

服务型政府与政府绩效评估体系创新

——基于德尔菲调查法的发现

杨宇谦 吴建南 马亮

内容提要: 文章采用德尔菲调查法,从政府绩效评估体系的定位、指标、利益相关者和结果使用四个方面考察了创新的现状与趋势。研究发现,从服务型政府理念出发,绩效评估体系创新的关键是突出战略思维、强化公众参与、鼓励创新导向和重视能力建设,但是调查结果显示,这些方面均非目前绩效评估体系创新的热点,也非未来可行的创新途径。分析表明,高度响应服务型政府思想的绩效评估体系创新可能尚处于“蛰伏”状态,需要外部环境的支持作为其产生条件。

关键词: 政府绩效评估 德尔菲调查法 服务型政府

一、引言

近年来,建设服务型政府成为推动我国行政体制改革的核心目标和关键使命(中国行政管理学会课题组,2005)。地方政府在职能定位、规制手段、服务模式、管理流程等方面不断创新,以增进公民、顾客和市场主体的满意度(陈振明,2008)。在此同时,地方政府绩效评估体系创新也在进行之中,体现出各种新做法不断涌现的态势(吴建南、杨宇谦,2009)。那么,目前政府绩效评估体系的创新在多大程度上响应了服务型政府理念的要求?针对这一问题,整体性的评估有待进行。

本文采用在2008年2~3月和同年8~9月期间,使用德尔菲调查法面向中国行政管理学会政府绩效管理研究分会(全国政府绩效管理研究会)理事发放问卷所获得的数据进行分析。所选问卷题目围绕市级政府绩效评估活动的定位、绩效评估指标、利益相关者和结果使用方式四个方面展开,共有16个城市26名政府绩效评估专家完整地参与了两轮调查。

论文先从目前使用优先权、未来需求度和未来可行性等方面构建了用于识别现有绩效评

作者简介:杨宇谦,西安交通大学博士生。吴建南,西安交通大学公共政策与管理学院副院长、全国政府绩效管理研究会副会长、博士生导师。马亮,西安交通大学公共政策与管理学院博士生。

本文是国家自然科学基金项目“目标责任考核、财政管理与组织创新——面向中国地方政府绩效改进的实证研究”(项目编号:70873092)的阶段性成果。作者感谢全国政府绩效管理研究会的领导和理事对本次德尔菲研究的大力支持。论文主要研究结果曾在2010年11月9日由国家行政学院公共管理研究中心、上海市奉贤区人民政府、中国行政管理学会教学研究会联合举办的“服务型政府建设的理论与实践”政府管理创新前沿论坛上被宣读。作者感谢到会领导和专家学者提供的建议,感谢2011年6月12日在西安交通大学主办的“华人公共管理学者研讨会”上为此研究提出建议的专家学者。

限于版面,文章发表时有删减,需要全文的读者请与编辑部联系。——编者注

估体系创新状况与趋势的分析框架;然后,基于德尔菲调查数据,描述市级政府绩效评估体系创新的基本特征,并与有关服务型政府理念倡导的创新要求进行比较。

通过识别目前已经广为采纳并将在未来持续下去的创新(本研究称之为“明日”的创新),目前并不普及但具有在未来进入实践潜力的创新(本研究称之为“明年”的创新),以及目前并不普及且在未来具有不确定性特征的创新(本研究称之为“蛰伏”的创新),本文发现,服务型政府的理念并未深入和广泛地影响市级政府绩效评估体系创新,其要求的创新内容尚处在“蛰伏”状态之中。本研究将结合如上发现,讨论研究的实践意义。

二、创新内容的分析框架

政府绩效评估体系的发展历程,是适应政治、社会和经济条件而形成的产物。西方发达国家的经验表明,绩效评估体系存在不同的发展阶段,在不同的历史环境中,绩效评估体系创新实践主题也各具特色(蓝志勇、胡税根 2008)。一般认为,绩效评估体系中的创新虽然形式内容多样,但概括起来,主要是在三个方面具有显著差异。第一,一些创新的积极效果已经显现,持续创新还能够持续产生积极效果,而另一些创新目前尚未进入主流实践,而是处在试点或者论证阶段。换句话说,绩效评估体系的创新在实践之中的优先权是不同的,一些创新在帮助政府提高决策质量和管理水平、改进与公众的互信关系方面已经成为实践的主流,而另一些尚未推广普及(Cavalluzzo and Ittner 2004)。第二,利益相关者对于不同创新的需求程度存在差异。研究发现,公众参与绩效评估活动之所以在发达国家的创新过程中较为活跃,是同地方议会、非营利组织和公众个人的大力倡导和影响有关(Dalehite 2008)。而在不发达国家,利益相关者还难以合作促成类似绩效评估体系的创新。第三,不同创新在其产生所必须的资源积累方面存在差异。

