

“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 晏凤鸣

THOMSON

公共项目评估导论

第三版



Program Evaluation: An Introduction
third edition

[美] 戴维·罗伊斯 (David Royse)
布鲁斯·A·赛义 (Bruce A. Thyer) 著
德博拉·K·帕吉特 (Deborah K. Padgett)
T.K.洛根 (T.K. Logan)

王军霞 涂晓芳 译
杨爱华 王丽珍 校

中国人民大学出版社

CENGAGE
Learning™

“十五”国家重点图书出版规划项目

社会工作经典译丛

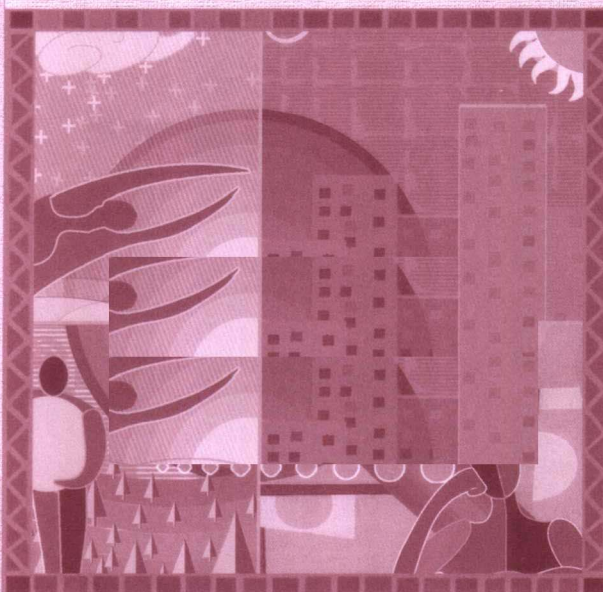
Social Work Classic Series

主编 晏凤鸣

THOMSON

公共项目评估导论

第三版



Program Evaluation: An Introduction
third edition

[美] 戴维·罗伊斯 (David Royse)
布鲁斯·A·赛义 (Bruce A. Thyer) 著
德博拉·K·帕吉特 (Deborah K. Padgett)
T.K.洛根 (T.K. Logan)

王军霞 徐晓芳 译
杨爱华 王丽珍 校

中国人民大学出版社

图书在版编目 (CIP) 数据

公共项目评估导论 (第三版) / [美] 罗伊斯等著; 王军霞, 涂晓芳译
北京: 中国人民大学出版社, 2007

(社会工作经典译丛)

ISBN 978-7-300-08151-9

I. 公…

II. ①罗…②王…③涂…

III. 公共管理: 项目管理-项目评价

IV. F062.4

中国版本图书馆 CIP 数据核字 (2007) 第 078117 号

社会工作经典译丛

公共项目评估导论 (第三版)

戴维·罗伊斯

[美] 布鲁斯·A·赛义 著
德博拉·K·帕吉特

T. K. 洛根

王军霞 涂晓芳 译

杨爱华 王丽珍 校

出版发行 中国人民大学出版社

社 址 北京中关村大街 31 号

邮政编码 100080

电 话 010-62511242 (总编室)

010-62511398 (质管部)

010-82501766 (邮购部)

010-62514148 (门市部)

010-62515195 (发行公司)

010-62515275 (盗版举报)

网 址 <http://www.crup.com.cn>

<http://www.ttrnet.com>(人大教研网)

经 销 新华书店

印 刷 河北涿州星河印刷有限公司

规 格 185mm×235mm 16开本

版 次 2007年7月第1版

印 张 23.5 插页 2

印 次 2007年7月第1次印刷

字 数 469 000

定 价 35.00 元

版权所有 侵权必究

印装差错 负责调换

总序

随着我国从传统社会向现代社会、从农业社会向工业社会、从封闭社会向开放社会的转型，必然产生众多的社会问题。随着人们对解决这些社会问题的必要性和紧迫性的认识，逐步改革和完善我国的各种社会福利制度和建立一支专业化的社会工作服务队伍成为必然的选择之一。这也是建立社会主义和谐社会的标志之一。因此，自20世纪80年代末期我国恢复社会工作专业教育以来，尤其是近几年来，我国的社会工作专业教育快速发展起来。但是，我们的社会工作专业教育质量还不高。虽然其中的原因多种多样，但是缺乏高质量的教材是重要原因之一。为了应对我国社工专业教育的快速发展，国内也编写了一些专业教材，但是，其理论的系统性和可操作性都有所欠缺，难以满足专业教学的需要。因此，系统地翻译一些来自欧美国家的社工专业教材是十分必要的。

