

**Harnessing Opportunities for Productivity Enhancement (HOPE) of Sorghum and Millets  
in Sub-Saharan Africa and South Asia**

# **Manual on survey design and data collection**

**Prepared by Dr. Christin Schipmann, ICRISAT Nairobi**

**based on**

**‘Designing Household Survey Questionnaires for Developing Countries, World Bank, 2000**

**and**

**‘Manual for the Implementation of USAID Poverty Assessment Tools’, USAID, 2008.**

**Nairobi, June 2010**

# Content

1. SURVEY DESIGN .....	3
1.1 Define objectives and decide the survey design.....	3
1.2 Deciding which topics (modules) to include in the survey .....	4
1.3 Draft modules.....	4
1.4 Modul combination .....	6
1.5 Translation, training and pre-test .....	7
2. DATA COLLECTION.....	10
2.1 Sampling .....	10
2.1.1 Introduction.....	10
2.1.2 The Sampling Process.....	11
2.2 Logistics of data collection .....	14
2.2.1 Personnel.....	14
2.2.2 Equipment.....	16
2.2.3 Other considerations .....	16
2.3 The interview .....	16
2.4 Quality control in the field .....	18
3.0 REFERENCES .....	18

## 1. SURVEY DESIGN

The five steps of survey design.

First: **define the fundamental objectives** of the survey and **choose an overall survey design** that best fits these objectives.

Second: **deciding which topics to include in the survey**, specifying the objectives of each module, and proposing an approximate length for the modules.

Third: work out, question by question, **draft questionnaires** for each module that will be included in the survey.

Fourth: **compare the modules to each other** to ensure that they are consistent and well integrated.

Fifth: translate (if necessary) and field test the draft questionnaires.

### 1.1 Define objectives and decide the survey design

#### Organizing a survey design team

The most important factor ensuring the success of a multitopic household survey is the involvement of the right people in the process. Designing the survey questionnaires is much too large a task for one person. Instead, a team of experts must be involved, including members of the organization implementing the survey as well as research analysts from other institutions. They define the overall objectives of the survey.

HOPE: The survey was first designed by socio-economists from ICRISAT from various regions. In a second step the survey has to be adapted to the country specific situation. This is done in cooperation with local partners. It is one of the main objectives of this training week.

#### Factors for deciding among various survey designs

The most important factors for determining the design of a proposed survey are: the kinds of policy issues the survey aims to address; the information available from existing surveys and other data sources; the country's institutional capacity for collecting data; and the financial and other resources available for implementing the survey, including any constraints on how these resources can be used. Possible survey designs include among others mail, phone or personal interviews, focus group discussions.

HOPE: We decided to conduct personal interviews using a structured questionnaire. Mail or phone interviews would not work in Ethiopia and as we need precise information of individual households we cannot do focus group discussions.

## **1.2 Deciding which topics (moduls) to include in the survey**

Survey designers must decide which modules to include in the household questionnaires and must also define specific objectives for each module and decide on each modul's approximate length. The tasks of choosing moduls, defining their objectives, and setting their approximate size are all closely related and thus must be done simultaneously rather than sequentially.

HOPE: The survey has the following objectives. Collecting information about variety adoption; farm income and profatibility of in particular sorghum and finger millet; information of utilization of sorghum and finger millet; information on commerzialisation of sorghum and finger millet; welfare effects of adoption of new sorghum and finger millet varieties; socio-economic household characterists. The objectives are set according to the vision of sucesss of the project. The objectives are reflected in the various modules of the questionnaire. Generally spoken, typical questionnaire modules are a household module, crop production module, expenditure module, financial and social asset module.

## **1.3 Draft modules**

Once the objectives for each module are finalized (at least tentatively), survey designers can begin to develop detailed draft modules for the household questionnaires. It always is attractive to put many questions to get very detailed and precise information. However, any questions not relevant to the objectives defined before should be removed. Otherwise the module will be too long. Once all modules are drafted they should be put together to estimate the length of the survey. In case the questionnaire is too long any questions that are needed only for the analysis of the least important of the policy issues also need to be removed. This process should continue until the questionnaire meets the length constraint. The general principle is that the most important policy issues should be addressed first, and additional ones should be included only if space allows. This approach is a good start, but much more remains to be done.

