United States Government Accountability Office

GAO

Report to the Chairman, Subcommittee on Government Management, Finance, and Accountability, Committee on Government Reform, House of Representatives

October 2005

PERFORMANCE BUDGETING

PART Focuses Attention on Program Performance, but More Can Be Done to Engage Congress

GAO
Accountability ★ Integrity ★ Reliability

PERFORMANCE BUDGETING

PART Focuses Attention on Program Performance, but More Can Be Done to Engage Congress

Why GAO Did This Study

GAO was asked to examine (1) the Office of Management and Budget's (OMB) and agency perspectives on the effects that the Program Assessment Rating Tool (PART) recommendations are having on agency operations and program results; (2) OMB's leadership in ensuring a complementary relationship between the PART and the Government Performance and Results Act of 1993 (GPRA); and (3) steps OMB has taken to involve Congress in the PART process. To do this, we also followed up on issues raised in our January 2004 report on the PART.

What GAO Recommends

GAO recommends that OMB solicit congressional views on the performance issues and program areas most in need of review; the most useful performance data and the presentation of those data; and select PART reassessments and crosscutting reviews based on factors including the relative priorities, costs, and risks associated with clusters of related programs, and reflective of congressional input. GAO also suggests that Congress consider a structured approach to articulating its perspective and oversight agenda on performance goals and priorities for key programs.

OMB generally agreed with our findings, conclusions, and recommendations, and outlined several actions it is taking to address some of the issues raised in the report.

www.gao.gov/cgi-bin/getrpt?GAO-06-28.

To view the full product, including the scope and methodology, click on the link above. For more information, contact Susan J. Irving at (202) 512-9142 or irvings@gao.gov.

What GAO Found

The PART process has aided OMB's oversight of agencies, focused agencies' efforts to improve program management, and created or enhanced an evaluation culture within agencies. Although the PART has enhanced the focus on performance, the PART remains a labor-intensive process at OMB and agencies.

Most PART recommendations are focused on improving outcome measures and data collection, and are not designed to result in observable short-term performance improvements. Since these necessary first steps on the path to long-term program improvement do not usually lead to improved short-term results, there is limited evidence to date of the PART's influence on outcome-based program results. Moreover, as of February 2005—the date of the most recent available OMB data—the majority of follow-on actions have not yet been fully implemented. By design OMB has not prioritized them within or among agencies. Because OMB has chosen to assess nearly all federal programs, OMB and agency resources are diffused across multiple areas instead of concentrated on those areas of highest priority both within agencies and across the federal government. This strategy is likely to lengthen the time it will take to observe measurable change compared with a more strategic approach. OMB has used the PART as a framework for several crosscutting reviews, but these have not always included all relevant tools, such as tax expenditures, that contribute to related goals. Greater focus on selecting related programs and activities for concurrent review would improve their usefulness.

OMB has taken some steps to clarify the PART-GPRA relationship but many agencies still struggle to balance the differing needs of the budget and planning processes and their various stakeholders. Unresolved tensions between GPRA and the PART can result in conflicting ideas about what to measure and how to measure it. Finally, we remain concerned that the focus of agencies' strategic planning continues to shift from long-term goal setting to short-term executive budget and planning needs.

OMB uses a variety of methods to communicate PART results, but congressional committee staff we spoke with had concerns about the tool itself, how programs were defined, and the usefulness of goals and measures. Most said that the PART would more likely inform their deliberations if OMB consulted them early on regarding the selection and timing of programs; the methodology and evidence to be used; and how PART information can be communicated and presented to best meet their needs. It is also important that Congress take full advantage of the benefits arising from the executive reform agenda. While Congress has a number of opportunities to provide its perspective on specific performance issues and performance goals, opportunities also exist for Congress to enhance its institutional focus to enable a more systematic assessment of key programs and performance goals.

_____ **United States Government Accountability Office**

Contents

October 28, 2005

The Honorable Todd R. Platts
Chairman
Subcommittee on Government Management,
Finance, and Accountability
Committee on Government Reform
House of Representatives

Dear Mr. Chairman:

Given current trends and challenges facing the nation—including the federal government's long-term fiscal imbalance—we must take advantage of opportunities to enhance performance, ensure accountability, and position the nation for the future. To this end, it is critical to reexamine the relevance of federal programs and their fit with national priorities, while maximizing program performance within current and expected resource levels.[1]

In the 1990s, Congress and the executive branch laid out a statutory and management framework that provides the foundation for strengthening government performance and accountability, with the Government Performance and Results Act of 1993[2] (GPRA) as its centerpiece. GPRA is designed to inform congressional and executive decision making by providing objective information on the relative effectiveness and efficiency of federal programs and spending. A key purpose of the act is to create closer and clearer links between the process of allocating scarce resources and the expected results to be achieved with those resources—also known as performance budgeting. Performance budgeting holds promise for reexamining existing federal government programs and aligning them with current needs. It can increase the government's capacity to assess competing claims for federal dollars and has the potential to enrich the budget debate by arming decision makers with credible information on both individual programs and the relationship between policies, programs, and other tools designed to address related goals.

---

[1] For more information on reexamination of federal programs, see GAO, *21st Century Challenges: Reexamining the Base of the Federal Government*, GAO-05-325SP (Washington, D.C.: February 2005).

[2] Pub. L. No. 103-62 (1993).

This administration has made budget and performance integration one of five governmentwide management priorities under its President's Management Agenda (PMA).[3] The Executive Branch Management Scorecard, which employs a red-yellow-green stoplight grading system, tracks how well departments and major agencies are executing the five governmentwide management initiatives. A central element in this initiative is the Office of Management and Budget's (OMB) Program Assessment Rating Tool (PART), which OMB describes as a diagnostic tool meant to provide a consistent approach to evaluating federal programs as part of the executive budget formulation process.

At your request, we reported in January 2004 on OMB and agency experiences during the first year of PART implementation.[4] Specifically, we examined how the PART changed OMB's decision-making process in developing the President's fiscal year 2004 budget request; the PART's relationship to GPRA; and the PART's strengths and weaknesses as an evaluation tool. Our findings and recommendations from the January 2004 report are detailed later in this report. In light of the challenges we identified, you asked us to examine (1) OMB's and agencies' perspectives on the effects the PART recommendations are having on agency operations and program results, including issues encountered in responding to the PART recommendations; (2) OMB's leadership in ensuring an integrated, complementary relationship between the PART and GPRA, including instances in which multiple programs or agencies are involved in meeting goals and objectives; and (3) steps OMB has taken to involve Congress in the PART process. To help us fulfill these objectives, we also followed up on the issues and challenges we identified during our January 2004 review. Lastly, we agreed to issue a separate report that explores (1) progress agencies have made in responding to OMB's PART recommendations that they obtain program evaluations, (2) what factors have facilitated or impeded agency progress toward obtaining the program evaluations that OMB recommended, and (3) whether the evaluations appear designed to yield the desired information on program results.[5]

---

[3]In addition to budget and performance integration, the other four priorities under the PMA are strategic management of human capital, expanded electronic government, improved financial performance, and competitive sourcing.

[4]GAO, *Performance Budgeting: Observations on the Use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget*, GAO-04-174 (Washington, D.C.: Jan. 30, 2004).

[5]GAO, *Program Evaluation: OMB's PART Reviews Increased Agencies' Attention to Improving Evidence of Program Results*, GAO-06-67 (Washington, D.C.: Oct. 28, 2005).

To address the first two objectives, we reviewed OMB materials on the implementation, application, and revision of the PART for calendar years 2002 through 2004.[6] We also interviewed OMB branch chiefs and OMB staff on the Performance Evaluation Team whose role is to provide guidance to budget examiners and help ensure consistent application of the PART across OMB offices. To better understand OMB's experience with crosscutting reviews, we interviewed OMB staff responsible for coordinating the Community and Economic Development and Rural Water crosscutting reviews conducted for the fiscal year 2006 President's budget request. To obtain agency perspectives on the relationship between the PART and GPRA and their interactions with OMB concerning that relationship, we interviewed department and agency officials, including senior managers, and program, planning, and budget staffs at (1) the Department of Health and Human Services (HHS), (2) the Department of Energy (DOE), (3) the Department of Labor (DOL), and (4) the Small Business Administration (SBA). We also interviewed officials from these departments and agencies concerning their perspectives and activities in response to the PART recommendations and the effects of implementing those recommendations on operations and results.

We selected these three departments and one independent agency for a number of reasons. Collectively, they offered examples of all seven of the PART program types (e.g., block/formula grants, competitive grants, direct federal, and research and development) for review. These examples covered about a fifth of all the programs subject to the PART as of 2004 and thus could provide a broad-based perspective on how the PART was applied. Our selection of these four agencies was also influenced by our intent to integrate this work with our related work examining progress in addressing the PART program evaluation recommendations. Approximately half of the evaluation recommendations in the 2002 PART were encompassed in our four case selections.

As part of our work on the second objective, we also performed various analyses of the PART recommendations made in all 3 years to discern possible changes or trends in recommendations over time and relationships between the type of recommendations made, type of program, overall rating, the total PART score, and answers to selected PART questions. In addition, we also examined relevant OMB and agency

---

[6]This period covers fiscal years 2004-2006.

documents to help determine how recommendations are tracked and their impact evaluated by OMB and the selected agencies.

To address our third objective of examining the steps OMB has taken to involve Congress in the PART process, we interviewed OMB and agency officials and asked questions about the steps OMB and agencies have taken to involve Congress in the PART process or in using the results of the PART. To obtain documented instances of Congress's uses and views of the PART, we interviewed House and Senate committee staff (minority and majority) for the authorization and appropriations subcommittees with jurisdiction over our selected agencies as well as OMB and officials from the four selected agencies. Finally, we reviewed fiscal years 2005 and 2006 House and Senate congressional hearing records and reports as well as conference reports for mentions of the PART. In addition, where possible, we corroborated testimonial evidence with documentary evidence of OMB and agencies' strategies for involving the Congress as well as evidence of collaboration and coordination, such as planning documents, briefing materials, or other evidence of contact with Congress.

While our summary analyses include all or almost all programs subject to the PART for the years 2002 to 2004 or all or almost all programs within a specified subset of programs (e.g., program type, specific year, specific rating, etc.), the information obtained from OMB, congressional, and agency officials as well as documentary materials from the selected agencies is not generalizable to the PART process for all years or all programs. The information obtained from various officials reflects their views and experiences with the PART. However, the consistency and frequency with which officials identified certain issues and concerns in our prior review and in this one suggest that they are both significant and continuing aspects of the PART's use as a budget and evaluation tool.

We conducted our audit work from January 2005 through September 2005 in accordance with generally accepted government auditing standards. Detailed information on our scope and methodology appears in appendix I.

Results in Brief

The PART process continues to aid OMB's oversight of agencies and encourage improvements in executive budget formulation and agency program management. The PART has helped to structure and discipline OMB's use of performance information for internal program analysis and budget review, and made its use of this information more transparent. Many agency officials told us that the PART helped to create or strengthen

an evaluation culture within agencies by providing external motivation for program review and focused attention on performance measurement and its importance in daily program management. Some officials said that the PART and the PMA helped them move away from "analysis by anecdote" and refocused their attention on the impact their programs have, instead of largely on output measures. Others echoed a similar sentiment—one indicated that the PART scores helped to create "a new sense of urgency" about performance measures and completing the changes to performance systems that were already underway. Although the PART has enhanced the focus on performance, the PART remains a labor-intensive process at OMB and agencies.

As we reported in our January 2004 report, senior OMB managers recognized the increased workload the PART initially placed on examiners; however, they expected the workload to decline as both OMB and agency staff became more familiar with the PART tool and process, and as issues with the timing of the PART reviews were resolved. During this review, we found that while the learning curve appears to have flattened, it did not compensate for the time needed to assess and reassess programs. Although a senior OMB official acknowledged continuing capacity issues regarding the PART, he said that the PART is still a better way for examiners to accomplish their traditional program assessment responsibilities because it is more objective and transparent.

OMB has said that a primary purpose of the PART is program improvement. Our analysis supports OMB's statements that most of the PART recommendations to date were aimed at improving outcome and efficiency measures, and in collecting performance data. Implementing these recommendations can lay the groundwork for long-term program improvement by improving managers' ability to assess program outcomes, identify information gaps, and assess next steps—these are necessary first steps on the path to long-term program improvement, but are not expected to result in observable program improvement in the short term. According to OMB's most recent available data (February 2005), the majority of the PART recommendations have not yet been fully implemented; nearly half of the fully implemented recommendations are of the type described above. For these reasons, it is too soon to tell whether the PART has fully produced the intended results.

Although agency officials appreciated the flexibility OMB provided by not making prescriptive recommendations, some were so general that it was difficult for them to understand what change was required or how progress

could be measured. OMB has made a conscious decision not to prioritize the recommendations within each agency or across the more than 1,700 recommendations governmentwide. Because OMB has chosen to assess nearly all federal programs, OMB and agency resources are diffused across multiple areas instead of concentrated on those areas of highest priority both within agencies and across the federal government. This strategy is likely to lengthen the time it will take to observe measurable change compared with a more strategic approach.

Federal agencies are increasingly expected to demonstrate effectiveness in achieving agency or governmentwide goals. Our previous work has shown that the accuracy and quality of evaluation information necessary to make the judgments called for when rating programs is highly uneven across the federal government. To help explain linkages between program activities, outputs, and outcomes, program evaluation designs are tailored to address various types of programs and questions. For example, a process evaluation reviews various aspects of program operations to assess the quality or efficiency of services. Alternatively, an impact evaluation depends on scientific research methods to establish a causal connection between program activities and outcomes, and, in light of important influences in the program environment, to isolate the program's contributions to the observed outcomes.

OMB and agencies' expectations sometimes differed on the scope and purpose of evaluations conducted in response to the PART recommendations. Some of the difficulties seemed to derive from OMB expecting to find, in the agencies' external evaluation studies, the types of comprehensive judgments about program design, management, and effectiveness found in the PART assessments. Because evaluations designed for internal and external audiences often have a different focus, differences of opinion on the usefulness of evaluations is perhaps not surprising. This raises the strong possibility that—without consultation—the evaluations that agencies conduct may not provide OMB with the information it wants.

The PART's focus on individual programs can yield useful information. At the same time, it is also often critical to understand how individual programs fit within a broader portfolio of tools and strategies—such as regulations, credit programs, and tax expenditures—to accomplish federal

goals.[7] OMB reported on the two crosscutting PART assessments—Rural Water programs and Community and Economic Development (CED) programs—for the fiscal year 2006 budget and it plans to conduct three additional crosscutting reviews on block grants, credit programs, and small business innovation research for the fiscal year 2007 budget. Cognizant agency officials involved with the CED were pleased with the collaborative interagency process. They found value in leveraging existing efforts within agencies and benefited from OMB staff consultation. To date, however, crosscutting assessments have not always addressed all programs identified by OMB as relevant to a common goal. Also, little information on the CED crosscutting assessment was initially available beyond the brief description contained in the *Analytical Perspectives* volume of the Fiscal Year 2006 President's budget request. This caused some confusion among budget stakeholders, particularly Congress.

Both the PART and GPRA aim to improve the focus on program results. However, the different purposes and time frames they serve continue to create tensions. Some agencies have made progress over the past several years in reconciling the two processes, but OMB and agencies often have conflicting ideas about what to measure, how to measure it, and how to report program results. Although less of a problem than it was during our January 2004 review of the PART process, difficulties in defining a unit of analysis useful for both OMB's budget analysis and program management purposes remain. For example, agency officials said that the PART sometimes defined a "program" in overly narrow terms which did not provide an understanding of how the entire program works. Agency officials also noted there can be difficulties if a PART "program" is defined so broadly that unrelated or loosely related programs and/or programs with uneven success levels are combined.

