Article:
Jabkowski, P., Cichocki, P. and Piekut, A. (2026) Switching to self-completion protocols impacts item nonresponse patterns. Lessons from the European Social Survey rounds 9, 10 and 11. International Journal of Social Research Methodology. ISSN: 1364-5579

https://doi.org/10.1080/13645579.2026.2680975

Switching to Self-Completion Protocols Impacts Item Nonresponse Patterns. Lessons from the European Social Survey rounds 9, 10 and 11

Version accepted for publication, April 2026.

Please cite as:

Jabkowski Piotr, Cichocki Piotr, Piekut Aneta (2026). Switching to Self-Completion Protocols Impacts Item Nonresponse Patterns. Lessons from the European Social Survey rounds 9, 10 and 11. *International Journal of Social Research Methodology*. https://doi.org/10.1080/13645579.2026.2680975.

Abstract:
Surveys are increasingly pivoting from interviewer-assisted to self-completion protocols, yet the consequences for data quality remain unclear. Guided by the survey satisficing theory, this is the first study that inspects the COVID–19 mode change in the European Social Survey Round 10, where eight countries switched to self-completion. At the same time, 19 countries maintained interviewer-assisted modes, with all 27 entirely implementing them in Rounds 9 and 11. We analyse item nonresponse on three items with complex response formats and three standard opinion questions with historically high missingness. We use hierarchical Bayesian beta regression to estimate mode effects. Our findings bring significant considerations for survey studies transitioning from interviewer-assisted to online self-completion mode. Switching to self-completion resulted in a marked increase in nonresponse for the cognitively demanding items, but a decline for the standard opinion items; both effects disappeared when interviewer-assisted modes were used. Our study supports the survey satisficing interpretation: self-completion fosters strong satisficing behaviour on complex survey items, yet it results in weak satisficing on opinion items.

Keywords:
item nonresponse; self-administrated push-to-web surveys; interviewer-assisted face-to-face surveys; mode-change effect

Piotr Jabkowski*, [1], Piotr Cichocki [2], Aneta Piekut [3]

1) Corresponding author: piotr.jabkowski@amu.edu.pl; Adam Mickiewicz University, Poznan, Poland

2) piotr.cichocki@amu.edu.pl; Adam Mickiewicz University, Poznan, Poland

3) a.piekut@sheffield.ac.uk; Sheffield Methods Institute, University of Sheffield, UK

Summary of the main argument and key methodological contribution

The article draws on the natural experiment created when eight of the twenty-seven European Social Survey (ESS) countries switched from face-to-face to self-completion data collection in Round 10 (2020) and then reverted to face-to-face data collection in Round 11 (2022). It demonstrates that the survey mode affects item nonresponse in opposite directions for different types of questions. Switching to self-completion sharply increases item nonresponse on cognitively demanding household roster items that use unfamiliar response formats, yet simultaneously decreases item nonresponse on simple political opinion items that use the ESS's familiar 0–10 rating scale. Reverting to face-to-face in Round 11 returns the item nonresponse rates to the baseline of Round 9. This "spike-and-dip" pattern is explained by survey-satisficing theory. Without an interviewer, respondents engage in strong satisficing (skipping) when faced with complex tasks and weak satisficing (quickly selecting any scale point) when asked routine opinion questions. When the seven countries returned to interviewer administration, both effects disappeared, reinforcing the causal inference that mode, rather than time or country differences, drives the observed pattern.

Methodologically, the study presents a rigorously specified hierarchical Bayesian beta-regression framework that models item nonresponse rates (bounded by 0 and 1) while capturing cross-national heterogeneity via random intercepts and random, round-specific slopes. A novel position-weighted indicator of item nonresponse for household roster items ensures that households of varying sizes contribute proportionately. By interacting survey round with a dichotomous "mode-switch" country grouping, the model isolates mode effects net of pre-existing country differences and provides full posterior uncertainty intervals for those effects. Thus, the design offers a template for evaluating mode changes in other longitudinal or cross-national studies. The spike-and-dip pattern uncovered by this study adds new empirical leverage to debates on mode-induced satisficing.

Switching to Self-Completion Protocols Impacts Item Nonresponse Patterns. Lessons from the European Social Survey rounds 9, 10 and 11

Introduction

Understanding the effects of mode change on survey quality and measurement equivalence constitutes a major current concern in survey methodology. The COVID-19 disruption accelerated a pre-existing trend in cross-national survey projects, i.e., shifting away from interviewer-assisted modes, with self-completion becoming the new normal in social surveys (Coffey et al., 2024; Maslovskaya et al., 2022; Nursel & Stähli, 2025). The European Social Survey (ESS), a methodology-driven comparative project conducted biennially since 2002, provides a prime example of the survey mode switch (Fitzgerald & Aizpurua, 2024). Interviewers conducted all ESS surveys as in-person interviews in Rounds 1 through 9; however, the pandemic-disrupted Round 10 (2020) required a mode change in some countries from interviewer-assisted to self-completion protocols. Subsequent Round 11 (2023) reverted to implementing interviewer-assisted modes in all participating countries. However, the ESS is switching to self-completion only, starting from round 13 (2027), with the transitional round 12 (2025) embracing a costly double-mode execution of two parallel surveys in each country. In selected countries, such as the Czech Republic, Finland, France, Hungary, and the United Kingdom, these double-mode parallel experiments have already been completed alongside the ESS Rounds 10 and 11. Furthermore, Ireland and Bulgaria implemented a pilot study of the self-completion-only approach in preparation for the ESS Round 12 (e.g., Berteen & Phillips, 2021; ESS, 2025; Lugtig, 2024; Singh, 2023).

The protocol changes in ESS10 provide original empirical foundations for investigating the mode-switching effects, whose understanding would be crucial for ensuring measurement comparability in future ESS Rounds and other similar cross-national surveys. Twenty-seven countries participated in all three Rounds, 9, 10 and 11, with interviewer-assisted fieldwork execution uniformly implemented in all countries in ESS9 and ESS11. However, in ESS10, eight countries switched to self-completion, while the remaining nineteen countries continued with interviewer-assisted fieldwork execution. Hence, a quasi-experimental condition arose, allowing for an evaluation of the effect of switching the mode of survey execution. The switching countries were not chosen randomly; however, they do not seem to represent any obvious regional or socio-economic grouping in Europe, as they comprise Austria, Cyprus, Germany, Latvia, Spain, Poland, Serbia, and Sweden. Therefore, these eight mode-switchers can be compared against the group of nineteen non-switchers, while ESS9 and ESS11 provide an interviewer-assisted baseline for all twenty-seven countries.

This study advances methodological discussion on the dissimilarity of mode effects across the questionnaire content, specifically differential item nonresponse patterns. Our analysis focuses on the effects of switching to self-administered fieldwork execution on item nonresponse rates for two types of ESS questionnaire items. On the one hand, we focused on questions with complex and unfamiliar response options, relative to other parts of the ESS questionnaire. These questions, selected from the multistage household roster initiating Module F of the ESS questionnaire, also employed complicated instructions. Three items were scrutinised: gender, year of birth and relationship to the respondent of household members other than the respondent. They place demands on the respondent’s factual recall and present response options that require substantial cognitive effort to navigate. On the other hand, we selected standard opinion items that have historically been associated with the highest item nonresponse rates in the ESS: left-right self-placement, trust in the United Nations, and views on European Unification. All three opinion questions use an 11-point end-labelled scale, uniformly implemented across ESS instruments. They probe abstract political constructs, which may be challenging for some respondents to judge; however, any ESS respondent would be thoroughly familiar with their response format when they appear in the questionnaire sequence.