第一，在欧美国家，社会工作专业教育已有近百年的发展，专业教育制度完善发达，专业教育水平和专业教材的水平都较高。

第二，作为一种服务性专业，不仅要求系统准确的理论指导，也要求规范科学的实务技巧的引导。

第三，作为一种新兴发展的专业，在其初创时期，系统引进发达国家的专业教材几乎是难以回避的选择，这也是使一个新兴专业迅速成熟的必由之路。

第四，尤为重要的是，我国在改革、建设和完善社会福利制度和建立一支专业化社会服务队伍的道路上并不会一帆风顺，社会工作专业被社会接纳仍需时日，专业培养的人才一定程度的浪费仍难以避免，因此，扎扎实实地做一些基础性的准备工作，争取把人才培养的工作做得更扎实，争取以较高的教育质量服务于有志于社工事业的人们是十分有意义的事情。

第五，引进高水平的社工专业教材不仅只是针对社工专业教育的要求，对于实践中正在从事各种类型的社工事业的人来说，也具有积极的意义。

正是基于上述理由，我们决定组织系统翻译一套社工专业教材。本套教材主要针对国内各大学社会工作专业所开基础课程，同时，考虑到国内专业教育的发展，我们也选择了提高社会工作技能方面的书。在翻译的种类确定以后，我们要求所选择的书籍要经典、要权威，要在欧美各大学被广泛地采用，要尽可能地采用最新的版本以反映最新的学术成果。为了保证该套丛书的质量，我们组织了一支有社会工作教育经验且外语水平较高的中青年学者的翻译队伍，这批译者不仅有乐于推动我国社会工作教育的热情，也有为此而贡献他们智慧的意愿，要在平日承担繁重的教学科研任务之余高质量地完成该项任务，可以想见他们的艰辛。

经过近两年的努力，凝聚着译者、编审人员和出版社巨大汗水的社会工作经典译丛由中国人民大学出版社出版。我们诚挚地向一切有志于我国社工事业的专家、学者、学生和各类社工实践者推荐此套丛书。我们相信，此套丛书定能为我国的社会工作专业教育和社工实践起到积极的推动作用。

当然，应当指出的是，任何一个国家的社会福利制度和社会工作教育都是由这个国家的特定的历史时期和其特定的文化环境所决定的。我们所翻译的这套系列教材也不例外。因此，我们希望，人们在使用这套教材时，不应忘记我们的国情，不应忘记我们的使命是要创造性地使之本土化。另外也要指出的是，由于社会工作专业在我国是一个新兴专业，从事本套教材翻译的学者又都是近年来逐步成长起来的中青年学者，翻译中势必存在这样那样的缺点，我们诚挚地希望广大的专家和朋友们不吝赐教，以不断地把此项事业推向完善。

最后，非常感谢中国人民大学出版社及其各位编辑，尤其是公共管理事业部的刘晶主任，没有她的真知灼见和对于推动社工专业教育的热情及富有实质性的帮助，本套丛书将难以出版。

晏凤鸣

2005年1月

序 言

本书重点

本书对项目评估作了全面介绍，适合从事社会工作、咨询服务、教育工作的人们和从事社会学研究、心理学研究的研究生以及其他行业从事与项目评估相关的工作的人员阅读，也可以当作研究方法论课程的补充教材。对行政管理人员以及刚刚或者即将被任职为检查和评估特定项目的人员，此书也值得一读。

我们主要集中于讨论项目评估的本质和精髓，提供项目检查和绩效评价需要的基本工具和必要知识。我们相信即使是那些已经离开课堂若干年的人也不会对此书的内容感到畏难。作为教育家，我们知道一个适当的例子常常比过多的解释更有价值。我们并不希望向读者灌输我们所了解的全部内容，只是期望读者能更容易地理解和掌握基本概念。

而且，我们认为，此书的大多数读者并不依靠项目评估谋生。很多读者可能对这个题目没有太多的热情，甚至察觉不到此项研究和评估工作的有趣性所在。如果你也有着同样的想法，希望本书可以改变你的观点。

