

“十五”国家重点图书出版规划项目

社会工作经典译丛 Social Work Classic Series

主编 隋玉杰 副主编 范燕宁

CENGAGE
Learning

项目评估

循证方法导论

第六版



Program Evaluation
An Introduction to an Evidence-Based Approach (Sixth Edition)

戴维·罗伊斯 (David Royse)

[美] 布鲁斯·A. 赛义 (Bruce A. Thyer) 著

德博拉·K. 帕吉特 (Deborah K. Padgett)

王海霞 王海洁 译

中国人民大学出版社

“十五”国家重点图书出版规划项目

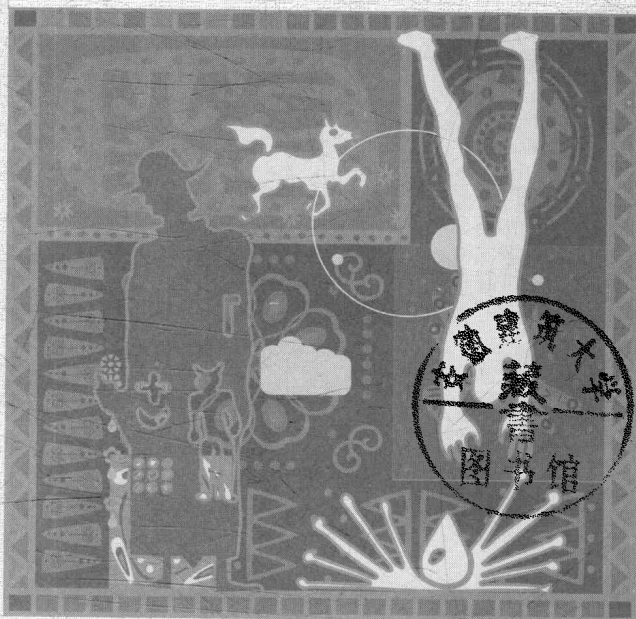
社会工作经典译丛 Social Work Classic Series

主编 隋玉杰 副主编 范燕宁

项目评估

循证方法导论

第六版



Program Evaluation
An Introduction to an Evidence-Based Approach (Sixth Edition)

戴维·罗伊斯 (David Royse)

[美] 布鲁斯·A.赛义 (Bruce A. Thyer) 著

德博拉·K.帕吉特 (Deborah K. Padgett)

王海霞 王海洁 译

中国人民大学出版社

· 北京 ·

图书在版编目 (CIP) 数据

项目评估: 循证方法导论: 第六版/(美) 戴维·罗伊斯等著; 王海霞等译. —北京: 中国人民大学出版社, 2018. 4

(社会工作经典译丛)

ISBN 978-7-300-25602-3

I. ①项… II. ①戴… ②王… III. ①项目评价 IV. ①F224.5

中国版本图书馆 CIP 数据核字 (2018) 第 032716 号

“十五”国家重点图书出版规划项目

社会工作经典译丛

主编 隋玉杰 副主编 范燕宁

项目评估——循证方法导论

第六版

戴维·罗伊斯 (David Royse)

[美] 布鲁斯·A. 赛义 (Bruce A. Thyer) 著

德博拉·K. 帕吉特 (Deborah K. Padgett)

王海霞 王海洁 译

Xiangmu Pinggu

出版发行 中国人民大学出版社

社 址 北京中关村大街 31 号

电 话 010-62511242 (总编室)

010-82501766 (邮购部)

010-62515195 (发行公司)

网 址 <http://www.crup.com.cn>

<http://www.ttrnet.com> (人大教研网)

经 销 新华书店

印 刷 北京七色印务有限公司

规 格 185mm×235mm 16 开本

印 张 24.5 插页 2

字 数 508 000

邮政编码 100080

010-62511770 (质管部)

010-62514148 (门市部)

010-62515275 (盗版举报)

