

Instructions for the Program Assessment Rating Tool

TABLE OF CONTENTS

GENERAL GUIDANCE.....	2
Overview.....	2
Standards of a <i>Yes</i>	2
Question Weighting.....	2
Relationship of the PART to the Government Performance and Results Act.....	3
Selecting Programs.....	3
The "P" in PART.....	3
Combining Programs for the PART.....	4
PART Crosscuts.....	5
Selecting Performance Measures.....	6
Basic Concepts.....	7
Categories of Performance Measures.....	7
Targets and Baselines.....	9
Long-term v. Annual Performance Goals.....	10
Sections of the PART.....	11
PART Instrument Linkages.....	11
Types of Programs.....	12
I. PROGRAM PURPOSE & DESIGN.....	13
II. STRATEGIC PLANNING.....	18
Specific Strategic Planning Questions by Program Type.....	28
Regulatory Based Programs.....	28
Capital Assets and Service Acquisition Programs.....	29
Research and Development Programs.....	29
III. PROGRAM MANAGEMENT.....	30
Specific Program Management Questions by Program Type.....	36
Competitive Grant Programs.....	36
Block/Formula Grant Programs.....	38
Regulatory Based Programs.....	39
Capital Assets and Service Acquisition Programs.....	41
Credit Programs.....	42
Research and Development Programs.....	43
IV. PROGRAM RESULTS/ACCOUNTABILITY.....	44
Specific Results Questions by Program Type.....	48
Regulatory Based Programs.....	48
Capital Assets and Service Acquisition Programs.....	49
Research and Development Programs.....	49
APPENDIX A: R&D INVESTMENT CRITERIA.....	50

GENERAL GUIDANCE

OVERVIEW

The Program Assessment Rating Tool (PART) is a systematic method of assessing the performance of program activities across the Federal government. The PART is a diagnostic tool; the main objective of the PART review is to improve program performance. The PART assessments help link performance to budget decisions and provide a basis for making recommendations to improve results.

The PART is composed of a series of questions designed to provide a consistent approach to rating programs across the Federal government, relying on objective data to assess programs across a range of issues related to performance. The PART also examines factors that the program or agency may not directly control but may be able to influence. For example, if statutory provisions impede effectiveness, legislative changes may be proposed. The formalization of performance assessments through this process is intended to develop defensible and consistent program ratings.

The questions that comprise the PART are generally written in a *Yes/No* format. They require the user to explain the answer briefly and to include relevant supporting evidence. Responses must be evidence based and not rely on impressions or generalities. The completed PART will be made available for public scrutiny and review. A *Yes* answer must be definite and reflect a high standard of performance. Where hard evidence is unavailable, assessments will rely more on professional judgment. No one question will determine a program's assessment; and in some instances, *Not Applicable* may be an appropriate answer.

This guidance document and the worksheets used to complete the assessments will be available on OMB's website at <http://www.omb.gov/part/> the week of March 22, 2004.

STANDARDS OF A YES

The PART holds programs to high standards. Simple adequacy or compliance with the letter of the law is not enough. Rather, a program must show it is achieving its purpose and that it is well managed. The PART requires a high level of evidence to justify a *Yes* response. Sections I through III are scored in a *Yes/No* format. In Section IV, a four-level scale (*Yes, Large Extent, Small Extent, and No*) permits answers to reflect partial achievement of goals and evidence of results. Answers must be based on the most recent credible evidence.

QUESTION WEIGHTING

As a default, individual questions within a section are assigned equal weighting that total 100 percent for each section. However, the user may alter the question weighting to emphasize key factors of the program. To avoid manipulation of the total score, weights

must be adjusted prior to responding to any questions. If a question is not relevant to the program, the user may rate it as *Not Applicable*. In these cases, the user would assign a weight of zero to the question but must provide an explanation of this response.

RELATIONSHIP OF THE PART TO THE GOVERNMENT PERFORMANCE AND RESULTS ACT (GPRA)

The GPRA statute provides a framework under which agencies prepare strategic plans, performance plans, and performance reports that set goals and report on the extent to which they are achieved. The PART is a systematic method of assessing performance of program activities, focusing on their contribution to an agency's achievement of its strategic and program performance goals.

The PART strengthens and reinforces performance measurement under GPRA by encouraging careful development of performance measures according to the outcome-oriented standards of the law and by requiring that agency goals be appropriately ambitious. . Therefore, performance measures included in GPRA plans and reports and those developed or revised through the PART process must be consistent.

The PART also extends the usefulness of performance information by defining programs that comprise activities about which management and budget decisions are made. As a matter of sound manage practice, agencies will integrate operational decisions with strategic and performance planning by:

- improving performance measures over time through the PART review,
- aligning budgets with programs, and
- aligning programs and measures with GPRA goals.

SELECTING PROGRAMS

The “P” in PART

One feature of the PART process is flexibility for OMB and agencies to determine the unit of analysis - “program”- for PART review. The structure that is readily available for this purpose is the formal budget structure of accounts and activities supporting budget development for the Executive Branch and the Congress and, in particular, Congressional appropriations committees. This structure is also linked to financial reporting requirements throughout the Executive Branch. Although the budget structure is not perfect for program review in every instance – for example, “program activities” in the budget are not always the activities that are managed as a program in practice – the budget structure is the most readily available and comprehensive system for conveying PART results transparently to interested parties throughout the Executive and Legislative Branches, as well as to the public at large. If OMB and the agency agree, programs can be identified as aggregations of program activities in budget accounts, or in other meaningful budget breakouts used for display to Executive and Legislative Branch decision makers.

Regardless of the program aggregation selected for analysis, agencies must be able to link that unit -- through cross-walks or other clear depictions -- with the agency's budget and GPRA strategic and performance plans / performance budgets. Likewise, agencies must be able to link PARTed program goals to GPRA goals. These linkages are designed to encourage greater use of performance information by agencies to manage programs better. It is at these levels that management and budget decisions are made.

Combining Programs for the PART

The PART should assess the appropriate unit to inform improved program performance and management decisions.

There may be situations in which it makes sense to look at more than one “program” (as recognized in existing structures, such as the Budget) with a single PART. This might be the case when programs share a common goal and are so interdependent that it makes more sense to review them as a unit than separately. For example, several grant programs within the Administration on Aging were combined within a single PART assessment. They each support a similar program goal of providing supports to older individuals to maintain a level of independence that enables them to stay in their homes, rather than move to an institution. Often the same entity receives multiple grants, and it is the combination of these supports that enables individuals to remain at home. The data reporting for these grants is combined and the services are so inter-related, that assessing the programs independently would provide less meaningful information than examining them as a group.

Analysis of the performance of these programmatic goals is more informative when the programs are considered as a package rather than as separate programs. Agencies and OMB should avoid, however, assessing overly aggregated units such that the PART is unable to illuminate meaningful management distinctions among programs that share common goals or forms but are managed differently.

The following list of criteria should be considered in determining whether to combine more than one program for PART assessment:

1. Program Purpose. The proposed combination of programs should share a common purpose and mission. Issues to consider include: program mission, beneficiary characteristics, target populations, grant recipients, etc.
2. Program Design/Administration. Programs should be similarly designed and administered. For example the grantee role, grant type (e.g. competitive, formula), benefit structure, oversight roles, data collection, and/or link to performance measurement should be common across the programs.
3. Budgeting. The identification of programs should relate to budget decisions because one goal of assessing programs is to develop information that will be useful to the budget process. This does not necessarily mean the definition of programs should be restricted to accounts and activities in the budget (i.e., a

single budget line), but rather that the programs are managed as a single unit. If programs are chosen that are not aligned with the budget structure, budget justifications must include a cross-walk between the program(s) and the budget structure for public presentation.

0. **Performance.** The performance measures incorporated into the PART review should be aligned with the GPRA strategic goals. The PART is intended to operationalize GPRA into the agency's management and budget decisions, so budget justifications/performance plans must also demonstrate how the PART programs fit within the GPRA strategic framework. Therefore, programs considered within a single PART should support similar long-term outcome goals. The same conclusions/recommendations on program performance should apply to each program.

In cases where a PART encompasses multiple programs, each program must:

1. Be addressed by each question. **In order to receive a *Yes* on a given question, each of the programs included in the PART must receive a *Yes* to that question unless evidence is provided on why the question is not relevant to one of the programs.**
2. Have related long-term outcome measures. In addition, each program should have annual performance measures. In the case of shared measures, to the extent possible, each program should be able to demonstrate how it contributes to the outcome or output measured.
2. Be addressed in the recommendations.

NOTE: More often programs that have similar goals are not interdependent. In those cases, there may be an advantage to program comparisons in the context of a "crosscut" that reviews the programs' different administration practices and differential contributions to a set of outcomes. **Programs in a crosscut should not be assessed within a single PART.**

PART Crosscuts

Instead of combining **interdependent** programs into a single PART, OMB and agencies may compare **independent** programs that have similar purposes or goals through a "crosscut". An internal crosscut (programs within the same agency) or external crosscut (programs across multiple agencies) would identify exemplary goals and practices, common measures of performance, possible tradeoffs in management and budget decisions, and opportunities for better coordination among programs.

Unlike a combined PART, a crosscut would still examine programs in individual PARTs. An individual assessment would allow recognition of the distinctive features among

programs. However, the following common themes could be used to compare the programs in a meaningful crosscut:

- What is the target population for each program?
- What products and services are provided (common output measures as appropriate)?
- How well are products and services provided (common output efficiency measures as appropriate)?
- What impact have the programs achieved (common outcome measures and common outcome efficiency measures as appropriate)?

An internal crosscut need not require production of additional materials, as each of the aforementioned corresponds to questions already posed in the PART. Individual PART summaries and recommendations would reference results of crosscutting issues as appropriate. The results of crosscutting analyses would summarize common strengths and opportunities for improvement, drawing upon the individual PART assessments. Ultimately, the analysis could lead to a common framework of performance measures and accountability if it doesn't already exist.

OMB and agencies are encouraged to identify opportunities to conduct internal crosscuts. For example, separate PART analyses were conducted for the following Small Business Administration technical assistance programs: Small Business Development Centers, Business Information Centers, and the Service Corps of Retired Executives (SCORE). All of these programs are administered independently and have some distinctive features, but have similar purposes in providing information and education assistance to help entrepreneurs start, grow, and sustain small businesses.

OMB has begun to examine opportunities to conduct external crosscutting analyses in areas such as community and economic development and rural water that extend across agencies.

SELECTING PERFORMANCE MEASURES

The purpose of this section is to provide guidance on how to define or select meaningful performance measures for programs. While this guidance includes measures as examples, they are illustrative and may not be appropriate for every program. **Agencies and OMB must assess the quality of a program's measure--if it is meaningful in the context of a specific program, if the measurement methodology is sound, and if the measure can be verified with reliable data.**

The key to assessing program effectiveness is measuring the right things. Performance measures must capture the most important aspects of a program's mission and priorities, not merely ones for which there are data. Individuals must be able to understand what is being measured and why it is relevant to the program.

The performance measures must reflect program priorities, be few in number, and provide information to inform budget and management decisions. Measures should reflect desired outcomes; however, sometimes a narrower approach is appropriate and output measures or interim milestone outcomes may be acceptable. Because of the importance of performance measures, OMB and agencies must agree on appropriate measures early to allow for review with relevant stakeholders, if needed. If the agency intends to revise its strategic goals as the result of a PART assessment, GPRA requires that relevant stakeholders be involved in the review.

Basic Concepts

The terms below will be used repeatedly throughout this guidance. Although further explained throughout the document, they are defined here as an introduction.

- **Strategic Goal or Strategic Objective:** A statement of aim or purpose included in a strategic plan (required under GPRA). In a performance budget/performance plan, strategic goals should be used to group multiple program outcome goals; the program outcome goals should relate to and in the aggregate be sufficient to influence the strategic goals or objectives and their performance measures.

Performance Goal: Sets a target level of performance over time expressed as a tangible, measurable objective, against which actual achievement can be compared, including a goal expressed as a quantitative standard, value or rate. **A performance goal is comprised of a performance measure with targets and timeframes.**

- **Performance Measures:** Indicators, statistics or metrics used to gauge program performance.
- **Target:** Quantifiable or otherwise measurable characteristic that tells how well a program must accomplish a performance measure.

Program Purpose Defines Performance Measures

The performance measurement process starts with the definition of the program's purpose and the **strategic goal(s) to which the program contributes**. Strategic goals may be measurable (e.g., lives saved due to tornado warning systems) or they may be very difficult to measure (e.g., improved relations between countries).

The PART encourages agencies to develop specific, operational **performance goals that align with strategic goals**. These may be program-specific goals, which must be consistent but not necessarily included in a GPRA strategic plan or performance plan/performance budget.

Categories of Performance Measures

The PART emphasizes robust outcome, output, and efficiency measures, because each kind of measure provides valuable information about program performance. Collectively, these measures convey a comprehensive story regarding what products and services agencies provide, how well they do so, and with what result.

