Monitoring and Evaluation in the United States Government
An Overview
No. 26 | October 2011

Katharine Mark and John R. Pfeiffer
Monitoring and Evaluation in the United States Government: An Overview

Katharine Mark and John R. Pfeiffer
IEG: Improving Development Results Through Excellence in Evaluation

The Independent Evaluation Group is an independent unit within the World Bank Group; it reports directly to the Bank’s Board of Executive Directors. IEG assesses what works, and what does not; how a borrower plans to run and maintain a project; and the lasting contribution of the Bank to a country’s overall development. The goals of evaluation are to learn from experience, to provide an objective basis for assessing the results of the Bank’s work, and to provide accountability in the achievement of its objectives. It also improves Bank work by identifying and disseminating the lessons learned from experience and by framing recommendations drawn from evaluation findings.

IEG’s Evaluation Capacity Development Working Papers are an informal series to disseminate the findings of work in progress to encourage the exchange of ideas about development effectiveness through evaluation.

The findings, interpretations, and conclusions expressed here are those of the author(s) and do not necessarily reflect the views of the Board of Executive Directors of the World Bank or the governments they represent, or IEG management.

IEG cannot guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply on the part of the World Bank any judgment of the legal status of any territory or the endorsement or acceptance of such boundaries.


Contact: IEG Communication, Learning and Strategies
e-mail: ieg@worldbank.org
Telephone: 202-458-4497
Facsimile: 202-522-3125
http://ieg.worldbankgroup.org
Contents

Abbreviations
Acknowledgements

1. Introduction: Purpose and Scope
2. Key Institutions and Their Roles in Monitoring and Evaluation
   2.1 Executive Branch
      2.1.a Office of Management and Budget
      2.1.b Operating Agencies
   2.2 Legislative Branch
      2.2.a United States Congress
      2.2.b Government Accountability Office
   2.3 Civil Society
      2.3.a Advocacy and Funding for Government Oversight
      2.3.b Private Sector Researchers and Evaluators
   3.1 Government Performance and Results Act
   3.2 Performance and Accountability Reports
   3.3 The Program Assessment Rating Tool
      3.3.a Reporting to the Public
      3.3.b Performance Improvement Council
   3.4 Obama Administration Performance Initiatives
   3.5 Program Evaluation
      3.5.a New Initiatives to Promote Evaluation
      3.5.b An Increased Focus on Rigorous Impact Evaluation
4. Impacts of Recent United States Monitoring and Evaluation Efforts
   4.1 Strengths
   4.2 Challenges
5. Lessons Learned

Bibliography
Abbreviations

CBO  Congress ional Budget Office
GAO  U.S. Government Accountability Office
GPRA  Government Performance and Results Act
HPPG  High Priority Performance Goal
M&E  Monitoring and evaluation
OMB  Office of Management and Budget
PAR  Performance and Accountability Report
PART  Program Assessment Rating Tool
PIC  Performance Improvement Council
Acknowledgements

The authors thank the Independent Evaluation Group (IEG) of the World Bank Group and the Urban Institute for their support of this work. This report was informed and enriched by interviews with Denise Fantone, Director of Strategic Issues for the Government Accountability Office (GAO), Elizabeth Curda, Deputy Director of Strategic Issues for GAO, Robert Shea of Grant Thornton (and formerly at the Office of Management and Budget), and Thomas Skelly of the Department of Education. Harry Hatry and Sharon Cooley provided (Urban Institute) valuable ideas and observations and a thoughtful review of the report.

The report also benefitted from comments by Nidhi Khattri and Ximena Fernandez Ordonez (IEG) and Keith MacKay and Manual Fernando Castro Quiroz (consultants, IEG) and editorial help from Heather Dittbrenner (IEG). Denise Fantone (GAO) provided helpful comments on an earlier draft. The task managers were Nidhi Khattri and Manual Fernando Castro. Katharine Mark is a Senior Research Associate at the Urban Institute. John Pfeiffer is a Program Examiner at the Office of Management and Budget. None of the views expressed herein by Mr. Pfeiffer should be construed to represent the policies or positions of the Office of Management and Budget or the U.S. government.
1. INTRODUCTION: PURPOSE AND SCOPE

In his inaugural address on January 20, 2009, President Barack Obama inspired his administration’s efforts to improve government performance with these words: “The question we ask today is not whether our government is too big or too small, but whether it works—whether it helps families find jobs at a decent wage, care they can afford, a retirement that is dignified. Where the answer is yes, we intend to move forward. Where the answer is no, programs will end. And those of us who manage the public's dollars will be held to account—to spend wisely, reform bad habits, and do our business in the light of day—because only then can we restore the vital trust between a people and their government.”

Thus, under a new President, the U.S. government’s long history of concern with accountability and oversight continues and evolves in new directions. The last two decades have seen an increased interest in outcomes-based performance monitoring and evaluation (M&E). Important new legislation, special attention by the administrations of Presidents Bill Clinton and George W. Bush, greater emphasis on transparency and results, and the introduction of initiatives such as the Program Assessment Rating Tool (PART) all have brought performance measurement more prominence. These developments build on older and well-established traditions and institutions of oversight, assessment, and review, such as the U.S. Government Accountability Office (GAO), a lively nonprofit sector keeping watch on government programs, a growing reservoir of performance management experience at state and local levels, and a Congress with a strong oversight role.

President Obama’s administration has introduced additional initiatives addressing performance-based management and promoting program evaluation, to encourage more advances in the system. Still, there are a number of ways in which the U.S. government can improve its use of M&E to increase the effectiveness of its programs.

To be successful, an M&E system must not only set clear objectives, identify and measure indicators, and carry out evaluations, but it must also support the use of the information derived from those efforts to improve performance. To that end, this report will examine the ways in which the U.S. government carries out the following functions:

- Identifying clear and specific missions, goals, and outcomes to ensure that programs are well targeted
- Creating performance measures to track progress toward key outcomes
- Setting targets for progress in accomplishing the outcomes
- Reporting performance information to enhance government accountability
- Using performance measurement data in managerial decisions to improve programs
- Carrying out evaluations of government programs
- Using evaluation findings to improve programs
This report is divided into five parts. Following this introduction, Section 2 provides an overview of the institutions and most important features in the landscape of M&E at the federal level in the United States. Section 3 details the actual systems for performance M&E that are now in place in the Executive Branch and coordinated (or led) by the Office of Management and Budget, including a look at their evolution and expected future trends. The focus is on the executive system, because it directly supports management and budgeting decisions, and because it provides a key basis for evaluation and research conducted by other agencies (such as the GAO and CBO). Section 4 discusses the strengths and particular challenges faced by these systems, and Section 5 concludes the report with lessons that may be useful to other countries.

We should note that this review distinguishes between program and performance monitoring or assessment and program evaluation, even though in practice the concepts are often confused, and the latter concept, in particular, has often been poorly understood and poorly practiced. Monitoring or assessment—the majority of those program analyses conducted in the U.S. government—typically reflects the ongoing, routine collection, analysis, and reporting of program performance measures to determine progress toward established goals and objectives. In contrast, program evaluations are more in-depth individual studies conducted periodically or on an ad hoc basis to assess how well a program is working, why it is achieving the results it is, and what adjustments might be pursued to improve performance. Although monitoring is usually done in-house, evaluations are often conducted by outside experts, both to ensure the independence of the findings and to apply greater technical expertise. Both sources of program information can be useful to policy makers and managers; ideally, their contributions to program understanding and improvement are complementary (cf. U.S. GAO 2011d). Most of the M&E activities described in the following pages focus on systematic performance monitoring and assessment activities. Recent governmentwide attempts to promote rigorous program evaluation will chiefly be addressed in a separate section.

