
出版说明

公共管理是一个涉及众多学科的重大理论与实践领域。随着全球化、信息化和知识经济时代的来临,当代国外的公共管理尤其是政府改革与治理的理论和实践已经发生了深刻变化,公共部门尤其是政府治理与改革的实践催生了大量的公共管理新思潮、新理论、新主题及新学科分支。在我国改革开放和现代化建设事业的发展过程中,公共管理尤其是在政府改革与治理方面产生了大量亟须解决的重大问题,新的改革实践呼唤新的理论,巨大的现实与理论需求必将有力地推动我国公共管理学科的迅速发展。这就要求公共管理学者不断跟踪研究公共管理的理论与实践前沿问题,拓展公共管理的学科视野,进行公共管理的知识创新。

有鉴于此,我社决定约请著名公共管理学者、厦门大学公共事务学院陈振明教授担任主编,组织编写了这套“公共管理学科前沿丛书”,每本书选择公共管理的一个分支学科,尤其是新研究领域或重大主题,基于国内外的最新文献与事实资料,介绍该学科、领域或主题的基础理论、学术前沿及实践进展,以较小的篇幅,勾画出这一学科、领域或主题的基本轮廓。这套丛书将分期分批陆续出版,既可以作为公共管理、政治学和社会学等学科各专业的本科生、研究生以及MPA学员的相对应课程的教材,也可以作为公务员培训的教学参考书,还可以作为其他对公共管理感兴趣的读者的学习参考书籍。

上海人民出版社

2006年2月

目 录

前言	1
1 导论	1
1.1 政府绩效评估的核心概念	1
1.2 政府绩效评估的基础理论	9
1.3 美国政府绩效评估的研究意义与研究现状	18
1.4 本书的研究体系	23
2 美国政府绩效评估的历史演进	26
2.1 投入、产出与效率:效率评估阶段	28
2.2 产出、结果(outcome)与顾客满意:结果(outcome)评估阶段	39
2.3 小结	50
3 美国政府绩效评估的主要基础	51
3.1 理论基础:实用主义与科学理性主义	51
3.2 价值观基础:质疑政府与金钱至上	59
3.3 政治制度基础:分权化政治架构	66
3.4 小结	78
4 美国政府绩效评估的基本环节	80
4.1 绩效评估的发起	80
4.2 绩效评估的操作设计	85
4.3 绩效评估的方案实施	105

4.4	小结	115
5	美国政府绩效评估的主要特征	116
5.1	绩效评估的静态特征	116
5.2	绩效评估的动态特征	121
5.3	小结	131
6	中国政府绩效评估的实践探索	132
6.1	中国政府绩效评估的现状	132
6.2	中国政府绩效评估的基础	136
6.3	中国政府绩效评估的主要特征	148
6.4	小结	154
7	美国政府绩效评估的经验借鉴	156
7.1	美国政府绩效评估经验的普适性局限	156
7.2	中国政府借鉴美国经验的基本思路	161
7.3	小结	175
	结束语	177
	附录	179
	主要参考文献	193
	后记	201

前 言

自 20 世纪七八十年代以来, 西方各国通过新公共管理运动对政府的内部运作模式以及外部管理机制都进行了重大的调整与变革, 最终在相当大程度上重塑了政府以及政府与社会的关系。在这一声势浩大的改革运动中, 各国政府的改革“魔杖”都是将工商管理工具引入政府实践, 其中, 绩效管理与评估是在新公共管理运动中被各国政府倍加推崇的政府改革工具。

作为管理主义实践的重要产物, 政府绩效评估是当前公共管理领域的前沿课题, 无论是实际工作部门还是理论界, 人们对政府绩效评估的热忱都处于不断升温的状态中。只是, 从目前世界各国的情况看, 不同国家在政府绩效评估实践方面的进展存在很大的差异。相比之下, 美国政府绩效评估早在 20 世纪早期就已经开始了, 并且延续至今, 对美国政府的管理实践产生了深刻的影响。但是, 从当前的情况看, 美国政府绩效评估的实践效果却无法与英国等后起之秀相比。一个非常典型的表现就是, 当我们进入英国任何一个层级的政府网站时, 都可以非常容易地查阅到有关政府绩效测量指标以及多个层次的绩效目标的信息, 但是, 要想在美国各层级的政府网站上查找相关信息却困难得多, 只有在堪称绩效测量典范的一些州或地方政府的网页上, 才能够比较容易地找到这样的信息。美国政府绩效评估的发展状态与趋势所显现出的特殊性, 促使我们去思考诸多问题, 如为什么美国政府绩效评估开始最早却落后于后起之秀的英国等国家? 对这些问题的解答除了能够帮助我们更为深刻地认识美国政府绩效评估之外, 还可以引导我们思考如何才能有效地借鉴美国政府绩效评估的先进经验。

我国政府绩效评估的实践开始于 20 世纪 90 年代中后期, 经过十几年的发展, 已经在从形式到内容各方面都取得了一些进展, 但是, 这一实践的推进却一直相对缓慢, 这与当前我国社会的飞速发展是相当不协调的。这是因为, 作为一种政府监督机制, 政府绩效评估在构建绩效型和服务型政府的过程中起着至关重要的作用。虽然我国当前一些地方政府在建设服务型政府的过程中, 也尝试着将对服务型政府建设

