

# Switching a Face-to-Face Panel to Self-Administered Survey Modes: Experimental Evidence on Effects of Mode Assignment on Response and Selectivity

Jette Schröder<sup>1</sup>  · Claudia Schmiedeberg<sup>2</sup>  · Josef Brüderl<sup>2</sup>  ·  
Christiane Bozoyan<sup>2</sup> 

<sup>1</sup>GESIS - Leibniz Institute for the Social Sciences

<sup>2</sup>LMU Munich, Department of Sociology

Based on a mode experiment implemented in pairfam, a large, established German panel study, we investigate whether switching a panel from face-to-face to push-to-web survey mode leads to increased attrition and selectivity. We find that the redesign increases overall attrition by almost six percentage points, and causes even larger losses among full-time employed, self-employed, and less educated respondents. Two of the Big Five personality traits moderate the mode effect: conscientiousness and openness, while no differences are found for agreeableness, neuroticism, and extraversion. These results suggest that mode changes in a panel study bear risk for data quality in terms of sample size and selectivity.

*Keywords:* panel; mode; self-administered; face-to-face; response; bias; experiment

## 1 Introduction

Face-to-face surveys have long been the “gold standard” in survey data collection (Allum, Conrad and Wenz 2018; Mackeben and Sakshaug 2023; Watson and Wooden 2009). However, given the high costs (National Research Council 2013; Villar and Fitzgerald 2017; Wolf et al. 2021), declining response rates, and increasing difficulties in reaching subgroups of the population at home (Luiten, Hox and De Leeuw 2020; Williams and Brick 2018), as well as the spread of the Internet among the population and the growing technical capabilities of web surveys (Biffignandi and Bethlehem 2021; Eurostat 2022), web surveys are becoming an attractive complement or alternative to face-to-face

interviews (e.g., Allum et al. 2018; Atkeson, Adams and Alvarez 2014).

This is the case also for panel studies. Many panels, such as the U.S. Health and Retirement Study, Understanding Society in the U.K., the Canadian Labour Force Survey, and the U.S. National Longitudinal Study of Adolescent to Adult Health, have been experimenting with additional web mode or have already added web interviewing to their mixed-mode strategy (Biemer et al. 2022; Domingue et al. 2023; Francis and Laflamme 2015; Jäckle et al. 2017).<sup>1</sup> The German Family Panel pairfam (Brüderl et al. 2023a), on which we report in the present study, switched from face-to-face interviews to self-administered push-to-web mode combining web and mail in its 14<sup>th</sup> wave in 2021.

Including the online mode in the survey strategy may be appealing for panels (Bianchi, Biffignandi and Lynn 2017): In the first wave face-to-face interviewers introduce the panel and establish commitment to the survey so that respondents may be willing to participate in the self-adminis-

---

**Supplementary Information** The online version of this article (<https://doi.org/10.18148/srm/2025.v19i1.8427>) contains supplementary material.

---

Corresponding author: Jette Schröder, GESIS - Leibniz Institute for the Social Sciences, Mannheim, Germany (Email: [jette.schroeder@gesis.org](mailto:jette.schroeder@gesis.org))

---

<sup>1</sup> Although some examples of web-only panels can be found, sequential web/mail mixed-mode panels are more common as coverage error is an issue with web-only surveys and adding a mail option increases response rates and reduces selectivity (Bandilla, Couper and Kaczmirek 2014; Cornesse and Bosnjak 2018).

tered mode without further personal contact and persuasion in later waves (Jäckle, Lynn and Burton 2015). Moreover, information collected in prior waves may facilitate efficient and targeted mode strategies allowing, for instance, to invite respondents to their preferred survey mode (Al Baghal and Kelley 2016; Bianchi et al. 2017).

On the other hand, changes in survey mode may have undesirable effects. First, panel attrition may increase, as response rates are typically lower for self-administered surveys, in particular web surveys, compared to interviewer-administered modes (Daikeler, Bosnjak and Lozar Manfreda 2020; De Leeuw 2005; Klausch, Hox and Schouten 2015; Wagner et al. 2014), but evidence is mixed in this regard. For example, Wolf et al. (2021) report higher response rates in the self-administered mixed-mode (web, mail) than in the face-to-face mode in an experiment conducted in the European Values Survey. Second, selective participation may be an issue as different types of respondents may be more inclined to participate in a survey depending on the mode offered (Allum et al. 2018; Bretsch and Weiß 2022; Cornesse and Bosnjak 2018; Fitzgerald et al. 2019). Thus, adding a new survey mode or substituting survey modes may lead to an altered composition of the sample. Third, measurement effects may bias panel estimates, for example due to reduced social desirability effects or more satisficing in self-administered than interviewer-administered mode (Allum et al. 2018; Cernat and Revilla 2021; Cernat and Sakshaug 2022; Sakshaug, Beste and Trappmann 2023).

Accordingly, a thorough understanding of the effects of changes in survey mode on panel attrition, selectivity, and measurement is needed, as such a mode switch in an ongoing panel study may affect time-series estimates and comparability across waves. In the present study, we examine how such a switch from interviewer-administered to self-administered mode affects response rates and selectivity of participation. An examination of measurement effects is beyond the scope of this paper, as it would require a different analytical approach.

The study is based on data from the German Family Panel pairfam, a panel study of young and middle-aged individuals in Germany running since 2008 (Brüderl et al. 2023a). Due to institutional changes, the panel switched from annual CAPI (computer-assisted personal interview) to biannual self-administered mixed-mode waves in wave 14, with web mode as the primary mode and a paper-and-pencil questionnaire (PAPI) as an alternative sent with the second reminder (Brüderl et al. 2023b). To investigate the impact of this mode change on data quality, an experiment was implemented in this wave, in which panel members were randomly assigned to two groups: One group ( $N = 1200$ ) was interviewed in the usual CAPI mode, and the other group ( $N = 6226$ ) was interviewed

using the new push-to-web design. Both groups were interviewed with the same instrument of approximately 20 min duration.

