

SCORING SUMMARY

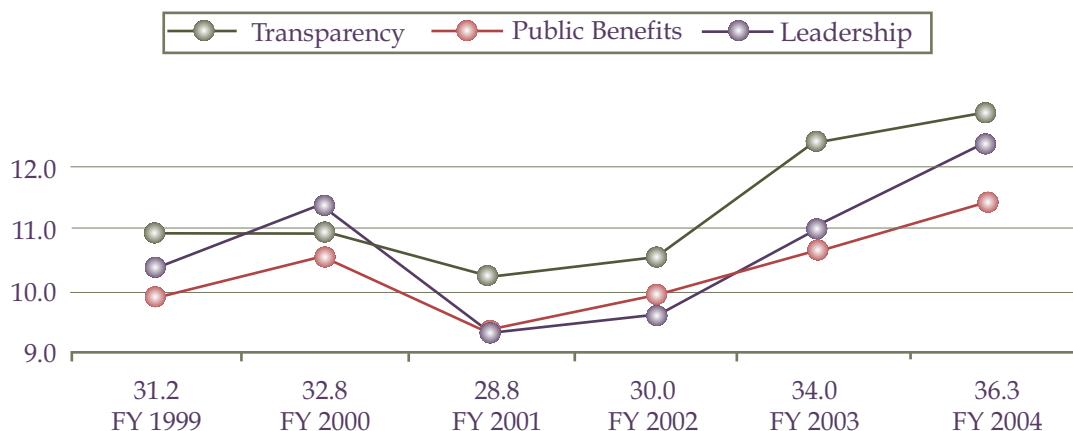
TRENDS AND OBSERVATIONS

Average scores on each of the three dimensions continued to improve in fiscal 2004, as Figure 3 shows. In fiscal 2004, the average transparency, public benefits, and leadership scores improved by 3 percent, 6 percent, and 12 percent respectively. The first two dimensions improved by a much smaller percentage than in fiscal 2003, when transparency improved by 16 percent and public benefits improved by 10 percent. Leadership improved by about the same percentage in fiscal 2004 as in fiscal 2003.

The average total score, which is the sum of the three dimensional scores, increased 7 percent to 36.3 for fiscal 2004. This is a significant milestone, because a report that received the satisfactory score of 3 on each criterion would receive a 36. This is the first time in the six-year history of the Mercatus Center's Scorecard that the average score exceeded satisfactory.

As Figure 4 shows (see p.17), nine criteria showed increased average scores. In particular, agencies made substantial improvements in showing how their activities make the country a better place to live (criterion 9), explaining failures to achieve goals (criterion 10), and describing changes to improve performance in the future (criterion 12). Average scores on these criteria jumped by more than 15 percent. A principal reason for improvements in criterion 9 is stronger transmittal letters from agency heads, which should paint a broad vision of the agency's value to taxpayers and summarize key performance highlights. Criterion 10, explaining failures, has been the lowest-scoring category since fiscal 2001, averaging only 2.3 in fiscal 2003. Therefore, there was a lot of room to improve. Improvement on this criterion, in turn, helps explain the improvement in criterion 12. Agencies that do a better job of disclosing failures also tend to do a better job of explaining how they intend to remedy them, and agencies that fail to disclose failures rarely disclose remedies either.

FIGURE 3: IMPROVEMENT CONTINUED IN FISCAL 2004



Reviewing the results, the research team was surprised to see only modest improvement in criteria 4 and 5—articulation of outcome-oriented goals and measures. More improvement was expected, since many agencies are now operating under new strategic plans. A review of the evidence revealed that most agencies adopted new strategic plans in fiscal 2003, and these new plans apparently drove scores on criteria 4 and 5 up by more than 20 percent in the fiscal 2003 Scorecard.¹¹

The average scores on criteria 1, 7, and 11 fell by 4, 5, and 6 percent respectively. These reductions suggest that at least some reports failed to improve sufficiently to keep pace with our tightening criteria:

CRITERION 1: To receive the highest score of 5 for web accessibility, agencies had to make their reports available via a direct link from the home page and offer contact information so that readers could follow up with questions, comments, or requests for a hard copy.

CRITERION 7: Although many reports offered anecdotes or logical explanations illustrating how their actions affected results, relatively few have backed this up with solid evidence from program assessments that demonstrate how the agency's actions have *caused* the measured outcomes. Without proof of causality, it is difficult to say that a report truly demonstrates an agency's contribution.

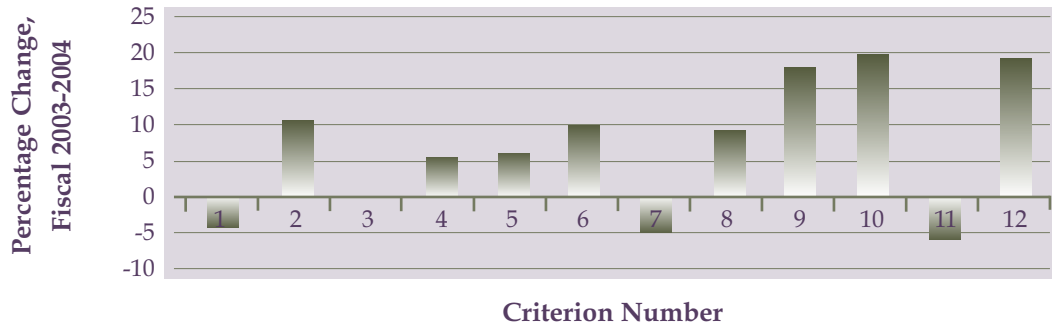
CRITERION 11: This criterion assesses whether the report adequately addresses major management challenges. This year, the research team looked for evidence beyond mere expressions of good intentions that the agency takes major management challenges seriously and is making specific, substantive progress on resolving them. Key factors include specific assessments of progress by the inspector general (rather than just assertions by the inspector general that the agency is "making progress") and responses by the agency that indicate what, in particular, it is doing to deal with each major management challenge.

MOST RECENT STRATEGIC PLAN FOUND ON AGENCY WEB SITES

AGENCY	YEAR
Agriculture	2002
Commerce	2004
Defense (Quadrennial Defense Review)	2001
Education	2002
Energy	2003
EPA	2003
GSA	2002
Homeland Security	2004
HUD	2003
Interior	2003
Justice	2003
Labor	2003
NASA	2003
NRC	2004
NSF	2003
OPM	2002
SSA	2003
State	2004
Transportation	2003
Treasury	2003
USAID	2004
Veterans' Affairs	2003

¹¹ The Government Performance and Results Act requires that strategic plans must cover a five-year period, and they should be updated every three years.

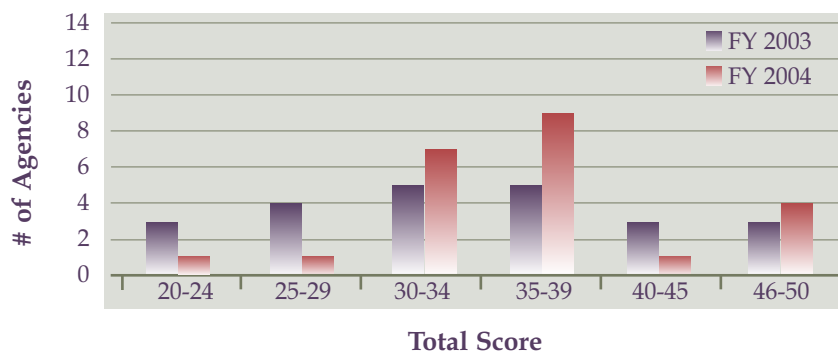
FIGURE 4: SCORES ON ALL BUT 4 CRITERIA IMPROVED IN FISCAL 2004



A YEAR OF UNEVEN IMPROVEMENT

The average score increased 7 percent, from 34 in fiscal 2003 to 36.3 in fiscal 2004. Figure 5 shows a reduction in the number of reports scoring below 30 and an increase in the number of reports scoring between 30 and 40. This increase reflects uneven improvement. About half of the agencies – 12 – increased their scores from fiscal 2003 to fiscal 2004. Scores fell for seven agencies and remained the same for three agencies. Scores for two agencies – Homeland Security and Health & Human Services – are missing in one year or the other because their reports were not available in time to meet our evaluation deadlines.

