ONLINE COGNITIVE INTERVIEWING. A CONTRIBUTION TO DATA QUALITY

Sabrina Barcherini, Katia Bontempi, Serena Liani, Barbara Lorè, Simona Rosati

1. Introduction

Survey data quality depends on the choices made in the questionnaire design: which concepts are to be measured in the population and how these concepts are translated into survey questions. Conducting a questionnaire pretest allows to identify the flaws in the design which could cause response error.

Cognitive interviewing is a pretesting method for investigating sources of errors in the sequence and wording of questions, and in the structure of responses. The interview is usually face-to-face, but the spread of the Covid-19 pandemic has imposed to experiment innovative solutions.

In this paper, we report on one of the first experiences of online cognitive pretest conducted at the Italian National Institute of Statistics (Istat), which took place in 2020 for the start of a new survey.

Within the framework of a research project on discrimination at work for sexual orientation and gender identity, from December 2020 to March 2021 Istat, along with the National Office against Racial Discrimination (UNAR), has run a survey on discrimination at work against people who were or had been in civil partnership.

The aim of this survey was to provide an informative framework on the perception and the diffusion of discrimination, threats, and aggression that LGBT people in civil partnership (or who have been in the past) may have suffered in the Italian labour market.

Several aspects of the work environment were explored: job search, work activity, climate and relationships in the workplace or actions taken following any episodes of discrimination.

Since the topic was sensitive, self-completion was chosen as survey mode. The survey was therefore carried out using Computer Assisted Web Interviewing (CAWI) technique.

Due to the novelty and sensitiveness of the topic, particular attention was devoted to the questionnaire design. Subject matter experts drew the first version of the questionnaire based on the needs of data users, then survey methodologists revised it with the aim of reducing respondents burden and making it easier to fill it in. A

first phase was aimed at improving the sequence of the questions, their wording and response categories; it was followed by an accurate pretest activity to understand how people would react to the questionnaire (Balì *et al.*, 2019). Before the questionnaire was released, a cognitive test was carried out to check whether researchers and respondents shared concepts, language, and definitions (Converse and Presser, 1986).

To cope with the social distancing rules due to the health emergency from Covid-19, an innovation was introduced: the cognitive interviews were conducted online instead of face-to-face.

2. Cognitive interviewing practice

Cognitive interviewing is a pretesting method to identify and detecting any difficulties with the questionnaire items and the survey questions, in order to improve data quality. As Beatty states (2004, p. 45) «cognitive interviewing is the practice of administering a survey questionnaire while collecting additional verbal information about the survey responses; this additional information is used to evaluate the quality of the response or to help to determine whether the question is generating the sort of information that its author intends». In this perspective, cognitive interviewing provides questionnaire designers with significant insights about the effects of some questionnaire design decisions, the advantages and the disadvantages of asking questions in a specific way.

Two main techniques can be used to carry out a cognitive interview: think-aloud interviewing and verbal probing technique. In the first technique, respondents are requested to spontaneously verbalize all their thought processes as they answer survey questions; in the second one, survey questions are followed by a set of probe questions to understand more about respondents' thought processes. These techniques involve a different role of the interviewer. In the think-aloud interviewing the interviewer intervenes as little as possible, in order to facilitate the verbalization of the participants' thought processes. In the verbal probing technique, the interviewer «guides the interaction more proactively, generally asking additional, direct questions about the basis for responses» (Beatty and Willis, 2007).

Although the two approaches seem distant, «in practice, think-aloud and verbal probing actually fit together very naturally» (Willis, 2005, p. 57). It is appropriate to think about these techniques not as alternative but complementary and consider the possibility to adopt both methods (Conrad and Blair ,2001; Willis, 2005). In fact, not all interviewees have the same verbal ability and different questions of a same questionnaire could need to be tested in different ways. There are some questions that, by their nature, spontaneously encourage verbalizations from the respondents

(for example the questions that require the recall of past behaviours), while for others it is more difficult to follow the thinking strategy (for example the questions that detect opinions). For these reasons, it would be necessary to foresee an alternation of techniques already in the phase of the cognitive interview design because the best choice depends on the type of data that are being collected and the role of the interviewer in that process.

Furthermore, the cognitive pretesting design requires other relevant decisions involving the definition of the sampling plan, the number of interviews to be performed, the procedures for selecting and training the interviewers, the methods of administering the cognitive interviews and the analysis techniques to be adopted. Each of these decisions contributes to influence the results of the pretest and the quality of the subsequent data collection process.

