CONDUCTING GENERAL SOCIAL SURVEYS AS SELF-ADMINISTERED MIXED-MODE SURVEYS

CHRISTOF WOLF*
PABLO CHRISTMANN
TOBIAS GUMMER
CHRISTIAN SCHNAUDT
SASCHA VERHOEVEN

Abstract   This article discusses the current challenges of conducting a General Social Survey (GSS) in face-to-face mode and evaluates the alternative of fielding these surveys in self-administered mixed-mode (web, mail) instead. Based on data from Germany, it first illustrates the stark decline of participation in face-to-face surveys since 2002 and reports a strong increase in the cost of conducting these surveys over the same period. It then discusses the possibility of implementing GSS-type surveys in a self-administered design and reports results from a mode experiment implemented in the German part of the European Values Study (EVS) 2017/18. The results of the experiment indicate that self-administered mixed-mode surveys are a viable alternative for cross-sectional general population surveys in Germany; they shorten the fieldwork period and lead to higher response rates, while being more cost-efficient than face-to-face surveys. Despite the finding that the sample composition deviates from the general population in both modes, the face-to-face mode represents the population slightly better.

Christof Wolf is president of GESIS—Leibniz Institute for the Social Sciences, Mannheim, Germany, and holds the chair of Social Stratification at the University of Mannheim, Mannheim, Germany. Pablo Christmann is a senior researcher at GESIS—Leibniz Institute for the Social Sciences, Mannheim, Germany. Tobias Gummer is a team leader and senior researcher at GESIS—Leibniz Institute for the Social Sciences, Mannheim, Germany, and a private lecturer at the University of Mannheim. Christian Schnaudt is a research associate at the University of Mannheim, Mannheim, Germany. Sascha Verhoeven is Census 2022 project manager at the City of Heidelberg, Office of City Development and Statistics, Heidelberg, Germany. *Address correspondence to Christof Wolf, GESIS—Leibniz Institute for the Social Sciences, B6, 4-5, 68159 Mannheim, Germany; email: christof.wolf@gesis.org.

Introduction

General Social Surveys (GSS), that is, multi-thematic cross-sectional surveys of the general population, are a core part of the social science data infrastructure in many countries. The GSS[1] of the United States (est. 1972) is the oldest example of this type of survey and the model for all that followed, including the German ALLBUS[2] (est. 1980), the British Social Attitudes Survey (est. 1983), and the Swiss MOSAiCH[3] (est. 2005). In addition to these national surveys, several international survey programs with similar design, content, and target groups have been established over the last four decades, most notably the European Values Study (EVS), the World Values Survey (WVS), the International Social Survey Programme (ISSP), the European Social Survey (ESS), and the East Asian Social Survey (EASS). Furthermore, various barometer surveys have been initiated that focus on social, political, and economic topics: the Eurobarometer, the Latinobarómetro, the Afrobarometer, the Asian Barometer, the AmericasBarometer, the Arab Barometer, and the Eurasia Barometer.

In recent years, GSS-type surveys have come under pressure because of decreasing response rates (Brick and Williams 2013; de Leeuw, Hox, and Luiten 2018). Response rates of less than 50 percent are quite common in many European countries today, and in some countries, such as Germany, they dropped to 30 percent or below, despite intensified fieldwork efforts to counter this trend (Beullens et al. 2018). This development raises questions concerning the quality of the data, as it might increase the *risk* of nonresponse bias.[4] In addition, "the escalating costs of face-to-face interviewing" (O'Muircheartaigh 2018, p. 15) raise the question if the investment is worthwhile.

Against this background, we explore alternative methods of collecting GSS-type data and assess how they compare in terms of costs, response rates, and further quality indicators. We use an experimental design implemented in the German part of the 2017/18 EVS to investigate the viability of conducting general population surveys as self-administered mixed-mode surveys (web, mail) based on probability samples. To this end, we contrast a face-to-face survey with (1) a mixed-mode survey with a split questionnaire, designed to reduce the overall response burden (Raghunathan and Grizzle

---

1. See https://gss.norc.org/ (retrieved July 27, 2020).

2. Allgemeine Bevölkerungsumfrage der Sozialwissenschaften. See https://www.gesis.org/en/allbus/allbus-home (retrieved July 27, 2020).

3. Measurement and Observation of Social Attitudes in Switzerland (CH). See https://forscenter.ch/projects/mosaich/ (retrieved July 27, 2020).

4. Though Groves (2006) and Groves and Peytcheva (2008) have shown that there is no strong statistical relationship between response rate and nonresponse bias, "the phenomenon of declining responses is sufficiently widespread that it has generated growing concerns about the potential for nonresponse bias" (Czajka and Beyler 2016, p. vii).

1995; Peytchev and Peytcheva 2017), and (2) a mixed-mode survey covering the full-length EVS questionnaire. This strategy provides innovative insights regarding the suitability of self-administered mixed-mode survey designs as alternatives to "traditional" face-to-face GSS-type surveys.

Our study is structured as follows. First, we discuss the main challenges facing GSS-type surveys and provide a brief overview of the development of response rates, refusal rates, non-contact rates, and fieldwork duration for the US GSS, the ALLBUS, and the German part of the ESS. Second, we illustrate how this development has impacted the costs of conducting face-to-face surveys in Germany during the last two decades, using the ALLBUS and the German part of the ESS as examples. Third, we discuss several mixed-mode designs as possible alternatives. Fourth, we compare data collection efforts of probabilistic face-to-face and mixed-mode (web, mail) surveys conducted as part of the EVS 2017/18 in Germany. Our focus here is on fieldwork, survey costs, and data quality. Finally, we evaluate the viability of self-administered mixed-mode surveys to conduct GSS-type surveys.

Challenges of the Current Model

The current operational model of GSS-type surveys—large nationwide surveys (whether of specific national scope or as part of a cross-national survey program) covering a broad range of topics conducted as face-to-face interviews—is under pressure because of two partly related developments: dwindling response rates resulting in increased fieldwork efforts and rising costs (Tourangeau 2017), both of which we discuss in turn.

DEVELOPMENT OF RESPONSE RATES

Declining survey response rates have been debated for several decades. One of the earliest systematic analyses of nonresponse trends looked at changes in response rates from the American National Election Studies and the Survey of Consumer Attitudes from 1952 to 1979 and concluded that there has been a substantial increase in nonresponse, which was mainly driven by increased refusal rates (Steeh 1981). In his 1992 address "A response to the nonresponse problem" as president of the AAPOR, Norman Bradburn noted that "We all believe strongly that response rates are declining and have been declining for some time" (Bradburn 1992, p. 392). Though Smith (1995) questions a universal trend of increasing nonresponse rates from the 1950s through the early 1990s, Brick and Williams (2013, p. 52) conclude "that nonresponse rates in U.S. household surveys have been increasing over time."

The problem of declining response rates is limited neither to the United States nor to academic surveys. Analyzing changes in response rates of 48 surveys from official statistics in 16 different countries, de Leeuw and de Heer

(2002) conclude that response rates are declining internationally due to growth in both non-contact rates and refusal rates. In a recent follow-up, de Leeuw, Hox, and Luiten (2018) provide evidence for a prolonged continuation of these trends in the twenty-first century as well. Czajka and Beyler (2016) report similar results for U.S. Health and Human Services-sponsored surveys, for which they find consistent and accelerating declines in response rates.

