

Foundations for Evidence-Based Policymaking Act of 2018

Social Security Administration Evaluation Policy



Securing today
and tomorrow



CONTENTS

Introduction	3
Evaluation Standards	4
Relevance and Utility	4
Rigor.....	4
Independence and Objectivity	5
Transparency	6
Ethics.....	7
Appendix A: Evaluation Definitions	8
Impact Evaluation	8
Outcomes Evaluation.....	8
Process Evaluation	8
Formative Evaluation	8
Descriptive Studies	8
Appendix B: OMB Guidance on Evidence Act Officials.....	9
Chief Data Officer	9
Evaluation Officer.....	9
Statistical Official.....	9

Selected Abbreviations

OMB	Office of Management and Budget
SSA	Social Security Administration

INTRODUCTION

This document sets forth our policy on planning, implementing, and using evaluations. Evaluation, for the purposes of this policy, refers to an assessment using systematic data collection and analysis of one or more programs, policies, and organizations intended to assess their effectiveness and efficiency. There are many types of evaluations that meet this definition, including formative, process, outcome, and impact evaluations. The definition of evaluation covers a variety of different types of analysis, such as quantitative analyses, qualitative analyses, and mixed methods. The definition includes a variety of purposes, for example, assessing operational needs, determining if policy is being implemented correctly, or determining if new programs or policies work and should be implemented.

All of our offices should follow this policy for all evaluations. By following a consistent policy, we demonstrate our commitment to conduct rigorous evaluations and to use evidence to inform our policies and practices. This document highlights five specific evaluation standards:

1. Relevance and Utility
2. Rigor
3. Independence and Objectivity
4. Transparency, and
5. Ethics.

These standards have a strong evidence base and are used throughout the Federal Government. These standards allow for office- and evaluation-specific needs and are not intended to prescribe any specific evaluation methods. Additionally, the overarching principle of Equity is woven into each of these standards and supports Executive Orders [13985](#) and [14035](#) and professional evaluation [principles](#). We use the definition of Equity established in Executive Order 13985, which is “the consistent and systematic fair, just, and impartial treatment of all individuals, including individuals who belong to underserved communities that have been denied such treatment, such as Black, Latino, and Indigenous and Native American persons, Asian Americans and Pacific Islanders and other persons of color; members of religious minorities; lesbian, gay, bisexual, transgender, and queer (LGBTQ+) persons; persons with disabilities; persons who live in rural areas; and persons otherwise adversely affected by persistent poverty or inequality.”

Offices that are unfamiliar with evaluation can consult with our [Evaluation Officer](#) for more information on evaluation in general and for assistance identifying ways to plan and implement evaluations. Additional detailed information on the above evaluation standards is in Office of Management and Budget circular [M-20-12](#).

EVALUATION STANDARDS

Relevance and Utility

As a steward of government funds and in support of our service to the public, our evaluations must address relevant questions, serve specific agency interests, and address program or policy priorities, as identified in the Agency Strategic Plan, Budget, or other documents.

Our offices should plan evaluations to provide timely input into the decision-making process—be it regulatory, operational, managerial, or other decisions. Our offices should work to integrate data collection and evaluation planning as part of their project or program design. To ensure the relevance and utility of evaluations, the data necessary for evaluation purposes must be included as part of the agency's information technology projects or other data collection activities. While evaluations are sometimes possible after a policy or program has been in place, such evaluations are not always possible and may be subject to limitations that affect the relevance and utility of the findings.

We have multiple stakeholders, including beneficiaries and claimants, legislators, the administration, internal offices, and others; it is important for us to consider multiple perspectives when developing and implementing an evaluation. We should design our evaluations to address the diverse needs of our stakeholders. Evaluators are encouraged to include a diversity of stakeholders on evaluation teams, including individuals with lived experiences related to the topic studied. Such inclusion helps ensure the questions being asked are meaningful, and it helps identify erroneous assumptions or mischaracterizations of stakeholder experiences.

Additionally, our evaluations should be designed in such a way that there are practical actions that can be taken with the results. Different stakeholders consume information in different ways; evaluations must be tailored to the needs of the user.