FIGURE 5: MORE REPORTS GOT BETTER SCORES IN FISCAL 2004



Each agency score is a sum of the three dimensional scores: transparency, public benefits, and leadership. Figures 6 through 8 show that changes in these scores reflect the same pattern as changes in the total scores.

FIGURE 6: MORE REPORTS GOT HIGHER TRANSPARENCY SCORES IN FISCAL 2004

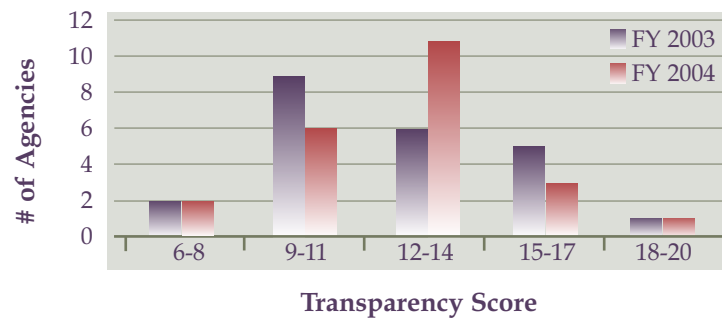


FIGURE 7: MORE REPORTS GOT HIGHER BENEFITS SCORES IN FISCAL 2004

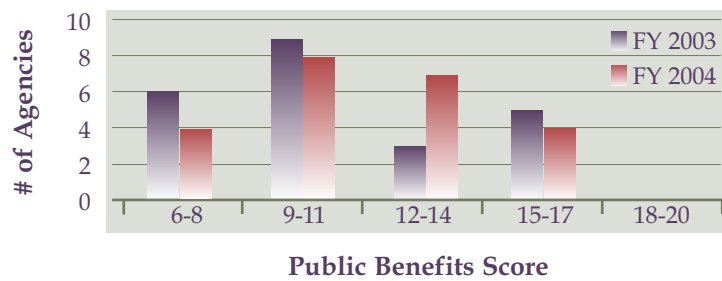
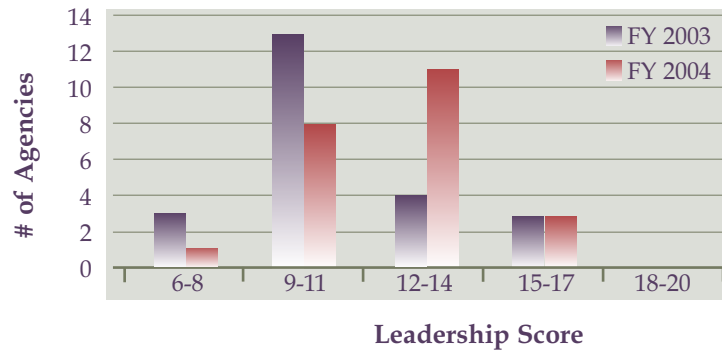
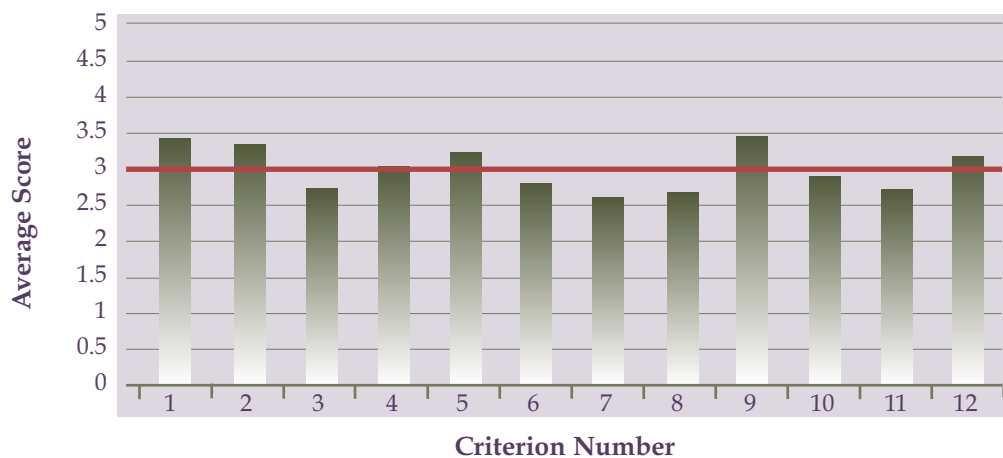


FIGURE 8: MORE REPORTS GOT HIGHER LEADERSHIP SCORES IN FISCAL 2004



Despite this year’s improvements, there is still substantial room for progress. Figure 9 reveals that average scores in six categories are below 3. The areas with scores below 3 (in order of severity of the problem) are demonstration that the agency’s efforts actually affected achievement of outcomes, linkage of results with costs, ensuring reliability and timeliness of data, discussion of major management challenges, articulation of outcome-oriented measures that accurately reflect the agency’s impact on its goals, and explanation of failures to achieve goals.

FIGURE 9: AVERAGE SCORES ON INDIVIDUAL CRITERIA FOR FISCAL 2004



TOP REPORTS

The Department of Labor's report tied for 1st place with Transportation's in fiscal 2003. This year, Labor's report captured top honors all on its own, scoring a record-high of 50 out of 60 possible points. Labor has demonstrated steady improvement in its reporting through the years, moving up from 5th for fiscal 1999 to first capture the top ranking in fiscal 2002. Labor has kept its leading position since then.

This year's 2nd place report, State, continues a remarkable improvement story. State's report ranked 20th in fiscal 1999 but has risen in the rankings every year. In addition, State's initiative to develop a joint strategic plan with USAID helped increase USAID's ranking by 12 places, from 21st place in fiscal 2003 to 9th place in fiscal 2004.

Two perennially strong reports, Transportation and Veterans, shared 3rd place. Six points separated the top four reports from the 5th place report. We encourage other agencies to learn from best practices and incorporate the approaches that have been successful at these agencies.

UPS AND DOWNS

In addition to State and USAID, four agencies made significant improvements in their reports: Commerce, Justice, Energy, and NSF. Each of these reports substantially increased its score and ranking for fiscal 2004. Commerce leapt 11 spots in the rankings to capture 5th place in fiscal 2004. This reversed its drop from 5th place in fiscal 2002 to 16th place in fiscal 2003. Justice and Energy rose nine and six places, respectively, to tie with SBA for 6th place. NSF captured 10th place in fiscal 2004, up from 17th in fiscal 2003. Another noteworthy climber is NASA, which rose from 20th in 2003 to 16th in 2004 after falling from 12th in fiscal 2002.

Four reports fell significantly in the rankings—Agriculture, Interior, HUD, and GSA. Agriculture fell from 4th place to 13th, Interior fell from 6th place to 13th, and HUD and GSA both fell from 10th place to 16th. All of these reports also had lower scores in fiscal 2004 than in fiscal 2003.

