Center for Effective Global Action
University of California, Berkeley

Module 3.1: Questionnaire Design

CONTENTS

1. INTRODUCTION

Up through Module 2.6, this course has concentrated on theoretical and practical considerations regarding Impact Evaluation techniques. Modules 3.1 through 3.4 will refer to a different yet related issue: survey design, field work and other field research methods. In this module, we will be focusing on the proper design of questionnaires and other survey instruments. Given these are often the main instruments to collect the data to be used in an impact evaluation, the quality of the survey and the clarity of the specific questions asked will dictate our ability to truly understand and accurately estimate the effects that our program or intervention are having on the population of interest.

Indeed, it is clear that the main objective of any Impact Evaluation is to assess the results of a particular intervention, and researchers will hopefully be confident that the intervention could be repeated in some other context. Good quality data and analysis are thus required to establish the magnitude of the impact, allowing researchers to extend their findings to these other settings. Data quality relies on the information provided by a survey.

Generating good quality data from a survey is not an easy task, because survey quality depends on many factors that, while unrelated in appearance, in fact heavily dependent on each other. This is why survey design and implementation should be the conceived under the overarching framework of *Total Survey Quality*. This framework demands the integration of Sampling Design, Instrument Design, Fieldwork, and Data Management.

Questionnaire design is a very sensitive matter in survey development for two reasons. In the first place, it is one of the most cost-effective ways of improving the quality of survey data (Fowler, 1995). Secondly, it has a direct impact on the analytical stage of the survey: decisions on how to collect the data will define what can be assessed and, maybe even most importantly, what will be left out of the analysis.

In spite of its fundamental role in surveys, questionnaire design is often disregarded as a less-important aspect of RCTs. Contrary to popular wisdom, however, ensuring a successful questionnaire design is not always easy. Indeed, it is a complex process the potential length of which should not be underestimated.

In what follows, we present some basic guidelines for achieving a successful questionnaire. The module is organized as follows. Section 2 stresses the importance of questionnaire design. The following section concentrates on the three fundamental aspects of questionnaire design: contents, drafting and testing. The last two sections refer to other survey instruments (such as manuals) and to considerations regarding paperless interviews.

Although there are many heuristics and procedures related to survey design, we have tried to concentrate on the most important ones according to our experience on the field. For further reading, please refer to the readings listed in the last section.

2. POPULAR WISDOM REGARDING QUESTIONNAIRE DESIGN

Questionnaire design seldom occupies the central place it deserves in Impact Evaluations. Common wisdom usually identifies questionnaire design as an operative task. However, it involves a very complex, long and iterative process in which the final outcome –the questionnaire—represents the convergence of hundreds of decisions made along the way.

One way of clearly undermining this 'popular wisdom' approach to questionnaire design is by comparing it to another relevant process of survey design: power calculations, which were covered in Module 2.6. Indeed, in that module you learned how to ensure that the size of the sample chosen was large enough so as to detect sufficiently-large effects of an intervention. The following chart shows some of the differences between the two processes:

Popular Wisdom: Power Calculations vs. Questionnaire Design

Power Calculations	*Questionnaire Design*
✓ *Supposed to be complex*	✓ *Supposed to be easy*
✓ *Theory-based*	✓ *No Theory*
✓ *Objective Evaluation*	✓ *Evaluation Unfeasible at Times*
✓ *Specialists*	✓ *Many Actors*

In the first place, power calculations have a reputation for being a complex activity, which is better left to specialists, whereas questionnaire design appears at the eyes of beginners as a simple task, which can be carried out using common sense. This in part derives from the fact that power calculations rely heavily on theory and formulas, while the theory of questionnaire design appears almost inexistent or badly-documented. In fact, such theory has been developed for decades based on experience and experiments, but in spite of the significant progress in cognitive psychology, it does not constitute the kind of formalized theoretical framework that has been presented in the last seven modules. In fact, almost all of the 'rules' that apply to questionnaire design have exceptions, and in many cases even the formulators of such rules sometimes opt to not follow them.

The clear-cut theoretical structure that permits the evaluation of an impact evaluation's sampling and experimental designs is largely missing from the design of questionnaires: it isn't always clear whether questionnaire design has been properly conducted. Finally, the 'black box' characteristics of power calculations imply that few actors will be involved in the process – only specialists – while questionnaire design is essentially a lengthy iterative process in which many actors – the implementing agency, financiers, the academic community, and many other stakeholders–may have a say. Furthermore, the interests of all these actors do not necessarily converge.

> *"Those who have never had to analyze data from a questionnaire they have developed themselves may think that designing a questionnaire is easy. It is not."*
> *(Grosh & Muñoz, 1996)*

In summary, contrary to what it may seem, Questionnaire Design may prove to be much more difficult than what is commonly assumed. The lack of well-documented theory combined with unfeasibility of evaluation and the involvement of many actors makes this process very complex; it ought to be conducted by the higher levels of the survey team.

3. THE THREE PHASES OF QUESTIONNAIRE DESIGN

Questionnaire design must ensure three key traits of the final product: relevance, accuracy and feasibility (Iarossi, 2005). The quality of the data will ultimately rely on the quality of the questionnaire. This requires that the survey designer is intimately familiar with the aim of the survey and thus the information that needs to be collected; that the information is collected in a reliable and valid manner; and that the questionnaire motivates the respondent to provide an answer. How is this achieved? By *(i)* involving the right kinds of people in the questionnaire's design; *(ii)* allowing enough time and repeated iterations during the process; and *(iii)* conducting careful testing of the questionnaire (Grosh & Muñoz, 1996).

Ensuring that all of these conditions are met involves a process with three related, yet independent, activities: defining the questionnaire's contents, drafting the questions, and pre-testing the draft questionnaire to ensure that it works. All three play an important role in determining the quality of the information, which will be collected in the field. In what follows, we describe each of these phases and provide some practical guidance.

3.1 Content

Defining the content of a survey is the first step in questionnaire design. The task is bounded by the context of Impact Evaluation, since the content will (or should) be essentially determined as measures of the chosen impact indicators. Nevertheless, the success of the survey will heavily rely on the adequate definitions adopted in at least five fronts:

- ✓ *Effects.* The most important task is to sharply clarify which are the expected results of the intervention. This implies adequately defining which outcomes need to be measured.
  For instance, is the aim of the intervention to reduce malnutrition, or to enhance teacher's knowledge on sanitary conditions, or to improve water quality? Perhaps all three are aims of the intervention? Once this conceptual building block is clear-cut, the following issues have to be considered.
- ✓ *Observation Units.* A natural next step will be to define which impacted entities should be observed. Following our previous example: Is it children? Teachers? Water sources?
- ✓ *Correlation with Other Factors.* In order to correctly measure the effects of the intervention, it will be necessary to measure the correlation between the measured impact and other factors. The most obvious factor to be registered should be whether the observational unit was in the treatment or control group (and, if there are multiple treatment groups, identifying the group she

was part of). If this data is not collected, the exact impact of the treatment may not be revealed in the survey data. Furthermore, in many cases it will also be interesting to assess whether other factors—such as geographic location, income levels, or weather conditions—may be partly responsible for the impact. If this is the case, they should also be part of the questionnaire. The implications of including such correlates are far from trivial: for instance, measuring income can be a somewhat arduous process.

