2

Tailoring Evaluations

Chapter Outline

What Aspects of the Evaluation Plan Must Be Tailored?

What Features of the Situation Should the Evaluation Plan Take Into Account?

The Purposes of the Evaluation

Program Improvement
Accountability
Knowledge Generation
Hidden Agendas

The Program Structure and Circumstances

Stage of Program Development
Administrative and Political Context of the Program
Conceptual and Organizational Structure of the Program

The Resources Available for the Evaluation

The Nature of the Evaluator-Stakeholder Relationship

Evaluation Questions and Evaluation Methods

Needs Assessment

Assessment of Program Theory

Assessment of Program Process

Impact Assessment

Efficiency Assessment

Every evaluation must be tailored to a specific set of circumstances. The tasks that

******ebook converter DEMO Watermarks*******

evaluators undertake depend on the purposes of the evaluation, the conceptual and organizational structure of the program being evaluated, and the resources available. Formulating an evaluation plan requires the evaluator to first explore these aspects of the evaluation situation with the evaluation sponsor and other key stakeholders. Based on this reconnaissance, the evaluator can then develop a plan that identifies the evaluation questions to be answered, the methods for answering them, and the relationships to be developed with the stakeholders during the course of the evaluation.

No hard-and-fast guidelines direct the process of designing an evaluation. Nonetheless, achieving a good fit between the evaluation plan and the program circumstances involves attention to certain critical themes. It is essential that the evaluation plan be responsive to the purposes of the evaluation as understood by the evaluation sponsor and key stakeholders. An evaluation intended to provide feedback to program decisionmakers for improving a program will take a different approach than one intended to help funders determine whether a program should be terminated. In addition, the evaluation plan must reflect an understanding of how the program is designed and organized so that the questions asked and the data collected will be appropriate to the circumstances. Finally, of course, any evaluation will have to be designed within the constraints of available time, personnel, and funding.

Although the particulars are diverse, the situations confronting the evaluator typically present one of a small number of variations. In practice, therefore, tailoring an evaluation is usually a matter of selecting and adapting one or another of a set of familiar evaluation schemes to the circumstances at hand. One set of evaluation schemes centers around the nature of the evaluator-stakeholder relations. Another distinct set of approaches is organized around common combinations of evaluation questions and the usual methods for answering them. This chapter provides an overview of the issues and considerations the evaluator should take into account when tailoring an evaluation plan.

One of the most challenging aspects of evaluation is that there is no “one size fits all” approach. Every evaluation situation has a different and sometimes unique profile of characteristics. The evaluation design must, therefore, involve an interplay between the nature of the evaluation situation and the evaluator’s repertoire of approaches, techniques, and concepts. A good evaluation design is one that fits the circumstances while yielding credible and useful answers to the questions that motivate it. We begin
******ebook converter DEMO Watermarks*******

our discussion of how to accomplish this goal by taking inventory of the aspects of the evaluation plan that need to be tailored to the program and the context of the evaluation.

What Aspects of the Evaluation Plan Must Be Tailored?

Evaluation designs may be quite simple and direct, perhaps addressing only one narrow question such as whether using a computerized instructional program helps a class of third graders read better. Or they may be prodigiously complex, as in a national evaluation of the operations and effects of a diverse set of programs for reducing substance abuse in multiple urban sites. Fundamentally, however, we can view any evaluation as structured around three issues that will be introduced in this chapter and elaborated in the chapters to come:

The questions the evaluation is to answer. A large number of questions might be raised about any social program by interested parties. There may be concerns about such matters as the needs of the **targets** (persons, families, or social groups) to which a program is directed and whether they are being adequately served, the management and operation of the program, whether the program is having the desired impact, and its costs and efficiency. No evaluation can, nor generally should, attempt to address all such concerns. A central feature of an evaluation design, therefore, is specification of its guiding purpose and the corresponding questions on which it will focus.

The methods and procedures the evaluation will use to answer the questions. A critical skill for the evaluator is knowing how to obtain useful, timely, and credible information about the various aspects of program performance. A large repertoire of social research techniques and conceptual tools is available for this task. An evaluation design must identify the methods that will be used to answer each of the questions at issue and organize them into a feasible work plan. Moreover, the methods selected must be practical as well as capable of providing meaningful answers to the questions with the degree of scientific rigor appropriate to the evaluation circumstances.

The nature of the evaluator-stakeholder relationship. One of the most important lessons from the first several decades of experience with systematic evaluation is that there is nothing automatic about the use of evaluation findings by the relevant stakeholders. Part of an evaluation design, therefore, is a plan for working with stakeholders to identify and clarify the issues, conduct the evaluation, and make effective use of the findings. This interaction may be collaborative, with the evaluator serving as a consultant or facilitator to stakeholders who take primary responsibility for

******ebook converter DEMO Watermarks*******

planning, conducting, and using the evaluation. Alternatively, the evaluator may have that responsibility but seek essential guidance and information from stakeholders. An evaluation plan should also indicate which audiences are to receive information at what times, the nature and schedule of written reports and oral briefings, and how broadly findings are to be disseminated beyond the evaluation sponsor.

What Features of the Situation Should the Evaluation Plan Take Into Account?

In developing an evaluation plan, the evaluator must be guided by a careful analysis of the context of the evaluation. The most significant features of the situation to consider have to do with (1) the purposes of the evaluation, (2) the program's structure and circumstances, and (3) the resources available for the evaluation.

The Purposes of the Evaluation

Evaluations are initiated for many reasons. They may be intended to help management improve a program; support advocacy by proponents or critics; gain knowledge about the program's effects; provide input to decisions about the program's funding, structure, or administration; or respond to political pressures. One of the first determinations the evaluator must make is just what the purposes of a specific evaluation are. This is not always a simple matter. A statement of the purposes generally accompanies the initial request for an evaluation, but these announced purposes rarely tell the whole story and sometimes are only rhetorical. Furthermore, evaluations may be routinely required in a program situation or sought simply because it is presumed to be a good idea without any distinct articulation of the sponsor's intent (see Exhibit 2-A).

The prospective evaluator must attempt to determine who wants the evaluation, what they want, and why they want it. There is no cut-and-dried method for doing this, but it is usually best to approach this task the way a journalist would dig out a story. The evaluator should examine source documents, interview key informants with different vantage points, and uncover pertinent history and background. Generally, the purposes of the evaluation will relate mainly to program improvement, accountability, or knowledge generation (Chelimsky,1997),but sometimes quite different motivations are in play.

******ebook converter DEMO Watermarks*******

Program Improvement

An evaluation intended to furnish information for guiding program improvement is called a **formative evaluation** (Scriven, 1991) because its purpose is to help form or shape the program to perform better (see Exhibit 2-B for an example). The audiences for formative evaluations typically are program planners, administrators, oversight boards, or funders with an interest in optimizing the program's effectiveness. The information desired may relate to the need for the program, the program's design, its implementation, its impact, or its efficiency. The evaluator in this situation will usually work closely with program management and other stakeholders in designing, conducting, and reporting the evaluation. Evaluation for program improvement characteristically emphasizes findings that are timely, concrete, and immediately useful.

EXHIBIT 2-A
Does Anybody Want This Evaluation?

Our initial meetings with the Bureau of Community Services administrators produced only vague statements about the reasons for the evaluation. They said they wanted some information about the cost-effectiveness of both New Dawn and Pegasus and also how well each program was being implemented… . It gradually became clear that the person most interested in the evaluation was an administrator in charge of contracts for the Department of Corrections, but we were unable to obtain specific information concerning where or how the evaluation would be used. We could only discern that an evaluation of state-run facilities had been mandated, but it was not clear by whom.

SOURCE: Quoted from Dennis J. Palumbo and Michael A. Hallett, "Conflict Versus Consensus Models in Policy Evaluation and Implementation," *Evaluation and Program Planning,* 1993, 16(1):11-23.

EXHIBIT 2-B
A Stop-Smoking Telephone Line That Nobody Called

Formative evaluation procedures were used to help design a "stop smoking" hotline for adult smokers in a cancer control project sponsored by a health maintenance organization (HMO). Phone scripts for use by the hotline counselors and other

******ebook converter DEMO Watermarks*******

aspects of the planned services were discussed with focus groups of smokers and reviewed in telephone interviews with a representative sample of HMO members who smoked. Feedback from these informants led to refinement of the scripts, hours of operation arranged around the times participants said they were most likely to call, and advertising of the service through newsletters and “quit kits” routinely distributed to all project participants. Despite these efforts, an average of less than three calls per month was made during the 33 months the hotline was in operation. To further assess this disappointing response, comparisons were made with similar services around the country. This revealed that low use rates were typical but the other hotlines served much larger populations and therefore received more calls. The program sponsors concluded that to be successful, the smoker’s hotline would have to be offered to a larger population and be intensively publicized

SOURCE: Adapted from Russell E. Glasgow, H. Landow, J. Hollis, S. G. McRae, and P. A. La Chance, “A Stop-Smoking Telephone Help Line That Nobody Called,” *American Journal of Public Health,* February 1993, 83(2): 252-253.

