Evidence Act Toolkit

Evaluation Planning Tips

OES GSA

Annual Evaluation Plan Overview

Each significant evaluation listed on the Annual Evaluation Plan must include a description that, at a minimum, addresses the following components described by the Office of Management and Budget (OMB) in its Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance (OMB M-19-23):

- Questions to be Answered
- Information Needed for Evaluations
- Methods to be Used
- Anticipated Challenges
- Dissemination

This guide provides tips and considerations related to each component.

Questions to Be Answered

Key Consideration: What question(s) does your agency need to answer to inform learning priorities?

After completing your agency's learning agenda, you likely already have a head start on developing evaluation questions. Some learning agenda questions might already be specific enough to guide a study, while others might need to be refined with more detail.

Involving agency staff and leadership to develop evaluation questions can be critical to developing studies that are useful to the agency and feasible to implement. Key stakeholders might include sub-agency representatives, leaders of offices that house relevant programs or initiatives, and program managers. Agencies may choose to involve external stakeholders in evaluation planning, especially in targeted ways. For example, representatives of communities served by your agency can provide valuable feedback on the focus and design of potential evaluations. Your agency can then use this feedback to develop studies that are supported by and useful to its constituents. Similarly, academic researchers may offer insights into the needs of a particular research field and advise your agency on potential evaluation questions or strategies.

Tips for Evaluation Planning

- **Choose an approach for stakeholder engagement that is best suited to your agency.** Though agencies are strongly encouraged to involve stakeholders in the development of evaluation questions, their approaches for doing so may differ. Some potential approaches include:
    - *Developing a formal process for soliciting evaluation ideas from program staff.* See the description of the Small Business Administration's approach in Section 2 of *A Guide to Developing Your Agency's Annual Evaluation Plan* (available for download on the Evidence Act Toolkits Homepage).
    - *Taking a more informal approach to planning evaluations.* For example, evaluation staff might engage regularly with agency leadership and program managers about their learning priorities and opportunities for evaluation. The box on the following page describes an iterative approach used by the Department of Education to document evaluation needs and develop evaluation ideas from conversations with leadership and program managers.
    - *Developing potential evaluations with a core team.* This team could reach out to staff within program offices with specific ideas. They could also gather feedback on logistics (such as the optimal timing for an evaluation), develop study designs that will inform program improvement, and explore options for funding evaluations.
- **Develop evaluation questions that clearly state what your agency seeks to understand.** A carefully crafted evaluation question will help you to determine the most appropriate and rigorous research method. Consider the criteria for a strong evaluation question on the following page.

Key Consideration: What question(s) does your agency need to answer to inform learning priorities?

- *Evaluative:* The question supports an investigation of a program or elements of a program based on information collected. It allows you to assess the value or significance of a program.
- *Appropriate:* The question is linked to program goals and is appropriate for the program's stage of development.
- *Relevant:* The question aligns with the information needs of key stakeholders.
- *Specific:* The question clearly outlines what will be investigated, including what you are evaluating, what you will measure, and the population of interest.
- *Objective:* The question does not presume to know the answer in advance.
- *Measurable:* The question uses assessable terms (e.g., quantity, quality, frequency).
- *Feasible:* The question reflects real-world constraints (e.g., budget, data availability).

How the Department of Education Develops Ideas for Evaluations

An iterative process can help you develop an evaluation that aligns with learning priorities, has buy-in from agency staff, and is feasible within time and budget constraints. At the Department of Education, evaluation staff from the Institute of Education Sciences have facilitated discussions by creating Summary Study Options (SSOs).

- **What are SSOs?** SSOs are short, written sketches of potential evaluations that respond to agency priorities. These summaries provide a brief overview of questions, data, methods, timeline, estimated cost, and other considerations for each option.
- **Why are SSOs beneficial?** SSOs help program staff and leadership stay engaged and think concretely about evaluation options. They also allow evaluation staff to develop options efficiently by starting with short sketches.
- **How are SSOs used?** As conversations unfold, evaluation staff gradually add detail to the options that are of greatest agency interest.