Outcome Measures. Outcomes describe the intended result from carrying out a program or activity. They define an event or condition that is external to the program or activity and that is of direct importance to the intended beneficiaries and/or the public. For a tornado warning system, outcomes could be the number of lives saved and property damage averted. While performance measures must distinguish between outcomes and outputs, there must be a reasonable connection between them, with outputs supporting (i.e., leading to) outcomes in a logical fashion.

Output Measures. Outputs describe the level of activity that will be provided over a period of time, including a description of the characteristics (e.g., timeliness) established as standards for the activity. Outputs refer to the internal activities of a program (i.e., the products and services delivered). For example, an output could be the percentage of warnings that occur more than 20 minutes before a tornado forms.

Outputs v. Outcomes. Outcome measures are best, because these are the ultimate results that benefit the public. Programs must try to translate existing measures that focus on outputs into outcome measures by focusing on the ultimate goal of the program, as shown here:

Outputs	Outcomes
Number of housing units rehabilitated.	Increases in equity (property value) of rehabilitated houses for low-income families as a result of targeted assistance.
Number of businesses assisted through loans and training.	Percent of businesses that remain viable 3 years after assistance.
Number of people served by water/sewer projects.	Increased percent of people with access to clean drinking water.
Number of acres of agricultural lands with conservation plans.	Percent improvement in soil quality; dollars saved in flood mitigation.

Programs that cannot define a quantifiable outcome measure – such as programs that focus on process-oriented activities (e.g., data collection, administrative duties or survey work) – may adopt a “proxy” outcome measure. For example, a program whose purpose is to make forecasts may use increases in the accuracy of predictions as a “proxy” outcome measure. Similarly, a program whose purpose focuses on processing claims or applications may measure increases in accuracy and timeliness of service delivery without a reduction in customer satisfaction. **In such cases, the program must demonstrate a clear link between the “proxy” measure and the outcomes or “public goods” the program is trying to achieve.**

Efficiency measures. While outcome measures provide valuable insight into program achievement, more of an outcome can be achieved with the same resources if an effective program increases its efficiency. The President’s Management Agenda (PMA) encourages agencies to develop efficiency measures. Sound efficiency measures capture skillfulness in executing programs, implementing activities, and achieving results, while

avoiding wasted resources, effort, time, and/or money¹. **Simply put, efficiency is the ratio of the outcome or output to the input of any program².**

The best efficiency measures capture improvements in program outcomes for a given level of resource use. For example, a program that has an outcome goal of “reduced energy consumption” may have an efficiency measure that shows the value of energy saved in relation to program costs. However, it may be difficult to express efficiency measures in terms of outcomes. In such cases, acceptable efficiency measures could focus on how to administer the program better. However, this approach should not shift incentives toward quick, cheap methods at the expense of program effectiveness as measured by better outcomes.

Meaningful efficiency measures consider the benefit to the customer and serve as indicators of how well the program performs. For example, reducing processing time means little if error rates increase. A balanced approach is required to enhance the performance of both variables in pursuit of excellence to customers. In these instances, one measure (e.g., increase in customer satisfaction) may be used in conjunction with another complementary measure (e.g., reduction in processing time).

In all cases, efficiency measures must be useful, relevant to program purpose, and help improve program performance. An efficiency measure for a federal program tracks the ratio of total outputs or outcomes to total inputs (federal plus non-federal). **Leveraging** as a policy can be good because it leads to risk or cost sharing; however, it is not an acceptable efficiency measure, because the leveraging ratio of non-federal to federal dollars represents only inputs. Although improving leveraging may stretch federal program dollars further, it does not measure improvements in the management of total program resources, systems, or outcomes.

Targets and Baselines

Once measures are defined, ambitious and achievable targets must be set, building off of a reliable baseline.

Baselines are the starting point from which gains are measured and targets are set. The baseline year shows actual program performance or prior condition for the given measure in a specified prior year.

Targets refer to improved levels of performance needed to achieve the stated goals. These targets must be *ambitious* (i.e., set at a level that promotes continued improvement) and *achievable* given program characteristics. Each target must have a timeframe (e.g., year(s) in which the target level is to be achieved). In most instances, these targets should be quantifiable. However, OMB recognizes that in some cases measures and their

¹ www.cogsci.princeton.edu/cgi-bin/webwn

² www.cogsci.princeton.edu/cgi-bin/webwn

targets may need to be qualitative and supported by peer review (e.g., expert panels or Inspectors General) or other means. When a target is not quantitative, it must still be verifiable.

Long-Term v. Annual Performance Goals

The PART also distinguishes between **long-term** performance goals and **annual** performance goals. While **long-term performance goals** address performance that is generally several years or more in the future; **annual performance goals** are stated in yearly increments. Similar to determining what a meaningful performance measure is for a program, the appropriate timeframe for a long-term performance goal should also reflect program characteristics.

Long-term is defined as covering a long period of time considering the nature of the program, but is consistent with the periods for strategic goals used in the Agency Strategic Plan.

Annual performance goals are the measures and targets affected by an activity in a particular (generally near-term) year.

Long-term and annual measures should be linked. Indeed a long-term performance goal could be an annual performance goal in the future. For example, a program may have a goal of handling 15 million transactions in 2010 for the same cost of handling 10 million transactions in 2005. This type of goal can demonstrate increased expectations for cost efficiencies over time.

Alternatively, a long-term performance goal could reflect the cumulative effect of annual activities. This type of goal can indicate when the program's mission is accomplished or how it should evolve over time. A climate change research and development (R&D) program might have as annual goals the reduction in uncertainty of various physical factors (parameters) in a climate model, with a 10 year goal of reducing overall model uncertainty by a specific numerical factor.

For a more detailed discussion please see "Performance Measurement Challenges and Strategies" (June 2003) as well as additional examples of performance measures on OMB's website at http://www.whitehouse.gov/omb/part/challenges_strategies.html and http://www.whitehouse.gov/omb/part/challenges_strategies.pdf

SECTIONS OF THE PART

Each PART is divided into four sections:

1. Program Purpose & Design to assess whether the program’s purpose and design are clear and sound.
2. Strategic Planning to assess whether the program has valid long-term and annual measures and targets
3. Program Management to rate agency management, including financial oversight and program improvement efforts
4. Program Results/Accountability to rate program performance on measures and targets reviewed in the strategic planning section and through other evaluations

PART INSTRUMENT LINKAGES

Building on the GPRA framework, establishing appropriate long-term measures and targets lays the groundwork both for annual measures and targets and for assessing program results relative to those targets. Because of the strong focus on strategic planning and performance measurement, certain questions in Sections II (Strategic Planning) and IV (Results) are linked. Specifically, a program cannot get full credit for meeting performance targets in Section IV, if the relevant questions in Section II indicate that the long-term or annual targets are not sound. Section IV scoring allows for partial achievement of performance goals (i.e., *Yes, Large Extent, Small Extent, and No*).

Results Not Demonstrated

A program that has not been able to establish long-term *and* short-term performance measures or does not have data to indicate how it has been performing under measures that have been established will receive a rating of Results Not Demonstrated.

In particular, a program that has a “no” on question 2.1 or question 2.3 should receive a Results Not Demonstrated rating. A program that received *Yes* to these questions, but has a “no” on question 4.1 and 4.2 because it does not have data on its performance measures (as opposed to receiving a “no” due to failure to meet its goals) should also receive a Results Not Demonstrated rating.

TYPES OF PROGRAMS

Although most PART questions are the same, the PART divides all programs into seven categories for the purpose of asking additional questions unique to a particular type of program. These categories apply to both discretionary and mandatory programs.

1. Direct Federal Programs Programs where services are provided primarily by employees of the Federal government, such as the National Weather Service and the Visa and Consular Services.
2. Competitive Grant Programs Programs that provide funds to State, local and tribal governments, organizations, individuals and other entities through a competitive process, such as Health Centers.
3. Block/Formula Grant Programs Programs that provide funds to State, local and tribal governments and other entities by formula or block grant, such as Weatherization Assistance and the Ryan White program.
4. Regulatory Based Programs Programs that accomplish their mission through rulemaking that implements, interprets or prescribes law or policy, or describes procedure or practice requirements, such as the Food Safety and Inspection Service.
5. Capital Assets and Service Acquisition Programs 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) , for example, Defense Shipbuilding and the Bonneville Power Administration.
6. Credit Programs Programs that provide support through loans, loan guarantees and direct credit, such as Export-Import Bank/Long Term Guarantees.
7. Research and Development (R&D) Programs Programs that focus on knowledge creation or its application to the creation of systems, methods, materials, or technologies, such as the Department of Energy/Solar Energy and NASA/Mars Exploration programs.

There is a separate PART form for each of the seven types of Federal programs, though most of the questions are common across the seven forms. The vast majority of Federal programs fit into one of the seven categories of programs for which there is a PART. However, some programs use more than one mechanism to achieve their goals (e.g., grants and credit). Even in these cases, using one PART is likely to be sufficient. To enable this for R&D programs, which can use one of the other program types (e.g., competitive grants) as a means of funding R&D, the R&D PART has been designed to enable R&D programs that fund research through grants, contracts, cooperative agreements or other transactions to answer questions from the Competitive Grants instrument. Similarly, R&D programs that construct or operate equipment or facilities will answer some questions from the Capital Assets and Service Acquisition PART. There may be other cases in which drawing questions from two different PARTs – i.e., creation of a “mixed” form – yields a more informative assessment. In those instances, the PART type that most closely reflects the core functions of the program should be chosen as a base, and then, if necessary, selected questions from another PART can be added. *The OMB examiner should consult with a member of the OMB Performance Evaluation Team, if considering this approach.*

For new programs for which it is impractical to expect results, it is possible to complete only Sections 1 through III of the PART. However, performance measures, targets and related information should still be provided (and scored) in Section IV for new programs where practical.

Question-specific instructions are attached to help explain the purpose of each question and general standards for evaluation. The individual PART worksheets also contain this guidance as well as instructions on the technical aspects of using the worksheets. These instructions will not cover every case, and it is up to the user to bring relevant information to bear in answering each question that will contribute to the program's assessment.

NOTE: Instructions on collecting budget data (accounts and funding levels) will be provided in OMB Circular A-11, which is expected to be released by July, 2004.

I. PROGRAM PURPOSE & DESIGN

This section examines the clarity of program purpose and soundness of program design. It looks at factors including those the program, agency, or Administration may not directly control but which are within their influence, such as legislation and market factors. Programs should generally be designed to address either an efficiency matter, such as a public good or externality, or a distributional objective, such as assisting low-income families in the least costly or most efficient manner. A clear understanding of program purpose is essential to setting program goals, measures, and targets; maintaining focus; and managing the program. Potential source documents and evidence for

answering questions in this section include authorizing legislation, agency strategic plans, performance plans/performance budgets, and other agency reports.

Options for answers are *Yes*, *No* or *Not Applicable*. Design flaws in the underlying legislation can and should be considered and supported by evidence, and are grounds for a *No*. *Not Applicable* answers are likely to be rare, particularly for items 1.1-1.4, as these questions should apply to virtually all programs. (For R&D programs, most of the questions in this section help address program “relevance,” one of the three fundamental issues of the R&D Investment Criteria (see Appendix A).)

1.1 Is the program purpose clear?

Purpose of the question: to determine whether the program has a focused and well-defined mission. Determining this purpose is critical to determination of useful performance measures and targets.

Elements of a *Yes* answer: A *Yes* answer would require a clear and unambiguous mission. Considerations can include whether the program purpose can be stated succinctly. A *No* answer would be appropriate if the program has multiple conflicting purposes.

Evidence/Data: Evidence can include a statement of the purpose and supporting objectives from the program’s authorizing legislation, program documentation or mission statement.

1.2 Does the program address a specific and existing problem, interest, or need?

Purpose of the question: to determine whether the program addresses a specific problem, interest, or need that can be clearly defined and that currently exists.

Elements of a *Yes* answer: A *Yes* answer would require the existence of a relevant and clearly defined interest, problem or need that the program is designed to address. A *Yes* answer would also require that the program purpose is still relevant to current conditions (i.e., that the problem the program was created to address still exists). Considerations could include, for example, whether the program addresses a specific market failure. A *No* should be given if there is no clear need for the program. Programs may receive a *Yes* to question 1.1 and a *No* on question 1.2 and vice versa.

For example, Housing and Urban Development’s HOPE VI program had a clear purpose with a goal of demolishing 100,000 of the most severely distressed public housing units. However, the program has surpassed this goal and addressed the need for which the program was originally created. Therefore, the program received a *Yes* in 1.1 and a *No* in 1.2. In addition, the Department of Education Vocational Education program had an unclear mission caused by multiple and overlapping objectives. It was able to document, however, that a significant number of students are graduating from high school and community college without the necessary

academic and technical skills. Therefore, the program received a *No* for 1.1 and a *Yes* for 1.2.