Measuring outcomes has long been challenging, especially for certain types of programs, and can complicate efforts to arrive at goals useful to multiple parties for multiple purposes. During the PART process, this was particularly true for programs whose outcomes take many years to observe, areas in which the federal government is one of many actors, and basic research programs. Differences of opinion about efficiency measures

---

[7] For more information on tax expenditures and their role in achieving federal objectives, see GAO, *Government Performance and Accountability: Tax Expenditures Represent a Substantial Federal Commitment and Need to be Reexamined*, GAO-05-690 (Washington, D.C.: Sept. 23, 2005).

highlighted the challenges that can result from the difficult but potentially useful process of comparing the costs of programs related to similar goals. Although these issues are clearly not new or unique to the PART, they are aggravated by the difficulties of developing measures useful for multiple purposes and audiences and often remain a point of friction in agencies and sometimes within OMB.

The PART was designed for and is used in the executive branch budget preparation and review process; as such, the goals and measures used in the PART must meet OMB's needs. GPRA is a broader process involving the development of strategic and performance goals and objectives to be reported in strategic and annual plans. Compared to the relatively open-ended GPRA planning process any executive budget formulation process is likely to seem closed. Without truly integrating agency budget formulation and strategic planning in a way that considers the differing needs of these processes and their various stakeholders, OMB's ability to strengthen and further the performance-resource linkages for which GPRA laid the groundwork will continue to be hampered. Although OMB has made efforts to ensure that the PART and GPRA are complementary processes, we continue to find evidence that the closed nature of the executive budget formulation process may not allow for the type of stakeholder involvement in strategic and annual planning envisioned by GPRA. Because the executive budget formulation process—to which the PART belongs—is "predecisional," information required in agencies' annual GPRA plans is embargoed within the executive branch until the President's budget request is transmitted to Congress. Several of our case study agencies described instances in developing annual and long-term goals where interaction and/or consultation with key stakeholders was limited or absent. They said because of timing issues or the explicit instructions of OMB, they had to present new or revised program goals and measures to their stakeholders after the fact, and in some cases stakeholders disagreed with the goals but had no choice but to accept them. We remain concerned that the focus of agency strategic planning is shifting from long-term goal setting to short-term budget and planning needs.

OMB uses a variety of tools to communicate the PART assessment results to Congress, including the President's budget request documents, OMB's Web site, and meetings with some congressional staff. OMB said that they also sent packages including the PART summary sheets for programs that fell within each committee's jurisdiction and a list of the programs OMB planned to review for the fiscal year 2006 budget request to all relevant House and Senate committees, but received little response. Committee

staff told us that OMB's efforts have not been sufficient to convince Congress of the value of the PART. Most congressional committee staff we spoke with did not find either the PART information or the way it was communicated suited to their needs. Many had concerns about the usefulness of the goals and measures OMB used to assess program performance and some questioned the "units of analysis" used for the PART as well as the design of the tool itself. Most reported that they would be more likely to use the PART results to inform their deliberations if OMB consulted them early in the PART process regarding the selection and timing of programs to assess, the methodology and evidence used or to be used, and how the PART information can best be communicated and presented to meet their needs.

That said, OMB has sometimes engaged Congress when it has communicated the selected PART results through traditional means of signaling executive branch priorities, such as legislative proposals. Congress has also held hearings in response to the PART, and has mentioned the PART in committee reports. Clearly, certain PART reviews have captured congressional attention and contributed to the policy debate. While Congress has a number of opportunities to provide its perspective on specific performance issues and performance goals, opportunities also exist for Congress to enhance its institutional focus to enable a more systematic assessment of key programs and performance goals.

We make recommendations and suggestions in this report to OMB and Congress, respectively, to promote greater discussion with congressional stakeholders and emphasize the PART's role in assessing programs related to common outcomes. We believe that implementing these recommendations will help ensure an integrated, complementary approach between GPRA and the PART.

In commenting on a draft of this report, OMB generally agreed with our findings, conclusions, and recommendations. OMB outlined several actions it is taking to address some of the issues raised in the report, including implementing information technology solutions to make application of the PART less burdensome and more collaborative. OMB also suggested some technical changes throughout the report that we have incorporated as appropriate. OMB's comments appear in appendix IV. We also received technical comments on excerpts of the draft from the Departments of Labor and Health and Human Services, which are incorporated as appropriate.

Background

Federal interest in performance information and its potential relationship to budgeting practices has existed to varying degrees for over 50 years.[8] More recently, this interest culminated in the passage of GPRA and related management reforms of the 1990s. GPRA mandates that federal agencies develop performance information describing the relative effectiveness and efficiency of federal programs as a means of improving the congressional and executive decision-making processes. Among other statutory obligations, GPRA requires federal agencies to publish strategic and annual plans describing specific program activities with the intention of establishing a more tangible link between performance information for these programs and agency budget requests.[9]

The current administration has taken several steps to strengthen performance-resource linkages for which GPRA laid the groundwork. The budget and performance initiative of the PMA contains the criteria agencies must meet in order to achieve "green" status on the initiative. The criteria include elements relating to budgeting and strategic planning and also ties those elements to individual performance. As we have previously reported, creating a "line of sight" between individual performance and organizational success is a leading practice used in public sector organizations to become more results-oriented.[10]

Central to the budget and performance integration initiative of the PMA, the PART is a means to strengthen the process for assessing the effectiveness of programs by making that process more robust, transparent, and systematic. The PART is a series of diagnostic questions designed to provide a consistent approach to rating federal programs. (See

---

[8]For a detailed examination of previous federal performance budgeting initiatives, see GAO, *Performance Budgeting: Past Initiatives Offer Insights for GPRA Implementation*, GAO/AIMD-97-46 (Washington, D.C.: Mar. 27, 1997).

[9]5 U.S.C. § 306 (strategic plans); 31 U.S.C. § 1115 (performance plans). For more information on federal efforts to implement GPRA, see GAO, *Results-Oriented Government: GPRA Has Established a Solid Foundation for Achieving Greater Results*, GAO-04-38 (Washington, D.C.: Mar. 10, 2004).

[10]For more information on using performance management systems as a strategic tool to drive internal change and achieve desired results, see GAO, *Results-Oriented Cultures: Creating a Clear Linkage between Individual Performance and Organizational Success*, GAO-03-488 (Washington, D.C.: Mar. 14, 2003) and *Managing for Results: Enhancing Agency Use of Performance Information for Managerial Decision Making*, GAO-05-927 (Washington, D.C.: Sept. 9, 2005).

app. II for the PART questionnaire.) Drawing on available performance and evaluation information, OMB staff use the questionnaire to rate the strengths and weaknesses of federal programs with a particular focus on individual program results. The PART asks, for example, whether a program's long-term goals are specific, ambitious, and focused on outcomes, and whether annual goals demonstrate progress toward achieving long-term goals. It is designed to be evidence-based, drawing on a wide array of information, including authorizing legislation; GPRA strategic plans, annual performance plans, and reports; financial statements; inspectors general and GAO reports; and independent program evaluations. The PART questions are divided into four sections; each section is given a specific weight in determining the final numerical rating for a program. Table 1 shows an overview of the four PART sections and the weights OMB assigned.

In addition, each PART program is assessed according to one of seven major approaches to delivering services. Table 2 provides an overview of these program types and the number and percentage of programs covered by each type in the 2002 through 2004 performance assessments.

Table 1: Overview of Sections of the PART Questions

Section	Description	Weight (percentage)
I. Program Purpose and Design	To assess whether • the purpose is clear, and • the program design makes sense.	20
II. Strategic Planning	To assess whether the agency sets valid programmatic • annual goals, and • long-term goals.	10
III. Program Management	To rate agency management of the program, including • financial oversight, and • program improvement efforts.	20
IV. Program Results/Accountability	To rate program performance on goals reviewed in • the strategic planning section, and • through other evaluations.	50

Source: GAO.

Table 2: Overview of the PART Program Types, Cumulative Data 2002-2004

Program type	Description	Number/percentage of programs[a]		
		2002	2003	2004
1. Direct federal	Programs in which support and services are provided primarily by federal employees.	67 29%	117 29%	179 29%
2. Block/formula grant	Programs that distribute funds to state, local, and tribal governments and other entities by formula or block grant.	41 18%	71 18%	98 16%
3. Competitive grant	Programs that distribute funds to state, local, and tribal governments, organizations, individuals, and other entities through a competitive process.	37 16%	62 16%	105 17%
4. Capital assets and service acquisition	Programs in which the primary means to achieve goals is the development and acquisition of capital assets (such as land, structures, equipment, and intellectual property) or the purchase of services (such as maintenance and information technology) from a commercial source.	34 15%	43 11%	58 10%
5. Research and development	Programs that focus on creating knowledge or applying it toward the creation of systems, devices, methods, materials, or technologies.	32 14%	42 11%	54 9%
6. Regulatory based	Programs that employ regulatory action to achieve program and agency goals through rule making that implements, interprets, or prescribes law or policy, or describes procedure or practice requirements. These programs issue significant regulations, which are subject to OMB review.	15 6%	23 6%	32 5%
7. Credit	Programs that provide support through loans, loan guarantees, and direct credit.	4 2%	14 4%	23 4%
8. Mixed[b]	Programs that contain elements of different program types.	4 2%	27 7%	58 10%

Source: GAO analysis of the PART data.

[a]Percentages do not add to 100 percent because of rounding.

[b]OMB noted that in certain cases, drawing questions from two or more of the seven main PART program types—that is, creation of a "mixed" program type—yields a more informative assessment.

As of February 2005, the PART ratings have been published for 607 programs (according to OMB, this represents about 60 percent of the federal budget). Each program received one of four overall ratings: (1) "effective," (2) "moderately effective," (3) "adequate," or (4) "ineffective" based on program design, strategic planning, management, and results. A fifth rating, "results not demonstrated," was given— independent of a program's numerical score—if OMB decided that a program's performance information, performance measures, or both were insufficient or inadequate. Table 3 shows the distribution of ratings for 2002-2004. During the next 2 years, the administration plans to assess all remaining executive branch programs with limited exceptions.[11]

Table 3: The Cumulative PART Program Results by Rating Category, 2002-2004

Ratings	2002	2003	2004
Effective	6%	11%	15%
Moderately effective	24%	26%	26%
Adequate	15%	20%	26%
Ineffective	5%	5%	4%
Results not demonstrated	50%	38%	29%
Total programs	**234**	**399**	**607**

Source: OMB.

In our January 2004 report on the PART, you asked us to examine (1) how the PART changed OMB's decision-making process in developing the President's fiscal year 2004 budget request; (2) the PART's relationship to the GPRA planning process and reporting requirements; and (3) the PART's strengths and weaknesses as an evaluation tool, including how OMB ensured that the PART was applied consistently. We found that the PART helped structure OMB's use of performance information for its internal program and budget analysis, made the use of this information more transparent, and stimulated agency interest in budget and performance integration. Our analysis confirmed that one of the PART's major impacts was its ability to highlight OMB's recommended changes in program management and design. We noted that while much of the PART's potential

---

[11]The administration is considering alternative methods and timelines for assessing programs with limited impact and large activities where it is difficult to determine an appropriate unit of analysis (generally programs within the Department of Defense).

value lies in the related program recommendations, realizing these benefits would require sustained attention to implementation and oversight to determine if desired results are achieved, and that OMB needs to remain congnizant of this as it considers capacity and workload issues in the PART.

We also recognized that while there are inherent challenges in assigning a single rating to programs having multiple purposes and goals, OMB devoted considerable effort to promoting consistent ratings but challenges remain in addressing inconsistencies among OMB staff, such as interpreting the PART guidance and defining acceptable measures. OMB senior officials recently told us that inconsistencies in the PART process could also be attributed to agency staff, given the shared agency-OMB responsibilities in the PART process. Limited credible evidence on results also constrained OMB's ability to rate program effectiveness, as evidenced by the almost 50 percent of programs rated "results not demonstrated."

We also found that the PART is not well integrated with GPRA—the current statutory framework for strategic planning and reporting. We said that by using the PART process to review and sometimes replace GPRA goals and measures, OMB substituted its judgment for a wide range of stakeholder interests. The PART/GPRA tension was further highlighted by challenges in defining a unit of analysis useful for both program-level budget analysis and agency planning purposes. Although the PART can stimulate discussion on program-specific measurement issues, it cannot substitute for GPRA's focus on thematic goals and department- and governmentwide crosscutting comparisons, and was not used to evaluate similar programs together to facilitate trade-offs or make relative comparisons.

Lastly, we said that while the PART clearly must serve the President's interests, the many actors whose input is critical to decisions will not likely use performance information unless they feel it is credible and reflects a consensus on goals. Our work showed that it if OMB wanted to expand the understanding and use of the PART beyond the executive branch, it would be important for OMB to discuss in a timely fashion with Congress the focus of the PART assessments and clarify the results and limitations of the PART and the underlying performance information. On the other hand, we noted that a more systematic congressional approach to providing its perspective on performance issues and goals could facilitate OMB's understanding of congressional priorities and thus increase the PART's usefulness in budget deliberations.

In light of these issues, we recommended that OMB address the capacity demands of the PART, strengthen the PART guidance, address evaluation information availability and scope issues, focus program selection on crosscutting comparisons and critical operations, broaden the dialogue with congressional stakeholders, and articulate and implement a complementary relationship between the PART and GPRA. We also suggested that Congress consider the need for a structured approach to articulating its perspective and oversight agenda on performance goals and priorities for key programs.

OMB took several steps to implement many of our recommendations. For example, OMB clarified its PART guidance on defining the PART programs, using outcome and output measures, and expanded the discussion of evaluation quality; began to use the PART as a framework for crosscutting assessments; and expanded its discussion about the relationship between the PART and GPRA. The guidance notes that the PART strengthens and reinforces performance measurement under GPRA by encouraging the careful development of performance measures according to GPRA's outcome-oriented standards. It also requires that PART goals be "appropriately ambitious" and that GPRA and the PART performance measures be consistent. They have also begun reporting on the status of each program's recommendations and implemented PARTWeb, a Web-based data collection tool to, among other things, improve collaboration between OMB and agencies and centrally track the implementation and status of the PART recommendations.

The PART Encourages a Focus on Performance Measurement and Program Review

The PART process has aided OMB's oversight of agencies, and has focused agencies' efforts to improve performance measurement. According to OMB, the PART is a framework for program assessment and informs its budget decisions. Many agency officials told us that the PART helped either create or strengthen a culture of evaluation within the agencies by providing external motivation for program review. Not surprisingly, agency officials used the PART results to make a case for increased resources in general and for program evaluation specifically. This increased focus on performance is often reflected in improved ratings when "results not demonstrated" programs get reassessed by the PART—86 percent of programs previously rated "results not demonstrated" were subsequently rated adequate, moderately effective, or effective when reassessed. This focus is not without cost, however; the PART remains a labor-intensive process for both OMB and agencies.

OMB senior officials describe the PART as providing a consistent framework for assessing federal programs, and as a means to inform its budget decisions. As a major component of the PMA, OMB clearly relies on the PART—a significant component of the PMA—as a major oversight tool and finds information from the PART reviews useful. As we previously reported, the PART has helped to structure and discipline OMB's use of performance information for internal program analysis and budget review, and made their use of this information more transparent. Given the PART's use in the budget process, the high profile of the PMA scorecard, and the strong connection between the PART and successful performance on the PMA's budget and performance integration initiative, agencies have clear incentives to take the PART seriously.

