

CHAPTER 11

Evaluation Research

What Is the History of Evaluation Research?

What Is Evaluation Research?

What Are the Alternatives in Evaluation Designs?

Black Box or Program Theory
Researcher or Stakeholder Orientation
Quantitative or Qualitative Methods
Simple or Complex Outcomes

What Can an Evaluation Study Focus On?

Needs Assessment
Evaluability Assessment
Process Evaluation
Impact Analysis
Efficiency Analysis

Ethical Issues in Evaluation Research

Conclusion

Drug Abuse Resistance Education (D.A.R.E.), as you probably know, is offered in elementary schools across America. For parents worried about drug abuse among youth and for any concerned citizens, the program has immediate appeal. It brings a special police officer into the schools once a week to talk to students about the hazards of drug abuse and to establish a direct link between local law enforcement and young people. You only have to check out bumper stickers or attend a few PTO meetings to learn that it's a popular program. It is one way many local governments have implemented antidrug policies.

And it is appealing. D.A.R.E. seems to improve relations between the schools and law enforcement and to create a positive image of the police in the eyes of students.

It's a very positive program for kids . . . a way for law enforcement to interact with children in a nonthreatening fashion . . . DARE sponsored a basketball game. The middle school jazz band played. . . We had families there. . . DARE officers lead activities at the [middle school]. . . Kids do woodworking and produce a play. (Taylor, 1999:1, 11)

For some, the positive police-community relationships created by the program are enough to justify its continuation (Birkeland, Murphy-Graham, & Weiss, 2005:248), but most communities are concerned with its value in implementing antidrug policies. Does

D.A.R.E. lessen the use of illicit drugs among D.A.R.E. students? Does it do so while they are enrolled in the program or, more important, after they enter middle or high school? Evaluations of D.A.R.E. using social science methods led to the conclusion that students who participated in D.A.R.E. were no less likely to use illicit drugs than comparable students who did not participate in D.A.R.E. (Ringwalt et al., 1994).

If, like us, you have a child who enjoyed D.A.R.E., or if you were yourself a D.A.R.E. student, this may seem like a depressing way to begin a chapter on evaluation research. Nonetheless, it drives home an important point: To know whether social programs work, or how they work, we have to evaluate them systematically and fairly, whether we personally like the programs or not. And there's actually an optimistic conclusion to this introductory story: Evaluation research can make a difference. A "new" D.A.R.E. program has now been implemented, building in part on the problems identified by early evaluation researchers (Toppo, 2002).

Gone is the old-style approach to prevention in which an officer stands behind a podium and lectures students in straight rows. . . . New DARE officers are trained as "coaches" to support students using research-based refusal strategies in high-stakes peer-pressure environments. (McConnell, 2006:33)

And so, of course, this approach is now being evaluated, too.

In this chapter, you will read about a variety of social program evaluations, alternative approaches to evaluation, and the different types of evaluation research and review ethical concerns. You should finish the chapter with a much better understanding of how the methods of applied social research can help improve society.

WHAT IS THE HISTORY OF EVALUATION RESEARCH?

Evaluation research is not a method of data collection, like survey research or experiments; nor is it a unique component of research designs, like sampling or measurement. Instead, **evaluation research** is conducted for a distinctive purpose: to investigate social programs (such as substance abuse treatment programs, welfare programs, criminal justice programs, or employment and training programs). For each project, an evaluation researcher must select a research design and method of data collection that are useful for answering the particular research questions posed and appropriate for the particular program investigated.

Evaluation research: Research that describes or identifies the impact of social policies and programs

So you can see why we placed this chapter after most of the others in the text. When you review or plan evaluation research, you have to think about the research process as a whole and how different parts of that process can best be combined.

