

Interviewer Effects in Standardized Surveys (Version 1.0)

Ackermann-Piek, Daniela; Schröder, Jette; Kluge, Rebekka; Bieber, Ina

Erstveröffentlichung / Primary Publication

Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Ackermann-Piek, D., Schröder, J., Kluge, R., & Bieber, I. (2019). *Interviewer Effects in Standardized Surveys (Version 1.0)*. (GESIS Survey Guidelines). Mannheim: GESIS - Leibniz-Institut für Sozialwissenschaften. https://doi.org/10.15465/gesis-sg_en_027

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC Lizenz (Namensnennung-Nicht-kommerziell) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by-nc/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC Licence (Attribution-NonCommercial). For more information see: <https://creativecommons.org/licenses/by-nc/4.0>

gesis

Leibniz Institute
for the Social Sciences

Interviewer Effects in Standardized Surveys

Daniela Ackermann-Piek, Jette Schröder, Rebekka Kluge, & Ina Bieber

September 2019, Version 1.0

Abstract

Concerns about interviewer effects in interviewer-mediated surveys have accompanied survey research for a long time. As interviewers are involved in nearly all aspects of the survey implementation process, they can affect almost all types of survey errors, including sampling error, nonresponse error, measurement error, and, to a lesser extent, error resulting from the coding and editing of survey responses. Building on the existing literature, this survey guideline provides an overview of interviewer effects and their estimation. It consists of two parts: first, an introductory text using the total survey error (TSE) paradigm as a theoretical framework to provide a general overview of interviewer effects; second, a brief introduction to calculating interviewer effects using multilevel analyses.

Citation

Ackermann-Piek, D., Schröder, J., Kluge, R., & Bieber, I. (2019). Interviewer Effects in Standardized Surveys. Mannheim, GESIS - Leibniz Institute for the Social Sciences (GESIS- Survey Guidelines).

DOI: [10.15465/gesis-sg_en_027](https://doi.org/10.15465/gesis-sg_en_027)

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



Interviewer Effects: Theoretical Concept and Empirical Findings

Surveys are an important tool for systematically gathering information about people's opinions, attitudes, knowledge, and behaviors. Survey data can be collected through different survey modes: Surveys can be interviewer-administered, as in the case of face-to-face and telephone modes, or they can be self-administered by respondents, as in the case of postal mail and online modes. Although the number of online surveys has increased over the last ten years, over 50% of all surveys conducted by ADM members in Germany in 2018 were still administered by interviewers (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. (ADM), 2018).¹

During the administration of a survey, interviewers' tasks can range from generating sampling frames, through gaining sample persons' cooperation and maintaining their motivation during the interview, to recording answers and other measurements. Compared to self-administered modes, the administration of a survey by interviewers can be advantageous in several ways. For example, interviewers reduce unit nonresponse (e.g., Durrant, Groves, Staetsky, & Steele, 2010; Lord, Friday, & Brennan, 2005; Webster, 1996). In addition, they positively affect response quality – for example, by probing adequately when respondents provide inadequate responses, or by clarifying questions when respondents have problems understanding them, thereby reducing the number of inappropriate answers (Belli, Weiss, & Lepkowski, 1999; Hanson & Marks, 1958; Schober & Conrad, 1997). Although the survey process and the role of the interviewers are standardized in order to minimize potential interviewer error in interviewer-administered standardized surveys,² interviewers may also have negative effects on data quality, for example by introducing measurement error (e.g., Ackermann-Piek, 2018; Mangione, Fowler Jr., & Louis, 1992; Schnell & Kreuter, 2005). Moreover, interviewers may differ in the extent to which they are able to realize positive effects, for example by gaining respondents' consent to link survey data with administrative data (e.g., Ackermann-Piek, 2018; Korbmacher & Schroeder, 2013; Sakshaug, Tutz, & Kreuter, 2013). Strictly speaking, all positive and negative interviewer effects can be subsumed under the concept of interviewer effects. In practice, however, only negative interviewer effects or interviewer variability regarding different aspects of data quality are usually addressed in this context. Thus, we follow this approach in the present survey guideline.

