Evaluating Environmental Education and Outreach Programs

Workshop Materials Developed for the Washington State
Department of Ecology Coordinated Prevention Grant Recipients

May, 2003

APPLIED RESEARCH NORTHWEST

WASHINGTON STATE
DEPARTMENT OF
ECOLOGY

Pamela M.M. Jull, Ph.D.
Applied Research Northwest

Table of Contents

INTRODUCTION

In Winter 2003 the Washington State Department of Ecology (DOE) contracted with Applied Research Northwest (ARN) to conduct a set of trainings for Coordinated Prevention Grant (CPG) recipients who are conducting environmental education and outreach programs.[1] The purpose of the workshops was two-fold:

1. CPG grant officers and administrators were responding to the Joint Legislative Audit and Review Committee's (JLARC) recommendations that CPG require evaluation of funded programs.
2. Grant recipients needed to be prepared to implement some evaluative measures for the upcoming grant cycle.

The specific JLARC recommendations are intended to lead to the following outcomes:
- An increase in systematic collection and sharing of information
- A shift from a focus only on program processes (outputs) to include program outcomes (results).
- The streamlining and integration of program services
- Ensuring that funding agencies work together

-JLARC, 2001[2]

The purpose of adding evaluative measures to current and often on-going programs is to improve the program managers' understanding of the programs' outcomes themselves and *not* to convert the grant funding into a competitive process. Because CPG monies require mandatory distribution, a competitive application is inappropriate to use.

The value of evaluating programs is to better understand which program components are most effective and to respond to evaluative information with program changes that may improve the effectiveness of the programs by providing better education.

The following pages outline the information presented during the four-hour training sessions. It is by no means a comprehensive guide to program evaluation and assessment; however, it provides a useful overview and some tools for implementing and assessing current programs. In addition, it provides references to supplementary guides and texts that are less specific to Environmental Education program assessment, yet more detailed with regard to research and assessment methods.

---

[1] CPG is part of Washington's Solid Waste Financial Assistance Program.
[2] Investing in the Environment: Environmental Grant and Loan Programs Performance Audit Report 01-1. January 2, 2001. http://jlarc.leg.wa.gov/Reports/01-1_EnvironmentalGrantsandLoans.pdf

Why do environmental education?

Education is often conceived of as an opportunity to provide information to people. The information is intended to increase people's knowledge and understanding, much like adding information to a database. However, when real learning takes place, information is not simply added to a compendium, it is integrated into people's conceptualization of their world. At it's best, education is a transforming process that challenges people's assumptions and preconceptions, provides new evidence to consider and results in a stronger understanding of our ideas about reality. Education requires people to rearrange their current frameworks, to rethink and test their assumptions with new information.

People involved in environmental education programming are fighting a cultural battle against the social norms of convenience that produce mounds of packaging and other waste-producing behaviors that burden our social and political systems. But without education, cultural change can't happen, and the burden is unlikely to be lifted.

What is assessment for?

Assessment is not accountability. It is not a system of grading – grades are for external audiences. Assessment is more akin to comments on a drafted term paper. Consider your program to be a draft in need of revision – anyone can always do better with a little constructive feedback. Assessment is to provide constructive feedback on a program that in turn will provide information on how to better the program.

What do we learn from comments on draft term papers? All feedback is geared toward how to improve the paper, though the elements they address can vary to include style, ideas, grammar, and organization. In program assessment and evaluation, we intend to learn:
1. What the most effective components of our programs are,
2. What components are weak,
3. And how our programs affect people – opportunity to study humans and really solve problems in changing behavior.

Generally, programs that are put together by thoughtful people often show in the assessment that there is no need for dramatic changes – no requirement to return to the drawing board. Instead, assessments enable program administrators to recognize small changes that can result in better assessment over time. In all cases, the goal is to make a bigger difference in what we know and what happens in our communities.

The results of education and outreach programs are often difficult to assess. The process of learning is an internal one, where the learner's thoughts and ideas are often hidden from the educator. Without explicit program elements designed to elicit responses from learners or enable observation from others, the outcomes of any education effort can only be inferred based on the materials presented.

For example, in a program designed to educate school children about recycling, the curriculum may be expected to:

1. Produce greater understanding of the process of recycling a particular material (e.g. paper or metal)
2. Give kids better identification skills – knowing what types of materials can be recycled
3. Produce a broader understanding of the implications of waste for their community and the environment.

However, knowing how many children take in the information provided, how well they have mastered the material and whether or not the information leads to a change in their behavior cannot be determined without developing specific processes that elicit that information in a systematic and unbiased way from the students.

Structure of this Document's Contents

CPG's solid waste program grants represent a wide variety of programs, as well as a wide range of resources and processes already established for evaluating programs. These materials use examples of four different types of education programs and discuss evaluation techniques that are useful in each setting. The program types are:

1. Environmental Workshops and classroom activities/presentations
2. Community-wide outreach/information campaigns
3. Information hotlines and special materials collection events
4. Special events (i.e., county fair, Earth Day, etc.)

With this documentation, we hope to provide clear information for people who have never had to implement an evaluation of a program, and some new ideas and challenges for people who have. A big part of doing assessment is figuring out where to start. We'll do so by showing you steps to:

- Focus on how you would use the information, including starting with information you already have and making it useful.
- Structure your assessment so evaluation is easy by embedding evaluative components in your program rather than trying to do assessment as a separate effort in addition to your program.
- Identify the kinds of information you want to get and which methods you can use to get it.

Throughout this documentation are several opportunities for group work to discuss your programs and reflect on the information presented. By the time you finish reading and working through the worksheets provided, you should have a plan for how to start assessing your program.

The documents here include some tools for conducting evaluation research including sample databases, survey items and guidelines, scripts for quick phone surveys and sample forms for eliciting feedback from your program participants. It is *not* a comprehensive workbook and does not include step-by-step instructions for the different methods presented. Rather, it provides an overview and some important considerations to help your organization get started in thinking about how to implement an evaluation.

In time, you should be able to develop a detailed plan for modifying your programs and producing objective information to describe how well the program is working. By engaging in

assessment on a regular basis, you'll be sure that your limited resources are being spent in the most effective manner – that you can make the money go as far as possible in producing the kinds of outcomes your communities need.

Section 1: Thinking about the evaluation

Any program assessment must start with some careful thought about how the program is expected to function. Assessment is not an end point, but a part of a process that influences design and implementation, as shown in the diagram below. Program design and implementation are followed by examination of evaluative measures – perhaps information that is collected in due course alongside the program such as counts of participants or workshop evaluation forms, or it is information that is gathered in a special effort such as a follow up of randomly selected participants or a community survey.

The Assessment cycle
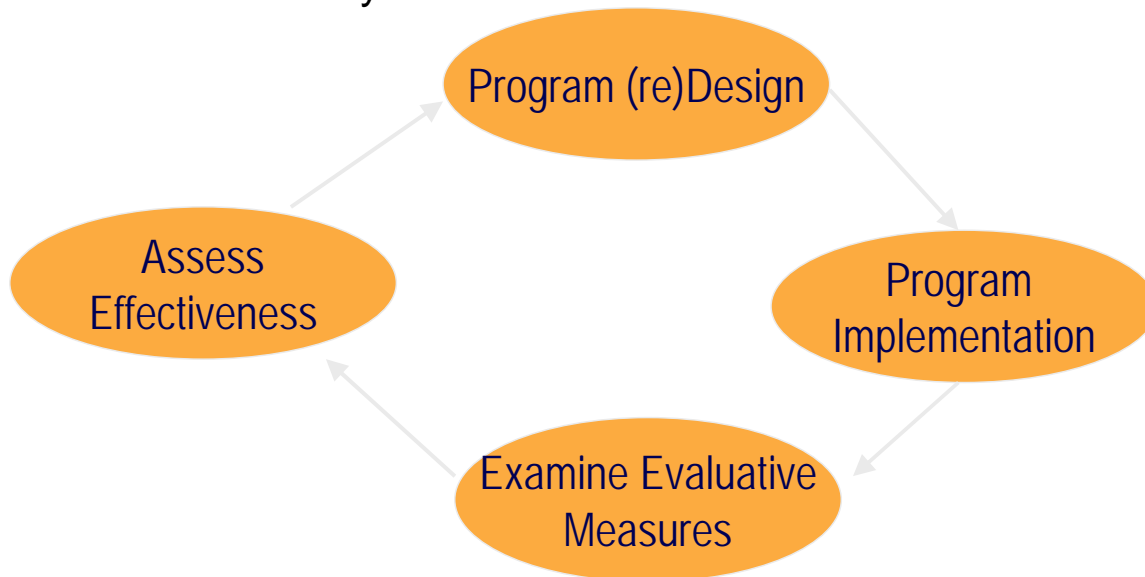
Once the measures are examined, program managers can make some determinations about the effectiveness of their work – identifying the strengths and weaknesses of the program. The strengths can be preserved, while the cause of the weaknesses can be discussed and new tactics introduced that modify the program and may improve the outcomes of the program. The program is then re-implemented, evaluated, assessed and more design enhancements can follow.

Thinking Evaluation: What kind of change to expect

Education programs are designed to teach, but different teaching methods produce different kinds of results.

Traditional Teaching

In Traditional Teaching, a teacher presents information in the form of a lecture, reading materials, or advertising information, sometimes presented with props and demonstrations, but with little interaction or input from the learner. As a result, traditional methods of teaching primarily provide people with new information in a relatively passive setting. This method relies heavily upon a good match between the level of the material presented and the ability of the learner to understand the material. It also depends on the learner to be amply engaged to be able to record or recall the information at a later date when it is needed.

Active Learning

Active learning methods include the learner in discussion, problem solving and interaction with the teacher – much like what is expected in a seminar classroom setting. Active learning techniques are known to produce better retention of information since there are more opportunities for the learner to check in with the teacher and check their understanding of what is being discussed. It also engages the learner more directly with problem solving and puzzling through appropriate examples relevant to the topic, which also enhances memory.

Hands-on Skill building

Educational settings like these generally take the form of workshops or project-based learning (e.g. creating recycled paper in a classroom, composting workshops, etc.). Skill-building settings include all the elements of an active learning setting, but also require the learner to practice a method or technique that enables them to understand the process being taught. The practice of the method is believed to be transferable to other settings. By practicing, the learner is expected to be more willing and able to try new methods, for example by composting at home or choosing appropriate materials for recycling. Education that involves hands-on skill building helps people retain information and learn new behaviors.

Behavioral Psychology Speaks to Environmental Education

Other program elements can also help in encouraging behavior change when embedded in any of the educational settings described above. Published research in psychology has been reviewed and applied to Environmental Education programs with an added focus on assessing the outcomes. Work by researchers such as Doug McKenzie-Mohr, put the varied findings of behavior modification together with assessment and program re-design strategies and call it Community Based Social Marketing (CBSM).

Social Marketing principles succinctly identify elements that may promote behavior change in environmental programs. While there is no room for a complete overview of these principles in this text, the information is very well presented in *Fostering Sustainable Behavior: An Introduction to Community Based Social Marketing* (McKenzie-Mohr & Smith,1999).

