

# **EVALUATING EXTENSION PROGRAM OUTCOMES**

Agricultural Extension Service  
Institute of Agriculture  
University of Tennessee

Prepared by:  
Richard L. Poling  
Associate Professor  
Agricultural and Extension Education

## Evaluating Extension Program Outcomes

	<b>Page</b>
Extension Program Evaluation Worksheet .....	2
“Asking the Right People” .....	3
“For the Right Information” .....	5
“The Right Way” .....	14
Data Collection Methods .....	19
Mail Surveys .....	19
Group Administered Surveys (Post-Program Surveys) .....	20
Telephone Surveys .....	20
Face-to-Face Interviews .....	21
Group Interviews (Focus Group Interviews) .....	22
Observation .....	22
Case Studies .....	23
Analysis of Existing Information .....	23
Documentation .....	24
Sources .....	24

## EVALUATING EXTENSION PROGRAM OUTCOMES

by

Richard L. Poling

Associate Professor, Agricultural and Extension Education  
University of Tennessee Agricultural Extension Service

In determining the effect of Extension educational programs, the objectives of the program are usually the guides from which potential program outcomes are determined. However, many outcomes of Extension programs are serendipitous in nature. We don't always plan on all the outcomes that actually do occur. Keeping that in mind, the program objectives are still the best places to start in developing a program evaluation plan.

The best time to begin developing the evaluation plan is when the program planning process itself begins. In each step of the program planning process, decisions are made that will affect how the program evaluation will be developed. The answers to the questions - What will the program subject matter include? Who will be the recipients of the program? What are the expected outcomes of the program? and What methodologies will be used in the delivery of the program? - all contribute to the and implementation of the program evaluation. Well-written program objectives address design most of these questions and become the cornerstones of the evaluation.

An easy way to approach developing an evaluation plan is to remember to do the following:

Ask the **RIGHT PEOPLE**  
for the **RIGHT INFORMATION**  
in the **RIGHT WAY**.

This means that, to evaluate a program, one must first identify the right sources of information that will provide information to determine the outcomes of a program. Once the sources have been identified, the information needed from those sources that address the expected program outcomes must be determined. Then, that information must be collected using data collection techniques that are appropriate for the sources of information and the environment in which the program is being conducted.

A more structured outline of items that need to be addressed when planning a program evaluation is shown in Figure 1. As each item is addressed using this form, more information is provided as to how the evaluation will be conducted, including: what are the expected outcomes of the program (objectives), what questions need to be addressed, what sources of information will be used, what methods of information collection will be used, and how will the results of the evaluation be used.

FIGURE 1. EXTENSION PROGRAM EVALUATION WORKSHEET

Use the following as guidelines in developing an evaluation plan for Extension programs.

- 1. Program, project, or activity to be evaluated: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_
- 2. Objective(s) of program [including level(s) of Bennett/Rockwell's TOP Model Hierarchy]:  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_
- 3. What questions are you trying to answer with the evaluation?  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_
- 4. Who are the stakeholders for the program being evaluated?  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_
- 5. What information is needed by stakeholders from the evaluation?  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

6. What/who are the information sources for the evaluation?

- 7. What information collection method(s) is/are appropriate?  
\_\_\_\_\_  
\_\_\_\_\_
- 8. How and to whom will evaluation results be reported?  
\_\_\_\_\_  
\_\_\_\_\_
- 9. What resources will be needed to conduct the evaluation? (Money, personnel, equipment, etc.)  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Prepared by Richard L. Poling, Evaluation Specialist, University of Tennessee Agricultural Extension Service.

## "ASKING THE RIGHT PEOPLE"

Asking the right people means identifying the best sources that can provide the information needed to evaluate a program. Program objectives are critical tools in identifying information sources. Objectives should identify program recipients, subject matter, and outcomes. With this knowledge, the program evaluator can identify the most appropriate sources of information. The following should also be taken into consideration when identifying the right sources of information.

1. A Source - the unit from which information is collected.

An information source may be a person or persons or the evidence left by a person. There are a variety of different sources that may be identified for a particular program evaluation. The selection of sources needs to be based upon certain criteria. How an information source is chosen should be based on the following:

- a. The source should be able to provide information related to evaluation needs.
- b. The source should be available. (May not be accessible.)
- c. Possibility of involvement. (May refuse to participate.)
- d. Cost restrictions to getting information. (Information collection costs may be very expensive.)
- e. Time involved.
- f. Resources needed.
- g. Credibility of information received.
- h. Usability of information received.
- i. Evaluator's skills.