Once each draft module has been written out in its entirety, the next task is to verify that the design of each module reflects the economic and institutional structures of the country in question. Questions and response codes need to be checked and, if necessary, modified to reflect local institutions and terminology.

### Rules for formatting modules

Before drafting the modules, the team has to select a formatting convention. For example, uppercase and lowercase letters can be used to distinguish words spoken aloud during the interview from instructions to the interviewer, but this could also be done in other ways, such as

using different colors or different fonts. Once a convention is selected, it is extremely important to use it consistently throughout the whole questionnaire (rather than choosing the "best" convention from among several possible options for each module). The convention chosen should be the one that is clearest and most likely to minimize the possibility of errors.

HOPE: Phrases that are instructions for the enumerator and should not be read out loud are put in italics.

### *Precoding*

All of the potential responses to almost all of the questions in the questionnaire should be given code numbers so that the interviewer records only code numbers, as opposed to words or phrases, on the questionnaire. In most cases these response codes should be printed directly in the box where the question appears, or next to the question if there is no box around it. Where the list of codes is lengthy and applies to several questions, it should be placed in a special box on the border of each page for which it is needed. Alternatively, if a list is very long it can be printed on the back of the preceding page (making it visible when the interviewer fills out the page in question).

Precoding requires that response codes be clear, simple, and mutually exclusive, that they exhaust all likely answers, that respondents will not all provide the same response, and that none of the codes apply to only a handful of respondents. Designing adequate response codes requires extensive knowledge of the phenomenon being studied as well as careful field testing.

### *Verbatim questions with simple answers*

All questions are written out in their entirety and are meant to be read out verbatim by the interviewer. This is done to ensure that questions are asked in a uniform way, since different wordings may elicit different responses.

When choosing the wording of questions, it is important to use terms that reflect the language as it is commonly spoken. Using language that is too formal or academic will make the interview stilted and unnatural.

The answers to the questions must be kept simple. Many survey designers are tempted to shorten the questionnaire or simplify the skip pattern in a way that results in complex questions and answers. This should be avoided since it will confuse respondents and is unlikely to save time.

Skip codes are codes that tell the interviewer which question to proceed to after finishing the current question. Some skip codes apply only when a particular answer is given. In such cases an

arrow and the number of the question to skip to are positioned in parentheses next to or below the individual response to which the code applies.

#### **1.4 Modul combination**

Once draft versions of each of the individual modules have been written, these drafts must be combined to form complete household questionnaires. Merely stapling the various modules together will not produce a well-designed questionnaire; but the modules of the household questionnaire must be made consistent with each other. Important aspects to consider are the following:

##### Gaps and overlaps

Ensure that all information that is needed is collected. It sometimes happens that one aspect can be covered in different modules. E.g. expenditure for land rent can be either asked for in the crop module or in the expenditure module. However, it needs to be ensured that it is covered in one of the two. This leads to overlaps. If an information is already collected in one module it does not need to be asked for again in another module.

##### Length

When combining modules the relative length of each module needs to be considered. In case a module that is considered as minor important is much longer than one considered as major important one should check these modules again and might remove some questions from the minor important module. Moreover the total length of the questionnaire needs to be considered when combining modules.

##### Recall periods/consistency

The recall periods for all information need to be consistent throughout the questionnaire. It has to be decided once for which period data should be collected. This should neither change between modules nor be changed once the survey has started. Typical recall periods are a period of the last 12 month or the last planting season/s. For some data, e.g. expenditure, shorter recall periods are needed to get precise information. In this case follow up questions are needed to ensure that the information can be afterwards used to calculate data for a 12 month period.

Example: Expenditure for tobacco might be easier to collect for a one week period. (How much did you spent on tobacco in the last seven days?) In this case a follow up question could be: Is this the average weekly amount you spent throughout the last 12 month? In case the answer is no

another follow up question is needed to get information about the average amount that is spent on tobacco consumption.

Moreover information needs to always be collected for the same units.

Example one: One household states its input use for a specific crop in kg, the other one in bottle. It is not possible to afterwards compare the input use of these two households. A way to overcome this problem is to have a list of standardized measurements of the area. This list needs to be obtained before the survey starts and it needs to be ensured that the list is correct and applies to the measurements used by the farmers.

