PN-ACK-302

Improving Educational Quality (IEQ) Project

Multiple Method Evaluations

IEQI undertaken by:
Institute for International Research
in collaboration with
Juárez and Associates, Inc.
The University of Pittsburgh

Report prepared by:
Regino Chávez

Contract No. DPE-5836-Q-00-1043-00

1996

Contents

C

Introduction

Over the decades, program evaluation has evolved into an increasingly complex process. A variety of issues confront researchers in designing an evaluation to see how a program is doing or whether it has met its goals. Demands from policy makers, program administrators, practitioners, users, and funding sources have led to more sophisticated designs in evaluation. Additionally, an increasing empowerment of communities has led to more constraints on who and how evaluation studies are conducted. Demands for rapid information on how a program is doing and whether it can be used in different settings has given rise to evaluations of programs implemented in different locales and other contexts. However, difficult issues emerge in assessing a program situated in different parts of an organization, city, or country. This module provides an introduction to some basic issues regarding the conduct of evaluations using multiple methods in a variety of locations.

The module can be used by any person unfamiliar with evaluations where multiple strategies are used to collect data or by program managers and other people who are affected by the evaluation process. The document is not intended to be a comprehensive and detailed examination of every procedure and issue relevant to multimethod and multisite evaluations. Rather, it is intended as a synopsis of essential concepts and strategies in the conduct of multimethod evaluation research. The module provides the reader with a good working introduction to this type of work and to the challenges in effectively carrying it out. Suffice to say that this should not be the sole source for material on this complex subject. Other IEQ modules are available that address in more detail other aspects of the research and evaluation process for educational settings. The document is designed to permit readers to acquire the following:

- Get a working understanding of multimethod and multisite evaluations;
- Get an overview of the complexities involved in designing and implementing this type of evaluation; and
- Show readers how to use these experiences to avoid major pitfalls in the conduct of program evaluations, in general.

Several important assumptions are made with regard to the training of staff for multimethod, multisite evaluations. First, training in evaluation skills-acquisition is a slow process. Also, skills training is effective if it is a cumulative process that is best conducted in small, doable projects. Finally, the training is best if it is broken down into basic components (e.g., defining evaluation objectives, identifying indicators, instrument development, etc.). Training works best when it is empowering and enabling in demystifying the evaluation process. Ideally, this resource document is best used along with a skills-based training course that includes a mini-project where trainees work together on a single project that incorporates defining objectives and indicators, instrument development, data collection and analysis and report writing. The mini-projects will allow readers to apply new techniques at each different phase of the effort as the group is guided through an entire evaluation project.

The module introduces the technical issues involved in implementing evaluations utilizing multimethods. The module addresses when an evaluation should be done, designing an evaluation, and what activities will facilitate carrying out an evaluation where multiple methods are used. Also addressed in the report are issues related to staffing and managing a research team, entree into the research setting, data analysis and role management.

Overview of Evaluation

The Framework

Program evaluation involves examining an intervention to determine how it has met its goals, whether it has met its goals or whether it is worth continuing. Program evaluators then focus on any of the following in carrying out that task:

- Measuring the results of a program relative to its goals;
- Assessing the degree to which a program fulfills identified goals; and
- Providing information for continuing, modifying or terminating program.

In conducting an evaluation in an educational setting, practitioners rely on formal and systematic procedures for looking at events in the schools. In the past, practitioners used methods from a number of disciplines and incorporated new ones as the practice of evaluation evolved. Now, practitioners have a number of methods available to use in looking at educational practices within the classroom context.

A key issue here is the perspective that educational practices in schools are social events constructed through interaction among students, teachers, parents and others involved in the educational process. In examining these social events, evaluators need to see what the action and interaction mean for the actors in the study and examine that dynamic to appreciate the complexities of the interaction and to identify the meaning attached by the different actors to the social event. One can only do so by looking at it from multiple perspectives- students, teachers, parents, slow readers, fast readers, etc. Each perspective has its own rules, systems of relevance and meaning. The goal in research is to examine what Geertz (1973:5) called "...the web of significance that [the actors] themselves have spun". No one evaluation instrument can be sensitive to all perspectives nor appropriate for use with all respondents.

In conducting evaluation research, evaluators need to be sensitive to the meaning that actors within the school setting attach to different events and activities. Thus, the first step in any evaluation is to focus the effort and set the objectives. Focusing an evaluation will assist an organization to decide whether an evaluation should be undertaken. Focusing also assures that a consensus on information needs exists on the part of all stakeholders in the evaluation. The focusing process helps in framing the evaluation questions to be asked. In focusing the evaluation, practitioners assure that the right information needed by decision-makers is obtained from the most appropriate

sources using the most effective data collection methods. Finally, establishing the objectives of the evaluation assures that the resources allocated for the assessment effort are used effectively and efficiently. For example, a cost-benefit analysis may focus solely on the financial end of whether a teacher training module will be bought or not; however, an impact assessment examining concept learning may look as much at student behaviors that led to change as to teachers' behaviors that contributed to that learning.

To focus an evaluation, evaluators need to know the system they are to examine, pose evaluation questions, assure that there is conceptual clarity in the aspects of the program to be assessed and that the appropriate goals of the program are translated into behaviors that can be measured.

To Evaluate or Not to Evaluate

In some instances, an organization may have no choice but to carry out an impact evaluation as funding for the program may be contingent on the effort. However, no impact evaluation should be carried out until a program has been given a chance to develop fully and staff implementing the program are knowledgeable and familiar with how to implement it as specified by developers. Other factors to consider in deciding whether to implement an evaluation are as follows:

- Will the resources available for the evaluation effort permit an adequate research design?
- Will the managers actually use the findings?

Formative evaluations, however, should be a component of every educational intervention. Formative evaluations permit the developers and implementing staff to identify processes and procedures that can be improved. Readers may wish to consult Shadish, et. al. (1991) for other issues important in deciding whether to carry out an evaluation.

Goal Definition

The need to define the evaluation goals for the any program evaluation are critical. The goals, to a large extent, will influence and, in some cases, determine the following:

- the number of sites selected;
- the time to conduct the research and data collection resources;
- who will be the focus of the evaluation, incorporating what perspectives, for what reasons;
- the data collection strategies;
- the data analysis techniques; and
- how the findings will be disseminated.

The Need to Understand the System under Study

Usually educational program evaluations are conducted in a number of schools implementing a program. The complexity of the effort is increased as one recognizes that the phenomenon or social events within each school include a number of actors. Because of the multiple sites in which the evaluations take place, there is a need to understand the system and how this will influence the design of the evaluation including program implementation or data collection strategies. For example, where there is a unity to the system being studied, this may permit easier implementation of standardized instruments (Hendrick, 1991). In the case of Guatemala, for example, standard qualitative and quantitative instruments were developed to examine the implementation of a curriculum model in two different contexts and with two different subgroups of the Guatemalan population. The unity of the system, though, was afforded by the educational system within which both of the sites operated. However, where a system differs, there may be issues of accepting standardized instrumentation. In South Africa, the IEQ project undertook an evaluation of pre-school (i.e., educare) teacher training programs. Though all the programs were in field of educare, there was no unity to interventions thus making the development and use of a standardized instrument to examine teacher training difficult. The politics of funding as well as the differences in emphasis of the training organizations made difficult the acceptance of one standard instrument to examine outcomes.

Tip

Prior to design, ask yourself:

- Is the evaluation needed?
- What information is needed?
- What factors will facilitate conducting the evaluation?
- What barriers might I encounter?
- Is the program sufficiently mature to warrant an evaluation?

An evaluation also takes place within a political context whether at the micro- or macro-level (Finsterbusch and Motz, 1980; Skogan and Lurigio, 1991; Turpin and Sinacore, 1991). For example, evaluation research that would require teachers to fill out observation forms, questionnaires, or administer standardized tests would generally not be welcomed by program personnel or teachers as it interferes, however slightly, with service provision (the goal of the staff). This is especially true when a highly charged political context exists, where an evaluation would be perceived as a major threat by implementers or where the program implementing organization is highly centralized and fragmented. In these circumstances, an evaluation creates an uncomfortable to hostile climate for evaluators as it could result in any of the following:

- Competition for resources: A conflict over resources is created. For example, the time allocated for teacher trainers to go into the field and lead workshops may go instead to data collection in the form of testing,

observing or responding to interviews. Intervention program staff may want that time used for providing service and not for carrying out research tasks.

- Inconsistency in Program Implementation: This occurs during the course of study and impacts on the evaluation. For example, modifications may be made as staff discover new ways to do something. These "new" ways are not part of the intervention model yet will affect program outcomes.

- Conflict in Perspectives: Situations may arise in which viewpoints differ between administrators and evaluators on the use of findings, goals of the research, and merits of the program. Evaluator and practitioner viewpoints may also conflict as practitioners associate evaluators with their supervisors or those who have power over them.

Evaluators can address these concerns by gaining an understanding of the system within which the evaluation is to be implemented and assuring the following:

a) *Stakeholder Benefits*: Evaluation managers need to nurture ownership of the evaluation at the local site to ensure that the locals have a stake in producing a high-quality product.

b) *Political climate*: The system may be composed of multiple political environments. There may be key persons at each site who can influence data quality. Thus, evaluation managers and field team members need to identify the political climate at each site and respond accordingly.

Tip

Focusing an evaluation assures that the right information is obtained from the appropriate sources using the most effective methods.

To focus an evaluation,

- ask what type of evaluation is sought: monitoring? effectiveness? impact?
- define with the stakeholders the objectives and acceptable levels for success;
- translate them into behavioral indicators;
- know the system.

c) *Cultural differences*: Evaluators need to take account of the cultural differences within a local area. This is especially important during instrument development as it needs to be in the local language and in local dialect of region (Whyte, 1984). Additionally, there may be instances where there are cultural differences by region as well as within a local implementing organization. For example, IEQ researchers need to account for cultural differences among the different ethnic groups within the indigenous populations of South Africa as well as for the manner in which the distinct Departments of Education structure policies for schooling. Another example may be that the different pace of life in rural areas of Guatemala may affect the data collection effort for an evaluation that is on a tight schedule if this is not taken into account (vacations, etc.).

