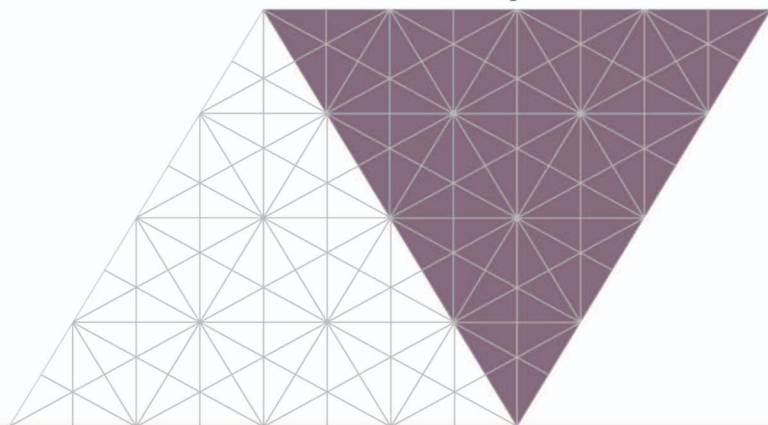


马亮 著

第三方评估、绩效排名 与政府循证管理

Third-Party Evaluation, Performance Ranking,
and Evidence-Based Management of the Government



江苏人民出版社

图书在版编目(CIP)数据

第三方评估、绩效排名与政府循证管理 / 马亮著.

—南京: 江苏人民出版社, 2021. 4

ISBN 978 - 7 - 214 - 23012 - 6

I. ①第… II. ①马… III. ①国家行政机关—行政管理—研究—中国 IV. ①D630. 1

中国版本图书馆 CIP 数据核字(2018)第 289073 号

书 名 第三方评估、绩效排名与政府循证管理
著 者 马 亮
责 任 编 辑 戴亦梁 胡天阳
装 帧 设 计 陈 婕
责 任 监 制 陈晓明
出 版 发 行 江苏人民出版社
地 址 南京市湖南路 1 号 A 楼, 邮编: 210009
网 址 <http://www.jspph.com>
照 排 江苏凤凰制版有限公司
印 刷 江苏凤凰通达印刷有限公司
开 本 652 毫米×960 毫米 1/16
印 张 13.75 插页 3
字 数 186 千字
版 次 2021 年 4 月第 1 版
印 次 2021 年 4 月第 1 次印刷
标 准 书 号 ISBN 978 - 7 - 214 - 23012 - 6
定 价 45.00 元

(江苏人民出版社图书凡印装错误可向承印厂调换)