Nevertheless, for those who are familiar with CBSM and using the principles in your programs, you will want to include them in your consideration of your program as you review what you expect the outcomes to be. For those who don't use CBSM, the research can be very helpful in redesigning programs that you find are weak in some areas or ineffective in producing behavior change. In particular, CBSM principles suggest that the use of new skills or information is

conditioned on four factors: information, motivation, barriers, and norms, most of which need to be present to a degree in order for the behavior to take place.

Information
The target program audience requires sufficient information to enact the behavior. For example, what can be recycled and what cannot? Where can they take toxic materials? Clarity of the information provided is key if the target is expected to understand it.

Motivation
The target must be willing and interested in engaging in the desired behavior. They see its value and are interested in responding to the need. Often information can produce motivation, but if it cannot, norm building can help (see Norms, below) as well as commitment strategies (e.g. asking participants to take a pledge, promise to participate, etc.).

Barriers
There must be limited infringements on the ability of the target to respond to their motivation – that is, do they have enough time, money, information, skills, transportation and can they remember to execute the behavior when needed? For example, if they have a reusable shopping bag, do they remember to take it into the store? If they are interested in purchasing products with less packaging, can they find the products? If such products are more expensive, can they afford them?

Norms
To some extent, the target needs to have a sense that the behavior is accepted and practiced by others. Often, the more people demonstrating the behavior (e.g. curbside recycling) the more likely others are to take it for granted that they should also engage in the behavior. Norm building can often be used to motivate people who otherwise have limited barriers and plenty of information.

How much change do you expect?
Because different teaching methods have different strengths, some programs may have a more powerful effect than others on participants' knowledge and/or behavior. Retention of information may last a longer time with more skill building, or some information may stick with people because of the effectiveness of the method used to communicate it.

In thinking about what kinds of changes to expect from people who encounter your program, you should also consider how much change to expect. It is useful to simply use three categories for considering the level of change your program produces for knowledge or behavior – low, moderate and high.

Sample Programs to consider

Before working on your own program, take a few minutes to practice thinking about evaluation with these examples.

The Recycling Roadshow

The Recycling Roadshow is a program implemented in Kensington England, where the overall recycling rate was 9%. The program sent workers to canvass areas while wearing noticeable shirts and hats. They distributed pamphlets, stickers and badges to people who they met, going door to door and talking to people in the street. In addition, the program workers gave specific information to people with questions about recycling. They managed to visit with about 20% of households in the target area.

Given this program's design, what kinds of change would you expect to occur in the area? How would households that were visited by the Recycling Roadshow differ from those that were not visited? List the changes you think the program might produce here. Consider both learning of information and changes in behavior:

_____

_____

_____

_____

For each change you identified, consider whether you think the program would produce a small, moderate or large effect for the households visited. Write in an (s) for small, (m) for moderate or (l) for large.

_____

In Concert with the Environment

In Concert with the Environment is a school-based program in which students take home a workbook that they go over with their parents. The workbook helps students and parents determine how resources like power and water are used in the home. Once completed, the parents and kids commit to one or more conservation measures.

Given this program's design, what kinds of change would you expect to occur in the area? How would school children that completed the workbook differ from those that did not? List the changes you think the program might produce here. Consider both learning of information and changes in behavior:

For each change you identified, consider whether you think the program would produce a small, moderate or large effect for the school children participating. Write in an (s) for small, (m) for moderate or (l) for large.

What will you do if you complete your assessment and find no effects?

The first rule is, Don't Panic! More than likely you have a solid program; it just needs some adjustment to elicit the kinds of change you want to produce. Knowing the program's weaknesses enables you to create a stronger program. The following items make a very nice system for examining your current program if your program shows no effects or weaker effects than you hoped for. Be sure you make a clear list of the strengths and weaknesses of the current design. Be sure not to change the strengths!

If information gain is your goal, you will want to:
❑ Consider the teaching methods you use. If they are more passive, consider how to add a component that better engages your target learners.
❑ Consider what methods you employ to encourage recall of the information. Clear messages, catchy phrases, and interesting illustrations are key to helping people learn the information you want to get across. Sometimes a simple sticker reminding people of the message and strategically placed (i.e. on the kitchen trash can lid), can go a long way to helping people remember what they have learned.

If behavior change is your goal, you'll want to:
❑ Consider what your priorities are, what result you want – whether you want to change a lot of people's common behaviors or a few people's highly destructive behaviors (e.g. dumping oil or paint - very hazardous, but relatively infrequent - versus using organic lawn fertilizer – more frequent, commonly practiced, somewhat less toxic).
❑ Consider what the most malleable issue would be (what will be easy for those folks to change? What's hard to change?).
❑ Identify the strategy you think best addresses that issue (information, motivation, barrier removal, norm building, clarifying the message).

Like the experimental method, it is helpful to limit adjustments to only a few aspects of the program. If you make too many changes to your current program, you won't be able to pin down which element was the most effective, and you may produce more work than the program has resources to handle.

Program Work

Before completing this section, please answer the following questions for the CPG program that you are planning to assess.

Your goals will be:

1. To choose a program on which to focus.
   - ❑ Is there a program for which you have not yet constructed evaluative measures?
   - ❑ Is there a program you think may be in particular need of evaluation and redesign?
   - ❑ Is there a program for which it is particularly difficult to assess its effectiveness?

Any of these criteria are useful in choosing the program or set of programs you would like to focus on.

2. Review the questions on this worksheet carefully and develop answers to the questions.

This worksheet is intended to help key program personnel assess the current status of their CPG and other environmental education programs. By working through these questions, people involved in implementing the programs as well as any current evaluation will be able to bring to the training a clear view of all the elements that are currently in place.

If there is more than one program you want to review, just work through the questions on this worksheet separately for each program. If you work with others, a 1-1/2 to 2 hour meeting should be sufficient to identify the answers to all the elements. If you work alone, be sure to take the time you need to think through your answers to the questions and document them.

Program Information

Please respond to the following questions about the program. Some of the items may not seem appropriate or relevant for your program. If so, please skip those items after giving them careful consideration.

Program Overview

1. What is the name of the program?_____

2. What is its current budget?_____

3a. What is/are the primary goal(s) of the program?

_____

_____

_____

3b. Most programs have several goals, or a single goal that is met through several different implementation methods (i.e. a recycling hotline, classroom presentations, media campaigns).

   How complex do you think this program is in terms of its goals? (e.g. Very complex? Moderately? Not very complex?)

   _____

4a. How complex is this program in terms of its implementation? (e.g. Very complex? Moderately? Not very complex?)

   _____

4b. Of all the elements that go into the program's implementation, which are the most demanding in terms of time, complexity or resources?

   _____

5. Which program elements do you think people are most responsive to? That is, are there some that are very popular, memorable, or well known?

_____

_____

6. Which program elements do you think have the greatest impact in terms of…

   a. Learning something important about environmental issues?

   _____

b. Changing people's behavior with regard to environmental issues?

_____

7. Overall, how would you describe the impact this program has in your community?

_____

_____

_____

8. On what kind of information do you base your assessment in question #7 (i.e. stories people tell you, group meetings/discussions, observations of change in behaviors like higher recycling rates, survey information you collect, etc.)?

_____

_____

_____

_____

Evaluation Information

1. One of the simplest ways to enable evaluation research to happen is by keeping lists or databases with information about the people who came into contact with the program. For example, when people call your hotline, do you take down any information about them or their call? When you visit schools, do you keep a listing of what classes were visited, the dates and what program materials were presented?

    Please list the types of information, if any, that you currently have with regard to the program and, if not self-evident, how that information is collected. For each information source, also describe how it is recorded (i.e. on paper, in a database, computer document etc.)

    _____

    _____

    _____

2. Are there any questions you have about your program's effectiveness that you would like to try to answer with some evaluation research? If so, please list them.

    _____

    _____

    _____

    _____

    _____

    _____

3. Finally, doing evaluation research can be quite straightforward once a few pieces are put into place, but usually, evaluation requires some specific resources. What resources do you currently have as far as:
    a. Computers [number and age of machines, amount of free hard drive space]

    _____

    b. Software [especially spreadsheet applications and database management software]

    _____

    c. Volunteers (do you have any dedicated staff time for organizing your volunteers?)

    _____

d. Funds to pay for additional materials – especially forms for participants to fill out

_____

e. Staff time devoted to the program (or is it catch-as-catch can?)

_____

f. Telephone lines – how many do you have?

_____

g. After hours access to phones (specifically, 5-9 p.m.)

_____

h. Staff with knowledge of statistics, excel and/or SPSS, Access and other database programs

_____

4. Please add any comments that you think are important in considering the information you have discussed above.

_____

_____

_____

Exercise 1: Thinking About Evaluation

Considering the above documented information, and your understanding of how educational setting can influence learning and behavior. Make a list of the change(s) you would expect people to experience as a result of encountering your program. It may be helpful at this phase to think of how people who had encountered your program might be different from people who had not. If you have a team that implements your program, each member should write down their thoughts independently, before meeting as a group to discuss them. These most likely align with your program's goals and objectives.

If you have a particularly large or complex program, you may want to select one element of the program to focus on. To help you decide, you may want to refer to these items from the Program Overview above to help you get started:

5. Which program elements do you think people are most responsive to? That is, are there some that are very popular, memorable, or well known?

6. Which program elements do you think have the greatest impact in terms of…

      a. Learning something important about environmental issues?

      b. Changing people's behavior with regard to environmental issues?

Information (What should they know after encountering your program?)

Behavior (What should they be more/less likely to do after encountering your program?)

For each of the program effects you listed above, which would you expect to be large effects? Moderate? Small?

Section 2: Getting Information

Now that you have an idea of what you want to do with your program, you'll need to consider how you're going to find out if you've accomplished it. This section describes a variety of methods for gathering information about your program. You will want to choose the method or combine methods to best meet the needs and design of your program, as well as the resources of your office or organization.

The programs that CPG funds tend to fall into four categories:
1. Environmental Workshops and classroom activities/presentations
2. Community-wide outreach/information campaigns
3. Information hotlines and special materials collection events
4. Special events (i.e. county fair, Earth Day, etc.)

To assess these programs, the following tools may be useful:
- ❑ Focus groups/Community meetings
- ❑ Participant Surveys or Evaluation forms
- ❑ Community Surveys
- ❑ Latent data collection

Focus groups, Community meetings

Focus groups – formal discussion groups, and community forums (or public meetings) are excellent tools for beginning the assessment of programs that have little or no prior assessment completed.

Focus groups require you to recruit a group of people from your target learning population. They will participate in a structured, 1-2 hour discussion about issues related to your program. Generally the groups consist of 6-8 members, a size that allows everyone to participate, but is not so small that few ideas are generated. The goal of focus groups is to allow open discussion of sensitive issues, to hear the kind of language people use to talk about the issues relating to your program, and to learn how they think about and practice the behaviors that are relevant to your program.

