



PART Backgrounder

April 2005

Adam Hughes

Budget Policy Analyst

J. Robert Shull

Director of Regulatory Policy

Celebrating 20 years: Promoting Government Accountability and Citizen Participation - 1983 - 2003.

1742 Connecticut Ave NW
Washington, DC 20009



tel: 202.234.8494
fax: 202.234.8584



email: ombwatch@ombwatch.org
web: <http://www.ombwatch.org>

PART: Background

Introduction/Background on PART

For more than five years, the Bush administration has focused a good portion of its rhetoric on performance, accountability and results. To that end, in 2001, the Office of Management and Budget (OMB) began to develop a mechanism called the Program Assessment Rating Tool (PART) to help budget examiners and federal managers measure the effectiveness of government programs.

PART was designed under the “Budget and Performance Integration Initiative,” one of five Presidential management initiatives, and its stated intention is to build upon the Government Performance and Results Act (GPRA), although it remains unclear how exactly it relates to and interacts with GPRA.

OMB initially tested the tool on a limited number of programs (67) in the spring of 2001 in preparation for the FY 2003 presidential budget. The tool was then reviewed by the Performance Measurement Advisory Council, an ostensibly independent committee that nonetheless was slanted and lacked stakeholder representation or public interest perspectives. OMB made some minor revisions in the tool as a result of this review process, mostly to the requirements for data, but the format remained generally the same.

Each year since, the PART has been used on 20 percent of all government programs/activities (approximately 200/year). To date, it has been used to review 60 percent of federal programs, and all programs will have been reviewed once by 2007.

The Nuts and Bolts of PART

The PART consists of six questionnaires designed for different government activities — competitive grant programs, block/formula grant programs, regulatory-based programs, capital assets and service acquisition programs, credit programs, research and development programs, and direct federal programs. Essentially, it consists of “yes” and “no” questions, although in the “results” section of the tool there are some additional gradations.

The programs are given a raw score based on four areas (program purpose and design, strategic planning, program management, and program results). The actual program results themselves constitute only half of the total score, the rest of which is based on internal processes. Once the raw score is calculated for each of these four areas, the final score is translated into one of five ratings — either “effective,” “moderately effective,” “adequate,” “ineffective,” or “results not demonstrated.”

Interaction of PART and GPRA

The stated objective of the PART is to get consistent and unbiased performance data from government programs and use that information to make budget decisions. Unlike GPRA, which has many objectives and represents a more thoughtful and comprehensive approach to improving government, the PART is very narrowly focused in many different ways.

Arguing that GPRA has fallen short of its well-intentioned goals — most specifically in failing to link performance data and budget decisions — President Bush rapidly moved forward with the PART. In some ways, PART seems to be an effort by the President to “fast track” the more complicated and iterative GPRA process to quickly arrive at an oversimplified rating of government programs that can justify budget decisions. In the worst case scenario, PART is a tool designed to by-pass the broad effort of GPRA and Congress and go directly to making budget decisions based on a narrowed, simplistic, and potentially biased measures of performance.

PART and GPRA appear to be redundant functions in the federal government — an ironic twist as each was meant to promote accountability and efficiency. The Government Accountability Office has concluded

that PART is a parallel and competing approach with GPRA's Performance Management Framework and expressed concerns the increasing emphasis on the PART would shift agency focus and come to drive the strategic planning processes in the federal government (GAO-04-174, January 2004). GAO concluded, "Many [agency officials] view PART's program-by-program focus and the substitution of program measures as detrimental to their GPRA planning and reporting processes." The relationship between the PART and GPRA is not well defined, is often confusing to program officials and agency managers, and undermines the efforts of both to promote efficiency and accountability.

Arbitrariness, Subjectivity, and Inconsistency in Scoring

The GAO has also found, despite efforts by OMB throughout the development and implementation of the PART, the assessments still ended up requiring a large amount of judgment and interpretation by OMB budget officers. Their report concludes, "By using the PART process to review and sometimes replace GPRA goals and measures, OMB is substituting its judgment for a wide range of stakeholder interests." This leads to inconsistencies in the implementation of the PART and therefore an inherent unfairness in PART scores across programs and budget officers. GAO found that design flaws in the tool made these problems worse.

