# CHAPTER ONE:    EVALUATION  PROTOTYPES

The purpose of this chapter is to help Principal Investigators and Project Evaluators think practically about evaluation and the kinds of information evaluations can provide. We start with the assumption that the term "evaluation" describes different models or prototypes that suit different purposes at different stages in the life of a project.  A major goal of this chapter is to help Principal Investigators and Project Evaluators understand what some of these different prototypes are and to assist them in using different approaches to evaluation to meet these varying needs.

## What is Evaluation?

The notion of evaluation has been around a long time—in fact, the Chinese had a large functional evaluation system in place for their civil servants as long ago as 2000 B.C.  Not only does the idea of evaluation have a long history, but it also has varied definitions.  Evaluation means different things to different people and takes place in different contexts. Thus, evaluation can be synonymous with  tests, descriptions, documentation, or management. Many definitions have been developed, but a comprehensive definition is presented by the Joint Committee on Standards for Educational Evaluation (1981):

*Evaluation means different things to different people and takes place in different contexts.*

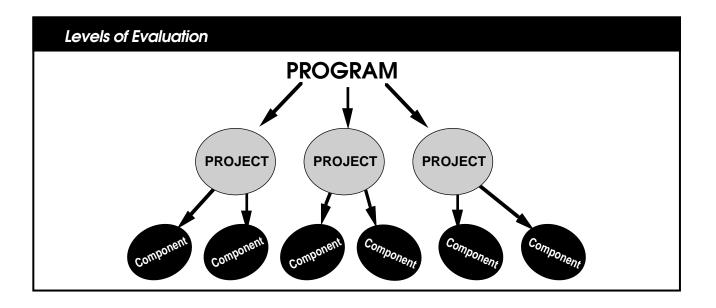**"Systematic investigation of the worth or merit of an object. . ."**

This definition centers on the goal of using evaluation for a purpose. Evaluations should be conducted for action-related reasons, and the information provided should facilitate deciding a course of action.

Over the years evaluation has frequently been viewed as an adversarial process.  Its main use has been to provide "thumbs-up" or "thumbs down" about a program or project.  In this role, it has all too often been considered by program or project Directors as an external imposition which is  threatening, disruptive, and not very helpful to Project staff. Our contention is that while this may be true in some situations, this is not the case in all, nor even in most, evaluation efforts. And, today in contrast to a decade or two ago, the view is gaining ground that evaluation should be a tool that not only measures, but can contribute to, success.

## What are the Different Kinds of Evaluations?

Within NSF, there are two basic levels of evaluation: Program Evaluation and Project Evaluation. Project Evaluation is sometimes further subdivided into specific project components as shown in Exhibit 1.

*Exhibit 1*

**Levels of Evaluation**



Let's start by defining terms and showing how they relate. First, let's define what we mean by a "program" and a "project." A **program** is a coordinated approach to exploring a specific area related to NSF's mission of strengthening science, mathematics, engineering, and technology. A **project** is a particular investigative or developmental activity funded by that program. NSF initiates a program on the assumption that a policy goal (for example, strengthening minority student development) can be attained by certain educational activities and strategies (for example, exposing students in inner-city schools to science presentations targeted at the interests and concerns of young African Americans). The Foundation then funds a series of discrete projects to explore the utility of these activities and strategies in specific situations. Thus, a program consists of a collection of projects that seek to meet a defined set of goals and objectives.

Now, let's turn to the terms "Program" and "Project Evaluation." A **Program Evaluation** determines the value of this collection of projects. It looks across projects, examining the utility of the activities and strategies employed, in light of the initial policy goal. It is carried to completion after the projects have become fully operational and adequate time has passed for expected outcomes to be achieved. Frequently, the initiation of a Program Evaluation is deferred until 3 to 5 years after program initiation. Other times, reporting may be deferred, while data collection is begun simultaneously with program onset. Under this latter alternative, the evaluation could draw upon information collected on an annual basis that is aggregated across projects and summarized at an appropriate check point. The program evaluator is, in this case, usually an experienced, external evaluator, selected by NSF.

*In NSF, a Program Evaluation determines the value of a collection of projects. Project Evaluation, in contrast, focuses on the individual projects funded under the umbrella of the program.*

**Project Evaluation**, in contrast, focuses on an individual project funded under the umbrella of the program. The evaluation provides information to improve the project as it develops and progresses. Information is collected to help determine whether it is proceeding as planned; whether it is meeting its stated program goals and project objectives according to the proposed timeline. Frequently these evaluation findings are also used to assess whether the particular project merits continued funding as it is currently operating, or if it needs modifications. Ideally in a Project Evaluation, evaluation design and data collection begin soon after the project is funded. Data collection occurs on a planned schedule, e.g., every 6 months or every year; and may lead to and support recommendations to continue, modify, and/or delete project activities and strategies. Frequently, although not universally, the Project Evaluator is a member of the project staff, is selected by, and reports to the Project Director.

Project Evaluations may also include examination of specific components. A component of a project may be a specific teacher training approach, a classroom practice, or a governance strategy. An evaluation of a component frequently looks to see the extent to which its goals have been met (these goals are a subset of the overall project goals), and to clarify the extent to which the component contributes to the success or failure of the overall project.

