

政府绩效第三方评估中主体意向的分化及其协调

——基于Q方法的分析

包国宪, 戴正

(兰州大学管理学院, 甘肃兰州, 730000)

摘要: 运用Q方法探究中国第三方政府绩效评估中主体意向的分化现象及其协调路径, 发现政府部门存在管理效能型、资源争取型和问责督导型三种意向类型, 第三方机构表现出专业价值型、市场生存型和公共责任型三种取向。这种意向分化根植于制度二元性、利益格局失衡和评估情境多元化等深层机制, 导致了“功能异化”现象, 表现为评估目标偏离、过程形式化和结果应用选择性等问题, 对评估的科学性和有效性构成了挑战。进而提出系统性的协调策略: 在宏观层面重构评估的制度逻辑框架, 将评估从“政府管理的外包服务”转向“公共治理的专业支撑”; 在中观层面构建多元制衡的治理架构, 破解委托代理关系中的权力失衡问题; 在微观层面完善专业技术标准体系, 压缩意向分化的操作空间。

关键词: 委托第三方评估; Q方法; 意向分化; 绩效评估

中图分类号: D630

文献标识码: A

文章编号: 1672-3104(2025)05-0111-15

20世纪80年代中期以来, 政府绩效评估作为重要的管理实践在世界各国被广泛应用^[1]。传统政府自评模式存在着评估主体单一、标准内化等问题, 容易带来结果片面性和过程封闭性问题^[2, 3]。为克服这些问题, 许多国家引入了第三方力量参与政府绩效评估, 以提高评估的客观性和公正性^[4]。第三方评估是指由政府委托但独立于政府的研究机构、专业公司等对政府绩效进行评价的外部评估方式。与内部评估相比^[5], 第三方评估具有评估主体独立、评估视角客观、评估方法专业等优势^[6]。中国自2007年以来日益重视对第三方绩效评估的运用^[7]。当前, 其已成为政府绩效管理的“标配”, 在提高政府绩效、回应社会关切中发挥着重要作用。

然而, 评估活动涉及多元利益相关者的复杂互动, 评估主体的价值取向、利益诉求和角色定位等因素都可能影响评估的过程与结果^[8]; 政策目标的不确定性和政府运作的复杂性, 使政府绩效测量面临着根本性的技术难题^[9]; 信息不对称易产生寻租空间, 评估结果更易受到操纵^[10]。因此, 引入第三方评估是把双刃剑: 对政府而言, 虽能提高评估的独立性和客观性, 但也带来了管控困难和标准不统一的问题; 对第三方而言, 虽能扩大业务范围、提升影响力, 但也面临着专业能力不足、角色定位不清等困境。随着实践的发展, 第三方评估的有效性逐渐凸显^[11, 12]。

现有研究主要从宏观制度环境和微观主体特征两个角度探讨影响第三方评估有效性的因素。在宏观制度环境方面, 主要关注: 评估体系规范化建设, 如评估指标体系科学性、程序标准化和结果应用估的制度协同机制, 涉及评估制度与其他相关制度的衔接程度^[16, 17], 以及不同层级政府间评估制度的

收稿日期: 2024-07-06; 修回日期: 2024-10-31

基金项目: 国家社会科学基金重点项目“国家治理效能目标导向下的政府职责体系优化研究”(20AZD032)

作者简介: 包国宪, 男, 甘肃庆阳人, 兰州大学管理学院教授、博士生导师, 主要研究方向: 政府绩效治理、战略管理; 戴正, 男, 湖北孝感人, 兰州大学管理学院博士研究生, 主要研究方向: 政府绩效管理、行政哲学, 联系邮箱: daizh21@lzu.edu.cn

协调程度^[18,19]。在微观主体特征方面,主要关注:评估主体的能力建设,包括第三方评估机构的自律机制、专业能力和道德操守,以及评估人员的专业背景、经验积累^[6,20],评估主体间的互动关系,如不同话语体系的影响,以及政府与第三方之间的权力不对称和信息不对称问题^[21-23];评估主体的动机和利益取向,包括第三方的策略性行为和政府部门的政绩考虑^[24-26]。

现有研究多将第三方评估面临的问题归结为“制度失效”或“第三方独立性缺失”^[27,28],这种研究取向隐含着一个重要的理论预设,即将某种理想的评估状态作为判断标准,偏离这一标准即被视为“失效”。然而,这种二元化的分析框架不仅带有明显的规范性色彩。更重要的是,它将复杂的评估实践简化为成功与失败的标签,忽视了评估活动本身的动态性和情境依赖性。本研究认为有必要超越“失效”与否的价值判断,采取更为中立和动态的视角来理解第三方评估的实践演变。这种视角转换的核心在于:承认意向分化作为一种客观存在,具有其内在的合理性和必然性,但也要认识到,当意向分化超出一定边界时,可能会对评估的整体功能产生负面影响,导致评估偏离其基本的公共价值追求。因此,关键问题不在于消除意向分化,而在于如何在承认其合理性的基础上,寻求适度的协调机制,使不同主体的意向能够在良性互动中实现评估的公共价值最大化。

为此,本研究引入“意向分化”概念开展分析。意向分化指评估参与各方在价值理念、角色定位和利益诉求等方面存在的显著差异。这一概念植根于建构主义评估理论,该理论强调评估是一个社会建构的过程,不同主体的主观认知和期待会深刻影响评估实践^[29]。同时,基于功能主义视角,任何社会行动或制度安排虽然源于行动者的主观意向,但其客观后果却不一定符合主观意向,甚至可能背道而驰^[30]。基于上述理论认识,本研究运用Q方法深入分析政府与第三方在绩效评估中的主观意向差异,旨在揭示不同利益主体对第三方绩效评估功能的预期差异及其成因,探讨这种差异对评估实践的影响机制,并在此基础上提出协调意向分化的优化策略,为完善第三方绩效评估机制、提升评估有效性提供理论参考和实践指导。

一、第三方评估主体意向的类型及分化

(一) 研究设计

Q方法是一种用于系统考察人们对特定议题主观态度和观点的研究方法^[31],要求被调查者对一组陈述进行主观排序,通过因子分析等统计方法提取反映人们观点分化的因子类型^[32],最后对各类型的特征及其异同进行定性解读^[33,34]。Q方法在公共管理领域已被广泛应用于主观性问题研究^[35,36]。本研究采用理论推演与实证归纳相结合的方式,形成研究所需的Q陈述集。陈述来源包括两个方面:一是理论演绎,从相关文献中提炼出反映第三方绩效评估中不同价值取向、角色定位的理论观点;二是深度访谈,从政府、第三方等利益相关方收集对第三方评估的主观认知与期待的表述。通过对理论观点与经验表述的系统梳理与整合,最终形成兼具理论相关性与情境贴切性的Q陈述集^[37,38]。

具体而言,在理论观点方面,研究借鉴Dryzek等^[39]的4×4话语分析矩阵,构建Q语句集。该矩阵包含两个维度:一是话语元素,包括本体论、能动性、动机和自然关系;二是主张类型,包括定义、制定、评判和倡导^[34]。本研究基于此矩阵对文献资料进行系统的分类与归并。在本体论维度,收集关于第三方评估性质定位的陈述;在能动性维度,整理关于评估主体权责界定的表述;在动机维度,归纳反映不同主体利益诉求的论述;在自然关系维度,梳理有关评估主体互动规范的观点。对于每一个矩阵单元,本研究均提取了多条具有代表性的陈述。在经验表述方面,本研究对15个政府部门、12个第三方的主要负责人或业务骨干进行了深度访谈。访谈聚焦受访者参与第三方评估的经历与体会,以及受访者对第三方评估的功能定位、理想图景的主观认知。

