

# How to reduce Item Nonresponse in Face-to-Face Surveys? A Review and Evidence from the European Social Survey

Malte Grönemann

University of Mannheim, Mannheim, Germany

4th July 2022

## **Abstract**

I review the literature on item nonresponse in surveys and based on this review, I extend the satisficing model with respondents' privacy concerns to incorporate all relevant aspects of the response process for item nonresponse. I review proposed strategies to reduce item nonresponse and test selected strategies. Results suggest that boosting respondents' use of showcards, shorter interview durations and conducting the interview in the respondents' primary language might be promising ways to reduce item nonresponse. Other people present during the interview only has a small association with the number of refusals. Matching age and gender of respondents and interviewers appears not to be a worthwhile strategy.

Keywords: item nonresponse, survey design, survey error, face-to-face surveys, fixed-effects negative binomial regression

## 1 Introduction

Missing data pose a problem to the analysis of survey data. They decrease the effective sample size and can introduce bias to estimates if the causes for missingness are related to the item or respondent characteristics (de Leeuw, Hox and Huisman, 2003). Although missing data are relatively rare in most single items, they can add up to a considerable loss of observations in multivariate analyses. Due to these harmful effects of missing data, one objective of survey researchers is to keep their prevalence as low as possible. Therefore, it is important to understand the processes that can lead to missing data.

I focus on the respondent as a source of missing data by answering with “Don’t know” (DK) or refusing to answer. (*item nonresponse*). Item nonresponse can also be seen as an indicator of overall data quality since it can be the result of *satisficing* (Krosnick, 1991). Satisficing means that the respondent is giving a satisfactory answer instead of the best one possible.

I review the existing literature on item nonresponse and extend the satisficing model based on this review to include privacy concerns resulting in an encompassing model of item nonresponse: the probability of item nonresponse depends on the task difficulty of the item(s) divided by the product of ability and motivation of the respondent or the respondents’ privacy concerns. This means that higher difficulty results in more item nonresponse, higher ability and motivation in less item nonresponse. Difficulty, ability and

motivation are separate from privacy concerns, e.g. due to item sensitivity or general mistrust. Privacy concerns are relevant when deciding whether to disclose an already formed answer and does not influence the cognitive burden of retrieving the answer.

In the empirical part of the paper, I explore a few practical ideas to reduce item nonresponse. Promoting the use of showcards, shorter interviews and translating questionnaires might be promising strategies. All of those reduce the cognitive burden of respondents. The effects are not causal though because I am not able to sufficiently control for respondent ability. Matching respondents' and interviewers' gender and age seem not to be efficient tools to reduce item nonresponse. They could have influenced item nonresponse if respondents are more willing to share private information with interviewers that are similar to them.

This introduction follows a short theoretical discussion presenting a model for the probability of item nonresponse based on satisficing. Structured by this model, I review the literature on strategies how to reduce item nonresponse and test a selection using the *European Social Survey*. The rest of the paper is devoted to this test of strategies, describing data and methods and presenting results. Finally, I summarise and discuss my review and results.

## 2 A Model of Item Nonresponse

The *survey response process* (Tourangeau, Rips and Rasinski, [2000](#)) involves multiple steps on behalf of the respondent. They need to comprehend the question, retrieve information from memory, eventually they need to judge

these information, map them onto the response options and perhaps edit the response due to sensitivity or social desirability. Respondents can take these steps in order or jump back and forth, e.g. if they need to form an opinion on the spot. But in all of these steps, item nonresponse can be introduced (de Leeuw, Hox and Huisman, 2003).

The two types of item nonresponse might be related to different steps of the survey response process. Refusals are more likely to stem from the editing step when respondents do not want to answer although they could. They might feel an item is too sensitive or do not trust the interviewer with this information. DK is more likely to come about when the respondent cannot answer in the first place. Either the respondent cannot answer because they do not know about the content of the question or are unable to remember an event (Beatty and Herrmann, 2002; Turner, Sturgis and Martin, 2015). Or they do not perceive it worth the effort to come up with an answer and satisfice (Krosnick, 1991). But Shoemaker, Eichholz and Skewes (2002) have shown that higher mental effort is related to more refusals as well. And conversely, it is plausible that DK is used as a more polite way to refuse. I, therefore, do not distinguish between them in the following sections even though they often show different strengths of predictors (Silber et al., 2021).

Ability has been the explanation for the differences often found in item nonresponse by education, age, physical and mental health (Colsher and Wallace, 1989; Pickery and Loosveldt, 1998; de Leeuw, Hox and Huisman, 2003; Messer, Edwards and Dillman, 2012; Silber et al., 2021). Women and ethnic minority members are often found to have more item nonresponse (Kupek, 1998; Pickery and Loosveldt, 1998). For the latter, on average lower liter-

acy and lower command of the majority language are hypothesised to cause more item nonresponse (Kupek, 1998). Meitinger and Johnson (2020) conclude that item nonresponse reflects broader social inequalities in abilities and access to information. The ability hypothesis is also directly supported by correlations between item nonresponse and measures of intelligence (Hedengren and Stratmann, 2012).

