



Chapter 4

EVALUATION

Evaluations are useful in making decisions about the value of a program or the effectiveness of a technique for delivering services. By assessing programs and improving them, a community can greatly increase the quality of service delivery. The community-conducted evaluation can demonstrate program progress according to the goals and views of the community itself. Since most programs are eventually evaluated by outside persons or agencies, an organization's own evaluation is one safeguard against a possible culturally-biased evaluation later on. And, for those programs developed completely within the community, evaluations can be used to increase the effectiveness of community resources.

Evaluation is a powerful research tool for improving or judging the effectiveness of new or existing approaches to community development. It is a process that can bring about improvement or change by providing information for decision making. Conducting an evaluation involves looking at and assessing a program or a concept, and comparing it to some standard. Or, it might measure the degree to which the program objectives were accomplished, and compare these accomplishments to certain criteria. Since evaluation is a technique that lends itself easily to cultural bias, the community-based effort is particularly valuable in structuring the culturally appropriate evaluation and avoiding such bias.

Negative impressions sometimes come from an outside evaluator who disrupts a program and passes on culturally inappropriate judgements that determine the fate of a project. And, in some cases, these impressions may be true. The best way to ensure that such unfortunate experiences don't happen is through the community-based evaluation. On the community level, evaluation activities can vary from long-term, continuing assessments of programs to short-term, cooperative assessments conducted with an outside evaluator. The outside evaluator may have some advantages in being objective or less firm in opinion because he or she is not directly involved with a program. The insider often is more familiar with the cultural meaning of activities, the real difficulties in carrying out program objectives, and has interest in applying the evaluation results.

These are a few examples of topics for evaluation:

- * Social service delivery techniques
- * Program management
- * Cost-effectiveness
- * Staff development

Particularly of high priority today is the testing of service delivery techniques that combine dominant-culture program structures with traditional Indian approaches. Evaluation is a core component of demonstration projects where a priority is put on finding out the effectiveness of the new program in comparison to those already tried. For example, many alcoholism and drug treatment programs are currently testing the effectiveness of traditionally-based methods of treatment within the structure of federally-funded projects. Mental health and educational programs are also testing new alternative methods that were previously not allowed by funding agencies. A successful research and demonstration project can frequently open the door for more culturally appropriate program structures in the future.

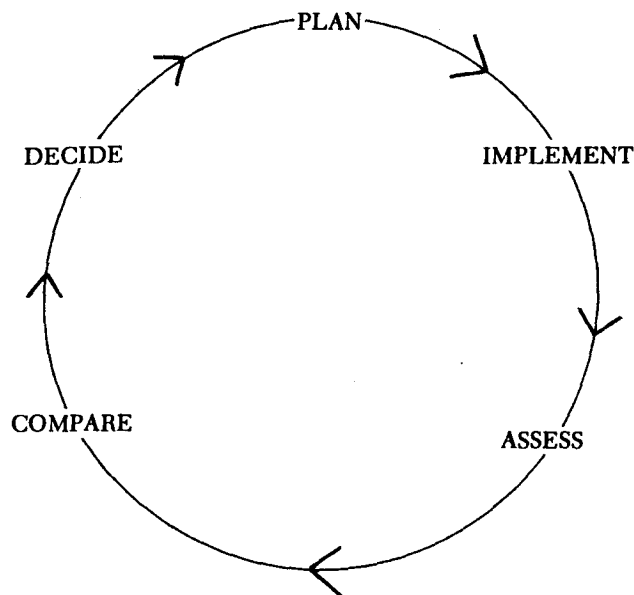
An important question to keep in mind before conducting an evaluation is whether the evaluation refers to this one particular program only, or



"The evaluation is complete, but they asked the wrong questions!"

whether the results can be generalized to a larger population. If the researcher draws conclusions that will extend to the tribe, region, or American Indian population as a whole, then the results could have an impact upon other communities. The studies of many research and development projects are intended to form general statements. Ask if this is what you want to do before undertaking a project. The basic steps for evaluation described below apply to program evaluation, and then these steps are carried further in the section on research and evaluation. Research and demonstration can be extremely valuable to the future of other programs, yet great care must be taken in defining variables and following the original research plan.

Figure 4.2 THE EVALUATION PROCESS



There are two basic types of evaluation relating to purpose, formative and summative¹ evaluations:

Formative evaluation is an ongoing process throughout the development of a project. It is usually conducted by the program staff or else by an evaluator working closely with the program. Data are collected, while the program is in process, to determine the continual effective-

ness of the program and to make changes in the program activities if necessary. Using this approach can save resources and improve results, by "forming" the program as effectiveness is determined. Results are reported directly to the community organization.

Summative evaluation occurs after the program is final or completed, although pretesting of the data collection instruments and the data collection usually occur while the program is in progress. The results of a summative evaluation are often used in making decisions as to keeping a particular program structure for future use. The summative evaluation is particularly useful in comparing different types of programs and deciding which type is the most effective. For example, a demonstration project may compare the use of dominant-culture therapies to the use of Indian traditional therapies and evaluate the effectiveness of both on a particular treatment group or similar groups. In addition to cost-effectiveness, the long-term success rate of treatment is an important factor in the evaluation.

