



DEFENSE TECHNICAL INFORMATION CENTER

Information for the Defense Community

DTIC® has determined on 10/28/2010 that this Technical Document has the Distribution Statement checked below. The current distribution for this document can be found in the DTIC® Technical Report Database.

DISTRIBUTION STATEMENT A. Approved for public release; distribution is unlimited.

© COPYRIGHTED; U.S. Government or Federal Rights License. All other rights and uses except those permitted by copyright law are reserved by the copyright owner.

DISTRIBUTION STATEMENT B. Distribution authorized to U.S. Government agencies only (fill in reason) (date of determination). Other requests for this document shall be referred to (insert controlling DoD office)

DISTRIBUTION STATEMENT C. Distribution authorized to U.S. Government Agencies and their contractors (fill in reason) (date of determination). Other requests for this document shall be referred to (insert controlling DoD office)

DISTRIBUTION STATEMENT D. Distribution authorized to the Department of Defense and U.S. DoD contractors only (fill in reason) (date of determination). Other requests shall be referred to (insert controlling DoD office).

DISTRIBUTION STATEMENT E. Distribution authorized to DoD Components only (fill in reason) (date of determination). Other requests shall be referred to (insert controlling DoD office).

DISTRIBUTION STATEMENT F. Further dissemination only as directed by (inserting controlling DoD office) (date of determination) or higher DoD authority.

Distribution Statement F is also used when a document does not contain a distribution statement and no distribution statement can be determined.

DISTRIBUTION STATEMENT X. Distribution authorized to U.S. Government Agencies and private individuals or enterprises eligible to obtain export-controlled technical data in accordance with DoDD 5230.25; (date of determination). DoD Controlling Office is (insert controlling DoD office).

H-14
30002097

United States General Accounting Office
Report to Congressional Requesters

REFERENCE COPY

January 2004

PERFORMANCE BUDGETING

Observations on the
Use of OMB's Program
Assessment Rating
Tool for the Fiscal
Year 2004 Budget

20101015319

H-14
30002097



G A O

Accountability * Integrity * Reliability

30002097



Highlights of GAO-04-174, a report to congressional requesters

Why GAO Did This Study

The Office of Management and Budget's (OMB) Program Assessment Rating Tool (PART) is meant to provide a consistent approach to evaluating federal programs during budget formulation. To better understand its potential, congressional requesters asked GAO to examine (1) how PART changed OMB's fiscal year 2004 budget decision-making process, (2) PART's relationship to the Government Performance and Results Act of 1993 (GPRA), and (3) PART's strengths and weaknesses as an evaluation tool.

What GAO Recommends

GAO recommends that OMB (1) address the capacity demands of PART, (2) strengthen PART guidance, (3) address evaluation information availability and scope issues, (4) focus program selection on crosscutting comparisons and critical operations, (5) broaden the dialogue with congressional stakeholders, and (6) articulate and implement a complementary relationship between PART and GPRA.

OMB generally agreed with our findings, conclusions, and recommendations and stated that it is already taking actions to address many of our recommendations.

GAO also suggests that Congress consider the need for a structured approach to articulating its perspective and oversight agenda on performance goals and priorities for key programs.

www.gao.gov/cgi-bin/getrpt?GAO-04-174.

To view the full product, including the scope and methodology, click on the link above. For more information, contact Paul Posner at (202) 512-9573 or posnerp@gao.gov.

PERFORMANCE BUDGETING

Observations on the Use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget

What GAO Found

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. OMB and agency staff said this helped OMB staff with varying levels of experience focus on similar issues.

Our analysis confirmed that one of PART's major impacts was its ability to highlight OMB's recommended changes in program management and design. Much of PART's potential value lies in the related program recommendations, but realizing these benefits requires sustained attention to implementation and oversight to determine if desired results are achieved. OMB needs to be cognizant of this as it considers capacity and workload issues in PART.

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 PART guidance and defining acceptable measures. 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."

PART is not well integrated with GPRA—the current statutory framework for strategic planning and reporting. By using the PART process to review and sometimes replace GPRA goals and measures, OMB is substituting 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 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. Moreover, PART does not currently evaluate similar programs together to facilitate trade-offs or make relative comparisons.

PART clearly must serve the President's interests. However, 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. It will be important for OMB to discuss timely with Congress the focus of PART assessments and clarify the results and limitations of PART and the underlying performance information. 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 PART's usefulness in budget deliberations.

Contents

Letter

Results in Brief	1
Background	4
OMB Used the PART to Systematically Assess Program Performance and Make Results Known, but Follow-up on PART Recommendations Is Uncertain	8
Despite OMB's Considerable Efforts to Create a Credible Evaluation Tool, PART Assessments Require Judgment and Were Constrained by Data Limitations	11
Subjective Terms and a Restrictive Format Contributed to Subjective and Inconsistent Responses	17
There Were Inconsistencies in Defining Acceptable Measures and in Logically Responding to Question "Pairs"	20
Disagreements on Performance Information Led to Creation of a "Results Not Demonstrated" Category	21
The Fiscal Year 2004 PART Process Was a Parallel, Competing Approach to GPRA's Performance Management Framework	25
Defining a "Unit of Analysis" That Is Useful for Program-Level Budget Analysis and Agency Planning Purposes Presents Challenges	27
Conclusions and General Observations	29
Matter for Congressional Consideration	33
Recommendations for Executive Action	36
Agency Comments	36

Appendixes

Appendix I: Scope and Methodology	39
Appendix II: The Fiscal Year 2004 PART and Differences Between the Fiscal Year 2004 and 2005 PARTs	48
Section I: Program Purpose & Design (Yes, No, N/A)	48
Section II: Strategic Planning (Yes, No, N/A)	48
Section III: Program Management (Yes, No, N/A)	48
Section IV: Program Results (Yes, Large Extent, Small Extent, No)	50
Appendix III: Development of PART	53
Fiscal Year 2003	62
Fiscal Year 2004	62
Fiscal Year 2005	64
Appendix IV: Comments from the Office of Management and Budget	64

Appendix V: GAO Contacts and Staff Acknowledgments	67
GAO Contacts	67
Acknowledgments	67

Tables

Table 1: Overview of Sections of PART Questions	9
Table 2: Overview of PART Program Types	10
Table 3: The Effect of Overall PART Score on Proposed Funding Changes (Discretionary Programs)	43
Table 4: The Effect of Overall PART Score on Proposed Funding Changes (Small Discretionary Programs)	44
Table 5: The Effect of Overall PART Score on Proposed Funding Changes (Medium-Size Discretionary Programs)	44
Table 6: The Effect of Overall PART Score on Proposed Funding Changes (Large Discretionary Programs)	44
Table 7: The Effect of PART Component Scores on Proposed Funding Changes (All Discretionary Programs)	45
Table 8: The Effect of PART Component Scores on Proposed Funding Changes (Small Discretionary Programs)	45
Table 9: Side-by-Side of the Fiscal Year 2005 PART and the Fiscal Year 2004 PART Questions	54

Figures

Figure 1: Fiscal Year 2004 PART Recommendations	13
Figure 2: Number of Discretionary PART Programs by Rating and Funding Result, Fiscal Years 2003-2004	14
Figure 3: The PART Process and Budget Formulation Timelines	64

This is a work of the U.S. government and is not subject to copyright protection in the United States. It may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



G A O

Accountability • Integrity • Reliability

United States General Accounting Office
Washington, D.C. 20548

January 30, 2004

The Honorable George V. Voinovich
Chairman
Subcommittee on Oversight of Government Management, the Federal
Workforce and the District of Columbia
Committee on Governmental Affairs
United States Senate

The Honorable Todd R. Platts
Chairman
Subcommittee on Government Efficiency and Financial Management
Committee on Government Reform
House of Representatives

The Honorable Sam Brownback
United States Senate

The Honorable Todd Tiahrt
House of Representatives

Since the 1950s, the federal government has attempted several governmentwide initiatives designed to better align spending decisions with expected performance—what is often commonly referred to as “performance budgeting.” Consensus exists that prior efforts—including the Hoover Commission, the Planning-Programming-Budgeting-System (PPBS), Management by Objectives, and Zero-Based Budgeting (ZBB)—failed to significantly shift the focus of the federal budget process from its long-standing concentration on the items of government spending to the results of its programs.

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 1993¹ (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

¹ Pub. L. No. 103-62 (1993).

and the expected results to be achieved with those resources. This type of integration is critical, as we have learned from prior initiatives that failed in part because they did not prove to be relevant to budget decision makers in the executive branch or Congress.² GPRA requires not only a connection to the structures used in congressional budget presentations but also consultation between the executive and legislative branches on agency strategic plans, which gives Congress an oversight stake in GPRA's success.³

In its overall structure, focus, and approach GPRA incorporates two critical lessons learned from previous reforms. First, any approach designed to link plans and budgets—that is, to link the responsibility of the executive to define strategies and approaches with the legislative “power of the purse”—must explicitly involve both branches of our government. PPBS and ZBB failed in part because performance plans and measures were developed in isolation from congressional oversight and resource allocation processes.

Second, the concept of performance budgeting has and likely will continue to evolve. Thus, no single definition of performance budgeting encompasses the range of past and present needs and interests of federal decision makers. The need for multiple definitions reflects the differences in the roles various participants play in the budget process. And, given the complexity and breadth of the federal budget process, performance budgeting must encompass a variety of perspectives in its efforts to link resources with results.

This administration has made the integration of performance and budget information one of five governmentwide management priorities under its President's Management Agenda (PMA).⁴ A central element in this initiative is the Office of Management and Budget's (OMB) Program Assessment Rating Tool (PART) that OMB describes as a diagnostic tool meant to provide a consistent approach to evaluating federal programs as part of the

² U.S. General Accounting Office, *Performance Budgeting: Past Initiatives Offer Insights for GPRA Implementation*, GAO/AIMD-97-46 (Washington, D.C.: Mar. 27, 1997).

³ See Pub. L. No. 103-62, § 2 (1993), 5 U.S.C. § 306 (2003), and 31 U.S.C. §§ 1115-1116 (2003).

⁴ 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.

executive budget formulation process. The PART is the latest iteration of 50 years of federal performance budgeting initiatives. It applies 25 questions to all “programs”⁵ under four broad topics: (1) program purpose and design, (2) strategic planning, (3) program management, and (4) program results (i.e., whether a program is meeting its long-term and annual goals) as well as additional questions that are specific to one of seven mechanisms or approaches used to deliver the program.⁶

To better understand the PART’s potential as a mechanism for assessing program goals and results, 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.

To respond to your request, we reviewed OMB materials on the development and implementation of the PART as well as the results produced by the PART assessments. To assess consistency of the PART’s application, we performed analyses of OMB data from the PART program summary and assessment worksheets for each of the 234 programs OMB reviewed for fiscal year 2004, including a statistical analysis of the relationship between the PART scores and funding levels in the President’s Budget. We also identified several sets of similar programs that we examined more closely to determine if comparable or disparate criteria were applied in producing the PART results for these clusters of programs. We reviewed 28 programs in nine clusters covering food safety, water supply, military equipment procurement, provision of health care, statistical agencies, block grants to assist vulnerable populations, energy research programs, wildland fire management, and disability compensation. We also interviewed OMB officials regarding their experiences with the PART in the fiscal year 2004 budget process.

⁵ There is no standard definition for the term “program.” For purposes of PART, OMB described the unit of analysis (program) as (1) an activity or set of activities clearly recognized as a program by the public, OMB, and/or Congress; (2) having a discrete level of funding clearly associated with it; and (3) corresponding to the level at which budget decisions are made.

⁶ The seven major categories are competitive grants, block/formula grants, capital assets and service acquisition programs, credit programs, regulatory-based programs, direct federal programs, and research and development programs. Tax programs were not addressed for the fiscal year 2004 PART process.

As part of our examination of the usefulness of the PART as an evaluation tool and also to obtain agency perspectives on the relationship between PART and GPRA, 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), and (3) the Department of the Interior (DOI). We selected these three departments because they had a variety of program types (e.g., block/formula grants, competitive grants, direct federal, and research and development) that were subject to the PART and could provide a broad-based perspective on how the PART was applied to different programs. With the exception of our summary analyses of all 234 programs, the information obtained from OMB and agency officials and our review of selected programs is not generalizable to the PART process for all 234 programs. However, the consistency and frequency with which similar issues were raised by OMB and agency officials suggest that our review reliably captures several significant and salient aspects of the PART as a budget and evaluation tool.

Our review focused on the fiscal year 2004 PART process. We conducted our work from May 2003 through October 2003 in accordance with generally accepted government auditing standards. Detailed information on our scope and methodology appears in appendix I. OMB provided written comments on a draft of this report that are reprinted in appendix IV.

Results in Brief

The PART has helped to structure and discipline OMB's use of performance information for its internal program analysis and budget review, made the use of this information more transparent, and stimulated agency interest in budget and performance integration. Both OMB and agency staff noted that this helped ensure that OMB staff with varying levels of experience focused on the same issues, fostering a more disciplined approach to discussing program performance with agencies. Several agency officials also told us that the PART was a catalyst for bringing agency budget, planning, and program staff together since none could fully respond to the PART questionnaire alone.

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. Over 80 percent of the recommendations made for the 234 programs assessed for the fiscal year 2004 budget process were for improvements in program design, assessment, and program management; less than 20 percent were related to funding issues. As OMB and others

recognize, performance is not the only factor in funding decisions. Determining priorities—including funding priorities—is a function of competing values and interests. Although OMB generally proposed to increase funding for programs that received ratings of “effective” or “moderately effective” and proposed to cut funding for those programs that were rated “ineffective,” our review confirmed OMB’s statements that funding decisions were not applied mechanistically. That is, for some programs rated “effective” or “moderately effective” OMB recommended funding decreases, while for several programs judged to be “ineffective” OMB recommended additional funding in the President’s budget request with which to implement changes.

Much of the potential value of the PART lies in the related program recommendations and associated improvements, but realization of these benefits will require sustained attention to implementation and oversight in order to determine if the desired results are being achieved. Such attention and oversight takes time, and OMB needs to be cognizant of this as it considers the capacity and workload issues in the PART. Currently OMB plans to assess an additional 20 percent of all federal programs annually. Each year, the number of recommendations from previous years’ evaluations will grow—and a system for monitoring their implementation will become more critical. OMB encouraged its Resource Management Offices (RMO) to consider many factors in selecting programs for the fiscal year 2004 PART assessments, such as continuing presidential initiatives and programs up for reauthorization. While all programs would eventually be reviewed over the 5-year period, selecting related programs for review in a given year would enable decision makers to analyze the relative efficacy of similar programs in meeting common or similar outcomes. We recommend that OMB centrally monitor and report on agency implementation and progress on PART recommendations to provide a governmentwide picture of progress and a consolidated view of OMB’s workload in this area. In addition, to target scarce analytic resources and to focus decision makers’ attention on the most pressing policy issues, we recommend that OMB reconsider plans for 100 percent coverage of federal programs by targeting PART assessments based on such factors as the relative priorities, costs, and risks associated with related clusters of programs and activities. We further recommend that OMB select for review in the same year related or similar programs or activities to facilitate such comparisons and trade-offs.

