To survey or not to survey?

Surveys require a significant programmatic investment. A survey questionnaire should:

- Have a clear objective
- Only collect data that cannot be found elsewhere
- Measure outcomes that have a reasonable chance of changing between baseline and the final survey.

Avoiding Common Pitfalls

1) Matching the method to the evaluation question: is a survey the best match?
2) Identifying content and writing questions: what exactly should I ask about? Are there questionnaires that I can use in full or in part?
3) Physical design and layout of the questionnaire.
4) Pre-testing, piloting, and translation.

Figure 1. Survey Decision Tree and Survey Alternatives
A questionnaire or survey is not the best data collection tool to:

- Get in-depth qualitative information
- Get quotations or stories
- Find out general information about the whole district or country
- Understand actual behavior rather than reported behavior
- Ask about sensitive topics

**Questionnaire development: Identifying content and writing questions**

**Identifying content**

The content of a questionnaire should include:

- Main outcomes you expect will change, and
- Factors that might influence how much people benefit from the program (e.g. literacy, gender).

Use the program theory of change to identify which content areas will answer important questions about your program. These content areas become **modules** in your questionnaire.

In the example below (Figure 2), we have a theory of change that maps out how the psychological well-being of women who have survived violence improves through participating in a Village Savings and Loans Association (VSLA). The dark gray box on the left describes the program activity (women participating in VSLAs), which causes the intermediate outcome in the black box (women's psychological well-being improves) and ultimate outcome in the yellow box (women recover and heal from sexual violence) through the mechanism described by the light gray boxes. Each of the light gray boxes is measured in some way. The purple box (women feel less stigmatized) became a stigma scale in the final version.

![Figure 2. Mapping questionnaire modules from a theory of change](image-url)
questionnaire. The green box (women’s economic functioning improves) became three modules on food consumption, household assets, and the types of work done by a woman.

Once you define the main content area of a questionnaire, choose questions for each content area. Search for existing questionnaires that have been used for high-quality research. Using existing questionnaires benefits your survey because existing research questionnaires were field-tested and should contain high-quality questions, and you will be able to compare your data to existing research.

If you have determined that you need to write your own questions from scratch, please review the following sections.

**Writing Original Questions**

A final question in a questionnaire looks like this:

<table>
<thead>
<tr>
<th>D.1.F</th>
<th>Question</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>What method do you use to irrigate the rice crop on this farm?</td>
<td>None (rain only) ................ 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hand-watering .................. 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Canals/sluices .................. 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pump/mechanized ................ 4</td>
</tr>
</tbody>
</table>

But how did we get here? To get to the final question as it looks in the questionnaire, we have made a series of decisions detailed in this section.

**Open-ended vs. closed-ended questions**

Open ended questions allow the respondent to give any answer s/he wishes, while closed ended questions provide a pre-set list of possible answers that the respondent chooses from. Closed ended questions are more common for survey questionnaires.

If your questionnaire is full of open-ended questions, this could indicate a methods mismatch. Consider a qualitative data collection method, like key informant interviews or focus groups, to collect the data.

**Phrasing questions**

Write short, precise, and clear questions. Use neutral words and present the question without any bias. Write a question that has only one answer.

*Instead of asking this:* We are very proud to run this program. How has it helped you?  
*Ask this:* Please give me some examples of how you have participated in this program.

*Instead of asking this:* Was the doctor you saw yesterday professional and helpful?  
*Ask this:* 1) Was the doctor you saw yesterday professional? 2) Was the doctor you saw yesterday helpful?

If people cannot understand what you are asking or are offended by the way you ask it, they will not give you good information.

**Developing answer categories**

The next step is to develop the answer categories. Closed ended questions have a pre-set list of possible answers or a finite answer, like a number. The possible answers should be **exhaustive** (cover all possible answers) and **mutually-exclusive** (should not overlap with one another). A respondent should be able to see their precise answer reflected in the list of possible answers. For example, if I ask someone in DR Congo how they get from place to place, possible answers could be:
Include “other,” “no response,” and “don’t know” as possible answer categories when applicable. In this case, “other” and “no response” are possible answers, but it’s not logical that someone wouldn’t know how she moves from place to place.

You should always pre-test and pilot answer categories. Pre-testing answer categories means you ask colleagues, data collection staff, program staff, and other experts in the context or content to review the answer choices.

