Performance Management Comes to Washington:

A Status Report on the Government Performance and Results Act

Dall Forsythe
Senior Fellow
Rockefeller Institute of Government

February 25, 2000

The Nelson A. Rockefeller Institute of Government

State University of New York

Contents

Preface iii

Executive Summary 1

Introduction 5

The Results Act: Progress in Implementation 11

- **Threats to Further Progress** 15
- **Leadership Transitions and Management Reform** 16
- **Building GPRA into the Routine Work of the Executive Branch** 17
  - **Performance Management in the Operating Agencies** 17
  - **Passing the "Influence Test"** 20
  - **Incorporating Performance Measures in Budgeting** 21

Making GPRA Work for Congress 23

- **Overselling Performance Management: The Several Dimensions of Management Reform** 25
- **Avoiding Measurement Myopia** 27
- **Managing Performance of Independent Agents** 29

Rewarding Performance 33

Conclusions and Recommendations 35

Appendix A: More on GPRA Legislation and Implementation 39

Appendix B: Great Expectations: The History of Management Reform 41

Appendix C: A Typology of Performance Measures 43

Appendix D: Excerpt from Veterans Administration Departmental Performance Plan, FY 2001 49

Selected Bibliography 51

Preface

Seven years ago, the U.S. Congress enacted an ambitious law to reform government to focus on getting results — the Government Performance and Results Act. The law is now at a critical juncture. The history of the law, its purposes, rationale and possible future are ably described in this report by Dall Forsythe, who directs the Rockefeller Institute's Project on Performance Management and Measurement in Government. Forsythe, a former New York State budget director and former lecturer at Harvard's Kennedy School of Government, is a Senior Fellow at the Rockefeller Institute.

After extensive study of the subject as well as interviews and meetings with a large number of federal officials, Dr. Forsythe's principal finding is that the federal government has made surprisingly good progress in implementing the Government Performance and Results Act (GPRA). It will be some time before we can be sure that those efforts have in fact produced results in the form of improvements in services, but the data needed to begin to make that assessment will be available this spring. All agencies will make their initial reports on performance by the end of March 2000, and this data will give Congress and the public an initial impression about how well those agencies are managing for results.

In spite of these steps forward, GPRA still faces significant hurdles if progress is to continue. The push for performance management could falter as a new president takes office. It could also bog down if the federal government fails to create stronger incentives for high-performing managers and agencies. It is also true that progress has been slower than expected in pilot programs linking performance and resource allocation and for programs that operate through intergovernmental grants-in-aid. GPRA-mandated performance budgeting experiments in five agencies are behind schedule.

These and other findings are contained in this progress report on the Government Performance and Results Act, which is the first published product of the Rockefeller Institute's Project on Performance Management and Measurement, funded by the Pew Charitable Trusts. An edited volume on this topic is scheduled for publication at the end of the year. To provide advice and guidance for this project, the Pew Trusts and the Rockefeller Institute of Government assembled a Task Force of distinguished public management specialists. In addition to myself and Dall Forsythe, Task Force members include:

- Walter Broadnax, Dean of the School of Public Affairs at American University in Washington.
- Patricia Ingraham, Professor at the Maxwell School at Syracuse University, and Director of the Government Performance Project.
- Donald Kettl, professor at the University of Wisconsin and Director of the LaFallotte Institute.
- Paul Light, Vice President and Director, Governmental Studies, at the Brookings Institution.

- Allen Schick, Professor at the University of Maryland.

Members of the Task Force met twice to discuss these issues, and several members provided written comments. Additional comments were received from Paul Posner and J. Christopher Mihm at the General Accounting Office; Jonathan Breul of the Office of Management and Budget; Barry White of the Council on Excellence in Government; and Donald Moynihan of the Maxwell School. Elaine Casey of the Pew Charitable Trusts provided assistance far beyond that expected of a foundation program officer. At the Rockefeller Institute, comments were provided by Donald Boyd, Joseph Burke, James Fossett, Thomas Gais, and David Wright. Project management assistance for the Task Force and its activities was provided by Miriam Trementozzi and Ellen Blake. Sandra Hackman provided editorial advice, and Melinda Mowery assisted in checking references. Lisa Taylor provided clerical support.

Richard P. Nathan
Director
Rockefeller Institute of Government

Executive Summary

Congress passed the Government Performance and Results Act (GPRA) into law in 1993.[1] The Act requires federal agencies to develop strategic plans and goals, and to create performance measures to track progress toward those goals. More broadly, the Results Act aimed to improve management in federal agencies, provide better information for congressional decisionmaking, and make agencies more accountable to the public for service delivery and program outcomes.

> ***GPRA Requirements***
>
> - Strategic Plans
> - Annual Performance Plans
> - Annual Performance Reports
> - Performance Budgeting Pilots
>
> See Appendix A on page 39 for more details on GPRA.

Although mixed across agencies, progress in the development of the performance management system required by the Results Act overall has been good — indeed, much better than students of the history of management reform might have expected. The Act has given impetus to much wider use of performance measurement, and new standing to the efforts of federal agency heads to manage programs better. Properly used, its strategic planning activities can provide needed focus for policy-makers and top managers. As additional performance data become available in March of 2000, GPRA will also create new accountability mechanisms, as its reports finally begin to provide timely and consistent information to citizens about government's activities and its progress in improving services. If the GPRA performance management system works as expected, decisionmakers will spend more of their time considering strategically important goals, and operating managers will focus more of their time and energy on improving performance in the same critical areas.

However, implementation of the Government Performance and Results Act stands at a critical juncture. During the next two years, several factors could undercut the program's momentum. Risks include:

- The election of a new president, who may place less priority on managing for results.
- Sluggishness by the executive branch in building GPRA's required plans and measures into the day-to-day work of managing government operations.
- Inadequate progress in the use of those plans and measures for policy direction and budgeting in the executive branch.
- Lack of interest by Congress in using GPRA data in its oversight and appropriations activities.

The next president may embrace performance management and GPRA, and Results Act champions are certain to continue their push to build its routines into the day-to-day activities of the

---

1 Public Law 103-62. The text of the Act can be found on the web at http://thomas.loc.gov/.

executive and legislative branches of government. But even if those challenges are met, management reformers must still guard against several longer-term threats.

For example, the Results Act could be derailed by unrealistic expectations about what it can accomplish. At this stage in its implementation, proponents seem to have abandoned their most grandiose claims, but the program can still be undermined if its champions fail to recognize that managing for results is only one of several important dimensions of good management. Alternatively, GPRA implementation may tail off into a sterile technical debate about performance measurement, sinking into a sea of detail and losing sight of its central goal of improving strategically important government outcomes.

For many decades, federal programs have used private and not-for-profit contractors to fulfill their purposes, and many more have delegated responsibility for service delivery to state and local governments. Moving from Washington to statehouses and city halls, GPRA implementation faces another formidable hurdle — the challenge of using performance management and measurement systems to monitor and improve contracts as well as grants and other intergovernmental programs. Federal officials have been grappling with this problem for years, and GPRA's architects simply underestimated the difficulties of getting nonprofits or other levels of government to report on and comply with performance goals.

To meet these challenges, reformers must sharpen their agenda and lay out more clearly their goals for the transitional period that will follow the next election. The newly elected president needs to understand that there is a constituency for improving management in the federal government, and that continued progress will enhance his standing as a strong manager in 2004. Congress must also embrace performance management in its own routines of oversight and budgeting. Reformers must be realistic about both the virtues and the limitations of GPRA, and must present a management agenda that is not limited to performance management. This broader framework for management improvement should include:

- ❖ A renewed push for experiments in performance budgeting — using performance data to allocate resources — accompanied by new impetus for the stalled federal cost-accounting initiative.
- ❖ A concerted effort to use evaluation research in tandem with performance measures to assess the impact and benefits of complex federal programs.
- ❖ Creation of a broader spectrum of incentives for high-performing individuals and organizational units in the federal government.
- ❖ New attention to the problems and possibilities of measuring and managing performance of contracts, grants, and other intergovernmental programs.
- ❖ A wider view of strategic plans, which must include systematic initiatives to improve the management of agency finances, technology, capital assets, and human resources.

More broadly, the federal government must change its philosophy of management. As a comprehensive system of performance measurement develops under GPRA, top managers and congressional overseers should proceed more confidently to devolve authority, giving operating managers more flexibility to manage their programs even as they hold those managers to higher standards of actual performance. At the same time, the federal government should continue to look for opportunities to use private-sector and nonprofit providers to increase competition and provide practical benchmarks for comparison with federal programs.

Introduction

Since the passage of the Government Performance and Results Act (GPRA) in 1993, executive branch agencies in the federal government have been working their way through a menu of interlinked activities required by the Results Act. With varying degrees of success, agencies have now formulated:

- Strategic plans laying out long-term goals.
- Operational strategies to reach those goals.
- Performance indicators to track progress in implementing those strategies.
- Annual targets for showing improvement in those indicators.

March 31, 2000, marks a significant milestone in this process. For the first time, every executive branch agency will publish reports on actual performance. These reports will help citizens and federal overseers gauge how successful agency managers have been reaching the targets in their annual performance plans. This essay gives the reader the history, background and vocabulary to understand the debate that those performance reports will engender.

At the heart of the performance management movement in Washington and around the world is a simple set of insights — managers cannot act consistently to improve services without data to track performance, and policymakers and citizens cannot make judgments about the relative value of government activities without measures of program performance and impact.

These two goals — management improvement and better accountability — are the rationale for the Results Act. Some federal policymakers, especially members of Congress, also hope that clearly stated goals and indicators of performance will help them better oversee and coordinate what they call "cross-cutting programs." Cross-cutting programs, which are housed in different agencies, aim at the same constituency or provide similar and potentially overlapping services. For example, four different agencies in two cabinet departments manage the vast bulk of federal lands: the Forest Service in the Department of Agriculture; and the National Park Service, the Bureau of Land Management and the Fish and Wildlife Service, all located in the Department of the Interior. The organizational structure in Congress multiplies the confusion, with oversight responsibility split between appropriating and authorizing committees in each chamber.

Proponents of GPRA believe that consistent use of its tool kit of plans and indicators will teach the federal government to "manage for results." Will the Results Act revolutionize management practice in Washington? Alternatively, is performance management just another fad, dressed up in a new set of buzzwords? Or is it something in between — less than revolutionary, but more substantive and enduring than many earlier management reform initiatives?

In the private sector, performance measurement has been a standard management tool since Frederick W. Taylor published *The Principles of Scientific Management* in 1911. Since then, managers in business have tried to improve their ability to measure inputs (labor, dollars, and materials) as well as the outputs (products) of their manufacturing processes. In conjunction with the tools of cost accounting, managers have used performance measurement to help reduce the cost of producing a unit of output. More recently, under challenge from foreign manufacturers, U. S. corporations have realized that they needed to move beyond measuring gross inputs and outputs and turn their focus to improving product quality and customer satisfaction.

In much the same way, the Results Act aims higher than the simple measurement of government inputs and outputs, which has been a focus of management reformers since the first Hoover Commission in 1949. The current initiative aims to focus the attention of managers and decisionmakers on the most important government services and activities by combining performance measurement with strategic planning.