Influence of Program Implementation

Understanding the system also includes understanding the nature of program implementation. Implementing organizations and community contexts influence how a program is implemented and the outcomes of the intervention. These contextual effects may result not only in adaptations of the program model but also in substantially different programs being implemented in different locations (Mowbray and Herman, 1991). For example, local implementing organizations can differ in philosophy, comprehensiveness of services or community resources. IEQ researchers, for example, identified preschool teacher training programs that offered teachers technical skills for use in the classroom; however, others also sought community change and provided teachers with skills for organizing their community and incorporating them in the preschool setting. These differences result in evaluation managers having unique rather than similar communities. Strategies for mediating the effects under these circumstances include the following:

a) Collect process data so that any threats to validity are known and quantified. Develop implementation measures to measure degree of implementation and to examine contextual effects (Mowbray & Herman, 1991).

b) Develop a theoretical model of implementation and expected outcomes.

c) Develop a model as much of the treatment as of its variations. This can be done by a review of literature, interviews with developers and staff. The South African IEQ research team used concept mapping where staff described their educare teacher training "treatment", the important features of their models and how these impacted outcomes. Evaluation managers can extend the concept by describing acceptable and unacceptable variations of treatment keeping in mind that fidelity of treatment is multidimensional (number of components of the preschool teacher training model that are in place as well as extent of variation at each site).

d) Develop definitions of success and failure with the program implementers prior to data collection; measures may need adapting to local contexts or identifying new performance measures in instances where new programs are being tested (Skogan and Lurigio, 1991).

Designing Program Evaluations

Program evaluation is a complex undertaking that can be reduced to several basic steps. Each step consists of a number of components each deserving of special consideration beyond the scope of this training module. Readers are referred to Posavac and Carey (1989), Shadish, Cook and Leviton (1991) or the Sage Publication series on program evaluation for more information on the subject. The basic steps are provided here as a means to help organize the remainder of the content for this module.

Steps in Designing an Evaluation

1. Focus the Evaluation
2. Set the objectives
3. Develop the instruments to collect data.
4. Collect the data
5. Analyze the data
6. Report the findings

As previously noted, focusing an evaluation clarifies the type of evaluation desired to assure that expectations for all stakeholders are addressed in the design. It assists program implementers and developers or funders to clarify information needs and to assure that the evaluation will meet those needs. The subsequent subsections of the module address issues related to setting objectives, data collection and analysis and reporting of findings. Additionally, the module highlights issues related to training of field workers who are to collect the data.

Translating Objectives into Behavioral Indicators

Educational programs usually set program goals in highly philosophical or global terms that are lofty and idealistic. Evaluators find it difficult to take those goals as they are stated and use them to focus an evaluation. Many times, such goals are non-specific and nonmeasurable. For example, a program goal may be to improve the quality of education in a community. A program objective to accomplish such a goal may include providing students with basic competencies and enable students to develop positive self-esteem. At this stage, the program goal has been specified somewhat by the two program objectives. However, the latter are still lack tangible content that allows an evaluator to measure whether they have been accomplished. Readers can refer to Isaac and Michael (1995), Mager (1962), Kibler, Barker and Miles (1970) for more information on defining behavioral objectives.

A major component of any evaluation is translating the objectives of the program into behavioral indicators. That is, the evaluation team must assure that the program

objectives are specified as behaviors, tasks, processes or activities that are observable and measurable. The evaluation team will be looking for such behaviors in seeking to determine what effect the program may have had on the target group. Another way to see this constellation of behaviors is by viewing them as definitions of the concepts to be measured by the instruments. Indicators are the skills, attitudes, feelings, aptitudes, knowledge that the program wishes to develop and assess in the target group(s).

Tip

- Program goals must be turned into clear, discrete, and measurable behavioral objectives.
- Behavioral objectives have four characteristics: 1) description of the learning task; 2) specify who is to carry out the action; 3) use action verbs that specify observable activities or processes; and 4) specify an outcome. Objectives may also define an acceptable level of performance.
- Evaluators must turn program goals and objectives into operational definitions.
- The operational definitions become behavioral indicators of the activities, skills, attitudes, feelings, aptitudes, and knowledge that the program wishes to develop and evaluate in the target group(s).

For example, the adjacent graphic depicts objectives for a domain of knowledge comprehension at increasing levels of specificity:

Exhibit 1. Specifying Objectives

Knowledge Comprehension:
Student understands the material being communicated without reference to other materials

Knowledge Comprehension:
Fourth grade level student demonstrates an understanding of the material being communicated by *answering questions* without reference to other materials.

Knowledge Comprehension:
Given 5 questions, the fourth grade- level student demonstrates an understanding of the material being communicated *by responding correctly in written form to 4 out of 5 questions in less than ten minutes.*

For illustrative purposes, let us consider the case of a teacher training program with the objective of training teachers in Learner-centered Teaching. Evaluators first need to understand what are the elements of Learner-centered Teaching intervention according to the specific model. Once each element of the model is specified, the behaviors are assigned to their respective actors. For example, the model would expect teachers to engage in certain behaviors, tasks or activities; it would suggest different sets of behaviors for learners; finally, it would set specific arrangements for the physical setting within which the interaction is to occur. The following exhibit demonstrates the results

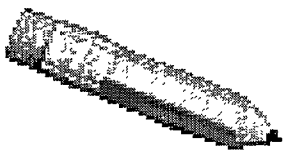
of a training exercise where an evaluation team translated the objectives to indicators. Note that the objectives and indicators are organized by subject of the action. The indicators are very specific behaviors that the research team will examine within the classroom context.

Exhibit 2. Intended Classroom Outcomes of the Learner-Centered Program

Unit of Focus	Educational objective by Domain	Behavioral Indicators: Intended Classroom Outcomes
Teacher	Teaching strategies	
	Encourages learner- teacher interaction	Teacher responds to child-initiated questions. Uses small groups
	Uses Language Home Sensitive	Teacher uses language that children understand in learning situations. Teacher uses classroom or home experience of children in learning situations. Teacher uses parents to deliver classroom lesson. Teacher uses local materials in conducting classroom lesson.
	Learner-centered learning	Teacher works with individuals, pairs or small groups. Teacher asks open-ended questions. Teacher uses probes (Uses key words or asks for explanations). Teacher allows students to assume different roles. Teacher allows students to choose own activity. Teacher permits child to select activity. Teacher permits children to select peer leaders.
	Role: Teacher as learner	Teacher describes knowledge gained from child.
	Classroom Management	Teacher specifies, orally or written, the objectives of a lesson. Teacher identifies skills to be developed, orally or written. Teacher carries out activities as planned. Teacher permits children to select peer leaders.

Unit of Focus	Educational objective by Domain	Behavioral Indicators: Intended Classroom Outcomes
Learner	Hands-on/Minds-on	Children move among different groups and learning centers. Children work in small groups without teachers. Children direct learning activities. Children use materials to carry out task.
	Communicative use of language	Children describe classroom or home experience, orally or written.
	Clear self-expression	Child verbally expresses opinion to adult or another child.
	Peer Interaction	Children sharing learning materials.
	Meaningful Material Interaction	Children working together on lesson. Children refer to materials during an activity. Children handle variety of materials in lesson. Children make their own materials.
Setting	Classroom Organization promotes interaction	Children work space arranged in flexible clusters.
	Presence of Stimulating materials	Materials have at least three different colors. Materials are relevant to context (rural or urban). Materials from home culture are available in the classroom. All children are working with materials, individually or in groups.
	Learner-friendly Physical Environment Available	Classroom has child-sized furniture. Children's current work displayed in classroom. Learning materials are safe for use.

The process of identifying indicators also clarifies the particular data collection strategies that can be employed by evaluators. For example, indicators could also include attendance, presence and type of books, and completion of homework assignments. Each of these could be recorded through different data collection strategies to obtain the needed information. Attendance could be obtained through a review of records; presence of books or types of books through a checklist; and completion of homework assignment through interviews.



Practice Exercise 1. Focusing an Evaluation

I. Focus the Evaluation.

- Is this evaluation needed? If so, what is to be accomplished?

  _____
  _____

- What questions will the evaluation answer?

  _____
  _____

- What type of evaluation is best given the information needs?

  _____
  _____

- Who are my stakeholders? Do all key stakeholders agree that this is the type of information they need?

  _____
  _____

II. Know the System

- What key individuals will influence how the evaluation is carried out?

  _____
  _____

- Are there conflicts in the goal of the evaluation?

  _____
  _____

- Is there consensus on what will be viewed as successful intervention?

  _____
  _____

III. Translating Objectives into Indicators

- What are the goals of the educational program?

  _____
  _____

- Who is to be affected?

  _____
  _____

- What behaviors are targeted by the intervention?

  _____

- Can you phrase the objectives in clear, specific and measurable terms? Use the verbs below or others to identify expected outcomes that are to take place for each target group. Specify the level and the target group (teacher, learner, classroom, parents, community, etc.)

Write	Produce	Recall	Recognize	Contrast	Acquire
Separate	Use	Provide	Identify	Increase	Describe
Discuss	Order	Complete	Classify	Categorize	Answer

On the next page, identify the goal of the evaluation you are to undertake. Specify the goals and objectives by area of the program. That is, break down the components of the program to identify the areas of behavior that the program seeks to influence, e.g., teaching strategies, peer interaction, classroom arrangement, use of material, etc. Specify the objectives using the verb statements you developed previously. Finally, construct a question for each indicator.

Practice Linking Goals, Objectives, Indicators and Questions

Evaluation Goal	Program Goal	Objective	Indicators	Target Group	Question (Number the Question)
		Teaching strategies Classroom Management		Teacher	
		Learner skills: Cognitive Socioemotional Physical		Learners	
		Organization Materials Physical Environment		Classroom or Site	

Multiple Methods in Evaluation

Educational evaluators are faced with the need for work that responds to specific questions raised by program developers, policy makers, program staff, clients or funders. Findings are usually needed to make timely decisions on improving or continuing a program. In designing an educational evaluation, researchers must be clear on the following:

- the intent of recording the information;
- what action to focus on;
- who to focus on; and
- how to collect the information on the lesson.

Tip

Evaluation methodology must be tied to the purpose of the evaluation, its objectives, the phenomenon to be examined and the resources available to carry out the assessment.