Many agency officials told us that the PART helped either create or strengthen a culture of evaluation within the agencies by providing external motivation for program review. The PART question that asks whether a program has undergone regular, independent evaluations sends the message that program assessment and evaluation is an important management tool. For example, according to one agency official at the Health Resources Services Administration in HHS, this requirement encouraged staff to think more broadly about using different types of program evaluations and how they could get the most out of their evaluation dollars. Another HHS official reported that the PART provided an impetus for finishing strategic and evaluation plans for his program, which in turn helped inform the division's planning process. Our companion report on the PART evaluation recommendations reports similar findings, noting that most program officials interviewed for that report said that the PART recommendations directed senior management's attention to the need for evaluation.[12]

Not surprisingly, agency officials used the PART results in some cases successfully—to argue for increased resources in general as well as specifically for program evaluation. For example, officials in one agency said that a good rating on the PART is a powerful aid in gaining bipartisan support for budget increases. DOL agency officials told us that absent the PART, they might not have received funding to evaluate the Youth Activities

---

[12]GAO-06-67.

program—a program they felt had been in need of an evaluation for a long time.[13]

Agency officials we spoke with credited the PART with increasing attention to the use of performance measurement in day-to-day program management, which they considered to be of greater consequence than the PART's bottom-line ratings and recommendations. For example, agency officials at DOL credited the first year's PART assessments with encouraging managers to take steps prior to assessments to identify and address program weaknesses, develop and improve performance measures, and train staff on the PART. Officials from DOL's Trade Assistance program said that the PART forced them to look at the program in a new light, and be objective about what they are doing and how they are doing it. SBA officials said that the PART and the PMA helped them move away from "analysis by anecdote" and refocused their attention on the impact their programs have on small businesses, instead of largely on output measures such as the number of loans they have made. One official at HHS said that the PART allowed him to "evangelize" on the importance of good performance data and the perils of bad data. Other officials echoed a similar sentiment, one of them indicating that the PART scores helped to create "a new sense of urgency" about performance measures and completing the changes to performance systems that were already underway. The link between program ratings and the PMA scorecard also provided an incentive for change.

"Results not demonstrated" ratings have implications beyond the PART. For an agency to achieve "green" on the Performance and Budget Integration initiative of the PMA scorecard, less than 10 percent of its programs could have received a results not demonstrated PART rating for more than 2 years in a row. According to OMB's PMA scorecard update as of June 30, 2005, only nine agencies have met this particular criterion.[14]

---

[13]Although DOL received funding for this evaluation, it has been postponed until after the program is reauthorized because reauthorization is expected to result in an increased focus on out-of-school youths and a significant change in program activities.

[14]The nine agencies are the Departments of Energy, Labor, State, and Transportation; U.S. Agency for International Development; National Aeronautics and Space Administration; National Science Foundation; Small Business Administration; and Social Security Administration.

This increased focus on performance is often reflected in improved ratings when programs originally rated "results not demonstrated" are reassessed. When reassessed, 86 percent of programs previously rated "results not demonstrated" were rated adequate, moderately effective, or effective. Because programs were only reassessed when OMB determined that significant changes had been made to address deficiencies, this result is not surprising. However, it was, on average, the "results not demonstrated" programs with initially higher section IV scores (section IV measures program results) that, when reassessed, showed the greatest improvement in rating. While there were programs with low section IV scores that received an "adequate" rating when reassessed, lower scoring programs generally remained in the "results not demonstrated" category or received an "ineffective" rating when reassessed.

PART Remains a Labor-Intensive Process at OMB and Agencies

Although the PART has enhanced the focus on performance, this has not come without a cost. As we reported in our January 2004 report, senior OMB managers recognized the increased workload the PART initially placed on examiners; however, they expected the workload to decline as both OMB and agency staff became more familiar with the PART tool and process, and as issues with the timing of the PART reviews were resolved. During this review we found that while the learning curve did appear to flatten, it did not seem to compensate for either the increased workload due to the sheer number of programs being assessed or reassessed each year or the amount of time an individual assessment takes. This finding is consistent with views expressed by OMB staff during our 2004 review. They told us that they were surprised that reassessments took almost as long as assessing programs for the first time. OMB limited the scope of reassessments to include only those programs where there is significant evidence of change. Programs that received a "results not demonstrated" rating received priority for reassessment. According to OMB officials, this change was made partly due to resource constraints.

Officials in some of our case study agencies expressed concern that OMB's growing workload affects how the PART programs are defined. They said that as more programs are assessed OMB has less time to focus on the PART units that "make sense" and instead is creating larger PART units to help control the number of the PART assessments that need to be completed. One official recognized that getting into too much detail can be time consuming, but nonetheless noted that reviewing a larger "program" can lead to missing some important details; another said it can lead to recommendations that are not specific enough to be useful to a program.

One agency official said that the PART assessments can be thoughtful when OMB is knowledgeable about a program and has enough time to complete the reviews, but the assessments are less useful when OMB staff are unfamiliar with programs or have too many of the PART assessments to complete. Officials across all of our case study agencies reported these types of issues. For example, one official said that although the PART reviews were to be completed by the cognizant OMB examiner for the program, this was not always the case. He said that due to turnover at OMB, programs in his department were assessed by several people even within a single PART cycle, resulting in a lack of continuity. In several cases, agencies reported that OMB was not able to reassess programs because of resource constraints. Some officials told us that the heavy workload meant that the PARTs were not completed in a timely enough fashion to allow agencies to appeal ratings or present new performance measures, sometimes resulting in lower PART scores. OMB officials noted that OMB policy permits agencies to appeal answers to individual questions, not entire ratings, and that in practice, ratings may be appealed at any time during the PART process whether the ratings are in draft form or completed.

Although a senior OMB official acknowledged continuing capacity issues regarding the PART, he said that the PART is still a better way for examiners to accomplish their traditional program assessment responsibilities because it is more objective and transparent. He noted that OMB is devoting more people to help administer the PART tool and that PARTWeb, OMB's new on-line Web-based data collection system for PART, is also designed to ease the management of the process. For example, the official said that PARTWeb will automate the production of PART summary sheets.

The PART is a resource-intensive process for agencies as well. Some agency officials at DOL noted that the PART process is "one size fits all" in that a small program at DOL is supposed to have the same resources to devote to helping the budget examiner through the process and have the same analytic and evaluation resources as a large organization like the Social Security Administration. They said that agency staff is diverted from mission work to the PART work and in some cases is spending significant time on helping OMB staff understand the history and context of the programs.

To Date, the PART Appears to Have Had Limited Impact on Outcome-Based Program Results

OMB has said that a major purpose of the PART is program improvement. Our analysis supports OMB's statements that most of the PART recommendations to date were aimed at improving performance assessment, such as developing outcome and efficiency measures, and in collecting performance data. Improving managers' ability to assess program outcomes, identify information gaps, and assess next steps are necessary first steps on the path to long-term program improvement, but are not expected to result in observable program improvement in the short term. Moreover, as of February 2005—the date of the most recent available OMB data—the majority of the PART recommendations have not yet been fully implemented. Consequently, there is limited evidence to date of outcome-based program results. Implementing the PART recommendations has proven challenging. Although some agency officials appreciated the flexibility OMB provided by not making prescriptive recommendations, some follow-on actions were so general that it was difficult to understand what change was required or how progress could be measured. Some agencies did not discuss with OMB the evaluation plans created in response to the PART recommendations; combined with different expectations on the scope and purpose of the evaluations and the quality of evaluation designs, it is not certain whether these evaluations will meet OMB's needs. Lastly, OMB has used the PART as a framework for several crosscutting reviews, but these generally do not include all tools, such as tax expenditures, that contribute to related goals. Greater focus on selecting related programs and activities for concurrent review would improve their usefulness.

Most PART Recommendations Will Not Result in Observable Short-term Performance Improvements

For each program assessment, the PART summary worksheets were published in a separate volume with the President's fiscal year 2004 budget request. For the fiscal year 2005 and 2006 budgets, similar information was provided in an accompanying CD-ROM. The detailed, supporting worksheets for each program were posted on OMB's Web site.[15] The PART summary sheets display the program's key performance measures, budget information, significant findings, and follow-up actions (also known as recommendations. See fig. 1 for examples of follow-on actions). In the fiscal year 2006 budget, summary sheets for programs that have been previously assessed also include information on when the program was last assessed and the status of the follow-up actions. Status markers include

---

[15]http://www.whitehouse.gov/omb/part/index.html.

"no action taken," "action taken but not completed," and "completed." (See appendix III for examples of summary worksheets for programs assessed for the first time and for programs that were reassessed.)

Figure 1: Selected PART Follow-on Actions

The PART recommendations are generally aimed at improving program design, assessment, management, and/or funding. They can be general or very specific. Examples of recommendations for each of the four categories include:

Program Design:

- "Reduce unnecessary subsidies to lenders and other…program participants."
- "The 2006 Budget proposes to restructure the grant allocation process, providing the Secretary with greater discretion to award funds based on risks, threats, and vulnerabilities."

Program Management:

- "Work with the agency to coordinate efforts with other federal agencies to improve violent-crime reducing services."
- "Increase the number of accounts supporting this program to quicken the transfer of funds with contractors and increase management flexibility to address changing security conditions and mission priorities. This will significantly improve the obligation and costing process of funds."

Funding:

- "Maintaining funding at the 2005 enacted level until the agency can show how it will use additional funds to improve performance."
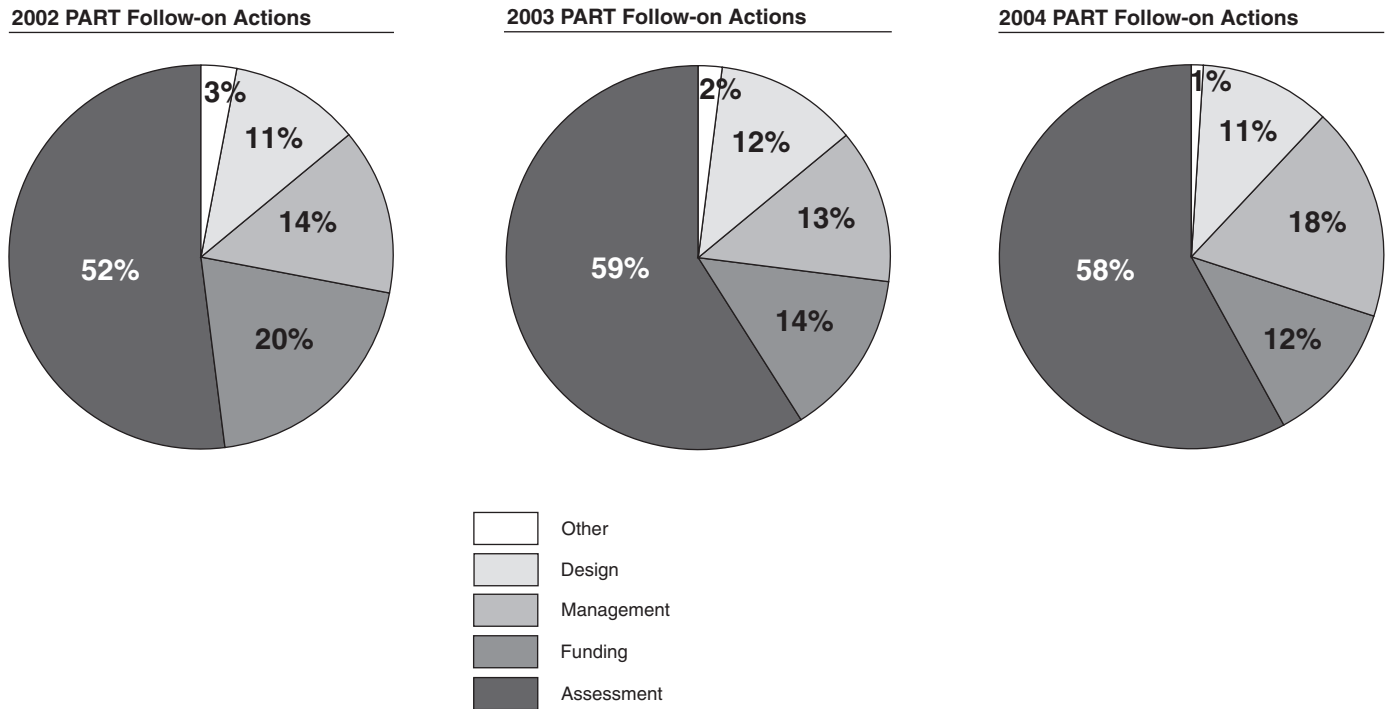- "Directly related to the PART findings, the Budget includes $37 million, a $3 million decrease."

Program Assessment:

- "Develop a means of regularly performing independent evaluations to examine program effectiveness."
- "Develop baselines for its proposed long-term measures. Without baselines for the measures, it was impossible to verify the performance of the program."

Source: OMB.

As figure 2 shows, the distribution of recommendations between program management, assessment, and design is fairly consistent over the 3 years; however, the percentage of recommendations with explicit funding references in a given year have steadily declined since the PART's inception from 20 percent in 2002 to 12 percent in 2004.

Figure 2: PART Recommendations by Type, 2002-2004

2002 PART Follow-on Actions

3%
11%
14%
52%
20%

2003 PART Follow-on Actions

2%
12%
13%
59%
14%

2004 PART Follow-on Actions

1%
11%
18%
58%
12%

Legend:
- Other
- Design
- Management
- Funding
- Assessment

Source: GAO analysis of OMB data.

A major goal of the PART is to identify program strengths and weaknesses and make recommendations to improve program results. However, we found that the link between problems identified by the PART assessments and the recommendations intended to address them is sometimes unclear. Regardless of what types of deficiencies were identified by the PART, the most frequent recommendations in each of the three years were related to performance assessments, such as developing outcome measures and/or goals, and improving data collection. While especially true for "results not demonstrated" programs, it also held true for programs rated "effective" and "moderately effective." Moreover, programs assessed for the first time in 2004—the most recent year for which data is available—received recommendations to improve performance assessments, such as outcome measures, as frequently as programs assessed during the first PART cycle.

More than half of all the PART recommendations made since the PART's inception are aimed at improving the "process" of program assessment. This includes developing meaningful and robust performance goals and

collecting quality data to measure progress against those goals. Of the 797 follow-on recommendations made in the first 2 years[16] for which OMB provided status information, 30 percent were considered fully implemented.[17] Of these, 47 percent are geared toward performance assessment. For example, the Animal and Plant Health Inspection Service Plant and Animal Health Monitoring Programs within the Department of Agriculture received three recommendations, one of which would create efficiency measures and the other which would update the program's measures and accomplishments. Such measures improve managers' ability to assess program outcomes, identify information gaps, and assess next steps, but are not expected to result in observable program improvement in the short term. OMB claims that many programs are getting better every year—which it defines as improving program outcomes, taking steps to address the PART findings, improving program management, and becoming more efficient—but, as noted above, these claims have not yet been fully born out.

Some recommendations are aimed at changing a program's purpose or design and/or implicitly or explicitly require action by Congress. For example, the Department of Agriculture's Commodity Credit Corporation's Marketing Loan Payments program received a recommendation to have the "House and Senate Agricultural committees examine the issue of payment limits for marketing loan and LDP gains and how they could be tightened." A Department of Education special education program was told to "work with Congress on the IDEA[18] reauthorization to increase the act's focus on accountability and results, and reduce unnecessary regulatory and administrative burdens." Even in cases where there is general agreement that legislative action or statutory changes are needed, making them takes time.

OMB has said that if statutory provisions impede effectiveness, one result of a PART review could be recommendations for legislative changes. The responsibility to implement the PART recommendations lies with agency

---

[16]Follow-on actions made for programs assessed in 2004 would not be expected to have status updates.

[17]Sixty-four percent have been partially implemented and the remaining 6 percent have not yet been acted on.

[18]IDEA (Individuals with Disabilities Education Act) was reauthorized by the Individuals with Disabilities Education Act of 2004, P. L. No. 108-446 (Dec. 3, 2004).

and program managers. Successfully implementing recommendations that require legislative action or statutory changes requires the additional step of positively engaging Congress. A perceived disconnect between what one is held accountable for and what one has the authority to accomplish is not unusual. Our 2003 survey of governmentwide federal managers supports this view.[19] We found that while 57 percent of non-Senior Executive Service (SES) managers and 61 percent of SES mangers believed they were held accountable for results to a great or very great extent, only 38 percent and 40 percent, respectively, believed that managers at their level had the decision-making authority they needed to achieve agency goals.

