Chapter Two: Training Satisfaction and Reaction Measurement

The assessment of satisfaction with the training process, program content, faculty, and venue is the most common type of training evaluation, with the focus on gathering participants’ reactions to their training experience. As previously discussed, it is important to move beyond mere satisfaction measurement in dependency court training evaluation to include other levels of measurement – skill or learning acquisition, behavior or attitude change, and impact or outcome measurement. That being said, the assessment of training participants’ reactions to, and satisfaction with, dependency court trainings should be included in any evaluation strategy as it provides critical information about the training process and implementation. In order to be useful, training satisfaction and reaction measurement must be tailored to the training’s goals and learning objectives, and collected in a reliable and valid manner.

Chapter Two of this Guide provides recommendations and sample strategies for ensuring that the measurement of training participants’ satisfaction is precise, valid, and provides useful information to training organizers. Some of the topics covered in this Chapter (e.g., the survey process, good questionnaire construction, building valid response categories, increasing response rates, the use of online measures, etc.) also have relevance to the assessment of learning, behavior and attitude change, and outcome measurement covered in other Sections. Much of what is presented in this Chapter on satisfaction and reaction measurement therefore, is foundational to all levels of good training evaluation practice.
Evaluating participants’ satisfaction is an essential component of program evaluation. Measuring trainee satisfaction is critical to ensure the vitality of any training program. Unsatisfactory training programs will not meet the goals set forth by the organizers, the facilitators or the trainees in need of training, calling into question the utility of a given training program. This chapter focuses on the specific reasons why measuring satisfaction is important, and discusses the techniques required to measure satisfaction successfully. Because the primary methods of measuring satisfaction and reaction involve survey, interview and focus group procedures which are important tools for the other levels of measurement, this Section of the Guide should also be referenced when building learning acquisition, behavior change, and impact or outcome evaluation strategies.

Measuring participants’ satisfaction allows training organizers to estimate how effective faculty was at engaging the audience. One way to ensure training has motivated trainees to learn is to measure their satisfaction with the program. If trainees are generally satisfied, it is an indication that they were willing to consider the curriculum being presented to them. If trainees are generally unsatisfied, it is an indication that barriers were present which prevented the trainees from considering the curriculum. To gather people’s true impressions, it is imperative to allow people the opportunity to express their level of satisfaction anonymously. This is typically done by administering short questionnaires or surveys directly after a training program has concluded, or soon thereafter, without asking for respondents to provide their name.

Another reason it is important to measure satisfaction is to illustrate to trainees that their perspectives are important to training program organizers. Asking for feedback about the training demonstrates that you value participants’ perspectives. The opportunity to express an opinion makes people feel as though they will have an influence on future training programs. Measuring satisfaction informs participants that training organizers and faculty want to do a good job and are interested in obtaining the perspective of trainees to ensure that they are doing a good job.

Measuring satisfaction also provides training organizers with quantitative information that can be used to establish standards of performance for a training program. If a training program is ongoing, it is important for organizers to know generally how satisfied participants are with the program. Determining the general satisfaction level allows organizers to determine whether trainings are continuing to meet the needs of trainees. It also allows organizers to determine whether specific faculty or training facilitators are equally as satisfying, or to determine whether specific training methods help or hinder a training program.
Satisfaction measurement may be used to address concerns and criticism voiced by individuals. Typically, people feel a need to express displeasure, an impulse not equally felt when something is pleasurable.\(^{32}\) In a training scenario, there are often a few vocal people who will readily express their concerns with some aspect of a training program, or perhaps their concerns with the entire training program itself. Measures of satisfaction can determine whether these feelings generalize to other trainees. In more instances than not, the majority of trainees do not share the criticism of others. By continuously and comprehensively measuring reactions to the training experience from as many participants as possible, training organizers will have a greater amount of information available to delineate valid from invalid criticisms.

In summary, measuring satisfaction is beneficial to both trainees and training organizers. For trainees, it gives them a voice to express themselves and their perceptions of the training, which includes them in the training process. For training organizers, measuring satisfaction provides general information regarding whether trainees received the training positively or negatively. The more areas of a training program that are measured the more information available to training organizers to decide how to improve on a training program.

*Keep in Mind ...*

While end-of-course participant satisfaction questionnaires provide valuable information on participants’ level of satisfaction with the training (their interest, their attitudes about the usefulness of the content, views on the quality of the training instructors and materials, etc.), satisfaction measurement does not provide valid information on workplace performance or the organizational impact of the training. In fact, most academic studies of satisfaction assessment or “reactionary” measurement have found that there is little correlation between participant satisfaction and learning or workplace performance results.\(^{33}\) In other words, just because your evaluation found high trainee satisfaction with the course does not mean that trainees have actually learned or that they will apply what they have learned on the job, to change behavior or practice. You will need measurement at the other levels of evaluation to answer those questions.

**General Guidelines for Evaluating Satisfaction**

*Determine What You Want to Find Out.* Training organizers need to determine the key areas of the training in which they would like to measure satisfaction. There are several different components of a training that can be analyzed for participant satisfaction, including:

- The curriculum/material in general

---


Chapter Two: Training Satisfaction and Reaction Measurement

- The organization of topics covered in the curriculum
- Whether participants’ expectations for the training were met
- Facilitators/Presenters/Faculty
- Location of the training
- Atmosphere (facilities, meals, room temperature)

Of course, there is value to training organizers in obtaining reactions to every component of the training program, particularly with newer training programs. However, this may create a very cumbersome survey or extremely lengthy feedback form and decrease the likelihood that participants will complete the form or respond thoughtfully. Instead, training organizers should consider a more focused approach to satisfaction measurement for both new and more established training programs – examining the specific components of the training program for which they feel information about satisfaction would be most critical. Or, training organizers can consider a two-phase process: 1) provide a shorter, more focused satisfaction feedback form to all participants; and then 2) follow-up with a sample of training participants and engage them in a focused, more in-depth discussion about their training experience (e.g., via a focus group or interview process, or via a more detailed survey).

Quantify Participants’ Responses. When designing participant reaction or feedback forms, it is helpful to design an instrument that will quantifying participants’ reactions. Quantifying satisfaction responses simplifies data analysis. Turning people’s responses into numbers increases training organizers’ ability (and other people exposed to the data) to understand and discuss what is going on in a training program. Quantifying responses is generally accomplished in surveys using measurement scales. For example, a Likert scale, which is a rating continuum anchored by dichotomous concepts, such as ‘agree’ or ‘disagree,’ is a common method for quantifying responses. The number of intervals on the scale can vary, but research finds that people map their perceptions best onto 5 or 7 point scales34 (See subsequent Sections of this Chapter for more detailed information on the use of measurement scales). With respect to quantification, we are not advocating abandoning open-ended questions (see below). However, we do recommend limiting them, as drawing out qualitative themes during data analysis is extremely time consuming if there are a large amount of respondents. The ideal evaluation form provides the maximum amount of information and requires the minimum amount of time.