2. KEY INSTITUTIONS AND THEIR ROLES IN MONITORING AND EVALUATION

The U.S. federal government has three independent branches—the Executive, the Legislative, and the Judicial. Although it is considered a presidential system, as there is a strong President heading the Executive Branch, the directly elected bicameral legislature—the Congress—is also strong. The Congress is comprised of a 100-member Senate with two representatives elected for an unlimited number of six-year terms from each of the 50 states and a 435-member House of Representatives, whose members are elected for an unlimited number of two-year terms and apportioned among the states by population. Congress holds final authority over budgetary matters, including the exclusive power to appropriate funds and the authority to investigate and oversee the Executive Branch.

Overall, the United States’ governance system is highly decentralized. As a result, many functions that are carried out in other nations by central government agencies are primarily state or local government responsibilities in the United States. For example, the federal government provides only about one-tenth of funding for primary and secondary education,
with the rest coming from state and local governments and private institutions. In the United States, state and local governments build schools, develop curricula, hire teachers, set graduation standards, and more. This division of roles and responsibilities both reduces the scope of the federal role in primary and secondary education and complicates M&E at the national level. (See http://www2.ed.gov/about/overview/focus/whattoc.html for more information on this division of labor.)

The U.S. government’s size and complexity make it impossible to have a single unified and comprehensive M&E framework, and, as a result, there are a wide range of entities monitoring and evaluating programs throughout the 15 major departments and many smaller federal agencies and institutions. As in most countries, the more formal, routine work of M&E in the U.S. government is conducted by the Executive Branch, both from within the Executive Office of the President, which provides direct staff support to the President, and from within the many departments, agencies, and programs through which most of the administrative work of the government is carried out. The Congress also engages in M&E of government programs—both directly through its internal committee hearing process and indirectly through requirements that it imposes on the Executive Branch in the legislation it enacts and through its own support agencies, principally the GAO, the Congressional Budget Office (CBO), and the Congressional Research Service of the Library of Congress.

In addition, in the United States, think tanks (organizations that study public issues and sometimes advocate for their solution), independent consultants, and universities play significant roles in the performance measurement and program evaluation work of the federal government. Often these organizations work under contract to an Executive Branch agency, but they also may work for other public and private stakeholder entities with an interest in the outcomes of federal policies and programs. Similar efforts are undertaken at state and local levels of government as well.

2.1 Executive Branch

Within the Executive Branch, the key players in M&E are the Office of Management and Budget (OMB) and the individual federal agencies, including the 15 cabinet departments whose senior policy officials are appointed by the President and approved by the Senate and a number of smaller agencies with varying degrees of independence from the President (see Figure 1).
2.1.a Office of Management and Budget

The OMB is the largest component of the Executive Office of the President and is staffed chiefly by several hundred career civil servants who serve regardless of the party in power; its director and top officials are appointed by and report to the President. The budget side of OMB, which is comprised of a central coordinating office, the Budget Review Division, and groups of career program examiners that roughly mirror the agencies and programs they oversee, chiefly helps to prepare the budget for the federal government that the President submits to Congress every year (cf. Dame and Martin, 2009). That budget is made up of funding requests for every federal department, agency, and program. The OMB examiners review annual agency funding requests and new policy proposals and advise senior officials on their merits, including likely effectiveness and efficiency and consistency with Presidential policy objectives. The budget staff plays a similar role with regard to agency legislative proposals. They also track Congressional actions on all spending and legislative proposals that the President sends to Congress or that Congress itself initiates, and they play a similar advisory role for those actions. Finally, because of their larger numbers and more regular contact with agency program officials, the budget staff also play a key role in helping oversee and coordinate major Executive Branch management improvement initiatives, as described below.

OMB’s management side includes several smaller offices, some established by statute, that provide governmentwide policy and oversight of specific functions, such as financial management, procurement, and information technology, and others that provide OMB-wide support, including economic policy, legislative and regulatory review, and governmentwide management improvement. It is in these offices that OMB’s governmentwide program performance M&E initiatives take shape under the leadership of politically appointed senior policy officials, who often come to these positions with extensive backgrounds in public and
private sector management. In the past two decades, such M&E initiatives have included the implementation of the Government Performance and Results Act of 1993 (GPRA) under President Clinton, the President’s Management Agenda and PART under President George W. Bush, and most recently, the High Priority Performance Goals (HPPG) Initiative and the Program Evaluation Initiative under President Obama.

2.1.b. Operating Agencies

The individual federal departments and agencies directly collect most federal performance data and manage most program evaluations conducted by the government. The determination of performance data to be collected, used, and reported by an agency stems not only from its own operational imperatives but also from governmentwide and program-specific mandates and requirements imposed by OMB and Congress.

Both Congress and the Executive Branch select evaluation topics to be examined. Congress periodically mandates specific evaluations to be undertaken in legislation, including appropriations acts, and in accompanying reports. When Congress does this, it generally provides funding for that evaluation. Individual federal agencies may include in their budgets their own proposals for evaluations that they or OMB feels are important. Most major agencies, even some small ones, have some form of unit responsible for evaluations. Until recently, agency program and evaluation offices often conducted their own evaluations; however, this practice has declined since President Bush’s administration began emphasizing the need for evaluations that are free of possible challenges to their independence and integrity, which could arise from programs evaluating themselves. Most in-depth program evaluations are contracted to experts outside of government at independent consultancies, think tanks, or universities. Most of these external evaluation contracts involve a competitive process and the winning proposals chosen on merit, including those mandated by Congress.

Certain evaluations are not contracted out, including those undertaken by the GAO at the request of Congress. Other assessments are undertaken by the Congressional Research Service and the CBO, both arms of Congress that do special studies but generally do not examine program outcome results in depth. CBO, for example, focuses primarily on estimating the costs of proposed legislation and producing reports on major fiscal issues.

2.2 Legislative Branch

2.2.a United States Congress

The U.S. Congress, through its legislative authorization and appropriations committees in both the House and the Senate, plays a powerful role in overseeing and shaping Executive Branch policies and programs. The tenure of committee and subcommittee leaders and their staffs often is far longer than that of the President, who may only be elected for two 4-year terms, and senior agency officials appointed by the President, whose average tenure is only about two years, and the approval of these committees is necessary for such major decisions as the creation or termination of programs and the funding of program operations. Surrounded by stakeholder organizations interested in the outcomes of their deliberations,
Congress possesses myriad sources of information to aid decision making. As public pressures mount for the government to focus its scarce resources on what works, a challenge for both the Executive Branch and the Congress is to find more effective ways to incorporate the results of formal performance assessment and evaluation efforts into decision making.

Congress has played an important direct role in performance reform in the Executive Branch by passing legislation that mandates the development of new information and reporting, most notably the GPRA and the GPRA Modernization Act of 2010. GPRA included a number of reporting requirements that provide Congress with regular performance information on every agency and its programs in five-year strategic plans and annual performance reports. The Modernization Act amended the GPRA to place more emphasis on the use of performance information, including data analysis and regular reviews by agency leaders, to establish federal cross-cutting Priority Goals and codify agency priority goals set every two years and to require a single federal Web site where performance information will be made publicly available.

2.2.b Government Accountability Office

The GAO, with its annual output of approximately 1,000 reports and testimonies and a staff of 3,300, provides the Legislative Branch with a major means of assessing the effectiveness of government programs. In this way, it serves as something of a counterweight to OMB, which serves the Executive Branch. The GAO was established as the General Accounting Office in 1921, focusing on financial oversight of the Executive Branch. In 2004, it was renamed the Government Accountability Office to reflect a broader understanding of its mission, which now expressly includes a commitment “to help improve the performance and ensure the accountability of the federal government” (U.S. GAO 2004a). The Comptroller General explained, “The days of accountants in green eyeshades…are long gone” and the GAO’s activities now encompass a broad range program evaluation, policy analysis, and legal opinions about all government programs and agencies (Walker 2004).

