



An Evaluation Roadmap for a More Effective Government

Revised October 2013



The American Evaluation Association (AEA) is an international professional association of evaluators devoted to the application and exploration of program evaluation, personnel evaluation, technology, and many other forms of evaluation. Evaluation involves assessing the strengths and weaknesses of programs, policies, personnel, products, and organizations to improve their effectiveness. AEA has over 7,500 members representing all 50 states in the United States and more than 75 foreign countries.

Mission:

- *To improve evaluation practices and methods*
- *To increase evaluation use*
- *To promote evaluation as a profession*
- *To support the contribution of evaluation to the generation of theory and knowledge about effective human action*

**An Evaluation Roadmap
for a
More Effective Government**

In keeping with our mission, the American Evaluation Association hereby describes its vision of the role of evaluation in the federal government. We provide a roadmap for improving government through evaluation, outlining steps to strengthen the practice of evaluation throughout the life cycle of programs.

Evaluation is an essential function of government. It can enhance oversight and accountability of federal programs, improve the effectiveness and efficiency of services, assess which programs or policies are working and which are not, and provide critical information needed for making difficult decisions about them.

An Evaluation Roadmap for a More Effective Government

The Challenge

Like all governments, the U.S. government faces challenges in both foreign and domestic policy arenas. Today, these challenges span subject areas such as national security, foreign aid, energy, the environment, health care, education, and the economy. In response, the government typically develops program or policy interventions in an effort to mitigate, resolve, or understand the problems involved.

To determine the merit, quality, and usefulness of these interventions, credible information is needed about what the program or policy in question has achieved and at what cost. Such information is crucial if government officials are to ensure that the chosen interventions are working, that taxpayers' money is being spent wisely, and that the government is accountable to the public for its interventions and their results.

Why Program Evaluation Is Essential

Within a government context, the legitimacy of evaluation can be seen as deriving from the structure of the government it serves and from the functions it fills.

In the United States, evaluation can serve information needs that arise within any of the three branches of government. For example, evaluation can provide information about a new program's initial outcomes, allowing for better management within the Executive Branch. It can also be used to assess the relative merits of a set of alternative policy options, informing legislative deliberations. The courts may cite evaluation findings as a basis for their judgments. More fundamentally, evaluation can contribute the evidence needed to support the system of checks and balances established by the U.S. Constitution. It enhances congressional oversight and executive accountability and, by its very nature, supplies publicly accessible information that is at the heart of transparency and open government.

More specifically, evaluation involves assessing the strengths and weaknesses of individual programs, policies, and organizations to improve their effectiveness, efficiency, and worth. For this purpose, it provides a useful and important tool to bring credible, well-grounded information to bear on a broad range of government decisions. It supports the development of new knowledge, innovation, and organizational learning in both the Legislative and Executive Branches. It uses systematic data collection and analysis to address questions about how well government programs and policies are working, whether they are achieving their objectives, and, no less importantly, why they are or are not effective. Evaluation evidence can be used to compare alternative programs, guide program development and decision making, and reveal effective practices.

Since the inception of modern program evaluation, federal agencies have conducted many evaluations and applied their results to make reasoned program decisions. But for the most part, these evaluations have been sporadic, applied inconsistently, and supported inadequately. Too often, the units formed to conduct evaluations are short lived and under-resourced. Training and capacity building for evaluation have been inconsistent across agencies and, in many cases, insufficient to achieve the needed evaluation capacity and sustain it over time.

Yet there is a strong case to be made for a commitment to evaluation as an integral feature of good government, whether the goal is better performance, stronger oversight and accountability, or more data-informed and innovative decision making. The lessons learned in agencies that have applied evaluation constitute a solid knowledge base upon which to build.

The U.S. government would benefit significantly from using program evaluation in order to:

- Address questions about current and emerging problems
- Inform program and policy planning efforts
- Monitor program performance
- Provide timely feedback to decision makers enabling them to make changes when needed
- Increase accountability and transparency
- Reduce waste and enhance efficiency
- Improve programs and policies in a systematic manner
- Support major decisions about program reform, expansion, or termination
- Identify program implementation and outcome failures and successes
- Identify innovative solutions that work and in what contexts
- Inform the development of new programs where needed
- Examine the requirements for the transfer of promising programs to new sites
- Share information about effective practices across government programs and agencies
- Re-examine program relevance and effectiveness over time.

