Article

Symposium 2008:
Data Collection: Challenges, Achievements and New Directions

Reducing the Number of Cognitive Interviews by Adding Other Cognitive Methods of Testing

by Patricia Houle

2009

Statistics Canada   Statistique Canada

Canada

Reducing the Number of Cognitive Interviews by Adding Other Cognitive Methods of Testing

Patricia Houle[1]

Abstract

The Questionnaire Design Resource Centre (QDRC) is the focal point of expertise at Statistics Canada for questionnaire design and evaluation. As it stands now, cognitive interviewing to test questionnaires is most often done near the end of the questionnaire development process. By participating earlier in the questionnaire development process, the QDRC could test new survey topics using more adaptive cognitive methods for each step of the questionnaire development process. This would necessitate fewer participants for each phase of testing, thus reducing the cost and the recruitment challenge.

Based on a review of the literature and Statistics Canada's existing questionnaire evaluation projects, this paper will describe how the QDRC could help clients in making appropriate improvements to their questionnaire in a timely manner.

Key Words: Questionnaire design, Cognitive methods, Testing and evaluation.

1. Background

The Questionnaire Design Resource Centre (QDRC) was created in 1986 and is the focal point of expertise at Statistics Canada for questionnaire design and evaluation. The Centre provides a variety of services including planning and conducting questionnaire testing using qualitative methods such as focus groups and cognitive interviews.

Qualitative testing has been used to evaluate survey content at Statistics Canada for almost 20 years. However, before 1992, all of the testing was conducted on behalf of the government agency by external private firms who specialized in market research. Starting in 1992, the QDRC started to conduct questionnaire testing. The QDRC researchers continued testing survey questionnaires as had been done earlier by the private firms. This testing primarily involved focus groups, as this technique had been used with success by the private firms and was recognized as a good approach according to the literature (Floyd J. Fowler, 1995). At that time, qualitative testing was carried out only after the survey questionnaires had been completely developed.

Shortly after the QDRC began conducting the cognitive testing, Statistics Canada created a Policy on the Review and Testing of Questionnaires in 1994. With the arrival of an increasing number of computer-assisted surveys, the content of surveys also expanded. As a result, questionnaires took more time to develop. In addition to more elaborate content, many survey planners also had to deal with issues such as fixed collection periods as well as allowing time for programming questionnaire specifications <u>before</u> the questionnaire was even developed.

2. QDRC current mandate

Currently, for a variety of reasons arising during survey development, questionnaire review and testing done by the QDRC is often delayed until the end of questionnaire development. At this late stage of development, there are times when certain findings or recommendations from the testing cannot be incorporated into the questionnaire (or be subjected to further testing) before the instrument goes to production. Rather than making changes directly to the

---

[1] Patricia Houle, Questionnaire Design Resource Centre, Statistics Canada, 100 Tunney's Pasture Driveway, Ontario, Canada, K1A 0T6, patricia.houle@statcan.gc.ca

questionnaire, notes are provided for the interviewers in the application and training manual to guide the respondents.

While the policy states that all new and redesigned questionnaires at Statistics Canada are to be reviewed and tested by the QDRC, the policy does not dictate that survey managers must make changes to questionnaires based on the recommendations or findings from the testing. The QDRC plays a consultation role for its clients; the subject matter areas. Changes cannot always be applied because of a desire to maintain comparability or because it could jeopardize the performance of the data collection application. For all these reasons, QDRC staff needs to thoroughly document their findings so they can be considered or retrieved for a future cycle or for another survey that would want to use the same questions.

The QDRC conducts qualitative testing for a wide range of social, business and agriculture surveys. When testing social and agriculture surveys, frequently, survey managers favour the use of focus groups. This is due primarily to the large number of participants this technique can reach (Gower et al., 1999). Managers feel that this is a way to obtain as much feedback as possible. Typically, a series of cognitive interviews (perhaps averaging around 50 per project) are also done during the qualitative testing. The QDRC encourages members of the survey development team to observe each of the focus groups and interview sessions to hear participant feedback directly, to get a grasp of the problems encountered and to understand why some changes are recommended.

