



公共政策经典译丛

FOURTH GENERATION
EVALUATION

第四代评估

埃贡·G·古贝
[美] Egon G. Guba 著
伊冯娜·S·林肯
Yvonna S. Lincoln
秦霖 蒋燕玲 等 译校
杨爱华

中国人民大学出版社



公共政策经典译丛

FOURTH GENERATION EVALUATION

第四代评估

埃贡·G·古贝
Egon G. Guba 著
[美] 伊冯娜·S·林肯
Yvonna S. Lincoln
秦霖 蒋燕玲 等 译
杨爱华 校

中国人民大学出版社
· 北京 ·

图书在版编目 (CIP) 数据

第四代评估/ [美] 古贝, [美] 林肯著; 秦霖等译.
北京: 中国人民大学出版社, 2008
(公共政策经典译丛)
ISBN 978-7-300-09906-4

- I. 第…
- II. ①古…②林…③秦…
- III. 公共政策-研究
- IV. D0

中国版本图书馆 CIP 数据核字 (2008) 第 173066 号

公共政策经典译丛
第四代评估

[美] 埃贡·G·古贝 伊冯娜·S·林肯 著
秦霖 蒋燕玲 等 译
杨爱华 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	河北三河市新世纪印务有限公司		
规 格	170 mm×228 mm 16 开本	版 次	2008 年 12 月第 1 版
印 张	16.25 插页 2	印 次	2008 年 12 月第 1 次印刷
字 数	250 000	定 价	39.00 元

版权所有 侵权必究

印装差错 负责调换

《公共政策经典译丛》

总 序

政策科学是第二次世界大战后在西方兴起的一个综合性和应用性很强的全新学科，它的出现被誉为当代社会科学和管理科学尤其是行政学、政治学发展的一个重大突破。政策科学以其一系列独特、新颖的范式以及它对公共决策的科学化、民主化和对社会经济发展的巨大促进作用，而备受各国学界和政界的共同关注，成为当代国外社会科学和管理科学中的一个重要的而又充满活力的新兴学科，成为20世纪70年代以后西方公共管理研究领域的一个新方向。

20世纪70年代末80年代初，伴随着改革开放的伟大历史脚步，西方政策科学传入我国。一些高校学者和实际部门的政策研究者注意到了国外社会科学中的这个新领域，着手进行介绍、引进和初步的研究工作；80年代中期，特别是1986年，万里在全国软科学工作座谈会上做了《决策民主化科学化是政治体制改革的一个重要课题》的报告，明确提出要做“政策研究”这一重大课题，促使我国的政策科学研究逐步走上正轨；到了90年代末，我国政策科学发展的重要性和迫切性已被越来越多的人所认识，政策科学的研究与教学逐步体制化，它作为我国社会科学和管理科学研究的新兴领域的地位得以确立。经过了学界与政界二十多年的共同努力，我国的政策科学由无到有，迅速发展，无论是学术研究及学科建设方面，还是在人才培养和知识应用方面都取得了显著的成就。特别是近几年，许多高校、党校和行政学院纷纷开设这一领域的课程，开办相关的本科专业和研究生专业方向以及MPA专业学位，使我国政策科学学科的教学和科研呈现迅猛发展的态势。

但是，我国政策科学的研究与教学刚起步不久，水平不高，仍有不少问题需要解决，有不少薄弱环节亟待加强。一个突出的问题是：学术界对于西方政策科学的引进、消化和吸收的工作做得不够。迄今为止，国内政策科学方面的译著尚为零星出版，种类和数量不多。目前国内学界仍缺乏对西方政策科学理论和方法





的系统了解以及对它的最新趋势的跟踪研究,批判、消化和吸收工作也就难以深入展开。因此,当务之急是要紧密跟踪国外政策科学发展的最新趋势,大胆借鉴其新理论和新方法成果。

