



DOJ Policy Statement

EVALUATION POLICY OF THE DEPARTMENT OF JUSTICE

PURPOSE: Establishes key principles that guide the Department of Justice’s (DOJ or Department) use of evaluation

SCOPE: All DOJ components

ORIGINATOR: Strategic Policy and Performance Staff

CATEGORY: (I) Administrative (II) 400 Policy, Directives, & Guidance

AUTHORITY: 5 U.S.C. §§ 311, 313; Office of Management and Budget (OMB) Memorandum M-19-23, Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance; OMB Memorandum M-20-12, Phase 4 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Program Evaluation Standards and Practices; OMB Memorandum M-21-27, Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance

CANCELLATION: None

DISTRIBUTION: Electronically distributed to those referenced in the “SCOPE” section and posted on the DOJ directives electronic repository (SharePoint) at <https://doj365.sharepoint.us/sites/jmd-dm/dm/SitePages/Home.aspx>

APPROVED BY: Christopher Alvarez *Christopher Alvarez* CHRISTOPHER
Evaluation Officer ALVAREZ
2025.01.06 16:54:43
-05'00'

ACTION LOG

Issuing component must review its DOJ directives, at minimum, every 5 years and make revisions as necessary. The action log records dates of approval, recertification, cancellation, and revisions to this directive and provides a summary. If this directive is cancelled, superseded, or supersedes another directive, annotate the action log.

Action	Authorized by	Date	
Initial Document	Christopher Alvarez, Evaluation Officer	Date of digital signature	Establishes an evaluation policy for the Department of Justice, as required by the Foundations for Evidence-Based Policymaking Act of 2018

TABLE OF CONTENTS

ACTION LOG	2
DEFINITIONS	4
ACRONYMS	6
I. Introduction.....	7
II. Policy	7
A. Rigor	7
B. Relevance and Utility.....	8
C. Independence and Objectivity.....	9
D. Transparency.....	9
E. Ethics.....	10
III. Roles and Responsibilities	10

DEFINITIONS

Term	Definition
Evaluation	“[A]n assessment using systematic data collection and analysis of one or more programs, policies, and organizations intended to assess their effectiveness and efficiency.” 5 U.S.C. § 311(3). Types of evaluation include formative evaluation, impact evaluation, outcome evaluation, and process/implementation evaluation.
Evaluators	Those who plan, design, conduct, manage, and/or report on evaluations. In general, these will be federal staff and associated partners who are trained—through advanced education and evaluation experience (e.g., quantitative, qualitative and/or mixed-method evaluation specializations) —to properly plan, implement, manage, and/or oversee evaluation activities and evaluations.
Findings	The principal outcomes of an evaluation; what the evaluation suggested, revealed or indicated. Findings are results, conclusions, and recommendations that are systematically generated through analyzing and interpreting data.
Formative evaluation	An evaluation 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 outcome and impact evaluations, which seek to answer whether the program, policy, or organization met its intended goals or had the intended impacts, a formative evaluation focuses on learning and improvement and does not aim to answer questions of overall effectiveness.
Impact evaluation	An evaluation that assesses the causal impact of a program, policy, or organization, or aspect thereof, on outcomes relative to those of 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?"

Term	Definition
Outcome evaluation	An evaluation that 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 typically cannot discern causal attribution. An outcome evaluation can help answer the question "were the intended outcomes of the program, policy, or organization achieved?"
Personally identifiable information (PII)	Information that can be used to distinguish or trace an individual's identity, either alone or when combined with other information that is linked or linkable to a specific individual
Pre-specification/pre-registration	The practice of documenting and committing to a specific evaluation design and analysis plan before collecting and analyzing data. Sometimes this includes a full study design and a comprehensive plan for analyzing data, and sometimes it focuses on certain key elements such as research questions and outcomes to be analyzed. The basic purpose is to commit to (1) using specific methods and (2) reporting the results that come from those methods.
Process/implementation evaluation	An evaluation that 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?"
Program	A set of projects or activities that support a higher-level objective or goal. For the purpose of this policy, program includes processes, projects, interventions, policies, operations, activities, entities, and functions.
Quasi-experimental design (QED)	An impact evaluation design that attempts to isolate and measure the causal impact of an intervention by constructing comparable groups and using statistical adjustments to compensate for any initial lack of equivalence between the groups.
Randomization	A research method involving the random assignment of individuals or other units to experimental conditions in order to measure the causal impact of one or more treatments or interventions. This technique is used in randomized evaluations, also known as experimental evaluations and randomized controlled trials (RCTs).

Term	Definition
Research participants	Living individuals about whom information is obtained within the scope of an evaluation, or any other systematic investigation, that is intended to produce generalizable (externally valid) findings.