More difficult or unclear questions typically have higher levels of item nonresponse (Holbrook, Cho and Johnson, 2006; Messer, Edwards and Dillman, 2012; Holbrook, Johnson et al., 2016; Olson, Smyth and Ganshert, 2019). Since demographic questions can often be recalled very simply by respondents, they often show less item nonresponse than attitudinal and behavioural questions where the answers often do not exist before before getting asked the question (Olson, Smyth and Ganshert, 2019; Silber et al., 2021).

Even when able to perform the task at hand, respondents not necessarily want to. They need to perceive the effort as worth it. Differences in *motivation* are an explanation for less item nonresponse of respondents that are more interested in the topic of the survey (Koch and Blohm, 2009; Silber et al., 2021). Item nonresponse is related to measures of conscientiousness as well (Hedengren and Stratmann, 2012).

Motivation can also affect the editing process. For example, respondents (and interviewers) might deliberately not answer screening and filtering questions to reduce survey burden (Tourangeau, Kreuter and Eckman, 2015). When editing an answer, concerns of privacy<sup>1</sup> come into play. Am I going

---

<sup>1</sup>I use this label to subsume general concerns about privacy towards the survey (e.g. about data processing, anonymity) and not wanting to answer single sensitive questions.

to be judged by the interviewer when answering honestly? Are my data really secure and confidential? This is a particular problem for questions perceived as more intrusive (Tourangeau and Yan, 2007). Most prominently, this affects questions on income (Yan, Curtin and Jans, 2010) and sexual behaviour (Kupek, 1998), which often show particularly high levels of item nonresponse. High item nonresponse is often used as an indicator for question sensitivity (Tourangeau and Yan, 2007). Reluctancy and overall skepticism towards surveys and science, privacy concerns and generalised mistrust are traits that result in more nonresponse (Silber et al., 2021).

I have thus identified four constructs related to the probability of item nonresponse under which most research can be subsumed: cognitive ability, task difficulty, motivation and privacy concerns. However, as Krosnick (1991) already hypothesised, these concepts are interrelated in their effect on item nonresponse. Very easy questions can be answered by less able respondents and very hard questions might even cause the most able to struggle. The resulting fraction of difficulty by ability represents the relative mental effort to answer a question. To what point the relative mental effort is worth it depends on the respondents' motivation. A highly motivated respondent even answers the difficult questions. The probability of nonresponse decreases with motivation. Krosnick (1991, p. 225) has summarised these relations into a formula of the probability of satisficing. Satisficing refers to respondents lowering their standards in answering a question and only giving a satisfactory instead of an optimal answer to cut down on cognitive effort. Item nonresponse is such a satisficing strategy. There are situations

where DKs are genuine though (Turner, Sturgis and Martin, 2015). This is nonetheless a product of the respondents' ability and thereby included in the formula.

Additionally, higher privacy concerns lead to more item nonresponse. Since this relates to another step in the survey response process, namely editing rather than comprehension, retrieval or judging, I postulate it to be independent from the other concepts. I therefore suggest extending the formula by Krosnick as follows:

$$P_{INR} = \max \left( \frac{difficulty}{ability * motivation}, privacy \right)$$

*The probability of a nonsubstantive answer by the respondent is given by the task difficulty divided by the product of ability and motivation or the respondents' privacy concerns, whichever is higher.*

This formula specifically addresses item nonresponse for a single item but its meaning is adaptable to every level of a survey.

### 3 How to reduce Item Nonresponse

Based on these four concepts, one can theorise how to decrease item nonresponse. Some of the following possibilities may seem obvious and are established standards not only because of their potential relationship to item nonresponse but to ensure quality of substantial answers as well. Others might reduce item nonresponse but they could have negative consequences for other parts of *total survey error*, the combined effect of all error sources

in a survey (Groves and Lyberg, 2010). They require a trade-off before implementation.

I have structured this review of approaches to reduce item nonresponse by the respective concepts they target.

