

*WHICH FEDERAL AGENCIES
BEST INFORM THE PUBLIC?*

8th ANNUAL
PERFORMANCE
REPORT SCORECARD

MAURICE MCTIGUE

HENRY WRAY

JERRY ELLIG

APRIL 2007

MERCATUS CENTER
GEORGE MASON UNIVERSITY

ABOUT THE MERCATUS CENTER AT GEORGE MASON UNIVERSITY

The Mercatus Center at George Mason University is a 501(c)(3) education, research, and outreach organization that works with scholars, policy experts, and government officials to bridge academic learning and real-world practice.

Our mission is to generate knowledge and understanding of how institutions affect the freedom to prosper and hold organizations accountable for their impact on that freedom. The aim of our work is to enable individuals to live free, prosperous, and peaceful lives.

The Mercatus Center is located on George Mason University's Arlington Campus, along with the George Mason University School of Law, the Law and Economics Center, and our sister organization, the Institute for Humane Studies.

ABOUT THE GOVERNMENT ACCOUNTABILITY PROJECT

Created in 1997, the Mercatus Center's Government Accountability Project uses a variety of tools to help improve the public sector management process, by which government decision makers identify and quantify the specific public benefits their actions and agencies generate. The program's goal is to help improve government funding and policy decisions by making transparent these benefits produced with citizens' resources. Full transparency brings praise and criticism of results — and eventually change — based upon maximizing outcomes and minimizing expenditures. Mercatus Center scholars use consulting opportunities, written analysis, testimony, educational programs, opinion pieces, and other products (such as the *Annual Performance Report Scorecard*) to affect the management of government agencies and to generate quality information to demonstrate the effectiveness of agencies.

For more information about the Mercatus Center and the Government Accountability Project, visit our adjoining Web sites (www.mercatus.org and www.GovernmentAccountability.org) or you may telephone (703) 993-4930.

The analysis, interpretations, and conclusions in this study are those of the authors and the research team and are not the official positions of the Mercatus Center or George Mason University.

EXECUTIVE SUMMARY

The Government Performance and Results Act of 1993 requires agencies to produce strategic plans, annual performance plans, and annual performance reports. Performance reporting started in fiscal 1999.

Researchers at the Mercatus Center at George Mason University initiated this Scorecard in fiscal 1999 to foster continuous improvement in the quality of disclosure in agencies' annual performance reports. This study is our eighth annual evaluation of the performance and accountability reports produced by the 24 agencies covered under the Chief Financial Officers Act. These agencies accounted for 99 percent of federal outlays in fiscal 2006. We employed the same criteria used in previous Scorecards. Our scoring process evaluates (1) how *transparently* an agency discloses its successes and failures; (2) how well an agency documents the tangible *public benefits* it claims to have produced; and (3) whether an agency demonstrates *leadership* that uses annual performance information to devise strategies for improvement. An expert team evaluated each agency's report on 12 criteria—four each for transparency, public benefits, and leadership.

By assessing the quality of agencies' reports, but *not* the quality of the results achieved, we seek to learn which agencies are supplying the information that citizens and their elected leaders need to make informed funding and policy decisions.

Key findings in this year's Scorecard include:

UPSET AT THE TOP. For the first time since fiscal 2001, Transportation is back at the top of the rankings, earning 53 out of a possible 60 points. Labor and Veterans Affairs tied for second place, with 51 points apiece. State, the number 2 report last year, remained in the top four with a 50.

BIGGEST SPENDERS DISCLOSE LESS. Agencies with reports receiving average scores below the satisfactory level accounted for 87 percent of non-interest federal spending in fiscal 2006. The 10 reports receiving a satisfactory score (36 or better out of 60 possible points) accounted for 13 percent of non-interest federal spending in fiscal 2006, down from 15 percent in fiscal 2005.

IMPROVEMENT AT THE TOP. For fiscal 2006, the top four reports scored above 48, versus just two in fiscal 2005. These agencies account for eight percent of non-interest spending.

QUALITY GAP. A significant quality gap emerged between the top four reports and the rest. Just three points separated the top three reports. The fifth place report, USAID's, earned 42 points—eight less than the fourth place report.

SUBSTANTIAL CHURN. Eleven reports improved their scores in fiscal 2006 compared to 2005; 11 had lower scores; and two were unchanged. Reports that had significantly improved rankings include GSA (+11 places in the ranking), Education (+9), Social Security (+6), EPA (+5), Defense (+5), and NASA (+4). Reports whose rankings significantly declined include SBA (-11), Treasury (-6), HUD (-6), OPM (-6), Interior (-5), and Energy (-4).

AVERAGE SCORE STAGNANT. The average total score has remained at about 36 for the past three years.

WEIGHTED AVERAGE SCORE SHOWS IMPROVEMENT. Since the size of federal agencies varies greatly, we also calculate a "weighted average" score, which weights each report's score by its agency's spending. Some of the larger agencies improved their reports, raising the weighted average score from 30.3 in fiscal 2005 to 32.2 in fiscal 2006.

TWO CRITERIA IMPROVE SIGNIFICANTLY. Scores on two criteria improved by more than five percent. Scores on criterion 1, accessibility, improved because most agencies have now mastered the mechanics of posting reports on their Web sites in a visible place and user-friendly format. Scores on criterion 8, linkage of results to cost, improved because more reports are using the best practices established over the past several years.

ONE CRITERION DETERIORATES NOTICEABLY. Scores on criterion 9, which assesses whether the report shows how the agency makes this country a better place to live, fell by five percent because few agencies have kept pace with the new best practice of backing up narratives with substantive performance metrics.

SUBSTANTIAL ROOM FOR IMPROVEMENT. Average scores on six of our 12 criteria are still below 3 (out of a possible 5), suggesting there is still substantial room for improvement.

UP FROM THE ASHES. Three reports that ranked poorly in recent years improved their rankings noticeably in fiscal 2006: Defense (16th), NASA (16th), and Social Security (15th).

MIREN IN THE MUD. Three reports that often rank poorly continued that tradition in fiscal 2006: HHS (24th), OPM (23rd), and Homeland Security (21st).

This Scorecard evaluates only the quality of agency reports, not the quality of the results the agencies produced for the public. Actual agency performance may or may not be correlated with report rankings in this Scorecard.

AGENCY NAMES AND ABBREVIATIONS USED IN THIS SCORECARD

AGENCY NAME	SHORT NAME	COMMONLY USED ABBREVIATION
Department of Agriculture	Agriculture	USDA
Department of Commerce	Commerce	DOC
Department of Defense	Defense	DOD
Department of Education	Education	DOEd
Department of Energy	Energy	DOE
Environmental Protection Agency	EPA	EPA
General Services Administration	GSA	GSA
Department of Health & Human Services	Health & Human Services	HHS
Department of Homeland Security	Homeland Security	DHS
Department of Housing & Urban Development	HUD	HUD
Department of the Interior	Interior	DOI
Department of Justice	Justice	DOJ
Department of Labor	Labor	DOL
National Aeronautics & Space Administration	NASA	NASA
National Science Foundation	NSF	NSF
Nuclear Regulatory Commission	NRC	NRC
Office of Personnel Management	OPM	OPM
Small Business Administration	SBA	SBA
Social Security Administration	SSA	SSA
Department of State	State	State
Department of Transportation	Transportation	DOT
Department of the Treasury	Treasury	Treasury
U.S. Agency for International Development	USAID	USAID
Department of Veterans Affairs	Veterans	VA

TABLE 1

SCORECARD SUMMARY & RANKING FOR FISCAL YEAR 2006

Highest Rank = 1; Lowest = 24. Maximum Possible Score = 60; Minimum = 12.

	TRANSPARENCY	PUBLIC BENEFITS	LEADERSHIP	TOTAL	RANK
Transportation	19	17	17	53	1
Labor	16	18	17	51	2
Veterans	19	16	16	51	2
State	17	18	15	50	4
USAID	16	15	11	42	5
GSA	15	14	11	40	6
Justice	15	10	12	37	7
Commerce	13	10	13	36	8
Education	14	10	12	36	8
EPA	12	11	13	36	8
Agriculture	12	10	13	35	11
Treasury	14	9	12	35	11
Energy	12	10	12	34	13
NRC	12	11	11	34	13
SSA	13	7	13	33	15
Defense	14	8	10	32	16
NASA	12	7	13	32	16
Interior	14	8	9	31	18
NSF	13	8	10	31	18
SBA	10	9	12	31	18
HUD	11	10	9	30	21
DHS	12	8	10	30	21
OPM	13	7	8	28	23
HHS	9	8	8	25	24
AVERAGE	13.6	10.8	12.0	36.4	
MEDIAN	13.0	10.0	12.0	34.5	

TABLE 2

FISCAL YEAR 2006 SCORES & RANKINGS COMPARISON TO FISCAL YEAR 2005

Highest Rank = 1; Lowest = 24. Maximum Possible Score = 60; Minimum = 12

	FY 2006		FY 2005		CHANGE IN SCORE	CHANGE IN RANKING
	TOTAL SCORE	RANK	TOTAL SCORE	RANK		
Transportation	53	1	47	3	6	2
Labor	51	2	51	1	0	-1
Veterans	51	2	46	4	5	2
State	50	4	48	2	2	-2
USAID	42	5	38	8	4	3
GSA	40	6	31	17	9	11
Justice	37	7	36	11	1	4
Commerce	36	8	41	5	-5	-3
Education	36	8	31	17	5	9
EPA	36	8	34	13	2	5
Agriculture	35	11	37	9	-2	-2
Treasury	35	11	41	5	-6	-6
Energy	34	13	37	9	-3	-4
NRC	34	13	35	12	-1	-1
SSA	33	15	29	21	4	6
Defense	32	16	29	21	3	5
NASA	32	16	30	20	2	4
Interior	31	18	34	13	-3	-5
NSF	31	18	33	15	-2	-3
SBA	31	18	40	7	-9	-11
HUD	30	21	33	15	-3	-6
DHS	30	21	27	24	3	3
OPM	28	23	31	17	-3	-6
HHS	25	24	25	24	0	0
AVERAGE	36.38		36.00		0.39	
MEDIAN	34.50		34.50		1.00	

TABLE OF CONTENTS

Introduction	1
Evaluation Criteria	5
Scoring Standards	9
Scoring Summary	13
Beyond Performance Reports: Comprehensive Accountability	21
Strongest and Weakest Scores	29
Transparency	29
Public Benefits	44
Forward-Looking Leadership	52
Agency-by-Agency Scoring Summaries	61
Department of Transportation	62
Department of Labor	63
Department of Veterans Affairs	64
Department of State	65
U.S. Agency for International Development	66
General Services Administration	67
Department of Justice	68
Department of Commerce	69
Department of Education	70
Environmental Protection Agency	71
Department of Agriculture	72
Department of the Treasury	73
Department of Energy	74
Nuclear Regulatory Commission	75
Social Security Administration	76
Department of Defense	77
National Aeronautics and Space Administration	78
Department of the Interior	79
National Science Foundation	80
Small Business Administration	81
Department of Housing and Urban Development	82
Department of Homeland Security	83
Office of Personnel Management	84
Department of Health and Human Services	85
Research Team and Project Design	86

INTRODUCTION

This year, America celebrates the 400th anniversary of the settlement of Jamestown—a colony that almost perished due to a crisis of accountability. The colonists faced starvation because many refused to work. The “gentlemen” believed physical labor was beneath their dignity, and the less gentle preferred to pilfer tools, guns, and other items to barter with the natives for food. Neither group was held accountable for the negative consequences of its behavior.

In 1608, Captain John Smith became president of the colony’s governing council. His solution to the accountability problem is legendary: “You must obey this now for a Law, that he that will not worke shall not eate (except by sickness he be disabled) for the labours of the thirtie or fortie honest and industrious men shall not be consumed to maintaine an hundred and fiftie idle loyterers.”¹

Smith pioneered performance reporting. “He made also a Table, as a publicke memorial of every mans deserts, to incourage the good, and with shame to spurre on the rest to amendment. By this many became industrious, yet more by punishment performed their businesse, for all were so tasked, that there was no excuse could prevaile to deceive him”²

Smith employed these seemingly harsh accountability measures not just to reward those who did work, but also for a significant public purpose: to provide for those who could not care for themselves.

You cannot deny but that by the hazard of my life many a time I have saved yours, when (might your owne wills have prevailed) you would have starved; and will doe still whether I will or noe; But I protest by that God that made me, since necessitie hath not power to force you to gather for your selves those fruits the earth doth yeeld, you shal not onely gather for your selves, but those that are sicke The sick shall not starve, but equally share of all our labours; and he that gathereth not every day as much as I doe, the next day shall be set beyond the river, and banished from the Fort as a drone, till he amend his conditions or starve.³

Positive outcomes quickly ensued: “This order many murmured was very cruell, but it caused the most part so well to bestirre themselves, that of 200 (except they were drowned) there died not past seven.”⁴

Though predating the Government Performance and Results Act by 385 years, Smith’s initiative satisfies many of the prerequisites for successful performance management. Beneficial outcomes were well defined, performance measures were clear, causation was well understood, and performance was transparently linked to consequences. Performance reporting was public, accessible, jargon free, and verifiable.

¹ Philip H. Barbour, ed., *The Complete Works of Captain John Smith (1580-1631)*, vol II. (Chapel Hill: University of North Carolina Press, 1986): 208.

² *Ibid.*, 208.

³ *Ibid.*, 213-14.

⁴ *Ibid.*, 214.

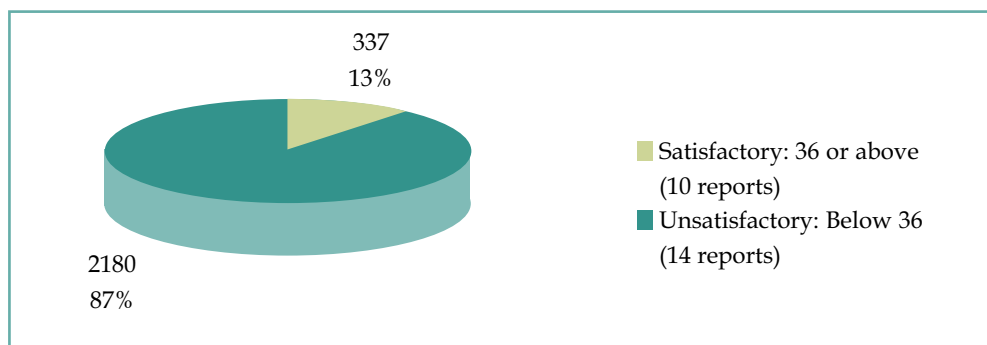
Today's federal government is rather more complex than the original Jamestown settlement. That makes effective accountability more difficult, but the same basic principles apply: define outcomes, define measures, understand how activities affect outcomes, and link outcomes to costs and consequences.

That's the genius of the Government Performance and Results Act (GPRA). Enacted in 1993, this legislation directs federal agencies to define the outcomes their agencies seek to produce, identify measures that show whether they are making progress on these outcomes, and disclose the results to Congress and the public.⁵

Federal agencies are supposed to be accountable to elected policy makers. Elected policy makers are supposed to be accountable to citizens. Surely the GPRA-mandated disclosure of goals and results is a first crucial step toward both types of accountability.

Ever since agencies issued their first performance reports for fiscal year 1999, the Mercatus Center has assembled a research team to assess the quality of their disclosure. As in past years, this Scorecard assesses the reports issued by the 24 agencies covered by the Chief Financial Officers Act (CFO Act). These agencies accounted for 99 percent of federal outlays in fiscal 2006.⁶ The 24 comprise all the Cabinet departments plus the largest independent agencies.

**FIGURE 1: FISCAL 2006 SPENDING (\$BILLIONS)
COVERED BY SATISFACTORY DISCLOSURE**



⁵ For a brief account of the evolution of federal performance reporting and the laws that influenced it, see Maurice McTigue, Henry Wray, and Jerry Ellig, *7th Annual Performance Report Scorecard: Which Federal Agencies Best Inform the Public?* (Arlington, VA: Mercatus Center at George Mason University, 2006), 21-22, http://www.mercatus.org/publications/pubid.2265/pub_detail.asp.

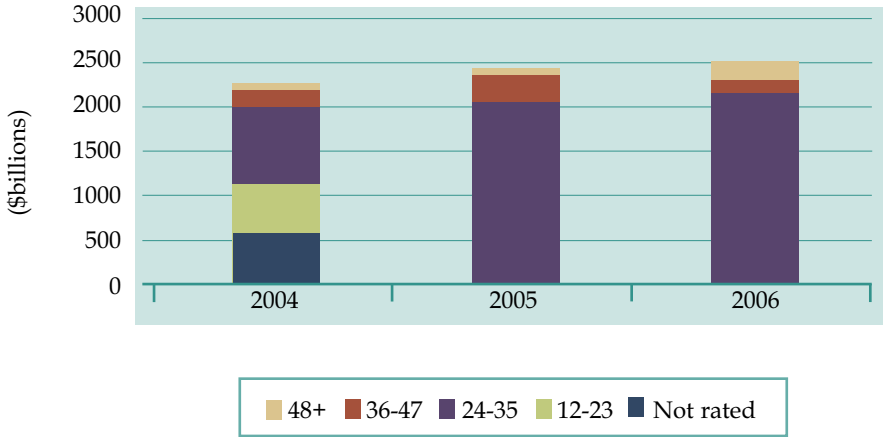
⁶ The principal parts of government not included in these 24 agencies are the judiciary, the legislative branch, the executive office of the president, and the independent agencies not among the 24 CFO Act agencies. See *Budget of the United States Government, Fiscal Year 2008*, "Historical Tables," Table 4.1, 78. Outlays for these agencies actually exceed the "total outlays" figure, but they account for 99 percent of total outlays plus undistributed offsetting receipts.

The quality of disclosure in these reports has improved, but the improvement has not been as rapid or as widespread as we had hoped. Comparing the quality of agency performance and accountability reports with federal expenditures provides a rough idea of how much federal spending is devoted to programs with well-documented outcomes. In fiscal 2006, after eight years of reporting, only 13 percent of non-interest federal spending by the 24 agencies was covered by reports with total scores of “satisfactory” or better in our evaluation.⁷ Though they covered only one-seventh of non-interest expenditures, these reports came from 10 of the 24 CFO Act agencies, as Figure 1 shows.

Figure 2 suggests that the best reports have improved their scores. More spending is now covered by reports with scores of very good (48 or better).⁸ Nevertheless, the percentage of expenditures covered by satisfactory disclosure remains virtually unchanged between fiscal 2005 and 2006, and more than \$2 trillion of the budget is still covered by unsatisfactory disclosure.

The four public benefits criteria most directly measure the quality of disclosure of outcomes. Figure 3 shows that in fiscal 2006 about 8 percent of the federal budget was covered by reports that earned a satisfactory score

FIGURE 2: MOST SPENDING COVERED BY DISCLOSURE SCORING BELOW SATISFACTORY (BELOW 36)

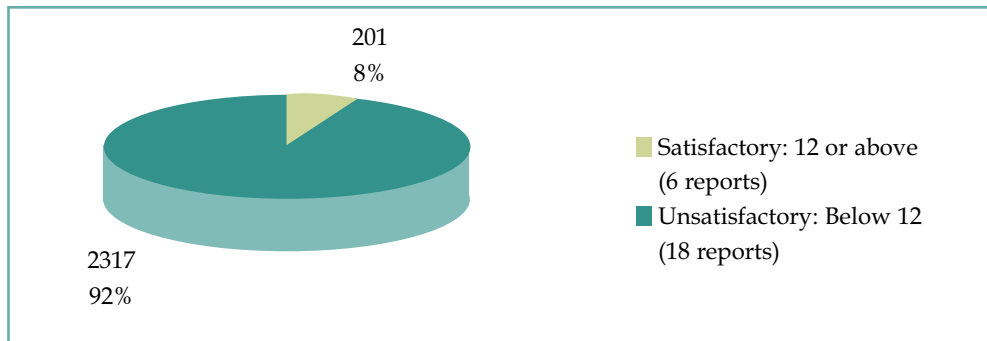


⁷ Reports can earn a maximum of five points on each of 12 criteria. A score of 3 on a criterion corresponds to a “satisfactory” rating, and a report with an average score of 3 across all 12 criteria would earn a 36. Thus, a report must earn a score of 36 or higher to be classified as “satisfactory.” For further explanation, see page 9.

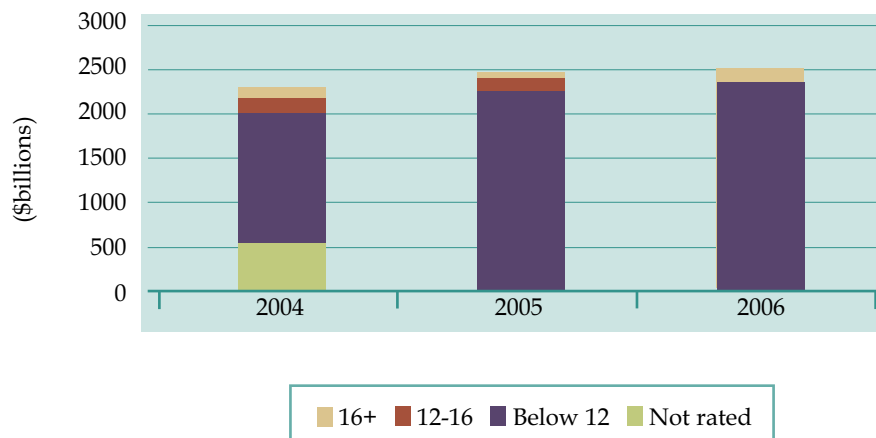
⁸ A score of 4 on a criterion corresponds to a “very good” rating. A report with an average score of 4 across all 12 criteria would earn a 48. Thus, a report must earn a score of 48 or higher to be classified as “very good.” For further explanation, see page 9.

on the public benefits criteria—12 or better. This is virtually the same as in fiscal 2005. However, almost all of the best reports actually scored 16 or better, thus earning a score of very good on the public benefits criteria. As Figure 4 shows, about \$190 billion of non-interest spending, or 7.6 percent, was covered by reports rated very good on the public benefits criteria. In contrast, just 3 percent of spending was covered by reports earning a very good public benefits score in fiscal 2005.

**FIGURE 3: FISCAL 2006 SPENDING (\$BILLIONS)
COVERED BY REPORTS WITH SATISFACTORY PUBLIC BENEFITS SCORE (12 OR ABOVE OUT OF 20)**



**FIGURE 4: MOST SPENDING COVERED BY PUBLIC BENEFITS
DISCLOSURE SCORING BELOW SATISFACTORY (BELOW 12)**



EVALUATION CRITERIA

The purpose of this assessment is not to evaluate or make judgments about the quality of actual results the agencies produced. Rather, our goal is simply to ascertain *how well the agencies' reports disclose to the public* the results they produced, so that policy makers and citizens may make informed judgments about the agencies' results. We review the reports solely from this perspective and not as accountants, government insiders, or experts on the functions of particular agencies.

Each agency's performance report necessarily addresses different audiences. Some readers are "stakeholders" who have expertise in the agency's work. They seek an extensive level of understanding about the agency's performance and may be willing to plow through a lengthy, detailed, and technical report to get it. From our perspective, though, the most important stakeholders are the ordinary citizens who pay the bills and deserve to know what the agency accomplished. A report will not score well in our evaluation if it does not do a good job of informing the average citizen, even if it is informative for experts, insiders, or others who have more specialized knowledge.

Reports that score high on our evaluation communicate important performance results in a way that lay readers—ordinary citizens and taxpayers—can understand. This key trait is relevant to most categories in our Scorecard, and the best reports tend to score well across the board. Reports that consistently score poorly do little to inform ordinary members of the public about important outcomes. Reports ranking in the middle may serve some audiences well, but they could do a better job of demonstrating the agency's value to ordinary citizens.

Specifically, in order to rank highly in this Scorecard, a report must:

- use clear, concise presentation formats and language throughout that a lay person can follow and understand;
- present a set of performance metrics that captures important public outcomes that a lay reader can relate to and appreciate;
- reinforce these performance metrics with clear narratives illustrating public benefits that flow from the agency's work;
- enable the lay reader to readily grasp and assess progress toward outcomes;
- provide confidence that the agency has adopted challenging measures, forthrightly acknowledges performance shortfalls, and takes steps to correct them; and
- provide confidence that the agency serves as a good steward of taxpayer resources by taking effective steps to resolve major management challenges.

Our research team used 12 evaluation factors grouped under three general categories of transparency, public benefits, and leadership.

1. Does the agency disclose its accomplishments in a *transparent* (easily understood) fashion?
2. Does the report focus on disclosing tangible *public benefits* (valued results) the agency produced?
3. Does the report show evidence of forward-looking *leadership* (guidance) that uses performance information to devise strategies for improvement?

TRANSPARENCY

Reports should be accessible, readable, and useable by a wide variety of audiences, including Congress, the administration, the public, news media, and other stakeholders. If a report fails to disclose significant achievements and problems to stakeholders, benefits or failures arising from agency activities will remain secret to all but a few insiders, and citizens will have no real opportunity to indicate their approval or disapproval.

PUBLIC BENEFITS

An agency's value to the public becomes clear only when its goals and measures are expressed in terms of the benefit produced or harm avoided for a particular set of clients or the public at large. To demonstrate openly how agency activities produce meaningful results for the community, reports should focus on outcomes (i.e., tangible benefits that matter in the lives of citizens) rather than on programs or activities. The reports should also clearly present the costs of achieving those results. The ultimate objective of such reporting is to match outcomes with costs, so that policy makers and the public understand what citizens are paying to achieve various outcomes.

TRANSPARENCY

1. Is the report easily accessible via the Internet and easily identified as the agency's Annual Performance and Accountability Report?
2. Is the report easy for a layperson to read and understand?
3. Are the performance data valid, verifiable, and timely?
4. Did the agency provide baseline and trend data to put its performance measures in context?

PUBLIC BENEFITS

5. Are the goals and objectives stated as outcomes?
6. Are the performance measures valid indicators of the agency's impact on its outcome goals?
7. Does the agency demonstrate that its actions have actually made a significant contribution toward its stated goals?
8. Did the agency link its goals and results to costs?

LEADERSHIP

9. Does the report show how the agency's results will make this country a better place to live?
10. Does the agency explain failures to achieve its goals?
11. Does the report adequately address major management challenges?
12. Does it describe changes in policies or procedures to do better next year?

Goals and measures that merely document agency activities, such as counts of checks processed or number of people enrolled in a program, assume that such activities automatically provide public benefits. Such an assumption can be incorrect for a wide variety of reasons. Thus reports must highlight achievement of results; otherwise, they will not inform the public of the success or failure of the programs. Budget decisions that rely on such flawed information will fail to reflect realistic assessments of what agencies can accomplish with appropriations.

FORWARD-LOOKING LEADERSHIP

Agencies should use the performance information produced by their organizations to identify solutions to problems and to change future plans accordingly. The reports should inspire confidence in the agencies' abilities to enhance citizens' quality of life commensurate with the resources they have entrusted to them. Among the factors that give such confidence is tangible evidence that the agencies are using performance and financial data to improve management of their programs.

WHAT DID THE AGENCIES KNOW, AND WHEN DID THEY KNOW IT?

As in past years, the Mercatus Center notified federal agencies of the deadlines our research team would follow in evaluating the reports and the evaluation criteria we would employ.