This element of subjectivity in the PART evaluation process is very distressing. The individual OMB reviewers exercise such unbridled discretion in interpreting and developing the PART assessments that they are able to wield significant power to covertly alter the focus and priorities of the federal government in incremental steps. This slight but steady shift in power, oversight, and jurisdiction from the legislative to executive branch will become increasingly problematic as Congress begins to consider codifying the PART in legislation and if PART scores gain irrefutable status as an unbiased reflection of a program's or agency's worth.

Problems in Implementation of PART

President Bush began 2005 by continuing to lay the groundwork for popular understanding and acceptance of his budget as results-oriented. In his State of the Union address, Bush said, “My budget substantially reduces or eliminates more than 150 government programs that are not getting results.” This was the first of many instances to follow of Bush reiterating this idea during speeches, press conferences, and other public statements.

But there is little evidence in the FY 2006 budget to support the presidential rhetoric that results are the basis of funding decisions. The president’s rhetorical focus on performance and results seem to be just that — merely a smokescreen providing political cover for a Bush agenda that seeks to promote particular ideological policies while drastically reducing the size of the federal government.

Popular acceptance of the results rhetoric as fact, both in Congress and government, as well as in the minds of the American public is essential to Bush’s efforts to provide legitimacy for the PART as an unbiased evaluative tool even when the facts showed quite different outcomes.

Inconsistencies in Budget Requests and PART Ratings in FY 2006

OMB Watch has analyzed the much-publicized list of 154 programs and other sections of the FY 2006 budget and compared program funding requests to the ratings received under the PART. This analysis has yielded some inconsistent and puzzling results. **Out of the list of 154 programs to be cut or eliminated, more than two thirds have never even been reviewed by the PART.** It is unclear what kinds of determinations, if any, the president used to identify these failing programs when the White House budget staff has yet to assess them.

A quick review of programs rated under PART since its inception finds no logical or consistent connections with budget requests.

- Of the 85 programs receiving a top PART score in 2006, the president proposed cutting the budgets of more than 38 percent, including the National Center for Education Statistics and a land management program run by the Tennessee Valley Authority.
- Of all the programs reviewed on the list of 154, **nearly 20 percent of programs receiving an “effective” or “moderately effective” PART score** — the two highest ratings — **were targeted for elimination.** Further, 46 percent of programs receiving the middle rating of “adequate” were proposed to be eliminated.
- Even stranger, some programs receiving the lowest score were not cut. For instance, the Substance Abuse Prevention and Treatment Block Grant, a program that provides grants to states to address addiction problems, was given the lowest possible rating of “ineffective” but received no reduction in funding. Moreover, **the Earned Income Tax Credit Compliance Program** — which targets poor people who have claimed the EITC and double-checks their eligibility for the credit — **was rated ineffective, yet it received a funding increase.**

In the 2006 budget, programs rated “ineffective” apparently were targeted for elimination if they were housed in HUD or the Department of Education; of the programs that were rated ineffective and targeted for elimination, 78 percent came from HUD or the Department of Education

Case Studies of PART in the 2006 Budget

Attack on Grant Programs

An alarming trend in the PART assessments is a bias against programs that operate through grants, whether competitive grants or block grants. Below is a chart comparing the overall breakdown of PART scores in competitive grant programs, block grant programs, and all other programs. As is evident, grant programs rate significantly lower in PART reviews than all other programs on average. Further, of the programs rated “ineffective” that were zeroed out completely in the 2006 budget, 89 percent were competitive or block/formula grants.

Comparison of Grant Programs and All Other Programs in PART (percentage of programs rated in each category)			
	Competitive Grant	Block Grant	All Other Programs
Effective or Moderately Effective	24%	27%	49%
Adequate or Ineffective	36%	36%	26%
Results Not Demonstrated	40%	37%	25%

Grant programs largely send money to the state and local governments, many of which have countered the administration’s attack on public safeguards with progressive regulatory policy initiatives. Some, like the Community Development Block Grant, are particularly important to improving the economies of poor and rural communities across the country.

The odds are stacked against block grants because performance review is an oversight mechanism, whereas the premise of block grants is that funds are sent to the states with certain freedoms from complex federal oversight requirements. Many states and local governments have their own performance and accountability review processes; overlaying federal PART reviews has the effect of overriding state and local government self-management, contrary to the intent of block grant projects.

