S. Hrg. 109–923

AUTOPILOT BUDGETING: WILL CONGRESS EVER RESPOND TO GOVERNMENT PERFORMANCE DATA?

HEARING

BEFORE THE

FEDERAL FINANCIAL MANAGEMENT, GOVERNMENT INFORMATION, AND INTERNATIONAL SECURITY SUBCOMMITTEE

OF THE

COMMITTEE ON HOMELAND SECURITY AND GOVERNMENTAL AFFAIRS UNITED STATES SENATE

ONE HUNDRED NINTH CONGRESS

SECOND SESSION

JUNE 13, 2006

Available via http://www.access.gpo.gov/congress/senate

Printed for the use of the Committee on Homeland Security and Governmental Affairs

U.S. GOVERNMENT PRINTING OFFICE

29–503 PDF WASHINGTON : 2007

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov Phone: toll free (866) 512–1800; DC area (202) 512–1800
Fax: (202) 512–2250 Mail: Stop SSOP, Washington, DC 20402–0001

COMMITTEE ON HOMELAND SECURITY AND GOVERNMENTAL AFFAIRS

SUSAN M. COLLINS, Maine, *Chairman*

TED STEVENS, Alaska
GEORGE V. VOINOVICH, Ohio
NORM COLEMAN, Minnesota
TOM COBURN, Oklahoma
LINCOLN D. CHAFEE, Rhode Island
ROBERT F. BENNETT, Utah
PETE V. DOMENICI, New Mexico
JOHN W. WARNER, Virginia

JOSEPH I. LIEBERMAN, Connecticut
CARL LEVIN, Michigan
DANIEL K. AKAKA, Hawaii
THOMAS R. CARPER, Delaware
MARK DAYTON, Minnesota
FRANK LAUTENBERG, New Jersey
MARK PRYOR, Arkansas

MICHAEL D. BOPP, *Staff Director and Chief Counsel*
MICHAEL L. ALEXANDER, *Minority Staff Director*
TRINA DRIESSNACK TYRER, *Chief Clerk*

FEDERAL FINANCIAL MANAGEMENT, GOVERNMENT INFORMATION, AND INTERNATIONAL SECURITY SUBCOMMITTEE

TOM COBURN, Oklahoma, *Chairman*

TED STEVENS, Alaska
GEORGE V. VOINOVICH, Ohio
LINCOLN D. CHAFEE, Rhode Island
ROBERT F. BENNETT, Utah
PETE V. DOMENICI, New Mexico
JOHN W. WARNER, Virginia

THOMAS CARPER, Delaware
CARL LEVIN, Michigan
DANIEL K. AKAKA, Hawaii
MARK DAYTON, Minnesota
FRANK LAUTENBERG, New Jersey
MARK PRYOR, Arkansas

KATY FRENCH, *Staff Director*
SHEILA MURPHY, *Minority Staff Director*
JOHN KILVINGTON, *Minority Deputy Staff Director*
LIZ SCRANTON, *Chief Clerk*

CONTENTS

	Page
Opening statements:	
Senator Coburn	1
Senator Carper	6

WITNESSES

TUESDAY, JUNE 13, 2006

Hon. Clay Johnson III, Deputy Director for Management, U.S. Office of Management and Budget	8
Eileen Norcross, Government Accountability Project, Mercartus Center at George Mason University	9
Adam Hughes, Director for Federal Fiscal Policy, OMB Watch	11

ALPHABETICAL LIST OF WITNESSES

Hughes, Adam:	
Testimony	11
Prepared statement with attachments	89
Johnson, Hon. Clay III:	
Testimony	8
Prepared statement with attachments	25
Norcross, Eileen:	
Testimony	9
Prepared statement	49

APPENDIX

"A Working Paper in Government Accountabiity" by Eileen Norcross and Kyle McKenzie	61

AUTOPILOT BUDGETING: WILL CONGRESS EVER RESPOND TO GOVERNMENT PERFORMANCE DATA?

TUESDAY, JUNE 13, 2006

U.S. SENATE,
SUBCOMMITTEE ON FEDERAL FINANCIAL MANAGEMENT,
GOVERNMENT INFORMATION, AND INTERNATIONAL SECURITY
OF THE COMMITTEE ON HOMELAND SECURITY AND
GOVERNMENTAL AFFAIRS,
Washington, DC.

The Subcommittee met, pursuant to notice, at 2:57 p.m., in room 342, Dirksen Senate Office Building, Hon. Tom Coburn (Chairman of the Subcommittee) presiding.

Present: Senators Coburn and Carper.

OPENING STATEMENT OF SENATOR COBURN

Senator COBURN. Good afternoon. The Federal Financial Management Subcommittee of the Homeland Security and Governmental Affairs Committee will come to order. Senator Carper will be here in a moment. We apologize for the delay. There was an official photo. We also have a conflict. There is a briefing ongoing now by the Secretary of Defense and the Secretary of State, which will limit Senator Carper's time with us. So we are going to go on and go forward so we have it in the record. I apologize for the conflicting schedules.

Americans have a crazy idea, that they should get something for their money, even when the money is spent by the government. It is a simple concept, and in policy-speak we call it performance-based budgeting. I know I am new in the Senate, but I am still surprised by the level of resistance in Washington to holding people accountable by measuring their performance. And it is a difficult thing to do. A multitrillion-dollar government imposing some sort of standardized outcome evaluation is difficult at best, and what it implies is that the tool will be very crude. But that does not say we should not attempt to make measurements, and I want to be one of many who should commend both Mr. Johnson and the Bush Administration, and the President himself, for being the first to attempt to do it.

It is not novel. It is required in the competitive business environment that we find ourselves worldwide. It is being used effectively in many State governments, and it is something that is long overdue. The Performance Assessment Rating Tool (PART) was first introduced by the President 4 years ago as a tool to review the

strengths and weaknesses of government programs to influence funding and programmatic decisions. The annual PART reports offer needed sunshine in government and provide good data for government managers to improve their programs.

Today, the Office of Management and Budget has reviewed 793 programs, which account for $1.47 trillion in taxpayer money. Almost a third of these programs have proven to not meet up to standards based on the PART analysis. I have already admitted that it is a blunt tool. One-third of $1.5 trillion is $500 billion. Maybe this is why PART scores so far have created a stir not only among the agencies but among the Members of Congress who make budgeting decision.

Some Members of Congress want to stick their heads in the sand and keep funding pet programs on autopilot year after year. To my amazement, just last week, the Appropriations Subcommittee that funds the Departments of Labor and Health and Human Services passed language prohibiting the use of PART assessments on those agencies. They may not like PART's message, but they should not shoot the messenger. This sort of Orwellian immunization against any hint that our favorite programs may not be performing up to the idealized, utopian goals of their Congressional champions is one of the reasons why Americans are mad at Congress.

The approval ratings for Congress are in the tank, and this prohibition of accountability for failing government is why the voters who fork over their hard-earned dollars every year may just have something new to say come this November. I am not sure why so many of my colleagues are afraid of assessment tools on performance. It may reflect their own performance.

As part of our investigation for this hearing, we learned that low PART ratings do not always mean that OMB will recommend a budget cut or a cut in the program or a recommendation to go on the terminations list. In some cases, programs rated ineffective have had budget reductions recommended. But in other cases, the reason they were low was because they were not funded appropriately to begin with, and therefore, they could not accomplish what they were intended to because they did not have adequate funding.

Each program is unique, and I do not know that a PART score should be the last word. But I do know that measurement of performance is something that every member of a Congressional authorizing or Sppropriations committee should be reading and using to inform their oversight work. Congress consistently neglects the duty to conduct oversight of Federal programs and spending. Instead, we spend most of the time passing spending bills that ignore PART ratings, the President's termination list, or any other performance data as if the spending were on autopilot. Congress might as well write a blank check.

By 2008, OMB will have applied PART to the entire government. In the last 4 years, OMB has scored 793 government programs. Here are the results: 15 percent were found to be effective; 29 percent were found to be moderately effective; 28 percent were rated adequate; 4 percent were found to be ineffective—that is one in every 25 programs—25 percent could not demonstrate results to get a rating and were labeled results not demonstrated.

I do not believe the spin that results not demonstrated can mean that the program is either good or bad; we just do not have enough information to tell. On the contrary, the results not demonstrated designation is a red flag marking a program so poorly conceived by us or so directionalist that that unaccountability seems to have been built into it by design. Programs rated ineffective or results not demonstrated account for $152 billion in budget authority. Imagine what we could do with $152 billion right now. The ideas are endless.

Outside of Washington, DC, any business or family with finite resources sets priorities and creates a budget based on the actual amount of bang they get from their buck. It is only inside the Beltway where that kind of information is not considered relevant, and in fact, some are even attempting to ban the collection of such information. But then, it is only Washington where you never have to declare bankruptcy, and debt is allowed to grow on the backs of future generations with impunity.

Let me give you one case study, and my co-chairman on this will disagree, but my firm believing is the following: We held a hearing last year on the Advanced Technology Program that was created in 1988 to subsidize high-risk research and development. This program has never demonstrated results. What it has demonstrated is corporate welfare. Its 2002 PART report, that the majority of ATP grants go to multibillion dollar corporations and that the GAO has found that ATP projects are very similar to private sector R&D undertaken without a government subsidy. An amendment to eliminate this funding that was offered last year lost by a vote of 68 to 29. In the end, Congress wasted a portion of $79 million last year for that program. The 2007 Senate budget resolution promises to fund the program at almost twice that amount.

It would be one thing if we were operating in a surplus. Then, we could have a legitimate debate about whether to keep failing programs, hoping that they would improve, or to give that surplus back to the taxpayers. But that is not where we are today. With a debt burden of $25,000 per man, woman, and child, we simply cannot afford to keep funding programs that cannot prove their worth. Non-defense discretionary spending has increased 45 percent since 2001. The President has requested a $2.8 trillion budget, and that does not include any of the so-called emergency, "supplemental bills in our future," nor does it include the late night pork barrel frenzy each time Congress schedules an appropriations bill vote.

Entitlement spending will tank our economy if we do not do something to get spending under control. The question remains: How do we get Congress to act? I would like to see OMB sell their PART terminations list more aggressively, forcibly sell the reforms and savings to Congress, fight for the cuts by taking the terminations list to the American people with the power of the bully pulpit. The President should veto spending bills that continue to issue blank checks for failing programs.

There is a bit of hope on the horizon. I was encouraged to see that the House Appropriations Committee wrote in their 2006 budget savings report that the only way to establish accountability in the budget process is to stop spending on programs that have

outlived their usefulness or could be delivered more effectively at the State or local level. I will believe that when I see it, but I welcome any help that we can get.

The best place to start is by immediately defunding all programs on the termination list and adopting other PART recommendation reductions. Granted, the list only cuts $20 billion from a $2.8 trillion budget, but we have got to start somewhere. What is more, we should suspend the creation of any new program until further notice or it is compared to the existing programs that it is meant to supplement. We need sunset legislation that would phase out government agencies on a timed basis, where we force ourselves to look at them and to reauthorize them.

These are challenging times, and we can no longer afford to run on a budget that is on cruise control. I want to thank our witnesses for being here.

[The prepared statement of Chairman Coburn follows:]

OPENING PREPARED STATEMENT OF CHAIRMAN COBURN

Americans have a crazy idea: They should get something for their money, even when the money is spent by government. It's a simple concept—in policy-speak, we call it "performance-based budgeting." I know I'm new in the Senate, but I'm still surprised by how much resistance there is in Washington to performance-based budgeting.

Now, to be fair, taking a multi-trillion dollar government and imposing some sort of standardized outcome evaluation on it is difficult at best. So I concede that any instrument we use will be a blunt instrument. But I want to commend President Bush for being the first to try.

The Performance Assessment Rating Tool (PART) was first introduced by the President 4 years ago as a tool to review the strengths and weaknesses of government programs to influence funding and programmatic decisions.

The annual PART reports offer needed sunshine in government and provide good data for government managers to improve their programs. To date, the Office of Management and Budget has reviewed 793 programs which account for $1.47 trillion in taxpayer money. Almost a third of these programs have proven either totally ineffective or are not demonstrating results. One-third of $1.5 trillion is $500 billion.

Maybe this is why the PART scores have created a stir—not only among the agencies, but among the Members of Congress who make budgeting decisions. Some Members of Congress want to stick their head in the sand and keep funding their pet programs, as if on autopilot, year after year.

Just last week the House Appropriations subcommittee that funds the Departments of Labor, Education and Health and Human Services passed language prohibiting the use of PART assessments on those agencies. They may not like PART's message, but they shouldn't shoot the messenger. This sort of Orwellian immunization against any hint that our favorite programs may not be performing up to the idealized utopian goals of their Congressional champions is why Americans are mad at Congress. The approval ratings for Congress are in the tank, and this prohibition of accountability for failing government is why the voters who fork over their hard-earned dollars every year may just have something to say come November.

I'm not sure why some of my colleagues are so afraid of PART. As part of our investigation for this hearing, we learned that low PART ratings don't always mean that OMB will recommend a budget cut or put the program on the Terminations List. In some cases, programs rated "ineffective" had budget reductions, but in other cases their budgets increased. Each program is unique and I don't know that a PART score should be the last word, but I do know that the PART is something every member of a Congressional authorizing or Appropriations committee should be reading and using to inform their oversight work.

You see, Congress consistently neglects the duty to conduct oversight of Federal programs and spending. Instead, we spend most of the time passing spending bills that ignore PART ratings, the President's terminations list and any other performance data. It is as if we're spending on "auto pilot"—Congress might as well just write a blank check.

By 2008, OMB will have applied PART to the entire government. In the last 4 years OMB has scored 793 government programs. Here are the results: Just 15 per-

cent were found to be "effective"; 29 percent were rated "moderately effective"; 28 percent were rated "adequate"; 4 percent were found to be "ineffective"; and 24 percent cannot demonstrate results to even get a rating and were labeled "results not demonstrated"! Don't believe the spin that "results not demonstrated" could mean that the program is either good or bad, we just don't have enough information to tell. On the contrary—the "results not demonstrated" designation is a red flag marking a program so poorly conceived or directionless that unaccountability seems to have been built into it by design.

Programs rated "ineffective" or "results not demonstrated" account for $152 billion in budget authority. Imagine what we could do with $152 billion.

Outside of Washington DC, any business or family with finite resources sets priorities and creates a budget based on the actual amount of bang they get for their hard-earned buck. It is only inside the beltway where that kind of information isn't considered relevant and in fact, some are trying hard to ban the collection of such information. But then, it's only in Washington where you never have to declare bankruptcy and debt is allowed to grow on the backs of future generations with impunity.

Let me give you one case study. We held a hearing last year on the Advanced Technology Program. The program was created by Congress in 1988 to subsidize high-risk research and development. The program cannot demonstrate results. It is corporate welfare. The 2002 PART reported that the majority of ATP grants go to multimillion dollar corporations and that the GAO has found that ATP projects are very similar to private sector R&D undertaken without a government subsidy. An amendment to eliminate funding for ATP that I offered last year was voted down in the Senate 68–29. In the end, Congress wasted another $79 million last year for the program. The 2007 Senate budget resolution promises to fund the program at almost twice that amount.

It would be one thing if we were operating in a surplus. Then we could have a legitimate debate about whether to keep funding failing programs hoping they will improve or to give that surplus back to the taxpayers. But that's not where we are today, with a debt burden of $25,000 per man, woman and child in America. We simply cannot afford to keep funding programs that cannot prove their worth.

Nondefense discretionary spending has increased over 45 percent since 2001. The President has requested a $2.8 trillion budget and that doesn't include any so called "emergency" supplemental spending bills in our future, nor does it include the late-night pork-barrel frenzy each time Congress schedules an Appropriations bill vote. Entitlement spending will tank our economy if we don't do something to get spending under control.

The question remains, how do we get Congress to act? I would like to see OMB sell their PART and Terminations List more aggressively:

- Forcefully sell these reforms and savings to Congress.
- Fight for these cuts, by taking the terminations list to the American people with the power of the bully pulpit.
- The President should veto spending bills that continue to issue blank checks to failing programs.

There's a bit of hope on the horizon—I was encouraged to see that the House Appropriations Committee wrote in their 2006 Budget Savings report that "the only way to establish accountability in the budget process is to stop spending on programs that have outlived their usefulness or could be delivered more effectively at the State or local level." I'll believe it when I see it, but I welcome any help we can get.

The best place to start is by immediately defunding all programs on the Terminations List and adopting the other PART reduction recommendations. Granted, the list only cuts $20.4 billion from a $2.8 trillion budget, but we've got to start somewhere. What's more, we should suspend the creation of any new program until further notice. We need "sunset" legislation that would phase out every single government agency, department or program after a certain deadline if the Congress fails to act or if the program consistently performs poorly. These are challenging times and we can no longer budget on cruise control.

I want to thank our witnesses for being here today and for the time they spent preparing testimony.

Again, I apologize for the lateness of our attendance, and Senator Carper, you are recognized.

OPENING STATEMENT OF SENATOR CARPER

Senator CARPER. Thank you.

Senator COBURN. And I have already explained that you will probably have to attend the briefing that is ongoing.

Senator CARPER. Thanks, Mr. Chairman. To our witnesses, welcome. It is good to see each of you. We appreciate you joining us and providing your testimony today. As the Chairman mentioned, Secretary of State Rice and Secretary Rumsfeld are briefing us as we speak over in the Capitol, and I want to slip out in a little bit and hear what they have to say and hopefully rejoin you before you leave.

Mr. Chairman, thank you for holding this hearing. It is an important hearing, as we both know. And as we have discussed in any number of our similar hearings in the past over the last couple of years, our country is facing a large budget deficit for as far as the eye can see, and we are just about to embark on another appropriations season here in Congress, where we will be called on to make some difficult decisions about what to do with relatively scarce Federal resources.

At the same time as GAO and other observers have pointed out again, and again, we are at a crossroads in our history, where we need to decide what we want our government to do in the 21st Century. Nearly 5 years after the attacks of September 11, 2001, we have a whole new set of needs, a whole new set of priorities that must be balanced against some of our older needs and priorities in scores of popular programs. And with the challenge of retiring baby boomers, guys like me, our generation on the horizon, we just cannot afford to do all of the things that we might want to do.

That is why initiatives like OMB's Program Assessment and Rating Tool (PART) are interesting and, I think, important. We should never be afraid of taking a hard look at Federal programs, my programs, Senator Coburn's programs, whatever, to determine whether or not they are accomplishing what was intended for them to accomplish when we first created them. And in this day and age, we simply cannot afford to allow either poorly conceived or poorly managed programs to continue without reform or, frankly, for a program that has run its course and achieved its goals, to continue draining resources from other, newer priorities.

That said, we need to be certain that PART or whatever mechanism we use to make these evaluations is in itself effective. I think to be effective, a program like PART must be totally separated from politics and ideology, at least to the extent we can make that happen. It must be closely coordinated with existing mechanisms agencies and Congress use to align the budget with program goals and outcomes such as the older government Performance and Results Act. And perhaps just as importantly, we also need to make sure that a program's intended beneficiaries outside of Washington have a say before an evaluation is actually completed.

Let me just add in closing, if I could, Mr. Chairman, that we are not going to close the budget deficit, we know, by reducing spending on a program here or eliminating a program there, although every little bit helps. But even if a program were to eliminate every single one of the programs receiving failing grades through PART,

I still think the savings would cover just a fraction of our budget deficit, but they would cover a portion of our budget deficit.

Non-defense discretionary spending, which is the target of many of the spending reductions and program eliminations in the President's budget proposals, make up a relatively small percentage of the Federal budget. I am sure we can find ways to improve the management of some of the funding in that 16 percent or even to find and eliminate waste and inefficient use of resources within that 16 percent.

If we truly want to tackle the fiscal problems facing us right now, however, we, and that is the Congress and I think the Administration needs to take a look at the entire budgetary picture. We need to look on both the spending and on the revenue side, and we need to make some tough choices.

Thank you, Mr. Chairman. Again, we look forward to your testimony today.

Senator COBURN. Thank you, Senator Carper.

I am going to ask the witnesses to limit their verbal testimony to 5 minutes. Your complete written statements will be made a part of the official hearing record, and we will hold our questions until you have given your testimony.

Let me first introduce Clay Johnson III, Deputy Director for Management at OMB, and in his capacity, he has provided the government-wide leadership to the Executive Branch agencies to improve agency and program performance. Formerly, he served as Assistant to the President for Presidential Personnel, responsible for the organization that identifies and recruits approximately 4,000 senior officials, middle management personnel, and part-time board and commission members. At OMB, he oversees PART process.

Eileen Norcross, Senior Research Fellow, Government Accountability Project, The Mercatus Center at George Mason University; she joined that center as a research fellow in January 2003. Her research areas include the U.S. budget, the use of performance budgeting in the Federal Government, tax and fiscal policy, and environmental regulation. She is one of the leading experts on performance-based budgeting, and her scholarship plays a vital role in the debate on PART and the importance of measuring outcomes.

Adam Hughes is the Director for Federal Fiscal Policy at OMB Watch. He oversees Federal budget and tax policy, income and wealth trends, and government performance issues at OMB Watch. Senator Carper and myself very much appreciate the work that OMB Watch has done in their pursuit of transparent and accountable government and for the support of the Federal Funding Accountability and Transparency Act that we both authored. This bill would create an online public database that itemizes Federal funding so taxpayers can see how their money is being spent.

I want to welcome you all. I will recognize Mr. Johnson first.

TESTIMONY OF CLAY JOHNSON III,[1] DEPUTY DIRECTOR FOR MANAGEMENT, U.S. OFFICE OF MANAGEMENT AND BUDGET

Mr. JOHNSON. Mr. Chairman, Senator Carper, thank you very much.

The title of this hearing is Will Congress Ever Respond to Program Performance Data? In preparing my response, I rephrased that to Does Congress Care Whether Programs Work or Not? My answer is "I am not sure, but I sure hope so." I believe that taxpayers want Congress to ensure that they, the taxpayers, get what they pay for. I believe that we all, to widely varying degrees, however, want Federal programs to do what they are supposed to do and get better every year.

I believe that money is tight, as you all have pointed out, and the biggest opportunity we have to add new services and expand existing services to more citizens is through causing our existing programs to work better, not spend more money. I believe that career employees want to be held accountable for how their programs perform. They tell me this in focus groups. And I also believe that career employees care about how their programs perform.

Because of this, I believe it is important to have certain things. I believe it is important to have clear outcome goals for each Federal program. We do not have that now. I believe it is important to have Federal program performance information that is objective, as objective and reliable as possible. I believe that we need to have lots of transparency about how well programs are performing. If we do all of this in the dead of night, it cannot be used to hold people accountable.

I believe that we need lots of debate about these performance assessments and how to make them better. As you said, Mr. Chairman, program assessment is going to be a blunt instrument, particularly in the early years. And it will only get better every year, but a blunt instrument is better than no instrument at all. I also believe it is important to have lots of discussion about how to help programs work better. We talk a lot about using the PART to make budget decisions. I believe the primary use of PART information is to help programs get better. If we cut programs, we might save $10 billion here or $15 billion there per year. If we cause 1 percent improvement in program performance each year, that is $28 billion a year. Two percent is obviously twice that.

After 5 years of effort, not 5 months, comprehensive program performance information is still time consuming and very hard to come by. We have program outcome goals, performance information, and lots of transparency, which other countries and several States are working to adopt, and most good government groups applaud. What we do not have from most Members of Congress is a lot of constructive debate about these assessments and how to improve and use them to improve program performance. We have asked for feedback. We have asked for engagement by Congress but have not gotten it.

Currently, a majority of Appropriations subcommittees have no objection to the way agencies use performance information to jus-

[1] The prepared statement of Mr. Johnson with attachments appears in the Appendix on page 25.

tify their budgets. Some of these subcommittees actually use the PART to justify program funding in their bills. A few Members of Congress have advanced greater use of performance information in decisionmaking. Congressmen Platts and Tanner have proposed separate pieces of legislation, while Senators like you, Senator Coburn, and Senators Carper, Ensign, and Allard have spoken out on the subject, and Congressmen Cuellar, Conaway, and Diaz–Balart have spoken out on it as well.

But these expressions of interest in program performance are the exceptions. There is a big, a huge opportunity for Congress to challenge programs to clearly define success and their plan for achieving it, and then to hold agencies accountable for doing what they said they were going to do.

That concludes my remarks, and I look forward to any questions.

Senator COBURN. Ms. Norcross.

TESTIMONY OF EILEEN NORCROSS,[1] SENIOR RESEARCH FELLOW FOR THE GOVERNMENT ACCOUNTABILITY PROJECT, THE MERCATUS CENTER AT GEORGE MASON UNIVERSITY

Ms. NORCROSS. Thank you, Chairman Coburn, Senator Carper, for inviting me to testify today on Autopilot Budgeting: Will Congress Ever Respond to Government Performance Data? Our work in the Government Accountability Project at the Mercatus Center at George Mason University focuses closely on performance information in government, and I note that the views expressed in my testimony are not an official position of the university.

I would like to submit for the record our paper on the results of the fiscal year 2007 PART for your reference.

Senator COBURN. Without objection, the document will be included in the record.[2]

A program is a tool to achieve a policy goal. Do economic development programs lead to prosperous communities? Are homeland security programs protecting the Nation? Congress needs to know the answers to these questions in order to make decisions about how to spend resources. Without performance information, Congress cannot reliably accomplish its policy aims. Not knowing its consequences, Congress has created anywhere from 180 to 342 programs dealing with economic development in over 24 agencies; 44 job training programs in nine agencies.

Program duplication on this scale tells us that Congress is not sure which programs are reaching their goals. It has no way of comparing programs around common outcomes. Not knowing if a job training program is employing people means not spending money on programs that are employing people. Not evaluating programs on a regular basis prevents the program from effectively reaching grantees or delivering results; performance information from its dialogue between agencies, the Executive Branch, and Congress around jointly defined objectives.

Congress took the initiative in 1993, when it passed GPRA. GPRA has encouraged the development of performance measures and data, but it was not until OMB's Program Assessment Rating

[1] The prepared statement of Ms. Norcross appears in the Appendix on page 49.
[2] The Working Paper in Government Accountability appears in the Appendix on page 61.

Tool that real progress towards developing measures was made. That is because the Administration does not just require the information; it uses it. Congress has identified the need for performance information. It must now commit to using it. Otherwise, measuring and gathering data is a paper exercise.

For the past 2 years, the President has issued a major savings and reforms report detailing his reasons for terminating or reducing funding for programs. Of the 154 recommended for termination or reduction in funds last year, 54 were PARTed. The document indicates where PART played a role. Other factors include lack of a Federal role, obsolescence, or completion of mission.

The Administration uses PART along with other information and does not limit itself to the evaluations. It does not automatically reward satisfactory programs or cancel underperforming ones. By contrast, the House Committee on Appropriations report "On Time and Under Budget" lists 53 programs that were terminated. It only offers explanations for three of the terminations. We do not know if the remainder were terminated because they were underperformers or politically easy choices.

The Administration's report gives a rationale for each recommendation. The House report only provides a list. Ultimately, the goal is not to randomly kill programs. Making judgments about how to fund agency activity should be constructive, not destructive. Performance information helps make policy effective. We want to know what works, what does not, and why.

The only way to give budgetary decisions credibility is to base them on a reliable evaluation of their performance. Is PART that system? PART's methodology has been criticized. Improvements can and should be made. But what is important about PART is not the ratings; it is the Management 101 questions PART asks of agency activity. Is the program purpose clear? Is it effectively targeted? Has it demonstrated progress towards its goals? These questions are the substance of PART. These are the questions Congress should be asking before allocating resources.

PART has a few virtues. It has identified and catalogued agency activity. It is transparent. It holds programs accountable to the same standards. It measures outcomes. Once strength often cited as a weakness: PART rates programs on statutory limitations. Though a source of frustration for agencies, here, PART provides a service by identifying those aspects of a program that are barriers to success. The hope is that Congress review the statute to see if it is preventing the program from meeting its objectives.

Some limitations of PART: It rates programs against their own performance. We would like to see PART advanced to compare like activities. In some cases, scores may not fully reflect program performance, and there is a potential for different budget examiners to reach different conclusions. We do not believe Congress should adopt PART wholesale. We hope Congress would consider using the kinds of questions PART is asking as the basis of developing its own method of evaluating agency activity based on common outcomes.

Indiscriminate cancellation of programs discredits the budgetary process. We leave program managers confused about why their programs failed. Programs need to deliver according to clear expecta-

tions and be given a chance to perform. When you do not meet the expectations, reduction in funding or termination should be the result. It should not be a surprise.

We believe performance information is best used in conjunction with other criteria. All of these form the basis against which Congress should continually scrutinize agency activity. Efforts to advance what PART has set in motion can only aid Congress in its work and give the American people confidence that our Nation's problems are being solved. Thank you.

Senator COBURN. Mr. Hughes.

TESTIMONY OF ADAM HUGHES,[1] DIRECTOR OF FEDERAL FISCAL POLICY, OMB WATCH

Mr. HUGHES. Chairman Coburn, thank you for having me here today and for holding this hearing. As you mentioned, I am the Federal Fiscal Policy Director at OMB Watch. OMB Watch was founded in the 1980s and has spent over 20 years advocating for government accountability, transparency, and access to government information, and citizen participation in governmental processes.

OMB Watch believes citizens must take an active role in holding their government accountable and that the Federal Government, when supported by sensible fiscal policy, can develop effective programs and safeguards that meet the public's needs.

The issue of government performance, as you mentioned earlier, has taken on added importance during the Bush Administration, as a combination of factors, some avoidable and some not, have plunged the Federal Government into debt. Large and sustained deficits over the past 5 years have made efficient use of government resources all the more important.

In light of the anticipated budget crunch due to the baby boomers' retirement over the coming decades, the fiscal situation in this country will only deteriorate further. Performance measurement can therefore become a particularly attractive alternative for those who want to set Federal priorities based on the current fiscal prospects of a strained and shrinking revenue base.

OMB Watch has been commenting on government performance issues for the better part of its existence. We have spent increased time and resources analyzing the Government Performance and Results Act and the Program Assessment Rating Tool over the last 10 years, as government itself has spent more time focusing on performance and results. We are strongly supportive of improving the Federal Government's capacity to meet the public's needs. OMB Watch has worked to protect and improve that capacity, and we have been open to the possibility of using performance measurement as one means for achieving those ends.

We bring a strong belief in the importance and potential of government itself to the work we do, and because of that belief, we, perhaps maybe more than anyone else, want government to be responsible to community needs, spend money wisely, and accomplish its goals. We are advocates for government and therefore have a strong motivation to see government programs succeed.

[1] The prepared statement of Mr. Hughes with attachments appears in the Appendix on page 89.

PART, however, is a very poor mechanism for measuring program performance and results, introducing biases and skewed ideological perspective into a model claiming to present consistent and objective performance data and evaluations of government programs. Oftentimes, the PART actually decreases the efficiency and effectiveness of government through increased administrative burdens, distracted managers, and compliance costs.

Ironically, we feel the PART mechanism itself does not produce the right type of results to support and improve government. We believe PART ratings should not be directly connected to the budgeting process of Congress because of significant deficiencies within the mechanisms, namely, the substantial biases and limitations embedded within the tool and the additional distortion and manipulation we have observed in OMB's actual application of the PART.

Based on our studies of the PART and our longstanding commitment to open, accountable government that is responsive to the public's needs, I would like to make three points today. First, we feel the PART continues a troubling trend we have seen in other recent Executive Branch proposals and even some Congressional proposals, namely, a trend towards increasing the power of the White House and the Executive Branch even into some areas that have been constitutionally designed to be committed to Congress.

Second, the PART is a limited and distorted tool that should not be used for either management of programs or for budget and appropriations decisions. In both the design of the tool and the process by which the tool is implemented, PART systematically ignores the reality and the complexity of Federal programs and judges them based on standards that are often deeply incompatible with the purposes those very programs are expected to serve. As one agency contact memorably explained to us, PART assessments are akin to a baseball coach walking to the mound to remove his star player and then chastising him for not kicking enough field goals.

My third point is that there is a better way. Specifically, Congress already has the means to investigate and produce far more sophisticated analyses of the usefulness, effectiveness, and results of government programs in a deliberative way, including the opportunity for input from a wide array of stakeholder interests. The openness of the Legislative Branch allows the Congress to be informed and make better decisions, but it also serves to balance competing agendas and perspectives from both inside and outside Congress.

The oversight and evaluation process is one of the primary if not the primary role for the Legislative Branch. While the oversight function of Congress may not be as robust as it once was because of significantly shorter legislative sessions and delays due to sharply divided political climates, the capacity to judge the results of government programs already exists within the existing structures, structures that we feel do not carry the significant limitations, biases, and negative consequences of the PART.

In conclusion, we all agree that everyone in government, the President, agencies and departments, and their staffs, and especially Congress, needs to be focused on achieving results in a fair, effective, and balanced way. However, this job should most of all fall on Congress, which already has the necessary tools and re-

sources in place to do the most robust and equitable review of the entire Federal Government.

Relying too heavily on PART ratings will not only gradually remove Congress from its funding and oversight responsibilities but will also continue to close the door on opportunities for outside stakeholder interests, the views of the public, to be infused into the Congressional budgeting and evaluation process. The limited perspective of the PART is one of its most glaring deficiencies. While subjectivity and bias will almost always creep into any rating system, the PART does not have a mechanism for balancing out the results of its one-size-fits-all, Executive Branch-focused perspective.

While the expansion of the Executive Branch powers has been present in our government since the turn of the last century, the overreach of those powers into areas historically and constitutionally given to Congress, the structuring of programs, appropriating and authorizing of revenues, and oversight of government is a disturbing trend. Because of this, PART should not be taken with just a grain of salt or even a hefty dose of skepticism. Unless the tool design and implementation systems are significantly modified, the PART ratings should probably be largely ignored by Congress.

Thank you. I look forward to your questions.

Senator COBURN. Thank you.

I wonder if either of you might want to comment on Mr. Hughes' testimony. It is certainly different than what we heard from either Mr. Johnson or Ms. Norcross, and I have several questions for Mr. Hughes as well, but I thought—Mr. Johnson, would you like to comment?

Mr. JOHNSON. Yes; several other countries around the world think the PART is great; other States in America think the PART is great. Most good government groups think it is great. It is an instrument. It has had blunt; will get better every year.

Most people that observe Congress, that have been around Congress a long time, believe that the Executive Branch is more interested in how well programs work than Congress is. David Walker has said that in hearings; so have Dick Armey and others. I would bet you agree. It is very hard to produce performance information and program assessments. What the Administration has done with PART is a place to start. We have been working 5 years on this. I do not believe Congress is going to invest 5 years to put together the information that we have right now. The PART information is a starting point for building better mechanisms to holding agencies and programs accountable for what they do.

So I believe, in spite of its flaws, that PART is an excellent tool. It is a wonderful beginning. It is the product of 5 years of effort. I do not see this as a power grab by the Executive Branch. The subject of this hearing is why won't Congress pay attention to PART, so I don't think Congress is actually reeling with this onslaught of performance information from the Executive Branch. Our challenge is to get them to pay attention to it.

Senator COBURN. Thank you. Ms. Norcross.

Ms. NORCROSS. What is the alternative to not using performance information? PART has given us—at least we have moved the discussion away from the policy preferences of an administration to-

wards evaluating programmatic activity. So I do not know what the option would be. Should we revert back to a system where we simply do not use performance information, expect it, gather it, or analyze it? And if there is discomfort with OMB performing these assessments, perhaps Congress should undertake that.

I understand Congress only engages about 7 percent of its time in oversight. So the current legislative mechanisms that are supposed to be engaged in this activity are not working up to speed. So that would simply be my response is if Congress is supposed to be evaluating these programs, where is the evidence that it is, in fact, evaluating them and providing guidance to agencies along the lines of performance?

Senator COBURN. All right; thank you.

Mr. Hughes, you mentioned that there is significant bias and distortion and manipulation. Would you give me examples of bias, please?

Mr. HUGHES. Sure. There are a number of different types of biases that can be involved in this. One is the perspective of the OMB officer. The budget officer at OMB is the person who has the final say on what the language will be for the answers to the questions, how that language that is written will translate into a yes or a no or a few of the modified answers that are possible now under the PART and also how those yeses and noes get translated not only into the numeric raw score but also into the actual rating.

There are a lot of inconsistencies between the guidelines that have been laid down for what raw score equals what rating and what the programs that have been reviewed actually get. That is one type of bias, and that is from a kind of implementation perspective. There are other biases in the actual design of the tool. I think that the format under which it was designed, which was designed to be accessible to people who may not be policy experts or who want to just know, like you say, come and look and see whether the government is getting results and whether the program is working, that necessitates that certain things are left out.

One of those things is whether the Congress has designed a program to have multiple goals. Many programs in the Federal Government are designed to have multiple goals. That sort of thing is not taken into consideration within the PART. Oftentimes, those goals can be conflicting. That does not necessarily mean that it is a bad design. That just means that it is a complex program. And that kind of complexity is lost in the way the tool was designed to apply to people who may not be policy experts.

Senator COBURN. Are you saying that there could be another PART program that would take into account for that? What is wrong by demanding a clear program mission from agencies?

Mr. HUGHES. Certainly nothing.

Senator COBURN. And questioning how a program fulfills that goal; is there anything wrong with those two things?

Mr. HUGHES. No.

Senator COBURN. So you do not disagree that a PART program might be designed better to take out more bias, but you do not disagree with the fact that knowing what a program's goal is and measuring performance against that goal, it should be an effective tool. You would not disagree with that?

Mr. HUGHES. No, theoretically, I agree with you.

Senator COBURN. The one problem I had with your testimony is the problem I have with the rest of Congress is we are lazy. And the fact is that this is the 37th oversight hearing of this Subcommittee. Go find another one that has done that. And the point is that ideally, Congress does have the responsibility, but they do not live up to it. And so, what we are working with is in a vacuum, is Congress ideally should be doing this. I do not disagree with you, but they are not.

And to have a blunt tool that is getting better, even though it can be criticized, and I think Mr. Johnson would agree that it is subject to some criticism, as is any assessment tool when you first start using it. But to say we should not have them doing it because it is Congress' role—I agree; that is why I am doing it; that is why we have done 37 of them, begs the question of how do we motivate Congress to do oversight?

So if we are critical of this one, answer me the question how I motivate my peers to do the appropriate thing when it comes to authorizing a program, and in that authorizing, saying we are going to measure it and then having the incentive to have Congress do the oversight to see whether or not they have a goal, and they are meeting that goal.

Mr. HUGHES. That is, of course, a very difficult question, one that I will probably be very insufficient in answering, giving a satisfactory answer for. I think that the oversight role of Congress, and you are correct, of course, in citing the fact that Congress does not really do oversight any more. That is indicative of larger things about our political system, about the way that the electoral process works, about the importance of fundraising. There are multiple things that are in there that actually have nothing to do with whether Congress should do oversight or not that are enormous problems that would be difficult to tackle.

I think your leadership on this issue is important. I think we need to have more folks in Congress who are paying attention to these sorts of issues. I do not know if there is a magic bullet procedural change or a statute or something that we could do that would make it so that Congress would be forced to do oversight more. I do think that some of the suggestions that have been made in front of this Subcommittee in the past about taking the Program Assessment Rating Tool or a modified version of it outside of the Office of Management and Budget, perhaps maybe having the Government Accountability Office do it or establishing a committee within Congress that would provide oversight in that regard. I think those ideas are worth exploring.

I do not think that you can just remove the PART the way it exists now and give it to GAO and have it work well. I think there are design flaws that need to be corrected, that need to be adjusted, and I am, of course, sympathetic to the point that if you change it too much, the previous reviews would not be as useful. But it is not necessarily just a problem with the way that the tool gets done at OMB. We think there are deficiencies within the way it was designed as well.

Senator COBURN. Well, you would be agreeable, then, to submit to this Subcommittee the things that you think are deficient in the design so that we can look at that?

Mr. HUGHES. Sure, and that was reflected in my written testimony. There is a section on that.

Senator COBURN. One of the problems with oversight is that a lot of agencies do not respond to our questions. Let us say we had oversight, and they do not respond. The only way you can solve that is either have somebody who can squeeze them on their money, or we have to squeeze them until they respond. But that requires the sausage-making process to be able to accomplish that.

The thing that is disconcerting is I have little faith that Congress is going to step up to the bar until they are absolutely forced to through a financial disaster to make the hard choices. Congress wants to avoid hard choices, and as long as they do not feel the pinch, they will not make the hard choices. And that is why 2016 is going to be a very tough year for this country, because that is when the pinch starts, the big pinch. And so, having an assessment tool, blunt, maybe somewhat biased, maybe somewhat distorted is, in my mind, better than nothing at all.

Mr. Johnson, and you may not care to comment on this, but you might comment on the motivation behind it: The House Subcommittee on Labor/HHS put a prohibition in their bill this past week that precludes any money from being spent on the PART assessment. Any comments on the motivation behind that or what you see? I am not trying to create a problem for you with the Subcommittee, but how did we get there?

Mr. JOHNSON. Well, there is one unelected staff member who is opposed to the PART. He worked on the Treasury/Transportation bill last year and put a similar prohibition in there. He was at HUD before that, and he disagreed with HUD's use of the PART, and he was at OMB before that. One unelected staff member is responsible for the provision. The chairman of the committee had no knowledge that it was in the bill. It is inexplicable to me that language like that is in the bill. That is my only comment.

Senator COBURN. OK; one of the other things, Mr. Hughes, with your testimony which I find, well, less than congruent is the statement that the PART increases the White House's power. And the problem with that is Congress ignores the PART assessment. We have been able to do nothing with the PART assessment. Even when I look at all of it, and I look at the agencies, and I have done the oversight, and I try to get somebody to do something about it, Congress ignores it.

So there is not a power grab there, because Congress is not paying any attention to it. So explain to me your reasoning behind—is it a potential? Because it is certainly not, in fact, acted out. There is no effect of the PART right now on the Congress, because they ignore it.

Mr. HUGHES. I actually would agree with you, and I would say that would probably be a poor choice of words on my part. I do think it is a potential problem. Let us do a for instance. Let us suppose that Congress will appropriate funds according to whatever the rating on the PART is. Why do we even need Congress? Let's just let OMB do it. So I think it is a slippery slope. I think that

particularly with respect to budgeting, we have been working more in trying to explore the management side of it as well, of PART, and the usefulness within agencies.

I think there is more potential for a productive use of the information there. I do not think that you can look at a PART score and say, OK, well, I know how to fund programs now because of these problems. So the way I chose my words is probably poor. I do not think it is a problem right now; as you say, you are correct, that Congress does not pay attention to them.

Senator COBURN. Well, but let me create a scenario for you. Let us say that Congress is doing great oversight on everything. We are sunsetting things; we are reauthorizing them; we are bringing them back up; we really know what we are doing and that we are doing a good job of that. Let's make that assumption. That is an absolute lie, but let's make that assumption.

Would you deny the fact that the Administration should have a performance tool themselves to measure what the goal is of the program and whether or not they are meeting that goal as a management tool to become more effective in carrying out the will of the Congress?

Mr. HUGHES. No; I think the problem exists when the tool that the Administration designs, or it does not even have to be this one, the Executive Branch designs portrays itself as an unbiased, objective evaluation of how programs and management are going at agencies when, in fact, it is anything but that. So I do not think that—again, in theory, that this is necessarily a problem. But with this particular instance, it is kind of like a wolf in sheep's clothing. You have a situation where they are saying we are doing this; it is systematic; it is transparent; it is on the Web; the public can view it; this is an innate good.

But the kind of things that we worry about are the things that are not transparent within the PART, that you do not necessarily see up front when you look at the one-page review. That is where you get into a tricky situation, and it is perfectly fine for the Executive Branch to have their own systems and whatever they like, but the problem occurs when they try to sell that to Congress as the one objective evaluator.

Senator COBURN. But they have not. They have just said since you are not doing one, we are going to do one, and here is what we have found, and here is what our recommendation is. We still control the purse strings, and it is obvious from the PART assessment that Congress has totally ignored the Administration when it comes to evaluating programs. So that is not seen as a risk to me whatsoever.

Mr. HUGHES. Well, that is encouraging to hear.

Senator COBURN. Well, they have not.

Mr. HUGHES. Well, I would say that they have not succeeded.

Senator COBURN. I think it is very discouraging to hear, because they are not looking at the other as well.

Mr. HUGHES. Fair enough.

Senator COBURN. They are paying attention to nothing and continue it. One of the battles I have, and I will share it in the Subcommittee, is there are a lot of bills that I block; they are authorizing bills. And I go to Members of the Senate, and I say these are

the things that I have problems with. And they say, well, why do you have problems? And I say, well, you have not looked at the programs that are already there before you authorize another program, and you have not said we are going to eliminate this program and put this one in. You are authorizing another program to do the same thing that is already happening without deauthorizing another program.

And what I get told: Well, we do not do things that way. Well, the American people do things that way. Business does things that way. States do things that way. Why should Congress not do it? So, really, we are shooting the messenger here. The messenger—there is a vacuum in terms of oversight, and we now have an Administration that has attempted, whether we think their tool is good or not. And you do not doubt that the tool is getting better as they have used it? They are using a tool that is improving, that does have maybe some bias and does have some risk for manipulation in it, but the fact is it is the only thing available right now, especially since this Subcommittee has time getting even agencies to come and testify before it or to give us information.

Mr. HUGHES. I will respond with two points. One, your shooting the messenger analogy, I think that may be part of our criticism of it, but our problem with it is that when the messenger leaves with his message, and when he gets to his destination, he is carrying two different messages. There is a problem with the transmission along the way, and that is something that is important to realize, regardless of where the criticisms are being pointed at.

I think the second thing is, and I sympathize with your frustrations about oversight in Congress, and that is certainly something that we would like to see a ton more of. I think you can kind of get around some of the rhetoric around what government—we have all these programs, and they do not do anything that is important. If we had more oversight, if we had more openness about what the government actually does, I think people will actually have a greater appreciation of things.

Senator COBURN. Right.

Mr. HUGHES. So I think our criticism—try to be focused on this particular instance of PART, the way that this PART assessment works. I do not think that it should be thrown in the garbage can. I think that it is very important that people in Congress and people in the agencies and the public know that this should be, despite the fact that there is not a lot going on elsewhere, this should be a really tiny part about evaluating how government works. That would be my caveat about—I am sympathetic to the fact that it is not going on elsewhere, but try not to latch on to it and say this is the tool, and this is what is going to get us there.

Senator COBURN. Nobody has in Congress. Would all three of you agree that some type of assessment of goals and measurement against the goals changes expectations of program managers?

Mr. JOHNSON. I agree totally.

Senator COBURN. Ms. Norcross.

Ms. NORCROSS. Totally correct.

Senator COBURN. Mr. Hughes.

Mr. HUGHES. In my limited experience, I would say that is right.

Senator COBURN. Ms. Norcross, you have some experience with performance tools in New Zealand, and I also know that South Korea has adopted assessment programs. Could you comment on those two things?

Ms. NORCROSS. Morris McTeague, with whom I work at the Government Accountability Program, has direct experience with the New Zealand experience in developing performance information systems and applying them to remedy some of New Zealand's budget crises. And if I could answer that question later, I could get you more information in specific on some of the reforms that they have undertaken. We are right now doing an analysis of that.

