Evaluation Policy

[Insert city name] | [date]

| How to Use this Template  * All instructions are highlighted in orange. Once you have completed the section pertaining to the instructions, delete the instructions.  References This template was inspired by:   * [Office of Evaluation Services Evaluation Policy](https://oes.gsa.gov/assets/files/evaluationpolicy.pdf) * [State of Minneapolis Evaluation Policy](https://mn.gov/mmb-stat/results-first/evaluation-policy.pdf) * [Longmont Evaluation Policy](https://docs.google.com/document/d/1F6xl7RenDuh4o-zVZ4p_cEcnjTLHCujp/edit) * [Bogota Monitoring and Evaluation Policy](http://www.sdp.gov.co/gestion-socioeconomica/politicas-sectoriales/seguimiento-y-evaluacion) * [Bill and Melinda Gates of Foundation Evaluation Policy](https://www.gatesfoundation.org/about/policies-and-resources/evaluation-policy) * [Administration for Children and Families’ (ACF) Evaluation Policy](https://www.acf.hhs.gov/opre/report/acf-evaluation-policy) * [Department of Labor Evaluation Policy](https://www.dol.gov/agencies/oasp/evaluation/EvaluationPolicy) * [Department of Housing & Urban Development Evaluation Policy](https://www.federalregister.gov/documents/2021/08/13/2021-17339/hud-program-evaluation-policy-policy-statement) * [USAID Evaluation Policy](https://www.usaid.gov/evaluation/policy) * [Corporation for National and Community Service Evaluation Policy](https://americorps.gov/sites/default/files/document/CNCSEvaluationPolicy.pdf) * [Results for America Standards of Excellence for Evaluation (Federal & State)](https://results4america.org/tools/2021-invest-works-federal-standard-excellence/) |
| --- |

# 

# Table of Contents

[**How to Use this Template 1**](#_tyjcwt)

[References 1](#_17dp8vu)

[**Table of Contents 2**](#_26in1rg)

[Background 3](#_35nkun2)

[Purpose / Rationale 3](#_1ksv4uv)

[Definition of Evaluation 3](#_44sinio)

[Policies / Programs / Services 3](#_2jxsxqh)

[Core Principles 3](#_z337ya)

[Enforcement 3](#_3j2qqm3)

[**Sample Policy Language 4**](#_o8rvblidwuu)

[Background 4](#_4i7ojhp)

[● Sample I: Longmont, CO 4](#_2xcytpi)

[● Sample II: CNCSE 4](#_1ci93xb)

[● Sample III: DOL 4](#_3whwml4)

[● Sample IV: City of Buenos Aires, Argentina 4](#_2bn6wsx)

[Policy Purpose / Rationale 5](#_qsh70q)

[● Sample I: Paterson, NJ 5](#_3as4poj)

[● Sample II: CNCSE 5](#_1pxezwc)

[● Sample III: Gates 5](#_49x2ik5)

[Definition of Evaluation 5](#_2p2csry)

[● Sample I: Paterson, NJ 5](#_147n2zr)

[● Sample II: Bogota, Colombia 5](#_3o7alnk)

[● Sample III: Gates Foundation 5](#_23ckvvd)

[● Sample IV: HUD 6](#_ihv636)

[● Sample V: ACF HHS 6](#_32hioqz)

[Included Policies, Programs & Services 6](#_1hmsyys)

[● Sample I: BIT Suggestion 6](#_41mghml)

[● Sample II: Gates 6](#_2grqrue)

[Core Evaluation Principles 7](#_vx1227)

[● Sample I: OES Core Principles 7](#_3fwokq0)

[● Sample II: State of Minneapolis 7](#_1v1yuxt)

[● Sample III: City of Paterson 9](#_4f1mdlm)

# 

# Background

Describe why this policy is being created and how it relates to your city’s mission and leadership priorities. In most cases, this is 1-2 paragraphs. If you’ve done past evaluation work, it may be helpful to reference that work to build additional context.

# Purpose / Rationale

Describe the purpose of this policy. In most cases, this is 1-2 simple sentences that articulate what this policy is aiming to do. (i.e., “This policy aims to establish the City of XX evaluation guidelines and guiding principles to support widespread use of high-quality evaluations.)