Finally, at Statistics Canada, questionnaires must be reviewed and tested in both official languages: English and French. Typically, our clients start developing the English questionnaire. Frequently, it's only once the English questionnaire is deemed final, that it is translated into French.

In these situations, the sheer volume of sessions (focus groups and interviews), the timing of the testing in the developmental process, and the constraints of getting a questionnaire prepared for Computer Assisted Interview (CAI) programming often do not permit direct changes or improvements to the final questionnaire.

3. Suggested approach

For all of the previously mentioned reasons, the QDRC is proposing a new multi-step approach that includes different cognitive techniques, some of which are less commonly used or not used at all by Statistics Canada. This approach to testing questionnaires could improve the performance of all parties and further reduce non-sampling errors. The following paragraphs will explain how adopting this approach would help obtain better results for the review and testing of questionnaires.

3.1 What is it?

This approach aims towards the use of the appropriate cognitive technique for each step of survey development. First, start with focus group to gauge the population's interest. Second, involve QDRC in the questionnaire development. Third, review questions as they are finalized by using the help of linguists and experts. Fourth, qualitatively test questionnaires in phases and add behaviour coding and interviewer debriefing techniques.

3.2 Why suggesting a new approach?

The idea behind this new approach is to provide survey managers with a better picture of all the possible techniques involved in developing and testing a questionnaire. QDRC is a small unit which serves all divisions within Statistics Canada (SC), and also some external clients. The growing demand for its services is making it harder to manage, sometimes due to the timing of projects or the period of involvement requested. By changing some of their usual work habits, clients could benefit from QDRC's special expertise and get a better overview of what is being done in other areas. They could also look at the possibility of linking already-collected data with their survey data to reduce designing and testing time, as well as reducing response burden for the respondents. The QDRC has divided its proposed approach into five sections to better demonstrate how adopting this approach would help managers obtain better results.

4. Use of focus groups

There are subjects on what the Canadian population doesn't want to be questioned on or is reluctant to talk about. Over the years, the QDRC has observed that it is getting harder to obtain a high response rate or to recruit certain types of respondents in some regions (Allard, 2007). One of the reasons behind this lower participation rate is a fear of personal information being made public or the fact that people do not always trust the government (Allard et al., 2004). Meeting these participants will therefore help the team to better prepare their respondent relations material, to know which kind of information to target and what kind of response rate to expect for potentially sensitive topics. It will also help the team, based on the information, to identify possible analytical problems such as non-response, which could result in the non-publication of some estimates.

Using focus groups at the beginning of the development of a project makes it possible to take the pulse of the population regarding specific subjects (Krueger, 1994). These groups are designed to collect general information, which can guide the direction of certain themes or questions (Snijkers, 2002). They help to measure the ease or difficulty of selling a survey and the possible participation rate. Despite the fact that focus group participants are not statistically representative, they nevertheless provide a good idea of the trends. Furthermore, the participants also give their preferences on the type of information they want to receive, which can help them to decide whether or not to take part in a survey. What may seem important for a researcher to mention may not be what a respondent is looking for. People want to know how the data will be useful to them and whether they can benefit from the results. They often ask "What good will it do for me to participate?"

This initial meeting with the population in a focus group is the brainstorming stage that gives a better idea of people's openness to any given subject. Rather than travelling across the country, the creation of French and English groups chosen from the general population would be optimal to find out opinions on the procedures, such as to whom the survey should be addressed, how to proceed, what are the population's concerns or subjects of interest, how participants could be persuaded to take part in the survey, etc. This step helps define parameters with partners, such as the minimum age to sample children, for example. By knowing that parents would not give any information on children under the age of 3, the survey managers could raise this point with partners who want to have children participate from the age of 1. The first groups could serve as a barometer to find out where to begin and how far to go.

