

# PART I

# INTRODUCTION

Cambridge University Press  
978-1-108-83599-2 — Pragmatic Program Evaluation for Social Work  
Allen Rubin  
Excerpt  
[More Information](#)

---

# Chapter 1

## Introduction and Overview

### CHAPTER OUTLINE

- 1.1 Introduction 4**
- 1.2 Why Evaluate? 4**
- 1.3 Some Programs are Ineffective or Harmful 5**
  - Critical Incident Stress Debriefing 5
  - Scared Straight Programs 6
- 1.4 Historical Overview of Program Evaluation 7**
- 1.5 Evidence-Informed Practice 8**
- 1.6 Philosophical Issues: What Makes Some Types of Evidence Better Than Other Types? 9**
  - Contemporary Positivism 9
  - Interpretivism 10
  - Empowerment 10
  - Constructivism 10
- 1.7 Qualitative versus Quantitative Evaluations: A False Dichotomy 12**
- 1.8 Definitions 13**
- 1.9 Different Evaluation Purposes 14**
- 1.10 Types of Evaluation 15**
  - Summative Evaluation 15
  - Formative Evaluation 16
  - Process Evaluation 16
  - Performance Measurement Systems 18
  - Evaluating One's Own Practice 19
  - Accreditation 19
- 1.11 Chapter Main Points 20**
- 1.12 Exercises 22**
- 1.13 Additional Reading 22**

## 4 Introduction and Overview

### WHAT YOU'LL LEARN IN THIS CHAPTER

In this first chapter you'll get an overview of program evaluation and a foundation for reading the rest of the book. You'll learn about the history of program evaluation and the various reasons to evaluate, including the field's finding out that some programs that seemed like good ideas at the time turned out to be ineffective or harmful. Program evaluation is largely about obtaining evidence to inform program activities and decisions, and therefore it is particularly germane to the current era of evidence-informed practice. Consequently, you'll learn about the process of evidence-informed practice. You'll also learn about philosophical issues that bear on the quality of evidence. The chapter provides definitions of some key terms in the field of program evaluation – terms that appear throughout this book. The chapter will conclude by discussing different purposes and types of evaluation.

## 1.1 Introduction

---

If you are like most of the students taking a course on program evaluation, you are not planning to be seeking a position as a program evaluator. You might even doubt that you will ever help to plan or conduct an evaluation. In fact, you might be reading this book only because it is a required text in a course you had to take, but did not want to. You probably are looking forward to a career in an agency that prioritizes service delivery and that views program evaluation as one of its lowest priorities. Well, you are likely to be surprised. Even in service-oriented agencies that do not prioritize program evaluation, the need to evaluate often emerges. For example, questions might arise as to whether the agency's target population has unmet needs that the agency should address. Answering questions like those might indicate the need to conduct a needs assessment evaluation. (Needs assessment evaluations will be discussed in Chapter 3.) Perhaps some new services or treatment modalities should be developed to meet those needs. Perhaps the ways some existing services are provided need to be modified to make them more accessible to current or prospective clients. If the agency's caseload has recently experienced a large influx of clients from a minority culture, for example, there may be a need to evaluate the cultural sensitivity of agency practitioners as well as support staff (i.e., receptionists, intake interviewers, etc.).

## 1.2 Why Evaluate?

---

The most common impetus to evaluate often involves funding – either persuading an existing funding source to continue or to increase its funding or convincing a new funding source to fund a new service initiative. For example, early in

### 1.3 Some Programs are Ineffective or Harmful 5

her career one of my former students – let’s call her Beverly – was co-leading a community-based support group for women who were being verbally abused by their partners. The group was sponsored by a battered women’s program that was not well funded. Beverly strongly believed in the effectiveness of her support group and wanted to conduct an evaluation of its effectiveness in order to seek external funding for it so that it would not draw from the limited resources of the larger program. Another former student – we’ll call her Jennifer – was a clinical supervisor in a child welfare agency. Her administrator asked her to write a proposal for funding from a new federal family preservation initiative. One requirement for funding was including a design to evaluate the impact of the proposed program. With my help Jennifer was able to design the evaluation and secure funding. However, I hope in your career your motivation for evaluating the impact of your agency’s services is not based exclusively on funding considerations. Your motivation should also stem from your compassion for clients and your professional ethics, both of which should impel you to want to find out if your services are really helping – and not harming people.