Once the type of evaluative approach is determined, the next question may concern who will conduct the evaluation. The decisions usually involve whether a person internal to the program or whether an external or outside person will be hired. One major advantage of an outside evaluator is the increased chance that a person not involved with the outcomes or the political nature of the evaluation is likely to be more objective. Outside consultants may have specialized skills not readily available with the staff and may relieve the staff of the time and stress involved in conducting the evaluation. A major disadvantage of hiring an outside evaluator is that the person may not have the background needed to understand the goals of the project. This is even more likely when the evaluator is not a member of the culture being evaluated. The costs of hiring an external evaluator can be higher than when the evaluation is conducted internally. And, outsiders are sometimes seen as a threat, thus reducing staff cooperation on the evaluation (although this can also be true of an insider). Another disadvantage to the outside evaluator is a possible lack of continuity in applying the findings after the evaluation is completed.

The formative evaluation can be conducted with an outside evaluator only if close contact is maintained as the program progresses. If the evaluation is conducted internally, generally a program developer or a planner is the chosen person. Weighing the advantages and disadvantages of outside consultants versus capabilities of staff on board should lead to that decision of the right person to conduct the evaluation.

STRUCTURING FOR EVALUATION

An in-house evaluation design can be of great benefit in ensuring against an after-the fact, culturally biased evaluation. The first step toward accomplishing later evaluations comes with the stage of program planning, that is, in setting down a clear set of goals and objectives for the program. This enables progress to be measured later on. To review the discussion of goals and objectives in the first chapter, goals are long-term general outcomes and objectives are short-term and more specific. Objectives are often the specific steps that will be taken to accomplish the goal.

When goals, objectives, and methods of program delivery are set out at the start of the program, the staff is given guidelines for conducting the work. This planning has several advantages. First, the staff have a work plan to follow. This overview serves as staff training and helps reduce misunderstandings as to the activities the program should support. Secondly, the program director can review the objectives periodically to decide if progress is being made or if the objectives need to be modified. When changes or modifications are made, the reasons for the changes should be documented for later evaluation purposes. And thirdly, when goals and objectives are defined, there is less chance of a misunderstanding at the end of the project for all parties concerned. Those involved in this planning process may include the program staff, management (tribal or executive director overseeing several projects), community members, funding agencies, and evaluators. In other words, when objectives are defined, reviewed regularly, and changes are reported, then the chance of a misunderstanding over the program activities is greatly reduced.

If a formative evaluation is not conducted as a component of the program, a regular review of goals and objectives provides a kind of informal evaluation that can document changes. This process has immediate rewards in furthering a cooperative staff effort. When documented, a review also provides valuable information for a more formal evaluation later. In addition to reviewing objectives, another program procedure that greatly aids the evaluation process is the collecting of program data in a regular and complete way. Records can be developed according to a set structure or format, which document client characteristics, services delivered, program expenses, and case histories. Such records can provide both hard data, or measurements, and soft data, such as descriptions, for a later evaluation.

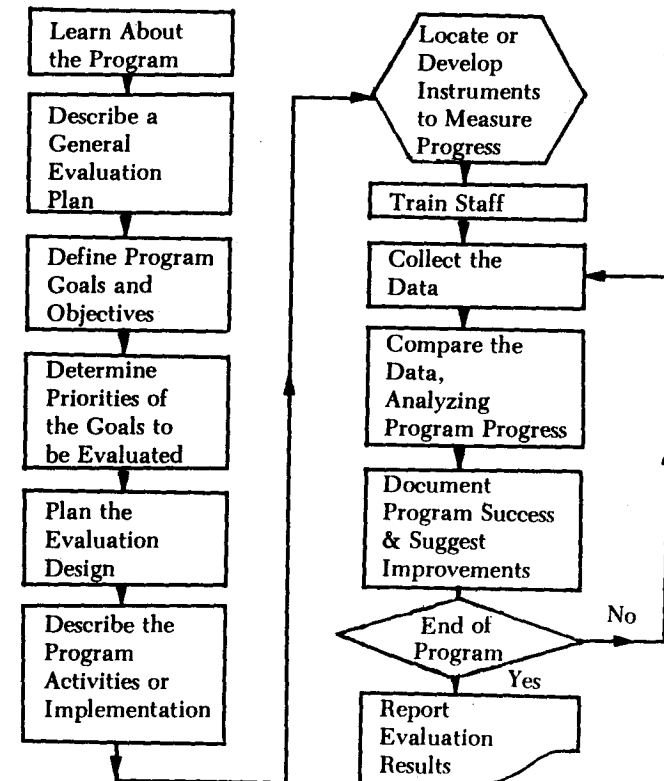
In the ideal evaluation setting, the complete evaluation is designed as part of the initial program plan. In other words, the program objectives are well defined, the methods for collecting measurements or data are developed, and the criteria for comparison or determining success are defined. This planning is usually completed by the program developer or

through cooperative efforts between evaluator and program developer. Although not an easy task to accomplish (and often not accomplished) in the program planning stage, this type of evaluation provides the maximum feedback to the program and establishes a support system to improve program effectiveness.

STEPS FOR CONDUCTING AN EVALUATION

Let us first look at an overview of the evaluation process. The formative approach to assessing performance or program impact as an ongoing process starts with program planning. Once the plan is implemented, comparisons are made, progress assessed, and decisions made. These decisions then can feed into the process of continued program planning. The diagram below² describes the steps for conducting a formative evaluation.

