

CHAPTER

1

INTRODUCTION

*The role of the program director is to believe,
the role of the evaluator is to doubt.*

—CAROLE WEISS

LEARNING OBJECTIVES

After reading this chapter, you should be able to

- Explain the primary and secondary goals of the evaluation
- Describe the steps of the evaluation framework

INTRODUCTION

An executive director of a medium-sized, nonprofit social service agency once told me, *“I know we’re doing good—I can see it in our clients’ eyes.”* I wish that this had been a unique encounter, but I have repeatedly had people responsible for running programs and delivering services—including board members, executive directors, and staff—say some version of the above. Sometimes, there is variation—*“Funders wouldn’t fund us if we weren’t doing good,”* or from the funders’ perspective, *“They must have done something good since they spent all of the money in the way that they said they were going to.”* Potential clients may be influenced by the popularity of a program, thinking that people would only flock to impactful programs.

Unfortunately, none of these sentiments tells us what would have occurred to clients if the program hadn’t existed. Clients may have been just as successful, or even more successful, if left to their own devices or if they were exposed to a different program. Or, perhaps clients who seemed not to have changed were actually better off, since had it not been for the program, they would have been worse off.

In all fairness, if those who deliver services didn’t believe that they are effectively improving people’s lives, they might have difficulty devoting themselves to their jobs. In the nonprofit sector, rather than being primarily motivated by wages and compensation, staff are often motivated by their commitment to their organization’s mission and belief in the effectiveness of the initiative on which they spend their time and resources.

An evaluator’s primary job is to examine rigorously the relevance, impact, efficiency, and sustainability of interventions. Evaluators typically ask questions such as:

- To what extent is the intervention the right way to address the problem it aims to address?
- What is the impact of the intervention on its target population?
- Does the intervention produce unintended consequences?
- To what extent is the intervention being delivered efficiently and how can it be delivered more efficiently?
- Should the intervention be continued and if so, how will it be continued? Who will support it?

By considering these broad questions and interacting with the program designers, current and potential beneficiaries, program governance structures (e.g., boards, trustees, and overseers), and staff, evaluators’ insights can result in program improvements for ongoing programs and improvements in program design for initiatives that have not yet started. Note that this set of questions can be asked retrospectively about interventions (e.g., “was the intervention the right way to address the problem?”) or prospectively (e.g., “to what extent does stakeholder feedback and a literature review indicate that the planned intervention will be the right way to address the problem?”).

An evaluator’s secondary job is to encourage program governance, staff, and planners to be more critical of their existing and future initiatives—encourage them to develop a mindset that, rather than assuming that initiatives are effective, instead asks how to make initiatives

more effective and whether an initiative should be continued or (gasp!) abandoned. To design the most impactful programs, boards, planners, and staff need to be more critical of the role that their services (or planned services) play (or will play) in their clients' lives. What assumptions are programs making about clients? Conversely, what do clients and potential clients assume about programs?

Although clients' eyes may reflect gratitude, they are not a substitute for rigorous evaluation. For example, they do not reveal whether a program reaches its intended target population or whether clients would have been just as well off without the program. A critical perspective is needed if the efforts are to continually improve over time.

For-profit organizations have metrics for success that typically involve profit margins, stock prices, and return-on-investment. Determining success for nonprofits (in the international sector referred to as nongovernmental organizations or NGOs) is a more nuanced enterprise—success ultimately rests on the degree to which the organization's mission has been accomplished.

In recent years, funders, management, fundraisers, and potential clients want more information on nonprofits' impacts and efficiencies. Competition for resources in the nonprofit sector is fierce. With the proliferation of nonprofit organizations, questions about which are the best, which are worth investing in, and which are strong and resilient enough to be sustainable become extremely important.

SIZE AND IMPORTANCE OF NONPROFIT SECTOR

The demand for evaluation services in recent years has exploded, not only because nearly all nonprofits and governmental organizations have realized that they need to critically examine their initiatives, but also because of the growth of the nonprofit sector. Consider that between 2010 and 2020, in the United States, the number of nonprofits (defined for this purpose as an organization that filed a required 990 form with the IRS) increased from 186,417 to 217,263 organizations. In 2020, these organizations held \$5.5 trillion in total assets, nearly \$2 trillion in total liabilities, and had total revenues of \$2.7 trillion. They spent \$2.1 trillion on program services.¹ To put this in perspective, consider that the USA Gross Domestic Product (GDP) in 2020 was nearly \$21 trillion. The nonprofit sector's spending on program services accounted for approximately 10% of the nation's GDP.

