



Government Performance Workshop: October 15, 2008

Should PART Be Continued?

Beryl A. Radin
Scholar in Residence
School of Public Affairs
American University

Paper prepared for the Government Performance Workshop
October 15, 2008

*Sponsored by the Georgetown Public Policy Institute (GPPI), Accenture and
OMB Watch*

Should PART Be Continued?

For the past 15 years, the federal government management reform agenda has included several efforts to develop performance measurement activities. These efforts have been designed as attempts to increase accountability in the use of federal funds and have tried to highlight the outcomes that have emerged from the expenditure of those dollars. There are two major initiatives related to performance measurement that illustrate this agenda: the Government Performance and Results Act (GPRA) of 1993 (one of the few management reform efforts that have emerged through legislative activity) and the Program Assessment Rating Tool (PART), the effort undertaken by the Bush administration and begun as a part of the Fiscal Year 2003 budget process. Both of these efforts have attempted to link performance measurement to the budget process.

It is difficult to argue against the goals of these two efforts. Determining whether public funds are actually accomplishing the outcomes expected of the programs and policies those dollars support is both important and commendable. To me, the issue is not whether one supports performance measurement in general (I clearly do) but whether the policies that have been put in place to carry out those goals have been effective. As we think about the arrival of a new administration in January 2009, it is important to determine whether the current approach to performance measurement should be continued. Before we can decide whether or not to continue the efforts, there are a number of questions that must be answered or at least explored.

This paper attempts to outline those questions and to make some suggestions that might be considered by the next administration. It does not deal with a discussion of the goals of performance activity but focuses on the means that have been put in place to accomplish those goals. Given the size and scope of the US federal government, it is not easy to characterize the experience with federal government performance requirements. That experience has produced a highly complex situation that is variable and often idiosyncratic to a particular program or policy. For the purposes of this discussion, I have emphasized the experience under the PART program which has largely eclipsed the requirements of GPRA over the past seven years.

1. Is PART appropriate to address the performance needs of all programs?

While there are a number of reasons that explain the limited ability of efforts such as PART to influence decision-making, one of the most important explanations lies in the inability of OMB to effectively acknowledge the diversity in structures of federal programs. The “one size fits all” approach of OMB attached to PART flies in the face of that diversity. It ignores that diversity (and the typology offered in James Q. Wilson’s classic book, *Bureaucracy*, that describes the differences between programs in terms of their ability to measure outputs and outcomes.) Wilson notes that agencies with the ability to define and measure both outputs and outcomes because of the nature of their work (he calls them production agencies) have the easiest time with performance measurement requirements. By contrast, agencies that have great difficulty devising information systems and agreements over the definition of outputs and outcomes (he calls them coping agencies) have significant problems meeting the OMB

requirements because they live with disagreement about goals of programs, the means of achieving even conflicting goals, and availability of data to measure either outputs or outcomes. Given this diversity, the PART process may be appropriate in some situations but not in others. My concern is with the application of the requirements across the board.

Others have noted that there are some types of programs that have particular problems meeting the PART stipulations. Two are particularly problematic: block grant programs and research and development programs. When one discusses the PART requirements with career civil servants, a number of their concerns revolve around these issues.

Block grant programs. Block grant programs are programs that are not expected to be implemented by federal officials. In fact, they (and other programs that involve third party players) are designed specifically to give discretion to others-- states, localities, and sometimes non governmental organizations. Yet PART assumes that federal officials should be held accountable for achieving a nationally defined set of standards and measures. It is not surprising, thus, that block grant programs have received lower ratings than many other program types. Programs such as the Community Development Block Grant received low scores because OMB did not support the discretion that is given to local officials to determine the approach to community development. Some communities focused on housing, others on infrastructure, others on economic development and still others on community organizing efforts. But OMB thought that this range of choice was too great and communities should be held accountable only for economic development approaches. This was so despite the congressional action to frame the program design in a discretionary fashion.

Research and development programs (R&D). Performance measurement in R&D programs has proven to be challenging for a number of federal agencies. The problems that have emerged have revolved around three areas: ability to measure the efficiency of their research, the appropriateness of using outcome-based measures to assess their contribution, and the ability of R&D agencies to provide information on an annual basis. The culture of science has relied on assessments by peers (e.g. peer review) to determine whether programs are effective or not; concepts of efficiency are not appropriate for those tasks. And we know that the calendar for research involves multiple years. In addition, some of the most effective research results from what might be called "failures." These problems were recently discussed in a study by the National Academy of Sciences entitled Evaluating Research Efficiency in the U.S. Environmental Protection Agency.

Both of these examples indicate how important it is for advocates of performance measurement and management to tailor specific requirements to meet the realities of each agency and program. Performance measures that are effective must make sense to officials within a particular agency; when requirements are devised government-wide they often evoke perverse responses that don't lead to increased attention to performance.