Senator COBURN. OK.

Ms. NORCROSS. So I could provide that for you.

Senator COBURN. Mr. Johnson, the question of bias in the instrument that you use, give us an example of three or four of the questions that PART asks about programs.

Mr. JOHNSON. Well, it asks if the program has a clear definition of—this is not exact wording, but it asks about do you have a clear definition of success? Does it have a good way of measuring your performance relative to that? Is it meeting its performance goals? It asks about the quality of management the program has. Do the program have an efficiency goal? Is it Management 101, or it is Accountability 101?

These assessments are put together by the agency and OMB, not by OMB alone. The agency and OMB are supposed to agree on the program performance goals. Just as agencies are afraid to disagree with Congress, agencies are sometimes afraid to disagree with OMB about its assessment, But if they really disagree with the assessment, agencies can submit their disagreements to an appeals board that I chair and that is made up of deputy secretaries from four or five agencies. We get a number of appeals every year, and we review them. Some of them, we approve, and some of them, we reject.

And we also conduct what we call consistency checks, where we review if the PART follows the rules we have for answering the questions. We also review whether the answers in a PART are consistent with each other.

As we look also at programs dealing with the same subject across agencies, we pull all the relevant program assessments together when we start doing a cross-cut analyses to make sure we are equally attentive to the issues, equally focused on the quality of the performance measures and so forth, because we are going to be using this information to compare one program to another. We use cross-cutting analyses to see if there is something an ineffective program can learn from an effective program dealing with the same topic.

So there is a lot of effort to make the assessments consistent; to make the information reliable; and to remove bias. There is bias in anything a human being does. So I have no doubt that these are not perfect instruments, but they will get better over time, and the assessments that we have done in the last 2 years are better than the assessments we did in the first 2 years.

Senator COBURN. So, in other words, the programs set their own outcome measurements.

Mr. JOHNSON. The program does. OMB and the program staff have to agree that the performance measures are acceptable.

Senator COBURN. And then, they measure themselves against it.

Mr. JOHNSON. They then determine the metrics they will use to measure performance and how to collect the data, and how often to collect that data.

Senator COBURN. And so, where is the bias in that?

Mr. JOHNSON. I do not know; just because——

Senator COBURN. If they are participating in setting the goal, and they are participating in setting the metrics, and they are the ones doing the measurement of the metrics, where is the bias?

Mr. JOHNSON. I do not know where there is unusual bias. I know that there is a bias in anything that human beings are involved in. So I do not know what specifically Mr. Hughes is talking about.

Senator COBURN. In teaching to the test, a problem across agencies as they respond to PART questions, Mr. Hughes wrote in his testimony that agency officials told him they gamed the system to avoid negative scores and consequences. Do you think that is true? Is there something in the program to help alleviate? We know everybody games when they are being measured to an extent. Are there things in the PART assessment system that take that into account?

Mr. JOHNSON. I know that agencies like to be green. They really like to be green, and they really like to have good PART scores. And so, they do a lot of things to please OMB and to get good scores and to look good on that scorecard.

Senator COBURN. Does that carry out into changed programs and changed management to make the programs more effective to deliver better process and therefore better response by the government to the very people they are supposed to be helping?

Mr. JOHNSON. Well, teaching to the test, gaming the PART to get a good score and providing only superficial analysis does not help the program work better. One approach that we have to improve the quality of program performance is shine a real big light on it all, which is one of the primary reasons we took all this assessment information, summarized it, and put it on the Website ExpectMore.Gov for all the world to see and for people to look up and say that is not the way I know the program works. An employee can look at it, or someone served by the program can look at it and say, well, that program does not work very well; it is ineffective as far as I am concerned, and they can complain to the agency or complain to OMB or complain to their Senator or Congressman.

Shining a lot of light on how the program is assessed, on what performance measures are used, and on what the performance information says can drive improvements in the measures that are used, the data that are used, and the quality assessment. So that is why I believe it is so important to have No. 1 on your list posted on this sign, which is transparency. You can have all of this information, but unless we shine a really big light on it, it will not get better over time. That is why we took it with all of its warts, with all of its dimples, put it out there. Now, let us begin the process with agencies and with Congress, I would hope, to improve this program performance information, to make these assessments bet-

ter and to make our plans to help program perform better more aggressive.

Senator COBURN. Let me follow up on that for a minute. So Mr. Hughes can go to every PART assessment and via the government Website can look at the goals, the metrics, the measurement of the metrics, the response, and the rating.

Mr. JOHNSON. Right.

Senator COBURN. In other words, nothing is hidden. Everything that comes to develop that, that can be accessed by OMB Watch, so they can see all that.

Mr. JOHNSON. Right.

Senator COBURN. Right; OK.

Mr. JOHNSON. There is a one-page summary of every PART on the ExpectMore.gov. There are links, at the bottom to the detailed PART, which is multiple pages. It is written in OMB-speak and has historical information and more detailed information. That is the meat of the assessment. The summary and all the details is available on ExpectMore.gov.

Senator COBURN. But they can get access——

Mr. JOHNSON. Yes, sir.

Senator COBURN [continuing]. If they need to as well.

And so, given your emphasis on transparency, are we to assume that you are going to be very accepting of our OMB transparency bill that Senator Carper and I have.

Mr. JOHNSON. That is on the contracting information.

Senator COBURN. Online grants, contracting, everything.

Mr. JOHNSON. We love transparency, and we are working very effectively with your staff to figure out——

Senator COBURN. I understand that.

Mr. JOHNSON [continuing]. The best way to get there as soon as possible. We are big on transparency and shining the light of day on performance to a strengthen accountability.

Senator COBURN. Any other comments from any of our panelists?

Mr. HUGHES. If I could just respond to some of the stuff that we have been talking about, the bias in the data and how the data, which data is important and which data is not important, there is a tension between outcomes and outputs in any type of performance management initiative. The PART focuses on outcomes, which is certainly a good goal. We think it is more of a broad government-wide goal, maybe something that should be included in something like the GPRA, the Government Performance and Results Act.

A lot of times, you cannot judge the effectiveness of programs based purely on outcomes, and I will give you a couple of examples. One program that is run out of, I believe it is the National Park Service, is an office that works as a consultant with local communities to transform the neglected or unused areas into public space: Parks, playgrounds, those sorts of things. They have collected, long before PART came along. And another thing that you should know is that agencies have collected performance data before PART came along. It is not like they were not doing this to begin with, and then, all of a sudden, OMB says you have to do it.

Senator COBURN. Some were not.

Mr. HUGHES. Some were not; that is correct; not all of them.

They had a couple of standards by which they judged whether they are doing a good job in this program that acts as a consultant. One is through surveys with the local communities that they consult with: Were you satisfied? Did people use the parks? Did you like the services we provided? Another way is they used to collect data about based on the amount of money that they were given, how many square acres of parks did they create? How many miles of jogging trails, those sorts of things.

Those are outputs, the second part. The survey part could be both. OMB, in the PART process, wanted them to focus on outcomes. And one of the things that they said should be an outcome was, Are the people living in the community healthier? And I think that is a perfectly good goal. I think people should be healthier. But the program in the National Park Service has no way to force people to go and jog in the park. All they can do is say this is the money we got to create parks for communities. These are the parks we created. These are the people we worked with and what they thought of what we did.

That is one instance, one example of multiple examples of the difference between outcomes and outputs and how certain programs are not necessarily structured to focus completely on outcomes, or maybe the outcomes are beyond their control. And one other example I will share about bias within the PART is the Appalachian Regional Commission. This is a program that Congress decided to be a patchwork, to cover the holes between other programs that were working in similar issue areas.

Senator COBURN. I have been trying to get rid of it for 10 years.

Mr. HUGHES. I am aware of that. [Laughter.]

And I do not have a personal perspective on the Regional Commission myself, but Congress designed this program to fill in the blanks, in the holes between programs, and the PART assessment said that this is not a unique program, which of course, it is not, because Congress designed it not to be unique. It was designed to be duplicative, because the evidence that Congress had seen at the time said that there are things that are being missed.

And we can talk more about, maybe we should have just pulled all of the programs together and redesigned them so that the holes are not missed, and that is certainly something that OMB is trying to do.

Senator COBURN. The whole point is you raise the question about what Congress has not done so that they will do it better. And to say that it is a blunt—it is a blunt tool, but it raises it up to a level so that somebody has to now—let's address this, and we have not addressed the Appalachian Regional Commission. What we have done is we have let it continue to do exactly what it does, and the danger with that is: One, we are not efficient; two, we could design a program that helps a whole lot more people with the same dollars; or three is we could help the same amount of people with a whole lot less dollars, which gives us dollars to help somebody else somewhere.

So the point is that is a commission that I am very well informed on, and I believe even the blunt tool will show that we could be much wiser as Congress to make the goals of that program more effective. I believe outcomes is the measure. I believe the American

people want outcomes as the measure. But part of it is laziness on our part. When we write a program, and this is something I am critical of Congress, we ought to be very specific about what our intentions are, and we are not. We ought to be very specific about what we expect, and we are not. We ought to be very specific on how we want to measure whether we got what we expected, and we are not.

So a lot of the problems do not have anything to do with you all in front of us; they have to do with Congress not being good legislators so that we design a system that can be looked at later and say did we go after what we intended to go after? Did we accomplish what we intended? And did we do it in a way or within the cost parameters that we thought it would?

And so, the real criticism is not at OMB. They are dealing with what we have dealt them. The real criticism is for us in not being specific enough in terms of—and you can ask staff: When I write a bill, I want it all the way down to the T. I want limited discretion, because if we are going to write a bill and do not know enough about it, we should not be writing the bill until we get the information to write a bill correctly. And I do not know many people who would disagree with that. It is just easier to write it loose and let somebody else worry with the details, and that is called lazy legislating.

Mr. HUGHES. That has been our experience working with you on the transparency bill as well. I think, though, that it is not necessarily as easy as you might make it out to seem. There is another example: The Consumer Product Safety Commission was ruled down on the PART review for not using cost-benefit analysis in its regulatory rulemaking. It is actually prohibited by Supreme Court decision from using cost-benefit changes like that in their rulemaking.

That is not something that the program can control. And I agree with you that it is good that even with the Appalachian Regional Commission example that these things are brought up to Congress. But as you have said many times, with the lack of the kind of investigatory role that Congress is playing now into how programs are made, we are concerned that the information that we all admit has some biases and those sorts of things will be taken as a snapshot, and the investigation will not be done to get underneath what the rating is.

So the ineffective for the Consumer Product Safety Commission, it should not be said we should get rid of it. But it is the bias in the tool that gives you the ineffective.

Senator COBURN. But experience tells us that is not happening, because Congress is not paying attention to PART, and they are not paying attention to their own. They are ignoring them both. So your fear is unfounded, because we are not using it.

Mr. HUGHES. Well, we would like to be vigilant.

Senator COBURN. We should do both, though. We should be using theirs plus our own, and that is the point. Outcomes, to me, is the measurement, not outputs. And outcomes, if we design something to have an outcome, then, we ought to know what that outcome measurement is, and then, we ought to hold agencies accountable to be to that outcome. And I will just give you a great example:

How about the incidence of HIV reduction in this country, which has not happened, and then, we spend money on flirting classes? And there is no connection between the two. In other words, if somebody is going to measure outcome, we ought to be asking why, with all of the money we are spending on HIV that we are not seeing a reduction in the incidence of new HIV cases in this country.

And yet, nobody is measuring the performance against that outcome, and that is an outcome that makes a difference in lives. It is not outputs; yes, we are spending a lot of money, but we are not measuring outcomes, and therefore, we are not getting the ability to make the programmatic changes that need to be made on the congressional side to accomplish that.

If the court prohibits a program from operating well, OK, if it prohibits a program from operating well, that tells us we have a problem in the design of the program.

Mr. HUGHES. Well, I would disagree with that classification. I do not think it is prohibiting it from operating well. I think it is saying that the program needs to take certain considerations into account when it does operate. It needs to say there are certain things, equity issues within programs in the Federal Government that are important to take a look at. It is not necessarily that the Supreme Court is putting up a roadblock in front of them getting the job done. The Supreme Court is making a value judgment about how the program should operate.

Senator COBURN. Which is not the Supreme Court's job. The Supreme Court's job is to interpret the laws and the Constitution and the treaties, not to tell Congress how to run the budget of the country, and that is——

Mr. HUGHES. And I would also say to you, too, that it is not OMB's job to tell you how to run the budget of the country either.

Senator COBURN. No, we agree.

Mr. HUGHES. Yes.

Senator COBURN. You will not disagree with me on that at all. I believe we have abdicated our responsibility, and the reason OMB is having to do this is because we have not. But I have no heartburn with somebody doing it somewhere. At least we have some information with which to make a decision.

I want to thank each of you for being here. I would like a little more formal response from your organization on specifics on how you would definitely change an assessment tool program for agencies and what we might be able to accomplish that would limit, and I want that as justifiable constructive criticism so that when we look at PART, we can have your thoughts in detail on how we can assess that and maybe make recommendations.

Mr. HUGHES. We look forward to that opportunity.

Senator COBURN. All right.

Thank you all so much for being here. The hearing is adjourned.

[Whereupon, at 3:43 p.m., the Subcommittee adjourned.]

APPENDIX

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF MANAGEMENT AND BUDGET
WASHINGTON, D.C. 20503

Testimony of the Honorable Clay Johnson III
Deputy Director for Management
Office of Management and Budget

before the

Subcommittee on Financial Management,
Government Information, and International Security
of the
Committee on Governmental Affairs
United States Senate

June 13, 2006

The Federal Government Wants to be Held Accountable

Americans deserve to have the government spend their hard earned tax dollars effectively, and better every year. The President, every member of Congress and all Federal employees need to be held accountable for getting results with the money they spend.

The PART – How We Figure Out What's Working and What's Not

To find out what's working and what's not, OMB and agency officials work together to determine whether a program:

- Has a clear purpose and a sound design;
- Sets outcome-oriented and suitably aggressive goals;
- Is well managed; and
- Achieves its goals.

This assessment is done systematically through the Program Assessment Rating Tool – PART. It is a set of common questions that are asked of every program, though it also includes additional questions for certain types of programs such as credit programs or competitive grants programs. The questions aim to identify a program's strengths and weaknesses so agencies can better identify actions needed to improve the program.

A key element of each assessment is defining the program's performance goals and determining whether they are being achieved. Performance goals are central to the PART. Through the PART process, OMB and agencies ensure all programs have a clear definitions of success and that they have outcome-oriented performance measures to judge their success. In order to achieve the most accurate program assessment, the PART process is collaborative. Agency and OMB staff work together and consider all available data in determining the answers to the questions. This supporting data is explained and cited in the detailed PART, which is available for public scrutiny at ExpectMore.gov.

The answers to the questions are used to generate an overall score for the program. Based on numeric ranges, the overall score is then translated into one of four qualitative ratings: Effective, Moderately Effective, Adequate, and Ineffective. If a program has not been able to develop outcome-oriented performance measures or collect performance information to measure performance against those goals, it receives a Results Not Demonstrated rating.

Whether a program is rated Effective or Ineffective, we are constantly looking for ways to improve its performance. Every program commits to taking steps to improve its performance and get more for taxpayer dollars every year. Some are more aggressive than others and we are working to strengthen these improvement plans.

ExpectMore.gov = Transparency = Accountability

Summary and detailed information about all assessed programs is posted to **ExpectMore.gov**, a website launched with the release of the President's FY07 Budget. The site is the most comprehensive source for information about programs we've assessed and their plans to improve. The purpose of this website is to provide easily understandable, candid information about which programs work, which programs don't, and what they are all doing to improve.

Currently, the ratings on **ExpectMore.gov** show that more than 70 percent of Federal programs are performing. A program that enhances highway safety provides a clear example of a program that demonstrates improved results. To reduce fatalities from automobile accidents, the National Highway Traffic Safety Administration promotes greater seat belt use among high-risk groups such as younger drivers, rural populations, pick-up truck occupants, 8–15 year-old passengers, occasional safety belt users, and motor vehicle occupants in states with secondary safety belt use laws. As a result, nationwide seat belt use increased from 73 percent in 2001 to 82 percent in 2005, which is an all-time high. This saves lives.

However, almost 30 percent of all programs are either ineffective or cannot demonstrate their success. A youth employment program created under the Workforce Investment Act demonstrates the need for improvement. The program awards grants for America's neediest youth to successfully transition to the workplace. The program is currently rated as ineffective. It does not have authority to target or reallocate resources to areas of greatest need and duplicates other programs. To remedy this problem, the Administration is working with Congress to gain increased authority to reallocate resources to areas of need. The Administration has also proposed legislation to consolidate this program with other Department of Labor job training grants. This will reduce overhead, ensure that more funds go directly to participants, and give States the flexibility to design processes that best serve their citizens.

We believe the transparency provided by **ExpectMore.gov** creates more constructive dialogue about how to improve program performance, and extra incentive to perform. **ExpectMore.gov** is not targeted to Democrats or Republicans, liberals or conservatives. Its audience is all Americans.

Program Assessments and the Federal Budget

This past year, the Administration assessed an additional 20 percent of the government's programs, marking the fourth year in our effort to find out what works, what doesn't, and what we need to do to improve. Program assessments are a factor in budgeting, but they are one among many factors. No budget decision is made automatically based on a program's rating. It may be that a highly rated program is not a priority for this Administration; therefore the President may propose to decrease funding for the program. A poorly rated program may need additional funds to address a weakness uncovered in the assessment. If we believe a program has been demonstrated to be ineffective and can't be fixed, or has outlived its usefulness, the Administration may recommend Congress spend the money on higher priority programs. The attached table shows the funding recommendation by program rating and by program.

This year's budget calls for major reductions in, or total eliminations of 141 Federal programs, saving nearly $15 billion. There are a variety of reasons for these reductions, primarily they were not getting results or not fulfilling essential priorities. Reductions in these areas do not mean Americans should expect less from Federal agencies or programs. On the contrary, they should expect the government to give them more for their tax dollars. They should expect the government to become more effective and efficient each year.

One program the Administration proposes to terminate is the Advanced Technology Program (ATP), a grant program for businesses that was intended to develop new technologies for commercial use. A PART analysis for this program noted that there are many non-governmental entities investing in early stage technology development, such as corporate research labs, venture capital firms, angel investors, and universities. The program is no longer warranted in today's research and development environment. Federal subsidies to industry for ATP projects are not appropriate or necessary, given the growth of venture capital and other financing sources for high-tech projects and the profit incentive private entities have to commercialize new technologies.

The Administration also proposes to eliminate the Even Start program and redirect funds to programs that are likely to be more effective at improving early childhood education including Title I. Even Start's poor results on national evaluations over a number of years and Ineffective PART rating provide strong justification for terminating the program. The children and adults who participate in the program do not make greater literacy gains than non-participants. The most recent evaluation concluded that, while Even Start participants made small gains, they did not perform better than the comparison group that did not receive Even Start services.

Because the National Assessment of Vocational Education found no evidence that high school vocational courses themselves contribute to academic achievement or college enrollment, the Administration proposes to terminate this program as well. Under the PART, Vocational Education State Grants was rated Ineffective because it has produced little or no evidence of improved outcomes for students despite decades of Federal investment. While the Administration has urged Congress to reform the Vocational Education program, neither the House nor Senate reauthorization bills adopted significant reforms to the current program.

Americans deserve better than to have their tax dollars invested in ineffective programs.

Congress and the Focus on Results

Like the Executive Branch, Congress wants programs to work. I believe the PART can be useful to Congress in its appropriation, authorization and oversight of programs. In some cases, Members of Congress are making use of the information to improve programs. Even Start is a good example. In 2004, the Administration proposed to fund only continuation awards, based on PART findings, and to begin phasing out the program. In 2005, the Administration proposed termination. Congress provided the first funding cut for the program in 2005 (-$22 million),

reducing it from $247 million to $225 million. The Congress reduced the program further in 2006 to $99 million.

Certainly, we can do a better job of making the information available in a form that is more useful to Congress. The report accompanying the Treasury, Transportation, and Housing and Urban Development Appropriations Act that recently passed the House Committee on Appropriations stated: "[M]ost [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." While a harsh assessment, I agree that we can improve. We must do a better job of more clearly articulating our objectives, not only for programs, but about how we expect information about program performance to be used. We also must do a better job of providing information about program performance in a way that is useful to you, the Congress. ExpectMore.gov is a first step in that effort. I would be grateful for the Committee's suggestions on how we might do more.

How has the PART changed?

Like programs, the PART process will improve over time. Although the Administration has tried to keep PART questions constant so the performance of programs can be compared over time, we have adopted changes in the PART process. We have implemented better information technology solutions to make application of the PART less burdensome and more collaborative. We review each newly completed PART to ensure the answers are consistent with PART guidance. If agencies disagree with a PART assessment, they can appeal to a panel of senior agency officials. These steps and others will make the PART more reliable, less of a burden, and hopefully, more focused on identifying what steps programs need to take to become more effective.

Conclusion

The message is simply that we want our citizens to expect more from their Federal government, and we want to be held accountable for how programs perform and how aggressively they improve. Of course, we do.

PART RATINGS AND PROGRAM FUNDING LEVELS
SUMMARY
(dollars in millions)

	2006 Enacted	2007 Budget	Change from 2006	
Effective (124)	231,853	231,968	115	0.0%
Moderately Effective (231)	433,843	443,435	9,592	2.2%
Adequate (219)	335,470	334,744	-726	-0.2%
Ineffective (28)	16,930	14,105	-2,825	-16.7%
Results Not Demonstrated (191)	139,147	135,991	-3,156	-2.3%

* The Medicare program is excluded in the Moderately Effective category calculation above. When it is included in the calculations, the percentage change in funding is 6.7%.

PART RATINGS AND PROGRAM FUNDING LEVELS
EFFECTIVE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
1	Engineering and Technical Services for International Broadcasting	BBG	138	133	-5	-3.6%
2	Inspector General Oversight of Federal Health Benefits Program	OPM	11	15	4	36.4%
3	Smithsonian Institution Facilities Capital	Smithsonian	99	114	15	15.2%
4	Capital Security Construction Program	State	1,000	1,165	165	16.5%
5	Educational and Cultural Exchange Programs in Near East Asia and South Asia	State	177	194	17	9.6%
6	Global Educational and Cultural Exchanges	State	426	474	48	11.3%
7	Secret Service Domestic Protectees	DHS	819	903	84	10.3%
8	National Nuclear Security Administration Naval Reactors	DOE	782	795	13	1.7%
9	National Institutes of Health - Buildings and Facilities	HHS	89	89	0	0.0%
10	Migration and Refugee Assistance -- Protection	State	125	123	-2	-1.6%
11	United Nations High Commissioner for Refugees	State	250	253	3	1.2%
12	New Currency Manufacturing	Treasury	276	375	99	35.9%
13	Secret Service Protective Intelligence	DHS	355	73	-282	-79.4%
14	Maritime Security Program	DOT	155	155	0	0.0%
15	GSA - New Construction	GSA	1,060	893	-167	-15.8%
16	Construction and Operations of Research Facilities	NSF	514	579	65	12.6%
17	Information Technology Research	NSF	160	160	0	0.0%
18	Nanoscale Science and Engineering Research	NSF	245	229	-16	-6.5%
19	Polar Research Tools, Facilities and Logistics	NSF	307	345	38	12.4%
20	Air Force Depot Maintenance	DOD	3,533	3,772	239	6.8%
21	National Nuclear Security Administration. Global Initiatives for Proliferation Prevention	DOE	38	0	-38	-100.0%
22	National Assessment for Educational Progress	ED	89	93	4	4.5%
23	Research on Biocomplexity in the Environment	NSF	83	58	-25	-30.1%
24	Basic Energy Sciences	DOE	1,135	1,421	286	25.2%
25	Asset Management of GSA-Owned Real Property	GSA	2,780	2,670	-110	-4.0%
26	Nuclear Materials Users Licensing & Inspection	NRC	66	74	8	12.1%
27	Fundamental Science and Engineering Research	NSF	2,240	2,387	147	6.6%
28	Humanitarian Demining	State	55	64	9	16.4%
29	Migration and Refugee Assistance -- Other Population, Refugee and Migration Programs	State	103	83	-20	-19.4%
30	South Asia Military Assistance	State	316	319	3	0.9%
31	Alcohol and Tobacco Tax and Trade Bureau Collect the Revenue Program	Treasury	44	45	1	2.3%
32	Financial Management Service Payments	Treasury	271	282	11	4.1%
33	U.S. Mint Protection Program	Treasury	37	36	-1	-2.7%
34	Bureau of Economic Analysis	DOC	76	76	0	0.0%
35	Strategic Petroleum Reserve	DOE	875	172	-703	-80.3%
36	Bureau of Reclamation - Hydropower	DOI	69	71	2	2.9%
37	FAA Research, Engineering & Development	DOT	137	130	-7	-5.1%
38	Federal Transit Administration - Formula Grant Programs	DOT	5,086	5,395	309	6.1%
39	Indian Health Service Health Care Facilities Construction	HHS	38	18	-20	-52.6%
40	Federally Funded Research and Development Centers	NSF	187	193	6	3.2%
41	Visa and Consular Services	State	1,035	1,139	104	10.0%
42	Worldwide Security Upgrades	State	808	914	106	13.1%
43	Financial Management Service Collections	Treasury	17	17	0	0.0%
44	Military Force Management	DOD	113,649	110,776	-2,873	-2.5%
45	National Center for Education Statistics	ED	91	94	3	3.3%
46	Regulation of Federal Credit Unions	NCUA	89	91	2	2.2%
47	Peace Corps International Volunteerism	Peace Corps	338	343	5	1.5%
48	Contribution to the United Nations Development Programme	State	95	95	0	0.0%
49	Humanitarian Migrants to Israel	State	40	40	0	0.0%
50	Financial Management Service Debt Collection	Treasury	50	66	16	32.0%

PART RATINGS AND PROGRAM FUNDING LEVELS
EFFECTIVE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
51	Thrift Institution and Savings Association Supervision	Treasury	216	221	5	2.3%
52	Census Bureau Economic Census	DOC	68	82	14	20.6%
53	Air Force Aircraft Operations	DOD	6,455	7,755	1,300	20.1%
54	Army Land Forces Operations	DOD	10,649	10,426	-223	-2.1%
55	Energy Conservation Investment	DOD	50	60	10	20.0%
56	Navy/Marine Corps Air Operations	DOD	5,795	5,689	-106	-1.8%
57	US Geological Survey - Geographic Research, Investigations, and Remote Sensing	DOI	129	142	13	10.1%
58	United States Trustees	DOJ	212	236	24	11.3%
59	Saint Lawrence Seaway Development Corporation - Operations and Maintenance	DOT	47	40	-7	-14.9%
60	National Institutes of Health - Intramural Research	HHS	2,956	2,946	-10	-0.3%
61	Office of Child Support Enforcement	HHS	3,322	3,953	631	19.0%
62	NASA Astronomy and Astrophysics Research	NASA	378	1,516	1,138	301 1%
63	Fuel Facilities Licensing & Inspection	NRC	38	38	0	0.0%
64	Reactor Inspection and Performance Assessment	NRC	204	222	18	8.8%
65	Spent Fuel Storage and Transportation Licensing and Inspection	NRC	24	27	3	12.5%
66	Support for Individual Researchers	NSF	496	519	23	4.6%
67	Export Control Assistance	State	43	45	2	4.7%
68	US Trade and Development Agency	TDA	61	51	-10	-16.4%
69	National Bank Supervision	Treasury	579	605	26	4.5%
70	Coast Guard Domestic Icebreaking Program	DHS	67	58	-9	-13.4%
71	Secret Service Foreign Protectees and Foreign Missions	DHS	130	132	2	1.5%
72	Department of Defense Depot Maintenance Ship	DOD	4,042	3,882	-160	-4.0%
73	Bureau of Justice Statistics	DOJ	35	59	24	68 6%
74	Maritime Administration - State Maritime Schools	DOT	11	10	-1	-9.1%
75	National Institutes of Health - Extramural Research Programs	HHS	21,223	21,249	26	0.1%
76	Support for Research Institutions	NSF	147	147	0	0 0%
77	Support for Small Research Collaborations	NSF	388	404	16	4.1%
78	Nonproliferation and Disarmament Fund	State	37	38	1	2.7%
79	Rural Water and Wastewater Grants and Loans	USDA	1,604	1,414	-190	-11 8%
80	Broadcasting to Latin America	BBG	42	42	0	0.0%
81	Navy Ship Operations	DOD	5,186	5,536	350	6.7%
82	Advanced Simulation and Computing (ASCI)	DOE	660	630	-30	-4.5%
83	Elimination of Weapons-Grade Plutonium Production Program	DOE	174	207	33	19 0%
84	Bureau of Reclamation - Safety of Dams Program	DOI	65	70	5	7.7%
85	Minerals Management Service - Outer Continental Shelf Minerals Regulation and Compliance	DOI	51	54	3	5.9%
86	Bureau of Labor Statistics	DOL	537	563	26	4.8%
87	Indian Health Service Resource and Patient Management System	HHS	45	55	10	22 2%
88	Support for East European Democracy/Freedom Support Act	State	864	715	-149	-17 2%
89	Customs and Border Protection Security Inspections and Trade Facilitation	DHS	2,273	2,343	70	3.1%
90	Science and Technology Biological Countermeasures	DHS	372	337	-35	-9 4%
91	Census Bureau Current Demographic Statistics	DOC	77	52	-25	-32.5%
92	Depot Maintenance - Naval Aviation	DOD	977	991	14	1 4%
93	National Nuclear Security Administration Weapons Activities - Readiness Campaign	DOE	217	211	-6	-2 8%
94	Nonproliferation and International Security	DOE	74	127	53	71 6%
95	Bureau of Reclamation - Science and Technology Program	DOI	17	10	-7	-41.2%
96	Refugee Transitional and Medical Services	HHS	268	282	14	5 2%
97	Homeless Assistance Grants (Competitive)	HUD	1,327	1,536	209	15 7%
98	Solar System Exploration	NASA	1,582	1,603	21	1.3%
99	Disaster Loan Program	SBA	1,261	218	-1,043	-82.7%
100	Administering the Public Debt	Treasury	175	178	3	1.7%

PART RATINGS AND PROGRAM FUNDING LEVELS
EFFECTIVE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
101	U S. Mint Numismatic Program	Treasury	1,282	1,321	39	3.0%
102	Economic Research Service	USDA	75	83	8	10.7%
103	Plant and Animal Health Monitoring Programs	USDA	286	346	60	21.0%
104	African Development Foundation	ADF	46	52	6	13.0%
105	Census Bureau Survey Sample Redesign	DOC	11	11	0	0.0%
106	Defense Basic Research	DOD	1,476	1,420	-56	-3.8%
107	Biological and Environmental Research	DOE	580	510	-70	-12.1%
108	Contributions For International Peacekeeping Activities	State	1,022	1,135	113	11.1%
109	Non-Security Embassy Construction Program	State	0	0	0	0.0%
110	Refugee Admissions to the US	State	159	223	64	40.3%
111	Pest and Disease Exclusion	USDA	156	182	26	16.7%
112	Preparedness – Grants and Training Office National Exercise Program	DHS	52	42	-10	-19 2%
113	National Institute for Standards and Technology Laboratones	DOC	569	535	-34	-6 0%
114	Department of Defense Training and Education Programs - Basic Skills and Advanced Training	DOD	4,957	4,984	27	0 5%
115	International Nuclear Materials Protection and Cooperation	DOE	423	413	-10	-2 4%
116	Nuclear Physics	DOE	367	454	87	23.7%
117	Department of Justice General Legal Activities	DOJ	767	795	28	3.7%
118	CDC Global Immunizations	HHS	138	138	0	0.0%
119	Health Centers	HHS	1,782	1,963	181	10.2%
120	Medicare Integrity Program	HHS	820	829	9	1.1%
121	Anti-Terrorism Assistance	State	122	136	14	11.5%
122	Terrorist Interdiction Program	State	5	12	7	140 0%
123	U.S. Mint Coin Production	Treasury	567	573	6	1.1%
124	TVA Resource Stewardship	TVA	83	84	1	1 2%
	Total funding for Effective programs [1]		231,853	231,968	115	0.0%

[1] *Funding levels represent program level and include BA, obligations, user fees, loan levels, etc as appropriate*

PART RATINGS AND PROGRAM FUNDING LEVELS
MODERATELY EFFECTIVE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
1	Formerly Utilized Sites Remedial Action Program	Corps	139	130	-9	-6.5%
2	Advanced Scientific Computing Research	DOE	235	319	84	35.7%
3	Directed Stockpile Work (DSW)	DOE	1,372	1,380	8	0.6%
4	High Energy Physics	DOE	717	775	58	8.1%
5	National Criminal History Improvement Program	DOJ	10	39	29	290.0%
6	EPA Acid Rain Program	EPA	19	20	1	5.3%
7	NASA Earth-Sun System Research	NASA	2,164	2,211	47	2.2%
8	Internal Revenue Service Taxpayer Advocate Service	Treasury	164	164	0	0.0%
9	TVA Power	TVA	8,632	8,946	314	3.6%
10	Corps of Engineers Regulatory Program	Corps	158	173	15	9.5%
11	Census Bureau: Current Economic Statistics and Census of Governments	DOC	140	156	16	11.4%
12	National Oceanic & Atmospheric Administration Navigation Services	DOC	93	91	-2	-2.2%
13	Vehicle Technologies	DOE	182	166	-16	-8.8%
14	Bureau of Indian Affairs - Dam Safety and Dam Maintenance	DOI	22	22	0	0.0%
15	National Park Service - National Historic Preservation	DOI	96	89	-7	-7.3%
16	National Park Service - Natural Resource Stewardship	DOI	210	215	5	2.4%
17	US Geological Survey - Energy Resource Assessments	DOI	24	26	2	8.3%
18	Hazardous Materials Transportation Safety - Emergency Preparedness Grants	DOT	14	28	14	100.0%
19	Highway Research and Development/Intelligent Transportation Systems	DOT	430	430	0	0.0%
20	Maritime Administration - Merchant Marine Academy	DOT	64	62	-2	-3.1%
21	GSA - Real Property Leasing	GSA	4,443	4,483	40	0.9%
22	HIV/AIDS Research	HHS	2,904	2,988	84	2.9%
23	Peace Keeping Operations - Organization for Security and Cooperation in Europe Programs	State	30	30	0	0.0%
24	Emergency Pest and Disease Management Programs	USDA	267	284	17	6.4%
25	Broadcasting to Near East Asia and South Asia	BBG	149	159	10	6.7%
26	Coast Guard Marine Environmental Protection	DHS	374	332	-42	-11.2%
27	Census Bureau Intercensal Demographic Estimates	DOC	9	10	1	11.1%
28	Fusion Energy Sciences	DOE	288	319	31	10.8%
29	National Nuclear Security Administration Weapons Activities - Science Campaign	DOE	277	265	-12	-4.3%
30	Weatherization Assistance	DOE	243	160	-83	-34.2%
31	Minerals Management Service - Outer Continental Shelf Environmental Studies	DOI	16	20	4	25.0%
32	US Geological Survey - Geologic Hazard Assessments	DOI	81	82	1	1.2%
33	Arson and Explosives Program	DOJ	188	196	8	4.3%
34	Federal Lands Highway Program	DOT	806	903	97	12.0%
35	Federal Transit Administration New Starts	DOT	1,488	1,466	-22	-1.5%
36	Export Import Bank - Long Term Guarantees	EX-IM	474	203	-271	-57.2%
37	GSA - Vehicle Leasing	GSA	1,000	1,030	30	3.0%
38	National Center for Health Statistics	HHS	109	109	0	0.0%
39	Inter-American Foundation	IAF	37	37	0	0.0%
40	Military Assistance to New NATO and NATO Aspirant Nations	State	155	154	-1	-0.6%
41	Research on Protection and Safety of Agricultural Food Supply	USDA	392	388	-4	-1.0%
42	Snow Survey Water Supply Forecasting	USDA	11	11	0	0.0%
43	Coast Guard Migrant Interdiction Program	DHS	464	488	24	5.2%
44	Immigration and Customs Enforcement Detention and Removal	DHS	1,450	2,076	626	43.2%
45	Department of Defense Training and Education Programs - Accession Training	DOD	829	878	49	5.9%
46	Secure Transportation Asset (STA)	DOE	210	210	0	0.0%
47	Fish and Wildlife Service - National Fish Hatchery System	DOI	60	61	1	1.7%
48	US Geological Survey - National Cooperative Geological Mapping	DOI	25	26	1	4.0%
49	DOT Pipeline Safety	DOT	72	76	4	5.6%
50	Mobile Source Air Pollution Standards and Certification	EPA	62	71	9	14.5%

PART RATINGS AND PROGRAM FUNDING LEVELS
MODERATELY EFFECTIVE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
51	Administration on Aging	HHS	1,363	1,335	-28	-2.1%
52	Child Care and Development Fund	HHS	4,979	4,979	0	0.0%
53	National Community Development Initiative	HUD	30	0	-30	-100.0%
54	Veterans Health Research and Development	VA	786	765	-21	-2.7%
55	US Geological Survey - Mineral Resource Assessments	DOI	53	31	-22	-41.5%
56	Federal Motor Carrier Safety Administration - Operations and Programs	DOT	211	223	12	5 7%
57	Railroad Safety Program	DOT	143	143	0	0 0%
58	Maternal and Child Health Block Grant	HHS	693	693	0	0 0%
59	Office of Transition Initiatives	USAID	40	50	10	25.0%
60	Agricultural Credit Insurance Fund - Guaranteed Loans	USDA	2,796	2,498	-298	-10.7%
61	Mutual Self-Help Housing -- Technical Assistance Grants	USDA	34	38	4	11.8%
62	Generation IV Nuclear Energy Systems Initiative	DOE	55	32	-23	-41.8%
63	Criminal Justice Services	DOJ	286	411	125	43.7%
64	Pension Benefit Guaranty Corporation	DOL	374	398	24	6.4%
65	EPA Lead-Based Paint Risk Reduction Program	EPA	23	25	2	8 7%
66	Indian Health Service Sanitation Facilities Construction Program	HHS	92	94	2	2.2%
67	Center for Talent Services - HR Products and Services for Federal Agencies	OPM	159	167	8	5 0%
68	Federal Crop Insurance	USDA	3,365	4,212	847	25 2%
69	National Agricultural Statistics Service	USDA	139	153	14	10.1%
70	Protection and Safety of Agricultural Food Supply (Grants)	USDA	238	207	-31	-13.0%
71	Veterans Burial Benefits	VA	153	161	8	5.2%
72	Broadcasting to Africa	BBG	12	13	1	8.3%
74	Federal Protective Service	DHS	1,648	1,706	58	3.5%
75	Science and Technology Homeland Security University Fellowships	DHS	62	52	-10	-16 1%
76	National Oceanic & Atmospheric Administration Climate Program	DOC	156	180	24	15 4%
77	Air Combat Program	DOD	13,737	11,784	-1,953	-14.2%
78	Defense Housing	DOD	17,047	16,545	-502	-2 9%
79	Department of Defense Recruiting	DOD	3,973	4,060	87	2.2%
80	Bonneville Power Administration	DOE	-19	-11	8	-42.1%
81	Facilities and Infrastructure	DOE	283	294	11	3.9%
82	US Geological Survey - Biological Research and Monitoring	DOI	140	136	-4	-2.9%
83	H-1B Work Visa for Specialty Occupations - Labor Condition Application Program	DOL	5	6	1	20.0%
84	National Highway Traffic Safety Administration Grant Program	DOT	572	584	12	2.1%
85	GSA - Real Property Disposal	GSA	55	52	-3	-5.5%
86	NASA Aeronautics Technology	NASA	884	724	-160	-18 1%
87	Child Survival and Health - Population	USAID	440	295	-145	-33 0%
88	Economic Development Administration	DOC	250	297	47	18.8%
89	Airlift Program	DOD	5,771	5,367	-404	-7.0%
90	Inertial Confinement Fusion Ignition and High Yield Campaign/NIF Construction Project	DOE	544	451	-93	-17.1%
91	National Nuclear Security Administration Nonproliferation and Verification Research and Development	DOE	319	269	-50	-15.7%
92	Safeguards and Security	DOE	766	721	-45	-5 9%
93	Western Area Power Administration	DOE	1,117	1,283	166	14 9%
94	US Geological Survey - Biological Information Management and Delivery	DOI	24	22	-2	-8 3%
95	Firearms Programs -- Integrated Violence Reduction Strategy	DOJ	663	765	102	15.4%
96	Railroad Research and Development	DOT	55	35	-20	-36 4%
97	Food and Drug Administration	HHS	1,876	1,947	71	3.8%
98	Pharmaceutical Outcomes	HHS	26	26	0	0.0%
99	Homeownership Voucher	HUD	5	14	9	180 0%
100	Security Assistance to Sub-Saharan Africa	State	113	148	35	31.0%

PART RATINGS AND PROGRAM FUNDING LEVELS
MODERATELY EFFECTIVE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
101	Development Credit Authority	USAID	38	53	15	39.5%
102	Single Family Housing Direct Loans	USDA	1,131	1,237	106	9.4%
103	Corps of Engineers Emergency Management	Corps	4	85	81	2025.0%
104	Federal Emergency Management Agency - Mitigation Programs	DHS	5,481	5,526	45	0.8%
105	Science and Technology Rapid Prototyping of Countermeasures	DHS	31	10	-21	-67.7%
106	Census Bureau: Decennial Census	DOC	454	512	58	12.8%
107	National Oceanic & Atmospheric Administration Weather and Related Programs	DOC	1,779	1,908	129	7.3%
108	Advanced Fuel Cycle Initiative	DOE	80	250	170	212.5%
109	Veterans' Employment and Training State Grants	DOL	161	161	0	0.0%
110	FAA Grants-in-Aid for Airports	DOT	3,415	3,000	-415	-12.2%
111	TRIO Student Support Services	ED	273	273	0	0.0%
112	HOME (Affordable Housing Block Grant)	HUD	1,757	1,917	160	9.1%
113	Nonproliferation of Weapons of Mass Destruction Expertise	State	52	56	4	7.7%
114	Bank Secrecy Act Data Collection, Retrieval and Sharing	Treasury	17	29	12	70.6%
115	Internal Revenue Service Criminal Investigations	Treasury	718	728	10	1.4%
116	Internal Revenue Service Examinations	Treasury	3,598	3,658	60	1.7%
117	Internal Revenue Service Submission Processing	Treasury	1,117	1,112	-5	-0.4%
118	Agricultural Export Credit Guarantee Programs	USDA	133	117	-16	-12.0%
119	Food Stamp Program	USDA	33,890	36,003	2,113	6.2%
120	Forest Service Forest Legacy Program	USDA	57	62	5	8.8%
121	Small Business Development Centers	SBA	107	109	2	1.9%
121	Manufacturing Extension Partnership	DOC	105	46	-59	-56.2%
122	Readiness in Technical Base and Facilities (RTBF), Operations	DOE	1,161	1,220	59	5.1%
123	Wind Energy	DOE	39	44	5	12.8%
124	Minerals Management Service - Outer Continental Shelf Minerals Evaluation and Leasing	DOI	51	55	4	7.8%
125	National Park Service - Visitor Services	DOI	465	502	37	8.0%
126	Federal Employees Compensation Act	DOL	2,719	2,775	56	2.1%
127	National Highway Traffic Safety Administration - Operations and Research	DOT	235	231	-4	-1.7%
128	Improving Teacher Quality State Grants	ED	2,887	2,887	0	0.0%
129	EPA New Chemicals Program	EPA	15	14	-1	-6.7%
130	Health - Data Collection and Dissemination	HHS	63	66	3	4.8%
131	Indian Health Service Federally-Administered Activities	HHS	1,886	1,979	93	4.9%
132	National Bone Marrow Donor Registry	HHS	25	23	-2	-8.0%
133	Projects for Assistance in Transition from Homelessness	HHS	54	54	0	0.0%
134	Manufactured Housing and Standards	HUD	13	13	0	0.0%
135	Neighborhood Reinvestment Corporation	Neighborhood	117	120	3	2.6%
136	Service Corps of Retired Executives Small Business Assistance	SBA	19	21	2	10.5%
137	Security Assistance for the Western Hemisphere	State	125	118	-7	-5.6%
138	Child Survival and Health for Latin America and the Caribbean	USAID	141	128	-13	-9.2%
139	Development Assistance to Latin America and the Caribbean	USAID	257	167	-90	-35.0%
140	Forest Service: Recreation	USDA	275	250	-25	-9.1%
141	Research/Extension Grants, Economic Opportunities for Producers	USDA	435	354	-81	-18.6%
142	USDA Foreign Market Development Programs	USDA	249	149	-100	-40.2%
143	Broadcasting to East Asia and Eurasia	BBG	91	91	0	0.0%
144	Coast Guard Fisheries Enforcement	DHS	815	808	-7	-0.9%
145	Southwestern Power Administration	DOE	55	63	8	14.5%
146	Prisons Operations	DOJ	4,830	4,987	157	3.3%
147	Unemployment Insurance Administration State Grants	DOL	2,549	2,650	101	4.0%
148	GSA - Personal Property Management	GSA	32	33	1	3.1%
149	Health and Human Services - Office for Civil Rights	HHS	35	36	1	2.9%
150	Medicare	HHS	407,249	453,890	46,641	11.5%