目 录

第一章 绪论 1

第一节 研究背景与问题 2

第二节 研究综述 4

第三节 本书的核心观点 14

第四节 本书的结构安排 18

第二章 第三方评估的理论视角 24

第一节 “外来的和尚好念经”？ 24

第二节 第三方评估如何影响政府绩效？ 26

第三节 第三方评估何时奏效？ 29

第三章 第三方评估的历史进程与国际经验 32

第一节 第三方评估的起源与演变 32

第二节 为什么第三方评估如此之多？ 36

第三节 国际第三方评估的演变与发展 38

第四章 第三方评估的评估框架与标准 46

第一节 研究问题 46

第二节 政府绩效评估及其主体 48

第三节 中国政府绩效评估的发展 51

第四节 评估框架 55

第五章 外部参与的政府绩效评估:个案研究 58

第一节 政府绩效评估:从内部封闭到外部开放 61

第二节 连氏项目的发起与组织 64

第三节 连氏项目的理论框架、评价体系和研究方法 67

第四节 从连氏项目看中国政府绩效外部评估 74

第五节 结论 76

第六章 第三方公共服务绩效评估:比较案例研究 78

第一节 引言 78

第二节 第三方公共服务绩效评估的两个案例 81

第三节 结果 85

第四节 讨论 90

第七章 第三方公共服务绩效评估:两个案例的纵贯比较 96

第一节 引言 96

第二节 研究案例与理论假说 99

第三节 数据与方法 104

第四节	结果	105
第五节	讨论与结论	111
第八章	第三方评估的多案例比较研究	117
第一节	研究问题	117
第二节	案例选取	118
第三节	研究方法	119
第四节	结果和讨论	121
第五节	结论	126
第九章	作为第三方评估的城市排行榜	130
第一节	引言:进入“榜时代”的中国城市	130
第二节	城市排行榜:定义、历史与研究述评	132
第三节	城市排行榜:一个供求分析框架	141
第四节	城市排行榜的生产	143
第五节	城市排行榜的消费	146
第六节	城市排行榜的问题与流弊	149
第七节	城市排行榜的未来展望	153
第十章	第三方评估与政府绩效信息使用	158
第一节	政府绩效信息使用:研究状况概览	159
第二节	政府绩效信息的特征与属性	161
第三节	政府绩效信息的使用特征	167
第四节	政府绩效信息(不)使用的影响因素	170
第五节	政府绩效信息使用的结果与影响	175

第六节 理论整合与研究展望 177

第十一章 第三方评估与循证管理 185

第一节 何谓“循证”? 185

第二节 循证:从医学到政策 187

第三节 循证决策与管理的证据库 189

第四节 第三方评估与循证决策 191

第五节 第三方评估如何推进循证决策 193

第十二章 结论与展望 198

第一节 第三方评估何去何从? 198

第二节 第三方评估的“内功”与“外力” 202

第三节 第三方评估的推进策略 206

第四节 第三方评估的未来研究展望 209

后记 213

第一章 绪论

第三方评估或外部评估,是由政府以外的第三方机构(如学术机构、媒体、咨询公司、智库、国际组织等)对政府管理与政策绩效进行的评估活动。近些年,国务院和各级政府部门都大力引入第三方评估,用于推动重大政策决策的落实与督查。李克强总理将第三方评估视为政府管理创新的举措,并在许多重大决策领域使用。李克强总理不仅要求中央政府部门委托和使用第三方评估,而且亲自参与评估工作的组织和评估报告的使用,特别关注第三方评估对政府工作的促进作用。各地政府也对第三方评估寄予厚望,如江苏省、广东省等地政府纷纷推出实施意见和支持举措。广东省率先试点财政支出第三方评估,用于推动政策绩效改进。江苏省政府出台法规,规范省内第三方评估活动。此外,还有许多省份和城市加强第三方评估,推动其在多个领域的应用。

与政府部门自评和自上而下的内部评估相比,第三方评估被认为在独立性、客观性和专业性等方面具有相对优势。第三方评估提供的证据,被认为是强化政府循证决策与管理的关键依据,有利于政府学习和绩效改进。但是,“外来的和尚会念经”吗?第三方评估能够担此大任吗?第三方评估的信度、效度和可用性如何?第三方评估结果得到政府部门的重视和使用了吗?目前对第三方评估的制度逻辑、作用机制与绩

效影响还缺少深入研究,而这些问题都值得深入的理论和实证研究。本章对第三方评估的研究背景与问题进行了梳理和提炼,并系统综述了该领域的研究文献。在此基础上,总结了本书的核心内容与观点,并介绍了本书的结构安排。

第一节 研究背景与问题

第三方评估也称为独立评估或外部评估,是由政府部门委托第三方机构对公共项目、政策或部门的绩效进行独立评估。2007年以来,包括兰州大学对甘肃省非公有制经济发展的评价^①、华南理工大学对广东省市县政府绩效的评价^②等在内的第三方评估在中国各地发展迅猛。2013年以来,国务院在督导重大政策的落实情况时引入第三方评估,并被李克强总理视为“政府管理方式的重大创新”。与此同时,在信息公开和财政透明、政府网站绩效、社会组织公信力、高等教育、医院服务、精准扶贫等方面,第三方评估被越来越广泛地使用。江苏、湖南、广东等地政府纷纷出台文件,鼓励通过政府与社会资本合作(PPP)的方式购买第三方评估服务,并在财政支出绩效评估领域推广。

与政府自评和自上而下的考核相比,第三方评估引入“外脑”,可以更客观、独立和专业地评估政府绩效,并为优化决策和改进绩效提供依据。但是,第三方评估真的是“外来的和尚好念经”?为什么政府部门使用第三方评估?什么样的第三方评估是有效的?第三方评估如何促进政策落实和提升政府绩效?在什么情况下第三方评估可以改进绩效?这些问题尚未得到深入研究,并可能使政府在推进第三方评估时收效甚微或误入歧途。

