Evaluation Guidelines

State of Tennessee

Tennessee Department of Finance & Administration | September 2023

Contents

Purpose

The purpose of these Evaluation Guidelines is to define and describe principles to guide the planning and execution of program evaluations within the State of Tennessee.

The State of Tennessee emphasizes the State's commitment to supporting rigorous, relevant evaluations and to evidence-based decisions to drive continuous improvement and careful, thoughtful use of taxpayer dollars. As the State's evaluation efforts mature, the State of Tennessee will review and update these guidelines as needed.

Scope

The intention is to foster a shared understanding of language and principles for all program evaluations conducted by the State of Tennessee'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 State has solicited in conjunction with the award of contracts, grants, or cooperative agreements.

Roles and Responsibilities

The Office of Evidence and Impact (OEI) assumes a central role in the State of Tennessee, acting as the driving force behind evidence-based decision-making. In the context of evaluation OEI will:

- Collaborate with the governor's office and state leaders to identify strategic evidence needs and develop a comprehensive Learning Agenda, guiding evidence generation and evaluation efforts across all state departments.
- Support agency leadership with identifying specific programs that would benefit from evaluation.
- Provide technical assistance and resources to state agencies to support their evaluation efforts including but not limited to funding, training, evaluation design consultation, external evaluation partner identification, and results dissemination planning.

OEI's Mission & Vision

Foster a culture of data and evidence-based policymaking and budgeting that pursues the best outcomes for Tennesseans.
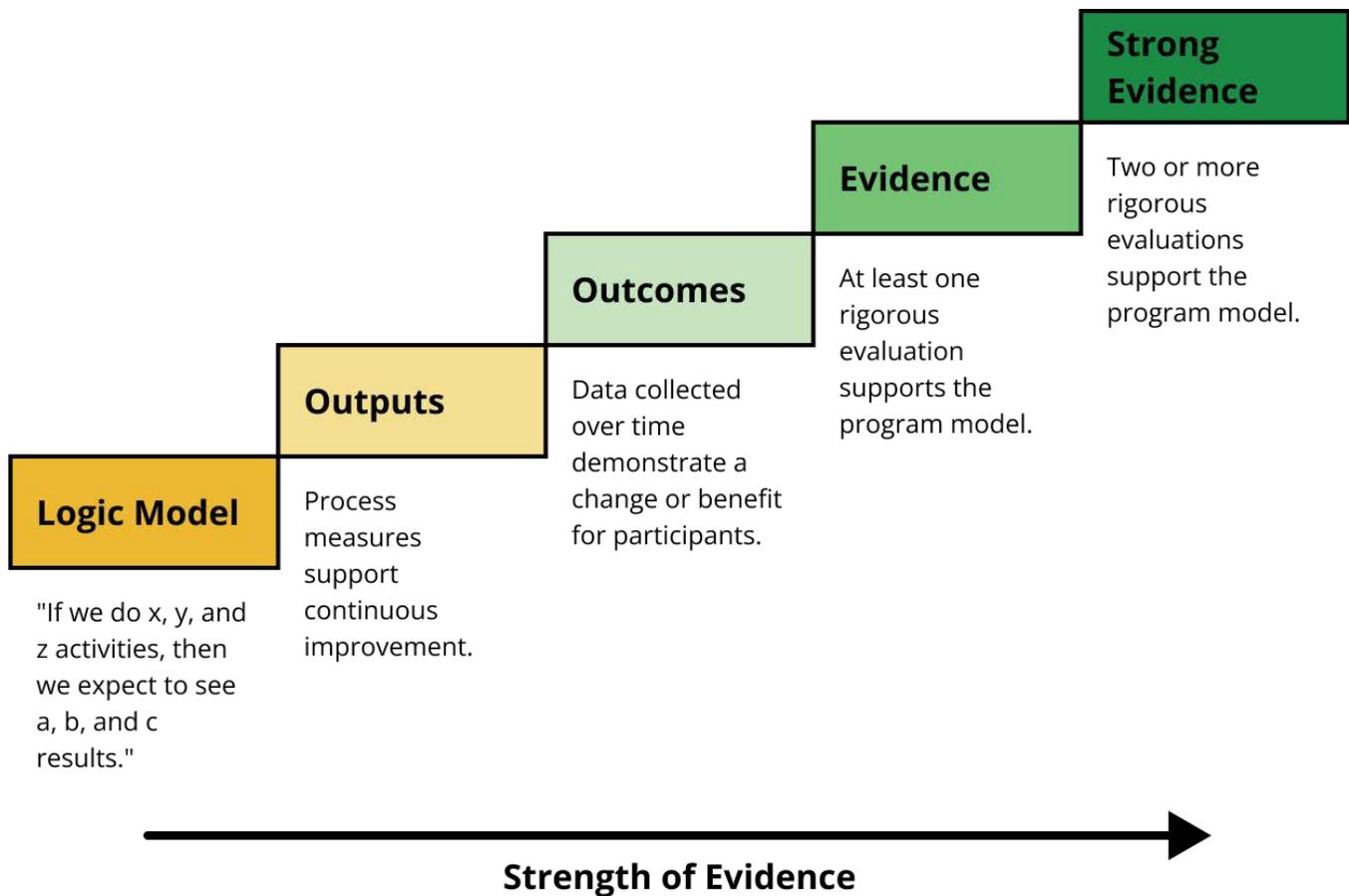
Better information empowers decision making that serves every Tennessean.