Drawing on the theory of survey satisficing (e.g., Krosnick et al., 1996; Mulligan et al., 2001; Roberts et al., 2019), we expect these two kinds of questions to be differentially affected by interviewer presence in terms of the respondent's propensity to opt for item nonresponse. For questions with complex response options, the rates of item nonresponse should rise in self-completion modes. Without interviewers offering explanations and enforcing compliance, low-motivation respondents are likely to be enticed by strong satisficing, skipping their way through without engaging with unfamiliar response options. Conversely, for opinion items with familiar response options, self-completion is likely to promote weak satisficing, resulting in lower item nonresponse rates. It seems easier and better fits the role of a good participant to just select any response option on a familiar scale than to skip the item. In interviewer-assisted interviews, weak satisficing should manifest disproportionately as item nonresponse on standard opinion items, i.e., either 'don't know' or refusal. The presence of an interviewer increases the interactional and social evaluative cost of providing a weakly substantially grounded response, making selecting the nonresponse option the most defensible option requiring the least effort. To evaluate our expectations, we fitted hierarchical Bayesian beta regression models to the item nonresponse rate outcomes. The baseline specification encompasses random intercepts for countries and random slopes for survey rounds, thereby encapsulating cross-national and cross-round heterogeneity in item nonresponse rates. In addition, a dichotomous country-group indicator distinguishes the eight ESS10 self-completion countries from the nineteen interviewer-assisted counterparts, and its interaction with the survey round directly tests the hypothesised mode-switch effect.

Theoretical frameworks

Answering survey questions constitutes a cognitive exercise whose difficulty depends on various factors, including respondents' knowledgeability (Alexander, 2017; Berinsky, 2004) and their levels of interest in and experience with the topic (Jabkowski & Piekut, 2024; Zhu, 1996). Respondents are known to engage in survey satisficing to minimise their cognitive burden (Alwin & Krosnick, 1991; Krosnick, 1991). Instead of providing thoughtful, substantive and valid answers, some participants employ optimisation strategies aimed at minimising their engagement with questions (Kraemer et al., 2023; Roberts et al., 2019). They prefer to complete the survey quickly, while remaining formally compliant with the instructions. Crucially, the presence of an interviewer, who provides support and encouragement to engage with the questions (West & Blom, 2017), directly affects respondents' motivation, compliance and cognitive effort. Therefore, the patterns of satisficing are likely to differ between self-administered and interviewer-assisted interviews (Holbrook et al., 2003; Ye et al., 2011).

The cognitive process underlying answer-formation involves multiple stages: 1) question assimilation and understanding, 2) information retrieval from memory, 3) estimation of the requested response, and finally, 4) fitting it into the given response options (Beatty & Herrmann, 2002). Skipping or rushing through some or all these stages indicates a satisficing approach on the part of the respondent, wherein the extent of survey satisficing depends on three interrelated factors: task difficulty, ability to answer the question, and the respondent's motivation to spend time and make effort to consider an answer possibly close to the true behaviour or attitude (Krosnick, 1991). Satisficing can be 'weak' when respondents apply all the cognitive stages but rush through them, or 'strong' if they skip some of them altogether. Item nonresponse is an example of a strong satisficing mechanism (Cornesse et al., 2020). Some respondents might genuinely not know how to answer, for example, if the survey asks about novel issues, unformed opinions or facts unavailable to them (Alwin & Krosnick, 1991). On the other hand, weak satisficing typically involves careless response styles (Blazek & Siegel, 2024; Stosic et al., 2024).

Low-motivation respondents have a higher propensity for speed-running self-administered surveys, resulting in random responses, a lack of substantive answers, or skipping questions by simply clicking the 'next button' (Cornesse & Blom, 2020). Speed-running seems much less likely in interviewer-assisted surveys, especially in face-to-face modes (Holbrook et al., 2003). The self-administering respondent can go through the questionnaire at their own pace and enjoy a higher sense of control over the response process (Bowling, 2005; West & Blom, 2017), which is beneficial for highly motivated participants as it allows for more time to engage in cognitive processing and careful reading (Ye et al., 2011). However, interviewer absence offers an easy path to weak satisficing for low-motivation respondents, who may readily translate their sense of control over the response process into a very high pace of answering without engaging with the questions (Heerwegh & Loosveldt, 2008; Jäckle & Lynn, 2008). This path would also be highly alluring to respondents with lower digital literacy, who might struggle with answering self-administered web-based surveys (Oliveri et al., 2021).

When it comes to cognitive complexity, neither questions nor response scales are born equal. Questionnaire items vary in the complexity of the measured constructs and of how they may be operationalised (Converse, 1976). Among the standard-module ESS questions, none touch upon truly complex constructs, except for the 21-item Portrait Value Questionnaire (PVQ), which measures the ten Schwartz values (Davidov et al., 2008; Schwartz, 2003). However, due to concerns about length, the PVQ was not included in the self-administered ESS10 questionnaires (ESS, 2023). Nevertheless, all question items selected for our study push respondents to engage in substantial cognitive effort. The household roster items probe for factual information about personal details of household members other than the respondent, while the three opinion questions touch upon somewhat abstract political concepts on which many respondents likely hold no firm prior opinions. Providing valid answers to both kinds of questions requires attention and some reflection, and is not straightforward.

While both types of items in our analysis involve cognitively burdensome questions, the design of their response options contrastively differ in terms of difficulty for the respondent. The opinion questions rely on the end-labelled 0-to-10 response formats, common in the ESS questionnaire, so respondents are exposed to such scales multiple times, and are likely to have been 'socialised' to use them through participating in various commercial surveys before (Groves 2011). However, the household's member enumeration questions, despite being concerned with potentially simpler factual information, involve a unique sequence and response options format. The respondent encounters such response options for the first time in the questionnaire completion process. For the former, responding is relatively easy even without reading the question, while the latter's response options require more attention on the part of the respondent and stronger memory recall, so they produce higher item nonresponse (Lipps & Monsch, 2022). Additionally, question formats that are not single-item format scales have a higher nonresponse rate (Messer et al., 2012). Arguably, for the opinion questions, the interviewer's presence may result in higher nonresponse rates, not because of the complexity of the response options, but due to social desirability biases (Kreuter et al., 2008). The three opinion items we selected for further analysis do not seem to count among very sensitive questions (they cover political ideology scale, political trust to UN, and EU expansion); however, they are rather typical 'no opinion' or 'do not know' items (e.g., Bearce & Jolliff Scott, 2019; Dassonneville & McAllister, 2025; Stoeckel, 2012), which means that they measure an opinion for which a significant proportion of respondents lack a sufficiently well-formed, accessible or justifiable evaluation. Thus, a common and meaningful response outcome is therefore a 'no opinion' or 'don't know' response (Krosnick, 1991). In addition, some people, such as those with lower incomes, who tend to have a lower perception of their political efficacy and less confidence in expressing their political views, have been found to be less likely to express their views in survey research (Berinsky, 2004; Laurison, 2015). Since the topics the opinion items cover concern political issues that some people might find difficult to assess (Wojcik et al., 2021) or consider private, interviewers might be less likely to probe about them, thereby improving respondents' motivation to answer.

Hypotheses

Nonresponse to household enumeration items

Interviewer assistance eases the cognitive burden of providing household demographics, as it allows for on-demand explanations of question wording, instructions, and response options. Furthermore, the presence of a well-trained interviewer facilitates the respondent's encounter with unfamiliar response formats. On the other hand, in self-administered surveys, strong satisficing is an efficient means of evading the burden of response to such factual questions with complex response options, as skipping items is inherently easier than navigating the response options.

H1: For questions with unfamiliar and complex response options, item nonresponse rates increase when the survey mode is switched from interviewer-assisted to self-completion protocols.

Nonresponse to standard opinion items

Opinion questions tend to have relatively uniform response formats in the ESS, with widespread implementation of end-labelled 11-point rating scales. In self-administered surveys, weak satisficing is a more likely pathway than strong satisficing for such question items, as respondents can quickly select any response on the scale provided without much cognitive engagement with the question semantics. A self-administering respondent is not under any immediate time or social pressure to respond truthfully and attentively to questions about political views, which makes providing any answer a more optimal solution than nonresponse.