目前,在中国地方政府绩效评估体系的发展中,以上几个方面的特征相互组合,形成了三类典型的绩效评估体系创新。

第一类绩效评估体系的创新具有较高的目前优先权、未来需求度和未来可行性,同时,人们对其未来可行性的判断比较确定。例如,目标责任考核是中国地方政府目前使用最为广泛的评估工具,它对于推动政府管理和提高行政效率的积极作用已经得到了各方面的认可。在未来发展中,这种创新将继续占据主要地位,并且在实践中探索与完善(吴建南、杨宇谦, 2009)。为了突出这种创新在未来出现的密集性和连续性,本研究将具有较高的目前优先权、未来需求度和未来可行性的绩效评估体系创新,称为“明日”的创新。

第二类绩效评估体系的创新,虽然目前优先权较低,但是具有较高的未来需求度和可行性。换句话说,这种创新已经具备了进入实践的必要性和可行性条件。例如,20世纪90年代开始尝试的“服务承诺制”将公众绩效评估作为重要环节。虽然在当时,地方政府并没有广泛开展公众参与政府绩效评估的先例,但是,地方政府主要领导和社会公众对于这种创新的需求较为迫切,同时在技术条件上并不存在太多的阻碍,因此,这种创新的扩散速度很快(Foster, 2006)。由于此类创新在目前并不流行,但创新潜力巨大,本研究将其称为“明年”的创新,表示同第一类创新相比,它们进入实践的起始时间距离目前可能稍远一些。

第三类创新,从目前优先权、未来需求度和未来可行性三方面来看不尽人意,人们对其未来可行性也并不确定。从强度和未来发生的可能性上看,这种类型的创新显然是较弱的,本研

究将这种创新称为“蛰伏”的创新。在实践中,网上评价政府绩效可能就属于这种创新。学者发现,由于在数据公开和结果实用性方面存在缺陷,很多地方政府并不把它作为高优先权的评估工具,对其未来的进一步发展需求也并不高,只有在地方领导充分重视并提高其可行性之后,这种创新才会进入实践(Hartford 2005)。该分析框架如表1所示。

表1:绩效评估体系创新的分析框架

类型	目前优先权	未来需求度	未来可行性	对未来可行性的确定程度
“明日”的创新	高	高	高	高
“明年”的创新	低	高	高	高或低
“蛰伏”的创新	低	低	低	低

通过识别如上三类创新的具体内容,可以对目前政府绩效评估体系创新在多大程度上符合服务型政府理念的要求进行比较。如果本研究界定的“明日”和“明年”的创新与服务型政府所要求的战略思维、公众参与、创新导向和能力发展等因素具有对应关系,就可以初步得出服务型政府理念对于绩效评估体系创新已经具备显著影响的判断。反之,如果在“明日”和“明年”的创新中,服务型政府要求的要素体现得并不明显,在“蛰伏”的创新中体现更多,则可能说明以服务型政府为导向的绩效评估体系创新趋势在未来可能难以获得突出的地位。在区分具体绩效评估体系创新究竟属于以上哪种创新类型时,本研究采用德尔菲调查方法,在全国范围内向熟悉和关心政府绩效评估体系的专家发放调查问卷。

三、研究方法

德尔菲调查法是在一组地域分散的专家之间进行组群通信的活动,它允许专家系统性地处理一个复杂的问题或任务(Ziglio,1996)。它是介于一次性问卷调查与其他介入性研究之间的调查方法。德尔菲方法实施的关键有三个方面,一是被调查者对所在政策领域和研究问题的熟悉程度,二是调查问题的覆盖面和重要性,三是调查者与被调查者之间的相互信任程度。