Correspondingly, the communication between the evaluator and the respective audiences may occur regularly throughout the evaluation and be relatively informal.

Accountability

The investment of social resources such as taxpayer dollars by human service programs is justified by the presumption that the programs will make beneficial contributions to society. Program managers are thus expected to use resources effectively and efficiently and actually produce the intended benefits. An evaluation conducted to determine whether these expectations are met is called a **summative evaluation** (Scriven, 1991) because its purpose is to render a summary judgment on the program’s performance (Exhibit 2-C provides an example). The findings of summative evaluations are usually intended for decisionmakers with major roles in program oversight, for example, the funding agency, governing board, legislative committee, political decisionmaker, or upper management. Such evaluations may influence significant decisions about the continuation of the program, allocation of resources, restructuring, or legal action. For this reason, they require information that is sufficiently credible under scientific standards to provide a confident basis for action and to withstand criticism aimed at discrediting the results. The evaluator may be expected to function relatively independently in planning, conducting, and reporting the evaluation, with stakeholders providing input but not participating directly in decision making. In these situations, it may be important to avoid premature or careless conclusions, so

******ebook converter DEMO Watermarks*******

communication of the evaluation findings may be relatively formal, rely chiefly on written reports, and occur primarily at the end of the evaluation.

Knowledge Generation

Some evaluations are undertaken to describe the nature and effects of an intervention as a contribution to knowledge. For instance, an academic researcher might initiate an evaluation to test whether a program designed on the basis of theory, say, an innovative science curriculum, is workable and effective (see Exhibit 2-D for an example). Similarly, a government agency or private foundation may mount and evaluate a demonstration program to investigate a new approach to a social problem, which, if successful, could then be implemented more widely. Because evaluations of this sort are intended to make contributions to the social science knowledge base or be a basis for significant program innovation, they are usually conducted using the most rigorous methods feasible. The audience for the findings will include the sponsors of the research as well as a broader audience of interested scholars and policymakers. In these situations, the findings of the evaluation are most likely to be disseminated through scholarly journals, research monographs, conference papers, and other professional outlets.

EXHIBIT 2-C
U.S. General Accounting Office Assesses Early Effects of the Mammography Quality Standards Act

The Mammography Quality Standards Act of 1992 required the Food and Drug Administration (FDA) to administer a code of standards for mammo-gram-screening procedures in all the states. When the act was passed, Congress was concerned that access to mammography services might decrease if providers choose to drop them rather than upgrade to comply with the new standards. The U.S. General Accounting Office (GAO) was asked to assess the early effects of implementing the act and report back to Congress. It found that the FDA had taken a gradual approach to implementing the requirements, which had helped to minimize adverse effects on access. The FDA inspectors had not closed many facilities that failed to meet the standards but, instead, had allowed additional time to correct the problems found during inspections. Only a relatively small number of facilities had terminated their mammography services and those were generally small providers located within 25 miles of another certified facility. The GAO concluded that the Mammography

******ebook converter DEMO Watermarks*******

Quality Standards Act was having a positive effect on the quality of mammography services, as Congress had intended.

SOURCE: Adapted from U.S. General Accounting Office, *Mammography Services: Initial Impact of New Federal Law Has Been Positive.* Report 10/27/95, GAO/HEHS-96-17 (Washington, DC: General Accounting Office, 1995).

Hidden Agendas

Sometimes the true purpose of the evaluation, at least for those who initiate it, has little to do with actually obtaining information about the program's performance. Program administrators or boards may launch an evaluation because they believe it will be good public relations and might impress funders or political decisionmakers. Occasionally, an evaluation is commissioned to provide a rationale for a decision that has already been made behind the scenes to terminate a program, fire an administrator, and the like. Or the evaluation may be commissioned as a delaying tactic to appease critics and defer difficult decisions.

Virtually all evaluations involve some political maneuvering and public relations, but when these are the principal purposes, the prospective evaluator is presented with a difficult dilemma. The evaluation must either be guided by the political or public relations purposes, which will likely compromise its integrity, or focus on program performance issues that are of no real interest to those commissioning the evaluation and may even be threatening to them. In either case, the evaluator is well advised to try to avoid such situations. If a lack of serious intent becomes evident during the initial exploration of the evaluation context, the prospective evaluator may wish to decline to participate. Alternatively, the evaluator might assume a consultant role at that point to help the parties clarify the nature of evaluation and redirect their efforts toward approaches more appropriate to their purposes.

EXHIBIT 2-D
Testing an Innovative Treatment Concept for Pathological Gambling

Pathological gambling is characterized by a loss of control over gambling impulses, lying about the extent of gambling, family and job disruption, stealing money, and chasing losses with additional gambling. Recent increases in the availability of gambling have led to corresponding increases in the prevalence of pathological gambling, but few treatment programs have been developed to help the victims of

******ebook converter DEMO Watermarks*******

this disorder. Research on the psychology of gambling has shown that problem gamblers develop an illusion of control and believe they have strategies that will increase their winnings despite the inherent randomness of games of chance. A team of clinical researchers in Canada hypothesized that treatment based on “cognitive correction” of these erroneous beliefs would be an effective therapy. Because excessive gambling leads to financial problems and interpersonal difficulties, they combined their cognitive intervention with problem-solving and social skills training.

To test their concept, the researchers recruited 40 pathological gamblers willing to accept treatment. These were randomly assigned to a treatment or control group and measures of pathological gambling, perception of control, desire to gamble, feelings of self-efficacy, and frequency of gambling were taken before and after treatment. The results showed significant positive changes in the treatment group on all outcome measures with maintenance of the gains at 6- and 12-month follow-up. However, the results may have been compromised by high attrition—8 of the 20 gamblers who began treatment and 3 of the 20 in the control group dropped out, a common occurrence during intervention for addictive problems. Despite this limitation, the researchers concluded that their results were strong enough to demonstrate the effectiveness of their treatment concept.

SOURCE: Adapted from Caroline Sylvain, Robert Ladouceur, and Jean-Marie Boisvert, “Cognitive and Behavioral Treatment of Pathological Gambling: A Controlled Study,” *Journal of Consulting and Clinical Psychology,* 1997, 65(5):727-732.

The Program Structure and Circumstances

No two programs are identical in their organizational structure and environmental, social, and political circumstances, even when they appear to provide the “same” service. The particulars of a program’s structure and circumstances constitute major features of the evaluation situation to which the evaluation plan must be tailored. Three broad categories of such particulars are especially important to evaluators: the stage of program development, the administrative and political context of the program, and the conceptualization and organizational structure of the program.

Stage of Program Development

The life of a social program can be viewed as a progression in which different
******ebook converter DEMO Watermarks*******

questions are at issue at different stages and, correspondingly, different evaluation approaches are needed (see Exhibit 2-E).Assessment of programs in the early stages of planning will be distinctly different from assessment of well-established programs. Similarly, assessment of an established program for which restructuring is contemplated or under way will be different from one for a program that is presumed to be stable in its basic operations and functions.

When new programs are initiated, especially innovative ones, evaluators are often called on to examine the social needs the program addresses, the program's design and objectives, the definition of its target population, the expected outcomes, and the means by which the program intends to attain those outcomes. The evaluator, therefore, may function as a planning consultant before the program is launched or when it is in an early stage of implementation by helping to assess and improve the program design.

Sometimes evaluations of new programs are expected to address questions of impact and efficiency, but the unsettled nature of programs in their beginning years most often makes those issues premature. It can easily take a year or more for a new program to establish facilities, acquire and train staff, make contact with the target population, and develop its services to the desired level. During this period, it may not be realistic to expect much impact on the social conditions that the program is intended to affect. Formative evaluation aimed at clarifying the needs of the target population, improving program operations, and enhancing the quality of service delivery, using approaches such as those discussed later in Chapters 4-6, is more likely to fit these cases.