H2: For opinion questions with familiar and simple response options, item nonresponse rates decrease when the survey mode is switched from interviewer-assisted to self-completion protocols.

Data and methods

Data

Our analysis encompasses data from 27 countries that participated in Rounds 9, 10, and 11 of the European Social Survey (ESS). In Round 9, all country surveys relied on the interviewer-assisted mode (CAPI). Due to COVID-19 disruptions, eight countries opted to switch to self-completion protocols in ESS10 with Push-to-Web priority, and in Round 11, all were back to interviewer-assisted modes. In Figure 1, we note the difference between interviewer-assisted fieldwork execution and self-completion. While we take note of internal mode distinctions (paper vs. web for self-completion, and CAPI vs. web-video for interviewer-assisted protocols), our analysis does not investigate their potential effects, as the respondents were not randomly split across the modes. In addition, we also note that the assignment into either the mode-switch or no-switch categories was not randomised, reflecting the specific needs and restrictions faced by the countries participating in the pandemic-disrupted Round 10. However, the eight mode-switchers are diverse geographically and in terms of broad socio-economic characteristics. Using the documented fieldwork durations and response rates from ESS Rounds 9–11, we noted that the Round 10 switchers do not constitute a distinct 'procedural cluster' compared to non-switchers. There is substantial overlap in fieldwork duration between the two groups in every Round (ESS9: 81–267 days for switchers vs 30–225 days for non-switchers; ESS10: 39–164 days vs 95–389 days; ESS11: 115–412 days vs 67–329 days) and in response rates (ESS9: 27.6–60.4% vs 34.9–69.4%; ESS10: 14.7–39.2% vs 20.9–72.5%; ESS11: 23.7–42.6% vs 27.0–76.6%, respectively for mode-switching and non-swithcing countries). Thus, the countries in both groups prove consistent with the ESS design principles and documentation structure, which are based on strict random probability sampling and standardised methodological requirements (including a minimum target response rate). Furthermore, in all countries, both pre- and post-pandemic ESS Rounds utilised interviewer-assisted modes, allowing us to control their baseline characteristics. Thus, our models compare survey outcomes in the two country categories for each of the three Rounds.

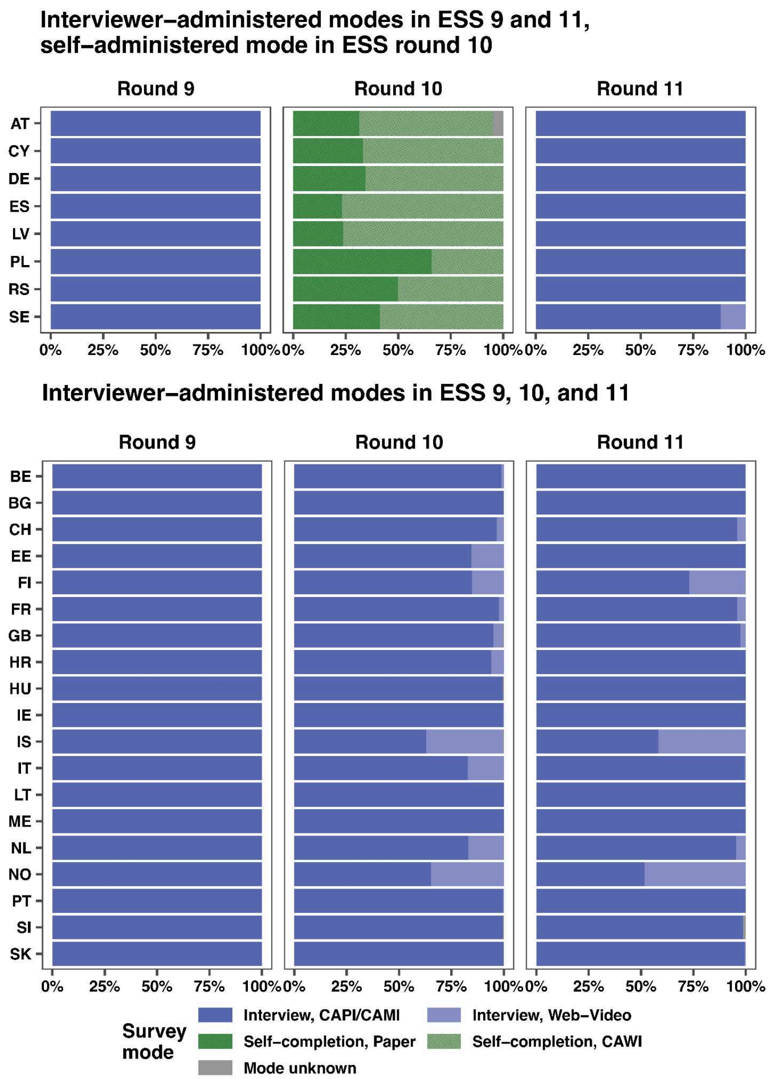

Figure 1. Modes of data collection in the ESS rounds 9, 10, and 11

We investigate item-nonresponse rates for two types of items in the ESS questionnaire. On the one hand, the items with complex response formats are principally featured in the household roster section. We selected three such question items: gender, year of birth, and the respondent's relationship to all other household members. On the other hand, we focused on three opinion questions with the highest rates of item nonresponse: left-right self-placement, trust in the United Nations, and attitudes towards European Unification.

As expressed in **H1**, for items with complex response formats, the item-nonresponse rates are expected to rise in the mode-switching countries in ESS10. However, there should be no contrast between their respective rates and the non-switching countries in ESS9 and 11, where all countries implemented interviewer-assisted modes. Conversely, as expressed in **H2** for standard opinion items, the nonresponse rates should be lower for mode-switchers in ESS10. Furthermore, in those eight countries, the item-nonresponse rates are expected to be similar for ESS9 and 11, which both implemented similar interviewer-assisted protocols, and in non-switching countries, the item-nonresponse rates are expected to be similar in all three rounds.

Outcome variables

Item nonresponse measures for complex response option items

Module F of the ESS questionnaire (household roster items) employs a multistage response process. Respondents are first required to report the total number of household members living in a dwelling and subsequently provide detailed, yet straightforward, factual data for each additional member. For each household member other than the respondent, the survey aims to collect information such as gender, year of birth, and the relationship of that person to the respondent.

Let r_k denote the observed item-nonresponse for the household composition items (gender, year of birth, relationship to respondent, respectively) corresponding to the household member k (with $k = 2, 3, ..., 8$) and let n_k denote the number of respondents for whom the question about the household member k is applicable (i.e., those households with at least k members). Note that the size of the household may exceed 8 (in the ESS the maximum is set up to 15); however, we excluded all such cases due to a limited number of respondents reporting more than 8 household members. For each applicable household member, the indicator r_k takes a value of 1 if the response is classified as a nonresponse (missed information) and 0 otherwise. To calculate item nonresponse outcomes for these complex format items; we first compute the average nonresponse rate for each household member position k, which we denoted $\bar{r}_k$. The overall item nonresponse rate for individual country-survey and all three outcomes related to complex questions, i.e., INR_{gndr}, INR_{year}, INR_{rshipa} is then calculated as a weighted average across the applicable households:

$$INR_{\{gndr;year;rshipa\}} = \frac{\sum_{k=2}^{8} n_k \bar{r}_k}{\sum_{k=2}^{8} n_k}$$

This formula ensures that each household member's position contributes to the overall rate proportionally to the number of respondents for whom that position is applicable. To illustrate this principle, consider a scenario where a small proportion of households include a fourth member. In this case, the nonresponse rate for the fourth member, denoted by $\bar{r}_4$ will contribute proportionally less to INR_{gndr}, INR_{year} or INR_{rshipa}, than the nonresponse rate for the second member $\bar{r}_2$, which applies to a larger group of respondents.