① 参见包国宪、董静、郎政《第三方政府绩效评价的实践探索与理论研究——甘肃模式的解析》,载《行政论坛》2010年第4期,第59—67页。

② 参见郑方辉《第三方评价地方政府整体绩效的实证研究——以广东省市、县两级政府为例》,载《中国行政管理》2008年第5期,第13—17页。

第三方评估无法取代自评或内部评估,使用不当还会导致意想不到的负面影响。目前第三方评估机构良莠不齐且发展滞后,无法做到独立、专业,很难满足政府部门的评估需求。一些部门出于合法性诉求,只是象征性地引入第三方评估以赢得公众信任,实际上难以促进政策优化和绩效改进。许多第三方评估“一评了之”或被“束之高阁”,在绩效反馈和评估使用方面存在欠缺。还有一些政府部门对第三方评估“袖手旁观”,无法“举一反三”,组织学习和创新乏力。第三方评估的投入较大,成本与收益往往不成比例或得不偿失。因此,第三方评估并非灵丹妙药,能否及如何提升政府绩效还有待于深入研究。围绕第三方评估,需要回答如下关键问题:

1. 为什么政府部门选择第三方评估? 哪些因素影响政府的评估选择(内部 VS 外部)?

2. 第三方评估具有哪些特征和属性? 第三方评估需要政府部门的哪些支持?

3. 第三方评估能否提升政府绩效? 通过哪些过程和机制提升政府绩效? 哪些因素和条件会制约第三方评估改善政府绩效?

对上述问题进行理论和实证研究,具有重要的理论价值和现实意义。概而言之,本书的理论价值与应用前景主要包括如下方面:

1. 理论价值。本书独辟蹊径地提出第三方评估影响政府绩效的理论框架,并使用实证方法对其进行研究,填补了本领域的研究不足,有利于理论创新。既有研究局限于就第三方评估本身谈第三方评估,而未能跳出第三方评估的范围去考察其对政府绩效的影响。本书关注回答第三方评估的“评估悖论”(为什么不奏效)和“评估迷思”(如何奏效),揭示第三方评估影响政府绩效的作用机制,以及第三方评估发挥作用的依存条件。本书跳出单纯的政策评估视角,将政策评估与绩效管理两大研究领域有机衔接,从更高层面去研究第三方评估的影响。上述研究为创新第三方评估理论提供了理论框架和经验依据。

2. 应用前景。从实践价值而言,本书有助于消除当前对第三方评估的盲从和迷信,廓清第三方评估影响政府绩效的路径和条件,为政府部

门更好地使用这个工具提供理论依据和实证证据。面对第三方评估这个新生事物,政府部门遇到要不要用、怎么用和如何好用的挑战,还需基于实证研究的价值规范和理论的引导。本书提出的理论框架及实证证据,可以解释第三方评估为什么、如何及在何种条件下影响政府绩效,据此提出的政策建议有利于政府部门合理使用第三方评估并改进政府绩效。

第二节 研究综述

国内研究方面,学者们对第三方评估的定义、特征和优点及面临的挑战等进行了理论探讨和案例研究。^{①②} 第三方评估可以由学术机构、咨询公司、社会代表和民众等参与,并会导致不同的评估模式。^③ 针对已有的第三方评估案例,学者们进行了多角度分析^{④⑤⑥},认为第三方评估在独立性^⑦、自律性^⑧、适用范围^⑨、制度环境^⑩等方面都遇到了难题。有

① 李志军:《第三方评估理论与方法》,中国发展出版社 2016 年版。

② 李文彬、卢扬帆、郑方辉:《财政专项资金绩效第三方评价》,光明日报出版社 2015 年版。

③ 徐双敏:《政府绩效管理中的“第三方评估”模式及其完善》,载《中国行政管理》2011 年第 1 期,第 28—32 页。

④ 包国宪、董静、郎玫:《第三方政府绩效评价的实践探索与理论研究——甘肃模式的解析》,载《行政论坛》2010 年第 4 期,第 59—67 页。

⑤ 郑方辉:《第三方评价地方政府整体绩效的实证研究——以广东省市、县两级政府为例》,载《中国行政管理》2008 年第 5 期,第 13—17 页。

⑥ Yu W., Ma L., “External government performance evaluation in China: A case study of the ‘lien service-oriented government project’”, *Public Money & Management*, 2015, 35 (6): 431 - 437.

