

10TH ANNUAL PERFORMANCE REPORT SCORECARD

WHICH FEDERAL AGENCIES BEST
INFORM THE PUBLIC?

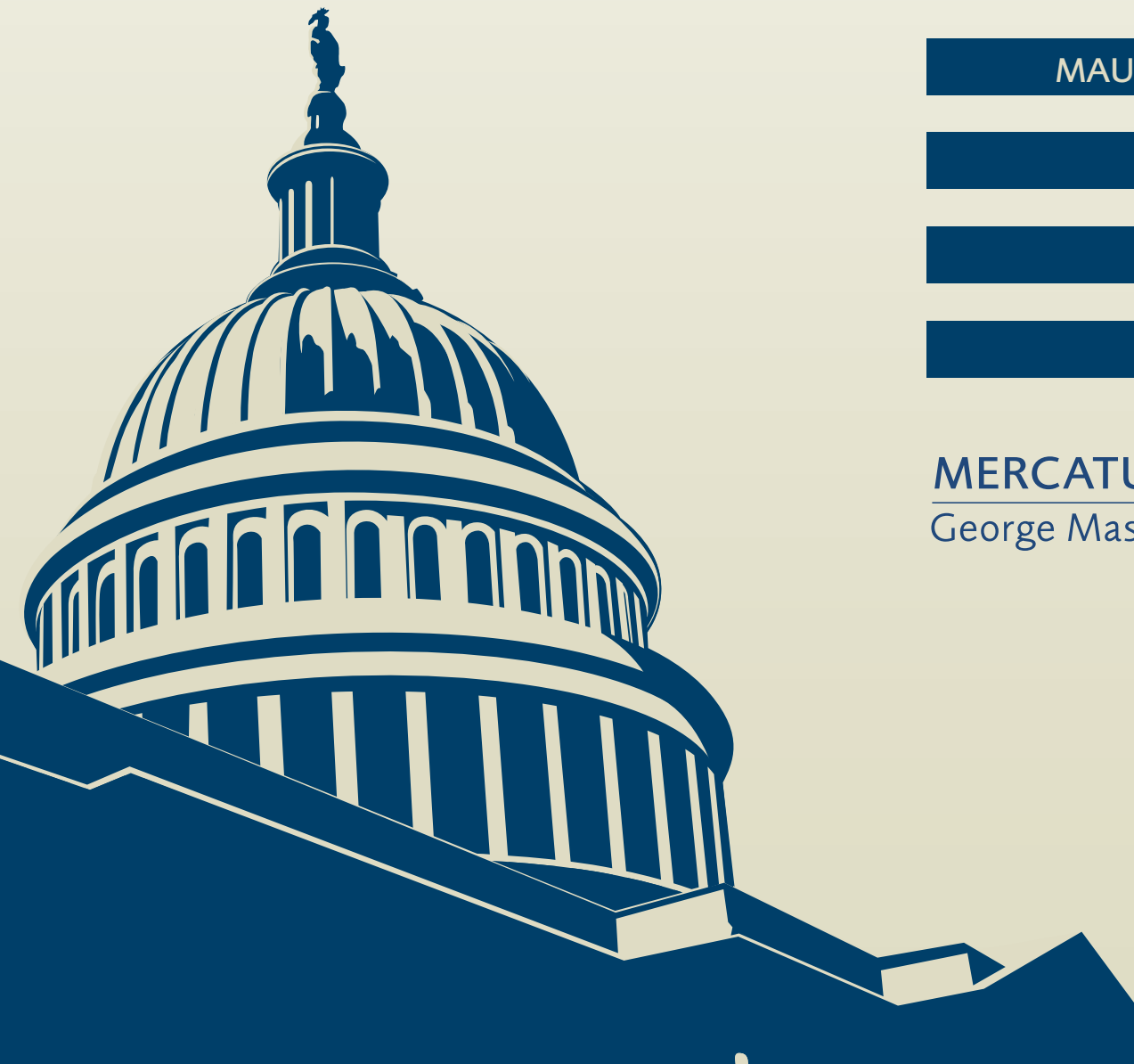
MAURICE MCTIGUE

HENRY WRAY

JERRY ELLIG

MAY 2009

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 to 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 the 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 that demonstrates 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. Fiscal year 2008 marks the tenth year agencies have produced annual performance reports. For fiscal year 2008, nine agencies again opted for a “pilot” format that allowed them to publish performance information separately from financial information and produce a shorter “citizens’ report” summarizing the two.

Researchers at the Mercatus Center at George Mason University initiated the Scorecard series in fiscal year 1999 to foster continuous improvement in the quality of disclosure in agencies’ annual performance reports. This study is thus our tenth annual evaluation of the performance and accountability reports produced by the 24 agencies covered under the Chief Financial Officers Act. 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 forward-looking leadership that uses annual performance information to devise strategies for improvement. An expert team evaluated each report on 12 criteria—4 each for transparency, public benefits, and leadership.

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

Best practices have improved substantially in ten years. A re-evaluation of the best four reports from fiscal year 1999 finds that these reports would rank well below average when judged on the same 12 criteria by fiscal year 2008’s higher standards. Qualitative analysis of best practices reveals substantial improvements since fiscal year 1999. Table 1 summarizes the current best practices and improvement in best practices over time. Quantitative analysis suggests that the average quality of performance reports may have improved by about 75 percent since fiscal year 1999.

Congress and OMB may affect the rate of progress. Between fiscal year 1999 and fiscal year 2008, report scores tended to improve more at agencies where lower percentages of managers surveyed by the Government Accountability Office (GAO) identify “lack of ongoing congressional commitment and support for using performance information” as a hindrance to performance management. Similarly, report scores tended to improve more at agencies where lower percentages of managers surveyed by GAO identify “concern that [the Office of Management and Budget] will micromanage programs” as a hindrance to performance management.

TABLE 1

Evolution of Best Practices, Fiscal Year 1999 – Fiscal Year 2008

FY 1999	FY 2008
Criterion 1: Accessibility <ul style="list-style-type: none"> Report online An obvious link from agency home page to the report 	Those features, plus <ul style="list-style-type: none"> Report posted in a timely fashion after due date Direct (single click) link from home page Downloadable as single or multiple files Contact information provided for questions/comments
Criterion 2: Readability <ul style="list-style-type: none"> Relative lack of jargon Lengthy reports easy to scan due to headings, sidebars, tables, and charts 	Those features, plus <ul style="list-style-type: none"> Performance section focuses on key measures Goals, measures, and results in tables Secretary's letter describes fiscal year 2008 results for each goal as well as achievements over time Citizens' reports include concise summaries of results and reader-friendly links to more information
Criterion 3: Verification and validation <ul style="list-style-type: none"> Discussion of validation procedures Acknowledgement of data deficiencies Outline of plans to develop more or better data 	Those features, plus <ul style="list-style-type: none"> Assessment of data quality for each goal Data source provided for each measure Data definitions, verification and validation information, and limitations discussed for each measure
Criterion 4: Baseline and trend data <ul style="list-style-type: none"> 5 to 10 years of trend data Several years of actual data combined with goals for next several years 	New standard <ul style="list-style-type: none"> Multiple years of data include targets, actual results, and costs Long-range targets or forecasts provided for each measure
Criterion 5: Outcome-oriented goals <ul style="list-style-type: none"> Most or all strategic goals are outcome oriented 	New standard <ul style="list-style-type: none"> Most goals are clear, measurable outcomes Outcomes are intuitively meaningful to the public
Criterion 6: Outcome measures <ul style="list-style-type: none"> Most measures are outcomes or related to outcomes 	New standard <ul style="list-style-type: none"> Most measures are final or intermediate outcomes

<p>Criterion 7: Agency-affected outcomes</p> <ul style="list-style-type: none"> • Report explains how activities contribute to results in specific cases • Report acknowledges other factors that might affect results 	<p>Those features, plus</p> <ul style="list-style-type: none"> • Report consistently describes how activities led to observed results • Performance metrics are highly outcome oriented
<p>Criterion 8: Results linked to costs</p> <ul style="list-style-type: none"> • Charts show costs, personnel counts, and performance measures for program areas 	<p>New standard</p> <ul style="list-style-type: none"> • Costs broken down by strategic goal and most individual performance measures • This information is provided for several years
<p>Criterion 9: Vision</p> <ul style="list-style-type: none"> • Each strategic goal identifies a result of interest to citizens and states how the department intends to accomplish it 	<p>Those features, plus</p> <ul style="list-style-type: none"> • Narratives cite major accomplishments that affect citizens' quality of life • Narratives linked to outcome-oriented performance measures demonstrating that the narratives describe typical results
<p>Criterion 10: Explanation of failures</p> <ul style="list-style-type: none"> • Failures to meet targets explained • Plans to remedy failures explained 	<p>New standard</p> <ul style="list-style-type: none"> • All performance shortfalls identified, along with plans and a timeline to remedy them • Improvement plans offered even when targets were met
<p>Criterion 11: Management challenges</p> <ul style="list-style-type: none"> • Thorough discussions of major management challenges identified by the agency inspector general and GAO 	<p>Those features, plus</p> <ul style="list-style-type: none"> • Inspector general's report lists major management challenges and assesses agency's progress on them • Agency provides self-assessment of progress and timeline for resolving each challenge • Report explains how each challenge affects strategic goals
<p>Criterion 12: Improvement plans</p> <ul style="list-style-type: none"> • Discussion of each measure includes fiscal year 1999 results, projected fiscal year 2000 performance, and actions planned for fiscal year 2001 	<p>New standard</p> <ul style="list-style-type: none"> • Improvement strategies presented for all shortfalls and major management challenges • Report describes broader challenges the agency faces and plans for addressing them

Note: "Best practices" are simply the best observed in that fiscal year, not necessarily the best that could imaginably be achieved. These best practices are usually found only in a minority of the reports for each fiscal year.

Other key findings in this year's Scorecard include the following:

Top three trade places: The top three reports—Labor, Veterans Affairs, and Transportation—finished in that order, within three points of each other. Labor's report received a 56 out of 60 possible points, the highest score ever awarded. Reports from Homeland Security and the Nuclear Regulatory Commission tied for fourth place, each earning 40 points.

Meaningful improvements: Five agencies made meaningful improvements in the content of their reports that led to higher scores in fiscal year 2008 than in fiscal year 2007. These were State (+6 points), Energy (+5 points), Interior (+5 points), Education (+4 points), and USAID (+4 points). The first four leapt by 8 to 12 places in the rankings, and USAID rose 4 places.

Average score rises: The average score increased by 1.5 points, from 34.6 in fiscal year 2007 to 36.1 in fiscal year 2008.

More satisfactory reports: Thirteen reports achieved a "satisfactory" score of 36 or better in fiscal year 2008, up from seven reports in fiscal year 2007.

Reports not satisfactory for a majority of spending: Sixty percent of federal spending was covered by reports scoring below satisfactory in fiscal year 2008—a slight drop from 65 percent in fiscal year 2007.

Concerns about Recovery Act accountability: Only 16 percent of appropriations in the American Recovery and Reinvestment Act, signed in February 2009, go to agencies whose reports scored "very good" (48 or above) in fiscal year 2008.

Pilot format reverses losses: Average scores for agencies using the pilot reporting format increased by 9 percent in fiscal year 2008, almost completely reversing their drop in fiscal year 2007. This suggests that the pilot format is a workable approach that has overcome many of the difficulties experienced in its first year.

Few agencies have inherent advantages: Larger agencies tend to score slightly higher on our Scorecard, but any size advantage disappears for the three largest: Defense, Social Security, and Health and Human Services. Agencies that provide more direct federal services seem to have no advantage, and grant-giving agencies seem to suffer no disadvantage. Agencies whose policy views were evaluated as more liberal, according to a recent expert survey, seem to score slightly better, but this difference is very small.

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 Affairs	VA

TABLE 2

Scorecard Summary & Ranking for Fiscal Year 2008

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

	TRANSPARENCY	PUBLIC BENEFITS	LEADERSHIP	TOTAL	RANK
Labor	20	19	17	56	1
Veterans Affairs	19	16	19	54	2
Transportation	16	20	17	53	3
Homeland Security	15	13	12	40	4
NRC	15	13	12	40	4
Education	14	12	11	37	6
Interior	16	10	11	37	6
State	15	10	12	37	6
Treasury	14	10	13	37	6
Energy	13	11	12	36	10
EPA	13	11	12	36	10
Health & Human Services	13	13	10	36	10
USAID	15	10	11	36	10
Commerce	15	10	10	35	14
Justice	15	8	11	34	15
Agriculture	12	10	11	33	16
GSA	11	12	9	32	17
NSF	15	7	10	32	17
Social Security	12	8	12	32	17
NASA	11	8	12	31	20
OPM	11	8	9	28	21
HUD	11	8	8	27	22
Defense	11	7	8	26	23
SBA	8	8	6	22	24
Average	13.8	10.9	11.5	36.1	
Median	14.0	10.0	11.0	36.0	

 Pilot Agencies

TABLE 3

Fiscal Year 2008 Scores & Rankings: Comparison to Fiscal Year 2007						
Highest Rank = 1; Lowest = 24. Maximum Possible Score = 60; Minimum = 12.						
	Fiscal Year 2008		Fiscal Year 2007		CHANGE IN SCORE	CHANGE IN RANKING
	TOTAL SCORE	RANK	TOTAL SCORE	RANK		
Labor	56	1	53	2	+3	+1
Veterans Affairs	54	2	51	3	+3	+1
Transportation	53	3	55	1	-2	-2
Homeland Security	40	4	37	5	+3	+1
NRC	40	4	39	4	+1	0
Education	37	6	32	14	+5	+8
Interior	37	6	32	14	+5	+8
State	37	6	31	18	+6	+12
Treasury	37	6	35	8	+2	+2
Energy	36	10	31	18	+5	+8
EPA	36	10	33	11	+3	+1
Health & Human Services	36	10	37	5	-1	-5
USAID	36	10	32	14	+4	+4
Commerce	35	14	37	5	-2	-9
Justice	34	15	35	8	-1	-7
Agriculture	33	16	33	11	0	-5
GSA	32	17	34	10	-2	-7
NSF	32	17	31	18	+1	+1
Social Security	32	17	33	11	-1	-6
NASA	31	20	31	14	0	-6
OPM	28	21	27	22	+1	+1
HUD	27	22	31	18	-4	-4
Defense	26	23	17	24	+9	+1
SBA	22	24	22	23	0	-1
Average	36.13		34.54		1.58	
Median	36.00		33.00		3.00	

 Pilot Agencies

Note: NASA used the pilot format in 2007 but not in 2008. OPM used the pilot format in 2008 but not in 2007.



CONTENTS

Introduction	1
Evaluation and Scoring	5
Scoring Summary	7
Update on the Pilot Format	17
Assessing the “Snapshots”	25
Do Some Agencies Have Inherent Advantages?	27
Strongest and Weakest Scores	33
Agency-by-Agency Scoring Summaries	57
Department of Labor	58
Department of Veterans Affairs	59
Department of Transportation	60
Department of Homeland Security	61
Nuclear Regulatory Commission	62
Department of Education	63
Department of the Interior	64
Department of State	65
Department of the Treasury	66
Department of Energy	67
Environmental Protection Agency	68
Department of Health & Human Services	69
U.S. Agency for International Development	70
Department of Commerce	71
Department of Justice	72
Department of Agriculture	73
General Services Administration	74
National Science Foundation	75
Social Security Administration	76
National Aeronautics & Space Administration	77
Office of Personnel Management	78
Department of Housing & Urban Development	79
Department of Defense	80
Small Business Administration	81
Appendix: Detailed Explanation of Evaluation and Scoring	83
Research Team and Project Design	89



INTRODUCTION

Outcome-oriented performance measurement isn't just a good idea; it's the law.

The Government Performance and Results Act (GPRA) requires federal agencies to produce strategic plans with performance measures, annual performance plans with performance goals, and annual performance reports that measure progress toward those goals. Measures are supposed to track the agency's "outputs, service levels, and outcomes."¹

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 by the 24 agencies covered by the Chief Financial Officers Act, which accounted for virtually all federal outlays in fiscal year 2008.² The 24 comprise all cabinet departments plus the largest independent agencies.

Compared to 1999, we've come a long way. In fiscal year 1999, an agency could receive the highest possible score for accessibility of its report simply by posting the report on its Web page and making it easy to find. Today, that is a minimal requirement; best practices include making the report available on time, downloadable as a single document or individual sections, and providing clear information about whom to contact with comments or questions. Similarly, in fiscal year 1999, a report received the top score on linking results to costs because it broke down costs and personnel by program area. Today, the best reports break costs down by individual performance measure and present the information for multiple years.

To celebrate the tenth anniversary of GPRA performance reporting, our research team reevaluated the top four reports from fiscal year 1999 using fiscal year 2008's standards. The best report from fiscal year 1999, USAID, would have ranked only 16th in fiscal year 2008. Transportation and Veterans Affairs would have tied for 20th place, and Education would have finished 22nd. Based on this scoring experiment, we think it is safe to say that the average quality of performance reporting has increased by at least 75 percent.³ Many reports have doubtless improved by much larger margins.

Despite this progress, outcome accountability for federal expenditures remains far short of ideal. In fiscal year 2008, 60 percent of federal expenditures (other than interest) were still covered by reports that received a score below the "satisfactory" level of 36, as figure 1 demonstrates.⁴ This is a slight improvement from fiscal year 2007, when 65 percent of spending was covered by unsatisfactory reports. Only 7 percent of non-interest spending, or \$207 billion, was covered by reports with a score of "very good"—48 or better.⁵ This percentage is unchanged from last year.

1. GPRA sec. 1115. Emphasis added.

2. The principal parts of government not included in these 24 agencies are the judiciary, the legislature, the executive office of the president, and the independent agencies not among the 24 CFO Act agencies.

3. Details of the calculation can be found on 13.

4. A score of 3 on a criterion corresponds to a "satisfactory" rating. A report with an average score of 3 across all 12 criteria would earn a 36. Thus, a report must score 36 or higher to be classified as "satisfactory." For further explanation, see the appendix.

5. 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 the appendix.

Record budget deficits prompted by the financial crisis and recession make outcome accountability all the more imperative. One large part of the expanded deficit stems from the \$787 billion worth of spending and tax reductions in the American Recovery and Reinvestment Act, passed in February 2009. The federal government has pledged “full transparency and accountability” in carrying out this legislation.⁶ Guidance from the Office of Management and Budget (OMB) instructs agencies to account for the results of Recovery Act spending using their GPRA performance measures to the maximum extent possible.⁷ Given this focus, it is worth considering how the quality of GPRA reporting matches up with Recovery Act spending and tax cuts.

Figure 2 classifies the Recovery Act’s \$334 billion in appropriations (listed in Division A of the legislation) according to the scores each agency received on this Scorecard for fiscal year 2008. Reports with scores in the “very good” range (48+ points) are most likely to achieve “full transparency and accountability.” But as figure 2 shows, only 16 percent of the appropriations in the Recovery Act go to agencies whose reports met this standard in fiscal year 2008.

About 71 percent of appropriations go to agencies whose reports received a “satisfactory” score of 36 or better in fiscal year 2008. This looks like a significant improvement compared to the percentage reported in a March 2009 *Mercatus on Policy* brief, which found that 74 percent of spending went to agencies whose reports scored *below* satisfactory in the fiscal year 2007 Scorecard.⁸ Most of the difference occurs because two agencies receiving large amounts of stimulus money improved their scores by several points in fiscal year 2008. Education’s score rose from 32 to 37, and Energy’s rose from 31 to 36.

The Recovery Act also makes direct grants to states and expands federal payments to individuals (in Division B of the legislation). It is not clear if the results of these expenditures, totaling approximately \$270 billion, will also be tracked using GPRA outcome measures. Another \$212 billion of the Recovery Act is tax provisions, which are still not covered by GPRA performance measures 16 years after passage of the law.⁹

For at least 787 billion new reasons, transparent reporting of outcomes is now more important than ever.

6. See <http://www.recovery.gov>.

7. Office of Management and Budget, “Memorandum for the Heads of Departments and Agencies,” February 18, 2009, <http://www.recovery.gov/files/Initial%20Recovery%20Act%20Implementing%20Guidance.pdf>.

8. Christina Forsberg and Stefanie Haefele-Balch, “Accountability and Transparency in the American Recovery and Reinvestment Act,” *Mercatus on Policy* 38 (March 2009), 3, <http://www.mercatus.org/PublicationDetails.aspx?id=26482>.

9. The 2009 Analytical Perspectives volume of the Budget of the United States Government explores some broad suggestions for measuring the outcomes of tax expenditures under GPRA and promises to keep working on it: “The above illustrative discussion, although broad, is nevertheless incomplete, omitting important details both for the provisions mentioned and the many that are not explicitly cited. Developing a framework that is sufficiently comprehensive, accurate, and flexible to reflect the objectives and effects of the wide range of tax expenditures will be a significant challenge. OMB, Treasury, and other agencies will work together, as appropriate, to address this challenge.” *Budget of the United States Government, Fiscal Year 2009, Analytical Perspectives*, 328.

Figure 1: Fiscal Year 2008 Spending (\$Billions) and Quality of Disclosure

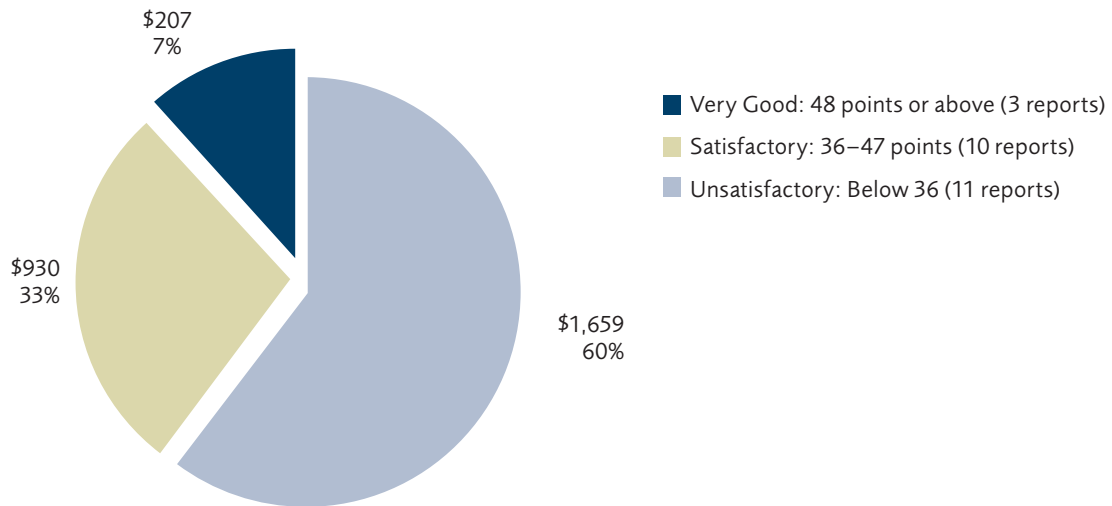
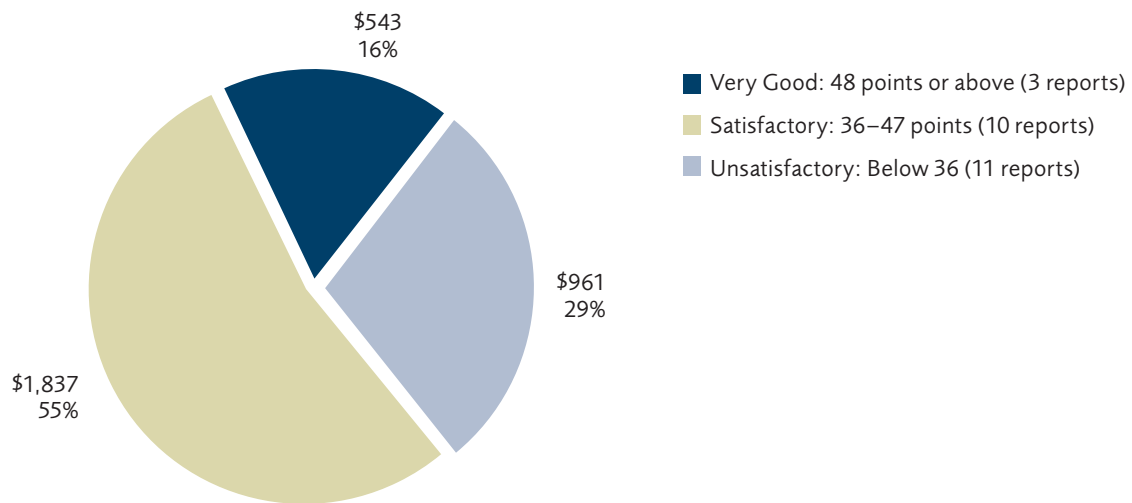


Figure 2: Recovery Act Appropriations (\$Billions) and Quality of Disclosure



EVALUATION AND SCORING

FOR THE PAST ten years, the Mercatus Center's Performance Report Scorecard has evaluated how well performance reports disclose to the public the outcomes the agencies produce. The Scorecard does not offer an opinion on the quality of agencies' performance, nor does it express views on what activities the government should or should not undertake. We assess the quality of disclosure, not the quality of results.

Reports are evaluated according to 12 criteria grouped into three categories:

1. **Transparency:** How easily can a non-specialist find and understand the report?
2. **Public benefits:** How well does the report document the outcomes the agency produces for the public and compare them with costs?
3. **Leadership:** How well does the report demonstrate that agency managers use performance information to make decisions?

On each of the 12 criteria, a report can achieve a score ranging from 1 point (no useful content) to 5 points (potential best practice). Thus, total report scores can range from 12 to 60 points. Over time, the research team tightens its scoring standards to reflect the development of new best practices.

A more detailed description of the scoring criteria and evaluation methods is available in previous editions of the Scorecard and is reproduced as an appendix this year.

Scorecard Criteria

TRANSPARENCY

1. **Accessibility:** Is the report easily accessible via the Internet and easily identified?
2. **Readability:** Is the report easy for a layperson to read and understand?
3. **Verification and Validation:** Are the performance data valid, verifiable, and timely?
4. **Baseline and Trend Data:** Did the agency provide baseline and trend data to put its performance measures in context?

PUBLIC BENEFITS

5. **Outcome Goals:** Are the goals and objectives stated as outcomes?
6. **Outcome Measures:** Are the performance measures valid indicators of the agency's impact on its outcome goals?
7. **Agency-Affected Outcomes:** Does the agency demonstrate that its actions have actually made a significant contribution toward its stated goals?
8. **Linkage to Costs:** Did the agency link its goals and results to costs?

LEADERSHIP

9. **Vision:** Does the report show how the agency's results will make this country a better place to live?
10. **Explain Failures:** Does the agency explain failures to achieve its goals?
11. **Major Management Challenges:** Does the report adequately address major management challenges?
12. **Improvement Plans:** Does the report describe changes in policies or procedures to do better next year?