The experimental design with random assignment of respondents to the CAPI and the mixed-mode group, respectively, allows for a causal interpretation of the difference in panel attrition between the two modes. In a second step, we examine whether the mode effect differs by respondent characteristics, based on information from previous panel waves: socio-demographic characteristics (education and employment status), but also personality traits (Big Five).

## 2 Theoretical Mechanisms

Different mechanisms are at work when a respondent is invited to a survey and decides to participate. According to Leverage-Salience-Theory (Groves, Cialdini and Couper 1992), participation depends on the survey design, the characteristics of the sampled person and the interviewer, and the interaction between respondent and interviewer. Respondents weigh the costs and benefits of participating versus refusing and follow compliance heuristics such as reciprocity, consistency, authority, social validation, scarcity, and liking (Groves et al. 1992). In a face-to-face survey, the interviewer contacts the respondent to make an appointment for the interview. To gain the respondent's cooperation, the interviewer may emphasize the authority of the research institute, create a situation of reciprocity, e.g., through doorstep incentives, and remind the respondent of previous participation, thus demanding consistency. In the absence of such interviewer-respondent interaction in self-administered modes, the reciprocity mechanism is reduced to incentives or brochures sent to the respondent, and the authority mechanism is activated only by the institution's logo on the invitation letter. In addition, self-administered surveys usually use only mail (postal or electronic) to invite respondents to participate, whereas interviewer-administered surveys use at least two forms of contact: the letter of invitation and the interviewer's phone call or visit to the respondent's home. Callbacks have been shown to reduce non-response, especially among hard-to-reach groups (Groves 2006). In contrast, a self-administered survey requires more action on the part of the respondent, and there is no alternative contact strategy if distracted or unconscious respondents overlook or fail to open the letter.

Several hypotheses about the effects of mode switching can be derived from Leverage-Salience-Theory. First, since the authority and reciprocity mechanisms are less salient in the self-administered mode, we expect lower participation rates. Basically, some respondents will be lost due to the loss of the motivational power of the interviewer.

Second, as self-administered modes require more self-discipline on the part of the respondents and as self-discipline varies amongst respondents, we can derive some moderation hypotheses. Especially, personality traits may be relevant here, in particular conscientiousness. We expect that less conscientious respondents will be less likely to participate in the self-administered mode. Moreover, face-to-face interviews are social activities involving conversations about personal matters, whereas self-administered surveys are solitary activities (Valentino et al. 2021). This may be particularly important in the case of face-to-face panel studies, where respondents are always contacted by the same interviewer, so that a relationship may have been established over the years (Hajek and Schumann 2018; Kühne 2018). Accordingly, differences in participation may arise as self-administered surveys may be more attractive to introverted individuals, while extraverted individuals may be more motivated to participate in an interviewer-administered survey (Valentino et al. 2021).

Moreover, in contrast to CAPI, where interviewers read questions to respondents and help them by explaining difficult concepts and questions, self-administered modes require sufficient language and reading skills (De Leeuw and Berzelak 2016; Heerwegh 2009; Tourangeau, Rips and Rasinski 2000). Accordingly, barriers to participation in a self-administered survey may exist for individuals with low levels of education. Consistent with this, there is evidence that individuals with lower levels of education participate less in self-administered than in interviewer-administered surveys (Roberts 2007; Wolf et al. 2021).

By contrast, self-administered surveys have the advantage of greater flexibility, as they can be completed when and where respondents prefer (Biffignandi and Bethlehem 2021). This reduces the perceived cost of participation, especially for individuals with time constraints, such as long working hours.

### 3 Data and methods

#### 3.1 Data

The mode experiment was implemented in wave 14 of the German Family Panel pairfam (Brüderl et al. 2023a), a panel study of a random sample of German residents from the birth cohorts 1971–1973, 1981–1983, 1991–1993, and 2001–2003, covering topics such as relationship quality, parenting, and intergenerational relationships (for details concerning survey methods, see Brüderl et al. 2023b; for a description of the study, see Huinink et al. 2011). Starting in 2008 with a sample size of approximately 12,000 respondents, an additional sample of Eastern German re-

spondents was added in wave 2 and a sample refreshment with approximately 5000 respondents in wave 11. Until wave 11, respondents were interviewed face-to-face every year, whereas in waves 12 and 13, due to contact restrictions during the COVID-19 pandemic, part of the sample was interviewed by CATI instead of CAPI.

In wave 14, the experimental group ( $N = 6226$ ) was invited to participate in a web survey instead of the usual CAPI survey via postal mail containing an unconditional incentive of €5 in cash. After two weeks, a reminder was sent to individuals who had not yet completed the web survey. After four weeks, the second reminder was sent together with a paper-and-pencil questionnaire (PAPI) as an alternative to the web survey. The fielding period started in October 2021. Anchor respondents who had not participated six weeks after the second reminder and whose telephone number was on file were contacted via telephone in a final attempt to motivate participation.

In contrast, the control group ( $N = 1200$ ) received a postal announcement letter also in October 2021 and was then contacted by their previous interviewer according to the usual procedure in the pairfam panel. Respondents in this group received a conditional incentive of €15 after completing the CAPI interview, as in previous waves. Respondents who had not been interviewed by February 2022 received the paper questionnaire in order to minimize attrition. These cases were coded as non-response for our analysis.

The gross CAPI subsample of 1200 respondents was randomly selected under two conditions. First, only respondents whose wave 13 interviewers were available to interview them in wave 14 were eligible for randomization to avoid confounding the effect of mode switching with the effect of interviewer switching. Second, after randomization, interviewers were required to have at least six addresses for face-to-face interviews in wave 14. All other interviewers' respondents were assigned to the CAWI group. (For more details on the experiment, see Brüderl et al. 2023b and Kantar Public 2022.)

