Researchers tend to gravitate towards two kinds of widely used research methods: surveys and focus groups. Surveys tend to provide quick, raw data that can easily be analyzed with little time constraints. Focus groups are a good way to dig deeper into the mechanics of a program. Analyzing data from a focus group can be much more time consuming, but it yields valuable, in-depth feedback since participants can provide unadulterated comments in a more stream of consciousness manner. Below are guidelines that one must keep in mind when designing these methods.

Survey Design
(Research Tips: [Research Tips])

In order to design a survey, one must first begin with 2-3 objectives that they are interested in measuring. A researcher can begin by asking, “What do you want to know about this program? How would one define success for this program? What will help improve it?” A survey should be tailored around 2-3 objectives. This way various questions can be used to measure one construct. A construct is the theme or concept that is being assessed in the analysis. If one were to measure 20 different constructs, then a survey could only contain one question per each concept. Hence, the survey would be exceptionally long and would result in survey fatigue. Respondents might blindly provide responses in order to complete the survey as quickly as possible as opposed to nulling each question to produce accurate results. Merely using one question to analyze each construct would not provide a very large picture of these constructs. It is important to measure one construct from multiple angles to ensure precision, accuracy, and thoroughness. One key term to remember is triangulation where multiple measures are utilized to measure one construct. This allows researchers to examine a construct through multiple lenses. One example of triangulation in an everyday, real world setting would be using several alarm clocks to wake oneself up in the morning. Under the circumstances that one alarm clock would fail, a person can be certain that they would still be awoken by an alternate alarm. Triangulation uses the same principle; attacking a construct from several avenues ensures that every possibility is examined. This way if one method misses a nook or cranny, the odds are that one of the alternate measures will detect it. This same idea can be applied to surveys. It is advantageous if a small number of constructs are selected because this facilitates the idea of being able to investigate a construct using multiple questions. This way the same construct can be looked at through more than one perspective. A survey’s results will then yield more meaningful information.
When designing a survey, questions should be lumped together according to common categories or themes, so that it is easier to follow. This helps the survey flow better, makes it more aesthetically pleasing, and requires less mental work from the survey respondent. It is imperative to leave a section for open-ended responses in a survey. This way respondents can provide any other feedback that they survey did not tap into. Open-ended responses allow an individual to free-write their thoughts on strengths or weaknesses of the program. These responses might also provide ideas for future modifications to the survey. If respondents continuously bring up the same issues in these open-ended questions, it might be helpful to construct Likert Scale items on these topics to further investigate them. A Likert Scale is, quite possibly, the most commonly used question format in a survey. This scale will be discussed in more detail later on.

A pragmatic open-ended question will be specific and concise. If a survey merely asks respondents if they have any additional comments, a respondent might be more inclined to leave the question blank or provide a one-word response. Respondents might write blanket statements such as, “good job!”, “the workshop was helpful”, or “thanks for holding the workshop”; however, they will not indicate what, particularly, was helpful about it. The purpose of open-ended questions is to further probe a respondent about key issues. It can ask, “what are some aspects of the program that you would like to change?” or “What component of the program have you been able to apply the most in your life?” If an open-ended question seeks to inquire about knowledge that was acquired as a result of a service that was received, it might be best to provide the respondent with a scenario. Scenarios are a great way to assess the range of knowledge that each participant attained. The respondent could be told that a colleague of theirs is struggling with “A”, “B”, and “C”. Then it can ask what advice they would provide them with or what features of a particular program could help them resolve the problem. This real-world application measures more of the “take-home” message of the program. It looks at how individuals can utilize a program’s resources and apply them to solve real-world problems.

The most commonly used type of survey question, as was previously mentioned, is the Likert scale item. Likert Scales usually consist of a statement or question that requires the respondent to provide some sort of rating. For example, a respondent might see the following: “Please rate your level of satisfaction with the workshop you participated in” (5- Very Satisfied, 4- Somewhat Satisfied, 3- Neutral, 2- Somewhat Dissatisfied, 1- Very Dissatisfied). Respondents might also be asked how strongly they agree or disagree with given statements. Most surveys use some sort of a Likert scale in which respondents must rate how strongly they agree/disagree with an item or how helpful/not helpful some tool might be.

One factor that can be tricky is developing the actual scale. It seems as though so many people have a preference for 5, 7, and even 10 point scales. When creating a Likert scale for survey instruments, it is best to use a 7-point Likert scale because it will give you more variance than a
5-point scale. A 5-point scale means that the question provides you with 5 possible response options that usually vary from strongly agree, agree, neutral, disagree, and strongly disagree. Five-point scales will not, necessarily, display much variation in responses. This means that the average scores for each survey question will generally lie around the same range due to the limited number of response options. A 7-point scale provides more wiggle room and allows individuals to voice their different opinions more precisely. Using anything higher than a 7-point scale might be too cognitively challenging, and people cannot mentally discriminate at such a precise level. In addition, it would reduce the reliability of your results. This means that an individual will be less likely to provide similar responses if they were to take the survey multiple times. It is far too arduous to pinpoint one’s opinion at such an intricate level with such a large scale. Bipolar scales (those with a neutral midpoint, which uses opposite ends such as “strongly agree” to “strongly disagree”) perform best with seven response options. Whereas, unipolar scales (zero scale at the end, such as “not at all important”) perform best with five response options. Plus, offering a midpoint on a bipolar scale, indicating a neutral position, increases reliability. In addition, it is best to provide neutrality, as a mid-point, because this does not force respondents to choose one camp over another. Respondents would be forced to choose a side even if the stand they take is weak. It is important to allow respondents to voice impartiality if they are “on the fence” about a particular topic. Also, they might be inclined to acquiesce to positively worded statements, and a neutral option might reduce this effect.

