

# STRONGEST AND WEAKEST SCORES

## T TRANSPARENCY

TABLE 4: TRANSPARENCY SCORES

	CRITERION				CATEGORY
	1	2	3	4	TOTAL
TRANSPORTATION	5	5	4	5	19
VETERANS	5	5	3	5	18
LABOR	5	5	3	4	17
STATE	5	5	3	4	17
TREASURY	5	4	3	4	16
COMMERCE	5	4	3	3	15
USAID	5	4	3	3	15
SBA	5	4	2	3	14
AGRICULTURE	2	4	4	4	14
JUSTICE	4	4	3	3	14
INTERIOR	5	4	3	2	14
ENERGY	5	3	3	2	13
EPA	5	2	2	4	13
HUD	5	3	2	3	13
NSF	3	3	4	3	13
DEFENSE	5	3	2	3	13
NRC	4	3	3	2	12
EDUCATION	5	2	2	3	12
GSA	2	4	2	3	11
OPM	4	3	2	2	11
NASA	4	3	2	2	11
SSA	1	3	2	3	9
DHS	3	2	2	2	9
HHS	2	2	2	2	8

## 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 members of the public can actually find out what benefits an agency provides. The annual report should be easily available to the public, stakeholders, and the media. Ideally, this means that the agency's home page displays a link clearly guiding the reader to the annual report for the most recent fiscal year. If one has to be an expert on performance management, the agency's structure, or the structure of the agency's web page to locate it, the spirit of accountability to the public is not satisfied. If the report is large, it should be divided into sections for more convenient reading and/or downloading. Making the report available in multiple formats is also desirable, since readers' needs vary and each format has its advantages and disadvantages (e.g., ease of printing, searching, etc.). Finally, 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.<sup>19</sup>

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 [insert link to 2005 Scorecard page].

For years we have been saying that criterion 1 should be the easiest criterion to satisfy, because any webmaster should be able to ensure a good score. This year, reports from 18 agencies proved us right. We received all 24 reports in time for evaluation this year (the first time in several years), and the research team was able to find all 24 reports on the agency web sites by the cutoff date. The scores confirm the widespread improvement here. Fourteen agencies scored a 5 this year, up from just six last year. Nine of the top ten reports received a 5. Perhaps encouragingly, several low-ranked agencies received a 5 on this criterion as well. Another four agencies scored a 4, meaning that the agency had a direct link to the report from its home page but lost a point for lack of contact information. In other words, 18 of the 24 agencies had direct links to the report from their home pages.

### **STRONGEST SCORES: Numerous.**

The Education report is particularly noteworthy. It is designed primarily as an interactive on-line product rather than a hard copy document. It includes a number of hyperlinks and drop-down menus. As far as we know, Education is the first agency to do this. Although it should not be viewed as a substitute for a clear, well-written document, the web-based version is a potential best practice for the future in terms of enhanced accessibility.

---

<sup>19</sup> 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.

### **WEAKEST SCORES: HHS, Social Security, Agriculture, GSA**

HHS and Social Security were the only two reports we had some difficulty finding, although we did eventually locate them on the web sites. Reports from Agriculture and GSA also got a 2 in this category, but not because their reports were particularly hard to find. Both lacked direct home page links, but had links from the home page that were easy to follow. The Agriculture report lacked contact information and the GSA report was downloadable only in a single file. Though “weakest” on criterion 1 this year, these reports are nevertheless much more accessible than some reports were a few years ago, when some agencies declined to put them on the web, and others declined to make reports available anywhere near the deadline.

#### **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.

There was slight improvement in this category. Four agencies got a 5 this year, up from three last year. It’s probably no coincidence that the agencies with a 5 were also the top scorers overall. Several agencies improved to a 4 in this category. Only four agencies got a 2 in this category, including those with the overall lowest scores: HHS and Homeland Security.

### **STRONGEST SCORES: Labor, State, Transportation, Veterans**

One feature that stands out for three of the four best reports is the limited number of performance measures. Labor has just 25 and Transportation has 35. While Veterans has 130 in all, the report focuses on 24 “key” measures. The exception is State, which has 195 measures, although that number is down slightly from 209 for the prior year.

A member of our advisory board noted that the performance reports are much too long, particularly when compared to private sector annual reports. A member of the public surely is more likely to read and understand a report that is relatively concise and built around a manageable number of key measures than the “kitchen sink” versions many agencies produce.

