Managerial Challenges of Performance Management: 
A Cross-national Exploration*

Dennis C. Smith
Robert F Wagner Graduate School of Public Service
New York University

Introduction

In the latest reinvention of American government, a central theme has been the idea of "managing for results." One of the requirements of the Government Performance and Results Act was that all federal agencies defined their objectives and develop measures that they will regular report. At the same time a variety of forces, especially the call for accountability and the emergence of performance-based contracting, have converged on the nonprofit sector in United States requiring the development of outcome measures. The starting point of this paper is that so much attention and energy has gone into the task on defining and measuring results or outcomes, the comparatively little attention or energy has been devoted to the managerial challenge of using outcome measures to improve results.

In England during 1990s the public sector was similarly engaged in a process of outcome measurement development, which began under Prime Minister Thatcher in the form of an emphasis on "value for money" and continued in the Blair governments own reinvention efforts. As a consequence there is considerable literature on performance measurement in the public sector in England. A review of that literature shows that, as in United States, managing results poses bigger challenges then measuring them.

In Latin America, the government of Chile, among others have also undertaken significant performance measurement and management initiatives. These are now being documented and their story will become part of the available history of this movement.

New York City began to measure the work of its government agencies in the 1970s in the wake of fiscal crisis. The 1990 study of the Mayors Management Planning Reporting System showed that pervasively city agencies were announcing performance targets and reporting their achievement or non-achievement, annually, but not using these reported measures in agency management. Beginning in 1994, with the New York City Police Department and later Department of Corrections, and Parks Department, began to use outcome measures as an integral part of management. Some structural changes were also involved, but a key feature of the reform was the dramatic change in the use of performance data in the process of management. For each of these agencies the performance outcomes were truly remarkable: crime in all categories and in all geographic districts was dramatically reduced, safety into jails was dramatically improved, and the safety and cleanliness of parks were very substantially increased. (See Charts 1-3). There are lessons to be learned from these experiences.

One consequence of the shrinkage of the world produced by the Internet and the World Wide Web is that developments in public administration in one part of the world are observed and potentially copied in others very quickly. In principle, the speed of travel on the information highway greatly accelerates the pace of the diffusion of innovation. Using the case of

* Preliminary Draft. For presentation of the meeting of the International Congress of CLAD on State and Administrative Reform October 24-27, 2000 Santo Domingo, Dominican Republic

NB: Not for publication or quotation without permission of the author
performance measurement and management, the major focus of this paper is to explore how political and cultural differences affect the process of technological borrowing in the field of management reform. By comparing the experience of the introduction of performance management in a number of agencies in a number of fields of public service in several countries, this paper aims to identify common and divergent patterns of managerial adaptation out of reform ideas.

Some of the major performance management ideas have been part of a discussion among a group of national leaders, notably U.S. President Bill Clinton and the Prime Minister of the United Kingdom Tony Blair, and including a variety of European leaders, in meetings hosted by New York University Law School, on the so-called "Third Way." In these discussions it is assumed that reinventing government means improving performance through clearer articulation of public sector goals, and objectives and logic for achieving them, better measurement of their achievement, and greater accountability of public officials charged with the achievement of those goals.

Not surprisingly, some of this rhetoric is now appearing in international fora dialogs, for example, at the United Nations, and in the call by the president of the World Health Organization for greater reliance on "evidence-based medicine."

However, in my informal survey of government practices in European countries as part of Wagner School course on Public Policy Analysis in Europe taught each summer in Lyon France and Brussels, it seems quite clear that in some public administrations, such as in France, the ideas of public sector performance measurement and management are not developing in the same way as for example in the United Kingdom and in the Netherlands.

The purpose of this paper is to provoke a discussion of whether the ideas of public sector performance management is built upon an infrastructure of institutional and cultural ideas, and is in fact dependent upon them for its operational success. A series of questions need to be raised:

1. Are the ideas is of causation underlying" theory based evaluation and performance management" specific to Western intellectual traditions and assumptions?
2. Are the concepts of evidence embedded in the key concepts of performance measurement in the public sector culturally or institutionally determined in ways that make their use variable across societies?
3. Are the ideas of public sector accountability in batted in performance management culturally or institutionally determined in ways that make their use variable across societies?

The lesson of the comparative public administration of the 1960s is that these questions tend not to be asked, and that failure to ask them runs the risk of generating patterns of diffusion of policy and management practices that result in substantial and not always positive, unintended consequences.

From Performance Measurement to Performance Management : the New York Experience

It therefore seems to me that analysts who want to help improve social service delivery should give high priority to developing and refining measures of performance. Relatively little effort has gone into devising such measures so far, despite their importance and the apparent intellectual challenge of the task... Performance measures for social service are not, of course, ends in themselves. They are prerequisites to
attempts both to find more effective methods of delivering social services and to construct incentives that will encourage their use.