我国改革开放和现代化建设事业尤其是市场经济的发展、加入WTO和政治—行政体制改革的深化,对政策科学的研究提出更高的要求。我国公共决策的科学化民主化以及依法行政、依法治国更需要政策—法律的相关知识,这也为政策科学的发展提供了极好的发展机遇。新形势要求我们迅速改变政策科学的研究与教学比较落后、水平不高的局面,克服目前政策科学发展中存在的问题与困难,迅速提高我国政策科学的研究及教学水平,以适应迅速变化的公共管理实践的需要,迎接新世纪的挑战。

政策科学的研究对象是政策实践、政策系统及其运行;它以行动取向,体现理论与实践的有机统一;它的目的和功能是提供政策相关知识,为现实政策实践服务。一方面,政策科学以各国具体的政策实践为基础,它要反映各国具体的政策系统、过程与经验,因而各国的政策科学具有自己的特色;另一方面,现代政策科学是西方的产物,西方政策科学的许多理论、范畴和方法反映了人类政策过程的本质或规律性,属于全人类的共同文化成果,可以为我所用。“他山之石,可以攻玉”。中国政策科学的发展必须既立足于对中国政策实践及现实政策问题的研究,又大胆借鉴西方的政策科学的理论和方法,充分吸收其积极成果。

正是基于上述考虑,我们组织翻译了这套《公共政策经典译丛》,以较系统、全面地反映西方政策科学发展的现状和理论成就,为我国公共政策分析学科的教学和科研提供参考资料与理论借鉴。

《公共政策经典译丛》选取目前在西方(主要是美国)最有影响、最新版本的公共政策领域的教科书或专著,提供当代国外公共政策学科及其分支研究的概貌,并展示出其新的学科框架、研究途径和知识体系。所选教材或专著有两个共同特点。一是学术水准高。作者基本上都是公共政策领域及其分支学科的名家,这些教科书或专著是其主要的代表作或成名作,凝结了作者多年的研究心得与教学经验;它们在西方的政策分析及公共管理学界产生了广泛影响,被经常引证,并被许多大学用作教材或教学参考书。这些教科书或专著大都经过较长时间的检验,有的经过多次的修订或再版。二是理论与实践密切结合,现实感、实践性和应用性强。这些教科书或专著面向的是当代西方的公共政策实践,探讨公共政策实践中出现的新课题;作者用大量的实践经验和案例材料来说明相关的理论问





题；所提出的理论与方法针对性、操作性较强，具有现实的应用价值。

《公共政策经典译丛》具有权威性、学术水平高和实践性强的特点。译丛将展示当代西方公共政策学科的新视野与新途径，它的出版将填补国内公共政策领域的一些空白点，为我国的读者特别是公共管理各专业的师生、研究人员提供公共政策学科的崭新知识体系，为我国公共政策知识体系的创新提供很好的参照和借鉴；它所提供的新理论、新方法以及新的概念框架和思维方式对于我国公共部门的决策者和管理者更新观念、开阔视野和增加理论素养，对于推进我国公共决策的科学化和民主化，具有现实的理论与实践意义。我们相信，《公共政策经典译丛》的问世，必将有力地推动我国公共政策及公共管理学科的教学与研究的发展，并对市场经济条件下和“入世”背景下的我国公共政策与公共管理实践产生积极的影响。

《公共政策经典译丛》编辑委员会

2003年7月



我仅知道社会的最高决策权不在别处，就在人民之中。如果我们认为人民不够开通，难以行使理智的决策权，那么，补救的方法不是将此决策权从他们手中收回，而是给他们指明方向。

——托马斯·杰斐逊 (Thomas Jefferson)

