LEARN HOW TO COLLECT EVALUATION DATA

K. S. U. (Jay) Jayaratne, Ph.D.
Professor and State Leader for Extension Evaluation
North Carolina State University

COLLECTING EVALUATION DATA

There are two methods of collecting evaluation data from Extension programs. The most common method is self-administered surveys, and the second method is observations.

SURVEY METHOD

When you use the survey method, you have two options to conduct the survey. You can use either a paper-based survey method or an online survey method. When you conduct face-to-face training programs, it is a good idea to use paper-based surveys to get a good response rate because you have an opportunity to collect evaluation data before your audience leave. Online surveys are appropriate for conducting distance delivering program evaluations and follow-up evaluations. However, if your audience has access to smartphones, the use of an online survey for collecting evaluation data by using online surveys is also a practical option. The following evaluation tools have been developed for you to use when conducting paper-based evaluations. However, if you want, you may use the given survey templates to build your data collecting survey online. If you intend to do so, first develop the data collecting survey by modifying the survey templates provided on this site. Next, use that template to build the online survey using an online survey building software. As an employee of NCSU, you have access to the *Qualtrics* online survey building software program. Here is the link to *Qualtrics* at NCSU: https://isa.ncsu.edu/surveys/qualtrics/accessing-qualtrics-at-nc-state/ This site explains how to create an account and build surveys on this software program.

1.1 How to Determine the Appropriate Evaluation Tool

The type of training program or activity will determine which tool to use. For the purpose of training evaluation, Extension programs and activities can be broadly divided into four categories:

1. Short training sessions (One-time educational activity or presentation conducted within 2 hours or less)
   Example: A training workshop on soil sampling or healthy lifestyle education

2. Long training sessions (One-time educational activity or presentation that is at least longer than 2 hours)
   Example: Food safety training for food service employees

3. Multi-session training programs (Educational program presented in series to an audience for more than once over an extended period of time)
   Example: a) 4-H club activities
   b) Nutrition and exercise education programs

4. Train-the-Trainer training programs
   Example: Master Gardener volunteer training

1.2 How Does the Length of a Program or Educational Activity Determine the Type of Evaluation Tool?

The type of evaluation varies with the training program or activity. The following guidelines are helpful for selecting the appropriate evaluation tool based on the type of Extension program or activity.

1.2a Evaluation of One-Time Short Training

If the training is a short educational presentation or an activity conducted within a few hours (less than 2 hours), you may not have time to conduct a pre- and post-evaluation. The best option is to use a retrospective pre- and post-evaluation tool to collect data at the end of the presentation or activity.

Advantage: The retrospective pre- and post-evaluation is easy to conduct.
Both pre- and post-test data are collected at once, eliminating the need to match data from two separate (pretest and posttest) evaluations.
Disadvantage: If the audience is not familiar with this method, it may cause some confusion, especially among low-literate audiences.

Measuring Outcomes:

a. Change in perceived knowledge,
b. Change in skills, and
c. Levels of aspiration (readiness to change)

1.2b Evaluation of One-Time Long Training

If the training is relatively long (Longer than 2 hours), the educator has adequate time to conduct a pre-evaluation at the beginning and a post-evaluation at the end. By conducting pre- and post-evaluations, you can compare participants' changes in knowledge and skills and document the training outcomes.

Advantage: More accurate than the retrospective method.
Since pre/post evaluation is less complicated, this method is appropriate for low-literate audiences.

Disadvantage: It takes more time to administer pre- and post-evaluations. Participants must be identified for matching pre- and post-surveys. Assigned identification numbers can replace names for matching surveys.

Measuring Outcomes:

a. Change in knowledge,
b. Change in skills, and
c. Levels of aspiration (readiness to change)

1.2c Evaluation of Multi-Session Programs

If the program is multi-session, “stages to change evaluation” is the best option to document the actual behavior changes of participants. In multi-session programs, the Extension educator has an opportunity to meet the same group of people more than once and observe changes in their behavior or practices. The stages to change evaluation is based on Procheska’s (1994) transtheoretical model and records the participant’s stage of change related to the planned behavior or practice before and after the program. By recording the stage of change, we can document the behavior change of the participants. Similarly, if the program is relatively long, we can record participants’ socio-economic data related to the program objectives before and after the program and compare them for the impacts. For example, in the case of an obesity reduction, nutrition, and exercise program, by recording participants’ body weight at the beginning and end of the program, we can document the reduced level of obesity as the program’s impact.

Advantages: Stages to change evaluation is helpful to document participants’ actual behavior or practice change as an intermediate outcome of the program. This method helps document even slight behavior changes of the participants.

Disadvantages: This type of evaluation is possible only with multi-session programs. Participants must be identified for matching pre- and post-surveys. Assigned identification numbers can replace names for matching surveys.

Measuring Outcomes:

a. Change in behaviors or practices
b. Change in skills

1.2d Evaluation of Train-the-Trainer Programs

Training workshops presented to individuals who use that information to educate others are called train-the-trainer programs. These Extension workshops are focusing on building participants’ knowledge about the particular subject and improving their teaching skills. Normally, train-the-trainer programs are relatively long. Therefore, pre- and post-evaluations can be considered as the best option to document the immediate outcome of train-the-trainer programs.

Advantage: More accurate than the retrospective method.

Disadvantage: It takes more time to administer pre- and post-evaluations. Participants must be identified for matching pre and post surveys. Assigned identification numbers can replace names for matching surveys.