Each state agency assumes the responsibility of conducting internal and external evaluations independently but is encouraged to do so in line with the principles laid out in the State of Tennessee's Evaluation Guidelines. The Evaluation Guidelines provide a framework for agencies to develop rigorous evidence for its programs such that innovative programs are structured for rigorous evaluation and ongoing programs meet the state's highest levels of evidence in the Tennessee Evidence Framework. The Evaluation Guidelines can enable an agency to develop a program evaluation to transition a program from "Outcomes" to "Evidence" in the Tennessee Evidence Framework.

Evidence at the State of Tennessee

What is evidence?

Evidence refers to a robust body of research that evaluates the effectiveness of existing programs or proposed pilots. Building evidence involves a range of qualitative and quantitative studies— including those in the Tennessee Evidence Framework. The pinnacle of evidence-building activities is rigorous program evaluation using experimental or quasi-experimental designs (see Figure 1 for more about evaluation types).

Logic Model

"If we do x, y, and z activities, then we expect to see a, b, and c results."

Outputs

Process measures support continuous improvement.

Outcomes

Data collected over time demonstrate a change or benefit for participants.

Evidence

At least one rigorous evaluation supports the program model.

Strong Evidence

Two or more rigorous evaluations support the program model.

Strength of Evidence →

What does evidence do?

Evidence helps us learn and make better decisions about programs and policies in our State. Evidence and the use of data play a pivotal role in informing the budget, as well as strategic and operational decision-making.

Existing evidence is catalogued and utilized via the state's annual evidence-based budgeting and program inventory processes. It is equally important to provide opportunities to create new evidence where it doesn't exist through data analytics and program evaluation.

Done consistently and in a focused manner, evidence building will result in a budget that is increasingly informed by solid data and research. These activities ensure the state allocates resources responsibly and improves outcomes for all Tennesseans.
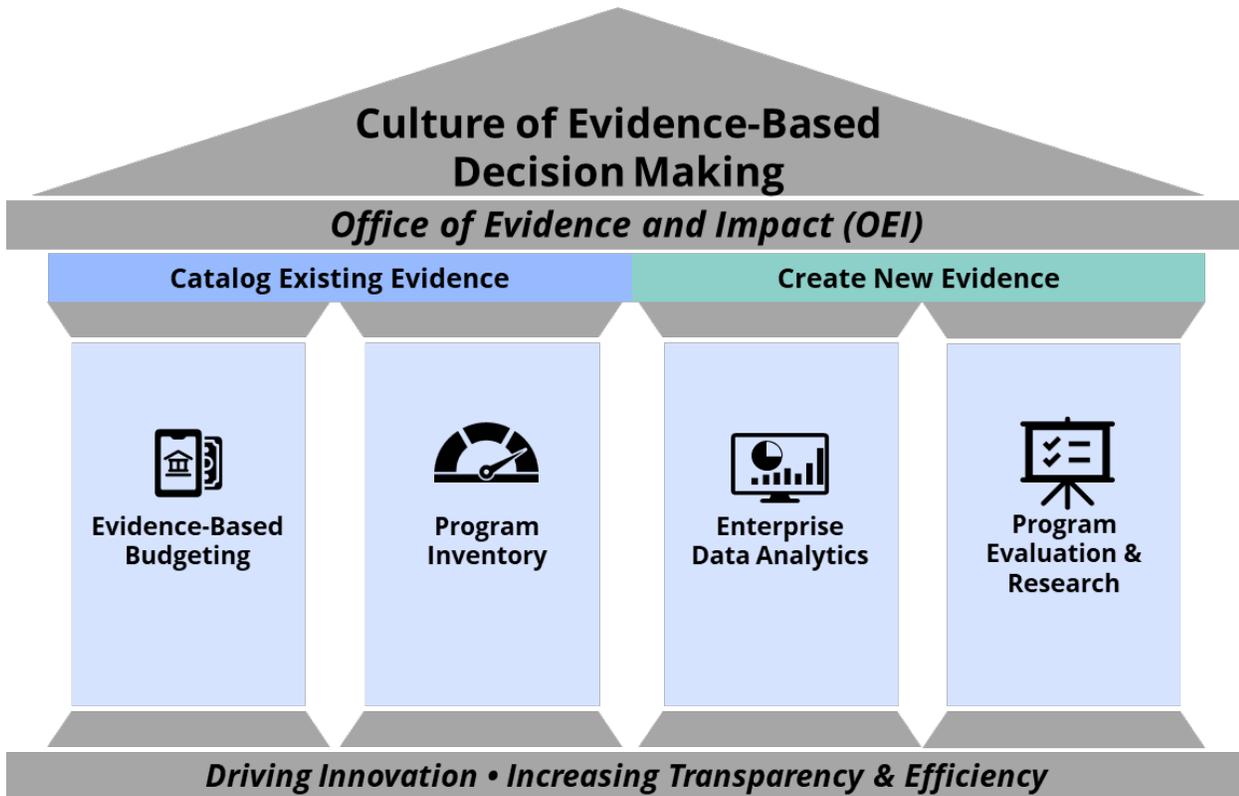