Figure 4.3 STEPS FOR CONDUCTING AN EVALUATION



Step 1: The first step in conducting an evaluation is **learning about the program**. This can be accomplished by reading through any written documentation, such as proposals, reports, or statements of purpose. In the ideal situation, the evaluation is designed by the program developer or by the evaluator as part of the program. In the situation where the program already exists, talking with the program staff is beneficial in learning about the program activities and possible tensions. Learning who requested the evaluation can alert the evaluator to sources of cooperation and possible sources of resistance. For example, the evaluation requested by a board of directors due to suspected poor program performance may not be a popular idea with the program director. On the other hand, the staff of a demonstration program may know that a successful program effort would pave the way for future programs and thereby would be eager to measure or demonstrate this success. In finding out about the program, the evaluator can ask about his role with the program. Is there a good chance that the results will be applied? How does the program see the responsibilities of the evaluator in assisting the program staff and in the collecting of data? Coming to an agreement on services and responsibilities early in the program can reduce tensions later on. And clear understandings are essential in estimating the cost of the evaluation.

Step 2: Secondly, describe a general evaluation plan for the program. There may be ways, unknown to the program staff, that the evaluation can benefit the program. Time may be required to collect data that the program staff is unaware of at this point in the evaluation. The effort taken initially in becoming familiar with the program can help define the time involved in carrying out the evaluation, such as working with the staff to define goals if these are not already present. Staff training in data collection procedures may be another time consuming need during the evaluation. Point out both the advantages and disadvantages of the evaluation, such as possible improvements versus extra responsibilities. Estimate some of the hidden costs of the evaluation, for instance, the staff time contributed. And a budget should be proposed before the start of the evaluation, so that any unrealistic expectations can be cleared up in advance. Modifications to the services and responsibilities may be necessary if funds are limited.

Step 3: The third step is **defining the program goals and objectives**. Program goals are often too general, and the evaluator may need to serve as an interpreter from the general to the more specific. For example, a very general goal might be to improve the educational opportunities in a community. A more specific goal might be to create an educational program that would increase certain skills. Becoming specific in defining goals is necessary in deciding the indicators for goal accomplishment.

In other words, to measure the effectiveness of a program in accomplishing a goal, there has to be a specific goal. Four characteristics of the well-defined goal³ are that the terms should be: 1) clear 2) specific 3) measurable 4) behavior-oriented. Goals are stated clearly, avoiding jargon or very general terms. For example, a goal that is not measurable would be: "To improve the educational level attained by community members." A clear, measurable goal would be: "To raise the level of educational attainment to a high school diploma or GED, from 40% to 75% of the community." Specific terminology indicates who will benefit from the program, for example, through increased employment or lowered cost of service delivery. If the goals are clear and specific, then the evaluator should be able to measure the progress toward accomplishment. And if the goals are directed toward the changes expected in the participants, then they are behavior-oriented, and the change to be measured is known. Objectives, or the more specific steps that will accomplish the goals, also need to be written in clear and measurable terms. Cooperative efforts in defining goals well can contribute to a quality evaluation that can be applied for program improvement.

Step 4: **Determining the priorities of the goals for evaluation purposes** is the next step. Although the program may have several goals, only one or a few may be the most preferred or even measurable. Also, funds may require that the evaluated goals are limited. Evaluation can be expensive, and one way to ensure that the funds do not run out before the task is done is to limit the task realistically. The cost of an evaluation generally totals five to ten percent of a project budget, with the percentage becoming smaller as the project size increases. By helping the program director to set priorities, or determine the most important goals for the evaluation, the evaluator can assist in the setting of realistic expectations. Deciding on which goals will be evaluated is often known as setting the boundaries for the evaluation.

Step 5: **Choosing the evaluation design** is an important step in planning the evaluation. There are several methods of gaining comparison, which are decided during the design of the evaluation project. One design, often considered the ideal, is the controlled experiment. With this design, similar groups are selected at the beginning of the evaluation and the same kind of data is collected for both groups. Usually these groups have similar characteristics but the services delivered are different. This enables the evaluator to compare the success of one service delivery strategy against another with similar groups of clients. Detailed steps for conducting an experimental design follow under the "Research and Evaluation" section. One important consideration when using this method in small communities is the difficulty in finding similar groups.

A second comparison design is the before vs. after, or time-series comparison of a particular program. This design can compare the assessment of clients before a program delivers services with an assessment after services are delivered or at some intermediate time, showing the changes in the client group. This technique works the best when a program is short, for the program conditions are less apt to change. Another way of using this design is the comparison of planned vs. actual performance of the program. This can be shown through client changes and program activities. A third comparison design is that of comparing client progress with other population segments not served by the program. For example, the rate of improvements in the client group may be compared generally to other similar groups. This approach differs from the experimental one, in that actual data are not collected for the other groups by the evaluation study. Although much less expensive and time-consuming, this approach is usually difficult due to the lack of groups that really compare in client characteristics and sampling procedures used to form the comparison group. It is particularly important that differences in the groups are identified and reported.

The summative evaluation may use the less desirable, but sometimes necessary retrospective design, where data are gathered after the program is over. This is often accomplished by talking with program staff, and carries the weakness of relying on people's memory of the activities. Records kept during the program for the purpose of its management are another source of data for the retrospective design. If the data were not collected earlier, this design is better than no evaluation at all. Examining case studies is another means of obtaining data on participant characteristics and services delivered during the course of the program.