The Lack of Prioritization of PART Recommendations Has Implications for Resource Allocation

Although OMB has given agencies discretion to define strategies to implement recommendations, OMB officials told us that, as a matter of policy, they have generally not prioritized the recommendations within each agency or across the more than 1,700 recommendations governmentwide because they do not want to dilute attention paid to any of the recommendations by deeming them a lower priority. Realistically, though, agencies cannot act on all of them concurrently. Because OMB has chosen to assess nearly all federal programs, resources are diffused across multiple areas instead of concentrated on those areas of highest priority both within agencies and across the federal government. This strategy is likely to lengthen the time it will take to observe measurable change.

Moreover, as we report in our companion report on the PART evaluation recommendations, agency officials questioned the PART's assumption that all programs should have evaluations. Agency officials in one of our case study agencies said that they were able to fund some evaluations for small programs without cutting into program budgets, but other agency officials pointed out that spending several hundred thousand dollars for an evaluation study was a reasonable investment for large programs; they questioned whether all programs, regardless of size or importance, need to be evaluated, especially in times of tight resources and suggested instead a risk-based approach to prioritizing programs to be evaluated.[20] We also noted that the lack of prioritization meant that agencies were free to choose which programs to evaluate first, and were likely to be influenced

---

[19]GAO-04-38.

[20]GAO-06-67.

by such factors as the potential effect of the PART reassessments on their PMA scores.

PART Recommendations Provide Wide Latitude for Agency Actions

OMB gives agencies wide latitude to implement the PART recommendations, which had both positive and negative effects on agency actions. Some officials appreciated the flexibility that OMB provided by not making prescriptive recommendations. They said that they were generally able to devise implementation strategies that suit programmatic needs and in most cases were able to devise implementation strategies that fit within existing agency plans and procedures. While they discuss their strategies with OMB, it is generally up to agency staff to determine the best course of action to implement the recommendations. In other cases, though, agency officials said that the recommendations were so broad as to be vague. This sometimes hampered implementation.

For example, a DOE program received a recommendation to "improve performance reporting by grantees and contractors by September, 2004." DOE officials told us that in this case, the desired result is unclear. The PART requires that they report grantee performance both aggregated on a programwide level and disaggregated at the grantee level. DOE officials said that because they already report grantee information in each of these ways for both the PART and their Performance and Accountability Report (PAR), and because the recommendation does not describe the deficiencies in the reporting, they are unsure how to change their reporting practices to meet OMB's needs. Our review of this program's PART worksheet supports this view. Although we found one reference to "inadequate communication in the PAR of program-level aggregate data on the impact of the grants program" in the detailed supporting worksheet for this program, the reason for the inadequacy is unclear. In cases such as these, it is difficult to see how OMB and agencies can monitor progress. Given the importance OMB places on implementing the PART recommendations, it is important that recommendations clearly identify deficiencies and provide a basis for determining whether they are complete.

OMB's and Agencies' Expectations Differed on Evaluations Conducted in Response to PART Recommendations

Federal agencies are increasingly expected to demonstrate effectiveness in achieving agency or governmentwide goals. Our previous work has shown that the accuracy and quality of evaluation information necessary to make the judgments called for when rating programs is highly uneven across the federal government. To help explain linkages between program activities, outputs, and outcomes, program evaluation designs are tailored to address various types of programs and questions. For example, a process evaluation reviews various aspects of program operations to assess the extent to which a program is operating as intended. Alternatively, an impact evaluation depends on scientific research methods to assess the net effect of a program by comparing program outcomes with an estimate of what would have happened in the absence of the program, in order to isolate the program's contributions to the observed outcomes. In other words, evaluations are useful to specific decisionmakers to the degree that the evaluations are credible and address their information needs.

Our companion report notes that although the evaluation recommendations provided agencies with flexibility to interpret what evaluation information OMB expected and which evaluations to fund, a few programs did not discuss their evaluation plans with OMB; combined with different expectations on defining the scope and purpose of evaluations and disagreements about the quality of evaluation designs, it is not certain whether these evaluations will meet OMB's needs.[21]

Disagreements about the Scope and Purpose of Evaluations

OMB and our case study agencies significantly differed in defining evaluation scope and purpose. Some of the difficulties seemed to derive from the OMB examiners' expecting to find, in the agencies' external evaluation studies, comprehensive judgments about program design, management, and effectiveness, similar to the judgments made in the PART examinations.

Because evaluations designed for internal and external audiences often have a different focus, differences of opinion on the usefulness of evaluations are perhaps not surprising. Evaluations that agencies initiate typically aim to identify how to improve the allocation of program resources or the effectiveness of program activities. Studies requested by program authorizing or oversight bodies, such as OMB, are more likely to

---

[21]GAO-06-67.

address external accountability—to judge whether the program is properly designed or is solving an important problem.

HHS departmental officials reported having numerous differences with OMB examiners over the acceptability of their evaluations. HHS officials were particularly concerned that OMB sometimes disregarded their studies and focused exclusively on OMB's own assessments. One program official complained that OMB did not adequately explain why the program's survey of refugees' economic adjustment did not qualify as an "independent, quality evaluation," although an experienced, independent contractor conducted the interviews and analysis. In the published PART review, OMB acknowledged that the program surveyed refugees to measure outcomes and monitored grantees on site to identify strategies for improving performance. In our subservent interview, OMB staff explained that the outcome data did not show the mechanisms by which the program achieved its outcomes and grantee monitoring did not substitute for obtaining an external evaluation, or judgment, of the program's effectiveness. Other HHS officials said that OMB had been consistent in applying the standards for independent evaluation, but these standards were set extremely high.

In reviewing a vaccination program, OMB did not accept the several research and evaluation studies offered, since they did not meet all key dimensions of "scope." OMB acknowledged that the program had conducted several management evaluations of the program to see whether the program could be improved but found their coverage narrow and concluded "there have previously been no comprehensive evaluations looking at how well the program is structured/managed to achieve its overall goals." The examiner also did not accept an external Institute of Medicine evaluation of how the government could improve its ability to increase immunization rates because the evaluation report had not looked at the effectiveness of the individual federal vaccine programs or how the program complemented the other related programs. However, in reviewing recommendation status, OMB credited the program with having contracted for a comprehensive evaluation that was focused on the operations, management, and structure of this specific vaccine program.

Disagreements about the Quality of Evaluation Designs

OMB and agencies differed in identifying which evaluation methods were sufficiently rigorous to provide high-quality information on program effectiveness. OMB guidance encouraged the use of randomized controlled trials, or experiments, to obtain the most rigorous evidence of program impact but also acknowledged that these studies are not suitable or

feasible for every program. However, without guidance on which—and when—alternative methods were appropriate, examiners and agency staff disagreed on whether specific evaluations were of acceptable quality. To help develop shared understandings and expectations, federal evaluation officials and OMB examiners held several discussions on how to assess evaluation quality according to the type of program being evaluated.

When external factors such as economic or environmental conditions are known to influence a program's outcomes, an impact evaluation attempts to measure the program's net effect by comparing outcomes with an estimate of what would have occurred in the absence of the program intervention. A number of methodologies are available to estimate program impact, including experimental and quasi-experimental designs. Experimental designs compare the outcomes for groups that were randomly assigned to either the program or to a nonparticipating control group prior to the intervention. The difference in these groups' outcomes is believed to represent the program's impact, assuming that random assignment has controlled for any other systematic difference between the groups that could account for any observed difference in outcomes. Quasi-experimental designs compare outcomes for program participants with those of a comparison group not formed through random assignment, or with participants' experience prior to the program. Systematic selection of matching cases or statistical analysis is used to eliminate any key differences in characteristics or experiences between the groups that might plausibly account for a difference in outcomes.

Randomized experiments are best suited to studying programs that are clearly defined interventions that can be standardized and controlled, and limited in availability, and where random assignment of participants and nonparticipants is deemed feasible and ethical. Quasi-experimental designs are also best suited to clearly defined, standardized interventions with limited availability, and where one can measure, and thus control for, key plausible alternative explanations for observed outcomes. In mature, full-coverage programs where comparison groups cannot be obtained, program effects may be estimated through systematic observation of targeted

measures under specially selected conditions designed to eliminate plausible alternative explanations for observed outcomes.[22]

Following our January 2004 report recommendation that OMB better define an "independent, quality evaluation," OMB revised and expanded its guidance on evaluation quality. The guidance encouraged the use of randomized controlled trials as particularly well suited to measuring program impacts but acknowledged that such studies are not suitable or feasible for every program, so it recommended that a variety of methods be considered. OMB also formed an Interagency Program Evaluation Working Group in the summer of 2004 which discussed this guidance extensively to provide assistance on evaluation methods and resources to agencies undergoing a PART review. Evaluation officials from several federal agencies expressed concern that the OMB guidance materials defined the range of rigorous evaluation designs too narrowly. In the spring of 2005, representatives from several federal agencies participated in presentations about program evaluation purposes and methods with OMB examiners. They outlined the types of evaluation approaches they considered best suited for various program types and questions. (See fig. 3.)[23] However, OMB did not substantively revise its guidance on evaluation quality for the fiscal year 2007 reviews beyond recommending that "agencies and OMB should consult evaluation experts, in-house and/or external, as appropriate, when choosing or vetting rigorous evaluations."[24]

---

[22]For further discussion see Peter H. Rossi, Howard E. Freeman, and Mark W. Lipsey, *Evaluation: A Systematic Approach*, 6th ed. (Thousand Oaks, California: Sage Publications, 1999). For additional examples of alternative evaluation designs, see GAO, *Program Evaluation: Strategies for Assessing How Information Dissemination Contributes to Agency Goals*, GAO-02-923 (Washington, D.C.: Sept. 30, 2002).

[23]The entire evaluation dialogue presentation is at http://www.epa.gov/evaluate/part.htm (Aug. 8, 2005).

[24]OMB, *Guidance for Completing the Program Assessment Rating Tool (PART)* (Washington, D.C.: March 2005) is at http://www.whitehouse.gov/omb/part.

GAO-06-28 Performance Budgeting

Figure 3: Federal Evaluators' Views on Tailoring Designs for Program Effectiveness Evaluations

Typical designs used to assess program effectiveness	Design features that help control for alternative explanations	Best suited for (typical examples)
Process and outcome monitoring or evaluation	Compares performance to pre-existing goal or standard. For example: • OMB R&D criteria of relevance, quality and performance. • Productivity, cost effectiveness and efficiency standards	Research, enforcement, information and statistical programs, and business-like enterprises with • few, if any, alternative explanations for observed outcomes. • ongoing programs producing goods and services • complete national coverage
Quasi-experiments – Single Group	Compare outcomes for program participants or entities before and after the intervention. • Multiple data points over time are necessary. • Control for alternative explanations by statistical adjustments and analyses such as modeling.	Regulatory and other programs with • clearly defined interventions with distinct starting times • complete national coverage • random assignment of participants or entities to groups is NOT feasible, practical, or ethical.
Quasi-experiments – Comparison Groups	Compares outcomes for program participants or entities with outcomes for a comparison group selected to closely match the "treatment" group on key characteristics. • Key characteristics are plausible alternative explanations for the outcome. • Measure outcomes before and after intervention (pretest, posttest).	Service and other programs with • clearly defined interventions that can be standardized and controlled • limited national coverage • random assignment of participants or entities to groups is NOT feasible, practical, or ethical.
Randomized experiments	Compares outcomes for randomly assigned program (treatment) participants or entities with outcomes for a randomly assigned "control" group prior to intervention. • Measure outcomes before and after intervention (pretest, posttest).	Service and other programs with • clearly defined interventions that can be standardized and controlled • limited national coverage • random assignment of participants or entities to groups is feasible and ethical.

Source: Eric Bernholz and others, *Evaluation Dialogue between OMB Staff and Federal Evaluation Leaders: Digging a Bit Deeper into Evaluation Science*, (Washington, D.C.: Apr. 2005).

The PART Framework Has Been Applied to Crosscutting Programs, but More Needs to Be Done

Although assessing programs in isolation can yield useful information, it is often critical to understand how individual programs fit within a broader portfolio of tools and strategies—such as regulations, credit programs, and tax expenditures—to accomplish federal goals. Such an analysis is necessary to capture whether a program complements and supports other related programs, whether it is duplicative and redundant, or whether it actually works at cross-purposes to other initiatives.

Although variations of the PART tool are meant to capture the different approaches to service delivery, such as grants versus direct federal

activities, not all approaches—such as tax expenditures—are systematically assessed by the PART. Tax expenditures may be aimed at the same goals as spending programs but little is known about their effectiveness or their relative efficacy when compared to related spending programs in achieving the objectives intended by Congress. Since in some years, tax expenditures are about the same order of magnitude as discretionary programs and in some program areas tax expenditures may be the main tool used to deliver services, this is a significant gap. We recently recommended that OMB require that tax expenditures be included in the PART and any future such budget and performance review processes so that tax expenditures are considered along with related outlay programs in determining the adequacy of federal efforts to achieve national objectives.[25]

Consistent with recommendations in our January 2004 report, OMB has begun to use the PART framework to conduct crosscutting assessments. OMB reported on two crosscutting PART assessments—Rural Water programs and Community and Economic Development (CED) programs—for the fiscal year 2006 budget and it plans to conduct three additional crosscutting reviews on block grants, credit programs, and small business innovation research for the fiscal year 2007 budget.

According to both OMB and agency participants in the cooperative CED assessment, several aspects worked well. For example, the CED effort leveraged federal governmentwide community and development expertise housed in the OMB Interagency Collaborative on Community and Economic Development (ICCED).[26] The group focused on four elements: (1) determining the programs to be included in such a comparison; (2) reaching agreement on the goals and objectives of similar programs; (3) identifying opportunities to better coordinate, target, leverage, and increase efficiency and effectiveness of similar programs; and

---

[25]GAO, *Tax Policy: Tax Expenditures Represent a Substantial Commitment of Federal Support and Need to Be Reexamined*, GAO-05-690 (Washington, D.C.: Sept. 19, 2005).

[26]ICCED participants included the Community Development Financial Institutional Fund, the Economic Development Administration within the Department of Commerce, the Department of Housing and Urban Development, and the Department of Agriculture. The Appalachian Regional Commission and the Neighborhood Reinvestment Corporation also participated. The ICCED mission is to "coordinate the federal government's economic development, housing and community improvement policies, activities and initiatives in distressed communities and targeted populations for maximum economic stimulus, efficiency, and impact."

(4) establishing a common framework of performance measures and accountability.[27] Cognizant agency officials were pleased with this collaborative interagency process. They found value in leveraging existing efforts within agencies and benefited from OMB staff consultation.

The CED crosscutting assessment examined the performance of 18 of the 35 federal community and economic development programs identified by OMB and account for the majority of the $16.2 billion OMB estimates is spent annually in this area.[28] Although OMB identified three tax expenditures in the CED portfolio, it did not assess all of them with the PART instrument even though the Department of the Treasury's (Treasury) estimate of their combined "cost" is nearly $1.4 billion, or about 57 percent of Treasury's revenue loss estimates for community development.[29]

Little information on the CED crosscutting assessment was initially available beyond the brief description contained in the *Analytical Perspectives* volume of the Fiscal Year 2006 President's budget request. Some OMB documents and administration officials stated that all 18 programs had been assessed by the PART. However, 8 of the 18 programs proposed for consolidation were actually assessed by the PART. Because PART programs do not always clearly align with the individual programs proposed for consolidation, it can be difficult to determine which programs were assessed with the PART and which were not. As the CED team itself recognized, the results of a crosscutting assessment need to be communicated to stakeholders and the public. Unless the scope, purpose,

---

[27]Such coordination is consistent with practices we identified in our recent report on agency coordination. See GAO, *Results-Oriented Government: Practices That Can Help Enhance and Sustain Collaboration among Federal Agencies*, GAO-06-15 (Washington, D.C.: Oct. 21, 2005).

[28]According to OMB, these 35 programs include grant and loan programs and tax incentives spread across 11 federal agencies.