The key is to make program evaluation integral to managing government programs at all stages, from planning and initial development through start-up, ongoing implementation, appropriations, and reauthorization. In short, what is needed is a transformation of the federal management culture to one that incorporates evaluation as an essential management function.

Recommendations

We recommend that each federal agency adopt the following framework to guide the development and implementation of its evaluation programs.

Scope and Coverage

- Conduct evaluations of public programs and policies throughout their life cycles and use evaluation to both improve programs and assess their effectiveness
- Evaluate federal programs and policies in a manner that is appropriate for program stewardship and useful for decision making
- Build into each new program and major policy initiative an appropriate evaluation framework to guide the program or initiative throughout its life

- For existing programs, assess what is already known and develop evaluation plans to support future decision making

Management

- Assign experienced, senior evaluation officials and managers to administer evaluation centers or coordinate evaluation functions at appropriately high levels of government agencies
- Prepare annual and long-term evaluation plans to guide decision making about programs
- Provide sufficient and stable funding to support evaluations and professional capacity building
- Coordinate and communicate about evaluation efforts across agencies with overlapping or complementary missions
- Develop written evaluation policies across and within federal agencies that can guide evaluation efforts and help ensure their quality
- Ensure that evaluation units and staff receive high-level support

Quality and Independence

- Develop and adopt quality standards to guide evaluation functions consistent with the American Evaluation Association's Guiding Principles for Evaluators
- Promote the use and further development of appropriate methods for designing programs and policies, monitoring program performance, improving program operations, and assessing program effectiveness and cost
- Safeguard the independence of program or policy evaluations with respect to study design, conduct, and results, while allowing for an appropriate level of consultation with and input from agency or staff personnel and other stakeholders
- Preserve and promote objectivity in examining program operations and impact

Transparency

- Consult closely with Congress and non-federal stakeholders in defining program and policy objectives and critical operations and definitions of success
- Disseminate evaluation findings and methods relating to public accountability to policy makers, program managers, and the public
- Create clearinghouses to share information about effective and ineffective program practices

In this Roadmap, we more fully develop these ideas. We describe the general principles that should guide a government-wide effort to strengthen evaluation functions. We propose broad administrative steps to institutionalize evaluation in federal agencies. Finally, we discuss how the Executive Branch and Congress can jointly make the most effective and efficient use of evaluation.

General Principles

The following general principles should guide efforts to integrate evaluation into program management.

Scope

Evaluation should be integral to planning, developing, managing, and implementing government programs at all stages. Evaluation activities should be used in order to:

- Make sure that program designs are appropriate to achieve program goals
- Identify problems during start-up and correct them before they become entrenched
- Identify and share promising approaches that emerge during program implementation
- Assess the extent to which programs and policies are being well implemented
- To the extent feasible, and in consultation with stakeholders, establish expectations and performance standards at program inception and refine them as the programs mature
- Develop data collection and reporting procedures for performance measurement systems that provide a continuing flow of timely information to policymakers and program managers
- Examine the extent to which programs reach their intended beneficiaries
- Examine selected program features periodically to improve their effectiveness and efficiency
- Assess program results and service quality periodically
- Examine systematically whether an apparently successful program can be expanded to another setting before scaling it up

In general, all federal programs and policies should be subject to evaluation.

Analytic Approaches and Methods

Which analytic approaches and methods to use depends on the questions addressed, the kind of program evaluated, the circumstances under which the program is implemented, its implementation status, the budgetary resources available, when the evaluation results are needed, the purpose of undertaking the evaluation, and the intended audience.

No simple answers are available to questions about how well programs work, and no single analytic approach or method can decipher the inherent complexities in the program environment and assess the ultimate value of public programs. Furthermore, definitions of “success” may be contested. A range of analytic methods is needed, and often several methods—including quantitative and qualitative approaches—should be used simultaneously. Some evaluation approaches are particularly helpful in a program’s early developmental stages, whereas others are more suited to ongoing and regularly implemented programs or to ex-post analysis of temporary programs upon their completion.

The broader policy and decision-making context also can influence which approach is most appropriate. Sometimes information is needed quickly, requiring studies that can use existing data or rapid data-collection methods; at other times, more sophisticated long-term studies are required to understand fully the dynamics of program administration and beneficiary behaviors. Moreover, different approaches can complement each other. For example, performance measurement systems can, on a routine basis, provide important, timely information to managers and policy makers about what has happened in programs, and this may, in fact, be sufficient for their immediate information needs. But performance measurement systems cannot in themselves say *why* any changes happened or what changes, if any, would be beneficial. Analyzing changes would require an additional evaluative approach capable of ruling out plausible causes of the results other than the program.