2. Should requirements for performance assessment be government wide?

The very structure of the US system makes it extremely difficult to devise detailed performance requirements that are appropriate for the entire federal government. Yet this is what has been undertaken through PART. Unlike a parliamentary system that provides a structure that allows the executive function to look at the government as a whole, the federal US system is devised to minimize the exertion of concentrated power. As a result, power and authority are separated and shared across all aspects of the political landscape. This occurs through the delineation of separate institutions charged with executive, legislative and judicial functions as well as through the assumption of shared as well as separate powers among the national, state and sometimes local levels of government. It is relevant for us to acknowledge that the US government produces programs and policies that are diverse not only in type but also in terms of the institutions involved. This creates conflict between the legislative and executive branches, fragmentation of responsibilities within the legislative branch through separate appropriations and authorizing bodies, and differentiated responsibilities and roles inside agencies and departments. Variations in the resources and authority available to achieve programmatic results as well as a variation in the level of complexity of tasks and outcome expectations emerge from this structure.

The result – as many have described – is a crazy quilt of program and policy design that reflects policy prescriptions at different points in time, inconsistencies of goals and expectations even across similar programs, and overlapping and conflicting strategies. When one examines the diversity of programs to be implemented in a single agency, it is rare to find consistency across those programs, making it difficult to apply a standardized framework to all elements.

While in a perfectly rational and efficient world one might want to eliminate some of these attributes, it would be necessary to change the structure of the US government in order to do this. The alternative approach would be to craft performance requirements that fit the substance and organizational and institutional context of each program (or, if possible, clusters of programs). Agencies could be required to devise their own performance plans but the expectation would be that significant variations would occur within the federal system.

3. How should performance requirements be implemented?

One of the strongest complaints waged against the implementation of the PART requirements involves the role of OMB in the process. This is an area in which PART implementation is significantly different from the GPRA process. GPRA attempted to involve both the executive branch and the Congress in the process. PART focuses only on the executive branch authority and OMB is the instrument of that authority. The GPRA process was a bottom up approach and most of the performance measures were devised through significant input of program units. PART, in contrast, is effectively a top down process with OMB playing the controlling role of approving the measures. GPRA requirements involved multiple submissions: a strategic plan every five years, annual performance plans, and annual performance reports. PART focused most of its emphasis on the performance measures and looked at programs every five years.

There is a range of problems that have emerged through the PART process. First, it essentially ignored the role of the Congress and operated as if the executive branch (through OMB budget examiners) had the major authority to make these determinations. The result is PART assessments that at times have actually preempted congressional decisions because they did not accord with the views of the White House.

Second, as the role of OMB increased, the specialized perspectives of experts in a program area often became lost in the debate. Focusing on programs through a budget lens is important but does not always allow for an appreciation of the nuances of program realities that are necessary to understand both the constraints and opportunities found in devising and measuring performance. Experience shows that centralization of authority supports a situation in which politics almost always trumps science and professional technical advice. The policy debate on global warming illustrates this problem.

Third, because OMB is an essential part of the White House, it narrows the access of the public to important information because OMB operates mostly out of public view reinforced by its habits of secrecy.

Finally, the centralization of these requirements demoralizes the career public service. These are decisions that should be made by individuals familiar with the details of programs and policies. One understated and under-appreciated story is the dedication and intelligence that resides in many parts of the civil service. A new administration will have to pay attention to this problem.

All of these issues suggest that any performance measurement program that is put in place by the next administration should be attentive to a balance between a centralized approach (the White House and OMB) and a decentralized approach that provides an important role for the program and agency level. Of course the White House has to be concerned about the performance of agencies and presidential policy agendas are crucial and appropriate. However, defining OMB as the controlling player in a complex system does not provide an effective way to achieve the desired ends. Nor can the White House behave as if it had complete control over the agencies, ignoring both Congress and sometimes the judicial branch.

4. Who do we assume will use the results of a performance assessment?

Both PART and GPRA share a confused set of expectations about the users of performance assessment. When GPRA was put into operation it never differentiated between the expectations of a range of potential users of the reports that were issued by agencies and departments in compliance with GPRA requirements. It assumed that a single type of report would meet the needs of program managers, political appointees within agencies, controllers of budgets both inside agencies and in OMB, policy designers in the White House, the range of actors within the Congress, and the general public. The required documents – strategic plans, performance plans and performance reports – were assumed to provide information that all of these players could use to determine whether programs (or agencies) were meeting expected outcomes.