SCORING SUMMARY

TOP THREE SWAP PLACES

THE TOP THREE reports—from Labor, Veterans Affairs, and Transportation—continue to outdistance the others by a large margin. But once again this year, the top three switched positions. Labor’s report took first place with a score of 56, the highest score ever awarded in the ten-year history of the Scorecard. Veterans Affairs took second with a score of 54, and Transportation earned third place with a score of 53. Thus, just three points separate these reports. Homeland Security and the NRC tied for fourth place with 40 points—a modest improvement over their scores for fiscal year 2007.

MOST-IMPROVED REPORTS

THE BIGGEST GAINERS among agencies that made meaningful improvements to their reports were State (+6 points), Energy (+5 points), and Interior (+5 points). Education and USAID each gained 4 points. Notably, three of these used the pilot reporting format that OMB first introduced in fiscal year 2007 instead of producing a traditional performance and accountability report (PAR). There is no striking pattern explaining the score increases for the pilot agencies; greater access to information from source documents (particularly performance reports) probably is the main factor. Education’s gain resulted primarily from a new strategic plan, which significantly enhanced the department’s performance metrics.

Defense was the biggest individual gainer, increasing its score 9 points from 17 last year to 26 this year. Nevertheless, Defense still produced one of the worst reports. The department’s score increased from 1 to 5 on criterion 1 (accessibility) by posting its citizens’ report on time and satisfying the other evaluation factors for that criterion. The remaining gains were increases from 1 (no relevant content) to 2 (minimal content) in five other categories. This report still has far to go before it regains the ground it has lost since fiscal year 2006 when it earned a score of 32.

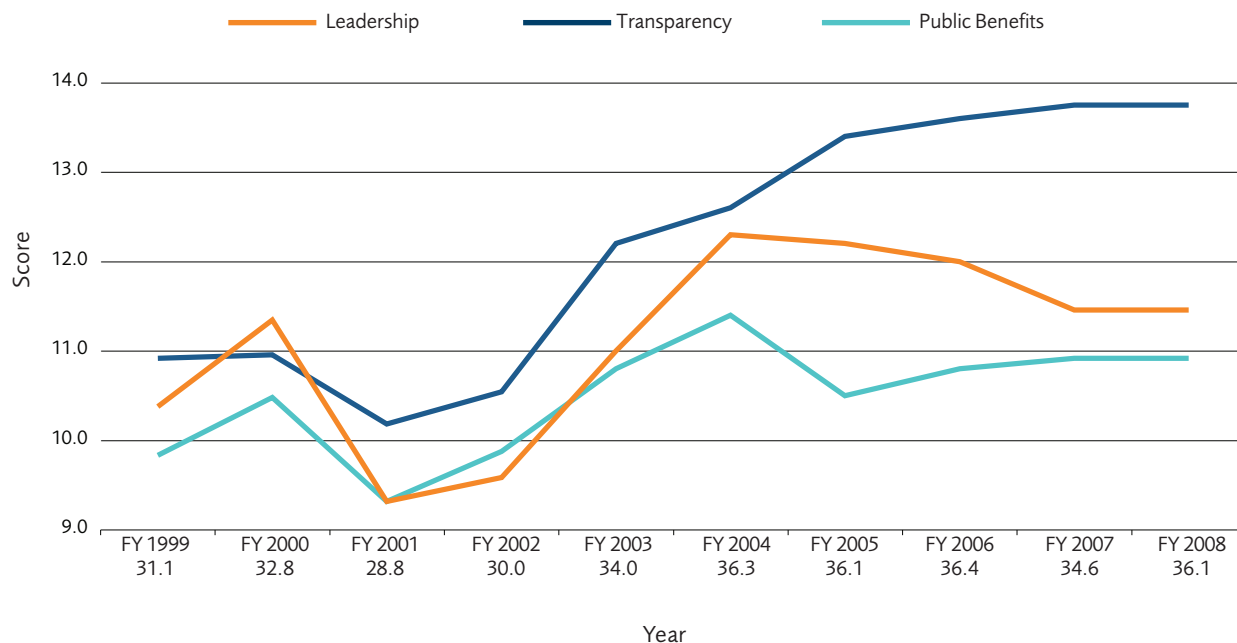
There were no dramatic individual score decreases. HUD lost the most ground (-4 points), dropping in four categories. Several agencies dropped 1 or 2 points.

CHURN IN THE RANKINGS

BELOW THE TOP five reports, fiscal year 2008 saw significant churn in the rankings. This occurred because one group of agencies made measurable improvements while others scored about the same as in fiscal year 2007. State, Energy, Interior, and Education leapt by 8 to 12 places in the rankings, and USAID rose by 4 places. Because these reports improved their scores, the others plummeted by between 4 and 9 spots in the rankings.

AVERAGE SCORES RISE

THE AVERAGE SCORE for all 24 agencies was 36.1. As figure 3 shows, this is an increase from last year’s average (34.6) and a slight decrease from fiscal year 2006 (36.4). The increase from fiscal year 2007 is attributable primarily to improvement in the pilot agency scores. The average score for traditional performance and accountability reports increased 1 point, from 37.3 last year to 38.3 this year. The average pilot agency score increased 2.6 points, from 30.0 last year to 32.6 this year.

Figure 3: A Year of Recovery

Report quality clusters in the middle

ALMOST HALF OF the report scores cluster around the satisfactory score of 36. Four reports earned 37 points, four earned 36 points, one earned 35 points, and one earned 34 points. As figure 4 shows, this led to a noticeable increase in the number of reports scoring satisfactory or better in fiscal year 2008—reversing the past several years’ trend. About half of the reports (13) achieved a satisfactory score for fiscal year 2008, up from just 7 reports in fiscal year 2007. Ten reports scored between 24 and 35 points, down from 15 in fiscal year 2007. Only one report scored below 24 in fiscal year 2008.

In general, transparency has improved at a faster rate and achieved a higher score than public benefits or forward-looking leadership. This probably indicates that it is easier to improve on factors like availability and readability than on substantive matters like definition of outcomes or analysis of performance results. Indeed, even many of the low-ranking agencies earned scores of 4 or 5 on some of the transparency criteria.

Opportunities for progress

There is still substantial room for progress. Table 4 reveals that average scores for eight criteria are below 3.

For most criteria, average scores increased in fiscal year 2008. Thanks largely to these improvements, below-satisfactory scores on many criteria are nevertheless close to 3. These include verification and validation (2.92 points), baseline and trend data (2.88 points), outcome-oriented measures (2.88 points), explanation of management challenges (2.92 points), and improvement plans (2.83 points).

Figure 4: More Reports Are Satisfactory

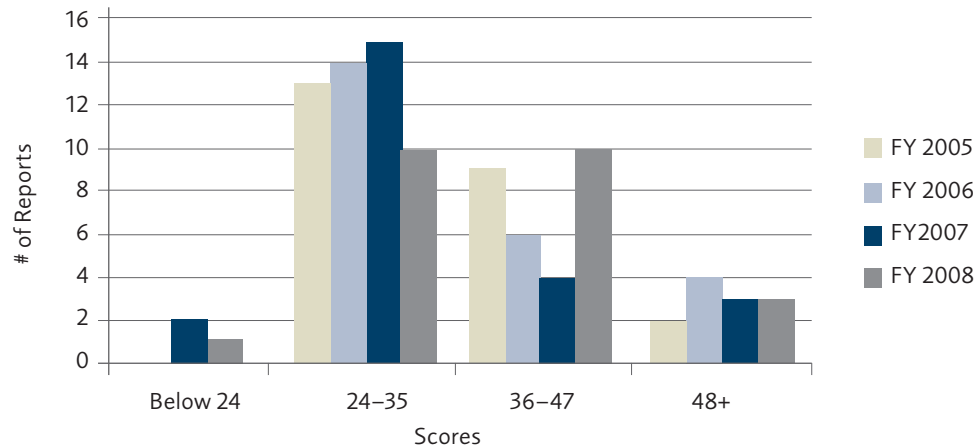


Table 4: Eight Criteria Below Satisfactory

CRITERION	FISCAL YEAR 2007	FISCAL YEAR 2008	CHANGE 2007-08	% CHANGE 2007-08
1. Accessibility	4.08	4.38	0.29	7
2. Readability	3.63	3.58	-0.04	-1
3. Verification & Validation	2.46	2.92	0.46	19
4. Baseline/Trend Data	2.67	2.88	0.21	8
5. Outcome Goals	3.13	3.21	0.08	3
6. Outcome Measures	2.67	2.88	0.21	8
7. Agency-Affected Outcomes	2.67	2.63	-0.04	-2
8. Linkage to Costs	2.13	2.21	0.08	4
9. Vision	3.00	3.08	0.08	3
10. Explain Failures	2.67	2.63	-0.04	-2
11. Management Challenges	2.83	2.92	0.08	3
12. Improvement Plans	2.67	2.83	0.17	6

Scores are rounded

YOU'VE COME A LONG WAY, BABY!

FEDERAL PERFORMANCE REPORTING has improved substantially since fiscal year 1999—much more than the 15-percentage-point increase in average score over the past ten years would indicate. Table 5 shows the change in each report’s score between fiscal year 1999 and fiscal year 2008. Large drops in scores for three reports—USAID, Defense, and SBA—help pull down the average score increase. For the 17 reports whose scores improved, the average increase was 8.94 points, almost double the average increase of 4.84 points for all 24 reports. Nine reports achieved double-digit increases in their scores.

All of these score data understate the full extent of improvement because the research team tightens the scoring criteria over time as new best practices emerge. To gain a better understanding of how the quality of disclosure has improved, we performed both a quantitative and a qualitative evaluation.

Table 5

Fiscal Year 2008 Scores & Rankings: Comparison to Fiscal Year 1999						
Highest Rank = 1; Lowest = 24. Maximum Possible Score = 60; Minimum = 12.						
	Fiscal Year 2008		Fiscal Year 1999		CHANGE IN SCORE	CHANGE IN RANKING
	TOTAL SCORE	RANK	TOTAL SCORE	RANK		
Labor	56	1	36	5	+20	+4
Veterans Affairs	54	2	48	3	+6	+1
Transportation	53	3	51	2	+2	-1
Homeland Security*	40	4	27	22	+13	+18
NRC	40	4	25	17	+15	+13
Education	37	6	37	4	0	-2
Interior	37	6	31	11	+6	+5
State	37	6	25	17	+12	+11
Treasury	37	6	36	5	+1	-1
Energy	36	10	27	14	+9	+4
EPA	36	10	31	11	+5	+1
Health & Human Services	36	10	24	20	+12	+10
USAID	36	10	52	1	-16	-9
Commerce	35	14	22	22	+13	+8
Justice	34	15	23	21	+11	+6
Agriculture	33	16	22	22	+11	+6
GSA	32	17	32	9	0	-8
NSF	32	17	21	24	+11	+7
Social Security	32	17	33	8	-1	-9
NASA	31	20	27	14	+4	-6
OPM	28	21	27	14	+1	-7
HUD	27	22	28	13	-1	-9
Defense	26	23	34	7	-8	-16
SBA	22	24	32	9	-10	-15
Average	36.13		31.29		4.83	
Median	36.00		29.50		6.50	

*Since DHS did not exist in 1999, the chart shows its score and rank from fiscal year 2003, the first year its report was included in the Scorecard.

BY THE NUMBERS

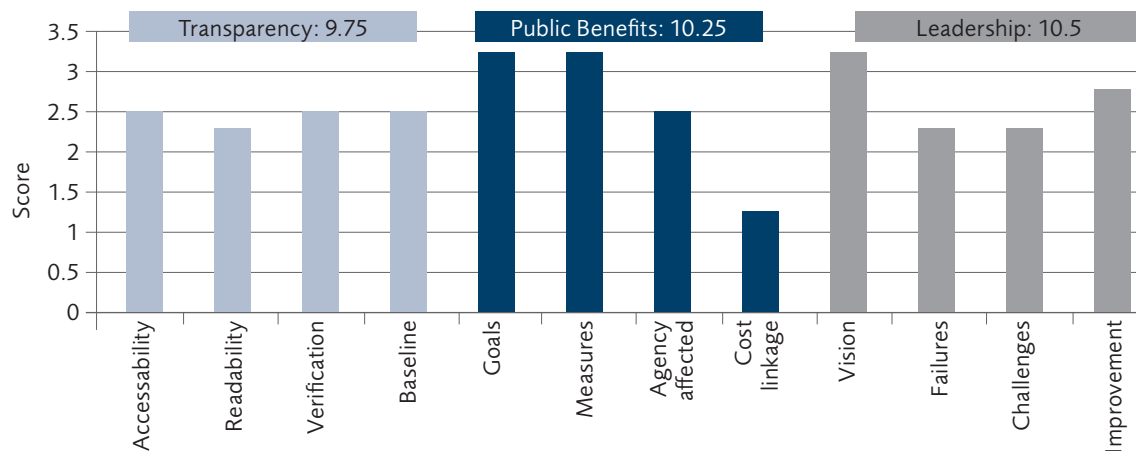
FOR ITS QUANTITATIVE evaluation, the research team re-examined the top four reports from fiscal year 1999 using the same standards that applied to the fiscal year 2008 reports.¹⁰ Comparing the scores the same report receives under two different years' standards should reveal how much the standards have tightened over ten years. Comparing reports the same agency produced in two different years, both evaluated under the fiscal year 2008 standards, should reveal how much that agency's report actually improved.

Table 6 shows the results. Evaluated by fiscal year 2008 standards, the best fiscal year 1999 report (from USAID) would have ranked 16th in fiscal year 2008, with just 33 points out of a possible 60. Transportation and Veterans Affairs received 51 and 48 points respectively in fiscal year 1999, ranking second and third. Evaluated under fiscal year 2008 standards, they both received 31 points and would have tied for 20th place in fiscal year 2008. Education, the fourth-place report in fiscal year 1999, received 37 points. Evaluated by fiscal year 2008 standards, it received 27 points and would have ranked 22nd in fiscal year 2008.

We can measure how these specific agencies' reports have improved by comparing the scores for their fiscal year 2008 reports with the scores on their fiscal year 1999 reports evaluated on the same 12 criteria under fiscal year 2008 standards. Using the scores reported in table 6, USAID's report improved by about 9 percent over ten years (from 33 to 36 points), Education's report improved by 37 percent (from 27 to 37 points), Transportation's report improved by 71 percent (from 31 to 53 points), and Veterans Affairs' report improved by 74 percent (from 31 to 54 points).

Figure 5 shows how these four reports scored on individual criteria. Note that average scores on criteria 5 (outcome-oriented goals), 6 (outcome-oriented measures), and 9 (vision) all exceeded the satisfactory score of 3—even when evaluated by fiscal year 2008 standards. This suggests that in the early days of GPRA, the higher-ranking agencies got off to a good start in formulating outcome-oriented goals and measures. Also, the average score on criterion 8 (linkage of results to costs) barely exceeds 1. This indicates that, compared to current practice, even the top scorers had little cost-related content in fiscal year 1999. Third, the readability score is also quite low—and the research team agreed that readability has improved substantially since fiscal year 1999.

Figure 5: Average Scores for Re-evaluated FY 1999 Reports



10. This was considered easier than transporting some fiscal year 2008 reports back in time so the original research team could evaluate them in 1999.

Table 6

Top Four 1999 Reports Would Rank Low Today					
Highest Rank = 1; Lowest = 24. Maximum Possible Score = 60; Minimum = 12					
	Fiscal Year 2008				
	Transparency	Public Benefits	Leadership	Total	Rank
Labor	20	19	17	56	1
Veterans Affairs	19	16	19	54	2
Transportation	16	20	17	53	3
Homeland Security	15	13	12	40	4
NRC	15	13	12	40	4
Education	14	12	11	37	6
Interior	16	10	11	37	6
State	15	10	12	37	6
Treasury	14	10	13	37	6
Energy	13	11	12	36	10
EPA	13	11	12	36	10
Health & Human Services	13	13	10	36	10
USAID	15	10	11	36	10
Commerce	15	10	10	35	14
Justice	15	8	11	34	15
Agriculture	12	10	11	33	16
USAID 1999	9	11	13	33	16
GSA	11	12	9	32	17
NSF	15	7	10	32	17
Social Security	12	8	12	32	17
NASA	11	8	12	31	20
Transportation 1999	9	12	10	31	20
Veterans Affairs 1999	11	10	10	31	20
OPM	11	8	9	28	21
HUD	11	8	8	27	22
Education 1999	10	8	9	27	22
Defense	11	7	8	26	23
SBA	8	8	6	22	24
Fiscal Year 2008 Average	13.8	10.9	11.5	36.1	
Fiscal Year 2008 Median	14.0	10.0	11.0	36.0	

On average, we estimate that report quality may have improved by about 75 percent. Scores for the four re-evaluated fiscal year 1999 reports averaged one-third lower under the fiscal year 2008 standards than under the fiscal year 1999 standards. If we assume that using the fiscal year 2008 scoring standards would have reduced all fiscal year 1999 scores by one-third, the average fiscal year 1999 score using fiscal year 2008 standards would have been 20.65 instead of the 31.29 shown in table 5. An increase from 20.65 to the average fiscal year 2008 score of 36.13 implies that the average quality of performance reports improved by at least 75 percent.¹¹ Of course, that means the quality of some agencies' reports increased by even more than that amount, others by less.

Qualitative evaluation

SINCE THE MERCATUS Scorecard explicitly identifies best practices every year, it is not difficult to compare some fiscal year 1999 best practices with their fiscal year 2008 counterparts. This comparison reveals how the “state of the art” in the best reports has advanced during the past decade. Except for criterion 1 (accessibility), the best practices in the list below are not usually typical practices. In most cases, only a few reports in each year used the best practices. The list below provides a summary; more detailed descriptions can be found in the fiscal year 1999 Scorecard (<http://www.mercatus.org/PublicationDetails.aspx?id=21144>) and the “Strongest and Weakest Scores” section of this report.

Criterion 1: Accessibility

Fiscal Year 1999—Labor and OPM had their reports available online with obvious links from their home pages. Sixteen agencies made their reports available online, but only four earned a score of 5 for clearly labeling the report as the annual performance report and making it easy to find.

Fiscal Year 2008—Thirteen agencies earned a score of 5 because they posted their reports on time, created a direct link on the home page, permitted downloads as both single and multiple files, and provided contact information for questions or comments.

Criterion 2: Readability

Fiscal Year 1999—USAID produced a lengthy report with a massive amount of detail, but the organization, headings, sidebars, tables, and charts made the report easy to scan and digest. Relative lack of jargon made the report understandable to readers who are not specialists in foreign aid.

Fiscal Year 2008—Justice and Labor are the only reports to score a 5. The Management’s Discussion and Analysis section and the Performance section of Justice’s report are much shorter than in most PARs. Justice’s report includes 25 key measures and presents goals, measures, and results in convenient tables. Labor’s report is visually appealing with helpful tables and graphics. The secretary’s letter describes key achievements over time as well as fiscal year 2008 results for each strategic goal. Both reports are complemented by excellent citizens’ reports that include the two-page snapshot, concise summaries of results, and reader-friendly links to additional information.

11. If the average fiscal year 1999 score using fiscal year 2008 scoring standards would have been 20.65, the percentage improvement is calculated by subtracting this score from the actual fiscal year 2008 score (36.13), then dividing this difference by 20.65.

Criterion 3: Verification and validation

Fiscal Year 1999—Education’s report presented a thorough discussion of validation procedures and data limitations. The report acknowledged data deficiencies and outlined plans to develop new databases, improve data quality, and ensure the accuracy of financial statements.

Fiscal Year 2008—Labor’s report assesses the quality of its data for each performance goal and provides data sources for each measure. Tables in the Veterans Affairs report provide data definitions, verification and validation information, and data limitations.

Criterion 4: Baseline and trend data

Fiscal Year 1999—Transportation’s report presented trend data for as far back as five or ten years. Defense juxtaposed three years of actual data with the goals for each program for the next three years. Veterans Affairs, Education, and USAID also provided useful baseline data.

Fiscal Year 2008—Labor and Veterans Affairs both present four years’ worth of baseline data for key measures. Data include targets, actual results, and (in Labor’s case) costs. Labor and Treasury use graphics to show performance trends. Veterans Affairs shows long-range targets for each measure, and Transportation presents a fiscal year 2009 forecast for each measure.

Criterion 5: Outcome-oriented goals

Fiscal Year 1999—All five of Transportation’s strategic goals were outcome oriented. Most of Veterans Affairs’ and USAID’s were, too.

Fiscal Year 2008—The vast majority of Labor’s and Transportation’s goals and objectives are clear, measurable outcomes that the general public can appreciate and understand, such as safer workplaces, employment of people who were in job training programs, reduced highway congestion, and reduced transportation accidents and fatalities.

Criterion 6: Outcome measures

Fiscal Year 1999—Most measures in the Veterans Affairs and Transportation reports were outcomes or closely related to outcomes.

Fiscal Year 2008—Labor’s measures for three of its four strategic goals are outcomes; the fourth has a mix of outcome, activity, output, and efficiency measures. Transportation’s measures for three of five strategic goals are all outcomes; for two other goals, about half the measures are end or intermediate outcomes.

Criterion 7: Agency-affected outcomes

Fiscal Year 1999—USAID explained how its activities contributed to results in specific cases, and also acknowledged other factors that may have contributed to the results.

Fiscal Year 2008—Four agencies consistently describe how their activities led to observed results on highly outcome-oriented performance metrics.

Criterion 8: Results linked to costs

Fiscal Year 1999—The Veterans Affairs report included charts that juxtaposed personnel and costs by program area with performance measures for each program area.

Fiscal Year 2008—Labor and Transportation both assign costs to strategic goals and most individual performance measures. Labor’s report contains this information for several years.

Criterion 9: Vision

Fiscal Year 1999—Transportation made a case for how it affects Americans’ quality of life by writing each strategic goal so that it identifies a broad result of interest to citizens and then explains the means by which the department hopes to accomplish the result.

Fiscal Year 2008—Narratives in both the Transportation and Veterans Affairs reports cite specific, major accomplishments that affect citizens’ quality of life. Narratives and vignettes are linked to outcome-oriented performance metrics, and performance metrics demonstrate that the narratives describe general results rather than isolated, anecdotal successes.

Criterion 10: Explanation of failures

Fiscal Year 1999—USAID’s report provided explanations when targets were not met and described planned remedies for unmet targets.

Fiscal Year 2008—Veterans Affairs’ report includes a “Performance Shortfall Analysis” that describes reasons for shortfalls and outlines steps and a timeline to remedy them. The report even offers performance improvement steps in cases where targets were met. NASA, Transportation, and Treasury also identify shortfalls on performance measures, explain the shortfalls, and outline plans to improve.

Criterion 11: Major management challenges

Fiscal Year 1999—Transportation, USAID, and Veterans Affairs all include thorough discussions of management challenges identified by the agency inspector general and the GAO.

Fiscal Year 2008—Best practices include an insightful inspector general report that lists major management challenges and assesses the agency’s progress toward resolving them, an agency self-assessment of progress, a description of actions taken, a timeline for resolving each challenge, and an explanation of how each challenge is related to the agency’s strategic goals. No agency did all of these things, but various combinations can be found in the reports from Labor, EPA, Transportation, and Veterans Affairs. The best example of an insightful inspector general report is the one for SBA. Unfortunately, the SBA citizens’ report does not reference this very useful document, which appears in the agency’s financial report.

Criterion 12: Improvement plans

Fiscal Year 1999—Discussion of each measure in Transportation’s report included fiscal year 1999 results, projected performance in fiscal year 2000, and actions planned to achieve goals in fiscal year 2001.

Fiscal Year 2008—Reports from Labor, Veterans Affairs, and Transportation present improvement strategies for performance shortfalls and major management challenges. They also describe broader challenges the agencies expect to face in the future and offer plans for addressing them.

This qualitative description of best practices is consistent with the quantitative score assessment: Performance reporting has made significant progress since fiscal year 1999.

UPDATE ON THE PILOT FORMAT

FOR THE SECOND year, the OMB gave agencies the option of submitting a “traditional” performance and accountability report (PAR) or participating in the reporting pilot. The PAR format combines the agency’s performance report and its financial report into a single document. The pilot format features three separate documents: a performance report, a financial report, and a citizens’ report. (Last year the citizens’ report was referred to as the “highlights” document.)

Mercatus researchers have followed the progress of the pilot format since its inception. The May 2007 OMB memo outlining the pilot encouraged agencies to consult previous editions of this Scorecard as well as the Association of Government Accountants’ guidelines for its Certificate of Excellence in Accountability Reporting.¹² We presented a preliminary analysis of the pilot format’s effectiveness in fiscal year 2007 at an open forum OMB sponsored at the National Academy of Public Administration in April 2008. Mercatus researchers also participated in an assessment of the first year of the pilot format.¹³

The Scorecard scoring process provides one way of evaluating the results of the pilot format. As was the case last year, 15 of the 24 agencies submitted PARs for fiscal year 2008, and the other 9 participated in the pilot. The breakdown of PAR versus pilot agencies was the same this year with two exceptions: OPM switched from the PAR to the pilot, and NASA made the opposite switch, submitting a PAR this year. All but 4 of the 15 PAR agencies voluntarily submitted citizens’ reports to complement their PARs, based on what our reviewers found at the time of their evaluations.¹⁴

PILOT SCORES IMPROVED

In fiscal year 2007, many agencies adopting the pilot format saw their scores decline, while average scores for agencies producing traditional performance and accountability reports remained largely unchanged. Pilot agencies reversed most of this decline in fiscal year 2008.

The average pilot agency score increased 2.6 points (9 percent) this year. The improvement was widespread. Six of the seven agencies that participated in the pilot for both years increased their score this year. The exception was HHS, which dropped just 1 point. It is also noteworthy that five of the nine pilot agencies scored average (36 points) or better this year. Only two pilot agencies accomplished this for fiscal year 2007. Tables 7, 8, and 9 document these developments in greater detail.

Table 7 shows that, just as in fiscal year 2007, traditional performance and accountability reports achieved higher average scores than reports produced in the pilot format. The difference narrowed, however, from 24 percent in fiscal year 2007 to 18 percent in fiscal year 2008. Moreover, this scoring difference may reflect factors other than the pilot format. Agencies that used the pilot format, on average, also scored lower in fiscal year 2006, before the pilot format was adopted. These agencies may have scored lower in fiscal year 2008 for reasons unrelated to the pilot format.

12. Memo from Clay Johnson, deputy director for management, “Fiscal Year 2007 Pilot Program for Alternative Approaches to Performance and Accountability Reporting” (May 17, 2007), http://www.whitehouse.gov/omb/assets/omb/memoranda/fy2007/par_alternative.pdf.

13. Valerie Richardson, *Increasing Transparency and Accountability in Federal Performance Reporting: Lessons from the OMB Pilot Program*. (Washington, DC: IBM Center for the Business of Government, 2009), http://www.businessofgovernment.org/publications/grant_reports/details/index.asp?gid=337.