Measuring Outcomes:

d. Change of knowledge,
e. Change of teaching skills, and
f. Levels of aspirations (readiness to apply)

1.3 What is Important in Designing Evaluation Instruments?

The quality of the evaluation instrument significantly contributes to the quality of evaluation data. Therefore, it is important to pay due attention to designing evaluation instruments. The following guidelines are helpful in designing quality evaluation instruments.

1.3.1 Determine the Required Evaluation Data and Information

Understand the purpose of your Extension evaluation. The main purpose of an evaluation is to document Extension impacts and improve Extension programs. We should decide what data and information are needed for these two evaluation purposes.

1.3.1a Data Needed for Extension Outcomes and Impacts

The extent to which participants have changed or benefited as a result of their participation in your Extension program or activity is referred to as the Extension outcome. Long term outcomes such as socio, economic, and environmental changes are called impacts. Depending on the type of your program or activity, you need to identify possible outcomes and impacts, as well as indicators for those outcomes and impacts. Collecting data for these indicators is necessary to document the outcome and impact of the program for accountability purposes. This set of evaluation data helps Extension document the program's value. Examples for outcome data are the participants' change in knowledge and skills. Examples for impact data are participants' practice changes, increased income, reduced health risks, and improved environmental conditions.

1.3.1.b Data and information Needed to Improve the Program

Identify strengths and weakness of your Extension program or activity to build on current strengths and eliminate existing weaknesses. Further improvement of Extension programs can be achieved only by evaluating the program delivery process to identify what worked and what didn't during the training. Also, you need to know for whom it worked and for whom it didn't. Such information and data can help you increase the cost-effectiveness of Extension programs. These data and information include some demographic data, instructor rating, educational material rating, program strengths, and weaknesses.

1.3.2 Limit the Number of Questions to the Needed Minimum

Questions should be limited to the minimum number necessary for collecting outcome and impact data and program improvement data. Short evaluation tools save time for both you and the participants. It is important to keep evaluation questions short and simple to facilitate data collection. Adding any unnecessary questions may confuse participants, impair the reliability of the instrument, and reduce the response rate.

You need to collect quantitative and qualitative data for Extension evaluation. Quantitative data refer to numbers, and qualitative information refers to non-numeric data such as stories, short written responses, and expressions made by the participants. There are strengths and weaknesses associated with both quantitative and qualitative data. For example, quantitative data generally provide strong evidence for program accountability purposes. However, quantitative data do not reveal reasons for not achieving or exceeding

program objectives. Qualitative data such as program strengths and weaknesses are needed to find reasons for the success or failure of your program.

1.3.3 Reliability and Validity of Data

High-level reliability and validity are essential characteristics of quality instruments for conducting effective evaluations. *Reliability* refers to your instrument's consistency in recording what you measure. It is the ability of your instrument to duplicate the measurement. If the evaluation instrument has a high level of reliability, it will record the same outcome each time you use it under the same condition. Clarity of questions and instructions directly contributes to the reliability of instruments. Make sure your questions and instructions are clear.

Validity refers to the accuracy of your measurement. Validity is the extent to which your evaluation instrument is actually measuring the outcome it should record. For example, if your instrument is measuring knowledge, it should not measure attitude or aspirations. There are different forms of validity. When you develop evaluation tools, content validity and *construct validity* are important determinants of the quality of measurements. The term construct validity refers to how accurately a concept, such as knowledge, is measured.

The instrument templates have been set up with adequate construct validity for you to measure knowledge, skills, aspirations, and behavior/practice changes. When you modify these templates, you must insert questions related to the content of your Extension program. For example, if your program is a food safety education program you should include food safety-related questions to test participants' knowledge. Content validity also ensures that you asked the right questions and that the questions you asked reflect what was taught in the Extension program or training session.

1.3.4 Write Questions Clearly and Concisely

The most common evaluation tools are self-responding surveys. Evaluation questions should be written clearly and concisely to avoid ambiguity and help the participant respond appropriately and establish the reliability of your evaluation instrument. Instructions should be clearly stated to help participants complete the survey easily. Generally, we use close-ended and open-ended questions in evaluation instruments. When closed-ended questions are used we need to provide answer choices. Multiple choice questions are examples for closed-ended questions. In open-ended questions, we don't provide answer choices. An open-ended question is asked in such a way that the respondent must use his or her own words to answer it. For example, what do you suggest to further improve this program?

When close-ended questions are used, make sure that the answer key contains all possible responses to prevent response errors. Closed-ended questions are easy to answer and analyze. Since open-ended questions are exploratory, they are appropriate for collecting information needed for program improvement. It is important to ask easy questions at the beginning and harder questions at the end to facilitate the response.

Collecting demographic information such as race and education level is somewhat challenging because generally, participants do not like to reveal demographic data. Therefore, it is not a good idea to include demographic questions at the beginning of your evaluation instrument. It is better to include demographic questions at the end. If you include

demographic questions at the beginning of the instrument, respondents may tend to withdraw from the survey without completing it.

When you write questions, think about the reading level of the target audience and write at or below the target audience's education level to avoid potential errors in collecting evaluation data.

1.3.5 Participants' Rights

When you collect evaluation data, pay special attention to protect the rights of the participants, and ensure that participants' privacy is not violated. Data must remain confidential. Participation in the evaluation should be voluntary, and the participant should be given the option to withdraw at any time without penalty. When you collect data from children below 18 years of age, make sure to get parental consent. This is important for 4-H youth development programs.

