

GAO

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

October 2005

# PROGRAM EVALUATION

## OMB's PART Reviews Increased Agencies' Attention to Improving Evidence of Program Results





Highlights of [GAO-06-67](#), a report to the Chairman, Subcommittee on Government Management, Finance, and Accountability, Committee on Government Reform, House of Representatives

## Why GAO Did This Study

The Office of Management and Budget (OMB) designed the Program Assessment Rating Tool (PART) as a diagnostic tool to draw on program performance and evaluation information for forming conclusions about program benefits and recommending adjustments to improve results. To assess progress in improving the evidence base for PART assessments, GAO was requested to examine (1) agencies' progress in responding to OMB's recommendations to evaluate programs, (2) factors facilitating or impeding agencies' progress, and (3) whether agencies' evaluations appear to be designed to yield the information on program results that OMB expects.

## What GAO Recommends

GAO recommends that OMB (1) encourage agencies to discuss evaluation plans with OMB and congressional and other program stakeholders to ensure that their findings will be timely, relevant, credible, and used; (2) discuss a risk-based allocation of evaluation resources with agencies and congressional stakeholders; and (3) continue to improve PART guidance and examiners' training to acknowledge a wide range of appropriate evaluation methods.

OMB agreed that evaluation methodology should be appropriate to the size and nature of a program, and noted they intended to provide additional guidance in this area.

[www.gao.gov/cgi-bin/getrpt?GAO-06-67](http://www.gao.gov/cgi-bin/getrpt?GAO-06-67).

To view the full product, including the scope and methodology, click on the link above. For more information, contact Nancy Kingsbury, (202) 512-2700, [KingsburyN@gao.gov](mailto:KingsburyN@gao.gov).

## PROGRAM EVALUATION

# OMB's PART Reviews Increased Agencies' Attention to Improving Evidence of Program Results

## What GAO Found

GAO examined agency progress on 20 of the 40 evaluations OMB recommended in its PART reviews at four federal agencies: the Department of Energy, Department of Health and Human Services, Department of Labor, and Small Business Administration. About half the programs GAO reviewed had completed an evaluation in the 2 years since those PART reviews were published; 4 more were in progress and 3 were still being planned. Program restructuring canceled plans for the remaining 2 evaluations.

Several agencies struggled to identify appropriate outcome measures and credible data sources before they could evaluate program effectiveness. Evaluation typically competed with other program activities for funds, so managers may be reluctant to reallocate funds to evaluation. Some agency officials thought that evaluations should be targeted to areas of policy significance or uncertainty. However, all four agencies indicated that the visibility of an OMB recommendation brought agency management attention—and sometimes funds—to get the evaluations done. Moreover, by coordinating their evaluation activities, agencies met these challenges by leveraging their evaluation expertise and strategically prioritizing their evaluation resources to the studies that they considered most important.

Because the OMB recommendations were fairly general, agencies had flexibility in interpreting the kind of information OMB expected. Some program managers disagreed with OMB on the purpose of their evaluations, their quality, and the usefulness of “independent” evaluations by third parties unfamiliar with their programs. Agency officials concerned about an increased focus on process said that they were more interested in learning how to improve program results than in meeting an OMB checklist. Since a few programs did not discuss their evaluation plans with OMB, it is not certain whether OMB will find their ongoing evaluations useful during the programs' next PART review.

GAO concludes that

- The PART review process stimulated agencies to increase their evaluation capacity and available information on program results.
- Agencies are likely to design evaluations to meet their own needs—that is, in-depth analyses that inform program improvement. If OMB wants evaluations with a broader scope, such as information that helps determine a program's relevance or value, it will need to take steps to shape both evaluation design and execution.
- Because agency evaluation resources tend to be limited, they are most usefully focused on important areas of uncertainty. Regular performance reporting is key to good management, but requiring all federal programs to conduct frequent evaluation studies is likely to result in superficial reviews of little utility and to overwhelm agency evaluation capacity.

---

# Contents

---

<b>Letter</b>		<b>1</b>
	Results in Brief	3
	Background	4
	About Half the Programs Completed Evaluations, and Three Evaluations Were Being Planned	6
	Management Attention, Caught by OMB's Recommendations, Overcame Measurement and Funding Barriers	13
	Where OMB and Program Managers Do Not Share Expectations, Evaluations May Not Meet OMB Needs	19
	Conclusions	28
	Recommendations for Executive Action	28
	Agency Comments	29
<b>Appendix I</b>	<b>Agency Programs OMB Recommended Evaluations For in PART Reviews</b>	<b>31</b>
<b>Appendix II</b>	<b>Related Agency Program Evaluation Reports</b>	<b>33</b>
<b>Appendix III</b>	<b>Comments from the Office of Management and Budget</b>	<b>36</b>
<b>Appendix IV</b>	<b>GAO Contact and Staff Acknowledgments</b>	<b>38</b>
<b>Related GAO Products</b>		<b>39</b>
<b>Tables</b>		
	Table 1: Status of Evaluations OMB Recommended in PART Reviews, by Agency	7
	Table 2: Federal Evaluators' Views on Tailoring Designs for Program Effectiveness Evaluations	25

---

---

## **Abbreviations**

DOE	Department of Energy
DOL	Department of Labor
GPRA	Government Performance and Results Act
HHS	Department of Health and Human Services
OMB	Office of Management and Budget
OSHA	Occupational Safety and Health Administration
PART	Program Assessment Rating Tool
SBA	Small Business Administration

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.



United States Government Accountability Office  
Washington, DC 20548

October 28, 2005

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

Dear Mr. Chairman:

In the 1990s, Congress and the executive branch laid out a statutory and management framework for strengthening government performance and accountability. The Government Performance and Results Act of 1993 (GPRA) was its centerpiece.<sup>1</sup> The act was designed to provide congressional and executive decision makers with objective information on the relative effectiveness and efficiency of federal programs and spending. The current administration has made integrating performance information into budget deliberations one of five governmentwide management priorities under its President's Management Agenda.<sup>2</sup>

A central element of this initiative is the Program Assessment Rating Tool (PART), designed by the Office of Management and Budget (OMB) to provide a consistent approach to assessing federal programs in the executive budget formulation process. PART is a standard series of questions meant to serve as a diagnostic tool, drawing on available program performance and evaluation information to form conclusions about program benefits and recommend adjustments that may improve results.

However, PART's ability to do this relies on OMB's access to credible information on program performance and on policy makers' confidence in the credibility of their analysis. In our January 2004 review of PART, we found that limited availability of credible evidence on program results

---

<sup>1</sup>Pub. L. No. 103-62 (1993).

<sup>2</sup>The agenda's four other priorities are strategic management of human capital, expanded electronic government, improved financial performance, and competitive sourcing. See [http://www.whitehouse.gov/omb/budintegration/pma\\_index.html](http://www.whitehouse.gov/omb/budintegration/pma_index.html) (Oct. 21, 2005).

---

constrained the ability of OMB staff to use PART to rate programs' effectiveness.<sup>3</sup> When OMB first applied PART, for the fiscal year 2004 budget, it judged fully half the programs it reviewed as not having adequate information on results. Moreover, although OMB's assessments recommended improvements in program design, management, and assessment, half the recommendations were to improve program assessment—to identify outcome measures and obtain improved performance data or program evaluations.

To examine progress in improving the evidence base for the PART assessments, you asked us to examine

1. progress agencies have made in responding to OMB's PART recommendations that they obtain program evaluations,
2. factors that facilitated or impeded agencies' progress in obtaining these evaluations, and
3. whether the evaluations appear to have been designed to yield the information on program results that OMB anticipated.

To answer these questions, we examined progress on 20 of the 40 evaluation recommendations in the President's fiscal year 2004 budget proposal. These 20 recommendations reflect a diverse array of programs concentrated in the Department of Energy (DOE), the Department of Health and Human Services (HHS), the Department of Labor (DOL), and the Small Business Administration (SBA). We reviewed OMB and agency documents and interviewed officials in the four agencies to learn the status of the evaluations and the factors that influenced how they were conducted. We also reviewed the available evaluation plans and reports to assess whether they were likely to yield the desired information on results. We conducted our review from December 2004 through August 2005 in accordance with generally accepted government auditing standards. A list of the programs reviewed and their evaluation recommendations appears in appendix I. OMB provided written comments on a draft of this report that are reprinted in appendix III.

---

<sup>3</sup>GAO, *Performance Budgeting: Observations on the Use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget*, [GAO-04-174](#) (Washington, D.C.: Jan. 30, 2004).

---

## Results in Brief

About half of the programs we reviewed (11 of the 20) had completed an evaluation by June 2005—2 years after the fiscal year 2004 PART reviews and recommendations were published. Four additional evaluations were in progress, and 3 were still being planned. Program restructuring canceled plans for the remaining 2 evaluations. The evaluations employed a variety of study designs, reflecting differences between the programs and the questions about their performance. For example, the quality of research project portfolios had been evaluated with external peer review, while occupational safety programs had been assessed on both the results of compliance investigations and reduction in workplace injuries.