本研究最终形成了 32 条陈述, 并形成了 Q 陈述集。其中, 本体论包括: S1——第三方评估是政府管理体系的内部组成部分; S2——第三方评估是独立于政府的外部监督机制; S3——第三方评估是市场化的专业服务; S4——评估的依据主要来自政府提供的数据; S5——评估的依据主要为实地调研获取的第一手资料; S6——评估的依据主要是专家的经验判断; S7——评估质量取决于指标设计的科学性; S8——评估质量取决于问题发现的深度; S9——评估质量取决于建议的可操作性。能动性包括: S10——评估机构主要是政府的参谋助手; S11——评估机构是社会监督的代言人; S12——评估机构是独立的第三方裁判者; S13——评估机构应主动发现问题并提出建议; S14——评估机构应完全按照委托方的要求开展工作; S15——评估机构应根据自身专业判断确定重点; S16——好的评估机构应以满足委托方需求为重; S17——好的评估机构应坚持独立客观的立场; S18——好的评估机构要善于平衡各方利益。动机包括: S19——评估的根本目的是改进政府管理; S20——评估的根本目的在于强化社会监督; S21——评估的根本目的是满足政策或政府的合法性要求; S22——评估的根本目的是满足各方的利益(包括第三方自身)。自然关系包括: S23——委托方与评估方是领导与被领导的关系; S24——委托方与评估方是买卖双方的关系; S25——评估方与被评估方是监督与被监督的关系; S26——评估过程中政府部门应全程参与指导; S27——评估主体之间分工明确且相对独立; S28——评估机构独立开展工作并定期汇报; S29——规制评估的关键在于形成政府购买服务与社会捐赠的双向供给机制; S30——规制评估的关键在于构建和完善评估机构资质认证制度; S31——规制评估的关键在于构建和完善评估行业的自律机制; S32——规制评估的关键在于构建和完善评估的信息公开制度。

研究采用目的抽样方法, 选取 80 名受访者作为 P 样本, 其中政府部门和第三方评估机构各 40 人。选择样本时以代表性为主要考量, 受访者均为各单位分管绩效评估工作的负责人或业务骨干。具体操作时, 研究者首先向受访者说明本研究的目的和要求, 澄清 Q 分类并非对错之分, 而是受访者主观判断的体现。随后, 要求受访者通读所有陈述, 并根据个人认同度将其分为“非常同意”“同意”“中立”“不同意”和“非常不同意”五堆, 然后要求受访者在强制正态分布评分表中, 将各堆中的陈述按照个人偏好由外向内依次排序, 最后形成完整的 Q 分类。在整个 Q 分类过程中, 研究者适时询问受访者进行关键陈述的排序理由, 这为随后的结果分析提供了重要参照。

(二) 第三方评估主观意向的类型分析

1. 数据分析方法

运用 PQ Method 软件对 80 位受访者的 Q 排序结果进行分析。数据处理过程包括样本数据录入、因子分析、因子旋转与因子得分计算四个主要步骤。在因子分析中, 采用主成分分析法计算特征值。两个主体均提取出 3 个因子(特征值大于 1, 累计解释量分别为 62.5%和 67.5%), 这些因子代表受访者对第三方绩效评估主观意向的主要类型。随后, 根据每条 Q 样本的 Z 值计算其在各因子中的得分值, 并通过比较 Q 样本在不同因子中的得分值, 探究不同群体对各陈述的认知差异。政府部门和第三方各自析出 3 个因子: F1~F6。本研究通过分析因子所包含的典型变量的 Q 陈述排序情况, 并结合受访者的选择依据, 对各主体的因子特征进行了描述与命名。

2. 政府部门的意向类型分析

F1 是管理效能型(与 10 位政府人员显著相关)。这类主体对评估的认知主要聚焦于其管理工具属性, 突出表现为对科学性指标和内部组成等陈述的高度认同(S7、S1)。在本体论维度, 他们将评估视为提升政府管理效能的内生性工具; 在能动性维度, 强调政府主导作用; 在动机维度, 突出评估对改进政府管理的促进作用; 在自然关系维度, 倾向于建立规范化的评估互动机制。此种取向表现为对全程参与和按需开展的支持(S26、S14)。政府受访者表示: “我们需要通过科学合理的评估体系来发现

问题、提升效率。这不仅是对第三方的要求，也是对我们自身管理水平的检验。”(访谈资料：ZF20241020-01)^①然而，这类政府人员对评估的理解并非简单的行政控制。他们对领导关系持否定态度，反映了对第三方专业性的尊重，但也不完全认同独立裁判等过分强调独立性的表述(S23、S12)。这种平衡反映了在追求管理效能与保持专业自主之间的辩证思维。

与其他类型相比，管理效能型政府人员对评估的社会监督功能关注较少，相对不认同社会监督和监督代言等陈述(S20、S11)。这表明他们更倾向于从内部管理而非外部监督的角度理解评估的功能定位。这种意向与部门背景、制度环境密切相关。持管理效能型意向的政府人员多分布在承担绩效管理和制度建设职责的综合部门，将第三方评估视为提升政府管理现代化水平的重要路径。分管绩效评估的主任科员表示：“我们要用好第三方评估这个工具，既要发挥专业机构的技术优势，也要确保评估始终服务于提升政府管理效能这个根本目标。”(访谈资料：ZF20241020-02)总之，管理效能型体现了一种将评估嵌入政府管理体系的制度化思维。这种意向类型虽然强调评估的工具理性，但并非简单的行政主导，而是试图在管理效能与专业独立之间寻求平衡。

F2是资源争取型(与8位政府人员显著相关)。这类主体对评估的认知聚焦于资源获取功能，突出表现为对合法性要求和参谋助手等陈述的高度认同(S21、S10)。在本体论维度，他们将评估视为一种政策工具；在能动性维度，强调评估的支持功能；在动机维度，突出评估对获取资源的促进作用；在自然关系维度，倾向于建立服务性的合作关系。此类政府人员特别关注评估的实用价值，高度认同可操作性建议和专业判断重点等陈述(S9、S15)。基层受访者表示：“好的评估要能找准工作短板，为我们争取项目资金和政策支持提供依据。”(访谈资料：ZF20241024-03)在评估过程中，这类官员倾向于强调评估的建设性作用，对平衡各方利益和实地调研等表示认同(S18、S5)。政府基层管理者表示：“评估要立足基层实际，提出有针对性的政策建议，帮助我们更好地向上级反映工作需求。”(访谈资料：ZF20241024-04)

资源争取型政府人员虽然看重评估的政策支持功能，但并不完全认同简单的配合性评估。他们对政府数据依据和完全按需开展评估持保留态度(S4、S14)，期待通过客观专业的评估获得更有说服力的支持。政府人员表示：“评估要真正深入基层，通过扎实的调研发现工作亮点和难点，这样形成的政策建议才更有分量。”(访谈资料：ZF20241025-05)这种意向与基层部门面临的资源约束和发展压力密切相关。这类政府人员通常来自基层执行部门或新兴领域(部门)，他们需要通过评估来展示工作成效、争取更多支持。科室负责人表示：“基层工作创新往往需要政策和资金支持，专业化的第三方评估能够帮助我们更好地总结经验。”(访谈资料：ZF20241024-06)