1.4 How to Design Evaluation Instruments

Your evaluation instrument should include the sections necessary for collecting the following three types of information:

I. Impact data
II. Program improvement data
III. Information to facilitate the data collection process

1.4.1 Step I. Collecting Impact Data

After you identified the outcome and impact indicators for your Extension program or activity, you should design sections to collect outcome data for those indicators. Depending on the type of impact indicators you identified, any of the following sections can be included in the design of the evaluation instrument (data collecting survey). You do not need to include all of the following sections, only those that are appropriate for your program evaluation.

1.4.1.1 How to Evaluate Participants' Change in Knowledge

The most common outcome of any Extension program or activity is a change in knowledge. Change of knowledge can be measured only if we record participants' knowledge about the subject before and after completing the program. There are two methods to document participants' change in knowledge.

1.4.1.1.a Retrospective Pre and Post Knowledge Evaluation

Retrospective evaluation measures participants' perception about their knowledge related to the subject or content covered in the program. Ask how the participant rates his or her knowledge about the training subject on a Likert-type scale. It is called *retrospective* because you ask participants to recall, at the end of the activity, what they knew when the program began to identify the change in knowledge. The

participant's knowledge can improve, stay the same, or decrease, depending on the effectiveness of the program or activity. Retrospective pre- and post-knowledge evaluation is appropriate for short training workshops or activities.

Advantages: Retrospective pre- and post-evaluation saves time because it is administered only once at the end of the program or activity. You don't need identifiers for matching pre- and post-evaluation surveys.

Disadvantages: Retrospective evaluation is too confusing to use with low-level readers.

Example: Retrospective Pre- and Post-Evaluation Tool for Recording Knowledge Change

Please circle the appropriate number to indicate your perceived level of knowledge about the following topics **before** and **after** completing the program. Please use the following key for rating:

1. Very Low = Don't know anything about this topic.
2. Low = Know very little about this topic
3. Moderate = Know about this topic but there are more things to learn
4. High = Have a fairly good knowledge but there are things to learn
5. Very High = Know almost everything about this topic

How do you rate your knowledge about:	BEFORE THIS WORKSHOP					AFTER THIS WORKSHOP				
	Very Low	Low	Moderate	High	Very High	Very Low	Low	Moderate	High	Very High
Planning a healthy meal.	1	2	3	4	5	1	2	3	4	5
My pyramid.	1	2	3	4	5	1	2	3	4	5
Making healthy food choices.	1	2	3	4	5	1	2	3	4	5
Saving money at the grocery store.	1	2	3	4	5	1	2	3	4	5
Nutritious value of beans.	1	2	3	4	5	1	2	3	4	5
Harmful effects of eating fried foods.	1	2	3	4	5	1	2	3	4	5
Nutritious value of fruits and vegetables.	1	2	3	4	5	1	2	3	4	5

1.4.1.1b Pre- and Post-Knowledge Evaluation

In this method, we ask a set of questions from the participants just before the program and ask the same set of questions at the end of the program and compare results to determine the change in participants' knowledge. You can use either multiple-choice or true/false questions. If you use multiple choice questions, your evaluation tool will be relatively long and need more time to complete it. This is not desirable in managing your training time. The true/false type question format is helpful to keep the evaluation tool brief and save respondents' time. This true/false answer format is appropriate for low-level reading audiences as well. Knowledge testing questions used with true and false answer format is appropriate to evaluate youth programs because knowledge acquisition is a reliable impact indicator for youth programs. The more questions used, the greater the reliability of

measurement; however, as the number of questions increases, the length of the survey and response time also increase. You need to balance reliability with practicality. Normally, 10 questions provide a reasonable level of accuracy for Extension evaluation. You may use the "don't know" answer choice to avoid the guessing error. It is a good idea to use true questions and well as false questions to increase accuracy.

Example: Pre-- and Post-Evaluation Tool for Recording Knowledge Change

Please circle the appropriate answer to each of the following questions.

1. Food high in saturated fat increases risk for heart disease. True False Don't Know
2. Fruits and vegetables contain fiber, which helps prevent constipation. True False Don't Know
3. Vitamin A found in many fruits and vegetables helps our bodies absorb iron. True False Don't Know
4. The Nutrition Facts label on foods tells you how many calories and nutrients are in one serving. True False Don't Know
5. Milk is a rich source of vitamin C. True False Don't Know

1.4.1.2 How to Evaluate Participants' Change in Skills

Someone's level of confidence to carry out a learned practice or task is a reflection of his or her level of skill to perform that task. Therefore, we use the levels of confidence as an indirect measure of someone's skills. Participants will be asked to rate their levels of confidence to carry out a learned task or practice before and after completing the program using a Likert-type scale. Change in skills is evaluated by comparing pre- and post-evaluation results.

Example: Pre- and Post-Evaluation Tool for Recording Skills

Please circle the number that best describes your* confidence *to do the following:

How confident are you in your ability to:	**Not confident**	**A little confident**	**Somewhat confident**	**Confident**	**Very confident**
1. Plan a healthy meal?	1	2	3	4	5
2. Reduce fat intake?	1	2	3	4	5
3. Make healthy food choices?	1	2	3	4	5
4. Interpret nutrition labels?	1	2	3	4	5
5. Fix healthy meals at home?	1	2	3	4	5

1.4.1.3 How to Evaluate Participants' Aspirations for Changes

Participants' aspirations means their inclination to taking charge of the behaviors or practices learned from an Extension program or activity. Aspiration is an appropriate indicator to document the immediate outcome of Extension programs and activities. If any Extension program or activity is effective, the evaluation will indicate participants' readiness to apply learned behaviors and practices. The following format can be used to record participants' aspirations to adopt learned behaviors and practices. Aspiration is recorded only at the end of the program or activity.