Evidence/Data: Evidence should describe the problem, interest or need that the program is designed to address and include relevant documentation. An example could be the number and income levels of uninsured individuals for a program that provides care to those without health insurance.

1.3 Is the program designed so that it is not redundant or duplicative of any other Federal, State, local or private effort?

Purpose of the question: to determine whether the program is designed to fill a unique role or whether it instead unnecessarily duplicates or even competes with other Federal or non-federal programs.

Elements of a *Yes* answer: A *Yes* answer would require that the program does not excessively overlap with other Federal or non-federal efforts, including the efforts of State and local governments or the private and non-profit sectors. A consideration can include whether the program serves a population not served by other programs.

A *No* answer should be given when there is more than one program that addresses the same problem, interest, or need, regardless of the size or history of the respective programs. For programs that partially overlap with others, a *No* should be given when major aspects of the program, such as its purpose, targeted beneficiaries, or mechanisms, are duplicative. If there are two programs that significantly overlap and one is large and another is small, both programs should receive a *No* for this question.

Similar programs might be justified in receiving a *Yes* if a strong case can be made that fixed costs are low and competition is beneficial (e.g., perhaps multiple laboratories) or if more than one service delivery mechanism is appropriate (e.g., block grants for base activities and competitive grants for demonstration projects). Also, the standard of evidence to receive a *Yes* should be higher where Federal programs overlap with each other than where a Federal program overlaps with private, local, or State programs. For example, two Federal programs to address training would face a high standard to receive a *Yes* to this question; in contrast, a *Yes* could be provided to a Federal program to address indigent medical care across the nation, even though there are many local and private programs that also address indigent medical care. The key would be whether the gaps in the non-Federal provision are large enough to warrant a Federal program and whether the Federal program is well designed to mesh with non-Federal efforts and responsibilities.

For credit programs, a *Yes* answer would require evidence of the market failure/absence or unwillingness of private-sector participation and an overview of the market, including all international, Federal, local, and private-sector participants.

For R&D programs, some degree of duplication is permissible, if it is well justified and coordinated. A *Yes* answer would require justification that the program provides

value beyond that of any similar efforts at the agency, efforts at other agencies, or efforts funded by State and local government, private and non-profit sectors, or other counties. Justification first requires due diligence in identifying similar past or present efforts.

Evidence/Data: Evidence should identify duplicative programs and their total expenditures and/or a description of efforts supported by those programs that address a similar problem in a similar way as the program being evaluated.

1.4. Is the program design free of major flaws that would limit the program's effectiveness or efficiency?

Purpose of the question: to determine whether there are major design flaws in the program that limit its efficiency.

Elements of a Yes answer: The program should be free from major design flaws that prevent it from meeting its defined objectives and performance goals. To receive a *Yes*, there should be no strong evidence that another approach or mechanism would be more efficient or effective to achieve the intended purpose. A consideration could be whether the government would get the same or better outcome by expending fewer total resources through a different mechanism. For example, there may be evidence that a regulatory program to ensure public safety would be more effective than a grant program. Analysis should consider whether the program structure continues to make sense given changing conditions in the field (e.g., changing threat levels or social conditions). Other considerations could include whether the program extends its impact by leveraging funds and contributions from other parties.

For credit programs, an additional consideration can include the extent to which a large number of borrowers would otherwise not have access to financial resources. Also consider whether the program costs are adequate, but not excessive, to achieve the policy goals. For example, a *Yes* answer could mean that the program effectively uses market mechanisms to reduce government risk and thus minimize program costs.

Regulatory programs should receive a *No* if the statute underlying the regulations is not designed to maximize net benefits.

For capital assets and service acquisition programs, a *Yes* answer requires that, in addition to the general criteria, the program is supported by an adequate capital asset management infrastructure that is consistent with the principles and techniques of effective capital programming, and has clear lines of authority, responsibility, and accountability for managing capital assets.

Evidence/Data: Evidence demonstrating efficient design can include cost effectiveness studies comparing alternative mechanisms (e.g., regulations or grants) with the current design (say, direct federal provision). Evidence on the relative benefits and costs of the activity are also useful.

Evidence for determining whether the threshold for capital programming has been met should include the documented program-relevant agency or bureau capital programming policies, directives, instructions, manuals, and assignment of authorities and responsibilities to agency personnel and organizational units.

1.5 Is the program design effectively targeted, so that resources will reach intended beneficiaries and/or otherwise address the program's purpose directly?

Purpose of the question: to determine whether the program is designed so that program resources will reach the intended beneficiaries efficiently and to avoid unintended subsidies. "Beneficiaries" refers to those who benefit from the favorable outcome of the program. "Reach" refers to the distribution of benefits.

Unlike Question 1.4, which addresses examination of alternatives to achieve a program's goals, this question asks whether program resources under the *chosen* alternative are oriented toward the effective achievement of the program's purpose.

Elements of a Yes answer: A *Yes* answer would require that a program demonstrates that the right beneficiaries are being targeted, activities that would have occurred without the program are not subsidized (or receive only warranted levels of subsidies), and program funds are targeted effectively to meet program purposes.

Acceleration of activities due to Federal funding can be grounds for a *Yes*, but there should be evidence that the acceleration warrants the subsidy or application of funding. Acceleration of an activity that increases profits for a business -- that the firm would or could have undertaken eventually without the subsidy or application of funding -- would not generally qualify for a *Yes*, unless there are significant external (i.e., social) benefits from the activity.

In the case of block-grant and credit programs, the assessment should also consider how well funds are targeted to meet the program purpose and whether funds are protected against supplantation or substitution.

For R&D programs, a *Yes* answer would require identification of relevance to specific national needs, agency missions, fields of science or technology, or other "customer" needs. A customer may be another program at the same or another agency, an interagency initiative or partnership, a firm, or an organization from another sector or country. For these programs, the question refers to awardees of contracts, cooperative agreements or other transactions, as well as grants.

Evidence/Data: Evidence should show that the program is designed to 1) reach the highest practicable percentage of target beneficiaries, and 2) have the smallest practicable share of funds going to unintended beneficiaries. Regarding item 1, a small program may only be able to reach a small number of beneficiaries, but it should be well targeted on some merit basis. On item 2, programs not designed to avoid unwarranted shares of funding going to beneficiaries who do not need or merit

the funding should receive a *No* answer. Programs that are designed in a way that is likely to result in significant levels of erroneous payments should receive a *No*.

II. STRATEGIC PLANNING

This section focuses on program planning, priority setting, and resource allocation. Key elements include an assessment of whether the program has a limited number of performance measures with ambitious -- yet achievable -- targets, to ensure planning, management, and budgeting are strategic and focused. Potential source documents and evidence for answering questions include strategic planning documents, agency performance plans/performance budgets and reports, reports and submissions from program partners, evaluation plans, and other program documents.

Options for answers are *Yes*, *No* or *Not Applicable*. While it is recognized that some programs may have greater difficulty than others in developing quantitative performance goals, programs must have meaningful and appropriate methods for demonstrating results. OMB and agencies should work together to develop approaches for programs where it is difficult to develop quantitative measures, and where qualitative, expert-review, or other measures are more appropriate. For R&D programs, most questions in this section help address the *prospective* aspects of the R&D Investment Criteria (see Appendix A).

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?

Purpose of the question: to determine if the program has long-term performance measures to guide program management and budgeting and promote results and accountability. This question seeks to assess whether the program measures are salient, meaningful, and capture the most important aspects of program purpose and appropriate strategic goals.

Elements of a *Yes* answer: A *Yes* requires two or three specific, easily understood outcome measures that directly and meaningfully support the program's purpose. "Long-term" means a long period relative to the nature of the program, likely 5-10 years, and consistent with time periods for strategic goals used in the agency's GPRA strategic plan.

The performance measures should focus mainly on outcomes, although in some cases output measures are permissible. The measures may be those developed to comply with GPRA, so long as they meet the "Performance Measures" section of this document. Otherwise, OMB and agencies should revise the measures to fully meet GPRA and PART standards. A *Yes* answer can also be given if OMB and the agency have reached agreement on a limited number of long-term measures that will be

added to the 2006 GPRA strategic plan or performance plan/performance budget. Significant changes to the GPRA *strategic plan* may require stakeholder consultation.

Output measures only meet the standards of a *Yes* answer if the program can produce sound justification for not adopting outcome measures. For example, a program that is exclusively focused on processing applications **and** is unable to adequately define a quantifiable outcome measure may use measures that focus on increases in accuracy and/or timeliness of service delivery. Whenever output measures are proposed, the program must clearly show how such measures reflect progress toward desired outcomes.

An example of an unacceptable long-term measure is a housing program that is proposing using a measure of “number of housing units constructed.” In this case, such an output is unacceptable as the program is expected to articulate and measure progress toward achieving outcomes (e.g., increases in homeownership rates, increases in housing equity in low-income communities).

For more detailed discussion on when output measures may be used as proxy measures for outcomes, please see “Selecting Performance Measures” section of the PART guidance or visit OMB’s website at <http://www.omb.gov/part/>.

A *No* must be given for long-term measures that do not directly and meaningfully relate to the program’s purpose or are unnecessarily focused on outputs and lack adequate justification. A program must not receive a *No* for having too many measures, if it has identified a few high-priority ones that represent important aspects of the program.

<p>Performance measures should be listed in the Measures tab of the PART worksheet. Only measures that meet the standards for a <i>Yes</i> should be entered on the worksheet.</p>

For block grant programs that support a wide range of purposes and allow grantees to set their own program priorities, measures that address the extent to which grantees meet their own goals or effectively target populations are options if no better measures are possible.

For R&D programs, OMB will work with agencies to assess existing or develop appropriate measures. Some R&D programs, especially in basic research, may not be required to provide long-term efficiency measures, due to the uncertainty of outcomes and the years it takes to achieve and recognize them.

Evidence/Data: Evidence will include the long-term measures established for the program either in the existing agency GPRA documents or other program documents or as agreed to by OMB and to be included in the 2006 GPRA documents. In the case of new measures, if targets and baselines are not defined, a plan for their

development (i.e., timeline, methods for data collection, responsible office and/or staff) must be agreed to by the agency and OMB.

2.2 Does the program have ambitious targets and timeframes for its long-term measures?

Purpose of the question: to determine if the program has challenging but realistic quantifiable targets and timeframes for the long-term measures.

Elements of a Yes answer: A *Yes* answer requires that specific quantified targets have been developed for measures evaluated in Question 2.1. (Where targets are not “quantitative,” they still must be verifiable, e.g., through a clean audit or outstanding ratings by an expert panel). Baselines from which to measure targets are encouraged for outcome measures and required for output measures. Targets and timeframes must be ambitious, that is, they must be set at a level that promotes continued improvement within achievable efficiencies. Where relevant, a *Yes* also requires that a program has defined an appropriate end target.

A *No* is appropriate if quantified targets or timeframes are not included for key measures or if the targets or timeframes are not ambitious or challenging.

If the program received a *No* in Question 2.1, the program must get a *No* for this question.

Targets must be listed in the Measures tab of the PART worksheet.

For R&D programs, a *Yes* answer would require that the program provides multi-year R&D objectives. Where applicable, programs must provide schedules with annual milestones, highlighting any changes from previous schedules. Program proposals must define what outcomes would represent a minimally effective program and a successful program.

Evidence/Data: Evidence will include targets in existing agency GPRA documents or other program documents or as agreed to by OMB and to be included in the 2006 GPRA documents.

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?

Purpose of the question: to determine whether a limited number of annual performance measures have been identified that directly support the long-term goals evaluated in Questions 2.1 and 2.2. The measures should be logically linked to the long-term goals in a manner that enables them to demonstrate progress toward achieving those long-term goals.

Elements of a *Yes* answer: A *Yes* answer would require annual performance measures that are discrete, quantifiable, and measurable. Most importantly, these annual measures should measure the program's progress toward reaching the long-term goals evaluated in Questions 2.1 and 2.2. The annual performance measures may focus on outputs and may or may not be those developed by the agency to comply with GPRA. A *Yes* answer can also be given, if OMB and the agency have reached agreement on a limited number of annual measures that will be included in the 2006 GPRA performance plan/performance budget.

To receive a *Yes* answer, programs must have at least one efficiency measure or be developing one. If a measure is under development, a schedule for adoption (i.e., timeline, methods for data collection, responsible office and/or staff) must be agreed to by the agency and OMB.

For more detailed discussion on defining acceptable efficiency measures please see “Selecting Performance Measure” section of the PART guidance or visit OMB’s website at <http://www.omb.gov/part/>.

Programs that are unable to provide an efficiency answer may only receive a *Yes* to this question if they are able to adequately explain why adoption of such a measure is not feasible (e.g., some basic R&D programs).

If the program received a *No* in Question 2.1, an explanation of how annual performance goals contribute to desired long-term outcomes and purpose of the program must be provided to receive a *Yes* for this question. Performance measures should be listed in the Measures tab of the PART worksheet. Only measures that meet the standards for a *Yes* should be entered on the worksheet.