In April 2006, when we released the fiscal year 2005 Scorecard, we also released a guidance document that explained how the best practices in the fiscal 2005 performance reports would affect the scoring of fiscal 2006 performance reports.⁹ We have incorporated the guidance into the "Strongest and Weakest Scores" section of this Scorecard.

For fiscal 2006, the Office of Management and Budget required agencies to submit their reports to the president and Congress by November 15—approximately six weeks after the fiscal year ended. In September 2006, the Mercatus Center notified each agency's chief financial officer via letter (and other individuals listed as agency GPRA contacts by e-mail) that the Mercatus research team would need a copy of the report by December 1 in order to include it in this year's evaluation. The letter also mentioned that reports would need to be available on the Internet by December 15 to earn credit on the first transparency criterion. A follow-up e-mail reminder was sent to the GPRA contacts on November 13. All but one of the agencies had their reports accessible on the Internet by December 15.

The September letter included an explanation of our evaluation criteria and noted that the quality of each year's reports raises the bar for subsequent years. It also thanked agencies for their continued participation in our study and iterated our goal of continuing to improve the quality of reporting. Finally, the letter invited agency personnel to contact Mercatus Center staff with questions or comments about the criteria, and many did so. Thus, agencies had ample notice about the criteria and deadlines.

⁹ "Annual Performance Report Scorecard Criteria Guidance" (Mercatus Center at George Mason University, April 2006), http://www.mercatus.org/publications/pubid.2264/pub_detail.asp.

SCORING STANDARDS

Each report had the opportunity to earn up to 20 points in each of the three categories, for a maximum score of 60 points. Each category included four equally weighted evaluation factors, and scores of 1 through 5 (from inadequate to outstanding) were awarded for each evaluation factor. Thus, an agency could achieve a minimum score of 12 merely by producing a report.

THE 5-POINT SCALE

The 5-point rating scale for individual criteria identifies distinct levels of quality. The research team used the accompanying table to guide its scoring. A report that adequately meets all requirements would receive the middle score of 3 on each factor, resulting in a total score of 36. A 2 indicates that the report accomplishes some but not all of the objectives under a given criterion. A 1 indicates failure to provide much relevant information. A 4 indicates unusually good practices that are better than most, and a 5 indicates an especially superior presentation.

Even when a report receives a 5 on a particular criterion, that does not mean there is no room for improvement. A 5 indicates a standard for best practice, but best practices should not be confused with perfection. We expect agency reporting practices to improve over time, and one of the goals of this Scorecard is to aid in the diffusion of best practices across agencies. Therefore, a practice that earned a 5 this year may only deserve a 4 or 3 in future years as it becomes standard for most agencies and new best practices emerge.

WEIGHTING THE EVALUATION FACTORS

To report the results of this study as transparently as possible, the researchers weighted the evaluation factors equally in calculating each agency's total score and rankings. Since the summary table reports scores for all three evaluation categories (transparency, public benefit, and forward-looking leadership) separately, readers who believe that one factor is more

WHAT DO THE SCORES MEAN?

5 Outstanding

- Substantially exceeds expectations
- Opens up a new field of information
- Sets a standard for best practice

4 Very Good

- Exceeds expectations
- Has potential to be a best practice
- Shows innovation and creativity
- Better than most

3 Satisfactory

- Meets expectations in all aspects
- Adequate, but does not exceed expectations

2 Unsatisfactory

- Fails to meet expectations
- May be adequate in some respects, but not all
- Produces partial information
- Does not fully disclose

1 Inadequate

- Fails to meet expectations
- Does not meet standard for adequate disclosure
- Shows no process or plans to overcome problems
- Omits critical information

important than others can apply whatever weights they wish to the separate scores and recalculate rankings accordingly.

In addition, in the interest of transparency, all reports were evaluated against a common scale even though different agency missions may make it inherently more difficult to develop results-oriented goals and measures or collect appropriate data. For example, agencies that provide direct measurable services, such as the General Services Administration or Department of Veterans Affairs, might find it easier to identify and quantify their contributions than an agency like the State Department, which deals in more intangible results. In reality, some agencies that seem to provide few services directly to members of the public, such as State, have eventually produced highly ranked reports, and some that arguably have a more direct effect on citizens' well-being, such as Health and Human Services, have produced low-ranked reports.

INTERPRETING OUR FINDINGS

It is important to emphasize that our research team evaluated only *the quality of reporting, not the quality of results*. Therefore, it would be a mistake to conclude that the agencies with the highest-scoring reports necessarily produce the most or best results for the country. Ideally, an agency's report reflects more about its managers' capabilities than just their ability to write reports. A high-scoring report reflects an agency's ability to explain its results in understandable and meaningful terms that Americans can appreciate.

Similarly, it would also be inappropriate to draw policy conclusions from our analysis. We offer no recommendations on whether the federal government should or should not be engaged in its current menu of activities.

So what do the findings in this study really mean? By assessing the quality of agency reports, we evaluate how the agencies are supplying the information that Congress and the public need to make informed funding, budgeting, and policy decisions.

An additional word on information quality is also in order. We assessed the quality of each report's disclosure of data verification and validation procedures. However, in the interest of producing a timely study, we did not independently verify the performance information cited in each agency's report. The reports themselves should inspire confidence by indicating how data are verified and validated.

OUR CONSISTENCY CHECK

Our research team employed the same criteria to assess the fiscal 2006 agency reports that we used to evaluate prior year reports. In each succeeding year however we have tightened our evaluation standards. There are two reasons for this approach. One, the highest achievable quality is unlimited because creative innovators can always find ways to improve reporting practices and set new standards. Two, each year agencies have an opportunity to learn from others' best practices. If we did not continually raise our expectations, most reports could eventually receive mostly 5s. This Scorecard would then convey little information about the quality of different agencies' reports, and it would give little recognition or credit to those agencies that continue to raise the bar for quality reporting.

For these reasons, an agency had to improve the absolute quality of its fiscal 2006 report in order to receive the same numeric score it received for its fiscal 2005 report. If an agency receives a higher score, that score is a reliable indicator that the quality of its report has indeed improved.

Several factors help ensure that the scoring criteria are applied consistently from year to year. The same Mercatus Center research team has evaluated the reports for the past four years. The team cross-checked the 2006 evaluations against the previous year's in several ways. For each report, the research team generated an extensive set of notes documenting the reasons for each preliminary score on each criterion. The head of the research team reviewed this documentation for both the fiscal 2006 and fiscal 2005 reports to ensure that differences in the actual contents of the reports justified any scoring differences across years. The team discussed instances in which proposed scores differed substantially from the previous year's scores.

Finally, for each report, a member of our outside advisory panel with extensive experience in performance reporting reviewed the report, scoring, and documentation. Some scores were modified when the advisor reached different conclusions from the research team and offered persuasive reasons for the difference. Final scores thus reflect a careful review to ensure that the results of the scoring process are consistent with the goal of raising standards.

SCORING SUMMARY

SAME FINAL FOUR—WITH AN UPSET

Reports from four agencies continue to set the pace: Transportation, Veterans Affairs, Labor, and State. These have been the four highest-ranking reports since fiscal 2003. Just three points separated these four reports in fiscal 2006.

For the first time since fiscal 2001, Transportation tops the rankings, earning 53 of a possible 60 points. Labor and Veterans Affairs tied for second place, earning 51 points each. State, the number 2 report last year, remained in the top four with a score of 50.

Fiscal 2006 saw a significant quality gap emerge between the top four and the rest of the pack. The fifth place report, submitted by USAID, earned 42 points. USAID's score improved by 3 points from fiscal 2005. (The biggest single factor was a jump in the score for criterion 8, linkage of results to costs, from 1 in 2005 to 4 in 2006. USAID accomplished this by linking costs with performance goals in the same manner as State.) Below the top four, there was significant churn in scoring and rankings between 2005 and 2006. Both "shooting stars" and "falling stars" abounded.

SHOOTING STARS

Several agencies made significant improvements in their reports for fiscal 2006. A nine-point improvement led GSA to jump 11 spots in the rankings, from 17th in fiscal 2005 to 6th in fiscal 2006. Education vaulted from 17th place in fiscal 2005 to 8th place in fiscal 2006 on the strength of a five-point gain in its score.

GSA made major improvements on criterion 1 (availability) and criterion 8 (linkage of results to costs). Its score on criterion 1 rose from 2 in fiscal 2005 to 5 in fiscal 2006. More impressively, its score on linkage of results to costs rose from 1 in fiscal 2005 to 5 in fiscal 2006. GSA's report includes a table that links projected fiscal year 2006 budget costs to all individual performance goals and, with very few exceptions, to the individual measures for the performance goals. This example of best practice is all the more impressive given that GSA's report for fiscal 2005 had no material linking results to costs.

Education's report improved in less dramatic fashion, with modest one-point gains in several criteria. The biggest improvement to this report was seen in criterion 2, readability. The report is well-organized, covers a reasonable number of measures, and is shorter than most agencies' reports. This is a significant improvement over fiscal 2005, when we noted that the report was difficult for the lay reader to understand.

Other agencies moving up significantly in the rankings, albeit with smaller score increases, were EPA (13th place to 8th), Social Security (21st place to 15th), and Defense (21st place to 16th). More modest improvements came in reports from USAID (8th place to 5th) and NASA (20th place to 16th).

Two of this year's improvers—Social Security and GSA—had steadily fallen in the rankings between 2002 and 2005. Their 2006 scores give hope that they've turned the corner.

FALLING STARS

The biggest decline occurred for SBA, which fell from 7th place in fiscal 2005 to 18th in 2006 as its score dropped by nine points. Four points of this drop, however, occurred because SBA did not have its report posted on the Internet by the deadline. Were it not for that glitch, SBA would have tied with Agriculture and Treasury for 11th place.

Several other agencies fell four or more places in the rankings: Treasury (5th to 11th), Interior (13th to 18th), HUD (15th to 21st), Energy (9th to 13th), and OPM (17th to 23rd). Scores for Treasury and Commerce dropped six and five points, respectively, reflecting one-point drops on multiple criteria. Scores for the other agencies fell by only a few points.

CELLAR DWELLERS

Two reports that have ranked consistently in the bottom half showed noticeable improvement in fiscal 2006:

- Defense: This report's 16th place finish is its best showing since it ranked 18th in 2000. Defense has risen in the rankings for two years in a row.
- NASA: NASA's ranking history may be the most erratic of all the 24 agencies. Its ranking went from 14th in fiscal 1999 to 23rd in fiscal 2000, rebounded over two years to 12th in fiscal 2002, plummeted to 20th in fiscal 2003, rose to 16th in fiscal 2004, then fell back to 20th in fiscal 2005. Whether the rise back to 16th in 2006 is the beginning of a comeback remains to be seen.

Several reports remain near the bottom and show little sign of improvement. These include:

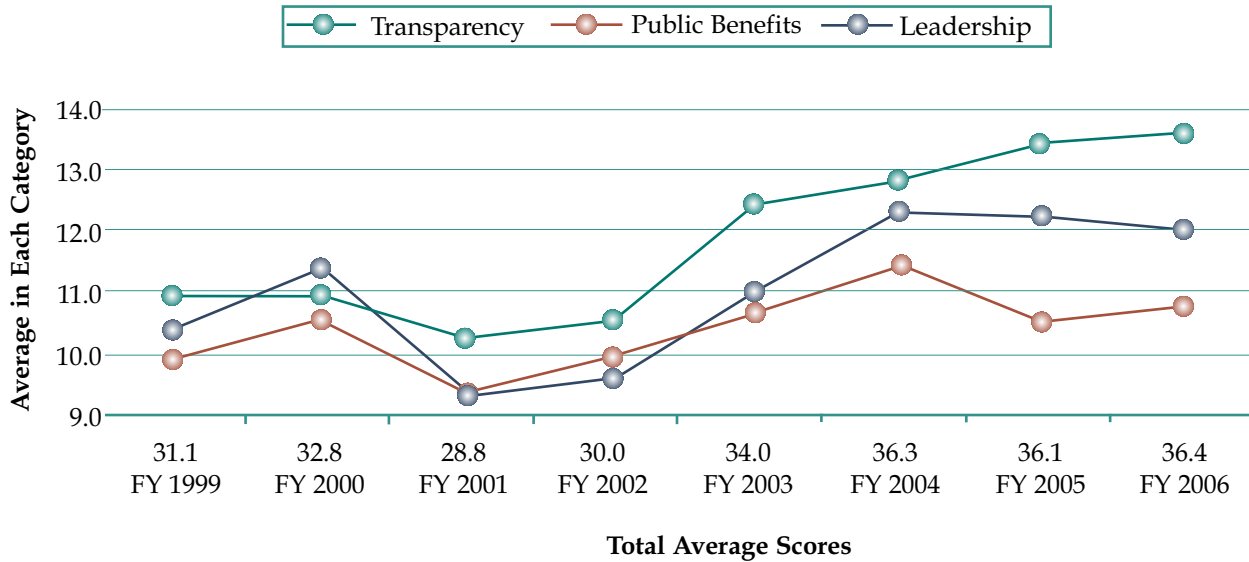
- Health and Human Services: Ranked 24th and has never ranked higher than 20th.
- Office of Personnel Management: Finished 23rd in fiscal 2006, 17th in fiscal 2005, and never finished higher than 12th.
- Homeland Security: Its 21st place finish in fiscal 2006 is the best in its four years of producing a report.

AVERAGE SHOWS SCANT IMPROVEMENT

As Figure 5 shows, the average total score has remained at about 36 since fiscal 2003. The average transparency score rose by one percent, the public benefits score rose by three percent, and the forward-looking leadership score fell by two percent.

Despite continual tightening of Scorecard criteria, many agencies demonstrated that it was possible to achieve a higher score in fiscal 2006 than in 2005. Eleven reports improved their scores in fiscal 2006 compared to 2005. Eleven received lower scores, and two were unchanged.

FIGURE 5: UNEVEN PROGRESS



The increase in the public benefits score resulted primarily from a 22 percent increase in the score on criterion 8, linkage of results to costs. This reverses a drop of similar size that occurred between fiscal 2004 and fiscal 2005. In fiscal 2005, scores on this criterion fell because they failed to keep pace with our research team’s rising expectations based on the prior year’s best practices. In fiscal 2004, some agencies found ways to link costs to performance goals or other levels below the strategic goal level. Linkage to some level below strategic goals became the new standard for satisfactory performance. After fiscal 2004, a report could not receive better than a 1 unless it broke costs down by more than just strategic goals. Similarly, a report could not receive a 5 on criterion 8 unless it actually supplied cost information that corresponded to outcome measures. Only a few agencies rose to the challenge in fiscal 2005, but many more did so in fiscal 2006, thus raising the average score on this criterion.

The principal factor driving the improvement in the transparency score was a 6 percent improvement in the average score on criterion 1, accessibility. Seventeen reports scored a 5 on this criterion, reflecting the fact that the report appeared on their Web site by our deadline, could be found via a direct link on the home page, could be downloaded as a single document or multiple sections, and was accompanied by contact information for readers having further questions or comments. For years, we have emphasized that these “mechanical” factors should be easy for any agency to execute, and in fiscal 2005, almost three-quarters of them proved our point. In fiscal 2004, only six reports scored a 5 on criterion 1. Scoring standards will tighten in the future when an agency introduces the next big innovation in transparency. After that only reports adopting the new best practice will receive a 5.

The average score on criterion 9 fell by 5 percent. This was the only criterion whose average score fell by more than a few percentage points. Three of the top four reports actually increased their score on this criterion from

4 in 2005 to 5 in 2006. This was more than offset, however, by eight reports that experienced one-point declines on this criterion.

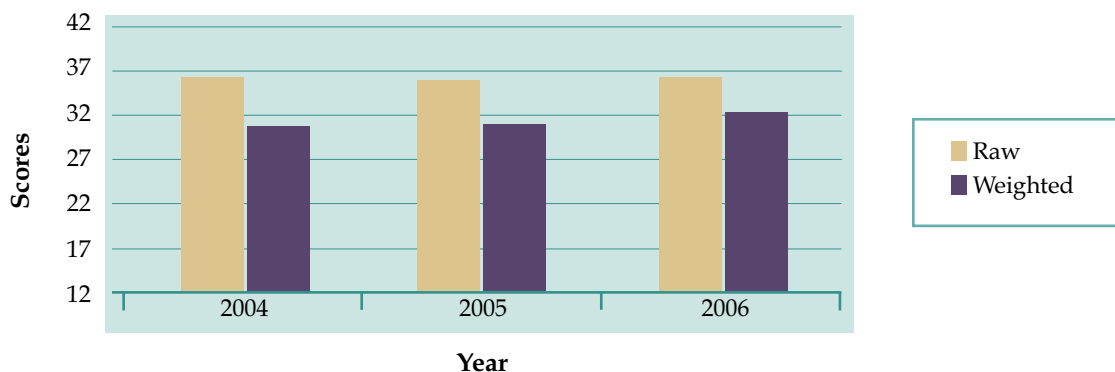
Improved best practices, and consequent tightening of scoring standards, also explain this drop. Criterion 9 gives agencies an opportunity to offer a broad vision of their value to the American people. The best reports integrate performance information with narratives that show how the agency’s actions affect real people. Anecdotes, however, are not a substitute for meaty performance information. To earn a high score on this criterion, a report must combine lofty vision with solid performance metrics.

WEIGHTED AVERAGE SHOWS MORE IMPROVEMENT

Figure 5 simply displays the averages of the scores from all 24 reports. Figures 6 and 7 offer a potentially different perspective on accountability by showing weighted average scores, using each agency’s reported “net cost of operations” as its weight.¹⁰

Why calculate a weighted average score? The sizes of the agencies’ budgets vary greatly – from more than \$500 billion for Defense, Social Security, and HHS to less than \$10 billion for each of the six smallest. The quality of

FIGURE 6: RAW VS. DOLLAR-WEIGHTED SCORES



¹⁰ The figures in each report labeled “net cost of operations” roughly correspond to the agency’s net outlays with a few exceptions. Veterans Affairs, for example, includes the present value of future benefit liabilities in its net cost of operations and adjusts the figure each year. Variations in interest rates can have a big effect on the present value of these liabilities and on Veterans’ reported net cost of operations. Therefore, for Veterans we use a figure that excludes the present value of future liabilities. GSA presents a different anomaly; its net cost of operations is usually positive because other agencies pay it for services. For weighting purposes we reverse the sign on GSA’s net cost of operations since we are just using the figure as a measure of GSA’s size.

disclosure for a very large agency affects a much larger portion of federal spending than for a small agency. The weighted average score may thus be a better indicator of overall accountability for federal spending. The scores of a large agency's report have a bigger effect on the weighted average than the scores of a small agency's report.

FIGURE 7: RAW VS. WEIGHTED SCORE CHANGE

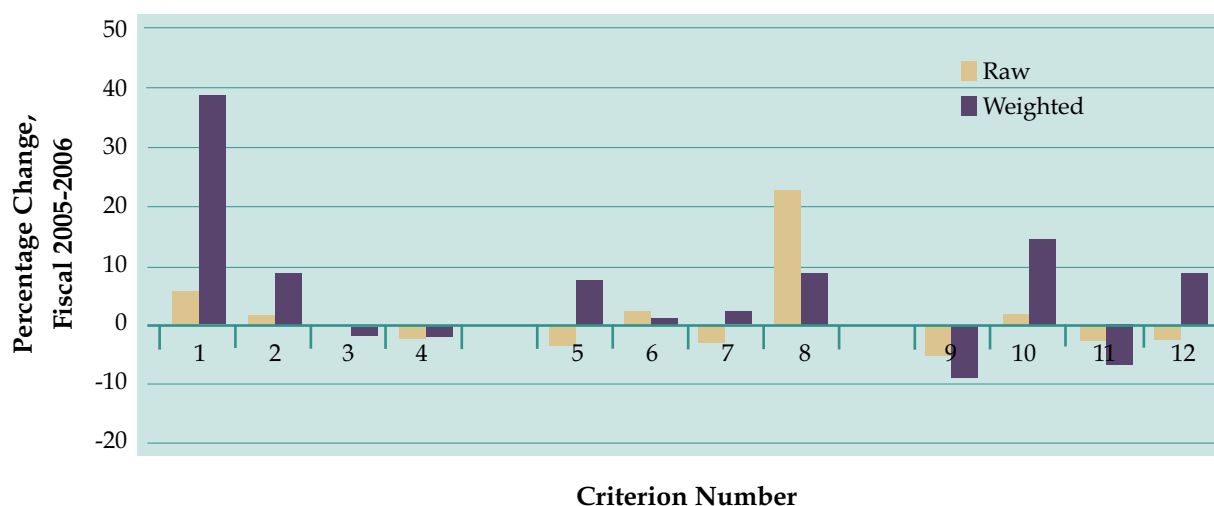


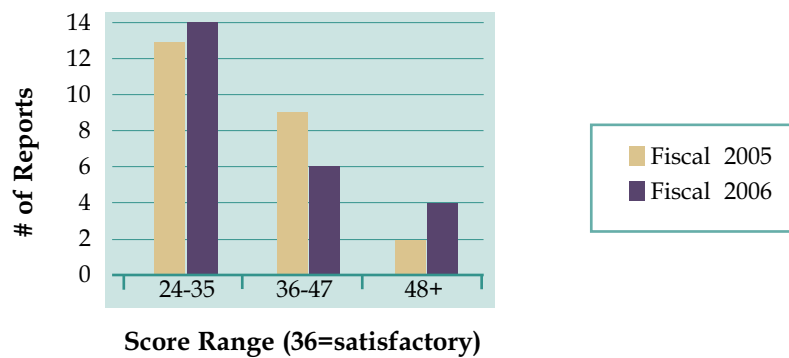
Figure 6 shows that the weighted average score improved at a somewhat faster rate than the raw average in fiscal 2006 after stagnating in fiscal 2005. Figure 7, which graphs scores for individual criteria, also shows that the weighted average improved much more than the raw average on several individual criteria. These numbers reflect the fact that some of the larger agencies demonstrated improvement in fiscal 2006 reports. In other words, the raw average masks some improvement in reports covering large amounts of the budget. These include a three-point increase for Defense (\$581 billion), a three-point increase for Social Security (\$588 billion), and a five-point increase for Education (\$97 billion).

REPORT QUALITY

Figure 8 reveals that 14 reports received scores below the satisfactory average of 36—one more than in 2005. Three agencies whose reports scored below satisfactory in 2005 improved their scores to 36 or above in 2006: Education, EPA, and GSA. Reports from four additional agencies joined the below-satisfactory group for fiscal 2006: Agriculture, Energy, Treasury, and SBA. In fairness, it's worth noting that Agriculture and Energy's scores fell by only a few points.

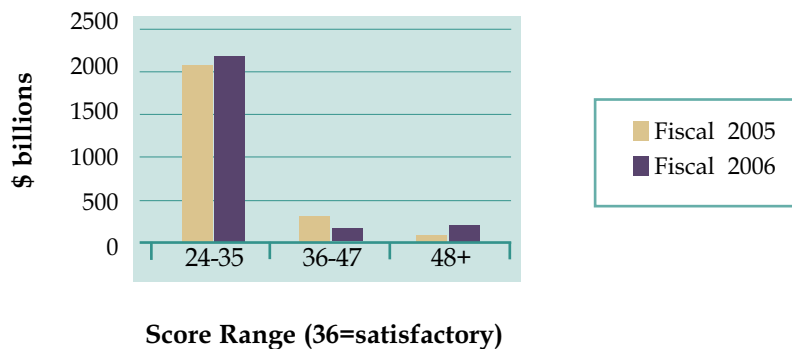
The graph does, however, show that more reports scored 48 or better – making it to the very good range. Only two reports scored 48 or better in fiscal 2005. For fiscal 2006, all of the top four reports scored above 48.

FIGURE 8: FOURTEEN REPORTS ARE BELOW SATISFACTORY



The figures in the introduction to this year’s Scorecard (Figures 1-4) show a very large percentage of the dollars in the budget are covered by below-satisfactory reporting and a very small percentage are covered by very good reporting. Figure 9 sheds additional light on this issue. There are two principal changes in 2006 compared to 2005: more spending is covered by reports that scored very good, and slightly more spending is covered by reports scoring less than satisfactory. As a result, the amount of spending covered by reports in the middle, satisfactory range declined.

FIGURE 9: LESS OF BUDGET IN MIDDLE SCORING RANGE



For several years, we have noted that reports from many of the larger agencies tend to score poorly. As Table 3 indicates this trend continues in fiscal 2006. Agencies with reports scoring below satisfactory spent \$2.18 trillion in fiscal 2006. Reports from all three agencies with the largest budgets scored below 36. However, two of them – Defense and Social Security – are approaching satisfactory scores on their reports, earning 32 and 33 respectively.

TABLE 3: REPORTS SCORING LESS THAN SATISFACTORY

DEPARTMENT	COST	TOTAL SCORE (36=SATISFACTORY)	PUBLIC BENEFITS SCORE (12=SATISFACTORY)
Agriculture	\$95.4 billion	35	11
Treasury	\$14 billion	35	9
Energy	\$63.9 billion	34	10
NRC	\$100 billion	34	11
Social Security	\$588 billion	33	7
Defense	\$581 billion	32	8
NASA	\$17.6 billion	32	7
Interior	\$13.4 billion	31	8
NSF	\$5.6 billion	31	8
SBA	\$1.5 billion	31	9
HUD	\$39.6 billion	30	10
DHS	\$54.3 billion	30	8
OPM	\$81.4 billion	28	7
HHS	\$624 billion	25	8
TOTAL	\$2.18 TRILLION		

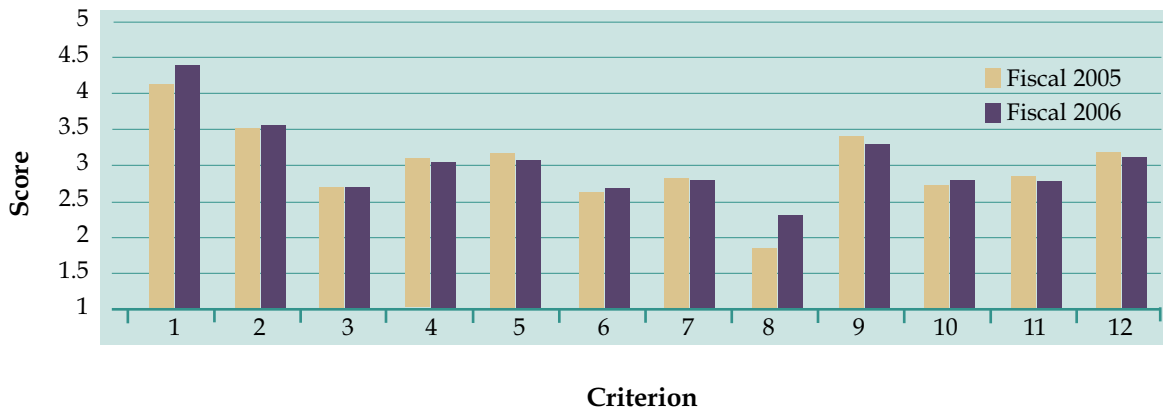
OPPORTUNITIES FOR PROGRESS

Despite this year’s improvements, there is still substantial room for progress. Figure 10 demonstrates that average scores in six categories are below 3. This is true regardless of whether one considers raw averages or weighted averages. The criteria with scores below 3 (in order of severity of the problem) are the following:

- Criterion 8: linkage of results to costs (average score: 2.29)
- Criterion 6: articulation of outcome-oriented measures that accurately reflect the agency’s impact on its goals (average score: 2.67)
- Criterion 3: ensuring reliability and timeliness of data (average score: 2.67)
- Criterion 7: demonstration that the agency’s efforts actually affected achievement of outcomes (average score: 2.75)
- Criterion 10: explanation of failures to achieve goals (average score: 2.79)
- Criterion 11: discussion of major management challenges (average score: 2.79)

Scores for these six criteria have averaged below 3 since fiscal 2004.