The information contained in this Handbook has been primarily prepared for the use of Project Evaluators

and Principal Investigators, although Program Evaluators may also find it useful. Our aim is to provide tools that will help those responsible for examination of individual projects gain the most from their evaluation efforts. Clearly, however, these activities will also benefit program studies and the work of the Foundation in general. The better the information about each of the NSF projects, the more we can all learn.

In the next section we describe three general types of evaluation studies: (1) Planning Evaluation, (2) Formative (Implementation and Progress) Evaluation, and (3) Summative Evaluation. While Summative Evaluation is frequently the notion that comes to mind when the term "evaluation" is used, each has its own contribution to make in understanding how well a project is doing. As each type of evaluation is discussed we present a brief definition of its purpose and some ideas of the kinds of questions that could be addressed.

## What is a Planning Evaluation?

The purpose of a Planning Evaluation is to assess understanding of a project's goals, objectives, strategies, and timelines. "Planning Evaluation" is not as commonly carried out as the other prototypes. In fact most project proposals typically mention only "Formative" and "Summative" Evaluation, defining these as activities to be performed once a project has been designed, written up, and funded. The evaluator enters the scene after the project has been put in place.

*The purpose of a Planning Evaluation is to assess understanding of projects' goals, objectives, strategies, and timelines.*

A strong argument can be made for a different approach. Rossi and Freeman (1993) argue strongly for the involvement of evaluators in diagnosing and defining the condition that a given project is designed to address, in stating clearly and precisely the goals of the project, and in reviewing the proposed procedures for accuracy of information and soundness of methods.

The Planning Evaluation will provide everyone—Program Directors, Principal Investigators, Project Directors/Managers, participants, and the public—with an understanding of what the project is supposed to do and the timelines and strategies for doing it. The product of the Planning Evaluation is a rich, context-laden description of a project, including its major goals and objectives, activities, participants and other major stakeholders, resources, timelines, locale, and intended accomplishments. The Planning Evaluation can also serve the purpose of describing the status of

key outcome indicators prior to the project to serve as a baseline for measuring success.

To conduct a Planning Evaluation, the evaluator should be present when the project is in its developmental phase. The Planning Evaluation is typically designed to address the following questions:

• Why was the project developed? What is the problem or need it is attempting to address?

• Who are the stakeholders (those who have credibility, power, or other capital involved in the project)? Who are the people interested in the project who may not be involved?

• What do the stakeholders want to know? What questions are most important to which stakeholders? What questions are secondary in importance? Where do concerns coincide? Where are they in conflict?

• Who are the participants to be served?

• What are the activities and strategies that will address the problem or need which was identified? What is the intervention? How will participants benefit? What are the expected outcomes?

• Where will the program be located (educational level, geographical area)?

• How many months of the school year or calendar year will the program operate? When will the program begin and end?

• How much does it cost? What is the budget for the program? What human, material, and institutional resources are needed? How much is needed for evaluation? for dissemination?

• What are the measurable outcomes which the project wants to achieve? What is the expected impact of the project in the short run? the longer run?

• What arrangements have been made for data collection? What are the understandings regarding record keeping, responding to surveys, and participation in testing?

These questions can become a checklist to determine

if all relevant elements are included in the description of the project or program. These questions also provide the basis for the formative and summative evaluative inquiries about the project.

## What is Formative Evaluation?

The purpose of Formative Evaluation is to assess ongoing project activities. Formative Evaluation begins at project start-up and continues throughout the life of the project. Its intent is to provide information to improve the project. It is done at several points in the developmental life of a project. According to evaluation theorist Bob Stake, Formative Evaluation, when contrasted with Summative Evaluation, is:

> *The purpose of Formative Evaluation is to assess ongoing project activities.*

**"When the cook tastes the soup, that's formative; when the guests taste the soup, that's summative."**

For most NSF projects, Formative Evaluation consists of two segments: Implementation Evaluation and Progress Evaluation.

## What is Implementation Evaluation?

The purpose of an Implementation Evaluation is to assess whether the project is being conducted as planned. It may occur once or several times during the life of the project. Recall the principle learned from the tale of the emperor who had no clothes and no one would tell him. The same principle applies to a new project or new program. Before you can evaluate the outcomes of a project, you must make sure the project is really operating, and if it is operating according to its plan or description. For example, in the description for Comprehensive Regional Centers for Minorities (CRCM), these Regional Centers must be comprehensive in their coverage of science, engineering and mathematics and focus on a span of educational levels—elementary through high school. An Implementation Evaluation of a CRCM project might begin by investigating whether or not the CRCM was indeed comprehensive in its coverage and whether its focus spanned elementary through senior high school. If these two essential conditions were satisfied, it could be concluded that the CRCM was initially implemented as intended and that evaluation of outcomes and impacts associated with the

implementation could proceed.

Implementation Evaluation collects information to determine if the program or project is being delivered as planned. A series of implementation questions is needed to guide the Implementation Evaluation. Examples of these questions are:

- Were the appropriate participants selected and involved in the planned activities?

- Do the activities and strategies match those described in the plan? If not, are the changes in activities justified and described?

- Were the appropriate staff members hired, trained, and are they working in accordance with the proposed plan?  Were the appropriate materials and equipment obtained?

- Were activities conducted according to the proposed timeline? by appropriate personnel?

- Was a management plan developed and followed?