## ANACRONYM-ISMS?

ACRONYM	APPARENT MEANING	PERPETRATOR
PRAGMA	Pacific Rim Applications and Grid Middleware Assembly	NSF
WARS	Worldwide Ammunition Reporting System	Defense
FSU	Former Soviet Union	State
PCP	Potomac Center Plaza	Education
NULO	Negative Unliquidated Obligations	Defense
OPDIVs	Operating Divisions	HHS
SORTs	Status of Resources and Training System	DHS
PIH	Office of Public and Indian Housing	HUD
WCC	White Collar Crime	Justice
HGJTI	High Growth Job Training Initiative	Labor
FBU	Funds Put to Better Use	Interior
BMENA	Broader Middle East and North America	Treasury
LGY	Loan Guaranty	Veterans
DFWP	Drug-Free Work Place	SBA
IWMS/DOWR	Integrated Work Measurement System/District Office Workload Report	SSA
COTS	Commercial Off-the-Shelf	GSA
AO	Abnormal Occurrence	NRC
ABC	Abstinence, Being Faithful and Using Condom Approach	USAID
HCLMSA	Human Capital Leadership and Merit System Accountability Division	OPM
EEIs	Exemplary Ecosystem Initiatives	Transportation
Unbelievably Numerous (Acronym includes numbers that are especially unintuitive)		
C2	Command and Control	Energy
P2	Pollution Prevention	EPA
Daily Double (Acronym with two different meanings)		
SAR	Search and Rescue Selected Acquisition Report	DHS Defense
Trifecta (Acronym with three different meanings)		
IP	Intellectual Property Internet Protocol Improper Payment	Commerce Commerce HHS

*List is intended merely to be illustrative, not exhaustive.*

Some agencies also enhanced transparency by tailoring versions of their reports for different audiences. Education’s interactive on-line report is one example. Another example is State’s publication of a “highlights” version, a brochure, and an interactive CD version in addition to the full hard copy version. Several other agencies also produce separately a summary version of their reports.

### **WEAKEST SCORES: EPA, Education, HHS, Homeland Security**

Discussion in EPA’s report is often quite technical. It takes a four-page appendix to list the acronyms.

Education’s report makes heavy use of acronyms, which, while defined in the text and in an appendix, are still a distraction. More fundamentally, it is difficult to get a handle on the department’s overall performance results since they are scattered throughout the 135-page section of the report. Although a useful format is used to present results for 58 individual measures, there is a lot of detail to go through. A further readability challenge is the plethora of tables that are organized by program goals rather than performance goals or measures. The report could benefit greatly by a summary table that ties the goals and results together in a more holistic and readily understandable way. The introductory portions of the report do highlight some major accomplishments, but they are somewhat selective and not consistently results-oriented. The secretary’s transmittal letter is quite brief and says little about specific departmental accomplishments.

The HHS report presents a number of readability challenges. It is organized more as a collection of reports on the department’s components, which it refers to as “OPDIVs,” instead of a unified and cohesive department level report. It is rife with acronyms (seven pages worth in an appendix) (IV.I.1). The report is text-heavy and often written in technical terms and jargon that are unhelpful to the lay reader. It uses a complex pagination system that makes it hard to navigate between sections.

Homeland Security’s complex organization and plans for further reorganization make its report especially difficult to follow. Annual performance goals sometimes support several intermediate performance goals, but it is not clear what outcome the intermediate performance goal supports. The number of measures (113) presents another challenge, requiring 90 pages to cover them all. Many of the measures are highly technical and difficult to relate to outcomes.

### **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 performance and financial data in the report, disclose any material weaknesses, and explain what the agency is doing to remedy any identified inadequacies.<sup>20</sup> 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.

---

<sup>20</sup> 31 USC 3516 (e).

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

This seems to be one of weaker categories each year. There were no 5s this year (or last year) and only three 4s (versus four last year). By contrast, there were 11 scores of 2 (versus nine 2s and a 1 last year). The reports pay somewhat more attention to data issues now than they did a few years ago. However, many agencies still face major data challenges and could do a much better job of explaining how they are addressing them. On the positive side, most agencies were able to report current results for the great majority of their measures again this year, although they often had to rely on estimates.

#### **STRONGEST SCORES: Agriculture, Transportation, NSF**

Each of these reports received a 4.