Grade Deflation

It is not clear how OMB determined that programs should be shifted into the category of “Results Not Demonstrated.” OMB assigns weights to the scores from each of the four PART sections and then assesses those scores on a grading scale to determine whether a program passes (“Effective,” “Moderately Effective,” or “Adequate”) or fails (“Ineffective”). The category of “Results Not Demonstrated” is supposed to be reserved for programs that do not generate sufficient data or information upon which a passing or failing score can reasonably be assigned. Although explained accordingly as an *indeterminate* category, programs relegated to the “Results Not Demonstrated” bin are often characterized in White House rhetoric as failing programs.

Whether it is indeterminate or a failing score, it is uncertain how the score is ultimately determined.

- In fact, **72 of the 178 programs (40 percent) categorized as “Results Not Demonstrated” had scores that, according to OMB’s own grading scale, should have been granted passing scores.**
- Of these 72 programs, **12 should have received the high score of “Moderately Effective.”**

- ❖ These 12 programs have higher scores for section 4 — the section that notionally measures actual results — than the average score for all programs actually rated “moderately effective.”
- ❖ **Three** — the Consumer Product Safety Commission, a USDA program for rural water treatment loans, and the National Credit Union Administration’s Community Development Revolving Loan Fund — **actually scored above 75 percent for producing results** (substantially higher than the 60 percent average for all programs rated “Moderately Effective”).
- The remaining 60 of the 72 otherwise passing programs should have received the middle passing score of “Adequate.”
 - ❖ **More than half of them (31/60) scored higher for section 4 than the average section 4 score for all programs that actually received the score of “Adequate.”**
 - ❖ Almost as many (24/60) had overall scores that bested the average overall score for programs that OMB allowed to receive the “Adequate” score.

Usurping Congressional Authority: Punishing Programs for Doing What Congress Intended

Many of the stated reasons for scoring programs negatively reflect nothing more than OMB’s disagreement with the way Congress designed a program or with the funding priorities laid out in annual appropriations. OMB does not merely suggest to Congress ways a program can be, in its view, improved; instead, OMB scores a program negatively and imposes consequences against it, such as reduced budget requests, simply for following the law. OMB then justifies its decision using the rhetoric of results rather than a direct statement of its disagreement with Congress.

Some examples:

- The Consumer Product Safety Commission, Occupational Safety and Health Administration, and Mine Safety and Health Administration were all penalized for failing to use economic analysis in their rulemaking processes — even though they are forbidden by law and Supreme Court precedent from doing so.
 - ❖ The CPSC is instructed by Congress not to use cost-benefit analysis when issuing rules specifically required by law, such as the rules governing garage door openers and bicycle helmets. **CPSC (which, despite an otherwise high passing score, was categorized “Results Not Demonstrated”) was penalized for following the law and not conducting cost-benefit analyses for those rules.**
 - ❖ CPSC was also scored down for not complying with OMB’s tests for net benefits, even though CPSC’s authorizing legislation instructs the agency to take a different approach in order to maximize public safety.
 - ❖ The same is true for OSHA and MSHA. **OMB scored these programs negatively for failing to do “cost-benefit comparisons or monetiz[ing] human life,” even though their organic acts and Supreme Court precedent forbid these practices.**
- OMB criticized the Appalachian Regional Commission (and flat-lined its budget request) for not being a “unique” program, because other existing agencies provide the same services. **OMB completely misses the point of the Appalachian Regional Commission**, which Congress created precisely because the existing patchwork of programs was failing to meet the needs of the extraordinarily impoverished population of that region.
- Another program serving rural populations, **HHS’s Rural Health Activities program, was likewise penalized for following the very law that created it.** OMB’s criticism speaks for itself: “The major flaw of the Office’s portfolio *stems from the programs’ authorization.*” [emphasis added]. The program was targeted for a drastic cut.