### 1.3 Some Programs are Ineffective or Harmful

---

It might seem far-fetched to you to suppose that the services you believe in and are devoting your work to every day are not helping people and much more far-fetched to imagine that they possibly could be harming people. But you might be surprised to learn that some programs that seemed to be well grounded in theory and that were widely embraced by leading theorists and practitioners in the helping professions were found to be ineffective. Some were even found to be harmful. Let’s take a look at two of them now.

**Critical Incident Stress Debriefing.** Critical incident stress debriefing (CISD) has been one of the most popular group modalities of crisis intervention for survivors of mass traumatic events such as natural disasters. In it, the group participants discuss the experiences they had during the traumatic event in small details such as where they were when the disaster occurred, what they saw, what they heard, how they felt, and so on. Current emotions and physical symptoms are also discussed. The group leaders normalize those reactions to reassure the participants that such reactions are commonly experienced by survivors and do not mean that there is something abnormal about them per se. The leaders also provide advice regarding what things survivors should or should not do to help alleviate their (normal) distress (Housley & Beutler, 2007). That all sounds pretty reasonable, doesn’t it? It sure does to me! Indeed, the CISD approach incorporates various generic principles of trauma treatment (Rubin, 2009). Well, here’s

## 6 Introduction and Overview

a surprise. Multiple evaluations have concluded that CISD is not helpful and can actually be harmful in that it slows down the normal recovery process over time for many trauma survivors. Speculations regarding why CISD is harmful note that people who are not particularly vulnerable to longer-term posttraumatic stress disorder (PTSD) (which includes approximately 70 percent of natural disaster survivors – those who will not develop PTSD) would benefit sooner from the healing effects of time if the CISD did not engage them in imagining a reliving of the traumatic experience (Bisson *et al.*, 1993; Carlier *et al.*, 1998; Mayou *et al.*, 2000; Rose *et al.*, 2002). But until the evaluations were completed such harmful effects were not anticipated, and as recently as 2018 it was portrayed as helpful on the *WebMD* Internet website.

**Scared Straight Programs.** During the 1970s a program was introduced that aimed to prevent criminal behavior by juveniles. It brought youths who already had been adjudicated as criminal offenders as well as other youths thought to be at risk of becoming offenders into prisons where convicts would describe the sordid nature of life in prisons and thus attempt to scare the youths so much that they would be turned away from criminal behavior. The program, called *Scared Straight*, seemed to make sense. It became very popular and was portrayed in an Oscar-winning film in 1979. The film’s narrator claimed that the program was successful in scaring juveniles “straight.” The narrator was wrong. In fact, an evaluation at that time found that the offenders who participated in the program ended up committing more crimes than did similar offenders who did not participate. Despite its reasonable-sounding premise, therefore, the program was not only ineffective – it was harmful. In trying to explain why it was harmful, the author of the evaluation speculated that rather than scare the “tough” youths it motivated them to commit more crimes to prove that they were not afraid (Finckenauer, 1979).

Thus, among the various reasons to conduct program evaluations, perhaps the most compelling one is our desire to help people coupled with our awareness that some programs that sound great in theory – even some that have been widely embraced by experts – can be ineffective and even harmful. Our desire to help people can motivate us to conduct evaluations for other reasons, as well – reasons that can bear on program effectiveness but do not examine effectiveness *per se*. These reasons – or evaluation foci – pertain to obtaining answers to questions that can help us improve the way our program is being implemented. One such question is “What unmet needs of our clients or prospective clients should we be addressing?” Another pertains to whether the agency’s staff members are implementing the program appropriately and whether there are any problems in how they are performing their roles. Box 1.1 lists some of the most common reasons for evaluating programs.