Over the years, the evaluation field has developed an extensive array of analytic approaches and methods that can be applied and adapted to a wide variety of programs, depending on the program's characteristics and implementation stage, the way the results will be used, and the kinds of decisions that will be made. All evaluation methods should be context-sensitive, have cultural relevance, and be methodologically sound. Evaluation approaches and methods include, but are not limited to, the following:

- Case studies
- Surveys
- Quasi-experimental designs
- Randomized field experiments
- Cost-benefit and cost-effectiveness analyses
- Needs assessments
- Early implementation reviews
- Logic models and evaluability assessments
- Evaluation synthesis and meta-analysis
- Systems analysis

Resources

Evaluation should be supported through stable, continuous funding sources and through special one-time funds for evaluation projects of interest to Executive Branch and congressional policy makers. The stable, continuous evaluation funds should be provided through appropriations or program fund set-asides. These methods can also be combined to support viable evaluation programs. Agencies should authorize and require periodic evaluations of each program throughout its life to provide rich evaluative information to policy makers during annual appropriation and cyclical reauthorization and amendment discussions. Under situations of severe budgetary constraints, where limited resources are available for evaluation, agency funding should go first to the highest-priority needs, with careful sequencing and prioritization of ongoing work in terms of program and agency evaluation needs as well as the potential of evaluations to provide information and insights that can guide action.

Professional Competence

Evaluations should be performed by professionals with appropriate training and experience for the evaluation activity (such as developing the study's design, performing a statistical analysis, critiquing evaluation results, and writing actionable reports). Evaluation is an interdisciplinary field that encompasses many areas of expertise. Many evaluators have advanced degrees in economics, political science, applied social research, sociology, anthropology, psychology, policy analysis, statistics, and operations research and often work collaboratively with colleagues in allied fields. Federal agencies should ensure that the required diversity of disciplines, including necessary expertise in the subject area being evaluated, is appropriately represented in internal and independent evaluation teams.

Evaluation Plans

Each federal agency should require its major program components to prepare annual and multiyear evaluation plans and to update these plans annually. The planning should take into account the need for evaluation results to inform program budgeting; reauthorization; agency

strategic plans; ongoing program development and management; and responses to critical issues concerning program effectiveness, efficiency, and waste. These plans should include an appropriate mix of short- and long-term studies to produce results of appropriate scope and rigor for short- or long-term policy or management decisions. To the extent practical, the plans should be developed in consultation with diverse program stakeholders.

Evaluation questions can spring up unexpectedly and urgently in response, for example, to a sudden need for information to address a presidential initiative, a management problem, or questions raised by Congress. Therefore, evaluation plans should allow for flexibility in scheduling evaluations.

Dissemination of Evaluation Results

The results of all evaluations related to public accountability should be made available publicly and in a timely manner (except where this is inconsistent with the Freedom of Information Act, Privacy Act, national security, or other relevant laws or policies). These results should be easily accessible and usable through the Internet. Similarly, evaluations of promising and effective program practices should be systematically and broadly disseminated to potential beneficiaries and to potential evaluation users in federal agencies. Evaluation data and methods should also—to the extent feasible and with sufficient privacy protections—be made available to professionals and the public to enable secondary analysis and assure transparency.

Evaluation Follow-Up and Tracking

To refine and improve the relevance of evaluation findings to policy- and program-based questions, agencies should track and re-examine the use made of those findings over time. For example, were the findings used by policy makers or program managers? If so, how? If not, why not? Did the program and its evaluation contribute over the longer term to resolving the public problem addressed? What could or should have been done differently? Again, the findings of such tracking should be shared with professionals and the public to the extent possible.

Evaluation Policy and Procedures

Each federal agency and its evaluation centers or coordinators (discussed below) should publish policies and procedures and adopt quality standards to guide evaluations within its purview. Such policies and procedures should identify the kinds of evaluations to be performed and the criteria and administrative steps for developing evaluation plans and setting priorities, selecting evaluation approaches and methods to use, consulting subject matter experts, ensuring evaluation product quality, publishing evaluation reports, ensuring independence of the evaluation function, using an appropriate mix of staff and outside consultants and contractors, focusing evaluation designs and contracts appropriately, and promoting the professional development of evaluation staff.