- ✓ *Hard-Core Indicators.* Translating the conceptual definition of the expected effects, the observational units, and the correlated factors into explicit indicators that can be used to measure the impact of the intervention will be the core of the questionnaire content. Keeping with our previous example, possible indicators could be: Height, Age, Standardized Test Scores; Microbial Content, Region, Season, etc.
- ✓ *Other Measures.* The survey's mechanics may impose other measures that may need to be taken into account when defining the content of the questionnaire. These include, among others, household and individual identifiers, survey conditions, field-workers' identification codes, and perhaps many others. Recording sufficient information to properly track the observed entities in follow-up surveys is particularly important when the impact evaluation relies on panel data, which is often the case.

Being successful in adopting adequate definitions on each of these five fronts will involve a series of iterative actions that go beyond choosing a set of questions. In the first place, extensive discussions will have to be held among different actors of the survey: the project implementation unit, the impact evaluation team, and the survey experts. It will also require some examination of the pre-existing literature as well as assessing other questionnaires to check for comparability issues.

Qualitative studies may represent a valuable input in defining the content of the questionnaire. While qualitative studies should not be considered a competitive alternative to quantitative impact evaluations, they may be introduced as complementary to questionnaire design, becoming an extremely valuable input in terms of detecting questions, issues and concerns that could otherwise go without notice.

Given the iterative and multi-party characteristics of this process, it is of fundamental importance that it is managed by a leader of the impact evaluation. He or she should be in charge of coordinating the process as a whole and mediating among opposing interests.

The leader should have a final word in establishing the questionnaire contents. Otherwise it is possible to end up with an extremely long questionnaire that, in satisfying each of the different actors, fails to accomplish the essential task of becoming a central piece of the impact evaluation.

The end product of this phase of the questionnaire design – i.e., defining its content- is a list of questions, which is still a long process away from being a complete questionnaire. The complex mechanics of converting the list of questions that emerged from the content design into a questionnaire is called 'drafting' and is developed in the following section.

Real World Example – Content Definition

A survey in Panama aimed at analyzing the impact of attending pre-school on educational attainment by comparing the number of years of education between those who had attended pre-school and those who had not. The problem is that the survey only asked two standard education questions: (1) the current level and grade of those currently attending school, and (2) the number of years of education for those currently not attending. A third an essential question was missing for the analysis: "did you attend pre-school?" Without this information, analysts only knew pre-school attendance for those of pre-school age, but not for the rest of the population.

3.2 Drafting

As stated before, drafting refers to the process of giving concrete form to the questions identified in the content design phase. In particular, it requires making decisions about how to elicit the necessary information from study participants.

This should be carried out with maximum care, since the accuracy of the data (and thus the possibility of adequately assessing the impact of the intervention) heavily relies on its success. For this reason, drafting should be carried out by the survey planners and not relegated to lower-level or clerical staff. The focus of drafting is to ensure that differences in answers only reflect actual differences in peoples' behaviors and beliefs rather than reflecting participants' different interpretations of the same question.

In what follows, we list a series of recommendations that should guide the drafting process. It is important to bear in mind that this is not a complete set of guidelines, but rather a selection of the most important guidelines that should be born in mind. For further reference, we recommend Iarossi (2006) and Lohr (1999).

3.2.1. Wording

Defining the way the question is formulated is key in the drafting process. A few guiding rules follow:

- ✓ *Keep it simple.* Avoid technical jargon and academic language or unfamiliar words incomprehensible to the respondent. Try to keep the question as brief as possible: long questions become complex and confusing for both the respondent and the interviewer.

  Avoid Double Negative Questions. Double negatives are not only confusing, but may also prime the respondent towards certain answers

  Avoid Double-barreled questions. Each question should ask about one **and only one** issue

Adopt the same definitions throughout the form. Avoid the use of different terms with the same meaning, as well as using interchangeably different definitions.

Example 1 – Wording: simplicity is key

Complex Question:	***Better:***
"What is the source of energy most frequently used by the members of this household for the preparation of meals?"	*"What fuel do you use for cooking?"*
Double Negative Question:	***Better:***
"Why didn't you not stop going to the doctor regularly?"	*"What did you stop going to the doctor regularly?"*
Double Barreled:	***Better: split the question in two***
"Are you satisfied with your pay and job conditions?"	*"Are you satisfied with your pay?"* *"Are you satisfied with your working conditions?*

✓ *Respect the Cultural Context.* Both the reliability and validity of concepts depend upon cultural context. This not only applies to differences across regions and countries but also among different groups within the same society. The same words –even very simple ones, as shown in the following example – may mean different things to different people. Be sure to use words and concepts, which mean the same thing to the respondents as they do to the survey designers.

Example 2 – Cultural Context: Same Words, Different Meanings

Question: *"Do you own a car?"*

Answer: *Depends on how the respondent interprets:*

"you" (the individual or the household?)

"own" (what if buying on credit?)

"car" (does a pickup truck count?)

- ✓ *Be objective.* Certain types of questions should be avoided:

  Leading questions. Take care to draft questions in such a way that do not influence interviewees' responses through their wording, content, or structure.

  Loaded Questions. Avoid emotionally charged questions that could push the respondent towards certain responses through loaded words, stereotypes or images (for instance: "fair", "honest", "colonialism", etc.).

  Built-in Assumptions. Questions should not assume that the respondent has familiarity with a complex topic (for instance, immigration laws)

Example 4 – Be Objective

Leading:	***Better:***
"How far away is the health facility you attend?"	*"What is the distance (in kilometers) to the health facility you attend?"*

- ✓ *Be Specific.* Avoid terms which are vague, ambiguous, too academic, or which the respondent might misunderstand. Use simple words and phrases that can be identically interpreted across respondents.

  Avoiding ambiguity in wording is especially important when the question involves reference periods and units of measurement. It is extremely important that explicit time-spans and units are used. Also **avoid overly general questions**.