14. Our reviewers could not locate citizens’ reports for Agriculture, NASA, Transportation, or SSA.

Table 7: Traditional Format Outscores Pilot Format

Criterion	1	2	3	4	Transparency
PAR 2008 Average	4.33	3.67	3.13	3.13	14.27
Pilot 2008 Average	4.44	3.44	2.56	2.44	12.89
Difference	-0.11	0.22	0.58	0.69	1.38
% Difference	-3	6	23	28	11
Criterion	5	6	7	8	Public Benefits
PAR 2008 Average	3.27	3.13	2.80	2.47	11.67
Pilot 2008 Average	3.11	2.44	2.33	1.78	9.67
Difference	0.16	0.69	0.47	0.69	2.00
% Difference	5	28	20	39	21
Criterion	9	10	11	12	Leadership
PAR 2008 Average	3.33	2.87	3.00	3.13	12.33
Pilot 2008 Average	2.67	2.22	2.78	2.33	10.00
Difference	0.67	0.64	0.22	0.80	2.33
% Difference	25	29	8	34	23
					TOTAL SCORE
PAR 2008 Average					38.27
Pilot 2008 Average					32.56
Difference					5.71
% Difference					18

Table 8 controls for this difference by calculating the average change in scores for agencies using two different reporting formats in fiscal year 2006 and fiscal year 2008. The average total score for agencies producing traditional performance and accountability reports was virtually identical in both years. Scores on some individual criteria rose or fell, but scores for the three categories of criteria held fairly steady. Reports using the pilot format in fiscal year 2008 had scores averaging 1 point lower than those same agencies' traditional reports received in fiscal year 2006.

Of greatest concern is the 20 percent drop in scores on criterion 8 (linkage of results to costs). This likely happened due to the tightening of the research team's standards on this question between fiscal year 2006 and fiscal year 2008, and not for any reason unique to the pilot format.

Scores on criterion 12 (discussion of plans for improvement) also dropped by 16 percent between fiscal year 2006 and fiscal year 2008. However, just three reports account for this drop: State and Defense each fell 1 point, and SBA lost 2 points because no useful content could be found for fiscal year 2008. All the other pilot agencies achieved the same score on this criterion in both years.

Table 8: Pilot Format Closing the Gap

Criterion	1	2	3	4	Transparency
PAR 2006–08 change	-0.20	0.07	0.47	-0.13	0.20
% change	-4	2	18	-4	1
Pilot 2006–08 change	0.33	0.00	-0.11	-0.22	0.00
% change	8	0	-4	-8	0
Criterion	5	6	7	8	Public Benefits
PAR 2006–08 change	0.13	0.27	-0.07	0.13	0.47
% change	4	9	-2	6	4
Pilot 2006–08 change	0.11	0.11	-0.22	-0.44	-0.44
% change	4	5	-9	-20	-4
Criterion	9	10	11	12	Leadership
PAR 2006–08 change	-0.13	-0.13	0.07	-0.20	-0.40
% change	-4	-4	2	-6	-3
Pilot 2006–08 change	-0.22	-0.22	0.22	-0.44	-0.67
% change	-8	-9	9	-16	-6
					TOTAL SCORE
PAR 2006–08 change					0.27
% change					1
Pilot 2006–08 change					-1.11
% change					-3

Despite these small shortcomings, pilot agencies in fiscal year 2008 regained most of the ground they lost in fiscal year 2007. Fiscal year 2007 pilot agency scores averaged 4.11 points lower than the same agencies’ scores in fiscal year 2006; fiscal year 2008 pilot scores averaged about 1 point lower. Table 9 documents the fiscal year 2008 turnaround.

Scores improved on 11 out of 12 criteria. Huge increases occurred on two criteria that saw huge drops in fiscal year 2007: criterion 3 (verification and validation of data) and criterion 4 (baseline and trend data). The much improved pilot scores for criteria 3 and 4 resulted mainly from the research team’s ability to access and use more source documents this year—particularly the free-standing performance reports, which usually contain the details on data quality and prior-year results.

Table 9: Pilot Reports Improve Substantially in Fiscal Year 2008

Criterion	1	2	3	4	Transparency
PAR 2007–08 change	0.20	-0.13	0.20	-0.07	0.20
% change	5	-4	7	-2	1
Pilot 2007–08 change	0.44	0.11	0.89	0.67	2.11
% change	11	3	53	38	20
Criterion	5	6	7	8	Public Benefits
PAR 2007–08 change	0.13	0.27	-0.07	0.07	0.40
% change	4	9	-2	3	4
Pilot 2007–08 change	0.00	0.11	0.00	0.11	0.22
% change	0	5	0	7	2
Criterion	9	10	11	12	Leadership
PAR 2007–08 change	0.13	0.00	0.00	0.20	0.33
% change	4	0	0	7	3
Pilot 2007–08 change	0.00	-0.11	0.22	0.11	0.22
% change	0	-5	9	5	2
					TOTAL SCORE
PAR 2007–08 change					0.93
% change					2.00
Pilot 2007–08 change					2.56
% change					9

Last year’s Scorecard concluded, “At least for FY 2007, the pilot approach to performance reporting seriously impaired the accessibility and transparency of performance information for the general public and detracted from its value to the public.”¹⁵ Given another year of results, it appears that most of the fiscal year 2007 drop in pilot agencies’ scores reflects the inevitable difficulties of adjusting to a new reporting format rather than a problem inherent in the pilot format.

REASONS FOR PILOT SCORES’ IMPROVEMENT

1) Information was much more accessible

THE MAIN REASON for the improvement in pilot scores this year is that our reviewers found it much easier to identify, access, and use the three source documents—the citizens’ report, the fiscal year 2008 performance report, and the fiscal year 2008 financial report. Last year, we were able to find and access all three source documents for only two of the nine pilot agencies. This year, we found and accessed all three documents for seven of the nine pilots.

Although SBA was late in posting its report, we found all nine citizens’ reports online. Furthermore, unlike last year, most of the pilot agency citizens’ reports clearly explained that the financial and performance reports

15. Maurice McTigue, Henry Wray, and Jerry Ellig, *9th Annual Performance Report Scorecard: Which Federal Agencies Best Inform the Public?* (Arlington, VA: Mercatus Center at George Mason University, May 2008), 26, <http://www.mercatus.org/PublicationDetails.aspx?id=16102>.

were part of their performance reporting and provided links to those documents. The foreword page of the Energy report offers a good example of clearly explaining the different pilot documents and leading the reader to them. Unlike last year, none of the agencies buried their performance reports in their congressional budget justifications—a practice that made the reports virtually useless to the general public. Instead, all but one pilot agency issued a separate, free-standing performance report this year. The exception was HHS. While we found separate performance reports for HHS component organizations, we could not locate a department-wide fiscal year 2008 performance report.

There were still some glitches, however. While Defense and SBA apparently produced fiscal year 2008 performance reports, these reports were not posted online at the time of our reviews so we could not consider them. Also, even when all three documents could be found, the usefulness of the links between them and the ease with which they could be opened and navigated varied considerably among the pilot agencies.

2) "Snapshots" improved readability

ANOTHER FEATURE THAT generally improved the pilot reports this year was inclusion of the two-page snapshots in their citizens' reports. Each snapshot follows a consistent format. The first page of each snapshot presents background information on its organization, functions, and budget as well as a brief summary of its performance accomplishments, management challenges, and the results of its financial statement audit. The second page lists the agency's strategic goals. For each goal, it shows the budget resources for the goal, aggregate results for all of its performance measures, and one or more representative performance measures. The snapshots had the benefit of providing certain basic background, financial, and performance information in a concise and consistent format. However, the quality of the performance-related snapshot content varied greatly. Not surprisingly, the content was far better for those agencies that had stronger performance metrics and produced better overall reports.

We of course considered the snapshots as part of the content of the citizens' reports that contained them. The snapshots per se were not usually a significant factor in our scoring since their quality varied in relation to the quality of the underlying reports. In the case of the Defense and SBA reports, however, the snapshots provided virtually the only content relevant to some of our criteria and did affect our scores to some extent. For example, the SBA snapshot had the only content on major management challenges that a lay reader could reasonably find, limited though it was. (The full inspector general presentation on SBA management challenges was contained in the agency's fiscal year 2008 financial report, but the citizens' report made no reference to it.)

ENHANCEMENTS TO CITIZENS' REPORTS NEEDED

WHILE MOST PILOT agency reports are better this year, there still is much room for improvement to make the reports more user-friendly. The pilot format forces the reader to cross-reference and navigate among three separate documents. This can be challenging and distracting. The financial and performance reports tend to be long, and many can be opened online only as large, single files. Thus, it is important that the citizens' report be more than a "bare bones" document that serves mainly as a reference point to other sources for substantive information. Rather, the citizens' report should have value as a self-contained document that provides a substantive and comprehensive overview of the agency's performance. When the citizens' report does refer to other sources, it should provide specific, user-friendly links to the relevant information in those documents. The quality of the links in the citizens' reports to other source documents is uneven.

Access to inspector general presentations on major management challenges—an important piece in performance and accountability reporting—illustrates this point. All of the pilot agencies included full inspector general presentations as part of their fiscal year 2008 financial reports. The ease with which a reader could identify and locate this presentation from the citizens' report varied considerably. Some citizens' reports (State, Homeland Security, OPM, and USAID) highlighted the availability of the inspector general presentation in the financial report and cited the specific page where it was located. The HHS citizens' report provided a link to the financial report for the inspector general presentation, but the link did not lead directly to the presentation or clearly indicate where to find it in the financial report. The worst example was SBA. Its snapshot lists some challenges the agency faces, but the citizens' report made no mention of the inspector general presentation on major management challenges in the financial report.

The citizens' report is mandatory for pilot agencies but optional for PAR agencies. This year, 11 of the 15 PAR agencies voluntarily issued citizens' reports to complement their full reports. That is an increase over last year, when 8 of the PAR agencies did so. Once again this year, the best citizens' reports came from the PAR agencies. We raised our "readability" (criterion 2) scores for 7 of the PAR agencies based on their voluntarily produced citizens' reports. Two PAR agency citizens' reports stood out as models for other agencies: Labor and Justice. (They are described in the "Strongest and Weakest Scores" section under criterion 2.)

PARS VS. PILOTS: CONCLUSIONS

TWO YEARS OF experience with the pilot reveals that it is a workable format that can produce good reports. While there remains much room for improvement, this year's round of pilot reports addressed many of the shortcomings we identified last year. This still leaves the issue of which format is better: the PAR or the pilot? Or are there better alternatives to either of these formats?

In terms of our scores, the PAR reports still fare better. The average score once more this year is significantly higher for PARs than pilot reports. Seven of the nine reports scoring above satisfactory (36 points) are PARs. The highest scoring format is a PAR accompanied by a citizens' report. The top two reports as well as six of the nine reports scoring above average used this format. However, the pilot reports narrowed the gap this year. Since most agencies participating in the pilot tended to score lower over the years than the non-pilot agencies, the differences in PAR and pilot agency scores can be attributed at least in part to factors other than the reporting format.

The intrinsic advantage of the PAR format is that it combines all performance, management, and financial information into a single document. The intrinsic disadvantage is the length of the resulting document. The length of the 15 PAR reports we reviewed this year ranged from an oppressive 567 pages (EPA) to a low of 172 pages (NRC). The average length for all PARs was 318 pages. Of course, coupling the PAR with a (well done) citizens' report can compensate for its length. Also, there are ways to reduce the size of the PARs substantially. (We offer some specific suggestions in the "Strongest and Weakest Scores" section under criterion 2.) One more fundamental alternative would be to discontinue the PAR format and revert to the pre-PAR practice of separate performance and financial reporting for all agencies, not just the pilots. Dropping the detailed financial data from the performance document, for example, would eliminate about 100 pages (on average) of content that holds little interest for the general public. Transparency would not suffer as long as the report contains direct links to the complete financial data for readers who want more detail.

The main potential advantages of the pilot format are brevity for a lay audience and flexibility for all audiences.

The premise is that the mandatory citizens' report provides the general public with a concise overview of an agency's performance, which is probably all most members of the general public need or want. In theory, the pilot format also affords flexibility so that readers with more in-depth or specialized interests can access other source documents for the additional information they need. Of course, these potential advantages depend on the citizens' report being a substantive stand-alone document that has user-friendly links to other source information.

In the final analysis, however, any report's value to the public depends much more on the quality of its content (understandable, results-oriented performance metrics; good organization; concise and insightful narratives; and so forth) than on its format. Our reviews have identified both high- and low-quality reports in both PAR and pilot formats. We have also highlighted a number of ways in which the presentational features and substantive content of both types of reports can be enhanced to better serve the general public.

ASSESSING THE "SNAPSHOTS"

IN A NEW feature this year, the OMB required all agencies to prepare a two-page "snapshot" document, which provided a high-level overview of each agency's mission as well as its performance and financial results. In early January 2009, the OMB issued *The FY 2008 Performance Report of the Federal Government*, which compiled the snapshots for the 24 agencies covered by the Chief Financial Officers Act. Also, most agencies that produced citizens' reports incorporated the snapshots, in whole or in part, into their citizens' reports. We evaluated an agency's snapshot if it was included in the citizens' report.

Agency snapshots—which use a consistent format—are very helpful. The background information presented on the first page of each agency snapshot is concise and informative. A particularly useful feature is a listing of the top five agency programs or activities by budget cost. Another interesting feature many agencies included is a dollar figure showing the cost of the agency's annual budget for each person in the United States.

The snapshots also contain descriptions of major management challenges, which vary from agency to agency. Some are specific and informative. While all are necessarily brief, some provide links to other sources for more detailed information. Several focus on one particularly significant challenge. Most, however, simply list a number of challenges in a way that is too general to convey much useful insight. The descriptions of management challenges should all link to more detailed sources of information. It would also be useful to include an independent assessment by the inspector general of the agency's progress in addressing major management challenges. Another approach might be to include, where applicable, agency-specific "high-risk" problems listed by the GAO and how long each problem has been on the high-risk list.

The usefulness of the performance information on the second page of each snapshot varies greatly. This is largely a function of whether and to what extent each individual agency's underlying performance metrics are clearly stated and outcome oriented. Some performance pages have a "cut-and-paste" quality as a result of simply lifting information from the agency performance report. For example, the Commerce performance page is laden with undefined acronyms. A more widespread problem is that the representative performance measures for some agencies are too technical or arcane to be understood by the general public. In addition, some agency measures set performance targets significantly below actual performance levels for the past three years.

Finally, the inclusion of aggregate performance results, such as percentage of measures met or missed for each goal, is not very meaningful given the wide variation in the quality of agency performance metrics. At worst, it can be misleading. An agency that meets a high proportion of weak or non-challenging measures may appear to have better results than an agency with a lower success rate that uses more challenging outcome-oriented measures. It might be more informative to omit aggregate results and include more individual representative measures. In this regard, the snapshots might include at least one representative measure for each of the agencies' five costliest programs. Since the snapshot is aimed at the general public, care should be taken to ensure that the representative measures it describes are ones that a lay audience can understand.

For those agencies that produced citizens' reports this year, the snapshot usually appears at the outset of the citizens' report.¹⁶ Incorporating the snapshot into the citizens' report ameliorates some of the issues mentioned above that result from viewing each snapshot in isolation.

More detailed comments on each agency's snapshot are available on the Mercatus Center Web site at <http://www.mercatus.org/ResearchAreaLanding.aspx?Category=74>.

16. All agencies participating in the performance reporting pilot this year included the full snapshot in their mandatory citizens' reports. Non-pilot agencies that voluntarily published citizens' reports usually included at least some content from their snapshot. Some incorporated the full snapshot.

DO SOME AGENCIES HAVE INHERENT ADVANTAGES?

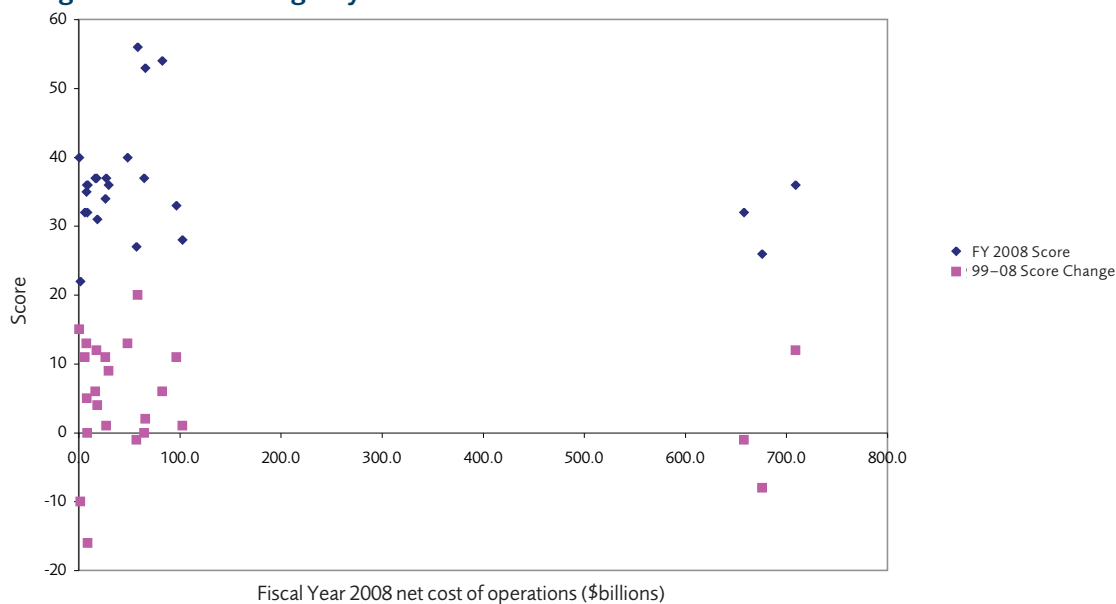
DIFFERENT FEDERAL AGENCIES of different sizes have different missions and utilize different means to achieve them. Do these differences make it easier for some agencies to identify and measure outcomes, report on results, and hence score well on our Scorecard?

Federal agencies that provide direct services to the public, such as Veterans Affairs or Social Security, might find it easier to define and gather data on outcomes than agencies that mainly make grants to states or other third parties, such as Education or Agriculture.¹⁷ Larger agencies might produce better reports because they can afford to put more resources into performance measurement, or they might do worse because they have more programs and are inherently harder to manage due to their size. Ideology could also play a role. Critics suggest that GPRA has been used to advance a variety of ideological and partisan agendas and that performance measurement tilts attention toward “efficiency values” at the expense of “equity values.”¹⁸ If either GPRA itself or the evaluation process in our Scorecard systematically disadvantages some agencies based on the ideological orientation of their missions or policies, we would certainly want to know.

The following figures test for agencies’ inherent advantages or disadvantages by comparing agency scores with various institutional characteristics that could potentially influence scores. Two sets of Scorecard data are used: the agencies’ fiscal year 2008 scores and the change in agency scores between fiscal year 1999 and fiscal year 2008. The former provides the most recent snapshot of report quality; the latter measures improvement during the past ten years.

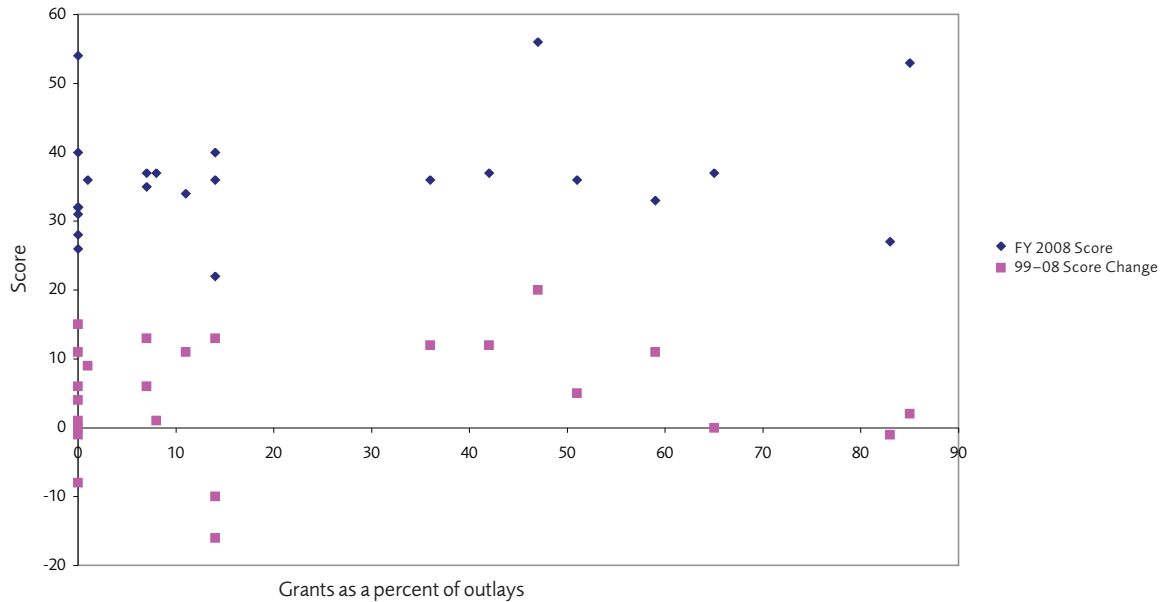
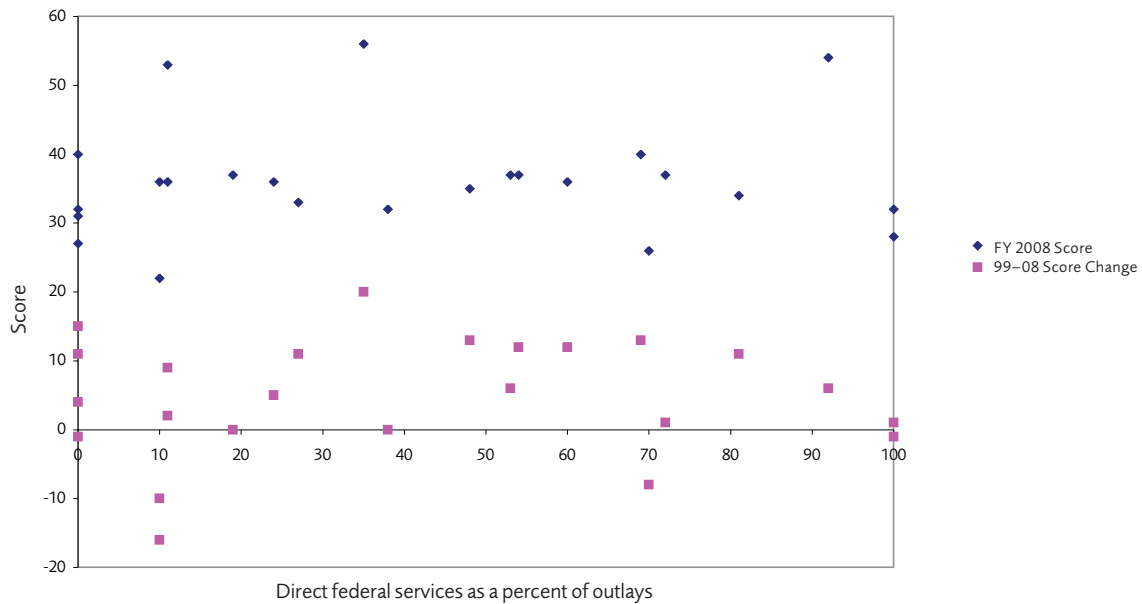
Figure 6 plots scores and score changes against agency size, measured by the agency’s net cost of operations in fiscal year 2008. The points on the far right represent the three very large agencies: Defense, Social Security, and Health & Human Services. Their size does not seem to provide an advantage, as many smaller departments

Figure 6: Scores vs. Agency Size



17. A prominent exponent of this view is Beryl Radin, *Challenging the Performance Movement* (Washington, DC: Georgetown University Press, 2006), 42–50.

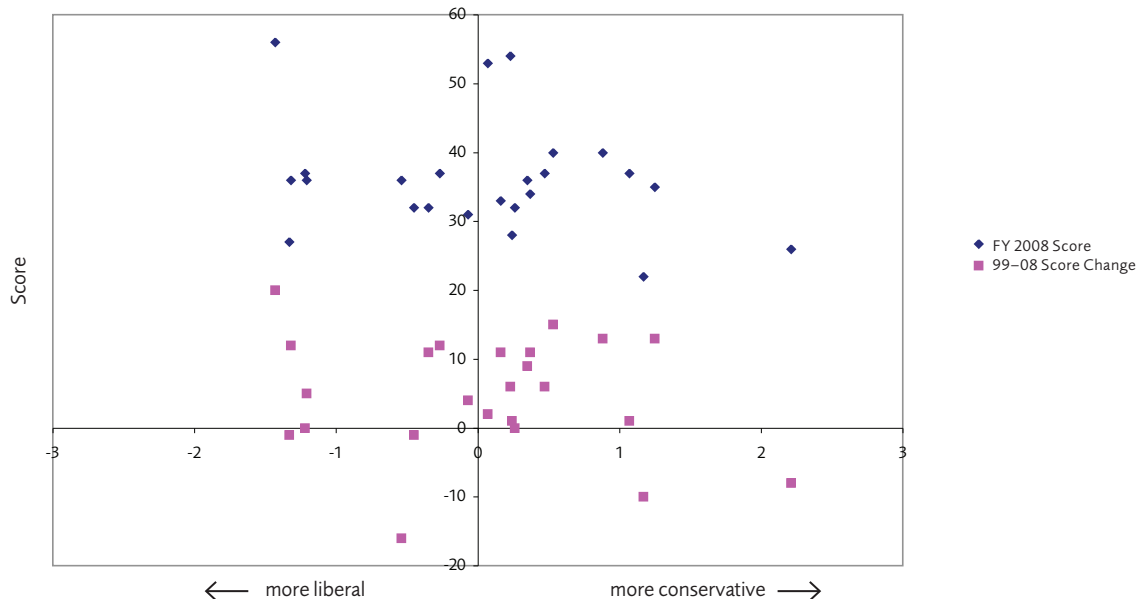
18. See Radin, *Challenging the Performance Movement*, 91–114, 120–21, 189.

Figure 7: Scores vs. Grants as a Percent of Outlays**Figure 8: Scores vs. Direct Federal Services as a Percent of Outlays**

scored much higher and improved at a faster rate. For the rest of the departments, however, there appears to be a weak positive relationship between size, fiscal year 2008 score, and the ten-year change in score. For departments with a net cost of operations below \$100 billion, a larger size might provide a slight advantage.

Figure 7 examines the relationship between scores and grants as a percentage of agency outlays. Agencies that make no grants, plotted on the vertical axis, have a wide range of scores and score improvements. For agencies

Figure 9: Scores vs. Agency Ideology



that make grants, there appears to be little relationship between the percentage of budget devoted to grants and the fiscal year 2008 Scorecard score or the ten-year score change. This suggests that either grant-making agencies have little inherent disadvantage in measuring and reporting on performance or that other factors provide offsetting benefits that counter any disadvantages they have.

Figure 8 gets at a similar issue by plotting score data against the percentage of the agency's budget spent on the direct provision of federal services. Agencies that provide services directly might arguably find performance measurement easier than agencies that have less control over the activities and outputs that produce outcomes. However, fiscal year 2008 scores and ten-year score improvements seem evenly scattered across the range of percentages, suggesting that agencies that provide more direct services have no tendency to score better.