作者对项目评估所带来的巨大回报感触颇深。从个人层次来说，带给你的可能是兴奋、挑战抑或精神上的激励与满足。项目评估会给社会带来巨大收益。无论面临什么问题或者它有多么特

殊，此书都可以帮助实际工作者获得更有实效的实践活动。最为重要的是，它可以帮助你谋生，获得高薪，这就是本书可能作出的贡献。

在越来越多的使用计量方法的时代，项目评估技巧不仅是必要的而且是必需的。即使你不希望成为全职的评估专家，我们相信这本书中所介绍的正确的评估概念和方法可以最终帮助你为案主提供更好的服务，更好地管理服务案主时所使用的资源。

关于第三版

第三版的各章都进行了广泛的修改，增加了新的章节（例如评估中的定性方法）和新的主题（例如诊断的逼真度），并对其他章节进行了扩展。旧的参考文献被删去，使用了最新的研究成果。在更新的材料中，我们突出了最新的有关项目评估如何被运用于实践的文献，再次强调评估技能是注重实际和实效的。

两位学者成为该版新的合著者。《社会学研究的定性方法》（*Qualitative Methods in Social Work Research*）（Sage, 1998）一书的作者德博拉·K·帕吉特为本书撰写了新的一章，解释了评估的非定量化的一面。另一位新作者 T. K. 洛根是一名对吸毒、酗酒、暴力、艾滋病等主题感兴趣的专家，也是一位迅速崛起的评估大师。他对过程评估一章进行了修订，并从其他方面对其进行了重要的补充。《社会工作实践研究》（*Research on Social Work Practice*）杂志的主编布鲁斯·A·赛义，同时也是美国最具献身精神和最多产的从事社会工作的学者，为本书修订工作的顺利展开提供了非常有用的建议。

我们希望本书能够最大程度地减轻你对评估工作的恐惧，并激发你对此项工作的兴趣，使你更加了解项目评估的多种方式和方法。同时，我们确信，读完本书后你会对评估工作的复杂性有更多的了解，同时也会理解开创性研究中信息缺乏的困难。

无论你是自愿还是不自愿踏上项目评估这条道路的，我们的目标是为你的行程做好准备。祝你阅读愉快！

目 录

第 1 章	绪论	1
	项目评估的重要性	1
	什么是项目	5
	优秀项目的特征	5
	项目评估的定义	10
	项目评估的原因	11
	克服主观性	15
	项目评估的哲学假设	19
第 2 章	项目评估的道德问题	28
	历史背景	28
	道德原则	30
	怎样把握道德原则	35
	对特殊人群的研究	37
	执行道德研究时的潜在问题	38
	实践者的伦理责任	41
	道德窘境	42

第 3 章	需求评估	46
	什么是需求评估	46
	需求评估的必要性	47
	需求的定义	48
	需求评估计划的制定	50
	需求评估方法的选择	52
	需求评估中的创造性思维	63
	团体意愿	66
第 4 章	评估中的定性方法	74
	什么是定性评估	75
	定性评估何时有用?	75
	定性评估方法和单系统评估	77
	设计一个定性评估	78
	行动(参与)评估	80
	进入评估	81
	互惠、回报和反馈	81
	定性评估中的取样	82
	数据收集	83
	分析定性数据	87
	质量控制	88
	撰写报告	89
	定性评估案例	91
第 5 章	构成评估和过程评估	99
	构成评估	99
	实施构成评估	101
	构成评估的其他用途	104
	过程评估	105
	项目监控	109
	使命陈述、目标和目的	112
	撰写项目方针	114
	监控的内容	117

	质量保证	121
	全面质量管理	126
第 6 章	单系统研究设计	136
	什么叫单系统研究设计?	136
	挑选产出指标	137
	反复评定的指标	139
	需求评估	139
	符号和一般规则	141
	构成评估	142
	质量保证研究	145
	累积性评估设计	146
	实验设计	149
	外部有效性	154
	推理分析	155
	如何准备图表	157
	单系统研究设计的伦理守则	158
第 7 章	目标实现测评法	164
	什么是目标实现测评法	164
	目标实现测评法的例子	166
	目标实现测评法和项目评估	167
	目标实现测评法的变异	168
	目标实现测评法的问题	169
第 8 章	案主满意度	173
	关注案主满意度的理由	173
	案主满意研究中的问题	175
	近来案主满意研究的抽样调查	176
	关于案主满意的解释	178
	高评价的解释	179
	对案主满意研究的建议	179
	样本数目	183