PART RATINGS AND PROGRAM FUNDING LEVELS
MODERATELY EFFECTIVE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
151	Lead Hazard Grants	HUD	76	92	16	21.1%
152	Examining Compliance with Securities Laws	SEC	223	228	5	2.2%
153	Agricultural Marketing Loan Payments	USDA	5,124	4,444	-680	-13.3%
154	USDA Research Economic Opportunities for Producers	USDA	392	314	-78	-19.9%
155	Department of Veterans Affairs- General Administration	VA	296	313	17	5.7%
156	Distributed Energy Resources	DOE	56	30	-26	-46.4%
157	Hydrogen Technology	DOE	156	196	40	25.6%
158	Solar Energy	DOE	83	148	65	78.3%
159	Southeastern Power Administration	DOE	38	40	2	5.3%
160	Bureau of Indian Affairs - Job Placement and Training	DOI	8	8	0	0.0%
161	Bureau of Reclamation - Site Security	DOI	50	50	0	0.0%
162	US Geological Survey - Water Information Collection and Dissemination	DOI	64	66	2	3.1%
163	US Geological Survey - Water Resources Research	DOI	141	138	-3	-2.1%
164	Hazardous Materials Transportation Safety	DOT	26	27	1	3.8%
165	Adoption Assistance	HHS	1,883	2,044	161	8.6%
166	Reactor Licensing	NRC	264	341	77	29.2%
167	President's Emergency Plan For AIDS Relief: Focus Countries	State	1,774	2,776	1,002	56.5%
168	Natural Resources Conservation Service Plant Materials Research Centers	USDA	10	11	1	10.0%
169	Veterans Life Insurance	VA	4	4	0	0.0%
170	Corps of Engineers: Recreation Management	Corps	283	267	-16	-5.7%
171	Preparedness – Infrastructure Protection National Communications Service	DHS	286	286	0	0.0%
172	Bureau of Industry and Security	DOC	75	79	4	5.3%
173	U.S. Patent and Trademark Office - Trademarks	DOC	188	203	15	8.0%
174	Department of Defense Education Activity	DOD	1,776	1,821	45	2.5%
175	Federal Energy Management Program	DOE	19	17	-2	-10.5%
176	Bureau of Indian Affairs - Indian Land Consolidation	DOI	34	59	25	73.5%
177	Federal Motor Carrier Safety Administration Grant Program	DOT	279	294	15	5.4%
178	TRIO Talent Search	ED	143	0	-143	-100.0%
179	EPA Support for Cleanup of Federal Facilities	EPA	31	31	0	0.0%
180	Superfund Removal	EPA	240	250	10	4.2%
181	Family Planning	HHS	283	283	0	0.0%
182	Human Trafficking	HHS	10	15	5	50.0%
183	Strategic National Stockpile	HHS	525	592	67	12.8%
184	Fair Housing Assistance Program	HUD	20	25	5	25.0%
185	Housing Vouchers	HUD	15,418	15,920	502	3.3%
186	Human Systems Research and Technology	NASA	624	275	-349	-55.9%
187	Historically Underutilized Business Zone - HUBZone	SBA	7	10	3	42.9%
188	Protection of Foreign Missions and Officials	State	9	9	0	0.0%
189	Agricultural Credit Insurance Fund Direct Loans	USDA	951	930	-21	-2.2%
190	Environmental Quality Incentives Program	USDA	1,017	1,000	-17	-1.7%
191	Multi-Family Housing Programs	USDA	899	745	-154	-17.1%
192	Natural Resource Base & Environment (Grants)	USDA	191	162	-29	-15.2%
193	Perishable Agricultural Commodities Act	USDA	10	10	0	0.0%
194	Corps of Engineers: Coastal Ports and Harbors	Corps	1,136	1,061	-75	-6.6%
195	District of Columbia: Pretrial Services Agency	Court Services	42	46	4	9.5%
196	Defense Air Transportation System	DOD	7,462	6,820	-662	-8.8%
197	Geothermal Technology	DOE	23	0	-23	-100.0%
198	Bureau of Land Management - Wildlife Habitat Restoration	DOI	148	150	2	1.4%
199	FBI Counterintelligence Program	DOJ	Classified	Classified		
200	Black Lung Benefits Program	DOL	1,401	1,374	-27	-1 9%

PART RATINGS AND PROGRAM FUNDING LEVELS
MODERATELY EFFECTIVE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
201	Employee Benefits Security Administration	DOL	134	144	10	7.5%
202	EPA Indoor Air Quality	EPA	37	39	2	5.4%
203	Surface Water Protection	EPA	189	192	3	1 6%
204	Temporary Assistance for Needy Families (TANF)	HHS	17,058	17,158	100	0.6%
205	Universal Newborn Hearing Screening and Intervention Program	HHS	10	0	-10	-100 0%
206	International Space Station	NASA	1,753	1,812	59	3 4%
207	Supplemental Security Income	SSA	41,108	40,702	-406	-1.0%
208	Economic Support Fund for the Western Hemisphere	State	122	152	30	24.6%
209	Conservation Reserve Program	USDA	2,021	2,192	171	8.5%
210	Soil Survey Program	USDA	87	89	2	2.3%
211	Immigration Services	DHS	1,889	1,986	97	5.1%
212	Science and Technology: Emerging Homeland Security Threat Detection	DHS	8	9	1	12.5%
213	Defense Applied Research Program	DOD	5,188	4,478	-710	-13.7%
214	DoD Unmanned Aircraft Systems (UAS)	DOD	1,588	1,785	197	12.4%
215	Future Combat Systems/Modularity Land Warfare	DOD	9,623	10,349	726	7.5%
216	High Temperature Superducting (HTS) Research and Development	DOE	50	45	-5	-10.0%
217	Bureau of Reclamation - Water Reuse and Recycling	DOI	20	10	-10	-50.0%
218	Job Corps	DOL	1,586	1,531	-55	-3 5%
219	FAA Aviation Safety	DOT	941	1,007	66	7.0%
220	Highway Emergency Relief Program	DOT	2,850	100	-2,750	-96 5%
221	Highway Infrastructure	DOT	34,215	37,650	3,435	10.0%
222	Advanced Placement	ED	32	122	90	281.3%
223	Childrens Mental Health Services	HHS	104	104	0	0 0%
224	National Health Service Corps	HHS	126	126	0	0 0%
225	Protection and Advocacy for Individuals with Mental Illness	HHS	34	34	0	0.0%
226	Substance Abuse Prevention Projects of Regional and National Significance	HHS	193	181	-12	-6.2%
227	Social Security Disability Insurance	SSA	92,989	100,051	7,062	7.6%
228	US Agency for International Development Administration and Capital Investment	USAID	770	836	66	8.6%
229	Federal Grain Inspection Services	USDA	60	66	6	10.0%
230	School Breakfast Program	USDA	2,076	2,204	128	6.2%
231	USDA Food Aid Programs	USDA	394	250	-144	-36.5%
	Total funding for Moderately Effective programs [1]		841,092	897,325	56,233	6.7%
	Total funding for Moderately Effective programs (excluding Medicare) [1]		433,843	443,435	9,592	2.2%

[1] *Funding levels represent program level and include BA, obligations, user fees, loan levels, etc as appropriate*

PART RATINGS AND PROGRAM FUNDING LEVELS
ADEQUATE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
1	Science and Technology Standards Development for Homeland Security Technology	DHS	33	22	-11	-33 3%
2	Missile Defense	DOD	7,695	9,318	1,623	21.1%
3	Nuclear Power 2010	DOE	66	55	-11	-16.7%
4	Department of the Interior - Central Utah Project	DOI	34	40	6	17.6%
5	National Institute of Justice	DOJ	54	56	2	3.7%
6	Federal Pell Grants	ED	13,045	12,739	-306	-2 3%
7	GSA - Vehicle Acquisition	GSA	1,213	1,233	20	1.6%
8	Tribally-Operated Health Programs	HHS	1,648	1,720	72	4.4%
9	Urban Indian Health Program	HHS	33	0	-33	-100.0%
10	Section 7 (a) Guaranteed Loan Program	SBA	97	83	-14	-14.4%
11	Agricultural Crops Counter Cyclical Payments	USDA	5,893	6,661	768	13.0%
12	Single Family Housing Loan Guarantees	USDA	3,745	3,564	-181	-4.8%
13	National Oceanic & Atmospheric Administration Protected Areas	DOC	56	42	-14	-25.0%
14	Space Launch	DOD	1,175	1,255	80	6.8%
15	Industrial Technologies Program	DOE	57	46	-11	-19.3%
16	Charter Schools Grant	ED	215	215	0	0.0%
17	National Institute on Disability and Rehabilitation Research	ED	107	107	0	0.0%
18	EPA Climate Change Programs	EPA	111	105	-6	-5.4%
19	Leaking Underground Storage Tank Cleanup Program	EPA	68	69	1	1.5%
20	CDC Environmental Health	HHS	146	141	-5	-3.4%
21	Poison Control Centers	HHS	23	13	-10	-43.5%
22	Overseas Private Investment Corporation - Finance	OPIC	48	50	2	4.2%
23	Overseas Private Investment Corporation - Insurance	OPIC	-211	-209	2	-0 9%
24	Appalachian Regional Commission	Appalachian	65	65	0	0.0%
25	U.S. Patent and Trademark Office - Patents	DOC	1,495	1,640	145	9.7%
26	National Security Space Weather Programs	DOD	394	437	43	10.9%
27	Fish and Wildlife Service - Partners for Fish and Wildlife	DOI	50	43	-7	-14.0%
28	National Park Service - Facility Maintenance	DOI	1,034	933	-101	-9.8%
29	Criminal Enterprises	DOJ	788	750	-38	-4.8%
30	Stratospheric Ozone Protection	EPA	15	19	4	26 7%
31	GSA - Global Supply	GSA	1,032	1,051	19	1 8%
32	CDC Infectious Diseases	HHS	227	245	18	7.9%
33	Ryan White HIV/AIDS	HHS	2,083	2,158	75	3.6%
34	FHA Single-Family Mortgage Insurance	HUD	414	414	0	0.0%
35	Section 504 Certified Development Company Guaranteed Loan Program	SBA	38	23	-15	-39.5%
36	Economic Support Fund for Africa	State	123	164	41	33.3%
37	Farmland Protection Program	USDA	74	50	-24	-32 4%
38	Food Safety and Inspection Service	USDA	829	863	34	4.1%
39	Coast Guard Marine Safety	DHS	685	678	-7	-1 0%
40	Advanced Technology Program	DOC	79	0	-79	-100.0%
41	National Park Service - Cultural Resource Stewardship	DOI	96	99	3	3.1%
42	US Marshals Service- Protection of the Judicial Process	DOJ	340	343	3	0.9%
43	Agency for Toxic Substances and Disease Registry	HHS	75	75	0	0.0%
44	State Children's Health Insurance Program	HHS	4,316	5,040	724	16.8%
45	Space and Flight Support	NASA	339	367	28	8.3%
46	Economic Support Fund - Human Rights and Democracy Fund	State	63	35	-28	-44.4%
47	USAID's Development Assistance for Sub-Saharan Africa	USAID	594	563	-31	-5 2%
48	Rural Electric Utility Loans and Guarantees	USDA	3,893	3,842	-51	-1 3%
49	Wetlands Reserve Program	USDA	250	403	153	61.2%
50	Office of Federal Contract Compliance	DOL	81	84	3	3.7%

PART RATINGS AND PROGRAM FUNDING LEVELS
ADEQUATE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
51	FAA Air Traffic Services	DOT	6,575	6,748	173	2.6%
52	Nonpoint Source Pollution Control Grants	EPA	204	194	-10	-4.9%
53	CDC Sexually Transmitted Diseases and Tuberculosis	HHS	296	294	-2	-0 7%
54	National Archives and Records Administration Records Services Program	NARA	413	430	17	4.1%
55	Small Business Surety Bonds	SBA	9	11	2	22 2%
56	Interagency Cooperative Administrative Support Services	State	1,269	1,359	90	7.1%
57	President's Emergency Plan For AIDS Relief Global Fund	State	545	299	-246	-45.1%
58	Treasury Technical Assistance	Treasury	60	48	-12	-20.0%
59	Agricultural Commodity Grading and Certification Programs	USDA	195	191	-4	-2.1%
60	Watershed Protection and Flood Prevention	USDA	111	15	-96	-86.5%
61	AmeriCorps State and National Grants	CNCS	268	259	-9	-3 4%
62	Bureau of Indian Affairs - Forestry Management	DOI	52	52	0	0 0%
63	Permanent Labor Certification Program	DOL	38	40	2	5.3%
64	Howard University	ED	237	237	0	0.0%
65	EPA Enforcement of Environmental Laws (Criminal)	EPA	51	51	0	0.0%
66	EPA Existing Chemicals Program	EPA	17	17	0	0.0%
67	CDC Occupational Safety and Health	HHS	255	250	-5	-2.0%
68	Animal Welfare	USDA	17	19	2	11.8%
69	Corps of Engineers: Hydropower	Corps	296	285	-11	-3 7%
70	Immigration and Customs Enforcement Office of Investigations	DHS	1,493	1,619	126	8.4%
71	Preparedness -- Grants and Training Office Training Program	DHS	210	92	-118	-56.2%
72	National Oceanic & Atmospheric Administration Ecosystem Research	DOC	201	167	-34	-16.9%
73	Navy Shipbuilding	DOD	13,778	13,280	-498	-3.6%
74	Bureau of Reclamation Water Management - Operation and Maintenance	DOI	338	346	8	2.4%
75	US Marshals Service - Apprehension of Fugitives	DOJ	242	260	18	7.4%
76	Childhood Immunization Program	HHS	520	509	-11	-2.1%
77	Refugee Social Services	HHS	210	203	-7	-3.3%
78	Rural Health Activities	HHS	160	27	-133	-83 1%
79	Alcohol, Tobacco, and Firearms Consumer Product Safety	Treasury	46	48	2	4.3%
80	Direct Crop Payments	USDA	5,237	5,237	0	0.0%
81	Rural Telecommunications Loan Programs	USDA	690	691	1	0.1%
82	Veterans Medical Care	VA	30,825	34,295	3,470	11.3%
83	Preparedness -- Grants and Training Office Technical Assistance Program	DHS	20	24	4	20.0%
84	National Telecommunications and Information Administration	DOC	18	23	5	27.8%
85	FBI Counterterrorism Program	DOJ	1,245	1,336	91	7.3%
86	Weed and Seed	DOJ	49	49	0	0.0%
87	Occupational Safety and Health Administration	DOL	472	484	12	2.5%
88	Assets for Independence	HHS	24	24	0	0.0%
89	CDC Buildings and Facilities	HHS	158	30	-128	-81.0%
90	Organ Transplantation	HHS	23	23	0	0.0%
91	Drug-Free Communities Support Program	ONDCP	79	79	0	0.0%
92	Forest Service Energy Resource Needs	USDA	91	95	4	4.4%
93	Intermediary Relending Program	USDA	34	34	0	0.0%
94	Defense Health Care	DOD	20,021	21,025	1,004	5 0%
95	Bureau of Indian Affairs - Economic Development Guaranteed Loans	DOI	6	6	0	0 0%
96	Bureau of Land Management - Recreation Management	DOI	65	64	-1	-1 5%
97	Cybercrime	DOJ	234	260	26	11.1%
98	Prison Construction	DOJ	89	-25	-114	-128.1%
99	Drinking Water State Revolving Fund	EPA	837	842	5	0.6%

PART RATINGS AND PROGRAM FUNDING LEVELS
ADEQUATE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
100	EPA Human Health Research	EPA	61	61	0	0.0%
101	EPA's Recycling, Waste Minimization, and Waste Management Program	EPA	180	183	3	1.7%
102	Pesticide Reregistration	EPA	59	55	-4	-6 8%
103	Food Aid for Emergencies and Development (Public Law 480 Title II)	USAID	1,138	1,218	80	7 0%
104	Corps of Engineers Environmental Stewardship	Corps	94	89	-5	-5 3%
105	Federal Emergency Management Agency Disaster Response	DHS	1,307	326	-981	-75.1%
106	Biomass and Biorefinery Systems	DOE	91	120	29	31.9%
107	Vaccine Injury Compensation Program	DOJ	10	10	0	0.0%
108	White Collar Crime	DOJ	708	674	-34	-4.8%
109	American Indian Vocational Rehabilitation Services	ED	0	0	0	0.0%
110	IDEA Special Education Grants to States	ED	10,583	10,683	100	0 9%
111	Pesticide Registration	EPA	44	43	-1	-2.3%
112	Developmental Disabilities Grant Programs	HHS	155	155	0	0 0%
113	Health Care Patient Safety	HHS	84	84	0	0.0%
114	Bioenergy	USDA	60	0	-60	-100.0%
115	Rural Business and Industry Guaranteed Loan Program	USDA	953	1,000	47	4.9%
116	Department of Defense Facilities Sustainment, Restoration, Modernization, and Demolition	DOD	11,366	11,518	152	1.3%
117	Building Technologies	DOE	69	77	8	11 6%
118	Environmental Management	DOE	6,590	5,828	-762	-11 6%
119	Drug Enforcement Administration	DOJ	2,285	2,276	-9	-0.4%
120	EPA Enforcement of Environmental Laws (Civil)	EPA	451	470	19	4.2%
121	Community Mental Health Services Block Grant	HHS	429	429	0	0.0%
122	Contributions to International Fisheries Commissions	State	24	21	-3	-12.5%
123	Internal Revenue Service Taxpayer Service	Treasury	2,179	2,134	-45	-2.1%
124	Pesticide Data/Microbiological Data Programs	USDA	21	15	-6	-28.6%
125	District of Columbia Community Supervision Program	Court Services	130	135	5	3.8%
126	Federal Law Enforcement Training Center	DHS	280	245	-35	-12 5%
127	International Trade Administration Market Access and Compliance	DOC	39	39	0	0 0%
128	US Attorneys	DOJ	1,580	1,637	57	3.6%
129	Trauma-EMS Systems Program	HHS	0	0	0	0.0%
130	President's Emergency Plan For AIDS Relief Other Bilateral Programs	State	935	915	-20	-2.1%
131	Africa Child Survival and Health	USAID	392	479	87	22.2%
132	US Agency for International Development Climate Change Program	USAID	**	**	**	**
133	Work Incentive Grants	DOL	20	0	-20	-100 0%
134	Impact Aid Construction	ED	18	18	0	0.0%
135	Student Aid Administration	ED	744	734	-10	-1.3%
136	Water Pollution Control Grants	EPA	216	222	6	2.8%
137	Chronic Disease - Diabetes	HHS	63	63	0	0.0%
138	Housing Counseling	HUD	42	45	3	7.1%
139	International Disaster and Famine Account	USAID	365		-365	-100 0%
140	National Forest Improvement and Maintenance	USDA	442	385	-57	-12.9%
141	Veterans Pension	VA	3,640	3,718	78	2 1%
142	International Trade Administration U S and Foreign Commercial Service	DOC	229	248	19	8.3%
143	Bureau of Indian Affairs - K-12 School Operations	DOI	535	519	-16	-3.0%
144	Bureau of Indian Affairs - Tribal Colleges	DOI	103	102	-1	-1.0%
145	Employment Service	DOL	827	753	-74	-8.9%
146	Federal Family Education Loans	ED	9,839	6,125	-3,714	-37.7%
147	William D. Ford Direct Student Loans	ED	599	36	-563	-94.0%
148	National Ambient Air Quality Standards and Regional Haze Programs	EPA	97	97	0	0.0%
149	Family Self-Sufficiency Program	HUD	48	48	0	0 0%

PART RATINGS AND PROGRAM FUNDING LEVELS
ADEQUATE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
150	Indian Community Development Block Grant Program	HUD	59	57	-2	-3.4%
151	International Development Association	Treasury	2,820	2,850	30	1.1%
152	Denali Commission	Denali	53	6	-47	-88 7%
153	National Marine Fisheries Service	DOC	687	649	-38	-5.5%
154	Coal Energy Technology	DOE	376	345	-31	-8 2%
155	Bureau of Land Management - Energy and Minerals Management	DOI	134	157	23	17.2%
156	Longshore and Harbor's Workers Compensation Program	DOL	248	248	0	0.0%
157	Mine Safety and Health Administration	DOL	278	288	10	3.6%
158	Javits Fellowships	ED	10	10	0	0.0%
159	Transition to Teaching	ED	44	44	0	0.0%
160	Vocational Rehabilitation State Grants	ED	2,693	2,837	144	5.3%
161	Drinking Water Research	EPA	45	49	4	8.9%
162	Ocean, Coastal, and Estuary Protection	EPA	37	31	-6	-16.2%
163	Toxic Air Pollutants - Regulations and Federal Support	EPA	97	99	2	2 1%
164	NASA Education Program	NASA	162	153	-9	-5 6%
165	Space Shuttle	NASA	4,776	4,057	-719	-15 1%
166	Federal Employees Retirement	OPM	58,501	61,427	2,926	5 0%
167	Homeland Security Operations Center	DHS	51	55	4	7.8%
168	Transportation Security Administration, Screener Training	DHS	88	88	0	0.0%
169	Bureau of Land Management - Realty and Ownership Management	DOI	104	98	-6	-5.8%
170	National Park Service - Concessions Management	DOI	56	58	2	3.6%
171	Office of Labor Management Standards	DOL	46	52	6	13.0%
172	Clean Water State Revolving Fund	EPA	887	688	-199	-22.4%
173	Endocrine Disruptors	EPA	19	17	-2	-10 5%
174	Resource Conservation and Recovery Act Corrective Action	EPA	38	41	3	7.9%
175	Superfund Remedial Action	EPA	601	592	-9	-1.5%
176	U. S.-Mexico Border Water Infrastructure	EPA	49	25	-24	-49.0%
177	Underground Injection Control Grant Program	EPA	11	11	0	0.0%
178	Children's Hospitals Graduate Medical Education Payment Program	HHS	297	99	-198	-66.7%
179	Chronic Disease - Breast and Cervical Cancer	HHS	204	204	0	0.0%
180	Foster Care	HHS	4,633	4,757	124	2.7%
181	Substance Abuse Treatment Programs of Regional and National Significance	HHS	399	375	-24	-6.0%
182	Tribal Housing Activities Loan Guarantees	HUD	2	2	0	0.0%
183	Merit System Compliance	OPM	17	17	0	0.0%
184	Nutrition Assistance for Puerto Rico	USDA	1,518	1,559	41	2 7%
185	Federal Emergency Management Agency Disaster Recovery	DHS	6,466	1,374	-5,092	-78.8%
186	National Park Service - Technical Assistance	DOI	12	12	0	0 0%
187	Workforce Investment Act - Adult Employment and Training	DOL	871	726	-145	-16.6%
188	Comprehensive School Reform	ED	8	0	-8	-100.0%
189	Gaining Early Awareness and Readiness for Undergraduate Programs	ED	303	0	-303	-100.0%
190	EPA Tribal General Assistance Program	EPA	57	57	0	0.0%
191	National Ambient Air Quality Standards Research	EPA	67	66	-1	-1.5%
192	Federal Employees Group Life Insurance	OPM	3,607	3,844	237	6.6%
193	International Narcotics Control and Law Enforcement Programs, Africa/Asia	State	15	31	16	106 7%
194	International Narcotics Control and Law Enforcement Programs, Western Hemisphere	State	61	55	-6	-9.8%
195	Forest Service Land Acquisition	USDA	42	25	-17	-40.5%
196	Chemical Demilitarization	DOD	1,387	1,408	21	1.5%
197	Bureau of Reclamation - Recreation and Concessions	DOI	16	15	-1	-6.3%
198	21st Century Community Learning Centers	ED	981	981	0	0 0%
199	EPA Oil Spill Control	EPA	12	13	1	8.3%

PART RATINGS AND PROGRAM FUNDING LEVELS
ADEQUATE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
200	Andean Counterdrug Initiative	State	735	722	-13	-1 8%
201	International Child Labor Program and Office of Foreign Relations	DOL	73	12	-61	-83.6%
202	Workforce Investment Act - Native American Programs	DOL	56	53	-3	-5 4%
203	National Technical Institute for the Deaf	ED	56	55	-1	-1 8%
204	Brownfields Revitalization	EPA	162	163	1	0.6%
205	Public Water System Supervision Grant Program	EPA	98	99	1	1 0%
206	Adoption Opportunities	HHS	27	27	0	0.0%
207	Federal Employees Health Benefits	OPM	32,126	34,625	2,499	7.8%
208	8(a) Business Development Program	SBA	38	35	-3	-7.9%
209	Financial and Technical Assistance	Treasury	35	0	-35	-100.0%
210	Civilian Radioactive Waste Management Program Yucca Mountain Project	DOE	495	544	49	9.9%
211	Workforce Investment Act - Dislocated Worker Assistance	DOL	1,200	976	-224	-18 7%
212	FAA Facilities and Equipment	DOT	2,555	2,503	-52	-2.0%
213	Education - State Assessment Grants	ED	400	400	0	0.0%
214	Magnet Schools	ED	107	107	0	0 0%
215	Projects with Industry for People with Disabilities	ED	20	0	-20	-100.0%
216	Troops-to-Teachers	ED	15	15	0	0 0%
217	Adoption Incentives	HHS	18	30	12	66.7%
218	Nursing Education Loan Repayment and Scholarship Program	HHS	31	31	0	0 0%
219	New Markets Tax Credit	Treasury	8	8	0	0.0%
	Total funding for Adequate programs [1]		335,470	334,744	-726	-0 2%

[1] *Funding levels represent program level and include BA, obligations, user fees, loan levels, etc as appropriate*

PART RATINGS AND PROGRAM FUNDING LEVELS
INEFFECTIVE PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
1	EPA Pesticide Enforcement Grant Program	EPA	19	19	0	0.0%
2	Oil Technology	DOE	32	0	-32	-100.0%
3	EPA Ecological Research	EPA	85	80	-5	-5.9%
4	Workforce Investment Act - Youth Activities	DOL	951	851	-100	-10.5%
5	Trade Adjustment Assistance	DOL	966	939	-27	-2.8%
6	Internal Revenue Service Earned Income Tax Credit Compliance	Treasury	167	168	1	0.6%
7	Natural Gas Technology	DOE	33	0	-33	-100.0%
8	Substance Abuse Prevention and Treatment Block Grant	HHS	1,759	1,759	0	0.0%
9	TRIO Upward Bound	ED	311	0	-311	-100.0%
10	HOPE VI (Severely Distressed Public Housing)	HUD	99	-99	-198	-200.0%
11	Health Professions	HHS	295	159	-136	-46.1%
12	Gallaudet University	ED	107	108	1	0.9%
13	Air Quality Grants and Permitting	EPA	194	160	-34	-17.5%
14	Workforce Investment Act - Migrant and Seasonal Farmworkers	DOL	81	2	-79	-97.5%
15	Project-Based Rental Assistance	HUD	5,037	5,676	639	12.7%
16	Healthy Community Access Program	HHS	0	0	0	0.0%
17	State Planning Grant Program	HHS	0	0	0	0.0%
18	Rural Housing and Economic Development	HUD	17	0	-17	-100.0%
19	Alaska Native Village Water Infrastructure	EPA	34	15	-19	-55.9%
20	Community Development Block Grant (Formula)	HUD	3,248	2,975	-273	-8.4%
21	AmeriCorps National Civilian Community Corps	CNCS	27	5	-22	-81.5%
22	Even Start	ED	99	0	-99	-100.0%
23	Community Service Employment for Older Americans	DOL	432	388	-44	-10.2%
24	Safe and Drug Free Schools State Grants	ED	347	0	-347	-100.0%
25	Juvenile Accountability Block Grants	DOJ	49	0	-49	-100.0%
26	Amtrak	DOT	1,294	900	-394	-30.4%
27	Vocational Education State Grants	ED	1,182	0	-1,182	-100.0%
28	Federal Perkins Loans	ED	65	0	-65	-100.0%
	Total funding for Ineffective programs [1]		16,930	14,105	-2,825	-16.7%

[1] *Funding levels represent program level and include BA, obligations, user fees, loan levels, etc as appropriate*

PART RATINGS AND PROGRAM FUNDING LEVELS
RESULTS NOT DEMONSTRATED PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
1	Enforcement of Commodity Futures and Options Markets	CFTC	27	27	0	0.0%
2	Credit Union Loan and Technical Assistance Grant Program	NCUA	13	13	0	0.0%
3	Consumer Product Safety Commission	CPSC	62	62	0	0.0%
4	Office of Surface Mining - State Managed Regulation of Surface Coal Mining	DOI	70	72	2	2 9%
5	Office of Surface Mining - State Managed Abandoned Coal Mine Land Reclamation	DOI	152	151	-1	-0.7%
6	Housing Opportunities for Persons with AIDS	HUD	286	300	14	4.9%
7	Energy Information Administration	DOE	86	91	5	5.8%
8	Drug Courts	DOJ	10	69	59	590.0%
9	Flood Damage Reduction	Corps	1,492	1,197	-295	-19.8%
10	Natural Resources Conservation Service National Resources Inventory	USDA	0	0	0	0.0%
11	Border Patrol	DHS	1,847	1,862	15	0 8%
12	Coast Guard Aids to Navigation	DHS	1,095	1,062	-33	-3.0%
13	Marine Corps Expeditionary Warfare	DOD	10,223	9,234	-989	-9.7%
14	Agriculture Marketing Service - Research and Promotion Programs	USDA	3	3	0	0 0%
15	Securities and Exchange Commission - Enforcement	SEC	286	292	6	2.1%
16	Securities and Exchange Commission - Full Disclosure Program (Corporate Review)	SEC	74	75	1	1.4%
17	Coast Guard Search and Rescue	DHS	886	880	-6	-0.7%
18	Public Defender Service for the District of Columbia	District of	30	32	2	6.7%
19	Office of Surface Mining - Federal Managed Regulation of Surface Coal Mining	DOI	67	67	0	0 0%
20	National School Lunch	USDA	7,415	7,667	252	3.4%
21	Fair Housing Initiatives Program	HUD	26	20	-6	-23 1%
22	Forest Service Invasive Species Program	USDA	166	165	-1	-0.6%
23	Economic and Trade Sanctions Program - Office of Foreign Assets Control	Treasury	22	23	1	4.5%
24	Mentoring Children of Prisoners	HHS	50	40	-10	-20.0%
25	Wildlife Habitat Incentives Program	USDA	43	55	12	27 9%
26	African Development Fund	Treasury	134	136	2	1 5%
27	Conservation Technical Assistance	USDA	723	643	-80	-11.1%
28	Fish and Wildlife Service - Migratory Bird Management and Conservation	DOI	124	132	8	6.5%
29	Community Facilities Program	USDA	550	522	-28	-5 1%
30	GSA - National Information Technology Solutions	GSA	6,782	6,401	-381	-5.6%
31	Tropical Forest Conservation Act	Treasury	20	*	*	*
32	Emergency Watershed Protection Program	USDA	300	0	-300	-100.0%
33	Rural Business Enterprise Grant Program	USDA	40	0	-40	-100.0%
34	Montgomery GI Bill- Veterans Education Benefits	VA	1,982	2,041	59	3.0%
35	Inland Waterways Navigation	Corps	934	871	-63	-6 7%
36	Transportation Security Administration Screener Workforce	DHS	2,529	2,618	89	3.5%
37	Minerals Management Service - Minerals Revenue Management	DOI	78	79	1	1.3%
38	Office of Disease Prevention and Health Promotion	HHS	7	7	0	0.0%
39	Dairy Payment Program	USDA	500	50	-450	-90.0%
40	Coast Guard Drug Interdiction	DHS	1,219	1,240	21	1.7%
41	Bureau of Reclamation - Rural Water Supply Projects	DOI	79	66	-13	-16 5%
42	GSA - Regional Information Technology Solutions	GSA	0	0	0	0 0%
43	Defense Communications Infrastructure	DOD	3,820	3,760	-60	-1.6%
44	Department of the Interior - Wildland Fire Management	DOI	755	770	15	2.0%
45	Developing Hispanic-serving Institutions	ED	95	95	0	0.0%
46	Strengthening Historically Black Colleges and Universities	ED	238	238	0	0.0%
47	Strengthening Historically Black Graduate Institutions	ED	58	58	0	0 0%
48	Federal Election Laws - Compliance and Enforcement	FEC	21	23	2	9.5%
49	Asian Development Fund	State	99	115	16	16.2%
50	Food and Nutrition Service - Child and Adult Care Food Program	USDA	2,159	2,274	115	5.3%

PART RATINGS AND PROGRAM FUNDING LEVELS
RESULTS NOT DEMONSTRATED PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
51	Science and Technology Threat and Vulnerability, Testing and Assessment	DHS	40	40	0	0.0%
52	International Trade Administration Import Administration	DOC	61	61	0	0.0%
53	Partnership for Advancing Technology in Housing	HUD	5	0	-5	-100.0%
54	Rural Distance Learning and Telemedicine Loan and Grant Program	USDA	50	25	-25	-50.0%
55	Afghanistan Health Initiative	HHS	6	6	0	0 0%
56	Health Care Fraud and Abuse Control	HHS	320	160	-160	-50 0%
57	Bank Enterprise Award	Treasury	12	0	-12	-100 0%
58	Internal Revenue Service Tax Collection	Treasury	2,106	2,131	25	1.2%
59	Enhancing Education Through Technology	ED	272	0	-272	-100.0%
60	National Archives and Records Administration Electronic Records Services	National	38	45	7	18 4%
61	International Information Programs	State	53	62	9	17.0%
62	Rural Business-Cooperative Service Value-Added Producer Grants	USDA	21	20	-1	-4.8%
63	National Nuclear Infrastructure	DOE	98	97	-1	-1.0%
64	State Energy Programs	DOE	36	49	13	36.1%
65	Fish and Wildlife Service - National Wildlife Refuge System	DOI	386	382	-4	-1.0%
66	Community-Based Child Abuse Prevention	HHS	42	42	0	0.0%
67	Public Diplomacy	State	703	762	59	8.4%
68	Transportation Security Administration Air Cargo Security Programs	DHS	44	40	-4	-9.1%
69	Bureau of Indian Affairs - K-12 School Construction	DOI	207	157	-50	-24.2%
70	Graduate Assistance in Areas of National Need	ED	30	30	0	0.0%
71	Child Abuse Prevention and Treatment Grants	HHS	27	27	0	0.0%
72	United States-Mexico Border Health Commission	HHS	4	4	0	0.0%
73	Global Environment Facility	Treasury	79	56	-23	-29 1%
74	USDA Wildland Fire Management	USDA	1,646	1,676	30	1.8%
75	Coastal Storm Damage Reduction	Corps	113	53	-60	-53.1%
76	Fish and Wildlife Service - Wildlife and Sport Fish Restoration	DOI	717	800	83	11.6%
77	International Education Domestic Programs	ED	92	93	1	1.1%
78	CDC. State and Local Preparedness Grants	HHS	824	822	-2	-0 2%
79	Dairy Price Support Program	USDA	42	21	-21	-50.0%
80	Transportation Security Administration Passenger Screening Technology	DHS	164	173	9	5 5%
81	Coastal Zone Management Act Programs	DOC	109	98	-11	-10.1%
82	Department of Defense Training and Education Programs - Other Training and Education	DOD	1,271	1,291	20	1.6%
83	Preparedness -- Grants and Training Office State Homeland Security Grants	DHS	564	668	104	18 4%
84	Pacific Coastal Salmon Recovery Fund	DOC	67	67	0	0.0%
85	Job Training Apprenticeship	DOL	21	21	0	0.0%
86	Child Care Access Means Parents in School	ED	16	16	0	0.0%
87	Education - Neglected and Delinquent State Agency Program	ED	50	50	0	0.0%
88	GSA - Office of Governmentwide Policy	GSA	62	65	3	4.8%
89	GSA - Travel Management	GSA	14	15	1	7.1%
90	Head Start	HHS	6,876	6,786	-90	-1.3%
91	Office of Minority Health	HHS	57	47	-10	-17.5%
92	Minority Business Development Agency	DOC	30	30	0	0.0%
93	Smaller Learning Communities	ED	94	0	-94	-100.0%
94	EPA Environmental Education	EPA	9	0	-9	-100.0%
95	Domestic HIV/AIDS Prevention	HHS	651	739	88	13 5%
96	National Bioterrorism Hospital Preparedness Program	HHS	474	474	0	0.0%
97	Transportation Security Administration Baggage Screening Technology	DHS	441	442	1	0.2%
98	Bureau of Reclamation - Water Management - Project Planning and Construction	DOI	222	175	-47	-21.2%
99	Community Oriented Policing Services	DOJ	387	-23	-410	-105.9%
100	Federal Work-Study	ED	980	980	0	0.0%

PART RATINGS AND PROGRAM FUNDING LEVELS
RESULTS NOT DEMONSTRATED PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
101	Physical Education Program	ED	73	26	-47	-64.4%
102	Ready to Learn Television	ED	24	24	0	0.0%
103	Mental Health Programs of Regional and National Significance	HHS	263	228	-35	-13.3%
104	Youth Anti-Drug Media Campaign	ONDCP	99	120	21	21.2%
105	Transportation Security Administration Federal Air Marshal Service	DHS	686	699	13	1.9%
106	Bureau of Land Management - Land Use Planning	DOI	50	49	-1	-2.0%
107	Bureau of Land Management - Southern Nevada Land Sales	DOI	738	735	-3	-0.4%
108	Emergency Medical Services for Children	HHS	20	0	-20	-100.0%
109	Family Violence Prevention and Services Program	HHS	129	0	-129	-100.0%
110	Runaway and Homeless Youth	HHS	103	103	0	0.0%
111	Transportation Security Administration Aviation Regulation and Enforcement	DHS	221	218	-3	-1.4%
112	Transportation Security Administration Flight Crew Training	DHS	30	30	0	0.0%
113	DOL - Women's Bureau	DOL	10	9	-1	-10.0%
114	Impact Aid Basic Support Payments and Payments for Children with Disabilities	ED	1,141	1,141	0	0.0%
115	High Intensity Drug Trafficking Areas	ONDCP	225	207	-18	-8.0%
116	Resource Conservation and Development	USDA	51	26	-25	-49.0%
117	Residential Substance Abuse Treatment	DOJ	9	0	-9	-100.0%
118	IDEA Special Education - Research and Innovation	ED	72	72	0	0.0%
119	GSA - Transportation Management	GSA	23	23	0	0.0%
120	Supplemental Educational Opportunity Grants	ED	771	771	0	0.0%
121	Teaching American History	ED	120	50	-70	-58.3%
122	Adolescent Family Life Program	HHS	30	30	0	0.0%
123	Independent Living Program	HHS	140	140	0	0.0%
124	Immigration and Customs Enforcement Automation Modernization Program	DHS	40	0	-40	-100.0%
125	Training and Advisory Services	ED	7	7	0	0.0%
126	Low Income Home Energy Assistance Program	HHS	2,162	2,782	620	28.7%
127	Bureau of Indian Affairs - Housing Improvement	DOI	28	28	0	0.0%
128	Bureau of Indian Affairs - Operation and Maintenance of Irrigation Projects	DOI	38	38	0	0.0%
129	College Assistance Migrant Program	ED	15	15	0	0.0%
130	Office on Women's Health	HHS	28	28	0	0.0%
131	Counterdrug Technology Transfer Program	ONDCP	16	0	-16	-100.0%
132	Adult Education State Grants	ED	564	564	0	0.0%
133	Traumatic Brain Injury	HHS	9	0	-9	-100.0%
134	FHA Multi-Family Mortgage Insurance	HUD	309	316	7	2.3%
135	Counterdrug Research & Development	ONDCP	14	10	-4	-28.6%
136	High School Equivalency Program	ED	19	19	0	0.0%
137	IDEA Special Education - Parent Information Centers	ED	26	26	0	0.0%
138	IDEA Special Education - Technical Assistance and Dissemination	ED	49	49	0	0.0%
139	Teacher Quality Enhancement	ED	60	0	-60	-100.0%
140	Community Services Block Grant	HHS	630	0	-630	-100.0%
141	Veterans Home Loans	VA	7,698	6,634	-1,064	-13.8%
142	Preparedness -- Grants and Training Office Assistance to Firefighters Grant Program	DHS	648	293	-355	-54.8%
143	Bureau of Indian Affairs - Operation and Maintenance of Roads	DOI	29	26	-3	-10.3%
144	Vocational Rehabilitation Demonstration and Training Programs	ED	7	7	0	0.0%
145	Coast Guard Polar Icebreaking Program	DHS	85	89	4	4.7%
146	Independent Living for People with Disabilities	ED	98	98	0	0.0%
147	Pesticide Field Programs	EPA	37	38	1	2.7%
148	Commodity Purchase Services (Section 32)	USDA	1,416	907	-509	-35.9%
149	Delta Regional Authority	Delta Regional	12	6	-6	-50.0%
150	Bureau of Land Management - Mining Law Administration	DOI	33	33	0	0.0%

PART RATINGS AND PROGRAM FUNDING LEVELS
RESULTS NOT DEMONSTRATED PROGRAMS
(dollars in millions)

	Program Name	Agency	2006 Enacted	2007 Budget	Change from 2006	
151	IDEA Special Education Grants for Infants and Families	ED	436	436	0	0.0%
152	IDEA Special Education Personnel Preparation Grants	ED	90	90	0	0.0%
153	The Emergency Food Assistance Program (TEFAP)	USDA	196	190	-6	-3 1%
154	Comprehensive Regional Assistance Centers	ED	56	56	0	0 0%
155	Leveraging Educational Assistance Partnership	ED	65	0	-65	-100.0%
156	Pollution Prevention and New Technologies Research	EPA	26	21	-5	-19.2%
157	Native American Housing Block Grants	HUD	622	624	2	0.3%
158	University Nuclear Education Programs	DOE	27	0	-27	-100.0%
159	Bureau of Indian Affairs - Law Enforcement	DOI	193	202	9	4 7%
160	Prevailing Wage Determination Program	DOL	10	10	0	0.0%
161	Multipurpose Law Enforcement Grants	DOJ	517	0	-517	-100 0%
162	Assistive Technology Alternative Financing Program	ED	4	0	-4	-100.0%
163	Public Housing	HUD	6,003	5,742	-261	-4.3%
164	Commerce Small Business Innovation Research Program	DOC	8	7	-1	-12 5%
165	American Printing House for the Blind	ED	18	18	0	0.0%
166	National Writing Project	ED	22	0	-22	-100 0%
167	Tech-Prep Education State Grants	ED	105	0	-105	-100 0%
168	Tribally Controlled Postsecondary Vocational and Technical Institutions	ED	7	7	0	0.0%
169	Preparedness -- Infrastructure Protection Cyber Security	DHS	146	178	32	21.9%
170	National Park Service - Land and Water Conservation Fund State Grants	DOI	30	2	-28	-93.3%
171	Commodity Supplemental Food Program	USDA	111	0	-111	-100 0%
172	Fish and Wildlife Service - Endangered Species	DOI	330	330	0	0.0%
173	Parental Information and Resource Centers	ED	40	0	-40	-100 0%
174	Social Services Block Grant	HHS	1,700	1,200	-500	-29.4%
175	Housing for Persons with Disabilities	HUD	239	119	-120	-50.2%
176	Housing for the Elderly	HUD	742	546	-196	-26.4%
177	Defense Small Business Innovation Research/Technology Transfer	DOD	1,264	1,282	18	1 4%
178	B.J. Stupak Olympic scholarships	ED	1	0	-1	-100 0%
179	Byrd Honors Scholarships	ED	41	0	-41	-100.0%
180	States Grants for Occupational and Employment Information	ED	0	0	0	0.0%
181	Universal Service Fund High Cost	FCC	3,982	4,367	385	9.7%
182	Impact Aid Payments for Federal Property	ED	64	64	0	0 0%
183	IDEA Special Education Preschool Grants	ED	766	766	0	0.0%
184	Universal Service Fund E-Rate	FCC	2,274	2,405	131	5.8%
185	State Criminal Alien Assistance Program	DOJ	400	0	-400	-100.0%
186	Education State Grants for Innovative Programs	ED	99	99	0	0.0%
187	Veterans Disability Compensation	VA	30,970	35,012	4,042	13.1%
188	Department of the Interior - Land and Water Conservation Fund Land Acquisition	DOI	103	112	9	8 7%
189	National Park Service - Heritage Partnership	DOI	13	7	-6	-46.2%
190	Bureau of Indian Affairs - Tribal Courts	DOI	12	12	0	0.0%
191	Health Care Facilities Construction and Other Miscellaneous Congressional Earmarks	HHS	0	0	0	0.0%
	Total funding for Results Not Demonstrated programs [1]		139,147	135,991	-3,156	-2.3%

[1] *Funding levels represent program level and include BA, obligations, user fees, loan levels, etc as appropriate*

* -- Tropical Forest Conservation Act -- Funding for 2007 will be provided within the amount appropriated for debt relief based on the program's ability to demonstrate results in 2006

Written Testimony Of

Eileen Norcross, M.A.
Senior Research Fellow for the Government Accountability Project
The Mercatus Center at George Mason University

Before the
Subcommittee on Federal Financial Management, Government Information
And International Security of the Senate Subcommittee on Homeland
Security and Governmental Affairs

June 13, 2006

Mr. Chairman and Members of the Subcommittee:

Thank you Chairman Coburn, Senator Carper, and Members of the Committee on Homeland Security and Governmental Affairs for inviting me to testify today on "Autopilot Budgeting: Will Congress Ever Respond to Government Performance Data?"

Our work in the Government Accountability Project at the Mercatus Center at George Mason University focuses closely on the use of performance information in government. I note that the views expressed in my testimony are not an official position of the University.

Before beginning I would like to submit for the record our paper on the results of the FY07 PART for your reference.

I. The importance of performance information

According to the Office of Management and Budget the federal government has created over 1000 programs to address a range of policy issues from alleviating poverty to curing dreaded diseases and protecting the nation from attack. Some programs have been in existence for decades and spend billions of dollars towards achieving their goals.

In many cases we do not know if they are working.

A program is a tool to achieve a policy goal. Do economic development programs result in prosperous communities? Do job training programs lead to increased employment? Are homeland security programs protecting the nation from attack?

Unless Congress knows the answers to these questions, it cannot make informed decisions about how to spend resources. More importantly, Congress cannot accomplish its policy aims.

Without information on program performance, agencies cannot meet their missions and goals. The public is left in the dark about whether the government is solving the problems Congress has identified as important.

Not knowing has several consequences:

a) Program duplication

Over the years Congress has created hundreds of programs addressing a single outcome. There are anywhere between 180 and 342 programs dealing with economic development in over 24 agencies. There are 44

job training programs in nine agencies; 130 programs serving at-risk youth, and 72 safe water programs, to name a few. Program duplication on this scale implies that Congress isn't sure which programs are reaching their goals. It has no way of comparing programs with common outcomes.

b) Fewer Public Benefits

Not knowing whether a program is performing means possibly not reaching those who are supposed to benefit. The real cost of an under-performing job training program is not merely the amount of money spent on the program; it is the lost opportunity to spend those funds on programs that are working. It is the number of people left unemployed by ineffectively spent dollars.

c) A barrier to the agency

When programs are not required to produce performance information they cannot know if their activities meet the program's ultimate policy objective. Not evaluating programs on a regular basis means that the program's statute may be preventing the program from properly targeting grantees, or delivering results. Performance information permits dialog between Congress and agencies on common grounds and a common understanding of joint objectives.

II. The state of performance information: PART and GPRA

Congress took the initiative in getting agencies to develop performance measures in 1993 when it passed the Government Performance and Results Act (GPRA) (P.L.103-62). GPRA requires that agencies produce three types of reports: strategic plans, annual performance plans, and annual reports on program performance.

The annual report is supposed to give the American people accurate, timely information and let them assess the extent to which agencies are producing tangible public benefits. GPRA has encouraged the development of performance measures and data. But it was not until the development of the Bush Administration's Program Assessment Rating Tool (PART), that real progress towards developing measures was made. Programs are now creating outcome measures because PART is holding them accountable for showing results.

Requiring the information is the first step. It helps agencies articulate goals. It identifies weaknesses in the statute or management of the program. It informs the Executive in making budget recommendations. Unless Congress uses performance information, attempts at holding programs accountable for results are merely a paper exercise.

Has PART played a role in the President's proposed budget or in Congress?
For two years in a row the president has issued a "Major Savings and Reform Report" to accompany the proposed budget.

For FY 2006, the president recommended that 154 programs be either terminated or reduced. Congress accepted 89 of the proposals partially, or in full, for a total savings of $6.5 billion.

Of the 99 programs recommended for termination, Congress agreed to terminate 24 of them; and reduce funding for 28.

Of the 55 programs proposed for reduction, Congress reduced funding for 37 of them.

Of the 154 programs, 54 were PARTed. The president's document indicates where PART played a role in the recommendation. Other factors were also taken into account: lack of a federal role, obsolescence, completion of mission, duplication with public or private efforts, policy priorities, and earmarking. The administration uses PART in conjunction with other information and does not limit itself to the evaluations. It does not automatically cancel programs with poor ratings; nor does it automatically reward satisfactory ones.

By contrast, Congress issued its own report on which programs it terminated in FY 2006. The "On Time and Under Budget" report from the House Committee on Appropriations lists 53 programs that were terminated. We could only identify three programs that were PARTed: Tech-Prep Education State Grants, Occupational and Employment Information, and Community Oriented Policing Services (COPS). We do not know if PART played a role in Congress's decisions to terminate these three programs.

