gesis

Leibniz Institute
for the Social Sciences

Web probing – **implementing probing techniques from cognitive interviewing in web surveys with the goal to assess the validity of survey questions**

GESIS Survey Guidelines

Dorothée Behr, Katharina Meitinger, Michael Braun, Lars Kaczmirek

These slides are based on the GESIS Survey Guideline paper about web probing:

Behr, D., Meitinger, K., Braun, M., & Kaczmirek, L. (2017). Web probing – implementing probing techniques from cognitive interviewing in web surveys with the goal to assess the validity of survey questions. Mannheim, GESIS – Leibniz-Institute for the Social Sciences (GESIS – Survey Guidelines). DOI: 10.15465/gesis-sg_en_023

Please cite the slides as:

Behr, D., Meitinger, K., Braun, M., & Kaczmirek, L. (2020). Slide Set: Web probing – implementing probing techniques from cognitive interviewing in web surveys with the goal to assess the validity of survey questions. GESIS Survey Guidelines. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences.

A complete list of all references used on these slides can be found in the above mentioned Survey Guideline paper.

Overview

- Introduction

- What is web probing?

- Implementation of web probing in surveys

- Stages of implementation and analysis potential

- Use cases: Errors and themes

- Analysis of probing data

- Cross-cultural web probing

Introduction

Introduction

- There is a necessity to pretest and evaluate survey questions before they go into the field.

- Traditional methods include expert review, pilot testing or cognitive interviewing.

- A prominent method in cognitive interviewing is probing.

  ▸ Interviewers ask follow-up questions after a closed-ended question (embedded) or at the end (retrospective).

  ▸ The aim is to learn about respondents' cognitive processes, their ways of understanding certain terms and of constructing their answers.

What is web probing?

What is web probing?

- It is the implementation of probing techniques from cognitive interviewing in web surveys with the goal to assess the validity of survey questions.

- It uses open-ended questions as follow-ups to closed-ended questions. The follow-ups are called *probes*.

- The next slide provides an example of a probe (screen).

What is web probing?

What ideas do you associate with the phrase "civil disobedience"? Please give examples.
The previous question was: How important is it that citizens may engage in acts of civil disobedience when they oppose government actions?

What is web probing?

- The aim is to use the open-ended answers to examine whether the closed-ended questions measure what they are supposed to measure.

- *When applied to the cross-national or cross-cultural context*, the qualitative data elicited through web probing allows checking for comparability across countries.

- Different sources can be used to recruit respondents, incl. probability-based panels, online access panels, crowdsourcing platforms or a pool of respondents that a researcher himself/herself recruits.

What is web probing?

Web probing aims to tackle some of the limitations of cognitive interviewing, in particular:

- small sample sizes that are usually used.

- With small samples, there is the danger to miss or overestimate errors or answer patterns.

- *When applied to the cross-national or cross-cultural context,* the comparability of cognitive interviewing results may suffer through "house effects" or "interviewer effects" across countries or cultural groups.

What is web probing?

Advantages of web probing:

- ease of recruitment of large sample sizes

- access to geographically diverse respondents

- elimination of interviewer effects by standardized probing

- anonymous survey environment and reduced social desirability

- no interviewer recruitment and training necessary

- time needed for data collection is shorter

- no need for transcriptions

What is web probing?

Disadvantages of web probing:

- restriction to population groups that can be reached online and that are sufficiently skilled in reading and writing

- lack of motivation by an interviewer and consequently an increase in probe nonresponse

- lack of interactivity, which would allow spontaneously acting on issues coming up or rephrasing a probe that turns out to be problematic

Implementation of web probing in surveys

Current and future research

- The implementation decisions, findings, and recommendations in these slides (and in the underlying GESIS survey guidelines) are based on research conducted in two DFG-funded research projects (Braun et al., 2010-2015).

- Many research gaps remain and we see an increase in studies on the topic.

Implementation of probing in web surveys

Access to respondents:

- Probability-based panels (e.g. Open-Probability-Based Panel Alliance) - https://openpanelalliance.org/

- Online access panels (constituted of respondents who have voluntarily signed up for taking part in surveys at regular intervals)

- Crowdsourcing platforms (e.g. Mturk)

- Respondents recruited by the researchers themselves

Implementation of probing in web surveys

Framing the research:

- Consider informing respondents on the existence of open-ended questions on the introduction screen – you may or may not want to explicitly frame the survey as a pretest.

  - Differently framed introductions (e.g. as a pretest) may have an effect on the number of probes one can ask.

- The respondent source (online access panel, MTurk, etc.) may also play a role concerning the willingness to answer probes.

Implementation of probing in web surveys

Burden-related aspects:

- Probes have a higher response burden than closed-ended questions since pre-coded answer categories do not exist that could guide respondents.

- Typing the response itself may prove annoying or strenuous for some respondents.

- A motivating interviewer is missing.