Two types of interviewer effects are differentiated in the literature. The first type is referred to as *interviewer bias* (Groves, 2004; Groves et al., 2009). Interviewer bias occurs when specific aspects of the data collection process are influenced by all interviewers in the same undesired direction. For example, the presence of interviewers can increase the social desirability bias of respondents' answers. In a study by Kreuter, Presser, & Tourangeau (2008), respondents claimed to have a higher grade point average (GPA) than they had actually achieved; and Chang & Krosnick (2009) found that respondents were more likely to say that the federal government should provide more help for African Americans if the survey was administered by an interviewer. Interviewer bias leads to biased estimators across all interviewers.

The second type of interviewer effect is referred to as *interviewer variance* (Groves, 2004; Groves et al., 2009). Even in standardized surveys, interviewers may differ in the extent to which they cause specific survey errors, as they vary in their characteristics, attitudes, and working style. Consequently, although estimators are not necessarily biased, these between-interviewer differences inevitably lead to greater variability of measurements and less efficient estimators. Interviewer variance can increase the variance

¹"ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. is a business association that represents the interests of private sector market and social research agencies in Germany. ADM members account for more than 80 percent of industry turnover. ADM is the only association of this kind in Germany." <https://www.adm-ev.de/en/>

²The present survey guideline focuses on standardized surveys with a standardized interviewing style. However, other interviewing styles, such as the conversational interviewing approach proposed by Schober and Conrad (2002) and the personal interviewing style proposed by (Dijkstra, 1987), are also implemented in practice.

of survey estimates of a simple random sample compared to that of a simple random sample without interviewer clustering; this is referred to as the interviewer design effect.

Most of the literature on interviewer effects focuses on interviewer variance – for example, in estimates of survey variables or in the response rates achieved (e.g., Biemer & Lyberg, 2003; Loosveldt, 2008; West & Blom, 2017). Only a few studies have investigated interviewer bias. The reason for this is that studying interviewer bias usually requires a true value as a reference, which is only rarely available. Therefore, the focus in the literature is not on interviewer performance problems per se, but rather on the variability among interviewers in introducing or reducing errors in different survey outcomes.

Interviewer Effects Within the Context of the Total Survey Error Framework

“Total survey error’ (TSE) is a conceptual framework describing statistical error properties of sample survey statistics” (Groves & Lyberg, 2010, p. 849). It incorporates sources of error that may occur throughout the lifecycle of a survey (Groves et al., 2009; Groves & Lyberg, 2010). TSE divides the sources of survey error into two main dimensions: measurement and representation: “The measurement dimension describes what data are to be collected about the observational units in the sample – what is the survey about? The representational dimension concerns what populations are described in the survey – who is the survey about?” (for further details regarding the TSE framework see Groves et al., 2009).

As interviewers can be involved in almost all aspects of the survey data collection process, interviewer effects can occur in nearly all of the sources of survey error described in the TSE framework (see Figure 1). Interviewers may contribute to coverage or sampling error if they are involved in developing frames for area probability samples for face-to-face surveys – for example, by building or refining address lists (e.g., Eckman, 2013; Kwiat, 2009; West & Blom, 2017). Furthermore, interviewers make contact with selected sample persons and gain their cooperation, which is why they may have a strong influence on nonresponse error (Blom, Leeuw, & Hox, 2011; Durrant, D’Addio, & Steele, 2013; Jäckle, Lynn, Sinibaldi, & Tipping, 2013). During the interview, measurement error can be introduced by interviewers if they vary, for example, in the extent to which they apply standardized interviewing techniques (e.g., Sinibaldi, Durrant, & Kreuter, 2013) or maintain respondent motivation (e.g., Tourangeau & Smith, 1996). Finally, interviewer effects on processing error can occur due to variability among interviewers in the coding and recording of respondents’ answers and measurements (e.g., Durrant et al., 2010; Rice, 1929; Smyth & Olson, 2015).