The House Appropriation Committee report only offers explanations for three of its terminations: Jobs-in-the-Woods (obsolete or completed mission), National Youth Sports (activity performed by private sector), and U.S. Capitol Mounted Police (no clear benefit or need.)

We do not know why the remaining 50 were terminated. Were they underperformers, or politically easy choices?

It is useful to compare the two reports. In the Executive's we are given a rationale for each recommendation. The House report merely provides a list. Ultimately, the goal is not to randomly kill programs. The process of making judgments about how to fund activities should be constructive and based on solid evidence, not destructive. If programs managers do not know why their program had its funding reduced, then no one has learned anything.

Performance information is not about how to kill programs. It is about how to make them effective. We have a stake in knowing what works and what doesn't, and why. It is about delivering public benefits in a transparent manner, and ensuring that agencies know to what standards and expectations they are performing.

The only way to give Congress' budgetary decisions credibility is to base them, in part, or in full, on a reliable evaluation of their performance. Congress should use performance information in conjunction with other criteria: e.g., is there a federal role for the activity? Has the program completed its mission or become obsolete? Is this activity a national

priority? Does this program do the best job of addressing a particular problem versus similar programs across the government?

Is PART that system?

PART's methodology has been validly criticized. Assigning quantitative scores to groups of question and aggregating the percentages into a single qualitative score may not fully reflect the program's performance. To illustrate, The Screener Training program in the Department of Homeland Security received a rating of adequate. They received a 100% in both the purpose and design category and the planning category, and an 86% in the management category but only a 13% in the results and accountability category. An adequate rating on its face may indicate to the reader that this program is satisfactorily meeting the objective of training airport screeners. However, according to the results section, this program has not acquired sufficient information to evaluate its performance.

Improvements can and should be made to the methodology. What is most important about PART, however is that it asks Management 101 questions of agency activity.

For example:

- "Is the program purpose clear?"
- "Does the program address a specific and existing problem, interest, or need?"
- "Is the program designed so it is not redundant or duplicative of any Federal, state, local, or private effort?"

- "Is the program design free of major design flaws that would limit the program's effectiveness or efficiency?
- "Is the program effectively targeted so program resources reach the intended beneficiaries and/or otherwise address the program's purpose directly?"
- Has it demonstrated adequate progress towards its long-term outcome performance goals?

This kind of logical process of questioning agency activity needs to be continued. Congress should be asking these kinds of questions before allocating resources.

The questions are the substance of PART. The ratings are based on PART's methodology of quantifying the answers to these questions. Improvements can and should be made to the methodology, but we should not disregard the contribution PART is making to getting agencies to critically examine their activities through the lens of outcomes.

PART's virtues are:

1) It has identified and cataloged agency activities, giving us a common unit of analysis.
2) It is transparent and accessible to the public.
3) It is systematically conducted. PART holds all programs accountable to the same standards.
4) By asking Management 101 questions of program performance, PART focuses agencies on measuring outcomes, not outputs.

5) There is one notable strength of PART that is frequently cited as a weakness. PART rates programs on statutory limitations. This is a source of frustration for agencies that are bound to follow statute, even when the statute may be working against the ultimate goals of the program. Here PART has provided a service. It is identifying those aspects of programs that are barriers to success. Congress should take up the work of reviewing program statutes and continually ask if they are designed to achieve the intended aims of the program.

Some limitations and areas for improvement:

1) PART currently rates programs against their own historical performance. It has not advanced to the stage of being able to compare like activities. Though it is attempting to do so through cross-cutting analyses and by asking if the program's objectives are being addressed elsewhere.
2) As we see with the Screener Training program, in some cases, the measures don't fully capture program performance.
3) Different budget examiners may reach different conclusions viewing the same set of data.

Congress is not bound to use PART ratings. But by ignoring what PART is trying to advance, Congress is missing an opportunity to meet the goals of GPRA. The administration can only take the development and usage of performance information so far. Congress has a responsibility to both agencies and the public to provide clear justification for its budgetary decisions. A systematic evaluation of their performance gives credibility and reliability to the budget process.

III. Congress's Opportunity: Advancing GPRA

Recent legislative proposals indicate that Congress is aware of the importance of establishing a system for evaluating programs.

Learning from PART, Congress has an opportunity to implement its own process to systematically review programs, based upon a logical process of questioning programs and holding them accountable for outcomes.

In addition to requiring and paying attention to agency performance evaluations Congress needs to consider the following:

1) Where the statute is a barrier to performance, Congress must work to update and change the statute so programs are able to meet their objectives.

2) Congress should articulate clear expectations of programs in the statute, including specific outcome-based measures of progress.

3) Outcome Based Scrutiny: Congress should be able to compare like programs that serve the same policy goal and ask which are producing results, and which aren't.

There are a few proposals in Congress to codify the systematic review of programs and get Congress to use performance information.

The Federal Agency Performance Review and Sunset Act introduced by Representative Kevin Brady (R-TX) would allow the president to give a list of programs to a Congressionally established Sunset Commission to review. This Commission then gives its findings to the president with its recommendations. And the president replies with his comments on those findings.

Sunset Commissions give Congress the task of reviewing agency data, eliminating concerns of Executive political influence.

One area for improvement is that Sunset Commissions should consider evaluating government activity according to common outcomes across agencies, rather than reviewing the activities of discrete agencies.

For example: if the policy is to alleviate urban blight. What programs or tools across the federal government currently exist? Which ones work best at addressing the problem and achieving results? Where might we move resources towards investing in programs that are successful in eliminating blight?

A second piece of legislation being considered is offered by Representative Todd Platts (R-PA), The Program Assessment and Results Act (H.R. 185). This bill would rely on OMB to conduct assessments of agencies programs at least once every five fiscal years. The legislation would also ensure that review criteria take into account programs performing similar functions. The results of the assessments are to be submitted to Congress.

This legislation identifies the need for comparing like activities and would advance PART as it currently exists.

We believe both of these pieces of legislation are steps in the right direction each with positive aspects.

We do not believe Congress need adopt OMB's PART wholesale. We hope improvements can be made to the methodology. Or alternatively, that Congress might consider using the questions PART asks as the basis for developing its own method of reviewing government performance and use the data to help inform its decisions. It is not beyond Congress' reach to create and administer such a system, building upon the kinds of questions PART is asking.

Indiscriminant cancellation of programs discredits the process. We leave program managers confused about why their program failed. Programs need to deliver according to clear expectations. When they do not meet them, reduction in funding or termination should be the result. It should not be a surprise. They should be given the chance to prove their effectiveness. And we must also recognize that performance information is best used in conjunction with other criteria: lack of a federal role, low-performance, duplicative, completion of mission. All of these form the basis against which Congress should continually scrutinize agency activity. Efforts to advance what PART has set in motion can only aid Congress in the budget process and give confidence to the American people that the problems our nation has identified are being solved.

A WORKING PAPER IN GOVERNMENT ACCOUNTABILITY

An Analysis of the Office of Management and Budget's Program Assessment Rating Tool (PART) for Fiscal Year 2007

by
Eileen Norcross and Kyle McKenzie[1]

May 2006

MERCATUS CENTER
GEORGE MASON UNIVERSITY

[1] Prepared by Eileen C. Norcross, senior research fellow, and Kyle McKenzie, research fellow, with the Government Accountability Project at the Mercatus Center at George Mason University. This paper is one in a series of working papers from the Mercatus's Government Accountability Project and does not represent an official position of George Mason University.

An Analysis of the Office of Management and Budget's Program Assessment Rating Tool (PART) for Fiscal Year 2007

Executive Summary

With the release of the Bush Administration's proposed budget for FY 2007, the Office of Management and Budget (OMB) has completed its fourth year of the Program Assessment Rating Tool (PART) for evaluating federal programs. Designed as a means of encouraging agencies to develop performance measures and data in order to show program results, PART is used, in conjunction with other information, to make recommendations in the president's budget as well as to inform Congress about agency progress towards goals.

This paper analyzes results of the PART to date and seeks to determine how agencies have fared over time according to PART's methodology. To this end, we examine, among other things, the proportion of agency budgets PARTed as results not demonstrated, or lacking in performance measures or data. We also consider how PART ratings are related to Congressional funding levels and the executive's funding recommendations.

According to OMB, the improvement of PART scores over time shows that many programs are improving in their ability to meet their goals offering relevant data and establishing measures to facilitate OMB's PART evaluation. The number of programs rated effective has risen from 6% in FY 2004, the first year of PART, to 16%. Overall, the number of programs moving from results not demonstrated (that is, not providing enough information to be evaluated), has gone from 50% in FY2004 to 24% in FY2007.

Those rated ineffective remain relatively steady at 4%. Some agencies have a larger proportion of their funding associated with ineffective scores. In particular, 22% of the Department of Housing and Urban Development's (HUD) funding is rated ineffective. Much of this is due to the fact that OMB rated two of HUD's largest programs—the Community Development Block Grant program ($4.1 billion), and Project Based Rental Assistance ($4.95 billion)—as ineffective.

To date, OMB has PARTed 64% percent of the budget, or $1.47 trillion. Six percent of the FY 2005 funding level for PARTed programs representing $143 billion falls into the results not demonstrated rating category.

Last year, the president issued a *Major Savings and Reform* report in which he recommended 154 programs for termination or reduction. The administration used PART, in some cases, to inform these decisions. Congress accepted 89 of these proposals at least partially, reducing spending by $6.5 billion.

This year, the president has again issued a *Major Savings and Reform* report, in which he is recommending 141 programs for either termination or reduction, representing $15

billion in spending. Like last year, the administration cited PART assessments as informing some of these decisions.

A new break-down included in this year's PART assessments isolates programs by "topic" or programmatic activity. According to this categorization, 47% of programs with an education focus are unable to show results, while 33% of foreign affairs programs are rated effective. The purpose of this new category is to facilitate comparison of similar activities across agencies. As last year, OMB applies PART data along with other information to perform crosscutting analyses of research and development programs, federal investment programs, credit and insurance and programs that provide aid to state and local governments.

Background

In February 2003, the Bush administration released with its proposed FY 2004 budget, a new method for evaluating the performance of federal programs called the Program Assessment Rating Tool (PART). PART represents the Bush administration's effort to get agencies to report consistently on their programmatic goals and results in order to improve performance and facilitate funding decisions. It is one of the five initiatives of the President's Management Agenda.

PART is an element of the Administration's Budget and Performance Integration initiative to link performance information to budgeting decisions, also known as "performance budgeting". A performance budget is "an integrated annual performance plan and annual budget that shows the relationship between funding levels and expected results. It indicates that a goal or set of goals should be achieved at a given level of spending." [2] The effort to get agencies to link budgets and performance information originated in 1994 with Congress' passage of the Government Performance and Results Act (GPRA).

I. PART's Methodology and Application

PART requires that agencies submit an assessment of their programmatic performance to OMB over a six year period. To date, OMB has rated 793 of roughly 1000 federal programs it has identified. By FY 2008, OMB will have assessed all identified programs at least once.

OMB bases PART ratings on program manager responses to a series of between 25 and 30 *Yes/No* questions. The questionnaire includes four sections—each weighted differently—dealing with an aspect of program performance: purpose and design (20%), strategic planning (10%), program management (20%) and results/accountability (50%). The individual assessments for each program are provided on OMB's interactive website, ExpectMore.gov. [3]

The results/accountability section (section four) of PART receives the greatest weight. This section's questions are designed to determine if the program has met or achieved efficiencies in its long-term performance goals and how the program compares with similar programs. It also asks if the program has been independently evaluated, and if so, what those evaluations determined. Section four also includes the program's relevant performance measures and data with suggestions for improvement.

A program may receive one of five ratings: ineffective, adequate, moderately effective, effective, and results not demonstrated. The latter rating means that a program does not have enough information (either measures or data) to be rated— not that the program is

[2] John Mercer, *Performance Based Budgeting for Federal Agencies*, AMS, Fairfax, 2002, p.2

[3] For a more detailed description of the assessment process see OMB's website, http://www.whitehouse .gov/omb/part/index.html.

ineffective. It is important to note that a program could receive an acceptable rating even if the results information suggests the program is ineffective. This is because only 50% of final rating depends on results information.

Though regarded as valuable management tool, some believe that PART's rating of programs based on statutory language is unfair and does not take into consideration that programs are bound to operate according to the statute as designed by Congress. Representative Todd Platts (R-PA) has introduced legislation, the Program Assessment Rating Act (H.R. 185), to require that a future program rating tool incorporate congressional intent[4]—something PART does not do. Currently, PART does not take into consideration that a program's authorizing statute may create barriers in achieving the program's intended outcomes. OMB argues this is intentional and is a means of encouraging agencies to consult with Congress on statutory language that may be impeding the agency's or the program's mission.

Other criticisms include the claim that PART is not consistently administered and that its results are too subjective. Assigning a numerical score is potentially inaccurate. Different budget examiners may rate a program differently when presented with the same set of information.

OMB has applied PART data (in conjunction with other information) to undertake crosscutting analyses of aspects of federal programmatic activity. These ongoing analyses compare programs across agencies on the basis of similar outcomes, or approaches to policy problems, with the intent of highlighting best practices, eliminating duplication, or improving coordination across agencies. These analyses include crosscuts of research and development programs, federal investment programs, credit and insurance, and aid to state and local governments. Last year, OMB applied PART data along with other information to analyze the performance of community and economic development programs across agencies. This produced the policy recommendation called the Strengthening America's Communities Initiative, and the suggestion that 18 similar programs be consolidated under one umbrella in the Commerce Department. The initiative was rejected by Congress.

Though PART scores and their application to budget decisions and policy remains the subject of debate in Congress and agencies, PART appears to have increased Congressional interest in evaluating programmatic activity for results, improving reliable performance information, and advancing the goals of GPRA.

Recent legislative efforts to codify the concept of an annual measurement of program performance (not the PART itself) include the Government Reorganization and Program Performance Improvement Act of 2005 sponsored by Representative Kevin Brady (R-TX).[5] The Act, which may come up for vote in the House during June 2006, would create sunset commissions to periodically review and phase out government programs that are obsolete, dysfunctional, duplicative, or unable to meet their goals.

---

[4] "OMB program assessments viewed as flawed budget tool" by Jenny Mandel, *Govexec.com*, April 4, 2006.
[5] A Senate version has also been introduced.

On May 25, 2006, Representative John Tanner (D-TN) introduced legislation, House Resolution 841, to hold Congress accountable for how it spends tax dollars. Provisions include requiring Congress to hold at least two hearings a year on performance reviews produced through PART.

Related to increased interest in the performance of federal dollars, the Federal Funding Accountability and Transparency Act (S. 2590) introduced by Senators Tom Coburn (R-OK) and Barack Obama (D-Ill.) in April 2006, would establish a public database to track the usage of federal grants.

II. Study Purpose and Previous Analysis

This study is an annual update of an analysis we undertook last year in order to examine the progression of PART scores over time, to classify the percentage of the federal budget represented by particular program ratings, and to explore the relationship between PART scores and appropriations. This study does not consider whether PART is *affecting* agency or legislative behavior and funding decisions. Rather, it describes correlations and trends in PART scores.

For the purpose of this analysis, we take PART ratings at face value. But that does not mean we necessarily agree with the methodology used or the conclusions arrived at in the individual assessments.

Many of the questions PART asks of agencies are valuable by themselves in that they focus program managers on their core missions and accomplishments, and areas that need improvement. However, assigning quantitative scores to groups of questions and then aggregating the percentages into a single qualitative score may not fully reflect the program's performance. For example, a program may receive a perfect score in three categories: purpose and design, strategic planning, and management, but fail in results and accountability, and still manage to receive a satisfactory rating. To illustrate, the Screener Training program in the Department of Homeland Security, received a rating of adequate. They received 100% in both the purpose category and the planning category, an 86% in the management category but only a 13% in the results and accountability category. An adequate rating on its face may indicate to the casual reader that this program is adequately meeting the objective of training airport screeners. However, according to the results section, this program, which is relatively new, has not acquired sufficient information in order to gauge its effectiveness. The PART assessment points to a GAO evaluation that shows the program has improved.

Criticisms of PART should not preclude us from studying it more closely. PART provides the first attempt to identify, measure, and aggregate performance data across agencies. PART is the start of a potentially valuable data source for decision makers seeking to understand the effects of individual programs, agency performance in given

policy areas, as well as possibly providing a window for the public into budgetary decision making.

Just as last year, the president's proposed budget for FY 2007 also includes a *Major Savings and Reforms* report. This supplement to the budget uses PART scores, in addition to other information, to make termination and funding decisions. We also analyze this document to find descriptive evidence of how the administration used PART in the FY 2007 proposed budget. This does not imply an endorsement or criticism of how PART was applied in making these decisions. We have updated last year's analysis by examining what Congress did in response to the president's request to terminate or reduce funding for 154 programs. Additionally, we include the programs that Congress terminated independent of the president's recommendations.[6]

We also examine the *Analytical Perspectives* of the FY2007 budget[7] in order to see how OMB is applying PART data in making its recommendations to agencies and policymakers.

1. How PART has rated programs cumulatively.

Table 1. Cumulative program results by ratings category

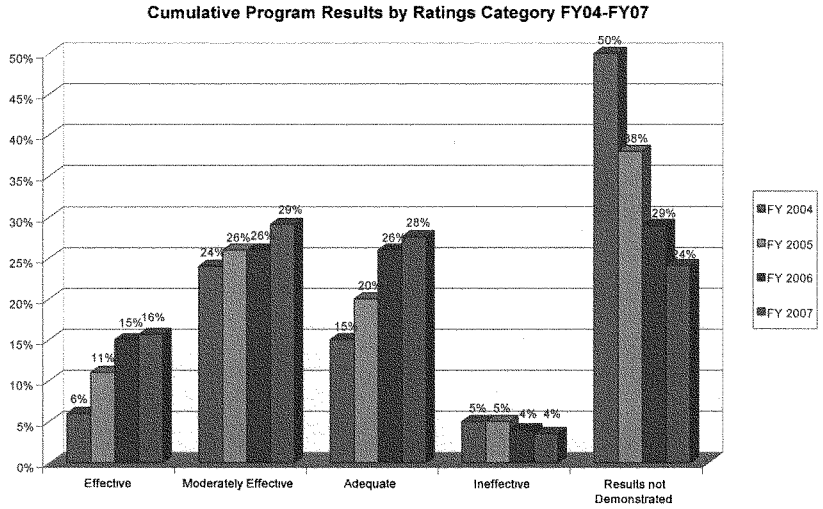
Cumulative Program Results FY 2004–FY 2007[8]				
	FY 2004	FY 2005	FY 2006	FY 2007
Effective	6%	11%	15%	16%
Moderately Effective	24%	26%	26%	29%
Adequate	15%	20%	26%	28%
Ineffective	5%	5%	4%	4%
Results not Demonstrated	50%	38%	29%	24%
Total	**234**	**395**	**607**	**793**

[6] United States House of Representatives Committee on Appropriations, "*On Time and Under Budget*" http://appropriations.house.gov/_files/OnTimeUnderBudget.pdf.
[7] See, http://www.whitehouse.gov/omb/budget/fy2007/pdf/spec.pdf
[8] In this paper we refer to the fiscal year of the budget in which the PART assessments appeared. That is, programs evaluated in 2005 appear in the president's FY 2007 budget proposal. This avoids confusion when trying to locate the PART assessments for a given year.

With each passing year of PART, there has been a steady decrease in the number of programs OMB has rated results not demonstrated. One in seven programs has improved its PART scores.[9]

The cumulative number of programs rated effective, moderately effective, and adequate has increased, while the number of programs rated ineffective remains the same as last year at 4%. OMB rated 16% of programs as effective and 28% as adequate. The later rating represents a 2% increase. The most significant change occurred for the number of moderately effective programs which increased from 26% to 29% and for results not demonstrated programs which dropped from 29% to 24% from last year. The improvement in cumulative program results may be due to a few factors: a) programs are improving their results information, b) evaluations by OMB are getting more, or less, accurate, c) OMB happens to be evaluating better-performing programs or, d) agencies are developing better performance measures.

Chart 1. Cumulative program results by ratings category



[9] See, *Analytical Perspectives of the U.S. Budget, FY2007*, p. 15.

2. Are there observable changes in program performance between FY 2004 and FY 2007 for reassessed programs?

OMB has reassessed 151 programs of the 793 programs it has assessed to date. Of these, 132 have been rated twice, 18 have been rated three times and one program—Missile Defense—has been rated four times.

Table 2. Ratings for reassessed programs

	Initial PART Rating	Most Recent PART Rating
RND	100	8
Ineffective	2	5
Adequate	17	59
Moderately Effective	29	49
Effective	3	30

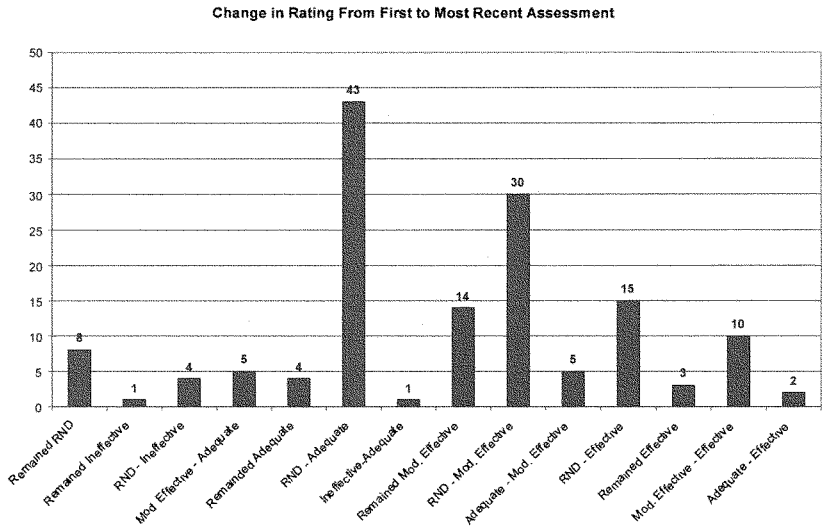
As last year, the greatest improvement among programs that have been evaluated more than once occurred in programs initially rated results not demonstrated. Of the100 programs initially receiving this rating, only eight retained their results not demonstrated upon their most recent reassessment. The number of reassessed programs rated effective increased significantly from three to 30. Of these 30 programs, 15 were initially rated results not demonstrated. Another significant change occurred for programs rated adequate. Initially 17 programs received this rating, upon reassessment, 59 were rated adequate. Improvements were also evident in the moderately effective category as its ranks increased from 29 to 49 programs.

Of the 151 programs reassessed to date, two were initially rated ineffective; OMB has since upgraded one of these to adequate. For all reassessed programs, five are currently rated ineffective; four of these moved out of the results not demonstrated category.

3. How did programs move within ratings categories?

The chart below shows how programs moved from their initial rating to their most recent. That is, of the 100 programs initially rated results not demonstrated, what is their current rating? Forty-three programs have moved from results not demonstrated to adequate; 15 have moved to effective; four are now rated ineffective; eight remain results not demonstrated; and 30 are now rated moderately effective. Only one program has remained ineffective—the Department of Energy's Oil Exploration and Production program—while four programs have moved from results not demonstrated to ineffective.

Chart 3. How reassessed programs moved within ratings categories from first to most recent assessment



4. Programs rated by program type/category

PART classifies programs according to seven categories:

1) Block/Formula Grants – Programs that provide funds to state, local, and tribal governments and other entities by formula block grant.
2) Capital Acquisition – Programs that achieve their goals through development and acquisition of capital assets (such as land, structures, equipment, and intellectual property) or the purchase of services (such as maintenance, and information technology).
3) Competitive Grants – Programs that provide funds to state, local and tribal governments, organizations, individuals and other entities through a competitive process.
4) Credit – Programs that provide support through loans, loan guarantees, and direct credit.
5) Direct Federal – Programs where services are provided primarily by employees of the federal government.
6) Regulatory Based – Programs that accomplish their mission through rulemaking that implements, interprets, or prescribes law or policy, or describes procedure or practice requirements.

7) Research and Development – Programs that focus on knowledge creation or its application to the creation of systems, methods, materials, or technologies.

Mixed programs are those that combine elements from two or more categories (e.g., a research and development program that uses grants as a means of funding research).

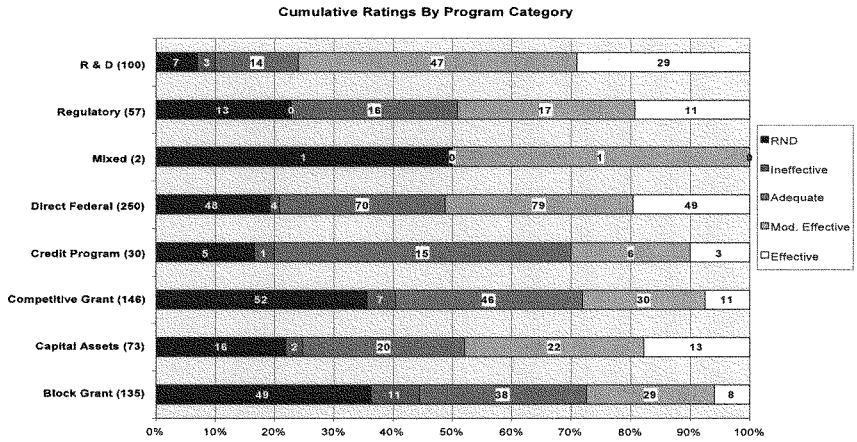
Examining PART data for FY 2004 through FY 2007 reveals that certain categories of programs fare better than others in the ratings

Table 4. Most recent PART ratings by program category

	RND	Ineffective	Adequate	Mod. Effective	Effective
Block Grant (135)	49 (36%)	11 (8%)	38 (28%)	29 (21%)	8 (6%)
Capital Assets (73)	16 (22%)	2 (3%)	20 (27%)	22 (30%)	13 (18%)
Competitive Grant (146)	52 (36%)	7 (5%)	46 (32%)	30 (21%)	11 (8%)
Credit Program (30)	5 (17%)	1 (3%)	15 (50%)	6 (20%)	3 (10%)
Direct Federal (250)	48 (19%)	4 (2%)	70 (28%)	79 (32%)	49 (20%)
Mixed (2)	1 (50%)	0 0%	0 0%	1 (50%)	0 0%
Regulatory (57)	13 (23%)	0 (%)	16 (28%)	17 (30%)	11 (19%)
R & D (100)	7 (7%)	3 (3%)	14 (14%)	47 (47%)	29 (29%)

Excluding mixed programs, which account for only two programs of the 793 PARTed, both block grant and competitive grant programs continue to have the largest percentage of programs rated results not demonstrated—36% each. And as was the case last year, both of these program types continue to have the largest percentage of programs rated ineffective, 8% and 5% respectively.

Direct federal and research and development programs by contrast have the greatest percentage of programs rated effective, 20% and 29% respectively. Regulatory programs at 19% and capital asset programs at 18% are not far behind.



Crosscutting analysis for credit programs and block grant programs

Credit programs

The ratings for program categories raise the question of why certain kinds of programs seem to operate more effectively than others. Included among OMB's crosscutting analyses are credit programs. OMB's analysis includes a detailed look at how credit programs perform within each of the four ratings areas (program purpose and design, strategic planning, management, and results.) Their analysis indicates that credit programs receive high scores for program purpose and design 77% on average although this is slightly lower than the average for all programs, 86%. Credit programs score low in program results (53%), yet compared to the average score for all programs, 47%, this is relatively high.

In terms of program purpose and design, OMB finds that though many of these programs have clear purposes, they are often duplicative of other programs or private sources, and have poor incentive structures, limiting their effectiveness, "For example, private lenders are generally better at screening borrowers, but they may not screen borrowers effectively

if the Government provides a 100% loan guarantee."[10] Thus, OMB suggests that these programs work more closely with private lending institutions.

In the area of strategic planning, OMB states that credit programs have good short-term measures, but are lacking in longer term metrics, such as linking their budgets to outcomes, and performing stringent performance evaluations.

OMB notes that in terms of program management, credit programs are strong in terms of basic finance and accounting practices, yet should incorporate more measures of risk analysis.

And in the most heavily weighted category, program results, OMB states that credit programs are weak, despite their higher than average score. Reasons for this include the difficulty of measuring the net outcome of the program, that is, what would have happened in the absence of the program? In addition, credit programs must also accurately estimate cost. OMB notes that the complexities and dynamic nature of financial markets make credit programs difficult to measure. As private entities reach more underserved populations, government credit programs may have decreased results. Conversely, if financial markets are in turmoil, government credit programs may become more effective. "A sub-par review could be related to financial market developments; the program might have failed to adapt to rapid changes in financial markets; or its function might have become obsolete due to financial evolution." [11]

Grant programs

Programs that provide grants to states and localities are also the subject of a crosscutting analysis in this year's budget. These 211 programs are a subset of block grant, and competitive grant programs, representing $209.8 billion in spending in 2005. Of these 211 programs, 41% are rated results not demonstrated, higher than the average for all programs (31%). OMB states that this is because grant programs have a broad purpose, and a general "lack of agreement among grantees and federal parties on the purpose and performance measures, and therefore lack of focused planning to achieve common goals."[12]

This marks the second year the OMB has been scrutinized block grant programs. OMB notes block grants are one of the most common tools used by the federal government, providing social service funding to states and localities. They are generally regarded as 'flexible' in that local grantees may determine how best to use the funds. However, OMB states that "accountability for results can be difficult when funds are allocated based on formulas and population rather than achievement or needs." Additionally, block grants pose performance management challenges, reflected in the high number of ineffective programs among block grants, 8%.

---

[10] See *Analytical Perspectives of the U.S. Budget, FY 2007* p. 68
[11] Op.cit. pp. 68-69
[12] *Analytical Perspectives of the U.S. Budget, FY 2007*, p.105

OMB notes that it intends to continue monitoring block grant programs to highlight best practices, sharing successful methods with low-performing programs.

5. PART Ratings by program topic

This year OMB budget examiners assigned a "topic" to PARTed programs during their evaluation based on the majority of the program's activities, based on a sub-category of the federal budget codes. This designation may be useful since it allows cross-agency analysis of programs based on common outcomes.

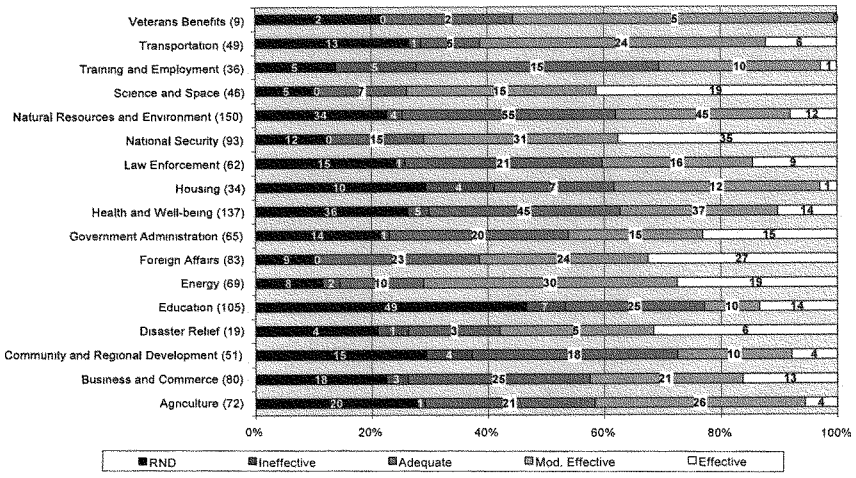
Table 5. Programs rated by topic

	RND	Ineffective	Adequate	Mod. Effective	Effective
Agriculture (72)	20 (28%)	1 (1%)	21 (29%)	26 (36%)	4 (6%)
Business and Commerce (80)	18 (23%)	3 (4%)	25 (31%)	21 (26%)	13 (16%)
Community & Regional Development (51)	15 (29%)	4 (8%)	18 (35%)	10 (20%)	4 (8%)
Disaster Relief (19)	4 (21%)	1 (5%)	3 (16%)	5 (26%)	6 (32%)
Education (105)	49 (47%)	7 (7%)	25 (24%)	10 (10%)	14 (13%)
Energy (69)	8 (12%)	2 (3%)	10 (14%)	30 (43%)	19 (28%)
Foreign Affairs (83)	9 (11%)	0 (0%)	23 (28%)	24 (29%)	27 (33%)
Government Administration (65)	14 (22%)	1 (2%)	20 (31%)	15 (23%)	15 (23%)
Health and Well-being (137)	36 (26%)	5 (4%)	45 (33%)	37 (27%)	14 (10%)
Housing (34)	10 (29%)	4 (12%)	7 (21%)	12 (35%)	1 (3%)
Law Enforcement (62)	15 (24%)	1 (2%)	21 (34%)	16 (26%)	9 (15%)
National Security (93)	12 (13%)	0 (0%)	15 (16%)	31 (33%)	35 (38%)
Natural Resources and Environment (150)	34 (23%)	4 (3%)	55 (37%)	45 (30%)	12 (8%)
Science and Space (46)	5 (11%)	0 (0%)	7 (15%)	15 (33%)	19 (41%)
Training and Employment (36)	5 (14%)	5 (14%)	15 (42%)	10 (28%)	1 (3%)
Transportation (49)	13 (27%)	1 (2%)	5 (10%)	24 (49%)	6 (12%)
Veterans Benefits (9)	2 (22%)	0 (0%)	2 (22%)	5 (56%)	0 (0%)

Assessing PART ratings according to topic shows that certain programmatic areas, across agencies, are getting better ratings than others. Nearly half, or 50, education programs are rated results not demonstrated. While more than a quarter, or 27 of 83 foreign affairs programs are rated effective. More than one-third, or 35, national security programs are rated effective. And 28% or 10 of 36 training and employment programs are rated either results not demonstrated or ineffective.

The relatively poor performance of education programs may be related to the fact that many of these are grant programs, which as OMB has noted tend to under perform relative to other types of programs.

Ratings by Topic



6. Programs rated by agency[13]

Some agencies have a higher percentage of programs that are rated results not demonstrated or ineffective than others. The agency with the greatest number and percent of programs rated results not demonstrated is the Department of Education at 55% or 41 programs of 74 rated to date. Last year they were second to the General Services Administration (GSA), but this year GSA has seen a drop in the number of programs rated results not demonstrated from eight to five, or from 61% to 37%.

Other agencies with relatively large proportions of their programs rated results not demonstrated include: Department of Homeland Security with 38%, Department of the Interior (37%), Housing and Urban Development (32%), Department of Agriculture (27%), and Health and Human Services (27%).

Housing and Urban Development has a high percentage of programs rated ineffective at 16%. Department of Labor follows with 14% or four of its programs rated ineffective. The Environmental Protection Agency also has four programs rated ineffective, or 9%.

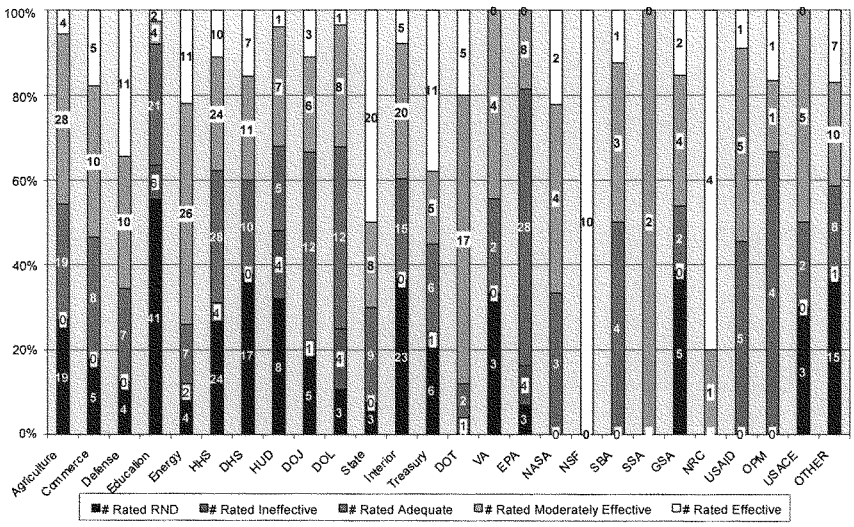
The highest rated agencies include the National Science Foundation with 100% of its programs rated effective. The Nuclear Regulatory Commission also has a high percentage of its programs rated effective at 80%. Other highly rated agencies include: Department of State (50%), Department of the Treasury (38%), NASA (22%) and Department of Transportation (20%).

[13] OMB includes a category for smaller agencies called "Other." We have extracted the five CFO agencies from this categorization for this analysis: Social Security Administration, General Services Administration, Nuclear Regulatory Commission, Office of Personnel Management and USAID. The remaining agencies in the other category include the following: Consumer Product Safety Commission, Corporation for National and Community Service, Office of National Drug Control Policy, Export-Import Bank of the U.S., Tennessee Valley Authority, Federal Communications Commission, Federal Election Commission, Public Defender of the District of Columbia, Securities and Exchange Commission, Armed Forces Retirement Home, Broadcasting Board of Governors, Trade and Development Agency, American Battle Monuments Commission, International Assistance Programs, National Archives and Records Administration, Commodity Futures Trading Commission, Delta Regional Authority, National Credit Union Administration, Court Services and Offender Supervision Agency for the District, Neighborhood Reinvestment Corporation, Appalachian Regional Commission, Denali Commission, and Smithsonian Institution.

Table 6. PART ratings according to agency

Agency	Results Not Demonstrated	Ineffective	Adequate	Moderately Effective	Effective
Agriculture (70)	19	0	19	28	4
	27%	0%	27%	40%	6%
Commerce (28)	5	0	8	10	5
	18%	0%	29%	36%	18%
Defense (32)	4	0	7	10	11
	13%	0%	22%	31%	34%
Education (74)	41	6	21	4	2
	55%	8%	28%	5%	3%
Energy (50)	4	2	7	26	11
	8%	4%	14%	52%	22%
HHS (90)	24	4	28	24	10
	27%	4%	31%	27%	11%
DHS (45)	17	0	10	11	7
	38%	0%	22%	24%	16%
HUD (25)	8	4	5	7	1
	32%	16%	20%	28%	4%
DOJ (27)	5	1	12	6	3
	19%	4%	44%	22%	11%
DOL (28)	3	4	12	8	1
	11%	14%	43%	29%	4%
State (40)	3	0	9	8	20
	8%	0%	23%	20%	50%
Interior (63)	23	0	15	20	5
	37%	0%	24%	32%	8%
Treasury (29)	6	1	6	5	11
	21%	3%	21%	17%	38%
DOT (25)	0	1	2	17	5
	0%	4%	8%	68%	20%
VA (9)	3	0	2	4	0
	33%	0%	22%	44%	0%
EPA (43)	3	4	28	8	0
	7%	9%	65%	19%	0%
NASA (9)	0	0	3	4	2
	0%	0%	33%	44%	22%
NSF (10)	0	0	0	0	10
	0%	0%	0%	0%	100%
SBA (8)	0	0	4	3	1
	0%	0%	50%	38%	13%
SSA (2)	0	0	0	2	0
	0%	0%	0%	100%	0%
GSA (13)	5	0	2	4	2
	38%	0%	15%	31%	15%
NRC (5)	0	0	0	1	4
	0%	0%	0%	20%	80%
USAID (11)	0	0	5	5	1
	0%	0%	45%	45%	9%
OPM (6)	0	0	4	1	1
	0%	0%	67%	17%	17%
USACE (10)	3	0	2	5	0
	30%	0%	20%	50%	0%
OTHER (41)	15	1	8	10	7
	37%	2%	20%	24%	17%

Ratings By Agency



Examining PART ratings by both agency and topic indicates that education programs tend to have a large number of programs that are either ineffective, or lacking in results. By contrast, foreign affairs and national security programs have a large number or percent of their programs rated effective or moderately effective.

Once more, the *Analytical Perspectives* section of the budget indicates that some of this may be due to the fact that many of the largest education and HUD programs, in terms of funding, are grant programs. OMB's analysis of grant programs shows that this type of program tends to lack in meaningful outcome data and has difficulty demonstrating results.

7. Agency program ratings as a percent of agency FY 2005 appropriations

What do these program ratings represent in terms of their proportion to the agency's total annual appropriation? Table 7 shows the ratio of the total of all FY 2005 appropriations of PARTed programs (grouped by rating) within an agency to the agency's total appropriations received, according to their FY 2005 financial statements.

Examining an agency's performance by analyzing the number of programs receiving a particular rating does not necessarily tell us about the effectiveness of budgetary resources. To get a clearer picture of agency performance according to PART, we look at the percentage of agency budgets receiving a particular rating. For example, as mentioned earlier, 55% or 41 of the Department of Education's programs are rated results not demonstrated. This represents 12% of the department's funding.

The Department of Homeland Security (DHS) and the Department of the Interior (DOI) both have relatively high percentages of their program appropriations rated results not demonstrated, 25% and 32% respectively. Veterans Affairs (VA) has 57% of its appropriations rated results not demonstrated. By contrast, the National Science Foundation (NSF) has 89% of its appropriations rated effective, corresponding to 100% of the ten programs PARTed in that agency to date. The Nuclear Regulatory Commission (NRC) also has a high percentage of its appropriations rated effective at 46%. Other high performers in terms of budget include the Department of Defense (DOD) with 29% of appropriations rated effective and NASA with 22%.

HUD stands out from all agencies as having the highest percentage of its program appropriations rated ineffective at 22%. This is not surprising considering that two of the four programs receiving this rating comprise a large portion of HUD's budget.[14]

Fifty percent of HHS's budget is rated moderately effective due to the presence of the Medicare program in this ratings category.

[14] These four programs include the Community Development Block Grant (CDBG) program, funded at $5 billion, HOPE IV, ($143 million), Project Based Rental Assistance ($4.95 billion), Rural Housing and Economic Development ($24 million).

Table 7. Percentage of agency funding levels according to ratings category

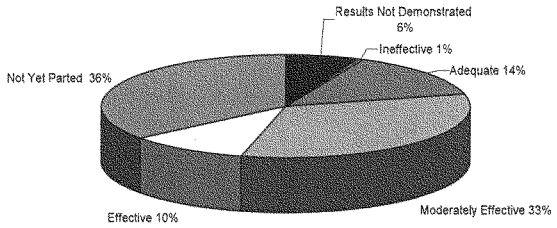
	Results Not Demonstrated	Ineffective	Adequate	Moderately Effective	Effective	Total Assessed as a percent of FY05 agency appropriations	Total Agency FY05 Appropriations Received ($mil)
Agriculture	17%	0%	22%	61%	2%	103%	89998
Commerce	5%	0%	45%	51%	11%	111%	6897
Defense	3%	0%	9%	11%	29%	53%	298656
Education	12%	3%	58%	5%	0%	78%	56678
Energy	1%	0%	34%	32%	19%	86%	21249
HHS	2%	0%	2%	50%	4%	59%	438004
DHS	25%	0%	14%	29%	9%	78%	34786
HUD	17%	22%	1%	39%	3%	82%	35448
DOJ	7%	0%	33%	24%	4%	68%	16016
DOL	0%	5%	9%	19%	1%	34%	16378
State	5%	0%	24%	15%	46%	90%	12993
Interior	32%	0%	18%	12%	2%	64%	9261
Treasury	7%	0%	14%	16%	8%	44%	15318
DOT	0%	2%	15%	75%	10%	103%	58618
VA	57%	0%	45%	2%	0%	104%	76380
EPA	1%	4%	52%	7%	0%	63%	5844
NASA	0%	0%	36%	37%	22%	95%	14903
NSF	0%	0%	0%	0%	89%	89%	4854
SBA	0%	0%	4%	4%	14%	23%	688
SSA	0%	0%	0%	22%	0%	22%	127272
NRC	0%	0%	0%	40%	48%	88%	569
USAID	0%	0%	38%	25%	0%	63%	4295
OPM	0%	0%	147%	0%	0%	148%	87998
USACE							3982
OTHER							17807
Total							1471939

8. What percentage of the budget is represented by PART ratings?

The total amount of money allotted to all of the 793 programs PARTed to date is $1.47 trillion. This represents 64% of total outlays in FY 2005 (excluding interest on the debt).[15] Breaking this out by ratings category, 6% of FY 2005 outlays are rated results not demonstrated, which amounts to $143 billion in FY 2005 appropriations. This may seem like a relatively small amount. However, some agencies have higher concentrations of results not demonstrated programs consuming a big part of some individual agency budgets as discussed in the previous section.

As noted earlier, 22% of HUD's appropriations for FY 2005 are rated ineffective or $9.5 billion of its $41 billion budget. Though ineffective programs account for only 1% of the overall federal budget, this represents $18.6 billion of all federal spending in FY 2005.

Percentage of FY05 Outlays by PART Rating

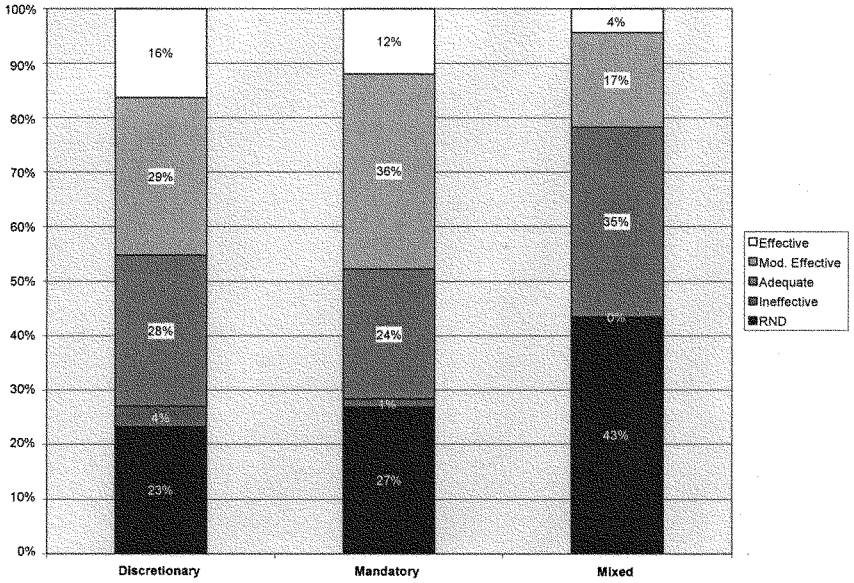


[15] Note that the budget amounts given in the PART for individual programs do not represent budget authority or outlays but instead represent 'funding levels'. This may include other kinds of spending such as fees and offsetting collections, therefore these figures are rough approximations. We take as our numerator the program budget figure or "funding level" reported in PART and calculate it as a percentage of the agency's total budget authority as reported in the agency's annual financial statement. Due to this mismatch, some fractions may exceed 100%.

9. Mandatory vs. discretionary

When we consider the budget in terms of mandatory, discretionary, and mixed spending, we are able to calculate the percentage of the budget that OMB has PARTed. Using the data for the most recent available year, FY 2006, we find that 27% of mandatory spending is rated results not demonstrated, while 23% of discretionary spending falls into this category. Forty-three percent of mixed spending (programs that have both a mandatory and discretionary component)[16] are rated results not demonstrated. Four percent of discretionary spending is ineffective, while 1% of mandatory spending is ineffective. The biggest mandatory program rated to date is Medicare, which is rated moderately effective and has a funding level of $407.2 billion in FY 2006.