Item nonresponse measures for standard opinion items

In the case of the opinion items, the approach was initiated by identifying a set of core ESS variables that were consistently administered across ESS Rounds 9, 10, and 11 (a complete list is available in the Supplementary Online Materials; hereafter SM). For each of these items, cases of nonresponse (missing answers) were identified, and a binary indicator I_{ij} was created for each respondent j on each item i. Here, $I_{ij} = 1$ indicates a nonresponse (i.e., missing or invalid answer) and $I_{ij} = 0$ denotes a valid response. The share of missing data for each item in each country-survey in ESS9 was then computed by averaging these binary indicators. Finally, the items were ranked according to their nonresponse rates, and the three items with the highest overall rates were selected for further analysis (namely, left-right political self-placement: $lrscale$, trust in the United Nations: $trstun$, and attitudes toward European unification: $euftf$).

Finally, for three selected items and country-surveys from ESS9, 10, and 11, separately, we calculated the item nonresponse rate by averaging binary indicators I_{ij} across all respondents. Mathematically, for each item i, the item nonresponse is defined as:

$$INR_{\{lrscale;trstun;euftf\}} = \frac{1}{N}\sum_{j=1}^{N} I_{ij}$$

where N represents the number of respondents in a given country. This straightforward approach yields the proportion of respondents who did not provide a valid answer for a particular item.

Analytical procedure

We employed a Bayesian modelling framework using the 'brms' package (Bürkner, 2017, 2018) in R (R Core Team, 2021). This approach enables a flexible specification of the likelihood function, hierarchical random effects, and the derivation of complete posterior distributions for the parameters of interest.

The analysis focused on item-nonresponse measures ranging from 0 to 1. We first attempted to distinguish the outcome variables that exhibited substantial zero inflation (i.e., a notable proportion of country surveys with no item nonresponse) from those that did not, but our preliminary analyses did not detect inflation of zeros in the data (see Table SM1 and Table SM7 in SM). Thus, we specified a standard beta family model for all items with complex response options and all three standard opinion items (we transformed outcome variables by adding a small constant 1×10^{-6} to any zero values, which is necessary as the beta distribution cannot accommodate exact zero but has no effect on the results).

To account for the nested structure of the data (surveys nested within countries), random effects were included, with random intercepts for each country and random slopes for the ESS Round within each country. This hierarchical approach acknowledges the possibility that countries may differ in their baseline levels of item nonresponse and that these levels may change across rounds. Modelling such heterogeneity explicitly allows for more precise estimates of the overall effects of changing data collection mode.

To evaluate the impact of the transition to self-completion protocols in Round 10 on item nonresponse, compared to countries maintaining interviewer-assisted modes of data collection, a 'country group' variable was introduced to distinguish countries using self-completion protocols in R10 from those using interviewer-assisted protocols in all three rounds. Model 1 (a baseline model) controlled for the basic covariates and estimated the main effects of the ESS Round and country group for each dependent variable. In contrast, Model 2 (the model of primary interest for testing the hypotheses) included the interaction between these two factors to determine whether the effects depended on the data collection mode. Rounds 9 and the group of countries utilising interviewer-assisted modes of data collection in all three rounds served as the reference categories. Detailed mathematical formulas for two Bayesian regression models under the beta family assumption are provided in the Appendix.

All models were fitted with four Markov Monte Carlo chains, each running for 4,000 iterations (1,000 of which were warmups). Two control parameters were set: "max treedepth" = 15 and "adapt delta" = 0.99. The purpose of these parameters was to improve sampler efficiency and reduce the risk of divergent transitions, thus enhancing the reliability of the posterior estimates. A fixed random seed was also used to ensure the reproducibility of all results presented in this paper (see the replication code in SM; however, for replicating the exact results in the "brms" package in R, the same operating system: macOS 15.3.1, rstan: 2.32.6, and same clang++ compiler version: 16.0.0 for Stan is required). Subsequently, a posterior summary and credibility intervals were employed to quantify the strength of evidence for differences in item nonresponse across rounds and data collection modes. Implementing a Bayesian approach offers several advantages, chief among them providing complete probability distributions for parameters. This is particularly valuable when estimating potentially small or uncertain differences across countries and rounds.

Results

To test the hypotheses, our analysis examines the interaction between ESS9, 10, and 11, and membership in a country group (eight mode-switching countries vs. nineteen non-switchers). In the results section, we provided a graphical visualisation of the interaction terms and the Bayesian regression parameters as well as tables with all regression terms and 95% credible intervals; however, the complete outputs, including convergence diagnostics, and trace plots for every model and each Markov chain, can be found in Sections 2.1 and 2.2 of the SM. Notably, all models converged, as evidenced by all R-hat convergence diagnostic values approximately equal to 1. Furthermore, the effective sample sizes (both Bulk and Tail) were sufficiently large for most parameters, indicating that the posterior estimates are founded on robust sampling and are less susceptible to instability in the tails or central region of the parameter's distribution.

Figure 2 provides a visual summary of interaction effects, displaying the predicted posterior average values of the probability of item-nonresponse occurrence on the y-axis for each combination of round and country groups. The three panels in the upper row show the effects of switching to self-completion for questions with complex response options, while the lower row does the same for the three standard opinion questions with end-labelled 0-to-10 response formats.

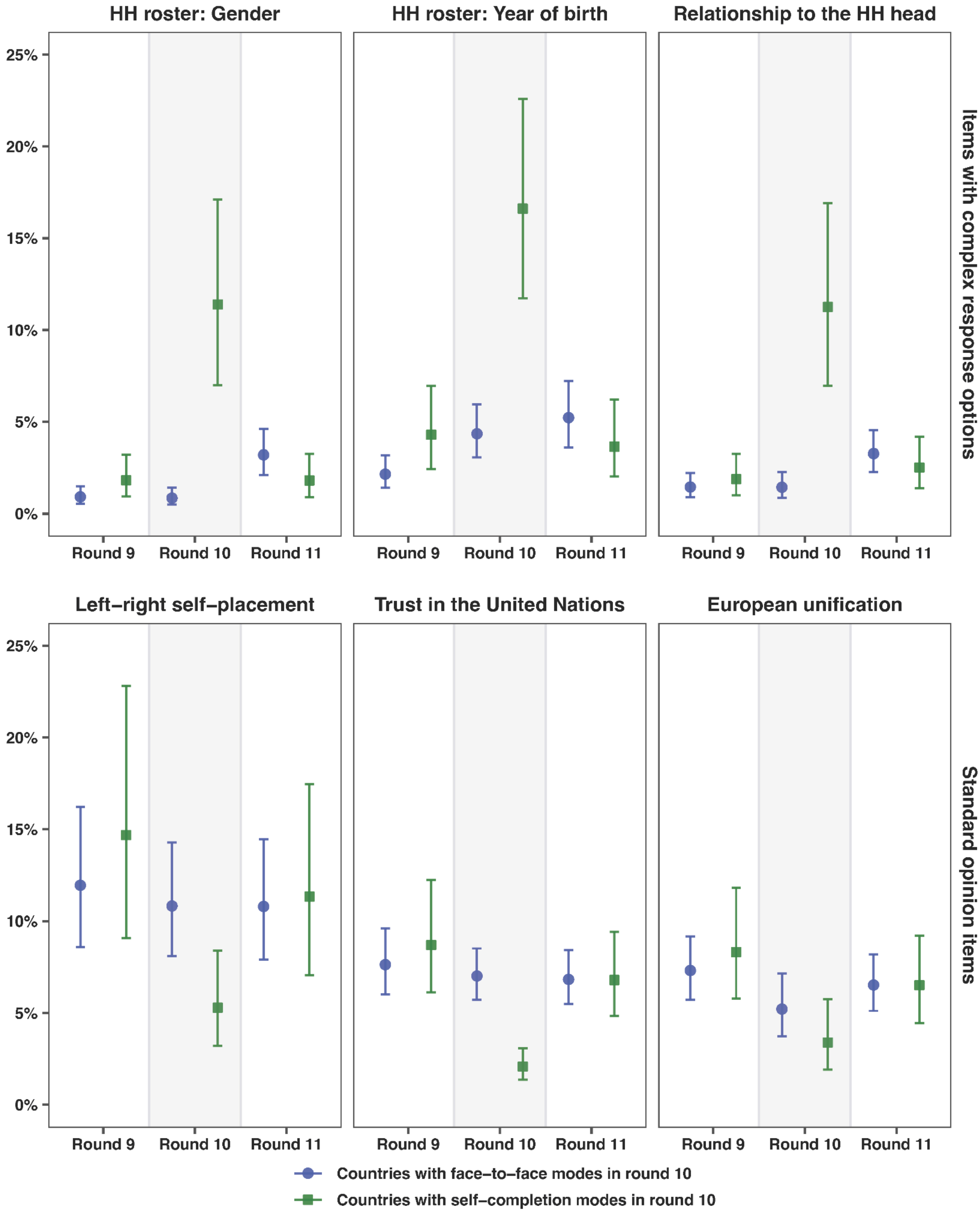

Figure 2. Interaction terms (ESS round x Country group) for testing the impact of switching to self-completion protocols in R10 on item nonresponse in complex response options and simple opinion items