1820年9月28日写给威廉·查尔斯·
贾维斯 (William Charles Jarvis) 的信

评估是对人民和进步的投资。

前 言



公共政策经典译丛

7 我们殷切地期待着读者读完此书后，会有一种截然不同的全新感受——它不同于您所读过的其他评估类书籍。在本书中随处可见评估学的发展历程，但我们并不是按这一历史过程来阐述这门学问的。此外，我们也时常简单而有选择地提到一些模型，但我们并不采用文献中大量存在的不同模型来进行我们的评估工作。随着对技术尤其是统计技术的强调，人们常常希望把评估当作咨询工作的一个技术过程。但我们并不这样认为，尽管我们也处理过程问题，也会提出我们认为是有效判断良好品质的标准。也许最令人惊讶的是，我们没有把评估当作一种纯粹的科学过程，因为我们确信这样做会完全忽略了评估学的社会、政治、价值取向等基本特征。

8 我们试图定义一种全新且成熟的评估方法，这种方法超越了纯粹的科学范畴（即仅为获得事实），涵盖





了人性的、政治的、社会的、文化的以及其他各种相关的因素。我们将这种新的方法称为**第四代评估**。创设这个概念是为了表明它大大超越了已经出现的过于注重测量、描述与判断的前几代评估方法，它将评估上升到一个以谈判协调为核心动力的新高度。这一新的评估方法有下述显著的特点：

第一，这种评估方法的基本立场是，评估行为的最后产出并不是对“事情是什么”、“事情如何进行”以及事物的某种“真实”状态进行描述，而是提出有意义的解释，即个体或者群体行为者为“理解”自身所存在的环境而作的建构。评估结果并非终极意义上的“事实”，而是由包括评估者（以确保客观公正）以及由于评估而处于风险之中的利益相关者通过互动而实际创造的一种结果。从这种互动过程产生的，是关于案例之真实性的一个或多个建构。

第二，这种评估方法认识到，人们借以理解自身环境而形成的建构在很大程度上是受建构者本身价值观影响的。当然，如果所有的建构者都拥有同样的价值系统，那么就不会存在什么问题。确实，在这样一种情况下，这种共同的建构可能是事件的真实状态，而且因为所有人都具有共同的信仰，所以不会造成任何损害。但事实上，每一个现代社会（包括那些通称的“初级”社会，即所谓的第三世界国家）都具有多元的价值观。因此，哪些价值观应得以考量和如何协调这些不同的价值观将成为最重要的问题。很明显，那种声称“价值中立”的方法论在这种情境中不起作用。

第三，这种评估方法认为，上述这些建构必然与具体的物质、心理、社会和文化相关联，正是在这些关联当中形成了这类建构，同时在这类建构中也会提到这些关联。这种关联构成的正是建构者所生活的“环境”——也是建构者努力理解的“环境”。同时，这种环境又是无形的，除非身处其中的人们给这种建构加上了具体的规范、特点和限制。从字面上看，这种关联赋予人们所形成并持有的建构以生命力，而这些建构反过来又赋予这种关联以意义。既然这些人彼此之间也构成了这种关联的一部分，也就不奇怪经过一段时间后他们会形成共同的建构；而那些处于特定的关联当中的人也会对这一关联的性质产生共识。但是，共识并不表示所赞成的事情发生的可能性更大，而只意味着那些达成共识的人已经拥有了一种共同的建构，这种建构对他们来说就是客观现实。而且，共识是人们





在特定环境下交换概念的结果；因而，它并不能够帮助其他人去理解他们自身的环境。最后要说明的是，正因为这些建构都是人类思想的产物，即使是共同的建构也无法避免人们出错。没有人可以提供一种绝对“真实”的认识，即使是近乎“真实”也不太可能。