Evaluators should solicit stakeholder input early and often throughout the evaluation to ensure both buy-in of the questions and ultimate utility of the evaluation.

Rigor

Rigor refers to an evaluation's ability to provide appropriate answers to the questions it is designed to answer, including the quality, completeness, and credibility of the evaluation. We are committed to conducting the most rigorous evaluations possible, subject to budgetary, legal, and other constraints, so that stakeholders can rely on findings produced by an evaluation.

The first opportunity to build rigor into an evaluation lies in developing a clear question(s) of interest; the evaluation design and methods are then chosen to answer the question that has been asked. To ensure equity in evaluation design, it is essential to engage in a diverse set of stakeholders inside and outside the agency, including customers and potential customers. Regardless of the evaluation type, it is important for all of our evaluations to have clear definitions of processes, policies, outcomes, data, and methods. No evaluation is without limitations, and evaluators must clearly identify these to the extent feasible.

To conduct rigorous evaluations, evaluators must be both technically and culturally competent, drawing on the full breadth of evaluation methods and applying them in a manner that appropriately matches the social context. While experimental methods are often preferred for impact evaluations that seek to answer questions of causality, other methods are appropriate for other types of questions. For

example, quasi-experimental methods offer an appropriate alternative for evaluations of previously implemented policies and programs, or when experimental methods are impractical. Qualitative methods are useful if the data available are not numeric or to understand the “why” or “how” that underlies the “what/how many/how much” of quantitative evaluation questions and methods. Qualitative methods are also particularly helpful uncovering inequities and other barriers that can be lost in quantitative analyses of administrative data or surveys. Ultimately, the questions asked determine the type of evaluation needed, and may require mixed-methods evaluations.

Evaluations, in general, must be on a sufficient scale to be meaningful; but we encourage small or pilot studies or other preliminary projects before launching large-scale evaluations. Such pilots allow sufficient time to synthesize the findings of formative or early assessments and make changes to the program or the evaluation, allowing the large-scale endeavor to be as close as possible to an implementable solution.

High-quality data are the foundation for rigorous evaluations. We will ensure that the high-quality data necessary to conduct evaluations will be created, collected, accessible, and maintained for all programs and policies, and related information technology projects. Data and privacy laws, as well as legal requirements, sometimes constrain our use of data both internally and externally. We will consult our Office of General Counsel, as needed, when we plan and implement evaluations to ensure that we are following the law.

We are committed to hiring and developing staff with the specialized knowledge and skills needed to conduct rigorous evaluations. While most of our offices evaluate policies and programs to some extent, we have several offices with significant expertise in evaluation. We are dedicated to expanding the evaluation capacity within these offices, which include, but are not limited to, the Office of Analytics, Review, and Oversight; the Office of the Chief Actuary; the Office of Retirement and Disability Policy; and the Office of Systems. Fully implementing the Evidence Act will require investment in staff, data, and tools to be prepared to provide evidence when needed. These offices can support, subject to resource availability, other components’ evaluation needs. The Evaluation Officer can also help components identify additional evaluation resources and options.

Independence and Objectivity

We are committed to evaluations free from actual or perceived biases and other influences that could bring into question the legitimacy of the results. Appointed leaders and other stakeholders participate in setting the evaluation agenda, identifying relevant questions, and determining how and whether we will ultimately use evaluation findings. We engage in activities, including both internal and external reviews of evaluation plans and findings, designed to promote the independence and objectivity of our research and evaluations.

Evaluations are strengthened when multiple reviewers from diverse perspectives provide insight in the planning, execution, and reporting stages of the evaluation. While evaluations should be protected from political interference, we expect evaluators to seek out and accept technical corrections from a wide range of stakeholders, including technical policy and operational experts, people with lived experiences, and others. However, post hoc changes to manipulate the design, methods, data, or results and findings or to ensure a certain use is not acceptable. The Evaluation Officer is ultimately accountable for ensuring the independence and objectivity of evaluations, and is responsible for arbitrating disputes about evaluation technical factors.

Our offices are generally able to conduct evaluations of their own programs, policies, and organization,

but we often use external evaluators to gain additional technical expertise, to provide a level of independence, and to identify weaknesses in practice or policy. If an office intends to use an external evaluator, our Office of Acquisitions and Grants has policies that we must observe to ensure no inappropriate influences, or perceptions of partiality, affect the selection of external contractors or awarding grants. We encourage offices to seek guidance from our Office of General Counsel on other legal issues related to external contracts.