Several agencies had struggled to identify appropriate outcome measures and credible data sources before they could conduct evaluations of program effectiveness. Evaluation generally competes with other program and department activities for resources, so managers may be reluctant to reallocate resources to evaluation. Some agency officials thought that evaluations should not be conducted for all programs but should be targeted instead to areas of policy significance or uncertainty. However, all four agencies indicated that the visibility of an OMB PART recommendation brought agency management attention—and sometimes funds—to getting these evaluations done. Moreover, by coordinating their evaluation activities, agencies met these challenges by leveraging their evaluation expertise and strategically prioritizing their evaluation resources to focus on the studies that they considered to be the most important.

Because the OMB evaluation recommendations were fairly general, it is not always clear—and agencies had flexibility in interpreting—what kind of information OMB expected. Some program managers disagreed with OMB on the scope and purpose of their evaluations, their quality, and the usefulness of evaluations by independent third parties unfamiliar with their programs. Agency officials concerned about an increased focus on process said that they were more interested in learning how to improve program performance than in meeting an OMB checklist. Since a few programs did not discuss their evaluation plans with OMB, it is not certain whether OMB will find their ongoing evaluations useful during the programs' next PART review.

To help ensure that agency program evaluations are timely, relevant, credible, and used, we reiterate and expand on our previous recommendations to OMB to encourage agencies to discuss their evaluation plans with OMB and congressional stakeholders, engage in dialogue with agency and congressional stakeholders on a risk-based

---

allocation of evaluation resources across programs, and continue to improve its PART guidance and training to acknowledge a wide range of appropriate evaluation methods.

---

## Background

PART's standard series of questions is designed to determine the strengths and weaknesses of federal programs by drawing on available program performance and evaluation information. OMB applies PART's 25 questions to all programs under four broad topics: (1) program purpose and design, (2) strategic planning, (3) program management, and (4) program results (that is, whether a program is meeting its long-term and annual goals).<sup>4</sup> During the fiscal year 2004, 2005, and 2006 budget cycles, OMB applied PART to approximately 20 percent of programs each year and gave each program one of four overall ratings: "effective," "moderately effective," "adequate," or "ineffective," depending on the program's scores on those questions. OMB gave a fifth rating of "results not demonstrated" when it decided that a program's performance information, performance measures, or both were insufficient or inadequate.

The summary assessments published with the President's annual budget proposal include recommended improvements in program design, management, and assessment. For example, a summary of the review's findings might be followed by the clause "the administration will conduct an independent, comprehensive evaluation of the program," or "the Budget includes [funds] to conduct independent and quality evaluations," both of which we interpreted as an OMB recommendation to the agency to conduct such an evaluation.<sup>5</sup> In our previous analysis of the fiscal year 2004 PART reviews, we analyzed over 600 recommendations made for the 234 programs assessed and found that half of those recommended improvements in program assessment.<sup>6</sup>

---

<sup>4</sup>"Program" has no standard definition. For purposes of PART, OMB described program, its unit of analysis, as an activity or set of activities (1) clearly recognized as a program by the public, OMB, 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.

<sup>5</sup> In subsequent PART reviews, OMB encouraged agencies to propose recommendations, which they refer to as "recommended follow-up actions" in the fiscal year 2006 PART summaries.

<sup>6</sup> [GAO-04-174](#), pp.12-13.

---

PART not only relies on previous program evaluation studies to answer many of the questions but also explicitly asks, in the strategic planning section, “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?” Program evaluations are systematic studies that assess how well a program is working, and they are individually tailored to address the client’s research question. Process (or implementation) evaluations assess the extent to which a program is operating as intended. Outcome evaluations assess the extent to which a program is achieving its outcome-oriented objectives; they focus on program outputs and outcomes but may also examine program processes to understand how outcomes are produced.<sup>7</sup>

OMB first applied PART to the fiscal year 2004 budget during 2002, and the assessments were published with the President’s budget in February 2003. In January 2004, we reported on OMB and agency experiences with PART in the fiscal year 2004 budget formulation process.<sup>8</sup> We noted that PART had helped structure OMB’s use of performance information in its budget review and had stimulated agency interest in budget and performance integration. However, its effectiveness as a credible, objective assessment tool was challenged by inconsistency in OMB staff application of the guidance and limited availability of credible information on program results. Moreover, PART’s influence on agency and congressional decision making was hindered by failing to recognize differences in focus and issues of interest among the various parties involved in programmatic, policy, and budget decisions. We noted that PART’s potential value lay in recommended changes in program management and design but would require sustained attention if the anticipated benefits were to be achieved.

To strengthen PART and its use, in our January 2004 report we recommended that OMB (1) centrally monitor and report on agency progress in implementing the PART recommendations; (2) improve PART guidance on determining the unit of analysis, and defining program outcomes and “independent, quality evaluation”; (3) clarify expectations regarding agency allocation of scarce evaluation resources among programs; (4) target future reviews based on the relative priorities, costs,

---

<sup>7</sup>See GAO, *Performance Measurement and Evaluation: Definitions and Relationships*, GAO-05-739SP (Washington, D.C.: May 2005).

<sup>8</sup>GAO-04-174.

---

and risks associated with clusters of programs; (5) coordinate assessments to facilitate comparisons and trade-offs between related program; (6) consult with congressional committees on performance issues and program areas for review; and (7) articulate an integrated, complementary relationship between GPRA and PART.

Requesting that we follow up on the findings in our January 2004 report, you asked that we examine (1) OMB and agency perspectives on the effects of PART recommendations on agency operations and results, (2) OMB's efforts at ensuring an integrated relationship between PART and GPRA, and (3) steps OMB has taken to involve Congress in the PART process. A companion report addresses all three objectives—including OMB's outreach to Congress—with regard to all PART reviews.<sup>9</sup> Because of the fundamental role that the availability of program evaluations plays in conducting PART assessments, we conducted an in-depth analysis of agencies' responses to OMB recommendations that they conduct program evaluations. These recommendations were identified through the analysis of recommendations for our January 2004 review. This report focuses on agencies' progress on those evaluations and the issues involved in obtaining them. For both analyses, we examined the same four agencies' experiences with PART. The four agencies were selected to represent a range of program types (such as research and regulatory programs), large and small agencies, and, for the purposes of this report, a large proportion of the OMB evaluation recommendations.

---

## About Half the Programs Completed Evaluations, and Three Evaluations Were Being Planned

All but two of the programs we reviewed had responded to some extent to OMB's recommendations to conduct an evaluation; agencies did not plan evaluations of the other programs because they were canceled or restructured. However, after 2 years, only about half the programs had completed evaluations, partly because of lengthy study periods and partly because of some lengthy planning phases. The evaluations used a variety of study designs, reflecting differences in the programs and in the questions posed about program performance.

---

<sup>9</sup>GAO, *Performance Budgeting: PART Focuses Attention on Program Performance, but More Can Be Done to Engage Congress*, [GAO-06-28](#) (Washington, D.C.: Oct. 28, 2005).

**All Programs Responded to OMB's Recommendations, but Only Half Completed Evaluations**

About half of the programs we reviewed (11 of the 20) had completed an evaluation by June 2005—2 years after the fiscal year 2004 PART reviews and recommendations were published. Four evaluations were in progress, while 3 were still in the planning stage. Agencies did not plan an evaluation of 2 programs because those programs had been canceled or restructured. (See table 1.) Most of OMB's evaluation recommendations asked for evaluation of the specific program reviewed, while some PART reviews at DOE and DOL asked the agencies to develop a plan for conducting multiple evaluations. At DOL, where two entire regulatory agencies had been assessed, these agencies had completed multiple studies.

**Table 1: Status of Evaluations OMB Recommended in PART Reviews, by Agency**

Agency (OMB recommendations)	Completed by June 2005 <sup>a</sup>	In progress	Being planned	None planned
DOE (7)	5 expert panel reviews		1 outcome evaluation	Program discontinued
DOL (5)	2 comprehensive evaluations, multiple process and outcome evaluations, 2 regulatory reviews		1 comprehensive evaluation Additional regulatory reviews scheduled	
HHS (4)	1 process evaluation, 1 outcome evaluation	2 comprehensive evaluations (1 interim report)		
SBA (4)		Customer outcome survey (1 interim report) 2 comprehensive evaluations	1 comprehensive evaluation	Program discontinued

Source: GAO analysis.

<sup>a</sup>Comprehensive evaluations combined assessment of program processes and outcomes.

OMB gave DOE seven evaluation recommendations in its fiscal year 2004 PART reviews. Six were for research programs in basic science and nuclear energy and one was for its formula grant program to weatherize the homes of low-income families. Since one research program in the Office of Science had previously been evaluated by a panel of external experts called a committee of visitors, OMB explicitly recommended that the other research programs in that office also institute such a process by September 2003.