While our study concentrates on face-to-face surveys, it is worth mentioning that strongly declining response rates have also been reported for RDD phone surveys using dual frame samples (Dutwin and Lavrakas 2016) and mail surveys from consumer research, at least those with national scope (Allen et al. 1997). In addition, academic mail surveys seem to struggle with falling response rates as well. For example, the Australian Election Study (AES) reported a drop in response rates from 62.8 percent in 1987 to 55.4 percent in 2001 to as low as 22.5 percent in 2016. Yet, a recent shift to a push-to-web strategy led the AES response rate to recover to 42.1 percent in 2019.[5]

As can be seen from table 1, response rates for different GSSs have been declining in the last two decades, though, at different paces and to different levels. While the decrease for the GSS in the United States occurred recently and was comparatively small, German response rates were much lower and have declined more steadily over the last 16 years.[6] ALLBUS 2002 reports a response rate of 47 percent, which has steadily declined since and currently amounts to just over 30 percent. A similar but even steeper decline of response rates occurred in the German part of the ESS, for which a drop from over 50 percent to under 30 percent is evident in the same period.[7] This decline in response rates is mainly due to respondents' increasing tendency to refuse to participate in social surveys. As the results in table 1 indicate, refusal rates for the US GSS as well as the German ALLBUS increased by almost 10 percentage points between 2002 and 2018. For the German part of the ESS, the refusal rate increased by more than 20 percentage points in the same period.

DEVELOPMENT OF COSTS

To counteract these trends, efforts have been intensified to reach the "last respondent" (Stoop 2005). For example, fieldwork monitoring has become stricter, the number of contact attempts to reach non-contacts was increased, as were interviewer remuneration and monetary incentives given to respondents. Costs for conducting such surveys have increased as a result. While face-to-face interviews were always more expensive than other survey

---

5. See https://australianelectionstudy.org/voter-studies (retrieved July 22, 2020).

6. It should be noted that other US surveys have suffered more from declining response rates (Czajka and Beyler 2016, p. A.7; Tourangeau 2017).

7. Though the German case is particularly pronounced, research has shown that response rates of the combined ESS have fallen over the first seven rounds (Beullens et al. 2018).

Table 1. Response rates, refusal rates, non-contact rates, and fieldwork duration for three social surveys, 2002 to 2018

Year	ALLBUS				ESS Germany				US-GSS			
	Response rate	Refusal rate	Non-contact rate	Field-work in days	Response rate	Refusal rate	Non-contact rate	Field-work in days	Response rate	Refusal rate	Non-contact rate	Field-work in days
2002	47.3	41.5	4.8	180	51.7	28.2	5.7	178	70.1	26.1	1.5	141
2004	45.7	42.4	6.7	134	51.0	32.8	7.0	144	70.4	22.5	2.4	140
2006	41.0	45.8	6.4	157	52.9	25.4	5.0	137	71.2	23.3	1.1	154
2008	40.3	48.2	6.6	176	42.7	32.6	6.4	158	70.4	24.1	1.2	150
2010	34.4	53.4	6.5	155	29.7	39.6	7.4	142	70.3	24.5	1.8	151
2012	37.6	49.2	6.5	161	33.7	45.9	4.7	139	71.4	21.0	2.3	170
2014	35.0	49.7	7.2	174	31.4	47.9	7.6	172	69.2	26.4	1.2	195
2016	34.9	47.0	10.3	167	30.6	48.0	3.3	216	61.3	32.7	2.0	229
2018	32.3	49.3	10.3	171	27.6	50.4	1.8	188	59.5	36.0	1.6	213

NOTE.—Data were kindly provided by the respective teams; we thank Michael Blohm (ALLBUS) and René Bautista (US-GSS). Data for the ESS is based on Matsuo and Loosveldt (2013) for ESS 2002–2010, Beullens et al. (2014) for ESS 2012, Beullens and Loosveldt (2016) for ESS 2014, Wuyts and Loosveldt (2019) for ESS 2016, as well as own calculations for ESS 2018.

modes, this development has increased the relative cost disadvantage of this mode even further. For example, prices for the German General Social Survey ALLBUS and the German part of the ESS nearly doubled between 2002 and 2018 (see figure 1), whereas consumer prices only increased by 25 percent during this period.

Two factors have mainly contributed to this development. As mentioned above, declining response rates have rendered fieldwork more difficult and work-intensive, requiring more thorough and detailed oversight of interviewer staff and more possibilities to directly intervene during fieldwork. Additionally, demands for improved surveys by (academic) clients have increased, such as minimizing interviewer effects, timely availability of datasets, and more transparent and comprehensive data documentation. Both developments complicate the process for survey agencies and lead to additional costs. Considering the enormous rise in costs for face-to-face surveys, organizations that regularly commission surveys are looking for alternative modes that are cheaper, yet ideally yield the same data quality as face-to-face surveys. Possible alternatives are surveys in self-completion modes, administered either as paper-based postal or as web survey.

Mixed-Mode Self-Administered Surveys as an Alternative to Face-to-Face Surveys

Self-administered web surveys have become a standard survey mode since the advent of the internet.[8] In contrast to market research, however, academic research was slower to take up the new mode. Flagship surveys of the GSS type were particularly reluctant to adopt this new survey mode. Two reasons primarily shaped this reluctance: ensuring representativeness and interview length—problems to which solutions exist today.

First, for GSS-type surveys, it is critical to capture a random sample that covers the general population well and thus produces as little bias as possible (with given resources). For a long time, it was unclear how random samples could be recruited into web surveys because the population is not evenly covered and there is no register of internet users that could be used as a sampling frame. By combining web surveys with mail invitations, the German Internet Panel (Blom, Gathmann, and Krieger 2015), the GESIS Panel (Bosnjak et al. 2018), and the Swiss MOSAiCH[9] have developed models to overcome these problems.

---

8. In 2018, 40 percent of all interviews in quantitative studies in market research in Germany were conducted online, 28 percent over the phone, and 23 percent face-to-face. See https://www.adm-ev.de/die-branche/mafo-zahlen (retrieved July 29, 2020).

9. Since 2018, the Swiss MOSAiCH has switched from a biennial face-to-face survey to an annual mixed-mode survey (web, mail) with two to three panel waves each; see https://forscenter.ch/projects/mosaich/ (retrieved July 29, 2020).
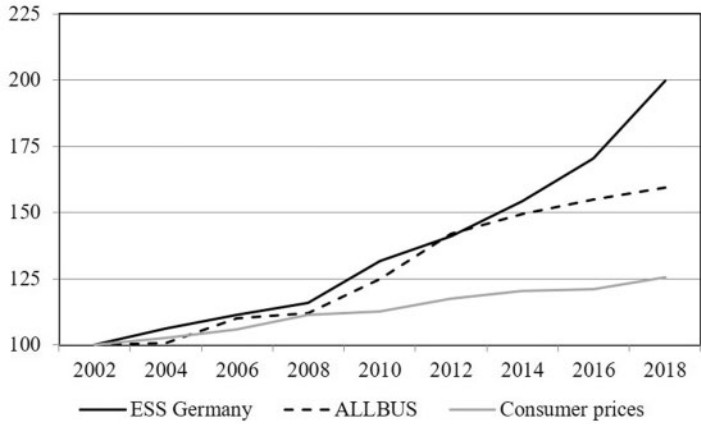
Figure 1. Costs for ALLBUS and ESS Germany as well as consumer prices, 2002 = 100. Price estimates are based on offers from field institutes to carry out the respective survey. Consumer price indices provided by German Federal Statistical Office (source: destatis, https://www.destatis.de/DE/ZahlenFakten/GesamtwirtschaftUmwelt/Preise/Verbraucherpreisindizes/Tabellen_/VerbraucherpreiseKategorien.html?cms_gtp=145110_slot%253D2, retrieved July 27, 2020).

Second, whereas the typical length of GSS-type surveys is 60 minutes or more, recommendations suggest that online surveys not exceed 20 minutes (Callegaro, Lozar Manfreda, and Vehovar 2015, p. 101). At least two solutions seem promising to this problem. One option is to choose a panel design and split the questionnaire content across waves asking respondents to answer a second or even third round of questions after completing the initial survey. By inviting respondents to multiple surveys, total survey length can be increased, while keeping individual surveys short. The downside of this solution is that it increases fieldwork costs and leads to many cases with incomplete data due to panel attrition.