的问题纳入到对政府绩效的考查中去,从而使得政府绩效评估与地方政府的战略发展导向取得了一致性,但是,这种努力还没能取得明显的成效。同时,在十六届四中全会上,党中央提出要坚持科学发展观、构建社会主义和谐社会,这种要求对当前单纯以GDP增长为主导的评估体系形成了挑战,需要各级政府从经济、社会、政治、生态、文化等多方面重构现有的绩效评估体系。对于正处于政府绩效评估探索阶段的各级政府来说,要迎接这一挑战,可选择的方案之一就是学习西方国家政府绩效评估的做法。

基于对以上问题的思考,本书的目的就是:分析美国政府绩效评估并对比中国的实践,最终对中国政府如何借鉴美国的经验提供基本的思路。围绕着这样的目的,本书在以下方面进行了努力,从而也形成了本书的特点:

首先,对美国政府绩效评估的分析采取环环相扣的方式,逐步展开。本书尽量避免对美国实践进行简单的平面扫描,而是将看似平面的几个重要方面之间的关联性充分挖掘出来,并将其综合在一起,即先从历史发展的叙述中概括出这一实践的基础,进而分析由这一基础引导的过程,最后基于过程的分析提炼这一实践的特征,从而使得各部分内容一气呵成。

其次,为了有针对性地提出中国政府借鉴美国经验的思路,本书在讨论中国政府绩效评估的基础时以与美国一一对应的方式展开,然后推演出中国实践的特征,继而在中美两国特征对比的前提下找到经验借鉴的切入点,最终由这一切入点返回到基础层面,确定中国政府借鉴美国经验应首先对哪些不相适应的基础进行调整。在以上分析的基础上,本书最后讨论了中国政府绩效评估的基础调整之后还应该注意的问题。

最后,因为中美两国的实践存在着较大的差异,因此,本书还尽量提供了大量的案例,借助于对案例的分析,加深和确保读者对两国实践的准确认识。如在美国的实践讨论中,提供了得克萨斯州、衣阿华州等的案例,而在讨论中国政府绩效评估时,则提供了福建莆田、厦门和山东青岛等地的政府绩效评估案例。

由于政府绩效评估既是公共管理学科的理论前沿也是实践性较强的课题,因此,本书的写作也必然是以理论联系实际的方法展开,这就使得本书既可以对从事政府绩效评估教学与理论研究的人员具有一定的参考价值,也对实际工作部门人士了解美国政府绩效评估的做法以及思考中国政府借鉴美国经验的问题有所帮助。

由于政府绩效评估是一个相当艰涩的研究课题,加之作者水平有限,因此,书中错误在所难免,请读者不吝赐教!

孟 华

2005年11月

1

导 论

在当今世界,政府机构对自身的组织绩效实施评估与测量是政府管理领域的热门话题,西方各国都以特定的方式推进着这一实践。美国政府组织绩效评估^①最早可以上溯至 20 世纪二三十年代,当时美国地方政府就开始了对好政府的追求并且以效率为核心对政府的运作进行测量与评估^②,后来特别是 50 年代绩效预算实施后,这一实践更是不断地发展。目前,设立绩效指标与标准、测量与评估政府绩效成为吸引美国实践界与学术界的显学。

对于美国政府绩效评估,有诸多问题值得我们深思:为什么早在 20 世纪早期美国政府就开始对政府绩效评估感兴趣?为什么这一实践发展到今天却无法超过绩效评估的后起之秀英国?诸如此类的问题都具有足够的吸引力,并引导我们对美国政府绩效评估进行深入的研究。

1.1 政府绩效评估的核心概念

在研究美国政府绩效评估之前,必须预先界定一些基本的概念,其中最为核心的就是绩效与绩效评估。

到目前为止,一个得到大家公认的绩效概念并不存在,这一方面是因为公共机

^① 绩效评估可分为组织绩效评估与个人绩效评估,本书只讨论政府组织绩效评估而不涉及政府中个人绩效评估问题即公务员绩效考核问题。选择这种研究视角主要是因为,我国政府部门中的绩效评估在个人绩效评估方面已经有着长时间的经验积累,而组织绩效评估却只是处于起步阶段,因此,组织绩效评估的研究也就更具有相应的迫切性。本书以下部分在单独只使用“绩效评估”一词时也均指组织绩效评估。

^② 请参阅尼古拉斯·亨利:《公共行政与公共事务》(第七版),华夏出版社 2002 年版,第 166 页。

构缺乏像私营部门的利润那样具有普适性的评估尺度,不同的利益相关者对于成就的认识缺乏一致性^①,另一方面是因为绩效概念还会随着研究与实践的发展而在内涵与外延上发生一定的变化。正是基于这些原因,确定一个适当的绩效概念成为推动人们研究公共部门绩效评估的三大因素之一。^②