34 Schwarz & Sudman (1996).
Encourage Open Responses. While quantification simplifies data collection and analysis, turning responses into numbers does not allow trainees to express any personal perspectives not accounted for on a questionnaire, survey, or interview. Qualitative responses allow participants the opportunity to share their perceptions, allowing participants' the opportunity to bring new ideas to the attention of the training organizers – ideas or concerns that organizers may not have been aware of when creating the evaluation form. Therefore, it is always optimal to include some open-ended questions on basic evaluation forms to encourage participants to share any comments or suggestions. Additional open-ended questions should be included and tailored to any area of a training program where organizers believe it is important in order to gauge trainees' perspectives. In addition, opportunities for respondents to clarify their responses are critical to your understanding of the data. For example, if respondents report that they were not satisfied with their training experience but they are not given an opportunity to tell you why they felt that way, training organizers may not have sufficient data to make appropriate program modifications based on this feedback. Be sure to include opportunities for respondents to comment or further explain their responses, in order to build an explanatory model for your survey findings - the “why’s” behind the satisfaction judgments.

Encourage Immediate Responses to Ensure Good Response Rates. Getting participants to complete evaluation forms as part of the training program allows organizers to measure the perceptions of the participants while the information is still fresh in their minds. It also ensures that most of the participants’ perspectives will be accounted for, by increasing the likelihood that forms will be completed and returned. Alternatively, training organizers may ask participants to complete and return evaluation forms at a later date, or may have participants complete an on-line survey at a later date. If this method is selected, evaluation forms should be returned within two weeks of concluding the training (organizers should follow-up with reminder emails to encourage completion and return of the evaluation forms at one week to ten days). The longer you wait to assess satisfaction the more likely that response rates will decrease. In addition, the accuracy of satisfaction measures decreases the further away in time the evaluation is conducted from the completion of the training. Even if evaluators are planning on conducting a follow-up survey 3 – 6 months after the training (i.e., to measure learning retention or behavior change), an initial evaluation of satisfaction should be assessed

Tip from Field-Test Sites: Some of the sites participating in field tests of the recommendations and tools in this Guide provided training participants with a short version of the evaluation form at the beginning of the training program. Participants were encouraged to use the form to record their immediate thoughts, impressions, ideas, etc. while they were still fresh in their minds. At the conclusion of the training, participants were directed to complete a full online version of the evaluation form. Participants could use the short form to refresh their memories if needed, but training organizers benefited from the ease of data collection and analysis afforded by the use of a web-based online survey.

36 Supra note 35.
as soon as possible after the training session. Tips for increasing survey response rates are covered in later chapters of this Section and in the Tools and Resources for this Section.

**Encourage Honest Responses.** Clearly, it is important to obtain honest responses from participants about their training experience. If you provide respondents with information about the goals of the survey and how critical their feedback is to the improvement of future training programs, you can motivate respondents to provide honest feedback. But the best way to ensure that participants’ satisfaction responses are genuine is to ensure their anonymity or confidentiality. This can be accomplished by collecting data in a manner which does not tie the information provided to the person supplying the information. For example, don’t have participants put their names on questionnaires or surveys (you can assign a code to the completed survey if you need to track their responses for pre- and post-test comparisons). Have interviewers erase any identifiable information that may have been collected during interviews (i.e., name and contact phone number) before handing over the information for data analysis. However, even though you have implemented these procedures, if training participants aren’t aware of them they may still feel that they can’t provide open and honest responses. Be sure to include a statement on evaluation feedback forms that lets training participants know that their responses will be anonymous (e.g., “Your name will not be associated with the responses you provide”). Another way to ensure anonymity is to report the findings as aggregate numbers, and abstain from presenting data in a way that may allow people to identify someone with a particular response. For example, you can help maintain anonymity by reporting findings by the role of respondent. But be careful ... if there was only one or two judges attending the training, don’t report findings as “judicial officers attending the training felt that ...” as this may identify the individual respondents. If it is critical to supply information that can be tied to an individual for your evaluation report, be sure to ask that individual if you may report that finding, quote them, or otherwise refer to their feedback. Ensuring participants of their anonymity, and taking lengths to ensure their responses are confidential, are the best ways to get an honest measure of trainee satisfaction.

**Develop Acceptable Standards for Trainings.** It is important to develop criteria of acceptable standards for your training program. Establish a baseline of how you would expect trainees to respond to a training program – for example, consider an acceptable average satisfaction level that you would expect from trainees (e.g., “at least 80% of trainees will express satisfaction with the training materials”). Knowing the acceptable average level of satisfaction among trainees allows organizers to determine what changes are necessary to maintain or improve trainee satisfaction. In addition to developing your own acceptable standards for training, be sure to determine whether training funders or other relevant administrative bodies have established standards that can guide your understanding of training satisfaction.

**Compare Current Participants’ Responses to Standard.** Measuring satisfaction at each training allows organizers to generate not only a baseline average from which to compare future trainings, but also allows organizers to determine what standards they feel are acceptable for a given training. Once realistic standards have been established, you should continue to evaluate trainee satisfaction with various components of the
training program and compare your findings with the standards. If the standards are not met, make a change and modify the training program accordingly.

**Communicate Reactions as Appropriate.** Finally, communicate the results to the appropriate audience. Typically the appropriate audience for training evaluation results includes the training organizer, facilitator, faculty, training advisory committee, and training funders. In some instances, other parties may be interested in the results. You may want to make an abbreviated report of the training evaluation results available to your training participants -doing so provides them with an opportunity to see how their feedback was used and may serve to encourage participation in your future or follow-up evaluation activities. To determine exactly how you should communicate trainee satisfaction, consider who wants to see the results, and whether it is appropriate for them to do so. See Section Seven of the Guide for recommendations for analysis and reporting of training evaluation findings.

Consideration of the guidelines for evaluating satisfaction above can help training organizers determine the level of inquiry desired out of a satisfaction measure. Knowing what you want to find out, how you want to find out about it (i.e., quantitative or qualitative), and how the information will be used provides organizers a general framework for developing their evaluation. If the goal is to measure satisfaction on only a few components of a training program, the questionnaire can be focused and brief. If the training is new, or organizers are interested in every component of the training, then a more in-depth survey with greater detail may be required.

**Types of Questions to Ask When Evaluating Satisfaction**

Clearly training organizers can ask satisfaction-based questions about almost any aspect of a training program. As a result, it’s helpful to group satisfaction questions into “types” or categories of questions and then to prioritize among those questions. The following text provides some sample questions under each general category or domain of training satisfaction measurement. Many of the sample questions below also lend themselves to follow-up, open-ended questions should training organizers require more detailed explanation for a specific response. The list of sample questions is illustrative only and is not meant to be exhaustive of the satisfaction questions that might be asked under each category. See later Chapters of this Section for more detailed instruction about question construction.

**General Satisfaction Questions**

General satisfaction questions are designed to capture trainees’ overall impression of the training. Some sample general satisfaction questions are provided below. Unless otherwise noted, these sample questions would include a rating scale for responses from “Highly Satisfied” to “Not at all Satisfied” – more about the use of rating scales and determining appropriate anchors for those scales in later chapters of this Section.

- How satisfied are you with the overall training?
• How would you rate the training program overall? [Provide rating scales from “Very Useful” to “Not at all Useful,” and “Very Interesting” to “Not at all Interesting”]
• How satisfied are you that the time you spent at this training is beneficial to your job?
• To what extent do you feel your personal learning objectives have been achieved? [Provide rating scale from “Fully” to “Not at All”]
• To what extent do you feel your expectations for this training have been met? [Provide rating scale from “Fully” to “Not at All”]
• Which of your personal learning objectives were not achieved at this training and why?
• Which of your expectations for the training were not met and why?