Although the evaluation of new programs represents an important activity, by far the greater effort goes into assessing established programs, usually in terms of implementation issues. Evaluation of established, stable programs rarely focuses on assessing the underlying conceptualization of the program, that is, the rationale that went into the original design of the program. Stakeholders in well-established social programs are generally very reluctant to alter the programs' traditional forms and approaches unless some crisis compels them to consider fundamental change. So, for example, stakeholders take for granted the value of such well-entrenched programs as Social Security pensions, guidance counselors in schools, vocational programs for disabled persons, parole supervision for released convicts, and community health education for the prevention of diseases. In addition, long-standing programs that provide services to virtually the entire eligible population can be difficult to evaluate for their impact and efficiency. In such cases, the evaluator has limited ability to compare the outcomes and costs of the program to alternative situations that indicate what things would be like in the absence of the program. Accordingly, the evaluation of universally available programs is often directed toward assessing the extent to which their objectives are explicit and relevant to the interests of program sponsors, staff, and other stakeholders; whether the programs are well implemented and conform to plan; and whether they actually reach all of their target population. For example, the U.S.

******ebook converter DEMO Watermarks*******

Department of Agriculture conducts periodic studies of the food stamps program to measure the extent to which eligible households are enrolled and to guide outreach efforts to increase participation (Trippe, 1995).

EXHIBIT 2-E
Stages of Program Development and Related Evaluation Functions

Stage of Program Development	Question to Be Asked	Evaluation Function
1. Assessment of social problems and needs	To what extent are community needs and standards met?	Needs assessment; problem description
2. Determination of goals	What must be done to meet those needs and standards?	Needs assessment; service needs
3. Design of program alternatives	What services could be used to produce the desired changes?	Assessment of program logic or theory
4. Selection of alternative	Which of the possible program approaches is best?	Feasibility study; formative evaluation
5. Program implementation	How should the program be put into operation?	Implementation assessment
6. Program operation	Is the program operating as planned?	Process evaluation; program monitoring
7. Program outcomes	Is the program having the desired effects?	Outcome evaluation
8. Program efficiency	Are program effects attained at a reasonable cost?	Cost-benefit analysis; cost-effectiveness analysis

SOURCE: Adapted from S. Mark Pancer and Anne Westhues, "A Developmental Stage Approach to Program Planning and Evaluation," *Evaluation Review*, 1989, 13(1):56-77.

Sometimes, however, evaluation is sought for established programs because the program status quo has been called into question. This may result from political attack, competition, mounting program costs, changes in the target population, or dissatisfaction with program performance. When that happens, restructuring may be an option and evaluation may be sought to guide that change. Under these circumstances, the evaluation may focus on any aspect of the program—whether it is needed, its conceptualization and design, its operations and implementation, or its impact and efficiency.

The federal food stamps program, for instance, has been a national program for more than three decades. It is intended to increase the quantity and quality of food consumed by providing poor households with food stamps redeemable for approved foods purchased at grocery stores. At one point, the Department of Agriculture contemplated issuing checks instead, thereby eliminating the high costs of printing,
******ebook converter DEMO Watermarks*******

distributing, and redeeming the food stamps. To test this concept, it launched four experiments comparing the food consumption in households receiving food stamps with that in households receiving checks for the same dollar amount of benefits (Fraker, Martini, and Ohls, 1995). The results showed that households receiving checks purchased less food than those receiving food stamps. The Department of Agriculture therefore decided to retain food stamps despite their higher processing costs.

Administrative and Political Context of the Program

Except possibly for academic researchers who conduct evaluation studies on their own initiative for the purpose of generating knowledge, evaluators are not free to establish their own definitions of what the program is about, its goals and objectives, and what evaluation questions should be addressed. The evaluator works with the evaluation sponsor, program management, and other stakeholders to develop this essential background. Different perspectives from these various groups are to be expected. In most instances, the evaluator will solicit input from all the major stakeholders and attempt to incorporate their concerns so that the evaluation will be as inclusive and informative as possible.

If significant stakeholders are not in substantial agreement about the mission, goals, or other critical issues for the program, evaluation design becomes very difficult (see Exhibit 2-F). The evaluator can attempt to incorporate the conflicting perspectives into the design, but this may not be easy. The evaluation sponsors may not be willing to embrace the inclusion of issues and perspectives from groups they view as adversaries. Furthermore, these perspectives may be so different that they cannot be readily incorporated into a single evaluation plan with the time and resources available.

Alternatively, the evaluator can plan the evaluation from the perspective of only one of the stakeholders, typically the evaluation sponsor. This, of course, will not be greeted with enthusiasm by stakeholders with conflicting perspectives and they will likely oppose the evaluation and criticize the evaluator. The challenge to the evaluator is to be clear and straightforward about the perspective represented in the evaluation and the reasons for it, despite the objections. It is not necessarily inappropriate for an evaluation sponsor to insist that the evaluation emphasize its perspective, nor is it necessarily wrong for an evaluator to conduct an evaluation from that perspective without giving strong representation to conflicting views.

EXHIBIT 2-F
Stakeholder Conflict Over Home Arrest Program

******ebook converter DEMO Watermarks*******

In an evaluation of a home arrest program using electronic monitoring for offenders on parole, the evaluators made the following comments about stakeholder views:

There were numerous conflicting goals that were considered important by different agencies, including lowering costs and prison diversion, control and public safety, intermediate punishment and increased options for corrections, and treatment and rehabilitation. Different stakeholders emphasized different goals. Some legislators stressed reduced costs, others emphasized public safety, and still others were mainly concerned with diverting offenders from prison. Some implementers stressed the need for control and discipline for certain "dysfunctional" individuals, whereas others focused on rehabilitation and helping offenders become reintegrated into society. Thus, there was no common ground for enabling "key policy-makers, managers, and staff" to come to an agreement about which goals should have priority or about what might constitute program improvement.

SOURCE: Dennis J. Palumbo and Michael A. Hallett, "Conflict Versus Consensus Models in Policy Evaluation and Implementation," *Evaluation and Program Planning,* 1993, 16(1): ll-23.

Suppose, for instance, that the funders of a job training program for unemployed persons have concerns about whether the program is mainly taking cases that are easy to work with and, additionally, is providing only vocational counseling services rather than training in marketable job skills. The sponsors may appropriately commission an evaluation to examine these questions. Program managers, in contrast, will likely have a sharply conflicting perspective that justifies their selection of clients, program activities, and management practices. A conscientious evaluator will listen to the managers' perspective and encourage their input so that the evaluation can be as sensitive as possible to their legitimate concerns about what the program is doing and why. But the evaluation design should, nonetheless, be developed primarily from the perspective of the program funders and the issues that concern them. The evaluator's primary obligations are to be forthright about the perspective the evaluation takes, so there is no misunderstanding, and to treat the program personnel fairly and honestly.

Another approach to situations of stakeholder conflict is for the evaluator to design an evaluation that attempts to facilitate better understanding among the conflicting parties about the aspects of the program at issue. This might be done through efforts to clarify the nature of the different concerns, assumptions, and perspectives of the parties. For instance, parents of special education children may believe that their children are stigmatized and discriminated against when mainstreamed in regular classrooms. Teachers may feel equally strongly that this is not true. A careful observational study by an evaluator of the interaction of regular and special education children may reveal that

******ebook converter DEMO Watermarks*******

there is a problem, as the parents claim, but that it occurs outside the classroom on the playground and during other informal interactions among the children, thus accounting for the teachers' perspective.

Where stakeholder conflict is deep and hostile, it may be based on such profound differences in political values or ideology that an evaluation, no matter how comprehensive and ecumenical, cannot reconcile them. One school of thought in the evaluation field holds that many program situations are of this sort and that differences in values and ideology are the central matter to which the evaluator must attend. In this view, the social problems that programs address, the programs themselves, and the meaning and importance of those programs are all social constructions that will inevitably differ for different individuals and groups. Thus, rather than focus on program objectives, decisions, outcomes, and the like, evaluators are advised to directly engage the diverse claims, concerns, issues, and values put forth by the various stakeholders.

Guba and Lincoln (1987, 1989, 1994), the leading proponents of this particular construction of evaluation, have argued that the proper role of the evaluator is to encourage interpretive dialogue among the program stakeholders. From this perspective, the primary purpose of an evaluation is to facilitate negotiations among the stakeholders from which a more shared construction of the value and social significance of the program can emerge that still respects the various ideologies and concerns of the different stakeholders.