The first step in creating a strategic plan usually entails defining an organization's mission — its fundamental focus and essential activities — more precisely. Next, the organization's leaders outline long-term goals consistent with that mission, and develop broad strategies or operational proposals to accomplish those goals. A well-articulated strategic plan can help all the managers in an organization, from first-level supervisors to top managers, steer an organization in the same direction. In a carefully laid-out series of steps, GPRA next requires agencies to develop indicators that can be used to measure progress toward those strategic goals, and to establish annual targets for performance improvements as measured by those indicators. Finally — in this case, by March of 2000 — the agencies must report actual progress against those targets.

The use in government of tools developed in the private sector often raises new questions and difficulties, and strategic planning can be more problematic in government agencies than in corporations. Businesses have relatively clear goals — to increase profitability and thus boost shareholder value. The goals of federal agencies are more complex, and those organizations often pursue several goals and serve several constituencies. For example, the Forest Service's timber-management activities produce needed revenue and profits for timber companies, but also affect the health of forest ecosystems and their use for recreation and wildlife conservation. When agencies serve competing constituencies, clarifying their long-term goals can exacerbate conflicts among those groups. Government agencies must also satisfy elected officials who are concerned about levels of employment and other economic impacts in the areas they serve. Finally, agency leaders must also obey federal rules and regulations enforced by the Office of Management and Budget, the General Accounting Office, and other overseers and control agencies.

Just as corporate managers had to shift their focus to quality and customer satisfaction, the architects of the federal performance management system have also raised their sights, asking agencies to look beyond inputs and simple outputs. Specifically, the Results Act requires agencies to try to measure the broader social, economic, and environmental impacts of governmental action — called *outcomes* in the technical jargon of performance management. To continue with the Forest Service example, simple outputs would include numbers of visitors to forests, and board-feet of timber cut. Long-term outcomes, on the other hand, would aim at "land health," or sustainable ecosystems. This focus on program outcomes helps distinguish GPRA from earlier efforts to measure performance in federal agencies.

Operational managers sometimes resist efforts to measure their success or failure by broad outcome measures. Managers are concerned about being held accountable for outcomes that seem only distantly related to their day-to-day activities, and that they can only partially control. Agency managers often feel threatened by other aspects of performance management systems as well. For example, they know that they will inevitably fall short on some measures of outputs. Unlike corporate managers, whose failures are not always widely reported, public-sector agency heads can reasonably expect that all their failures will be publicized and punished — and worry that their successes will go unheralded and unrewarded.

Some GPRA sponsors have put forth even more grandiose visions of the Act's impact, envisioning a time when budget and policy decisions would be closely linked to agency and program performance, and performance data would help decisionmakers eliminate marginal programs, allowing resources to be refocused on high performers. Thus, while GPRA's main focus is "managing for results," it also contemplates "budgeting for results." In fact, the Act required limited experiments with "performance budgets," tying resource allocation to specific performance targets, in an attempt to harness the power of the budget process in support of performance improvement. Those experiments are behind schedule, and progress in the budgetary arena has been slower than contemplated in the Results Act.

Looking at the GPRA apparatus, a skeptic might argue that performance management champions have simply invented new vocabulary for basic management practice. Don't federal officials already "manage for results"? In fact, top managers in federal agencies do not always focus on improving performance, but frequently aim at other goals. Their objectives can include:

- ❖ Complying with rules and regulations, seeking to avoid negative press and audits from overseers.
- ❖ Generating positive press coverage about their programs, a priority for political executives and the president.
- ❖ Meeting the demands of interest groups and congressional overseers — the other two parties in the famous "iron triangle" of Washington policymaking.
- ❖ Maximizing funding for and jobs in a bureau or department.

Reformers label this complex of objectives "the bureaucratic paradigm," and argue that performance management systems, linked to the budget process, can create a new results-oriented paradigm.

Indeed, a careful look at GPRA implementation reveals that the initiative has met many of its own management goals. Agencies have developed strategic plans, and some — the Department of Transportation, for example — are models of their kind. Working with Congress and the General Accounting Office (GAO), executive branch personnel have also developed indicators to track progress against those plans, and many are sophisticated tools to measure important outputs and fundamental program outcomes. In sum, GPRA seems to have moved federal management practice a significant step beyond its previous level, and has given recognition and standing to managers in federal agencies who care about improving the delivery of government services.

An important implementation milestone will be reached in March of 2000, when the entire federal government will, for the first time, report on actual performance for the prior fiscal year. In these initial reports, agencies are likely to show good progress in meeting many of their targets, in part because Congress and OMB did not push hard to establish difficult-to-attain goals for the first full set of performance plans. However, GAO and others are likely to have trouble verifying these performance data. Many agencies are still collecting and validating data for use in performance measurement, and they will be scrambling to meet GAO's standards of timeliness, accuracy, completeness, and consistency in this first set of performance reports.

- ❖ The initial set of agency reports on performance should produce a flurry of interest and analysis. Apart from their intrinsic interest, they will also produce baseline data against which citizens and members of Congress can measure future performance improvements. As each year's data extends this baseline, we will be able to make much more accurate judgments about whether federal agencies are succeeding in one of three basic performance goals — delivering higher quality services, on a timely basis, at the lowest possible cost.

That information, in turn, will help us determine whether the GPRA performance management system has succeeded in fostering widespread improvement in management practice in federal agencies, or whether its effects will be much less far-reaching. But while performance reports are an important milestone, they also raise important questions. How can the public and elected officials best use these new tools for accountability? Will the push to develop broad measures of social outcomes distract managers from the more straightforward tasks of improving basic services? Will OMB, the president, and Congress find these new tools helpful in deciding how to allocate resources in some or all federal programs?

These questions suggest that some of the most difficult terrain for implementation of the Results Act lies ahead, and that the program could still fall short of some of its goals. To maintain momentum at the federal level, GPRA advocates will have to walk a fine line, maintaining their optimism about performance management without promising more than GPRA can deliver, and emphasizing over and over again that improvements in service delivery and accountability — not improvements in measurement — are GPRA's goal.

The rest of this White Paper focuses on each of these issues in greater detail, analyzing the risks to continued success in implementing GPRA. These risks constitute a series of barriers to successful reform of management at the federal level:

- ❖ Transitions of executive leadership, since management reforms are often jettisoned when new chief executives take office.
- ❖ The deep-seated problems faced by proponents of reform as they seek to build performance management into day-to-day management and budgeting routines in both the executive branch and in Congress.
- ❖ The danger of inflated expectations about what performance management can accomplish without improvements in other areas of management.
- ❖ The contrary threat that managers will narrow their goals too much, myopically focusing on technical measurement issues and losing sight of GPRA's orientation toward results.
- ❖ The complex problems of measuring performance in programs where services are delivered by agents not under the direct control of the federal government. These

problems are especially difficult when state and local governments share responsibility for implementing programs.

- The risk that the federal government will aim too low as it develops incentives to reward high performance by its employees and its agents.

GPRA proponents need to work their way carefully through these barriers if performance management is to fulfill its potential for the federal government.

The Results Act: Progress in Implementation

With some fanfare, Congress passed the Government Performance and Results Act in August, 1993. Students of government reform immediately catalogued GPRA as the latest in a long line of ineffective management reform initiatives aimed at the federal government. Earlier federal efforts at systemic management reform — including Program Planning Budgeting Systems (PPBS), Management by Objective (MBO), and Zero-Based Budgeting (ZBB) — indeed underperformed expectations. Oversold and poorly designed, those management reforms were unrealistically scaled and difficult to sustain and advance. Largely as a result, they died relatively quickly, only to be reinvented in a seemingly inexorable cycle. Lacking limited and realistic expectations, reformers seem doomed to failure.

GPRA proponents agreed that the historical record was not reassuring. They argued, however, that even the failed reforms had improved management and budgeting practice, although those improvements were much less far-reaching than originally promised. They also pointed out that the drafters of the Results Act had taken steps to avoid some of the problems faced by earlier reforms,[1] and that GPRA's design was also informed by the successful experience in managing for results in several OECD countries.[2]

Post-War Initiatives to Reform Management and Budgeting

Reform Initiative	*Date*	*President*
Hoover I	1949	Truman
PPBS	1965	Johnson
MBO	1973	Nixon
ZBB	1977	Carter

See Appendix B on page 41 for more on the history of management reform.

Several key features of GPRA reflect those design decisions. Previous reforms had typically been initiated by presidents and were of limited interest to Congress. The Results Act was a legislative initiative, and it was designed to produce consultation between Congress and the Executive Branch in developing strategic plans and performance measures. While previous efforts were often rushed to government-wide implementation in the space of a year or two, GPRA provided for a long and careful timetable for implementation, with pilot programs as intermediate steps. This extended roll-out has helped avoid the confusion caused by hasty implementation of reforms like PPBS and ZBB. Moreover, the simple fact that the program has congressional support and a statutory base provided staying power that earlier executive initiatives could not muster.

Although GPRA was a congressional initiative, the Clinton administration embraced the Results Act with enthusiasm. Management improvement was an important priority for top Clinton officials, including the vice president, who headed the National Performance Review (NPR) established in 1993. While performance management and measurement were not the primary focus

1 GAO 1997.

2 Breul 1994.

of NPR, Gore's sustained focus on management improvement did give GPRA implementation high-level political backing, and helped offset the skepticism with which many bureaucrats greeted the Results Act.

These positive factors notwithstanding, the early history of GPRA implementation gave credence to skeptics. Members of Congress issued "report cards" on agencies as Results Acts deadlines were reached, and most of those report cards focused on failure. Majority Leader Dick Armey issued a preliminary review in summer of 1997. In Armey's grading system, a perfect score for an agency's strategic plan was 105. The top-ranking agency plan received 62 out of 105 possible points, while the worst received only 6.5 points. The average score was 30. The Departments of Labor (DOL) and Housing and Urban Development (HUD) were singled out for especially sharp criticism. Armey complained about agency disinterest in major management problems, and about the "scant attention" paid to "cross-cutting functions," Washington shorthand for program activities in different agencies that serve overlapping constituencies. Cynics complained that Congress was trying to use GPRA to shut down government, not improve its performance. Reformers worried that GPRA might bog down in partisan debate. Executive branch officials defended their efforts, sometimes escalating the rhetoric. Privately, they worried that NPR's personnel cutbacks of about 275,000, initially targeted at managerial ranks, could bleed the life out of management improvement programs.

While press attention focused on Armey's "report cards," the General Accounting Office began to produce detailed reports tracking on GPRA implementation in individual agencies. These reports also drew broader lessons about the impact of the program as a whole. Anchored by GAO's careful and professional analysis, the tenor of oversight discussion soon lost its partisan edge. Later in 1997, commenting on the final versions of agencies' strategic plans, Congressman Armey raised his grades, giving the top plan 75 points and raising the average grade to 46. As congressional rhetoric cooled, the executive responded in kind. In the FY 2000 Government-Wide Performance Plan, OMB stated that:

> The Administration has appreciated the interest of Congress in promoting the use of the strategic and annual performance plans in the appropriations process and in conducting authorization and oversight activities. We look forward to continuing to work with Congress and other stakeholders on expanding the usefulness of these tools to achieve better performance.[3]

At the same time, executive agencies were improving their GPRA products, producing better strategic plans and annual performance plans.