#### 3.2 Analytical Approach

Our estimand is the average treatment effect (ATE) of switching the wave 14 pairfam interviews from face-to-face (CAPI) mode to self-administered push-to-web mode on the probability of response. As the data come from a randomized experiment, identification is given by comparing response rates between the treatment and control groups. However, given the imperfect randomization described above, treatment assignment may be partially selective. To address this issue, we perform a randomization check

and several robustness checks controlling for potential confounders.

We use linear probability models (LPM) on the dichotomous outcome variable: response (1) versus no response (0). We apply LPM instead of logit models because LPMs are easier to interpret than logit models, in particular when including interaction terms (Breen, Karlson and Holm 2018).

To analyze the consequences of a mode shift on the selectivity of the resulting sample, we analyze whether the ATE is causally affected by specific moderator variables (causal moderation). Our estimand is the (average) moderation effect. If there is a moderation effect, the sample composition after the mode switch will differ from the sample composition under the control condition, i.e. the mode switch will produce a selective sample. To perform moderation analysis, constitutive and multiplicative interaction terms must be included in the regression. However, moderation analysis requires strong identification assumptions, even when based on experimental data (VanderWeele 2015). As the moderator variables are not randomized, the focal moderation effect may be confounded by moderation effects of other variables that are correlated with the focal moderator. This requires controlling for potential confounders of the effect of the focal moderator on the outcome, including interaction terms of these confounders with the treatment variable. Confounders must be measured before treatment, which in our case is ensured by using information from previous panel waves.

To estimate moderation effects, we add the focal moderator variable to the LPM one at a time. Gender, birth cohort, and pairfam sample are included as controls in all models. To estimate the moderating effect of employment status, we additionally control for education because education may be a potential confounder for the moderating effect of employment. The five Big Five are entered simultaneously, so we are essentially estimating the direct moderating effect of each trait, netting out the moderating effects of the other four.

Note that the number of observations included in the models varies according to the number of missing values in the independent variables. Analyses were performed using Stata 18.0.

### 3.3 Measures

The binary outcome variable indicates response at wave 14 (completed interview vs. non-contact, refusal, and break-off). Ineligibility (e.g., respondent moved to an unknown location or died) was coded as 0 because different data collection modes differ in their ability to detect different types of ineligibility.

The treatment variable is survey mode, which indicates the mode assigned for wave 14 with two categories: self-administered mode (experimental group) and face-to-face mode (control group).

The moderator variables are education, employment status and the Big Five. For education, we use the ISCED classification (Schneider 2008) collapsed into four groups (lower secondary or less, upper secondary, post-secondary, tertiary). Employment status is a categorical variable with seven groups: homemaker, part-time employed, full-time employed, self-employed, unemployed, enrolled in education, other. Personality is captured by the Big Five traits neuroticism, extraversion, agreeableness, conscientiousness, and openness, measured with the short version of the Big Five Inventory, and additive indices are computed ranging from 1 (low) to 5 (high) (Rammstedt and John 2005; Thönnissen et al. 2023). As the Big Five are not assessed in every wave, data from wave 10 are used for long-term respondents and from wave 11 for the refreshment sample.

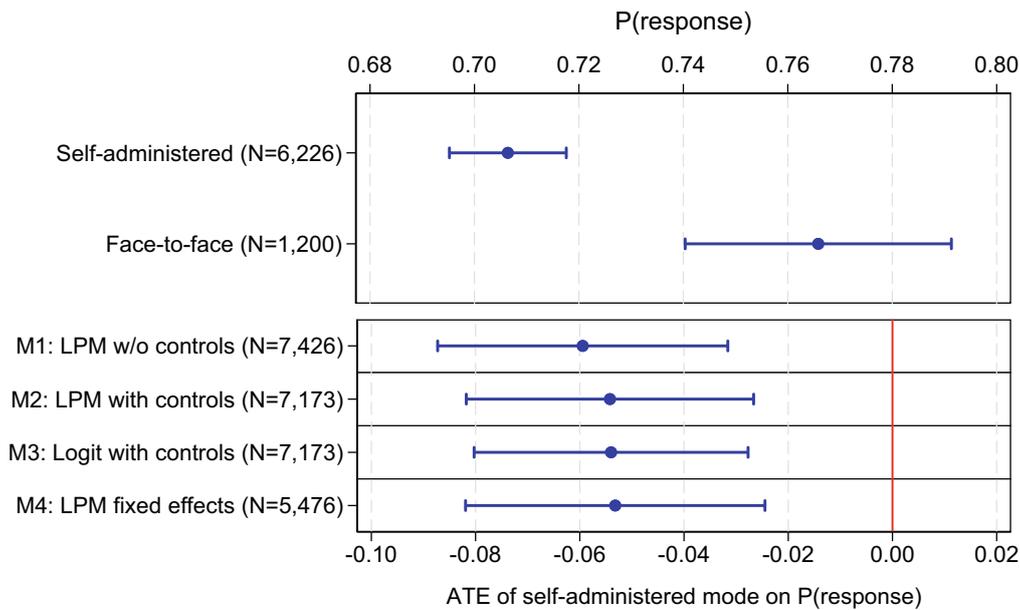
The control variables are gender, dummies for the pairfam birth cohorts, and dummies for the pairfam samples (pairfam main sample, East German oversample (DemoDiff) added in wave 3, refreshment sample added in wave 11).

Information on moderator and control variables were taken from the last panel wave a respondent participated in, i.e. wave 12 or 13. Descriptive statistics for all variables, as well as the number of missing values, can be found in Table A1 in the Appendix.

### 3.4 Randomization check

As the randomization of mode assignment was constrained, we checked the effectiveness of randomization by estimating a linear probability model with mode assignment as the dependent variable and all the variables used in the main analyses as independent variables. As Table A2 in the Appendix shows, the explanatory power of this regression is rather low:  $R^2$  is 0.006, meaning that the variables controlled in the regression explain less than 1% of the mode assignment. This means that the randomization worked well overall.