CELLAR DWELLERS

Several reports have consistently ranked in the bottom half for most of the past six years:

- Defense's report ranked 7th in fiscal 1999, fell to 18th in fiscal 2000, and has never ranked higher than 22nd since then.
- The Office of Personnel Management's report has never finished higher than the middle of the pack; its best ranking was 12th in fiscal 2002. It ranked 17th in fiscal 2003 and 21st in fiscal 2004.
- NASA's report has had a similar, though even more erratic, experience. Its ranking went from 14th in fiscal 1999 to 23rd in fiscal 2000, rebounded over two years to 12th in fiscal 2002, plummeted to 20th in fiscal 2003, and rose slightly to 16th in fiscal 2004.
- The Department of Health and Human Services did not provide a copy of its fiscal 2004 report in time for our evaluation. Its report has never ranked higher than 20th.

STRONGEST AND WEAKEST SCORES

TRANSPARENCY

1. Is the report easily accessible and easily identified as the agency's Annual Performance and Accountability Report?

Access to performance information is critical because public accountability can only be served if the public can actually find out what benefits an agency provides. The annual report should be easily available to the public, stakeholders, and the media. Ideally, this means that the agency's home page displays a link clearly guiding the reader to the annual report for the most recent fiscal year. If one has to be an expert on performance management, the agency's structure, or the structure of the agency's web page to locate it, the spirit of accountability to the public is not satisfied. If the report is large, it should be divided into sections for more convenient reading and/or downloading. Making the report available in multiple formats is also desirable, since readers' needs vary and each format has its advantages and disadvantages (e.g., ease of printing, searching, etc.). Finally, it is helpful to include contact information so that people can phone or e-mail questions, comments, or requests for a hard copy of the report.¹² In some sense, criterion 1 should be the easiest criterion to satisfy, because any webmaster should be able to ensure a good score.

Reports were due to the President and Congress by November 15. Reports are included in this evaluation if they were made available to Mercatus in some form by December 1. The Mercatus research team gave agencies several additional weeks, until December 15, before checking to see if the reports were available on agency web sites. Links to agency reports can be found on the Mercatus web site at <http://www.mercatus.org/governmentaccountability/article.php/1095.html>

Fifteen of the agency reports were linked directly to the home page or were readily identifiable by an intuitive search. Five were available as a home page link captioned "about the agency" or obtainable via a search under "performance report."

Strongest Scores: State, Transportation, Veterans, Energy, EPA, Education

Each of these reports received a 5 on this criterion. All of these agencies make their reports available via a direct link from the home page. All allow the reader to download the entire PDF document or separate sections. And all provide contact information in case readers have questions, comments, or further requests. A report had to be available via a direct link and include contact information in order to receive a 5. The most common reason an agency received only a 4 was that it failed to provide any contact information so readers could reach a responsible individual or office at the agency.

Several agencies provided additional features as well. State offers downloads in HTML as well as PDF. Energy's report web page offers individual employee names, phone numbers, and e-mail addresses for followup. Education offers the report in Word as well as PDF and includes information for obtaining the report

¹² Contact information as used here means, ideally, a named agency official or employee to contact or, at a minimum, a specific office within the agency.

in alternative forms such as Braille, large print, or computer disc. At least two agencies that failed to receive 5s nevertheless offered interesting improvements on accessibility. GSA and SBA both made their reports available in Spanish as well as English.

Weakest Scores: HUD, OPM, Homeland Security

All three agencies received a 1 on this criterion. The first two agencies were a day late and a dollar short in their web posting. HUD has a direct link to its report from its home page, but none of the research team could open the report from that link by December 16. Subsequent to the cutoff date, the report could be opened from the web site link. Similarly, researchers could not find OPM's report on line by December 16. The "About the Agency" link from the home page leads to a link for "Performance & Accountability Reports." However, the most recent report posted on that site as of the cutoff date was the fiscal 2003 version. Subsequently, the agency did post the fiscal 2004 report on its web site.

More timely and user-friendly web presentation could have benefited each of these reports substantially in the rankings. If each had scored a 5, HUD would have tied for 10th place instead of 16th, and OPM would have tied for 15th place instead of ranking 21st. Although the lower rankings are arguably an unfortunate result of the deadline chosen by the Mercatus research team, we note that agencies were informed of the deadline in June 2003, and 21 of the 23 agencies whose reports are evaluated this year managed to have working web links by December 16.

The researchers eventually found Homeland Security's report on the department's web site. However, the report is located under "budget," which is linked under "DHS organization" from the home page. This is not an intuitive search path for a lay reader. The report is downloadable in a single PDF or multiple files. No contact information is furnished.

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

The annual performance and accountability report is a communications device directed at non-specialist audiences. Therefore, its style, language, and subject matter must reflect that purpose. It should focus on an agency's mission, how it organizes efforts toward that end, and how much progress was made toward its achievement in the preceding fiscal year. Contents should be clear, logical, easy to navigate, and presented in such a way that the structure aids understanding. Consistent format, clarity of text, absence of jargon, and effective use of visual techniques like headings, graphs, tables, and photos are helpful. Acronyms can be helpful if they substitute for lengthy proper names that readers may be familiar with, but use of acronyms to refer to documents, processes, systems, nouns other than proper names, verbs, short names, or names of things known only to insiders inhibits understanding, even if the report provides a list of acronyms. Details can either inform or confuse, depending on how they are presented. Anecdotes can promote effective communication if they complement and illustrate, rather than substitute for, outcome-based results measures. Information necessary to evaluate the agency's performance should actually be present in the report, not just referenced as available in some other document, web page, or computer disk.

Readability was an area of widespread improvement. Nine agencies scored higher this year than last year. Two agencies (Commerce and Justice) improved from 2 to 4. Three agencies regressed (Education, EPA, and

ANACRONYM-ISMS?

ACRONYM	MEANING	SOURCE
MIDAS	Modernize and Innovate the Delivery of Agriculture System	USDA
EEEL	Electronics and Electrical Engineering Laboratory	Commerce
GWOT	Global War on Terrorism	Defense
HOUSSE	High, Objective, Uniform State Standard of Evaluation	Education
LUST	Leaking Underground Storage Tank	EPA
EWGPP	Elimination of Weapons Grade Plutonium Production	Energy
OCPO	Office of Chief People Officer	GSA
TOPOFF	Top Officials	Homeland Security
TOTAL	Technology Open to All Lenders	HUD
WAU	Whereabouts Unknown	Interior
WYC	Watch Your Car	Justice
LWDII	Lost Workday Injury and Illness	Labor
ESMF	Earth Science Model Framework	NASA
SDLCMM	System Development Life-Cycle Management Methodology	NRC
PITO	People, Ideas, Tools and Organizational Excellence	NSF
USSTR	U.S. Navy's Aircraft Carrier <i>Theodore Roosevelt</i>	OPM
JA2MS	Joint Accounting and Administrative Management System	SBA
EPOXY	Earnings Posted Overall Cross Total/Year-to-Date System	SSA
WW2BW	White Water to Blue Water	State
TE	Traffic Enforcement	Transportation
FMLOB	Financial Management Lines of Business	Treasury
GESAMP	Group of Experts on the Scientific Aspects of Marine Environmental Protection	USAID
SKIPPES	Skills, Knowledge, and Insurance Practices and Procedures Embedded in Systems	Veterans

List is intended merely to be illustrative, not exhaustive.

HUD). Only four agencies received a below-average (2) score on this criterion.

Strongest Scores: State, Labor, Veterans

All of these reports received a 5 on this criterion.