Figure 2 supports **H1**, as for household roster items with complex or unfamiliar response formats, we found that mode-switching countries demonstrate a clear spike in item nonresponse in Round 10 when self-completion protocols were in use, compared to their own interviewer-assisted baseline in Rounds 9 and 11, as well as compared to non-switching countries. Conversely, the pattern we found aligns with **H2** for standard opinion items. In self-completion countries, we found that item nonresponse is lower in Round 10 (particularly for left–right self-placement and trust in the UN), before rising again once interviewer-assisted modes resume in Round 11. We provide detailed results for testing **H1** and **H2** in the following sections.

Item nonresponse and items with complex response options

Tables 1-3 summarise Bayesian regressions (with posterior odds ratios and credible intervals with 95% posterior probability) for all three questions constituting a battery of cognitively demanding complex response format items, i.e., for the gender of household members, the year of birth of household members, and the relationship of household members to the respondent.

Table 1. Posterior Odds Ratios with 95% Credible Intervals for item nonresponse rate in questions on the gender of household members as the outcome variable

Predictors	**Model 1**		**Model 2**	
	Odds Ratios	*CI (95%)*	*Odds Ratios*	*CI (95%)*
Intercept	0.009	0.005 – 0.017	0.009	0.005 – 0.017
Round 10 vs. 9	1.833	0.879 – 3.751	0.919	0.448 – 1.950
Round 11 vs 9	2.308	1.245 – 4.359	3.561	1.778 – 6.994
Self-completion modes in Round 10	2.033	0.952 – 4.478	1.993	0.792 – 4.974
Round 10 x Self-completion modes in Round 10	-	-	7.559	2.268 – 24.267
Round 11 x Self-completion modes in Round 10	-	-	0.278	0.080 – 0.990
ICC	0.70		0.29	
N Countries	27		27	
N surveys	81		81	
Marginal R^2 / Conditional R^2	0.064 / 0.689		0.491 / 0.667	

Table 2. Posterior Odds Ratios with 95% Credible Intervals for item nonresponse rate in questions on the age of household members as the outcome variable

Predictors	Model 1		Model 2	
	Odds Ratios	*CI (95%)*	*Odds Ratios*	*CI (95%)*
Intercept	0.021	0.013 – 0.034	0.022	0.013 – 0.036
Round 10 vs. 9	2.716	1.698 – 4.321	2.058	1.195 – 3.739
Round 11 vs 9	1.759	1.053 – 3.072	2.500	1.385 – 4.484
Self-completion modes in Round 10	2.340	1.240 – 4.327	2.033	0.852 – 4.616
Round 10 x Self-completion modes in Round 10	-	-	2.164	0.895 – 5.424
Round 11 x Self-completion modes in Round 10	-	-	0.342	0.115 – 1.002
ICC	0.39		0.26	
N Countries	27		27	
N surveys	81		81	
Marginal R^2 / Conditional R^2	0.298 / 0.641		0.495 / 0.689	

Table 3. Posterior Odds Ratios with 95% Credible Intervals for item nonresponse rate in questions on the relationship of household members to the respondent as the outcome variable

Predictors	Model 1		Model 2	
	Odds Ratios	*CI (95%)*	*Odds Ratios*	*CI (95%)*
Intercept	0.014	0.007 – 0.023	0.015	0.008 – 0.025
Round 10 vs. 9	1.803	0.937 – 3.360	0.989	0.501 – 1.919
Round 11 vs 9	1.900	1.134 – 3.194	2.293	1.270 – 4.187
Self-completion modes in Round 10	1.679	0.826 – 3.605	1.300	0.525 – 3.134
Round 10 x Self-completion modes in Round 10	-	-	6.620	2.330 – 20.441
Round 11 x Self-completion modes in Round 10	-	-	0.584	0.191 – 1.757
ICC	0.69		0.30	
N Countries	27		27	
N surveys	81		81	
Marginal R^2 / Conditional R^2	0.056 / 0.691		0.415 / 0.671	

The magnitude of the interaction effects for the Round 10 × self-completion modes of data collection (see models with the second specification) varied across three items. For the gender of the household members, the odds of a missing answer incresed over 7 times (OR = 7.56, with 95% CI ranging from 2.28 to 24.27) In estimating the interaction term for the item missingness to the year of birth, the estimate was more modest, and the interval overlapped the odds ratio of 1 (OR = 2.16, 95% CI: 0.90–5.24). However, regarding the relationship with the respondent, the odds were over six times higher (OR = 6.61, 95% CI: 2.33–20.44). Notably, the comeback to the face-to-face modes in Round 11 resulted in a substantial decrease in the odds for each item, indicating a near-complete restoration of baseline nonresponse levels in Round 9.

Overall, the findings indicate that transitioning from interviewer-assisted to self-completion protocols in ESS10 significantly increased item nonresponse in all three household roster items. Specifically, in countries that employed interviewer-assisted protocols in all ESS rounds, the predicted levels of item nonresponse remained relatively stable, suggesting that interviewer-administered modes are not associated with substantial fluctuations across Rounds 9, 10, and 11. Conversely, countries' mode-switching to self-completion in ESS10 exhibited a substantial increase in nonresponse, as evidenced by the odds ratios for the interaction terms for all three variables above 1.0, with 95% posterior probability. This result suggests that self-completion modes of data collection may have a notable impact beyond baseline differences. When the eight countries reverted to interviewer-assisted protocols in ESS11, their predicted item nonresponse levels returned to the lower values observed in ESS9.

The results corroborate the hypothesis (**H1**) that switching the survey mode from interviewer-assisted to self-completion results in higher nonresponse rates for questions with unfamiliar and complex response options. Furthermore, the transition to self-administered surveys in one round generates a discernible 'dip-and-rebound' effect, whereby nonresponse rates for the eight mode-switching countries return to the baseline values of ESS9 when they switch back to interviewer-administered surveys in ESS11.

Item nonresponse and standard opinion items

For opinion questions with familiar and standard response scales, the three Bayesian regression models reveal a distinctly different effect of switching from interviewer-assisted to self-completion data collection modes (see Tables 4-6).