5. Involving QDRC in the questionnaire development

The ongoing turnover of Statistics Canada's staff due to promotions, retirement, leave or special assignments, leaves some working teams with little expertise or knowledge of the overall portrait of the survey development process. Not always knowing where or why some questions were used and using them in another context can create problematic situations or provide a false picture of a situation. For example, some questions were added to the 2001 Canadian Census questionnaire to create a sampling frame for a post-census survey on Participation and Activity Limitations. The only intention behind these questions was to establish a sample frame. Unfortunately, some managers added these questions to their own surveys to calculate the prevalence of disability among the population, which is not the purpose of these questions. As one would expect, the Participation and Activity Limitations Survey results on disability were much lower than any other survey results obtained for disability solely using the Census filter questions (Rietschlin, 2004). Consulting a QDRC resource person during the analytical proposal stage would have provided important insights about these questions and even if they were still used, it would have been evident not to use them to calculate a prevalence rate.

Having a QDRC consultant as part of the content development team would provide a larger scope of knowledge based on the fact that QDRC is aware of most of the new or revised questionnaires that are being developed at Statistics Canada. It can also provide expertise on existing questions and raise awareness of potential problems that could arise by using new modules or questions, whether validated or not, from other groups. Additionally, it can give examples of problems encountered during previous testing projects. The QDRC receives most of Statistics Canada's questionnaires for review. For this very reason, the resource person could also suggest other ways of

getting the desired information by finding validated questions from other Statistics Canada surveys. In addition, this person could be a Francophone resource for unilingual English teams, or vice versa.

By attending the team meetings, the QDRC resource person could stay up-to-date on the analytical needs and get a better understanding of the necessity of asking certain questions. He or she would know which data are sought and what they will be used for in the analysis phase, and could propose new versions of a question to avoid the usage of problematic questions. For more information, see Groves (1989) and Groves (1991).

6. Subject specialists and linguists

6.1 Subject specialists

Meetings with specialists or experts in the field should be planned to ensure that all of the field-related concepts are covered, and thereby ensure that the research objectives are attained. These experts could also review the questions individually or in a group (panel) to provide an indication of their value for obtaining data that will address the analytical needs of the survey team (Presser and Blair, 1994; Snijkers, 2002).

Taking part in the meeting sessions with specialists or experts could help identify issues that usually arise later in the questionnaire development process and of which content development team members are not always aware, or knowledgeable.

6.2 Linguists

Another technique for checking a questionnaire is to have the terminology reviewed by a specialist. Vocabulary and terminology change over time and it is important to consider this when developing a questionnaire. New words are added to our vocabulary and some words or expressions are replaced by others. Linguists are there to help us use the best words or the most commonly used ones. They can provide us with advice on when, and in which circumstance to use a specific word. Adding this additional step to the questionnaire review process could help identify problematic terminology and prepare cognitive interview follow-up questions to address it. In the last six months, to ensure that the service QDRC would be paying for was worthwhile, it has been sending entire modules or some questions which had been tested and for which problems were identified, to some linguists at Statistics Canada. The Centre wanted to see if they could identify potential problems related to question wording. The linguists questioned words that were identified by respondents as problematic and also served as informal testers. For example, "Is there public transportation in your local area", "your local area" was confusing for respondents who were not sure how close to home it meant when they were referring to "my area". A comment received from the linguists was 'local' is implied, when "your' is present. Having this information on hand would have helped to confirm whether "local" was necessary.

When questions come from other countries, whether English or French, a word's meaning can differ from one country to another. Therefore, it is important to ensure that the chosen terminology and vocabulary are adapted to the Canadian context. A linguist's expertise would be helpful in identifying terms whose proposed meaning is not the same in Canada. The linguist can raise possible comprehension problems and indicate the elements to look at during qualitative testing. A term's meaning must either be the same or as similar as possible so that the data collected has an equal value for all respondents.

For business surveys, it is important to make sure that the terminology is as close as possible to the respondents' own terminology and as accurate as possible. It makes sense to use accounting terms for accountants and technological or scientific terms for Research & Development managers. The vocabulary for business surveys doesn't have to be understood by the population in general. Therefore, unless a specialist is involved, the help of someone who knows the specific terminology is required to ensure the terms used are appropriate and current.

7.   Timeliness and testing

If the questionnaire is not received at the last minute, the QDRC will have time to review it thoroughly, send it to linguists to help identify potential problematic terms and suggest different testing strategies. During the review phase, the reference periods have to be looked at to make sure that recall will not be problematic or cause non-response errors. One must make sure possibly sensitive topics or questions are well introduced and placed towards the end of the questionnaire. Also ensure that each section flows smoothly from one to the other and look at the translation to ensure that the same concepts are measured in French and English.

