



EVALUATION METHOD OF EDUCATION PROJECT

教育项目评估方法

微观视界的评估研究

刘康宁 著

EVALUATION RESEARCH
FROM THE MICRO-PERSPECTIVE

云南出版集团公司
云南科技出版社

责任编辑 孙 琳
邓玉婷
责任校对 叶水金
责任印制 翟 苑



www.ynkjph.com

ISBN 978-7-5416-3023-1



9 787541 630231 >

ISBN 978-7-5416-3023-1/G·613

定价：39.00 元



EVALUATION METHOD OF EDUCATION PROJECT

教育项目评估方法

微观视界的评估研究

刘康宁 著

EVALUATION RESEARCH
FROM THE MICRO-PERSPECTIVE

云南出版集团公司
云南科技出版社
· 昆明 ·

图书在版编目 (CIP) 数据

教育项目评估方法：微观视界的评估研究/刘康宁著.
—昆明：云南科技出版社，2008.9
ISBN 978-7-5416-3023-1

I.教… II.刘… III.教育项目评估—研究 IV. G449

中国版本图书馆CIP数据核字 (2008) 第154659号

云南出版集团公司

云南科技出版社出版发行

(昆明市环城西路609号云南新闻出版大楼 邮政编码：650034)

昆明理工大学印务包装有限公司印刷 全国新华书店经销

开本：787mm × 1092mm 1/16 印张：9.625 字数：180千字

2008年9月第1版 2008年9月第1次印刷

印数：1~1000册 定价：39.00元

序

如果我们想要找有关教育评估专著的话，那走进书店可谓举目皆是。在今天的教育评估研究中，理论与实践研究受到了广泛重视。但如果我们细致地阅读一下这些专著，又会发现研究重点基本还是以评估基本理论和方法居多，而关于评估专题研究的论著并不多见。在从事教育评估的研究与实践过程中，已深切地感受到有很多复杂的评估现实问题，在现有基本理论与方法框架下难找到系统答案，或者说现有的方法在解决这些问题时还存在局限性。

在问题研究与研究生课程教学的双重驱动下，开始萌生了研究“教育项目评估方法”的动念。项目评估这一术语，在社会学、经济学中较为常见。在社会学领域，项目评估是一个活动激烈的应用研究，包括收集、分析、解释和沟通有关旨在改善社会环境的社会项目绩效。评估有各种使用目的：帮助决定项目是否应该继续、改善、扩展和缩减；评估新项目的用途和创意；提高项目管理和指导的绩效；满足项目各方的需求（Peter H.Rossi,1979）。项目是活动的一个基本单元，对项目的管理主要依靠项目评估，但项目评估却并不专属社会学，若历史溯源的话，系统的项目评估最早是出现在教育和公共健康领域。反观今天的教育评估研究，人们似乎淡化了教育中的项目元素。社会科学的历史告诉我们，理论的建构和科学方法是紧密联系的，方法的改善、进步在描述、预测、解释现象方面总是起着至关重要的作用。

浏览现有的评估研究专著，项目评估内容仅作为教育研究的一般应用被提及，对项目评估方法的研究更为少见。在实践中，操作者更多地习惯套用一些基本评估方法来进行项目评估，譬如对教学计划、新教学方法以及教育政策的评估等等，在教育项目的管理与干预方面表现出诸多的不适应和过于笼统。从教育评估理论的建构看，“项目评估”的研究还尚属盲区，现实状况是人们在经常实施项目评估，却很少重视项目评估的研究，操作需求和方法研究之间存在严重的不对称。

把研究指向了一个重要而少有人涉足的领域，并非要凸现新异，之所以选择

“项目评估方法”，更多的考虑还是能使教育评估贴近现实需要，使评估在教育项目管理上有更多针对性，同时也给评估实施者提供更广的思索空间和方法选择。尽管本书的研究重点是教育项目评估的基本问题和方法，但却并不囿于理论的探讨，把方法科学性与系统性的研究应用于教育问题的解决，特别是教育项目的管理，并帮助项目绩效的提高才是本书所关注的。