第四，这种新的评估形式认为，评估可以以各种方式给予或者剥夺利益相关者的特权，因此，与评估的设计与实施过程相关的利益相关者可以选择是否参与评估。如果仅仅是评估者和委托者有权决定调查哪些问题、使用何种手段和数据分析模型，以及做出何种解释等，那么很有可能剥夺其他利益相关者追求合法利益的权利。考虑到评估者和委托者之间通常的那种“暧昧关系”（cozy relationship）[感谢迈克尔·斯科里文（Michael Scriven）提出的这一术语]，就不会诧异于为何评估通常更容易揭露利益相关群体而不是委托者的低效或不法行为（比如老师常常被要求为学校公认的糟糕状况负责）；也就不会惊讶为何委托者的目标具有优先于利益相关者利益的权利（比如，在评估干预民事纠纷的警务程序备选方案时，通常采用惯犯案件减少率而不是家庭暴力案件减少率作为指标，就是因为前者是警察所渴望的状况，这能减少他们的风险、文件工作和出庭次数等麻烦，而后者对他们的工作负担几乎不会产生即时的影响）。通过有选择地公布评估结果能够给予或者剥夺利益相关者的某些权力。如果拥有信息本身就是权力，那么隐瞒信息就是对这种权力的削弱。如果委托者可以决定公开哪些信息、对谁公开、何时以及用何种方式公开，那么，这一规则明显地有利于那些早已掌握这些信息的人，甚至有助于扩大他们的权力，从而剥夺那些相对处于信息劣势的人的权力，甚至是知道得越少越无权。

10

第五，这种评估方法提出，评估必须具有行为导向性（action orientation），即要定义一系列的工作流程，激励利益相关者遵循这一程序，并且成立和维持一个自己的委员会来督促执行。评估工作者或许在自己同行心目中是“纯粹”的科学家，但他们通常都不屑于做后续工作，即评估工作本身的应用研究。种种有关提高评估利用率的相关文献都证明了这种后续工作确实很少进行。这一专业不被使用的事实令人震惊：不断有人责备，委托者不认同所推荐的评估之可信性，而非理性地一意孤行，其结果导致了评估失败，或者责备评估者不善于有效地“推销”评估产品。的确，在美国评估协会（American Evaluation Association）最近所做的一



次演讲中(1987),新任的会长迈克尔·昆·帕顿(Michael Quinn Patton)也督促协会成员要采取推销人员的姿态和技巧;当然他也告诫说,成功推销的第一要诀是“有一个值得自己信赖的产品”。通常,评估“产品”都有一套只符合评估者和委托者目的的赞美之辞,而很少注意其他利益相关者的合法权益,不关心其他相关群体提出的问题和争论,仅仅只是反映他们自己的价值观,尽管这只是整个情境中的多元价值观的一种。如果要达成一种大多数利益相关者都同意的行动计划,则必须在尊重各种价值观的基础上,通过不断地协商谈判使得他们都找到支持这一计划的理由,并积极地去实施该计划。如果要实现这一目标,那么评估人员必须担任一个更大的角色,而不只是信息收集的专家;评估者应该是这一协调过程的总指挥,这一点在最后的分析中将成为整个评估的实质所在。

11 最后,第四代评估坚持认为,由于评估是牵涉许多人的活动(比如委托人、利益相关者、信息提供者以及以其他方式参与评估的人),因此评估者有责任在与上述人员交往时尊重他们的身份、相信他们的诚信、保护他们的隐私。当然,传统的评估者也十分重视伦理问题,他们制定了各种关于维护信息的完整内容、防止利益受损、杜绝欺骗和保护个人隐私及秘密等行为规范,并充分理解、最大程度地执行这些规范。但是,前面所提到的尊重身份、诚信原则和隐私保护的全部内涵已经大大超出了上述传统规范的范畴,它提倡一种**全面的积极参与**,要求利益相关者和其他相关人在评估当中处于平等地位,无论是评估方案的设计、实施、解释还是最后的结论阶段,都要求以政治上完全的平等和控制来对待他们。这就意味着所有评估参与人都有权分享彼此的理解,并努力形成一种公认的、符合常理的、信息量更大的、成熟的共同建构——他们实际上拥有全面的**观念性平等**。当然,这也意味着在整个过程中,参与者都被当作人来看待,而不是实验的对象或研究的主体。