FIGURE 10: SIX CRITERIA REMAIN BELOW SATISFACTORY



BEYOND PERFORMANCE REPORTS: COMPREHENSIVE ACCOUNTABILITY

Performance and Accountability Reports provide information on agencies' outcomes and expenditures. The outcomes covered are those that flow from the agencies' activities, and the associated costs are budgetary expenditures. Comprehensive accountability requires a broader reckoning of costs and consequences. It also requires that decision makers actually use the information to guide the government's management, design, and budget decisions.

COMPREHENSIVE ACCOUNTING FOR OUTCOMES AND COSTS

Expenditures are only one way federal agencies achieve outcomes. The federal government also affects resource allocation in the economy through tax policy and regulation.¹¹ The performance sections of agency reports likely include the effects of regulation, but the financial section includes neither the costs of tax expenditures nor the full costs of regulation. Comprehensive accountability requires a full accounting of the costs and consequences of tax and regulatory policies, as well as on-budget spending.

TAX EXPENDITURES

In addition to on-budget spending, the federal government also seeks to influence outcomes through various targeted tax policies that reduce revenues, such as the exclusion of employer-provided health insurance from income, the home mortgage interest deduction, and the child tax credit. Agencies do not report directly on the outcomes associated with these tax expenditures. Tax expenditures are not explicitly mentioned in GPRA, but the report on GPRA from the Senate Governmental Affairs Committee indicates that the committee expected the Office of Management and Budget to take the lead in assessing the outcomes associated with them.¹²

The "Analytical Perspectives" volume of the fiscal 2008 budget presents an informative discussion of inputs, outputs, and outcomes from tax expenditures and suggests how the results of various categories of tax expenditures could be measured.¹³ Some of the more insightful ideas include:

¹¹ U.S. Office of Management and Budget, *Budget of the United States Government, Fiscal Year 2007, "Analytical Perspectives"* (Washington, DC, 2006): 324.

¹² The Report notes:

To increase significantly the oversight and analysis of tax expenditures, the Committee believes that the annual overall Federal Government performance plans should include a schedule for periodically assessing the effects of specific tax expenditures in achieving performance goals. (This schedule would be in addition to the primary content of the overall plan—the program performance goals tied to the direct expenditure of funds.) The Committee expects that annual performance reports would subsequently be used to report on these tax expenditure assessments. These assessments should consider the relationship and interactions between spending programs and related tax expenditures. The Committee hopes that such reports will foster a greater sense of responsibility for tax expenditures with a direct bearing on substantial missions and goals. See S. Rep. No. 103-58, pp. 27-28.

¹³ *Budget of the United States Government, Fiscal Year 2007, "Analytical Perspectives,"* 324-26.

- measuring the effects of tax expenditures on health insurance by assessing the effects of this tax expenditure on insurance coverage and then assessing how insurance coverage affects health outcomes, such as infant mortality or life expectancy;
- considering tax preferences for military benefits along with direct budget costs when making decisions about defense programs;
- evaluating tax expenditures targeted to particular types of investments (such as research, housing, and energy) by estimating their effects on the cost of capital, the additional amount of investment they create, and the public benefits created by that investment;
- comparing the results of tax expenditures intended to promote community development or higher education with the results of other programs, such as loans and grants, that have the same goal; and
- measuring the effects of tax expenditures related to retirement savings by estimating their effect on retirement incomes, private savings, and national savings.

The volume does not present actual outcome measures or data. It pledges that OMB and Treasury will develop data over the next several years.

The fiscal 2008 budget contains estimates of each tax expenditure for fiscal 2006.¹⁴ Some tax expenditures are quite substantial compared to related federal spending. Deductibility of home mortgage interest, for example, cost \$68.3 billion in fiscal 2006—more than HUD’s entire \$39.6 billion net cost of operations. Exclusion of pension and retirement savings contributions and earnings (such as corporate government and corporate pensions, 403(b)s, and IRAs) cost \$104 billion, equal to nearly one-fifth of the Social Security Administration’s expenditures. Similarly, tax expenditures related to education totaled \$22.8 billion, equal to about one-quarter of the Education Department’s budget. Favored tax treatment for medical insurance and medical expenses cost \$133 billion—equal to more than 20 percent of HHS expenditures.

The mere fact that something is a tax expenditure, and hence reduces federal revenues, does not automatically mean that it is good or bad tax policy. Similarly, tax credits or deductions are not inherently a more effective or less effective means of accomplishing the desired outcome. Good information on outcomes would help policy makers determine whether favored tax treatment or direct federal expenditures most effectively accomplish various public goals, such as promoting home ownership, retirement security, education, and economic growth.

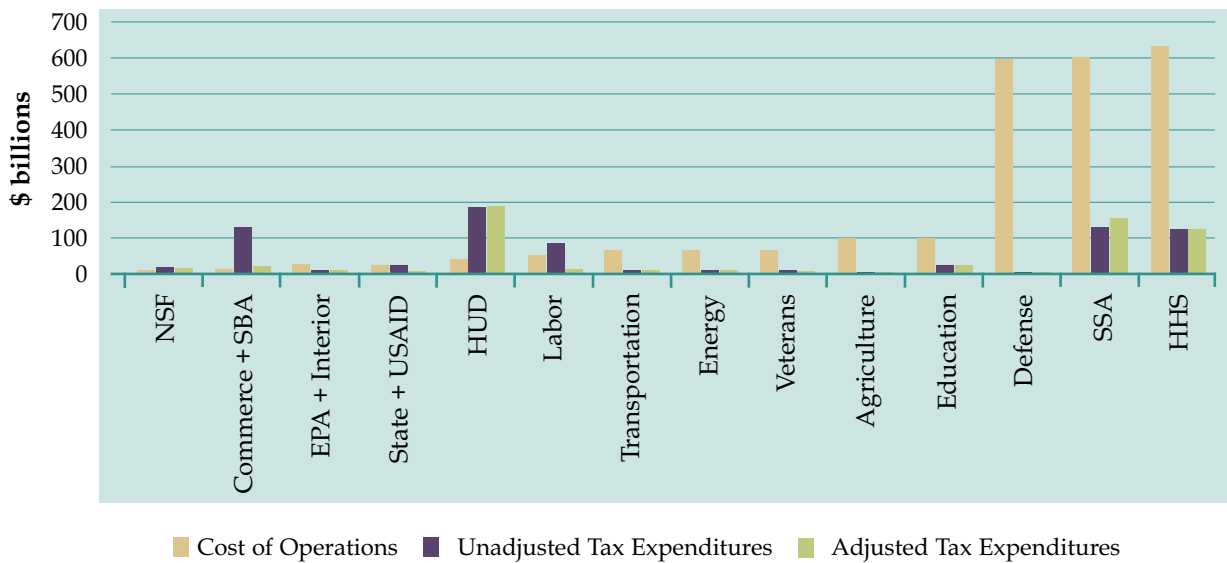
Figure 11 compares agency spending with tax expenditures that appear to be related to that agency’s outcomes. The “Unadjusted” numbers are largely the same as the tax expenditures by budget function that

¹⁴ U.S. Office of Management and Budget, *Budget of the United States Government, Fiscal Year 2008, “Analytical Perspectives,”* (Washington, DC, 2007) Table 19-1.

¹⁵ One principal difference is that we allocated items under Community and Regional Development (lines 84-90) in Table 19-1 of Analytical Perspectives to whichever budget function seemed to make the most sense. For example, Line 85, Exclusion of interest for airport, dock, and similar bonds, was allocated to Transportation. The other principal difference is that tax expenditures related to Income Security (lines 131-150) are included in the total of tax expenditures related to Social Security since the bulk of these tax expenditures are related to retirement plans and income security.

appear in Table 19-1 of the budget’s “Analytical Perspectives” volume.¹⁵ The “Adjusted” numbers alter four agencies’ figures to add or remove large items that arguably align the totals more closely with the outcomes for which the relevant agency is responsible.¹⁶

FIGURE 11: SPENDING VS. OUTCOME-RELATED TAX EXPENDITURES



¹⁶ Adjustments are as follows:

COMMERCE + SBA: Exclude items related to tax treatment of capital gains (lines 67-70, \$79 billion) and depreciation (lines 72-74, \$40.5 billion). These items arguably aim at outcomes more generally related to economic growth than the specific outcomes for which Commerce and SBA are responsible.

STATE + USAID: Exclude items related to international tax policy (lines 4-7, \$19.6 billion). Remaining tax expenditures involve tax treatment of federal employees abroad—arguably a cost of executing U.S. foreign and aid policies.

LABOR: Exclude child tax credit (line 114, \$30.4 billion) and deductibility of charitable contributions other than education and health (line 117, \$37 billion). While these items are related to the size of skill level of the labor force, they are arguably aimed at more general outcomes related to tax or social policy than the outcomes for which Labor is responsible.

SOCIAL SECURITY: Add exclusion of interest on life insurance savings (line 50, \$19.4 billion) since life insurance is obviously related to income security and many people use life insurance as a tax-deferred retirement savings vehicle.

Some tax expenditures are larger than the budgets of agencies with related outcomes. These include the tax expenditures related to science, business, and housing. Others are large in an absolute sense and equal to a noticeable fraction of the related agency's budget. These are the tax expenditures related to education, income security, and health. Performance and accountability reports are unlikely to provide a comprehensive picture of either costs or outcomes in these areas.

Other tax expenditures are small relative to the total federal budget and to the budget of the related agency. These include tax expenditures related to environment and natural resources, foreign affairs, training, employment, transportation, energy, veterans, agriculture, and defense. The federal government accomplishes most of its outcomes in these areas through direct spending or regulation. Some agencies are not even in the graph, because there appear to be no significant tax expenditures obviously related to their missions. These are OPM, GSA, NRC, NASA, Justice, Homeland Security, and Treasury.¹⁷ Performance and accountability reports should be expected to provide comprehensive accountability for outcomes in these areas, but cost disclosure may be incomplete unless the reports include the costs associated with regulation.

The budget chapter containing the tax expenditure estimates cautions that adding up the estimates of various tax expenditures may not create an accurate picture of their size or effects, since a simple summation ignores economic incentive effects and interactions by which the size or existence of one tax expenditure may affect the size of others.¹⁸ Though inexact, adding up the estimates does suit our purpose of providing a rough estimate of the amount of resource allocation the federal government accomplishes through tax expenditures. Tax expenditures totaled \$847 billion in fiscal 2006.¹⁹ The tax expenditure chapter of the budget indicates that data are not yet good enough to assess outcomes. Therefore, tax expenditures represent a very large additional avenue of federal resource allocation for which disclosure of outcomes is currently poor.

REGULATION

Another way the federal government seeks to achieve outcomes is through regulation. Regulation compels or influences consumers, businesses, and other levels of government to expend resources in various ways. The costs of regulation are substantial, often far exceeding the expenditures the agencies incur to promulgate and enforce them. The most recent estimate suggests that compliance with federal regulations costs approximately

¹⁷ Some substantial tax expenditures that are more related to general tax policy, such as the child tax credit, taxation of capital gains, depreciation, and deductibility of state and local income taxes, could arguably be related to Treasury's tax policy mission. In that case, Treasury would join the list of agencies whose related tax expenditures exceed their spending.

¹⁸ "An important assumption underlying each tax expenditure estimate reported below is that other parts of the tax code remain unchanged. The estimates would be different if tax expenditures were changed simultaneously because of potential interactions among provisions. For that reason, this chapter does not present an estimated grand total of the tax expenditures." *Budget of the United States Government, Fiscal Year 2008, "Analytical Perspectives,"* 285. A more detailed discussion appears on page 286.

¹⁹ *Budget of the United States Government, Fiscal Year 2007, "Analytical Perspectives,"* Table 19-1.

\$1.1 trillion—nearly half as large as the \$2.5 trillion in non-interest federal spending.²⁰ The agencies’ performance and accountability reports document many of the outcomes associated with regulation, but not the costs borne by consumers, businesses, and other levels of government. Nevertheless, an agency that engages in a lot of regulation but relatively little spending might direct substantial social resources toward achievement of the outcomes for which it is responsible.

TOTAL

The sum of non-interest spending, tax expenditures, and regulatory costs totaled \$4.45 trillion in fiscal 2006, as Table 4 shows. That’s equal to 33.6 percent of U.S. gross domestic product in 2006.²¹ Tax expenditures and regulatory costs together equal 78 percent of the \$2.5 trillion in non-interest spending.

AGENCIES’ NON-INTEREST NET COST OF OPERATIONS	\$2.5 trillion
ESTIMATED TAX EXPENDITURES	\$847 billion
COST OF REGULATION	\$1.1 trillion
TOTAL	\$4.45 trillion
AMOUNT OF SPENDING IN REPORTS SCORING 36+	\$337 billion (7.6% of total)

Tax expenditures and regulatory costs appear nowhere in agencies’ performance and accountability reports. The Senate Government Affairs Committee’s report on GPRA gives the Office of Management and Budget responsibility for evaluating the costs and results of tax expenditures.²² Agencies evaluate the costs of major regulations when they are proposed, and each year OMB’s Office of Information and Regulatory Affairs com-

²⁰ W. Mark Crain, “The Impact of Regulatory Costs on Small Firms,” Report Prepared for the Office of Advocacy, U.S. Small Business Administration, RFP No. SBHQ-03-M- 0522 (Sept. 2005): 1, <http://www.sba.gov/advo/research/rs264tot.pdf>.

²¹ Gross domestic product figure of \$13.248 trillion is from *Budget of the United States Government, Fiscal Year 2008*, Analytical Perspectives, Table 12-1.

²² See note 12 above.

piles these estimates for the previous ten years. However, this report presumes that outcomes and costs were correctly predicted when the agency adopted each regulation. It does not offer ex-post evaluation of actual outcomes or actual costs.

As a result, federal accountability would have serious deficiencies even if every agency achieved a high score on this Scorecard. Comprehensive accountability requires full analysis of the costs, outcomes, and consequences of spending, tax expenditures, and regulation. Comprehensive disclosure will require that some entity in the federal government take responsibility for integrating information from these diverse sources. Whether that is done by the executive branch or Congress is less important than ensuring that someone in the government actually does it.

TOWARD PERFORMANCE-BASED DECISION MAKING

Effective accountability requires that both the executive branch and Congress actually use performance information to make budget decisions. The executive branch has taken a key step toward performance-based budgeting by creating the Program Assessment Rating Tool (PART). PART assesses the design, management, and performance of individual programs.

One aspect of PART ratings sheds light on the quality of performance measurement for individual programs. If a program lacks sufficient measures or data to evaluate whether it is accomplishing intended results, it receives a rating of Results Not Demonstrated. This does not mean the program has failed to produce results; it means insufficient measures or data are available to determine whether the program is accomplishing its statutory goals. In effect, Results Not Demonstrated is similar to a financial audit that renders a disclaimer, wherein the auditor cannot find sufficient competent evidence to verify items in the financial statements.

A Results Not Demonstrated rating for a program can also be somewhat analogous to a low score in the public benefits section of this Scorecard. Our public benefits criteria examine whether a report expresses goals and measures as outcomes, demonstrates how the agency's actions have affected the outcomes, and links outcomes to costs to permit an evaluation of cost-effectiveness. PART seeks to develop similar measures and data. OMB data on PART ratings and budget authority for programs can be used to calculate the number of programs and percent of the federal budget for which OMB believes performance information is insufficient. This calculation is analogous to our calculation of the percent of the budget covered by reports that received poor scores on our public benefits criteria.

PARTed programs account for about \$2.3 trillion in fiscal 2006 spending, compared to total expenditures of about \$2.66 trillion.²³ About 5.5 percent of PARTed programs, accounting for about \$128 billion in spending, received a Results Not Demonstrated rating in their most recent PART evaluation.²⁴

²³ These two figures are somewhat different measures of federal spending. Spending on PART programs is fiscal 2006 budget authority. Our measure of non-interest expenditures is fiscal 2006 outlays, from *Budget of the United States Government, Fiscal Year 2008*, "Historical Tables," 78, Table 4.1.

²⁴ Authors' calculations based on OMB PART data, <http://www.whitehouse.gov/omb/budget/fy2008/sheets/part.xls>.

A Results Not Demonstrated rating might be analogous to a total score of 8 or below on our four public benefits criteria. This is the score a report would receive if it averaged a 2, “unsatisfactory,” on each of the four criteria. A score of 8 or below on the public benefits criteria suggests that disclosure and documentation of outcomes and cost-effectiveness are substantially incomplete. Seven agencies, representing \$1.96 trillion in federal spending in fiscal 2004, scored 8 or below on the public benefits section. Several large agencies’ reports received an 8, but reports covering \$687 billion worth of spending received less than 8. Thus, reports with substantially incomplete disclosure of outcomes and cost-effectiveness account for a very large share of federal non-interest spending.

Several factors might explain this large difference between our assessment and OMB’s assessment of the quality of outcome information. Perhaps OMB is more lenient in its view of what constitutes adequate results information. In addition, our research team and OMB are examining measures at different levels of aggregation. An agency could have quite good outcome information for some of its large programs and yet still score poorly on our public benefits criteria if many strategic goals and performance measures are not outcome-oriented.

PART may be having some effect on budget decisions. In fiscal 2006, programs rated Results Not Demonstrated or Ineffective were less likely to receive spending increases than other programs. Appropriations increased for 34 percent of programs rated Results Not Demonstrated and 18 percent of programs rated Ineffective. Fifty-nine percent of programs rated Effective received increases in funding. Conversely, 42 percent of Results Not Demonstrated programs and 79 percent of Ineffective programs received

TABLE 5: LOWER PART SCORES INCREASED ODDS OF BUDGET CUTS IN FISCAL 2006

	RESULTS NOT DEMONSTRATED	INEFFECTIVE	ADEQUATE	MOD. EFFECTIVE	EFFECTIVE
APPROPRIATIONS INCREASE	64 programs (34%)	5 programs (18%)	104 programs (47%)	122 programs (53%)	73 programs (59%)
NO CHANGE	47 programs (25%)	1 program (4%)	29 programs (13%)	28 programs (12%)	6 programs (5%)
APPROPRIATIONS DECREASE	80 programs (42%)	22 programs (79%)	86 programs (39%)	81 programs (35%)	45 programs (36%)

Source: Eileen Norcross and Kyle McKenzie, “An Analysis of the Office of Management and Budget’s Program Assessment Rating Tool (PART) for Fiscal Year 2007,” (working paper, Mercatus Center at George Mason University, May 2006): 24, http://www.mercatus.org/repository/docLib/20060831_20060605_mc_gap_partpaper07.pdf.

less funding in fiscal 2006 than in fiscal 2005. Though the correlation is suggestive, much more detailed analysis would be necessary to determine whether performance information affected budget decisions in any category of programs.

WHITHER ACCOUNTABILITY?

“Everyone is entitled to his own opinion, but not his own facts.”

– Sen. Daniel Patrick Moynihan

Since 1993, federal employees have done yeoman’s work defining outcomes and measures. Some are even close to linking results with costs. While the agencies’ performance and accountability reports have not all improved as fast as we had hoped, the improvements are real and tangible.

Effective accountability requires that agencies present full, concise, relevant, accurate, and reliable disclosure of the benefits created for the public, as well as the costs of producing those benefits. Disclosure of goals, outcomes, and costs should not be controversial. In fact, bipartisan majorities passed GPRA in Congress.

Similarly, the idea of using this information to make budget and program decisions should not be controversial. Indeed, there is little alternative if government decisions are genuinely to advance the public interest. If policy makers do not know what results a program is supposed to accomplish, whether the program accomplishes those results, and what alternatives were forgone to produce the results, they are flying blind.

Nevertheless, many congressional oversight and appropriations committees have shown scant interest in using the performance information to make decisions on program design and budgeting. Republicans and Democrats, liberals and conservatives, may rightfully disagree based on values, priorities, or honestly different assessments of whether particular results are worth the cost. But surely they could muster a bipartisan consensus to examine the performance information before they decide.

STRONGEST AND WEAKEST SCORES

T TRANSPARENCY

TABLE 6: TRANSPARENCY SCORES

	CRITERION				CATEGORY
	①	②	③	④	TOTAL
TRANSPORTATION	5	5	4	5	19
LABOR	5	4	3	4	16
VETERANS	5	5	4	5	19
STATE	5	5	3	4	17
USAID	5	5	3	3	16
GSA	5	4	3	3	15
JUSTICE	5	4	3	3	15
COMMERCE	5	3	2	3	13
EDUCATION	5	4	2	3	14
EPA	5	2	2	3	12
AGRICULTURE	2	4	2	4	12
TREASURY	5	4	2	3	14
ENERGY	5	3	2	2	12
NRC	2	4	3	3	12
SSA	5	3	2	3	13
DEFENSE	5	4	2	3	14
NASA	5	3	2	2	12
INTERIOR	5	3	3	3	14
NSF	3	3	4	3	13
SBA	1	3	3	3	10
HUD	4	2	3	2	11
DHS	5	3	2	2	12
OPM	5	3	3	2	13
HHS	3	2	2	2	9

This section lists the reports that scored the highest and lowest for each of our 12 evaluation criteria and briefly explains what made them the strongest or the weakest. The discussion of each criterion begins by describing the key factors we look for under that criterion and why we consider them important. Each ends by highlighting the main improvement opportunities based on our review of the reports this year.²⁵

1. IS THE REPORT EASILY ACCESSIBLE AND EASILY IDENTIFIED AS THE AGENCY'S ANNUAL PERFORMANCE AND ACCOUNTABILITY REPORT?

Access to performance information is critical because public accountability can only occur if members of the public can actually find out what benefits an agency provides. The annual report should be easily available to the public, stakeholders, and the media. Ideally, this means that the agency's home page displays a link clearly guiding the reader to the annual report for the most recent fiscal year. If one has to be an expert on performance management, the agency's structure, or the structure of the agency's Web page to locate it, the spirit of accountability to the public is not satisfied. If the report is large, it should be divided into sections for more convenient reading and/or downloading. Making the report available in multiple formats is also desirable since readers' needs vary and each format has its advantages and disadvantages (e.g., ease of printing, searching, etc.). Finally, the agency should include contact information so that people can mail, phone, or e-mail questions, comments, or requests for a hard copy of the report.

Reports were due to the president and Congress by November 15. The Mercatus research team gave agencies an additional month, until December 15, before checking to see if the reports were available on agency Web sites. Links to agency reports can be found on the Mercatus Web site at <http://www.mercatus.org/Scorecard/agencies>.

As the scores demonstrate, this is by far the easiest criterion for agencies to satisfy. With minimal effort and a decent webmaster, any agency can score a 5 here, and 18 of the 24 agencies did. Once again this year, we had little difficulty accessing and downloading the reports. All but one agency (SBA) posted their reports on their Web sites by our December 15 cutoff date. The SBA report was posted within a week after December 15. Twenty of the 24 agencies had an explicit link directly to the report from their home page. Four agencies

KEY EVALUATION FACTORS

- Did the agency post the report to its Web site on a timely basis?
- Is the report easy to locate on the Web site?
- Is it easy to navigate online and download?
- Does the agency provide prominent and useful contact information?

²⁵ See also the "Performance Report Scorecard Criteria Guidance" we issued last year, which further discusses our evaluation criteria and provides specific examples of good and not so good practices taken from prior year reports. That document can be found on our Web site, http://www.mercatus.org/publications/pubid.2264/pub_detail.asp.

lacked an explicit home page link, but their reports were easily found via intuitive home page links or searches. All reports but one could be downloaded in either a single file or in multiple smaller files. All but two of the agencies provided some sort of contact information online, in the hard copy of the report, or in both places. However, the usefulness of the contact information varied greatly.

STRONGEST SCORES: Numerous

Most agencies satisfied the basic elements of criterion 1 by posting their reports on time, creating a direct home page link to the report, permitting downloads in both single and multiple files, and providing at least some contact information. We continue to see improvement under this criterion since four more agencies than last year earned the maximum score. This includes the 10 agencies with the highest overall scores.

WEAKEST SCORE: SBA

As noted above, SBA did not post its report on its Web site by December 15. Four agencies lost credit for not having an explicit home page link to the report: Agriculture, HHS, NRC, and NSF. Our researchers could find no contact information for Agriculture and NRC either online or in the hard copy versions of their reports. The HUD report lost a point for being downloadable only in one large file.

OPPORTUNITIES FOR IMPROVEMENT

While agencies are doing well under criterion 1, areas for improvement still remain.

Better contact information. Most agencies could enhance the visibility and specificity of their contact information. Five agencies provided contact information only in the hard copy report, with no online contact information separate from the report itself. The hard copy contact information usually appeared at the very beginning or the very end of the report. However, in several cases, it was embedded in the body of the report and therefore hard to find. The contact information for several agencies consisted only of the agency's main address or phone number and had nothing specific to the report. A reader with questions about one of these agency reports probably would have to go through a number of referrals before reaching a knowledgeable agency source. It is much more helpful to provide contact information that steers readers with questions or comments directly to individual employees who worked on the report or at least to the specific agency organization that produced it.

Examples of agencies with prominent and specific contact information include Interior, EPA, and NSF. The Web page containing the Interior report has good organizational contact information via a link captioned "We'd Like to Hear from You." The last page of the hard copy report also encourages feedback and provides contact information by mailing address and e-mail. The Web page for the EPA report invites comments and provides organizational contact information specific to the report. The back cover of the hard copy version lists the same contact information. The NSF Web page containing its report identifies sources to contact for more information on different aspects of the report. The hard copy report provides the e-mail and mailing address of a named employee as a general contact for the report.

More prominent home page links. While most agencies have direct home page links to the report, some of the links are more prominent than others. Several agency report links are rather obscure. For example, the NASA link is located in fine print at the very bottom of the agency’s home page. Agencies with the most prominent home page links include Justice, Labor, Veterans Affairs, and USAID.

Access to prior year reports. Members of the public and others may want to see prior year versions of the agency’s report as well as the current version. We found that a search for “performance report” using each agency’s home page search feature produced at least some prior year versions for all agencies. The prior year reports for some agencies were easily identified in the search results. For other agencies, however, using the search feature was unwieldy and produced hundreds or even thousands of miscellaneous “hits” that needed to be sorted through in order to find a particular prior year report. We also found that, for half of the agencies, the Web page containing the fiscal year 2006 performance report also produced prior year versions, sometimes going all the way back to their fiscal year 1999 report. State, OPM, Veterans Affairs, and SSA are just a few examples. It would be helpful for the remaining agencies to include prior year versions in their report links as well.

2. IS THE REPORT EASY FOR A LAYPERSON TO READ AND UNDERSTAND?

“The UFMS will replace five legacy accounting systems (PSC’s CORE Accounting System, CDC’s TOPS, FDA’s GLAS, NIH’s CAS, and CMS’ FACS) . . .”