7.1 Questionnaire review and pre-testing

A first round of informal testing can be done simultaneously with the questionnaire review, or right after the review, to identify major comprehension problems. This is also the time to seek an external opinion on the questions from people who are not involved in the subject matter. Is the questionnaire easy to follow and to answer? Is it easy to recall the information? Etc. Experienced survey interviewers are also a source of expertise that survey areas should be consulting more often. They are the ones in contact with respondents and who know the type of comprehension or other problems that participants encounter. They can be consulted for their point of view from their experience in reading questions and the types of comments they might have received from respondents pertaining to certain words or open-ended types of questions. All identified problems should be taken into consideration for formal testing or be rectified, depending on the issue.

7.2 First round of testing

Following the review process, which should be as exhaustive as possible, a first round of testing with some participants can be conducted. This first testing with a targeted population is mainly aimed at testing the clarity of concepts and the appropriateness of terms. It also ensures that the question flow is appropriate and that the response categories are complete. According to the studies by Hess (1999) and Willis (2005), two to fifteen interviews are sufficient to find the major problems with questionnaires. The United States Office of Management and Budget also limits each round of testing for federal research to a maximum of nine participants (Willis, 2005). The recommended tests are in-depth interviews. This technique is useful in getting more specific information because the participant is the only one responding to the questions and potential problems are easily observed, especially if a term or question has a double meaning or is double-barrelled. Furthermore, cognitive interviewers can also ask which term or word a participant would use to designate a certain object, or suggest terms and see which one is chosen most often.

The first round of testing should address the understanding of general concepts. Since this process does not require the use of a long questionnaire, it can be done as the questions are developed. For example, this phase could be used to determine the ease or difficulty of remembering events that occurred during the reference periods used in the survey.

The importance of testing in both official languages at this stage is to make sure that the French and the English questions are comparable and make sure the right words to express a concept are used and not just the best literal translation.

Once the first phase of testing is completed, it is important to review and analyse the comments to help propose the necessary changes that will also respect the analytical plans. If some changes are not made, it will be possible to qualitatively identify the potential impact that this will have on analysis. Before further testing, the changes that were accepted by the client should be made and then preparations can be made to test the changes.

7.3 Second round of testing

Once the questionnaire has been improved from the first round of testing, it is time to start the second round of testing with targeted populations in two to three different regions. This round is to confirm concept comprehension and to make sure the vocabulary and terminology used is evenly understood across the country. It is also to confirm that the response categories are mutually exclusive and exhaustive.

When the survey is targeting a specific population and uses filter questions, it is important to make sure those questions are not too inclusive. For this reason, a test with the general population is advised. This will confirm the appropriateness of the filters or provide insights about how to modify them to exclude non-targeted respondents. This can be done locally beforehand, at the same time or after testing in different provinces, depending on time constraints.

8. Behaviour coding and interviewer debriefing

After the rounds of qualitative testing, a questionnaire pre-test on a representative sample of the target population is recommended. Pre-tests are commonly used at Statistics Canada to ensure that everything is working well, including computer applications (CAI), the processing system, the time it takes to administer the questionnaire, etc. An additional step in the qualitative testing could be to add behaviour coding, which consists of drawing up a list of items to observe about both the interviewer and the respondent, such as "Was the question read as it was written?", "Did the interviewer give any explanations other than those written?", "Did the respondent cut off the interviewer while he/she was talking?", "Did the respondent respond before hearing the end of the question?", "Did the respondent ask the interviewer to repeat the question?", etc. (Willis, 2005). All of these behaviours may indicate the degree of a defective or problematic question. This cognitive technique is valuable if there is enough time to make changes to the questionnaire before the start of data collection. It is also worthwhile because it permits quantitative analyses, given that a representative sample is used. However, this process is very demanding, because all of the interview must be analyzed thoroughly using a very strict and consistent protocol. Observers must be trained in advance so that the information collected is comparable. The last step of qualitative testing proposed is to do an interviewer debriefing to identify which kind of problems interviewers have encountered and resolve them or address them during training.