本书的撰写，以项目评估的程序为主线，循序渐进地探讨与之相适应的方法论、评估基本方法和具体操作技术。主要读者对象包括教育学相关专业的本科生、研究生、评估的执行者、评估研究人员、官员、教育项目的组织者。为了使内容与读者的实际需要相近，本书将致力于帮助读者认识项目理论、丰富方法论的评估知识。此外，还将讨论这种社会研究形式的一些要素，并考虑这种应用研究在实际中的影响以及如何有效地使用它。围绕评估研究的一般程序，讨论的内容主要以操作为主，从如何介入评估开始，到形成评估报告的一系列活动中，该选择什么样的方法，以及怎样运用这些方法，都是要解决的关键问题。与此同时，还将引入一些身边的案例，以加深读者对方法的理解，并希望能帮助他们提高解决实际问题的能力。

在强调操作性的同时，并不意味着要编一本教育评估的“食谱”。教育活动的复杂性，使得教育评估技术很难像烹调技术一样速成。事实也表明，在一项教育实践活动中，要恰当地运用好评估方法需要评估者自身的丰富知识和经验，至今也没什么良方能教别人“做”评估，本书唯一能做的也是在学理讨论上有一好的开端，果真如此的话，也算是大幸了。

写作是痛苦而有趣的工作。本书的形成过程中，常常会因为思绪不清而陷入混沌的方法世界中，曾经几度停笔，又几度重拾。幸好，前人在社会科学研究方法的探索和管理科学的贡献，使得我有了再次继续下去的勇气和信心。欣慰的是，在写作过程中，又体验了一种从未有过的学习乐趣。在13年的教育评估实践经历和10年的研究方法教学中，我总是把活动过程和备课当作是最好的学习，每次都有不一样的体验，感觉自己和“研究方法”、“教育评估”两个“朋友”越来越熟悉，感情越来越深，以至于再也舍不得离开。现在，让我更高兴的是，又认识了“项目管理”这一“朋友”，而且让它们都碰面了。虽然，我对“项目评估”的认识还远远不够，甚至是有失偏颇的，但这一阶段的学习积累对于我却至关重要，起码使我更坚定了今后的方向，而且让我体验到了持续做一件事所获得的乐趣。我想，我还会沿着评估研究这条路继续走下去，下一个议题已在酝酿之中。

我始终相信：“过程比结果重要，体验比知识重要，行动比空想重要”。把这句话作为这段序的结束，也把这句话送给我三岁半的儿子。

刘康宁

2008年9月17日于云南大学

目 录

第一章 有关项目的议题	1
1.1 识别项目	1
项目概念	1
教育项目的指标	3
1.2 项目评估研究发展史	4
社会学的历史起因	5
经济学的方法演变	7
教育学的应用延伸	9
项目评估在中国	10
1.3 什么是教育项目评估	13
研究与评估	13
项目管理与项目评估	14
项目评估的内容	15
教育评估的质量标准	15
优秀的项目评估	17
第二章 方法论	19
2.1 社会科学范式	19
结构主义	19
结构功能	21
网络分析	22
2.2 项目评估的分析方法	23

关于项目指标的分析	23
项目指标结构的研究	25
2.3 项目评估的哲学基础	28
2.4 项目理论与评估理论	29
第三章 评估规划与设计	30
3.1 项目评估的目标	30
项目改进	31
实施干预	32
绩效判断	33
优化资源	33
解释与知识	33
3.2 目标表达	34
3.3 厘清评估各方的关系	34
互动	34
角色	35
卷入	36
传播	36
3.4 评估项目概念化和结构化	37
概念化	37
结构化	38
3.5 适合于项目评估的问题	38
形成问题	39
“好的”评估问题标志	40
典型的评估问题	40
3.6 评估策略	41
独立评估	41
参与性评估	42
自主性评估	42
3.7 撰写评估计划书	43
第四章 需求评估	46
4.1 什么是需求评估	46
4.2 需求评估的目的	47
寻找足够的存在理由	47