However, three of the independent variables are significantly associated with mode assignment. Assignment to the self-administered mode was more likely for respondents in the refreshment sample and for those higher in agreeableness or lower in conscientiousness. The higher likelihood for the refreshment sample is most likely due to the constraints in selecting interviewers for randomization. To avoid potential bias, we will control for the pairfam sample in the moderation analyses. In addition, we will perform robustness checks with all controls when analyzing the ATE.



**Fig. 1**

*The effect of switching to self-administered mode on the response probability Upper panel: Predicted probability of response (including 95% confidence interval [CI]) in the control (face-to-face) and treatment (self-administered) group (LPM without control variables, M1). Lower panel: Average treatment effects (ATE) and their 95% CIs estimated by four different models. M1: LPM without controls. M2: LPM with controls. M3: Logit with controls (the coefficient reported is the average marginal effect). M4: LPM with interviewer fixed effects (estimated only with those 128 interviewers that were part of the randomization). Numerical results can be found in Table A3 in the Appendix.*

#### 4 Results

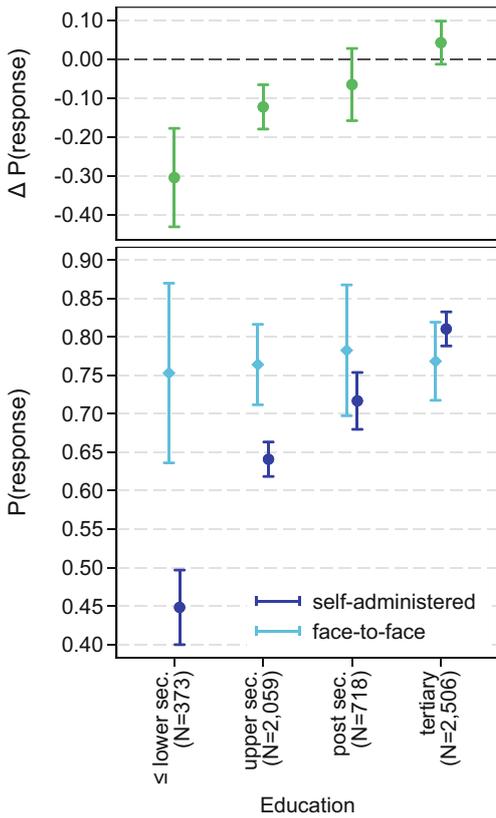
As Fig. 1 shows, the mode switch significantly and substantially reduces the response. The probability of response is 77% in the face-to-face group and 71% in the self-administered mode group (Fig. 1, upper panel). The ATE of switching to self-administered mode on response probability is  $-0.059$  (M1 in Fig. 1, lower panel), meaning that the probability of response is 5.9 percentage points lower in the experimental group.<sup>2</sup>

The lower panel of Fig. 1 additionally shows results from our robustness checks. First, we estimated a LPM with controls (M2). As expected, the resulting ATE is slightly lower ( $-0.054$ ). Second, we computed a logit model with controls (M3). After transforming the logit estimates to average marginal effects, results are virtually identical to those of the LPM. Third, as an alternative control strategy, we estimated a linear probability model with interviewer dummies. This implements an interviewer fixed effects approach

that essentially compares respondents assigned to the two modes within each interviewer. This allows for a within-interviewer estimation of the treatment effect, taking into account sampling point and interviewer effects. The resulting ATE is  $-0.053$  (M4). As including additional controls only marginally reduces the ATE, we continue with the simpler model without controls and add only control variables necessary for the unbiased estimation of the respective moderation (as explained above).

In the next step, we present the moderation analyses. Fig. 2 shows how the mode effect varies by respondent education. The bottom panel of Fig. 2 shows the predicted probabilities of response by mode and educational level. The top panel gives the resulting ATE for each educational group. We observe a strong interaction effect between mode and education. Among respondents with lower secondary education or less, the treatment effect of the self-administered mode is substantial ( $-31$  percentage points). The effect size diminishes with increasing educational levels and becomes positive, although not significant, for the group with tertiary education. The differences in the treatment effect between educational groups are significant in all but one case (the

<sup>2</sup> Due to rounding inaccuracies, the difference between the probabilities of response in the two mode groups is not 5.9, but 6.0.

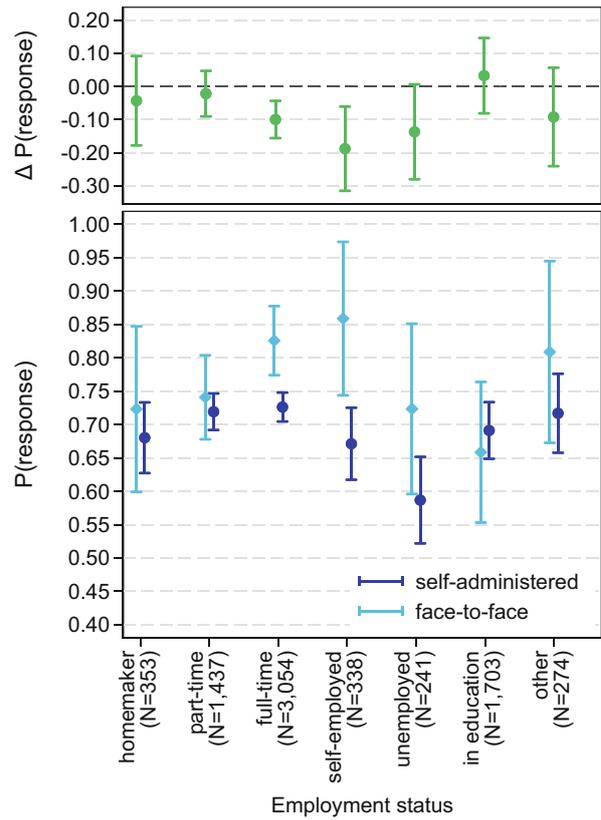


**Fig. 2**

*Mode effect on response by educational level (ISCED)  $N = 7408$ . Linear probability model; coefficients and 95% CI shown. Missing category “enrolled” (in education) not shown. Control variables: gender, birth cohort, and pairfam sample. Numerical results can be found in Table A4 in the Appendix.*

contrast between upper secondary and postsecondary education; results not shown, see replication file). Note that the main reason for the differences in treatment effect between groups is that response probability for the self-administered mode sharply declines with decreasing educational level. In contrast, face-to-face response does not significantly vary with educational level.