Step 6: The sixth step involves describing the program's activities, or implementation to see if the planned activities actually took place. This description might include details on the program's organizational structure, staff qualifications, staff training provided, staff turnover, the program setting, the program participants, and the services delivered. One of the program's activities may be the initial client interview to gather data on the participants. This information is very important to the program in deciding which services will be delivered to the client or participant. The characteristics of clients vary, and are therefore known as variables. For example, client variables could include tribe, age, sex, educational level, employment, residence, native language usage. The data used to assess the client upon entry to the program are called baseline data, and are very important later on when outcomes are measured against goals to determine if, and how much, change has occurred. The formative evaluation is often concerned with monitoring the implementation of the program.

Step 7: The next step is deciding how progress will be measured. This decision requires that instruments be located or developed to measure outcomes. These instruments are usually intended to assess outcomes or the changes in program participants. The term "instrument" refers to a tool for collecting the data according to a set format. This enables a consistent set of data to be collected for each participant. Some people may have a negative reaction to the word "instrument," in that it may imply more specific measurements (as in the sciences) than are actually possible when assessing human behavior. As long as the evaluator and the program staff understand the measurements are in most cases only approximate, then the flexibility of the evaluation should not be affected by using this term. In most fields, instruments exist that have already been tested for accuracy, completeness, and ease in gaining rapport with a client. If such an instrument can be located, it saves a great deal of time and money that would be required for development and testing. The problem that Native American communities often encounter with such instruments is that they lack cultural information or sensitivity. One way of overcoming this is to add a cultural section to the instrument that describes such variables as, tribe, blood quantum, type of residence, native language usage, religious belief, type of healing preferred, participation in traditional activities, and other cultural variables relevant to the project being evaluated. If the basic instrument does not contain questions that are offensive or insensitive, then the addition of a cultural section may provide the necessary extra information to enable a breakdown of the basic variables by the important cultural variables. Too often a data collection system is imposed upon a program by a funding source, and that system does not always gather even the most important cultural variables that could be used by the staff for service delivery! More detailed information on instruments is presented in the next section.

Step 8: Data collection is another important step in the evaluation process. Once the format or instrument for data collection is determined, a method for collecting the data in a continual and consistent way is decided. The quality of the results depends heavily on these two factors, continuity and consistency. This decision means that a schedule for completing the data needs to be met and that the data needs to be collected in the same or consistent manner. A good staff training program, with the major points for data collection written as a short manual that is frequently reviewed, works well for most programs where staff are responsible for collecting the data. One way in which community-based, data collection efforts sometimes run into trouble is through heavy staff turnover and inadequate training of new staff, resulting in different interpretations of the variables or instrument questions. Another common difficulty is the lack of a schedule for coordinating the collection of the data,

resulting in incomplete data collection. An important part of the data collection process is conducting periodic inventories to see if all the data have been collected, for, once the client or participant has left the program, the opportunity to gather the data are lost. Also, the time at which the data are gathered is critical, for a client may answer questions differently after being in the program a while than when he would have a the time of admission to the program.

Step 9: Although the method for comparison is decided before the data are collected as part of the evaluation design, the actual comparison of data occurs after the data are collected. The determination of program success rests on the comparisons that show whether progress is occurring. Since comparison is the key to accurate interpretation, it is accomplished by comparing the changes in the values of the criteria before, during, and after the program. In determining program success or failure, it is important to remember that there can be many different reasons for failure to accomplish program objectives. The reason that a new therapeutic technique did not succeed, for example, could be due to staff failure to implement the technique properly rather than due to shortcomings on the part of the technique itself. Reasons for poor program performance could vary from poor management to unavailability of adequately trained staff. Non-program factors need to be sorted out from program factors affecting success. In other words, the questions to be answered focus not only on the results but also on whether it is the program that is causing them. Comparison does not end with a statistic, but should pursue further the cause for success or failure. Such in-depth searching is essential in light of the enormous responsibility an evaluator has on the future of other similar programs.

Step 10: As a final step, applying the results usually involves: (1) the evaluator making recommendations about the program, (2) the administrator making decisions about the program, and (3) the administrator implementing changes in the program. With the formative evaluation process, the evaluator makes comparisons on a continual basis and reports to the administrator for possible improvement in the program. In the summative evaluation process, the evaluator's recommendations describe success at the end of the program.

The evaluation planning checklist,⁴ Figure 4.4 summarizes the choices that are presented in this chapter. This list may be valuable for your organization to copy and use.

Figure 4.4 EVALUATION PLANNING CHECKLIST

Project Name: _____ Date: _____

Check all items that apply to this project

1. Audience
 - Internal
 - Staff
 - Administrators
 - Board of Directors
 - External
 - Funding source
 - Community
 - Professional
 - Legislature
 - Public
2. Purpose
 - Feasibility
 - Improvement
 - Efficiency
 - Adoption
 - Dissemination
3. Type
 - Formative
 - Summative
4. Design type
 - Experimental
 - Time series
 - Retrospective
 - Case study
 - Other _____
5. Who does the evaluation
 - Internal
 - External
6. Subjects
 - (List specific groups)
7. Timing of data gathering
 - Pre-test
 - Interim-tests (during program)
 - Post-test
 - Follow-up tests
8. Instruments to collect data
 - Inventories
 - Records
 - Tests
 - Performance tests
 - Questionnaires
 - Interviews
 - Observations
9. Responses
 - Qualitative
 - Quantitative
 - Closed format
 - Open format
10. Analysis of data
 - Means, percentiles, etc.
 - Standard deviation, variance
 - Crosstabulations
 - Tables, graphs, etc.
 - Inferential statistics
11. Dissemination of results
 - Timing
 - Progress
 - Final
 - Media
 - Oral
 - Written
 - Other _____
 - Formality
 - Informal
 - Formal
12. Budget

INSTRUMENTS FOR EVALUATION

The instruments or tools for collecting data can vary greatly depending upon the purpose of the evaluation and the resources for data collection. Formats can vary from a set of highly structured questions to open-ended instruments that allow individual opinions to be recorded. Open-ended instruments can be particularly valuable for formative evaluations, in that they provide a way to ask participants for suggestions on improving the program.