如果这一新的评估概念能够实现上述意图和承诺,就必然迫切需要一种新的评估方法:这种方法能够识别调查结果的建构本质,能够全面考虑到不同的价值观和不同的关联(包括物质的、心理的、社会的和文化的),能够给予利益相关者能力和权力,能够将评估行为和后续工作融合成一个不可分割的整体;而且这种评估方法具有充分的参与性,因为它使所有利益相关者在政治和理念上都享有充分的平等。我们认为,我们所提出的第





四代评估模式适合上述挑战，至少达到了近似的初级阶段。当然，这种方法还需要进一步的发展。

正如我们即将说明的那样，第四代评估依赖于两个因素：第一个因素是**响应式聚焦**（responsive focusing），即在利益相关者参与的基础上决定要解决什么问题和收集什么信息；第二个因素是**建构主义方法论**（constructivist methodology），即在建构主义范式的本体论和认识论的假设基础上实施整个研究程序。^[1]

从发展历程来看，评估学主要关注以下几个方面：收集有效测试中的分数与均值，这些通常都是标准化的；决定测量绩效是否与预先制定好的目标相符；分析要做出的决策，收集需要告知的信息；测定一个评估对象的成本—收益效用；评估一项创新或者调解的效果，不论这种效果是有意的还是无意的；确认前期未满足的需求的实现或者改善程度；等等。上述所有方法都是基于这样一种假设前提，即认为委托人和评估者之间存在亲密的合作关系，其中由委托人指定（或至少是批准）所要采用的检测方法、要达到的目标、要做出的决策、要实现的成本—收益率、要试验的调解或者创新以及需要满足的需求等；但这种合作关系很容易忽视整个情境中的其他相关人员，尤其容易忽视他们的不同价值观和解释。罗伯特·斯泰克（Robert Stake）将自创的评估方法命名为“响应式评估”（responsive evaluation），这种方法强调：所有因为评估而承担了风险的利益相关者都有权公开讨论自己的主张、焦虑和争议，并且可以要求获得答复，无论他们持有什么样的价值观念。这种评估方法与以前通常使用的评估方法相反，因为以前的评估方法都认为评估者和委托人拥有足够的信息和合法权益，由他们双方就能完成评估的设计与实施而不必向其他团体咨询（尽管事实上也常常出于政治原因而进行这样的咨询）。12

支持第四代评估的第二个因素是建构主义方法论。过去，评估工作所使用的方法几乎都是科学式的，也就是实证假设的本体论（即认为存在一个由永恒不变的自然规律支配的客观事实）和矛盾假设的认识论（即坚持观察主体与观察客体二分法，观察主体能够完全站在客体范围之外进行观察，既不影响客体，也不为客体所影响）。依据这些假设，评估者就能断言自己找到了事实，即事物究竟是什么和它们究竟如何运作，而丝毫不用考虑评估者自身、委托人以及其他人的任何可能的影响。



第四代评估与传统观点截然不同。它的假设前提是，事实并非客观显然地摆在那里，而是依据人为的建构而存在的，并常常受到各种各样的社会和文化因素的影响，正是这些因素形成了共同的建构。但是，这种社会学意义上的“共同建构”并不等同于实证主义的“事实”；实际上，没有比这种由试图“理解”自身环境的人们所创造出来的共同观念更真实的“事实”了。很显然，这种社会建构的事实没有独立于“观察主体”（这里即指“建构者”）之外，而是绝对依赖于建构者（她或他）本人的存在。