*The purpose of an Implementation Evaluation is to assess whether the project is being conducted as planned.*

Sometimes the terms "Implementation Evaluation" and "Monitoring Evaluation" are confused. They are not the same. While Implementation Evaluation is an early internal check by the project staff to see if all the essential elements of the project are in place and operating, monitoring is an external check and should follow the Implementation Evaluation. The monitor comes from the funding agency and is responsible for determining progress and compliance on a contract or grant for the project. The monitor investigates proper use of funds, observes progress, and provides information to the funding agency about the project. Although the two differ, Implementation Evaluation, if effective, can facilitate and ensure that there are no unwelcome surprises during monitoring.

## What is Progress Evaluation?

The purpose of a Progress Evaluation is to assess progress in meeting the project's goals. Progress Evaluation is also formative. It involves collecting information to learn whether or not the benchmarks of participant progress were attained and to point out unexpected developments.  Progress Evaluation collects information to determine what the impact of the activities and strategies is on the participants at

various stages of the intervention. By measuring interim outcomes, project staff eliminate the risk of waiting until participants have experienced the entire treatment to assess outcomes. If the data collected as part of the Progress Evaluation fail to show expected changes, this information can be used to "fine-tune" or terminate the project. The data collected as part of a Progress Evaluation can also contribute to, or form the basis for, a Summative Evaluation study conducted at some future date. In a Progress Evaluation, the following questions could be asked:

- Are the participants moving toward the anticipated goals of the project or program?

- Which of the activities and strategies are aiding the participants to move toward the goals?

For example, one of the goals for the Alliances for Minority Participation (AMP) Program is to increase the size of the pool of underrepresented minority students eligible for Science, Engineering, and Mathematics (SEM) graduate study. One of the interim indicators which shows progress towards meeting the goal is the number (percent) of participants in the Summer Bridge program (a component of the AMP Program) who successfully complete calculus by their freshman year in college. Additional progress information could be scores from calculus classroom quizzes throughout the summer before the final exam and grades given for the course. Collecting this information on course completion, test scores, and grades, gives interested parties some idea of the rate and extent to which progress is being made toward the overarching goal of increasing the numbers of underrepresented minority students eligible for SEM graduate study. It gives some idea of the probability of achieving that final goal. If course completion or other indicators are not showing progress, significant project changes may be considered.

Another example of measuring progress can be drawn from Comprehensive Regional Centers for Minorities (CRCM). A goal is that, through workshops, teachers will learn to improve and enrich their teaching strategies when teaching classes such as high school chemistry. This interim goal is related to meeting the CRCM goal of retaining precollege students' interest in science. Progress findings could include teachers' ratings of their inservice training classes, and the independent

*The purpose of a Progress Evaluation is to assess progress in meeting the project's goals.*

appraisals by outside observers of the quality of their performance when using new strategies in the classroom. In addition, the opinions and attitudes of the participants (students and teachers) could be collected to determine whether the impact of the activities and strategies is negative or positive.

In Progress Evaluation, quantitative and qualitative information about the participants is collected to determine if parts of the project need to be changed or deleted to improve the project. Progress Evaluation is useful throughout the life of a project, but is most vital during the early stages when activities are piloted and their individual effectiveness or articulation with other project components is unknown.

## What is Summative Evaluation?

The purpose of a Summative Evaluation is to assess the project's success. Summative Evaluation takes place after ultimate modifications and changes have been made, after the project is stabilized and after the impact of the project has had a chance to be realized. (Another term frequently used interchangeably with "Summative Evaluation" is "Impact Evaluation.") Summative Evaluation answers these basic questions:

- Was the project successful? What were its strengths and weaknesses?

- To what extent did the project or program meet the overall goal(s)?

*The purpose of a Summative Evaluation is to assess the project's success.*

- Did the participants benefit from the project? In what ways?

- What components were the most effective?

- Were the results worth the project's cost?

- Is this project replicable and transportable?

Summative Evaluation collects information about processes and outcomes. The evaluation is an external appraisal of worth, value or merit. Usually this type of evaluation is needed for decisionmaking. The decision alternatives may include the following: disseminate the intervention to other sites or agencies; continue funding; increase the funding; continue on probationary status; or discontinue.

Summative Evaluation informs decisionmakers about whether the activities and strategies were successful in helping the project and/or its participants reach their goals. This evaluation also describes the extent to which each goal was attained. Sample Summative Evaluation questions for a project like AMP could include the following:

- Did the majority of the undergraduate students in the project graduate with majors in mathematics, engineering or science?

- What proportion of graduates pursued their education until they received doctorates in mathematics, engineering, or science?

- Which elements or combinations of elements (mentoring, counseling, tutoring, or financial support) were most effective in retaining students in the SEM pipeline?

An important idea to keep in mind in conducting a Summative Evaluation is what has been called "unanticipated outcomes." These are findings that come to light during data collection or data analyses that were never anticipated when the study was first designed. An example of an unanticipated finding comes from a study that started out to look at whether or not school buses should have seat belts. This study also looked at the cost of purchasing new buses that had seat belts, versus retrofitting old models. This study, prompted by a desire to assure the safety of students, was ultimately unable to reach definitive conclusions regarding the utility of seat belts from the data available. Along the way, however, it was found that buses manufactured before a certain date were missing other safety features and the safety of the transportation system could be greatly enhanced by replacing buses purchased before this date. This unanticipated outcome became the basis for significant changes in the system's transportation policy.