Key Consideration: What question(s) does your agency need to answer to inform learning priorities?

- **Understand how to write descriptive, normative, or causal questions.** Evaluation questions can be grouped into three types that provide different information about a program or policy: descriptive, normative, and causal. Each question type requires different design considerations. For guidance on how to write strong questions for each type, see the table below.

QUESTION TYPE	START WITH...	PAIR WITH...	AVOID...
Descriptive	*Who* *What* *When* *Where* *How* *How much*	*Do/does* *Is/are* *Many* *Much*	Words like *why, relate, influence, effect, cause, increase*, or *decrease*, as these indicate a relationship that is not being tested
Normative	*To what extent*	*Associated with* *Related to* *Correlated with*	Words like *impact, lead to, cause*, or *effect*, as these indicate directionality that is not being tested
Causal	*How* *When*	*Lead to* *Increase* *Decrease* *Cause* *Compared to*	Ambiguous langauge that does not clearly specify the outcome you expect to change or who you expect it to change for

Information Needed for Evaluations

Key Consideration: What Information will produce results to inform strategic or operational decision(s) at your agency?

Your agency will need to assess what data are available to address evaluation questions. When reviewing existing data, consider whether and how evaluation results could inform decision-making. If new information is needed, consider what agency resources are available to collect new data, from what sources the data needs to be collected, and whether the data could be collected in a timely and cost-effective manner.

Tips for Evaluation Planning

- **Consider how administrative data can be used in evaluations.** Administrative data, such as program data or outcomes data that your agency already tracks, are a low-cost alternative to gathering new data. Keep in mind that using administrative data might mean that your agency will need to make specific plans for accessing and using it, such as obtaining clearances for contractors or planning for cleaning data for external use. The box below highlights several approaches that agencies have taken to use administrative data for evaluations.

How Agencies Have Used Administrative Data for Evaluation Purposes

Federal agencies have used administrative data to:

- **Examine trends over time.** The Department of Housing and Urban Development (HUD) used administrative data from 1995 to 2015 to examine program and household variations in assisted housing, as well as trends in length of stay and factors affecting length of program participation.
- **Provide descriptive statistics.** The Department of Agriculture (USDA) used administrative data from a longitudinal study of low-income families in three U.S. cities to determine whether enrollment in the Food Stamp Program increased in households with U.S.-born children and foreign-born heads after legal immigrants/access to the program was restored under the Farm Bill Act of 2002.
- **Understand associations between program features and outcomes.** The Department of Labor (DOL) used administrative data to understand factors associated with disabled workers' return to work and to examine the extent to which the timing of disability management services resulted in positive outcomes for injured workers.

Key Consideration: What Information will produce results to inform strategic or operational decision(s) at your agency?

- **Evaluate effectiveness through a randomized controlled trial.** The Department of Education (ED) Institute of Education Sciences conducted an experiment to evaluate the impact of the Teacher Incentive Fund, which offered pay-for-performance bonuses and opportunities for teachers to take on new roles. The study used administrative data from participating districts.
- **Develop predictive models.** The Mine Safety and Health Administration (MSHA) of the DOL sought to predict at an early stage which mine operating firms with health and safety violations might be at risk of failing to pay their fines on time. The study used internal MSHA administrative data to develop predictive models to identify mine operators who were at a high risk of failing to make timely payments to MSHA.

Keep in mind the Paperwork Reduction Act (PRA) and research approvals that might be needed from other jurisdictions or entities. If administrative data are not sufficient for addressing an evaluation question, new data might be collected by administering surveys or assessments or conducting interviews or focus groups. Consult relevant laws and engage your agency's legal counsel prior to collecting new data.

Consider partnering with other federal agencies to share data and provide a more comprehensive picture of program need, implementation, or impact. Data sharing partnerships can be a low-cost way to receive information needed for evaluations. Keep in mind that data definitions may differ, and for an effective evaluation, your agency should collect data in consistent ways.[1]

[1] GAO. Designing Evaluations (2012).