Any number of research strategies are available to the evaluator given the purpose of the evaluation and the evaluation questions to be answered: school records, classroom maps, census data, surveys, observations, checklists, interviews, and experiments. No one strategy is perfect. Use of several strategies permits looking at the classroom phenomenon from several perspectives. The best research strategy seeks to establish a dynamic, productive interplay among data collection methods (Pelto & Pelto, 1978). Among the advantages of using combinations of methods are the following:

- Using more than one way to collect data will provide a means of looking at a social event in a fuller perspective as data is collected on more than one aspect of some phenomenon. For example, a language lesson has a beginning a middle and an end. It involves teachers, students and materials. It can take place while students are working in pairs, small groups or doing individual seat work. In conducting evaluations on language learning, researchers must collect information that captures all the dynamics involved in such a lesson. This is best done by observing interactions, noting the materials and physical setting and talking to the different actors about

Tip

Multiple research strategies provide opportunities for developing innovative ways of studying complex social phenomenon.

Multiple Research Strategies allow interpretation not possible if only one source of data is used.

Multiple research strategies increase reliability and validity of the data.

what the researcher saw. No one data collection method incorporates such dynamism.

- Multiple research strategies will allow the interpretation of data not possible if only one data collection method is used. For example, Weisz, et. al. (1995) had teachers use a rating form to examine 'problem behavior' of Thai and American children as well as having a team of researchers observe the same classrooms. It was found that teachers rated children as exhibiting much higher instances of problem behavior than did the researchers. They also pointed to an inverse relationship with teachers noting more Thai children exhibiting problem behavior whereas observers found more incidence of problem behavior among American children. The observations permitted examining the data to note that culture of the teachers and gender of the children were factors affecting the ratings by teachers.

- Using more than one source of data also increases the reliability and validity of the information collected. In using more than one data collection strategy, evaluators can corroborate findings from one source with others. For example, data from teacher interviews on use of specific teaching strategies such as small groups can be corroborated through observations in the classrooms. Use of multiple research strategies also permits placing specific findings within the broader sociocultural context and the fleshing out of abstract numbers as would be obtained through surveys.

- Multiple research strategies also permit examining phenomenon that does not easily lend itself to observation or discussion. Evaluators cannot collect quantitative information nor observe all aspects of some social phenomenon. Some topics may be difficult to discuss with strangers, e.g, teacher attitudes toward children of a different ethnic group. Some respondents may not have adequate communication skills to be able to respond to surveys or interview questions, e.g., three- or four-year olds, developmentally disabled persons, etc. Some phenomenon cannot be observed, e.g., beliefs, attitudes, feelings. Thus, multiple research strategies allow an evaluators to collect data on these phenomenon using the most appropriate instrument for the context, the respondent and the evaluation intent.

- Finally, use of multiple methods forces the evaluator to resolve conflicts between qualitative and quantitative data sets rather than discarding findings. The evaluator is forced to pose new hypotheses or propose new studies.

Tip

The combination of methods used is based on the evaluation objectives and the research context - resources and respondents.

Regardless of the strategy, all methods are used to collect information that describes and explains social phenomenon. In considering which methods to use, an evaluator must focus on selecting the best procedures for generating the required information given the content and respondents. As no one data collection strategy captures every dynamic, a combination of methods will strengthen any evaluation or research study (Reichardt and Cook, 1979). However, evaluators must take care in assuring that the various methods define and use the operational definitions or behavioral indicators in the same manner regardless of the data collection instrument used. Researchers have pointed to the problems inherent when concepts are used differently in a survey and observation form (Isaac and Michaels, 1995; Temple, 1994). Of special concern is that poorly-defined concepts will lead to evaluators examining different behaviors under the guise of the same concept and to contradictory findings. A basic question to ask is:

> *What constellation of methods will produce the kind of information that will address the [*evaluation*] questions and concerns?* (Pitman & Maxwell, 1992:761)

Once the sources of the information have been identified, a number of factors will influence the final selection of the strategies for collecting the data; a number of sources are readily available to assist the evaluator in selecting methods (Shadish, et.al. 1991; Isaac and Michael, 1995; Pitman and Maxwell, 1992; Worthen and Sander, 1987). Two critical factors to consider are the resources and the time available to the evaluator to conduct the study and the purpose of the evaluation. Others are that research methods selected must fit both the research and the research setting. Other appropriate questions include the following:

- Will the information to be collected provide a comprehensive picture of what is evaluated?
- Are the procedures legal? Ethical?
- Are the costs worthwhile given the amount and type of data to be provided?
- Will data collection disrupt the project?
- Can the data be collected in a timely manner?
- Will the data provided be valid? Reliable? (Worthen and Sanders, 1987:237).

Data Collection Methods

Having identified the purpose and objectives of an evaluation, the next step in the process is to identify what information is needed and how it is to be collected. A number of data collection strategies are available to an evaluator in assessing an educational program. As mentioned previously, the combination of methods used is based on the evaluation objectives and the research context. Important to note for the instruments used in the data collection effort are the following:

- Are they reliable? Will the observations be consistent over time? Will others find similar results with this instrument?
- Are they valid? Do they measure what they claim to measure?

Miller and Crabtree (1992) discuss the various research aims and how these influence the methods used for a study. Among the categories they use are survey, historical, field-based methods, and experimental. Each can be used to identify, describe, explain, test and predict/control social phenomenon. According to Bieger and Gerlach (1996), the methods can also be classified by the nature of the questions posed, as their table below shows.

Table 1. Classification of Research Approaches by Type of Data and Question*

Approach	Type of Research Question		
	Descriptive	Relational	Causal-comparative
Quantitative	Uses numerical data (e.g., mean, median) to describe a variable	Uses numerical data (e.g., correlation coefficient) to show a relationship among variables	Uses numerical data (e.g., critical ratio) to establish a cause and effect connection among variables
Qualitative	Uses verbal reports to describe a trait, characteristic, or phenomenon	Uses verbal reports to show how traits are related to each other	Uses verbal reports to establish a cause and effect connection among phenomena

*From Bieger, G.R. and G.J. Gerlach. (1996) Educational Research: A Practical Approach. Boston: Delmar Publishers.

Use of a combination of methods is gaining more and more popularity among researchers as it permits flexibility and increases reliability of the results. For example, an evaluator can use standardized tests to measure a child's performance on specific measures. The same study could include observations of the child in the classroom to identify what is actually being learned through the program, how the program is being implemented or what factors influence what is learned. By augmenting the standardized tests with classroom observations, a complete picture is provided of classroom dynamics and the effects of an educational intervention. Among the more common methods used in educational evaluations are the following:

Exhibit 3. Types of Methods in Evaluation

Method	Purpose	Description
Standardized Tests	Measure achievement or performance through scores on a set of indicators.	Tests can allow evaluators to compare achievement of the study group to some norm of a defined population (norm-referenced); or to compare the scores of the study group to the total body of knowledge that the test is designed to cover (criterion-referenced); or on performance on some exercise or activity (performance assessment).
Surveys	Measure a target group(s)' behavior through questionnaires or interviews	Use instruments with fixed-choice responses and specially sequenced questions to obtain data from respondents.
Field-based methods	Record observations on predetermined and/or emergent categories important to the evaluation objectives.	May include focus groups, classroom maps, observations or in-depth interviews with predetermined topics to explore but no special sequencing or specified responses.
Content Analysis	Used to record information obtained from secondary sources.	May include review of school records such as attendance reports, newspaper articles, grade point averages, enrollment records in assessing aspects of behavior pertinent to the objectives
Experimental	Can be applied in situations that permit tight control of the intervention settings.	Uses random assignment of respondents to treatment groups and interventions.

To continue with the illustrative example posed previously on learner-centered teaching, a team of evaluators assessing the teacher training program might decide to use teacher observation forms, teacher interviews and classroom environment checklists as the principal means of collecting information. Appendix A provides examples of instruments that were developed for a training program on evaluation methods in the IEQ-South Africa project. The exhibit below shows some items from the different instruments as these relate to the behavioral objectives.

Exhibit 4. Links between Objectives and Instruments

Behavioral Objective	Interview	Checklist	Observation form
Provide meaningful material interaction: Children handle a variety of materials Children make their own materials	13. What kinds of materials do you use in conducting your lesson? 14. How did you decide on the materials you used in your lesson?		
Provide learner friendly physical environment		Child current work is displayed Safe learning materials	
Promote hands-on, Minds on Learning		Learning centers/corners are present Flexible work space for children	Recording of child-initiated interaction during any of the four contexts Recording of child working alone during individual contexts

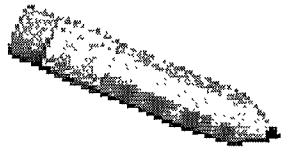
Standardization of Data Collection

Evaluations that use multiple methods require that at least some instrumentation and procedures be standardized to meet the goal of an evaluation of a program prototype. This need is made even more severe in multisite evaluations as long-term efforts that may have high staff turnover either of researchers or in implementing organizations (Turpin & Sinacore, 1991). Standardization of measures also contributes to mediating the effect of the potential flux in the field work context (Reiss & Boruch, 1991). However, MSEs are also burdened by the need to remain sensitive to local adaptations of program treatments or interventions. Where local adaptation of the program is needed, additional instrumentation is needed to reflect the unique aspects of those sites. Evaluation managers must remain sensitive to the differences in how the instruments are implemented by the members of the data collection team as this will differ even after training.

Standardization of Data Organization

Evaluation program managers need to establish procedures to standardize data organization and storage, data entry, and editing. Carrying out these procedures and processes in a central site minimizes any problems that might arise. For example, the data for the IEQ evaluation of the *Escuela Nueva* curriculum model was written up in the field by the team of researchers. They had been trained in naturalistic observation strategies, data coding and storage. Field notes from the team were then sent from the different sites to the IEQ central office in Guatemala City. Here, the coding was verified by the supervisor and data was catalogued and stored in central office files.

When data organization and storage procedures is carried out in at different sites, evaluation managers need to train individuals in the organization, filing, storing, computerizing, and editing of data (Turpin & Sinacore, 1991). Additionally, managers need to develop procedures as well as ensure the use of compatible equipment. Standardization of procedures will also help when there is staff turnover, in verifying data, and retrieval. When the procedures are standardized, anyone can be trained to retrieve the required data.



Practice Exercise 2: Selecting the Research Methods

Using the matrix below, fill the question number for the most appropriate research strategy for obtaining the information required to meet the evaluation goals. Use the questions developed in Practice Exercise 1. Add more rows or columns as appropriate. Use categories relevant to your program.