Finally, evaluators must realize that, despite their best efforts to communicate effectively and develop appropriate, responsive evaluation plans, program stakeholders owe primary allegiance to their own positions and political alignments. This means that sponsors of evaluation and other stakeholders may turn on the evaluator and harshly criticize the evaluation if the results contradict the policies and perspectives they advocate. Thus, even those evaluators who do a superb job of working with stakeholders and incorporating their views and concerns in the evaluation plan should not expect to be acclaimed as heroes by all when the results are in. The multiplicity of stake-holder perspectives makes it likely that no matter how the results come out, someone will be unhappy. It may matter little that everyone agreed in advance on the evaluation questions and the plan for answering them, or that each stakeholder group understood that honest results might not favor its position. Nonetheless, it is highly advisable for the evaluator to give early attention to identifying stakeholders, devising strategies for minimizing discord due to their different perspectives, and conditioning their expectations about the evaluation results.

Conceptual and Organizational Structure of the Program

It is a simple truism that if stakeholders do not have a clear idea about what a

******ebook converter DEMO Watermarks*******

program is supposed to be doing, it will be difficult to evaluate how well it is doing it. One factor that shapes the evaluation design, therefore, is the conceptualization of the program, or the **program theory**, that is, its plan of operation, the logic that connects its activities to the intended outcomes, and the rationale for why it does what it does. As we will discuss later in this chapter, this conceptual structure can itself be a focus of evaluation. The more explicit and cogent the program conceptualization, the easier it will be for the evaluator to identify the program functions and effects on which the evaluation should focus. If there is significant uncertainty about whether the program conceptualization is appropriate for the social problem the program addresses, it may make little sense for the evaluation design to focus on how well the conceptualization has been implemented. In such cases, the evaluation may be more usefully devoted to assessing and better developing the program plan. In the planning stages of a new program, an evaluator can often help sharpen and shape the program design to make it both more explicit and more likely to effectively achieve its objectives.

When a program is well established, everyday practice and routine operating procedures tend to dominate, and key stakeholders may find it difficult to articulate the underlying program rationale or agree on any single version of it. For instance, the administrators of a counseling agency under contract to a school district to work with children having academic problems may be quite articulate about their counseling theories, goals for clients, and therapeutic techniques. But they may have difficulty expressing a clear view of how their focus on improving family communication is supposed to translate into better grades. It may then become the task of the evaluator to help program personnel to formulate the implicit but unarticulated rationale for program activities.

At a more concrete level, evaluators also need to take into consideration the organizational structure of the program when planning an evaluation. Such program characteristics as multiple services or multiple target populations, distributed service sites or facilities, or extensive collaboration with other organizational entities have powerful implications for evaluation. In general, organizational structures that are larger, more complex, more decentralized, and more geographically dispersed will present greater practical difficulties than their simpler counterparts. In such cases, a team of evaluators is often needed, with resources and time proportionate to the size and complexity of the program. The challenges of evaluating complex, multisite programs are sufficiently daunting that they are distinct topics of discussion in the evaluation literature (see Exhibit 2-G; Turpin and Sinacore, 1991).

Equally important are the nature and structure of the particular intervention or service the program provides. The easiest interventions to evaluate are those that involve discrete, concrete activities (e.g., serving meals to homeless persons) expected to have relatively immediate and observable effects (the beneficiaries of the program are not hungry). The organizational activities and delivery systems for such

******ebook converter DEMO Watermarks*******

interventions are usually straightforward (soup kitchen), the service itself is uncomplicated (hand out meals), and the outcomes are direct (people eat). These features greatly simplify the evaluation questions likely to be raised, the data collection required to address them, and the interpretation of the findings.

EXHIBIT 2-G
Multisite Evaluations in Criminal Justice: Structural Obstacles to Success

Besides the usual methodological considerations involved in conducting credible evaluations, the structural features of criminal justice settings impose social, political, and organizational constraints that make multi-site evaluations difficult and risky. To begin, the system is extremely decentralized. Police departments, for example, can operate within the province of municipalities, counties, campuses, public housing, mass transit, and the states. The criminal justice system is also highly fragmented. Cities administer police departments and jails; counties administer sheriffs' and prosecutors' offices, jails, and probation agencies; state governments run the prisons. Agencies are embedded in disparate political settings, each with its own priorities for taxing and spending. In addition, criminal justice agencies foster a subculture of secrecy concerning their work that has serious consequences for evaluators, who are readily seen as "snoops" for management, the courts, or individuals with political agendas. Line staff easily adopt an "us against them" mentality toward outside evaluators. Also, criminal justice agencies generally exist in highly charged political environments. They are the most visible components of local government, as well as the most expensive, and their actions are frequently monitored by the media, who historically have assumed a watchdog or adversarial posture toward the system. Finally, the criminal justice system operates within a context of individual rights–legal constraint in procedural issues, an unwillingness to risk injustice in individual cases, and a stated (though not actually delivered) commitment to providing individualized treatment. This translates, for example, into a general aversion to the concept of random or unbiased assignment, the hallmark of the best designs for yielding interpretable information about program effects.

SOURCE: Adapted from Wesley G. Skogan and Arthur J. Lurigio, *Multisite Evaluations in Criminal Justice Settings: Structural Obstacles to Success,* New Directions for Evaluation, no. 50 (San Francisco: Jossey-Bass, summer 1991), pp. 83-96.

The most difficult interventions to evaluate are those that are diffuse in nature (e.g.,

******ebook converter DEMO Watermarks*******

community organizing), extend over long time periods (an elementary school math curriculum), vary widely across applications (psychotherapy), or have expected outcomes that are long term (preschool compensatory education) or indistinct (improved quality of life). For such interventions, many evaluation questions dealing with a program's process and outcome can arise because of the ambiguity of the services and their potential effects. The evaluator may also have difficulty developing measures that capture the critical aspects of the program's implementation and outcomes. Actual data collection, too, may be challenging if it must take place over extended time periods or involve many different variables and observations. All these factors have implications for the evaluation plan and, especially, for the effort and resources required to complete the plan.

The Resources Available for the Evaluation

Conducting a program evaluation requires resources of various kinds. Personnel must be allocated to the evaluation activities and arrangements must be made for materials, equipment, and facilities to support data collection, analysis, and reporting. These resources may be drawn from the existing capabilities of the program or evaluation sponsor, or they may be separately funded. An important aspect of planning an evaluation, therefore, is to break down the tasks and construct timelines for accomplishing them so that a detailed estimate can be made of the personnel, materials, and funds that are needed. The sum total of the resources required must then, of course, fit within what is available or changes in either the plan or the resources must be made. Useful advice on how to go about resource planning, budgeting, and determining timelines can be found in Hedrick, Bickman, and Rog (1992), Card, Greeno, and Peterson (1992), and Fink (1995, chap. 9).

Although available funding is, of course, a critical resource around which the evaluation must be planned, it is important to recognize that it is not the only resource that should concern the evaluator. Specialized expertise is often necessary and must be made available if the evaluation is to be done well and, in a large project, a considerable number of proficient evaluators, data collectors, data managers, analysts, and assistants may be required to do a quality job. Even with generous funding, it will not always be easy to obtain the services of sufficient persons with the requisite expertise. This is why large, complex evaluation projects are often done through contracts with research firms that maintain appropriate personnel on their permanent staffs.

Another critical resource for an evaluation is support from program management, staff, and other closely related stakeholders. For instance, the degree of assistance in data collection that program personnel will provide can have considerable influence on

******ebook converter DEMO Watermarks*******

how much an evaluation can accomplish. Barriers to access and lack of cooperation from the program or, worse, active resistance, are very expensive to the evaluation effort. It can take a substantial amount of time and effort to overcome these obstacles sufficiently to complete the evaluation. In the most severe cases, such resistance may compromise the scope or validity of the evaluation or even make it impossible to complete.

It is especially important for the evaluator to have access to program records, documents, and other such internal data sources that identify the number and characteristics of clients served, the type and amount of services they received, and the cost of providing those services. Information that can be confidently obtained from program records need not be sought in a separate, and almost certainly more expensive, data collection administered by the evaluator. However, program records vary in how easy they are to use for the purposes of evaluation. Records kept in writing often require considerable amounts of coding and processing while those kept in machine-readable databases can usually be used with relatively little adaptation. It is often advisable to inspect a sample of actual program records to try out the procedures for accessing and working with them and determine their completeness and quality.

The crucial point is that the evaluator must view cooperation from program personnel, access to program materials, and the nature, quality, and availability of data from program records as resource issues when planning an evaluation. The potential for misunderstanding and resistance can be lowered considerably if the evaluator spells out the resources and support needed for the evaluation in early discussions with the evaluation sponsors and program personnel (Hatry, 1994).