F3是问责督导型(与7位政府人员显著相关)。这类主体对评估的认知主要聚焦于监督和约束功能，突出表现为对监督关系和社会监督等陈述的高度认同(S25、S20)。在本体论维度，他们将评估定位为独立于政府的外部监督机制；在能动性维度，强调评估机构的主动监督作用；在动机维度，突出评估对社会监督的促进作用；在自然关系维度，倾向于构建规范的评估问责体系。此类政府人员特别强调评估的独立性和问题导向，高度认同主动发现问题和问题发现的深度等陈述(S13、S8)。财政部门的管理者表示：“第三方评估要严格保持独立立场，深入发现问题，提出切实可行的整改建议，这样才能真正发挥监督作用。”(访谈资料：ZF20241025-07)在评估结果运用方面，问责督导型政府人员格外重视信息公开和结果追踪，强烈支持信息公开制度和行业自律机制(S32、S31)。分管绩效评估的主任科员表示：“评估结果必须向社会公开，同时要建立问题整改台账，定期跟踪整改情况，确保评估发现的问题得到有效解决。”(访谈资料：ZF20241020-02)

这类政府人员虽然强调监督约束，但更注重监督的建设性作用，对具有可操作性的建议和专业判

断重点表示认同(S9、S15)。这反映出他们期待评估机构发挥积极的咨询作用,通过专业化的第三方评估推动问题解决。审计部门的管理者表示:“评估不仅要指出问题,更要提出有针对性的改进建议。只有这样,才能真正发挥监督的效果。”(访谈资料:ZF202410025-08)这种意向与问责督导型政府人员的工作重点有关。这类政府人员在日常工作中往往直接面对具体的管理问题和监督任务,他们深知单靠内部监督具有局限性,因此更加重视发挥第三方评估的外部监督作用。部门负责人表示:“内部评估存在固有局限,引入第三方评估能够带来新的视角,发现我们自己看不到的问题。”(访谈资料:ZF-20241025-09)

3. 第三方的意向类型分析

F4 是专业价值型(与 11 位第三方人员显著相关)。这类主体对评估的认知主要聚焦于专业价值属性,表现为对独立客观立场和市场化专业服务陈述的高度认同(S17、S3)。在本体论维度,他们强调评估的专业独立性;在能动性维度,注重专业判断的自主权;在动机维度,突出评估的专业价值;在自然关系维度,倾向于构建专业化的评估规范。此类评估人员特别重视评估的规范性和独立性,高度认同资质认证制度和行业自律机制等陈述(S30、S31)。评估机构片区负责人表示:“评估工作必须坚持专业标准,建立起完善的质量控制体系,这是评估机构的立身之本。”(访谈资料:QY20241015-01)在评估方法上,这类评估人员强调专业技术规范,支持科学性指标和专业判断重点等陈述(S7、S15)。企业评估工作人员:“专业的评估需要严谨的方法和规范的程序,这样才能确保评估结论的科学性和可信度。”(访谈资料:QY20241015-02)

这类评估人员虽然强调专业独立,但并不排斥与政府的建设性合作。他们对独立裁判者的表述并不认同,而是更支持主动发掘问题的积极姿态(S12、S13)。资深评估工作者表示:“我们要在保持专业独立的同时,为政府提供有价值的决策参考,这就要求我们既要有独立性,也要有建设性。”(访谈资料:GX20241022-01)这种意向取向主要与专业人员的学术背景和职业操守有关。这类评估人员往往具有深厚的研究功底和丰富的实践经验,他们将评估视为一项需要专业积累和持续创新的事业。评估机构主任表示:“评估不是简单的数据统计或问题诊断,而是需要扎实的专业功底和系统的分析框架,这样才能真正发挥评估的价值。”(访谈资料:GX20241022-02)

F5 是市场生存型(与 8 位第三方人员显著相关)。这类主体对评估的认知主要聚焦于服务属性,突出表现为对买卖双方关系和满足各方利益等陈述的高度认同(S24、S22)。在本体论维度,他们更倾向于将评估定位为市场服务;在能动性维度,重视回应委托方需求;在动机维度,注重多方利益平衡;在自然关系维度,更看重契约性合作关系。此类评估人员格外关注评估服务的实用性和接受度,强调满足委托方的需求,对政府数据依据表示认可(S16、S4)。项目负责人表示:“作为市场主体,我们必须考虑评估服务的实际效果和客户满意度。”(访谈资料:QY20241022-03)对于评估方式,这类评估人员采取务实立场,赞同完全按需开展和全程参与指导等做法(S14、S26)。机构负责人表示:“评估工作要立足实际需求,在保证基本专业性的同时,也要考虑委托方的具体期待。”(访谈资料:QY20241022-04)在实际运作中,市场生存型评估人员展现出务实态度,更重视在具体项目中寻找平衡点,对平衡各方利益普遍认可(S18)。评估机构项目主管:“在保证评估基本质量的前提下,我们也要考虑项目的可持续性和机构的长远发展。”(访谈资料:QY20241022-05)

当然,市场导向并不意味着完全放弃专业原则,他们试图在市场需求和专业要求之间找到平衡点,体现为对资质认证制度的支持(S30)。资深评估师表示:“在激烈的市场竞争中,专业资质和服务能力是我们的核心竞争力。”(访谈资料:QY20241019-06)这种意向主要源于评估机构所面临的市场压力和生存挑战。作为市场主体,这类机构需要在竞争中谋求发展,这促使他们更多地从服务供给和市场

需求的角度思考评估工作。

F6是公共责任型(与8位第三方人员显著相关)。这类主体对评估的认知主要聚焦于其公共价值实现功能,表现为对社会监督代言人和双向供给机制等陈述的高度认同(S11、S29)。在本体论维度,他们强调评估的公共属性;在能动性维度,注重评估的社会效应;在动机维度,突出评估的公共责任;在自然关系维度,倾向构建多元协同的评估机制。此类评估人员特别关注评估的公共价值导向,强调实地调研资料和独立工作汇报(S5、S28)。高校评估团队负责人表示:“作为学术机构,我们开展评估工作不仅要服务政府决策,更要承担起推动公共治理进步的责任。”(访谈资料:GX20241022-03)在评估方法上,这类评估人员格外注重科学性和系统性,认同明确分工和专家经验判断等做法(S27、S6)。这种倾向与高校评估机构的学术追求密切相关。资深学者表示:“评估不只是一项服务,也是一个重要的研究领域。我们希望通过严谨的评估研究,为改进绩效提供理论支撑。”(访谈资料:GX20241022-04)其中,公共责任型评估人员与政府部门的问责督导型评估人员在某些方面表现出趋同性,都强调评估的独立性和社会监督功能。但两者之间存在区别:问责督导型更侧重于通过评估强化行政问责,而公共责任型则更关注评估对公共治理的促进作用。这种差异源于双方在制度身份和价值取向上的不同:政府部门基于行政监督的需要推动评估,而学术机构则从学术研究和社会责任的角度开展评估。评估专家表示:“我们既要保持学术独立性,又要考虑研究的社会价值。通过评估研究推动公共治理创新,这是高校智库的责任。”(访谈资料:GX20241022-05)这种意向来自高校评估机构的学术属性和公共使命。作为兼具学术性和公共性的机构,它们在评估中既要维护学术标准,又要推动公共价值的实现。评估中心主任表示:“高校开展第三方评估,既是一项学术工作,也是一种社会服务。我们要把评估做成一个推动公共治理进步的平台。”(访谈资料:GX20241022-02)

二、第三方评估主体意向分化的成因及影响

第三方评估主体意向的分化是一个复杂的动态过程,受多种因素影响。为深入理解这一现象,本节将从宏观和微观两个层面剖析评估主体意向分化的生成机制及其影响。

(一) 宏观制度环境对评估主体意向的塑造

第三方评估主体意向的形成深深嵌入特定的制度语境和结构约束之中。在第三方评估语境下,这一理论视角揭示了评估主体意向分化的宏观根源。本研究从制度二元性、利益格局和评估情境类型三个维度剖析第三方评估主体意向分化的宏观生成机制。制度二元性反映了正式规制与非正式规范的交互作用;利益格局揭示了评估主体间的权力分配和资源控制状况;评估情境类型体现了不同评估形式下的权力结构和角色期待差异。