The State Department's report is certainly one of the most (if not the most) visually appealing of the performance reports. It is also well organized and highly readable. The report makes excellent use of graphics. It con-

tains many features that assist the reader in readily grasping the department's functions and performance. To cite just a few examples: colorful and succinct tables summarize the department's fiscal 2004 results and compare them to fiscal 2003 (pp. 16-20); narratives describe the most important results and continuing challenges for each strategic goal (pp. 22-25); and the following pages provide examples of significant achievements. Also of note is the 60-page highlights report available for the general reader. Both the full report and the highlights report have an accompanying CD for readers interested in greater depth. The report makes liberal use of acronyms, but this is not a significant distraction for the reader. The only challenge in terms of readability and understandability is that the department has 209 performance measures. Its already excellent presentation could be even better if it focused on significantly fewer measures.

Labor's report is well-organized, easy to read, and very informative. Among its many excellent features, the secretary's transmittal letter (pp. 6-8) highlights the department's core missions and selected accomplishments. The executive summary includes a table (pp. 14-18) that presents the department's performance results at a glance and a narrative summary thereafter (pp. 19-20) that succinctly highlights major performance trends and plans for the future. The report also makes good use of graphics, as well as vignettes that highlight the impact of its programs on individual citizens. It includes a list of Internet sites for more detailed information on subjects covered in the report (p. 341). This report clearly reflects a strong interest in communicating the department's performance to the public.

The Veterans' Affairs report is well-organized and easy to understand, both for the casual reader and one seeking greater depth. It makes excellent use of graphics, particularly in tables and graphs that present the department's performance results. The department has 127 performance goals, but the report focuses on 22 key goals selected by the department's leadership as being critical to the department's success (p. 7). The performance results for these key goals are presented in an excellent table (p. 6). The report covers all of the department's performance goals on visually appealing, color-coded tables (pp. 139 ff.) The format makes it very easy to digest the performance results at a glance. Straightforward narratives contribute to making the report a model of clarity.

Weakest Scores: EPA, Education, Homeland Security, Defense

Each received a 2. All of these reports have some good features, but each also posed some major readability challenges.

EPA has descriptions of its many goals and measures scattered throughout the report and presented mainly by narrative. It is hard to identify performance results and to interpret trends. The report is overwhelming in detail and would benefit greatly from overviews and summaries of goals and specific performance results. It could also make much better use of graphics. The report uses numerous acronyms and abbreviations (e.g., CO, SO, NO, Pb, PM, NAAQS), which requires the reader to flip back and forth to determine their meanings and deflects the reader's focus from the topic being discussed. Appendix C, which describes Performance Assessment Rating Tool results, would be more helpful if there was an easy method to tie the measures used for the programs assessed with the strategic and annual performance goals/measures. One good feature of the report is a narrative "Overview of Performance Results," which succinctly captures some key results using an effective "headline" format (pp. I-2 - I-7).

Education's report is visually appealing and has good photographs and graphics. It starts with a brief preface entitled "A Journey from Access to Excellence," which is well done. However, the main content is not easy to read or understand. The report makes heavy use of acronyms. While it includes a glossary to explain them (p. 261), they are still distracting. The performance portions of the report are decentralized and somewhat disjointed. Performance results are presented in one place (e.g., p. 39) and the related information necessary to put those results in context, including baseline and trend data, is presented elsewhere (e.g., p. 194). This forces the reader to cross-reference continuously. Even when the reader cross references, the measures are sometimes incomplete and cryptic. For example, the measure for "highly qualified teachers" for objective 2.4.1 (p. 206) doesn't define the term, but simply refers the reader to a section of the Elementary and Secondary Education Act. The reader must look in still another place (p. 29) for an explanation of the methodology for reporting performance results, which is in itself complex. The report would be improved considerably by the use of summary tables to highlight performance results.

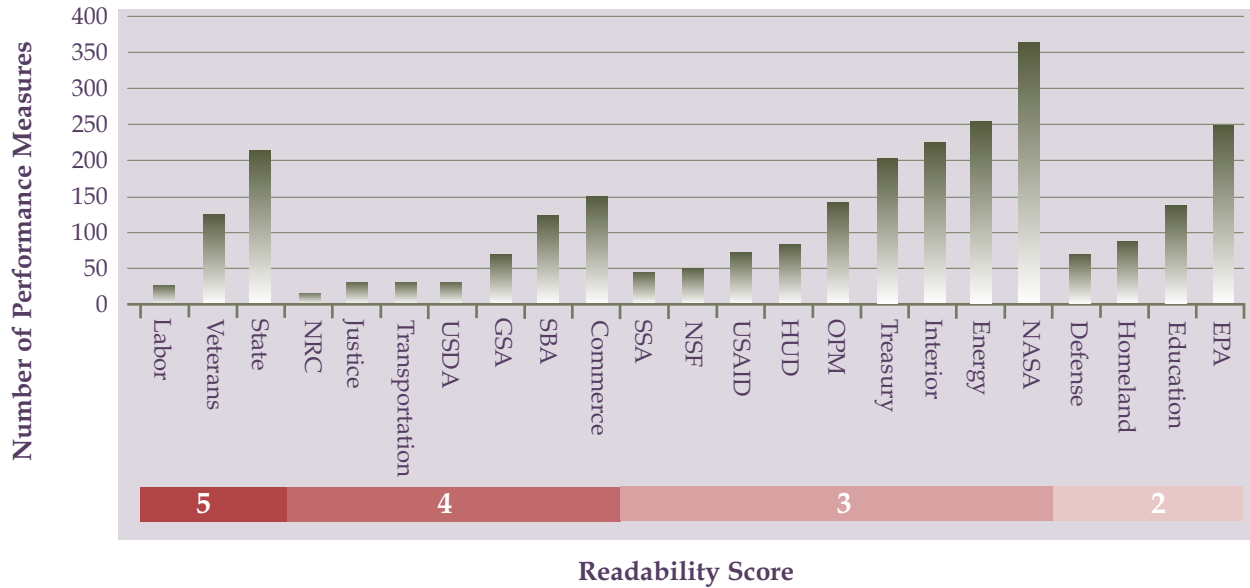
Homeland Security's report is mainly text and could benefit from more graphics—particularly in Part III, the detailed performance section, which consists entirely of text (pp. 192 ff.). A confusing numbering system and summary reports require considerable digging by the reader and cross-referencing between the un-numbered Executive Summary Performance Scorecard in Part I and the detailed numbered Performance Information in Part III. With that effort, the reader can get some sense of what the specific goal was for 2004 and whether or not it was met. However, it would enhance the understandability of the report if these strategic objectives were mentioned earlier and integrated with the description of the strategic goals themselves. Moreover, the report comes across as a series of reports on its individual components that lack overall cohesion. On the positive side, the secretary's transmittal letter does a good job of highlighting what the department does and the importance of its work to the American people. The "performance scorecard" presents the department's goals and results in tables that are easy to digest (pp. 7-17).

The Defense Department report is a hard read for one interested in understanding the department's performance beyond the most obvious aspects. One reason for this is that the department evidently still has not fully adapted to the Government Performance and Results Act as the primary framework for its performance accountability and reporting. The report says that the Quadrennial Defense Review Report serves as the department's strategic plan (p. 6). Thus, its report is a confusing mix of risk-management strategic objectives, "policy goals," and "performance metrics." Moreover, the portions of the report detailing the department's program performance do not deal comprehensively with its missions and activities (pp. 48-112). Finally, it takes a close reading to ascertain how well the department performed against the limited measures that it had for fiscal 2004.