Developing a credible evidence-based rating tool to provide bottom-line ratings for programs was a major impetus in developing the PART.

However, inherent challenges exist in assigning a single “rating” to programs that often have multiple purposes and goals. Despite the considerable time and effort OMB has devoted to promoting consistent application of the PART, the tool is a work in progress. Additional guidance and considerable revisions are needed to meet OMB’s goal of an objective, evidence-based assessment tool. In addition to difficulties with the tool itself—such as subjective terminology and a restrictive yes/no format—providing flexibility to assess multidimensional programs with multiple purposes and goals often implemented through multiple actors has led to a reliance on OMB staff judgments to apply general principles to specific cases. OMB staff were not fully consistent in interpreting the guidance for complex PART questions and in defining acceptable measures. In addition, the limited availability of credible evidence on program results also constrained OMB staff’s ability to use the PART to rate programs’ effectiveness. Almost 50 percent of the 234 programs assessed for fiscal year 2004 received a rating of “results not demonstrated” because OMB decided that program performance information, performance measures, or both were insufficient or inadequate. OMB, recognizing many of the limitations with the PART, modified the PART for fiscal year 2005 based on lessons learned during the fiscal year 2004 process, but issues remain. We therefore recommend that OMB continue to improve the PART guidance by (1) clarifying when output versus outcome measures are acceptable and (2) better defining an “independent, quality evaluation.” We further recommend that OMB both clarify its expectations regarding the nature, timing, and amount of evaluation information it wants from agencies for the purposes of the PART and consider using internal agency evaluations as evidence on a case-by-case basis.

The PART is not well integrated with GPRA—the current statutory framework for strategic planning and reporting. According to OMB officials, GPRA plans were organized at too high a level to be meaningful for program-level budget decision making. To provide decision makers with program-specific, outcome-based performance data useful for executive budget formulation, OMB has stated its intention to modify GPRA goals and measures with those developed under the PART. As a result, OMB’s judgment about appropriate goals and measures is substituted for GPRA judgments based on a community of stakeholder interests. Agency officials

we spoke with expressed confusion about the relationship between GPRA requirements and the PART process. Many view PART's program-by-program focus and the substitution of program measures as detrimental to their GPRA planning and reporting processes. OMB's effort to influence program goals is further evident in recent OMB Circular A-11 guidance⁷ that clearly requires each agency to submit a performance budget for fiscal year 2005, which will replace the annual GPRA performance plan.

The tension between PART and GPRA was further highlighted by the challenges in defining a unit of analysis that is useful both for program-level budget analysis and agency planning purposes. Although the PART reviews indicated to OMB that GPRA measures are often not sufficient to help it make judgments about programs, the different units of analysis used in these two performance initiatives contributed to this outcome. For the PART, OMB created units of analysis that tied to discrete funding levels by both disaggregating and aggregating certain programs. In some cases, disaggregating programs for the PART reviews ignored the interdependency of programs by artificially isolating them from the larger contexts in which they operate. Conversely, in other cases in which OMB aggregated programs with diverse missions and outcomes for the PART reviews, it became difficult to settle on a single measure (or set of measures) that accurately captured the multiple missions of these diverse components. Both of these "unit of analysis" issues contributed to the lack of available planning and performance information.

Although the PART can stimulate discussion on program-specific performance measurement issues, it is not a substitute for GPRA's strategic, longer-term focus on thematic goals and department- and governmentwide crosscutting comparisons. GPRA is a broad legislative framework that was designed to be consultative with Congress and other stakeholders and allows for varying uses of performance information, while the PART applies evaluation information to support decisions and program reviews during the executive budget formulation process. Moreover, GPRA can anchor the review of programs by providing an overall strategic context for programs' contributions toward agency goals. We therefore recommend that OMB seek to achieve the greatest benefit from both GPRA and PART by articulating and implementing an integrated, complementary relationship between the two. We further recommend that OMB continue to improve the PART guidance by expanding the discussion

⁷ OMB Circular A-11, *Preparation, Submission, and Execution of the Budget*, Section 220.

of how programs—also known as “units of analysis”—are determined, including recognizing the trade-offs, implications, or both of such determinations.

As part of the President’s budget preparation, the PART clearly must serve the President’s interests. However, experience suggests that efforts to integrate budget and performance are promoted when Congress and other key stakeholders have confidence in the credibility of the analysis and the process used. It is unlikely that the broad range of players whose input is critical to decisions will use performance information unless they believe it is relevant, credible, reliable, and reflective of a consensus about performance goals among a community of interested parties. Similarly, the measures used to demonstrate progress toward a goal, no matter how worthwhile, cannot appear to serve a single set of interests without potentially discouraging use of this information by others. We therefore recommend that OMB attempt to build on the strengths of GPRA and PART by seeking to communicate early in the PART process with congressional appropriators and authorizers about what performance issues and information are most important to them in evaluating programs. Furthermore, while Congress has a number of opportunities to provide its perspective on performance issues and goals through its authorization, oversight, and appropriations processes, we suggest that Congress consider the need for a more structured approach for sharing with the executive branch its perspective on governmentwide performance matters, including its views on performance goals and outcomes for key programs and the oversight agenda.

In commenting on a draft of this report, OMB generally agreed with our findings, conclusions, and recommendations. OMB outlined actions it is taking to address many of our recommendations, including refining the process for monitoring agencies’ progress in implementing the PART recommendations, seeking opportunities for dialogue with Congress on agencies’ performance, and continuing to improve executive branch implementation of GPRA plans and reports. 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 provided to the Departments of the Interior, Energy, and Health and Human Services, which are incorporated as appropriate.

Background

The current administration has taken several steps to strengthen and further performance-resource linkages for which GPRA laid the groundwork. Central to the budget and performance integration initiative, the PART is meant to strengthen the process for assessing the effectiveness of programs by making that process more robust, transparent, and systematic. As noted above, the PART is a series of diagnostic questions designed to provide a consistent approach to rating federal programs. (See app. II for a reproduction of the PART.) Drawing on available performance and evaluation information, the questionnaire attempts to determine 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 and performance plans and reports, financial statements, inspector general and GAO reports, and independent program evaluations. 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.

Table 1: Overview of Sections of PART Questions

Section	Description	Weight
I. Program Purpose and Design	To assess whether <ul style="list-style-type: none">• the purpose is clear, and• the program design makes sense.	20%
II. Strategic Planning	To assess whether the agency sets valid programmatic <ul style="list-style-type: none">• annual goals, and• long-term goals.	10%
III. Program Management	To rate agency management of the program, including <ul style="list-style-type: none">• financial oversight, and• program improvement efforts.	20%
IV. Program Results/Accountability	To rate program performance on goals reviewed in <ul style="list-style-type: none">• the strategic planning section, and• through other evaluations.	50%

Source: GAO analysis of the *Budget of the United States Government, Fiscal Year 2004, Performance and Management Assessments* (Washington, D.C.: February 2003).

In addition, each PART program is assessed according to one of seven approaches to service delivery. Table 2 provides an overview of these program types and the number and percentage of programs covered by each type in the fiscal year 2004 President's Budget performance assessments.

Table 2: Overview of PART Program Types

Program type	Description	Number/percentage of programs ^a
1. Direct federal	Programs in which support and services are provided primarily by federal employees.	67 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%
3. Competitive grant	Programs that distribute funds to state, local, and tribal governments, organizations, individuals, and other entities through a competitive process.	37 16%
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%
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%
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%
7. Credit	Programs that provide support through loans, loan guarantees, and direct credit.	4 2%
8. Mixed ^b	Programs that contain elements of different program types.	4 2%

Source: GAO summary and analysis of the *Budget of the United States Government, Fiscal Year 2004, Performance and Management Assessments* (Washington, D.C.: February 2003).

^aPercentages do not add to 100 percent due to rounding.

^bOMB noted that in rare cases, drawing questions from two of the seven PART program types—that is, creation of a "mixed" program type—yields a more informative assessment.

During the fiscal year 2004 budget cycle, OMB applied the PART to 234 programs (about 20 percent of the fiscal year 2004 President's Budget request to Congress⁸), and gave each program 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. The administration plans to assess an additional 20 percent of the budget each year until the entire executive branch has been reviewed. For more information on the development of the PART, see appendix III.

OMB Used the PART to Systematically Assess Program Performance and Make Results Known, but Follow-up on PART Recommendations Is Uncertain

The PART clarified OMB's use of performance information in its budget decision-making process and stimulated new interest in budget and performance integration. OMB generally proposed budget increases for programs that received ratings of "effective" or "moderately effective" and decreased funding requests for those programs that were rated "ineffective," but there were clear exceptions. Moreover, the more important role of the PART was not in making resource decisions but in its support for recommendations to improve program design, assessment, and management. OMB's ability to use the PART to identify and address future program improvements and measure progress—a major purpose of the PART—is predicated on its ability to oversee the implementation of PART recommendations. However, it is not clear that OMB has a centralized system to oversee the implementation of such recommendations or evaluate their effectiveness.

The PART Made Budget and Performance Integration at OMB More Transparent

The PART helped structure and discipline the use of performance information in the budget process and made the use of such information more transparent throughout the executive branch. According to OMB senior officials and many of the examiners and branch chiefs, the PART lent structure to a process that had previously been informal and gave OMB staff a systematic way of asking performance-related questions. Both

⁸ OMB defined 20 percent of the budget as either 20 percent of programs or their funding levels so long as all programs are assessed over the 5-year cycle for fiscal years 2004 through 2008 budget requests.

agency and OMB staff noted that this helped ensure that OMB staff with varying levels of experience focused on the same issues, fostering a more disciplined approach to discussing performance within OMB and with agencies. Agency officials told us that by encouraging more communication between departments and OMB, the PART helps illuminate both how OMB makes budget decisions and how OMB staff think about program management. The PART also provided a framework for raising performance issues during the OMB Director's Reviews. OMB managers and staff reported that it led to richer discussions on what a program should be achieving, whether the program was performing effectively, and how program performance could be improved.

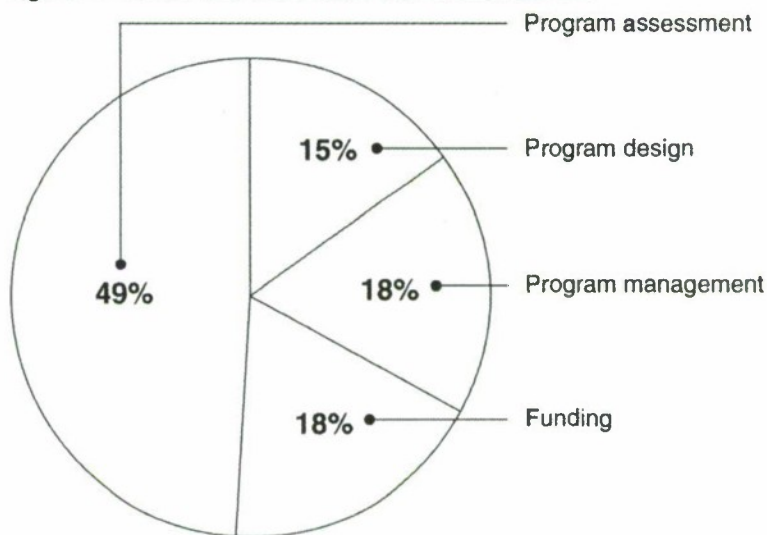
Agencies also reported that the PART process expanded the dialogue between program, planning, and budget staffs, and stimulated interest in budget and performance integration. Several agency officials stated that the PART worksheets were a catalyst for bringing staffs together since none could fully respond to the questionnaire alone. OMB and agency officials agreed that the PART led to more interactions between OMB and agency program and planning staff and, in turn, increased program managers' awareness of and involvement in the budget process. According to OMB and several agency officials, the PART process—that is, responding to the PART questionnaire—involved staff outside of the performance management area. Additionally, both agency and OMB officials said that the attention given to programs that were not routinely reviewed was a positive benefit of the PART process.

Use of Performance Information Was Evident in OMB's Recommendations

OMB senior officials told us that one of the PART's most notable impacts was its ability to highlight OMB's recommended changes in program management and design. As shown in figure 1, we found that 82 percent of PART recommendations addressed program assessment, design, and management issues; only 18 percent of the recommendations had a direct link to funding matters.⁹

⁹ The 234 programs assessed for fiscal year 2004 contained a total of 612 recommendations.

Figure 1: Fiscal Year 2004 PART Recommendations



Source: GAO analysis of OMB data.

The majority of recommendations relate to changes that go well beyond funding consideration for one budget cycle. For example, OMB and HHS officials agree that the Foster Care program as it is currently designed does not provide appropriate incentives for the permanent placement of children; the program financially rewards states for keeping children in foster care instead of the original intent of providing temporary, safe, and appropriate homes for abused or neglected children until children can be returned to their families or other permanent arrangements can be made. The PART assessment provided support for OMB's recommendation that legislation be introduced that would create an option for states to participate in an alternate financing program that would "better meet the needs of each state's foster care population."