⑦ 郑方辉、陈佃慧:《论第三方评价政府绩效的独立性》,载《广东行政学院学报》2010 年第 2 期,第 31—35 页。

⑧ 包国宪、张志栋:《我国第三方政府绩效评价组织的自律实现问题探析》,载《中国行政管理》2008 年第 1 期,第 49—51 页。

⑨ 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》,载《理论探讨》2016 年第 3 期,第 12—18 页。

⑩ 徐双敏、陈尉:《“第三方”评估政府绩效的制度环境分析》,载《学习与实践》2013 年第 9 期,第 22—27 页。

学者基于两项第三方评估案例,对其信度和效度进行比较。^① 还有研究结合财政支出绩效评估,分析了第三方评估的效度、效力和效用。^② 有研究者提出了第三方评估的质量标准,并综合比较了十余个典型案例。^③ 上述文献为理解和廓清第三方评估提供了启示,但是仍然有必要围绕第三方评估是否奏效及如何影响政府绩效,构建理论框架并进行实证检验。在这些方面,国外研究积累了一定文献,并对外部评估影响组织绩效的主要路径和权变因素进行了研究。

一、第三方评估的动因与特征

第三方评估或外部评估是一种评估模式,与之对立的是内部评估。内部评估与外部评估并非泾渭分明,一般依据评估结论是由组织内部还是外部机构得出加以区分。依据评估者和目标客户是内部还是外部,还可以将评估进一步分为四种。内部人指同项目执行及其影响直接有关的个人或组织。外部人指项目以外的实体,但对项目实施及其影响有既得利益。据此可以划分四种评估模式:评估者和目标受众都是内部人(如组织自评),评估者是内部人而目标受众是外部人(如年度报告),评估者是外部人而目标受众是内部人(如委托评估),评估者和目标受众都是外部人(如审计报告)。^④ 如果将委托机构、评估机构和客户三者予以综合考虑,则可以进一步对内部和外部评估予以细分。

内部评估和外部评估各有利弊,并取决于不同的条件,如组织能力、

① 马亮、于文轩:《第三方公共服务绩效评价的评价:一项比较案例研究》,载《南京社会科学》2013年第5期,第55—63页。

② 李文彬、黄怡茵:《基于逻辑模型的财政专项资金绩效评价的理论审视——以广东省人大委托第三方评价为例》,载《公共管理学报》2016年第3期,第111—121、158—159页。

③ Yu W., Ma L., “External government performance evaluation in China: Evaluating the evaluations”, *Public Performance & Management Review*, 2016, 39 (1):144 - 171.

④ Owen J. M., Rogers P. J., *Program Evaluation: Forms and Approaches*, London: Sage, 1999.

评估需求、制度环境、政策议题和组织路径。^① 组织在选择内部还是外部评估时,会综合考虑许多因素,如成本、可及性、项目及其运营的知识、对情境的知识、收集信息的能力、灵活程度、专业技能、客观性、财政问责、批评的意愿、评估的使用、结果的扩散、伦理议题和组织投入等。^②

政府在引入外部评估时如何使其同既有内部评估相契合和兼容,可能是至关重要的问题。内部评估的支持者认为,自评估组织有持续的生命力,特别是当组织同其利益相关者建立了较高级别的相互信任时更是如此。^③ 虽然“外来的和尚好念经”,但恰恰只有组织内部成员才最知晓内情并能执行变革。当外部评估与内部评估相兼容时,外部评估对组织绩效的影响更强。所以,内外评估的契合性或平衡性至关重要。^④ 如果组织需要的是内部评估,却选择了外部评估,可能评估效果会大打折扣甚至适得其反。

二、第三方评估对组织绩效的作用机制

除了关注如何优化评估和促进评估使用以外,更为重要的是如何通过评估改进政府绩效。评估不仅会影响被评估的某个具体项目或政策,而且经由“举一反三”的外溢效应,会对这些项目或政策所在的政府部门产生显著影响。因此,评估与绩效管理之间存在兼容性,评估者与管理者之间的建设性对话,有助于通过评估去改进绩效。^⑤ 结合已有文献,可

① Pattyn V., Brans M., “ Outsource versus in-house? An identification of organizational conditions influencing the choice for internal or external evaluators”, *The Canadian Journal of Program Evaluation*, 2013, 28 (2):43 - 63.