– HHS Report, p. I-27.

The annual performance and accountability report is fundamentally a communications device that addresses different audiences with different needs. Some readers are subject matter experts in an agency’s work. They seek an in-depth understanding of the agency’s performance and may be willing to plow through a lengthy, detailed, and technical report to get it. From our perspective, however, the most important readers are the ordinary citizens who pay the bills. They have a right to know in clear and concise language how well each agency is accomplishing the important public outcomes it seeks to achieve on their behalf and in return for their tax dollars. We review the reports solely from this perspective, not as accountants, subject matter experts, or government “insiders.”

Therefore, to do well on our Scorecard, the report’s style, language, and subject matter must reflect the central purpose of communicating to the public. The hallmark of the reports that score the highest in our evaluations year after year is that they effectively communicate important performance results in a way that lay readers can understand. This key trait is relevant to most categories of our scorecard, not just criterion 2. Thus, the best reports tend to score well across the board.

The report should focus on an agency’s mission, how it organizes efforts toward that end, and how much progress was made toward its achievement in the preceding fiscal year. Contents should be clear, logical, easy to navigate, and presented in such a way that the structure aids understanding. Consistent format,

clarity of text, absence of jargon, and effective use of visual techniques like headings, graphs, tables, and photos are helpful. Acronyms can be helpful if they substitute for lengthy proper names that readers may be familiar with, but the use of acronyms to refer to documents, processes, systems, nouns other than proper names, verbs, short names, or names of things known only to insiders inhibits understanding, even if the report provides a list of acronyms. Details can either inform or confuse depending on how they are presented. Anecdotes can promote effective communication if they complement and illustrate, rather than substitute for, outcome-based results measures. Information necessary to evaluate the agency's performance should actually be present in the report, not just referenced as available in some other document, Web page, or computer disk.

Overall, the scores in the category improved modestly. Four reports again earned 5s, and eight reports scored 4s, up one from last year. Last year four reports earned 2s, while this year only three reports did so.

STRONGEST SCORES: State, Transportation, USAID, Veterans Affairs

State and USAID collaborate closely on their reports. In addition to using most of the same goals, their reports also share a number of excellent presentational formats. Both reports are visually appealing and make effective use of tables and graphics. The well organized, concise, and informative tables used to present results in the joint performance section for each report are particularly notable. The narrative portions of the reports are generally clear and easy for the lay reader to understand. USAID uses a manageable total of 35 performance measures. State's 94 measures are fewer than it had last year but are still too many, the report's one negative from a readability standpoint.

The Transportation report has many of the same positive readability attributes as the State and USAID reports. For example, an introductory portion of the report highlights fiscal year 2006 performance accomplishments. The performance section begins with a helpful "Reader's Guide." A table summarizes the department's performance results in a reader-friendly format. The secretary's transmittal letter is exceptionally substantive and informative. The department uses a manageable 34 performance measures in all.

KEY EVALUATION FACTORS

- Is the report written in plain language that a lay reader can understand?
- Does it use clear presentational formats that the reader can easily navigate and comprehend, including effective tables and graphics?
- Are the narratives clear and concise?
- Does the report feature a manageable number of performance goals and measures that capture outcomes whose significance is apparent to a lay reader?
- Does it limit the use of technical terms, jargon, and acronyms?

The Veterans Affairs report also begins with an excellent transmittal letter from the secretary. Like the others, the report makes excellent use of tables and graphics. The color coding it uses to highlight information presented is particularly effective. Another noteworthy feature is the “Performance Scorecard,” which presents at a glance the department’s strategic goals, objectives, and key performance measures as well as results for the key measures and cross references to the pages of the report dealing with each key measure in more detail. The department has a large number of performance measures (150), but the main body of the report concentrates on 23 key measures.

While the above-cited reports are the best under this criterion, a number of other reports clearly were written with the general public in mind and scored well. Each of the top seven scorers overall and 10 of the top 12 received at least a 4 here.

ACRONYMS WE WISH WE HADN'T HEARD . . .

ACRONYM	MEANING	PERPETRATOR
DASIS-TEDS	Drug Abuse Services Information System Treatment Episode Data Set	HHS
IAQTS	Indoor Air Quality Tools for Schools	EPA
FEAPMO	Federal Enterprise Architecture Program Management Office	Interior
MEPNAB	Manufacturing Extension Partnership National Advisory Board	Commerce
IFMSCFM	Integrated Financial Management System Core Financial Module	NASA
PCECGF	The Pollution Control Equipment Contract Guarantee Fund	SBA
FEDS-HEAL	Deployment-Based Medical Service Contract	Defense
HDHP	High Deductible Health Care Plan	OPM
PEPFAR	President’s Emergency Plans for AIDS Relief	State
SDNT	Specially Designated Narcotics Traffickers	Treasury
APPAS	Associate Performance Planning and Appraisal System	GSA
VOI/TIS	Violent Offender Incarceration and Truth-in-Sentencing	Justice
FLPIDS	Farm Loan Program Information Delivery System	Agriculture
BABAR	B and B-bar Experiment	Energy
RIRIP	Risk-Informed Regulation Implementation Plan	NRC
RZ SDO	Redeterminations Service Delivery Objective	SSA

WEAKEST SCORES: HHS, HUD, EPA

The HHS report is an example of how not to connect with the general public. It is text-heavy and includes few graphics. The pagination system, in which each section has separate numbers, is distracting and inhibits navigation between the sections. The report is rife with acronyms but, unlike most reports, it does not include an appendix defining them. Thus, the reader must search back to the first use of each acronym to discover its meaning. The report covers only 35 performance measures, which is quite manageable. However, it appears that these are only “highlighted” measures taken from a much larger set of performance measures that the department uses. The full set of departmental measures is not described anywhere in the report.

While the HUD report has some good features, it is lengthy and text-heavy as a whole. The performance section consumes over 160 pages. Furthermore, the text is generally verbose. The goals are designated by acronyms rather than names, consecutive letters, or numbers, which is somewhat distracting. The descriptions

IS AN ACRONYM REALLY NECESSARY FOR THIS?

ACRONYM	MEANING	PERPETRATOR
IP	Improper Payments	Several
P3	People, Prosperity, and the Planet	EPA
FBU	Funds can be put to better use	Interior
AD	Antidumping	Commerce
BD	Business Development	SBA
PERSTEMPO	Personnel Tempo	Defense
PIH	Office of Public and Indian Housing	HUD
MINT	United States Mint	Treasury
CW	Cooperating Witness	Justice
RTE	Ready-to-Eat	Agriculture
ROL	Rule of Law	USAID
VOT	Victims of Torture	USAID
CSW	Commercial Sex Workers	USAID
NUREG	Nuclear Regulatory Commission Regulation	NRC
ASA	Average Speed of Answer	SSA
FI	Financial Investigations	DHS

NOT WHAT YOU THINK IT MEANS . . .

ACRONYM	MEANING	PERPETRATOR
LUSTs	Leaking Underground Storage Tanks	EPA
FIST	Facilities Instructions Standards and Techniques	Interior
ASAP	Automated Standard Application for Payments	Commerce
STEREO	Solar Terrestrial Relations Observatory	NASA
COTS	Commercial Off the Shelf	Several
COOP	Continuity of Operations Plan	Several
MANPADS	Man Portable Air Defense System	Several
COLA	Certificates of Label Approval	Treasury
FATS	Foreclosure avoidance through servicing	Veterans
FLITE	Financial and Logistics Integrated Technology Enterprise	Veterans
PC	Plum Curculio	Agriculture
NORAD	Norwegian Aid	USAID
FECA	Federal Employees' Compensation Act	Labor
ROAR	Recovery of Overpayments, Accounting and Reporting System	SSA
SOW	Statement of Work	DHS

NOT THEIR FAULT . . .

(Acronyms that appear in the reports but may have been invented outside the agency.)

ACRONYM	MEANING
SAFETEA-LU	Safe, Accountable, Flexible, Efficient, Transportation Equity Act: A Legacy for Users
USA-PATRIOT	Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism
FREEDOM	Freedom for Russia and Emerging Eurasian Democracies and Open Markets Support Act of 1992

for some performance indicators are obscure, making it difficult to grasp their significance. There are a number of apparent incongruities or typographical errors in the report.

The EPA report is improved over last year’s version in conveying information to lay readers and has some good features. However, it still presents major readability challenges. It is long (over 400 pages), text-heavy, and laden with acronyms. The agency’s performance metrics are spread out over the report, making it difficult for the reader to gain an overall grasp of them. The report uses a complex and somewhat distracting system of pagination. Each section of the report begins with a new sequence of numbers, and the detailed performance section is further divided into two subsections that also designate the goal being discussed on that page. (Examples of page numbers include “Section I-9,” “Section II.1-Goal 1-9,” and “Section II.2-Goal 1-124.”) Perhaps the most pervasive challenge for the lay reader is that the report seems to assume a higher level of expertise in environmental science than most members of the public possess. Many of the measures, as well as the accompanying narratives, are technical and hard to understand. The relationship between the performance goals and their accompanying measures is not always clear.

OPPORTUNITIES FOR IMPROVEMENT

The scores under this criterion have increased significantly in recent years, but opportunities for further improvement remain plentiful.

TOO CUTE!		
ACRONYM	MEANING	PERPETRATOR
GEAR UP	Gaining Early Awareness and Readiness for Undergraduate Programs	Education
TRI-ME	Toxic Release Inventory Made Easy	EPA
REGARDS	Reasons for Geographic and Racial Differences in Stroke	HHS
CIOT	Click It or Ticket	DOT
SMART	State Messaging and Archive Retrieval Toolset	State
MILC	Milk Income Loss Contract	Agriculture
AGILE	Accelerated Growth, Investment, and Liberalization with Equity	USAID
DOLAR\$	Department of Labor Accounting and Related Systems	Labor
eMerge ²	Electronically Managing Enterprise Resources for Government Efficiency and Effectiveness	DHS

Keep the general public in mind. Agencies face many challenges in crafting their performance reports, but writing a report that ordinary citizens can read and understand should not be one of them. Report writers should affirmatively take into consideration how well their product will communicate to lay readers. It is easy to tell which reports were or were not written from this perspective. Reports may need technical content to satisfy the legitimate interests of specialized audiences. However, there are ample opportunities to reach the general public as well. Two good venues for this are the transmittal letter and the Management's Discussion and Analysis section.

Avoid complicated designations of goals and measures and complex pagination. Simply using a straightforward system to designate goals and measures and number pages would significantly help the lay reader (or probably any reader) navigate through the report. It is hard to imagine an easier fix to a wholly unnecessary readability burden.

Focus on a clear and concise set of performance metrics. Some agencies have been quite successful in presenting their performance through a concise set of measures that captures important public outcomes. As noted above, Transportation is a good example. Other examples are Agriculture (38 total measures) and NRC (16). Agencies with more numerous performance indicators can ease the reader's burden by focusing their reports on a smaller set of key measures. As noted above, Veterans Affairs does this; so do GSA and Justice.

Minimize lengthy, text-heavy narratives and maximize tables and graphics. Many reports make effective use of tables, vignettes with photographs, graphs, and similar features to highlight aspects of their performance or drive home key points. One particularly helpful feature is a "report card" or "scorecard" table that summarizes the agency's overall performance at a glance. Examples are Agriculture, HUD, and Veterans Affairs.

Incomplete reports. Four of the reports had significant gaps in the coverage of agency performance. The Defense, HHS, and SBA reports cover only their agencies' key performance measures and have no content for the remaining measures. The GSA report covers all agency measures but includes fiscal year 2006 results for only the key measures, saying that the rest will be reported later. While focusing the report on a set of key measures is generally a good feature, each report should still have *some* content on all Government Performance and Results Act measures for the applicable fiscal year. Gaps in reporting limit the reader's overall understanding of the agency's performance. They also deprive the reader of full information concerning specific aspects of the agency's performance, such as performance trends and shortfalls.

3. ARE THE PERFORMANCE DATA RELIABLE, CREDIBLE, AND VERIFIABLE?

The Reports Consolidation Act requires that the agency head's transmittal letter assess the completeness and reliability of the report's performance and financial data. The transmittal letter must also describe any material inadequacies in the completeness and reliability of the data and what the agency is doing to resolve them.²⁶ Like a similar requirement imposed on the chief executive officers of publicly-held companies, the purpose of this requirement is to ensure that the quality of the underlying data is a management priority and that deficiencies are acknowledged and corrected as quickly as possible.

²⁶ See *Reports Consolidation Act*, U.S. Code 31 (2000), 3516(e).

More generally, the report should indicate the agency's confidence in the quality of the data used to document its results. Since the purpose of gathering these data is to manage programs strategically, one test of their adequacy is whether they are relevant, timely, complete, accurate, and consistent enough to use as the basis for decision-making. Data should be independently validated (i.e., certified as appropriate for the associated performance measure) and verified (i.e., assessed as reliable). Outside verifiers should be able to access the data with relative ease. Sources and descriptions should be provided for all outcome data. If some data cannot be provided to the public in a way that would permit verification, the report should still inspire confidence that some type of independent verification took place.

Individual and overall scores under criterion 3 are among the weaker year after year. Again this year, there were no 5s and only three 4s. Eleven reports again received scores of 2 this year. The reports pay more attention to data issues now than they did a few years ago. However, most agencies still face serious data challenges, and many could do a better job of explaining how they are addressing those challenges. On the positive side, most agencies reported current results for most of their measures again this year although they often had to rely on preliminary data.

STRONGEST SCORES: NSF, Transportation, Veterans Affairs

As described in its report, NSF undertakes the most extensive efforts of any agency to validate and verify its performance measures and data. Specifically, the agency contracts each year for a comprehensive external validation and verification review. The results of the review for fiscal year 2006 verified the reliability of the agency's measurement processes and validated the accuracy of data for all of the agency's strategic goals and 21 of its 22 annual performance goals. The external reviewer could only partially verify the reliability of the process for the remaining goal, but expressed the belief that the reported outcome for this goal was consistent with the data collected.

The Transportation report includes a Performance Data Completeness and Reliability section that gives a useful overview of the department's system for assessing data. The report provides considerable detail concerning the data for each measure. While the department received a qualified opinion on its financial statements for fiscal year 2006, the qualification does not seem to cast doubt on the department's financial or performance data in general. In fact, the inspector general credits the department with substantially improving its ability to track and report financial results.

The Veterans Affairs report provides comprehensive background on data used for its 23 key measures, including data definitions, sources, any limitations, and methods to verify data and validate the measures. The report includes fiscal year 2006 results for all key measures.

WEAKEST SCORES: Numerous

Typical weaknesses include vague or nonexistent descriptions of data verification and validation procedures, vague or ambiguous assessments of data quality in transmittal letters, data assurances by agency heads that are inconsistent with Office of Management and Budget guidance or apparently contradicted by material presented elsewhere in the report, missing data, excuses rather than solutions for data quality problems, failure to cite data sources, disclaimers or qualified opinions from auditors, and data quality problems identified by agency inspectors general.

KEY EVALUATION FACTORS

- Does the report fully and candidly assess data quality?
- Does it indicate that the agency has major data challenges?
- If so, does it provide confidence that the challenges are being addressed?
- Does the report include at least preliminary results for all measures?
- Does it provide data definitions and sources for its performance measures?

The Defense transmittal letter, for example, states that its report contains “the most complete, reliable financial and performance information available.” The body of the report reveals massive weaknesses in the department’s financial management systems that not only result in disclaimers of opinion on its financial statements each year but also detract from the quality of performance data. The HHS report lacks fiscal year 2006 results for over half of the few performance measures it covers because of data unavailability. The independent auditor’s report states: “The control processes in place to ensure the accuracy of the HHS Performance and Accountability Report are not working as intended by management.” The Agriculture transmittal letter asserts that the report’s financial performance data are complete and accurate in accordance with applicable guidance, but the report offers few specifics concerning measures taken to ensure the quality of the performance data. The Education report lacks 2006 results for 25 percent of its key measures, and the EPA report lacks 2006 results for about 40 percent of all of its measures.

OPPORTUNITIES FOR IMPROVEMENT

Most federal agencies face serious challenges in developing complete and reliable performance data. This is particularly true of agencies that seek to achieve broad national outcomes that are difficult to measure and attribute, especially in annual increments. Indeed, the better an agency’s performance metrics are in outcome orientation, the more difficult the measurement challenges may be. Another challenge confronting many agencies is the need to rely on third party data sources.

Recognizing that these challenges exist, our evaluations under criterion 3 place considerable weight on the agency’s candor in disclosing data weaknesses and whether the report provides confidence that the agency is doing its best to address them. From this perspective, two recurring problem areas raise red flags and call out for improvement.

More candid agency head data assessments. The agency head transmittal letters for almost half the reports asserted that the report’s data were complete and reliable in accordance with OMB’s guidance notwithstanding the fact that the report lacked even preliminary fiscal year 2006 results for some measures or revealed what appeared to be significant data reliability issues. This suggests that the OMB guidance is being widely misapplied and

may need to be clarified or tightened up.²⁷ The following are some examples of reports with transmittal letters claiming complete and reliable data consistent with OMB’s guidance:

- Agriculture (inspector general’s list of major management challenges includes performance data problems affecting one component; one result missing)
- Education, EPA, and Interior (many results missing)
- Homeland Security (disclaimer of opinion; inspector general reports performance data problems at one component)
- Treasury (inspector general reports major data reliability problems at the Internal Revenue Service; results missing for one measure)
- SSA (inspector general audits found data unreliable for some performance measures; results missing for two measures)

Fuller data disclosure. Many reports simply do not provide enough information to give the reader a sense of how reliable the data are or even what data the agencies are using. Agencies can enhance the credibility and usefulness of the reports by providing data definitions and sources for individual measures. Readers might also find it helpful if agencies described in at least general terms what they do to verify and validate performance data. Labor, State, Justice, Transportation, and Veterans Affairs are examples of reports that already do this. The Justice report’s narratives for each performance indicator include a useful box that discusses data definitions, validation and verification, and any applicable data limitations. Labor takes it one step further and has developed a system for rating the completeness and reliability of the data underlying each performance goal. The results are discussed in the performance section, along with any shortcomings and planned improvements. This is a potential best practice.

4. DID THE AGENCY PROVIDE BASELINE AND TREND DATA TO PUT ITS PERFORMANCE MEASURES IN CONTEXT?

The “bottom line” for citizens is whether an agency’s actions make a given situation better or worse. To provide this information, agencies must design measurement systems that facilitate analysis of trends over time. Data should be displayed in a way that allows readers to detect and understand their significance easily. Both quantity of data (years of data included) and presentation matter. Good performance measures that have limited data (due to newness or revision) may convey more information than inferior measures with more data points that are not clearly linked to an agency’s results.

Multiple years of data help identify trends, but they do not by themselves show how close the agency is to achieving its goals, or explain why the agency will produce a significant level of public benefits if it hits its targets. Reports should explain rationales behind the selection and alteration of quantitative targets, so the reader can understand the magnitude of the agency’s goals in relation to the size of the problem.

²⁷ OMB guidance on performance report data is in section 230.2(e) of OMB Circular No. A-11. Among other things, it states that in order for performance data to be considered complete, the report must include at least preliminary results for every measure.

There was little movement this year in the scores for criterion 4. The two reports which received a 5 this year also were the only two to get 5s last year. Six agencies received scores of 2, a modest improvement over last year when seven agencies scored 2s.

STRONGEST SCORES: Transportation, Veterans Affairs

Both of these reports are rich in baseline and trend data as well as other information and analysis to put the data in context. The Transportation report has a summary table that provides prior year data going back to fiscal year 2000. The tables accompanying individual performance measures show prior year targets as well as actual results. The narratives in the performance section provide additional useful analyses of performance results and trends, such as a “performance forecast” indicating whether the targets for individual measures are likely to be met in fiscal year 2007. A review of the baseline and trend data indicates that the department sets more ambitious targets from year to year.

In addition to showing the fiscal year 2006 targets and results for its measures, the Veterans Affairs report states whether or not performance improved from the prior fiscal year. Tables in the performance section show performance data going back to fiscal year 2002 for all 150 performance measures. The tables include the department’s long-term strategic target for each measure, which provides additional valuable context. The narratives in the performance section elaborate on the significance of the targets and performance trends over time for the key measures and also for some of the supporting measures. A review of this information indicates that the department challenged itself with rigorous fiscal year 2006 targets in relation to prior year results.

WEAKEST SCORES: Numerous

We see many recurring weaknesses from year to year. Some agencies simply report whether or not performance measures were met in prior years without providing quantified data on actual prior results. Examples are Energy and NASA. Such information, standing alone, does nothing to enlighten the reader about per-

KEY EVALUATION FACTORS

- Does the report provide prior year results for comparable measures so that the reader can identify performance trends and track progress from year to year?
- Does it describe long-term goals so that the reader can assess the extent of progress?
- Does it explain the rationale for annual targets, particularly where there are significant differences between current targets and prior year results?
- Do the trend data indicate that the agency is setting challenging targets in relation to past performance?

formance trends. Some reports, such as Homeland Security and OPM, have limited prior year data that are relevant to their current measures. Several agencies set current performance targets for some measures below prior year performance levels. Absent an explanation, this practice suggests that they are not challenging themselves to improve. Examples are HUD and USAID. Similarly, targets for some NSF measures have remained unchanged for a number of years; in one case, the annual target is consistently lower than the prior year results.

OPPORTUNITIES FOR IMPROVEMENT

Data showing prior year actual performance results for comparable measures provide the minimal information needed to identify performance trends and afford the reader some sense of how challenging the agency's current targets are. The better reports have two additional features that make the baseline and trend data more insightful and put the agency's performance in fuller context.

Explanations for targets. One is an explanation of how the agency sets performance targets. Explanations are particularly important where the trend data indicate that the agency's current targets are not challenging in relation to past performance or where there are large fluctuations from year to year between targets and actual results. The Education report's presentations of individual performance results use an excellent format that includes an analysis of progress, a description of target context, and often additional background information relevant to the measure.

Long-term perspective for performance trends. The second feature is providing a sense of the agency's long-term goal in relation to each performance measure. Simple trend data can indicate whether the agency is making progress from one year to the next. However, it does not tell the reader where the agency hopes to go over the long term and thus enable the reader to assess the *extent* of progress toward its destination. The EPA report is an example of one that could use more information on long-term goals in order to put annual incremental progress in context. Transportation and Veterans Affairs do a good job of this; so do Labor and State. In the Labor report, each goal is accompanied by a discussion, captioned "Program Perspective and Logic," which provides background concerning the significance of the goal as well as performance trends and future plans. These discussions also cover external trends. The narratives in the performance section of the State report, particularly those under the headings "Justification" and "Impact," provide additional useful context for assessing performance results and progress over time.

B PUBLIC BENEFITS

TABLE 7: PUBLIC BENEFITS SCORES

	CRITERION				CATEGORY
	5	6	7	8	TOTAL
TRANSPORTATION	5	4	4	4	17
LABOR	5	4	4	5	18
VETERANS	4	4	4	4	16
STATE	5	4	5	4	18
USAID	5	3	3	4	15
GSA	3	3	3	5	14
JUSTICE	3	3	3	1	10
COMMERCE	2	2	2	4	10
EDUCATION	3	3	3	1	10
EPA	3	3	3	2	11
AGRICULTURE	3	3	3	1	10
TREASURY	3	2	2	2	9
ENERGY	2	2	2	4	10
NRC	4	3	3	1	11
SSA	2	2	2	1	7
DEFENSE	3	2	2	1	8
NASA	2	2	2	1	7
INTERIOR	2	2	2	2	8
NSF	2	2	3	1	8
SBA	3	2	2	2	9
HUD	3	3	3	1	10
DHS	2	2	2	2	8
OPM	2	2	2	1	7
HHS	3	2	2	1	8

5. ARE THE GOALS AND OBJECTIVES STATED AS OUTCOMES?

An “outcome goal” is defined as the intended benefit (or harm avoided) that results from an agency’s programs or activities. It should be articulated in clear and simple terms that describe benefits to the community rather than activities that are presumed to be of value. Vague verbiage that emphasizes things an agency does instead of why it is doing them should be avoided. This admonition applies at all goal levels—strategic goals, objectives, and annual performance goals.

Strategic goals should be few in number (three to five). Management goals (including financial, human resources, information technology, etc.) exist to support the achievement of genuine outcome goals. A majority of agencies now enunciate at least one strategic goal focused on management improvement, but these should not be confused with actual outcome goals.

Methods, measures, and targets are different from goals, since they are expected to change. By comparison, goals (if selected and stated properly) are likely to remain valid over several years at least. Overly specific goal statements make trend analysis difficult since goals (not just quantitative targets) might change each year.

This particular criterion is crucial for effective performance reporting, and it is also crucial for earning a high score. Agencies typically enunciate strategic goals, performance goals or objectives, and measures. A report’s success on this and the next two criteria depends in large part on whether it has enunciated outcome-oriented goals and measures; it is very difficult to do well on these three criteria without having outcome-oriented goals and measures.

Criteria 5, 6, and 7 are quite closely related. It is difficult to achieve a high score on criterion 7, which assesses whether the report has demonstrated that the agency has an effect on outcomes, if the report has not articulated clear outcomes (criterion 5) and outcome-oriented performance measures (criterion 6). Similarly, it is difficult, though not impossible, to score high on criterion 6 without scoring high on criterion 5. A high score on 6 combined with a low score on 5 would indicate that the report managed to articulate good, outcome-oriented measures even though its goals were not stated as outcomes.

KEY EVALUATION FACTORS

- Are strategic (long-term) and annual performance goals expressed as measurable outcomes that a lay reader can understand and relate to public benefits?
- If strategic goals are expressed at a high level of generality, are they accompanied by strategic objectives that provide more specific and measurable outcomes?
- Do the goals cover all of the agency’s core missions?

The scores for criterion 5 are largely the same as last year. The four reports that scored 5s this year are the same as last year. Eight agencies—the same number as last year—received 2s on this criterion.

STRONGEST SCORES: Labor, State, Transportation, USAID

Labor’s strategic and performance goals are highly outcome-oriented. All four strategic goals are clearly stated as outcomes. Of Labor’s 28 performance goals, 25 capture end outcomes or intermediate outcomes. State clearly phrases its programmatic strategic goals and annual performance goals as intermediate or end outcomes. Even its management-related strategic goal and some of its performance goals focus directly on enhancing the department’s capacity to accomplish its mission outcomes. One management-related outcome performance goal reads: “Personnel are safe from physical harm and national security information is safe from compromise.” USAID shares a common set of programmatic strategic and performance goals with State. Four of Transportation’s five programmatic strategic goals and most of the strategic objectives under them are outcome-oriented. The report states that transportation safety is the department’s top priority, and indeed, the safety strategic goal and its strategic objectives contain the most clearly stated outcomes. While the security strategic goal is not outcome-oriented, the strategic objectives under this goal are stated as intermediate outcomes.

WEAKEST SCORES: Numerous

Some reports have strategic goals that refer to outcomes but at too high a level to enable the reader to understand how and what the agency contributes. These reports are also short on results-oriented annual performance goals that might flesh out the measurable outcomes.