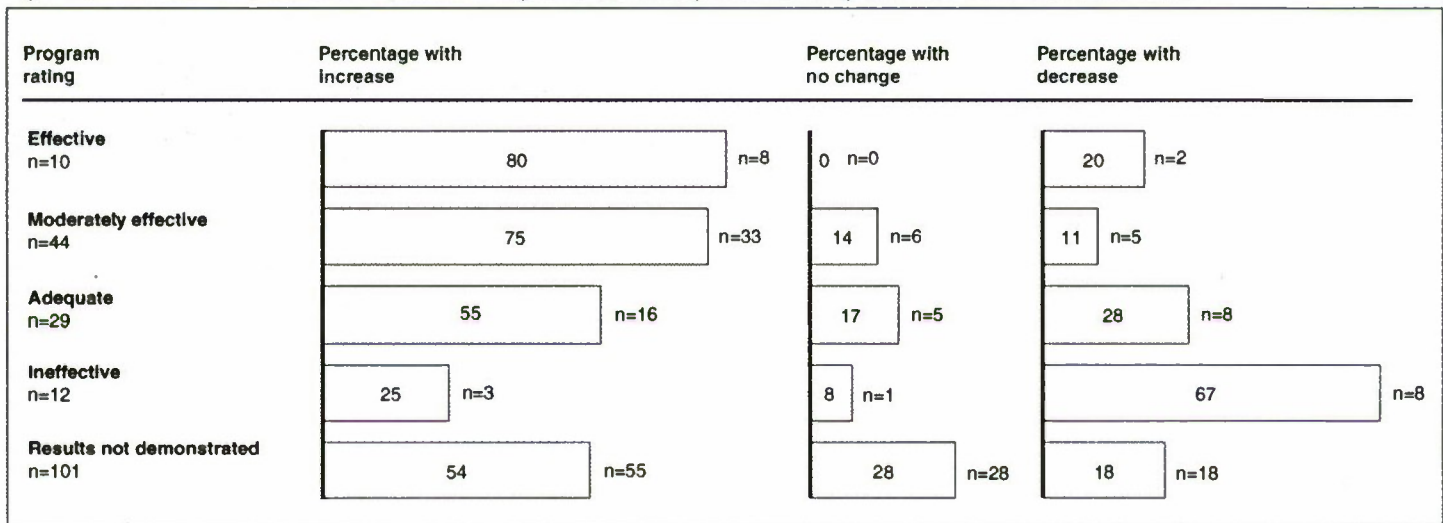
Performance information included in the PART for the Department of Labor's (DOL) Community Service Employment for Older Americans program helped to shape OMB's recommendation to increase competition for the grants. OMB concluded that although the Older Americans Act of 2000 amendments¹⁰ authorize competition for grants in cases in which grantees repeatedly fail to perform, the programs' 10 national grantees

¹⁰ Pub. L. No. 106-501 (2000).

have historically been the sole recipients of grant funds regardless of performance. OMB recommended that DOL award national grants competitively to strengthen service delivery and open the door to new grantees.

As OMB and others recognize, performance is not the only factor in funding decisions. Determining priorities—including funding priorities—is a function of competing values and interests. As seen in figure 2, we found that PART scores were generally positively related to proposed funding changes in discretionary programs but not in a mechanistic way. In other words, PART scores did not automatically determine funding changes. OMB proposed funding increases for most of the programs rated “effective” or “moderately effective” and proposed funding decreases for most of the programs rated “ineffective,” but there were clear exceptions. Programs rated as “results not demonstrated”—which reflected a range of PART scores—had mixed results.

Figure 2: Number of Discretionary PART Programs by Rating and Funding Result, Fiscal Years 2003-2004



Source: GAO analysis of OMB data.

Note: Discretionary programs refer to those programs with budgetary resources provided in appropriation acts. Because Congress controls spending for mandatory programs—generally entitlement programs such as food stamps, Medicare, and veterans’ pensions—indirectly rather than directly through the appropriations process, we excluded them from our analysis. Of the 234 programs, we could not classify 11 as being either predominantly mandatory or discretionary; these programs are excluded from our analysis as well, and are listed in appendix I.

A large portion of the variability in proposed budget changes could not be explained by the quantitative measures reported by the PART. Regressions of PART scores never explained more than about 15 percent of the proposed budget changes. For only the one-third of discretionary programs with the smallest budgets, we found that the composite PART scores had a modest but statistically significant effect on proposed budget changes (measured in percentage change) between fiscal years 2003 and 2004. For a fuller discussion of the statistical methods used, see appendix I.

The relationship between performance levels and budget decisions was not one-dimensional. For example, OMB rated the Department of Defense's Basic Research program as "effective," but recommended a reduction in congressionally earmarked projects that it stated did not meet the program's merit review process. OMB also recommended reducing funding for DOE's International Nuclear Materials Protection and Cooperation program (rated "effective") because difficulties in obtaining international agreements had resulted in the availability of sufficient unobligated balances¹¹ to make new funding unnecessary. However, OMB sometimes proposed funding increases for programs that were rated "ineffective" to implement improvement plans that had been developed, such as the Internal Revenue Service's new Earned Income Tax Credit compliance initiatives and DOE's revised environmental cleanup plans for its Environmental Management (Cleanup) program.

Capacity Issues Could Affect OMB's Ability to Use the PART to Drive Program Improvements

OMB has said that a major purpose of the PART is to focus on program improvements and measure progress. Effectively implementing PART recommendations aimed at program improvements will require sustained attention and sufficient oversight of agencies to ensure that the recommendations are producing desirable results. However, each year, the number of recommendations from previous years' evaluations will grow. Currently, OMB plans to assess an additional 20 percent of all federal programs annually such that all programs would eventually be reviewed over a 5-year period. OMB encouraged its RMOs to consider a variety of factors in selecting programs for the fiscal year 2004 PART assessments, including continuing presidential initiatives and programs up for reauthorization. Strengthening the focus on selecting related programs for review in a given year would enable decision makers to analyze the relative

¹¹ Unobligated balances are defined as portions of available budget authority that the agency has not set aside to cover current legal liabilities.

efficacy of similar programs in meeting common or similar outcomes. As our work has shown, unfocused and uncoordinated programs waste scarce funds, confuse and frustrate program customers, and limit overall program effectiveness. Therefore it is prudent to highlight crosscutting program efforts and clearly relate and address the contributions of alternative federal strategies toward meeting similar goals.

Although OMB has created a template for agencies to report on the status of their recommendations and has reported that agencies are implementing their PART recommendations, OMB has no central system for monitoring agency progress or evaluating the effectiveness of changes. While RMOs are responsible for overseeing agency progress, OMB senior managers will not have a comprehensive governmentwide picture of progress on the implementation of PART recommendations, nor will they have a complete picture of OMB's workload in this area. As OMB has recognized, following through on the recommendations is essential for improving program performance and ensuring accountability.

Senior OMB managers readily recognized the increased workload the PART placed on examiners—in one public forum we attended, a senior OMB official described many examiners as being very concerned about the additional workload. However, OMB expects the workload to decline as OMB and agency staff become more familiar with the PART tool and process, and as issues with the timing of the PART reviews are resolved. Agency officials told us that originally, there was no formal guidance for reassessing PART programs—it varied by RMO. When issued, OMB's formal PART guidance limited reassessments to (1) updating the status/implementation of recommendations from the fiscal year 2004 PART and (2) revisiting specific questions for which new evidence exists. OMB expected that in most reassessments, only those questions in which change could be demonstrated would be "reopened." OMB officials acknowledged that this formal guidance is at least partly due to resource constraints.

OMB staff were divided on whether the PART assessments made an appreciable difference in time spent on its budget review process. Many of those we spoke with told us that their workloads during the traditional budget season have always been heavy and that PART did not add significantly to their work, especially since the PART generally formalized a process already taking place. Those who did acknowledge workload concerns said that they were surprised at the amount of time it was taking to reassess programs. In fact, more than one OMB official told us that

reassessing programs was taking almost as long as brand-new assessments, despite the fact that OMB scaled back the scope of these reassessments.

Despite OMB's Considerable Efforts to Create a Credible Evaluation Tool, PART Assessments Require Judgment and Were Constrained by Data Limitations

OMB went to great lengths to encourage consistent application of the PART in the evaluation of government programs, including pilot testing the instrument, issuing detailed guidance, and conducting consistency reviews. However, while the instrument can undoubtedly be improved, any tool that is sophisticated enough to take into account the complexity of the U.S. government will always require OMB staff to exercise interpretation and judgment. Providing flexibility to assess multidimensional programs with multiple purposes and impacts has led to a reliance on OMB staff judgments to apply general principles to specific cases. Accordingly, OMB staff were not fully consistent in interpreting complex questions about agency goals and results. In addition, the limited availability of credible evidence on program results also constrained OMB's ability to use the PART to rate programs' effectiveness.

Inherent Performance Measurement Challenges Make It Difficult to Meaningfully Interpret a Bottom-Line Rating

OMB published a single, bottom-line rating for the PART results as well as individual section scores, which are potentially more useful for identifying information gaps and program weaknesses. For example, one program that was rated "adequate" overall got high scores for purpose (80 percent) and planning (100 percent), but did poorly in being able to show results (39 percent) and in program management (46 percent). Thus, the individual section ratings provided a better understanding of areas needing improvement than the overall rating alone. Bottom-line ratings inevitably force choices on what best exemplifies a program's mission—even when a program has multiple goals—and encourages a determination of the effectiveness of the program even when performance data are unavailable, the quality of those data is uneven, or they convey a mixed message on performance.

Many of the outcomes for which federal programs are responsible are part of a broader effort involving federal, state, local, nonprofit, and private partners. We have previously reported that it is often difficult to isolate a particular program's contribution to an outcome and especially so when it involves third parties.¹² This was reinforced by the results of the fiscal year 2004 PART reviews. One of the patterns that OMB identified in its ratings was that grant programs received lower than average ratings. To OMB this suggested the need for greater effort by agencies to make grantees accountable for achieving overall program results. However, grant structure and design play a role in how federal agencies are able to hold third parties responsible and complicate the process of identifying the individual contributions of a federal program with multiple partners. In particular, block grants present implementation challenges, especially in those instances in which national goals are not compatible with state and local priorities.

OMB Employed Numerous Tools and Techniques to Promote and Improve Consistent Application of the PART

OMB went to great lengths to encourage consistent application of the PART in the evaluation of government programs. These efforts included (1) testing the PART in selected agencies before use in the fiscal year 2004 assessment, (2) issuing detailed guidance and worksheets for use by PART teams, (3) making the Performance Evaluation Team (PET) available to answer PART implementation questions, (4) establishing an Interagency Review Panel (IRP) to review consistency of PART evaluations, and (5) making improvements to the fiscal year 2005 process and guidance based upon the fiscal year 2004 experience.

OMB conducted a pilot test of the PART and released a draft of the PART questionnaire for public comment prior to its use for the fiscal year 2004 budget cycle. During Spring Review in 2002, OMB and agency staff piloted the draft PART on 67 programs. The PART was also shared with and commented on by the Performance Measurement Advisory Council and other external groups. According to OMB, the results of the Spring Review and feedback from external groups were used to revise the draft version of the PART to lessen subjectivity and increase the consistency of reviews.

¹² See GAO-03-595T and U.S. General Accounting Office, *Managing for Results: Efforts to Strengthen the Link Between Resources and Results at the Administration for Children and Families*, GAO-03-9 (Washington, D.C.: Dec. 10, 2002).

OMB issued detailed guidance to help OMB and agency staff consistently apply the PART and created electronic “templates” or worksheets to aid in completing PART assessments. This guidance explains the purpose of each question and describes the evidence required to support a “yes” or “no” answer. In order to account for different types of programs, several questions tailored to the seven program types were added to the PART (primarily in Section III—Program Management). While the PART guidance cannot be expected to cover every situation, the instructions established general standards for PART evaluations.

PET addressed in “real time” questions and issues that OMB staff that were completing the PART evaluations repeatedly raised. PET consisted of examiners drawn from across the OMB organization representing a variety of programmatic knowledge and experiences. It served as a sounding board for OMB staff and a source for sharing experiences, issues, and useful approaches and also provided training to OMB and agency staff on the process. For example, in one OMB branch, staff were grappling with how to apply the PART to a set of block grants. They went through the instrument with the PET member from their RMO and continued to consult with that individual throughout the process.

OMB also formed IRP, which consisted of both OMB and agency officials, to conduct a consistency check of the PART reviews and to review formal appeals of the process or results for particular questions. During the fiscal year 2004 budget process, IRP conducted a consistency review of 10 percent of the PART evaluations using a subset of the PART questions that OMB staff identified as being the most subjective or difficult to interpret. IRP also reviewed formal agency appeals to determine whether there was consistent treatment of similar situations.

As an Evaluation Tool, the PART Has Weaknesses in Its Design and, as a Result, Its Implementation

Despite the considerable time and effort OMB has devoted to promoting consistent application of the PART, difficulties both with the tool itself (such as subjective terminology and a restrictive yes/no format) and with implementing the tool (including inconsistencies in defining acceptable measures and contradictory answers to “pairs” of related questions) aggravated the general performance measurement challenges described earlier.

Subjective Terms and a Restrictive Format Contributed to Subjective and Inconsistent Responses

Many PART questions contain subjective terms that are open to interpretation. Examples include terminology such as “ambitious” in describing sought-after performance measures. Because the appropriateness of a performance measure depends on the program’s purpose, and because program purposes can vary immensely, an ambitious goal for one program might be unrealistic for a similar but more narrowly defined program. Some agency officials claimed that having multiple statutory goals disadvantaged their programs. Without further guidance, subjective terminology can influence program ratings by permitting OMB staff’s views about a program’s purpose to affect assessments of the program’s design and achievements.

Although OMB employed a yes/no format for the PART because OMB believes it aided standardization, the format resulted in oversimplified answers to some questions. OMB received comments on the yes/no format in conducting the PART pilot. Some parties liked the certainty and forced choice of yes/no. Others felt the format did not adequately distinguish between the performance of various programs, especially in the results section (originally in the yes/no format). In response to these concerns, OMB revised the PART in the spring of 2002 to include four response choices in the results section (adding “small extent” and “large extent” to the original two choices “yes” and “no”), while retaining the dichotomous yes/no format in the other three sections. OMB acknowledged that a “yes” response should be definite and reflect a very high standard of performance, and that it would more likely be difficult to justify a “yes” answer than a “no” answer. Nonetheless, agency officials have commented that the yes/no format is a crude reflection of reality, in which progress in planning, management, or results is more likely to resemble a continuum than an on/off switch.

Moreover, the yes/no format was particularly troublesome for questions containing multiple criteria for a “yes” answer. As discussed previously, we conducted an in-depth analysis of PART assessments for 28 related programs in nine clusters and compared the responses to related questions. That analysis showed six instances in which some OMB staff gave a “yes” answer for successfully achieving some but not all of the multiple criteria, while others gave a “no” answer when presented with a similar situation. For example, Section II, Question 1, asks, “Does the program have a limited number of specific, ambitious, long-term performance goals that focus on outcomes and meaningfully reflect the purpose of the program?” The PART defines successful long-term goals by multiple, distinct characteristics

(program has long-term goals, time frames by which the goals are to be achieved, etc.), but does not clarify whether a program can receive a “yes” if *each* of the characteristics is met, or if *most* of the characteristics are met. This contributed to a number of inconsistencies across program reviews. For example, OMB judged DOI’s Water Reuse and Recycling program “no” on this question, noting that although DOI set a long-term goal of 500,000 acre-feet per year of reclaimed water, it failed to establish a time frame for when it would reach the target. However, OMB judged the Department of Agriculture’s and DOI’s Wildland Fire programs “yes” on this question even though the programs’ long-term goals of improved conditions in high-priority forest acres are not accompanied by specific time frames. In another example, OMB accepted DOD’s recently established long-term strategic goals for medical training and provision of health care even though it did not yet have measures or targets for those goals. By breaking out targets and ambitious time frames separately from the question of annual goals, agencies have an opportunity to get credit for progress made.