本文采用2008年“中国市级政府绩效评估体系‘德尔菲’调查”的数据进行分析。该调查以全国政府绩效管理研究会理事(共151名专家)作为专家库,他们在政府绩效评估领域的专业背景和知识可以充分保证调查的有效性。采用邮寄(快递、挂号信、平信相结合)的方式将问卷发放给选定专家。第一轮调查在2008年2~3月进行,在全国范围内回收34份有效问卷,学者与官员数量平均分布。第二轮在2008年8~9月进行。在第二轮调查中有26位专家给予回复。在系统检索和考察国内市级政府绩效评估体系创新实践的基础上,本文选取了德尔菲调查的四个方面进行分析(如表2所示)。

专家调查问卷采用5级量表的形式,询问了专家四个方面的问题:(1)“现在的使用率”题目,请专家填答其所在城市政府目前在绩效评价活动中某个题目在实践中具有怎样的使用情况,“1”表示完全不用、“3”表示介于两者之间、“5”表示到处使用。(2)“未来的需求度”题目,请专家填答其所在城市在未来三年中,使用某个题目所描述的内容在多大程度上是该城市政府所期望的,“1”表示没有需求、“3”表示介于两者之间、“5”表示极为需要。(3)“未来的可行性”题目,请专家填答其所在城市在未来三年中,使用某个题目所描述内容的可行性如何,“1”表示低、“3”表示中等、“5”表示高。(4)“未来可行性之确定性”题目,请专家填答其所在城市

在未来三年中,对某个方面所描述的内容可行性的判断在多大程度上是确定的,“1”表示低、“3”表示中等、“5”表示高。

表 2: 绩效评估创新的考察清单

组成部分	题数	题项设置
评价定位	12	服务创新、合作创新、流程创新、激励、晋升(面向上级领导和同级人大满意)、晋升(改善可持续发展)、晋升(面向公民满意)、控制下级、问责、组织学习、绩效预算、降低成本
绩效指标	10	数量、质量、单位成本、公平、效果、性价比、市民满意度、服务对象满意度、问责、廉洁
利益相关者	19	市民、公共服务对象(公民个人)、国有企业、私营企业、区县政府、财政局、人事局、审计局、监察局、省级政府、党的纪律检查部门、党的组织部门、专家学者、工会、省长、市长、市委书记、人大、政协
结果使用	11	透明公开、绩效改进、沟通和协商工作目标、诊断和控制问题、组织学习与发展、激励下属、实施奖惩、调整工作安排、配置组织资源、为预算做准备、识别培训机会

根据专家的得分,如果所调查题目的得分均值超过该领域中所有调查题目得分的均值时,认为该题目所反映的创新在特定评判标准上的评价较高。例如,在有关绩效指标创新的可行性评估方面,如果调查结果表明,专家认为“数量”在未来三年中作为绩效指标的可行性较之于“数量、质量、单位成本、公平、效果、性价比、市民满意度、服务对象满意度、问责、廉洁”在内的全部题目的可行性评估均值更高,则认为“数量”作为绩效指标的将来可行性较高,说明“数量”指标成为未来绩效指标的可行性更高;反之,则认为“数量”成为未来创新的可行性较低。如果“数量”这一指标在目前优先权、未来需求度、未来可行性与不确定性的评价方面同时符合三类创新中任一类的判断标准时,则认为“数量”这一绩效指标的创新属于某类创新。

四、绩效评估体系创新的现状与趋势

(一) 评估定位的创新

表 3 报告了绩效评估定位各种创新的目前优先权、未来需求度、未来可行性,以及对可行性的确定性评价。

表 3: 绩效评估定位

评估定位	CP	FD	FF	C
晋升: 使上级领导确信该政府做得很好	3.71	3.96	3.72	3.00
晋升: 通过促进可持续发展来增加晋升机会	3.59	4.08	3.54	3.00
控制: 确保下属做正确的事	3.48	3.77	3.77	3.08
流程创新: 政府通过新的方式来提供服务	3.48	4.00	4.00	3.16
晋升: 通过提高公民满意度来增加晋升机会	3.38	3.81	3.54	3.12
服务创新: 政府实施面向公民的新服务	3.36	4.23	3.73	3.04
激励: 激励员工做有利于改进绩效的事	3.30	3.81	3.50	2.96
问责: 为行为(尤其是错误)承担责任	3.32	3.96	3.68	3.04
降低成本: 确保当地行政成本不致太高	3.24	4.00	3.65	2.96
合作创新: 政府与其他机构或实体合作以进行创新	2.76	3.50	3.04	2.69
预算: 在特定领域、群体和项目上规划公共资金使用	2.72	3.69	3.50	2.92
学习: 找出工作中干得好坏的内在原因	2.63	3.58	3.42	2.84