9. Costs

Table 9-1 provides a visual representation of the cost of the current way of testing versus the cost of the proposed new way. Right now, the number of days allowed for a typical project is around 40 days which are mostly spent on travelling and doing interviews by one, two or sometimes three different QDRC consultants. If major problems arise with questions or answer categories while testing, it is harder to make the appropriate changes and coordinate with the survey manager who might not be at the testing location or with another consultant on the project, if testing is done simultaneously in different languages or locations.

Due to budget constraints many QDRC clients are waiting until fiscal year-end to test their questionnaires to make sure they will have enough money to cover the testing costs. Testing strategies are often based on budget and schedules and not on questionnaire testing requirements. This is also true for clients who budgeted a large amount on testing and want an enormous number of focus groups or one-on-one interviews, which unfortunately bring no new feedback or information to improve questions; on the contrary, it frustrates the consultants who hear the same comments over and over.

Table 9-1
Cost of testing (an example)

	Salary	Non-Salary	Total
Current way	(40 days) $24,480	(3 locations) $24,145	$48,625
New way	(55 days) $33,660	(4 locations) $16,440	$50,100
Difference from current to new	+ 15 days	$-7,705	$1,475

The proposed new way of testing will allow the involvement of a consultant from start to finish on the project. The rationale is that better results can be achieved with less traveling time and fewer interviews if the QDRC is involved

earlier and different techniques are implemented in the testing of a questionnaire. In addition, more work can be done locally, thus reducing non-salary expenditures and leaving the total project cost at the same budget level as it is now.

Reducing the number of interviews and the number of person-days has a huge impact on non-salary costs. In this estimate, four sites were chosen for testing on different dates, allowing time for changes. It also allows the spreading of costs over a longer period, even over two different fiscal years.

10. Benefits of the new approach to managers

There are a wide range of cognitive testing approaches and each technique provides different results. The important issue is to highlight the needs of each project in order to propose the cognitive tests that will be most useful to each survey. The main point remains to include a QDRC member from the very beginning of the development process, so that he/she can suggest different tests to conduct. These tests should be done throughout the development process so that the questionnaire is ready earlier and has received input from all available experts.

As explained earlier, it is not necessary to have many participants in a test. The important thing is to do different tests to obtain a final product that contains the fewest errors possible. Testing both languages early could eliminate the discrepancy between Francophone and Anglophone responses.

There are real advantages for survey managers in adopting this new proposed testing approach. They will receive more consultation time and expertise from QDRC consultants. They will be aware of the population's information needs. Doing thorough questionnaire reviews with different experts and several rounds of cognitive testing should eliminate most of the comprehension problems and reduce non-sampling errors. If other cognitive techniques are used, it will be possible to reduce the number of cognitive interviews and therefore, reduce the number of days for traveling and other expenses. Not waiting until the last minute for testing will allow time to spread tasks out and could potentially reduce the high level of stress to meet deadlines. And as shown in Table 9-1, the dollar amount is very similar to the current testing method, which should therefore encourage managers to use this approach.

11. Conclusion

The Policy on the Review and Testing of Questionnaires already outlines some of the proposed ways of testing, but it is not followed by survey managers mainly due to a lack of time, knowledge and proof that it works. The policy describes the best practices to achieve better results, but does not prescribe involving QDRC consultants during the questionnaire development process. Reviewing and testing questionnaires is an obligation for managers and not always seen as a helpful tool in achieving the best results, therefore recommendations are not always followed and the review and testing of questionnaires is done at the last minute.

Even though the literature shows that adding cognitive methods to test questionnaires helps achieve better results, as shown in Willis et al. (2005) and Campanelli (1997), managers would like proof of the validity that such an approach works in the context of Statistics Canada before using it.

The QDRC needs to gradually change the testing mentality of survey managers by suggesting and adding cognitive methods to proposals. The Centre also needs to validate these techniques and clearly demonstrate to managers that their survey could benefit from such methods. Being part of a new project from the beginning and applying the proposed approach could validate the proposed alternative methods, while keeping the actual testing method in place and then comparing the results. The QDRC has a role in educating survey managers, but it also needs to find support from management to validate and encourage this new approach to testing questionnaires and survey projects at Statistics Canada.