*An important idea to keep in mind while conducting a Summative Evaluation is to be vigilant of "unanticipated outcomes."*

## Summary

Evaluations can serve many different needs and provide critical data for decision-making at all steps of project development and implementation. Although some people feel that evaluation is an act that is done to a project, if done well, an evaluation is really done for the project.

It is important to remember that evaluation is not a

single thing, it is a process. When done well, evaluation can help inform the managers of the project as it progresses, can serve to clarify goals and objectives, and can provide important information on what is, or is not, working, and why.

This chapter has presented information to help Principal Investigators and project evaluators understand the various types of evaluation, the different stages in the evaluation process at which they occur, and the different kinds of information they provide. To summarize this information, a restatement of the important issues has been developed (see pages 12 and 13) to serve as a "shorthand" guide. For additional discussion of the various types of evaluation prototypes see Rossi and Freeman (1993). Chapter Five in this Handbook presents some additional examples of evaluations that further illustrate these roles and their differences.

## REFERENCES

Rossi, P. H. & Freeman, H. E. (1993). *Evaluation:—A Systematic Approach (5th Edition).* Newbury Park, CA: Sage.

Joint Committee on Standards for Educational Evaluation. (1981). *Standards for Evaluation of Educational Programs, Projects, and Materials.* New York, NY: McGraw-Hill.

## *Overview of Evaluation Prototypes*

### *Planning Evaluation:*

A Planning Evaluation assesses the understanding of project goals, objectives, strategies and timelines.

It addresses the following types of questions:

♦ Why was the project developed?  What is the problem or need it is attempting to address?

♦ Who are the stakeholders? Who are the people involved in the project? Who are the people interested in the project who may not be involved?

♦ What do the stakeholders want to know?  What questions are most important to which stakeholders? What questions are secondary in importance? Where do concerns coincide? Where are they in conflict?

♦ Who are the participants to be served?

♦ What are the activities and strategies that will involve the participants? What is the intervention?  How will participants benefit?  What are the expected outcomes?

♦ Where will the program be located (educational level, geographical area)?

♦ How many months of the school year or calendar year will the program operate?  When will the program begin and end?

♦ How much does it cost? What is the budget for the program? What human, material, and institutional resources are needed?  How much is needed  for evaluation? for dissemination?

♦ What are the measurable outcomes? What is the expected impact of the project in the short run?  the longer run?

♦ What arrangements have been made for data collection?  What are the understandings regarding record keeping, responding to surveys, and participation in testing?

### *Formative Evaluation*

A Formative Evaluation assesses ongoing project activities. It consists of two types:  Implementation Evaluation and Progress Evaluation.

#### *Implementation Evaluation*

An Implementation Evaluation assesses whether the project is being conducted as planned.  It addresses the following types of questions:

- **Were the appropriate participants selected and involved in the planned activities?**

- **Do the activities and strategies match those described in the plan? If not, are the changes in activities justified and described?**

- **Were the appropriate staff members hired, and trained, and are they working in accordance with the proposed plan? Were the appropriate materials and equipment obtained?**

- **Were activities conducted according to the proposed timeline? by appropriate personnel?**

- **Was a management plan developed and followed?**

### *Progress Evaluation*

**A Progress Evaluation assesses the progress made by the participants in meeting the project goals. It addresses the following types of questions:**

- **Are the participants moving toward the anticipated goals of the project?**

- **Which of the activities and strategies are aiding the participants to move toward the goals?**

### *Summative Evaluation*

**A Summative Evaluation assesses project success—the extent to which the completed project has met its goals. It addresses the following types of questions:**

- **Was the project successful?**

- **Did the project meet the overall goal(s)?**

- **Did the participants benefit from the project ?**

- **What components were the most effective?**

- **Were the results worth the project's cost?**

- **Is this project replicable and transportable?**

# CHAPTER TWO: THE EVALUATION PROCESS — AN OVERVIEW

In the preceding chapter, we outlined the types of evaluations that Principal Investigators and Project Evaluators may want to carry out. In this chapter we talk further about how to carry out an evaluation, expanding, in particular, on two types of studies, Formative and Summative. In the sections that follow we provide an orientation to some of the basic language of evaluation, as well as some hints about technical, practical, and political issues that should be kept in mind in conducting evaluation efforts. Our goal is to capture a snapshot of the various pieces that make up the evaluation process from planning to report writing.

This overview is limited to topics pertinent to content and technique and does not cover other practical issues, such as budget planning, time tables, etc. Information on such topics can be found in the *Project Evaluation Kit* described in Chapter 8 (Bibliography). We have also limited the discussion to the types of evaluation most frequently carried out for NSF.

## What are the Steps in Conducting a Formative or Summative  Evaluation?

Whether they are Summative or Formative, evaluations can be thought of as having five phases:

- Develop evaluation questions

- Match questions with appropriate information-gathering techniques

- Collect data

- Analyze data

- Provide information to interested audiences.

All five phases are critical for provision of useful information. If the information gathered is not perceived as valuable or useful (the wrong questions were asked) or the information is not credible or feasible (the wrong techniques were used), or the report is presented too late or is written inappropriately, then the evaluation will not contribute to the decisionmaking process.

In the sections below we provide an overview of each of these phases, describing the activities that need to

take place in each. This overview is intended to provide a basic understanding of conducting an evaluation from start to finish.  In Chapters Three and Four we provide greater detail in selected areas. A checklist summarizing the complete process is presented at the end of this chapter.