It is not necessary to have extensive training to moderate a focus group as long as the moderator understands the basic principles and holds a neutral position on the issues. Staffing requires just a moderator and a note-taker/observer.   There are several excellent books to refer to if you decide to do a focus group to assess your program. Please see the appendix for some suggestions. These and similar sources are readily available at most bookstores.

A community forum (public meeting) is also useful. It allows all interested parties to come voice their opinions and provide input on the topics at hand. As in the focus group, a good deal of information can be presented for participants to respond to. By recording the meetings and reviewing what was said, (cassette rather than video is often helpful for maintaining people's sense of confidentiality and encouraging candor) a synopsis of the meeting can be reported

easily. Transcription can be useful if the budget allows, so that specific, representative quotes can be used in the reports.

Generally, meetings and focus groups require some incentives – refreshments should always be provided, public recognition is sometimes useful – and care should be taken to hold the meeting at a convenient time and place. It is often helpful to choose neutral territory for the meeting as well, i.e. a public school, library, senior center, etc., rather than a county or city office building.

Strengths of Focus groups and Community forums

These two research methods are excellent for exploring people's ideas and obtaining new information. The quality of the information is rich and varied – entirely valid and reliable as long as the moderator is unbiased in the way information is solicited.   They are relatively easy to implement – the largest investment is the time and resources required to recruit participants. Both settings readily stimulate new ideas for programs.

Limitations and Challenges in using Focus groups and Community forums

Two key weaknesses may steer you away from using these methods for assessing your program. First, it is difficult to get people to attend such meetings. Focus group participants are notorious for committing to attending and then failing to show. Community forums sometimes draw few or no attendees. Sufficient publicity and hints of controversy can improve attendance. However, a recent trend for many organizations is to seek audiences with existing groups during regular meetings, for example visiting with neighborhood associations, PTA, garden clubs, and other community associations.

The second weakness of such methods is that the information is never[3] representative of the target population. People with the strongest opinions tend to be the most likely to attend such meetings. If a majority of participants are either critical or highly supportive, it is very important that you do not infer that most of the community agrees with them. They most certainly represent the critics or fans in the community, but the preponderance of attitudes cannot be determined by the opinions of those who show up at a forum or focus group. The value of the discussion is that you learn about a variety of perspectives, not that you learn how most people feel or think on the issues.

Participant Surveys or Evaluation Forms

If your program is primarily a workshop, hotline or school-based program in which the names and contact information for the people participating are relatively easy to obtain, you may want to consider a survey or evaluation form as the tool to use in your assessment. A survey is a structured set of questions, usually with a fixed set of response categories for people to use. Evaluation forms are just a special type of survey, and are often given to participants at the close of a project. You can provide surveys or forms to participants immediately after participating, or you can contact them at a reasonable time after participating to check in on how much they remember and how they feel about the program after having some time away to reflect on what they did or learned.

---

[3] Never, never ever - this weakness cannot be emphasized enough.

Timing can have some interesting effects. People are often most critical of services they receive immediately after their experience. With time, their negative memories tend to fade, and the benefits take on more weight. However, with programs that attempt to teach new skills the reverse may be true. People may feel quite capable upon leaving a workshop only to get home and find they cannot get past some unanticipated problem. In these cases, it can be useful to wait for the program evaluation since the program administrators can use the evaluation to learn the more common problems people encounter after the workshop. When the evaluation happens over the phone or face to face, they also have the chance to provide the past participants with the information they need to execute their home- or business-based project. In addition, program administrators can use what they learn to augment their curriculum and address commonly encountered post-program problems up front, thereby improving their program.

Surveys are good for confirming whether certain beliefs or behaviors are common among your target population. When executed properly, they yield highly systematic and representative information about your targets. The information can be recorded in a simple database using spreadsheets or more sophisticated programs, and analysis can be as extensive and on going as the instrument allows.

However, surveys and forms are rarely a good place to explore what people's beliefs or behaviors are. For example, how do people think about pesticide application? Or how they choose a particular product? Unless you have some very good ideas about how people might respond to questions of pesticide use or product selection, you will not be able to get good feedback in a survey. If you have important issues to explore, do not leave them to a survey.

Strengths of Participant Surveys and Evaluation Forms

Participant surveys are particularly easy to implement since they can consist of brief, self-administered, point-of-service feedback forms. If you use the telephone to follow-up with folks, it can be quite easy to get people to cooperate since they are already invested in your program. Often they are appreciative of the conscientiousness that such efforts reflect. Surveys like these are easy to revise and can provide feedback on items that are highly relevant to the program. Little additional effort is required to do this kind of evaluation, if planning is taken care of in advance.

Limitations and Challenges of using Participant Surveys and Evaluation Forms

Surveys are not a very useful tool for getting new ideas for your program, and determining what directions you could take it. However, running survey results past a group of community members can be fruitful in obtaining innovative information.

Good surveys are difficult to construct – developing questions that are succinct, appropriate, easy to understand and easy to answer while still getting at the information you want can take a fair amount of time. Ideally, the questions are developed, revised, tested, and revised again. Writing good survey questions takes a great deal of practice and a high degree of familiarity with both the program and with the constituency to create reliable, valid questions. Simple changes to question format and wording can have big implications for your findings – what you can and cannot find out in an analysis.

If you have a creative analyst who understands a lot of the nuances of the program and its target population, you can learn a great deal from survey responses. But without a good data-manager and analyst, interpreting the information you get from these forms can be time consuming and challenging.

Keeping track of who has participated and who has been contacted for surveys, coding the forms and creating databases with the information are a challenge in most evaluative work. Finally, it is important to note that the information you get from participant surveys do not reflect your community in general. Most programs draw particular types of community members.

Community Surveys

Surveys of your community are an excellent way to find out if your program is recognized community-wide and to estimate how frequently particular beliefs or behaviors occur. They are especially useful if your program consists of a community-wide information campaign, or if you want to know how much of an impact your program has had on the community overall.

As with all types of surveys, they are not appropriate tools to use if you want to explore people's ideas or identify new directions for your program, and writing good questions, data management and analysis are challenging. Organizations typically use one of three methods to conduct community surveys: Mail, Telephone, and Intercept surveys.

Mail surveys

Mail surveys were once very popular and inexpensive compared to telephone surveys. But they have fallen out of favor as response rates have declined resulting in more cumbersome administration.

Ideally, mail surveys include a mailing of the survey with a cover letter, a reminder postcard 7-10 days later, a second mailing of the survey to non-responding households, and a final reminder. In order to keep from doing multiple mailings to all the households, a tracking system is required to identify which households returned surveys and which did not. In the days when 40-50% of households could be counted on to respond to the initial survey mailing, reminders and follow up mailings were not so costly to do. But with the growth in junk mail and market surveys, most mail surveys now garner only 10-20% of initial survey recipients responding, meaning that as much as 90% of the initial households require reminders. Reminders often add another 5-10% to the response rate, but then as much as 70-80% of the initial households require a follow up mailing.

Mail surveys are time consuming in terms of calendar time as well. Aside from survey development, printing and envelope stuffing, by the time all the mailings have gone out and surveys are returned, as much as 6-8 weeks may have passed. Once the surveys are returned, additional resources are required to record the information received in a database. Once that is completed, analysis can ensue. I recommend a minimum of 3 months to initiate and execute a mail survey.

A final draw back of mail surveys is that the data quality is sometimes compromised. Poorly formatted documents can lead to skipped questions, limited literacy can inhibit people's ability to respond to questions, people may skip items they don't understand, or provide written answers in margins rather than responding using the answer categories you provided. All of these issues lead to difficulties in deciding how to interpret the findings and record the responses.

Telephone surveys

Telephone surveys are somewhat more efficient than mail surveys. In a telephone survey, an interviewer calls a household and recruits a respondent. If a computer-aided-telephone-interviewing (CATI) lab is available, the interviewer can record the responses as the interview progresses, eliminating a separate data entry step. Interviewers can also clarify confusion, identify problems in the survey based on their interaction with respondents, and provide additional information if needed. Unfortunately such services can be quite expensive, ranging from $5,000-$30,000 depending on the number of people interviewed, the length of the survey and the amount of analysis required.

Calendar time on a telephone survey depends on the number of people who are calling and the number of households you hope to reach. It's best to contract with a professional firm, as telephone survey management can be technically complex. Many market research and polling firms will do the work quickly, but with very low response rates (20% is quite common). If you contract with such a firm, you'll want to specify a minimum 50% response rate. It will be more costly, but the findings will be trustworthy.

Intercept Surveys

A final method that is known for its ease of implementation is the Intercept survey – surveys that are conducted by approaching people at an event or commonly attended location (such as a grocery store or mall). Intercept surveys can be conducted over a very brief period of time, allowing for a quick turnaround. They have many of the benefits of the telephone interview (interviewers can clarify confusion and provide needed information) and they tend to have high response rates.

Intercept surveys – face-to-face structured interviews - are particularly easy to implement. The key to getting valid information using intercept surveys is to be aware of how respondents are recruited to answer the questions. Conducting an intercept survey can be intimidating, and it is important to have outgoing, committed and trained staff or volunteers in order to execute it properly. The tendency of your staff will be to only approach people who appear to be interested or who they feel comfortable approaching – often people of the same age, sex and social background as themselves. If staff interviews only those who show an interest or those they are comfortable approaching, the results of your survey will surely be biased. Your goal is to solicit responses from a representative group of individuals at the event. In that way, the findings will also be representative.

Since many social characteristics are not apparent, the only mechanism that will ensure representativeness is randomization. Random is not the same as haphazard; it is a systematic way of being certain there is no pattern in the recruitment. Some of the methods used include flipping a coin to decide whether or not to approach a person, approaching every 5th or 10th

person (or any other arbitrarily chosen number). If your program hosts a booth at a public event, it is important to move away from the booth and solicit responses from people who do not come your way. In that way, you will not be introducing bias by interviewing only those who show an interest in your program.

Representativeness also depends on a high response rate. That is, it is important that a majority of those you approach actually agree to answer your questions. As long as the interviewers understand the importance of getting a high response rate, are forthright, expecting cooperation or willing to ask people to participate again even if they initially refuse (for example, by telling them how important the survey is, or convincing them it will not take much of their time, or offering them a token gift in return for their time and effort). Typically, response rates are not a problem in intercept surveys. People are not often willing to refuse an earnest interviewer and will acquiesce if the interviewer is moderately persistent. To assess how representative your intercept survey data are, it is helpful for interviewers to keep a tally of how many people they approach and how many complete surveys. Interviewers can also be spurred on with promises of rewards for particularly high productivity.

Latent Data collection

Latent data collection means that you set up a system for collecting information that is relevant to your program, but that does not require a significant additional effort. For example: reporting forms from schools on recycling volumes, county waste records, or workshop registration forms that ask for some program-related information. The goal with latent data collection is to find points of reporting and embellish upon them, making them more useful in assessing your program.

Often latent data come from objective measures rather than self-reporting of opinions or behavior. In that sense, the data may be considered more reliable. Yet, the real quality of the data depends upon how consistently it is reported, and how accurate it is. For example, if people use different methods for estimating volume of waste, then the variation in the reports may be too large to detect any change as a result of the program.