ACRONYMS

Acronym	Meaning
DOJ	Department of Justice
PII	Personally identifiable information

I. Introduction

The Department of Justice (DOJ or Department) recognizes evaluation as an important tool for assessing and improving the effectiveness of its programs and policies. This directive affirms and reflects the Department's commitment to conducting rigorous and relevant evaluations and to using evidence from those evaluations to inform decisions about programs and policies. It identifies key principles that guide the Department's use of evaluation and provide a foundation for a culture of evidence-informed, data-driven decision making.

For purposes of this directive, evaluation is defined as in the Foundations for Evidence Based-Policymaking Act of 2018 (Evidence Act): "An assessment using systematic data collection and analysis of one or more programs, policies, and organizations intended to assess their effectiveness and efficiency" (5 U.S.C. §311(3)). This directive applies to all forms of evaluation conducted, sponsored, or funded by the Department, including impact evaluations, process/implementation evaluations, and formative evaluations.

II. Policy

In planning, conducting, and reporting on evaluations, the Department adheres to the following principles: rigor, relevance and utility, independence and objectivity, transparency, and ethics. These principles have been recognized as essential to high-quality evaluation in the Federal Government.

This directive provides a framework to guide the Department's components in conducting high-quality, credible, reliable evaluations while also allowing appropriate flexibility for components to identify and implement specific evaluation practices consistent with their capacities and needs.

A. Rigor

Evaluations should produce accurate findings that the Department and its stakeholders can confidently rely upon to inform decisions about programs and policies. Further, evaluation findings should be accompanied by clear information about methods and limitations.

The quality of an evaluation depends heavily on the underlying design and methods. When designing and conducting an evaluation, the Department will use methods that address the evaluation's objectives as rigorously as possible within statutory, budgetary, and other constraints and in accordance with ethical principles. When evaluating the causal impact of a program or policy, the Department will use methods that, to the greatest extent possible, isolate the impact of that program or policy from other influences such as contextual factors, preexisting trends, or preexisting demographic or geographic differences. For this purpose, the most rigorous evaluation design may include random assignment of individuals (or other appropriate units of analysis) to experimental and control conditions. In cases where randomization is not feasible, evaluators should use a strong quasi-experimental design that can address the risk of selection bias.

Unless results will be reported as exploratory and treated as tentative for decision-making purposes, an evaluation's design and methods should be pre-specified before data collection begins. Pre-specification reduces the risk that methods are tailored to achieve specific findings or that results are reported selectively; it thus enhances the rigor, transparency, and credibility of an evaluation. There are multiple ways of pre-specifying an evaluation's design and methods, involving different levels of transparency and accountability. For a given evaluation, a method should be chosen that affords the greatest transparency and accountability within legal, ethical, national security, law enforcement, or other constraints on disclosing information. If pre-specification is not feasible, then the reason should be clearly stated in project reports. If an evaluation's pre-specified design and methods are later revised, then project reports should include an explanation of the revisions. For evaluations conducted by external partners under grants, contracts, or other agreements, these requirements concerning pre-specification should be included in those agreements.

All evaluations will adhere to widely accepted scientific principles and will make use of measures that accurately address the evaluation's objectives (for example, accurately represent a program's intended outcomes or specific aspects of program implementation). Final reports and presentations of results will use precise language to characterize findings, design, methods, assumptions, and limitations – including, whenever relevant, clear information about the extent to which conclusions about cause and effect are well founded (internal validity) and the extent to which findings can be generalized to other populations, settings, or circumstances (external validity). Those responsible for designing, conducting, and reporting evaluations should have demonstrated expertise in the appropriate evaluation methods and practices.

When evaluations are conducted by external partners on behalf of the Department, appropriate steps will be taken to apply the same standards of rigor in methods that apply to evaluations conducted internally – for example, through carefully assessing the technical merit of proposals and the qualifications of contractors, grant applicants, and other prospective partners.

B. Relevance and Utility

Evaluations must address questions of importance to help the Department effectively and efficiently achieve its mission. The Department will prioritize evaluations that have potential to meaningfully inform agency decisions and activities in areas such as budgeting, program improvement, accountability, management, regulatory action, and policy development. Evaluations should have significant potential to yield results that the Department and its partners and stakeholders at the federal, state, tribal, territorial, and local levels can use to make informed decisions for the benefit of the public. Legislative requirements and congressional interests may also inform evaluation efforts.

In designing evaluations, the Department will incorporate the needs of decision makers and other stakeholders for relevant, actionable results. In planning evaluations, the Department will recognize the critical relationship between utility

and timeliness and the importance of delivering high-quality results in time to be actionable. In disseminating the results of evaluations, the Department will promote their utility by actively sharing findings with stakeholders and ensuring that findings are presented in clear, actionable terms for audiences of different backgrounds. Whenever feasible and appropriate, evaluators should advise agency decision makers on the translation of results into policy or practice, including by advancing specific recommendations and by clarifying any relevant limitations on how or how broadly results should be applied. This is especially important in cases where results might conflict with decision makers' expectations.