Figure 1 Tennessee Evidence Framework

How do agencies engage in evidence-building?

Agencies actively contribute to evidence-building by engaging with the components of the Tennessee Evidence Framework and continuously assessing their quality for each program, including **logic models** and **outcomes**, and working to move programs up the continuum of the framework. Agencies should regularly reflect on whether they are measuring the right outcomes and pursue access to data sources outside their department that increase the understanding of program performance.

Existing programs that have not yet reached the evidence or strong evidence steps in the Framework should assess how **program evaluation** can be used to rigorously assess and improve program efficiency, effectiveness, and quality. This often includes formative or process evaluations that examine the integrity of program implementation, identify areas for program improvement, and determine readiness for more rigorous evaluation. Figure 2 outlines the benefits of program evaluation.
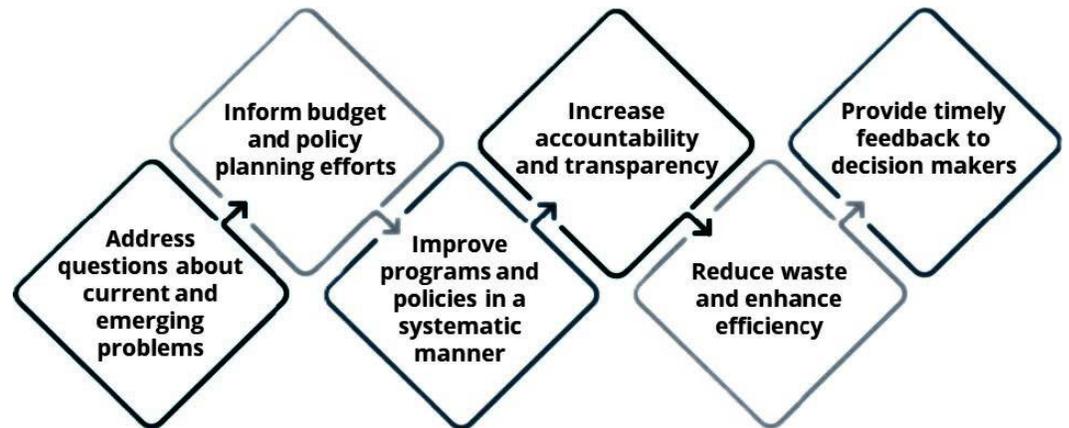

Figure 2: Benefits of Evaluation for the State Government

The State seeks to create an environment where agencies consider evidence-building activities to support innovative practices. When an agency introduces a **new program**, the evidence framework can provide a roadmap for ensuring that the correct measures are considered on the front end so that rigorous evaluation can assess the program's effectiveness. By creating an **evaluation plan** prior to program implementation, agencies proactively take ownership of building evidence right from the start.

