

## The Program Assessment Rating Tool and its Role in Evaluating the Effectiveness of U.S. Government Programs

Philip G. Joyce and Alice Levy  
The George Washington University

Over the past 50 years, the United States government has engaged in a number of efforts designed to focus on the results of federal agencies and programs. Starting in the 1950s with “performance budgeting” and continuing through the 1960s and 1970s program budgeting and zero-based budgeting efforts, the goal of these has been a closer linkage between resources provided and results achieved. In 1993, during the administration of President Bill Clinton, the Government Performance and Results Act (GPRA) was signed into law. This act required federal agencies to engage in strategic planning, performance planning, and performance reporting.

When the George W. Bush administration came into office, they inherited this legacy of attention to performance, but attempted to expand it, in particular to emphasize the use of performance data rather than its production. The Bush effort has two parts. The first, embodied in the President’s Management Agenda (PMA), is designed to encourage federal agencies to improve their management across five areas—financial management, human capital management, E-government, contracting out, and budget/performance integration. Agencies are evaluated according to standards established by the White House, in each of these areas. These scores are translated into “traffic lights” (green, yellow, and red) for each.<sup>1</sup> The second Bush administration initiative, called the Program Assessment Rating Tool (PART), is the focus of this article. The PART, unlike these past initiatives, focuses very much on evaluating programs, rather than whole agencies.

In the article that follows, we will describe the background of the PART, discuss its use and impacts, and offer conclusions as to how to make systems similar to the PART more effective in practice.

---

**Philip Joyce** is Professor of Public Policy and Public Administration at The George Washington University. Alice Levy is a Ph.D. student in public policy and administration, specializing in public budgeting, at The George Washington University.

<sup>1</sup> For more information on this and for recent scores, see <http://www.omb.gov>.

## PART: Overview and History

The United States' Bush administration's Program Assessment Rating Tool (PART) is an attempt to assess the performance of all federal programs using a single set of criteria. As such, the PART represents a departure from traditional performance measurement practices because it aims to facilitate the comparisons of programs with very different missions. Under traditional performance measurement and performance budgeting systems, policymakers must grapple with tough decisions such as how to compare homeland security attacks averted to infant mortality decreases. Even with programs that have similar missions, how does one compare decreases in mortality due to heart disease to decreases in mortality due to cancer? The PART attempts to surmount these hurdles by creating performance scores that can be compared across program areas and represents a positive step forward in performance budgeting practices. Before the lessons of the PART could be applied to any other country, it is necessary to understand how the PART works and what its limitations are.

The PART represents an effort to facilitate performance budgeting and as such, has roots in previous reforms of the American budgeting system. In the United States, the origins of performance budgeting systems lie in the progressive movement's efforts to promote the study of government programs and activities. Early reforms include: the 1921 Budget and Accounting Act, which represented an attempt to limit government spending while making budgeting more transparent; the Budget and Accounting Procedures Act of 1950 that created "performance budgeting" (which, contrary to its name, focused mostly on outputs, not results); and the planning-programming-budgeting system (PPBS) of the 1960's.<sup>2</sup> Successfully pioneered in the U.S. Department of Defense by then-Secretary, Robert McNamara, PPBS was President Lyndon Johnson's attempt to create a planning-programming-budgeting system. Although it was intended to be employed throughout the entire federal government, difficulties in transferring the practice to civilian agencies caused the requirements to lapse in 1971.<sup>3</sup> During the 1970's efforts to integrate information on governmental performance and the allocation of resources continued including management by objectives (MBO) and zero-based budgeting (ZBB), but in the 1980's government reform receded from public attention.<sup>4</sup> Recent years have witnessed additional federal reforms, most notably the Government Performance and Results Act of 1993<sup>5</sup>.

---

<sup>2</sup> Melkers, J.E. & Willoughby, K. G. "Budgeters' Views of State Performance Budgeting Systems: Distinctions Across Branches" *Public Administration Review*. (2001).

<sup>3</sup> Joyce, P.G., "Performance Based Budgeting." In *Handbook of Governmental Budgeting* edited by R. T. Meyers (San Francisco, California: Jossey Bass, Inc., 1999).