Table 4. Posterior Odds Ratios with 95% Credible Intervals for item nonresponse rate in questions on the left-right political orientation as the outcome variable

Predictors	**Model 1**		**Model 2**	
	Odds Ratios	*CI (95%)*	*Odds Ratios*	*CI (95%)*
Intercept	0.162	0.103 – 0.259	0.136	0.086 – 0.210
Round 10 vs. 9	0.688	0.517 – 0.921	0.893	0.699 – 1.168
Round 11 vs 9	0.850	0.691 – 1.052	0.890	0.700 – 1.153
Self-completion modes in Round 10	0.650	0.285 – 1.485	1.280	0.559 – 2.815
Round 10 x Self-completion modes in Round 10	-	-	0.359	0.217 – 0.603
Round 11 x Self-completion modes in Round 10	-	-	0.832	0.531 – 1.274
ICC	0.84		0.81	
N Countries	27		27	
N surveys	81		81	
Marginal R^2 / Conditional R^2	0.068 / 0.922		0.081 / 0.880	

Table 5. Posterior Odds Ratios with 95% Credible Intervals for item nonresponse rate in questions on trust in the United Nations as the outcome variable

Predictors	**Model 1**		**Model 2**	
	Odds Ratios	*CI (95%)*	*Odds Ratios*	*CI (95%)*
Intercept	0.101	0.072 – 0.147	0.083	0.060 – 0.113
Round 10 vs. 9	0.641	0.468 – 0.871	0.909	0.719 – 1.164
Round 11 vs 9	0.851	0.708 – 1.033	0.885	0.713 – 1.108
Self-completion modes in Round 10	0.561	0.285 – 1.083	1.154	0.657 – 2.057
Round 10 x Self-completion modes in Round 10	-	-	0.243	0.142 – 0.411
Round 11 x Self-completion modes in Round 10	-	-	0.864	0.586 – 1.264
ICC	0.64		0.67	
N Countries	27		27	
N surveys	81		81	
Marginal R^2 / Conditional R^2	0.068 / 0.922		0.081 / 0.880	

Table 6. Posterior Odds Ratios with 95% Credible Intervals for item nonresponse rate in questions on attitudes towards European unification as the outcome variable

Predictors	**Model 1**		**Model 2**	
	Odds Ratios	*CI (95%)*	*Odds Ratios*	*CI (95%)*
Intercept	0.082	0.060 – 0.110	0.079	0.057 – 0.107
Round 10 vs. 9	0.589	0.417 – 0.814	0.701	0.482 – 0.995
Round 11 vs 9	0.845	0.677 – 1.070	0.883	0.679 – 1.164
Self-completion modes in Round 10	1.029	0.612 – 1.770	1.151	0.655 – 2.090
Round 10 x Self-completion modes in Round 10	-	-	0.550	0.276 – 1.094
Round 11 x Self-completion modes in Round 10	-	-	0.869	0.526 – 1.398
ICC	0.58		0.57	
N Countries	27		27	
N surveys	81		81	
Marginal R^2 / Conditional R^2	0.105 / 0.659		0.125 / 0.661	

An investigation of the Bayesian posterior estimates led to the conclusion that the interaction term between Round 10 and self-completion modes has substantially decreased the odds of item nonresponse. Specifically, for the left-right self-placement, the odds ratio was reduced by approximately two-thirds (OR = 0.36, 95% CI: 0.21–0.60). For the United Nations, the reduction was three-quarters (OR = 0.23, 95% CI: 0.14–0.41), and for the European unification item, the reduction was around one-half (OR = 0.55, 95% CI: 0.28–1.09). In the latter case, however, the credibility interval included 1, so the effect was negligible with the 95% posterior probability.

The regression results indicate that implementing self-completion modes of data collection in ESS10 was associated with a lower predicted probability of item nonresponse occurrence (even if negligible in the last item). In countries that transitioned to self-completion protocols in Round 10, nonresponse declined for standard opinion items. However, when these countries returned to interviewer-assisted modes of data collection in Round 11, their item nonresponse levels returned to those closer to (or occasionally below) those in Round 9. These findings corroborate the hypothesis **(H2)** that switching to self-completion reduces item nonresponse rates for standard opinion questions with familiar and simple response formats. Notably, this effect constitutes a mirror image of the effect attested for questions with complex and unfamiliar response options, with a corresponding rebound effect in ESS11 to the baseline values of ESS9.

Discussion and Conclusions

Switching from an interviewer-assisted to a self-administered mode of survey execution constitutes a major challenge for cross-national projects aiming to provide reliable comparative data over time. Mode-divergent item nonresponse rates constitute only a part of a larger conundrum, and the two types of question formats in our analysis are merely two contrasting examples of differential nonresponse patterns. Our analysis demonstrated how the switch to self-completion mode in ESS10 affected data quality, while item nonresponse remained stable for countries that did not change from interviewer-assisted to self-completion mode. We focused on the contrast between items with unfamiliar and complex response options for the ESS respondents, where the incidence of nonresponse increases in the absence of the interviewer, and items with familiar and simple response formats, for which a corresponding decline in nonresponse rates can be observed. A plausible interpretation of this decline is that it results from weak satisficing, indicating that it is not a good sign for survey quality; respondents find it easier to be less attentive when self-administering than they would in the presence of an interviewer. Conversely, when it comes to complex and unfamiliar items, the absence of the interviewer makes it less likely for the respondent to engage with such parts of the questionnaire attentively. This was the case in the ESS household roster items, which are crucial for estimating many survey measures relating to a household situation. A higher item nonresponse rate in self-completion surveys for questions about household members poses a risk to the comparability of mode-mixing survey data over time and their internal reliability.

Study limitations

The limitations of our study need to be acknowledged. Although the analysis is based on a quasi-experimental design, in which countries that switched to self-completion in ESS10 come from a variety of European regions, economic contexts, and welfare regimes, their decisions to implement self-completion were not random. The potentially confounding but unmeasured factors in this design include unequal capacities of survey infrastructures, socio-cultural or other country-specific characteristics. These factors may be correlated with some unobserved characteristics of survey administration or respondent behaviour and cannot therefore be ruled out as potential sources of bias.

Unfortunately, the available ESS data does not allow for testing our satisficing-based interpretation of the differential item nonresponse patterns. Given that they constitute a problem for data quality, and given the much better capacity for monitoring the response process in computer-based self-administered surveys, high-quality projects should strive to make available a standardised paradata set including last-question time, item latency distribution, browser focus changes (Mathews et al., n.d.; Schenk & Reuß, 2024; Vehovar et al., 2024). Such paradata availability would identify hidden satisficing behaviours that raw item nonresponse cannot capture, thereby providing a more comprehensive cross-mode quality metric. Furthermore, as both strong and weak satisficing emerge in real time, leveraging latency patterns, scroll depth, and break-off traces allows survey organisations to intervene on the spot. A shift towards responsive and adaptive survey designs (Einarsson et al., 2024) should be considered when pivoting to self-completion modes. Some cross-national applications have already been piloted in large projects such as SHARE and the ESS. The real-time analysis of paradata can facilitate the identification of specific types of satisficing respondents, thereby moderating mode-specific response behaviours and enhancing the uniformity of data quality across diverse data collection methods in longitudinal, multi-wave, cross-national studies.

Study implications

Our study has implications for the new era of survey research, when survey data requires additional insight and support from new technology tools (Groves, 2011) - more recently referred to as ‘smart surveys’ (Schouten et al., 2025). Moving survey data collection online requires changes in the survey logic and embedding more ‘smart features’, including interactive digital tools replacing the interviewer that can be more effective in keeping respondents engaged, yet would not introduce social desirability bias, and crucially could make survey responding a less daunting task.

Just as interviewer presence once did, automated prompts, attention checks, or conversational bots alter the respondent’s involvement in surveys. To mitigate weak satisficing, survey execution protocols should implement real-time attention checks (Gummer et al., 2018; Newman et al., 2021; Ternovski & Orr, 2022). Timing attention checks and adaptive ‘speed-bump’ prompts are widely recommended to turn potential nonrespondents into engaged respondents, mitigating the risks highlighted by our finding related to self-administration effects on standard opinion items. This approach has been shown to identify cases of inattention and improve data quality (Oppenheimer et al., 2009) and combine them with timing-based ‘speed bump’ prompts that appear when paradata reveal ultra-fast completion or straight-lining (Hillygus & LaChapelle, 2022). These layered interventions can transform potential weak satisficing nonrespondents into attentive respondents. This directly addresses the spurious dip in item nonresponse as it prevents careless answers from silently replacing legitimate skips.