[29]While sufficiently reliable as a gauge of general magnitude, the sum of the individual revenue loss estimates has important limitations in that any interactions between tax expenditures will not be reflected in the sum. In addition, tax expenditure revenue loss estimates for specific provisions do not take into account potential behavioral responses to changes in these provisions on the part of taxpayers, and, in turn, no potential behavioral response would be reflected in the sum of the estimates. Thus, the revenue loss from all or several tax expenditures together might be greater or less than the sum of the estimated revenue losses from the individual tax expenditures, and no measure of the size or the magnitude of these potential interactions or behavioral responses to all or several tax expenditures is available.

and results are clear to stakeholders, the fruits of crosscutting assessments will likely not be realized.

In choosing programs to assess and reassess with the PART, OMB considers a variety of factors, including continuing presidential initiatives, whether a program is up for reauthorization, and whether a program received a rating of "results not demonstrated" in a previous PART review. Although these are reasonable criteria, a greater emphasis on selecting programs related to common or similar outcomes for review in a given year would enable decision makers to better analyze the efficacy of programs related to such outcomes, and improve the usefulness of crosscutting reviews conducted under the PART framework. Moreover, using PART assessments to review the relative contributions of similar programs to common or crosscutting goals and outcomes established through the GPRA process could improve the connection between the PART and GPRA. Developing a more strategic approach to selecting and prioritizing programs to be reassessed under the PART can also help ensure that OMB and agencies are using limited staff resources to the best advantage.

The PART-GPRA Relationship Does Not Adequately Consider the Different Needs of the Budget and Planning Processes and Their Various Stakeholders

Although both the PART and GPRA aim to improve the focus on program results, the different purposes and time frames they serve continue to create tensions. Some agencies have made progress over the past several years in reconciling the PART and GPRA, however, unresolved tensions can result in conflicting ideas about what to measure, how to measure it, and how to report program results. We continue to find evidence that the closed nature of the executive budget formulation process may not allow for the type of consultative stakeholder involvement in strategic and annual planning envisioned by GPRA. We remain concerned that the focus of agency strategic planning is shifting from long-term goal setting to short-term budget and planning needs.

Agencies Continue to Struggle with Integrating the PART and GPRA

OMB attempted to clarify the relationship between the PART and GPRA in its PART guidance in 2004. The guidance notes that the PART strengthens and reinforces performance measurement under GPRA by encouraging the careful development of performance measures according to GPRA's outcome-oriented standards. It also requires that the PART goals be "appropriately ambitious" and that GPRA and PART performance measures be consistent.

Some agencies have made progress over the past several years in reconciling the PART and GPRA. For example, DOE and SBA officials told us that their existing GPRA measures now relate to or are generally accepted for PART purposes. Officials from DOE's Office of Science and Labor's Employment Standards Administration told us that OMB actively encouraged them to use their GPRA measures in the PART. HHS's Breast and Cervical Cancer, Diabetes, and Foster Care programs as well as the Administration on Developmental Disabilities were able to use some existing GPRA measures as annual PART measures. These experiences are consistent with OMB's view that although the PART and GPRA are often focused on different sets of measures, the characteristics of both sets should be the same (e.g., outcome-oriented, ambitious) and support OMB's belief that they have adequately clarified the relationship between the PART and GPRA.

However, some agency officials we spoke with described persistent difficulties in integrating the two processes. Some described the PART and GPRA as duplicative processes that strain agency resources; others said they conflicted. As described below, defining a "unit of analysis" and performance measures useful for both budget and performance purposes remains a challenge. One official noted, "there is almost an unavoidable conflict between data that is useful from a governmentwide perspective and data that is useful to program managers." Although the Breast and Cervical Cancer and Diabetes programs had some success marrying their annual GPRA measures with short-term PART measures, they found that OMB did not consider their long-term GPRA goals to be aggressive enough; the measures were revised to meet OMB's needs.

Defining a "Unit of Analysis" Useful for Both Budget and Planning Purposes Remains a Challenge

OMB acknowledges that to improve performance and management decisions, OMB and agencies should determine an appropriate "unit of analysis" for a PART assessment. The PART guidance notes that although the budget structure is not perfect for program review in every instance, the budget structure is the most readily available and comprehensive system for conveying PART results. In response to our January 2004 report, OMB expanded its guidance on how the unit of analysis is to be determined. The guidance notes that interdependent programs or program activities can be combined for purposes of the PART as long as the aggregated unit of analysis for the PART is able to illuminate meaningful management distinctions among programs that share common goals but are managed differently. Moreover, it notes that several factors should be considered when deciding whether to combine programs, such as a

program's purpose, design, and administration; budgeting; and whether the programs support similar outcome goals.

Although less of a problem than it was during our January 2004 review of the PART process, difficulties in defining a unit of analysis useful for both OMB's budget analysis and program management purposes remain, and continue to highlight the tension between the PART and GPRA. Some agency officials acknowledged that overly disaggregating programs for the PART sometimes does not provide an understanding of how the entire program or service delivery system works before attempting to assess the performance of component pieces. One official described it as "putting the cart before the horse." Some agency officials noted that difficulties can also arise when unrelated programs and programs with uneven success levels are combined for the PART. For example, OMB combined programs authorized under titles VII and VIII of the Public Health Services Act to create the Health Professions PART program. As required by the PART guidance, the entire PART program received a "no" for each question where any of the PART program components did not meet the requirements for a "yes" answer. As agency officials recognized, assessing the programs separately would have made each program's successes and weaknesses more transparent. They felt this was important, as the individual programs have different underlying program authorities, goals, and attempt to address the maldistribution of health professionals in a variety of ways. In other words, although they complement each other, they serve different needs. OMB senior officials acknowledged that combining programs could theoretically make each component's successes and challenges less apparent, but that in this case it is hard to argue that programs authorized by different titles of the Public Health Services Act are unrelated to each other.

The goals and recommendations developed in a PART review, and hence the overall quality of the review, may suffer when the unit of analysis is not properly targeted. For example, the Centers for Disease Control and Prevention's (CDC) National Immunization Program (NIP) includes both the 317 program—which provides funding to support 64 state, local, and territorial public health immunization programs for program operations and vaccine purchase—and the Vaccines for Children (VFC) program— which provides publicly purchased vaccines to participating providers which are then given to eligible children without cost to the provider or the parent. Only the 317 program has been assessed by the PART to date. In its PART assessment of the 317 Program, OMB noted that the administration was including a legislative proposal in the fiscal years 2004 and 2005 budget

requests to "make it easier for uninsured children who are eligible for the CDC Vaccines for Children (VFC) program to receive immunizations in public health clinics, to improve program efficiency in the overall childhood immunization program. This proposal will expand the VFC program and result in $110 million in savings to the 317 discretionary childhood immunization program." According to HHS officials, these proposals are outside the scope of the 317 program. They said that the 317 program's stakeholders believe that OMB penalized the 317 program by recommending a change in that program that only the VFC program could accomplish. Program officials were unable to convince OMB to remove the VFC legislative proposal from the 317 program PART assessment summary sheet. Similarly, when OMB proposed a goal related to the global eradication of polio, HHS officials were unable to convince OMB that while global eradication of polio is a goal of the NIP overall, it is not within the scope of the individual 317 program, which is solely a domestic program. Although one of the program's annual measures is still the "number of polio cases worldwide," OMB responded to the agency's concern in the most recent PART summary sheet for the 317 program, noting that "the global polio measure will be tracked by the global immunization program, which will be assessed separately in the future, and not by the 317 program."

Concerns about PART Measures Reflect Long-standing Tensions over Performance Measurement

We have long recognized the difficulties of developing useful results-oriented performance measures for programs such as those geared toward long-term health outcomes and research and development (R&D) programs. However, in a June 1997 report discussing the challenges of GPRA implementation, we also said that although such performance measurement efforts were difficult, they have the potential to provide important information to decision makers. We noted that agencies were exploring a number of strategies to address these issues, such as using program evaluations to isolate program impact, developing intermediate measures, employing a range of measures to assess progress, and working with stakeholders to seek agreement on appropriate measures.[30] OMB recognizes several of these approaches as appropriate alternatives to outcome measures for PART purposes but as described below, agencies have had mixed success reaching agreement with OMB in these areas. Although these types of measurement challenges are clearly not new or unique to the PART, they are aggravated by the difficulties of developing

---

[30]GAO, *The Government Performance and Results Act: 1997 Governmentwide Implementation Will Be Uneven*, GAO/GGD-97-109 (Washington, D.C.: June 2, 1997).

measures useful for multiple purposes and audiences and often remain a point of friction in agencies and sometimes within OMB.

For programs that can take years to observe program results, it can be difficult to identify performance measures that will provide information on the annual progress they are making toward achieving program results. This can complicate efforts to arrive at goals useful to multiple parties for multiple purposes. For example, CDC officials told us that long-term health outcome measures favored by the PART are often not as useful to them as data on preventative measures, which tell managers where more efforts are needed and allows them to respond more quickly.

Programs where the federal government is one among many actors present similar challenges—when an outcome is beyond the scope of any one program, any changes made to a single federal program will not necessarily have an immediate effect. For example, for the Diabetes program OMB expressed interest in a long-term health outcome measure that tracks changes in national blindness and amputation rates. Program officials said that these types of changes generally cannot be attributed solely to the Diabetes program because it serves a relatively small portion of the population and works with many partners. The Breast and Cervical Cancer program—which screens low-income women and provides public education, quality assurance, surveillance, partnerships, and evaluation regarding breast cancer screening among low-income women—has similar concerns about OMB's interest in linking the program to changes in the overall mortality rates of cancer patients.

Agency experiences with the PART's emphasis on efficiency measures presented a more varied picture. Some programs had success by defining efficiency in terms of program administration rather than program operations. For example, HHS's foster care officials said that because children's safety could have been compromised by moving children too quickly out of foster care, OMB agreed that an administrative efficiency measure would be appropriate instead of the type of outcome-oriented efficiency measure cited above. DOE officials told us that the Strategic Petroleum Reserves program is well suited to the PART's view of "efficiency" because it can show (1) how every dollar from its budget is spent, (2) that it is spent efficiently, and (3) that the program results related to spending those dollars.

In other cases, differences of opinion about efficiency measures highlighted the challenges that can result from the difficult but potentially

useful process of comparing the costs of programs related to similar goals. For example, DOL agency officials told us that since Job Corps is a self-paced program, participants are permitted to remain in the program for up to 2 years (or up to 3 years with special approval). They consider this to be adequate time for students to complete their education and/or vocational training, which, as studies indicate, generally results in higher wages. DOL agency officials noted that since costs per participant increase the longer a student remains in the program, Job Corps appears less "efficient" compared with other job training programs, which reflects poorly in Job Corps' PART assessment. They suggested cost per student day as a cost measure with less inherent perverse incentive, but OMB did not accept the suggestion.

Similarly, DOL agency officials explained that whereas Job Corps' current GPRA measures track the percentage of job/education placements for *program exiters who graduate*, the common measures—which OMB uses to gauge performance across all job-training programs—track entered employment/education for *all* program exiters, regardless of their graduate status at exit.[31] Although there are significant differences in the time frames, the placement criteria, and the pool of participants for these measures, these officials told us that the measures are treated as interchangeable in the PART review. In other words, the same benchmark set for the "graduate placement" GPRA indicator is also used for the "placement of all participants" common measure indicator. Consequently, agency officials said, Job Corps is in the position of either (1) failing to meet the common measure goal or (2) being labeled un-ambitious by OMB if the goal is changed to be attainable yet—in DOL's view—still aggressive. Either way, the agency officials said that their PART assessment is negatively affected.

Several R&D officials noted that prior to the PART, there had been a collaborative effort to develop OMB's R&D investment criteria to better assess the value of R&D programs. However, these managers believed that the investment criteria—which R&D program managers find useful to manage their programs—have been overshadowed by the PART—which OMB finds useful in its budget development process. Part of the trouble seems to be that the PART explicitly requires all programs to have or be

---

[31]DOL agency officials described graduates as those who complete a vocation and/or attain a Graduate Equivalent Degree or High School diploma, stay in the program for 60 or more days, and do not violate the zero-tolerance policy.

developing an efficiency measure, while the investment criteria do not. The investment criteria focus on improving the management of basic research programs. One agency official noted that this is a management efficiency question, not a cost question; therefore it should be captured in the PART's management section instead of the results section. Such a change could affect a program's PART score because the management section represents 20 percent of the total weighted score whereas the results section represents 50 percent of the total weighted score.

In the investment criteria published with the 2004 PART guidance, OMB noted that it had worked to clarify the implementation of the investment criteria, stating that the investment criteria are broad criteria for all R&D programs while the PART is used to determine compliance with the investment criteria at the program level. OMB also recognized that while programs must demonstrate an ability to manage in a manner that produces identifiable results, taking risks and working toward difficult-to-attain goals are important aspects of good research management, especially for basic research. They further note that the intent of the investment criteria is not to drive basic research programs to pursue less risky research that has a greater chance of success, and that the administration will focus on improving the management of basic research programs.

Disagreements over when and how to revise and communicate information about federal programs further highlight tensions between OMB and agencies. OMB Circular A-11 states that the performance targets included in the PARTs and congressional justifications need to be updated to reflect the budgetary resources and associated performance targets decided for the President's budget, and that budget and performance reports should identify changes to performance goals that primarily stemmed from assessing actual performance. However, several agency officials reported problems with adjusting or retiring goals. For example, agency officials told us that sometimes goals need to be retired or consolidated, and cited instances in which they were not permitted to do so even after intense negotiation with OMB. They said that changing goals disrupts the ability to observe historical trends, making it hard for OMB to measure against a baseline. We recognize the value of baseline information and that changing goals and measures can limit the ability to observe trends over time. However, this is not always possible. Revised performance information can also further enhance performance assessments.

Nature of the Budget Formulation Process Limits Stakeholder Involvement

As we have previously reported, successful organizations base their strategic planning to a large extent on the interests and expectations of their stakeholders, since they recognize that stakeholders will have a lot to say in determining whether their programs succeed or fail. Congress, the executive branch, and other stakeholders may all strongly disagree about a given agency's missions and goals—in fact, full agreement among stakeholders on all aspects of an agency's efforts is relatively uncommon because stakeholders' interests can differ significantly. Still, stakeholder involvement is important to help agencies ensure that their efforts and resources are targeted at the highest priorities. Just as important, involving stakeholders—especially Congress—in strategic planning efforts can help create a basic understanding among stakeholders of the competing demands that confront most agencies. Because of Congress's constitutional power to create and fund programs, congressional involvement is indispensable to defining each agency's mission and establishing its goals.

Some tension between the level of stakeholder involvement in the development of performance measures in the GPRA strategic planning process and the process of developing performance measures for the PART excutive is inevitable. Compared to the relatively open-ended GPRA process, any executive budget formulation process is likely to seem closed. An agency's communication with stakeholders, including Congress, about goals and measures created or modified during the formulation of the President's budget, is likely to be less than during the development of the agency's own strategic or performance plan.

Although OMB's PART guidance discusses the need to integrate the PART and GPRA, we continue to find evidence that the closed nature of the executive budget formulation process may not allow for the type of stakeholder involvement in strategic and annual planning envisioned by GPRA. Beginning with the fiscal year 2005 budget submission, OMB required agencies to submit a performance budget, which is expected to satisfy all statutory requirements of the GPRA annual performance plan. It is generally expected to include the PART performance goals (including annual and long-term performance measures with targets and time frames) for programs that have been assessed by the PART.[32] The PART guidance recognizes stakeholder involvement in strategic planning as required by

---

[32]For programs not yet assessed by PART, OMB expects the measures and targets to meet the standards set in the PART guidance.

GPRA by saying agencies must consult with Congress and solicit and consider the views of interested and potentially affected parties.