For block grant programs that support a wide range of purposes and allow grantees to set their own program priorities, measures that address the extent to which grantees meet their own goals or effectively target populations are options if no other measures are possible.

For R&D programs, a *Yes* answer would require that the program has annual performance measures to track how the program could improve scientific understanding and its application. For R&D programs that have multi-year schedules, the annual measure should tie into the longer term milestones, as appropriate. Some basic research programs may not be able to define meaningful annual outcome or efficiency measures. In such cases, these programs may use process-related indicators (see Question 3.4), if the program can explain how those processes are directed toward the intended long-term goals. OMB will work with agencies to address appropriate measures.

For capital assets and service acquisition programs, a *Yes* answer requires that, in addition to the general criteria, annual performance measures include those that are

sufficient to track achievement of the cost, schedule, and performance goals of asset acquisitions as they relate to the overall program.

Evidence/Data: Evidence will include the annual measures established for the program in the agency GPRA performance plan/performance budget or other program documents, or they may be new measures as agreed to by OMB and which will be included in the 2006 GPRA performance plan/performance budget. Evidence for capital asset acquisition programs includes agency acquisition and project management working documents, contract performance measures and metrics, and business cases (OMB Circular A-11 Exhibit 300s).

2.4. Does the program have baselines and ambitious targets for its annual measures?

Purpose of the question: to determine if the program has baselines and challenging but realistic quantified targets for the annual measures.

Elements of a Yes answer: A *Yes* answer would generally require that baselines have been established for most of the annual measures evaluated in Question 2.3. A *Yes* answer also requires that specific quantified annual targets have been developed for most of the program's annual measures evaluated in Question 2.3. These targets provide a specific value with which performance can be compared. These targets must be ambitious, that is they must be set at levels that ensure continued improvement and realization of efficiencies. They also should be within reason for the program to achieve.

A *No* answer would be appropriate if quantified targets or timeframes are not included for most measures or if the targets are not ambitious or challenging.

If the program received a *No* in Question 2.3, the program must get a *No* for this question. Targets should be listed in the Measures tab of the PART worksheet.

Evidence/Data: Evidence will include targets in the agency GPRA performance plan/performance budget or other program documents or as agreed to by OMB and will be included in the 2006 GPRA performance plan/performance budget.

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?

Purpose of the question: to determine whether program efforts carried out by program partners also support the annual and long-term performance goals of the program.

Elements of a Yes answer: A *Yes* answer would require that program managers strive to ensure that partners support the overall goals of the program and measure and report on their performance as it relates to accomplishing those goals. For example, a program that requires all grant agreements and contracts to include performance

measures that will help the program achieve its goals and monitor those measures would receive a *Yes*.

If, however, a program does not through a performance requirement or some other means get program grantees to link their activities to the program's goals, a *No* would be appropriate. The most obvious example of a partner is an entity receiving program funding. While a program cannot always control the activities of its partners, it can exert influence through a number of various mechanisms.

If the program received a *No* for both Questions 2.1 and 2.3, the program must receive a *No* for this question.

In the case of regulatory programs, all regulated entities are not necessarily defined as program partners.

For programs funding grants for basic research, a *Yes* answer can be achieved if the program solicitation explicitly includes the program goals, and grant applications and progress reports provide sufficient means for the program manager to assess performance and continuing relevance.

Evidence/Data: Evidence can include contracts and other documents that tie contractor performance to program goals, as well as other procedures the program uses to get partners to commit to, measure, and report on performance related to the program's goals.

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?

Purpose of the question: to ensure that the program (or agency) conducts non-biased evaluations on a regular or as-needed basis to fill gaps in performance information. These evaluations should be of sufficient scope to improve planning with respect to the effectiveness of the program. (For R&D programs, this question is central to prospective planning to address all of the R&D investment criteria (see Appendix A).)

Elements of a *Yes* answer: A *Yes* answer would require regularly scheduled objective, high quality, independent evaluations that examine how well the program is accomplishing its mission and meeting its long-term goals. A *Yes* answer also would require that program evaluations address the following elements of quality, scope, and independence.

Not Applicable is not an option for this question; given the flexibility in determining what constitutes an evaluation, all programs should undergo an evaluation that meets the following elements of quality, scope, and independence.

Quality. Evaluations should be sufficiently rigorous to provide information on the effectiveness of the program and, for programs that support or employ a range of services and approaches, information on the effectiveness of the various services and approaches. To receive a *Yes*, agencies should provide evidence that they have chosen and applied evaluation methods that provide the most rigorous evidence of a program's effectiveness that is appropriate and feasible. A program may receive a *Yes* for this question if the agency and OMB determine that existing evaluations do not provide the most rigorous evidence possible, but that the program is in the process of developing new evaluation approaches that will provide the most rigorous evidence possible by a specified future date.

The purpose of the question is to find out whether agencies have evaluations that assess the effectiveness of the program. The most significant aspect of program effectiveness is *impact*—the outcome of the program, which otherwise would not have occurred without the program intervention. Where it is feasible to measure the *impact* of the program, randomized controlled trials are generally the highest quality, unbiased evaluation to demonstrate the actual impact of the program. However, these studies are not suitable or feasible for every program, and a variety of evaluation methods may need to be considered because Federal programs vary so dramatically. Other types of evaluations, including well-designed quasi-experimental studies, may provide useful information about the impact of a program (but should be scrutinized given the increased possibility of an erroneous conclusion) and/or can help address *how* or *why* a program is effective (or ineffective) (i.e., meeting performance targets, achieving efficiency, fulfilling stated purpose).

Overall, *evaluations must be appropriate to the type of program*. Agencies and OMB should consult evaluation experts, either in-house and/or external, as appropriate, when deciding what type of evaluation will provide the most rigorous evidence appropriate and feasible. The following are several links to references on program evaluation located on the internet that are not intended to be exhaustive, but which may be helpful:

- Program Evaluation Methods: Measurement and Attribution of Program Results; Treasury Board of Canada, Secretariat; 1998. (a book available online)
http://www.tbs-sct.gc.ca/eval/pubs/meth/pem-mep_e.pdf
- Understanding Impact Evaluation; The World Bank Group. (a web site)
<http://www.worldbank.org/poverty/impact/index.htm>
- “Program Evaluation: An Evaluation Culture and Collaborative Partnerships Build Agency Capacity;” GAO-03-454; U.S. General Accounting Office; May 2003.
<http://www.gao.gov/docdblite/summary.php?recflag=&accno=A06797&rptno=GAO-03-454>

- “Performance Measurement and Evaluation: Definitions and Relationships;” GAO/GGD-98-26; *U.S. General Accounting Office*; April 1998.
<http://www.gao.gov/docdblite/summary.php?recflag=&accno=160204&rptno=GGD-98-26>
- “Designing Evaluations;” GAO/PEMD-10.1.4; *U.S. General Accounting Office*; May 1991.
<http://161.203.16.4/t2pbat7/144040.pdf>
- Randomized Controlled Trials: A User’s Guide; Jadad, Alejandro A.; BMJ Books; 1998. (*a book available online*)
<http://www.bmjpub.com/rct/contents.html>
- Research Methods Knowledge Base; Trochim, William M.; Cornell University. (a web site)
<http://www.socialresearchmethods.net/kb/>
- “Identifying and Implementing Educational Practices Supported By Rigorous Evidence: A User Friendly Guide;” U.S. Department of Education; December 2003.
<http://www.ed.gov/rschstat/research/pubs/rigorousetid/rigorousetid.pdf>

Scope. This question looks directly at whether there are evaluations on the program’s achievement of performance targets, and that these evaluations examine the underlying cause and effect relationship between the program and the target. In cases where a comprehensive evaluation is unnecessary based on the known effectiveness of an intervention and performance data on the program, evaluations that fill gaps in performance information can meet the elements of a *Yes* answer. A program’s *effectiveness*, including *impact*, also may be considered. A program may receive a *Yes* for this question if the agency and examiner determine that existing evaluations are inadequate, but the program is in the process of developing an appropriate evaluation to be completed by a specified future date.

Evaluations also should include recommendations on how to improve the program's performance. To ensure the program continues to meet its performance targets, an evaluation should be scheduled on a periodic basis, such as every two to five years, or whatever time schedule is reasonable based on the specific program and its mission and goals.

R&D programs also should undergo independent reviews of relevance to their agencies, fields of science or technology, or customers, (e.g., of process) in addition to assessing questions of performance. These reviews should conclude with reports documenting the findings and recommendations. A “customer” may be another program at the same or another agency, an interagency initiative or partnership, or a firm, an organization from another sector or country, or the general public. Industry-

relevant programs may use industry cost-sharing of associated projects as an indicator of market-relevance, and they should incorporate industry in planning and prioritization. Reviews should be rigorous and methodical and be a critique of the program's methods, results and findings by others in the field with requisite training, expertise, and independence.

Independence. To be independent, non-biased parties with no conflict of interest would conduct the evaluation. Evaluations conducted by the program itself should generally not be considered "independent;" however, if the agency or program has contracted out the evaluation to a third party this may qualify as being sufficiently independent. Evaluations conducted by an agency's Inspector General or program-evaluation office might also be considered "independent." OMB examiners and agency staff will determine if a specific evaluation can be considered "independent" for this question.

If a program has had previous evaluations that meet the elements of independence and scope, but that were not sufficiently rigorous, and the program is developing or about to conduct a new program evaluation using the most rigorous method that is feasible and appropriate, then the program would receive a *Yes*.

In the absence of ANY independent evaluations, a program would receive a *No*, unless it is planning to carry out a rigorous program evaluation in the near future. A *No* answer would also be appropriate for a program that has insufficient independent evaluation data or has evaluations that address process and not performance.

Evidence/Data: Evidence should include a program evaluation plan or schedule of program evaluations and program documentation describing the type of evaluation, including scope and quality, and the criteria for selecting an independent evaluator.

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?

Purpose of the question: to establish whether the performance-planning and budget-planning processes are integrated so that 1) resource allocation decisions reflect desired performance levels (given resource constraints) and 2) the effects of funding and other policy changes on results are clear.

Elements of a Yes answer: A *Yes* answer should reflect effective program budgeting based on sound levels for 1) annual and long-term performance targets and 2) budget resources. To receive a *Yes*, the program must provide a presentation that makes clear the impact of funding, policy, or legislative decisions on expected performance and explains why the requested performance/resource mix is appropriate. A program with budget planning that is not tied to performance or strategic planning would receive a *No*.

A *Yes* answer would also require that the program report all direct and indirect costs needed to attain the performance results, including applicable agency overhead, retirement, and other costs that might be budgeted elsewhere. The exclusion of minor amounts of services provided from central departmental offices (e.g., Office of the Secretary) from program costs does not require a program to receive a *No* answer. However, a program that generates significant costs – which might range from radioactive waste disposal to attorneys’ salaries – that must be addressed by another program should budget for these costs or, at a minimum, provide this information in clear display tables that display the full costs of attaining results.

For capital assets and services acquisition programs, a *Yes* answer requires that, in addition to the general criteria, programs explain the relationship of asset acquisitions to overall program performance goals and would be able to identify impacts of changes on program performance (for example, the effect of a change in the quantity acquired).

Evidence/Data: Evidence can include documentation of how the budget request directly supports achieving performance targets. Budget documents should also clearly indicate the full costs of achieving performance goals, even if some of these costs do not appear in the specific account or activity line of the program.

Also, evidence can include an agency program budget estimate that identifies all spending categories in sufficient detail to demonstrate that all relevant costs are included or a report that shows the allocation of all significant program overhead costs to the program.

2.8. Has the program taken meaningful steps to correct its strategic planning deficiencies?

Purpose of the question: to determine whether the program is on track to correct any strategic planning deficiencies that have been identified.

Elements of a *Yes* answer: A *Yes* answer would require that the program has acted to correct strategic planning deficiencies. A program that does not review planning efforts or does not make corrections to eliminate identified deficiencies would receive a *No*. The question addresses any deficiencies identified in this section. Particular emphasis, however, should be placed on whether the program is working to adopt a limited number of specific, ambitious long-term performance goals and a limited number of annual performance goals that demonstrate progress toward achieving the long-term goals, if they do not already have these measures or associated baselines, targets, and timeframes.

For Capital Assets and Service Acquisition programs (and relevant R&D programs), one strategic planning deficiency to be addressed is if the program has in the past received a *No* to Question 4.CA1.

Evidence/Data: Evidence can include a description of how deficiencies in the strategic planning of a program are identified and corrected, as well as examples of such changes.

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?

Purpose of the question: to determine whether the program had developed regulations with clearly specified goal(s). It should be determined whether (1) the program is only issuing those rules absolutely necessary to achieve long-term program goals and is not over-regulating, (2) all of the rules necessary to meet the program goals have been issued, and (3) the regulations clearly indicate how they help to meet the program goals.

