GPRA@DOE.SC

Report to NSAC on the BESAC Subpanel and Related Matters

J. Schiffer March 14, 2002

OUTLINE

1. What is GPRA *(briefly)*

2. The 2001 COSEPUP report *(briefly)*

3. The BESAC Subpanel on GPRA in the Office of Science of DOE – what happened there for the FY2003 budget *(briefly)*

4. The Report of the Subpanel

5. The OMB Feb. 2002 ‘Discussion Draft’ *(issued after the Panel Report)*

Government Performance and Results Act of 1993 (GPRA)

Findings and Purpose:

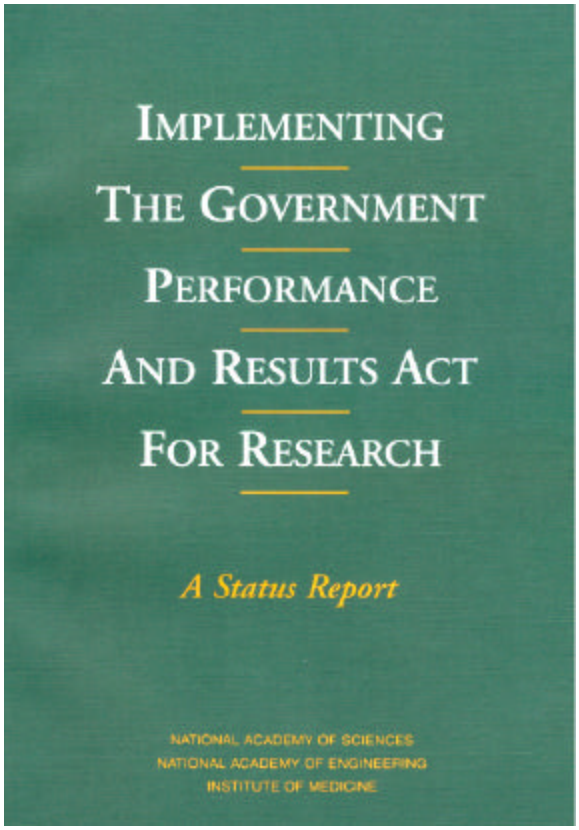
?? Waste and inefficiency
?? Insufficient program goals and inadequate information on program performance
?? Congressional policy … handicapped

Strategic Planning:

?? Comprehensive mission statement for each agency
?? Cover at least 5 years, revised at least every 3

Annual Performance Plans:

?? Establish objective, quantifiable, measurable goals
?? Establish performance indicators
?? If not feasible, include separate descriptive statements, or
?? Such alternative as authorized by the Director of OMB



Conclusion 3: The most effective technique for evaluating research programs is review by panels of experts using the criteria of quality, relevance, and, when appropriate, leadership. Agency approaches to GPRA research programs demonstrate the utility of expert review using the same criteria of quality and relevance as outlined in COSEPUP's original report. The international leadership criterion is generally not evaluated by most federal agencies at this time, although several are interested in such a measure. However, given the diversity in mission, complexity, culture, and structure of federal agencies that support research, it is not surprising that their approaches to GPRA have varied. One size definitely does not fit all.

Conclusion 6: The development of human resources as an agency objective sometimes does not receive explicit emphasis or visibility in GPRA plans and reports.

Conclusion 7: Agencies often receive conflicting messages from oversight bodies about the desired format, content, and procedures to be used in GPRA compliance.

Recommendation 1: Federally supported programs of basic and applied research should be evaluated regularly through expert review, using the performance indicators of quality, relevance, and, where appropriate, leadership.

Recommendation 2: Agencies should continue to improve their methods of GPRA compliance and to work toward the goals of greater transparency, more-realistic reporting schedules, clear validation and verification of methods, and the explicit use of human-resources development as an indicator in performance plans and reports.

Recommendation 3: Agencies and oversight bodies should work together as needed to facilitate agencies integrating their GPRA requirements with their internal planning, budgeting, and reporting processes. In addition, they should work together to adjust the timing of GPRA reporting to capitalize on the value of the planning process.

Recommendation 4: Agencies should strive for effective communication with oversight groups on the implementation of GPRA. For their part, oversight bodies should clarify their expectations and meet more often among themselves to coordinate their messages to agencies.