<sup>4</sup> Ibid.

<sup>5</sup> Irving, S. "Performance Budgeting: PART Focuses Attention on Program Performance, but More can be Done to Engage Congress." United States Government Accountability Office. (2005).

<http://www.gao.gov/new.items/d0628.pdf>

As part of the fiscal year 2002 budget process, President Bush initiated the Presidential Management Agenda (PMA), a series of reforms aimed at improving the performance of the U.S government. These reforms include the following:

- Strategic Management of Human Capital: Reforms under the strategic management of human capital include restructuring workforces to increase the number of front-line employees, reducing organizational levels, and utilizing information technology to create knowledge management systems that can pass on the knowledge of retiring employees to new employees.<sup>6</sup>
- Competitive Sourcing: These reforms include identifying government activities that could be outsourced, increasing the number of activities that are outsourced, and improving the competitive environment by subjecting contracts to renewed competition every three to five years.<sup>7</sup>
- Improved Financial Performance: To improve financial performance, the Bush administration is requiring agencies to assess the full extent of their erroneous payment rates and implement accounting and reporting reforms to reduce these rates and make financial information more timely and accessible.<sup>8</sup>
- Expanded Electronic Government: The administration is committing to increase the use of information technology in government with practices such as e-procurement, e-grants, e-regulations, and e-signatures. These reforms should increase the accessibility of government to citizens and business, improve information sharing between governments, and reduce internal costs by automating internal processes.<sup>9</sup>
- Performance Improvement: In addition to working with the Office of Management and Budget (OMB) on identifying programmatic objectives, agencies are expected to develop and track data on outcome measures, integrating these measures with cost information. The ultimate goal of the initiative is to align performance measures with funding, eliminating or reforming programs that are not performing, while reinforcing successful programs.<sup>10</sup>

A major component of the PMA, the goal of the PART process is to assess and improve program performance. By identifying program strengths and weaknesses, the PART

---

<sup>6</sup> Executive Office of the President Office of Management and Budget. 2002. "The President's Management Agenda." <http://www.whitehouse.gov/omb/budget/fy2002/mgmt.pdf>, p.14.

<sup>7</sup> Ibid, p.18.

<sup>8</sup> Ibid, p.20.

<sup>9</sup> Ibid, p.23.

<sup>10</sup> Ibid, p.27.

is designed to be used as a management tool and to allocate resources between programs more effectively.

The creation of the PART required the Office of Management and Budget to come up with a listing of federal programs. Perhaps surprisingly, there was no established listing of “programs” in the federal budget. The budget is organized variously by broad “functions” (national defense, health, etc.) or agency (the Department of Defense or Department of Health and Human Services, etc.) or appropriation account (there are about 1200 of those, some of which are very large and contain many programs, and other of which may be small, single-program, accounts). None of these equated precisely to programs. For this reason, OMB initially developed a listing of approximately 1000 programs after discussion with the relevant Departments and agencies. The listing of programs as developed by OMB has continued to be refined and updated, which means that in some cases, “programs” have either expanded or ceased to exist over time. The Congressional Research Service (CRS) notes that in defining programs OMB relied on the budget accounts used in the president’s budget, an approach that has been criticized as many program definitions are inconsistent with agency organization and strategic planning structures.<sup>11</sup>

Once this list was developed, OMB stated its intent to evaluate approximately one-fifth of these programs each year; by year 5, all 1000 would be evaluated.<sup>12</sup> The PART assessment is done using a questionnaire, which is divided into four parts and then scored in the following categories:<sup>13</sup>

1. program purpose and design, attempting to determine whether program objectives and missions are clear and appropriate (weighted 20 percent);
2. performance measurement, evaluations, and strategic planning, to find out the extent to which the agency sets itself valid goals, both long-term and annual (weighted 10 percent);
3. program management, which focuses on assesses financial management and program improvements (weighted 20 percent); and.
4. program results, which program results, looks at the actual results agencies have achieved for their performance measures (weighted 50 percent).<sup>14</sup>

---

<sup>11</sup> Brass, C. “The Bush Administration’s Program Assessment Rating Tool (PART).” Congressional Research Service. November 2004.