Chart 9. PART ratings by mandatory and discretionary funding



[16] This should not be confused with the designation of "mixed" under program category, which defines the mechanism (e.g., a loan or a grant) by which programs allocate money.

10. Presidential funding trends

How has the president used PART in making FY 2007 budget decisions? By considering the difference between the president's funding request for FY 2007 and what Congress appropriated in FY 2006 to the 793 programs PARTed to date, we see that there is a tendency for the president to recommend funding decreases for programs with ineffective ratings (75%), while recommending increases for a large percentage of effective programs (61%). The same percentage (42%) of programs rated results not demonstrated and adequate were recommended for funding decreases. A relatively large percentage of moderately effective programs, (56%) were recommended for funding increases.

Chart 11. Difference between president's FY07 request and FY06 actual

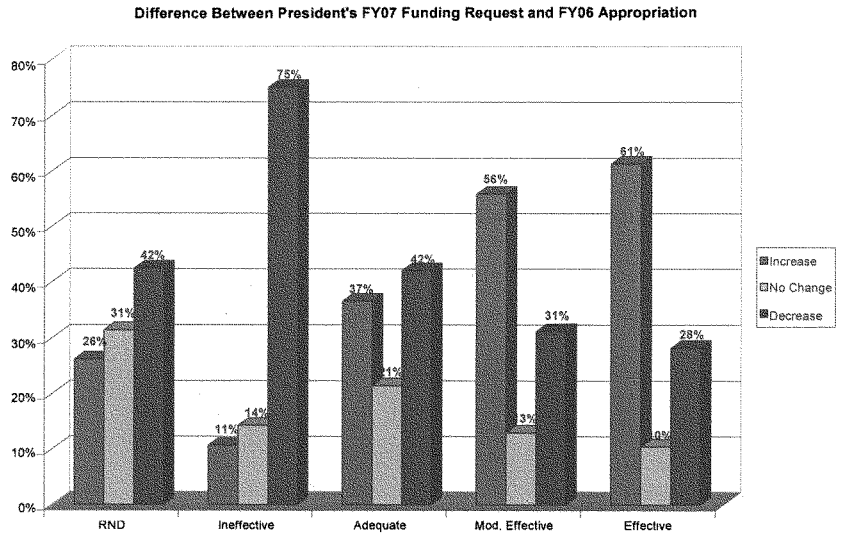


Table 10. Difference between president's FY07 request and FY06 actual

	RND	Ineffective	Adequate	Mod. Effective	Effective
Increase	50 (26%)	3 (11%)	80 (37%)	129 (56%)	76 (61%)
No Change	60 (31%)	4 (14%)	47 (21%)	30 (13%)	13 (10%)
Decrease	81 (42%)	21 (75%)	92 (42%)	72 (31%)	35 (28%)

11. How did Congress appropriate money to PARTed programs (FY 05–FY 06)?

Programs rated results not demonstrated and ineffective received fewer increases from Congress, 34% and 18%, respectively, than those rated adequate, moderately effective, and effective, while 59% of effective programs received increases in funding. Conversely, 42% of results not demonstrated programs and 79% of ineffective programs were given funding decreases. In the case of ineffective programs, the percent of programs recommended for funding decreases is slightly more than what was recommended by the president. We are not able to say if PART scores were used in making these decisions. Table 11 and Chart 11 illustrate the change in congressional appropriations between FY 05 and FY06 for PARTed programs.

Chart 11. Difference between Congress FY06 and FY05 actual appropriation

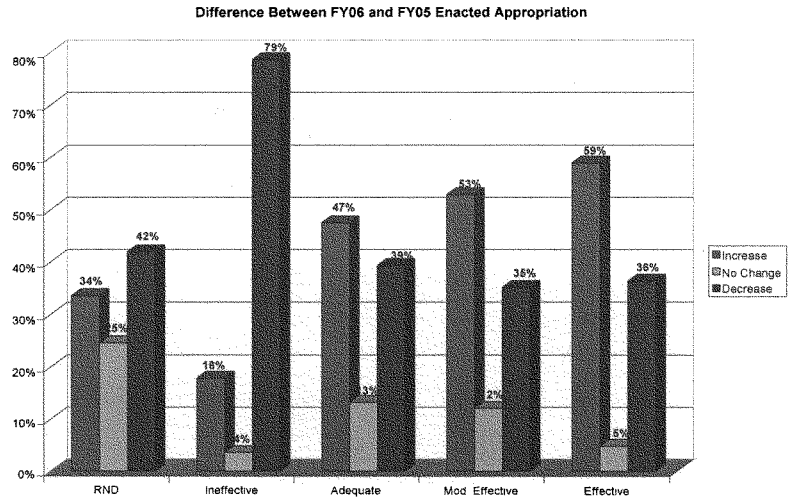


Table 11. Difference between Congress FY06 and FY06 actual appropriation

	Results not Demonstrated	Ineffective	Adequate	Mod. Effective	Effective
Increase	64 (34%)	5 (18%)	104 (47%)	122 (53%)	73 (59%)
No Change	47 (25%)	1 (4%)	29 (13%)	28 (12%)	6 (5%)
Decrease	80 (42%)	22 (79%)	86 (39%)	81 (35%)	45 (36%)

12. The president's *Major Savings and Reforms* report for FY 2007

The FY 2007 budget marks the second year that the Bush Administration has issued its *Major Savings and Reforms* report.[17] This supplemental document to the president's recommended budget contains all of the programs that the administration recommends for termination, reduction, or reform. This year the president is recommending the termination or reduction in funding for 141 programs, representing a potential $15 billion in savings. Of these programs, 91 are suggested for termination ($7.3 billion), and 50 programs are recommended for reduction ($7.4 billion). Sixteen programs are recommended for reform.

13. Ratings for PARTed programs selected for termination in FY07

Of the 91 programs recommended for termination in the FY07 budget, OMB has PARTed 32. OMB rated 15 of the programs as results not demonstrated, seven as ineffective, eight as adequate, and two as moderately effective.

In addition to poor PART scores, reasons for terminating programs include a lack of an appropriate federal role, the program completing its mission, overlap with existing programs, earmarking, and a change in budget priorities based on policy decisions.

Appendix 1 located at the end of this paper includes a chart of all 141 programs and the reason given by the administration for its recommendation.

Table 13. PART ratings and current funding levels for suggested terminations in the FY 2007 Budget

($ Mil)	RND	Ineffective	Adequate	Mod. Effective	Effective
Terminations	15	7	8	2	0
Dollar amount proposed for termination	-$2348	-$1843	-$419	-$62	$0

[17] See, http://www.whitehouse.gov/omb/budget/fy2007/pdf/savings.pdf

14. Ratings for PARTed programs suggested for reductions in the FY07 Budget

Of the 50 programs the administration recommended for reduced funding, OMB has PARTed 14. Three are rated results not demonstrated and three more are rated ineffective. Six programs are rated adequate, and two are rated moderately effective.

Table 14. Ratings for PARTed programs recommended for reduction in FY07

($ Mil)	Results Not Demonstrated	Ineffective	Adequate	Moderately Effective	Effective
Reductions	3	3	6	2	0
Dollar amount proposed for Reduction	-$620	-$819	-$1246	-$101	$0

In addition to programs recommended for termination and reduction, President Bush has proposed 16 major reforms amounting to $5.7 billion reduced spending. These reforms include re-proposing the Strengthening America's Communities Initiative. First introduced in the FY 2006 budget, the proposal would consolidate 17 existing community and economic development programs under one program in the Department of Commerce.

15. What did Congress do in response to last year's *Major Savings and Reforms* report?

In FY 2006, the president recommended that 154 programs be terminated or allotted less funding. Congress accepted 89 of the president's recommendations, in full or in part, for a total reduction in spending of $6.5 billion.

Of the 99 programs recommended for termination last year, Congress terminated 24 of them and reduced funding for 28, yielding a total savings of $2.7 billion.

Of the 55 programs proposed for reduction, Congress reduced funding for 37 programs, leading to a savings of $3.78 billion.

16. Did PART play a role?

Of these 154 programs recommended for termination or reduction for FY 2006, OMB PARTed 54. Congress agreed to terminate or reduce funding for 21 of the 54 PARTed programs. Whether the PART evaluation played a role in Congress's decision on these programs is not certain. Congress does not detail whether PART evaluations were considered in their decisions to terminate or reduce funding for these programs. *Appendix 2* provides a full listing of the programs and their associated Congressional action.

It should be noted that Congress terminated or reduced funding for additional programs not included in the president's recommendations. According to the U.S. House of Representatives Committee on Appropriations, Congress eliminated a total of 53

programs for a savings of $3.5 billion. Some of these (24) are in response to the president's recommendations, while Congress eliminated the remainder at its own prerogative. These programs are also included in *Appendix 2.*

This is an increase over previous years. In FY 2005, the president proposed terminating 65 programs but Congress only adopted seven of these recommendations, reducing spending by $366 million.

III. Conclusion

The purpose of this study was to apply PART data in order to answer some basic questions about agency and budgetary performance. Overall, programs have moved from not having performance measures and data, to developing information to enable periodic evaluation of their performance. The number of programs rated results not demonstrated has decreased from 50% in FY 2004 to 24% in FY 2005. Though an improvement, this still represents 6% of federal outlays, meaning we do not have sufficient information to judge the performance of $143 billion of the federal budget. One percent of total outlays are rated ineffective representing $18.6 billion in spending in FY 2005.

As last year, Department of Education programs continue to have the largest number of results not demonstrated (55%), representing 12% of its funding in FY 2005. The Department of Housing and Urban Development also has a large number of its programs rated ineffective, at 16%, representing 22% of its funding in FY 2005. This is due to the fact that two of its largest programs: the Community Development Block Grant program and Project-Based Rental Assistance, received $4.1 billion and $4.95 billion in funding in FY 2005, representing a large portion of HUD's annual funding level.

According to the president's *Major Savings and Reforms* report, PART continues to inform some, but not all, Executive decisions in the proposed budget. Of the 141 programs proposed for either termination or reduction in FY 2007, 46 have been PARTed.

Calculating the difference between what the president proposed for funding in FY 2007 with what Congress appropriated to the program in FY 2006, we find that 75% of programs rated ineffective are recommended for funding decreases, while 61% of programs rated effective are recommended for funding increases. There is not a perfect correlation however. Eleven percent of ineffective programs are recommended for increases, and 28% of effective programs are recommended for decreases.

This mirrors congressional action. When we consider the difference between what Congress appropriated to programs in particular ratings categories in FY 2005 with what it appropriated to programs in those ratings categories in FY 2006 we find that 79% of programs rated ineffective were given funding decreases, while 59% of effective programs were given funding increases. Conversely, 18% of ineffective programs were given funding increases, while 36% of effective programs were given funding decreases.

In the case of ineffective programs, Congress gave funding decreases to more programs, than recommended by the president. We are not able to say if PART played a role in Congress's decisions to terminate or reduce funding for programs.

The Committee on Appropriations notes that, "the only way to establish accountability in the budget process is to stop spending on programs that have outlived their usefulness or could be delivered more effectively at the state or local level."

PART, it should be noted, is the Executive's attempt to advance performance budgeting. Trying to link budgets with performance information is an idea that originated in 1994 under GPRA. Though PART has advanced a particular method for evaluating government activity, using PART to make congressional decisions is not the goal, rather it is to encourage agencies to gather and report on program activity by establishing and using reliable outcome measures. This also means open and frequent dialog between program managers and Congress on the policy aims and intent of programs Congress has established to achieve its goals. Imparting increased transparency, and consistency, to the budget process means Congress and the Executive must systematically evaluate program activity and show taxpayers how public benefits are being achieved by either funding or de-funding activities that Congress has deemed a federal responsibility.

If Congress is to truly implement GPRA, i.e. to link budget and performance information in order to strategically allocate resources, it must first require reliable, consistent, performance information from agencies, and then it must use it, in conjunction with other information. This also means moving the appropriations debate from one of dollars spent to one of public benefits sought and achieved.

PART's methodology should continue to be subject to criticism and scrutiny, but this should not detract from PART's main contribution, which is to forward performance budgeting within agencies, while bringing increased transparency and accountability to the budget process inside the Executive Branch.

Testimony to the Senate Homeland Security and Government Affairs Subcommittee on Federal Financial Management, Government Information, and International Security

June 13, 2006

Adam Hughes, MA
Director of Federal Fiscal Policy
OMB Watch

Chairman Coburn, Ranking Member Carper, members of the subcommittee: My name is Adam Hughes and I am the Director of Federal Fiscal Policy at OMB Watch — an independent, nonpartisan watchdog organization. Thank you for inviting me to testify today on what we all can agree is a crucial cause — making our government the most effective and responsive it can absolutely be.

OMB Watch was founded in the early 1980s and has spent over twenty years advocating for government accountability, transparency and access to government information, and citizen participation in governmental processes. OMB Watch believes citizens must take an active role in holding their government accountable and that the federal government, when supported by sensible fiscal policy, can develop the programs and safeguards that meet the public's needs.

This issue has taken on added importance during the Bush administration as a combination of factors, some avoidable, some not, have plunged the federal government into debt. Large and sustained deficits over the past five years have made efficient use of government resources all the more important. In light of the anticipated budget crunch due to the baby boomers retirement over the coming decades, the fiscal situation of this country will only deteriorate further. Performance measurement can therefore become a particularly attractive alternative for those who want to set federal priorities based on the current fiscal prospects of a strained and shrinking revenue base (that is, without expanding that base to fund longstanding programmatic commitments).

OMB Watch has been commenting on government performance issues for the better part of its existence. We have spent more time analyzing the Government Performance and Results Act (GPRA) and the PART over the last ten years as government itself has implemented multiple initiatives and mechanisms to attempt to gauge whether goals are being met.

We are supportive of the concept of improving federal capacity to meet the public's needs. OMB Watch has worked for over 20 years to protect and improve that capacity, and we have been open to possibility of using performance measurement as a means for achieving those ends. We bring a strong belief in the importance and potential of government itself to the work we do, and because of that belief, we want government to be responsive to community needs, spend money effectively, and accomplish it goals. We are advocates for government and therefore have a very strong self-interest in seeing government programs get results.

PART, however, is a very poor mechanism for measuring program performance and results, introducing biases and a skewed ideological perspective into a model claiming to present consistent and objective performance data and evaluations of government programs. Often times, the PART actually decreases the efficiency and effectiveness of government through increased administrative burdens, distracted managers, and compliance costs.

Ironically, the PART mechanism itself does not produce the right type of results to further support and improve government. We believe PART ratings should not be directly connected with the budgeting process of Congress because of significant deficiencies — mainly the substantial biases and limitations embedded within the tool and the additional limitations we have observed in OMB's actual application of PART.

Based on our studies of PART and our longstanding commitment to an open, accountable government that is responsive to the public's needs, I come to you today with three points to make:

(1) PART continues a troubling trend we have seen in other executive branch initiatives and even congressional proposals—namely, a trend to arrogate increasing power to the White House, even in areas that by constitutional design have been committed to Congress.

(2) PART is so limited and distorted a tool that it should be used neither for management nor for budget and appropriations decisions. Both by the design of the tool and as the mechanism is implemented, PART systematically ignores the reality of federal programs and judges them based on standards that are deeply incompatible with the purposes that federal programs are expected to serve. As one agency contact memorably explained to us, PART assessments are tantamount to a baseball coach walking to the mound to remove his pitcher and then chastising him for not kicking enough field goals as he brings in a reliever.

(3) There is a better way. Specifically, Congress already has the means to investigate and produce far more sophisticated analyses of the usefulness, effectiveness, and results of government programs. In fact, this is one of the primary, if not *the* primary role of the legislative branch. While the oversight function of Congress may not be as robust as it once was due to significantly shorter legislative sessions and delays due to a sharply divided political climate, the capacity to judge the results of government programs already exists within the existing structures of Congress — structures that do not carry with them the significant limitations and negative consequences of the PART.

I. PART: EXAMPLE OF BROADER SHIFTS IN POWER IN GOVERNMENT

Before I discuss some of the specific weaknesses and negative consequences of the PART, I want to point out a larger trend in government over the last few years that we believe PART is connected to. Since the Bush administration came into office and after the terrorist attacks in September 2001, we have seen a steady shifting of power to the executive branch in many different facets of our government — particularly security and military policy.

Yet this larger trend toward increased executive power has spilled over into other areas outside security and defense. Some of the "budget process" changes currently being considered by Congress also have a tendency to consolidate yet more power in the White House. Specifically, the president's enhanced recession proposal scheduled for debate this month in Congress and a proposal to establish sunset commissions gaining traction in the House are indicative of this larger trend by allowing the President increased power over spending priorities and program authorizations— activities that are the proper domain of Congress These proposals represent a disturbing trend.

In some ways PART is even worse than those proposals for two distinct reasons. First, PART is more insidious: whereas the other proposals openly seek to arrogate power to the White House, PART portrays itself as an unbiased evaluator of results and performance while serving the White House's political priorities. As I will discuss today, PART is anything but an even-keeled evaluator of government programs.

Second, the White House is using PART to supplant Congress's role and even to contravene long-settled Supreme Court precedent. By instituting the president's or OMB's subjective policy preferences and biases for those of the other branches of government, the PART is a seemingly innocuous tool for the executive to manipulate the balance of power across all of the federal government and remove some of the checks and balances that are an integral part of our representative political system. For this reason alone, PART should be approached extremely cautiously by those outside the administration.

II. TWO TYPES OF BIASES LEAD TO FLAWED TOOL

I would like to focus on two main aspects of the biases inherent in PART. I believe these biases are significant and numerous enough to discredit the PART from being heavily or directly involved with both budget requests and appropriations and also management of programs.

Both by the design of the tool and as the mechanism is implemented, PART systematically ignores the reality of federal programs and judges them based on standards that are deeply incompatible with the purposes that federal programs are expected to serve. As one agency contact memorably explained to us, PART assessments are tantamount to a baseball coach walking to the mound to remove his pitcher and then chastising him for not kicking enough field goals as he brings in a new pitcher.

A. STRUCTURAL BIASES EMBEDDED IN TOOL DESIGN

1. Overly Simplistic Model Fails to Capture Diversity, Complexity, and Possibilities of the Federal Government

The intricacies of the federal legislative process, the necessity of crafting coalitions to pass legislation, and the shifting face of congressional representation often lead Congress to create and later amend a wide diversity of federal programs with multiple, and at times conflicting, goals. The PART tool — because of its crude design and over-simplified rating system — is not robust enough to capture the complexity inherent in the federal government.

First and foremost, the black and white rating scale (ranging from effective to ineffective) ignores the multiple and diverging reason a program could be succeeding or failing. Different program have different problems for different reasons. Perhaps a program is struggling to achieve its mission because it is underfunded and an ineffective program deserves more resources. The PART ratings are unable to convey such complexity.

The one-size-fits-all approach of the PART review process often minimizes or ignores important differences in purpose and design between varying types of government programs, possible intentionally overlapping goals between programs and departments, and even multiple goals Congress has charged a single program with achieving.

Social problems are complex and diverse, and federal programs must accordingly take many shapes, attempt many approaches, and address a wide range of needs. The assumptions embedded in the very design of PART—that all that can be meaningfully known about programs is quantifiable; that programs have a single, unitary purpose that never adjusts to changing circumstances; that the only meaningful work performed by federal programs leads to a single outcome—are short-sighted assumptions that embody a narrow and simplistic vision of the role of government. It is simply too crude to serve as a useful guide for government management.

Perhaps the most obvious failure of the tool in this regard is its narrow insistence on outcome measures as the benchmark of programmatic success. The outcome measurement straightjacket is problematic because it is inadequate to the task of informing the management of programs that can only be measured in terms of outputs or that are difficult to measure in terms of either outputs or outcomes. This blind adherence to outcome measures in the tool design fails to accommodate some very important types of programs. For example:

- Multiple programs with varying approaches to the same problem, block grants, competitive grants, and demonstration grants are all ways to experiment with solutions to complex social problems. Grants to state and local governments, for example, attempt to take advantage of the fabled "laboratories of democracy" to experiment with ways to attack persistent and often intractable social problems. For some issues, such as foster care, Congress has decided that multiple programs in multiple agencies and departments — including the Title IV-E entitlement, the Adoption Assistance program, the Chafee Independence Living Program grants, Medicaid, special education services, and more — are needed to meet the needs of abused and neglected children. PART's rigid criteria for uniqueness and unitary performance goals ignore the value of multiplicity and overlap and create perverse incentives to recentralize in the federal government what Congress has decided to shift to the states.

- Research programs, such as the National Toxicology Program and the IRIS database, are intended to close gaps in our knowledge rather than lead to immediately measurable outcomes such as reduced incidence of cancer or decreases in lifetime fatality risks from exposure to toxic substances. In these cases, improvements in what we know and what we can reasonably determine are valuable in and of themselves, not because they lead to other measurable consequences. The PART tool fails to recognize the value in pure research programs and the like; not only does PART therefore fail to

offer anything of value to the management of such programs, but it also threatens to lead to reduced funding and distorted priorities for no justifiable reason.

- Research programs are the canary in the coalmine for another limitation of the tool: its bias for short-term impacts rather than long-term efforts. Every EPA research program PARTed as of the FY06 budget was assessed as "Results Not Demonstrated," (RND) 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 to the very nature and benefits of research: often we can learn as much from failure as from any success. This bias is built into the design of the tool itself, according to a member of EPA's Science Advisory Board, who testified "it appears that the weighting formula in the PART favors programs with near-term benefits at the expense of programs with long-term benefits. Since research inevitably involves more long-term benefits and fewer short-term benefits, PART ratings serve to bias the decision-making process against programs such as STAR ecosystem research, global climate change research, and other important subjects."[1]

- Many programs are created to address concerns that are broader and deeper than PART, with its insistence on quantifiable outcome measures, can begin to accommodate. The Americorps National Civilian Conservation Corps, for example strives to achieve the goals of "strengthening communities" and "increasing civic responsibility." It is not possible to establish quantifiable measures of community strength, but that impossibility does not mean that the communities themselves cannot attest to their strength. In such cases, the real measure of success will have to be subjective and narrative — and must include outside stakeholder input in order to balance competing perspectives and viewpoints. PART contains no avenues for stakeholder input into the program review process. Using PART as a management guide will threaten such programs and lead to a government that has no vision and fails to embody the public's most cherished values.

A management tool that disapproves of visionary, values-driven, future-oriented, or knowledge-creating programs is a tool for mismanagement, which would detract from what a federal government is uniquely situated to do.

2. PART Creates Increased Management, Compliance, and Data Burdens

Over the years since the PART was first introduced, the review process has often forced program managers and agencies to alter their existing management and performance review practices, institute new and costly data collection structures and systems, generate independent reviews and analyses from outside the government and overlay this performance initiative with previous government efforts. These alterations to program management have created an entire compliance system within itself that distracts energy and resources from achieving program goals.

---

[1] Testimony of Dr. Genevieve Matanoski, EPA Science Advisory Board, before Subcomm. on Environment, Technology, and Standards, House Comm. on Science, Mar. 11, 2004, *available at* 2004 WL 506081.

PART often conflicts or complicates other government wide reform initiatives. Collecting new types of data within agencies for OMB in order to comply with the PART rating system is often constrained by the Paperwork Reduction Act, which requires agencies to reduce the number of data elements collected. Further, the PART and the Government Performance and Results Act (GPRA), which attempts to develop strategic goals and department and government cross-cutting comparison for the federal government through a much more open and accessible process than the PART mechanism, are often in conflict with each other, creating added management difficulties and increased compliance burden within agencies.

Furthermore, there are significant obstacles to the data collection that PART demands. Agency data collection is constrained by the Paperwork Reduction Act, which requires agencies to obtain OMB approval before conducting any information collection that asks the same questions of ten or more people. Additionally, data collection efforts, especially the independent evaluations PART expects programs to rely on, can be expensive, but PART does not excuse programs that cannot collect the expected level of data because of a lack of funding. OMB itself is responsible for these obstacles, even as it penalizes programs for running into them.

Between this Catch-22 and the sometimes absurd mismatch of PART measures and actual program purposes, program staff have learned to treat PART as a compliance exercise instead of a guide to better management. OMB Watch has conducted extensive, in-depth interviews with agency staff involved in PART assessments at the program level. We have heard repeatedly that agency staff have spent considerable time "gaming" the PART system—learning the pressure points and pitfalls to avoid negative scores and consequences. A performance appraisal system so widely regarded as a mere compliance exercise offers little diagnostic benefit for agency program managers and is another indication PART scores should not be related to budget allocations for programs.

3. PART's Bias Toward Specific Program Types

The extreme biases against block grant programs within the PART process are perhaps the most egregious and the most obvious example of the problems embedded in the very design of the tool.

Programs that operate through grants, whether competitive grants or block grants, are rated lower on average than all other programs. When OMB rated block/formula grant programs (a category that includes both block grants and entitlements) in FY 2005 process, it found no block/formula grant programs were "effective" while finding 11 percent of programs in general were "effective." In addition, OMB found 43 percent of block/formula grant programs to be ineffective while determining only 5 percent of programs overall were "ineffective."

The chart below compares the overall breakdown of PART scores in competitive grant programs, block grant programs, and all other programs after the reviews were completed for FY 2006. 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 president's FY 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%

There is an easy explanation for this trend. Federal grant programs largely send money to the state and local governments, a system established intentionally by Congress because they have realized that in some instances it is vastly more efficient to allow individual states the flexibility to tailor their respective programs and initiatives to suit local and regional needs. Some, like the Community Development Block Grant, are particularly important to improving the economies of poor and rural communities across the country with locally designed projects and programs that address specific community needs. Entitlements, meanwhile, will always fail the PART's demand for linkages between performance goals and revenue allocation, because entitlements are automatic distributions to entitled populations — in other words, PART scores them negatively for being exactly what Congress intended them to be.

The odds are stacked against block/formula grants within the PART 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.

B. POLITICAL BIASES EMERGE THROUGH IMPLEMENTATION

1. Inconsistencies with Presidential Budget Request Cast Doubt on Purpose of Tool

A quick glance at PART ratings and budget requests should dissuade anyone from trying to find a logical or consistent pattern between them — there is no pattern. Even after reviewing almost every federal program and being used to develop multiple budget requests, it remains unclear if even the Bush administration uses the PART ratings to inform their budgeting decisions at the start of each year.

OMB Watch has conducted analyses of the list of programs highlighted by the President in each of the last two State of the Unions for not achieving the required results, as well as the broader list of programs reviewed under the PART and found some puzzling results. A few examples:

- 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 grant program run by the Tennessee Valley Authority.

- Of the programs rated "ineffective," in the 2006 budget that were targeted for elimination, more than 78 percent came from the Departments of Housing and Urban Development or the Department of Education.

- The Substance Abuse Prevention Block Grant, a program that provides grants to state 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 lower income working Americans who have claimed the EITC and double checks their eligibility for the credit — was rated ineffective, yet received a substantial funding increase.

The examples above are not used to cherry pick arbitrary cases, but underscore a larger pattern of inconsistency. Most troubling, in each of the last two years, of those programs singled out by the president for failing to produce results, **more than two-thirds had yet to be reviewed by the PART questionnaire**. In many more cases than not, it is unclear what kinds of determinations, if any, the president used to identify these supposed failing programs when the White House budget staff has not even used their own performance review tool to assess them.

While other analysts have criticized the failure of the PART to establish a toehold in the budget formulation process, we believe these facts point to a larger problem that underscores the need for Congress to be highly dubious of the usefulness of using PART scores to inform budget decisions. The lack of consistency among ratings and the president's own budget requests points to the possibility that the PART is merely a rhetorical tool to support pre-ordained political conclusions.

2. PART Sends Management Signals that Would Distorts Federal Priorities

OMB uses PART to alter the management of federal programs in troubling ways. The PART mechanism allows for OMB perspectives and policy preferences to be inserted into the oversight and management structures of federal programs without congressional approval. Agency staff implementing federal programs are subordinate to OMB within the construction of the survey answers in PART, and experience concrete consequences — such as flat or decreased budget requests and, if the administration is successful with pay-for-performance proposals, even the inability to receive an annual salary increase — if they fail to heed the management signals OMB sends through PART. As a result, PART has enormous potential to distort federal priorities in ways that Congress has never permitted.

OMB is, unfortunately, taking advantage of that potential. Many of the stated reasons for scoring programs negatively reflect nothing more than OMB's disagreement with the way Congress

designed a program by law. 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 (CPSC), Occupational Safety and Health Administration (OSHA), and Mine Safety and Health Administration (MSHA) 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 demand for using net benefits as a criterion for regulatory decisions, 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) in FY 2006 by claiming through the PART review that it was not 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 from the PART review 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 (83 percent) in the president's budget this year.

Interestingly enough, these examples are no longer necessary; in a recent hearing before this very subcommittee, an OMB official was asked point-blank whether it is possible for programs to receive low PART score simply because it follows the law, and OMB answered, simply, "Yes."[2]

This distortion of priorities is also happening in a host of more subtle and indirect ways. Buried in the small type of the specific program assessments, the standards actually used to measure program "effectiveness" or "results" very often fail to focus on what is most meaningful or relevant about a program. One particular example is that 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."

[2] Add citation to Carper/Johnson colloquy.

This measure is not a scientifically appropriate measure of actual water quality: as EPA recently 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.[3] 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.

These conflicts between the statutory mandates imposed by Congress and the willful arbitrariness of OMB are waived away when the assessments are offered to Congress, and the scores are attributed to the program's "ineffectiveness" or failure to demonstrate results rather than OMB's decision to measure programs with inapposite criteria or include subjective judgments about a program's worth.

3. Grade Deflation Allows OMB to Manipulate Levers of Congressional Spending

While OMB has gone to great lengths to advertise the PART as having an unprecedented level of transparency for the public by unearthing vast amounts of government information for the public. While it is certainly true OMB has marketed the PART to the public as an open government initiative, the most crucial decisions, value judgments, and processes for arriving at the final product of a PART rating still remain largely hidden. These can often be the most important aspects of the entire process, masking a biased or manipulated product.

One of the most glaring examples of this is the "Results Not Demonstrated" (RND) rating. It is not clear how OMB determines which programs should be shifted into the category of RND. 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 in PART materials accordingly as an *indeterminate* category, programs relegated to the RND bin are often characterized in White House rhetoric as failing programs.

Indeed this fact has been confirmed during our interviews with agency employees as all those interviewed told us the RND rating was the absolute worse one a program could receive under the PART — far worse than an "Ineffective" rating.

The RND score is based on failure in a couple of specific questions. It is interesting to observe, however, that many of the programs scored RND otherwise score more highly in the section for producing *results*. In fact, 72 of the 178 programs (40 percent) categorized as "Results Not Demonstrated" by FY 2006 had scores that, according to OMB's own grading scale, would have been granted passing scores if not for failure on the specific RND-determining questions. 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 programs — 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 —

[3] Add citation to Aug 2004 announcement – this probably came from BNA

actually scored above 75 percent for producing results (substantially higher than the 60 percent average for all programs rated "Moderately Effective").

Moreover, the remaining 60 of the 72 otherwise passing programs would have received the middle passing score of "Adequate" — again, if not for failing the specific RND-determining questions. 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.

There is no explanation given for the weighting assigned to any of the particular questions or sections, nor for the absurd results once these weights are assigned. This inconsistency highlights an important point that emerged from our agency interviews as well. Implementation of the PART survey is highly dependent on the individual program officer at OMB, and working with different officers can not only completely alter the process by which the survey is completed but also the final rating for the program. The Government Accountability Office has also concluded that the PART gives a high level of influence to budget officers at OMB and leads to inconsistent application of the tool across the federal government.[4]

III. THERE IS A BETTER WAY

In much the same way other "budget process reform" proposals seek to increase the executive's control over federal revenues and spending priorities, PART also attempts to alter the balance of power within the federal government. The tool gives the executive a mechanism by which to impose its budgetary preferences, however political or biased, on Congress in a seemingly benign way by wrapping them in good government and results rhetoric.

While the President is certainly free to classify federal programs in whatever way he believe is best and recommend those programs be supported with increased funding or eliminated according to his own preferences, it is disingenuous to attempt to pass off subjective and, at times, politically motivated policy conclusions as unbiased program reviews.

There does seem to be some usefulness for the PART review process to serve as a diagnostic tool for program managers and agency employees. In particular, a process known as a PART cross-cut undertaken by OMB has shown significant promise as a model to improve efficiency of management and stewardship of specific programs across different agencies and departments. To our knowledge, this process was devoid of attempts to connect the results to significant alterations in budget priorities or alterations to the management agenda for implementing policy decisions. It true, I believe these are certainly aspects of the cross-cut that allowed it to be a productive exercise.

In order for the variety of actors whose input is needed to make formulate budgeting decisions to use any type of performance review mechanism, it is crucial for those actors to believe the information is credible and constitutes a consensus on objectives and goals.

---

[4] See Government Accountability Office, *Performance Budgeting: Observations on the Use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget,* No. GAO-04-174 (Jan. 2004)

This has not been the case with the PART. Many individuals both inside and outside of Congress remain highly skeptical of this tool and the process by which the ratings are determined by OMB. Perhaps the biggest reason for this belief is because the PART is attempting to reinvent the wheel from a new perspective. Congress already has the structural and institutional capacity to develop a rigorous system of determining results and effectiveness of government programs through the appropriating and authorizing processes. The vast resources of the Congress available within the committee and personal staff structure as well as in the offices such as the Congressional Budget Office, Government Accountability Office, and Congressional Research Service are more than sufficient to provide far more robust information about program performance and results.

Most importantly, relying too heavily on the PART ratings not only will gradually remove Congress from its funding and oversight responsibilities granted under the Constitution, but also will continue to close the door on opportunities for outside stakeholder interests to be infused into the congressional budgeting and evaluation process. This limited perspective on programs and goals is a crucial deficiency of the PART. By limiting the perspective of the reviews, the subjectivity and bias that will almost always creep into any time of rating does not have a counterbalance from a wide range of outside stakeholder interests.

While the expansion of the executive branch powers has been present in our government, particularly during times of war, since the turn of the last century, the overreach of those powers into areas historically and constitutionally given to Congress — the structuring of programs, appropriating and authorizing of revenues, and oversight of government — is a disturbing trend. Because of this, PART scores should be taken with more than just a grain of salt or even a hefty dose of skepticism by Congress. Unless the tool design and implementation system are significantly modified, they should probably be largely ignored.

Thank you for the opportunity to share our views with you here today. I look forward to your questions.

FISCAL POLICY PROGRAM

www.ombwatch.org/budget

Appendix

Appendix to Testimony
Before the Subcommittee on Federal Financial Management, Government Information, and International Security of the Senate Committee on Homeland Security and Government Affairs

Case Examples of Absurd Results from PART Scores

ExpectMoreOfTheSame.gov

From ThinkProgress, http://thinkprogress.org/2006/02/06/omb-website/, Feb. 6, 2006

Appendix to Statement of

Adam Hughes
Director of Federal Fiscal Policy

ExpectMoreOfTheSame.gov

The White House released its 2007 budget today, and budget director Joshua Bolten unveiled a new website – ExpectMore.gov — that "allows taxpayers to review the [White House] assessments of nearly 800 federal programs." "Here, you can see the exhaustive work that goes into each one of these assessments," Bolten said at today's press conference. "I expect that this website will be a useful tool for everyone who care about how tax dollars are spent."

Their "exhaustive work" produced a delusion-riddled website that showcases the White House's inability to assess its own problems and weaknesses. Katrina offers a real-world illustration of the new site's inaccuracies:

1) "Federal Emergency Management Agency: Disaster Recovery":

> The Department of Homeland Security's Recovery program ensures that individuals and communities affected by disastes [SIC] of all sizes, including catastrophic and terrorist events, are able to return to normal function with minimal suffering and disruption of services.
> **PERFORMING: Adequate (one star)**

Reality — Reuters:

> With **no clear recovery plan in sight five months after Hurricane Katrina**, many victims are simply hanging on, waiting anxiously for signs that their neighborhoods are either reviving or turning into permanent ghost towns.

2) "Preparedness — Grants and Training Office National Exercise Program":

> Prepare Federal, state, and local responders to prevent, respond to, and recover from acts of terrorism by providing the tools to plan, conduct, and evaluate exercises. **PERFORMING: Effective (three stars)**

Reality — GAO:

> Although the [National Response Plan] framework envisions a proactive national response in the event of a catastrophe, **the nation does not yet have the types of detailed plans needed** to better delineate capabilities that might be required and how such assistance will be provided and coordinated.

3) "Federal Emergency Management Agency: Disaster Response":

> The Department of Homeland Security's Response program is designed to quickly, efficiently and effectively provide support to State, Tribal, and local governments, and Federal response teams in the event of a natural or manmade disaster, emergency or terrorist event
> **PERFORMING: Adequate (one star)**

Reality — Washington Post:

> Four years after the Sept. 11, 2001, attacks, administration officials did not establish a clear chain of command for the domestic emergency; disregarded early warnings of a Category 5 hurricane inundating New Orleans and southeast Louisiana; and did not ensure that cities and states had adequate plans and training before the Aug. 29 storm, according to the Government Accountability Office.

Filed under: Katrina

Posted by Payson February 6, 2006 5:28 pm

FISCAL POLICY PROGRAM

www.ombwatch.org/budget

Appendix

Appendix to Testimony
Before the Subcommittee on Federal Financial Management, Government Information, and International Security of the
Senate Committee on Homeland Security and Government Affairs

Problems of PART for Research Programs

Excerpt from Fiscal 2005 Budget: EPA: Hearing before the Subcomm. on Envt., Tech. & Stands. of the House Comm. on Science, 108th Cong. (2004), 2004 WL 506081 (F.D.C.H.) (statement of Dr. Genevieve Matanoski, EPA Science Advisory Board).

Appendix to Statement of

Adam Hughes
Director of Federal Fiscal Policy

2. Program Planning and Measurement

a. Program Assessment

Each year, the [Environmental Protection Agency's Science Advisory] Board tries to evaluate EPA's research priorities and their role in meeting the Agency's goals. As part of the current review, the Board was given information resulting from the application of a new survey tool, the Program Assessment Rating Tool (PART) that was used to evaluate selected EPA programs. The Board is concerned that decisions are being made about research program funding on the basis of the application of this new tool.

To be clear, the Board did not receive or review information on the rating instrument itself; however, after evaluating PART summaries for several research programs, our conclusion is that PART may, at this time, have a limited capacity to inform budget decisions on research programs. The Board is concerned with the manner in which the weighting formula in PART seems to influence the full analysis and thus favor programs with short-run results over those having long term results. There is also concern that an evaluator's subjective considerations might be able to bias those weights and the rating itself.

Specifically, it appears that the weighting formula in the PART favors programs with near-term benefits at the expense of programs with long-term benefits. Since research inevitably involves more long-term benefits and fewer short-term benefits, PART ratings serve to bias the decision-making process against programs such as STAR ecosystem research, global climate change research, and other important subjects. The PART seems to be intended as a formula for predictions about likely program success. However, the weights that the PART assigns to different program characteristics do not seem to have been validated systematically against the contribution of each program characteristic to any independent objective measure of program success. If the weights in the tool are arbitrarily assigned, the PART may have characteristics that could lead to biases in evaluation that are related to the subjective judgments of its designers. We believe that the tool should be reviewed to determine its adequacy for its use in supporting budget decisions.

As the Board observed significant decreases in science and research funding, it also noted a substantial resource increase in the State and Tribal Assistance Grant account (STAG) for an initiative for retrofitting school busses. The Board does not challenge the worthiness of this program, rather it notes that it has no information on the science supporting this initiative. The Board trusts that the benefits of this program have been rigorously reviewed.

The real issue here is how research programs (and others) are to be evaluated and whether a different metric is necessary for basic vs. applied research programs. Also, of interest is whether research results should be evaluated separately from the outcomes of programs they are intended to support? Although the Board did not directly evaluate the PART itself, it is of obvious difficulty to conceive of a simple quantitative metric that could be applied across the broad areas of ecosystem quality, human health effects, endocrine effects, and technology development. The question is even more complex when you consider that some research is intended to develop limited data in the short-run to fill a specific knowledge gap and other research is intended to provide an understanding of whole systems in the long-term. Research program measurement is even more difficult because the knowledge and methods developed by EPA, especially ORD's

researchers, are not usually directly applied by ORD, rather they are often used by others to support decisions on a broad suite of diverse statutory mandates. Thus, we believe that evaluations of the performance of research programs will need to consider the specific factors of each program that the research is intended to support. Further, it is unlikely that simple formulas will be able to handle this task well. It is more likely that realistic research program performance assessment will need to be a combination of quantitative metrics and other information and analyses which is then evaluated by groups of experts with relevant knowledge.

I note that the NAS, in its review of STAR, also had concerns with quantitative routines used in performance assessments and noted that "The Committee judges that expert review by a group of people with appropriate expertise is the best method of evaluating broad research programs, such as the STAR program."

— Statement of Dr. Genevieve Matanoski, EPA Science Advisory Board, to Subcommittee on Environment, Technology, and Standards, House Committee on Science, March 11, 2004, 2004 WL 506081

FISCAL POLICY PROGRAM

www.ombwatch.org/budget

Appendix

Appendix to Testimony
Before the Subcommittee on Federal Financial Management, Government Information, and International Security of the Senate Committee on Homeland Security and Government Affairs

Limitations of Performance Measures & PART's Federalism Problems

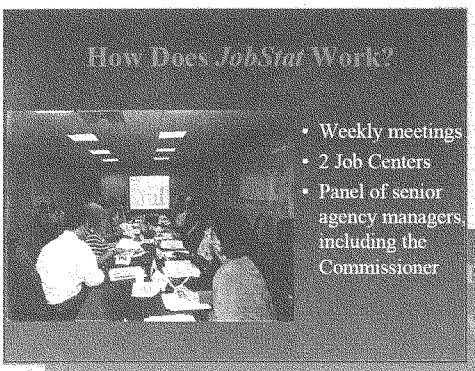
From Performance Management in State and Local Government (Nelson A. Rockefeller Institute of Government, 2005)

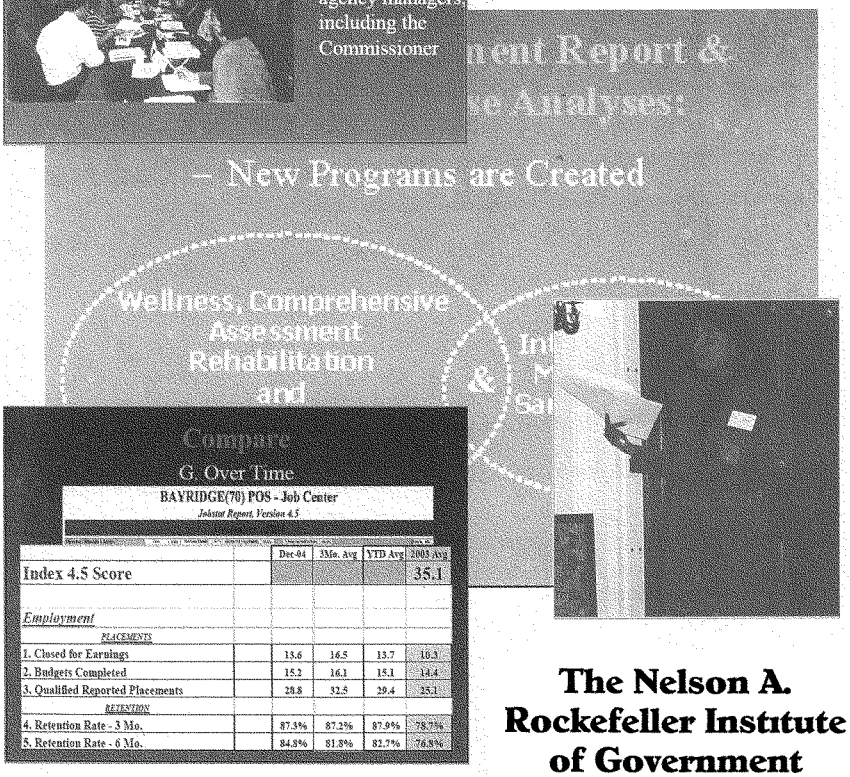
Appendix to Statement of

Adam Hughes
Director of Federal Fiscal Policy

1742 CONNECTICUT AVE. NW • WASHINGTON, DC 20009 • T/202.234.8494 • F/202.234.8584
www.ombwatch.org

Performance Management in State and Local Government





Compare
G. Over Time

BAYRIDGE(70) POS - Job Center
Jobstat Report, Version 4.5

		Dec-04	3Mo. Avg	YTD Avg	2003 Avg
Index 4.5 Score					35.1
Employment					
PLACEMENTS					
1. Closed for Earnings		13.6	16.5	13.7	10.3
2. Budgets Completed		15.2	16.1	15.1	14.4
3. Qualified Reported Placements		28.8	32.5	29.4	25.1
RETENTION					
4. Retention Rate - 3 Mo.		87.3%	87.2%	87.9%	78.7%
5. Retention Rate - 6 Mo.		84.8%	81.8%	82.7%	76.8%

The Nelson A. Rockefeller Institute of Government

Introduction

On February 8, 2005, the Rockefeller Institute held a public policy forum on the state and local role in performance management in New York State. The forum was co-sponsored by the Rockefeller Institute, the New York State Division of the Budget, and the Manhattan Institute. This introduction is organized bottom-up, beginning with the local level and then discussing the state and federal levels.

Speakers at the forum made me feel good. All six of the speakers presented constructive, upbeat reports on what they are doing. Their statements reflected a positive view of what can be done, and at the same time demonstrated a needed strong dose of realism on how hard it is to get good performance data that can influence state and local public management.

The speakers to a person stressed using performance management systems to monitor and ratchet up performance to achieve clear goals on a timely basis — not annually, but much more regularly (preferable on a monthly basis) — with extensive interaction between agency leaders and the managers of agency programs.

In the Dall Forsythe edited volume published by the Rockefeller Institute Press on performance management,[1] one of the major chapters on state and local performance management (of which there are several in this volume) is on the CompStat performance management system in New York City for the New York Police Department. Crime reduction is the main goal. Dennis Smith, who is a co-author with William Bratton of the chapter on CompStat, presented an update of this chapter and an appraisal of how other performance management systems, outgrowths of CompStat, are being implemented in New York City.

Swati Desai moderated the panel and presented a talk on how the JobStat system in New York City works to monitor and manage the performance of the City's 26 Job Centers for welfare and related human services. I have attended Thursday morning meetings on JobStat where the commissioner and his/her chief aides meet and interact with the heads of two of the City's job centers. I was, and continue to be, impressed by this demonstration of performance management in action — where it matters most, at the front lines.

Also at the afternoon session, Fred Wulczyn, a leader nationally on performance management for child welfare programs (foster care, adoptions, family preservation and abuse prevention), described his role in designing and helping to operate New York City's EQUIP system. This performance management system, which relies on techniques developed at the University of Chicago, has had extensive practical application. Because it has

1 *Quicker, Better, Cheaper?: Managing Performance in American Government*, edited by Dall W. Forsythe. See Appendix A for the Table of Contents. Go to http://www.rockinst.org/publications/ripress_books.html for more information.

been field-tested and operates with carefully scrubbed data, EQUIP is used for ranking and decision making about the sponsorship and funding of child welfare services.