判定存在的问题	48
描述目标人口的服务需求特征	48
4.3 教育项目的“需求”	49
需求概念	49
需求类型	50
4.4 需求评估的步骤	50
进行需求评估的准备	51
识别需求评估的团体成员	51
如何确定鉴别“需求”的标准	51
结果的应用	52
4.5 需求评估方法	52
既有资料估计	52
趋势分析	53
需求预测	54
4.6 需求评估计划的制定	54
4.7 决策的需求定向评估	55

第五章 计划评估

57

5.1 计划评估的发端与形成	57
5.2 教育计划	59
精心编制的计划	59
教育项目计划及其特征	60
5.3 什么是教育项目计划评估	62
计划评估可以做什么?	63
“自上而下”和“自下而上”的切入点	63
5.4 评估模型	64
成绩——目标一致法	64
交互式评价	65
应答式评价	65
决策——管理法	66
判断定向的方法	66
对手式方法	67
多数——直觉法	68
5.5 项目可行性评估——计划评估应用实例	69
项目可行性评估的内容	69

项目可行性评估的程序	71
案例——高校教育项目可行性评估报告	72

第六章 过程评估

76

6.1 项目监控	76
关键人员	77
需要回答的问题	78
有效的参照标准	79
6.2 获取评估监测信息	81
参与式观察	81
访谈	82
调查	83
项目参与者分析	84
管理信息系统	84
6.3 目标实现评估	85
项目目标的确定	85
对项目目标进行管理和评估	86
目标网络	87
测评方法	87
6.4 进度控制	88
控制的目的	88
主要内容和步骤	89
工作细分结构	90
进度报告方法	91
甘特图	92
计划评审法（PERT）和关键路径法（CPM）	93
6.5 预算与非预算控制	93
预算控制	93
非预算控制	94
6.6 质量管理	95
来自国际标准化组织的定义	95
全面质量管理模式	96
项目的质量控制环节	96
质量评估技术	96

第七章 结果影响评估	98
7.1 影响评估的基本问题	98
实验方法的评估应用	98
项目影响与干预	99
实验室观察与实地实验研究	100
评估中的伦理	100
7.2 项目干预影响的实验控制	101
项目效果	101
评估结果的准确性和概括性干扰	102
“最大最小”控制	104
对照组的设立	106
7.3 前实验评估设计	109
单组后测设计	109
单组前测后测设计	110
静态组设计	111
7.4 实验设计	111
“典型的”随机化设计	112
所罗门四组设计	113
仅有后测的对照组设计	113
7.5 准实验研究设计	114
时间序列设计	114
多重时间序列设计	114
第八章 评估操作化工具和实例	116
8.1 量表设计	116
量表类型	116
量表编制	117
8.2 量表建立的实证研究	125
大学教师品质测量的瑟斯顿量表研究	125
8.3 教学评估问卷研究示例	128
高职高专评估问卷设计研究	128
8.4 满意度评估问卷研究示例	133
教育改革教师满意度评估问卷	133
学生满意度评估问卷(部分)	136
8.5 研究性问卷研究示例	137
大学生创业行为与态度调查问卷	137
参考文献	140

第一章 有关项目的议题

项目评估发端于公共社会项目和经济项目，继而引申到教育领域。在从西方工业发达国家传入中国的近几十年，项目评估理论和实践都得到了极大丰富，特别是教育领域，人们已越来越意识到评估的重要性，也意识到了对教育项目研究的重要性。评估正在从国家行政管理的手段向下转移，有更多的教育机构、学校甚至教育者产生了评估管理的需求，而他们关心的就是教育项目。本章旨在加深人们对教育项目及项目评估的认识，以及与之相关要素的理解。

