SURVEYS · QUESTIONNAIRE DESIGN

Respondent Debriefing Techniques for Improving Future Survey Knowledge Collection

Respondent debriefing asks people, after they have answered a survey, how they understood and answered its questions. This article explains the main debriefing techniques, how they differ from other pretesting methods, how to analyse the results and how to use them to improve later surveys.

ModQu research article · 2,368 words · 24 references (APA 7th edition)

A survey question can be answered by thousands of people and still produce misleading data, because respondents understood it differently from the way its author intended. Most of the time, the researcher never finds out. **Respondent debriefing** is a set of techniques for finding out. After respondents have answered some or all of the survey questions, they are asked additional questions about how they understood the questions, how they arrived at their answers and how they experienced the survey (Martin, 2004). The answers tell researchers what their questions actually measured and how to improve them for the next round of data collection.

This article explains the techniques used, how debriefing compares with other ways of testing questionnaires, how to design and analyse debriefing questions and how organisations can build debriefing into a cycle of continuous improvement.

What respondent debriefing is

Respondent debriefing is usually carried out in a field setting, with respondents drawn from the same population as the main survey. This distinguishes it from laboratory methods such as cognitive interviewing, in which a small number of volunteers are interviewed in depth (Willis, 2005). Debriefing questions are typically structured, so they can be asked of many respondents and analysed quantitatively. They are often asked at the end of the questionnaire, so they do not affect the answers to the main questions, though some designs place them immediately after the question they concern (Converse & Presser, 1986).

The method has a long history. Belson (1981) conducted intensive follow-up interviews with respondents who had just answered common survey questions and found that many questions were misunderstood by a large share of people. Schuman (1966) proposed the "random probe", in which interviewers ask a randomly selected subset of respondents to explain one of their answers in their own words, providing a check on whether closed questions are measuring what they are meant to. Survey agencies, notably the United States Census Bureau, later made debriefing a standard part of questionnaire development (Campanelli et al., 1991; Esposito, 2004).

Why debriefing is needed

The process of answering a survey question has several stages: understanding the question, retrieving relevant information from memory, forming a judgement and reporting it in the format offered (Tourangeau et al., 2000). Problems can occur at any stage, and most are invisible in the final data. A respondent who interprets "household" more narrowly than intended still gives an answer, and nothing in that answer reveals the misunderstanding.

Even familiar words cause problems. In a series of studies, Fowler (1992) showed that terms such as "exercise" or "servings of fat" were understood in very different ways, and that clarifying them changed the answers. Debriefing exposes such differences at a scale large enough to estimate how common they are.

The main techniques

Comprehension probes

The most common debriefing questions ask what a term or question meant to the respondent. These can be open ("What did you think of when you read the word 'income'?") or closed, listing possible interpretations and asking which the respondent had in mind. Closed versions are easier to analyse and allow the researcher to estimate the share of respondents holding each interpretation (Martin, 2004).

Vignettes

In a vignette debriefing, respondents are given short descriptions of hypothetical situations and asked how they would answer the survey question for each. For example, a question about whether someone "worked last week" can be tested with vignettes describing unpaid work in a family business or a few hours of casual work. The answers show where respondents draw the boundaries of a concept and whether these match the intended definition (Martin et al., 1991; Martin, 2004).

Category-selection and confidence probes

Other probes ask why a respondent chose a particular answer category, or how confident they are in a figure they reported. For factual questions, asking whether an answer was an exact figure or an estimate, and how it was arrived at, helps to judge the accuracy of the data (Converse & Presser, 1986).

Sensitivity and burden questions

Debriefing can also ask about the respondent's experience: whether any questions felt intrusive, whether they think others
would answer honestly and whether the survey was too long or repetitive. Answers to these questions help explain item nonresponse and identify sections that cause burden (Galesic & Bosnjak, 2009).

Open feedback

Finally, many surveys end with an open question inviting comments. These verbatim comments are often ignored, but they can reveal problems that the researcher did not anticipate, from confusing instructions to technical faults in a web survey (Singer & Couper, 2017).