At least two sources now provide non-partisan and systematic assessments of agency efforts at managing for results. In addition to GAO's work, the Government Performance Project (GPP), headquartered at the Maxwell School in Syracuse and funded by the Pew Charitable Trusts, "grades" federal agencies on their management activities. In February 1999 the GPP issued reports on fifteen federal agencies. The Veterans Health Administration received an A, eight others got grades of B, and six agencies received C grades. Overall, then, 60 percent of federal grades for managing for results fell in the A and B range, and the other 40 percent received C grades. The GPP also graded states, and 42 percent received grades in the A and B range, 46 percent fell in the C

3 OMB 1999, p. 2

range, while 12 percent got D or F grades. The comparison speaks well for federal efforts under GPRA, since the states have been working longer on performance management systems, and the services state agencies provide are arguably easier to measure than federal programs.[4]

As early as 1997, both the GAO and congressional leaders praised the strategic plans developed by the Department of Education and the Social Security Administration (SSA).[5] The plans of the Department of Justice, the Department of Energy, and the Nuclear Regulatory Commission were also held up as models.

Inside and outside of government, new forums developed to discuss and disseminate agency successes in planning and measurement. The National Academy of Public Administration created a Center for Improving Public Performance and published a series of booklets on best practices. The President's Management Council, which consists of agency and department chief financial officers, developed a training program for executive branch personnel, and the nonprofit Congressional Institute initiated training sessions for members of Congress and their staff. At the same time, private consultants and conferences on these topics proliferated.

By mid-1999, when the GAO reviewed annual performance plans of 24 agencies, the Department of Transportation and the Social Security Administration garnered high praise, with the Departments of Education, Justice, and Labor and the General Services Administration close behind. Laggards included NASA and the Small Business Administration, near the bottom of the pack, and the Departments of State and Treasury, which ranked last. GAO summed up progress as follows:

> On the whole, agencies' fiscal year 2000 performances plans show moderate improvements over the fiscal year 1999 plans and contain better information and perspective. However, key weaknesses remain, and important opportunities exist to improve future plans.[6]

Although most agencies had strengthened their goals and measures, they had failed to make much progress analyzing cross-cutting programs or establishing multi-agency goals and measures to track progress in those programs. Looking ahead to performance budgeting, agencies were supposed to reconfigure their budget requests to show linkages to performance goals, but only a few agency plans did this successfully. GAO singled out improvements in SSA's ability to show such linkages, and highlighted different approaches used by several agencies, including the U.S. Geological Survey in the Department of the Interior and the Nuclear Regulatory Commission.

Most of the plans received their worst marks on data verification and validation. As GAO described the problem:

> Congressional and executive branch decisionmakers must have assurance that the program and financial data being used will be sufficiently timely, complete, accurate, useful and consistent if these data are to inform decisionmaking.[7]

---

4 GPP's assessment of federal agencies is published in *Government Executive,* while their report on state efforts appears in *Governing* magazine. The data are also available on GPP's website, at http://www.maxwell.syr.edu/gpp/.

5 Armey 1997.

6 GAO July 20, 1999.

7 Ibid, p. 28.

While agencies had designed performance indicators that made sense, many had not completed collection of the data required by those indicators, and others were using data that GAO had previously criticized as unreliable.

Although problems remain, some agencies have reached high standards of performance planning and reporting. The Department of Transportation, battered by Armey's report card in 1997, is now widely praised for its GPRA plans, including its analysis of cross-cutting programs. DOT's current strategic plan is a tightly packed 15-page statement of goals in five key areas, and strategies to achieve those goals.[8] For each major goal, DOT provides strategies to reach it in all industry segments, such as highway, railroad, mass transit, air transportation, and marine safety. The 135-page performance plan includes specific quantitative goals for several programs within a bureau, such as the Coast Guard's aim to reduce recreational boating fatalities by 22 percent. The Coast Guard has also been widely lauded for initiatives to shrink staffing and refocus and improve their commercial marine safety program, reducing fatalities by focusing on high-risk industry segments. The DOT document provides baseline historical data for key goals in simple graphic form. Appendices describe in detail the sources of data that will be used to measure performance, and provide a cross-walk between budget accounts and the five categories of strategic goals.

DOT's Strategic Goals, FY 2000

- **Safety** — Promote the public health and safety by working toward the elimination of transportation-related deaths, injuries, and property damage.
- **Mobility** — Shape America's future by ensuring a transportation system that is accessible, integrated, efficient, and offers flexibility of choices.
- **Economic Growth and Trade** — Advance America's economic growth and competitiveness domestically and internationally through efficient and flexible transportation.
- **Human and Natural Environment** — Protect and enhance communities and the natural environment affected by transportation.
- **National Security** — Advance the nation's vital security interests in support of national strategies such as the National Security Strategy and National Drug Control Strategy by ensuring that the transportation system is secure and available for defense mobility and that our borders are safe from illegal intrusion.

Even before the statutory deadline for performance reports, agencies have begun to present early examples of actual improvements in results.

- The Social Security Administration, which uses survey data to track customer satisfaction, reported that 85 percent of customers rated service good, very good, or excellent in FY 1999, compared with 80 percent in FY 1998. SSA has established a target of 88 percent for FY 2000. The agency also reported significant increases in the number of customers using Social Security Online — from 4.2 million in FY 1998 to 5.6 million in FY 1999, with a target of 7 million in FY 2000. The agency

8 DOT's strategic and performance plans are available on the Web at http://ostpxweb.dot.gov/budget/8measure.htm.

continued to meet its goals for 800-number responsiveness and waiting times in regional offices.[9]

- The National Oceanic and Atmospheric Administration has been praised for establishing — and sometimes achieving — clear and aggressive goals for producing better weather forecasts. For example, NOAA has set quantitative targets requiring longer lead time for severe weather warnings for tornados, flash floods, and severe thunderstorms; better 24-hour predictions about where hurricanes will hit land; and more accurate predictions for heavy snowfalls.[10]
- The Veterans Health Administration has increased performance on its Chronic Disease Care Index (CDCI), which measures how well its caregivers follow nationally recognized clinical guidelines for heart disease, hypertension, chronic obstructive pulmonary disease, diabetes mellitus, and obesity. The CDCI has risen from 44 percent in 1996 to 91 percent in 1999 — with a target of 93 percent in 2000. More rigorous attention to treatment standards has led to improved mortality and morbidity rates. The VHA has also improved preventive care, including inoculations, and has reported sharp increases in the percentage of inpatients and outpatients rating the agency's service as very good or excellent.[11]
- The Food and Drug Administration, an early pilot for GPRA, has improved the speed and efficiency of its drug review process, exceeding its own performance goals.[12]

Success stories notwithstanding, government-wide performance is still uneven, and complete and systematic information on real results — actual performance compared with plan targets — will not be available for all agencies until March 31 of 2000. However, at this stage of GPRA implementation, it seems fair to conclude that the development of performance management systems in federal agencies has proceeded much more successfully than many students of management reform had predicted, and that GPRA has produced widespread improvements in strategic planning and performance measurement throughout the federal government. This performance management infrastructure will improve accountability to Congress and key government constituencies, including citizens and the press, and these plans and data may set the stage for more widespread improvements in services in years to come.

Threats to Further Progress

Over the next few years, Results Act proponents must surmount several hurdles if progress is to continue. Some of those obstacles will appear in the next two years. Leadership transitions create risks for management reform, and GPRA may stumble when a new president takes office. Other

9 The Social Security Administration's performance plan is on the Web at http://www.ssa.gov/ob/00appfin.htm.

10 The Annual Performance Plan for the Department of Commerce, which includes NOAA, can be found on the web at http://www.doc.gov/bmi/budget/PB2000/apptoc.htm.

11 The Veterans Administration Performance Plan is available on-line at http://www.va.gov/Partners/stratcap/index.htm.

12 See http://www.fda.gov/ope/FY00plan/intro00.htm for the FDA Performance Plan.

problems will arise as Congress and the executive branch grapple with the less tractable problems of performance-based management and budgeting that have been postponed to the end of the long implementation process. In state and local governments, performance management systems have often lost their impetus when they encountered these difficulties, and GPRA implementation may falter, too. Specifically, GPRA will face challenges as it encounters several problems:

- Transitions in elected and appointed management.
- Building performance management into the day-to-day routines of management and budgeting in the executive branch.
- Making GPRA work for Congress as it oversees federal programs and makes budget decisions.

If GPRA implementation stays on track during this next phase, it will still face risks and challenges over the longer term. These include:

- Inflated expectations about how much improvement performance management alone can produce.
- Over-emphasis on technical measurement.
- Management of the performance of independent agents delivering services, especially state and local governments.
- Creation of a system of incentives powerful enough to sustain interest in improving performance in federal programs.

Leadership Transitions and Management Reform

In January 2001, a new president will take office, accompanied by a fresh cadre of political executives. If history is any guide, a new president may take some time to recognize the importance of management skills and experience in his appointments. In this transition, management reform may lose its standing as a priority, and GPRA implementation and related tasks may be relegated to routine status. It is even possible that the incoming administration, wielding the proverbial "new broom," will actively reject initiatives favored by the Clinton administration.

More generally, in Washington and in many statehouses, a significant threat to sustained reform has been the revolving door of leadership in America's governments. Especially at the national level, but also in large states and localities, top leaders — including appointed officials — tend to serve for relatively short periods of time, typically two to three years. When new leaders arrive on the scene, they naturally try to put their own stamp on government. One way to do that is to create new management systems and procedures and jettison the ones established by their predecessors. As a result, management reforms in too many cases are evanescent, and performance management systems in particular have been short-lived in statehouses as governors come and go. GPRA's statutory underpinning will help keep it alive through the next change of administration, but cannot guarantee that it will receive the same political backing as during the Clinton administration.

In March, when federal agencies report their results in meeting the GPRA performance goals, those results could turn into campaign issues. Indeed, in state and local governments, the political risks associated with performance shortfalls have sometimes undermined performance management. For this reason, some fear that the timing of these reports may force performance management too quickly and too high onto the political agenda, recreating the "gotcha" atmosphere of a few years ago. Others argue that political visibility for the Results Act will be a plus for reform, ratcheting up interest in government performance. After all, reformers have complained for years about how difficult it is to get management issues onto the political agenda. GAO's close attention to GPRA reporting should prevent significant political interference with accurate reporting in the short term, although agencies may be slow to report new data series if they show performance shortfalls. However, campaign disputes about performance in the 2000 election might provide incentives for political executives to try to insulate the president from political damage in 2004 by undermining performance measurement.

A Simple Typology of Performance Measures

- *Inputs*: Measures of dollars or staff used in programs.
- *Outputs*: Measures of process activities or program activities.
- *Outcomes*: Measures of changes in conditions produced by a program. *Intermediate outcomes* are usually shorter term and easier to measure than *end outcomes*.

See Appendix C on page 43 for more on this topic.

Building GPRA into the Routine Work of the Executive Branch

As required by GPRA, agencies have been producing strategic plans and targets, but these reports and data are often the product of high-level planning staff and may not be affecting day-to-day operations. Officials at the federal Office of Management and Budget (OMB) say that building GPRA into the daily routines of the executive branch is a top priority. This emphasis is reflected in the most recent version of OMB's Circular A-11 — the annual document issued by the budget office to give guidance to agencies on key budgetary and management issues. Even this strong bureaucratic push, however, may be insufficient to bridge some of the gaps between planners and managers that have arisen as the operating agencies implement the Results Act.

Performance Management in the Operating Agencies

The dilemma for federal managers is one that is familiar to officials at other levels. Top officials can sketch out far-reaching strategic goals, and management analysts can construct high-level outcome measures to track those goals, but the goals and measures may end up divorced from agency operations.