### EXAMPLES OF INFORMATION SOURCES:

#### HUMAN SOURCES

Program participants	Relatives/friends
Teachers/Supervisors	Program providers/staff
Community leaders/officials	Program administrators
"Experts"	General public
Advisory committees	Professional colleagues
Fund sources	

#### SOCIAL INDICATORS

Studies	Reports
Complaints	Agency records
News articles/reports	Public documents

## PHYSICAL EVIDENCE

Observation of outcomes of:

- skill changes
- practice changes
- end results

In deciding on information sources, especially people, the individual receiving the program will often be the source of information about program outcomes (Ex. knowledge or skill change or practice adoption). There may be times when people other than the program participant would be the appropriate source of information. For example, if an objective of a program was to have young children adopt new behaviors, parents or teachers of the children might be good sources of information as to whether or not those behaviors have been adopted.

**"FOR THE RIGHT INFORMATION"**

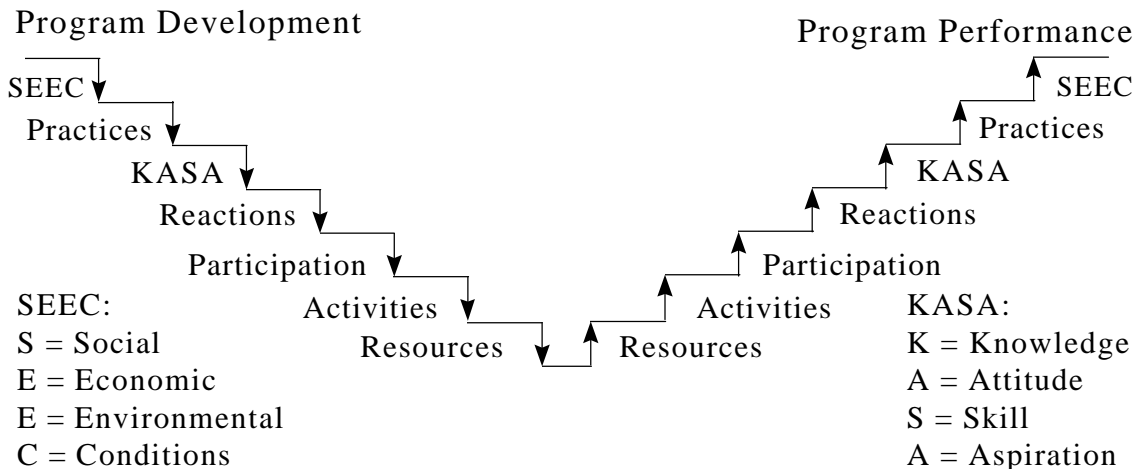
Clear, measurable program objectives are the best source about what information should be collected in a program evaluation. Objectives should be written so that they identify what are the expected outcomes of the program and at what level of the Bennett/Rockwell Targeting Outcomes of Programs (TOP) Model Hierarchy are the outcomes. The TOP Hierarchy is a model for program development, objective writing and evaluation widely used in Extension work. The TOP Model consists of seven levels of expected inputs and outcomes that take place through Extension programming efforts. The hierarchy and brief descriptions of program inputs and outcomes at each level are shown in Figure 2.

Looking at the hierarchy, one can see that levels 1 (Resources) and 2 (Activities) describe what Extension does to produce and conduct educational programs. Levels 3 (Participation) through 7 (SEEC) are representative of outcomes that result from participation of clientele in Extension educational programs. Levels 5 (KASA), 6 (Practices) and 7 (SEEC) are more often seen as the real "impact" levels of Extension programming. In general, the following statements are found to be true in relation to the levels of the Bennett/Rockwell TOP Model Hierarchy:

1. Evidence of program impact becomes stronger as the hierarchy is ascended.
2. The difficulty and cost of obtaining evidence on program accomplishments increases as the hierarchy is ascended. (NOTE: Costs include time and personnel, as well as dollars.)
3. Evaluations are strengthened by assessing Extension programs at several levels of the hierarchy.

Figure 2.

**Bennett/Rockwell TOP Model  
(Targeting Outcomes of Programs)**



To determine the degree of attainment of program objectives, specific criteria are needed. These criteria identify what measures or data will be collected to demonstrate that the objective has been reached. These measures may vary in quality. This variation in data quality is often referred to as "hard" versus "soft" data. "Hard" and "soft" data represent opposite ends of a continuum, not a dichotomy.

The "hardness" of data is based upon the three qualities of validity, quantification, and representativeness. Validity of data is the extent that they truly reflect the characteristics of subjects or situations being studied. The dimension of quantification relates to the amount or degree of differences that exist between data. Differences are usually represented through the assignment of numerical values to the observations that can then be used to show how much the differences are among subjects on the measures used in the evaluation. Representativeness is the extent to which observations of subjects or situations being evaluated are applicable to a larger population or situation. Representativeness can be enhanced by studying the entire populations or a representative sample of the population.