Example two: The crop output data is collected on plot level. The crop input data is collected per crop. This leads to biased estimates as it might be that the inputs were not equally distributed around the crops. Productivity and efficiency estimated might lead to over or underestimates.

#### Nomenclatur and coding schemes

Ensure that the nomenclatur and coding schemes are consistent. The questionnaire should stick to one format for each (compare ‘Rules for formatting modules’ above).

#### Order of the modules

Modules should follow a logical order. A questionnaire should not start with a crop module, followed by a household module and then a crop marketing module. Modules about similar topics should be grouped. Moreover the questionnaire should begin with some warm up questions which are easy to answer, like information about the village, the address, name, phone number, etc.

Each questionnaire must collect some metadata in the beginning. This is usually data about the geographical location of the household, the date and the name of the interviewer, the respondent name and the length of the interview. Moreover, every questionnaire needs an unique ID code.

### **1.5 Translation, training and pre-test**

#### Translation

In case the questionnaire needs to be translated from English into one or more local Languages, the entire questionnaire must be translated before the interviewer training. Allowing the interviewers to translate during the interviews will bias the results, since each interviewer might phrase a question a bit differently. Translations needs to be checked in particular for the following regular mistakes:

- *Incomplete set of responses.* Bias is introduced when some of the multiple choice response alternatives available to the respondent are left out.
- *Leading questions.* A leading question is one that seems to lead or encourage the respondent to answer in a certain way. For example, “You like the way this manual is written, don’t you?” makes it harder for a person to say that he or she does not like the manual. In contrast, “Do you like the way this manual is written, yes or no?” does not lead the respondent. When you translate questions it is important to maintain the intended structure and not introduce or rephrase the question so that it suggests the preferred response.
- *Unfamiliar terms and jargon.* Wherever possible, words that are likely to be familiar to the respondent should be substituted for unfamiliar terms or technical jargon.
- *Poor grammatical format.* Weak grammatical format can introduce bias by confusing the respondent.

Training of enumerators

All future enumerators for the survey need to participate in the training. It is highly recommended to **train more than minimum number of interviewers**. There are always illnesses or family and work emergencies that pull trainees away at the last minute and there usually is not time to train substitutes adequately. Therefore, it is best to train more people than needed. Also, during or at the end of the training, the Manager or Trainers want the ability to recommend that certain trainees be discontinued because they did not prove during the pre-test to have the skills needed to be good interviewers.

The supervisors and all of the other people who are going to be involved in the process should go through the same training so that they know what is being asked and expected of the interviewers. This will allow them to be of better assistance to the interviewers during the days of interviewing.

**Table 1: Example of an enumerator training**

<b>Day 1</b>	<ul style="list-style-type: none"> <li>*Introductions; agenda for data collection; agenda and logistics for training week.</li> <li>*Purpose, objectives and key concepts of this research and the tools.</li> <li>*Interviewers’ roles and responsibilities.</li> <li>*Reviewing the survey in the first language.</li> <li>*Techniques of survey interviewing and practice first survey.</li> </ul>
<b>Day 2</b>	*Review the questionnaires in each of the languages and make consensus changes in writing.

	*Practice each of the surveys in the local languages.
<b>Day 3</b>	*Finish any remaining practice from Day 2 and practice exercises. *Learn sampling and interviewee replacement techniques. *Prepare for going to field to interview early in the morning on day 4.
<b>Day 4</b>	*Pre-test the survey with sample hh; *Reflect on pre-test experience and lessons learned.
<b>Day 5</b>	*Make final changes to each of the tools in each language. *Print up the final versions each language and make enough copies *The interviewers should practice again.

In case many problems occur during the pre-tests, the questionnaire should be adapted according to the problems that occurred and need to be pre-tested again until it works.

*Pre-testing of the questionnaire*

Pre-testing has to be done with people who are similar to the sampled respondents. Pre-testing is an essential step in finalizing the questionnaire, because it always bring out potential problems, inconsistencies or other sources of bias and error. The pre-test should include respondents of the same socio-economic, ethnic group and location to those respondents who are in the final sample (none of the pre-test respondents should be included in the final results). If the survey is going to be implemented in a rural area, the pre-test is also conducted in a rural area.