② Conley-Tyler M., “ A fundamental choice: Internal or external evaluation?” *Evaluation Journal of Australasia*, 2005, 4 (1/2):3 - 11.

③ Wildavsky A., “ The self-evaluating organization”, *Public Administration Review*, 1972, 32 (5):509 - 520.

④ Blok H., Slegers P., Karsten S., “ Looking for a balance between internal and external evaluation of school quality: Evaluation of the SVI model”, *Journal of Education Policy*, 2008, 23 (4):379 - 395.

⑤ Nielsen S. B., Ejler N., “ Improving performance? Exploring the complementarities between evaluation and performance management”, *Evaluation*, 2008, 14 (2):171 - 192.

以发现如下中介变量和作用机制是联结第三方评估与绩效改进的关键路径,包括绩效反馈、绩效问责、组织学习和组织创新。

首先,第三方评估为组织提供了专业、独立的绩效信息,而这些绩效反馈有利于组织学习和绩效改进。绩效反馈理论认为,基于评估获取的信息为双环组织学习和持续绩效改进提供了条件,使组织可以对标调适。^①特别是来自组织外部的非常规绩效信息,有利于组织获得有价值的绩效反馈。^②来自第三方评估的把脉问诊,有助于政府发现问题和不足,从而为绩效改进提供靶点。

其次,第三方评估可以强化绩效问责,使参评部门有更强的压力和动力去改善公共服务绩效。^③第三方评估形成的外部监督,以及评估结果的社会公开,都会对参评部门产生问责压力,并驱动其采取措施改进公共服务绩效。^④实验研究发现,如果预期到会有外部评估,人们会努力完成可以胜任的目标。但如果目标难以实现,那么外部评估则会打击人们的士气。^⑤第三方评估通常由上级主管部门委托,被评部门因此还会受到自上而下的官僚问责,这会进一步督促其执行政策和改进绩效。

再次,第三方评估引入了外部力量和新鲜血液,使管理层和组织成员能够获取新的视角和知识,这有助于促进组织学习,^⑥进而提升公共服务绩效。特别是第三方评估可以“举一反三”,使组织成员学以致用,将

① Mausolf C. , “ Learning from feedback in performance measurement systems ”, *Public Performance & Management Review*, 2004, 28 (1):9 - 29.

② Kroll A. , “ The other type of performance information: Nonroutine feedback, its relevance and use ”, *Public Administration Review*, 2013, 73 (2):265 - 276.

③ Van der Meer F. - B. , Edelenbos J. , “ Evaluation in multi-actor policy processes: Accountability, learning and co-operation ”, *Evaluation*, 2006, 12 (2):201 - 218.

④ Schyve P. M. , “ The evolution of external quality evaluation: Observations from the joint commission on accreditation of healthcare organizations ”, *International Journal for Quality in Health Care*, 2000, 12 (3):255 - 258.

⑤ Shalley C. E. , Oldham G. R. , “ Effects of goal difficulty and expected external evaluation on intrinsic motivation: A laboratory study ”, *Academy of Management Journal*, 1985, 28 (3): 628 - 640.

⑥ Balthasar A. , Rieder S. , “ Learning from evaluations: Effects of the evaluation of the swiss energy 2000 programme ”, *Evaluation*, 2000, 6 (3):245 - 260.

第三方评估结果及其启示应用到其他领域,从而扩大第三方评估效应。^① 绩效评估只有将学习嵌入组织文化,才能实现政府绩效的持续提升。^② 因此,如果第三方评估可以促进参评部门的知识共享和集体学习,将有助于公共服务绩效提升。^③

最后,第三方评估可以促进组织创新,特别是为组织提供新的视角、观念、技术和实践,这都有助于评估能力建设和组织绩效改进。^④ 评估能力指“一种组织能力,即获取、匹配与维持组织的目标、结构、流程、文化、人力资本与技术,借此提供可以指导实践和支持决策的评估知识,进而提升组织绩效”^⑤。评估能力可以分为执行评估和使用评估这两大类能力,并具体分为六类:人力资源、组织资源、评估规划与活动、评估知识、组织决策和学习收益。^⑥ 评估能力受到组织环境和组织特征的影响,并会通过组织学习和组织创新影响组织绩效。^⑦ 评估能否真正奏效,取决于能否使组织发生改变,而理解其过程至关重要。^⑧

三、第三方评估影响组织绩效的依存条件

第三方评估要想发挥作用并改进政府绩效,取决于一定的环境和条

① Moynihan D. P. , “ Goal-based learning and the future of performance management”, *Public Administration Review*, 2005, 65 (2):203 - 216.