反映上面这一假设前提的评估方法必须建立在一种截然不同于实证科学的质疑式研究范式的基础之上。我们认为这样一种范式已经包含在我们称之为建构主义的范式当中（这种范式，从不同的角度来看，也被称为解释学范式，有时也被错误地当作定性分析范式）。这种范式在本体论上更遵从于相对主义而非现实主义；在认识论上更多地体现了一元论和主观主义，而不是二元论或客观主义。这种范式使用的是解释学辩证的方法，建立一种评估者和利益相关者的互动模式并确保质量，并由此得到共同的评估结果。而且，与传统的评估方法明显有别的是，此种评估产品并不是一系列的结论、建议或价值判断，而是关于那些主张、焦虑和争议的协商程序，这些通常都是在辩证学解释交流中没有得到解决的问题。

下面是本前言所要做的重要声明，尽管写得很简要，而且几乎得不到支持，但全是为本书考虑的。在写作本书时，我们实际上拥有明确针对的阅读对象，而且我们也要劝说一些读者不必来阅读本书。

首先，也是最重要的，我们称评估者为专业团队。如果说过去评估并不像一个专业的话，那么近十年来，这应该已经是显而易见的了。这首先表现为两个独立的专业协会的成立——评估网络和评估研究会；三年前这两个协会合并成了现在的美国评估协会。另外，人们还设计了多种评估标准体系来指导评估者的工作（同时也使委托人和使用者清楚评估流程），其中最有影响的是《教育规划、项目及资料评估的联合委员会标准》（*the Joint Committee Standards for the Evaluation of Educational Programs, Projects, and Materials*, 1981）以及相应的《个人评估标准》（*Personnel Evaluation Standards*, 1988）。此外，大量的期刊随之产生，¹⁴比如《评估实践》（*Evaluation Practice*）、《评估评论》（*Evaluation Review*）、《教育评估与政策分析》（*Educational Evaluation and Policy A-*





analysis), 以及《新项目评估指南》(New Directions in Program Evaluation) 等, 这些都迅速积累了丰富的文献资料。我们的经验是, 越来越多的职业评估人员不再只受到传统评估方法的影响。这些专业人员可以从本书中找到可供选择的理论基础以及具体实施上的有效建议。而那些不是把评估当作主要兴趣的研究者和政策分析人士也能从书中找到合适的材料和理念用于重新建构新的研究指导范式。

本书也是为利益相关者而写的, 他们有时并非情愿地被拉进一场评估当中, 却因此而承担了风险。要相信有些利益相关者的主张、焦虑和争议是值得高度重视和认真对待的! 这可是公理。他们扮演着各种各样的角色, 其中有各种代理人, 如策划、实施、管理、资助、监督以及其他对评估事业的发展、建立和运作有所作用的人; 也有假定的受益人(也称作“目标人”), 他们会直接或间接地从评估行为中获益; 还有和这些人相关的间接受益人(secondary beneficiaries), 比如戒酒者的同伴、孩子要在学校补习的父母, 等等; 还有一些是受害者, 他们由于评估而直接或者间接受到损害, 或者被剥夺了某些权益, 包括先前的某些机会。利益相关者常常注意不到某项对自己有影响的评估正在进行, 某些对自己可能十分重要的信息被隐瞒了, 或者只有那些与其他群体(比如竞争对手)有关的问题才会受到重视。他们通常都不知道自己在评估活动中应该扮演什么样的角色, 甚至都不清楚自己完全有权利要求参与评估过程。我们恰恰希望能让这些利益相关者明白自己在评估中的地位、自己可以提出哪些问题, 并使他们坚信自己有权成为评估公司的完全合作者, 不论是名义上的还是实际操作当中的。

我们也希望本书能够引起评估委托人或赞助商的兴趣, 他们委托进行评估并(或者)提供评估的经费和实施评估所需要的其他资源。大多数委托人和赞助商实际上并不清楚自身希望从评估当中获得什么; 我们则希望本书可以给予他们一些明确的提示。但是, 我们更加希望能够说服委托人或者赞助商, 请他们相信, 那种传统的评估模式, 即给委托人和赞助商保留实权和决策权, 不仅仅在道义和商业伦理上是错误的, 而且在政治上过于天真, 在概念上过于狭隘。传统的评估信奉“委托人和赞助商永远是明智的”, 却丝毫不顾公众的意见; 的确, 评估活动常常倾向于要说服利益相关者相信: 一项评估的全部要点就是收集足够的信息, 以使“早已存在