There was very little evidence that many of these players actually considered the GPRA documents as they performed their roles. Rarely were these documents cited in appropriations committee hearings and it was difficult to determine whether either career or political staff within agencies did use them to either modify programs or policies or to justify budget requests. Although there were a few exceptions to this pattern, overall the use of GPRA documents was minimal anywhere in the decision-making process. And there was very little evidence that the public (or even the constituency communities) gave them much attention.

A similar pattern was found in the PART process. Although PART did focus exclusively on the executive branch and not on the Congress, the situation was not much better. It was not surprising that PART assessments were basically invisible in either the appropriations or the authorizing processes in the Congress. But it was also very difficult to determine how OMB and the White House itself used the findings in the process of constructing a budget. There was great variability in the way that OMB budget examiners approached their task and arguments that were made in some program areas that led to budget decreases actually seem to have justified budget increases in other situations. Indeed, the basis for ratings that were given to programs was difficult to figure out.

What is clear is that use of performance information through PART or GPRA by program managers is rare. Indeed, the need for program managers to think about program outcomes seems to have been largely ignored. Low ratings have been used to accuse careers managers of incompetency and not as diagnostic information that could help them make changes to improve the way they implement programs. Although OMB has developed a website that does make the PART ratings available to the public, the rationale for these ratings is hard to determine and most of the assessments have been attentive solely to efficiency measures (ignoring effectiveness and equity program goals).

It is my impression that OMB staff have been unable to point to any examples of program or budget decisions that emerged from the PART process.

If this situation were not complicated enough, the motivation for embarking on a performance assessment journey is also complex. There are at least three agendas that seem to be present: a negative agenda (a search for information that will justify eliminating programs); a neutral agenda (information focusing on changes in the environment that require program modifications); and a positive agenda (information that will provide evidence of program effectiveness). It is often difficult to disentangle the predominant motivation behind establishing a performance assessment agenda.

5. Are resources available that will allow agencies to develop effective performance measurement activities?

Both the GPRA experience as well as that in PART have indicated that the current public service does not have adequate expertise to identify appropriate performance measures for specific programs and determine what information sources are available and required. Despite more than a decade of experience with these two sets of performance requirements, training has not been available to improve the ability of agency staff to meet performance measurement requirements. Performance measurement requires a combination of knowledge of specific

programs as well as familiarity with the technical aspects of performance measurement. Too often agencies depend on outside consultants to perform this work and while they may be of assistance, at least part of the process could be viewed as an intrinsically governmental function that should not be contracted out.

In addition to failures to provide training, other resource limitations are clear. It is rare for an agency to have an opportunity to make a case for the creation of data systems that could produce appropriate information to assess performance. Not only are these systems expensive but the requirements of the Paperwork Reduction Act have provided obstacles to agencies that try to argue for them. And it is usually assumed that agencies will take existing resources from their budgets to pay for the cost of addressing performance requirements.

6. Does it make sense to attach the performance requirements to the budget process?

Both GPRA and PART follow the tradition of some past federal reform efforts that link assessment of performance to the budget process. The GPRA requirements were linked directly to agency submission of budget requests and responsibility for PART was given to the OMB budget examiners who put together program budgets. PPBS, MBO, and Zero-Based budgeting were past efforts to link some form of evaluation of programs to budget decisions. The budget process has its own regularity and is one of the most automatic of decision processes. But, as GAO has noted, some of these past efforts failed because they developed performance plans and measures in isolation from congressional processes.

However, there is little evidence that either GPRA or PART played much of a role in congressional processes. There is an argument to be made that the annual budget calendar does not provide a venue that allows for a consideration of the details of program implementation. When the conversation is around the size of budgets, it is difficult to raise substantive policy issues. Much of the performance rhetoric speaks of the budget process as if it were a simple, well-defined set of activities that produces rational allocation patterns. This approach removes issues of values and political choices from the process and also ignores the multiple functions of budgeting (include both budget execution and budget creation).

There are, however, other ways to use performance information outside of a direct link to the budget process. The system that has been put in place in the United Kingdom to rate the performance of local government through the Audit Commission is not linked to the budget process. Instead the information that is developed is released to “name and shame” the government unit that has been examined.

It would be useful for a new administration to at least consider the possibility of decoupling the budget process and the performance assessment process.

Conclusion

This paper has attempted to raise a number of issues that should be considered before a new administration decides how to structure a performance measurement process. It is clearly

important that some process be put into place but the experience with both PART and GPRA suggests that a number of issues should be examined before embarking on that same path.

As I have noted, an effective system should have the ability to respond to a number of issues:

- The diversity of program designs and forms
- The importance of providing flexibility for agencies to craft measures and identify data sources
- Development of a system that will be useful to program managers
- Methods of dealing with appropriate congressional players
- Devising White House roles through use of interagency task forces, training programs, venues for agencies to share experience, and transparency