OEI can support agencies with identifying specific programs that would benefit from evaluation, consult on evaluation design, and under certain conditions can support agencies by funding evaluations. These activities enable a broader use of evaluation as part of building evidence, further explained below.

Effective Program Evaluation

What is evaluation?

Evaluation is the "assessment[s] using systematic data collection and analysis of one or more programs, policies, and organizations intended to assess their effectiveness and efficiency."[1] An evaluation can assess an entire program or focus on an initiative within a program. Evaluations should follow a systematic and mutually agreed-on plan. **Evaluation plans will typically include the following**:

- Determining the goal: What is the evaluation question? What decisions will the evaluation inform?
- Applying appropriate methods: What quantitative or qualitative methods will be used to increase confidence that the program is having the intended effect?
- Making the results useful: How will the results be communicated so that they can be used by the organization to make improvements?

---

[1] H.R.4174 - Foundations for Evidence-Based Policymaking Act of 2018

What are the types of evaluation?

TYPE	GOALS	KEY FEATURES	ANSWERABLE QUESTIONS
Formative Evaluation	• Assesses whether a program is feasible, appropriate, and acceptable before it is fully implemented. • May be used to learn more about suspected needs and to uncover new needs. • May be used to obtain baseline data that can be monitored for changes over time.	• Can include qualitative and quantitative data collection. • Often does not require advanced statistical methods. • Focuses on learning and improvement and does not answer questions of overall effectiveness.	? Is the program appropriate for the context? ? Does the program feasibly address the identified need? ? Can the program be implemented as designed?
Process or Implementation Evaluation	• Documents what the program is doing. • Documents to what extent and how consistently the program has been implemented as intended. • Informs changes or improvements in the program's operations.	• Does not require a comparison group. • Includes qualitative and quantitative data collection. • Does not require advanced statistical methods.	? How is the program being implemented? ? How do program beneficiaries describe their program experiences? ? What resources are needed for implementing the program?
Outcome Evaluation	• Identifies the results or effects of a program. • Measures program beneficiaries' changes in knowledge, attitude(s), and/or behavior(s) that result from a program.	• Typically requires quantitative data. • Often requires more advanced statistical methods. • May include a comparison group (impact evaluation).	? Did program beneficiaries change their (knowledge, attitude, behavior, or condition) after program completion? ? Did all types of program beneficiaries experience this change or only specific subgroups
Cost Benefit/Cost Effectiveness Evaluation	• Cost-effectiveness analysis (e.g., ROI) assesses the costs of a program to its key outcomes or benefits. • Cost-benefit analysis takes that process one step further, comparing costs with the dollar value of all (or most) of a program's many benefits.	• Primary focus of the evaluation involves program cost.	? Has the program been cost effective (compared to alternatives)? ? What is the most cost-effective option? ? Do the benefits of program X outweigh program costs?
Impact Evaluation	• Assesses the causal impact of a program, policy, or aspect thereof, on outcomes relative to those of a comparison group.	• Impact is defined as a comparison between: What actually happened AND what would have happened, had the program not been introduced (i.e., a counterfactual analysis).	? "Can cause and effect be attributed to the program and its outcomes?" ? What are outcomes for program participants compared to a similar group not in the program?

How do we know when to evaluate and select the right type?

Each type of evaluation serves an important role in evidence-building. The age, logic model, scope, scale, and available data can drive the type of evaluation that can be reasonably pursued. Under the right circumstances, agencies should pursue Impact Evaluations in collaboration with independent evaluators (e.g., university researchers or research organizations). High-quality Impact Evaluations enable the agency to describe the program as falling within the "Evidence" step in the Tennessee Evidence Framework and justify its inclusion in Evidence-Based Budgeting activities.

Guiding Principles for Evaluation

An effective evaluation culture utilizes common principles to increase the likelihood that evaluations are valuable and inform key program decisions. In Tennessee, our guiding principles are rigor, relevance and utility, transparency, independence, and ethics.

Rigor

All evaluations will use methods that generate the highest quality and most credible evidence that corresponds to the questions asked, within 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. The State's evaluations should:

- 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 State's context (external validity); and
- Use measures that accurately (measurement validity) and consistently (measurement reliability) capture the intended information.

Relevance and Utility

Ultimately, evaluations should address questions that are important and provide findings that are actionable in nature. Evaluation designs should consider Governor and Department leadership priorities, legislative requirements, and the needs of stakeholders, including public- and private-sector customers and program administrators. Information resulting from evaluations should be presented in ways that are understandable and promote the use of findings. Evaluations should aim to inform how a program fits within the Tennessee Evidence Framework.