第 9 章	小组研究设计	192
	什么是小组研究设计	192
	启动结果评估	193
	成果评估设计	194
	小组研究设计的基本原理	196
	实验前研究设计	196
	准实验研究设计	202
	准实验设计	203
	内部有效性的威胁	207
	防止其他可能解释	208
	实验设计	209
第 10 章	成本效益和成本分析设计	221
	作为评价标准的成本	221
	成本效果示例	223
	如何进行成本效果研究	223
	谁的观点?	226
	成本效益分析	227
第 11 章	测评工具和策略	234
	测评的重要性	234
	决定测评什么	236
	信度	238
	证明信度	239
	效度	244
	确定合适的工具	247
	建立“好的”评价工具	249
	测评水平	255
第 12 章	测评工具说明	262
	虐待儿童倾向测评量表	263
	临床焦虑量表	266

	抑郁量表	268
	希望量表	270
	UCLA 孤独量表 (第三版)	272
	罗森伯格自尊量表	274
	成人及青少年亲职问卷	276
	简明症状问卷	278
	HIV 心理调节量表	280
	知觉社会支持量表	283
	健康量表	286
第 13 章	实际问题	290
	实施 (过程) 的精确性	290
	精确性的梦魇	292
	计划的偏离和转变	294
	评估的政治性	295
	评估的“威胁”	297
	政治舞台中评估的指导方针	298
	文化敏感的评估实践	300
第 14 章	数据分析意味着什么?	307
	数据分析和计算机	308
	单变量分析	309
	双变量分析	315
	数据的多变量分析	322
	关于统计显著性的神话	323
	理解趋势	325
	在报告中使用统计数据	327
第 15 章	撰写评估建议书、报告和期刊文章	333
	评估建议书和评估报告的组成	334
	正确撰写评估报告	343
	为评估报告的写作和评价准备清单	346
	评估报告的应用	348

6 公共项目评估导论 (第三版)

为专业出版物写作	349
索 引	353

第 1 章 绪 论

项目评估的重要性

欢迎你进入项目评估的殿堂。此方面的职业训练旨在帮助你
将研究和实践技能有机结合起来，利用前者来加强后者。必须承
认，社会工作、心理学、咨询服务、教育、公共管理和护理工作
这些公共事业领域的实践者，很少拥有从事科学研究的经历。然
而，这些领域的任何人都会涉及为不同层次的案主（个人、家
庭、夫妇、小组、组织或者社会团体）提供服务，并且每位公共
事业的实践者都会有强烈的兴趣去了解他们所提供的服务是否真
正有助于他们将要服务的对象。因此，即使你并不喜欢这项研
究，你将会发现，用科学可信的研究工具来评估服务和项目是一
种具有重大价值的专业技能。除了你亲自参与项目评估的设计和
执行，你还必定了解其他人所进行的评估研究以及发表在学术刊
物和著作中的相关知识，并对它们进行批判性的分析。你对发表
的和未发表的有关项目评估研究的价值的判断能力有助于增强你
的实践能力。否则，你如何判断哪种类型的服务对你的案主来说
是有价值的呢？

与公共事业有关的科学研究的目的是从广义上讲主要可以分
为三类（但它们常常相互交错）：客观地描述事情、评价服务
的有效性和有效地解释事情。描述性研究将案主看作一个特殊

的对象，此类研究可以更好地理解案主的特征和需求。评价性研究有助于判断案主的需求是否被满足或者案主的目标是否达到。而解释性研究则旨在发现心理方面的原因或者找出某个起阻碍和干预作用的过程（因此分别对病理学和干预理论都有所贡献）。

项目评估可以被看作是那些所谓的研究活动的一个子集，可以被简单定义为“用来寻找事实和原理的系统化程序”（Barker, 1999, p. 410）。评估研究即“对决定特殊项目成功的因素的系统化调查”（Barker, 1999, p. 161）。项目评估是一项实践工作，而非学术性活动，也不是要尝试构建一些理论或者一定要创造社会科学知识（如果能做到这一点，当然了不起）。正如 Tripodi 在《社会工作百科全书》（*Encyclopedia of Social Work*）中指出的那样，“在社会工作中，项目评估的任务是提供可以提高社会项目绩效的信息”（Tripodi, 1987, p. 366）。

