Annual Performance Report Scorecard Evaluation Criteria

The purpose of the Mercatus Center's assessment of federal agencies' annual performance reports is to ascertain *how well agency reports inform the public about the results they produced.* It is not intended to evaluate the quality of the actual results that federal agencies produced or to determine if the reports adhere to reporting guidelines issued by the Office of Management and Budget. Our focus is entirely on the document's usefulness to the public and to decision makers who are not familiar with (and not necessarily interested in) the details of the Government Performance and Results Act (GPRA) or agency procedures. The scorecard evaluates agencies' annual performance reports according to how well they demonstrate: (a) transparency, (b) benefits to the community, and (c) forward-looking leadership.

Transparency is the key to improving accountability in government. Inherent in GPRA reporting requirements is the principle that the reports should be accessible, readable, current, and useable by a wide variety of audiences, including Congress, the Administration, the general public, the news media, stakeholders, and interest groups. If a report does not make significant achievements and problems apparent, benefits to the community of agency activities will remain unknown to all but a few insiders, and citizens will have no real opportunity to indicate whether or not they approve.

The clear intent of the Government Performance and Results Act is to focus on *benefits to the community*. An agency's value to the public becomes evident only if goals and measures are expressed in terms of actual results. A result is the benefit produced or harm avoided for a particular set of clients or the public at large.

In the absence of information about public benefits, budgets reflect decision makers' wishes rather than a sober assessment of the government's ability to make those wishes come true. To openly demonstrate how agency activities produce meaningful results for the community, performance reports should focus on *outcomes,* i.e. , benefits of programs and activities. They should also clearly present the full cost of achieving those results.

In the past, goals and measures have emphasized activities and procedures, such as levels of effort, expenditures, or internal project milestones. This implicitly assumes that the amount of effort spent on activities automatically translates into results, an assumption that can be incorrect for various reasons. An agency performance report must highlight achievement of results. Otherwise, it does not inform the public of the success or failure of government programs, and budget decisions that rely on reported performance information will fail to reflect realistic assessments of what agencies can accomplish with appropriates.

Finally, *forward-looking leadership* means that the agency uses the information in the performance report to identify solutions to perceived problems and change future plans to capitalize on opportunities for improvement. The report should inspire confidence in the ability of the agency to enhance citizens' quality of life commensurate with the resources and trust they have invested in the enterprise.

For the purpose of judging the individual reports and enabling comparison, these three categories are defined by twelve criteria. To develop the agency score from these criteria, each is worth a minimum of one point and a maximum of five points (total minimum of 12 and maximum of 60). For each criterion, 1 means "poor," 2 is "fair," 3 means "acceptable," 4 is "good," and 5 is "excellent."

The 5-point scale

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

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

Scores reflect the quality of each agency's report relative to other agencies' reports for the same fiscal year. Because scoring is relative, standards rise annually with the level of performance of the "role model agencies" for each criterion. Each year, the best reports raise the bar, establishing a new standard to which all should aspire. Thus, our criteria remain essentially the same from year to year, but the quality of reporting required to achieve a good score continually increases (provide, of course, that the best in each criterion betters the previous year's leader).

What do the scores mean?

5

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

4

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

3

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

2

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

1

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

Explanations of the principles applied in the evaluation process follow the list.

Transparency

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

Benefits to the Community

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

Forward-looking Leadership

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

TRANSPARENCY

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

Access to performance information is critical because public accountability can only be served if the public can find out what benefits the agency is providing. Ideally, this means that the home page at the agency's main web site has a link that clearly guides the reader to the annual performance report for the most recent fiscal year. If one has to be an expert on GPRA and the agency's structure to locate it, the spirit of accountability to the public is not satisfied. If the report is large, it ought to be divided into sections that are easier to follow and read and/or download. Making the report available in alternative formats is also desirable, since readers' needs vary and each version has its advantages and disadvantages (ease of printing, searching, etc.).

Other factors such as timeliness and outreach efforts may affect scores in this area as well. It is highly desirable that electronic and print publications occur simultaneously with official delivery of the report to Congress. Just as a corporation provides an annual report to shareholders, any citizen calling the agency should be able to obtain a copy of the report mailed to them upon request.

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

This annual report is a device for communicating performance to non-specialist audiences. Therefore, its style, language, and subject matter must reflect that purpose. Readers should understand quickly the agency's mission, how it organizes efforts toward that end, and how much progress was made toward its achievement in the recently completed fiscal year. It should read like a corporate annual report to shareholders.

For scoring purposes, the most significant element in this criterion is general organization – whether the contents are clear, logical, easy to navigate, and presented in such a way that their structure aids understanding. Other key elements are consistent format, clarity of text, absence of jargon, minimal use of acronyms, and effective use of visual techniques such as headings, graphs, tables, and photos. It is very important to recognize that details can either inform or confuse, depending on how they are presented.