One caveat is in order for figures 7, 8, and some of the subsequent graphs that seem to show no correlation between Scorecard score and other factors. There still may be some degree of correlation after controlling for other factors that influence scores, but graphing data on scores versus a single factor might not reveal this relationship. A two-dimensional graph that appears to show no relationship is really telling us that if there is a relationship, it is either not obvious or not simple.

Figure 9 examines the effect of ideology on scoring by juxtaposing score data against results from a recently published survey that asked scholars, journalists, and Washington think tank experts to evaluate whether individual agencies tend to have more conservative or liberal policy views.¹⁹ Mercatus Center researchers were not involved in this survey, and we might not agree with all of the individual results, but it is one of the more innovative attempts we have seen to quantify agency ideology. A higher positive number means the agency ranked as more "conservative;" agencies with a more negative number were considered more "liberal." There

19. Joshua D. Clinton and David E. Lewis, "Expert Opinion, Agency Characteristics, and Agency Preferences," *Political Analysis* 16 (2008), 3-20.

Figure 10: Scores vs. Perceived Congressional Commitment

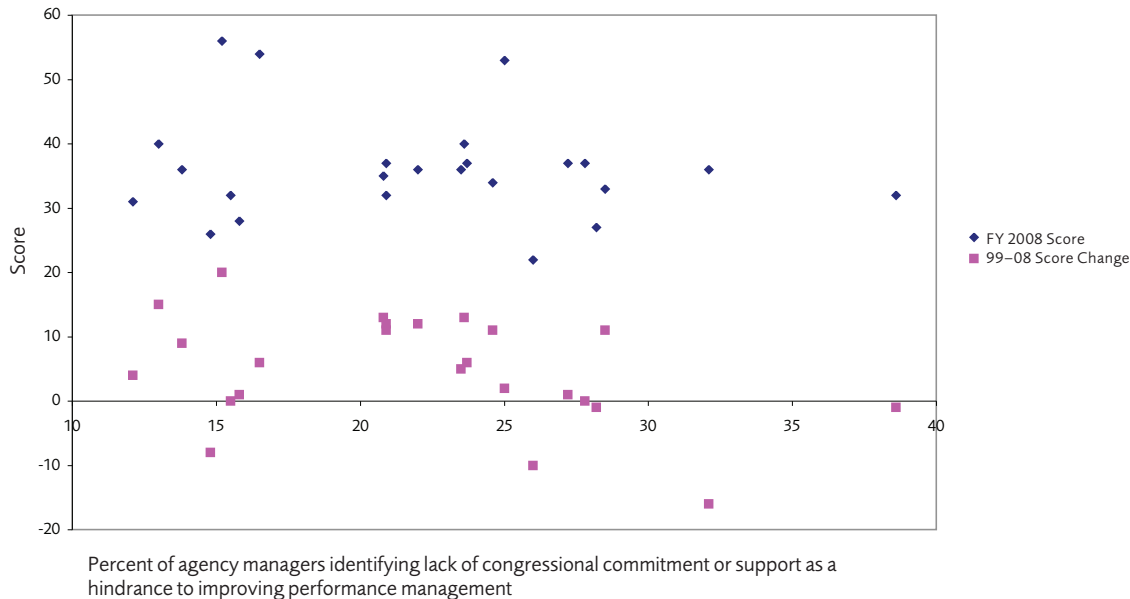
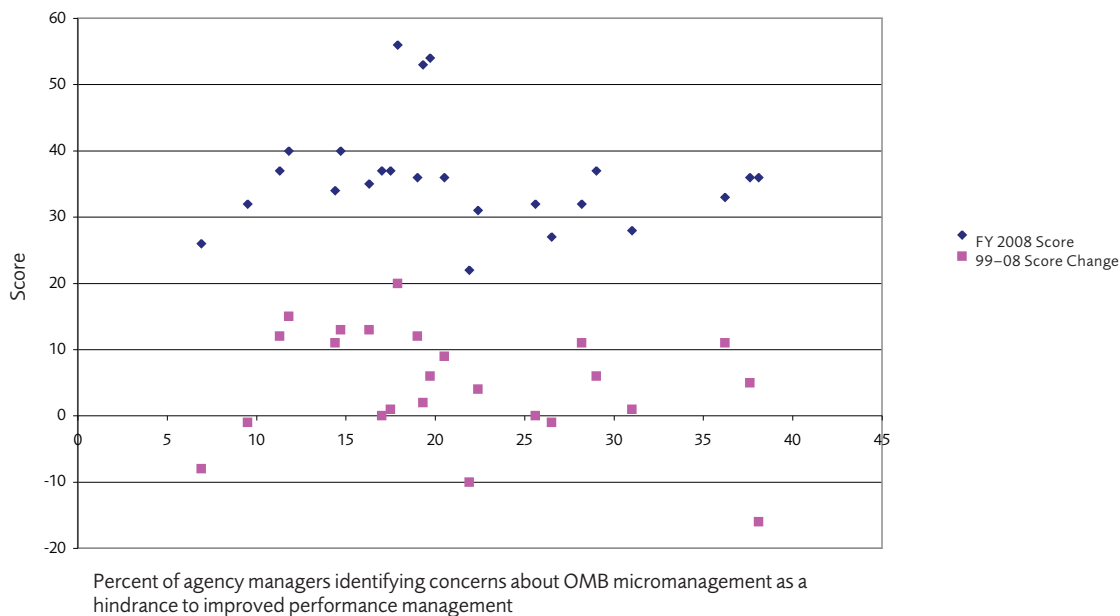


Figure 11: Scores vs. Concern about OMB Micromanagement



appears to be little relationship between ideology and score, or perhaps a slight negative one: more “liberal” agencies may tend to score slightly higher or improve at a faster rate.

Several theories of bureaucracy imply that an agency’s tendency to adopt good performance measures depends in part on whether agency managers believe that the president or Congress wants them to have good perfor-

mance measures.²⁰ Figures 10 and 11 investigate whether agency managers' perceptions of Congress and the OMB are systematically related to Scorecard scores.

In 2007, the GAO conducted a survey of federal managers on the availability and use of performance information, as well as obstacles to more widespread use. One question asked managers whether lack of congressional commitment or support hinders performance measurement or the use of performance information in their agency. Another asked managers whether concerns about micromanagement by OMB hinder performance measurement or the use of performance information in their agencies.²¹

Using this survey data, figure 10 suggests that there may be little relationship between congressional commitment and agencies' fiscal year 2008 scores. However, there appears to be a negative correlation between congressional commitment and improvement in scores over the past ten years. An agency seems less likely to improve if a higher percentage of its managers believe lack of congressional commitment or support is a barrier to improvement.

A similar pattern appears in figure 11. Concern about OMB micromanagement does not seem to be closely related to fiscal year 2008 scores, but concerns about micromanagement do seem to be correlated with the ten-year score change. Agencies with a higher percentage of managers identifying OMB micromanagement as a barrier to better performance management are somewhat less likely to have improved their reports during the past ten years.

In fairness, it is not completely clear what the survey responses about Congress and OMB measure. Taken literally, they measure the extent to which federal managers believe congressional disinterest or OMB micromanagement discourage their agencies from improving performance measurement or performance management. Interpreted more broadly, though, these findings may merely indicate managers' perceptions of the state of their agencies' relationship with Congress and OMB. Responses identifying lack of congressional interest or concerns about micromanagement as a hindrance may simply indicate that managers believe there are problems in the agencies' relationships with their overseers.

Nevertheless, the patterns in figures 10 and 11 suggest that an agency's relationship with Congress and OMB might affect the rate at which it improves the quality of its performance reporting. Perhaps a positive relationship with Congress and OMB leads to more rapid improvement—or more rapid improvement leads to a better relationship with Congress and OMB.

Overall, the graphs in this section suggest that agencies have little inherent advantage related to their size or the means by which they carry out their missions. Either that, or the relationship is much more complicated than a two-dimensional graph can identify. Agencies with more liberal policy views may score a little higher, though the reason is not obvious. Finally, two factors that do seem to matter are perceived congressional and OMB interest in the agency's performance.

20. Anthony Downs, *Inside Bureaucracy* (Boston: Little, Brown and Co., 1967); Gordon Tullock, *The Politics of Bureaucracy*, reprinted in Charles K. Rowley, Ed., *The Selected Works of Gordon Tullock* (Indianapolis, IN: Liberty Fund, 2005 [1965]).

21. Full survey results are available at <http://www.gao.gov/special.pubs/gao-08-1036sp/toc.html>.

STRONGEST AND WEAKEST SCORES

THIS SECTION LISTS the agencies, the performance and accountability reports, and the pilot reports that scored the highest and lowest for each of our 12 evaluation criteria. We briefly explain what made these reports 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 we would offer based on our review of the reports this year.

TRANSPARENCY

1. Is the report easily identified and easily accessible online?

ACCESS TO PERFORMANCE information is critical because public accountability can only be served if members of the public can actually determine the benefits an agency produces. The annual report should be easily available to Congress, 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, agency structure, or the agency's Web site 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 downloading. Making the report available in multiple formats is also desirable since readers' needs vary and each format has its advantages and disadvantages (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.

We apply the following key evaluation factors for criterion 1:

- 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?
- Are prior-year versions of the report readily available?

Strongest Scores: Thirteen agencies (5 points)

MERCATUS RESEARCHERS REGARDED a PAR as "timely" if it was posted by December 1, and we regarded a citizens' report as "timely" if it was posted by January 19. 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. With one exception, the reports were posted on time. Indeed, few agencies needed until December 1 or January 19. Ten of the 15 PAR agencies posted the reports on their submission date, November 17, and 2 more agencies posted their reports within a day or two after November 17. All were posted by our December 1 deadline. Likewise, 8 of the 9 pilot agencies posted their citizens' reports to their Web sites either on the submission date of January 15, 2009 or the following day.

Weakest Score: SBA (2 points)

THE ONE EXCEPTION to timely Web posting was the SBA, which failed to post its citizens' report until well after the Mercatus deadline of January 19. The SBA also failed to post its performance report in time for us to consider it in our evaluation. Finally, the SBA citizens' report failed to identify, reference, or link to the inspector general's presentation on major management challenges in its fiscal year 2008 financial report. This is particularly unfortunate since the SBA inspector general's presentation is one of the most insightful of all such presentations.

Opportunities for Improvement

Make home page links more prominent. Two of the 15 agencies filing PARs lacked a direct home page link to the report. As in past years, the prominence of the links varied considerably. Some agencies (e.g., Interior, Justice, and Labor) highlighted the report via a prominent home page link. Others (e.g., NASA and SSA) had a relatively obscure link in small print at the bottom of their home pages. All nine pilot agencies had a direct home page link to the citizens' report. A number of the pilot agencies had prominent links, including Defense, HHS, OPM, and USAID. In a new feature this year, two agencies (NRC and USAID) included an introductory video by the agency head on the Web page that contains the citizens' report.

Provide better contact information, particularly online. As in past years, most agencies provided contact information for readers interested in obtaining paper copies of the report or asking questions about it. However, the usefulness of the contact information varied. EPA had particularly good contact information in its report. The report's table of contents page and the inside back cover identify extensive report-specific contacts. In addition, an e-mail address for submitting comments or questions about the report appears at the bottom of each page. Some agencies had no contact information online, forcing the reader to peruse the body of the report for such information. In some cases, the online contact information was not specific to the report. A number of agencies (e.g., HUD and Treasury) had online general contact information, such as an "About Us" link, for personnel in the office that produced the report but not necessarily those officials specifically knowledgeable about the report. In our view, the best practice is to post contact information online that puts the reader directly in touch with agency officials who have specific knowledge of the report. Examples of agencies having online, report-specific contact information are Education, Energy, Labor, and Treasury.

Improve links to prior-year reports. Prior-year performance reports for most agencies could be found online, but the ease of accessing them varied. In some cases, it took a series of searches to access the prior reports. For a number of agencies, the Web pages containing their current reports readily produced or linked to the prior-year versions as well: Commerce, Energy, Homeland Security, HUD, Interior, NASA, NSF, OPM, SBA, SSA, and Veterans Affairs. This is a best practice that all agencies could easily emulate.

2. Is the report easy for a layperson to read and understand?

The report should focus on what the agency's mission is, how the agency organizes efforts toward that end, and how much progress was made. The report's 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 with which readers may be familiar, but using acronyms to refer to documents, processes, systems, nouns other than proper 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. We apply the following key evaluation factors for criterion 2:

- Is the report written in plain language that a lay reader can understand?
- Is it well organized and easy for the reader to navigate and identify key information?
- Does it use clear presentational formats that the reader can readily 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?

Strongest Scores: Justice, Labor (5 points)

BOTH JUSTICE AND Labor produced highly readable PARs, which were improved by the addition of excellent citizens' reports. The key portions of the Justice report are much shorter than most PARs. At 22 pages, the Management's Discussion and Analysis section is the shortest of all reports, and the detailed performance section is a relatively modest 54 pages. The report is well formatted and lays out the department's performance results clearly. Justice uses a manageable number of performance indicators: 25 key measures with a total of 33 targets. The report includes reader-friendly tables that present at a glance the department's goals, performance measures, and fiscal year 2008 results.

Justice's citizens' report obviously was prepared with the general public in mind. Its sections have captions such as "The Department of Justice at a glance" and "How do we serve you?" It has excellent, user-friendly links to other source information. It includes the department's "snapshot" as well as summaries of performance results by strategic goals that are concise and well suited to a lay audience. The report features representative performance goals, which, while not consistently outcome oriented, are clear and focused on subjects of obvious importance to the public. It highlights accomplishments in each strategic goal area and includes vignettes to illustrate some of them.

Labor's report is also reasonably short by PAR standards. It is visually appealing and makes effective use of tables and graphics. The secretary's transmittal letter highlights the department's key achievements over time as well as fiscal year 2008 results for each strategic goal area. Each strategic goal presentation begins with an overview of the department's role and the public benefits it provides. The report includes a number of vignettes that illustrate the impact of the department's work on individuals.

Labor's citizens' report includes the department's "snapshot" and uses a shorter version of the secretary's transmittal letter from the PAR that retains the content on performance results but omits the technical content. The report is concise and well organized. It incorporates a number of the strong presentational features from the full report and has user-friendly direct links to other information sources. The main body consists of reader-friendly, insightful summaries of the department's results by strategic goal, using representative and primarily results-oriented performance measures.

TMA (Too Many Acronyms)

In past years, we've been treated to acronyms like LUST (Leaking Underground Storage Tanks) and GWOT (Global War on Terror). These acronyms caught our eyes in the fiscal year 2008 reports.

Acronym	Meaning	Perpetrator
COOL	Country of Origin Labeling <i>Will terrorists think this is cool?</i>	Agriculture
NIPLECC	National Intellectual Property Law Enforcement Coordination Council <i>A very intuitive one. Not!</i>	Commerce
YRBSS	Youth Risk Behavior Surveillance System <i>Sounds like it might be useful in the GWOT.</i>	Education
STARS	Standard Accounting Reporting System	Energy
STRIPES	Strategic Integrated Procurement Enterprise System <i>What a cute pair!</i>	Energy
WHAT If	Watershed Health Assessment Tools Investigating Fisheries <i>Does it include any "What if?" questions?</i>	EPA
IDAMWIG	Identify and Access Management Working Group <i>We wiggled out over this one!</i>	GSA
SPITS	Service's Permit Issuance and Tracking System <i>And what's the output from this system called?</i>	Interior
AFF/SADF	Assets Forfeiture Fund and Seized Asset Deposit Fund <i>Sounds sadistic.</i>	Justice
DOLLAR\$	Department of Labor Accounting and Related Systems <i>Might lead to worker's comp injuries when we stretch to reach the \$ sign on the keyboard.</i>	Labor
CAROLFIRE	Cable Response to Live Fire <i>Is this an instruction or a condition?</i>	NRC
PAAT	Performance Appraisal Assessment Tool <i>Useful for giving employees a PAAT on the back.</i>	OPM
SO	Strategic Objective <i>Almost takes longer to pronounce the acronym than the words.</i>	SSA
DASHR/CHCO	Office of Deputy Assistant Secretary of Human Resources/Chief Human Capital Officer <i>Is Treasury's Bureau of Acronyms under this officer?</i>	Treasury
SMART	Standardized Monitoring and Assessment Relief and Transition <i>Acronyms can be smart as long as you avoid the dumb ones.</i>	USAID
HIS	Indian Health Service <i>Obviously a victim of Microsoft Word's overactive spell-checker.</i>	Veterans Affairs

Honorable mentions:

HHS, for using nine acronyms in the one-page transmittal letter for its citizens' report.
OPM, for using "CR" as an acronym for its citizens' report.

Weakest Scores: HUD, SBA (2 points)

THE HUD REPORT is over 400 pages long and consists primarily of text, with only a few tables interspersed among its lengthy narratives. The report covers 6 strategic goals, 27 strategic objectives, and more than 100 individual performance measures. The detailed content for each strategic goal starts with a “performance scorecard” listing the performance measures and results for each measure applicable to that goal. However, the reader must delve into the following text to discover what some of the measures mean. The secretary’s transmittal letter is brief and says little substantive about the department’s performance. To its credit, HUD voluntarily produced a citizens’ report to complement its PAR. However, this document has little content on the department’s specific performance results.

SBA, a pilot agency, has a citizens’ report that has virtually no useful content within the document and lacks useful links to other information sources. While brief, it contains little information about the agency’s program performance. There is no agency head transmittal letter. Few specific performance measures are described. There are no accompanying narratives to explain the agency’s performance on the few measures that are included. Less than half of one page of the report is devoted to describing fiscal year 2008 accomplishments even in general terms. It does not mention the major management challenges identified by the inspector general and does not reference the inspector general presentation in the financial report. In sum, it falls far short of providing the overview of agency performance that a citizens’ report is supposed to convey. While SBA’s fiscal year 2008 performance report may supply much of the missing content, it was not accessible online at the time of our review.

Opportunities for Improvement

Shorten the PARs. A basic challenge for PARs is to convey information to the general public as concisely as possible through a document that has much required (and often highly technical) content and which must serve audiences with different needs. Size certainly can be a barrier to readability. Given its scope, the PAR will always be a fairly large document. However, the fact that PARs vary greatly in size indicates that agencies have considerable flexibility and that the reports need not be oppressively long. PAR agencies can (and do) use many techniques to kill fewer trees and package information more concisely. One common practice to reduce report length and enhance readability is to use fewer performance measures. This is a good approach for all agencies, PAR and pilot alike, and is discussed further below. Another technique, discussed below, is to use tables and graphics to summarize performance results. This saves space and enhances readability. Also, the reports might omit some technical content (e.g., data verification and validation details) in favor of providing a link to it online for those readers who are interested. Above all, report drafters should be less verbose. For example, it is hard to believe that the Management’s Discussion and Analysis sections of the reports, which are supposed to be overviews, need to consume an average of more than 50 pages as they now do. It also is hard to believe that descriptions of major management challenges need to consume over 80 pages (e.g., Transportation).

Make the citizens’ reports of pilot agencies more substantive and user friendly. While the PAR reports contain too much, we found that the citizens’ reports of pilot agencies often contained too little. For a pilot agency, the citizens’ report is the key document to inform the general public. It needs to serve as a self-contained, substantive overview of the agency’s performance and provide user-friendly links to other source documents for more detail. Three of the pilot citizens’ reports (Defense, OPM, and SBA) fell well short of this standard. The others were better, but could be improved. In particular, all need more user-friendly links directly to the relevant portions of their agency performance and financial reports. The Justice and Labor citizens’ reports serve as models for the pilot agencies both for their content and their direct links to other sources.

The following suggestions to improve readability apply to both PARs and pilots.

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. Since 1998, federal agencies have been obligated to write documents in “plain language” with logical organization, everyday words, active voice, short sentences, and easy-to-read tables and graphics, among other features.²² Report writers should consider 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. Good venues for this are the citizens’ report, the transmittal letter, and the Management’s Discussion and Analysis section of the PAR. As discussed previously, Justice’s citizens’ report provides an excellent example of a document written for the general public. By contrast, the performance section of the GSA report offers an example of how not to write for the public. It is organized primarily by the agency’s components rather than by its strategic goals. This organization may make it easy for agency insiders to use the report, but it makes the report disjointed and difficult for the general public to navigate.

Focus on a clear and concise set of performance metrics. Some agencies have been quite successful in presenting their performance through a limited set of measures that capture important public outcomes. For example, Justice uses a total of 25 performance measures; Agriculture gets by with 33 measures. Agencies that find it necessary to have many more performance indicators can ease the reader’s burden by focusing their reports on a smaller set of key representative measures. Agencies that accomplished this include Commerce, GSA, Homeland Security, Interior, and Veterans Affairs.

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 performances 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 Veterans Affairs and SSA.

3. Are the performance data reliable, credible, and verifiable?

THE REPORTS CONSOLIDATION Act of 2000 (Public Law 106-531) provides that the agency head transmittal letter must include “an assessment by the agency head of the completeness and reliability of the performance and financial data used in the report” and that the assessment must “describe any material inadequacies in the completeness and reliability of the data” as well as what the agency can do and is doing to resolve them.²³

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. More generally, the report should indicate the agency’s confidence in the quality of the data used to document its results. Data sources and descriptions should be provided for all performance measures.

22. See President Clinton's *Memorandum on Plain Language in Government Writing* (June 1, 1998), <http://www.plainlanguage.gov/whatisPL/govmandates/memo.cfm>, and “How to Comply with the President’s Memo on Plain Language” (July 1998), <http://www.plainlanguage.gov/howto/guidelines/PresMemoGuidelines.cfm>.

23. These requirements are codified at *U.S. Code* 31 § 3516(e).

We apply the following key evaluation factors for criterion 3:

- 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?

Strongest Scores: Labor, Veterans Affairs (5 points)

THE DATA QUALITY Act obligates federal agencies to ensure and maximize “the quality, objectivity, utility, and integrity of information (including statistical information) disseminated by the agency. . . .”²⁴ The Labor report systematically assesses and rates the quality of its data for each performance goal, a best practice for other agencies to emulate. While the report identifies some concerns, it states that performance data are accurate and relevant for all goals. The performance section elaborates on data quality for each goal and provides data sources for each measure.

The Veterans Affairs report includes a detailed discussion of data quality, which has been expanded this year to include more specifics on verification techniques. Tables in the performance section provide detailed background on the data supporting the key measures, including data definitions, data validation and verification, and any data limitations. Additional data definitions and background appear later in the report.

Weakest Scores: Defense (1 point); Eight agencies (2 points)

THE DEFENSE REPORT has no content on performance data. The transmittal letter does not include the agency head data assessment required by the Reports Consolidation Act, and the body of the report does not address data quality in any way. The department received a disclaimer of opinion on its fiscal year 2008 financial statements. A disclaimer of opinion means the financial statements were not complete or accurate enough to allow the auditors to judge whether the statements fully and fairly present the agency’s financial position.

Many reports lack sufficient content to give the reader confidence in their supporting data. Some reports’ contents (or lack thereof) seem to contradict assurances that data are complete and reliable. The HHS secretary’s transmittal letter states that the report’s financial and performance data are reliable, complete, and the latest available except where otherwise noted. However, data completeness obviously is a major issue for HHS, because its report indicates fiscal year 2008 results are available for only 40 percent of the department’s performance targets overall and only about one-third of the measures featured in the report. The transmittal letter for the OPM report does not include the agency head data assessment required by the Reports Consolidation Act. Instead, the agency’s “performance improvement officer” gives a data assessment in the transmittal letter, stating that the performance information contained in that report “is reasonably complete and reliable, as defined by the Government Performance and Results Act.” However, the report later cites questions about the validity of two of its own measures as a reason for failing to meet them.

24. Cited in Office of Management and Budget, *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies*, January 3, 2002, <http://georgewbush-whitehouse.archives.gov/omb/fedreg/reproducible.html>.

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 (such as the results of multi-year basic research and development programs). Indeed, the more outcome-oriented an agency's performance metrics, 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 these weaknesses. From this perspective, two recurring problem areas call for improvement.

More candid agency-head data assessments. Unfortunately, many reports respond to the Report Consolidation Act's data assessment requirement with a perfunctory statement that the report's data are complete and reliable in accordance with OMB guidance or except as otherwise noted. Agency-head transmittal letters occasionally make such assertions even though the reports lack even preliminary current results for some measures or present what appear to be significant data reliability issues. This suggests that agency heads need to be more forthright, that the OMB guidance needs to be tightened up, or both. The EPA administrator's transmittal letter is one of the few that provides a candid and informative data assessment. He states:

Data used to report progress are reliable and as complete as possible. Inherent to the nature of our work is a time lag between when we take action to protect human health or the environment and when we can measure a result from that action. Therefore, for the reporting year, we cannot provide results data for several of our performance measures; however, we portray trend data, when possible, to show progress toward results over time, and we present final results for prior years when data have become available in fiscal year 2008.²⁵

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 agency uses. The credibility and usefulness of the report can be enhanced by providing data definitions and sources for individual measures. It is also helpful to describe what the agency does to verify and validate performance data. As noted above, Labor and Veterans Affairs are examples. Other examples of reports with good content on data verification and validation are Commerce, NRC, and NSF.

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 are making a given situation better or worse—by how much and at what cost. 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

25. Stephen L. Johnson, "Administrator's Letter", *U.S. Environmental Protection Agency Performance and Accountability Report Fiscal Year 2008* (Washington, D.C.: EPA, 2008), vii.

achieving its goals, or explain why the agency will produce a significant level of public benefits if it hits its targets. A report should explain rationales behind the selection and alteration of quantitative targets so readers can understand the magnitude of the agency's goals in relation to the size of the problem.

We apply the following key evaluation factors for criterion 4:

- 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?
- Do the report narratives assist the reader by analyzing performance results and trends?

Strongest Scores: Labor, Veterans Affairs (5 points)

THE PERFORMANCE SECTION of the Labor report includes baseline and trend data for each measure generally going back four fiscal years or program years. These data include the prior-year targets, actual results, and costs. The report uses graphics to show further trend data for some goals and measures. In a best practice, the performance section includes for each goal a "Program Perspective and Logic" discussion, which provides valuable context to assist the reader in understanding the department's performance.

The Veterans Affairs report is rich in baseline and trend data and does a good job of integrating these data into its presentations. Performance summaries in the report show prior-year performance targets and results for key measures and representative supporting measures, generally going back to fiscal year 2004. The tables in the performance section show the strategic (long-range) target for each measure. This provides the reader with valuable context to assess the extent of the department's progress and how far it needs to go.

The Transportation and Treasury reports also feature good presentations of baseline and trend data. The Transportation report includes a performance "forecast" for each measure for fiscal year 2009. The narratives in the performance section elaborate on the department's results, adding valuable context for assessing its performance. In an innovative feature and potential best practice, the Treasury report uses graphics to show the trends for target achievement and actual performance over time.