At the same time, the executive budget formulation process—to which the PART belongs—is "predecisional." This means that information from the annual budget process, including information required in agencies' annual GPRA plans, is embargoed within the executive branch until the President's budget request is transmitted to Congress. Agencies may therefore be prevented from consulting with their stakeholders when developing annual and long-term goals and measures. Some of our case study agencies described similar experiences. Their interaction with key stakeholders was limited. Sometimes they had to present new or revised program goals and measures to their stakeholders after the fact, and in some cases stakeholders disagreed with the goals, or had no choice but to accept them after the fact.

Discussions of how performance information is being used are important because GPRA performance reports are intended to be one of Congress's major accountability documents. As such, these reports are to help Congress assess agencies' progress in meeting goals and determine whether planned actions will be sufficient to achieve unmet goals, or, alternatively, whether the goals should be modified. Because predecisional performance information must be excluded from the reports, their potential as a source of information to Congress is limited. However, this embargo conflicts with OMB's own reporting requirements regarding the issuance of agency Performance and Accountability Reports (PAR). OMB's Circular A-11 guidance notes that the transmittal date for an agency's PAR is November 15, and that because this precedes the transmittal of the President's budget, an agency may need to omit certain "privileged" information from its PAR.[33] As described in Circular A-11, this privileged information includes specifically required elements of agency PARs, including an evaluation of performance for the current fiscal year; schedules for achieving established performance goals; and, if a performance goal is impractical or infeasible, an explanation of why that is the case and what action is recommended. However, OMB senior officials

---

[33]OMB's November 15 deadline for the PAR submission is earlier than the statutory deadline. OMB's recent memorandum to agency heads on this subject made this accelerated deadline permanent, and notes that the purpose of accelerated reporting is to better ensure that timely and accurate financial and performance information is made available to federal agency managers as soon as possible after the end of the fiscal year and throughout the year.

told us that the only information that cannot be included in a PAR is that related to target levels of funding and/or policy changes.

Tailoring Outreach to Meet Congressional Needs Is Key to Increasing the Likelihood of Congress's Considering the PART in Its Deliberations

While the PART has been useful to OMB to achieve its own budget formulation goals, OMB acknowledges the need to work to gain congressional acceptance of the tool and its results. In response to our January 2004 report on the first year of implementing the PART, OMB said that it was working to "generate, early in the PART process, an ongoing, meaningful dialogue with congressional appropriations, authorization, and oversight committees about what they consider to be the most important performance issues and program areas warranting review." Although OMB uses a variety of methods to communicate the PART assessment results, congressional committee staff said these methods have not facilitated this early consultation on the PART. An absence of early consultation has contributed to several areas of disagreement between OMB and Congress about this executive branch tool, resulting in most congressional staff we spoke with not using the PART information. Most congressional staff reported that they would more likely use the PART results to inform their deliberations if OMB (1) consulted them early in the PART process regarding the selection and timing of programs to assess, (2) explained the methodology and evidence used or to be used to assess programs, and (3) discussed how the PART information can best be communicated and leveraged to meet their needs. Although OMB will be less likely to demonstrate the value of the PART beyond executive branch decision making without early consultation, OMB has had some success in engaging Congress when it has communicated selected PART results through legislative proposals and other traditional methods that clearly signal an executive branch priority. Although Congress currently has a number of opportunities to provide its perspective on specific performance issues and performance goals, opportunities also exist for Congress to enhance its institutional focus to enable a more systematic assessment of key programs and performance goals.

OMB Used a Variety of Methods to Communicate PART Scores, but Congressional Staff Raised Concerns about These Methods

OMB uses a variety of methods to communicate PART results to both the public and to Congress, primarily through the President's budget request documents, OMB's Web site, and meetings with some congressional staff. For example, OMB provides the single, bottom-line PART ratings in the *Analytical Perspectives* volume of the President's budget request, while the one-page PART summary sheets are available on a CD-ROM accompanying the President's budget request or on OMB's Web site. The Web site also contains the detailed supporting worksheets as well as other information about the tool itself. Certainly, OMB has provided more extensive information on program performance than in the past.

OMB's PART guidance also directed agencies to address the PART findings—from both current and prior year's PARTs—in their fiscal year 2006 budget submissions to OMB and budget justifications to Congress, as well as in testimony to Congress, in particular when a key budget or policy recommendation was influenced by a PART analysis. Agency witnesses testifying before the appropriations subcommittees did in fact include the results of the PART assessments in their written statements, and in some instances the PART was discussed during the "Q&A" portions of these hearings.

In addition to requiring agencies to inform Congress about the PART, OMB offered to brief congressional committees about the PART in 2004. According to OMB, packages including the PART summary sheets for programs that fell within each committee's jurisdiction and a list of the programs OMB planned to review for the fiscal year 2006 budget request were sent to all relevant House and Senate committees. An OMB senior official also said he followed up on these packages with phone calls, but received very little response. His records show that between February 2005 and June 2005 there were about 21 congressional meetings (bicameral and bipartisan) about the PART. In February 2005, upon the release of the *Major Savings and Reforms in the President's 2006 Budget* document, OMB held what it termed a briefing on the PART, inviting all appropriations staff.

OMB has set an ambitious benchmark for involving Congress in the PART process. In recent testimonies,[34] OMB's Deputy Director for Management stated that OMB's responsibility is to convince Congress that the PART assessments have correctly identified whether a program is working and, if not, what to do about it. In the past 3 years OMB states it has conducted 607 the PART assessments (about 60 percent of federal programs) that have generated nearly 1,800 recommendations. However, it is not clear that the PART has had any significant impact on congressional authorization, appropriations, and oversight activities to date. Moreover, it is unlikely that performance information will be used unless it is believed to be credible and reliable and reflects a consensus about performance goals among a community of interested parties. The PART likely has required a significant additional commitment of OMB's and agencies' resources, but demonstrating the value of the assessments beyond the executive branch will require further efforts.

Despite OMB's Efforts, Congressional Staff Said There Is Little Early Consultation on the PART

According to OMB senior officials, OMB's efforts generally focused on providing an overview of the PART process or communicating program assessment results to Congress rather than seeking early consultation about how the tool can best serve congressional needs. For example, upon the release of the *Major Savings and Reforms in the President's 2006 Budget* document, OMB said they invited leadership, appropriations, and budget committee staff to a presentation about it. However, some subcommittee staff said that the presentation was primarily intended to provide the *Major Savings* document that proposed program funding reductions and terminations, some of which were based on the PART assessments. Although some subcommittee staff said that they met with OMB and that OMB officials asked for their input about the PART, they did not see subsequent evidence that their views had been considered. Overall, most committee staff said that OMB generally did not involve them in the PART process.

The need for early consultation is clearly demonstrated by the strong interest House appropriators expressed in being consulted early in the PART process about the programs, activities, or projects that OMB intends

---

[34]*Accountability and Results in Federal Budgeting, Hearing before the Senate Comm. On Homeland Sec. & Governmental Affairs, Subcomm. On Fed. Fin. Mgmt, Gov't Info. & Int'l Sec.*, 109th Cong. (June 14, 2005) (Testimony of Clay S. Johnson, III, Deputy Dir., OMB). See also *Performance-Based Budgeting, Hearing before the House Budget Comm.*, 109th Cong. (July 20, 2005) (Testimony of Clay S. Johnson, III, Deputy Dir., OMB).

to assess for the fiscal years 2007 and 2008 budget requests, including approval of the methodology to be used to conduct each assessment. Congress went so far as to express these concerns in committee report language related to OMB's fiscal year 2006 appropriations.[35] Similar views were also echoed by many appropriations and authorizing committee staff we spoke with. As we have noted, some tension about the amount of stakeholder involvement in the internal deliberations surrounding the development of the PART measures and the broader consultations more common to the GPRA strategic planning process is inevitable.[36] Compared to the relatively open-ended GPRA process, any executive budget formulation process is likely to seem closed. However, if the PART is to be accepted as something more than an executive branch budget formulation tool, congressional understanding and acceptance of the tool and its analysis will be critical.

Because of Limited Agreement between OMB and Congress about the PART, Most Congressional Staff We Spoke with Do Not Use PART Information

A lack of early consultation has contributed to both congressional skepticism about the PART and to several areas of disagreement between OMB and Congress. As a result, most congressional staff we spoke with do not use PART information. Many committee staff we spoke with expressed frustration with the lack of available detail on how OMB arrived at their ratings of a program's performance. Many had concerns about the goals and measures used to assess program performance. Some subcommittee staff questioned the "unit of analysis" for the purposes of the PART as well as the design of the tool itself. The PART is OMB's tool of choice for assessing program performance and as such serves the administration's needs. However, it is only one source of information available to congressional committees.

Concerns Raised about Lack of Detail in PART Summary Worksheets and Unconvincing PART Assessments

Several committee staff were frustrated with the lack of detail provided on the PART summary sheets as to why a program was rated a certain way. They were unlikely to accept conclusions about a program's performance without seeing the evidence used to support them, particularly when the rating was contrary to what they believed to be true about a program. For example, some appropriations subcommittee staff expressed concerns

---

[35]*Depts. of Trans., Treasury, & Housing and Urban Dev., the Judiciary, Dist. of Columbia, & Indep. Agencies Appropriations Bill 2006*, H.R. Rep. 109-153, at 137-138 (2005).

[36]GAO, *21st Century Challenges: Performance Budgeting Could Help Promote Necessary Reexamination*, GAO-05-709T (Washington, D.C.: June 14, 2005).

about the "ineffective" PART rating given to the Health Professions program, which assists in paying for health professionals' education in exchange for their working in underserved areas. They said OMB could have made a stronger case for this rating if it had provided information showing that the program is unsuccessful in placing participating health professionals in underserved areas. In general, many committee staff we spoke with said they do not use the Web site containing the detailed supporting worksheets, primarily because finding this information on the Web site is too time-consuming, or the Web site is difficult to use.[37]

Although the detailed supporting worksheet for the Health Professions program notes that the agency has not conducted evaluations necessary to measure the program's performance—thus a factor for the "ineffective" rating—OMB's explanation of this rating is not clearly stated on the one-page summary sheet. Several committee staff said they wanted detailed information or criteria used to evaluate the program so that they could reach their own conclusions about program effectiveness. Some subcommittee staff felt that if OMB intends to request funding reductions or program eliminations based on PART assessments, a special burden exists to prove that these programs are ineffective.

In other cases, committee staff remained unconvinced about the PART ratings and the evidence used to support them. House appropriations subcommittee staff said that the Agricultural Credit Insurance Fund—Direct Loans, which they had held hearings on, was rated "moderately effective;" however, the subcommittee staff questioned the basis on which this program was given this rating since the agency has written off many of its loans. Committee staff also cited a PART assessment that stated that SBA's 7(a) loan program and its 504 program overlap because both provide long-term financing for similar borrowers. The committee staff disagreed with this assessment.[38]

---

[37]An OMB official recently testified that OMB recognizes that its Web site must more clearly communicate the PART information and that OMB is currently working on improvements to its Web site and the PART presentations.

[38]SBA officials said although on the surface the two programs appear to overlap, internal assessments show the differences between the programs. However, they said that OMB is constructively challenging SBA to articulate these differences using an independent, outside evaluation. SBA is seeking resources to pursue these independent evaluations.

Lack of Agreement about PART Measures and Defining an Appropriate "Unit of Analysis"	A lack of consultation early in the PART process has contributed to congressional committee staff not agreeing with or not finding useful OMB's choice or use of certain measures to determine the effectiveness of certain programs. Some committee staff reported that not all programs are well suited to being assessed by a tool like the PART. For example, a House subcommittee held a hearing in March 2004 that addressed concerns about defining acceptable PART measures for environmental research programs.[39] Hearing witnesses noted that OMB permitted some research programs to use output or process measures while it held similar programs to stricter standards, requiring them to use outcome measures. During a recent House Budget Committee hearing on performance budgeting, an OMB senior official agreed with committee members that the PART needs a set of goals and measures useful to OMB and Congress.[40] He added that consulting Congress early in the PART process, including discussions about how to make the PART useful for Congress, can better take place now that the PART has generated a critical mass of performance information.
	Some congressional staff were troubled by OMB's definition of certain programs—the "unit of analysis"—used for the PART assessments. They noted that what was useful for congressional budget deliberations sometimes differed from the unit of analysis OMB used to assess program performance in the PART. For example, appropriations subcommittee staff said that they often look at the performance of a particular project in determining how much funding to provide it. When OMB combines projects that are only loosely related by their authorizing statutes and rates them all as "ineffective" or "effective," this arrangement does not help Congress make trade-offs among those projects.
Committees Use a Variety of Performance Information	A few committee staff we talked with said that they use the PART information as one of many sources of information about program performance, including inspectors general reports, agency-commissioned evaluations, National Academy of Sciences reports, GAO reports, and National Academy of Public Administration reports. Several indications of congressional attention to the PART results were reflected in recent

---

[39]*Fiscal Year 2005 EPA Budget, Hearing before the House Comm. on Science, Subcomm. on Environment, Technology & Standards,* 108th Cong. (March 14, 2004).

[40]*Performance-Based Budgeting, Hearing before the House Budget Comm.,* 109th Cong. (July 20, 2005) (Testimony of Clay S. Johnson, III, Deputy Dir., OMB).

appropriations committee reports. For example, a House Appropriations Committee report on fiscal year 2006 appropriations cites a PART assessment stating among other things that performance measures have still not been developed and that effects on Pacific salmon stocks are still unknown.[41] The same committee applauds the Department of State's educational and cultural exchange programs (ECA), noting that "ECA received from the Office of Management and Budget Program Assessment Rating Tool [PART] ratings of 98 percent and 97 percent, the highest in the State Department and in the top one-percent in the Executive Branch."[42] Another House Appropriations Committee report for fiscal year 2006 noted that DOE's natural gas and petroleum/oil research and development programs received a poor PART score. In response, the committee encouraged the department to develop a strategic planning process that "demonstrates a clear path of investment that will yield demonstrable results, and better reflect the successes of these programs."[43]

The PART's focus on outcome measures may not fully appreciate congressional needs for other types of measures, such as output and workload information. Committee staff said they consider a variety of performance information such as outcome, output, and input measures to help gauge program performance. We have previously reported that congressional staff are interested in using a diverse array of information to address key questions on program performance, such as recurring information on spending priorities within programs; the quality, quantity, and efficiency of program operations; as well as the populations served or regulated.[44] Our recent work examining performance budgeting efforts at both the state and federal levels also bears this out. We found that appropriations committees consider workload and output measures important for making resource allocation decisions.[45] Workload and output

---

[41]*Science, State, Justice, Commerce, & Related Agencies Appropriations Bill, Fiscal Year 2006*, H.R. Rep. No. 109-118, at 100 (2005).

[42]Ibid, p. 125.

[43]*Energy and Water Dev. Appropriations Bill, 2006*, H.R. Rep. No. 109-86, at 112 (2005).

[44]GAO, *Managing for Results: Views on Ensuring the Usefulness of Agency Performance Information to Congress*, GAO/GGD-00-35 (Washington, D.C.: Jan. 26, 2000).

[45]GAO, *Performance Budgeting: States' Experiences Can Inform Federal Efforts*, GAO-05-215 (Washington, D.C.: Feb. 28, 2005); and *Performance Budgeting: Efforts to Restructure Budgets to Better Align Resources with Performance*, GAO-05-117SP (Washington, D.C.: February 2005).

measures lend themselves to the budget process because workload measures, in combination with cost-per-unit information, can help relate appropriation levels to a desired level of service.