Separating “nice to know” from “need to know” questions
“Nice to know” questions are related to the outcomes of interest or the general content area, but do not give you information that directly address the goal of the questionnaire. Getting rid of “nice to know” questions is difficult because we want to learn about so many different things; however, cutting “nice to know” questions increases the chance that we will actually use the collected data, saves time and money during survey administration, data entry and analysis, and it protects human rights.

Questionnaire length
Generally, questionnaires should not last more than 1 hour for an adult respondent and less than 45 minutes for a child. If it is necessary to conduct a long questionnaire, put a break in the middle, move a more interesting part of the questionnaire to the middle (if there is a “game” or anthropometric section), providing small incentives throughout (e.g. small snacks), or come back the next day.

Data analysis plan
A data analysis plan usually looks like a spreadsheet or table where we list each question in the questionnaire, and how we will analyze the data we receive from that question. For example, if we need to know the number of women in your program that are using contraception, should we ask question 1 or question 2?

<table>
<thead>
<tr>
<th>1. What kind of contraception are you currently using?</th>
<th>2. Are you currently using some form of contraception?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not currently using contraception .... 0</td>
<td>Yes ................. 1</td>
</tr>
<tr>
<td>Condoms.......... 1</td>
<td>No ................. 0</td>
</tr>
<tr>
<td>Birth control pills................. 2</td>
<td>I don’t know ....... 888</td>
</tr>
<tr>
<td>Rhythm method..... 3</td>
<td>No response ...... 999</td>
</tr>
<tr>
<td>Long-acting form of birth control ....... 4</td>
<td></td>
</tr>
<tr>
<td>Breastfeeding....... 5</td>
<td></td>
</tr>
<tr>
<td>I don’t know........ 888</td>
<td></td>
</tr>
<tr>
<td>No response......... 999</td>
<td></td>
</tr>
</tbody>
</table>
If all we need to know is the number of women in our program that are currently using contraception, and we cannot use information regarding the types of contraception currently in use in any way, then we should ask question #2, not question #1.

**Questionnaire design: Organization, Formatting, and Layout**

**Organizing the questionnaire**

Two rules of thumb are:
1) Logically group questions into modules and logically order modules, and
2) Start and finish with non-sensitive questions that are easy to answer.

Grouping questions into modules makes it easier for enumerators to administer the questionnaire and for participants to answer questions. In the questionnaire, add a transition sentence when moving from one module to the next. (i.e. “I have finished asking questions about your house, and now I’m going to ask you some questions about your family.”)

Starting and finishing with non-sensitive questions allows the enumerator to build rapport with the respondent and increases the probability that respondents will answer the sensitive questions fully and truthfully. Start and end with straightforward and easy questions, like demographic questions, and put difficult or more sensitive questions in the middle.

**Instructions**

Write instructions for each module, even if the instructions are repetitive, explaining the exact procedure for each module or question, any new skip patterns that enumerators should follow or new answer categories that respondents must follow (e.g. “now I will read a series of statements and ask you to tell me whether the statement is true, not true, or sometimes true”). Instructions should serve as a script for the enumerator, so that s/he never has to guess or make a judgment about how to proceed during data collection.

**Skip patterns**

When writing instructions, we should also review which questions require skip patterns. Skip patterns are instructions to the enumerator to skip one or a series of questions that become irrelevant if a particular response is given to the previous question.

For example, if a question asks if a respondent attended secondary school, and the respondent answers “no,” then there is no reason to ask questions like “Where did you attend secondary school?” “When did you attend secondary school?” and “How much was your tuition?” A skip pattern instructs the enumerator to skip these follow-up questions, which are not relevant for someone who did not attend secondary school. Skip patterns ensure that the questionnaire is streamlined and respondent time is not wasted.

Standard formatting should be used throughout a questionnaire for noting a skip pattern. These instructions should be expressed as succinctly as possible but be thoroughly reviewed and practiced in the enumerator training. When making changes to the questionnaire, review and update the skip patterns.