Another possibility is to use a split-questionnaire or matrix design in a cross-sectional survey. In this design, a questionnaire is broken down into several modules that then are distributed over different parts of a sample. Despite being well known in survey-based assessments (Childs and Jaciw 2002), this design has not been widely used for GSS-type surveys. While the matrix design avoids issues related to panel attrition, it also introduces systematic gaps into the dataset that need to be dealt with afterwards, for example through Bayesian multiple imputation (Adiguzel and Wedel 2008). Yet, the issue of imputing large GSS-type surveys is a research challenge on its own.

Despite the conventional wisdom to keep self-administered (web) surveys short, mode experiments implemented in the ESS in Estonia indicate that it is feasible to have respondents answer an hour-long online questionnaire (Villar and Fitzgerald 2017, p. 272), and that the measurement quality of (selected) items and composite scores hardly change in comparison with the face-to-face mode (Revilla 2010, 2015).[10] Yet, these findings are based on data collected roughly a decade ago—a time in which the abovementioned problems associated with conducting face-to-face surveys (i.e., decreasing response rates and increasing costs) were clearly less severe than today. What is more, the ESS findings emanate from a relatively small number of three countries (the Netherlands, Slovenia, and the UK), requiring further investigation of their generalizability to different countries and survey contexts (Revilla 2015, p. 1222).

Against this backdrop, this study complements and expands upon previous findings on the feasibility of mixed-mode designs in at least four distinct ways: First, by relying on data collected during the most recent wave of the EVS in 2017/18, it provides an up-to-date analysis that takes into account the aggravated survey climate as well as the current problems and obstacles in conducting face-to-face surveys as evident in many countries. Second, by focusing on Germany as a country that is particularly hampered by both decreasing response rates and increasing costs of face-to-face surveys, it offers much-needed evidence on an empirical case for which mixed-mode surveys might establish a more promising alternative than for countries still exhibiting comparatively high response rates in face-to-face surveys, such as the Netherlands, Slovenia, or the UK in the ESS experiments. Third, by comparing a face-to-face survey with a mixed-mode (web, mail) survey in a matrix design as well as a full-length mixed-mode survey, it provides novel insights on the feasibility of conducting mixed-mode surveys as a general alternative to face-to-face surveys, as well as on the feasibility of different mixed-mode designs (i.e., matrix versus full-length) in GSS-type studies such as the ESS or the EVS. Fourth, in comparison to previous studies investigating the feasibility of mixed-mode surveys (cf. Villar and Fitzgerald 2017, pp. 282–87), it presents a more detailed comparison and analysis of the survey costs resulting from different survey-mode choices.

---

10. Between 2003 and 2012, the ESS conducted six (mixed-)mode studies, comparing combinations of telephone and web surveys with the face-to-face mode in selected European countries (for an overview, see Villar and Fitzgerald 2017, pp. 262–64). Two of these studies, similar to our study, specifically investigated the feasibility of mixed-mode designs as an alternative to face-to-face surveys (for details, see Revilla 2010, 2015).

Mixed-Mode Experiments in the European Values Study

The EVS follows a repeated cross-national survey design aimed at capturing the general population's views on topics like the environment, national identity, family, politics, society, religion, morale, and work. The EVS questionnaire is traditionally administered as a face-to-face interview (CAPI and/or PAPI). In 2017/18, a group of countries (i.e., Denmark, Finland, Germany, Iceland, the Netherlands, Switzerland) decided to implement additional self-administered modes (mail and/or web).[11] In their overview of the recent EVS surveys, Luijkx et al. (2021) highlight that each country adapted the design of the self-administered survey to their country-specific contexts based on prior experiences and expectations. Thus, implementations differed between countries with respect to whether and when the mail mode was offered, how incentives were used, which questionnaires were provided in self-administered modes (see below), and whether the country experimentally tested different ways of implementing the surveys. The analyses in this study are based on the data from Germany that included extensive experiments on survey modes regarding mode choice sequence, incentives, and questionnaire length. All analyses presented here are unweighted.

All survey data collected for the EVS 2017/18 in Germany is based on the same two-stage random sample of persons drawn from the residential registers of 162 sample points in Germany. To allow for an experimental comparison of data collection modes, we randomly split the addresses of the gross sample into three experimental groups. Three surveys were conducted based on these samples (see figure 2). First, a face-to-face survey was conducted that featured the full EVS questionnaire with an average length of 59 minutes. The face-to-face fieldwork was organized in two phases, to which the gross sample was randomly allocated before fieldwork began. In the first phase, a postpaid incentive of 10 euros (conditional on participation in the survey) was announced in the advance letter to respondents. After an evaluation in early January 2018, the initial postpaid incentive was increased to 20 euros for those respondents who had refused to participate so far (soft refusals) or for those who had not been reached. For the second phase of the face-to-face fieldwork (starting in mid-January 2018), the initial postpaid incentive was replaced by a 5 euro unconditional prepaid incentive.

Second, a mixed-mode survey was conducted with a split questionnaire design[12] to reduce the overall response burden (Raghunathan and Grizzle 1995;

---

11. Data from the EVS mixed-mode surveys are freely available from the GESIS data archive; see doi:10.4232/1.13314.

12. The EVS questionnaire was divided into six different split questionnaires for this purpose. Each split questionnaire consisted of a common core module and two out of four additional modules A, B, C, and D. The gross sample was randomly allocated to one of the six split questionnaires. For further details, see Luijkx et al. (2021).

Figure 2. Experimental comparison of survey modes for EVS Germany. AAPOR Response Rate 6 reported.

Peytchev and Peytcheva 2017). The use of this matrix design reduced the average questionnaire length for the self-administered mixed-mode survey to 38 minutes (for the web interview). The mixed-mode matrix survey was implemented in a responsive design with two phases. In the first phase, respondents were randomly allocated to two experiments on incentive strategies and mode choice sequence (sequential, concurrent), resulting in a 2×2 factorial experimental design. In the sequential mode choice sequence, the web mode was always offered first and mail questionnaires were only provided with a second reminder, whereas in the concurrent mode choice sequence respondents received both the mail and web questionnaires from the outset. The incentive strategies tested here include a 5 euro prepaid and a 10 euro postpaid incentive. After evaluation of the first phase, the second phase of the matrix survey started, for which we provided respondents with a 5 euro prepaid incentive in a concurrent mode choice sequence because this was the condition with the highest projected response rate. However, a matrix design results in a large
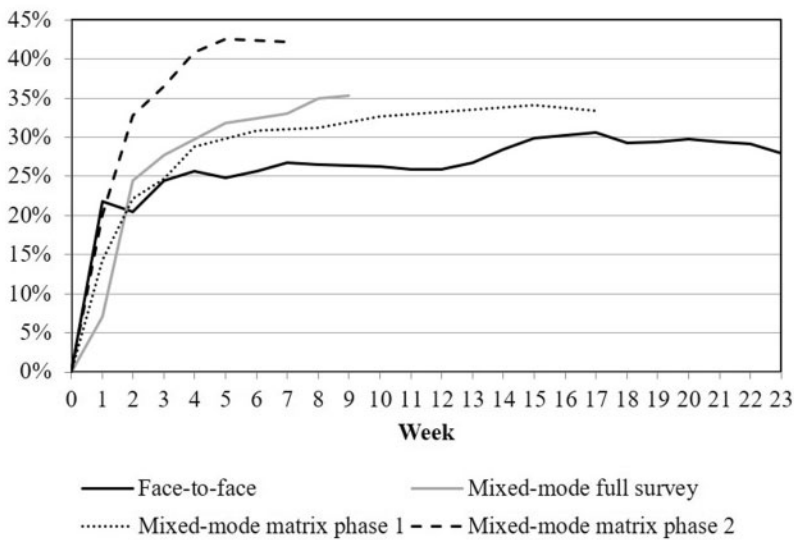
Figure 3. Response rate (AAPOR RR6) by week of fieldwork and survey mode. Response rates are estimated by each week of fieldwork and are based on final disposition codes.

amount of information missing by design for each case, because each respondent receives only a part of the full questionnaire. Consequently, while a matrix design reduces the length of the questionnaire, the reduction comes at the price of increased missing information. To analyze the data, the missing information needs to be either imputed or collected in follow-up surveys. In our data, about 33 percent of the answers are missing by design when using the matrix questionnaire instead of the full questionnaire.