In terms of employment status (Fig. 3), the picture is more mixed, with a significant and negative ATE on response for two of the seven employment status groups: full-time employed (−10 percentage points) and self-employed (−19 percentage points). The unemployed also show a lower, but not significant, response in the self-administered mode (−14 percentage points,  $p = 0.061$ ). However, most of the contrasts are not significant, with the exception of the ATE for the self-employed, which is significantly different from those for the part-time employed and those in education, both of which are close to zero and insignificant

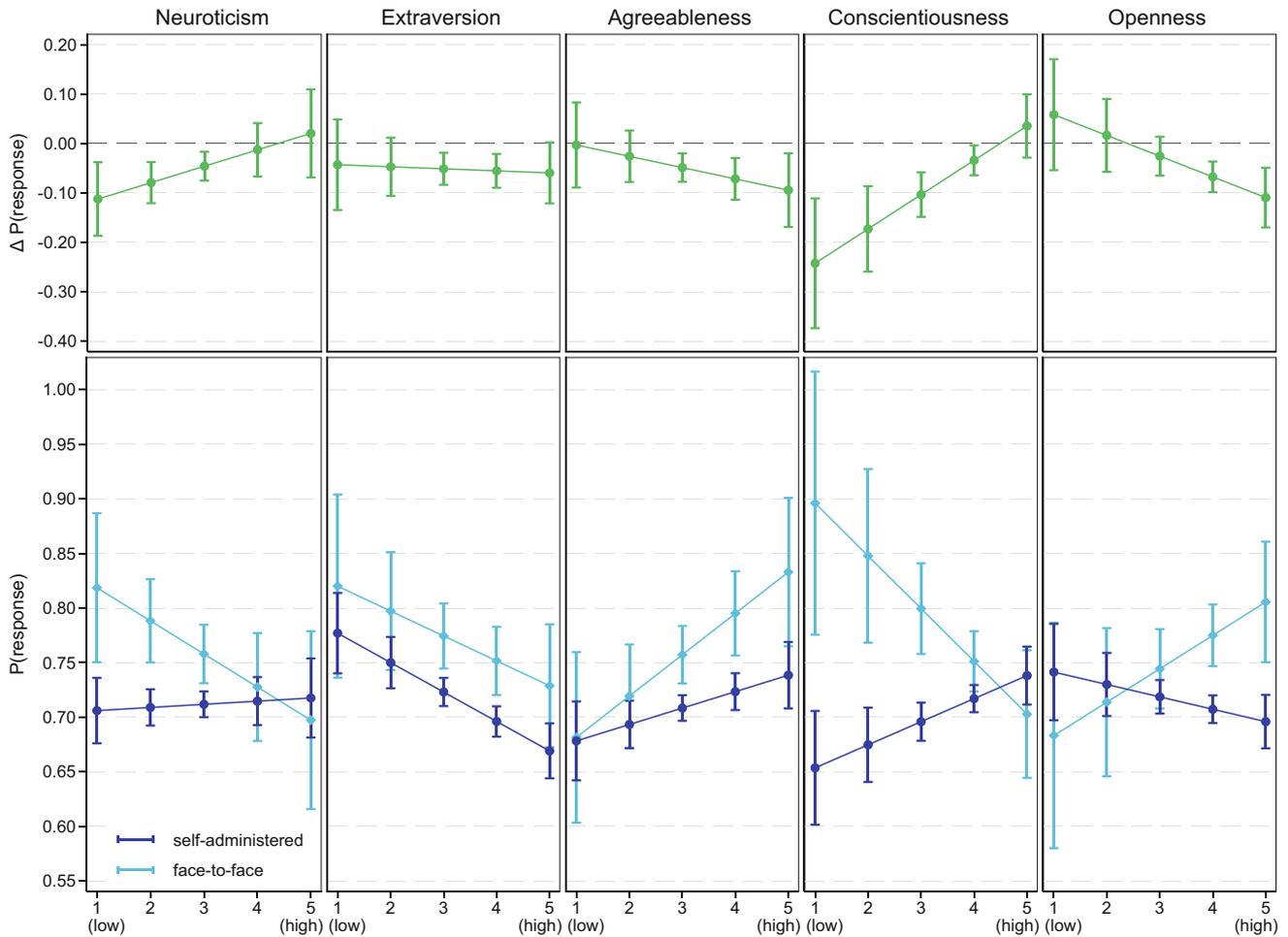


**Fig. 3**

*Mode effect on response by employment status  $N = 7400$ . Linear probability model; coefficients and 95% CI shown. Control variables: gender, birth cohort, pairfam sample, and educational level. Numerical results can be found in Table A5 in the Appendix.*

(results not shown, see replication file). Nevertheless, the results may provide some insight into the mechanisms at work. Above, we argued that the greater flexibility of the self-administered mode might increase response, especially among time-constrained groups. This is certainly not the case, as the strongly negative ATEs of the full-time employees and the self-employed show. The relevant mechanism seems to be the loss of motivational power of the interviewer: both groups have an exceptionally high likelihood of participating in the face-to-face mode. Interestingly, it is the unemployed who have the lowest response probability in the self-administered mode. This can also be interpreted as a consequence of the absence of the interviewer’s push.