1.1 识别项目

对项目的基本属性和特征进行分析，并把项目管理作为有效的管理方法体系，已成为几乎所有产业领域的重要管理手段。现代项目管理方法形成于美国20世纪的曼哈顿计划，经过半个多世纪的发展，已发展成理论和方法都相对完整的管理科学和艺术。

项目概念

各种教科书中，对项目的定义至今也未能达成共识，这也许来自于学者对概念认识的不断推陈出新，更可能是项目本身的复杂性和不确定性使然。对项目概念的认识，笔者认为可以从五个关键词语来理解：

一次性工作 多数学者都认同项目是“一种一次性的复合任务”¹，项目是一次性的工作，单件性或非重复性工作，但其本质是一个多任务复合体。一次性明确了项目的时间维度，即在约定的时间范围内完成一项工作，有明确的开始时间和完成时间，一项不断重复和反复完成某一种单一任务不能算一个项目，“这也是项目不

¹ [美]詹姆斯·刘易斯. 赤向东译. 项目计划、进度与控制(第3版)[M]. 北京: 清华大学出版社, 2002. 5

同于日常运营的一个方面。”¹同时，一次性还表明项目的任务属性是“一件”，尽管这件任务可能是复杂的，与其他任务有关联，但这项任务应该是独特的，与某一目标相联系的。

一个目标 项目必须有一个特定目标，目标的不同，可以让我们能区分任务特点，以便划分不同的项目，项目目标也是我们对项目进行评估的前提，是项目管理的重要依据之一。项目目标有条件性和成果性目标，条件性目标是限制性条件，包括项目实施的条件、成本及质量目标；而成果性目标则是对项目最终的预期价值和功能的效益要求。一个目标下可能有多元与多层次的子目标，不仅有条件性的，还有限制性的，但都与该项目的独特要求有关。

预期价值 项目在确立之初，项目各方都在期望着项目实施以后的成果价值，价值载体是成果，而成果是目标制定时所规定的，当然不排除在项目实施过程中，价值发生转移。项目价值是项目在实施过程及完成后，对项目利益相关者需求的满足程度，价值是测量项目质量经常要考虑的因素。

成本与收益估算 项目需要投入资源，包括人力资源和非人力资源，在“资源有限”的限制下，需要在进度控制中，合理分配和利用资源。任何一个项目都面临着在固定资源中，将实施成本降至最低，追求目标价值及收益的最大化。

关联性 关联意味着项目活动间存在着逻辑和技术关系。首先，项目不会孤立存在，总是与其他活动有不同程度的联系；其次，如果我们把项目看作由各个部分——目标、资源、时间、环境、过程、思想、机构、人员、原则等，相互作用而形成的一个可以识别的整体，那么这个整体就有活动的输出、输入，项目与子项目等多种关联关系。

就教育项目而言，对概念的归纳总结目前还不是第一位的，更重要的工作恐怕是对教育项目属性的理解。既然对项目的界定还存在分歧，那就没必要把过多精力浪费在“教育项目”的概念定义上，更何况我们已经被越来越多的教育学概念所混淆，以至于很多学生在学习过后，感叹从概念到概念的无趣。

专栏1-1

“一个项目到底是什么？”

一个项目就是在一个事先确定的期限和预算范围内，必须完成的，且要达到特定的预期效果的工作。

它的一致性本质将它与我们日常做的、缓慢而无价值、但我们大多数人大部分时间依然苦干的工作区别开。

设计新网站是一个项目。一旦设计结束，上了网，并清除了初期的小毛病，项目就宣告结束。每周用新资料维持该网站就不是一个项目，它是日常工作。假如下一年重新设计，它就是一个新项目。

