**Cognitive interviewing in the Disability Pilot Survey in Spain**

Rogelio Pujol Rodriguez, National Statistics Institute of Spain, rogelio.pujol.rodriguez@ine.es

Angeles Lora-Tamayo D’Ocon, National Statistics Institute of Spain, angeles.lora.tamayo@ine.es

**Abstract**

*Cognitive interviewing is a technique that helps ensure that a survey questions capture successfully the scientific aims by identifying potential measurement errors and question wording problems (overall understanding and interpretation of the question, convenience of the examples given, temporary references, type of scoring at an item level, etc.)*

*This method has proven to increase the quality of the collected data. Hence we conducted several cognitive interviews on two questionnaire models prior to the fieldwork of the Disability Pilot Survey. The aim was to analyse the comprehension and quality of the questions and correctly collect those homes where people with disabilities live, and their limitations.*

*For this purpose voluntary subjects with specific characteristics of interest were recruited and interviewed in a laboratory environment. Some of them were asked to respond to a previously assigned questionnaire model via web (CAWI: Computer-assisted web interviewing) while others were telephoned (CATI: Computer-assisted telephone interviewing). All of them were asked afterwards about their responses and feelings while answering the questionnaire.*

*As a result of these interviews, most of the participants preferred the second model questionnaire (the one with multiple response format), several questions changed their wording (the known GALI question was rephrased) and we confirmed the importance of the examples given at the end of each one of the specific disability questions.*

**Keywords:** Pretest, data collection, survey, disability, health

**1. Qualitative methodology in the evaluation of questionnaires: Cognitive pretest and cognitive interviews.**

In the process of evaluating and optimizing the information obtained through a questionnaire, it is becoming increasingly common to study the cognitive behavior of the respondent. The administration of the questionnaire is thus conceived through the *question-response model* (see Figure 1).

**Figure 1. Question-response model: the cognitive process in answering questions**

Source: own elaboration

The so-called Cognitive Pretest groups together a set of procedures (Cognitive Interviews, Behavioral Coding, Discussion Groups) that allow us to provide evidence on the achievement of this model, specifically on problems in the understanding of key terms, the retrieval of information to elaborate an answer, the choice of the answer option, etc. These problems are mostly due to measurement errors (deficiencies of the questionnaire, insufficient definitions), which belong to the set of non-sampling errors and, within them, to those not related to fieldwork.

For the purpose of this article, we centred on Cognitive Interviews. This procedure consists of administering a draft questionnaire focusing on the verbalization of participants' thoughts while they are responding (think-aloud method), conducting probing tests on selected questions (probing-based method) or based on coding schemes (the results are summarized in different categories).

**2. Cognitive interviewing in the Disability Pilot Survey in Spain**

***2.1. Objective***

The Survey on Disability, Personal Autonomy and Dependency Situations (EDAD in Spanish) is a survey carried out by the National Statistics Institute at the state level, designed to obtain data on the characteristics and situation of persons with disabilities who live in private households or dwellings. Taking into account that since the last survey (EDAD 2008) new data collection methods have been introduced, a Pilot Test is being carried out to study their application to a two-stage procedure (phase 1: Internet (CAWI) and/or telephone (CATI) collection; phase 2: Computer Assisted Personal Interviewing (CAPI)). To this end, 4 questionnaires with different filter questions (two models) and different answers (dichotomous and polytomous or multiple) were tested. Each questionnaire consists of 4 blocks:

* Block A: Composition of the dwelling
* Block B: Socio-demographic data on residents in the dwelling
* Block C: Questions about Limitations (GALI) and Certification of Disability
* Block D: Questions about disabilities (filter questions)

Before carrying out the corresponding pilot survey, it was decided to perform a qualitative study (Cognitive Pretest) on the four models of questionnaires that will, in principle, be used. These models were generically called MXRY, with X=1,2 and Y=1,2. The M1R1 model was not considered.

The MX models differ in the number of questions in Block D, and the RY models differ in the type of answer (dichotomous if Y=1, multiple or politomous if Y=2) to the questions in Block D.