② Sanderson I. , “ Performance management, evaluation and learning in ‘ modern ’ local government”, *Public Administration*, 2001, 79 (2):297 - 313.

③ Fletcher G. , Dyson S. , “ Evaluation as work in progress: Stories of shared learning and development”, *Evaluation*, 2013, 19 (4):419 - 430.

④ Carman J. G. , Fredericks K. A. , “ Evaluation capacity and nonprofit organizations: Is the glass half-empty or half-full?”, *American Journal of Evaluation*, 2010, 31 (1):84 - 104.

⑤ Nielsen S. B. , Lemire S. , Skov M. , “ Measuring evaluation capacity-results and implications of a danish study”, *American Journal of Evaluation*, 2011, 32 (3):324 - 344.

⑥ Bourgeois I. , Cousins J. B. , “ Understanding dimensions of organizational evaluation capacity”, *American Journal of Evaluation*, 2013, 34 (3):299 - 319.

⑦ Cousins J. B. , Goh S. C. , Elliott C. J. , “ Framing the capacity to do and use evaluation”, *New Directions for Evaluation*, 2014, 2014 (141):7 - 23.

⑧ Henry G. T. , Mark M. M. , “ Beyond use: Understanding evaluation’s influence on attitudes and actions”, *American Journal of Evaluation*, 2003, 24 (3):293 - 314.

件,包括评估目的、评估参与、评估使用、评估支持等。

首先,评估目的不同,第三方评估的效果也会不同。有学者将评估分为两大类:致力于改进绩效的形成性评估(formative evaluation)和旨在得到评估结论的总结性评估(summative evaluation)。也有人将评估目的二分为评估和改进,还有学者指出评估的目的无法穷尽。^① 有学者指出,绩效评估的终极目标是改进,学习、激励和控制等其他目标都是从属于该目标的。^② 凡此种种,都表明评估目的对于引导评估过程至关重要,并会影响乃至左右评估的奏效。基于评估的目的是问责、改进还是获取基础知识,内部评估和外部评估的适宜性也不尽相同。对于问责和获取基础知识而言,外部评估具有相对优势;对于绩效改进来说,内部评估则表现相对较好。^③

其次,第三方评估能否对政府绩效产生影响,至关重要的是评估结果能否得到合理有效的使用。^④ 绩效信息使用是绩效管理的“大问题”,因为它关乎绩效评估向绩效管理跨越的“惊险一跳”。^⑤ 如果第三方评估产生的绩效信息得不到实质性使用,就很难对组织行为和绩效产生影响。^⑥ 如果参评部门只是表姿态象征性地使用评估结果,那么就可能使

① Chen H.-T., “A comprehensive typology for program evaluation”, *Evaluation Practice*, 1996, 17 (2):121 - 130.

② Behn R. D., “Why measure performance? Different purposes require different measures”, *Public Administration Review*, 2003, 63 (5):586 - 606.

③ Vedung E., “Six models of evaluation”, in Araral E., Fritzen S., Howlett M., Ramesh M., Wu X., eds. *Routledge Handbook of Public Policy*, London: Routledge, 2013: 387 - 400.

④ Contandriopoulos D., Brousselle A., “Evaluation models and evaluation use”, *Evaluation*, 2012, 18 (1):61 - 77.

⑤ Moynihan D. P., Pandey S. K., “The big question for performance management: Why do managers use performance information?” *Journal of Public Administration Research and Theory*, 2010, 20 (4):849 - 866.

⑥ Kroll A., “Drivers of performance information use: Systematic literature review and directions for future research”, *Public Performance & Management Review*, 2015, 38 (3): 459 - 486.