The Commerce report’s first strategic goal, for example, refers to such vague and high-level outcomes as “maximize” United States competitiveness and “enable” economic growth. It is just as difficult to find clear and measurable outcomes in the remaining two programmatic strategic goals. Few of Commerce’s performance goals are stated as outcomes, and some are vague: “Serve society’s needs for weather and water information” and “Understand climate variability and change to enhance society’s ability to plan and respond.”

Energy and Homeland Security are additional examples of reports whose strategic goals capture some outcomes but at high levels of generality that do not clearly specify measurable outcomes. Homeland Security’s management-related strategic goal and its objectives are exceptionally weak, reading more like a string of clichés: “Value our most important resource, our people. Create a culture that promotes a common identity, innovation, mutual respect, accountability and teamwork to achieve efficiencies, effectiveness and operational synergies.” The Energy and Homeland Security reports have few outcome-oriented performance goals.

Other reports lack strategic goals that could pass for even high-level outcomes. Interior’s strategic goals are stated primarily as activities rather than outcomes, for example, “Protect the Nation’s natural, cultural and heritage resources” and “Provide recreational opportunities for America.” NASA’s strategic goals are primarily statements of specific tasks to be accomplished, as are most of the agency’s 37 “multi-year outcomes.” OPM’s strategic goals read more like marketing slogans than objectively measurable goals of any kind. For example: federal agencies will be “employers of choice;” they will have “exemplary” human resources management practices; OPM will be “a model of performance” for other agencies; it will be a “leader in the human resources professional community” and will “have positive name recognition;” OPM will have “constructive and positive relationships with external stakeholders.”

OPPORTUNITIES FOR IMPROVEMENT

The Government Performance and Results Act has been on the statute books for over a decade. Agencies have completed eight annual cycles of planning and reporting under the Act, including several rounds of strategic plan updates. At this stage, it is disappointing that only one-quarter of the 24 agencies have strategic and performance goals that could be rated above average and that the goals for one-third of the agencies are seriously deficient. The scores are even lower for the following two criteria, which, along with criterion 5, go to the heart of the performance reports.

Any agency should be able to develop a set of long-term strategic goals that capture the important public outcomes they exist to achieve in a clear and specific way that an ordinary citizen can grasp. Some of the agencies that score poorly here could improve by breaking out their current high-level strategic goals into more specific and measurable outcomes that could be tied with some credibility to the agency's contributions. Others, such as OPM, may need to rethink their basic approach to strategic planning and performance accountability.

It is challenging for agencies to develop performance goals that convert outcomes into annual measurable increments. One approach that some agencies use is to include intermediate outcome goals. For example, increased use of seat belts is an intermediate outcome toward the end outcome of fewer highway deaths and injuries. Not all annual performance goals need to be outcomes or even intermediate outcomes. Output, activity, and efficiency goals have their place and may be important to some stakeholders. However, agencies should strive to have at least a readily identifiable set of outcome-oriented performance goals in order to demonstrate the agency's value to the general public in a credible way.

6. ARE THE PERFORMANCE MEASURES VALID INDICATORS OF THE AGENCY'S IMPACT ON ITS OUTCOME GOALS?

Performance measures selected by an agency should relate directly to its outcome goals. Activity measures, such as number of participants, studies completed, facilities built, projects funded, etc. may contribute to achievement of a result, but do not constitute results of interest to the public at large. Including these measures in the report may actually detract from the report's effectiveness in demonstrating the agency's impact. Data measuring levels of output can support a claim for success, but only if the agency makes a compelling case for a causal link between the output and results achieved.

The same four reports received 4s this year as last year. Once again, no report earned a 5. Twelve reports scored 2s on this criterion, a slight improvement over last year's unlucky 13.

STRONGEST SCORES: Labor, State, Transportation, Veterans Affairs

About 60 percent of Labor's annual performance measures are outcome-oriented. The outcome orientation is particularly strong for the measures applicable to strategic goals 1 (prepared workforce) and 3 (quality workplaces). State's performance measures consistently align well with its goals, and about half of the measures capture intermediate or end outcomes. Of the 29 programmatic measures in the Transportation report, 23 capture end outcomes or intermediate outcomes. One good example is: "Percent of days in the shipping season that the U.S. portion of the St. Lawrence Seaway is available." Veterans Affairs uses 23 key measures of which about one-third are clear

KEY EVALUATION FACTORS

- Do the measures and targets relate directly to outcome goals and are they expressed as outcomes?
- Where measures are not stated as outcomes, do the accompanying narratives explain their relevance to mission outcomes?
- Where measures and targets are technical or otherwise unclear to a lay reader, do the accompanying narratives explain their significance?

outcomes. A number of the non-key measures relating to timely and accurate benefit determinations of various kinds could be regarded as outcome-oriented, since making such determinations is a core departmental mission.

WEAKEST SCORES: Numerous

The common thread running through the lower-scoring reports is that most of their performance measures address activities, outputs, processes, and efficiencies rather than outcomes. Some agency measures, like the SBA's, focus heavily on outputs, such as levels of assistance provided, but have less to demonstrate when they come to what public benefits the assistance achieves. Some reports use activity measures that seem unrelated to outcomes. With respect to the SSA report, for example, it is unclear how measures expressed as raw numbers (e.g., number of disability claims or hearings processed) affect the agency's outcome goals. Efficiency, workload, productivity, and customer satisfaction measures may be valuable for certain purposes, such as internal agency management. However, they bear little direct relationship to public outcomes. In some instances, operational measures are so modest and routine that it is hard to understand how they qualified for inclusion in a performance and accountability report. For example, OPM's "key" measures include joining two human resources professional organizations and acknowledging receipt of congressional inquiries within 24 hours.

OPPORTUNITIES FOR IMPROVEMENT

As the scores indicate, there is much room for improvement in the outcome orientation of agency performance measures. Not all performance measures need to be outcome-oriented. Different types of measures may be important for different agency stakeholders. However, in order to persuasively demonstrate the agency's value to the general public, the report should have a core set of specific and understandable outcome measures to implement a set of specific and understandable outcome goals.

In addition to developing more outcome measures, agencies might consider eliminating from their reports measures that have little if any relevance to the public or other external stakeholders. For example, workload, productivity, and many other process-related measures seem relevant primarily, if not exclusively, to agency

officials for internal management purposes. Such measures contribute only to the size of the performance report. “Customer satisfaction” measures may fall into this category. They sometimes appear to be treated as a substitute for true outcome-oriented measures that get more directly at whether an agency is meeting the public’s needs. For example, the SBA and Treasury reports classify customer satisfaction measures as “outcome” measures. Over-reliance on such measures is another potential issue. Many agencies have customer satisfaction measures, and some, such as GSA, use them extensively.

7. DOES THE AGENCY DEMONSTRATE THAT ITS ACTIONS HAVE ACTUALLY MADE A SIGNIFICANT CONTRIBUTION TOWARD ITS STATED GOALS?

The report should show whether things improved because of what an agency did, and if so, how much of the improvement can be attributed to its actions. Program evaluations or other empirical evidence of a cause/effect relationship should support claims of impact. A less desirable alternative would be to logically connect outcome measures, output measures, and anecdotal evidence. A case that rests on merely assumed cause/effect relationships is unsatisfactory. The report should explain how agency outputs create or enhance outcomes for the public and describe the nature and extent of influence so that outcomes can be attributed (at least in part) to specific agency actions.

Discussion of the operating environment and the extent of the agency’s influence are helpful in keeping expectations ambitious, yet realistic. External factors, however, should be treated as influences that must be controlled for in order to identify the agency’s contribution—not excuses for a failure to demonstrate performance.

State was the only agency to earn a 5 this year, down from four agencies last year. Eleven reports received scores of 2 on this criterion, compared to last year’s nine.

KEY EVALUATION FACTORS

- Does the report have overall performance metrics, strong goals (criterion 5), and strong measures (criterion 6), which combine to provide a clear basis for assessing agency performance toward end outcomes?
- Do the narratives in the performance section help the reader understand the impact of the performance results, particularly where their significance is not self-evident?
- Does the report demonstrate a causal connection between the agency's actions and the reported results?
- Does the report provide sufficient current results, prior year data, and other information to give the reader a complete picture of the agency's performance?

STRONGEST SCORE: State

State’s performance metrics, particularly the goals component, are strongly outcome-oriented. The narratives accompanying the individual measures consistently do a good job of demonstrating the significance of the performance results toward achieving the department’s goals. Throughout the performance section, the report lists specific accomplishments and explains how those results affect the department’s missions. The three reports that scored 4s on this criterion—Labor, Transportation, and Veterans—also combine strong performance metrics with informative narratives in the performance section to demonstrate accomplishments in a persuasive way.

WEAKEST SCORES: Numerous

The predominant feature of the low-scoring reports is weak performance metrics—specifically, the absence of significant outcome orientation in the agency’s goals or measures. Most of the reports lack outcome orientation in both their goals and measures. Other problem areas include weak narratives to accompany the performance metrics (e.g., SBA, Treasury), limited baseline and trend data to put performance in context (e.g., Energy, Homeland Security, NASA), and technical measures that are difficult for a lay person to understand (e.g., NASA, Treasury). The incomplete reports also pay a price here.

OPPORTUNITIES FOR IMPROVEMENT

The improvements suggested in a number of other evaluation criteria will pay dividends here as well. The most directly relevant are criteria 5 and 6, but criteria 2 through 4 are also relevant.

8. DID THE AGENCY LINK ITS GOALS AND RESULTS TO COSTS?

In some ways, this is the most crucial criterion of all for effective government accountability. When results are linked to costs, decision makers know how much they have to pay per successful result. Armed with this information, they can assess the opportunities forgone when resources are allocated to less effective programs, and they can estimate how much more could be accomplished if resources were reallocated to the most effective programs.

This year’s scoring roughly corresponds to the following standards: 1, the report has no relevant content or links budget resources only to strategic goals; 2, the linkage is carried down only to the strategic objectives; 3, there is some useful content beyond the strategic objective level but short of the performance goal level; and 4, the linkage is carried down to the individual performance goals. To earn a 5, the report needs to link budget costs to individual performance measures.

KEY EVALUATION FACTORS

- Does the report link its budget resources to individual performance measures, thereby permitting cost-effectiveness analysis?
- Do the linkages include all budget resources, including indirect costs?

As in past years, the scores on this criterion remain the lowest—an average of just 2.29 points. On the positive side, however, some reports are beginning to show progress on linking budget costs to individual performance measures. For the first time, two agencies earned 5s on this criterion by linking budget costs to a number of their individual performance measures. One other agency links its costs to a few individual measures. Eleven reports, on the other hand, received scores of 1 on this criterion. Another five reports received 2s. These results at the low end represent only slight improvement from last year. Most agencies are far from able to tie their costs to their performance results in a meaningful way. Unfortunately, many do not appear to be making a serious effort to do so.

STRONGEST SCORES: Labor, GSA

These two agencies made major strides to link their budget resources to individual performance measures this year. The Labor report links budget resources to many, but not all, individual performance measures. The report states that additional linkages will be made in the future, but that some measures do not lend themselves to individual cost linkages. The GSA report has a table linking projected fiscal year 2006 budget costs to all individual performance goals and, with very few exceptions, to the individual measures for the performance goals. Most of the performance goals have only one measure, but a few have multiple measures. The table notes that the cost allocations are preliminary. It does not specify whether the cost allocations account for the entire fiscal year 2006 budget.

One other agency, Veterans Affairs, allocates budget costs to a few individual performance measures. Its report links costs to three key performance measures and states that future reports will increase the number of measures for which cost data are provided. Five agencies link budget resources to their performance goals but not to individual performance measures: Commerce, Energy, State, Transportation, and USAID.

WEAKEST SCORES: Numerous

Three of the reports that scored 1s on this criterion fail to link budget resources to their performance metrics at any level: Defense, Education, and HHS. The Defense and HHS reports simply have no content on budget-cost linkages. The Education report gives several methodological reasons why the department believes it cannot link performance to funding. The other agencies that scored 1s link budget costs only to their strategic goals. The agencies with scores of 2 take the linkage down to the strategic objective level but no farther.

OPPORTUNITIES FOR IMPROVEMENT

With respect to this criterion, agencies can be divided into two camps: those that are making good progress toward linking costs to performance metrics in a meaningful way and those that are not. The former have scores of 5 or 4; the latter have scores of 1 or 2. There are no scores of 3 under this criterion. For the two-thirds of the agencies that earned 1s or 2s, the most obvious improvement strategy is to at least make a serious effort. Developing linkages to the point of permitting meaningful and credible comparison of costs with outcomes presents methodological challenges. It may be impractical or even impossible for some agencies to allocate costs to every individual performance measure. The goal should be to find the right unit of analysis for the linkage so that the agency can assess performance results in relation to funding. Ideally, the agencies that are leading the way now will resolve or ameliorate the challenges and develop best practices for the rest to follow, and the rest will try harder in the future.

L FORWARD-LOOKING LEADERSHIP

TABLE 8: LEADERSHIP SCORES

	CRITERION				CATEGORY
	9	10	11	12	TOTAL
TRANSPORTATION	5	4	4	4	17
LABOR	4	4	5	4	17
VETERANS	5	4	3	4	16
STATE	5	4	2	4	15
USAID	3	2	3	3	11
GSA	4	2	2	3	11
JUSTICE	3	3	3	3	12
COMMERCE	3	3	3	4	13
EDUCATION	3	3	3	3	12
EPA	3	3	3	4	13
AGRICULTURE	3	3	3	4	13
TREASURY	3	3	3	3	12
ENERGY	3	3	3	3	12
NRC	4	2	2	3	11
SSA	3	3	4	3	13
DEFENSE	3	2	2	3	10
NASA	3	4	2	4	13
INTERIOR	3	2	2	2	9
NSF	3	2	3	2	10
SBA	3	2	4	3	12
HUD	3	2	2	2	9
DHS	2	3	2	3	10
OPM	2	2	2	2	8
HHS	2	2	2	2	8

9. DOES THE REPORT SHOW HOW THE AGENCY'S RESULTS WILL MAKE THIS COUNTRY A BETTER PLACE TO LIVE?

Does an agency realize and articulate the value it provides to the country? The report should speak directly to the public about how the agency produces benefits that are important to citizens. Politics have no place in this report. The public's interests are paramount, not individual or partisan credit or blame. Just as the best corporate reports feature communication directly from the chief executive, agency reports should demonstrate accountability of agency heads for their organization's performance. Outcome orientation, sound strategies, and successful achievement discussions should support lofty ideals. The report should create confidence in an agency's ability to improve America's future. Anecdotes and success stories can be important communication strategies in this regard, but their value is limited if solid performance data does not back them up.

KEY EVALUATION FACTORS

- Do the report narratives describe public benefits that flow from the agency's work in a way that lay readers can understand and appreciate?
- Does the report illustrate the agency's public benefits with specific outcome-oriented examples of accomplishments?
- Does the report back up its narratives with strong performance metrics that affirmatively demonstrate the public benefits the agency achieves?

This year's evaluations under criterion 9 produced a wider range of scores than last year, when all reports earned either 3s or 4s. Three reports scored 5s this year, and three received 2s. The research team's evaluation under this criterion tends to focus on the narrative portions of the report, such as the transmittal letter, the Management's Discussion and Analysis (MDA) section, and other features that highlight specific agency accomplishments in a persuasive and reader-friendly way. The highest scores this year reflect the many excellent examples we found in several reports. The lower scores for some reports reflect our view that narratives, no matter how good, cannot offset weak performance metrics. Thus, a report with below average scores under criteria 5 through 7 is very unlikely to earn an above average score here.

STRONGEST SCORES: State, Transportation, Veterans Affairs

The State report's combination of strong performance metrics and highly informative narratives does a very effective job of demonstrating the public benefits that the department achieves. To drive this home, the narratives for each strategic goal include specific "public benefit" descriptions that are informative and persuasive. The narrative portions of the Transportation report are informative and persuasive in showing the benefits

that flow from the department's work, particularly in the areas of transportation safety. Even where results fell short of their targets, the report demonstrates the department's commitment to improve. The secretary's transmittal letter, the performance highlights section, and the descriptions of in-depth accomplishments accompanying the performance measures are particularly noteworthy. The department's strong performance metrics re-enforce the strong narratives. The narrative portions of the Veterans Affairs report do an excellent job of laying out the department's results in a clear and compelling way. The transmittal letter is very strong in this regard, as are the table on most important achievements, the "public benefit" sections for each strategic goal, and the vignettes illustrating specific accomplishments. The generally strong performance metrics reinforce the narratives.

Though it received a 4 rather than a 5, the GSA report deserves mention as one that does a commendable job of demonstrating public benefits for an agency that has little direct contact with the general public. The narratives describe how GSA's efforts enhance the ability of its customer agencies to achieve results that directly benefit the public and emphasize GSA's efforts to maximize return to the taxpayers. (This contrasts with the report from the other government support agency, OPM, which had one of the weakest scores on this criterion.)

WEAKEST SCORES: HHS, Homeland Security, OPM

The performance metrics in the HHS report do little to demonstrate public benefits since they are incomplete, generally weak in outcome orientation, and lack fiscal year 2006 data for many measures. The narratives describing highlights under each strategic goal focus more on processes and activities than results.

The importance of Homeland Security's mission is obvious. However, the lack of results orientation—particularly in the performance metrics, but also in the narratives—leaves the reader with little sense of concrete results that flow from the department's work.

The OPM report does virtually nothing to demonstrate that it is accomplishing its core mission of ensuring that the federal government has an effective workforce. The agency's performance metrics are exceptionally weak, and the narratives are not much better. Even the director's transmittal letter calls into question the agency's commitment to its core mission by stating that her top priority is improving the administration of the Federal Retirement System. This re-enforces the impression left by the remainder of the report that OPM is buried in administrative process rather than providing leadership to achieve important federal workforce outcomes.

OPPORTUNITIES FOR IMPROVEMENT

Maximize opportunities to tell the agency's story to the public. The narrative portions of the report, whose formats are less constrained than some other portions, offer many opportunities to address the general public. The higher-scoring agencies have found effective ways to take advantage of these opportunities. They use the agency head's transmittal letter as more than a perfunctory document that recites the necessary legalisms. The best transmittal letters highlight important outcome-oriented accomplishments. Other narrative portions in the better reports describe performance highlights in a way that the general public can understand and appreciate. They frequently include vignettes, which illustrate how the agency's programs have benefited specific individuals. The narratives can be particularly useful for agencies that have difficulty translating their performance accomplishments into clear and specific annual outcomes. This includes research and development

agencies, agencies with highly technical missions, and agencies whose broad outcome accomplishments take years to mature.

Improve performance metrics. As noted previously, even the most compelling narratives are not a substitute for solid performance goals and measures. While an agency’s performance metrics are covered primarily by other evaluation criteria, a report cannot earn a strong score in criterion 9 with weak performance metrics.

10. DOES THE AGENCY EXPLAIN FAILURES TO ACHIEVE ITS GOALS?

If an agency cannot identify reasons for failure, its ability to claim credit for success is suspect. Successes and failures that really matter occur at the strategic goal and objective level. The report should aggregate performance goal results and assess their impact on high-level goals. These summaries should take into consideration the fiscal year’s priorities and relative significance of different goals, measures, and actual results. Transparency and accountability are ill-served by merely listing detailed measures and data from which the reader is expected to draw conclusions.

It should be clear why specific targets were chosen. What are the upper and lower limits of acceptable and achievable performance, and why? The effects of unexpected events or barriers—both internal and external—should be explained, and solutions revealed or suggested. Special care should be taken with resource explanations to indicate precisely how more or different resources would fix the problem and why reallocations were not made internally.

Criterion 10 saw very modest improvement this year. Once again this year, no report scored a 5. Five reports earned 4s, one more than last year. Nine agencies scored 2s, about the same as last year.

STRONGEST SCORES: Labor, NASA, State, Transportation, Veterans Affairs

The NASA report clearly discloses performance shortfalls and uses the best format of all reports to describe them. A table captioned “NASA’s FY 2006 Performance Improvement Plan” describes each shortfall, the reasons for it, and plans for improving in the future. The content is generally informative. The report also provides the same information as a follow-up for fiscal year 2005 performance shortfalls.

KEY EVALUATION FACTORS

- Does the report clearly and candidly disclose performance shortfalls?
- Does it explain shortfalls in a way that gives the reader useful insight into their cause?
- Does it describe planned remedial actions in a way that gives the reader confidence that the agency is working to improve its performance?

The Transportation secretary's transmittal letter candidly acknowledges disappointing results, particularly the increase in highway fatalities, and briefly describes the department's strategies to improve. The detailed performance section provides informative explanations for each target that was missed. The report even includes cogent explanations and future projections for some measures that were met.

The Veterans Affairs transmittal letter acknowledges last year's highly publicized theft of a computer containing sensitive data and describes what it is doing to prevent a recurrence of that problem. The Labor report consistently provides useful information on why it missed goals and what it will do to improve. The State report uses color-coded graphics to clearly disclose its performance results. Where performance fell below target, significantly or otherwise, the narratives consistently provide a brief but informative explanation as well as a description of steps to improve.

WEAKEST SCORE: Numerous

The Interior report attributes many performance shortfalls to data glitches or methodological issues and provides no insight into how the department can improve its performance. For example, one narrative states, unhelpfully: "The goal was not met because of the inherent difficulties in setting precise goals for this measure." Another indicates that the goal was missed because the department is now using more accurate data. The HUD and USAID reports use good formats to describe performance shortfalls, but the content leaves much to be desired. Most explanations in the USAID report simply state that analyses of the reasons for shortfalls and improvement steps are "pending" or lack adequate data. The NRC report does not highlight performance shortfalls. The reader must look closely to find them, although they are disclosed. NSF uses a system that is unlikely to surface negative performance examples. The four incomplete reports (Defense, GSA, HHS, and SBA) also did poorly here, since they do not fully disclose performance results.

OPPORTUNITIES FOR IMPROVEMENT

Straightforward descriptions of performance results. Some agencies use multiple categories to describe their performance results, such as "fell well short," "almost met," or "substantially exceeded." Refining performance results in this way can be informative. However, it is important that the reports provide a clear "bottom line" that clearly indicates whether each performance target was or was not met. Classifications that obscure this bottom line detract from the report's credibility. For example, the Commerce report classifies a goal as "significantly met" if the agency achieved 75-99 percent of the applicable performance targets. Treating goals as "significantly met" if they were missed by as much as 25 percent across the board suggests that the agency is less than transparent in its reporting or less than demanding in its performance.

Reporting perfect or near perfect performance is cause for skepticism, not celebration. No reasonable person would expect a major federal agency (or any organization with complex and inherently difficult missions) to have *no* performance shortfalls. Rather than assuming that the organization was performing flawlessly, the reader would more likely suspect that the agency either was not reporting candidly or was not challenging itself in its performance measures. Two reports serve as examples here. OPM reports meeting all 58 of its fiscal year 2006 performance measures, but its measures are exceptionally weak—probably the least significant and least challenging of any agency. NSF reports no performance shortfalls with respect to its three programmatic strategic goals this year, nor has it reported any such shortfalls at least as far back as fiscal year 2002. However,

NSF uses a performance assessment system that may be fundamentally appropriate but is highly unlikely to disclose any examples of unsuccessful projects.

11. DOES THE REPORT ADEQUATELY ADDRESS MAJOR MANAGEMENT CHALLENGES?

The report should describe how an agency minimizes the risks to its success so as to maximize results for citizens. The impact of management issues is clearest in a context of specific goal achievement. It should be clear which challenges are “mission-critical” and why. Major management challenge discussions should include full disclosure of the background, comments of the agency’s inspector general and Government Accountability Office, agency responses indicating an appreciation of threats to its mission and goals, and an anticipation of future risks.

This year, we continued to scrutinize the reports for specific evidence that agencies are addressing major management challenges, rather than merely expressing good intentions. As in past years, the research team focused on the inspector general’s (IG) presentation of major management challenges that each report must contain; the agency’s response, if any, to the inspector general’s presentation; and any discussion of management issues raised by GAO. Once again this year, there is considerable variation in the quality of the inspector general’s presentations.

As we did last year, the research team also considered the agency’s scores on the President’s Management Agenda, which are included in each report. The “status” rather than the “progress” scores were what counted. The President’s Management Agenda scores did not affect our scores in any mechanical way. Rather, we used them as a source of information on the agency’s progress, in the same way that we examined the inspector general’s discussion of major management challenges. Good scores increased confidence that the agency is taking management issues seriously. Evaluation focused on the areas of human capital, financial management, e-government, and budget-performance integration. The research team did not focus on competitive sourcing, which is more controversial than the other areas and perhaps more of a policy initiative than a management issue.

KEY EVALUATION FACTORS

- Does the report include the inspector general (IG) summary and assessment of the agency’s most serious management and performance challenges?
- Does it include a response by agency management to the IG’s presentation?
- Does the report provide confidence that the agency is taking specific and effective actions to resolve major management and performance challenges?
- Does the agency have good “status” scores on the President’s Management Agenda?

As a whole, the scores changed little from last year and remain low. There was one 5 and there were three 4s, the same as last year. Eleven agencies earned 2s, versus ten last year.

STRONGEST SCORE: Labor

Once again this year, Labor was the only report to receive a 5 on this criterion. The table containing Labor management's response to the IG's summary of major management challenges is more specific and informative than the IG's summary. It covers major management challenges identified by the Government Accountability Office and the department's chief financial officer as well as the IG. It breaks down each challenge into the specific issues that need to be addressed, lists actions taken in the fiscal year to address them, and also lists any further actions that need to be taken and their estimated completion date. Labor was one of two agencies that scored "green" in each status category of the President's Management Agenda as of September 30, 2006. (State also earned all green scores, but subsequently received a disclaimer of opinion on its financial statements.)

WEAKEST SCORES: Numerous

The fundamental shortcoming in the low-scoring reports is that they provide the reader with little or no confidence that the agency is serving as a good steward of taxpayer dollars by taking effective steps to resolve its major management challenges. The reports often are long on statements of commitment to resolve the challenges and general assertions of progress, but short on evidence of specific and significant remedial actions. This is a particular concern since many of the challenges agencies face have persisted for years. Defense is the leading example of a report that provides few specifics about remedial actions in the face of pervasive and long-standing management challenges.

OPPORTUNITIES FOR IMPROVEMENT

The Reports Consolidation Act requires performance reports to include a statement by the agency's inspector general that (1) summarizes what the inspector general considers to be the agency's most serious management and performance challenges and (2) briefly assesses the agency's progress in addressing them. The law goes on to provide: "The agency head may comment on the inspector general's statement, but may not modify the statement."²⁸ Two main issues arise concerning the required inspector general statements:

Incomplete inspector general statements. Four reports do not include the full original text of the inspector general's statement on major management challenges: EPA, Interior, OPM, and Veterans Affairs. Arguably, failure to include the full inspector general statement violates the law. In any event, the absence of the full statement raises a red flag concerning what, if anything, is omitted. It can also create confusion over what part of the discussion of major management challenges is attributable to the inspector general and what part comes from agency management.

Inadequate inspector general assessments of progress. While the inspector general statements generally do a good job of describing major management challenges, their assessments of agency progress often leave much to be

²⁸ *Reports Consolidation Act*, U.S. Code 31 (2000), 3516(d).

desired. Some inspector general statements have little if any content that could pass for an assessment of progress. A number of others simply list remedial actions that the agency is taking with no analysis of the significance of those actions in terms of resolving the challenge and no insight into what remains to be done. In a number of reports, agency management's assessments of the challenges and progress toward resolving them is more informative and insightful (and more concise) than the inspector general's statement. Labor and Transportation are two such agencies.