# Definition of Evaluation

Include your city’s definition of evaluation. This definition should include a general statement about what evaluation is and may be expanded to include specific types of methods. Additionally, it may be helpful to describe how evaluation supports other quality and performance management tools that the city uses.

# Policies / Programs / Services

Describe what types of policies, programs, and services your city will require/encourage the use of evaluation. This section is quite flexible to fit your city’s needs. Some organizations identify policies/programs/services for evaluation by potential impact or cost, while others provide a more general statement encouraging widespread use.

# Core Principles

Describe your city’s guiding principles that will support the high-quality evaluation work and descriptions of what these principles will mean within your specific city context. BIT recommends adopting the five principles developed by the Federal Office of Evaluation Services (rigor, relevance, transparency, independence, ethical practice) in addition to any city-specific principles.

# Enforcement

Describe how the policy will be made a reality in your city, including any processes that will be altered. If it is not feasible for you to include enforcement, we recommend passing a policy without this section. Articulating the vision and definitions can be impactful on their own and/or with implementation actions or other commitments that are not included within the policy itself.

# Sample Policy Language

This section includes language from cities, states, and federal agency evaluation policies.

## Background

| Sample I: Longmont, CO The City of Longmont is committed to the continual improvement of our services and programs. As part of this commitment, we encourage program and service managers to conduct rigorous, relevant evaluations and to use evidence from evaluations to inform policy and practice. Sample II: CNCSE This evaluation policy statement presents key principles that govern the Corporation for National and Community Service’s (CNCS’s) planning, conduct, and use of program evaluations. The policy expresses our commitment to conducting rigorous, relevant evaluations and using evidence from evaluations to inform policy and practice. CNCS seeks to promote rigor, relevance, transparency, independence, and ethics in the conduct of evaluations. This policy addresses each of these principles. The mission of the Corporation for National and Community Service is to “…improve lives, strengthen communities and foster civic engagement through service and volunteering.” The importance of these goals demands that we continually innovate and improve and that we evaluate the performance and effectiveness of our programs and activities. Through evaluation, CNCS and our partners can learn systematically so that we can make our services as effective as possible. Sample III: DOL The mission of the Department of Labor is to "...foster, promote, and develop the welfare of the wage earners, job seekers, and retirees of the United States; improve working conditions; advance opportunities for profitable employment; and assure work-related benefits and rights." The importance of these goals demands that we continually innovate and improve, and that we evaluate the performance and effectiveness of our programs and activities. Through evaluation, DOL can learn systematically so that we can make our services as effective as possible. Sample IV: City of Buenos Aires, Argentina The objective of the Direction for Evaluation and Quality is to “measure and evaluate the quality of the services and benefits provided by the different areas of the Government of the City of Buenos Aires in a centralized, independent and multidimensional way, generating a comprehensive vision, with a focus on the neighbor.” |
| --- |

## Policy Purpose / Rationale

| Sample I: Paterson, NJ The purpose of this policy document is to establish the City of Paterson’s evaluation guidelines so that evaluation practices become more consistent and widely used across the organization. Sample II: CNCSE The purpose of this policy is to articulate the key principles that govern the agency’s program evaluations. Sample III:Gates Our evaluation policy is intended to help foundation staff, and our partners align their expectations in determining why, when, and how to use evaluation. |
| --- |