## How Do You Develop Evaluation Questions?

The development  of evaluation questions consists of several steps:

• Clarify goals and objectives of the evaluation

• Identify and involve key stakeholders and audiences

• Describe the intervention to be evaluated

• Formulate potential evaluation questions of interest to all stakeholders and audiences

• Determine resources available

• Prioritize and eliminate questions.

*Getting started right can have a major impact on the progress of an evaluation all along the way.*

Although it may sound trivial, at the outset of an evaluation it is important to describe the project or intervention briefly and clarify goals and objectives of the evaluation. Getting started right can have a major impact on  the progress of an evaluation all along the way. Patton (1990) suggests considering the following questions in developing an evaluation approach:

• Who is the information for and who will use the findings?

• What kinds of information are needed?

• How is the information to be used?  For what purpose is evaluation being conducted?

• When is the information needed?

• What resources are available to conduct the evaluation?

• Given the answers to the preceding questions, what methods are appropriate?

A critical component of clarifying goals and objectives is  the identification of the evaluation's focus.  Is the

evaluation to be Formative, looking at, for example, whether or not a teacher enhancement activity has been implemented as planned?  Or is it Summative, looking at the impact of the program on teaching practices and, ultimately, student learning?

Equally important is the identification of either stakeholders in the project or potential audiences for the evaluation information. In all projects, multiple audiences are likely to be involved. Being clear about your audience is very important as different audiences will have different information needs.  For example, the kinds of information needed by those who are concerned about the day-to-day operations of a project will be very different from those needed by policymakers who may be dealing with more long-term issues or who have to make funding decisions.

The next step is a goal-oriented description of the project including the rationale given for its existence and its goals and objectives as seen by the stakeholders.  The essence of the intervention  should also be documented: where it is situated, who is involved, how it is managed, and how much it costs.  An in-depth understanding of the intervention is usually necessary to determine the full range of evaluation questions.  This type of goal-centered description  is often a significant part of the evaluation effort.

After the purpose and stakeholders are identified and the project is described, specific questions about the project should be formulated. The process of identifying target audiences and formulating potential questions will usually result in many more questions than can be addressed in a single evaluation effort.  This comprehensive look at potential questions, however, makes all of the possibilities explicit to the planners of the evaluation and allows them to make an informed choice among evaluation questions.  Each potential question should be considered for inclusion on the basis of the following criteria:

• Who would use the information

• Whether the answer to the question would provide information not now available

• Whether information is important to a major group or several stakeholders

• Whether information would be of continuing interest

*It is critical to identify the major stakeholders, their questions, and their needs for information.*

- Whether it would be possible to obtain the information, given financial and human resources.

- Whether the time span required to obtain the information would meet the needs of decision makers.

These criteria determine the relevance of each potential question.

*A general guideline is for evaluations to be funded at about 5-10 percent of the total project cost.*

The final selection of questions depends heavily on the resources available. Some evaluation activities are more costly than others. For example, it may be that the only way to answer the question: "How has a project designed to enhance teachers' classroom activities affected classroom practices?" is through extensive classroom observations, an expensive and time-consuming technique. If sufficient funds are not available to carry out observations, it may be necessary to reduce the sample or use a different data collection method such as a survey. A general guideline is to allocate 5-10 percent of project costs for the evaluation of large-scale projects (those exceeding $100,000); for smaller projects, the percentage may need to be higher to meet minimum costs of fielding evaluation activities.

## How Do You Determine the Information-Gathering Techniques?

The next stage is the determination of the appropriate information-gathering techniques, including several steps:

- Select a general methodological approach

- Determine what sources of data would provide the information needed and assess the feasibility of the alternatives

- Select data collection techniques that would gather the desired information from the identified sources

- Develop a design matrix.

After the evaluation questions have been formulated, the most appropriate methods for obtaining answers must be chosen. In determining what approach to use, some initial questions need to be answered. First, is it better to do case studies, exploring the experiences

*If you are limited in your evaluation resources, it is best to stick to the simpler approaches.*

of a  small number of participants in depth or is it better to use a survey approach? In the latter case, do you need to survey all participants or can you select a sample? Do you want to look only at what happens to project participants or do you want to compare the experiences of participants with those of some appropriately selected comparison group of nonparticipants? How you answer some of these questions will affect the kinds of conclusions you can draw from your study. Rigorous, "controlled" designs are not always needed for Formative (Process) Evaluations, although they are always the preferred design. But Summative Evaluations, or Impact Assessments, gain a great deal from being based on experimental or quasi-experimental designs. For more information on this and the implications of different evaluation designs, see Cook and Campbell (1979).

Next you need to determine the kinds of data you want to use. Some alternatives are listed in Exhibit 2. Which one or ones to use depends on a number of factors, including the questions, the timeline and the resources available. Another factor to take into account is the technical skill level of the evaluator or evaluation team. Some of the techniques require more skills than others to design and analyze. If you are limited in your evaluation resources, it is best to stick to the simpler approaches. For example, observational techniques can produce a rich database which, analyzed properly, can be highly informative. The trick here is to design instruments which are either suitable for statistical analysis, or for other analytic strategies which have been developed for case study evidence (Yin, 1989). In the absence of careful advance planning for  the analysis, many an evaluator has wound up with a massive investment (both in time and in money) of data collected via observation that elude reasonable analysis.