GAO serves as the principal research and investigative arm of Congress. Any member of Congress can request an assessment of government programs, with priority given to Congressional mandates and leadership requests. The range of work conducted by the GAO includes:

- “Evaluations of federal programs, policies, operations, and performance;
- Management and financial audits to determine whether public funds are spent efficiently, effectively, and in accordance with applicable laws;
- Investigations to assess whether illegal or improper activities are occurring;
- Analyses of the financing for government activities;
- Constructive engagements in which GAO works proactively with agencies, when appropriate, to help guide their efforts towards transformation and achieving positive results;
- Legal opinions to determine whether agencies are in compliance with applicable laws;
- Policy analyses to assess needed actions and the implications of proposed actions; and
- Additional assistance to the Congress in support of its oversight, appropriations, legislative, and other responsibilities” (U.S. GAO 2004a).
In addition to its own analyses, GAO has long been a strong proponent of efforts to build and strengthen Executive Branch M&E capabilities and practices. GAO recently noted, for example, that “the GPRA Modernization Act can offer opportunities to help make tough choices in setting priorities as well as reforming programs and management practices to better link resources to results” (U.S. GAO 2011a). GAO makes all its products available to Congress and to the public and posts them on its Web site—www.gao.gov.

The 15-year term of the Comptroller General, the head of the agency, ensures both continuity and independence, and GAO’s mandate to be bipartisan ensures respect from the entire Congress. GAO’s reputation for fairness also is buttressed by its policy of giving relevant agencies and other directly affected parties the opportunity to comment on draft reports and of including those comments in the reports along with GAO’s assessment. GAO’s recommendations are not formally binding, but it has found that over 80 percent of them are implemented.

Examples of GAO Performance Assessment

- Biennially, GAO highlights major problems that are at high risk for waste, fraud, abuse, and mismanagement or that are in need of broad reform. Among the 30 topics on GAO's High-Risk list is protecting public health through enhanced oversight of medical products. GAO’s Web site contains a clear, concise text and video summary of its many Congressional reports on the topic, explaining why it is high risk, what its studies have found, and what it has recommended (www.gao.gov/highrisk/risks/safety-security/public_health.php).

- GAO’s recommendations to OMB for improving its oversight and transparency of federal information technology investments led OMB to develop and implement improved oversight by deploying a “dashboard,” a Web site clearinghouse of information that provides details on all major federal information technology investments (U.S. GAO 2008b, 2008c, 2008d, 2009d).

- GAO identified areas in which the Department of Housing and Urban Development could be more proactive in promoting energy efficiency and green building. Since then the Department has added strong incentives for energy efficiency and green building in competitive grants programs that were funded through the American Recovery and Reinvestment Act of 2009 (U.S. GAO 2009b).

2.3 Civil Society

2.3.a Advocacy and Funding for Government Oversight

A key feature of the U.S. M&E environment is the vigorous role played by nongovernmental organizations both in overseeing and assessing government policies and programs—including funding the work of experts who are monitoring and evaluating government programs—and also in working to influence government decision making in both the Executive Branch and Congress. Some of the most influential evaluations of government
programs in the United States have been the product of collaboration among government agencies, private philanthropic foundations, and research institutions.

The wealth of resources supporting M&E in the United States results, in part, from such significant political factors as the constitutional separation of powers, the complexity of the government, an Executive Branch that draws on unelected officials to staff policy-making positions, and the openness of the American system to external sources of advice. A plethora of advocacy and watchdog organizations represents virtually every societal group affected by government activity that is interested in assessing the results of government programs. Many of these organizations promote, fund, or demand monitoring or evaluation of specific government programs. In some cases, the purpose is to call attention to problems that are otherwise being ignored by the government; in other cases, the goal is to supplement analysis and evaluation by government agencies.

A Recent Example of Foundation Support: Evaluating Health Reform in Massachusetts

After major health reform passed in Massachusetts in 2006, various stakeholders became concerned that existing data sources were not sufficient to support tracking its impacts. A group of foundations, led by the Blue Cross Blue Shield of Massachusetts Foundation (http://bluecrossfoundation.org/), funded a survey to evaluate the reform. After the first year of the survey, the state rebid its own survey as well. Although the state initially did not see the need for another survey, in the end it welcomed the findings from the foundations’ work.

Major findings were that health reform led to significant gains in insurance coverage and health care access, use, and affordability. The findings from the evaluation helped move the policy debate in Massachusetts—and at the national level—from the question of whether it was possible to achieve near universal coverage to an ongoing discussion of how to maintain and pay for the coverage expansion. Study findings were also an important part of the state’s discussions with the federal government on continued financial support for the initiative, as it provided independent information on the impacts of reform.

Source: Correspondence with Sharon Long, PhD, Urban Institute, 6/18/10.

2.3.b Private Sector Researchers and Evaluators

Partly because of the widespread interest in analysis of public policy, many researchers and analysts who specialize in evaluating government programs work in the private sector in the United States. The organizations that employ them include think tanks (research organizations, often focused on analysis of public policy), universities, and private, for-profit consulting firms that are available to conduct high-quality evaluation. The organizations fill a host of functions from evaluation to public education to (in some cases) advocacy. Although they share an orientation around policy issues, think tanks examine diverse sectors, ranging from social policy, political strategy, and economic programs to scientific issues, technological growth, and military advice. In addition to their subject matter interests, they represent a wide array of political views – conservative, liberal, and nonpartisan. The
organizations provide valuable forums for discussion and debate, often encourage dialogue via conferences and lecture series, and publish books and papers. Many engage resident “experts” in specific fields and actively collaborate with academic institutions, government agencies, and private sector partners.

Universities whose individual professors long have undertaken evaluation and research projects for the government increasingly launch think-tank-like research entities of their own (see www.hks.harvard.edu/library/research-guides/think-tanks-directory.htm). This rich institutional framework provides resources that greatly enhance the ability of the federal government to pursue analyses and evaluations of its programs that are beyond its own capabilities.

3. MONITORING AND EVALUATION: SYSTEMS AND TOOLS

Although each new presidential administration initiates its own management reform initiatives, this section will focus primarily on describing those of the past two decades that most directly affect current performance M&E practices (U.S. GAO 1997). These efforts include, principally, the GPRA, the President’s Management Agenda, the PART, the GPRA Modernization Act, HPPG, and program evaluation initiatives under President Obama.

3.1 Government Performance and Results Act

Congress enacted the GPRA out of concern that decision making in both the Executive and Legislative Branches was placing greater emphasis on program inputs and activities than on program outcomes and results (U.S. GAO 1997, 2004b). Prior to enactment of GPRA, the GAO studied the state of program performance measurement in 103 large federal agencies with more than 1,000 employees or annual budgets greater than $500 million (U.S. GAO 1992). The study found that about two-thirds of the agencies had a long-term strategic plan and collected performance measures, but that only about half of those used their performance measures to assess progress in accomplishing the goals in their plans, and few used the data to prepare regular reports on their progress in achieving their strategic goals. Most used performance data for internal management purposes and not for strategic decision making. Very few agencies reported such information externally to OMB, the Congress, or the public.
GAO also found that only nine agencies possessed the characteristics deemed necessary to integrate measurement with strategic planning and use it for policy decision making: a single strategic plan with measurable goals, a central office that collected performance data, regular consolidated reports, and top management support.

The GPRA aimed to correct these deficiencies. It sought to increase public confidence in the government by holding agencies accountable for results; to improve program effectiveness and accountability by focusing on results, service quality, and customer satisfaction; to help federal managers improve service delivery by requiring plans to meet program objectives and providing them with information on results and service quality; to assist Congressional decision makers by providing more objective data on program performance and on programs’ relative effectiveness and efficiency; and to improve internal government management (U.S. OMB 2010c).

The GPRA required every federal agency to develop five-year strategic plans and goals, to develop annual performance plans identifying near-term goals aligned with the long-term strategic goals, and to prepare annual performance reports on their accomplishments. For any goal not met, the reports were to explain why and to outline actions needed to address the failure. Annual performance plans tied to agency budget submissions were required to present results-oriented annual goals and performance measures aligned with the long-term strategic goals and linked to programs contained in the budget (U.S. GAO 2004b, p. 5). Six months after the end of each fiscal year, agencies also were required to produce an annual performance report with the latest values for each performance indicator, along with data on previous years’ values.