In addition, data available through latent methods sometimes does not speak to key goals of the program, may not provide insights for new directions for the program, or may be complex to set up.

A Final Consideration: Comparison Groups

Program evaluation can be very powerful if information is provided for people or places that warrant comparison. For example, interviews with people before and after they encounter a program can suggest changes that are a direct result of the program. Comparing people who participate in a program with people who do not; Examining areas that are targeted for an information campaign and those that are not.

Choosing a method

There are a lot of different directions you can choose to go in an evaluation. Before you tackle your own program, try discussing some sample programs to see what issues arise in making a choice over how you will do your evaluation.

Sample Programs to Review

Given each programs' design, what methods do you think would be most useful in conducting an evaluation? What would the benefits or drawbacks be of using the methods? See if you can envision what might happen with each method, then see if any rise to the top as being the most beneficial or easiest to implement.

The Recycling Roadshow

The Recycling Roadshow is a program implemented in Kensington England, where the overall recycling rate was 9%. The program sent workers to canvass areas while wearing noticeable shirts and hats. They distributed pamphlets, stickers and badges to people who they met, going door to door and talking to people in the street. In addition, the program workers gave specific information to people with questions about recycling. They managed to visit with about 20% of households in the target area.

In Concert with the Environment

In Concert with the Environment is a school-based program in which students take home a workbook that they go over with their parents. The workbook helps students and parents determine how resources like power and water are used in the home. Once completed, the parents and kids commit to one or more conservation measures.

The Wastemobile/Toxic Taxi

This program advertised a three-day event where people could bring batteries, paints and other household hazardous waste to a convenient location for free disposal. In addition to collecting the waste, the staff disseminates information materials.

Exercise 2: Choosing a Method

Consider the five methods and the program work you did in Exercise 1. Which method would be most appropriate for your situation? (choose one or more) What would be the strengths of the evaluation? What would be the weaknesses?

Section 3: Sources of Error in Evaluation

The most common problems that arise in evaluating programs are those that are rarely studied by program managers. This section discusses problems that are referred to as "error," meaning characteristics of the research that compromise what can be learned from it: its validity and reliability. Error can come from a variety of sources, and some types of error are more problematic than others.

Error 1: Coverage

Coverage refers to how much your evaluation allows for every potential target member to be included in the study. The U.S. Census is notable for its goal of reaching 100% of all U.S. residents – complete and comprehensive coverage.

In evaluation of solid waste programs, you want to be sure that there are few or no limits to your coverage, within what is reasonable to obtain – that is, that you have information about all your workshop participants, or nearly every person who contacts your hotline, that you had a chance to intercept a reasonable portion of your event's attendees (not just those who came to the opening or closing), that your database includes information from all schools that recycle, not just the ones that you have the most contact with, or that your telephone survey includes people from most of the neighborhoods in your community.

If there are missing participants, your main concern is whether they are systematically different from the included participants in any way. For example, are the schools located in wealthier or poorer neighborhoods? Is it a particular workshop or trainer who obtains the best class lists? Do you tend to get more information about participants when your office is not busy? Systematic differences in coverage are referred to as "bias" in your information. For example, if the schools are in wealthier neighborhoods, then your information is biased toward the wealthy and under-represents schools in less well-off areas.

Limitations on your coverage are reasonable to expect. Since you are unlikely to get everyone, you will want to make a reasonable determination of what portion are included and how that might influence the outcomes of the evaluation – whether the lack of coverage introduces any bias into your study.

If you have severe limitations in your coverage, you have an instantly available avenue for significant program improvement: Is there any way you could improve the documentation of the participants? Which program elements do you think warrant the extra effort of documenting participants? Are there any points of contact that aren't being taken advantage of? By improving your documentation of participants you make it easier to assess the effectiveness of your program when you need to.

Error 2: Sampling

Once your listing of potential respondents is relatively complete (or the limitations are well understood), you will likely want to assess the experiences of only a portion of them. It is perfectly legitimate to do so. Few organizations can afford the resources that enable them to get

feedback from every person they come into contact with. Sampling error refers to problems that arise in the selection of the people who will participate in your evaluation.

Fortunately, sampling error is easy to control. Once you have made a determination about who will be a part of the study (i.e. composting workshop participants in the past 2 years, hotline callers in the past 2 months, school classrooms visited in the prior year), you will want to select those who will be pursued for evaluative information.

The simplest method to use in avoiding sampling error is also extremely reliable – randomization. Simple random sampling means that every case has an equal chance of being selected – it is closely akin to throwing all the names in a hat and drawing them out one at a time. Systematic random sampling is a variation on simple random sampling: If you have a list of participants, or a list of community households (i.e. the phone book) you can simply choose every 5th or 10th person on your lists until you have enough for your evaluation (just as you would to decide who to approach for an intercept survey). This is the same method you would use in an intercept survey, except that instead of choosing from a list, you would randomize by choosing every N^{th} person you see (where "N" is an arbitrary but constant number).

If you have names in a database, you can do simple random sampling by assigning them a random number, then sorting the cases according to the random numbers. Then you can select as many as you need from a section of the list.

In Microsoft Excel, you can create random numbers by entering this formula in a cell: =rand(). The formula will produce a number between zero and one – sometimes with as many as 10 decimal spaces. **Important:** Once you have generated a random number for each case, copy the column of random numbers, and use "paste special" from the "edit" menu. Choose "values" so that the formula is replaced by the actual random number value that was generated. If you do not do this step, each time you sort, the number will be recalculated.

Step 1: Database of names and phone numbers, in alphabetical order

First Name	Last Name	Telephone
Allison	Applegate	756-8596
Burt	Bender	542-6548
Candace	Collins	245-8551
Doug	Dimple	123-4569
Earl	Eggers	321-3322
Frances	Finner	555-6523
Greg	Google	854-5698
Hillary	Highstreet	456-8956

Step 2: Database with Random Numbers Assigned, still in alpha order

First Name	Last Name	Telephone	Random number
Allison	Applegate	756-8596	0.117584915
Burt	Bender	542-6548	0.900168883
Candace	Collins	245-8551	0.76993124
Doug	Dimple	123-4569	0.706037187
Earl	Eggers	321-3322	0.202540481
Frances	Finner	555-6523	0.36139212
Greg	Google	854-5698	0.443342235
Hillary	Highstreet	456-8956	0.275420656

Step 3: Database Sorted by Random Number

First Name	Last Name	Telephone	Random number
Allison	Applegate	756-8596	0.117584915
Earl	Eggers	321-3322	0.202540481
Hillary	Highstreet	456-8956	0.275420656
Frances	Finner	555-6523	0.36139212
Greg	Google	854-5698	0.443342235
Doug	Dimple	123-4569	0.706037187
Candace	Collins	245-8551	0.76993124
Burt	Bender	542-6548	0.900168883

For classroom or workshop-based programs, you can select a random sample of the presentations or participants. If you choose to sample the presentations, it may be more efficient to choose all those done in one month, all classrooms in one school, or all those done by a particular staff person. For any of those considerations you will want to consider if the month or school or staff person might bias the results – that is, that the responses you get from those groups would be significantly different from those that meet at other times of the year, in different schools or taught by different staffers.

Comparison groups can be of help at this stage. Working with staffers or teachers, you can randomize the selection of classrooms within a school to receive the program, and then test the students in all classes on their knowledge or behavior for issues relevant to the program materials. If you are working to improve your program, you can select some workshops to be conducted using your usual methods, and others (at random) to get a revised program. Check the results and see which methods do a better job at producing the knowledge or behavior you want.

Error 3: Margins of Error - How many cases is enough?

One of the first questions people ask when starting a new project is "how many people do I have to get to do my survey?" The answer is not entirely straightforward. It depends on the size of the population (did you teach 2,000 or 20,000 people in the past year?) and how much precision you need in order to measure change.

The number of cases you collect information from will determine the margin of error in your study. By now, you should recognize that error margins are only one source of potential error in your research. They are useful for knowing roughly how accurate your findings are. For

example, if a national study with 1200 survey respondents finds that 43% support legislation to protect the environment, the margin of error would suggest that if everyone in the nation were interviewed, between 40% and 46% (a 3% error margin) would support legislation to protect the environment.

Some simple rules of thumb for large populations (over 5,000 people) are that 300 cases will produce a 6% error margin, which is usually enough for most programs evaluations. If you collect information from 400 people, your margin drops to 5%, and with 800 you reduce the margin to 4%.[4] You can see that it takes substantial extra effort to reduce margins of error.

It is important to note that if you plan to split your respondents into groups – for example, males and females, aged and young, Westside and Eastside, then it is important to note that the error margin will be larger for those subgroups than for the sample overall. For example, if you collect data from 400 cases, and 150 are males, the error margin for their responses will be 8%.

Many of you will be working with smaller populations – fewer than 5,000 total target cases – especially if you do participant follow-up. In these cases, you can drop your sample size to 250.[5] If you only have 1000 cases, 150 is sufficient and at 400-1000 cases, 80-150 is sufficient. Once your population drops below 400, you can produce reasonable estimates with as few as 40-80 cases. The more cases you can collect, the more accurate your estimates will be and the more complex your analysis can be. However, there are diminishing returns to interviewing more people than needed to estimate the effects of your program.

Important note: Since some people will not be willing or able to respond to your evaluation, you have to plan for the non-response as well as the response in preparing to draw your random sample (see Error 4, below).

Error 4: Non-response
Non-response is a very serious problem in evaluation research and one of the most difficult to control. Non-response can refer to the lack of participation in a survey (unit non-response) or the propensity people have to skip questions (item non-response). Unit non-response is the most disconcerting since often we don't know the characteristics of the people who failed to respond. Because we don't know their characteristics, we often cannot determine whether the people who responded are systematically different than those who did not respond. Ultimately, non-response leaves researchers questioning whether the findings are biased or representative of the target population.

It can be advantageous to collect information on participant sex, age, education, and homeownership since much of that information is readily available from the Census on the Internet. If you are doing a community survey and your surveys are returned from people with relatively similar characteristics as those in the community overall, it suggests that your sample may resemble the entire community in other ways. If your sample does not resemble your

---

[4] An easy formula that helps provide estimates of error margins is 1/sqrt(N) where N=number of cases.
[5] These sample sizes are not based on a specific formula as they exceed sample sizes needed for reasonable error estimates. Instead, they are based on the ease with which analysis can be performed in meaningful ways on the information that is produced.

community in some ways, you can identify the bias and qualify your findings. For example, if you find the people who responded to your survey are more likely to be female, older and better educated than the community overall, you know that your findings may under-represent younger, less educated, and male community members.

What's a response rate?

Unit non-response is most efficiently described by reporting the evaluation's response rate. The response rate is the percentage of surveys, forms or interviews that were completed relative to the number in your original sample. For example, if 200 people called your information line in March and all were given the opportunity to evaluate the program, but only 100 of them gave feedback, your response rate would be 50%.