### 3.1 Task Difficulty

Task difficulty is predominantly influenced by design choices regarding single questions and their type of expected response. For the design of questions, much typical advice from question design aims to reduce task difficulty: asking short questions with simple structures, avoiding ambiguity and fuzzy concepts and not asking respondents to do multiple things at once. For a more comprehensive review on advice on question and questionnaire design, see for example Smyth (2016). But the type of question is also associated with task difficulty: open-ended questions, being allowed to select multiple options, more response categories and being asked to order categories are associated with higher item nonresponse (Schuman and Presser, 1979b; Holbrook, Cho and Johnson, 2006; Holbrook, Johnson et al., 2016; Olson, Smyth and Ganshert, 2019; Silber et al., 2021). Additional measures to decrease difficulty can be visual helpers like images or showcards. With showcards, respondents do not need to remember all response categories while thinking about the question. But there is not much empirical work on the effect of showcards on item nonresponse. Holbrook, Johnson et al. (2016) found that questions using showcards had more item nonresponse although this is likely because they are only used with more difficult questions. In the context of the ESS,



showcards do not seem to affect the distribution of substantive responses in experiments (Jäckle, Roberts and Lynn, 2010) but item nonresponse was not investigated.

The question level is very well understood and constitutes the largest part of the variance in item nonresponse. This highlights the importance of the single question for overall data quality (Olson, Smyth and Ganshert, 2019).

But not only single questions but the whole questionnaire needs to be designed as simply as possible. Questionnaires with changes in response categories, routing and filtering are associated with higher nonresponse (Messer, Edwards and Dillman, 2012). Grouping questions by topic could reduce mental effort and therefore item nonresponse but it also increases the likelihood of non-differentiation between the items (Krosnick, 1991). Offering DK and refusal explicitly increases their use because respondents are more aware of the “easy way out” (Schuman and Presser, 1979a; Beatty and Herrmann, 2002).

Task difficulty can be reduced for language minorities by providing the questionnaire in more languages. However, this is expensive and can reduce comparability of cases. For a review on comparability in cross-cultural surveys, see e.g. Behr and Shishido (2016).

## **3.2 Ability**

To make sure that the respondent is able to answer the questions, interviews should be conducted in situations where understanding and concentration is as good as possible. Other people present during the interview, for example,

could be a source of distraction. Empirical investigations have not been conclusive though (Kupek, 1998; Tu and Liao, 2007; Silber et al., 2021). With longer interviews, respondents might get tired (Holbrook, Johnson et al., 2016; Olson, Smyth and Ganshert, 2019).

### 3.3 Motivation

Motivation could decrease over the course of the interview as well. Different page layouts and progress bars might help with that in web surveys but evidence is mixed (Peytchev et al., 2006; Yan, Conrad et al., 2011; Sarraf and Tukibayeva, 2014). The literature on cooperation enhancement, e.g. using incentives, has focussed on unit nonresponse. Some of these ideas might be applicable to increase respondent motivation within the interview as well. After all, unit and item nonresponse are linked: respondents that initially refused to participate have higher levels of item nonresponse (Yan and Curtin, 2010; Fricker and Tourangeau, 2010).

### 3.4 Privacy Concerns

To tackle reluctance to answer due to privacy concerns, survey researchers need to make sure that the respondents' data are secure and anonymous and communicate that. Explaining why the data are collected, how they are processed and what is done to protect their privacy is also ethically advisable and often legally required.

In face-to-face surveys, establishing trust depends on the respondent-interviewer dynamics. Based on the hypothesis that respondents might per-

ceive interviewers similar to themselves as more trustworthy, scholars have assessed the impact of matching their characteristics. Vercruyssen, Wuyts and Loosveldt (2017) find less item nonresponse when matching age of interviewers and respondents and matching gender reduces item nonresponse for males and increases it for females. Matching education has no effect (Vercruyssen, Wuyts and Loosveldt, 2017; Silber et al., 2021). Tu and Liao (2007) on the other hand find age and education matching are potent predictors of item nonresponse.

## 4 Strategies to be Tested

In this review, I have stressed the role of the respondent and how design choices can increase the probability of a substantial answer. But in face-to-face surveys, the interviewer can be a source of item nonresponse as well. This is left for further research. I proceed with testing a few proposed strategies for item nonresponse reduction based on this review.

Over time, respondents might lose the ability to concentrate and motivation. I therefore expect more item nonresponse, the longer the interview takes. If this is true, making questionnaires only as long as necessary is a strategy to reduce item nonresponse and maintain high data quality. However, the duration of the interview might also be dependent on the respondents' ability. Overall ability of the respondent needs to be controlled for in the model.

Interferences and other people being present during the interview are expected to increase item nonresponse because they can be a source of dis-

traction and respondents might not want to answer some questions in front of people they know. If true, interviews should be conducted without others present if possible.

Showcards are intended to make the process of answering less burdensome because respondents do not need to remember the response options. If this has an effect on item nonresponse, showcards should be provided and their use promoted to ensure a high data quality. In this operationalisation, the interviewer assesses afterwards to what extent the respondent used the showcards. Respondent ability might influence how well they understand the procedures, ability needs to be controlled for again. The interviewer assessment is likely influenced by the overall impression the interviewer has from the interview and is therefore endogenous. The overall impression of the interviewer needs to be controlled for as well.