**Pre-testing a questionnaire in the field should involve everyone in the survey team.** The pre-test is essential for finding weak points in translations and errors in the logistical plan, as well as identifying the need for additional field staff training. Pre-testing provides an opportunity to make corrections before doing the actual survey and to learn more about the time and resources required to locate and interview respondents. Pretesting is done as part of the interviewer training. A pre-test should address the adequacy of the draft questionnaires at three levels:

*Questionnaire as a Whole:* Is the full range of required information collected? Is the information collected in different parts of the questionnaire consistent? Are any variables unintentionally double-counted?

*Individual Module:* Does the module collect the intended information? Are some questions missing? Are some questions redundant or irrelevant?

*Individual Questions:* Is the wording clear? Do any questions allow for ambiguous responses? Are there multiple interpretations? Have all responses been anticipated and coded?

Each interview during the field test should include, at minimum, the respondent, the interviewer, and an analyst or senior survey specialist. During the field test it is acceptable for the analyst or survey specialist to interrupt the interview tactfully in order to refine the wording of a question or the responses coded for it. Of course, in the actual survey the interviews should be conducted in private, and the interviewers should adhere to the wording of the questionnaire. The data from the field test should not be entered in the computer or examined for any analytical purposes. Pre-tests also need a follow up day, where pre-tests and needed modifications of the questionnaire are discussed.

## **2. DATA COLLECTION**

### **2.1 Sampling**

#### **2.1.1 Introduction**

The sample must be large enough and truly random to be representative of the population in the research area. This is every household MUST have the same chance of being selected in the sample. The best way to ensure that your sample is random would be to create a list of all households in the research area, and then to randomly select the number required for your sample (a small-scale example would be to put everyone's name on a piece of paper, cut the names apart, put them all in a hat, and then draw a few names). However, sampling from a master list can be impractical if households are very spread out, geographically, so that it would be very time consuming and costly to have the interviewers travel so much between interviews; moreover, with a large number of households, it can be very difficult to generate a list of ALL households (hh) that is up-to-date.

To avoid these problems, two sampling methods are available to use instead: cluster sampling and the random walk. In most instances, cluster sampling will be used to divide the clients into smaller groups (such as regions or branches), which will allow the sample selection to be done using a shorter list of hhs. If it is impossible to generate any lists of hhs, then the random walk method may be used.

#### ***Cluster Sampling***

Cluster sampling is a method whereby you divide the population into geographic units, sample the units, and then divide the selected units into smaller geographic units and resample, repeating the process until the actual people to be interviewed are selected. For instance, in the first stage of sampling, different research areas may be sampled/defined (in our case the three rings: target,

diffusion and control area), then villages within these areas are sampled and finally hhs. So, instead of taking a random sample of hhs from a very long list spread out over a wide geographic area, the sampling is divided into stages. Now that you have a much more manageable list (of hhs in just a few communities or cities), you can easily number them, and then randomly select the desired number from a random number list.

### **2.1.2 The sampling process**

#### *Step 1: Define the population and sampling method*

Determine the geographical area to be sampled. Identify any problem areas that may not be feasible to survey. In general, the implementation team will need to determine a standardized rule or set of rules for filtering out unfeasible survey areas and then follow this rule consistently. However, the process of eliminating areas from consideration needs to be carefully scrutinized to avoid any unintentional introduction of bias. Before setting any rules to limit the possible survey area, you need to consider whether the implementation team's proposed rules would result in a possible selection bias. The team may have reasons for excluding operational areas from the survey. In summary, determining the feasible area for the survey requires good information and careful judgment to avoid bias.

HOPE: Generally the survey areas are determined by the project location. There are three project locations per each crop (sorghum and finger millet). For the survey, out of these three two will be selected by the project management team for each crop.

#### *Step 2: Construct the sampling frame*

The sampling frame refers to lists of the number and distribution of hhs within the research area. These lists can be used to determine which localities (and the number of hhs in each) will be surveyed. In most cases, a list of all qualifying households is not needed. Instead, information on the distribution of hhs within the research area can be used to randomly select a handful of smaller geographic areas, making it necessary to prepare hh lists only for these areas (see 'cluster sampling' above).