There Were Inconsistencies in Defining Acceptable Measures and in Logically Responding to Question “Pairs”

In particular, our analysis of the nine program clusters revealed three instances in which OMB staff inconsistently defined appropriate measures—outcome versus output—for programs. Officials also told us that OMB staff used different standards to define measures as outcome oriented. This may reflect, in part, the complexity of and relationship between expected program benefits. Outcomes are generally defined as the results of outputs—products and services—delivered by a program. But in some programs, long-term outcomes are expected to occur over time through multiple steps. In these cases, short-term outcomes—immediate changes in knowledge and awareness—might be expected to lead to intermediate outcomes—behavioral changes in the future—and eventually result in long-term outcomes—benefits to the public.

In the employment and training area, OMB accepted short-term outcomes, such as obtaining high school diplomas or employment, as a proxy for long-term goals for the HHS Refugee Assistance program, which aims to help refugees attain economic self-sufficiency as soon as possible after they arrive. However, OMB did not accept the same employment rate measure as a proxy for long-term goals for the Department of Education’s Vocational Rehabilitation program because it had not set long-term targets beyond a couple of years. In other words, although neither program contained long-term outcomes, such as participants gaining economic self-sufficiency, OMB accepted short-term outcomes in one instance but not the other.

Similarly, OMB gave credit for output measures of claims processing (time, accuracy, and productivity) as a proxy for long-term goals for the Social Security Administration's Disability Insurance program, but did not accept the same output measures for the Veterans Disability Compensation program. OMB took steps to address this issue for fiscal year 2005.

We also found that three "question pairs" on the PART worksheets are linked, yet in two of the three "pairs," a disconnect appeared in how OMB staff responded to these questions for a given program.¹³ For example, 29 of the 90 programs (32 percent) judged as lacking "independent and quality evaluations of sufficient scope conducted on a regular basis" (Section II, Question 5) were also judged as having "independent and quality evaluations that indicated the program is effective and achieving results" (Section IV, Question 5). There is a logical inconsistency in these two responses. In another instance, there was no linkage between the questions that examine whether a program *has* annual goals that demonstrate progress toward achieving long-term goals and whether the program actually *achieves* its annual goals. For example, 15 of the 75 programs (20 percent) judged *not* to have adequate annual performance goals (Section II, Question 2) were nevertheless credited for *having made progress* on their annual performance goals (Section IV, Question 2). However, the guidance for the latter question clearly indicates that a program must receive a "no" if it received a "no" on the existence of annual goals (Section II, Question 2). It seems that some raters held programs to a higher standard for the quality of goals than for progress on them.

¹³ In the third question pair, a question in the planning section asks about whether the program has long-term goals, and a question in the results section asks whether the agency has made progress in achieving the program's long-term goals. Yet, in 6 of the 115 programs (5 percent) judged not to have adequate long-term goals, credit was given for making progress on their long-term goals even though the guidance again clearly states that a program must receive a "no" if the program received a "no" on the existence of long-term outcome goals.

The Lack of Performance Information Creates Challenges in Effectively Measuring Program Performance

According to OMB, 115 out of 234 programs (49 percent) lacked “specific, ambitious, long-term performance goals that focus on outcomes” (Section II, Question 1). In addition, OMB found that 90 out of 234 programs (38 percent) lacked sufficient “independent, quality evaluations” (Section II, Question 5). While the validity of these assessments may be subject to interpretation and debate, our previous work¹⁴ has raised concerns about the capacity of federal agencies to produce evaluations of program effectiveness.

The lack of evaluations may in part be driven by how OMB defined an “independent and quality evaluation.” To be independent, nonbiased parties with no conflict of interest would conduct the evaluation, but agency officials felt that OMB staff started from the default position that agency-sponsored evaluations are, by definition, biased. However, our detailed review of 28 PART worksheets found only 7 instances in which OMB explicitly noted its rejection of evaluations: 1 for being too old, 3 for not being independent (of the 3, 1 was an internal agency review and 2 were conducted by industry groups), and the remaining 3 for not assessing program results. OMB officials have acknowledged that this issue was a point of friction with agencies and that beyond GAO, inspectors general, and other government reports that were automatically presumed to be independent, the independence standard was considered on a case-by-case basis. In these case-by-case situations, OMB staff told us that they looked for some degree of detachment and objectivity in the evaluations. For example, in the case of one DOE-sponsored evaluation, the OMB examiner attended the meetings of the review group that conducted the evaluation in order to see firsthand what sorts of questions the committee posed to the department officials. In OMB’s estimation, there was clear independence. While OMB changed the fiscal year 2005 guidance to recognize evaluations contracted out to third parties and agency program evaluation offices as possibly being sufficiently independent, the new guidance generally prohibits evaluations conducted by the program itself from being considered “independent.”

Other reasons evaluation data may be limited include (1) constraints on federal agencies’ ability to influence program outcomes and reliance on states and others for data for programs for which responsibility has

¹⁴ U.S. General Accounting Office, *Program Evaluation: Agencies Challenged by New Demand for Information on Program Results*, GAO/GGD-98-53 (Washington, D.C.: Apr. 24, 1998).

devolved to the states and (2) the lack of a statutory mandate or dedicated funds for evaluation, which agency officials told us can hamper efforts to conduct studies or to improve administrative data collection.

As we have previously noted, program evaluations can take many forms and agencies may obtain evaluations in a variety of ways.¹⁵ Some evaluations simply analyze routinely collected program administrative data; others involve special surveys. The type of evaluation can greatly affect evaluation cost. Net impact evaluations compare outcomes for program participants to those of a randomly assigned control group and are designed for situations in which external factors are also known to influence those outcomes. However, the adequacy of an evaluation design can only be determined relative to the circumstances of the program being evaluated. In addition, agencies can obtain evaluations by having program or other agency staff collect and analyze the data, by conducting the work jointly with program partners (such as state agencies), or by hiring contract firms to do so. Our survey of 81 federal agency offices conducting evaluations in 1995 of program results found they were most commonly located in administrative offices at a major subdivision level or in program offices (43 and 30 percent, respectively). Overall, they reported conducting 51 percent of their studies in-house, while 34 percent were contracted out. Depending on the sensitivity of the study questions, agencies can conduct credible internal evaluations by adopting procedures to ensure the reliability and validity of data collection and analysis.

¹⁵ GAO/GGD-98-53.

Disagreements on Performance Information Led to Creation of a “Results Not Demonstrated” Category

During the PART process OMB created an additional rating category, “results not demonstrated,” which was applied to programs regardless of their score if OMB decided that one or both of two conditions pertained: (1) OMB and the agency could not reach agreement on long-term and annual performance measures and (2) there was inadequate performance information. Almost 50 percent of the 234 programs assessed for fiscal year 2004 received this rating of “results not demonstrated,” ranging from high-scoring programs such as the Consumer Product Safety Commission (83) to low-scoring programs such as the Department of Veterans Affairs Disability Compensation program (15). OMB officials said that this rating was given to programs when agreement could not be reached on long-term and annual performance measures and was applied regardless of the program’s PART score. Our own review found that OMB generally assigned the “results not demonstrated” rating as described above.¹⁶

It is important for users of the PART information to interpret the “results not demonstrated” designation as “unknown effectiveness” rather than as meaning the program is “ineffective.” Having evidence of poor results is not the same as lacking evidence of effectiveness. Because the PART guidance sets very high standards for obtaining a “yes,” a “no” answer can mean either that a program did not meet the standards, or that there is no evidence on whether it met the standards. In some readily measured areas, lack of evidence of an action may indicate that the standard probably was not met. However, because effectiveness is often not readily observed, lack of evidence on program effectiveness cannot be automatically interpreted as meaning that a program is ineffective. Furthermore, an agency might have results for goals established under GPRA, but if OMB and the agency could not reach agreement on new or revised goals or measures, then OMB gave a program the rating “results not demonstrated.”

¹⁶ However, we found 8 cases (out of 118) programs that were rated as “results not demonstrated” despite having both annual and long-term performance goals and evidence that these goals were being met.

Changes to the PART and Related Guidance for Fiscal Year 2005 Are Meant to Address Previously Identified Problems

OMB, recognizing many of the issues we have just discussed, made modifications to the PART instrument and guidance in time for the fiscal year 2005 process. OMB said these changes were based upon lessons learned during the fiscal year 2004 process and input from a variety of sources, such as PET, IRP, and agency officials, although we were unable to determine which changes resulted from which recommendations. Although the PART as used for fiscal year 2005 is very similar to that for fiscal year 2004, several questions were added, dropped, merged with other questions, or divided into two questions. For example, a research and development question used in the fiscal year 2004 PART that received “not applicable” answers in 13 out of the 32 cases in which it was applied was dropped from the fiscal year 2005 PART. According to OMB officials, several of the multicriteria questions were split into separate questions in order to reduce inconsistency, as described earlier in this report. Appendix II provides more complete information on the guidance changes between fiscal years 2004 and 2005. To complement the fiscal year 2005 PART guidance and offer strategies for addressing common performance measurement challenges, many of which were encountered during the fiscal year 2004 process, OMB released a separate document, titled *Performance Measurement Challenges and Strategies*, which was the result of a workshop in which agencies participated and identified measurement challenges and shared best practices and possible work-arounds.

Instead of reestablishing IRP (which included both agency and OMB representatives) for the fiscal year 2005 process, OMB officials told us that PET (which included only OMB representatives) would conduct a consistency review of 25 percent of all PART evaluations, with at least one consistency check per OMB branch. OMB also told us that it has asked the National Academy of Public Administration (NAPA) to review PET’s consistency review for the fiscal year 2005 process; the scope and results of that review were not available to us during our audit work.¹⁷ OMB senior officials cited resources, timing, and the differing needs of the fiscal year 2004 and 2005 PART processes as reasons for dropping the IRP review. The absence of agency participation in this important phase of the PART could hamper ensuring crucial transparency and credibility.

¹⁷ Because our audit focused on the fiscal year 2004 PART process, our engagement was not limited by OMB’s decision to not share its reasoning for shifting the consistency review from IRP to PET or our lack of access to the NAPA review.

The Fiscal Year 2004 PART Process Was a Parallel, Competing Approach to GPRA's Performance Management Framework

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. However, GPRA—the current statutory framework for strategic planning and reporting—is a broader process involving the development of strategic and performance goals and objectives to be reported in strategic and annual plans. OMB's desire to collect performance data that better align with budget decision units means that the fiscal year 2004 PART process was a parallel competing structure to the GPRA framework. Although OMB acknowledges that GPRA was the starting point for the PART, as we explain below, the emphasis is shifting such that over time the performance measures developed for the PART and used in the budget process may come to drive agencies' strategic planning processes.

Agencies told us that in some cases, OMB is replacing PART goals and measures for those of GPRA. Effective for fiscal year 2005, OMB's Circular A-11 guidance states that performance budgets are to replace GPRA's annual performance plans. Agencies see the change as detrimental to planning and reporting under GPRA and as a resource drain since they have to respond to both GPRA and PART requirements. Some agency officials told us that although the PART can stimulate discussion on program-specific performance measurement issues, it is not a substitute for GPRA's outcome-oriented, strategic look at thematic goals and departmentwide program comparisons. Moreover, while the PART does not eliminate the departmental strategic plans created under GPRA, many OMB and agency officials told us that the PART is being used to shape the strategic plans.

OMB's Efforts to Link Performance Information with the Budget Often Conflict with Agencies' GPRA Planning Efforts

OMB guidance and officials made clear that GPRA goals, measures, and reports needed to be modified to provide decision makers with program-specific, outcome-based performance data that better aligned with the budget presentation in the President's Budget. According to OMB, such changes were needed because performance reporting under GPRA had evolved into a process separate from budget decision making, with GPRA plans organized at too high a level to be meaningful for program-level budget analysis and management review. Furthermore, according to OMB officials, GPRA plans had too many performance measures, which made it difficult to determine an agency's priorities. However, as some officials pointed out, the cumulative effect of adding new PART measures to GPRA plans may actually increase the number of measures overall; both agency and OMB officials recognize that this is contrary to goals issued by an OMB official previously responsible for the PART, indicating his desire to reduce the number of GPRA measures by at least 25 percent in at least 70 percent of federal departments.¹⁸ As a result of these sometimes-conflicting perspectives, agency officials said that responding to both PART and GPRA requirements increased their workloads and was a drain on staff resources.

OMB's most recent Circular A-11 guidance clearly requires that each agency submit a performance budget for fiscal year 2005 and that this should replace the annual GPRA performance plan.¹⁹ These performance budgets are to include information from the PART assessments, where available, including all performance goals used in the assessment of program performance done under the PART process. Until all programs have been assessed using the PART, the performance budget will also include performance goals for agency programs that have not yet been assessed using the PART. OMB's movement from GPRA to PART is further evident in the fiscal year 2005 PART guidance stating that while existing GPRA performance goals may be a starting point during the development of PART performance goals, the GPRA goals in agency GPRA documents are to be revised significantly, as needed, to reflect OMB's instructions for developing the PART performance goals. Lastly, this same guidance states that GPRA plans should be revised to include any new performance measures used in the PART and unnecessary measures should be deleted from GPRA plans.

¹⁸ Memorandum to the President's Management Council, "Where We'd Be Proud To Be," May 21, 2003.

¹⁹ OMB Circular A-11, *Preparation, Submission, and Execution of the Budget*.

OMB's interest in developing more useful program goals is further evident in its PART recommendations. Almost half of the fiscal year 2004 PART recommendations related to performance assessment—developing outcome goals and measures; cost or efficiency measures; and increasing the tracking/monitoring of data, improving the tracking/monitoring of data, or both. GPRA was generally the starting point for PART discussions about goals and measures, and many agency officials told us that OMB used the PART to modify agencies' existing GPRA goals and measures. Agency officials reported that the discussions about goals and measures were one of the main areas of contention during the PART process. At the same time, agency officials acknowledged that (1) sometimes OMB staff accepted current GPRA measures and (2) sometimes the new PART measures and goals were improvements over the old GPRA measures—the PART measures were more aggressive, more outcome-oriented, more targeted, or all of the above.