Alice Rivlin, Systematic Thinking for Social Action (1973)

For more than two decades, the City of New York has utilized a planning and management tool called Mayor’s Management Plan and Report System (MMPRS).

The MMPRS is carried out by individual agencies under the direction of a designated Management Plan Coordinator, and monitored and administered by the Mayor’s Office of Operations. The MMPRS allows for development of an annual agency plan, including a review of agency mission and programs which define the job of the agency, performance plans which determine how well and how much of the job is to be done, and planned improvement projects which detail efforts being taken to upgrade service delivery or operations management.1

In fact, as several studies documented, for much of its history very little of the system’s(MMPRS) potential was realized. Colloquially the system is referred to as the “MMR” because there was little that was operational city-wide except the Mayor’s Management Report that is issued twice a year. The city government, which had pioneered the development of performance measurement in large cities, with rare exceptions, was not using those measures in any systematic way to manage performance.2 The important lesson of this experience is the following: Without an explicit commitment to make it happen, performance measurement, while arguably useful in its own right, does not automatically produce systematic performance management. From a distance, the enormous effort in the Federal government to implement the Government Performance and Results Act (GPRA) by 1999 seems to be following the New York experience of focussing on measurement rather than management. In Donald Kettl’s testimony to the Congress about the value of performance measurement in reducing the “performance deficit” and the importance of Congress using the measures produced, there is no discussion of Congress insisting that federal agencies use their measures in managing. 3

The introduction of a new management approach, widely known as COMPSTAT, in the New York City Police Department (NYPD) in the mid-1990s, and its dramatic success in bringing down crime and increasing citizens’ perceptions of public safety, led the way to a new emphasis on performance management. Smith (1997) reported in “What public managers can learn from police reform in New York City: COMPSTAT and the promise of performance management” that other city agency leaders in New York were adapting the approach developed by the NYPD and using it in managing their performance with notable success. Using its version of COMPSTAT, the New York City Department of Correction has, for example, significantly reduced inmate violence and cut overtime expenditures. Using PARKSTAT, the New York City Department of Parks and Recreation has significantly improved park facilities safety and cleanliness. Now other agencies, such as the City’s Human Resource Administration (HRA) and its Department of Homeless Services (DHS), are developing their own versions of COMPSTAT.

In the NYPD’s own exposition of how COMPSTAT works, there are four principles: (1) Accurate and

---

1 The Mayor’s Office of Operations produced a guide in 1986, “Mayor’s Management Plan and Report System Manual.” Consistent with the low level of investment by the City in training managers to use the MMPRS, by 1989, in my study of the MMPRS, I found that virtually no actor interviewed, including staff in the Mayor's Office of Operations, had ever heard of this highly informative Manual.

2 The process of public sector performance measurement has reached a new level of refinement in New York with the publication by the Fund for the City of New York of an independent study of the quality of streets and roads, serving as a check on the performance measures reports in the MMR. New York Times. October 19, 1998, p. 1.

3 Donald Kettl, 1996. Kettl, a leading student of GPRA, comes closest to addressing this issue when he recommends that performance measures be related to strategic plans required of agencies, but his conception of the elements of strategic plans does not include the explication of a programs logic or theory, nor therefore, the idea of tracking progress along the causal chains. According to Kettl, “Good strategic plans have six steps: Define the mission... Frame the goals... Set the objectives.... Assign responsibility for achieving objectives...Specify output/outcome measures.... Compare results with goals.”
timely intelligence, (2) Effective tactics, (3) Rapid deployment of personnel and resources, and (4) Relentless follow-up and assessment. While these are all critical elements of the model, as was shown in the model of COMPSTAT elaborated by Smith (1997) there are a number of other necessary components to this successful approach to performance management. Devolution of discretion to act on timely intelligence, and accountability for acting effectively, combined with experience tested strategies for achieving results are perhaps the most crucial additional elements of the model. (See Model I) These strategies, while typically not formally explicated, are themselves causal or logic models of interventions that can be used to track performance. They are a set of implicit “if–then” assumptions (A causes B, B causes C and D, and so on) in which the last element in the causal chain is “improved performance.” This causal thinking is a key to successful performance management.

**The Case for “Theory-Based” Performance Management**

Carol Weiss and other advocates of theory-based evaluation (TBE) see significant advantages for evaluators.

The TBE approach is intended to help evaluators focus the study on key questions. With inevitably limited resources and time, that evaluator always has to make choices on what to study. TBE can help to sharpen the aspects of the program that are of interest to the study’s audiences.