Experienced evaluators also know that one of the most precious resources is time. The time allotted for completion of the evaluation and the flexibility of the deadlines are essential considerations in evaluation planning but are rarely determined by the evaluator's preferences. The scheduling imperatives of the policy process usually control the time allowed for an evaluation because results typically must be available by a given date to play a role in a decision; after that they may be relatively useless. Further complicating the situation is the tendency for evaluation sponsors to underestimate the time needed to complete an evaluation. It is not uncommon for sponsors to request an evaluation that encompasses an imposing range of issues and requires considerable effort and then expect results in a matter of a few months.

The trade-offs here are quite significant. An evaluation can have breadth, depth, and rigor but will require proportionate funding and time. Or it can be cheap and quick but will, of necessity, deal with narrow issues or be relatively superficial (or both).All but the most sophisticated evaluation sponsors usually want evaluations that have breadth, depth, and rigor and that are also cheap and quick. Too often the result is overburdened evaluators working frantically against deadlines with inadequate resources and frustrated evaluation sponsors perturbed about delays in receiving the product they have

******ebook converter DEMO Watermarks*******

paid for. An especially direct relationship exists between the time and technical expertise available for the evaluation and the methods and procedures that can be realistically planned. With few exceptions, the higher the scientific standard to be met by the evaluation, the greater the time, expertise, effort, and program cooperation required.

It is generally better for an evaluation to answer a few important questions well than a larger number poorly. The best way for the evaluator to prevent disappointment is to negotiate very explicitly with the evaluation sponsor about the resources to be made available to the evaluation and the trade-offs associated with the inevitable constraints on those resources.

The Nature of the Evaluator-Stakeholder Relationship

Every program is necessarily a social structure in which various individuals and groups engage in the roles and activities that constitute the program: program managers administer, staff provides service, participants receive service, and so forth. In addition, every program is a nexus in a set of political and social relationships among those with an association or interest in the program, such as relevant policymakers, competing programs, and advocacy groups. Early in the planning process, evaluators should give explicit attention to the nature of their relationship with these and other stakeholders who may participate in the evaluation or who have an interest in the evaluation process or results. More specifically, the primary stakeholders the evaluator may need to consider include the following:

- *Policymakers and decisionmakers:* Persons responsible for deciding whether the program is to be started, continued, discontinued, expanded, restructured, or curtailed.
- *Program sponsors:* Organizations that initiate and fund the program. They may also overlap with policymakers and decisionmakers.
- *Evaluation sponsors:* Organizations that initiate and fund the evaluation (sometimes the evaluation sponsors and the program sponsors are the same).
- *Target participants:* Persons, households, or other units that receive the intervention or services being evaluated.
- *Program managers:* Personnel responsible for overseeing and administering the intervention program.
- *Program staff:* Personnel responsible for delivering the program services or in supporting roles.
- *Program competitors:* Organizations or groups that compete with the program for

******ebook converter DEMO Watermarks*******

available resources. For instance, an educational program providing alternative schools will attract the attention of the public schools because the new schools are seen as competitors.

- *Contextual stakeholders:* Organizations, groups, and individuals in the immediate environment of a program with interests in what the program is doing or what happens to it (e.g., other agencies or programs, public officials, or citizens' groups in the jurisdiction in which the program operates).
- *Evaluation and research community:* Evaluation professionals who read evaluations and pass judgment on their technical quality and credibility and researchers who work in areas related to a program.

All these groups or only a few may be involved in any given evaluation. But, whatever the assortment of stakeholders, the evaluator must be aware of their concerns and include in the evaluation planning appropriate means for interacting with at least the major stakeholders (Exhibit 2-H provides suggestions about how to do that).

At the top of the list of stakeholders is the evaluation sponsor. The sponsor is the agent who initiates the evaluation, usually provides the funding, and makes the decisions about how and when it will be done and who should do it. Various relationships with the evaluation sponsor are possible and will largely depend on the sponsor's preferences and whatever negotiation takes place with the evaluator. A common situation is one in which the sponsor expects the evaluator to function as an independent professional practitioner who will receive guidance from the sponsor, especially at the beginning, but otherwise take full responsibility for planning, conducting, and reporting the evaluation. For instance, program funders often commission evaluations by publishing a request for proposals (RFP) or applications (RFA) to which evaluators respond with statements of their capability, proposed design, budget, and time line, as requested. The evaluation sponsor then selects an evaluator from among those responding and establishes a contractual arrangement for the agreed-on work.

Other situations call for the evaluator to work more collaboratively with the evaluation sponsor. The sponsor may want to be involved in the planning, implementation, and analysis of results, either to react step by step as the evaluator develops the project or to actually participate with the evaluator in each step. Variations on this form of relationship are typical for internal evaluators who are part of the organization whose program is being evaluated. In such cases, the evaluator generally works closely with management in planning and conducting the evaluation, whether management of the evaluation unit, the program being evaluated, someone higher up in the organization, or some combination.

EXHIBIT 2-H

Stakeholder Involvement in Evaluation: Suggestions for Practice

******ebook converter DEMO Watermarks*******

Based on experience working with school district staff, one evaluator offers the following advice for bolstering evaluation use through stake-holder involvement:

- *Identify stakeholders:* At the outset, define the specific stakeholders who will be involved with emphasis on those closest to the program and who hold high stakes in it.
- *Involve stakeholders early:* Engage stakeholders in the evaluation process as soon as they have been identified because many critical decisions that affect the evaluation occur early in the process.
- *Involve stakeholders continuously:* The input of key stakeholders should be part of virtually all phases of the evaluation; if possible, schedule regular group meetings.
- *Involve stakeholders actively:* The essential element of stakeholder involvement is that it be active; stakeholders should be asked to address design issues, help draft survey questions, provide input into the final report, and deliberate about all important aspects of the project.
- *Establish a structure:* Develop and use a conceptual framework based in content familiar to stakeholders that can help keep dialogue focused. This framework should highlight key issues within the local setting as topics for discussion so that stakeholders can share concerns and ideas, identify information needs, and interpret evaluation results.

SOURCE: Adapted, from Robert A. Reineke, "Stakeholder Involvement in Evaluation: Suggestions for Practice," *Evaluation Practice*, 1991, 12(1):39-44.

In some instances, the evaluation sponsor will ask that the evaluator work collaboratively but stipulate that the collaboration be with another stakeholder group. For instance, private foundations often want evaluations to be developed in collaboration with the local stakeholders of the programs they fund. An especially interesting variant of this approach is when it is required that the recipients of program services take the primary role in planning, setting priorities, collecting information, and interpreting the results of the evaluation.

The evaluator's relationship to the evaluation sponsor and other stakeholders is so central to the evaluation context and planning process that a special vocabulary has arisen to describe various circumstances. The major recognized forms of evaluator-stakeholder relationships are as follows:

******ebook converter DEMO Watermarks*******

Independent evaluation. In an **independent evaluation**, the evaluator takes the primary responsibility for developing the evaluation plan, conducting the evaluation, and disseminating the results. The evaluator may initiate and direct the evaluation quite autonomously, as when a social scientist undertakes an evaluation for purposes of knowledge generation with research funding that leaves the particulars to the researcher's discretion. More often, the independent evaluator is commissioned by a sponsoring agency that stipulates the purposes and nature of the evaluation but leaves it to the evaluator to do the detailed planning and conduct the evaluation. In such cases, however, the evaluator generally confers with a range of stakeholders to give them some influence in shaping the evaluation.

Participatory or collaborative evaluation. A **participatory or collaborative evaluation** is organized as a team project with the evaluator and representatives of one or more stakeholder groups constituting the team (Greene,1988; Mark and Shotland,1985).The participating stakeholders are directly involved in planning, conducting, and analyzing the evaluation in collaboration with the evaluator whose function might range from team leader or consultant to that of a resource person called on only as needed. One particularly well-known form of participatory evaluation is Patton's (1986, 1997) utilization-focused evaluation. Patton's approach emphasizes close collaboration with those who will use the evaluation findings to ensure that the evaluation is responsive to their needs and produces information that they can and will actually use.

Empowerment evaluation. Some evaluators have advanced a view of evaluator-stakeholder relations that emphasizes the initiative, advocacy, and self-determination of the stakeholders (Fetterman, Kaftarian, and Wandersman, 1996). In an **empowerment evaluation**, the evaluator-stakeholder relationship is participatory and collaborative. In addition, however, the evaluator's role includes consultation and facilitation directed toward developing the capabilities of the participating stakeholders to conduct evaluations on their own, to use the results effectively for advocacy and change, and to experience some sense of control over a program that affects their lives. The evaluation process, therefore, is directed not only at producing informative and useful findings but also at enhancing the self-development and political influence of the participants. As these themes imply, empowerment evaluation most appropriately involves those stakeholders who otherwise have little power in the context of the program, usually the program recipients or intended beneficiaries.