The most common shortcoming for agencies that did not score well on this criterion is too many performance goals and measures. As Figure 10 shows, more measures do not mean that a report is more informative. The graph ranks reports from left to right based on readability scores, with the highest-scoring reports on the left. Few reports with large numbers of performance measures received better than a satisfactory (3) score.

The most extreme example is NASA, which has 365 measures. Other examples of agencies with a host of measures are Energy (255), Interior (226), EPA (200-300), and Treasury (203). Commerce has 153 measures but notes that it plans to reduce the number in the future. By contrast, Transportation has only 31 measures.

FIGURE 10: DATA IS NOT INFORMATION



The two high-scoring reports with large numbers of measures—Veterans and State—focus the reader on a more manageable number of “key” goals and measures. It also helps to summarize performance results in an easily comprehensible table, as Labor does.

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

The Reports Consolidation Act requires that the head of each agency assess the quality of the data in the report and disclose any material weaknesses. Like a similar requirement imposed on the chief executive officers of publicly-held companies by the Securities and Exchange Commission, the purpose of this requirement is to ensure that the quality of the underlying data is a management priority, and the deficiencies are acknowledged and corrected as quickly as possible.

It appears that agencies were more conversant with the data assessment requirements of the Reports Consolidation Act this year. Last year, seven agencies offered no assessment of data quality; this year, only one failed to do so (Defense). However, the transmittal letters for several agencies (Commerce, Energy, and Interior) were somewhat vague in their assessments. A more serious problem is that some agency heads (HUD, SSA) provided data assessments that appeared to ignore or understate data problems identified in other portions of the report, such as the internal auditor’s report or the inspector general’s presentation of major management problems.

More generally, the report should indicate the agency's confidence in the quality of the data used to document its results. Since the purpose of gathering these data is to manage programs strategically, one test of their adequacy is whether they are relevant, timely, complete, accurate, and consistent enough to use as the basis for decision-making. Data should be independently validated (i.e., certified as appropriate for the associated performance measure) and verified (i.e., assessed as reliable). Outside verifiers should be able to access the data with relative ease. Sources and descriptions should be provided for all outcome data.

For fiscal 2004, the Mercatus research team paid special attention to opinions offered by agency inspectors general, the Government Accountability Office, and other third party auditors to assess the credibility of information the agencies offered in their reports. For example, an agency head's certification that data in the report are "complete and reliable" carried less weight if the inspector general or other auditor presented findings undermining that claim.

One interesting result for fiscal 2004 is that the accelerated reporting deadline of November 15 does not appear to have diminished the availability or completeness of data in the reports. Fourteen of the 23 agencies reported actual, preliminary, or estimated results for almost all of their measures. Four reports had significant problems with missing data, but for three of the four, the problems clearly result from factors other than the earlier reporting deadline. Education, for example, has always had problems with lagging data because much of its performance data comes from the states, and OPM lacked data on a large number of measures because it did not implement the customer service surveys that were supposed to generate the data. Only one report, EPA, had significant amounts of missing data not explained by other factors.

Strongest Scores: Agriculture, NSF, State, Labor, Transportation

No agency received a 5 on this criterion, but five agencies received 4s.

Although Agriculture's report ranked 13th overall, it offers a potential best practice in this category. The secretary's transmittal letter (p. ii) certifies that the report's data are valid, reliable, and accurate except where specific limitations are discussed in the body of the report. The report provides extensive detail concerning the data for each individual performance measure (pp. 116-37). This includes a discussion of data sources and their completeness, reliability, and quality. The only shortcoming is that the department deferred reporting even preliminary or estimated results for three of its 31 performance goals (about ten percent) due to unavailable data (p. 8).

Similar to Agriculture, NSF ranked around the middle of the pack overall. Nevertheless, NSF has another idea that could be a best practice. The agency head's transmittal letter (pp. I-1 and I-2) asserts that the report's financial and performance data are complete and reliable and notes that this assertion is based on an independent external verification and validation review of the agency's performance results (pp. I-1 - I-2). The report elaborates on this review (pp. I-21 and I-22). The only negative here is the inspector general's comment, as part of her description of major management challenges, that the agency's current information systems do not readily provide the cost accounting information needed to link costs to program performance (p. IV-9).

The Secretary of State's transmittal letter certifies that the report's financial and performance data are complete and reliable in keeping with Office of Management and Budget guidance (p. IV). Unlike most reports,

it also describes this guidance so the reader can actually understand what the certification means (p. 12). Other portions of the report provide further information about steps the department takes to assess data veracity and validity (p. 13). Also, the narratives for each performance measure provide a data source (at least in general terms), as well as a statement as to the verification confidence level and a validation indicator. (See, e.g., p. 89) Since some of the department's measures are of necessity subjective, an arrangement for third-party assessments (rather than exclusive reliance on department officials) would enhance reader confidence in the data. The department reports fiscal 2004 results for all 209 of its measures (pp. 17-20).

Labor's report provides interesting examples of ways to deal with common data problems. The secretary's transmittal letter states that the department's performance data are complete and reliable, under the Office of Management and Budget criteria, with one exception (p. 7). The transmittal letter briefly describes the exception, which affects youth employment and retention rates, and refers to a more detailed discussion later in the report. The body of the report provides more information and indicates that the department is taking steps to address the problem (pp. 69-70). The department reports results for all of its goals and measures that are subject to fiscal year reporting. Some goals and measures are reported on a program year basis (July 1 to June 30). For these, program year 2003 results are included in this report and program year 2004 results will be reported next year (pp. 14, 16-17). Appendix 1 includes data sources for all performance measures (pp. 292-314). It points out that some results are reported on the basis of estimates and are so identified (p. 292).

Transportation's report contains much useful information on performance data. The report discusses performance data completeness and reliability in general terms (pp. 83-84). However, Appendix C provides greater detail on each performance measure that includes entries on data sources, completeness and reliability and also references a web site that provides additional helpful source and accuracy statements concerning the department's performance data (p. 265). The department had enough data to report fiscal 2004 results for all of its performance measures. About 60 percent of the results (19 of 31) are based on projections or estimates. The report is clear when less than complete data are used. There is nothing in the report (including the inspector general's input) that calls into question the credibility of the department's performance data. In accordance with the Reports Consolidation Act, the transmittal letter states that the performance and financial data "are substantially complete and reliable," and refers to a section of the Executive Summary for a detailed assessment of limits to the performance data and remedies for them (p. vi).

Weakest Score: Defense

While there are many tables throughout the report providing numbers, the relationship of these to performance objectives is often tenuous, and the department admits that correction efforts for reported weaknesses in accurate, reliable, and timely financial and performance information are not yet in place (p. 19). Furthermore, the report does not even arguably comply with the Reports Consolidation Act requirement that the secretary's transmittal include an assessment of the completeness and reliability of the financial and performance data used in the report and what is being done to address material inadequacies. (In fact, the secretary did not transmit the report, leaving this task to the deputy secretary.) In any event, other portions of the report make clear that the department faces severe data problems. The report states: "The Department of Defense financial and business management systems and processes are costly to maintain and operate, not fully integrated, and do not provide information that is reliable, timely, and accurate." (p. 24) Later, the report admits: "The DoD

systemic deficiencies in financial management systems and business processes result in the inability to collect and report financial and performance information that is accurate, reliable, and timely.” (p. 43) The Defense Department’s report received a 1 on this criterion.