The nonprofit sector provides services ranging from functions provided by universities, private schools, hospitals, arts and cultural institutions, legal assistance, food assistance, and other safety net and social organizations. Given the magnitude and reach of the nonprofit sector and its importance to the US economy, it is imperative that its entities and funders continue to develop an evaluative mindset. As nonprofits carry out more functions and greater reliance is placed on them, there is increased demand for information about their

¹ *Source:* Author's analysis of data from the Internal Revenue Service, Statistics of Income Division, Exempt Organizations (Except Private Foundations), November 2023.

ability to accomplish their missions efficiently and their impact on the persons, entities, and/or communities they serve.

In recent years, the role of government versus the nonprofit sector has notably changed. The US government is relying more on nonprofits to provide assistance, sometimes with governmental funding and sometimes without. Government increasingly relies on the nonprofit sector, with its army of volunteers and relatively low-paid staff, to compensate for gaps in governmental programs. This becomes painfully evident during the aftermath of natural disasters, when local food banks and feeding centers run by nonprofits are asked to provide food, the American Red Cross provides shelters, and other nonprofits step up in other ways to serve populations in crisis. But even when there is not an acute emergency, the nonprofit sector is increasingly comprising the safety net for lower-income persons. Rather than providing services directly, federal, state, and local governments increasingly contract with nonprofits to provide services.

With more reliance placed on the nonprofit sector comes the demand for all nonprofit organizations and especially those that compete with other nonprofits for “market share” to show their effectiveness. Government, consumers, and donors would like to know which nonprofits are most effective at providing comparable services. In this respect, evaluation is viewed as essentially a means of providing information to rank programs and organizations. Evaluation reports are used to answer the questions “*which agency should a government contract for a needed service?*” and “*which nonprofits should foundations and individual donors fund?*”

This book is intended to provide leaders, program creators, and program implementers in the nonprofit/NGO/governmental arena with enough knowledge of the evaluation field to empower them to engage with evaluation activities in meaningful ways. Many nonprofit leaders lack training in the nuances of evaluation research. Nonprofit leaders can come to their positions in a variety of ways. For example, the Executive Director of a nonprofit organization that provides guardianship services to the elderly and/or disabled was a noted accountant, but never received any training in evaluation. The same goes for the Executive Director of a local food bank who had been a passionate anti-hunger advocate and the Executive Director of a local library who excelled as a head librarian.

Many influential foundations lack an evaluation policy that would speak to the size of investments that must be evaluated and the types of evaluation activities that would occur when considering whether to fund an initiative. Some leaders and staff of foundations who have decided that evaluation is important lack the background to determine fair criteria by which to judge and reward effective nonprofit organizations and programs.

Working in the background behind service deliverers and funders are people who call themselves evaluators. Many evaluators are primarily trained in fields other than evaluation, such as public health, sociology, economics, statistics, or public policy. They may have had some formal training in evaluation, but they often have moved into evaluation over time and many have acquired only on-the-job training. For this reason, although the majority of evaluators have in common a desire to learn the “truth” about interventions, they bring to the table a divergent set of skills and may place varying emphases on different aspects of programs, policies, and evaluatory practices.

My perspective is aligned with that of Michael Quinn Patton’s (1996)—the goal of evaluation is to assist with continuous programmatic improvement and introspection. I readily

acknowledge that how to achieve this goal is an art informed by rigorous social science methods. There are tools that the evaluator needs to have mastered and with these tools, the evaluator must create the *least intrusive, least burdensome, and least costly evaluation plan* that will reveal the relevance, impact, efficiency, and sustainability of an initiative. To do so, the evaluator needs to understand the initiative's context and develop knowledge of how others have approached the issue that the initiative aims to address.

THE EVALUATION FRAMEWORK

The evaluative framework for conducting an evaluation that examines an initiative's effects is a 10-step process in which the evaluator and the "evaluand" (the entity, program, or initiative being evaluated) should engage.

1. **Determine the audience and accountability structure for the evaluation activities.**

Who is the audience for evaluation activities? To whom will the evaluator be reporting? Who will receive reports or memos from the evaluator? Who will endorse the evaluation recommendations? Who will be responsible for ensuring that recommendations are implemented?

I encourage clients to create an Evaluation Reference Group (ERG) consisting of program administrators, funders, and program beneficiaries. Substantive interactions that the evaluator has with the "client" go through the ERG. The ERG ensures that the evaluation answers relevant and important questions. The ERG includes people who will assist the evaluator in getting access to people and materials during the data collection process and helps bring attention to the evaluation process and final evaluation report. The ERG provides an opportunity for the representatives of the community that the initiative is to benefit to participate in the evaluation process.