于幕后”的决策更加合理化和合法化，但实际情况却恰好相反。在这里，公众常常被征求意见却很少真正得到重视。如果本书不只是提供一些有用的信息，而是能够毫无保留地揭示传统评估方法的缺陷，并且能够达成政治上的共识，那么就可以说我们写作的目的达到了！

我们希望能够对上述三类读者产生作用，但也不得不声明，我们相信有些人是不会从这本书中获益的。我们称那些坚信自己已经拥有了管理一项评估的方法的人为“真正的信徒”。迈克尔·昆·帕顿曾提到过他在一个加勒比村庄指导评估时的经历。当时是夏季最热的时候。一天，因为调查需要他来到甘蔗地里，村民们正在砍甘蔗——你肯定能想象得到这是最消耗体力的劳动。为了尽量凉快些，村民们不论男女都脱掉了衣服，只剩下一条缠腰布。突然间，就好像接到某种信号一样，妇女们扔掉手中的大砍刀，飞快地跑到田边放东西的地方，迅速穿上宽大的长罩衫或者任何可以遮住胸部的东西。帕顿是直到明白了她们这一奇怪行为的原因才真正注意到这一现象的。原来是当地的牧师开着他的吉普车从田边经过，耳尖的村民在帕顿意识到这点之前就听出了是牧师的车子。妇女们清楚地知道如果牧师发现她们袒胸露背的话，会招致他的一顿恶骂甚至是惩罚，因此她们决定赶紧穿上衣服。但是牧师刚一离开，她们就立刻脱掉外衣，重新开始干活！

16 帕顿指出，由于这些甘蔗收割者除了信奉天主教之外还信奉巫术，对他们来说穿不穿衣服并不是一件很困扰的事情，只是为了适应环境罢了。与之相反，牧师则是一个真正的信徒，他无法容忍那些不为自己的信仰体系所宽恕的行为以及这些行为背后的信条；在他眼里，这些行为天生就是不对的，必须遭到谴责和惩罚。

我们认为，实证主义的信徒就像那个牧师一样。他们坚信存在着由自然法则支配的单一的、真实的事实，这些事实等待着科学的方法去发现和利用。实证主义者拒绝一切相对论的观点，比如建构主义。他们认为这种观点不仅完全错误而且遗患无穷、令人讨厌。他们甚至认为那些信奉建构主义的人不过就是一些骗子、奸商而已。因此，从这个意义上讲，这种“真正的信徒”不大可能从本书中发现任何能对他们的观念有所触动的东西。然而，相信相对论的建构主义者，尽管也会不太赞同本书中的实证主义论述，但却能把本书当作某种可能的建构而接受下来。建构主义者会发





现实证主义的论点并不是错误的或者虚假的，只是它们孤陋寡闻、不够成熟罢了。

当然，我们并不认为本书对所提出的问题提供了“正确的”或者是“最后的”答案，也不认为本书解决了“评估是什么”和“评估如何进行”的问题。我们只是把自己的工作当作对评估这门学问的另一种建构，希望读者会发现本书的论述合理而且条理清晰，尽管它绝非普遍真理。的确，并没有什么普遍真理，但我们的建构却或多或少是对真理的一种接近。我们相信，对于我们所说的持续的、辩证的对话将导致具有更大力量 and 价值的重新建构，但很难产生更高的真理。在最后的分析中，所有这些建构，包括那些我们已经提到和将要提到的，都是人类心智的产物，因此难免会有某些消损进取心的错误和缺陷。但如果我们都坚守职业道德的义务，事情就会变得简单。我们将继续尽力寻求每一种意见和批评并对之负责，也会努力去处理这些意见和批评，即使为此要彻底放弃我们现有的建构去接受另一种完全不同的新观念（我们认为这就是所谓的“范式革命”）。

