OXFORD

The European Values Study 2017: On the Way to the Future Using Mixed-Modes

Ruud Luijkx [1,2,*], Guðbjörg Andrea Jónsdóttir[3], Tobias Gummer[4],
Michèle Ernst Stähli[5], Morten Frederiksen[6], Kimmo Ketola[7],
Tim Reeskens[1], Evelyn Brislinger[4], Pablo Christmann[4],
Stefán Þór Gunnarsson[3], Árni Bragi Hjaltason[3], Dominique Joye[5,8],
Vera Lomazzi[4], Angelica M. Maineri[1], Patricia Milbert[5], Michael Ochsner[5],
Alexandre Pollien[5], Marlène Sapin[5], Ivet Solanes[4], Sascha Verhoeven[4], and
Christof Wolf[4,9]

[1]Department of Sociology, Tilburg University, Tilburg, Netherlands, [2]Department of Sociology and Social Research, University of Trento, Trento, Italy, [3]Social Science Research Institute, University of Iceland, Reykjavík, Iceland, [4]GESIS—Leibniz Institute for the Social Sciences, Mannheim & Cologne, Germany, [5]FORS—Swiss Foundation for Research in Social Sciences, Lausanne, Switzerland, [6]Department of Sociology and Social Work, Aalborg Universitet, Aalborg, Denmark, [7]Church Research Institute, Tampere, Finland, [8]Institute of Social Sciences, University of Lausanne, Lausanne, Switzerland and [9]School of Social Sciences, University of Mannheim, Mannheim, Germany

*Corresponding author. Email: R.Luijkx@tilburguniversity.edu

Abstract

The European Values Study (EVS) was first conducted in 1981 and then repeated in 1990, 1999, 2008, and 2017, with the aim of providing researchers with data to investigate whether European individual and social values are changing and to what degree. The EVS is traditionally carried out as a probability-based face-to-face survey that takes around 1 hour to complete. In recent years, large-scale population surveys such as the EVS have been challenged by decreasing response rates and increasing survey costs. In the light of these challenges, six countries that participated in the last wave of the EVS tested the application of self-administered mixed-modes (Denmark, Finland, Germany, Iceland, the Netherlands, and Switzerland). With the present data brief, we will introduce researchers to the latest wave of the EVS, the implemented mode experiments, and the EVS data releases. In our view, it is pivotal for data use in substantive research to make the reasoning behind design changes and country-specific implementations transparent as well as to highlight new research opportunities.

Introduction

The European Values Study (EVS) was first conducted in 1981 by the European Values Systems Study Group (EVSSG). The EVSSG was a group of academics who, at the time of the first European parliamentary elections, sought to analyse the moral and social values underlying the social and political institutions, asking questions such as do Europeans share common values; are values changing and, if so, in what direction. The first survey conducted in 1981 collected data from 16 countries. Since then, EVS was carried out every 9 years (1990, 1999, 2008, 2017), with the highest number of countries/regions (47) in 2008. In the most recent survey, at the time of writing this data brief, 34 countries across Europe participated. The EVS is primarily funded by its participating members' institutions, such as universities, research institutes, national science foundations, and private sponsors.

The EVS was founded to respond to questions emerging in political and social debates at the end of the 1970s, such as whether Europe was one cultural entity or converging to one, whether Christian values were central in the lives of Europeans, whether new belief systems were emerging that might ultimately replace Christian belief, and what the implications of all these changes were for the unity of Europe (see www.europeanvaluesstudy.eu for more details). To adequately address these research questions, the EVS incorporated questions measuring values and attitudes in several domains of life: family, work, religion and morale, politics and society, environment, as well as national identity, tolerance, and social solidarity. Over the waves, the substantive scope of EVS expanded, but in the last wave, the questionnaire has been shortened not to overload interview time. Since the early 1980s, EVS data sets have been used in at least 2,100 publications in the social sciences and related fields (EVS, 2020a).

With its emphasis on surveying values and attitudes on several domains of life, the EVS inspired the development of the World Values Survey (WVS, see www.worldvaluessurvey.org) also committed to research questions on values change, e.g. the shift from material to post-materialist values (Inglehart, 1997) and more recently the idea of a 'Cultural Backlash' (Norris and Inglehart, 2019). Also in its dissemination, there is a close relation with WVS, which is in existence since the early 1990s. Since the third wave of EVS, EVS and WVS released longitudinal EVS and WVS files with common dictionaries that facilitated easy merging of the two data sets and so makes it possible to undertake global analyses.

The EVS is traditionally carried out as a probability-based interviewer-administered face-to-face (F2F) survey that takes around 1 hour to complete. In recent years, large-scale population surveys such as the EVS have been challenged by decreasing response rates and increasing survey costs (de Leeuw and De Heer, 2002; Groves and Peytcheva, 2008; Brick and Williams, 2013; de Leeuw, Hox and Luiten, 2018). Non-response is most likely not random and might eventually correlate with opinions that are at the core of the EVS, such as trust (Billiet et al., 2007), making that traditional modes of data collection potentially jeopardize the response to substantial research questions. In the light of these developments, the EVS Methodology Group, which oversees data collection standards and quality, allowed the use of self-administered modes in EVS 2017 under the condition that they were based on the same kind of probability sample, according to the same standards as for the F2F survey. Six countries decided to (experimentally) test the implementation of self-administered mixed-mode designs as part of EVS 2017 with respect to their country-specific demands and contexts: Denmark, Finland, Germany, Iceland, the Netherlands, and Switzerland.

With our data brief, we want to familiarize readers with the latest EVS data release, the recent methodological innovations, and promote the use of EVS for social science research. In our view, this brief should prove helpful to researchers who aim to use the data for substantive research as well as for methodologists who are interested in getting insights into self-administered mixed-modes in a large-scale cross-cultural survey programme.