Finally, you need to decide on the appropriate mix of data collection techniques, including both quantitative and qualitative  approaches.

In a broad sense, quantitative data can be defined as any data that can be represented numerically, whereas qualitative data are more frequently expressed through narrative description. Quantitative data also are useful in measuring the reactions or skills of large groups of people on a limited set of questions, whereas qualitative data provide in-depth information on a smaller number of cases (Patton, 1990). These distinc-

*Exhibit 2*

## Sources and Techniques for Collecting Evaluation Information

**I.  Data Collected Directly From Individuals Identified as Sources of Information**

A. Self-Reports: (from participants and control group members)

1. Diaries or Anecdotal Accounts

2. Checklists or Inventories

3. Rating Scales

4. Semantic Differentials

5. Questionnaires

6. Interviews

7. Written Responses to Requests for Information (for example, letters)

8. Sociometric Devices

9. Projective Techniques

B. Products from participants:

1. Tests

a.  Supplied answer (essay, completion, short response, and problem-solving)

b.  Selected answer (multiple-choice, true-false, matching, and ranking)

2. Samples of Work

**II.  Data Collected by an Independent Observer**

A. Written Accounts

B. Observation Forms:

1. Observation Schedules

2. Rating Scales

3. Checklists and Inventories

**III.  Data Collected by a Mechanical Device**

A. Audiotape

B. Videotape

C. Time-Lapse Photographs

D. Other Devices:

1. Graphic Recordings of Performance Skills

2. Computer Collation of Student Responses

**IV.  Data Collected by Use of Unobtrusive Measures**

**V.  Data Collected from Existing Information Resources**

A. Review of Public Documents (proposals, reports, course outlines, etc.)

B. Review of Institutional or Group Files (files of student records, fiscal resources, minutes of meetings)

C. Review of Personal Files (correspondence files of individuals reviewed by permission)

D. Review of Existing Databases (statewide testing program results)

From:  *Education Evaluation:  Alternative Approaches and Practical Guidelines.* by Blaine R. Worthen and James R. Sanders. Copyright ©1987 by Longman Publishing Group.

tions are not, however, absolute. Rather, they can be thought of as representing two ends of a continuum rather than two discrete categories. Furthermore, in some instances qualitative data can be transformed into quantitative data using judgmental coding (for example grouping statements or themes into larger broad categories and obtaining frequencies). Conversely, well-designed quantitative studies will allow for qualitative inputs.

Both types of data can provide bases for decisionmaking; both should be considered in planning an evaluation. And evaluations frequently use a mix of techniques in any one study. Further details on data collection and analysis techniques and the pros and cons of different alternatives are presented in Chapter Three of this Handbook.

Once these decisions are made it is very helpful to summarize them in a "design matrix." Although there is no hard and fast rule, a design matrix usually includes the following elements:

- General evaluation questions

- Evaluation subquestions

- Variables to be examined and instruments/ approaches for gathering the data

- Respondents

- Data collection schedule.

Exhibit 3 presents an example of a design matrix for a study of the effects of a teacher enhancement program.

## How Do You Conduct Data Collection?

Once the appropriate information-gathering techniques have been determined, the information must be gathered. Both technical and political issues need to be addressed. The technical issues are discussed in Chapter Three. The political factors to be kept in mind are presented below:

- Obtain necessary clearances and permission

- Consider the needs and sensitivities of the respondents

## Exhibit 3:  Summary of the Design for a Study of Project SUCCEED

| Question 1:  Did project SUCCEED change teachers' mathematics instructional practices? | | | |
|---|---|---|---|
| **Subquestion** | **Data Collection Approach** | **Respondents** | **Schedule** |
| 1a. Did teachers use different materials? | Questionnaire | Teachers Supervisors | Pre/post training |
| | Observation | NA | 3 x during year |
| 1b. Did teachers change  testing practices? | Questionnaire | Teachers | End of year |
| 1c. Was cooperative learning increased? | Questionnaire | Teachers | End of year |
| | Observation | NA | 3 x during year |
| Question 2: What impact did project SUCCEED have on teachers' use of planning time? | | | |
| **Subquestion** | **Data Collection Approach** | **Respondents** | **Schedule** |
| 2a. Did teachers spend more time plan- ning for instruction? | Questionnaire | Teachers | End of year |
| 2b. Did teachers de- velop lesson plans reflecting new approaches? | Questionnaire | Teachers | End of year |
| | Review of plans | NA | 3 x during year |
| 2c. Did teachers use the sample lessons as models? | Questionnaire | Teachers | End of year |
| | Review of plans | NA | 3 x during year |

• Make sure your data collectors are adequately trained and will operate in an objective, unbiased style

• Cause as little disruption as possible to the ongoing effort.

First, before data are collected, the necessary clearances and permission must be obtained. Many groups, especially school systems, have a set of established procedures for gaining clearance to collect data on students, teachers, or projects. This may include who is to receive/review a copy of the report, restrictions on when data can be collected, or procedures to safeguard the privacy of students or teachers. Find out what these are and address them as early as possible, preferably as part of the initial proposal development. When seeking cooperation, it is always helpful to offer to provide feedback to the participants on what is learned. Personal feedback or a workshop in which findings can be discussed is frequently looked upon favorably. If this is too time-consuming, a copy of the report or executive summary may well do. The main idea here is to provide incentives for people or organizations to take the time to participate in your evaluation.