Example: Evaluation Tool for Recording Aspirations (Taking Charge)

Taking Charge

Please circle the number that best describes your answer.

As a result of this program, do you intend to:	No	Maybe	Yes	Already doing this
1. Eat recommended servings from five food groups?	1	2	3	4
2. Plan meals ahead of time?	1	2	3	4
3. Consume reduced- or non-fat milk and dairy products?	1	2	3	4
4. Eat whole grain bread and cereals?	1	2	3	4
5. Eat baked or grilled food rather than fried food?	1	2	3	4

1.4.1.4 How to Evaluate Participants' Behavior or Practice Changes

In a multi-session Extension program, participants' actual behavior change is evaluated by recording their behavior at the beginning of the first session and at the end of the last session. For all the other Extension programs and activities, follow-up evaluation is needed to assess participants' actual behavior changes.
In a multi-session Extension program, the agent meets the same group of participants more than once. As a result, he or she has an opportunity to record participants' actual behavior before and at the end of the program to document changes. The evaluation tool designed to record participants' actual behavior change is based on the Transtheoretical Model.

1.4.1.4a The Transtheoretical Model of Behavior Change

The Transtheoretical Model of human behavior change was developed by Dr. James Prochaska and Dr. Carlo C. DiClemente at the University of Rhode Island. This model has been used extensively to promote healthy behavioral changes in areas such as smoking, diet, substance abuse, and eating disorders. According to the Transtheoretical model, people go through a series of changes before they adopt the desired behavior. Prochaska

and DiClemente describe five stages to change in their Transtheoretical Model and called those five stages *precontemplation, contemplation, preparation, action, and maintenance*.

Precontemplation stage – An individual in this stage has no interest in changing his or her behavior in the near future. Those who are in this stage are not aware of the need for a change toward the desired direction. For example, someone who doesn't eat fruits and vegetables will not show any interest in eating them at this stage. The main characteristic of the individuals in this stage is their ignorance about the need for a change. This situation can be captured by using this statement: "*I'm not considering this behavior or practice.*"

Contemplation stage – Those who are in this stage are aware of the problem and are considering whether there is a need to correct the problem. At this stage, individuals compare the benefits of change versus maintaining the current behavior. For example, a person who is not eating fruits and vegetables will tend to assess the pros and cons of eating them. If we ask a person in this stage, he or she might say, "*I'm considering this dietary habit.*" If someone is convinced that the change is more beneficial than maintaining the current behavior, he or she will move to the next stage of change.

Preparation stage – Individuals in this stage have been convinced about the need for change and are looking for information and alternatives to correct the problem. These people are in the teachable stage for facilitating the change. For example, a person who is not eating fruits and vegetables will begin eating them occasionally and say, "*I'm doing this sometimes.*"

Action stage – Those who are in this stage take actions to change their behavior or practice toward the desired direction. They actively improve their behavior or practice toward that direction. For example, a person in this stage who was not eating fruits and vegetables will eat them and say, "*I'm doing this most of the time.*"

Maintenance stage – A person in this stage accepts the change is important and continues the change as a desirable behavior or practice. When a person moves into this stage, the change has been fully integrated into his or her life. For example, someone who previously didn't eat fruits and vegetables regularly will continue eating them and say, "*I'm doing this all of the time.*"

The multi-session Extension program evaluation tool uses these five stages to document the actual behavior change of program participants. The following tool is used to record the participants' stages of change related to desired behavior or practice.

Example for Multi-Session Program Evaluation – Recording Behavior/Practice Change

For each of the following practices, please circle the number that best describes your current behavior.

Practices	I am not considering this	I am considering this	I am doing this sometimes	I am doing this most of the time	I am doing this all of the time
1. Eat the recommended servings from the five food groups daily.	1	2	3	4	5
2. Eat 2 1/2 cups or more of vegetables per day.	1	2	3	4	5
3. Eat at least 2 cups of fruit per day.	1	2	3	4	5
4. Eating dried beans or peas.	1	2	3	4	5
5 Eat fruit for dessert and snacks more often than cookies, cakes, pies, ice cream or other high fat, high sugar foods.	1	2	3	4	5

The Transtheoretical Model of change helps you understand and facilitate the behavior change process. At the beginning of the training, if you record the stage of the participants' behavior, you will be able to understand the composition of the group in terms of their behavior change stages. If most of the participants are close to column one, the instructor has a reasonably challenging task in changing participants' behavior. If the majority is in the middle, it is relatively easy to facilitate the desired change.
By recording participants' stages of change on this scale at the beginning, in the middle, and at the end of a multi-session program, you can compare the behavior changes at three different stages of the program and document the impact of the program in terms of participants' actual behavior change. The assessment data at the beginning and in the middle of the program can be used to determine the change stage of the majority of participants and select the appropriate instructional strategies to facilitate the change.

1.4.1.4b Example for Multi-Session Program Evaluation – Recording Social, Economic, and Environmental Changes

If the multi-session program is long enough to create social, economic, and/or environmental changes such as health gains, cost reductions, money savings, and/or environmental condition improvements, the following evaluation tool can be used to document the impact of your program. This is based on recording the participant's socio, economic, or environmental situation at the beginning and end of the program. Examples:

1. Nutrition and exercise education program to prevent obesity – In this program, if you record participants' body weight at the first session and at the end of the last session, you will be able to document participants' weight losses and report the number of participants who were able to reduce their excess body weight. This is an example for the social impacts of a family and consumer sciences Extension program.