有趣的是，尽管几十年来公共事业专业性地设计和实施项目评估所具有的异乎寻常的重要性已经被广泛认可，但这一领域的项目评估研究并不多见。Rosen 等人（Rosen, Proctor, & Staudt, 1999）对 1993 年至 1997 年上半年期间 13 种社会工作重要期刊上刊登的所有 1 849 篇文章进行了统计，发现只有 863 篇（47%）文章是研究型文章。在这 863 篇研究型文章中，314 篇是描述性研究，423 篇是解释性研究，仅有 126 篇文章与某种类型的评估（比如输出）有关。总之，项目评估方面的文章仅占有所有已发表的以社会工作研究为主题的文章的 3%。你可以借助框 1—1 中所引用的文献来分析这个事实。

框 1—1 关于评估研究重要性的观念

我对测度、估价、估计、评价结果很感兴趣。这项工作在某形式上超越了以信念、主张和例证性案例为依靠的术语。它陈述你的目标和与你与目标间的距离……我相信通过评估可以给案主提供更好的服务。（我呼吁你们……用某种形式测量、评估、估算或评价你们的成果，在任何情况下，该评估都要超越信念、主张以及“例证性案例”等主观性事物。描述你的目标并介绍你目前完成了多少……出于这种评估的结果，我相信能更好地为案主服务。）（Cabot, 1931）

第三种研究类型，对公共事业和实践者活动的评估研究，是最为重要的（Angell, 1954, p. 169）。

社会工作不是一门旨在创造知识的科学，而是旨在运用知识进行目标控制的技术。因此，连续化的社会工作研究更接近于应用结果，因为它的目的是将知识付诸实践（Greenwood, 1957, p. 315）。

评估和案主反馈不仅对提供有效的服务有帮助，也是考察职业道德所要求的。必须形成系统化的方法来评价社会工作者是否帮助或妨碍了他们所服务的人群，或者根本无所作为（Rosenberg & Brody, 1974, p. 349）。

任何关于其他社会工作方法的研究都没有评估实践的成果和政策选择的效果那么重要。

如果能够成功地以机构的实践效率的评估为中心，并在实践者和研究者之间建立起至关重要的合作关系，那么小规模、以机构为基础的研究是有意义的（Mullen, 1995, pp. 282 - 283）。

有关心理学干预效率评价的研究是必要的，包括在理想的受控条件下测试干预行为，以及现实的卫生保健系统（Ell, 1996, p. 589）。

对现实中的服务性干预行为的研究是关键要素，它涉及职业化实践者所应用的知识基础……因此，评估研究超越了描述性和解释性的研究，是一种对社会工作干预动机的设计性调查（Austin, 1998, pp. 17, 43）。

我们试图建立社会工作的研究议程，干预性研究在此研究议程中占据了很高的优先地位（Proctor, 1998, p. 19）。

从这一点出发，我们发现这样一个奇怪的现象：尽管很多权威专家认为项目评估是我们可以进行的最有价值的研究，但目前这方面的研究非常缺乏，发表的文章也非常有限。相反，大多数研究致力于描述性和解释性研究，而这些研究最多与公共事业的实践存在潜在的关系，并不具有像构思精妙的产出研究所拥有的那种直接应用的潜力。

为什么对已经设立的服务进行评估是必要的？因为一般情况下都会有多种可供选择的方案，有时候可以找到解决问题的更好的方案。例如，在一篇名为《大学咨询中心的限时服务：限时服务和非限时服务的中心有什么不同吗？》（Time-limited Therapy in University Counseling Centers: Do Time-limited and Time-Unlimited Centers Differ?）的文章（Gyorky, Royalty, & John, 1994）中，作者发现只能提供限时服务的咨询中心有更长的等待名单，并且接受其服务的学生占有所有学生的比例比非限时服务的咨询中心更低。现在，如果你是大学咨询中心的负责人，你会阅读这篇文章吗？理解这篇文章的作者的方法论和样本选择会很重要吗？