The SBA report provides the best example by far of an insightful inspector general assessment and should serve as a model for other inspectors general. The inspector general's assessment breaks down each challenge into specific recommended actions to resolve it. It includes tables that use a color-coded system to rate the agency's status with respect to each action and that list remaining actions needed for fiscal year 2007. There is also a table summarizing the agency's overall progress for each action and challenge.

The Agriculture inspector general's presentation of major management challenges is another good one. It breaks down each challenge into specific components and includes for each challenge an analysis of agency accomplishments and improvement plans as well as further actions needed to address the challenge.

More specifics from agency management. Of course, agencies also need to improve their own presentations on major management challenges and provide more specifics on the remedial actions they are taking, as well as explaining the significance of those actions.

12. DOES THE REPORT DESCRIBE CHANGES IN POLICIES OR PROCEDURES TO DO BETTER NEXT YEAR?

The intent of the Government Performance and Results Act is not just reporting for its own sake. The law's intent is to hold agencies accountable for results rather than for activities. The idea is to gather information on results and then to use that information in a strategic manner – as a guide to future decisions. The most important improvement will therefore occur at the highest level rather than in individual program goals or with the adjustment of measures. Is it evident that the agency is actually using knowledge gained from the reporting

KEY EVALUATION FACTORS

- Does the report do a good job of describing improvement plans to address specific performance shortfalls and major management challenges?
- Does it describe how to do better even where current performance met expectations?
- Does it evidence a strategic focus by describing more general future challenges that the agency faces in accomplishing its missions and how to address them?
- Does it explain efforts to enhance performance accountability and reporting?

process to revise its priorities and guide its activities? What is the potential for an agency to make a positive difference in the future? How will it realize that potential?

Eight agencies scored 4s this year; seven earned scores of 4 last year. Five reports received 2s this year, compared to three last year. Notably, the four agencies with scores of 4 are at the top of our overall rankings, and those with scores of 2 are clustered at and near the bottom of the rankings.

STRONGEST SCORES: Numerous

While no single report excelled on this criterion, a number did well. From start to finish, the Transportation report demonstrates a commitment to improve on both the programmatic and management sides. The department consistently sets challenging targets and thoughtfully describes how it can do better with respect to individual measures. The report also outlines broader programmatic improvement strategies with respect to its strategic goals.

The Labor report discusses future programmatic and management challenges for each performance goal and how they will be addressed. The department continues to innovate and improve on many features of its performance reporting, such as linking costs to performance measures.

From the secretary's transmittal letter on, the State report evidences strategic thinking and serious efforts to improve performance. The introduction this year of a joint performance section with USAID is further evidence of the department's strategic focus. The descriptions of efforts to improve on specific programmatic performance shortfalls are generally strong.

The Veterans Affairs report is rich in information on efforts to improve programmatic and managerial performance. One noteworthy feature is a table that highlights the department's most significant current challenges under each strategic goal. As discussed previously, Veterans Affairs is one of the agencies moving to link budget costs to individual performance measures.

WEAKEST SCORES: HUD, HHS, Interior, NSF, OPM

Many of the annual targets in the HUD report do not appear to be challenging when compared to prior year actual performance levels. The many incongruities and apparent errors in the report call into question how much attention HUD is devoting to its performance reporting. The Interior report provides generally weak explanations of programmatic performance shortfalls, and the descriptions of actions to address management challenges are not particularly insightful. The report has little additional content on improvement strategies. The lackluster and meager contents of the OPM report give the impression that the agency is focused on short term operational concerns. It lacks any evidence of strategic thinking.

OPPORTUNITIES FOR IMPROVEMENT

Improvements in a number of other evaluation criteria would pay dividends here. In particular, agencies could benefit by enhancing their descriptions of remedial actions for performance shortfalls and management challenges. Innovative efforts to enhance the quality of the reports in any category would help here as well.

AGENCY-BY-AGENCY SCORING SUMMARIES

IN FISCAL 2006 RANK ORDER

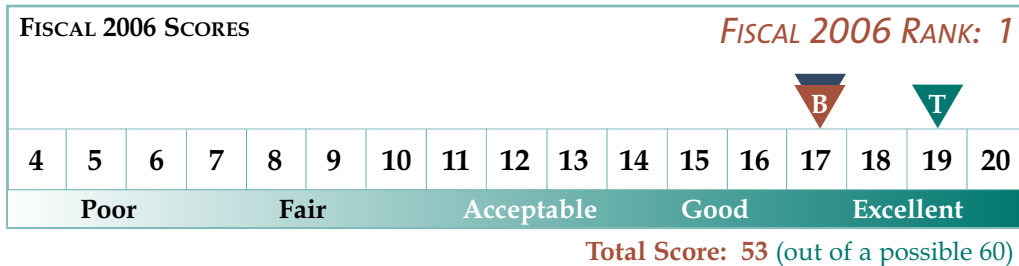
This section summarizes the scores received by each agency in the three major scoring categories: Transparency, Public Benefits, and Leadership. Each agency summary appears on a separate page in rank order from highest to lowest. The graphic at the top of each page displays the scores each agency received in the three categories this year, fiscal 2006. The graph at the bottom shows the rankings each agency has earned on the Scorecard for fiscal 1999 thru fiscal 2006.

For example, the Department of Transportation's report this year earned scores of 19, 17, and 17 on the Transparency, Public Benefits, and Leadership criteria respectively. The total of these scores, 53, gave this report the top ranking for fiscal 2006.

Significant strengths and weaknesses of each agency's report are then summarized in bullet form. These summaries correspond to the 12 evaluative factors and are organized according to the three evaluative categories: Transparency, Public Benefits, and Leadership.

These 1-page descriptions summarize extensive notes compiled by the research team that explain the reasons for each report's score on each criterion. The full sets of notes for each report are available on the Mercatus Center's Web site at <http://www.mercatus.org/scorecard/agencies>.

DEPARTMENT OF TRANSPORTATION



T TRANSPARENCY

- Home page has a direct link to the report, which is downloadable as a single file or in multiple files.
- Report makes excellent use of tables and graphics, has clear and informative narratives, and uses a very manageable number of performance measures.
- A useful overview of the department’s system for assessing data is included, as well as considerable detail on the data for each performance measure.
- Report is rich in baseline and trend data.

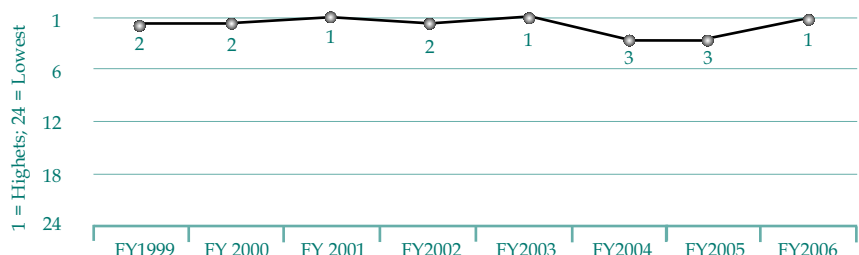
B PUBLIC BENEFITS

- Most strategic goals and strategic objectives are outcome-oriented.
- 23 of 29 strategic goal measures appear to be ultimate or intermediate outcomes.
- Informative performance metrics and accompanying narratives demonstrate the department’s accomplishments.
- Report allocates budget resources to each strategic and performance goal.

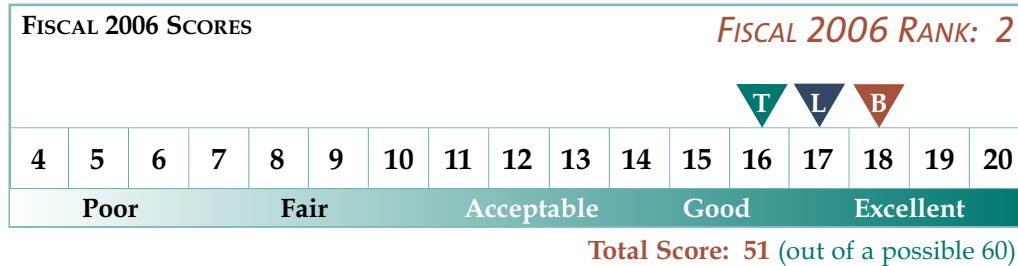
L LEADERSHIP

- Report clearly discloses the department’s performance results, including shortfalls, in reader-friendly tables.
- “Progress Meter” icon, used to classify the status of each management challenge, indicates moderate or better progress on most of the rated items.
- Report discusses several broad improvement strategies for programs.

RANKING HISTORY



DEPARTMENT OF LABOR



T TRANSPARENCY

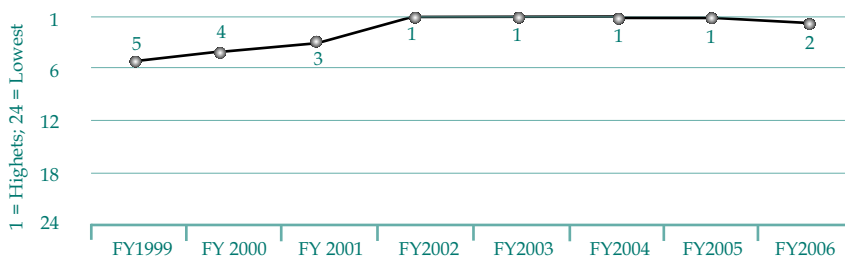
- Department’s home page has a prominent link, and report is downloadable in a single file or multiple files.
- Report is well organized and easy to follow.
- Report uses an excellent system for rating the completeness and reliability of data although significant data issues remain.
- Performance section shows prior year results and provides background on the significance of each performance goal, as well as performance trends and future plans.

B PUBLIC BENEFITS

- All four strategic goals and 25 of 28 performance goals capture end outcomes or at least intermediate outcomes.
- A significant majority of the 85 annual performance measures are outcome oriented.
- Strong performance metrics demonstrate contributions toward goals.
- Major advance made by allocating costs to all strategic and performance goals and to many individual performance measures.

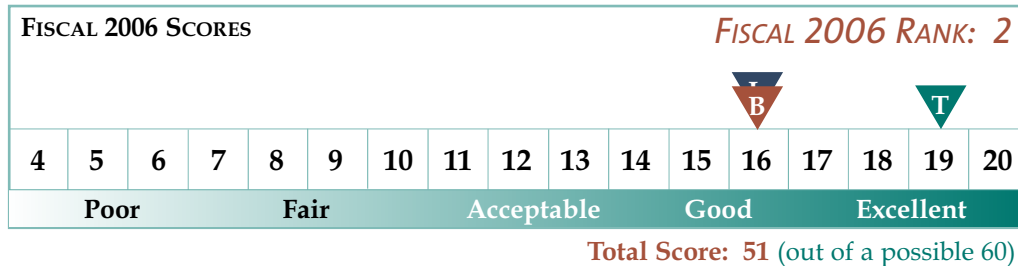
L LEADERSHIP

- Outcome-oriented performance metrics, thorough narratives, and personal vignettes demonstrate public benefits.
- Transmittal letter could highlight more specific accomplishments.
- Performance section consistently provides useful information on why goals were missed and what will be done to improve.
- Department’s presentation of major management challenges is more informative than inspector general’s list since it describes specific steps to resolve challenges and estimated completion dates.
- Narratives provide insightful descriptions of improvement plans even when targets were achieved.



RANKING HISTORY

DEPARTMENT OF VETERANS AFFAIRS



T TRANSPARENCY

- Home page has a prominent link to the report, which is downloadable in a single file or in multiple files.
- Noteworthy “Performance Scorecard” table presents at a glance strategic goals, objectives, and key performance measures and results.
- Report includes results for all 23 key measures, but data are incomplete for about 15 percent of the non-key measures.
- Report includes a wealth of baseline and trend data, and the narratives elaborate on the performance trends.

B PUBLIC BENEFITS

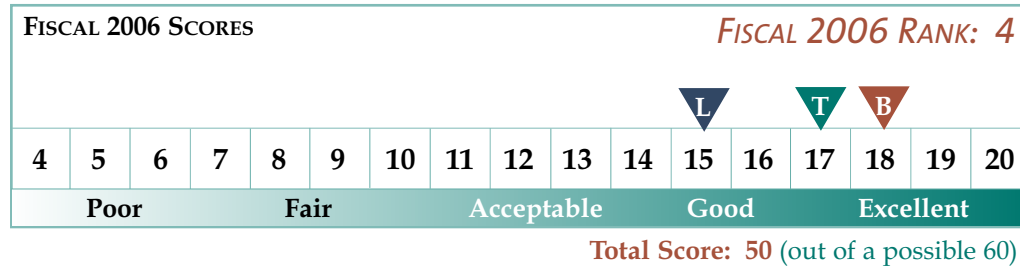
- About half of the strategic goals and two-thirds of the strategic objectives are stated as outcomes.
- All but four key measures are outcome-oriented. Many of the non-key measures relate to timely and accurate benefit determinations.
- Narratives and comprehensive baseline and trend data demonstrate the significance of the measures in relation to the department’s goals.
- Report allocates budget costs to strategic goals for 21 strategic objectives, and, for the first time, begins to allocate costs to individual performance measures.

L LEADERSHIP

- Transmittal letter and narratives lay out results in a clear and compelling way.
- Performance shortfalls are presented clearly.
- If the inspector general provided a more informative presentation on major management challenges, it should have been included.
- The report is rich in information on efforts to improve on programmatic and managerial performance.

RANKING HISTORY





T TRANSPARENCY

- Home page has a direct link to the report (although not very prominent), and report is downloadable in a single file or in separate files.
- Four versions accommodate a variety of audiences: the full report, a summary version, an eight-page overview, and an interactive CD.
- Report has good background information on performance data, but department got a disclaimer of opinion on its financial statements.
- Baseline and trend data are present for the individual measures, and narratives provide additional useful context for assessing progress over time.

B PUBLIC BENEFITS

- Virtually all programmatic strategic goals and annual performance goals are clearly stated as intermediate or end outcomes.
- The 94 performance measures align well with the goals and include many end outcomes or intermediate outcomes.
- Narratives accompanying the individual measures consistently demonstrate the significance of the performance results toward achieving the department’s goals.
- Budget and personnel resources are allocated to strategic goals and annual performance goals, but not individual performance measures.

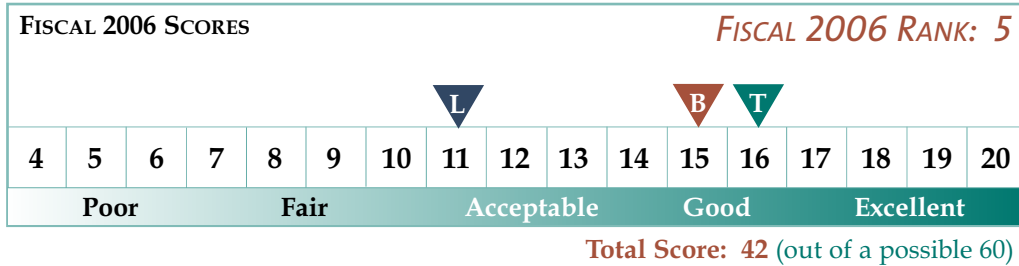
L LEADERSHIP

- The “public benefit” descriptions of each strategic goal are generally informative and persuasive.
- Color coded graphics clearly disclose the department’s performance and shortfalls.
- No separate response to the inspector general’s presentation of management challenges.
- Descriptions of efforts to improve on specific programmatic performance shortfalls are generally strong.



RANKING HISTORY

U.S. AGENCY FOR INTERNATIONAL DEVELOPMENT



T TRANSPARENCY

- Home page has a prominent link to the report, which is downloadable in a single file or multiple files.
- Report makes effective use of tables and graphics and uses a manageable number of performance measures.
- Report includes results and a data source and quality assessment for all 35 performance measures, but the inspector general suggests possible “weaknesses” in the agency’s performance reporting system.
- Trends and target selection for some measures could use more explanation.

B PUBLIC BENEFITS

- All strategic and performance goals are stated as outcomes.
- Management goals are stated less as outcomes but are relevant to ensuring capacity to achieve mission outcomes.
- Measures are much less outcome-oriented than goals.
- The report allocates budget and personnel resources to the strategic goals and to performance goals, but not to performance measures.

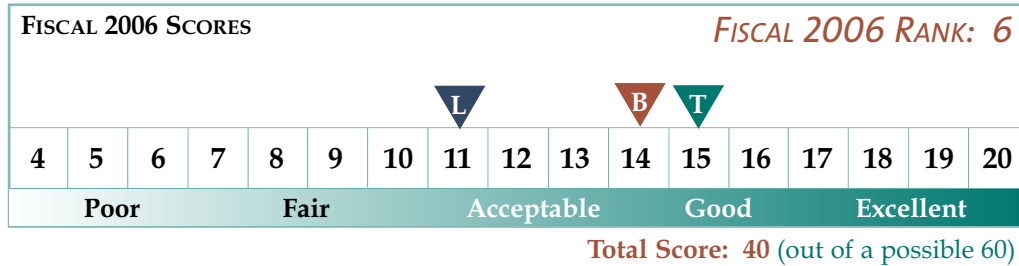
L LEADERSHIP

- Informative, specific descriptions of public benefits accompany each strategic goal.
- While the format of the tables describing performance and shortfalls is excellent, the content often leaves much to be desired.
- Useful table lists specific actions taken to address management challenges, as well as remaining actions needed and expected completion dates.
- Weak in describing improvement plans for specific performance shortfalls.

RANKING HISTORY



GENERAL SERVICES ADMINISTRATION



T TRANSPARENCY

- Agency's home page has a direct link to the report, which is downloadable as a single file or multiple smaller files.
- Although the report is highly self-congratulatory in tone, the performance-related portions of the report are concise and easy to follow.
- Report includes fiscal year 2006 results for all 21 key performance measures, but no results are provided for any of the other measures.
- Appendix shows prior year baseline and trend data but does not include prior year targets.

B PUBLIC BENEFITS

- Of the six listed strategic goals, only two are outcome-oriented. About half of the performance goals focus to some extent on outcomes or intermediate outcomes.
- Measures are relevant and valid indicators for listed goals, but few of the measures address outcomes, and many are customer satisfaction or efficiency measures.
- Although all key measures have results, well over half of the agency's fiscal year 2006 results are absent.
- Table links projected fiscal year 2006 budget costs to all individual performance goals and, with very few exceptions, to the individual measures for the performance goals.

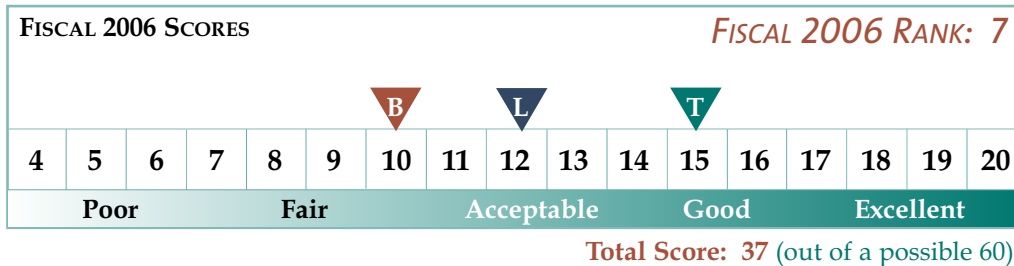
L LEADERSHIP

- Narratives describe how GSA's efforts enhance the ability of its customer agencies to achieve results that directly benefit the public.
- Neither the inspector general nor agency management provides useful insight into how much progress the agency is making to resolve major management challenges.
- The report, especially the transmittal letter, indicates that the agency is undergoing fundamental changes designed to improve its processes and performance results.



RANKING HISTORY

DEPARTMENT OF JUSTICE



T TRANSPARENCY

- Home page has a prominent direct link to the report, which is downloadable as a single file or multiple files.
- Brevity of the performance section and the generally clear and succinct narratives mitigate the text-heavy nature of the report.
- Report acknowledges some data limitations, which are described in narratives accompanying individual performance indicators.
- Boxes accompanying individual performance measures include prior year results, but only a few show prior year targets.

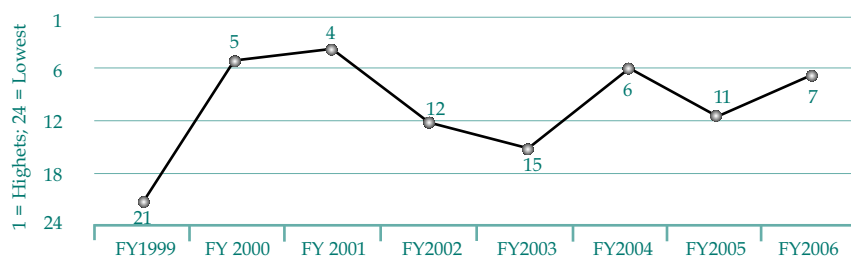
B PUBLIC BENEFITS

- Two of the four strategic goals and eight of the 18 strategic objectives are clearly stated as outcomes.
- As many as 15 of the 28 performance measures could be considered to capture end or intermediate outcomes, but more would be outcome-oriented if stated as percentages rather than raw activity numbers.
- Budgetary resources are linked only to the strategic goals; there is no linkage to the strategic objectives or to individual performance measures.

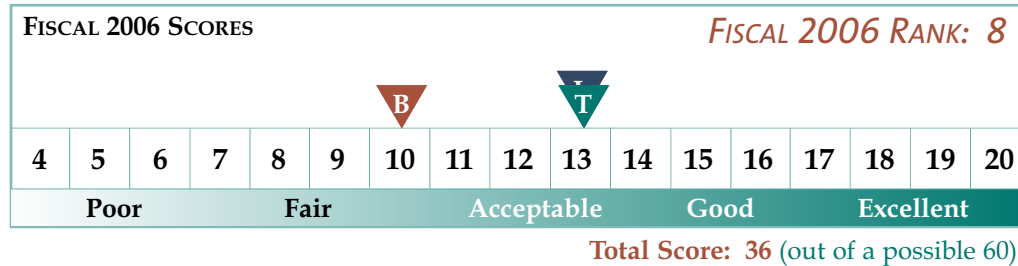
L LEADERSHIP

- Performance metrics and narratives generally do not convey a clear sense of whether and to what extent the overall conditions being addressed are improving.
- Explanations of specific performance shortfalls are often weak, but narratives usually provide some discussion of what the department is doing to improve performance.
- Discussion of management challenges breaks down each problem and outlines specific remedial steps.

RANKING HISTORY



DEPARTMENT OF COMMERCE



T TRANSPARENCY

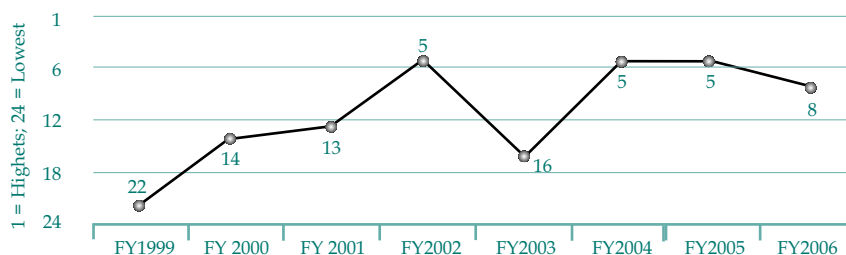
- Home page provides a direct link to the report, which is downloadable in a single file or in multiple files.
- Report is text-heavy, but also contains some good tables and graphics.
- No data definitions or sources for individual performance measures are available, but the report does have fiscal year 2006 results for all 131 measures.
- Tables indicate through color-coded symbols whether the prior year results were met although they do not show prior year targets.

B PUBLIC BENEFITS

- Strategic goals and objectives are overwhelmingly not outcome-oriented. Majority of the performance goals are either not stated as outcomes or vague.
- A few key measures capture outcomes, but most do not.
- Narratives and vignettes are generally informative and compensate somewhat for the weak performance metrics.
- Report links the budget resources and costs to strategic goals, strategic objectives, and down to each performance goal.

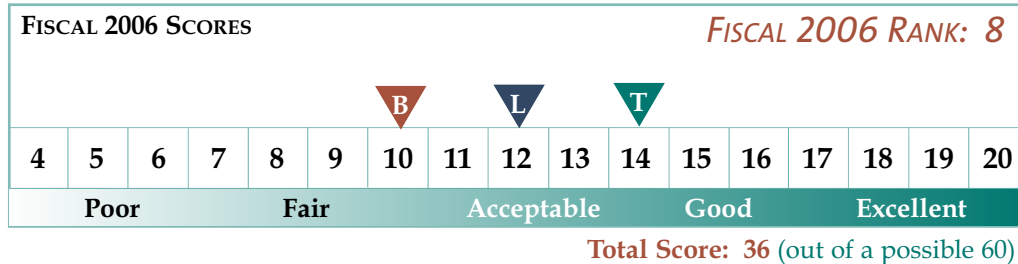
L LEADERSHIP

- Practice of classifying performance measures as “significantly met,” even if missed by as much as 25 percent, raises a red flag.
- Response to inspector general’s management challenges outlines specific actions that the department is taking.
- Narratives describe general strategies for the future with respect to each strategic objective, including anticipated challenges and plans to address them.



RANKING HISTORY

DEPARTMENT OF EDUCATION



T TRANSPARENCY

- Direct link to the report at the top of the department’s home page is easy to miss; report is downloadable in a single file or in multiple files.
- Report is well organized and uses an excellent format to present the department’s performance results.
- Improvements are needed in data timeliness and accuracy.
- Presentation consistently shows whether progress was made over the prior year.

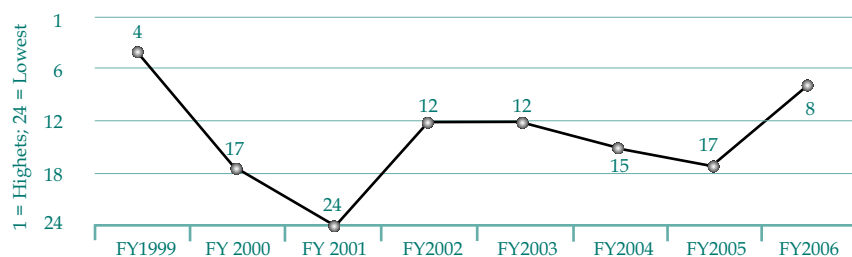
B PUBLIC BENEFITS

- Three of the five programmatic strategic goals are outcomes.
- Although none of the measures under “strategic goal 1” are outcome-oriented, a majority of the remaining measures appear to capture end or intermediate outcomes.
- Metrics are average, but most of the narratives usefully elaborate upon the department’s contributions.
- Report does not link costs to results.

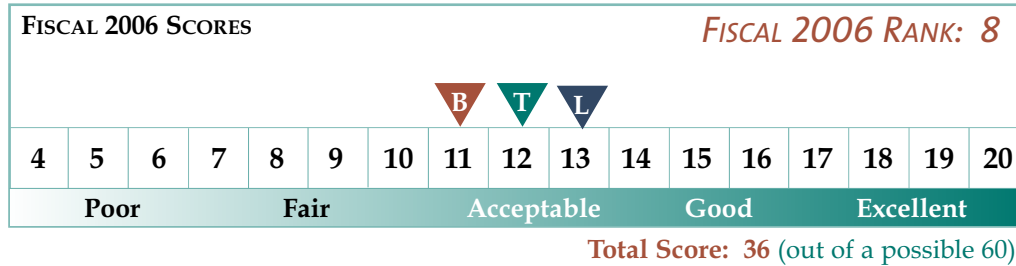
L LEADERSHIP

- Report clearly discloses performance results, including shortfalls.
- There is no agency management response directly to the inspector general’s concise and explicit report on management challenges.
- Little content on planned corrective actions for specific performance shortfalls, but the department does seem intent on improving its performance data.