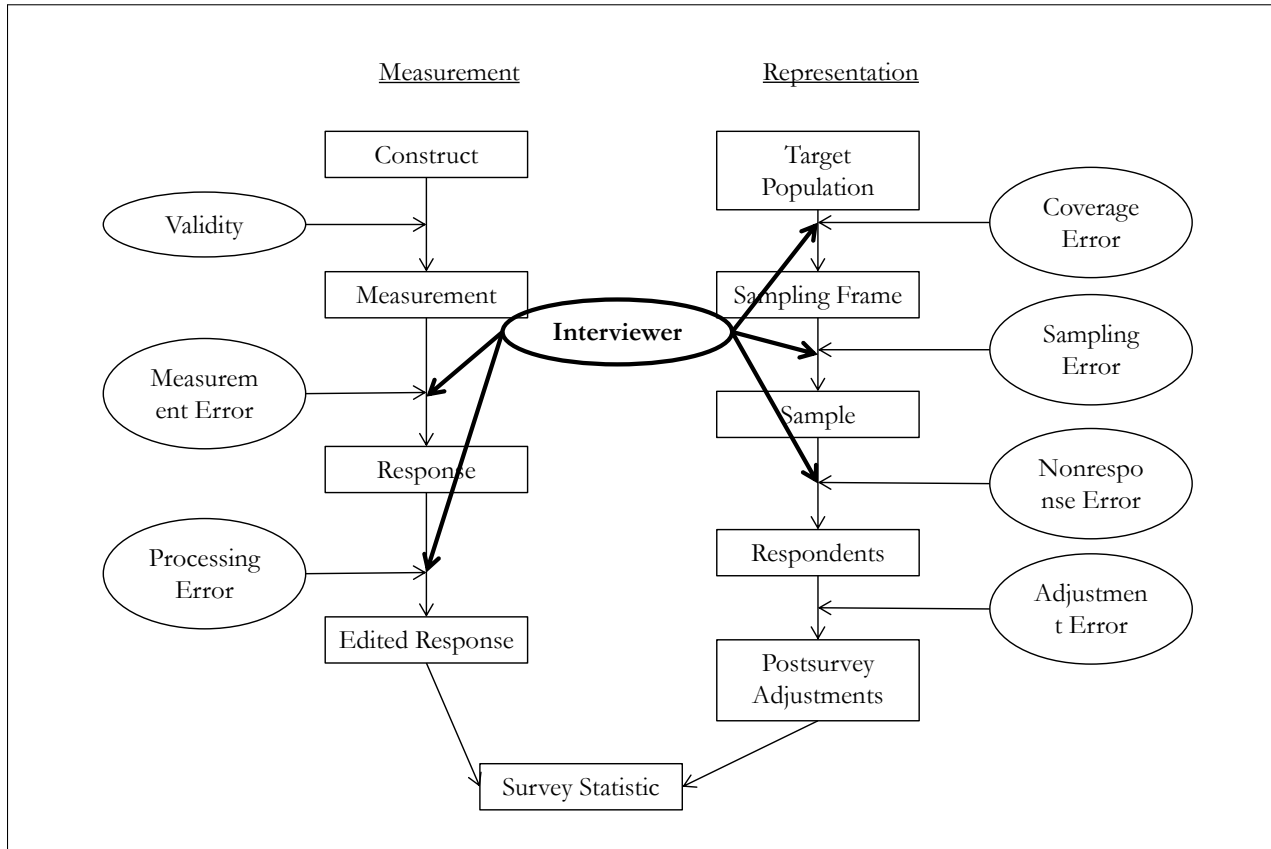


Figure 1: Interviewer effects in terms of the total survey error framework from (Ackermann-Piek, 2018)
 Note: Adapted from *Survey Methodology* (Vol. 2, p. 48) by R. M. Groves, F. J. J. Fowler, M. P. Lepkowski, J. M. Singer, and R. Tourangeau (2009). Hoboken, NJ: Wiley & Sons.