从当前学界对绩效的研究来看,大致有四种绩效概念的界定方式。第一种以韦氏为代表,强调绩效所蕴含的结果(results)层面的意义。根据韦氏词典,绩效是完成某种任务或达到某个目标。^③美国的《政府绩效与结果法案》(Government Performance and Results Act, GPRA)直接将绩效所体现的结果(results)归结为产出(output)与结果(outcome)^④两个层面。经济合作与发展组织(OECD)在1997年有关结果(results)导向的绩效管理实践的报告中,根据多国的实践将绩效所体现的结果(results)层面的意义界定为几种不同的表达方式,认为英文 results 所代表的结果可以包含产出、由 outcome 指代的结果以及服务质量(service quality)、顾客满意度(customer satisfaction)。^⑤由于服务质量和顾客满意度最终也可以归结为产出与结果(outcome)^⑥,因此,这一观点与《政府绩效与结果法案》的界定基本上是一致的。

第二种则是从过程的角度界定绩效。如科内(Kearney)和伯曼(Berman)认为,绩效是为实现结果(results)而管理公共项目,并指出,“公共绩效的一个显著特点就是它由效益、效率和公正的多个同等重要的标准引导和评估”。^⑦

① See AL GORE. Serving the American Public; Best Practices in Performance Measurement. <http://govinfo.library.unt.edu/npr/library/review.html>. 1997-6.

② JOS L. T. BLANK ed. Public Provision and Performance; Contributions from Efficiency and Productivity Measurement. Amsterdam, the Netherlands; Elsevier Science B. V., 2000. 3.

③ 转引自汪家常、魏立江:《业绩管理》,东北财经大学出版社2001年版,第11页。

④ 这里的结果(outcome)包括预期与非预期的结果(outcome)以及中期与最终的结果(outcome)。

⑤ OECD. In Search of Results; Performance Management Practices. Paris; OECD Publications. 1997. 18.

⑥ 根据美国会计准则委员会(GASB)的服务提供与成就(Service Effort and Accomplishment, SEA)报告系统以及国家公共生产力中心(NCPP)的理解,服务质量和顾客满意度其实也可以通过产出或结果(outcome)表达出来。在SEA中,服务质量可以部分地通过达到某种质量要求的产出的数量(产出)来测量,也可以通过顾客对服务质量的感知的调查[结果(outcome)]来测量,顾客满意度本身就是一个结果(outcome)指标。NCPP则直接将质量归入结果(outcome)。请参阅 GASB. Concept Statement No. 2 of the Governmental Accounting Standards Board; Service Efforts and Accomplishments Reporting. <http://www.seagov.org>. 2002-10-15 访问,以及 National Center for Public Productivity. A Brief Guide for Performance measurement in Local Government. <http://www.rugers.edu/~ncpp>. 1997.

⑦ RICHARD C. KEARNEY, EVAN M. BERMAN ed. Public Sector Performance; Management, Motivation, and Measurement. Boulder, Colorado; Westview Press, 1999. 1—2.

第三种以经济合作与发展组织为代表,主要是从绩效所反映出的能力的角度来认识绩效。如在 2000 年经济合作与发展组织给出的定义中,绩效被视为资源获取与使用上的能力,是“一个机构或当局经济地获取资源以及高效率(投入-产出)、高效益[产出-结果(outcome)]的利用那些资源实现绩效目标的熟练性”。^①

第四种是通过列举绩效内容给出绩效的概念。如格里泽(Grizzle)认为绩效是一个多维度的概念,它包括效率(将成本与直接产出相关联),成本-效益(将成本与收益或影响相关联),服务提供的质量与公正性、政府财政稳定性和政府政策的一致性。^②臧乃康则将政府绩效分解为政绩、政府成本、政府效率、政治稳定、社会进步、发展预期等。他同时指出,从框架上看,政府绩效的主要部分是经济绩效、社会绩效和政治绩效。^③布朗克(Blank)和洛弗尔(Lovell)更深入一步,他们从测量与评估的角度出发,指出绩效具有相对性,公共部门的绩效通常是相对于其他组织或其他时间点等的效率、效益和生产力变化。^④

从学者们的概念界定中,我们可以看出,人们对绩效的认识其实并不存在矛盾之处,只是看待绩效的视角不同而已。大致而言,绩效是一个机构或组织的相关活动或项目的投入、产出和结果(outcome),它强调产出与结果(outcome),并表现为能够反映组织所具有的特定能力的效率、效益、公正、质量等。评估其实就是对这一能力的认识过程。^⑤对于政府来说,政府绩效又包括“政府的项目、服务与操作绩效”。^⑥其中,项目一般包括政府直接或间接向社会提供的项目以及管制项目(regulatory program)^⑦,操作指的是政府的日常运作,而服务既可以是机构职能操作的过程与结果,也可以

① See OECD. Working Definitions. <http://www.oecd.org>. 2000.

② GLORIA A. GRIZZLE. Measuring State and Local Government Performance: Issues to Resolve Before Implementing a Performance Measurement System. in RICHARD C. KEARNEY, EVAN M. BERMAN. Public Sector Performance: Management, Motivation, and Measurement. Boulder, Colorado: Westview Press, 1999. 329—340.

③ 臧乃康:“政府绩效的复合概念与评估机制”,《南通师范学院学报》2001年第3期,第25—29页。

④ JOS L. T. BLANK. Public Provision and Performance: Contributions from Efficiency and Productivity Measurement. Amsterdam, the Netherlands: Elsevier Science B. V., 2000. 3—11.