1. 制度二元性的影响:正式规制与非正式规范的张力

第三方评估活动嵌入了正式规制与非正式规范交织的制度环境中。这种制度的二元性为评估主体的行动提供了基本规则和边界条件,也构成了意向分化的重要根源。正式规制主要包括相关法律法规、政策文件等明文规定,为评估活动提供了基本框架和行为准则。例如,《中华人民共和国政府采购法》对第三方评估机构的遴选程序做出了明确规定^[40],财政部门出台的绩效评价指引文件为评估实践提供了操作指南^[41]。与之并行的是一系列根深蒂固的非正式规范,包括行政文化传统、组织惯例、隐性行为准则等^[42, 43]。在我国的政治行政环境里,科层制文化、官本位思想等对评估实践产生了深远影响。

制度二元性通过塑造评估主体的认知框架和行为动机来影响意向的形成,主要体现在评估目标设定、过程管理和结果运用三个关键环节上。

第一, 在评估目标设定上, 正式规制强调客观性和独立性, 而非正式规范则倾向于满足行政需求。这种张力直接影响了评估主体对评估目的的理解和定位。例如, G 省发改委的产业绩效第三方评价招标文件明确要求第三方评估机构独立、客观、公正地开展评估工作, 但在实际操作中, 评估项目的立项却主要是为了配合特定的政策推进或行政决策。高校专家表示: “有时候我们接到的评估任务, 实际上是要给已经确定的政策方向‘背书’。在这种情况下, 如何平衡专业操守和委托方需求就成了一个很大的挑战。”(访谈资料: GX20231105-06)

第二, 在评估过程管理上, 正式规制要求建立规范化的评估流程, 而非正式规范则强调灵活应对领导意图。这种二元性导致评估主体在具体实践中面临着操作规范与实际需求的矛盾。这种情况在 L 市“放管服”绩效评估项目中得到了充分体现。该项目的正式评估方案强调要全面评估“放管服”改革在各个部门的落实情况和效果, 但在实际执行过程中, 第三方常常需要根据市领导的关注点临时调整评估重点。评估专家表示: “我们制订的评估方案原本计划对市级各部门的‘放管服’改革进展进行全面评估。但在实施过程中, 我们多次接到要求重点关注某些特定部门或某些具体政策效果的指示。这导致我们不得不在保持评估全面性和回应领导关切点之间不断权衡, 有时甚至需要临时调整评估指标和权重。”(访谈资料: GX20231105-07)这种现象反映了正式规制与非正式需求在评估过程管理中的张力。政府部门的资源争取型意向倾向于将评估视为绩效展示和获取支持的工具, 而第三方的专业价值型意向则更强调评估的客观性和全面性。这种意向差异进一步加剧了评估过程中的制度二元性, 导致评估实践中常常出现方法严谨性和操作灵活性并存的现象。评估主体往往需要在遵循正式评估程序和响应非正式行政需求之间寻找平衡点, 这不可避免地影响了评估的独立性和客观性。

第三, 在评估结果运用上, 正式规制强调结果的客观性和公信力, 而非正式规范则更注重结果的实用性和政治正确性。例如, Q 市专项资金的绩效评估结果公开制度要求客观全面地公布评估结果, 但在实践中, 负面评价往往会被淡化或以内部参考的形式呈现。资深评估专家表示: “评估报告的措辞常常需要反复斟酌, 既要保证专业性, 又要考虑各方的接受度。”(访谈资料: GX20231105-06)

制度二元性不仅导致了评估主体间的意向分化, 还引发了同一主体内部的认知冲突。Q 方法分析显示, 部分政府人员在排序过程中对独立客观立场和满足委托方需求这两个陈述给予了相近的高分(S17、S16), 这反映出他们在遵守正式评估规则和服从行政惯例之间的矛盾心理。同样, 一些第三方人员在排序时给予了“专业判断重点”和完全“按需开展”这两个陈述较高的评分(S15、S14), 这体现出他们在坚持专业标准和适应行政文化之间的权衡。

2. 利益格局的影响: 委托代理关系中的权力失衡

利益格局是指评估主体间的权力分配和资源控制状况, 构成了行动者策略选择的基础。第三方评估本质上是一种委托代理服务, 其中政府部门是委托方, 第三方是代理方。这种关系内嵌于特定的利益格局之中, 反映了评估主体间的权力分配和资源控制状况。利益格局的复杂性是导致评估主体意向分化的又一重要因素。在理想状态下, 委托方和代理方应该基于契约精神, 共同追求评估的客观性和有效性。然而, 现实中的利益格局往往呈现出明显的权力不对称, 主要体现在以下三个方面。

第一, 在评估资源控制上, 政府部门占据主导地位, 掌控着评估项目的立项、经费分配和结果应用等关键资源。这种资源控制权使政府部门在评估过程中处于相对优势地位。在大部分评估中, 第三方评估机构的遴选和续聘往往取决于政府部门的决策, 这直接影响了评估机构的生存和发展。这种资源依赖导致部分评估机构在实践中更倾向于迎合委托方需求, 而非坚持独立立场。第三方人员表示: “我们常常需要考虑如何维持与政府部门的良好关系, 这关系到我们能否获得人家后续的评估项目。”(访谈资料: QY20231023-07)

第二,在评估标准制订方面,政府部门拥有更大的决定权,主要体现在关键指标的設置和权重分配上。政府部门往往基于既有的正式规制和管理需求来确定评估标准,而第三方在这一过程中的影响力有限。评估专家表示:“很多关键指标实际上是由政府部门预先设定的,我们作为第三方评估机构,能够参与修改的空间非常有限。这些指标可能更多地反映了行政管理的需求,而不完全符合专业评估的标准。”(访谈资料:GX20240105-08)这种标准制订权的不平衡可能导致评估指标过于行政化,忽视了专业评估的特殊需求。

第三,在评估信息获取方面存在明显的不对称性。政府部门作为委托方,掌握着大量的内部数据和背景信息,而第三方往往处于信息劣势。同时,被评估对象对信息的选择性提供也加剧了评估的复杂性。第三方项目经理表示:“我们在评估过程中经常遇到信息获取的困难。一方面,一些重要的背景资料需要多次向政府申请;另一方面,被评估单位有时会选择性地提供有利于自己的数据,这给客观评估带来了很大挑战。”(访谈资料:QY20240420-08)评估专家表示:“我们有时不得不花费大量时间来核实和补充信息,这无疑提高了评估的难度和成本。”(访谈资料:GX20240420-09)这种信息获取上的权力失衡可能导致评估结果的片面性,影响评估的客观性和全面性。

权力失衡通过多重机制影响评估主体的意向形成,导致政府部门和第三方产生不同的策略性考量。对于政府部门而言,权力优势强化了其行政主导意向,使其倾向于将评估视为行政管理的延伸,形成了对领导关系和完全按需开展的高度认同(S23、S14)。同时,作为最终决策者,政府部门更注重结果导向,倾向于将评估结果用于验证政策有效性和展示工作成效。这种权力优势还体现在信息掌控上,政府部门可以凭借信息优势主导评估方向,进一步强化其在评估过程中的主导地位。相对而言,第三方面临着更为复杂的处境。权力劣势迫使其在专业操守和市场生存之间寻求平衡,这种资源依赖转化为策略性妥协,导致专业价值型和市场生存型意向并存。为消除信息上的劣势,第三方不得不采取更加灵活和更具适应性的策略,这种信息不对称转化为认知差异,进一步加深了其与政府部门的意向分化。