2. **Determine the universe of possible evaluation questions.**

With most initiatives, a nearly infinite set of evaluative questions could be asked. However, there is never enough time or resources to answer them all. When listing the universe of evaluation questions, the evaluator should map out the stakeholders who comprise the audience for each question. Broadly, the questions will pertain to the program's relevance, effectiveness, efficiency, and sustainability.

3. **Rigorously describe the program.**

The evaluator will describe the initiative by reading its documents, learning the history of the program verbally, and with a group of stakeholders diagramming its *Theory of Change* (ToC) (also known as the *program theory*) and charting out its *program logic model*. These two tools not only force evaluators to become acquainted with the program (theoretically and operationally) but also get the stakeholders in agreement about the program's description and goals. Any evaluation that is done without a thorough understanding of the program can be of no constructive use. For this reason, this book devotes considerable space to describing programs and initiatives.

THE EVALUATION FRAMEWORK

1. Determine the audience and accountability structure for the evaluation activities.
2. Determine the universe of possible evaluation questions.
3. Rigorously describe the program by developing the program theory and using a program logic model.
4. Conduct a literature review around the Theory of Change and to learn about similar initiatives.
5. Revisit and narrow the universe of evaluation questions.
6. Develop the evaluation plan.
7. Carry out the evaluation plan.
8. Report on draft evaluation findings and recommendations to the Evaluation Reference Group.
9. Finalize the evaluation report.
10. An accountability mechanism (e.g., board) endorses the final evaluation report's recommendations and ensures that recommendations are implemented.

4. **Conduct a literature review around the Theory of Change and to learn about similar initiatives.**

The literature review will inform as to the certainty of the ToC's causal connections actually occurring, once the initiative's activities occur. The literature review will also inform as to the experience of similar programs—their successes, impact, and challenges. The literature review will reveal the measures that other programs have used when considering impact. Essentially, the literature review will allow the evaluator to assess the veracity of the ToC and allow the evaluator to focus data collection efforts on the areas of the ToC that are not certain to occur.

5. **Revisit and narrow the evaluation questions.**

After rigorously describing the program, the evaluator and ERG should revisit the universe of evaluation questions and narrow them. It is unlikely that any evaluation will have the time and resources to answer all possible questions. Thus, the universe of evaluation should be narrowed according to the resources available for the evaluation, the timeline of the evaluation, the audience for each evaluation question, and the evaluator's access to the information that would allow for the evaluation question to be answered.

The ERG must be involved in narrowing the evaluation questions and developing an understanding of the limitations of the evaluation. The group should be open to the idea that outcome-based evaluation is not the only type of evaluation that could be conducted. Examining the processes in implementing an initiative can also have value. In fact, for unstable programs, process-based rather than outcome-based evaluation will be appropriate.

In narrowing the evaluation questions, thought should be given to who the audience is for each evaluation question and whose interest each question represents. If a group's interest is left out of the list of narrowed questions, the evaluator should encourage the ERG to discuss this omission so that they realize and appreciate the evaluation's focus and direction.

6. **Develop an evaluation plan.**

The evaluation plan, often called an **"Inception Report,"** will cover the evaluation questions that the evaluation will answer and include the data collection and data analysis plan. Without a developed blueprint for the evaluation, there is a risk that the evaluation activities will continue *ad infinitum*. Before the evaluation team proceeds with the evaluation activities, the ERG must provide feedback on the Inception Report and the evaluator should come to consensus with the ERG on how the evaluation will be conducted.

The evaluation plan will include a data collection and data analysis plan. If qualitative data will be collected, the inception report will indicate details such as the number of focus groups and/or interviews that will be conducted, who will be included in focus groups and interviews, and where and when they will occur. The report will include draft interview and focus group protocol. If data collection includes conducting a survey, the inception report will indicate who will be surveyed, when the survey will occur, how data from the survey will be used, and will include a draft survey instrument. The inception report will also identify any secondary (pre-existing) data sources (e.g., surveys conducted by others, census data) that will be relied upon and the analysis approach.

The inception report will include the **quasi-experimental design(s)** that the evaluation will rely upon. The report will show how the data collected and the data analysis techniques proposed will be used to make the argument regarding the program's role in causing any changes that occurred.

The report will also include a timeline for the evaluation activities. This timeline should be reasonable and have built into it time for the ERG to review materials.

It is important that the inception report address the limitations of data collection and of the evaluation, if there are any such limitations. It is better for the ERG to be aware and understand the limitations at the outset rather than after the evaluation activities are complete.