A significant contribution of the participatory and empowerment perspectives is to call into question the assumption that an independent evaluation is always appropriate. Participation by the evaluation sponsor or other stakeholder groups may ensure that the

******ebook converter DEMO Watermarks*******

evaluation results will more closely address their concerns and be useful to them. Moreover, it can create a sense of ownership in the evaluation that amplifies the significance of its findings and reduces its potential to engender resistance. And, as the empowerment theorists point out, when stakeholder groups with little formal power are able to conduct and use an evaluation, it can alter the balance of power in a program context by enhancing their influence and sense of efficacy. It is thus appropriate for the evaluation sponsors and the evaluator to explicitly consider the way evaluation responsibilities will be assigned and the arrangements for organizing the evaluator-stake-holder collaboration.

Whether the evaluation is planned and conducted by an independent evaluator or a team of stakeholders has a considerable effect on the nature of the decision making, the evaluator's role, and, most likely, the focus and character of the evaluation. The result, nonetheless, should represent an application of established evaluation concepts and methods to the program. We thus distinguish the process of working with stake-holders, whether as an independent evaluator, collaborator, facilitator, or resource person, from the evaluation plan that results from that process. The features of a good plan for the evaluation context and the program at issue should be much the same whatever the process through which the planning and implementation is done. In the remainder of this chapter, therefore, we will discuss general planning issues and, when reference is made to the evaluator's role, assume that can mean either an independent evaluator or a collaborative team.

In addition to the evaluator-stakeholder relationship, the evaluation plan should make some provision for the communication and dissemination of the findings of the evaluation. To be useful, evaluation findings must be communicated to those with interest in the program, especially to those with responsibility for making important decisions. It is difficult to communicate evaluation findings in fine detail and there is often inherent uncertainty about what information will be of most interest to stakeholders. It is usually best, therefore, to discuss this issue with the major stakeholders and develop an organized communication and dissemination plan from the beginning. Useful advice for planning effective communication and dissemination activities is found in Torres, Preskill, and Piontek (1996; also see Exhibit 2-I).

Evaluation Questions and Evaluation Methods

A program evaluation is essentially information gathering and interpretation to answer questions about a program's performance and effectiveness. An important step in designing an evaluation, therefore, is determining the questions the evaluation must answer. This is sometimes done in a very perfunctory manner, but we advocate that it be

******ebook converter DEMO Watermarks*******

given studious and detailed attention. A carefully developed set of **evaluation questions** gives structure to the evaluation, leads to appropriate and thoughtful planning, and serves as a basis for essential discussions about who is interested in the answers and how they will be used. Indeed, constructing such questions and planning how to answer them is the primary way in which an evaluation is tailored to the unique circumstances associated with each program that comes under scrutiny.

EXHIBIT 2-I
Successful Communication With Stakeholders

Torres, Preskill, and Piontek (1996) surveyed and interviewed members of the American Evaluation Association about their experiences communicating with stakeholders and reporting evaluation findings. The respondents identified the following elements of effective communication:

- Ongoing, collaborative communication processes were the most successful. Periodic meetings and informal conversations can be used to maintain close contact throughout the evaluation, and interim memos and draft reports can be used to convey findings as they develop.
- It is important to use varied formats for communication. These might include short reports and summaries, verbal presentations, and opportunities for informal interaction.
- The content of the communication should be tailored to the audience and be easy for them to understand. Communication should use clear language, graphs and charts, and vivid, concrete illustrations. It should present contextual information about the program and the evaluation, cover both positive and negative findings, and be specific about recommendations.

SOURCE: Adapted from Rosalie T. Torres, Hallie S. Preskill, and Mary E. Piontek, *Evaluation Strategies for Communicating and Reporting: Enhancing Learning in Organizations* (Thousand Oaks, CA: Sage, 1996), pp. 4-6.

Because the evaluation questions to be addressed are so pivotal to evaluation planning, Chapter 3 is devoted entirely to discussing the form they should take, how they are generated, and how they are winnowed, organized, and integrated to provide the structure for the evaluation design. For present purposes, we will assume that an appropriate set of evaluation questions has been identified and consider some of the

******ebook converter DEMO Watermarks*******

broader implications of their character for tailoring and planning the evaluation. In this regard, it is useful to recognize that evaluation questions generally fall into recognizable types according to the program issues they address. Five such types are commonly distinguished:

- *Needs assessment:* Questions about the social conditions a program is intended to ameliorate and the need for the program.
- *Assessment of program theory:* Questions about program conceptualization and design.
- *Assessment of program process (or process evaluation):* Questions about program operations, implementation, and service delivery.
- *Impact assessment (impact evaluation or outcome evaluation):* Questions about program outcomes and impact.
- *Efficiency assessment:* Questions about program cost and cost-effectiveness.

These forms of evaluation are discussed in detail in Chapters 4-11.Here we will only provide some guidance regarding the circumstances for which each is most appropriate.

Needs Assessment

The primary rationale for a social program is to alleviate a social problem. The impetus for a new program to increase literacy, for example, is likely to be recognition that a significant proportion of persons in a given population are deficient in reading skills. Similarly, an ongoing program may be justified by the persistence of a social problem: Driver education in high schools receives public support because of the continuing high rates of automobile accidents among adolescent drivers.

One important form of evaluation, therefore, assesses the nature, magnitude, and distribution of a social problem; the extent to which there is a need for intervention; and the implications of these circumstances for the design of the intervention. These diagnostic activities are referred to as **needs assessment** in the evaluation field but overlap what is called social epidemiology and social indicators research in other fields (McKillip, 1987; Reviere et al., 1996; Soriano, 1995; Witkin and Altschuld, 1995). Needs assessment is often a first step in planning a new program or restructuring an established one to provide information about what services are needed and how they might best be delivered. Needs assessment may also be appropriate to examine whether established programs are responsive to the current needs of the target participants and provide guidance for improvement. Exhibit 2-J provides an example of one of the several approaches that can be taken. Chapter 4 discusses the various aspects of needs assessment in detail.

******ebook converter DEMO Watermarks*******

Assessment of Program Theory

Given a recognized problem and need for intervention, it does not follow that any program, willy-nilly, will be appropriate for the job. The conceptualization and design of the program must reflect valid assumptions about the nature of the problem and represent a feasible approach to resolving it. Put another way, every social program is based on some plan or blueprint that represents the way it is "supposed to work." This plan is rarely written out in complete detail but exists nonetheless as a shared conceptualization among the principal stakeholders. Because this program plan consists essentially of assumptions and expectations about how the program should conduct its business in order to attain its goals, we will refer to it as the program theory. If this theory is faulty, the intervention will fail no matter how elegantly it is conceived or how well it is implemented (Chen, 1990; Weiss, 1972).

EXHIBIT 2-J
Needs for Help Among Homeless Men and Women

A representative sample of 1,260 homeless men and women were interviewed in New York City's municipal shelters for single adults to determine their perception of their needs. The interview covered 20 items, each indicating need for help in a particular area. Most respondents identified multiple needs, averaging 6.3. The need for help in finding a place to live and having a steady income were the most commonly cited needs overall, closely followed by the need for help in finding a job and improving job skills. Compared to women, men more often reported needs for help with drinking problems, drug problems, learning how to handle money, getting veterans benefits, problems with the police, getting along better with other people, and finding a place to live. Women more frequently reported needs for help with health and medical problems and learning self-protection skills. The evaluators pointed out that for programs to be truly responsive to these multiple needs, they must have the capacity to deliver or broker access to a comprehensive range of services.

SOURCE: Adapted by permission from Daniel B. Herman, Elmer L. Struening, and Susan M. Barrow, "Self-Reported Needs for Help Among Homeless Men and Women," *Evaluation and Program Planning,* 1994, 17(3):249-256. Copyright © 1998, John Wiley & Sons, Inc.

An **assessment of program theory** focuses on questions relating to the way the

******ebook converter DEMO Watermarks*******

program is conceptualized and designed. This type of assessment involves, first, describing the program theory in explicit and detailed form. Then, various approaches are used to examine how reasonable, feasible, ethical, and otherwise appropriate it is. The sponsors of this form of evaluation are generally funding agencies or other decisionmakers attempting to launch a new program. Exhibit 2-K describes an examination of the conceptual foundation for family preservation programs that indicated that programs based on those notions had little prospect for success. Chapter 5 provides further discussion of program theory and the ways in which it can be evaluated.