Defining a “Unit of Analysis” That Is Useful for Program-Level Budget Analysis and Agency Planning Purposes Presents Challenges

The appropriate unit of analysis or “program” is not always obvious. What OMB determined was useful for a PART assessment did not necessarily match agency organization or planning elements. Although the units of analysis varied across the PART assessments, OMB's guidance stated that they should be linked to a recognized funding level in the budget. In some cases, OMB aggregated separate programs for the purposes of the PART, while in other cases it disaggregated programs. Aggregating programs to tie them to discrete funding levels sometimes made it difficult to create a limited, but comprehensive, set of measures for programs with multiple missions. Disaggregating programs sometimes ignored the interdependence of programs by artificially isolating programs from the larger contexts in which they operate. Both contributed to the lack of available planning and performance information. For example, aggregating rural water supply projects as a single unit of analysis may have been a logical choice for reviewing related activities, but it created problems in identifying planning and performance information useful for the PART since these projects are separately administered. In another case, HHS officials told us that the PART program Substance Abuse Treatment Programs of Regional and National Significance is an amalgamation of activities funded in a single budget line, not an actual program. They said it was a challenge to make these activities look as if they functioned as a single program.

Disaggregating a program too narrowly can create problems by distorting its relationship to other programs involved in achieving a common goal.

For example, agency officials described a homeless program in which outreach workers help homeless persons with emergency needs and refer them to other agencies for housing and needed services. They said that their OMB counterparts suggested that the program adopt long-term outcome measures indicating number of persons housed. Agency officials argued that chronically homeless people require many services and that this federal program often supports only some of the services needed at the initial stages of intervention. The federal program, therefore, could contribute to, but not be primarily responsible for, affecting late stages of the intervention process, like housing status.

These issues reveal some of the unresolved tensions between the President's budget and performance initiative—a detailed budget perspective—and GPRA—a more strategic planning view. In particular, agency officials are concerned with problems in trying to respond to both and overwhelmingly agreed that the PART required a large amount of agency resources to complete. Moreover, some agency officials said that the PART (a program-specific review) is not well suited to one of the key purposes of strategic plans—to convey agencywide, long-term goals and objectives for all major functions and operations. In addition, the time horizons are different for the two initiatives—PART assessments focus on program accomplishments to date while GPRA strategic planning is long-term and prospective in nature.

Changes Made to GPRA in
the PART Process Create
Uncertainty About
Opportunities for
Substantive Input by
Interested Parties and
Congressional Stakeholders

As noted above, PART goals and measures must meet OMB's needs, while GPRA is a broader process involving the development of strategic and performance goals and objectives to be reported in strategic and annual plans. As a phased reform, GPRA required development of the planning framework first, but also explicitly encouraged links to the budget.²⁰ Our work has shown that under GPRA agencies have made significant progress.²¹ Additionally, GPRA requires agencies to consult with Congress and solicit the views of other stakeholders as they develop their strategic plans.²² We have previously reported²³ that stakeholder involvement appears critical for getting consensus on goals and measures. Stakeholder involvement can be particularly important for federal agencies because they operate in a complex political environment in which legislative mandates are often broadly stated and some stakeholders may strongly disagree about the agency's mission and goals.

The relationship between the PART and its process and the broader GPRA strategic planning process is still evolving. 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 is inevitable. Compared to the relatively open-ended GPRA process any 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. Since different stakeholders have different needs and no one set of goals and measures can serve all purposes, the PART can complement GPRA but should not replace it.

²⁰ 31 U.S.C. § 1115(a) (2003).

²¹ U.S. General Accounting Office, *Managing for Results: Agency Progress in Linking Performance Plans With Budgets and Financial Statements*, GAO-02-236 (Washington, D.C.: Jan. 4, 2002).

²² 5 U.S.C. § 306(d) (2003).

²³ U.S. General Accounting Office, *Agencies' Strategic Plans Under GPRA: Key Questions to Facilitate Congressional Review (Version 1)*, GAO/GGD-10.1.16 (Washington, D.C.: May 1997).

Although these tensions between the need for internal deliberations and broader consultations are inevitable, if the PART is to be accepted as a credible element in the development of the President's budget proposal, congressional understanding and acceptance of the tool and its analysis will be important. In order for performance information to more fully inform resource allocations, decision makers must also feel comfortable with the appropriateness and accuracy of the performance information and measures associated with these goals. It is unlikely that decision makers will use performance information unless they believe it is credible and reliable and reflects a consensus about performance goals among a community of interested parties. Similarly, the measures used to demonstrate progress toward a goal, no matter how worthwhile, cannot serve the interests of a single stakeholder or purpose without potentially discouraging use of this information by others.

While it is still too soon to know whether OMB-directed measures will satisfy the needs of other stakeholders and GPRA's broader planning purposes, several appropriations subcommittees have stated, in their appropriations hearings, the need to link the PART with congressional oversight. For example, the House Committee on Appropriations, Subcommittee on the Department of the Interior and Related Agencies notes that while it supports the PMA, the costs of initiatives associated with it have generally not been requested in annual budget justifications or through reprogramming procedures.²⁴ The Subcommittee, therefore, has been unable to evaluate the costs, benefits, and effectiveness of these initiatives or to weigh the priority that these initiatives should receive as compared with ongoing programs funded in the Interior Appropriations bill. Similarly, the House Report on Treasury and Transportation Appropriations included a statement in support of the PART, but noted that the administration's efforts must be linked with the oversight of Congress to maximize the utility of the PART process, and that if the administration treats as privileged or confidential the details of its rating process, it is less likely that Congress will use those results in deciding which programs to fund. Moreover, the Subcommittee said it expects OMB to involve the House and Senate Committees on Appropriations in the development of the PART ratings at all stages in the process.²⁵

²⁴ H.R. Rep. No. 108-195, p. 8 (2003).

²⁵ H.R. Rep. No. 108-243, pp. 168-69 (2003).

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 new program, during the annual appropriations process, and in its oversight of federal operations, opportunities exist for Congress to more systematically articulate performance goals and outcomes for key programs of major concern and to allow for timely congressional input in the selection of the PART programs to be assessed.

Conclusions and General Observations

OMB, through its development and use of the PART, has more explicitly infused performance information into the budget formulation process; increased the attention paid to evaluation and performance information; and ultimately, we hope, increased the value of this information to decision makers and other stakeholders. By linking performance information to the budget process, OMB has provided agencies with a powerful incentive for improving data quality and availability. The level of effort and involvement by senior OMB officials and staff clearly signals the importance of this strategy in meeting the priorities outlined in the PMA. OMB should be credited with opening up for scrutiny—and potential criticism—its review of key areas of federal program performance and then making its assessments available to a potentially wider audience through its Web site.

While the PART clearly serves the needs of OMB in budget formulation, questions remain about whether it serves the needs of other key stakeholders. The PART could be strengthened to enhance its credibility and prospects for sustainability by such actions as (1) improving agencies' and OMB's capacity to cope with the demands of the PART, (2) strengthening the PART guidance, (3) expanding the base of credible performance information by strategically focusing evaluation resources, (4) selecting programs for assessment to facilitate crosscutting comparisons and trade-offs, (5) broadening the dialogue with congressional stakeholders, and (6) articulating and implementing a complementary relationship between PART and GPRA.

OMB's ambitious schedule for assessing all federal programs by the fiscal year 2008 President's Budget will require a tremendous commitment of OMB's and agencies' resources. Implementation of the PART recommendations will be a longer-term and potentially more significant result of the PART process than the scores and ratings. No less important will be OMB's involvement both in encouraging agency progress and in signaling its continuing commitment to improving program management and results through the PART. OMB has created a template by which

agencies report on the status of the recommendations and left follow-up on the recommendations to each RMO. However, there is no single focal point for evaluating progress and the results of agency efforts governmentwide; without this it will be difficult for OMB to judge the efficacy of the PART and to know whether the increased workload and trade-offs made with other activities is a good investment of OMB and agency resources.

The goal of the PART is to evaluate programs systematically, consistently, and transparently, but in practice, the tool requires OMB staff to use independent judgment in interpreting the guidance and in making yes or no decisions for what are often complex federal programs. These difficulties are compounded by poor or partial program performance data. Therefore, it is not surprising that we found inconsistencies in our analysis of the fiscal year 2004 PART assessments. Recognizing the inherent limitations of any tool to provide a single performance answer or judgment on complex federal programs with multiple goals, continued improvements in the PART guidance, with examples throughout, can nonetheless help encourage a higher level of consistency as well as transparency.

The PART requires more performance and evaluation information than agencies currently have, as demonstrated by the fact that OMB rated over 50 percent of the programs for fiscal year 2004 as “results not demonstrated” because they “did not have adequate performance goals” or “had not yet collected data to provide evidence of results.” In the past, we too have noted limitations in the quality of agency performance and evaluation information and in agency capacity to produce rigorous evaluations of program effectiveness. Furthermore, our work has shown that few agencies deployed the rigorous research methods required to attribute changes in underlying outcomes to program activities. However, program evaluation information often requires large amounts of agency resources to produce, and the agency and OMB may not agree on what is important to measure, particularly when a set of measures cannot serve multiple purposes. Agreement on what are a department or agency’s critical, high-risk programs and how best to evaluate them could help leverage limited resources and help determine what are the most important program evaluation data to collect.

Federal programs are designed and implemented in dynamic environments where competing program priorities and stakeholders’ needs must be balanced continually and new needs must be addressed. GPRA is a broad legislative framework that was designed to be consultative with Congress and other stakeholders and allows for varying uses of performance

information, while the PART applies evaluation information to support decisions and program reviews during the executive budget formulation process. While the PART reflects the administration's management principles and the priority given to using performance information in OMB's decision-making process, its focus on program-level assessments cannot substitute for the inclusive, crosscutting strategic planning required by GPRA. Moreover, GPRA can anchor the review of programs by providing an overall strategic context for programs' contributions toward agency goals. Although PART and GPRA serve different needs, a strategy for integrating the two could help strengthen both.

Opportunities exist to develop a more strategic approach to the selection and prioritization of areas to be assessed 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 could not only help ration scarce analytic resources but could also focus decision makers' attention on the most pressing policy and program issues. Moreover, such an approach could facilitate the use of PART assessments to review the relative contributions of similar programs to common or crosscutting goals and outcomes.

As part of the President's budget preparation, the PART clearly must serve the President's interests. However, it is unlikely that the broad range of actors whose input is critical to decisions will use performance information unless they believe it is credible and reliable and reflects a consensus about performance goals among a community of interested parties. Similarly, the measures used to demonstrate progress toward a goal, no matter how worthwhile, cannot appear to serve a single set of interests without potentially discouraging use of this information by others. If the President or OMB wants the PART and its results to be considered in the congressional debate, it will be important for OMB to (1) involve congressional stakeholders early in providing input on the focus of the assessments; (2) clarify any significant limitations in the assessments as well as the underlying performance information; and (3) initiate discussions with key congressional committees about how they can best take advantage of and leverage PART information in authorizations, appropriations, and oversight processes.

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 new 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

In order to facilitate an understanding of congressional priorities and concerns, we suggest that Congress consider the need for a strategy that could include (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 have seven recommendations to OMB for building on and improving the first year's experience with the PART and its process. We recommend that the Director of OMB take the following actions:

- Centrally monitor agency implementation and progress on PART recommendations and report such progress in OMB's budget submission to Congress. Governmentwide councils may be effective vehicles for assisting OMB in these efforts.
- Continue to improve the PART guidance by (1) expanding the discussion of how the unit of analysis is to be determined to include trade-offs made when defining a unit of analysis, implications of how the unit of analysis is defined, or both; (2) clarifying when output versus outcome measures are acceptable; and (3) better defining an "independent, quality evaluation."
- Clarify OMB's expectations to agencies regarding the allocation of scarce evaluation resources among programs, the timing of such evaluations, as well as the evaluation strategies it wants for the purposes of the PART, and consider using internal agency evaluations as

evidence on a case-by-case basis—whether conducted by agencies, contractors, or other parties.

- Reconsider plans for 100 percent coverage of federal programs and, instead, target for review a significant percentage of major and meaningful government programs based on such factors as the relative priorities, costs, and risks associated with related clusters of programs and activities.
- Maximize the opportunity to review similar programs or activities in the same year to facilitate comparisons and trade-offs.
- Attempt 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.
- Seek to achieve the greatest benefit from both GPRA and PART by articulating and implementing an integrated, complementary relationship between the two.

Agency Comments

We provided a draft of this report to OMB for its review and comment. OMB generally agreed with our findings, conclusions, and recommendations. In addition, OMB outlined actions it is taking to address many of our recommendations, including refining the process for monitoring agencies' progress in implementing the PART recommendations, seeking opportunities for dialogue with Congress on agencies' performance, and continuing to improve executive branch implementation of GPRA plans and reports. OMB officials provided a number of technical comments and clarifications, which we incorporated as appropriate to ensure the accuracy of our report. OMB's comments appear in appendix IV. We also received technical comments on excerpts of the draft provided to the Departments of the Interior, Energy, and Health and Human Services. Comments received from the Departments of Energy and the Interior were incorporated as appropriate. The Department of Health and Human Services had no comments.

OMB noted that performance information gleaned from the PART process has not only informed budget decisions but has also helped direct program management, identified opportunities to improve program design, and promoted accountability. We agree. As shown in figure 1 in our report, we

found that 82 percent of PART recommendations addressed program assessment, design, and management issues; only 18 percent of the recommendations had a direct link to funding matters.

We are sending copies of this report to the Director of OMB, appropriate congressional committees, and other interested members of Congress. We will also make copies available to others upon request. In addition, the report will be available at no charge on the GAO Web site at <http://www.gao.gov>.

If you or your staff have questions about this report, please contact Paul Posner at (202) 512-9573 or posnerp@gao.gov. An additional contact and key contributors to this report are listed in appendix V.

A handwritten signature in black ink, appearing to read "D. M. Walker", with a horizontal line extending to the right.

David M. Walker
Comptroller General
of the United States

Scope and Methodology

To address the objectives in this report, we reviewed Office of Management and Budget (OMB) materials and presentations on the development and implementation of the Program Assessment Rating Tool (PART) as well as the results of the PART assessments. Our review of materials included instructions for using PART, OMB's testimony concerning PART, and public remarks made by OMB officials at relevant conferences and training. We also reviewed PART-related information on OMB's Web site, including the OMB worksheets used to support the assessments, and attended OMB's PART training for the fiscal year 2004 process.