Independence

Consultation with stakeholders is generally desirable on such matters as the questions to be addressed in an evaluation and the schedule for study completion. In addition, the leadership of government agencies and their component organizations can, and in many cases should, play a role in establishing general evaluation agendas, budgets, schedules, and priorities. At the same time, the real and perceived independence of the evaluation office must be preserved. Especially for evaluations performed in the service of public accountability, the evaluation organization must retain control of the evaluation's questions, design and methods, and most importantly, its

results and distribution. Without such independence, the credibility of evaluations is at risk. However, the degree of shared decision making, and in some cases stakeholder participation in the conduct of an evaluation, may be greater for evaluations aimed at the improvement of ongoing programs and certain kinds of capacity building.

Institutionalizing Evaluation

Significant progress has been made in establishing evaluation as an integral component of government program management. However, additional steps are needed.

Background

Some federal agencies, such as the Department of Defense, established evaluation offices in the late 1950s and early 1960s. Others, including the Departments of Education and of Health and Human Services, developed their evaluation functions in the 1970s within the then-Department of Health, Education, and Welfare. The authorizing statutes for some of these agencies set aside a fixed percentage of program funds for evaluation. Other departments have added evaluation offices to their organizations, although these offices have grown and shrunk over the intervening years.

One relatively stable evaluation organization has been the Government Accountability Office (GAO), previously known as the General Accounting Office. It has remained the largest single government agency producing evaluations at Congress's request.

One of the most enduring evaluation-related functions has been the Government Performance and Results Act (GPRA) of 1993, recently strengthened by the GPRA Modernization Act of 2010. This law encourages each agency to develop an agency-wide strategic plan and mission statement and also requires them to determine whether their programs achieve their goals and objectives. GPRA defines evaluation as assessing the "manner and extent to which" agencies achieve their goals, thus addressing both implementation and results. In practice, government agencies have implemented GPRA by using performance indicators and measurement to determine whether they have reached a goal and have conducted relatively few evaluation studies that might shed light on how programs reached their goals, why programs do or do not meet their goals and objectives, and how to improve programs. As a result, while the GPRA process does track performance, giving managers an "early warning system" about what is happening in their programs, it generally cannot explain *why* any changes are happening, and thus, it often fails to produce the information needed for guiding programmatic or policy action.

We propose that government agencies, policy makers, and program managers build on the progress already made by embracing evaluation as an integral feature of good government. Agencies should consistently use program evaluation and systematic analysis to improve program design, implementation, and effectiveness and to assess what works, what does not work, and why. This comprehensive vision recognizes that evaluation is more than simply "looking in the rearview mirror" and needs to be used throughout a program's life as an integral part of managing government programs at all stages.

For this approach to work, the Executive Branch and Congress will need to take action, as described below.

Executive Branch Role

As noted earlier, the infrastructure and practice of program evaluation in federal agencies is something of a mixed story. Some agencies have well-developed and stable evaluation offices; others do not. The same can be said for evaluation funding, scope, policies, planning, and dissemination.

Different federal agencies and programs have different evaluation needs, and the maturity and breadth of their evaluation programs vary. In addition, the evaluation function might be a component of other offices focused on such functions as management, planning, research, and policy development, including legislative or regulatory development. For example, several agencies have offices of planning and evaluation, research and evaluation, or monitoring and evaluation, and some inspectors general have offices of inspections and evaluations. In some agencies, the evaluation function is highly centralized or covers a large program area; in other agencies, the evaluation function is scattered across small offices throughout the agency.

No single best practice exists for organizing evaluation offices and functions. All the arrangements described above have emerged in response to such factors as substantive area, kind of agency, or type of evaluation focus. They may or may not be the most effective models for current circumstances. Whatever model is chosen, the evaluation office must include the functions and possess the attributes described above under general principles.

Based on the general principles discussed in the previous section, we propose that agencies in the Executive Branch establish one of the following organizational frameworks to support evaluation.