Weakest Scores: Nine agencies (2 points)

THE ENERGY REPORT'S presentations are an incomplete and confusing mix of prior-year targets and results. Many of the measures are technical, and the narratives do little to help a lay reader understand the measures or put the department's performance in context. NASA and NSF largely use non-quantifiable performance measures that afford little basis for assessing progress from year to year. The Commerce and SSA reports have baseline and trend data suggesting that the agencies tend to set "lowball" performance targets; current targets are significantly below prior-year performance levels. There may be good reasons for this, but the reports do not explain them. SSA offers a particularly striking example. The administrator's transmittal letter describes

eliminating the hearing backlog for disability claims as a “moral imperative.” However, the backlog will not be eliminated until fiscal year 2013 under the agency’s new strategic plan.

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 an agency’s current targets are. The better reports have two additional features that make the baseline and trend data more insightful and put the agencies’ performances in fuller context.

Explanations for targets. One additional feature is an explanation of how performance targets are set. 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. As noted above, the Labor report is strong in this respect. The Education report also includes a good “target context” feature.

Long-term perspective for performance trends. The second feature is providing a sense of each of the agency’s long-term goals in relation to its respective 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 agency’s progress toward its destination. The Veterans Affairs and EPA reports do provide long-term strategic targets for this purpose. Several other agencies (e.g., Homeland Security and USAID) provide performance targets at least one year into the future.

PUBLIC BENEFITS

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 terms that describe benefits to the community rather than activities that are presumed to be of value. For example, the number and percentage of children learning to read at grade level is an outcome; the number of children enrolled in a remedial reading program, or the number of schools receiving grants to fund these programs, are measures of activity, not outcomes. 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, strategic objectives, and annual performance goals.

Strategic goals should be few in number. Management goals (financial, human resource, information technology, etc.) exist to support the achievement of genuine outcome goals. A majority of agencies now include a strategic goal focused on management improvement, but these should not be confused with actual programmatic outcome goals.

Criteria 5, 6, and 7 are closely related and all are crucial for effective performance reporting. 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 earn a high score on criterion 6 without a high score on criterion 5.

We apply the following key evaluation factors for criterion 5:

- 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?

Strongest Scores: Labor, Transportation (5 points)

WHAT MAKES THESE two reports stand out is that the great majority of the goals and objectives are stated as clear, measurable outcomes of obvious importance that the general public can readily understand and appreciate. For example, Labor's goals focus on such outcomes as safer workplaces and employment training programs that result in trainees obtaining and retaining jobs. Transportation's first strategic goal emphasizes increasing the safety of highway, aviation, and rail transportation by reducing accidents, deaths, and injuries. Its second goal aims at increasing transportation efficiency by reducing highway congestion and aviation flight delays. Other reports featuring strong outcome-oriented goals are Education, Health & Human Services, Homeland Security, NRC, and Veterans Affairs.

Weakest Scores: Six agencies (2 points)

THE WEAKER REPORTS lack goals with clear, measurable outcomes. Some have strategic goals, but they refer to outcomes at too high a level for the reader to understand how and what the agency contributes. Commerce's first strategic goal, for example, is to "Maximize U.S. competitiveness and enable economic growth." Reports frequently obscure the agency's role by using vague terms, such as "facilitate" or "enable." These reports also are short on results-oriented strategic objectives or annual performance goals that might flesh out the measurable outcomes.

Opportunities for Improvement

SCORES FOR THIS criterion have changed little from last year. Since few agencies implemented new strategic plans for fiscal year 2008, their strategic goals and objectives, as well as the annual performance goals, remained largely the same. This is not surprising given that this was the last year before a transition of administrations. One agency that did significantly change its strategic goals and objectives this year was Education, which made major improvements to the clarity and outcome orientation of its goals and objectives. Also, while retaining the same strategic goals as last year, State added descriptions to them for fiscal year 2008 that specify more clearly the measurable outcomes envisioned for each goal. State added text to its strategic objectives as well to clarify its intended outcomes.

The coming of a new administration offers an opportunity for agencies to develop new strategic plans and, in the process, improve the clarity and outcome orientation of their goals. Any agency should be able to develop a set of long-term strategic goals that capture the important public outcomes the agency exists 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 down their current high-level strategic goals into more specific and measurable outcomes that could be tied with some credibility to the agency's contributions.

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 that promotes the end outcome of fewer highway deaths and injuries. It is a good intermediate indicator as long as the agency has factual evidence that increased use of seat belts does indeed lead to the desired outcome. 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, or projects funded, 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 achievement.

We apply the following key evaluation factors for criterion 6:

- 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?

Strongest Scores: Labor, Transportation (5 points)

NOT SURPRISINGLY, THE Labor and Transportation performance measures are highly outcome oriented since both reports have strong outcome goals. Labor's performance measures track well with their respective goals and are highly outcome oriented for the agency's strategic goals relating to a prepared workforce, safe and secure workplaces, and strengthened economic protections. The measures applicable to the competitive workforce strategic goal also include strong outcomes along with activity, output, and efficiency-type measures. All of the Transportation measures relating to the agency's programmatic strategic goals for safety, congestion, and security capture end or intermediate outcomes. At least half the measures for the global connectivity and environmental stewardship strategic goals focus on end outcomes or intermediate outcomes.

Weakest Scores: Eleven agencies (2 points)

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. For example, although State made its goals more outcome oriented this year, its measures have become less outcome oriented. By our count, only 2 of the 16 programmatic performance measures featured in the State report capture outcomes. Most are raw

number measures. Many of the numbers deal with activities or outputs, such as number of election observers trained with U.S. assistance, number of host-country officials trained in counterterrorism, or number of foreign military personnel trained. Energy and NASA compound this problem by using arcane activity measures that are hard for a lay reader to understand, such as Energy's measure of "Efficiency of 'white light' solid-state lighting in a lab device, in lumens per watt (LPW)" and NASA's "Design, build, and deliver for flight two Foton M3 experiments." Many agency measures focus heavily on outputs, such as levels of assistance provided, but fail to capture what public benefits the assistance achieves. HUD and SBA are examples. Efficiency, workload, productivity, and customer satisfaction measures may be valuable for certain purposes, such as internal agency management. However, they tell the public little about outcomes.

Opportunities for Improvement

Convert raw numbers measures into outcome form. Many agency measures have the potential to capture outcomes but are not stated in an outcome-oriented manner. Raw-number measures are a recurring case in point. They may relate to outcomes but are expressed in a way that only measures activity levels. For example, one highlighted measure used by Homeland Security is the number of border miles under effective control. Border security certainly is an outcome. However, simply reporting on the number of border miles under control gives the reader little sense of progress toward that outcome. The measure would be more meaningful and outcome oriented if expressed as a percentage of border miles under control. There are many opportunities, particularly in law enforcement areas, to convert raw-number measures into percentages or other forms that capture progress toward the desired outcome. For example, such measures as the number of organized criminal enterprises dismantled or quantities of illegal substances seized do not tell the reader whether the outcome-related conditions they address (prevalence of criminal gangs or illegal drugs) are getting better or worse.

Eliminate measures lacking external significance. In addition to developing more outcome-oriented 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 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 also fall into this category. Indeed, they are sometimes treated as a substitute for true outcome-oriented measures that express more directly the degree to which an agency is meeting the public's needs.

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. Claims of impact should be supported by program evaluations or other empirical evidence of a cause-and-effect relationship. A less desirable alternative would be to logically connect outcome measures, output measures, and anecdotal evidence. A case that rests on merely assumed cause-and-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 as excuses for a failure to demonstrate performance.

We apply the following key evaluation factors for criterion 7:

- 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 provide sufficient current results, prior-year data, and other information to give the reader a complete picture of the agency's performance?

Strongest Score: Transportation (5 points)

THOSE AGENCIES WITH strong, outcome-oriented performance goals and measures will do well under this criterion, particularly if the narratives describing specific performance results are also strong. As noted previously, Transportation's performance metrics are predominantly results-oriented and focus on outcomes of obvious public importance. The narratives in the performance section complement the strong metrics and are highly effective in elaborating upon the department's performance results. Labor, NRC, and Veterans Affairs likewise scored well here based on their strong performance metrics and insightful descriptions of performance results.

Weakest Scores: Fourteen agencies (2 points)

JUST AS STRONG performance metrics largely determine success on this criterion, the predominant feature of the low-scoring reports is weak performance metrics—specifically, the absence of significant outcome orientations in the agencies' goals or measures. Of the 14 reports that scored a 2 here, all but one also scored a 2 for their goals (criterion 5), their measures (criterion 6), or both. The exception was GSA, which received a score of 2 on this criterion for incomplete reporting. In addition to weak performance metrics, some of the low-scoring agencies had weak descriptions of performance results (e.g., SBA); appeared to set non-challenging performance targets (e.g., Commerce and SSA); or had technical measures that were hard to understand (e.g., Energy and NASA). Three agencies (Energy, NASA, and NSF) constructed their measures in such a way that they say virtually nothing about value to the general public.

The absence of current (fiscal year 2008) results for a high proportion of performance measures is a significant limitation for four agencies: Education, EPA, GSA, and Health & Human Services. There is no apparent reason why GSA fails to report current results for many of its measures. For the other three agencies, the problem stems from data lags that may be largely unavoidable. Education and HHS deal with outcomes that take years to materialize and do not readily lend themselves to measurement in annual increments. To its credit, HHS mitigates this problem somewhat by reporting the most recent performance results available for those measures lacking fiscal year 2008 results.

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, of course, but criteria 2 through 4 also come into play.

8. Did the agency link its goals and results to costs?

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 corresponded to the following standards: 1 point if the report had no relevant content or linked budget resources only to strategic goals, 2 points if the linkage was carried down only to the strategic objectives, 3 points in cases where there was some useful content beyond the strategic objective level but short of the performance goal level, and 4 points if the linkage was carried down to the individual performance goals. To earn 5 points, the report needed to link budget costs to individual performance measures.

We apply the following key evaluation factors for criterion 8:

- 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?

Strongest Scores: Labor, Transportation (5 points)

THE LABOR REPORT links costs to the strategic goals and performance goals for the current year and the two previous fiscal or program years. In addition, the performance section links costs for the current year and several prior years to most of the department's individual performance measures. Likewise, Transportation has links to its individual performance measures. Other agencies are making progress here by linking costs to a portion of their performance measures or by developing innovative techniques to relate costs to performance.

Weakest Scores: Eleven agencies (1 point)

THE REPORTS THAT scored 1 for this criterion either failed to link budget resources to their performance metrics at any level or provided linkages only at the strategic goal level.

Opportunities for Improvement

WITH RESPECT TO this criterion, agencies can be divided into two camps: those that are making progress toward linking costs to performance metrics in a meaningful way and those that are not. The former have scores of 3 or higher. The latter have scores of 1 or 2. For the agencies that fall into the latter camp, the most obvious improvement strategy is to at least start trying. Developing linkages to the point of permitting meaningful cost-benefit analysis 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 investments. 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.

Some agencies are pursuing innovative techniques here that may provide a model for others. The Treasury report presents information primarily on the basis of "performance cost" as opposed to budgetary resources. As the report explains, this approach provides a more accurate picture of the actual costs related to performance

since it includes “imputed costs, depreciation, losses, and other expenses not requiring budgetary resources.” It also covers costs for programs funded by sources other than general-fund appropriations, such as user fee-financed or non-appropriated fund activities. Interior allocates costs to the 25 representative performance measures featured in its report and, through activity-based costing graphs, shows performance results in relation to funding levels. The narratives also provide some analysis of the impact of funding levels on performance. This innovative feature is a potential best practice for performance reports.

FORWARD-LOOKING LEADERSHIP

9. Does the report show how the agency's results will make this country a better place to live?

AGENCIES SHOULD REALIZE and articulate the value they provide to the country. Reports should speak directly to the public about how agencies produce benefits that are important to citizens. Just as the best corporate reports feature communication directly from the chief executives, agency reports should demonstrate accountability of agency heads for their organizations' performances. Lofty ideals must be supported by outcome orientation, sound strategies, and successful achievement discussions. The reports should create confidence in agencies' abilities to improve America's future. Anecdotes and success stories can be important communication strategies in this regard, but their value is limited if they are not backed up by solid performance data.

We apply the following key evaluation factors for criterion 9:

- 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?

Strongest Scores: Transportation, Veterans Affairs (5 points)

COMPLEMENTING ITS STRONG performance metrics, the Transportation report's narratives do an excellent job demonstrating the important public benefits that flow from the department's work. For example, the secretary's transmittal letter and the Management's Discussion and Analysis section both highlight major performance accomplishments. The secretary's transmittal letter (p. 3) notes that increased seat belt use in 2008, largely due to high-visibility enforcement campaigns, resulted in an estimated 270 lives saved. The performance highlights section (p. 18) states that the fatality rate from highway accidents has dropped more than 13 percent from 1998 to 2007. The Veterans Affairs report narratives do an excellent job of describing the department's accomplishments and the public benefits that flow from them. The secretary's transmittal letter highlights a number of specific accomplishments. The Management Discussion and Analysis section details the most important accomplishments for each strategic goal area. The personal vignettes further contribute in this regard. These and other narrative portions of the report are backed up by solid performance metrics.

Labor also scores well on this criterion. A vignette in the Labor report (p. 102) describes how a low-income widower at age 80 was able to obtain work through training provided under a department-funded mature workers program. Another (p. 111) concerns a disabled Iraq War veteran who was unable to return to his prior

civilian job but acquired enhanced information-technology skills through a department-funded training program. Through this training, he qualified for a new job teleworking from home. In yet another example (p. 141), the department funded grants in 20 states to use debit cards to make faster and more secure unemployment insurance payments to claimants who lacked bank accounts.

Weakest Scores: Defense, HUD, OPM, SBA (2 points)

DEFENSE, OPM, AND SBA are pilot-report agencies whose inadequate, bare-bones citizens' reports convey little insight into their performance results or public benefits achieved. The lengthy HUD report has some useful narrative descriptions of performance highlights and accomplishments. However, they do not compensate for the department's weak performance metrics. Indeed, a reader might wonder about the credibility of the department's metrics when it reports achieving over 85 percent of its targets in a year that featured a nationwide housing crisis. The secretary's transmittal letter has little substantive content in terms of performance accomplishments or strategies to address future challenges.

Opportunities for Improvement

Maximize opportunities to tell the agency's story to the public. The narrative portions of the report, which have less constrained formats 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 transmittal letters as more than perfunctory documents that recite the necessary legalisms. The best transmittal letters highlight important outcome-oriented accomplishments. Examples include Agriculture, Interior, Labor, and Transportation.

Highlight accomplishments to illustrate performance results. The narrative portions in the better reports describe performance highlights in a way that the general public can understand and appreciate. They frequently include vignettes illustrating 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. In particular, NSF should describe more examples of public benefits that result from the projects it funds since its performance metrics provide little insight about effects on the public.

Improve performance metrics. 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 the 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 the 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.

We apply the following key evaluation factors for criterion 10:

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

Strongest Scores: Veterans Affairs (5 points); NASA, Transportation, and Treasury (4 points)

THE VETERANS AFFAIRS report clearly discloses performance results, including shortfalls. It contains a specific “Performance Shortfall Analysis” section, which describes causes and resolution strategies for significant performance shortfalls. This is a useful feature and a potential best practice. Under the heading “How VA Leadership Uses Results Data” in the detailed performance section, the report briefly describes performance improvement steps even where targets were met. Likewise, a table in the NASA report, captioned “FY 2008 Performance Improvement Plan,” lists each measure that was not fully achieved, explains why, and outlines plans for achieving the measure in the future. The Treasury report specifies gradations of performance results, both in the percentage of achievement for individual targets and the level of success. Where targets were missed, an appendix consistently provides an explanation under the heading “Future Plans/Explanation for Shortfall.” The explanations are brief but usually informative, and they are accompanied by a general description of remedial steps. The Treasury report clearly discloses performance shortfalls and offers concise and generally informative explanations where targets were missed or are projected to be missed. The secretary’s transmittal letter candidly highlights several major performance concerns.

Weakest Scores: SBA (1 point), 12 agencies (2 points)

THE SBA REPORT provides no explanations for the four shortfalls shown, all of which missed their targets by substantial margins ranging from 20 to 35 percent. SBA’s citizens’ report does discuss economic factors influencing loan volume, but it does not link these factors to the missed targets. Most of the reports scoring 2 for this criterion had incomplete or weak explanations of performance shortfalls. For example, the Commerce report offered no explanation for two of the four key measures that were missed. The explanation for another missed key programmatic measure was vague and unsatisfying, especially considering that the shortfall involved highly publicized problems with a proposed automated census system. The Education report had weak explanations for the three shortfalls reported. One suggested that the measure was arbitrary; another suggested that the target was too ambitious; the third really provided no explanation. The explanations of shortfalls in the OPM report sometimes seem to raise more questions than they answer. For example, the report attributes one shortfall to the fact that the performance data used did not align with the measure. The explanation of another shortfall calls into question the validity of the measure.

Opportunities for Improvement

Explanations, not excuses. Explanations for shortfalls are most valuable as a diagnostic tool to identify their causes and devise improvement strategies. Too often, however, the explanations amount to no more than excuses that provide no useful insight. Common examples are statements to the effect that the target was challenging or too ambitious, the agency came close and will try harder next year or the shortfall was attributable to factors beyond the agency's control. These may sometimes be true, but if so, the report needs to explain how the agency will correct its planning process so that those kinds of mistakes do not recur. The most useful shortfall explanations provide specifics that point the way to future improvement.

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. NSF serves as an example: The agency reported 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, while perhaps technically sound, is highly unlikely to surface any instances of unsuccessful or even marginal projects. In NSF's assessment system, outside experts base their reviews primarily on best examples of projects provided to them by NSF program managers, rather than using a random sample or another more comprehensive or representative slice of the projects the agency funds. Thus, evaluators are not routinely exposed to anything but the agency's best work.

11. Does the report adequately address major management challenges?

THE REPORT SHOULD describe how risks to an agency's success are being minimized 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 the GAO, agency responses indicating an appreciation of threats to its mission and goals, and an anticipation of future risks.

The Reports Consolidation Act requires performance reports to include a statement by the agency's inspector general that (1) summarizes the agency's most serious management and performance challenges and (2) "briefly assesses the agency's progress in addressing those challenges."²⁶ These statements play an important part in our evaluations for criterion 11. As we have done in prior years, the research team also considered the agency's "status" scores on the President's Management Agenda.

The inspector general statements play an important part in our evaluations for criterion 11. We also carefully review the agency management's response, if there is one, to the statement. It is important for the agency response to identify any differences it may have with the inspector's assessment. The inspectors, of course, do not have a monopoly on truth, and agencies need not agree with their assessments of the challenges or their view of priorities. The reader is best informed by a complete and candid exchange between the inspector and agency management. We found several instances of differences between inspectors and agency management. While HUD management generally agreed with the challenges identified by its inspector general, it rejected

26. U.S. Code 31 §3516(d).

some aspects of the inspector’s problem statements and explained why. For example, management disagreed that the department lacked an adequate database to implement a program for the redevelopment of abandoned and foreclosed homes (see pp. 378 and 386). Also, Treasury countered its inspector’s assessment of progress in improving performance and financial data at the Internal Revenue Service (see pp. 257–258 and 268). The most striking example is the GSA report (pp. 186–188), in which agency management takes issue with most of the inspector general’s assessment.

We apply the following key evaluation factors for criterion 11:

- Does the report include the inspector general’s summary and assessment of the agency’s most serious management and performance challenges?
- Does it include a response by agency management to the inspector general’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?

Strongest Scores: Labor (5 points); EPA, Transportation, and Veterans Affairs (4 points)

LABOR’S INSPECTOR GENERAL does a good job of summarizing the challenges and assessing the department’s progress toward resolving them. The department itself uses a “traffic light” system to assess its own progress on each challenge. The assessments are presented in tables that discuss each challenge, the actions taken in fiscal year 2008 to address it, and the remaining actions needed along with their expected completion dates. Both the inspector’s assessments and the department’s self-assessments provide confidence that the department is working seriously to address each challenge. Another source of confidence is that the department has achieved “green” ratings on each of the core items of the President’s Management Agenda for the past two fiscal years.

The EPA report describes material weaknesses and other significant deficiencies identified by the OMB, the GAO, and the agency’s inspector. Like Labor, EPA achieved “green” scores on all the core items of the President’s Management Agenda. The Management’s Discussion and Analysis section of the Transportation report breaks down the challenges identified by the department’s inspector general into their component parts and uses a “progress meter” graphic to self-assess its progress toward resolving each component. The department reports “moderate” or better progress for most items. Ironically, departmental management’s discussion of the challenges is much more specific and insightful than the inspector’s verbose and at times confusing presentation. The Veterans Affairs report has a table showing each inspector-identified major management challenge, its underlying components, and the strategic goal to which it most closely relates. The table also gives an estimated time frame for resolution of each challenge—although many are open ended (“2009 and beyond”)—and it provides a page reference for a detailed discussion of the challenge. The report uses the same format to describe high-risk areas applicable to the department as identified by the GAO.

Weakest Scores: Seven agencies (2 points)

THE FUNDAMENTAL COMMON shortcoming in the low-scoring reports is that they provide the reader with little or no confidence that the agencies are serving as good stewards of taxpayer dollars by taking effective steps to address their major management challenges. The reports often are long on statements of commitment to resolve the challenges and general assertions of progress, but short on specific descriptions of 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. Inadequate inspector general assessments often contribute to the weakness of the reports in this area.

Opportunities for Improvement

Better inspector general assessments of progress. The Reports Consolidation Act requires reports to include a statement by the agency inspector general that not only summarizes the most serious management and performance challenges, but also assesses the agency's progress in addressing them. The reports consistently include the inspector general's descriptions of major management challenges. What is often missing, however, is the required assessment of progress, which Congress considers an important "barometer."²⁷ Too many inspector general statements are lengthy tomes that describe the challenges in great detail but offer little if any insight into what key steps are needed to remedy them, whether the agencies are taking those steps, and when they are likely to produce a solution. Indeed, agency management's self-assessments of the challenges and progress toward resolving them can be more informative (and more concise) than the inspector general's statement. Transportation is an example.

For a number of years, the SBA inspector general presentation has provided the best example of an insightful assessment that should serve as a model for other inspectors general. The SBA inspector'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 list remaining actions needed to resolve it. It includes a table summarizing the agency's overall progress for each action and challenge. Other agency inspectors general are now starting to follow suit. Like SBA, the Homeland Security inspector uses a "scorecard" to assess the department's overall progress on four of the nine challenges identified as "Limited," "Modest," "Moderate," or "Substantial." His scorecard also breaks down the four selected challenges into components and rates the department's progress on the components as well. The inspectors for Education and Labor also provide good progress assessments, although they are not as specific as those of the SBA and Homeland Security.

More specifics from agency management. Of course, agencies also need to improve their own presentations on major management challenges and give more specifics on the remedial actions they are taking as well as the timing and significance of those actions. More specific inspector assessments probably would encourage the agencies to be more specific in their responses. A few agencies do not respond at all to the inspector general presentations; they need to do so.

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 using knowledge gained from the reporting process to

27. S. Rep. no. 106-337 (2000), 4.

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?

We apply the following key evaluation factors for criterion 12:

- 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 evidence efforts to enhance performance accountability and reporting?

Strongest Scores: Labor, Veterans Affairs

BOTH THE LABOR and Veterans Affairs reports contain thorough, substantive, and thoughtful descriptions of improvement strategies for specific programmatic shortfalls and major management challenges. They also do an excellent job of describing future challenges that transcend individual measures and how the department plans to meet them. Finally, both departments continue to be leaders in developing innovations and improvements with respect to performance accountability and reporting. The Transportation report shares many of the same features.

Weakest Scores: SBA (1 point); Seven agencies (2 points)

OUR REVIEWERS COULD find no content on this subject in the SBA report. The other reports with weak scores tend to have little insightful content on improvement strategies for specific programmatic or management shortcomings and little to say about the broader mission-related challenges facing the agencies or how those challenges will be addressed. Some are among the agencies that seem reluctant to set challenging performance targets. The HUD report is an example. One striking feature is the report's lack of content on the current housing crisis. By contrast, the Treasury report highlights the department's actions and strategies to address the current financial crisis.

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 YEAR 2008 RANK ORDER

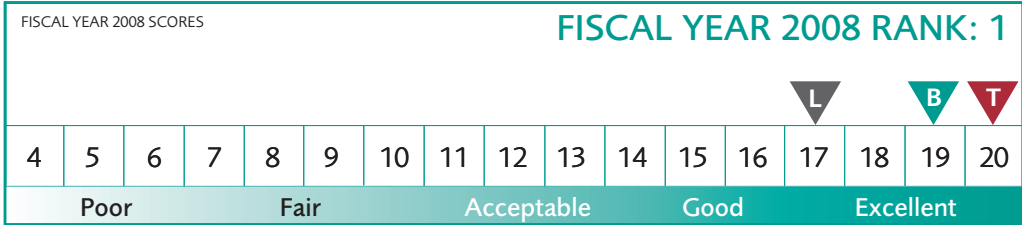
THIS SECTION SUMMARIZES the scores received by each agency in the three major scoring categories: Transparency, Public Benefits, and Leadership. Each agency summary appears on a separate page in rank order from highest to lowest. The graphic at the top of each page displays the scores each agency received in the three categories this year, fiscal year 2008. The graph at the bottom shows the rankings each agency has earned on the Scorecard for fiscal year 1999 through fiscal year 2008.

For example, Labor's report this year earned scores of 20, 19, and 17 on the Transparency, Public Benefits, and Leadership criteria respectively. The total of these scores, 56, gave this report the top ranking for fiscal year 2008.

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 one-page descriptions summarize the 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/ResearchAreaLanding.aspx?Category=74>.

DEPARTMENT OF LABOR



Total Score: 56 (out of a possible 60)

Transparency (T)

- The department’s home page has a direct link to the report, which is downloadable in a single file or multiple files. Report-specific contact information is in the report and online. The secretary’s transmittal letter highlights the department’s key achievements over time as well as fiscal year 2008 results for each strategic goal area.
- The accompanying citizens’ report is concise, well organized, and reader friendly.
- Baseline and trend data are shown for each measure. As a best practice, there is a “Program Perspective and Logic” discussion on performance for each goal.

Public Benefits (B)

- All four strategic goals and many performance goals are strongly outcome oriented.
- Performance measures also include some activity, output, and efficiency-type measures.
- Strong, outcome-oriented performance metrics superbly demonstrate contributions in a compelling way that an ordinary citizen can easily appreciate.
- The report links costs to strategic and performance goals, as well as to numerous individual performance measures for the current and two previous fiscal years.

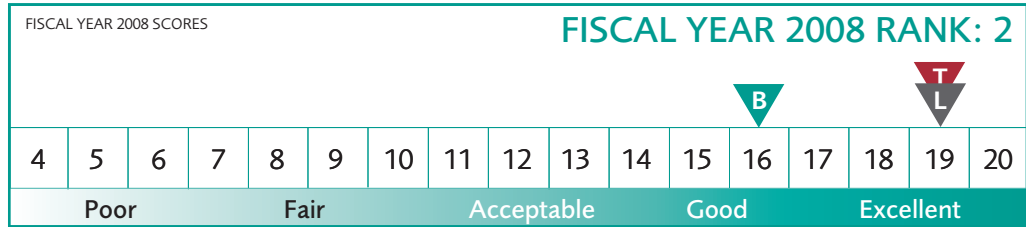
Leadership (L)

- The transmittal letter and narratives describe the public benefits that the department achieves.
- The report clearly discloses performance results, including shortfalls, and consistently offers explanations of shortfalls, many of which are insightful.
- The inspector general and department assessments provide confidence that the department is working seriously and effectively to address major management challenges.
- The report does an excellent job of describing future challenges that transcend individual goals and measures and how the department plans to meet them.
- The Department of Labor is a leader in developing innovations and improvements on many fronts with respect to performance accountability and reporting.