Effective Communication and Consultation between OMB and Congress Could Increase the Role That the PART Plays in Congressional Deliberations

Despite its efforts, OMB has had limited success in engaging Congress in the PART process. For example, in June an OMB senior official testified that the PART had some effect on congressional authorizations, appropriations, or the oversight, but that OMB could clearly do a better job convincing Congress of the usefulness of performance information generated by the PART. Many majority and minority staff of House and Senate committees we talked with said that OMB should communicate the PART results in a way that meets individual committee needs. Most congressional committee staff said they would be more likely to use the PART results relevant to their committee responsibilities if OMB consulted with them early in the PART process and made PART information more useful for their work. They said it is important that such discussions also address performance information congressional committees find most useful. According to some staff, consulting them about congressional program priorities for PART assessments could be useful for linking these assessments to the authorization and appropriations processes by informing OMB about the committees' planned legislative agenda and informing Congress about programs OMB plans to assess in the near future. In discussing options for increasing congressional staff's access to performance information, we have previously noted that improved communication could go a long way to ensuring that congressional needs are understood and, where feasible, met.[46]

While some House and Senate committee staff stated that it would be difficult to conveniently time these consultations for both OMB and congressional staff, most agreed that they were a necessary step if Congress were to be able to use the PART to inform its deliberations. However, several majority and minority staff questioned how OMB could provide policy-neutral assessments given its institutional role. A couple of congressional subcommittee staff suggested that for any assessment to be considered credible it would have to be conducted or reviewed by an independent entity, such as a commission or a nonpartisan organization.

---

[46]GAO/GGD-00-35.

OMB has sometimes been able to engage Congress when it has communicated selected PART results through traditional means of signaling executive branch priorities, such as legislative proposals. For example, as discussed previously, the administration recently proposed to consolidate 18 federal CED programs, including the Community Development Block Grant (CDBG), into a single block grant, citing as one factor the low PART scores received in a crosscutting review of CED programs. The proposal led to hearings by several committees, involving administration officials, programs' stakeholders, and experts. Although the full House and Senate Appropriations Committee rejected the President's proposal to transfer the CDBG program to the Department of Commerce and instead kept the program at the Department of Housing and Urban Development, the House and Senate reduced the funding level for the CDBG formula grants by $250 million and $347 million, respectively, from last year's level.

Congress has initiated other hearings in which the PART has been a central subject of discussion. For example, OMB proposed funding cuts for the Environmental Protection Agency's science research grant programs (STAR) for the fiscal year 2005 budget because, according to a PART assessment, parts of STAR did not have adequate outcome measures and therefore could not demonstrate results. The Subcommittee on Environment, Technology, and Standards, House Committee on Science, held a hearing to discuss competing claims about whether these programs were contributing to their stated goals.

The fact that Congress has held such hearings indicates that certain PART reviews have captured congressional attention and contributed to the policy debate. As we have previously noted, success in performance budgeting should not be defined only by its effect on funding decisions but by the extent to which it changes the kinds of questions raised in Congress and executive agencies.[47] That is, performance budgeting helps shift the focus of congressional debates and oversight activities by changing the agenda of questions asked.

Congress has a number of opportunities to provide its perspective on specific performance issues and performance goals—when it establishes or reauthorizes a new program, during the annual appropriations process, and in its oversight of federal operations. Opportunities also exist for Congress

---

[47]GAO-05-709T.

to enhance its institutional focus to enable a more systematic assessment of key programs and performance goals. For example, identifying the key oversight and performance goals that Congress wishes to set for its own committees and for the government as a whole, perhaps for major missions such as budget functions could be useful. Collecting the "views and estimates" of authorization and appropriations committees on priority performance issues for programs under their jurisdiction and working with such crosscutting committees as the House Committee on Governmental Reform and the House Committee on Rules could be an initial step. Such a process might not only inform and better focus congressional deliberations, but could allow for more timely input into the PART.

It is important that Congress take full advantage of the benefits arising from the reform agenda under way in the executive branch. As we have suggested in the past, one approach to achieving the objective of enhancing congressional oversight is to develop a congressional performance resolution by modifying the current congressional budget resolution, which is already organized by budget function. Ultimately, what is important is not the specific approach or process, but rather the intended result of helping Congress better promote improved fiscal, management, and program performance through broad and comprehensive oversight and deliberation.

Conclusions and General Observations

The federal government is in a period of profound transition and faces an array of challenges and opportunities to enhance performance, ensure accountability, and position the nation for the future. A number of overarching trends—including the nation's long-term fiscal imbalance—drive the need to reexamine what the federal government does, how it does it, who does it, and how it gets financed. Performance budgeting holds promise as a means for facilitating a reexamination effort and bringing the panoply of federal activities in line with the demands of today's world. It can help enhance the government's capacity to assess competing claims for federal dollars and has the potential to better inform the budget debate.

PMA and its related initiatives, including the PART, demonstrate the administration's commitment to improving federal management and performance. Calling attention to successes and needed improvements is certainly a step in the right direction. The PART has helped perpetuate and sustain the performance culture ushered in by the management reforms of the 1990s. The PART has lent support to internal agency initiatives and—whatever criticism may be made regarding the value of scorecards and

bottom-line ratings—has highlighted the need for improvements and motivated agencies to do more.

There is no doubt that creating a closer link between the resources expended on programs and the results we expect from them is an important goal. The PART made a significant contribution by demonstrating one way to make a direct connection between performance and resource considerations. However, without truly integrating the PART and GPRA in a way that considers the differing needs of the budget formulation and strategic planning processes and their various stakeholders, OMB's ability to strengthen and further the performance-resource linkages for which GPRA laid the groundwork will continue to be hampered.

Successful integration of the inherently separate but interrelated GPRA strategic planning and the PART performance budgeting processes is predicated on (1) ensuring that the growing supply of performance information is credible, useful, reliable, and used; (2) increasing the demand for this information by developing goals and measures relevant to the large and diverse community of stakeholders in the federal budget and planning processes; and (3) taking a comprehensive and crosscutting approach.

By linking performance information to the budget process OMB has provided agencies with a powerful incentive for improving data quality and availability and has increased the potential for using performance information to inform the resource allocation process. To be effective, however, this information must not only be timely—to measure and affect performance—and reliable—to ensure consistent and comparable trend analysis over time and to facilitate better performance measurement and decision making—but also useful and used in order to make more informed operational and investing decisions.

Improvements in the quality of performance data and the capacity of federal agencies to perform program evaluations will require sustained commitment and investment of resources. However, evaluations can be very costly; opportunities exist to carefully target federal evaluation resources such that the American people receive the most benefit from each evaluation dollar spent. Moreover, the question of investment in improved evaluation capacity is one that must be considered in budget deliberations both within the executive branch and in Congress. Importantly, it is critical that budgetary investments in this area be viewed

as part of a broader initiative to improve the accountability and management capacity of federal agencies and programs.

Some program improvements related to the PART's success—such as improving program outcomes, taking steps to address PART findings, improving program management, and becoming more efficient—can often come solely through executive branch action, but for the PART to meet its full potential the assessments it generates must also be meaningful to and used by Congress and other stakeholders. For the PART to result in congressional action on the PART's funding and policy recommendations as OMB desires, the PART must hold appeal beyond the executive branch. The PART was designed for and is used in the executive branch budget preparation and review process; as such, the goals and measures used in the PART must meet OMB's needs. Because OMB has not developed an effective strategy for connecting the PART process to congressional needs, Congress generally does not use the PART in its deliberations. Without developing an effective strategy for obtaining and acting on congressional views on what to measure, how to measure it, and how to best present this information to a congressional audience, it is more likely that PART will remain an executive branch exercise largely ignored in the authorization, appropriations, and oversight processes. Infusing a performance perspective into budget decisions may only be achieved when we discover ways to reflect both the broader planning perspective that can add value to budget deliberations and foster accountability in ways that Congress considers appropriate for meeting its roles, responsibilities, and interests. Congress also can facilitate the use of performance information by enhancing its focus on performance in budget, authorizing, appropriations, and oversight processes.

Looking forward, opportunities exist to develop a more strategic approach to selecting and prioritizing areas for assessment under the PART process. Targeting PART assessments based on such factors as the relative priorities, costs, and risks associated with related clusters of programs and activities addressing related strategic and performance goals not only could help ration scarce analytic resources but also could focus decision makers' attention on the most pressing policy and program issues. Moreover, key outcomes in areas ranging from low income housing to food safety to counterterrorism are addressed by a wide range of discretionary, entitlement, tax, and regulatory approaches that cut across a number of agencies. Some tax expenditures amount to hundreds of billions of dollars of annual expenditures—the same order of magnitude as total discretionary spending, yet relatively little is known about the effectiveness

of tax incentives in achieving the objectives intended by Congress. Broadening the PART to assess complete portfolios of tools used to achieve key federal outcomes is absolutely critical. A crosscutting approach could also facilitate the use of the PART assessments to review the relative contributions of similar programs to common or crosscutting goals and outcomes established through the GPRA process.

As we have previously reported, effective congressional oversight can help improve federal performance by examining the program structures agencies use to deliver products and services to ensure that the best, most cost-effective mix of strategies is in place to meet agency and national goals. While Congress has a number of opportunities to provide its perspective on performance issues and performance goals, such as when it establishes or reauthorizes a program, during the annual appropriations process, and in its oversight of federal operations, a more systematic approach could allow Congress to better articulate performance goals and outcomes for key programs of major concern. Such an approach could also facilitate OMB's understanding of congressional priorities and concerns and, as a result, increase the usefulness of the PART in budget deliberations.

Matter for Congressional Consideration

To facilitate an understanding of congressional priorities and concerns, Congress should consider the need for a strategy that includes (1) establishing a vehicle for communicating performance goals and measures for key congressional priorities and concerns; (2) developing a more structured oversight agenda to permit a more coordinated congressional perspective on crosscutting programs and policies; and (3) using such an agenda to inform its authorization, oversight, and appropriations processes.

Recommendations for Executive Action

We make three recommendations to OMB. We recommend that the Director of OMB take the following actions:

- Ensure that congressional leadership and key committees are given an opportunity to provide input early in the PART process on the performance issues and program areas they consider to be the most important and in need of review.

- Seek input from congressional committees on the performance information they find useful and how that information could best be presented to them.

- Target individual programs to be reassessed based on factors such as the relative priorities, costs, and risks associated with clusters of related programs, and in a way that reflects the congressional input described above.
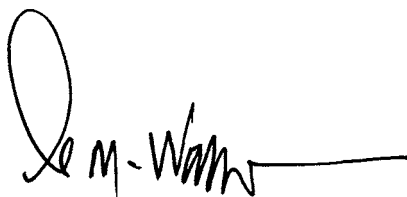
Agency Comments and Our Evaluation

In commenting on a draft of this report, OMB generally agreed with our findings, conclusions, and recommendations. OMB outlined several actions it is taking to address some of the issues raised in the report, including implementing information technology solutions to make application of the PART less burdensome and more collaborative. OMB also suggested some technical changes throughout the report that we have incorporated as appropriate. OMB's comments appear in appendix IV. We also received technical comments on excerpts of the draft from the Departments of Labor and Health and Human Services, which are incorporated as appropriate.

We are sending copies of this report to the Director of the Office of Management and Budget and other interested parties. We will also make copies available to others upon request. In addition, the report will be available at no charge on GAO's Web site at http://www.gao.gov.

If you or your staff have any questions about this report please contact Susan Irving at (202) 512-9142 or irvings@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the

last page of this report. GAO staff making key contributions to this report are listed in appendix V.

Sincerely yours,

David M. Walker
Comptroller General
of the United States

Scope and Methodology

To address the first two objectives, we reviewed the Office of Management and Budget's (OMB) materials on the implementation, application, and revision of the Program Assessment Rating Tool (PART) for calendar years 2002 through 2004.[1] We also interviewed OMB branch chiefs and OMB staff on the Performance Evaluation Team (PET). The PET's role is to provide guidance to budget examiners and help ensure consistent application of the PART across OMB offices. To better understand OMB's experience with crosscutting reviews, we interviewed OMB staff responsible for coordinating the Community and Economic Development and Rural Water crosscutting reviews conducted for the fiscal year 2006 President's budget request. To obtain agency perspectives on the relationship between the PART and the Government Performance and Results Act of 1993 (GPRA) and their interactions with OMB concerning that relationship, we interviewed department and agency officials, including senior managers, and program, planning, and budget staffs at (1) the Department of Health and Human Services (HHS), (2) the Department of Energy (DOE), (3) the Department of Labor (DOL), and (4) the Small Business Administration (SBA). We also interviewed officials from these departments and agencies concerning their perspectives and activities in response to the PART recommendations and the effects of implementing those recommendations on operations and results.

We selected these three departments and one independent agency for a number of reasons. Collectively, they offered examples of all seven PART program types (e.g., block/formula grants, competitive grants, direct federal, and research and development) for review. These examples covered about a fifth of all the programs subject to the PART as of 2004 and thus could provide a broad-based perspective on how the PART was applied. We also chose to return to HHS and DOE—two of the departments included in our previous study on the PART.[2] To broaden our coverage of agency perspectives we selected DOL and SBA because they had received a "green" score on their President's Management Agenda Executive Branch Management Scorecard for the budget and performance integration initiative and were considered good candidates for showing progress. Our selection of these four agencies was also influenced by our intent to integrate this work with our related work examining progress in addressing the PART program evaluation recommendations. Approximately half of the

---

[1]This period covers fiscal budget years 2004-2006.

[2]See GAO-04-174.

evaluation recommendations in the 2002 PART were encompassed in our four case selections.

As part of our work on the second objective, we also performed various analyses of the PART recommendations made in all 3 years to discern possible changes or trends in recommendations over time and relationships between the type of recommendations made, type of program, overall rating, total PART score, and answers to selected PART questions. To do these analyses, we classified the recommendations OMB made into the same four categories we used in our prior report, i.e., program assessment, program design, program management, and funding. We employed a slightly modified classification procedure from our previous review, which included the addition of an "other" category for recommendations that did not fit within any of the four categories.[3] We then combined the results of our recommendation classifications with selected data we downloaded from PART summaries and worksheets posted on OMB's PART Web sites, data developed for our previous report of the 2002 PART, and a data set provided by OMB of programs covered in the 2004 PART.[4] In addition, we also examined relevant OMB and agency documents to help determine how recommendations are tracked and their impact evaluated by OMB and the selected agencies.

To address our third objective of examining the steps OMB has taken to involve Congress in the PART process, we interviewed OMB and agency officials and asked questions about the steps OMB and agencies have taken to involve Congress in the PART process or in using the results of the PART. To obtain documented instances of Congress' uses and views of the PART, we interviewed House and Senate committee staff (minority and majority) for the authorizing and appropriations subcommittees with jurisdiction over our selected agencies as well as OMB and officials from the four selected agencies. Finally, we reviewed fiscal years 2005 and 2006 House and Senate congressional hearing records and reports as well as conference reports for mentions of the PART. In addition, where possible,

---

[3]As a result, the percentage distribution of recommendation categories for 2002 differs somewhat from our prior report. See GAO-04-174.

[4]While this combined data set includes all programs covered in the 2004 PART, it does not include those programs subject to the PART in 2002 or 2003 and subsequently merged or incorporated into other programs or dropped. These programs were excluded to help ensure that our analyses did not include programs that, as defined in either 2002 or 2003, no longer exist.

we corroborated testimonial evidence with documentary evidence of OMB's and agencies' strategies for involving Congress as well as evidence of collaboration and coordination, such as planning documents, briefing material, or other evidence of contact with Congress.

We did not independently verify the PART assessments as posted on OMB's Web sites; however, we did take several steps to ensure that we reliably downloaded and combined the various data files with our recommendation classifications. Our steps included (1) having the computer programs we used to create and process our consolidated dataset verified by a second programmer; (2) performing various edit checks on the data and (3) selecting computer-processed data elements checked back to source files for a random sample of programs and also for specific programs identified in our analyses or through edit checks. We determined that the data were reliably downloaded and combined, and sufficient for the purposes of this report.

While our summary analyses include all or almost all programs subject to the PART for the years 2002 to 2004 or all or almost all programs within a specified subset of programs (e.g., program type, specific year, specific rating), the information obtained from OMB, congressional and agency officials, as well as documentary material from the selected agencies is not generalizable to the PART process for all years or all programs.

We conducted our audit work from January 2005 through August 2005 in accordance with generally accepted government auditing standards. OMB provided written comments on this draft that are reprinted in appendix IV.