In the following, we describe the methodological challenges that motivated the EVS Methodology Group to allow for the use of self-administered modes. We will look at the implications of this decision for the (country-specific) survey designs and illustrate how each country has (experimentally) implemented the EVS. We then describe how it performed in each instance. We continue by providing information on the data release, structure of data files, documentation, and data access to facilitate the use of the most recent EVS wave that includes several experiments and different designs. We close with discussing limitations and opportunities for research with the EVS data.

Methodological Challenges

In many countries, survey researchers are facing declining response rates (de Leeuw and De Heer, 2002; Brick

and Williams, 2013). Neither is this trend new (e.g., Steeh, 1981) nor is it limited to individual countries as declining response rates have been reported in various countries (e.g., Stoop et al., 2010). Albeit low response rates are not necessarily linked to non-response bias (Groves and Peytcheva, 2008) non-participation may lead to non-response bias. As a consequence, research has focused on meeting the non-response challenge in recent decades by advancing survey methodology in this regard (Kreuter, 2013) with incentives likely being the most prominent and well-researched tool in the field (for an overview see, e.g., Pforr et al., 2015) and the adaptive and responsive design paradigms (Groves and Heeringa, 2006; Wagner, 2008; Tourangeau et al., 2017), which advocate the use of more flexible data collection protocols, being the newest approach for tackling non-response.

The downward trend in response rates appears to be continuing and is leading researchers either to invest more effort in fieldwork or to work with larger gross samples. If outcome rates are not stabilized, the resulting sample size may be too small and the statistical power of the data may be too low, or larger gross samples must be used. As a consequence, increasing survey costs have been reported (e.g., Massey and Tourangeau, 2013).

While this development is a formidable challenge for most surveys, it becomes an even more pressing issue for large cross-national surveys such as the EVS, European Social Survey (ESS), International Social Survey Programme (ISSP), WVS, and others. These surveys' added value depends on their ability to cover a large and diverse number of countries over time. Thus, enabling cross-national comparison (over time) requires a trade-off between the need to ensure high data quality and avoiding that countries are unable to participate because they cannot cover the costs of a survey complying with the methodological requirements of the study (e.g., using a probability-based sample, achieving a given target net sample size, or reaching specific outcome rates). Consequently, survey researchers—and especially those in charge of cross-national studies—must cope with the rising costs in F2F surveying to maintain the availability of high-quality survey data for social science research.

Design of the Mixed-Mode Experiments

For the purpose of testing methodological innovations, especially with regard to increasing response rates and lowering survey costs, in EVS 2017 countries had the possibility to implement self-administered surveys in parallel to the traditional F2F survey. Specific guidelines were developed by the EVS Central Team in cooperation with those country teams that expressed interest in employing self-administered modes.

The suggested main alternative to fielding a F2F survey was a self-administered web-based survey (Computer Assisted Web Interview, or CAWI) of a probability-based sample. Countries were encouraged to find suitable strategies (e.g., for contact modes or incentives) to make sure that the response rate of the survey reached at least 40 per cent. For this reason, some of the countries also decided to offer a postal self-administered paper-and-pencil survey, which allowed them to reach 'offline' segments of the population and accommodate respondents who did not want to answer a CAWI survey.

Other general guidelines referred to the preferred layout of the web survey and were designed with the purpose of maintaining comparability with the F2F mode. For instance, the item-by-item presentation of questions was recommended for the CAWI mode; progress bars were not allowed; the design had to be adaptive to allow for completion on mobile devices; scrolling had to be avoided. For a complete overview of the self-administered mixed-mode guidelines, please refer to the EVS 2017 Methodological Guidelines (EVS, 2020e).

Matrix Questionnaire

The EVS 2017 was based on a 1-hour questionnaire designed for F2F mode. Compared to F2F surveys, where interviewers can motivate and guide respondents, self-administered surveys are assumed to require a larger effort from the respondent (Klausch, Hox and Schouten, 2013). For this reason, it is common practice to reduce the length of surveys when moving from an interviewer-administered to a self-administered mode (cf. Olson et al., 2020), and the recommended threshold is usually around 20 minutes for a web-based survey (e.g., Revilla and Ochoa, 2017). The EVS not only encouraged testing the use of a 1-hour questionnaire, but also proposed the use of matrix questionnaires to reduce the overall length of the questionnaire.

A matrix questionnaire design (or split questionnaire design, Raghunathan and Grizzle, 1995) consists of splitting the questionnaire into different, shorter versions, each version including a selection of questions from the full questionnaire. Respondents are then randomly assigned to the different versions of the questionnaire, ensuring that—over the entire sample—all

Table 1. Summary of thematic blocks

Blocks	Themes	Maximum number of ticks per block
Core	Socio-demographics, questions often used as control or correlating highly with other items of the thematic blocks A–D	58
A	Family, work, socio-demographic questions about parents and partner	55
B	Religion, morality, and social identity	54
C	Society	53
D	Politics	53

Table 2. An illustrative example for a matrix questionnaire design

Respondent groups	Blocks				
	Core	A	B	C	D
1	300	300	300		
2	300	300		300	
3	300	300			300
4	300		300	300	
5	300		300		300
6	300			300	300

Notes: 1,800 respondents overall, 900 respondents for each substantive question; at least 300 respondents for each binary combination of questions.

questions have been answered by a sufficient large share of respondents.

In the EVS, the survey was divided into four thematic blocks (A, B, C, and D) and one 'core' block with background variables. Each respondent had to answer two of the thematic blocks and the core. Each block contained no more than 60 ticks. Table 1 provides a summary of thematic blocks created in this process.

Splitting the F2F questionnaire resulted in six versions of the self-completion questionnaire, representing all possible combinations of blocks. Table 2 represents an example of allocating respondents to matrix groups for a sample size of 1,800 respondents. With this design, each block is answered by 900 respondents and at least 300 answers can be used for analyses between blocks. In the final data set, answers are missing by design, because each respondent is answering only parts of the questionnaire depending on the respective matrix group.

Table 3. Matrix design and follow-up

Respondent groups	Blocks				
	Core	A	B	C	D
1	X	X	X	O	O
2	X	X	O	X	O
3	X	X	O	O	X
4	X	O	X	X	O
5	X	O	X	O	X
6	X	O	O	X	X

Note: X: matrix first round; O: matrix follow-up.