Cognitive interviewing is usually conducted in person but the spread of the Covid-19 pandemic, combined with the availability of current web technologies, have encouraged the use of videoconferencing systems. Indeed, in the last times some researchers have begun to experiment the videoconferencing systems to carry out online cognitive interviews both in qualitative and in quantitative research.

3. The cognitive interview track

As it was mentioned earlier, in 2020 Istat engaged in remote interviewing to pretest the questionnaire on discrimination against LGBT people.

After an initial review, in which the focus was on the questions order, the standardization of the classifications and the wording of all the textual elements (i.e., questions, instructions for completion, warnings, definitions, etc.), some questions still showed some critical issues. Some of them were somewhat ambiguous, others were difficult to respond because of their structural complexity, others posed a challenge to memory processes. They were likely to lead respondents to misinterpret the question content, to underreport events or to provide careless responses. For all these reasons, these questions were selected to be tested.

The cognitive test had several aims:
- to make sure that survey questions were easily and properly understood;
- to check whether respondents and researchers shared concepts and definitions;
- to find out if survey questions contained inappropriate assumptions;
- to assess the capability of respondents to recall events and to map them onto the response categories;
- to verify whether response categories were exhaustive and mutually exclusive;

- to explore whether questions content or wording were perceived as embarrassing or too intrusive;
- to detect any other possible issues and to make sure that nothing relevant was missed.

As cognitive interviewing is based on a semi-structured in-depth interview, it is quite demanding. Thus, it is necessary to select only the questions that are considered more problematic or that are expected to be more insightful when submitted to the respondents' judgement.

To the aim of this pretest, twenty-two questions of different formats were selected: nine multiple choice, seven multi-response, three grids, and three open-ended questions. For each question, an interview track summarized the aim of the test, suggested a cognitive technique and included a number of concurrent probes.

In choosing the questions, an effort was also made to ensure a logical flow and to avoid exceeding a reasonable interview duration.

Throughout the interview we asked participants:
- to think-aloud while answering the questions;
- to explain what certain words meant to them;
- to tell how they got to answer memory questions;
- to rate the degree of confidence they had in their answer;
- to rephrase the questions in their own words.

The cognitive interview track included an introduction to explain to respondents the pretest aims and its role within the entire survey process, to obtain the participants' consent to take part in it and to explain what they were expected to do. Before starting the interview, respondents were also reassured that their identity would remain anonymous, and the information provided confidential.

4. Sampling and recruitment

Since the questionnaire was targeted at people entered in the Civil Union Registry, we needed to identify subjects who were representative of the target population, for which the survey questions were designed.

To reach the suitable people, we used snowballing, or chain sampling method, that is «a useful approach to implement when you need to find quite specific, or even hidden, populations» (Collins and Gray, 2015, p. 92).

We started by approaching some people in our circle of acquaintances who were in a civil union and asking them to help us with snowballing. We sent them the following WhatsApp message to explain the reason why they were contacted and how to make their contribution to the study:

> *Hello,*
> *The Italian institute of statistics is about to start a survey on discrimination at work. We are looking for people who are, or have been, in a civil union to take part in an online exploratory interview. If you want to make your contribution, <u>click this link</u> to learn more. Please, forward this message to any people who might be interested.*
> *Thank you!*

The link pointed to an online screening questionnaire[1] that asked for information useful to check the eligibility of the respondents. We were looking for people of 18 years or older, who were or had been in a civil union, and who had a job in Italy or had had it in the past. The screening questions also served to ensure a diverse range of participants to the cognitive testing. As recommended by Willis, our focus should have been «on subject variation across a range of characteristics, as opposed to statistical representativeness» (2005, p. 140). The subjects to be interviewed had to reflect the heterogeneity of the target population and allow us to evaluate those survey questions targeted at employees or self-employees. Thus, the screening questionnaire also asked for the respondent's age, gender, occupational status, and geographic location (Northern, Central or Southern Italy).

Respondents found eligible were asked to provide an e-mail address or a phone number where we could reach them for further contacts. In total, the screening questionnaire was completed by twenty-seven people, four of whom were not eligible for the cognitive interview. A phone or an e-mail message was sent to the remaining twenty-three people to arrange an appointment for the cognitive interview. The e-mail provided detailed information about the study and what taking

---

[1] The screening questionnaire was developed using the LimeSurvey system, a tool to create online surveys.

part in it involved; those who did not enter an e-mail address got the explanation on the phone.

