

Designing your evaluation – some tips

Demographic data collection

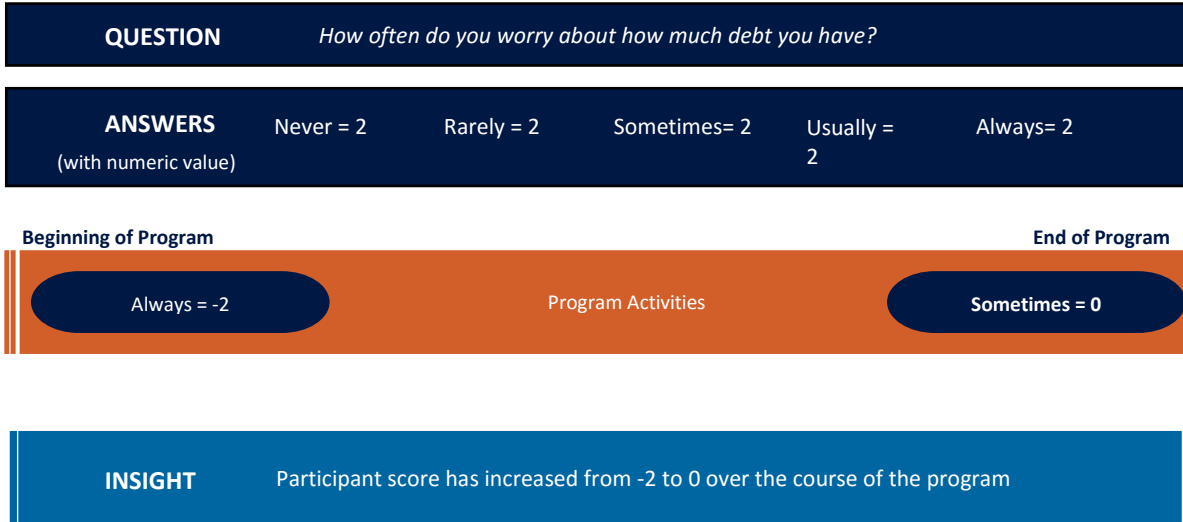
You will most likely want to collect some demographic information about your participants, such as age, gender, cultural background, or immigration status. It is up to you to decide what kind of demographic data is the most relevant to your program. However most instruments recommend that at least age and gender information be collected. Additionally, some scales require demographic data to be collected in order for the scale to be properly used. One example of this is the Consumer Financial Protection Bureau (CFPB) Financial Well-Being Scale, which requires age to be collected as part of the scoring method.

Try to limit the number of demographic questions to only those that would be helpful for understanding your program. Many participants may feel uncomfortable about giving personal information, so in general limit your survey to only the questions you can meaningfully use.

Pre/post design

A pre/post evaluation design asks participants the same set of questions at two different periods of time during the program, generally before a form of program participation has happened (“pre”), and after (“post”). The purpose of this kind of evaluation design is to be able to compare the results before and after the intervention to see how an individual’s scores have changed as a result of participating in your program. For example, by asking the same question about household spending habits related to income during “pre” and “post” surveys, you would be able to demonstrate change in household spending patterns over the course of participation in your program. This could then be reported as a measurable outcome of your program. Please note, it is important for you to consider the realistic impact of your program. For example, shorter programs (e.g., one hour sessions) may demonstrate changes in knowledge and confidence, but behaviour changes may not be achievable within the timeframe of delivery.

Once evaluation questions have been determined, in most cases a number value must then be assigned to each potential answer, to allow the responses to be measured and compared. The numeric results between the “pre” and “post” numeric values are compared and analyzed to see what kind of change has occurred. Individual scores are assigned to each personal response to each individual question, and an aggregate calculation is used to show the combined change from all program participants. The graphic below provides an example of this process. Please note, it is important to assign a unique identifier for each of your participants, in order to analyze the difference in their pre and post scores.