Specific Questions about the Training Curriculum

An important component of trainee satisfaction involves gauging impressions of the training material. If trainees are not stimulated by the material, they will not incorporate it into practice. Therefore, training evaluations should assess trainee satisfaction with the material presented during training. Examples include:

• How satisfied were you with the way the training program was organized (i.e., the order of topics)?
• Which parts of the training curriculum will be most useful to you in your job?
• Which parts of the training curriculum will be least useful, or not at all useful, to you in your job?
• Please indicate your agreement with the following statement: The materials provided to me at this training were tailored to meet my training needs. [Provide a rating scale from “Totally Agree” to “Totally Disagree”].
• Are there any subjects you would like to have seen included that were not included?
• To make way for additional curriculum materials, what would you omit from the training program?

Specific Questions about Presenters and Presentations

A key component of trainee satisfaction involves impressions of the presenters or faculty. Measuring satisfaction with the presenter provides an indication as to whether the presenter was successful in establishing his or her credibility and holding the audience. Presenting information in a clear and concise way that is easy to follow is essential for facilitating adult learning. Training evaluations should incorporate items investigating trainee satisfaction with both the presenters and the presentations. Examples include:

• How satisfied were you with the presenter overall? [Insert specific presenter or presentation for these questions as appropriate]
• How satisfied were you with the presenter’s style?
• How satisfied were you with the way the subject matter was presented?
• How satisfied were you with the pace of the presentation?
• How satisfied were you with the organization of the presentation?
• Please rate the presenter’s ability to engage the audience [Provide rating scale from “Highly Skilled” to “Not at all Skilled”]
• How useful was the exercise to you as a means of illustrating the application of concepts? [Provide rating scale from “Very Useful” to “Not at all Useful”]
• How satisfied were you with the materials handed out to participants?
• Please rate the usefulness of the materials handed out to participants [Provide rating scale from “Very Useful” to “Not at all Useful”]

Specific Questions about Facilities

It’s often valuable to obtain the reactions of training participants to matters outside of the evaluation of the learning itself (e.g., training facilities, training administration, etc.). By using a well-constructed and effective feedback form – not one that is skewed to prompt favorable comments – useful data can be obtained to help plan future training. Clearly, the training environment or atmosphere may also impact trainee satisfaction levels. Trainees will not be motivated to listen if they are not comfortable with the facilities. To assess satisfaction with training environments, questions should be included about satisfaction with the location of training, the hotel where trainees stayed, the rooms in which the training occurred, as well as the food that was served. Examples include:

• How satisfied were you with the location of the training?
• How satisfied were you with the room in which the training was held?
• How satisfied were you with the sound quality in the room in which the training was held?
• How satisfied were you with the food served during the training?
• How would you rate the quality of the food served during the training? [Provide rating scale from “Excellent” to “Poor”]
• How would you rate the comfort of the training room? [Provide rating scale from “Excellent” to “Poor”]
• How would you rate the ease of travel to the training location? [Provide rating scale from “Excellent” to “Poor”]

In summary, measuring satisfaction is important to ensure that the audience is accepting of the training - providing organizers with some evidence that the program was able to overcome the barriers associated with adult learning mentioned in Section Two of this Guide. Measuring satisfaction also provides training organizers with an estimate of how well the training was received.

See the Chapter Two Tools and Resources for additional sample satisfaction measurement questions and sample or template satisfaction measurement forms. See also the Question Bank Tool included with the Tools and Resources materials for more sample satisfaction and reaction measurement questions.
Chapter Two: Training Satisfaction and Reaction Measurement

Section 2.2
The Survey Process

Clearly a major tool in the assessment of satisfaction with a training program is the survey or questionnaire. However, in discussions on surveying generally, and on satisfaction or reaction measurement specifically, the focus is often incorrectly placed only on the survey instrument and not on the entire survey process. The entire survey process includes defining the survey objectives, developing a sample frame (a list of the population of individuals of interest from which a sample to receive the survey will be drawn), designing questions, specifying the strategy for data collection, and conducting the appropriate analyses. This entire process is important to achieving acceptable response rates, obtaining reliable and valid findings from your survey, and receiving acceptance for the recommendations that are generated from your results.

The Formal Survey Process

1. Define the Survey Objectives:
   a. Specify the population of interest (who you need to survey)
   b. Determine the type of data to be collected
   c. Determine the desired precision of the results

2. Determine who will be sampled:
   a. Specify the method of sample selection as either probability-based or convenience-based.
      For training evaluations, this is almost always the entire group of individuals who attended the training program
   b. Create a sampling frame if necessary (a list of individuals from whom you will draw the survey sample)
   c. Select the sample

3. Create and test the survey instrument:
   a. Choose the response mode (handed out at training, mailed, web-based)
   b. Draft the questions
   c. Pre-test and revise the survey instrument

4. Contact respondents throughout the survey process:
   a. Notify respondents that the survey is coming
   b. Provide instructions for completing and returning the survey at the delivery stage
   c. Remind respondents to complete the survey through post-delivery
   d. Send post-delivery thank-you’s for completing the survey
   e. Conduct non-response follow-up for those who did not return the survey

5. Collect data, data reduction, and data analysis

---

37 The Total Design Method, Dillman (1978; 2000); see also Fowler (1988).
38 In a probability sample, the probability with which an individual is selected into a sample can be computed. When the probability cannot be computed, the sample is called a convenience sample (it is more “convenient” to not have to worry about the ability to compute probabilities of selection).
Important Considerations in Planning and Designing a Survey

Increasing Survey Response Rates

The response rate is used to evaluate the success of your data collection effort. It is simply the number of people responding (or interviewed) divided by the total number of all of the people sampled. The denominator in this equation includes all of the people who were selected to receive a survey (e.g., attended the training) but did not respond for whatever reason. With respect to surveys in training evaluations, your response rate is calculated by taking the total number of completed and returned surveys and dividing that number by the total number of training participants (those individuals who received the survey in the first place).

Clearly, the greater the response rate, the more confidence you can have that evaluation findings derived from the survey represent the experiences of your training participants. If your response rate is low, then your findings may be peculiar to those individuals who were more inclined to complete a survey - and not truly reflective of the experience of the majority of training participants. Response rates vary widely for different types of surveys. Customer satisfaction surveys and market research surveys often have response rates in the 10% - 30% range. Employee surveys typically have a response rate of 25% - 60%. An important incentive to survey respondents is that their opinions will be heard and action will be taken based on their feedback. If respondents believe that participating in a survey will result in real improvements, response rates may increase, as will the quality of the feedback. Response rates can soar past 85% (about 43 responses for every 50 invitations sent) when the respondent population is motivated and the survey is well-executed. Response rates can also fall below 2% (about 1 response for every 50 invitations sent) when the respondent population is less-targeted, when contact information is unreliable, or where there is less incentive or little motivation to respond. Regardless of the type of survey you are conducting, you can have a major effect on the number of respondents who complete your survey. In order to increase your response rates, attention needs to be paid both at the survey design stage and at the collection stage.

To motivate respondents to agree to answer a survey and to complete it accurately, survey research demonstrates that the length of a survey is not as important as its design. Follow the strategies listed below to help increase your survey response rates:

- Particularly important when a respondent first views a survey is that it looks easy to do and the instructions are clear and consistent. To format a survey to maximize responses, follow some simple principles that give visual signals that the survey is easy to complete, such as creating a sense of space in the survey by eliminating the unnecessary use of lines - a long survey with lots of white space looks easier to complete than a two page survey filled with lots of print and lines.