TBE also enables the evaluator to provide information about interim markers of progress. Many programs aim to achieve goals that are far in the future...Many years must elapse before such outcomes can be expected. In the interim, this approach to evaluation provides information on short-term and intermediate outcomes—not just any intermediate outcomes available, but outcomes that are linked, according to the best available, knowledge, to the long-term outcomes of interest. When the evaluation is completed,... [with] data that follow the chains of program theory, [evaluators] can say more useful, relevant and practical things.

Weiss also envisions other advantages to policy-makers, program funders and managers: If they can understand the what, how, and why of program success (and failure), they can undertake new ventures better prepared to reproduce those elements of the program that are associated with the successful transition to the next link in the chain of assumptions and to rethink and rework those program elements that do not lead to the next program marker.

Carol Weiss and other proponents of TBE are not oblivious to the challenges posed by the approach. Her recent article “How can theory-based evaluation make greater headway?” is devoted to trying explain why an approach that is “cogent and plausible” and “promises to bring greater explanatory power to evaluation,” “appears to be having only marginal influence on evaluation practice.” Many of the problems she identifies would be addressed if, instead of being a perspective imposed on programs from the outside by evaluators--and sometimes retrospectively, the focus on theory were integral to the operation of the program-- and developed for their own use by the managers. If the approach to performance management presented below were adopted, not only would the advantages Weiss promised accrue to managers, but their use of a theory-based performance-management approach would provide a solid foundation for theory-based evaluation.

---

4 Howard Safir, Police Commissioner, The COMPSTAT Process, Prepared by the Office of Management Analysis and Planning, the Police Department of the City of New York, n.d.
5 As addressed in Smith, 1997, as a corollary of the idea that “what gets counted --counts,” the failure to include police misconduct as one of the outcomes tracked by COMPSTAT when it was introduces may have contributed to rising civilian complaints and some infamous cases involving allegations of use of excessive force by NYPD officers.
6 Weiss regards the discussions in the literature of “theory of change” and intervention models as addressing the same issues as TBE. In the General Accounting Office's Prospective Evaluation Synthesis methodology, the term "conceptual model" is synonymous with the term “theory” in TBE.
In an analysis of COMPSTAT and other similar approaches to performance management in New York City, Smith (1997) argues that the greatest successes came when managers had an explicit theory to guide their collecting and tracking of data on performance. The early breakthrough in reducing subway crime came when police leaders determined that fare beating (jumping turnstiles), while technically a minor offense, was probably a predictor or more serious criminal intent. When enforcement of fare beating led to the discovery of weapons and outstanding arrest warrants, and led to offenders giving evidence against others to reduce their now more serious charges, the original “theory” was validated and refined. At NYPD, COMPSTAT performance review meetings “strategies” for effective action against particular crimes guide the data analysis and reporting process. These explicit strategies serve the function of program theories.

Similarly, the breakthrough in reducing inmate violence came largely through the discovery of certain patterns in the escalation of conflict involving the importation of weapons by visitors and the extension of street gang connections into the facilities. Using these insights, intervention strategies (theories of cause and effect) were developed and tracked by their COMPSTAT-like system called TEAMS.8

Public management, having only recently begun to measure the performance of its own agencies and programs without entirely mastering the task, is now faced with a more challenging demand on its competence. Increasingly government agencies, regardless of their capacity to deliver goods and services directly, are being charged under the prevailing policy wisdom to contract out much of their missions. In a review of a recent Brookings Institute report on the newly renamed National Partnership for Reinventing Government (formerly the National Performance Review), Washington Post writer Stephen Barr cites the report’s author’s observation that the fact that the Federal government increasingly provides federal services through partnerships with nonprofit organizations and state and local governments, makes the challenge of reinventing especially daunting.9

The provision of services by the federal government through nonprofits and state and local governments is not new. What is new in the era of reinvention is the requirement that performance management focus on results, on outputs and impacts, not inputs. While there can be no denying that government must be held accountable for what it actually produces in terms of valued goods and services, it is equally undeniable that measuring the results of public investments is harder than measuring expenditures. It is far easier to count the number of staff than to gauge how much service they produce, and even harder to determine the effects of those services. The services often are delivered in a complex bundle in a complex environment, and in many cases their intended effects are expected at some indefinite and perhaps far-in-the-future date. Nevertheless, as Alice Rivlin observed more than two decades ago, privatization of government services requires the development and refinement of performance measures.10

