
实证理论篇

- 第一章 第三方评估的厘清
- 第二章 第三方评估的意义
- 第三章 第三方评估的原则
- 第四章 第三方评估的理论基础
- 第五章 政府绩效第三方评估的实证理论
- 第六章 政府购买公共服务第三方评估的实证理论
- 第七章 精准扶贫工作成效第三方评估的实证理论

第一章 第三方评估的厘清

第一节 第三方评估概述

一、第三方评估的概念

第三方的概念是由美国莱维恩教授最早提出的，他认为以往将社会组织划分为政府公共组织和私营组织太粗陋，忽略了很大一批处于政府公共组织与私营组织之间的其他社会组织。^① 这些社会组织所从事的活动有别于政府和私人企业的活动，或者从事政府和私人企业都不喜欢做或者做不好的事。这些组织就是第三部门也叫第三方。美国约翰·霍普金斯大学第三部门比较研究中心认为凡符合组织性、民间性、非营利性、自治性和志愿性等五个特性的组织都可被视为第三方。^② 由于“第三方”与“第一方（委托方）”、“第二方（承担方）”都既不具有行政隶属关系，也不具有经济利益关系，所以一般也会被称为“独立第三方”。多数情况下是由独立于政府的，包括社会组织、专家团队、专业的评估机构或研究机构、公民个体等作为第三方的主体组成。其独立性、专业性、权威性比较可信。

关于评估的概念，彼得·罗西在其著作《项目评估——方法与技术》一书中提出，“从广义来看，评估的定义涵盖所有探讨事件、过程或人的价值的努力活动。”^③ 艾尔·巴比在《社会研究方法》中认为，作为一种应用型研究，评估研究的是它进行社会干预的结果和效果问题。^④ 总的来说，他们都认为评估的核心、内容是对评估客体的价值进行判断和评价，或是对社会干预的效果进行考察和研究。因此，评估是指由特定的组织或个人（评估主体）对指定的对象（评估客体）依

① 余梦. 我国第三方政府绩效评估的独立性研究[D]. 成都: 西南财经大学, 2009.

② 邓彩霞. 第三方政府绩效评估主体再认识[J]. 社科纵横, 2015, 30 (2): 49.

③ 彼得·罗西等. 项目评估巧法与技术[M]. 北京: 华夏出版社, 2002: 241 - 253.

④ 艾尔·巴比. 社会研究方法[M]. 北京: 华夏出版社, 2005: 65 - 78.

据某种目标、标准、技术或手段（评估度量尺度），按照一定的程序和方法（评估方法）进行分析、研究、比较、判断、评估和预测其效果、价值、趋势或发展的一种活动（评估活动），是人们认识、把握事物或活动的价值或规律的行为。^① 评估的类型可以依据不同的标准，划分为以下几种类型。根据评估的时段不同，分为前期评估、中期评估和后期评估；根据被评估对象的特征和评估目标不同，分为政府绩效评估、企业绩效评估和非营利组织公益评估；另外还有学者提出形成性评估、累计性评估、独立性评估、参与性评估等评估类型。作为一种管理工具和监督手段，评估的作用和影响力日益凸显，客观科学的评估能够使管理层在政策制定和执行过程中充分考虑到社会、政治、技术和金融等各方面因素，让不同的利益相关者参与其中，推动制定出更好的政策，更有效地实施这些政策，帮助管理者诚实地向社会呈现和报告这些政策的执行效果和影响。

