

UNESCAP Project on Disability Statistics

Field testing protocol



- Four studies:
 - One: Translation
 - Two: Cognitive testing
 - Three: Field test
 - Four: Test/Retest



1. Translation

- Objective:
 - Examine translatability of Ext Sets
 - Identify problematic concepts
- Process:
 - Committee translation: experts in linguistics, health and disability and interviewers
 - Review by 2 independent people
 - Comments reviewed and incorporated



2. Cognitive testing

- See presentation by Mitch Loeb and Kristen Miller



3. Field testing (1)

- Objectives: to test.....
 - Relationship between Short and Extended Sets
 - Cross country comparability
 - Individual hypotheses arising from Cognitive testing and translation



3. Field testing (2)

- Process:
 - Background documentation – field test protocol, interviewer training manual, question by question guide.
 - Selection of sample and interviewers
 - Training of interviewers
 - Data collection
 - Interviewer debriefing
 - Data capturing
 - Analysis



3. Field testing (3)

- **Sampling:**
 - as representative as possible in selection of area and within area – cover urban and rural, etc. Recommend around 1000 individuals
- **Interviewers:**
 - as experienced as possible and to provide information on the experience of the interviewers.
- **Information on response rates –**
 - total sample selected and final data collected – i.e. get a questionnaire for each selected individual even if only name/number and address.



3. Field testing (4)

- **Selecting respondents:**
 - Household member 5 years and older
 - In a household with children and adults – randomly select 1 child (5 – 17 years) and 1 adult 18 yrs +
 - In a household with only adults select 1 adult 18 yrs +
 - Selection of respondents: get names and dates of birth of all household members; determine who is eligible for selection; select adult and child with next birthday.



3. Field testing (5)

- **Proxy interviews:**

- all children (under 18 years)
- adults who cannot reply for themselves (e.g. very ill, unable to communicate or are cognitively impaired),
- Note reasons for doing proxy interview.
- Analyse reasons for proxy and relationship with responses given on the questions – e.g. communication, self care, cognition.



3. Field testing (6)

- **Interview process:**

- Provide translation together with English in questionnaire (?)
- Try and ask question to same as it is written down.
- Make notes on questionnaire when person struggles with a question



3. Field testing (7): Question sets

- **Question Set 1** The WG short set (expanded)
- **Question Set 2** The WG proposed extended question set
- **Question Set 3** Selected WHODASII and the three ABS 'Need for Assistance' questions.
- **Question Set 4 Country specific questions (if applicable)**

- Order effect: use of two or more versions



3. Field testing (8)

- Content of questionnaire:
 - Socio-demographic data
 - Responses for Question Set 1
 - Responses for Question Set 2
 - Time to administer the questionnaire



4. Test/Retest (1)

Objective

- to assesses reliability of individual questions by a standard rest-retest procedure. The aim is to flag unreliable questions in terms of inconsistency of response for later attention in the production of a recommended question set.



4. Test/Retest (2)

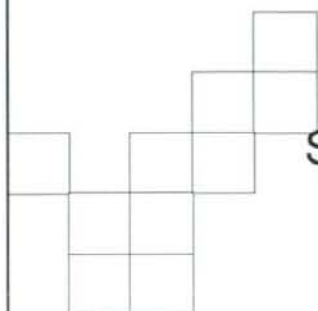
Procedure

- **Step 1:** Familiarise interviewers with the objective and documents to be used in Study 2.
- **Step 2:** Select 100 respondents who agreed to be interviewed again at the end of Study 1, ensuring that 50 used Set 1 and 50 Set 2.
- **Step 3:** Administer the questionnaire after a period of one week using a different interviewer. Ensure that the same version of the questionnaire is used for each respondent.
- **Step 4:** Submission of data recorded on the data entry program submitted to UNESCAP.