Evaluation data can be obtained through using the following:

- * Interviews
- * Observations
- * Analysis of documents
- * Questionnaires
- * Ratings or rankings (by staff, participants, managers)
- * Journals by staff, or case histories of participants
- * Organizational record keeping
- * Clinical examinations
- * Information tests
- * Standardized attitude inventories
- * Ability tests
- * Observation

In review of the detail on instruments given in Chapter 3, the questionnaire and the interview are both usually a structured set of questions. In use of the questionnaire, the participant receives the questions on paper and indicates the answers on the paper; whereas, in the interview setting, the questions are asked by an interviewer and the responses are noted by that person. The disadvantages of the questionnaire method are that the process of data gathering is less personal, and that the reliability of the data may be unknown due to possible respondent difficulties with interpreting the questions. Literacy levels are an important consideration in deciding whether to use the questionnaire in native communities. An advantage of the questionnaire method of data collection is the privacy that it affords to the respondent. Privacy and confidentiality can be particularly valuable in obtaining health data. Advantages of the interview schedule lie in the personal contact by the interviewer. The person conducting the interview can interpret the questions in the tribal language if necessary (with training to ensure that all interviews interpret the same way) and can collect additional information if the person interviewed offers a more lengthy response.

The evaluator can analyze documents, such as program reports and proposals to identify program goals and obtain baseline data in retrospect. In an ideal situation, the evaluator helps with the development of

the record-keeping system, for example with a management and information system (MIS) that documents data for management purposes in addition to client progress. Organizational records can yield data on participants and participation as well as the details of program activities. Cost-benefit evaluations can greatly utilize the good organizational record keeping system. And a last advantage of this type of data collection is that the program keeps data over a long period of time.

At specified intervals, the evaluator can conduct ratings with the potential for suggesting improvement. Staff suggestions may be invaluable, particularly if program methods seem to be culturally inappropriate. Ratings are a useful technique for the formative evaluation. Another source of data, journals kept by staff or participants, is useful in documenting the qualitative side of progress and seeing suggestions for program improvements. The evaluator's journal is extremely useful in recording program changes and interpretations.

Clinical examinations are useful records for health program and substance abuse evaluations. When information is collected in a systematic way (not always the case with standard medical records--a new form may be needed) at the pre-program and post-program intervals, there may be a concrete basis for measuring progress.

Existing information tests, attitude inventories, and ability tests are particularly likely to contain cultural bias. These should be examined with care by a culturally balanced committee before use and either modified or rejected if bias exists in the instrument. Many research and evaluation projects have the testing of instruments as a critical problem to address, and consideration of this as a component of the research design can save interpretation difficulties later on. If expertise or funds are a problem in testing instruments, it may be possible to locate another research effort that will cooperate in the development and testing of instruments. Additional resources can often be located by getting in touch with other programs with similar funding and purposes.

Comparison is the critical factor in collecting data for evaluation purposes. A balanced design often includes both instruments to gather hard data and soft data. Quantitative or hard data (number and measurements) are more exact and can be less political. Instruments to collect qualitative data, such as descriptions and examples, can capture some of the more difficult measures of program success, like, for instance, the cultural appropriateness of services delivered. Evaluations vary from the use of one instrument to several, depending on the purpose of the evaluation and the resources available.

RESEARCH AND EVALUATION

Research projects using an evaluation methodology are usually intended to test a new technique for service delivery or treatment and to gener-

alize about its effectiveness for widespread use. Because of the generalizations that may be drawn from the research conclusions, a great deal of emphasis is usually placed on control of the conditions in the program. The standard research approach, then, is usually summative in nature and not flexible to change in services, for change in conditions affects the ability to measure client progress.

The basic research method is the experimental model. Two identical groups are set up or selected for the research. With the ideal experimental design, participants are chosen at random (see STATISTICS chapter for method) for the two groups. This method of selecting participants is not always possible in the real world, particularly when the participants are chosen from treatment groups. During the project, the experimental group would receive the treatment or the special services being tested, while the control group would not receive the special services. The step (in addition to the basic steps of developing a research design) for a controlled experiment usually involve the following:

- Step 1: Measuring the pre-program performance or collecting baseline data;
- Step 2: Applying the new program activities to the experimental, but not to the control group;
- Step 3: Monitoring the program activities to see if the planned program activities actually were carried out;
- Step 4: Measuring the post-program performance of each group on the goals and objectives;
- Step 5: Comparing the pre-program conditions to the post-program performance in order to identify changes for each of the evaluation criteria, or comparing post-performance between two groups;
- Step 6: And seeking explanations for differences in the two groups, rather than the treatment or special services, to see whether there are factors affecting change other than the intended special program activities. For example, there may be changes in the environment (health factors, economic factors, political factors), or other changes in the participants over time. When participants are aware of the evaluation in progress, there may be a special incentive for improvement of performance or conditions that can have a more short-term effect.