Third, we conducted an additional mixed-mode survey using the full EVS questionnaire which—compared to the face-to-face survey—yielded a slightly shorter average interview duration of 55 minutes (for the web interview).

FIELDWORK AND RESPONSE RATE

The fieldwork of the face-to-face survey took place between October 23, 2017, and April 4, 2018. As an indicator for survey outcome, we calculated AAPOR Response Rate 6 (AAPOR 2016).[13] Overall, participation in the face-to-face survey remained relatively low with a response rate of 28 percent,

13. AAPOR Response Rate 6 was chosen, as it includes both complete and partial interviews in the calculation. This allowed us to track a consistent response rate throughout the whole survey period, as respondents could finish their web interviews at a later point in time. Following the AAPOR standard definitions (AAPOR 2016), interviews were defined as partially complete when respondents

despite respondents being offered monetary incentives. However, the outcome is comparable to experiences reported by ALLBUS and the German ESS (see table 1). Despite increasing the postpaid incentive and then switching to prepaid incentives during the field period, the completion of fieldwork was slow (6 months) and took much longer than the fieldwork for the mixed-mode surveys (6–8 weeks). Figure 3 provides a summary of the response rates and fieldwork periods for each of the survey modes.

The fieldwork of the mixed-mode matrix survey was conducted between November 16, 2017, and March 20, 2018. The mixed-mode survey in the matrix design achieved an overall response rate of 36.1 percent. As the mixed-mode matrix survey was implemented in a responsive design with two phases, response rates of each experimental group are detailed in table 2.[14,15] In the first phase, providing respondents with a 5 euro prepaid incentive in a concurrent mode choice sequence was the most efficient for the matrix survey. Thus, in the second phase of the survey, which started in January 2018, every respondent was provided with a 5 euro prepaid incentive in a concurrent mode choice sequence.

Encouraged by the promising outcome and the comparatively high response rate of the already lengthy mixed-mode matrix design survey, we decided to field the full-length EVS questionnaire in a self-administered mixed-mode survey as well. The fieldwork period for the full mixed-mode survey was between September 20 and November 28, 2018. The full-length EVS instrument achieved a response rate of 35.3 percent, relying on a concurrent mode choice sequence and offering a prepaid incentive of 5 euros.

COSTS

The fieldwork of the EVS 2017/18 face-to-face and mixed-mode experiments in Germany was conducted by the same survey institute relying on the same sample drawn from the population registers of the German population. Based

provided valid answers for 50 percent to 80 percent of the survey. When less than 50 percent of the applicable questions were answered, the interview was coded as a break-off.

14. Table 2 also considers the ratio of achieved web interviews to mailed interviews on paper questionnaires for sequential and concurrent designs. Here, it is noteworthy that the concurrent design resulted in only 20 percent web interviews. Thus, respondents showed a clear preference for using paper questionnaires when provided with a simultaneous mode choice. This preference was most pronounced for respondents aged 60 years or older, of which more than 90 percent of the respondents chose to reply by mail (not shown in the table). The preference for paper questionnaires does not seem to be a phenomenon only encountered in Germany; Denmark, Finland, and Switzerland report similar results (Luijkx et al. 2021).

15. Offering a 5 euro prepaid incentive resulted in response rates that were between 12.4 and 15.1 percentage points larger compared to offering a 10 euro postpaid incentive. This finding replicates previous studies on incentives in Germany that were solely focused on face-to-face surveys (Pforr et al. 2015), although the magnitude of the effect was larger in our study.

Table 2. Relative survey costs across modes for $N = 3,000$ interviews (face-to-face $= 100$ percent), share of web respondents, and response rates

Survey design	Total costs (%)	Share of web respondents (%)	Response rate (%)
Face-to-face	ref	–	28.0
Mixed-mode full survey (5€ prepaid, concurrent)	44.5	16.1	35.3
Mixed-mode matrix (5€ prepaid, sequential)[+][*]	37.9	50.7	36.0
Mixed-mode matrix (10€ postpaid, sequential)[+][*]	39.8	46.3	23.6
Mixed-mode matrix (5€ prepaid, concurrent)[+]	38.7	18.7	39.4
Mixed-mode matrix (10€ postpaid, concurrent)[+]	43.9	23.4	24.3

NOTE.—Ref = reference category. Face-to-face incentives are described in section "Mixed-mode experiments in the European Values Study."

[*]Respondents aged 60 or older also receive a concurrent mode choice. All costs in percent, relative to the reference category.

[+]Estimates are based on survey experiments implemented in phase one of the responsive mixed-mode matrix survey.

on our own calculations and information provided by the survey institute, we could disentangle the project's shared costs and estimate the total costs for each survey mode with sufficient precision. The actual costs cannot be reported here due to confidentiality. Instead, we provide estimates of the relative costs of each mode, using the face-to-face survey as the reference category.

For comparison, we projected the fixed and relative costs for a net sample of $N = 3,000$ realized interviews for each survey mode (see table 2). As we also conducted incentive and contact experiments in the first phase of the *responsive* mixed-mode matrix survey, we were also able to provide separate estimates for each experimental condition, again projected to $N = 3,000$ interviews in a static mixed-mode matrix design. As table 2 shows, costs for the face-to-face survey were more than twice the associated costs for both the mixed-mode survey using a matrix design and the full mixed-mode survey.

When comparing the full-length mixed-mode survey with the mixed-mode survey in a matrix design relying on 5 euro prepaid incentives and a concurrent contact strategy, it should be noted that the former was approximately 15 percent more expensive than the latter. There was also an advantage regarding response rates (+4 percentage points) for the shorter matrix design

($\varnothing$38 minutes) over the full survey ($\varnothing$55 minutes), which can be interpreted as the effect of the difference in the survey's length. This advantage should be put into perspective, however. The matrix design also implies relatively higher (data) preparation efforts, as several different questionnaires ultimately had to be programmed, tested, and administered. Furthermore, given the more complicated data structure including the systematic occurrence of missing values as implied by the matrix design (see above; roughly 33 percent of answers are missing in the matrix design compared to the full-length questionnaire), fielding the full-length survey seems to be a more viable and efficient strategy from a cost-benefit perspective.

SAMPLE COMPOSITION

To determine the sample's balance with respect to sociodemographic characteristics known from official population registers and, thus, the risk of non-response bias (Schouten et al. 2016), we calculated the Adjusted Coefficient of Variation (CV) as suggested by Schouten, Cobben, and Bethlehem (2009; see Supplementary Material table B.1, for more detailed information on the variables used). The CV integrates multivariate information on the risk of nonresponse bias in one indicator, facilitating a comparison between different experimental groups. For this purpose, we fitted nonresponse models to predict response propensities for each person in the gross sample based on the available information from the population register (age, sex, citizenship, size of municipality, East-West Germany). The CV is defined as the standard deviation of a sample's response propensities divided by the response rate. A CV close to zero indicates higher sample balance and, thus, a lower risk of nonresponse bias, whereas a higher CV indicates a less balanced sample.

Figure 4 presents the adjusted CV for each survey mode. While there were differences in the magnitude of the CVs (lowest CV [better sample balance] for the face-to-face survey), these differences were only significant when compared with the shortened mixed-mode matrix survey, but were not statistically significant for the mixed-mode full survey. From these findings, we conclude that the face-to-face mode represents the population slightly better than the mixed-mode matrix survey, but the different survey designs resulted in a similar sample balance and a similar risk of nonresponse bias with regard to the full-length mixed-mode survey.