When it comes to complex and unfamiliar question items and other questionnaire parts that traditionally require interviewer elicitation, the tendency to employ strong satisficing can also be mitigated by existing and emerging technologies, such as virtual assistants capable of providing instructions and explanations to confused respondents. In this respect, the recent applications of experimental 'AI interviewers,' e.g., Wuttke et al. (2025), seem particularly important. Preliminary field trials have demonstrated that large language model AI interviewers can replicate a substantial proportion of the clarifying role typically performed by human interviewers, while maintaining web efficiency. This approach has been shown to reduce nonresponse to cognitively demanding items without incurring additional costs; however, the key design issues identified are transparency and the ability to fall back on human assistance.

Acknowledgments

None

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

None

References

Alexander, E. C. (2017). Don't Know or Won't Say? Exploring How Colorblind Norms Shape Item Nonresponse in Social Surveys. *Sociology of Race and Ethnicity*, *4*(3), 417–433. https://doi.org/10.1177/2332649217705145

Alwin, D. F., & Krosnick, J. A. (1991). The Reliability of Survey Attitude Measurement: The Influence of Question and Respondent Attributes. *Sociological Methods & Research*, *20*(1), 139–181. https://doi.org/10.1177/0049124191020001005

Bearce, D. H., & Jolliff Scott, B. J. (2019). Popular non-support for international organizations: How extensive and what does this represent? *The Review of International Organizations*, *14*(2), 187–216. https://doi.org/10.1007/s11558-019-09351-3

Beatty, P., & Herrmann, D. (2002). To answer or not to answer: Decision processes related to survey item nonresponse. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & J. R. Little (Eds.), *Survey nonresponse* (pp. 71–85). Willey.

Berinsky, A. J. (2004). *Silent voices: Public opinion and political participation in America*. Princeton University Press.

Berteen, E., & Phillips, natasha. (2021). *European Social Survey Round 10 self-completion paper questionnaire - User testing report*. https://www.europeansocialsurvey.org/sites/default/files/2024-09/ESS_SC_paper_questionnaire_NatCen_user_testing_report.pdf

Blazek, D. R., & and Siegel, J. T. (2024). Preventing satisficing: A narrative review. *International Journal of Social Research Methodology*, *27*(6), 635–648. https://doi.org/10.1080/13645579.2023.2239086

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health*, *27*(3), 281–291. https://doi.org/10.1093/pubmed/fdi031

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, *10*(1), 395. https://doi.org/10.32614/RJ-2018-017

Coffey, S., Maslovskaya, O., & McPhee, C. (2024). Recent Innovations and Advances in Mixed-Mode Surveys. *Journal of Survey Statistics and Methodology*, *12*(3), 507–531. https://doi.org/10.1093/jssam/smae025

Converse, J. M. (1976). Predicting no opinion in the polls. *Public Opinion Quarterly*, *40*(4), 515–530.

Cornesse, C., & Blom, A. G. (2020). Response Quality in Nonprobability and Probability-based Online Panels. *Sociological Methods & Research*, *52*(2), 879–908. https://doi.org/10.1177/0049124120914940

Cornesse, C., Blom, A. G., Dutwin, D., Krosnick, J. A., De Leeuw, E. D., Legleye, S., Pasek, J., Pennay, D., Phillips, B., Sakshaug, J. W., Struminskaya, B., & Wenz, A. (2020). A Review of Conceptual Approaches and Empirical Evidence on Probability and Nonprobability Sample Survey Research. *Journal of Survey Statistics and Methodology*, *8*(1), 4–36. https://doi.org/10.1093/jssam/smz041

Dassonneville, R., & McAllister, I. (2025). The Direction and Meaning of Left-Right in Postcommunist Societies. *International Journal of Public Opinion Research*, *37*(2), edaf024. https://doi.org/10.1093/ijpor/edaf024

Davidov, E., Schmidt, P., & Schwartz, S. H. (2008). Bringing Values Back In: The Adequacy of the European Social Survey to Measure Values in 20 Countries. *Public Opinion Quarterly*, *72*(3), 420–445. https://doi.org/10.1093/poq/nfn035

Einarsson, H., Cernat, A., & Shlomo, N. (2024). Responsive and Adaptive Designs in Repeated Cross-National Surveys: A Simulation Study. *Journal of Survey Statistics and Methodology*, *12*(4), 906–931. https://doi.org/10.1093/jssam/smad038

European Social Survey. (2023). *Note on ESS Round 10 data collection modes, edition 1.0*. https://stessrelpubprodwe.blob.core.windows.net/data/round10/methods/ESS10_note_on_data_modes_e01_0.pdf

European Social Survey. (2025). *Data Collection: Switch to Self-Completion*. https://www.europeansocialsurvey.org/methodology/methodology/data-collection-sc

Fitzgerald, R., & Aizpurua, E. (2024). 20 Years of the European Social Survey: An Interview with Rory Fitzgerald. *Survey Practice*, *17*. https://doi.org/10.29115/SP-2024-0025

Groves, Robert M. (2011). Three eras of survey research. Public opinion quarterly, 75(5), 861-871. https://doi.org/10.1093/poq/nfr057

Gummer, Tobias, Roßmann, Joss, & Silber, Henning. (2018). Using Instructed Response Items as Attention Checks in Web Surveys: Properties and Implementation. *Sociological Methods & Research*, *50*(1), 238–264. https://doi.org/10.1177/0049124118769083

Hillygus, D. S., & LaChapelle, T. (2022). Chapter 2: Diagnosing survey response quality. In T. Rudolph (Ed.), *Handbook on Politics and Public Opinion* (pp. 10–25). Edward Elgar Publishing Limited.

Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Response Bias*. *Public Opinion Quarterly*, *67*(1), 79–125. https://doi.org/10.1086/346010

Jabkowski, P . & Piekut, A. (2024) Between task complexity and question sensitivity: nonresponse to the income question in the 2008–2018 European Social Survey. Survey Research Methods, 18(2), 113-135. https://doi.org/10.18148/srm/2024.v18i2.8117

Kraemer, F., Silber, H., Struminskaya, B., Weiß, B., Bosnjak, M., Koßmann, J., & Sand, M. (2023). atisficing Response Behavior Across Time: Assessing Negative Panel Conditioning Using an Experimental Design with Six Repetitions. *Survey Research Methods*, *17*(3), 269–300. https://doi.org/10.18148/srm/2023.v17i3.7986

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*(3), 213–236.

Krosnick, J. A., Narayan, S., & Smith, W. R. (1996). Satisficing in surveys: Initial evidence. *New Directions for Evaluation*, *1996*(70), 29–44. https://doi.org/https://doi.org/10.1002/ev.1033

Laurison, D. (2015). The willingness to state an opinion: Inequality, don't know responses, and political participation. Sociological Forum. 30(4). 925-948. https://onlinelibrary.wiley.com/doi/full/10.1111/socf.12202

Lipps, Oliver, & Monsch, Gian-Andrea. (2022). Effects of Question Characteristics on Item Nonresponse in Telephone and Web Survey Modes. *Field Methods*, *34*(4), 318–333. https://doi.org/10.1177/1525822X221115838

Lugtig, P. (2024). *ESS round 10 mode experiments in Great Britain and Finland. Findings on mode effects*. https://www.europeansocialsurvey.org/sites/default/files/2024-10/round-10-experimental-comparison-final.pdf

Maslovskaya, O., Struminskaya, B., & Durrant, G. (2022). The Future of Online Data Collection in Social Surveys: Challenges, Developments and Applications. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *185*(3), 768–772. https://doi.org/10.1111/rssa.12895

Mathews, M., Elliott, M. N., Parast, L., Williamson, S., Joestl, S. S., Stark, D. S., Lehrman, W., & Becker, K. (n.d.). What Paradata Can Tell Us Beyond Response Rates. Exploring Device Use and Push-To-Web Access Methods in a Survey of

Emergency Department Patient. *Survey Practice*, *16*(1). https://doi.org/10.29115/SP-2023-0006

Messer, B. L., Edwards, M. L., & Dillman, D. A. (2012). Determinants of Item Nonresponse to Web and Mail Respondents in Three Address-Based Mixed-Mode Surveys of the General Public. *Survey Practice*, *5*, 3086. https://api.semanticscholar.org/CorpusID:166782509

Mulligan, K., Krosnick, J. A., Smith, W., Green, M., & Bizer, G. (2001). Nondifferentiation on attitude rating scales: A test of survey satisficing theory. *Unpublished Manuscript*.