- In sum, everything needs **to be designed in a manner that keeps the response burden low and makes the research intention clear.**

Probe types

Category-selection probing:

- asks respondents for their reason(s) for having chosen a specific response category (e.g. "Please explain why you selected [chosen answer value]"; e.g. 'completely agree').

- is useful for checking whether the categories make sense to respondents.

- can reveal silent misunderstandings of an item (that make sense to the respondents but are not in line with the research goals).

Probe types

Comprehension probing:

- asks respondents to define how they understand a certain term (e.g. "What do you consider to be a 'serious crime'?").

- The responses typically take the form of a definition or a list of things (themes) that respondents think of in the context of the requested term.

- is ideal for testing whether a term is understood as intended by the researcher.

Probe types

Specific probing:

- focuses on a particular detail of a term (e.g. "Which type of immigrants were you thinking of when you answered the question?").

- is useful for getting an understanding of the breath that certain terms can have.

- Responses typically contain a list of themes.

- For example, the term "immigrant" triggers associations of many different concepts such as specific countries and reasons to immigrate, etc.

Probe placement

- Implementation of the probe on a separate screen following the respective closed-ended question.

- This disentangles the response process for the closed-ended questions from the probing process and keeps the 'usual' survey experience of closed-ended questions as stable as possible.

Probe placement

1st screen

In general, how would you rate the current state of the economy in Britain?

○ very good

○ good

◉ partly good, partly poor

○ poor

○ very poor

_____

○ *can't choose*

2nd screen

Please explain why you selected "partly good, partly poor".

The question was: "In general, how would you rate the current state of the economy in Britain?"
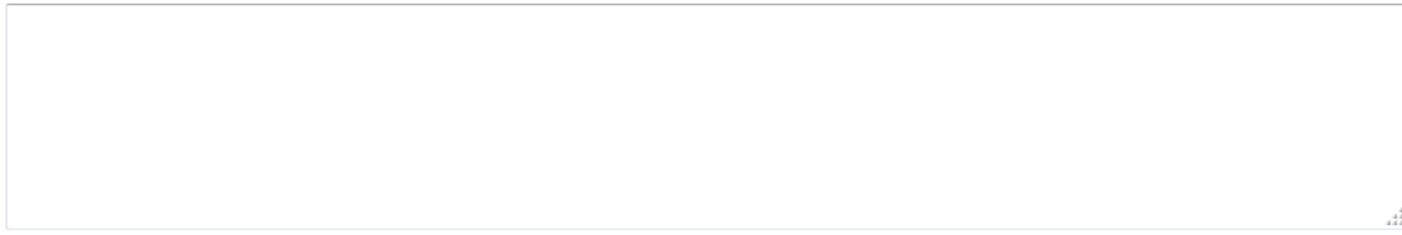
Probe presentation

- To alleviate response burden, respondents should be provided with the corresponding closed-ended item and their answers on the probe screen.

- This way, recall is aided, respondents can concentrate on the probe itself, and probe nonresponse is decreased.

Probe presentation

Probe screen for a probe including a numerical value

Please explain why you selected "3".

The question was: "And how important is it that people convicted of serious crimes lose their citizen rights?" Your answer was "3" on a scale from 1 (not at all important) to 7 (very important).

Number of probes

- 8 or 9 probes in a 15-minute survey work well, but there is certainly still much to learn how many probes can be asked.

- Consider having probes and a certain number of closed-ended items take turns to keep the burden low .

- When all items from a longer scale are relevant, consider implementing splits to have only subgroups of respondents receiving a given set of probes (which allows testing many more items).

Sequence of probes

- The distribution of probes *across the entire survey* should avoid unintended habituation effects.

- Regarding the sequence of probes *for the same item,* a design with category-selection probe coming first compared to comprehension or specific probes fares best (increases response rates for the probes, decreases mismatching answers which do not fit to the probe type asked).

Text box size

- The text box size gives cues to the respondents as to the desired answer depth, length or format.

- The text box size should fit the desired answer type or format (an explanation requires a larger text box than a 'mere' listing of activities).

- The text box for a category-selection probe should be larger than a text box for a specific probe.

- Adjusting the text box size to the probe type can also help to prevent mismatching answers.

Examples text box size

Please explain why you selected "3".

The question was: "And how important is it that people convicted of serious crimes lose their citizen rights?"
Your answer was "3" on a scale from 1 (not at all important) to 7 (very important).

What particular citizen rights did you have in mind when you were answering the question?

The question was: "And how important is it that people convicted of serious crimes lose their citizen rights?"

Nonresponse reduction & tool support

- Probes in the web context are particularly prone to nonresponse.

- Kaczmirek, Meitinger, & Behr (2017) have developed a tool to automatically detect different types of nonresponse *during* the survey and to follow up with a suitable follow-up probe and a tailored motivational statement.

- This tool is freely available in German, English, and Spanish.