注(以下各表相同): CP = 目前的优先权; FD = 未来需求度; FF = 未来可行性; C = 对可行性的确定程度。

目前优先权中得分最高的创新均与官员晋升有关。德尔菲专家认为,用于“使上级领导以及同级人大确信该政府做得很好”,以及通过“促进可持续发展来增加晋升机会”的创新目前优先权最高。“控制”和“流程创新”并列第三。目前优先权最低的分别是“学习”和“预算”,两者得分的均值均低于3分。

在未来需求度方面,“通过促进可持续发展来增加晋升机会”这一创新的未来需求度最高。“服务创新”和“降低成本”同样需求度较高。“合作创新”的未来需求度最低。所有需求度的得分都在3分以上。

在未来可行性方面,“服务创新”的未来可行性最高,其次为“控制”,也就是“确保下属做正确的事”。学习即“找出工作中干得好坏的内在原因”的得分较低。“合作创新”的未来可行性最低。所有创新的未来可行性都获得了高于3分的评价。

在未来可行性的确定性方面,德尔菲专家最确定的是“流程创新”的可行性。专家最不确定的是“政府与其他机构或实体合作以进行创新”成为评估定位的可行性。

(二) 绩效评估指标的创新

由表4可以看出,在不同的绩效指标中,目前优先权占前三位分别是“正式的效果”、“廉洁”和“市民满意度”。它们的得分都在3.5分以上。“回应性”、“公平性”和“单位产出的成本”得分处在最后三位。其中,“单位产出”的成本得分低于3分。

表4: 绩效指标

评估指标	CP	FD	FF	C
正式的效果: 例如 GDP 实际增长; 通过考试学生百分比	3.96	4.00	3.92	2.92
廉洁: 例如接受贿赂的领导或官员的数目	3.70	4.00	3.65	2.69
市民满意度: 例如市民对于当地政府的总体满意度	3.58	4.38	3.92	3.04
问责性: 例如对于当地反腐倡廉人员的投诉状况	3.38	4.15	3.62	2.58
数量: 例如公共服务接受者的数目	3.38	3.88	3.77	2.81
质量: 例如在过度拥挤的学校上学的百分比	3.25	3.69	3.73	2.73
用户满意度: 例如读者对于图书馆员和开放时间的满意度	3.21	3.88	3.77	2.88
回应性: 例如急救服务的响应时间	3.17	4.12	3.85	2.88
公平性: 例如不同阶层居民对于文化休闲设施满意的百分比	3.13	4.00	3.54	2.62
单位产出的成本: 例如每个家庭垃圾收集的成本	2.46	3.77	3.31	2.65

在未来需求度方面,“市民满意度”成为绩效评估指标的需求得分最高,其次为“问责”和“回应性”。其他题目的得分也在3分以上。其中,“廉洁”、“正式的效果”和“公平性”的均值都为4分。“质量”在所有绩效指标中的未来需求度的得分最低。

在未来可行性方面,“市民满意度”和“正式的效果”作为指标创新的可行性最高,但是均未超过4分。“问责性”平均分为3.62,在所有题目中排名倒数第三。“公平性”在未来成为绩效指标创新的可行性同样偏低,但在3.5分以上。“单位产出的成本”的可行性最低。

在未来可行性的确定性方面,德尔菲专家对于“市民满意度”在未来成为绩效评估指标的可行性评价最高。对于其他题目,确定性的得分均小于3分。其中,“问责”成为未来政府绩效评估指标的确定程度最低。

(三) 绩效评估利益相关者的创新

表5报告了绩效评估体系中利益相关者的情况。在目前优先权方面,获得4分以上的利益相关者从高到低依次是“党委”、“纪律检查部门”、“党的组织部门”、“人民代表大会”和“市长”。这些利益相关者的共同特征是掌握或部分掌握领导干部的人事管理权限。“市民”和“工会”的目前优先权最低。