Using Measures That Are Incompatible With the Program's Purpose

In the core questions applied to every program during a PART review, there are several questions about economic efficiency but not a single question about equity. The specific standards used to measure program “effectiveness” or “results” are buried deep in documents that are sometimes several inches thick when printed out, but they follow the trend established by the overall questions themselves and, generally, do not focus on what is most meaningful or relevant about a program. Examples:

- Head Start was designated “Results Not Demonstrated,” which is peculiar for a program so widely perceived by both Congress and the public to be effective. **PART used measures that do not correspond with measures used in academic literature, from which education scholars and childhood development experts conclude the program is effective.** PART uses measures such as increased vocabulary and percentage of parents reporting they read to their children three times a week or more, whereas the scholarly literature relies on more logical indicators of success, such as the likelihood a child will repeat a grade, increases in test scores, and access to preventive health care.
- The Clean Water Revolving Fund was given a low passing score and slated for deep budget cuts, in part because PART measured success based on the “percentage of water miles/acres with fish consumption advisories removed.” **This measure is not a scientifically appropriate measure of actual water quality:** as EPA recently (Aug. 2004) announced, the number of rivers and lakes with mercury fish advisories increased in the last ten years even though the amount of mercury emissions actually declined by 100 tons. An increase in the number of advisories can actually be a sign of *success*, as it could mean the government is doing a better job of monitoring pollution and informing the public.
- **Every EPA research program PARTed to date was assessed as “Results Not Demonstrated,” based on rationales that are deeply incompatible with the purposes of those programs.** OMB criticized these programs for failing to link their research activities with the accomplishment of outcomes, but such criticism is willfully blind of the very nature and benefits of research: often we can learn as much from failure as from any success.

Political and Ideological Implications of the PART

The PART and government performance rhetoric have been increasingly prominent throughout the Bush presidency. Each year, more emphasis is placed on this idea in the federal budget document, in the president's public comments, and throughout the federal government.

PART “Data” Are Meaningless

Despite all the hype, the PART cannot be characterized as a refined method of gauging program performance. Rather, it is no more sophisticated than a grade-school sticker method used to reward good or bad work. However, its very simplicity makes it a potentially powerful way to justify budget cuts or increases, and, in fact, is the key to effectively accomplishing undisclosed political and ideological goals — namely the reduction of the federal role.

While OMB has stated repeatedly the PART scores will not be used, in and of themselves, to slash program budgets, there is a danger of having the PART score become the sole determinant of a program's value. In fact, OMB's website states, “In general, because the PART is an evidence-based, objective analysis, the Administration believes it is a useful tool in budget and policy formulation for all branches of government.”

OMB Watch's current analyses of the PART have uncovered more questions than answers about its value and purpose. It is unclear how the PART scores impact budgeting decisions within OMB as there are no consistent patterns to follow. It is hard to determine whether the PART is measuring programs accurately, consistently, and in an unbiased and value-neutral way. Even if it achieves these goals, it appears there has been little, if any, attention paid to whether the PART is measuring the right kinds of outcomes. While a lot of emphasis is placed on efficiency and cost-savings measures, the PART does not appear to incorporate equity issues within the rating process — such as protecting and assisting specific groups like minorities, the disabled, and women, or protecting the public interest.

PART Is Arbitrary and Political

Worst of all, the PART does not include the statutory intent of programs (Congress' will) or input and feedback from program participants, service providers, beneficiaries, or the public at large (the people's will). These are the very people and intents the government is supposedly representing and any attempt to improve programs should be inclusive of those perspectives.

Underneath all the smoke and behind the rhetoric, we find a disturbingly large number of the PART evaluations to be based on the dispositions of White House budget bean counters, not some truly objective measure such as information collected from program beneficiaries or statutory standards. Conducting program evaluations in this way is a costly exercise generating meaningless numbers that do nothing except back up pre-ordained political conclusions.

PART Is Costly

PART comes with opportunity costs. Agency resources that should be focused on addressing the public's unmet needs are instead diverted in this time-intensive exercise that ultimately provides no meaningful information whatsoever. These opportunity costs, which include but are not limited to the costs from lowered agency morale (characterized elsewhere as “reform fatigue”), are not justified by the outcome. The PART is but a thin veneer of accountability and good government deflecting attention and criticism from controversial and politically biased judgments. In this sense the PART mechanism itself, ironically, fails to demonstrate successful results.

Despite Its Failures, the Performance Movement Proceeds

Despite failing to produce any meaningful results for the public interest, PART and performance rhetoric are being pushed in a number of policy proposals. A bill recently reported out of committee in the House would essentially codify the PART (despite its many flaws), and some proposals for civil service reform have called for using PART “data” in determining staff salaries. Uncritical acceptance of the PART is a recipe for disaster.