The secretary of agriculture's transmittal letter provides reasonable assurance that the data in the report are valid and reliable. Nothing in the body of the report casts doubt on this assurance. The performance section of the report provides a detailed assessment of the data supporting each performance measure that covers data completeness, reliability, and quality (e.g., p. 133).

The secretary of transportation's transmittal letter states that the report's financial and performance data are "substantially complete and reliable" (p. 3). It refers to another section of the report for a detailed assessment of data inadequacies and what the department is doing to remedy them. This section of the report (pp. 129-131) does address these issues, although in somewhat general terms. It references a departmental web site for data source and accuracy statements for each measure used in the report. Another section of the report (pp. 238-272) provides further extensive detail on the completeness and reliability of data for each measure. With respect to timeliness, the report includes fiscal 2005 results for all 35 measures (pp. 66-70). It clearly indicates when results are based on estimates.

NSF appears to do the most thorough job of data validation and verification, although its measurement system is unique. The director's transmittal letter refers to a statement (I-18) that provides assurance that the report's financial and performance data are complete and reliable as defined by OMB guidelines. Moreover, the report describes extensive efforts that NSF undertakes to verify and validate its performance data. Specifically, the agency has, for the past six years, retained an outside firm to conduct a verification and validation review. The report presents the results of the review, which were favorable for all measures reviewed. (See pp. I-13-13; II-87-89; II-92 et seq.) These extensive efforts are particularly important since NSF relies primarily on a qualitative system of performance assessment with input from external reviewers. Nevertheless, there are clearly best practices here that other agencies could adopt, even if they have less direct control over the process that generates their performance data.

### WEAKEST SCORE: Almost Half

Eleven reports received a 2 on this criterion, including all of the reports ranked 17<sup>th</sup> or lower, one ranked in the middle (EPA, 13<sup>th</sup>), and one ranked in the top third (SBA, 7<sup>th</sup>).

Problems include vague or nonexistent descriptions of data verification and validation procedures, vague or nonexistent attestations to data quality in the letter from the agency head, an attestation of data quality that is contradicted by material presented elsewhere in the report, missing data, excuses rather than solutions for data quality problems, failure to cite data sources, disclaimers or qualified opinions from auditors, and data quality problems identified by the agency's inspector general. HHS, for example, reported fiscal 2005 results for only one-third of the limited measures covered in its report, and EPA did not report results for 33 of its 84 annual performance goals. HUD's report states that it "can assess outcomes of a number of programs only in limited ways because of statutory provisions, potential reporting burdens, and privacy concerns" (p. 60). Compliance with GPRA is identified as one of Homeland Security's material weaknesses, based on a lack of supporting documentation for performance measures (p. 93). NASA has received disclaimers on its financial statements for the past three years (p. 230). Social Security's inspector general found the data for five of the seven performance measures audited were unreliable (p. 178). SBA and Defense do not assess data shortcomings and describe steps to resolve them, as required by the Reports Consolidation Act.

#### 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: Transportation, Veterans

Both reports received a 5. Transportation's report illustrates the value of clear and consistent measures over time. The report is rich in baseline and trend data that often go back to fiscal 1999, particularly for the key transportation safety measures. This makes it easy to identify performance trends. The report's narratives also do an excellent job of providing additional context for the measures and targets. The Veterans' report also is rich in baseline and trend data. Another excellent feature is that the tables on performance results in the Veterans' report incorporate the strategic goal for the relevant measures, thus further assisting the reader in assessing progress toward the long-term goal.

### WEAKEST SCORE: Numerous

Seven agencies received scores of 2. In general, here are some recurring problems:

- (1) Measures tend to be in a state of flux from year to year. This lack of consistency in the measures obviously impedes identification of performance trends.
- (2) Some agencies report on whether or not prior year measures were met, but do not describe those measures in terms of targets and actual results. This provides no basis for assessing progress.
- (3) Agencies fail to provide adequate background information to put their annual performance targets in context. This limits the reader's ability to assess progress toward the ultimate goal.
- (4) Significant unexplained disparities between the targets and actual results for the current year (and sometimes prior years) raise questions about the credibility of the measures.