Option 1: Evaluation Centers. Agencies could establish one or more evaluation centers to promote evaluation capacity and provide stable organizational frameworks for planning, conducting evaluations, or procuring evaluation advice or studies from outside organizations. Every program in the agency should be assigned to one of the centers for program evaluation. The heads of these evaluation centers should report directly to the senior executive of their center's organizational component. Each of these centers would have the following responsibilities:

- Have a stable budget with sufficient funds to plan and carry out both evaluations and the needed evaluation capacity building over several years
- Issue policies and procedures to guide its evaluation work, including guidance on appropriate methods for conducting formative and summative evaluations, as well as developmental evaluations to improve evaluative capabilities within agencies
- Strategically plan a body of evaluation work for the agency and each agency component for which it has evaluation responsibility
- Consult with agency program and budget offices and, in concert with the agency's legislative liaison office, with Congress in developing evaluation plans
- Hire professional evaluators or engage consultants or contractors with the diverse skills necessary to plan and execute (or procure) independent evaluation studies
- Publish the results of evaluations related to public accountability of the programs within their jurisdictions
- Share information about effective programs and evaluation methods with other

government agencies

- Promote and facilitate the ongoing training and professional development of the center's evaluators

Option 2: Evaluation Coordinators. Agencies that choose to distribute their evaluation offices, associating small evaluation offices with individual programs or small groups of programs, should promote evaluation capacity and performance by appointing one or more senior officials to accomplish the following:

- Advise the agency head or senior officials on matters pertaining to evaluation
- Ensure that each program or program group in the agency has a current annual evaluation plan
- Promote, facilitate, and coordinate the development of evaluation plans for programmatic issues that cut across agency lines
- Facilitate the preparation of evaluation budgets
- Establish appropriate standards, frameworks, and procedures for evaluation activities in the agency
- Facilitate the development and efficient and effective production of evaluation plans, designs, instruments, and reports by government agency staff or outside evaluators
- Facilitate the dissemination of evaluation reports related to public accountability
- Share with other agency groups information about effective programs and evaluation methods
- Promote and facilitate the ongoing training and professional development of evaluators in the agency

Option 3: Combined Approach. Federal agencies may find it advantageous to use Option 1 and Option 2—evaluation centers for large programs, program groupings, and overall evaluation support, and evaluation coordinators for distributed evaluation offices—to ensure the viability of the evaluation function.

Congress's Role

The GAO, Congressional Research Service, Congressional Budget Office, National Academies, and temporary commissions carry out evaluation and analysis for Congress, usually in fulfillment of their oversight role. Congressional committees or subcommittees also conduct some evaluative studies or investigations.

We do not propose to change these organizational structures. Instead, we offer recommendations to strengthen the connection between evaluations and the laws that Congress passes. This can be done by building evaluation expectations into authorizing legislation and explicitly setting aside adequate resources for evaluation.

Authorizing legislation. Program authorization and periodic reauthorization provide opportunities for Congress to establish frameworks for systematic evaluation of new and continuing programs. Congressional committees can, through authorizing legislation, provide guidance on or mandate such activities and products as:

- Early implementation reviews to identify start-up problems in such areas as scheduling, contracting, and grant making and to correct them before they become more serious

- Requirements for developing evaluation plans
- Evaluation of promising approaches to share among program implementers
- Development of performance indicators and the means to collect meaningful data on them once the program starts and, ideally, before
- Studies reviewing the efficiency of federal program management as well as the fidelity of program implementation to the congressional mandate that instituted the program
- Studies assessing program effects and identifying why programs are or are not effective
- Evaluations of topics of interest to Congress and reports on the results to Congress in support of its oversight and appropriations functions and to inform future reauthorizations
- Establishment, expansion, or amendment of ongoing surveys or other data-collection mechanisms to become permanent sources of reliable data
- Establishment of evaluation centers or evaluation coordinators, as described above
- Funding for evaluation activities, including strengthening evaluation capacity, where needed (in terms of personnel, professional development, etc.)

Collaboration Between the Executive Branch and Congress

The utility of evaluation results may be maximized if Congress and the Executive Branch jointly specify broad evaluation expectations and concerns in authorizing statutes and appropriations.

We recognize that such collaboration will not always be easy or even possible to achieve.

Nevertheless, experience suggests that, when possible, a partnership of this kind can help increase the benefits that evaluation provides.

Looking to the Future

The U.S. government faces major challenges in the years to come as well as significant opportunities to improve lives, protect the planet, and create efficiencies. With adequate resources; more thoughtful and systematic integration of evaluation into the planning, management, and oversight of programs; and the application of evaluation results to planning and decision making, we can both improve the performance of today's programs and ensure that tomorrow's programs reap the lessons learned today. Institutionalizing evaluation can also help achieve a more accountable, open, and democratic system of governance for future generations.