RANKING HISTORY



DEPARTMENT OF VETERANS AFFAIRS



Total Score: 54 (out of a possible 60)

Transparency (T)

- The report is posted via a prominent home page link, which includes previous year versions and multiple download options.
- The report is lengthy yet provides useful tables and is accompanied by a more concise and informative citizens' report.
- There is a detailed discussion on data quality and verification, as well as data limitations.
- The report includes considerable baseline and trend data, with discussions of performance trends and evidence of increasingly challenging performance targets from year to year.

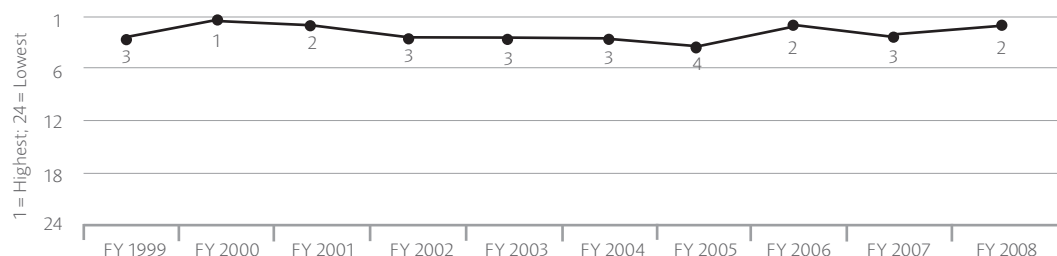
Public Benefits (B)

- Three out of four strategic goals and their related objectives are stated as end or intermediate outcomes.
- The most relevant performance measures are outcome oriented or capture efficiencies that are central to the department's missions.
- The report links budget costs to the department's strategic goals and objectives as well as to 7 of the 25 key performance measures.

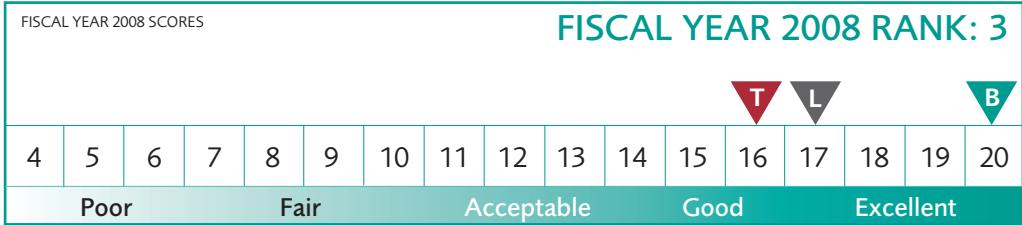
Leadership (L)

- The transmittal letter and report narratives do an excellent job of describing the department's accomplishments and the public benefits that flow from them. The narratives are backed up by solid performance metrics.
- The report contains a specific section titled "Performance Shortfall Analysis" that describes the causes of and the resolution strategies for significant performance shortfalls.
- The report has extensive and elaborate content on major management challenges and specific remedial actions for this fiscal year and the next, along with their anticipated impacts.
- The report thoroughly discusses performance and management shortfalls, is generally forward looking in content, and evidences continued enhancement to reporting formats.

RANKING HISTORY



DEPARTMENT OF TRANSPORTATION



Total Score: 53 (out of a possible 60)

Transparency (T)

- The report begins with an excellent transmittal letter that highlights major performance achievements in a results-oriented way and candidly highlights several concerns.
- The report conveys substantive information about the department's performance, but parts could be condensed and enhanced by greater use of summary tables.
- The transmittal letter lacks the statutorily required data assessment, but the body of the report provides detailed backup on data completeness and reliability.
- The performance section narratives add valuable context for assessing performance, but the department set targets below current performance levels for some key safety outcome measures.

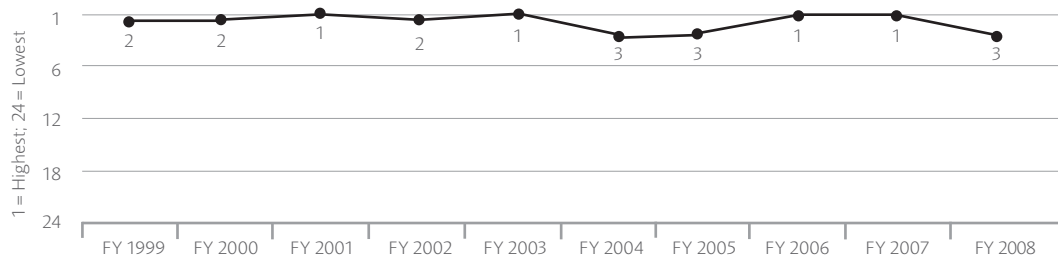
Public Benefits (B)

- The department's programmatic strategic goals are highly outcome oriented; most of their accompanying objectives also focus on end or intermediate outcomes.
- Most of 32 programmatic performance measures depict end or intermediate outcomes.
- The narratives in the performance section complement the highly results-oriented performance metrics and are highly effective in elaborating upon the department's performance results.

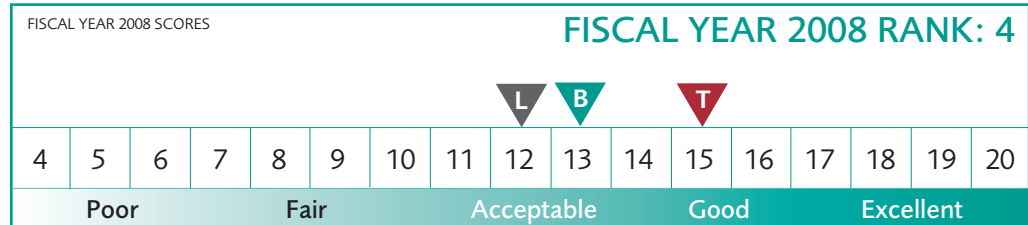
Leadership (L)

- Complementing the strong performance metrics, the report's narratives do an excellent job of demonstrating the key public benefits that flow from the department's work.
- The report consistently provides insightful commentary on performance results and describes improvement strategies even where targets were not missed.
- The inspector general's presentation on major management challenges is weak, but the department uses a "progress meter" to self-assess its steps to resolving each challenge.
- The report is thorough on improvement strategies, providing ample evidence that the department is forward looking and focused on enhancing its performance.

RANKING HISTORY



DEPARTMENT OF HOMELAND SECURITY (PILOT)



Total Score: 40 (out of a possible 60)

Transparency (T)

- A prominent link on the home page leads to the citizens' report, other relevant documents, and prior years' reports. There is contact information in the report but not online.
- The report is well organized, concise, and easy to read, focusing on 31 "highlighted" measures. However, some highlighted measures have limited value to the public because they use arcane terms or their results are classified.
- The narratives in the citizens' and performance reports show prior-year targets, but they could be expanded to provide more insight into performance trends and the extent of progress.

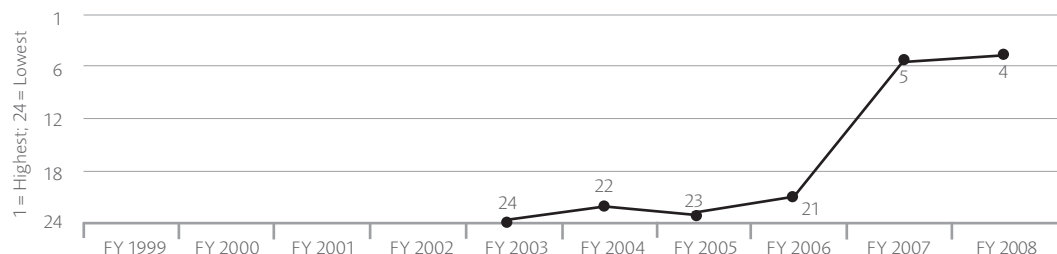
Public Benefits (B)

- The department's strategic goals and objectives are highly outcome oriented.
- Its measures are much less outcome oriented than the goals. Some measures would be more outcome oriented if expressed as percentages (e.g., proportion of secure national borders) instead of raw numbers (e.g., border miles under effective control).
- The narratives on performance results need to better explain the significance of the measures and performance trends relating to them.

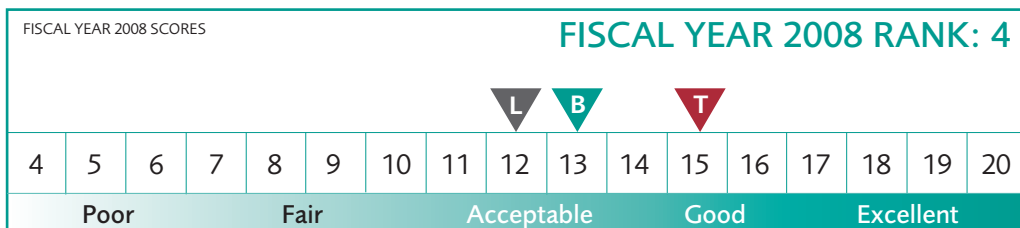
Leadership (L)

- The citizens' report and transmittal letter highlight public benefits from the department's work, but they could use more narrative content on results for the key measures.
- The detailed performance report systematically covers the department's results, including shortfalls; the citizens' report could use more content on results for the key measures.
- The financial report has extensive, specific content on major management challenges, including a strong presentation by the inspector general. It evidences a firm commitment to address the challenges, although the department has far to go to resolve most of them.
- The department's financial and performance reports describe improvement strategies to tackle specific management challenges and programmatic performance shortfalls well.

RANKING HISTORY



NUCLEAR REGULATORY COMMISSION



Total Score: 40 (out of a possible 60)

Transparency (T)

- The department's home page has a prominent direct link to the report. The report is downloadable in a single file or multiple files. There is no report-specific contact information online or in the report.
- The report is concise and its presentations are straightforward.
- The chairman's transmittal letter states that the report's financial and performance data are reliable and complete; the report contains substantial background on data.
- While some of the programmatic targets reflect zero tolerance for shortfalls, others are set significantly above current and historical performance levels.

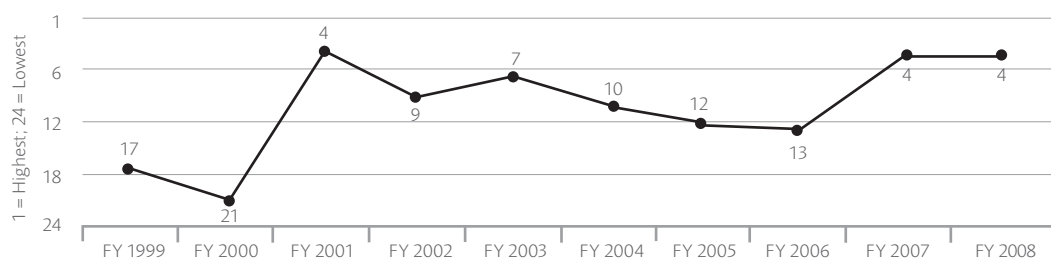
Public Benefits (B)

- The programmatic strategic goals and objectives are strongly outcome oriented but could be improved with the addition of intermediate outcomes.
- The safety and security performance measures are highly outcome oriented, and the operations and management measures focus on timeliness, efficiency, and customer service.
- The report links budgetary resources only to the programmatic strategic goals, yet indicates that the agency is working to better align costs with outcomes.

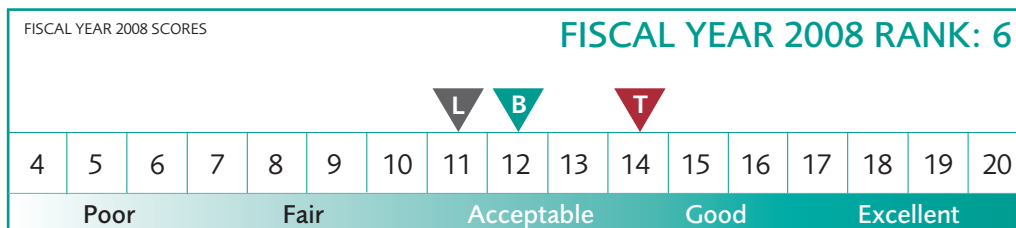
Leadership (L)

- The performance measures and narratives establish the public benefits the agency achieves, but the transmittal letter could be improved by highlighting some specific accomplishments.
- The report disclosed shortfalls and minimal explanations for operational and management metrics.
- The inspector general presents management challenges and indicates the agency is working on addressing them. The report could be more specific in assessing progress.
- The report discusses future performance challenges and ideas for improvement.

RANKING HISTORY



DEPARTMENT OF EDUCATION



Total Score: 37 (out of a possible 60)

Transparency (T)

- The home page links current and prior-year reports with report-specific contact information.
- The report is well formatted, clearly written, and effectively uses tables and graphics.
- The department faces major data challenges due to the long-term nature of its outcomes and its need to rely on external data, but the report indicates the department is working to ameliorate these challenges.
- The report contains best practice “target context” descriptions and good presentations on educational results and trends over many years, but has incomplete data for many measures.

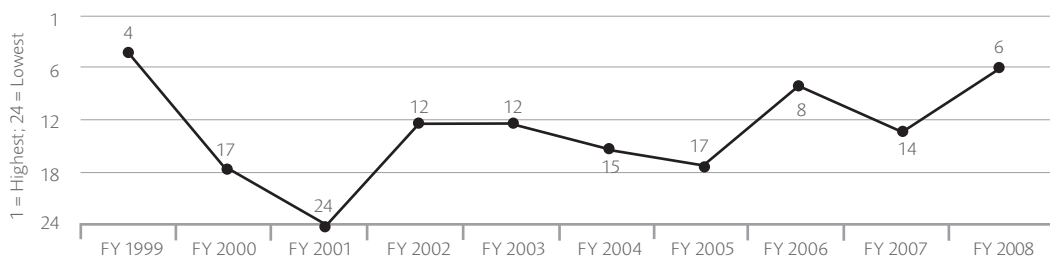
Public Benefits (B)

- The department’s strategic goals are new and much improved this year, clearly capturing end or intermediate outcomes of obvious public importance.
- Most measures are very outcome oriented, supporting outcome goals and objectives.
- The measures reported as raw numbers could be improved if they were restated as percentages of outcomes sought.
- The report does not allocate budget resources to the performance metrics, but explains why the department finds it difficult to link annual funds to performance results.

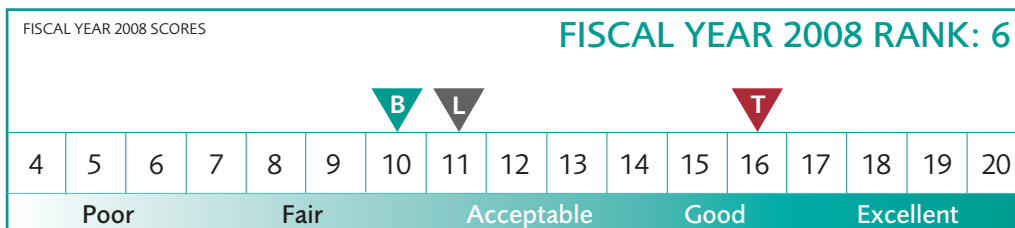
Leadership (L)

- The report’s narrative portions describe the benefits that flow from the department’s work. The secretary’s letter should highlight specific education-related results.
- The report shows performance shortfalls for three measures but gives weak explanations for the shortfalls.
- The inspector general’s concise presentation on major management challenges includes an assessment of the department’s progress on each of the six challenges and gives the department credit for taking positive steps to address them.
- The report has little content on changes to improve on specific performance shortfalls, perhaps because few results are reported; the narratives in the performance section do describe improvement strategies more generally.

RANKING HISTORY



DEPARTMENT OF THE INTERIOR



Total Score: 37 (out of a possible 60)

Transparency (T)

- The home page links directly to the report, which has multiple download options.
- Interior has more than 200 performance metrics; the report looks at 25 “representative” measures.
- The secretary’s substantive transmittal letter provides interesting background and highlights several specific accomplishments in each mission area.
- The citizens’ report is well organized and rich in vignettes, photographs, and graphics.
- The performance section concisely describes the department’s performance measurement system as well as its data validation and verification processes. It also includes baseline and trend data back to fiscal year 2005, but it lacks accompanying narratives for most of the measures.

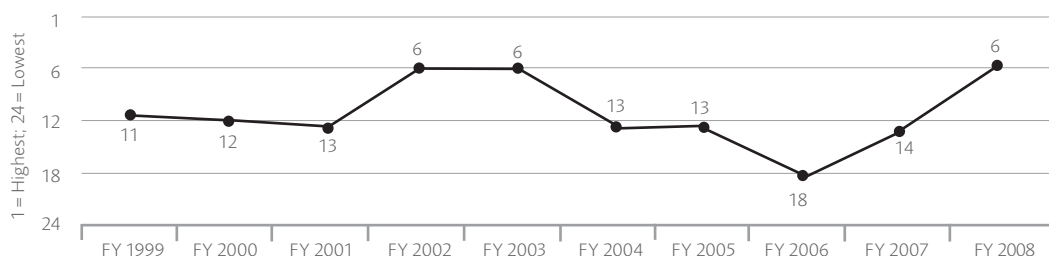
Public Benefits (B)

- Most of the strategic goals and objectives fall short of supplying clear, measurable outcomes and could be improved with increased specificity.
- As with last year, about half of the representative measures (13 of 25) are outcome oriented.
- The report allocates budget costs to the 25 representative performance measures and analyzes the impact of funding levels on performance for them. This is a potential best practice and should be carried over to the bulk of the department’s performance measures.

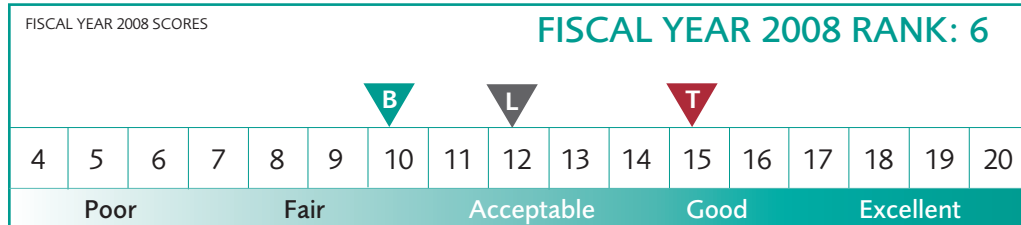
Leadership (L)

- The department’s performance metrics are limited in demonstrating public benefits, but the transmittal letter and report narratives show the public benefits from Interior’s work.
- The report clearly discloses results and explains performance shortfalls for key measures.
- The department’s progress in dealing with major management challenges is lacking, but the transmittal letter candidly highlights and addresses some serious management challenges.
- The department is improving performance reporting with innovative and useful techniques. The report provides insight into improvements for key measures but contains limited discussion for non-representative measures.

RANKING HISTORY



DEPARTMENT OF STATE (PILOT)



Total Score: 37 (out of a possible 60)

Transparency (T)

- The department's Web page containing the report had other relevant fiscal year 2008 documents and linked to prior-year versions of the report; there was no report-specific contact information online.
- The report is concise, well organized, and informative; it uses tables and graphics very well.
- The report has data source and verification and validation descriptions for its 18 measures.
- The prior-year data are limited since many measures date back only to fiscal year 2007. However, the report provides useful narrative information on the significance of the measures and performance trends.

Public Benefits (B)

- The report has elaborated on its strategic goals this year to describe the intended outcomes more clearly and with greater specificity.
- In contrast to its goals, the department's performance measures are less outcome oriented than last year. Most are raw number measures dealing with activities or outputs.
- The department's predominant output and activity measures do not provide a strong foundation to demonstrate contributions toward its outcome goals.
- The report links budget costs to the strategic goals and to the strategic objectives.

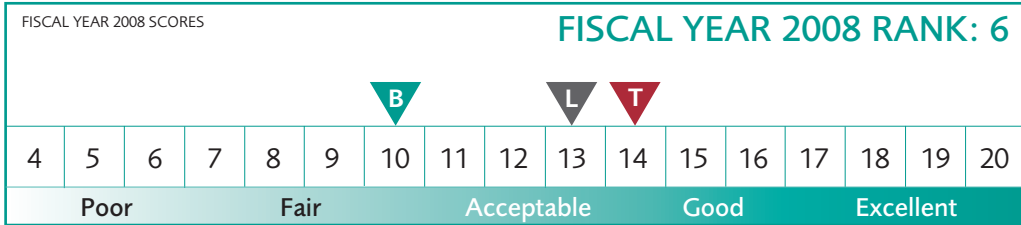
Leadership (L)

- The transmittal letter and the narrative portions of the report highlight a number of specific accomplishments and describe how they serve the public.
- The report provides brief explanations for missed measures, but their quality varies.
- The report summarizes major management challenges identified by the inspector general along with current and future remedial actions. The department improved this year from a disclaimer to an unqualified opinion on its financial statements, which means auditors found the financial statements fully and fairly represent the agency's financial condition.
- The report contains some descriptions of changes to address specific programmatic and management shortcomings and has some content on broader challenges.

RANKING HISTORY



DEPARTMENT OF THE TREASURY



Total Score: 37 (out of a possible 60)

Transparency (T)

- The department's home page does not link directly to the report. Previous year versions and report-specific contact information are available online.
- The report highlights the agency's actions and accomplishments, particularly with respect to the current financial crisis, and performance accomplishments by strategic goal.
- The report contains innovative features but has some readability challenges.
- The report gives a thorough presentation of baseline and trend data. The use of graphics to show trends for target achievements and actual performance over time is an innovative feature and a potential best practice.

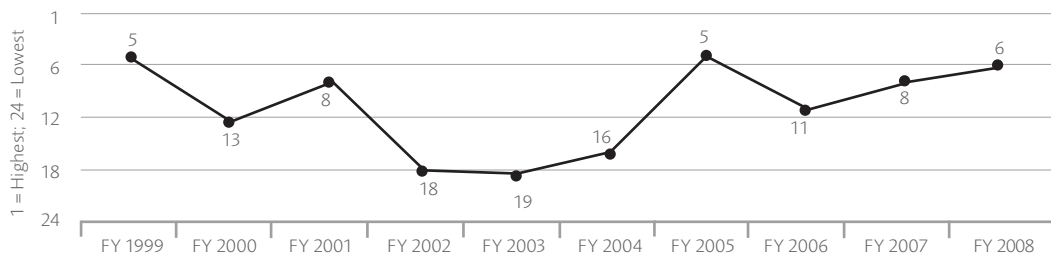
Public Benefits (B)

- The programmatic strategic goals and their related objectives are fairly outcome oriented in relation to the department's functions, but further specificity could improve them.
- Less than 20 percent of the performance measures clearly capture outcomes. The majority focus on efficiency, cost effectiveness, or customer satisfaction.
- The department's weak performance metrics, particularly the measures, have limited effectiveness in demonstrating public benefits.
- The report innovatively allocates performance costs to strategic goals, strategic objectives, and performance goals.

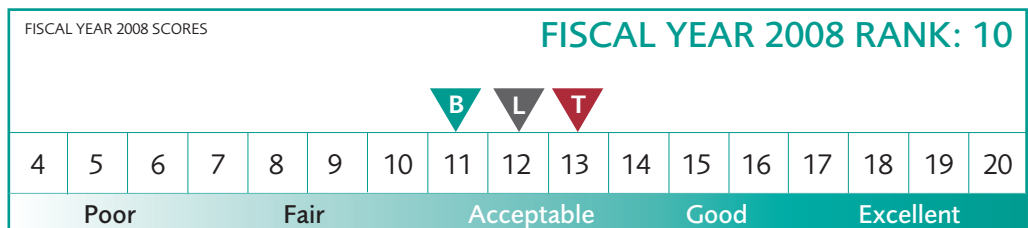
Leadership (L)

- The report's narrative portions describe the department's functions and their importance and achievements, with particular attention to the current financial crisis.
- The report discloses shortfalls and provides brief but informative explanations.
- The report does a thorough job of describing major management challenges and the remedial actions taking place and evinces some progress from the previous year's issues.
- The forward-looking report describes future programmatic and management improvements.

RANKING HISTORY



DEPARTMENT OF ENERGY (PILOT)



Total Score: 36 (out of a possible 60)

Transparency (T)

- The department's home page has a link to the current and prior-year reports. There is contact information in the citizens' report, but the information is not listed separately on the Web site.
- The report is visually appealing, well organized, and user friendly.
- The secretary's transmittal letter is substantive and assures the reader that the data are complete and reliable. It references the fiscal year 2008 performance report for further information on data.
- The citizens' report has prior-year data for only a handful of measures. The performance report has more prior-year data, but the presentations are often technical and confusing.

Public Benefits (B)

- The strategic goals and objectives are mainly outcome oriented, capturing clear outcomes of obvious public importance.
- The department has only a few outcome-oriented performance measures; most of the 220 measures are technical and activity oriented.
- The department's budget is linked to its programmatic strategic goals and also to the annual performance goals.

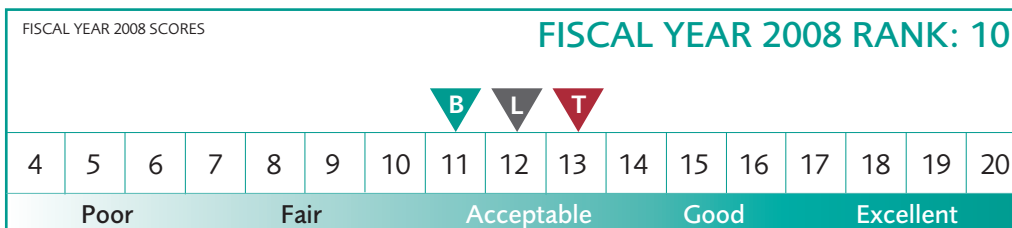
Leadership (L)

- The transmittal letter and other report narratives adequately describe the department's accomplishments and the public benefits they achieve.
- The citizens' report and detailed performance report clearly disclose performance shortfalls and offer explanations for them.
- The citizens' report summarizes major management challenges from the inspector general's presentation and the GAO.
- The report provides overviews of future challenges and strategies to address them. The performance and financial reports, respectively, discuss improvements for individual performance and management shortcomings.

RANKING HISTORY



ENVIRONMENTAL PROTECTION AGENCY



Total Score: 36 (out of a possible 60)

Transparency (T)

- The home page link produces the report, prior-year versions, and contact information.
- The inclusion of an e-mail address at the bottom of each page of the report for submitting comments or questions about the report is a useful and unique feature.
- The performance report is lengthy, technical, laborious to read, and could be greatly enhanced with more summary presentations.
- Although it was difficult to find online, the voluntary citizens' report is concise, visually appealing, and clearly written with ordinary citizens in mind.
- The administrator's transmittal letter is a candid and informative assessment of the agency's serious data challenges and time lags.