**B PUBLIC BENEFITS**

**TABLE 5: PUBLIC BENEFITS SCORES**

	CRITERION				CATEGORY
	5	6	7	8	TOTAL
LABOR	5	4	4	4	17
STATE	5	4	4	3	16
TRANSPORTATION	5	4	4	1	14
VETERANS	4	4	4	2	14
SBA	4	2	3	3	12
USAID	5	3	3	1	12
COMMERCE	2	2	3	4	11
TREASURY	3	3	3	2	11
AGRICULTURE	3	3	3	2	11
NRC	4	3	3	1	11
ENERGY	2	2	2	4	10
JUSTICE	3	3	3	1	10
EPA	3	2	3	2	10
HUD	4	2	3	1	10
INTERIOR	2	3	2	2	9
NSF	2	2	3	2	9
EDUCATION	3	3	2	1	9
GSA	3	2	3	1	9
OPM	2	2	2	3	9
DHS	3	2	2	1	8
HHS	3	2	2	1	8
NASA	2	2	2	1	7
DEFENSE	2	2	2	1	7
SSA	2	2	2	1	7

## 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 articulate 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 articulated outcome-oriented goals and measures; it is very difficult to do well on these three criteria without having outcome-oriented goals and measures.

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

### **STRONGEST SCORES: Labor, State, Transportation, USAID**

Labor’s report has four strategic goals, each of which is stated as an outcome (p. 13). Each strategic goal is accompanied by several strategic objectives, which the report refers to as “outcome goals” (pp. 41, 75, 97, 123). All of the outcome goals under the first three strategic goals also are clearly stated as outcomes. The outcome goals under the fourth strategic goal are somewhat less clear but could also be described as at least intermediate outcomes (p. 123). The report lists 25 annual performance goals (pp. 14-17). Most of these are likewise stated as outcomes or at least intermediate outcomes.

Transportation’s report has five programmatic strategic goals and one management-related strategic goal. Each is accompanied by strategic objectives. (See pp. 71, 89, 99, 105, 109, and 117.) (The department refers to its strategic goals as “strategic objectives” and its strategic objectives as “strategic outcomes.”) The programmatic strategic goals and their accompanying objectives are highly outcome-oriented. The security goal (“Balance homeland and national security transportation requirements with the mobility needs of the Nation

for personal travel and commerce”) could be made more outcome-oriented, but the objectives under this goal are stated as intermediate outcomes.

State’s report has 11 programmatic strategic goals and one management-type strategic goal. All of the 11 programmatic goals as well as almost all of the 32 programmatic annual strategic goals are stated as outcomes. Moreover, the goals are expressed in clear, non-technical terms so that their nature and importance can be easily grasped by the lay reader. (The strategic and annual performance goals are listed by name on page 21; they are spelled out in more detail in the performance section.)

USAID shares a joint strategic plan with State and thus uses many of the same strategic goals. (State has some additional strategic goals that USAID does not use.) Most of USAID’s annual performance goals are also the same as State’s (pp. 38-40). These strategic and annual performance goals are highly outcome-oriented.

#### **WEAKEST SCORES: Numerous**

Eight departments received a 2 on this criterion: Commerce, Energy, Interior, NSF, OPM, NASA, Defense, and Social Security.

The principal problem in the low-scoring reports under this criterion is that they simply fail to express many of their goals and objectives as outcomes. NASA, for example, pursued 14 strategic goals during fiscal 2005, but virtually all of them are activities; similarly, 78 “multi-year outcomes” are mostly activities. OPM’s measures focus on providing services to other federal agencies rather than assessing whether the services actually achieved the outcomes the client agencies sought. (In contrast, the report from the other agency that primarily provides services to the rest of the federal government—GSA—scored a 3 because the strategic goals at least suggest ultimate outcomes sought by the client agencies.) Commerce’s score of 2 is somewhat unusual, given that its report ranked fifth overall. The strategic goals and objectives allude to some outcomes, but only 10 of 28 performance goals are clearly outcomes.

#### **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**

Each of these reports received a 4; no report received a 5. All four reports present many solid examples of outcome-oriented measures, and any of these reports could have earned a 5 if it had carried the same practice through for all measures.

Of the 29 performance measures under Transportation’s programmatic strategic goals, 26 are outcome-

oriented. All eight measures under the safety strategic goal are solid end outcomes. The measures under the remaining programmatic strategic goals consist primarily of intermediate outcomes, although the measures under the Global Connectivity goal could use further refinement. All of the measures relate directly to their goals and demonstrate impact due to their highly outcome-oriented nature.

