

Appendix 4: Field test approaches, instruments and instructions

The purpose of this field test is to determine how the short set of questions, as described in the question battery and the question by question specifications, are working in different countries. There are four sets of questions and three approaches to their implementation. The data collected in the field test will provide the basis for the analysis described in Appendix 6 (“Analytic Plan”).

This section does not discuss the processes of pre-testing (including expert review) and pilot testing. Please refer to the relevant section on “Sample design” in the implementation protocol for details on these processes.

The two questions that countries participating in field testing should ask themselves are:

1. Which set of questions should be used in the field test?
2. What approach should be used for the field test?

Questions sets:

Four different sets of question options are proposed. Within certain limits, each country can decide which of the four sets of questions to use in their field test. The first is compulsory, the second is highly recommended, while the remaining two are optional depending on an individual country's needs and resources. .

Whichever sets of questions are chosen, the order of administration is prescribed. The WG short set should always be first followed by the detailed questions for each domain in the WG set. The order of any other question sets can be decided by the individual countries.

A. Short set (Compulsory)

The short set as set out in the question battery are the main focus of the question-testing exercise. These should be administered in all of the options for field testing as they represent the minimum set. There are 6 questions covering 6 domains of human functioning: seeing, hearing, communication, walking, self care and cognition. The questions are set out below.

Short set:

Introductory phrase:

The next questions ask about difficulties you may have doing certain activities because of a HEALTH PROBLEM.

1) Do you have difficulty seeing, even if wearing glasses?

No - no difficulty

Yes – some difficulty

Yes – a lot of difficulty

Cannot do at all

2) Do you have difficulty hearing, even if using a hearing aid?

No - no difficulty

Yes – some difficulty

Yes – a lot of difficulty

Cannot do at all

3) Do you have difficulty walking or climbing steps?

No - no difficulty

Yes – some difficulty

Yes – a lot of difficulty

Cannot do at all

4) Do you have difficulty remembering or concentrating?

No - no difficulty

Yes – some difficulty

Yes – a lot of difficulty

Cannot do at all

5) Do you have difficulty (with self-care such as) washing all over or dressing?

No - no difficulty

Yes – some difficulty

Yes – a lot of difficulty

Cannot do at all

6) Because of a physical, mental or emotional health condition, do you have difficulty communicating, (for example understanding others or others understanding you)?

No - no difficulty

Yes – some difficulty

Yes – a lot of difficulty

Cannot do at all

B. Additional questions on core domains (highly recommended)

In order to have a better sense of how the short set of questions proposed by the WG are performing, it is necessary to ask a further set of questions that can then be compared with the short set. These questions should be asked after the short set.

The structure of this option comprises asking -

- 1i the short set of questions as set out in the section on the questions battery followed by

1ii the additional questions on core domains as set out below:

Vision

- i. Do you have difficulty seeing and recognizing a person you know from 7 meters (or 20 feet) away?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

- ii. Do you have difficulty seeing and recognizing an object at arm's length?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

Hearing

- i. Do you have difficulty hearing someone talking on the other side of the room in a normal voice?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

- ii. Do you have difficulty hearing what is said in a conversation with one other person in a quiet room?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

Mobility

- i. Do you have difficulty moving around inside your home?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

- ii. Do you have difficulty going outside of your home?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

- iii. Do you have difficulty walking a long distance such as a kilometer (or equivalent)?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

- iv. Do you have difficulty in using your hands and fingers, such as for picking up small objects or opening and closing containers?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

Remembering

- i. Do you have difficulty concentrating on doing something for ten minutes?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

- ii. Do you have difficulty remembering to do important things?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

Self-care

- i. Do you have difficulty washing your whole body?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

- ii. Do you have difficulty getting dressed?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

iii. Do you have difficulty feeding yourself?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

iv. Do you have difficulty staying by yourself for a few days?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

Communicating

i. Do you have difficulty generally understanding what people say?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

ii. Do you have difficulty starting and maintaining a conversation?

No difficulty
Some difficulty
A lot of difficulty
Can't do at all

C. Additional questions:

In addition to the short set and the detailed questions for each domain, if space permits, questions on a further two domains can be added to enlarge the scope of human functioning already covered. The two additional domains proposed are learning and applying knowledge, and interpersonal interactions. These questions are:

Learning

- i Do you have difficulty learning a new task, for example learning how to get to a new place? (*Countries to think of appropriate examples of a new task*)
- ii Do you have difficulty analyzing and finding solutions to problems in day to day life?

Interpersonal interactions

- i Do you have difficulty dealing with people you do not know?
- ii Do you have difficulty maintaining a friendship?
- iii Do you have difficulty getting along with people who are close to you?
- iv Do you have difficulty making new friends?

The same responses should be used for these additional domains;

- No difficulty
- Some difficulty
- A lot of difficulty
- Can't do at all

Finally, it would be useful to be able to include questions that tap into the psychological domain if that is culturally acceptable. The following set of questions can be used for that purpose.

How much, during the past 4 weeks....