One major difficulty with the experimental model is the assumption that participants in the research effort will be chosen at random. In real-life program situations this is often impossible, for the clients of a program are usually those who seek out the treatment or services. In communities where programs are often lacking, the ideal of locating an experimental group and a control group that are identical may well be impossible. If two groups are set up for the duration of the research project, this is a luxury situation due to high costs. The lack of flexibility with the experimental approach is often a cause of tension between the research and program staff members. Control groups may be easier to locate in educational programs than in any other evaluation setting.

An alternative to the experimental, the quasi-experimental model is more often used in an applied setting where the variables and conditions cannot be as rigidly controlled. This research design follows the steps of the experimental design, but indicates where the limitations of the study exist when controls could not be met. Such documentation as to the weaknesses and validity of the study are important if generalizations are to be made.

When control group is not available, the time series or before versus after design mentioned above is an alternative. Planned vs. actual performance comparisons, the retrospective design, or the case study approach are other possibilities. Again, if the evaluation is a research and demonstration project with generalizations expected to result, then extra care must be taken in following the research design and in collecting the data. Staffing considerations are particularly important in demonstration projects. Extra funds to hire qualified staff can make the critical difference to which a trained staff contributes in a research effort.

Since research and evaluation designs are often quite complex, more details cannot be presented in this text. We recommend the sources listed at the end of the chapter for additional details and examples.

In summary, research is an important means of initiating change, for, rather than accepting what programs are available, new concepts can be tested and their effectiveness documented. Although research is a demanding task, the efforts of clearing a new path for improved resources to the community is well worth the efforts for the coming generations.

PRESENTING THE EVALUATION

An important step in completing a useful evaluation is communicating the results to those involved in the evaluation. Careful and thorough presentation is an important part of remaining objective during the evaluative process. For example, the summative evaluation can have an impact on the decision to continue the program. Written reports of the evaluation are more useful when all of the data are presented. Through the use of tables and displays, these reports can be summarized for con-

cieness, yet the availability of certain types of data may greatly assist decision makers. The readers of the summative evaluation may also be removed from the program, and not be very familiar with program activities.

Since the formative evaluation is an ongoing process, several reports may be needed to adequately convey the results. Progress reports, verbal or written, can be useful to decision makers by suggesting program changes that would improve effectiveness. Verbal reports accompanied by charts, diagrams, and other visual displays may help program personnel obtain a clear picture of where the program is going in comparison to the original objectives. If the evaluator outlines a timetable for progress reports, the program staff becomes aware of points during the project duration when changes and improvements are likely to occur.

The thorough evaluation report may contain the following items,⁵ although parts of this list may be deleted if unnecessary for a particular project:

Title Page

- * Title of program
- * Location of program
- * Name of evaluator
- * Name of people receiving the evaluation
- * Period covered by the evaluation
- * Date report submitted

Summary

- * What program was evaluated
- * Why evaluation conducted
- * Who requested the evaluation
- * Audience addressed by the report
- * Brief statement of the findings and recommendations
- * Decisions that were to be made on the basis of the evaluation

Background Information

- * Origin of the program---where is the community, tribe(s) or bands, groups in the community; how the program started; was a needs assessment conducted?
- * Goals of the program, and priorities
- * Participants in the program

- * Characteristics of the program---resources, activities, materials, administrative structure, staff, rationale underlying the planned structure of the program, difference between the planned structure and how the program is actually structured

Description of the Evaluation Study

- * Purposes of the evaluation---who requested the evaluation, evaluation formative or summative, audience, information needed for a decisions; who are the decision makers; does the evaluation pertain just to the program evaluated, or were there research questions involved and limitations placed on the evaluation?
- * Evaluation design---details of the design(s); why design chosen; limitations of the design
- * Instruments---program results measured; data collected for each measurement; examples of instruments used; how instruments were tested; how instruments were developed
- * Data collection---schedule for data collection; who collected data; training provided for those collecting data; sampling techniques
- * Implementation---implementation described; what aspects of the program were observed; any difficulties with implementation

Results

- * Presenting the data---summaries, tables, graphs
- * Participating groups---how many took pretests and posttests; did groups change?
- * Comparisons---differences, are the statistically significant?
- * Implementation---did the program end up as planned?; reasons for modifications; effects of changes on participants
- * Informal results---comments or summaries that describe or support the data findings
- * Discussion of results---certainty that the program caused the results; does the staff feel that the program could be improved with changes?

Costs and Benefits (optional)

- * Definition of costs and benefits
- * Method used for calculating costs and benefits
- * Dollar costs---extra funds required to implement program
- * Non-dollar costs---staff overtime, volunteer, stress to participants
- * Dollar benefits---income received for the program
- * Non-dollar benefits---program value to participants, other benefits to community

Conclusions

- * Major conclusions about the program as a whole
- * Conclusions about different components, if applicable
- * Weakness of the evaluation
- * Recommendations concerning the program
- * Recommendations concerning future evaluations

APPLYING RESULTS

In the past, one of the major down falls of program evaluation has been the failure to apply the results. This omission may be due in part to the fact that evaluation is rarely a smooth process. Often times the evaluator is seen as an outsider and a potential threat to the future of the program rather than as someone to give guidance. When programs have had only negative experiences with evaluators, such as at the point of defunding decisions, it may be difficult for a community to see the positive benefits of evaluation. Conducting a sensitive and useful evaluation involves a great deal of compromise between ideal methods and flexibility to meet program needs. The unfortunate result of the insensitive evaluation is the report that sits on a shelf and is not applied.