Fig. 4 shows the results for the five Big Five. We include all five scales simultaneously in one model. In contrast to education and employment status, we estimate linear (rather than categorical) moderation effects. Among the Big Five personality traits, only conscientiousness and openness sig-



**Fig. 4**

*Mode effect on response by Big Five traits (neuroticism, extraversion, agreeableness, conscientiousness, openness)  $N = 7196$ . Linear probability model; coefficients and 95% CI shown. Control variables: gender, birth cohort, and pairfam sample. All five linear scales are entered simultaneously in one model. Numerical results can be found in Table A6 in the Appendix.*

nificantly moderate the mode effect (see the respective coefficients of the interaction terms in Table A6). The other three traits do not significantly alter the mode effect, i.e., the probability of responding in the self-administered mode is lower regardless of how neurotic, extraverted, or agreeable respondents are. These results are only partly in line with our expectations: Regarding extraversion, we have expected a larger negative ATE for extraverted than for introverted respondents assuming that they would be more likely to participate in a face-to-face interview. For conscientiousness, we find the expected large negative ATE for respondents with the lowest conscientiousness score ( $-24$  percentage points), in contrast to an insignificant ATE for respondents with the highest scores. However, this is only partly due to the fact that respondents with low conscientiousness are more likely to participate in the face-to-face mode than in the self-administered mode. The main factor is, unex-

pectedly, that individuals high in conscientiousness are less likely to participate face-to-face than those low in conscientiousness.

For openness, we observe opposite effects: The response probability in the self-administered group decreases slightly with openness, while it increases in the face-to-face group. As a result, the ATE is significant and substantial ( $-10$  percentage points) among highly open respondents, and insignificant among those with low openness scores.

## 5 Conclusion

This study examines the effects of switching a face-to-face panel to a self-administered mode on response rate and selectivity of response. The analysis draws on an experiment conducted in wave 14 of the German Family Panel pairfam,

in which a subset of respondents was surveyed in a self-administered mode instead of the usual face-to-face mode.

The response rate in the self-administered group is 5.9 percentage points lower than in the face-to-face group, indicating that switching to the self-administered survey mode results in significant panel attrition. Furthermore, the effect of the self-administered mode varies by respondent characteristics, leading to selective participation.

The negative effect of the self-administered mode on response is particularly pronounced for respondents with lower levels of education, consistent with the expectation that self-administered surveys act as a barrier to participation for the less educated due to the language and reading skills required and the lack of interviewer assistance. As the response of respondents with high levels of education is unaffected by mode, switching to the self-administered mode reduces the proportion of less educated respondents in the sample.

In terms of employment status, we observe a strong negative effect of the self-administered mode for full-time-employed, self-employed and unemployed (although not significant) individuals, while the response of homemakers, part-time employed, and students is not affected by survey mode. Again, this leads to an altered sample composition. These findings contradict the initial expectation that the flexibility offered by the self-administered mode would facilitate survey participation among respondents with severe time constraints.

Personality is found to play a role as well. Neuroticism, extraversion, and agreeableness do not moderate the effect of survey mode on response, but conscientiousness and openness do. Differences between face-to-face and self-administered mode are found for respondents with low levels of conscientiousness. While this difference is not unexpected, the pattern is somewhat surprising as the difference does not arise due to a particularly low response of less conscientious individuals in self-administered mode but their high response in face-to-face mode. Respondents with high scores on openness are more likely to participate when approached face-to-face than when approached in the self-administered mode. In contrast, no mode effect is found among respondents with low scores on openness.

Concerning the impact of transitioning a face-to-face panel to the self-administered mode on data quality, these results indicate substantial negative effects due to higher attrition and selection bias, especially concerning the educational bias of the sample. This raises the question of whether it is a good strategy to recruit a panel face-to-face and then switch to self-administered mode, as this approach might introduce a double bias. Both the existing literature and our analyses indicate that there are groups of individuals who are more likely to participate in self-administered surveys than face-to-face surveys and vice versa. Conse-

quently, by opting for face-to-face recruitment, individuals with low likelihood of face-to-face participation may be lost. And with the switch to the self-administered mode, respondents without the necessary skills and competencies for the self-administered survey might drop out. Therefore, it might be advisable to offer the self-administered mode as an alternative during face-to-face recruitment to avoid mode-specific selectivity during the recruitment process.

Our analysis may also be informative concerning cross-sectional self-administered surveys. It suggests that such surveys may not only suffer from bias in the distribution of standard socio-demographic characteristics, such as education, which can be easily detected by comparing the net sample with official statistics, but there may also be a systematic bias in variables typically not analyzed in such bias analyses such as personality.

This study does not come without limitations: Due to restrictions in the face-to-face fieldwork, the randomization of the experiment was not perfect, as shown by our randomization check. Moreover, as the experiment was implemented within a larger redesign of the panel, the treatment and control groups differed not only in survey mode but also in the incentive structure. Therefore, we cannot exclude the possibility that the different incentive structure (€5 prepaid compared to €15 postpaid) contributed to the lower response in the self-administered mode. Future research may overcome these shortcomings by implementing a pure mode experiment, keeping all other conditions constant.

**Acknowledgements** This paper uses data from the German Family Panel pairfam, coordinated by Josef Brüderl, Sonja Drobnič, Karsten Hank, Johannes Huinink, Bernhard Nauck, Franz J. Neyer, and Sabine Walper. From 2004 to 2022 pairfam was funded as priority program and long-term project by the German Research Foundation (DFG).

We thank the editor and reviewers for helpful comments. Katrin Auspurg provided valuable advice concerning our identification strategy. We received further valuable input from participants at the ESRA Conference (Milan, July 2023) and at the FReDA User Conference (Mannheim, October 2024).