All three systems — CompStat, JobStat, and EQUIP — have developed over time and operate in real time. They are success stories where success is most critical.

The morning session on state-level performance management was organized by the New York State Division of the Budget. The first speaker was Chauncey Parker, Director of New York's Division of Criminal Justice Services. He concentrated on New York City's CompStat system, praising its architect, Jack Maple, and noting that he had attended upwards of 150 CompStat review meetings. Parker stressed what he called "the three Ds" — Defining goals clearly, having timely accurate Data, and holding people accountable in well-organized Deliberation processes. He described the CrimeStat system his office has established to partner with 15 major urban counties in New York to create similar performance management systems, focused like CompStat on crime reduction.

The second state-level speaker was Robert Fleury, First Deputy Commissioner of the Office of General Services, assisted by Rebecca Meyers. An important contribution Fleury made was to emphasize the way the mission of an agency affects its goals and management system. The Office of General Services, he said, is "a decidedly operational organization that builds, fixes, and maintains state facilities." Its performance management system is necessarily inward looking — a tool for agency management.

In the discussion of Fleury's presentation, Edward Ingoldsby, Division of Budget Chief Budget Examiner, highlighted points brought out by Fleury. Ingoldsby noted that performance management works best "on an agency-by-agency basis where you have strong commissioner level support." He added that it is difficult "to link performance management with the formal budgeting system." Performance management is not well suited as a tool for budgeting. Doing this, he said, can undermine its efficacy as a management tool.

A good example of how hard it is to avoid problems in performance management if the budgetary stakes are tempting was brought out by John Reed, New York State Department of Civil Service. He cited a mis-specified goal for the sanitation system in New York City, the amount of refuse collected. Reed said, "they discovered they were hosing down the truck to increase the weight people were delivering."

The third speaker at the morning session was Andrew Eristoff, Commissioner of the New York State Department of Taxation and Finance. He previously served as both a City Council member and agency head at the local level in New York City, so he brought a multi-level perspective to the discussion. We use performance management "to manage our state-of-the-art taxpayer and collection call centers, to reduce waiting times, allocate resources, adjust hours, and match employee skills to caller issues." This, he said, is "embedded in our culture." Eristoff described the agency's "compliance continuum" and talked about the challenges involved in making such a system work well, which he said requires that it be "a continuing process." The latter point reflects an important generaliza-

tion — namely, that performance management has to be dynamic, with frequent adjustments of goals and measures to reflect changed conditions and policy preferences.

In the question-and-answer sessions, there was discussion about how agency executives can pull together and showcase performance management systems. The Mayor's Management Report in New York City was discussed — how it has been slimmed down over time, how it is sometimes viewed too much as a political document, and the reasons why governments have to be careful not to "over-integrate" and over simplify performance management conceptually and operationally.

Although it was not the subject of the forum, it is appropriate to add a discussion of the federal role in performance management. For both the federal and state role, my view is that their role should be primarily a leadership, catalytic, and teaching role, except for agencies where the federal government or the state has operating responsibility. (In the Forsythe volume, the chapter on performance management by the Social Security Administration is a demonstration of this point.)

Unfortunately, there is a strong tendency at the national level for the federal government to design and require the implementation of elaborate performance management systems that fail because they misunderstand the federalism terrain. Both the 1994 law passed by Congress, the Government Performance and Results Act (GPRA) and the results measurement system adopted by the Bush Administration, focused on what are called Program Assessment Rating Tools (PART), have this problem.[2]

The Bush administration frequently stresses "results" in budget documents, using PART scores to justify budget changes, which in the current fiscal environment are mostly expenditure reductions. This is unfortunate. For one thing, it can cause the kind of gaming and distortions that undermine the idea of smarter, stronger, data-driven management to improve program performance. For another, it misses a critical point. The fact that a program is underperforming doesn't mean its goals are unimportant. Maybe, to the contrary, the purposes involved are so important that more money is needed along with better managerial capability to carry them out. Performance management is best suited, as its name indicates, to managing performance. It is strongest and most useful if carried out at the level of operational responsibility.

When we decided to publish this report in hard copy, we asked all of the participants to work with us on editing their presentations and I thank them for doing so. Michael Cooper, Director of Publications, supervised the preparation of this report; Irene Pavone in my office worked with us to organize and review the material presented. I thank both of them for their help.

Richard P. Nathan

2 Richard P. Nathan "Presidential Address: 'Complexifying' Performance Oversight in America's Governments," *Journal of Policy Analysis and Management*, vol. 24, no. 2 (2005): 207-215. See Appendix B.

Appendix B: Presidential Address: "Complexifying" Performance Oversight in America's Governments

Presidential Address: "*Complexifying*" Performance Oversight in America's Governments

Richard P. Nathan

In keeping with my focus as President of APPAM on the "M" in our name, this talk deals with the performance management movement in American government.

Senator Daniel Patrick Moynihan, in a moment of frustration with me at a 1989 Senate hearing on welfare reform, said I am a "complexifier." In this spirit, the way I see it is that the performance management movement in American government is on the right track, but that it *oversimplifies*. I want to serve today in the role of a "complexifier" for the performance management movement. My aim is to be constructive—to suggest ways in which efforts to improve government performance can be reconciled with the pluralistic setting of U.S. public management.

Leaders in the federal government over the past 40 years have oversold simplistic systems for fulfilling public policy goals as expressed in the alphabet soup of systems like PPBS, MBO, ZBB, NPR, and GPRA.

The Alphabet Soup

PPBS stands for Lyndon Johnson's Planning-Programming-Budgeting System, adopted with much fanfare and based on private industry and Defense Department systems to assess and compare public spending options. MBO was Nixon's successor approach for Management by Objectives. ZBB was Carter's more radical initiative for zero-based budgeting to rank all spending options from the ground up in allocating government funds. NPR was Clinton's National Performance Review, which sought to focus government management and budgeting on achieving results. In 1993, a system for strategic results-based management and budgeting was enacted into law under the Government Performance and Results Act (GPRA).

For shock value, at the end of October of this presidential election year, I want to say something good about the newest alphabetically named performance management system. No matter what happens in the presidential election next week, I suggest this slogan, "*Let's not part with PART*," referring to the George W. Bush Administration's "Program Assessment Rating Tool" to compile effectiveness ratings for all federal programs.

THE PART SYSTEM

According to the FY 2005 budget, the Administration has "PART*ed*" (this verb form is used frequently in the budget) 400 programs representing 40 percent of the

Journal of Policy Analysis and Management, Vol. 24, No. 2, 207–215 (2005)
© 2005 by the Association for Public Policy Analysis and Management
Published by Wiley Periodicals, Inc. Published online in Wiley InterScience (www.interscience.wiley.com)
DOI. 10.1002/pam.20085

budget. Two hundred programs will be added in FY 2006. These evaluations are based on 30 questions to rate each program that is assessed in four areas: 20 percent for its purposes and design being clear, 10 percent for strategic planning, 20 percent for program management, and 50 percent for results. The ultimate goal is to evaluate the performance of all federal programs (over 1,000) in this way.

PART assessments draw one of five conclusions: effective, moderately effective, adequate, ineffective, or results not demonstrated. They are available online.[1] Illustrative ratings are for Head Start "results not demonstrated," for Medicare "moderately effective," and for the Community Development Block Grant (CDBG) "ineffective." Reading PART reports (most of them are about eight to ten pages long), one is struck both by their brevity and their variation. Some, for example on food stamps, are evenhanded and draw on a range of sources from within government and outside. Others are analytically weak.

DISCUSSION

The critical challenge for the PART system and efforts like it is setting performance goals. Where do goals come from? The answers vary. In some cases, it is wishful thinking in the form of political over-promising. In other cases, and these are the ones I care about, performance goals are based on research, policy analysis, and expertise. For this big category, I believe there should be broader consultation and collaboration involving policy officials, program managers, policy analysts, academic policy researchers, and public management experts.

In situations in which the results of definitive social science experiments based on random-assignment studies can be drawn upon, they materially aid policymakers by giving them a high level of confidence about impacts. *Such studies, however, are not available, and indeed could not be conducted, for all types of public programs and for all types of affected groups and policy conditions and needs.*

Terminology is important here. An *impact* means we can show that a particular public program caused something to happen that would not have happened otherwise. An *outcome* is the word customarily used to express a program's performance goals, regardless of whether we know if its activities are additive. Frequently, two types of people are involved: policy researchers, who generally favor experimental studies of program impacts, and management experts, who focus on outcomes. The two groups often have different mindsets and skill sets. As I see it, public policy and public management researchers (especially active members of APPAM among them) should contribute to knowledge about *both* impacts and outcomes.

Leaders in the public service are called upon in many (indeed, most) situations to set and periodically adjust outcome goals for performance management based on the most pertinent public policy knowledge available, and also drawing on expertise, experience, and observation. And yet even with the best of such efforts, governments often do not deal wisely with another critical challenge discussed in this paper, the need to devise politically acceptable and workable performance goals that can't be gamed for undesirable purposes.

While I think the PART system is on the right track in focusing on individual programs as the basic building blocks for assessing managerial performance, major problems are that it is too centralized, too insular, and not sufficiently discriminating. It does not adequately take into account the great differences that exist in the

[1] See www.whitehouse.gov/omb. Accessed September 7, 2004. See also Rodriguez, 2004, pp. 56–61.

size, importance, operational character, and settings of different public programs. I have a particular *federalism* problem in this context.

THE FEDERALISM CHALLENGE OF PERFORMANCE MANAGEMENT

Most domestic programs of the federal government operate by indirection through state and local governments that in many cases contract with nonprofit and private corporations to provide public services. It is beyond the scope of this talk to survey the backward and forward bounces of decentralization in American intergovernmental relations. Suffice to say, as Martha Derthick emphasizes, there are many ways in which members of the Congress and federal Executive Branch officials attempt to influence domestic affairs by adopting narrowly focused grant-in-aid programs and imposing and enforcing conditions, regulations, and guidelines on their operation. On the other hand, there is also a long history of devolutionary efforts to broaden grants (for example, by creating block grants), loosening or not enforcing regulations, granting waivers of federal requirements that are sought by state and local governments and interest groups, and generally by virtue of the fact that policymakers often disagree on public purposes and as a result adopt vague or even contradictory goals for domestic public policies.[2]

Wherever the devolution ball bounces, the point that stands out for me is that state and local governments have to be involved in assessing and improving program performance, especially under federal grant-in-aid programs that are broad gauged and have multiple purposes, as is so often the case. Federal agency officials should work with state officials in ways that are not heavy handed. They should adopt approaches that are continuous, user-friendly, candid, and appropriately intergovernmental.

This recommendation gets me into the consideration of differences between inputs, outputs, and intermediate and end outcomes as the goals of performance management.[3] It is arguable as a general principle that under broad-gauged federal grants-in-aid, the federal government should care most about *organizational outputs*, and in turn work with states to stimulate them to assess and help them in the best ways they can to define and measure the *outcomes for individual participants* of these federally aided domestic programs in the different environments in which they operate. This is what we actually do in more cases than is acknowledged.

Governmental programs have literally thousands of iterations. Reconciling two values, the flexibility (which is inherent to the diverse governmental environment of American federalism) with accountability (which policymakers and administrators should care about and achieve), requires that the PART system be *intergovernmentally sophisticated*. But this is not the only way the performance management movement needs to be complexified.

[2] For useful essays on this subject, see Derthick (2001) and also her other writings.

[3] See the attached definitions by Hatry, 2001, p. 19. This part of the discussion concerns "units of analysis," as well as measured goals. For some government programs, performance management goals involve *individuals* as the units of analysis (students, job seekers, sick people). For others, and for reasons that often reflect inherent limitations of measurement, we settle for (or may even decide we prefer) *organizational* goals as the units of analysis: Did service providers do what they were supposed to (for example, social security offices, state food stamp programs)? The subject of organizational performance is treated in Nathan (2000), chapters 6, 8–10, which describe institutional evaluations of the implementation of national welfare reform policies.

THE GAMING CHALLENGE OF PERFORMANCE MANAGEMENT

The Office of Management and Budget in its on-line documentation on the PART system says it is not primarily a budget tool. I agree with this idea. The PART system should not be sold solely as a method for deciding that such and such a program doesn't work so we should cut it, or that it does work and we should add resources. That is not to say that weak programs should be retained or that they should receive more resources. There are situations in which performance findings should influence budgeting. But if PART is used mainly for this purpose, the likelihood is that it will lose its managerial efficacy.

This possibility of gaming raises difficult questions for performance management involving how to set performance goals so that they influence agency behavior in the desired ways. Doing this requires both hard and soft accountability measures. In private industry, soft decision factors are often taken into account in rewarding managers. Economists have affirmed the wisdom of blending objective and subjective performance measures in this way (Baker, Gibbons, and Murphy, 1994). Unfortunately, however, the preponderance of attention and literature on managerial oversight in government has focused on rigid numeric goal setting.

In this intense 2004 pre-election period, it is not easy to envision the kind of consultative and interactive goal setting that could wisely and in evenhanded ways bring qualitative and policy related decision factors into play in performance oversight. This is made harder right now by cuts that have taken place in the managerial staffing of domestic programs, because overseeing such goals is necessarily labor intensive. Moreover, countering the rise in rating-mania by bringing qualitative and political variables into play in performance oversight is hard to prescribe. Yet, the truth is that, like speaking prose, we do it all the time, as suggested by some of the examples discussed next.

SOME EXAMPLES

In the field of employment and training, random-assignment demonstrations have shown that focusing on jobs is effective in aiding low-income, low-skilled people, especially women. However, performance-management goals focused on these dependent variables involving, for example, job placement and tenure can have problems. They can produce cream skimming, that is, selecting the most job-ready people to be aided, people who would have found jobs anyway. Corrections for this problem have been attempted, but can backfire. In one case I know of, the result was to undermine the public employment programs of the 1970s, leading to their eventual demise (Cook et al., 1985).

In the field of K–12 education, the emphasis on numeric goal achievement has caused unintended consequences that have required political jockeying and managerial recalibration. To paraphrase Tip O'Neil, all education is local. No matter how elegantly formed, the nature of the educational process requires that national oversight goals for local schools be focused on simple and general indicators, such that they tell only part of the story about school performance. For this reason it is important that performance goals for local schools be viewed as motivational every bit as much as being viewed as tools for the close and fulsome calibration of institutional effectiveness.

In the field of health policy, experts have urged the adoption of financial incentives for effective practice and quality enhancement, but at the same time acknowledge that the processes for providing and overseeing such standards face impedi-

ments (Epstein et al., 2004). Furthermore, we are told that such goals are unlikely to achieve their intended results unless substantial amounts of new money are provided (Epstein et al., 2004). Where is the money going to come from?

Citing these examples is not meant to suggest that setting performance goals and assessing how well they are achieved is unwise. The point is that performance management is difficult, can be expensive, and worst of all can backfire. It has to be smart and it has to be flexible, adaptive, and subtle. Stimulating efforts to ratchet up program performance on the part of the armies of governmental and quasi-governmental workers at every level of government who implement public policies requires setting and treating performance goals so that they serve *both* as targets for managers and as symbols for the public, in the latter sense, symbols that are well and widely understood and accepted. Raising student performance by using performance goals to focus attention on shared values about what our schools should teach has a good effect, and yet in the process of doing this we have learned about pitfalls of over-specification and have had to make adjustments. This is as it should be; undue rigidity in setting performance goals can undermine their effectiveness.

A good illustration of the latter point is shown by the recent success of welfare reform in reducing dependency and facilitating employment. The 1996 national welfare reform law dramatically changed bureaucratic behavior. Seemingly, this was rooted in strict and specific goals about jobs, hours of work, etc. But in reality, there was a big loophole, the "caseload reduction credit." This provision of the law, plus others, enabled states to advance their work-first purposes in a manner that reflected varied state and local values and conditions. The process was incremental and typically American. The success achieved was not the result of crafty planning by calculating policymakers. It came about serendipitously in ways that were surprising to us in our implementation studies of the 1996 national welfare reform law (Nathan and Gais, 1999; Gais et al., 2001, pp. 35–69).

MODIFYING THE PART SYSTEM

To reiterate, what is appealing to me about the PART system is that it focuses on individual programs as the basic building blocks, more so than on strategic and often overly elaborate purposes as advocated under previous federal government management reforms (Radin, 2000, pp. 111–135). What is needed, however, is a more candid and flexible treatment of goals in ways that involve a range of scholarly and expert perspectives.

Ann Blalock and Burt Barnow have called for a *partnership* for connecting demonstration and evaluation studies conducted by academic experts with performance management systems:

> Our recommendation is that competent evaluation research, or applied social science research, must be coordinated with or integrated within performance management systems if precise, valid, reliable information about social programs is to be made available to decisionmakers. (Blalock & Barnow, 2001, pp. 487–519)

Blalock and Barnow point out that what they call the "evaluation movement" was developed in the crucible of academia, while the "performance management movement" has its roots in public administration and in administrative bureaucracies. They believe, and I agree, that collaboration between these two movements would yield important benefits.

> We recommend that the major direction for the future is to coordinate evaluation research with performance management systems more fully, moving toward full integration of evaluations within performance management. Such integration will require that performance management systems treat evaluators not as aliens from outer space, who land only periodically to study and give advice, but as part of an interdisciplinary team. It will require that evaluators become more sensitized to managers' needs, to have ongoing information for tracking outcomes, and to express the benefits of their professional roots with greater humility. (Blalock & Barnow, 2001, pp. 487–519)

The two "movements" as described by Blalock and Barnow have different disciplinary bases. The impact/experimental movement is essentially Weberian in its assumptions about bureaucracy and implementation: Programs should be precisely controlled and replicated. By contrast, performance management seeks to establish goals as reference points for managers who are encouraged to innovate and change what they are doing in response to continuous feedback.

Just as there are many players in the bargaining processes for policymaking in American political pluralism, multiple players should be involved in deciding and adjusting the outcome goals for performance management.

Who should be the multiple players in outcome goal setting? Program managers at the appropriate levels of government (federal, state, and local) bring a needed perspective to bear about the likely effects of different performance goals in different settings. Researchers and policy analysts can also play an important role drawing on research that shows what programs are likely to work best and what their effects are when they do. But, it isn't enough to know if a program works. The responsible goal-setting officials also need to work their way through the hard questions raised earlier about how to express such findings in politically acceptable ways in performance goals that can't be gamed for undesirable purposes. Another important group of players is elected and appointed political officials; they clearly have a role to play in setting performance goals that reflect executive branch and legislative policies.

There is still another overarching aspect of this question about who should participate in performance oversight involving *auspices*. The GPRA law seeks the joint role of Congress and the Executive, whereas the Bush Administration's PART system gives the strong lead role to the Office of Management and Budget. While someone (that is, some agency) has to be in the lead, the challenge, as I see it, is not so much a challenge involving agency roles as one involving *transparency*. There needs to be a high level of transparency in sharing information about performance oversight. This is needed in the case of the PART system as it applies to other control agencies besides OMB, so that they can participate in performance oversight. The Congressional Budget Office, the General Accounting Office, and the Congressional Research Service, along with state and local budget and management offices and outside evaluators, need to have access to the underlying data that are used and understand how they are used in PART performance evaluations.

In line with the "complexifying" theme of my talk, this challenge of sharing data is made more difficult as more data are brought to bear, and more data *should be* brought to bear in performance management. An obvious and important opportunity is the availability of administrative data. We are virtually drowning in administrative data. Almost all domestic programs have reporting requirements that are extensive and detailed. Yet, despite this fact, and despite the fact that the quality of administrative data varies, not enough effort has been made to clean up these data, to scrutinize them and compile the best sources. Here, the underlying federalism

condition that states are different—that is, laboratories not just of democracy, but also of data availability—creates an environment in which multi-state studies can take advantage of opportunities to use administrative data wisely and more extensively.[4] Making fuller use of the best existing administrative data sets can have the important additional advantage of enabling evaluators to break out data for different groups of program participants in different kinds of program settings.

Another step that I believe would advance the art and practice of performance management would be to *relax* a little, loosen up the rating games. Instead of widely advertised *U.S. News and World Report*–type scorecards and report cards, people inside America's governments would do well to present performance results in more detailed and nuanced ways, explaining why and how they operate under different programs and conditions.

As a political scientist, I feel obliged at this juncture to go more deeply into my reason for putting so much emphasis in this talk on the dynamism, diversity, and fragmentation of the political milieu in which public services are provided. The reason is that so often in government *administration is policy*. Policymaking and administration are intrinsically linked. The day-to-day promulgation of rules and guidelines; the issuance of policy interpretations; the approval of grants; appointments and staffing decisions, all involve the transmission of values in ways that are more than routinely ministerial (Nathan, 1983).[5] The public administration literature is deficient in recognizing that policies and programs are changing all the time. Performance management must take this dynamism into account. It has to be seen as a *continuous* process. It has to be carried out by trial and error. It cannot be accomplished by fixed "one-size-fits-all" managerial formulas.

This is not to say that we should stop working hard on performance rating, only that these ratings should be presented in thorough and nuanced ways, and that they should be explained in statements that describe the strengths and limitations of the technology and data on which they are based.

It was not my intention in this talk to make specific recommendations for modifying the PART system. Moreover, my sense is that this should not be done in a rigid and formal way. Discussions about who is consulted and what kinds of inputs are used in setting performance are necessarily *situational*.[6]

Opening up the PART system (and state and local PART-like performance management systems) to more players and greater variations can also achieve something that is very practical and low tech. It can help deal with the information explosion in public affairs. We are virtually flooded with reports, e-mail messages, books, articles, conferences, and harangues. Performance management systems like PART, which pull together what is known about the effects of public programs, can inform and clarify debates about what America's governments do and the immense, diverse, and complex ways their actions affect social and economic conditions.

I realize that this plea for nuanced discourse, greater selectivity, flexibility, and transparency is wishful. "Complexifying" performance management in these ways

[4] Some large cities, notably New York City, have innovative performance management systems, which can provide valuable lessons and insights.

[5] On management tracking, I suggest Heinrich, 2002, pp. 712–725.

[6] The best book that shows how public management is *situational* is *Bureaucracy* by James Q. Wilson (2000 edition). Wilson describes organizations as needing "to acquire sufficient freedom of action." Based on his teaching and research, Wilson said he discovered that what was most needed was to permit an organization "to define its tasks as it saw best and to infuse the definition with a sense of mission." See chapter 2, "Organization Matters," pp. 14–28.

requires the application of judgment, and this invariably opens opportunities for political maneuver. Still, the fact of the matter is that we can never depoliticize performance management, nor should we even think of trying to do so. In the final analysis, what is needed is to bring more players into the process in ways that create a wise balance of expertise and politics. The long-term result would be to create a setting in which there is a more realistic focus on program results that gives citizens an honest, believable, and realistic portrayal of what their governments do and how well they do it.

I consulted with many experts with varied perspectives and experiences in writing this paper. Suggestions and comments were received from David Balducchi, Burt Barnow, Douglas Besharov, Patricia Billen, Jonathan Breul, Martha Derthick, Swati DeSai, Erik Devereux, Thomas Gais, William Grinker, Judy Gueron, Harry Hatry, Carolyn Heinrich, Michael Lipsky, Irene Lurie, Lawrence Lynn, Gerald Marschke, Lawrence Mead, Mark Nadel, Sonia Ospina, Beryl Radin, Justine Rodriguez, Frank Thompson, and Barry White. Irene Pavone worked ably and patiently with me in this consultation process. In the usual way, the ideas and points made in this talk are my responsibility alone.

RICHARD P. NATHAN, director, Nelson A. Rockefeller Institute of Government, State University of New York.

ATTACHMENT

Performance Management Definitions

* Inputs: Resources (that is, expenditures or employee time) used to produce outputs or outcomes.
* Outputs: Products and services delivered. Output refers to the completed products of internal activity: the amount of work done within the organization or by its contractors (such as a number of miles of road repaired or number of calls answered).
* Intermediate Outcomes: An outcome that is expected to lead to a desired end but is not an end in itself (such as service response time, which is of concern to the customer making a call but does not tell anything directly about the success of the call). A service may have multiple intermediate outcomes.
* End Outcomes: The end result that is sought (such as the community having clean streets or reduced incidence of crimes or fires). A service may have more than one end outcome.
* Efficiency, or Unit-Cost Ratio: The relationship between the amount of input (usually dollars or employee-years) and the amount of output or outcome of an activity or program. If the indicator uses outputs and not outcomes, a jurisdiction that lowers unit cost may achieve a measured increase in efficiency at the expense of the outcome of the service.
* Performance Indicator: A specific numerical measurement for each aspect of performance (for example, output or outcome) under consideration.

Source: Hatry, 1997, pp. 1–4.

REFERENCES

Baker, G., Gibbons, R., & Murphy, K.J. (1994). Subjective performance measures in optimal incentive contracts. The Quarterly Journal of Economic, 109(4) November, 1125–1156.

Blalock, A.B., & Barnow, B.S. (2001). Is the new obsession with performance management masking the truth about social programs? In D.W. Forsythe (Ed.), Quicker better cheaper: Managing performance in American government. Albany, NY: The Rockefeller Institute Press, pp. 487–519.

Cook, R.F., Adams, C.R., & Rawlins, V.L. (1985). Public service employment: The experience of a decade. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Derthick, M. (2001). Keeping the compound republic: Essays on American federalism. Washington, DC: The Brookings Institution Press.

Epstein, A.M., Lee, T.H., & Hamel, M.B. (2004). Paying physicians for high-quality care. The New England Journal of Medicine, 350(4), 406–410.

Gais, T.L., Nathan, Richard P., Lurie, Irene, & Kaplan, T. (2001). Implementation of the Personal Responsibility Act of 1996. In The World of Welfare, R., Blank and R. Haskins (Eds.), Washington, DC: Brookings Institution Press.

Hatry, H.P. (1997). Comparative performance measurement: FY 1996 data report (adapted). Washington, D.C.: International City/County Management Association.

Hatry, H.P. (2001). What types of performance information should be tracked. In D.W. Forsythe (Ed.), Quicker better cheaper: Managing performance in American government, Albany, NY: The Rockefeller Institute Press.

Heinrich, C.J. (2002). Outcome-based performance management in the public sector: implications for government accountability and effectiveness. Public Administration Review, 62(6), 712–725.

Nathan, R.P. (1983). The administrative presidency. New York, NY: Macmillan.

Nathan, R.P. (2000). Social science in government: The role of policy research. Albany, NY: The Rockefeller Institute Press (reprinted).

Nathan, R.P., & Gais, T.L. (1999). Implementing the Personal Responsibility Act of 1996: A first look. Albany, NY: The Nelson A. Rockefeller Institute of Government.

Radin, B.A., (2000). The Government Performance and Results Act and the tradition of federal management reform: Square pegs in round holes. Journal of Public Administration, Research and Theory, 10(1), 111–135.

Rodriguez, J.F. (2004). The arrival of performance budgeting. The business of government, pp. 56–61 (summer).

Wilson, J.Q. (2000). Bureaucracy. New York, NY: Basic Books.

FISCAL POLICY PROGRAM

www.ombwatch.org/budget

Appendix

Appendix to Testimony
Before the Subcommittee on Federal Financial
Management, Government Information,
and International Security of the
Senate Committee on
Homeland Security and Government Affairs

"The Tool Is a Work in Progress"

Limitations and Subjectivity

From GAO, Performance Budgeting: Observations on the Use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget, No. GAO-04-174 (Jan. 2004)

Appendix to Statement of

Adam Hughes
Director of Federal Fiscal Policy

United States General Accounting Office

GAO

Report to Congressional Requesters

January 2004

PERFORMANCE BUDGETING

Observations on the Use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget

GAO-04-174

January 2004

GAO
Accountability · Integrity · Reliability

Highlights

Highlights of GAO-04-174, a report to congressional requesters

PERFORMANCE BUDGETING

Observations on the Use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget

Why GAO Did This Study

The Office of Management and Budget's (OMB) Program Assessment Rating Tool (PART) is meant to provide a consistent approach to evaluating federal programs during budget formulation. To better understand its potential, congressional requesters asked GAO to examine (1) how PART changed OMB's fiscal year 2004 budget decision-making process, (2) PART's relationship to the Government Performance and Results Act of 1993 (GPRA), and (3) PART's strengths and weaknesses as an evaluation tool.

What GAO Recommends

GAO recommends that OMB (1) address the capacity demands of PART, (2) strengthen PART guidance, (3) address evaluation information availability and scope issues, (4) focus program selection on crosscutting comparisons and critical operations, (5) broaden the dialogue with congressional stakeholders, and (6) articulate and implement a complementary relationship between PART and GPRA.

OMB generally agreed with our findings, conclusions, and recommendations and stated that it is already taking actions to address many of our recommendations.

GAO also suggests that Congress consider the need for a structured approach to articulating its perspective and oversight agenda on performance goals and priorities for key programs.

www.gao.gov/cgi-bin/getrpt?GAO-04-174.

To view the full product, including the scope and methodology, click on the link above. For more information, contact Paul Posner at (202) 512-9573 or posnerp@gao.gov.

What GAO Found

PART helped structure OMB's use of performance information for its internal program and budget analysis, made the use of this information more transparent, and stimulated agency interest in budget and performance integration. OMB and agency staff said this helped OMB staff with varying levels of experience focus on similar issues.

Our analysis confirmed that one of PART's major impacts was its ability to highlight OMB's recommended changes in program management and design. Much of PART's potential value lies in the related program recommendations, but realizing these benefits requires sustained attention to implementation and oversight to determine if desired results are achieved. OMB needs to be cognizant of this as it considers capacity and workload issues in PART.

There are inherent challenges in assigning a single rating to programs having multiple purposes and goals. OMB devoted considerable effort to promoting consistent ratings, but challenges remain in addressing inconsistencies among OMB staff, such as interpreting PART guidance and defining acceptable measures. Limited credible evidence on results also constrained OMB's ability to rate program effectiveness, as evidenced by the almost 50 percent of programs rated "results not demonstrated."

PART is not well integrated with GPRA—the current statutory framework for strategic planning and reporting. 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. The PART/GPRA tension was further highlighted by challenges in defining a unit of analysis useful for both program-level budget analysis and agency planning purposes. Although PART can stimulate discussion on program-specific measurement issues, it cannot substitute for GPRA's focus on thematic goals and department- and governmentwide crosscutting comparisons. Moreover, PART does not currently evaluate similar programs together to facilitate trade-offs or make relative comparisons.

PART clearly must serve the President's interests. However, the many actors whose input is critical to decisions will not likely use performance information unless they feel it is credible and reflects a consensus on goals. It will be important for OMB to discuss timely with Congress the focus of PART assessments and clarify the results and limitations of PART and the underlying performance information. A more systematic congressional approach to providing its perspective on performance issues and goals could facilitate OMB's understanding of congressional priorities and thus increase PART's usefulness in budget deliberations.

Contents

Letter

Results in Brief	1
Background	4
OMB Used the PART to Systematically Assess Program Performance and Make Results Known, but Follow-up on PART Recommendations Is Uncertain	8
Despite OMB's Considerable Efforts to Create a Credible Evaluation Tool, PART Assessments Require Judgment and Were Constrained by Data Limitations	11
Subjective Terms and a Restrictive Format Contributed to Subjective and Inconsistent Responses	17
There Were Inconsistencies in Defining Acceptable Measures and in Logically Responding to Question "Pairs"	20
Disagreements on Performance Information Led to Creation of a "Results Not Demonstrated" Category	21
The Fiscal Year 2004 PART Process Was a Parallel, Competing Approach to GPRA's Performance Management Framework	25
Defining a "Unit of Analysis" That Is Useful for Program-Level Budget Analysis and Agency Planning Purposes Presents Challenges	27
Conclusions and General Observations	29
Matter for Congressional Consideration	33
Recommendations for Executive Action	36
Agency Comments	36
	37

Appendixes

Appendix I:	**Scope and Methodology**	39
Appendix II:	**The Fiscal Year 2004 PART and Differences Between the Fiscal Year 2004 and 2005 PARTs**	48
	Section I: Program Purpose & Design (Yes, No, N/A)	48
	Section II: Strategic Planning (Yes, No, N/A)	48
	Section III: Program Management (Yes, No, N/A)	50
	Section IV: Program Results (Yes, Large Extent, Small Extent, No)	53
Appendix III:	**Development of PART**	62
	Fiscal Year 2003	62
	Fiscal Year 2004	62
	Fiscal Year 2005	64
Appendix IV:	**Comments from the Office of Management and Budget**	65

Appendix V: GAO Contacts and Staff Acknowledgments 67
GAO Contacts 67
Acknowledgments 67

Tables

Table 1: Overview of Sections of PART Questions 9
Table 2: Overview of PART Program Types 10
Table 3: The Effect of Overall PART Score on Proposed Funding Changes (Discretionary Programs) 43
Table 4: The Effect of Overall PART Score on Proposed Funding Changes (Small Discretionary Programs) 44
Table 5: The Effect of Overall PART Score on Proposed Funding Changes (Medium-Size Discretionary Programs) 44
Table 6: The Effect of Overall PART Score on Proposed Funding Changes (Large Discretionary Programs) 44
Table 7: The Effect of PART Component Scores on Proposed Funding Changes (All Discretionary Programs) 45
Table 8: The Effect of PART Component Scores on Proposed Funding Changes (Small Discretionary Programs) 45
Table 9: Side-by-Side of the Fiscal Year 2005 PART and the Fiscal Year 2004 PART Questions 54

Figures

Figure 1: Fiscal Year 2004 PART Recommendations 13
Figure 2: Number of Discretionary PART Programs by Rating and Funding Result, Fiscal Years 2003-2004 14
Figure 3: The PART Process and Budget Formulation Timelines 64

This is a work of the U.S. government and is not subject to copyright protection in the United States. It may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.

United States General Accounting Office
Washington, D.C. 20548

January 30, 2004

The Honorable George V. Voinovich
Chairman
Subcommittee on Oversight of Government Management, the Federal Workforce and the District of Columbia
Committee on Governmental Affairs
United States Senate

The Honorable Todd R. Platts
Chairman
Subcommittee on Government Efficiency and Financial Management
Committee on Government Reform
House of Representatives

The Honorable Sam Brownback
United States Senate

The Honorable Todd Tiahrt
House of Representatives

Since the 1950s, the federal government has attempted several governmentwide initiatives designed to better align spending decisions with expected performance—what is often commonly referred to as "performance budgeting." Consensus exists that prior efforts—including the Hoover Commission, the Planning-Programming-Budgeting-System (PPBS), Management by Objectives, and Zero-Based Budgeting (ZBB)—failed to significantly shift the focus of the federal budget process from its long-standing concentration on the items of government spending to the results of its programs.

In the 1990s, Congress and the executive branch laid out a statutory and management framework that provides the foundation for strengthening government performance and accountability, with the Government Performance and Results Act of 1993[1] (GPRA) as its centerpiece. GPRA is designed to inform congressional and executive decision making by providing objective information on the relative effectiveness and efficiency of federal programs and spending. A key purpose of the act is to create closer and clearer links between the process of allocating scarce resources

[1] Pub. L. No. 103-62 (1993).

and the expected results to be achieved with those resources. This type of integration is critical, as we have learned from prior initiatives that failed in part because they did not prove to be relevant to budget decision makers in the executive branch or Congress.[2] GPRA requires not only a connection to the structures used in congressional budget presentations but also consultation between the executive and legislative branches on agency strategic plans, which gives Congress an oversight stake in GPRA's success.[3]

In its overall structure, focus, and approach GPRA incorporates two critical lessons learned from previous reforms. First, any approach designed to link plans and budgets—that is, to link the responsibility of the executive to define strategies and approaches with the legislative "power of the purse"—must explicitly involve both branches of our government. PPBS and ZBB failed in part because performance plans and measures were developed in isolation from congressional oversight and resource allocation processes.

Second, the concept of performance budgeting has and likely will continue to evolve. Thus, no single definition of performance budgeting encompasses the range of past and present needs and interests of federal decision makers. The need for multiple definitions reflects the differences in the roles various participants play in the budget process. And, given the complexity and breadth of the federal budget process, performance budgeting must encompass a variety of perspectives in its efforts to link resources with results.

This administration has made the integration of performance and budget information one of five governmentwide management priorities under its President's Management Agenda (PMA).[4] A central element in this initiative is the Office of Management and Budget's (OMB) Program Assessment Rating Tool (PART) that OMB describes as a diagnostic tool meant to provide a consistent approach to evaluating federal programs as part of the

[2] U.S. General Accounting Office, *Performance Budgeting: Past Initiatives Offer Insights for GPRA Implementation*, GAO/AIMD-97-46 (Washington, D.C.: Mar. 27, 1997).

[3] See Pub. L. No. 103-62, § 2 (1993), 5 U.S.C. § 306 (2003), and 31 U.S.C. §§ 1115-1116 (2003).

[4] In addition to budget and performance integration, the other four priorities under the PMA are strategic management of human capital, expanded electronic government, improved financial performance, and competitive sourcing.

executive budget formulation process. The PART is the latest iteration of 50 years of federal performance budgeting initiatives. It applies 25 questions to all "programs"[5] under four broad topics: (1) program purpose and design, (2) strategic planning, (3) program management, and (4) program results (i.e., whether a program is meeting its long-term and annual goals) as well as additional questions that are specific to one of seven mechanisms or approaches used to deliver the program.[6]

To better understand the PART's potential as a mechanism for assessing program goals and results, you asked us to examine (1) how the PART changed OMB's decision-making process in developing the President's fiscal year 2004 budget request; (2) the PART's relationship to the GPRA planning process and reporting requirements; and (3) the PART's strengths and weaknesses as an evaluation tool, including how OMB ensured that the PART was applied consistently.

To respond to your request, we reviewed OMB materials on the development and implementation of the PART as well as the results produced by the PART assessments. To assess consistency of the PART's application, we performed analyses of OMB data from the PART program summary and assessment worksheets for each of the 234 programs OMB reviewed for fiscal year 2004, including a statistical analysis of the relationship between the PART scores and funding levels in the President's Budget. We also identified several sets of similar programs that we examined more closely to determine if comparable or disparate criteria were applied in producing the PART results for these clusters of programs. We reviewed 28 programs in nine clusters covering food safety, water supply, military equipment procurement, provision of health care, statistical agencies, block grants to assist vulnerable populations, energy research programs, wildland fire management, and disability compensation. We also interviewed OMB officials regarding their experiences with the PART in the fiscal year 2004 budget process.

[5] There is no standard definition for the term "program." For purposes of PART, OMB described the unit of analysis (program) as (1) an activity or set of activities clearly recognized as a program by the public, OMB, and/or Congress; (2) having a discrete level of funding clearly associated with it; and (3) corresponding to the level at which budget decisions are made.

[6] The seven major categories are competitive grants, block/formula grants, capital assets and service acquisition programs, credit programs, regulatory-based programs, direct federal programs, and research and development programs. Tax programs were not addressed for the fiscal year 2004 PART process.

As part of our examination of the usefulness of the PART as an evaluation tool and also to obtain agency perspectives on the relationship between PART and GPRA, we interviewed department and agency officials, including senior managers, and program, planning, and budget staffs at (1) the Department of Health and Human Services (HHS), (2) the Department of Energy (DOE), and (3) the Department of the Interior (DOI). We selected these three departments because they had a variety of program types (e.g., block/formula grants, competitive grants, direct federal, and research and development) that were subject to the PART and could provide a broad-based perspective on how the PART was applied to different programs. With the exception of our summary analyses of all 234 programs, the information obtained from OMB and agency officials and our review of selected programs is not generalizable to the PART process for all 234 programs. However, the consistency and frequency with which similar issues were raised by OMB and agency officials suggest that our review reliably captures several significant and salient aspects of the PART as a budget and evaluation tool.

Our review focused on the fiscal year 2004 PART process. We conducted our work from May 2003 through October 2003 in accordance with generally accepted government auditing standards. Detailed information on our scope and methodology appears in appendix I. OMB provided written comments on a draft of this report that are reprinted in appendix IV.

Results in Brief

The PART has helped to structure and discipline OMB's use of performance information for its internal program analysis and budget review, made the use of this information more transparent, and stimulated agency interest in budget and performance integration. Both OMB and agency staff noted that this helped ensure that OMB staff with varying levels of experience focused on the same issues, fostering a more disciplined approach to discussing program performance with agencies. Several agency officials also told us that the PART was a catalyst for bringing agency budget, planning, and program staff together since none could fully respond to the PART questionnaire alone.

Our analysis confirmed that one of the PART's major impacts was its ability to highlight OMB's recommended changes in program management and design. Over 80 percent of the recommendations made for the 234 programs assessed for the fiscal year 2004 budget process were for improvements in program design, assessment, and program management; less than 20 percent were related to funding issues. As OMB and others

recognize, performance is not the only factor in funding decisions. Determining priorities—including funding priorities—is a function of competing values and interests. Although OMB generally proposed to increase funding for programs that received ratings of "effective" or "moderately effective" and proposed to cut funding for those programs that were rated "ineffective," our review confirmed OMB's statements that funding decisions were not applied mechanistically. That is, for some programs rated "effective" or "moderately effective" OMB recommended funding decreases, while for several programs judged to be "ineffective" OMB recommended additional funding in the President's budget request with which to implement changes.

Much of the potential value of the PART lies in the related program recommendations and associated improvements, but realization of these benefits will require sustained attention to implementation and oversight in order to determine if the desired results are being achieved. Such attention and oversight takes time, and OMB needs to be cognizant of this as it considers the capacity and workload issues in the PART. Currently OMB plans to assess an additional 20 percent of all federal programs annually. Each year, the number of recommendations from previous years' evaluations will grow—and a system for monitoring their implementation will become more critical. OMB encouraged its Resource Management Offices (RMO) to consider many factors in selecting programs for the fiscal year 2004 PART assessments, such as continuing presidential initiatives and programs up for reauthorization. While all programs would eventually be reviewed over the 5-year period, selecting related programs for review in a given year would enable decision makers to analyze the relative efficacy of similar programs in meeting common or similar outcomes. We recommend that OMB centrally monitor and report on agency implementation and progress on PART recommendations to provide a governmentwide picture of progress and a consolidated view of OMB's workload in this area. In addition, to target scarce analytic resources and to focus decision makers' attention on the most pressing policy issues, we recommend that OMB reconsider plans for 100 percent coverage of federal programs by targeting PART assessments based on such factors as the relative priorities, costs, and risks associated with related clusters of programs and activities. We further recommend that OMB select for review in the same year related or similar programs or activities to facilitate such comparisons and trade-offs.

Developing a credible evidence-based rating tool to provide bottom-line ratings for programs was a major impetus in developing the PART.

However, inherent challenges exist in assigning a single "rating" to programs that often have multiple purposes and goals. Despite the considerable time and effort OMB has devoted to promoting consistent application of the PART, the tool is a work in progress. Additional guidance and considerable revisions are needed to meet OMB's goal of an objective, evidence-based assessment tool. In addition to difficulties with the tool itself—such as subjective terminology and a restrictive yes/no format—providing flexibility to assess multidimensional programs with multiple purposes and goals often implemented through multiple actors has led to a reliance on OMB staff judgments to apply general principles to specific cases. OMB staff were not fully consistent in interpreting the guidance for complex PART questions and in defining acceptable measures. In addition, the limited availability of credible evidence on program results also constrained OMB staff's ability to use the PART to rate programs' effectiveness. Almost 50 percent of the 234 programs assessed for fiscal year 2004 received a rating of "results not demonstrated" because OMB decided that program performance information, performance measures, or both were insufficient or inadequate. OMB, recognizing many of the limitations with the PART, modified the PART for fiscal year 2005 based on lessons learned during the fiscal year 2004 process, but issues remain. We therefore recommend that OMB continue to improve the PART guidance by (1) clarifying when output versus outcome measures are acceptable and (2) better defining an "independent, quality evaluation." We further recommend that OMB both clarify its expectations regarding the nature, timing, and amount of evaluation information it wants from agencies for the purposes of the PART and consider using internal agency evaluations as evidence on a case-by-case basis.

The PART is not well integrated with GPRA—the current statutory framework for strategic planning and reporting. According to OMB officials, GPRA plans were organized at too high a level to be meaningful for program-level budget decision making. To provide decision makers with program-specific, outcome-based performance data useful for executive budget formulation, OMB has stated its intention to modify GPRA goals and measures with those developed under the PART. As a result, OMB's judgment about appropriate goals and measures is substituted for GPRA judgments based on a community of stakeholder interests. Agency officials

we spoke with expressed confusion about the relationship between GPRA requirements and the PART process. Many view PART's program-by-program focus and the substitution of program measures as detrimental to their GPRA planning and reporting processes. OMB's effort to influence program goals is further evident in recent OMB Circular A-11 guidance[7] that clearly requires each agency to submit a performance budget for fiscal year 2005, which will replace the annual GPRA performance plan.

The tension between PART and GPRA was further highlighted by the challenges in defining a unit of analysis that is useful both for program-level budget analysis and agency planning purposes. Although the PART reviews indicated to OMB that GPRA measures are often not sufficient to help it make judgments about programs, the different units of analysis used in these two performance initiatives contributed to this outcome. For the PART, OMB created units of analysis that tied to discrete funding levels by both disaggregating and aggregating certain programs. In some cases, disaggregating programs for the PART reviews ignored the interdependency of programs by artificially isolating them from the larger contexts in which they operate. Conversely, in other cases in which OMB aggregated programs with diverse missions and outcomes for the PART reviews, it became difficult to settle on a single measure (or set of measures) that accurately captured the multiple missions of these diverse components. Both of these "unit of analysis" issues contributed to the lack of available planning and performance information.

Although the PART can stimulate discussion on program-specific performance measurement issues, it is not a substitute for GPRA's strategic, longer-term focus on thematic goals and department- and governmentwide crosscutting comparisons. GPRA is a broad legislative framework that was designed to be consultative with Congress and other stakeholders and allows for varying uses of performance information, while the PART applies evaluation information to support decisions and program reviews during the executive budget formulation process. Moreover, GPRA can anchor the review of programs by providing an overall strategic context for programs' contributions toward agency goals. We therefore recommend that OMB seek to achieve the greatest benefit from both GPRA and PART by articulating and implementing an integrated, complementary relationship between the two. We further recommend that OMB continue to improve the PART guidance by expanding the discussion

[7] OMB Circular A-11, *Preparation, Submission, and Execution of the Budget*, Section 220.

of how programs—also known as "units of analysis"—are determined, including recognizing the trade-offs, implications, or both of such determinations.

As part of the President's budget preparation, the PART clearly must serve the President's interests. However, experience suggests that efforts to integrate budget and performance are promoted when Congress and other key stakeholders have confidence in the credibility of the analysis and the process used. It is unlikely that the broad range of players whose input is critical to decisions will use performance information unless they believe it is relevant, credible, reliable, and reflective of a consensus about performance goals among a community of interested parties. Similarly, the measures used to demonstrate progress toward a goal, no matter how worthwhile, cannot appear to serve a single set of interests without potentially discouraging use of this information by others. We therefore recommend that OMB attempt to build on the strengths of GPRA and PART by seeking to communicate early in the PART process with congressional appropriators and authorizers about what performance issues and information are most important to them in evaluating programs. Furthermore, while Congress has a number of opportunities to provide its perspective on performance issues and goals through its authorization, oversight, and appropriations processes, we suggest that Congress consider the need for a more structured approach for sharing with the executive branch its perspective on governmentwide performance matters, including its views on performance goals and outcomes for key programs and the oversight agenda.