表5: 利益相关者

利益相关者	CP	FD	FF	C
学术专家	3.33	4.46	4.19	3.35
企业: 国有企业	3.29	3.65	3.35	3.00
企业: 私有企业	2.72	3.62	3.35	2.92
个人用户: 如医院患者、学生家长	2.52	3.92	3.50	3.04
市民: 如一般公众	2.48	3.92	3.35	2.96
工会	2.28	3.31	3.00	3.04
市级党委	4.32	4.27	4.31	3.23
党的纪律检查部门	4.20	4.27	4.46	3.44
党的组织部门	4.16	4.36	4.40	3.28
人民代表大会	4.04	4.23	4.23	3.27
市长	4.00	4.23	4.23	3.27
审计局	3.92	4.12	4.08	3.40
人事局	3.92	3.65	3.84	3.24
更高一级政府: 省政府	3.92	4.08	4.15	3.19
统计局	3.76	4.35	4.35	3.42
监察局	3.72	4.35	4.15	3.19
政协	3.52	4.04	3.88	3.15
县或区政府(含县级市)	3.44	3.77	3.81	3.31
财政局	3.17	4.12	3.92	3.19

在未来需求度方面,平均得分最高的利益相关者依次为“学术专家”、“党的组织部门”、“统计局”和“监察局”。这些利益相关者的未来需求度都高于3分。未来需求度得分最低的利益相关者是“国有企业”、“私营企业”和“工会”。

在未来可行性方面,“党的纪律检查部门”、“党的组织部门”、“统计局”和“市级党委”的可行性最高。其他利益相关者也获得了较高的可行性评价,均在3分以上。“工会”是未来可行性得分最低的利益相关者。

在未来可行性的确定性方面,多数利益相关者的得分在3分以上。其中“党的纪律检查部门”、“统计局”、“审计局”和“学术专家”的得分较高。德尔菲专家对于两类利益相关者参与绩效评估体系创新的确定性评价较低,一是“一般公众”、二是“私有企业”。

(四) 结果使用

绩效评估结果具有不同的使用方式。表6报告了统计结果。在目前优先权方面,“实施奖惩”、“改进绩效”与“诊断与控制问题”的得分最高。“为预算做准备”的目前优先权最低,并且在3分以下。

表 6: 结果使用

结果使用	CP	FD	FF	C
实施奖惩	3.80	4.35	4.19	3.15
改进绩效	3.56	4.46	4.19	3.23
诊断与控制问题	3.48	4.15	3.65	2.92
与下级沟通和协调工作目标	3.32	4.15	3.92	3.08
激励员工	3.32	3.96	3.77	3.12
调整工作安排	3.20	3.69	3.69	3.12
配置组织资源	3.12	3.81	3.62	2.92
组织学习与发展	3.04	3.81	3.54	3.00
向公众公开评价结果	3.00	4.54	4.19	3.19
提供培训机会	2.96	3.50	3.27	2.96
为预算做准备	2.88	3.77	3.38	2.77

在未来需求度方面,“向公众公开结果”、“改进绩效”和“实施奖惩”的需求度最高。所有题目的得分均在 3 分以上,这说明在未来,评估结果使用方式的创新需求是多方面的。“为预算做准备”、“调整工作安排”和“提供培训机会”成为结果使用方式的需求度最低。

在未来可行性方面,“向公众公开结果”、“改进绩效”和“实施奖惩”的可行性最高。所有题目的得分都在 3 分以上,说明绩效评估结果使用方式的创新空间较广。“为预算做准备”和“提供培训”机会的可行性最低。

在未来可行性的确定性方面,德尔菲专家同样对“向公众公开结果”、“改进绩效”和“实施奖惩”的可行性最为确定。多数题目获得了大于 3 分的确定性得分。“提供培训机会”、“诊断与控制问题”、“配置组织资源”和“为预算做准备”的可行性得分均在 3 分以下。

(五) “明日”、“明年”与“蛰伏”的创新分布

分阶段创新点的识别结果如表 7 所示。表 7 显示了“明日”的创新、“明年”的创新和“蛰伏”的创新分节,分别从评估定位、评估指标、评估中的利益相关者和评估的结果使用四个方面列举了未来各阶段的创新内容。