*The main idea here is to provide incentives for people or organizations to take the time to participate in your evaluation.*

Second, the needs of participants must be considered. Being part of an evaluation can be very threatening to participants. Participants should be told clearly and honestly why the data are being collected and the use to which the results will be put. On most survey type studies, assurances are given and honored that no personal repercussions will result from information presented to the evaluator and, if at all possible, individuals and their responses will not be publicly associated in any report. This guarantee of anonymity frequently makes the difference between a cooperative and a recalcitrant respondent. There may, however, be some cases when identification of the respondent is deemed necessary, perhaps to enforce the credibility of an assertion. In such cases, the evaluator should seek informed consent before including such information. Informed consent may also be advisable where a sensitive comment is reported which could be identified with a given respondent, despite the fact that the report itself includes no names. Common sense is the key here.

Third, data collectors must be carefully trained and supervised, especially where multiple data collectors are used.  They must be trained to see things in the same way, to ask the same questions, to use the same

prompts. Periodic checks need to be carried out to make sure that well-trained data collectors do not "drift" away from the prescribed procedures over time. (More details on training of data collectors are presented in Chapter Three.)

*Checks need to be carried out to make sure that well-trained data collectors do not "drift" away from the prescribed procedures.*

In addition, it is important to guard against possible distortion of data due to well intentioned but inappropriate "coaching" of respondents —an error frequently made by inexperienced or overly enthusiastic staff. They must be warned against providing value-laden feedback to respondents or to engage in discussions that might well bias the results. One difficult but important task is understanding one's own biases and making sure that they do not interfere with the work at hand. This is a problem all too often encountered when dealing with volunteer data collectors, such as parents in a school or teachers in a center. They volunteer because they are interested in, advocates for, or critics of, the project that is being evaluated.  Unfortunately, the data they produce may reflect their own perceptions of the project, as much or more than that of the respondents, unless careful training is undertaken to avoid this "pollution."  Bias or perceived bias may compromise the credibility of the findings and the ultimate use to which they are put. An excellent source of information on these issues is the section on accuracy standards in Standards for Evaluation of Educational Programs, Projects and Materials (Joint Committee on Standards for Educational Evaluation, 1981).

Finally, the data should be gathered causing as little disruption as possible. Among other things, this means being sensitive to the schedules of the people or the project, as well as the schedule of the evaluation itself. It also may mean changing approaches as situations come up.  For example, instead of asking a respondent to provide data on the characteristics of project participants—a task that may require considerable time on the part of the respondent to pull the data together and develop summary statistics—the data collector may have to work from raw data, applications, monthly reports, etc. and personally do the compilation.

## How Do You Analyze the Data?

Once the data are collected they must be analyzed and interpreted. The steps to be followed in preparing the data for analysis and interpretation differ, depending on the type of data. The interpretation of qualitative data may in some cases be limited to descriptive

narratives, but other qualitative data may lend themselves to systematic analyses through the use of quantitative approaches such as thematic coding or content analysis. Analysis includes several steps:

• Check the raw data and prepare data for analysis

• Conduct initial analysis based on the evaluation plan

• Conduct additional analyses based on the initial results

• Integrate and synthesize findings.

The first step in quantitative data analysis is the checking of data for responses that may be out of line or unlikely. Such instances include: selecting more than one answer when only one can be selected; always choosing the third alternative on a multiple-choice test of science concepts; reporting allocations of time that add up to more than 100 percent; inconsistent answers, etc. Where such problematic responses are found, it is frequently necessary to eliminate the item or items from the data to be analyzed.

*Data analysis is an iterative process.*

After this is done, the data are prepared for computer analysis; usually this involves coding and entering (keying) the data with verification and quality control procedures in place.

The next step is to carry out the data analysis specified in the evaluation plan. While new information gained as the evaluation evolves may well cause some analyses to be added or subtracted, it is a good idea to start with the set of analyses that seemed to be of interest originally. For the analysis of both qualitative and quantitative data there are statistical programs currently available on easily accessible software that make the data analysis task considerably easier today than it was 25 years ago. These should be used. Analysts still need to be careful, however, that the data sets they are using meet the assumptions of the technique being used. For example, in the analysis of quantitative data, different approaches may be used to analyze continuous data as opposed to categorical data. Using an incorrect technique can result in invalidation of the whole evaluation project. (See Chapter Three for more discussion of alternative analytic techniques.)

It is very likely that the initial analyses will raise as many questions as they answer. The next step, there-

fore, is conducting a second set of analyses to address these further questions. If, for example, the first analysis looked at overall teacher performance, a second analysis might want to subdivide the total group into subunits of particular interest—i.e., more experienced versus less experienced teachers—and examine whether any significant differences were found between them. These reanalysis cycles can go through several iterations as emerging patterns of data suggest other interesting avenues to explore. Sometimes the most intriguing of these are results which emerge from the data; ones that were not anticipated or looked for. In the end, it becomes a matter of balancing the time and money available, against the inquisitive spirit, in deciding when the analysis task has been completed.

The final task is to choose the analyses to be presented, to integrate the separate analyses into overall pictures, and to develop conclusions regarding what the data show.  Sometimes this integration of findings becomes very challenging as the different data sources do not yield completely consistent results. While it is always preferable to produce a report that is able to reconcile differences and explain apparent contradictions, sometimes the findings must simply be allowed to stand as they are, unresolved and thought-provoking.