EXHIBIT 2-K
A Flaw in the Design of Family Preservation Programs

As part of an evaluability assessment (see Chapter 5), evaluators working under contract to the U.S. Department of Health and Human Services reviewed the design of family preservation programs (FPPs). FPPs are time-limited, intensive home-based services to families in crisis that are intended to prevent the placement of children in foster care. The evaluators held discussions with the staff of federal and national private sector agencies about the definition of FPPs, reviewed available literature, obtained descriptions of state and local programs, and made site visits to four programs. From this information they developed "models" of how the programs were supposed to operate and then obtained the views of policymakers, program managers, and operating-level staff on four key dimensions: (a) program goals, (b) aspects of the child welfare system that affect the programs, (c) the target population, and (d) the characteristics that distinguish FPPs from other home-based services. Based on their own analysis and discussions with an expert advisory committee, the evaluators concluded that, as currently designed, FPPs could not achieve the policymakers' primary goal of preventing placement in foster care. The major flaw found in the program design was the practical difficulty of identifying children at "imminent risk" of placement; this meant that programs could not consistently target families with children truly at risk of placement.

SOURCE: Adapted from Joseph S. Wholey, "Assessing the Feasibility and Likely Usefulness of Evaluation," in *Handbook of Practical Program Evaluation,* eds. J. S. Wholey, H. P. Hatry, and K. E. Newcomer (San Francisco: Jossey-Bass, 1994), pp. 29-31. Wholey's account, in turn, is based on Kaye and Bell (1993).

******ebook converter DEMO Watermarks*******

Assessment of Program Process

Given a plausible theory about how to intervene in an accurately diagnosed social problem, a program must still be implemented well to have a reasonable chance of actually improving the situation. It is not unusual to find that programs are not implemented and executed according to their intended design. A program may be poorly managed, compromised by political interference, or designed in ways that are impossible to carry out. Sometimes appropriate personnel are not available, facilities are inadequate, or program staff lack motivation, expertise, or training. Possibly the intended program participants do not exist in the numbers required, cannot be identified precisely, or are not cooperative.

A basic and widely used form of evaluation, **assessment of program process**, assesses the fidelity and effectiveness of a program's implementation. Such process assessments evaluate the activities and operations of the program and are commonly referred to as **process evaluation** or, when the evaluation is an ongoing function, **program monitoring**. Process evaluation investigates how well the program is operating. It might examine how consistent the services actually delivered are with the goals of the program, whether services are delivered to appropriate recipients, how well service delivery is organized, the effectiveness of program management, the use of program resources, and other such matters (Exhibit 2-L provides an example).

EXHIBIT 2-L

Failure on the Front Lines: Implementing Welfare Reform

Work Pays is a state-level welfare reform demonstration program in California designed to establish incentives to work and disincentives for staying on the Aid to Families with Dependent Children (AFDC) welfare program. The program administrators recognized that to realize the policymakers' intent, the workers in local welfare offices would have to inform their clients about the new policy and present this information in a positive, individualized way that would reinforce clients' understanding of their obligations and choices about work and welfare. An implementation assessment was therefore conducted in which researchers interviewed welfare workers about the Work Pays program and observed a number of meetings with clients. This information revealed that the type of transaction expected between welfare workers and their clients under the new policy was exceedingly rare. In more than 80% of their interviews with clients, workers did not provide and interpret information about the new policy. Most workers continued their routine patterns of collecting and verifying eligibility information and

******ebook converter DEMO Watermarks*******

providing scripted recitations of welfare rules. However, the evaluators also found that the workers had been given only minimal information about the Work Pays program and no additional time or resources for educating their large caseloads about the changes. These findings demonstrated that welfare reform was not fully implemented at the street level in California and revealed some of the reasons why it was not.

SOURCE: Adapted from Marcia K. Meyers, Bonnie Glaser, and Karin MacDonald, "On the Front Lines of Welfare Delivery: Are Workers Implementing Policy Reforms?" *Journal of Policy Analysis and Management,* 1998, 17(1):1-22.

Process evaluation is the most frequent form of program evaluation. It is used both as a freestanding evaluation and in conjunction with impact assessment (discussed below) as part of a more comprehensive evaluation. As a freestanding evaluation, it yields quality assurance information, assessing the extent to which a program is implemented as intended and operating up to the standards established for it. When the program model employed is one of established effectiveness, a demonstration that the program is well implemented can be presumptive evidence that the expected outcomes are produced as well. When the program is new, a process evaluation provides valuable feedback to administrators and other stakeholders about the progress that has been made implementing the program plan. From a management perspective, process evaluation provides the feedback that allows a program to be managed for high performance (Wholey and Hatry, 1992), and the associated data collection and reporting of key indicators may be institutionalized in the form of a management information system (MIS) to provide routine, ongoing performance feedback.

In its other common application, process evaluation is an indispensable adjunct to impact assessment. The information about program outcomes that evaluations of impact provide is incomplete and ambiguous without knowledge of the program activities and services that produced those outcomes. When no impact is found, process evaluation has significant diagnostic value by indicating whether this was because of implementation failure, that is, the intended services were not provided hence the expected benefits could not have occurred, or theory failure, that is, the program was implemented as intended but failed to produce the expected effects. On the other hand, when program effects are found, process evaluation helps confirm that they resulted from program activities, rather than spurious sources, and identify the aspects of service most instrumental to producing the effects. Process evaluation is described in more detail in Chapter 6.

Impact Assessment

******ebook converter DEMO Watermarks*******

An **impact assessment**, sometimes called an impact evaluation or outcome evaluation, gauges the extent to which a program produces the intended improvements in the social conditions it addresses. Impact assessment asks whether the desired outcomes were attained and whether those changes included unintended side effects.

The major difficulty in assessing the impact of a program is that usually the desired outcomes can also be caused by factors unrelated to the program. Accordingly, impact assessment involves producing an estimate of the *net effects* of a program—the changes brought about by the intervention above and beyond those resulting from other processes and events affecting the targeted social conditions. To conduct an impact assessment, the evaluator must thus design a study capable of establishing the status of program recipients on relevant outcome measures and also estimating what their status would be had they *not* received the intervention. Much of the complexity of impact assessment is associated with obtaining a valid estimate of the latter status, known as the *counterfactual* because it describes a condition contrary to what actually happened to program recipients (Exhibit 2-M presents an example of impact evaluation).

Determining when an impact assessment is appropriate and what evaluation design to use present considerable challenges to the evaluator. Evaluation sponsors often believe that they need an impact evaluation and, indeed, it is the only way to determine if the program is having the intended effects. However, an impact assessment is characteristically very demanding of expertise, time, and resources and is often difficult to set up properly within the constraints of routine program operation. If the need for outcome information is sufficient to justify an impact assessment, there is still a question of whether the program circumstances are suitable for conducting such an evaluation. For instance, it makes little sense to establish the impact of a program that is not well structured or cannot be adequately described. Impact assessment, therefore, is most appropriate for mature, stable programs with a well-defined program model and a clear use for the results that justifies the effort required. Chapters 7-10 discuss impact assessment and the various ways in which it can be designed and conducted

EXHIBIT 2-M

No Impact on Garbage

Taiwan is a high-density island country with a garbage problem. Garbage accumulation has increased exponentially in recent years, 26 rivers are polluted by garbage, and the number of landfill sites is increasingly limited. Consequently, in 1993 a garbage reduction demonstration program (GRD) was launched in Nei-fu, a suburb of Taipei, and evaluated for its impact on the amount of waste produced. Garbage is collected daily in Taiwan and the plan of the GRD was to disrupt this

******ebook converter DEMO Watermarks*******

routine by suspending Tuesday collections. The theory was that requiring residents to store garbage one day a week in their homes, which are ill equipped for that function, would create sufficient inconvenience and unpleasantness to raise awareness of the garbage problem. As a result, it was expected that residents would make efforts to reduce the volume of garbage they produced. A process evaluation established that the program was implemented as planned.

The impact assessment was conducted by obtaining records of the daily volume of garbage for Nei-fu and the similar, adjacent suburb of Nan-kan for a period beginning four months prior to the program onset and continuing four months after. Analysis showed no reduction in the volume of garbage collected in Nei-fu during the program period relative to the preprogram volume or that in the comparison community. The evidence indicated that residents simply saved their customary volume of Tuesday garbage and disposed of it on Wednesday, with no carryover effects on the volume for the remainder of each week. Interviews with residents revealed that the program theory was wrong—they did not report the inconvenience or unpleasantness expected to be associated with storing garbage in their homes.