Methods to be Used

Key Consideration: Does your evaluation design match the evaluation question(s)?

Once you have created evaluation questions, a next step is developing a plan for how you will answer them. A strong evaluation design uses methods that are well-matched to the evaluation questions, which means that the questions should drive the choice of methods, not the other way around. The table below illustrates how question types map to possible evaluation designs.

QUESTION TYPE	PURPOSE	EXAMPLE QUESTION	EVALUATION DESIGNS
Descriptive	Describe or understand a program or process	*Who received which services?*	◆ Process/implementation evaluation ◆ Descriptive tabulations ◆ Trend lines
Normative	Measure against previously established criteria	*To what extent was the target met?*	◆ Outcome evaluation ◆ Pre-post design
Causal	Determine the impact of an intervention	*Do participants have improved outcomes as a result of the intervention?*	◆ Impact study (e.g., experiment, quasi-experiment)

Key Consideration: Does your evaluation design match the evaluation question(s)?

Tips for Evaluation Planning

- **Use the most rigorous methods possible within budget and other resource constraints.** Rigorous methods have several features – see the box to the right for the definition of "rigor" provided in OMB Circular A-11.

- **Consider when results will be needed.** If the need for information is pressing, long-term methods will not be useful.

- **Assess the maturity of a program.** Is the program ready for an impact study? If not, performance tracking or a simple outcome study may be more appropriate.

- **For each evaluation design, select the appropriate method(s) for answering questions.** These can be qualitative (such as interviews or case studies) or quantitative (such as analysis of existing administrative data). See the box below for more examples of evaluation methods.

Excerpt from OMB Circular A-11

What is rigor?

Rigorous evaluations meet the following requirements:

- Inferences about cause and effect are well-founded (internal validity)
- There is clarity about the populations, settings, or circumstance to which results can be generalized (external validity)
- Measures accurately capture the intended information (meausurement reilability and vaildity)
- Samples are large enough for meaningful inferences
- Evaluations are conducted with an appropriate level of independence by experts external to the program either inside or outside an agency

- **Surveys**
- **Observations**
- **Focus Groups**
- **Document Review**
- **Pre-post analysis**
- **Statistical analysis**
- **Quasi-experimenal study**
- **Randomized controlled trial (RCT)**

- **Understand the benefits of quantitative and qualitative data.** Quantitative data is more generalizable than qualitative data, assuming a representative sample, but it provides limited understanding related to program context. Qualitative data is better suited to addressing the question of "why". Given the benefits and limitations of both methods, using several methods could be the best way to answer your key questions.

Anticipated Challenges

Key Consideration: How could challenges affect the quality or timing of each evaluation?

All evaluation designs have strengths and weaknesses and come with challenges related to data and methods. Time constraints and resource limitations are also common challenges.

Tips for Evaluation Planning

- **Consider challenges across several categories.** The table below presents a brief, not exhaustive, list of challenges that may apply to an evaluation, and suggested mitigation strategies.

Evaluation Challenge Example	Potential Mitigation Strategies
External factors, such as economic conditions, could influence the evaluation	◆ Measure contextual factors to the extent possible (e.g., qualitative assessments can help in estimating how external factors affect program processes and results)
Lack of access to data	◆ Identify pre-existing data sources and explore opportunities to link data with other agencies ◆ Consider whether new data collection is necessary and feasible
Lack of appropriate measures	◆ Reach out to stakeholders (e.g., academic experts) to assist with identifying or developing measures
Challenges of finding a comparison group	◆ Consider alternative designs ◆ Look for opportunities to use within-group comparisons
Conflicting results	◆ Consider accuracy of methods, your confidence in the theory of change, the statistical significance of findings, assumptions made by statistical tests, and the match between methods and evaluation questions
Desire for quick production of results and need to inform real-time decision-making	◆ Develop plans for sharing interim results with key stakeholders

**Table adapted from CDC Overview of Policy Evaluation, Appendix C.*

Dissemination

Key Consideration: Who are the primary users of evaluation results?