Note, however, that the adjusted CVs rely on the limited amount of information we could acquire for the gross sample from the registers. Consequently, our nonresponse prediction includes a substantial degree of uncertainty, leading to large confidence intervals for the adjusted CVs. For this reason, we also compared the sample composition of the net samples against official population statistics for a more diverse set of sample
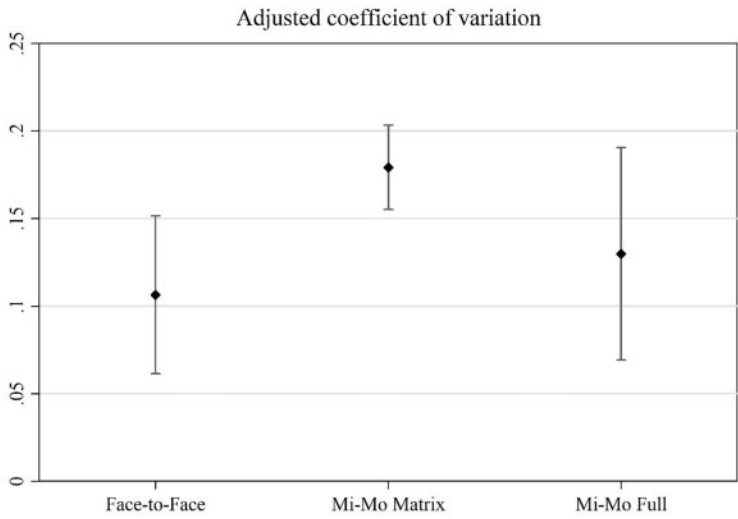
Figure 4. Adjusted coefficient of variation (CV) by survey mode. Adjusted Coefficient of Variation (CV) of predicted response propensities (logistic regression) dependent on age, gender, nationality, size of municipality, and East-/West Germany. With 95%-CI. Gross sample face-to-face = 5,314; gross sample mixed-mode matrix = 8,941; gross sample mixed-mode full = 1,908. Differences in the *N* for the gross samples compared to figure 2 are due to missing information on citizenship, not fully provided by the municipality registers.

characteristics and use the population statistics as a reference to calculate χ^2 tests of difference (see table 3).

Overall, all modes performed reasonably well; however, the face-to-face mode represents the population slightly better. First, there was no misrepresentation of gender observed for any mode. Second, the mixed-mode surveys underrepresented the "hard to reach" young age group, while no big difference exists for the face-to-face mode. Somewhat surprisingly, the face-to-face mode overrepresented the 60- to 69-year-old age bracket, while underrepresenting the very old (70 years and older). The mixed-mode surveys, on the other hand, had no issues reaching older age groups. Third, all modes strongly overrepresented the highly educated. Yet, the overrepresentation of the highly educated was more pronounced in the mixed-mode surveys compared to the face-to-face mode. Fourth, all modes underrepresented non-German citizens and, again, the differences were more pronounced for the mixed-mode surveys than for the face-to-face survey, indicating a language barrier that some respondents might not be capable of bridging in some instances, but that an interviewer might help to cross. Fifth, concerning

Table 3. Representation of basic sociodemographic variables by survey mode

	Population (%)	Face-to-face (%)	Mixed-mode matrix design (%)	Mixed-mode full design (%)
Gender				
Men	49.1	49.9	49.8	50.6
Women	50.9	50.1	50.2	49.4
Age				
18–29 years	16.8	16.5	12.7***	13.4*
30–39 years	15.1	14.0	13.9*	11.2**
40–49 years	15.5	15.5	15.1	17.2
50–59 years	19.3	19.3	20.3	20.9
60–69 years	14.6	18.3***	17.0***	17.6*
70+ years	18.8	16.3*	21.1***	19.8
Education				
ISCED level 0–2	18.7	11.5***	17.6	13.6***
ISCED level 3–4	57.0	53.7**	41.9***	47.4***
ISCED level 5+	24.3	34.8***	40.4***	39.0***
Household size				
1 hh member	20.9	20.4	20.8	20.8
2 hh members	33.5	39.7***	42.1***	41.5***
3 hh members	18.0	17.7	15.5***	15.8
4 hh members	18.5	15.3**	15.2***	16.1
5+ hh members	9.0	6.9**	6.4***	5.7**
Citizenship				
German citizenship	87.9	91.2***	94.3***	93.5***
Non-German citizenship	12.1	8.8***	5.7***	6.5***
Employment status[a]				
Active	56.5	57.8	60.2***	59.2
Unemployed	2.1	3.7***	2.1	2.1
Not in labor force	41.3	38.6*	37.7***	38.8
Urbanity (DEGURBA)				
Cities	35.9	33.7	37.3	36.0
Towns and suburbs	41.5	39.7	39.7*	41.3
Rural areas	22.7	26.6***	22.9	22.7

NOTE.—χ^2 Test of difference,
*$p < 0.05$
**$p < 0.01$
***$p < 0.001$
[a]For 20 years and older. See Supplementary Material table A, for the sources of population data.

employment status, the distribution in the face-to-face mode was closer to the overall population compared to the mixed-modes regarding those active or not in the labor force. In contrast, the face-to-face mode overrepresented the unemployed. Sixth, the picture reverses for urbanity, where the face-to-face mode performed slightly worse than the mixed-mode survey, overrepresenting the rural population. Seventh and finally, no big differences between the modes exist in terms of household size: all modes overrepresented two-person households and underrepresent large households with four or more household members.

ITEM NONRESPONSE AND RESPONSE QUALITY

An experimental comparison of nonresponse or response quality indicators between different survey modes is generally difficult, as there are several inherent differences in the design of the questionnaire (interviewer-administered versus self-administered), such as "don't know" options or using grid questions (de Leeuw 2018, pp. 82–83). This also applies for the EVS surveys, in which the usability of the survey was optimized for each mode, leading to important differences in the design of the questionnaires. For instance, the item batteries were presented differently in the different survey modes: while question batteries were presented item-by-item in the face-to-face mode, in the mail mode grids were used when possible to increase the questionnaire's readability. For large screens in the web mode, we also showed grids, whereas for mobile devices, we presented only single items.

Figure 5 provides an overview of frequently employed indicators for unit nonresponse (break-off cases, partial interview rate), item nonresponse (no answer, don't knows), and response quality (acquiescence, straightlining, middle categories, extreme response style; see Supplementary Material table B.2, for more details about these variables). However, the differences in the design of each questionnaire make it impossible to experimentally isolate the effect of a specific mode; rather, figure 5 can be understood as a combined test for differences between the data collection modes. As it is often considered best practice to remove break-off cases prior to releasing a dataset (e.g., latest published dataset of the EVS 2017), we also analyzed item nonresponse and response quality indicators without break-off cases.