Stages of implementation and analysis potential

Pretesting stage

- At the pretesting stage, cognitive interviewing is the method of choice for *in-depth* exploration, due to the possibility to follow up on probes.

- Web probing is advisable when researchers are interested in answer patterns and their prevalence, when the probe types to use are known, and a certain geographical spread is needed.

- Combinations are possible depending on the research question.

Main survey stage

- Probing can be implemented at *selected* questions in the main survey itself.

- The random probe method conceived by Schuman (1966) may be a guiding principle whereby random splits of respondents receive a probe rather than all respondents.

Post-hoc evaluation

- Web probing can be implemented post-hoc to shed light on existing survey data to (1) explain anomalies in the data or to (2) assess problematic questions in general.

- Post-hoc evaluation may be especially interesting for surveys with multiple waves or rounds, and where feedback is needed to take decisions on whether items should remain in a survey or not.

Use cases: Errors and themes

Use cases: Errors and themes

- In an error perspective, errors may be coded along the components of the response process, that is, comprehension, retrieval, judgment, and response.

- When substantive themes are relevant, the coding scheme will have to be developed (inductively, deductively or both ways) specifically for the item of interest.

Analysis of probing data

Analysis of probing data

- Inductive development of coding schema starts from a large sample of probe answers.

- Deductive development of coding schema starts from theory and translates hypotheses into codes.

- Coding schemata include definitions of codes, coding rules, and example answers for the different codes.

- The schemata should be consistently understood by coders not involved in schema development.

- A finalization of a coding schema should occur only after a few trial runs.

Analysis of probing data

- Coders need to be trained on the final coding schema (explanation and several rounds of exercises and feedback).

- One coder can code the entire data set, a second coder only a random sample of probe answers to produce a basis for the establishment of intercoder reliability.

Cross-cultural web probing

Cross-cultural web probing

- If a coding scheme is developed inductively, responses from all languages should be taken into account.

- Research teams should ideally consist of members from all countries so that coding and analysis can be done by native speakers.

- In the case of external translation: Translators should not change the message of the responses (e.g. by making unambiguous what is truly ambiguous in the original response, by rendering clear what is not intelligible, etc.)

Cross-cultural web probing

- Regarding translation: Besides a close meaning transfer (which should still be grammatically and semantically correct, of course) commenting on cultural allusions, cultural facts, persons, etc. is helpful to understand responses against their cultural backdrop.

- There should be an open communication channel to the translators so that linguistic and cultural queries can be raised.

Selected publications

Hands-on guidelines:

- Behr, Dorothée, Katharina Meitinger, Michael Braun, and Lars Kaczmirek. 2017. *Web probing – implementing probing techniques from cognitive interviewing in web surveys with the goal to assess the validity of survey questions*. GESIS – Survey Guidelines. GESIS – Leibniz-Institute for the Social Sciences . doi: http://dx.doi.org/10.15465/gesis-sg_en_023.

- Kaczmirek, Lars, Katharina Meitinger, and Dorothée Behr. 2017. *Higher data quality in web probing with EvalAnswer: a tool for identifying and reducing nonresponse in open-ended questions*. GESIS Papers 2017/01. Köln: GESIS - Leibniz-Institut für Sozialwissenschaften. urn: http://nbn-resolving.de/urn:nbn:de:0168-ssoar-51100-0.

Selected publications:

- Behr, Dorothée, Michael Braun, Lars Kaczmirek, and Wolfgang Bandilla. 2014. "Item comparability in cross-national surveys: results from asking probing questions in cross-national web surveys about attitudes towards civil disobedience." *Quality & Quantity* 48 (1): 127-148. doi: http://dx.doi.org/10.1007/s11135-012-9754-8.

- Braun, Michael, Dorothée Behr, and Lars Kaczmirek. 2013. "Assessing cross-national equivalence of measures of xenophobia: evidence from probing in web surveys." *International Journal of Public Opinion Research* 25 (3): 383-395. doi: http://dx.doi.org/10.1093/ijpor/eds034.

- Meitinger, Katharina (2017). " Necessary but Insufficient - Why measurement invariance tests need online probing as a complementary tool. " *Public Opinion Quarterly* 81 (2): 447-472). Doi: https://doi.org/10.1093/poq/nfx009.

- Meitinger, Katharina, Michael Braun, and Dorothée Behr. 2018. "Sequence matters in online probing: The impact of the order of probes on response quality, motivation of respondents, and answer content." *Survey Research Methods* 12 (2): 103-120. doi: http://dx.doi.org/10.18148/srm/2018.v12i2.7219.

- Meitinger, Katharina, Dorothée Behr, and Michael Braun. 2019. "Using apples and oranges to judge quality?: Selection of appropriate cross-national indicators of response quality in open-ended questions." *Social Science Computer Review* online first 1-22. doi: https://doi.org/10.1177/0894439319859848.