Example 3 – Be Specific: Precision is Essential

Ambiguous:	***Better:***
"When did you move into this dwelling?" *Answer 1: "in 1999"* *Answer 2: "10 years ago"* *Answer 3: "After I finished school"* *Answer 4: "In the summer"* *Answer 5: "When we bought it"*	*"In what year did your household move into this dwelling?"*
Confusing:	***Better:***
"How much wine do you consume?" *Per month? Per week? Yesterday?* *Bottles? Glasses? Ounces?*	*"How many glasses of wine did you drink in the past week? "*
General:	***Better:***
"Did the household suffer any negative shock in the last six months?"	*"Did the household suffer a negative shock in the last six months, such as: a) flood; b) death of any member; or c) household head fired from job?*

3.2.2 Organizing

The order in which questions are asked on a questionnaire is also important. The flow affects the time it takes to complete the questionnaire, the ability of the interviewer to engage the interest of the respondent, and thus the quality of the information collected. The questionnaire should flow smoothly from one question to the next, in a logical sequence that helps the respondent remember his or her activities or think about which other household member might be able to help. Here, we list several issues that should be considered when organizing the questionnaire:

- ✓ *Nature of statistical entities.* If possible, all questions regarding the same statistical unit should be grouped together (the units can be households, household members (e.g. all of them, children, women 15-49), farms, crops, fields, consumption items, and more).

- ✓ *Respondents.* The survey designer must decide who should answer each section, or each group of questions within a section; this needs to be made explicit to the interviewers in the questionnaire itself, in the field manual, and also during the field staff's training. Filters and jumps should be used as a guide for the interviewer to exclude respondents from questions or sections that do not apply to them.

- ✓ *Reference periods.* Questions with a common reference period should be grouped together (if possible) so as to prevent the respondent from getting confused (reference periods often used are "now", "yesterday", "the past 7 days", "the past month", "the past 3 months", and "the past 12 months").

- ✓ *Logical order of topics.* Questions should be grouped according to the topic to which they refer. The flow of these sections should be logical. Abrupt changes in topic can create confusion and may frustrate the respondent.

- ✓ Instructions. Use a consistent case convention to distinguish what the interviewer reads from the instructions for the interviewer. Instructions are key for both questionnaire administration and the collection of accurate data. It is good practice to use capital letters for instructions and regular font for questions that are to be read. Instructions should be clearly placed at the beginning of each section or directly above the question to which they refer. You should avoid including instructions only at the beginning of the questionnaire or in manuals.

- ✓ *Skips.* One of the more relevant instructions for the interviewer is the skipping pattern. There are two ways of demarking this pattern: establishing a skip code convention (e.g., use arrows "►" pointing to the next question to be asked depending on a previous answer or other circumstances), or including explicit instructions in the question (e.g. "ONLY FOR WOMEN AGED 14-49"). Though the former is generally preferable, all that's important is that it's clear who should answer each question (and who shouldn't!). Explicit skip codes present several advantages:
    - interviewers do not have to make decisions;
    - interviewers will not to need to remember complicated rules that might or might not have been provided in the manual or during training;
    - it ensures that instructions will be followed uniformly; and
    - irrelevant questions will be asked with less frequency (avoiding respondent irritation, wasting interview time and confusing analysis).

- ✓ *Open-ended Questions.* For open-ended questions (occupations, consumption products and so forth), the list of codes to be used for coding answers should be prepared in advance, taking into consideration the needs of the analysts and the expected frequency of answers.

- ✓ *Recall and embarrassing topics.* The order of questions may be a powerful tool to aid individual's memory and/or to gradually introduce embarrassing questions.

- ✓ *Sensitive questions.* Leave sensitive questions for the last sections; otherwise, if the interviewee feels uncomfortable about answering them, this might affect his or her attitude towards the whole survey.

3.2.3 Formatting

The physical appearance of the questionnaire will also contribute to the successful development of each interview. A good format minimizes potential interviewer and data entry errors, improving data quality, and hastens the availability of the data. Because the format of the questionnaire is so closely linked to the data-entry program and the way the database will be organized, it is highly advisable to involve the survey's Data Manager in its drafting process. A few guidelines are proposed in what follows:

> *"Formatting a questionnaire is a complex art and proper formatting is critical to survey success. It must be done by the survey planners and not relegated to lower-level staff".*
>
> *(Grosh & Muñoz, 1996)*

- ✓ Identification. It is extremely important to assign unique identification numbers on each questionnaire. Both the interviewer and the household should also be clearly identified. Each section of the questionnaire should be given a title and should be identified with a unique number or letter.

- ✓ Numbering. Questions should be numbered sequentially, without any repetition. It is good practice to use progressive numbers even across sections. This is of utmost relevant on many fronts: helping interviewers carry out organized interviews; facilitating the data entry process; and maintaining consistency in referring unambiguously to the questions (among the survey team, in questionnaire manuals, and during training).

  Consistent question numbering will also facilitate the analysts' work if it is also adopted for labeling the corresponding variables in the survey datasets. (A word of caution may be useful here: as tempting as it may be to use mnemonic data labels (such as "age" and "gender") in the survey datasets, this should be avoided in general– mnemonics only work when the database is very small, but they soon become uncontrollable and confusing. It is better to simply replicate the question-numbering convention established in the questionnaire.)

- ✓ Layout. Space between questions should not be saved by compromising the questionnaire administration. Although it may appear to be economically convenient, this may jeopardize data accuracy.

- ✓ Fonts and formats. Different forms and formats (bold letters, underlying) may be used to ensure that the same words are emphasized each time the question is asked, thereby inducing uniform interpretation by respondents. This may be also useful when the time reference changes between questions. Typographical conventions –typically upper- and lower-case letters—are also used to differentiate:
    - questions which the interviewer is supposed to read aloud;
    - alternative responses which should be read aloud;
    - alternative responses which should *not* be read aloud; and
    - instructions for the interviewer.

- ✓ Symbols. Using different symbols (circles, arrows, boxes and even color shades) is a useful visual tool to guide the interviewer through the questionnaire administration, as well as to facilitate the task of supervisors and survey operators. It helps identify particular types of answers and questionnaire flow. For instance, skip patterns may be signaled with an arrow, while boxes may be used when the answer requires inserting a number.

- ✓ Types of Questions. The questionnaire will certainly include different types of questions: quantitative, qualitative (aka "categorical") and typographic (aka "alphabetic", or "textual"). This matters in terms of physically organizing the questionnaire.

- ✓ Lists vs. Grids. Two main resources are used to physically lay out different types of questions: lists and grids. The basic rule for choosing between these options is the following: when the information is recoded for only one observational unit within the household, a list should be used, while, if more than one observational unit is involved, grids typically prove much more convenient.

  It should be noted that grids can always be transformed into lists, through the inclusion of one additional list for each observation within the household for which the information should be recoded. The drawback of this procedure is that additional pages must be added to *all* questionnaires, and the number of pages to be added should cover the maximum size of observations that could be found in a given household. For instance, in the example shown below, if the survey designer chooses a list over a grid, one page per each household member should be included, each of them replicating the list that is shown.

  Grids, then, often imply higher costs in printing and in carrying questionnaires around. This option, however, could be preferred in case the questions require a certain degree of intimate context for the respondent to feel comfortable in answering. The converse –i.e. transforming lists into grids- never occurs.