Evidence of this problem appeared at the state level as Oregon established its heralded Oregon Benchmarks.[13] The Benchmarks are a set of social indicators and targets that set out broad strategic goals for the state. As agency managers tried to use those indicators to run state agencies and programs, however, they often found it difficult to connect those long-term measures with their day to day work. To use the jargon of performance measurement, the Oregon Benchmarks focus on *end outcomes* — the social, economic, and environmental changes that government hopes to produce over many years. The end outcomes highlighted in the Benchmarks include: cutting crime rates, infant mortality rates, and the number of Oregonians with incomes below the poverty rate; increasing the number of adults who have completed college; and raising air and water quality standards. While those outcomes are unquestionably important, many social and economic factors contribute to success or failure in reaching these targets, and state government's role in achieving those goals is limited.

In contrast, operating managers trying to improve performance in their agencies typically focus their attention on measuring outputs, including key *process measures* of agency bureaucratic activities and the more immediate *outputs* produced by those processes. For a job-training agency, process indicators might include caseload numbers and counts of agency contacts with job providers. Outputs might include numbers of clients who graduate from training programs. Agency managers may also be able to measure *intermediate outcomes* that might stem from those outputs, like job placements for clients. When operating managers have not been included in the development of outcome indicators, managerial measures of output and intermediate outcomes may not link easily to the high-level outcome indicators — unemployment rates, for example — developed by planners.

Federal agencies face similar problems. Top managers and their planning offices are developing high-level strategic plans and devising indicators of end outcomes, often with little input from operating managers. The emphasis on end outcomes has been reinforced by the GAO, which continues to push hard for far-reaching outcome-oriented measures. Front-line managers, however, are still struggling to climb up the performance measurement ladder from inputs and process indicators to output measures and intermediate outcomes. The Forest Service provides a useful example. In the most recent revision of the agency's strategic plan,[14] the Forest Service states its mission:

> The mission of the USDA Forest Service is to sustain the health, diversity, and productivity of the Nation's forests and grasslands to meet the needs of present and future generations.

To carry out that mission, the agency proposes four general or strategic goals. The first two elaborate on the Forest Service's motto of "caring for the land and serving people":

- ❖ Goal I: Ecosystem Health — promote ecosystem health and conservation using a collaborative approach to sustain the Nation's forests, rangelands and watersheds.
- ❖ Goal II: Multiple Benefits to People — provide a variety of uses, values, products, and services for present and future generations for present and future generations by managing within the capability of sustainable ecosystems.

---

13 The Oregon Benchmarks are on the Web at http://www.econ.state.or.us/opb/benchmark_tables/index.htm.

14 See Department of Agriculture 1999.

Goals III and IV commit the agency to developing the best possible scientific underpinning for its mission, and to improving its "corporate infrastructure," with the aim of delivering services more efficiently.

So far, however, the Forest Service has found it difficult to define ecosystem health and to develop baseline data about the current condition of its forests and other properties.[15] In the forests themselves, agency personnel continue to cut timber, put out fires, maintain roads, assist visitors, and otherwise perform their day-to-day work. Supervising foresters track these activities with straightforward output measures. However, without clear and workable measures of ecosystem health, the foresters who perform the agency's daily activities find it difficult to understand how and whether their work is achieving the agency's broader strategic outcomes described in the agency's statement of goals. This problem is endemic in many federal agencies, where top-down outcome measures established by planners do not tie to the more mundane output measures that lower-level managers use every day.

In theory, the tools exist to solve this problem. As the agency develops its strategic plan and settles on its most important goals, it should lay out a detailed *logic model* linking inputs, outputs and intermediate outcomes under the control of agency managers. This model should also show the link between those intermediate outcomes and the higher-level end outcomes established as strategic goals for the agency.[16] When all levels of an operating agency understand these relationships, the performance management system becomes a vehicle for goal-setting and communication. Top managers develop what Senator Moynihan calls "walking around numbers" — a few key indicators that help them track their agency's progress. At the same time, lower-level managers and workers come to understand the agency's strategic goals and develop a clearer sense of how their own work contributes toward reaching them. With better internal operating information for top managers and strategic direction for workers, the agency has presumably laid a solid foundation for improving results.

A few federal agencies are held up as models of this theory in action. At the Veterans Administration, for instance, the head of Veterans Health has developed a well-articulated strategic plan and goals and a consistent set of clear performance measures. (See Appendix D on page 49 for a summary table of the VHA's performance goals and measures.) Those measures are used to write performance contracts with managers, and through that device, as well as other communication channels, those goals and measures cascade down through the whole organization.

Despite those successes, experience at the state and local level suggests that the problem of linkage is not easily resolved. Almost by definition, high-level outcome measures track social changes that are influenced by factors that are not under the direct control of operating agencies. In theory, the logic model will again provide the answer, describing the external factors besides agency action that will affect outcomes and outlining hypotheses about the strength of those effects. In practice, however, getting past this barrier will be difficult, and will require close involvement of operating managers, and surmounting what might be called the "influence test."

15 For an in-depth critique of the Forest Service's accountability efforts, see GAO October 1999.

16 See Hatry 1999, chapters 5 and 6, for a helpful discussion of the role of logic models in developing performance measures.

Passing the "Influence Test"

To make a performance management system work, goals have to be measurable. They also have to be goals that the entity being assessed can appreciably influence. A good example is the Head Start program, which provides developmental childcare for poor children. Political leaders can reasonably ask for data about outputs from Head Start programs — how many children participate in Head Start, their age and other characteristics, and the duration of their participation, for example — and managers may be willing to be evaluated on their progress in meeting goals tied to those output measures.

Under the Results Act, however, programs are supposed to develop measures of outcomes, in addition to outputs. The outcomes that Head Start is supposed to achieve are quite far-reaching, and at their broadest could simply be described as improved well-being of participating children. While no one doubts that this goal is important, thoughtful officials may find it difficult to agree on its definition. Nor, of course, is Head Start the only agent of improvement for children. Other factors contribute to their well-being, including their families, churches, and other community organizations, not to mention the economy and other government programs.

Managers quite naturally prefer to be held accountable only for results that are within their control. If they are being evaluated by a benchmark that they cannot appreciably influence, they are likely to resist using it and fall back on lower-level outputs or intermediate outcomes as more appropriate measures of their success. A thoughtful response to this problem can be seen in the measures and reports devised by the performance measurement consortium organized by the Urban Institute and the International City/County Managers Association (ICMA).[17] Several dozen local governments, with populations over 200,000, have banded together to develop common measures for performance, with results scheduled for annual publication. The consortium's performance reports include *explanatory data* explicitly discussing factors other than government performance that can influence the reported results.

For example, these localities publish data on crime rates — an important end outcome of police work. To help readers understand the full range of factors affecting crime rates, the performance reports analyze external variables that have been shown to have a significant impact on them, including median household income and the percent of housing that is owner-occupied. The goal is to help readers understand that police staffing and strategy are not the only influences on crime rates.

Federal agencies share similar problems as they move toward measuring end outcomes. The Social Security Administration may report on its performance in answering phones, processing checks, and resolving complaints, and those outputs are undoubtedly important to recipients. More broadly, however, the Social Security program is also one important instrument in reducing poverty among the aged and disabled. But success or failure in meeting that broader objective depends not only on the Social Security program but also on the economy, the success of other pension systems, and the tax codes of the federal as well as state and local governments. SSA officials should be willing to be evaluated in part on their role reducing poverty among target populations, but they can fairly request that those evaluations include explanations of other variables that can be important in

17 Urban Institute/ICMA 1998.

achieving those outcomes, so the public can understand that the link between the Social Security Administration and achievement of that of broad goal is only partial.

While managers may worry about being evaluated using measures that do not pass the "influence test," elected officials may have a contrary fear. In devising new programs or marshaling support for existing ones, elected officials may feel obliged to exaggerate the impact those programs can have on deep-seated social problems. When broader forces, especially economic growth, improve social and economic conditions, program champions want to take as much credit as possible for the role of their own initiatives in achieving those successes. In tougher times, it might be argued, elected executives and legislators may be grateful for a realistic understanding of what those programs can and cannot accomplish. In good times, however, elected officials — anxious to take credit for good results — do not welcome analysis that emphasizes the role of external forces, as opposed to program activities, in achieving successes. Citizens and the media should find it easier to make their own judgments about program impact if performance data is coupled with better explanatory data, and both are buttressed by additional findings from evaluation studies.

Incorporating Performance Measures in Budgeting

Performance budgeting is an approach to resource allocation that seeks to use the power of the budget process to reinforce performance management. Performance budgeting also focuses on linkage — in this context, between resources and results. Some have hoped — or feared — that GPRA would put budgeting on "automatic pilot," to use a phrase ascribed to one senator who helped enact the Results Act. With a formulaic linkage between performance and resources, elected officials could simply reward high-performing programs with more funding and cut back on programs that were not meeting their goals.

States have experimented with formulas to link performance and budget allocations. The most extensive application has been in higher education, where sixteen states now distribute modest portions of funding for their public universities using performance-based formulas. In those jurisdictions, between one half to six percent of state support is allocated based on the success of a campus in meeting performance targets, which may include graduation and retention rates, job placements, and student and alumni satisfaction rates. While the use of formula budgeting for higher education has been spreading, it has also been controversial, and six states have abandoned the effort since 1996.[18] More recently, a few states have begun to experiment with formulas which distribute extra funding to elementary and secondary schools based on performance standards.

In the case of Congress, however, budgetary allocations reflect the preferences of elected officials and their constituents. It is difficult to see how performance-based formula budgeting could play a significant role in this process. Would Congress be comfortable with funding formulas that automatically cut poor-performing programs, and automatically raised budgets for high performance? Consider a poor-performing program with strong political support. The likely outcome is more funding and a management push to fix the program — not cuts to cripple it. We return to this question in discussing the risks of overselling performance measurement to the federal government.

18 Burke 1999.

At the moment, then, the "automatic pilot" model is of theoretical interest only. The federal debate about linking performance and resource allocations focuses on more mundane goals — especially that of restructuring budget accounts to tie spending totals more directly to the strategic goals and performance measures established over the last six years. The Results Act specifies that the measures in an agency performance plan must link directly to the budget account structure used in the president's budget. The goal is to encourage decisionmakers to consider resources and results at the same time, and to help agencies show the impact of increases or cuts in funding on increases or cuts in services.

The Results Act required that the Office of Management and Budget initiate performance budgeting pilots in five agencies for fiscal years 1998 and 1999. Agencies participating in those pilots must develop presentations showing specifically how increases or cuts in funding would affect outcome measures of performance in one or more of the major functions of the agency. According to the Act, the president's budget for fiscal year 1999 would include these pilot performance budgets as alternative presentations. However, no agencies volunteered for the pilots, and OMB postponed this deadline. The director of OMB is also scheduled to report to the president and Congress by March of 2001 on the feasibility and advisability of including a broader performance budget as part of the president's budget.

Restructuring Budget Accounts

Agency budgets are divided into accounts and then into program activities. The Results Act requires an agency's annual performance plan to link directly to the accounts and activities presented in the president's budget.

The most straightforward linkage uses one or more performance measures for each activity. GAO chose the budget of the Health Resources and Services Administration as an illustration:

Activity		Performance Goal
AIDS emergency relief grants	ά	Increase number of clients served by 7.5 percent + 3 other performance goals

For illustration of more complex linkages, see GAO Report GAO/T-AIMD/GGD-99-67, April 1999.