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. To provide this information, agencies must design a measurement system that facilitates analysis of trends over time. Data should be displayed in a way that allows readers to detect and understand their significance easily. Both quantity of data (years of data included) and presentation matter. Good performance measures that have limited data (due to newness or revision) may convey more information than inferior measures with more data points that are not clearly linked to an agency’s results.

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

Strongest Score: Numerous

Seven reports received a 4 on this criterion. Two of the best reports in this category are Agriculture and Veterans. The Agriculture report provides detailed backup for its performance measures that includes rationales for each performance target. This is very helpful for putting the measures in context. Unfortunately, Agriculture is the only agency to provide such information systematically. The Veterans report includes strategic (i.e., long-term) performance targets, which show the reader where the department is trying to go as an important aid to assessing performance and progress.

Most agencies could improve in this category by (1) presenting both prior year targets and performance results; (2) presenting long-term targets for each measure; and (3) providing rationales for annual performance targets. A rationale is particularly important where there is a significant disparity between the target and result. Even when actual performance significantly exceeded the target, some explanation may be necessary, because the target may not have been very challenging.

Weakest Score: NASA

For several reasons, the report provides virtually no baseline or trend data that help the reader put its measures in context. First, a large proportion of the measures are not expressed in quantifiable terms. Second, the agency’s performance measures tend to change considerably from year to year; there are many new measures for fiscal 2004. Third, NASA’s method of reporting prior year results where there is continuity of measures indicates only whether or not the prior year target was met; it does not identify prior year targets or prior year results even in the case of quantifiable measures. (See, e.g., pp. 68-69) To NASA’s credit, the report recognizes the limitations of its approach in terms of identifying trends and notes that “performance outcomes” have been added to its performance metrics this year in an effort to help track trends over time (p. 18). This should provide for better baseline and trend data in the future.

B 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 and simple terms that describe benefits to the community rather than activities that are presumed to be of value. Vague verbiage that emphasizes things an agency does instead of why it is doing them should be avoided. This admonition applies at all goal levels—strategic goals, objectives, and annual performance goals.

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

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

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

Some agencies did a better job of articulating outcomes in the transmittal letter from the agency head or in overview sections than in the detailed performance review. Agencies seeking to develop more outcome-oriented goals would do well to examine those prefatory parts of their report for inspiration.

Strongest Scores: State, USAID, Transportation, Labor

All of the State Department’s 11 programmatic strategic goals (see p. 11) are outcome-oriented, as are the great majority of the 38 annual performance goals (pp. 11, 17-20). Of necessity, many are stated at a high level of generality, such as “advance the growth of democracy” and “strengthen world economic growth.” The department has done a very commendable job in meeting the substantial challenge of describing its functions through such outcome-oriented goals.

USAID and the State Department have developed a new joint strategic plan. Thus, USAID’s eight strategic goals (p. 31) are the same as the State Department’s, although the State Department has some additional strategic goals. Likewise, USAID’s annual performance goals track State Department performance goals. The agency’s development of joint goals with State has resulted in a major improvement over last year in this category. Like State’s goals, USAID’s strategic and performance goals now are very outcome-oriented, although many are necessarily stated at a high level of generality.

JOINT STATE DEPARTMENT/USAID STRATEGIC GOALS

REGIONAL STABILITY: Avert and resolve local and regional conflicts to preserve peace and minimize harm to the national interests of the United States

COUNTERTERRORISM: Prevent attacks against the United States, our allies, and our friends, and strengthen alliances and international arrangements to defeat global terrorism

INTERNATIONAL CRIME AND DRUGS: Minimize the impact of international crime and illegal drugs on the United States and its citizens

DEMOCRACY AND HUMAN RIGHTS: Advance the growth of democracy and good governance, including civil society, the rule of law, respect for human rights, and religious freedom

ECONOMIC PROSPERITY AND SECURITY: Strengthen world economic growth, development, and stability, while expanding opportunities for U.S. businesses and ensuring economic security for the nation

SOCIAL AND ENVIRONMENTAL ISSUES: Improve health, education, environment, and other conditions for the global population

HUMANITARIAN RESPONSE: Minimize the human costs of displacement, conflicts, and natural disasters

MANAGEMENT AND ORGANIZATIONAL EXCELLENCE: Ensure a high quality workforce supported by modern and secure infrastructure and operational capacities

ADDITIONAL STATE DEPARTMENT STRATEGIC GOALS

HOMELAND SECURITY: Secure the homeland by strengthening arrangements that govern the flows of people, goods, and services between the United States and the rest of the world

WEAPONS OF MASS DESTRUCTION: Reduce the threat of weapons of mass destruction to the United States, our allies, and our friends

AMERICAN CITIZENS: Assist American citizens to travel, conduct business, and live abroad securely

PUBLIC DIPLOMACY AND PUBLIC AFFAIRS: Increase understanding for American values, policies, and initiatives to create a receptive international environment

All of Transportation’s strategic goals are results-oriented, particularly when read in conjunction with their accompanying strategic objectives, termed “strategic outcomes.” The strategic goals for mobility and economic growth (p. 48) use somewhat vague, non-results verbs—“shape” and “support”—but clearly address measurable results in the accompanying text and strategic outcomes. All of the 31 performance measures are stated as outcomes. The national security strategic goal (p. 65) is stated in a more results-oriented way than last year, although its strategic outcomes remain the same.

Labor’s report has four programmatic strategic goals (pp. 19-20). Each strategic goal has several accompanying “outcome goals.” (See pp. 35-36, 75, 97-98, and 129.) For example, Strategic Goal 1, “A Prepared Workforce,” has the following three outcome goals: “Increase Employment, Earnings, and Assistance”; “Increase the Number of Youth Making a Successful Transition to Work”; and “Improve the Effectiveness of Information and Analysis on the U.S. Economy.” All of the strategic goals are stated as outcomes, as are all the outcome goals for the first three strategic goals. The two outcome goals under the fourth strategic goal are less clearly stated as outcomes (p. 129). There are 30 annual performance goals under the four strategic goals, almost all of which are clearly outcome-oriented (pp. 14-16). There are 12 additional annual performance goals under departmental management (pp. 16, 137-138). The first eight of these are also expressed as outcomes.

Weakest Scores: Numerous

Among the more problematic agencies are Agriculture, OPM, and Treasury. While Agriculture restructured and redefined its goals for this year, the net effect of these changes seems to have been to make the goals even *less* outcome-oriented than they were before. OPM defines its goals primarily in terms of the assistance it provides other agencies to enhance their workforces, not whether the other agencies actually have enhanced their workforces. Treasury’s goals seem fundamentally out of balance in relation to its missions. The goals place much greater emphasis on macroeconomic policy than on effective administration of the tax laws, although the Internal Revenue Service accounts for about 70 percent of the department’s budget and almost 90 percent of its staff resources. This is particularly ironic since the department’s span of influence over economic policy is much less clear than its ability to affect the performance of the Internal Revenue Service.

6. Are the performance measures valid indicators of the agency’s impact on its outcome goals?

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

Strongest Scores: Labor, State, Transportation, Veterans

Labor’s report offers several performance measures (“indicators”) for most annual goals (e.g., pp. 58, 64, 67). Most of these measures are expressed in the form of target outcomes that clearly relate to their accompanying goals. The narratives reinforce the relationship between the measures and the goals.

In contrast to State's broadly stated strategic and performance goals, many of its 209 individual performance measures are highly specific and tailored to address particular current needs or circumstances (e.g., pp. 89-93, 98, 107). Many also are expressed in narrative rather than quantitative form, thus leaving room for subjective judgment. However, they usually bear an obvious relationship to the goal. The report includes a brief description of the results and their impact for each measure, which further drives home the connection between the measure and its goal.