3. Are the performance data valid, verifiable and timely?

The report should indicate the agency's confidence in the quality of its data. Since the purpose of gathering this data is to manage programs strategically – from a perspective that sees the mission and the execution – the test of adequacy is whether the data are relevant, timely, complete, accurate, and consistent enough to use as the basis for decision making.

Timely performance data should relate to the fiscal year under review. Data not current enough to make appropriate management decisions is problematic. Data should be independently validated (determined appropriate for the associate performance measure) and verified (assessed as reliable). Anyone outside the agency should be able to access the data with relative ease (i.e., Internet access is good, but availability only via Freedom of Information Act request is not). Sources and descriptions should be provided for all outcome data.

In cases where data cannot be released to the public because they are considered confidential under the law, the agency can instill confidence in the quality of the data by indicating that an objective party such as the Inspector General or Government Accountability Office has assessed the quality of the data.

An effective way of demonstrating usability of data is to describe its use for internal management and budget decisions.

4. Did the agency provide baseline and trend data to put its performance measures in context?

The "bottom line" for citizens is whether the situation of concern is getting better or worse. To provide this information, agencies must design a measurement system that facilitates analysis of trends over time. The data should be displayed in a way that allows readers to easily detect and understand their significance. This means that the quality of measures (see #6) will affect this aspect of the reports, and the quality of data presentation will affect the readability of the report (see #2).

To assign scores under this criterion, we consider both quantity (years of data included) and presentation. Meaning matters most in our scoring system, however, so good performance measures that have limited data (due to being new or revised) may score better than inferior measures with more data points. New outcome measure "baselines" may convey more useful information than a decade of data that are not clearly linked to the agency's results. Adoption of better outcome measures is good; replacing invalid measures with other invalid ones is not.

BENEFITS TO THE COMMUNITY

5. Are the goals and objectives stated as outcomes?

An "outcome goal" is defined as the intended benefit (or harm avoided) that results from the 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 verbs like "promote" that emphasize things the agency does instead of why it is doing them should be avoided. This applies at all levels – strategic goals, objectives, and annual performance goals.

Strategic goals should be few in number (three to five). Management goals (including financial performance, human resources, information technology, etc.) exist to support achievement of program goals, so it is inappropriate to place them alongside goals that address public benefits more directly.

Methods, measures, and targets are separate issues that have no place in performance goal statements; they are expected to change, whereas goals (if selected and stated properly) are likely to remain valid over at least several years' time. Overly specific goal statements make trend analysis impossible.

In this area, scoring is holistic, i.e., it involves subjective impressions of results focus as expressed by the simplicity, quality, and "fit" of goal choices and statements with each other and with the mission/vision, rather than a binary system of good or bad and a percentage count.

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

Performance measures selected by the agency should relate directly to its outcome goals. Especially important for our scoring is objective validation of these measures (e.g., by the Inspector General or another qualified, independent entity).

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 performance report may actually detract from its effectiveness in demonstrating an agency's impact. Data measuring levels of output can support a claim for success, but only if the agency makes a compelling case that there is a causal link between the activity and the achievement of results (see #7).

Direct measures, surrogate measures and indicators should tell whether or not intended results were achieved. For instance, if the goal is to reduce suffering from disease, a direct measure would quantify suffering in terms of losses in time, money, and "quality of life;" a surrogate measure might be a nationwide study that included physicians' assessments of severity; and an indicator would be incidence of the disease. None are perfect, but all deal with results. Of course, the causal link must be established here too, if the measure is of an outcome that is only partially or indirectly influenced by agency activity.

Scores on two other criteria (#7, causation, and #3, data) depend heavily on how well this job is done, because if agencies are measuring the wrong things, it will not matter how correct the numbers are. It will be impossible to show that they are making a difference.

7. Does the agency demonstrate that its actions have actually made a significant contribution toward its stated goals?

Claims must be credible. Correlation or association of one activity with another is not causation. The question the report should answer is what effect the agency had, i.e., did things improve because of what the agency did, and if so, how much of the improvement can be attributed to the agency? Causation may be established by program evaluations or other studies, enabling reasonable claims regarding the contribution or agency activity to outcomes. A less desirable alternative would be to furnish a combination of results measures, output measures, and anecdotal evidence that logically support the connection(s).

The agency's case should not rest on merely assumed cause/effect relationships between the agency's work and results valued by the public. The report should explain how agency outputs create or enhance outcomes for the public. It should also describe the nature and extent of influence so that measurements of outcomes can be attributed (at least in part) to specific agency activity. The layman should not be expected to rely on faith alone; it should be easy to understand that the agency's programs influence the results in a positive way. Frank discussion of the operating environment and the extent of the agency's influence is helpful in keeping expectations ambitious, yet realistic.