The legislation also required five-year strategic plans, with reviews at least every three years. No cross-agency strategic plan was mandated, and none has yet been prepared by the Executive Branch. The legislation also referred to in-depth program evaluations but did not explicitly require them. GPRA did require that a summary of the findings from any program evaluations completed during the fiscal year covered by the annual performance report be included in those annual reports however.

All these requirements still hold and are being followed by the federal government. They form the framework particularly for the performance measurement work of the federal government. The Executive Branch is responsible for providing these reports as called for and has been doing so since the Act became fully operational in fiscal year 1999.

---

1 The GAO has recommended to OMB that cross-cutting analyses be conducted to ascertain potential redundancies in resource allocation and effort or whether programs are possibly working at cross purposes. See U.S. GAO 2008a. The GAO also has recommended that OMB develop a cross-agency strategic plan to capture the entire spectrum of U.S. investment in one place.

2 A note on terminology: In the United States, the term “performance measurement” is used for the regular reporting of performance information, particularly outcome and output information. Thus, it appears to be equivalent to the use of the term “monitoring” more commonly used by the World Bank.
Under GPRA, OMB is responsible for ensuring that agency plans and reports are consistent with the President’s budget and policies. OMB issues annual guidance for the plans and reports in its annually updated guidance memorandum on preparation of agency budget requests, “Circular No. A-11: Preparation, Submission, and Execution of the Budget,” and OMB’s program examiners review and approve the documents at approximately the same time that they review the agencies’ annual budget requests.

3.2 Performance and Accountability Reports

GPRA required all federal agencies to produce annual performance reports containing performance targets and results achieved against their targets during the fiscal year just ending. (The U.S. government’s fiscal year runs from October 1 through September 30.) The reports are submitted to the President, Congress, and OMB in November following the end of the respective fiscal year. The Performance and Accountability Reports (PARs) have been an important part of GPRA reporting since the passage of the act, and they continue to be prepared.

The key elements of the reports include annual outcome-based performance targets for the next fiscal year, comparison of actual performance with the targets for prior years, and an analysis and explanation of the causes of any missed targets or changes in trends, including a graphic depiction of historical performance trends. A description of whether and how the agencies use their data to promote improved outcomes, or to highlight promising practices, is also required.

An OMB circular regarding the preparation of the PARs for fiscal year 2010 describes it as a transition year, with changes to be expected in the form and content of the report for next year. The circular describes the elements required and calls for a summary of performance information “in a brief, user-friendly format that is easily understood by a reader with little technical background in these areas,” with the stated aim to increase accountability by making the information more transparent and accessible to Congress and the public (U.S. OMB 2010c). Both the circular and further guidance from OMB in summer 2010 demonstrated that the Obama administration would expect the PARs to emphasize communicating clear, concise analyses to the public through the Internet, while also making detailed background information accessible through the same medium. It directed that agency PARs coherently display performance trends and progress toward targets, explain key causal factors affecting trends, align actions with outcomes, and link (via the Internet) to more detailed explanatory analyses of the data, evaluations, and relevant studies (for example, about the size and characteristics of a problem, of the regulated community to be influenced, and/or of the population eligible to be served) (Metzenbaum 2010).

Figure 2 shows a screen of a page from the PAR of the Department for Veterans Affairs for fiscal 2010 (U.S. Department of Veterans Affairs 2010).
3.3 The Program Assessment Rating Tool

In its 2004 Budget, the Bush administration introduced PART, partly out of a belief that implementation of the GPRA had “fallen short of its authors’ hopes” (U.S. OMB 2002, p. 49). The Bush administration believed that agency plans suffered from performance indicators that were vague and too numerous and that the plans were usually ignored in budget decision making. With the creation of the PART, the Bush administration attempted to carry out the spirit as well as the letter of GPRA. The administration wanted to shift the focus of federal goal setting and measurement from the agency level, as was the case with GPRA, to the program level. Over the course of the next five years, virtually every federal

---

3 For more detailed information about the PART, see www.whitehouse.gov/omb/performance_past/.
program was assessed—more than 1,000 individual programs, representing 98 percent of all federal spending.

The PART was designed by OMB career staff as a diagnostic tool for assessing program performance and driving performance improvements. Of special relevance to this review, the PART included questions about whether the program has established a performance monitoring system and whether it conducts evaluations. Although most of the PART focused on the processes of strategic planning, performance management, and evaluation, only the fourth section asked about the actual results of the program.

Once completed, PART reviews were to help inform budget decisions and identify actions for agencies to take to improve results. The PART was intended to provide a consistent approach to assessing and rating programs across the federal government. PART assessments reviewed overall program effectiveness, from design to implementation and results. The assessments were intended to be completed by the agencies in collaboration with their individual OMB program examiners before OMB’s annual budget season began in the fall. They also were intended to provide a standard, comprehensive, objective basis for comparative analyses of agency program budget requests that could improve the quality and consistency of OMB’s budget review process, as well as to provide a basis for ongoing initiatives to address program weaknesses and improve performance.

The PART was actually a multipart questionnaire about a Federal program to be completed by professionals within each program being assessed. Each team completed the survey instrument with the assistance and oversight of OMB program examiners whose regular responsibilities largely focus on oversight of budget and policy development and execution for the program.

Although agencies were given the lead responsibility for completing the PART assessments, the process was intended to involve both agency personnel and OMB staff, typically the relevant program examiners. Agencies involved personnel both from the program being reviewed and from their own central budget and planning offices. The requirement that both the agencies and OMB had to agree on the answers to each question provided incentive for collaboration throughout the process of completing the questionnaire. Agencies that took advantage of the opportunity to consult with OMB before and during the assessment process generally were more satisfied with the results, as providing satisfactory justification for the seemingly simple “Yes” or “No” responses could sometimes prove quite difficult.

Even before conducting an assessment, OMB and agencies first needed to select and define the programs that were to be subject to review, that is, to determine the unit of analysis, or “program,” to be assessed. Although the budget structure did not always clearly define all programs (“program activities” in the budget are not always the same activities that are managed as a program in practice), the budget structure and associated documents, for example, the “Congressional Justification” volumes that the agencies sent to Congress in support of their annual budget requests, were often comprehensive and readily available as an inventory of agency programs. A “program” could also be a collection of programs or activities that were managed as one entity or that had a clear set of goals.
For example, several grant programs within the Administration on Aging of the Department of Health and Human Services were combined for consideration under a single PART assessment. Each of the programs contributed to a similar program goal of providing support to older individuals to enable them to stay in their homes rather than move to a long-term care facility. Often the same organization received multiple grants, and it was the combination of those grants that enabled individuals to remain at home. The data reporting for these grants was combined, and the services were so interrelated that assessing the programs together yielded more meaningful information for understanding and improving performance.

The PART questions were divided into four sections as follows, the first three primarily focused on internal management rather than program outcomes.

**Figure 3. Distribution of PART Questions**

<table>
<thead>
<tr>
<th>PART Questionnaire Section</th>
<th>Section Focus</th>
<th>Relative Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Program Purpose and Design</td>
<td>Clarity and targeting of the program and redundancy with other government programs</td>
<td>20</td>
</tr>
<tr>
<td>II. Strategic Planning</td>
<td>Management practices related to strategic planning, performance measurement, and evaluation, including whether appropriate annual and long-term outcome indicators existed, the nature of targets, whether evaluations are carried out, and the relationship between budget requests and performance goals</td>
<td>10</td>
</tr>
<tr>
<td>III. Program Management</td>
<td>Whether the program was effectively managed to meet program performance goals. Key areas included financial oversight, evaluation of program improvements, performance data collection, and program manager accountability.</td>
<td>20</td>
</tr>
<tr>
<td>IV. Program Results/Accountability</td>
<td>Addressed actual results. Asked about performance, progress in achieving goals, improved efficiencies, and evaluation findings.</td>
<td>50</td>
</tr>
</tbody>
</table>

OMB examiners and agencies were encouraged to work together to develop alternative measurement approaches for programs where it was difficult to develop quantitative measures and where qualitative, expert review, or other measures were more appropriate. Measures were expected to focus on outcomes, although in some cases output measures were permissible. Agencies could appeal disagreements with OMB on PART questions to an interagency appeals board made up of senior officials from multiple agencies.

