sider these issues later in some detail, but for now let us begin by exploring the two main types of evaluation research and the evaluator’s role.

THE EVALUATOR’S ROLE

The evaluator’s avowed purpose may be FORMATIVE (trying to improve the program) or SUMMATIVE (rendering a judgment regarding the program’s mission and/or effectiveness) (Chambers, 1994; Posavac and Carey, 1997:14). In each of these two general roles, the evaluator typically asks a different range of research questions and uses a distinctive methodological toolkit.

Formative Evaluation

In formative research there are three major questions:

- What is the definition and scope of the problem or issue?

To answer this question the researcher works together with clients to elicit ideas, for example, by using focus groups or stakeholder analysis. Conceptualizing methods, such as brainstorming or creating visual maps of ideas and their connections to each other (Deshler, 1990), are often employed in this type of exercise.

- Who needs the program, how great is the need, how big or serious is it?

To answer these questions a NEEDS ASSESSMENT — a comprehensive evaluation of the demand for some new program or service — is performed. Need assessments typically rely on existing data sources, surveys, and in-depth interviews.

- How is the program being run, and how can program delivery be improved?

These questions are answered via PROCESS EVALUATION, which investigates the actual implementation of a program, including possible alternative delivery procedures. Some of the methods appropriate to process evaluation

---

3 See the sections on “Focus Groups” and “Sampling” in this chapter.
are multivariate statistics and causal modeling, simulation and gaming techniques, and the examination of organizational flowcharts and project schedules as well as the lines of authority in decision making. Management information systems are an ideal source of data for process evaluation.

**Summative Evaluation**

In summative evaluation there are two major questions:

- What is the effectiveness of the program?

To answer this question one may select observational methods, or perform statistical correlations to demonstrate whether desired effects occurred. Quasi-experimental and experimental designs can establish whether observed effects can reasonably be attributed to the intervention and not to other sources\(^4\) (Trochim, 2002).

- What is the net impact of the program?

**Impact Evaluation** is typically more detailed and covers a longer time period than just measuring program effectiveness. It looks at both intended and unintended consequences of the whole program. For example, we might measure the effectiveness of placing some high school juniors in an “honors track” by testing whether they achieve higher scores on a standardized test of mathematics. However, in a more thorough impact evaluation, we would ask whether the tracking of these high school mathematics students favorably influenced their performance in college-level math courses. Or, we might ask whether students not placed in the highest track suffered a loss of self-esteem as an unintended consequence (Loveless, 1999).

Other types of summative evaluation, cost-effectiveness and **cost-benefit analysis**, address questions of efficiency by standardizing outcomes in terms of their dollar costs and values (Boardman et al., 2000; Trochim, 2002).

---

\(^4\) See Chapter 18 for coverage of these methodologies.

\(^5\) Observational, correlation, and experimental techniques are explained in Chapters 5, 10, and 12, respectively.

**Resistance to the Evaluation: Outsiders and Insiders**

Whether the evaluator’s role is primarily formative or summative, resistance to his or her presence or approach may develop. One reason for conflict is that investigators and program personnel are not in a naturally cooperative situation. Many researchers are at least somewhat concerned with the relationship of the study’s findings to the growth of knowledge in a particular academic discipline. Program personnel, however, may have a more nuts-and-bolts attitude. The academic training of some researchers encourages a more detached and analytical posture, whereas program personnel may be more sensitive to the issue of servicing people’s needs right now. Closely associated with these potential personality conflicts are differences in the respective roles. Program personnel are usually committed to the policies and strategies currently in use, whereas the researcher is in the position of asking how effective these strategies are. Indeed, it is likely that evaluation researchers are always viewed with some skepticism and often as a threat. The personnel see the researcher as taking time and money from the program while offering a possibly negative report in return. Furthermore, they may perceive the researcher as unfavorably judging their work, competence, and personalities. They may be uncertain regarding the criteria being used for making judgments or suspicious concerning the “real” motives of the evaluator. These conflicts of interest are compounded when the investigator is not actually hired by the agency or program concerned but by the government or another supervisory body.

Social scientists who do evaluation research are usually hired from outside an organization that is in the process of developing its own assessment plan or has constituted its own task force to work on solving a problem. The researcher is retained as an “expert” to assist in the implementation of the assessment or to help reach a solution. In this situation, the outsider must be sensitive to the way in which the problem has been defined and the preexisting efforts to solve it. Failure to approach these tasks...
The Evaluator's Role

in the spirit of collaboration can easily produce a sense of suspicion and hostility toward the outsider. Careful inquiries should be made about the perceived causes of the problem and the rationale for the evaluation plan. To be most effective, these inquiries should be posed not only to the most senior decision makers in the organization but also to a variety of people at all levels. Here, the outsider's role resembles the participant-as-observer in fieldwork. Even if the evaluator is using quantitative methods to collect data, the principles of effective rapport that apply to qualitative fieldwork nonetheless apply. Without basic trust and mutual understanding of the purpose of the intervention, even the best evaluation research design will fail.

An increasing number of evaluation research exercises are not being conducted by outsider "experts," but by regular employees of organizations hired to perform IN-HOUSE RESEARCH (Gill and Johnson, 2002). Self-study for the purpose of solving problems or assessing program effectiveness may be more economical than hiring outside consultants, especially if these employees are not full-time evaluators. Insiders are generally more knowledgeable about the functioning of the program being investigated. However, these individuals face a special set of constraints: (1) they may be less well trained than professional researchers; (2) they may be less likely to "push back" when cost-cutting serves as an incentive to do "quick and dirty" research with inadequate sampling or analysis; and (3) their scientific objectivity may be threatened to a greater extent than outsiders because, as regular employees, they may be only too aware of hidden agendas. Sometimes the real purpose of conducting an in-house study is to provide evidence that senior management's pet project is a success; in this situation, to design a study that may show it to be a failure may put one's own position at risk.

Let us examine the tensions inherent in the outsider's and insider's evaluator roles by looking at a large corporation that has launched a program to publicize what it is doing to hire and retain more female executives and wants the initiative evaluated. Is the organization sincerely interested in adding significant numbers of qualified women and acknowledging its own part in creating the problem or is it merely trying to convince the public that it is doing something valuable by showcasing a few, token female hires to bolster its own image? It is precisely this kind of focusing that is needed at the outset of the evaluation for the researcher to be able to decide whether the project is desirable from a personal and ethical point of view.

If the answer is that problems of hiring discrimination against women are actually being ignored, the outsider as evaluator may expose the initiative for what it is and recommend a more effective strategy that, in the long term, would lead to the hiring of more women. For example, to sensitize male executives to the overall issue, a teaching team composed of a psychologist, a sociologist, an American historian, and an economist—all with special knowledge in the area of women's studies could be contracted to organize and conduct intensive seminars on the topic of Inclusion. Alternatively, if the evaluation is being sponsored by an unsympathetic corporation, a potential evaluator may choose not to participate at all rather than become part of a deceptive public relations campaign. An insider taking the role of an evaluator, and who wants to be a "team player," may have less freedom to select research approaches and make recommendations.

As the twenty-first century begins, an effort is under way to develop some alternative evaluator roles that can avoid or reduce tensions such as these by decentering the evaluator as "expert." These new roles encourage the active participation of those being evaluated in all aspects of the work, including the planning of the engagement, data collection, the process of making recommendations and implementing them. The impetus for this innovation in evaluation research comes from the same source as the feminist and postmodern critiques of conventional research methodology. 

The argument for involving evaluators more closely with program people (and other interested parties)

---

6 See Chapter 9 for a full explanation of this role.

7 Refer to Chapter 10 for a review of these critiques.