7. **Carry out the evaluation plan.**

In carrying out the evaluation plan, the ERG should be kept apprised of the unexpected changes to and/or deviations from the plan. It would be wise to have regular meetings with the ERG (e.g., biweekly) to discuss progress.

8. **Report on draft evaluation findings and recommendations to the Evaluation Reference Group.**

Once the evaluation team has completed data collection, the team should meet with the ERG and present tentative evaluation results and recommendations to it. This meeting and follow-up to it becomes the ERG's opportunity to provide feedback before the report is written. Such a discussion can be very fruitful and can save time (and frustration) in the report finalization process.

9. **Finalize the evaluation report.**

The evaluator should provide the ERG a written draft evaluation report and give the ERG an opportunity to circulate the draft report among colleagues. The evaluator should ask for feedback on the draft report, especially feedback that corrects (or updates) facts in the report. Remember that errors in the report could impact reliance on recommendations. An important part of the finalization process is bringing attention to the ERG and implementers findings and results in the draft report. Sometimes, there will be an objection to a recommendation included in the draft. The evaluator may decide that there are issues that once raised in a draft report, do not need to be included in the final report.

10. **The Evaluation Reference Group should ensure that an accountability mechanism endorses the final evaluation report's recommendations.**

After the report is finalized, the evaluator's role in a retrospective evaluation is done. It is then up to the evaluand and/or its governance, with the ERG, to act on the evaluation's recommendations and to communicate the evaluation's results. For example, the ERG could bring the evaluation report to the attention of the organization's board for the endorsement of recommendations and press for there to be a follow-up mechanism on the recommendations. The organization could place the evaluation report on its website, circulate its Executive Summary, and offer a public session regarding the evaluation.

Separating the evaluation activities into these ten discrete "tasks" allows evaluators to pause and to think about the direction that their activities will take and how they will interact with the evaluand.

High-quality evaluations can improve ongoing initiatives and inform the field about lessons learned from similar initiatives. When planning initiatives, planners have an obligation to review the evaluations conducted on similar initiatives so that planners can create the most impactful, relevant, and efficient initiatives. Funders have an obligation to support and read evaluation reports so that their resources can be put to the best use. The evaluations of an initiative should eventually become a part of the base of knowledge about a field of initiatives.

STRUCTURE OF BOOK

The remainder of this book is devoted to more fully explaining each of these 10 steps. Regardless of whether you will conduct evaluations yourself or contract with evaluation consultants or firms, these 10 steps must be understood.

The first half of the book focuses on developing the appropriate background for evaluation and the second half addresses the use of quasi-experimentation and data collection in an evaluation context. Unlike many evaluation and economics texts, this one emphasizes that one must rigorously describe the program prior to evaluating it.

Chapter 2 introduces the Program Logic Model and Theory of Change as tools to use when describing the program. The chapter closes with examples of these tools applied to various initiatives, hoping to make you comfortable enough with these tools so that you will be able to apply them to one of your own initiatives.

Chapter 3 discusses the framing of evaluation questions, elaborating upon the pre-evaluation steps that one must consider before embarking on conducting the evaluation activities. Since the ultimate goal of initiatives is to cause positive change, Chapter 4 delves into the meaning of causation in an evaluation context.

To engage in high-quality evaluation activities, one must understand what it means to do valid evaluation research. There are four types of validity and Chapter 5 explains what these are, their importance when conducting literature reviews, and their importance when creating a robust evaluation plan.

Chapter 6 introduces the reader to the quasi-experimental designs used and the trade-offs made between quasi-experimental designs and types of validity. Chapter 7 presents approaches to data collection—interviews, focus groups, surveys, and pre-existing data (archival, administrative, and census data). While the chapter does not substitute for a good course in statistics, it discusses the areas particularly relevant to evaluation research.

Chapter 8 concludes the text, discussing the similarities between grant proposals and evaluation plans. The chapter also includes a template for a Terms of Reference to use when hiring an evaluator or evaluation firm.

SUMMARY

The primary job of the evaluator is to examine the impact of program interventions. The secondary job is to provide a critical perspective to improve programs. The goal of evaluation is to assist with continuous program improvement and introspection. To successfully evaluate a program, the evaluator must follow the steps of the evaluation framework.

10 Evaluation Essentials

KEY TERMS

evaluation

evaluators

formative evaluations

program logic model

program theory

quasi-experimental design

summative evaluations

target population

DISCUSSION QUESTIONS

1. What are the goals of program evaluation?
2. Why is it important to have an evaluation plan?
3. What is the evaluation framework?