For this report, we focused on the process and final results of the fiscal year 2004 PART process, but also looked at the initial stages of the fiscal year 2005 process. We compared the PART guidance for both years and asked agency and OMB staff to discuss generally the differences between the 2 fiscal years. We did not review the final results for the fiscal year 2005 PART, which are embargoed until the publication of the President's fiscal year 2005 budget request. For the same reasons, we did not review the results of any reassessments conducted for fiscal year 2005 on programs originally assessed for fiscal year 2004. This report presents the experiences of staff from the three departments and OMB officials who we interviewed. We did not directly observe the PART process (for either year) in operation nor did we independently verify the PART assessments as posted on OMB's Web site or the program or financial information contained in the documents provided as evidence for the PART assessments. We did, however, take several steps to ensure that we reliably downloaded and combined the PART summaries and worksheets with our budget and 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) having transcribed data elements from all programs checked back to source files; and (3) having selected, computer-processed data elements checked back to source files for a random sample of programs and also for specific programs identified in our analyses.

To better understand the universe of programs OMB assessed for fiscal year 2004, we developed overall profiles of PART results and examined relationships between such characteristics as type of program, type of recommendation, overall rating, total PART score, and answers for each question on PART. This review enabled us to generally confirm some information previously reported by OMB, for example, that PART scores do not automatically determine proposed funding and that grant programs scored lower overall than other types of programs. It also allowed us to

select a sample of programs for more in-depth review, and this sample was used to determine which OMB and agency officials we interviewed.

To gain a better understanding of the PART process at both OMB and agencies, to inform our examination of the usefulness of PART as an evaluation tool, and to obtain various perspectives on the relationship between PART and GPRA, we interviewed officials at OMB and three selected departments. At OMB, we interviewed a range of staff, such as associate directors, deputy assistant directors, branch chiefs, and examiners. Specifically, we interviewed staff in two Resource Management Offices (RMO). In the Human Resources Programs RMO, we spoke with staff from the Health Division and the Education and Human Resources Division. In the Natural Resources, Energy and Science RMO we interviewed staff from the Energy and Interior Branches. In addition, we obtained the views of two groups within OMB that were convened specifically for the PART process: the Performance Evaluation Team (PET) and the Interagency Review Panel (IRP). The IRP included agency officials in addition to staff from OMB.

The three departments for which we reviewed the PART process were the Department of Energy (DOE), the Department of Health and Human Services (HHS), and the Department of the Interior (DOI). We selected these three departments based on our data analysis of program types. The departments selected and their agencies had a variety of program types (e.g., block/formula grants, competitive grants, direct federal, and research and development) that were subject to PART and could provide us with a broad-based perspective on how PART was applied to different programs employing diverse tools of government. We also chose these three departments because they had programs under PART review within the two RMOs at OMB where we did more extensive interviewing, thus enabling us to develop a more in-depth understanding of how the PART process operated for a subset of programs. We used this information to complement our broader profiling of all 234 programs assessed. Within DOE we studied the experiences of the Office of Science, the Office of Energy Efficiency and Renewable Energy, and the Office of Fossil Energy. Within HHS, we studied the experiences of the Administration for Children and Families, the Health Resources and Services Administration, and the Substance Abuse and Mental Health Services Administration. Within DOI, we studied the experiences of the Bureau of Land Management, the Bureau of Indian Affairs, and the National Park Service. We interviewed planning, budget, and program staff within each of the nine agencies as well as those

at the department level. We also reviewed relevant supporting materials provided by these departments in conjunction with these interviews.

To allow us to describe how PART was used in fiscal year 2004 to influence changes in future performance, we created a consolidated dataset in which we classified recommendations OMB made by three areas in need of improvement: (1) program design, (2) program management, and (3) program assessment. A fourth category was created for those recommendations that involved funding issues. We created a consolidated dataset of information from our analysis of recommendations and selected information from the PART program summary page and worksheet for each program.¹

In addition, for approximately 95 percent of the programs, we identified whether the basis for program funding was mandatory or discretionary. It was important to separate discretionary and mandatory programs in our review of PART's potential influence on the President's budget proposals because funding for mandatory programs is determined through authorizations, not through the annual appropriations process. Of the 234 programs that OMB assessed for fiscal year 2004, we identified 27 mandatory programs and 196 discretionary, but could not categorize 11 programs as solely mandatory or discretionary because they were too mixed to classify.²

For discretionary programs, we explored the relationship between PART results and proposed budget changes in a series of regression analyses.³ Using statistical analysis, we found that PART scores influenced proposed

¹ The PART program summary sheets are included in the *Budget of the United States Government, Fiscal Year 2004, Performance and Management Assessments* (Washington, D.C.: February 2003). The summary sheets and worksheets for the 234 programs are on OMB's Web site: <http://www.whitehouse.gov/omb/budget/fy2004/pma.html>.

² These 11 programs are animal welfare, food aid, multifamily housing direct loans and rental assistance, rural electric utility loans and guarantees, and rural water and wastewater grants and loans programs in the Department of Agriculture; the nursing education loan repayment and scholarship program in HHS; the methane hydrates program in DOE; the reclamation hydropower program in DOI; the long-term guarantees program in the U.S. Export-Import Bank; and the climate change and development assistance/population programs in the Agency for International Development.

³ We tested the regression on mandatory programs and as expected the results showed no relationship between the PART scores and the level of funding proposed in the President's Budget.

funding changes for discretionary programs; however, a large amount of variability in these changes remains unexplained. We examined proposed funding changes between fiscal years 2003 and 2004 (measured by percentage change) and the relationship to PART scores for the programs assessed in the fiscal year 2004 President's Budget. These scores are the weighted sums of scores for four PART categories: Program Purpose and Design, Strategic Planning, Program Management, and Program Results and Accountability. The corresponding weights assigned by OMB are 0.2, 0.1, 0.2, and 0.5, respectively.⁴ Tables in this appendix report regression results obtained using the method of least squares with heteroskedasticity-corrected standard errors.⁵ The same estimation method is used throughout this analysis.

Overall PART scores have a positive and statistically significant effect on discretionary program funding. The programs evaluated by OMB include both mandatory and discretionary programs. Regression results for mandatory programs showed—as expected—no relationship between → PART scores and the level of funding in the President's Budget proposal. Assessment ratings, however, can potentially affect the funding for discretionary programs either in the President's Budget proposal or in congressional deliberations on spending bills.⁶ Table 3 reports the regression results for discretionary programs.

⁴ *Budget of the United States Government, Fiscal Year 2004, Performance and Management Assessments*, 10.

⁵ For a discussion of this method, see W.H. Greene, *Econometric Analysis*, Section 10.3 (Upper Saddle River, N.J.: Prentice Hall, 2003).

⁶ *Budget of the United States Government, Fiscal Year 2001, A Citizen's Guide to the Federal Budget* (Washington, D.C.: February 2000), <http://w3.access.gpo.gov/usbudget/fy2001/guide03.html>, (downloaded April 2003), 2.

Table 3: The Effect of Overall PART Score on Proposed Funding Changes (Discretionary Programs)

Variable	Coefficient estimate	Robust standard error	t-Statistic	P-value
Overall PART score	0.536	0.159	3.38	0.001
Constant	-25.671	8.682	-2.96	0.003

Source: GAO analysis of OMB data.

Notes: R-squared = 0.058, Prob-F = 0.001, N = 196. Originally we identified 197 discretionary programs. However, no fiscal year 2004 budget estimate is reported for the Disclosed Worker Assistance program due to grant consolidation at the Department of Labor. (*Budget of the United States Government, Fiscal Year 2004, Performance and Management Assessments* (Washington, D.C.: February 2003), 191.) This reduced the number of discretionary programs to 196.

The estimated coefficient of the overall score is positive and significant. These results show that the aggregate PART score has a positive and statistically significant effect on the proposed change in discretionary programs' budget, suggesting that programs with better scores are more likely to receive larger proposed budget increases.

To examine the effect of program size on our results, we divided all programs equally into three groups—small, medium, and large—based on their fiscal year 2003 funding estimate. Regressions similar to those reported in table 3 were then performed for discretionary programs in each group. The results, reported in tables 4, 5, and 6 suggest that the statistically significant effect of overall scores on budget outcomes exists only for the smaller programs. The estimated coefficient of the overall score for large programs, which is significant but only at the 10 percent level, reflects an outlier.⁷ Once this outlier is dropped, the estimated coefficient becomes statistically insignificant.

⁷ The outlier is the Community Oriented Policing Services program with an estimated 77 percent reduction in funding (see OMB, *Budget of the U.S. Government, Fiscal Year 2004, Performance and Management Assessments*, (Washington, D.C.: February 2003), 178). The outlier in this case is identified using scatter plot and estimating with and without the outlier. The reported results for small and medium programs are not outlier driven.

Table 4: The Effect of Overall PART Score on Proposed Funding Changes (Small Discretionary Programs)

Variable	Coefficient estimate	Robust standard error	t-Statistic	P-value
Overall PART score	1.074	0.404	2.66	0.010
Constant	-50.523	21.155	-2.39	0.020

Source: GAO analysis of OMB data.

Note: R-squared = 0.092, Prob-F = 0.01, N = 71.

Table 5: The Effect of Overall PART Score on Proposed Funding Changes (Medium-Size Discretionary Programs)

Variable	Coefficient estimate	Robust standard error	t-Statistic	P-value
Overall PART score	0.306	0.188	1.62	0.109
Constant	-17.984	12.480	-1.44	0.154

Source: GAO analysis of OMB data.

Note: R-squared = 0.039, Prob-F = 0.109, N = 67.

Table 6: The Effect of Overall PART Score on Proposed Funding Changes (Large Discretionary Programs)

Variable	Coefficient estimate	Robust standard error	t-Statistic	P-value
Overall PART score	0.194	0.109	1.77	0.082
Constant	-8.216	7.778	-1.06	0.295

Source: GAO analysis of OMB data.

Note: R-squared = 0.057, Prob-F = 0.082, N = 58.

The statistical analysis suggests that among the four components of the PART questionnaire, program purpose, management, and results have statistically significant effects on proposed funding changes, but the effects of program purpose and results are more robust across the estimated models. The overall score is a weighted average of four components: Program Purpose and Design, Strategic Planning, Program Management, and Program Results and Accountability.⁸ To identify which of the four components contribute to the significant relationship observed here, we

examined the effect of each on proposed changes in programs' funding levels. Tables 7 and 8 show estimates from regressions of the proposed funding change on purpose, planning, management, and results scores for all discretionary programs as well as small discretionary programs alone.

Table 7: The Effect of PART Component Scores on Proposed Funding Changes (All Discretionary Programs)

Variable	Coefficient estimate	Robust standard error	t-Statistic	P-value
Purpose	0.325	0.127	2.56	0.011
Plan	-0.259	0.199	-1.30	0.194
Management	0.191	0.117	1.63	0.105
Results	0.363	0.205	1.77	0.078
Constant	-33.096	14.136	-2.34	0.020

Source: GAO analysis of OMB data.

Note: R-squared = 0.087, Prob-F = 0.003, N = 196.

Table 8: The Effect of PART Component Scores on Proposed Funding Changes (Small Discretionary Programs)

Variable	Coefficient estimate	Robust standard error	t-Statistic	P-value
Purpose	0.223	0.274	0.81	0.419
Plan	-0.671	0.543	-1.24	0.221
Management	0.547	0.304	1.80	0.077
Results	0.956	0.534	1.79	0.078
Constant	-42.455	34.800	-1.22	0.227

Source: GAO analysis of OMB data.

Note: R-squared = 0.149, Prob-F = 0.043, N = 71.

These results suggest that among the four components, program purpose, management, and results are more likely to affect the proposed budget changes for discretionary programs. When all discretionary programs are

⁸ *Budget of the United States Government, Fiscal Year 2004, Performance and Management Assessments*, 10.

included, the estimated coefficients are positive and significant for results (at the 10 percent level) and purpose. When only the small discretionary programs are included, the estimated coefficients are positive and significant for both management and results (at the 10 percent level). We also estimated the above regression for medium and large programs, but coefficient estimates were not statistically significant, except for the estimated coefficient of purpose for medium programs.

PART scores explain at most about 15 percent of the proposed funding changes, leaving a large portion of the variability in proposed funding changes unexplained. This suggests that most of the variance is due to institutional factors, program specifics, and other unquantifiable factors. The coefficient of determination (or R^2) is used to measure the proportion of the total variation in the regression's dependent variable that is explained by the variation in the regressors (independent variables).⁹ The maximum value of this measure across all estimated regressions is about 15 percent.

Similar analyses were carried out for changes in the proposed budget for fiscal year 2004 and congressionally appropriated amounts in fiscal year 2002. Results were qualitatively similar to those reported here.

To assess the strengths and weaknesses of PART as an evaluation tool and the consistency with which it was applied, we analyzed data from all 234 programs that OMB reviewed using PART for fiscal year 2004. As part of our examination of the consistency with which PART was applied to programs, we also focused on a subset of programs to assess the way in which certain measurement issues were addressed across those programs. The issues were selected from those identified in interviews with officials from the selected agencies described above and our own review of the PART program summaries and worksheets. Measurement issues included acceptance of output versus outcome measures of annual and long-term goals, types of studies accepted as program evaluations, acknowledgment of related programs, and justifications for judging a PART question as "not applicable." Programs were selected that formed clusters, each addressing a similar goal or shared a structural similarity pertinent to performance measurement, to examine whether PART assessment issues were handled similarly across programs when expected. We reviewed the worksheets and compared the treatment of assessment issues across specific questions

⁹ See Greene, 33.

within and across programs in a cluster to identify potential inconsistencies in how the tool was applied. We reviewed a total of 28 programs in nine clusters. The nine clusters are food safety, water supply, military equipment procurement, provision of health care, statistical agencies, block grants to assist vulnerable populations, energy research programs, wildland fire management, and disability compensation.

With the exception of our summary analyses of all 234 programs, the information obtained from OMB and agency interviews, related material, and review of selected programs is not generalizable to the PART process for all 234 programs reviewed in fiscal year 2004. We conducted our review from May through October 2003 in accordance with generally accepted government auditing standards.

The Fiscal Year 2004 PART and Differences Between the Fiscal Year 2004 and 2005 PARTs

Below we have reproduced OMB's fiscal year 2004 PART instrument. We have also included the comparison of fiscal year 2004 and fiscal year 2005 PART questions that appeared in the fiscal year 2005 PART guidance (see table 9).

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

1. Is the program purpose clear?
2. Does the program address a specific interest, problem or need?
3. Is the program designed to have a significant impact in addressing the interest, problem or need?
4. Is the program designed to make a unique contribution in addressing the interest, problem or need (i.e., not needlessly redundant of any other Federal, state, local or private efforts)?
5. Is the program optimally designed to address the interest, problem or need?