This Pretest, carried out by the General Subdirectorate for Data Collection (SGRD in Spanish), should shed light on whether the understanding and quality of the filtering questions of the previous models via CAWI and/or via CATI meets the same objectives as the household questionnaire used in EDAD 2008, that is, it allows us to capture the households where people with disabilities live and their disability or disabilities. For example, respondents were asked about their knowledge of the disability certificate (Block C); their preference between dichotomous (Yes; No) or multiple answers (Yes, total difficulty; Yes, a lot of difficulty; Yes, some difficulty; No, no difficulty) to the questions in Block D; also on the understanding of certain texts and temporal references in certain questions, such as "keeping the body in the same position for the necessary time" (mobility question), "activities that people usually do" (GALI indicator question in block C), the concept of "temporarily absent person" (Block A), etc.

***2.2. Planning process***

The informants who have collaborated in this study came from different sources: collaborators of the Continuous Household Survey, collaborators of the Employment of Persons with Disabilities Survey, contacts facilitated by SGRD personnel and others facilitated by different associations of persons with disabilities.

This qualitative sample of people, unlike the statistics samples, follows a structural sampling criterion and does not seek to be statistically representative but rather to identify profiles of people whose specific characteristics may have an impact on their attitudes and skills in completing the questionnaire and on their view of disability and dependency. Traditionally used characteristics (age and gender) as well as the level of studies were taken as influential characteristics.

The respondent was first contacted by telephone and invited to participate in the study, without giving the respondent any indication of the subject matter of the questionnaire. In case of contacting for a CATI interview, they were also informed that they would have to go to the INE headquarters to continue the interview with the interviewer who attended them on the telephone.

At all times, the informants were informed about the confidentiality of the procedure guaranteed by the Statistical Secret, which obliges the statistical services not to disseminate, under any circumstances, the personal data of the informants, whatever their origin (Chapter III of the Law on Public Statistical Services).

***2.3. Execution***

The first interview was conducted on 28 April 2017 and the last one on 25 June 2017. In the interviews, a mix of probing-based method and coding schemes was followed. The nature of the interviews was "retrospective", the cognitive inquiry was subsequent to the administration of the questionnaire.

**CAWI Interviews**

Once the appointment was accepted and arranged for completion via the Internet, the collaboration developed in two parts:

1º- The informant carried out the CAWI questionnaire in INE headquarters, in a computer room with access to the IRIA portal via the Internet. They were given a letter like the one they would receive at home, with the Web Portal address and the access codes. While the informant was filling in the questionnaire, in an adjacent room, the team involved visualized the completion, using video-conferencing software, in order to see the difficulties and doubts that the informant had.

2º- At the end of the online questionnaire, a personal interview was conducted asking about the feelings, difficulties and opinions on some of the questions of the questionnaire and the usability of the CAWI application (length of the questionnaire, difficulty in accessing the application, interruptions that occurred, difficulties in navigating between screens, etc.).

**CATI Interviews**

Once the telephone appointment was accepted and arranged, the collaboration developed in two parts:

1º- The informant was telephoned and the CATI questionnaire was carried out.

2º- The same day or the day after the telephone interview, the informant went to INE headquarters to conduct a personal interview, asking about the feelings, difficulties and opinions on some of the questions of the CATI questionnaire.

Both CAWI and CATI interviews were recorded (video and audio) for further analysis.

22 informants (12 women and 10 men) attended the appointments, one of the informants ended with a refusal (CATI); in addition, a CATI interview was conducted but was not included in the final count as the informer did not show up for the personal interview.

The average age of the informants was 48 years old, living in an average household size of three people. Most of them had university or higher education. The average time to complete the questionnaire was 21 minutes (20 minutes on CATI channel and 22 on CAWI channel). Tables 1 to 3 present the distribution of informants by origin, level of studies and type of questionnaire provided and channel.

**Table 1. Origin of the informants**

|  |  |
| --- | --- |
| **Origin** | **Frequency** |
| Contacts facilitated by SGRD personnel | 15 |
| Collaborators of the Continuous Household Survey | 4 |
| Contacts facilitated by associations of persons with disabilities | 2 |
| Collaborators of the Employment of Persons with Disabilities Survey | 1 |
| **Total** | **22** |

Source**:** own elaboration

**Table 2. Level of education of the informants**

|  |  |
| --- | --- |
| **Level of education** | **Frequency** |
| University or higher education (PhD) | 11 |
| Vocational Training | 3 |
| Baccalaureate | 3 |
| Compulsory Secondary Education | 2 |
| School Graduate or Basic General Education | 3 |
| **Total** | **22** |

Source**:** own elaboration

**Table 3. Number of interviews by model of questionnaire provided and channel**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model\Channel** | ***CATI*** | ***CAWI*** | **Total model** |
| ***M1R2*** | 3 | 4 | **7** |
| ***M2R1*** | 0 | 2 | **2** |
| ***M2R2*** | 6 | 7 | **13** |
| **Total channel** | **9** | **13** | **22** |

Source**:** own elaboration

***2.4. Results***

Table 4 shows the average completion times by type of questionnaire supplied and channel.