Hello,
we are writing in replay to your application for the preliminary phase of the Survey on discrimination at work. We would like to meet you on the Whereby platform for an online interview. You don't need to download any program, just click the link that we will send you by email on the day of the interview. You will be requested to register, even with a fake name, and to click the Knock button to enter the chatroom. We suggest that you use a PC for a better view of the questions. We will have a chat of around 45 minutes about the questionnaire we are designing. Please, replay to this email to indicate days and times of the next week when you would prefer to take the interview.
Thank you.

Twelve people replied, but two of them declined the invitation shortly before the cognitive interview. The recruited subjects were distributed as follows: six men and four women; aged between thirty-six and fifty-six years (average age 46.5 years); all of them had a job in Italy, seven as employees and three as self-employees; eight lived in Central Italy and two in the North.

5. Conducting online cognitive interview

The cognitive interviews took place on the Whereby[2] videoconferencing system. Compared to other similar systems, this tool is very easy to use because it does not require to download any software nor to sign up for an account; respondents can also connect to the chatroom just by clicking a link. These features were important in order to minimize the respondents' burden and increase their participation. Whereby also includes audio and video recording functionalities, but they were not exploited to ensure confidentiality. Furthermore, the Whereby system is GDPR-compliant.
    Each interview was conducted by two researchers, with different roles:
- an interviewer, who led the conversation according to the semi-structured track;
- an observer, who managed all the technical issues.

The interviewer welcomed interviewees, gave them an idea of the research goals, explained what they would be asked to do, and then proceeded to administer the survey questions and the concurrent probes.

---

[2] https://whereby.com/

The observer shared the questionnaire on the screen, noted all the interviewees' answers and reactions, both verbal and non-verbal, and if necessary asked additional probes at the end of the interview (retrospective probing). The presence of the observer was very important because it allowed the interviewer to focus on the interaction, to easily grasp evidences of interviewees' difficulties and to explore them further with appropriate in-depth questions.

Four researchers joined the interviews and took turns in the roles of interviewer and observer. The interviews lasted about 40 minutes and researchers and interviewees were always visible to each other. A great effort was made to build a good relationship with participants: the objective was to make them feel comfortable and to gain their collaboration. This goal was met: all participants were relaxed and helpful, despite the topic. Therefore, every interview was a pleasant experience both for researchers and for participants.

6. Results

Our experience has shown that online cognitive interviewing is feasible and helpful.

As some scholars have pointed out, it can be more convenient for researchers and participants, and it could make it easier to recruit a diverse range of people. As seen before, thanks to the method used for recruiting and interviewing participants, we could reach people living throughout Italy and finish a cognitive pretest round in shorter time than usual.

Since interviewees and interviewers did not have to be in the same physical location, we managed to involve people that would have otherwise be excluded, such as one person who was on vacation and agreed to take part in the interview from the hotel lobby. The online cognitive interviewing ensures greater flexibility to respondents in choosing where and when to participate and it gives them the possibility to use their own devices (Geisen and Murphy, 2020). Even in our interviews, some of the interviewees connected via mobile phone.

In addition, the opportunity for people to participate in the study from a more natural setting improves the perception of anonymity and can facilitate discussions on sensitive topics. Our participants had no fear to provide private information about their sexual orientation and discrimination episodes in their life.

Participating from home or from another familiar place, along with the physical absence of the interviewer, seem to increase the perception of privacy. «Both the researcher and the researched are able to remain in a 'safe location' without imposing on each other's personal space. Interviewees can remain in the comfort of their home without the sense the researcher is encroaching on their personal space, while the

researcher avoids the feeling of physically imposing themselves within the participant's personal space. Thus, a neutral yet personal location is maintained for both parties throughout the process» (Hanna, 2012, p. 241).

7. Conclusion

In conclusion, we recommend some guidelines that can be helpful for researchers when setting up an online cognitive interview:
- the platform should require neither the registration of an account nor the download of a software;
- it can be useful, but not essential, a voice and video recording functionality;
- the screen sharing allows to monitor participants' responses to the questionnaire in real time and to probe them if necessary (Shepperd et al. 2021);
- as regards the setting of the interview, it is recommended a PC instead of a mobile, to better focus participants' attention on the cognitive task;
- a branded background could improve the perception of legitimacy and authority, in addition to remove possible source of distraction;
- although we have no control over the place where the participants are, we suggest choosing a quiet, comfortable and distraction-free environment;
- an interviewer alone is not enough to conduct an online cognitive interview; you need the help of someone who deal with the technical aspects;
- it is important to invest in building a good relationship before the interview begins and reassure the participants about the privacy and confidentiality of their answers;
- a good video connection is essential to monitor participants' reactions, facial expressions and other nonverbal responses which can reveal confusion or discomfort.