In a recent report to Congress, GAO identified 14 agencies that had developed indicators and account structures that allowed them to relate resources to results, and suggested that the approaches used by these agencies could be useful building blocks for the performance budgeting pilots required by GPRA. The president's FY 2001 budget did earmark agencies and programs for pilot treatment, and renewed the promise to report on them by March of 2001.

Restructuring budget accounts is a task of some technical complexity, made all the more difficult by a lack of interest from appropriators in Congress, who show little enthusiasm for changing familiar and long-established account structures. While careful reconfiguration of account structures can make it easier for decisionmakers to understand the linkage between budget allocations and results, revised account structures are only one of the steps required to develop budgets that relate resources to results. Detailed cost analysis must also be carried out before budget analysts can show decisionmakers how different levels of resources will affect performance levels.

Cost accounting is well developed in the private sector, where it gives corporate executives the information they need to set prices and make decisions regarding plant investments and modifications. In contrast, budget officials in most U.S. government agencies focus most of their attention on analyzing and controlling levels of expenditures, not on scrutinizing the cost factors underlying those expenditures. This problem is perhaps more acute in budget offices like the OMB, which pride themselves on their role in formulating policy. In its initial stages, cost accounting is anything but policy relevant. Long and painstaking efforts to collect data are required before the information is of interest to the policy-oriented budget office, and the experts in cost analysis are accountants, not budget professionals. However, budget professionals will be called upon to unravel difficult questions about appropriate allocation methodologies for capital, support services, retirement benefits, and other overhead costs. The federal government committed itself to development of reliable cost information in the Chief Financial Officer (CFO) Act of 1990, and the Federal Accounting Standards Board (FASAB) has recommended standards for agency cost-accounting systems. However, little progress has been made in implementation of those systems.[19] Without cost information, the linkage between budget allocations and results cannot be established with confidence.

Thoughtful budget officials go further to question the fundamental assumption underlying the performance budgeting model, which holds that cost analysis and performance measurement can establish and quantify the relationship between changes in resources and changes in outcomes. As discussed earlier, end outcomes are affected by other factors besides agency activities and resources. The more difficult the social problem addressed by a program, skeptics would argue, the less likely it is that performance measurement and cost accounting can establish a linear relationship between resources and results. Moreover, as the earlier discussion of Head Start indicates, results may not appear quickly in many programs, and expectations that changes in outcomes will be possible in the near term may be misguided. In those circumstances, data from high-powered and expensive evaluation studies may be a necessary complement to performance data and cost analysis before sound judgments can be made about the impacts and benefits of such programs.

Even if performance data and the findings of evaluation research are more easily available, presidents and their staff will still make budget decisions based in large part on politics and programmatic preferences. When presidential candidates run for office, they outline their programmatic views, and those preferences are an important element in voters' choices. The president's budget puts forth that program in concrete financial terms. Performance data may help refine and modify those preferences, but it will not — and should not — supplant a president's political and programmatic priorities.

Making GPRA Work for Congress

As suggested, in its early stages of implementation the Results Act presented an irresistible opportunity for partisan skirmishing. Undergirded by GAO's analytic work, congressional staff and leaders have now stepped back from that bickering, and GPRA's proponents are struggling to find new ways to make performance reports and data provided under the Results Act politically real and relevant for members of Congress. Some congressional staffers have followed GPRA

19 Ibid., pp. 7-8.

implementation closely. They confess, however, that without encouragement and greater interest from members, staff will have few incentives to spend their time on performance management and measurement. And without congressional interest, it will be difficult to sustain enthusiasm for performance management in operating agencies of the executive branch.

In theory, Results Act data should be valuable to Congress as it conducts two of its key functions — oversight and budgeting. According to staffers, members of Congress are beginning to make modest use of annual performance plans (APPs) in oversight. Several authorizing committees, for instance, conducted hearings on draft strategic plans in 1997, and a few members have begun to key their questions directly to annual performance plans, focusing on the three big issues that the plans are designed to address:

- ❖ What is this agency trying to achieve over the next year?
- ❖ How is it going to achieve it?
- ❖ Will the agency be able to provide performance data that will tell Congress how well the agency is performing?

However, the oversight plans formulated at the beginning of every session of Congress do not yet routinely include review of the APPs or other GPRA data.

Members of Congress complain about the problem of overseeing cross-cutting programs — separate programs, often in different cabinet departments, aimed at similar goals and constituencies. In theory, performance data in GPRA reports should be helpful in monitoring those programs, and some committees are beginning to use those data for this purpose. For example, the committees looking at land management agencies have begun using GPRA data to develop some cross-cutting questions, reflecting the obvious reality that ecosystems do not fall easily within the jurisdictional boundaries of agencies and subcommittees. Those committees are asking, for example, how the Forest Service, the Bureau of Land Management, and the National Parks Service are working together to develop a consistent set of goals for properties that are sometimes contiguous. Until the agencies themselves develop compatible indicators and goals, however, performance data are going to be of limited value for oversight of cross-cutting programs.

The release of performance reports in March of 2000 may spark new interest in the use of these data for monitoring programs. So far, however, members of Congress have made limited use of GPRA data in their oversight activities.

If congressional use of Results Act reports for oversight is spotty, staffers and other close observers of Congress are even less optimistic about the use of performance data in budgeting. Right now, budget discussions in Congress reflect strong partisan disagreements between the executive and legislative branches on basic priorities, not finely honed questions about agency performance. Even in review of programs that are not salient to these partisan battles, budget debate has been dominated by the constraints of budget caps, leaving little room to discuss — much less reward — performance.

At a more technical level, appropriators have shown little willingness to revise account structures. Often an agency's basic budget structure reflects long years of haggling with interest groups about expenditure categories, and changes would upset those constituencies. Members of Congress and their staff are also familiar with existing account structures, and changes would add to their workload. OMB has asked for more help from Congress on this point, stating in the FY 2000 Government-Wide Performance Report that:

> An obstacle that has been identified by many, including the Chief Financial Officers, is the current structure of budgetary accounts. We believe that a dialogue with the Appropriations Committees about the alignment of results with appropriations would make performance plans more useful in connecting resources to results.[20]

Some appropriators have expressed a willingness to look at what budgeteers call "cross-walks" — complex tables showing the relationship between old and new proposed budget structures. There is no certainty, however, that use of those cross-walks will persuade congressional staffers and members that the rewards for restructuring accounts outweigh the additional work required to make those changes.

Even when appropriators are willing to consider revising and simplifying budget structures, they are not certain that performance data are sufficiently well developed to replace the detailed financial data they now use for oversight. Early in 2000, the House Subcommittee on Interior and Related Agencies began to review a proposal for a simplified account structure for the Forest Service. Recent studies by GAO and the National Academy of Public Administration had found that the current account structure was not providing useful information about the work actually performed by foresters. In an artificial effort to show how the goal of multiple uses was reflected in day-to-day activities, individual projects were charged to many different accounts without backup about the logic of those charges. Subcommittee members did use output data — timber sales, for example — for oversight and seemed less troubled than GAO by the difficulty presented in linking those data to the agency's strategic goals. However, some subcommittee members questioned whether even that output data was detailed and reliable enough to meet their needs.

Federal agencies are in a quandary. GAO reports press them to develop new account structures that link to performance data. Without hope that Congress will approve budget restructuring, however, agencies are not likely to maintain two sets of account structures for long. And without revised accounts, long-term progress in performance budgeting will be modest. Members of Congress, like presidents, may reasonably believe that their political and programmatic preferences should outweigh performance data in budget decisions. For their part, GPRA proponents have not shown elected officials convincing evidence that results-oriented data can be a useful tool in effectuating those political and programmatic priorities.

Overselling Performance Management: The Several Dimensions of Management Reform

As Congress debated the Results Act, its proponents speculated that its enactment would spark far-reaching transformations in management practice and improvements in government performance. GPRA has given managers in federal agencies new standing and new tools, and

20 OMB 1999, p. 2

created a performance management infrastructure that will help citizens and legislators track improvements in government services. However, the Results Act has not transformed service delivery in federal agencies. While excessive claims for GPRA are less frequently heard now, it is still the case that performance management proponents sometimes overstate the ability of performance management and measurement systems to improve management in and of themselves.

Performance measurement is a necessary but by no means sufficient condition for producing far-reaching management improvement. Demonstrating this point, the Government Performance Project (GPP) at the Maxwell School uses a research design that aims to encompass several key dimensions of management, not just performance management. The GPP is measuring management capacity broadly, and managing for results is but one dimension. Other dimensions include managing an organization's finances, its information technology, and its capital assets, and developing its human resources. This multifaceted model implies that significant improvement in management is not likely without attention to all five major dimensions, and initiatives to improve performance must address all of these factors in an integrated fashion.

It is also clear that leadership is required to change organizations. Leadership adds energy, focus, and commitment to technical management capacity, and that combination is essential to producing major changes in complex organizations. The development and maintenance of managerial capacity, including performance information, is insufficient to produce organizational change on its own.

In short, GPRA cannot transform government by itself. Performance management needs to be accompanied by far-reaching initiatives in other key areas, and this broadened management reform agenda must be championed by energetic and skilled management leaders in key positions throughout the executive branch.

If performance management cannot singlehandedly transform management in federal agencies, neither is performance information by itself a sufficient basis for allocating resources and making decisions on many programs. The linkages between inputs, outputs, and outcomes in many federal programs are too complex to be captured by performance measures alone. More sophisticated evaluation studies will be needed to assess the impact of those programs, analyze their benefits, and create appropriate incentives for program managers and recipients.[21]

However, none of these caveats mean to suggest that performance measurement and management is valueless. GPRA has given managers in federal agencies new standing and new tools, and created plans, targets, and data that will help citizens and legislators track improvements in government services. As several years of performance data become available, agency managers should find it difficult to duck informed questions about the magnitude and speed of managerial improvement in key programs. And performance data can be combined with the findings of evaluation research studies to produce indicators with greater explanatory power.

21 See Burt S. Barnow's discussion of this problem using the Job Training Partnership Act as an example (Barnow 2000).

Avoiding Measurement Myopia

For more than three decades a subgroup of management specialists has been writing about the technical aspects of performance measurement. This valuable work has been recently summarized by Harry Hatry, a pioneer in this field, in *Performance Measurement: Getting Results*.[22] Other analysts have also provided provocative treatment of those technical issues; two recent examples are *Organizational Report Cards*, by William Gormley and David Weimer, and *The Balanced Scorecard*, by Robert Kaplan and David Norton.[23] Kaplan and Norton's work, originally aimed at corporations, has been especially influential in federal management circles, serving as a staple for training sessions at NAPA's Center for Improving Performance in Government. Within the federal government, analysts at the General Accounting Office continue to evaluate agency reports and performance measures, and GAO's technical work on agency account structures has helped keep the idea of performance budgeting from slipping off the management agenda altogether.

This technical work has allowed managers working with performance data to speak with a common vocabulary. Consensus on terminology may seem unremarkable, until we look closely at programs where managers and clients cannot agree on terms and concepts — in the case of Empowerment Zones, for instance. In that federal initiative, local organizations found themselves facing a shifting set of reporting and performance requirements, repackaged in changing vocabulary as the program was implemented.

The three basic types of performance measures are *inputs, outputs*, and *outcomes*. Many analysts go further to divide outputs into process measures and programmatic outputs, and outcomes into intermediate and end outcomes, as outlined earlier. The idea that governmental agencies and their contractors can be held accountable for results can involve either of the second two kinds of measures outputs or outcomes, but not the first, inputs.[24] A focus on measuring inputs is a symptom of the old "bureaucratic paradigm," which focused on process and regulations, not performance.