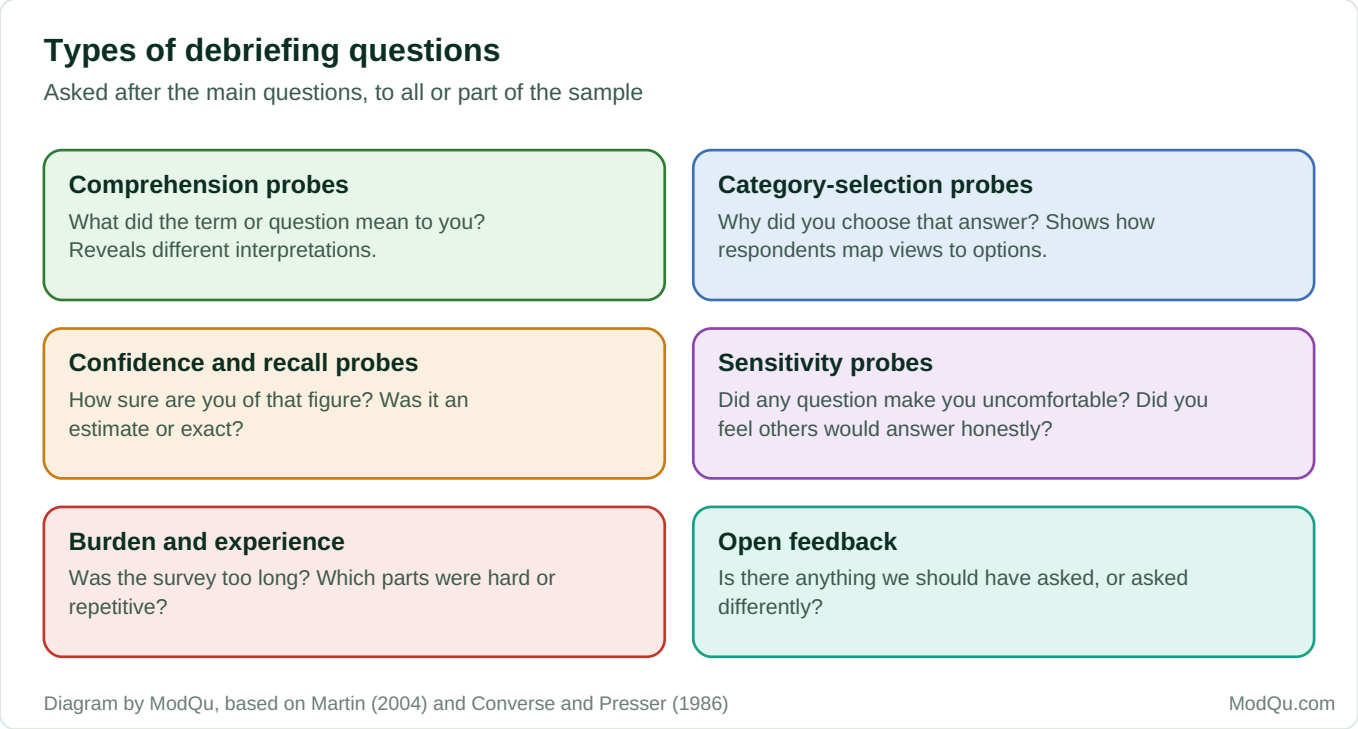


Figure 1. Types of debriefing questions. Diagram by ModQu, based on Martin (2004) and Converse and Presser (1986).

Debriefing in web surveys: online probing

Web surveys have made debriefing cheaper and easier to apply on a large scale. In **online probing**, follow-up questions are added to a web questionnaire, often directly after the target question, and respondents type their answers. Studies have found that the quality of answers to such probes depends on factors such as the wording of the probe and the effort respondents are willing to make (Behr et al., 2012). A comparison of online probing with cognitive interviewing found that the two methods identified many of the same problems, with online probing offering larger samples and cognitive interviewing offering more depth (Meitinger & Behr, 2016).

Web surveys also produce paradata, such as response times and changes of answer, which can point to questions that cause difficulty (Kreuter, 2013). Paradata do not explain why a question is difficult, but combined with debriefing questions they help to locate and interpret problems.

How debriefing compares with other methods

Questionnaire testing uses several methods, and each finds different problems (Presser et al., 2004). Expert review relies on the judgement of experienced methodologists. Cognitive interviews explore the response process in depth with small samples. Behaviour coding records how interviewers and respondents interact during interviews (Oksenberg et al., 1991). Field pretests try the full questionnaire under realistic conditions.