2. Youth entrepreneurship and financial education – In this program, if you record participants' savings at the beginning and end of a program, you will be able to document how many youths opened saving accounts and the amount of money they saved. This is an example for the economic impacts of a 4-H financial education program.
3. New agronomic recommendation education – If your program educates farmers about reduced soybean plant population (Farmers tend to use more than the required amount of seed to plant their fields.), you need to record the farmers' current seed rate at the beginning of the program, and compare it to the new seed rate they used at the end of the season to assess the amount of seed saved and their value. This is an example for the economic impacts of an agricultural Extension program.

The following evaluation tool format is useful to record end results of multi-session programs.

Examples for Recording End-Results

Please indicate your financial position based on where you currently are in the program.

Recording Condition	**At the beginning of the program**	**At the end of the program**	**Change**
Example 1: Obesity prevention Body weight	205 pounds	195 pounds	10 pounds
Example 2 : Youth entrepreneurship Saving amount	$ 0	$ 150.00	$150.00
Example 3: New agronomic recommendation Value of seed saved	pounds/acre	pounds/acre	17pounds/acre

---

1.4. 2 Step II. Development of sections to collect program improvement data

This section includes program delivery process evaluation questions and some selected demographic questions. It is important to limit the demographic questions to the required minimum to manage evaluation time wisely. For example, if the program has to reach a targeted ethnic group or needs to justify reaching every ethnic group, you must collect the ethnicity of participants. If you plan to select the best delivery method or educational materials for the participants with different academic levels, you need to collect data related to educational levels. If the program is targeted at a low-income audience, you may need to collect participants' household income levels. Generally, Extension needs to collect ethnicity and gender-related demographic data to ensure that our educational programs are not discriminating against any segment of society and meet the civil rights expectations.

Generally, the program delivery process evaluation section evaluates clients' satisfaction with the instructors, educational materials, and the overall program. The following is a sample format.

Example:

Satisfaction

Please circle the appropriate number for your level of response.

How satisfied are you with:	Not Satisfied	Somewhat Satisfied	Satisfied	Very Satisfied
The relevance of information to your needs?	1	2	3	4
Presentation quality of instructor(s)?	1	2	3	4
Subject matter knowledge of instructor(s)?	1	2	3	4
Training facilities?	1	2	3	4
The overall quality of the training workshop?	1	2	3	4

Did the training workshop meet your expectation? ___Yes ____No

Would you recommend this training workshop to others? ___Yes ____No
If not,
why:_____

What did you like most about this training workshop?

What did you like least about this training workshop?

How could this training be further improved?

1.4.3 Step III. Development of sections to facilitate the data collection process

Instructions about data collection should be included to complete the evaluation tool. This information includes the following:

a. The program title and your county name at the top of the tool.
b. Identification number or name of the participant when pre- and post-evaluations are conducted to match those evaluations.
c. Date of the program.
d. Instructions to complete the survey.
e. Return address if it is a follow-up evaluation.
f. Explain the purpose of the evaluation to help participants understand why we ask them to complete it.
g. Thank you note to appreciate their input and time.

1.5. Conducting Follow-Up Evaluations

Follow-up evaluations collect intermediate and ultimate impact data. Intermediate impact data include participants' behavior/practice changes. Ultimate impact or end results include socio, economic, and environmental improvements. Follow-up evaluations are rarely conducted due to lack of adequate resources and low return rates. However, it is useful to identify such challenges and find alternatives to document intermediate and ultimate impacts of Extension programs. The following is a brief summary of typical obstacles and their alternatives.

1.5.1 Insufficient Resources

Follow-up evaluations require more time and money than evaluations conducted during the program. If you plan to conduct a follow-up, be sure to allocate adequate resources to complete the evaluation. Because follow-ups document the intermediate and ultimate impact of Extension programs, it is important to make sure that the program is comprehensive enough to anticipate long-term impact. Since the follow-up evaluation demands more time and money, review the output and immediate outcome of the program before deciding to conduct the follow-up evaluation. First, review the potential of the program to generate intermediate and ultimate impact. Second, review the availability of resources to conduct the follow-up. If the evidence indicates that the program has a great potential to create intermediate and long-term impact, then it is appropriate to plan and conduct a follow-up evaluation.

1.5.2 Difficult to Locate Participants Due to Relocation

Follow-up evaluations are conducted with participants after sometime completing the program. Some of the participants might have relocated, especially in low-income audiences such as migrant farm workers. The longer the time gap to carry out follow-up after the program, the greater the difficulty of locating participants and getting a good response rate. It is therefore more practical to conduct a three- or six-month follow-up rather than a one-year follow-up.

1.5.3 Collecting Data from a Large Number of Participants

Follow-up evaluations demand much time and effort to collect data. If the number of individuals who completed your Extension program is fairly large, you don't have time to collect data from every participant. When your program has a fairly large number of participants, select a representative sample to collect follow-up data. The following steps are tips to help you select a representative sample from your program participants.

1.5.3.1 How to Select a Representative Sample

The main purpose of sampling is to study a part of a population and make conclusions about the entire population based on the sample data. The validity of these conclusions depends mostly upon how well the sample represents the overall population.

1.5.3.1a Identify the Extension program for which you plan to conduct a follow-up

evaluation. Compile a list of all program participants, making sure to include everyone who completed the program. This is referred to as the *target population* for your follow-up evaluation. You may not have contact information for all participants.