One nuance of computing response rates is to be certain to drop cases that were selected but which did not have current contact information or were otherwise unable to participate. For example, if 20 of the folks who called your information line gave you phone numbers that are now out of order or for residences where they no longer live, your response rate would be 100/(200-20) or 56%. The remaining 44% either were not home when you called or refused to participate in the survey.

An optimal response rate is 65% or better, though 50% or better is sufficient for making program decisions. At 65% you can be fairly certain that the 35% who did not respond are unlikely to completely change your results, or to differ remarkably from the 65% who did respond. Response rates will vary depending on how engaging your survey is and how interesting the topics are to the people you call.

How does response rate affect sample size?

If non-response exists, obtaining a large enough sample to minimize your error margins will require you to select more cases for attempted contact than the number of cases you want to end up with. For a participant survey or evaluation, you will want to double the number of cases you select for contact in order to ensure a 50-65% minimum response rate. For example, if you want to collect information from 200 people, select 400 to participate. Some will have outdated contact information, some will be unavailable to respond, and others will be uninterested in responding. With 400 cases to start with, you have a very good chance of getting complete information from 200.

For a community telephone or mail survey, you will want to triple your sample size since non-response and bad contact information tends to be much higher in these groups. If you send your mail survey first class in an envelope printed with "address correction requested" under the area where the return address is printed, the post office will send you a copy of the address the mail is forwarded to. In addition, by sending it first class, you will get the mail returned if the address is not valid. By spending the extra money to send your mail first class (at least for the first round of the mail survey) you will know how many of the prospective recipients never received your mail. By eliminating those cases from the response rate computation, you can be more certain of how many valid contacts you started with. If your survey is sent using bulk mail, the post office will not notify you if the address is invalid.

Error 5: Measurement

Constructing reliable and valid measures for evaluation is more of an art than a science. In survey research, a good deal of evidence has been compiled to guide some of the decisions survey writers make in choosing their question formats. Because there is a substantial amount of information to share around survey question writing, a separate section is devoted to it in Section 4 below.

Error 6: Timing

Timing errors occur when the evaluation happens too early or late after the program so that it impacts the outcomes of the assessment. Picking your timing carefully and being consistent with each round of evaluation is very important if you want to track your program's improvement over time.

Too soon

If evaluation happens too soon after a program (particularly a workshop or school program) people will feel very competent and efficacious and will likely give the curriculum high marks. These feelings are the result of the information being easy to recall at the time, and knowing what is expected immediately after hearing the new information. If your goal is to assess perceptions of the curriculum, evaluation forms filled out before people exit a workshop or class can be very helpful in getting feedback, whether there were particularly confusing or engaging segments, and whether their expectations were met.

However, this feedback happens before participants have had time for reflection, prior to their integrating the learning into their existing beliefs and practices, and before they have encountered problems or opposition to the new information. If your goal is to measure retention of information or behavior and attitude change it is best to wait at least a few weeks before asking participants to provide information about their assessment of a program.

Too late

Once some time has passed, participants may find they forget a good deal of what was discussed or taught, or they may encounter difficulties in executing the skills they learned in the workshop or classroom setting. If evaluation happens too long after the program, a good deal of new information may interfere with the participant's perception of what they learned or what they changed. The age of your participants should be taken into consideration as adults and children differ in their recall patterns.

Wrong time of year

A final consideration in timing is to consider whether the time of year might make a difference in people's responses. For example, a community survey about yard waste practices conducted in the dead of winter, will probably produce different results than it would at the beginning, middle or end of the growing season when people are more engaged in their yard maintenance.

Exercise 3: Evaluating your data

You may have already gone about collecting some information about your program – contact information, latent data measuring outputs or inputs, or evaluation forms and surveys. To begin, make a list here of any of the information you already have about your program. You may want to refer to your responses to Item 8 in Exercise 1.

_____

_____

_____

_____

_____

_____

The goal of the last two sections was to present what some of the dangers and pitfalls are in collecting information. Now that you have seen that information, you may have some critical assessments of the information you have already collected. Based on what you know so far, evaluate the quality of the information you currently have.

1. How comprehensive is it? Do you have good coverage?

_____

_____

2. How systematic is it? Is there any reason to suspect bias exists? If so, how is your information biased?

_____

_____

_____

_____

3. How accurate do you think it is? Are there sufficient cases?

_____

_____

4. What are the other potential sources of error in your current data, if any? Consider how you might remedy them.

_____

_____

Section 4:Designing Questions

Designing a Questionnaire: The Process

Many people start a research process by trying to make up questions for their forms or surveys. However, the effectiveness of a question can only be assessed if you know what the explicit information is that you're seeking. For some information this may seem obvious – for example if you want to know if people attended your event. But for others, the question will be too complex unless you know what the information goal is, and how you will use it.

The first step in designing questions is to identify the information goals of the project. What is it you want to know and why? Try not to phrase this information as a question someone would answer, but instead as information you could use in the program.   For example:
- ❑ Do people remember our advertising campaign?
- ❑ How many attended our event?
- ❑ What did people learn at our event?

These questions are great for helping to identify information goals but are too broad and general to be used as survey items.

The second step is to consider how you would use the information you would get from the research you do. Do you plan a major overhaul? Can you afford to respond to everyone's comments? Are there some things you absolutely would not do, for example, start a curbside recycling program for toxics? By knowing your limits you can eliminate some questions that otherwise might be interesting but not useful.

Finally, you will want to choose your method (consider if there is a comparison group), construct your questions (be sure each one ties to an information goal), get feedback from coworkers about the questions, revise them, then test them with a few people from the population you would be using. Once you have a feel for how well people can understand and answer your questions, you will be ready to use your form or survey in a broader study.

Designing Questions

Constructing questions is very challenging and takes a lot of practice, but if you think you are ready to give it a try, here are some guidelines to help you make some choices about how you present the questions. Most of these recommendations come from the work of Jon Krosnick and his associates (see his web page at http://www.psy.ohio-state.edu/social/krosnick.htm). Some of his publications are listed in the appendix to this report. Another great source is Norbert Schwarz's work – also listed in the appendix.

Of primary importance is the level of difficulty you introduce by how you ask your questions. You want to consider your needs simultaneously with the needs of the respondent. Respondent burden – that is, the difficulty people may have answering questions, the length of the survey and their level of interest in the project you are doing are all very important. For any question, consider that a respondent needs to:

- ❑ Interpret what you have asked

- ❑ Retrieve the information they need from memory in order to answer the question.
- ❑ Decide what their response is
- ❑ Figure out which answer category (i.e. strongly agree, disagree etc.) their response fits into.

Respondent burden increases when:

1. People have a hard time understanding what you are asking – you have used ambiguous language or jargon, or you have asked for too much information with one question.
2. People cannot remember the experience, behavior or feelings they have around an issue.
3. People cannot decide how to respond to your question based on their experience.
4. Your answer categories do not fit with how they think or feel.

Here are some hints for easing the burden:

1. Ask relatively simple questions. Do not ask people to answer more than one item at a time (a double-barreled question). For example "How useful and interesting was the information you learned at the workshop?" Some people may find it interesting but not useful and vice versa. By putting both descriptors together in one question, you may shorten the survey by a question, but you also make it harder for people to answer! Ultimately, the difficulty will lengthen the time people spend on the question and give you less reliable feedback.

2. Rather than asking yes/no questions, use a range of responses. Often people's response is a "sort of" or "maybe" or "I'm not sure." Asking them to decide if their answer is mostly no or mostly yes can lead them to skip the question or just choose one or the other because they cannot decide. Random choices lead to lousy data and difficulty tracking changes from year to year or program to program.

3. Use an odd number of categories – either 3 or 5 when you ask for a response. For example:

   - ❑ Yes/Maybe/No;
   - ❑ Agree/Neutral/Disagree;
   - ❑ Strongly Agree/Somewhat Agree/Neither Agree nor Disagree/Somewhat disagree/ Strongly disagree;
   - ❑ Excellent/Good/Fair/Poor/Very Poor;
   - ❑ Extremely/Very/Somewhat/A little/Not at all

4. Allow everyone an out – just in case you misunderstood the process or your population, it is helpful to have a "don't know," "not applicable" or a "no opinion" response option. Without these options, people will randomly choose a category or skip the question all together.

5. Be sure to make your language age appropriate. Generally, 8th grade level writing is considered accessible to all adults. You can check the level of the language you are using with a tool in Microsoft Word. Though the best source is a classroom teacher's feedback, by using Word's spell-checker on individual sentences, you can estimate how easily your questions can be read.

1. Highlight the question
2. From the "Tools" menu choose "Spelling and Grammar".
3. Once the spell check is complete click "Ok" and the readability scores will be shown.

See the information on Readability Scores in Word's Help Menu for more information on the scores. Most adults can read easily between an 8th and 10th grade level. For kids, try to hit the grade level for the group you will be addressing. Ideally, run the instrument past a few kids and teachers before finalizing it.

Do not use numbers without meaning. For example, "please rate your satisfaction with the workshop on a scale from one to ten, where one is 'very dissatisfied' and 10 is 'very satisfied.'" Without definitions for the numbers 2 through 9, you will get thoroughly ambiguous responses. People may find it easy to answer, but there is no way to assess the values people gave to the unlabeled parts of the scale.

6. Try not to use numbers to represent categories (i.e. circle the number that best represents your response):

Please rate your satisfaction with the workshop.

Extremely satisfied	Very satisfied	Somewhat satisfied	A little satisfied	Not at all satisfied	Don't Know
5	4	3	2	1	0

Instead, use check boxes or parentheses for the respondent to check. That way people do not infer that there is more value to the "extremely satisfied" response. In this case, the numbers may bias people's responses toward giving more positive answers. Here is the alternative:

Please rate your satisfaction with the workshop.

Extremely satisfied	Very satisfied	Somewhat satisfied	A little satisfied	Not at all satisfied	Don't Know
( )	( )	( )	( )	( )	( )

7. If you are asking about experiences or behavior, be sure to give folks a time frame to consider, i.e. in the past month, the past 12 months, the past week or ever in your lifetime.

8. Try to keep the time required to complete your form less than 10 minutes. If you must, 15 minutes is acceptable, but fewer people will be willing to finish, and they may give less thought about the answers towards the end. A five-minute survey or form is really optimal for ease of response. Unfortunately, most of us are seeking more information than can be answered in five minutes.

9. Be careful to give each answer category the same space on the page. If there are big gaps between some and small ones between others, it appears that the former has more weight. For example:

Please rate your satisfaction with the workshop.

Extremely satisfied	Very satisfied	Somewhat satisfied	A little satisfied	Not at all satisfied	Don't Know
( )	( )	( )	( )	( )	( )

Because the first three categories have more space around them, the eye is more easily drawn to them and respondents are more likely to choose one of those categories as their answer.

10. Use white space to make it easier for people to read your document. Sometimes just small adjustments can help rather than double spacing. In Word's Format menu, choose "Paragraph" and add 2-4 points either before or after the paragraph. By doing so you can change something that looks like this:

Please rate your satisfaction with the workshop.