The degree of "hardness" of data depends upon trade-offs between ideal data and the data needed and resources available to collect data. "Hard" data are more expensive and difficult to obtain and should only be collected when the benefits to decision-making from superior evidence clearly outweigh the costs of obtaining such evidence. Stakeholders in the Extension program should be involved in the determination of the "hardness" of data needed and what trade-offs will be made to collect the type of data needed. Stakeholders include program planners, supervisors, and program funders.

There are many situations where "soft" data on accomplishment of objectives are all that can be obtained. Program participants are often unwilling or unable to be observed or to respond to instruments which require detailed answers and/or extensive time to complete. A guide for determining the "hardness" of measures used as evidence of achievement of objectives is to collect evidence of a "hardness" that will meet the needs of stakeholders and the resources available to collect the evidence. Table 1 shows examples of "hard" and "soft" data at each level of the TOP Model Hierarchy.

#### What Questions to Ask:

At each level of the TOP Model Hierarchy, there are certain questions that need to be answered in the evaluation process to determine if objectives have been achieved. The questions' structures will depend on at what level of the hierarchy is the objective and , possibly, on whether the data is to be "hard" or "soft." Below are examples of types of questions or items that will collect data to address the evaluation questions at different hierarchy levels. The examples presented are not the only ways to ask the questions.

Table 1. Examples of "hard" and "soft" evidence in the Bennett/Rockwell TOP Model Hierarchy

LEVEL	EXAMPLES	
	"Hard" Evidence	"Soft" Evidence
7. SEEC Change	Trends in profit-loss statements, live birth weights of babies, reduction of pesticide residues in well water sources.	Perceptions of changes in quality of economy, health or the environment.
6. Practices Change	Direct observation of use or evidence of clientele use of recommended practices.	Self-reported use of recommended practices by clientele.
5. KASA Change	Changes in scores on validated measures of knowledge, attitudes, skills and aspirations. Direct observation of skills.	Self-reported perceptions of change in participants' knowledge, attitudes, skills, and aspirations.
4. Reactions	Direct observation or measurement of all or a representative sample of participants reactions during and/or after a program.	Recording the views of only those who volunteer reactions about the program (non-representative sample).
3. Participation	Use of sign in sheets, recorded observation of individuals participating or holding leader positions.	Casual observation of attendance and leadership given by participants.
2. Activities	Pre-structured observation of activities and social processes through participant observation, use of video or audio tapes, etc.	Staff recall of how activities were conducted and the extent to which they were completed.
1. Resources	Documentation of staff time and resource allocations through a formal process.	Staff's subjective reports regarding time allocation.

## RESOURCES LEVEL

*What were the total number of staff days expended in the planning and conducting of this program?*

*How many dollars were expended in the planning and delivery of the program?*

## ACTIVITIES LEVEL

*How many educational activities and programs were conducted in \_\_\_\_\_?*

*How many news articles were published related to the \_\_\_\_\_ program?*

## PARTICIPATION

*How many individuals participated in educational programs on \_\_\_\_\_ as indicated on sign-in sheets?*

*How many people served as members of the program committee for the \_\_\_\_\_ program?*

## REACTIONS

Reactions are the responses of clientele as to the value or quality of programs, teachers, teaching methods used, materials, facilities and other factors involved in the program. Reactions should not be confused with KASA changes. A person indicating that the program was well presented and met their needs is not documentation for knowledge or skill change.

### "Hard" Data

Observation of reactions of program participants during the program presentation. Experimental designs can be used to provide "hard" data of the value of programs to participants (e.g. providing a "distraction" to participants that might lure them away from participation if they are not really interested in the program).

### "Soft" Data

Asked of participants using either a check list (Check All That Apply), a dichotomous scale (ex. YES/NO) or a Likert-type scale of three or more levels (ex. STRONGLY AGREE to STRONGLY DISAGREE):

(NOTE: Blanks in items are for the subject matter or topic that was taught.)

*This workshop helped provide a good background on \_\_\_\_\_?*

## REACTIONS (Cont.)

*The program provided little useful information about \_\_\_\_\_.*

*The presenter appeared to be knowledgeable about \_\_\_\_\_.*

*The written materials used in this program were appropriate.*

*This was one of the best programs on \_\_\_\_\_ that I have attended.*

Open-ended reaction questions can also be asked, but remember that you will probably not get as many responses to open-ended questions and open-ended data will be more difficult to summarize and to analyze the results.

*What were the three things that you liked the most about the program?*

*How could the program have been improved?*

## KASA CHANGE

KASA changes are measured in several ways depending on whether "hard" or "soft" data are desired. "Hard" measures of knowledge, attitude (opinion) and skills changes are obtained either through validated measurement instruments or observations and pre-test, posttest measurements "Soft" measures of knowledge change are self-reported perceptions of participants as to how they feel their knowledge has changed. "Soft" measures could also use pre-test, posttest measurements to provide slightly "harder" data.