Respondent debriefing complements these. Its main strength is that it can estimate how widespread a problem is in the real population. A cognitive interview may show that two of twelve participants misunderstood a question; a debriefing in a field pretest of 500 respondents can show whether that problem affects 2 per cent or 30 per cent of the target population. Its main weakness is that structured debriefing questions can only find problems the researcher thought to ask about. For this reason, debriefing is often used after cognitive interviews have identified possible issues, to test how common they are (Esposito, 2004; Beatty & Willis, 2007).

Designing debriefing questions

Decide what to learn

Each debriefing question should address a specific concern about a specific survey question. Vague questions such as "Were any questions unclear?" produce little useful information, because respondents are poor at noticing their own misunderstandings. Specific probes about specific terms work better (Martin, 2004).

Place the questions carefully

Debriefing questions asked straight after a target question can change how respondents answer later questions, because they draw attention to the meaning of terms. Asking them at the end of the survey avoids this but relies on respondents remembering what they thought earlier. A common compromise is to put debriefing at the end and to repeat the target question's wording in the probe (Converse & Presser, 1986).

Ask a subsample

Debriefing adds to respondent burden. In large surveys it is often enough to ask debriefing questions of a random subsample, or to rotate different debriefing modules across subsamples, so that each respondent receives only a few extra questions. Schuman's random probe used this logic (Schuman, 1966).

Debrief interviewers too

In interviewer-administered surveys, interviewers see problems that never appear in the data: hesitations, requests for clarification and visible confusion. Structured interviewer debriefings, through questionnaires or group discussions, add this evidence and were used alongside respondent debriefings in early Census Bureau studies (Campanelli et al., 1991; DeMaio & Rothgeb, 1996).

When to debrief

Debriefing can be used at three points in the life of a survey. In a **field pretest**, it is part of the development process: the questionnaire is tried on a few hundred respondents, and the debriefing results guide revisions before the main fieldwork. In the **main survey**, a debriefing module asked of a subsample provides evidence about the quality of the data that are actually being collected, which can be used to interpret results and to adjust analysis. **After a survey**, debriefing studies can follow up a sample of respondents, sometimes with a second interview, to investigate specific problems in depth (Campanelli et al., 1991).

Each point serves a different purpose. Pretest debriefing changes the questionnaire before it is used. Main-survey debriefing documents how well it worked. Post-survey debriefing investigates puzzles in the results, such as figures that do not match other sources.

Debriefing in the research sense and the ethical sense

The word "debriefing" has a second meaning in research. In experiments that involve deception, participants are told afterwards what the study was really about. This **ethical debriefing** is a requirement of research ethics and serves to protect participants. Respondent debriefing in survey methodology is different: its purpose is to learn about the questions, not to inform participants about a hidden aim. The two can occur in the same study, for example in a survey experiment where respondents saw a version of a question chosen at random, but they should be designed separately, and survey reports should make clear which kind is meant.

Analysing debriefing data

Closed debriefing questions are analysed like any other survey questions. The main output is usually the distribution of interpretations, as in Figure 2, together with a comparison of answers to the target question across interpretations. If respondents with a broader interpretation report systematically different answers, the question is producing measurement error.