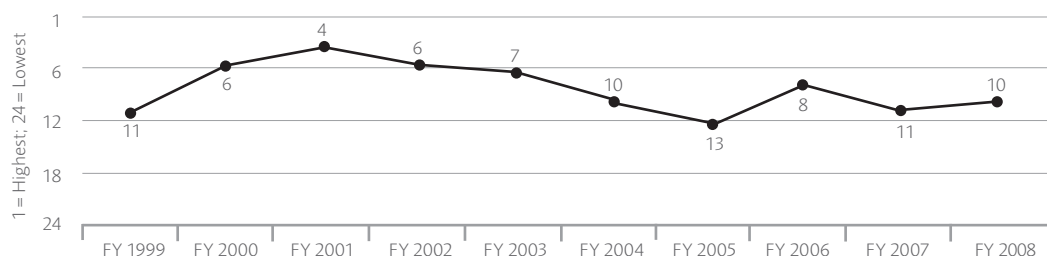
Public Benefits (B)

- The strategic goals capture outcomes of clear public importance. Some strategic objectives are outcome oriented; some are too abbreviated to indicate measurable outcomes.
- While many of the agency's performance measures are stated in outcome terms, they are often highly technical and difficult for a non-expert to comprehend.
- The report links budget costs to the agency's strategic goals and objectives but not to individual performance measures.

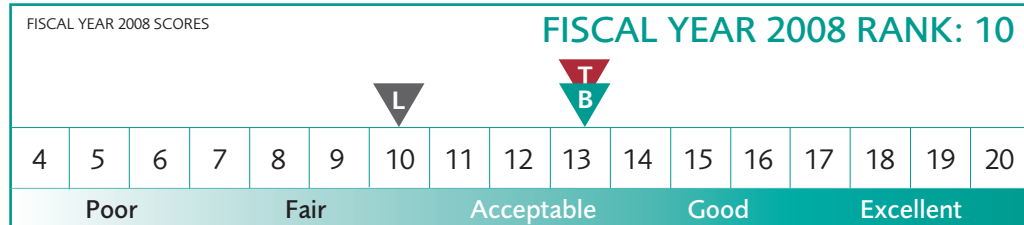
Leadership (L)

- The narratives in the report highlight the public benefits that flow from EPA's work.
- The performance section provides some explanations for missed measures although they generally are not very insightful and often lack any description of improvement steps.
- The report treats major management challenges comprehensively. It includes descriptions of corrective actions the agency is taking as well as time estimates for completion.
- Improvement actions for management challenges are more useful than specific programmatic shortfalls. Mission-related programmatic challenges are generally covered.

RANKING HISTORY



DEPARTMENT OF HEALTH & HUMAN SERVICES (PILOT)



Total Score: 36 (out of a possible 60)

Transparency (T)

- The department’s home page has a prominent link to the citizens’ report, related documents, and a contact for the office producing the report.
- The report is somewhat text heavy, and the graphics can be hard to interpret due to the absence of numbers to accompany the bars.
- The completeness of data is a major issue. Fiscal year 2008 results are available for well under half of the agency’s performance targets. The agency uses the most recent results to fill in the gaps.
- The report provides prior-year baseline and trend data generally going back to fiscal year 2005 as well as longer-term future targets for most measures.

Public Benefits (B)

- The strategic goals and objectives are, for the most part, clear and outcome oriented and capture public benefits of obvious public importance.
- Three-quarters of the 16 “spotlight” measures capture intermediate or end outcomes; 75 percent of the remaining 24 measures also appear to be at least intermediate outcomes.
- The department’s performance metrics are highly outcome oriented; data lags, narrow measures, and targets set below performance levels are negative features.
- The report links funding to each strategic goal and strategic objective.

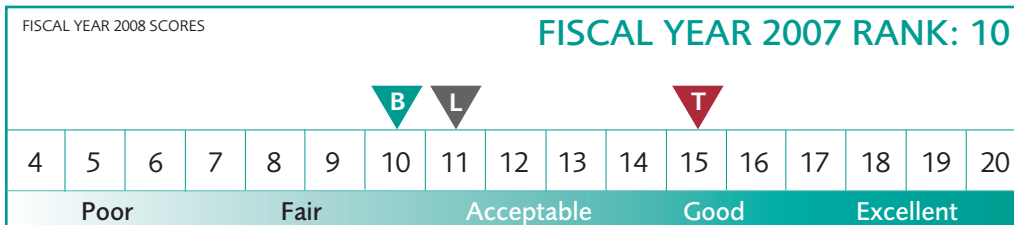
Leadership (L)

- The report effectively demonstrates the important public benefits of the department’s work. The secretary’s transmittal letter could highlight performance outcomes better.
- The report does not highlight performance shortfalls and explanations for them, so it can be hard to find these in the report. Moreover, once found, the explanations for missed measures tend to be weak.
- The report identifies management challenges addressed by the inspector general and provides an assessment of progress, a response by departmental management, and a description of future remedial plans.

RANKING HISTORY



U.S. AGENCY FOR INTERNATIONAL DEVELOPMENT (PILOT)



Total Score: 36 (out of a possible 60)

Transparency (T)

- The link to the report also includes an introductory video by the agency head. The hard copy has report-specific contact information, but the online version does not.
- The citizens' report is concise and informative; the stand-alone performance report details the agency's overall performance under each strategic goal and performance measure.
- The transmittal letter states the report's performance and financial data are complete and reliable. Though it raises some red flags about data quality, it provides no details.
- The report is formatted to show prior-year results, but not targets, back to fiscal year 2005, but less than half of the measures have results going back that far. The report does include preliminary targets going forward to fiscal year 2009.

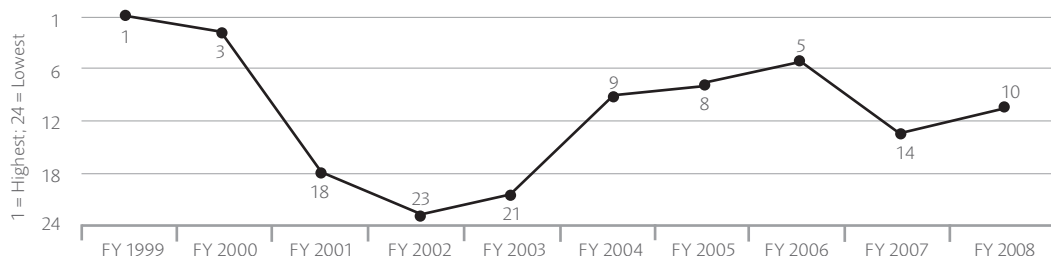
Public Benefits (B)

- The agency added text to better describe its strategic goals, but it lacks specific and measurable outcomes. Adding strategic objectives would help the agency flesh out its goals.
- One-third of the agency's measures appear to be outcomes, mainly intermediate ones. Raw number measures could be improved if the report expressed them as a percentage of the end achievement.
- The report links budget costs only to the strategic goals.

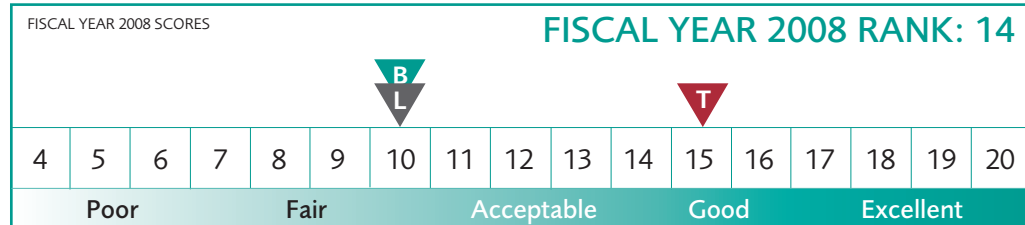
Leadership (L)

- The transmittal letter and narratives explain the public benefits of the agency's work.
- The report clearly discloses performance results, including shortfalls, but does not consistently provide explanations for missed targets.
- The report summarizes the major management challenges identified by the inspector general, the remedial actions taken, and future actions to take, but it does not methodically address the extent of progress.
- The report should have more content on specific performance shortfalls since USAID missed 60 percent of fiscal year 2008 targets. It is stronger in describing changes to address major management challenges and broader programmatic challenges.

RANKING HISTORY



DEPARTMENT OF COMMERCE



Total Score: 35 (out of a possible 60)

Transparency (T)

- A link from the home page produces the report in single or multiple files, along with prior reports, but a direct and report-specific link for contact information would be useful.
- The report is lengthy, but it is well formatted and makes good use of tables and graphics.
- The reader could better grasp the department's overall performance if the goals, measures, and results were brought together in a summary table. Less jargon would also improve readability.
- The appendix on performance measures shows baseline and prior years' data, but these data show that the department frequently set current targets well below prior-year performance levels.

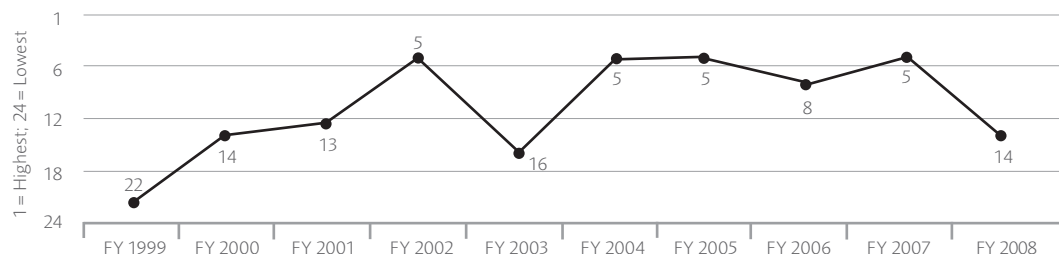
Public Benefits (B)

- The strategic goals and objectives need more specificity about intended outcomes.
- About one-third of the 42 programmatic performance measures seem outcome oriented. The rest are activity, output, efficiency, and customer-satisfaction measures.
- The department's performance metrics are slightly more outcome oriented than last year, but still leave much room for improvement.
- The report links budget costs to the department's strategic goals and objectives as well as to its annual performance goals.

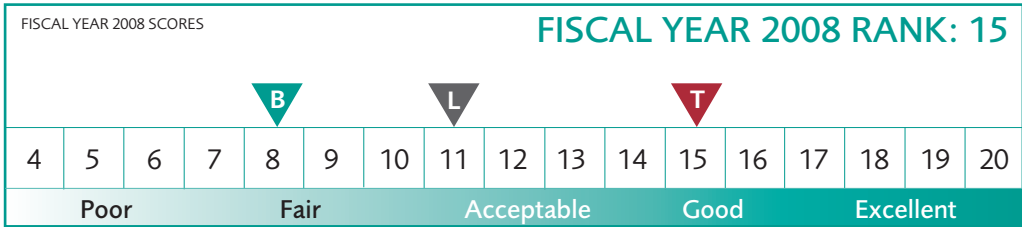
Leadership (L)

- The secretary's transmittal letter and the narrative portions of the report generally do a good job of describing the department's accomplishments.
- The report clearly discloses performance results, including shortfalls, but does not consistently explain the shortfalls.
- The inspector general's presentation on major management challenges and the department's response show little progress in addressing the challenges.

RANKING HISTORY



DEPARTMENT OF JUSTICE



Transparency (T)

- The department's home page has a prominent direct link to the report, which offers single file or multiple file downloads. There is no report-specific contact information in the report or online.
- The report is well formatted and lays out the department's performance results clearly.
- The department voluntarily produced an excellent citizens' report, which reads like a document specifically prepared for the general public.
- Performance baselines and trend data for each measure are provided in reader-friendly graphs. The information includes longer-term future goals as well.

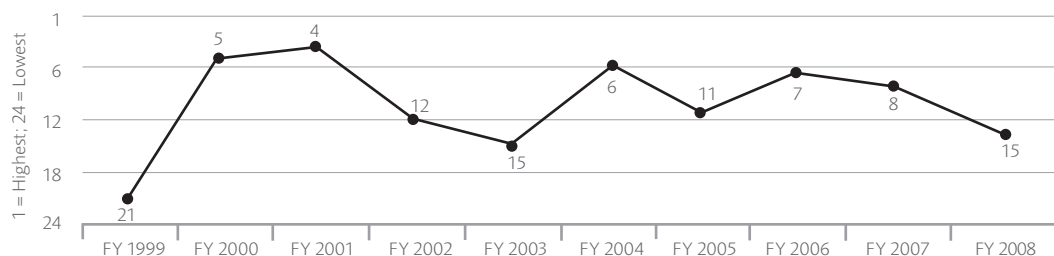
Public Benefits (B)

- All of the strategic goals and half the strategic objectives reference outcomes, but the goals would be more compelling if they were stated with greater specificity.
- The performance measures generally relate directly to the applicable goals and incorporate important outcomes. However, pass/fail metrics and raw-number targets do not provide insight to assess annual progress.
- The report links performance metrics to costs only at the strategic goal level.

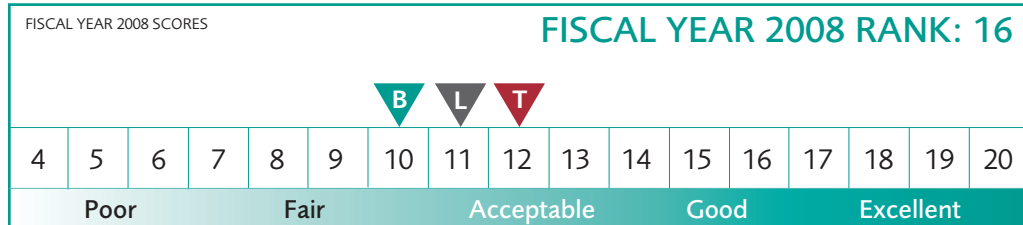
Leadership (L)

- The report narratives do a good job of describing the department's missions and the public benefits it seeks to achieve.
- Performance results, including shortfalls, are clearly disclosed and accompanied by summary explanations.
- The inspector general's presentation thoroughly describes management challenges but lacks insight on resolution progress. The department's response is more specific.
- The department's continued failure to develop measures related to illegal drugs is a concern.

RANKING HISTORY



DEPARTMENT OF AGRICULTURE



Total Score: 33 (out of a possible 60)

Transparency (T)

- A prominent link leads to the report, but the report is downloadable only in a large single file. There is no report-specific contact information in the report or online.
- The report is well organized and features a summary “performance scorecard.” It uses reader-friendly formats and tables to clarify the department’s performance results.
- The performance section provides data sources for measures and, along with the transmittal letter, provides an assessment of data completeness, reliability, and quality.
- The performance section tables include baseline, trend data, and narratives back to fiscal year 2004.

Public Benefits (B)

- The strategic goals are stated in outcome terms, but the outcomes are at such a high level that measurement and attribution to the department’s contributions pose challenges.
- The 34 annual performance goals mix outcomes and intermediate outcomes, activity and output, and efficiency measures.
- The department’s performance metrics would be more effective in demonstrating results if they provided more outcomes expressed with greater clarity for lay readers.
- The report links program obligations and staff years to the strategic goals only.

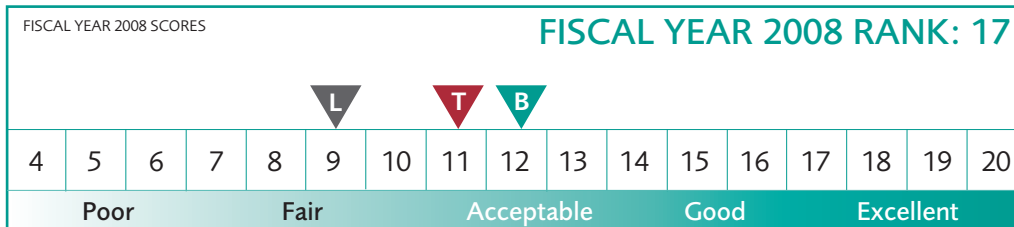
Leadership (L)

- The performance section narratives and transmittal letter provide a useful supplement to the performance metrics regarding the public benefits that flow from the department’s work.
- The report clearly discloses performance results, including shortfalls, but it offers weak explanations for the four missed measures.
- The inspector general’s presentation on major management challenges could be more concise and pointed, but it gives credit for the remedial actions the department is taking.
- Although the explanations for missed measures are weak, the report includes analyses of results for each measure and future challenges that are generally informative.

RANKING HISTORY



GENERAL SERVICES ADMINISTRATION



Total Score: 32 (out of a possible 60)

Transparency (T)

- The report is not directly linked on the home page and is downloadable only as a very large single file. There is no report-specific contact information online.
- The report seems verbose and redundant in parts, and the performance section is organized primarily by agency component rather than by strategic goal, making it somewhat disjointed for a lay reader. The citizens' report adds little value to the full report.
- The report provides little detail on performance data validation and verification. The reported results are incomplete, and the inspector general raises questions about the agency's data quality.
- The performance section shows baseline and trend data for the key measures from prior years, but it does not include prior-year targets. No results are given for any non-key measures.

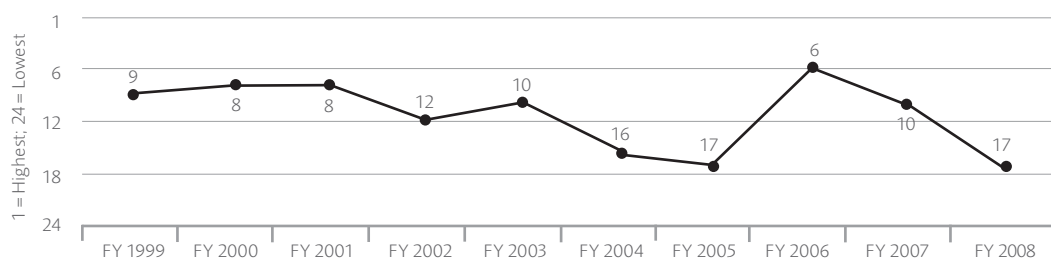
Public Benefits (B)

- The report has the same four strategic goals as last year, but they are reasonably outcome oriented in terms of the value GSA offers American taxpayers, especially given GSA's largely intra-governmental functions.
- About half of the 16 key measures focus on cost savings or public service enhancements that might be regarded as outcomes or intermediate outcomes of direct public relevance.
- The report links agency's key and non-key budget costs to individual performance measures.

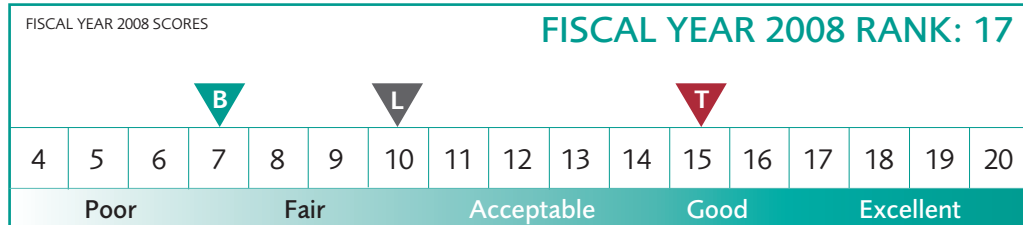
Leadership (L)

- Narratives describe the agency's operations and their benefits to the public.
- The report is incomplete since it discloses fiscal year 2008 results for only 16 of over 90 performance measures. It clearly discloses key measure results, including shortfalls.
- It is difficult to determine whether the agency is adequately addressing its major management challenges. The agency disagrees with the inspector general on most points.
- The report provides little insight on improvement plans for specific program performance or management shortcomings.

RANKING HISTORY



NATIONAL SCIENCE FOUNDATION (PILOT)



Total Score: 32 (out of a possible 60)

Transparency (T)

- The agency's home page links to the citizens' report, related documents, and to prior-year reports.
- The report is brief, easy to read, and makes good use of tables.
- NSF's use of a qualitative, instead of quantitative, system to measure program performance provides no baseline or other prior-year data that would enable the reader to assess performance trends over time or put the agency's performance in context.

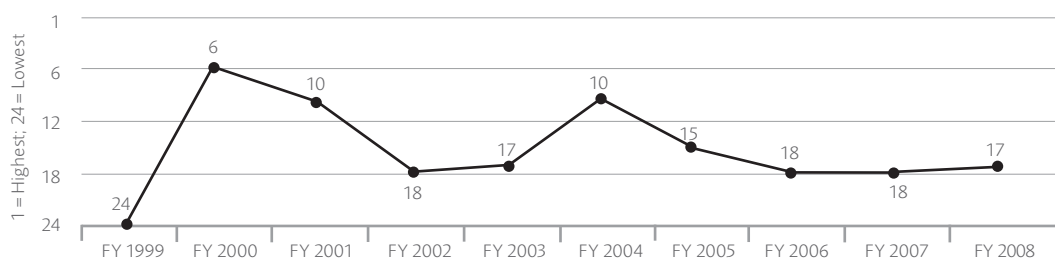
Public Benefits (B)

- Strategic goals are stated as high-level outcomes, but they lack annual goals and outcomes.
- Performance measures, where present, focus on processes and activities rather than results.
- The agency's performance metrics are not geared to demonstrate results to the general public.
- The report links costs only at the strategic goal level.

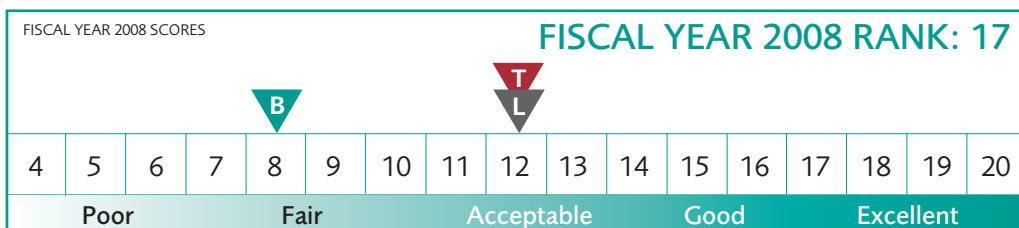
Leadership (L)

- The report refers and links to a Web site that describes additional project accomplishments. More examples of this nature would enhance the report's value to lay readers.
- The citizens' report indicates one performance target shortfall, but does not explain or describe it.
- The inspector general's presentation and the agency's responses detail management challenges and current and future remedial actions.
- The report has some improvement strategies for management but no discernable content on changes to improve program performance.

RANKING HISTORY



SOCIAL SECURITY ADMINISTRATION



Total Score: 32 (out of a possible 60)

Transparency (T)

- The report has multiple download options, but no report-specific contact information online or in the hard copy. A series of links provides prior-year reports back to fiscal year 1998.
- The report is reasonably concise by federal reporting standards and is well formatted.
- Six of the 26 performance measures are missing fiscal year 2008 results due to data lags, and an inspector general report's audits found data unreliable for 6 of 8 performance measures.
- The baseline and trend data are limited since many of the measures are fairly new. The agency does not set ambitious targets for many measures although narratives do explain the reasons for some low targets.

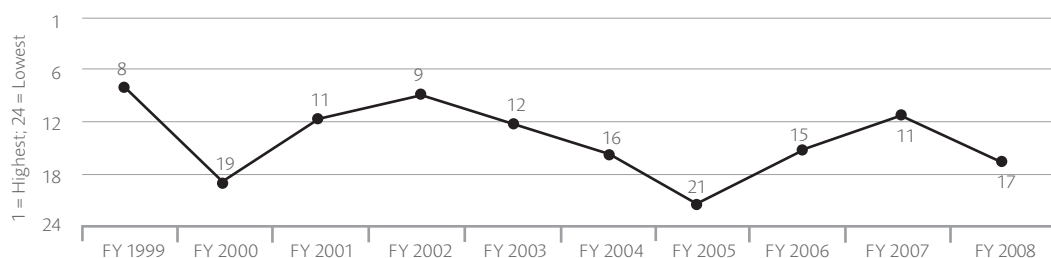
Public Benefits (B)

- The agency's current goals and objectives reference some outcomes but they could be refined to specify more precisely the intended results directly attributable to the agency's work.
- More of the agency's 26 performance measures are outcome oriented than last year, but some measures that are expressed as raw numbers would be more outcome oriented if they had been stated as percentages.
- Performance is deteriorating in some key areas even though the agency is meeting its targets.
- The report does not link budget costs to the agency's performance metrics.

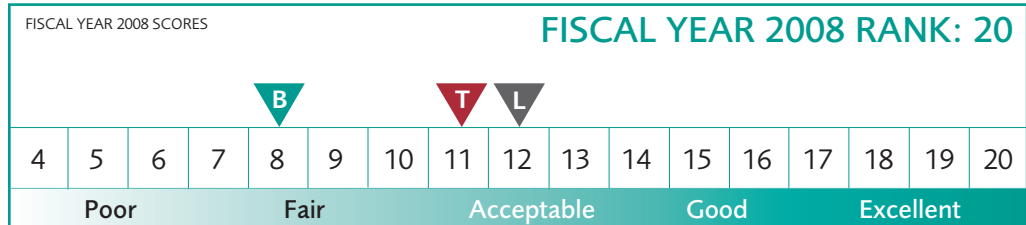
Leadership (L)

- The narratives are stronger than the performance metrics in explaining public benefit successes.
- The report clearly discloses fiscal year 2008 performance results, including shortfalls, attributing shortfalls to high workloads and inadequate staff.
- The inspector general's presentation on major management challenges does not convey how much progress the agency is making or when it might resolve the challenges.
- The forward-looking narratives describe plans to improve performance for individual measures even where targets were met.

RANKING HISTORY



NATIONAL AERONAUTICS & SPACE ADMINISTRATION



Total Score: 31 (out of a possible 60)

Transparency (T)

- Although bookmarks permit online navigation through various sections, the report is downloadable as a single file only. There is report-specific contact information in the report.
- The Management’s Discussion and Analysis section provides the public with a useful overview of the agency’s undertakings, but the rest of the report is less useful to a lay reader.
- The transmittal letter provides “reasonable assurance” that the report’s performance data are complete and reliable; the body of the report offers little backup for this assurance.
- The report offers little baseline or trend data to assist the reader in assessing progress over time since many measures are judgmental rather than quantifiable.

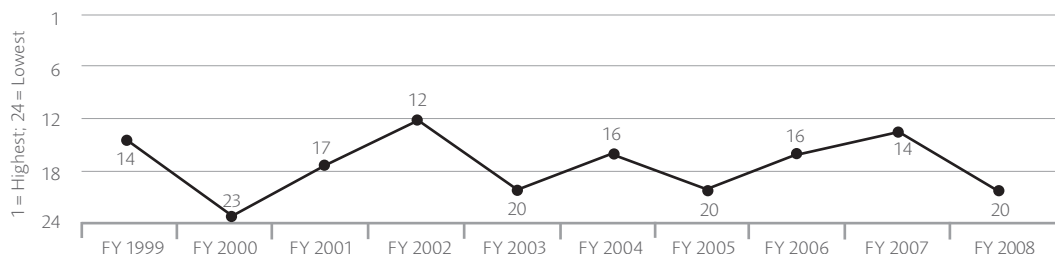
Public Benefits (B)

- The agency’s strategic goals are a mix of outcomes and activities.
- Most performance measures address highly technical activities.
- The performance metrics and assessment system afford little basis for the public to evaluate the agency’s progress from year to year or to appreciate its accomplishments.
- Efforts to explain the link between budget costs and goals are confusing, and the performance section shows cost linkages only to the strategic goals and subgoals.