下面列举一个由很多行为学科学家和社会工作者所做的另外一个研究实例（Lynam, Milich, Novak, Logan, Martin, Leukfeld, & Clayton, 1999）。他们评估了一个广泛推广的预防吸毒的项目（被称作 DARE 项目）。该项目跟踪研究了采用 DARE 项目的 1 000 多个学生 10 年的状况，这些学生就读于 23 所学校，分布在 15 个年级。也研究了非结构化的毒品教育项目（在 8 个不同的学校进行）。进行 DARE 项目的学校是随机选定的（在 17 周内由警察对学生进行每周一小时的课程教育），或者实施健康教育培训（由健康教育者每周进行 30 到 45 分钟的健康教育，持续时间为 2 到 4 周），经过 10 年的训练和教育，发现这两组学生之间并不存在明显的差异。对接受 DARE 项目培训的学生和接受健康教育训练的学生的大麻的使用情况、违禁药品的使用状况以及学生的自尊心的评估，都表明两组学生没有差别。

假设你是一个学校的顾问或者是一名社会工作者，你被授权选择一种项目来预

防学校里的毒品泛滥。考虑到 DARE 项目的高成本及健康教育项目时间和资金成本较低，以及 Lynam 企图发现 DARE 项目较高的长期效应却遭到了失败，先前的评估研究表明 DARE 项目对防范毒品没有长期的效应（比如 Dukes, Ullman, & Stein, 1996）这些因素，你还会为你的学校选择 DARE 项目吗？具备常识的人会选择不——直到通过实证研究证明 DARE 项目的效果更明显。

假设你被要求为因受过性侵犯而精神紊乱的妇女提供心理健康方面的咨询服务，你会怎样帮助这些案主？一项由社会工作者和心理学家 Foa, Rothbaum, Riggs 和 Murdock (1992) 所从事的治疗评估研究会帮助你选择提供何种服务。45 个因性侵犯而精神受挫的女性被随机选定为接受或者等待接受认知行为的治疗、支援性咨询等服务的对象。接受治疗的 3 个半月后发现，认知行为治疗比支援性咨询以及任由时间流逝而不采取任何措施（等待治疗而未采取任何措施）更有助于帮助妇女消除由于性侵犯造成的心理创伤。而且，最近进行的采用更大样本量的研究（样本量为 96）再次证明认知行为治疗比不采取任何治疗措施的效果要好得多（Foa, Dancu, Hembree, Jaycox, Meadows, & Street, 1999）。诸如此类的评估研究对你选择何种治疗方式提供服务具有重要意义。当然，实证研究不仅仅是考虑采纳何种治疗模式，但我们认为它是一个很重要的问题。

这里还有另外一个强迫性研究（Weiss, Catron, Harris, & Phung, 1999）的例子。随机选取了 160 个接受由诊所的社会工作者、心理学家、护士等专业人士提供的临床式心理治疗或者由大学生和初中教师提供的家庭式心理健康教育。临床式心理治疗平均进行了 60 个课时，家庭式心理健康教育进行了大约 53 课时，两种治疗都维持了大约 2 年时间。许多来自精神病理学、适应性机能和行为问题方面的可靠而有效的指标被用来评价治疗的效果。接受职业化的心理治疗的学生在为期 2 年的治疗过程中在各方面有明显的提高。这当然是一个令人愉悦的结果，然而，接受家庭式治疗的孩子显然花费了比职业治疗少得多的成本！实际上，作者们得出的结论是，传统的临床式心理学治疗的效果并不比非职业化的唾手可得的家庭式教育好。现在，假设你是当地一个心理健康诊所的行政管理人员，你被授权为需要接受服务的孩子们选择服务项目，你可以选择昂贵的有官方许可证的心理治疗师，或者花很少的钱聘请可能获得相近效果的大学生执行治疗。你将会怎么做？好，这将会使一个优秀的自信的管理者宣布放弃执行孩子的心理学治疗（验证了非科学因素对决策过程的潜在影响）。并且，事实上，一个博识的管理人员应该随时了解是否这项发现被其他的研究所重现——这会进一步证明所做的选择并非侥幸。

而且，诸如此类的研究应该被鼓励，比如对心理治疗方面的文献的仔细检索以找出代替传统的心理学治疗的新方法，或者去发现是否还有其他更有意义的研究支持传统的治疗方法。此外，其中隐含的意义是很清楚的。在设计机构的服务项目时，先前发表的、可以相信的评估研究应该作为主要的输入资源。我们希望通过阅