---

39 See for example, Herberlien & Baumgartner, 1978; Steele, Schwendig & Kilpatrick, 1992; Yammarino, Skinner & Childers, 1991
40 Ibid.
41 Ibid.
• While thinking about including space in your survey’s design, be sure to make your survey as professional-looking as possible. This will increase your credibility and help to ensure that people invited to take the survey believe that their responses are important and will be used.

• Use the logic model, objectives and purpose for the evaluation that you developed to base your evaluation questions on - that is a useful way to cut back on extraneous evaluation questions. Focus your questions on evaluating the pieces of your training program that are most important to you and to your stakeholders.

• If your survey is not completed immediately after the training, don’t give participants too much time to complete the survey. While the measurement of the application of training content and resulting behavior and practice change benefits from delay, response rates do not. With satisfaction measurement you also want respondents to be able to recall their impressions of the training experience accurately - this is facilitated by short turn-around times between the completion of training and satisfaction measurement. We recommend 2 weeks as a run time for training surveys in which it is important to get a full response (as measured from sending out the survey to closing the survey for analysis).

• Be sure to use follow-up reminders to non-completing survey recipients after the original invitation is sent. A follow-up within 10 days after the initial invitation is optimal.

• Offer an incentive to complete and return the survey. Research results demonstrate that incentives will typically increase response rates by 10-15% (depending on the quality and attractiveness of the incentive to your target audience). 43

• Personalize the survey - invite respondents to take surveys that are sent to them after the training by addressing the invitation with their name. Research has shown that personalization of e-mailed survey invites (‘Dear Kathy’ instead of ‘Dear Training Participant’) can increase response rates by 8% or more. 44

---

44 Heergweth (2005).
• With online surveys, give consideration to the time you send your survey link and invitation out. Research indicates that response rates and times are best for surveys sent out between 6:00 AM and 9:00 AM, at the beginning of the work day – but not on Monday morning.45

• Make sure you make the link between the survey and the needs of the individuals who you have invited to take the survey - make the survey salient. Reinforce for training participants (during the training itself and during the survey invitation) how the evaluation findings are used to design and implement trainings that assist them in performing their job functions and achieving their court improvement goals. Salience of an issue to the population receiving the survey has been found to have a strong positive correlation with response rate for mail and web-based surveys.46 The greater the importance of the topic to the individual the more likely they are to complete and return the survey. If individuals attach little interest or importance to the content of the survey, then it won’t matter if the survey is short and appears to be easy; the person is still unlikely to respond.47

Question Order

Question sequence requires careful thought to reduce the likelihood of bias in surveys, interviews or focus groups. Order can affect results when a general question and a more specific question are asked in sequence. Research using “split-ballot experiments,” in which the order of questions is rotated, suggests that results from a general question are likely to be affected depending on whether it comes before or after a specific question.48 For example, consider the order of questions in this hypothetical training needs assessment:

[A Specific Question] How would you describe the quality of child abuse and neglect permanency hearings in your jurisdiction? (Please circle the appropriate number below):

[General Question] Overall, how would you describe the quality of child abuse and neglect hearing practice in your jurisdiction? (Please circle the appropriate number below):

In the examples above, placing the specific question first may bias respondents’ answers to the more general question that follows because their thoughts about the permanency hearing have become salient as they consider their “overall” rating of hearing practice.

Clearly, ensuring that the order of your questions doesn’t potentially bias responses is an important consideration when putting together the questions for your survey. Also important, is a logical sequence for the questions. Questions tapping the same domain, for example, are best grouped together (e.g., questions about specific faculty and workshops should be grouped together, as should questions about training logistics, venue, etc.). Feedback about the logical order for questions, ease of understanding and navigation through the survey, should also be obtained before finalizing the survey instrument (see pre-testing below).

45 Ibid.
47 Martin (1994).
Pre-Testing and Piloting Evaluation Questions
Before finalizing your survey, it’s important to pre-test or pilot the survey instrument. Pre-testing or piloting is an important component of the evaluation process. It affords researchers the opportunity to identify any problems with instrumentation before beginning formal data collection. Ideally, your evaluation instrumentation pre-test or piloting procedures should involve having your survey reviewed by experts who are knowledgeable about survey question construction and the subject matter, as well knowledgeable about the objectives of the training program itself. Experts can tell you which questions appear too complex to be administered easily and which are too long or too difficult to be answered accurately. Experts can also help you prioritize among questions in order to ensure your final survey instrument is concise and lacking in superfluous questions – that every question is logically tied to the objectives of the training and will provide useful information to training organizers. Reach out to your evaluation research community, local universities or the National Child Welfare Resource Center on Legal and Judicial Issues for technical assistance with survey construction and training evaluation instrument design.

Pre-testing or piloting of your instrumentation should also involve having instruments reviewed by potential respondents. In the training context, potential respondents are people who are eligible to be part of the survey sample – the kind of people you want to hear from. For example, if your training is focused on judges, then a review of your instrument by judges who will not be involved in the training would be useful. A review by potential respondents helps to guarantee that the survey’s questions are understood by your target audience, and are meaningful and inclusive of all important ideas.

In addition to a review of the survey content, be sure to pre-test the time it takes to complete survey instruments. This is especially important when you are suggesting timeframes to your respondents for the completion of surveys. Have a sample of individuals who are unfamiliar with the survey (i.e., have not been involved in the survey’s design) take the survey and time their completion. This will provide you with a fairly accurate estimate of time that respondents can expect it to take to complete the survey. You can then cite this time in your invitations to take the survey (web-based or paper forms).
Section 2.3

Question Construction

Above all, the questions you include on any evaluation instrument should be purposeful. Purposeful questions are those that are logically related to your objectives. Construction of your training evaluation questionnaire, whether it is primarily designed to gauge satisfaction with the training experience or is more comprehensively aimed (as this Guide recommends) to assess satisfaction, learning, and behavior change, begins with an identification of the following:

- How will the survey or questionnaire fit with other training evaluation data collection methods?
- What kinds of questions need to be asked, based on the training needs assessment, identification of learning objectives and training program logic model?
- Who will receive the questionnaire and when in the training and evaluation process will responses be sought?
- What types of decisions will be made from the collected data?
- What kinds of questions or items will elicit useable data?
- Are there existing questions that have been used successfully in other surveys that we can adapt?
- What format will be used for the questionnaire – keeping in mind that the format should be as easy as possible for the user and for analysis?
- How will the data be analyzed and collated with other evaluation data to establish training priorities and to make determinations about training impacts and outcomes?

Types of Questions - Closed vs. Open-Ended Questioning

There are two types of survey questions: open-ended questions and closed-ended questions. Each type has advantages and disadvantages. Closed-ended questions provide a range of answer categories or options from which participants can choose. Open-ended questions provide no answer categories but allow participants to express ideas in the language they choose. Closed-ended questions are considered more efficient for both the respondent and for analysis. Respondents experience fewer burdens in completing closed-ended questions than opened-ended ones, thus insuring you a greater response and less missing data. Additionally, you can tally questions with answer categories more quickly. However, open-ended questions provide respondents with the opportunity to tell you more about their opinion or experiences, and to clarify close-ended responses (e.g., by providing answers to “please explain” or “why do you feel that way?”). Open-ended questions provide the necessary explanatory models for quantitative survey findings.

Clearly, question types should depend largely on the needs and goals of the training, the length of survey or questionnaire, and the type of analyses planned. All types of questions and responses can be correct, if used
appropriately. Three types of questions are typically used to measure satisfaction and reaction to dependency court trainings: closed-ended scale item questions, closed-ended forced choice (i.e., true/false) questions, and open-ended questions.