References

BALÌ N., BARCHERINI S., FAZZI G., GRASSI D., LORÈ B., MACCHIA S. 2019. Metodologia integrata per l'ottimizzazione del disegno dei questionari di indagine. *Working Papers N. 16/2019*, Istat.

BEATTY P. 2004. The Dynamics of Cognitive Interviewing. In Presser S. *et al*. (*eds*.), *Methods for Testing and Evaluating Survey Questionnaires*, New York: John Wiley and Sons, pp. 45-66.

BEATTY P., WILLIS G.B. 2007. Research Synthesis: the practice of cognitive interviewing, *Public Opinion Quarterly*, Vol. 71, No. 2, pp. 287-311.

COLLINS D., GRAY M. 2015. Sampling and Recruitment, in Collins D. (*eds*.), *Cognitive Interviewing Practice*, London: Sage, pp. 80-100.

CONRAD F., BLAIR J. 2001. Interpreting Verbal Reports in Cognitive Interviews: Probes Matter. In *Proceedings of the Annual Meeting of the American Statistical Association*, American Statistical Association.

CONVERSE J. M., PRESSER S. 1986. *Survey questions: Handcrafting the standardized questionnaire*. Sage Publications, Inc, Quantitative Applications in the Social Sciences.

GEISEN E., MURPHY J., 2020. A Compendium of Web and Mobile Survey Pretesting Methods. In Beatty P., Collins D., Kaye L., Padilla J.L., Willis G. and Wilmot A. (Eds.) *Advances in Questionnaire Design, Development, Evaluation and Testing*, John Wiley & Sons, Inc, pp. 287-314.

HANNA P. 2012. Using internet technologies (such as Skype) as a research medium: A research note, *Qualitative Research*, Vol. 12, No. 2, pp. 239–242.

WILLIS G. B. 2005. *Cognitive Interviewing. A Tool for Improving Questionnaire Design*. Thousand Oaks: Sage.

SHEPPERD J. A., POGGE G., HUNLETH J.M., RUIZ S., WATERS E.A. 2021. Guidelines for Conducting Virtual Cognitive Interviews During a Pandemic, *Journal of medical Internet research*, Vol. 23, No. (3).

SUMMARY

Survey data quality depends on the choices made during the questionnaire design process. Flaws in the questionnaire design can be often identified by conducting a pretest prior to the survey. Cognitive interviewing is a pretesting method useful for exploring hidden and overt problems respondents have in answering survey questions because of the sequence of questions, their wording, or the structure of responses. This interview is usually face-to-face, but the spread of the Covid-19 pandemic has imposed to find innovative solutions.
In this paper we report on one of the first experiences of online cognitive interviewing run at the Italian National Institute of Statistics, which took place in 2020 in preparation for the Survey on discrimination at work against LGBT people.

Subjects were recruited by the chain sampling method, starting from some acquaintances. They were asked to complete a screening questionnaire, sent via WhatsApp, and to invite other people to do the same. Eligible respondents were contacted by e-mail or by phone to arrange an appointment for the interview. The Whereby videoconferencing system was chosen, as it requires neither installation nor account registration, and therefore minimizes the response burden.

Based on this experience, conclusions have been drawn about the benefits of conducting online cognitive interviews. Relying on videoconferencing system allows to reach people spread throughout the territory in a short time and at low cost. In addition, it encourages the participation of those who, due to lack of time, might otherwise refuse. The freedom in choosing the time and the place of the interview, along with the physical distance from the interviewer, also foster a perception of confidentiality and can reduce the interviewees' reluctance to deal with sensitive topics. Finally, under certain setting conditions, the interaction dynamics do not seem to differ significantly from those in presence. Therefore, the spontaneity of the response process and the interviewee's ability to faithfully reconstruct the thought processes are not compromised.

_____

Sabrina BARCHERINI, Italian National Institute of Statistics, sabrina.barcherini@istat.it
Katia BONTEMPI, Italian National Institute of Statistics, katia.bontempi@istat.it
Serena LIANI, Italian National Institute of Statistics, serena.liani@istat.it
Barbara LORÈ, Italian National Institute of Statistics, lore@istat.it
Simona ROSATI, Italian National Institute of Statistics, srosati@istat.it