Language difficulties are likely to inhibit the response process. For respondents that primarily speak a different language at home than the interview is conducted in, the task might be more difficult. To ensure data quality from these respondents, translating the questionnaire is an option (although costly and comes with problems of comparability). Respondents for whom this is true are likely migrant groups that have lower levels of education than the native population. Therefore, ability needs to be controlled for.

Previous research suggests that respondents might be more willing to answer to socially similar interviewers. I test this proposition for gender and age. If true and the sampling frame contains some demographic information, survey agencies could assign interviewers based on these demographics. Whether the gender of respondent and interviewer matches can be seen as

random while the probability of the respondent being older or younger than the interviewer is dependent on the respondents' age, which is correlated with ability. Therefore, respondent age needs to be controlled for <sup>2</sup>.

Table 1 summarises the strategies to be tested. Necessary controls are respondents' ability and specifically age and the interviewers' overall assessment of the interview. Based on the theory outlined above, I do not see that inclusion of any of these control variables or the other strategies are likely to distort the effect of another variable of interest. Therefore, I am going to test all effects in a single model.

Table 1: Reduction Strategies to be Tested

Strategy	Mechanism	Expectation
Duration of the Interview	Ability, Motivation	positive
Interference of the Interview	Ability, Privacy	positive
Use of Showcards	Difficulty	negative
Interview not primary Language	Difficulty	positive
Gender Matching	Privacy	negative
Interviewer more than 10 years older	Privacy	positive
Respondent more than 10 years older	Privacy	positive

## 5 Data and Methods

### 5.1 Data

Following this review and synthesis, I want to test the effectiveness of some potential strategies to reduce item nonresponse. I use the *European Social Survey (ESS)* Round 9 collected in 2018 (ESS ERIC, 2019). The ESS is a biannual face-to-face survey on attitudes and beliefs towards social and

---

<sup>2</sup>Interviewer age is accounted for in the fixed effects.

political topics in Europe. The data release 3.1 includes data from 49,519 respondents from 29 countries.

## **5.2 Dependent Variables**

The three dependent variables are the sum of DK, the sum of refusals and the total sum of item nonresponse for every respondent. Item nonresponse is counted only in variables that are presented to all respondents, i.e. that are not affected by filtering questions. DK and refusal are not presented as response options in the ESS but are explicitly recorded as such. Interviewers are instructed to accept them without probe.

## **5.3 Independent Variables**

The independent variables of interest are discussed in the previous section.

To control for ability and the interviewers' assessment of the interview, I include education (operationalised by the ISCED scale), age and squared age of the respondent as well as the interviewers' assessment of how well the respondent understood the questions, to what extent the respondent answered to the best of their ability and how often they asked for clarifications.

More information on all variables including summary statistics can be found in the [supplemental material](#).

## 5.4 Model

I analyse the data using a negative binomial regression with interviewer fixed effects (Allison and Waterman, 2002)<sup>3</sup>. Standard errors are clustered by interviewer. Since the population of interest in this case are interviews and not societies, no weighting is applied. Missing data are deleted listwise.

The analyses are carried out in *R* (R Core Team, 2021) using the *Tidyverse* (Wickham et al., 2019) for data handling and graphics and the *fixest* package (Bergé, 2018) to estimate the regressions. The code is available in the [supplemental material](#) as well.

## 6 Results

Figure 1 shows the incidence rate ratios (exponentiated coefficients) of the tested strategies. The exact coefficients of the tested strategies and of the control variables as well as standard errors and coefficients of model fit can be found in the table in the appendix.

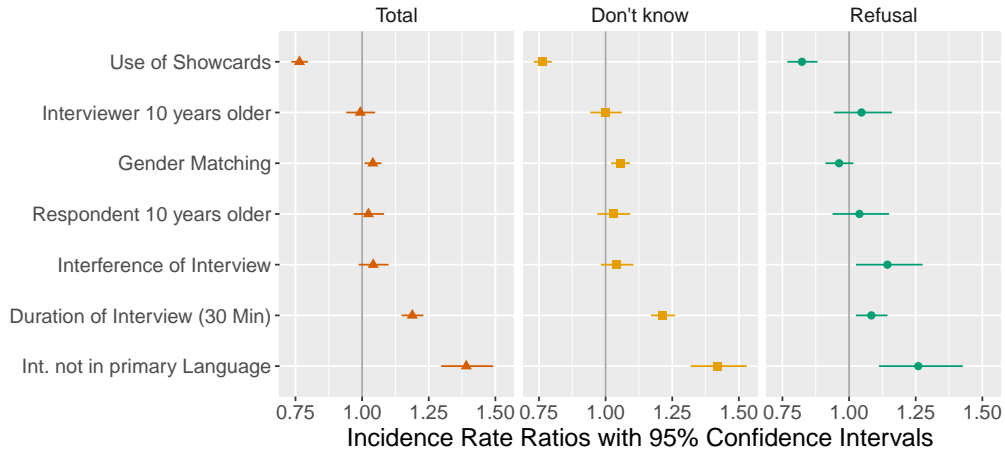
The most promising ways to reduce item nonresponse seem to be boosting the use of showcards, shorter interviews and translating questionnaires.