Also, one does not want to ask too many repetitive questions. For example, it is useful to ask 5-7 questions that relate to confidence (e.g., “I feel more confident about taking a math exam”, “I feel confident when faced with a challenging math problem”, “I know that I possess the skills to complete difficult homework assignments”) but one should not rephrase the same question. It leads to survey fatigue and disinterest (e.g., “I can solve any math problem” and “I have enough knowledge to solve different kinds of math problems”). These two aforementioned statements, generally, ask the same question. Also, each item should not ask two questions at once; this is called a double-barreled statement. For example, “do you think this workshop was useful, and the location was easy to find?” A respondent might feel positively about the workshop but negatively about its location. This could not be expressed in a Likert scale item where a respondent is to select only one response.

Insofar as the pool of survey participants, random sampling is important. This asserts that each individual in the population has an equal opportunity of being selected for the survey. One also wants to ensure that there is nothing specific or unique about the pool of respondents that would skew the results. One must use a representative sample of the population, so that the generalizability of one’s results are enhanced. This means that the results can be generalized to the entire population and are not restricted to one, particular group. The number of respondents to use is contingent on the greatest amount of error that you are willing to accept in your study. Each study is prone to some error unless if you sample the entire population at large, which is, oftentimes, impossible. In order to determine how large a sample size should be, one should keep in mind the following factors: how much sampling error can be tolerated, the population size that the sample will be drawn from, how much diversity the population has
research regarding the characteristic of interest, and the amount of confidence one wishes to have in the chosen sample. The sample size also depends on whether the researcher feels that the population will be evenly split on certain issues. For example, if it is believed that 50% of the population will answer “yes” and 50% will answer “no”, a different sample size should be used than if the population were expected to have an 80/20 split. The table below exemplifies this notion and was retrieved from Don Dillman’s Mail and Internet Surveys (Dillman, 2007, p. 207).
It is important to place oneself in the shoes of the survey respondent. Keep in mind that a survey should never be too lengthy, convoluted, or use technical jargon if the audience comes from varied backgrounds. Choose 2-3 constructs and stick to those themes throughout the survey. If new issues emerge based on participant responses, they can be added to future surveys. Surveys should also ensure anonymity and confidentiality. Respondents should be told that no names will be tied to their survey, and individual responses will not be shared. The key point to remember is that evaluation research is cyclical, and results can change through time. It is essential to continue to assess programs throughout the duration of their existence.
Focus Groups (For more information on focus groups on our website: Focus Groups Article)

Focus groups are a good way to obtain feedback in an interactive setting where participants are at liberty to talk candidly. Focus groups are like face-to-face interviews, but they take place in a group setting. The purpose of a focus group is to obtain in-depth feedback and allow participants to expand on any key issues that are of concern to them. Also, participants can use each other as resources. This is more of an open dialogue forum between participants. One participant might bring up an issue, which then snowballs into another topic of discussion. Participants can also elaborate or clarify topics that are brought up by other individuals in the focus group. Common problems, concerns, and thoughts are addressed. Participants can also be probed if responses are ambiguous or unclear. The saturation or prominence of key issues can come to light in this type of setting.

Generally speaking, it is best to use approximately 10-12 participants in a focus group. Not everyone would have the opportunity to voice their opinion if a focus group has far too many participants. About 10 questions should be asked, and the session, itself, typically lasts about 45 minutes to an hour and a half. Homogeneity is desired in focus groups because obvious differences in education, authority, prestige, income, etc. can lead to hostility between participants or withdrawal. If input from different groups is needed, then it is best to hold separate focus groups for each group.

A focus group moderator should be present, so that he or she can facilitate the process. Employees or someone known to the focus group participants should never serve as a moderator. Their presence can have a large impact on the direction of the discussion. The moderator should have similar characteristics as the group in terms of age, race, gender, or ethnicity although it is not necessary. A moderator should, at least, have an understanding of the culture or lifestyle of the group. This way they can more effectively facilitate the session and understand comments better. It can also be useful to have another individual take note of the participants’ interactions and body language throughout the session.

The moderator should assure participants that feedback is confidential and that no one’s names will be tied to any of the responses provided. Participants should be reassured that they will not be penalized for negative comments because the purpose of the focus group is to improve the quality of services that are given. As such, one must point out both the strengths and weaknesses in the program in order to ensure that it is efficiently serving its intended purpose. Participants must also be warned that the moderator might interrupt them at any given time simply to assure that each person has the opportunity to speak and that all of the questions in the focus group are addressed. If time begins to run out, the moderator might also have to interrupt participants mid-sentence. This will only occur as a last resort, but participants should be cautioned that it can happen.

In addition to a moderator, a note-taker should also be present. It is preferred that the note-taker has a laptop or some other electronic device to type notes on. Handwritten comments are not as efficient and will be difficult to transcribe. It is also beneficial to audio record the focus group; in case that the note-taker misses any key points, they can easily fill in the gaps in
their notes at a later time. It is helpful to play back the audio recording once the focus group has ended in order to guarantee that each critical point was documented.

The notes can then be examined for common themes. Comments can then, most likely, be lumped into main categories. It is best to create about five categories because one does not want to make the categories far too broad nor so numerous that they become too complex to follow. Commonalities will emerge from these qualitative analyses, and the underpinnings of the program will become more readily apparent as a result. One must remember that if focus group participation is voluntary, these individuals tend to be more motivated than others, so their experiences in the program might be slightly different. Nevertheless, there are still universal strengths and weaknesses that will surface regardless of motivational levels, backgrounds, or other demographic characteristics.

Reference