Performance management specialists have gone further, combining input measures with output or outcome measures to produce what are sometimes called unit-cost ratios. Dividing inputs by outputs or outcomes produces *efficiency measures* — cost per high school graduate, for example. Dividing outputs or outcomes by inputs generates *productivity measures*, such as meals served/school lunchroom worker. To link performance measurement systems to budgets, additional data from cost analysis must be factored in.

As a performance management system grows, support from other technical disciplines, especially information technology, becomes more important. Indeed, the explosion in computing power has helped build a stronger technical base for performance management and measurement, allowing easier storage of and access to large amounts of data, as well quicker and more precise analysis. Finally, for complex programs, performance data can be supplemented by the findings from evaluation studies, allowing more sophisticated inferences about program impacts and benefits.

---

22 Hatry 1999.

23 Gormley and Weimer 1999. Kaplan and Norton 1996.

24 See Appendix C for a fuller discussion of types of performance measures.

While a clear understanding of the vocabulary, conceptual framework, and technical underpinnings of performance management and measurement is important, a focus on such issues can bog down GPRA implementation in a sterile debate about measurement, and prompt managers to lose sight of their operational objectives. Key management priorities may include expanding programs of critical strategic importance, or improving areas with serious management problems. For example, GAO has complained that agency performance plans demonstrate technical improvements in the design of strategic plans and performance measures, but do not consistently focus on "mission-critical management challenges and program risks."[25] The latter include 300 areas where federal operations are at high risk for waste, fraud and mismanagement. Don Kettl, a professor at the University of Wisconsin and a member of the Task Force on Performance Management and Measurement, has stated the problem more broadly:

> The biggest mistake in pursuing performance-based management is to conceive of performance as primarily a *measurement* problem. It would be deceptively easy to allow government performance to degenerate into a process-based, numbers-driven exercise. In fact, that is precisely what undermined previous federal experiments with tactics like the planning-programming-budgeting system, management by objectives, and zero-based budgeting.[26]

In short, technical improvements in performance measures do not guarantee that top management will spend its time and resources on the most important targets for improvement.

In some state and local governments, performance management systems have drifted away from their management focus, concentrating on technical measurement issues instead of results. They have also extended their reach too widely, trying to measure everything that occurs in government, burying information about strategic services and their improvement in large volumes of performance indicators. The impulse to measure everything is understandable. Elected officials and political executives may find it politically difficult to tell constituencies that their favorite program is not strategically important enough to warrant measurement, and the Results Act requires agencies to develop measures for every program activity. The tendency to measure everything has also been reinforced by squadrons of consultants now working on implementing GPRA in the agencies. These so-called Beltway bandits have created a huge industry based on GPRA, and will continue to try to extend that line of business.

It may not seem immediately obvious why measuring too many activities can become an impediment to strategic management. But as long as managerial talent is in short supply in government, top officials will have to make decisions about where to focus managerial resources. New York City is a case in point. Shortly after the fiscal crisis in 1975, the newly created Mayor's Office of Operations set out to develop performance measures for all city agencies and services. The result is the *Mayor's Management Report*, a thick compendium of performance data published twice a year. Initially, those data were a stimulus to debate about the quality and direction of city services, but by now these data are produced so routinely that they serve little function in managing strategically.[27]

25 GAO July 20, 1999, P. 3.

26 Kettl 1998, p. 47.

27 Smith 1993.

In an effort to increase the salience of performance data, some successful managers have developed their own separate performance management systems apart from the *Mayor's Management Report*. An important example is the COMSTAT system in the New York Police Department (NYPD), which has served as a model for other operating agencies in New York City and for police departments around the country. Precinct commanders and top NYPD officials meet monthly to review COMSTAT data on criminal activity and police response, and major shifts of resources and police tactics result from these discussions.[28] Support from the mayor and leadership by the police commissioner were major factors in developing this high-intensity performance management approach.

In some other state and local sectors, notably elementary and secondary education, we see signs that the drift toward measurement myopia may be reversible. For decades, schools and state departments of education have collected massive amounts of data, but have failed to use that information successfully to change schools. In many parts of the county, state governments and others are pushing a new emphasis on testing as an accountability measure so that parents, taxpayers, and citizens can assess the results produced by this crucial public enterprise. And in more and more jurisdictions, some rewards — higher pay, larger budgets — are linked to those measures.

Developing sound and reliable tests is difficult, and unraveling the relationship of test scores to social and economic factors still more complex. In some states — notably Tennessee — efforts are underway to refine performance measures by adjusting test scores for external factors. Those adjustments are aimed at allowing citizens and educators to make more sophisticated comparisons between performance in schools with a large number of low-income children from multi-problem neighborhoods and schools in wealthy suburbs where strong parental and community support for education are widespread. The researchers who develop these measurement schemes are attempting to isolate and quantify the impact of the schools, apart from the impact of other factors. However, testing is only one part of an effective management system for schools, and tests alone are not likely to be as useful for management as they are for research.

Managing Performance of Independent Agents

As the Results Act has been implemented, the tendency in Washington has been to concentrate performance management and performance measurement efforts on *directly provided* public services, a small and diminishing subset of federal programs. Not surprisingly, therefore, successes in performance management have clustered in those few federal agencies providing direct services. In its National Performance Review (NPR), for example, the Clinton administration has focused most of its high-level management efforts on what the administration calls *high-impact agencies*. These are the 32 agencies with the most direct involvement and contact with citizens, including the Social Security Administration (SSA), the Immigration and Naturalization Service (INS), the passport office of the Department of State, and the Postal Service. NPR's goal has been improvements in customer service goals for these 32 agencies, reflecting the expectation that such improvements will help boost citizens' approval of government as a whole. Although the first

28 Smith 1997.

publication of performance data for these and other federal agencies is due in spring of 2000, results from surveys of citizen "customers" of the high-impact agencies began to appear in the fall of 1999.

While some progress in measuring performance has been achieved in agencies that deliver direct services, the federal government has been slower to show similar gains in managing performance in cross-cutting programs that share responsibility for service delivery. One positive example is drug policy. The Clinton administration has stepped up efforts to coordinate and improve programs aimed at controlling illegal drugs, and the Office of National Drug Control Policy (ONDCP) manages this effort, but delivers no services itself. ONDCP annually publishes what it calls *Performance Measures of Effectiveness: A System for Assessing the Performance of the National Drug Control Strategy*. These measures are extensive, and some are even directed at foreign governments, establishing goals for cutting drug production or increasing interdiction overseas. The agency's performance management tools have been supplemented by other formal powers to coordinate the elements of drug control policy, most notably the ability to veto an agency's budget.

Managing performance by foreign nations is not a typical problem, but the federal government does regularly find itself setting performance goals for services that are delivered by private or not-for-profit contractors. The difficulties the federal government faces in improving performance grow as it becomes more and more reliant on contractors to deliver services. Operating outside of the immediate constraints of the bureaucracy, the roles of these agents — who they are and what they do — are highly fluid. Indeed, elected officials sometimes decide to use nongovernmental agents to deliver programs precisely because they are seen as more flexible and responsive than traditional bureaucracies. However, those very characteristics increase the difficulty of collecting timely, accurate, and verifiable performance data, as the contractors have different information systems and management approaches.

We can see this problem more clearly by returning to our discussion of the Head Start program. Head Start programs are conducted under contracts, mostly with nonprofit organizations but also with private firms. In theory, service delivery by non-governmental organizations could actually facilitate performance management because managers in government might be able to devise real performance incentives for contractors and replace them if they do not perform. The hard question is not so much one of theory as of practice. Politicians play politics with contracts, blunting incentives for high performance. And Head Start advocates have resisted long term evaluations of the program. But even where there is a commitment to excellence, it is much harder to measure the results of third-party action than when governments themselves directly control some activity. Research has repeatedly shown that the "nonprofitization" of public services has made it extremely hard to get accurate, timely, and consistent data to assess and reward performance.

These problems are exacerbated when the agents are state and local governments, proud of their independent status within the federal system and run by their own elected leadership. These issues pose another significant barrier to management reform: efforts to achieve results cannot be undertaken simultaneously at every level of government — they take time to design and implement. Moreover, elected officials and political executives must agree on goals before performance management systems can be designed to reflect the disparate aims of the parties that provide services in the complex U.S. intergovernmental system. Many of the governments and nongovernmental agencies involved in those programs disagree on who should be held accountable for achieving desired results, and how those results should be measured. Finally, the programs that are devolved to state and local governments often aim to achieve far-reaching outcomes, but their impact on those end outcomes often can be difficult to assess without detailed evaluation studies.

Welfare reform demonstrates the full range of difficulties encountered in managing and measuring performance in the federal system. Politicians often advance welfare reforms with the aim of transmitting values and changing attitudes — for example, emphasizing work, reducing long-term dependency, and discouraging teen-age pregnancy and out-of-wedlock births. One familiar problem in managing performance in this area involves understanding and isolating the impact of government in achieving those aims. One might think that the goals of welfare reformers could be measured by the number of single-parent welfare recipients who have jobs or the number of teenagers who do not have more children, but on examination such tests fall short. Many factors, ranging from AIDS education to improved contraception, affect teen birth rates. And people's decision to work is anything but simple. If moving from welfare to work is the test used, how can we know whether people would have gone to work anyway? Lots of people do. Did welfare reform produce this outcome? Moreover, if work is to be the test, how does one measure the way the *signals* transmitted by welfare policies discourage or prevent some people from coming forward for benefits and encourage them to work instead? New workers may join the workforce precisely because of the value-changing "successes" of welfare reform laws that emphasize work, but we don't know who they are. When local governments or private agencies are the direct service providers, it is difficult to determine who was *diverted* or *deflected* from applying for welfare benefits in order to measure positive policy outcomes.

Moreover, the new block grant structure of welfare was specifically designed to encourage experimentation with different approaches to moving welfare recipients to work. While this allows state and local agencies to be more responsive to local conditions, it also complicates efforts at oversight and accountability. This shifting mix of tools and tactics also makes it difficult to determine which government actions are responsible for reductions in welfare rolls or increases in job placements.[29] Taken together, these factors create measurement problems of the highest order, and point to even more difficult management problems.

29 Gais 2000.

Rewarding Performance

Incentives for improving performance are an important part of any performance management system. The federal government has developed some non-financial incentives, including forums for special recognition for high-performing managers. Top federal managers may find recognition a significant incentive. In agencies like the Social Security Administration, for example, where customer service is demonstrably better than it was five years ago, managers are proud to lead a national benchmark organization and enjoy showing private-sector visitors, including executives from L.L. Bean and Southwest Airlines, how they manage and monitor their 800-number unit, based in Baltimore.

While these intangible rewards can help motivate top managers produce results, the federal government will also have to develop tangible incentives for high-performing individuals and organizations. Simply cycling through the second full round of strategic plans, annual performance plans, and performance reports may not be enough to sustain high levels of interest on the part of agencies and their employees. GPRA has helped get performance management in the federal government off the ground, but continued progress may require a second stage booster in the form of concrete incentives.

Devising a workable system of incentives is not easy. Private sector firms often use bonuses to link pay to performance, and they have years of experience grappling with the problems entailed. These include difficulties defining performance for organizational units and for individuals, as well as analytical problems in distinguishing organizational impacts on financial results from broader economic factors. In government, these issues are complicated by the difficulty of measuring the impact of agency programs on complex social outcomes.