Elements of a Yes answer: A *Yes* answer would require that only those regulations that are absolutely necessary to accomplish the program mission and goals are promulgated or are in the process of being promulgated; this determination should be based on the standards laid out in Executive Order 12866 on the identification of the need for the regulation and the identification of the market failure (if applicable). Additionally, the public should be able to understand how the regulations fit into the overall achievement of the program goals. A *Yes* response indicates that there are no superfluous regulations, that regulations are planned or in the process of being promulgated to cover regulatory gaps where new regulations are required to accomplish program goals, and that the Preamble of each program regulation indicates how the rule contributes to the achievement of specific program goals.

A program would receive a *No* if it has 1) obvious regulatory gaps or outdated regulations in effect and 2) not initiated planned actions to rectify these problems in a timely manner.

Evidence/Data: Evidence can include legislation that indicates specifically or generically what regulations need to be promulgated as well as the rules themselves, especially the preambles. It can also include internal agency guidance indicating a process is in place to ensure rulemaking involves a clear linkage of the need for the rule to a stated goal. In this case, the agency should be able to clearly articulate this association.

NOTE: Questions 2.RG1, 3.RG3, and 4.RG1 address the progression of reviews for a program's rules. Specifically, 2.RG1 addresses targeted *development* of regulations,

3.RG3 addresses effective *implementation* of regulations, and 4.RG1 addresses *maximization* of potential benefits during the regulation's implementation.

Capital Assets and Service Acquisition Programs

R&D programs addressing the acquisition, construction or operation of facilities or other capital assets should answer the Capital Assets and Service Acquisition question 2.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?

Purpose of the question: to determine whether the agency is investing in an asset or service that provides the best value to the government.

Elements of a Yes answer: To receive a *Yes* response, the agency should have conducted analyses of alternatives and use those analyses. Each analysis should include the baseline assessment/status quo, non-material solutions (e.g., data compression in lieu of a new data cable), consideration of alternatives (including alternatives to capital assets and benefit-cost analysis in accordance with OMB Circular A-94), and trade-offs between cost, schedule, and performance goals. The program should be able to demonstrate that the analysis is credible (e.g., by having it reviewed and validated by an independent entity outside the program). If an independent entity's analysis differs from the program's analysis, the program should defend differences.

Evidence/Data: Evidence can include a summary of the analysis of alternatives, and documentation of any independent reviews of the analysis. As one source of data, an agency capital asset plan or business case (Exhibit 300) documentation may also be used.

Research and Development Programs

R&D programs addressing the acquisition, construction or operation of facilities or other capital assets should answer the Capital Assets and Service Acquisition question (2.CA1).

2.RD1. If applicable, does the program assess and compare the potential benefits of efforts within the program and (if relevant) to other efforts in other programs that have similar goals?

Purpose of the question: to determine whether applicable programs are evaluating their efforts with respect to their relative potential benefits. Applicable programs include those applied R&D programs that pursue multiple options toward achieving

similar public benefits. (This question addresses the first of the industry-related R&D criteria, regarding the articulation of program benefits (see Appendix A).)

Elements of a *Yes* answer: To receive a *Yes* rating, the program (or agency) should conduct periodic comparisons of the potential benefits of its proposals with alternatives. The program should be able to demonstrate that the analysis is credible (e.g., by having it reviewed and validated by an independent entity outside the program). If an independent entity’s analysis differs from the program’s analysis, the program should defend differences.

Evidence/Data: Evidence can include an analysis comparing proposed approaches with alternative strategies. OMB will work with agencies as needed to assist in the content and structure of these assessments.

2.RD2. Does the program use a prioritization process to guide budget requests and funding decisions?

Purpose of the question: to determine whether the program has clear priorities and uses them in budget requests and funding decisions. (This question addresses the R&D “relevance” criterion (see Appendix A).)

Elements of a *Yes* answer: A *Yes* answer would require a documented process to identify priorities and use them, as well as an identified set of current priorities among program goals, objectives, and activities.

R&D programs are encouraged to work with independent advisory bodies to help prioritize in ways that benefit the larger science and technology enterprise.

Evidence/Data: Evidence can include clear statements of program priorities in program documentation or mission statements, as well as documentation of the priorities identified by any qualified independent advisory bodies.

III. PROGRAM MANAGEMENT

This section focuses on a variety of elements related to whether the program is effectively managed to meet program performance goals. Key areas include financial oversight, evaluation of program improvements, performance data collection, and program manager accountability. Additionally, specific areas of importance for each program type are also explored. Potential source documents and evidence for answering questions in this section include financial statements, GAO reports, IG reports, performance plans, budget execution data, IT plans, and independent program evaluations.

Options for answers are *Yes*, *No* or *Not Applicable*. (For R&D programs, some of the questions in this section help address the *prospective* aspects of program “quality” and

“performance” of the R&D Investment Criteria, in addition to addressing general program management issues (see Appendix A.)

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?

Purpose of the question: to determine whether the program collects data on performance and the performance of its partners and uses the data to inform program management, resource decisions, and program performance.

Elements of a Yes answer: A Yes answer would require that the program agency regularly collect high-quality performance data relating to key program goals and use that information to adjust program priorities, allocate resources, or take other appropriate management actions. When key program activities are carried out by other entities, such as grantees, agencies should also consider their performance as well. A Yes also requires that the agency has collected the baseline performance data necessary to set meaningful, ambitious performance targets.

Program partners are other agencies or intermediaries responsible for carrying out different aspects of the program and might include partner agencies, grant recipients, participating financial institutions, regulated bodies, and contractors. Timely performance information is information that reflects current performance and is current enough to be useful in program management. Credible performance information is information that is collected through a systematic process with quality controls to confirm the validity of the data.

For capital assets and service acquisition programs, a consideration is whether the program uses an earned value management system or similar system.

For credit programs, consider whether the agency or program managers regularly collect and update loan performance information to effectively and consistently manage the portfolios.

Some long-term basic research programs may not be able to define meaningful annual outcome performance measures, aside from process measures. In such cases, these programs may use process-related measures, especially those that can be conceptually linked to long-term research goals.

Evidence/Data: Evidence can include a description of how the agency uses performance information in managing the program, as well as illustrative examples of recent management actions based on performance information. Evidence can also include steps taken by a program to enact necessary improvements cited by a specific evaluation.

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?*

Purpose of the question: to determine whether the program managers and partners are accountable for achieving program results.

Elements of a Yes answer: A Yes answer would require that the program agency identify the managers who are responsible for achieving key program results and establish performance standards for those managers. When program partners contribute to the achievement of program goals, a Yes would also require those partners to achieve specific performance standards.

In the case of block and formula grant programs, elements of a Yes are not confined to complying with the law. Elements of a Yes can include the presence of incentives for managers and program partners that would encourage corrections in deficient programs. For block-grant programs which support a wide range of purposes and allow grantees to set their own program priorities, this question should be interpreted as whether the grantees hold managers accountable and encourage corrections in deficient programs.

For capital assets and service acquisition programs, a consideration is whether contracts include minimum performance thresholds, incentives for good performance, or other mechanisms to increase accountability.

Some long-term basic research programs may not be able to define meaningful annual outcome performance measures, aside from process measures. In such cases, these programs may use process-related measures, especially those that can be conceptually linked to long-term research goals.

Evidence/Data: Evidence can include the use of performance management contracts with program managers, or some other mechanism for incorporating program performance into personnel performance evaluation criteria. Evidence of partners' accountability can include requiring grant and contract awards and renewals to consider past performance.

3.3 *Are funds (Federal and partners') obligated in a timely manner and spent for the intended purpose?*

Purpose of the question: to determine whether funds are administered efficiently and obligated in accordance with planned schedules and spent for the intended purposes.

Elements of a Yes answer: A Yes answer would require that the program funds be obligated consistently with the overall program plan and that a limited amount of unobligated funds remain at the end of the year. A Yes answer would also require that programs and partners establish schedules for obligations that properly correspond to

the resource needs of the program plan. In addition, a *Yes* answer requires that adequate procedures exist for reporting actual expenditures, comparing them against the intended use, and taking timely and appropriate action to correct single audit findings when funds are not spent as intended.

A program would receive a *No* if it had significant erroneous payments or was in violation of the Anti-Deficiency Act.

Evidence/Data: Evidence can include periodic and year-end spending reports from the program and its partners. Evidence on expenditures can include spending reports that draw intended purpose from the Congressional Justifications, Appropriations, and program operating plans and match them against actual spending. For grantees, evidence can include grantee audit reports under the Single Audit Act, including data captured in the Federal Audit Clearinghouse, and the existence of an established procedure for reviewing actual expenditures against budgets in grant awards or appropriate Federal guidelines.

A *No* answer is strong evidence that a *No* may be warranted for Question 3.6, which examines financial management more generally.

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?

Purpose of the question: to determine whether the program has effective management procedures in place to ensure the most efficient use of each dollar spent on program execution.

Elements of a *Yes* answer: The answer to this question should focus on whether the program has regular procedures in place to measure and achieve efficiencies and cost effectiveness, recognizing that various types of procedures may be acceptable. A *Yes* would require that the program's performance plans include efficiency measures and targets, such as per-unit cost of outputs, timing targets, and other indicators of efficient and productive processes germane to the program. In addition, there are other acceptable ways to demonstrate that a program has established procedures for measuring and achieving efficiencies and cost effectiveness. For example, a program that regularly uses competitive sourcing to determine the best value for the taxpayer, invests in IT with clear goals of improving efficiency, etc., could receive a *Yes*. A de-layered management structure that empowers front line managers and that has undergone competitive sourcing (if necessary) would also contribute to a *Yes* answer. For mandatory programs, a *Yes* could require the program to seek policies (e.g., through review of proposals from States) that would reduce unit costs. Also consider if, where possible, there is cross-program and inter-agency coordination on IT issues to avoid redundancies.

For R&D programs, efficiency measures can include program overhead costs, average times to fund competitive awards, etc.

Evidence/Data: Evidence can include efficiency measures, competitive-sourcing plans, and IT improvement plans designed to produce tangible productivity and efficiency gains, or IT business cases that document how particular projects improve efficiency.

3.5. Does the program collaborate and coordinate effectively with related programs?

Purpose of the question: to determine whether a Federal program collaborates with other related program(s) in a meaningful way.

Elements of a *Yes* answer: A *Yes* answer would require that the program collaborate with related Federal programs and, to the extent appropriate or possible, with related State, local, and private programs. A *Yes* answer would require that the program show evidence of collaboration leading to meaningful actions in management and resource allocation. For example, the existence of a coordinating council would not by itself constitute meaningful collaboration. (This question applies to programs that have interrelated, but separately budgeted, efforts. An example of an interrelated Federal program is the shared effort of the Department of Veterans Affairs and the Medicare Program to provide care for aging veterans.) Meetings, discussions groups, and task forces are not sufficient for a *Yes*. A *Yes* requires evidence of collaboration leading to meaningful actions in management and resource allocation.

Evidence/Data: Evidence of meaningful collaboration could include joint grant announcements, planning documents, performance goals, or referral systems.

3.6. Does the program use strong financial management practices?

Purpose of the question: to determine whether the program uses effective financial management practices in administering program funds.

Elements of a *Yes* answer: A *Yes* answer would require that, at a minimum, the program be free of material internal control weaknesses reported by auditors. In addition, depending on the particular risks inherent to the program, a *Yes* may require meeting some or all of the following criteria:

- The program has procedures in place to ensure that payments are made properly for the intended purpose to minimize erroneous payments.
- Financial management systems meet statutory requirements.
- Financial information is accurate and timely.
- Integrated financial and performance systems support day-to-day operations.
- Financial statements receive a clean audit opinion and have no material internal control weaknesses.

If an agency-wide material weakness exists that is unrelated to the program, then a *Yes* response would be appropriate. However, if an agency-wide material weakness has a direct relation to the program (e.g., a lack of systems that support day-to-day operations), then the program would receive a *No* for this question.

For block-grant programs which support a wide range of purposes and allow grantees to set their own program priorities, a *Yes* answer should also require that grantees (e.g., States and localities) meet the criteria for effective financial management identified above.

For credit programs, a *Yes* answer would also require that the program consistently meets the requirements of the Federal Credit Reform Act of 1990, the Debt Collection Improvement Act and applicable guidance under OMB Circular A-129.

Evidence/Data: Evidence can include recent audit reports and existence of procedures to identify the above-listed criteria, such as the ability to measure improper payments.

Unlike Question 3.3, the presence of significant erroneous payments does not automatically generate a *No* for this question. Specifically, if an agency has instituted strong financial management controls that are demonstrating measurable improvements in erroneous payments over time, then a *Yes* may be appropriate. Nevertheless, a *No* on Question 3.3 is strong evidence that a *No* may be appropriate for this question.

For credit programs, evidence can include actual reports detailing the performance of the agency's portfolio management, subsidy calculations, re-estimates, modifications, etc. Other evidence can include independent evaluations of the program's performance.