Extremely satisfied	Very satisfied	Somewhat satisfied	A little satisfied	Not at all satisfied	Don't Know
( )	( )	( )	( )	( )	( )

Into something that is much easier on the eyes, like this:

Please rate your satisfaction with the workshop.

Extremely satisfied	Very satisfied	Somewhat satisfied	A little satisfied	Not at all satisfied	Don't Know
( )	( )	( )	( )	( )	( )

11. Focus on knowledge or behavior rather than opinion and "feeling" questions. People's attitudes can sometimes change with their mood, so asking a lot of questions about things people know or do rather than what they think or feel can yield more reliable information.

12. Ask questions about information you would expect people to know better if they had been exposed to your program. If you have a mixed group – some participants and some not – you can infer that differences between the groups may be due to your program.

13. Allow space for comments at the end. It is always nice for respondents to be able to vent or qualify their responses. Sometimes you will get very useful information about your survey or form as well as the program.

Exercise 4: Designing Questions

Consider the goals, changes and research methods decisions you have made so far in the prior exercises. Take some time to fine tune the goals into a more detailed list of information goals here. For example, if you said "Increasing public awareness of recyclable materials," you might want to detail the type of materials you teach people about, and what you'd expect them to know as a result (that is, what do you mean by "awareness?").

_____

_____

_____

_____

_____

_____

_____

Next, consider the guidelines for asking some questions. Choose a single information goal and construct 2-3 questions that you could ask in order to learn the information. The different questions should work to answer the same information goal item. By practicing different ways of getting the same information, you will have the opportunity to consider how people's responses might differ, and how what you would learn might differ depending on how you asked the question.

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Section 5: Working with Information

The goal of this section is to help you get a start putting your information into useful formats for assessing your programs. Once you have got some completed forms, latent data reports, surveys, or quizzes, this will provide you with some tools of making sense of what you have. It is often helpful to try and set up your system before you make a final decision about how you are going to collect information. Sometimes by setting up the system, you have an opportunity to head off complications that will arise after the information is collected.

Setting up a Database

Expensive, complex database software is not required to make good use of information. All of the examples presented here use only Microsoft Excel, though any spreadsheet application will work.

Why set up a database?

Working with a database can be intimidating, frustrating and time consuming. But ultimately, getting your evaluation materials into a computer database is extremely *efficient*. If you can put the information in a database, you have the opportunity to analyze and reanalyze, answering different program questions as they arise, rather than having to re-review all the information you have.

Setting up your database the first time may take the better part of a day (or several hours over the course of a week), but once you have done it, you will find it easier every time after. If you use the same reports or forms over time, you can simply add to your database over time, enabling you to compare information from different seasons, different years, etc.

Having a database is also extremely *effective* for writing up your program evaluation. Many organizations try to tally responses by hand, or read over the information they have and summarize it. But reports are more powerful if you can include specific counts, averages, differences or charts and graphs to back your discussion. It can also help you avoid being biased by a few negative or positive pieces of information, giving you a clear and realistic understanding of your data.

Steps in setting up a database

This section outlines the steps you will go through to set up your database and begin working with the information. Each step is described briefly here, and then elaborated upon in the next section.

Step 1: Before you get on your computer, take some time to work with the report, form or questionnaire for which you want to develop the database. In order to create an effective database, you will want to use numbers rather than words to represent the information as often as possible. The key reason to do so is that many of the automatic functions that Excel can perform must be used with numerical information. It also keeps your spreadsheet columns narrower, making it easier to view on the computer screen. The process of creating a numeric scheme to

represent the information on your documents is called "coding," and the numbers themselves are the "codes."

Step 2: After developing a coding scheme, you will apply your codes to the documents you have – a process called "hand coding." When you hand code, you simply interpret the information on the form according to the coding scheme you have created, writing down the code for each segment of the form as you go. Although you may be tempted to just work from the form without coding it first, most data entry errors happen as people try to read, interpret and input information in a single step. More accurate data entry is done when each part of the process is divided up into a separate step.

Step 3: The third step is to set up your database in Excel, or whatever database program you will be using. The format I describe here is a matrix where each column represents an item from your document, and each row represents all the information from a single document. In Access it is called a "flat file." In statistics it is referred to as "raw data."   Most organizations jump immediately to summarizing information into a table or a report. By creating a raw data file first, you can have much more flexibility in developing tables, graphs, charts and reports.

Step 4: Enter your data.

Step 5: Explore your data. Generate some statistics, sort the file by a specific column, and look over what you see (some patterns are obvious in a flat file and require no statistics). It is often helpful to generate a list of questions you want to answer before you dig in to the analysis. You can refer back to the information goals of the program to guide your work.

Step 6: Consider the implications of your findings. In the analysis, ask questions about information you would expect people to know better if they had been exposed to your program.

Creating a Coding Scheme

Coding schemes are somewhat arbitrary, but here you will learn some of the conventions used in most statistical research, as well as seeing some examples of how they can be applied. By creating a master document that captures all the codes you want to apply, you have a reference for anyone who helps you and an explicit description for anyone who wants to use the database as well. That document, called a "coding master" is shown below.

The quiz below might be used to follow up on a classroom program. Next to each item is a code assigned to the item in red font.   You will note a few of the conventions here. First, binary items, where something is checked, circled or simply has a yes/no answer are often recorded as a zero (0) if they are not checked or circled or if the response is "no." If they *are* checked, a one (1) is used as the indicator.

For scale responses, larger numbers are used to indicate the positive or high end of the scale, and small numbers are used for the negative or low end of the scale. If someone skipped an item, it is conventional to enter an eight (8) for "no answer." If you have a "don't know" response category, those are usually coded as nine (9), and "not applicable" is often seven (7).

If you are asking for information that might include a value of 7, 8 or 9 (such as how many years the respondent has lived in your county), you can use the conventions of 77, 88 and 99 for not applicable, no answer and don't know responses. Even 777, 888 or 999 can be used, but be careful of how many keystrokes you require for each non-response entry.

<u>Sample Coding Master</u>
Recycling Class Feedback Form

Which of these kinds of paper and cardboard can you recycle at home? Circle your answers.   0 if not circled/ 1 if circled

Envelopes   Food containers
Flat cardboard Magazines
Ice Cream Cartons Newspapers
Colored paper Wrapping Paper

2. At home, how often do you put paper into the recycling instead of throwing it in the trash? Choose just one answer.   8 – no answer

All of the time   5
Most of the time   4
Half of the time   3
Rarely   2
Never   1

Here is a set of mock responses and the coding that would be applied to them.

Recycling Class Feedback Form

Which of these kinds of paper and cardboard can you recycle at home? Circle your answers.

Envelopes   0   Food containers   0

Flat cardboard   Magazines

~~Ice Cream Cartons~~ 0   1   Newspapers

Colored paper   0   Wrapping Paper

2. At home, how often do you put paper into the recycling instead of throwing it in the trash? Choose just one answer.

All of the time
Most of the time
☑ Half of the time   3
Rarely
Never

Use a clean version of your document to create your coding master, and refer to it as you code the information on the documents you have. The coding master becomes invaluable once you have set up your database – it is helpful to create a place in your document to describe what the codes mean in case you lose track of the master.

In addition to the codes, you will want to create names for each of the items that will help you associate the columns of information with the content on the document. For questions you can just use the question number (i.e. Q1, Q2 etc.) or you can use a single word that helps you recognize the content (keep it short!).

Finally, as you code your documents, you will want to assign each document a unique ID number. If you have an ID number on the document and an ID number in the database, then you can correct noticeable data-entry problems quite easily. For example, if you have a cell with nothing entered, or with a number you wouldn't expect (i.e. a "4" where values should only be 0, 1 or 7, 8 or 9), you can look at the ID number in the database, find the document in your file (it helps to keep them sorted by ID), and check what the code should be.

It is sometimes helpful to assign ID numbers according to some scheme. For example if you have data coming monthly from various locations, you can give one location all the 100-level ID numbers, another the 200-level ID numbers and so on. That way, each ID would represent one report from the location (i.e. 101 would be the first report from location X, 201 would be the first report from location Y).

A final consideration in your coding may be to assign a group ID to the document. For example, if you are administering a program at several locations, you can give each location a number (this would be in lieu of using the ID to identify a location). You could assign a group number to a set of kids in one classroom, to a particular program administrator's workshop participants' responses, or to a comparison group – people who did not participate in a program or who got a modified version of the program.

Below is an example of what the coded information might look like in a database. In this example, we decided that we could learn some interesting information by counting how many of the recyclable items in question 1 the students correctly circled and how many were incorrectly circled. For this example, envelopes, colored paper, magazines, newspaper, and cardboard are correct answers, while food containers, ice cream cartons and wrapping paper are not.

A	B	C	D	E	F	H	I	J	K	L	M	N	O
2	1000= Fall 2003	1=did program	# of correct	# of incorrect	1=circled item	Q2 codes 5=always 4=often 3=sometimes 2=rarely 1=never							
3	Case ID	Group ID	Q1 - correct	Q1 - incorrect	Envelope	Colored paper	Food cont.	Maga-zines	News-papers	card-board	Ice Cream	Wrap	Q2
4	1001	1	2	0	0	0	0	1	1	0	0	0	2
5	1002	0	5	3	1	1	1	1	1	1	1	1	5
6	1003	1	3	1	1	0	0	0	1	1	1	0	4
7	1005	1	3	0	0	0	0	1	1	1	0	0	3
8	1006	1	2	0	0	0	0	1	1	0	0	0	4
9	1007	0	5	3	1	1	1	1	1	1	1	1	5
10	1008	1	3	1	1	0	0	0	1	1	1	0	1

One trick to using Excel for data entry is learning how to keep the column headings visible (Case ID, Group ID, Q1 – correct, etc.) as you move down the page. Once you have entered about 25 cases (depending on the size of your computer screen) the headings will scroll off the page, making it more difficult to see the item for which you're supposed to be entering information. If look on the right hand side of the document, at the top of the scroll bar, just above the up arrow, you will see a small horizontal bar. If you click and drag that bar you will see the bar stretch across the whole spreadsheet. By dragging it down, you can drop it just below the row showing the column headings. Notice that now you have two scroll bars – one for the top of the sheet and one for the remainder of the sheet.

At this point, with the data you want showing in the top window, go to "Window" menu and choose "Freeze Panes." The bar will disappear, as will the upper scroll bar, but when you scroll down the page, the top row of cells will stay put. To remove the effect, go back to the Window menu and choose "Unfreeze Panes."

There is a similar vertical bar just to the right of the horizontal scroll bar. You can use that bar to freeze the panes on a set of rows that you want to be able to view as you do your data entry, if needed.

Analyzing your data

Once your data are entered into a spreadsheet, you can do some simple manipulations that will provide you with the summaries you need to understand what it all means and what you can report about your program's efficacy.

Imagine for this program you wanted to examine:
- Whether kids who experienced the classroom program gave more correct answers than those who did not.
- Which materials kids who got the program were most likely to identify correctly and which were incorrectly identified (such information might suggest you change your curriculum to clarify)
- Whether kids who got the program were more likely to recycle than kids who did not.