The 2004 PART Questionnaire

Section I: Program Purpose & Design (Yes, No, N/A)

1. Is the program purpose clear?

2. Does the program address a specific and existing problem, interest, or need?

3. Is the program designed so that it is not redundant or duplicative of any other Federal, State, local or private effort?

4. Is the program design free of major flaws that would limit the program's effectiveness or efficiency?

5. Is the program design effectively targeted, so that resources will reach intended beneficiaries and/or otherwise address the program's purpose directly?

Section II: Strategic Planning (Yes, No, N/A)

1. Does the program have a limited number of specific long-term performance measures that focus on outcomes and meaningfully reflect the purpose of the program?

2. Does the program have ambitious targets and timeframes for its long-term measures?

3. Does the program have a limited number of specific annual performance measures that can demonstrate progress toward achieving the program's long-term goals?

4. Does the program have baselines and ambitious targets for its annual measures?

5. Do all partners (including grantees, sub-grantees, contractors, cost-sharing partners, and other government partners) commit to and work toward the annual and/or long-term goals of the program?

6. Are independent evaluations of sufficient scope and quality conducted on a regular basis or as needed to support program improvements and evaluate effectiveness and relevance to the problem, interest, or need?

7. Are Budget requests explicitly tied to accomplishment of the annual and long-term performance goals, and are the resource needs presented in a complete and transparent manner in the program's budget?

8. Has the program taken meaningful steps to correct its strategic planning deficiencies?

Specific Strategic Planning Questions by Program Type

Regulatory-Based Programs

RG1. Are all regulations issued by the program/agency necessary to meet the stated goals of the program, and do all regulations clearly indicate how the rules contribute to achievement of the goals?

Capital Assets and Service Acquisition Programs

CA1. Has the agency/program conducted a recent, meaningful, credible analysis of alternatives that includes trade-offs between cost, schedule, risk, and performance goals and used the results to guide the resulting activity?

Research and Development Programs

RD1. If applicable, does the program assess and compare the potential benefits of efforts within the program and (if relevant) to other efforts in other programs that have similar goals?

RD2. Does the program use a prioritization process to guide budget requests and funding decisions?

Section III: Program Management (Yes, No, N/A)

1. Does the agency regularly collect timely and credible performance information, including information from key program partners, and use it to manage the program and improve performance?

2. Are Federal managers and program partners (including grantees, sub-grantees, contractors, cost-sharing partners, and other government partners) held accountable for cost, schedule and performance results?

3. Are funds (Federal and partners') obligated in a timely manner and spent for the intended purpose?

4. Does the program have procedures (e.g., competitive sourcing/cost comparisons, IT improvements, appropriate incentives) to measure and achieve efficiencies and cost effectiveness in program execution?

5. Does the program collaborate and coordinate effectively with related programs?

6. Does the program use strong financial management practices?

7. Has the program taken meaningful steps to address its management deficiencies?

Specific Program Management Questions by Program Type

Competitive Grant Programs

CO1. Are grants awarded based on a clear competitive process that includes a qualified assessment of merit?

CO2. Does the program have oversight practices that provide sufficient knowledge of grantee activities?

CO3. Does the program collect grantee performance data on an annual basis and make it available to the public in a transparent and meaningful manner?

Block/Formula Grant Programs

BF1. Does the program have oversight practices that provide sufficient knowledge of grantee activities?

BF2. Does the program collect grantee performance data on an annual basis and make it available to the public in a transparent and meaningful manner?

Regulatory-Based Programs

RG1. Did the program seek and take into account the views of all affected parties (e.g., consumers; large and small businesses; State, local and tribal

governments; beneficiaries; and the general public) when developing significant regulations?

RG2. Did the program prepare adequate regulatory impact analyses if required by Executive Order 12866, regulatory flexibility analyses if required by the Regulatory Flexibility Act and SBREFA, and cost-benefit analyses if required under the Unfunded Mandates Reform Act; and did those analyses comply with OMB guidelines?

RG3. Does the program systematically review its current regulations to ensure consistency among all regulations in accomplishing program goals?

RG4. Are the regulations designed to achieve program goals, to the extent practicable, by maximizing the net benefits of its regulatory activity?

Capital Assets and Service Acquisition Programs

CA1. Is the program managed by maintaining clearly defined deliverables, capability/performance characteristics, and appropriate, credible cost and schedule goals?

Credit Programs

CR1. Is the program managed on an ongoing basis to assure credit quality remains sound, collections and disbursements are timely, and reporting requirements are fulfilled?

CR2. Do the program's credit models adequately provide reliable, consistent, accurate and transparent estimates of costs and the risk to the Government?

Research and Development Programs

RD1. For R&D programs other than competitive grants programs, does the program allocate funds and use management processes that maintain program quality?

Section IV: Program Results/Accountability (Yes, Large Extent, Small Extent, No)

1. Has the program demonstrated adequate progress in achieving its long-term performance goals?

2. Does the program (including program partners) achieve its annual performance goals?

3. Does the program demonstrate improved efficiencies or cost effectiveness in achieving program goals each year?

4. Does the performance of this program compare favorably to other programs, including government, private, etc., with similar purpose and goals?

5. Do independent evaluations of sufficient scope and quality indicate that the program is effective and achieving results?

Specific Results Questions by Program Type

Regulatory-Based Programs

RG1. Were programmatic goals (and benefits) achieved at the least incremental societal cost and did the program maximize net benefits?

Capital Assets and Service Acquisition Programs

CA1. Were program goals achieved within budgeted costs and established schedules?

Sample PART Summary Worksheets

Program: *Agricultural Credit Insurance Fund - Guaranteed*

Agency: *Department of Agriculture*
Bureau: *Farm Service Agency*

Rating: *Moderately Effective*

Program Type: *Credit*

Last Assessed: *1 year ago*

Key Performance Measures from Latest PART	Year	Target	Actual
Long-term Measure: Maintain a low loss rate on guaranteed loans	2002	2%	1%
	2003	1%	1%
	2004	<1%	
	2005	<1%	
Long-term Measure: Increase the percent of loans to beginning and socially disadvantaged farmers/ranchers	2002		32%
	2003	34%	33%
	2004	38%	
	2005	36%	
Annual Measure: Decrease in loan average processing times (days)	2002		16
	2003	15.5	14
	2004	14	
	2005	14	

Recommended Follow-up Actions	Status
Assess performance targets to ensure they are ambitious.	Action taken, but not completed
Conduct a performance-focused review that will include, but is not limited to: analysis of program participants; length of time borrowers remain in program; number of borrowers who 'graduate' and return to the program; effectiveness of targeted assistance; and the potential to reduce subsidy rates.	Action taken, but not completed
Develop an efficiency measure such as 'cost per loan processed' to track administrative expenses and allow comparison among loan programs.	Completed
Revise long-term performance measure to better assess progress toward meeting the goal of improving economic viability of farmers/ranchers.	Action taken, but not completed

Update on Follow-up Actions:

FSA participated in the USDA Credit Programs Common Efficiency Measure initiative along with FAS, RD, OBPA, and OMB to develop an efficiency measure to be used by al USDA agencies with credit programs: Maintain or reduce operating expense ratio for average loan portfolio. In addition, the PART evaulation contained a recommendation to conduct a performance-focused review of the farm loan program. This review is being completed by an independent contractor and the results will be used to assess effectiveness of guaranteed loans, as applicable. Estimated completion date is 7/30/2006. FSA is developing new, outcome oriented performance measures as part of the agency's strategic planning process and the development of the new FSA Strategic Plan.

Program Funding Level (in millions of dollars)

2004 Actual	2005 Estimate	2006 Estimate
2,402	2,763	2,866

Program: *Migratory Bird Program*

Agency: *Department of the Interior*
Bureau: *Fish and Wildlife Service Division of Migratory Birds*

Purpose	**100**
Planning	**70**
Management	**85**

Results / Accountability	**22**

0   **100**

Key Performance Measures from Latest PART	*Year*	*Target*	*Actual*
Long-term Measure: Percent of all migratory bird species that are at healthy and sustainable levels.	2001	61.8%	Baseline
	2005	61.8%	
	2008	62.3%	
Long-term Measure: Percent of adult Americans who participate in bird-related recreation.	2001	29.8%	Baseline
	2005	29.8%	
	2011	30%	
Annual Measure: Percent of bird population management needs met to achieve healthy and sustainable populations of birds listed on the Birds of Management Concern list. (Baseline and targets under development.)	2005	Baseline	

Rating: *Results Not Demonstrated*
Program Type: *Direct Federal*

Program Summary:

The Migratory Bird Program of the U.S. Fish and Wildlife Service is responsible for maintaining healthy migratory bird populations for the benefit of the American people. The program accomplishes this by conserving and restoring migratory bird populations, restoring and acquiring migratory bird habitat, surveying and monitoring migratory birds, and regulating the take of migratory birds. The program works closely with many partners to ensure the conservation of the birds.

The assessment revealed that the program has a clear mission and has undertaken some valuable strategic planning activities with partners. Specific findings include:

- While the program has a strategic plan that identifies three strategic goals and supporting strategies, the program did not have specific long-term outcome or annual output performance goals. Through the PART process, specific long-term outcome or annual output performance goals were developed.
- There are no regular objective, independent program performance evaluations of the entire program.
- Budget requests have not been explicitly tied to long-term performance goals.
- Program regulations have not been systematically reviewed to ensure consistency in accomplishing program goals or if the program is using the least intrusive and most efficient approach.
- While the program is working to incorporate performance goals into specific employee performance plans, the program needs to complete this task to ensure full accountability for achieving specific program goals.

In response to the PART findings, the Administration will:
1. Adopt long-term outcome and annual output goals developed during PART process. Accomplishment of the outcome goals will depend on the efforts of many and will require the program to continue to work with partners to achieve these goals.
2. Request additional funding in the Budget to develop and implement management plans for five migratory bird species to help achieve the program's new long-term goal to increase the percentage of migratory birds that are healthy and sustainable.
3. Develop baseline data and revise targets as necessary for new performance measures.
4. Schedule and carry out independent program evaluations, including the regulatory part of the program.
5. Link individual employee performance plans with specific goal-related performance targets for each year.

Program Funding Level (in millions of dollars)

2004 Actual	**2005 Estimate**	**2006 Estimate**
119	129	141

Comments from the Executive Office of the President

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF MANAGEMENT AND BUDGET
WASHINGTON, D.C. 20503

DEPUTY DIRECTOR
FOR MANAGEMENT

October 17, 2005

Ms. Susan J. Irving
Director for Federal Budget Analysis
Government Accountability Office
441 G Street, NW
Washington, DC 20548

Dear Ms. Irving:

Thank you for the opportunity to comment on the draft GAO report on program evaluation (*Performance Budgeting: PART Focuses Attention on Program Performance, But More Can Be Done to Engage Congress*, GAO-06-28).

We appreciate GAO's continued interest in the Program Assessment Rating Tool (PART) and our determination to assess federal programs in a consistent fashion through it. As is acknowledged in your conclusion, "There is no doubt that creating a closer link between the resources expended on programs and the results we expect from those is an important goal."[1] We fervently believe that the PART has helped do just that, and we are grateful for any guidance you can provide that will help us achieve even better results.

In this same spirit, OMB and agencies continue to search for ways to make PART assessments more rigorous and consistent. Additionally, we are implementing information technology solutions to make application of the PART less burdensome and more collaborative. Moreover, we reviewed each newly completed PART this year to ensure the answers were consistent with PART guidance. These steps and others will make the PART more reliable, less of a burden, and hopefully, more focused on identifying what steps programs need to take to become more effective.

In many cases, it takes only administrative actions to address weaknesses in program efficiency and effectiveness and the PART process has helped do just that. But where Congressional action is required to ameliorate a program flaw, GAO correctly points out that PART has been less successful. OMB and agencies are grateful for any specific suggestions GAO may have to obtain greater Congressional support for our initiative to improve the performance of all programs.

OMB notes the particular interest that GAO has taken in the Administration's standards for measuring performance. Thank you for your continued enthusiasm about the PART, as well as for your willingness to take our oral and written comments into consideration in the final

---

[1] *See*, draft report "Conclusions and General Observations," p. 58.

draft. I look forward to working with you to improve the ways in which we are making the Federal Government more results-oriented.

Sincerely,

Clay Johnson III

GAO Contact and Staff Acknowledgments

GAO Contact

Susan Irving (202) 512-9142

Acknowledgments

In addition to the contact named above, Denise Fantone (Assistant Director), Thomas Beall, Kylie Gensimore, Joseph Leggero, Patrick Mullen, Jacqueline Nowicki, Stephanie Shipman, Katherine Wulff, and James Whitcomb made significant contributions to this report. Dianne Blank and Amy Rosewarne also provided key assistance.

Related GAO Products

21st Century Challenges: Performance Budgeting Could Help Promote Necessary Reexamination. GAO-05-709T. Washington, D.C.: June 14, 2005.

Management Reform: Assessing the President's Management Agenda. GAO-05-574T. Washington, D.C.: April 21, 2005.

Performance Budgeting: States' Experiences Can Inform Federal Efforts. GAO-05-215. Washington, D.C.: February 28, 2005.

21st Century Challenges: Reexamining the Base of the Federal Government. GAO-05-352T. Washington, D.C.: February 16, 2005.

Long-Term Fiscal Issues: Increasing Transparency and Reexamining the Base of the Federal Budget. GAO-05-317T. Washington, D.C.: February 8, 2005.

21st Century Challenges: Reexamining the Base of the Federal Government. GAO-05-325SP. Washington, D.C.: February 2005.

Performance Budgeting: Efforts to Restructure Budgets to Better Align Resources with Performance. GAO-05-117SP. Washington, D.C.: February 2005.

Performance Budgeting: OMB's Performance Rating Tool Presents Opportunities and Challenges for Evaluating Program Performance. GAO-04-550T. Washington, D.C.: March 11, 2004.

Results-Oriented Government: GPRA Has Established a Solid Foundation for Achieving Greater Results. GAO-04-38. Washington, D.C.: March 10, 2004.

Performance Budgeting: OMB's Program Assessment Rating Tool Presents Opportunities and Challenges for Budget and Performance Integration. GAO-04-439T. Washington, D.C.: February 4, 2004.

Performance Budgeting: Observations on the Use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget. GAO-04-174. Washington, D.C.: January 30, 2004.

Results-Oriented Government: Using GPRA to Address 21st Century Challenges. GAO-03-1166T. Washington, D.C.: September 18, 2003.

Performance Budgeting: Current Developments and Future Prospects. GAO-03-595T. Washington, D.C.: April 1, 2003.

Performance Budgeting: Opportunities and Challenges. GAO-02-1106T. Washington, D.C.: September 19, 2002.

Managing for Results: Views on Ensuring the Usefulness of Agency Performance Information to Congress. GAO/GGD-00-35. Washington, D.C.: January 26, 2000.

Performance Budgeting: Fiscal Year 2000 Progress in Linking Plans With Budgets. GAO/AIMD-99-239R. Washington, D.C.: July 30, 1999.

Performance Budgeting: Initial Experiences Under the Results Act in Linking Plans With Budgets. GAO/AIMD/GGD-99-67. Washington, D.C.: April 12, 1999.

Managing for Results: Measuring Program Results That Are Under Limited Federal Control. GAO/GGD-99-16. Washington, D.C.: December 11, 1998.

The Government Performance and Results Act: 1997 Governmentwide Implementation Will Be Uneven. GAO/GGD-97-109. Washington, D.C.: June 2, 1997.

Managing for Results: Analytic Challenges in Measuring Performance. GAO/HEHS/GGD-97-138. Washington, D.C.: May 30, 1997.

Performance Budgeting: Past Initiatives Offer Insights for GPRA Implementation. GAO/AIMD-97-46. Washington, D.C.: March 27, 1997.

Executive Guide: Effectively Implementing the Government Performance and Results Act. GAO/GGD-96-118. Washington, D.C.: June 1996.

Managing for Results: Achieving GPRA's Objectives Requires Strong Congressional Role. GAO/T-GGD-96-79. Washington, D.C.: March 6, 1996.