Target Group	Data Collection Strategy				
	Interview	Observation form	Survey	Records	
Teacher					
Learner					
Classroom					
Community					
Principal					
Board Member					

Data Analysis

Data analysis is the process of bringing order to the information collected in the field. The process consists of organizing the information into patterns and categories. The process is fluid to the extent that analysis begins with the posing of the evaluation question(s) and flows back and forth during the remainder of the project. The researcher in the field plays a critical part in the analysis of qualitative and quantitative data. As the "instrument" for qualitative data collection, the researcher not only gathers data but is also the one immersed in the context from which it flows. Evaluators need to plan beforehand, how each question will be analyzed usually specifying the technique to be employed for the analysis and the means for conducting the analysis; that is, will the data be described or counted? Will one use a calculator or a computerized analysis program?

Data interpretation involves giving meaning to the analysis. The interpretation process consists of explaining the descriptive patterns that emerge and identifying the significance of the associations or linkages among the various categories or variables. A number of texts provide information on data analysis techniques and are useful to examine. Among the more user-friendly texts are those by Bryman and Burgess, 1994; Crabtree and Miller, 1992; Goetz and LeCompte, 1984; Miles and Huberman, 1994; Patton, 1987; Spradley, 1980; Strauss, 1987; and Strauss and Corbin, 1990.

Data Analysis Strategies

The choice of analytical procedures will be determined by the type of question asked. Some studies may have no need for elaborate statistical analysis of the data. It is always preferable to use statistical procedures that are simple, if these are required. Usually, program sponsors or users have little understanding of statistical procedures. Planning ahead assures that the appropriate procedures can be applied to the information collected through a particular question or research strategy. Among the procedures are the following:

Tip

Develop a data analysis plan by deciding ahead of time how information from each question will be analyzed. Think through the best way of integrating quantitative and qualitative data sets.

Descriptions: The task of describing incorporates descriptions of beliefs and values of the participants, descriptions of their behaviors, and descriptions of the physical setting. The setting includes as much the community, as the school and the classroom. The focus of description is to ascertain what happened in the classrooms.

Counting: Another task involved in data analysis is simply counting the incidence of a behavior, value, belief or other phenomena of importance to the objectives of the evaluation. One can simple count the frequency of occurrence of a phenomena. If one is to compare the incidence across categories, it is importance to convert the instances of observed or recorded behavior into Relative Frequencies or percentages. For example, one can count the types of teaching strategies used, the instances of child-initiated interactions, etc.

Comparison/Contrast: Most times, it is not sufficient to simply count the number of times a phenomena occurs. It is important to examine whether there are specific conditions under which it occurs. Two strategies for elucidating conditions are comparing and contrasting the phenomena observed. In comparing behavior, the evaluator is interested in how are things similar. For example, a good comparison type question is "What types of strategies are used by all teachers?" In examining contrasts, the evaluator is interested in how are things different. For example, s/he might ask, "Do some types of teachers rely on some type of strategy?"

Identify patterns: It is in carrying out tasks related to comparison and contrasts that the analyst begins to identify patterns or structures in the data. For example, the analyst may begin to look for patterns of behavior organized by relationship within the classroom, e.g., Are some teachers doing some things more than others? Do boys interact differently with the teacher than girls? Does the teacher with more experience use different strategies than one with less experience? What categories emerge? Is there some order to how these appear? Additionally, one may begin to find exceptional cases; these are the cases that do not fit into the pattern found. It may be the one teacher that uses many strategies; the one classroom where interaction is mainly initiated by children; the one teacher who relies mainly on small groups in classroom organization; etc.

Data display, reduction and linking processes all permit the researcher to order and manage the information collected for the evaluation. The display and other processes of data analysis are then used to produce a narrative description of the evaluation findings.

Training in data analysis is a difficult and arduous process. Most information on data analysis seem to imply that patterns will magically emerge from the data. Thus, it is important that any training program emphasize practical experience in data analysis using data collected by trainees and analyzed by the group as a whole.

As trainees examine the data collected, many will feel intimated and hopelessly lost in attempting to find patterns and linkages among the different data sets collected through the various instruments. This fear of data analysis can be overcome by initiating data analysis procedures. Data are displayed using the various forms developed for this. The group then begins the analysis process by identifying global patterns in the data and moving on toward discovering links between the different variables under study.

The nature of multiple methods and multisite evaluations sets out a unique constellation of analysis concerns for trainers and program managers. Among the issues relevant to data analysis are developing an appropriate analysis plan, the pooling of data, testing for site interaction, and integrating quantitative and qualitative data sets.

Integrating Qualitative and Quantitative Data Sets

Traditionally, qualitative and quantitative research methods have been in antagonistic positions relative to their practitioners. More recently, increasingly complex interventions implemented in varying settings have placed stringent demands for more refined evaluation designs. One result has been the utilization of both qualitative and quantitative methods in educational evaluations utilizing multiple methods. However, the integration of data sets has not always been easy. In fact, the integration of data sets has not taken place in many evaluations (Greene, Caracelli and Graham, 1989). Researchers have noted that integrating qualitative and quantitative data sets may be appropriate only under specific circumstances (Greene, Caracelli & Graham, 1989), only for specific types of evaluations (Caracelli & Greene, 1993) or may be dependent on the subject matter under investigation and on certain assumptions (House, 1994). Whatever the circumstance or outlook, integration of data sets can take place at the following levels:

Tip

Integrate data sets through operational definitions-use the same indicators.

Use qualitative work to define or refine quantitative instrument development.

Use qualitative work to describe quantitative findings-describe a difficult concept.

Use qualitative work to interpret quantitative findings.

Use qualitative work to describe a phenomenon difficult to quantify.

- Transformation of one data set into another, such as through coding;
- Creation of new variables or data sets through the integration of data sets;
- Follow-up of extreme cases via one method though found through another;
- Typology formation from one method tested through another strategy;
- Findings: through explanations or extensions of quantitative findings, program descriptions or identifying circumstances under which an intervention was or was not effective;
- Content: through providing the story of the intervention;

Appendix B provides examples of data display, data analysis and data linking using information from data collected for the training workshop focused on Learner-Centered Teaching. The data is used for illustrative purposes only.

Pooling Data

In the case of evaluations using multiple sites, program managers need to determine whether data will be analyzed by site or pooled to test for overall effect (Turpin & Sinacore, 1991; Isaac & Michael, 1995). If one aggregates and obtains the average of scores on outcomes, they may differ due to a number of factors such as differences in demographics or the needs of participants. Differences due to program implementation are even more of a concern given the systematic nature of their influence on participants. If meaningful differences in implementing the intervention occur, one should not pool data, as findings are not interpretable.

Turpin & Sinacore (1991) cite two principal strategies for pooling data from multiple sites: 1) Pooling by Averaging can be conducted by averaging within-site differences where the effect of a program is assessed at each site, then its effects are averaged across the sites. 2) Pooling by Lumping takes place by adding together all the data from participants in a similar treatment group, irrespective of site; managers then examine the outcomes as though these are originating from one location. Pooling by lumping is the more common method in multisite evaluations though it assumes that, theoretically, participants at all site location receive the same treatment.

Test for Program-by-Site Interaction

MSE managers need to look for program-by-site interaction to see if overall analysis is going to be interpretable. Turpin and Sinacore (1991) suggest the use of an Analysis of Variance framework with site and treatment group as two factors. For example, one could examine teacher training in three schools where one group is offered the Escuela Nueva curriculum while another is offered the regular curriculum. Thus, one has two treatment groups at each school and three sites. Use of two-by-three ANOVA may reveal that participants responded similarly at all sites, thus permitting combining of outcomes for an overall analysis. If one site responds differently, a manager's options are to combine those sites where participants responded similarly and to look to why other site(s) are different.

However, the pooling of data does mask site specific impact (Cottingham, 1991). Also, the pooling of data may give the impression that the impact of an intervention was similar across all sites when one site may account for most of the impact. Also, pooling of data limits the possibility of finding treatment differences that work best in certain contexts or with specific types of participants.

Use of Multiple Methods, Multiple Sites and Training Research Teams

One defining characteristic of evaluations that use multiple methods is that they inevitably require a research team to carry out the evaluation as the program is usually implemented in different schools. A number of concerns are closely linked between the use of multiple methods and multiple sites in an evaluation.

According to Turpin & Sinacore (1991), distinguishing characteristics of multisite evaluations include the analysis of original data as well as the following:

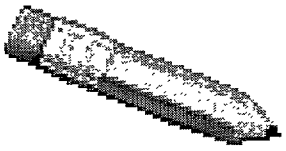
- Multisite evaluations can be prospective or retrospective. The research team designing the evaluation may plan to use multiple sites as part of exploring the effects of a program model on its subjects. Most IEQ evaluations have incorporated this characteristic in their evaluations usually exploring how an innovation (e.g, Escuela Nueva in Guatemala) or service (e.g., the effects of different levels of educare teacher training in South Africa) has had an impact on teachers and/or students. Or, multisite evaluations can be retrospective. Different evaluations on the same topic are brought together and an analysis is performed on the data.
- Multisite evaluations can focus on a program implemented in same way at different geographical sites.
- Multisite evaluations can focus on a program implemented in different ways at different geographical locations (program to examine different instructional methods in different settings like lectures vs. Independent readings vs. question/answer).

A major component of the multisite evaluation strategy involves the use of teams of researchers. No one evaluator can meet the increasingly complex demands of an evaluation. As program evaluations designs get more complex or demands from different stakeholders (administrators; teachers; students; parents; funders) get more intense, program evaluators are required to focus on special areas in evaluation. Some may be experts in sampling, others in data collection strategy, others in program administration or instrumentation. Additionally, the multi-year nature of this type of evaluations coupled with the need to investigate numerous contexts in a variety of settings places too great a strain on one lone researcher. The result is that no one person can meet the multiplicity of demands of the MSE. As Douglas (1976) notes, the Lone Ranger approach to program evaluation is no longer feasible and is giving way to the use of teams of researchers. In fact, Douglas goes on to describe team field research and in so doing, points out the various issues relevant to the conduct of multimethod and multisite evaluation when he states,

> "Team field research involves a number of people working together in a flexibly planned and coordinated manner to get at the multiperspectival realities of a group, constructing the team to achieve the research goals of the project in the concrete setting, utilizing the specialized abilities and opportunities of the various team members, providing both support and cross-checks on the work of each member by the other members, and all

> members (ideally) providing creative inputs to the research, the grasping, the understanding and the final report."(Douglas, 1976:194).