Just as governments at all levels are involved in the reinvention-based quest for better measures of performance, the nonprofit sector has also seen the handwriting on the wall. Margaret Plantz and her colleagues have summarized recent efforts to add outcome measurement to the performance measurements used by nonprofit organizations.11 They identify the many pressures pushing nonprofits to measure outcomes, and spell out thirty “lessons learned” from nonprofit organizations’ early experience in using outcome measurement. They cite three lessons about the value of outcome measurement, eleven lessons about implementing outcome measurement systems, six about “useful

---

roles for funders,” eight about using outcome measures in resource allocation, and two about the limitations of outcome measures. Taken together these lessons offer a lot of practical wisdom about measuring outcomes. However, their focus is on managing measurement rather than on the broader issue of using outcome measurements to manage performance. As a result, the role of program theory is quickly passed over. Lesson Seven (implementing outcome measures): “Creating a written logic model of program inputs, activities, outputs, and outcomes is a helpful way to think through the changes participants experience during and after the program and to check the logic of the if then influence the program intends to set in motion.” If, as the authors state earlier, Lesson Three (the value of outcome measurement): “The most important reason for implementing outcome measurement is that it helps programs improve services,” then the role of logic model is more than just one helpful tool among many. Understanding and tracking “the changes participants experience” is central to whole concept of public service accountability.

**Self- Assessment as a Management Tool for Non-Profits**

While clearly not a panacea for the litany of problems summarized so well by Smith and Lipsky, we will explore the utility of a new approach to performance assessment developed by Human Service Solutions and currently managed by Seedco and its recent affiliate, the Non-Profit Assistance Corporation (Seedco/N-PAC), to support a special initiative of the Charles Hayden Foundation. The Charles Hayden Foundation General Support Grant Initiative (the Initiative), a unique combination of funding, technical assistance, and evaluation, was devised by the Foundation to support capacity building projects in nineteen youth serving organizations in New York City and Boston. As part of the Initiative, nonprofit organizations’ leaders have received training in performance management using key elements of the COMPSTAT model. In addition, the aspect of a theory-based approach to performance management, which is implicit in COMPSTAT, is made explicit in the contract between the Charles Hayden Foundation and the youth serving organizations selected to participate in the Initiative. As a result, the Foundation and Seedco/N-PAC have developed and are now using an approach to performance-based grants management that is a potential model for improving the prospects of performance-based human services grants and contracting in general.

**Conclusions**

In Policy Paradox: The Art of Political Decisionmaking (W. W. Norton, 1997) Deborah Stone develops perhaps the most elaborate critique of the theory and practice of policy analysis in the literature. Some the concepts central to her critique, not surprisingly have a central place in the approach to performance management discussed here. Theory based performance management requires the explication of causal assumptions embedded in policies, programs, and even organizational practices. Stone cautions against oversimplifying the concept of causation. Using two kinds consequences, intended and unintended, and two kinds of actions, unguided and purposeful, she maps out for kinds of causal theory: mechanical, accidental, intentional and inadvertent. Policy analysis occurs in an explicitly political context and assumptions about causation have significantly more than scientific implications. Causal assumptions underlie all discussions of administrative--and political--accountability. Ideas about cause and effect are central to political debates and the fate of politicians in democracies. Therefore, ideas about cause and effect are heavily freighted.

In another potentially relevance parallel with performance management, Stone also singles out for critical scrutiny the prominent place of "numbers" or quantification in policy analysis. Again her critique is complex but the basic point is that numbers never speak for themselves but almost always reflect unstated assumptions and submerged (often value) premises.

---

12 Seedco is a national nonprofit community development intermediary that is dedicated to rebuilding and revitalizing impoverished communities. The Non-Profit Assistance Corporation (N-PAC), an affiliate of Seedco, designs and implements strategies and programs to support community based organizations.
These criticisms are thought-provoking and cautionary when applied to policy analysis as practiced in a single culture and political system, it seems to me they take on even greater resonance in the context of the transnational diffusion of public administration reform ideas.

In American public administration the idea of public sector performance management, as presented here, seems to offer the prospect of improved, more accountable public services. Within the context of its American application there may be no reason to ponder the importance by the underlying structure of checks and balances in the American political system, or the relative place of transparency in the hierarchy values of the American political culture, because those factors are, relatively speaking, constants in the equation. In the transport of these ideas and practices to other polities and cultures factors such as these may be of paramount importance. That is a conversation this paper is intended to provoke.

Bibliography


Kathryn E. Newcomer in *New Directions for Evaluation*, No. 75, Fall, 1997.


Peter Smith, editor, *Measuring Outcome in the Public Sector* (Taylor and Francis Ltd., 1996)


Professor Dennis C. Smith, Director
Office of International Programs
Robert F. Wagner Graduate School of Public Service
New York University
624 Tisch Hall
NY., NY 10012 212-998-7425
fax 212-995-4162
smithde@is.nyu.edu