Results of a “pre” survey can establish a baseline set of data for your program. This is important because it gives you a score against which future results may be compared. Based on what you have determined for your program as a measure of success, you will then be able to track this level of success against the original baseline scores. For example, if your program is concerned with increasing participants’ level of savings rather than elimination of debt, and if the “post” scores showed a statistically significant increase in their savings even if their debt stayed level, these survey results would be able to demonstrate that your program has been successful.

These kinds of questions can give you valuable data about the outcomes of your program. However, they cannot determine what your program defines as success or progress – you must decide these things as a part of your program, and evaluation tools may then be used to measure the success of program outcomes you have defined.

Some questions in our registry already have a recommended scoring system. You can find the scoring instructions in the ‘scoring’ section of each individual question. If a question does not have scoring instructions, you can assign them as described below.

Scoring questions

Evaluation questions can be written in several different ways. These are:

- **Yes/No questions** (Participant can respond “yes” or “no” only)
- **Questions with multiple response options** (Participant can choose one or more specific response)

- **Open-ended, qualitative questions** (Participant responds to the question in their own words)
- **Scales** (The survey assigns a single score based on the participant's response to a series of questions delivered in a specific sequence)

Yes/No questions

Sample Question: Does your household have a budget, spending plan, or financial plan that you use to guide how your money gets spent each month?

- a) *Yes, or, We do not need one because we always have enough money to meet our needs*
- b) *No*

Design: Assign a value to each of the potential responses, with the higher number representing the more positive outcome(s). Compare this number between “pre” and “post” responses to determine whether a positive outcome has happened – i.e. whether an increase in value has occurred between the two sets of responses.

For the sample question above, this could be done as follows:

- Yes, OR, We do not need one because we always have enough money to meet our needs = 1*
- No = 0*

Questions with multiple response options

Sample question: In a typical month, how difficult is it for your household to cover your expenses and pay all your bills?

- a) *Not at all difficult*
- b) *Somewhat difficult*
- c) *Very difficult*

Design: Assign a value to each of the potential responses, with the higher number representing the more positive outcome(s). Compare this number between “pre” and “post” responses to determine whether a positive outcome has happened – i.e. whether an increase in value has occurred between the two sets of responses.

For the sample question above, this could be done as follows:

- Not at all difficult = 2*
- Somewhat difficult = 1*
- Very difficult = 0*

Open-ended, qualitative questions

Sample question: *During the last month, in your opinion, what was the most important change that took place for you during the program?*

Design: In this case, a numeric value cannot be assigned to the potential responses because it is open-ended. The possible answers will depend uniquely on the individual person and the program in which they have participated. In this example, program staff would review all answers to this question and end up with a collection of vignettes to share with funders and stakeholders. The answers allow the staff to check on the ongoing significance of the program and whether this matches the logic model of the program.

Scales

A scale is a set of evaluation questions that must stay together, in the order in which they are offered. They are groups of questions that provide you with a combined performance score as a result of a specific scoring method.

A scale is a set of questions that have been grouped together in order to collectively measure a distinct concept, such as an attitude or ability. For example, a single question may not be able to measure the concept of ‘self-confidence’. However, there may be a group of questions that, when put together, can provide one with a ‘self-confidence score’.

A scale assumes that the concept that it is measuring does not change, and it reliably measures what it is supposed to be measuring.

Example:

The CFPB Financial Well-Being Scale is an example of a scale. The scale analyzes individuals’ perceptions of their personal financial wellbeing to deliver a single score that captures four elements of financial wellbeing (control over day-to-day finances, capacity to absorb a financial shock, financial freedom to make choices to enjoy life, being on track to meet financial goals). Financial well-being is assessed from the scale as a whole. If someone were to isolate only the questions on “capacity to absorb a financial shock,” they would receive answers on the participants’ ability to absorb a financial shock successfully, but would NOT be able to confidently measure the participants’ financial well-being. This is why the scale is needed.

The CFPB Financial Well-Being Scale contains 10 questions in the full version (five questions in the abbreviated version). This scale can be used in a variety of ways, including as an initial assessment tool, an outcomes assessment tool, and to track individual progress over time.