Multisite evaluations incorporate a team research-approach which allows use of the specialized characteristics, interests, talents and specialized knowledge of its members to the benefit of the project. (Douglas, 1976). This approach also provides for a division of labor; protects against biases of a single individual; creates a synergy due to discussions among team members makes for learning; leads to improved quality of data as team members carry out cross-checks; and can stimulate new areas of investigation based on findings by one team (Whyte, 1984). The important issues of planning, coordination/management, achievement of research goals, data collection and quality are all important aspects to the conduct of a successful evaluation. Each is addressed in subsequent subsections of this document.



Practice Exercise 3. Planning for Data Analysis

1. What type of data will be collected for your evaluation?

_____

_____

2. How are the data going to be summarized?

_____

_____

3. How will your data sets be linked?

_____

_____

4. How will your data be displayed to make it easy to analyze?

_____

_____

Look at Appendix B. Answer the following given the information provided in those pages.

A. How is the interview data displayed?

__ A. Table form
__ B. Frequency
__ C. Words

B. How is the data summarized?

___ 1. Not summarized
___ 2. Raw numbers
___ 3. Percentages

C. What strategies were used to link the information from the observations and teacher interviews?

__ 1. Interview data describes observation data
__ 2. Observation data elaborates interview data
__ 3. Observation data explains interview data
__ 4. Data sets are not linked

Training for Multiple Method Evaluations

Training of the field team is critical in any evaluation effort but especially in the use of multiple methods for multisites. Training will assure that all staff know the goals of the evaluation, the rationale for the methods to be used, the appropriate use of instruments, and the rules each member of the team is expected to follow for the duration of the project. Additionally, Turpin and Sinacore (1991) point out that training may be needed when there is a new topic for evaluation as well as when new activities required of evaluators. Also, training will provide the field team with the diverse expertise required to bring back appropriate and accurate data using on the topics of concern and using the data collection methods appropriately. Finally, training provides the team with "...a broad range of experience and a kit with many tools" (Anderson and Ball, 1978:183). Readers are referred to the IEQ's evaluation module on instrumentation (Chesterfield, 1996) for thorough discussion of training issues for evaluation. This subsection will touch on some issues of importance to multimethod, multisite evaluations in particular.

Program Implementation

There are some challenges to the use of the multisite evaluation strategy in program evaluations. One important difficulty is that the implementation of an intervention, especially in schools, is not identical at all sites. There is no standard evaluation situation: programs vary in activities, duration, number of objectives, clarity of goals, timing of effects, consistency of program over time, etc. which all impact on the design of the evaluation. A number of factors including the attitudes of implementers, the support of administrators, the demographics of the context, the training provided to teachers, and any other number of concerns will influence how staff perceive a program, how they understand it, how they implement it and the resultant outcomes.

Tip

Training of field teams must permit practice with:

- Gaining entree into the field
- Managing different roles
- Recognizing the components of an intervention

Additionally, a program may have a number of components to its overall intervention. Program staff may incorporate all, some or adjust some program delivery strategy "component" in attempting to adapt it to the local context. Thus, not only may a difference arise in how a program is implemented, but there are also differences in the degree to which an intervention is employed within a setting. Use of multiple methods for collecting information on a program intervention may afford practitioners with a better understanding of the outcomes and the processes that produced such results.

Entree and Cooperation

Given the varying conditions in which a multisite evaluation may be undertaken, it is critical that the evaluation team gain entree into the sites and the cooperation of the staff if the data to be collected is to be valid, reliable and meaningful. One difficulty in entering a site and gaining cooperation lies in how the evaluator is perceived by those implementing the program. Evaluator and practitioner viewpoints may be in conflict as practitioners associate evaluators with their supervisors or administrators (Finsterbusch and Motz, 1980). Thus, a single researcher going into an evaluation setting will almost inevitably be stigmatized as a member of one group by those in the conflicting groups (Douglas, 1976; Becker, 1970). Also, viewpoints will inevitably differ between administrators and evaluators on the use of findings, goals of the research, and the merits of the program.

The use of a team of researchers tends to alleviate some of the problems of gaining entree and cooperation during the initial phases of the evaluation. When team researchers are used, they enter the field together thus meeting key people at the same time. If one has difficulty with acceptance, the other(s) may be able to switch tasks so that data is not lost or imperiled. In those instances where a team cannot be assigned jointly, the "sociability specialist" is the first in the field making the initial contacts with local program practitioners and administrators. Traditionally, they have been women as they are perceived as less threatening in gaining entree with either gender. (Douglas, 1976). This sociability specialist is usually the friendly, well-liked team member with good social skills. Other specialists that may emerge are local project leaders and the finders (good at moving through an organization to find key people).

Another strategy for gaining cooperation is to develop contacts with national and regional gatekeepers and enlisting their support (Hendrick, 1991). In this instance, officials in the Departments of Education would serve to introduce the research team to the local sites at times via letters explaining the purpose of the research and the important role the local staff will play in policy development by participating in the study. In this instance cooperation is centralized in the team of evaluators.

One strategy that has worked in the IEQ is the naming of one researcher as a link to the site project directors. This evaluator will serve to keep abreast of any issues that arise at the local level that may influence program implementation or data collection. For example, during the educare evaluation, the South Africa IEQ team was divided into three distinct sub-teams. Each researcher was linked to several projects in a region of the country and was responsible for a) securing the cooperation of a site in participating in the evaluation; b) training of program staff in the use of the data collection instrument; and c) calling their local sites weekly to track data collection progress and problems. Additionally, each researcher served as resource person to the site providing information on evaluation and other issues related to educare. Training and retraining workshops of local staff were conducted to enhance cooperation as well as to upgrade skills, motivate site coordinators and promoted identification with study goals and the project.

Role Management

An integral component to gaining entree and cooperation is the ability to manage the multiplicity of roles that an evaluator may play. The researcher may be called on to select persons for the program, to assign the respondents to different treatments or interventions, to select sites from a pool, to be a confidant of the staff, to report unethical behavior (e.g., child abuse). The researchers may be called on to do any of the following:

A. To help administrators delineate program components & effective means for measuring outcomes (Weiss, 1972).
B. To present findings to maximize likelihood of utilization (Weiss, 1972)
C. To gain influence and build rapport to minimize possible resistance by staff & administrators to the evaluation efforts (Tharp & Gallimore, 1979).
D. To inform measures and outcomes as well as the design of who to include (who benefits) (Brown, 1980).

The role of evaluator, by its very nature is conflictive. Chesterfield (undated) notes the reactivity looming for an evaluator with regard to disclosure rules, confidentiality, unethical behavior, advocacy, or inappropriate discussion of field notes. For example, the researcher-interviewer may not intervene in situations yet will undoubtedly be asked by program staff for their opinion of a program component, the structure of a lesson or person's ability. Role conflict is especially likely to arise when the evaluator as part of the program operations staff is called upon to sell the intervention or treatment program then watch as the potential participant is rejected from the program because he or she may not have met the criteria and had to be left out of the study (Hendrick, 1991). A more likely scenario, however, is the role conflict that emerges when a site coordinator might be responsible for assuring that data is collected but has no authority over the researcher who is an employee of program as was the case with the South African IEQ teams.

Role conflict as well as many of the other issues that will arise in the process of carrying out the evaluation can be managed through regularly scheduled and ad hoc phone calls to each of the field researchers. Additionally, training provided to the field team should prepare them with the knowledge of the different experiences they will be likely to confront and the skills for handling them appropriately. Training or field manuals serve to reinforce the knowledge and as references once the researcher is in the local sites. Role conflict can be managed by training of field researchers designed to develop knowledge and skills for the following:

- What their roles will be and why;
- What evaluation design decisions have been made and why; and
- A delineation of expectations and guidelines for the roles.

Administration of Evaluations Using Multiple Methods

Staffing

Appropriate staffing is critical to the successful implementation of an evaluation using multiple methods. The demands of such efforts are reflected in the project goals, its needs for specialized knowledge or special opportunities (e.g. contacts within a central bureaucracy or program operation) (Douglas, 1976). Additionally, the composition of the field team requires flexibility to meet the demands of the context (e.g., special language requirements) and project (specialized knowledge of child development; sampling of specialized populations; etc.). Strategies for forming a team abound each with its own advantages and disadvantages.

Tip

Staffing may require fielding a team with special skills if specialized knowledge or instruments are need for the evaluation.

Turpin & Sinacore (1991) encourage the establishment of a staffing plan to permit evaluation managers to identify evaluation needs and the skills required of the field team. The elaborate nature of the plan is contingent on budget concerns, time constraints in each site, personal motivation of candidates and evaluation project leaders, and the authority that can be exercised over each site. Among the factors to consider in making staffing decisions are the following:

- Knowledge of the system in which the program is to be implemented;
- Length of field-period;
- Potential attrition;
- Difficulty of data collection instruments;
- The available pool of applicants; and
- Cost.

Additionally, Reiss and Boruch (1991) point to the need to have an administrative structure that can sustain the effort over time. The multi-year nature of this type of evaluation points also to the need for highly specialized or skilled project managers to design the study and supervise data collection. An administrative structure also enables efficient handling of daily organizational problems, needs and decisions (Turpin & Sinacore, 1991). The multi-year nature of the evaluation also points to the likelihood that any long-term projects will lose local researchers as they may begin to look for

other jobs toward the end of a project (Cottingham, 1991). Basically, though, the team is formed in one of two ways: one hires within system or out side of it.

One common manner of forming a team of field researchers is the use of existing staff of the implementing organization at the local site as site coordinators and data collectors (e.g., teachers, counselors, etc.). A variation on this strategy is to hire research staff at each site though the researchers are not part of the implementing organization. Members of a local research team may add a dimension of sensitivity to local context. The idea of local researchers also addresses the issue of going "cold" into a research setting, contacting one group and being stigmatized by the others as a member of that one clique or group (administrator, teacher, district person) (Douglas, 1976; Becker, 1970). In the long term, the local researcher also assists and advises other team members from the research organization on entree into the site, e.g., pecking order or who to talk with first. The incorporation of site staff into the evaluation is made easier when they are motivated, enthusiastic about the project and when local program administrators have the resources (including staffing) to allocate to the project. However, staff of implementing organizations may have high turnover rates due to fluctuations in program funding cycles and program restructuring. This will require training of alternates or constant retraining as new members are incorporated into the team.