### **Knowledge:**

#### "Hard" Data

Pre- and post-program scores on a knowledge achievement test are compared to determine knowledge change. The test must be validated to ensure that it is actually measuring the knowledge you want. Responses to test items can be True/False, multiple choice, or fill in the blank. Open-ended essay questions can be used, but validation, analysis, and summary of such questions will be difficult.

#### "Soft" Data

Items for responses from participants. Responses may be similar to those described for reaction level items: True/False, Yes/No, Strongly Agree to Strongly Disagree, Check All That Apply.

*As a result of my participation in this program, my knowledge of \_\_\_\_\_ has increased.*

**Knowledge (Cont.):**

*The information provided by the agent has increased my knowledge about*  
\_\_\_\_\_.

*As a result of my participation in this program, my knowledge of the following topics has increased:*

*(LIST TOPICS COVERED AND HAVE PARTICIPANT RESPOND TO EACH TOPIC)*

**Attitude (Opinion):**

"Hard" Data

Pre- and post-test scores on a validated instrument that measures the attitudinal construct being addressed by the program. A number of such instruments have been developed and are available for use (often at a cost) when appropriate for the program. NOTE: Attitude change is often difficult to achieve through short-term programs. Objectives that state a change of attitude as an outcome with minimal contact with participants are very difficult to achieve and should be given much thought before being adopted.

"Soft" Data

Items for responses from participants. Responses may be similar to those described for reaction level items: Yes/No, Strongly Agree to Strongly Disagree, Check All That Apply.

*I feel more positive about my abilities to succeed in school as a result of this program.*

*I believe more in the importance of collaboration in achieving group goals than I did before participating in this program.*

**Skills:**

Skills can be measured the same way as knowledge, by perceptions or by actual testing of skills. If you are teaching skills that can be performed, why not have participants perform the skills for you at the end of the program?

"Hard" Data

Observation of participants demonstrating learned skills serves as evidence of skill development. A checklist of skills demonstrated is used for documentation. Pre-program skill levels should also be demonstrated, if possible, to document change.

**Skills (Cont.):**

"Soft" Data

Items for responses from participants. Responses may be similar to those described for reaction level items: Yes/No, Strongly Agree to Strongly Disagree, Check All That Apply.

*As a result of my participation in this program, my ability to \_\_\_\_\_  
\_\_\_\_\_ has increased.*

*I can \_\_\_\_\_ better than I could before participating in this class.*

**Aspirations:**

"Hard" measures of aspiration are difficult to imagine. Aspirations are really self-reported perceptions ("soft" data) of participants as to the likelihood that they will do something or change the way they do something as a result of participating in the program. NOTE: Aspirations are not the same as Practice Change. Aspirations are indications that participants intend to use or adopt what was learned, not that they have already used it.

Items for responses from participants. Responses may be similar to those described for reaction level items: Yes/No, Strongly Agree to Strongly Disagree, Check All That Apply.

*As a result of participating in this program, what practices or behaviors do you  
plan to adopt, or utilize more often? \_\_\_\_\_ (Practices/Behaviors)*

*I plan to adopt, or utilize more often, the following practices that I learned in this  
program:*

*(LIST PRACTICES/BEHAVIORS INCLUDED IN PROGRAM)*

**PRACTICES (CHANGE):**

Practice change is a measure of changes in or adoption of behaviors or practices by participants. Measures can be actual observed behavior/practice change or self-reported perceptions of change/adoption.

"Hard" Data

Observed adoption of new practices, techniques, or behaviors that were learned through participation in the program. Practice adoption can be documented on a checklist by the observer. Secondary evidence of adopted practices/behaviors can also be observed.

PRACTICES (CHANGE): (Cont.)

"Soft" Data

Items for responses from participants. Responses may be similar to those described for reaction level items: Yes/No, Check All That Apply.

*As a result of participating in the \_\_\_\_\_ program, I am now using the following practices:*

*(LIST PRACTICES PRESENTED IN PROGRAM)*

*Based on what I learned in the EFNEP program, I now prepare more nutritious meals for my family.*

SOCIAL, ECONOMIC, AND ENVIRONMENTAL CONDITIONS (SEEC)

SEEC changes are things that happen to clientele and/or communities as a result of knowledge and practice/behavior changes of program participants. SEEC changes are sometimes hard to document. But, in some cases, they might be fairly easy to document, especially when quantitative data has already been collected as a part of the system being measured (ex. farm records, demographic statistics). Sometimes, linking SEEC changes to Extension programs appears to be difficult. However, by demonstrating efforts and results at other levels of the hierarchy, a better case can be made that Extension programs have been a factor in the SEEC change achieved.

