

EVALUATION FRAMEWORK



**U.S. Department
of Transportation**

Published February 2025

PURPOSE

The purpose of this Evaluation Framework is to define and describe principles to guide the planning and execution of program evaluations at the U.S. Department of Transportation (DOT, the Department), defined as “assessment[s] using systematic data collection and analysis of one or more programs, policies, and organizations intended to assess their effectiveness and efficiency.”¹ This framework emphasizes the Department’s commitment to supporting rigorous, relevant evaluations and to evidence-based strategic and operational decisions that result in continuous improvement across the Department. As the Department’s evaluation efforts mature, DOT will review and update this framework as needed.

SCOPE

This framework will govern all program evaluations, as defined above, conducted by the Department’s staff and its contractors (see *What is an evaluation?* below for details). It does not apply to evaluations of proposals, applications, quotations, or similar submissions that the Department has solicited in conjunction with the award of contracts, grants, or cooperative agreements.

EFFECTIVE DATE

This framework is effective immediately for evaluations that begin after the issuance date of this document. Evaluations in progress on the framework issuance date should review the framework and make reasonable efforts to incorporate the guiding principles.

AUTHORITIES

This framework is established under the Foundations for Evidence-Based Policymaking Act of 2018, [Public Law No. 115-435](#) (Evidence Act), the Office of Management and Budget’s (OMB) [Circular A-11 Section 51.9](#), and OMB Memoranda [M-19-23](#) and [M-20-12](#), and [M-21-27](#).

ROLES AND RESPONSIBILITIES

The Department’s three designated officials for Evidence Act implementation (Evaluation Officer, Chief Data Officer, and Statistical Official) promote the collection and use of evidence throughout the Department. The Evaluation Officer is responsible for administering this framework. The Evidence Act and [OMB Memorandum M-19-23](#) describe the responsibilities for the designated officials.

¹ The term “evaluation” as used throughout this document is synonymous with “program evaluation.” See the Foundations for Evidence-Based Policymaking Act of 2018 for more information.

EVIDENCE AT THE DEPARTMENT OF TRANSPORTATION

What is evidence?

The Evidence Act defines evidence as “information produced as a result of statistical activities conducted for a statistical purpose.”

How is evidence created?

[OMB Memorandum M-19-23](#) describes four types of activities that create evidence: foundational fact-finding, performance measurement, policy analysis, and program evaluation. The Department encourages continuous learning and improvement through the building and use of many types of evidence (including evidence generated in everyday operations) such as descriptive research studies, performance measures, financial and cost data, survey results, and program administrative data.

How is evidence used?

Evidence should inform strategic and operational decisions, resulting in improved mission accomplishment across the Department. The intent is to integrate the use of evidence into all critical decisions about programs, policies, and regulations. Evidence will serve as a feedback loop in the Department’s management and analytical functions, allowing for adjustments in data collection or analysis, as well as potential program refinements.

Investing in and focusing on the management and use of data and evidence across the Department will support decision makers in investing in high value activities, by linking spending to program outputs, delivering on mission, better managing enterprise risks, and promoting stakeholder engagement and transparency.

The Department is committed to ongoing, data-driven reviews of performance and progress to support decision-making. With improved and high-quality data, the Department will be equipped to make informed decisions on requirements vetting and prioritization, allocate resources, and improve the efficiency and effectiveness of programs to achieve mission results.

How can programs support evidence-building?

The Department encourages the use of logic models for all programs. These logic models should link program spending and inputs, activities, and outputs to short-, medium-, and long-term outcomes. New discretionary programs and initiatives should consider data collection needs in their initial design. Existing programs should consider how evaluation requirements can be used to rigorously assess and improve program efficiency, effectiveness, and quality. This includes improving existing data collections and identifying needs for new collections.



DEFINITIONS AND GUIDING PRINCIPLES FOR EVALUATION

What is an evaluation?

The Evidence Act defines evaluation as an “assessment using systematic data collection and analysis of one or more programs, policies, and organizations intended to assess their effectiveness and efficiency.” Rigorous evaluations could establish a causal relationship between an activity and the outcomes, address questions about implementation, or measure variations in effectiveness across different settings or populations. As described in [OMB Memorandum M-20-12](#), different types of evaluations include:

- **Formative Evaluation:** 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 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.
- **Process or Implementation 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 questions such as, “Was the program, policy, or organization implemented as intended?” or “How is the program, policy, or organization operating in practice?”
- **Outcome 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, it typically cannot discern causal attribution. Importantly, it is distinct from, but complementary to, performance measurement. An outcome evaluation can help answer the question, “Were the intended outcomes of the program, policy, or organization achieved?”
- **Impact Evaluation:** Assesses the causal impact of a program, policy, or organization, or aspect thereof, on outcomes relative to those of a counter-factual. 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?”

What are the Department's principles for evaluations?