Specific Program Purpose & Design Questions by Program Type

Research and Development Programs

6. (RD. 1) Does the program effectively articulate potential public benefits?
7. (RD. 2) If an industry-related problem, can the program explain how the market fails to motivate private investment?

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

1. Does the program have a limited number of specific, ambitious long-term performance goals that focus on outcomes and meaningfully reflect the purpose of the program?
2. Does the program have a limited number of annual performance goals that demonstrate progress toward achieving the long-term goals?
3. Do all partners (grantees, subgrantees, contractors, etc.) support program-planning efforts by committing to the annual and/or long-term goals of the program?

4. Does the program collaborate and coordinate effectively with related programs that share similar goals and objectives?
 5. Are independent and quality evaluations of sufficient scope conducted on a regular basis or as needed to fill gaps in performance information to support program improvements and evaluate effectiveness?
 6. Is the program budget aligned with the program goals in such a way that the impact of funding, policy, and legislative changes on performance is readily known?
 7. Has the program taken meaningful steps to address its strategic planning deficiencies?
-

Specific Strategic Planning
Questions by Program Type

Regulatory-Based Programs

8. (RD. 1) 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

8. (Cap. 1) Are acquisition program plans adjusted in response to performance data and changing conditions?
9. (Cap. 2) Has the agency/program conducted a recent, meaningful, credible analysis of alternatives that includes trade-offs between cost, schedule and performance goals?

Research and Development Programs

8. (RD. 1) Is evaluation of the program's continuing relevance to mission, fields of science, and other "customer" needs conducted on a regular basis?
9. (RD. 2) Has the program identified clear priorities?

**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 (grantees, subgrantees, contractors, etc.) held accountable for cost, schedule and performance results?
3. Are all funds (Federal and partners') obligated in a timely manner and spent for the intended purpose?
4. Does the program have incentives and procedures (e.g., competitive sourcing/cost comparisons, IT improvements) to measure and achieve efficiencies and cost effectiveness in program execution?
5. Does the agency estimate and budget for the full annual costs of operating the program (including all administrative costs and allocated overhead) so that program performance changes are identified with changes in funding levels?
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

8. (Co. 1) Are grant applications independently reviewed based on clear criteria (rather than earmarked) and are awards made based on results of the peer review process?
9. (Co. 2) Does the grant competition encourage the participation of new/first-time grantees through a fair and open application process?
10. (Co. 3) Does the program have oversight practices that provide sufficient knowledge of grantee activities?
11. (Co. 4) Does the program collect performance data on an annual basis and make it available to the public in a transparent and meaningful manner?

Block/Formula Grant Programs

8. (B. 1) Does the program have oversight practices that provide sufficient knowledge of grantee activities?
9. (B. 2) 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

8. (Reg. 1) Did the program seek and take into account the views of affected parties including state, local and tribal governments and small businesses, in drafting significant regulations?
9. (Reg. 2) Did the program prepare, where appropriate, a Regulatory Impact Analysis that comports with OMB's economic analysis guidelines and have these RIA analyses and supporting science and economic data been subjected to external peer review by qualified specialists?
10. (Reg. 3) Does the program systematically review its current regulations to ensure consistency among all regulations in accomplishing program goals?
11. (Reg. 4) In developing new regulations, are incremental societal costs and benefits compared?
12. (Reg. 5) Did the regulatory changes to the program maximize net benefits?
13. (Reg. 6) Does the program impose the least burden, to the extent practicable, on regulated entities, taking into account the costs of cumulative final regulations?

Capital Assets and Service Acquisition Programs

8. (Cap. 1) Does the program define the required quality, capability, and performance objectives of deliverables?
9. (Cap. 2) Has the program established appropriate, credible, cost and schedule goals?

10. (Cap. 3) Has the program conducted a recent, credible, cost-benefit analysis that shows a net benefit?
11. (Cap. 4) Does the program have a comprehensive strategy for risk management that appropriately shares risk between the government and contractor?

Credit Programs

8. (Cr. 1) Is the program managed on an ongoing basis to assure credit quality remains sound, collections and disbursements are timely and reporting requirements are fulfilled?
9. (Cr. 2) Does the program consistently meet the requirements of the Federal Credit Reform Act of 1990, the Debt Collection Improvement Act and applicable guidance under OMB Circulars A-1, A-34, and A-129?
10. (Cr. 3) Is the risk of the program to the U.S. Government measured effectively?

Research and Development Programs

8. (RD. 1) Does the program allocate funds through a competitive, merit-based process, or, if not, does it justify funding methods and document how quality is maintained?
9. (RD. 2) Does competition encourage the participation of new/first-time performers through a fair and open application process?
10. (RD. 3) Does the program adequately define appropriate termination points and other decision points?
11. (RD. 4) If the program includes technology development or construction or operation of a facility, does the program clearly define deliverables and required capability/performance characteristics and appropriate, credible cost and schedule goals?

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

1. Has the program demonstrated adequate progress in achieving its long-term outcome goal(s)?
 - Long-Term Goal I:
Target:
Actual Progress achieved toward goal:
 - Long-Term Goal II:
Target:
Actual Progress achieved toward goal:
 - Long-Term Goal III:
Target:
Actual Progress achieved toward goal:
2. Does the program (including program partners) achieve its annual performance goals?
 - Key Goal I:
Performance Target:
Actual Performance:
 - Key Goal II:
Performance Target:
Actual Performance:
 - Key Goal III:
Performance Target:
Actual Performance:

Note: Performance targets should reference the performance baseline and years, e.g. achieve a 5% increase over base of X in 2000.

3. Does the program demonstrate improved efficiencies and cost effectiveness in achieving program goals each year?
4. Does the performance of this program compare favorably to other programs with similar purpose and goals?
5. Do independent and quality evaluations of this program indicate that the program is effective and achieving results?

Appendix II
 The Fiscal Year 2004 PART and Differences
 Between the Fiscal Year 2004 and 2005 PARTs

Specific Results Questions
 by Program Type

Regulatory-Based Programs

- 6. (Reg. 1) 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

- 6. (Cap. 1) Were program goals achieved within budgeted costs and established schedules?

Research and Development Programs

- 6. (RD. 1) If the program includes construction of a facility, were program goals achieved within budgeted costs and established schedules?

Table 9: Side-by-Side of the Fiscal Year 2005 PART and the Fiscal Year 2004 PART Questions

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
I. Program purpose & design				
1.1	Is the program purpose clear?	1	Same	
1.2	Does the program address a specific and existing problem, interest, or need?	2	Does the program address a specific interest, problem or need?	Wording clarified.
		3	Is the program designed to have a significant impact in addressing the interest, problem or need?	Dropped; "significant" worked against small programs and was not clear.
1.3	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 designed to make a unique contribution in addressing the interest, problem or need (i.e., is not needlessly redundant of any other Federal, state or, local or private effort)?	Wording clarified.
1.4	Is the program design free of major flaws that would limit the program's effectiveness or efficiency?	5	Is the program optimally designed to address the national, interest, problem or need?	Minor change to clarify focus; "optimally" was too broad.
1.5	Is the program effectively targeted, so that resources will reach intended beneficiaries and/or otherwise address the program's purpose directly?			New question to address distributional design.

**Appendix II
The Fiscal Year 2004 PART and Differences
Between the Fiscal Year 2004 and 2005 PARTs**

(Continued From Previous Page)

This year's question (fiscal year 2005 PART)	Last year's question (fiscal year 2004 PART)	Comment
<i>Specific Program Purpose and Design Questions by Program Type</i>		
Research and Development Programs		
RD.1	Does the program effectively articulate potential public benefits?	Dropped; covered by 1.2.
RD. 2	If an industry-related problem, can the program explain how the market fails to motivate private investment?	Dropped; covered by 1.2 and 1.5.
II. Strategic planning		
2.1	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?	1
Does the program have a limited number of specific, ambitious long-term performance goals that focus on outcomes and meaningfully reflect the purpose of the program?	Splits old II.1 into separate questions on existence of (1) long-term performance measures and (2) targets for these measures. Together, the measures and targets comprise the long-term performance goals addressed in last year's question.	
2.2	Does the program have ambitious targets and timeframes for its long-term measures?	Splits old II.1; see above.
2.3	Does the program have a limited number of specific annual performance measures that can demonstrate progress toward achieving the program's long-term goals?	2
Does the program have a limited number of annual performance goals that demonstrate progress toward achieving the long-term goals?	Splits old II.2 into separate questions on existence of (1) annual performance measures and (2) targets for these measures. Together, the measures and targets comprise the annual performance goals addressed in last year's question.	
2.4	Does the program have baselines and ambitious targets for its annual measures?	Splits old II.2; see above.
2.5	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?	3
Do all partners (grantees, sub-grantees, contractors, etc.) support program planning efforts by committing to the annual and/or long-term goals of the program?	Wording clarified.	
4	Does the program collaborate and coordinate effectively with related programs that share similar goals and objectives?	Moved to question 3.5.

Appendix II
The Fiscal Year 2004 PART and Differences
Between the Fiscal Year 2004 and 2005 PARTs

(Continued From Previous Page)

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
2.6	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?	5	Are independent and quality evaluations of sufficient scope conducted on a regular basis or as needed to fill gaps in performance information to support program improvements and evaluate effectiveness?	Wording clarified.
2.7	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?	6	Is the program budget aligned with the program goals in such a way that the impact of funding, policy, and legislative changes on performance is readily known?	Modified.
2.8	Has the program taken meaningful steps to correct its strategic planning deficiencies?	7	Same.	
<i>Specific Strategic Planning Questions by Program Type</i>				
<i>Regulatory Based Programs</i>				
2.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?	Reg. 1	Same.	
<i>Capital Assets & Service Acquisition Programs</i>				
		Cap. 1	Are acquisition program plans adjusted in response to performance data and changing conditions?	Dropped; covered in 2.CA1 and 3.CA1.
2.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?	Cap. 2	Has the agency/program conducted a recent, meaningful, credible analysis of alternatives that includes trade-offs between cost, schedule and performance goals?	Minor change.
<i>R&D Programs</i>				
R&D programs addressing technology development or the construction or operation of a facility should answer 2.CA1.				
2.RD1	If applicable, does the program assess and compare the potential benefits of efforts within the program to other efforts that have similar goals?	RD. 1	Is evaluation of the program's continuing relevance to mission, fields of science, and other "customer" needs conducted on a regular basis?	Modified.
2.RD2	Does the program use a prioritization process to guide budget requests and funding decisions?	RD. 2	Has the program identified clear priorities?	Modified.

**Appendix II
The Fiscal Year 2004 PART and Differences
Between the Fiscal Year 2004 and 2005 PARTs**

(Continued From Previous Page)

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
III. Program management				
3.1	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?	1	Same.	
3.2	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?	2	Same.	
3.3	Are funds (Federal and partners') obligated in a timely manner and spent for the intended purpose?	3	Same.	
3.4	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?	4	Same.	
3.5	Does the program collaborate and coordinate effectively with related programs?			Same as old question 2.4.
		5	Does the agency estimate and budget for the full annual costs of operating the program (including all administrative costs and allocated overhead) so that program performance changes are identified with changes in funding levels?	Now covered by guidance for question 2.7.
3.6	Does the program use strong financial management practices?	6	Same.	
3.7	Has the program taken meaningful steps to address its management deficiencies?	7	Same.	
<i>Specific Program Management Questions by Program Type</i>				
Competitive Grant Programs				
3.CO1	Are grants awarded based on a clear competitive process that includes a qualified assessment of merit?	Co. 1	Are grant applications independently reviewed based on clear criteria (rather than earmarked) and are awards made based on results of the peer review process?	Modified. Guidance also captures former question Co. 2.

**Appendix II
The Fiscal Year 2004 PART and Differences
Between the Fiscal Year 2004 and 2005 PARTs**

(Continued From Previous Page)

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
		Co.2	Does the grant competition encourage the participation of new/first-time grantees through a fair and open application process?	Now considered in guidance for answering 3.CO1, above.
3.CO2	Does the program have oversight practices that provide sufficient knowledge of grantee activities?	Co. 3	Does the agency have sufficient knowledge about grantee activities?	Wording clarified.
3.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?	Co. 4	Same.	
Block/Formula Grant Programs				
3.BF1	Does the program have oversight practices that provide sufficient knowledge of grantee activities?	B. 1	Same.	
3.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?	B. 2	Same.	
Regulatory Based Programs				
3.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?	Reg. 1	Did the program seek and take into account the views of affected parties including state, local and tribal governments and small businesses in drafting significant regulations?	Wording clarified.
3.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?	Reg. 2	Did the program prepare, where appropriate, a Regulatory Impact Analysis (RIA) that comports with OMB's economic analysis guidelines and have these RIA analyses and supporting science and economic data been subjected to external peer review, as appropriate, by qualified specialists?	Minor change.
3.RG3	Does the program systematically review its current regulations to ensure consistency among all regulations in accomplishing program goals?	Reg. 3	Same.	
		Reg. 4	In developing new regulations, are incremental societal costs and benefits compared?	Merged into new 3.RG4.

**Appendix II
The Fiscal Year 2004 PART and Differences
Between the Fiscal Year 2004 and 2005 PARTs**

(Continued From Previous Page)

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
3.RG4	Are the regulations designed to achieve program goals, to the extent practicable, by maximizing the net benefits of its regulatory activity?	Reg. 5	Did the regulatory changes to the program maximize net benefits?	Combines former questions Reg. 4, 5, & 6.
		Reg. 6	Does the program impose the least burden, to the extent practicable, on regulated entities, taking into account the costs of cumulative final regulations?	Merged in to new 3.RG4.
Capital Assets and Service Acquisition Programs				
3.CA1	Is the program managed by maintaining clearly defined deliverables, capability/performance characteristics, and appropriate, credible cost and schedule goals?			New question, covers old Cap. 1, 2, 3, and 4.
		Cap. 1	Does the program clearly define the required quality, capability, and performance objectives for deliverables and required capabilities/performance characteristics?	Merged into new 2.CA1 and 3.CA1.
		Cap 2.	Has the program established appropriate, credible, cost and schedule goals?	Merged into new 2.CA1 and 3.CA1.
		Cap 3.	Has the program conducted a recent, credible, cost-benefit analysis that shows a net benefit?	Merged into new 2.CA1 and 3.CA1.
		Cap 4.	Does the program have a comprehensive strategy for risk management that appropriately shares risk between the government and contractor?	Merged into new 2.CA1 and 3.CA1.
Credit Programs				
3.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?	Cr. 1	Same.	
		Cr. 2	Does the program consistently meet the requirements of the Federal Credit Reform Act of 1990, the Debt Collection Improvement Act and applicable guidance under OMB Circulars A-1, A-11, and A-129?	Merged into new 3.CR2.