Purpose of Panel

To review:

?? **The Office of Science's methods of performance measurement;**
?? **The appropriateness and comprehensiveness of the methods;**
?? **The effects on science programs; and**
?? **SC's integration of performance measurements with the budget process as required by GPRA.**

Agenda
January 24, 25 meeting

Overview	**Decker SC**
Performance Measurement in SC	**Dehmer SC**
Integrating Performance Measurement & the Budget in SC	**Valdez SC**
Integrating GPRA and the Budget Process – General	**Powers CFO**
Administration Expectations	**Holland OMB**
Congressional Expectations	**Nazzaro GAO**

Round Table Panels

1. **Effects of Performance Measurement on Science Programs in SC**
2. **Effects of Performance Measurement on Facility Construction & Operation**
3. **Alternatives and Other Agency Experiences**

from

Science Corporate Context

in front of FY2003 DOE Budget submission

Strategic Objectives

SC2: By 2015, describe the properties of the nucleon and light nuclei in terms of the properties and interactions of the underlying quarks and gluons; by 2010, establish whether a quark gluon plasma can be created in the laboratory and, if so, characterize its properties; by 2020, characterize the structure and reactions of nuclei at the limits of stability and develop theoretical models to describe their properties, and characterize using experiments in the laboratory the nuclear processes within stars and supernovae that are needed to provide an understanding of nucleosynthesis. *(NP)*

SC7: **Provide major advanced scientific user facilities where scientific excellence is validated by external review; average operational downtime does not exceed 10% of schedule; construction and upgrades are within 10% of schedule and budget; and facility technology research and development programs meet their goals.** ***(Crosscutting all major programs.)***

from the Budget Summary for the Office of Science

Nuclear Physics

Program Mission

The mission of the Nuclear Physics (NP) program is to foster fundamental research in nuclear physics that will provide new insights and advance our knowledge on the nature of matter and energy and develop the scientific knowledge, technologies and trained manpower that are needed to underpin the DOE's missions for nuclear-related national security, energy, and environmental quality. The Program provides world-class, peer-reviewed research results and operates user accelerator facilities in the scientific disciplines encompassed by the NP mission areas under the mandate provided in Public Law 95-91 that established the Department.

Strategic Objectives

SC2: By 2015, describe the properties of the nucleon and light nuclei in terms of the properties and interactions of the underlying quarks and gluons; by 2010, establish whether a quark-gluon plasma can be created in the laboratory and, if so, characterize its properties; by 2020, characterize the structure and reactions of nuclei at the limits of stability and develop the theoretical models to describe their properties, and characterize using experiments in the laboratory the nuclear processes within stars and supernovae that are needed to provide an understanding of nucleosynthesis.

SC7: Provide major advanced scientific user facilities where scientific excellence is validated by external review; average operational downtime does not exceed 10% of schedule; construction and upgrades are within 10% of schedule and budget, and facility technology research and development programs meet their goals.

Progress toward accomplishing these Strategic Objectives will be measured by Program Strategic Performance Goals, Indicators and Annual Targets, as follows:

from the Budget Summary for the Office of Science

Nuclear Physics

SC2-2: Determine the behavior and properties of hot, dense nuclear matter as a function of temperature and density. Discover and characterize the quark-gluon plasma. (Heavy Ion Nuclear Physics and Nuclear Theory subprograms)

Performance Indicators

Results of external and internal reviews of quality, relevance and leadership of research activities and facility operations number of significant scientific discoveries.

Performance Standards

As discussed in Corporate Context/Executive Summary.

Annual Performance Results and Targets

FY 2001 Results	FY 2002 Targets	FY 2003 Targets
Produced first heavy-ion collisions at the Relativistic Heavy Ion Collider (RHIC - construction completed FY1999) at 10% of its design luminosity, as planned, with four experimental detectors. Published first results of heavy-ion collisions. [Met Goal]	Complete first round of experiments at RHIC at full energy; achieve the full design luminosity (collision rate) of 2 x 10^{26} cm^{-2} s^{-1} for heavy ions. (SC2-2)	Initiate first round of experiments with collisions with other ions to compare to results of gold-gold collisions. (SC2-2)
Continued major accelerator improvement projects at RHIC in order to improve machine reliability and efficiency. [Met Goal]	Complete Helium Storage addition and liquid nitrogen standby cooling system at RHIC leading to better cost effectiveness ($ DELETED savings) and operational efficiency (10% increase). (S2-2/SC7-2)	Upgrade the RHIC cryogenics system by replacing turbine oil skids and removing seal gas compressor, eliminating a single point failure. (S2-2/SC7-2)