For "cross-cutting" issues (those involving more than one agency), whatever knowledge exists regarding synergies (or conflicts) and combined impacts ought to be shared with all those whose goal achievement is affected and reported in the appropriate sections of each performance report.

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

Even the simplest breakdown of appropriations according to goals adds powerful information to the managing for results equation, because it shows where priorities lie. Strategic reallocation, however, requires more detail – costs broken down by outcome measures that tell the reader what we are paying per unit of success. Cost information should ultimately be available by goal and objective.

Though admittedly difficult to implement, this is not dependent on wholesale reorganization of accounting procedures and/or acquisition of highly sophisticated costing systems. In other words, this is a matter of allocating total costs, not tracking individual expenditures. At the strategic goal level, these data should be relatively easy to obtain. The rest can be filled in as linkages are made between programs and lower-level objectives and goals. Scores in this area reflect the quantity, timeliness, and apparent utility of cost data.

FORWARD-LOOKING LEADERSHIP

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

This is really about vision – for America by way of the agency. Does the agency know what its "main thing" is? Does the agency realize and articulate the value it provides to the country? The report should speak directly to the public about how the agency's work produces important benefits to each citizen. Just as the best corporate reports feature communication directly from the chief executive, agency reports should demonstrate accountability of senior executives to citizens for their organizations' performance.

This means that although a good letter from the agency head at the front of the report is necessary, it is not sufficient. The lofty ideals must be supported by outcome orientation, sound strategies and successful achievement discussed in the rest of the document. It is impossible to do well here if goals are not good (#5) and/or the agency cannot show its contribution to public benefit (#7). Taken as a whole, the report should create excitement about the opportunities ahead – regardless of actual past performance.

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

Successes and failures that really count are those at the strategic goal and objective level – seldom indicated by single measures. The report should clarify the meaning of lower level achievements or missed targets by including an assessment by agency management of the consequences in terms of higher-level aims. These summaries should take into consideration the fiscal year's priorities and relative significance of different goals, measures and actual results. This is truly transparent disclosure, in contrast to merely listing detailed measures and data from which the reader is expected to draw a conclusion.

Success should be explained so that readers understand why specific targets are chosen. What are the upper ad lower limits of acceptable and achievable performance, and why? If factors outside the agency's control have changed so as to increase or decrease expected benefits, they should be quantified as much as possible. Is this a one-time occurrence, or does it call for a different strategy? Barriers to success in key areas should be identified as internal or external, and solutions revealed if they are of the former variety and suggested if they require action by Congress or other partners.

Special care should be taken with resource explanations to indicate precisely how this would fix the problem and why reallocations were not made internally. If the fiscal year's appropriation did not match or exceed the requested amount, targets should have been adjusted accordingly in the revised performance plan.

If an agency cannot identify reasons for failure, its ability to claim credit for success is suspect. Therefore, scores in this area are linked with #7 (causation). Management challenges may contribute to the internal barriers, so there is a link to #11, also. And since much learning comes from understanding reasons for failure, #12 (changed policies and procedures) is affected, too.

11. Does the report adequately address major management challenges?

Management flaws can affect performance in all programs and activities. The report should describe how risks are being minimized in order to maximize results for citizens. Ideally, the impact of management issues would be integrated with discussions of all relevant factors in achievement of specific goals. Prioritization – identifying which are "mission-critical" – and determination of causal factors should be evident in the discussion.

The Government Accountability Office and Inspectors General routinely identify such issues. However, responding to their published comments merely complies with a minimum expectation. An agency will serve its customers, stakeholders and partners best if it is pro-active in this area, i.e., working to minimize aggravators that might become serious threats if preventative action is not taken. Scoring in this area credits full disclosure, appreciation of threats to agency mission and goals, quality of response and anticipation of future risks.

12. Does it describe changes in policies or procedures aimed at achieving an improved level of results next year?

The intent of the Government Performance and Results Act is not just reporting for its own sake. The legislation enunciates a different way of managing federal agencies that holds them accountable for results rather than activities; the idea is to gather information that informs future action – and use it. Is it evident that knowledge gained from this process is actually being used to make decisions on future activities and priorities? Is innovation occurring? Does the report incite dialogue about future improvement?

Questions that should be answered in this report are: What does the future hold for this area of concern? What is the potential of this agency to enhance public benefit? How is this potential going to be achieved?

The really important evidence is at the highest level, rather than how individual program goals and measures have been adjusted. A report that consists of detailed plans for performance goals but does not show strategic thinking at higher levels will not score well in this area. Significant links exist between this criterion and #7 (causation), #9 (vision), #10 (failures) and #11 (management challenges). Without solid cause/effect knowledge it is arguably impossible to know what adjustments are needed and what will result. It is important to understand what did not work in the past (and why), as well as what barriers exist, before plowing ahead.