**Table 4. Average completion time (minutes) by model of questionnaire provided and channel**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model\Channel** | ***CATI*** | ***CAWI*** | **Total model** |
| ***M1R2*** | 16 | 17 | **17** |
| ***M2R1*** | - | 20 | **20** |
| ***M2R2*** | 22 | 25 | **24** |
| **Total Channel** | **20** | **22** | **21** |

Source**:** own elaboration

**Understanding and quality of questions**

Most of the informants preferred the multiple-response model (R2 model) to the dichotomous response model (R1) because they saw the possibility of indicating the degree of their limitation better (in the R1 model many of those who would not claim to have a limitation in principle responded to have one).

The question in the Learning section, in model M1, has been understood, in some cases, as a "Mobility" problem. This is due to the concatenation of verbs, which makes the informant unable to understand the main action. The initial text of this question was: Does anyone in the household have difficulty in learning or developing tasks, without help or supervision? For example, paying attention, learning to add, reading, handling everyday utensils, passing on a message or keeping an appointment.

It was questioned, a priori, whether the example of "newspaper letter" (in the Vision questions of both models) could be misinterpreted by digital newspaper. All informants understood a newspaper to mean a newspaper on paper.

The importance of giving some examples after a question has been noted. This is most evident in the M1 model, where the number of examples in each question is higher. When presenting the question and the examples, it should be decided in the future whether to give more importance to the question, by presenting a small number of examples, or more importance to the examples, including, therefore, a list that is as complete as possible.

There have been informants who, because they did not see their specific case in any example, have marked NO in the corresponding question in block D on disabilities. Others have answered the question according to the level of each of the examples, and answered by having in mind the one that is most difficult for them.

It was found that proxies, and especially parents, play down their child's disability.

The cognitive interviews also allowed to improve the usability and navigation aspects of the web questionnaire to be improved: the order of some screens was changed, the size of the button panel to move through the questions of the questionnaire was increased.

**Changes made during the pretest**

During the first interviews, some problems were noted in some of the questions. Therefore, it was decided to modify the wording of some of them, which are shown below:

Both models, Block C, question C1 (GALI question)

* Previous text: 'Due to a health problem or disability, has anyone in the household been limited for a period of 6 or more months to perform the activities that people usually do?’
* New text: 'For at least the last 6 months, has anyone in the household been limited in activities people usually do because of a health problem?’

Model M1, Block D, question D2

* Previous text: '... to understand what is said in a conversation between several people or to hear alarms or sirens...'.
* New text: '... to listen to what is said in a conversation between several people or to hear alarms or sirens...'.

Model M2, Block D, Question D4

* Previous text: '... to hear and understand what is said in a conversation between several people...'.
* New text: '... to hear what is said in a conversation between several people...'.

**3. References**

Conrad, F.G. and Blair, J. (2009). Sources of error in cognitive interviews. Public Opinion Quarterly, 77, 32-55.

Eurostat. (2015). Global Activity Limitation Indicator (GALI) as a core variable. Directorate F: Social Statistics. Luxembourg: European Commission. DSS/2015/Sept/04.3, 42

Foddy, W.H. (1992). Constructing Questions for Interviews and Questionnaires. Melbourne: Cambridge University Press.

Groves, R. M. (1989), Survey errors and survey costs. New York: Wiley.

Napoles-Springer, A.M., Santoyo-Olsson, J.,O’Brien, H. and Stewart, A.L. (2006), Using cognitive interviews to develop surveys in diverse populations. Medical Care, 44, 21-30.

Willis, G. B. (2005), Cognitive interviewing: a tool for improving questionnaire design, Thousand Oaks, Calif.: Sage Publications.

World Health Organization. (2001). International classification of functioning, disability and health: ICF. Geneva, Switzerland: Author.