Eighth Meeting of the Washington Group on Disability Statistics

資料9



Summary of Annual Reports on National Activities Related to Disability Statistics

Cordell Golden
United States

Manila, Philippines
29-31 October 2008



Responding countries (n=43)

- **Africa (6):** *Ivory Coast, Kenya, Mauritius, Sierra Leone, Zambia, Zimbabwe*
- **Asia/Pacific (14):** *Cambodia, China, China (Macao), Fiji, Hong Kong SAR, India, Indonesia, Maldives, Mongolia, New Zealand, Philippines, Singapore, Sri Lanka, Vietnam*
- **Europe (10):** *Armenia, France, Germany, Hungary, Italy, Lithuania, Poland, Romania, Slovenia, Spain*
- **Central and South America (4):** *Brazil, Panama, Paraguay, Peru*
- **Middle East (4):** *Israel, Jordan, Sultanate of Oman, United Emirates*
- **North America (5):** *Bermuda, Canada, Cuba, Mexico, United States*



New or updated information for 2008

- **New information (18):** *Armenia, Cambodia, China, China Macao, Fiji, Germany, Indonesia, Israel, Ivory Coast, Maldives, Mexico, Peru, Sierra Leone, Singapore, Sri Lanka, Sultanate of Oman, Vietnam, Zambia*
- **Update of last year's information (17):** *Bermuda, Brazil, Canada, Cuba, France, Hungary, India, Italy, Kenya, Lithuania, Mauritius, Mongolia, Philippines, Poland, Romania, Slovenia, Spain*
- **No update from previous year's report (8):** *Hong Kong (SAR), Jordan, New Zealand, Panama, Paraguay, United Arab Emirates, United States, Zimbabwe*



Included short set of WG questions on national data collection activity

Yes 22 (51.2%)
No 21 (48.8%)



Plans to use short set of WG Questions in upcoming Census

Kenya (2009)

Philippines (2009 – Pilot Survey for 2010 Census)

Vietnam (2009)

Brazil (2010)

Mexico (2010)

Paraguay (2010)

United Arab Emirates (2010)



Reason short set of WG questions not included:

- *Disability is considered a sensitive subject respondents would be hesitant to answer*
- *Not responsible for deciding what questions are included on national surveys and census*
- *Disability data already collected from administrative sources*
- *Pre-testing within the country needs to be conducted*
- *Plans to use WG short set in future*
- *Use 1 or 2 questions from WG short set combined with other measures of disability*



Recent and upcoming national data collection activities related to disability statistics



Summary Statistics (# countries)	Type of Data Collection							Total
	C	S	AR	C/S	C/AR	S/AR	C/S/AR	
Plans to conduct national data collection activities on disability within the next 3 years	7	10	3	1	9	9	3	42
No plans to conduct national data collection activities on disability within the next 3 years	1							1

C = Census

S = Survey

AR = Administrative Records



Mode of Data Collection

Face to face interview	Self-completed questionnaire	Telephone interview	Direct observation and measurement	Other	Administrative Records only
35	7	8	2	6	2



Date of next data collection

2008

Bermuda
France
Ivory Coast (*June*)
Kenya (*August*)
Spain (*October*)
United States (*December*)
Israel (*December*)

2010

Singapore
United Arab Emirates
Zambia
Mongolia (*January*)
Mexico (*February*)
Vietnam (*May / September*)
Sultanate of Oman (*April*)
Mauritius (*June*)
Indonesia (*June*)
Brazil (*August*)

2009

Fiji
Germany
Sri Lanka (*January*)
Cambodia (*January*)
Vietnam (*April*)
Philippines (*May*)
Italy (*September*)
Sierra Leone (*September*)
Poland (*October*)
China (*October*)
Hungary (*December*)

2011

Lithuania
New Zealand
Maldives (*March*)
India (*March*)
Canada (*May*)
China Macao (*August*)

2013

Romania (*June*)



Frequency of data collection

- **One-time data collection (3):** *Bermuda, Cuba, Lithuania*
- **Annually (3):** *Cambodia, China, Mongolia*
- **Every 2-4 years (7):** *Germany, Indonesia, Israel, Ivory Coast, Sierra Leone, Vietnam*, Zambia*
- **Every 5-10 years (20):** *Brazil, Canada, China Macao, Fiji, France, Hong Kong SAR, Hungary, India, Kenya, Mauritius, New Zealand, Peru, Philippines, Poland, Romania, Singapore, Slovenia, Sultanate of Oman, Vietnam*, Zimbabwe*
- **Unknown (10):** *Armenia, Italy, Jordan, Maldives, Mexico, Paraguay, Spain, Sri Lanka, United Arab Emirates, United States*