Explaining Interviewer Effects

Besides studies that have simply reported evidence of variance in interviewer effects on different components of TSE, various studies have sought to explain this variance on the basis of interviewer characteristics (for an overview, see West & Blom, 2017). Most of these studies have focused on interviewer characteristics such as gender, age, or experience to explain the variability among interviewers in different types of survey errors (e.g., Loosveldt & Beullens, 2014; Purdon, Campanelli, & Sturgis, 1999). This focus is due mainly to the fact that these interviewer characteristics are usually easy to access. There are also findings relating to other interviewer characteristics, such as speech rate (e.g., Steinkopf, Bauer, & Best, 2010), reading behaviors (e.g., Bergmann & Bristle, 2016), motivation and attitudes (Schröder, Jette and Schmiedeberg, Claudia and Castiglioni, Laura, 2016), and physical attractiveness (e.g., Nedelec, 2017). However, it has been noted that interviewer effects cannot be explained solely by characteristics of interviewers (West & Blom, 2017). Interviewer variance can also be influenced by the characteristics of the sample persons, or, more specifically, by the match between the characteristics of interviewers and sample persons. For example, a gender and education match between the interviewer and the sample person has been found to increase cooperation rates (Durrant et al., 2010). However, most characteristics of sample persons are not known prior to the survey, which makes matching quite complicated for most surveys.

Overall, the literature suggests that the effects of interviewers' characteristics are not consistent across

surveys (for an overview see Ackermann-Piek, 2018; West & Blom, 2017). Instead, interviewer effects seem to be survey-specific, and other survey characteristics, such as the topic or the sponsor of the survey, may moderate the identified interviewer effects (Ackermann-Piek, Blom, Korbmacher, & Krieger, 2019). However, the following interviewer characteristics have been found to contribute to the explanation of interviewer effects in the majority of the studies: race/ethnicity, working experience as an interviewer, interviewing styles, and sociodemographic similarities between the interviewer and the sample person or the interviewer and the respondent.³

Calculation of Interviewer Variance

In this section, we provide a short introduction on how multilevel analysis can be used to model interviewer variance on different survey outcomes, and how interviewer design effects can be calculated. Obviously, this brief overview cannot replace textbooks on multilevel models.

To avoid confusion, a note on our terminology before we continue: The more specific term “interviewer variance” is seldom used in the literature on interviewer effects. Instead, when referring to interviewer variance, the broader term “interviewer effects” is used. This is probably due to the fact that the vast majority of studies on interviewer effects deal with interviewer variance, whereas the other type of interviewer effects, interviewer bias, is rarely studied. Similar to the practice in the literature, we use in what follows the specific term “interviewer variance” and the broader term “interviewer effects” as synonyms.

General Approach to Calculating Interviewer Variance

Typically, survey data from interviewer-administered surveys are characterized by their hierarchical structure: Sample persons or respondents (level one) are nested within interviewers (level two). The theoretical assumption behind multilevel analysis is that these levels are not independent of each other: Survey outcomes of respondents interviewed by the same interviewer may be more similar to each other than survey outcomes of respondents interviewed by different interviewers. Due to the nested data structure, variance can be produced at each level. In contrast to ordinary, single-level analysis, multilevel analyses account for joint effects of the different levels (Hox & Kreft, 1994). In some cases, the data structure is even more complex. For example, sample persons or respondents (level one) are nested within geographical areas (level two), and areas are nested within interviewers (level three). If an additional level does not contribute to the explained variance, the simpler model should be implemented in order to reduce complexity. Another example is a cross-classified data structure, which is most common in computer-assisted personal interviewing (CAPI) studies of the general population: Sample persons or respondents (level one) are nested within areas (level two) and interviewers (level three), but areas and interviewers are crossed with each other. In other words, one interviewer works in more than one area, and one area is worked by more than one interviewer.