This example instructs the enumerator to proceed to question A.20 if the response to question A.17 is “no.” If the response is “yes,” the surveyor proceeds to the next question (A.18).

| Instructions for Enumerator: Read the following question out loud to the respondent. Circle the answer that the respondent gives. |
|---|---|---|---|
| Question Code | Question | Coded responses | Skip |
| A.17 | Is attending student in secondary school now? | Yes …………………..1 | To A.20 |
| | | No …………………..0 | |
Coding answer categories
The majority of answers on a quantitative survey are coded, meaning that a number is assigned to the response. For example:

<table>
<thead>
<tr>
<th>Question</th>
<th>Response</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is the student in secondary school now?</td>
<td>Yes</td>
<td>……..1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>……..0</td>
</tr>
<tr>
<td></td>
<td>No response</td>
<td>……..888</td>
</tr>
<tr>
<td></td>
<td>Don’t know</td>
<td>……..999</td>
</tr>
</tbody>
</table>

Coding answers makes data entry and analysis easier and quicker and reduces errors. Codes should be consistent across questions within a questionnaire. Maintain a list of codes in a codebook to keep track of all the codes used for each question as well as the dates codes were used.

Start your code book during questionnaire development and piloting in order to develop a comprehensive list of codes. Always have consistent codes for “other”, “don’t know”, and “no response.”

When creating a new code, make sure it is necessary: the answer you want to code should appear frequently and should not fit into an earlier code. If you create a new code after data collection has already started, you will need to recode answers from previous questionnaires.

Formatting the questionnaire
A well-formatted questionnaire minimizes potential enumerator and data entry errors, improves the accuracy of data, and reduces the time needed to check or clean the data before analysis.

Rules of Thumb for Questionnaire Formatting

1) Use a table format to keep the positions of questions, answers, and instructions consistent and easy-to-read.
   The example below shows a standard table format. Each question should include:
   - Instructions for the enumerator
   - Instructions for the respondent
   - The question (and a code for the question)
   - All possible answers
   - Answer codes
   - Any directions for skip patterns. A skip pattern instructs the enumerator to skip questions that are not relevant for a respondent. For example, A.18 and A.19 are about secondary school, so if this respondent answered “no” to A.17, the enumerator does not need to ask more questions about secondary school, s/he can jump immediately to A.20. (See the annex for an example of a clearly formatted questionnaire)

<table>
<thead>
<tr>
<th>Question Code</th>
<th>Question</th>
<th>Coded responses</th>
<th>Skip</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.17</td>
<td>Is attending student in secondary school now?</td>
<td>Yes</td>
<td>To A.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

2) Use consistent style choices to distinguish between parts of the questionnaire that are instructions or questions, and for the enumerator or for the respondent.
   - Put all instructions in uppercase and bold font or inside a box.
   - Different formatting helps distinguish between the various content.

3) For numeric responses, use boxes to indicate how many digits the response should be.
   - For example, if you know the response must be a three-digit number, use |__|__|__| to force the enumerator to write three digits.

4) Leave enough space for open-ended responses where needed.

5) The first page is always the informed consent page. The second page of a questionnaire collects identifying information.
• Make sure that these two pages can be detached from the rest of the questionnaire to preserve privacy and confidentiality.
• Make sure that the questionnaire can always be linked to one particular respondent by a unique identification number.
• Include a box at the top of every page for the unique identification number for the survey in case the front page becomes separated from the rest of the survey.

Pre-testing, piloting, translating

Translating
Translation can change the specific wording or meaning of certain questions, so it is important to back-translate into the original language (by a different person than the person who originally translated) to verify the translation. A third person who was involved in the questionnaire design and development compares the two translations to make sure they maintained the same questions and meaning.

If the local language is commonly read, then the questionnaire can be translated and printed in the local language. If the local language is not commonly read, then the questionnaire can be verbally translated during enumerator training and translated in real time during questionnaire administration. If necessary, a sheet of common or problem words and their phonetic translation can be drafted to assist enumerators in remembering the agreed upon translation for key terms.

Pre-testing and piloting can help improve the final translation, but back translation should be done at least once before piloting the questionnaire.

Pre-testing and Piloting
Pre-testing the questionnaire means you informally test out the questions with colleagues, collaborators or enumerators. Do the questions and answer categories make sense to them? Are the instructions (for enumerators and for participants) clear? This should be done at least during enumerator training.

Piloting your questionnaire is like a dress rehearsal for data collection. All enumerators should go through the questionnaire from start to finish with several test participants who are not part of your survey population.

Piloting can help catch problematic skip patterns, reword awkward or unclear questions, update or add to answer categories, document the time it takes to actually administer the questionnaire (hint: It often takes longer than you think!), make sure the questionnaire flows well, and catch any other mistakes that could result in poor data collection or mistakes during data entry. Piloting and pre-testing can also help you identify which enumerators are prepared for data collection. If you pilot and pre-test the questionnaire with a larger number of enumerators than you will ultimately hire, you can use this process to identify the best enumerators.