 20

第 3 章 测量 / 43

第 4 章 绩效测量与标杆 / 62

第二部分 实验设计 / 75

第 5 章 前-后测控制组设计 / 77

改善长期贫困儿童的认知能力 / 80

哈里斯·迈凯、利奥纳多·辛尼斯坦拉、阿琳·迈凯、赫纳多·

戈门齐、帕斯柯拉·劳瑞德 / 80

点评 / 105

第 6 章 所罗门四组设计 / 109

大学生管理技能开发项目多元方法的评估 / 110

玛丽安·M·伊柯斯泰特、J·本杰明·福布斯、约纳坦·E·

史密斯 / 110

点评 / 128

第 7 章 单后测控制组设计 / 132

社区出院后续服务 / 134

安·索尔伯格 / 134

点评 / 146

第三部分 准实验设计 / 149

第8章 前-后测比较组设计 / 151

节约项目评估：自我选择偏差的控制 / 153

梯姆·M·纽考姆 / 153

点评 / 166

第9章 间断时间序列比较组设计 / 169

减少工作场所受伤害的规制战略 / 172

加莱特·E·莫兰 / 172

点评 / 185

第10章 单后测比较组设计 / 188

小学儿童能力早期培养的效果 / 190

马肖·B·布劳森、唐纳德·E·佩尔森、

泰勒斯·提维南 / 190

点评 / 203

第四部分 反身设计 / 207

第11章 单组前-后测设计 / 209

营养行为转变：一种教育方法的结果 / 210

帕崔希·K·爱德华兹、艾兰·C·埃考克、罗伯特·L·约翰

斯顿 / 210

点评 / 227

第12章 简单时间序列设计 / 230

小有成效：旧金山租金控制的影响 / 231

爱德华·G·高兹 / 231

点评 / 241