1 威安邦. 项目评估学[M]. 天津: 南开大学出版社, 2006. 3

你上小学时为自然科学课准备彗星剪贴簿——制作彗星剪贴簿就是一个项目。上高中时、每周阅读10页自然科学课本并回答问题——这就是日常工作。

资料来源：[美]多娜·迪普罗丝，汪开虎译，项目管理[M]，上海：上海远东出版社，2002. 9

教育项目的指标

指标常被用来说明概念的属性，马歇尔和罗斯曼（Catherine Marshall & Gretchen Rossman, 1995）用“概念漏斗”的方法，帮助我们认识事物的根本属性。借助这一方法，我们可以更好地认识教育项目，为选择适当的项目评估方法奠定基础。

时间参数

大多数人可能会因为教育活动的延续性，比如一个专业的建设，某项技能的训练等例子来否认项目的一次性时间特征。事实上，无论项目有什么样的独特性，项目完成都有生命周期，即项目从开始到结束所经历的各个阶段，项目周期及其各阶段的时间活动序列是客观存在的，项目中各项活动必须按特定序列或顺序进行。不论怎么创新项目周期模式，项目的构想与准备、项目计划、项目执行与控制、项目报告五个环节都是必不可少的。其中，某项活动开始需要以上一序列活动的结束作为起始，反过来又可能成为下一周期活动的条件，各阶段相互衔接，不同的环节构成了项目的循环周期。

时间参数除了有阶段循环的周期外，还有不可储存、自行流逝的资源特征。很多时候，教育项目的发起者和执行者似乎淡忘了项目时间指标的存在，没有人会在意项目是否在按进度执行，也没有人会为弥补进度而努力或不安。教育项目与其他项目所共有的这一特点，决定了教育项目存在效率的问题。意识到这一点，可以提醒我们在评价教育项目时，应考虑到效率和效益的因素。

成本 - 收益

有学者感言，现在的教育已越来越不计成本了。其中有两层含义：一是教育活动经常对经济成本大而划之，二是对时间成本忽略不计，这在经济领域几乎是不可原谅的。相对于其他资源，教育项目资源更显得稀缺，估算成本、合理分配资源，本身就是教育项目实施的重要条件，对经济成本的有效控制，体现了对教育项目价值的尊重。从时间成本角度看，教育项目往往是在没有多大竞争压力的环境中实施的，没有压力，自然也不需要挤压时间。不论是一项教育课题，还是一门课程建设，时间约定形同摆设，即使是按期完成，也很少能与项目成果质量联系，以凸出项目的效率。项目理论中，最优的时间控制，也是降低成本的办法，时间延长，项目成本会因为缺乏效率而提高；反之，时间缩短便会碰到边际收益缩减点，也就是“人海战术”不会因为人数不断增加，而让项目时间无限减少。教育项目的成本 - 收益问题与一般项目是一致的，这一属性因为资源稀缺性而更显得重要，教育项目评估方法需要把项目占有

资源和时间放在一起,来综合分析 with 成果的匹配程度。

质量指标

汉谟拉比法典规定,如果一幢建筑物倒塌了,那么设计师和施工人员都要被处以死刑,来自古巴比伦的惩罚性管理成为了有史可考的最早质量管理方法。而现代质量概念则出自于工业社会人们对产品不稳定性产生的忧虑,更强调过程的控制以及预防问题发生。20世纪20年代,贝尔实验室沃尔特·休哈特(Walter Shewhart)团队通过控制图、抽样技术和经济分析工具等新理论和统计方法,来评价不稳定的电话服务质量,从而奠定了现代质量评价的基础。此后,美国和欧洲的一些企业和组织,例如美国国家标准学会(ANSI)、国际标准化组织(ISO)、欧洲质量委员会等纷纷研制各领域的质量标准,以推进质量管理水平。