⑤ 当然,由于价值考虑的干预,认知绩效所反映出的组织能力是一个经过放大或缩小处理后的能力。

⑥ See JOHN MAYNE, EDUARDO ZAPICO-GOÑI ed. Monitoring Performance in the Public Sector: Future Directions from International Experience. New Brunswick, New Jersey: Transaction Publishers, 1997. 161.

⑦ 对项目的这种分类参阅 Harry P. Hatry etc. How Federal Programs Use Outcome Information: Opportunities for Federal managers. <http://www.businessofgovernment.org>. 2003-5 访问。在此我们可以看出,项目、操作与服务三个概念有着交叉之处。因此,总审计署干脆将三者统称为项目,这在后面将会涉及。

是项目本身的目的所在。

确定“评估”的内涵与外延是定义绩效评估的关键一步,然而,在英文中,measurement、evaluation、review、assessment、appraisal 等都被人们用来指称对于绩效的评估,我国学者在翻译时,倾向于将它们都译为“评估”。那么,它们到底是一种怎样的关系呢?根据笔者收集的有关政府绩效评估的英文文献,人们经常用到的几个指代组织绩效评估的词语是 measurement、evaluation 和 assessment。同时,就美国历史上人们对前两个词的使用情况来看,它们具有相对的阶段性和限定性,主要表现为人们主要将 evaluation 用于 program/policy evaluation^①,该词组出现频率最高的时期是在林登·约翰逊的“伟大社会”政策实施之后的 20 世纪 60 年代中期至 70 年代;而 measurement 经常与 performance 相联,并且自 20 世纪 80 年代末^②以来充斥在美国学者和实践家的口头与文字中。对于它们之间的关系,尤其是 measurement 与 evaluation 之间的关系,人们有着多样化的认识。

首先,performance assessment 包含 performance measurement 与 program evaluation。这种认识的典型代表是总审计署(GAO)。该机构将 performance measurement 与 program evaluation 合称为 program performance assessment。其中,“项目”(program)可以是“任何具有可识别的目的或一系列目标的活动、计划(project)、职能或政策”。^③因此,这里的“项目”比我们前面所使用的项目概念在范围上要宽泛得多,包括了政府绩效所涉及的项目、服务与操作等多个方面。

其次,performance measurement 是 program evaluation 的一个组成部分。这种看法的典型代表是彼得·罗西等,他们认为 performance measurement 是 program evaluation 中的过程评估或者说项目监测的一种形式,而且,与其他项目监测方式相比,它特别强调对服务结果(results)的评估。^④

① 虽然有些学者在使用 evaluation 时经常将项目与政策区分开来,但是,在大多数情况下它们两者所指代的评估或评价方法与内容是相似的,而且,项目通常就是政策的直接产物。本书中的 program evaluation 就包括对两者的评价或评估。这种研究方式也是较为常见的,如戴伊就采用这一研究方式。请参阅托马斯·R.戴伊:《自上而下的政策制定》,中国人民大学出版社 2002 年版,第 210 页。

② See ROBERT KRAVCHUK, RONALD W. SCHACK. Designing Effective Performance Measurement Systems under the Government Performance and Results Act of 1993. Public Administration Review 56, 1996, (4): 348—359.

③ UNITED STATES GENERAL ACCOUNTING OFFICE. Performance Measurement and Evaluation: Definitions and Relationships. GAO/GGD-98-26, April 1998. 3.

④ 彼得·罗西,霍华德·弗里曼和马克·李普希:《项目评估:方法与技术(第六版)》,华夏出版社 2002 年版,第 148—151 页。

再次, performance measurement 包含 program evaluation。这一观点的代表人物是科内和伯曼,他们在其主编的一本有关公共部门绩效问题的书中,将 program evaluation 与 performance measurement 作为同义语看待,将涉及到 program evaluation 与 performance measurement 的文章统一编入“measurement”这一部分。^①另外,格里泽在讨论 performance measurement 时也将对绩效的 measurement 锁定在对政府项目的 evaluation 上。^②

国内学者在研究绩效评估时,通常采用第三种认识,并且将政府绩效评估译作 government performance measurement。但是,由于美国政府绩效评估的实践更多地受到了总审计署给出的概念的限定,而非理论界研究成果的约束,所以,笔者倾向于接受总审计署的观点,并主张将 assessment 译作评估,而将 measurement 和 evaluation 分别译作测量和评价。^③这样,绩效评估就是由绩效测量与项目评价两部分组成的,当然,此处的项目评价与社会学和政策科学领域的项目(政策)评估虽然有相通之处,但是,根据总审计署的界定,本书所研究的项目评价包括过程(或执行)评估、结果(outcome)评估、影响力评估以及成本-收益和成本-效益评估^④,而不包括社会学和政策科学领域对项目(政策)在价值层次的评估,即无需实施需求评估、可评估性评估以及项目理论评估等。

现在的问题是要确定绩效测量与项目评价所指代的事物。而有关绩效测量与项目评价的概念解释也是多种多样的,我们可以将它们分开来梳理。

对于绩效测量,第一种界定方式主要是从绩效测量的不同内容入手给出绩效测量的概念。在这种界定方式下,人们表现出的认识差异是与对绩效内容的不同认识直接相关的。其中,涉及内容较为全面的绩效测量概念是由格里泽给出的。