** Vietnam has multiple data collection activities scheduled*



Sampling frame

- **Most recent or upcoming Census**
- **Country population registries**
- **Government administrative records**
- **Physician or long-stay facility patient lists**



Languages in which survey offered / administered

- *Arabic* (3)
- *English (with translation)* (18)
- *Chinese* (2)
- *French* (4)
- *German* (2)
- *Italian* (1)
- *National language of country* (26)
- *Portuguese* (2)
- *Samoan* (1)
- *Spanish* (5)



Other national activities related to disability data collection

- Pre-test of disability survey questions
- Training for data collectors
- Collection of administrative data
- Dissemination and publication of disability data
- Discussions with government agencies about disability data collection activities
- Training workshop at ISI (Durban - August 2009)

Development of an Internationally Comparable Disability Measure for Censuses Washington Group on Disability Statistics (WG)

A new set of questions on disability for use on national Censuses has been developed, tested and adopted by the Washington Group on Disability Statistics (WG). The WG is a United Nations (UN) sponsored City Group commissioned to improve the quality and international comparability of disability measures. The approach taken by the WG represents a break with methods used in the past. The questions reflect advances in the conceptualization of disability and use the World Health Organization's International Classification of Functioning, Disability, and Health (ICF) as a conceptual framework. The focus is on functioning in basic actions in contrast to approaches that are based on impairments or bodily functions. The UN Principles and Recommendations for Population and Housing Censuses incorporates the approach taken by WG. (See: Section VI-8: Disability Characteristics pages 178-183, and Tabulations on Disability Characteristics pages 292-294; available online at: http://unstats.un.org/unsd/demographic/sources/census/docs/P&R_Rev2.pdf). A more detailed discussion of the conceptual framework and data collection objectives of the WG can be found in the Washington Group Position Paper: Proposed Purpose of an Internationally Comparable General Disability Measure (WG3.6), available online at: <http://www.cdc.gov/nchs/about/otheract/citygroup/meeting3.htm#papers>.

Background:

In June of 2001, the United Nations International Seminar on the Measurement of Disability identified the need for comparable population-based measures of disability for individual country use and for international comparisons. This determination was based on the scarcity and general poor quality of data on disability, especially in developing countries, and the lack of internationally comparable measures, even among developed countries. The Washington Group on Disability Statistics (WG) was formed to address this urgent need.

The main purpose of the WG is to promote and co-ordinate international co-operation in the area of health statistics focusing on disability measures suitable for censuses and national surveys. The major objective is to develop tools to collect the basic data necessary to provide information on disability that is comparable throughout the world. The first priority of the WG was to guide the development of a short set of disability measures suitable for use in censuses, sample-based national surveys, or other statistical formats, for the primary purpose of informing policy on equalization of opportunities for the population with disabilities. A second priority is to recommend one or more extended sets of survey items that elaborate the measurement of the multiple concepts associated with disability and can be used as components of population surveys, as supplements to surveys or as the core of a disability survey. There will be a connection between these extended sets of survey items and the short set of disability measures. The disability measures recommended by the group will be accompanied by descriptions of their technical properties, and methodological guidance will be provided for their implementation and their applicability to all population subgroups.

I. Recommended Short Set of Questions on Disability for Censuses

At the 6th Annual Meeting of the WG in Kampala, Uganda in 2006, test results were reported and the short set of questions on disability was endorsed by the 23 countries and 5 international agencies in attendance. Some minor wording modifications were suggested based on pre-test results presented at the meeting (see Section V below). The set comprises questions on six core functional domains: seeing, hearing, walking, cognition, self care, and communication. In countries where resources do not permit inclusion of 6 questions on a census, the first four domains are recommended for inclusion (seeing, hearing, walking, and cognition); however the WG strongly endorses the use of the six questions.