根据评估结果的应用，可将评估理论的发展划分为三个阶段：首先是强调评估结果使用阶段（1970—1986年）：此阶段对评估结果的使用主要关注其对被评估者决策的直接影响，在部分研究中仅意味着评估建议是否被采纳。Leviton（2004）对这种评估的批判在于其有时会依赖有缺陷的证据标准，尤其是不加批判地使用自报告测量而缺乏三角验证或其他能够确认行为变革的方式。Alkin, Daillak 和 White（1979）认为，一个项目的影响效应可能需要几个月甚至几年才能表现出来，因此评估不应选择短期参照点。Cousins 和 Leithwood（1986）将评估结果定义为与评估产出相关联的任何信息，并将评估结果的使用划分为4种类型：用于支持独立决策、用于教育决策者、构成心理过程的一部分以及其他潜在应用。其次是强调评估结果效用阶段（1987—2000年）：此阶段的显著特点在于对评估结果不再仅仅关注其直接效应，而更关注情境，即评估结果如何与其他影响决策的要素相互作用，评估者又如何介入这一作用的过程中。这种观点一定程度上推进了参与式与协作式评估的进展。Patton（1988）主张评估者应通过提供满足评估专员需求的信息与流程，积极介入对评估结果的使用。这一阶段的另一重要影响来自于“过程应用”概念的提出，即评估中进行参与获得的经验与最终的评估建议和报告会影 响人们的行为。这一概念促使评估中平等权力关系、组织学习和利益相关者赋权等观念的产生。最后是强调评估结果影响阶段（2001年至今）：2000年以来，评估研究者开始讨论评估影响的框架。评估影响的概念包含以往的主要观点，但提供了一个更为广

^① 李志军. 第三方评估理论与方法[M]. 北京：中国发展出版社，2016：3.

泛的视角，即不仅强调评估中有意识活动的效应，而且考虑超出目标之外的非预期效应，非直接目标（如组织学习与赋权）以及评估者通过其工作对社会产生更多的贡献。这里的影响包括“所有的可能引致或偏离更佳社会效应的评估结果”，它可能是预期的、直接的、正向的、短期的、显在的，也可能是非预期的、间接的、负向的、长期的、隐含的。

第三方评估起源于西方国家。1906年，Bruere等人发起成立了纽约市政研究院，1907年，该研究院率先开始了对纽约市政府进行初步的绩效（效率）评估。这是能够检索到的最早的绩效评估、绩效测量、绩效核算的实践，它本质上是由一个外在于政府、以探究政府管理为宗旨的研究机构发起的“第三方评估”。^①因此，第三方评估的概念是与政府绩效管理、政府绩效评估联系在一起的。对于第三方评估，我国目前没有统一的定论，学者们的看法不一。彭国甫认为：第三方评估又称社会评估，主要包括公民个人、社会团体、社会舆论机构、中介评估机构等通过一定程序和途径，采取各种方式，直接或间接，正式或非正式地评估政府绩效。^②程样国、李志认为：第三方评估是指由政策制定者与执行者以外的人员进行的正式评估。包括受行政机构委托的研究机构、专业评估组织（包括大专院校和研究机构）、中介组织、舆论界、社会组织和公众，特别是利益相关者等多种评估主体。^③包国宪教授将第三方评估的概念解释为：第一方评价是指政府部门组织的自我评价；第二方评价是指政府系统内，上级对下级做出的评价，这都属于内部评价。而第三方评价是指由独立于政府及其部门之外的第三方组织实施的评价，也称外部评价，通常包括独立第三方评价和委托第三方评价。^④石富覃认为：所谓第三方评估就是指由与政府无隶属关系和利益关系的第三部门和民间机构所组织实施的评估政府及其部门绩效的活动。^⑤

由此可见，学者们更多地认为，第三方评估是一种区别于政府内部自评的外部评估方式。是由具有专业性的机构，按照一定的评估标准，采取科学的评估方法，对政府、社会组织或者社会服务项目进行评估，得出具有说服力的评估结果与报

① 尚虎平，王春婷. 政府绩效评估中“第三方评估”的适用范围与限度——以先行国家为标杆的探索[J]. 理论探讨, 2016 (3): 12-18.

② 彭国甫. 对政府绩效评估几个基本问题的反思[J]. 湘潭大学学报(哲学社会科学版), 2004 (3).

③ 程样国，李志. 刍议第三方政策评估对我国的启示[J]. 行政与法, 2006 (3).

④ 包国宪. 我国第三方政府绩效评估组织的自律实现问题探析[J]. 中国行政管理, 2008 (1): 49-50.

⑤ 石富覃. 政府绩效评价组织的自律与监管研究[D]. 兰州: 兰州大学, 2006.