3.7. Has the program taken meaningful steps to address its management deficiencies?

Purpose of the question: to determine whether the program has developed a system of evaluating program management and correcting deficiencies when they are identified. This question should include, but is not limited to, financial management or other Presidential Management Agenda deficiencies. However, the focus of the question is program-level deficiencies, as opposed to agency-level deficiencies that may not directly affect the program.

Elements of a *Yes* answer: A *Yes* answer would require that the program has a system for identifying and correcting program management deficiencies and uses the system to make necessary corrections. A program that does not review program management activities and make corrections to eliminate identified deficiencies would receive a *No*.

Evidence/Data: Evidence can include a description of how deficiencies in the program management are identified and corrected as well as examples of such changes.

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?

Purpose of the question: to determine whether or not grant funds are distributed according to a competitive process so that the most meritorious applications are awarded and so that new applicants of merit will be able to compete fairly with previous grant recipients. (For R&D competitive grants programs, this question is central to addressing the R&D “quality” criterion (see Appendix A).)

Elements of a Yes answer: A Yes answer would require that the overwhelming majority of awards (95 percent or more of funding) are distributed according to a competitive process. Elements of the process can include independent merit review and ranking of applications, as well as a limit on the percentage of funds that are earmarked. A Yes answer would also require that the program operate a fair and open competition and provide a reasonable amount of outreach to encourage the participation of new grantees. Fair and open does not mean that data should not be protected (e.g., to meet classification needs, to protect personal data or, for an appropriate span of time, intellectual property.) or that other reasonable qualification standards can't be imposed (e.g. meets status as an academic organization). Considerations can include whether the program tends to provide grants to the same list of grantees year after year. Awards may be renewed and considered competitive if the original award was competitively awarded, renewals only extend for a short and definite period of time and the possibility of renewal is announced in the original competition..

For R&D programs, the same competitive hurdle should apply for external projects, although a tightly defined internal program can be maintained if the agency can clearly show that it is needed to provide for the maintenance of scientific expertise in-house. Unsolicited, potentially breakthrough ideas and out-of-the-box proposals can be part of the portfolio if they are merit-reviewed from the beginning. If such proposals compete fairly against a significant number of other proposals for a share of funding, they may be considered to meet the competitive standard even if they don't respond to a specific request for proposal. In addition to grants, allowable transactions may include mechanisms such as contracts or other transactions. (Merit review processes and ranking of applications should also apply to these transaction vehicles.)

Evidence/Data: Evidence can include a description of the awards process, percentage of funds earmarked, and percentage of funds subject to peer review. Evidence can also include the relative number of new awardees per award cycle and technical assistance and outreach efforts of the agency.

3.CO2. Does the program have oversight practices that provide sufficient knowledge of grantee activities?

Purpose of the question: to determine whether or not the program has an understanding of how its funds are utilized by grantees.

Elements of a Yes answer: A *Yes* answer would require that a program have sufficient oversight capacity. This capacity may be demonstrated by a program that has a reporting system in place to document grantees' use of funds in eligible activity categories, conducts site visits to a substantial number of grantees on a regular basis, audits grantee performance, and tracks actual expenditures to verify that funds are used for their designated purpose. A program with a strong relationship to its grantees and a high level of understanding of what grantees do with the resources allocated to them would receive a *Yes*.

A program with a poor reporting system to track expenditures by grantees would receive a *No* rating.

For R&D programs, the above standards would apply, but in addition to grants and grantees they would apply to contracts, cooperative agreements and other transactions and their awardees.

Evidence/Data: Evidence can include the reporting structure, oversight techniques, audit or site visit schedule, and/or an assessment of program data quality.

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?

Purpose of the question: to determine whether or not the program has a system in place to collect and present publicly information that captures the most important impacts of program performance.

Elements of a Yes answer: A *Yes* would require that the program collects, compiles and disseminates grantee performance information in an accessible manner, such as via a web site or widely available program reports. Data would be both aggregated on a program-wide level and disaggregated at the grantee level. The assessment about the appropriate level of aggregation of results may depend upon needs to protect certain data, such as classified data, personal data or, for a limited span of time, intellectual property.

A program would receive a *No* if grantee performance information is not available to the public, or if it is only aggregated at a high level. Similarly, a program could receive a *No* response if the data it presents are not related to the impact of the program.

For R&D programs, the above standards would apply, but in addition to grants and grantees would apply to contracts, cooperative agreements and other transactions and their awardees.

Evidence/Data: Evidence can include citations of the types of data that are collected and disseminated as well as a description of how these data are made available.

Block/Formula Grant Programs

3.BF1. Does the program have oversight practices that provide sufficient knowledge of grantee activities?

Purpose of the question: to determine whether or not the program has an understanding of how its funds are utilized by grantees.

Elements of a *Yes* answer: A *Yes* answer would require that a program have sufficient oversight capacity. This capacity may be demonstrated by a program that has a reporting system in place to document grantees' use of funds in eligible activity categories, conducts site visits to a substantial number of grantees on a regular basis, audits grantee performance, and tracks actual expenditures to verify that funds are used for their designated purpose. A program with a strong relationship to its grantees and a high level of understanding of what grantees do with the resources allocated to them would receive a *Yes*.

A program with no reporting system to track expenditures by grantees would receive a *No*.

Evidence/Data: Evidence can include the reporting structure, oversight techniques, audit or site visit schedule, and/or an assessment of program data quality.

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?

Purpose of the question: to determine whether or not the program has a system in place to collect and present publicly information that captures the most important impacts of program performance.

Elements of a *Yes* answer: A *Yes* answer would require the program collects, compiles and disseminates grantee performance information in an accessible manner, such as via a web site or widely available program reports. Data would be both aggregated on a program-wide level and disaggregated at the grantee level.

A program would receive a *No* if grantee performance information is not available to the public, or if it is only aggregated at a high level. Similarly, a program could receive a *No* response if the data it presents are not related to the impact of the program.

Evidence/Data: Evidence can include citations of the types of data that are collected and disseminated as well as a description of how these data are made available.

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?

Purpose of the question: to determine the level of coordination, during the rulemaking process, with parties affected by the regulations.

Elements of a Yes Answer: A *Yes* would require the program solicits the opinions of affected parties on significant regulations and thoroughly evaluates the concerns and suggestions raised by these entities. For example, a program that seeks the opinions of affected parties and incorporates their suggestions or explains why other suggestions were not incorporated during the rule making process could receive a *Yes*.

If the program drafts its rules in a vacuum without consulting any of the potentially affected parties, it would likely receive a *No*. While the element of seeking views is mandated by law, the assessment should consider the extent to which the program takes those views into account.

Evidence/Data: Evidence can include notices seeking public comment and addressing comments in final rules, regulation preambles which discuss compliance with the Regulatory Flexibility Act, Unfunded Mandates Act of 1995, Small Business Regulatory Enforcement Fairness Act (SBREFA) of 1996, E.O. 13132, and National Environmental Policy Act, and detailed preamble language discussing how public comments were considered and addressed.

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?

Purpose of the question: to determine whether the program, in justifying its rules, prepares sound analyses (i.e., cost-benefit analysis, risk analysis) that are rigorous, thorough, and based upon the best available data and consistent with OMB's economic analysis guidelines.

Elements of a Yes Answer: A *Yes* answer should generally include, but is not limited to, a statement of need of the proposed action, an examination of alternative approaches, and an analysis of the incremental benefits and costs of the proposed action. In accordance with OMB's economic guidelines, programs' regulatory actions should maximize net benefits; consider regulatory and non-regulatory alternatives; monetize regulatory costs and benefits; and choose the alternative that maximizes net benefits in their regulatory analyses. For example, programs that fully document the impacts on public health and safety and the regulated industry through a thorough

benefit, cost and risk analysis based upon the best possible available data, and examine other regulatory and non-regulatory alternatives would receive a *Yes*. A program may receive a *Yes* if its analyses are subjected to peer review by government entities, academia, industry, or non-profit research organizations; however, peer review is not required for a program to receive a *Yes*.

If a program's impact analyses fail to include a discussion of the costs of restrictions on the regulated industry, a *No* response to this question would be appropriate. If the program certifies that regulations would not have a significant economic impact on a substantial number of small entities, the program must provide adequate justification for that conclusion to get a *Yes* answer.

Evidence/Data: Evidence can include regulatory impact analyses, regulatory flexibility analyses, and cost-benefit analyses for the program's rules, any reports or feedback generated by outside reviewers, and coordination between reviewers and the sponsoring agency or program.

3.RG3. Does the program systematically review its current regulations to ensure consistency among all regulations in accomplishing program goals?

Purpose of the question: to determine whether the agency met the goal intended when developing the regulation. It should be clear that the program consists of only those regulations that are: (1) necessary in achieving its goals, (2) relevant to the current societal and economic situation, and (3) complementary and consistent with each other.

Elements of a *Yes* Answer: A *Yes* answer would require a program to review its regulations periodically (e.g., every two years) to ensure that they were consistent with program policies. A consideration would include whether the program makes attempts to minimize regulatory burden through constant review of regulations, with an eye toward streamlining, if possible. An additional factor to consider is whether the program ensures that every regulation is consistent with the program's goals. An example of a *Yes* could be a program that conducts look-back studies every third year on all of its significant regulations to ensure that they are all current, consistent, and relevant to the program goals, and, if the review concludes that a regulation is no longer necessary, the program proposes or takes action to remedy the situation.

If a program, however, continues to enforce regulations that are no longer justified and/or necessary, the program would receive a *No*. In addition, a program that has not reviewed its regulations for consistency and necessity in a significant time period would receive a *No*.

Evidence/Data: Evidence should include:

- a program plan or process to conduct this exercise on a regular basis;
- documentation (such as assessments, evaluations, or examinations; planned or completed) generated as a result of the above processes or plans; and/or

- as a result of this type of review, any changes made to or eliminations from the program or its regulations, or a justification that no change is needed.

NOTE: Questions 2.RG1, 3.RG3, and 4.RG1 address the progression of reviews for a program's rules. Specifically, 2.RG1 addresses targeted development of regulations, 3.RG3 addresses effective implementation of regulations, and 4.RG1 addresses maximization of potential benefits during the regulation's implementation.

3.RG4. Are the regulations designed to achieve program goals, to the extent practicable, by maximizing the net benefits of its regulatory activity?

Purpose of the question: to determine whether the program, as it promulgates regulations, ensures that its regulatory requirements, in total, maximize net benefits. (Note that this question relates to the promulgation of regulations, as opposed to their implementation.)

Elements of a Yes Answer: A Yes answer would require the program to maximize overall net benefits to the greatest extent practicable. The program should show that it makes the best effort to assess how each additional regulation adds to the current level of regulatory requirements and keeps regulatory compliance burden at a minimum, including the burden associated with information collection. Any additional compliance burdens should be shown to increase benefits substantially. An important consideration for this question is whether in promulgating its regulations, for a given level of benefits, the agency allows alternative methods for compliance, record keeping, and reporting to minimize the cost burden on regulated entities (including electronic means). Programs should be judged within their statutory framework; the program should maximize net benefits to the extent allowed by statute in order to get a Yes.

Evidence/Data: Evidence can include statistics on compliance reporting burden and the costs of the program's requirements on regulated industries in total, or evidence from a thorough cost-benefit analysis.

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?

Purpose of the question: to determine whether the agency has clearly identified and defined the required quality, capability, and performance characteristics or objectives expected of the end product/result of the asset or service acquisition. This element is critical because it assures that all parties (government, contractor, etc) are working toward the same end-product and result.

Elements of a Yes answer: If acquiring a capital asset, a *Yes* would require the program to validate planning decisions (procurement phase step III.1 of the [Capital Programming Guide](#)), document the capabilities or characteristics that are expected, and to make management decisions based on whether milestones are being met. For example, a weapon system that has defined key performance parameters and operational requirements would get a *Yes*, one that is proceeding without such definition should receive a *No*. For services, a *Yes* would require the program made adequate use of performance-based and fixed-priced type contracts. A program that acquires services through other than performance-based and fixed-priced type contracts should receive a *No*, unless there is a legitimate reason for not using such contracts.

Evidence/Data: Evidence can include program documentation of planning decision validation for asset acquisitions, and documentation describing key performance characteristics and/or deliverables and demonstration that this information is used appropriately in management decisions.

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?

Purpose of the question: to determine whether the program agency and its partners manage the financial performance of their credit programs.

Elements of a Yes answer: A *Yes* answer would require managing the program based on the results of an effective monitoring system that tracks the financial performance of each credit facility. Collection and analysis of borrower repayment streams should be part of the evaluation process and could be coupled with reports from or trips to the field. The evaluation should also include an analysis of agency efforts to reduce default rates, and maximize collections and recoveries.

Evidence/Data: Evidence can include quarterly financial statements from the program, agency, Treasury, the guaranteed lender, loan servicing agent; internal evaluations, external independent performance evaluations; reports from field representatives or trips to the field on the borrowers' performance.