"Hard" Data

"Hard" measures of end results would include documented effects of changes program participants have made. Examples would include: increased profits for business/farm operations, documented reduction in specific health problems in communities where programs have been implemented, reduction in environmental problems in areas where program participants have adopted new practices.

"Soft" Data

Perceptions (not supported by quantifiable data) of participants, community members or officials as to changes that have occurred as a result of programs. Items for responses from participants. Responses may be similar to those described for reaction level items: Yes/No, Strongly Agree to Strongly Disagree, Check All That Apply.

*The economy of the community has improved as a result of Extension community development programming.*

*Extension water quality programs have resulted in fewer contaminated wells in the county.*

SOCIAL, ECONOMIC AND ENVIRONMENTAL CONDITIONS (SEEC): (Cont.)

*The highways and roads are a lot cleaner as a result of the Keep America Beautiful program in \_\_\_\_\_ County.*

*No matter what questions you determine need to be answered at what level of the hierarchy, it is extremely important that you ask only for the information that is necessary to the evaluation and that you ask for the information in the simplest fashion possible!*

## **"THE RIGHT WAY"**

When the right information sources have been identified and the right information to be asked has been determined, the next step is to make sure that we ask for the information in the right way. Data collection is as much an art as it is a science. What data collection methods to use and how to apply them in particular situations is a multi-layered set of decisions that the evaluator must make for each program being evaluated.

There are several considerations that have to be addressed in order to select the most appropriate data collection methods. These include:

### The Kinds and Amount of Data That are to be Collected

- Do they imply specific methods or instruments?  
  
ex. Actual knowledge change implies use of an achievement test.
- Do they imply a particular evaluation design? (comparison groups, multiple measurements, random assignment or random sampling of participants)  
  
ex. If you are interested in evaluating the performance of participants versus non-participants, you would collect data from both groups.
- Are they "hard" or "soft" data?  
  
ex. "Hard" measures of Practice Change require direct observation, "soft" Practice Change measures may be self-reported by participants using a survey instrument.

### The Needs of Evaluation Users/Stakeholders

- What do they expect? Noted evaluation expert, Michael Patton says, "It is critical that intended users participate in the making of measurement and methods decisions so that they understand the strengths and weaknesses of the data -- so they believe in the data." (p.224, Utilization-Focused Evaluation, 1986)  
  
ex. Supervisors who participate in the development of program evaluation plans will accept the evaluation measures that result from that plan.
- What will they accept?  
  
ex. Quantitative results or Qualitative results, "hard" or "soft" data, sample of participants or all participants.

## Information Sources

- Are there any special attributes of the information source(s) that would indicate that certain methods would be more appropriate than others. Such attributes could include:

Size of the Population or Sample - A larger population or sample size could result in the use of methods that result in less direct contact with subjects due to costs of collecting larger amounts of data. (Ex. A mail survey instead of face-to-face interviews.)

Location or Dispersion of Subjects (Near or Far Away, Clustered or Widely-Dispersed) - Subjects located near to the evaluator or located in close vicinity to each other can be contacted using face-to-face or telephone interview methodology easier and at less cost than if they were at a distance or scattered over a wide area, in which case mail surveying might be the cost-effective method. Participants meeting as a group for a program offer the easiest access for the evaluator through the use of a group administered survey or interview.

Age (Very Young or Very Old) - Very young children may not be able to read a mail survey, so interviews may be a better choice. Responses such as Strongly Agree or Strongly Disagree may not have much meaning to children, so "smiley" faces may be used to represent responses of children. Parents or teachers may be good sources of information about knowledge or behavior changes in children.

Older people may have problems with sight or hearing. Surveys may have to be produced in large print. Interviewers may have to speak louder or plan on additional time in which to complete the interview. Older people may be more suspicious of strangers trying to collect information about them. More time and effort may be needed to explain the purpose of the evaluation and provide adequate identification.

Ability (Mental, literacy, physical, language) -

How do you measure individuals with widely differing mental abilities (both high and low)? May have to develop instruments and collection methods differently for each group.

Adjustments in instrumentation and collection methodology may have to be made for populations with low reading/writing abilities. Somewhat similar to young children, but you must be careful not to treat adults like children.

Adjustments in data collection may be necessary to insure that physically disabled individuals have the same opportunities to participate as information sources as those who are not disabled. Orally administered surveys for blind participants, visually administered or signed interviews for deaf participants, or access to facilities for focus group interviews for a person in a wheelchair.

When collecting data from populations who do not speak English, instruments and interview schedules will need to be translated carefully into the subjects' native languages. Those administering the instruments will also have to speak and understand the language. Translations into other languages need to be done with caution as certain phrases have different meanings in other languages.

Cultural Differences -

Concepts and terms used in survey instruments or interviews may not be easily interpretable across cultures. Testing of instruments and evaluation questions with individuals from the culture is a good idea.