HOPE: Our sampling frame follows a stratified sampling approach. Stratified means that different groups of households are purposely selected. In the project we are interested in three different hh groups. (i) hhs that are directly involved in the project because they are living in the villages where the project takes place and participate at project activities. (ii) hhs that live in the area that surrounds the project area and can indirectly benefit from the project through spillover effects. This area is called the diffusion area. (iii) hhs that live in an area that is that far away from the

project area that no direct spillover effects are expected to happen there. This area is called the control area.

For reasons of comparability all three areas should be in the same agro-ecological zone and sorghum or finger millet have to be the most important crop in all three areas.

#### *Determining required clustering stages/ stratified sampling*

The first stage is the definition of areas where the survey takes place (three areas). The second decision is how many villages/hh should be included in the survey. These villages will then be randomly sampled within each area. Third, a list of all households in the selected villages is needed and a random selection method to sample hhs from this list.

HOPE: Our project wants to sample an equal number (80+5) of households in the project area and in counterfactual areas. Villages in the project areas are already selected, namely two. A list of all hhs in the two villages has to be obtained. Villages in the other two areas are not yet selected. In a first step a list of all villages in the diffusion area and in the control area is needed. From these lists two villages are randomly selected for each area. The next step is to get lists of all hhs in the selected villages. The hhs lists from villages in the same area (project, diffusion and control) need to be combined. The needed number of hh per area (80 in the project area and 40 in each of the two other areas) is randomly selected from the combined list. A mistake would be to select purposely the same number of hhs per village in a given target area. Assuming that each village has a different number of hhs, choosing the same number of hhs per village would introduce bias by giving clients different likelihoods of being selected. Those in the smaller villages would have a greater chance of being selected than those in the larger villages.

#### *Step 3: Determine appropriate sample size*

Calculating sample sizes is one of the most technically demanding aspects of survey design. On a practical level, sample size is partly determined by the time and resources available for the survey. On a technical level, four factors affect the decision on sample size: (i) the desired precision of the survey, (ii) the probability distribution of the variable that the survey seeks to measure in the population, (iii) the choice of sampling design (i.e., single random sampling or multi-stage random sampling), and (iv) the number of variables (in this case, poverty indicators) that the survey seeks to capture. A rule of thumb is a minimum number of 30 hhs per target area.

#### *Step 4: Select the actual sample*

Selecting geographic clusters is convenient, but not all clusters will necessarily be the same size, so it is more likely that not everyone will have the exact same chance of being selected.

Example: Cluster A (diffusion area) has 50 hh and cluster B (control area) has 100 hhs. Selecting 5 out of 50 gives everyone in A 10% chance of being selected, while selecting 5 out of 100 gives everyone in B a 5% chance of being selected.

When clusters are of different sizes, the perfect situation would be to select a different number of hhs from each cluster. Another method that we apply in HOPE is to oversample the project area and afterwards in our analysis add weights to the clusters to ensure representative results. It is therefore necessary to collect information about how many hhs are in total living in each of the three areas.

#### *Random sampling with excel*

Open an excel file and write the numbers of farmers from your household list in the first column (column A). Start with 1 in A1 and continue downwards in the table until the final number on your hh list. Mark B1 and go to INSERT in the menu above and there on FUNCTION. Enter RANDOM in the first field SEARCH FOR A FUNCTION. From the list given below select the possibility RANDBETWEEN. A new window appears. In the first column of this window you enter A1 in the second column you enter the last A row with a number. In case of 100 hhs this would be A100. You put \$ before the A and before the number (\$A\$1) to keep these fixed when you copy the formula to other columns. You enter ok and a random number will be generated in B1. You copy the formula to as many rows in column B as many hh need to be sampled. If you need to sample 20 you copy the formula until B20. The numbers given in the 20 rows of B are the hh IDs from column one that need to be selected for the sample.

#### *Step 5: Selecting “extras” for the sample*

In addition to randomly sampling hhs to interview, you should prepare a second list of randomly sampled hhs to place on a reserve list in the event that a sampled hh does not qualify for an interview or is unable to be interviewed. As a rule of thumb, the reserve list should contain an additional 4 reserve names for each 10 sampled names. Once the survey process is underway, the first name on the reserve list is taken to replace the first sampled hh that cannot be interviewed. All additional replacements are made in the order in which they appear on the reserve list.