A set of decision rules was applied to split the questionnaire into blocks. These decision rules aimed at providing respondents with a questionnaire that made sense in its thematic order and resulted in a positive survey experience. In addition, the selection of questions for the core and the thematic blocks was intended to enable the later use of imputation methods to handle the substantial share of values that were missing by design (see Table 2). Accordingly, the core included key items from each thematic block. The following decision rules were applied:

1. Each questionnaire had to be meaningful to the respondents in its question order.
2. The question order from the F2F questionnaire had to be retained.
3. If variables were known to be often analysed together, they were placed in the same block.
4. The block of background variables (core) included not only the major socio-demographic information but also some widely used substantial variables. Less used socio-demographic variables were placed in the thematic blocks.

The assignment of EVS items to blocks is reported in the ZA7500 Matrix Design Blocks (EVS, 2020b).

Follow-up Survey

The application of a matrix design resulted in many missing values by design. A potential solution to this issue that required testing—apart from fielding the whole 1-hour questionnaire—was to conduct a follow-up survey in which each respondent was asked the remaining blocks. Table 3 illustrates such a follow-up design, where circles symbolize blocks answered in a second survey that were missing by design in the initial survey.

The EVS Methodology Group proposed the use of a follow-up survey in countries that employed the matrix

questionnaire. However, it needs to be noted that respondents were free to refuse to participate in that survey or to not participate due to other reasons (e.g., non-contact). Using a follow-up survey, thus, introduced a further form of missingness in the data: panel attrition.

Country-Specific Designs

Six countries (Denmark, Finland, Germany, Iceland, the Netherlands, and Switzerland) chose to field the EVS 2017 in self-administered mode(s) in parallel to the F2F mode. Due to the various ways of fielding, the self-administered survey outlined above and lacking research in this regard, the countries explored different implementations. Four of them (Germany, Iceland, the Netherlands, and Switzerland) administered matrix questionnaires. Three of these countries (Iceland, the Netherlands, and Switzerland) further implemented a follow-up survey to complement the matrix design. Three countries (Germany, Iceland, and Switzerland) fielded full-length questionnaires in parallel to a matrix questionnaire. Two countries (Denmark and Finland) decided to only field the full-length questionnaire.

While the design of the matrix was coordinated, some decisions on survey characteristics such as contact procedures, mode choice sequence, and incentives were

left to the countries, based on their resources and previous field experience to reflect the country-specific context and survey climate. Table 4 summarizes the design per country.

Denmark

In Denmark, the sampling and fieldwork of EVS were carried out by Statistics Denmark. A stratified simple random sample was drawn from the national population register, and sampled individuals were then randomly assigned to be invited to participate either in the F2F or self-administered survey.

While the fieldwork for the F2F mode took place between 27 September 2017 and 31 January 2018, the fieldwork of the self-administered survey started later (11 December 2017) and ended alongside the F2F one. No matrix design was implemented. Thus, respondents only received the full-length EVS questionnaire. Respondents assigned to the self-administered mode received an advance letter and leaflet with the link to the web survey. The first postal reminder included also the paper-based questionnaire. The second reminder, as a final contact attempt, was made by telephone. A monetary incentive conditional upon the completion of the survey was offered to respondents. Among the 4,004

Table 4. Main country-specific design features of the EVS 2017 mixed-mode field

Design Feature	DK	FI	DE	IS	NL	CH
Mode(s)	Web and paper	Web and paper	Web and paper	Web and paper	Web only	Web and paper
Contact design	Invitation by letter; push to web: paper with 1st reminder	Invitation by letter, push to web, paper on request.	Invitation by letter; *MM matrix:* simultaneous (paper, web) vs. sequential (push to web) 3/4 contacts (phase 1/2) *MM full:* simultaneous (paper, web) 3 contacts	Invitation by letter; paper only on request	Invitation by e-mail;	Invitation by letter, push to web: paper with 2nd reminder
Type of sample	Random, individuals, register data	Random individuals, register data	Random, individuals, register data	Random, individuals, register data	Random selection among LISS panel participants	Random, individuals, register data
Incentive	Conditional monetary incentive	Conditional lottery of gift cards (10 × €100 and 1 × €500)	*MM matrix:* 5€ unconditional vs. 10€ conditional *MM full:* 5€ unconditional	Conditional lottery (10 × 10,000 ISK/ 63€ and 1 × 100,000ISK/ 635€)	Standard LISS panel (15€ per hour of survey completion)	9€ (10 CHF) unconditional + conditional lottery for FU (3 iPads)

respondents selected for the self-administered mode, 1,255 completed the survey via web, and 411 respondents completed the paper-based questionnaire.

Finland

In Finland, the F2F survey was complemented by a self-administered CAWI survey that featured the whole 1-hour questionnaire. The sampling frame in both cases was the Finnish population register. For the F2F survey, a two-staged stratified random sampling was used. The strata were living area and degrees of urbanization. The CAWI survey relied on one-stage stratified random sampling with the same strata.

All respondents received an invitation letter to participate in the survey. The F2F participants were informed that they will be contacted by interviewers. The CAWI participants were asked to use the provided web link and fill in the survey. As an incentive, a lottery was organized so that 11 gift cards (10 × 100€; 1 × 500€) were drawn among all who completed the interview.

The gross sample for the CAPI was 1,037 and the fieldwork took place between 3 January and 10 July 2018. With a minimum of four contact attempts, 388 interviews were achieved. The gross sample for the CAWI was 4,209 and the fieldwork period was 24 November 2017 to 10 July 2018. There were three rounds of invitation letters and two rounds of phone reminders. An additional option to respond by paper-based questionnaire, sent by mail, was provided for those who requested it. In the end, 811 self-administered interviews were achieved (CAWI $N = 668$ and paper-based $N = 143$).