With more showcard use as indicated by the interviewer (respondent used no showcards vs. used some showcards vs. used all showcards), the amount of DK reduces by about a quarter and the amount of refusal by about 18% on average. With every additional thirty minutes interview duration, the number of DKs rises by about 21% and the number of refusals rises by about

---

<sup>3</sup>Count data models with fixed effects are quite debated (Wooldridge, 1999; Allison and Waterman, 2002; Cameron and Trivedi, 2013). For a justification of model choice, see the supplementary material.

Figure 1: Coefficients of the tested Strategies



8% on average. And compared to interviews conducted in the language the respondent primarily speaks at home, interviews conducted in a language different from the respondents' primary language at home show on average 42% more DKs and 26% more refusals.

As a general observation for all variables, the effects on the total number of item nonresponse closely mirror the effect on DK. This is not surprising since there are many more DKs than refusals. The effects on the number of refusals is typically weaker than the effect on DK but still present. This supports the idea that refusals and DKs are not perfectly separate in their meaning but not identical as well. For the variables presented so far, stronger effects on DK make substantial sense as well since they are all based on respondents' ability or difficulty of the task.

Matching respondents' gender has a small positive effect on the number of DKs. This is contrary to expectations, which suggested that matching socio-demographics of interviewers and respondents leads to a more trusting



interview situation and reduces item nonresponse. The effect on refusals is not significant but should be pronounced since this strategy is partly based on the privacy mechanism. Matching by age has no significant effect. Matching respondents and interviewers therefore seems not to be a promising strategy to reduce item nonresponse.

Other people present during the interview raises the amount of refusals by 14% in line with the reasoning that respondents do not want to answer some questions in presence of others they know. The effect on DK is not significant. Other people present might therefore influence item nonresponse more via privacy than distraction. However, due to the relatively small number of refusals, the effect on the total item nonresponse is not significant.

## 7 Discussion

I have reviewed the literature on item nonresponse and extended the cognitive *satisficing* model (Krosnick, 1991) with concerns about privacy to encompass all aspects that can interfere with the response process in survey interviews. Based on this new model, I have reviewed possibilities to reduce item nonresponse in face-to-face surveys.

In an empirical analysis using data from the *European Social Survey Round 9*, I found that boosting the respondents' use of showcards, shortening interview durations and conducting the interview in the respondents' primary language might be promising ways to reduce item nonresponse in face-to-face surveys. These strategies reduce the cognitive effort on behalf of respondents. Other people present during the interview is moderately asso-

ciated with more refusals as well. Respondents are probably not willing to disclose private information in front of people they know. However, my hypothesis that respondents might trust interviewers more and share more private information if interviewer and respondent are socially similar has received no support: matching socio-demographic characteristics of interviewers and respondents seems not to be a worthwhile strategy.

I want to stress though that I do not claim causal effects. Although I carefully selected control variables, I cannot rule out violations of the *conditional independence assumption* which is necessary to identify causal effects with regression analysis. Most variables are influenced by respondents' ability (e.g. to understand survey procedures) as is item nonresponse. Respondents' ability is notoriously hard to measure in surveys and proxies like education and age that I have used as controls are not perfect.

A second threat to the results is the endogeneity of some variables of interest, in particular showcard use and others present during the interviews. They are measured in the interviewer questionnaire after the interview and are likely biased by the interviewers' overall assessment of the interview, including the amount of item nonresponse. I tried to control for that using other variables from the interviewer questionnaire but again, I cannot verify the conditional independence assumption to be true.

In summary, this review and analysis can only serve as a *current best guess* in how to think about item nonresponse and its reduction. Future research might be able to overcome some of the shortcomings of this and other previous studies.

## Acknowledgements

I want to thank Achim Koch, Johanna Mehltritt and especially Tobias Gummer for helpful comments in multiple stages of this project. I want to thank Marcel Kappes and Thomas Gautschi for methodological discussions.

This work was supported by the University of Mannheim's Graduate School of Economic and Social Sciences.