**Closed-Ended** – Forced-choice or closed-ended question formats ask respondents to select one of two options – typically a *true* or *false* or *yes* or *no*. Forced choice response questions are often used when determining *if’s* – if training was informative, or if the materials were clear. Forced choice (i.e., yes or no) questions do not allow for variation in responses and are only preferable when there should be no variation in response. Because individuals are forced to select between two options (such as yes or no), evaluators have no way of knowing why the response was selected. For example, if a respondent indicated that the information provided was not useful, without a follow-up question evaluators would not know why the respondent felt that way. This means the information could have been highly informative or only slightly informative. Other forced-choice question formats include rankings, paired comparisons, and ‘most’ or ‘least’ important questions.

- **Rankings** – Items are ranked by respondents in order of importance or preference, with only one rank assigned per item. This format works best with 15 items or fewer.
- **Paired Comparisons** – Items are presented in pairs, and respondents choose one from each pair.
- **Most and Least Important** – Respondents choose the three to five most important (or most desirable) and the three to five least important (or least desirable) from a list of statements.

**Open-Ended** – Open-ended questions allow respondents to give answers in their own words. These questions are useful if you are interested in getting unanticipated answers or in learning about the world as your respondents really see it. Also, some respondents prefer to state their views in their own words, which sometimes results in quotable material that can be useful for your evaluation report. The responses to open-ended questions, however, are often more difficult to compare and interpret.
Table 2-1

<table>
<thead>
<tr>
<th>When to Use Open- and Closed-Ended Questions</th>
<th>If Yes, Use Open-Ended Questions</th>
<th>If Yes, Use Closed-Ended Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>Respondents own words are essential (to give voice, to obtain quotes).</td>
<td>You want data that are rated or ranked and you have a good idea of how to order the ratings (or response categories) in advance.</td>
</tr>
<tr>
<td>Respondent Characteristics</td>
<td>Respondents are capable of providing answers in their own words. Respondents are willing to provide answers in their own words.</td>
<td>You want respondents to answer using a pre-specified set of response choices to facilitate interpretation and comparison across respondents.</td>
</tr>
<tr>
<td>Asking the Question</td>
<td>Potential choices or possible response categories are unknown.</td>
<td>You prefer that respondents choose among known choices.</td>
</tr>
<tr>
<td>Analyzing the Results</td>
<td>You have the skills to analyze respondents’ comments even though answers may vary considerably.</td>
<td>You prefer to quantify or count the number of responses per specified choice or category.</td>
</tr>
<tr>
<td>Reporting the Results</td>
<td>You will provide individual or grouped verbal responses.</td>
<td>You will report statistical data (e.g., percentages, frequencies, average ratings).</td>
</tr>
</tbody>
</table>

Pitfalls to Avoid When Designing Questions

**Conversational Language** – A survey is not a conversation. To get accurate information, survey questions rely on standard grammar, punctuation and spelling. You should use words that maximize understanding for everyone involved in the survey. All questions should be reviewed and tested by people who are proficient in reading and speaking the language in which the survey is written, by content experts, and ideally, by potential respondents. Complete sentences should be used, whether as statements or questions, expressing clear and complete thoughts. In addition, avoid using slang and colloquialisms because they appear unprofessional, go out of fashion quickly, and may not be familiar to all of your survey respondents. However, it may be appropriate to use phrases that are “terms of art” or are familiar to your respondents if they come from a homogeneous group (e.g., individuals who share a common professional language).

**Poor** – Stakeholder role?

**Better** – Please identify your role in the dependency court system in the space provided (e.g., judge, child welfare worker, etc.).

**Poor** – Any thoughts on curriculum content?

**Better** – Please use the space provided below to include any thoughts you might have for improving the training curriculum content.

**Abbreviations and Acronyms** – Avoid using abbreviations and acronyms in questions unless you are sure that they are commonly understood. Most people are familiar with
Chapter Two: Training Satisfaction and Reaction Measurement

ASFA and CFSR, for example, in the dependency court system context. If in any doubt, it’s important to spell out abbreviations and acronyms.

Jargon and Technical Expressions – It is best to avoid the use of jargon and technical terms in your questions unless you have good reason to believe that all of your respondents are familiar with the terms.

Ambiguous Questions – Concrete questions are precise and unambiguous. Questions may be defined as precise and unambiguous when, without prompting, two or more potential respondents agree on the meaning of the words used in the question. In the training context, for example, words like “quality” may be difficult to interpret as individual respondents may have different perceptions or definitions of “quality.” The more detail you can provide in a question, the more reliable the answer is likely to be. To help make questions more precise and concrete, you may also consider whether adding a time frame might be appropriate.

<table>
<thead>
<tr>
<th>Use the following scale to rate each question:</th>
<th>Poor</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambiguous – Quality of written material provided</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Better – Relevance of written material provided to my role</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less precise – Will you apply the training learned?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More precise – Will you apply the training learned in your job in the next month?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Ambiguous wording – Use of words like quality of material may be difficult to interpret by respondents and lead to difficulty in interpreting responses |

Biased Words, Phrases and Response Sets – Biasing words or phrases elicit emotional responses that may have little to do with the issues addressed by the survey. They are considered biasing because they trigger emotional responses or prejudice. Bias may also arise in your survey if you do not fully understand the culture and values of the respondents and ask questions that are inadvertently offensive. To guard against this possibility, you need to have all questions reviewed and pilot-tested before you use them. In addition, bias may enter into your survey through your choice of response sets – a biased response set is created when the response categories offered do not include all relevant options. The omission of the extreme on the negative end of a rating scale, for example, biases responses.

| Biased – In your opinion, did the training do a good job of getting judges to care more about timely permanency for children? |
| Better – In your opinion, did the training do a good job of focusing judges on the importance of timely permanency for children? |

| Biased – Do you believe that the new protocol introduced at this training will finally get judges to exercise their oversight role? |
| Better – Do you believe that the new protocol introduced at this training will assist judges in exercising their oversight role? |

| Biased – Overall, how would you rate the presentation on timely permanency? |
| Excellent | Good | Fair | Poor |
| Better – Overall, how would you rate the presentation on timely permanency? |
| Excellent | Good | Fair | Poor |

Biased Response Set – The responses do not include an option for something worse than fair, such as poor, and are thus skewed toward a positive response.
Chapter Two: Training Satisfaction and Reaction Measurement

towards favorable ratings.

**Double-barreled Questions** - A double-barreled question contains two separate ideas. An example is “Do you think we should continue to train on ASFA and strategies for effective court-agency collaboration?” This question is really twofold: “Do you think we should continue to train on ASFA?” and “Do you think we should continue to train on strategies for effective court-agency collaboration?” Some respondents may endorse continued education on ASFA, some may endorse training on collaboration, some may endorse one but not the other, and some may choose not to endorse either type of training. No matter what the respondent answers to a double-barreled question, however, you will not know exactly what he or she means. To avoid asking double-barreled questions, check the use of the word “and” in your questions.