Within each section, weights were distributed by question to total 100 percent. When answers were added, they provided a numeric score between zero and 100 for each of the four sections. These were translated into a final rating of effective, moderately effective, adequate, or ineffective. An overriding rating of “results not demonstrated” was given when
OMB and the agency could not agree on performance indicators and/or if OMB deemed performance information inadequate. This rating applied even if the score otherwise would be relatively high.

**Figure 4. PART Ratings**

<table>
<thead>
<tr>
<th>Program Rating</th>
<th>Meaning</th>
<th>2008 Ratings (1,015 Programs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective</td>
<td>Program had ambitious goals, achieved results, was well managed, and had improved efficiency</td>
<td>193</td>
</tr>
<tr>
<td>Moderately Effective</td>
<td>Program had ambitious goals and was well managed but likely needed to be more efficient or to address other design or management problems.</td>
<td>326</td>
</tr>
<tr>
<td>Adequate</td>
<td>Program likely needed more ambitious goals, better results, improved accountability, or stronger management practices.</td>
<td>297</td>
</tr>
<tr>
<td>Ineffective</td>
<td>Program was unable to achieve results because of a lack of clear purpose or goals, poor management, or some other significant weakness.</td>
<td>26</td>
</tr>
<tr>
<td>Results Not Demonstrated</td>
<td>Program was unable to develop acceptable performance goals or collect data to determine whether it was performing.</td>
<td>173</td>
</tr>
</tbody>
</table>

There were fears among the agencies that poor PART ratings would lead OMB to recommend funding cuts, but within OMB, staff were directed to use the PART results to inform whatever recommendations seemed appropriate, given the larger circumstances of the program, which might indicate that a funding increase or management reform effort to improve program outcomes was the preferable course of action.

3.3.a Reporting to the Public

A key feature of the PART was the transparency built into its design. From the outset, completed assessments were made available to the public at [www.ExpectMore.gov](http://www.ExpectMore.gov) (see Figure 5), where one could locate programs by subject matter, agency, and performance rating. In addition to the completed, detailed PART assessment, which included scores by section and performance indicators with targets and actual achieved values, the Web site contained a one-page summary report on each program that described the program, the program’s PART rating, plans for improvement, and a link to the program’s own Web site.
Plans for program improvement during the following year were developed collaboratively by the agency and OMB during the assessment process. Their type and scope varied; some focused on one or two specific areas needing improvement, and others could be much broader. Typical plans included developing new outcome indicators, improving use of performance indicators, and conducting new evaluations. Most improvement plans focused on strengthening performance management or evaluating a project rather than improving actual outcomes. In a few cases where the program duplicated a better program or had fulfilled its original purpose, the plan might have been to try to persuade Congress to terminate the program’s funding.
3.3.b Performance Improvement Council

Near the end of his tenure, President Bush issued an Executive Order on improving government performance that formally established an interagency Performance Improvement Council (PIC) comprised of agency performance improvement officers appointed to the position by the agency head. These officers were to be responsible for coordinating the performance management activities of the agency, including development and improvement of strategic plans and annual performance plans and annual performance reports. They also were to ensure use of performance information in agency budgets and in agency personnel performance appraisals, particularly for program managers, in order to promote accountability for program effectiveness.

The PIC itself was envisioned as a collaborative and consultative body that would work with OMB to establish program performance standards and evaluation criteria, exchange information among agencies, coordinate and monitor performance assessments, keep the public informed, obtain information and advice from stakeholders, and make policy recommendations.

---

3.4 Obama Administration’s Accountable Government Initiative

President Obama has promised a government that holds itself accountable for results. With leadership from OMB, long the central point of coordination for federal management reforms, the Obama administration is introducing several performance initiatives that variously differ from those of its predecessors. Although it is too early to assess the initiatives’ results, it is possible to describe their dimensions and to explain how they either continue or differ from earlier attempts to make government programs more performance based, data driven, and results oriented.

In February 2009, the first indications of how the new administration’s approach to performance-based management would work were outlined in the proposed 2010 budget. Under the rubric of “putting performance first,” that document said that a focused team would be created in the White House to work with agency leaders and OMB to improve results and outcomes for federal government programs while eliminating waste and inefficiency. The team would be headed by a new Chief Performance Officer, who would work with federal agencies to set tough performance targets and hold managers responsible for progress. The President himself would meet regularly with cabinet officers to monitor their agencies’ progress toward meeting performance improvement targets.

The budget also said that the administration would “fundamentally reconfigure” the PART to open up the performance measurement process to the public, the Congress, and outside experts and to set program performance goals based on congressional intent and feedback from program constituencies. Programs would be assessed in the context of other programs serving the same population or meeting the same goals.

More specifically, the new administration committed to replace the PART with a new performance improvement and analysis framework that would switch the government’s performance management focus from grading programs as successful or unsuccessful to requiring agency leaders to set priority goals, demonstrate progress in achieving goals, and explain performance trends. Cross-program and cross-agency goals would receive as much or more focus as program-specific ones, and the administration would engage the public, Congress, and outside experts to develop a better and more open performance measurement process. As a first step, OMB would ask each major agency to identify a limited set of high-priority goals, supported by meaningful measures and quantitative targets, and would identify ongoing opportunities to engage the public, stakeholders, and Congress in this effort.

In addition, the administration proposed:

- Placing more emphasis on program evaluation.
- Reforming program assessment and performance measurement processes to emphasize the reporting of performance trends, explanations for the trends, mitigation of implementation risks, and plans for improvement with accountable leads.

• Streamlining reporting requirements under GPRA and PART to reduce the burden on agencies and OMB.
• Improving the communication of performance results to Congress, the public, and other stakeholders through better data display in agency reports and on the Internet.
• Launching a comprehensive research program to study the comparative effectiveness of different program strategies to ensure that programs achieve their ultimate desired outcomes.

In June 2009, the OMB Director sent a memorandum to federal agencies outlining guidance for preparation of the 2011 budget and “performance plans.” The memorandum asked agency heads to identify and commit to three to eight priority goals of high value to the public and strategies for achieving them. The goals should be ambitious but realistic, be attainable in 18 months to two years without new resources or legislation, and have well-defined, outcomes-based measures of progress that would be reviewed by OMB and initiative leaders on a quarterly basis. The goals also were to present significant management challenges, including across agency boundaries, and to require “a concerted focus of agency resources” for their attainment.

Agencies were encouraged to consult widely with their employees, Congress, and other stakeholders in developing the goals. Proposals were due at the end of July 2009, more than a month before their budget requests were due, and for each one, the agency was to identify the problem to be addressed, the goal, measures with targets and timeframes, contributing programs inside and outside the agency, persons in and out of the agency to be accountable, and strategies for achievement.

The fiscal 2011 budget introduced in February 2010 described more than 100 performance goals across two dozen agencies that their top leaders were committed to achieving. OMB stated, “By making agencies’ top leaders responsible for specific goals that they themselves have named as most important, the administration is dramatically improving accountability and the chances that government will deliver results on what matters most” (U.S. OMB, 2010b, 74ff).