With all these problems in mind, government human resource managers and decisionmakers may be daunted at the prospect of developing more powerful incentives for improved performance, but they should not abandon that effort. The federal pay system already allows high-performing individuals in the very top ranks of the bureaucracy to receive modest bonus payments, and the amounts that can be awarded through that mechanism have just been raised. A few of the school incentive programs mentioned earlier also provide financial rewards to teachers, and some local governments have worked out gain-sharing arrangements with unions and front-line workers who help achieve significant reductions in the cost of services.

Government, however, typically provides incentives to groups of workers and agencies in the form of additional appropriations or special recognition, not payments to individual workers. In part, this bias against financial rewards for individuals stems from union influence in the government workforce. But it also reflects the very real and difficult measurement issues faced by human resource specialists as they devise systems to parcel out individual financial incentives. Elementary and secondary education provides a case in point. Student test scores may successfully track school performance, but school-level testing data are not specific enough to determine which individual teachers should receive performance bonuses. More generally, economists argue that incentive pay programs can change behavior, but that those programs are most effective when workers have small number of easily measurable tasks. Teachers, like many government workers, do not fit that description.

Some have suggested that the incentives that can serve as the second-stage booster for performance management might be additional budget allocations for high-performing agencies and programs, not individual bonuses. Although strict performance-based budgeting formulas are unlikely to work in the federal government, some state and local governments have experimented with other kinds of budgetary incentives. These include:

- ❖ Access to a portion of any surpluses generated by an agency.
- ❖ Access to funds set aside for such purposes, such as Philadelphia's Productivity Bank.
- ❖ Additional funds for special administrative needs, like New York City's allocation to managers of $2,000 gift certificates to Staples for office supplies and equipment.

These kinds of financial incentives may be useful in federal agencies as well.

State and local governments have also given high-performing agencies regulatory relief, in the form of freedom from some budget and administrative controls. Such programs may give high-performing agencies more flexibility to hire without prior approvals by control agencies, or to reallocate funds between agency programs. Similar "flexibility waivers" were part of the original GPRA design, but have not been widely used as GPRA has been implemented.

The essence of this idea is neatly captured by Thomas J. Peters and Robert H. Waterman in their book *In Search of Excellence,* where they discuss what they call the "tight-loose" approach to managing private corporations. Peters and Waterman note that the most successful corporations are specific about goals and flexible about means. The idea is to give managers clear objectives, and then step aside and let them decide how they will carry out the steps required to meet those goals.[1] Similarly, the performance paradigm in government suggests that results-oriented leaders focus on outcomes while letting operations managers decide how to deploy resources and organize key business processes to produce those results. The Veterans Health Administration, which has won widespread praise for its success in results-oriented management, has used this approach to give its managers wide discretion on means, but hold them strictly accountable for results. To make flexibility incentives work, Congress will have to step back from its continuing tendency to micro-manage federal programs.

1 Peters and Waterman 1997.

Conclusions and Recommendations

Performance management is alive and well in the federal government. It has been most effective in programs where goals are clear, the link between outputs and outcomes is understandable, and responsibility for success does not depend on multi-agency coordination. However, the next steps of implementation present some formidable problems, and the upcoming presidential transition creates additional risks.

One element of the original plan outlined in the Results Act is behind schedule — the performance budgeting pilots. OMB's professionals argue that they are already taking account of performance in budget decisions, and doing so with more thought and sophistication than the mechanistic approach envisioned in the pilots. In retrospect, it seems clear that GPRA's designers, looking at overly optimistic reports from the states about their progress in performance budgeting, underestimated the difficulty entailed in linking resources to goals and performance. States as well as most large local governments, grappling with services and constituencies less complex than those in federal programs, have had great difficulty transforming their traditional budget processes into results-oriented budgeting. As outlined earlier in this paper, federal agencies must surmount several technical impediments, including restructuring budget account structures and collecting cost data on services, before performance budgeting can even be tested fairly.

Those problems are serious, but OMB is pushing ahead to find services in several agencies that can use and evaluate the performance budgeting model. Customer surveys in so-called high-impact agencies can probably help isolate key services to test the effect of different resource levels on performance, and accelerated initiatives should collect data on the cost structure of those services. These pilots may lead to a more informed debate about *where* performance budgeting might or might not work in the federal government, not *whether* it can be useful at all.

Scholars and program managers alike have expressed concern that many federal programs cannot be fairly assessed using performance measures alone, and thoughtful budget officials are similarly skeptical about the ability of performance measures to capture the full range of information about many federal programs. They are right, and the response to the problem should be clear-cut — performance data in those areas needs to be supplemented with evaluation studies. Agency heads and budget officials should identify areas where such studies would be most valuable. If possible, when evaluation findings are complete, program impact data can be used to supplement measures of gross outcomes.[1]

Human resources specialists in the Office of Personnel Management and in the agencies, working with OMB, should also craft a broader spectrum of incentives that can be used to reward high-performing individuals and organizational units in federal agencies. In addition to pay-for-performance, these incentives should include access to special pools of funding, and more aggressive grants of administrative flexibility to managers in bureaus and offices with records of performance improvements.

1 Barnow 2000.

Managing for results in the intergovernmental arena presents special difficulties. In implementing key programs like welfare reform, state and local governments often have interests and objectives that are very different from federal goals. In the search for performance measures that can encompass those different objectives, program overseers must sometimes resort to crude measures, like reductions in welfare rolls, that do little to illuminate the complexities of these programs. Additional study and thoughtful experimentation is needed in the design of federal performance partnerships with states and localities.

Top managers and overseers must also understand that performance management systems by themselves cannot produce significant improvements in service delivery. Performance targets may create incentives for improvement, but to reach those goals, leaders and managers must mount systematic initiatives to enhance management capacity for dealing with agency finances, information technology, capital assets, and human resources. These initiatives must be part of broader strategies to improve performance, and should be detailed in strategic plans and tracked with milestones and other qualitative performance measures.

These issues should be addressed in the framework of an updated and expanded agenda for improving management in the federal government, prepared for consideration by the next president. Continued progress on performance management by itself will not suffice.

All things considered, the challenge of managing for results in the federal government is formidable. In American politics, leaders must overpromise to have their goals accepted in laws and other policy pronouncements. These promises, in turn, make it more likely that results will fall short. Management reformers must also contend with important policymaking features of American government, including the diversity and complexity of the federal system, and the political nature of the decisionmaking and implementation processes. Even when the goals chosen are realistically selected and articulated, and even when the agents assigned responsibility for achieving them are appropriate, the political calendar is crowded, and elections distract politicians from management reform.

Many observers, buoyed by successes abroad, suggest that these and other performance management reforms can spawn more far-reaching changes in the underlying strategy of management improvement in the federal government. These changes entail several steps, sometimes summarized under the catch-phrase “let the managers manage”:

- Deregulate managers, eliminating rules and regulations imposed by control agencies like budget and personnel offices.
- Shift the focus of measurement from inputs to outcomes, and overall emphasis from control and compliance to freedom and flexibility.
- Complete the system by providing rewards to those who succeed and imposing sanctions on those who fail.

This perspective, pioneered in Australia, assumes that managers are competent and caring and have failed because they have been hog-tied by rules and regulations. This approach to reform might be labeled *managerialism*.

Others envision a different strategy, suggesting that government leaders "make managers manage," as the government of New Zealand has done. Under this model, top officials try to create competition by establishing market-like mechanisms, such as outsourcing some portion of service delivery to alternative providers. This strategy assumes that, on their own, managers cannot or will not produce, and have too many agendas of their own. The most elegant formulation of this model is found in the principal-agent literature in institutional economics. The goal, which might be labeled *marketization*, is to create competition through privatization or internal markets through contracts with managers.[2]

These two approaches can be mutually reinforcing, and should be used in combination. To change organizations, elected officials need to give managers flexibility, then contract for needed performance improvements. Competition with other providers can also be a powerful motivating force. But the federal government will have to develop its own version of managerialism, marketization, or some combination of the two consistent with the U.S. political culture and government institutions. The goal should not be slavish imitation of an imported management philosophy, but a renewed commitment to move management of the federal government out of the old bureaucratic paradigm and into the future, and to do so as quickly as possible.

2 Kettl 1998, pp. 13-14.

Appendix A: More on GPRA Legislation and Implementation

A number of web-sites provide useful information on GPRA and its implementation. OMB's page includes the full text of the bill and the Senate committee report which accompanied it, as well as memos providing guidance to agencies during implementation. See http://www.whitehouse.gov/OMB/mgmt-gpra/index.html.

The Congressional Institute and the National Partnership for Reinventing Government (the re-named National Performance Review) also have useful information at their websites:

- http://server.conginst.org/conginst/results/
- http://www.npr.gov/initiati/mfr/

Most reports of the General Accounting Office are available through their website at http://www.gao.gov/.

Appendix B: Great Expectations — The History of Management Reform

A number of authors have recounted the discouraging history of earlier reform initiatives, reaching generally similar conclusions.[1] Begun with great fanfare, those programs have usually ended up achieving much less than expected.

A GAO publication also looks back, aiming more specifically to relate those earlier initiatives to GPRA and its design. Paul Posner's fine short history of post-war management reform can be found in *Performance Budgeting: Past Initiatives Offer Insights for GPRA Implementation*, published by GAO in March 1997 (GAO/AIMD-97-46) and available on GAO's website (www.gao.gov).

1 Garvey 1995; DiIulio et al. 1993; Light 1997.

Appendix C: A Typology of Performance Measures

Practitioners and academics in the performance management field generally agree on the terminology and typology of performance indicators. For a complete review of this vocabulary, together with a clear and accessible discussion of many of the technical issues of performance measurement, see Harry Hatry, *Performance Measurement: Getting Results* (Washington: Urban Institute Press, 1999).

Table C-1 **General Types of Performance Indicators**
Types of Indicators
Input Indicators
Output Indicators
Outcome Indicators
Unit-Cost Ratios

Tables C-2-a through C-2-c give additional detail about these types of indicators. Examples are drawn from elementary and secondary education. Table C-3 outlines some additional types of data often provided in performance reports, including per capita ratios and explanatory data.

Table C-2-a
Input and output indicators

Types and Sub-Types of Performance Indicators	*Definition*	*Examples*	*Comments*
Input Indicators	Inputs are resources, usually expenditures or employee time, used in producing an output or outcome.		
l Financial		Expenditures or budget allocations devoted to schools. Can be expressed in current or constant dollars.	Some budget data can be difficult to disaggregate into operating units
l Other resources		Teacher full-time equivalents (FTEs)	Numbers of personnel (FTEs) are regularly reported. Other resources may include vehicles, space, or other equipment.
Output Indicators	Outputs are completed activities of a governmental unit.		
l Process Measures	Intermediate activities completed by the governmental unit.	Assessments completed for special education classification.	Important for agency managers, but customers and overseers may be more interested in outcomes.
l Workload measures	Units produced or services provided by the governmental unit.	Students graduated. Dropout or absenteeism rates.	Important to agency managers, but less policy relevant than outcome measures.