1.5.3.1b Compile a list of participants who have provided their contact information. This list of participants is called the *accessible population*. When you select a sample from the accessible population and make conclusions for the target population, there is an error factor, because everyone on your participant list may not be accessible. If the listing difference between the target population and the accessible population is almost equally distributed among all the segments of the participants, this error is not significant. However, if the listing difference is systematic, such as many of the low-literate participants are not accessible, then a generalization of your evaluation results based on a sample drawn from the accessible population will not be accurate for the target population.

1.5.3.1c Determine what size sample you need to draw based on the number of participants in your population. The following table will help you determine the appropriate sample size for your target population.

Table 1

Table for Determining Sample Size from a Given Population

Size of Population	Sample Size	Size of Population	Sample Size	Size of Population	Sample Size
10	10	220	140	1,200	291
15	14	230	144	1,300	297
20	19	240	148	1,400	302
25	24	250	152	1,500	306
30	28	260	155	1,600	310
35	32	270	159	1,700	313
40	36	280	162	1,800	317
45	40	290	165	1,900	320
50	44	300	169	2,000	322
55	48	320	175	2,200	327
60	52	340	181	2,400	331
65	56	360	186	2,600	335
70	59	380	191	2,800	338
75	63	400	196	3,000	341
80	66	420	201	3,500	346
85	70	440	205	4,000	351
90	73	460	210	4,500	354
95	76	480	214	5,000	357
100	80	500	217	6,000	361
110	86	550	226	7,000	364
120	92	600	234	8,000	367
130	97	650	242	9,000	368
140	103	700	248	10,000	370
150	108	750	254	15,000	375
160	113	800	260	20,000	377
170	118	850	265	30,000	379
180	123	900	269	40,000	380
190	127	950	274	50,000	381
200	132	1,000	278	75,000	382
210	136	1,100	285	1,000,000	384

Source: Krejcie, & Morghan, (1970, p. 608)

1.5.3.1d Draw your sample randomly from the accessible population. You may use any of the following sampling techniques, which are appropriate for Extension evaluation due to their practicality.

i. Systematic Sampling:

First, you need to determine your population size and the required sample size. For example, if your population size is 400, you need to draw a random sample of 196 from this population. Next, determine the sampling interval to apply for selecting participants from the accessible population list. The sample size of 196 is roughly equal to 200. Therefore, the sampling interval would be = 400/200 = 2. Take the accessible population list and select the first individual randomly from the first five listings, and then select every

second individual thereafter until you get 196 individuals. If the first selecting number is 3, then the subsequent selections would be 5, 7, 9, 11, likewise till you reach 400.

ii. Stratified Random Sampling:

When your target population consists of identifiable subgroups such as different racial groups or different income groups, use this technique to ensure that every segment is included in your random sample. In stratified sampling, you first identify the strata or segments of interest, and then randomly draw a specified number of individuals from each segment of your population. If you want to get a cross-section of program participants, draw your samples from each stratum or subgroup proportionate to the number of individuals in each group. For example, suppose you need to draw a sample of 196 from a population of 400 participants. Of the population of participants, 40 were Hispanics, 160 were African Americans, and 200 were Whites. You may use the following approach to select a proportionate random sample. Your sample size, 196, is roughly equal to 200. Next, calculate the percentage distribution of each ethnic group in your population.
Percentage of Whites in your population =200X100/400 = 50%
Number of Whites randomly drawn for a sample of 200 would be = 50X200/100 = 100 Whites
Percentage of African Americans in your population =160X100/400 = 40%
Number of African Americans randomly drawn for a sample of 200 would be = 40X200/100 = 80 African Americans
Percentage of Hispanics in your population =40X100/400 = 10%
Number of Hispanics randomly drawn for a sample of 200 would be = 10X200/100 = 20 Hispanics

1.5.4 Low Rate of Response

A low rate of response to follow-up evaluations is a common problem. Normally, a considerable number of participants do not respond to follow-up evaluations. The following tips help maximize the response rate.

Keep the evaluation tool short, clear, and easy to complete – Research shows that as follow-up surveys become longer and more complex, their return rate decreases.

Limit the number of open-ended questions to a minimum – Open-ended questions usually demand more time to respond compared to close-ended questions. Sometimes responding to open-ended questions can become a challenging task, especially if the audience is a low literacy group. The use of more open-ended questions will lead to a lower response rate.

Conduct follow-up before participants forget about the program – Research shows that the longer the time lapse between the program and the follow-up, the lower the return rate. It is better to conduct three-month follow-ups for less comprehensive Extension training workshops and six-month follow-ups for comprehensive programs. Make sure that you allow adequate time for participants to apply what they learned before you conduct the follow-up evaluation.

Contacting non-respondents with reminders – Telephone or mail reminders help increase the response rate.

Personalize the request letter – The follow-up survey should include a personalized letter to the participant with a clear explanation of the purpose and the contact information of the Extension agent to help participants understand the importance and purpose of the evaluation. It is always advisable to include a stamped, addressed envelope to return the evaluation. If the target audience has access to the internet, the use of an online survey provider is a cost and time-saving strategy. As an employee of NCSU, you have access to the *Qualtrics* online survey development program. You may use it to conduct the follow-up evaluation. Here is the link to the NCSU Qualtrics access site: https://isa.ncsu.edu/surveys/qualtrics/accessing-qualtrics-at-nc-state/

Essential Parts of a Follow-up Extension Evaluation Tool:

Recording Behavior/Practice Improvement
Please check the appropriate box for your response.