<table>
<thead>
<tr>
<th>Poor</th>
<th>Overall, the presenter was prepared and organized</th>
<th>5 4 3 2 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Better</td>
<td>Overall the presenter was prepared</td>
<td>5 4 3 2 1</td>
</tr>
<tr>
<td></td>
<td>Overall the presenter was organized</td>
<td>5 4 3 2 1</td>
</tr>
</tbody>
</table>

**Negative Questions** - Negative questions are difficult for many respondents to answer because they require an exercise in logical thinking. For example, suppose a question asks respondents if they agree or disagree with the statement “Prior to the adjudication stage, dependency mediation should not be used.” Some respondents will fail to read the word “not.” Others will mistakenly translate the negative into the positive and believe the question is “Do I think mediation should be used prior to adjudication in dependency cases?” If you use a negative question, be sure to emphasize the negative word: “Prior to the adjudication stage, dependency mediation should **not** be used.” Be careful not to use double negative wording as well, such as “Do you disagree with those that do not want to expand the mediation program?” Double negatives are not only grammatically incorrect but they are also very confusing.

<table>
<thead>
<tr>
<th>Negative Question</th>
<th>Do you agree with the statement that prior to the adjudication stage, dependency mediation should not be used?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Better</td>
<td>Do you agree with the statement that dependency mediation should <strong>not</strong> be used prior to adjudication?</td>
</tr>
<tr>
<td>Better</td>
<td>Do you agree with the statement that dependency mediation should only be used after adjudication?</td>
</tr>
</tbody>
</table>

Negative Questions - are often confusing for respondents
Measurement Choice: Choosing Response Categories

Response choices, or the choice given to respondents from among which they select their answers, take several forms. Deciding which kind of response choices you should use first involves determining what type of information you need. A categorical or nominal response choice involves categories, such as male and female, yes and no, ‘applies’ or ‘does not apply.’ These types of responses have no numerical or preferential values – they are simply correct or incorrect, true or false. A second type of response choice is ordinal, in which respondents are asked to rate or order the items in a list (e.g., from very positive to very negative). Numerical response choices call for numbers, such as years of experience or number of trainings attended.

**Rating Scales** - By far the most common types of questions for training evaluation are closed-ended rating scale items (e.g., Likert scales).\(^49\) Rating scales are usually 5 point scales with descriptions (anchors) at each end of the scale. Questions may be asked about frequency (never – always), amount (least – most) or satisfaction (not at all – very). Alternatively, you can use adjectives for each point of the scale (never, occasionally, some of the time, frequently, always). While 5 point scales are the most common, 4 point or 6 point scales can be used to prevent artificial clustering around a midpoint. Rating scales may also determine the strength or intensity of judgment regarding satisfaction with training delivery, course content, or other features or processes involved in the training program implementation. Some research indicated that the anchors of scales should be alternated (e.g., leading with the positive anchor for some sets of questions and then leading with the negative anchor for other sets of questions) in order to prevent a “response-pattern” or the tendency for respondents to go down a list of questions, not paying attention to those questions and mechanically circling scores.\(^50\)

**Example of a Rating Scale:**

Please indicate your degree of agreement with the following statement, by circling the appropriate number on a scale from “5” Strongly Agree to “1” Strongly Disagree.

The presentation of information was easy to understand.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

\*Poor balance* – Yes, constantly; Yes very often; Yes, once; No, never.

\*Better balance* – Yes, constantly; Yes, very often; Yes, fairly often; Yes, a couple of times; Yes, once; No, never.

**Guidelines for Determining Ordered Responses, Scales or Rankings**

- **Use a meaningful scale.** A meaningful scale is one that makes sense in terms of the survey’s specific objectives. To choose among potential scales, you may test one or more scales on a preliminary or pre-test basis and select the one that gives a good “spread” of answers and is the most meaningful to respondents.

---

\(^{49}\) Likert scales are probably the most widely used response scale featured in surveys – Created by Rensis Likert in the 1930s, his original scale featured five points.

\(^{50}\) Fowler, F.J. (1988); Fowler, F.J. (2001).
• **Balance all responses.** A scale is balanced when the two endpoints mean the opposite of each other and the intervals between the points on the scale are about equal. “Much worse” is the opposite of “much better,” and the meaning of the interval between “much worse” and “somewhat worse” is similar in degree to the interval between “somewhat better” and “much better.”

• **Use a neutral response category only if it is valid.** Provide a neutral category only when you are sure it is a valid response. A neutral category is either a middle point (“neither satisfied nor dissatisfied”) or a “no opinion” or “neutral” or “don’t know” option. Some survey research has indicated that providing neutral choices gives respondents an excuse for not answering questions. 51 If you think your respondents might react this way to neutral choices, pre-test your questions with and without neutral choices and compare the results. How many responses cluster around the middle? Alternatively, some survey researchers suggest that respondents may resent not having a neutral option, particularly a “don’t know” response category. 52 As part of the pre-testing process, ask respondents about the scales used – would another set of responses be more appropriate? While research regarding the use of a “don’t know” option is inconclusive, your training evaluation survey should include a “don’t know” option if it’s plausible that people may simply not know, not have an opinion, or be offended by the forcing of a response.

• **Use 5-7 point rating scales.** Over time, there have been many discussions and disagreements in survey research focused on one central question: What works best with the Likert scale to give you the most accurate responses? Most measurement scholars agree that more than seven points on a scale are too much. Studies have shown that people are not able to place their point of view accurately on a scale greater than seven. 53 What is the perfect number? Studies are inconclusive, but the most commonly recommended scales are five, four or three point scales. Current thinking suggests that 5-7 point scales are adequate for the majority of surveys that use ordered responses or rating scales. However, conclusive evidence for the superiority of either odd or even numbered scales is currently unavailable. You should use whichever best suits your survey’s needs, but as mentioned above, 4 point or 6 point scales can be used to prevent artificial clustering around a midpoint. Regardless of which scale you use, be sure to pre-test your questions and response categories!

• **Use rankings only if respondents can see or easily remember all choices.** Rankings or rank-order scales are a type of ordinal measure in which choices are placed in a list and respondents are asked to order them from the highest to the lowest (or the other way around). The rank of training priorities among a list of training options, for example. In self-administered written surveys or online survey questionnaires, it’s important that the list of options be readily visible in their entirety and not be broken by a page. If lists of alternatives are too long, you may want to consider asking respondents to choose the top two or three and the bottom two or three.

---

52 Babbie, E. (2009)
Web-based or internet surveys are becoming more common training evaluation tools. Survey instruments that respondents complete on the internet look like other self-administered questionnaires. The rules for writing questions for online surveys are almost exactly the same as those that apply to other self-administered questionnaires. They include making sure that you have a specific purpose for the question you ask, that you understand the needs of the survey's users, and that you have the resources available to conduct and complete all survey activities. Also, when you conduct any type of survey, you must respect the cultural and sociopolitical beliefs of respondents and their ability to understand and complete each question.

On the whole, the types of questions that are appropriate for online surveys are fairly similar to those used in other self-administered questionnaires. In fact, an important and ongoing area of survey research is whether online surveys are simply a technological advance in self-administered survey design or if they are actually a different type of survey altogether.

When Should you Consider an Online Survey?

Internet surveys may be preferable to written surveys in the following cases:

- The survey can be conducted with a convenience sample. In a convenience sample, the probability with which a respondent is selected into the sample may not be known – the respondents “self-select” into the survey.

- You have a list of e-mail addresses for the target population. The benefits in terms of cost and timeliness are greatest when the target population can be contacted to receive the survey via e-mail.

- The sample size is relatively large. Generally, web-based surveys have a larger initial start-up cost than written self-administered surveys. However, web-based surveys are more cost effective at the analysis stage with larger target populations, as web-based survey programs have automated data tracking, coding and analysis features.