告，提供给决策者作为主观价值判断的一系列工作活动。第三方评估实质是一种更客观的社会监督，是要在政府管理运行机制中引入多元参与，更多更好地发挥社会组织参与政府决策、政策执行、效果评价、公开监督等方面的作用。评估活动完全由第三方负责主持和实施，独立决定评估的标准、内容、测量、数据处理、结果发布等各个环节。第三方评估通常包括独立第三方评估和委托第三方评估。从设立方面看，独立第三方的设立是组织依法或者自发设立的，而且评估过程都是组织自发进行，跟政府没有隶属关系，主要包括中介组织、受行政机关委托的考核机构、社会组织和公众、专业考核组织（包括大专院校和研究机构）、特别是利益相关者等多种考核主体；委托第三方评估是政府委托第三方成立绩效评估组织，组织与政府只存在委托与被委托的关系，没有政治和经济利益上的直接关系。在经费来源上，独立第三方的经费则来源于与政府无任何直接利益关系的组织，如基金会、营利性组织或个人的捐赠，还可来源于组织内部成员的收费等；委托第三方评估收取的费用是政府部门提供的。

美国著名学者尼古拉斯·亨利对第三方评估给予了高度评价，认为“大多数评估专家赞同评估由组织中负责项目的最高层委托和支持，为了制衡起见，由外部机构或是第三方来进行评估也是可取的。”^①从美国第三方评估的实践来看，第三方不仅在利益关系上与作为政策制定者的政府以及相关政策制定者保持了距离，更主要的是他们在评估中的专业性受到人们的推崇。^②

二、 第三方评估的内涵

可以从以下几个方面理解第三方评估的内涵：^③

从评估主体上看，第三方评估是指由独立于政府及其部门之外的第三方组织实施的评价。^④也就是说，第三方评估主体相对于政府是独立存在的，这些主体可以是社会组织、高等院校、民间智库、民意评价机构等。

① [美] 尼古拉斯·亨利. 公共行政与公共事务[M]. 张昕, 等, 译. 北京: 中国人民大学出版社, 2002: 320.

② 尚虎平. 非政府评估的政府管理效果: 美国高校基于利益相关主体参与的第三方评估对我国的启示[J]. 内蒙古社会科学, 2007 (5).

③ 吴佳惠、王佳鑫、林誉. 论作为政府治理工具的第三方评估[J]. 中共福建省委党校学报, 2015 (6).

④ 包国宪. 绩效评价: 推动地方政府职能转变的科学工具——甘肃省绩效评价活动的实践与理论思考[J]. 中国行政管理, 2005 (7).

从评估对象上看，第三方评估包括政府绩效评估和公共服务评估，也包括政策制定、执行、效果评估，评估的目的是改善和优化政府行为。

从评估结果上看，第三方评估通过问卷调查法、随机抽样、统计技术等方法，对于结果分析采用目标图形转化法、价值澄清、价值评论、交互影响分析、折扣等技术手段，有利于将利益相关者潜在的目标明确化。^①

从评估方式上看，第三方评估倡导公民参与的多样化途径，如通过问卷或访谈、电话调查、现场测评等，评议政府机构及其工作人员。随着大数据时代的来临，还可以是网上评议，通过在线系统填写调查问卷或以其他方式参与评估。

三、 第三方评估的特点

(1) 客观性。独立于政府之外的第三方评估，要求评估者在评估的过程中必须克服感情因素的影响，避免评估活动主观上的随意性。当然，评估本质上就是一种价值判断，涉及价值判断就有很强的主观性。不可否认的是，基于评估者所采用的评估标准和评估的指导思想不同，再加上评估主体在评估过程中不可避免地会受到某些主观因素的影响，特别是涉及评估者与被评者之间的感情因素，这些原因综合起来都很容易导致评估偏离客观性。因此，只有第三方评估主体的立场客观、公正，评估的结果才具有可信度。

(2) 公正性。作为独立于政府和社会特定利益相关者群体的第三方评估，其权力运行往往是为了实现公众的问责和利益保障。公正性是第三方评估运行的根本价值要求，即第三方评估主体要公正地对待所有的对象，保障所有利益相关者的利益，正确看待评估的价值，理解评估的重要性，对评估结果有强烈的责任感，做出客观真实的判断，而不能“变公为私”，狭隘地重点照顾某些群体的利益或忽视其他群体的利益。

(3) 权威性。第三方评估是独立于政府而存在的，它与政府没有直接的利益关系，能够客观、公正地开展评估。它使评估主体更加多元化，评估更加科学全面，评估结果更加有效，有利于政府管理理念的转变，引导政府向服务型政府转变。同时，第三方评估拥有专业的人才队伍，在评估的知识、方法、技能的掌握，

^① [美] 威廉·N·邓恩. 公共政策分析导论[M]. 谢明, 等, 译. 北京: 中国人民大学出版社, 2002: 342-343.