In response, DOE completed evaluations of five of the six research programs, but it did not plan to evaluate the sixth, the Nuclear Energy Research Initiative, because it considered this not a stand-alone program

---

but, rather, a source of funding for follow-up projects to other nuclear energy research programs. DOE revised this program's objective and now authorizes funds for its projects through the other nuclear energy research programs; thus it is no longer considered a separately funded program to be evaluated. Finally, DOE officials indicated that they had only recently gained funding for planning the evaluation of the Weatherization Assistance program. (A bibliography of related agency evaluation reports appears in app. II.)

OMB gave DOL five evaluation recommendations for fiscal year 2004. Two were for evaluations of specific DOL programs: grants to state and local agencies to provide employment-related training to low-income youths and administration of the Federal Employees Compensation Act regarding work-related injuries and illnesses. The three others were regulatory enforcement offices or agencies of DOL that were reviewed in their entirety: the Office of Federal Contract Compliance Programs, regarding equal employment opportunity; the Employee Benefits Security Administration; and the Occupational Safety and Health Administration (OSHA). OMB recommended that the last, which is a large regulatory agency, develop plans to evaluate the results of its regulatory and nonregulatory programs.

The two DOL regulatory administrations each completed several evaluations of their enforcement activities by spring 2005, as did two of the three other DOL programs we reviewed. DOL is waiting to conduct an evaluation of the fifth—the youth employment program—until after its reauthorization because that is expected to result in an increased focus on out-of-school youths and a significant change in program activities. In addition, OSHA completed two regulatory “lookback” reviews—assessing the cumulative effects of a regulation over time—one in 2004 and another in 2005. Program officials indicated that they had developed a plan for conducting lookback reviews of employee benefit regulations beginning in fiscal year 2006.

OMB recommended evaluations for four diverse HHS programs: (1) grants and technical assistance to states to increase childhood disease immunization, (2) grants to states to help recently arrived refugees find employment, (3) education loan repayment and scholarships for nurses in return for serving in facilities facing a nursing shortage, and (4) direct assistance in constructing sanitation facilities for homes for American Indians and Alaskan Natives. Evaluations of the two state grant programs were still in progress during our review, although an interim report on the immunization program was available. Reports from the two other program

---

evaluations had recently been completed and were under departmental review.

OMB recommended evaluations for four SBA programs: (1) support for existing Business Information Centers that provide information and access to technology for small businesses; (2) use of volunteer, experienced business executives to provide basic business counseling and training to current and prospective entrepreneurs; (3) Small Business Development Centers that provide business and management technical assistance to current and prospective entrepreneurs; and (4) the small business loan program that provides financing for fixed assets. OMB also asked all three counseling programs to develop outcome-oriented annual and long-term goals and measures.

SBA is conducting customer surveys and had recently initiated a comprehensive evaluation of one its counseling programs, and is planning one for the other in fiscal year 2006. Another evaluation has begun to compare the costs, benefits, and potential duplication of its business loan programs. SBA planned no evaluation of the Business Information Centers program because the program was canceled, partly as a result of the PART review and an internal cost allocation study. In reassessing the need for the program, SBA decided that because of the increase in commercially available office supplies and services and the accessibility of personal computers over the years, such a program no longer needed federal government support.

---

### Evaluation Design and Focus Differed, Reflecting Different Program Purposes and Structures

Because evaluations are designed around programs and what they aim to achieve, the form of the evaluations reflected differences in program structure and anticipated outcomes. The evaluations were typically multipurpose, including questions about results as well as the agency processes that managers control in order to achieve those results, and designed to respond to OMB and yield actionable steps that programs could take to improve results.

The Nursing Education Loan Repayment and Scholarship programs aim to increase the recruitment and retention of professional nurses by providing financial incentives in exchange for service in health care facilities that are experiencing a critical shortage of nurses. The ongoing evaluation of the two programs combined was shaped by the reporting requirements of the

---

Nurse Reinvestment Act of 2002.<sup>10</sup> The act requires HHS to submit an annual report to Congress on the administration and effect of the programs. Each yearly report is to include information such as the number of enrollees, scholarships, loan repayments and grant recipients, graduates, and recipient demographics to provide a clear description of program beneficiaries. Program beneficiaries are compared with the student, nurse applicant and general populations to assess success in outreach. Information pertaining to beneficiaries' service in health care facilities is important for determining whether program conditions and program goals have been met.<sup>11</sup> The number of defaulters, default rate, amount of outstanding default funds, and reasons for default are reported for each year. These data as well as follow-up data on whether beneficiaries remain in targeted facilities after their term of commitment will be important in assessing the overall cost-benefit of the program. Subsequent data collection will establish trends and allow for a cost-benefit analysis in the future.

The Indian Health Service Sanitation Facilities Construction delivers construction and related program services to provide drinking water and waste disposal facilities for American Indian and Alaska Native homes, in close partnership with tribes. Among other issues, the evaluation examined key areas of service delivery, while the health benefits of clean water were assumed. Specifically, project needs identification and project portfolio management were evaluated to see how well construction efforts are prioritized and targeted to areas of greatest need, and whether facilities construction projects are competently designed, timely, and cost-effective. The completed evaluation recommended that the agency consider integrating its separate data systems into a single portfolio management system to represent all projects or, at least, to adopt standardized project management and financial tracking systems.

The primary responsibility of DOL's Office of Federal Contract Compliance Programs is to implement and enforce rules banning

---

<sup>10</sup> Pub. L. No. 107-205 (2002).

<sup>11</sup> The Nursing Education Loan Repayment Program offers registered nurses financial assistance to repay educational loans in exchange for service in a critical shortage facility. Participants contract to work full-time in a critical shortage facility. For 2 years of service, the program pays up to 60 percent of the total qualifying loan balance. For the Nursing Scholarship Program, participants incur a year of full-time obligated service for each full or partial year of support, with a minimum of a 2-year service obligation of full-time clinical service at a health facility with a critical shortage of nurses.

---

discrimination and establishing affirmative action requirements for federal contractors and subcontractors. Because of the time and expense involved in conducting compliance reviews and complaint investigations, the office is attempting to target establishments for review based in part on an analytic prediction that they will be found to discriminate. The focus of its effectiveness evaluation, therefore, was on identifying a targeting approach and measuring change in the rate of discrimination among federal contractors during the period of oversight. The logic for this choice of outcome measure was based on the expectation that overall rates of discrimination would decrease if the oversight programs were effective. Using data on the characteristics of establishments that had already been reviewed, evaluators used statistical procedures to estimate a model of the probability of discrimination. The coefficients from that model were then used to predict rates of discrimination among contractors who had not been reviewed and among noncontractors. The analysis showed that the office effectively targeted selected establishments for review, but there was no measurable effect on reducing employment discrimination in the federal contractor workforce overall. To improve the office's effectiveness, the evaluators recommended that the office focus on establishments with the highest predicted rates of discrimination rather than employ its previous approach, targeting larger establishments that are likely to affect a greater number of workers.

The DOE Office of Science used a peer review approach to evaluating its basic research programs, adapting the committee of visitors model that the National Science Foundation had developed. Because it is difficult to predict the findings of individual basic research projects, science programs have adapted the peer review model they use for merit selection of projects to evaluate their portfolios of completed (and ongoing) research. The Office of Science convenes panels of independent experts as external advisers to assess the agency's processes for selecting and managing projects, the balance in the portfolio of projects awarded, and progress in advancing knowledge in the research area and in contributing to agency goals. Panel reviews generally found these programs to be valuable and reasonably well-managed and recommended various management improvements such as standardizing and automating documentation of the proposal review process, adopting program-level strategic planning, and increasing staffing or travel funds to increase grantee oversight.

OSHA, pursuant to section 610 of the Regulatory Flexibility Act and section 5 of Executive Order 12866, must conduct lookback studies on OSHA standards, considering public comments about rules, the continued

---

need for them, their economic impacts, complexity, and whether there is overlap, duplicity, or conflict with other regulations.<sup>12</sup> OSHA recently concluded a lookback review on its Ethylene Oxide standard and issued a final report on another lookback review that examined the Presence Sensing Device Initiation standard for mechanical power presses.<sup>13</sup> A press equipped with a sensing device initiates a press cycle if it senses that the danger zone is empty, and if something should enter the zone, the device stops the press. Accidents with mechanical presses result in serious injuries and amputations to workers every year.

In the sensing device lookback review, OSHA examined the continued need for the rule, its complexity, complaints levied against the rule, overlap or duplication with other rules, and the degree to which technology, economic conditions, or other factors have changed in the area affected by the rule. Typically, once a standard is selected for a lookback review, the agency gathers information on experience with the standard from persons affected by the rule and from the general public through an announcement in the Federal Register. In addition, available health, safety, economic, statistical, and feasibility data are reviewed, and a determination is made about any contextual changes that warrant consideration. In conducting such reviews, OSHA determines whether the standards should be maintained without change, rescinded, or modified. OSHA found that there was a continued need for the rule but that to achieve the expected benefits of improved worker safety and employer productivity, the rule needed to be changed. Although the technology for sensing device systems had not changed since their adoption in 1988, the technology for controlling mechanical presses had changed considerably, with press operation now often controlled by computers, introducing hazards that were not addressed initially by the standard.