As a result of this training, have you made any improvement with regard to the following behaviors or practices?	**Yes**	**No**
Eating recommended servings from five food groups daily?		
Applying integrated pest management practices?		
Following food safety guidelines?		

Have you made any other change as a result of what you learned from this training? ____Yes
_____No

If Yes, please list the most significant change you have made after participating in this program:_____

_____

Recording end results/ultimate outcomes

As a result of the training, were you able to achieve your goal(s)?
(Examples: reducing or preventing cost, increasing income, saving money, buying a home, starting a new business, following a healthy lifestyle, serving the community, gaining and applying life skills, and conserving the environment) _____Yes
_____No

If yes, please identify the type of goals you achieved. Check all that apply to your situation.

_____Economic goal(s) such as cost savings or preventions, increased earnings, farm profitability, home buying

_____Social goals such as achieving a healthy lifestyle, reducing the risk of chronic diseases, preventing food-borne illnesses, development of youth, children, and community

_____Environmental expectations such as safety and sustainability of food supply system, conserving soil and water, improving the quality of indoor air

What was the single most important goal you achieved as a result of this training?_____

As a result of this training were you able to save money *(example: reduce the cost of production or living)* or earn money *(example: Increased production, value addition, or new enterprise)*? ____Yes

_____No

_____Don't know

If yes, roughly how much in total? About $_____

If this program helped you achieve your purposes, would you please share your success story with us:

_____

_____

_____

_____

Recording diffusion of information to other users:

Have you shared what you learned from this training with others? ____ Yes ____No
If yes, how many people did you share this information with? _____

---

2. SYSTEMATIC OBSERVATION AS A TECHNIQUE FOR DATA COLLECTION

Systematic observation is a practical method for collecting evaluation data, especially from multi-session programs. In a multi-session program, you have an opportunity to observe the behavior or practice changes taking place in your target audience. If you record your observations systematically before and after the program, you will be able to document the changes taking place during your program. This technique can be used to document participants' observable behavior/practice changes and end results. Observation is appropriate for educational programs presented to young audiences. The following steps are helpful for planning observations as an evaluation strategy.

2.1 Determine the Impact Indicators You Can Observe

You can observe only indicators that you can see. You can visually observe participants' behaviors, practices, and skills. You can't observe participants' knowledge and attitudes. Sometimes, you can observe social, economic, and environmental changes as well.

Therefore, use observation to record skills, practices, behaviors, and end results.

2.2 Develop a Recording Scale for Observations

In order to document the change of participants' skills, behavior, practices, or end results, you need to have data recorded at two points in time. Systematic recording of observations is possible only if you have a recording scale developed for your program. The following instrument can be modified to record participants' skills and behavior/practices.

Sample Instrument for Observations

Recording Skills

Please use the following descriptions of the levels of demonstrated ability of participants when you record their respective skills

1. I doubt he/she could do it: Someone in this category will not demonstrate any ability to carry out this activity.
2. He/She might do it but it would be very hard: One in this category will demonstrate very slight ability with great effort to carry out this activity.
3. He/She could do it but it would be a little hard: Someone in this category will demonstrate ability with some effort to carry out this activity.
4. He/She could do it easily: Someone in this category will demonstrate a fairly high level of ability to carry out this activity without difficulty.
5. He/She could do it very easily: Someone in this category will demonstrate a very high level of ability to carry out this activity.

Please check(X) the column which most describes the observations you made about this participant's ability to carry out each of the following activities.

How do you feel about this person's ability to:	1 I doubt he/she could do it	2 He/She might do it but it would be very hard	3 He/She could do it but it would be little hard	4 He/She could do it easily	5 He/She could do it very easily
1. Help others?					
2. Lead others?					
3. Work with a diverse group?					
4. Volunteer for community service?					
5. Plan health meals?					
6. Calibrate a sprayer?					
7.					
8					
9.					
10.					

Recording Behavior/Practices

Please use the following descriptions of the levels of participants' behavior when you record their respective behaviors/practices.

1. He/She is not doing this: One in this stage will not do or show any interest in this behavior/practice.
2. Looks like he/she is considering about this: One in this stage will show some interest about this but he/she will not do it.
3. Looks like he/she is trying this sometimes: One in this stage will show interest and try this behavior/practice sometimes.
4. Looks like he/she is doing this most of the time: One in this stage will do this behavior/practice most of the time.
5. Looks like he/she is doing this regularly: One in this stage will do this behavior/practice regularly.

For each of the following practices, please check the column that best describes this person's current behavior based on your observation.

Practices	He/She is not doing this	Looks like he/she is considering about this	Looks like he/she is trying this sometimes	Looks like he/she is doing this most of the time	Looks like he/she is doing this regularly
1. Looking for opportunities to help others.					
2. Leading others.					
3. Working with people who are different from me.					
4. Volunteering for community service.					
5. Planning healthy meals					
6.					
7.					
8					
9.					
10.					