版 次 2018 年 5 月第 1 版

印 次 2018 年 5 月第 1 次印刷

定 价 79.90 元

版权所有 侵权必究 印装差错 负责调换

主编简介

隋玉杰 女，香港大学哲学博士，中国人民大学社会与人口学院社会工作系副教授、主任。作为专家组成员参与了民政部及国家卫生健康委员会推动社会工作职业化、专业化的多项工作。主要研究领域为老年社会工作、心理健康与老年评估工具、社会工作职业化与专业化、社会工作教育等。兼任中国社会工作教育协会副秘书长、首届全国社会工作者职业水平评价专家委员会委员、西北民族大学客座教授。翻译出版了多部社会工作领域的经典著述。

副主编简介

范燕宁 女，首都师范大学政法学院社会学与社会工作系主任、教授。中国社会工作教育协会理事、中国人学研究会理事。主要研究方向为：社会发展理论、社会问题、社会工作价值观、矫正社会工作等。先后在《中国社会科学》《哲学研究》《哲学动态》《社会主义研究》《光明日报》等重要报刊发表研究论文数十篇。主要代表性论著有：《新时期中国发展观》《邓小平发展理论与科学发展观》《社会工作专业的历史发展与基础价值理念》《社区矫正的基本理念和适用意义》等。主要译著、译文有：《社会工作概论》（第九版）、《契合文化敏感性的社会工作课程》等。

社会工作正面临着前所未有的发展契机。

所谓契机，一是大的社会背景为社会工作的发展提供了舞台。随着改革的深入，中国在取得举世瞩目的成就的同时，如一些社会学家所言，也出现了“发展困境”的苗头或“类发展困境”的现象。新千年，政府在工作报告和政策文件中明确提出要关心弱势群体、加强就业和社会保障工作。与社会工作传统的工作对象，如贫困者、残疾人、妇女、儿童、老年人相关的一系列政策法规纷纷出台。这些都为开展社会工作提供了良好的政策环境。

二是社会工作专业本身已经步入组织化、规范化的轨道。中国社会工作者协会、中国社会工作教育协会等组织开始发挥行业指导和自律的作用。此外，经过多年的酝酿，2004年劳动和社会保障部制定的《社会工作者国家职业标准》在上海出台，明确了社会工作者的专业人员地位，一改多年来社会工作人员师出无名的状况，同时也为社会工作者在专业上不断发展提供了方向和路径。社会工作职业化、专业化有了突破性进展，在政府认可上迈出了坚实的一步。

进入新千年后，许多迹象表明，社会工作正在朝着进入新的发展时期的方向迈进。

然而，社会的需要和认可也给社会工作带来了挑战。社会工作是否已经拥有了完备的知识储备，成了一个羽翼丰满的专业，能发挥社会所期待的作用呢？

在今天，对许多中国的社会工作者来说，社会工作发展伊始弗莱克希纳提出的问题“社会工作是一个专业吗？”仍是个具有挑战性的问题。弗莱克希纳之所以断言社会工作不具备一个专业的资格，是因为他认为社会工作不是建立在科学知识的基础上。按照格林伍德提出的著名观点，成为一个专业应该具备五个特性：拥有自己的理论体系、具有权威性、得到社会的认可、有专门的伦理守则以及专业文化。其中排在第一位的就是专业知识的建构。

应当说，自1986年国家教育委员会同意北京大学、中国人民大学、吉林大学等高校设置社会工作与管理专业以来，中国社会工作理论与实务知识的建构已经有了可喜的收获。然而，在总体上，社会工作的专门知识仍然十分匮乏，对国外的社会工作仍缺乏系统的介绍，而本土的理论仍未形成。拿知识建构的领军团体社会工作教育界来说，情况也不容乐观。社会工作教育协会开展的中国社会工作教育发展状况调查的结果表明，以在学