SOURCE: Adapted from Huey-Tsyh Chen, Juju C. S. Wang, and Lung-Ho Lin, "Evaluating the Process and Outcome of a Garbage Reduction Program in Taiwan," *Evaluation Review,* 1997, 21(1):27-42.

Efficiency Assessment

Finding that a program has positive effects on the target problem is often insufficient for assessing its social value. Resources for social programs are limited so their accomplishments must also be judged against their costs. Some effective programs may not be attractive because their costs are high relative to their impact in comparison to other program alternatives (Exhibit 2-N presents an example).

EXHIBIT 2-N
The Cost-Effectiveness of Community Treatment for Persons With Mental Disabilities

If provided with supportive services, persons with mental disabilities can often be maintained in community settings rather than state mental hospitals. But is such community treatment more costly than residential hospital care? A team of researchers in Ohio compared the costs of a community program that provides

******ebook converter DEMO Watermarks*******

housing subsidies and case management for state-certified severely mentally disabled clients with the costs of residential patients at the regional psychiatric hospital. Program clients were interviewed monthly for more than two years to determine their consumption of mental health services, medical and dental services, housing services, and other personal consumption. Information on the cost of those services was obtained from the respective service providers and combined with the direct cost of the community program itself. Costs for wards where patients resided 90 or more days were gathered from the Ohio Department of Mental Health budget data and subdivided into categories that corresponded as closely as possible to those tabulated for the community program participants. Mental health care comprised the largest component of service cost for both program and hospital clients. Overall, however, the total cost for all services was estimated at $1,730 per month for the most intensive version of community program services and about $6,250 per month for residential hospital care. Community care, therefore, was much less costly than hospital care, not more costly.

SOURCE: Adapted from George C. Galster, Timothy F. Champney, and Yolonda Williams, "Costs of Caring for Persons With Long-Term Mental Illness in Alternative Residential Settings," *Evaluation and Program Planning,* 1994,17(3):239-348.

An **efficiency assessment** takes account of the relationship between a program's costs and its effectiveness. Efficiency assessments may take the form of a **cost-benefit analysis** or a **cost-effectiveness analysis**. Typical issues include whether a program produces sufficient benefits in relation to its costs and whether other interventions or delivery systems can produce the benefits at a lower cost.

Efficiency assessment can be tricky and arguable because it requires making assumptions about the dollar value of program-related activities and, sometimes, imputing monetary value to a program's benefits. Nevertheless, such estimates are often essential for decisions about the allocation of resources to programs and identification of the program models that produce the strongest results with a given amount of funding.

Like impact assessment, efficiency assessment is most appropriate for mature, stable programs with a well-structured program model. This form of evaluation builds on process and impact assessment. A program must be well implemented and produce the desired outcomes before questions of efficiency become relevant. Given the specialized financial expertise required to conduct efficiency assessments, it is also apparent that it should be undertaken only when there is a clear need and identified user for the information. With the high level of concern about program costs in many contexts, however, this may not be an unusual circumstance. Chapter 11 discusses efficiency assessment methods in more detail.

******ebook converter DEMO Watermarks*******

Summary

■ Every evaluation must be tailored to a specific set of circumstances so that it will be capable of yielding credible and useful answers to the questions at issue while still being sufficiently practical to actually implement within the resources available.

■ Key aspects of the evaluation plan that must be tailored include the questions the evaluation is to answer, the methods and procedures to be used in answering those questions, and the nature of the evaluator-stakeholder relationship.

■ Three principal features of the evaluation context must be taken into account in an evaluation plan: the purpose of the evaluation, the structure and circumstances of the program being evaluated, and the resources available for the evaluation.

■ The overall purpose of the evaluation necessarily shapes its focus, scope, and construction. Evaluation is generally intended to provide feedback to program managers and sponsors, establish accountability to decisionmakers, or contribute to knowledge about social intervention.

■ The evaluation plan must also be responsive to a program's structure and circumstances, including how new or open to change the program is, the degree of consensus or conflict among stakeholders about the nature and mission of the program, the values and concepts inherent in the program rationale, and the way in which the program is organized and administered.

■ The evaluation plan must accommodate the inevitable limitations on the resources available for the evaluation effort. The critical resources include not only funding but also the time allowed for completion of the work, pertinent technical expertise, program and stakeholder cooperation, and access to important records and program material. A balance must be found between what is most desirable from an evaluation standpoint and what is feasible in terms of available resources.

■ An often neglected but critical aspect of an evaluation plan involves spelling out the appropriate relationship between the evaluator and the evaluation sponsor and other major stakeholders. The three major types of evaluator-stakeholder relationships are (1) independent evaluation, in which the evaluator takes primary responsibility for designing and conducting the evaluation; (2) participatory or collaborative evaluation, in which the evaluation is conducted as a team project involving stakeholders; and (3) empowerment evaluation, in which the evaluation is designed to help develop the capabilities of the participating stakeholders in ways that enhance their skills or

******ebook converter DEMO Watermarks*******

political influence.

- The questions an evaluation is designed to address generally fall into recognizable categories. Evaluators have developed relatively distinct conceptual and methodological approaches for these different issues. The main types of concerns addressed by evaluations and the associated methods are (1) the need for services (needs assessment), (2) the conceptualization and design of the program (assessment of program theory), (3) the implementation of a program (assessment of program process, also called process evaluation or program monitoring), (4) the program's outcomes (impact assessment), and (5) the program's efficiency (efficiency assessment). In practice, much of evaluation planning consists of identifying the approach corresponding to the type of questions to be answered, then tailoring the specifics to the program situation.

Key Concepts

Assessment of program process

An evaluative study that answers questions about program operations, implementation, and service delivery. Also known as a process evaluation or an implementation assessment.

Assessment of program theory

An evaluative study that answers questions about the conceptualization and design of a program.

Cost-benefit analysis

Analytical procedure for determining the economic efficiency of a program, expressed as the relationship between costs and outcomes, usually measured in monetary terms.

Cost-effectiveness analysis

Analytical procedure for determining the efficacy of a program in achieving given intervention outcomes in relation to the program costs.

Efficiency assessment

******ebook converter DEMO Watermarks*******

An evaluative study that answers questions about program costs in comparison to either the monetary value of its benefits or its effectiveness in terms of the changes brought about in the social conditions it addresses.

Empowerment evaluation

A participatory or collaborative evaluation in which the evaluator's role includes consultation and facilitation directed toward the development of the capabilities of the participating stakeholders to conduct evaluation on their own, to use it effectively for advocacy and change, and to have some influence on a program that affects their lives.

Evaluation questions

A set of questions developed by the evaluator, evaluation sponsor, and other stakeholders; the questions define the issues the evaluation will investigate and are stated in terms such that they can be answered using methods available to the evaluator in a way useful to stakeholders.

Formative evaluation

Evaluative activities undertaken to furnish information that will guide program improvement.

Impact assessment

An evaluative study that answers questions about program outcomes and impact on the social conditions it is intended to ameliorate. Also known as an impact evaluation or an outcome evaluation.

Independent evaluation

An evaluation in which the evaluator has the primary responsibility for developing the evaluation plan, conducting the evaluation, and disseminating the results.

Needs assessment

An evaluative study that answers questions about the social conditions a program is intended to address and the need for the program.

Participatory or collaborative evaluation

******ebook converter DEMO Watermarks*******

An evaluation organized as a team project in which the evaluator and representatives of one or more stakeholder groups work collaboratively in developing the evaluation plan, conducting the evaluation, and disseminating and using the results.

Process evaluation

A form of program monitoring designed to determine whether the program is delivered as intended to the target recipients. Also known as implementation assessment.

Program monitoring

The systematic documentation of aspects of program performance that are indicative of whether the program is functioning as intended or according to some appropriate standard. Monitoring generally involves program performance related to program process, program outcomes, or both.

Program theory

The set of assumptions about the manner in which a program relates to the social benefits it is expected to produce and the strategy and tactics the program has adopted to achieve its goals and objectives. Within program theory we can distinguish *impact theory,* relating to the nature of the change in social conditions brought about by program action, and *process theory,* which depicts the program's organizational plan and service utilization plan.

Summative evaluation

Evaluative activities undertaken to render a summary judgment on certain critical aspects of the program's performance, for instance, to determine if specific goals and objectives were met.

Target

The unit (individual, family, community, etc.) to which a program intervention is directed. All such units within the area served by a program comprise its target population.

******ebook converter DEMO Watermarks*******