2.3 Recording End Results

First, we need to identify the end results before planning to record them. For example, if you are introducing a new farming technology, you will be able to observe whether the farmer has practiced the technology and make any economic or environmental improvement. If it is a weight reduction Extension program, you will be able to observe participants' excess body weight reduction as a social impact. The following instrument can be modified for recording end results of your Extension program. Guidelines for selecting indicators for the end results of your Extension program:

1. Imagine all possible end results of your Extension program.
   Example: Obesity reduction program will have reduced body weight, Body Mass Index (BMI), blood sugar levels, cholesterol levels, blood pressure.
2. Identify a few obvious and easily observable indicators.
   Example: Bodyweight reduction can be easily and accurately recorded.
3. Determine how to record the levels of identified indicators.
   Example: Record participant's body weight in pounds

4. Record the values of the levels at two different stages (One at the beginning and one at the end of implementing learned behaviors/practices)
   Example: Participant's body weight at the beginning (If it is 240 pounds) and end of the program (If it is 230 pounds)

Examples for FCS Programs

Indicators:	Before Program	After Program	Change
Body weight			
BMI			
Number of vegetable servings per day			
Number of fruit servings per day			
Weekly food expenditure ($/week)			
Number of asthma attacks/month			
Cost of utility bill ($/month)			
Number of food safety violations			
Number of child abuse incidences			
Number of people who opened saving accounts and amount saved ($)			
Medical bills ($/three months)			

Examples for Agriculture and Natural Resources Programs

Indicators:	Before Program	After Program	Change
Seed rate per acre (Bushels/Acre)			
Amount of herbicides used (Pounds/Acre)			
Amount of insecticides used (Pounds/Acre)			
Yield (Unit of production/Acre)			
Water quality (Amount of nitrogen in water)			
Extent of land under conservation tillage (Acres)			
Value of farm loss prevention ($/Unit of operation)			
Cost of production ($/Acre)			
Gross income ($/Acre)			

Examples for 4-H Programs

Indicators:	Before Program	After Program	Change
Days of community volunteer service (Days of community service/month)			
Body Mass Index (If the focus is reducing childhood obesity)			
Amount of money in saving accounts after completing a financial education program.			
Total number of achievement awards per school year			
Number of presentations made to groups			

2.4 Train Yourself for Recording Observations

It is important to understand the evaluation criteria for accurately recording observations. Train yourself or any volunteer who records observations.

2.5 Collecting Data

You need to collect observational data at two points in time to evaluate the change. Record first observations before your Extension intervention. When you record observations, you must remain unbiased. When you make the second observation, do not look at your first observation results to avoid any influence on your rating process. Another option is asking someone else familiar with the program such as program assistants or volunteers to observe, and compare those ratings with your observation ratings to ascertain the objectivity of your rating.

3. TIPS FOR COLLECTING EVALUATION DATA

Collecting evaluation data is a challenge because many people do not like to participate in evaluations—even people who enjoy educational programs. The following tips will help you to overcome this challenge.

3.1 Build your Evaluation Capacity

Preparing yourself as an evaluator is essential. Understanding evaluation concepts and learning how to apply them to Extension programs is essential for conducting meaningful evaluations. Building evaluation capacity also helps you respond to participants' questions about the evaluation instruments they are being asked to complete.

3.2 Plan Your Evaluation in Advance

Early planning helps you develop necessary evaluation tools and allocate adequate time for collecting data. You should allocate the necessary time and resources at the planning stage—well in advance of program delivery, when you are considering program content, duration, and outcomes—to ensure the collection of evaluation data. If you plan to conduct pre- and post-evaluations, you need to develop tools for both. To avoid confusion in data collection and analysis processes, it is better to print pre-evaluation and post-evaluation on different colored papers. To match pre- and post-evaluations of each participant, include an identification number on surveys. Reasonably adequate time should be allocated to complete evaluation surveys. If the literacy level of the target audience is low, it is important to design very simple tools and allow extra time to complete the evaluation surveys.

3.3 Provide Clear Instructions to Complete the Survey

Clear instructions help avoid confusion and minimize missing data. Participants should receive the necessary instructions as part of the survey. If the audience is a low literacy group, it may be helpful to read or explain instructions before collecting data.

3.4 Motivate Participants to Complete the Evaluation

Normally, participants are hesitant to be evaluated—especially adults. Sometimes, this apprehension is because they think they are being evaluated as individuals. The best way to address this concern is to explain that the purpose of Extension program assessment is not to evaluate individual participants, but to evaluate the educator, program, materials, and outcome. By keeping participants informed about the purpose of the evaluation, it is easier to receive their cooperation in data collection.

This communication is significant in pre- and post-evaluations and follow-ups. If you do not convince participants about the purpose of evaluation, they may not include identification numbers to match pre- and post-evaluations or agree to participate in follow-up evaluations. It is also important to explain how you maintain the confidentiality of data. Identification numbers (could be random numbers that the participant will know but which will not be meaningful to the educator) help ensure the confidentiality of participants' identities and responses. Explain that identification numbers are used only to match pre- and post-evaluation surveys and not to identify individual participants. Sometimes, low-income audiences might think that if they reveal some information, such as family income, they will lose their eligibility for public support programs. Therefore, you must explain the purpose of the evaluation and assure the confidentiality of data and information to their satisfaction. Awarding door prizes for those who complete evaluations is a good strategy for motivating participants; however, the disadvantage of this strategy is that it requires additional resources.

References

Ary, D., Jacobs, L. C., & Razavieh, A. (1996). *Introduction to research in education* (5th ed.). Orlando, Florida: Harcourt Brace College Publishers.

Cook, T.D. and Campbell, D.T. (1979). Quasi-Experimentation: Design and Analysis for Field Settings. Rand McNally, Chicago, Illinois.

Krejcie, R. V., & Morghan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement 30,* 607-610.

Jayaratne, K. S. U., Lyons, A. C., & Palmer, L. (2007). *Financial education evaluation manual.* Englewood, Colorado: National Endowment for Financial Education.

Prochaska, J. O. and DiClemente, C. C. (1994). *The Transtheoritical Approach: Crossing Traditional Boundaries of Therapy*. Malabar, Florida: Kerieger Publishing Company.