3. 评估情境类型的影响:互动模式与角色期待的差异

第三方评估呈现出多元化的情境特征,这种多元性源于评估目的、对象和方法上的差异。不同的评估情境塑造了特定的互动模式和角色期待,进而影响评估主体意向的形成。根据评估的组织形式和互动模式,可以将第三方评估分为三种典型情境:上级评下级、同级部门评估和自评。

上级评下级通常发生在上级政府部门委托第三方对下级部门或地方政府进行评估的场景中。在这种情境下,评估往往带有较强的督察和问责色彩,其主要目的是加强监管、推动改进和促进政策落实。其互动模式呈现出明显的等级特征。委托方(上级部门)处于主导地位,能够决定评估的框架、标准和结果的应用。第三方则需要在委托方的要求和专业标准之间寻求平衡,既要满足上级部门对评估严格性的期望,又要考虑到下级的实际情况。被评估对象(下级部门)往往处于相对被动的位罝上,可能会采取各种策略来应对评估,如选择性地提供信息或强调特殊情况^[18,44]。这种情境下的评估往往更加正式和结构化,评估结果的影响力也较大。

同级部门评估则呈现出不同的特征。当政府部门委托第三方评估同级的另一个部门时,评估更多地体现为一种横向比较和绩效监督。评估的氛围可能更加敏感和复杂,因为涉及部门间的利益博弈和权力平衡。互动模式呈现出一种相对平等但潜在竞争的特征,第三方评估机构需要在保证评估客观性的同时,考虑部门间的关系和潜在的影响。评估的重点可能更多地放在比较分析和经验共享上,而非简单的优劣判断上。这种情境下的评估往往需要更高的政治敏锐性、沟通和协调能力。

自评情境,即当政府委托第三方对自身工作进行评估时,评估的独立性和客观性面临着最大挑战。

委托方既是评估主体, 又是评估对象, 这种双重身份导致了评估目的模糊化和操作复杂化。其互动模式呈现出内部化特征, 委托方对评估拥有更大的控制力, 可能影响评估范围、方法和结果的呈现。第三方评估机构需要在保持专业独立性和满足委托方需求之间寻找平衡点^[19]。评估焦点更多地放在内部改进和问题诊断上, 而非外部问责上。这种情境下的评估更加灵活和具有探索性, 但面临着被异化为自我辩护工具的风险。

在这三种情境下, 评估主体的意向分化表现出不同的特征。在上级评下级的情境中, 政府部门倾向于强调评估的监督和问责功能。Q 方法分析显示, 政府部门更认同满足委托需求导向(S16), 反映出其将评估视为管理工具的倾向。相比之下, 第三方更强调利益平衡取向(S18), 体现出其在满足上级要求和考虑下级实际之间寻求平衡的努力。评估专家表示: “我们需要在评估中既反映上级部门的要求, 又考虑到基层的实际情况, 这种平衡很关键。”(访谈资料: GX20231105-10)在同级部门评估情境中, 评估主体的意向更趋向于谨慎和策略性。政府部门更关注建议的可操作性(S9), 以提高评估的实用性和影响力。第三方则更倾向于利益平衡(S18), 反映出其在复杂部门关系中寻求中立立场的策略。评估机构负责人表示: “在这种情境(同级部门评估)下, 我们需要特别注意平衡各方利益关系, 避免成为部门博弈的工具, 关键是不要得罪人。”(访谈资料: QY20240420-08)在自评情境中, 评估主体的意向分化最为明显。政府部门更关注全程参与指导(S26), 反映出其希望维护自身利益的倾向。而第三方则相对强调分工独立原则(S27), 表明其试图在复杂情境中坚守专业立场。评估专家表示: “在自评情境下, 保持独立性是最大的挑战, 我们需要通过专业方法和多源数据来确保评估的客观性。”(访谈资料: GX20240511-11)这种意向分化反映出了不同评估情境下主体间不同的互动模式和利益考量。

在上级评下级情境中, 权力关系明确但存在上下级张力; 同级部门评估涉及更复杂的利益平衡问题, 需要更高的政治敏锐性; 在自我评估情境下, 独立性与回应性的矛盾最突出, 考验评估主体的专业操守和策略平衡能力。这种情境依赖的意向分化, 不仅影响了评估的具体实施, 也深刻塑造了评估实践的整体生态。情境多元性对评估主体意向分化的影响主要体现在两个方面。一是不同评估情境塑造了独特的权力结构, 直接影响着评估主体的行为策略和价值取向。在上级评下级情境中, 明确的等级关系使评估过程更加正式化和标准化, 评估主体意向聚焦于如何在严格执行和灵活应对之间寻求平衡。同级部门评估和自评情境则是更复杂的互动模式, 评估主体的意向形成过程更加动态和多元。例如, 在同级部门评估中, 由于缺乏明确的权力等级, 评估主体更倾向于将评估视为一种协商和学习的过程, 而非简单的考核。二是不同评估情境下的角色期待和责任分配机制显著影响着评估主体意向的形成。在上级评下级情境中, 第三方被赋予了更多的问责和监督职能, 导致其更注重评估的严谨性和结果的可量化性。相反, 在自评情境中, 第三方更多地被视为咨询顾问, 期望其提供具有建设性的改进建议, 从而促使其更关注评估发展性和参与性。这种角色期待的差异不仅影响了评估主体的工作方式, 也深刻塑造了他们对评估本质和价值的理解。例如, 在同级部门评估中, 第三方可能更倾向于扮演中立调解者的角色, 这种定位使其更加注重平衡各方利益, 而非追求单一的评判标准。

这种情境依赖的意向分化机制揭示了评估实践中的复杂性和动态性。评估是技术过程, 更是深嵌于特定社会政治环境中的实践活动。评估主体的意向并不是固定的, 而是在不同情境下不断调整和重构的。这要求在设计 and 实施评估时, 要充分考虑具体情境特征, 采取更加灵活和具有适应性的策略。同时, 这种视角促使我们超越传统技术理性范式, 更多地关注评估的社会建构过程和权力动态。

(二) 微观意向对宏观结构的反作用与重构

评估主体的微观意向并非单向受制于宏观结构, 而是通过日常实践和策略性行动不断影响和改变着既有的制度环境、权力结构、实施情境和社会功能。这种反作用与重构过程体现了评估主体的能动

性和创造性,构成了第三方评估实践演进的内在动力。这一观点与吉登斯(Giddens)的结构化理论相呼应,强调社会结构与个体行动的辩证关系^[45]。在第三方评估语境中,评估主体既受制于既有的制度和权力结构,又通过日常实践重塑着这些结构。评估主体对宏观结构的反作用与重构主要体现在三个相互关联的方面上:对制度环境的能动性回应、对互动模式的动态重塑、对评估情境的拓展与重新定义。