> ***Tip***
>
> *If the grid spans several pages, split the lines in groups of three, to reduce the risk of the interviewer switching lines between pages.*

Example 5 – List vs. Grid

List

SECTION 6: HEALTH
ALL HOUSEHOLD MEMBERS

6.00 ID CODE

6.01 ID CODE OF THE PERSON PROVIDING THE INFORMATION

6.02 Do you suffer from any continuous, medically diagnosed disability that is expected to continue for six month or more?
YES 1
NO 2 ▶ 6.05

6.03 What kind of disabilty is this?
OLDEST DISABILITY
ARTHRITIS 1
ASTHMA 2
CANCER 3
COPD 4
DIABETES 5
OTHER 6

6.04 How many years ago did you become disabled? OLDEST DISABILITY YEARS

6.05 APART FROM THIS CHRONIC DISABILITY Were you ill or sick in the past 28 days?
YES 1
NO 2 ▶ 6.23

Grid

SECTION 6: HEALTH
ALL HOUSEHOLD MEMBERS

ID CODE	6.01 WRITE THE ID CODE OF THE PERSON PROVIDING THIS INFORMATION	6.02 Do you suffer from any continuous, medically diagnosed disability that is expected to continue for six month or more?	6.03 What kind of disabilty is this? OLDEST DISABILITY	6.04 How many years ago did you become disabled? OLDEST DISABILITY	6.05 APART FROM THIS CHRONIC DISABILITY Were you ill or sick in the past 28 days?
	ID CODE OF RESPONDENT	YES 1 NO 2 ▶ 6.05	ARTHRITIS 1 ASTHMA 2 CANCER 3 COPD 4 DIABETES 5 OTHER 6	YEARS	YES 1 NO 2 ▶ 6.23
01					
02					
03					
04					
05					
06					
07					
08					
09					
10					
11					
12					

✓ *"Doesn´t Know" / "Not Applicable" / "Other" Categories.* In many questionnaires these are categories included as possible answers to multiple-choice questions. It is important to bear in mind that they mean very different things, and that they should generally be avoided or made more explicit.

The "doesn't know" category should be discouraged as much as possible. Explicitly including it as an alternative may be seen by the interviewer as an invitation to not thoroughly investigate the respondent's knowledge; in general, even an approximate answer will be preferred by analysts to a "doesn´t know" answer. That said, it is still possible that the respondent actually *does not know* the answer to a question. One way to get around this issue is to include the "do not know" response but to make it costly to select.

For instance, the interviewer could be instructed to write "doesn't know" in red, or in capital letters, and he/she should be warned that the supervisor will diligently check this type of response. In this way, the option –which is inevitable- is still included but discouraged.

The "not applicable" category should never be necessary in a well-designed questionnaire. If the questionnaire contains explicit instructions and skipping patterns, this category should never be needed by the interviewer.

The "other" category should be offered as infrequently as possible. Even though, as with the "doesn't know" category, this alternative is often unavoidable, it should be rarely used. Analysts should not be expected to find more than 5 percent of answers concentrated in this category (and even 5 percent is a generous figure). Larger percentages may reveal one of two failures (or both): either the questionnaire has not been well designed (specifically, the pilot survey did not serve its ends properly; see Section C on testing) or fieldwork quality controls were not properly implemented (the interviewer is failing to thoroughly enquire the respondent – see Module 3.3). As in the "doesn't know" case, one way to ameliorate this issue is to make the option costly, for instance by compelling the interviewer to write down the precise content of the "other" response. This can be done by adding a "specify" instruction and leaving room for the interviewer to complete it.

✓ *Multiple answers.* Many categorical variables may in principle be considered multiple alternative answers (i.e. more than one category may be marked). Deciding whether to allow for multiple answers is difficult but relevant, since it determines how the analyst will be able to use that information.

The risk of multiple answers is that they implicitly add a certain degree of subjectivity to the interviewing process: the number of alternatives chosen will depend upon the assiduousness of the interviewer and the loquaciousness of the respondent. These are not of issue in the case of unique answers because they force both parties to record a single response. For this reason, while multiple answers may occasionally be needed, they should only be deployed when necessary and after careful consideration.

> ***Tip***
>
> *Record multiple answers only if multiplicity is an essential analytical need. Otherwise, you can always avoid multiplicity by adequately wording the question.*

It is important to remember that multiple-answer questions can *always* be reduced to non-multiple answer questions simply by way of an adequate wording. Adding either an adjective or an adverb to the question will often do the trick: "What sport do you practice?" may be easily transformed into the unique-answer questions "What was the last sport you practiced?" or "What sport do you mainly practice?"

Example 6 – Multiple Answers

6.01	6.02	6.03						6.04	6.05
WRITE THE ID CODE OF THE PERSON PROVIDING THIS INFORMATION	Do you suffer from any continuous, medically diagnosed disability that is expected to continue for six month or more?	What kind of disabilty is this? CROSS ALL THAT APPLIES						How many years ago did you become disabled?	APART FROM THIS CHRONIC DISABILITY Were you ill or sick in the past 28 days?
		ARTHRITIS	ASTHMA	CANCER	COPD	DIABETES	OTHER	OLDEST DISABILITY	YES 1
ID CODE OF RESPONDENT	YES 1								NO 2
	NO 2 ► 6.05							YEARS	► 6.23

This all having been said, if a multiple-answer question is included in a questionnaire, some issues should be considered. In the first place, how are the responses going to be recorded? Two main options are used: either to write a cross under each of the answers chosen (as in Example Six) or coding the possible answers and recording "1st reason"/"2nd reason"/… Secondly, the question's analytical implications should be born in mind. Unique answers to categorical questions result in variables that are easily analyzed by way of frequency distributions, but the analysis of multiple answers is not always trivial. Rather than having a single variable, analysts are left with a certain number of dummy variables, which results in the tricky issue of not having cumulative percentages adding up to 100%.

- ✓ *Proxy Answers.* It is also important to be explicit with interviewers about whether proxy answering (i.e. when one member of the household responds on another's behalf) is permitted and/or recorded. Proxy answers are usually less valuable and less desirable than answers from the target respondent, since proxies tend to know less about the actual respondents than the respondents themselves. Bardasi et al. (2010) provide compelling evidence on this issue. They carry out an experiment in Tanzania to assess the role of questionnaire design and proxy answering on labor statistics estimated from the data by conducting a survey in which households are randomly assigned to different treatments (specifically, whether the respondent is able to self-report or whether he/she answers on behalf of someone else). Striking results arise from their experiment: allowing for proxy responses reduces measured Labor Force Participation Rates for both men and women by 13 percent and 8 percent respectively. The age distance between the proxy and the respondent seems to play a central role in these differences, while gender and educational differences appear to be less relevant.

It is clear, then, that avoiding proxy answers should be encouraged. In case proxy responses cannot be avoided, it is important to record the characteristics of the proxy respondent. This record may serve many purposes: as a way of discouraging the interviewer to accept proxy respondents; to monitor the interviewer's assiduousness (it may be used as a quality indicator – see Module 3.3 on Data Quality Assurance for further detail); and perhaps for analytical ends.