If a clustered data structure is not modeled correctly by applying a three-level model, interviewer and area effects are confounded, and thus the variance located at the interviewer level is overestimated. In other words, part of the variance in the outcome measure allocated to the interviewer level may actually be located at the area level. Thus, when modeling interviewer effects, we recommend using all information available.

Furthermore, like single-level regression, multilevel models can be applied for dependent variables of different scale levels, such as dichotomous or metric variables. Independent of the type of model, the

³The term “sample person” refers to a person selected from the sampling frame to participate in a survey, whereas the term “respondent” refers to a person who actually participates in the survey.

basic idea of using multilevel models to estimate interviewer effects is to differentiate which part of the variance of the dependent variable is located at the interviewer level.

When the focus is on disentangling interviewer effects and area effects in interviewer-administered surveys, the random assignment of sample persons to interviewers – a so-called interpenetrated design – is required (Campanelli & O’Muircheartaigh, 1999). However, fully interpenetrated designs, which typically ensure that there is no systematic difference between the assigned sample persons, are not feasible in practice – for example, because interviewers work only in specific areas (and the budget is not sufficient to eliminate this restriction). Even in telephone surveys that are conducted in centralized computer-assisted telephone interview (CATI) studios, practical obstacles such as time restrictions on interviewers’ work schedules may result in the implementation of an imperfectly interpenetrated design. In the case of face-to-face surveys, the problem of non-random assignment of sample persons to interviewers is often addressed during fieldwork by assigning more than one interviewer to one area in order to be able to disentangle area and interviewer effects. However, even in this case, it is essential to specify multilevel models carefully by including controls for sample composition effects, such as area size and sample person characteristics (Schaeffer, Dykema, & Maynard, 2010; West & Blom, 2017).

The Intra-Class Correlation Coefficient: The Concrete Measure of Interviewer Variance

The concrete measure that quantifies interviewer variance – that is, the proportion of total variance in survey outcomes caused by differences between interviewers – is the intraclass correlation coefficient ([ICC], West & Blom, 2017). In the case of interviewer effects, the ICC informs about the proportion of the total variance in the specified survey outcome explained by interviewer clustering (Hox, 2010). In a two-level model, with sample persons or respondents at level one and interviewers at level two, the ICC can be calculated as follows:

$$ICC = \rho_{int} = \frac{VAR_{BETWEEN}}{VAR_{BETWEEN} + VAR_{WITHIN}} \quad (1)$$

where $VAR_{BETWEEN}$ is the between-interviewer variance – that is, the variance in the outcome variable that is due to the differences between the interviewers. VAR_{WITHIN} is the within-interviewer variance between sample persons or respondents – that is, the variance in the outcome variable that is due to the differences between sample persons or respondents within the interviewers (Hox, 2010, p. 15). $VAR_{BETWEEN}$ and VAR_{WITHIN} add up to the total variance of the specified survey outcome, VAR_{TOTAL} . As a rule of thumb, an ICC under .05 ($\cong 5\%$) can be interpreted as a low interviewer effect for the respective outcome variable where no action is needed. An ICC greater .10 ($\cong 10\%$) indicates that a considerable amount of variance in the outcome variable is located at the interviewer level. Thus, when an ICC is greater than .10 ($\cong 10\%$), we recommend taking a closer look at possible reasons for interviewer effects. Interviewer effects between .05 and .10 can be interpreted as moderate, and decisions on further action can be made on a case-by-case basis.