- The survey contains questions of a particularly sensitive nature. There may be a bias toward socially acceptable answers (as opposed to more honest answers that may be less socially acceptable) in surveys in which the respondent has direct contact with the researcher.\(^{54}\)

- The survey contains a large number of important open-ended questions. Web-based surveys incur no coding or editing costs because responses are received electronically. There is also some evidence

that respondents give longer answers to open-ended questions in electronic surveys than open-ended questions in written self-administered surveys.55

- The survey includes graphics or contains interactive elements. A web-based survey easily allows the use of color and simple graphics to make your survey more attractive and professional-looking. Although online surveys rely on the same principles of question writing as do other self-administered surveys, the question and response formats you can use online are more varied. For example, questions in online surveys can offer responses in dropdown lists. In addition, while complicated skip patterns may appear confusing to respondents to paper surveys, surveys that require complicated skip patterns can be easily built into web-based surveys and you don’t have to rely on the respondent’s ability to follow your skip instructions.

- The survey requires some responses. There may be a reason, based on your evaluation objectives, to require a response to a question before an individual can move on to the next question. This is easily accomplished in an online survey.

**Guidelines for Designing and Implementing a Web-based Survey**

The recommendations for designing effective surveys (question construction, order, response categories, pre-testing, etc.) that we’ve already discussed apply to online or web-based surveys as well. However, to implement an effective online training evaluation survey, we also recommend the following:

- List only a few questions per screen - don’t require the respondent to scroll down too much on a single page.
- Eliminate unnecessary questions - use your logic model and training objectives to delete any extraneous questions.
- Use graphics and color sparingly – too many graphics and too many colors can be distracting.
- Use matrix (table or grid) questions sparingly. Or, if a matrix is used, don’t use too many lines in the matrix table - this may make your survey look unnecessarily difficult.
- Reduce response errors by restricting response choices. Make sure any critical questions to the analysis require responses before individuals can move on. Require responses for open-ended questions as well if they are critical to providing you with further clarification or explanation of responses.
- Be careful not to over-use the “force” choice option - use it only when a response is required for assessment of a specific learning objective.
- Ensure that respondents’ privacy and their perception of privacy are protected.
- Provide some indication of survey progress (e.g., a bar graph that shows the percentage of the questions completed or the percentage of the survey remaining to complete).

- Provide respondents with an estimate of how long it will take to complete the survey. Use pre-test information to make this estimate as accurate as possible.
- Allow respondents to interrupt and re-enter the survey.
- Inform respondents of the deadline for completing the survey (e.g., let them know when the survey will close).
- Remember to send a follow-up email to encourage respondents to complete and return the survey.
- Use automatic skip patterns (build these into the program rather than only in the instructions to respondents).
- Thoroughly pre-test the survey.
- Provide a means for respondents to report problems they are having with accessing or completing the survey.
Interview and Focus Group Methods

Qualitative methods such as interviewing and focus groups are important training evaluation tools that not only can provide more detailed information about satisfaction with the training experience, but also can provide detailed information about learning acquisition and behavior and practice change. By listening to individuals, you can learn about important aspects of participants’ experience, learning, and application of knowledge that you may not have known about before or can capture through quantitative means.

Focus groups and interviews are used in evaluation research methods to collect qualitative data. The utility of including interviews and focus group methods in your training evaluation depends on asking good questions. While interviews can vary in formality (from informal to structured interviews), interviews, just as surveys, need to carefully frame questions and ask all the respondents the same questions in the same words in the same order to provide reliable data for your evaluation.

Focus groups can be especially helpful to providing training organizers with more in-depth information about training experiences (as well as learning and application of new skills, policies and procedures in the field). The purpose of a focus group is to develop an understanding of the participants' experiences or views about an activity, a program, or a product or service. The purpose of a focus group is not to reach consensus, provide recommendations, or make decisions. Typically a focus group is composed of 8 to 10 people, who are selected because they have certain characteristics or experiences in common. Participants are asked to share ideas and perceptions about a particular area of interest in a relaxed, comfortable atmosphere. The discussion is carefully planned around a few open-ended questions. The focus group moderator plays an unobtrusive role, encouraging comments, both positive and negative, and being careful not to make judgmental comments. At the opening, the moderator ensures that participants understand that they are expected to express opinions and feelings and that alternative explanations are expected. Participants are assured that there are no right or wrong answers, only alternative points of view. The moderator promotes interaction and assures that the discussion remains on topic. The interview follows pre-determined questions arranged in a logical order.

In collecting data needed for training evaluations, focus groups have a number of advantages over surveys or individual interviews. Focus groups are a rich source of information because participants' comments build on the insights of other participants. A wider range of insight is provided because a comment by one person often triggers a chain of responses. Focus groups provide a quick and cost effective method of gaining a large amount of information.
Important Considerations when Implementing Interview and Focus Group Methods

**Be Consistent and Neutral.** In conducting interviews for evaluation purposes, it is important to be consistent and neutral. Just as you want to reduce bias in quantitative measurement instruments, you want to reduce bias when conducting qualitative interviews. Structured interviews (with pre-set questions, including agreed upon probes to encourage detail) facilitate treating all participants in the same manner during the interview. However, whether your interview is structured, semi-structured, or more informal, to reduce bias, all interviewers need to be consistent and neutral in their interviews of individuals.\(^{56}\) If you do not treat all training participants the same way in an interview, your conclusions about participants’ experiences may be undermined by any differences in how you treated the interviewees. For example, you may find that program participants who mentioned positive aspects of the program talked in more detail during interviews than participants who mentioned negative aspects. But your findings may be undermined by the fact that you may have, even inadvertently, encouraged participants to talk more when they mentioned something positive about the training program by using reinforcing statements such as “I think you’ve mentioned a really good point.” On the other hand, you may have discouraged participants from expanding on their points when they mentioned something negative by using statements such as “I’m disappointed to hear that.”

**To be Consistent, interviewers should:**
- Always read the instructions to each participant as stated.
- Ask every question as stated.
- Ask the questions in the order stated.
- Use standard phrases in response to a participant if he/she does not understand a question, or gives tangential or overly lengthy answers.

**To be Neutral, interviewers should:**
- Avoid agreeing or disagreeing with a participant.
- Avoid indicating that a participant’s answer is ‘right,’ ‘wrong,’ ‘good,’ ‘poor,’ or ‘interesting.’
- Avoid suggesting an answer or interpreting a question for a participant.
- Avoid giving your own opinion.
- Use standard phrases in response to a participant who is ‘chatty’ or asks for the interviewer’s opinion.

---

\(^{56}\) Patton (1987).
Chapter Two: Training Satisfaction and Reaction Measurement

Ensure a Good Flow for Questioning. One challenge in reading all of the questions in the same way, is the need to use a smooth and conversational tone in all parts of the interview that are read to respondents (including instructions, probes and prompts). Even formal interviews should flow like a conversation. Avoid organizing interview items in a way that requires the interviewer to go back and forth in the interview form. Generally, the first items in your interview should establish the purpose for the interview and maintain the respondent’s interest. As a result, any demographic questions that need to be obtained during the interview should not be among the first questions. When the questions flow logically from the introduction, respondents are drawn into the interview rather than being distracted and annoyed by questions they may consider irrelevant. A smooth start also sets the tone for the rest of the interview, establishing a “rapport effect” that builds trust and enhances the respondent’s willingness to participate fully in the interview. Include transition statements that tell the respondent what topic the interviewer is going to address next. Transition statements give respondents a sense of movement through the interview and establish an overall coherence among the parts of the interview. Again, when designing your interview instrument, you must consider the possibility of question-order effects, or situations in which answers to certain questions may influence respondents, consciously or unconsciously, in their responding to later items.