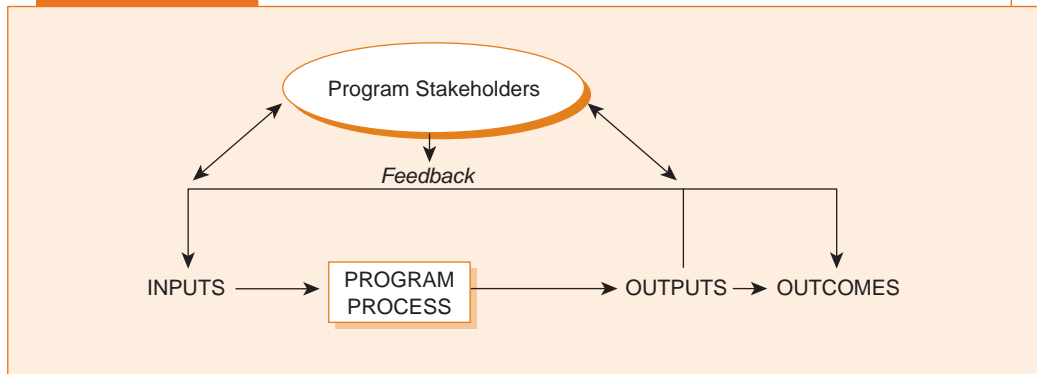
The development of evaluation research as a major enterprise followed on the heels of the expansion of the federal government during the Great Depression and World War II. Large Depression-era government outlays for social programs stimulated interest in monitoring program output, and the military effort in World War II led to some of the necessary review and contracting procedures for sponsoring evaluation research. However, not until the Great Society programs of the 1960s did evaluation begin to be required when new social programs were funded (Rossi & Freeman, 1989:34). The World Bank and International Monetary Fund (IMF) began to require evaluation of the programs they fund in other countries (Dentler, 2002:147). More than 100 contract research and development firms began in the United States between 1965 and 1975, and many federal agencies developed their own research units. The RAND Corporation expanded from its role as a U.S. Air Force planning unit into a major social research firm; SRI International spun off from Stanford University as a private firm; and Abt Associates in Cambridge, Massachusetts, which began in a garage in 1965, grew to employ more than 1,000 employees in five offices in the United States, Canada, and Europe.

With the decline of many Great Society programs in the early 1980s, many such evaluation research firms closed down. But recently, with more calls for government “accountability,” the evaluation research enterprise has been growing again. The Community Mental Health Act Amendments of 1975 (Public Law 94–63) required quality assurance (QA) reviews, which often involve evaluation-like activities (Patton, 2002: 147–151). The Government Performance and Results Act of 1993 required some type of evaluation of all government programs (Office of Management and Budget, n.d.). At century’s end, the federal government was spending about \$200 million annually on evaluating \$400 billion in domestic programs, and the 30 major federal agencies had between them 200 distinct evaluation units (Boruch, 1997). Recently, the new Governmental Accounting Standards Board urged that more attention be given to service efforts and accomplishments in standard government fiscal reports (GASB, 2008).

The growth of evaluation research is also reflected in the social science community. The American Evaluation Association was founded in 1986 as a professional organization for evaluation researchers (merging two previous associations) and is the publisher of an evaluation research journal. In 1999, evaluation researchers founded the Campbell Collaboration to publicize and encourage systematic review of evaluation research studies. Their online archive contains 10,449 reports on randomized evaluation studies (Davies, Petrosino, & Chalmers, 1999).

WHAT IS EVALUATION RESEARCH?

Exhibit 11.1 illustrates the process of evaluation research as a simple systems model. First, clients, customers, students, or some other persons or units—cases—enter the program as **inputs**. (Notice that this model regards programs as machines, with clients—people—seen as raw materials to be processed.) Students may begin a new school program, welfare

EXHIBIT 11.1 A Model of Evaluation

recipients may enroll in a new job-training program, or crime victims may be sent to a victim advocate. Resources and staff required by a program are also program inputs.

Inputs: Resources, raw materials, clients, and staff that go into a program

Next, some service or treatment is provided to the cases. This may be attendance in a class, assistance with a health problem, residence in new housing, or receipt of special cash benefits. This process of service delivery—the **program process**—may be simple or complicated, short or long, but it is designed to have some impact on the cases as inputs are consumed and outputs are produced.

Program process: The complete treatment or service delivered by the program

Program **outputs** are the direct product of the program's service delivery process. They could include clients served, case managers trained, food parcels delivered, or arrests made. The program outputs may be desirable in themselves, but primarily they indicate that the program is operating.

Outputs: The services delivered or new products produced by the program process

Program **outcomes** indicate the impact of the program on the cases that have been processed. Outcomes can range from improved test scores or higher rates of job retention to fewer criminal offenses and lower rates of poverty. There are likely to be multiple outcomes of any social program, some intended and some unintended, some viewed as positive and others viewed as negative.