- ✓ *Precoding.* Pre-code as much as possible, ideally using numeric codes, in order to ease the transition from data-collection to analysis. Precoding requires that choices be clear, simple, and mutually exclusive; that they exhaust all likely answers; and that categories will not contain too-few respondents to be meaningful. A standard technique to ensure that codes include all possible answers is to add an "Other (specify)" category. This however, entails some risks (as we discussed above).

3.2.4. Other Issues

- ✓ *Recording Consumption.* In previous sections we stated that when designing the content of the questionnaire, it is often necessary to record information regarding factors that may correlate with the outcomes of the experiment. One particularly important such factor is consumption, since the bottom line of many interventions is to improve the living conditions of the treated group, and consumption is the international standard for measuring welfare. Recording consumption is not easy, quick, or cheap. Here we provide some basic guidelines that should be followed for a successful consumption record. We divide consumption into food consumption and other consumption. Let's begin with the former.

- ✓ *Food consumption***.** Food consumption is not the same as expenditure on food. Food can often be purchased at a market, but many households themselves cultivate or raise some of the food they eat. Food can sometimes be obtained for free (from food aid programs, for instance) or as payment for work. Measuring food consumption thus implies an effort to directly record food quantities. There are two main methods for recording quantities: diaries and recall.
    - ***Diaries*****.** Each day over a certain period (minimum 7 days), quantities of everything eaten and drunk in the household are recorded. Either the interviewer weighs the food to establish the quantities, or household members make a daily note of what has been consumed, or the interviewer records the quantities based on information given by the household members.

      Most survey practitioners agree that the diary method, when it is properly applied, provides the most accurate available measure of food consumption. One disadvantage is that participating households must be visited often – ideally every day – which is expensive. Even if households are recording their own consumption, experience has shown that if they are not visited frequently, they stop recording.

- **Recall.** An alternative to the diary method is for the interviewer to ask the household to recall the quantities of foods consumed in the days preceding the interview. Seven days is the standard period, although reference periods of 14 days or even 12 months can sometimes be used for infrequently-consumed types of food. The interviewer reads food items from a list and records the quantities consumed on the questionnaire. There are two types of measurement error associated with this method:
    - the respondent might include items which were in fact consumed before the start of the recall period (this is called 'telescoping')
    - some items may be not included as the respondent does not remember they were consumed in the reference period (the 'memory effect')

✓ *Valuing food consumption***.** Unit prices are needed to value food consumption. There are three ways to find out how much food costs, each of which has advantages and disadvantages:

- Ask the households how much they paid for the food (or, in the case of self-consumption or food received without payment, how much they would have paid).
- Ask the households to record their food purchases in a separate diary
- Conduct a price survey in the local market

Units of measurement: Unless the interviewer is weighing the household's food, the question arises of how to deal with local units of weight or volume. A good solution is to allow the household to refer to the weight or volume in whichever units they prefer, but to instruct the interviewer to convert these into a standard unit for that food item. In this way, the respondent might say that the household consumed 'two bags of onions', but the interviewer will record the equivalent weight in kilograms. Ideally, only kilograms and liters should be used, apart from eggs (which are recorded as individual units). If this option is followed, the best approach is for the interviewer to carry a set of scales with which to weigh the food items in different households. This implies more training, but facilitates the analytical phase by providing standard measures.

Dealing with small quantities: For some items it is difficult to be precise, especially about items used in small quantities, such as salt, pepper or oil. Housekeepers tend to declare having used "a bit of salt", "some oil" and so on. Both the diary and recall approaches are vulnerable to these inaccuracies. The best recommendation is not to register those items' consumption but only theiir expenditure, since these items are generally purchased, rather than self-produced, and their nutritional value is less important than those of other items.

Measuring non-food consumption: Unlike foodstuffs, it is generally assumed that the purchase price of non-food items is equivalent to their consumption. In currently-conducted surveys, the recall method is always used with non-food consumption, and only the financial cost needs to be recorded (some survey practitioners also record the quantity of various items purchased, not so much because of its analytic interest but as a control variable). Generally, the questionnaire is divided into a few sections, one for each of several recall periods.

A common choice is to use the following three periods:

- one month (e.g. housing, education, transportation)
- three months (e.g. footwear, clothing, recreational services).
- one year (e.g. maintenance of dwelling, expenditure on furniture).

Items are sorted into these periods on the basis of the frequency with which they are normally bought in the survey region.

Some practical considerations: Breaking down records of consumption to a satisfactory level for analysis brings with it a range of practical problems, which affect the survey's cost. This includes deciding how many items will be included on the consumption list, how to measure expenditures on health and education, how to measure housing costs, and several other issues. For further reference on these issues, see Muñoz (2006).

✓ *Paneling.* It is very frequent for impact evaluation surveys to be carried out in the form of panels in order to follow the outcome of the intervention. A very harmful disadvantage of panels is ***attrition***, which is the practical inability to locate all households in each round of the survey. **As discussed in previous modules, attrition is very undesirable for measuring the true impact of an intervention.** Attrition is partly due to inflexible realities (such as households relocating or refusing to collaborate after one or two surveys) and partly due to the practical difficulties of managing a panel (difficulty keeping a core staff for many years, bad documentation, lost files, etc.). The second variety of problem can and should be explicitly addressed by researchers conducting panel analysis. Some practical advice may help reduce these avoidable failures:

- Record and digitize. As much tracking information as possible should be recorded and entered in a database during the baseline survey (GPS coordinates, individual telephones, relatives and neighbors, etc.) Even though this information lacks analytical value, it is a very cheap and useful way to keep track of households.
- Record cell-phone numbers of all household members. This serves both for follow-up interviews as well as for supervision within the same survey.

Another important issue related to paneling is how to design follow-up questionnaires to account for changes in the household's composition since the baseline survey. Indeed, it is typical for some member to have left the dwelling, babies may have been born, a new spouse may have been added, or a member may have died. At times, there may be a new household head, and this will change all recorded relationships (since they are usually recorded relative to the head). These possible situations must be taken into account when designing follow-up questionnaires.

✓ *Translation.* Even in countries with a single language, translating questionnaires may be necessary to interact with international partners at the design and analysis phases of the survey. However, translation becomes essential in countries where more than one language is spoken (Guarani in Paraguay, Quechua and Aimara in Bolivia, etc.). It has been demonstrated that interviewer errors are up to four times higher with oral field interpretations than with written translations of the questionnaire.

✓ Translation is not easy, quick, or cheap, but it is essential to recording accurate data in many countries. In what follows, some practical advice regarding translation:

- *Translator.* It is important to be very careful about who is in charge of translation. Particularly in countries where these languages are politicized, it is important to refrain from trusting the task to academic language departments. Grammar rules and vocabulary are often developed that do not necessarily reflect the actual language used by the individuals that will be interviewed.
- The best option is to assign the task to a statistician who is both familiar with surveys and with the languages of interest.