Germany

In addition to the F2F survey, Germany conducted two surveys in self-administered mixed-mode (CAWI and paper-based): the first survey was performed using the matrix design and carried out using a responsive survey design with experiments on mode choice sequence (simultaneous vs. sequential) and incentives (10€ conditional vs. 5€ unconditional). The second survey was administered using the full-length questionnaire. In this latter case, a simultaneous mode choice sequence was offered to respondents along with a 5€ unconditional incentive; this approach turned out to reach the highest response rate in the first survey. Respondents in the F2F survey received a 10€ conditional incentive what was later changed to 5€ unconditional to increase outcome rates.

In both surveys, respondents in the simultaneous mode choice sequence were offered paper-based and CAWI questionnaires right from the start. Respondents in the sequential mode choice sequence ('push to web') were offered CAWI questionnaires first, and paper-based questionnaires later with the second reminder.

All three surveys used the same probability-based sample that was drawn from German municipalities' population registers and then randomly split up into separate samples for each survey. The gross sample sizes for each survey were 5,833 (CAPI), 9,369 (mixed-mode matrix), and 2,106 (mixed-mode full length) cases.

The German EVS surveys were carried out from 23 October 2017 to 4 April 2018 (F2F, $N = 1,494$), from 16 November 2017 to 20 March 2018 (mixed-mode matrix, $N = 3,237$), and from 20 September 2018 to 28 November 2018 (mixed-mode full length, $N = 676$).

Iceland

In Iceland, a stratified random sample of 2,322 individuals was drawn from the national population register for the F2F survey and another simple random sample of 3,500 individuals for the CAWI survey. The sample for the CAWI survey was randomly split into seven groups of 500 individuals each. One group was asked to finish the full-length questionnaire, the remaining six groups received different versions of the matrix design questionnaire. Those who completed a matrix design questionnaire were then invited to participate in a follow-up survey.

Advance letters were sent to all 5,822 individuals in the samples with a web link and invitation to the CAWI sample to answer online. Following a given time allowing participants to respond, contact attempts via phone were made to those who had not at that time answered the questionnaire. After four unsuccessful contact attempts made to non-respondents, a final reminder was sent to their address along with a full-length paper-based questionnaire.

In Iceland, there was no unconditional incentive, however, 11 participants were randomly drawn from those that completed the questionnaire: 10 receiving a lottery prize of 10,000 ISK (~63€) each and one receiving 100,000 ISK (~635€).

F2F interviews were administered between 19 June 2017 and 28 March 2018 ($N = 915$), whereas the self-administered interviews were collected between 27 September 2017 and 4 April 2018 ($N = 1,379$ matrix design; $N = 217$ full length). The follow-up survey was completed by 472 respondents.

The Netherlands

In the Netherlands, the F2F survey was complemented by a self-administered CAWI survey that featured matrix design questionnaires. Those respondents who

completed the matrix questionnaire were invited to a follow-up survey.

The Central Bureau of Statistics (CBS) provided a stratified random sample of 1,500 respondents, extracted from the national population register, for the F2F survey. For the CAWI survey, the country team relied on the Longitudinal Internet Studies for the Social Sciences (LISS) online panel, managed by CentERdata. A sample of 2,515 respondents for the CAWI survey was drawn from the members of the LISS panel (which consisted of 5,000 households and approximately 7,000 individuals as of July 2017). The LISS panel relies on probability samples of households which were drawn, similarly to the F2F EVS sample, from the national population register and provided by CBS. For households which are not online at the time of recruitment, an Internet connection and a computer are provided. For more information on sampling and recruitment strategies, see: www.lissdata.nl.

The F2F interviews were conducted between 31 August 2017 and 28 February 2018. The incentive strategy changed during the fieldwork: initially, all persons within the sample received an unconditional 5€ gift card together with the advance letter, and a 10€ gift card was sent to them after successful completion of the interview. At about three-quarters of fieldwork, interviewers were granted the possibility to award respondents with a 20€ gift card if they deemed it necessary to enhance co-operation. In total, 686 respondents participated in the F2F survey.

Between 11 September and 31 October 2017, 2,053 respondents took part in the CAWI data collection. Contact attempts (invitation and up to two reminders) were only made via email, no advance letters or brochures were used. In accordance with the LISS default implementation, conditional monetary incentives were offered to respondents. More specifically, panel members were rewarded with 15€ per hour of completed survey time.

Respondents who participated in the CAWI survey were later invited to participate in the follow-up survey to complete the matrix, which was carried out between 1 January and 30 January 2018. The contact design was the same as in the first round of data collection. Out of the 2,053 respondents who completed the CAWI survey, 1,722 respondents completed also the follow-up survey.

Switzerland

In Switzerland, additionally to the traditional F2F survey, respondents who were selected to participate via self-administered modes (CAWI and paper-based)

received a matrix questionnaire or the full-length questionnaire. Those who received a questionnaire in matrix design then received a follow-up survey. Moreover, there were two versions of the full-length questionnaire: one with the original question order, and one with a modified question order.

Separate simple random samples were drawn from the individual-based register frame of the national statistical office for each survey. The F2F survey was based on a gross sample of 1,400 units, and the self-administered surveys on independent gross samples of overall 6,800 units (4,800 for the matrix and 2,000 for the full-length questionnaire).

The contact procedure was held as similar as possible among the different groups: all sample units received an unconditional 10 CHF (∼9€) value incentive in form of a postal check along with the invitation letter. This letter announced the visit of an interviewer or gave the credentials for the web participation. In case of non-response, at least four contact attempts were being made or reminders sent. The paper-based questionnaire was sent to all CAWI non-respondents along with the second reminder. In order to motivate respondents to also answer the follow-up questionnaire, three iPads were drawn among the respondents. The CAWI data collection spanned between 15 September and 28 December 2017, while the paper-based questionnaires were distributed and collected between 30 September 2017 and 22 February 2018 ($N = 2,129$ matrix design; $N = 858$ full length; 1,661 cases participated in the follow-up survey). F2F interviews were conducted between 11 September 2017 and 22 February 2018 ($N = 673$).