In commenting on a draft of this report, OMB generally agreed with our findings, conclusions, and recommendations. OMB outlined actions it is taking to address many of our recommendations, including refining the process for monitoring agencies' progress in implementing the PART recommendations, seeking opportunities for dialogue with Congress on agencies' performance, and continuing to improve executive branch implementation of GPRA plans and reports. OMB also suggested some technical changes throughout the report that we have incorporated as appropriate. OMB's comments appear in appendix IV. We also received technical comments on excerpts of the draft provided to the Departments of the Interior, Energy, and Health and Human Services, which are incorporated as appropriate.

Background

The current administration has taken several steps to strengthen and further performance-resource linkages for which GPRA laid the groundwork. Central to the budget and performance integration initiative, the PART is meant to strengthen the process for assessing the effectiveness of programs by making that process more robust, transparent, and systematic. As noted above, the PART is a series of diagnostic questions designed to provide a consistent approach to rating federal programs. (See app. II for a reproduction of the PART.) Drawing on available performance and evaluation information, the questionnaire attempts to determine the strengths and weaknesses of federal programs with a particular focus on individual program results. The PART asks, for example, whether a program's long-term goals are specific, ambitious, and focused on outcomes, and whether annual goals demonstrate progress toward achieving long-term goals. It is designed to be evidence based, drawing on a wide array of information, including authorizing legislation, GPRA strategic plans and performance plans and reports, financial statements, inspector general and GAO reports, and independent program evaluations. PART questions are divided into four sections; each section is given a specific weight in determining the final numerical rating for a program. Table 1 shows an overview of the four PART sections and the weights OMB assigned.

Table 1: Overview of Sections of PART Questions

Section	Description	Weight
I. Program Purpose and Design	To assess whether • the purpose is clear, and • the program design makes sense.	20%
II. Strategic Planning	To assess whether the agency sets valid programmatic • annual goals, and • long-term goals.	10%
III. Program Management	To rate agency management of the program, including • financial oversight, and • program improvement efforts.	20%
IV. Program Results/Accountability	To rate program performance on goals reviewed in • the strategic planning section, and • through other evaluations.	50%

Source: GAO analysis of the *Budget of the United States Government, Fiscal Year 2004, Performance and Management Assessments* (Washington, D.C.: February 2003)

In addition, each PART program is assessed according to one of seven approaches to service delivery. Table 2 provides an overview of these program types and the number and percentage of programs covered by each type in the fiscal year 2004 President's Budget performance assessments.

Table 2: Overview of PART Program Types

Program type	Description	Number/percentage of programs[a]
1. Direct federal	Programs in which support and services are provided primarily by federal employees.	67 29%
2. Block/formula grant	Programs that distribute funds to state, local, and tribal governments and other entities by formula or block grant.	41 18%
3. Competitive grant	Programs that distribute funds to state, local, and tribal governments, organizations, individuals, and other entities through a competitive process.	37 16%
4. Capital assets and service acquisition	Programs in which the primary means to achieve goals is the development and acquisition of capital assets (such as land, structures, equipment, and intellectual property) or the purchase of services (such as maintenance and information technology) from a commercial source.	34 15%
5. Research and development	Programs that focus on creating knowledge or applying it toward the creation of systems, devices, methods, materials, or technologies.	32 14%
6. Regulatory-based	Programs that employ regulatory action to achieve program and agency goals through rule making that implements, interprets, or prescribes law or policy, or describes procedure or practice requirements. These programs issue significant regulations, which are subject to OMB review.	15 6%
7. Credit	Programs that provide support through loans, loan guarantees, and direct credit.	4 2%
8. Mixed[b]	Programs that contain elements of different program types.	4 2%

Source. GAO summary and analysis of the *Budget of the United States Government, Fiscal Year 2004, Performance and Management Assessments* (Washington, D.C.: February 2003)

[a]Percentages do not add to 100 percent due to rounding.

[b]OMB noted that in rare cases, drawing questions from two of the seven PART program types—that is, creation of a "mixed" program type—yields a more informative assessment.

During the fiscal year 2004 budget cycle, OMB applied the PART to 234 programs (about 20 percent of the fiscal year 2004 President's Budget request to Congress[8]), and gave each program one of four overall ratings: (1) "effective," (2) "moderately effective," (3) "adequate," or (4) "ineffective" based on program design, strategic planning, management, and results. A fifth rating, "results not demonstrated," was given—independent of a program's numerical score—if OMB decided that a program's performance information, performance measures, or both were insufficient or inadequate. The administration plans to assess an additional 20 percent of the budget each year until the entire executive branch has been reviewed. For more information on the development of the PART, see appendix III.

OMB Used the PART to Systematically Assess Program Performance and Make Results Known, but Follow-up on PART Recommendations Is Uncertain

The PART clarified OMB's use of performance information in its budget decision-making process and stimulated new interest in budget and performance integration. OMB generally proposed budget increases for programs that received ratings of "effective" or "moderately effective" and decreased funding requests for those programs that were rated "ineffective," but there were clear exceptions. Moreover, the more important role of the PART was not in making resource decisions but in its support for recommendations to improve program design, assessment, and management. OMB's ability to use the PART to identify and address future program improvements and measure progress—a major purpose of the PART—is predicated on its ability to oversee the implementation of PART recommendations. However, it is not clear that OMB has a centralized system to oversee the implementation of such recommendations or evaluate their effectiveness.

The PART Made Budget and Performance Integration at OMB More Transparent

The PART helped structure and discipline the use of performance information in the budget process and made the use of such information more transparent throughout the executive branch. According to OMB senior officials and many of the examiners and branch chiefs, the PART lent structure to a process that had previously been informal and gave OMB staff a systematic way of asking performance-related questions. Both

[8] OMB defined 20 percent of the budget as either 20 percent of programs or their funding levels so long as all programs are assessed over the 5-year cycle for fiscal years 2004 through 2008 budget requests.

agency and OMB staff noted that this helped ensure that OMB staff with varying levels of experience focused on the same issues, fostering a more disciplined approach to discussing performance within OMB and with agencies. Agency officials told us that by encouraging more communication between departments and OMB, the PART helps illuminate both how OMB makes budget decisions and how OMB staff think about program management. The PART also provided a framework for raising performance issues during the OMB Director's Reviews. OMB managers and staff reported that it led to richer discussions on what a program should be achieving, whether the program was performing effectively, and how program performance could be improved.

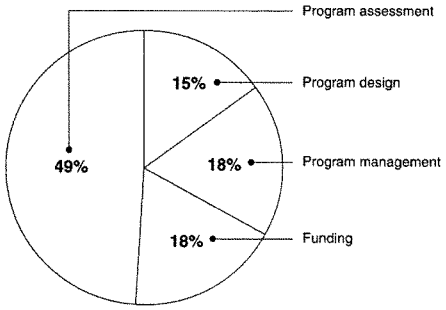
Agencies also reported that the PART process expanded the dialogue between program, planning, and budget staffs, and stimulated interest in budget and performance integration. Several agency officials stated that the PART worksheets were a catalyst for bringing staffs together since none could fully respond to the questionnaire alone. OMB and agency officials agreed that the PART led to more interactions between OMB and agency program and planning staff and, in turn, increased program managers' awareness of and involvement in the budget process. According to OMB and several agency officials, the PART process—that is, responding to the PART questionnaire—involved staff outside of the performance management area. Additionally, both agency and OMB officials said that the attention given to programs that were not routinely reviewed was a positive benefit of the PART process.

Use of Performance Information Was Evident in OMB's Recommendations

OMB senior officials told us that one of the PART's most notable impacts was its ability to highlight OMB's recommended changes in program management and design. As shown in figure 1, we found that 82 percent of PART recommendations addressed program assessment, design, and management issues; only 18 percent of the recommendations had a direct link to funding matters.[9]

[9] The 234 programs assessed for fiscal year 2004 contained a total of 612 recommendations.

Figure 1: Fiscal Year 2004 PART Recommendations



Source. GAO analysis of OMB data.

The majority of recommendations relate to changes that go well beyond funding consideration for one budget cycle. For example, OMB and HHS officials agree that the Foster Care program as it is currently designed does not provide appropriate incentives for the permanent placement of children; the program financially rewards states for keeping children in foster care instead of the original intent of providing temporary, safe, and appropriate homes for abused or neglected children until children can be returned to their families or other permanent arrangements can be made. The PART assessment provided support for OMB's recommendation that legislation be introduced that would create an option for states to participate in an alternate financing program that would "better meet the needs of each state's foster care population."

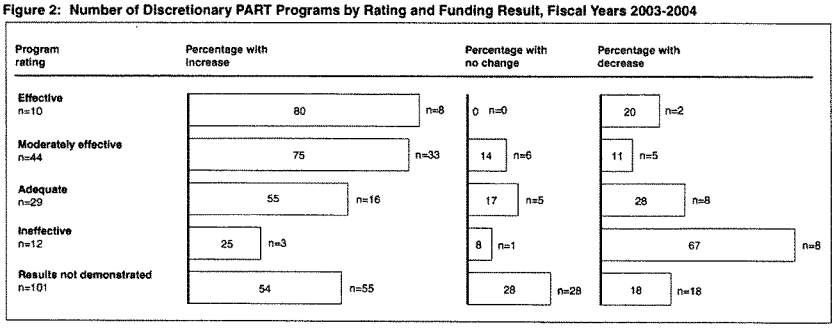
Performance information included in the PART for the Department of Labor's (DOL) Community Service Employment for Older Americans program helped to shape OMB's recommendation to increase competition for the grants. OMB concluded that although the Older Americans Act of 2000 amendments[10] authorize competition for grants in cases in which grantees repeatedly fail to perform, the programs' 10 national grantees

[10] Pub. L. No. 106-501 (2000).

have historically been the sole recipients of grant funds regardless of performance. OMB recommended that DOL award national grants competitively to strengthen service delivery and open the door to new grantees.

As OMB and others recognize, performance is not the only factor in funding decisions. Determining priorities—including funding priorities—is a function of competing values and interests. As seen in figure 2, we found that PART scores were generally positively related to proposed funding changes in discretionary programs but not in a mechanistic way. In other words, PART scores did not automatically determine funding changes. OMB proposed funding increases for most of the programs rated "effective" or "moderately effective" and proposed funding decreases for most of the programs rated "ineffective," but there were clear exceptions. Programs rated as "results not demonstrated"—which reflected a range of PART scores—had mixed results.

Figure 2: Number of Discretionary PART Programs by Rating and Funding Result, Fiscal Years 2003-2004



Source GAO analysis of OMB data

Note: Discretionary programs refer to those programs with budgetary resources provided in appropriation acts. Because Congress controls spending for mandatory programs—generally entitlement programs such as food stamps, Medicare, and veterans' pensions—indirectly rather than directly through the appropriations process, we excluded them from our analysis. Of the 234 programs, we could not classify 11 as being either predominantly mandatory or discretionary; these programs are excluded from our analysis as well, and are listed in appendix I.

A large portion of the variability in proposed budget changes could not be explained by the quantitative measures reported by the PART. Regressions of PART scores never explained more than about 15 percent of the proposed budget changes. For only the one-third of discretionary programs with the smallest budgets, we found that the composite PART scores had a modest but statistically significant effect on proposed budget changes (measured in percentage change) between fiscal years 2003 and 2004. For a fuller discussion of the statistical methods used, see appendix I.

The relationship between performance levels and budget decisions was not one-dimensional. For example, OMB rated the Department of Defense's Basic Research program as "effective," but recommended a reduction in congressionally earmarked projects that it stated did not meet the program's merit review process. OMB also recommended reducing funding for DOE's International Nuclear Materials Protection and Cooperation program (rated "effective") because difficulties in obtaining international agreements had resulted in the availability of sufficient unobligated balances[11] to make new funding unnecessary. However, OMB sometimes proposed funding increases for programs that were rated "ineffective" to implement improvement plans that had been developed, such as the Internal Revenue Service's new Earned Income Tax Credit compliance initiatives and DOE's revised environmental cleanup plans for its Environmental Management (Cleanup) program.

Capacity Issues Could Affect OMB's Ability to Use the PART to Drive Program Improvements

OMB has said that a major purpose of the PART is to focus on program improvements and measure progress. Effectively implementing PART recommendations aimed at program improvements will require sustained attention and sufficient oversight of agencies to ensure that the recommendations are producing desirable results. However, each year, the number of recommendations from previous years' evaluations will grow. Currently, OMB plans to assess an additional 20 percent of all federal programs annually such that all programs would eventually be reviewed over a 5-year period. OMB encouraged its RMOs to consider a variety of factors in selecting programs for the fiscal year 2004 PART assessments, including continuing presidential initiatives and programs up for reauthorization. Strengthening the focus on selecting related programs for review in a given year would enable decision makers to analyze the relative

[11] Unobligated balances are defined as portions of available budget authority that the agency has not set aside to cover current legal liabilities.

efficacy of similar programs in meeting common or similar outcomes. As our work has shown, unfocused and uncoordinated programs waste scarce funds, confuse and frustrate program customers, and limit overall program effectiveness. Therefore it is prudent to highlight crosscutting program efforts and clearly relate and address the contributions of alternative federal strategies toward meeting similar goals.

Although OMB has created a template for agencies to report on the status of their recommendations and has reported that agencies are implementing their PART recommendations, OMB has no central system for monitoring agency progress or evaluating the effectiveness of changes. While RMOs are responsible for overseeing agency progress, OMB senior managers will not have a comprehensive governmentwide picture of progress on the implementation of PART recommendations, nor will they have a complete picture of OMB's workload in this area. As OMB has recognized, following through on the recommendations is essential for improving program performance and ensuring accountability.

Senior OMB managers readily recognized the increased workload the PART placed on examiners—in one public forum we attended, a senior OMB official described many examiners as being very concerned about the additional workload. However, OMB expects the workload to decline as OMB and agency staff become more familiar with the PART tool and process, and as issues with the timing of the PART reviews are resolved. Agency officials told us that originally, there was no formal guidance for reassessing PART programs—it varied by RMO. When issued, OMB's formal PART guidance limited reassessments to (1) updating the status/implementation of recommendations from the fiscal year 2004 PART and (2) revisiting specific questions for which new evidence exists. OMB expected that in most reassessments, only those questions in which change could be demonstrated would be "reopened." OMB officials acknowledged that this formal guidance is at least partly due to resource constraints.

OMB staff were divided on whether the PART assessments made an appreciable difference in time spent on its budget review process. Many of those we spoke with told us that their workloads during the traditional budget season have always been heavy and that PART did not add significantly to their work, especially since the PART generally formalized a process already taking place. Those who did acknowledge workload concerns said that they were surprised at the amount of time it was taking to reassess programs. In fact, more than one OMB official told us that

reassessing programs was taking almost as long as brand-new assessments, despite the fact that OMB scaled back the scope of these reassessments.

Despite OMB's Considerable Efforts to Create a Credible Evaluation Tool, PART Assessments Require Judgment and Were Constrained by Data Limitations

OMB went to great lengths to encourage consistent application of the PART in the evaluation of government programs, including pilot testing the instrument, issuing detailed guidance, and conducting consistency reviews. However, while the instrument can undoubtedly be improved, any tool that is sophisticated enough to take into account the complexity of the U.S. government will always require OMB staff to exercise interpretation and judgment. Providing flexibility to assess multidimensional programs with multiple purposes and impacts has led to a reliance on OMB staff judgments to apply general principles to specific cases. Accordingly, OMB staff were not fully consistent in interpreting complex questions about agency goals and results. In addition, the limited availability of credible evidence on program results also constrained OMB's ability to use the PART to rate programs' effectiveness.

Inherent Performance Measurement Challenges Make It Difficult to Meaningfully Interpret a Bottom-Line Rating

OMB published a single, bottom-line rating for the PART results as well as individual section scores, which are potentially more useful for identifying information gaps and program weaknesses. For example, one program that was rated "adequate" overall got high scores for purpose (80 percent) and planning (100 percent), but did poorly in being able to show results (39 percent) and in program management (46 percent). Thus, the individual section ratings provided a better understanding of areas needing improvement than the overall rating alone. Bottom-line ratings inevitably force choices on what best exemplifies a program's mission—even when a program has multiple goals—and encourages a determination of the effectiveness of the program even when performance data are unavailable, the quality of those data is uneven, or they convey a mixed message on performance.

Many of the outcomes for which federal programs are responsible are part of a broader effort involving federal, state, local, nonprofit, and private partners. We have previously reported that it is often difficult to isolate a particular program's contribution to an outcome and especially so when it involves third parties.[12] This was reinforced by the results of the fiscal year 2004 PART reviews. One of the patterns that OMB identified in its ratings was that grant programs received lower than average ratings. To OMB this suggested the need for greater effort by agencies to make grantees accountable for achieving overall program results. However, grant structure and design play a role in how federal agencies are able to hold third parties responsible and complicate the process of identifying the individual contributions of a federal program with multiple partners. In particular, block grants present implementation challenges, especially in those instances in which national goals are not compatible with state and local priorities.

OMB Employed Numerous Tools and Techniques to Promote and Improve Consistent Application of the PART

OMB went to great lengths to encourage consistent application of the PART in the evaluation of government programs. These efforts included (1) testing the PART in selected agencies before use in the fiscal year 2004 assessment, (2) issuing detailed guidance and worksheets for use by PART teams, (3) making the Performance Evaluation Team (PET) available to answer PART implementation questions, (4) establishing an Interagency Review Panel (IRP) to review consistency of PART evaluations, and (5) making improvements to the fiscal year 2005 process and guidance based upon the fiscal year 2004 experience.

OMB conducted a pilot test of the PART and released a draft of the PART questionnaire for public comment prior to its use for the fiscal year 2004 budget cycle. During Spring Review in 2002, OMB and agency staff piloted the draft PART on 67 programs. The PART was also shared with and commented on by the Performance Measurement Advisory Council and other external groups. According to OMB, the results of the Spring Review and feedback from external groups were used to revise the draft version of the PART to lessen subjectivity and increase the consistency of reviews.

[12] See GAO-03-595T and U.S. General Accounting Office, *Managing for Results: Efforts to Strengthen the Link Between Resources and Results at the Administration for Children and Families*, GAO-03-9 (Washington, D.C.: Dec. 10, 2002).

OMB issued detailed guidance to help OMB and agency staff consistently apply the PART and created electronic "templates" or worksheets to aid in completing PART assessments. This guidance explains the purpose of each question and describes the evidence required to support a "yes" or "no" answer. In order to account for different types of programs, several questions tailored to the seven program types were added to the PART (primarily in Section III—Program Management). While the PART guidance cannot be expected to cover every situation, the instructions established general standards for PART evaluations.

PET addressed in "real time" questions and issues that OMB staff that were completing the PART evaluations repeatedly raised. PET consisted of examiners drawn from across the OMB organization representing a variety of programmatic knowledge and experiences. It served as a sounding board for OMB staff and a source for sharing experiences, issues, and useful approaches and also provided training to OMB and agency staff on the process. For example, in one OMB branch, staff were grappling with how to apply the PART to a set of block grants. They went through the instrument with the PET member from their RMO and continued to consult with that individual throughout the process.

OMB also formed IRP, which consisted of both OMB and agency officials, to conduct a consistency check of the PART reviews and to review formal appeals of the process or results for particular questions. During the fiscal year 2004 budget process, IRP conducted a consistency review of 10 percent of the PART evaluations using a subset of the PART questions that OMB staff identified as being the most subjective or difficult to interpret. IRP also reviewed formal agency appeals to determine whether there was consistent treatment of similar situations.

As an Evaluation Tool, the PART Has Weaknesses in Its Design and, as a Result, Its Implementation

Despite the considerable time and effort OMB has devoted to promoting consistent application of the PART, difficulties both with the tool itself (such as subjective terminology and a restrictive yes/no format) and with implementing the tool (including inconsistencies in defining acceptable measures and contradictory answers to "pairs" of related questions) aggravated the general performance measurement challenges described earlier.

Subjective Terms and a Restrictive Format Contributed to Subjective and Inconsistent Responses

Many PART questions contain subjective terms that are open to interpretation. Examples include terminology such as "ambitious" in describing sought-after performance measures. Because the appropriateness of a performance measure depends on the program's purpose, and because program purposes can vary immensely, an ambitious goal for one program might be unrealistic for a similar but more narrowly defined program. Some agency officials claimed that having multiple statutory goals disadvantaged their programs. Without further guidance, subjective terminology can influence program ratings by permitting OMB staff's views about a program's purpose to affect assessments of the program's design and achievements.

Although OMB employed a yes/no format for the PART because OMB believes it aided standardization, the format resulted in oversimplified answers to some questions. OMB received comments on the yes/no format in conducting the PART pilot. Some parties liked the certainty and forced choice of yes/no. Others felt the format did not adequately distinguish between the performance of various programs, especially in the results section (originally in the yes/no format). In response to these concerns, OMB revised the PART in the spring of 2002 to include four response choices in the results section (adding "small extent" and "large extent" to the original two choices "yes" and "no"), while retaining the dichotomous yes/no format in the other three sections. OMB acknowledged that a "yes" response should be definite and reflect a very high standard of performance, and that it would more likely be difficult to justify a "yes" answer than a "no" answer. Nonetheless, agency officials have commented that the yes/no format is a crude reflection of reality, in which progress in planning, management, or results is more likely to resemble a continuum than an on/off switch.

Moreover, the yes/no format was particularly troublesome for questions containing multiple criteria for a "yes" answer. As discussed previously, we conducted an in-depth analysis of PART assessments for 28 related programs in nine clusters and compared the responses to related questions. That analysis showed six instances in which some OMB staff gave a "yes" answer for successfully achieving some but not all of the multiple criteria, while others gave a "no" answer when presented with a similar situation. For example, Section II, Question 1, asks, "Does the program have a limited number of specific, ambitious, long-term performance goals that focus on outcomes and meaningfully reflect the purpose of the program?" The PART defines successful long-term goals by multiple, distinct characteristics

(program has long-term goals, time frames by which the goals are to be achieved, etc.), but does not clarify whether a program can receive a "yes" if *each* of the characteristics is met, or if *most* of the characteristics are met. This contributed to a number of inconsistencies across program reviews. For example, OMB judged DOI's Water Reuse and Recycling program "no" on this question, noting that although DOI set a long-term goal of 500,000 acre-feet per year of reclaimed water, it failed to establish a time frame for when it would reach the target. However, OMB judged the Department of Agriculture's and DOI's Wildland Fire programs "yes" on this question even though the programs' long-term goals of improved conditions in high-priority forest acres are not accompanied by specific time frames. In another example, OMB accepted DOD's recently established long-term strategic goals for medical training and provision of health care even though it did not yet have measures or targets for those goals. By breaking out targets and ambitious time frames separately from the question of annual goals, agencies have an opportunity to get credit for progress made.

There Were Inconsistencies in Defining Acceptable Measures and in Logically Responding to Question "Pairs"

In particular, our analysis of the nine program clusters revealed three instances in which OMB staff inconsistently defined appropriate measures—outcome versus output—for programs. Officials also told us that OMB staff used different standards to define measures as outcome oriented. This may reflect, in part, the complexity of and relationship between expected program benefits. Outcomes are generally defined as the results of outputs—products and services—delivered by a program. But in some programs, long-term outcomes are expected to occur over time through multiple steps. In these cases, short-term outcomes—immediate changes in knowledge and awareness—might be expected to lead to intermediate outcomes—behavioral changes in the future—and eventually result in long-term outcomes—benefits to the public.

In the employment and training area, OMB accepted short-term outcomes, such as obtaining high school diplomas or employment, as a proxy for long-term goals for the HHS Refugee Assistance program, which aims to help refugees attain economic self-sufficiency as soon as possible after they arrive. However, OMB did not accept the same employment rate measure as a proxy for long-term goals for the Department of Education's Vocational Rehabilitation program because it had not set long-term targets beyond a couple of years. In other words, although neither program contained long-term outcomes, such as participants gaining economic self-sufficiency, OMB accepted short-term outcomes in one instance but not the other.

Similarly, OMB gave credit for output measures of claims processing (time, accuracy, and productivity) as a proxy for long-term goals for the Social Security Administration's Disability Insurance program, but did not accept the same output measures for the Veterans Disability Compensation program. OMB took steps to address this issue for fiscal year 2005.

We also found that three "question pairs" on the PART worksheets are linked, yet in two of the three "pairs," a disconnect appeared in how OMB staff responded to these questions for a given program.[13] For example, 29 of the 90 programs (32 percent) judged as lacking "independent and quality evaluations of sufficient scope conducted on a regular basis" (Section II, Question 5) were also judged as having "independent and quality evaluations that indicated the program is effective and achieving results" (Section IV, Question 5). There is a logical inconsistency in these two responses. In another instance, there was no linkage between the questions that examine whether a program *has* annual goals that demonstrate progress toward achieving long-term goals and whether the program actually *achieves* its annual goals. For example, 15 of the 75 programs (20 percent) judged *not* to have adequate annual performance goals (Section II, Question 2) were nevertheless credited for *having made progress* on their annual performance goals (Section IV, Question 2). However, the guidance for the latter question clearly indicates that a program must receive a "no" if it received a "no" on the existence of annual goals (Section II, Question 2). It seems that some raters held programs to a higher standard for the quality of goals than for progress on them.

[13] In the third question pair, a question in the planning section asks about whether the program has long-term goals, and a question in the results section asks whether the agency has made progress in achieving the program's long-term goals. Yet, in 6 of the 115 programs (5 percent) judged not to have adequate long-term goals, credit was given for making progress on their long-term goals even though the guidance again clearly states that a program must receive a "no" if the program received a "no" on the existence of long-term outcome goals.

The Lack of Performance Information Creates Challenges in Effectively Measuring Program Performance

According to OMB, 115 out of 234 programs (49 percent) lacked "specific, ambitious, long-term performance goals that focus on outcomes" (Section II, Question 1). In addition, OMB found that 90 out of 234 programs (38 percent) lacked sufficient "independent, quality evaluations" (Section II, Question 5). While the validity of these assessments may be subject to interpretation and debate, our previous work[14] has raised concerns about the capacity of federal agencies to produce evaluations of program effectiveness.

The lack of evaluations may in part be driven by how OMB defined an "independent and quality evaluation." To be independent, nonbiased parties with no conflict of interest would conduct the evaluation, but agency officials felt that OMB staff started from the default position that agency-sponsored evaluations are, by definition, biased. However, our detailed review of 28 PART worksheets found only 7 instances in which OMB explicitly noted its rejection of evaluations: 1 for being too old, 3 for not being independent (of the 3, 1 was an internal agency review and 2 were conducted by industry groups), and the remaining 3 for not assessing program results. OMB officials have acknowledged that this issue was a point of friction with agencies and that beyond GAO, inspectors general, and other government reports that were automatically presumed to be independent, the independence standard was considered on a case-by-case basis. In these case-by-case situations, OMB staff told us that they looked for some degree of detachment and objectivity in the evaluations. For example, in the case of one DOE-sponsored evaluation, the OMB examiner attended the meetings of the review group that conducted the evaluation in order to see firsthand what sorts of questions the committee posed to the department officials. In OMB's estimation, there was clear independence. While OMB changed the fiscal year 2005 guidance to recognize evaluations contracted out to third parties and agency program evaluation offices as possibly being sufficiently independent, the new guidance generally prohibits evaluations conducted by the program itself from being considered "independent."

Other reasons evaluation data may be limited include (1) constraints on federal agencies' ability to influence program outcomes and reliance on states and others for data for programs for which responsibility has

[14] U.S. General Accounting Office, *Program Evaluation: Agencies Challenged by New Demand for Information on Program Results*, GAO/GGD-98-53 (Washington, D.C.: Apr. 24, 1998).

devolved to the states and (2) the lack of a statutory mandate or dedicated funds for evaluation, which agency officials told us can hamper efforts to conduct studies or to improve administrative data collection.

As we have previously noted, program evaluations can take many forms and agencies may obtain evaluations in a variety of ways. [15] Some evaluations simply analyze routinely collected program administrative data; others involve special surveys. The type of evaluation can greatly affect evaluation cost. Net impact evaluations compare outcomes for program participants to those of a randomly assigned control group and are designed for situations in which external factors are also known to influence those outcomes. However, the adequacy of an evaluation design can only be determined relative to the circumstances of the program being evaluated. In addition, agencies can obtain evaluations by having program or other agency staff collect and analyze the data, by conducting the work jointly with program partners (such as state agencies), or by hiring contract firms to do so. Our survey of 81 federal agency offices conducting evaluations in 1995 of program results found they were most commonly located in administrative offices at a major subdivision level or in program offices (43 and 30 percent, respectively). Overall, they reported conducting 51 percent of their studies in-house, while 34 percent were contracted out. Depending on the sensitivity of the study questions, agencies can conduct credible internal evaluations by adopting procedures to ensure the reliability and validity of data collection and analysis.

[15] GAO/GGD-98-53.

Disagreements on Performance Information Led to Creation of a "Results Not Demonstrated" Category

During the PART process OMB created an additional rating category, "results not demonstrated," which was applied to programs regardless of their score if OMB decided that one or both of two conditions pertained: (1) OMB and the agency could not reach agreement on long-term and annual performance measures and (2) there was inadequate performance information. Almost 50 percent of the 234 programs assessed for fiscal year 2004 received this rating of "results not demonstrated," ranging from high-scoring programs such as the Consumer Product Safety Commission (83) to low-scoring programs such as the Department of Veterans Affairs Disability Compensation program (15). OMB officials said that this rating was given to programs when agreement could not be reached on long-term and annual performance measures and was applied regardless of the program's PART score. Our own review found that OMB generally assigned the "results not demonstrated" rating as described above.[16]

It is important for users of the PART information to interpret the "results not demonstrated" designation as "unknown effectiveness" rather than as meaning the program is "ineffective." Having evidence of poor results is not the same as lacking evidence of effectiveness. Because the PART guidance sets very high standards for obtaining a "yes," a "no" answer can mean either that a program did not meet the standards, or that there is no evidence on whether it met the standards. In some readily measured areas, lack of evidence of an action may indicate that the standard probably was not met. However, because effectiveness is often not readily observed, lack of evidence on program effectiveness cannot be automatically interpreted as meaning that a program is ineffective. Furthermore, an agency might have results for goals established under GPRA, but if OMB and the agency could not reach agreement on new or revised goals or measures, then OMB gave a program the rating "results not demonstrated."

[16] However, we found 8 cases (out of 118) programs that were rated as "results not demonstrated" despite having both annual and long-term performance goals and evidence that these goals were being met.

Changes to the PART and Related Guidance for Fiscal Year 2005 Are Meant to Address Previously Identified Problems

OMB, recognizing many of the issues we have just discussed, made modifications to the PART instrument and guidance in time for the fiscal year 2005 process. OMB said these changes were based upon lessons learned during the fiscal year 2004 process and input from a variety of sources, such as PET, IRP, and agency officials, although we were unable to determine which changes resulted from which recommendations. Although the PART as used for fiscal year 2005 is very similar to that for fiscal year 2004, several questions were added, dropped, merged with other questions, or divided into two questions. For example, a research and development question used in the fiscal year 2004 PART that received "not applicable" answers in 13 out of the 32 cases in which it was applied was dropped from the fiscal year 2005 PART. According to OMB officials, several of the multicriteria questions were split into separate questions in order to reduce inconsistency, as described earlier in this report. Appendix II provides more complete information on the guidance changes between fiscal years 2004 and 2005. To complement the fiscal year 2005 PART guidance and offer strategies for addressing common performance measurement challenges, many of which were encountered during the fiscal year 2004 process, OMB released a separate document, titled *Performance Measurement Challenges and Strategies*, which was the result of a workshop in which agencies participated and identified measurement challenges and shared best practices and possible work-arounds.

Instead of reestablishing IRP (which included both agency and OMB representatives) for the fiscal year 2005 process, OMB officials told us that PET (which included only OMB representatives) would conduct a consistency review of 25 percent of all PART evaluations, with at least one consistency check per OMB branch. OMB also told us that it has asked the National Academy of Public Administration (NAPA) to review PET's consistency review for the fiscal year 2005 process; the scope and results of that review were not available to us during our audit work.[17] OMB senior officials cited resources, timing, and the differing needs of the fiscal year 2004 and 2005 PART processes as reasons for dropping the IRP review. The absence of agency participation in this important phase of the PART could hamper ensuring crucial transparency and credibility.

[17] Because our audit focused on the fiscal year 2004 PART process, our engagement was not limited by OMB's decision to not share its reasoning for shifting the consistency review from IRP to PET or our lack of access to the NAPA review.

The Fiscal Year 2004 PART Process Was a Parallel, Competing Approach to GPRA's Performance Management Framework

The PART was designed for and is used in the executive branch budget preparation and review process; as such, the goals and measures used in the PART must meet OMB's needs. However, GPRA—the current statutory framework for strategic planning and reporting—is a broader process involving the development of strategic and performance goals and objectives to be reported in strategic and annual plans. OMB's desire to collect performance data that better align with budget decision units means that the fiscal year 2004 PART process was a parallel competing structure to the GPRA framework. Although OMB acknowledges that GPRA was the starting point for the PART, as we explain below, the emphasis is shifting such that over time the performance measures developed for the PART and used in the budget process may come to drive agencies' strategic planning processes.

Agencies told us that in some cases, OMB is replacing PART goals and measures for those of GPRA. Effective for fiscal year 2005, OMB's Circular A-11 guidance states that performance budgets are to replace GPRA's annual performance plans. Agencies see the change as detrimental to planning and reporting under GPRA and as a resource drain since they have to respond to both GPRA and PART requirements. Some agency officials told us that although the PART can stimulate discussion on program-specific performance measurement issues, it is not a substitute for GPRA's outcome-oriented, strategic look at thematic goals and departmentwide program comparisons. Moreover, while the PART does not eliminate the departmental strategic plans created under GPRA, many OMB and agency officials told us that the PART is being used to shape the strategic plans.

OMB's Efforts to Link Performance Information with the Budget Often Conflict with Agencies' GPRA Planning Efforts

OMB guidance and officials made clear that GPRA goals, measures, and reports needed to be modified to provide decision makers with program-specific, outcome-based performance data that better aligned with the budget presentation in the President's Budget. According to OMB, such changes were needed because performance reporting under GPRA had evolved into a process separate from budget decision making, with GPRA plans organized at too high a level to be meaningful for program-level budget analysis and management review. Furthermore, according to OMB officials, GPRA plans had too many performance measures, which made it difficult to determine an agency's priorities. However, as some officials pointed out, the cumulative effect of adding new PART measures to GPRA plans may actually increase the number of measures overall; both agency and OMB officials recognize that this is contrary to goals issued by an OMB official previously responsible for the PART, indicating his desire to reduce the number of GPRA measures by at least 25 percent in at least 70 percent of federal departments.[18] As a result of these sometimes-conflicting perspectives, agency officials said that responding to both PART and GPRA requirements increased their workloads and was a drain on staff resources.

OMB's most recent Circular A-11 guidance clearly requires that each agency submit a performance budget for fiscal year 2005 and that this should replace the annual GPRA performance plan.[19] These performance budgets are to include information from the PART assessments, where available, including all performance goals used in the assessment of program performance done under the PART process. Until all programs have been assessed using the PART, the performance budget will also include performance goals for agency programs that have not yet been assessed using the PART. OMB's movement from GPRA to PART is further evident in the fiscal year 2005 PART guidance stating that while existing GPRA performance goals may be a starting point during the development of PART performance goals, the GPRA goals in agency GPRA documents are to be revised significantly, as needed, to reflect OMB's instructions for developing the PART performance goals. Lastly, this same guidance states that GPRA plans should be revised to include any new performance measures used in the PART and unnecessary measures should be deleted from GPRA plans.

[18] Memorandum to the President's Management Council, "Where We'd Be Proud To Be," May 21, 2003.

[19] OMB Circular A-11, *Preparation, Submission, and Execution of the Budget.*

OMB's interest in developing more useful program goals is further evident in its PART recommendations. Almost half of the fiscal year 2004 PART recommendations related to performance assessment—developing outcome goals and measures; cost or efficiency measures; and increasing the tracking/monitoring of data, improving the tracking/monitoring of data, or both. GPRA was generally the starting point for PART discussions about goals and measures, and many agency officials told us that OMB used the PART to modify agencies' existing GPRA goals and measures. Agency officials reported that the discussions about goals and measures were one of the main areas of contention during the PART process. At the same time, agency officials acknowledged that (1) sometimes OMB staff accepted current GPRA measures and (2) sometimes the new PART measures and goals were improvements over the old GPRA measures—the PART measures were more aggressive, more outcome-oriented, more targeted, or all of the above.

Defining a "Unit of Analysis" That Is Useful for Program-Level Budget Analysis and Agency Planning Purposes Presents Challenges

The appropriate unit of analysis or "program" is not always obvious. What OMB determined was useful for a PART assessment did not necessarily match agency organization or planning elements. Although the units of analysis varied across the PART assessments, OMB's guidance stated that they should be linked to a recognized funding level in the budget. In some cases, OMB aggregated separate programs for the purposes of the PART, while in other cases it disaggregated programs. Aggregating programs to tie them to discrete funding levels sometimes made it difficult to create a limited, but comprehensive, set of measures for programs with multiple missions. Disaggregating programs sometimes ignored the interdependence of programs by artificially isolating programs from the larger contexts in which they operate. Both contributed to the lack of available planning and performance information. For example, aggregating rural water supply projects as a single unit of analysis may have been a logical choice for reviewing related activities, but it created problems in identifying planning and performance information useful for the PART since these projects are separately administered. In another case, HHS officials told us that the PART program Substance Abuse Treatment Programs of Regional and National Significance is an amalgamation of activities funded in a single budget line, not an actual program. They said it was a challenge to make these activities look as if they functioned as a single program.

Disaggregating a program too narrowly can create problems by distorting its relationship to other programs involved in achieving a common goal.

For example, agency officials described a homeless program in which outreach workers help homeless persons with emergency needs and refer them to other agencies for housing and needed services. They said that their OMB counterparts suggested that the program adopt long-term outcome measures indicating number of persons housed. Agency officials argued that chronically homeless people require many services and that this federal program often supports only some of the services needed at the initial stages of intervention. The federal program, therefore, could contribute to, but not be primarily responsible for, affecting late stages of the intervention process, like housing status.

These issues reveal some of the unresolved tensions between the President's budget and performance initiative—a detailed budget perspective—and GPRA—a more strategic planning view. In particular, agency officials are concerned with problems in trying to respond to both and overwhelmingly agreed that the PART required a large amount of agency resources to complete. Moreover, some agency officials said that the PART (a program-specific review) is not well suited to one of the key purposes of strategic plans—to convey agencywide, long-term goals and objectives for all major functions and operations. In addition, the time horizons are different for the two initiatives—PART assessments focus on program accomplishments to date while GPRA strategic planning is long-term and prospective in nature.

Changes Made to GPRA in the PART Process Create Uncertainty About Opportunities for Substantive Input by Interested Parties and Congressional Stakeholders

As noted above, PART goals and measures must meet OMB's needs, while GPRA is a broader process involving the development of strategic and performance goals and objectives to be reported in strategic and annual plans. As a phased reform, GPRA required development of the planning framework first, but also explicitly encouraged links to the budget.[20] Our work has shown that under GPRA agencies have made significant progress.[21] Additionally, GPRA requires agencies to consult with Congress and solicit the views of other stakeholders as they develop their strategic plans.[22] We have previously reported[23] that stakeholder involvement appears critical for getting consensus on goals and measures. Stakeholder involvement can be particularly important for federal agencies because they operate in a complex political environment in which legislative mandates are often broadly stated and some stakeholders may strongly disagree about the agency's mission and goals.

The relationship between the PART and its process and the broader GPRA strategic planning process is still evolving. Some tension between the level of stakeholder involvement in the development of performance measures in the GPRA strategic planning process and the process of developing performance measures for the PART is inevitable. Compared to the relatively open-ended GPRA process any budget formulation process is likely to seem closed. An agency's communication with stakeholders, including Congress, about goals and measures created or modified during the formulation of the President's budget is likely to be less than during the development of the agency's own strategic or performance plan. Since different stakeholders have different needs and no one set of goals and measures can serve all purposes, the PART can complement GPRA but should not replace it.

[20] 31 U.S.C. § 1115(a) (2003).

[21] U.S. General Accounting Office, *Managing for Results: Agency Progress in Linking Performance Plans With Budgets and Financial Statements*, GAO-02-236 (Washington, D.C.: Jan. 4, 2002).

[22] 5 U.S.C. § 306(d) (2003).

[23] U.S. General Accounting Office, *Agencies' Strategic Plans Under GPRA: Key Questions to Facilitate Congressional Review (Version 1)*, GAO/GGD-10.1.16 (Washington, D.C.: May 1997).

Although these tensions between the need for internal deliberations and broader consultations are inevitable, if the PART is to be accepted as a credible element in the development of the President's budget proposal, congressional understanding and acceptance of the tool and its analysis will be important. In order for performance information to more fully inform resource allocations, decision makers must also feel comfortable with the appropriateness and accuracy of the performance information and measures associated with these goals. It is unlikely that decision makers will use performance information unless they believe it is credible and reliable and reflects a consensus about performance goals among a community of interested parties. Similarly, the measures used to demonstrate progress toward a goal, no matter how worthwhile, cannot serve the interests of a single stakeholder or purpose without potentially discouraging use of this information by others.

While it is still too soon to know whether OMB-directed measures will satisfy the needs of other stakeholders and GPRA's broader planning purposes, several appropriations subcommittees have stated, in their appropriations hearings, the need to link the PART with congressional oversight. For example, the House Committee on Appropriations, Subcommittee on the Department of the Interior and Related Agencies notes that while it supports the PMA, the costs of initiatives associated with it have generally not been requested in annual budget justifications or through reprogramming procedures.[24] The Subcommittee, therefore, has been unable to evaluate the costs, benefits, and effectiveness of these initiatives or to weigh the priority that these initiatives should receive as compared with ongoing programs funded in the Interior Appropriations bill. Similarly, the House Report on Treasury and Transportation Appropriations included a statement in support of the PART, but noted that the administration's efforts must be linked with the oversight of Congress to maximize the utility of the PART process, and that if the administration treats as privileged or confidential the details of its rating process, it is less likely that Congress will use those results in deciding which programs to fund. Moreover, the Subcommittee said it expects OMB to involve the House and Senate Committees on Appropriations in the development of the PART ratings at all stages in the process.[25]

[24] H.R. Rep. No. 108-195, p. 8 (2003).

[25] H.R. Rep. No. 108-243, pp. 168-69 (2003).

While Congress has a number of opportunities to provide its perspective on performance issues and performance goals, such as when it establishes or reauthorizes a new program, during the annual appropriations process, and in its oversight of federal operations, opportunities exist for Congress to more systematically articulate performance goals and outcomes for key programs of major concern and to allow for timely congressional input in the selection of the PART programs to be assessed.

Conclusions and General Observations

OMB, through its development and use of the PART, has more explicitly infused performance information into the budget formulation process; increased the attention paid to evaluation and performance information; and ultimately, we hope, increased the value of this information to decision makers and other stakeholders. By linking performance information to the budget process, OMB has provided agencies with a powerful incentive for improving data quality and availability. The level of effort and involvement by senior OMB officials and staff clearly signals the importance of this strategy in meeting the priorities outlined in the PMA. OMB should be credited with opening up for scrutiny—and potential criticism—its review of key areas of federal program performance and then making its assessments available to a potentially wider audience through its Web site.

While the PART clearly serves the needs of OMB in budget formulation, questions remain about whether it serves the needs of other key stakeholders. The PART could be strengthened to enhance its credibility and prospects for sustainability by such actions as (1) improving agencies' and OMB's capacity to cope with the demands of the PART, (2) strengthening the PART guidance, (3) expanding the base of credible performance information by strategically focusing evaluation resources, (4) selecting programs for assessment to facilitate crosscutting comparisons and trade-offs, (5) broadening the dialogue with congressional stakeholders, and (6) articulating and implementing a complementary relationship between PART and GPRA.

OMB's ambitious schedule for assessing all federal programs by the fiscal year 2008 President's Budget will require a tremendous commitment of OMB's and agencies' resources. Implementation of the PART recommendations will be a longer-term and potentially more significant result of the PART process than the scores and ratings. No less important will be OMB's involvement both in encouraging agency progress and in signaling its continuing commitment to improving program management and results through the PART. OMB has created a template by which

agencies report on the status of the recommendations and left follow-up on the recommendations to each RMO. However, there is no single focal point for evaluating progress and the results of agency efforts governmentwide; without this it will be difficult for OMB to judge the efficacy of the PART and to know whether the increased workload and trade-offs made with other activities is a good investment of OMB and agency resources.

The goal of the PART is to evaluate programs systematically, consistently, and transparently, but in practice, the tool requires OMB staff to use independent judgment in interpreting the guidance and in making yes or no decisions for what are often complex federal programs. These difficulties are compounded by poor or partial program performance data. Therefore, it is not surprising that we found inconsistencies in our analysis of the fiscal year 2004 PART assessments. Recognizing the inherent limitations of any tool to provide a single performance answer or judgment on complex federal programs with multiple goals, continued improvements in the PART guidance, with examples throughout, can nonetheless help encourage a higher level of consistency as well as transparency.

The PART requires more performance and evaluation information than agencies currently have, as demonstrated by the fact that OMB rated over 50 percent of the programs for fiscal year 2004 as "results not demonstrated" because they "did not have adequate performance goals" or "had not yet collected data to provide evidence of results." In the past, we too have noted limitations in the quality of agency performance and evaluation information and in agency capacity to produce rigorous evaluations of program effectiveness. Furthermore, our work has shown that few agencies deployed the rigorous research methods required to attribute changes in underlying outcomes to program activities. However, program evaluation information often requires large amounts of agency resources to produce, and the agency and OMB may not agree on what is important to measure, particularly when a set of measures cannot serve multiple purposes. Agreement on what are a department or agency's critical, high-risk programs and how best to evaluate them could help leverage limited resources and help determine what are the most important program evaluation data to collect.

Federal programs are designed and implemented in dynamic environments where competing program priorities and stakeholders' needs must be balanced continually and new needs must be addressed. GPRA is a broad legislative framework that was designed to be consultative with Congress and other stakeholders and allows for varying uses of performance

information, while the PART applies evaluation information to support decisions and program reviews during the executive budget formulation process. While the PART reflects the administration's management principles and the priority given to using performance information in OMB's decision-making process, its focus on program-level assessments cannot substitute for the inclusive, crosscutting strategic planning required by GPRA. Moreover, GPRA can anchor the review of programs by providing an overall strategic context for programs' contributions toward agency goals. Although PART and GPRA serve different needs, a strategy for integrating the two could help strengthen both.

Opportunities exist to develop a more strategic approach to the selection and prioritization of areas to be assessed under the PART process. Targeting PART assessments based on such factors as the relative priorities, costs, and risks associated with related clusters of programs and activities could not only help ration scarce analytic resources but could also focus decision makers' attention on the most pressing policy and program issues. Moreover, such an approach could facilitate the use of PART assessments to review the relative contributions of similar programs to common or crosscutting goals and outcomes.