## References

- Al Baghal, T., & Kelley, J. (2016). The Stability of Mode Preferences: Implications for Tailoring in Longitudinal Surveys. *methods, data, analyses*, 10(2), 1–24. <https://doi.org/10.12758/mda.2016.012>.
- Allum, N., Conrad, F., & Wenz, A. (2018). Consequences of mid-stream mode-switching in a panel survey. *Survey Research Methods*, 12(1), 43–58. <https://doi.org/10.18148/srm/2018.v12i1.6779>.
- Atkeson, L. R., Adams, A. N., & Alvarez, R. M. (2014). Nonresponse and mode effects in self- and in-

- interviewer-administered surveys. *Political Analysis*, 22(3), 304–320. <https://doi.org/10.1093/pan/mpt049>.
- Bandilla, W., Couper, M.P., & Kaczmirek, L. (2014). The effectiveness of mailed invitations for web surveys and the representativeness of mixed-mode versus internet-only samples. *Survey Practice*, 7(4), 1–9. <https://doi.org/10.29115/SP-2014-0020>.
- Bianchi, A., Biffignandi, S., & Lynn, P. (2017). Web-face-to-face mixed-mode design in a longitudinal survey: effects on participation rates, sample composition, and costs. *Journal of Official Statistics*, 33(2), 385–408. <https://doi.org/10.1515/JOS-2017-0019>.
- Biemer, P.P., Kathleen, M.H., Liao, B.J.B.D., & Halpern, C.T. (2022). Transitioning a panel survey from in-person to predominantly web data collection: Results and lessons learned. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(3), 798–821. <https://doi.org/10.1111/rssa.12750>.
- Biffignandi, S., & Bethlehem, J. (2021). *Handbook of web surveys* (2nd edn.). John Wiley & Sons.
- Breen, R., Karlson, K.B., & Holm, A. (2018). Interpreting and understanding Logits, Probits, and other non-linear probability models. *Annual Review of Sociology*, 44(1), 39–54. <https://doi.org/10.1146/annurev-soc-073117-041429>.
- Bretsch, D., & Weiß, B. (2022). How do internet-related characteristics affect whether members of a German mixed-mode panel switch from the mail to the web mode? *Social Science Computer Review*. <https://doi.org/10.1177/08944393221117267>.
- Brüderl, Josef, Sonja Drobnič, Karsten Hank, Franz. J. Neyer, Sabine Walper, Christof Wolf, Philipp Alt, Irina Bauer, Simon Böhm, Elisabeth Borschel, Christiane Bozoyan, Pablo Christmann, Rüdiger Edinger, Felicitas Eigenbrodt, Madison Garrett, Svenja Geissler, Tita Gonzalez Avilés, Nicolai Gröpler, Tobias Gummer, Kristin Hajek, Michel Herzig, Renate Lorenz, Katharina Lutz, Timo Peter, Richard Preetz, Julia Reim, Barbara Sawatzki, Claudia Schmiedeberg, Philipp Schütze, Nina Schumann, Carolin Thönnissen, Katharina Timmermann, and Martin Wetzel. 2023a. *The German Family Panel (pairfam)*. GESIS Data Archive, Cologne. ZA5678 Data file Version 14.2.0. <https://doi.org/10.4232/pairfam.5678.14.2.0>
- Brüderl, J., Schmiedeberg, C., Castiglioni, L., Arránz Becker, O., Buhr, P., Fuß, D., Ludwig, V., Schröder, J., & Schumann, N. (2023b). *The German Family Panel: Study design and cumulated field report (Waves 1 to 14)*. Munich. [https://www.pairfam.de/fileadmin/user\\_upload/redakteur/publis/Dokumentation/TechnicalPapers/pairfam\\_TP\\_01\\_2023.pdf](https://www.pairfam.de/fileadmin/user_upload/redakteur/publis/Dokumentation/TechnicalPapers/pairfam_TP_01_2023.pdf)
- Cernat, A., & Revilla, M. (2021). Moving from face-to-face to a web panel: impacts on measurement quality. *Journal of Survey Statistics and Methodology*, 9(4), 745–763. <https://doi.org/10.1093/jssam/smaa007>.
- Cernat, A., & Sakshaug, J.W. (2022). The impact of mixing survey modes on estimates of change: a quasi-experimental study. *Journal of Survey Statistics and Methodology*. <https://doi.org/10.1093/jssam/smac034>.
- Cornesse, C., & Bosnjak, M. (2018). Is there an association between survey characteristics and representativeness? A meta-analysis. *Survey Research Methods*, 12(1), 1–13. <https://doi.org/10.18148/srm/2018.v12i1.7205>.
- Daikeler, J., Bosnjak, M., & Manfreda Lozar, K. (2020). Web versus other survey modes: an updated and extended meta-analysis comparing response rates. *Journal of Survey Statistics and Methodology*, 8(3), 513–539. <https://doi.org/10.1093/jssam/szm008>.
- De Leeuw, E.D. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, 21(2), 233–255.
- De Leeuw, E.D., & Berzelak, N. (2016). Survey mode or survey modes? In C. Wolf, D. Joye, T.E.C. Smith & Y. Fu (Eds.), *The SAGE handbook of survey methodology* (pp. 412–156). London, Thousand Oaks, New Delhi, Singapore: SAGE.
- Domingue, B.W., McCammon, R.J., West, B.T., Langa, K.M., Weir, D.R., & Faul, J. (2023). The mode effect of web-based surveying on the 2018 U.S. Health and retirement study measure of cognitive functioning. *Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 78(9), 1466–1473. <https://doi.org/10.1093/geronb/gbad068>.
- Eurostat (2022). Digital economy and society statistics—households and individuals. [https://ec.europa.eu/eurostat/statistics-explained/images/3/35/Internet\\_access\\_of\\_households%2C\\_2017\\_and\\_2022\\_%28%25\\_of\\_all\\_households%29\\_14-12-2022.png](https://ec.europa.eu/eurostat/statistics-explained/images/3/35/Internet_access_of_households%2C_2017_and_2022_%28%25_of_all_households%29_14-12-2022.png). Accessed 23 Oct 2023.
- Fitzgerald, D., Hockey, R., Jones, M., Mishra, G., Waller, M., & Dobson, A. (2019). Use of online or paper surveys by Australian women: longitudinal study of users, devices, and cohort retention. *Journal of Medical Internet Research*, 21(3), e10672. <https://doi.org/10.2196/10672>.
- Francis, J., & Laflamme, G. (2015). Evaluating web data collection in the Canadian Labour Force Survey. Washington, DC. [https://nces.ed.gov/fcsm/pdf/h2\\_francis\\_2015fcsm.pdf](https://nces.ed.gov/fcsm/pdf/h2_francis_2015fcsm.pdf). Accessed 20 Oct 2023.

- Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70(5), 646–675. <https://doi.org/10.1093/poq/nfl033>.
- Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56(4), 475. <https://doi.org/10.1086/269338>.
- Hajek, K., & Schumann, N. (2018). Continuity trumps? The impact of interviewer change on item nonresponse. *Survey Research Methods*, 12(3), 211–230. <https://doi.org/10.18148/SRM/2018.V12I3.7230>.
- Heerwegh, D. (2009). Mode differences between face-to-face and web surveys: an experimental investigation of data quality and social desirability effects. *International Journal of Public Opinion Research*, 21(1), 111–121. <https://doi.org/10.1093/ijpor/edn054>.
- Huinink, J., Brüderl, J., Nauck, B., Walper, S., Castiglioni, L., & Feldhaus, M. (2011). Panel analysis of intimate relationships and family dynamics (pairfam): Conceptual framework and design. *Zeitschrift für Familienforschung*, 23(1), 77–101. <https://nbn-resolving.org/urn:nbn:de:0168-ssaar-376463>
- Jäckle, A., Lynn, P., & Burton, J. (2015). Going online with a face-to-face household panel: effects of a mixed mode design on item and unit non-response. *Survey Research Methods*, 9(1), 57–70. <https://doi.org/10.18148/srm/2015.v9i1.5475>.
- Jäckle, A., Tarek Al Baghal, A. G., Burton, J., & Lynn, P. (2017). Understanding society: the UK household longitudinal study innovation panel, waves 1–9, user manual. [https://www.understandingsociety.ac.uk/sites/default/files/downloads/documentation/innovation-panel/user-guides/6849\\_ip\\_waves1-9\\_user\\_manual\\_June\\_2017.pdf](https://www.understandingsociety.ac.uk/sites/default/files/downloads/documentation/innovation-panel/user-guides/6849_ip_waves1-9_user_manual_June_2017.pdf). Accessed 20 Oct 2023.
- Kantar Public (2022). *Beziehungen und Familienleben in Deutschland (pairfam). Methodenbericht Welle 14 (2021 / 2022)*. Munich.
- Klausch, T., Hox, J., & Schouten, B. (2015). Selection error in single- and mixed-mode surveys of the Dutch general population. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(4), 945–961. <https://doi.org/10.1111/rssa.12102>.
- Kühne, S. (2018). From strangers to acquaintances? Interviewer continuity and socially desirable responses in panel surveys. *Survey Research Methods*, 12(2), 121–146. <https://doi.org/10.18148/SRM/2018.V12I2.7299>.
- Luiten, A., Hox, J., & De Leeuw, E. D. (2020). Survey non-response trends and fieldwork effort in the 21st century: results of an international study across countries and surveys. *Journal of Official Statistics*, 36(3), 469–487. <https://doi.org/10.2478/JOS-2020-0025>.
- Mackeben, J., & Sakshaug, J. W. (2023). Transitioning an employee panel survey from telephone to online and mixed-mode data collection. *Statistical Journal of the IAOS*, 39(1), 213–232. <https://doi.org/10.3233/SJI-220088>.
- National Research Council (2013). *Nonresponse in social science surveys*. Washington, D.C.: National Academies Press.
- Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K). *Diagnostica*, 51(4), 195–206. <https://doi.org/10.1026/0012-1924.51.4.195>.
- Roberts, C. (2007). *Mixing modes of data collection in surveys: a methodological review*. ESRC National Centre for Research Methods, NCRM Methods Review Paper 008.
- Sakshaug, J. W., Beste, J., & Trappmann, M. (2023). Effects of mixing modes on nonresponse and measurement error in an economic panel survey. *Journal for Labour Market Research*. <https://doi.org/10.1186/s12651-022-00328-1>.
- Schneider, S. L. (2008). Applying the ISCED-97 to the German educational qualifications. In S. L. Schneider (Ed.), *The International Standard Classification of Education (ISCED-97): An evaluation of content and criterion validity for 15 European countries* (pp. 76–102). Mannheim: MZES.
- Thönnissen, C., Reim, J., Geissler, S., Alt, P., Sawatzki, B., Böhm, S., & Walper, S. (2023). *Pairfam scales and instruments manual, release 14.0*. Munich: LMU Munich.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.
- Valentino, N. A., Zhirkov, K., Hillygus, D. S., & Guay, B. (2021). The consequences of personality biases in online panels for measuring public opinion. *Public Opinion Quarterly*, 84(2), 446–468. <https://doi.org/10.1093/poq/nfaa026>.
- VanderWeele, T. J. (2015). *Explanation in causal inference: methods for mediation and interaction*. New York: Oxford University Press.
- Villar, A., & Fitzgerald, R. (2017). Using mixed modes in survey research: Evidence from six experiments in the ESS. In M. J. Breen (Ed.), *Values and identities in Europe* (pp. 299–336). Routledge.
- Wagner, J., Arrieta, J., Guyer, H., & Ofstedal, M. B. (2014). Does sequence matter in multi-mode surveys: results from an experiment. *Field Methods*, 26(2), 141–155. <https://doi.org/10.1177/1525822X13491863>.

- Watson, N., & Wooden, M. (2009). Identifying factors affecting longitudinal survey response. In P. Lynn (Ed.), *Wiley series in survey methodology, Methodology of longitudinal surveys* (pp. 157–181). Chichester: John Wiley.
- Williams, D., & Brick, J.M. (2018). Trends in U.S. face-to-face household survey nonresponse and level of effort. *Journal of Survey Statistics and Methodology*, 6(2), 186–211. <https://doi.org/10.1093/jssam/smx019>.
- Wolf, C., Christmann, P., Gummer, T., Schnaudt, C., & Verhoeven, S. (2021). Conducting general social surveys as self-administered mixed-mode surveys. *Public Opinion Quarterly*, 85(2), 623–648. <https://doi.org/10.1093/poq/nfab039>.