<sup>12</sup> This is a cumulative assessment. In other words, in Year 1 200 programs would be evaluated, then 400 in year 2 (the Year 1 programs plus 200 more), 600 in Year 3, etc.

<sup>13</sup> Office of Management and Budget. “Program Assessment Rating Tool.” <http://www.whitehouse.gov/omb/part/index.html#background> (accessed November 7, 2007).

<sup>14</sup> Office of Management and Budget. 2004. “Rating the Performance of Federal Programs.” <http://www.gpoaccess.gov/usbudget/fy04/pdf/budget/performance.pdf> (accessed November 7, 2007).

## The Program Assessment Rating Tool & its Role in Evaluating the Effectiveness of US Government Programs

Scores are assigned for each segment on a scale from 0 to 100 and totaled. Based on the numeric score, programs can be classified as effective, moderately effective, adequate, or ineffective, as indicated in Table 1:

Table 1: Ranges for PART Ratings

<b>Rating</b>	<b>Range</b>
Effective	85-100
Moderately Effective	70-84
Adequate	50-69
Ineffective	0-49

In addition, regardless of the score, OMB evaluates a program as “Results Not Demonstrated” in cases where the program does not have acceptable long term and annual performance measures.<sup>15</sup>

Table 2 shows the number of programs evaluated, and the distribution of ratings of these programs, over the five years of the PART thus far. As the table indicates, a progressively larger number of programs have been evaluated each year, and the scores, in general have been improving. Of particular note is the substantial reduction in the “Results Not Demonstrated” category (50 percent of all programs were in this category in 2002, down to 22 percent in 2006). By 2006, almost half of all programs evaluated received a score of moderately effective or higher, up from 30 percent in 2002.<sup>16</sup> All PART assessments are readily available to the public through the OMB’s database, ExpectMore.gov - <http://www.whitehouse.gov/omb/expectmore/>.

---

<sup>15</sup> Ibid.

<sup>16</sup> Office of Management and Budget, PART Training Slides, at [www.whitehouse.gov/omb/part/training/2007\\_training\\_slides.pdf](http://www.whitehouse.gov/omb/part/training/2007_training_slides.pdf).

**Table 2:** PART Ratings, 2002-2006 (FY2004-FY2008 Budgets)

Year	Programs	Effective	Moderately Effective	Adequate	Ineffective	Results Not Demonstrated
2002	234	6%	24%	15%	5%	50%
2003	407	11%	26%	20%	5%	38%
2004	607	15%	26%	26%	4%	29%
2005	793	15%	29%	28%	4%	24%
2006	977	17%	30%	28%	3%	22%

In addition, when OMB compared the 2006 scores with the 2002 scores for those 234 programs evaluated in 2002, they found that the Results Not Demonstrated category had declined from 50% to 14%, and the Effective and Moderately Effective categories, taken together, comprised 48% of all programs in 2006, up from 31% in 2002.<sup>17</sup>

In addition to the distribution of the ratings themselves there is the question of how the ratings have been used in the budgeting process. In the Bush administration’s proposed 2008 budget, there is circumstantial evidence suggesting a relationship between PART scores and funding levels. An analysis of this relationship suggested that higher PART scores translated into more money requested in the budget. This can be demonstrated in two ways. First, 62.7 percent of “effective” programs, and 55.5 percent of “moderately effective” programs were recommended for an increase in 2008 (over 2007), compared to only 15.3 percent of “ineffective” programs. Second, the mean recommended percentage increase for effective and moderately effective programs was in excess of 9 percent. Ineffective programs, on average, had a recommended DECREASE of 34 percent.<sup>18</sup>

Relying on similar data, but using regression analysis researchers have also found evidence that there is a relationship between executive funding recommendations and PART scores. Gilmour and Lewis found that PART scores are positively associated with traditionally Democratic programs, although they did not observe a relationship between traditionally Republican programs and the scores.<sup>19</sup> In a study for Mathematica Policy Research, Inc. the authors found a substantively and statistically significant relationship

<sup>17</sup> Ibid.

