

gesis

Leibniz Institute
for the Social Sciences

Documentation of Face-to-Face Surveys

Jette Schröder & Anouk Zabal

August 2021, Version 1.0

Abstract

Documentation is an integral element of good scientific practice. Transparent and reproducible research with survey data requires a comprehensive and careful documentation of the data collection process and data processing. This guideline gives an overview of key information on the survey, data collection, and data processing that should be included in the documentation of face-to-face surveys. The guideline is most useful when read and incorporated already in the planning phase of a survey to ensure all relevant information for the final documentation is collected during survey implementation.

Citation

Schröder, Jette, & Zabal, Anouk (2021). Documentation of Face-to-Face Surveys. Mannheim, GESIS - Leibniz Institute for the Social Sciences (GESIS - Survey Guidelines).

DOI: 10.15465/gesis-sg_en_041

Acknowledgements

We are very grateful for the helpful comments and suggestions we received from Achim Koch, Silke Martin, and Ines Schaurer.

This work is licensed under a Creative Commons Attribution – NonCommercial 4.0 International License (CC BY-NC).



Documentation is an integral element of good scientific practice. Transparent and reproducible research with survey data requires a comprehensive and careful documentation of both the data collection process and data processing. Setting up an efficient documentation structure and process at the onset of any survey should be considered best practice. Documentation is resource-intensive and often regarded as a tiresome task of little glory. However, without high-quality documentation, it is not possible for external data users or recipients of research results to assess the informative value of the data, and it is difficult, or even impossible, for researchers from outside the project context to fully exploit the data's potential in data analysis.

We would like to emphasize the importance of documentation regardless of whether the data is only analyzed within the context of the project or published for use by third parties. Of course, a small data collection project cannot be expected to produce the same level and detail of documentation as a large-scale survey program. Nevertheless, and especially from an open science perspective, documenting and sharing information is important independent of the size and scope of the survey project. This is essential to allow the data's informative value to be assessed or to replicate the study.

This guideline is most useful when considered already in the planning phase of a survey to ensure all relevant information for the final documentation is collected during survey implementation. A detailed response metric can, for example, only be provided in the final report if the contact documentation form used by interviewers entails all necessary details right from the beginning of the field period and if interviewers document all contacts carefully. Thus, documentation requirements also need to be addressed in the survey protocols and interviewer training. From our own experience we know that establishing a certain discipline in documenting as you go along is in the end more efficient and more accurate than reconstructing certain details months and years after the fact.

This guideline is the third in a series of survey guidelines on documentation recommendations for surveys published in the GESIS Survey Guidelines series. The first two focused on the online (Schaurer, Kunz, & Heycke, 2020) and paper mode (Stadtmüller & Beuthner, 2020), whereas this guideline concentrates on the documentation of face-to-face surveys.¹ The documentation of interviewer-administered surveys – and especially face-to-face surveys – is actually even more demanding than the documentation of self-administered surveys, since, in addition to all the other survey features that need to be covered, information on the interviewers and their work needs to be included. The goal of this survey guideline is to give an overview of key information on the survey, data collection, and data processing that could, and often should, be included in the documentation of face-to-face surveys. As we understand it, survey documentation includes not only reports, but also the original survey materials, for example the advance letter, contact protocols, or interviewer instructions. Data sets with data from contact protocols or paradata are an additional important documentary resource. Beyond the key aspects to be covered by documentation that we focus on in this guideline, there are numerous further possibilities for more detailed survey documentation. These include, for example, the analysis and documentation of various quality indicators, for example on interviewer effects and non-response bias, or the publication of detailed data sets with contact information (an example of a survey with very extensive documentation, including extensive analyses of data quality, is the ESS²).

¹For those interested in the documentation of mixed-mode surveys, we suggest you consult all three guidelines. Beyond the documentation of the different survey modes implemented, mixed-mode studies need to provide additional information on the mixed-mode-nature of the study, for example, whether the modes were offered simultaneously or sequentially and which questionnaire decisions were taken with respect to the different modes (see Stadtmüller, Beuthner, & Silber, 2021; Tourangeau, 2017).

²<https://www.europeansocialsurvey.org>

In the following you will find a comprehensive list of elements that, in our view, should ideally be documented for face-to-face surveys. Not all of them are applicable for every survey, so that survey practitioners and researchers need to extract their own list according to their specific situation and framework, omitting aspects that are not appropriate. For example, if there are no filter questions in the questionnaire, there is no routing information to be specified. However, we do recommend specifying that certain aspects are omitted, and why, both for completeness and because it is often informative for users. For example, if no incentives were offered to target persons, it is helpful to mention this explicitly in the documentation. Of course, the list of key documentation components is not exhaustive either; it may be necessary to add specific elements that may not have been considered in this rather generic list but are relevant in a certain survey context.