The final short question set is:

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?¹
2. Do you have difficulty hearing, even if using a hearing aid?¹
3. Do you have difficulty walking or climbing steps?
4. Do you have difficulty remembering or concentrating?
5. Do you have difficulty (with self-care such as) washing all over or dressing?
6. Using your usual (customary) language, do you have difficulty communicating, (for example understanding or being understood by others)?

Each question has four response categories: (1) No, no difficulty, (2) Yes, some difficulty, (3) Yes, a lot of difficulty and (4) Cannot do it at all. The severity scale is used in the response categories in order to capture the full spectrum of functioning from mild to severe.

II. Rationale for Choice of Questions

Disability represents a complex process and is not a single, static state. It refers to the outcome of the interaction of a person and his/her environment (physical, social, cultural or legislative) and represents a measure of the negative impact of environmental factors on one's ability to function. The complexity of the concept has resulted in the proliferation of statistics on disability that are neither comparable nor easy to interpret. Furthermore, disability data are collected for different purposes such as to estimate the prevalence of physical impairments or to plan for the provision of services. Each purpose elicits a different statistic and even when the intention is to measure the same concept, the actual questions used differ in ways that severely limit comparability. The conclusion is not that some estimates are right and others are wrong, but that they are measuring different things. The WG chose to develop questions that would address the issue of whether persons with disability participate to the same extent as persons without disabilities in activities such as education, employment or family/civic life. A major reason for this choice is the pivotal importance of the issue of social participation and equal rights from a policy perspective as illustrated by the recently ratified UN Convention on the

¹ The inclusion of assistive devices was considered for two domains only, seeing and hearing, as limitations in these domains can often be overcome with the use of glasses or hearing aids.

Rights of Persons with Disabilities². In addition, there was agreement that it would be possible to develop a question set to meet this objective, that could be administered using Census methodology and that could produce internationally comparable data.

One approach to measuring social engagement is to ask directly if a disability has impacted participation. An example of such a question is “Are you limited in the kind or amount of activities that you can do because of on-going difficulties due to long term physical, mental or health problems?” Such questions are difficult to ask in a way that produces comparable data. An alternative approach is to obtain information on difficulty in functioning in basic actions (c.f. the WG short set of 6 questions above) since these actions form the building blocks for more complex activities and, when restricted by the environment, can result in disparities in participation. The task is then to determine whether persons with difficulties or limitations in basic actions have participation rates equal to those without these limitations.

The WG questions were designed to provide comparable data cross-nationally for populations living in a variety of cultures with varying economic resources. While the ideal would be to collect information on **all** aspects of the disablement process and to identify every person with a disability within every community, this would not be possible given the limited number of questions that can be asked on a National Census. The basic actions represented in this question set are those that are most often found to limit an individual and result in participation restrictions. Domains were selected using the criteria of simplicity, brevity, universality and comparability. It is expected that the information that results from the use of these questions will, a) represent the majority of, **but not all**, persons with limitation in basic actions, b) represent the most commonly occurring limitations in basic actions, and c) be able to capture persons with similar problems across countries.

III. How to Use the Data

As mentioned above, this short set of questions will identify the majority of persons with a disability – but not all; and they will identify most limitations – but not all. Psychological limitations, for example, are not covered by the short set of questions. Neither were the WG questions developed specifically for use with children. Obtaining information on disability in children, especially in young children, using a limited number of questions is very difficult. These limitations would preclude the use of the short set of questions as screeners for disability in a population.

However, the recommended set of questions identifies the majority of the population with difficulties in functioning in basic actions that have the **potential** to limit independent participation in society. The intended use of these data is to compare levels of participation in, for example, employment, education, or family life for those with disability versus those without disability and thereby to assess equitable access to opportunities as mandated by the UN Convention. In addition, the data can be used to monitor prevalence trends for persons with limitations in the specific basic action domains.

² United Nations Enable: Rights and Dignity of Persons with Disability, <http://www.un.org/disabilities/>

The WG recognizes that the short set of questions for censuses may not meet all the needs for disability statistics, nor will it replicate a survey of the population that can collect information across a wider range of disability domains. While a census can provide valuable information on disability especially for local areas, other data collection mechanisms are necessary to obtain a more complete understanding of disability nationally and internationally.