首先,在制度环境方面,评估主体通过灵活解释和创造性执行正式规制来应对制度压力。数据显示,满足委托方需求导向(S16)得到了较多认同,反映出评估主体在制度执行中的灵活性。评估从业者表示:“我们会在报告的框架和形式上严格遵守官方要求,但在内容的组织和呈现上,我们会采用一些创新的方法。比如,我们会在正面评价中巧妙地融入一些建设性意见,或者通过案例分析的形式提出一些敏感问题。”(访谈资料:QY20240420-09)这种“创造性遵从”的做法,实际上是评估主体在制度约束下的一种能动性策略。同时,评估主体还通过参与制度设计来影响制度环境。高校评估专家表示:“最近几年,我们越来越多地被邀请参与政府部门组织的评估标准制订讨论会。我们利用这些机会提出了很多来自实践的建议,比如增加对评估过程的质量控制要求,强化对评估结果应用的跟踪等。有不少建议被(相关部门)采纳,成为新的评估规范。”(访谈资料:GX20240511-12)这种参与制度设计的做法,使得评估主体从单纯的制度执行者转变为制度的共同缔造者。其次,在互动模式方面,评估主体通过专业化提升、服务范围拓展和利益相关方合作重塑着委托代理关系。数据显示,问题发现深度(S8)和行业自律机制(S31)得到较多支持。高校评估专家表示:“我们不再局限于单一类型的评估服务,而是开始提供从前期调研、中期评估到后期跟踪的全流程服务。这种全面的服务能力让我们在与政府部门谈判时有了更多筹码。”(访谈资料:GX20240105-08)通过这些策略,评估主体不仅增强了市场竞争力,也在一定程度上改变了在委托代理关系中的弱势地位。最后,在评估情境方面,评估主体通过实践创新拓展了评估边界并重新定义了评估关系。数据显示,满足各方利益(S22)得到第三方较多的认同。评估机构经理表示:“我们现在不仅与政府部门保持沟通,还积极与媒体、学术机构和社会组织建立联系。这种多方合作丰富了我们的信息来源,也增强了评估结果的社会影响力。”(访谈资料:QY20231023-10)这种做法不仅拓展了评估边界,也重新定义了评估主体与各利益相关方的关系。

这些微观层面的创新实践和策略性行为通过累积效应和扩散机制,逐渐影响和改变了宏观层面的制度环境、权力结构和评估情境。例如,创新性评估方法和工具被纳入官方评估指南,改变了正式规制的内容;评估机构通过专业化和网络化行为获得了更大话语权,推动了评估政策调整,重塑了互动模式;新的评估理念和实践拓展了评估的边界和功能,重新定义了评估在公共管理中的角色和价值。然而,这种从微观到宏观的影响过程并不是单向或平稳的,往往伴随着冲突、协商和妥协。政府评估管理人员表示:“我们欢迎第三方提升专业能力,拓展服务范围,但也希望他们能够理解并尊重政府的管理需求。”(访谈资料:ZF20240510-10)这种平衡需求反映了宏观结构重构过程中的复杂性和挑战性。

(三) 意向分化的系统影响与长期效应: 功能异化的视角

研究发现,意向分化不仅影响着单次评估的效果,还对整个评估体系的功能发挥产生了深远影响。为了更好地理解和描述这种系统性影响,本研究提出“功能异化”这一概念。功能异化是指评估活动在实际运作过程中偏离其原有目标和功能,转而服务于其他可能与初衷相悖的目的的现象。第三方政府绩效评估的初衷是提升政府治理能力、改进公共管理方式,但在实际运作中,由于各方主体意向的差异和相互作用,评估活动常常呈现出与原初设想不同的功能和效果。

1. 目标异化: 评估宗旨的偏离与拓展

目标异化主要体现为评估的根本目的被扭曲或被重新定义。这种异化源于不同评估主体对评估目标的差异化理解。正如前文的分析,政府部门的管理效能型和资源争取型意向倾向于将评估视为管理

工具和政绩展示的手段。这种倾向导致评估目标从客观评价和改进管理转变为维护部门利益或展示政府成就。例如, 在资源争取型政府部门的实践中, 评估往往被设计为验证既定政策有效性的工具, 而非发现问题、促进改进的机制。这种目标异化不仅限制了评估批判性功能的发挥, 还可能导致系统性的管理盲点。从长期来看, 这种异化可能削弱评估制度的公信力, 影响其在公共管理中的地位。然而, 目标异化并非全然消极。第三方评估机构, 尤其是专业价值型和公共责任型评估机构, 通过坚持专业标准和回应民众需求, 实际上拓展了评估的目标范畴。这种拓展使得评估从单一的行政考核工具逐渐发展为集问责、学习、改进和民众参与等多重功能于一体的综合性机制。这种目标的多元化虽然提高了评估的复杂性, 但也提高了其在复杂公共管理环境中的适应性和价值。

2. 过程异化: 评估实施的形式化与创新

过程异化主要表现为评估实施过程中的形式主义倾向和实质性偏离。这种异化源于政府部门的行政控制意向与第三方的专业自主性追求之间的矛盾。这种矛盾导致评估过程中出现了一些形式主义做法, 如过分强调程序完整性而忽视实质调查, 或者为迎合特定预期而调整评估方法。这种过程异化不仅降低了评估的实际效用, 还可能掩盖真实问题, 阻碍管理的实质性改进。从长期来看, 它可能导致评估成为一种例行公事, 失去发现问题、促进创新的功能。但是, 意向的分化也推动了评估过程的创新。不同主体的多元需求激发了评估方法和工具的不断更新。特别是专业价值型评估机构通过引入新的技术和方法, 如大数据分析、参与式评估等, 提高了评估的科学性和适用性。这种创新不仅提升了评估质量, 还拓展了评估的应用范围, 为解决复杂的公共问题提供了新的思路和工具。

3. 结果异化: 评估应用的选择性与综合化

结果异化主要体现在评估结果的选择性使用和解释偏差上。这种异化源于不同主体对评估结果的差异化诉求。如前文所示, 资源争取型政府部门倾向于选择性地使用有利于自身的评估结果, 而忽视不利的发现。这种做法不仅偏离了评估的客观性原则, 还可能导致决策失误和资源错配。从长期来看, 结果异化可能导致评估失去其作为决策支持和管理改进工具的核心价值。它可能强化某些既有的管理偏见, 阻碍真正的组织学习和创新。更严重的是, 持续的结果异化可能导致评估制度的信誉危机, 削弱其在公共管理中的地位和影响力。然而, 意向分化也促进了评估结果应用的综合化。不同主体的多元诉求推动着评估结果在更广泛的领域内得到应用。例如, 一些地方政府不仅将评估结果用于内部管理改进, 还用于政策宣传、民众沟通和跨部门协调。这种综合应用提高了评估的影响力和价值, 使其成为连接政府、市场和社会的重要桥梁。

总之, 功能异化反映了第三方政府绩效评估在实践中面临的深层次挑战。它既是不同主体意向分化的结果, 也是评估制度适应复杂公共管理环境的一种表现。虽然功能异化在某些方面偏离了评估的初衷, 但它也推动了评估功能的拓展和创新。因此, 应对功能异化的挑战, 需要在认识其复杂性和两面性的基础上, 采取系统性的应对措施。这些措施包括但不限于: 重新定义和澄清评估的核心目标, 平衡各方诉求; 优化评估过程管理, 既要防范形式主义, 又要鼓励方法创新; 建立评估结果的多元化应用机制, 既要防止选择性偏差, 又要拓展应用范围。

三、第三方评估主体意向分化的协调

面对第三方评估主体意向分化导致的功能异化问题, 协调的关键不在于消解差异或寻求包容, 而在于重构制度逻辑, 通过系统性的制度变革消除意向分化的制度根源。意向分化本质上源于制度供给缺陷、治理结构失衡和技术标准模糊, 需要在宏观制度逻辑、中观治理结构、微观技术标准三个层面