For the most part, Transportation's performance measures bear a clear relationship to its goals. This is due in large part to the highly outcome-oriented nature of both the goals and the measures. The relationships are most compelling under the safety goal and least clear under the national security goal. As was the case last year, there still are no performance measures that relate to the very important second strategic outcome under the security goal: "Reduce the vulnerability of the transportation system and its users to crime and terrorism" (p. 65). While the Department of Homeland Security has primary responsibility in this area, Transportation presumably still retains an important role in view of the stated strategic objective.

In the Department of Veterans' Affairs report, the key measures are clearly related to their outcome goals. The narrative portions of the report elaborate upon and reinforce these relationships. Most of the remaining measures appear related to their strategic goals and objectives, although there is little narrative on these measures.

Weakest Scores: Numerous

Eight different reports received a score of 2 on this criterion, including several that ranked relatively high on the strength of their overall scores: Commerce (ranked 5th) and SBA and Energy (tied for 6th with Justice). The *average* score on this criterion improved in fiscal 2004 only because 12 agencies received a 2 or a 1 in fiscal 2003.

This criterion examines whether reports show that the agency's actions actually caused the outcomes that it reports. The most common factor leading to low scores on this criterion is failure to measure outcomes in the first place. For example, one of Commerce's performance goals is to ensure competition in international trade (p. 293). However, the measures under this goal deal only with numbers of cases completed on time, initiated, or concluded. Most of NASA's goals simply describe actions to be taken, often in highly technical terms. The Social Security Administration uses a total of 45 performance measures, of which 17 are designated "key" measures (pp. 83-89). At most, six of the 17 key measures are stated as outcomes. A higher proportion of the remaining measures (about 16) could be described as outcomes. Many of the agency's measures simply track the volume of various activities.

In some cases, the measures focus on activities because the goals focus on activities:

- The Energy Department's 255 annual performance targets are primarily activity measures. In general, it is difficult to see how they relate to outcomes since in most cases neither the measures nor the goals to which they apply are stated as outcomes.
- Like OPM's goals, its performance measures are predominately activity or output-oriented in most areas, including those relating to OPM's core human capital missions.

- Similarly, some of Defense’s existing measures relate to outcome goals; many others are activity measures. However, the main problem is that the department’s existing measures do not come close to capturing the full range of the department’s operations. To its credit, this year’s report indicates that the department is in the process of developing many new performance measures.

In some cases, reports state measures as outcomes, but they need to explain why these measures actually benefit the public. SBA, for example, includes two outcome measures which seek to reduce federal regulatory enforcement actions and penalties against small businesses (p. 85). It is not clear from the text, however, why this would benefit the general public.

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/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/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 is helpful in keeping expectations ambitious, yet realistic. External factors, however, should be treated as influences that must be controlled for in order to identify the agency’s contribution — not excuses for a failure to demonstrate performance.

Strongest Scores: Labor, State, Transportation

No report received a 5 on this criterion; three received a 4. These agencies adopted outcome measures whose links to their goals were either intuitively clear or could be logically explained.

Such links were clear even when agencies faced substantial external factors that affected their ability to achieve their goals. In Labor’s case, for example, economic conditions obviously have a major impact. Nevertheless, the highly outcome-oriented nature of the department’s goals and measures and the narratives combine to do a good job of demonstrating the department’s contributions toward achieving its goals. In particular, the narratives usually provide brief explanations of the relevant departmental programs and activities and how they contribute.

In the future, agencies will likely have to offer credible program evaluations or other solid evidence of causality to receive a 5 on this criterion.

Weakest Scores: Defense

This report provides a good example of how weakness in articulating outcome-oriented measures can lead to

a low score on this criterion as well. The limited performance goals and measures that the department now uses do not provide an adequate basis for demonstrating its contributions. The new goals and measures under development offer some hope for the future.

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

Knowledge of resource allocation and linkage to strategic goals, objectives, and performance goals is useful because it clarifies priorities. Managing for results requires more, however. *Strategic* reallocation of resources becomes possible only when financial information allows one to calculate the cost per unit of success and to compare alternative methods of achieving the same goal. An agency cannot obtain the highest score on this criterion unless it breaks cost information down sufficiently to make such calculation possible.

Strongest Score: Numerous

Seven reports received a 4 on this criterion, though none was outstanding enough to earn a 5.

All of these reports broke down costs by more than just strategic goals and objectives. Some, such as Commerce, USAID, and Energy, even allocated costs among lower-level performance or program goals. The key missing ingredient is allocation of costs to individual performance measures. This information would make it possible to understand what citizens pay for various types of successes.

Weakest Scores: GSA, NASA, Homeland Security, Defense

Each of these reports earned a 1 on this criterion. Reports from GSA, NASA, and Defense have no content on this subject. NASA's omission is surprising, since its report states that it is a leader in implementing full-cost budgeting (p. 5). For Homeland Security, the only linkage between costs, goals, and results is a pie chart that shows net cost by strategic goal, but only by percentage—not actual budget numbers (p. 35). Later portions of the report seem to provide inconsistent figures for the department's fiscal 2004 resources (e.g., p. 65 vs. pp. 308-09).

L FORWARD-LOOKING LEADERSHIP

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

Does an agency realize and articulate the value it provides to the country? The report should speak directly to the public about how the agency produces benefits that are important to citizens. Politics have no place in this report. The public's interests are paramount, not individual or partisan credit or blame. Just as the best corporate reports feature communication directly from the chief executive, agency reports should demonstrate accountability of agency heads for their organization's performance. Lofty ideals must be supported by an outcome orientation, sound strategies, and successful achievement discussions. The report should create confidence in an agency's ability to improve America's future. Anecdotes and success stories can be important communication strategies in this regard, but their value is limited if not backed up by solid performance data. Indeed, the most common comment under this criterion from the research team was that many reports with good anecdotes or strong letters from the agency head never-

theless lacked the credible performance information necessary to show that the stories told in the report typify the results achieved.

Strongest Scores: Numerous

No report earned a 5, but 12 reports earned a 4.

Most of the reports receiving this score possessed three ingredients: (1) outcome-oriented measures that made the agency's contributions clear, (2) vignettes and stories that showed how these results mattered to real people, and (3) a strong letter from the agency head that highlights key results and themes.

Weakest Score: Defense

A "Year in Review" section recounts the department's most obvious accomplishments, but very little else in the report makes a case that benefits have been achieved for the American public.

10. Does the agency explain failures to achieve its goals?

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

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

Strongest Scores: Numerous

Six agencies received a 4 on this criterion.

The best agencies in this category are Energy and NASA. The Energy report uses color-coded tables that clearly disclose whether it met or failed to meet its goals. Performance shortfalls are further categorized by the extent to which a goal was missed. Where a target was missed, the report consistently provides a brief explanation as well as a corrective plan of action. The NASA report includes a section captioned "NASA's Performance Improvement Plan" that lists all performance shortfalls, briefly explains the reasons for the shortfall, and describes what will be done to achieve the goal in the future. Transportation's report is particularly strong in explaining performance shortfalls in safety areas and what will be done to redress them.

Weakest Score: Numerous

Nine agencies received a 2 on this criterion. Two types of factors generally contribute to low scores.

First, if a report fails to clearly and adequately document an agency's performance, then it is unlikely to disclose failures. SBA, USAID, EPA, and Defense all have this kind of difficulty.