<sup>18</sup> We are grateful to our colleague Joseph Cordes for this analysis of the effect of PART ratings in the fiscal year 2008 President’s budget (based on PART scores for calendar year 2006 compared to recommended funding levels in the budget for fiscal year 2008).

<sup>19</sup> Gilmour, J. B., & Lewis, D. E.. “Does Performance Budgeting Work? An Examination of OMB’s PART Scores.” *Public Administration Review*, (September/October 2006).

between PART scores and executive funding. For a one standard deviation increase in PART scores, funding increased by 9 percentage points, although this effect is dependent on program size – with average effects of 20 percentage points for small programs and average effects of only 3 percentage points for large programs. The authors estimates are consistent with earlier research from the Government Accountability Office (GAO) and remain robust when including control variables for the type of program, the sponsoring agency or department, and estimated funding of the previous year (included to account for administration priorities).<sup>20</sup>

While the forgoing evidence suggests that PART scores have an impact on the executive budget request, in the U.S. budget system, the power of the purse rests with the separate legislative branch where evidence of PART score utility is much lower. A GAO study found that the use of PART scores by Congress was hampered due to OMB's failures to consult with Congress early in the process, explain its methodology, and communicate the information in an appropriate manner.<sup>21</sup> While PART proponents have often viewed the process as an augmentation to strategic planning and performance budgeting requirements mandated in GPRA, opponents argue that the PART is a political tool designed to shift power to the executive branch.<sup>22</sup> The fact that OMB's program definition is inconsistent with the program definitions used in GPRA strategic planning requirements suggests that the ability of the PART process to augment GPRA requirements may be limited.

One major obstacle to integrating PART scores with the Congressional budget process is the incompatibility of programs as defined by OMB and the appropriations accounts that drive Congressional budgeting. For the current article, the authors attempted to look at the relationship between 2006 PART scores and the Fiscal Year 2008 appropriations bills, as passed by the House and Senate in calendar year 2007. Out of a random sample of 90 evaluated programs, there are only 13 cases in which there is an obvious, direct link between programs as defined by OMB and the appropriations accounts described in Congressional Committee reports.<sup>23 24</sup>

Challenges in linking Congressional funding to the PART evaluations come in a number of forms. In some cases, the OMB has divided programs according to their purpose, while Congress divides programs jurisdictionally. For example, the OMB defined

---

<sup>20</sup> Levy, D. & Olsen, R. "Program Performance and the President's Budget: Do OMB PART Scores Really Matter?" Mathematica Policy Research, Inc. October 2004.

<sup>21</sup> Irving, S.. "Performance Budgeting: PART Focuses Attention on Program Performance, but More Can Be Done to Engage Congress." Government Accountability Office. October 2005

<sup>22</sup> Brass, C. "The Bush Administration's Program Assessment Rating Tool (PART)." Congressional Research Services. November 2004.

<sup>23</sup> We are grateful to Robert Shea from the OMB for assistance in compiling a list of all PART evaluated programs.

<sup>24</sup> U.S. Congressional committee reports attained from Congressional Quarterly's CQ.com on Congress. Committee reports were searched electronically using CQ's keyword search tool.

“International Narcotics Control and Law Enforcement Programs, South Asia” as a State Department program, but in the appropriations bills, funding is listed either by nation or in the aggregate. Domestically, in the Department of the Interior, OMB defined “Fish and Wildlife Service - Habitat Conservation” as a program, but in the appropriations accounts, funding is divided by state or local government recipients. In other cases, problems in linking congressional funding to PART scores arise because the House and Senate report appropriations accounts differently. Finally, linking funding to PART scores is particularly problematic in the case of programs with multiyear authorizations that are not subject to the annual appropriations process, as is the case with the Bank Secrecy Act Administration, which was another program as defined by OMB.