The WG is currently developing extended question sets for use on surveys.

IV. Context for the work of the WG

The finalization of the short question set will facilitate the inclusion of the questions in the 2010 census round. The questions were developed according to the Fundamental Principles of Official Statistics³ and are consistent with the International Classification of Functioning, Disability and Health⁴. Most importantly, however, the endorsed questions support the UN Convention on the Rights of Persons with Disabilities. The short set addresses equalization of opportunities for persons with disabilities which is one of the General Principles listed in Article 3 and is the focus of Article 5. It is also particularly relevant to the collection of data for policy purposes outlined in Article 31 and will facilitate the monitoring of participation in cultural life, leisure, recreation, work, and employment that is called for in Articles 27 and 30 of the Convention.

V. Question Testing Protocol

Pre-test process

In addition to developing the short question set, the WG also developed a plan for both cognitive and field testing of the questions.

Standardized testing was undertaken in 15 countries, including 13 that were funded through the World Bank⁵ to assess the validity of the questions and to better understand how they operate in a variety of regions, cultures, and languages. Cognitive tests were designed to provide insight about how respondents comprehend, retrieve, judge, and respond to questions. Cognitive tests were also intended to identify response errors related to question design and to reveal how and why these errors may have occurred. These tests involved in-depth, face to face interviews with a small sample of respondents representing the group of interest. For the purpose of the WG test, the interviews were semi-structured and the analysis of results was largely qualitative. Field tests were conducted to determine whether the questions were being interpreted as intended by the developers in that they captured the most relevant key aspects of the functional domains selected (i.e. whether the single question per domain captured a reasonable proportion of those with functioning difficulties in that domain). Field tests were intended to simulate the conditions of an actual survey. The objectives of the WG field test were to determine whether the single question per domain was representative of that domain, whether the questions produce comparable data across countries, and how the WG questions work as a set in comparison with other questions used by the country.

³ See *Statistical Commission, Report on the Special Session (11-15 April 1994)*, Economic and Social Council, Official Records, 1994, Supplement No.9, Series No. E/CN.3/1994/18, United Nations, New York, 1994, para.59.

⁴ *International Classification of Functioning, Disability and Health* (World Health Organization, Geneva, 2001)

⁵ Funding secured through the World Bank Development Grant Facility

Summary of results

Based on combined test results from all countries, the WG questions were well understood and interpreted consistently across countries. Only a few cases of inconsistent response patterns for each of the six WG questions were observed. Most notable were 'literal' translation difficulties from English into local languages and it was suggested that better conceptual translation was needed since certain phrasing is not easily interpretable in some languages. Countries are instructed to translate these phrases in a way that is culturally appropriate to capture the concepts in the question. In earlier versions, the communication question (Question 6) had an introductory text that was considered somewhat cumbersome. In order to reduce some erroneous responses and to improve precision, this particular text was removed since it repeated the introduction to the entire set of questions. Furthermore, some respondents to this question reported a second language problem (i.e. they had communication problems because the primary language in the country was not their native language), therefore, a reference to communicating in one's normal language was added in the preface to this question.

Correspondence between each WG question and the extended questions for each domain was generally good. In some instances respondents answered affirmatively to detailed questions about a particular domain of functioning, but did not respond affirmatively to the WG question for that domain, and *visa versa*. For instance, a respondent may have reported some difficulty with near or far vision, but did not report difficulty seeing in general as captured by the WG question on vision. Misidentification can occur for a number of reasons and it is important to determine whether false positives or negatives are occurring systematically (true errors) or randomly. For cases identified as true errors, it is furthermore important to determine if these are associated with gender, country, age, disability or health status. True error was identified for the vision question but the error was related to the glasses clause rather than the result of misunderstanding of the question. There is a potential for false positives in extended questions on cognition as a high rate of inconsistencies was found; however, unlike the vision question, the inconsistencies are more likely to be a result of interpretation issues and not blatant misunderstanding.