进行递进式重构。

在宏观制度逻辑层面,当前第三方评估面临的核心问题是制度二元性导致的行为预期不一致。正式制度强调独立客观,非正式规范却要求服务行政,这种制度逻辑的内在矛盾为意向分化提供了制度空间。解决这一问题需要构建统一的制度逻辑框架,将评估定位从“政府管理的外包服务”转向“公共治理的专业支撑”。这种制度逻辑的转换意味着评估不再是政府部门的私有物品,而是具有公共产品属性的治理工具。在这一制度逻辑下,政府部门作为委托方的核心职责是确保实现评估的公共价值,而非满足部门的特定需求;第三方作为专业服务提供者的核心职责是维护评估的科学性和客观性,而非迎合委托方的偏好。基于这一理论重构,可以建立“评估公共性保障机制”,通过法律法规明确评估的公共产品地位^[46],设立独立的评估监管机构,制订统一的评估基本准则,从制度源头上消除意向分化的合法性空间。

在中观治理结构层面,委托代理关系中的权力失衡是意向分化得以延续的结构性根源。传统的一对一委托模式赋予政府部门过大的控制权,第三方缺乏制约委托方的有效手段,导致专业判断让位于行政意志。解决这一问题需要引入多元制衡的治理结构,将单一委托关系转变为多方参与的治理网络。理论上,这种治理结构应当包含委托方、代理方、监督方、受益方四个基本角色,形成相互制衡的权力配置。委托方负责提出评估需求和提供基础条件,但不能干预评估过程;代理方负责进行专业化评估,享有方法选择和结论形成的自主权;监督方负责评估质量控制和程序合规性检查;受益方(包括被评估对象和社会公众)有权获得真实客观的评估信息并参与评估结果的应用监督。基于这一理论框架,可以构建“多元制衡治理架构”,在重要评估项目中引入独立监督主体,建立评估质量的同行评议制度,设立评估结果的社会监督平台,通过治理结构的多元化制衡消解权力失衡问题。

在微观技术标准层面,评估方法和质量标准的模糊性为不同意向的实现提供了操作空间。当评估缺乏统一的技术规范时,不同主体可以根据自身意向选择有利的标准和方法,从而导致评估结果具有可操纵性。解决这一问题需要建立高度专业化和标准化的技术体系,压缩意向实现的技术空间。理论上,这种技术标准体系应当具备科学性、统一性、可操作性三个基本特征。科学性要求标准建立在扎实的理论基础和充分的实证研究之上;统一性要求标准在全国范围内具有普遍适用性,避免因地区差异导致的标准不一;可操作性要求标准具体明确,能够指导实际操作并便于质量检验。同时,技术标准体系还应当具备动态调整能力,能够根据实践发展和学术进步不断完善提升。基于这一理论认知,可以完善“专业技术标准体系”,组织专业力量制订涵盖评估设计、实施、报告、应用全流程的技术规范,建立评估方法和工具的认证制度,设立评估质量的量化评价体系,通过技术标准的刚性约束减少主观意向对评估质量的影响。

这一系统性制度重构的深层逻辑是将第三方评估从服务政府转向专业自主,从部门工具转向公共产品,从经验判断转向科学规范。通过制度逻辑的统一、治理结构的平衡和技术标准的规范,不仅能够有效协调主体意向分化,而且可以推动第三方评估的发展,使其成为政府绩效管理现代化的重要支撑力量。

四、结语

(一) 研究结论与理论反思

本研究运用 Q 方法揭示了中国第三方政府绩效评估中一个被既往研究忽视的关键现象——主体意向分化。这种分化表现为政府部门内部的管理效能型、资源争取型和问责督导型意向并存,第三方

机构内部的专业价值型、市场生存型和公共责任型取向分化, 深层地反映了评估活动作为一项制度性实践所内含的复杂性。不同主体基于各自的制度位置、利益考量和价值追求, 对评估活动形成了迥异的功能期待和行为逻辑。这一发现呼应了建构主义评估理论关于评估主体能动性的核心论点^[29], 同时拓展了利益相关者分析框架对评估情境复杂性的理论解释^[47]。

意向分化的深层影响在于它触发了评估功能的系统性重构——从追求单一的客观评价转向服务多元化的利益诉求, 进而产生了“功能异化”现象。这种异化体现为评估目标的偏离、过程的形式化和结果应用的选择性, 也意味着评估制度在适应复杂治理环境过程中的自我调适。因此, 协调意向分化的关键并非消除差异, 而在于构建能够容纳多元意向并引导其良性互动的制度框架。

基于这一认知, 本研究提出的协调策略遵循“制度逻辑重构—治理结构平衡—技术标准规范”的递进逻辑: 宏观层面的制度逻辑重构旨在为多元意向提供统一的价值基础, 将评估从部门工具转向公共产品; 中观层面的治理结构平衡致力于破解权力失衡导致的意向冲突, 通过多元制衡实现利益协调; 微观层面的技术标准完善则通过专业化约束压缩意向分化的操作空间, 确保评估的科学性和一致性。

(二) 研究局限与未来展望

本研究的样本代表性还需进一步提高, 对来自不同层级、不同部门的评估主体的兼顾有所不足。此外, 还可进一步探讨主体意向分化的生成逻辑, 尤其是外部制度环境、内在心理机制的交互影响有待理论化。未来研究应进一步扩大研究样本, 提升结论的普适性; 深化情境化分析, 揭示意向分化的动态演化规律; 拓展跨案例、跨层级的比较视野, 探讨不同制度、文化语境下的异同。需要指出的是, 个体评估意向的形成并非独立于情境脉络而存在。前文分析的制度语境、利益格局及具体评估情境, 既是塑造个体意向的客观条件, 也在不断地被行动者建构着。个体行动受到情境规则的约束, 但也在与情境的互动中形成着对规则的反思性认知, 进而实现着行动策略的调整。对此, 未来有必要在案例分析的基础上进一步厘清制度、利益、情境等因素与个体认知的交互影响, 以期刻画“个体—情境”互构的生成逻辑。这对深化行动的过程性理解, 把握评估意向分化的动态性, 具有重要的理论价值。

注释:

- ① 访谈记录编码规则: 英文字母为受访者身份编码(ZF 代表政府; GX 代表高校; QY 代表企业), 前六位数字是访谈日期(年月), 最后两位数字是受访者编号, 如: ZF202312-15 意为 2023 年 12 月访谈的 15 号政府人员。

参考文献:

- [1] 中国行政管理学会课题组, 龚禄根, 包国宪, 等. 政府部门绩效评估研究报告[J]. 中国行政管理, 2006(5): 11-16.
- [2] 贞杰. 中国政府绩效管理 40 年: 路径、模式与趋势[J]. 重庆社会科学, 2018(6): 5-13, 2.
- [3] 何文盛, 姜雅婷. 系统建构视角下政府绩效评估结果偏差生成机理的解构与探寻[J]. 兰州大学学报(社会科学版), 2015, 43(1): 104-111.
- [4] 徐双敏. 政府绩效管理中的“第三方评估”模式及其完善[J]. 中国行政管理, 2011(1): 28-32.
- [5] 张菊梅. 县级政府绩效评估的发展及其典型模式[J]. 中南大学学报(社会科学版), 2013, 19(6): 165-169.
- [6] 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[J]. 中国行政管理, 2008(1): 49-51.
- [7] 尚虎平, 孙静. 失灵与矫治: 我国政府绩效“第三方”评估的效能评估[J]. 学术研究, 2020(7): 50-58, 177.
- [8] 王春业, 邓盈. 重要立法事项第三方评估机制研究[J]. 中南大学学报(社会科学版), 2017, 23(6): 27-33.
- [9] BEHN R D. Why measure performance? Different purposes require different measures[J]. Public Administration Review, 2003, 63(5): 586-606.
- [10] 周志忍, 徐艳晴. 政府绩效评估中博弈行为的防范之道: 理论与国际经验[J]. 中国行政管理, 2015(6): 140-145.