-
- ① RICHARD C. KEARNEY, EVAN M. BERMAN ed. *Public Sector Performance: Management, Motivation, and Measurement*. Boulder, Colorado: Westview Press, 1999. 263.
- ② GLORIA A. GRIZZLE. *Measuring State and Local Government Performance: Issues to Resolve Before Implementing a Performance Measurement System*. in RICHARD C. KEARNEY, EVAN M. BERMAN. *Public Sector Performance: Management, Motivation, and Measurement*. Boulder, Colorado: Westview Press, 1999. 329—340.
- ③ 一些著名的美国学者也接受了总审计署的这种概念界定。如尼古拉斯·亨利在其第八版的《公共行政与公共事务》、格罗弗·斯塔林在其第六版的《公共部门管理》中都采用了这一概念界定方式。而且,译者在翻译斯塔林的著作时也明确地将 assessment 译作评估,而将 measurement 译作测量。请分别参阅尼古拉斯·亨利:《公共行政与公共事务(第八版)》,中国人民大学出版社2002年版,第283—287页,以及格罗弗·斯塔林:《公共部门管理》,上海译文出版社2003年版,第201—203页。
- ④ UNITED STATES GENERAL ACCOUNTING OFFICE. *Performance Measurement and Evaluation: Definitions and Relationships*. GAO/GGD-98-26, April 1998. 5.

她认为综合性的绩效测量体系可以提供有关一个政府或单个的政府机构在效率、成本-效益、服务提供的质量、服务提供的公正性、政府财政稳定性和与政府政策的一致性这些绩效内容方面的运作情况的信息。^①经济合作与发展组织^②、GASB的服务提供与成就报告系统^③等则强调绩效测量是对投入、产出和结果(outcome)的测量活动。戈尔领导下的国家绩效评审委员会(NPR)关注效率、质量和效益,认为绩效测量提供以下信息:资源被转化为产品和服务(产出)的效率,产出的质量(它们被提供给客户的质量以及客户满意的程度如何),结果(outcome)的质量[与预期目的相比,一个项目活动的结果(results)],以及政府的操作在其对项目目标的特定贡献方面的效益。^④总审计署1980年给出的绩效测量概念与此类同。^⑤也有人从绩效测量的核心测量内容入手给出了绩效测量的简单定义。如哈特里(Hatry)的概念关注效率和效益^⑥,而爱泼斯坦(Epstein)则重视效益^⑦。虽然由于人们对绩效内容的认识不统一导致了绩效测量内容认识的差异,但是,绩效测量的传统内容——三E(经济、效率、效益)却得到了绝大多数学者的认可。当然,还是有些学者认为,只测量三E是不够的,并为其加上了公平(equity)、企业家精神(entrepreneurship)、卓越(excellence)、功效(efficacy)和伦理(ethics)等。^⑧

第二种是从绩效测量的过程出发,认为“应该把绩效评估^⑨作为一个衡量程序的过程或系统来看待”^⑩,通过对测量过程所涉及的活动与因素进行描述和概括给出绩效测量的定义。如彼得·罗西、霍华德·弗里曼和马克·李普希认为“绩效测

① GLORIA A. GRIZZLE. Measuring State and Local Government Performance: Issues to Resolve Before Implementing a Performance Measurement System. in RICHARD C. KEARNEY, EVAN M. BERMAN. Public Sector Performance: Management, Motivation, and Measurement. Boulder, Colorado: Westview Press, 1999. 329—340.

② See OECD. Working Definitions. <http://www.oecd.org>. 2000.

③⑤⑦ NATIONAL CENTER FOR PUBLIC PRODUCTIVITY. A Brief Guide for Performance Measurement in Local Government. <http://www.rutgers.edu/~ncpp>. 1997.

④ AL GORE. Serving the American Public: Best Practices in Performance Measurement. <http://govinfo.library.unt.edu/npr/library/review.html>. 1997.

⑥ HARRY P. HATRY. Performance Measurement Principles and Techniques: An Overview for Local Government. in Richard C. Kearney, Evan M. Berman. Public Sector Performance: Management, Motivation, and Measurement. Boulder, Colorado: Westview Press, 1999. 304.

⑧ See JOHN MAYNE, EDUARDO ZAPICO-GOÑI. Monitoring Performance in the Public Sector: Future Directions from International Experience. New Brunswick, New Jersey: Transaction Publishers, 1997. 5.

⑨ 此处译者将“measurement”译作“评估”。

⑩ 马克·霍哲:“公共部门业绩评估与改善”《中国行政管理》,2000年第3期,第36—40页。

量是绩效指标的搜集、报告和解释,这些指标与项目执行得如何相关,特别是与服务(产品)送达和最终达到的结果(outcome)相关”。^①与他们的看法相似的是费舍的观点,他认为绩效测量是发现问题是什么的过程,它需要确定并收集绩效标准方面的资料,这些标准能够说明一个项目或一项服务的绩效。^②

第三种是从绩效测量的特点入手界定绩效测量。如总审计署强调绩效测量的“进行中的”(ongoing)特性,它在1998年将绩效测量界定为“对项目成就尤其是趋向于既定目标的进展情况的进行中的监控与报告”,并指出它通常由机构或项目管理层实施。^③国家绩效评审委员会给出的绩效测量概念与此类似。^④