Results

In this section, we provide readers with a brief introduction on how the different modes of the EVS performed in the countries that participated in the mixed-mode experiments with respect to outcome rates, representation of socio-demographic variables, and data quality.

Outcome Rates

We report the AAPOR Response Rate 6 (AAPOR, 2016) as an indicator of how well the different surveys performed with respect to achieving the required number of cases (Table 5). The EVS has its own categories for (non)response, the interested reader finds the correspondence between EVS and AAPOR categories in the Appendix, Table A1.

The F2F survey achieved different response rates across countries: While Finland and Germany showed

Table 5. AAPOR Response rates (RR6) and size of final sample for the EVS 2017 mixed-mode field

Country	DK	FI	DE	IS	NL	CH
CAPI	50.6% $N = 1,696$	27.2% $N = 388$	28.0% $N = 1,494$	41.0% $N = 915$	47.8% $N = 686$	50.6% $N = 673$
MM matrix	—	—	36.1% $N = 3,237$	46.8% $N = 1,379$	81.6% $N = 2,053$	44.4% $N = 2,129$
MM[a] matrix FU	—	—		(16% of resp) $N = 472$	(85% of resp) $N = 1,722$	(78% of resp) $N = 1,661$
MM full length	41.9% $N = 1,666$	19.4% $N = 811$	35.3% $N = 676$	44.5% $N = 217$	—	42.9% $N = 858$
Share of paper[b]	25%	18%	74%[c]	5%	—	29%

[a]Matrix FU: People who answered to the MM matrix were invited in a second phase to complete the remaining part of the questionnaire as a follow-up. Here, we considered only those who answered to this second part.

[b]Share of paper: the way of offering the paper mode strongly differed from country to country, depending on respective survey climate and practices (Table 4).

[c]Differs strongly by experimental condition: 83.9 per cent for MM full-length (all simultaneous contact mode), 72.7 per cent for MM matrix (simultaneous and sequential/push to web contact modes combined), and 82.5 per cent for MM matrix (only simultaneous, all age groups), respectively, 32.0 per cent for MM matrix (only sequential, 18-59 years).

the lowest response rates (below 30 per cent), Denmark, the Netherlands, and Switzerland report response rates of around 50 per cent, and Iceland lies in-between with 40 per cent response rate for F2F. Comparing response rate outcomes between F2F and self-administered modes, we found that some countries achieved a higher response rate for self-administered modes than for F2F (DE, IS, NL), while others reported a lower response rate for self-completion (DK, FI, CH). The full-length mixed-mode survey achieved lower response rates than the matrix mixed-mode in all countries that fielded both designs; however, the differences were very small.

Two differences stand out: In the Netherlands, the matrix mixed-mode achieved a much higher response rate than the F2F survey. This outcome was the result of using panelists from the LISS panel and is not strictly comparable to the other countries. In Finland, the response rate of the F2F survey was already low but participation in the full-length mixed-mode survey was even lower, resulting in the lowest response rate across countries and designs.

All self-completion designs except for the Netherlands included both web and paper-based versions of the questionnaire. The way of offering paper-based questionnaires to respondents differed between countries (see Table 4). Comparing the share of paper in those mixed-mode designs, we found a considerable difference between countries. Depending on when the paper-based questionnaire was offered led to different shares of participations via paper-based questionnaires. The earlier it was offered and the easier it was to get the paper-based questionnaire, the more respondents participated this way (note that based on our study, we

cannot rule out country-differences in general preference for the paper mode by country).

Representation of Socio-Demographics

Another aspect of data quality is non-response bias, which can differ between modes. In the following tables, we compare the distributions of basic socio-demographic information of the achieved samples across countries and designs to the same distribution in the population. Pearson's Chi-square tests were performed in order to detect significant differences between the samples and the population for each category of the socio-demographic characteristic (sources for population data are reported in the Appendix, Table A2). With this analysis, we aimed at identifying under- or over-represented groups.

Age

Overall, there seemed to be an over-representation of older age groups and an under-representation of younger generations (see Table 6). In Denmark, Germany, and the Netherlands, the F2F sample performed better in terms of representation of age than the mixed-mode samples, as there were fewer significant differences from the population distribution. In Iceland, the mixed-mode full-length sample provided the best representation of the population in terms of age categories.

Sex

Overall, samples appeared to be balanced when it comes to the distribution of sex. Only in the Netherlands and

Table 6. Representation bias of age categories. Chi-square test comparison (baseline: population data)

Country/age categories	Population	Face-to-face	Mixed-mode full	Mixed-mode matrix	Mixed-mode matrix FU
Denmark		N = 1,695	N = 1,663		
18–24	11.5%	10.0%*	8.1%***		
25–34	15.7%	12.6%***	8.7%***		
35–44	15.5%	15.9%	13.6%*		
45–54	17.7%	17.4%	21.7%***		
55–64	15.4%	14.9%	19.4%***		
65–74	14.0%	18.9%***	18.8%***		
Over 75	10.1%	10.3%	9.7%		
Finland		N = 388	N = 776		
18–24	10.3%	5.9%**	10.3%		
25–34	15.9%	10.6%**	11.1%***		
35–44	15.1%	10.6%*	11.5%**		
45–54	16.1%	12.4%*	15.6%		
55–64	16.6%	20.6%*	18.9%		
65–74	14.7%	25.3%***	22.6%***		
Over 75	11.3%	14.7%*	10.1%		
Germany		N = 1,494	N = 667	N = 3,198	
18–24	8.9%	9.1%	6.4%*	7.5%**	
25–34	15.5%	14.3%	13.6%	12.1%***	
35–44	14.8%	14.6%	12.4%	14.8%	
45–54	19.5%	18.4%	19.6%	18.9%	
55–64	16.7%	18.3%	20.7%**	18.3%*	
65–74	12.1%	14.7%**	16.5%***	15.6%***	
Over 75	12.5%	10.6%*	10.6%	12.7%	
Iceland		N = 915	N = 207	N = 1,362	N = 472
18–24	17.4%	10.7%***	13.0%	11.8%***	7.0%***
25–34	18.1%	16.6%	14.0%	16.9%	11.0%***
35–44	16.7%	16.3%	17.9%	17.3%	15.3%
45–54	15.7%	18.6%*	16.9%	19.6%***	18.6%
55–64	14.7%	17.6%*	20.3%*	17.3%**	23.3%***
65–74	10.1%	11.9%	12.6%	12.0%*	19.1%***
Over 75	7.4%	8.3%	5.3%	5.2%**	5.7%
The Netherlands		N = 686		N = 2,053	N = 1,722
18–24	10.8%	9.0%		6.0%***	5.2%***
25–34	15.7%	15.3%		13.5%**	11.3%***
35–44	15.0%	17.3%		15.5%	14.3%
45–54	18.5%	18.1%		14.4%***	13.6%***
55–64	16.6%	16.9%		18.6%*	20.2%***
65–74	13.6%	15.5%		21.8%***	24.3%***
Over 75	9.9%	7.9%		10.2%	11.0%
Switzerland		N = 671	N = 858	N = 2,129	N = 1,664
18–24	9.8%	10.9%	11.8%*	10.1%	9.2%
25–34	16.9%	16.1%	13.9%*	16.5%	16.1%
35–44	17.0%	13.3%*	16.4%	17.6%	17.1%
45–54	19.0%	20.1%	21.9%*	18.5%	18.3%
55–64	15.3%	18.3%*	15.6%	17.1%*	17.8%**
65–74	11.8%	10.9%	12.4%	11.7%	12.6%
Over 75	10.3%	10.4%	8.1%*	8.6%*	8.9%