**Appendix II
The Fiscal Year 2004 PART and Differences
Between the Fiscal Year 2004 and 2005 PARTs**

(Continued From Previous Page)

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
3.CR2	Do the program's credit models adequately provide reliable, consistent, accurate and transparent estimates of costs and the risk to the Government?	Cr. 3	Is the risk of the program to the U.S. Government measured effectively?	Combines former Cr. 2 and 3.
	Research and Development Programs			
	R&D programs addressing technology development or the construction or operation of a facility should answer 3.CA1. R&D programs that use competitive grants should answer 3.CO1, CO2 and CO3.			
3.RD1	For R&D programs other than competitive grants programs, does the program allocate funds and use management processes that maintain program quality?	RD. 1	Does the program allocate funds through a competitive, merit-based process, or, if not, does it justify funding methods and document how quality is maintained?	Modified.
		RD. 2	Does competition encourage the participation of new/first-time performers through a fair and open application process?	Covered by 3.CO1.
		RD. 3	Does the program adequately define appropriate termination points and other decision points?	Covered by 2.CA1 and 3.CA1.
		RD. 4	If the program includes technology development or construction or operation of a facility, does the program clearly define deliverables, capability/performance characteristics, and appropriate, credible cost and schedule goals?	Covered by 2.CA1 and 3.CA1.
IV. Program results				
4.1	Has the program demonstrated adequate progress in achieving its long-term performance goals?	1	Has the program demonstrated adequate progress in achieving its long-term outcome goal(s)?	Minor change.
4.2	Does the program (including program partners) achieve its annual performance goals?	2	Same.	
4.3	Does the program demonstrate improved efficiencies or cost effectiveness in achieving program goals each year?	3	Same.	
4.4	Does the performance of this program compare favorably to other programs, including government, private, etc., with similar purpose and goals?	4	Does the performance of this program compare favorably to other programs with similar purpose and goals?	Minor change.

**Appendix II
The Fiscal Year 2004 PART and Differences
Between the Fiscal Year 2004 and 2005 PARTs**

(Continued From Previous Page)

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
4.5	Do independent evaluations of sufficient scope and quality indicate that the program is effective and achieving results?	5	Same.	
<i>Specific Results Questions by Program Type</i>				
<i>Regulatory Based Programs</i>				
4.RG1	Were programmatic goals (and benefits) achieved at the least incremental societal cost and did the program maximize net benefits?		Same.	
<i>Capital Assets and Service Acquisition Programs</i>				
4.CA1	Were program goals achieved within budgeted costs and established schedules?	Cap. 1	Same.	
<i>Research and Development Programs</i>				
	R&D programs addressing technology development or the construction or operation of a facility should answer 4.CA1.	RD. 1	If the program includes construction of a facility, were program goals achieved within budgeted costs and established schedules?	Simplified.

Source: OMB Web site, <http://www.whitehouse.gov/omb/part/bpm861.pdf> (downloaded Apr. 7, 2003), 6-12.

Development of PART

Fiscal Year 2003

This administration's efforts to link budget and performance began with the fiscal year 2003 budget, in which the administration announced the "Executive Branch Management Scorecard," a traffic-light grading system to report the work of federal agencies in implementing the President's Management Agenda's five governmentwide initiatives. Each quarter, OMB assessed agencies achievement toward the "standards of success"—specific goals articulated for each of the five initiatives. Since some of the five initiatives require continual efforts, OMB also assessed agencies' progress toward achieving the standards. The fiscal year 2003 President's Budget also included OMB's assessments of the effectiveness of 130 programs and a brief explanation of the assessments. According to OMB, the assessments were based on OMB staff's knowledge of the programs and professional judgments; specific criteria were not publicly available with which to support OMB's judgments.

Fiscal Year 2004

During the spring of 2002, an internal OMB task force—PET—consisting of staff from various OMB divisions, created PART to make the process of rating programs robust and consistent across government programs. During the development of PART, OMB solicited input from interested parties both inside and outside the federal government, including GAO and congressional staff. PART was tested on 67 programs during a series of Spring Review meetings with the OMB Director. Based on these results and other stakeholder feedback, PET recommended a series of refinements to PART, such as using a four-point scale in the Results section as opposed to the "yes/no" format. Another key change was revising the Program Purpose and Design section (Section I) to remove the question "Is the federal role critical?" because it was seen as subjective—based on an individual's political views.

In July 2002, OMB issued PART in final and accompanying instructions for completing the assessments for the President's fiscal year 2004 budget submission. Later that month, OMB provided a series of training sessions on PART for staff from OMB and agencies. Agencies received completed PART assessments during early September 2002 and submitted written appeals to OMB by mid-September. OMB formed the IRP, comprising OMB and agency officials, to conduct consistency reviews¹ and provide recommendations on selected PART appeals. The IRP also provided OMB with a broad set of recommendations aimed at improving the PART based on IRP's experience with the consistency audit and appeals. OMB was to finalize all PART assessments by the end of September 2002, although both agency and OMB officials told us that changes and appeals continued through the end of the budget season. RMOs within OMB provided draft summaries of PART results to the Director of OMB during the Director's review of agencies' budget requests. The President's fiscal year 2004 budget (issued February 3, 2003) included a separate volume containing one-page summaries of the PART results for each of the 234 programs that were assessed.²

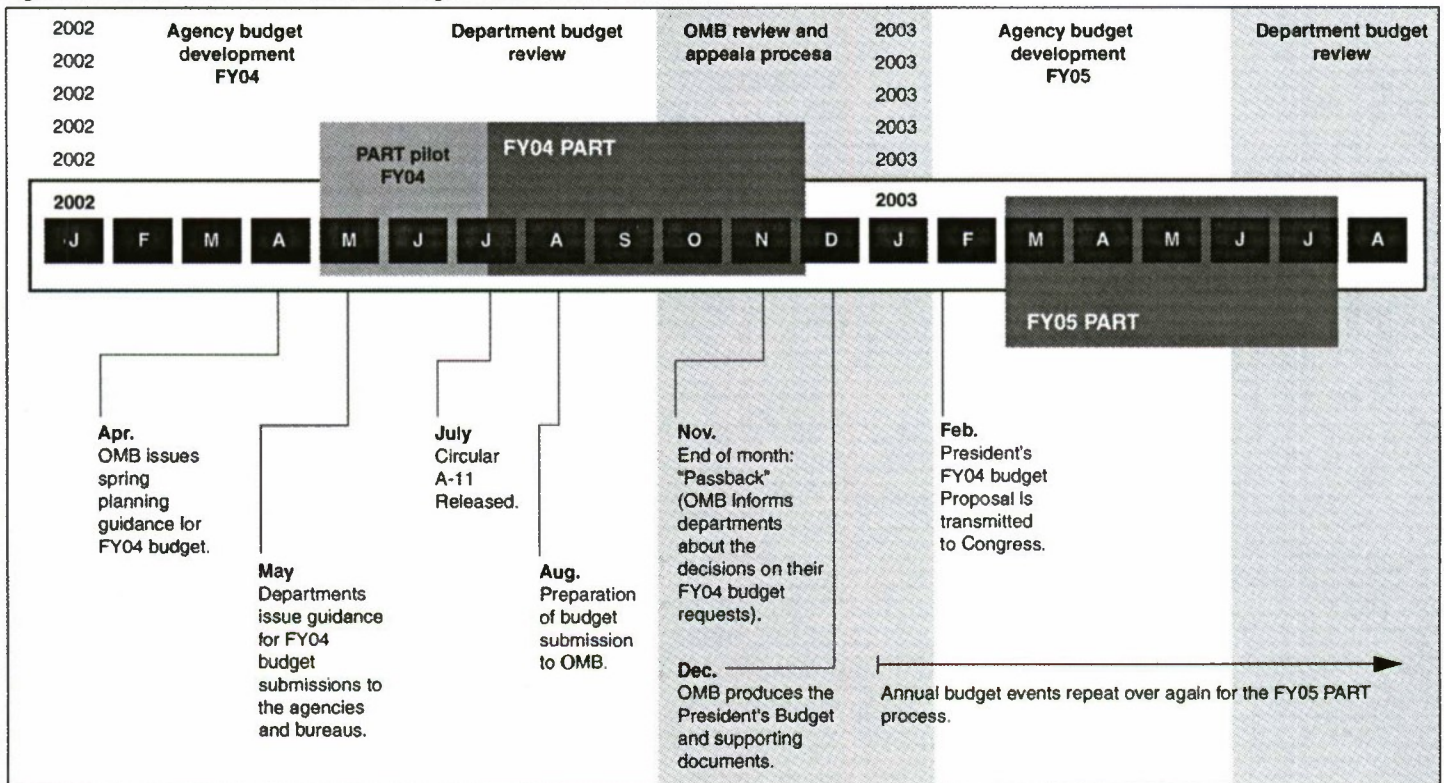
The relationship between PART and the administration's proposals was presented in agencies' budget justification materials sent to Congress. In an unprecedented move, OMB also posted PART, one-page rating results, and detailed supporting worksheets on its Web site. OMB also included its Web address in the Performance and Management Assessments volume of the budget and, in the budget itself, also described PART and its process and asked for comments on how to improve PART.

Figure 3 depicts a time line of the events related to the formulation of the President's budget request, including the key stages of PART development.

¹ According to OMB, IRP performed consistency reviews on a stratified random sample of programs that completed the PART in preparation for the fiscal year 2004 budget. While IRP made recommendations regarding its findings, it did not have the authority to enforce them.

² *Fiscal Year 2004 Budget of the United States Government, Performance and Management Assessments*, (Washington, D.C.: February 2003).

Figure 3: The PART Process and Budget Formulation Timelines



Source: GAO.

Fiscal Year 2005

For the fiscal year 2005 PART, OMB moved the entire assessment process from the fall to spring. OMB told us that the change was meant to help alleviate the burden of having the PART process overlap the end of the budget season, when workload is already so heavy. Another difference between the 2 years was that agency officials reported that OMB was more collaborative with the agencies in selecting the programs for the fiscal year 2005 PART.

Training on the PART assessments to be included in the President's fiscal year 2005 budget began in early May 2003. Agencies submitted PART appeals in early July, and OMB aimed to resolve the appeals and finalize the PART scores by the end of July. In December of 2003, RMOs were to finalize the summaries of PART results, which will be published in February along with the fiscal year 2005 President's Budget.

Comments from the Office of Management and Budget



DEPUTY DIRECTOR
FOR MANAGEMENT

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

January 16, 2004

Mr. Paul Posner
Managing Director
Federal Budget and Intergovernmental Relations
General Accounting Office
441 G Street, NW
Washington, DC 20548

Dear Mr. Posner:

Thank you for the opportunity to comment on the draft GAO report on the PART (*Performance Budgeting: Observations on the use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget*, GAO-04-174).

We appreciate GAO's extensive review of the PART process. We are particularly pleased that your report recognizes the unprecedented transparency of the PART process and materials that we have posted on our website¹ and the extensive efforts OMB has taken to make the PART process consistent across the government. We will continually strive to make the PART as credible, objective, and useful as it can be and believe that your recommendations will help us do that. As you know, OMB is already taking actions to address many of them. For instance:

- With respect to centrally monitoring PART recommendations, we have provided a simple format for agencies to follow when reporting the status of recommendation implementation to OMB and I receive these reports semi-annually. We will continue to refine this process so that sufficient attention is given to recommendation follow-up.
- As the PART relies on separate evaluations of evidence of a program's success, we agree that the judgment about what constitutes a sufficient evaluation should be based on the quality of the evaluation.
- Except for programs of insignificant size or impact, we are committed to assessing 100 percent of programs using the PART. We are sensitive to the impact this will have on OMB staff workload and will manage it accordingly.
- One of the greatest opportunities for the PART is to compare the performance of, and share best practices among, like programs across government. We will continue to use the PART for that purpose.
- We are working diligently to generate the meaningful dialogue with Congress you describe in your recommendations.

¹ See, draft report Appendix III, p. 2.

Appendix IV
Comments from the Office of Management
and Budget

- We will continue to improve agency and Executive Branch implementation of GPRA by insisting GPRA plans and reports meet the requirements of this important law and the high standards set by the PART.

Your report makes valuable conclusions and recommendations about the PART and our overall effort to create a more results-oriented government. I want to note that the PART was designed for and is used in many ways other than just budget formulation. Performance information gleaned from the PART process has not only informed budget decisions, but has also helped direct management, identified opportunities to improve program design, and promoted accountability. We believe that the PART will also greatly improve the goals and measures adopted through the GPRA strategic and performance planning processes.

Thank you for the opportunity to review and comment on your draft report. I appreciate your willingness to take our oral and written comments into consideration in the final draft. I look forward to working with you to improve the ways in which we are creating a results-oriented government.

Sincerely,



Clay Johnson III

GAO Contacts and Staff Acknowledgments

GAO Contacts

Paul Posner, (202) 512-9573
Denise Fantone, (202) 512-4997

Acknowledgments

In addition to the above contacts, Kristeen McLain, Jackie Nowicki, and Stephanie Shipman made significant contributions to this report. Thomas Beall, Joseph Byrns, Hashem Dezhbakhsh, Evan Gilman, Patrick Mullen, David Nicholson, and Mark Ramage also made key contributions to this report.

GAO's Mission

The General Accounting Office, the audit, evaluation and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through the Internet. GAO's Web site (www.gao.gov) contains abstracts and full-text files of current reports and testimony and an expanding archive of older products. The Web site features a search engine to help you locate documents using key words and phrases. You can print these documents in their entirety, including charts and other graphics.

Each day, GAO issues a list of newly released reports, testimony, and correspondence. GAO posts this list, known as "Today's Reports," on its Web site daily. The list contains links to the full-text document files. To have GAO e-mail this list to you every afternoon, go to www.gao.gov and select "Subscribe to e-mail alerts" under the "Order GAO Products" heading.

Order by Mail or Phone

The first copy of each printed report is free. Additional copies are \$2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to:

U.S. General Accounting Office
441 G Street NW, Room LM
Washington, D.C. 20548

To order by Phone: Voice: (202) 512-6000
 TDD: (202) 512-2537
 Fax: (202) 512-6061

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm

E-mail: fraudnet@gao.gov

Automated answering system: (800) 424-5454 or (202) 512-7470

Public Affairs

Jeff Nelligan, Managing Director, NelliganJ@gao.gov (202) 512-4800
U.S. General Accounting Office, 441 G Street NW, Room 7149
Washington, D.C. 20548