为此，我们欢迎任何有意义的交流，不论对方秉持何种信仰体系，只要他们认同解释学的辩证法，对其他的观点持开放的态度，并愿意随着知识的增长和不断成熟而完善自我，向着那种被认为是有所收获的交流的方向前进。同样，我们也将促使自己朝着这一方向努力。 17

我们已经讨论过，按照目前的实证主义与建构主义的定义，这两种信仰体系之间无法协调；我们看不到二者互相适应的任何可能，除非其中一种范式以自身论点的绝对力量彻底推翻另一种范式，或者出现一些起补充作用的范式，再或者就是能够证明一种范式是另一种的一个特例。但我们看到了这样一种可能，即，通过一个解释学的辩证过程，一种新的建构将会产生，这种建构不能说比前两者“更好”或“更正确”，而只能说是比它们在理论上更成熟、在知识范畴上更广阔。我们相信，这样一种交流代表着“理念的学术市场”这一术语的最佳阐释。

遵从这种精神，我们希望读者不要带着从书中找到进行评估的“公式”的目的来阅读此书；我们也不希望有人看完此书后会说，“我们现在知道要做什么，也知道该怎么做了。”玛丽·赫西（Mary Hesse, 1980）早就敏锐地指出，正如所有的科学理论迟早都会被证明是假的一样，我们现在所接受的任何理论都逃不出同样的命运。我们并不想建立一种新的正





统学说，也不希望我们的读者利用本书去做这样一件事情。

下面我们将简要地概括一下全书内容：

第1章首先从历史的角度回顾了评估学的发展，我们称之为前三代评估方法；然后分析了对这些早期的方法进行过的批判，进而指出，罗伯特·斯泰克提出的响应式评估法很好地解决了这些方法中的大部分问题；而第四代评估方法则作为有效协助响应式评估的一种形式而被引入。

第2章着重讨论对新范式的需要。我们简要阐述了实证主义哲学的问题，并将建构主义作为一种有用的替代方案。

18 第3章具体解释了建构主义范式的内涵，并比较了它和实证主义范式的基本信念体系的区别，引用了大量的文献和论证来支持放弃实证主义而采用建构主义的观点。我们还检验了作为协调实证主义和建构主义两种范式的各种方案，最终证明它们是行不通的。

第4章探讨伦理学和政治学在评估中的意义。我们从实证主义和建构主义两种范式的优势来考察伦理的和政治的问题，最后的结论是，两种范式在这些领域都有一定的困难，但建构主义范式似乎能够具备更多的优势，它至少可以赋予所有利益相关者以权力和投票权。

第5章转而详细分析解释学的逻辑论证过程。我们首先讨论了建构是一种创造出来的事实，然后探讨了建构是如何形成、受到挑战、发生变化并达到妥协的。我们还说明了解释学的逻辑论证有效性的条件，并详细考察了论证过程的各个步骤。

第6章探讨范式是否会影响方法论的问题。方法论或者使用的工具和技术（方法）的整体战略问题，通常由于方法自身而变得混乱不清。这是因为，举个例子来说明可能更清楚，比如建构主义有时被称作“定性分析”的范式，使得这种范式好像更偏好于使用定性分析的方法似的。这一章具体区分了所谓的“第一种自然主义（建构主义）评估”（这种方法其实并没有超出实证主义范式中的定性评估范畴）和“第二种自然主义（建构主义）评估”（这种方法来自于建构主义的基本信仰体系，因此被当作一种十分不同的方法论）。同时我们在某些方面详细探讨了传统范式和建构主义范式的方法论，以使它们之间的区别更加清楚。

第7章阐述第四代评估法所偏好的方法论。读者会看到每一个具体步骤是如何实施的，包括合同的签订；识别利益相关者群体及他们的各种主