Identify Information-Rich Participants. Think carefully about the characteristics of the people you should have in your interviews or focus groups. Who will give you the most in-depth information to supplement any quantitative data you have about the training? Are there specific groups that you need to hear more from? It sometimes helps to think of this as

Sample Interview transition statements:
First, I’d like to start by asking you some general questions about your training experience ...

Now, I’d like to learn more about specific workshops you attended ...

Next, I’d like to ask you for your recommendations for how the training program can be improved ...

We’re almost finished with the interview; I just have a few questions about ...

This completes our interview. Thank you for taking the time to participate - do you have any comments you’d like to add?

The general rule of thumb for the number of focus groups in evaluation is to plan for 3-4 focus groups. Once you have conducted these, determine if you have reached “saturation.” Saturation describes the situation when you have heard the range of ideas or issues and aren’t getting new information. If you are still getting new information after 3-4 groups, you may need to consider additional groups.
identifying the “information-rich” respondents. After determining what you need to discover through the interview or focus group process, ask yourself “Who has the greatest amount of insight on this topic?” Include those individuals in interview and focus group methods.

Steps in the use of Focus Group Interviews

1. **Developing a questioning route.** Questioning route is known as the order or sequence in which questions will be administered. The questioning route should be based on the objectives you have defined for the focus group. Brainstorming with colleagues and prospective users of the information can then help you to generate the questions.

2. **Recruiting the participants.** Participants for the focus group should be those individuals who are most likely to provide you with in-depth information on the target issue, or those individuals from whom you have not had input (i.e., an under-represented group). Participants should also be selected on the basis of their ability to discuss freely in a group, and their interest in the topic.

3. **Planning resources.** Develop a timetable for the focus group, including planning, implementation, and evaluation and reporting. Consider the resources both in staff, time and cost that will need to be marshaled for the focus group. Secure a comfortable facility for the focus group - one that includes sufficient space for flip-charting the discussion.

4. **Moderating interviews.** Ensure the moderator is someone that is comfortable and skilled at facilitating group discussion and is well-versed in the objectives of the focus group study. Moderators must keep the discussion on track. Moderation involves bringing the conversation back on target when irrelevant topics are introduced. This guidance has to be provided without reducing group enthusiasm and interest in the discussion. There are several personal attributes of a good moderator, including: familiarity with group process either from previous experience in working with groups or through training in group dynamics; good listening skills; adequate background knowledge on the topic of discussion; well-developed written and oral communication skills; and a sense of humor.

5. **Data analysis and reporting.** Data analysis and reporting can follow an interpretative summary format, whereby the data are not only described but also interpreted. This analysis produces an interview summary including key incidents, strong statements, and frequently occurring responses. Next, the key incidents, strong statements, and frequent responses are classified by question, coded, and grouped. The coding and grouping helps to identify the general themes in the responses. The identified themes are then compared across interviews (or series of focus groups) in order to develop a general picture on a question by question basis and to draw conclusions related to the objectives of the focus group study.

---

58 Krueger & Casey (2000)
59 Ibid.
60 Miles & Huberman (1994)
Ideally, surveys of training participants about their satisfaction with and reactions to the training would include the entire population of training attendees. In this situation, a sample would not be needed. However, it may be the case that you have to “sample” the training population for survey measurement - and also to sample the training participants to receive additional follow-up measures that cannot realistically be applied to all of the training attendees (due to cost, time constraints, etc.). All samples are subject to error (called “sampling error”), though this error can be minimized by including larger samples and samples of individuals that are homogenous (similar).

**Convenience and Probability Samples**

There are two main types of survey samples: convenience samples and probability samples. Convenience samples arise from uncontrolled instrument distribution (not everyone who could receive a survey receives the survey) or self-selection (i.e., individuals self-select to take and complete the survey). Probability samples, also called random samples, are samples in which the probability with which an individual was selected into the sample can be determined. Probability samples are purposely random - each person in the survey population has an equal chance of being included or excluded from the survey sample. Probability samples can be classified into three types: those taken from closed populations (such as the list of individuals who attended a training), general populations, and pre-recruited panels. Convenience samples are often less costly to generate than probability samples, but the statistical inference needed to generate conclusions from those samples becomes problematic.

While the nuances of survey sampling are beyond the scope of this Guide, the following matrix provides definitions and examples of when different approaches might be used.
## Table 2-2
### Probability Sampling Designs

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Random Sampling (SRS)</td>
<td>Each person in the survey population is assigned a random number, which is then randomly selected for inclusion in the sample.</td>
</tr>
<tr>
<td>Systematic Sampling</td>
<td>The survey population is systematically numbered 1-X, and each Nth person is selected. For example, each 10(^{th}) person is selected for inclusion in the sample. This differs little from SRS, but it is more accurate in many instances. One danger to consider is periodicity: if the survey population is coded cyclically, (e.g. - in an order), the sample may be biased. A second danger to be aware of is implicit stratification: if the persons in the survey population are arranged in a certain pattern (e.g. - alphabetically by last name).</td>
</tr>
<tr>
<td>Stratified Sampling</td>
<td>Instead of drawing the sample from the entire survey population, the survey population is broken down into categories that are then sampled appropriately (though not necessarily equally). For example, a population might be broken down by age before being sampled. Stratified sampling produces a more representative sample from the survey population, markedly reducing sampling error.</td>
</tr>
<tr>
<td>Multistage Cluster Sampling</td>
<td>Cluster sampling is best used when the survey population is too large to list individually for sampling (e.g. – all judges in the United States). It involves listing possible data sources (or clusters) and either sampling or stratifying those clusters until a representative, usable sample is defined. For example, you might list all states in the United States and then sample or stratify the states into one cluster; then you might list all counties in each selected state and sample or stratify the counties into another cluster; and so forth, until you arrive at a reasonable and representative sample.</td>
</tr>
<tr>
<td>Probability Proportionate to Size (PPS)</td>
<td>In some instances of Multistage Cluster Sampling, very large populations stand a chance of being overrepresented, and very small populations stand a similar chance of being underrepresented. In such cases, combining several small populations into one larger cluster is common practice, as is pre-sampling larger populations to account for their overrepresentation.</td>
</tr>
</tbody>
</table>

### Non-probability Sampling Designs

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purposive Sampling</td>
<td>Purposive sampling involves the selection of a sample based on the researcher’s knowledge of the larger community or sampling frame. Though this provides very well-articulated results, they often fall victim to biases of the researcher.</td>
</tr>
<tr>
<td>Quota Sampling</td>
<td>For quota sampling, the researcher must have considerable demographic data about the sample population. The sample population is broken down into demographically-delineated groups which are then appropriately weighted according to their portion of the total population, resulting in a reasonable representation of the sample population.</td>
</tr>
</tbody>
</table>
Chapter Two: Training Satisfaction and Reaction Measurement

TOOLS AND RESOURCES – CHAPTER TWO

- Checklist to Determine if a Formal Instructional Design Approach is Needed
- Curriculum Design Worksheet
- Framework to Design a Training Plan
- Overview of Data Gathering Methods for Needs Assessment
- Personal Learning Styles Inventory
CHAPTER TWO: References and Resources