## References

- Allison, Paul D. and Richard P. Waterman (2002). 'Fixed-Effects Negative Binomial Regression Models'. In: *Sociological Methodology* 32.1, pp. 247–265. DOI: [10.1111/1467-9531.00117](https://doi.org/10.1111/1467-9531.00117).
- Beatty, Paul and Douglas Herrmann (2002). 'To Answer or Not to Answer: Decision Processes Related to Survey Item Nonresponse'. In: *Survey Non-response*. Ed. by Robert M. Groves. New York, NY: Wiley.
- Behr, Dorothe and Kuniaki Shishido (2016). 'The Translation of Measurement Instruments for Cross-Cultural Surveys'. In: Wolf, Christof et al. *The SAGE Handbook of Survey Methodology*. London: SAGE Publications Ltd, pp. 269–287. ISBN: 978-1-4739-5789-3. DOI: [10.4135/9781473957893.n19](https://doi.org/10.4135/9781473957893.n19).
- Bergé, Laurent (2018). 'Efficient Estimation of Maximum Likelihood Models with Multiple Fixed-Effects: The R Package FENmlm'. In: *CREA Discussion Papers* 13.
- Cameron, A. Colin and Pravin Trivedi (2013). *Regression Analysis of Count Data*. 2nd ed. Cambridge: Cambridge University Press. ISBN: 978-1-139-01356-7. DOI: [10.1017/CB09781139013567](https://doi.org/10.1017/CB09781139013567).
- Colsher, Patricia L. and Robert B. Wallace (1989). 'Data Quality and Age: Health and Psychobehavioral Correlates of Item Nonresponse and Inconsistent Responses'. In: *Journal of Gerontology* 44.2, pp. 45–52. DOI: [10.1093/geronj/44.2.P45](https://doi.org/10.1093/geronj/44.2.P45).
- De Leeuw, Edith D., Joop Hox and Mark Huisman (2003). 'Prevention and Treatment of Item Nonresponse'. In: *Journal of Official Statistics* 19.3, pp. 153–176.
- ESS ERIC (2019). 'European Social Survey (ESS), Round 9 - 2018'. In: DOI: [10.21338/NSD-ESS9-2018](https://doi.org/10.21338/NSD-ESS9-2018).

- Fricker, Scott and Roger Tourangeau (2010). ‘Examining the Relationship Between Nonresponse Propensity and Data Quality in Two National Household Surveys’. In: *Public Opinion Quarterly* 74.5, pp. 934–955. DOI: [10.1093/poq/nfq064](https://doi.org/10.1093/poq/nfq064).
- Groves, Robert M. and Lars E. Lyberg (2010). ‘Total Survey Error: Past, Present, and Future’. In: *Public Opinion Quarterly* 74.5, pp. 849–879. DOI: [10.1093/poq/nfq065](https://doi.org/10.1093/poq/nfq065).
- Hedengren, David and Thomas Stratmann (2012). ‘The Dog That Didn’t Bark: What Item Nonresponse Shows about Cognitive and Non-Cognitive Ability’. In: *SSRN Electronic Journal*. DOI: [10.2139/ssrn.2194373](https://doi.org/10.2139/ssrn.2194373).
- Holbrook, Allyson L., Young I. Cho and Timothy P. Johnson (2006). ‘The Impact of Question and Respondent Characteristics on Comprehension and Mapping Difficulties’. In: *Public Opinion Quarterly* 70.4, pp. 565–595. DOI: [10.1093/poq/nf1027](https://doi.org/10.1093/poq/nf1027).
- Holbrook, Allyson L., Timothy P. Johnson et al. (2016). ‘Do Interviewer Errors Help Explain the Impact of Question Characteristics on Respondent Difficulties?’ In: *Survey Practice* 9.2, pp. 1–11. DOI: [10.29115/SP-2016-0009](https://doi.org/10.29115/SP-2016-0009).
- Jäckle, Annette, Caroline Roberts and Peter Lynn (2010). ‘Assessing the Effect of Data Collection Mode on Measurement’. In: *International Statistical Review* 78.1, pp. 3–20. ISSN: 03067734, 17515823. DOI: [10.1111/j.1751-5823.2010.00102.x](https://doi.org/10.1111/j.1751-5823.2010.00102.x).
- Koch, Achim and Michael Blohm (2009). ‘Item Nonresponse in the European Social Survey’. In: *ASK: Research & Methods* 18.1, pp. 45–65.
- Krosnick, Jon A. (1991). ‘Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys’. In: *Applied Cognitive Psychology* 5.3, pp. 213–236. DOI: [10.1002/acp.2350050305](https://doi.org/10.1002/acp.2350050305).
- Kupek, Emil (1998). ‘Determinants of Item Nonresponse in a Large National Sex Survey’. In: *Archives of Sexual Behavior* 27.6, pp. 581–594.
- Meitinger, Katharina M. and Timothy P. Johnson (2020). ‘Power, Culture and Item Nonresponse in Social Surveys’. In: *Understanding Survey Methodology*. Ed. by Philip S. Brenner. Vol. 4. Frontiers in Sociology and Social Research. Cham: Springer, pp. 169–191. DOI: [10.1007/978-3-030-47256-6\\_8](https://doi.org/10.1007/978-3-030-47256-6_8).
- Messer, Benjamin L., Michelle L. Edwards and Don A. Dillman (2012). ‘Determinants of Item Nonresponse to Web and Mail Respondents in Three Address-Based Mixed-Mode Surveys of the General Public’. In: *Survey Practice* 5.2, pp. 1–9. DOI: [10.29115/SP-2012-0012](https://doi.org/10.29115/SP-2012-0012).
- Olson, Kristen, Jolene D. Smyth and Amanda Ganshert (2019). ‘The Effects of Respondent and Question Characteristics on Respondent Answering