3.CR2. Do the program's credit models adequately provide reliable, consistent, accurate and transparent estimates of costs and the risk to the Government?

Purpose of the questions: to determine whether the program uses a reliable method for estimating program costs.

Elements of a Yes answer: A *Yes* answer would require that the agency employ a rigorous cost-estimation model that adequately accounts for the government's risk and generates dependable cost estimates for each cohort. A program whose cost estimates routinely differ markedly from actual results should receive a *No*.

Evidence/Data: Evidence can include a description of any problems or advantages of the program's cost estimation model. For example, a strong model could employ well established statistical estimation techniques that have a solid track record for predicting borrower activity. However, a weaker model could fail to account for potential contingencies that increase government risk.

Research and Development Programs

R&D programs addressing the acquisition, construction or operation of facilities or other capital assets should answer the Capital Assets and Service Acquisition question (3.CA1).

R&D programs that use competitive grants, contracts, cooperative agreements or other transactions should answer the Competitive Grants questions (3.CO1, CO2, 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?

Purpose of the question: to determine whether the program uses a clearly stated, defensible method for allocating its R&D funding. This question is central to addressing the R&D “quality” criterion (see Appendix A).

Elements of a Yes answer: A *Yes* answer would require that the program allocate funding using a broadly competitive process based on merit, or that it has compelling justifications for R&D funding allocated through other means. Interpretations of competition and merit review should be consistent with the definitions in Circular A-11: “...intramural and extramural research programs where funded activities are competitively awarded following review for scientific and technical merit.” All program funds allocated through means other than unlimited competition must document the processes they use to distribute funds to each type of R&D performer (e.g., federal laboratories, federally funded R&D centers, universities, etc.). Programs are encouraged to use external assessment of the methods they use to allocate R&D and maintain program quality.

Evidence/Data: Evidence can include a description of the awards process, percentage of funds earmarked, percentage of funds subject to competitive peer review, and results of external assessments.

IV. PROGRAM RESULTS/ACCOUNTABILITY

This section considers whether a program is meeting its long-term and annual performance goals. This section also assesses how well the program compares to similar programs and how effective the program is based on independent evaluations. Potential source documents and evidence for answering questions in this section include GPRA performance reports, evaluations, GAO reports, IG reports and other agency documents. Assessments of program results should be based on the most recent reporting cycle or other relevant data. The Measures tab in the PART worksheet contains data fields for a performance targets and results, and should be completed to the greatest extent possible for all measures agreed to by OMB and the agency for Section II.

Answers in this section are rated as *Yes*, *Large Extent*, *Small Extent*, and *No*. Like Sections I-III, the scoring system in this section remains on a 0 to 1 point scale. Scoring for this section differs by including the option of partial credit between 0 and 1 in increments of 0, .33, .67, and 1.

In general, *Not Applicable* answers are not appropriate for Questions 4.1 and 4.2. While it is recognized that some programs may have great difficulty developing quantitative performance goals, programs are strongly encouraged to have some meaningful and appropriate methods for demonstrating results. OMB and agencies should work together to develop approaches for programs where it is difficult to develop quantitative measures, and where qualitative, expert-review, or other measures are more appropriate. *Not Applicable* is a potential answer for Question 4.3 if the program is already operating at very high efficiency levels. It may also be a possible answer for Question 4.4 and, with specific justification, for Question 4.5. (For R&D programs, most of the questions in this section help address the *retrospective* aspects of the R&D Investment Criteria, with emphasis on the “performance” criterion (see Appendix A).)

4.1. Has the program demonstrated adequate progress in achieving its long-term performance goals?

Purpose of the question: to determine whether the program is meeting or making measurable progress toward meeting the long-term performance goals evaluated in Questions 2.1 and 2.2. The question also seeks to determine whether the program's partners are meeting long-term goals evaluated in Question 2.5, if partner performance is critical to the program achieving its goals. Examples of partners can include grantees, participating financial institutions, regulated bodies, or suppliers.

Elements of a *Yes* answer: A *Yes* answer (i.e., full credit) would require that the program is on track to meet all the long-term performance goals – including ambitious targets and timeframes – evaluated in Questions 2.1 and 2.2. A program would not receive a *Yes* answer by simply meeting any one of its long-term targets, or by having performance measures but no ambitious targets and timeframes. A *Yes* answer would also require that, where applicable, partners commit to long-term

outcome targets and achieve them as well. Where relevant, a *Yes* answer would also require that a program has addressed appropriately any predefined end targets.

Partial credit, such as *Large Extent* or *Small Extent*, should be given in cases where there is partial, but notable, achievement of long-term targets. A program could receive a *No* if it had received a *Yes* for achieving its annual targets (next question), but is not making progress toward meeting its long-term goals.

Additional rating guidance:

- **If adequate outcome (or output) measures are not available and a program received a *No* in Question 2.1, the program must receive a *No* answer to this question.**
- **If the program received a *Yes* in Question 2.1 and a *No* in Question 2.2, then the program cannot receive a rating higher than *Small Extent*.**

The only exceptions to this guidance are in cases in which OMB has approved the use of alternative forms of assessment, as discussed in the Section IV overview.

Evidence/Data: Evidence can include data from the agency's GPRA performance report, a strategic plan, or other Administration goals and objectives. Reports detailing customer satisfaction with program performance, program reports detailing rates of utilization or participation, or independent evaluations of the program's performance may also be considered as relevant evidence. In cases where targets are not met, additional evidence can include an explanation of the main reasons.

Space is provided in the Measures tab of the PART worksheet to list and document goals, targets and achieved results. Only measures that meet the standards for a *Yes* should be entered on the worksheet.

4.2. Does the program (including program partners) achieve its annual performance goals?

Purpose of the question: to determine whether the program is meeting the targets evaluated in Question 2.4. The question also seeks to determine whether the program's partners are meeting annual targets evaluated in Question 2.5, if partner performance is critical to the program achieving its overall targets. Examples of partners can include grantees, contractors, participating financial institutions, regulated bodies, or suppliers.

Elements of a *Yes* answer: A *Yes* answer (i.e., full credit) would require that the program meet all the annual performance targets evaluated in Question 2.4. A *Yes* answer would also require the program received a *Yes* for Questions 2.1 and 2.3, and a *Yes* or *Not Applicable* for Question 2.5. A program would not receive a *Yes* answer by simply meeting any one of its annual targets. A *Yes* answer would also require that, where applicable, partners commit to annual targets and achieve them as well.

Partial credit such as *Large Extent* or *Small Extent*, should be given in cases where there is partial, but notable, achievement of targets.

Additional rating guidance:

- **If a program received a *No* in Question 2.3, the program must receive a *No* answer to this question.**
- **If the program received a *Yes* in Question 2.3 and a *No* in Question 2.4, then the program cannot receive a rating higher than *Small Extent*.**

The only exceptions to this guidance are in cases in which OMB has approved the use of alternative forms of assessment, as discussed in the Section IV overview.

Evidence/Data: Evidence can include data from the agency's annual GPRA performance report, a strategic plan, or other Administration goals and objectives. In cases where targets are not met, additional evidence can include an explanation of the main reasons.

Space is provided in the Measures tab of the PART worksheet to list and document goals, targets and achieved results. Only measures that meet the standards for a *Yes* should be entered on the worksheet.

4.3. *Does the program demonstrate improved efficiencies or cost effectiveness in achieving program goals each year?*

Purpose of the question: to determine whether management practices have resulted in efficiency gains over the past year.

Elements of a *Yes* answer: A *Yes* would require that the program demonstrate improved efficiency or cost effectiveness over the prior year. Efficiency improvements should generally be measured in terms of dollars or time. Programs that complete an A-76 competition – an indicator of cost-efficient processes – would also likely be eligible for a *Yes* answer, provided that the competition addresses the program's key cost and performance drivers. Also, programs that clearly demonstrate very high levels of efficiency through other means may receive a *Yes* without documenting increasing efficiency over time. **A program would normally not be eligible for a *Yes* answer to this question if it received a *No* in Question 3.4.**

Evidence/Data: Evidence can include meeting performance targets to reduce per unit costs or time, meeting production and schedule targets; or meeting other targets that result in tangible productivity or efficiency gains. Efficiency measures may also be considered in Questions 4.1 and 4.2.

4.4. Does the performance of this program compare favorably to other programs, including government, private, etc., with similar purpose and goals?

Purpose of the question: to determine how well the program performs relative to other programs engaged in a similar activity.

Elements of a Yes answer: A *Yes* answer would require the program compare favorably to other programs with similar purpose and goals. Programs are not limited to Federal government and can include State and local government and the private sector. The user should consider relevant evaluations that allow a comparison of programs with similar purpose and goals. A *Not Applicable* rating is appropriate if 1) no comparable federal, state, local government, or private sector programs exist, or 2) the comparison would be too inherently difficult and costly to perform for the foreseeable future. The explanation for *Not Applicable* should explain why the comparison is inherently too difficult to perform. (This question is not limited to comparisons of programs with explicitly coordinated “common measures.”)

For capital assets and service acquisition programs, review of performance should include cost/schedule adherence, quality, and quantity of deliverables.

Evidence/Data: Evidence can include evaluations and documentation comparing similar programs.

4.5. Do independent evaluations of sufficient scope and quality indicate that the program is effective and achieving results?

Purpose of the question: to determine whether the program is effective based on independent and comprehensive evaluations. This question may be particularly important for programs that have substantial difficulty formulating quantitative performance measures. (For R&D programs, this question is central to retrospective assessment of all of the R&D criteria (see Appendix A).)

Elements of a Yes answer: A *Yes* answer would require that independent program evaluations indicate that the program is effective. (If the evaluations of Question 2.6 address past effectiveness and accomplishments, they may be used as evidence for this question as well.) The quality of evaluations of the program presented in Question 2.6 should strongly be considered in answering this question. The most definitive data supporting a program’s overall effectiveness would be from a randomized controlled trial, when appropriate and feasible. Data from other evaluation methods, such as quasi-experimental and non-experimental, can be considered as detailed in Question 2.6, but should be scrutinized given the increased possibility of an erroneous conclusion. If a program is taking necessary steps to correct deficiencies uncovered by the evaluation, the user should address this effort in Question 3.7.

Relevant evaluations would be at the national program level, rather than evaluations of one or more program partners, and would not focus only on process indicators such as the number of grants provided, or hits on a web site. Relevant evaluations would consider a program's impact and effectiveness. Evaluations conducted by the program itself should not be considered “independent.” However, if the program has contracted out the evaluation to a third-party, it might be considered independent. Evaluations conducted by an agency’s Inspector General or program-evaluation office also might be considered “independent.” OMB examiners and agency staff will determine if a specific evaluation can be considered “independent” for this question.

Not Applicable is not an option for this question; given the flexibility in determining what constitutes an evaluation, all programs should undergo an evaluation that meets the elements of quality, scope, and independence detailed in Question 2.6.

Evidence/Data: Evidence can include findings of an evaluation conducted by the General Accounting Office, Inspectors General, academic and research institutions, agency contracts, or other independent entities.

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?

Purpose of the question: to determine whether the program met its goals in the most efficient way possible. It should be determined whether the program maximized net benefits through implementation of its regulatory actions (as opposed to regulatory development). In calculating the incremental costs of a new regulation, these costs should be compared to a baseline or, in a small number of cases, a less stringent alternative. This question deals with the actual implementation of the regulatory action, not just the conception and promulgation of the regulatory action.

Elements of a Yes answer: A *Yes* answer would require that the program’s implementation of the regulatory action maximizes net benefits. For example, a Department of Transportation maximum load regulation that demonstrates that 1) the realized benefits to health and safety outweigh the realized incremental costs of compliance, and 2) the net benefits of the regulatory approach adopted are higher than reasonable alternatives, would receive a *Yes*.

If a program’s regulations result in greater incremental costs than benefits, or if the program does not analyze alternative approaches to demonstrate that its regulatory approach maximized net benefits, the program should get a *No*.

Evidence/Data: Evidence can include Regulatory Impact Analyses (RIA) or other supporting programmatic analyses (with supporting data after implementation that assumptions were correct), look-back studies, independent evaluations, or additional impact analyses using retrospective data. If a *No* answer is attributable to statutory requirements to regulate despite the fact that incremental costs exceed benefits, these statutory requirements should be included in the evidence section.

NOTE: Questions 2.RG1, 3.RG3, and 4.RG1 address the progression of review for a program's rules. Specifically, 2.RG1 addresses targeted development of regulations, 3.RG3 addresses effective implementation of regulations, and 4.RG1 addresses maximization of potential benefits during the regulation's implementation.

Capital Assets and Service Acquisition Programs

4.CA1. Were program goals achieved within budgeted costs and established schedules?

Purpose of the question: to determine whether valid program goals were achieved within budgeted costs and established schedules and whether the program spends funds as planned and budgeted. For capital assets procured and in operation, this question also addresses management-in-use (i.e., the operations and disposal costs).