表 7: 各类创新的具体分布

	“明日”的创新	“明年”的创新	“蛰伏”的创新
评估定位	使上级领导满意从而晋升,流程创新,问责	降低成本	学习、合作创新、预算
评估指标	正式的效果,市民满意度	回应性	单位产出的成本
评估中的利益相关者	市级党委、党的纪律检查部门、党的组织部门、人民代表大会、市长、审计局、更高级政府、统计局、监察局	学术专家	国有企业、私有企业、个人用户、市民、工会
评估的结果使用	实施奖惩,改进绩效,同下级沟通协调工作目标	向公众公开评价结果	配置组织资源,组织学习与发展,提供培训机会,为预算做准备

五、结 论

通过德尔菲调查法,本文对中国市级政府绩效评估体系的创新现状和未来趋势进行了初步分析,以期识别整体的政府绩效评估体系创新在多大程度上回应了服务型政府理念的要求。

从现状来看,绩效评估的定位、指标、利益相关者和结果使用方式等方面的主流创新实践,并没有体现服务型政府所要求的战略思维、公众参与、创新导向和能力发展等要素。同时,一些目前优先权不高,但是具有较高未来需求度和可行性的创新,同样未能较为充分地回应服务型政府的要求。

本研究进一步发现,服务型政府所倡导的创新,具有在目前优先权偏低、未来需求度和可行性较低,同时可行性的不确定性较高的特点。按照本文的说法,它们尚处在“蛰伏”的状态之中,只有在现有行政环境不断变化,条件充分成熟之后,才有可能产生。由于目前绩效评估体系的创新并没有直接有力地回应服务型政府的要求,服务型政府和绩效评估体系的创新可能并未充分衔接,为了推动两者充分融合,地方政府应进一步考虑从服务型政府的理念出发,完善现有的绩效评估体系,充分借鉴国外政府在战略规划、公众参与等方面的经验,提高绩效评估对于推动服务型政府的管理价值。

本文的局限在于,从数量上看,调查覆盖的样本城市较少,专家学者的数量也不多,可能存在抽样上的偏差;另外,本文所提出的创新类型识别框架,仅仅考虑了目前优先权、未来需求度和未来可行性和确定性,未能纳入更多复杂的因素,因此还需要不断扩展。后续研究可以采用案例研究方法,深入分析某地绩效评估体系变迁的历史经验。另外,可以采用大范围样本调查,分析绩效评估体系的差异和服务型政府建设成效之间的关系,以提供更多的经验和证据。

参考文献:

- 薄贵利 2005 “完善政府公共服务职能,加强服务型政府建设”,《国家行政学院学报》2005 6: 23。
- 陈振明 2008 “深化行政体制改革,加快服务型政府建设——中国政府改革与治理的新趋势透视”,《福建行政学院学报》2008 4: 5—12。
- 蓝志勇、胡税根 2008 “中国政府绩效评估:理论与实践”,《政治学研究》2008 3: 106—115。
- 唐铁汉 2008 “建设服务型政府与基本公共服务均等化”,《国家行政学院学报》2008 2: 8—12。
- 王川兰 2005 “论服务型政府的困境、超越与建构”,《公共管理学报》2005 2(4): 19—25。
- 吴建南、杨宇谦 2009 “地方政府绩效评估创新:主题、特征与障碍”,《经济社会体制比较》2009 5: 152—158。
- 中国行政管理学会课题组 2005 “服务型政府是我国行政改革的目标选择”,《中国行政管理》2005 4: 5—8。
- Cavalluzzo S. K. and D. C. Ittner 2004. “Implementing Performance Measurement Innovations: Evidence from Government.” *Accounting, Organization and Society*. 2004 29: 243—67.
- Foster K. W. 2006. “Improving Municipal Governance in China: Yantai’s Pathbreaking Experiment in Administrative Reform.” *Modern China*. 37(2): 221—50.
- Hartford K. Dear Mayor 2005. “Online Communications with Local Governments in Hangzhou and Nanjing.” *China Information*. 19: 217—60.
- Ziglio E. 1996. “The Delphi Method and Its Contribution to Decision-Making.” M. Alder and E. Ziglio eds. *Gazing into the Oracle: The Delphi Method and Its Application to Social Policy and Public Health*. London: Jessica Kingsley.

(责任编辑: 郜继红)