Note: χ^2 Test of difference: $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, see Appendix for the sources of population data.

Table 7. Representation bias of gender. Chi-square test comparison (baseline: population data)

Country/sex categories	Population	Face-to-face	Mixed-mode full	Mixed-mode matrix	Mixed-mode matrix FU
Denmark		$N = 1,695$	$N = 1,663$		
Men	49.4%	49.5%	47.4%		
Women	50.6%	50.5%	52.6%		
Finland		$N = 388$	$N = 809$		
Men	48.8%	51.5%	45.9%		
Women	51.2%	48.5%	54.1%		
Germany		$N = 1,494$	$N = 676$	$N = 3,225$	
Men	49.1%	49.9%	50.6%	49.8%	
Women	50.9%	50.1%	49.4%	50.2%	
Iceland		$N = 915$	$N = 208$	$N = 1,363$	$N = 471$
Men	48.7%	49.6%	44.7%	46.5%	46.5%
Women	51.3%	50.4%	55.3%	53.5%	53.5%
The Netherlands		$N = 686$		$N = 2,053$	$N = 1,722$
Men	49.2%	50.6%		44.4%***	45.1%***
Women	50.8%	49.4%		55.6%***	54.9%***
Switzerland		$N = 673$	$N = 858$	$N = 2,129$	$N = 1,664$
Men	49.2%	49.2%	49.4%	46.2%**	45.8%**
Women	50.8%	50.8%	50.6%	53.8%**	54.2%**

Note: χ^2 Test of difference: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, see Appendix for the sources of population data.

in Switzerland, the mixed-mode matrix design samples slightly over-represented women (see Table 7).

Higher educational level

We found a representation bias regarding education in all countries and almost all samples in the self-administered modes. Persons with a high education level (tertiary education)[1] were over-represented while persons with a low and/or persons with a middle level of education are under-represented. Noticeably, this also applied to the F2F sample, albeit less pronounced compared to self-administered modes (see Table 8). The Icelandic mixed-mode full-length sample and the Swiss F2F sample performed well in terms of representation.

Attrition in Follow-up Surveys

Panel surveys are challenged with attrition (Lynn, 2009). Panel attrition is the non-participation of respondents who have participated in previous waves of the survey. It not only lowers the statistical power of the data due to smaller net sample sizes, but it might also introduce a bias if respondents systematically differ from non-respondents (e.g., Kreuter and Olson, 2011; Gummer and Roßmann, 2019).

Panel attrition might be a challenge for the follow-up surveys conducted as part of the EVS. Therefore, we estimated logistic regression models with participation

in the follow-up survey (0 = no, 1 = yes) as a dependent variable for those cases who had received a matrix questionnaire in self-administered modes in Iceland, the Netherlands, and Switzerland. As independent variables, we selected a set of variables, some of which we have presented in the previous section and which we assumed to relate to the likelihood of participating in the follow-up (descriptive statistics for these variables are provided in the Appendix, Table A3). The results of each regression are presented in Table 9.

Across all countries, we found older respondents to be more likely to participate in a follow-up survey compared to younger respondents. In both Switzerland and Iceland, we further found that respondents who held the national citizenship were more likely to participate in the follow-up survey. Apart from these effects, some country-specific effects were significant (e.g., education in Iceland, paper-based surveys in Switzerland). Generally, we observe a tendency for selective participation in the second wave of the mixed-mode samples, based on the variables we investigated. While our analyses revealed a few significant effects, the models' Pseudo R^2 hinted towards a rather low explanatory power of the selected variables. In other words, with respect to the variables we included in our models, the degree of systematic attrition should not be over interpreted.