The list of aspects to be documented for face-to-face data collections has been organized in nine groups: 1) key facts, 2) survey ethics and data privacy 3) sampling, 4) measurement instruments, 5) paradata and auxiliary data, 6) fieldwork staff, 7) fieldwork, 8) survey metrics, and 9) data processing. The key facts are broad descriptors of the survey, so aspects found here may at times be reiterated with further detail in the other sections. It is neither necessary to provide the required information in the order in which they are listed below nor to provide it in a single document. Large-scale surveys often produce a technical or fieldwork report. Small-scale surveys may produce technical information in a much more succinct form. It is recommended to additionally provide a separate documentation of the survey instruments and codebooks. Note that this is a requirement if the data is to be archived in the GESIS data archive (see Struminskaya, Gauly, Daikeler, Khorshed, & Jedinger, n.d. for recommendations of data documentation based on the regulations of the GESIS data archive). When planning the documentation of a face-to-face survey, we recommend that you also consult the published documentation of large-scale surveys for further inspiration (e.g., ALLBUS³ and PIAAC⁴ for their detailed technical reports, ESS⁵ for its especially comprehensive survey documentation, and pairfam⁶ for its excellent data documentation).

³<https://www.gesis.org/allbus>

⁴<https://www.gesis.org/piaac>

⁵<https://www.europeansocialsurvey.org>

⁶<https://www.pairfam.de>

1 Key Facts

- Project title**
- Key objectives of study**
- Embedding of survey in larger project context**
e.g., a national survey within an international project
- Principal investigators, project team members and affiliations**
- Funding institution**
- Fieldwork agency**
alternative: carried out by survey team
- Survey design**
e.g., cross-sectional, repeated cross-sectional or panel survey (incl. number of waves and time interval between waves)
- Target population**
- Fieldwork period**
start, end, duration
- Survey mode(s)**
CAPI, CASI, Audio-CASI etc.
- Realized sample size**
- Average interview duration**
- Data access**
whether and, if so, where and how data can be obtained by other researchers

2 Survey Ethics and Data Privacy

- Ethics approval**
- Consent**
e.g., informed consent to survey participation, supplementary consent for data linkage, incl. how target person is informed (e.g., data protection sheet) and specifics of consent procedure

3 Sampling

- Target population**
- Type of sampling**
probability, non-probability
- Sample size**
gross sample, target net sample size

Probability samples

- sampling frame and coverage error
- detailed description of sampling stages and selection procedure incl. questions used to select/identify target person(s)

Non-probability samples

- recruitment methods and selection criteria

4 Measurement Instruments

Instrument components and topic areas

Language versions and translation method

- documentation of translation procedures⁷
- quality control

Specifics of measurement instrument(s)

- detailed documentation of measurement instrument(s); best practice is publication of survey questionnaire in all language versions (including showcards) with
 - questions
 - response alternatives, including e.g., information on which alternatives are to be read out, when showcards are administered
 - interviewer instructions and help instructions
 - routing or randomisation specifications
 - plausibility and consistency checks
- question/item source
e.g., questionnaire items/modules from other surveys, published scale, newly developed items/modules
- if possible, screenshots of CASI-elements, screen capture of multimedia or interactive elements
- design elements, e.g., experimental design, assignment of respondents to groups

Codebook

variable names, response codes (may be documented together with specifics of measurement instrument in one file)

Survey device and software

Information on pretest/cognitive pretest/pilot studies

e.g., number and sampling of cases, timing, procedure, objective, consequences

Interview duration (questionnaire/other instrument components)

⁷For guidelines on the documentation of survey translation procedures, see Behr and Zabal (2020).