术期刊上公开发表论文的数量、出版专著数、编写教材数、承担课题数等数据来衡量，社会工作教育院校教师的科研情况总体上水平不高。面对这一形势的却是社会工作教育在经过十几年的缓慢发展后，在世纪之交进入了高速扩张期。据中国社会工作教育协会统计的数据，截至2000年，协会的团体会员只有32个，到2003年12月已经达到148个。近80%的会员都是在2000年之后的三年新加入的。于是有了这样的景象，一方面是知识提供和传输上的不足；另一方面是嗷嗷待哺的跨入社会工作之门的莘莘学子。这便有了策划和出版社会工作经典译著的最初动因。我们希望通过这一系列书籍能够较为全面地介绍在西方已有上百年历史的社会工作专业的核心知识，为建立中国自己的社会工作知识体系作参考。

在整体结构上，“社会工作经典译丛”由三类书籍构成，即社会工作的基础理论、社会工作的基本方法和社会工作的价值观。这也是基于对社会工作知识体系构成的基本共识。具体来讲策划这套书主要有以下几点考量：

其一，完整性。整个译丛力图完整地呈现社会工作作为一个学科的全貌。译丛精选了社会工作理论、人类行为与社会环境、社会政策、个案工作、小组工作、社区工作、社会工作督导、社会工作研究和社会工作伦理等方面的书籍，全面涵盖了社会工作专业知识的三大组成部分，即基础理论、工作方法和价值观。考虑到价值观方面的教学一直是专业教育中非常重要的一部分，但也是专业教育中的难点，所以本套丛书特别精选了再版7次的专门用来帮助学生认识伦理问题和困境，并适当加以处理的有关社会工作伦理的专著。其中涉及的保密原则和隐私权问题、当事人的知情权和自决权问题、临终关怀问题、艾滋病问题等在中国的社会工作实践中已经出现，由于处理不当而引发的争端和法律诉讼也曾见诸报端。这方面的论述相信不仅对于社会工作学生，对于社会工作从业人员也不无借鉴作用。

其二，经典性。所选书籍都是广受好评的教材或论著，对社会工作的知识有精到的描述和评说。作者都是各自领域的专家和知名学者，有着丰厚的积累，在书中详细展现了与所述主题相关的专业知识，特别是融合了许多最新研究成果和实务动态，对读者来说极具参考价值。这些书在许多国家都被社会工作教育者采用。几乎每本书都是再版过多次，最多的再版次数已达9次。经过了使用者的检验和编写者的不断完善，这些书非常适合做社会工作专业教学的配套教材使用。

其三，適切性。为了能更好地配合教育部高等教育司组织制定的对社会工作专业主干课程教学的基本要求，译丛所选择的书籍基本都是社会工作专业主干课程的教材或论著。各书的框架也多与国内教学所要求的主体结构相契合，更能配合教学用途。

其四，实用性。一方面，所选书籍在内容的编排上注重方便读者使用。受以实证为本的工作方法的影响，大部分书籍都穿插了与所涉及内容相关的研究结果和案例讲解，将理论与实践相结合。在语言上也大多深入浅出，贴近读者，减少了他们在消化吸收知识上的障碍。另一方面，书籍所涉及的内容也多是国内社会工作界涉足和关心的领域。如通才社会工作实务模式，操作层面的社会工作方法，社会政策的研究、分析与应用，身为社会工作

教育和高层次管理人员开展督导的方法等。书中推荐的一些专业网站更可以帮助读者找寻更多的资源,丰富对书中相关内容的理解和把握。

其五,时代性。丛书中的每本书都是近两年来的最新版本,书中的内容涉及社会工作实务领域的一些最新发展,整套书如同一个多棱镜折射出社会工作学科的发展现状。大到社会福利体制管理上的变革,小至一些新的工作方法的使用,都有鲜明的时代特点。比如其中谈到的管理型卫生保健制度,个案管理,基因技术对社会工作的影响,网络技术对社会工作的影响,以实证为本的实践,私人执业,充实生活性质的社会工作等。一些实验性的工作方案在书中也有所介绍。这些无疑会拓展读者的视野。