RANKING HISTORY



ENVIRONMENTAL PROTECTION AGENCY



▼ TRANSPARENCY

- Home page has a direct link to the report and provides good contact information for those with questions or comments.
- Although improved, the report still presents major readability challenges.
- Discussion of efforts to improve data quality is present, but lagging data are a significant impediment to performance reporting.
- Baseline and trend data are presented, but readability could be improved.

▼ PUBLIC BENEFITS

- All of the strategic goals are stated as high-level outcomes. About half of the strategic objectives are outcome-oriented.
- Many performance measures are highly technical and difficult for a non-expert to interpret.
- Narratives help to connect EPA's contributions to goal achievement.
- There is cost linkage to strategic goals and objectives as well as major programs, but not to annual performance goals or measures.

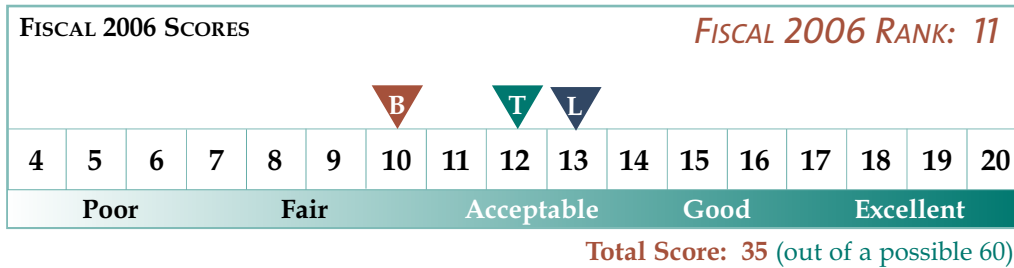
▼ LEADERSHIP

- Performance highlights describe important public benefits that flow from the agency's work.
- Explanations for specific performance shortfalls and remedial actions are provided.
- Inspector general's presentation indicates progress in managing for results, but is less specific concerning progress on the other challenges.



RANKING HISTORY

DEPARTMENT OF AGRICULTURE



▼ T TRANSPARENCY

- Homepage lacks a direct link to the report. Contact information could not be found.
- Secretary's transmittal letter is succinct and highlights the department's major missions in a results-oriented way.
- Report uses a manageable number of performance measures, and narratives are generally clear and understandable for the lay reader.
- No data sources or data assessments with respect to individual performance measures.
- Provides baseline and trend data. Accompanying narratives are quite thorough.

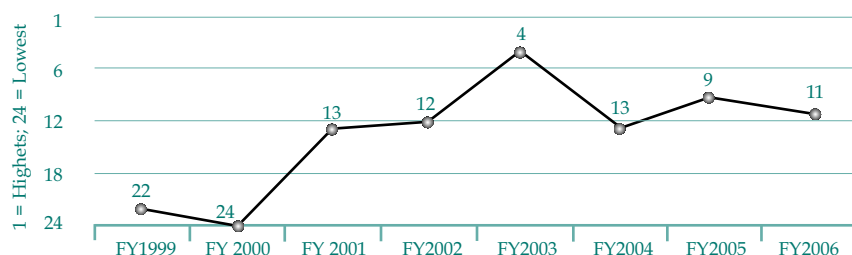
▼ B PUBLIC BENEFITS

- All strategic goals are stated as outcomes, although often at a high level of generality.
- Less than half of the 38 performance goals, which double as measures, are stated as outcomes.
- Budget and staff resources are linked to the strategic goals, but not to any other level of the performance metrics.

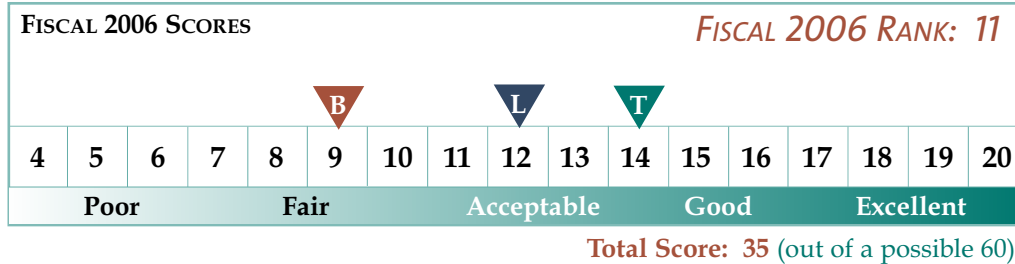
▼ L LEADERSHIP

- Narratives describe the public benefits produced, but would be stronger if backed by more outcome-oriented performance metrics.
- Quality of the explanations of shortfalls and improvement plans varies.
- Inspector general's presentation of major management challenges is thorough and informative.
- Narratives analyze challenges in each strategic objective area.

RANKING HISTORY



DEPARTMENT OF THE TREASURY



T TRANSPARENCY

- Home page has a link, albeit obscure, to the report, and report is downloadable in a single file or multiple files.
- Good blend of narratives, graphs, and tables.
- Report provides little detail on data verification and validation.
- Performance section contains graphics illustrating multi-year trends in a number of categories.

B PUBLIC BENEFITS

- Some of the strategic goals and objectives capture outcomes, but at very high levels.
- Many of the measures classified as outcome measures are not.
- Narratives are more effective than the performance metrics in capturing the department's accomplishments.
- There is no cost allocation below the level of the strategic objectives.

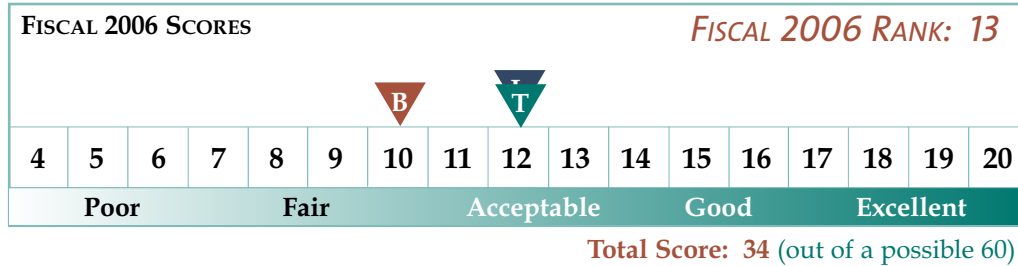
L LEADERSHIP

- Report clearly discloses performance shortfalls and consistently offers explanations for them.
- Reported management challenges are the same as last year, with indications of progress in some areas but not others.
- Performance section includes a discussion of future plans with respect to each strategic goal.



RANKING HISTORY

DEPARTMENT OF ENERGY



T TRANSPARENCY

- Although reduced from last year, the number of performance goals and measures (53 annual goals and 204 measures) creates a readability problem.
- Many goals and measures are stated in technical terms that are difficult for the lay reader to grasp.
- There are considerable deficiencies in data completeness and reliability.
- Report shows whether measures were met in prior years, but does not provide trend or baseline data to put performance in context.

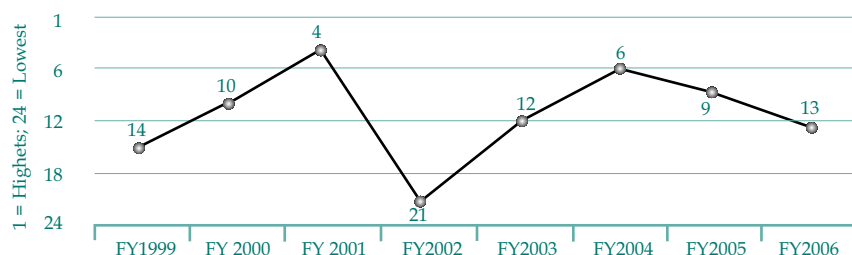
B PUBLIC BENEFITS

- Strategic goals and strategic objectives capture some clear outcomes, but at a high level of generality.
- Vast majority of performance measures address outputs, processes, and efficiencies rather than outcomes.
- Report allocates costs to the department’s strategic goals as well as to its strategic objectives and annual performance goals.

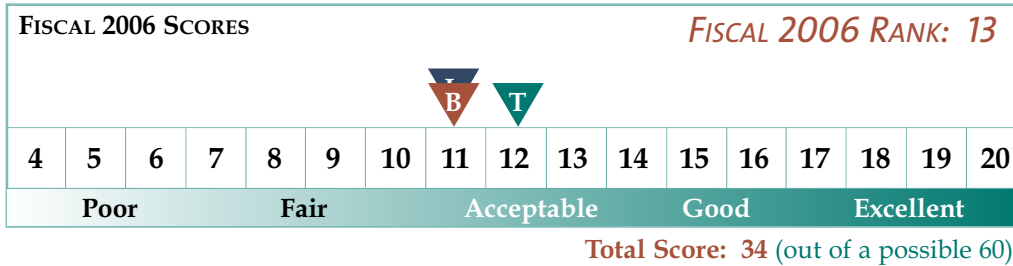
L LEADERSHIP

- “How We Serve the Public” sections do a good job of explaining the importance and public benefits of the department’s programs.
- Difficult for the lay reader to gain much insight into the “action plans” that describe remedies to performance shortfalls, due mainly to the technical nature of many measures.
- Report discusses major management challenges and the specific actions taken to address them.

RANKING HISTORY



NUCLEAR REGULATORY COMMISSION



T TRANSPARENCY

- Home page contains direct link to the report; no contact information could be found.
- Body of the report, while text-heavy, is well-organized, relatively concise, and generally easy to understand.
- Very manageable number of performance indicators: 5 strategic goals, 10 strategic outcomes, and 16 performance measures.
- New measures do not have significant prior year data as of yet.

B PUBLIC BENEFITS

- Safety and security strategic goals are clearly stated as outcomes while the rest, although focused on important aspects of the regulatory process, are less outcome-oriented.
- Like their associated goals, the safety and security measures are outcome-oriented, but most of the rest are not.
- Pass-fail nature of the key goals and measures makes it difficult to assess whether the agency is enhancing its performance from one year to the next.
- Report allocates costs to its safety and security goals; the agency chose not to allocate costs to the other goals, but does not explain why.

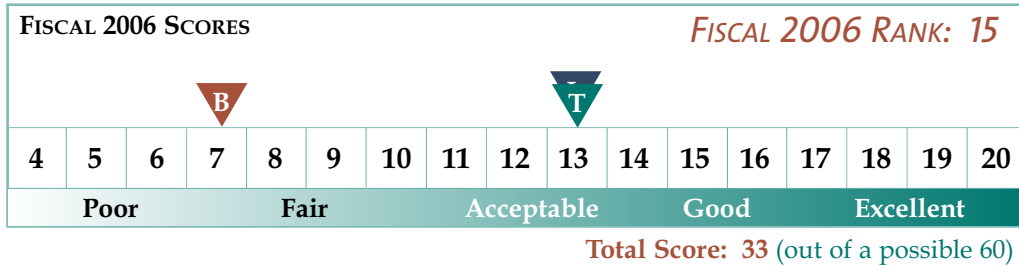
L LEADERSHIP

- Strong performance metrics in the agency’s key mission areas and informative narratives show the agency is effectively carrying out safety and security functions.
- Report does not highlight performance shortfalls, and the reader must look closely to find them.
- No specific response to the inspector general’s presentation of management challenges.



RANKING HISTORY

SOCIAL SECURITY ADMINISTRATION



T TRANSPARENCY

- Home page has a direct, but hard to find, link to the report; the report is downloadable in a single file or in multiple files.
- Report has good use of tables and graphics.
- Audits of 21 performance measures determined that five were reliable and seven were unreliable.
- Report clearly presents baseline and trend data for individual measures, using graphs to highlight them.

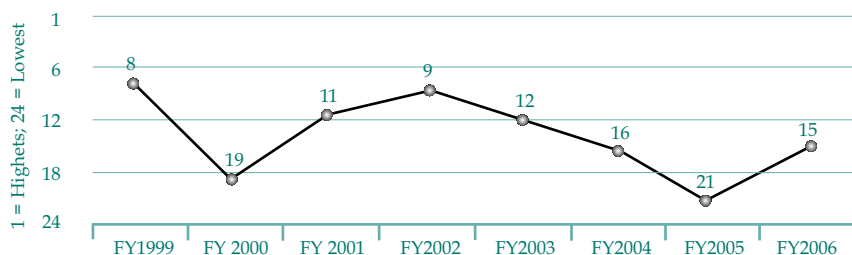
B PUBLIC BENEFITS

- Few strategic goals are outcome-oriented.
- Overall, the agency's performance metrics are weak; it is unclear how some measures and targets relate to outcomes, particularly those expressed as raw numbers of activities or outputs.
- Report provides limited context for understanding to what levels of performance the agency aspires.
- Costs are linked to the performance metrics only at the strategic goal level.

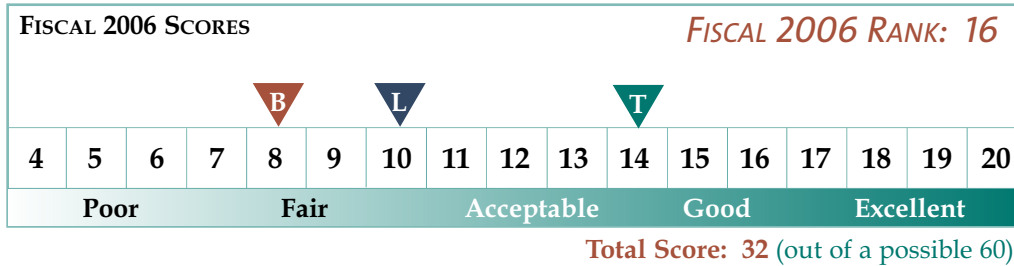
L LEADERSHIP

- Explanations are provided where shortfalls occurred, and some discussion of improvement strategies is usually included.
- Report provides an extensive description of actions being taken to address major management challenges.
- Report conveys that SSA is working to improve its timeliness and accuracy.

RANKING HISTORY



DEPARTMENT OF DEFENSE



▼ T TRANSPARENCY

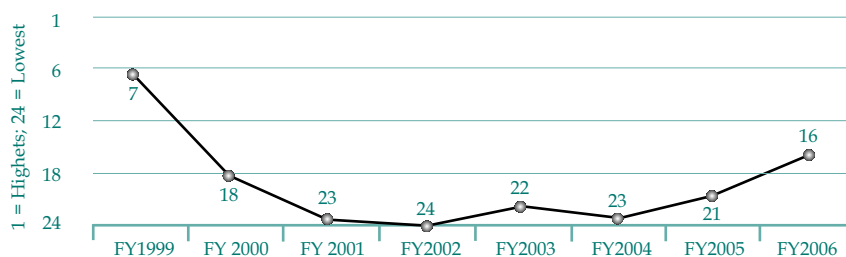
- Report is visually appealing, concise, and understandable, but incomplete due to omitted performance measures.
- Department has serious data shortcomings that result in disclaimers of opinion on financial statements each year and affect the quality of performance data.
- Report has no content on slightly over half of the department’s measures.

▼ B PUBLIC BENEFITS

- All four of the strategic objectives are highly results-oriented, as are the first two risk-management strategic goals.
- There are very few outcome-oriented measures among the 66 selected measures featured in the report.
- There’s no content linking costs to goals or measures at any level.

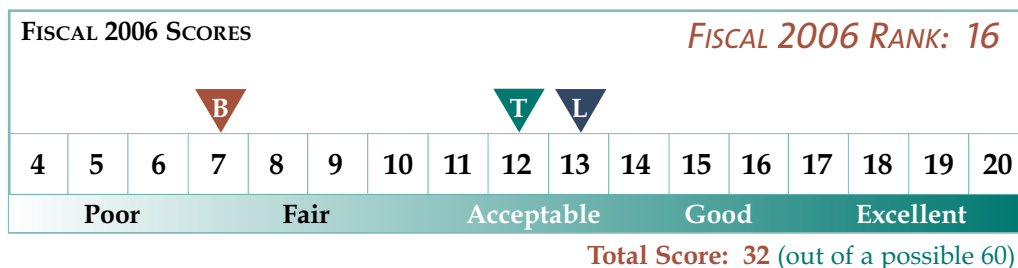
▼ L LEADERSHIP

- Narratives effectively describe the department’s accomplishments in relation to its missions.
- The report gives no explanation of shortfalls for those measures not covered in the report, but the narratives describe shortfalls for the covered measures.
- The report has a perfunctory response to pervasive management challenges identified by the inspector general and GAO.
- Narratives highlight and discuss the many transformational changes undertaken in the department, although in somewhat general terms.



RANKING HISTORY

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION



T TRANSPARENCY

- Home page has a direct link to the report, which is downloadable in a single file or in multiple files.
- Lengthy, text-heavy, and at times quite technical report poses a readability challenge.
- No specific content in the main body of the report dealing with performance data verification and validation.
- Report does not provide baseline or trend data that are useful for putting performance in context or making comparisons from year to year.

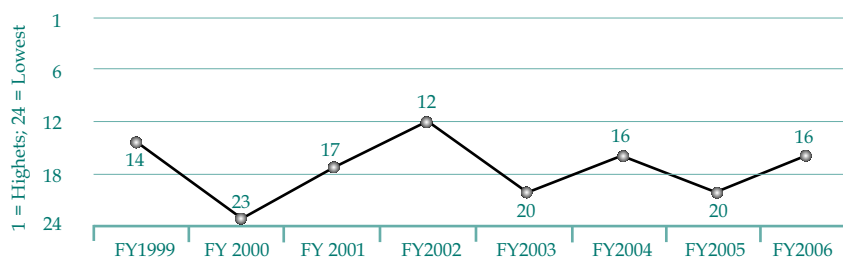
B PUBLIC BENEFITS

- The six strategic goals and most of the strategic objectives are not stated as outcomes.
- Performance measures follow the same pattern as the goals. Most are activities, tasks, or qualitative measures rather than outcomes.
- Report links costs only to the strategic goals and subgoals.

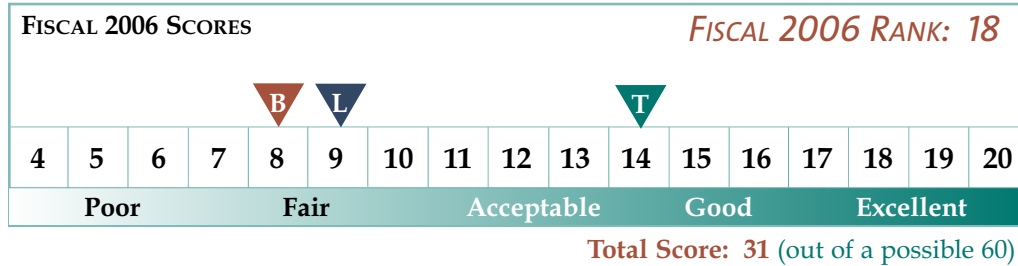
L LEADERSHIP

- Performance overview section does a good job of linking space exploration and other agency functions to enhancing human life.
- Report clearly discloses performance shortfalls and uses an excellent format to describe them and remedial actions.
- No specific response to the inspector general's presentation of management challenges.
- While report describes improvement plans for specific programmatic shortfalls, it says little about long-standing but critical management shortcomings, such as financial management and contract management.

RANKING HISTORY



DEPARTMENT OF THE INTERIOR



T TRANSPARENCY

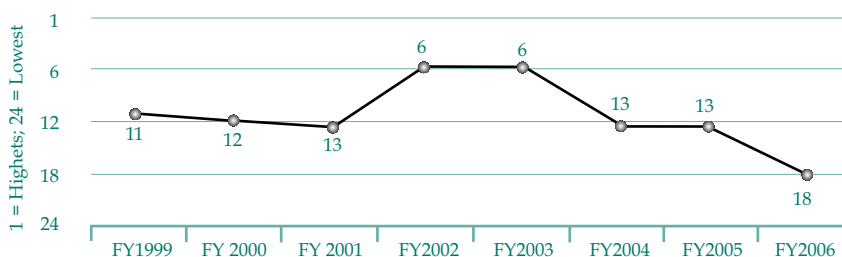
- Home page has an explicit direct link to the report and provides good report-specific contact information online and in the body of the report.
- The 209 performance measures are complex and sometimes confusing.
- No fiscal year 2006 results for about 18 percent of the measures, but a much higher percentage of results are reported this year than previous years.
- Prior year data for each performance measure are included, along with narratives that elaborate on performance trends.

B PUBLIC BENEFITS

- None of the strategic goals are stated as outcomes.
- Programmatic measures are somewhat more outcome-oriented than the goals, although less than half address outcomes.
- Narratives in the report provide some examples of significant contributions to goals, but stronger performance metrics are needed to back them up.
- Costs are linked to the “end outcome goals,” but it does not appear that costs are linked at a higher level or at a lower level.

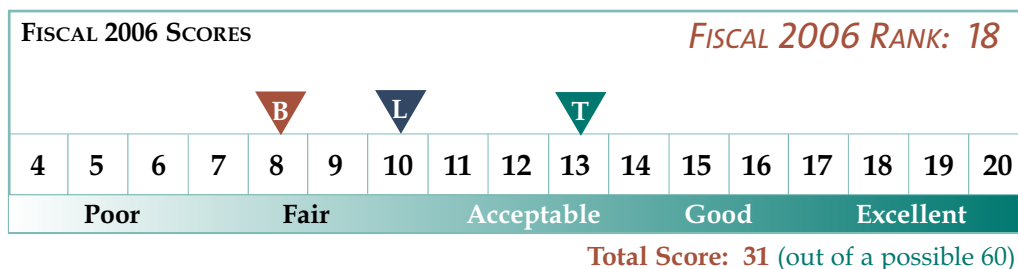
L LEADERSHIP

- Some explanations attribute shortfalls to data glitches or methodological issues and provide no insight into the department’s performance or how to improve it.
- Major management challenges are covered, but it is difficult to gain much insight into how much progress is being made.
- Report has little insightful content on improvement strategies, either on the programmatic or management side.



RANKING HISTORY

NATIONAL SCIENCE FOUNDATION



T TRANSPARENCY

- Home page does not have an explicit direct link to the report, and only the HTML version is downloadable in multiple files.
- Report has some good tables and graphics, but is generally text-heavy and makes extensive and excessive use of acronyms.
- The agency undertakes extensive efforts to verify and validate its performance measures and data.
- No meaningful baseline or trend data for the non-quantifiable measures.

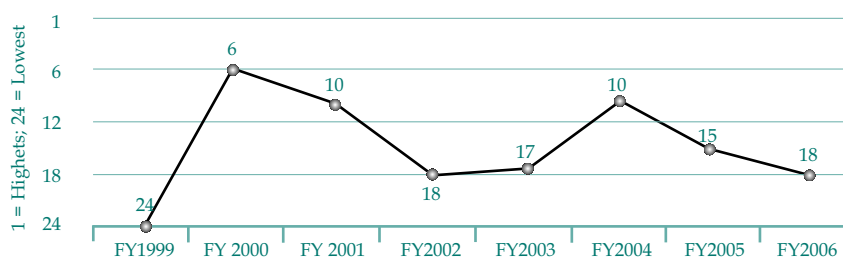
B PUBLIC BENEFITS

- Performance under the four strategic goals that are stated as high-level outcomes is assessed qualitatively or judgmentally rather than through quantifiable measures.
- The indicators, or measures, are generally not stated as outcomes.
- Research highlights and the narratives provide useful information regarding agency impact, compensating somewhat for the weak measures.
- Budget resources are only allocated to the agency's strategic goals and to "investment categories" within the three programmatic strategic goals.

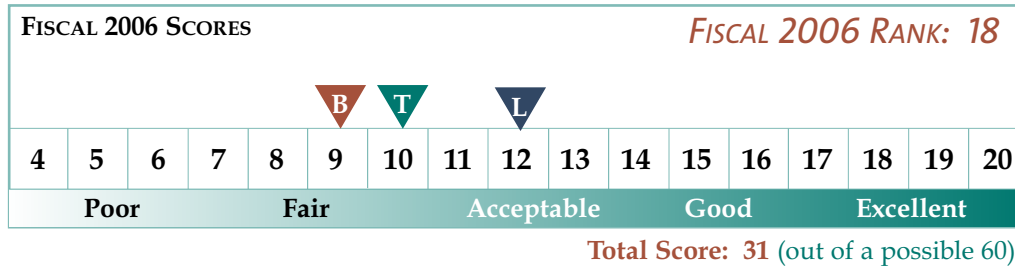
L LEADERSHIP

- Narratives would be more compelling for lay readers if they clearly demonstrated the connection between research results and important public benefits.
- Report does not provide a basis for assessing whether the agency's performance is improving over time.
- Agency's response to management challenges lists actions NSF is taking to address each challenge.
- Report conveys the sense that the agency's leadership is largely satisfied with NSF's performance and not particularly focused on efforts to improve.

RANKING HISTORY



SMALL BUSINESS ADMINISTRATION



T TRANSPARENCY

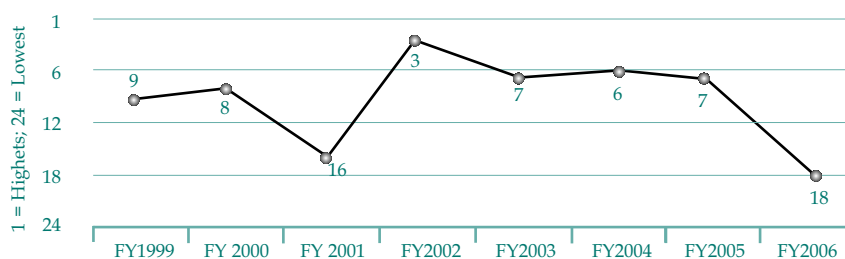
- The fiscal year 2006 report was not posted to Web site by the December 15, 2006 deadline.
- Report uses reader-friendly tables with color-coded graphics to summarize performance results.
- It is difficult for the reader to gain an overall sense of the agency's performance since the report covers only key measures and the presentation of even these measures is disjointed.
- Report provides a succinct but substantive discussion of steps taken to ensure data quality, some significant data issues that exist, and steps being taken to address them.
- Baseline and trend data are present, but additional context information concerning performance targets and trends would be useful.

B PUBLIC BENEFITS

- Some goals are stated as outcomes, but at such high levels that measurement is challenging.
- About half of the 11 agency-wide performance measures are truly outcome measures.
- Narratives accompanying the specific descriptions of performance results provide little perspective to assist in assessing the agency's performance.
- The report allocates budget resources to the strategic goals and strategic objectives.