from the Budget Summary for the Office of Science

High Energy Physics

Annual Performance Results and Targets

FY 2001 Results	FY 2002 Targets	FY 2003 Targets
Delivered sufficient luminosity (25 fb-1) to double total BaBar data set. Added one new RF station.	Increase the total data recorded by BaBar at the SLAC B-factory by delivering 35 fb-1 of total luminosity. (SC1-2)	Increase the total data delivered to BaBar at the SLAC B-factory by delivering 50 fb-1 of total luminosity. (SC1-2)
	Add one new RF station. (SC1-2)	{DELETED}
BaBar collaboration published first unambiguous observation of CP violation in B meson decays. Errors on the measurement are +/- 0.15.	Measure CP violation in B mesons with an uncertainty of +/- 0.12. (SC1-2)	Measure CP violation in B mesons with an uncertainty of +/- 0.10. (SC1-2)

National Institutes of Health Research Subgoals
Government Performance and Results Act

FY 1999-2001	FY 2002-2003
Goals	Subgoals
A: Add to the body of knowledge of normal and abnormal biological processes and behavior	A1) Discover innovative approaches for identifying and measuring genetic and environmental factors that contribute to common, complex diseases across populations.
	A2) Develop model systems (animal models, cell lines, etc.) that will advance our understanding of disease processes.
B: Develop new or improved instruments and technologies.	B1) Develop new technologies to enable greater understanding of genomic and proteomic information.
	B2) Develop biocompatible materials for use in replacing or repairing damaged and non-functioning or missing tissue.
C: Develop new or improved approaches for preventing or delaying the onset or progression of disease and disability.	C1) Identify modifiable risk factors for disease/disability.
	C2) Identify, develop, and test new/improved medications for the prevention of disease/disability.
D: Develop new or improved methods for diagnosing disease and disability.	D1) Develop and apply powerful new imaging, genetic, and biological technologies to enable early and more precise diagnosis and intervention.
	D2) Identify and apply knowledge about factors, including gender, race, ethnicity, and socioeconomic status, to improve diagnostic reliability and treatment response.
E: Develop new or improved methods for treating disease and disability.	

NSF Strategic Outcomes

- Evaluate programs regularly through expert review, using COSEPUP's 'Quality, Relevance, Leadership
  (International Leadership not evaluated at this time)
- NSF Goals for Strategic Outcomes:
  - **People –** *"a diverse, internationally competitive and globally engaged workforce of scientists, engineers, and well-prepared citizens".*
  - **Ideas –** Enabling *"discoveries across the frontier of science and engineering, connected to learning, innovation and service to society."*
  - **Tools –** Providing *"Broadly accessible, state-of-the-art shared research and education tools."*

Outcome Goals are Mostly Qualitative

- NSF awards lead to important discoveries; new knowledge … as judged by independent external experts.
- Judged either successful or unsuccessful for the program.
- Results from prior investment accepted in the year reported
- NSF is successful when,
  - *Growing fundamental knowledge enhances progress*
  - *Discoveries advance frontiers of science, engineering …*
  - *Partnerships connect discovery to innovation, learning, and societal advancement*
  - *Research & education processes are synergistic*

Report
of the Subpanel on

Performance Measurement
in the Office of Science

U.S. Department of Energy

Eugene Bierly	*American Geophysical Union*	**BER**
Roscoe Giles	*Boston University*	**ASCR**
Fred Gillman	*Carnegie Mellon University*	**HEPAP**
John Roberts	*California Institute of Technology*	**BES**
Ned Sauthoff	*Princeton Plasma Physics Lab.*	**FES**
John Schiffer	Argonne Nat. Lab. and U. of Chicago	**NSAC**
John Stringer	*EPRI*	**chair**
Nicholas Vonortas	*George Washington Univ.*	**Economics**

Introduction

…….

New knowledge that leads to a better understanding of our world is the direct product of basic scientific research. It is widely recognized that advances in basic science also underlie and propel developments in applied science and technology that are needed for national security, economic competitiveness, new sources of energy, the environment, and improved health care in the United States.

Over the past 50 years, the Department of Energy and its predecessor agencies have been a major source of support for long-range basic research programs in the United States, especially in the physical sciences. Many of these have been and continue to be the envy of the world.