> ***Tip***
>
> *Questionnaires should be worded in simple terms used in the language as commonly spoken rather than in academic or formal language.*

- *Back-translation.* An advisable practice when translating questionnaires is to translate the questionnaire from the original language into the second language and then back to the original, after which the two versions in the original language are compared. The back-translation should be done by someone who was not closely involved in the development of the questionnaire to avoid contaminating the back-translation with prior knowledge. Do not assume that differences between the original version and the back-translation imply that the translated version which is wrong; also check the questionnaire in the main language in case there is an ambiguity which needs to be corrected.
- *If opting for no translation.* If researchers decide not to translate the questionnaire into every local language, it is essential that the interviewer's manual provides full and complete explanations of each of the questions in the questionnaire, and that these are also covered thoroughly during the training of the field staff. The interviewer will only be able to translate the questions accurately during the interview if she completely understands the nature of the questions. Local language versions of the questionnaire should be integrated into the manuals before the field test has been completed, so that they can be referred to during training.

Real World Example – Questionnaire Translation

For the 2007 Timor-Leste LSMS, the questionnaire was developed concurrently in Tetum, Portuguese and English. The 2007 Iraq IHSES saw concurrent development of the questionnaire in Arabic, English and Kurdish. The problem for the survey designers was how to keep the physical format of the questionnaire the same in the different language versions.

First, a single dictionary was created which consisted of all the questions, options and other text in the questionnaire in all three languages. Next, the questionnaire form was purged of its text, leaving a blank form with everything except text. Because the questionnaires had been drafted in Excel, it was then a relatively straightforward matter to program an Excel macro which slotted the appropriate text in the right language into the spaces on the questionnaire form. (The macro even turned the questionnaire right-to-left or left-to-right depending on the language)

- ✓ *Software issues.* Drafting a questionnaire requires a tool which will number questions and skips in an intelligent way, explicitly identify the statistical units observed, and insert control sub-totals. Although the perfect software for this purpose does not exist, we recommend using Microsoft Excel instead of Microsoft Word of desktop publishing programs. There are several reasons for this:
    - Excel is much more structured than word processors, and thus gives the drafter much better control over questionnaire design. Its cell structure makes it easy to insert and delete rows and columns.
    - Excel greatly facilitates inserting and extracting questions, a very common practice during the design stage. Renumbering questions becomes trivial when using formulas.
    - Excel macros facilitate the administration of the translation process. Instead of running the risk of actually translating the questionnaire in the excel worksheets, the texts can be extracted, translated in a separate page and then reinserted into the matching spaces on the questionnaire form. Furthermore, question or answer categories that appear more than once may be translated only once, and if new versions of the questionnaire arise only the new questions need to be translated.
- ✓ *Who should draft.* Drafting represents a tedious an iterative process, but it is also a very difficult and relevant task to be left in the hands of a low-ranked survey official. It is highly recommended that this stage be carried out by someone who is trained and recognizes the relevance of the task.

3.3. Testing the questionnaire

Once the iterative process of generating a version of the questionnaire which satisfies all those involved (end-users, policy-makers, the survey team, etc.), the testing process may be started. Testing represents a critical element of questionnaire design. Questionnaires should be tested at three levels:

- ✓ *As a whole:* is the full range of required information collected? Is the information across modules consistent? Are there unintentional double counts of some variables?
- ✓ *At module level*: does it collect the intended information? Are some questions irrelevant?
- ✓ *At question level:* is the wording clear? Is the question ambiguous? Have all responses been anticipated?

This should be done ideally in a two-stage process: a pre-test followed by a field test. Survey developers should resist the temptation to save time or money by skimping on either of these essential stages of the survey process.

It is important to distinguish between these tests. Pre-testing aims at assessing whether the questionnaire as a whole and its individual modules collect the full range of information required in a consistent way, and whether the wording of the questions is clear and unambiguous. Once pre-testing is finished, a new version of the questionnaire is developed and field testing begins. The main purpose of the field test is to check that all of the field procedures are properly understood and will work adequately in the real operation to follow

3.3.1. Pre-Testing.

The final version of the questionnaire is tested to check whether it fulfills the purposes of the survey. The underlying idea is to detect problems and correct them *before* the field test is conducted.

- ✓ *Who should conduct it?* It is sometimes said that pre-tests should be carried out by the same interviewers that will carry out the survey. This is false: the aim of the pre-test is both to detect problems in the questionnaire and to find ways to correct them. Even though interviewers will generally perform the former task very well, they are in no position to provide solution to questionnaire problems. This is why the pre-test should be conducted by the survey core team, the analysts, policy-makers and, most importantly, the impact evaluation team, since they have been those most closely involved with the development of the questionnaire, as well as a few experienced interviewers.

- ✓ *Sample.* Contrary to popular wisdom, the sample does not need to be either random or large. Instead, it should ensure diversity among respondent households, since the aim of the pre-test is to detect possible survey problems. Different types of households should be included so that the various situations likely to be experienced during the survey are observed during the test. For example, if the final survey will be conducted at a national scale, we would like the sample to include urban and rural households, men and women, children and babies, employed and self-employed individuals, etc.

- ✓ *Cyclic Process.* Pre-testing is a cyclic process. Problems are to be expected, which will inevitably result in multiple iterations. After one round of pre-testing, the questionnaire can be revised and

refined before submitting it to a new iteration of pre-testing. Heuristically, a pre-test of a small number of households will usually reveal about half of the deficiencies in the draft questionnaire during the first couple of days.

- ✓ *Interview time.* Interview time will typically be overestimated in pre-tests. If, however, the time exceeds by far what was expected, the questionnaire may need some revision.

3.3.2. Field-Testing.

From the team's discussions about the results of the pre-test, researchers generate a questionnaire to be field-tested. This version is used to produce the manuals and the first version of the data-entry program. When all these instruments are completed, selected staff are trained to conduct the field test. This is a test of the field operation as a whole: materials, working methods and all the logistical support.

- ✓ *Who should conduct it?* The teams who conduct the field test should be similar to those who will really work in the field later, without forgetting the data-entry operators who are an integral part of the teams.
- ✓ *Sample.* The field test should be carried out in a small number of sample points in the population. All the questionnaires, including the community and prices questionnaires (if they are part of the survey design), should be field tested at the same time as the household questionnaire.
- ✓ *Manuals.* Manuals can start to be drafted immediately after the pre-test, but they need to be reality-checked as a part of the field test. This task is neither insignificant nor rapid, which means that the time allowed for it should be specified carefully in the calendar of activities. High-quality documents can be quite lengthy (it is not unusual for manuals to have a hundred pages or more), but efforts should be made to restrict their length without jeopardizing the fullness and accuracy of the explanations and instructions. The manuals should have been completely finalized by the time the training of field staff gets under way. The issue is further discussed in the following section.
- ✓ *Data-Entry Programs.* Data-entry programs should be prepared immediately following the pre-tests, and should have been refined by the time of the field test. Once the modifications arising from the field test have been made, very few changes should be made to the data-entry program; most of these will be clustered in the first weeks of fieldwork.
- ✓ *Questionnaire modification.* Experience shows that even though the questionnaire should be close to its final state before the field test, it may nevertheless be modified afterwards. If so, the manuals and data-entry program must be adjusted immediately to reflect the changes. Clear and robust conventions should be adopted to track changes to the questionnaires; this will help to identify the alterations which consequently need to be made to the manuals and programs.