- Behaviors in Telephone Interviews’. In: *Journal of Survey Statistics and Methodology* 7.2, pp. 275–308. DOI: [10.1093/jssam/smy006](https://doi.org/10.1093/jssam/smy006).
- Peytchev, Andy et al. (2006). ‘Web Survey Design: Paging versus Scrolling’. In: *Public Opinion Quarterly* 70.4, pp. 596–607. DOI: [10.1093/poq/nfl028](https://doi.org/10.1093/poq/nfl028).
- Pickery, Jan and Geert Loosveldt (1998). ‘The Impact of Respondent and Interviewer Characteristics on the Number of “No Opinion” Answers’. In: *Quality & Quantity* 32, pp. 31–45.
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Version 4.0.5. Wien: R Foundation for Statistical Computing. URL: <https://www.R-project.org/>.
- Sarraf, Shimon and Malika Tukibayeva (2014). ‘Survey Page Length and Progress Indicators: What Are Their Relationships to Item Nonresponse?’ In: *New Directions for Institutional Research* 2014.161, pp. 83–97. DOI: [10.1002/ir.20069](https://doi.org/10.1002/ir.20069).
- Schuman, Howard and Stanley Presser (1979a). ‘The Assessment of “No Opinion” in Attitude Surveys’. In: *Sociological Methodology* 10, pp. 241–75. DOI: [10.2307/270773](https://doi.org/10.2307/270773). JSTOR: [270773](https://www.jstor.org/stable/270773).
- (1979b). ‘The Open and Closed Question’. In: *American Sociological Review* 44.5, pp. 692–712. DOI: [10.2307/2094521](https://doi.org/10.2307/2094521). JSTOR: [2094521](https://www.jstor.org/stable/2094521).
- Shoemaker, Pamela J., Martin Eichholz and Elizabeth A. Skewes (2002). ‘Item Nonresponse: Distinguishing between Don’t Know and Refuse’. In: *International Journal of Public Opinion Research* 14.2, pp. 193–201. DOI: [10.1093/ijpor/14.2.193](https://doi.org/10.1093/ijpor/14.2.193).
- Silber, Henning et al. (2021). ‘The Effects of Question, Respondent and Interviewer Characteristics on Two Types of Item Nonresponse’. In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, pp. 1–18. DOI: [10.1111/rssa.12703](https://doi.org/10.1111/rssa.12703).
- Smyth, Jolene D. (2016). ‘Designing Questions and Questionnaires’. In: Wolf, Christof et al. *The SAGE Handbook of Survey Methodology*. London: SAGE Publications Ltd, pp. 218–235. ISBN: 978-1-4739-5789-3. DOI: [10.4135/9781473957893.n16](https://doi.org/10.4135/9781473957893.n16).
- Tourangeau, Roger, Frauke Kreuter and Stephanie Eckman (2015). ‘Motivated Misreporting: Shaping Answers to Reduce Survey Burden’. In: *Survey Measurement Techniques, Data Quality and Sources of Error*. Ed. by Uwe Engel. Frankfurt: Campus Verlag.
- Tourangeau, Roger, Lance J. Rips and Kenneth Rasinski (2000). *The Psychology of Survey Response*. 1st ed. Cambridge University Press. ISBN: 978-0-521-57246-0. DOI: [10.1017/CB09780511819322](https://doi.org/10.1017/CB09780511819322).