A common cultural stigma is a reluctance to divulge information to those outside of the culture. Subjects may refuse to participate or may provide invalid responses. The use of trained individuals from within a culture to collect evaluation data may be necessary. Additional time and effort may be needed to develop trust with participants.

## Socio-Economic Status -

Accessibility may be a problem. Lack of telephone or frequent moving of subjects may make it difficult to contact individuals through mail surveys or phone interviews. Interviewers may be reluctant to go to subjects' location for face-to-face interview due to fear of possible harm, either real or perceived. Again, the use of trained individuals from the socio-economic background of the subjects may help to make the necessary contacts.

Socio-economic status (SES) is very similar in nature to a cultural difference. Participants from one socio-economic level may be suspicious of those outside of their SES. They may refuse to participate as an information source or give socially acceptable or invalid responses. Similar to cultural differences, the use of individuals for data collection from similar SES backgrounds as the participants may be appropriate. Otherwise, additional time and effort will be needed to develop trust between the participants and the evaluator.

The adjustments in data collection methods, use of additional individuals as data collectors and the development of instrumentation that fits the special needs of the information sources will result in the collection of more valid and complete data. However, it must also be kept in mind that these adjustments will result in higher costs in conducting the evaluation in terms of both dollars and time. The trade-off of more complete and valid data versus costs is a major decision to be made by the evaluator.

## Limitations

The following are limitations that may or may not hamper an effective program evaluation. Each limitation or a combination of limitations will result in decisions about what evaluation methods will be most appropriate for a given situation. Understanding these limitations during the evaluation planning stage will help to avoid an evaluation beyond the means available to conduct it.

- Money (includes costs of hiring additional personnel, travel, telephone, printing instruments, mailing surveys, etc.)
- Time (time involved in developing and testing instruments, training personnel, collecting data, entering results, analyzing data, and preparing reports)
- Resources (in addition to money and time, facilities and equipment use are resources)

- Skills/Abilities of the Evaluator (certain methods of data collection require certain skills and abilities that the program evaluator may not possess or have the ability to acquire)

*The key to selecting appropriate data collection methods is to identify a method or methods that best addresses the above considerations. A combination of methods could end up being the most appropriate choice.*

## DATA COLLECTION METHODS

Data collection methods can be categorized as quantitative or qualitative. However, these classifications are made on the basis of the type of data collected rather than on the method itself. A particular method can be used to collect either quantitative data, qualitative data or both. There are situations where one method is more appropriate for collecting a certain types of data. The following information will identify some of those situations.

### SURVEYS

Surveys can be used to collect information on attitudes, beliefs, behaviors, reactions, and attributes of respondents. They can be relatively inexpensive ways to collect information from a large number of people. Surveying a sample of a population can provide information about that population. However, samples must be carefully selected to avoid misinterpretation of results. Surveys can also be used to collect data from an entire population and give many persons in that population a feeling of being involved in the process.

Results from surveys may also be misinterpreted depending on how questions and response categories are designed. There is also a tendency for survey data (especially quantitative data) to be limited in nature, often reporting what "is" without addressing the underlying reasons "why" or actual behavior patterns.

Survey questions can range from structured close-ended responses (yes-no, true-false, agree-disagree, demographics) that provide quantitative information to unstructured open-ended responses that provide qualitative data.

The most commonly used survey methods are: mail surveys, group administered surveys (after meeting survey), telephone surveys, and face-to-face interviews.

### **Mail Surveys**

The mail survey is a cost-efficient method for the volume of data collected. However, the costs are increasing. People are likely to provide responses to a well-designed and conducted mail survey. Even though low response rates can be a problem (especially with specialized groups), good management of the mail survey process, including follow-up of non-respondents, can result in response rates of 70% and higher. Use of the mail survey technique by private organizations and businesses to solicit money and sell products has made some people reluctant to respond to mail surveys. The intent and importance of the survey must be clearly outlined to these individuals to assure them that the surveyor is not "selling" something. Assurances of confidentiality or anonymity are very important, especially when the information collected in the survey is sensitive or controversial.

The survey instrument used in a mail survey must be developed with care. The respondent will have no one to clarify information for them as they fill out a mail survey, so the instrument

must be simple and easy to understand. This means that a lot of work must be done in the development of the survey instrument to guarantee its validity, reliability, and clarity. Pre-testing of the instrument is a must.

Mail surveys are not very effective in the collection of qualitative data. Open-ended questions can result in responses that often depart from the intent of the question due to misinterpretation by respondents. In these cases, it is very difficult to summarize responses to open-ended questions.