To increase transparency and public understanding and confidence in government operations, OMB announced that the administration would identify and eliminate performance measurements and documents that are not useful and used and make those that remained much more understandable and accessible. In particular, it said that the administration would create an online federal performance dashboard or portal, www.performance.gov, that provides a “clear, concise picture of federal goals and measures by theme, by agency, by program, and by program type,” with links to similar dashboards supporting other important management functions throughout the government (U.S. OMB2010b, 74).

In September 2010, the OMB’s Deputy Director for Management told senior federal managers that the administration would use Performance.gov as a “one-stop shop for federal

performance information” that would “provide in-depth information on agency priority goals and key performance indicators, measures, and milestones… and help create the clarity and the culture of accountability required to create meaningful improvement” (Zients 2010). The administration would use the site to inform “regular, data-driven program reviews focused on achieving rapid results” both to achieve near-term improvements and to generate the momentum needed for ongoing improvement (Zients 2010). Performance.gov went online to the public in August, 2011; a copy of its homepage is shown as Figure 7 below, and below that, Figure 8 shows a copy of the introduction to the “Performance Improvement” section of the Web site.
Driving Federal Performance

“If we believe the government can make a difference in people’s lives, we have the obligation to prove that it works – by making government smarter, and leaner and more effective…”

PRESIDENT BARACK OBAMA
APRIL 13, 2011

Responding to the President’s challenge to cut waste, save money, and better serve the American people, Performance.gov provides a window on the Obama Administration’s approach to improving performance and accountability. Performance.gov shows progress on the Administration’s efforts to create a government that is more effective, efficient, innovative, and responsive.

Reforming how Washington works is an ongoing effort that demands vigilance and leadership. The Administration is strongly committed to investing in what works and fixing or cutting what does not. As part of this effort, the Administration is leading the “Campaign to Cut Waste,” an initiative to eliminate wasteful spending and get the most from taxpayer dollars throughout the government. Whether driving progress on top priorities or cutting the cost of delivery, government leaders at all levels are accountable for achieving results. We invite your ideas and suggestions.

Featured Story: Disposing of Excess Property

The government is identifying ways to reduce excess and underutilized properties. This map helps to depict excess properties identified across the country. Learn More

Areas of Focus

- Acquisition
- Financial Management
- Human Resources
- Technology
- Performance Improvement
- Open Government
- Sustainability
- Customer Service

Explore metrics on the government’s efforts to:
- Save money on contracting
- Decrease contracting risk
- Develop the acquisition workforce

Additionally, we have information on efforts to expand strategic sourcing.

Learn More
The budget also announced that the administration would build on the existing PIC and the capabilities of new information technologies to make extensive use of new and existing networks in and out of government to solve problems and share and diffuse knowledge. It said that the PIC would function as the hub of the performance management network and that OMB would work with it to create and advance a new set of federal performance management principles, refine a governmentwide performance management implementation plan, and identify and tackle specific problems as they arise. The PIC would also serve as a home for federal communities of practice that would develop tools and provide expert advice and assistance on effective performance management practices.

3.5 Program Evaluation

U.S. government agencies have undertaken in-depth evaluations of individual federal programs for decades. But while most agencies use performance measurement to track progress toward program goals, few agencies regularly conduct rigorous independent, expert evaluations of their programs’ effectiveness (U.S. GAO 2011d, p. 1). In addition, until very recently, there has been no centrally driven initiative to ensure that all agencies conduct such evaluations or have the capacity to produce them. Answers to a PART item on agency evaluations revealed wide variation in agency evaluation practices and in the ability and inclination of agencies to conduct and use quality evaluations. The PART encouraged agencies to carry out high-quality, independent evaluations of their programs’ effectiveness, and singled out randomized controlled trials as the preferred methodology, and program
improvement plans often called for the conduct of such studies, but there were few guidelines and no real requirements for their conduct and use (Metzenbaum 2010, p. 7 ff).

3.5.a New Initiatives to Promote Evaluation

The Obama administration’s fiscal 2011 budget launched a new push for rigorous evaluation with an allocation of approximately $100 million to support such strong evaluations, and the fiscal 2012 proposed a similar suite of evaluations and efforts to strengthen agency capacity to do them. OMB said that although performance measurement was critical to improving performance, it often could not “conclusively answer questions about how outcomes would differ in the absence of a program or if a program had been administered in a different way.” The budget said that the administration is committed to building an evaluation infrastructure that complements and is integrated with performance measurement and management.

The 2011 and 2012 budgets noted that President has made it very clear that policy decisions should be driven by evidence—evidence about what works and what does not and evidence that identifies the greatest needs and challenges. As a result, they described how the administration is using a new evaluation-based funding approach for several grant-based initiatives (U.S. OMB 2010, 2011, pp. 84-85). Under this three-tiered approach more money is first proposed for programs and practices that generate results backed up by strong evidence; for a second group of programs with some supportive evidence, funding is proposed on condition that the programs will be rigorously evaluated. If the programs are unsuccessful, their funds will be directed elsewhere. Finally, the approach encourages agencies to innovate and to test ideas with strong potential as indicated by preliminary research findings or reasonable hypotheses. The administration assumes that as more programs come forward with strong supporting evidence, they will be motivated toward further performance improvement. The administration also intends to work with agencies to make information on program impact evaluations available on line.

Figure 9 below provides a summary comparison of the PART and the Priority Goals Initiative.

**Figure 9. Comparison of PART and Priority Goals Initiative**

<table>
<thead>
<tr>
<th>Dimension</th>
<th>PART</th>
<th>Priority Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>To assess programs in a systematic, consistent, and transparent manner so that the Federal government can improve program performance.</td>
<td>To use specific challenging agency goals to focus performance improvement and cut costs in areas of importance to the administration.</td>
</tr>
<tr>
<td>Method</td>
<td>Agencies use online questionnaire to record PART answers and evidence, performance data, and follow-up actions.</td>
<td>Agencies report at least quarterly on progress measurement, diagnostic analysis, and senior management review and follow-up.</td>
</tr>
<tr>
<td>Basis for selection</td>
<td>Virtually all Federal programs were assessed. In practice, about 200 assessments were done</td>
<td>Agencies chose 5–8 priority goals achievable within 18–24 months and dependent on strong execution</td>
</tr>
<tr>
<td><strong>annually, selected by OMB and the agency; Choice based largely on policy importance. Potential for reassessment every five years.</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rather than additional resources or legislation.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Scope of analysis</strong></td>
<td><strong>Quarterly agency and OMB review of progress toward goal on key measures and milestones and revision of action plans as necessary. Goals and deadlines can be adjusted upon review.</strong></td>
<td></td>
</tr>
<tr>
<td>All factors that affect and reflect program performance, including program purpose and design, performance measurement, evaluations, and strategic planning, program management, and results.</td>
<td>Each agency has designated a senior accountable official, a Goal Leader, to drive progress on each priority goal.</td>
<td></td>
</tr>
<tr>
<td><strong>Leadership involvement</strong></td>
<td><strong>Frequency of review</strong></td>
<td></td>
</tr>
<tr>
<td>Limited involvement by individual program managers.</td>
<td>Consolidated quarterly review at overall agency level with OMB.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Goal by goal quarterly review with OMB.</td>
<td></td>
</tr>
<tr>
<td><strong>Outputs</strong></td>
<td><strong>Follow-up actions</strong></td>
<td></td>
</tr>
<tr>
<td>Annual program ratings based on weighted scoring of PART responses: Effective, Moderately Effective, Adequate, Ineffective, and Results Not Demonstrated.</td>
<td>All programs must have Improvement Plans based on assessment findings and must report on results. Overall agency efforts monitored quarterly in context of larger Performance Improvement Initiative.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Goal Leaders report progress on their goals on Performance.gov quarterly to support cross-agency coordination and learning and to inform review by senior agency management and OMB.</td>
<td></td>
</tr>
<tr>
<td><strong>Public information dissemination</strong></td>
<td><strong>Public information dissemination</strong></td>
<td></td>
</tr>
<tr>
<td>ExpectMore.gov provided public access to detailed PART assessments and summaries.</td>
<td>Performance.gov went online in August 2011. Currently displays only high-level overview of the initiative, promising more detail in 2012, as required under the GPRAMA.</td>
<td></td>
</tr>
</tbody>
</table>

