

PROGRAM EVALUATION STANDARDS

OMB M-20-12, Phase 4 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Program Evaluation Standards and Practices

The Office of Management and Budget (OMB) Memorandum [M-20-12](#), “Phase 4 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Program Evaluation Standards and Practices,” provides program evaluation standards to guide agencies in developing and implementing evaluation activities, evaluation policies, and in hiring and retaining qualified staff. OMB M-20-12 [Appendix B](#) includes a more detailed explanation of the standards, and [Appendix D](#) includes a bibliography of sources that directly influenced the identification of the standards.

RELEVANCE AND UTILITY: Federal evaluations¹ must address questions of importance and serve the information needs of stakeholders in order to be useful resources. Evaluations should present findings² that are actionable and available in time for use. Information should be presented in ways that are understandable and that can inform agency activities and actions such as budgeting, program improvement, accountability, management, regulatory action, and policy development.

RIGOR: Federal evaluations must produce findings that Federal agencies and their stakeholders can confidently rely upon, while providing clear explanations of limitations. The quality of an evaluation depends on the underlying design and methods, implementation, and how findings are interpreted and reported. Credible evaluations must be managed by qualified evaluators with relevant education, skills, and experience for the methods undertaken. An evaluation must have the most appropriate design and methods to answer key questions, while balancing its goals, scale, timeline, feasibility, and available resources.

INDEPENDENCE AND OBJECTIVITY: Federal evaluations must be viewed as objective in order for stakeholders, experts, and the public to accept their findings. This depends on the independence and objectivity of the evaluators. Federal agencies should enable evaluators to, and evaluators should, operate with an appropriate level of independence from programmatic, regulatory, policymaking, and stakeholder activities. While stakeholders have an important role in identifying evaluation priorities, the implementation of evaluation activities, including how evaluators are selected and operate, should be appropriately insulated from political and other undue influences that may affect their objectivity, impartiality, and professional judgement. 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.

TRANSPARENCY: Federal evaluation must be transparent in the planning, implementation, and reporting phases to enable accountability and help ensure that aspects of an evaluation are not tailored to generate specific findings. Decisions about the evaluation’s purpose and objectives (including internal versus public use), the range of stakeholders who will have access to details of the work and findings, the design and methods, and the timeline and strategy for releasing findings should be clearly documented before conducting the evaluation. These decisions should take into consideration any legal, ethical, national security, or other constraints for disclosing information. Once evaluations are complete, comprehensive reporting of the findings should be released in a timely manner and provide sufficient detail so that others can review, interpret, or replicate/reproduce the work.

ETHICS: Federal evaluations must be conducted to the highest ethical standards to protect the public and maintain public trust in the government’s efforts. Evaluations should be planned and implemented to safeguard the dignity, rights, safety, and privacy of participants and other stakeholders and affected entities. Evaluators should abide by current professional standards pertaining to treatment of participants. Evaluations should be equitable, fair, and just, and should take into account cultural and contextual factors that could influence the findings or their use.

¹ “Program evaluation” and “evaluation” are synonymous. The term “evaluation” signifies “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). “Evaluation,” “program,” and other key terms are further described in OMB M-20-12 [Appendix A](#).

² “Findings” refers to results, conclusions, and recommendations that are systematically generated through analyzing and interpreting data. “Design and methods” is used to collectively address the structure of an evaluation, inclusive of evaluation approach; variables for, conditions under, timing of, and sources from which data are used or collected; and quantitative and qualitative data collection and analysis methods.