5 Paradata and Auxiliary Data

- Include key auxiliary data in final data set**
e.g., date of interview, interviewer identification number
- Document additional data (type and content)**
 - process data, e.g., time stamps
 - observational data, e.g., interviewers' observations of neighborhood conditions
 - data from external sources, e.g., administrative data, data from registries

6 Fieldwork Staff

- Interviewers**
e.g., selection criteria, number, attrition, characteristics (incl. experience), payment scheme
- Supervisors**
e.g., number of supervisors, training
- Interviewer training**
e.g., length, mode, sessions, main modules
- Interviewer materials and support**
e.g., interviewer manual

7 Fieldwork

- Data collection period**
start, end, duration
- Fieldwork phases**
incl. re-issues
- Incentives**
e.g., monetary/non-monetary, prepaid/ postpaid, value, handling
- Advance contact and information material for target person**
e.g., advance letter, study flyer/brochure, data protection sheet, study website, hotline
- Contact procedures and contact protocol**
- Major problems affecting fieldwork**
incl. strategies implemented to cope with problems
- Fieldwork monitoring**
e.g., key performance indicators and monitoring intervals
- Procedural changes during fieldwork**
e.g., incentive change, use of traveling interviewers
- Quality control**
e.g., validation of interviews (in-person, phone, mail), checks of interview length, times

8 Survey Metrics

- Realized sample size**
- Outcome rates**
response rate, cooperation rate, refusal rate, contact rate, ideally calculated following AAPOR-scheme⁸
- Distribution of final disposition codes for gross sample**
ideally following AAPOR-scheme or an appropriate adaptation of AAPOR-scheme⁹
- Contact attempts**
e.g., average number of contact attempts for respondents and nonrespondents, overall distribution of total contact attempts
- Interviewer case load**
average case load/average number of interviews per interviewer
- Sample characteristics**
basic sociodemographics, representation of target population, nonresponse bias measures

9 Data Processing

- Data entry**
for paper questionnaires/instruments: automatic or manual data entry, software, number people entering data, data entry rules, quality control
- Data cleaning and editing**
e.g., checks and handling of errors and inconsistencies (such as accepting, flagging, and recoding variables, or excluding cases)
- Coding of open-ended questions**
e.g., coding institution, number of coders, coding scheme, coder training, quality control, documentation of reliability measures (such as Cohen's Kappa)
- Derived variables**
purpose and construction of additional variables included in final dataset; if possible, publish syntax
- Imputation procedure of missing values**
variables and description of procedure for variables included in final dataset; if possible, publish syntax
- Weighting variables**
purpose and calculation; if possible, publish syntax
- Anonymization of variables**
description of confidentiality edits, e.g., suppression or coarsening of variables
- Linkage of additional data**
linkage procedure and results

⁸See American Association for Public Opinion Research (2016).

⁹See American Association for Public Opinion Research (2016); for a possible German adaptation see Stadtmüller et al. (2019).

References

- American Association for Public Opinion Research. (2016). *Standard definitions: Final dispositions of case codes and outcome rates for surveys*. Retrieved from https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf
- Behr, D., & Zabal, A. (2020). *Documenting survey translation*. Mannheim, GESIS – Leibniz Institute for the Social Sciences (GESIS – Survey Guidelines). https://doi.org/10.15465/gesis-sg_en_035
- Schaurer, I., Kunz, T., & Heycke, T. (2020). *Documentation of online surveys*. Mannheim, GESIS – Leibniz Institute for the Social Sciences (GESIS – Survey Guidelines). https://doi.org/10.15465/GESIS-sg_en_031
- Stadtmüller, S., & Beuthner, C. (2020). *Documentation of mail data collection*. Mannheim, GESIS – Leibniz Institute for the Social Sciences (GESIS – Survey Guidelines). https://doi.org/10.15465/gesis-sg_en_032
- Stadtmüller, S., Beuthner, C., & Silber, H. (2021). *Mixed-Mode Surveys*. Mannheim, GESIS – Leibniz Institute for the Social Sciences (GESIS – Survey Guidelines). https://doi.org/10.15465/gesis-sg_en_038
- Stadtmüller, S., Silber, H., Daikeler, J., Martin, S., Sand, M., Schmich, P., ... Zabal, A. (n.d.). *Adaptation of the AAPOR final disposition codes for the German survey context*. Mannheim, GESIS – Leibniz Institute for the Social Sciences (GESIS – Survey Guidelines). https://doi.org/10.15465/gesis-sg_en_026
- Struminskaya, B., Gauly, B., Daikeler, J., Khorshed, J., & Jedinger, A. (n.d.). *Survey data documentation*. Mannheim, GESIS – Leibniz Institute for the Social Sciences (GESIS – Survey Guidelines). https://doi.org/10.15465/gesis-sg_en_24
- Tourangeau, R. (2017). Mixing modes: Tradeoffs among coverage, nonresponse, and measurement error. In P. P. Biemer, E. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L. E. Lyberg, ... B. T. West (Eds.), *Total survey error in practice* (pp. 115--132). <https://doi.org/10.1002/9781119041702.ch6>