Dissemination strategies are an integral component of evaluation planning. It is important to identify the stakeholders whom you will engage with throughout the course of the evaluation and share findings. Complete **Worksheet #2** in the *Annual Evaluation Plan Workbook* (available for download on the Evidence Act Toolkits Homepage) to document a landscape of stakeholders and determine the best reporting approaches for each group.

Tips for Evaluation Planning

- **Consider communication formats tailored to stakeholders' needs and preferences.** Remember to think broadly about how to inform. For example, if you submit a final evaluation report exclusively to program leadership upon conclusion of the study, you might miss an opportuinty for feedback from program staff about critical next steps. The table below describes suggested communication approaches for different stakeholder groups.

Stakeholder Group	Reporting Approach	Example Format	Example Channel
Program Office	Clear, program improvement-relevant takeaways or action items	List of recommendations	◆ In-person meeting
Evaluation Office/ Academic Researchers	Precise wording and technical details	Final Report	◆ Conference
Agency Leadership	Clear, policy-relevant takeaways	One-page brief	◆ In-person meeting ◆ Email
Congress	Emphasis on rigor and evidence around changes in outcomes	Executive summary or memo	◆ Email
Grantees	Clear, program improvement-relevant takeaways or action items	Summary of takeaways or presentation	◆ Newsletter ◆ Webinar ◆ Listserv
Media/Public	Easy to understand, visually appealing	Infographics	◆ Social media ◆ Blog ◆ Website

Key Consideration: Who are the primary users of evaluation results?

Employ the suite of communication formats and channels available at your agency. Consider involving a communications expert in the evaluation planning process for dissemination support. In the tables below, you will find a list of potential communication formats and channels for disseminating evaluation results. Regardless of the format or channel, unless you are writing for an academic audience or publishing in a journal, aim to write clear evaluation reports without jargon.[2] Tables and charts should be accessible to multiple audiences and communicate key takeaways.

Suggested Formats

Dashboards	Scorecards	Recommendations
Executive Summaries	Briefs	Findings Tables
Interim Reports	Brochures	PowerPoint Presentations
Final Reports	Postcards	Infographics
Memos	Flipcharts	Data Visualizations

Suggested Channels

Websites	In-Person meetings	News Releases
Conferences	Teleconferences	Newsletters
Social Media	Videos	Bulletins
Blogs	Emails/Listservs	Podcasts
Webinars		Displays/Exhibits

Look to other agencies for leading practices. Some federal agencies, such as the Department of Housing and Urban Development (HUD), have robust dissemination capabilities and processes. The box on the following page includes dissemination tips and strategies used by HUD, including suggestions on how to share findings about a budget.

[2] Torres, Rosalie, Preskill, Hallie, Piontek, Mary. Evaluation strategies for communicating and reporting: Enhancing learning in organizations (1996).

Key Consideration: Who are the primary users of evaluation results?

Dissemination Strategies from the Department of Housing and Urban Development

The HUD Office of the Assistant Secretary for Policy Development and Research (PD&R) has robust dissemination practices for its research results. This includes the HUD research library, quarterly public events, the newsletter Evidence Matters, and a biennial PD&R report. When planning how your agency will disseminate results, keep in mind the following tips from PD&R:

- **Engage the network of your agency's practitioners.** It is important to know who is going to find specific results useful, and practitioners in your agency will likely know best. For example, work with program staff to create an email listserv of their key partners to disseminate results on specific topics.
- **Make your research consumable.** To reach a broad audience, long and complex evaluation and research results should be summarized for key highlights. For an example, see the Spring/ Summer 2019 edition of Evidence Matters to read a summary of "Evaluating Place-Based Incentives."
- **Employ low-cost communications platforms.** Not every agency has the resources to develop a magazine or biennial report. In addition to listservs, consider other low-cost communication platforms to disseminate results, such as hosting webinars or partnering on existing conferences or events. For example, HUD holds its quarterly events via webcast to enable regional practitioners to participate.