Counts and Statistics

To answer the first question, we can start by sorting the list according to which group the kids are in. Once sorted, you can use the average function to find out how many items kids in each group got right or wrong. The average function looks like this:

=average(cell range)

Where cell range is just the column and row number for the start of the list of answers and the end of the list, separated by a colon. First, sort the whole data set (it is important not to sort just one column, or you mix up the answers people have given to other questions – to avoid that error, it is helpful to save the data under a new name before you begin making changes. One sort can do irreparable harm!). Then count how many responses were from kids in classrooms that experienced the presentation. The formula is just Excel's "count" formula and it looks like this:

=count(cell range, value)

or more explicitly:

=count(c4:c77, 1)

Which means count all the cells from column C row 4 through column C row 77 that have the value of "1" in them.

Here is what the result looks like:

A	B	C	D
2	1000= Fall 2003	1=did program	# of correct
3	Case ID	Group ID	Q1 - correct
69	1010	0	1
70	1011	0	4
71	1012	0	2
72	1013	0	2
73	1014	0	3
74	1015	0	4
75	1016	0	3
76	1017	0	2
77	1018	0	3
78			
79			
80	# of participants	57	
81	# of nonparticipants	17	

The figure shows that column C, row 80, where the formula was entered, is equal to 57, meaning that 57 people were participants in the program. Below it, in row 81, the formula reads:

=count(c4:c77, 0)

and shows a value of 17, meaning that 17 cases were not participating in the program – about 1 classroom for comparison.

It is helpful to know the counts, but what you may really care about is whether they differed in their responses to the questions. To determine this, you can use the average function in Excel, which is:

=average(cell range)

Because I want to compare the average of the participants to the average of the non-participants, I want to generate two averages – one for each group. Because the data are sorted, you can select the cell range that includes each group for each average. So the following formulas are used:

=average(d4:d60)
=average(d61:d77)

As shown in rows 82 and 83 below, the average of correct responses for the participant group is .4 higher than that of the non-participant group, and the average of incorrect answers is 1.1 lower for the participants. You can also find the median and standard deviation for a set of scores using similar formulas (=median(cell range); =stddev(cell range)).

43

If you have had any basic statistics, you know that differences in averages can be an artifact of one or more extreme scores, so it is always helpful to test the differences. Excel allows you to do so with a t-test formula. The answer it returns is the probability that the differences occurred by chance. Generally, if the answer is less than .05, there are significant differences between the average scores, though there are several details of your research design to consider. Refer to an introductory statistics text for more information.

A	B	C	D	E
2	1000= Fall 2003	1=did program	# of correct	# of incorrect
3	Case ID	Group ID	Q1 - correct	Q1 - incorrect
75	1016	0	3	3
76	1017	0	2	3
77	1018	0	3	2
78				TOTAL
79				
80	# of participants	57		
81	# of nonparticipants	17		
82	Average Participants		3.3	0.9
83	Average Non Participants		2.9	2.0
84	Median Participants		3.0	0.0
85	Median Non Participants		3.0	2.0
86	Standard Deviation P		1.04	1.20
87	Standard Deviation NP		1.14	0.94
88	T-test P vs. NonP		0.20604	0.000588

The t-test formula is:

*=**TTEST**(cell range group 1,cell range group 2,number of tails (usually 2),type of test)*

If you are comparing two separate groups, as in this example, the type of test is an independent samples t-test, which is coded two (2) or three (3) in the Excel formula (see the help menu for more information. If you are checking to see if there was change over time for a single unit of analysis (i.e. schools at time A versus time B) you'll use a one (1) for the test type (paired samples t-test). So the formula for this problem looks like this:

=TTEST(D4:D60,D61:D77,2,2)

The result is .20 – nowhere near the .05 cutoff. The t-test result for the differences in the average incorrect answers is also unambiguous at .0006 – a highly significant result. What does it mean? These results suggest that students may not learn that much more about which of the five items are *correct* to recycle, but the program is very effective in teaching kids what they should *not* put into their recycling bins. It is very likely the non-program kids were guessing at their answers and happened to get most right since five of the eight items were recyclable. But the true test of their knowledge was in how many they misidentified as being recyclable.

Using Percents

Another way to examine this information is by looking at the percentage of kids who identified each item in the quiz as something that could be recycled. To compute the percentages, you can use the formula:

$$=(SUM(cell\ range)/number\ of\ cases)$$

In this case, for the participants, the formula looks like:

$$=(SUM(G4:G60)/57)$$

This formula will return a number between zero and one. If you highlight the cell and click on the percent symbol on the toolbar (%), it will convert the number to a percent.

The figure below shows the percent of participants and non-participants who circled each of the items. The items are color-coded; green represents recyclable items, and red represents non-recyclables for this example. Column F, row 81 shows that All 17 non-participants selected Envelopes compared to just 49% of participants. Since Envelopes are something that can be recycled in this example, it may suggest that the curriculum does not make it clear that envelopes are recyclable, or includes some confusing information that makes participants less likely to choose envelopes as a recyclable category.

A	B	F	H	I	J	K	L	M	N
2	1000= Fall 2003	1=circled item	Q2 codes 5=always 4=often 3=sometimes 2=rarely 1=never						
3	Case ID	Envelope	Colored paper	Food cont.	Maga-zines	News-papers	card-board	Ice Cream	Wrap
75	1016	1	0	1	0	1	1	1	1
76	1017	1	0	1	0	0	1	1	1
77	1018	1	0	1	0	1	1	0	1
78		45	27	25	47	69	51	37	26
79									
80	# of participants	49%	35%	23%	74%	100%	74%	49%	23%
81	# of nonparticipants	100%	41%	71%	29%	71%	53%	53%	76%

A minority of both groups chose colored paper as a recyclable category, which might also suggest a need for more information on that item. Only 23% of participants selected "food containers" as a recyclable item, but 71% of non-participants selected this item. This is clear information about how it is that the program is making a difference. Looking at the other non-recyclable items, it is clear that wrapping paper is another often-misidentified item unless the student has experienced the program information and training. Looking at the recyclable items, newspapers are the most often correctly identified item by both groups. If the program includes a substantial component on newspapers, the program managers may want to consider spending less time on newspapers (which even non-participants seem to recognize as recyclable) and more time on other items.

Distributions

A final method for examining your data is especially useful with categories like those in question 2. Although we have assigned numbers to the categories, remember that the categories are not

implicitly numeric. That is, Always (5) minus Often (4) does not equal Never (1)! Most organizations will report the average score on an item like this, but often they will miss the most important information by stopping there!

The figure below shows that in this case, the average scores between the two groups is quite similar, and the median is exactly the same. Another problem of using statistics inappropriately like this comes with the description of the results. Would you say that the participants averaged "sometimes" while non-participants were between "sometimes" and "rarely?" You can see that the numbers and the meaning of the categories don't go together very easily.

	Participants	Non participants
Average	3.1	2.6
Median	3	3

Rather than using statistics like these, try using the count function again. By counting the number of times participants' responses were coded with each number, you can build a table like this one:

	Participants	Non participants
Always	11	0
Often	12	4
Sometimes	21	8
Rarely	0	0
Never	13	5
Total	57	17

The formulas in the cells are as follows:

	Participants	Non participants
Always	=count(o4:o60,5)	=count(o61:o77,5)
Often	=count(o4:o60,4)	=count(o61:o77,4)
Sometimes	=count(o4:o60,3)	=count(o61:o77,3)
Rarely	=count(o4:o60,2)	=count(o61:o77,2)
Never	=count(o4:o60,1)	=count(o61:o77,1)

Clearly, it is difficult to compare the counts of the participants with those of the non-participants because the groups are such different sizes. However, if you create a table with the percentages, it is easier to examine.
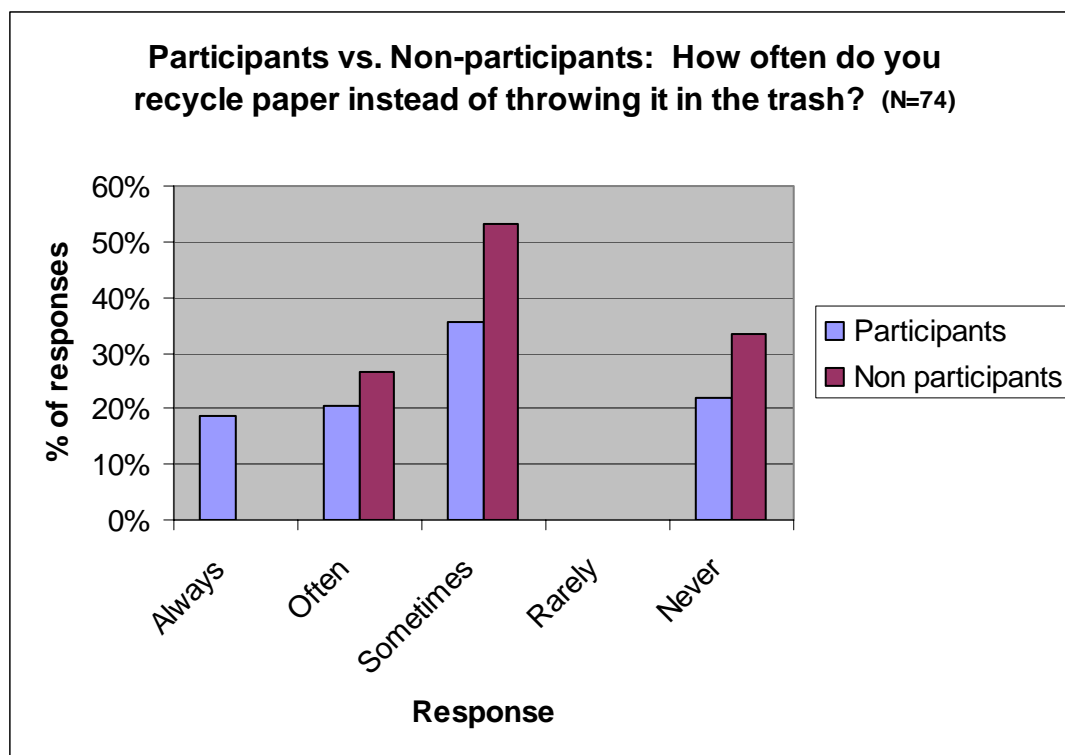
	Participants	Non participants
Always	19%	0%
Often	20%	27%
Sometimes	36%	53%
Rarely	0%	0%
Never	22%	33%

From this information, it becomes clearer where the differences are in the data. In particular, 19% of participants reported "always" recycling while none of the non-participant group did so. Most of the non-participants reported "sometimes" recycling.

Graphics

Although this table is fine for your own use, you may want to create something easier to read for external audiences (county commissioners, community groups, reports, etc). If you have not experimented with Excel's Chart Wizard, it is a very straightforward tool for making a chart or graph of your findings.