---

<sup>12</sup> The Regulatory Flexibility Act, 5 U.S.C. § 610, and Executive Order No. 12866, Regulatory Planning and Review, Sept. 30, 1993, 58 Fed. Reg. 51735 (Oct. 4, 1993), require certain regulatory agencies to conduct such periodic reviews of their rules.

<sup>13</sup> 29 C.F.R. §§ 1910.1047, 1910.217 (2005).

---

## Management Attention, Caught by OMB's Recommendations, Overcame Measurement and Funding Barriers

Agency officials described two basic barriers to completing the evaluations that OMB recommended: obtaining valid measures of program outcomes to assess effectiveness and obtaining the financial resources to conduct independent evaluations. Although most of the program officials claimed that they had wanted to conduct such evaluations anyway, they noted that the visibility of an OMB recommendation brought evaluation to the attention of their senior management, and sometimes evaluation funds, so that the evaluations got done. Indeed, in response to the PART reviews and recommendations, two of the agencies initiated strong, centrally led efforts to build their evaluation capacity and prioritize evaluation spending.

---

## Measurement Challenges Delayed Evaluation Starts

To evaluate program effectiveness, agencies needed to identify appropriate measures of the outcomes they intended to achieve and credible data sources for those measures. However, as noted in our previous report, many programs lacked these and needed to develop new outcome-oriented performance measures in order to conduct evaluations.

Agency officials identified a variety of conceptual and technical barriers to measuring program outcomes similar to those previously reported as difficulties in implementing performance reporting under GPRA.<sup>14</sup> SBA officials acknowledged that before the PART reviews, they generally defined their programs' performance in terms of outputs, such as number of clients counseled, rather than in outcomes, such as gains in small business revenue or employment. SBA revised its strategic plan in fall 2003 and worked with its program partners to develop common definitions across its counseling programs, such as who is the client or what constitutes a counseling session or training. Since SBA had also had limited experience with program evaluation, it contracted for assistance in designing evaluations of the economic impact of its programs.

DOL had difficulty conceptualizing the outcomes of regulations in monetary terms to produce the cost-benefit analyses that PART (and the Regulatory Flexibility Act) asks of regulatory programs. For instance, OSHA has historically considered the likely controversy of quantifying the value of a human life in calculating cost-benefit ratios for developing

---

<sup>14</sup> GAO, *Results-Oriented Government: GPRA Has Established a Solid Foundation for Achieving Greater Results*, [GAO-04-38](#) (Washington, D.C.: Mar. 10, 2004), p. 88 noted these previously reported challenges: developing outcome-oriented measures, isolating the impact of a program, and obtaining timely, useful performance data.

---

worker health and safety regulations. OSHA officials explained that the Assistant Secretary had helped to mitigate such a controversy by issuing a July 2003 memorandum that directed OSHA staff to identify costs, benefits, net benefits, and the impact of economically significant regulations and their significant alternatives, as well as discuss significant nonmonetized costs and benefits.

DOL officials noted that designing a cumulative assessment of the net benefits of employer reporting requirements for pension and health benefit plans was complicated. For example, a primary benefit of reporting is to aid the agency's ability to enforce other benefit plan rules and thereby protect or regain employees' benefits. They also pointed out that although health and safety regulations are mandatory, employers are not required to offer benefit plans, so a potential cost of regulators' overreaching in their enforcement actions could be discouraging employers from offering these pension and health benefits altogether.

DOE officials acknowledged that they could not continue to use state evaluations to update the national estimates of energy savings from a comprehensive evaluation of weatherization assistance conducted a decade ago. They recognized that assumptions from the original national evaluation could no longer be supported and that a new, comprehensive national evaluation design was needed. They noted new hurdles to measuring reductions in home heating costs since the previous evaluation: (1) monthly electric bills typically do not isolate how much is spent on heating compared with other needs, such as lighting, and (2) the increased privatization of the utility industry is expected to reduce government access to the utilities' data on individual household energy use.

Other barriers were more operational, such as the features of a program's data system that precluded drawing the desired evaluative conclusions. For one, regulations need to be in place for a period of years to provide data adequate for seeing effects. HHS officials noted that their databases did not include the patient outcome measures OMB asked for and that they would need to purchase a longitudinal study to capture those data. They also noted that variation in the form of states' refugee assistance programs and data systems, as well as regional variation in refugees' needs, made it difficult to conduct a national evaluation. Their evaluation especially relied on the cooperation of state program coordinators. DOL officials pointed out that the federal employees' compensation program's data system was developed for employee and management needs and did not lend itself to making comparisons with the very different state employee compensation programs.

---

Agencies, with Limited Funds, Delayed or Narrowed Evaluations and Questioned the Need to Evaluate All Programs

Evaluation generally competes for resources with other program and department activities. Contracts for external program evaluations that collect and analyze new data can be expensive. In a time of tight resources, program managers may be unwilling to reallocate resources to evaluation. Agencies responded to such limitations by delaying evaluations or cutting back on an evaluation's scope. Some agency officials thought that evaluations should not be conducted for all programs but should be targeted instead to areas of uncertainty.

HHS's Office of Refugee Resettlement—which was allotted funds especially for its evaluation—is spending \$2 million to evaluate its refugee assistance program over 2 years. Costs are driven primarily by the collection of data through surveys, interviews, and focus groups and the need for interpreters for many different languages. Given the size and scope of the program, even with \$2 million, program officials would have liked to have more time and money to increase the coverage of their national program beyond the three sites they had selected.

DOL program officials explained that although they had had a large program evaluation organization two decades ago, the agency downsized in 1991, the office was eliminated, and now they must search for program evaluation dollars. The program spent \$400,000 for an 18-month evaluation of the Federal Employees Compensation Act program, which relied heavily on program administrative data, but they also spent a large amount of staff time educating and monitoring the contractor. Program officials were disappointed with the lack of depth in the evaluation. They believed that their evaluation contractor did not have enough time to plan and conduct a systematic survey, and consequently, their selective interview data were less useful than they would have liked.

DOE program officials indicated that they have been discussing an evaluation of Weatherization Assistance since spring 2003, but not having identified funds for an evaluation, they have not been able to develop a formal evaluation plan. They had no budget line item for evaluation, so they requested one in their fiscal year 2005 appropriations. Although there was congressional interest in an evaluation, additional funds were not provided in fiscal year 2005. DOE instructed program officials to draw money for evaluation from the 10 percent of the program's funds that are set aside for training and technical assistance, increase the federal share from 1.5 percent to 2 percent, and reduce the states' share to 8 percent. Program officials indicated that the amount from the technical assistance account would cover only planning and initial implementation activities, not the bulk of the evaluation itself. And they were concerned about

---

displacing existing training, so they were still looking for an evaluation funding commitment.

Agency officials also questioned PART's assumption that all programs should have evaluations. SBA officials indicated that some agency appropriations generally precluded SBA's spending program funds on any but specifically identified program activities. Thus, evaluations had to be funded from agency administrative funds. They thought that it was unreasonable to ask a small agency to finance several program evaluations, as might be expected of a larger agency. SBA dealt with this by conducting evaluations sequentially as funds became available. DOL program officials also thought that spending several hundred thousand dollars for a comprehensive evaluation study was a reasonable investment for a \$2.5 billion program but not for small programs. They did not believe that all programs need to be evaluated—especially in a time of budget deficits. They recommended that OMB and agencies should “pick their shots” and should be more focused in choosing evaluations to conduct. They suggested a risk-based approach, giving higher priority to evaluating programs for which costs are substantial and effectiveness uncertain.

---

## OMB's Recommendations Increased Management Attention and Investment in Evaluation

Most of the agency officials we interviewed declared that they valued evaluation. For example, HHS and DOE officials described evaluation as part of their culture. Many said they had already been planning to do something similar to the evaluation that OMB had recommended. In a couple of cases, OMB's recommendation appeared to have been shaped by planned or ongoing activities. However, officials in all four agencies indicated that the visibility of a PART recommendation and associated OMB pressure brought management attention, and sometimes funds, to getting the evaluations done.

HHS departmental officials said that the agency was a federal leader in terms of evaluation capacity, and that they spend approximately \$2.6 billion a year on agency-initiated research, demonstrations, and evaluation. They stated that it is part of their culture to conduct evaluations—because their program portfolio is based in the physical and social sciences. DOE officials said that they embraced the PART process because, as an agency with a significant investment in advancing science and technology, DOE had already been using similar processes, such as peer review, to evaluate its programs. DOE officials noted that DOE had developed a basic evaluation mechanism—independent peer review—that all its research programs undertake. Officials in the Office of Energy Efficiency and Renewable Energy developed a corporate peer review

---

guide summarizing best practices in this field and considered their peer review process as “state of the art,” as it is used as a model nationally and globally.<sup>15</sup>

In other cases, agency or congressional interest in evaluation seemed to set the stage for OMB evaluation recommendations. For example, while OMB was reviewing the Nursing Education Loan Repayment program, the Nursing Reinvestment Act of 2002 was enacted, expanding the program and instituting a requirement for annual reports after the first 18 months. The reports were to include data on the numbers of loan applicants and enrollees, the types of facilities they served in, and the default rates on their loans and service commitments and an evaluation of the program’s overall costs and benefits. OMB then recommended that the agency evaluate the program’s impact, develop outcome measures, and begin to track performance against newly adopted benchmarks. To respond to OMB’s request for a long-term outcome measure, the agency agreed to also collect information on how long beyond their service commitment nurses stay in service in critical shortage facilities. In another example previously discussed, the DOE Office of Science had already initiated committee of visitors reviews for its Basic Energy Sciences program, which OMB then recommended for other research programs in that office.