State's performance measures consistently relate directly to their goals and to a large extent are outcome-oriented. However, they are not as consistently outcome-oriented as the goals and include a significant proportion of activity or process measures (e.g., pp. 98, 121). Many of the targets describe actions or conditions to meet current needs and require the exercise of judgment in determining the extent to which the target was met.

The Veterans' report characterizes the nature of its key measures and describes only nine of the 24 as "outcome" measures (pp. 6-7). However, a number of what the report describes as "efficiency" measures could be regarded as at least intermediate outcomes, such as timeliness and accuracy of determinations. In any event, most of the key measures relate directly to the outcome and other goals and serve to demonstrate the department's impact. The other measures are a mix of outcomes and activities.

For the most part, Labor's performance measures and targets track outcomes that relate directly to the outcome goals. Some process and intermediate outcome measures can be directly related to improved end outcomes in the future (e.g., pp. 72, 79). However, there is still room to develop more outcome measures.

#### WEAKEST SCORES: More than Half

Thirteen reports earned a 2 on this criterion, usually because most measures focus on activities or processes.

The vast majority of the 246 measures in Energy's report, for example, address activities and outputs. Fewer than 20 of Homeland Security's 113 measures clearly capture outcomes or intermediate outcomes; most of the outcome-oriented measures relate to the Coast Guard or Secret Service. For two of HUD's strategic goals, less than one-third of measures are outcomes, and three other strategic goals have no outcome measures. Fewer than one-third of HHS's measures are ultimate or intermediate outcomes. For example, targets for a measure "to improve health care quality across Medicaid and SCHIP" (the State Children's Health Insurance Program) are to refine strategies and work plans and continue collecting performance measurement data (p. I.23).

Some reports indicate that people at the agency are aware of the problem. EPA's report pledges that for fiscal 2006, about 65 percent of its annual performance measures will track environmental or intermediate outcomes (pp. 16-20). Defense states about half its measures as outcomes but notes that many measures are still under development. Homeland Security notes that it is moving from output to outcome measures (pp. 142-44).

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

### **STRONGEST SCORES: Labor, State, Transportation, Veterans**

Clearly the top four reports all start with outcome-oriented strategic goals and then work downward to design measures and report on performance. This approach yields much clearer reporting than a "bottom up" approach which assumes everything the agency is doing fits somewhere, and then tries to craft strategic goals broad enough to make everything fit. The latter approach is mere marketing, not performance reporting.

State's goals are highly results-oriented. The measures are somewhat less outcome-oriented overall, but they include many outcomes that are consistently related to the goals. The narratives accompanying the individual measures, particularly the description of "impact" for each measure, do a good job of demonstrating the significance of the measure and re-enforcing its connection to the goal.

Transportation's performance metrics are almost exclusively results-oriented, probably more so than any other department or agency. Thus, they serve the department well in demonstrating that its actions have made significant contributions. The narratives in the performance section do a good job of re-enforcing the connection between the department's actions and its intended goals.

The highly outcome-oriented nature of Labor's goals and measures, combined with the excellent narrative portions of the report, combine to produce a very effective demonstration of the department's impact. Of course, as the report acknowledges, the department's impact on the results it seeks to achieve is affected by many external factors such as economic conditions. Interestingly, Labor's report notes that of the 21 department programs that have been assessed under PART, only one was rated "Results Not Demonstrated" — the indicator that a program lacked sufficient performance measures and data to reach a conclusion (pp. 37-38).

Performance metrics in the Veterans' report do a good job of documenting substantive, results-oriented contributions. They yield many specific accomplishments that are described in the secretary's transmittal letter and highlighted elsewhere in the report.

### **WEAKEST SCORES: Numerous**

Nine reports received a score of 2 on this criterion.

Social Security's report illustrates some common problems as well as opportunities for improvement. The agency's performance metrics generally lack a sufficient focus on outcomes to earn a good score on this criterion. Social Security's main outcomes deal with effective customer service and stewardship of Social Security programs. It has some measures that relate directly to these outcomes, such as timeliness and

accuracy of eligibility determinations and payments. However, such measures are limited in number and quality. Some outcome or intermediate outcome measures could be improved to provide a more meaningful indication of contributions. For example, measures focusing on employment of people with disabilities (p. 62) could be expressed as percentages rather than raw numbers. Even some of the activity measures could be improved if they were calculated on a basis other than raw numbers.