### 4. IMPACTS OF RECENT U.S. MONITORING AND EVALUATION EFFORTS

Widespread and persistent public concerns about government deficits, a lack of confidence in government, and calls for more accountability helped create a broad climate of support for the kinds of results-oriented and performance-based reforms, of which GPRA and PART are probably the best known and most substantial initiatives to date. GAO has argued that such efforts can produce information essential to assess agency performance, improve management decision making, and hold programs accountable for results (U.S. GAO 2005a).
Despite nearly two decades of experience with these performance management tools, however, the record of success at “ensuring that performance information is both useful and used” is relatively modest (U.S. GAO 2008a, p. 2). Among its many reports on GPRA, a 2004 GAO review of the law’s implementation found that:

- The law had established a statutory foundation for results-oriented management in federal agencies.
- The data collected and reported permits agencies to make better-informed decisions and to make the results of their operations more visible to the public.
- The law made the relationships between program resources and results clearer.
- The law improved the bases for assessing program performance through improvements in strategic and annual plans and performance reports (U.S. GAO 2004b).

Although GAO found that agencies were beginning to use performance data to inform budget formulation, it also found that:

- GPRA information was not a significant driver of program funding decisions.
- GPRA reports were not as clear and usefully constructed as desired.
- Data quality issues, such as timeliness, meaning, and attribution of impacts, remained.
- Top leadership support and commitment to a focus on results remained inconsistent across the government.

GAO believed that transformations in organizational culture to make agencies “more results-oriented, customer-focused, and collaborative” could only be achieved through human capital management strategies that used performance management systems to align “institutional, unit, and employee performance” (U.S. GAO 2004b, p. 102). GAO said that such institutional change requires investments in the skills necessary to implement it, including skills in rigorous program evaluation methodology, and it faulted agencies for inconsistent support of such efforts.

GAO also called for:
- OMB leadership and support for transition to a results-oriented federal government.
- Reports more tailored to the information needs of those who should be concerned with results-oriented management.
- More meaningful outcome-oriented performance goals and measures
- More in-depth program evaluations.
- A long-term governmentwide performance integrating federal programs across agencies to replace the annual budget as OMB’s main planning document.

GAO’s 2007 survey, possibly reflecting the effects of both GPRA and the PART, found that managers had access to much more performance data than a decade before in 1997. Significantly more managers reported having the specific kinds of performance data required by GPRA and PART—measures of output, efficiency, customer service quality, and outcomes. The proportion of program managers who reported using performance data to a great or very great extent for such purposes as setting program priorities, allocating
resources, and setting employees’ job expectations, however, remained at about half; this lack of progress raised questions about the real impact of the PART (U.S. GAO 2008a, 1–7).

In contrast, GAO did find significant differences among agencies. Where data usage was greater, managers attributed that behavior to agency leaders’ commitment to achieving results and using performance management principles to do so, as well as to other management improvements, including investment in staff capacity for using performance information and in information technology to support them (U.S. GAO 2009a, 2009b). One analysis of the GAO surveys also found that, regardless of size, agencies with a more focused mission and with greater Congressional interest were more likely to make greater use of performance data (Ellig 2010).

The Bush administration framed the PART initiative as an improvement on the GPRA that would provide a more thorough, objective, and transparent basis for understanding federal programs and their impacts; for making comparisons among them; and for informing program management, improvement, and resource allocation decisions. GAO found that PART contributed to improvements in executive budget formulation and helped to strengthen OMB’s use of performance information for internal program analysis and budget review (U.S. GAO 2007).

PART results also were intended to be used to hold managers accountable for program improvements flowing from the assessments and to justify subsequent budget requests, management decisions, and legislative proposals. Detailed and summary reports on all individual assessments were posted on www.ExpectMore.gov to increase public awareness of government performance, program goals, achievements, and plans for improvements, as well as to stimulate demand for even better and timelier performance information (US Government 2008) Whether the Obama “new performance improvement and analysis framework,” including more frequent and focused reviews of performance trends of a much smaller number of priority goals supported by the performance.gov Web site, will significantly improve the use and usefulness of performance measurement systems remains to be seen (U.S. OMB 2010a).

Following is a summary of the key strengths of the U.S. federal government’s M&E efforts to date in the Executive Branch and of the challenges to their effectiveness that remain.

4.1 Strengths

- Large amounts of performance data are being collected. Most government agencies are collecting relevant performance data; most are setting targets. More evaluations are being conducted, including rigorous experimental design evaluations.

- Performance orientation is strongly embedded in the institutional framework. In addition to continued collection of input and process measures, the identification of quantitative program performance goals and the gathering of quantitative outcome, impact, or effectiveness measures has grown.

- The importance of analyzing and using performance data to manage programs and to justify policy and resource allocation decisions is widely acknowledged. Congress recently
enacted the GPRA Modernization Act of 2010, which should strengthen these practices, and the Obama administration initiatives to build a high-performing government should do so as well.

- Recognition of the need for transparency and accountability for performance has been a hallmark of both the Bush and the Obama administrations. Most government agencies include performance indicators and values on their Web sites. Both administrations have created sites to provide the public with better information about the performance of the governments’ programs. The need to conduct and use the results of many more high-quality performance evaluations across the government has been explicitly recognized and made a policy priority, as has the need to develop agencies’ capacity to plan and manage such studies. Fortunately, the United States has a wealth of university-trained and experienced researchers and research institutions available to carry out such evaluations in most, if not all, policy areas.

- Both administrations have recognized the potential of the interagency Performance Improvement Council to help promote best practices and address common concerns, including making better use of performance and evaluation data for program management and decision-making.

4.2 Challenges

- Focused, consistent analysis and use of performance data is not yet a significant focus of routine management or policy decision making. Many agencies still lack the resources, ability, and—most importantly—the leadership commitment to analyze the data collected and apply that analysis for policy, management, and accountability purposes. GAO has argued that “building analytic capacity to use performance information and to ensure its quality—both in terms of staff trained to do the analysis and availability of research and evaluation resources—is critical” to meaningful use of performance information (U.S. GAO 2005, p. 26, 2008a, p. 15).

- The use of performance information and evaluation findings by Congress has been limited. OMB and Executive Branch agencies could make better efforts to determine how this information can be made useful to Congress. More could be done to solicit congressional input in the development and production of performance-related information products to increase congressional interest in them. Congress has played a significant role in establishing requirements for monitoring performance, but it could make much greater use of the products of those mandates for oversight, resource allocation, and other decision making.

- Data collection is too often perceived as onerous and not useful. In some instances, the burden of compliance with data collection requirements overwhelms staff resources that might otherwise be available for analysis. Too often, new data requirements have been added to old ones rather than efficiently and effectively integrated with them.

- The potential for using performance information to look at cross-cutting issues or groups of programs that address similar outcomes has been largely ignored, as reflected in the limited use of comparative program analysis to enhance effectiveness and efficiency, even within agencies like OMB and GAO, which have governmentwide oversight responsibilities.
GAO’s recent report, *Opportunities to Reduce Potential Duplication in Government Programs, Save Tax Dollars, and Enhance Revenue*, outlined many areas where such analyses could be fruitfully pursued (U.S. GAO 2011b).

- Wider acceptance of M&E efforts has often been stymied, because they are often perceived as potentially punitive rather than constructive activities. Collaborative stakeholder participation, including Congress, in the design, analysis, and use of M&E data could help build more support for them.

- Citizen involvement has been limited. Despite the emphasis on performance and monitoring information and the widespread availability of new technologies like the Internet, greater efforts are needed to make high-quality performance information accessible and appealing to the public. For example, few agencies use online methods like citizen surveys to elicit direct feedback from beneficiaries or discussion forums to encourage stakeholder dialog on issues of common concern.