Another strategy is to hire evaluation staff for the project in some central location. The staff then travel to sites for data collection. though this necessitates a large budget for evaluation. However, when staff researchers are data collectors they provide for greater consistency to what is studied. Data collection by staff researchers facilitate cross-site impact comparisons as they become familiar with all sites and have collected data across sites.

Training Program

The training program for multimethod, multisite evaluations should be field-based or structured in a manner that permits trainees to apply the techniques in as real-like a setting as possible. Some good points to assist in structuring the training are as follows:

- The instruction should take place in a seminar-like environment and enrollment should be limited to 15 trainees to permit uninhibited discussions as well as appropriate monitoring of trainees' work. The instruction should take place in a workshop- or working group-type environment that permits trainees to work together on tasks.

Tip

Make the training program content specific to your evaluation, field-based and focused on acquiring the skills needed for using instruments pertinent to your evaluation.

- Each new topic may be introduced by the seminar instructor to lay the theoretical foundation for the topic and to lay the logic underlying the research method. The goal of the instructor is not abstract discussions of research techniques but rather how the techniques are applied in real evaluation situations.
- Trainees need to practice with the instruments developed for the project. Initial exposure to the distinct data collection strategies can be made through common assignments. For example, trainees may be asked to observe events in a single setting. Ideally, the setting is accessible although unfamiliar where many activities may be taking place. Beauty salons, parking lots, department stores, grocery stores and such serve as likely places where behavior is repetitive and observable. (See Spradley, 1980; Twitchell, 1989).
- Practice with the protocols, once developed, should take place in the field, replicating the situations trainees are likely to encounter. For example, teacher interview protocols, classroom observations forms and parent interviews may be piloted in local schools. Where local schools are not accessible, videotapes of public schools that closely match those of the evaluation project's sample may be used. Trainees may be allowed to practice interviewing on one another during the workshop or by interviewing a teacher or other relevant person brought in for that purpose.

At a minimum, the training program should be structured to allow the members of the evaluation team to acquire the knowledge and skills to carry out their tasks in an responsible manner. Generally, the training program would want to touch on topics such as the following (Lohr, 1980; Mertens, 1994; Lewis, 1980; Sechrest, 1980):

- Evaluability of a specific program including its goals, objectives and legislative history;
- Definition of the intervention or treatment that includes sensitizing the team to its complexity or lack of specificity;
- Identification of independent and dependent variables; e.g., the treatment and outcomes (anticipated effects) as well as an examination of the magnitude and direction of effectiveness;
- Measurement strategies; validity, reliability and the relevance of the measures being used;
- Sampling procedures to be used in selecting subsamples for intense observation;
- Problems associated with confounding factors & potential plausible rival hypotheses;
- Analytic techniques available given the data to be collected;
- Problems in data collection, quality, and relevance, e.g. how they reflect valid measurement dimensions, are reliable and exhibit uniform quality;
- Inferential power inherent in the design;
- Timeliness of evaluation- when to begin and when to reach conclusions;
- Political context of the evaluation and the climate within each site;
- Perceived threats from management;
- Cost of evaluation;
- Need, if any, for rapid dissemination of results; and

- Ethics including privacy and confidentiality of the field notes, questionnaires, etc.

The actual training of field workers can take place in several ways. Usually, the training in the IEQ project has been through a workshops, as this is most efficient for achieving high quality. This strategy requires that a central location be identified where trainees of varying backgrounds will be comfortable. Using a five- to ten-day training schedule, the participants are exposed to experiential tasks designed to provide them with the skills that they will need to carry out the tasks in the 'real life' settings. Variable skill levels among the field team members, though, may require that those with limited research experience have more intensive training. Chesterfield (undated) notes that the structure of the training experience should provide for the following elements:

- Simulations of the actual fieldwork, e.g., conduct of observations, interviews, etc.;
- Experience with events that the field team will need to examine, e.g. observations of classroom behaviors relevant to the evaluation goals;
- Synchronized tasks so that no one skill will be learned in isolation; and
- Incorporation of peer sharing of field experience.

Training in the IEQ project has been of a short-term nature geared towards practitioners already involved in educational research or educational settings. This type of training model assumes that participants have a need to know the material; have problem to investigate; and have a setting to apply it (Wortman, Cordray and Reis, 1980). Retraining sessions are undertaken to reinforce the skills, the use of standard procedures or to solve problems.

Tip

Use a training model consisting of short-term training workshops; narrowly-focused topics covered in-depth; focus on skills acquisition and competency in using the evaluation instruments and procedures.

An important component to the training is the production and distribution of a training manual at workshops to ensure collection of high quality data and uniformity in data collection. Handbooks can be developed that address role descriptions and appropriate interaction by those in distinct positions within the structure of the evaluation project (Project manager, local project directors, field team members, etc.); procedures and guidelines regarding how to carry out tasks; and expectations from central office. The manual should also address all the topics of the training and serve as a reference to the field team members when confronting issues in the field. Training may be extended through the use of conference calls (Turpin & Sinacore, 1991). The IEQ/Washington project managers effectively used regularly scheduled conference calls to monitor the projects in all sites. With the advent of electronic mail, the Guatemala IEQ project made efficient use of the electronic mail system to monitor progress, exchange ideas among field staff, and provide rapid feedback on documents produced for the project.

Data Quality Control and Training of Field Teams

Data that is of high quality refers to having the researcher gather information that accurately describes the phenomenon under examination. The training of field researchers increases the probability of obtaining higher quality data. However, data quality is also related to research techniques. In selecting data collection strategies for an MSE, important to consider are the characteristics of the participants, the nature of the research situation and the types of measures used in the evaluation (Krause, 1996). In fact, Worthen and Sanders (1987) note that the success of site visits is dependent on the following factors:

- Identifying the specific information needed;
- Developing evaluation questions to be asked;
- Developing on-site instruments (interview forms, checklists, etc.);
- Selecting on-site researchers;
- Making pre-visit contact and arrangements; and
- Structuring the conduct of the visit including allocating an appropriate amount of time on site; having an initial on-site team meeting with program staff; arranging for an initial briefing by site administrators; structuring efficient use of team members; having interspersed team meetings; and conducting exit interviews with program staff.

Quality Control and Data Verification

Multimethod evaluations will inevitably rely on the collection of data by teams of researchers as this makes for more efficient use of resources. However, the use of teams of researchers will affect the nature and quality of the data obtained. Circumstances may emerge where sites may even be missing specific data sets thus limiting the possibility of cross-site analyses. Caracelli and Greene (1993) report on a multisite, longitudinal evaluation of the Research and Development Utilization program using mini-ethnographies, standard site visit reports, surveys and other data collection strategies. She notes that less than 20 percent of the sites had a complete data set. Thus, MSE managers need periodic audits to measure quality of data collection (Turpin & Sinacore, 1991). The audits can include computerized checks on data responses outside of expected range or random checks of data for accuracy & completeness. As was noted in the IEQ Guatemala evaluation, field notes were sent to the central office where a supervisor verified coding of the data. Data audits of this type, when carried out during an evaluation, permit redirecting field researchers to appropriate topics in line with the evaluation goals and purposes. Additionally, the audits will also increase confidence in the quality of the data. During the educare study, the IEQ team members used regularly scheduled phone calls to provide data collectors with feedback on problems and to correct data collection errors. In many instances, the phone calls were followed-up with faxes to demonstrates corrections or extensions of a concept.

Teams of researchers also serve as a means of data verification. Observation of the same setting by different team members allows for cross-checks of data. Also, team

research provides a manner for implementing triangulation strategies and thus viewing a phenomenon from multiple perspectives, e.g., teacher interviews, observations or observations of the classroom and home. Also, researchers will inevitably be confronted with circumstances where they will experience discomfort and a reticence to observe or interview. When there are research teams, its members can provide emotional support thus increasing the likelihood that a task will be carried out appropriately. Team members also can provide stimulation based on data queries that can lead to important site specific or cross-site findings.

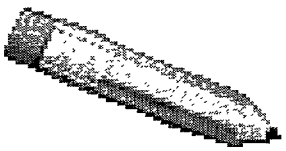
Quality Control and Supervision

Monitoring the team in the field is of utmost importance in the multimethod evaluation context. The supervision will influence the nature of the data as well as the type of data collected and analyzed. This in turn will affect the analysis procedures that can be carried out with the data. Supervision also includes providing the team with motivation when energies run low, when they have a bad day in the field or towards the last days of a particularly lengthy stretch of data collection. First-time interviewers may be interested in the interview for the first series of these. Then, they may be tempted to go through the motions as the process becomes repetitive and dull. They may fill out a questionnaire without first establishing rapport or securing the cooperation of a respondent. Whyte (1984) suggests a number of strategies for providing adequate supervision in the field. Among them were the following:

- Appoint and make sure that local team knows who is in charge;
- Have the field team turn in the instruments at end of day to a supervisor;
- Check each for omissions and discuss problems in field work;
- Makes spot checks in the field, re-interviewing certain respondents; and
- Secure local field supervisors to provide close supervision.

Quality Control and Cross-site Communication

Cross-site communication is needed to ensure consistency in data collection and to provide emotional and logistical support for members of the evaluation team stationed in the field (Hendrick, 1991). Evaluation managers need to balance supervision with need for field researchers and local program implementers to feel ownership of the procedures and data being collected. In essence, there is a need to balance a directive approach with one that guides the field research and motivates the team members. One manner of doing this is to work out an appropriate frequency and type of communication with team members. During the initial forming of teams for fieldwork, the educare group in South Africa defined how contact would be maintained and how often the team members would call each other. Phone numbers for hotels were exchanged and fax number made available to ensure that documents could be exchanged while the team was in the field. In turn, the field team communicated daily with the project managers in Washington, D.C. and Guatemala though electronic mail as a means of keeping them abreast of project progress.



Practice Exercise 4. Quality Control

1. Identify the strategies that you will use for maintaining quality control of the data.

Field Team Category	Communications	How often	How	Type of training Needed
Interviewer				
Observer				
Tester				

2. How will you monitor data collection?

_____

_____

3. How will you assure that hypotheses from different sites get known among field workers?

_____

_____

Report Writing

Trainees face daunting tasks in addressing the issue of writing of research reports. In many circumstances, they will be unfamiliar with tasks other than data collection. Thus, there may be little to no development of data analysis and report writing skills. In fact, trainees may note little relationship between the data collection task and the writing of the final report. Evaluation managers then need to address the different concerns related to the communication of evaluation findings. Among the principal issues to address in report writing are identifying the audience, identifying what they need to know; determining when results are needed; and determining the best format for communicating the findings. Several texts are available to guide trainees in writing of evaluation findings; among the texts are those of Morris, Fitz-Gibbons and Freeman, 1987; Herman, Morris and Fitz-Gibbons, 1987; and Elbow, 1973.