In some cases, agency performance discussions appeared to contain contradictions. For example, OPM declares that the goal for a customer service measure—“positive stakeholder feedback”—was met, but the reported result seems more ambiguous: “Stakeholders applaud the new flexibilities, but some complain that the process is outdated” (p. 67).

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

In some ways this is the most crucial criterion of all for effective government accountability. When results are linked to costs, decisionmakers 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 is the criterion on which scores are lowest – a raw average of 1.88, and a weighted average of just 1.26. This is also the category where we have seen the least progress from year to year.

This year’s scoring roughly corresponded to the following standards: 1 if the report had no relevant content or linked budget resources only to strategic goals; 2 if the linkage was carried down only to the strategic objectives; 4 if the linkage was carried down to the individual performance goals; and 3 in cases where there was some useful content beyond the strategic objective level but short of the performance goal level. To earn a 5, the report would need to link budget costs to individual performance measures. No report did this.

#### **STRONGEST SCORES: Commerce, Energy, Labor**

Each of these reports earned a 4.

The body of the Commerce report links costs (and personnel levels) to strategic goals and strategic objectives. Appendix A takes the linkage down to the performance goal level (p. 265). The only missing element is a link between costs and individual performance measures. Energy’s report links budget resources to the seven strategic goals and then extends the linkages to the strategic objectives (“general goals”) and, most importantly, to each annual performance goal (pp. 13, 19, 27, 35). Similarly, Labor’s report links its budget costs to its strategic goals, its outcome goals, and its annual performance goals. All that remains is to link costs to individual performance measures.

#### **WEAKEST SCORES: Half of Them**

Fully half of the reports received the lowest possible score of 1 on this criterion. Another six reports received a 2. The scores in this category were low last year and are even lower this year since we were a little more demanding. Several of our advisors have suggested that we are still too generous in this category. Federal agencies are far from being able to tie their costs to their performance results in a meaningful way.



## FORWARD-LOOKING LEADERSHIP

TABLE 6: LEADERSHIP SCORES

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

### 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.

The research team's evaluation tends to focus here on the introductory portions of the report, such as the transmittal letter, and the performance aspects of the "Management's Discussion and Analysis" (MDA), such as a narrative description of "performance highlights" (if there is one).

#### STRONGEST AND WEAKEST SCORES: Numerous

Every report received either a 3 or a 4 on this criterion. Defense, the only report receiving a 2 in fiscal 2004, moved up to a 3 in fiscal 2005.

Several examples of good transmittal letters stand out: Labor, State, VA, Treasury, and Transportation. In these transmittal letters, the agency heads took the opportunity to highlight how their agencies serve the public and usually included specific examples of fiscal 2005 accomplishments. By contrast, the transmittal letters for other reports are perfunctory and mainly limited to the boilerplate legalisms and required assurances. Examples include EPA, Education, and GSA. The GSA letter even lapses into jargon that is lost on the lay reader. The worst example is Defense. The transmittal letter is not only perfunctory but it is signed by the acting deputy secretary. The secretary of defense remains the only agency head who does not bother to personally sign the performance report. This sends an unfortunate signal.

The reports are generally quite weak in describing the extent of progress toward the fundamental public benefits the agencies exist to achieve. While agencies report on specific accomplishments and success or failure in relation to their specific measures, they rarely address the broader question of whether overall conditions are improving—e.g., are poverty levels and homelessness declining, are Americans healthier, are we safer and more secure from domestic crime and terrorist and other international threats? A few reports provide this broader focus, at least in part, such as Transportation's report on transportation safety trends. However, most do not. A reader could much better assess how well agencies are achieving public benefits if the reports were not just internally focused but provided more information on relevant national trends in their key mission areas.

### 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: NASA, Commerce, Veterans, Labor**

On this criterion, NASA’s report is the strongest. It contains detailed tables that systematically assess each missed target and describe remedial actions (pp. 132-42). If a goal is deemed impractical or infeasible, the table explains why that is and what if any action will be taken.

The Commerce report contains a summary table that clearly discloses the department’s performance results using “traffic light” symbols (pp. 10-13). The body of the report also displays the results in a clear manner using the traffic lights. Appendix A displays the results for each performance measure, again using traffic lights. The appendix also includes for each missed measure a brief statement explaining why performance was not met and describing strategies for improvement.