### **Group Administered Surveys (Post-Program Surveys)**

Group administered surveys are similar to mail surveys except they are administered directly to individuals who are participating in a group activity. Everyone attending a meeting, etc. completes a survey form. This method allows for the administration of pre- and posttests to identify changes, as well as the traditional post-meeting survey. This method is an inexpensive way to collect data from individuals. A high response rate is usually possible with this method, especially if the participants are made aware of the importance of the survey and how it will be used to improve programs and provide accountability for the program. Because the surveyor is present, it is much easier to clarify items and respond to questions from individuals, which can allow for the inclusion of more open-ended items. It is also easier to guarantee anonymity for respondents due to the lack of a need for identity of respondents on the survey.

This method does allow for possible influence of the surveyor on results and the mood of the meeting may influence individual responses, either positively or negatively. Group administered surveys allow only for the collection of data from those present at that activity.

### **Telephone Surveys**

Conducting a survey by telephone is competitive in cost per response with mail surveys. Generally, response rates are high. Once people answer the telephone, they are likely to agree to respond to a well-designed telephone survey process. Telephone surveys usually do not allow for the collection of data from as many individuals as do mail surveys, depending on the amount of resources available and the size of the population being surveyed. The actual time to collect data is much shorter for telephone surveys than for other methods, but as much or more time is needed to develop a telephone survey instrument and train additional interviewers, if used, than is needed for mail surveys.

Telephone numbers used in telephone surveys can be obtained from several sources. Phone numbers of program participants can be obtained from sign-in sheets from the program or activity. Telephone directories are a good source of numbers for the general public, but unpublished numbers can prevent certain individuals from being included. Random digit dialing (RDD) can help by accessing all working numbers in the area of the survey. RDD can result in contacts with non-usable numbers (businesses, public and private agencies, etc.). If a specific

subgroup is of a population is being surveyed, it would be necessary to have phone numbers for those individuals.

The presence of an individual asking questions over the phone can provide clarification for respondents. Telephone interviewers must be careful, however, that their voice inflections or statements don't bias responses. Training of interviewers is a necessity to avoid this from happening.

Telephone interviewing forms can be more complex than a mail survey instrument. More complex skip patterns can be included in a telephone interview form that can simplify the elimination of non-applicable questions for respondents. There are limitations to the lengths of questions and number of response categories per question on a telephone survey form because the respondent isn't able to visualize the survey form. Telephone interviewing is a good technique to use in the collection of qualitative data through open-ended questions, especially if the interview is tape recorded for accuracy of responses. In this sense, a telephone survey is similar to a face-to-face interview. Computer assisted telephone interviewing (CATI) is also making the use of telephone surveys much easier.

### **Face-to-Face Interviews**

Face-to-face interviewing as a survey method is the most expensive method when compared to mail and telephone surveys. It is also the slowest method in terms of actual data collection. Face-to-face interviewing is not very appropriate for collecting data from a large number of individuals in a population due to the high cost of resources (interviewing personnel, travel, data management and analysis) and time involved. It is, however, a good way to collect in-depth qualitative data from individuals. For specific audiences (the poor, the mobile, high status individuals, etc.), face-to-face interviewing may be the only way to collect data. When special segments or individuals of a population have been identified as information sources, it may be easier to interview these individuals face-to-face. Response rates for this method of data collection are likely to be high as individuals are more reluctant to refuse to talk to someone face-to-face.

For collecting in-depth qualitative data from individuals, a face-to-face interview is preferable to other methods. Quantitative data from close-ended questions can also be collected using face-to-face interviews, but this same information can often be collected quicker and at a lower cost by either mail survey or telephone survey.

The presence of the interviewer in face-to-face interviews can both help and hinder the data collection. The interviewer can use probes to gather more in-depth information and pick up on non-verbal cues from the respondent. Some people prefer the personalized touch of having a "real person" to respond to. However, the interviewers presence and personal characteristics may be a biasing factor for how some people respond to questions. Individuals may be less frank or thoughtful about responses than they might if they were filling out a mail survey or responding to an unseen individual in a telephone survey. The lack of anonymity may bother some respondents.

If individuals other than the program evaluator are conducting the face-to-face interviews, they must be trained in how to avoid biasing responses by their actions or questioning. If possible, face-to-face interviews can be tape recorded to insure accuracy of responses, as well as to include some idea of the intensity and personality in responses.

## OTHER METHODS

In addition to surveys, there are other methods for collecting data, both quantitative and qualitative. These methods can be used by themselves or, more likely, in conjunction with one of the survey methods to collect data accurately representing people's attitudes, beliefs, behaviors, reactions and attributes.

### **Group Interviews (Focus Group Interviews)**

Group interviews, sometimes called focus group interviews (FGI) if "focused" on a particular topic, provide in-depth qualitative data similar to that of face-to-face interviewing. Because more individuals are involved in each interview, group interviews are usually less expensive and collect data faster than the face-to-face process. Group interviews are usually not used for the collection of quantitative data.