- [11] 徐双敏, 崔丹丹. 完善社会组织第三方评估工作机制研究: 基于 5 市调查数据的分析[J]. 中南财经政法大学学报, 2016(6): 52-57.
- [12] 陆明远. 政府绩效评估中的第三方参与问题研究[J]. 生产力研究, 2008(15): 121-122, 155.
- [13] 李春, 王千. 政府购买养老服务过程中的第三方评估制度探讨[J]. 中国行政管理, 2014(12): 38-42.
- [14] 徐双敏, 陈尉. “第三方”评估政府绩效的制度环境分析[J]. 学习与实践, 2013(9): 22-27.
- [15] 李晓杰, 毕金平. 公平竞争审查的实践困境与优化路径: 以第三方评估制度为例[J]. 当代经济, 2023, 40(10): 105-111.
- [16] 何阳, 孙萍. 精准扶贫第三方评估流程再造: 理论依据、现实动因与政策设计: 对民族地区精准扶贫第三方评估实践的反思[J]. 内蒙古社会科学(汉文版), 2018, 39(5): 177-183.
- [17] 祁中山. 扶贫绩效第三方评估: 价值与限度: 以 2016 年国家精准扶贫工作成效考核第三方评估为参照[J]. 信阳师范学院学报(哲学社会科学版), 2017, 37(6): 38-44.
- [18] 邓亚当. 我国政府绩效“第三方”评估存在的问题与对策分析[J]. 中共珠海市委党校珠海市行政学院学报, 2016(4): 23-28.
- [19] 张圣. 政府连环抉择机制: 行政委托第三方评估实践差异现象的一种解释[J]. 公共行政评论, 2022, 15(3): 119-139, 199.
- [20] 韩江风. 政府购买服务中第三方评估的内卷化及其优化: 以 Z 市 S 区社会工作服务评估项目为例[J]. 四川理工学院学报(社会科学版), 2019, 34(2): 20-37.
- [21] 陈那波, 陈永宏, 谭荷丹. 政府绩效评估中的“控制权”分配与行动逻辑[J]. 科研管理, 2022, 43(5): 191-199.
- [22] 何文盛, 姜雅婷. 政府绩效评估“三权”视角下第三方评估权力组织模式与评估结果偏差: 基于多案例的阐释[J]. 江苏行政学院学报, 2020(4): 98-105.
- [23] 周跃辉. 基于信息不对称理论的第三方评估问题研究[J]. 经济研究参考, 2015(48): 5-9, 21.
- [24] 阎波. 公共组织第三方评估何以可为? 一项双案例比较研究[J]. 经济社会体制比较, 2018(3): 125-135.
- [25] 王卓, 丁洪彬. 脱贫摘帽第三方评估的地方逆向因应实证研究: 基于纳什均衡理论视角[J]. 经济理论与经济管理, 2021, 41(9): 85-98.
- [26] 耿磊磊. 第三方评估视域下精准扶贫中乡村干部行为探析: 以湖北省 H 乡为例[J]. 湖北经济学院学报(人文社会科学版), 2020, 17(8): 19-22.
- [27] 盛明科, 彭国甫. 用元评估分析政府绩效评估偏差[J]. 湘潭大学自然科学学报, 2004, 26(1): 136-139.
- [28] 吴建南, 白波. 评估政府绩效评估: 元评估方法的探索性应用[J]. 行政论坛, 2009, 15(6): 32-37.
- [29] GUBA E G, LINCOLN Y S. Fourth generation evaluation[M]. California: Sage Publications, 1989.
- [30] MERTON R K. The unanticipated consequences of purposive social action[J]. American Sociological Review, 1936, 1(6): 894-904.
- [31] STEPHENSON W. Technique of factor analysis[J]. Nature, 1935, 136(3434): 297-297.
- [32] BROWN S R. A primer on Q methodology[J]. Operant Subjectivity, 1993, 16(3/4): 91-138.
- [33] RAMLO S E, NEWMAN I. Q methodology and its position in the mixed methods continuum[J]. Operant Subjectivity, 2011, 34(3): 172-191.
- [34] 李延伟. 公共行政研究中的主观性探索: Q 方法的方法论价值[J]. 中国行政管理, 2022(9): 132-139.
- [35] 马佳铮. 执行力认知的群体偏好聚合与公共政策场域的适配性研究: 以 44 名处级官员为样本的 Q 方法探索[J]. 中国行政管理, 2022(8): 82-89.
- [36] JEFFARES S, SKELCHER C. Democratic subjectivities in network governance: A q methodology study of English and Dutch public managers[J]. Public Administration, 2011, 89(4): 1253-1273.
- [37] ZABALA A. Q method: A package to explore human perspectives using Q methodology[J]. The R Journal, 2014, 6(2): 163-173.
- [38] WEBLER T, DANIELSON S, TULER S. Using Q method to reveal social perspectives in environmental research[J]. Greenfield MA: Social and Environmental Research Institute, 2009, 54(1): 45.
- [39] DRYZEK J S, BEREJIKIAN J. Reconstructive democratic theory[J]. American Political Science Review, 1993, 87(1): 48-60.
- [40] 财政部修订《政府采购货物和服务招标投标管理办法》[J]. 预算管理会计, 2017(9): 2-10.

- [41] 财政部预算司. 深化财政改革加快建立现代财政制度[J]. 中国财政, 2020(21): 11-14.
- [42] 蒋悟真, 李其成, 郭创拓. 绩效预算: 基于善治的预算治理[J]. 当代财经, 2017(11): 26-35.
- [43] 李燕, 王宇龙. 论绩效预算在我国实施的制度约束[J]. 中央财经大学学报, 2005(6): 11-14.
- [44] 蒋天佑. 扶贫绩效第三方评估客观性消解及其应对: 基于 A 县个案和相关者利益博弈视角的分析[J]. 湖南农业大学学报(社会科学版), 2018, 19(5): 68-73.
- [45] GIDDENS A. The constitution of society: Outline of the theory of structuration[M]. Berkeley: Univ of California Press, 1984.
- [46] 包国宪, 戴正. 中国政府绩效管理的演进逻辑、面临困境与治理转型[J]. 行政论坛, 2024, 31(4): 101-111.
- [47] BRYSON J M. What to do when stakeholders matter: Stakeholder identification and analysis techniques[J]. Public Management Review, 2004, 6(1): 21-53.

Consensus and divergence of subject intentions in the third-party evaluation of government performance: Analysis based on Q methodology

BAO Guoxian, DAI Zheng

(School of Management, Lanzhou University, Lanzhou 730000, China)

Abstract: This study, by employing Q methodology to investigate the phenomenon of stakeholder intention divergence and coordination pathways of the subject intention in China's third-party government performance evaluation system. The research identifies such three distinct intention types among government departments as management-efficiency-oriented, resource-acquisition-oriented, and accountability-supervision-oriented, while third-party institutions demonstrate profession-value-oriented, market-survival-oriented, and public-responsibility-oriented orientations. This intention divergence is rooted in deep-seated mechanisms including institutional dualism, asymmetric interest configurations, and pluralistic evaluation contexts, resulting in "functional alienation" characterized by goal displacement, procedural formalization, and selective utilization of evaluation outcomes, posing challenges to the scientific rigor and effectiveness of performance evaluation. The study further proposes the following systematic coordination strategies: at the macro level, reconstructing the institutional logic framework by reorienting evaluation from "outsourced government services" to "professional governance support"; at the meso level, establishing multi-stakeholder checks-and-balances governance architecture to address power imbalances inherent in principal-agent relationships; and at the micro level, enhancing professional technical standards systems to constrain the operational space for intention divergence.

Key words: commissioned third-party evaluation; Q methodology; intention divergence; performance evaluation

[编辑: 郑伟]