	All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	None of the time
mhi1: Did you feel very nervous?	1	2	3	4	5	6
mhi2: Have you felt so down in the dumps, nothing could cheer you up?	1	2	3	4	5	6
mhi3: Have you felt calm and peaceful?	1	2	3	4	5	6
mhi4: Have you felt down-hearted and depressed?	1	2	3	4	5	6
mhi5: Have you been happy?	1	2	3	4	5	6

D. Country's own set of questions

The fourth set of questions comprise those which meet the country's own concerns, for example, questions on disability used by that country in a previous census or as part of a national household survey. Responses to these questions could be compared with those from the short set.

In summary, the structure of the questionnaire will be:

1. Background questions (age, sex, educational level, employment status, and other socio-demographic questions the country might want to add.) Countries can use their usual wording and response categories for these questions.
2. Section A: Short measurement set (4 or 6)
3. Section B: Additional questions for 6 domains
4. Section C: Further questions on the additional domains of learning, social interactions, and psychological functioning
5. Section D: Country specific questions on disability

Field test approaches

The choice of question sets is dependent to some extent on the approach to the field test. Three approaches are discussed. The first is to piggy-back the chosen question set onto a planned census or national survey; the second is to design a special study to test the questions using a small sample of purposively chosen respondents; and the third is to carry out a survey to test the questions using a population based sample.

Each approach has advantages and disadvantages and most importantly implications for the choice of sampling strategy.

Census or survey approach

The first approach is to insert the short set of questions in an existing census run or national survey.

The advantage of this approach is that it provides a large sample from which to obtain test results¹. The approach uses principles of sampling for a survey or asking the whole population as in a census. The sampling strategy would have been determined by the original survey design and not dictated by the disability question test. A further advantage would be that there would be minimal cost of doing the disability question test as the data collection would be happening anyway.

The disadvantage of this approach is that there may not be sufficient space to add all the additional questions that would make the test most effective. If space is limited and only sufficient for the short measurement set, a sub-sample should be identified and the additional question set and cognitive testing administered to this sub-sample. This would have some costing implications as there would be additional data collection time involved.

¹ This assumes that the national survey would involve a large sample of respondents, e.g. between 5000 and 10,000 households.

Special study approach using a small purposively selected sample²

A second approach is to test the questions on a sample that is selected based on the probability of responding 'yes' to one or more of the Census questions (same as a quota sample). This would involve designing a special study that will focus primarily on testing the short set of disability questions (i.e. comparing the short set to other questions on disability that cover the same domains).

The sample selected is purposive where the disability status of the respondent is known prior to administering the question set, although this status would not necessarily be known by the interviewer to avoid bias. A large enough sample of people with and without a disability would be selected through various networks.

The questions are administered and a 2 X 2 table constructed to measure the rate of identification of true positives and negatives using the known status of the respondent as a 'gold standard'.³ A suggested sample size is 200 true positives (with disability) and 200 true negatives (without disability). There must be a clear definition provided for who counts as a true positive or negative and some means of assessing this status beyond the use of the questions.

The advantage of this approach is that it allows for a larger set of questions to be administered and the situation is more controlled in terms of being able to observe more

² See discussion of "quota sample" under "Field Test Instrument and Instructions" section of main implementation document.

³ 'Known status' – a more detailed assessment, a self-report, or some form of confirmation by a rehabilitation professional, physician, or observation.

interviews to better understand how the questions behave in the field. A further advantage is that the cost should not be prohibitive.

Special study approach using a sufficiently large population based sample

A third approach is to devise a special study using a population based sample. The sample size should be calculated based on the expected prevalence rates for the different types of disabilities in the country population. For example, in a country with a 20% prevalence rate for visual difficulties where you interviewed 1000 cases, you would get roughly 200 from whom you could obtain more detailed information about reported visual difficulties. The advantage of this approach is that it provides a large data set with possibility of including a very detailed set of questions on disability.

The disadvantage is that it is costly and time consuming an exercise.

Relationship between the cognitive test and field test

The cognitive test and the field test are administered to separate samples. The short set is included in both the cognitive and field tests. Additional questions about each functioning domain are also included in the field test and the cognitive test but, in addition, the cognitive test includes some questions that are not on the field test. In other words, the cognitive test contains some questions that are unique to the cognitive test and there are some questions that overlap between the two tests.

Treatment of non-response and refusals

For results from a data collection exercise to be regarded as valid, responses must be obtained from a sufficiently large proportion of the sample to keep possible sample loss bias to an acceptable level, and the final sample achieved must be large enough to produce the test significance required. This is a particular problem for any data collection where response is dependent on the willing cooperation of randomly selected household respondents.