上述概念正好从不同的角度刻画出了绩效测量的全貌。概括起来,所谓绩效测量就是根据既定的组织目标在政府活动的投入、产出、结果(outcome)以及由它们反映出的经济、效率、效益、质量等内容上设置绩效指标和标准,并在日常的管理工作中围绕绩效指标不断地收集有关政府活动趋向于既定目标的进展情况的信息,从而通过与绩效标准的比较随时确定组织的绩效表现情况。通常情况下,绩效测量只关注对有限的一些指标(即核心指标,如《政府绩效与结果法案》只要求联邦政府机构设计产出与结果(outcome)指标)的信息的收集与测量,关注政府的活动或项目是否实现了目标或达到了可测量的绩效标准。指标的限定性以及将进展情况与标准比较的约束,决定了绩效测量并不关心指标以外信息或者与标准无关的信息的收集。

与绩效测量不同的是,项目评价首先是一种系统的评估,是“针对一个项目或政策以及项目或政策的价值的重要方面的系统解析式评估,也是寻求研究结果的可靠性与可用性的系统解析式评估”^⑤。它不像绩效测量那样局限于进展情况与标准或目标的比较,而通常要对有关项目绩效与其背景中更为广泛的信息进行检查,因此,项目评价可以检查项目操作问题、阻碍或促进项目成功的环境因素,可以评价项目产生的超出预期目标的影响,可以评价项目的净影响,也可以比较趋向于

① 彼得·罗西、霍华德·弗里曼和马克·李普希:《项目评估:方法与技术(第六版)》,华夏出版社2002年版,第326页。

② RICHARD J. FISCHER. An Overview of Performance Measurement. *Public Management* 76, 1994, (9): S2—S8.

③ UNITED STATES GENERAL ACCOUNTING OFFICE. Performance Measurement and Evaluation: Definitions and Relationships. GAO/GGD-98-26, April 1998. 3.

④ AL GORE. Serving the American Public: Best Practices in Performance Measurement. <http://govinfo.library.unt.edu/npr/library/review.html>. June 1997.

⑤ See OECD. Improving Evaluation Practices. Paris: PUMA/PAC(99)1, 1999.

同一目标的不同项目活动的效益。其次,项目评价也不是随时进行的,而是为了评估一个项目的进展情况而定期或专门进行的、单独的系统性研究,并且,它通常由项目外部的(可以是机构内部的,也可以是机构外部的)专家实施,当然,也可以由项目管理者实施。因此,绩效测量可以作为一种管理的预警系统,而项目评价则可以通过对项目的作用进行全面评估,帮助组织找到改善项目结果(results)的调整措施。^①

现在我们可以考虑将先前对绩效的界定、对项目评价与绩效测量关系的考虑以及对绩效测量和项目评价的概念理解综合起来,从整体上来理解政府绩效评估。简单地说,所谓政府绩效评估是一些特定的评估活动的总称,它既包括将政府的项目或活动的成就与预期目标或绩效标准相比较的日常测量活动,也包括对项目或政策的重要方面进行客观可信的定期的系统评价活动。具体而言,绩效评估的对象应该是暗示着一定的组织能力的投入、产出与结果(outcome);绩效评估既包括政府活动进行中的日常测量,也包括定期的或特别举行的评价;绩效评估既强调项目、服务与操作的进展情况和既定目标的比较,也关注对项目所发挥作用的全面评价;由于绩效是一个相对的概念,绩效评估就要涉及一定的指标与标准的设置与选择。

具体到美国政府的绩效评估,需要特别注意的是它的评估焦点,这是将其与其他发达国家区分开来的一个重要因素。从本质上讲,美国政府的绩效评估是对投入与过程控制的反叛,它的关注焦点在于产出和结果(outcome)。当然,从历史发展的情况来看,产出(以及用投入/产出表达的效率)测量与评估已经有很长的历史,但是,结果(outcome)测量与评估则主要是 90 年代以来绩效评估重点发展的内容。当前美国各级政府都力图在原有的产出测量之外,更多地增加结果(outcome)测量的成分^②。关于这一点,经济合作与发展组织 1997 年的报告中已经给予了明确的表达。该报告指出,与其他国家相比,美国政府的绩效评估非常明显地表现为同时重视产出和结果(outcome),这与美国 90 年代推行的结果(results)管理是直接相关的(具体比较请参阅附表 1)。

另外,在本书的行文中,政府一词仅指行政系统,在美国,它所指代的范围可上至联邦政府,下至州、郡、市和县政府。政府绩效评估虽然可以由两类不同的力量来实施,即由政府自己实施和由行政部门以外的机构(如议会)、学者或中介组织等

① See UNITED STATES GENERAL ACCOUNTING OFFICE. Performance Measurement and Evaluation: Definitions and Relationships. GAO/GGD-98-26, April 1998. 3—4.