**Box 1.1 Reasons to Evaluate****Required for funding**

- To seek new or continued funding for agency
- To seek new or continued funding for a new unit or intervention

**To answer these questions**

- 1 Are we really helping (harming?) people? (Are we effective?)
- 2 How can we improve the program?
  - What needs of our clients or prospective clients should we be addressing?
  - Are staff members implementing the program as intended?
  - Are clients satisfied with our program, and if not, why not?

**1.4 Historical Overview of Program Evaluation**

Although some have traced program evaluation as far back to 2200 BC in China (Shadish *et al.*, 2001), it emerged in the United States early in the twentieth century with evaluations of the effectiveness of alternative educational approaches that compared outcomes on student standardized test scores. During the ensuing early decades of the century, as concerns emerged regarding the impact of industrialization, evaluations examined worker morale and the impact of public health education on personal hygiene practices. The growth of evaluation accelerated with the emergence of New Deal social welfare programs during the 1930s, and included evaluations of the effectiveness of financial relief policies, public housing, and programs to combat juvenile delinquency. After World War II that growth accelerated even more as public expenditures increased to alleviate problems in public health, housing, family planning, community development, and juvenile delinquency. Rubin & Babbie (2017) note that “by the late 1960s, textbooks, professional journals, national conferences, and a professional association on evaluation research emerged” (p. 322). Interest in program evaluation continued to accelerate during the 1970s, but less so in response to the development of new programs and more in response to increasing skepticism being expressed by conservative pundits and politicians, and consequently the public overall, regarding whether the (dubious) effects of existing public programs were worth the amount of public expenditures being invested in them.

## 8 Introduction and Overview

Growing public skepticism about the value of social welfare and human service programs, coupled with a trend toward a more conservative electorate in the 1980s and 1990s, resulted in substantial funding cuts to those programs and put more pressure on programs to complete evaluations in the hope of obtaining results that would justify their proposals for renewed funding. That pressure intensified the political context of program evaluation, as program personnel wanted their programs to be evaluated in ways that maximized the likelihood of obtaining findings that would portray their programs as effective and therefore worthy of more funding. (Chapter 9 will discuss the political context of program evaluation.)

Another major historical development affecting demands for program evaluation during the late twentieth century was the emergence and popularity of the concept of managed care. This concept emphasized various ways to control the rising costs of health and human services. One way to control those costs is to fund only those programs or treatments that have had their effectiveness supported by the findings of very rigorous experimental outcome evaluations that randomly assign participants to treatment and control groups – evaluation designs known as randomized control trials (RCTs).

A similar influence on evaluation during the first decades of the twenty-first century was the growth of the evidence-based practice (EBP) movement, as leading thinkers in the helping professions promoted the idea that practice decisions about programs, policies, and interventions should take into account the best research and evaluation evidence regarding which policies, programs, and interventions are most effective and how best to implement them. The EBP concept, however, was not without controversy, as some opposed it on the grounds that it was nothing more than a way for managed care efforts to reduce costs by reducing the range of efforts to help people that would be funded. While recognizing that implementing the most effective policies, programs, and interventions offers the opportunity to do more with less, proponents of EBP countered that its priority was not cost saving and pointed out its consistency with the professional ethic of seeking to ensure that clients received what is known to be effective and not harmful. This book is being written during the evidence-informed practice era when program evaluation – despite historical fluctuations influencing its prevalence and nature – continues to be an important concern among human service programs and agencies, both for utilitarian reasons and for humanitarian reasons.