评估工具的使用方面都具有相对的优势，能够科学设计指标体系，分配权重位序，选择科学的评估工具和方法，提高评估结果的精确度和有效度，并能够从较为专业的视角提出具有针对性、科学性的整改方案和措施，帮助政府在实践中不断提高治理绩效。受到政府、社会组织、社会大众的认可，由此建立起了自身的权威。权威性是第三方评估机构自身良好的社会信誉，权威性是建立在独立性与专业性的基础之上的，独立性、专业性、权威性三者是相辅相成的。

第二节 第三方评估的国外经验

起源于西方国家的第三方评估，由于起步较早，发展比较成熟。特别是美国、英国这两个国家在第三方评估方面的发展和完善程度尤为值得我们去研究和学习。

一、美国第三方评估的发展

美国第三方评估的发展是随着政府绩效评估的发展而发展的，是政策评估、绩效管理的重要形式。

19世纪末20世纪初，美国的第二次工业革命进入尾声，经济高速发展，就工业生产而言，已经超越英、德成为世界头号强国。与此同时，美国政府效率低下，甚至腐化无能，政府的权力达到了最大化，政府规模日益庞大，政府支出大幅增长，导致了国家垄断和对社会回应性的下降。在这样的背景下，公众呼吁美国政府采用企业的一些有益做法来提高政府效率，政府也开始重视效率问题。纽约市政研究院1907年对纽约市政府进行的绩效（效率）评估，这是最早的绩效评估实践，它本质上就是由一个外在于政府、以探究政府管理为宗旨的研究机构发起的第三方评估，开创了第三方评估政府绩效的先河，在第三方评估历史上具有划时代的意义。1911年，费城新当选的改革派市长决定聘请科学管理创始人之一莫里斯·库克（M. Cooke）采用成本会计思想对费城的公共事业进行预算规划、管理效果测量。在这种绩效规划、绩效测量与评估的作用下，库克4年内为费城节省了100多万美元的垃圾收集成本，降低了125万美元的公共事业费用，解雇了1000名低效率工作人员，建立了养老基金和社会福利基金，把“市政府管理从烟雾缭绕的房间变成

了阳光明媚的场所”。这就是由政府引入外部第三方智慧的“体制内绩效评价模式”。^① 20世纪50年代,美国在国内推行政策与政府绩效评估,随着评估规模与范围的扩大,委托第三方开展评估逐步成形,出现了一批较为知名的第三方评估机构。

随着社会的进步,民主、公平理念越来越引起人们的关注,受此影响,政策评估与政府绩效评估开始以民主、公平为价值取向,更加重视结果、质量和公众满意度。20世纪六七十年代,美国政府普遍存在着财政危机和信任危机,公民对政府的不满情绪日益高涨,为了消除危机和安抚民众的不满情绪,美国政府实施了一系列重塑政府形象的改革运动。其中最著名、影响最大的是被称为“美国行政改革重要里程碑”的两届“胡佛委员会”的改革方案。该方案改变了以前只重视投入的弊端,强调要注重产出及成本问题,要更加注重政府的管理能力和水平的改善。并且强调在集权控制的同时要适当放权给下属部门,以实现资源的合理配置。胡佛委员会通过绩效预算的方式实现改革的目标,用数字量化的形式评估政府的政策和绩效,使政府的绩效在公众面前清晰的展现出来,让民众对政府有直观的认识。两届胡佛委员会的绩效评估探索,也属于政府之外的第三方评估。1973年,在尼克松总统执政期间,美国政府出台了“联邦政府生产率测定方案”,开始探讨政府绩效评估的规范化。此时的美国政府绩效第三方评估整合了政府预算、政府责任、政府目标,更加重视结果。