## Definition of Evaluation

| Sample I: Paterson, NJ Evaluation is an independent, systematic investigation into how, why, and to what extent objectives or goals of a program have been achieved. It can help the City of Paterson answer key questions about policies, grants, initiatives, or strategies. (Paterson, NJ) Sample II: [Bogota, Colombia](https://www.sdp.gov.co/sites/default/files/guia_para_el_seguimiento_final_impresa.pdf) “Evaluation is the systematic process of observation, measurement, analysis and interpretation aimed at understanding a public intervention, whether a policy, a plan, a program or a project, to reach a judgement assessment based on evidence regarding its design, implementation, effects, results and impacts. The purpose of the evaluation is to be useful for decision-making and as a mechanism of accountability to citizens.” Sample III: [Gates Foundation](https://www.gatesfoundation.org/about/policies-and-resources/evaluation-policy) Evaluation is the systematic, objective assessment of an ongoing or completed intervention, project, policy, program, or partnership. Evaluation is best used to answer questions about what actions work best to achieve outcomes, how and why they are or are not achieved, what the unintended consequences have been, and what needs to be adjusted to improve execution. When done well, evaluation is a powerful tool to inform foundation and partner decision making about how to optimize scarce resources for maximum impact. It is distinct from other forms of measurement that focus only on observing whether change has occurred, not why or how that change occurred. Sample IV: [HUD](https://www.federalregister.gov/documents/2021/08/13/2021-17339/hud-program-evaluation-policy-policy-statement) Section 101 of the Evidence Act defines “evaluation” to mean “an assessment using systematic data collection and analysis of one or more programs, policies, and organizations intended to assess their effectiveness and efficiency.” Sample V: [ACF HHS](https://www.acf.hhs.gov/opre/report/acf-evaluation-policy) Evaluation produces one type of evidence. A learning organization with a culture of continuous improvement requires many types of evidence, including not only evaluation but also descriptive research studies, performance measures, financial and cost data, survey statistics, program administrative data, and feedback from service providers, participants, and other stakeholders. Further, continuous improvement requires systematic approaches to using information, such as regular data-driven reviews of performance and progress. Although this policy focuses on evaluation, the principles and many of the specifics apply to the development and use of other types of evidence as well. |
| --- |

## Included Policies, Programs & Services

| Sample I: BIT Suggestion Evaluation is strongly suggested for newly funded programs or services and for renewed funding opportunities for existing programs or services where there is no significant evidence base. Sample II: Gates Evaluation is particularly warranted in the following instances:   * + When evidence is needed to fill a knowledge gap or evaluate a significant policy decision. Evaluation can help resolve uncertainty and determine the relative cost-effectiveness of different interventions, models, or approaches.   + When our partners and we need a better understanding of how a cluster of important investments or a specific program or project is performing.   + When an organization, intermediary, or consortium that we work with is at a critical stage of development and can benefit from an independent performance assessment.   + When a program team needs to assess the progress of a new operational model or approach. Evaluation provides reliable, independent feedback about what needs to be improved to strengthen our approach and partner relationships. |
| --- |