We recommend that researchers using data from the follow-up surveys should evaluate, based on their

Table 8. Representation bias of educational attainment. Chi-square test comparison (baseline: population data)

Country/education categories	Population	Face-to-Face	Mixed-mode full	Mixed-mode matrix	Mixed-mode matrix FU
Denmark		N = 1,684	N = 1,599		
ISCED Level 0–2	27.0%	17.9%***	18.7%***		
ISCED Level 3–4	41.2%	32.5%***	30.3%***		
ISCED Level 5+	31.8%	49.6%***	51.0%***		
Finland		N = 388	N = 801		
ISCED Level 0–2	25.8%	21.4%*	14.7%***		
ISCED Level 3–4	42.5%	41.8%	37.5%*		
ISCED Level 5+	31.7%	36.9%*	47.8%***		
Germany		N = 1,493	N = 662	N = 3,192	
ISCED Level 0–2	18.7%	11.5%***	13.6%***	17.6%	
ISCED Level 3–4	57.0%	53.7%**	47.4%***	41.9%***	
ISCED Level 5+	24.3%	34.8%***	39.0%***	40.4%***	
Iceland		N = 910	N = 207	N = 1,357	N = 471
ISCED Level 0–2	31.0%	22.9%***	25.6%	23.8%***	22.1%***
ISCED Level 3–4	35.5%	33.4%	36.7%	36.0%	32.9%
ISCED Level 5+	33.5%	43.7%***	37.7%	40.2%***	45.0%***
The Netherlands		N = 683		N = 1,994	N = 1,687
ISCED Level 0–2	31.9%	29.0%		29.5%*	30.6%
ISCED Level 3–4	39.0%	31.2%***		27.1%***	26.3%***
ISCED Level 5+	29.1%	39.8%***		43.4%***	43.1%***
Switzerland		N = 671	N = 812	N = 2,039	N = 1,641
ISCED Level 0–2	16.3%	16.2%	16.7%	15.7%	15.0%
ISCED Level 3–4	48.6%	47.1%	41.4%***	40.1%***	39.9%***
ISCED Level 5+	35.1%	36.7%	41.9%***	44.1%***	45.1%***

Note: χ^2 Test of difference: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, see Appendix for the sources of population data.

research questions, if the attrition process relates to their selected dependent and independent variable(s) to assess whether correction methods such as weighting might be applied (cf. Kreuter and Olson, 2011).

Data Quality

F2F surveys are assumed to be producing high-quality data—partly due to a low share of item non-response. The assumption is that interviewers can motivate respondents to answer and help when a question is ambiguously phrased or hard to answer. However, interviewer presence might also lead to response errors (i.e., interviewer effects) such as socially desirable responding, which in turn might bias the data.

In self-administered surveys, no interviewer is present. Especially when answering a survey via a paper-based questionnaire, questions can easily be skipped. Moreover, no interviewer is available to record spontaneous reactions of respondents regarding the questionnaire or to answer questions of the respondent.

Previous research has used item non-response as an indicator of low data quality that is the result of a lack of motivation or ability of the respondent to answer a

question (Toepoel, Das and Van Soest, 2009; Lenzner, 2012). Similar arguments have been put forward with regard to survey break-offs where respondents provide only partially completed questionnaires (Roßmann, Blumenstiel and Steinbrecher, 2015).

To investigate whether the use of self-administered modes resulted in larger amounts of missing answers and, thus, might indicate issues with data quality, we calculated average item non-response rates for each design for each country. For the purpose of this analysis, we counted 'don't know' and 'I prefer not to answer' as item non-response. The item non-response *rates* were calculated by dividing the amount of item non-response by the actual number of items a respondent received, varying between 144 and 271 (depending on the survey design—full length or matrix—the respective matrix questionnaire version, and filtered questions). Table 10 summarizes our findings.

Overall, item non-response was low in most instances. Across all countries, the paper-based mode had a larger share of item non-response compared to the CAWI. In Germany, where the mixed-mode had a high share of participants via paper-based questionnaire, this

Table 9. Panel attrition—logit coefficients of logistic regression models on completing the follow-up in Iceland, the Netherlands, and Switzerland

Independent variables	IS		NL		CH	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Constant	−2.736***	(0.541)	0.182	(0.462)	0.267	(0.303)
Male	Ref.		Ref.		Ref.	
Female	0.030	(0.122)	−0.005	(0.133)	0.158	(0.117)
18–24 years	Ref.		Ref.		Ref.	
25–34 years	0.060	(0.280)	−0.026	(0.260)	0.199	(0.224)
35–44 years	0.370	(0.297)	0.378	(0.271)	0.528*	(0.251)
45–54 years	0.553	(0.294)	0.584*	(0.293)	0.712**	(0.257)
55–64 years	1.169***	(0.298)	1.581***	(0.320)	1.055***	(0.272)
65–74 years	1.583***	(0.311)	1.939***	(0.331)	1.317***	(0.305)
75 and more years	0.815*	(0.395)	1.640***	(0.374)	1.334***	(0.326)
Lower educational level	Ref.		Ref.		Ref.	
Medium educational level	0.236	(0.167)	0.036	(0.185)	0.167	(0.164)
Higher educational level	0.551**	(0.172)	0.153	(0.175)	0.245	(0.189)
Holding non-national citizenship	Ref.		Ref.		Ref.	
Holding national citizenship	1.323**	(0.446)	0.673	(0.347)	0.555***	(0.146)
Town under 5000 inhabitants	Ref.				Ref.	
5000–20,000	−0.292	(0.182)			−0.128	(0.138)
20,000–100,000	−0.283	(0.188)			−0.052	(0.177)
City over 100,000 inhabitants	−0.139	(0.167)			−0.147	(0.184)
Married/partnership	Ref.		Ref.		Ref.	
Widowed/Divorced/Separated	−0.029	(0.182)	−0.395*	(0.181)	−0.072	(0.174)
Never married	0.177	(0.205)	0.073	(0.169)	0.355*	(0.168)
CAWI mode	Ref.				*Ref.*	
Mail mode	−0.037	(0.343)			−0.786***	(0.128)
McFadden's Pseudo R^2	0.058		0.044		0.071	
Observations	1,316		2,028		1,970	

*Note: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$; coeff., logit; SE, standard error.*

Table 10. Average item non-response rates across 277 EVS questions

Subsample	DK	FI	DE	IS	NL	CH
	$N = 3,362$	$N = 1,199$	$N = 5,407$	$N = 2,503$	$N = 2,728$	$N = 3,606$
Mode						
F2F	0.76%	2.31%	2.07%	2.26%	1.93%	2.77%
CAWI only[a]	0.54%	0.67%	5.07%	1.31%	4.66%	1.82%
Paper-based only[a]	6.95%	12.51%	8.30%	12.27%	—	3.50%
Mixed-mode design						
MM full length	2.12%	2.76%	8.01%	2.86%	—	2.25%
MM matrix	—	—	7.37%	1.69%	4.66%	2.34%
MM matrix FU	—	—	—	1.40%	4.34%	1.78%

[a]The web and paper mode can be compared only within countries because of design differences. Figures are averaged item non-response rates across the whole questionnaire. Break-offs were excluded.

resulted in a comparatively large overall share of item non-responses. However, in Switzerland, Iceland, Denmark, and Finland, the share of item non-response is higher in the F2F mode compared to the CAWI. When comparing the different questionnaire designs, we found that the full-length questionnaire generally resulted in a higher share of item non-response compared to the matrix design. The low degree of item non-response was

replicated for the follow-up surveys conducted in Switzerland, Iceland, and the Netherlands.