Elements of a *Yes* answer: A *Yes* answer would require that the program achieved the goals evaluated in Section II on budget and on schedule. An example of a program that could receive a *No* rating could be an acquisition program that has experienced 60 percent cost growth and is behind schedule. If a program's cost and schedule targets were changed in the last 12 months specifically due to failure to achieve previous goals, the program should get a *No*.

Evidence/Data: Evidence can include a comparison of the contract schedule, deliverables, and costs with the final outcomes for that fiscal year.

Research and Development Programs

R&D programs addressing the acquisition, construction or operation of facilities or other capital assets should answer the Capital Assets and Service Acquisition question (4.CA1).

Appendix A

For Research and Development Programs: R&D Investment Criteria

As another initiative of the President's Management Agenda, the development of explicit R&D investment criteria builds on the best of the planning and assessment practices that R&D program managers use to plan and assess their programs. The Administration has worked with experts and stakeholders over the past two years to build upon lessons learned from previous approaches.

This year, the content of the R&D investment criteria has not changed from last year, but OMB has worked to clarify its implementation. Specifically, agencies should use the criteria as broad guidelines that apply at all levels of federally funded R&D efforts, and they should use the PART as the instrument to periodically evaluate compliance with the criteria at the program level. To make this possible, the R&D PART has been modified to clarify its alignment with the R&D criteria. The R&D criteria are reprinted here as a guiding framework for addressing the R&D PART. More details will follow on other aspects of R&D criteria implementation this year.

The R&D criteria address not only planning, management, and prospective assessment but also retrospective assessment. Retrospective review of whether investments were well-directed, efficient, and productive is essential for validating program design and instilling confidence that future investments will be wisely invested. Retrospective reviews should address continuing program relevance, quality, and successful performance to date.

While the criteria are intended to apply to all types of R&D, the Administration is aware that predicting and assessing the outcomes of *basic* research in particular is never easy. Serendipitous results are often the most interesting and ultimately may have the most value. Taking risks and working toward difficult-to-attain goals are important aspects of good research management, and innovation and breakthroughs are among the results. However, there is no inherent conflict between these facts and a call for clearer information about program goals and performance toward achieving those goals. The Administration expects agencies to focus on improving the management of their research programs and adopting effective practices, and not on predicting the unpredictable.

The R&D investment criteria have several potential benefits:

- Use of the criteria allows policy makers to make decisions about programs based on information beyond anecdotes, prior-year funding levels, and lobbying of special interests.
- A dedicated effort to improve the process for budgeting, selecting, and managing R&D programs is helping to increase the return on taxpayer investment and the productivity of the federal R&D portfolio.
- The R&D investment criteria will help communicate the Administration's expectations for proper program management.

- The criteria and subsequent implementation guidance will also set standards for information to be provided in program plans and budget justifications.
- The processes and collected information promoted under the criteria will improve public understanding of the possible benefits and effectiveness of the federal investment in R&D.

The R&D Investment Criteria

The Relevance, Quality, and Performance criteria apply to all R&D programs. Industry- or market-relevant applied R&D must meet additional criteria. Together, these criteria can be used to assess the need, relevance, appropriateness, quality, and performance of federal R&D programs.

I. Relevance

R&D investments must have clear plans, must be relevant to national priorities, agency missions, relevant fields, and “customer” needs, and must justify their claim on taxpayer resources. Programs that directly support Presidential priorities may receive special consideration with adequate documentation of their relevance. Review committees should assess program objectives and goals on their relevance to national needs, “customer” needs, agency missions, and the field(s) of study the program strives to address. For example, the Joint DOE/NSF Nuclear Sciences Advisory Committee’s Long Range Plan and the Astronomy Decadal Surveys are the products of good planning processes because they articulate goals and priorities for research opportunities within and across their respective fields.

OMB will work with some programs to identify quantitative metrics to estimate and compare potential benefits across programs with similar goals. Such comparisons may be within an agency or among agencies.

A. Programs must have complete plans, with clear goals and priorities.

Programs must provide complete plans, which include explicit statements of:

- specific issues motivating the program;
- broad goals and more specific tasks meant to address the issues;
- priorities among goals and activities within the program;
- human and capital resources anticipated; and
- intended program outcomes, against which success may later be assessed.

B. Programs must articulate the potential public benefits of the program.

Programs must identify potential benefits, including added benefits beyond those of any similar efforts that have been or are being funded by the government or others. R&D benefits may include technologies and methods that could provide new options in the future, if the landscape of today’s needs and capabilities changes dramatically. Some programs and sub-program units may be required to quantitatively estimate expected benefits, which would include metrics to permit meaningful comparisons among programs that promise similar benefits. While all

programs should try to articulate potential benefits, OMB and OSTP recognize the difficulty in predicting the outcomes of basic research. Consequently, agencies may be allowed to relax this as a requirement for basic research programs.

C. Programs must document their relevance to specific Presidential priorities to receive special consideration.

Many areas of research warrant some level of federal funding. Nonetheless, the President has identified a few specific areas of research that are particularly important. To the extent a proposed project can document how it directly addresses one of these areas, it may be given preferential treatment.

D. Program relevance to the needs of the Nation, of fields of Science & Technology, and of program “customers” must be assessed through prospective external review.

Programs must be assessed on their relevance to agency missions, fields of science or technology, or other “customer” needs. A customer may be another program at the same or another agency, an interagency initiative or partnership, or a firm or other organization from another sector or country. As appropriate, programs must define a plan for regular reviews by primary customers of the program’s relevance to their needs. These programs must provide a plan for addressing the conclusions of external reviews.

E. Program relevance to the needs of the Nation, of fields of S&T, and of program “customers” must be assessed periodically through retrospective external review.

Programs must periodically assess the need for the program and its relevance to customers against the original justifications. Programs must provide a plan for addressing the conclusions of external reviews.

II. Quality

Programs should maximize the quality of the R&D they fund through the use of a clearly stated, defensible method for awarding a significant majority of their funding. A customary method for promoting R&D quality is the use of a competitive, merit-based process. NSF’s process for the peer-reviewed, competitive award of its R&D grants is a good example. Justifications for processes other than competitive merit review may include “outside-the-box” thinking, a need for timeliness (e.g., R&D grants for rapid response studies of *Pfisteria*), unique skills or facilities, or a proven record of outstanding performance (e.g., performance-based renewals).

Programs must assess and report on the quality of current and past R&D. For example, NSF’s use of Committees of Visitors, which review NSF directorates, is an example of a good quality-assessment tool. OMB and OSTP encourage agencies to provide the means by which their programs may be benchmarked internationally or across agencies, which provides one indicator of program quality.

A. Programs allocating funds through means other than a competitive, merit-based process must justify funding methods and document how quality is maintained.

Programs must clearly describe how much of the requested funding will be broadly competitive based on merit, providing compelling justifications for R&D funding allocated through other means. (See OMB Circular A-11 for definitions of competitive merit review and other means of allocating federal research funding.) All program funds allocated through means other than unlimited competition must document the processes they will use to distribute funds to each type of R&D performer (e.g., federal laboratories, federally funded R&D centers, universities, etc.). Programs are encouraged to use external assessment of the methods they use to allocate R&D and maintain program quality.

B. Program quality must be assessed periodically through retrospective expert review.

Programs must institute a plan for regular, external reviews of the quality of the program's research and research performers, including a plan to use the results from these reviews to guide future program decisions. Rolling reviews performed every 3-5 years by advisory committees can satisfy this requirement.

Benchmarking of scientific leadership and other factors provides an effective means of assessing program quality relative to other programs, other agencies, and other countries.

III. Performance

R&D programs should maintain a set of high priority, multi-year R&D objectives with annual performance outputs and milestones that show how one or more outcomes will be reached. Metrics should be defined not only to encourage individual program performance but also to promote, as appropriate, broader goals, such as innovation, cooperation, education, and dissemination of knowledge, applications, or tools.

OMB encourages agencies to make the processes they use to satisfy the Government Performance and Results Act (GPRA) consistent with the goals and metrics they use to satisfy these R&D criteria. Satisfying the R&D performance criteria for a given program should serve to set and evaluate R&D performance goals for the purposes of GPRA. OMB expects goals and performance measures that satisfy the R&D criteria to be reflected in agency performance plans.

Programs must demonstrate an ability to manage in a manner that produces identifiable results. At the same time, taking risks and working toward difficult-to-attain goals are important aspects of good research management, especially for basic research. The intent of the investment criteria is not to drive basic research programs to pursue less risky research that has a greater chance of success. Instead, the Administration will focus on improving the management of basic research programs.

OMB will work with some programs to identify quantitative metrics to compare performance across programs with similar goals. Such comparisons may be within an agency or among agencies.

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. Within DOE, the Office of Science's formalized independent reviews of technical cost, scope, and schedule baselines and project management of construction projects ("Lehman Reviews") are widely recognized as an effective practice for discovering and correcting problems involved with complex, one-of-a-kind construction projects.

A. Programs may be required to track and report relevant program inputs annually.

Programs may be expected to report relevant program inputs, which could include statistics on overhead, intramural/extramural spending, infrastructure, and human capital. These inputs should be discussed with OMB.

B. Programs must define appropriate output and outcome measures, schedules, and decision points.

Programs must provide single- and multi-year R&D objectives, with annual performance outputs, to track how the program will improve scientific understanding and its application. Programs must provide schedules with annual milestones for future competitions, decisions, and termination points, highlighting changes from previous schedules. Program proposals must define what would be a minimally effective program and a successful program. Agencies should define appropriate output and outcome measures for all R&D programs, but agencies should not expect fundamental basic research to be able to identify outcomes and measure performance in the same way that applied research or development are able to. Highlighting the results of basic research is important, but it should not come at the expense of risk-taking and innovation. For some basic research programs, OMB may accept the use of qualitative outcome measures and quantitative process metrics. Facilities programs must define metrics and methods (e.g., earned-value reporting) to track development costs and to assess the use and needs of operational facilities over time. If leadership in a particular field is a goal for a program or agency, OMB and OSTP encourage the use of benchmarks to assess the processes and outcomes of the program with respect to leadership. OMB encourages agencies to make the processes they use to satisfy GPRA consistent with the goals and metrics they use to satisfy these R&D criteria.

C. Program performance must be retrospectively documented annually.

Programs must document performance against previously defined output and outcome metrics, including progress toward objectives, decisions, and termination points or other transitions. Programs with similar goals may be compared on the basis of their performance. OMB will work with agencies to identify such programs and appropriate metrics to enable such comparisons.

IV. Criteria for R&D Programs Developing Technologies That Address Industry Issues

The purpose of some R&D and technology demonstration programs and projects is to introduce some product or concept into the marketplace. However, some of these efforts engage in activities that industry is capable of doing and may discourage or even displace industry investment that would occur otherwise. For the purposes of assessing federal R&D investments, the following criteria should be used to assess industry-relevant R&D and demonstration projects, including, at OMB discretion, associated construction activities.

OMB will work with programs to identify quantitative metrics to measure and compare potential benefits and performance across programs with similar goals, as well as ways to assess market relevance.

- A. Programs and projects must articulate public benefits of the program using uniform benefit indicators across programs and projects with similar goals.**
In addition to the public benefits required in the general criteria, *all* industry-relevant programs and projects must identify and use uniform benefit indicators (including benefit-cost ratios) to enable comparisons of expected benefits across programs and projects. OMB will work with agencies to identify these indicators.
- B. Programs and projects must justify the appropriateness of federal investment, including the manner in which the market fails to motivate private sector investment.**
A lack of market incentives discourages private firms from investing in research where the benefits may occur far in the future, the risks may be too great for non-federal participants, or the benefits accrue to the public rather than private investors. Programs and projects must demonstrate that industry investment is sub-optimal and explain in what way the market fails that prevents the private sector from capturing the benefits of developing the good or service.
- C. Programs and projects must demonstrate that investment in R&D and demonstration activities is the best means to support the federal policy goals, compared to other policy alternatives.**
When the federal government chooses to intervene to address market failures, there may be many policy alternatives to address those failures. Among the other tools available to the government are legislation, tax policy, regulatory and enforcement efforts, and an integrated combination of these approaches. In this context, projects to address issues of genuine federal concern should be able to illustrate how R&D and demonstration activities are superior to other policy tools in addressing federal goals, either by themselves or as part of an integrated package.

D. Programs and projects must document industry or market relevance, including readiness of the market to adopt technologies or other outputs.

Programs must assess the likelihood that the target industry will be able to adopt the technology or other program outputs. The level of industry cost sharing is one indicator of industry relevance. Before projects move into demonstration or deployment stages, an economic analysis of the public and private returns on the public investment must be provided.

E. Program performance plans and reports must include “off ramps” and transition points.

In addition to the schedules and decision points defined in the general criteria, program plans should also identify whether, when, and how aspects of the program may be shifted to the private sector.