Explaining in advance the steps that can be taken to work cooperatively can help establish a relationship of confidence and mutual trust with the program staff. For example, it is not unusual for evaluators to respect the needs of the program by keeping confidential many of the details upon which recommendations were based. Some staff members may not realize that the purpose of the evaluation is to suggest improvements, rather than to find fault. This can be explained in advance. Evaluation is usually a new process for those involved and it creates considerable demands on staff time. Estimates of the time required should be explained in advance. The items that will be reported on, emphasizing both the positive and the critical, can be agreed upon at the start of the evaluation. Reports are also often reviewed by staff before being submitted to the funding agency or administrator requesting the evaluation. Building a trusting relationship

for the evaluation is an important part of creating the evaluation setting. And the trust relationship must continue after the evaluation is completed. Suggestions are given in the last chapter for developing cooperative agreements.

One of the most important tasks in interpreting the evaluation results is determining if the right questions were asked. If the evaluation results is negative, the reasons could vary from a poor "cultural fit" of the methods, the methods themselves, the structure of the program, or the program's ability to carry out the methods. It is often difficult to determine the cause-and-effect of certain program results. For example, factors outside of the program may have influenced the evaluation outcome. With the formative evaluation, the timing of questions asked is critical, as in allowing the program to have an adequate startup before looking for results.

Another important aspect of interpreting the evaluation lies in connecting the particular program to the rest of the culture. Ways in which the program activities are related to cultural values and activities can have an impact on the effectiveness of the program. In many instances, a great deal of staff effort goes into the design of a culturally appropriate program structure, a detail that should be considered in measuring the overall progress of the program. The cultural appropriateness of the program can also be related to the long-term effectiveness of the program. One weakness of many evaluation studies is that they measure only the short-term effects. Future assessments should be built into the evaluation plan whenever this is possible, and, if this is not possible, the likely weaknesses of the short-term study should be identified.

Several challenges face the evaluators of Native American programs in developing **culturally appropriate** measures of progress:

First, new data bases or sources for "baseline data" need to be developed for comparison purposes. The importance of baseline data lies in the question "where were we at when we started?" For example, a small amount of change in substance abuse may be a milestone in a population where patterns have been set and learned for generations, as compared to another population with recent histories of substance use problems. The "social indicators" that have been developed for health, education, and employment for the general population may not be applicable to specific population segments, such as Native Americans. Creating a research network to collect baseline data, to avoid duplication of studies, to share new methodologies, and to locate comparison groups is a first step in meeting this challenge.

Second, new comparison models must be developed that present alternatives to the experimental model. As mentioned previously, there is at

present a difficulty with appropriate comparison groups for the experimental approach. With the diversity of Native American communities, this may well continue to be an evaluation problem even after more evaluations have been conducted. Due to the lack of comparison groups, the "before vs. after" model or other approaches may be more appropriate to the future of the evaluation with a good cultural fit.

And third, the development of models for improving programs rather than those aimed at "yes/no" decisions may be more beneficial to the future of effective programs. Improvement as a goal for evaluation contributes to the idea of community development where the results have specific suggestions for improvement and change. Relating the evaluation results to decisions that need to be made about the program is one of the team efforts between the evaluator and the program staff.

In a community where resources are limited and the needs are many, evaluation can assist the community effort for developing the most effective, low-cost community assistance. Evaluation is not only a tool for improvement but also a means of documenting effectiveness according to culturally appropriate standards.

NOTES

1. This distinction made by Michael Scriven "The Methodology of Evaluation," in *Perspectives of Curriculum Evaluation*, edited by R. W. Tyler, R. M. Gange and M. Scriven (Chicago: Rand McNally, 1967) is commonly followed by evaluators.
2. The evaluation process diagram is adapted from a model presented by John Van Maanen, "The Process of Program Evaluation," *The Grantsmanship Center News*, January/February, 1979.
3. Same reference, p. 44.
4. Adapted from evaluation planning checklist in "Writing the Evaluation Section of a Proposal" by Dave Churchman, *Grants Magazine*, Vol. 4, No. 3, Sept. 1981, pp. 175-185.
5. Content for this outline is adapted from Lynn Lyons Morris and Carol Taylor Fitz-Gibbon, *How to Present an Evaluation Report* (Beverly Hills: Sage Publications, Inc., 1978).

ADDITIONAL SOURCES

Borg, Walter and Meredith Gall, "Evaluation Research," in *Educational Research* (New York: Longman, Inc., 1979).

This description of evaluation research is intended primarily for educators. Topics include: the role of evaluation in decision making, formative and summative evaluation, goal-free evaluation, the similarities and differences between educational evaluation. Explanations are very clear at an intermediate level.

Datta, Lois-Ellin and Robert Perloff, editors, *Improving Evaluations* (Beverly Hills: Sage Publications, Inc., 1979).

Topics covered in this collection of articles include: obstacles to evaluation, improving evaluation designs (control groups, nonexperimental evaluation research, approaches to assessing change, applying time series strategies), improving measurement (measuring impact, behavioral scales of job performance, costs and benefits of new health technologies), the role of evaluation (problem solving, evaluation and change, emerging issues for evaluators and evaluation users). The text is written on an intermediate to advanced level and concerns practical, applied questions.