在教育领域,“教育质量”是各类从事教育工作的学者、领导甚至普通民众时常脱口而出的流行词,可没有人能真正说清楚教育质量到底是什么,怎样才能实施有效的评价。ISO在1994年对质量的定义“与满足明确的或隐含的需要有关的实体的全部特征。”¹也难以引申到对教育质量的理 解。要知道,教育活动有目标、有内容、有需求,那就必定有质量可循。国内一些学校也曾将西方的全面质量管理方法引入到教育管理,大多无果而终,但不可因此而简单地归咎于方法错了,或教育不适合质量评估。我们承认,由于教育活动时间边界的模糊性和教育效益的滞后性,以致于教育的质量标准难以明确。但也不可否认,现在对教育质量的评价方法和思路过于笼统,在缺少项目评估方法支撑的情况下,开展人才培养质量评估和办学水平评估,无疑是一种舍弃细节、追求整体的行为。现行的本、专科院校评估受到如此多的质疑,正是方法缺陷的一个映照。简单将ISO9000族标准复制到教育质量管理行不通,那么关注教育活动的项 目构成,建立项目的质量标准和评估方法,尝试从教育项目的评估入手,兴许会对我们认识教育的质量指标有所启发和帮助。

1.2 项目评估研究发展史

“项目评估”是社会研究的一般应用,在社会科学研究方法体系中也称评估研究,是指试图评估社会项目的有效性的一种应用性研究,目的在于理解和分析社会问题,以及评估研究项目的有效性。

评估研究的历史不算久远,随着社会研究方法的发展和改进,特别是20世纪社会问题的凸显,引起了人们解决问题的兴趣与探索。在美国,政府为了城市发展、军事等需要,实施了很多大规模的公共项目,实施前需要知道“项目是否可行”,而实施后又需要知道“结果如何”。正因为如此,社会项目评估研究成为了政府、

1 International Organization for Standardization (ISO). Quality Management and Quality Assurance. Geneva, Switzerland: ISO Press, 1994. 转自[美]蒂莫西J·克洛彭博格,约瑟夫A·佩特里克.项目质量管理[M].北京广联达慧中软件技术有限公司译.北京:机械工业出版社,2005. 5

学者和公众关注的焦点，现代评估研究在特定的社会背景下得以迅速发展和繁荣。

“系统的社会项目评估最早出现在教育和公共健康领域。在第一次世界大战以前，最大的努力就是扫盲、职业培训以及为降低死亡率和流行病发病率所实施的公共健康项目。20世纪30年代，各学科的社会科学家都曾致力于用严格的研究方法评估社会项目，由此，使得系统的评估活动变得越来越频繁。”¹可以说，是社会项目的实施绩效分析催生了评估研究，而作为一种方法的完善与成熟则得益于经济学的成本-收益分析法。

社会学的历史起因

追溯社会项目评估的萌芽，是伴随着美国政府公共社会项目的出现而开始的。随着20世纪30年代西方经济大萧条的到来，政府承担的社会责任越来越大，包括改善社会和环境条件以及国民生活的素质等，这期间美国的公共服务获得了飞速发展。公共社会项目数量多、内容广、规模大，这就要求有科学的管理方法，不仅满足政府机构对项目实施管理的需要，同时也是满足公众对社会项目的“知情权”，公共项目的决策者、计划者、行政执行者和消费者（公众）在各自领域共同推动了项目评估的技术研究。

作为社会研究的应用，项目评估一开始就与社会科学研究方法的发展密切联系，尤其是社会测量方法和社会实验方法。迈克尔·斯克里文（Michael Scriven，1991）认为：“评估有两个重要方面，一方面是搜集资料；另一方面是整理和甄别资料的相对价值和标准。”如何搜集资料，对社会行为特征进行准确描述，是进行价值甄别的基础，也是社会研究方法的难点。虽然社会科学方法仍然无法做到自然科学方法的有效控制与准确，但社会学家对社会抽样、系统观察、测量、统计分析等技术的研究与发展使得项目评估变得更加有效和可靠。斯蒂芬（S. Smith Stevens）把社会测量分为定类（nominal）、定序（ordinal）、定距（interval）、定比（ratio）四种尺度；把测量定义为“根据规则对物体和事件进行的数字分配”²，奠定了社会测量与评估的理论分析框架。心理测量学的发展也为社会项目评估提供了方法路径，信度和效度的定义成为项目评估的重要内容，因素分析被广泛采用，以及瑟斯顿量表（Thurstone scaling）建立的题项形式和反应选择模式在评估中的推广等等。评估研究领域的发展还离不开组织社会学中人际关系研究和社会实验的贡献。例如，1927~1932年，在美国西方电气公司霍桑分场进行的劳动生产率实验，以及二战期间斯托福及其同事对不同军种士兵士气的评估与监测（参见专栏1-2），都使得评估研究技术日臻完善。