C. Independence and Objectivity

Evaluations have the potential to inform important Department decisions about programs and policies. For this reason, evaluations must be viewed as credible for the public and other stakeholders to accept their findings. The credibility of evaluations is dependent on the actual and perceived independence and objectivity of those managing and conducting the evaluations.

While Department and program leadership, program staff, and other stakeholders should participate in setting evaluation priorities and identifying evaluation questions, those responsible for designing, conducting, and reporting evaluations should operate with an appropriate level of independence from programmatic, regulatory, policymaking, and stakeholder influences. Evaluators should strive for objectivity in the planning and conduct of evaluations and in the interpretation and dissemination of findings, avoiding conflicts of interest, bias, and other partiality.

As evaluation is a scientific activity, evaluations should be conducted in accordance with the Department's Scientific Integrity Policy, which discusses adherence to professional practices, ethical behavior, and the principles of honesty and objectivity in scientific matters.

D. Transparency

The Department recognizes the importance of transparency about both planned and completed evaluations, to promote accountability and help ensure that no aspects of an evaluation are tailored to generate specific findings. Transparency in the evaluation, research, and analysis that inform Department decisions engenders public trust in those decisions and in government more broadly.

Whenever possible, information about planned and ongoing evaluations will be made easily accessible and will cover each evaluation's purpose and objectives, design and methods, expected timeline for reporting results, and any contractors, grantees, or other external partners conducting the work. Mechanisms for communicating about planned and ongoing evaluations include the Department's Annual Evaluation Plans and the websites of the Department and its components.

When evaluations are completed, all results should be reported—whether favorable, unfavorable, or neutral—along with sufficient detail on methods and limitations so that others can review, interpret, and learn from the work. Evaluation findings should be shared in concise formats, using plain language and supported by data, so they are useful to decision makers. Findings should be disseminated in contextually appropriate ways to policymakers, program partners, and, when applicable, the public. In the event an evaluation does not adequately address its research objectives, this fact should be reported, along with sufficient detail on challenges encountered so that others can learn how to avoid or address these challenges in future work. The Department will release results of evaluations in a timely manner, whenever such results do not pertain to legal or enforcement procedures and are not otherwise prohibited from disclosure due to national security concerns, law enforcement sensitivity, or privacy or other considerations.

E. Ethics

Federal evaluations must be conducted to the highest ethical standards to protect the public and maintain public trust in the government's efforts. Heads of Components will assist in ensuring that evaluations are conducted in an ethical manner and safeguard the dignity, rights, safety, and privacy of participants and other stakeholders and affected entities.

Evaluations that involve research participants, including participants from vulnerable populations, will be conducted in accordance with applicable laws and rules governing the protection of human subjects.

Evaluations should be equitable, fair, and just. The Department will strive to ensure equity in setting research priorities (which policies or programs are prioritized for evaluation) and in developing research questions and evaluation designs that are free of bias. When evaluations make use of existing data, the Department should consider whether historically underrepresented populations are adequately represented in the data or whether new data collection might be required to ensure representativeness.

To safeguard the privacy of individuals whose data is used for evaluation, the Department will employ data management practices that ensure appropriate collection, processing, storage, and dissemination consistent with applicable laws and regulations governing the use of personally identifiable information (PII). Data elements collected during an evaluation should be appropriately tailored to the evaluation's objectives in order to avoid unnecessary collection of PII. If emerging technologies, including artificial intelligence, are employed in the design, conduct, analysis, or reporting of an evaluation, care will be taken to ensure compliance with all relevant policies and guidance on the safe, ethical use of such technologies.

III. Roles and Responsibilities

The DOJ Evaluation Officer has authority and responsibility for providing leadership over the Department's evaluation activities, including by overseeing the implementation of this directive.

Additionally, the Evaluation Officer will: execute the duties and responsibilities of the Evaluation Officer as described in the Foundations for Evidence-Based Policymaking Act of 2018, OMB Memorandum M-19-23, and OMB Memorandum M-21-27; serve as an agency champion for and educator of agency staff and leaders about evaluation; and serve as a senior advisor to agency leaders on issues of evaluation policy and practice, such as designing and undertaking evaluations, interpreting results, and integrating evaluation findings into day-to-day agency operations, management processes, budgeting, strategic planning, and other decision making.

Heads of Components will ensure compliance with this directive within their components and will consult with the Evaluation Officer, as needed.

All DOJ staff involved in planning, conducting, and reporting on evaluations will adhere to the five principles articulated in this directive.