Veterans’ report discloses performance results, including shortfalls, very clearly. Narratives in the detailed performance section provide some elaboration where targets were missed. The report has a table that goes into greater detail on significant deviations between actual and targeted performance, including causes and resolution strategies (p. 60).

Labor’s report clearly discloses the department’s overall performance results, including missed goals (in tables on pp. 13-16). Additional tables and narratives in the performance section of the report further disclose performance results in a clear and candid manner. The narratives consistently provide explanations for performance shortfalls.

### **WEAKEST SCORE: Numerous**

Eight agencies scored a 2 and HHS scored a 1. HHS fails to report on most of the department’s performance measures and does a poor job of explaining those few performance shortfalls that it does disclose.

### **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.

This year, we continued to scrutinize the reports for specific evidence that major management challenges were being addressed rather than mere expressions of good intentions. As a whole, the scores are low. There was one 5 and there were three 4s.

As in past years, the research team focused on the inspector general (IG) presentations of major management challenges that each report must contain; the agency's response, if any, to the IG presentation; and any discussion of management issues raised by GAO. Once again this year, there was considerable variation in the quality of the IG presentations.

For the first time this year, the research team also considered the agency's scores on the President's Management Agenda, which are included in each report. The "status" rather than the "progress" scores are what count. Our evaluation focused on the areas of human capital, financial management, e-government, and budget-performance integration.

#### **STRONGEST SCORE: Labor**

Labor received a 5 on this criterion – the first time any agency has received a 5 on this criterion since fiscal 2000. The Labor report responds to management challenges identified by the IG and GAO and lists actions it took to address them in fiscal 2005 as well as remaining actions needed and expected completion dates. For the most part, this information is quite substantive and specific. Labor also is the only agency to receive "green" status scores for all government-wide areas on of the President's Management Agenda.

State, SBA, and SSA were the only agencies to score a 4 in this category. The SBA report again constitutes the "gold standard" for an IG presentation. The SBA IG breaks down each major management challenge into the steps needed to resolve it and then uses a color coded system to rate the agency's progress under each of these steps. The assessments are presented in concise tables that are exceptionally informative. The presentations by the Agriculture and EPA IGs are also quite good.

The Energy report also does a very good job of describing major management problems and what the department is doing to resolve them. Its score suffered in this category because the department received a disclaimer of opinion on its fiscal 2005 financial statements. This appears to be a temporary problem.

#### **WEAKEST SCORES: Numerous**

Nine agencies earned a 2. Defense is probably the weakest of all. It has the most pervasive and persistent management problems of any agency. Unfortunately, the report conveys no sense of urgency to resolve the problems. Projected completion dates for some of the problems are 10 years into the future.

Some reports may have been hurt by poor-quality presentations of management challenges by their IGs. Some IGs simply provided laundry-list narratives of management challenges with limited analysis of the challenges and little if any assessment of the agency's progress in addressing them. Interior and GSA are two examples of this.

## 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 agencies scored a 4, including the six top overall scorers. Of these, Energy, Labor, State and Treasury probably are the best.

The Energy report generally does a good job of describing improvement plans both for missed program measures and in terms of resolving major management problems. Another useful feature of the report is a section that follows up on the status of remedial efforts for measures that were missed in fiscal 2004 (p. 171).

Labor's report thoroughly demonstrates that the department takes performance accountability and reporting very seriously and that it is thoroughly engaged in improving its performance results. The report contains thoughtful assessments of results under each performance goal and consistently describes how it plans to do better in the future. Even where targets were met, the report discusses potential future concerns and how they will be addressed (e.g., p. 56). The report demonstrates the same focus on, and commitment to, improvement on the management side.

The secretary of state's transmittal letter is particularly noteworthy in describing her strategic vision for the department. The report follows through by describing changes to improve on specific performance shortfalls as well as management challenges.

Treasury's performance tables consistently describe steps to improve on performance shortfalls. More generally, the report has a good section on broad future challenges facing the department and strategies to address them (p. 26). A section entitled "Moving Forward" accompanies each strategic objective (e.g., p. 36). Appendix F provides good analyses of the department's results under the Program Assessment Rating Tool (PART) and what it is doing to improve them.

### **WEAKEST SCORES: HUD, HHS, Defense**

All three reports received a 2. There is relatively little in these reports describing future changes to improve on specific performance shortfalls.