Whether internal or external to the agency, evaluators must take care to only draw conclusions that are supported by the findings of the study. It is rare that a single evaluation will assess all relevant concepts and contexts of an issue. Evaluators should integrate the findings of any single evaluation with other evidence to provide an unbiased assessment of how any one specific evaluation fits into the broader evidence base and policy and social context.

Transparency

We will make evaluations available, as appropriate, primarily by posting evaluation plans and findings on our website, to ensure that others may learn from our work. Prior to initiating an evaluation, we will consult with all relevant stakeholders to determine whether to make an evaluation public or otherwise limit the release of any part of the evaluation activities. The outcomes of the evaluation—favorable, unfavorable, or null—do not determine whether we will release it. In addition to making evaluation plans and findings publicly accessible, evaluators should proactively seek ways to share the findings with a diverse group of stakeholders, including internal staff, advocates, subject communities, and others.

In addition to reporting comprehensive findings, the agency should make publicly available the data, methods, and statistical programs used in evaluation, subject to any legal, security, or other constraints. We will only disclose evaluation data consistent with applicable laws, regulations, and policies to ensure the proper protection of interests such as the security, privacy, and integrity of the data and participants. If we determine that we cannot release certain content or details of an evaluation, we will include sufficient detail in published reports so that a knowledgeable consumer can replicate or understand the evaluation design, methods, and data. Statistical programs and data should be archived in a way that supports internal or external, as appropriate, review and replication, and should be maintained in accordance with Federal privacy and records laws. Evaluation contracts should include, when possible, public use data file deliverables and other requirements to support replicability of the evaluation.

While we implement many evaluations as planned, issues such as unexpected data limitations or changes in other policies, may affect the evaluation. If necessary, evaluation reports should identify and discuss changes to the implementation or the evaluation design. Additionally, we will make stakeholders aware as early as possible if issues arise or of potential changes in evaluation scope. Evaluators should not expand the scope of an evaluation for the sole purpose of seeking a specific result. If we decide to terminate an evaluation before achieving pre-specified milestones, we will provide notice to all relevant stakeholders. For instances where we published a public notice announcing the start of an evaluation, we will publish a notice of the reasons for the decision to terminate.

Evaluators should share findings in accessible formats and consider equity in all components of the research and evaluation process, including the dissemination of findings. While we must comply with federal requirements (e.g., Section 508), we should consider additional formats that will meet the needs of our diverse stakeholders.

Ethics

We will conduct our evaluations in an ethical manner and will safeguard the dignity, rights, safety, and privacy of participants, stakeholders, and affected entities. Evaluators must be aware of and mitigate existing inequities, structural barriers, and historic disadvantages to participation in an evaluation by various groups. Additionally, evaluators should consider equity when identifying topics evaluations address, focusing not just on the agency's needs, but also the overall well-being of the customer and society. We will conduct evaluations in accordance with our regulations on protection of human subjects' research, which adopt the Common Rule. Additionally, we acknowledge that different cultures have different expectations about privacy, participation, dignity, and equity. Our evaluators, both internal and external, must respect any person participating in an evaluation and comply with any relevant Federal Government or other requirements. Communicating appropriately with stakeholders is important to ensure evaluations remain relevant, useful, and transparent, and to ensure we do not engage in unintentionally burdensome, intrusive, disrespectful, or otherwise inappropriate evaluations or actions.

APPENDIX A:

EVALUATION DEFINITIONS¹

Impact Evaluation

This type of evaluation assesses the causal impact of a program, policy, or organization, or aspect of them on outcomes, relative to a counterfactual. In other words, this type of evaluation estimates and compares outcomes with and without the program, policy, or organization, or aspect thereof. Impact evaluations include both experimental (i.e., randomized controlled trials) and quasi-experimental designs. An impact evaluation can help answer the question, “does it work?” or “did the intervention lead to the observed outcomes?”