Second, a report may do a good job of documenting results but fail to identify performance shortfalls. NRC, for example, identified only one shortfall, perhaps because of the "pass/fail" nature of many of its measures (e.g., zero nuclear accidents). NSF subjects its performance to external review, but it appears that the reviewers only assess research projects that NSF staff has already judged successful. Obscure measurement of shortfalls, such as counting a target as "met" even when performance falls somewhat short of the target, can also inhibit disclosure; Education, HUD, and Interior all counted targets as "met" even when results fell short of the target.

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 Government Accountability Office, agency responses indicating an appreciation of threats to its mission and goals, and an anticipation of future risks.

Strongest Scores: SBA, Labor, Energy, OPM

All of these reports received a 4 on this criterion. The reports from SBA and Energy in particular illustrate a number of practices that other agencies may want to emulate.

SBA's report is rich in its treatment of major management challenges. Again this year, the inspector general provides an excellent description of the most serious management challenges, breaking down each challenge into specific actions needed and evaluating the agency's progress on each action using color codes (Appendix 5, pp. 423 ff.). The other portions of the report also devote considerable attention to management challenges. Appendix 5 starts with a crosswalk between the inspector general's management challenges and the strategic objectives to which they apply. Appendices 3 and 4 also describe inspector general and Government Accountability Office reports and recommendations. Page 74 provides a crosswalk between inspector general reports and strategic objectives, and these reports (as well as relevant Government Accountability Office reports) are briefly discussed in the narratives accompanying the applicable strategic objectives. While the report deserves high marks for disclosure in this area, it also shows that the agency continues to have serious management problems. According to the inspector general's assessments, the agency is making good progress in some areas but not in others.

The Department of Energy's discussion of major management challenges uses a format that constitutes a best practice and a model for other agencies (pp. 15-23). The department lists management challenges identified by the inspector general and Government Accountability Office; parses them into significant issues; describes actions being taken to address them as well as actions remaining; and provides an expected time frame for

completion. While the format is outstanding, the content could be improved in some areas. For example, the actions listed to address information technology (p. 19) and human capital (p. 21) challenges are largely process-oriented. The actions listed to improve contract administration (p. 16) seem to be basic steps that any federal procurement operation would take routinely, such as clearly defining needs and expectations for contractors. The inspector general's presentation of major management challenges is concise and includes an assessment of the department's progress in addressing each challenge (pp. 385-87). In all but one case, the inspector general's assessments provide confidence that the department is taking positive specific steps to deal with the challenge.

Weakest Scores: Numerous

Ten reports received a score of 2 on this criterion.

It would help if the agency inspector general presentations of major management challenges were more pointed in their assessment of agency progress. Too many of the inspector general presentations give the agency credit for "making progress" and list steps the agency is taking, but do not really assess the extent of progress or the significance of the agency remedial actions. The most egregious example is the Defense inspector general's presentation, which states that Defense "has made progress" on most listed challenges but provides few specifics to support the statements of progress. By contrast, the Government Accountability Office's recently-updated high-risk report demonstrates that Defense suffers from monumental management problems, many of which have persisted for decades. This report lists a new record of eight individual high-risk problems at Defense as well as six government-wide problems that affect the department. It also lists significant problems at several other agencies. As the table shows, most of these agencies' reports have scored poorly on this Scorecard's "major management challenges" criterion.

Many agency reports fail to respond directly to the inspector general's presentation. Those that do frequently are similar to the inspector general's presentations, reciting a number of remedial actions but not providing much insight into their significance.

12. Does it 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—that is, 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 knowledge gained from the reporting process is actually being used by the agency to revise its priorities and guide its activities? What is the potential for an agency to make a positive difference in the future? How will it realize that potential?

Strongest Scores: Numerous

Seven reports earned a score of 4 on this criterion. There is a correlation between scores on this item and item 10. Agencies that do a good job of disclosing and explaining their performance shortfalls tend to do well in this category—and vice versa.

**AGENCY-SPECIFIC HIGH-RISK AREAS IDENTIFIED
BY THE GOVERNMENT ACCOUNTABILITY OFFICE**

AGENCY	HIGH-RISK AREA (YEAR FIRST LISTED)	SCORE ON THIS SCORECARD'S "MAJOR MANAGEMENT CHALLENGES" CRITERION (1=LOWEST, 5=HIGHEST)
DEFENSE	Supply Chain Management (1990) Weapons System Acquisition (1990) Contract Management (1992) Financial Management (1995) Business Systems Modernization (1995) Support Infrastructure Management (1997) Approach to Business Transformation (2005) Personnel Security Clearance Program (2005)	2
HOMELAND SECURITY	Implementing and Transforming the Department of Homeland Security (2003)	2
HUD	Single Family Mortgage Insurance and Rental Housing Assistance Programs (1994)	2
LABOR	Pension Benefit Guarantee Corporation Single-Employer Insurance Program (2003)	4
TRANSPORTATION	FAA Air Traffic Control Modernization (1995)	2
TREASURY	Enforcement of Tax Laws (1990) IRS Business Systems Modernization (1995)	3
NASA	Contract Management (1990)	2
HHS	Medicare (1990) Medicaid (2003)	Not reviewed

Source: Government Accountability Office, *High-Risk Series: An Update*, GAO-05-207 (January 2005).

Reports scoring highly on this criterion usually offered a discussion of initiatives to improve performance for every performance measure or key performance measure (DOT, Labor, Veterans), or at least for every performance shortfall (Commerce, NASA).

The report from only one agency, Interior, scored well on this criterion while scoring poorly on disclosing failures to achieve goals. The principal reason was Interior's extensive discussion of significant problems that transcend individual performance measures. The secretary's very well written three page transmittal letter acknowledges the many longstanding and difficult issues facing the department, particularly in the financial management areas. The progress in removing 70 percent of the previous year's material weaknesses in the financial report adds credibility to the seriousness of top management's commitment to continuing efforts to resolve financial management problems. Additionally, development of a new departmental strategic plan that embodies meaningful department-wide cohesion for the first time further indicates a commitment to continuing emphasis on improving performance and reporting. The report also includes a discussion of major programmatic challenges the department faces and, to some extent, how the department plans to meet them (pp. 49-53). None of these items directly addresses the types of specific performance shortfalls that would be disclosed under criterion 10, but they are all surely significant issues that deserve discussion.

Weakest Scores: NRC, Defense

The NRC report's score on this criterion exemplifies what can happen when performance measures are constructed in a way that makes shortfalls of a lesser magnitude than, almost literally, a nuclear disaster, virtually impossible. With the possible exception of one missed deadline for an internal decision, the agency reports no failures for any measures. More fundamentally, unless the agency has achieved virtual perfection, its current performance metrics do not effectively serve one of the basic purposes of the Government Performance and Results Act—enabling the agency, its stakeholders, and the public to assess its progress and identify areas for improvement. Overall, the report does not provide confidence that the agency feels a need to scrutinize more specifically or to improve its performance.

The Defense report raised serious questions about how much improvement can be expected in the areas where opportunities for improvement were discussed. In describing specific objectives, there are several instances where action is underway with specific due dates in 2005. However, the message from the deputy secretary is not reassuring.¹³ He states, "During fiscal year 2004, we corrected 11 management control weaknesses, and except for the unresolved weaknesses noted in the Management Discussion and Analysis section of this report, the Department has reasonable assurance that its management controls are effective." Since that section then lists 46 unresolved weaknesses (which is an increase over the number of unresolved weaknesses from the previous year), and the report does not appear to describe the 11 that were corrected, assurances regarding continuing efforts to resolve the remaining issues seem overstated.

¹³ Of all the departments and agencies we reviewed, Defense is the only one that delegated the opening message to the deputy secretary.