2003年的一场“非典”像是对整个社会运行机制的一次检测,留下了许多宏观层面的问题,有待社会工作者去思考和解决。比如,社会危机处理机制、弱势群体保障机制、社会捐赠机制、基层社区的疾病预防和康复机制、志愿者的动员与使用机制等。而2004年的马加爵杀人案则给开展微观层面的社会工作提出了许多问题。比如,如何更有效地建立个人的社会支持系统、如何筛查处于危机边缘的人、如何提供更有效的危机防范与干预方法等。

德国著名哲学家恩斯特·卡西尔在《人论》中说:“当领悟了一门外语的‘神韵’时,我们总会有这样的感觉:似乎进入了一个新的世界,一个有着它自己的理智结构的世界。这就像在异国进行一次有重大发现的远航,其中最大的收获就是学会了以一种新的眼光来看待我们自己的母语。”歌德也说过:“谁不懂得外国语,谁也就不了解本国语。”我们希望“社会工作经典译丛”的面世能起到这样的作用,让读者能有一次异国社会工作之旅,看到社会工作在专业发展比较成熟的国度里的情况。虽然译丛中谈到的都是美国社会工作的状况以及他们的问题与处理方法,但对我们反观自身,处理中国的问题应当说不无启示。

译丛的策划得到了中国人民大学出版社潘宇博士,首都师范大学教授、博士生导师范燕宁和中华女子学院教授刘梦的鼎力相助。在甄选书籍的过程中,笔者同她们进行了反复的讨论,最后确定的书目是笔者与她们共同斟酌的结果。丛书的译者队伍也都是各高校的教师,有较丰富的社会工作专业积累,为翻译质量提供了保证。在此对上述参与本丛书策划和翻译等工作的人员一并表示衷心感谢。

虽然参与本丛书的人都倾尽了心力,但仍难免挂一漏万,希望广大读者对不当之处能给予指正。

隋玉杰

2004年10月14日

一套好教材所提供的信息，首先必须能够帮助读者更好地理解主题。这些信息必须是可信、易于访问和易于消化的。这本书的读者可以从文献中找到许多例子来说明要点和概念。你会发现一个新的重点，那就是循证实践。我们不断强调在项目和政策的有效性上找到证据的重要性。我们认为，项目评估人员和研究生学习如何评估，必须熟悉与他们感兴趣的项目相似的最重要的评估和当代评估。这可以为有效和无效的项目和实践提供有用的信息，还可以为评估同样设计类型的项目提供范例，其中后者可以为制定自己的项目评估提供指导。由于审稿人的意见和读者的反馈，出版过多个版本的教科书会得到不断改进，而我们这本书也不例外。回顾我们付出的点点滴滴，第六版是建立在前五个版本的坚实基础上的，但也添加了新的材料。 xi

有时候很难确定哪些部分需要“重做”，哪些部分需要保留。突出经典研究固然重要，但更重要的是，总有些最新观点和研究需要突出。有些内容必须常变常新，才能跟上时代的步伐。但是有些教育家总是喜欢选择他们以前用过的教材，因为他们对内容很熟悉，也很喜欢。如果一本书变化得太快，就不像最初选择时那样适用了。知道哪些需要更新、哪些需要摒弃，这多少也是一个挑战。所以亲爱的读者朋友们，这也是为什么要请你告诉我们，哪些起作用了而哪些却没有。如果项目评估的有些章节我们的解释不太好，你用起来不顺手，烦请你把自己的想法发邮件告诉我们。

对于刚刚发现这本书的读者来说，《项目评估》这本书就是要成为社会工作、社会学、心理学、公共管理、咨询、教育、护理及其他相关学科研究生的基础教材。重要的是要知道，干预是否对焦点个体、家庭、小组或社区有预期影响。它也可以作为研究方法的补充教材。我们希望本书已经写得很清楚，那些刚刚担任督导或评估具体项目或政策的从业人员、管理人员和其他人员也许会发现，这本书在对项目进行计划评估或需求评估时，是非常适用的。 xii