## Core Evaluation Principles

| Sample I: OES Core Principles  * + Rigor: Our findings should be credible and mean what they purport to mean. Evaluations are conducted to the highest standards; our statements about statistical significance are clear and correct, and the limitations of our findings are clear. All OES analyses go through an internal replication, and any discrepancies between the two analyses are addressed before the results are finalized. Bottom line: Policymakers and program designers should be able to act on our findings with confidence.   + Relevance: Project selection should take into account (1) the policy or program priorities of our agency collaborators and (2) potential impact on a priority outcome-based. To OES, relevance is also ensured by exclusively utilizing existing administrative data as the outcome of our evaluations.   + Transparency: We are committed to ensuring that researchers, agency collaborators, policymakers, and the public at large are able to learn from our work. We keep a public record of all evaluations fielded and publicize all of our findings (including null results and those that run counter to our own prior expectations and goals). We conduct our work in such a way that we can verify our results, and we document our work thoroughly so that others with required data access can reproduce the results.   + Independence: We retain control over decisions about project selection and the ability to release our evaluation results.   + Ethical Practice: OES evaluations will be conducted in an ethical manner and safeguard the dignity, rights, safety, and privacy of participants. Evaluations will comply with both the spirit and the letter of relevant requirements, such as regulations governing research involving human subjects.  Sample II: State of Minneapolis  * + Rigor: We are committed to using the most rigorous methods available given our evaluation questions, budget, and other constraints. Rigor is relevant in all forms of evaluation, including process, outcome, and impact evaluations that use qualitative and/or quantitative data. Rigor requires ensuring that inferences about cause and effect are well-founded (internal validity); requires clarity about the populations, settings, or circumstances to which results can be generalized (external validity); and requires the use of measures that accurately capture the intended information (measurement reliability and validity). In order to assess the effects of programs or services in a rigorous way, our evaluations use methods that isolate to the greatest extent possible the impacts of the programs or services from other influences such as trends over time, geographic variation, or pre-existing differences between participants and non-participants. For causal questions, experimental approaches are preferred. When experimental approaches are not feasible, high-quality quasi-experiments offer an alternative. Accordingly, we employ staff with academic training and experience in a range of relevant social science disciplines. We provide professional development opportunities so that staff can keep their skills current. We also consult with external advisors who are leaders in relevant fields.   + Relevance: Evaluation priorities should reflect the interests, needs, and sensitivities of the populations served; legislators and other state leaders; partners such as other state agencies, tribes, local governments, and grantees; researchers; and other stakeholders. There should be strong partnerships amongst evaluation staff, frontline program staff, policy-makers, target populations, and service providers. These stakeholders should have the opportunity to influence the questions that will be answered by evaluations. For new initiatives, evaluations are more feasible and useful when planned with stakeholders from the outset of the initiative rather than as an afterthought. We disseminate findings in ways that are accessible and useful to policy-makers and practitioners. We also strive to provide this information at the right time, place, and format to facilitate use. We integrate both uses of existing evidence and opportunities for further learning into all of our activities. Where prior evidence exists, we use it. Where an evidence base is lacking, we seek to build evidence through strong evaluations. The emphasis on evidence is to support—not inhibit—innovation, improvement, and learning.   + Independence: Independence and objectivity are core principles of evaluation. Agency and program leadership, program staff, service providers, and others should participate actively in setting evaluation priorities, identifying evaluation questions, and assessing the implications of findings. However, it is important to insulate evaluation functions from undue influence and from both the appearance and the reality of bias. In order to objectively assess programs, we protect independence in the design, conduct, and analysis of evaluations. To this end, the Results Management team has the authority to approve the design of evaluation projects and to publish evaluation findings.   + Transparency: To promote accessibility and replicability, we pre-register evaluation hypotheses, analysis plans, and limitations in advance (MN OSF page). We release evaluation results regardless of the findings. Final evaluation reports describe the methods used, including strengths and weaknesses, and discuss the generalizability of the findings. These reports present comprehensive results, including favorable, unfavorable, and null findings. We release results in a timely manner, usually within two months of a report’s completion. When permitted by data privacy laws and feasibility, we archive evaluation data for secondary use by interested researchers.   + Ethics: We conduct high-quality evaluations in an ethical manner and safeguard the dignity, rights, safety, and privacy of participants and their data. We seek to uphold the Belmont Report principles of respect for persons (protecting participant autonomy), beneficence (minimizing participant risks and maximizing benefits), and justice (fair distribution of any risks and benefits in participating in the study). To support these aims, we seek appropriate Institutional Review Board approval for evaluations classified as Human Subjects Research, uphold both the spirit and the letter of relevant regulations requirements, and adhere to professional guidelines (such as the American Evaluation Association’s Ethical Guiding Principles).  Sample III: City of Paterson  * + We lead with purpose. We design evaluation trials with actions and decisions in mind. We ask, “How and when will we use the information that comes from this evaluation?” By anticipating our information needs, we will design evaluations that are useful to evaluate the effectiveness of a policy change or a program design. Too often, program evaluations are commissioned without a clear purpose, and then shelved without generating useful insights. We do not want to fall into that trap.   + Evaluation is fundamentally a learning process. As we engage in evaluation planning, implementation, and use of results, we actively learn and adapt. We stay agile. As we implement our strategies, we use evaluation as a key vehicle for learning and bringing new insights to our work.   + Evaluation is tied to strategic goals. We cannot evaluate everything, so we choose strategically. Department and mayoral priorities should guide decisions about where to focus our evaluation efforts, including the opportunity for learning, any urgency to make course corrections, the potential for organization and/or reputational risk.   + We choose methods of measurement that allow us to maximize rigor. We seek to match methods to our questions and do not routinely choose one approach or method over others. We seek to use multiple methods and data sources when possible in order to strengthen our evaluation design and reduce bias. All evaluations should clearly articulate methods used and their limitations.   + We share our intentions to evaluate our findings with appropriate audiences. As we plan evaluations, we consider and identify audiences for the findings. We communicate early with our grantees and our residents about our intention to evaluate and involve them as appropriate in issues of design and interpretation. We presumptively share the results of our evaluations so that others may learn from our successes and failures. We will make principled exceptions on a case-by-case basis, with care given to issues of confidentiality.   + We use data. We take time to reflect on the results, generate implications for policy or practice, and adapt as appropriate. We recognize the value in combining the insights from evaluation results with the wisdom from our own experiences. |
| --- |