While the above complications do not render a link between Congressional funding and PART scores prohibitive, they vastly increase the amount of time required for such comparisons. Given the crowded agenda of Congressional budgeters, particularly those committees with yearly appropriations responsibilities, the potential for members of Congress to use PART scores during appropriations deliberations is seriously compromised as a result of the way OMB has chosen to define programs for its analysis. Several committee reports, both from the 109<sup>th</sup> Congress (when the Republicans were still in charge) and the 110<sup>th</sup> (when the Democrats has assumed control), indicated that PART details in agency budget justifications had made the review of yearly budget requests more difficult as PART data has replaced and not supplemented traditional data with which members are more familiar:

“Unless specifically exempted, no funds are provided in this Act to conduct or participate in the conduct of a PART analysis or study unless the Committees on Appropriations of the House and Senate have approved of the study, inclusive of the data on which the analysis will be based, the methodology to be employed and the relative weight of each of the four factors that will be assigned to the study in determining a final score.”<sup>25</sup>

“The [PART] process has failed largely through the inability of the administration to establish meaningful benchmarks and program goals that can be used as a valid measure for the success of a program and its funding requirements/need.”<sup>26</sup>

---

<sup>25</sup> U.S. House, Committee on Appropriations. “Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriations Bill, Fiscal Year 2007”. Accessed at <http://www.cq.com.proxygw.wrlc.org/display.do?dockey=cqonline/prod/data/docs/html/commreport/109/commreport109-000002308390.html@allbillsarchive&metapub=CQ-COMRPTS&searchIndex=0&seqNum=9>

<sup>26</sup> U.S. Senate Committee on Appropriations Report on Fiscal Year 2008 Appropriations for Transportation, Housing, Urban Development, and Related Agencies. July 16, 2007.



“OMB and Federal agencies have tended to accommodate an increasing amount of PART performance data in the budget justifications by eliminating fundamental and objective programmatic budget data that is critical to the work of the [Appropriations] Committee. This trend has made it increasingly difficult for the Committee to perform a meaningful review of budget justifications, including the ability to conduct necessary budget oversight work as well as the ability to reach valid and comprehensive funding decisions”<sup>27</sup>

“Most [agency budget] justifications continue to be filled with references to the Program Assessment Rating Tool (PART), drowning in pleonasm, and yet still devoid of useful information... The Committee finds little use for a budget justification which does not reveal specific details of the measurable indicators and standards used to evaluate a program’s performance, relevance, or adherence to underlying authorization statute. Further, the Committee has little patience for secretaries and administrators who cannot explain the rationale behind a program’s funding level other than ‘the PART score,’ ‘getting to green,’ or ‘this is what OMB provided.’”<sup>28</sup>

Given the forgoing expressions of Congressional resentment of the PART, it is unlikely that there is a strong relationship between PART scores and legislative funding. Table 3 lists the 13 programs for which appropriations data could be readily traced to 2006 PART scores, along with the House and Senate recommended changes in funding. While the three programs rated “effective” all received recommended funding increases, there does not appear to be a relationship between the other scores and recommended funding levels. Given the small sample size, these examples are not intended to be representative but rather a venue for future research.

## **Conclusions**

There are many possible conclusions that might be drawn from the U.S. government’s experience with the PART so far, but four seem particularly salient in the current context.

1. It is important to attend to the relationship between performance measurement and budget structure in devising any new system. The PART process is well aligned with the executive budget in the United States and evidence suggests that OMB and the president use PART

---

<sup>27</sup> U.S. Senate Committee on Appropriations Report on Fiscal Year 2008 Appropriations for Financial Services and General Government and Related Agencies. June 13, 2007.

<sup>28</sup> U.S. House Committee on Appropriations Report on Fiscal Year 2008 Appropriations for Financial Services and General Government and Related Agencies. June 22, 2007.

scores when making funding decisions. However, there is limited evidence that Congress finds the PART scores at all useful, largely due to the fact that reporting processes are incompatible with existing budget structures. For other countries with constitutionally fragmented budget processes, the design of a system similar to the PART should begin by asking who is the most important intended audience.