DOE is a mission-oriented agency. The Department, through its Office of Science, supports research at both its National Laboratories and at universities. As part of its mission, the DOE constructs and operates major user facilities (light sources, neutron sources, and a range of accelerators) that are essential to the research communities across a broad range of basic and applied sciences. The part of DOE's mission that relates to the Office of Science is best described by the following Goal, which is taken from the FY 2003 Congressional Budget Request:

"Deliver the scientific knowledge and discoveries for DOE's applied missions; advance the frontiers of the physical sciences and areas of the biological, environmental, and computational sciences; and provide world-class research facilities and essential scientific human capital to the Nation's overall science enterprise."

Recommendations

1. ***The Subpanel recommends that the Office of Science complete its Strategic Plan as soon as possible.***

   This is a key part of the GPRA process, and is particularly important in relation to developing criteria for basic research, since its five-year scope allows for longer-term planning, and the review on a three-year basis allows for the introduction of new discoveries into the research planning process. Furthermore, since the Strategic Plan must relate to the Office's Science Goal, and through that to DOE's mission, this gives criteria against which the 'relevance' criterion can be measured.

2. ***The Subpanel recommends that the general principles of the performance assessment methods that have been used by the Office of Science in the past should continue to be followed***.

   The success of the Office of Science in maintaining a very effective program of world-class research and the development of a significant number of world-class facilities has been recognized by independent reviews in the recent past.

3. ***The Subpanel recommends that the Office of Science's performance measurement criteria be aligned with those that have been developed by the National Academies' Committee on Science, Engineering, and Public Policy (COSEPUP), and with their ongoing studies on the development of criteria for Basic Research, to allow a common basis for the different Federal Agencies that support basic research programs.***

4. ***The Subpanel recommends that the discussions between the Office of Science and the Office of Management and Budget as to appropriate criteria for the assessment of the progress of basic science programs be continued, to allow the development of appropriate metrics.***

   This should take into account the considerable qualitative component in measuring the quality of basic research, and the intrinsically longer time scales involved. To the extent that other Federal Agencies are supporting basic research, discussions should include considerations as to the extent to which similar methods of assessment might be appropriate.

5. ***The Subpanel recommends that criteria to assess the 'world leadership' element in the assessment of the Office of Science's research should be developed.***

6. ***The Subpanel recommends that work-force issues, including the development of succession plans for the research staffs, and the education and training of a technically sophisticated personnel reservoir for the future of the nation, be incorporated into the GPRA goals of the Office of Science.***

   The DOE should describe in their strategic and performance plans the goal of developing and maintaining adequate human resources in fields critical to their mission. Human resources should become a part of the evaluation process.

Analysis of the Major Issues

The research program of the Department of Energy's Office of Science is an outstanding program, which has been remarkably successful in advancing basic research in the U.S., developing world-leading research in a number of important areas, and developing both an important research infrastructure and a remarkable set of major user facilities. **The processes being developed in the GPRA management plans should help to make these contributions better understood by the stakeholders; and assist the Office of Science in managing the existing program and developing the case for further advancements.** All the parties involved in this exercise are in alignment with this view, and are trying to develop procedures which will help improve this valuable program, and avoid introducing processes which would harm it. Our discussion points below are not intended to criticize any of the contributors to this exercise, but to help in pointing directions that seem to us to need attention. In particular, our concern is with the development and maintenance of a world-leading program in basic research within a mission-oriented agency.

The Use of Assessment Methods by the Office of Science

.....

Recent Experience in the Application of GPRA to the Office of Science

The Subpanel was shown some of the procedures for the implementation of GPRA in the 2003 Budget Submission for the Office of Science. The Subpanel was told that it is recognized by most of the participants that it did not meet several of the GPRA requirements, including for example that the program descriptions should give a comprehensive description of the program.

The overall Science Goal is supported by eight Strategic Objectives. Each of these has related Program Strategic Performance Goals (PSPGs): there are a total of 22 of these. The Subpanel members from the Office of Science Advisory Committees considered that the set of these for the parts of the programs with which they are familiar **distorted the aims and accomplishments of SC research programs**. With PSPGs that are only representative and not at all comprehensive, the Office's programs are portrayed as significantly less than they truly are. ….. The Budget Submission then fails as an effective communication tool, which is one of its most important roles.

….. the Subpanel believes that the opening Executive Summary should be consistent with the GPRA wording that "an agency may aggregate, disaggregate, or consolidate program activities, except that any aggregation or consolidation may not omit or minimize the significance of any program activity constituting a major function or operation for the agency."