第三方评估大打折扣,甚至不会对政府绩效产生实质性影响。^① 评估使用包括可用性(可否使用)和能用性(能否使用)两个维度:可用性(usability)涉及评估方案设计和实施是否会影响评估结果、是否适合使用,能用性(use)指基于参评部门所处的组织情境和能力是否有条件使用评估结果。^② 评估使用还可以分为过程使用(process use)和结果使用(use of evaluation findings)。^③ 政府期望获得评估结果,并可以将其用于决策,这会表现为工具性、概念性和象征性使用。^④ 但是,更为重要的是评估过程使用,即利益相关者在参与评估的过程中获得的收益,如认识、学习和互动。^⑤

最后,第三方评估并非完全由第三方机构完成,需要参评部门的配合、参与和支持,如提供资料和数据、接受访谈和调查,乃至评估方案设计和调整。实际上,外部评估的引入对内部评估提出了许多新要求。^⑥ 因此,参评部门的参与方式和程度,会影响第三方评估对政府绩效的作用。特别是组织领导层的关注和支持,对于提升第三方评估的战略重要性具有很强的垂范作用。此外,社会利益相关者的参与和透明公开,有助于提升第三方评估质量,进而有利于评估得到关注和重视。

评估机构与被评部门之间的相互信任关系,是决定评估成败及其绩

① Weiss C. H. , “Have we learned anything new about the use of evaluation?” *American Journal of Evaluation* , 1998, 19 (1):21 - 33.

② Saunders M. , “The use and usability of evaluation outputs: A social practice approach”, *Evaluation* , 2012, 18 (4):421 - 436.

③ Amo C. , Cousins J. B. , “Going through the process: An examination of the operationalization of process use in empirical research on evaluation”, *New Directions for Evaluation* , 2007 (116):5 - 26.

④ Leviton L. C. , “Evaluation use: Advances, challenges and applications”, *The American Journal of Evaluation* , 2003, 24 (4):525 - 535.

⑤ Shulha L. M. , Cousins J. B. , “Evaluation use: Theory, research, and practice since 1986”, *American Journal of Evaluation* , 1997, 18 (3):195 - 208.

⑥ Volkov B. B. , “Beyond being an evaluator: The multiplicity of roles of the internal evaluator”, *New Directions for Evaluation* , 2011, 2011 (132):25 - 42.

效影响的关键因素之一。^①“如果问责信息不被参与的主体视为可信的——即其发现不被人相信——那么它将被不会使用,或者更糟的是它被误用,因为不准确或有偏的信息会被用于牟取私利。结果可能是这种信息的来源名声扫地。”^②即便是严谨的随机对照实验(RCT),也可能因为项目开发人员的参与而导致严重的利益冲突,并使评估结果出现偏差。特别是项目委托方与评估方之间的财政利益冲突就很难避免,并可以通过许多“蛛丝马迹”予以揭露。换句话说,如果循证决策的证据基础受到利益侵蚀,就会影响证据效力及据此做出的决策的质量。特别是在医药、犯罪等领域,这种有偏的评估结果会造成难以想象的负面影响。^③

第三方评估的可信度可以通过两种途径得到确保:提供高质量的不被视为有偏的信息;有声望的第三方确定信息是高质量和无偏的。得到第三方的质量认证,是提高评估可信度的重要措施。特别是在财务报表领域,外部独立审计师出具的审计报告,可以表明财务报表符合特定标准。在政府绩效评估领域,越来越多的国家和地区要求绩效信息审计的第三方认证。^④许多第三方评估都是以研究课题形式进行的,因此评估使用也意味着研究与实践的关联。一直以来,学术界与实务界、科学界与政策界之间存在很深的鸿沟,使研究转化和知识利用得不到充分发

① Burt R. S., Knez M., “Kinds of third-party effects on trust”, *Rationality and Society*, 1995, 7 (3): 255 - 292.

② Mayne J., “Evaluation for accountability: Myth or reality”, in Bemelmans-Videc M.-L., Lonsdale J., Perrin B., eds., *Making Accountability Work: Dilemmas for Evaluation and for Audit*, New Brunswick, NJ: Transaction Publishers, 2007: 63 - 84.

③ Gorman D. M., Conde E., “Conflict of interest in the evaluation and dissemination of ‘model’ school-based drug and violence prevention programs”, *Evaluation and Program Planning*, 2007, 30 (4): 422 - 429.

④ Mayne J., “Evaluation for accountability: Myth or reality”, in Bemelmans-Videc M.-L., Lonsdale J., Perrin B., eds., *Making Accountability Work: Dilemmas for Evaluation and for Audit*, New Brunswick, NJ: Transaction Publishers, 2007: 63 - 84.