我们的重点也是沟通要点，即概念化项目评估所需的基本工具和知识，并执行与审查和评估项目绩效相关的任务。本书旨在帮助学生和实务工作者理解并促进循证实践。至关

2 项目评估——循证方法导论（第六版）

重要的是，我们的专业人员和专业培训人员不断检查我们方案的有效性及其对我们的服务对象的影响。这种改进方式既有利于服务对象也有利于社会。我们希望本书能为协助你承担这一重要责任尽绵薄之力。

戴维·罗伊斯 (David Royse) (droyse@uky.edu)

布鲁斯·A. 赛义 (Bruce A. Thyer) (bthyer@fsu.edu)

德博拉·K. 帕吉特 (Deborah K. Padgett) (deborah.padgett@nyu.edu)

第 1 章

绪论 / 1

- 一、项目评估的重要性 / 1
- 二、什么是项目? / 7
- 三、“优质”项目的特征 / 8
- 四、基于证据的实践 / 11
- 五、项目评估的定义 / 16
- 六、为什么要进行项目评估? / 17
- 七、项目评估的动机 / 18
- 八、克服主观性观点 / 22
- 九、项目评估的哲学假设 / 26
- 十、实证主义的更多信息 / 30
- 十一、本章小结 / 32

第 2 章

项目评估中的伦理问题 / 38

- 一、伦理审查委员会的角色 / 39
- 二、专业伦理准则 / 39
- 三、如何使用伦理准则? / 44
- 四、对特殊群体的研究 / 47
- 五、涉及不同文化与群体的伦理问题 / 48
- 六、行为研究和参与项目的伦理界限 / 50
- 七、评估工作者的伦理责任 / 53
- 八、伦理问题与评估研究总结 / 56

第3章

需求评估 / 59

- 一、需求评估的类型 / 60
- 二、如何进行需求评估？ / 64
- 三、需求评估涉及哪些步骤？ / 65
- 四、如何选择需求评估方法？ / 67
- 五、关于调查我需要了解什么？ / 73
- 六、特殊情况：合并数据 / 76
- 七、在需求评估中发挥创造性思维 / 78
- 八、什么是社区意愿研究？ / 79

第4章

定性评估法与混合评估法 / 87

- 一、导言 / 87
- 二、什么是“定性评估法”？ / 88
- 三、什么是“混合评估法”？ / 89
- 四、定性评估法何时可发挥作用？何时采用混合评估法？ / 89
- 五、定性法与从业人员的现场评估 / 92
- 六、常用于评估研究的定性法 / 92
- 七、定性评估法 / 93
- 八、混合评估法的设计 / 93
- 九、管理和组织数据 / 98
- 十、数据分析 / 98
- 十一、混合数据分析法和综合分析法 / 101
- 十二、质量控制 / 102
- 十三、撰写报告 / 103
- 十四、推广和倡议 / 104
- 十五、定性评估案例 / 106

第5章

什么是形成性评估及过程评估？ / 113

- 一、如何建构逻辑模型？ / 114
- 二、真实情景 / 116
- 三、形成性评估 / 116
- 四、形成性评估的操作 / 117
- 五、什么是过程评估？ / 120
- 六、过程评估：项目描述 / 122