Quantitative and Qualitative Assessment Criteria

... There is no doubt that it is easier for the sort of comparative assessments that have to be made in a budgeting process if the annual results of the programs can be expressed in objective quantitative terms; but it is clear from the description of the peer review process above that these assessments are generally qualitative; attempts to make them quantitative, for example by making reviewers score projects on a scale of one to ten, is artificial, and scarcely objective. GPRA requires the plan to "establish performance goals to define the level of performance to be achieved by a program activity;" and "to express such goals in an objective, quantifiable, and measurable form *unless authorized to be in an alternative form under section (b)*". (Our italics). Such an alternative form may be authorized by the Director of the Office of Management and Budget.

The Subpanel believes that much basic research is better assessed in qualitative terms. While this offers challenges to the concept of being 'measurable' this should not lead to the imposition of quantitative goals. To do this would have significant negative effects on basic research, and would certainly not be consistent with the principle that application of GPRA should "do no harm"; a principle which is agreed to by all the participants in this exercise. In its ongoing discussions with OMB, this issue should be reviewed.

Experience in Other Related Federal Agencies

Other Federal Agencies also support basic research, to a greater of lesser degree, notably the National Science Foundation (NSF), the National Institutes of Health (NIH), the National Aeronautics and Space Administration (NASA), and the Department of Defense. The Subpanel heard presentations from NIH and NSF outlining their responses to the GPRA directives. All of these agencies are different, and the Subpanel recognizes that this will lead to differences in the ways in which OMB will wish to see the performance assessed.

However, there will be some overlap in the character of specific basic research programs, and the Subpanel believes that it would be worthwhile in the Office of Science's ongoing discussions with OMB on procedures for this aspect to be reviewed in relation to the development of appropriate goals and metrics.

OMB Preliminary Investment Criteria for Basic Research

Background

Predicting the outcome of worthwhile basic research should not be easy. Serendipitous results are often the most interesting and ultimately may have the most value. Taking risks and working towards difficult-to-attain goals are important aspects of good research management.

However, there is no inherent conflict between the difficulty of predicting the success of basic research programs and the call for R&D investment criteria and budget-performance integration in the President's Management Agenda. Bringing clearer information about program performance to bear upon resource allocation decisions lies at the heart of these initiatives. The Administration will focus on improving the management of basic research programs, not on predicting the unpredictable. Reinforcing good management practices and the adoption of best practices by all basic research programs across the federal government is the goal. Not all programs will meet all the criteria initially, but we expect that over time they will.

Principles behind proposed OMB investment criteria for basic research programs

OMB proposes using the guidelines: Quality, Relevance and Performance, combining both the COSEPUP and ARL models. OMB retains COSEPUP's Leadership concept, but as a potential indicator to demonstrate Quality and not as an independent goal. Adapting the ARL Productivity metric provides a means of coupling investment criteria for basic research programs to the President's Management Agenda initiative for budget and performance integration.

The criteria include clear distinctions between prospective assessment and retrospective evaluation. Research agencies fund a mix of contracts, grants, and in-house activities, which means that program management often entails placing "bets," monitoring contractors, and managing internal research activities. It is tremendously important that basic research programs are able to demonstrate responsible management of their inputs, in addition to clearly articulating and demonstrating progress towards expected outputs. Yet, outcomes still matter. Retrospective review of whether investments were productive is essential for validating program design and instilling confidence that future investments will be wisely invested as well. Retrospective reviews should address both technical excellence and the relevance of program outputs to others. In practice, Quality, Relevance, and Performance are more readily demonstrated separately for prospective information but are highly interrelated in the retrospective analysis.

Next Steps

- ***Discuss criteria with stakeholders in government, academia, and industry.***

- ***Validate general criteria against relevant programs at specific agencies.***

- ***Develop larger strategy and framework for an integrated approach assessing basic and applied programs across the agencies.***

- ***Determine how, where, when, and at what levels the criteria will be applied.***

- ***Work with agencies to determine how best to implement criteria at each agency.***

- ***Provide more detailed guidance to agencies to help them develop more meaningful performance metrics.***

Draft Criteria for Basic Research

All basic research programs must meet all of the following criteria. (Appropriate levels of applicability remain to be determined.)

1. Quality

A. Prospective Review of Awards

- Clearly define how much of the requested funding will be directed to specific research performers, open to a limited subset of research performers through merit-reviewed competitions, or open to all potential research performers through merit-reviewed competitions
- Provide a compelling justification for research funding that is to be directed to specific performers or open to a limited subset of research performers.