② 纽科默、詹宇斯等在对各州将绩效信息引入预算的努力进行研究时也得出这一结论,请参阅凯瑟琳·纽科默,爱德华·詹宇斯,谢里尔·布鲁姆和艾伦·洛马克斯:《迎接业绩导向型政府的挑战》,中山大学出版社 2003 年版,第 81 页。

实施的对政府绩效的评估,但是,本书的研究视角将限定在美国政府作为评估者的绩效评估活动上。在这种情况下,仍然可以再细分为两种政府绩效评估的形式,一种是上对下的评估或者由其他行政机构实施的评估,如管理与预算局对各联邦机构实施的评估;另一种是机构自己对自己实施的绩效评估。在美国,虽然两种形式都存在,但是,在更多的情况下,第二种形式更为普遍,第一种形式通常是以第二种形式的监督机制出现的。当然,我们还应该注意的是,任何评估实践都不可能在一个封闭的系统中运行。在美国政府的绩效评估中,政府部门所起的作用最为重要或者说是主导性作用,但是,这并不能排除其他的一些力量(民选官员、公众等)通过一定的方式对这一过程施加影响。

1.2 政府绩效评估的基础理论

绩效测量与项目评价虽然都是对政府操作与项目的评估,但它们除了概念的差异之外,在操作过程、表现形式以及技术手段等各方面都有着明显的不同。

1.2.1 绩效测量

虽然绩效测量自 20 世纪 80 年代以来才日益兴盛起来,但是,学者们对于绩效测量的讨论仍然建构出了一个相对完整的绩效测量理论框架。

绩效测量是一个过程,它包括设置绩效目标、开发绩效指标、收集绩效信息、分析绩效信息以及汇报绩效结果等几个环节,它与绩效管理不同仅在于,绩效管理还必须将对组织使命、战略的考虑以及如何使用绩效结果信息改善组织绩效包括在内。因此,绩效测量是绩效管理的重要组成部分。

实施绩效测量的第一步是设置绩效目标(或称之为绩效标准),无可辩驳的是,绩效目标的设置不是任意的,它与组织的战略直接相关,也就是说它是对长期的组织战略的短期化表达。绩效目标的设置首先需要对组织的战略进行分析,从而确定组织工作的核心,识别组织要测量的核心活动。在确定了绩效测量的核心活动后,就可以根据机构过去的绩效、组织内类似部门的绩效、其他地区类似机构的绩效、专业化标准或全国的标准等来确定绩效目标的水平。要设置好绩效目标必须遵循以下原则:绩效目标必须与机构所提供的服务以及工作目标相关;绩效目标具有可实现性但又要对机构形成压力;绩效目标的表达必须清晰易懂;绩效目标必须

具有可测量性等。在这些原则的指导下,绩效目标又可以划分为多种类型。第一种就是始终如一的目标(all-the-time targets),它承诺向公众自始至终提供达到某种质量水平的服务或产品,如“每次您联系我们时,您都将得到礼貌、高效和专业的服务”;第二种是百分比目标(percentage achievement targets),即承诺达到某百分比水平的服务绩效,如“90%的电话在响后三声之内就要予以接听”;质量目标(qualitative targets)是对人们所接受的服务作出描述性的承诺,如“我们将会以热情、及时的服务接待任何一位到咨询处询问有关信息的旅客”;时限目标(time-bound targets),即对服务提供给出明确的时间范围,如“自房屋维修的申请提出后,1 000 元以下的维修将在两日内完成”;全国性目标(national targets),它的最常见的形式是依据全国性法律或政策的规定设定服务的要求,如“居民身份证的办理期限最长不能超过法定时限”等等。^①当然,有时同一个目标可以同时归属于不同的类别,如“每次您联系我们时,您都将得到礼貌、高效和专业的服务”就既是一个始终如一目标也是一个质量目标。

10 在设置了绩效目标之后,就可以比较容易地开发出绩效指标。通常情况下,绩效目标与绩效指标是一一对应的。上面所设置的百分比目标“90%的电话在响后三声之内就要予以接听”就可以产生这样的绩效指标:电话响后三声内予以接听的百分比。这一指标进而可用公式表达为:(电话响后三声内予以接听的数量÷全部电话的数量)×100%。通常情况下,指标也可以划分为几种不同的类型:投入指标、产出/工作量指标、结果(outcome)/效益指标、效率指标和生产力指标。一般情况下,投入指标涉及机构运作或项目耗费的资源;产出指标汇报一个项目生产的产品或提供的服务,而工作量指标表明完成的工作数量或接受的服务数量;结果(outcome)/效益指标报告服务的结果(results)(包括质量),其中,效益指标指服务回应公众需求的程度;效率指标检查政府的工作情况如何,特别是所提供的服务的数量(产出)与成本(投入)的比率,与结果(outcome)/效益指标不同的是,效率指标不考虑政府所做的工作是否是它应该做的;生产力指标是效率与效益指标的综合,如将“一小时维护的路面里程数”(效率指标)与“经过维护的路面里程中6个月内无需返工的百分比”(效益或结果(outcome)指标)相结合,而形成“经维护而在6个月内无需返工的每一公里路面的单位成本或劳动时间”(生产力指标)。有时,有些指标的信息是难以获得的,如一些结果(outcome)或质量指标,这时可以用其他的

^① 这是英国审计署倡导的绩效目标划分方式,请参阅 AUDIT COMMISSION of U. K. Setting And Monitoring Local Performance Targets. London: Audit Commission. 1999. 20—22.