As part of the President's budget preparation, the PART clearly must serve the President's interests. However, it is unlikely that the broad range of actors whose input is critical to decisions will use performance information unless they believe it is credible and reliable and reflects a consensus about performance goals among a community of interested parties. Similarly, the measures used to demonstrate progress toward a goal, no matter how worthwhile, cannot appear to serve a single set of interests without potentially discouraging use of this information by others. If the President or OMB wants the PART and its results to be considered in the congressional debate, it will be important for OMB to (1) involve congressional stakeholders early in providing input on the focus of the assessments; (2) clarify any significant limitations in the assessments as well as the underlying performance information; and (3) initiate discussions with key congressional committees about how they can best take advantage of and leverage PART information in authorizations, appropriations, and oversight processes.

As we have previously reported, effective congressional oversight can help improve federal performance by examining the program structures agencies use to deliver products and services to ensure that the best, most cost-effective mix of strategies is in place to meet agency and national

goals. While Congress has a number of opportunities to provide its perspective on performance issues and performance goals, such as when it establishes or reauthorizes a new program, during the annual appropriations process, and in its oversight of federal operations, a more systematic approach could allow Congress to better articulate performance goals and outcomes for key programs of major concern. Such an approach could also facilitate OMB's understanding of congressional priorities and concerns and, as a result, increase the usefulness of the PART in budget deliberations.

Matter for Congressional Consideration

In order to facilitate an understanding of congressional priorities and concerns, we suggest that Congress consider the need for a strategy that could include (1) establishing a vehicle for communicating performance goals and measures for key congressional priorities and concerns; (2) developing a more structured oversight agenda to permit a more coordinated congressional perspective on crosscutting programs and policies; and (3) using such an agenda to inform its authorization, oversight, and appropriations processes.

Recommendations for Executive Action

We have seven recommendations to OMB for building on and improving the first year's experience with the PART and its process. We recommend that the Director of OMB take the following actions:

- Centrally monitor agency implementation and progress on PART recommendations and report such progress in OMB's budget submission to Congress. Governmentwide councils may be effective vehicles for assisting OMB in these efforts.

- Continue to improve the PART guidance by (1) expanding the discussion of how the unit of analysis is to be determined to include trade-offs made when defining a unit of analysis, implications of how the unit of analysis is defined, or both; (2) clarifying when output versus outcome measures are acceptable; and (3) better defining an "independent, quality evaluation."

- Clarify OMB's expectations to agencies regarding the allocation of scarce evaluation resources among programs, the timing of such evaluations, as well as the evaluation strategies it wants for the purposes of the PART, and consider using internal agency evaluations as

evidence on a case-by-case basis—whether conducted by agencies, contractors, or other parties.

- Reconsider plans for 100 percent coverage of federal programs and, instead, target for review a significant percentage of major and meaningful government programs based on such factors as the relative priorities, costs, and risks associated with related clusters of programs and activities.
- Maximize the opportunity to review similar programs or activities in the same year to facilitate comparisons and trade-offs.
- Attempt to generate, early in the PART process, an ongoing, meaningful dialogue with congressional appropriations, authorization, and oversight committees about what they consider to be the most important performance issues and program areas warranting review.
- Seek to achieve the greatest benefit from both GPRA and PART by articulating and implementing an integrated, complementary relationship between the two.

Agency Comments

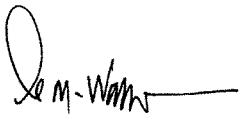
We provided a draft of this report to OMB for its review and comment. OMB generally agreed with our findings, conclusions, and recommendations. In addition, OMB outlined actions it is taking to address many of our recommendations, including refining the process for monitoring agencies' progress in implementing the PART recommendations, seeking opportunities for dialogue with Congress on agencies' performance, and continuing to improve executive branch implementation of GPRA plans and reports. OMB officials provided a number of technical comments and clarifications, which we incorporated as appropriate to ensure the accuracy of our report. OMB's comments appear in appendix IV. We also received technical comments on excerpts of the draft provided to the Departments of the Interior, Energy, and Health and Human Services. Comments received from the Departments of Energy and the Interior were incorporated as appropriate. The Department of Health and Human Services had no comments.

OMB noted that performance information gleaned from the PART process has not only informed budget decisions but has also helped direct program management, identified opportunities to improve program design, and promoted accountability. We agree. As shown in figure 1 in our report, we

found that 82 percent of PART recommendations addressed program assessment, design, and management issues; only 18 percent of the recommendations had a direct link to funding matters.

We are sending copies of this report to the Director of OMB, appropriate congressional committees, and other interested members of Congress. We will also make copies available to others upon request. In addition, the report will be available at no charge on the GAO Web site at http://www.gao.gov.

If you or your staff have questions about this report, please contact Paul Posner at (202) 512-9573 or posnerp@gao.gov. An additional contact and key contributors to this report are listed in appendix V.



David M. Walker
Comptroller General
of the United States

Appendix I

Scope and Methodology

To address the objectives in this report, we reviewed Office of Management and Budget (OMB) materials and presentations on the development and implementation of the Program Assessment Rating Tool (PART) as well as the results of the PART assessments. Our review of materials included instructions for using PART, OMB's testimony concerning PART, and public remarks made by OMB officials at relevant conferences and training. We also reviewed PART-related information on OMB's Web site, including the OMB worksheets used to support the assessments, and attended OMB's PART training for the fiscal year 2004 process.

For this report, we focused on the process and final results of the fiscal year 2004 PART process, but also looked at the initial stages of the fiscal year 2005 process. We compared the PART guidance for both years and asked agency and OMB staff to discuss generally the differences between the 2 fiscal years. We did not review the final results for the fiscal year 2005 PART, which are embargoed until the publication of the President's fiscal year 2005 budget request. For the same reasons, we did not review the results of any reassessments conducted for fiscal year 2005 on programs originally assessed for fiscal year 2004. This report presents the experiences of staff from the three departments and OMB officials who we interviewed. We did not directly observe the PART process (for either year) in operation nor did we independently verify the PART assessments as posted on OMB's Web site or the program or financial information contained in the documents provided as evidence for the PART assessments. We did, however, take several steps to ensure that we reliably downloaded and combined the PART summaries and worksheets with our budget and recommendation classifications. Our steps included (1) having the computer programs we used to create and process our consolidated dataset verified by a second programmer; (2) having transcribed data elements from all programs checked back to source files; and (3) having selected, computer-processed data elements checked back to source files for a random sample of programs and also for specific programs identified in our analyses.

To better understand the universe of programs OMB assessed for fiscal year 2004, we developed overall profiles of PART results and examined relationships between such characteristics as type of program, type of recommendation, overall rating, total PART score, and answers for each question on PART. This review enabled us to generally confirm some information previously reported by OMB, for example, that PART scores do not automatically determine proposed funding and that grant programs scored lower overall than other types of programs. It also allowed us to

select a sample of programs for more in-depth review, and this sample was used to determine which OMB and agency officials we interviewed.

To gain a better understanding of the PART process at both OMB and agencies, to inform our examination of the usefulness of PART as an evaluation tool, and to obtain various perspectives on the relationship between PART and GPRA, we interviewed officials at OMB and three selected departments. At OMB, we interviewed a range of staff, such as associate directors, deputy assistant directors, branch chiefs, and examiners. Specifically, we interviewed staff in two Resource Management Offices (RMO). In the Human Resources Programs RMO, we spoke with staff from the Health Division and the Education and Human Resources Division. In the Natural Resources, Energy and Science RMO we interviewed staff from the Energy and Interior Branches. In addition, we obtained the views of two groups within OMB that were convened specifically for the PART process: the Performance Evaluation Team (PET) and the Interagency Review Panel (IRP). The IRP included agency officials in addition to staff from OMB.

The three departments for which we reviewed the PART process were the Department of Energy (DOE), the Department of Health and Human Services (HHS), and the Department of the Interior (DOI). We selected these three departments based on our data analysis of program types. The departments selected and their agencies had a variety of program types (e.g., block/formula grants, competitive grants, direct federal, and research and development) that were subject to PART and could provide us with a broad-based perspective on how PART was applied to different programs employing diverse tools of government. We also chose these three departments because they had programs under PART review within the two RMOs at OMB where we did more extensive interviewing, thus enabling us to develop a more in-depth understanding of how the PART process operated for a subset of programs. We used this information to complement our broader profiling of all 234 programs assessed. Within DOE we studied the experiences of the Office of Science, the Office of Energy Efficiency and Renewable Energy, and the Office of Fossil Energy. Within HHS, we studied the experiences of the Administration for Children and Families, the Health Resources and Services Administration, and the Substance Abuse and Mental Health Services Administration. Within DOI, we studied the experiences of the Bureau of Land Management, the Bureau of Indian Affairs, and the National Park Service. We interviewed planning, budget, and program staff within each of the nine agencies as well as those

at the department level. We also reviewed relevant supporting materials provided by these departments in conjunction with these interviews.

To allow us to describe how PART was used in fiscal year 2004 to influence changes in future performance, we created a consolidated dataset in which we classified recommendations OMB made by three areas in need of improvement: (1) program design, (2) program management, and (3) program assessment. A fourth category was created for those recommendations that involved funding issues. We created a consolidated dataset of information from our analysis of recommendations and selected information from the PART program summary page and worksheet for each program.[1]

In addition, for approximately 95 percent of the programs, we identified whether the basis for program funding was mandatory or discretionary. It was important to separate discretionary and mandatory programs in our review of PART's potential influence on the President's budget proposals because funding for mandatory programs is determined through authorizations, not through the annual appropriations process. Of the 234 programs that OMB assessed for fiscal year 2004, we identified 27 mandatory programs and 196 discretionary, but could not categorize 11 programs as solely mandatory or discretionary because they were too mixed to classify.[2]

For discretionary programs, we explored the relationship between PART results and proposed budget changes in a series of regression analyses.[3] Using statistical analysis, we found that PART scores influenced proposed

[1] The PART program summary sheets are included in the *Budget of the United States Government, Fiscal Year 2004, Performance and Management Assessments* (Washington, D.C.: February 2003). The summary sheets and worksheets for the 234 programs are on OMB's Web site: http://www.whitehouse.gov/omb/budget/fy2004/pma.html.

[2] These 11 programs are animal welfare, food aid, multifamily housing direct loans and rental assistance, rural electric utility loans and guarantees, and rural water and wastewater grants and loans programs in the Department of Agriculture; the nursing education loan repayment and scholarship program in HHS; the methane hydrates program in DOE; the reclamation hydropower program in DOI; the long-term guarantees program in the U.S. Export-Import Bank; and the climate change and development assistance/population programs in the Agency for International Development.

[3] We tested the regression on mandatory programs and as expected the results showed no relationship between the PART scores and the level of funding proposed in the President's Budget.

funding changes for discretionary programs; however, a large amount of variability in these changes remains unexplained. We examined proposed funding changes between fiscal years 2003 and 2004 (measured by percentage change) and the relationship to PART scores for the programs assessed in the fiscal year 2004 President's Budget. These scores are the weighted sums of scores for four PART categories: Program Purpose and Design, Strategic Planning, Program Management, and Program Results and Accountability. The corresponding weights assigned by OMB are 0.2, 0.1, 0.2, and 0.5, respectively.[4] Tables in this appendix report regression results obtained using the method of least squares with heteroskedasticity-corrected standard errors.[5] The same estimation method is used throughout this analysis.

Overall PART scores have a positive and statistically significant effect on discretionary program funding. The programs evaluated by OMB include both mandatory and discretionary programs. Regression results for mandatory programs showed—as expected—no relationship between PART scores and the level of funding in the President's Budget proposal. Assessment ratings, however, can potentially affect the funding for discretionary programs either in the President's Budget proposal or in congressional deliberations on spending bills.[6] Table 3 reports the regression results for discretionary programs.

[4] *Budget of the United States Government, Fiscal Year 2004, Performance and Management Assessments*, 10.

[5] For a discussion of this method, see W.H. Greene, *Econometric Analysis*, Section 10.3 (Upper Saddle River, N.J.: Prentice Hall, 2003).

[6] *Budget of the United States Government, Fiscal Year 2001, A Citizen's Guide to the Federal Budget* (Washington, D.C.: February 2000), http://w3.access.gpo.gov/usbudget/fy2001/guide03.html, (downloaded April 2003), 2.

Table 3: The Effect of Overall PART Score on Proposed Funding Changes (Discretionary Programs)

Variable	Coefficient estimate	Robust standard error	t-Statistic	P-value
Overall PART score	0.536	0.159	3.38	0.001
Constant	-25.671	8.682	-2.96	0.003

Source. GAO analysis of OMB data

Notes: R-squared = 0.058, Prob-F = 0.001, N = 196. Originally we identified 197 discretionary programs. However, no fiscal year 2004 budget estimate is reported for the Disclosed Worker Assistance program due to grant consolidation at the Department of Labor. (*Budget of the United States Government, Fiscal Year 2004, Performance and Management Assessments* (Washington, D.C.: February 2003), 191.) This reduced the number of discretionary programs to 196.

The estimated coefficient of the overall score is positive and significant. These results show that the aggregate PART score has a positive and statistically significant effect on the proposed change in discretionary programs' budget, suggesting that programs with better scores are more likely to receive larger proposed budget increases.

To examine the effect of program size on our results, we divided all programs equally into three groups—small, medium, and large—based on their fiscal year 2003 funding estimate. Regressions similar to those reported in table 3 were then performed for discretionary programs in each group. The results, reported in tables 4, 5, and 6 suggest that the statistically significant effect of overall scores on budget outcomes exists only for the smaller programs. The estimated coefficient of the overall score for large programs, which is significant but only at the 10 percent level, reflects an outlier.[7] Once this outlier is dropped, the estimated coefficient becomes statistically insignificant.

[7] The outlier is the Community Oriented Policing Services program with an estimated 77 percent reduction in funding (see OMB, *Budget of the U.S. Government, Fiscal Year 2004, Performance and Management Assessments*, (Washington, D.C.: February 2003), 178). The outlier in this case is identified using scatter plot and estimating with and without the outlier. The reported results for small and medium programs are not outlier driven.

Table 4: The Effect of Overall PART Score on Proposed Funding Changes (Small Discretionary Programs)

Variable	Coefficient estimate	Robust standard error	t-Statistic	P-value
Overall PART score	1.074	0.404	2.66	0.010
Constant	-50.523	21.155	-2.39	0.020

Source: GAO analysis of OMB data.

Note: R-squared = 0.092, Prob-F = 0.01, N = 71.

Table 5: The Effect of Overall PART Score on Proposed Funding Changes (Medium-Size Discretionary Programs)

Variable	Coefficient estimate	Robust standard error	t-Statistic	P-value
Overall PART score	0.306	0.188	1.62	0.109
Constant	-17.984	12.480	-1.44	0.154

Source: GAO analysis of OMB data.

Note: R-squared = 0.039, Prob-F = 0.109, N = 67.

Table 6: The Effect of Overall PART Score on Proposed Funding Changes (Large Discretionary Programs)

Variable	Coefficient estimate	Robust standard error	t-Statistic	P-value
Overall PART score	0.194	0.109	1.77	0.082
Constant	-8.216	7.778	-1.06	0.295

Source: GAO analysis of OMB data.

Note: R-squared = 0.057, Prob-F = 0.082, N = 58.

The statistical analysis suggests that among the four components of the PART questionnaire, program purpose, management, and results have statistically significant effects on proposed funding changes, but the effects of program purpose and results are more robust across the estimated models. The overall score is a weighted average of four components: Program Purpose and Design, Strategic Planning, Program Management, and Program Results and Accountability.[8] To identify which of the four components contribute to the significant relationship observed here, we

examined the effect of each on proposed changes in programs' funding levels. Tables 7 and 8 show estimates from regressions of the proposed funding change on purpose, planning, management, and results scores for all discretionary programs as well as small discretionary programs alone.

Table 7: The Effect of PART Component Scores on Proposed Funding Changes (All Discretionary Programs)

Variable	Coefficient estimate	Robust standard error	t-Statistic	P-value
Purpose	0.325	0.127	2.56	0.011
Plan	-0.259	0.199	-1.30	0.194
Management	0.191	0.117	1.63	0.105
Results	0.363	0.205	1.77	0.078
Constant	-33.096	14.136	-2.34	0.020

Source: GAO analysis of OMB data

Note: R-squared = 0.087, Prob-F = 0.003, N = 196.

Table 8: The Effect of PART Component Scores on Proposed Funding Changes (Small Discretionary Programs)

Variable	Coefficient estimate	Robust standard error	t-Statistic	P-value
Purpose	0.223	0.274	0.81	0.419
Plan	-0.671	0.543	-1.24	0.221
Management	0.547	0.304	1.80	0.077
Results	0.956	0.534	1.79	0.078
Constant	-42.455	34.800	-1.22	0.227

Source. GAO analysis of OMB data.

Note: R-squared = 0.149, Prob-F = 0.043, N = 71.

These results suggest that among the four components, program purpose, management, and results are more likely to affect the proposed budget changes for discretionary programs. When all discretionary programs are

[8] *Budget of the United States Government, Fiscal Year 2004, Performance and Management Assessments*, 10.

included, the estimated coefficients are positive and significant for results (at the 10 percent level) and purpose. When only the small discretionary programs are included, the estimated coefficients are positive and significant for both management and results (at the 10 percent level). We also estimated the above regression for medium and large programs, but coefficient estimates were not statistically significant, except for the estimated coefficient of purpose for medium programs.

PART scores explain at most about 15 percent of the proposed funding changes, leaving a large portion of the variability in proposed funding changes unexplained. This suggests that most of the variance is due to institutional factors, program specifics, and other unquantifiable factors. The coefficient of determination (or R^2) is used to measure the proportion of the total variation in the regression's dependent variable that is explained by the variation in the regressors (independent variables).[9] The maximum value of this measure across all estimated regressions is about 15 percent.

Similar analyses were carried out for changes in the proposed budget for fiscal year 2004 and congressionally appropriated amounts in fiscal year 2002. Results were qualitatively similar to those reported here.

To assess the strengths and weaknesses of PART as an evaluation tool and the consistency with which it was applied, we analyzed data from all 234 programs that OMB reviewed using PART for fiscal year 2004. As part of our examination of the consistency with which PART was applied to programs, we also focused on a subset of programs to assess the way in which certain measurement issues were addressed across those programs. The issues were selected from those identified in interviews with officials from the selected agencies described above and our own review of the PART program summaries and worksheets. Measurement issues included acceptance of output versus outcome measures of annual and long-term goals, types of studies accepted as program evaluations, acknowledgment of related programs, and justifications for judging a PART question as "not applicable." Programs were selected that formed clusters, each addressing a similar goal or shared a structural similarity pertinent to performance measurement, to examine whether PART assessment issues were handled similarly across programs when expected. We reviewed the worksheets and compared the treatment of assessment issues across specific questions

[9] See Greene, 33.

within and across programs in a cluster to identify potential inconsistencies in how the tool was applied. We reviewed a total of 28 programs in nine clusters. The nine clusters are food safety, water supply, military equipment procurement, provision of health care, statistical agencies, block grants to assist vulnerable populations, energy research programs, wildland fire management, and disability compensation.

With the exception of our summary analyses of all 234 programs, the information obtained from OMB and agency interviews, related material, and review of selected programs is not generalizable to the PART process for all 234 programs reviewed in fiscal year 2004. We conducted our review from May through October 2003 in accordance with generally accepted government auditing standards.

The Fiscal Year 2004 PART and Differences Between the Fiscal Year 2004 and 2005 PARTs

Below we have reproduced OMB's fiscal year 2004 PART instrument. We have also included the comparison of fiscal year 2004 and fiscal year 2005 PART questions that appeared in the fiscal year 2005 PART guidance (see table 9).

Section I: Program Purpose & Design (Yes, No, N/A)

1. Is the program purpose clear?
2. Does the program address a specific interest, problem or need?
3. Is the program designed to have a significant impact in addressing the interest, problem or need?
4. Is the program designed to make a unique contribution in addressing the interest, problem or need (i.e., not needlessly redundant of any other Federal, state, local or private efforts)?
5. Is the program optimally designed to address the interest, problem or need?

Specific Program Purpose & Design Questions by Program Type

Research and Development Programs

6. (RD. 1) Does the program effectively articulate potential public benefits?
7. (RD. 2) If an industry-related problem, can the program explain how the market fails to motivate private investment?

Section II: Strategic Planning (Yes, No, N/A)

1. Does the program have a limited number of specific, ambitious long-term performance goals that focus on outcomes and meaningfully reflect the purpose of the program?
2. Does the program have a limited number of annual performance goals that demonstrate progress toward achieving the long-term goals?
3. Do all partners (grantees, subgrantees, contractors, etc.) support program-planning efforts by committing to the annual and/or long-term goals of the program?

4. Does the program collaborate and coordinate effectively with related programs that share similar goals and objectives?

5. Are independent and quality evaluations of sufficient scope conducted on a regular basis or as needed to fill gaps in performance information to support program improvements and evaluate effectiveness?

6. Is the program budget aligned with the program goals in such a way that the impact of funding, policy, and legislative changes on performance is readily known?

7. Has the program taken meaningful steps to address its strategic planning deficiencies?

Specific Strategic Planning Questions by Program Type

Regulatory-Based Programs

8. (RD. 1) Are all regulations issued by the program/agency necessary to meet the stated goals of the program, and do all regulations clearly indicate how the rules contribute to achievement of the goals?

Capital Assets and Service Acquisition Programs

8. (Cap. 1) Are acquisition program plans adjusted in response to performance data and changing conditions?

9. (Cap. 2) Has the agency/program conducted a recent, meaningful, credible analysis of alternatives that includes trade-offs between cost, schedule and performance goals?

Research and Development Programs

8. (RD. 1) Is evaluation of the program's continuing relevance to mission, fields of science, and other "customer" needs conducted on a regular basis?

9. (RD. 2) Has the program identified clear priorities?

Section III: Program Management (Yes, No, N/A)

1. Does the agency regularly collect timely and credible performance information, including information from key program partners, and use it to manage the program and improve performance?
2. Are Federal managers and program partners (grantees, subgrantees, contractors, etc.) held accountable for cost, schedule and performance results?
3. Are all funds (Federal and partners') obligated in a timely manner and spent for the intended purpose?
4. Does the program have incentives and procedures (e.g., competitive sourcing/cost comparisons, IT improvements) to measure and achieve efficiencies and cost effectiveness in program execution?
5. Does the agency estimate and budget for the full annual costs of operating the program (including all administrative costs and allocated overhead) so that program performance changes are identified with changes in funding levels?
6. Does the program use strong financial management practices?
7. Has the program taken meaningful steps to address its management deficiencies?

Specific Program Management Questions by Program Type

Competitive Grant Programs

8. (Co. 1) Are grant applications independently reviewed based on clear criteria (rather than earmarked) and are awards made based on results of the peer review process?
9. (Co. 2) Does the grant competition encourage the participation of new/first-time grantees through a fair and open application process?
10. (Co. 3) Does the program have oversight practices that provide sufficient knowledge of grantee activities?
11. (Co. 4) Does the program collect performance data on an annual basis and make it available to the public in a transparent and meaningful manner?

Block/Formula Grant Programs

8. (B. 1) Does the program have oversight practices that provide sufficient knowledge of grantee activities?

9. (B. 2) Does the program collect grantee performance data on an annual basis and make it available to the public in a transparent and meaningful manner?

Regulatory-Based Programs

8. (Reg. 1) Did the program seek and take into account the views of affected parties including state, local and tribal governments and small businesses, in drafting significant regulations?

9. (Reg. 2) Did the program prepare, where appropriate, a Regulatory Impact Analysis that comports with OMB's economic analysis guidelines and have these RIA analyses and supporting science and economic data been subjected to external peer review by qualified specialists?

10. (Reg. 3) Does the program systematically review its current regulations to ensure consistency among all regulations in accomplishing program goals?

11. (Reg. 4) In developing new regulations, are incremental societal costs and benefits compared?

12. (Reg. 5) Did the regulatory changes to the program maximize net benefits?

13. (Reg. 6) Does the program impose the least burden, to the extent practicable, on regulated entities, taking into account the costs of cumulative final regulations?

Capital Assets and Service Acquisition Programs

8. (Cap. 1) Does the program define the required quality, capability, and performance objectives of deliverables?

9. (Cap. 2) Has the program established appropriate, credible, cost and schedule goals?

10. (Cap. 3) Has the program conducted a recent, credible, cost-benefit analysis that shows a net benefit?

11. (Cap. 4) Does the program have a comprehensive strategy for risk management that appropriately shares risk between the government and contractor?

Credit Programs

8. (Cr. 1) Is the program managed on an ongoing basis to assure credit quality remains sound, collections and disbursements are timely and reporting requirements are fulfilled?

9. (Cr. 2) Does the program consistently meet the requirements of the Federal Credit Reform Act of 1990, the Debt Collection Improvement Act and applicable guidance under OMB Circulars A-1, A-34, and A-129?

10. (Cr. 3) Is the risk of the program to the U.S. Government measured effectively?

Research and Development Programs

8. (RD. 1) Does the program allocate funds through a competitive, merit-based process, or, if not, does it justify funding methods and document how quality is maintained?

9. (RD. 2) Does competition encourage the participation of new/first-time performers through a fair and open application process?

10. (RD. 3) Does the program adequately define appropriate termination points and other decision points?

11. (RD. 4) If the program includes technology development or construction or operation of a facility, does the program clearly define deliverables and required capability/performance characteristics and appropriate, credible cost and schedule goals?

Section IV: Program Results (Yes, Large Extent, Small Extent, No)

1. Has the program demonstrated adequate progress in achieving its long-term outcome goal(s)?

   - Long-Term Goal I:
     Target:
     Actual Progress achieved toward goal:

   - Long-Term Goal II:
     Target:
     Actual Progress achieved toward goal:

   - Long-Term Goal III:
     Target:
     Actual Progress achieved toward goal:

2. Does the program (including program partners) achieve its annual performance goals?

   - Key Goal I:
     Performance Target:
     Actual Performance:

   - Key Goal II:
     Performance Target:
     Actual Performance:

   - Key Goal III:
     Performance Target:
     Actual Performance:

Note: Performance targets should reference the performance baseline and years, e.g. achieve a 5% increase over base of X in 2000.

3. Does the program demonstrate improved efficiencies and cost effectiveness in achieving program goals each year?

4. Does the performance of this program compare favorably to other programs with similar purpose and goals?

5. Do independent and quality evaluations of this program indicate that the program is effective and achieving results?

Specific Results Questions by Program Type

Regulatory-Based Programs

6. (Reg. 1) Were programmatic goals (and benefits) achieved at the least incremental societal cost and did the program maximize net benefits?

Capital Assets and Service Acquisition Programs

6. (Cap. 1) Were program goals achieved within budgeted costs and established schedules?

Research and Development Programs

6. (RD. 1) If the program includes construction of a facility, were program goals achieved within budgeted costs and established schedules?

Table 9: Side-by-Side of the Fiscal Year 2005 PART and the Fiscal Year 2004 PART Questions

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
I. Program purpose & design				
1.1	Is the program purpose clear?	1	Same	
1.2	Does the program address a specific and existing problem, interest, or need?	2	Does the program address a specific interest, problem or need?	Wording clarified.
		3	Is the program designed to have a significant impact in addressing the interest, problem or need?	Dropped; "significant" worked against small programs and was not clear.
1.3	Is the program designed so that it is not redundant or duplicative of any other Federal, state, local or private effort?	4	Is the program designed to make a unique contribution in addressing the interest, problem or need (i.e., is not needlessly redundant of any other Federal, state or, local or private effort)?	Wording clarified.
1.4	Is the program design free of major flaws that would limit the program's effectiveness or efficiency?	5	Is the program optimally designed to address the national, interest, problem or need?	Minor change to clarify focus; "optimally" was too broad.
1.5	Is the program effectively targeted, so that resources will reach intended beneficiaries and/or otherwise address the program's purpose directly?			New question to address distributional design.

(Continued From Previous Page)

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
	Specific Program Purpose and Design Questions by Program Type			
			Research and Development Programs	
		RD.1	Does the program effectively articulate potential public benefits?	Dropped; covered by 1.2.
		RD. 2	If an industry-related problem, can the program explain how the market fails to motivate private investment?	Dropped; covered by I.2 and I.5.
II. Strategic planning				
2.1	Does the program have a limited number of specific long-term performance measures that focus on outcomes and meaningfully reflect the purpose of the program?	1	Does the program have a limited number of specific, ambitious long-term performance goals that focus on outcomes and meaningfully reflect the purpose of the program?	Splits old II.1 into separate questions on existence of (1) long-term performance measures and (2) targets for these measures. Together, the measures and targets comprise the long-term performance goals addressed in last year's question.
2.2	Does the program have ambitious targets and timeframes for its long-term measures?			Splits old II.1; see above.
2.3	Does the program have a limited number of specific annual performance measures that can demonstrate progress toward achieving the program's long-term goals?	2	Does the program have a limited number of annual performance goals that demonstrate progress toward achieving the long-term goals?	Splits old II.2 into separate questions on existence of (1) annual performance measures and (2) targets for these measures. Together, the measures and targets comprise the annual performance goals addressed in last year's question.
2.4	Does the program have baselines and ambitious targets for its annual measures?			Splits old II.2; see above.
2.5	Do all partners (including grantees, sub-grantees, contractors, cost-sharing partners, and other government partners) commit to and work toward the annual and/or long-term goals of the program?	3	Do all partners (grantees, sub-grantees, contractors, etc.) support program planning efforts by committing to the annual and/or long-term goals of the program?	Wording clarified.
		4	Does the program collaborate and coordinate effectively with related programs that share similar goals and objectives?	Moved to question 3.5.

(Continued From Previous Page)

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
2.6	Are independent evaluations of sufficient scope and quality conducted on a regular basis or as needed to support program improvements and evaluate effectiveness and relevance to the problem, interest, or need?	5	Are independent and quality evaluations of sufficient scope conducted on a regular basis or as needed to fill gaps in performance information to support program improvements and evaluate effectiveness?	Wording clarified.
2.7	Are budget requests explicitly tied to accomplishment of the annual and long-term performance goals, and are the resource needs presented in a complete and transparent manner in the program's budget?	6	Is the program budget aligned with the program goals in such a way that the impact of funding, policy, and legislative changes on performance is readily known?	Modified.
2.8	Has the program taken meaningful steps to correct its strategic planning deficiencies?	7	Same.	
	Specific Strategic Planning Questions by Program Type			
	Regulatory Based Programs			
2.RG1	Are all regulations issued by the program/agency necessary to meet the stated goals of the program, and do all regulations clearly indicate how the rules contribute to achievement of the goals?	Reg. 1	Same.	
	Capital Assets & Service Acquisition Programs			
		Cap. 1	Are acquisition program plans adjusted in response to performance data and changing conditions?	Dropped; covered in 2.CA1 and 3.CA1.
2.CA1	Has the agency/program conducted a recent, meaningful, credible analysis of alternatives that includes trade-offs between cost, schedule, risk, and performance goals and used the results to guide the resulting activity?	Cap. 2	Has the agency/program conducted a recent, meaningful, credible analysis of alternatives that includes trade-offs between cost, schedule and performance goals?	Minor change.
	R&D Programs			
	R&D programs addressing technology development or the construction or operation of a facility should answer 2.CA1.			
2.RD1	If applicable, does the program assess and compare the potential benefits of efforts within the program to other efforts that have similar goals?	RD. 1	Is evaluation of the program's continuing relevance to mission, fields of science, and other "customer" needs conducted on a regular basis?	Modified.
2.RD2	Does the program use a prioritization process to guide budget requests and funding decisions?	RD. 2	Has the program identified clear priorities?	Modified.

(Continued From Previous Page)

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
III. Program management				
3.1	Does the agency regularly collect timely and credible performance information, including information from key program partners, and use it to manage the program and improve performance?	1	Same.	
3.2	Are Federal managers and program partners (including grantees, sub-grantees, contractors, cost-sharing partners, and other government partners) held accountable for cost, schedule and performance results?	2	Same.	
3.3	Are funds (Federal and partners') obligated in a timely manner and spent for the intended purpose?	3	Same.	
3.4	Does the program have procedures (e.g. competitive sourcing/cost comparisons, IT improvements, appropriate incentives) to measure and achieve efficiencies and cost effectiveness in program execution?	4	Same.	
3.5	Does the program collaborate and coordinate effectively with related programs?			Same as old question 2.4.
		5	Does the agency estimate and budget for the full annual costs of operating the program (including all administrative costs and allocated overhead) so that program performance changes are identified with changes in funding levels?	Now covered by guidance for question 2.7.
3.6	Does the program use strong financial management practices?	6	Same.	
3.7	Has the program taken meaningful steps to address its management deficiencies?	7	Same.	
	Specific Program Management Questions by Program Type			
	Competitive Grant Programs			
3.CO1	Are grants awarded based on a clear competitive process that includes a qualified assessment of merit?	Co. 1	Are grant applications independently reviewed based on clear criteria (rather than earmarked) and are awards made based on results of the peer review process?	Modified. Guidance also captures former question Co. 2.

(Continued From Previous Page)

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
		Co.2	Does the grant competition encourage the participation of new/first-time grantees through a fair and open application process?	Now considered in guidance for answering 3.CO1, above.
3.CO2	Does the program have oversight practices that provide sufficient knowledge of grantee activities?	Co. 3	Does the agency have sufficient knowledge about grantee activities?	Wording clarified.
3.CO3	Does the program collect grantee performance data on an annual basis and make it available to the public in a transparent and meaningful manner?	Co. 4	Same.	
	Block/Formula Grant Programs			
3.BF1	Does the program have oversight practices that provide sufficient knowledge of grantee activities?	B. 1	Same.	
3.BF2	Does the program collect grantee performance data on an annual basis and make it available to the public in a transparent and meaningful manner?	B. 2	Same.	
	Regulatory Based Programs			
3.RG1	Did the program seek and take into account the views of all affected parties (e.g., consumers; large and small businesses; State, local and tribal governments; beneficiaries; and the general public) when developing significant regulations?	Reg. 1	Did the program seek and take into account the views of affected parties including state, local and tribal governments and small businesses in drafting significant regulations?	Wording clarified.
3.RG2	Did the program prepare adequate regulatory impact analyses if required by Executive Order 12866, regulatory flexibility analyses if required by the Regulatory Flexibility Act and SBREFA, and cost-benefit analyses if required under the Unfunded Mandates Reform Act; and did those analyses comply with OMB guidelines?	Reg. 2	Did the program prepare, where appropriate, a Regulatory Impact Analysis (RIA) that comports with OMB's economic analysis guidelines and have these RIA analyses and supporting science and economic data been subjected to external peer review, as appropriate, by qualified specialists?	Minor change.
3.RG3	Does the program systematically review its current regulations to ensure consistency among all regulations in accomplishing program goals?	Reg. 3	Same.	
		Reg. 4	In developing new regulations, are incremental societal costs and benefits compared?	Merged into new 3.RG4.

(Continued From Previous Page)

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
3.RG4	Are the regulations designed to achieve program goals, to the extent practicable, by maximizing the net benefits of its regulatory activity?	Reg. 5	Did the regulatory changes to the program maximize net benefits?	Combines former questions Reg. 4, 5, & 6.
		Reg. 6	Does the program impose the least burden, to the extent practicable, on regulated entities, taking into account the costs of cumulative final regulations?	Merged in to new 3.RG4.
	Capital Assets and Service Acquisition Programs			
3.CA1	Is the program managed by maintaining clearly defined deliverables, capability/performance characteristics, and appropriate, credible cost and schedule goals?			New question, covers old Cap. 1, 2, 3, and 4.
		Cap. 1	Does the program clearly define the required quality, capability, and performance objectives for deliverables and required capabilities/performance characteristics?	Merged into new 2.CA1 and 3.CA1.
		Cap 2.	Has the program established appropriate, credible, cost and schedule goals?	Merged into new 2.CA1 and 3.CA1.
		Cap 3.	Has the program conducted a recent, credible, cost-benefit analysis that shows a net benefit?	Merged into new 2.CA1 and 3.CA1.
		Cap 4.	Does the program have a comprehensive strategy for risk management that appropriately shares risk between the government and contractor?	Merged into new 2.CA1 and 3.CA1.
	Credit Programs			
3.CR1	Is the program managed on an ongoing basis to assure credit quality remains sound, collections and disbursements are timely, and reporting requirements are fulfilled?	Cr. 1	Same.	
		Cr. 2	Does the program consistently meet the requirements of the Federal Credit Reform Act of 1990, the Debt Collection Improvement Act and applicable guidance under OMB Circulars A-1, A-11, and A-129?	Merged into new 3.CR2.

(Continued From Previous Page)

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
3.CR2	Do the program's credit models adequately provide reliable, consistent, accurate and transparent estimates of costs and the risk to the Government?	Cr. 3	Is the risk of the program to the U.S. Government measured effectively?	Combines former Cr. 2 and 3.
	Research and Development Programs			
	R&D programs addressing technology development or the construction or operation of a facility should answer 3.CA1. R&D programs that use competitive grants should answer 3.CO1, CO2 and CO3.			
3.RD1	For R&D programs other than competitive grants programs, does the program allocate funds and use management processes that maintain program quality?	RD. 1	Does the program allocate funds through a competitive, merit-based process, or, if not, does it justify funding methods and document how quality is maintained?	Modified.
		RD. 2	Does competition encourage the participation of new/first-time performers through a fair and open application process?	Covered by 3.CO1.
		RD. 3	Does the program adequately define appropriate termination points and other decision points?	Covered by 2.CA1 and 3.CA1.
		RD. 4	If the program includes technology development or construction or operation of a facility, does the program clearly define deliverables, capability/performance characteristics, and appropriate, credible cost and schedule goals?	Covered by 2.CA1 and 3.CA1.
IV. Program results				
4.1	Has the program demonstrated adequate progress in achieving its long-term performance goals?	1	Has the program demonstrated adequate progress in achieving its long-term outcome goal(s)?	Minor change.
4.2	Does the program (including program partners) achieve its annual performance goals?	2	Same.	
4.3	Does the program demonstrate improved efficiencies or cost effectiveness in achieving program goals each year?	3	Same.	
4.4	Does the performance of this program compare favorably to other programs, including government, private, etc., with similar purpose and goals?	4	Does the performance of this program compare favorably to other programs with similar purpose and goals?	Minor change.

(Continued From Previous Page)

	This year's question (fiscal year 2005 PART)		Last year's question (fiscal year 2004 PART)	Comment
4.5	Do independent evaluations of sufficient scope and quality indicate that the program is effective and achieving results?	5	Same.	
	Specific Results Questions by Program Type			
	Regulatory Based Programs			
4.RG1	Were programmatic goals (and benefits) achieved at the least incremental societal cost and did the program maximize net benefits?		Same.	
	Capital Assets and Service Acquisition Programs			
4.CA1	Were program goals achieved within budgeted costs and established schedules?	Cap. 1	Same.	
	Research and Development Programs			
	R&D programs addressing technology development or the construction or operation of a facility should answer 4.CA1.	RD. 1	If the program includes construction of a facility, were program goals achieved within budgeted costs and established schedules?	Simplified.

Source: OMB Web site, http://www.whitehouse.gov/omb/part/bpm861.pdf (downloaded Apr. 7, 2003), 6-12.

Development of PART

Fiscal Year 2003

This administration's efforts to link budget and performance began with the fiscal year 2003 budget, in which the administration announced the "Executive Branch Management Scorecard," a traffic-light grading system to report the work of federal agencies in implementing the President's Management Agenda's five governmentwide initiatives. Each quarter, OMB assessed agencies achievement toward the "standards of success"—specific goals articulated for each of the five initiatives. Since some of the five initiatives require continual efforts, OMB also assessed agencies' progress toward achieving the standards. The fiscal year 2003 President's Budget also included OMB's assessments of the effectiveness of 130 programs and a brief explanation of the assessments. According to OMB, the assessments were based on OMB staff's knowledge of the programs and professional judgments; specific criteria were not publicly available with which to support OMB's judgments.

Fiscal Year 2004

During the spring of 2002, an internal OMB task force—PET—consisting of staff from various OMB divisions, created PART to make the process of rating programs robust and consistent across government programs. During the development of PART, OMB solicited input from interested parties both inside and outside the federal government, including GAO and congressional staff. PART was tested on 67 programs during a series of Spring Review meetings with the OMB Director. Based on these results and other stakeholder feedback, PET recommended a series of refinements to PART, such as using a four-point scale in the Results section as opposed to the "yes/no" format. Another key change was revising the Program Purpose and Design section (Section I) to remove the question "Is the federal role critical?" because it was seen as subjective—based on an individual's political views.

In July 2002, OMB issued PART in final and accompanying instructions for completing the assessments for the President's fiscal year 2004 budget submission. Later that month, OMB provided a series of training sessions on PART for staff from OMB and agencies. Agencies received completed PART assessments during early September 2002 and submitted written appeals to OMB by mid-September. OMB formed the IRP, comprising OMB and agency officials, to conduct consistency reviews[1] and provide recommendations on selected PART appeals. The IRP also provided OMB with a broad set of recommendations aimed at improving the PART based on IRP's experience with the consistency audit and appeals. OMB was to finalize all PART assessments by the end of September 2002, although both agency and OMB officials told us that changes and appeals continued through the end of the budget season. RMOs within OMB provided draft summaries of PART results to the Director of OMB during the Director's review of agencies' budget requests. The President's fiscal year 2004 budget (issued February 3, 2003) included a separate volume containing one-page summaries of the PART results for each of the 234 programs that were assessed.[2]

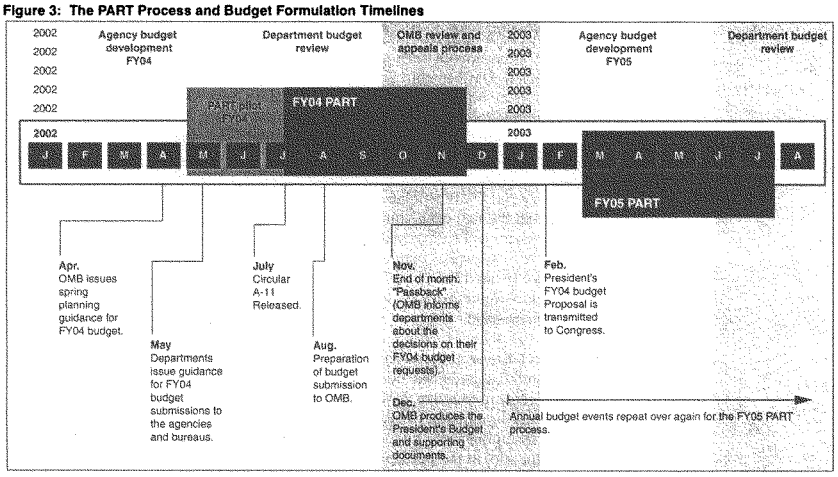
The relationship between PART and the administration's proposals was presented in agencies' budget justification materials sent to Congress. In an unprecedented move, OMB also posted PART, one-page rating results, and detailed supporting worksheets on its Web site. OMB also included its Web address in the Performance and Management Assessments volume of the budget and, in the budget itself, also described PART and its process and asked for comments on how to improve PART.

Figure 3 depicts a time line of the events related to the formulation of the President's budget request, including the key stages of PART development.

[1] According to OMB, IRP performed consistency reviews on a stratified random sample of programs that completed the PART in preparation for the fiscal year 2004 budget. While IRP made recommendations regarding its findings, it did not have the authority to enforce them.

[2] *Fiscal Year 2004 Budget of the United States Government, Performance and Management Assessments*, (Washington, D.C.: February 2003).

Figure 3: The PART Process and Budget Formulation Timelines



Source: GAO

Fiscal Year 2005

For the fiscal year 2005 PART, OMB moved the entire assessment process from the fall to spring. OMB told us that the change was meant to help alleviate the burden of having the PART process overlap the end of the budget season, when workload is already so heavy. Another difference between the 2 years was that agency officials reported that OMB was more collaborative with the agencies in selecting the programs for the fiscal year 2005 PART.

Training on the PART assessments to be included in the President's fiscal year 2005 budget began in early May 2003. Agencies submitted PART appeals in early July, and OMB aimed to resolve the appeals and finalize the PART scores by the end of July. In December of 2003, RMOs were to finalize the summaries of PART results, which will be published in February along with the fiscal year 2005 President's Budget.

Appendix IV

Comments from the Office of Management and Budget

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF MANAGEMENT AND BUDGET
WASHINGTON, D.C. 20503

DEPUTY DIRECTOR
FOR MANAGEMENT

January 16, 2004

Mr. Paul Posner
Managing Director
Federal Budget and Intergovernmental Relations
General Accounting Office
441 G Street, NW
Washington, DC 20548

Dear Mr. Posner:

Thank you for the opportunity to comment on the draft GAO report on the PART (*Performance Budgeting. Observations on the use of OMB's Program Assessment Rating Tool for the Fiscal Year 2004 Budget*, GAO-04-174).

We appreciate GAO's extensive review of the PART process. We are particularly pleased that your report recognizes the unprecedented transparency of the PART process and materials that we have posted on our website[1] and the extensive efforts OMB has taken to make the PART process consistent across the government. We will continually strive to make the PART as credible, objective, and useful as it can be and believe that your recommendations will help us do that. As you know, OMB is already taking actions to address many of them. For instance:

- With respect to centrally monitoring PART recommendations, we have provided a simple format for agencies to follow when reporting the status of recommendation implementation to OMB and I receive these reports semi-annually. We will continue to refine this process so that sufficient attention is given to recommendation follow-up.
- As the PART relies on separate evaluations of evidence of a program's success, we agree that the judgment about what constitutes a sufficient evaluation should be based on the quality of the evaluation.
- Except for programs of insignificant size or impact, we are committed to assessing 100 percent of programs using the PART. We are sensitive to the impact this will have on OMB staff workload and will manage it accordingly.
- One of the greatest opportunities for the PART is to compare the performance of, and share best practices among, like programs across government. We will continue to use the PART for that purpose.
- We are working diligently to generate the meaningful dialogue with Congress you describe in your recommendations.

[1] *See*, draft report Appendix III, p. 2.

- We will continue to improve agency and Executive Branch implementation of GPRA by insisting GPRA plans and reports meet the requirements of this important law and the high standards set by the PART.

Your report makes valuable conclusions and recommendations about the PART and our overall effort to create a more results-oriented government. I want to note that the PART was designed for and is used in many ways other than just budget formulation. Performance information gleaned from the PART process has not only informed budget decisions, but has also helped direct management, identified opportunities to improve program design, and promoted accountability. We believe that the PART will also greatly improve the goals and measures adopted through the GPRA strategic and performance planning processes

Thank you for the opportunity to review and comment on your draft report. I appreciate your willingness to take our oral and written comments into consideration in the final draft. I look forward to working with you to improve the ways in which we are creating a results-oriented government.

Sincerely,

Clay Johnson III

Appendix V

GAO Contacts and Staff Acknowledgments

GAO Contacts

Paul Posner, (202) 512-9573
Denise Fantone, (202) 512-4997

Acknowledgments

In addition to the above contacts, Kristeen McLain, Jackie Nowicki, and Stephanie Shipman made significant contributions to this report. Thomas Beall, Joseph Byrns, Hashem Dezhbakhsh, Evan Gilman, Patrick Mullen, David Nicholson, and Mark Ramage also made key contributions to this report.