Table C-2-b
Outcome Indicators

Types and Sub-Types of Performance Indicators	*Definition*	*Examples*	*Comments*
Outcome Indicators	An outcome is an event, occurrence, or condition of direct important to users, and outside the governmental agency or program itself.		
l Intermediate Outcomes (sometimes called program goals)	An outcome expected to lead to a desired end, but not the end in itself. Indicators often focus on customer satisfaction measures, timeliness of responses, or other quality measures.	Student test scores. Percentage of students graduating on time. Student or parent satisfaction rates. Waiting time for completion of special education assessments.	Can be valuable for assessing the progress of programs. It can be difficult to determine how much the program has contributed to an outcome, compared with other factors.
l End Outcomes (sometimes called policy goals or long-term outcomes)	The end result. End outcomes can range from broad social indicators, such as infant mortality rates, to much narrower measures like street cleanliness.	Literacy rates. Percentage of students accepted in to college. Job readiness of graduates.	These outcomes, while important to citizens, are not usually completely under the control of a government agency. Collecting information on outcomes may require specialized methods, such as citizen surveys, trained observers, or technology.

Table C-2-c Unit-Cost Ratios			
Types and Sub-Types of Performance Indicators	*Definition*	*Examples*	*Comments*
Unit-Cost Ratios			Allocating all appropriate resources to specific organizational units and activities cannot be done without detailed cost accounting.
l Efficiency measures	The amount of an input divided by the amount of output/outcome.	Cost per student graduated.	Efficiency measures using outputs as opposed to outcomes do not provide true performance information. Also, greater efficiency is often achieved at the expense of service quality.
l Productivity measures	The amount of output/outcome divided by the amount of input.	Special education assessment completed per psychologist.	Productivity measures using outputs as opposed to outcomes do not provide true performance information.

Table C-3
Other Performance-Related Data

Data Types	*Definition*	*Examples*	*Comments*
Per Capita Indicators	Dollars or staff resources applied to services in relation to the population served.	Dollars spent per student enrolled.	These data are helpful for simple comparisons between different jurisdictions but do not provide performance information.
Explanatory Information	Additional data to help users of performance measures, especially outcome measures, understand the full range of factors affecting an organization's performance.		Program managers need reassurance that these data will be presented with outcome measures of performance. Even so, they will worry that outcome indicators will be misinterpreted.
l Controllable factors	Staffing ratios and other factors under governmental control.	Contractual limits on class size. Percentage of teachers with master's degrees. Percentage of uncertified teachers. Entry-level and average teacher salaries.	
l External factors	Information on economic, environmental, and demographic factors, as well as data on program demand.	Growth in school-age cohorts. Socioeconomic characteristics of students. Percentage of children in subsidized lunch programs. Percentage of students with ESL. Student mobility.	

Appendix D: Excerpt from Veterans Administration Departmental Performance Plan, FY 2001

Key Performance Goals and Measures

This section of the FY 2001 Performance Plan presents the performance goals and measures VA leaders consider critical to the success of the Department. Some of these deal with program outcomes; others pertain to the management of our programs. For each of these key performance goals, which are to be achieved in FY 2001, we present the:

- strategic goal and objective it supports;
- performance measure or measures to gauge progress toward achieving the performance goal;
- target levels of performance in graphic form;
- principle means and strategies to achieve the goal;
- crosscutting activities ongoing with other federal and non-federal organizations that contribute to accomplishment of the goal;
- major management challenges identified by the General Accounting Office and the VA Office of Inspector General that may affect goal achievement; and
- source of the performance information and how it is validated.

Our planning and measurement processes are very dynamic. This year's Performance Plan incorporates new strategic goals and objectives. These new goals and objectives have been reviewed with VSOs, Congressional staff, and OMB as part of the Department's Four Corners Stakeholder Consultation Sessions. As a next step, we hope our stakeholders will provide their views on which performance goals and measures are most important to the success of veterans programs. This year, and in future plans, we will improve the alignment of our performance measures with our strategic goals and objectives.

VA's Key Performance Goals For FY 2001

Strategic Goal	*Performance Measure*	*1997 Actual*	*1998 Actual*	*1999 Actual*	*2000 Est.*	*2001 Est.*
Restore the capability of disabled veterans to the greatest extent possible and improve their quality of life and that of their families	Compensation and dependency and indemnity compensation (DIC) program outcomes	N/A	N/A	N/A	N/A	N/A
	Vocational rehabilitation and employment rehabilitation rate	N/A	42%	53%	60%	65%
Ensure smooth transition for veterans from active military service to civilian life	Montgomery GI Bill usage rate	52.8%	54.0%	55.6%	57.0%	60.0%
	Foreclosure avoidance through servicing (FATS) ratio	41.0%	37.0%	37.6%	39.0%	40.0%
Honor and serve veterans in life and memorialize them in death for their sacrifices on behalf of the Nation	Chronic disease care index	76%	85%	89%	89%	95%
	Prevention index	67%	79%	81%	89%	90%
	Pension program outcomes	N/A	N/A	N/A	N/A	N/A
	Insurance program outcomes	N/A	N/A	N/A	N/A	N/A
	Percentage of veterans served by a burial option	65.7%	65.5%	66.8%	75.1%	76.2%
Contribute to the public health, socioeconomic well being and history of the Nation	Research projects relevant to VA's health care mission	97%	99%	99%	99%	99%
	Percent of residents trained in primary care	39.3%	41.3%	46.0%	47.0%	48.0%
	Percent of customers who rate the appearance of national cemeteries as excellent	78%	77%	79%	82%	88%
Provide *One VA* world-class customer service to veterans and their families through effective management of people, technology, processes, and financial resources	Percent of patients able to schedule primary care appointment within 30 days	N/A	N/A	N/A	N/A	N/A
	Percent of patients able to schedule specialist appointment within 30 days	N/A	N/A	N/A	N/A	N/A
	Percent of patients with scheduled appointments at VA health care facilities seen within 20 minutes	55%	66%	68%	75%	79%
	Abandoned call rate for compensation and pension	9%	13%	9%	10%	7%
	Blocked call rate for compensation and pension	45%	52%	27%	15%	12%
	Percent of patients who rate VA health care service as very good or excellent: Inpatient Outpatient	 65.0% 63.0%	 65.3% 65.0%	 65.0% 65.0%	 67.0% 67.0%	 68.0% 68.0%
	Percent of compensation and pension claimants who are satisfied with the handling of their claims	58%	57%	57%	65%	70%
	Percent of customers who rate the quality of NCA service as excellent	86%	85%	84%	88%	90%
	National accuracy rate for core rating work	N/A	64%	68%	81%	85%
	Average days to process rating-related actions on compensation and pension claims	94	128	166	160	142
	Appeals resolution time (in days)	628	686	745	670	650
	Average days to complete: Original education claims Supplemental education claims	 19 11	 25 15	 26 16	 26 17	 20 13
	Percent reduction from 1997 in average cost (obligations) per patient	N/A	-10%	-16%	-16%	-16%
	Percent increase from 1997 in number of unique patients treated	N/A	9%	15%	21%	24%
	Percent of medical care operating budget derived from alternative revenue streams	<1%	4%	4%	4%	3%

Selected Bibliography

Dick Armey. Letter to Franklin D. Raines. "RE: Results Act Implementation." August 7, 1997.

Burt S. Barnow. "Exploring the Relationship between Performance Measurement and Program Impact." *Journal of Policy Analysis and Management* 19,1 (Winter 2000): 118-141.

Jonathan D. Breul. "How 'The Government Performance and Results Act' Borrows From the Experience of OECD Countries." Paper prepared for the Fulbright Symposium on Public Sector Reform, Brisbane, Australia, July 1994.

Joseph C. Burke and Shahpar Modarresi. P*erformance Funding and Budgeting: Popularity and Volatility — The Third Annual Survey.* Albany, NY: The Nelson A. Rockefeller Institute of Government, 1999.

John J. DiIulio, Gerald Garvey and Donald F. Kettl. *Improving Government Performance: An Owner's Manual*. Washington, D.C.: Brookings, 1993.

Thomas L. Gais. "Concluding Comments: Welfare Reform and Governance." *Learning From Leaders: Welfare Reform Politics and Policy in Five Midwestern States*, edited by Carol Weissert. Albany, N.Y.: Rockefeller Institute Press, 2000.

Gerald Garvey. "False Promises: The NPR in Historical Perspective," in Donald F. Kettl and John J. DiIulio, eds. *Inside the Reinvention Machine: Appraising Governmental Reform.* Washington, D.C.: Brookings, 1995.

William T. Gormley and David L. Weimer. *Organizational Report Cards.* (Cambridge: Harvard University Press, 1999.)

Harry P. Hatry. *Performance Measurement: Getting Results.* (Washington D.C.: The Urban Institute Press, 1999.)

Robert S. Kaplan and David P. Norton. *The Balanced Scorecard.* (Boston: Harvard Business School Press, 1996.)

Donald F. Kettl. Reinventing *Government: A Fifth-Year Report Card.* Washington, DC: Brookings Institution's Center for Public Management, September, 1998.

Paul C. Light. *The Tides of Reform: Making Government Work 1945-1995.* New Haven: Yale University Press, 1997.

Thomas Peters and Robert H. Waterman. *In Search of Excellence: Lessons From America's Best Run Companies.* (Thorndike, ME: G.K. Hall & Company*,* 1997.)

Public Law 62, 103rd Cong., 1st sess. (3 August 1993), *Government Performance and Results Act of 1993*.

Dennis C. Smith. "What Public Managers Learn From Police Reform in New York? COMSTAT and the Promise of Performance Management." Paper prepared for the annual meeting of the Association of Public Policy and Management, Washington, D.C., November 1997.

Dennis C. Smith. "Performance Management in New York City: The Mayor's Management Plan and Report System in the Koch Administration." Paper prepared for the annual meeting of the Association of Public Policy and Management, Washington, D.C., October 1993.

United States Department of Agriculture. *Draft USDA Forest Service Strategic Plan (2000 Revision)*, Washington, D.C: November, 1999.

United States General Accounting Office. *Forest Service: A Framework for Improving Accountability.* Washington, D.C., October 1999. GAO/RCED/AIMD-002.

United States General Accounting Office. *Managing for Results: Opportunities for Continued Improvements in Agencies' Performance Plans.* Washington D.C.: July 20, 1999, GAO/GGD/AMID-99-215.

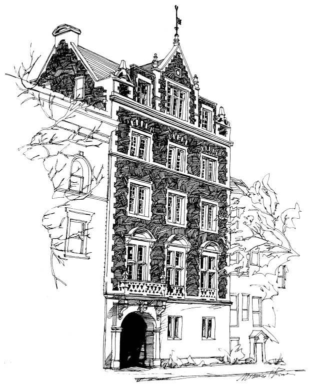
United States General Accounting Office. *Performance Budgeting: Initial Agency Experiences Provide a Foundation to Assess Future Directions.* Washington D.C.: July 1, 1999, GAO/T-AIMD/GGD-99-216.

United States General Accounting Office. *Performance Budgeting: Initial Experiences Under the Results Act in Linking Plans with Budgets.* Washington, D.C.: April 1999. GAO/AIMD/GGD-99-67.

United States General Accounting Office. *Performance Budgeting: Past Initiatives Offer Insights for GRPA Implementation.* Washington D.C.: March 1997. GAO/AIMD-97-46.

United States Office of Management and Budget. *FY 2000 Government-Wide Performance Plan*. Washington D.C.: February 1999.

The Urban Institute and the International City/County Management Association. *Comparative Performance Measurement: FY 1996 Data Report.* Washington. D.C.: Urban Institute Press, 1998.



NONPROFIT
ORG.
U.S. POSTAGE
PAID
ALBANY, N.Y.
PERMIT NO. 21

**The
Nelson A.
Rockefeller
Institute
of
Government**

The State University
of New York

411 State Street
Albany, New York 12203-1003