There are a number of factors that can significantly reduce the size of the final respondent non-response and refusal rates:

1. Use of an up-to-date sample frame will assist in reducing sample loss due to demolished, or non existent dwellings;
2. Having clearly defined reasons for the data collection and identifying benefits to be obtained will facilitate respondent cooperation;
3. Making initial contact with selected households prior to interviewer contact: where possible. This should be done by mailing an explanatory letter/brochure to selected households prior to interviewer contact, or in areas where this is not practical, contact with someone such as a village elder or other 'significant person' who can be informed of the nature and importance of the collection and can in turn use this information to assist in gaining cooperation of selected householders;

4. Adequate interviewer training, including ability to explain the purpose and importance of the interview, and sufficient practice interviewing and training to become proficient with the wording and sequencing;
5. Allowance within budget and time constraints for call-backs to households with no contact at initial visit, and spread of call-back times at households to cover different times of day/days of week to maximize possibility of contact;
6. Having a plan for further non-response follow-up where possible within time/cost constraints, if sufficient response/sample take has not been achieved within the initial interviewer timeframe.

It is important to identify the reasons for non-response. Separation of refusals/part-refusals from non-contacts and from sample loss (e.g. vacant/demolished dwellings) is important. Sample loss will not impact on data interpretation /accuracy other than by reducing the final effective sample size, and the initial sample should allow for sample loss to achieve the required sample size for the required precision. Non-contacts and refusals can potentially bias the results to the extent that non-contact and/or refusal respondents differ in their characteristics from the responding sample. For the initial test phase it will be very important for interviewers to record the reasons for a person's refusal, and in particular to identify if it was the subject matter and/or particular questions which resulted in the refusal. This information should be recorded in detail at the time of occurrence, reported back to the office, and also discussed at interviewer debriefing.

Interviewer debriefing

Well trained field interviewers are in a unique position to evaluate the merits of survey training, design, survey question structure and wording. Not only do they obtain very useful observed feedback from respondents in the course of administering questionnaires, more experienced interviewers can also draw on their accumulated knowledge to identify, during the pre-testing stage of questionnaire development, questions that are likely to be difficult for interviewers to read and/or for respondents to comprehend.

There are a variety of techniques that can be used to obtain information from interviewers about problems with a questionnaire. Interviewers can be debriefed in a group setting, through interviewer rating forms, or through standardized interviewer questionnaires (also referred to as structured post interview evaluations). These techniques are frequently used in conjunction with each other.

Debriefing interviewers in a group setting (similar to a focus group) is the most common method used during pretests. Interviewers who conduct the pretest are brought together after the interviewing is completed and asked about their experiences administering the questionnaire. Typically, the moderator of the debriefing will review the questionnaire item by item to identify any problems interviewers found with question wording, question sequencing and the overall flow of the interview.

If not all interviewers can participate in the debriefing, then where possible those that do participate should consist of interviewers with varying years of experience and levels of interviewing skill. This is important because newer interviewers, or interviewers who have not acquired good interviewing skills, may have different concerns about a questionnaire than experienced or well-skilled interviewers.

During interviewer focus groups it is critical that the moderator encourage participation of all attendees. He/she must ensure that a few participants do not dominate. It may be necessary for the moderator to solicit comments from the more timid participants, to insure that all views are represented. Group debriefing sessions should generally be held within a few days after interviewing is completed.

Another technique to obtain information from interviewers regarding pretest questionnaires is the use of interviewer rating or debriefing forms. After the pretest is completed, interviewers complete a standardized rating form and rate each question in the pretest questionnaire on selected characteristics of interest to the researchers. The focus is to quantify the extent of problems, and information about the reasons for the problems may or may not also be obtained.

The exact content of the debriefing form can vary according to the needs of the study. It would be beneficial to examine at a minimum the following three characteristics:

1. Interviewer has trouble reading the question as written;
2. Respondents don't understand words or ideas in the question;

3. Respondents have trouble providing answers to the question; and to obtain written feedback from the interviewers as to their perception of why problems/difficulties might have occurred. The frequency of occurrence should also be identified, as a rare problem would be less of an issue for correction than a common one.

Interviewer debriefing forms can be used in conjunction with group debriefings. When used in this way, they should be completed by interviewers prior to the group debriefing. Ideally the forms should be collected and examined before the face-to-face debriefing session, as this will give the debriefing staff time to use the information to frame discussion points for sessions within the debriefing meeting.

Debriefing questionnaires are designed to be self-administered and they are an extremely cost-efficient way to collect data from all interviewers participating in a field test. Both the design of the questionnaire and the structure of the debriefing process are affected by the number of interviewers involved in the field test. If the field test involves a small number of interviewers (less than 15), then the questionnaire can be designed with open-ended questions, since it will be possible for each questionnaire to be reviewed individually. Given the small number of interviewers and the open-ended design of the questions, the resulting analysis will be qualitative in addition to quantitative. However, if a field test involves a large number of interviewers and each questionnaire will not be individually reviewed, then the debriefing questionnaire should be designed with closed-ended questions, and the data should be entered into a database and quantitatively analyzed.

For the purposes of the WG testing, countries should instruct their interviewers to record their observations 'as they happen' to facilitate later summarizing and reporting in either a structured debriefing form, at a formal interviewer face-to-face debriefing session, or if time and funding allow, both procedures.