七、过程评估：项目监控 / 125

八、成为项目督导 / 126

九、宗旨、目标和目的 / 128

十、项目目标的撰写 / 131

十一、应该督导什么？ / 133

十二、质量保证 / 136

十三、整体质量的管理 / 138

十四、本章小结 / 140

第 6 章

单系统研究设计 / 144

一、什么是单系统研究设计？ / 144

二、结果测量的选择 / 146

三、评估测量的演变 / 149

四、符号及其基本原理 / 149

五、需求评估 / 152

六、形成性评估 / 153

七、质量保证研究 / 156

八、总结性评估设计 / 158

九、总结性实验性设计 / 161

十、外部效度 / 167

十一、推论统计 / 168

十二、如何准备曲线图？ / 170

十三、单系统研究设计的伦理道德 / 171

十四、如何批判性地审查单系统研究设计？ / 173

十五、本章小结 / 174

第 7 章

客户满意度 / 180

一、监控客户满意度的重要性 / 180

二、客户满意度调查中存在的争议 / 182

三、客户满意度研究中存在的问题 / 183

四、近期客户满意度的一次抽样调查 / 186

五、如何解释高回应率？ / 187

六、对客户满意度研究的建议 / 188

第8章

抽样 / 198

- 一、什么是非概率抽样设计？ / 198
- 二、什么是概率（科学的）抽样？ / 202
- 三、抽样的考量 / 205
- 四、样本规模应该有多大？ / 207
- 五、本章小结 / 209

第9章

小组研究设计 / 213

- 一、什么是小组研究设计？ / 213
- 二、从结果评估开始 / 214
- 三、结果评估设计 / 217
- 四、小组研究设计的一般原则 / 220
- 五、前实验研究设计 / 220
- 六、准实验研究设计 / 229
- 七、对内在效度的一些威胁 / 235
- 八、防止其他解释 / 238
- 九、实验研究设计 / 239
- 十、效能与有效性研究 / 247
- 十一、什么是负面结果的研究？ / 248
- 十二、对“实验”一词的注解 / 251
- 十三、实验项目评估的一些新进展 / 252
- 十四、本章小结 / 253

第10章

成本效益和成本分析 / 262

- 一、为什么要把成本作为一个评估标准？ / 262
- 二、成本效益评价实例 / 265
- 三、如何开始一项成本效益研究？ / 266
- 四、谁的观点？ / 269
- 五、成本效益分析 / 270
- 六、本章小结 / 275

第11章

测量工具及策略 / 280

- 一、为什么准确的测量对评估如此重要？ / 280
- 二、我们应该测量什么？ / 282

- 三、信度 / 285
- 四、关于信度我们有必要知道些什么? / 286
- 五、关于效度我应该知道些什么? / 291
- 六、怎样为我的评估项目找到一个测量工具? / 294
- 七、如何构建一个“好的”评估工具? 关于问卷设计, 我需要知道些什么? / 295
- 八、在编制调查问卷时, 有哪些常见的错误? / 297
- 九、关于设计评估工具的测量水平, 我还需要知道些什么? / 304

第 12 章

为项目选择最佳评估指标 / 308

- 一、选择评估工具的实用程序 / 309
- 二、临床焦虑量表 / 311
- 三、自我效能评估量表 / 313
- 四、社区对性犯罪者的态度量表 / 315
- 五、亲密暴力责任量表 / 317
- 六、学校支持量表 / 319
- 七、文化敏感性 / 321

第 13 章

务实的问题 / 325

- 一、什么是干预的忠实度? 它为什么这么重要? / 325
- 二、忠实度的噩梦 / 330
- 三、项目偏移和项目修改 / 333
- 四、政治问题是如何影响项目评估的? / 333
- 五、评估的“威胁” / 335
- 六、在政治舞台上进行评估的准则 / 336
- 七、内部评估人员和外部评估人员 / 339
- 八、文化敏感性评估实践 / 342
- 九、最后的思考 / 345

第 14 章

评估提案、报告以及期刊论文的撰写 / 350

- 一、评估提案和评估报告的组成部分 / 351
- 二、计划及撰写评估报告和论文初稿的注意事项 / 362
- 三、学生在撰写评估报告时常犯的错误 / 363
- 四、写作大纲及评估报告估价表 / 365

五、评估报告的使用 / 367

六、为专业期刊撰稿 / 368

译后记 / 373

为什么说“从业者应熟知如何进行项目评估研究，在某些情况下甚至要参与其中”是十分重要的呢？

只有通过项目评估，才能充分说明相应的服务是有用的还是无效的抑或是有害的。项目评估标志着一个专业的真实性，因为该专业的服务是以证据为基础的。所有的人文类服务领域都十分看重项目评估，并要求相关人员在实践及做出政策决定时，重视当前可靠的结果研究。