There was discussion about whether the WG questions should be more specific, for instance, by adding a clear reference to distance for the question on difficulty walking (Question 3). However, the overall level of error with the WG questions was small and bias in responses due to contextual differences across countries did not appear to be a problem. It was agreed that increasing the specificity of the questions would not be beneficial, particularly in a census format and given the potential for cultural and contextual differences across countries.

In certain instances the WG question represents only a portion of the entire domain of functioning. In terms of cognition (Question 4), for example, memory and concentration do not represent all aspects of mental functioning. In addition to memory and concentration, the extended questions included aspects of learning new tasks and finding solutions to problems.⁶ Field testing revealed good correspondence between the WG question and the extended

⁶ Extended questions: Do you have difficulty (i) concentrating on doing something for 10 minutes? (ii) learning a new task, for example, learning how to get to a new place? (iii) finding solutions to problems in day to day life? (iv) remembering the names of people or places? (v) remembering appointments? (vi) remembering how to get to familiar places? (vii) remembering important tasks, like taking medications or paying bills?

questions that captured *similar* aspects of functioning (memory and concentration) but poor correspondence between the WG question and the extended questions that captured *different* aspects of mental functioning (learning and finding solutions). This was interpreted as a limitation in terms of the aspects or scope of functioning in the domain captured by the WG question, but not a problem with the question itself.

It was anticipated that the question on self-care (Question 5) would be somewhat problematic since it was oriented toward more complex activities than the other WG questions (and therefore potentially more culturally influenced); however, this did not seem to pose a problem when data were compared across countries.

Overall, the inconsistencies identified in the field and cognitive tests were not considered to be significant problems; the number of false negatives and false positives were few and responses to the WG questions appeared to reflect respondents' overall abilities in that domain. The testing suggested that potential errors could be minimized by addressing issues of inconsistency at the country level, ensuring that questions are interpreted and translated according to the customary language of the country and providing country-specific cognitive testing in terms of question development. A single question per domain was deemed appropriate to capture those with functioning difficulties in that domain. The questions also captured the intended aspects of the functional domains selected, although not all aspects of those domains. It should be reiterated, however, that **not all** functional domains are covered by the six WG questions; for example, no information is collected that deals with psychological limitations or difficulties.

In further analyses using data from existing national surveys from developed countries where questions similar to the WG questions were included as well as questions on many other aspects of functioning, the domains captured by the WG questions represented the majority (but not all) of respondents with self-reported limitations in any aspect of functioning.

These tests, as well as studies in other countries employing the WG approach, show an improvement over use of more traditional impairment focused census questions on disability. For example, according to the 2000 Census in Zambia⁷, disability referred to individuals who were limited in the kind or amount of activities they could do because of on-going difficulties due to long term physical, mental, or emotional health problems. The actual questions used to capture disability in that Census, however, were: "Are you disabled in any way?" (Yes/No), and "What is your disability?" (Response categories included: blind, partially sighted, deaf/dumb, hard of hearing, mentally ill, ex-mental, mentally retarded, and physical handicapped.) This approach yielded a disability prevalence rate in Zambia of 2.7% and represented, in fact, a trebling of the 1990 prevalence rate of 0.9% which used the same approach but included only 4 impairment categories: blind, deaf/dumb, mentally retarded, and crippled^{7,8}. When the short set of WG questions was included in the 2006 Living Conditions Survey in Zambia⁹, 13.3% of

⁷ CSO, 2000 Census of Population and Housing, Available online at: <http://www.zamstats.gov.zm/census.php> see Chapter 9: Disability; and http://unstats.un.org/unsd/demographic/sources/census/quest/ZMB/Eng2000_ind.pdf

⁸ CSO, 1990 Census of Population, Housing and Agriculture, Available online at: http://international.ipums.org/international/world_census_forms/zambia_1990.pdf

⁹ Eide AH, Loeb ME (eds.) (2006) Living Conditions among people with activity limitations in Zambia: A national representative study. Report No. A262, SINTEF Health Research, Oslo. Available online at: <http://www.sintef.no/le>