In addition to getting responses from individuals to questions, group interviews also provide an atmosphere where responses from individuals stimulate additional thoughts by others in the group. This synergy usually results in gaining information beyond what even the interviewer expected. In fact, group interviews can be very challenging to the interviewer in terms of keeping the discussion and responses on track with the topic at hand without stifling the enthusiasm of group members to respond to the questions and each other.

In a group interview situation, more outgoing or aggressive individuals can dominate the responses and intimidate less dominant members of the group. The interviewer must control this situation and insure that all group members have the opportunity to present their own viewpoints. The skills of the group interviewer is very influential in how accurate the results of a group interview will be.

Several authors have developed procedures for conducting focus group interviews that address solutions to the potential problems that may bias interview results.

### **Observation**

Observation is the systematic gathering of behavioral actions and reactions through the use of specific instruments or professional impressions. Information can be collected in a number of ways: by watching, listening, and documenting what is seen and heard; by asking questions; by sharing activities and noting comments, behaviors, and reactions; or a combination of these. It is important, however, that the observation and documentation be carried out in a

systematic manner so that the information obtained is as free as possible of bias and is relevant to the focus of the study.

Observation of individuals can be done by the evaluator or others trained to systematically observe and record information. If more than one observer is used, inconsistencies between observers will bias results. Training of observers is very important. The observers may make their identities known to the subjects or may keep their identities a secret. They may choose to participate in the activity being observed or may observe passively. If individual(s) are aware that they are being observed, they may behave differently than normal, but the principles of right of privacy and informed consent may require that they be made aware. The observer can also lose objectivity as a result of actually being involved in the activity.

Observation is useful for studying small units such as a class or committee, but is not realistic for studying large groups. Observations may not be valid for the entire population. If generalization of results to the population is desired, a plan needs to be developed to insure that those observed are representative of the population. Results of observation can be combined with other techniques to add to the quality of data collected.

### **Case Studies**

A case study is a method to provide in-depth information about a single program or situation. A variety of other data collection methods, such as surveys, interviews and existing documents, are used to capture the total picture of the program. The case study technique often finds significant information about program processes or impacts that may not have been anticipated. This method is generally utilized with a small number of individuals or groups.

Case studies involve extensive amounts of data collection and can be very time consuming. Specific techniques for data collection will be tailored for the individual case. The results of a case study can be very subjective, requiring judgements by those using results. Results of a case study are best used as background information and as a guide for further study or action.

### **Analysis of Existing Information**

Before any new data are collected, a thorough check should be made of what information is already available through existing sources. Although extensive data on the program or individuals included in your study may not exist, there may be background information or particular variables available that could save time and costs in your data collection process. Existing information will usually support information collected through other methods.

Sources of existing data are reports, historical accounts, summaries of previous surveys or reviews, meeting minutes, budgets and staff records. These sources may include information that has been collected specifically for the purpose of documenting programs or has been collected for other purposes but serves as a measure of the program being studied or information that is

part of the administrative database of the organization being studied. Access to some of these sources may be limited due to privacy concerns if it relates to information about individuals.

Data from existing records, if available, is low cost and can be accessed on a continuing basis. The user may have to spend time selecting and sorting the data. Problems sometimes exist with incomplete data or estimated data as opposed to actual accounts. Data from these sources have a tendency to be descriptive rather than looking at relationships.

## DOCUMENTATION

One of the most important aspects of the program evaluation process is documentation of evaluation efforts. Documentation of what took place in the program evaluation serves as evidence of accountability to those who are interested in how the results were established. Documentation can also be used to describe the evaluation process to those who might disagree with the results or doubt the validity of the evaluation.

Documentation is a matter of having a record of the evaluation process from its development in the program planning process to its conclusion with the analysis of the data collected. Items that need to be a part of the documentation are:

- the program objectives, documents indicating the input and activities levels of the program (agendas, written materials used in the program
- notices used to announce the program, etc.)
- instruments used to collect evaluation data, and
- summaries of evaluation results.

Some of this information may already be collected through other existing data collection procedures (civil rights documentation, MIS, program planning documents).

## SOURCES

Bennett, C. F. (1976). Analyzing impacts of Extension programs (ESC-575). Washington, DC: Extension Service, USDA.

Bennett, C. F., & Rockwell, K. (1996, draft). Targeting outcomes of programs (TOP): an integrated approach to planning and evaluation. Washington, DC: CSREES, USDA.

Dillman, D. A. (1978). Mail and telephone surveys: the total design method. New York: John Wiley & Sons.

Patton, M. Q. (1986). Utilization-focused evaluation. Beverly Hills, CA: Sage.

Warner, P. D., & Maurer, R. C. (1984). Methods of program evaluation.  
Lexington, KY: University of Kentucky.