Field testing is an activity that lies in between two stages of the survey: questionnaire design and training. In fact, the questionnaire design stage will properly end only once training is concluded, since it is possible that valid suggestions are made during training, especially during the first few days.

The core team will have to decide whether the questionnaire should be changed in light of such suggestions, before it is sent to be printed.

With field-testing, then we move on to the following stage of the survey: Fieldwork Organization and Survey Management. This will be the focus of next module.

The three phases of questionnaire design - Check Questions

- ***3.3.2.1.*** *Is the Questionnaire asking the right questions?*
- ***3.3.2.2.*** *Are all the relevant topics covered?*
- ***3.3.2.3.*** *Is each topic adequately covered?*
- ***3.3.2.4.*** *Are the questions addressed to the right respondents?*
- ***3.3.2.5.*** *Will the respondents understand the questions?*
- ***3.3.2.6.*** *Is the flow of the questions adequate?*
- ***3.3.2.7.*** *Will the respondents be willing to answer the questions?*
- ***3.3.2.8.*** *Will the answers be recorded correctly on the questionnaire?*
- ***3.3.2.9.*** *Have all of these issues been "reality-checked"?*

Real World Example – How can Questionnaire Design Affect Analytical Purposes

Context. During 2013 and 2014, the ORPHEA (Optimizing the Response of Prevention: HIV Efficiency in Africa) project ran surveys in several African countries with the aim of estimating the costs of the health care system. The survey was originally developed for Zambia and then adjusted to be carried out in Kenya, Zimbabwe, and South Africa and is now being prepared for Nigeria. Even though the health system in Zambia presented some particularities (it is highly centralized and homogeneous in its functioning), adapting the survey to both Kenya and Zimbabwe did not yield major inconveniences.

What happened in South Africa? The development of the survey in South Africa, however, did not occur smoothly. As it turned out, the Health System in South Africa is by far more complex those of the countries where the survey had been run previously. The particular characteristics of the South African Health System, however, were not sufficiently taken into account when adapting the survey instruments.

What were the complications? Many problems arose during survey development. We will focus on one in particular: the inclusion of an "other" category for classifying personnel. Since nurses in South Africa may belong to two very different categories ("administrative" and "professional") and this was not explicitly marked in the questionnaire, the "Other" category was over-used.

Analytical implications. Within the "administrative" and "professional" categories of nurses, salary bands are very different. Thus, by observing the category "administrative nurse", it was almost impossible to deduce the cost of that particular nurse to the health system. This, however, was an essential goal of the survey analysis: estimating the cost of the health system. Analysts ended up having to calculate an average cost across all nurses, which did not comply with the goal of the survey.

Key Messages.

1. It is crucial to respect cultural context when developing survey instruments.
2. Apparently minor decisions in questionnaire drafting (in this case, its layout and formatting) may have significant analytical implications.
3. Pre-testing and Field testing determine the success of a questionnaire: indeed one of the purposes of these tests is to minimize the use of the "Other" category.
4. Close supervision of interviewers –using quality control indicators which will be discussed in Module 3.3—is fundamental to tracking potential problems and correcting them.

4. OTHER SURVEY INSTRUMENTS

Household surveys needs manuals, which explain the working procedures. Many surveys use community and other complementary questionnaires. These topics are covered in the following paragraphs.

Manuals

Manuals are usually written by local statistical agencies or survey firms in one (or sometimes more) of the local languages, which means that they are not usually revised by the international consultants advising on the survey. In an ideal world, they would be translated, like the questionnaires, into a major international language so that all users can refer to their definitions of the variables being collected. Many of the Key References at the end of this chapter provide further information about the content of these and other manuals.

- ✓ **Field manual.** The field manual – the interviewer's manual – is meant to be complementary to the questionnaire, not merely to reproduce it. The goal is not to produce a lengthy document, full of text copied-and-pasted directly from the questionnaire, but to create something which will help the interviewer when she is in the field. The field manual should provide the interviewers with the conceptual background to the survey and should define field procedures. If there are any sections of the questionnaire which are not self-explanatory, the field manual should cover them so that they are interpreted uniformly by all field staff. It should include general sections on the survey's objectives and methodology, the professional conduct expected of the interviewer, the structure of the questionnaire and the conventions used in it, and some information about the data outputs.
- ✓ **Team leader's manual.** The team leader's manual – sometimes called the supervisor's manual – should explain his or her tasks in each PSU: (1) completing the community, price and facilities questionnaires, and (2) meeting and working with the local authorities and the households selected. It should also cover how the team leaders are expected to deal with potential problems, such as difficulties in locating selected households or their refusal to participate.
- ✓ **Data-entry manual.** This should be written by the Data Manager and include instructions on how to use the data entry program, with a description of the program's error messages.

Supervision Forms

Supervision forms are documents written to support the team leaders in their quality control tasks. Using such forms, rather than leaving supervision to the team leader's personal initiative, gives formally defines these tasks and makes it easier for the core survey staff to supervise the team leaders' work.

The forms usually cover interviewer evaluation, questionnaire verification, and check-up interviews. Grosh and Muñoz (1996) reproduces examples of these forms from the 1991 Pakistan LSMS.

Community Questionnaire

The community questionnaire collects general information about the village or neighborhood where households are being interviewed. It is usually completed by the team leader. Some of the information is obtained from knowledgeable local people, such as village elders, police chiefs or mayors; other parts of the questionnaire deal with empirical matters such as road conditions.

Price Questionnaire

If local prices are being obtained by means of a prices questionnaire (see "Valuing food consumption" above), the team leader will be responsible for interviewing market traders and obtaining their prices. Food items are weighed, and both the weight and price are recorded on the questionnaire. Grosh and Muñoz (1996) reproduce (in "Annex IV") the instructions for completing prices questionnaires which were given to team leaders in the 1991–94 Kagera Health and Development Survey.

Ideally both the community and price questionnaires should be field tested at the same (or very nearly the same) time as the household questionnaire. This allows the analysts involved to treat the resulting data uniformly and to assess changes in one instrument that may have repercussions for other instruments.

5. PAPERLESS INTERVIEWS

The use of Computer-Assisted Personal Interview (CAPI) technology eliminates the need for a paper questionnaire, but does not directly ease the questionnaire development process. Indeed, all of the aforementioned issues and discussions in this module still apply. Furthermore, additional issues arise which may further complicate survey development. For example, how is it possible to carry out the questionnaire design process in a CAPI context? Many individuals will need to intervene, meet in order to discuss the content, etc. This implies that the questionnaire draft is often created initially on paper and then converted into CAPI.