As interviewer effects can lead to inflated variance in survey estimates, it is useful to report the interviewer design effect ($def f_{int}$), which is a function of the interviewer variance, ρ_{int} , and the average number of interviews across all interviewers, m (Biemer & Lyberg, 2003; Kish, 1962). It can be calculated in a two-level model as follows:

$$def f_{int} = 1 + \rho_{int}(m - 1) \quad (2)$$

For easier interpretation, the interviewer design factor is typically calculated (Schnell & Kreuter, 2005). It represents the factor by which interviewer variance increases the variance of survey estimates of a simple random sample compared to that of a simple random sample without interviewer clustering:

$$def t_{int} = \sqrt{def f_{int}} \quad (3)$$

The consequence of an increase in the variance is that the effective sample size is reduced. Following West & Blom (2017), the effective sample size can be calculated as follows:

$$n_{eff} = \frac{n}{1 + \rho_{int}(m - 1)} \quad (4)$$

For more details, please consult textbooks on multilevel models (Hox & Kreft, 1994; Hox, Moerbeek, & Van de Schoot, 2017; Rabe-Hesketh & Skrondal, 2012; Skrondal & Rabe-Hesketh, 2003).

Prevention of Interviewer Effects

It should be emphasized that the best strategy for dealing with interviewer effects is to prevent their occurrence as far as possible. Commonly, standardization is used to reduce interviewer effects in interviewer-administered surveys. Standardization applies to all aspects of the survey: the survey instrument, the implementation standards, and all aspects of interviewers' work. With regard to interviewers' work, one option for realizing standardization is to implement targeted and standardized interviewer training addressing all aspects of interviewers' involvement in the survey (Daikeler, Silber, Bosnjak, Zabal, & Martin, 2017; Stiegler & Biedinger, 2016). For example, Billiet & Loosveldt (1988) and Groves & McGonagle (2001) have shown that detailed interviewer training and adequate interviewer remuneration reduce interviewer variance in various survey outcomes. To reduce the effects that interviewers can have on estimates of survey outcomes, the number of interviews per interviewer should be limited to a maximum of between 10 and 50, depending on interviewer experience and training (Loosveldt, 2008; Schnell, 2012). In addition, to reduce social desirability, interviewer-administered parts of the survey can be supplemented with self-administered parts, for example for sensitive topics. Moreover, interviewers' work should be closely monitored to detect falsifications or any other problems that may affect data quality. Finally, as the advantages of interviewer involvement in standardized surveys described in section 1 outweigh the fact that, even with great effort, it is usually impossible to fully avoid interviewer effects, analyzing interviewer effects remains relevant.

References

- Ackermann-Piek, D. (2018). *Interviewer effects in PIAAC Germany 2012* (PhD thesis). Retrieved from <https://ub-madoc.bib.uni-mannheim.de/42772/>
- Ackermann-Piek, D., Blom, A. G., Korbmacher, J. M., & Krieger, U. (2019, February). *Explaining interviewer effects on unit nonresponse: A cross-survey analysis*. Poster session presented at the Interviewer Workshop, Lincoln, Nebraska.
- Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. (ADM). (2018). Retrieved from <https://www.adm-ev.de/en/>
- Belli, R. F., Weiss, P. S., & Lepkowski, J. M. (1999). Dynamics of survey interviewing and the quality of survey reports: Age comparisons. In N. Schwarz, D. Park, B. Knauper, & S. Sudman (Eds.), *Cognition, aging, and self-reports*. Hove, UK: Psychology Press/Erlbaum (UK) Taylor & Francis.
- Bergmann, M., & Bristle, J. (2016). *Do interviewers' reading behaviors influence survey outcomes? Evidence from a cross-national setting*. (SHARE Working Paper Series 23-2016). Retrieved from http://www.share-project.org/uploads/tx_sharepublications/WP_Series_23_2016_Bristle_Bergmann_01.pdf.
- Biemer, P. P., & Lyberg, L. (2003). *Introduction to survey quality* (Vol. 335). New York, NY: John Wiley & Sons.
- Billiet, J., & Loosveldt, G. (1988). Improvement of the quality of responses to factual survey questions by interviewer training. *Public Opinion Quarterly*, *52*, 190–211.
- Blom, A. G., Leeuw, E. D. de, & Hox, J. J. (2011). Interviewer effects on nonresponse in the European Social Survey. *Journal of Official Statistics*, *27*(2), 359–377.
- Campanelli, P., & O'Muircheartaigh, C. (1999). Interviewers, interviewer continuity, and panel survey non-response. *Quality & Quantity*, *33*, 59–76. <https://doi.org/10.1023/A:1004357711258>
- Chang, L., & Krosnick, J. A. (2009). National surveys via RDD telephone interviewing Versus the internet comparing sample representativeness and response quality. *Public Opinion Quarterly*, *73*(4), 641–678. <https://doi.org/doi:10.1093/poq/nfp075>
- Daikeler, J. X., Silber, H., Bosnjak, M., Zabal, A., & Martin, S. (2017). General interviewer training curriculum for computer-assisted personal interviews (GIT-CAPI; Version 1, 2017). *GESIS Survey Guidelines. Mannheim, Germany: GESIS – Leibniz-Institute for the Social Sciences*. https://doi.org/10.15465/gesis-sg_en_022
- Durrant, G. B., D'Addio, J., & Steele, F. (2013). Analysing interviewer call record data by using a multilevel discrete-time event history modeling approach. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *176*(1), 251–269. <https://doi.org/10.1111/j.1467-985X.2012.01073.x>
- Durrant, G. B., Groves, R. M., Staetsky, L., & Steele, F. (2010). Effects of interviewer attitudes and behaviors on refusal in household surveys. *Public Opinion Quarterly*, *74*(1), 1–36. <https://doi.org/10.1093/poq/nfp098>
- Eckman, S. (2013). Do different listers make the same housing unit frame? Variability in housing unit listing. *Journal of Official Statistics*, *29*(2), 249–259. <https://doi.org/10.2478/jos-2013-0021>
- Groves, R. M. (2004). *Survey error and survey costs*. Hoboken, NJ: John Wiley & Sons.