Working directly from the cells created above, simply highlight all the relevant cells (row and column headings included), and click on the Chart Wizard icon in the tool bar. When comparing groups side by side on the same measures, a vertical bar chart is useful, as shown below:

Participants vs. Non-participants: How often do you recycle paper instead of throwing it in the trash? (N=74)

A vertical bar chart showing % of responses on the y-axis (0% to 60%) and Response categories on the x-axis (Always, Often, Sometimes, Rarely, Never). Participants: Always 19%, Often 20%, Sometimes 36%, Never 22%. Non participants: Often 27%, Sometimes 53%, Never 33%.

It is fairly easy to see from this graph that only participants said they "always" recycle, while the non-participants' responses exceeded the participants on every other lesser category.

APPENDICES

Appendix: Print Resources

Evaluation:
Methods for Studying Programs and Policies, 2nd Edition, Carol H. Weiss, Prentice Hall. 1998.
Very thorough and accessible. More academic in tone than most.

F*ostering Sustainable Behavior: An Introduction to Community Based Social Marketing*, Doug
McKenzie-Mohr and William Smith, New Society Publishers1999.

Focus Groups
Out of print but a great standard: *Focus Groups : A Practical Guide for Applied Research 2nd
Edition.* Krueger's 3rd Edition of the same book is less comprehensive. This is out of print now
but runs about $15, compared to the $40 3rd Edition. Krueger is a co-author with David Morgan
on a comprehensive six-volume set called The Focus Group Kit ($140).

Billson, Janet Mancini. 2001. *The Power of Focus Groups: A Training Manual for Social and
Policy Research*. Skywood Press, Barrington, RI. ISBN 0-9700075-0-7. Available for order
online at http://www.focusgroupdimensions.com/manual.htm. Price: $75.
A valuable and very accessible overview of focus group uses and methods from recruiting to
writing the report.

Research Design Principles
Babbie, Earl. 2001. *The Practice of Social Research*, ninth edition, Belmont, CA: Wadsworth.
Also see his book with Lucia Benaquisto: *Fundamentals of Social Research*.

Analysis and Statistics
Aron, A., & Aron, E. (2002). *Statistics for the behavioral and social sciences: A brief course
(Second edition).* Upper Saddle River, NJ: Prentice Hall. ISBN 0-13-026186-6. My all time
favorite book – excellent examples and extensive answer solving keys. The only draw back is
their unconventional notation.

Doing Surveys, Research Methods and Writing Questions:
Among the best there is from our own guru at WSU:
How to Conduct Your Own Survey by Priscilla Salant (Author), Don A. Dillman (Author) John
Wiley & Sons; 1 edition (October 27, 1994) $25

Mail and Internet Surveys: The Tailored Design Method by Don A. Dillman. John Wiley &
Sons; 2nd edition (November 1999) **ISBN:** 0471323543

Work from Jon Krosnick and his colleagues, psychologist:
Petty, R. E., & Krosnick, J. A. (Eds.). (1995). Attitude strength: Antecedents and
consequences. Hillsdale, NJ: Erlbaum.

Weisberg, H., Krosnick, J.A., & Bowen, B. (1996). Introduction to survey research, polling,
and data analysis. Thousand Oaks, CA: Sage.

Seymour Sudman, Norman M. Bradburn, Norbert Schwarz (1996). Thinking about answers : the
application of cognitive processes to survey methodology. San Francisco : Jossey-Bass
Publishers, 1996.

Training Assessment Summary

In April of 2003 the Department of Ecology (DOE) contracted with Applied Research Northwest (ARN) to conduct a workshop on evaluating environmental education and outreach programs.

The workshop was designed to answer why we do environmental education and why environmental education assessment is so important. The workshop was structured into 7 main sections:

1) Thinking about the Evaluation
2) Getting Information
3) Examining What You Know
4) Sources of Error in Program Evaluation
5) Designing a Questionnaire
6) Working With the Information
7) Doing a Community-based Survey

Overall, the goal was to enable the participants to walk away from the workshop with a detailed plan for modifying their environmental education and outreach programs and to produce objective information to describe how well the program is working.

At the close of the workshop, the participants completed a feedback form as a tool to measure the successfulness of the workshop. The form was aimed to provide an evaluation of four primary facets.

1) how the participants felt about their time spent at the workshop
2) how worthwhile it was for them to attend
3) how appropriately the workshop content was targeted
4) what they would change about the workshop or do differently

Overall, the participants' feedback was very positive. Most participants reported that they would not change or do anything differently. In addition, the majority of participants felt that both their time on the materials was well spent and it was worthwhile to attend.

Detailed Results

When the data from the feedback forms was compiled and analyzed, it was quite clear that the training was a success.

Question 1 on the feedback form read, "Which sections of the presentation, if any, were confusing or difficult to follow?" As can be seen in Table 1 below, more than half of respondents (51%) reported that *none* of the sections were confusing or difficult to follow. Almost one quarter of respondents (22%) reported o*ther*[6] or a response not listed and18% reported working with information.

---

[6] Other Category Responses: "Had a little trouble at start figuring out how to apply this information to programs - overall okay!" "I don't know if people learned enough to undertake an assessment." "At the start of the workshop, we did not take time for intros. It was confusing on how to break into groups." "Towards the end of the presentation I needed a break. I'd suggest a stretching break in the room to stimulate the brain functions." "I suggest that the spreadsheet data be representative. I think that portion would be

Table 1: SECTION OF PRESENTATION CONFUSING OR DIFFICULT TO FOLLOW*		
None	28	51%
Other	12	22%
Working With Information	10	18%
Getting Information	5	9%
Needed More Time	3	5%
Thinking About Evaluation	2	4%
Group Work	2	4%

*N=53. Sum of percentages is greater than 100 due to multiple responses.

Question 2 on the feedback form read, "Which sections of the presentation, if any, did we spend too much time on?" Over half (58%) of respondents reported that *none* of the sections had too much time spent on them (See Table 2 below).

Table 2: SECTION OF PRESENTATION SPENT TOO MUCH TIME ON*		
None	31	58%
Other	7	13%
Getting Information	6	11%
Working With Information	4	8%
Needed More Time	4	8%
Group Work	2	4%
Thinking About Evaluation	1	2%
Needed More Examples	1	2%

*N=53. Sum of percentages greater than 100 due to multiple responses.

Question 3 read, "Which sections of the presentation, if any, did we spend too little time on?" A large portion (38%) of respondents reported that there was too little time spent on the *working with information* section (See Table 3 below). Nineteen percent (19%) of respondents reported that *none* of the sections had too little time spent on them. An additional, 19% reported that there was too little time spent on the *thinking about evaluation* section.

Table 3: SECTION OF PRESENTATION SPENT TOO LITTLE TIME ON*		
Working With Information	20	38%
None	10	19%
Thinking About Evaluation	10	19%
Other	9	17%
Getting Information	6	11%
Needed More Time	3	6%
Needed More Examples	3	6%

*N=53. Sum of percentages greater than 100 due to multiple responses.

Question 4 was a two-part question. The first part read, "Overall, how useful was the group interaction elements?" Forty-two percent (42%) of respondents reported that the group interaction was *very useful*, nearly half (45%) reported that it was *moderately useful*, and 11% reported that it was *not very useful* (See Table 5 below).

---

less confusing." "Usually people not familiar with evaluation already are hesitant to it. Starting with instrument design and coding at the beginning may create some confusion." "First half of program I was having trouble trying to relate your information to CPG reporting."

Table 4: HOW USEFUL THE GROUP INTERACTION ELEMENT WAS*		
Very Useful	22	42%
Moderately Useful	24	45%
Not very useful	6	11%

*N=53.

The second part read, "Overall, how useful was the information presented?" Nearly three quarters of respondents agreed that the information presented was *very useful* (73%) and the remaining respondents (27%) reported that it was *moderately useful*. None of the workshop participants reported that it was *not very useful* (0%).

Table 5: HOW USEFUL THE INFORMATION PRESENTED WAS*		
Very Useful	38	73%
Moderately Useful	14	27%
Not Very Useful	0	0%

*N=52.

Finally, participants were asked to make any additional comments on the backside of the feedback form. We are happy to report that most participants (67%) who had additional comments reported that the workshop presenter, Dr. Pamela Jull, did a *great job* (see table 6). Forty-two percent (42%) made recommendations[7] on how to improve the training. It is important to note here that the categories used on this form were not mutually exclusive. In other words, some people said that the training was great but also included a way to improve it.

Table 6. ADDITIONAL SUGGESTIONS OR COMMENTS*		
Great Job	16	67%
Made Recommendation	10	42%
Other	4	17%

*N=24. Sum of percentages greater than 100 due to multiple responses.

---

[7] Recommendations: "Longer time (maybe 8:00 to 4:00) with more frequent breaks." "Would have liked to have groups report back or final results." "We need more detailed follow-up on specific work in each county." "Wished there were more specifics about what others with similar programs have done and what has been most successful." "One note on user-friendliness and resource efficiency: Some of the slides on the printout were repetitive, using much more paper than necessary. Could have just printed the whole list just once (drew picture for example)." "Try to allow for more time to prepare for the training." "Some info/slides was missing from packet. Consider not having people from same county sit together - they seemed really interested in learning about what everyone else is doing." "We didn't talk about recycle and waste prevention programs and how we could implement what we learn very much. Would have liked to focus more." "Missing the last page of slides on notes." "Needed - designing a project to accommodate appropriate evaluation."

About the Presenter

Pamela M. Jull Ph.D.
Dr. Jull and her company, Applied Research Northwest have been the evaluators for multiple Watershed Pledge programs funded by the Department of Ecology and the Environmental Protection Agency. All aspects of research and administration fall under Dr. Jull's direction. She is an expert in both qualitative research and quantitative advanced statistical processes and techniques and has successfully managed research for a wide range of organizations and topics. With specialties in research methods, social policy, and statistical analysis, she adds her wealth of experience to all aspects of research, from project structure and development, to focus group moderation, data analysis, and reporting.

She has conducted, coordinated and directed program evaluations and survey research for the Department of Ecology's Watershed Pledge program for Lake Whatcom, Sumas Aquifer and the Padden Creek watersheds. She is also the point person for evaluating the upcoming expansion of the program to include other municipalities in Whatcom County. Her related work on environmental programs includes a membership survey for the Yakama Indian Nation's integrated resource management plan, an evaluation of Olympia's Wellhead Protection Project and a current project assisting Thurston County with an assessment of the Common Sense Gardening program.

She has conducted multiple projects for Washington State schools, libraries, city and state government agencies, including welfare to work programs developed by the Employment Security Department. Her work also includes community non-profits and health organizations. She has been responsible for the production of many polls, community surveys, academic presentations and technical reports, and volunteers her time to contribute to local efforts promoting sustainability and environmental programming. ARN is an active member of the Bellingham Sustainable Connections organization (www.sconnect.org).

Dr. Jull taught sociology and statistics and directs the Office of Survey Research at Western Washington University. She is currently designing educational assessment tools for a variety of departments and special university projects. Her primary role at the university is devising methods for assessing educational outcomes of programs in higher education.