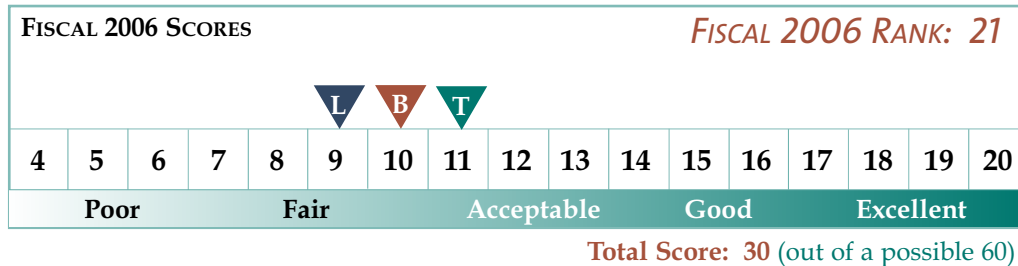
L LEADERSHIP

- General narrative portions of the report do a good job of articulating the public benefits that flow from the agency's work.
- Performance shortfalls are disclosed, but the explanations are not very informative and, in some cases, cryptic.
- Inspector general's presentation of major management challenges is exceptionally informative and useful.
- Transmittal letter and executive summary candidly emphasize lessons learned, as well as changes to be made.



RANKING HISTORY

DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT



▼ T TRANSPARENCY

- Home page has a prominent direct link to the report, but the report is downloadable only as a single large file.
- Text is verbose rather than succinct, focusing mainly on outputs, activities, and processes. It highlights few specific outcome-oriented results for the year.
- The inclusion of a “data discussion” to accompany many the individual performance indicators is a positive feature but sometimes could use more information.
- Report does not disclose prior year targets nor does it generally discuss how performance targets are set.

▼ B PUBLIC BENEFITS

- Goals and objectives are mixed in terms of outcome-orientation.
- Of the more than 90 measures under strategic goals, few are clearly stated as end outcomes.
- Narratives generally do a better job of demonstrating the department’s contributions than the department’s goals and measures.
- Budget resources are linked to goals and results only at the strategic goal level.

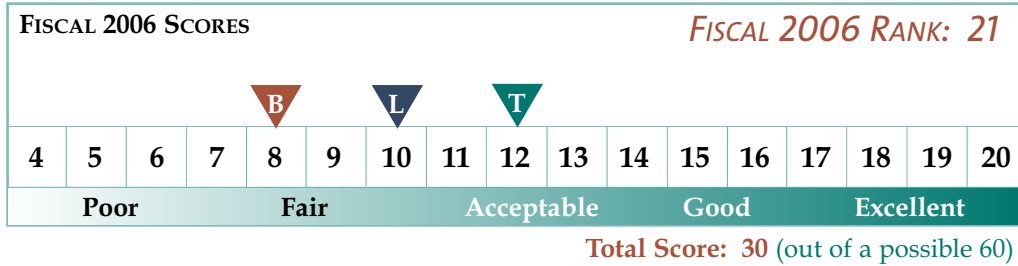
▼ L LEADERSHIP

- Report is formatted to clearly disclose performance shortfalls, but there are serious deficiencies in the actual data presented.
- Descriptions of management challenges present a mixed picture, indicating significant progress on at least one challenge but providing little evidence of specific improvement on others.
- Report does not convey a sense that the department is committed to improving its performance, due to unchallenging targets and slow-paced management improvements.

RANKING HISTORY



DEPARTMENT OF HOMELAND SECURITY



▼ T TRANSPARENCY

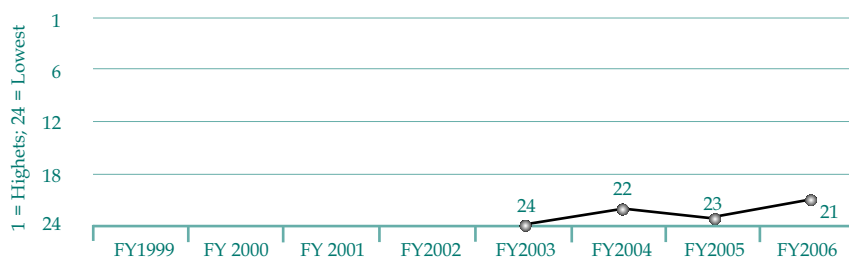
- While the report is lengthy and text-heavy, the narratives are generally clear and understandable.
- Little affirmative basis for confidence in the reliability of the data.
- Prior year data are missing for a number of measures. For others, results are not disclosed because they are sensitive or classified.

▼ B PUBLIC BENEFITS

- Only the first three of the six strategic goals are stated as outcomes and only then at high levels of generality.
- Only six of 26 strategic objectives are clearly stated as measurable outcomes.
- Most of the performance measures focus on outputs, activities, and processes as opposed to results.
- Costs are linked to the strategic goals and objectives, but not down to the performance goals or targets.

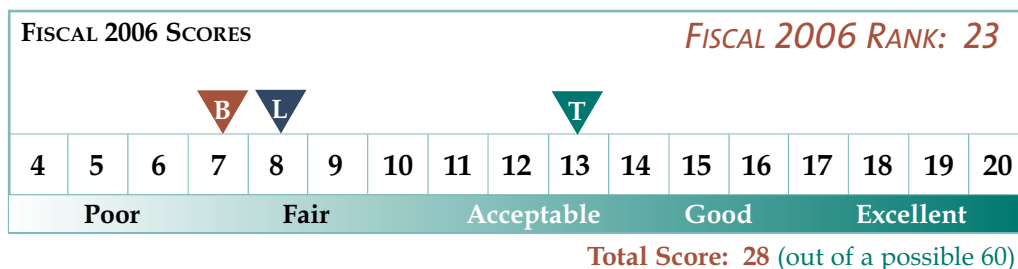
▼ L LEADERSHIP

- Lack of results-orientation in both the performance metrics and the narratives leaves the reader with little sense of concrete results that flow from the department's work.
- Report clearly discloses performance shortfalls in an initial table of aggregate results.
- Goals and objectives in the management area are exceptionally weak.
- Report describes generally adequate and informative planned improvements in both the programmatic and management areas.



RANKING HISTORY

OFFICE OF PERSONNEL MANAGEMENT



T TRANSPARENCY

- Home page has an explicit direct link to the report, which is downloadable in a single file or multiple files.
- Report's brevity reflects the limited substantive information that it conveys rather than its conciseness.
- Fiscal year results are included for all 58 goals. Data sources are provided, but the data sources and verification methods are sometimes unclear.
- Very little prior year baseline or trend data for the agency's goals.

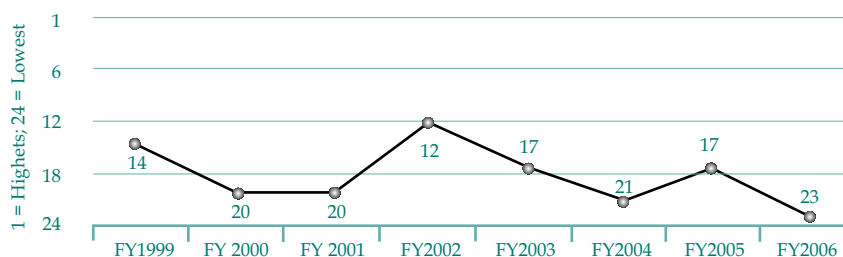
B PUBLIC BENEFITS

- None of the seven programmatic strategic goals are outcome-oriented.
- None of the goals, which double as measures, are stated as outcomes.
- Costs (and staff resources) are linked only to the strategic goals.

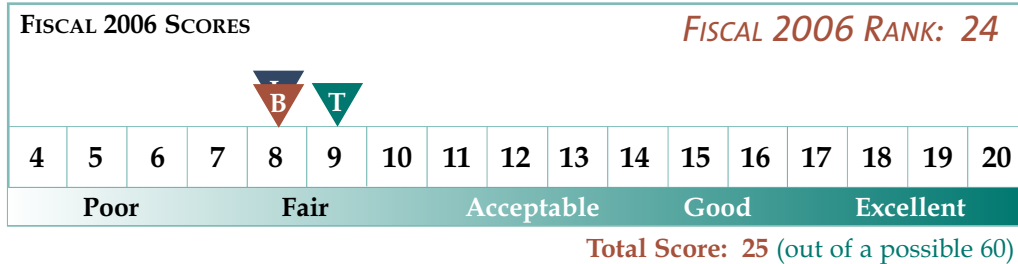
L LEADERSHIP

- Narratives do little to demonstrate that the agency is accomplishing its core mission.
- There are no performance shortfalls due to reported success on all 58 performance goals/measures; however, many goals/measures are mundane and unchallenging.
- No direct response by agency management to the inspector general's presentation on management challenges.
- Report fundamentally lacks evidence of forward-looking leadership or strategic thinking to enhance performance.

RANKING HISTORY



DEPARTMENT OF HEALTH AND HUMAN SERVICES



T TRANSPARENCY

- Home page does not have a direct link to the report. Only the hard copy version provides specific staff contacts.
- The report covers 35 performance measures, but the full set of departmental measures is not described anywhere in the report.
- The report lacks results for just over half of the measures it covers (18 of 35), and an independent auditor’s report raises issues about data accuracy.
- The report has useful prior year data for each covered measure, but the limited coverage of the department’s measures and the absence of many current results make it hard to assess performance trends.

B PUBLIC BENEFITS

- All of the programmatic strategic goals are stated as outcomes.
- Of the 35 measures, ten are clearly outcome-oriented.
- No costs allocated to the performance metrics at any level.

L LEADERSHIP

- Narratives describe some of the department’s accomplishments, but usually focus more on process and activities than results.
- Report discloses only one performance shortfall and offers no specific remedial plan.
- Inspector general’s presentation indicates progress on two management challenges, but it is difficult to determine progress on the others.
- Report has little discussion of improvement strategies for next year.



RANKING HISTORY

RESEARCH TEAM AND PROJECT DESIGN

The Scorecard project is headed by the Hon. Maurice McTigue, QSO, director of the Government Accountability Project and a distinguished visiting scholar at the Mercatus Center at George Mason University, and Dr. Jerry Ellig, a senior research fellow at the Mercatus Center.

Dr. Tyler Cowen, the Mercatus Center's general director, served as advisory director for the study. A three-member professional research team with extensive government experience completed all report review and analysis. A 10-member advisory panel, made up of experts in public management, former federal performance managers, corporate strategists, and communications experts, reviewed our evaluations and analysis.

PROJECT MANAGEMENT

Ms. Jen Wekelo, Associate Director, Government Accountability Project
Mercatus Center at George Mason University

Ms. Heather Hambleton, Program Associate, Government Accountability Project
Mercatus Center at George Mason University

ADVISORY PANEL

Jonathan D. Breul
IBM Global Business Services
Washington, DC

Jonathan D. Breul is currently a partner in IBM Global Business Services and executive director of IBM's Center for the Business of Government. He is a widely recognized expert on the policy and practice of improving government management and performance.

Formerly senior advisor to the deputy director for management in the White House Office of Management and Budget, Mr. Breul was OMB's senior career executive with primary responsibility for government-wide general management policies. He helped develop the President's Management Agenda, was instrumental in establishing the President's Management Council, and championed efforts to integrate performance information with the budget process. He led the overall implementation of the Government Performance and Results Act. In addition to his OMB activities, he helped Senator John Glenn (D-Ohio) launch the Chief Financial Officers Act. Mr. Breul also served as the U.S. delegate and vice chair of the Organization for Economic Cooperation and Development's Public Management Committee.

Mr. Breul is a fellow and member of the board of trustees of the National Academy of Public Administration and an adjunct professor at Georgetown University's graduate Public Policy Institute. He holds a master's in Public Administration from Northeastern University and a bachelor of arts from Colby College.

Veronica Campbell

*Independent Consultant
Falls Church, Virginia*

Veronica Campbell currently consults on performance, financial management, and general management issues. She retired in 2003 after a 30 year career in the federal service, where she contributed to the effectiveness of programs operated by the Departments of Labor, Interior, and Agriculture. As the director of the Center for Program Planning and Results (CPPR), she managed the Department of Labor's (DOL) strategic planning and performance-based initiatives, working with DOL executives and a dedicated inter-agency committee to foster a results-driven organizational culture. Ms. Campbell joined the CPPR after serving on the DOL Year 2000 Project Management Team from 1998 to 2000.

Prior to her administrative experience, Ms. Campbell managed and performed program evaluations and audits for 25 years in the Offices of Inspectors General (OIG) of three federal agencies. She served as the first director of the Office of Evaluations and Inspections at the Department of Labor, OIG, from 1992 to 1998. In this capacity, she offered departmental clients a range of new consultative services with an emphasis on collaborative program evaluation. From 1987 to 1992, Ms. Campbell was the regional inspector general for audit for the Labor Department's New York and Boston regions. Her early audit career at the Department of Agriculture, from 1973 to 1982, developed Ms. Campbell's analytical skills as she recommended improvements to USDA's international programs and evaluated a wide array of domestic programs.

Ms. Campbell holds a bachelor of arts in history from Barat College.

Mortimer L. Downey, III

*PB Consult, Inc.
Washington, DC*

Mortimer L. Downey, III is chairman of PB Consult, Inc., a Parsons Brinckerhoff subsidiary providing advisory and management consulting services to public and private owners, developers, financiers, and builders of infrastructure projects worldwide.

Prior to joining PB Consult, Mr. Downey served eight years as U.S. deputy secretary of transportation, the longest serving individual in that post. As DOT's chief operating officer, he developed the agency's highly regarded strategic and performance plans. During this period he also served on the President's Management Council, chaired the National Science and Technology Council's Committee on Technology, and was a member of both the Trade Promotion Coordinating Council and the National Railroad Passenger Corporation (Amtrak) board of directors. In addition to his federal service, Mr. Downey has served as executive director and chief financial officer of the New York Metropolitan Transportation Authority and as a senior manager at the Port Authority of New York and New Jersey. His legislative experience includes service on the staff of the U.S. House of Representatives' Committee on the Budget.

Mr. Downey has received numerous professional awards, including election to the National Academy of Public Administration, where he has served as chairman of the board of directors. He earned a master's in public administration from New York University, a bachelor of arts in political science from Yale University, and completed the Advanced Management Program at Harvard Business School. He has also served as an officer in the U.S. Coast Guard Reserve.

John Kamensky

*IBM Global Business Services
Washington, DC*

Mr. Kamensky is an associate partner with IBM Global Business Services and a senior research fellow for the IBM Center for the Business of Government. During 24 years of public service, he had a significant role in helping pioneer the federal government's performance and results orientation. He is passionate about creating a government that is results-oriented, performance-based, and customer-driven.

Mr. Kamensky served eight years as deputy director of Vice President Gore's National Partnership for Reinventing Government. Previous to his White House position, he worked at the U.S. Government Accountability Office for 16 years where, as an assistant director, he played a key role in the development and passage of the Government Performance and Results Act (GPRA).

Mr. Kamensky is a fellow of the National Academy of Public Administration. He earned a master's degree in public affairs from the Lyndon B. Johnson School of Public Affairs at the University of Texas.

Patricia Kelley, CISA

*SiloSmashers
Fairfax, Virginia*

Patricia Kelley is vice president for planning, measurement, and analysis for SiloSmashers, a management consulting firm that specializes in strategic planning and performance management. She held senior management positions with the Federal Reserve Board, advising the governors on policy issues that affect the efficiency and effectiveness of the board's operations. She also worked extensively with the Federal Reserve Banks on automation and payment system policy matters and acted as the liaison to other federal banking regulators.

Prior to joining the Federal Reserve, Ms. Kelley held various positions with the U.S. Government Accountability Office and evaluated the effectiveness of programs in the Departments of Defense, Treasury, Agriculture, the Government Printing Office, and the Overseas Private Investment Corporation. In 2000, Ms. Kelley co-authored *The Business of Government: Strategy, Implementation, and Results* with Dr. Thomas Kessler. She has provided management consulting support to more than 30 federal agencies.

She holds a master of science in computer systems management and a bachelor of science in accounting from the University of Maryland. She is also a graduate of its Stonier School of Banking. She is working on her doctor of public administration dissertation at the University of Baltimore.

Thomas Kessler, DBA, CISA

SiloSmashers

Fairfax, Virginia

Dr. Thomas Kessler is a SiloSmashers senior consultant. From 1983 to 1996, Dr. Kessler served as a manager at the Board of Governors of the Federal Reserve System. He frequently advised senior officials and provided recommendations for enhancing mission-critical business processes. Prior to joining the Federal Reserve, Dr. Kessler was employed at Westinghouse Electric Corporation's Manufacturing Systems and Technology Center and at the Maryland State Government's Judicial Information Systems.

Over the past several years, Dr. Kessler has trained and facilitated outcome-based performance measurement and planning sessions for many federal agencies, including the Departments of Labor, Treasury, Justice, Veterans Affairs, and Agriculture, as well as the Commodity Futures Trading Commission and the Federal Aviation Administration. He co-authored *The Business of Government: Strategy, Implementation, and Results* with Patricia Kelley and is a frequent speaker at professional conferences throughout the United States.

Dr. Kessler earned a doctor of business administration degree from Nova Southeastern University, has a master of business administration from University of Baltimore, and is a certified information systems auditor.

Sarah E. Nutter, PhD

Associate Professor of Accounting

George Mason University, School of Management

Fairfax, Virginia

Sarah Nutter is an associate professor of accounting in George Mason University's School of Management. Prior to joining the GMU faculty in 1995, she worked as an economist at the Internal Revenue Service. Professor Nutter teaches a variety of courses in accounting and taxation in undergraduate, MBA, and executive MBA programs. She recently received the Outstanding Faculty Member teaching award from George Mason's executive MBA program.

Professor Nutter's research interests include investigating the impact of changing decision rules on individual and business behavior. Her research focuses primarily on the impact of taxes and tax structures on individuals and businesses. She has written extensively and has published in the *Journal of the American Taxation Association*, *Journal of Accounting and Economics*, *Advances in Taxation*, and the *Statistics of Income Bulletin*. One of her articles won the 1998-99 American Taxation Association's Outstanding Manuscript Award.

Professor Nutter earned a bachelor of science from Ferris State University, and a master of business administration and PhD from Michigan State University.

John M. Palguta

*Partnership for Public Service
Washington, DC*

As vice president for policy and research at the Partnership for Public Service, John Palguta is responsible for the development and implementation of a comprehensive program of research and analysis on human capital issues in the federal government.

Prior to joining the Partnership in December 2001, Mr. Palguta was a career member of the federal senior executive service as director of policy and evaluation for the U.S. Merit Systems Protection Board (MSPB), the culmination of a federal career spanning almost 34 years of progressively responsible experience in federal human resources management and public policy issues.

Mr. Palguta earned a bachelor of arts in sociology from California State University at Northridge and a master of public administration degree from the University of Southern California. He is active in a number of professional associations and is a fellow of the National Academy of Public Administration, vice chair of the Coalition for Effective Change, and a past president of the Federal Section of the International Public Management Association for Human Resources. He has published a number of articles on federal human resources management issues and is a frequent speaker at professional conferences and other forums. He is a recipient of MSPB's highest honor, the Theodore Roosevelt Award.

Paul L. Posner

*Department of Public and International Affairs
George Mason University
Arlington, Virginia*

Paul L. Posner is the director of the Public Administration Program at George Mason University. He has completed a career at the U.S. Government Accountability Office (GAO) where he served as managing director for federal budget and intergovernmental issues. He led GAO's work on the long term outlook for the federal budget and emerging challenges for public sector finances at federal, state, and local levels.

Dr. Posner has published articles on public budgeting and federalism issues in various academic journals and books and is the author of *The Politics of Unfunded Federal Mandates*, published by Georgetown University Press in 1998. He earned his PhD in Political Science from Columbia University. He was elected a Fellow in the National Academy of Public Administration and chairs their Federal Systems Panel. Prior to his current position with George Mason, Dr. Posner was a senior adjunct lecturer at Johns Hopkins, Georgetown, and George Washington Universities. He received the James Blum Award for outstanding budget practitioner from the Association of Budget and Program Analysts. He was recently elected as vice president of the American Society for Public Administration and will become president of the organization in 2009.

John Sacco, PhD

*Department of Public and International Affairs
George Mason University
Fairfax, Virginia*

John Sacco is an associate professor at George Mason University's Department of Public and International Affairs. Prior to joining GMU, he was a program analyst for the U.S. Department of Housing and Urban Development.

Dr. Sacco is currently working on a government and nonprofit accounting and financial reporting textbook that will be accessible to students on the Internet. In 1999, along with several scholars, he published a policy paper about the major government reforms undertaken by New Zealand during the 1980s and 1990s. The paper compared New Zealand's integrated, business-like financial management system with the emerging attempts by the U.S. federal and state governments to use accounting and performance measures similar to those in private business. In 2000, Dr. Sacco published work in the *Association of Governmental Accountants' Journal* analyzing the most dramatic changes in state and local governmental accounting and finance practices that have taken place in the 20th century. His work on the evolution of end-user computing is forthcoming. In addition to his writing, Dr. Sacco has consulted for several state and local governments and Certified Public Accounting (CPA) firms, including contact work with CPA firms on finance and accounting for the Chinese government.

RESEARCH TEAM**Valerie J. Richardson**

*TreWyn, Inc
Germantown, Maryland*

Valerie Richardson is president and chief executive officer of TreWyn, a financial and strategic management practice based in Germantown, Maryland.

Formerly she served as associate director of the Center for Improving Government Performance at the National Academy of Public Administration. Ms. Richardson has been a practitioner of public sector performance management for well over a decade; she managed the Results Act of 1993 pilot and implementation efforts at the Patent and Trademark Office.

Ms. Richardson is widely published in public administration and accounting journals and proceedings and is the author of the book *Annual Performance Planning – A Manual for Public Agencies*. She was awarded the Best Paper Award—Highest Quality Rating at the Co-operation & Competition Conference in Vaxjö, Sweden in 2002 and is a contributing chapter author to the book *Co-Operation & Competition “Co-opetition” – The Organization of the Future*.

She is a graduate of the John F. Kennedy School of Government, Harvard University's Senior Executive Fellows Program. Ms. Richardson holds a master of arts in financial management from the University of Maryland and undergraduate degrees from Trenton State College in Public Administration and Political Science. She is currently pursuing her doctorate in public administration from the University of Baltimore.

Michael D. Serlin

Independent Consultant

Alexandria, Virginia

Michael D. Serlin is currently writing and consulting on public service change and participating in a number of volunteer activities after a 36-year federal career. He retired in 1994 from the U.S. Department of the Treasury's Financial Management Service. As an assistant commissioner, he directed innovations in electronic funds transfer and financial operations for the U.S. government and initiated entrepreneurial administrative support across agencies.

Mr. Serlin led the financial management team for the National Performance Review (Reinventing Government) Task Force, most of whose recommendations were incorporated in the Government Management Reform Act of 1994. Among other things, the law required audited financial statements for all major agencies and introduced franchising—competitive cross-servicing of agency administrative support.

A former senior executive service presidential rank award winner, Mr. Serlin is a principal with the Council for Excellence in Government; he has participated in its reviews of agency strategic performance draft plans. He has contributed frequent articles and research papers on entrepreneurial government, executive mobility, and other government change efforts to magazines and professional journals.

Mr. Serlin earned his bachelor of arts in political science from Stanford University.

Kyle McKenzie

Mercatus Center at George Mason University

Arlington, Virginia

Kyle McKenzie is a research fellow with the Mercatus Center at George Mason University. He is involved with both the Social Change Project and the Government Accountability Project. His issue areas include energy market regulation, homeland security policy, and creative ways to enhance transparency of government activity.

He graduated from Beloit College in Wisconsin with degrees in economics and psychology. After working at the Mercatus Center as a Charles G. Koch Summer Fellow, he migrated to the mid-Atlantic and accepted his current position.

ABOUT THE AUTHORS

DR. JERRY ELLIG (jellig@gmu.edu) has been a senior research fellow at the Mercatus Center at George Mason University since 1996. Between August 2001 and August 2003, he served as deputy director and acting director of the Office of Policy Planning at the Federal Trade Commission. Dr. Ellig has also served as a senior economist for the Joint Economic Committee of the U.S. Congress and as an assistant professor of economics at George Mason University.

Dr. Ellig co-authored several previous editions of the Mercatus Center's Performance Report Scorecard. He has also published numerous articles on government regulation and business management in both scholarly and popular periodicals, including the *Journal of Regulatory Economics*, *Federal Communications Law Journal*, *Business & Politics*, *CommLaw Conspectus*, *Managerial and Decision Economics*, *Antitrust Bulletin*, *Competitive Intelligence Review*, *Journal of Private Enterprise*, *Texas Review of Law & Politics*, *Wall Street Journal*, *New York Times*, *Barron's*, and *Washington Post*. He has edited and co-authored several books, including *Dynamic Competition and Public Policy* (Cambridge, 2001), *New Horizons in Natural Gas Deregulation* (Praeger, 1996), and *Municipal Entrepreneurship and Energy Policy* (Gordon & Breach, 1994).

Dr. Ellig earned his PhD and MA in economics from George Mason University in Fairfax, VA, and his BA in economics from Xavier University in Cincinnati, OH.

THE HONORABLE MAURICE P. MCTIGUE, QSO (mmctigue@gmu.edu) is the director of the Government Accountability Project and a distinguished visiting scholar at the Mercatus Center at George Mason University. He joined George Mason University in 1997 after completing his term as New Zealand's ambassador to Canada. Previously, as a cabinet minister and a member of Parliament, he led an ambitious and successful effort during the 1980s and 1990s to restructure New Zealand's public sector and to revitalize its stagnant economy. In 1999, in recognition of his public service, Her Majesty Queen Elizabeth II bestowed upon Mr. McTigue the prestigious Queen's Service Order during a ceremony at Buckingham Palace.

At the Mercatus Center, Maurice McTigue shares with U.S. policy makers his practical experience and lessons learned. Since 1997 he has worked with senior administration officials, members of Congress, and scores of senior agency executives on applying the principles of economics, transparency, and accountability in the public sector. He frequently speaks at conferences on performance issues, testifies before congressional committees on issues of government reform, and writes on the importance of transparency in reforming government.

HENRY WRAY (henrywray@verizon.net) is a visiting fellow with the Mercatus Center's Government Accountability Project. He recently completed a distinguished career in Washington, DC, where he served for over 30 years on the staff of the U.S. Government Accountability Office (GAO) and the United States Congress. At GAO he started in the Office of the General Counsel, where he served for many years as an associate general counsel, overseeing a group of attorneys who provided legal support to one of GAO's auditing divisions. He also served for four years as GAO's ethics counselor. In addition, for two years he headed the GAO audit group responsible for evaluations of the U.S. Department of Justice, the law enforcement components of the U.S. Department of Treasury, and the Federal Judiciary. He completed his GAO career while serving several years as a detailee to the U.S. Congress, during which time he was on the professional staff of the House Budget Committee, the House Committee on Government Reform, and the Senate Governmental Affairs Committee. After retiring from GAO, he served as counsel for the Senate Governmental Affairs Committee. In 2001, he became senior counsel to the House Subcommittee on Government Efficiency, Financial Management and Intergovernmental Relations.

Prior to coming to Washington, Mr. Wray served as deputy attorney general in the Commonwealth of Pennsylvania Department of Justice.

Henry Wray earned his bachelor of arts in political science from Washington & Jefferson College and a juris doctor with honors from the National Law Center, George Washington University. He remains an active member of the bar of the District of Columbia.

MERCATUS CENTER

GEORGE MASON UNIVERSITY

3301 North Fairfax Drive, Suite 450

Arlington, Virginia 22201

Tel: 703-993-4930

Fax: 703-993-4935

www.mercatus.org