Questionnaire Design

The CAPI application must permit the interviewer to move between modules, return to questions asked previously, and move ahead to specific questions (and then return to the present question). The interviewer must be able to interrupt an interview at the respondent's request, and return to the point at which the interview was interrupted at some point in the future to complete the interview.

Even though the use of a CAPI application eliminates the need to develop a paper questionnaire, some form of documentation is required in order to:

- ✓ develop the content of the CAPI application;
- ✓ distribute to the data users; and
- ✓ track the changes that are made over time.

Manuals and Supervision

In the case of a CAPI application, the interviewer's manual should include context-specific online help. Paperless interviews offer the possibility of additional supervision techniques, above and beyond the direct observation of interviews. For example, interviews can be digitally recorded; data can be time-stamped at the time of entry to ensure the integrity of the interview process; and information from multiple sources (reference tables, other respondents, other questionnaires) can be cross-checked automatically against the information as it is keyed in during the interview.

6. BIBLIOGRAPHY/FURTHER READING

1. Bardasi, E., Beegle, K., Dillon, A., Serneels, P. (2010) *Do Labor Statistics depend on how and to whom the questions are asked? Results from a survey experiment in Tanzania*, Policy Research Working Paper 5192. Washington D.C.: The World Bank.

2. Dubois, J.L. & Blaizeau, D. (1989) *Connaître les conditions de vie des ménages dans les pays en développement*. Ministère de la coopération et du développement, France (3 volumes).

3. Fowler, F. J. (1995). *Improving Survey Questions: Design and Evaluation.* Applied Social Research Methods Series, vol. 38. Thousand Oaks, CA: Sage Publications.
4. Grosh, M. and Glewe, P. (2000) *Designing household survey questionaires for developig countries: lessons from 15 years of the Living Standards Measurement Study*. Washington D.C.: The World Bank.
    i. **» see** *Vol I Chapter 5-Consumption (Angus Deaton and Margaret Grosh)*
    ii. **» see** *Vol I Chapter 1-Housing (Stephan Malpezzi)*
    iii. **» see** *Vol III Module for Chapter 9-Employment (Julie Anderson Schaffner)*

5. Grosh, M. and Muñoz, J. (1996) *A manual for planning and implementing the Living Standards Measurement Study Survey,* Living Standards Measurement Study, Working Paper 126. Washington D.C.: The World Bank.

6. Iarossi, G. (2006) *The power of survey design – A User's Guide for Managing Surveys, Interpreting Results, And Influencing Respondents*. Washington D.C.: The World Bank.
    i. **» see** *Chapter 3-How easy is to ask the wrong question*

7. Lohr, S.L.(1999) *Sampling: Design and Analysis. California: Brooks/Cole Publishing Company.*
    i. **» see** *Chapter 1. Section 1.5-Questionnaire Design*

8. Muñoz, J. (2007), Survey Quality Assessment Framework – SQAF, printing pending, draft available at http://lsms.adeptanalytics.org/course/fscommand/session2/ref/SQAF_pp_57-76.pdf
    i. **» see** *Chapter 6*

9. Scott, C., Vaessen. M., Coulibaly, S. and Verrall, J. (1988) *Verbatim Questionnaires versus Field Translation or Schedules: An Experimental Study*, International Statistical Review, 56(3): 259-278

10. Ravallion, Martin (1992), *Poverty Comparisons - A Guide to Concepts and Methods*, Living Standards Measurement Study, Working Paper 88 (or French version *Comparaisons de la pauvreté, Concepts et méthodes*, LSMS document de travail Nº 122). Washington D.C.: The World Bank.

11. United Nations / Statistics Division/Department of Economic and Social Affairs (2005) *Household surveys in developing and transition countries*. New York.
    i. **» see** *Chapter III-An overview of questionnaire design for household surveys in developing countries*

12. United Nations Population Division, Department of Economic and Social Affairs (2005), *Designing household survey samples: practical guidelines*, UN New York.
    i. **» see** *Chapter 2-Planning and Execution of Surveys, section on Questionnaire Design*

13. United Nation's Children Fund (2006) *MICS3 manual - Monitoring the situation of children and women - Multiple indicator survey manual*. New York.
    i. **» see** *Chapter 3-Designing the questionnaire*
    ii. **» see** *Appendix three-Instructions for interviewers*
    iii. **» see** *Appendix four-Instructions for supervisors and editors*
    iv. **» see** *Appendix five-Anthropometric techniques*

7. QUESTIONNAIRES OF SUCCESSFUL SURVEYS

Here is a short list of successful survey questionnaires:

1. Nepal Living Standards Survey (NLSS-III) 2008
    a. *The third round of the Nepal Living Standards Survey (NLSS-III) was conducted in 2008. The household questionnaire contains modules on all relevant aspects of living standards and inquired about food consumption using a recall technique. Detailed community-level information was collected with specialized urban and rural questionnaires.*
2. *http://lsms.adeptanalytics.org/course/fscommand/session2/ref/NLSS_III_Household.pdf*
3. *http://lsms.adeptanalytics.org/course/fscommand/session2/ref/NLSS_III_CommUrban.pdf*
4. *http://lsms.adeptanalytics.org/course/fscommand/session2/ref/NLSS_III_CommRural.pdf*
5. French Polynesia Household Budget Survey (EBF) 2000
    a. *The Household Budget Survey (EBF) in French Polynesia was conducted in 2000 under very challenging conditions (due to the extension of the territory), and it has served as a model for similar surveys in other Pacific Island Countries. Food consumption was collected using diaries.*
    b. *http://lsms.adeptanalytics.org/course/fscommand/session2/ref/Polynesie_Questionnaire_EBF_2000.pdf*
6. Malawi Third Integrated Household Survey (IHS3)
    a. *The third round of the Malawi Integrated Household Survey (IHS3) was conducted in 2011. It was part of the LSMS-ISA (Integrated Surveys on Agriculture) program, which uses the survey design of the LSMS but with a strong focus on agriculture.*
    b. *Approximately one in four households was designated as a panel household to be visited twice: after planting crops and after harvest. The rest of the households answered a cross-sectional agriculture questionnaire.*
    c. http://lsms.adeptanalytics.org/course/fscommand/session2/ref/IHS3.Household.Qx.FINAL.pdf
7. Living in Bosnia and Herzegovina Survey
    a. *This LSMS survey consisted of a 4-round panel conducted from 2001 to 2004. The panel monitored changes in living conditions over time and explored the dynamics of poverty in a way that is not possible with cross-sectional data, which was of special interest at a time of reconstruction and economic reform.*
    b. http://lsms.adeptanalytics.org/course/fscommand/session2/ref/BosniaLSMS_Eng.pdf