The PART and President’s Management Agenda pressed agencies to report progress on the recommendations. OMB published the cumulative set of completed PART review summaries, including the recommendations, in the President’s budget proposals for fiscal years 2004 through 2006. In the fiscal year 2006 budget, OMB reported on the status of its previous recommendations in the PART summaries, whether action had been taken or completed. OMB also asked agencies to report on their progress in implementing PART recommendations to provide input into its quarterly scorecards on agencies’ progress in implementing the President’s Management Agenda initiatives. In addition, OMB precluded agencies from being scored “green” on Budget and Performance Integration if more than 10 percent of their programs were rated “results not demonstrated” 2 years in a row. DOE and DOL program officials reported being asked to update the status of the recommendations every 2 to 3 months. HHS officials noted that since fall 2004, they have been reporting on PART

---

<sup>15</sup>DOE Office of Science also has a leading role in an international, informal professional organization—the Washington Research Evaluation Network, at <http://www.wren-network.net/>—exploring evaluation approaches for improving the management of public science and technology programs (Oct. 21, 2005).

---

recommendations to OMB twice a year, tracking approximately 100 PART recommendations (with about 200 separate milestones) for the 62 programs reviewed for fiscal years 2004 through 2006.

Most of the officials we interviewed believed that because of PART and the President's Management Agenda, their agencies were paying greater attention to program results and evaluation. Officials at DOL noted that the department spends much time and effort making sure it scores green on the next President's Management Agenda assessment; for example, the department's management review board, chaired by Labor's Assistant Secretary for Management and Administration, discusses these issues monthly. In addition, DOL's Center for Program Planning and Results reviews programs' progress on OMB's recommendations, scores programs internally on the Budget and Performance Integration scorecard, and provides agencies with training and preparation before their PART reviews. The SBA Administrator initiated a series of steps after August 2003 to increase the agency's focus on achieving results. SBA rewrote its strategic plan to focus on a limited number of strategic goals and integrated its strategic plan, annual performance plan, and performance report. The agency formed a central Office of Analysis, Planning, and Accountability to help each program office develop results-oriented performance measures and conduct program assessments.

Although HHS officials said that the department had invested in evaluation long before the PART reviews, Indian Health Service program officials indicated that they had not planned an evaluation of their sanitation facilities program before the PART review. However, they thought it was a good idea and said that the recommendation brought their lack of a recent evaluation to HHS's attention, making it easier to justify efforts to quantify their program's benefits.

---

### Centralized Coordination Helped Agencies Leverage Their Evaluation Resources

SBA and DOL responded to demands for more performance information by centrally coordinating their assessment activities, helping to address evaluation's measurement and funding challenges. Centralization helped the agencies to leverage their evaluation expertise throughout the agency and helped them prioritize spending on the evaluations they considered most important.

SBA program offices had little experience with outcome measurement and evaluation before the 2002 PART reviews. The central planning office was formed to help the program offices develop outcome measures linked to the agency's strategic goals and collect and validate their performance

---

data. The office also conducts an annual staff activity survey to support cost allocation across programs, a key step toward performance budgeting. This office took advantage of the similarity in outcome goals across SBA's programs and the evaluation methodology developed for the counseling programs to contract for the development of a standard methodology for assessing other SBA programs' economic impacts on small businesses. The central office is also funding the subsequent evaluations. For a small agency, this type of coordination can result in important savings in contract resources as well as staff time.

DOL, much larger than SBA, has measurement and evaluation experience, but capacity had declined over time. DOL established the Center for Program Planning and Results in 2001 to provide leadership, policy advice, and technical assistance to GPRA-related strategic and performance planning. The center was expanded in fiscal year 2003 to respond to the President's Management Agenda and manage the PART process. With a budget of \$5 million a year, the center solicits and selects evaluation proposals focusing on program effectiveness submitted by DOL's component agencies, funds the studies, and helps oversee the external contractors. The center's officials claimed that the Secretary's and Assistant Secretary's support for evaluation, combined with pressure from OMB, has led to increased interest by the component agencies in evaluation, resulting in \$6 million to \$7 million in proposals competing for \$5 million in evaluation funds. Some DOL agencies retained their evaluation expertise and design, fund, and oversee their own evaluations. In addition to helping program offices develop research questions and evaluation designs, the center helps develop agency evaluation capacity by holding "Vendor Days," when evaluation contractors are invited to exhibit for agency staff the specialized design, data collection, and analysis skills that could inform future studies.

---

## Where OMB and Program Managers Do Not Share Expectations, Evaluations May Not Meet OMB Needs

Because the OMB evaluation recommendations were fairly general, agencies had flexibility in interpreting the information OMB expected and the evaluations to fund. Some program managers disagreed with OMB on the scope and purpose of their evaluations, their quality, and the usefulness of evaluations conducted by independent third parties. Program managers concerned about an increased focus on process said that they were more interested in learning how to improve program performance than in meeting an OMB checklist. Since a few programs did not discuss their evaluation plans with OMB, it is not certain whether OMB will accept their ongoing evaluations.

---

## Agencies Have Flexibility in Determining Evaluation Timing and Content

Agencies had a fair amount of flexibility to design their evaluations. Except for the recommendations to the DOE Office of Science to conduct committee of visitors reviews, OMB's evaluation recommendations were fairly general, typically telling agencies to conduct an independent evaluation of a program's effectiveness. Agencies reported little guidance from OMB on how to conduct these evaluations, beyond the PART written guidance and the rationale the examiner provided for not accepting their previous evaluations or measures of program outcomes. They said that follow-up on previous PART recommendations was generally limited to providing responses to the OMB reporting template, unless OMB conducted a second formal PART review.

Agencies also had flexibility to determine the timing of their evaluations. Agency officials reported that OMB did not prioritize its recommendations within or among programs. Moreover, because evaluation resources were limited, DOL and SBA officials reported that they had to choose which evaluations to conduct first. The recommendations for the two DOL regulatory agencies explicitly acknowledged their need to balance responsibility for several programs. OMB asked these agencies to develop plans to evaluate their programs or expand existing efforts for more comprehensive and regular evaluation. In the reviews of recommendation status for the fiscal year 2006 budget, OMB credited both agencies with having conducted one or more program reviews and planning others. Agencies were free to choose which programs to evaluate but were likely to be influenced by the potential effect of PART reassessments on their President's Management Agenda scores and, thus, to attempt to reduce the number of programs rated "results not demonstrated." Research and development programs were held to a somewhat higher standard than other programs were, since their agencies could not be scored "green" on the separate R&D Investment Criteria Initiative if less than 75 percent of their programs received a score of "moderately effective" or better. DOE officials noted that their Office of Energy Efficiency and Renewable Energy now requires programs to outline their plans for evaluations in their multiyear plans.

---

## Disagreements on the Scope and Purpose of Evaluation

OMB and the agencies significantly differed in defining evaluation scope and purpose. Program officials were frustrated by OMB's not accepting their prior evaluations of program effectiveness in the PART review. Some of the difficulties seemed to derive from OMB expecting to find, in the agencies' external evaluation studies, comprehensive judgments about program design, management, and effectiveness, like the judgments made in the OMB PART assessments.

---

PART's criteria for judging the adequacy of agency evaluations are complex and may have created some tension as to the importance of one dimension over another. For example, question 2.6 read: "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?" OMB changed the wording of the question to help clarify its meaning and added the reference to "relevance." However, while OMB's revised guidance for this question defines quality, scope, and independence, it does not address the assessment of program "relevance." Specifically, sufficient scope is defined as whether the evaluation focuses on achievement of performance targets and the cause and effect relationship between the program and target—i.e., program effectiveness. This is different from assessing the relevance—i.e., appropriateness—of the program design to the problem or need. Instead, questions in section 1 ask whether the design is free of major flaws and effectively targeted to its purpose.

Another potential contribution to differences between OMB and agency expectations for program evaluations is that evaluations designed for internal audiences often have a different focus than evaluations designed for external audiences. Evaluations that agencies initiate typically aim to identify how to improve the allocation of program resources or the effectiveness of program activities. Studies requested by program authorizing or oversight bodies are more likely to address external accountability—to judge whether the program is properly designed or is solving an important problem.

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

---

applying the standards for independent evaluation, but these standards were set extremely high.