#### *Step 6: ID codes for selected hhs*

Once hh list for the primary sample and for the replacements are obtained, a hh ID has to be added to every hh name. The hh ID consists of a letter that specifies the cluster and a number that enables identification.

HOPE: hhs from the project area have the letter code P, from the diffusion area D and from the control area C. They are then numbered from one to the final number. Replacement hhs have the

letter code PR (project area replacement), DR (diffusion area replacement) and CR (control area replacement) and then again numbers. Once a hh replacement is interviewed it needs to be marked on the primary list and on the questionnaire, which hh is replaced by which respective hh.

## **2.2 Logistics of data collection**

### **2.2.1 Personnel**

The implementation team for the survey should include an overall project manager, field supervisors, enumerators, (a data processing coordinator, and data entry personnel).

Project manager (Dr. Adam): oversees and participates in planning and implementation of the field survey and the training of the staff. S/he supervises:

- translation of the survey
- verify that the sampling is being done correctly and that only sampled households are being interviewed
- monitor progress towards completing the survey
- monitor the team's progress in staying on schedule and within budget as the field work progresses.

Field supervisors (Socio-economists in the region): coordinates daily activities of interviewers, and quality control in the field; creates and maintains tracking system for survey implementation; ensures sampling plan is maintained. Moreover, s/he:

- helps with the interviewer training and translation of the survey
- coordinating the daily activities of the interviewers, including arranging movement to and from interviews and transport from one survey site to the next
- control that the questionnaires are filled out correctly and completely and that the information contained in them is accurate before leaving each survey area
- check the work of each interviewer on a daily basis to minimize the number of errors and missing values
- report regularly to the project manager on progress, costs incurred, and any irregularities in the field

Enumerators: thoroughly learn survey questions and interview techniques to minimize bias, and conduct interviews in a confident and relaxed manner. S/he:

- reports regularly to the manager on progress, costs incurred, and any irregularities
- participates in training and review the questionnaire

- be available throughout the training process and respect the schedule
- respect interviewees and explain questionnaire according to their level of education
- ask questions with precision and record information accurately
- use probing questions when necessary
- maintain the same quality of work throughout the process
- resolve any logistical or other problems or report to the field supervisor as necessary

(Data processing coordinator: identifies and trains data processors and maintains quality control in data processing. S/he:

- reports regularly to the manager on progress, costs incurred, and any irregularities
- participates in training and review the questionnaire
- sets up numbering system for all surveys: test surveys and actual surveys, including replacement selections for each sampling category
- participates in sampling selection and oversee the assignment of survey numbers
- confirm that all surveys have been handed in and are completely filled out
- compiles statistics totaling number of surveys completed for each sampling group and supervises a physical counting of documents to confirm the totals
- leads training on survey coding and data entry techniques
- supervising translation of data entry template into local language (if needed), personally test data entry
- directly oversee data entry of first few surveys to ensure everything is being done correctly
- supervises data input and cleaning, ensures that all surveys have been entered and re-entered
- ensures survey forms are filed and kept in a safe and accessible place, and final database is properly backed up.

Data entry personnel: enter and clean data collected in the survey. S/he:

- supervises data input and cleaning, ensures that all surveys have been entered and re-entered
- ensures survey forms are filed and kept in a safe and accessible place, and final database is properly backed up
- prepares equipment and test systems
- daily data entry and data back up
- reports problems/obstacles to data processing coordinator

- provides data processing coordinator with requested information)

### **2.2.2 Equipment**

*Transportation:* Vehicles should be large enough to carry the field team and supplies, and sturdy enough to withstand road conditions in survey areas.

*(Computers:* in case of data entry during the survey computers should be designated full-time for the persons responsible for data entry.)

*Photocopying:* Several copies of the survey form will need to be copied, collated and stapled.

*Office Space:* Office space and desks should be dedicated to the survey team to meet, discuss, store questionnaires, etc.

### **2.2.3 Other considerations**

\*The timing of the survey needs to be well planned. It should not take place at peak harvest seasons or other periods of intensive workloads of farmers. This holds true for the timing of the survey as well as for the timing of the interview. Moreover, the needs of the survey need to be considered. E.g. which planting period is covered.