20世纪八九十年代,美国面临着巨大的财政赤字,通货膨胀严重,失业率居高不下,公众对政府失去了信心。为了应付危机,改善政府管理,美国进行了一系列的政府改革运动。将私营企业的管理理念、管理方法和管理技术渗透到公共部门,引入市场机制,强调结果导向和顾客导向,关注公民满意度。还颁布了绩效评估的专门法律——《政府绩效与结果法案》。这是美国政府第一次为政府绩效评估设立法律,对政府绩效评估的推动起到了不可替代的作用。在美国,第三方评估机构参与政策和政府绩效评估是美国政府绩效评估的一个显著特点。第三方评估机构提出有关政策和政府绩效评估的理论、方法、指标体系和步骤等,并以这些理论指导评估政府的政策和绩效,有力地推动了政府绩效评估活动的开展。哈佛大学的肯尼迪学院是较早介入政策和政府绩效评估的第三方评估高校。自1984年,肯尼迪学院由福特基金会支持,开始创办“美国政府创新奖”,用于识别和培育政府创新性工作。美国坎贝尔研究所是独立第三方评估的典型代表。坎贝尔研究所经过广泛的调查和深入的研究,建立了一套完善的绩效评估体系。1998年起开始对美国的

^① COOKE, M. L. Our Cities Awake[M]. New York: Doubleday, 1918: 98.

州和城市政府进行绩效评估，在 1998 年与美国《政府管理》杂志合作，对全美国 50 个州政府展开了大规模的绩效测评活动，评估内容涉及政府人事、财政等方面，在美国社会上引起了巨大反响；在 1999 年，对美国财政收入最好的 35 个城市政府展开绩效评估；在 2000 年，坎贝尔研究所进行了第二次对 50 个州的政府绩效评估，并将评估结果与排名一五一十地向社会公开，在美国政府、美国民众心中得到了很高的认可率；2001 年对美国 40 个最大县进行了政府绩效评估。坎贝尔研究所是美国第三方评估政策和政府绩效的中坚力量。美国第三方评估机构发展较充分，除以上两个典型代表外，还有许多著名的机构，如美国的兰德公司、布鲁金斯研究所、现代问题研究所、政策研究所、总会计局等，它们都在某种程度上参与了政府战略、政策的制定以及政府绩效评估活动，它们的参与和介入奠定了政策和政府绩效评估的群众基础，体现了政府重视民意、体恤民情的管理理念，规范了政府人员的行政行为，提高了政府公共部门的管理绩效，对美国政府改革做出了重大贡献。

二、英国第三方评估的发展

在 20 世纪七八十年代经济发展“滞胀”的社会背景下，英国为了提高政府行政效率，改善政府服务水平，采取委托第三方评估的方法来进行政府绩效评估。英国第三方评估的发展，大致分为两个阶段：

一是通过绩效评估提高政府工作效率阶段。这一时期，英国政府信奉“效率至上”理念，政府的一切工作都是以经济和效率优先，追求小投入，大产出。在政府绩效评估中，最著名、影响最大的是雷纳评审，这是第三方评估的典型例子。雷纳评审是以“解决问题为导向”的“经验式调查”。它关注的重点是经济和效率，采用问题诊断的方式，从政府的日常工作中寻找出导致效率不佳的各种因素，包括工作任务、工作方法和工作程序，其目的在于通过提高效率来降低政府公共部门的运营成本。雷纳聘请人员担任评审员来完成评审任务，鉴于各个部门之间的差异性，评审员拥有较大的自主权限来决定评估工作的展开，但基本的评审内容和程序由雷纳确定。首先，评审对象的选取全由评审员自己决定，可以大到一个政府或政府的某个工作部门，也可以小到一个工作程序或某个工作环节，对象选定后由部门公务员承担大量的工作。其次，对正在进行的工作提出质疑，要清楚地了解为什么要开展这项工作；为什么要以现在这种方式开展工作；如何更有效率地完成工作。它强调评审员要深入到最基层、最前线调查，从实践调查中总结出方案，撰写出评估报告。最后在深入调查的基础上，评审员与被评审单位的负责人就有关报告