5. LESSONS LEARNED

Based on the U.S. experience, it is possible to draw lessons that can inform the efforts of other countries seeking to establish sound systems for M&E.

**Focus on a few key issues.** Don’t get spread too thin. As GAO has suggested and the Obama administration’s priority goal initiative reflects, consider cross-cutting reviews of a smaller number of high-priority issues that address a range of programs and policies potentially impacting goal achievement in an important issue area (U.S. GAO 2008a).

**Leadership with the right message is essential to success.** Leadership commitment to focusing on results and to using performance information to improve services, as well as to communicate more effectively with the legislature, is vital. However, the fact that the use of performance data still lags significantly behind the definition of performance indicators and collection of data suggests that the leadership’s message needs to be crafted in ways that motivate everyone to perform better. Otherwise, the performance management system can be viewed as a dispiriting tool of top management to keep tabs on and potentially penalize managers and staff. The message about why a performance management system is being created ought to stress management and service improvement, along with the added benefit of providing a framework to enhance accountability and transparency.

**A central coordinating unit helps keep the entire effort on track.** To plan and guide its performance management reform efforts, OMB has had a small, full-time central management team, supplemented by advice and feedback from other OMB and agency staff and overseen by at least one senior policy official. The structure has been essential to keeping the effort focused, coordinated with other government decision processes like formulation of the President's budget, and implemented consistently across multiple agencies. Individual agencies, similarly, have usually created an internal body to perform the same functions.
High-level, performance-focused interagency mechanisms like the PIC can be an effective means to inform members about and encourage them to adopt best practices and to engage in useful information sharing and problem solving. The Obama administration intends to make even greater use of the PIC as vehicle for capacity building, identification of best practices, and joint problem solving. The use of online technologies for communication and information sharing can be a particular aid to the ultimate effectiveness of such efforts. The PIC can be particularly helpful in tackling common systemic issues like improving agency capacity to plan and use quality evaluations and improving communication of performance information with Congress and other key stakeholders.

Train managers and staff how to analyze and use the data effectively early in the implementation process. In developing performance management systems, the early phases need significant effort to define goals and objectives and identify performance indicators by which progress can be measured and tracked. The process takes a longer time when stakeholders are brought into this part of the process, but this step should not be ignored. The core benefit of the system—using the data to improve performance—too easily can be overwhelmed by the data collection process. Focusing on a limited set of the most important data can help avoid this problem. Looking early at how others have successfully tackled these challenges can be very helpful. An ongoing investment in capacity building and the infrastructure required to identify good measures, collect the data, and use and report on performance is vital for the effectiveness of the system.

Plans for program evaluation need to be built into program design at the beginning. Managers need to be trained in the design and use of evaluations and to learn the importance of incorporating plans for their conduct at the outset of program operations to guide program and managerial priorities and information design, as well as to assess program impacts after implementation. Expert, independent professional associations like the American Evaluation Association in the United States, which presented the incoming Obama Administration with “An Evaluation Roadmap for a More Effective Government,” (AEA 2009), are a valuable resource that the government could use to help develop solid strategies for implementing such efforts.

Keep it simple. Avoid requirements to collect and report on too many indicators. Otherwise, the whole effort can become a burdensome compliance exercise rather than an effective management tool. It is important not to overload agencies. Both PART and GPRA have been criticized heavily for that (U.S. GAO 2009d). However, it is important not to discourage programs from tracking any indicators that they truly find useful.

Creative but carefully considered and managed use of technology can be very helpful in collecting and analyzing data and in communicating with stakeholders and others. For example, a significant spillover benefit of technologies like the Internet is the ability to share experiences and best practices with others, which can help them to move up the learning curve and implement constructive innovations much faster than ever before.

Performance communications must be designed to inspire confidence in the results. Thus, early on, a significant effort should be made to identify the performance information needs of
key stakeholders, like Congress, program staff, and delivery partners, and to make that information timely, accessible, and readily understandable (Metzenbaum 2010, p. 7). Training for all stakeholders on how such information could be used in decision making would also be very helpful (Metzenbaum 2010, p. 21).

To legislate or not to legislate? A balance should be sought between legislating the right things and making the framework too prescriptive and burdensome. Overly prescriptive legislation could stifle managerial creativity, flexibility, and innovation. Requirements for data collection and report production need to be balanced by emphases on data analysis and use and on effective and timely communication of results to stakeholders. Performance measurement lends itself well to oversight functions, but it also has an important role to play in motivating managers and front-line supervisors to focus on results and the impact of their work on citizens’ lives. GPRA was a major motivator for getting the U.S. government moving in the direction of results-oriented government. At the same time, the law’s administrative demands on each agency dampened enthusiasm and did not produce the degree of improvement in the analysis and use of data that advocates of performance-based decision making continue to strive for.
Bibliography


<table>
<thead>
<tr>
<th>#</th>
<th>Author</th>
<th>Year</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Keith Mackay</td>
<td>1998</td>
<td>Lessons from National Experience.</td>
</tr>
<tr>
<td>#3</td>
<td>Alain Barberie</td>
<td>1998</td>
<td>Indonesia’s National Evaluation System.</td>
</tr>
<tr>
<td>#4</td>
<td>Keith Mackay</td>
<td>1998</td>
<td>The Development of Australia’s Evaluation System.</td>
</tr>
<tr>
<td>#5</td>
<td>R. Pablo Guerrero O.</td>
<td>1999</td>
<td>Comparative Insights from Colombia, China and Indonesia.</td>
</tr>
<tr>
<td>#6</td>
<td>Keith Mackay</td>
<td>1999</td>
<td>Evaluation Capacity Development: A Diagnostic Guide and Action</td>
</tr>
<tr>
<td></td>
<td>Framework</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#7</td>
<td>Mark Schacter</td>
<td>2000</td>
<td>Sub-Saharan Africa: Lessons from Experience in Supporting Sound</td>
</tr>
<tr>
<td></td>
<td>Governance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#8</td>
<td>Arild Hauge</td>
<td>2001</td>
<td>Strengthening Capacity for Monitoring and Evaluation in Uganda: A</td>
</tr>
<tr>
<td></td>
<td>Results Based Management</td>
<td></td>
<td>Perspective.</td>
</tr>
<tr>
<td>#10</td>
<td>Arild Hauge</td>
<td>2003</td>
<td>The Development of Monitoring and Evaluation Capacities to Improve</td>
</tr>
<tr>
<td></td>
<td>Government Performance</td>
<td></td>
<td>in Uganda.</td>
</tr>
<tr>
<td>#11</td>
<td>Keith Mackay</td>
<td>2004</td>
<td>Two Generations of Performance Evaluation and Management System in</td>
</tr>
<tr>
<td></td>
<td>Australia.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#12</td>
<td>Adikeshavalu Ravindra</td>
<td>2004</td>
<td>An Assessment of the Impact of Bangalore Citizen Report Cards on the</td>
</tr>
<tr>
<td></td>
<td>Performance of Public Agencies.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#15</td>
<td>Keith Mackay</td>
<td>2006</td>
<td>Institutionalization of Monitoring and Evaluation Systems to Improve</td>
</tr>
<tr>
<td></td>
<td>Public Sector Management</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#16</td>
<td>Ariel Zaltsman</td>
<td>2006</td>
<td>Experience with Institutionalizing Monitoring and Evaluation Systems In Five Latin American Countries: Argentina, Chile, Colombia, Costa Rica and Uruguay</td>
</tr>
<tr>
<td>#17</td>
<td>Keith Mackay and others</td>
<td>2007</td>
<td>A Diagnosis of Colombia’s National M&amp;E System, SINERGIA.</td>
</tr>
</tbody>
</table>
#19: Rafael Goméz, Mauricio Olivera, and Mario A. Velasco. 2009. *Implementing a Subnational Results-Oriented Management and Budgeting System: Lessons from Medellín, Colombia.*