Outcomes Evaluation

This type of evaluation measures the extent to which a program, policy, or organization has achieved its intended outcome(s), and focuses on outputs and outcomes to assess effectiveness. Unlike impact evaluation above, it cannot discern causal attribution but is complementary to performance measurement. An outcomes evaluation can help answer the question, “were the intended outcomes of the program, policy, or organization achieved?”

Process Evaluation

This type of evaluation assesses how the program or service is delivered relative to its intended theory of change, and often includes information on content, quantity, quality, and structure of services provided. These evaluations can help answer the question, “was the program, policy, or organization implemented as intended?” or “how is the program, policy, or organization operating in practice?”

Formative Evaluation

This type of evaluation is typically conducted to assess whether a program, policy, or organizational approach—or some aspect of these—is feasible, appropriate, and acceptable before it is fully implemented. It may include process and/or outcome measures. However, unlike summative evaluation designs like outcome and impact evaluations, which seek to answer whether or not the program, policy, or organization met its intended goals or had the intended impacts, a formative evaluation focuses solely on learning and improvement and does not answer questions of overall effectiveness.

Descriptive Studies

These studies can be quantitative or qualitative in nature, and seek to describe a program, policy, organization, or population without inferring causality or measuring effectiveness. While not all descriptive studies are evaluations, some may be used for various evaluation purposes, such as to understand relationships between program activities and participant outcomes, measure relationships between policies and particular outcomes, describe program participants or components, and identify trends or patterns in data.

¹ See Office of Management and Budget circular [M-20-12](#) for additional evaluation-related definitions.

APPENDIX B:

OMB GUIDANCE ON EVIDENCE ACT OFFICIALS²

The Evidence Act requires agencies to designate three senior officials—Chief Data Officer, Evaluation Officer, and Statistical Official—who will oversee the use of data and evidence-building activities in agencies. We identify each of the three designated officials on our website at www.ssa.gov/data.

Chief Data Officer

The Chief Data Officer (CDO) shall have authority and responsibility for, among other things, data governance and lifecycle data management. While there are many roles in the Federal Government that relate to data management, over the last few years, Chief Data Officers (CDOs) have emerged to lead organizational development of processes to leverage the power of data. CDOs enable data-driven decision-making in a variety of ways, from providing and leveraging centralized agency analytics capacity to creating tools and platforms that enable self-service across their agencies and for the public. Successful data management must account for every stage of the data lifecycle. Among other things, it involves establishing effective procedures, standards, and controls to ensure quality, accuracy, access, and protection of data, as well as managing information technology and information security.

Evaluation Officer

The Evaluation Officer shall have authority and responsibility for providing leadership over the agency's evaluation and learning agenda activities. The Evaluation Officer shall be responsible for overseeing the agency's evaluation activities, learning agenda, and capacity assessment, as well as collaborating with, shaping, and contributing to other evidence-building functions within the agency. The Evaluation Officer is responsible for providing technical and methodological leadership to assess, improve, and advise on evaluation activities across the agency. For agencies that are less mature in their evaluation activities, or for those agencies without additional evaluation expertise distributed throughout the agency, the Evaluation Officer may also be responsible for conceptualizing, prioritizing, and designing the agency's evaluation activities.

Statistical Official

The Statistical Official shall have authority and responsibility to advise on statistical policy, techniques, and procedures. Statistical expertise allows organizations to ensure that data are gathered, processed, and curated so as to produce statistical products with the highest standards of data quality while protecting confidentiality, privacy, and security. Data quality has multiple dimensions, including credibility and accuracy, timeliness and relevance to valuable decision-making processes, the objectivity with which data are produced, and their accessibility to multiple users at an appropriate level of clarity and detail. Applying statistical expertise involves the maintenance and development of policies that anticipate and address the needs of data users and providers, the continual advancement and adoption of statistical techniques to maximize the quality of statistical outputs, and the delineation of procedures to ensure that we implement these policies and techniques in a rigorous and efficient manner.

² See Office of Management and Budget circular [M-19-23](#) for additional information.



Securing today
and tomorrow



[SSA.gov](https://www.ssa.gov)



Social Security Administration

April 2024

Foundations for Evidence-Based Policymaking Act of 2018 Documentation

Produced and published at U.S. taxpayer expense