2. Understanding the desirable relationship between funding and performance is critical to determining the success or failure of performance budgeting systems. In the literature on the use of performance information, a frequently used method (one also employed in the present study) is to attempt to link performance scores with funding levels. This approach, however, rests upon the premise that it is a *good* thing to create a direct linkage between performance and funding. The OECD characterizes performance budgeting as falling into one of three categories: presentational, informed, and direct linkage. With presentational performance budgeting, performance information is collected and reported in the budget process, but not used during resource allocation. Informed performance budgeting indicates that performance information is used, but that the relationship between funding levels and performance scores is not always positive. Direct linkage performance budget occurs when there is always a positive relationship between performance scores and funding levels; and is very rare among OECD countries.<sup>29</sup>

One of the reasons that direct performance budgeting is rare, is that its normative justification has yet to be established. First, poor performance may be an indicator of inadequate funding, a problem that will not be remedied when additional funding cuts are made. If the program is important to the citizens, increased funding is the most likely way to improve outcomes. An additional risk of direct performance budgeting is that it will evolve into incremental budgeting, though at a significantly higher cost. Each budget cycle agencies report incremental performance increases and receive incremental budget increases. Although no real evaluation or analysis of performance data occurs, the costs of collecting the performance information is still born by the government. Thus, directly tying performance information to funding levels may have negative unintended consequences if executed with insufficient attention to detail.

---

<sup>29</sup> Curristine, T. "Performance Information in the Budget Process: Results of the OECD Questionnaire." OECD Journal on Budgeting, 5(1608-7143) 2005.

**The Program Assessment Rating Tool & its Role in Evaluating the Effectiveness of US Government Programs**

Table 3: Comparisons of PART Scores and Congressional Funding Levels for 13 Comparable Programs

Program Name	PART Score	House Recommended Change in Funding (%)	Senate Recommended Change in Funding (%)
Black Lung Clinics	Ineffective	0.0	1.9
Child and Adult Care Food Program	Adequate	5.4	5.4
U.S. Fire Administration	Adequate	4.7	4.7
Education for Homeless Children and Youths	Adequate	8.1	8.1
Contributions to International Organizations	Moderately effective	13.1	17.4
Early Reading First	Moderately effective	-2.7	0.0
Government National Mortgage Association	Moderately effective	0.0	-9.8
Consumer Product Safety Commission	Effective	6.6	12.0
Indian Housing Loan Guarantees	Effective	24.2	24.2
International Boundary and Water Commission	Effective	200.1*	1590.1*
Commission on Civil Rights	Results not demonstrated	0.3	0.3
Compassion Capital Fund	Results not demonstrated	0.0	-16.7
Delta Regional Authority	Results not demonstrated	51.5	51.5

\* These very large increases can be traced to the fact that both the House and Senate included capital funding to build a fence between the United States and Mexico with the program's budget.

3. While it may seem desirable to employ a system that is as comprehensive as the PART (that is, one that looks at all or most programs), a “one size fits all” approach can impede the ability of government officials and citizens to understand the differences between programs. As noted above, the PART uses a standard questionnaire and set of scores to compare programs to one other with the benefit of allowing programs that are quite dissimilar to be compared. The down side of such an approach is that it may force programs into a methodological “box” that inhibits the understanding of complex differences between programs. This is particularly true under the PART as many of the questions used in the evaluation are of the “yes/no” variety, when the real answer in many cases is “maybe” or “sometimes”.

4. While the PART fits firmly within the recent management tradition of the federal government, it is not clear whether it will survive into the next U.S. Presidential administration. Since some of the leadership of the Congress has been openly hostile to the PART, and since it is so closely identified with the Bush administration, the next President may be hesitant to continue this specific program. This is particularly true because there is suspicion (borne out by some of the research cited earlier) that the PART is just a sophisticated way to collect ammunition to cut some domestic programs that are opposed by the Bush administration. On the other hand, recent trends in federal management would suggest that the next President (whoever he or she may be) is likely to embrace some management reform ideas that will be consistent with the stated goals of the PART—that is, improving the understanding of policymakers concerning what works and what doesn’t.

In the end, the PART is likely to be seen as another in the long line of performance-oriented reform efforts that has improved the availability of information on performance, and has also encouraged its use by both federal politicians and managers. But such an approach has inherent problems, not the least of which are a conceptual understanding of exactly how to link budget and performance data, and the creation of incentives for people to use the information that is produced by such systems. In the United States, performance-informed budgeting is very much a work in progress.