1 [美]彼得·罗西等. 项目评估方法与技术(第6版)[M]. 邱泽奇, 等译. 北京: 华夏出版社, 2002. 7

2 [美]罗伯特·德威利斯. 量表编制理论与应用(第2版)[M]. 魏勇刚, 等译. 重庆: 重庆大学出版社, 2004. 7

专栏1-2

人际关系理论研究——社会实验贡献

霍桑实验

“不管是什么问题（有关工人的生产效率），都部分地，有时完全取决于个体工人的态度”。这是组织社会学中人际关系理论的主要观点，1927~1932年在美国西方电气公司的霍桑分厂进行的实验，为这一理论提供了起点。

实验的目的是为了查明，要使雇佣员工的生产效率最大化，需要些什么条件。这一实验的基本假设是，在物质工作条件和工作效率之间存在明确的联系。研究者建立了实验室和对照室，开始改变实验组的照明条件，但人们所预计的物质工作条件同工作效率之间的明确联系没有得到证实。对照效果没有得到体现。要解释这个看起来自相矛盾的结果，工人对企业和实验的态度就作为一种解释项纳入考虑之中。自此开始，实验者的态度对研究项目造成的影响就被称为霍桑效应。

实验所展示的结果非常不符合当时的传统观点，结果在社会学领域内，激励了新鲜思想的出现。……研究者发现，员工对企业的态度取决于个体工人感到被欣赏和重视的程度和他们的要求得到满足的程度等等。这导致了人们在关于组织内部沟通和冲突观念上的变化，因此人们把注意力转向工人内部的团体建设和非正式组织上。在霍桑实验中揭示出的用来解释工作组织中行为的这种范式被称作人际关系理论。

霍桑效应，如同它被观察到的那样，在方法论层面提出了新的视野。在专门文献中，霍桑效应这一术语从此被用来表示实验环境本身对因变量的影响，而不管自变量被赋予的值。

资料来源：〔挪威〕斯坦因·U. 拉尔森. 社会科学理论与方法〔M〕. 任晓，等译. 上海：上海人民出版社，2002. 146~149

军人士气评估研究

在第二次世界大战期间，世界上最伟大的社会科学家之一萨弥尔·斯托福（Samuel Stouffer）及其同事对军人士气做了大量研究。

长久以来，人们认为晋升会影响军中士气。当有人获得晋升而且晋升制度看起来公平时，军人士气就会上升。而且获得晋升的人通常会认为晋升制度公平，但是和晋升擦肩而过的人，则会认为制度不太公平。……

研究者选择了晋升缓慢的宪兵部队和晋升最快的空军特种部队作为研究对象，结果与预期假设完全相反，宪兵部队的士兵认为晋升制度更公平。为什么这样呢？研究者通过深入观察，借助于参照群体理论得到了合理解释。参照群体理论说明，一般人评判自己生活的好坏，并不是根据客观的条件，而是和周围的人相比较，周围的人就构成所谓的参照群体。

军人士气研究对于社会评估方法的贡献在于，人们更进一步认识到对照组在社会实验的重要性，并在评估研究中积极地采用社会调查的方法，包括理论参照和假设检验等。

资料来源：整理分析自〔美〕艾尔·巴比. 社会研究方法（第八版上册）〔M〕. 邱泽奇译. 北京：华夏出版社，2002. 37~38