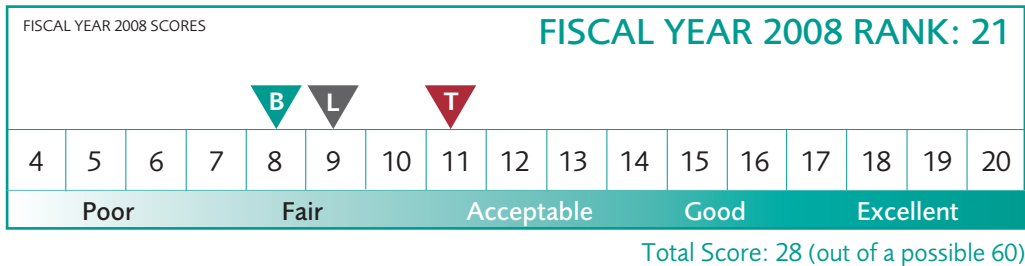
Leadership (L)

- The transmittal letter and performance highlights usefully describe NASA’s missions and accomplishments in a high-level way that a lay reader can understand.
- In what is a best practice, a table in the report lists each measure that was not fully achieved, explains why, and outlines plans for achieving the measure in the future.
- The inspector general’s presentation on management challenges does not methodically assess agency’s progress on the challenges, and the report shows little progress in key management areas.
- On the programmatic side, the report does a good job of describing changes for next year, but it is weaker on the management side.

RANKING HISTORY



OFFICE OF PERSONNEL MANAGEMENT (PILOT)



Transparency (T)

- There is a prominent direct link on the home page to the citizens' report and the fiscal year 2008 performance report, as well as prior-year versions. No report-specific contact information is online or within the report.
- The citizens' report is concise and makes good use of tables, but performance narratives are found only in the accompanying performance report.
- The citizens' report provides prior-year results for performance measures, but no narratives to explain the results. Some narratives in the performance report are confusing.

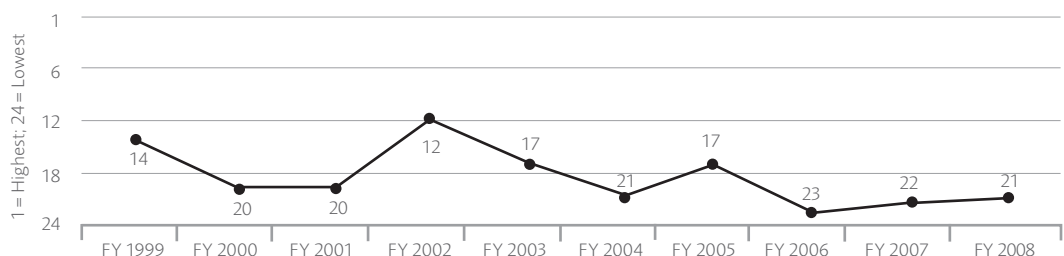
Public Benefits (B)

- The department's strategic goals and performance measures lack clear, measurable outcomes.
- Performance metrics do little to demonstrate public benefit contributions or indicate achievement of agency missions.
- The citizens' report links budget resources only to strategic goals while the performance report includes some links to strategic objectives.

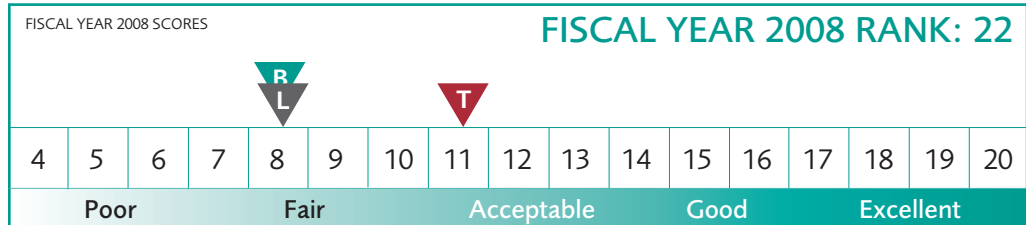
Leadership (L)

- Confusing narratives and weak performance metrics do not sufficiently demonstrate accomplishments and results.
- The citizens' report discloses but does not adequately explain performance shortfalls.
- The report lists major management challenges identified by the agency's inspector general and self-assesses the agency's progress in addressing them.
- The report contains improvement strategies to address management challenges but has little clear and insightful content on program performance improvement.

RANKING HISTORY



DEPARTMENT OF HOUSING & URBAN DEVELOPMENT



Total Score: 27 (out of a possible 60)

Transparency (T)

- The department's home page has a direct link to the report, which can be downloaded as a single file or multiple files.
- The performance report is lengthy and text heavy; the voluntarily produced citizens' report is much more concise and reader friendly.
- The secretary's transmittal letter states that the report's data are complete and reliable, but the contents of the report indicate otherwise.
- The department frequently sets non-challenging targets that aspire to performance levels lower than those of prior years.

Public Benefits (B)

- Two-thirds of the programmatic strategic goals and one-third of the cross-cutting strategic goals are clearly stated as outcomes.
- Most of the department's performance measures deal with outputs, activities, or efficiencies; few are outcome measures.
- Budget resources are linked to the performance metrics at the strategic goal level only.

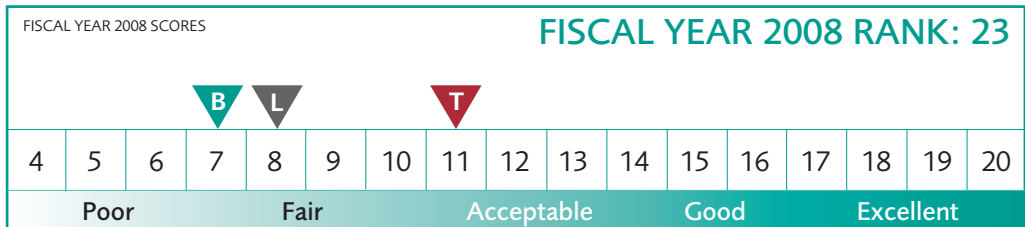
Leadership (L)

- The narratives in the report convey some useful information about the department's accomplishments, yet do not compensate for the department's weak performance metrics.
- The report clearly discloses performance shortfalls and offers narrative explanations for missed targets and future improvement plans. However, the information is rarely insightful.
- The inspector general's presentation on management challenges provides little specific assessment of progress.
- The report has some discussion of improvement strategies for programmatic shortfalls and management challenges, but it shows little ambition to enhance performance.

RANKING HISTORY



DEPARTMENT OF DEFENSE (PILOT)



Total Score: 26 (out of a possible 60)

Transparency (T)

- The department's home page has a prominent direct link to the citizens' report, but the department's fiscal year 2008 performance report was not posted at the time of our evaluation. No link was found for reports from previous years.
- The citizens' report was concise and easy to read, but had sparse performance content.
- There was no content on data sources or quality, and the department continues to have a disclaimer of opinion on its financial statements.
- Very little baseline or trend data are included.

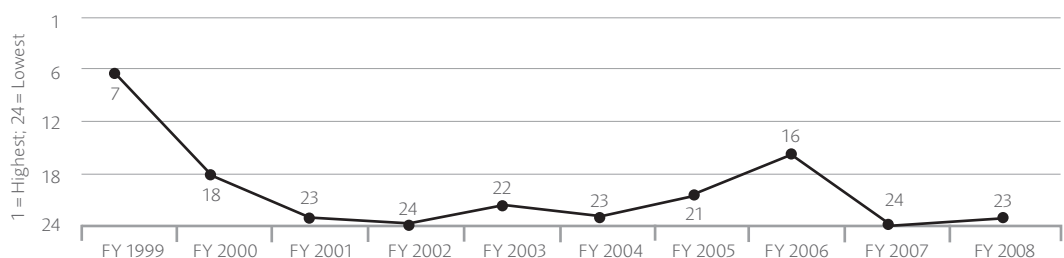
Public Benefits (B)

- No strategic goals are outcome oriented, and less than half of the strategic objectives have intermediate or end outcomes.
- Few of the 16 key performance measures capture outcomes. Most are raw numbers or efficiency measures.
- The performance metrics presented in the report are limited in scope and weak in outcome orientation.
- The report has no content linking performance to costs.

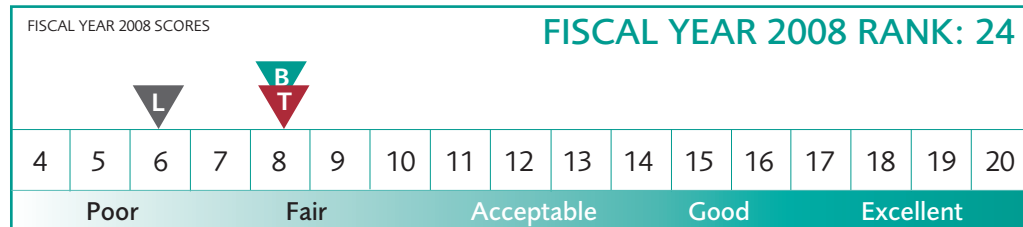
Leadership (L)

- The report narratives and performance metrics do not convey results and public benefits.
- The results for the 16 key measures are clearly disclosed, including performance shortfalls, but the shortfalls lack specific explanations.
- The citizens' report references the financial report for the inspector general's presentation on management challenges as well as the department's response. A number of steps are being taken to address the management challenges, but the challenges remain pervasive.
- While the report describes management-related improvements, there is virtually no content on changes to improve program performance.

RANKING HISTORY



SMALL BUSINESS ADMINISTRATION (PILOT)



Total Score: 22 (out of a possible 60)

Transparency (T)

- The citizens' report and performance report were not posted online in a timely manner.
- The report is brief and easy to read but falls far short of providing the overview of agency performance that a citizens' report is supposed to convey.
- The report has no transmittal letter, but the transmittal letter for the financial report asserts that the financial and performance data in that report are reliable and complete.
- The report shows prior-year results for the few performance measures it covers, but there is no accompanying explanation of performance results or trends.

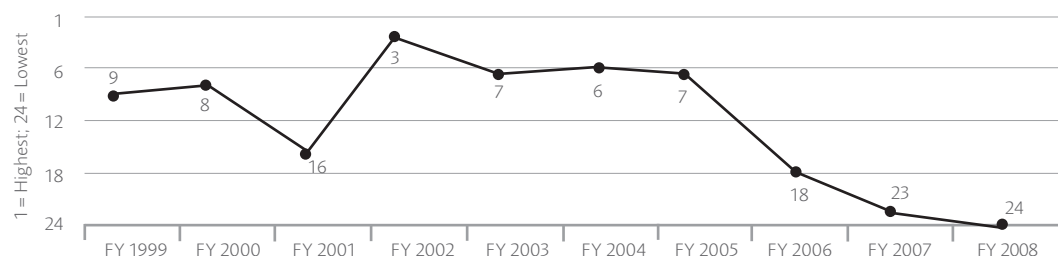
Public Benefits (B)

- Two of the agency's three programmatic strategic goals are outcomes, but they are at a high level of generality and are sometimes vague; the third is more specific but is stated in activity rather than outcome terms.
- The report describes most measures as "outcomes," but most are not.
- The report has little content about the agency's performance results or accomplishments.
- The report links the agency's budget costs to its strategic goals but not to lower levels of its performance metrics.

Leadership (L)

- The report clearly discloses fiscal year 2008 results, including shortfalls, for the few listed measures. There are no explanations for the shortfalls shown.
- The inspector general's presentation on major management challenges is one of the most informative, but the report makes no mention of it.
- The report has no content on changes in policies and procedures to do better next year.

RANKING HISTORY



APPENDIX: DETAILED EXPLANATION OF EVALUATION AND SCORING

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 outcomes the agencies produce 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.

TARGET AUDIENCE: THE GENERAL PUBLIC

A PERFORMANCE REPORT necessarily addresses different audiences. Some are “stakeholders” with expertise in the agency’s work, who 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. Others may be government insiders who know how to ferret out information that may be tucked away in congressional budget justifications or other esoteric documents.

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 do 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. Of course, we do not expect tens of millions of fellow citizens to rush to agency Web pages to read these reports. Journalists, bloggers, and other writers can also play an important role in making agency results more widely accessible to the public. But like the general public, these readers are not agency insiders. The information should be accessible and understandable for those who wish to examine it.

Reports that score high on our evaluation effectively 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 low do little to inform ordinary members of the public about important outcomes. Reports ranking in the middle may serve some audiences well but 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 do the following:

- use clear, concise presentation formats and language throughout that a lay person can follow and understand
- present a set of performance metrics that capture 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
- match outcomes with costs, so decision makers and stakeholders can assess tradeoffs
- provide confidence that the agency has adopted challenging measures, forthrightly acknowledge performance shortfalls, and take steps to correct them
- 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 the 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 as such. The reports should also clearly present the costs of achieving those results. These costs include both the agency's expenditures and other costs borne by the public at large. For regulatory agencies, much of this information should come from rigorous retrospective analysis of the actual effects of regulation, analogous to the information they are supposed to produce when conducting a Regulatory Impact Analysis prior to issuing a regulation. 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. 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. An agency report must highlight achievement of results; otherwise, it will not inform the public of the success or failure of its programs. Budget decisions that rely on 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 report should inspire confidence in an agency's ability to enhance citizens' quality of life commensurate with the resources they have entrusted to the agency. Among the factors that give such confidence is tangible evidence that the agency is using performance and financial data to improve management of its programs.

EVALUATING REPORTS UNDER THE PILOT FORMAT

AGAIN IN FISCAL year 2008, OMB allowed agencies to use an alternative reporting format under a pilot program. Instead of producing a performance and accountability report that combines performance and financial information, agencies had the option of publishing detailed performance information along with their congressional budget justifications, producing a separate financial report, and producing a shorter “citizens’ report” intended to summarize both the performance and financial results. Nine agencies opted for the pilot approach. Agencies producing traditional performance and accountability reports were also urged to produce a summary citizens’ report; 11 did so.

The research team evaluated the pilot reports for fiscal year 2008 the same way it evaluated them for fiscal year 2007. We started with the premise that our evaluation is intended to gauge the usefulness of these reports to the general public. The pilot format does not in any way absolve agencies of the responsibility to produce meaningful information for the public.

For agencies participating in the pilot program, the citizens’ report is the primary document available and accessible to the public that explains performance and financial information. We recognized, however, that the citizens’ report might not contain as much detailed information as a traditional performance and accountability report. Therefore, for pilot agencies, we included information from other documents if the citizens’ report clearly indicated where and how the reader could find this information and presented it in an accessible and understandable form.

We treated the citizens’ report differently for agencies that produced traditional performance and accountability reports. For those agencies, the performance and accountability report should still be the primary means of communication with the public. After evaluating this report, we then examined the agency’s citizens’ report to see if anything in that document merited a higher score. The principal way a citizens’ report might supplement a traditional performance and accountability report would be by improving readability. In fact, several agencies did receive extra points on our readability criterion due to the quality of their citizens’ reports.

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

AS IN PAST years, the Mercatus Center notified federal agencies of the deadlines we would follow in evaluating the reports and the evaluation criteria we would employ.

For fiscal year 2008, the Office of Management and Budget required agencies (except those participating in the pilot for alternative performance reporting) to submit their reports to the president and Congress by November 17—approximately six weeks after the fiscal year ended. OMB required those agencies participating in the pilot program to release their citizens’ report (the primary document the Mercatus research team evaluated for agencies opting into the pilot) by January 15, 2009.²⁸

In early October, the Mercatus Center notified each agency’s chief financial officer via letter (and other individuals listed as agency contacts by e-mail) that the Mercatus research team would regard a report as timely

28. See OMB Circular A-11, part 6, section 200, for these deadlines.

if it was posted on the Web by December 1 (for PAR agencies) or January 19 (for pilot agencies).

The October letter included an explanation of our evaluation process and criteria. It also referred agencies to a May 22, 2007 memo, detailing what the Mercatus research team would be looking for when reviewing citizens' reports. Finally, the letter invited agency personnel to contact Mercatus Center staff with questions or comments about the criteria and the project, and many did so. Thus, agencies had ample notice about the criteria and deadlines.

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 poor to excellent) were awarded on 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 factors is intended to identify 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 potential best practice, but best practices should not be confused with perfection. We expect agency reporting practices to improve continually over time, and one of the goals of this Scorecard is to aid in the diffusion of best prac-

WHAT DO THE SCORES MEAN?

5—OUTSTANDING

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

4—VERY GOOD

- Exceeds expectations
- Has the potential to become 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 standards for adequate disclosure
- Shows no process or plans to overcome problems
- Omits critical information

tices 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 separately (transparency, public benefits, and forward-looking leadership), readers who believe that one factor is more important than another 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 directly measurable services, such as the General Services Administration and the Department of Veterans Affairs, might find it easier to identify and quantify their contributions than an agency like the State Department or the National Science Foundation, whose results are less tangible. In reality, some agencies that seem to provide few services directly to members of the public have often produced highly ranked reports, and some that arguably have a more direct effect on citizens' well-being have produced poorly 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 produced the most or best results for the country. In fact, an agency can score well if it is performing poorly but reporting transparently. A high-scoring report reflects an agency's ability to explain its results in understandable and meaningful terms 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 are trying to evaluate whether the agencies are supplying the cost and performance information that the public needs to understand the agencies' work.

An additional word on information quality is also in order. Our researchers 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

THE MERCATUS RESEARCH team employed the same criteria to assess the fiscal 2008 agency reports that we used to evaluate prior-year reports. However, generally in each succeeding year, we have tightened our evaluation standards for two reasons. One, the highest possible quality that can be achieved is unlimited because creative innovators can always find ways to improve reporting practices and set new standards. Two, each

year gives agencies an opportunity to learn from each other's 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 year 2008 report in order to receive the same numeric score it received for its fiscal year 2007 report. That said, there were few new best practices in fiscal year 2008, so the standards did not change a great deal. Nevertheless, 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. Two members of the Mercatus Center research team that evaluated the reports—Henry Wray and Valerie Richardson—have done so for the past six years. Patricia Kelley, who joined the research team in fiscal year 2007, has been on the advisory panel to the Scorecard project since 2002. Next, the team cross-checked the 2008 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 then reviewed this documentation for both the fiscal year 2008 and fiscal year 2007 reports to ensure that any scoring differences across years were justified by differences in the actual content of the reports. The team then 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.

RESEARCH TEAM AND PROJECT DESIGN

THE SCORECARD PROJECT is headed by the Honorable 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. An 11-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

Heather Hambleton

*Program Associate, Government Accountability Project
Mercatus Center at George Mason University*

RESEARCH TEAM

Patricia Kelley

*Hodges University
Naples, Florida*

AFTER SERVING ON our advisory panel since 2002, Dr. Patricia Kelley joined the research team for the 9th Annual Performance Report Scorecard. She has extensive strategic planning, program evaluation, and operational management experience. In 2000 she co-founded a private consulting firm, Denali Associates, which provided a full range of management and IT consulting services to over 30 federal agencies, including support for the implementation of the Government Performance and Results Act and the Clinger Cohen Act. Prior to establishing Denali Associates, Dr. Kelley held senior management positions with the Federal Reserve Board, advising the governors on policy issues regarding efficiency and effectiveness of the board's operations. She also 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.

She recently co-authored the book *The Business of Government: Strategy, Implementation, and Results* with Dr. Thomas Kessler and currently serves on the faculty of Hodges University teaching graduate course in public administration.

Dr. Kelley has a doctorate in public administration from the University of Baltimore, an MA in computer systems management from the University of Maryland University College, and a BS in accounting from the University of Maryland at College Park. She is also a graduate of the Stonier School of Banking.

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 was 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 an MA 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.

Christina Forsberg

*Mercatus Graduate Fellow
George Mason University
Arlington, Virginia*

CHRISTINA FORSBERG IS a graduate student fellow at the Mercatus Center working with the Government Accountability Project and the Regulatory Studies Program. She is working toward her master’s degree in economics at George Mason University.

Stefanie Haeffele-Balch

*Mercatus Graduate Fellow
George Mason University
Arlington, Virginia*

STEFANIE HAEFFELE-BALCH IS a graduate student fellow at the Mercatus Center working with the Government Accountability Project and the Regulatory Studies Program. She is working toward her master’s degree in economics at George Mason University.

ADVISORY PANEL

Jonathan D. Breul

*IBM Global Business Services
Washington, DC*

JONATHAN D. BREUL is executive director of IBM's Center for the Business of Government and a partner in IBM Global Business Services. 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 Organisation for Economic Co-operation and Development's Public Management Committee.

Mr. Breul is a fellow of the National Academy of Public Administration and an adjunct professor at Georgetown University's Public Policy Institute. He holds an MPA from Northeastern University and a BA from Colby College.

Veronica Campbell

*Independent Consultant
Falls Church, Virginia*

VERONICA CAMPBELL IS currently consulting on performance, financial management, and general management issues after retiring in 2003 from 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, she managed the Department of Labor's strategic planning and performance-based initiatives from 2000 to 2003, working with the department's top executives and a dedicated inter agency committee to foster a results-driven organizational culture. Ms. Campbell joined the planning office after participating on the department's Year 2000 Project Management Team, where she served 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 the inspectors general 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, where 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 the USDA's international programs in 20 countries and evaluated a wide array of domestic programs.

Ms. Campbell holds a BA in history from Barat College.

John Kamensky

*IBM Global Business Services
Washington, DC*

MR. KAMENSKY IS a senior research fellow for the IBM Center for the Business of Government and an associate partner with IBM Global Business Services. 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, customer focused, and collaborative in nature.

Mr. Kamensky served eight years as deputy director of Vice President Al 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 an MPA from the Lyndon B. Johnson School of Public Affairs at the University of Texas.

Thomas Kessler

Central Michigan University

FROM 1983 THROUGH 1996, Dr. Kessler was 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 served as assistant director for systems and programming 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; the Commodity Futures Trading Commission; and the Federal Aviation Administration. In 2008 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 his doctor of business administration degree from Nova Southeastern University, an MBA from the University of Baltimore, and is a certified information systems auditor.

Sarah E. Nutter

*George Mason University
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. Dr. Nutter teaches a variety of courses in accounting and taxation in undergraduate, MBA, and executive MBA programs and has received a teaching award for Outstanding Faculty Member 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*, *Statistics of Income Bulletin*, and *Advances in Taxation*. One of her articles was granted the 1998–99 American Taxation Association’s Outstanding Manuscript Award.

Dr. Nutter earned a BS from Ferris State University and an MBA and PhD from Michigan State University.

John M. Palguta

*Partnership for Public Service
Washington, DC*

AS VICE PRESIDENT for policy at the Partnership for Public Service, John Palguta has responsibility for the development and implementation of a comprehensive program of review 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 BA in sociology from California State University at Northridge and an MPA 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 (IPMA-HR). 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 also an affiliated professor at Georgetown University’s Public Policy Institute. He is a recipient of MSPB’s highest honor, the Theodore Roosevelt Award, and the 2006 Warner W. Stockberger award, which is the highest honor presented annually by IPMA-HR to recognize an individual who has made outstanding contributions in the field of public-sector HR management.

Paul L. Posner

*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 with 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 the 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* (Georgetown University Press, 1998). He received his PhD in political science from Columbia University and was elected a fellow in the National Academy of Public Administration in 1996 where he chairs their Federal Systems Panel. Prior to his current position with George Mason University, Dr. Posner was a senior adjunct lecturer at Johns Hopkins, Georgetown,

and George Washington universities. He has received the James Blum Award for outstanding budget practitioner from the Association of Budget and Program Analysts. He is president of the American Society for Public Administration.

Michael Rosenbaum

*Rosenbaum Advisors, Inc.
Arlington Heights, Illinois*

MICHAEL ROSENBAUM IS currently president of Rosenbaum Advisors, an independent consulting firm specializing in investor relations and strategic communications.

During the past 20 years, he has counseled more than 150 public and private firms—including both start-ups and Fortune 500 companies—on issues from initial public offerings to mergers, acquisitions, bankruptcy, proxy, earnings restatements, and crises. He has written two books on investor relations—*Board Perspectives: Building Value through Investor Relations* (CCH, 2004) and *Selling Your Story to Wall Street: The Art and Science of Investor Relations* (McGraw-Hill, 1994)—and is co-author of *The Governance Game* (Aspatore, 2003). He is also widely published on investor relations, corporate governance, and management issues, and is a frequent speaker on a wide range of business topics. Mr. Rosenbaum holds an MBA from Roosevelt University and a BA in communications from the University of Illinois.

John Sacco

*George Mason University
Fairfax, Virginia*

JOHN SACCO IS currently 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. He has forthcoming work on the evolution of end-user computing. In addition to his writing, Dr. Sacco has consulted for several state and local governments and Certified Public Accountant (CPA) firms, including contact work with CPA firms on finance and accounting for the Chinese government.

John Sacco holds a PhD in political science, an MS in accounting, and a BS in data processing.

Michael D. Serlin

*Independent Consultant
Alexandria, Virginia*

MICHAEL D. SERLIN, previously a member of the Scorecard research team, is currently writing, speaking, 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. The law included requiring 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 and has participated in their 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 received his BA in political science from Stanford University.

Robert Shea

*Grant Thornton LLP
Alexandria, VA*

ROBERT SHEA IS a director in Grant Thornton's Global Public Sector. He was most recently at the Office of Management and Budget as associate director for administration and government performance. In addition to managing OMB's internal operations, Mr. Shea led the President's Performance Improvement Initiative, administered the Program Assessment Rating Tool, advised on government human capital policy, and led inter-agency collaborations in the areas of food safety and implementation of the Federal Funding Accountability and Transparency Act.

Before joining OMB, Mr. Shea served as counsel to the Senate Committee on Governmental Affairs where, in addition to general oversight of executive branch management, he advised committee leadership on the status of implementation of the statutory framework for performance-based government, including the Government Performance and Results Act and the Chief Financial Officers Act. He was legislative director for Rep. Pete Sessions (R-TX) from 1997 to 1999, where he organized the Results Caucus, a group of members of Congress dedicated to results-based management and solving many of the government's major management problems. Mr. Shea was a professional staff member with the House Committee on Government Reform from 1995 through 1996. There he examined the economy and efficiency of government programs and acted as liaison with the government's inspectors general.

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 Public Manager*, *Journal of Politics*, *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 the *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 received 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 Mercatus Center's Government Accountability Project. In 1997, after completing his term as New Zealand's ambassador to Canada, he joined George Mason University as a distinguished visiting scholar. 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, Mr. McTigue shares with U.S. policy makers his practical experience. 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, D.C., where he served for over 30 years on the staffs of the U.S. Government Accountability Office 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 and then in 2001 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.

Mr. Wray earned his BA in political science from Washington & Jefferson College and a JD with honors from the National Law Center at 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