References

Allard, B., Gower, A., Kelly P., and Walker C. (2004). Experiences in Testing Questionnaires with Specialized Populations, *Proceedings: Symposium 2004, Innovative Methods for Surveying Difficult-to-Reach Populations.* Statistics Canada.

Allard, B. (2007). Who Is More Likely To Attend? A Study of No-Shows in Qualitative Research, *Proceedings of the 2007 Workshop on Questionnaire Evaluation Standards (QUEST)*, 83-87.

Campenelli, P. (1997). Testing Survey Questions: New Direction in Cognitive Interviewing, *Bulletin de Méthodologie Sociologique*, 55, 5-17.

Fowler, F.J., Jr. (1995), *Improving Survey Questions: Design and Evaluation*, Applied Social Research Methods Series, Vol. 38, Sage Publications, Thousand Oaks, California.

Gower, A.R. and Haarsma G. (1997), A Comparison of Two Methods in a Test of the Canadian Census Questionnaire: Think-aloud Interviews vs. Focus Groups, *Proceedings of the Minimum Standards in Questionnaire Testing Workshop*, Statistics Sweden, Örebro, Sweden, 35-37.

Gower, A.R., McClure K., Paletta A., and Williams M.J., (1999), When to Use Focus Groups versus Cognitive Interviews in the Development and Testing of Questionnaires: The Statistics Canada Experience, *Proceedings of the 1999 Workshop on Questionnaire Evaluation Standards (QUEST)*, 51-66.

Groves, R.M. (1989). *Survey Errors and Survey Costs*. New York: John Wiley & Sons.

Groves, R.M., (1991), Measurement Errors Across the Disciplines, *Measurement Errors in Surveys* (Eds. P. Biemer, R. M. Groves, L.E. Lyberg, M.A. Mathiowetz, and S. Sudman). New York: John Wiley & Sons, 1-25

Miller, K. (2002). A Comparison of Focus Groups and One-on-One Cognitive Interviewing for Question Evaluation. *International Conference on Questionnaire Development, Evaluation, and Testing (QDET)*

Groves, R.M., Fowler, F.J. Jr., Couper, M.P., Lepkowski, J.M., Singer E. and Tourangeau, R. (2004). *Survey Methodology*, John Wiley & Sons, Inc., Hoboken, New Jersey.

Hess, J. (1999). Report on Cognitive Interviewing Results for the Survey of Program Dynamics Children's Residential History Module. Washington, DC: U.S. Census Bureau, Centre for Survey Research Methods.

Krueger, R. A. (1994). *Focus Groups: A Practical Guide for Applied Research* (2nd edition) Thousand Oaks, CA: Sage.

Presser, S. and Blair, J. (1994). Survey Pretesting: Do Different Methods Produce Different Results? *Sociological Methodology*, 24, Washington D.C., American Sociological Association, 73-104

Presser, S., Couper, M.P., Lessler, J.T., Martin, E., Martin, J., Rothgeb J.M. and Singer E. (2004), Methods for Testing and Evaluating Survey Questions, *Public Opinion Quarterly*, 68, 109-130.

Rietschlin, J. and MacKenzie, A. (2004) Variation in Disability Rates in Statistics Canada National Surveys: Building Policy on a Slippery Foundation. *Proceedings: Symposium 2004, Innovative methods for surveying difficult-to-reach populations*. Statistics Canada.

Rothgeb, J., G. Willis and B. Forsyth (2001). Questionnaire Pretesting Methods: Do Different Techniques and Different Organizations Produce Similar Results? *Proceedings of American Association for Public Opinion Research.* American Statistical Association.

Willis, G.B. (2005). *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Sage Publications, Inc., Thousand Oaks, California.

Willis, G.B. (2001). The Continuing Saga of Pretest Methods Comparions: Cognitive Interviewing, Expert Review, and Behaviour Coding. *Proceedings of the 2001 Workshop on Questionnaire Evaluation Standards (QUEST)*, 67-72