## How Do You Communicate Evaluation Results?

The final stage of the Project Evaluation is reporting what has been found. While reporting can be thought of as simply creating a written document, successful reporting rests on giving careful thought to the creation and presentation of the information. In fact, while funding agencies like NSF require a written report, many projects use additional strategies for communicating evaluation findings to other audiences.

The communication of evaluation findings involves several steps:

• Provide information to the targeted audiences

• Customize reports and other presentations to make them compelling

• Deliver reports and other presentations in time to be useful.

Providing the evaluation information should not present a problem if the evaluation has been successful so far, and if some simple steps are followed. Again, special attention should be given to the stakeholders and the constructive part they can play. The specification of questions and selection of data-gathering techniques should have already involved the stakeholders so that the information should be relevant and important to them. By also involving the stakeholders at the end of the study, the utility and probable attention given to the evaluation findings are sure to be increased. One way of accomplishing this is through a pre-release review of the report with selected stakeholder representatives. Such a session provides an important opportunity for discussion of the findings, for resolving any final issues that may arise and for setting the stage for the next steps to be taken as a result of the successes and failures that the data may show.

*By involving the stakeholders at the end of the study, the utility and attention given to the evaluation findings are sure to be increased.*

Second, the information must be delivered when it is needed. Sometimes there is leeway in when the information will be used; but the time of decision-making is often fixed, and information that arrives too late is useless. There is nothing so frustrating to a Principal Investigator than being told by a funding agency or community group:

**"Oh, I wish I had known that two months ago!  That's when I had to make some decisions about the projects we were going to support next year."**

Our earlier discussion stressed the importance of agreeing up front what is needed and when the needs must be met. As the evaluation is carried out, the importance of meeting the agreed-upon time schedule must be kept in mind.

Finally, the information needs to be provided in a manner and style that is appropriate, appealing, and compelling to the person being informed. For example, a detailed numerical table with statistical test results might not be the best way to provide a school board member with achievement data on students. Different reports may have to be provided for different audiences.  And, it may well be that a written report is not even the preferred alternative. While most evaluations will include some written product, other alternatives are becoming increasingly popular.

It should be noted that while discussions of commu-

nicating study results generally stop at the point of presenting a final report of findings, there are important additional steps that should be considered. Where a new product or practice turns out to be successful, as determined by a careful evaluation, dissemination is an important next step. This topic is covered in a separate NSF publication.

## Summary

There are several phases to conducting and implementing an evaluation. No one stage is more important than the rest. And, as can be seen from the discussion of the role of the stakeholders in both the first step—developing questions—and the last—provision of information—the groundwork laid at the earliest stages can have important implications for the success of the evaluation in the long run.

*While there are technical skills needed to do an evaluation, there is also "common sense" involved.*

Evaluation isn't easy, but there also is very little mystery about the steps that need to be taken and the activities that need to be carried out. While there certainly are technical skills needed to do an evaluation that is helpful and credible—and that is why trained evaluators are important—there is also a lot of "common sense" involved. Sound advice is to blend these two factors—technical skills and common sense. In the best evaluations, both of these inevitably exist.

# REFERENCES

Cook, T. D. & Campbell,  D. T. (1979). *Quasi-experimentation: Design and Analysis Issues or Field.* Settings, Chicago, IL:  Rand McNally.

Joint Committee on Standards for Educational Evaluation. (1981). *Standards for Evaluation of Educational Programs, Projects, and Materials.*  New York, NY: McGraw-Hill.

Patton, M. Q. (1990). *Qualitative Evaluation and Research Methods.* Newbury Park, CA:  Sage.

Worthen, B. & Sanders. J. (1987). *Educational Evaluation:  Alternative Approaches and Practical Guidelines.*  New York: Longman, Inc.

Yin, R. (1989). *Case Study Research.*  Newbury Park, CA: Sage.

## *Tips for Conducting an Evaluation*

*1.    Develop Evaluation Questions*

♦    **Clarify goals and objectives of the evaluation.**

♦    **Identify and involve key stakeholders and audiences.**

♦    **Describe the intervention to be evaluated.**

♦    **Formulate potential evaluation questions of interest to all stakeholders and audiences.**

♦    **Determine resources available.**

♦    **Prioritize and eliminate questions.**

*2.    Match Questions with Appropriate Information-Gathering Techniques*

♦    **Select a general methodological approach.**

♦    **Determine what sources of data would provide the information needed.**

♦    **Select data collection techniques that would gather the desired information from the identified sources.**

*3.    Collect Data*

♦    **Obtain the necessary clearances and permission.**

♦    **Consider the needs and sensitivities of the respondents.**

♦    **Make sure data collectors are adequately trained and will operate in an objective, unbiased manner.**

♦    **Cause as little disruption as possible to the ongoing effort.**

*4.    Analyze Data*

♦    **Check raw data and prepare data for analysis.**

♦    **Conduct initial analysis based on the evaluation plan.**

♦    **Conduct additional analyses based on the initial results.**

♦    **Integrate and synthesize findings.**

*5.    Provide  Information to Interested Audiences*

♦    **Provide information to the targeted audiences.**

♦    **Deliver reports and other presentations in time to be useful.**

♦    **Customize reports and other presentations.**