B. Retrospective Expert Review of Program Quality

- Clearly define a plan for regular, external reviews of the quality of the program's research and research performers. Explain how the results from these reviews will be used to guide future program decisions. Rolling reviews performed every 3-5 years by advisory committees can satisfy this requirement. Benchmarking is an effective means of assessing program quality relative to other programs, other agencies, and other countries.
- Provide a clear response to prior external reviews, including whether and how the program has responded or will respond to recommendations from those reviews.

Discussion -- Quality

Programs should **maximize quality** of the research they fund and periodically examine whether their portfolio of projects produces scientific and technical **excellence**.

Programs must use a clearly stated, defensible method for awarding a significant majority of their research grants and contracts. The customary method for promoting research quality is the use of a competitive, merit-based process. NSF's process for the peer-reviewed, competitive award of its research grants is a good example. Exceptions must be well justified, but they may include concerns for timeliness (e.g., research grants for rapid response studies of *Pfisteria*) or a proven record of outstanding performance (e.g., performance-based renewals).

Programs must assess and report on the quality of current and past research. For example, NSF's Committees of Visitors, which reviews NSF directorates, are one implementation of these reviews. Benchmarking programs in one agency against other federal programs is encouraged, as is international benchmarking, which provides a measure of leadership in a field of research. Leadership is the result of a well-defined, high-quality, well-managed program, one able to make identifiable contributions to a field within the confines of available resources. Not delivering world-class performance within existing resources is indicative of poor program execution, timidity, or an overly broad portfolio.

2. Relevance

A. Definition of Program Direction and Relevance

- .... Plans must identify and prioritize research goals within the program. Even programs that fund unsolicited proposals for investigator-initiated research should be able to clearly articulate what new knowledge, understanding, technology, or tools might result from the investment.
- .....

- For capital projects, include schedules with milestones for major future competitions, decisions, and termination points. Highlight any changes from previous schedules.

B. Retrospective Outcome Review to Assess Program Design and Relevance

- Provide a clear plan for external reviews of the program's relevance to future research, including how results from these reviews will be used to guide future program decisions. Committees should be composed of peers, experts in related fields, and potential users of research results. Retrospective reviews should be conducted on a 3-5 year, rolling basis.
- .....

Discussion – Relevance

Research programs need to set a general direction for their investments. These plans must identify and prioritize research opportunities within the program. **The Joint DOE/NSF Nuclear Sciences Advisory Committee's Long Range Plan and the Astronomy Decadal Surveys are the products of good planning processes.** Workshops may be an acceptable planning mechanism for "small science" programs.

Program objectives and goals should be assessed by their relevance to agency missions, national needs, and the field(s) of study the program intends to address. Review committees should provide an answer to the question: "Does the agency's research address subjects in which new understanding could be important in fulfilling the agency's mission?"

An example of a responsibility that goes beyond agency bounds is the operation of scientific facilities for the use and benefit of the entire research community. Committees will address whether programs are fulfilling these responsibilities appropriately or whether other stakeholders should bear those burdens.

3. Performance

(Metrics and assessments should be reflected in the agency's annual performance plan.)

A. Prospective Assessment of Program Inputs and Output Performance Measures

- ... Where possible, programmatic risk on individual capital projects should be quantified.
- ...
- For operational facilities, define appropriate metrics for the dependability, effectiveness, and use of those facilities over time.

B. Demonstration of Performance

- Document performance against previously defined output metrics, including any success in reaching one or more multi-year objectives.
- If a proposal includes significant capital projects, provide external, independent cost and schedule estimates. .....
- For operational facilities, report on metrics for dependability, effectiveness, and use.

Discussion – Performance

Programs must demonstrate attentiveness to the health of their research enterprise and an ability to manage their programs in a manner that produces identifiable results.

Input statistics help demonstrate to oversight bodies that agency heads are managing the inputs of their research enterprise. The range of these statistics is highly variable and should be tailored to address issues of concern to agency management and OMB.

Construction projects and facility operations will require additional performance metrics. Cost and schedule earned value metrics for the construction of R&D facilities must be tracked and reported. Formalized independent reviews for DOE's Office of Science of technical cost, scope, and schedule baselines and project management of construction projects ("Lehman Reviews") are a widely recognized "best practice" for discovering and correcting problems involved with complex, one-of-a-kind construction projects.

Any set of specific output milestones is unlikely to cover 100 percent of a program's research portfolio, nor should it. OMB will assume that basic research programs reporting on the results of specific output milestones as developed above combined with reporting the results retrospective portfolio reviews will have satisfied the requirement of GPRA.