Transparency

To the extent allowable by applicable law, including but not limited to the Tennessee Public Records Act, the State will share evaluation designs, methodologies, analyses, interpretations, and results. Evaluation results related to internal management, legal, risk, or enforcement procedures, or that are otherwise privileged or prohibited from disclosure will not be shared publicly. Evaluation reports will present comprehensive results, including favorable, unfavorable, and null findings.

Independence

Department 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.

Ethics

The State 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.

Appendix A: Terms & Key Concepts

- **Cost-Benefit Analysis**: A method of identifying and comparing relevant quantitative and qualitative costs and benefits associated with a program or activity, usually expressed in monetary terms.

- **Cost-Effectiveness Analysis**: A method of identifying the cost of achieving a single goal, nonmonetary outcome, or objective, which can be used to identify the least costly alternatives for meeting that goal.

- **Evaluation:** Evaluation refers to a systematic assessment of information from one or more programs, policies, and organizations intended to assess their effectiveness and efficiency. Evaluation and Program Evaluation are synonymous.

- **Evaluability Assessmen**t: A pre-evaluation examination of the extent to which a program can be evaluated in a reliable and credible fashion or to which an evaluation is worthwhile based on the evaluation's likely benefits, costs, and outcomes.

- **Evidence-Based Policymaking**: Policymaking which uses the best available research and information on program results to guide decisions at all stages of the policy process and in each branch of government.

- **Evidence:** Evidence is a rigorous body of research that speaks to the efficacy of existing programs or proposed pilots.

- **Impact Evaluation**: Often used for summative purposes, a type of evaluation that focuses on assessing the impact of a program or aspect of a program on outcomes by estimating what would have happened in the absence of the program or aspect of the program.

- **Learning Agenda:** A learning agenda is a plan to identify, prioritize, and establish strategies to develop evidence to answer important short- and long-term questions. These strategic and prioritized questions will enable an agency to build evidence, namely rigorous quantitative estimates of the effectiveness of specific programs.

- **Logic Model:** A logic model is a program planning tool designed to support an intended change; it is a visual representation of how a program works toward a set of outcomes. Logic models are typically written during the planning phase of a program and are often required by funders in grant applications and annual reports.

- **Office of Evidence and Impact (OEI):** A division of the Tennessee Department of Finance and Administration, OEI fosters a culture of data and evidence-based policymaking and budgeting that pursues the best outcomes for Tennesseans. With a vision to empower decision making that serves every Tennessean, OEI uses data to inform decision makers and ensure the State invests in programs that work.

- **Outcomes:** Changes in behavior or benefit, trends over time, contains directionality (e.g., increase, decrease, reduction, growth, maintain, improve, fall, etc.); It could be short-term, mid-, or long-term (system-level) changes. It can also contain the elements of S.M.A.R.T. (specific, measurable, action-oriented, realistic, and time bound).

- **Outcome Evaluation**: Often used for summative purposes, a type of evaluation that assesses the extent to which a program has achieved certain objectives, and how the program achieved these objectives.

- **Process Evaluation**: Often used for formative purposes, an evaluation that assesses the extent to which essential program elements are in place and conform to statutory and regulatory requirements, program design, professional standards, or customer expectations.

- **Program:** Systematic activity that engages participants in order to achieve desired outcomes. Terms often used in place of program can include service, intervention, or practice. If a systematic activity has participants and has its own name, it is likely a program.

- **Program Inventory:** A comprehensive list of programs in the base budget, including whether they are supported by rigorous evidence.

- **Rigorous Evaluations**
  - **Quasi-experimental Studies:** Involve subjects who are not randomly assigned to create the comparison groups. These studies utilize statistical controls to try to create equivalent comparison groups.
  - **Randomized Control Trials:** Involve randomly assigning subjects to treatment and control groups and comparing outcomes of interest.
  - **Systematic Reviews:** Statistical method to synthesize or combine the results from separate studies on a program, policy, or topic in order to estimate its effect on an outcome. These draw on multiple experimental studies to form conclusions and consider the quality of included studies.

- **Tennessee Evidence Framework:** The Tennessee Evidence Framework is designed to standardize the language that is used enterprise-wide to classify programs in Tennessee based on the level of available evidence supporting the program. In this model, we identify five possible evidence steps.