In the face-to-face mode, respondents typically do not break off a personal interview (0.0 percent)[16] or conduct partial interviews (0.1 percent). In contrast, the web survey software tracks any survey response respondents may give—even if they drop out early. For the mail mode, we assume that respondents will not make the effort to drop a partially filled-out

16. The field institute reported back 11 cases as "not valid," which may be break-off cases as well.
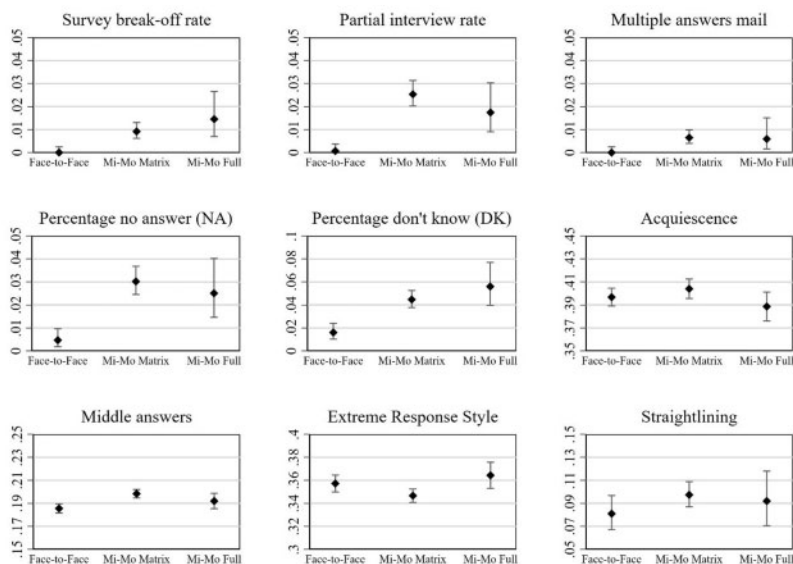
Figure 5. Unit nonresponse, item nonresponse, and response style indicators by survey mode. Break-off cases have been excluded from analysis except for indicators of unit nonresponse. Binomial 95 percent confidence intervals have been obtained for the survey break-off and partial interview rates and Poisson 95 percent confidence intervals for the share of the count of no answer, don't know, multiple answers, and straightlining. For the indicators measuring extreme response style, middle answers and acquiescence normal-based 95 percent confidence intervals have been computed.

questionnaire off at the postal office. Overall, we found only moderate levels of survey break-off and partial interviews in the self-administered mode (see figure 5). For the shortened mixed-mode matrix survey, we observed a 1.0 percent break-off rate and 2.5 percent partial interviews. For the mixed-mode full survey, we observed a 1.7 percent break-off rate and 1.5 percent partial interviews. As the confidence intervals in figure 5 indicate, differences between the mixed-mode matrix survey and the mixed-mode full survey were not significant for break-off rates and partial interview rates, but were highly significant in comparison to the face-to-face mode.

In the web mode, the survey software's programming ensures that respondents can only use the response scales as intended. For instance, respondents can only select one of the response options in single-choice closed-ended questions. In face-to-face mode, the programming of the interviewer's device (e.g., laptop, tablet) has a similar effect. Yet, when answering mail questionnaires, respondents can use scales in unintended ways, for instance, by selecting more

than one response option in single-choice questions. This behavior makes it impossible to assign one valid code for these respondents' answers. Hence, invalid answers are a challenge when fielding a mail survey. In our data, the amount of multiple answers to questions was well below 1 percent in the combined mixed-mode datasets (web and mail) when using either matrix (0.6 percent) or the full questionnaires (0.4 percent).

Another major difference relates to the usage of "don't know" answers. Following best practice, "don't know" options have not been mentioned by the interviewers and were not included in the show cards for the face-to-face mode. Yet, "don't know" was still introduced as an admissible response alternative in the interviewers' questionnaire if respondents spontaneously mentioned this option. For the web and mail modes, we have explicitly provided "don't know" options for most questions so as not to frustrate respondents and force them to give substantive answers (Bethlehem and Biffignandi 2012; Dillman, Smyth, and Christian 2014). However, providing "don't know" response options could increase satisficing, because respondents may not want to think about the questions and the "don't know" option provides an easy alternative (Krosnick et al. 2002). These design differences may well explain why we found far higher shares of "don't know" answers for the self-administered modes, where this option was explicitly offered. As figure 5 shows, about 4.6 percent of all items were answered with "don't know" for the mixed-mode matrix survey and 5.8 percent for the mixed-mode full survey. In face-to-face mode, the share was much lower, with only 1.7 percent of "don't know."

Similarly, the share of "no answer" was significantly lower in the face-to-face mode (0.5 percent) than in the self-administered modes (2.4 percent for the mixed-mode matrix survey and 2.1 percent for the mixed-mode full survey). Respondents in the mail mode can always choose to leave questions blank. For this reason, respondents in the web mode were also allowed to skip questions that they prefer not to answer and thereby decrease their response burden.

To address response styles that previous research has associated with satisficing or inattentive responding (Krosnick 1991, 1999; Weijters, Cabooter, and Schillewaert 2010; Gummer, Roßmann, and Silber 2021), we compared the share of acquiescent, mid-point responding, extreme responding, and straightlining in rating scales. The share of acquiescence and extreme response styles among respondents varied only slightly between the face-to-face survey and the full-length and matrix mixed-modes, with no significant differences. The prevalence of mid-point responding and straightlining in the mixed-mode full-length survey was in between face-to-face and the mixed-mode matrix survey. This was especially apparent for mid-point responding. In the face-to-face mode, respondents used the mid-point category of the rating scale in 18.5 percent of their answers on average, whereas this category

was used in 19.9 percent of the answers in the mixed-mode matrix (difference was statistically significant). When completing the full questionnaire in mixed-modes, 19.2 percent of respondents' answers were the mid-point category (no significant differences). In this context, the results seem encouraging in terms of straightlining in the self-administered modes. As noted above, in the self-administered modes, for large screens in the web and for paper-based questionnaires grid questions were used, which have been shown to increase the likelihood of straightlining response behavior in comparison to presenting question batteries item-by-item (e.g., Roßmann, Gummer, and Silber 2018). However, in our study, respondents to the self-administered surveys did not straightline more frequently than respondents in face-to-face surveys.

COMPARISON OF CORE VARIABLES

Turning to the outcomes of our surveys, we investigated how the distribution of key variables differed between experimental groups. While differences in response rates, costs, and data quality may interest researchers and survey practitioners in particular, substantive researchers' main concern will likely be the differences in the survey's key variables.

Figure 6 details the mean value with confidence intervals for substantive variables that the EVS Methodology Group considered key characteristics and were included in the core of the matrix questionnaire (i.e., fielded to every respondent regardless of questionnaire version; see Supplementary Material table B.3, for more information about these variables). This analysis considers every substantive core variable ($n = 18$) answered on a rating scale. For comparability, we rescaled all scales to a range from 0 to 1. Overall, the substantive differences between the experimental groups were negligible for most variables. Due to the large sample size, the confidence intervals are rather narrow, but the distance between the mean values remained very small.

The largest visible differences occurred in respondents' satisfaction with the political system in Germany. Here, respondents to the face-to-face survey reported higher satisfaction compared to the self-administered modes. This could be the result of social desirability, as respondents may be less likely to state perceived issues with the democratic system when an interviewer is present.

Conclusion

GSS-type surveys are increasingly difficult to conduct face-to-face: response rates are falling, while refusal and non-contact rates are both rising. In Europe, response rates of less than 50 percent are quite common today; in
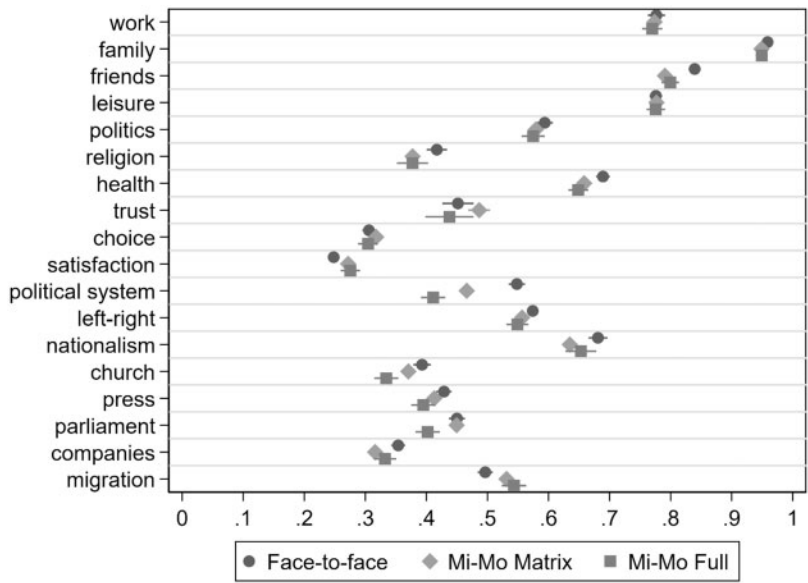
Figure 6. Outcomes for substantive key variables across experimental groups. All variables measure respondents' attitudes and position towards the respective topic on rating scales. For comparability, we rescaled all variables to range between 0 and 1.

some countries, such as Germany, it is increasingly difficult to even reach a 30 percent response rate, despite increased financial resources and intensified fieldwork efforts (Beullens et al. 2018).