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

DOE Office of Science officials described much discussion with OMB examiners about what was or was not a good committee of visitors review in following up on the status of the evaluation recommendations. Although OMB had revised and extended its guidance on what constituted quality in evaluation, program officials still found this guidance difficult to apply to research programs. They also acknowledged that their first committee of visitors reviews might have been more useful to the program than to OMB.

---

## Disagreements about the Quality of Evaluation Designs

OMB and agencies differed in identifying which evaluation methods were sufficiently rigorous to provide high-quality information on program effectiveness. OMB guidance encouraged the use of randomized controlled trials, or experiments, to obtain the most rigorous evidence of program impact but also acknowledged that these studies are not suitable or feasible for every program. However, as described above, without guidance on which—and when—alternative methods were appropriate, OMB and agency staff disagreed on whether specific evaluations were of acceptable quality. To help develop shared understandings and expectations, federal evaluation officials and OMB staff held several discussions on how to assess evaluation quality according to the type of program being evaluated.

When external factors such as economic or environmental conditions are known to influence a program’s outcomes, an impact evaluation attempts

---

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

Randomized experiments are best suited to studying programs that are clearly defined interventions that can be standardized and controlled, and limited in availability, and where random assignment of participants and nonparticipants is deemed feasible and ethical. Quasi-experimental designs are also best suited to clearly defined, standardized interventions with limited availability, and where one can measure, and thus control for, key plausible alternative explanations for observed outcomes. In mature full-coverage programs where comparison groups cannot be obtained, program effects may be estimated through systematic observation of targeted measures under specially selected conditions designed to eliminate plausible alternative explanations for observed outcomes.<sup>16</sup>

Following our January 2004 report recommendation that OMB better define an "independent, quality evaluation," OMB revised and expanded its guidance on evaluation quality for the fiscal year 2006 PART reviews. The guidance encouraged the use of randomized controlled trials as particularly well suited to measuring program impacts but acknowledged that such studies are not suitable or feasible for every program, so it recommended that a variety of methods be considered. OMB also formed

---

<sup>16</sup> For further discussion see Peter H. Rossi, Howard E. Freeman, and Mark W. Lipsey, *Evaluation: A Systematic Approach*, 6th ed. (Thousand Oaks, Calif.: Sage Publications, 1999). For additional examples of alternative evaluation designs, see GAO, *Program Evaluation: Strategies for Assessing How Information Dissemination Contributes to Agency Goals*, [GAO-02-923](#) (Washington, D.C.: Sept. 30, 2002).

---

an Interagency Program Evaluation Working Group in the summer of 2004 to provide assistance on evaluation methods and resources to agencies undergoing a PART review that discussed this guidance extensively. Evaluation officials from several federal agencies expressed concern that the OMB guidance materials defined the range of rigorous evaluation designs too narrowly. In the spring of 2005, representatives from several federal agencies participated in presentations about program evaluation purposes and methods with OMB examiners. They outlined the types of evaluation approaches they considered best suited for various program types and questions (see table 2).<sup>17</sup> However, OMB did not substantively revise its guidance on evaluation quality for the fiscal year 2007 reviews beyond recommending that “agencies and OMB should consult evaluation experts, in-house and/or external, as appropriate, when choosing or vetting rigorous evaluations.”<sup>18</sup>

---

<sup>17</sup> The entire evaluation dialogue presentation is at <http://www.epa.gov/evaluate/part.htm> (Oct. 21, 2005).

<sup>18</sup> Office of Management and Budget, *Guidance for Completing the Program Assessment Rating Tool (PART)*. (Washington, D.C.: March 2005) is at <http://www.whitehouse.gov/omb/part> (Oct. 21, 2005).

**Table 2: Federal Evaluators' Views on Tailoring Designs for Program Effectiveness Evaluations**

Typical designs used to assess program effectiveness	Design features that help control for alternative explanations	Best suited for (typical examples)
Process and outcome monitoring or evaluation	Compares performance to pre-existing goal or standard. For example: <ul style="list-style-type: none"> <li>• OMB R&amp;D criteria of relevance, quality and performance.</li> <li>• Productivity, cost effectiveness and efficiency standards.</li> </ul>	Research, enforcement, information and statistical programs, and business-like enterprises with <ul style="list-style-type: none"> <li>• few, if any, alternative explanations for observed outcomes.</li> <li>• ongoing programs producing goods and services</li> <li>• complete national coverage</li> </ul>
Quasi-experiments – Single Group	Compare outcomes for program participants or entities before and after the intervention. <ul style="list-style-type: none"> <li>• Multiple data points over time are necessary.</li> <li>• Control for alternative explanations by statistical adjustments and analyses such as modeling.</li> </ul>	Regulatory and other programs with <ul style="list-style-type: none"> <li>• clearly defined interventions with distinct starting times</li> <li>• complete national coverage</li> <li>• random assignment of participants or entities to groups is NOT feasible, practical, or ethical.</li> </ul>
Quasi-experiments – Comparison Groups	Compares outcomes for program participants or entities with outcomes for a comparison group selected to closely match the “treatment” group on key characteristics. <ul style="list-style-type: none"> <li>• Key characteristics are plausible alternative explanations for the outcome.</li> <li>• Measure outcomes before and after intervention (pretest, posttest).</li> </ul>	Service and other programs with <ul style="list-style-type: none"> <li>• clearly defined interventions that can be standardized and controlled</li> <li>• limited national coverage</li> <li>• random assignment of participants or entities to groups is NOT feasible, practical, or ethical.</li> </ul>
Randomized experiments	Compares outcomes for randomly assigned program (treatment) participants or entities with outcomes for a randomly assigned “control” group prior to intervention. <ul style="list-style-type: none"> <li>• Measure outcomes before and after intervention (pretest, posttest).</li> </ul>	Service and other programs with <ul style="list-style-type: none"> <li>• clearly defined interventions that can be standardized and controlled</li> <li>• limited national coverage</li> <li>• random assignment of participants or entities to groups is feasible and ethical.</li> </ul>

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

### Disagreements about Requiring Independent Third-Party Evaluations

A related source of tension between OMB and agency evaluation interests was the importance of an evaluation’s independence. PART guidance stressed that for evaluations to be independent, nonbiased parties with no conflict of interest, for example, GAO or an Inspector General, should conduct them. OMB subsequently revised the guidance to allow evaluations to be considered independent if the program contracted them out to a third party or they were carried out by an agency’s program evaluation office. However, disagreements continued on the value and importance of this criterion.

---

HHS officials reported variation among examiners in whether their evaluations were considered independent. Two programs objected to OMB examiners' claims that an evaluation was not independent if the agency paid for it. OMB changed the fiscal year 2005 PART guidance to recognize evaluations contracted out to third parties and agency program evaluation offices as possibly being sufficiently independent, subject to examination case by case. But HHS officials claimed that they were still having issues with the independence standard in the fiscal year 2006 reviews and that OMB's guidance was not consistently followed from one examiner to the next.

DOL program officials stated that using an external evaluator who was not familiar with the program resulted in an evaluation that was not very useful to them. In part, this was because program staff were burdened with educating the evaluator. But more important, they claimed that the contractor designed the scope of the work to the broad questions of PART (such as questions on program mission) rather than focusing on the results questions the program officials wanted information on. In combination, this led to a relatively superficial program review, in their view, that provided the external, independent review OMB wanted but not the insights the program managers wanted.

In reviewing the status of its PART recommendations, OMB did not accept advisory committee reviews for two research programs that DOE offered in response because OMB did not perceive the reviews as sufficiently independent. These two program reviews involved standing advisory committees of approximately 50 people who review the programs every 3 years. The OMB examiner believed that the committee was not truly independent of the agency. DOE program officials objected, noting the committee's strong criticisms of the program, but have reluctantly agreed to plan for an external review by the National Academies. Program officials expressed concern that because evaluators from the National Academies may not be sufficiently familiar with their program and its context, such reviews may not address questions of interest to them about program performance.

HHS program officials were also concerned about the usefulness of an evaluation of the sanitation facilities program if it was conducted by a university-based team inexperienced with the program. The agency deliberately guarded against this potential weakness by including two ex-agency officials (one an engineer) on the evaluation team, and by taking considerable effort with the team to define the evaluation questions.

---

## Agencies Not Consulting with OMB on Evaluation Plans May Not Meet OMB's Expectations

Agencies' freedom to design their evaluations, combined with differences in expectations between agencies and OMB, raises the strong possibility that the evaluations that agencies conduct may not provide OMB with the information it wants. Most of the agency officials we interviewed said that they had discussed their evaluation plans with their OMB examiners, often as part of their data collection review process. SBA and DOL, in particular, appeared to have had extensive discussions with their OMB examiners. However, a few programs have not discussed their plans with OMB, presumably on the assumption that they will meet OMB's requirements by following its written guidance.

Officials in SBA's and DOL's central planning offices described extensive discussions of their evaluation plans with their OMB examiners. SBA vetted the evaluation design for SBA's counseling programs with OMB in advance, as well as the questionnaire used to assess client needs. DOL planning and evaluation officials noted that they had worked with OMB examiners to moderate their expectations for agencies' evaluations. They said that OMB understands their "real world" financial constraints and is allowing them to "chip away" at their outcome measurement issues and not conduct net impact evaluations in program areas where they do not have adequate funds to do this type of evaluation.