- Tourangeau, Roger and Ting Yan (2007). ‘Sensitive Questions in Surveys.’ In: *Psychological Bulletin* 133.5, pp. 859–883. ISSN: 1939-1455, 0033-2909. DOI: [10.1037/0033-2909.133.5.859](https://doi.org/10.1037/0033-2909.133.5.859).
- Tu, Su-Hao and Pei-Shan Liao (2007). ‘Social Distance, Respondent Cooperation and Item Nonresponse in Sex Survey’. In: *Quality & Quantity* 41.2, pp. 177–199. DOI: [10.1007/s11135-007-9088-0](https://doi.org/10.1007/s11135-007-9088-0).
- Turner, Gosia, Patrick Sturgis and David Martin (2015). ‘Can Response Latencies Be Used to Detect Survey Satisficing on Cognitively Demanding Questions?’ In: *Journal of Survey Statistics and Methodology* 3.1, pp. 89–108. DOI: [10.1093/jssam/smu022](https://doi.org/10.1093/jssam/smu022).
- Vercruyssen, Anina, Celine Wuyts and Geert Loosveldt (2017). ‘The Effect of Sociodemographic (Mis)Match between Interviewers and Respondents on Unit and Item Nonresponse in Belgium’. In: *Social Science Research* 67, pp. 229–238. DOI: [10.1016/j.ssresearch.2017.02.007](https://doi.org/10.1016/j.ssresearch.2017.02.007).
- Wickham, Hadley et al. (2019). ‘Welcome to the Tidyverse’. In: *Journal of Open Source Software* 4.43, p. 1686. DOI: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686).
- Wooldridge, Jeffrey M. (1999). ‘Distribution-Free Estimation of Some Non-linear Panel Data Models’. In: *Journal of Econometrics* 90.1, pp. 77–97. DOI: [10.1016/S0304-4076\(98\)00033-5](https://doi.org/10.1016/S0304-4076(98)00033-5).
- Yan, Ting, Frederick G. Conrad et al. (2011). ‘Should I Stay or Should I Go: The Effects of Progress Feedback, Promised Task Duration, and Length of Questionnaire on Completing Web Surveys’. In: *International Journal of Public Opinion Research* 23.2, pp. 131–147. DOI: [10.1093/ijpor/edq046](https://doi.org/10.1093/ijpor/edq046).
- Yan, Ting and Richard Curtin (2010). ‘The Relation Between Unit Nonresponse and Item Nonresponse: A Response Continuum Perspective’. In: *International Journal of Public Opinion Research* 22.4, pp. 535–551. DOI: [10.1093/ijpor/edq037](https://doi.org/10.1093/ijpor/edq037).
- Yan, Ting, Richard Curtin and Matthew Jans (2010). ‘Trends in Income Nonresponse Over Two Decades’. In: *Journal of Official Statistics* 26.1, pp. 145–164.

Table 2: Regression Results

Dependent Variables: Model:	Don't know (1)	Refusal (2)	Total (3)
<i>Variables</i>			
Duration of Interview (30 Min)	0.194*** (0.019)	0.080** (0.028)	0.173*** (0.018)
Interference of Interview	0.041 (0.030)	0.135* (0.056)	0.041 (0.028)
Use of Showcards	-0.270*** (0.023)	-0.195*** (0.035)	-0.268*** (0.021)
Int. not in primary Language	0.351*** (0.038)	0.231*** (0.064)	0.330*** (0.036)
Gender Matching	0.054** (0.017)	-0.038 (0.028)	0.040* (0.016)
Respondent 10 years older	0.028 (0.031)	0.038 (0.052)	0.024 (0.029)
Interviewer 10 years older	$9.65 \times 10^{-5}$ (0.030)	0.045 (0.053)	-0.007 (0.028)
Education (ISCED)	-0.128*** (0.005)	0.065*** (0.009)	-0.097*** (0.005)
Age	-0.050*** (0.003)	0.010* (0.005)	-0.042*** (0.002)
Age squared	0.0005*** ( $2.43 \times 10^{-5}$ )	$-6.54 \times 10^{-5}$ ( $4.59 \times 10^{-5}$ )	0.0004*** ( $2.3 \times 10^{-5}$ )
Understood Questions	-0.337*** (0.017)	-0.046 (0.028)	-0.287*** (0.016)
Answered to best Ability	-0.035* (0.016)	-0.177*** (0.024)	-0.062*** (0.015)
Amount of Clarifications	0.254*** (0.011)	0.598*** (0.021)	0.328*** (0.011)
<i>Fixed-effects</i>			
Interviewer	Yes	Yes	Yes
<i>Fit statistics</i>			
Observations	41,512	34,632	41,795
Pseudo R <sup>2</sup>	0.11966	0.18119	0.12252
Within Pseudo R <sup>2</sup>	0.05809	0.08775	0.05937
BIC	189,641.8	86,863.5	204,151.4
Over-dispersion	0.89111	0.75064	1.0251

Clustered (Interviewer) standard-errors in parentheses

Signif. Codes: \*\*\*: 0.001, \*\*: 0.01, \*: 0.05