In light of such developments, we experimentally tested whether long GSS-type surveys can be conducted in self-administered modes (web, mail). To this end, we implemented the EVS 2017/18 in Germany as an interviewer-administered face-to-face survey, as a self-administered mixed-mode survey in a matrix design, and as a self-administered mixed-mode survey featuring the full-length questionnaire. Our findings indicate that self-administered modes require less fieldwork time (6 to 8 weeks) than the face-to-face survey (6 months) and come with much lower costs (savings of over 50 percent). What is more, they also achieve higher response rates (+8.1 percentage points for the mixed-mode matrix and +7.3 percentage points for the full-length survey).

With respect to sample composition, for all modes, we found small deviations from the population, although the face-to-face survey has an edge in reaching younger age groups. Every mode underrepresented non-German citizens, with the self-administered modes being worse than the face-to-face

survey. All modes strongly overrepresented the highly educated, but deviations from the population were smaller in the face-to-face mode. On the other hand, the face-to-face mode overrepresented the unemployed and those living in rural areas.

Considering nonresponse, the face-to-face mode exhibited lower break-off rates and fewer partial interviews. Furthermore, we observed a smaller share of "don't know" and no answers in face-to-face mode—certainly an effect explained by the interviewers' presence. Data quality indicators varied only slightly across modes. In terms of substantive answers to 18 key items of the EVS capturing attitudes toward various topics, our analysis revealed only minor differences between the three survey modes.

Overall, our results suggest that self-administered mixed-mode surveys can be a viable alternative to face-to-face surveys when conducting general population studies (in Germany). However, some critical remarks and suggestions are necessary when conducting such a survey: first, even though the mixed-mode matrix survey decreases the response burden for respondents, it requires considerably more effort for field institutes, data curators, and researchers. These increased efforts include creating several versions of the questionnaire, programming and testing each instrument, as well as designing several paper versions. In light of the results of this study, we recommend fielding a single but longer self-administered survey instead of multiple shorter ones.

Second, in our study the concurrent mixed-mode (i.e., simultaneously offering respondents the possibility to participate in the survey per web or mail) was associated with the highest response rate. This design induced 80 percent of the participants to respond per mail which, in turn, leads to higher data encoding costs and limits the possibility of using complex survey instruments (e.g., extensive routing, preloads, etc.). We opted for this design because our emphasis was on yielding a high response rate and, thus, a large net sample. However, a push-to-web strategy may be more expedient in other situations to minimize the proportion of mail surveys. Note that the success of a push-to-web strategy may depend on the availability of computers and other devices in the target population, such as tablets and smartphones.

Third, although self-administered mixed-mode surveys seem to be a viable alternative to face-to-face surveys for GSS-type programs, this conclusion relies on the availability of a suitable sampling frame: a register (or several) of named individuals covering the residents of a given country that can be accessed for research purposes. This condition is lacking in many countries, such as France, Ireland, or the UK.[17]

---

17. An overview on the use of sampling frames in Europe can be accessed under https://seriss. eu/wp-content/uploads/2017/01/SERISS-Deliverable-2.1-Report-on-the-use-of-sampling-frames-in-European-studies.pdf (retrieved July 29, 2020). Of course, such a "register" of named

Fourth, in this study, we focused on the implementation of self-administered mixed-mode surveys in Germany. This decision was motivated by the fact that the EVS Germany featured extensive experimentation on different ways of implementing such surveys compared to other EVS countries that also tested mixed-mode surveys. Studying a single country further allowed us to investigate a comprehensive set of outcome indicators: fieldwork effort, response rate, survey costs, sample composition, nonresponse and data quality indicators, and distribution of substantive core variables. As response rates in Germany are among the world's lowest, the German case is particularly interesting for investigating (new) approaches to addressing low response rates and increasing survey costs. If the global trend in declining response rates continues, other countries will eventually find themselves in a similar position and could gather valuable insights from a case such as Germany.

Note, however, that the singular focus on Germany may impair the generalizability of these findings to other countries. In their overview of the mixed-mode testing as part of EVS 2017/2018, Luijkx et al. (2021) illustrate discrepancies in how each country implemented the surveys. Based on the respective context and experiences, each country adopted different solutions with respect to whether and when to offer paper-based questionnaires in the self-administered surveys, which incentives to use, and so on. It remains an open question whether similar designs would perform (equally) well in each country when implemented in cross-national surveys. Thus, future research should extend the evaluation of a comprehensive set of quality and cost factors (as presented in this study) to a broader set of countries, if the research design offers an opportunity to do so. Given that comparability is a key feature of cross-national studies, design decisions are typically a trade-off between comparability across countries (i.e., a design not optimized for any country but that can be implemented similarly in many countries) and optimizing country-specific performance (i.e., each country selects the best design for its situation). A similar research gap exists for comparability over time when designs are modified for either cross-sectional or panel surveys. We consider this line of research a formidable challenge for future research. The tests conducted as part of the EVS 2017/18 establish a valuable and necessary starting point, but many questions will require more extensive data collection, especially for comparability. Our study, based on the EVS in Germany, provides some helpful indications about the different dimensions

---

individuals does not need to be an official register. However, it most likely has to go beyond an address-based sample, because it is unclear how a household should be selected at a sampled address and how an individual household member should be chosen from within that household. These issues are resolvable (while not being unproblematic) when doing face-to-face interviews and having interviewers in the field carrying out these selections. Without such an aid in the field, this problem does not seem to have a solution yet.

of survey outcomes and survey design decisions that may be crucial to consider when planning mixed-mode surveys in a cross-national setting as well.

Data Availability Statement

REPLICATION DATA are not available because of privacy concerns. The editors have waived POQ's replication policy for this manuscript. However, the analysis code is available at https://doi.org/10.7802/2150.

Supplementary Material

SUPPLEMENTARY MATERIAL may be found in the online version of this article: https://doi.org/10.1093/poq/nfab039.

References

AAPOR. 2016. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. *9th edition.* American Association for Public Opinion Research. https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf.

Adiguzel, F., and M. Wedel. 2008. "Split Questionnaire Design for Massive Surveys." *Journal of Marketing Research* 45(5):608–17.

Allen, Jeff, Jeffrey Scott, Michael Roberts, and Octavian Chiriac. 1997. "A Trend Analysis of Mail Survey Response Rates for Selected Marketing Journals, 1960–1995." Academy of Marketing Science (AMS) Annual Conference, Coral Gables, FL.

Bethlehem, Jelke G., and Silvia Biffignandi. 2012. *Handbook of Web Surveys, Wiley Handbooks in Survey Methodology*. Hoboken, NJ: Wiley.

Beullens, Koen, and Geert Loosveldt. 2016. "Interviewer Effects in the European Social Survey." *Survey Research Methods* 10(2):103–18.

Beullens, Koen, Geert Loosveldt, Caroline Vandenplas, and Ineke Stoop. 2018. "Response Rates in the European Social Survey: Increasing, Decreasing, or a Matter of Fieldwork Efforts?" *Survey Methods: Insights from the Field (SMIF)*. doi:10.13094/SMIF-2018-00003.

Beullens, Koen, Hideko Matsuo, Geert Loosveldt, and Caroline Vandenplas. 2014. *Quality Report for the European Social Survey, Round 6*. London: European Social Survey ERIC.