HHS program officials were concerned about whether OMB will accept their ongoing evaluation of the immunization program when they receive their next PART review. The evaluation recommendation was general, so they based their design on the fiscal year 2004 criteria and to provide information useful to the program. However, the officials had heard that the fiscal year 2007 evaluation quality criteria were more rigid than those previously used, so they were concerned about whether the program will meet OMB's evaluation criteria when it is reviewed again. They said they would have liked OMB to consider its evaluation progress and findings so far and to have given them input as to whether the evaluation will meet the current criteria. OMB officials denied that the PART criteria for evaluation quality had changed much in the past two years. They also expected, from their review of the design, that this new evaluation would meet current PART criteria, assuming it was carried out as planned.

Several program officials expressed the view that in designing their evaluations, they were more concerned with learning how to improve their programs than in meeting an OMB checklist. Program officials complained that OMB's follow-up on whether evaluations were being planned sent the message that OMB was more interested in checking off boxes than in having a serious discussion about achieving results. When one program

---

official was asked for the program’s new evaluation plan, he answered “Who needs a plan? I’ve got an evaluation.” DOE program officials indicated that they believe a comprehensive evaluation of Weatherization Assistance should include all the questions that state, regional, and local officials would like to ask and not just establish a new national energy savings estimate. Those questions—also of interest to DOE—include: Which weatherization treatments correlate with energy savings? Should they use their own crews or hire contractors? What are the nonenergy benefits, such as improved air quality or employment impacts? Program officials indicated that they had conducted a great deal of planning and discussion with their stakeholders over the past 5 to 6 months and expect to conduct five or six studies to meet those needs.

---

## Conclusions

The PART review process has stimulated agencies to increase their evaluation capacity and available information on program results. The systematic examination of the array of evidence available on program performance has helped illuminate gaps and has helped focus evaluation questions. The public visibility of the results of the PART reviews has brought management attention to the development of agency evaluation capacity.

Evaluations are useful to specific decision makers to the degree that the evaluations are credible and address their information needs. Agencies are likely to design evaluations to meet their own needs—that is, in-depth analyses that inform program improvement. If OMB wants evaluations with a broader scope, such as information that helps determine a program’s relevance or value, it will need to take steps to shape both evaluation design and execution.

Because agency evaluation resources tend to be limited, they are most usefully focused on illuminating important areas of uncertainty. While regular performance reporting is key to good program management and oversight, requiring all federal programs to conduct frequent evaluation studies is likely to result in many superficial reviews that will have little utility and that will overwhelm agency evaluation capacity.

---

## Recommendations for Executive Action

In light of our findings and conclusions in this report, we are making the following recommendations to OMB reiterating and expanding on recommendations in our previous report:

---

OMB should encourage agencies to discuss their plans for program evaluations—especially those in response to an OMB recommendation—with OMB and with congressional and other program stakeholders to ensure that their findings will be timely, relevant, and credible and that they will be used to inform policy and management decisions.

OMB should engage in dialogue with agencies and congressional stakeholders on a risk-based allocation of scarce evaluation resources among programs, based on size, importance, or uncertain effectiveness, and on the timing of such evaluations.

OMB should continue to improve its PART guidance and training of examiners on evaluation to acknowledge a wide range of appropriate methods.

---

## Agency Comments

We provided a draft of this report to OMB and the agencies for review and comment. OMB agreed that evaluation methodology should be appropriate to the size and nature of the program and that randomized controlled trials may not be valuable in all settings. It noted its intent to provide additional guidance in this area. OMB disagreed with the reference to the PART as a checklist. This view was not ours but the view of agency officials who expressed concern about the focus of the assessment process. OMB also provided a number of technical comments, which we incorporated as appropriate throughout the report. OMB's comments appear in appendix III. We also received technical comments from DOE, DOL, and HHS that we incorporated where appropriate throughout the report. SBA had no comments.

---

We are sending copies of this report to the Director of the Office of Management and Budget; the Secretaries of Energy, Labor, and Health and Human Services; the Administrator of the Small Business Administration; appropriate congressional committees; and other interested members of Congress. We will also make copies available to others on request. In addition, the report will be available at no charge on GAO's Web site at <http://www.gao.gov>.

---

If you or your staff have questions about this report, please contact me at (202) 512-2700 or KingsburyN@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made key contributions to this report are listed in appendix IV.

Sincerely,

A handwritten signature in black ink that reads "Nancy L. Kingsbury". The signature is written in a cursive style with a large, stylized initial 'N'.

Nancy Kingsbury  
Managing Director  
Applied Research and Methods

# Appendix I: Agency Programs OMB Recommended Evaluations For in PART Reviews

Agency	Program	Program type	OMB recommendation
DOE	Advanced Fuel Cycle Initiative	R&D	Establish plans for periodic independent evaluations to assess program progress and recommend program improvements
	Advanced Scientific Computing Research	R&D	Institute formal committee of visitors process by September 2003
	Generation IV Nuclear Energy Systems Initiative	R&D	Develop a plan for independent program evaluations to guide program managers and policy decision makers
	High Energy Physics	R&D	Institute formal committee of visitors process by September 2003
	Nuclear Energy Research Initiative	R&D	Will plan independent program evaluations to guide program management and development
	Nuclear Physics	R&D	Institute formal committee of visitors process by September 2003
	Weatherization Assistance	Block/formula grants	Recommends periodic independent evaluation of the program's cost-effectiveness
DOL	Employee Benefits Security Administration	Regulatory	Expand existing efforts for more comprehensive and regular program evaluation
	Federal Employees Compensation Act	Direct federal	An evaluation of strategic goals, the success of various program strategies, and state/industry best practices
	Occupational Safety and Health Administration	Regulatory	Develop a plan to evaluate the results and cost-effectiveness of its regulatory and nonregulatory programs
	Office of Federal Contract Compliance Programs	Regulatory	Complete in 2003 an external evaluation and staff analysis to measure and improve program performance
	Youth Activities	Direct federal	Plan and conduct an impact evaluation
HHS	317 Immunization Program	Competitive grants	Conduct a comprehensive evaluation of the structure, management, and operations of the immunization program
	Indian Health Service Sanitation Facilities Construction Program	Capital assets	Conduct an independent, comprehensive evaluation of the program
	Nursing Education Loan Repayment and Scholarship Program	Competitive grants	Evaluate impact, develop outcome measures, and track performance
	Refugee and Entrant Assistance	Block/formula grants	The budget includes funds for ORR to conduct independent and quality evaluations
SBA	Business Information Centers	Direct federal	Undertake an evaluation of the program's effectiveness and measure whether it duplicates other federal and nonfederal mentoring programs
	SCORE	Block/formula grant	Undertake an evaluation of the program's effectiveness and measure whether it duplicates other federal and nonfederal mentoring programs
	Section 504 Certified Development Company Loan program	Credit	The 2004 budget proposes to increase program evaluations to determine the factors that affect both demand and performance in the 504 and 7(a) programs

**Appendix I: Agency Programs OMB  
Recommended Evaluations For in PART  
Reviews**

<b>Agency</b>	<b>Program</b>	<b>Program type</b>	<b>OMB recommendation</b>
	Small Business Development Centers	Block/formula grants	Undertake an evaluation of the program's effectiveness and measure whether it duplicates other federal and nonfederal mentoring programs

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

Note: OMB = Office of Management and Budget, DOE = Department of Energy, R&D = Research and Development, DOL = Department of Labor, HHS = Department of Health and Human Services, ORR = Office of Refugee Resettlement, SBA = Small Business Administration.

---

# Appendix II: Related Agency Program Evaluation Reports

---

---

## Department of Energy Agency Reports

**Advanced Fuel Cycle Initiative:** Nuclear Energy Research Advisory Committee (NERAC) Evaluation Subcommittee. *Evaluation of DOE Nuclear Energy Programs*. Washington, D.C.: Sept. 10, 2004.

**Advanced Scientific Computing Research Program:** Advanced Scientific Computing Research. *Committee of Visitors Report*. Washington, D.C.: April 2004.

**Generation IV Nuclear Energy Systems Initiative:** Nuclear Energy Research Advisory Committee (NERAC) Evaluation Subcommittee. *Evaluation of DOE Nuclear Energy Programs*. Washington, D.C.: Sept. 10, 2004.

**High Energy Physics Program:** Committee of Visitors to the Office of High Energy Physics. *Report to the High Energy Physics Advisory Panel*. Washington, D.C.: Apr. 7, 2004.

**Nuclear Physics Program:** Committee of Visitors. *Report to the Nuclear Science Advisory Committee*. Washington, D.C.: Department of Energy, Office of Science, Feb. 27, 2004.

---

## Department of Health and Human Services Agency Reports

**317 Immunization Program:** RTI International. *Section 317 Grant Immunization Program Evaluation: Findings from Phase I*. Draft progress report. Atlanta, Ga.: Centers for Disease Control and Prevention, January 2005.