Newman, A., Bavik, Y. L., Mount, M., & Shao, B. (2021). Data Collection via Online Platforms: Challenges and Recommendations for Future Research. *Applied Psychology*, *70*(3), 1380–1402. https://doi.org/https://doi.org/10.1111/apps.12302

Nursel, A., & Stähli, M. E. (2025). Transitioning from interviewer-administered to self-administered survey modes: Implications for the quality of measures of political attitudes and behavior. *Frontiers in Political Science*, *7*. https://doi.org/10.3389/fpos.2025.1425888

Oliveri, S., Lanzoni, L., Petrocchi, S., Janssens, R., Schoefs, E., Huys, I., Smith, M. Y., Smith, I. P., Veldwijk, J., Wit, G. A. de, & Pravettoni, G. (2021). Opportunities and Challenges of Web-Based and Remotely Administered Surveys for Patient Preference Studies in a Vulnerable Population. *Patient Preference and Adherence*, *15*, 2509–2517. https://doi.org/https://doi.org/10.2147/PPA.S327006

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872. https://doi.org/https://doi.org/10.1016/j.jesp.2009.03.009

R Core Team. (2021). R: A language and environment for statistical computing. In *R Foundation for statistical computing, Vienna*. https://www.R-project.org/.

Roberts, C., Gilbert, E., Allum, N., & Eisner, L. (2019). Research Synthesis: Satisficing in Surveys: A Systematic Review of the Literature. *Public Opinion Quarterly*, *83*(3), 598–626. https://doi.org/10.1093/poq/nfz035

Schenk, P. O., & Reuß, S. (2024). Paradata in Surveys. In I. Huvila, L. Andersson, & O. Sköld (Eds.), *Perspectives on Paradata: Research and Practice of Documenting Process Knowledge* (pp. 15–43). Springer International Publishing. https://doi.org/10.1007/978-3-031-53946-6_2

Schwartz, S. (2003). A proposal for measuring value orientations across nations. In *Questionnaire package of the European SocialSsurvey*. http://www.europeansocialsurvey.org/docs/methodology/core_ess_questionnaire/ESS_core_questionnaire_human_values.pdf

Schouten, B., Lugtig, P., & Luiten, A. (2025). Can Smart Surveys Have a Positive Business Case? An Evaluation Based on Three Case Studies. Journal of Official Statistics, 41(2), 547-568. https://doi.org/10.1177/0282423X251321634

Singh, R. K. (2023). *ESS Mode Change: Response Scale Report*. https://www.europeansocialsurvey.org/sites/default/files/2024-09/ESS_mode_change_response_scale_report.pdf

Stoeckel, Florian. (2012). Ambivalent or indifferent? Reconsidering the structure of EU public opinion. *European Union Politics*, *14*(1), 23–45. https://doi.org/10.1177/1465116512460736

Stosic, Morgan D, Murphy, Brett A, Duong, Fred, Fultz, Amber A, Harvey, Summer E, & Bernieri, Frank. (2024). Careless Responding: Why Many Findings Are Spurious or Spuriously Inflated. *Advances in Methods and Practices in Psychological Science*, *7*(1), 25152459241231580. https://doi.org/10.1177/25152459241231581

Ternovski, J., & Orr, L. (2022). A Note on Increases in Inattentive Online Survey-Takers Since 2020. *Journal of Quantitative Description: Digital Media*, *2*(2), 1–35. https://doi.org/10.51685/jqd.2022.002

Vehovar, V., Berzelak, N., & Čehovin, G. (2024). Identifying a Set of Key Paradata Indicators in Web Surveys Based on the Relationship with Response Quality, Respondent Characteristics and Survey Estimates. *Survey Research Methods*, *18*(3), 225–249. https://doi.org/10.18148/srm/2024.v18i3.8304

West, B. T., & Blom, A. G. (2017). Explaining Interviewer Effects: A Research Synthesis. *Journal of Survey Statistics and Methodology*, *5*(2), 175–211. https://doi.org/10.1093/jssam/smw024

Wojcik, A. D., Cislak, A., & Schmidt, P. (2021). 'The left is right': Left and right political orientation across Eastern and Western Europe. The Social Science Journal, 1-17. https://www.tandfonline.com/doi/full/10.1080/03623319.2021.1986320

Wuttke, A., Aßenmacher, M., Klamm, C., Lang, M. M., Würschinger, Q., & Kreuter, F. (2025). *AI Conversational Interviewing: Transforming Surveys with LLMs as Adaptive Interviewers*.

Ye, C., Fulton, J., & Tourangeau, R. (2011). More positive or more extreme? A meta-analysis of mode differences in response choice. *Public Opinion Quarterly*, *75*(2), 349–365. https://doi.org/10.1093/poq/nfr009

Zhu, J. (1996). 'I don't know' in public opinion surveys in China: Individual and contextual causes of item non-response. *Journal of Contemporary China*, *5*(12), 223–244. https://doi.org/10.1080/10670569608724251

Appendix. Mathematical formulas for Bayesian regression models

As all outcomes did not display zero inflation, we used a simple model with the beta family. Let Y_{ij} denote the item-nonresponse measure for survey *i* in country *j*:

$$Y_{ij} \sim Beta\left(\alpha_{ij}, \beta_{ij}\right).$$

In the "brms" R package, the beta distribution is parametrised by a mean μ_{ij} and a precision term ϕ, such that: and a precision term ϕ, such that:

$$\alpha_{ij} = \mu_{ij}\phi$$

and

$$\beta_{ij} = \left(1 - \mu_{ij}\right)\phi$$

thus,

$$Y_{ij} \sim Beta\left(\mu_{ij}\phi\,, \left(1 - \mu_{ij}\right)\phi\right).$$

The link function for μ_{ij} is $logit\left(\mu_{ij}\right) = \eta_{ij}$, while ϕ is estimated as $log(\phi)$.

Specification of Model 1

Model 1 incorporates the main effects of the ESS round and country group (interviewer-assisted vs. self-completion), with random intercepts and random slopes for the ESS round:

$$\eta_{ij} = \beta_0 + \gamma_{0j} + \left(\beta_1 + \gamma_{1j}\right)I(Round = 10)_{ij} + \left(\beta_2 + \gamma_{2j}\right)I(Round = 11)_{ij} + \beta_3 I(SC\ G$$

where:

- β is a vector of regression parameters.
- $I(Round = 10)_{ij}$ and $I(Round = 11)_{ij}$ are dummy variables (Round 9 is the reference category).
- $I(SC\ Group)_j$ is 1 if country *j* used self-completion in Round 10, and 0 otherwise.

- γ_{0j}, γ_{1j} and γ_{2j} denote country-specific random effects drawn from a multivariate normal distribution

$$\left(\frac{\gamma_{0j}}{\gamma_{1j}\,\gamma_{2j}}\right) \sim N(0, \Sigma)$$

where Σ is the covariance matrix capturing correlations among the intercept and slope deviations across countries.

Specification of Model 2

Model 2 extends Model 1 by adding an interaction between the ESS Round and country group, thereby permitting the effect of Round 10 or Round 11 to differ for self-completion vs. face-to-face countries:

$$\eta_{ij} = \beta_0 + b_{0j} + \left(\beta_1 + \gamma_{1j}\right) I(Round = 10)_{ij} + \left(\beta_2 + \gamma_{2j}\right) I(Round = 11)_{ij} + \beta_3 I(SC\ G$$

where all symbols remain as previously.