Blom, Annelies G., Christina Gathmann, and Ulrich Krieger. 2015. "Setting Up an Online Panel Representative of the General Population: The German Internet Panel." *Field Methods* 27(4):391–408.

Bosnjak, Michael, Tanja Dannwolf, Tobias Enderle, Ines Schaurer, Bella Struminskaya, Angela Tanner, and Kai W. Weyandt. 2018. "Establishing an Open Probability-Based Mixed-Mode Panel of the General Population in Germany: The GESIS Panel." *Social Science Computer Review* 36(1):103–15.

Bradburn, Norman M. 1992. "A Response to the Nonresponse Problem." *Public Opinion Quarterly* 56(3):391–97.

Brick, J. Michael, and Douglas Williams. 2013. "Explaining Rising Nonresponse Rates in Cross-Sectional Surveys." *Annals of the American Academy of Political and Social Science* 645(1):36–59.

Callegaro, Mario, Katja Lozar Manfreda, and Vasja Vehovar. 2015. *Web Survey Methodology*. Los Angeles: SAGE.

Childs, Ruth A., and Andrew P. Jaciw. 2002. "Matrix Sampling of Items in Large-Scale Assessments." *Practical Assessment, Research & Evaluation & the Health Professions* 8:1–9.

Czajka, John L., and Amy Beyler. 2016. *Declining Response Rates in Federal Surveys: Trends and Implications*. Washington, DC: Mathematica Policy Research.

de Leeuw, Edith. 2018. "Mixed-Mode: Past, Present, and Future." *Survey Research Methods* 12(2):75–89.

de Leeuw, Edith, and Wim de Heer. 2002. "Trends in Household Survey Nonresponse: A Longitudinal and International Comparison." In *Survey Nonresponse*, edited by Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, 41–54. New York: John Wiley & Sons.

de Leeuw, Edith, Joop Hox, and Annemieke Luiten. 2018. "International Nonresponse Trends Across Countries and Years: An Analysis of 36 Years of Labour Force Survey Data." *Survey Methods: Insights from the Field*. doi:10.13094/SMIF-2018-00008.

Dillman, Don A., Jolene D. Smyth, and Leah Melani Christian. 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. 4th ed. Hoboken: Wiley.

Dutwin, David, and Paul Lavrakas. 2016. "Trends in Telephone Outcomes, 2008–2015." *Survey Practice* 9(3):1–9.

Groves, Robert M. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly* 70(5):646–75.

Groves, Robert M., and Emilia Peytcheva. 2008. "The Impact of Nonresponse Rates on Nonresponse Bias—A Meta-Analysis." *Public Opinion Quarterly* 72(2):167–89.

Gummer, Tobias, Joss Roßmann, and Henning Silber. 2021. "Using Instructed Response Items as Attention Checks in Web Surveys: Properties and Implementation." *Sociological Methods & Research* 50(1):238–64.

Krosnick, Jon A. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5:213–36.

———. 1999. "Survey Research." *Annual Review of Psychology* 50:537–67.

Krosnick, Jon A., A. L. Holbrook, M. K. Berent, R. T. Carson, W. M. Hanemann, R. J. Kopp, R. C. Mitchell, S. Presser, P. A. Ruud, V. K. Smith, W. R. Moody, M. C. Green, and M. Conaway. 2002. "The Impact of 'No Opinion' Response Options on Data Quality—Non-Attitude Reduction or an Invitation to Satisfice?" *Public Opinion Quarterly* 66(3):371–403.

Luijkx, Ruud, Guðbjörg A. Jónsdóttir, Tobias Gummer, Michèle Ernst Stähli, Morten Fredriksen, Kimmo Ketola, Tim Reeskens, Evelyn Brislinger, Pablo Christmann, Stefán Þ. Gunnarsson, Árni B. Hjaltason, Dominique Joye, Vera Lomazzi, Angelica M. Maineri, Patricia Milbert, Michael Ochsner, Alexandre Pollien, Marlène Sapin, Ivet Solanes, Sascha Verhoeven, and Christof Wolf. 2021. "The European Values Study 2017: On the Way to the Future Using Mixed-Modes." *European Sociological Review* 37(2):330–346.

Matsuo, Hideko, and Geert Loosveldt. 2013. *Report on Quality Assessment of Contact Data Files in Round 5: Final Report 27 Countries*. Leuven: European Social Survey, University of Leuven.

O'Muircheartaigh, Colm. 2018. "Sampling for Single and Multi-Mode Surveys Using Address-Based Sampling." In *The Palgrave Handbook of Survey Research*, edited by David L. Vannette and Jon A. Krosnick, 13–18. Cham (CH): Palgrave.

Peytchev, Andy, and Emilia Peytcheva. 2017. "Reduction of Measurement Error Due to Survey Length: Evaluation of the Split Questionnaire Design Approach." *Survey Research Methods* 11(4):361–68.

Pforr, Klaus, Michael Blohm, Annelies Geerte Blom, Barbara Erdel, Barbara Felderer, Mathis Fräßdorf, Kristin Hajek, Susanne Helmschrott, Corinna Kleinert, Achim Koch, Ulrich Krieger, Martin Kroh, Silke Martin, Denise Saßenroth, Claudia Schmiedeberg, Eva-Maria

Trüdinger, and Beatrice Rammstedt. 2015. "Are Incentive Effects on Response Rates and Nonresponse Bias in Large-Scale, Face-to-Face Surveys Generalizable to Germany? Evidence from Ten Experiments." *Public Opinion Quarterly* 79(3):740–68.

Raghunathan, Trivellore E., and James E. Grizzle. 1995. "A Split Questionnaire Survey Design." *Journal of the American Statistical Association* 90(429):54–63.

Revilla, Melanie. 2010. "Quality in Unimode and Mixed-Mode Designs: A Multitrait-Multimethod Approach." *Survey Research Methods* 4(3):151–64.

Revilla, Melanie. 2015. "Comparison of the Quality Estimates in a Mixed-Mode and a Unimode Design: An Experiment from the European Social Survey." *Quality & Quantity* 49 (3):1219–38.

Roßmann, Joss, Tobias Gummer, and Henning Silber. 2018. "Mitigating Satisficing in Cognitively Demanding Grid Questions: Evidence from Two Web-Based Experiments." *Journal of Survey Statistics and Methodology* 6(3):376–400.

Schouten, Barry, Fannie Cobben, and Jelke Bethlehem. 2009. "Indicators for the Representativeness of Survey Response." *Survey Methodology* 35(1):101–13.

Schouten, Barry, Fanny Cobben, Peter Lundquist, and James Wagner. 2016. "Does More Balanced Survey Response Imply Less Non-Response Bias?" *Journal of the Royal Statistical Society A* 179(3):727–48.

Smith, Tom W. 1995. "Trends in Non-Response Rates." *International Journal of Public Opinion Research* 7(2):157–71.

Steeh, Charlotte G. 1981. "Trends in Nonresponse Rates, 1952–1979." *Public Opinion Quarterly* 45(1):40–57.

Stoop, Ineke. 2005. *The Hunt for the Last Respondent: Nonresponse in Sample Surveys*. The Hague: Social and Cultural Planning Office.

Tourangeau, Roger. 2017. "Presidential Address: Paradoxes of Nonresponse." *Public Opinion Quarterly* 81(3):803–14.

Villar, Ana, and Rory Fitzgerald. 2017. "Using Mixed Modes in Survey Data Research: Results from Six Experiments." In *Values and Identities in Europe: Evidence from the European Social Survey*, edited by M. Breen, 273–310. London: Routledge.

Weijters, Bert, Elke Cabooter, and Niels Schillewaert. 2010. "The Effect of Rating Scale Format on Response Styles: The Number of Response Categories and Response Category Labels." *International Journal of Research in Marketing* 27(3):236–47.

Wuyts, Celine, and Geert Loosveldt. 2019. *Quality Matrix for the European Social Survey, Round 8*. London: European Social Survey ERIC.