**Indian Health Service Sanitation Facilities Program:** Department of Health and Human Services, U.S. Public Health Service, Federal Occupational Health Service. *Independent Evaluation Report Summary*. Prepared for Indian Health Service Sanitation Facilities Construction Program, Rockville, Maryland. Seattle, Wash.: Mar. 8, 2005.

**Nursing Education Loan Repayment and Scholarship Program:** Department of Health and Human Services, Health Resources and Services Administration, Bureau of Health Professions. *HRSA Responds to the Nursing Shortage: Results from the 2003 Nursing Scholarship Program and the Nursing Education Loan Repayment Program: 2002–2003*. First report to the United States Congress. Rockville, Md.: n.d.

---

Department of Labor  
Agency Reports

**Employee Benefits Security Administration Reports:**

- Mathematica Policy Research, Inc. *Case Opening and Results Analysis (CORA) Fiscal Year 2002: Final Report*. Washington, D.C.: Mar. 31, 2004.
- Royal, Dawn. *U.S. Department of Labor, Employee Benefits Security Administration: Evaluation of EBSA Customer Service Programs Participant Assistance Program Customer Evaluation*. Washington, D.C.: The Gallup Organization, February 2004.
- Royal, Dawn. *U.S. Department of Labor, Employee Benefits Security Administration: Evaluation of EBSA Customer Service Programs Participant Assistance Mystery Shopper Evaluation*. Washington, D.C.: The Gallup Organization, January 2004.
- Royal, Dawn. *U.S. Department of Labor, Employee Benefits Security Administration: Evaluation of EBSA Customer Service Programs Participant Assistance Outreach Programs Evaluation*. Washington, D.C.: The Gallup Organization, January 2004.
- Royal, Dawn. *U.S. Department of Labor, Employee Benefits Security Administration: Evaluation of EBSA Customer Service Programs Participant Assistance Web Site Evaluation*. Washington, D.C.: The Gallup Organization, January 2004.

**Federal Employees Compensation Act Program:** ICF Consulting. *Federal Employees Compensation Act (FECA): Program Effectiveness Study*. Fairfax, Va.: U.S. Department of Labor, Office of Workers' Compensation Programs, Mar. 31, 2004.

**Office of Federal Contract Compliance Programs:** Westat. *Evaluation of Office of Federal Contract Compliance Programs: Final Report*. Rockville, Md.: December 2003.

**Occupational Safety and Health Administration Reports:**

- ERG. *Evaluation of OSHA's Impact on Workplace Injuries and Illnesses in Manufacturing Using Establishment-Specific Targeting of Interventions*. Final report. Lexington, Mass.: July 23, 2004.
- Marker, David and others. *Evaluating OSHA's National and Local Emphasis Programs*. Draft Final Report for Quantitative Analysis of Emphasis Programs. Rockville, Md.: Westat, Dec. 24, 2003.

- 
- OSHA, Directorate of Evaluation and Analysis. *Regulatory Review of OSHA's Presence Sensing Device Initiation (PSDI) Standard [29 CFR 1910.217(h)]*. Washington, D.C.: May 2004.  
[www.osha.gov/dcsp/compliance\\_assistance/lookback/psdi\\_final2004.html](http://www.osha.gov/dcsp/compliance_assistance/lookback/psdi_final2004.html) (Oct. 21, 2005).

# Appendix III: 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

October 17, 2005

Ms. Nancy Kingsbury  
Managing Director  
Applied Research and Methods  
Government Accountability Office  
441 G Street, NW  
Washington, DC 20548

Dear Ms. Kingsbury:

Thank you for the opportunity to comment on the draft GAO report on program evaluation (*Program Evaluation: OMB's PART Reviews Increased Agencies' Attention to Improving Evidence of Program Results*, GAO-06-67).

We appreciate GAO's continued interest in the Program Assessment Rating Tool (PART) and our determination to use it to assess Federal programs in a consistent fashion. As is acknowledged in your report, the President's Management Agenda and the PART have raised awareness of the importance of program results.<sup>1</sup> Indeed, we believe that without measurement, there is no certainty about how Federal programs are performing for the American people.

Oftentimes, Federal programs proceed to carry out their business based on the assumption that their existing strategies and approaches are effective. High quality, independent evaluation may validate or disprove that assumption. Your draft report suggests that randomized controlled trials may not be valuable in every setting. OMB agrees and recognizes that the size and nature of the program should be considered when determining the appropriate evaluation methodology. We have tried to make this point clear and will provide additional guidance in this area.

The report suggests that the PART requires too high a standard for evaluations and discards other useful evaluations. The evaluation question in Section II of the PART, the primary subject of this report, focuses on evaluations of program impact which must rely on rigorous methodologies to be credible; however, OMB's guidance makes clear that the methodology should be appropriate for the program. It is also important to note that other evaluations, such as those assessing processes or smaller aspects of a program, provide valuable information and are regularly used as evidence for other questions in the PART, particularly in Section III.

The report also includes a few references to the PART as an OMB "checklist."<sup>2</sup> Sections I – III of the PART essentially are a checklist – one to ensure that the program has key elements

<sup>1</sup> See, draft report "Management Attention, Caught by OMB's Recommendations, Overcame Measurement and Funding Barriers," p. 12.

<sup>2</sup> See, for example, the report "Highlights."

---

**Appendix III: Comments from the Office of  
Management and Budget**

---

in place which are generally understood to make programs effective, such as a clear, focused mission; thoughtful strategic planning; and strong management practices. Section IV, on the other hand, measures whether these conditions and practices have produced results. In addition, after each PART assessment, programs commit to an improvement plan, key follow-up actions they will take to achieve performance results. PART is more than just a checklist.

The particular interest that GAO has taken in the Administration's standards for measuring performance is noteworthy. Thank you for your continued enthusiasm about program evaluation and the PART, as well as for 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 making the Federal Government more results-oriented.

Sincerely,



Clay Johnson III

---

# Appendix IV: GAO Contact and Staff Acknowledgments

---

## GAO Contact

Nancy Kingsbury (202) 512-2700 or [KingsburyN@gao.gov](mailto:KingsburyN@gao.gov)

---

## Acknowledgments

In addition to the contact named above, Stephanie Shipman, Assistant Director, and Valerie Caracelli made significant contributions to this report. Denise Fantone and Jacqueline Nowicki also made key contributions.

---

# Related GAO Products

---

*Performance Budgeting: PART Focuses Attention on Program Performance, but More Can Be Done to Engage Congress.* [GAO-06-28](#). Washington, D.C.: Oct. 28, 2005.

*Managing for Results: Enhancing Agency Use of Performance Information for Managerial Decision Making.* [GAO-05-927](#). Washington, D.C.: Sept. 9, 2005.

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

*Performance Measurement and Evaluation: Definitions and Relationships.* [GAO-05-739SP](#). Washington, D.C.: May 2005.

*Results-Oriented Government: GPRA Has Established a Solid Foundation for Achieving Greater Results.* [GAO-04-38](#). Washington, D.C.: Mar. 10, 2004

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

*Program Evaluation: An Evaluation Culture and Collaborative Partnerships Help Build Agency Capacity.* [GAO-03-454](#). Washington, D.C.: May 2, 2003.

*Program Evaluation: Strategies for Assessing How Information Dissemination Contributes to Agency Goals.* [GAO-02-923](#). Washington, D.C.: Sept. 30, 2002.

*Program Evaluation: Studies Helped Agencies Measure or Explain Program Performance.* [GAO/GGD-00-204](#). Washington, D.C.: Sept. 29, 2000.

*Performance Plans: Selected Approaches for Verification and Validation of Agency Performance Information.* [GAO/GGD-99-139](#). Washington, D.C.: July 30, 1999.

*Managing for Results: Measuring Program Results That Are Under Limited Federal Control.* [GAO/GGD-99-16](#). Washington, D.C.: Dec. 11, 1998.

---

## GAO's Mission

The Government Accountability 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 GAO's Web site ([www.gao.gov](http://www.gao.gov)). Each weekday, GAO posts newly released reports, testimony, and correspondence on its Web site. To have GAO e-mail you a list of newly posted products every afternoon, go to [www.gao.gov](http://www.gao.gov) and select "Subscribe to Updates."

---

## 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. Government Accountability 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](http://www.gao.gov/fraudnet/fraudnet.htm)

E-mail: [fraudnet@gao.gov](mailto:fraudnet@gao.gov)

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

---

## Congressional Relations

Gloria Jarmon, Managing Director, [JarmonG@gao.gov](mailto:JarmonG@gao.gov) (202) 512-4400  
U.S. Government Accountability Office, 441 G Street NW, Room 7125  
Washington, D.C. 20548

---

## Public Affairs

Paul Anderson, Managing Director, [AndersonP1@gao.gov](mailto:AndersonP1@gao.gov) (202) 512-4800  
U.S. Government Accountability Office, 441 G Street NW, Room 7149  
Washington, D.C. 20548