Fitz-Gibbon, Carol Taylor and Lynn Morris, *How to Design a Program Evaluation* (Beverly Hills: Sage Publications, 1978).

This educational research text is clearly written and covers both formative and summative designs, the elements of an evaluation design, control group design, time series designs, before-and-after designs and other techniques such as analysis of variance and selecting a random sample (intermediate to advanced level). A companion volume, *Evaluator's Handbook* (Morris and Fitz-Gibbon) provides a workbook format for constructing an evaluation design. It is a practical guide that includes valuable details on how to work with the people involved in an evaluation.

Hatry, Harry P., Richard E. Winnie, and Donald M. Fisk, *Practical Program Evaluation for State and Local Government Officials* (Washington, D.C.: The Urban Institute, 1973).

A text for evaluating government funded programs, valuable for its clear presentations of specific evaluation examples. Practical details are given for each example, such as time and funding considerations, and steps to accomplishing the evaluation. Advantages and disadvantages of outside evaluators are discussed. This is a good guide for gaining an understanding to the constraints that government evaluators work with, and good preparation material for that outside evaluation.

Morris, Lynn Lyons and Carol Taylor Fitz-Gibbon, *How to Present and Evaluation Report* (Beverly Hills: Sage Publications, Inc., 1978).

An excellent guide to preparing both verbal and written reports. Practical tips are given on communicating evaluation information. Many examples are given for different techniques of presenting data, such as graphs, tables, and preparing the audience to read the data.

Popham, W. James, *Educational Evaluation* (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1975).

The focus of this text is on measurement techniques for assessing educational progress. Alternative measurement techniques are presented as well as classical techniques, measurement of effect, evaluation designs, sampling strategies, analyzing evaluative data, reporting evaluation results, cost analysis considerations, and teacher evaluation. Written on an intermediate to advanced level.

Rutman, Leonard, editor, *Evaluation Research Methods: A Basic Guide* (Beverly Hills: Sage Publications, 1977).

This intermediate to advanced level text on evaluation presents such topics as planning of the evaluation study, evaluability assessment (determining which program components can be appropriately evaluated for their effectiveness), measurement in criminal justice, randomized and quasi-experimental designs, data analysis, information systems for evaluation and feedback in mental health organizations, and benefit cost evaluation. The focus of the content is on identifying the practical problems faced by the evaluator in carrying out an evaluation design, and is useful in planning an evaluation study.

Shortell, Stephen M. and William C. Richardson, *Health Program Evaluation* (St. Louis: The C. V. Mosby Company, 1978).

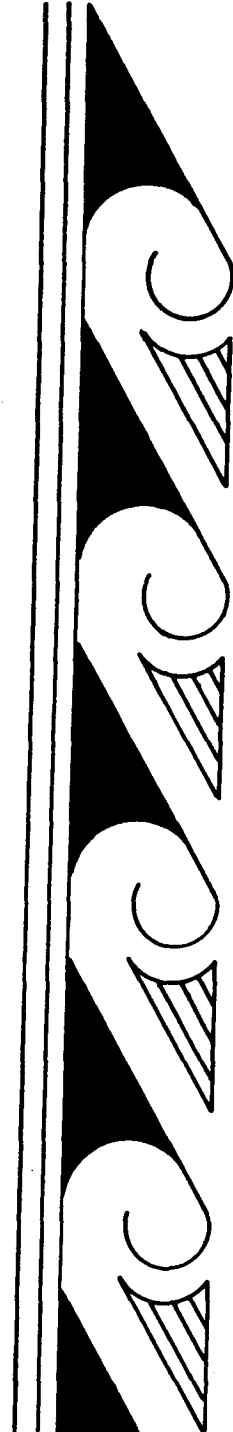
With an emphasis on health programs, the text outlines the evaluation process, evaluation designs, measurement, data collection, problems of implementation, and public policy issues. Presentation is at an intermediate to advanced level.

Van Mannen, John, "The Process of Program Evaluation," in *The Grantsmanship Center News* (January/February, 1979), pp. 29-74.

This is an excellent introductory article on formative evaluation. Topics covered include the evaluation model (processes), evaluation questions, measurement issues, comparison, confidentiality, and practical issues about working within the program setting. The material is written from an experienced viewpoint and covers many practical questions, including application of the results.

Wholey, Joseph S., *Evaluation: Promise and Performance* (Washington, D. C.: The Urban Institute, 1979).

This text, intended as a handbook for evaluating federal programs, is used by many agencies in monitoring program progress. Seeing the structure of measures can help a program to structure record keeping and data collecting in a manner that will maximize the ability to work with an evaluator. Topics covered include: collecting information, modeling, analysis, resource requirements, potential problems, feedback to the program, performance monitoring, and managing a useful evaluation program. Material is presented on a beginning to intermediate level.



Chapter 5 CULTURAL ARTS

Community-based research techniques are useful for preserving and continuing the traditional arts and culture. Attention to this type of development can balance change with tradition. In this chapter information is presented on organizing people, identifying resources, and defining a cultural or arts project. Steps are given for accomplishing a project. You will find specific techniques described for tape recording and photography, as well as an overview of video and film documentation. Other topics are covered that often prove difficult in culture and fine arts projects, such as protecting information, avoiding bias, and developing continued support for the arts.