In accordance with OMB Memoranda [M-20-12](#) and [M-21-27](#), the Department establishes the following guiding principles for evaluations:

RIGOR

All evaluations will use methods that generate the highest quality and most credible evidence that corresponds to the questions asked, within our statutory authority and the limits of time, budget, and other practical considerations. This application of rigor results in a robust and unbiased evaluation design, sound methodology, and analysis, interpretation, and reporting of results. All evaluations, regardless of method (e.g., qualitative, quantitative, or mixed) or type (e.g., impact, outcome, process) must adhere to widely accepted scientific principles and employ methods most appropriate for the evaluation’s objectives, consistent with the availability of resources and mission priorities. The Department’s evaluations will:

- Ensure that any inferences about cause and effect are well-founded (internal validity);
- Draw appropriate conclusions about the extent to which results can be generalized outside the Department’s context (external validity);
- Use measures that accurately (measurement validity) and consistently (measurement reliability) capture the intended information;
- Identify risks and challenges and develop mitigation strategies in all phases of the evaluation process; and
- Continue building staff capacity, efficient processes, and frameworks necessary to conduct high-quality evaluations.

These efforts described above align to the Department’s evidence-building plan to develop:

- Evaluation plans that support mission needs and consider cost and time to maximize the use of available resources;
- An evaluation workforce with appropriate expertise and experience (includes providing ongoing training and professional development opportunities that promote cutting-edge evaluation methods and tools);
- Partnership, funding, and procurement vehicles that support high-quality, efficient, and effective evaluations;
- A regulatory and statistical review infrastructure for all OMB Office of Information and Regulatory Affairs Information Collection Requests that are routinely required in order to conduct high-quality evaluations; and
- A data and quality assurance infrastructure for adeptly collecting, analyzing, and reporting necessary data.

RELEVANCE AND UTILITY

Evaluation designs will consider legislative requirements, leadership priorities, and the needs of the Department’s stakeholders, including public- and private-sector customers, partners, and vendors. Evaluation schedules should consider the relevant decision-making processes to deliver timely results.

TRANSPARENCY

In accordance with [OMB Memorandum M-20-12](#), Federal evaluations must be transparent in the planning, implementation, and reporting phases to promote accountability and help ensure that aspects of an evaluation are not tailored to generate specific findings. To the extent allowable by law, the Department will promptly and openly share evaluation designs, methodologies, analyses, interpretations, and results, consistent with the Freedom of Information Act and other relevant laws and regulations (studies related to internal management, legal, risk, or enforcement procedures will not be shared publicly). Please refer to the Department’s [Public Access Plan](#) for additional information. Evaluation reports will present comprehensive results, including favorable, unfavorable, and null findings. The Department will also archive evaluation data for secondary use by interested researchers.

INDEPENDENCE AND OBJECTIVITY

Agency leadership and program staff should participate in setting evaluation priorities, identifying questions, and assessing the implications of findings. Evaluators should operate with an appropriate level of independence, insulated from undue influences and both the appearance and reality of bias that may affect their objectivity, impartiality, and professional judgment. Staff should refer to and abide by the Department’s [Memorandum on](#)

[Implementation of the Departmental Scientific Integrity Policy](#) for additional guidance on political appointees, whistleblower protections, and use of objective science.

Objectivity of data collected by the Statistical Official is protected by OMB Statistical Policy Directives Nos. 3 and 4 and by the Bureau of Transportation Statistics authorizing language in [49 USC 6302\(d\)](#).

ETHICS

The Department will conduct evaluation activities in an ethical manner, adhering to all applicable human subject protection laws and policies and respecting the rights, safety, dignity, and privacy of all participants. The Department will adhere to the Federal Data Strategy Data Ethics Framework.

APPENDIX: DEFINITIONS, ACRONYMS, AND ABBREVIATIONS

Foundational Fact-Finding: Foundational research and analysis such as aggregate indicators, exploratory studies, descriptive statistics, and basic research.

Performance Measurement: Ongoing, systematic tracking of information relevant to policies, strategies, programs, projects, goals or objectives, and/or activities.

Policy Analysis: Analysis of data, such as general-purpose survey or program-specific data, to generate and inform policy, e.g., estimating regulatory impacts and other relevant effects.

Program: A set of projects or activities that support a higher-level objective or goal. Program includes processes, projects, interventions, policies, operations, activities, entities, and functions. Program operations are the strategies, processes, and activities management used to convert program inputs into program outputs.

Program Evaluation: An assessment using systematic data collection and analysis of one or more programs, policies, and organizations intended to assess their effectiveness and efficiency.

Statistical Activities: The collection, compilation, processing, or analysis of data for the purpose of describing or making estimates concerning the whole, or relevant groups or components within, the economy, society, or the natural environment. Includes the development of methods or resources that support those activities, such as measurement methods, models, statistical classifications, or sampling frames.

Statistical Purpose: The description, estimation, or analysis of the characteristics of groups, without identifying the individuals or organizations that comprise such groups, [which] includes the development, implementation, or maintenance of methods, technical or administrative procedures, or information resources that support these purposes.

CDO	Chief Data Officer
DOT	U.S. Department of Transportation
EO	Evaluation Officer
EVIDENCE ACT	Foundations for Evidence-Based Policymaking Act of 2018
OMB	Office of Management and Budget
SO	Statistical Official