Tip

Plan for report writing:

- Who is the audience?
- What do they want to know?
- How do they want the information?
- When do they need it?

In addressing report writing, evaluators must address the issue of the audience. Evaluation team members should be warned about assuming that the reader knows anything about the evaluation. Additionally, training in report writing must address the level of language appropriate for the audience. Audiences can be multifaceted and include board members, funders, teachers, parents, other community members, and learners. For example, reports can be in text, table or graphic form depending on the literacy level of the audience.

Not all interested parties want to know everything about the evaluation. Thus, there may be a need to have one-page summaries that basically provide the major findings as these respond to the evaluation questions. Executive summaries are brief three- to five-page summaries that frame the questions, methods and major findings. Oral presentations may be required for key audiences such as executive boards or funders.

An evaluation will do no one any good if the results are not delivered in time for key decision-makers to act on them. Thus, evaluators must plan ahead to assure that reporting requirements can be met by laying out a reporting schedule with the various types of reports required of the client. The reports may include quarterly report that inform on the progress of the evaluation, problems encountered and solutions to the problems. The schedule must also list the major technical preliminary and final reports to be submitted for the evaluation.

Conclusions

Multi-method evaluations are an appropriate response to the increasingly complex nature of program interventions as well as to the need of policy makers and program administrators for information on the effectiveness of an intervention. When coupled with evaluation designs incorporating multisites, evaluations of this type work for obtaining information quickly, for purposes of generalizing or replicating program models and for obtaining practical results. Careful planning must be part of the design and implementation of any multi-method evaluation. Planning begins with defining the purpose and goals of the evaluation, as goal definition will influence the methods to be used, the sites selected, the staffing requirements, the type and length of the training as well as the proposed data analysis procedures.

The issues explored in this document begin to chart the terrain of multisite evaluations. The concerns addressed here serve to assist program managers in navigating the territory laid out for multi-method evaluations and in laying out priorities in executing the strategy. Additionally, use of multiple sites requires that evaluators pay close heed to the goals of the evaluation and transmit these to the research team contracted for the lengthy duration of the program. Strategies for entering the sites and gaining the cooperation of program implementers are similar as those for any evaluation; however, the use of a team of researchers in the conduct of site visits presents unique circumstances in solving problems with role management, entree and rapport building not available with other program evaluation strategies.

Evaluations incorporating multimethods also pose their unique challenges in data collection, storage and analysis. The exigencies of data collection call for careful scrutiny in identifying and hiring evaluation staff that can meet the needs of the evaluation design as well as those posed by the local context. Training can provide some of the necessary skills required to meet the evaluation goals. Differences in implementation of programs across sites beckon the analyst to examine the effects of an intervention within sites prior to conducting any cross-site analyses.

Future research into multisite evaluations needs to examine what general principles are worth applying especially with regards to how to best deal with situations where evaluator values differ from those of administrators.

Bibliography

Anderson, S. B. and Ball, S. (1978). The Profession and Practice of Program Evaluation. San Francisco.

Becker, H.S. (1970) Sociological Work. Chicago: Aldine Press.

Bieger, G.R. and G.J. Gerlach. (1996) Educational Research: A Practical Approach. Boston: Delmar Publishers.

Boruch, R.F. and H. Gómez. (1977). *Sensitivity, bias and theory in impact evaluations.* Professional Psychology. 8:411-434.

Boruch, R.F. and J. Reis. (1980). The student, evaluative data, and secondary analysis. In Sechrest, L. (Ed.). Training Program Evaluators, No. 8. San Francisco: Jossey-Bass, Inc.: 59-72.

Brown, E. D. (1980). Effective training of program evaluators: A Mixture of art and science. Sechrest, L. (Ed.). Training Program Evaluators, No. 8. San Francisco: Jossey-Bass, Inc.: 79-87.

Bryman, A. and R. G. Burgess, Eds. (1994). Analyzing Qualitative Data. New York: Routledge.

Caracelli, V. J. and J. C. Greene. (1993). Data analysis strategies for mixed-method evaluation designs. Educational Evaluation and Policy Analysis. 15(2):195-207.

Chesterfield, R. (Undated). "Initial Thoughts on Training". Personal communication.

Cottingham, P. H. (1991). Unexpected lessons: Evaluation of job-training programs for single mothers. In Turpin, R.S. and J. M. Sinacore. (Eds.) (1991) Multisite Evaluations: New Directions for Program Evaluation, No. 50. San Francisco: Jossey-Bass Inc.:59-70.

Crabtree, B. F. and W. L. Miller, Eds. (1992). Doing Qualitative Research. Newbury Park: Sage Publications.

Douglas, J. D. (1976). Investigative Social Research: Individual and Team Field Research. Beverly Hills: Sage Publications.

Elbow, P. (1973). Writing without Teachers. London: Oxford University Press.

Fienberg, S. E. (1980). Quantitative training for evaluation research. In Sechrest, L. (Ed.). Training Program Evaluators, No. 8. San Francisco: Jossey-Bass, Inc.: 49-58.

Finsterbusch, K. and A. B. Motz. (1980). Social Research for Policy Decisions.

Belmont, CA: Wadsworth Publishing Company.

Geertz, C. (1973). The Interpretation of Cultures. New York: Basic Books.

Goetz, J.P. and LeCompte, M.C. (1984). Ethnography and Qualitative Design in Educational Research. San Diego: Academic Press.

Hedrick, S.C., J. H. Sullivan, J. L Ehreth, M. L. Rothman, R. T. Connis, and W. W. Erdly. (1991). Centralized versus decentralized coordination in the adult day health care evaluation study. In Turpin, R.S. and J. M. Sinacore. (Eds.) (1991) Multisite Evaluations: New Directions for Program Evaluation, No. 50. San Francisco: Josssey-Bass Inc: 19-31.

Herman, J.L., L.L. Morris, and C.T. Fitz-Gibbon. (1987). Evaluator's Handbook. Newbury Park: Sage Publications.

House, E. R. (1994) Integrating the quantitative and qualitative. In C.S. Reichardt and S. F. Rallis eds. The Qualitative-Quantitative Debate: New Perspectives. San Francisco: Jossey-Bass:17-22.

Isaac, S. and W. B. Michael. (1995) Handbook in Research and Evaluation for Education and the Behavioral Sciences. San Diego: Educational and Industrial Testing Services.

Kibler, R.J., L.L. Barker, and D. T. Miles. (1970). Behavioral Objectives and Instruction. Boston: Allyn and Bacon.

Krause, D. (1996). Effective Program Evaluation: An Introduction. Chicago: Nelson-Hall Publishers.

LeCompte, M.D. and J. P. Goetz (1982). *Problems of reliability and validity in ethnographic research.* Review of Educational Research, 52 (1):31-60.

LeCompte, M. D., W.L. Millroy, and J. Preissle, eds. (1992). The Handbook of Qualitative Research in Education. San Diego: Academic Press.

Lewin, K. 1946. *Action research and minority problems.* Journal of Social Issues. 2:34-36.

Lewis, J. H. (1980). Training evaluators of social experiments and social programs. In Sechrest, L. (Ed.). Training Program Evaluators, No. 8. San Francisco: Jossey-Bass, Inc.

Lohr, W. (1980). Qualifications of evaluators: A federal perspective. In Sechrest, L. (Ed.). Training Program Evaluators, No. 8. San Francisco: Jossey-Bass, Inc.: 73-88.

Mager, R. F. (1962). Preparing Instructional Objectives. Palo Alto, CA: Fearon.

Mertens, D.M. (1994). Training evaluators: Unique skills and knowledge. In, J.W. Altschuld, J. W. and M. Engle, eds. The Preparation of Professional Evaluators: Issues, Perspectives and Programs. San Francisco: Jossey-Bass:17-27.

Miles, M. B. and A. M. Huberman. (1994). Qualitative Data Analysis. Thousand Oaks: Sage Publications.

Miller, W.L. and B.F. Crabtree. (1992). *Primary Care Research: A Multimethod Typology and Qualitative Road Map.* In Crabtree, B. F. and W. L. Miller, Eds. Doing Qualitative Research. Newbury Park: Sage Publications.

Morris, L.L., C.T. Fitz-Gibbon and M.E. Freeman. (1987). How to Communicate Evaluation Findings. Newbury Park: Sage Publications.

Morell, J. A. and E. W. Flaherty. (1978). *The development of evaluation as a profession: current status and predictions.* Journal of Evaluation and Program Planning. 1:11-17.

Mowbray, C. T. and S. E. Herman. (1991) Using multiple sites in mental health evaluations: Focus on program theory and implementation. In Turpin, R.S. and J. M. Sinacore. (Eds.) (1991) Multisite Evaluations: New Directions for Program Evaluation, No. 50. San Francisco: Josssey-Bass Inc.: 45- 57.

Patton, M. Q. (1987). How to Use Qualitative Methods in Evaluation. Newbury Park, CA: Sage Publications.

Pelto, P.J. and G.H. Pelto. (1978). Anthropological Research: The Structure of Inquiry. Cambridge: Cambridge University Press.

Pitman, M.A. and J. A. Maxwell. (1992). Qualitative approaches to evaluation: Models and Methods. In LeCompte, M. D., W.L. Millroy, and J. Preissle, eds. The Handbook of Qualitative Research in Education. San Diego: Academic Press.

Reichardt, C.S. and T.D. Cook. (1979). Beyond qualitative versus quantitative methods. In Cook, T.D. and C. S. Reichardt, eds. Qualitative and Quantitative Methods in Evaluation Research. Newbury Park: Sage.

Reiss, R. J. and R. Boruch. (1991). The program review team approach and multisite experiments: The spouse assault replication program. In Turpin, R.S. and J. M. Sinacore. (Eds.) (1991) Multisite Evaluations: New Directions for Program Evaluation, No. 50. San Francisco: Josssey-Bass Inc.:33-44.