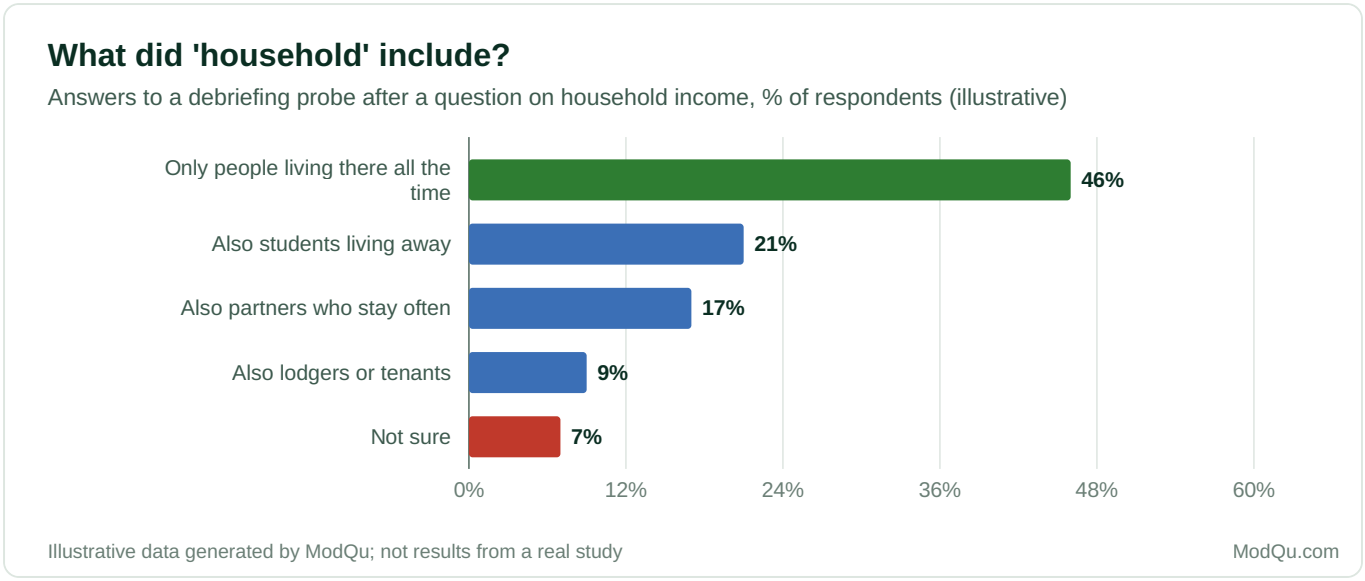


Figure 2. Example of a debriefing result showing different interpretations of "household" (illustrative data). Illustrative data generated by ModQu.

In the example in Figure 2, fewer than half the respondents used the intended definition. A researcher seeing this result would add a definition to the question, or ask separately about people who live elsewhere part of the time (illustrative figures).

Open answers require coding. A coding frame is developed from a sample of answers and applied by trained coders, with a check of agreement between coders. Open debriefing answers in large surveys can be numerous, and systematic approaches to analysing them have been described for cognitive interviews and can be adapted (Willis, 2015; Miller et al., 2014).

Using debriefing for continuous improvement

Debriefing is most valuable when it is part of a cycle. A question is tested, revised, fielded and debriefed again. Esposito (2004) described such an iterative, multiple-method process for evaluating a labour force questionnaire, in which debriefing results led to revisions that were themselves tested. For organisations that run surveys repeatedly, such as annual employee surveys or customer tracking studies, a small debriefing module in each wave can steadily improve the questionnaire. Changes to tracked questions must be handled carefully, though, because changing a question breaks comparability with earlier waves. Debriefing can help decide whether a change is worth that cost by showing how large the problem is.

An example

A firm conducts a quarterly survey of small business customers that asks, "How many employees does your business have?" Results look plausible, but they do not match figures from other sources. The research team adds a debriefing module, asked of a random fifth of respondents, with two questions: whether the respondent included part-time staff,

contractors and the owners themselves, and whether the figure was exact or an estimate. The results show that a third of respondents excluded part-time staff and a quarter excluded owners. In the next wave, the question is rewritten to include a short definition, and the debriefing is repeated to check that the new wording works (illustrative example).

In business surveys

Commercial and organisational surveys often run on tight timelines, and debriefing can seem like a luxury. In practice, the cost is small. Adding three or four debriefing questions for a random tenth of respondents rarely adds more than a minute to their survey, and it can reveal problems that would otherwise distort results for years. Customer surveys, for example, often use terms such as "service", "value" or "support" that customers interpret in different ways, and employee surveys use terms such as "my manager" or "senior leadership" that mean different people in different parts of an organisation. A short debriefing module can show whether these differences are large enough to matter. Where the same questionnaire is used across countries or languages, debriefing in each version helps to check that translated questions are understood in the same way.

Common problems

Several problems recur. Debriefing questions that are too general produce reassuring but uninformative answers. Probes that suggest an answer, such as "Did you include part-time staff, as you should have?", invite respondents to agree. Results are sometimes collected but never analysed, because no one was given responsibility for acting on them. Finally, debriefing is sometimes treated as a substitute for testing before fieldwork, when it works best as one method among several (Presser et al., 2004; Collins, 2015).

Checklist

- Identify the specific survey questions and terms that most need testing.
- Write a specific probe for each concern, avoiding leading wording.
- Place probes at the end of the questionnaire unless there is a strong reason not to.
- Use a random subsample or rotating modules to limit burden.
- Combine closed probes for measurement with an open question for unexpected problems.
- Debrief interviewers where interviewers are used.
- Analyse results promptly and decide on changes before the next wave.
- Document question changes so that trends can be interpreted (Groves et al., 2009).