### 1.5 Evidence-Informed Practice

---

The current era also is a time when evidence-informed practice is being emphasized in the educational preparation of social workers and other human service professionals (CSWE, 2015). The term *evidence-informed* practice has replaced



the term *evidence-based* practice in response to critics who misconstrued the use of the word *based* to mean that evidence should be the *only* thing that informs practice decisions. The original term, *evidence-based practice*, was never meant to imply that practice decisions should be determined only by research evidence. Those who coined the term, as well as its early pioneers (Sackett *et al.*, 1997; Gambrill, 1999; Gibbs & Gambrill, 2002), emphasized that the evidence-based practice model recognized that practice decisions should take into account *all* of the following considerations:

- Client characteristics, needs, values, and treatment preferences
- Practitioner expertise and treatment resources
- The environmental and organizational context
- The best available research evidence

Five steps were proposed for the model, as follows.

1. Formulate a question concerning practice needs.
2. Search for evidence to inform the answer to that question.
3. Critically appraise the methodological quality of the research supplying the evidence.
4. Answer the posed question in light of the best available evidence as well as client characteristics, needs, values, and treatment preferences; practitioner expertise and resources; and the organizational context.
5. Implement the action (intervention) implied in step 4.
6. Evaluate the action taken in step 5 and provide feedback based on the results of the evaluation.

## 1.6 Philosophical Issues: What Makes Some Types of Evidence Better Than Other Types?

---

Not all evaluators or leading thinkers agree about what makes some types of evidence better than other types. Leading thinkers representing different schools of thought – or *paradigms* – have argued about this question for decades. Three main paradigms are currently prominent regarding evidence in social work and the human services: contemporary positivism, interpretivism, and the empowerment paradigm.

**Contemporary Positivism.** Contemporary positivists emphasize objectivity and precision in measurement and the logic of causality in designing evaluations and appraising the quality of evidence. They recognize the difficulty of being purely objective and of conducting flawless evaluations whose findings and methods

## 10 Introduction and Overview

are immune from warranted criticism. They tend to notice the imperfections in any individual evaluation and see all forms of research (evaluation included) as an eternal and self-correcting quest for better evidence that involves a replication process in which any particular evaluation is followed by subsequent evaluations to see if the conclusions of the previous evaluation(s) are upheld or merit modification or rejection.

**Interpretivism.** Interpretivists do not focus on maximizing objectivity in measurement or on logically isolating the causes of social phenomena. Instead, they emphasize probing for a deeper, subjective understanding of people's experiences, the meanings and feelings connected to those experiences, and the idiosyncratic stated reasons for their behaviors. They believe that the best way to evaluate what people need and how well services meet their needs is to take a flexible and subjective approach to evaluation that seeks to discover how people experience things on an internal and subjective basis.

**Empowerment.** The empowerment paradigm puts less emphasis on the priorities of contemporary positivists and interpretivists and more emphasis on using evaluation to empower oppressed people. Thus, for adherents of this paradigm, the best evidence is that which best empowers people and advances social justice. Evaluators influenced primarily by the empowerment paradigm might use evaluation methods favored by contemporary positivists or methods favored by interpretivists, and the choice of which to use will depend largely on which are most likely to produce findings that are consistent with their empowerment and advocacy aims. One type of evaluation design geared specifically to the empowerment paradigm is *participatory action evaluation*. In this design the evaluation participants are members of disadvantaged or oppressed populations. Those participants make decisions about the goals and design of the evaluation. The evaluation's priority is not only to produce its ultimate findings, but also as a means of educating the participants, raising their consciousness about social injustice, and mobilizing them for action. Box 1.2 summarizes a participatory action evaluation that involved university students who were welfare recipients. The box also illustrates how an evaluation can focus on social policy.

**Constructivism.** Another paradigm, one that shares commonalities with the interpretivist and empowerment paradigms, is called *constructivism*. This paradigm has different versions, or types. Its most extreme version rejects the notion of absolute truths, or of an objective social reality, and argues that people only