Schulberg, H. C. and Baker, F. 1979. Evaluating health programs: Art and/or science? In. H.C. Schulberg and F. Baker (eds.), Program Evaluation in the Health Fields. Vol. 2. New York: Human Sciences Press.

Sechrest, L. (Ed.) (1980). Evaluation researchers: Disciplinary training and identity. In Sechrest, L. ed. Training Program Evaluators, No. 8. San Francisco: Jossey-Bass, Inc.

Sechrest, L. (Ed.)1980. Training Program Evaluators, No. 8. San Francisco: Jossey-Bass, Inc.

Shadish, Jr., W.R., T. D. Cook, and L. C. Leviton. (1991). Foundations of Program Evaluation: Theories of Practice. Newbury Park: Sage.

Skogan, W. G. and A. J. Lurigio. (1991) Multisite evaluations in criminal justice settings: Structural obstacles to success. In Turpin, R.S. and J. M. Sinacore. (Eds.) (1991) Multisite Evaluations: New Directions for Program Evaluation, No. 50. San Francisco: Josssey-Bass Inc.:83-95.

Spradley, J. (1980). Participant Observation. New York: Holt, Rinehart & Winston.

Strauss, A. (1987). Qualitative Analysis for Social Scientists. New York: Cambridge University Press.

Strauss, A. and J. Corbin. (1990). Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Newbury Park: Sage Publications.

Turpin, R.S. and J. M. Sinacore. (1991). Multiple sites in evaluation research: a survey of organizational and methodological issues. In Turpin, R.S. and J. M. Sinacore. (Eds.) (1991) Multisite Evaluations: New Directions for Program Evaluation, No. 50. San Francisco: Josssey-Bass Inc.: 5-18.

Temple, B. (1994). *Combining methods: Epistemological considerations in the study of families and households.* Journal of Family Issues. 15(4):562-573.

Turpin, R.S. and J. M. Sinacore. (Eds.) (1991) Multisite Evaluations: New Directions for Program Evaluation, No. 50. San Francisco: Josssey-Bass Inc

Twitchell, J.B. (1989). Preposterous Violence: Fables of Aggression in Modern Culture. New York: Oxford University Press.

Webb, R. B. and C. Glesne. (1992). *Teaching Qualitative Research.* In The Handbook of Qualitative Research in Education. M.D. LeCompte, W.L. Millroy, and J. Preissle, eds. San Diego: Academic Press.771-814.

Weiss, C. 1972. Evaluation Research: Methods of Assessing Program Effectiveness. Englewood Cliffs: Prentice-Hall.

Weiss, C. 1973. Where politics and evaluation research meet. Evaluation 1:37-45.

Weisz, J.R., W. Chaiyasit, B. Weiss, K.L. Eastman, and E.W. Jackson. (1995). *A multimethod study of problem behavior among Thai and American children in school*: Teacher reports versus direct observations. Child Development. 66:402-415.

Whyte, W.F. (1984) Learning from the Field: A Guide from Experience. Beverly Hills: Sage Publications.

Worthen, B. R. and J.R. Sanders. 1987. Educational Evaluation: Alternative Approaches and Practical Guidelines. New York: Longman.

Wortman, P.M. 1977. *Evaluation research: New training programs for new careers.* Professional Psychology. 8:189-192.

Wortman, P.M., D.S.Cordray and J. Reis. (1980). Training for evaluation research: Some Issues, in Sechrest, L. (Ed.)1980. Training Program Evaluators, No. 8. San Francisco: Jossey-Bass, Inc.

Appendix A:
Instruments for Learner-Centered Teaching Evaluation

Learner-Centered Teaching Observation Form

School: _____
Date: _____
Observer: _____

Teacher: _____
Grade: _____
Observation period: _____ to _____

Interactions	Initiator		Context				Teacher Behavior								Receiver Response		
	T	Stdnt	SmG	WhC	Pr	In	Ch Exp	Dircts	Explns	Op-En Q's	Demo	Probes		Lang	V	NV	NR
												V	NV				

Comments:

_____

_____

Learner-Centered Teaching Classroom Checklist

School: _____ Teacher: _____
Date: _____ Observer: _____
Classroom period: _____

Use this instrument in three different activities which occur during the visit to the classroom. Indicate **PRESENCE, ABSENCE**, or **USE** of each of the following in the classroom:

Classroom Organization	TIME 1		TIME 2		TIME 3	
	Present	**In Use**	**Present**	**In Use**	**Present**	**In Use**
Flexible work space for children						
Learning centers/corners are present						
Children move among desks, materials and learning centers						
Child current work is displayed						
Safe learning materials						

Yes = Present No = Absent

Comments:

_____

_____

Learner-Centered Study Teacher Interview Guide

Teacher: _____ Date: _____
School: _____ Observer: _____
Time: _____

Hello. Thanks for letting me sit in your classroom today. I saw many things and I would appreciate your help in sorting them out. I will not take much of your time- this should last for about 30 mins. Please know that what we talk about will be used only for looking at what's going on in the classroom, and you will never be identified in any way. Now, may I ask you a few questions, please?

THE CLASSROOM PROFILE

1. How many boys are enrolled in your class? _____ Boys

2. How many girls are enrolled in your class? _____ Girls

3. What is the age of the youngest child enrolled in your class? ___ Age of youngest

4. What is the age of the oldest child enrolled in your class? ___ Age of oldest

5. What is the dominant language spoken by the learners in your classroom?

6. Do you speak the language of the learners in the classroom?

CLASSROOM ACTIVITIES.

7. What did you want to achieve in today's lesson?

8. Were there two or three things you wanted the learners to acquire from today's lesson?

9. How did you hope to achieve this?

10. How do you identify or assess individual children's needs?

11. How do you assess individual children's performance in the classroom?

12. How do you assess your own performance ?

MATERIALS

13. What kinds of materials do you use in conducting your lesson?

14. How did you decide on the materials you used in your lesson?

CLASS MANAGEMENT

15. Who makes classroom rules?

16. How do you ensure that rules are followed?

17. Can you tell me how classroom activities are chosen?

18. Do you make use of peer leaders during your lesson? If yes, How are the leaders chosen?

19. Do you group the learners in your classroom? IF YES, How?

20. Have you gained any knowledge from your learners?

TEACHER PROFILE FORM

This information is just for purposes of analysis.

19. How long have you been teaching?

20. How long have you taught this particular subject?

21. How many inservice training programmes have you been involved in?

22. Was there anything different about today that you would like to talk about?

23. Is there anything you would like to ask me?

Thank you very much. I appreciate the time you took from your work to talk with me about the day's events.

Appendix B:
Data Analysis and Data Integration on the Learner-Centered Teaching Study

Data Analysis and Data Integration on the Learner-Centered Teaching Study

The following examples are used to illustrate the various strategies available for analyzing qualitative data. The examples are taken from the data collected during the training exercises conducted in South Africa for the working groups. Examples are provided for data display, data reduction, and for the linking of two data sets. For purposes of seeing how data is linked, the illustrative examples used are from the Learner-Centered Teaching workshop.

Data display is important in that it is the first step in organizing and managing the information collected from respondents. In this first step, responses to a question are placed in some order that permits description and analysis. The researcher, then, labels, codes, indexes, classifies or otherwise names the units of data. Data displays permit the researcher to initiate the cataloguing of responses to allow the identification of themes, patterns or other conceptual contents. Exhibit __ displays data from the responses to the question on the Teacher Interview in Learner-Centered Teaching Study asking whether peer leaders are used in the classroom and who selects them. The first column identifies the school from which the response came; the second column classifies or labels the response. The display below is in the form of a table. However, displays may consist of lists, matrices, diagrams, outlines or textual displays. In the example below, global statements about findings would be difficult to discern quickly. By reducing the response categories, the researcher is able to key in on the issues central to the evaluation objectives.

Exhibit 1. Data Display for Information from Teacher Interview

School	Have Peer Leaders and Who Selects Them
1	No
2	Yes, with teacher
3	Yes, rotation
4	Yes, rotation
5	Yes, with teacher
6	Yes, with teacher
7	No
8	No
9	Yes, with students
10	Yes, with students

Data reduction permits the researcher to manage the information in a more efficient manner and to initiate the process of identifying relevant categories, an important aspect of qualitative data analysis. Data reduction strategies are closely linked to the evaluation objectives. In the question above, the evaluation objective focused on whether learner-centered learning occurred in the classroom. As part of the objectives, children were to select their own peer

leaders in the classroom. As is evident above, four categories of responses can be easily identified. The researcher would need to keep separate those categories that are important to the evaluation objectives. Thus, one could reduce even further and have three basic categories: "Yes, with others"; "Yes, by students", and "No". In the example below, the four basic categories were not reduced. In this example, the categories emerged from the data. At other times, the researcher may have categories that have been identified as important through the conceptual questions posed and the review of literature undertaken to examine how other have approached similar studies.

Exhibit 2. Data Reduction

Schools	Original Response	New Response Category
2, 5, 6	Yes with teacher	1, Yes, with teacher
3, 4	Yes, rotation	2, Yes, rotation
9, 10	Yes, students	3. Yes, students
1, 7, 8	No	4, No

Data Linking occurs when information from two or more sources are brought together and analyzed to identify associations among the variables. During the various work groups, several instruments were developed to permit collection of data using the most appropriate means. Thus, workshop participants developed observation forms, teacher interview guides and classroom checklists. Linking data from several sources permits the testing of hunches or hypotheses that the researcher may have gleaned during the data collection phase. The following table links information from the teacher interview and from the observation form from one study. The workgroup examined learner involvement in the classroom. The example below links the question on years of experience teaching collected from the teacher interview and the item on the observation form recording the use of small groups by teachers.

Exhibit 3. Linking of Information from Interviews and Observation Form

School	Years Teaching - Teacher Interview	Use of Small Groups - Observation Form
1	8	Yes
2	15	No
3	5	Yes
4	1	Yes
5	20	Yes
6	10	No
7	2	Yes
8	20	No
9	5	Yes
10	15	No

The researcher now can identify patterns related to years of teaching experience and the use of small groups in the classroom. From the exhibit above, it is readily apparent that 80 percent of teachers with greater than 10 years of teaching experience do not use small groups in their classrooms.

Data display, reduction and linking processes all permit the researcher to order and manage the information collected for the evaluation. The display and other processes of data analysis are then used to produce a narrative description of the evaluation findings.