Conclusion

Respondent debriefing turns the people who answer a survey into a source of evidence about the survey itself. By asking respondents how they understood and answered the questions, researchers can find out how well their questions worked, how widespread problems are and what to change. It complements cognitive interviewing and other pretesting methods

by adding scale and realism. Used routinely, it makes each round of data collection better than the last, and it helps ensure that the knowledge produced by a survey reflects what respondents meant, not only what the researcher assumed they meant (Martin, 2004; Dillman et al., 2014).

References

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly, 71*(2), 287–311. https://doi.org/10.1093/poq/nfm006

Behr, D., Kaczmirek, L., Bandilla, W., & Braun, M. (2012). Asking probing questions in web surveys: Which factors have an impact on the quality of responses? *Social Science Computer Review, 30*(4), 487–498. https://doi.org/10.1177/0894439311435305

Belson, W. A. (1981). *The design and understanding of survey questions*. Gower.

Campanelli, P. C., Martin, E. A., & Rothgeb, J. M. (1991). The use of respondent and interviewer debriefing studies as a way to study response error in survey data. *The Statistician*, *40*(3), 253–264. https://doi.org/10.2307/2348278

Collins, D. (Ed.). (2015). *Cognitive interviewing practice*. SAGE. https://doi.org/10.4135/9781473910102

Converse, J. M., & Presser, S. (1986). *Survey questions: Handcrafting the standardized questionnaire*. SAGE. https://doi.org/10.4135/9781412986045

DeMaio, T. J., & Rothgeb, J. M. (1996). Cognitive interviewing techniques: In the lab and in the field. In N. Schwarz & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research* (pp. 177–195). Jossey-Bass.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method* (4th ed.). Wiley.

Esposito, J. L. (2004). Iterative, multiple-method questionnaire evaluation research: A case study. *Journal of Official Statistics*, *20*(2), 143–183.

Fowler, F. J., Jr. (1992). How unclear terms affect survey data. *Public Opinion Quarterly, 56*(2), 218–231. https://doi.org/10.1086/269312

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly, 73*(2), 349–360. https://doi.org/10.1093/poq/nfp031

Groves, R. M., Fowler, F. J., Jr., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Wiley.

Kreuter, F. (Ed.). (2013). *Improving surveys with paradata: Analytic uses of process information*. Wiley. https://doi.org/10.1002/9781118596869

Martin, E. (2004). Vignettes and respondent debriefings for questionnaire design and evaluation. In S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin, & E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires* (pp. 149–171). Wiley.

Martin, E., Campanelli, P., & Fay, R. E. (1991). An application of Rasch analysis to questionnaire design: Using vignettes to study the meaning of 'work' in the Current Population Survey. *The Statistician, 40*(3), 265–276. https://doi.org/10.2307/2348279

Meitinger, K., & Behr, D. (2016). Comparing cognitive interviewing and online probing: Do they find similar results? *Field Methods, 28*(4), 363–380. https://doi.org/10.1177/1525822X15625866

Miller, K., Willson, S., Chepp, V., & Padilla, J.-L. (Eds.). (2014). *Cognitive interviewing methodology*. Wiley. https://doi.org/10.1002/9781118838860

Oksenberg, L., Cannell, C., & Kalton, G. (1991). New strategies for pretesting survey questions. *Journal of Official Statistics, 7*(3), 349–365.

Presser, S., Couper, M. P., Lessler, J. T., Martin, E., Martin, J., Rothgeb, J. M., & Singer, E. (2004). Methods for testing and evaluating survey questions. *Public Opinion Quarterly, 68*(1), 109–130. https://doi.org/10.1093/poq/nfh008

Schuman, H. (1966). The random probe: A technique for evaluating the validity of closed questions. *American Sociological Review, 31*(2), 218–222. https://doi.org/10.2307/2090907

Singer, E., & Couper, M. P. (2017). Some methodological uses of responses to open questions and other verbatim comments in quantitative surveys. *Methods, Data, Analyses, 11*(2), 115–134. https://doi.org/10.12758/mda.2017.01

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press. https://doi.org/10.1017/CBO9780511819322

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. SAGE. https://doi.org/10.4135/9781412983655

Willis, G. B. (2015). *Analysis of the cognitive interview in questionnaire design*. Oxford University Press.