- Groves, R. M., Fowler, F. J. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (Vol. 2). Hoboken, NJ: John Wiley & Sons.
- Groves, R. M., & Lyberg, L. (2010). Total survey error past, present and future. *Public Opinion Quarterly*, 74(5), 849–879. <https://doi.org/10.1093/poq/nfq065>
- Groves, R. M., & McGonagle, K. (2001). A theory-guided interviewer training protocol regarding survey participation. *Journal of Official Statistics*, 17(2), 249–265.
- Hanson, R. H., & Marks, E. S. (1958). Influence of the interviewer on the accuracy of survey results. *Journal of the American Statistical Association*, 53(283), 635–655. <https://doi.org/10.1080/01621459.1958.10501465>
- Hox, J. J. (2010). *Multilevel analysis - Techniques and applications* (2nd ed., pp. 1–368). <https://doi.org/10.4324/9780203852279>
- Hox, J., & Kreft, I. G. G. (1994). Multilevel analysis methods. *Sociological Methods & Research*, 22, 283–299. <https://doi.org/10.1177/0049124194022003001>
- Hox, J. P., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: techniques and applications* (3rd ed.). New York, NY: Routledge.
- Jäckle, A., Lynn, P., Sinibaldi, J., & Tipping, S. (2013). The effect of interviewer experience, attitudes, personality and skills on respondent co-operation with face-to-face surveys. *Survey Research Methods*, 7(1), 1–15.
- Kish, L. (1962). Studies of interviewer variance for attitudinal variables. *Journal of the American Statistical Association*, 57(297), 92–115.
- Korbmacher, J. M., & Schroeder, M. (2013). Consent when linking survey data with administrative records: The role of the interviewer. *Survey Research Methods*, 7(2), 115–131.
- Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys: The effects of mode and question sensitivity. *Public Opinion Quarterly*, 72(5), 847–865.
- Kwiat, A. (2009). Examining blocks with lister error in area listing. *Proceedings of the Joint Statistical Meetings of the American Statistical Association, Section on Survey Research Methods*, 2546–2557.
- Loosveldt, G. (2008). Face-to-face interviews. In E. D. De Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 201–220). New York, NY: Routledge.
- Loosveldt, G., & Beullens, K. (2014). A procedure to assess interviewer effects on nonresponse bias. *SAGE Open*, 4(1), 1–12. <https://doi.org/10.1177/2158244014526211>
- Lord, V. B., Friday, P. C., & Brennan, K., Pauline. (2005). The effects of interviewer characteristics on arrestees' responses to drug-related questions. *Applied Psychology in Criminal Justice*, 1(1), 36–55.
- Mangione, T. W., Fowler Jr., F. J., & Louis, T. A. (1992). Question characteristics and interviewer effects. *Journal of Official Statistics*, 8(3), 293–307. Retrieved from <http://www.jos.nu/Articles/article.asp>
- Nedelec, J. L. (2017). A multi-level analysis of the effect of interviewer characteristics on survey respondents' reports of sensitive topics. *Personality and Individual Differences*, 107, 96–101. <https://doi.org/10.1016/j.paid.2016.11.036>
- Purdon, S., Campanelli, P., & Sturgis, P. (1999). Interviewers' calling strategies on face-to-face interview surveys. *Journal of Official Statistics*, 15(2), 199–216.

- Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modeling using Stata* (3rd ed.). Lake-way Drive, College Station, TX: Stata Press.
- Rice, S. (1929). Contagious bias in the interview: A methodological note. *American Journal of Sociology*, 35(3), 420–423.
- Sakshaug, J. W., Tutz, V., & Kreuter, F. (2013). Placement, wording, and interviewers: Identifying correlates of consent to link survey and administrative data. *Survey Research Methods*, 7(2), 133–144.
- Schaeffer, N. C., Dykema, J., & Maynard, D. W. (2010). Interviewers and interviewing. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (pp. 437–470). Bingley, UK: Emerald Group Publishing Limited.
- Schnell, R. (2012). *Survey-Interviews. Methoden standardisierter Befragungen*. Wiesbaden: VS Verlag.
- Schnell, R., & Kreuter, F. (2005). Separating interviewer and sampling-point effects. *Journal of Official Statistics*, 21(3), 389–410.
- Schober, M. F., & Conrad, F. G. (1997). Does conversational interviewing reduce survey measurement error? *Public Opinion Quarterly*, 61(4), 576–602.
- Schröder, Jette and Schmiedeberg, Claudia and Castiglioni, Laura. (2016). The effect of interviewers motivation and attitudes on respondents consent to contact secondary respondents in a multi-actor design. *Survey Methods: Insights from the Field*, 1–14.
- Sinibaldi, J., Durrant, G. B., & Kreuter, F. (2013). Evaluating the measurement error of interviewer observed paradata. *Public Opinion Quarterly*, 77(Special Issue), 173–193.
- Skrondal, A., & Rabe-Hesketh, S. (2003). Multilevel logistic regression for polytomous data and rankings. *Psychometrika*, 68(2), 267–287.
- Smyth, J., & Olson, K. (2015). *Recording what the respondent says: Does question format matter? Paper presented at the 2015 Annual Conference of the American Association for Public Opinion Research, , Hollywood, FL.*
- Steinkopf, L., Bauer, G., & Best, H. (2010). Nonresponse und Interviewer-Erfolg im Telefoninterview: Empirische Untersuchungen zum Einfluss stimmlicher Eigenschaften der Interviewer. *Methoden, Daten, Analysen*, 4(1), 3–26.
- Stiegler, A., & Biedinger, N. (2016). Interviewer skills and training (Version 2.0). *GESIS Survey Guidelines. Mannheim, Germany: GESIS – Leibniz-Institute for the Social Sciences*. https://doi.org/10.15465/gesis-sg_en_01
- Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and questions context. *Public Opinion Quarterly*, 60, 275–304. <https://doi.org/10.1086/297751>
- Webster, C. (1996). Hispanic and Anglo interviewer and respondent ethnicity and gender: The impact on survey response quality. *Journal of Marketing Research*, 33(1), 62–72.
- West, B. T., & Blom, A. G. (2017). Explaining interviewer effects: A research synthesis. *Journal of Survey Statistics and Methodology*, 5(2), 175–211.