我需要具备哪些能力？

所有的人文类服务专家都需要具备阅读、理解以及批判性评价项目评估研究的能力，并且能够决定是否将这些研究运用到个人情况中去。有些从业者发现，正如当代伦理规章所提倡的，乐意去帮助他人设计、执行以及翻译项目评估是非常有必要的。

一、项目评估的重要性

欢迎进入项目评估这一领域，在这里我们将进行专业训练，我们的主要目标是帮助大家将研究与实践技能相结合，并通过相关研究更好地提高实践技能。研究发现，从事社会工作、心理学、咨询、教育、精神病学、公共卫生、公共管理以及护理行业等人文类服务的人员，鲜有人在其工作中致力于科学研究。然而，以上这些领域的每一个专家都专注于为不同的对象提供一定程度的服务（包括个人、家庭、情侣、小群体、组织以及社区的形式），每一位人文类服务从业者也都十分想知道自己的服务是否帮到了他的服务对象。因

此，即使做研究不是你的主要专业活动，你也会发现，运用科学、可靠的研究工具对服务和项目进行评估是一项十分有用的专业技能。除了可能要参与到你自己设计及操作的评估项目中之外，你也需要学会阅读并批判性分析他人操作的，或是发表在本学科或相关学科的期刊、书籍或网页上的评估研究。有能力去评判发表了或还未发表的项目评估的价值，实在难能可贵。在面对单个受助者或是团体受助者时，哪种类型的服务值得向他们提供，你该如何做出抉择？

在人文类服务领域，广义的科学研究主要有三大用途（它们彼此间相互关联）：客观描述事物、凭实际经验评估服务的有效性、合理解释事物。描述性研究有助于更好地理解特定机构所服务对象的特征及需求；评估性研究有助于确认客户的需求或目的是否得以实现；而解释性研究是为了揭示社会心理问题的成因以及干预因素产生作用的过程，在此过程中形成了两大理论——病因理论和干预理论。

项目评估可被视为研究的一部分，或被简单地定义为“用以寻求事实或准则的系统化程序”（Barker, 2014, p. 365）。美国教育研究协会对研究（research）给出了更宽泛的定义，详见表 1-1，其中有几点值得强调。研究一定要严格精确、系统周到并找出客观信息。真正的科学研究不是未经认真计划就随意进行的，它致力于找出事情的真相，而非去证明早已存在的预期。研究一定要符合逻辑且方法多样，这些方法将贯穿本书始末。在评估研究中，工具（instruments）这一术语通常指由受助者或项目中其他利益相关者（如看护人、监督人、老师等）使用的铅笔和纸尺。有时项目评估也涉及对可观察行为的直接测量，在这种情况下，声音研究包括采用一些适当的方法来评估这些观察性方法的可靠性及有效性。在极少情况下，项目评估会使用物理仪器，如生物反馈仪、血压或心率监测仪等。项目评估偶尔会评估客户的生理变量，如血糖、胆固醇含量以及测压。

3 表 1-1 科学研究的定义

1. “科学研究原则”这一术语指的是使用严格、系统、客观的方法，以获得可靠有效的知识。具体而言，这样的研究需要具备：
A. 发展一个具有逻辑性、依靠证据的推理过程；
B. 根据提出的问题采用合适的方法；
C. 拥有可观察的或可进行实验的设计和工具，以提供可靠、可归纳的研究结果；
D. 有充足的数据和分析来支持研究结果；
E. 清楚详细地说明研究程序和结果，包括可以归纳结果的特定人口数量；
F. 遵循同行评审的职业规范；
G. 传播研究结果，为科学知识做出贡献；
H. 能够重新分析、复制数据，并有机会根据研究结果加以修改。
2. 检查临时的需要运用随机评估或准实验研究及其他实验设计，这从实质上减少了对已有结果做出其他看似合理的解释的可能性。其中包括纵向设计、病例对照研究、测量与多元分析、时序分析，但不仅限于此。这一标准尤其适用于评估政策及项目对教育结果影响的研究。