\*A survey schedule needs to be designed that states who will be interviewed where at which day.

\*Check with local field staff to make sure time estimates for travel between interview locations reflect current road and traffic conditions.

\*Communication methods and back-up plans should be identified (cell phones, phone cards, etc.)

\*Route planning should include access to petrol stations, food, accommodations, if necessary.

\*Consider variations in schedule to accommodate local customs, holidays and political circumstances to avoid problems.

## **2.3 The interview**

### *Step 1: Identify yourselves*

Hhs will be more cooperative if they know who is conducting the study. Please introduce yourself with your name and a short notice about your professional background.

### *Step 2: Show letters of introduction and endorsement*

In some countries it is expected that outsiders will first seek permission from local leaders before approaching hhs in a given locality. In addition to introducing the survey, these courtesy visits also can provide you an opportunity to collect important information about the community being surveyed. In some cases, a letter of introduction from your headquarters to hhs and from local authorities might be needed.

### Step 3: The interview

#### *Begin the interview*

- Fill in the date of the interview, the time when the interview starts and other information that might be required in the very first part of the questionnaire.
- Begin your conversation with an introductory dialogue explaining the organization's interest in learning more about the hhs and ensuring confidentiality. Most hhs will not fully understand the methodology used for this study. However, you should be able to explain the overall objective of the study in easy words so that hhs understand the importance of the interviews. Ask the client permission to interview her before asking her any survey questions.

#### *During the interview*

- Ask the exact question (as written) and listen/determine the relevant information.
- Show interest. Pause. Repeat question if necessary. Repeat her/his reply to stimulate the hh to say more, or to recognize an inaccuracy
- Record answers in boxes/correct places
- Probe (*not prompt*) to increase accuracy/clarity & completeness
- Avoid unnecessary reinforcement: "oh, that's very good!"
- Never suggest an answer

#### *Ending the Interview*

- Thank the hh and tell her/him that s/he has provided important/helpful information for the study
- Answer any questions s/he has
- Quickly proofread completed questionnaire. Find and correct errors. Clarify answers with hh if necessary.

#### Additional tips:

- Do not get involved in long explanations of the study and keep interruptions to a minimum
- Do not deviate from the sequence of questions or question wording
- Do not rush the respondent; give her/him adequate time to answer and clarify anything s/he does not understand
- If a hh is unresponsive, then politely conclude the interview when appropriate and eliminate her incomplete survey at a later time
- The interviews should take place in the home of the hh, if possible. Seasonal demands on the time of the hhs should be taken into account as well

## **2.4 Quality control in the field**

The best time to correct mistakes is immediately after they occur. *Thus, the most important step in the quality control process is carried out by the interviewers and field supervisors while they are still out in the field.* Once the survey forms have returned to the central processing location, it will be very expensive to try and return to the field to re-interview anyone. The basic steps of the quality control process in the field are:

- Interviewer confirms that the person to be interviewed is on the sample list before beginning the interview.
- Interviewer looks over the survey form immediately after completing the interview, looking for missing, incomplete, or inconsistent entries, or any writing which a data entry staff member might not be able to read. If the interviewer sees that any information is missing, he or she might be able to remember what the answer was or be able to ask the interviewee before they leave. Every question should be completed.
- Field supervisor meets the interviewers after each one or two interviews to collect the surveys and look them over to make sure they are complete, easy to read, and have no inconsistencies. If the field supervisor sees a number which looks like a “0” but could be a “6,” instruct them to clarify it ASAP. Even if s/he can tell the difference, the data entry staff member who enters the data may not. The supervisor also should check the sampling list to make sure the sampled person was interviewed or that the substitution was done correctly.

### Document Control and Tracking

It is critical to ensure that there is a system in place for collecting and tracking all documents, including survey numbers that are written on each questionnaire prior to the interviews, and a log book for tracking the status of each document.

## **3.0 REFERENCES**

World Bank, 2000. *Designing Household Survey Questionnaires for Developing Countries*, Vol. 1-3. Edited by Margaret Grosh and Paul Glewwe, Washington, DC.

USAID, 2008. *Manual for the Implementation of USAID Poverty Assessment Tools*. Prepared by The Iris Center, University of Maryland College Park.