指标来替代。以顾客满意度为例。如果直接询问“你对某机构的服务是否满意？”所得到的回答肯定会因为个人感知与评价标准差异而非常主观,这时可以用顾客的投诉率或者顾客的回头率来替代对顾客直接询问满意情况。我们可以以公共卫生部门为例具体说明这些指标的设计情况。(见表 1-1)

表 1-1 绩效测量指标的类型示例

职 能 指 标	公 共 卫 生
投入指标	公共卫生部门预算;交通工具数量
产出/工作量指标	收集的垃圾的吨数;清扫的路面的里程数;服务的顾客的人数
效率指标	每收集一吨垃圾占用的雇员时间;清运一公里路面积雪花费的资金
效益指标	干净路面的百分比(通过定期视察或市民调查得到)
生产率指标	每一公里干净路面所耗费的成本(干净路面的总成本/干净路面的总里程)

资料来源:National Center for Public Productivity. A Brief Guide for Performance Measurement in Local Government. <http://www.rutgers.edu/~ncpp>. 1997。

当然,绩效指标的选择与设计也必须遵循一些必要的原则与要求,通常情况下,有效性/准确性、可理解性、时效性、全面性以及独特性等都会对绩效指标的选择产生重要的指导作用。其中,准确性要求评估指标能够有效地反映它要评估的内容;可理解性则是为了让政府官员或民众能够理解指标所要展示的信息;时效性强调评估指标所要展示的信息能够及时地收集到,如一些评估长期影响的指标就可能一时间无法得到相关信息,从而因缺乏时效性而无法正常使用;全面性则要求指标体系能够覆盖一个机构绩效的大部分或全部;独特性要求指标之间不重复。另外,绩效评估指标是否会鼓励一些负面行为的发生以及指标所涉及的资料收集的成本等都是机构在选择指标时不可回避的问题。^①

需要注意的是,政府部门的差异性是相当明显的,这会对绩效指标的选择产生直接的影响。威尔逊曾经从管理学的视角提出了机构产生差异性的两个主要方面,即“操作者的活动是否能够被观察到?那些活动的结果能否被观察到?”其中第一个方面与对产出的测量有关,而第二个方面与对结果(outcomes)的测量有关。

^① RICHARD C. KEARNEY, EVAN M. BERMAN ed. Public Sector Performance: Management, Motivation, and Measurement. Boulder, Colorado: Westview Press, 1999. 305.

由此可以将机构划分为四种不同的类型：产品型(production)组织即机构有可见的产出与结果(outcomes)；程序型(procedural)组织即产出可见但结果(outcomes)不可见；技巧型(craft)组织即产出不可见但结果(outcomes)可见；顶盖型(coping)组织即产出与结果(outcomes)均不可见。^①机构这种差异性的客观存在意味着，组织在设计绩效指标时情况不一，技巧型组织很难设计出产出指标，而程序型和顶盖型组织则很难开发出结果(outcomes)指标。

12 绩效目标与指标设计完成之后，还必须为绩效目标的实现确定责任人，即每一项绩效目标都要有专门的人员负担其是否实现的责任。通常情况下，目标实现责任可以借助于一些责任工具落实到人。威尔·阿特雷(Will Artley)介绍了一些常用的工具：绩效协议、责任报告、自我评估、绩效回顾、绩效会议等。其中，绩效协议是与管理层或那些实际参与服务提供活动的人员所签订的用于测量绩效的一种协议，协议通常明确指出对协议签定各方的期望，确定协议人职责。责任报告是定期(通常为一年)出版的有关绩效责任履行情况的报告，它通常包括有关项目或财务的一些信息，能够帮助人们了解绩效目标的实现情况。自我评估是一个不断进行的过程，它帮助组织不断监测自身的绩效并评价自身实现绩效目标的能力。绩效回顾也是不断进行的过程，它通过将一特定时期内的事实绩效与计划绩效进行比较找出未来的发展方向。绩效会议的目的不在于评估绩效，而是要通过会议的形式理解当前的缺陷、吸取以前的教训，从而改善未来的绩效。^②

绩效资料的收集是资料分析的基础。为了更好地收集绩效资料，通常需要预先构建完善的资料收集方案。这一方案的第一步就是确定所需收集的信息。这些信息的确定与绩效指标有着直接的关系，一般可以通过对指标的分析确定每个指标所需要的绩效信息。如前面所设计的“电话响后三声内予以接听的百分比”这一指标就需要收集“电话响后三声内予以接听的电话数量”和“全部电话的数量”这两方面的资料。第二步就是要确定这些信息出自于什么地方，可以由什么人负责收集。电话接听服务的信息来源自然是机构中的“热线中心”或者“电话接收中心”等部门，从电话接听记录中应该可以得到这两方面的信息，如果有网络进行记录，信息的质量无疑会更高。当然，组织可能需要对热线中心等的工作记录作一些相应的调整以确保这一绩效指标所需的信息能够被准确记录下来。接下来就是进行资料的收集与汇报。在这一过程中需要注意的是，机构必须时刻关注资料的有效性或

① J. WILSON. Bureacracy. New York: Basic Books. 1989. 158—159.

② WILL ARTLEY. Establishing Accountability for Performance. <http://www.orau.gov/pdm>. 2001-9.