EVS Data Releases

The EVS dataset is made freely available for teaching and research purposes via the data archive at GESIS—Leibniz Institute for the Social Sciences. Datasets are complemented with comprehensive documentation allowing researchers to discover and reuse the data. The documentation covers information on project, study, and variable level. With respect to the self-administered mixed-modes in several countries, the documentation further includes scenarios and recommendations on how to select the appropriate subsamples for analyses.

The de-facto anonymized EVS 2017 survey data are available as two off-site Scientific-Use File:

(1) Integrated Dataset (EVS 2017): *This dataset (EVS, 2020b) contains only cases who have completed the questionnaire. More specifically, the following cases are included: (i) all F2F interviews from currently 34 EVS 2017 countries; (ii) all the full-length interviews collected via self-administered mixed-mode (Denmark, Finland and full-length questionnaires from Switzerland, Iceland, and Germany); (iii) all respondents who received a matrix questionnaire and completed the follow-up survey (a selection of cases from Iceland, the Netherlands, and Switzerland who received the full questionnaire due to the combination of matrix and follow-up questionnaires). The current version of this dataset contains data from 56,491 respondents and over 450 variables, including calibration weights for each country (based on age, sex, education, and region).*
(2) Integrated Dataset (EVS 2017)—Matrix Design Data: *This dataset (EVS, 2020c), which can be easily merged to dataset (1), contains all data coming from the self-administered mode in countries that implemented the matrix design, that is 10,598 interviews from Germany, Iceland, the Netherlands, and Switzerland. Respondents who filled in the full-length questionnaire are included too, as well as—for methodological purposes—break-off cases (namely cases that filled in less than 50 per cent of the questions). Cases that appear in both (1) and (2) are marked with a flag variable (fduplicate). Instructions on how to merge the datasets are provided in the Guide to the Mixed-mode Approach and Matrix Design (EVS, 2020f), alongside a description of variables that can be used to navigate the design features*

(e.g., mixed-mode design, matrix design group, etc...) presented in this data brief.

Additionally, sensitive data will be made accessible under contractual regulations. The Secure Data Center (SDC) at GESIS will grant access to the Scientific-Use File via 'contracted off-site' or 'on-site' usage.

A further dataset, the *European Values Study Longitudinal Data File 1981–2008* (EVS, 2020d), gathers data from all the EVS waves from 1981 to 2008; an updated version including all data from the (1) EVS 2017 Integrated Dataset will be made available in 2021. Additionally, a joint dataset containing the EVS 2017 data together with the last wave of WVS will be available at the end of 2020.

Discussion

With the present data brief, we introduced researchers to methodological innovations implemented in the latest wave of EVS. Unlike previous waves, the EVS 2017 includes the experimental test of self-administered modes that aimed at investigating possibilities to tackle the ongoing challenges of decreasing response rates and increasing survey costs. Coping with these challenges seems essential for a large-scale cross-cultural survey programme such as the EVS because these survey programmes heavily depend on the participation of a large and diverse set of countries. The increasing fieldwork and data processing costs risk limiting countries' participation in the future. Furthermore, participation in surveys is related with items central to the EVS, including trust, inspiring alternative modes of data collection. In our data brief, we outlined the methodological challenges that motivated an openness for change in the EVS and facilitated testing in several participating countries. In our view, it is pivotal to make the reasoning behind design changes as well as the country-specific implementations transparent and show whether and how they influenced survey outcomes. The recent challenges for F2F data collection imposed by the COVID-19 pandemic (e.g., Gummer et al., 2020; Sakshaug et al., 2020) demonstrate the necessity for openness to methodological changes.

Yet, our contribution is not only targeted on other survey infrastructures but especially on users of EVS data and those researchers interested in learning about new developments. Indeed, the use of mixed-mode data collection in a large, cross-national survey opened up new possibilities for research. For instance, by incorporating the EVS in online panel surveys representative of

national populations, the potential arises to supplement existing repeated cross-sectional design into a panel structure. Such shift would not only allow for continuing the study of social change with great detail but also enables to assess the effect of certain events, for instance the aforementioned COVID-19 pandemic, on relevant values and attitudes.

Nevertheless, we urge EVS data users to pay careful attention to the survey design features described in this data brief when analysing EVS data. Potential mode, selection, and order effects should be considered when investigating substantive topics. As described in the survey documentation (EVS, 2020f)) a number of variables provided in the datasets can help acknowledging such effects and/or defining relevant subsamples for analyses.

Our findings highlight that employing self-administered modes was successful in most of the participating countries. Even fielding a 1-hour questionnaire worked reasonably well in those countries that tried it—in Germany, Iceland, and Switzerland, similar response rates could be achieved as in a shorter version of the questionnaire. Employing a matrix design yielded a substantive share of missingness by design. It remains to be seen whether this will impair the use of the data for substantive research, and standards to deal with this challenge still need to be established. In Iceland and Germany, the cheaper self-administered modes outperformed F2F mode in terms of outcome rates. Interestingly, in those four countries in which we widely distributed paper-based questionnaires, a non-neglectable proportion of respondents chose to use them. Our analyses on representation of socio-demographic distributions and item non-response further suggested that self-administered modes show slightly higher bias on average but—given the lower cost—still may yield acceptable data quality. However, non-response bias and more generally data quality of self-administered surveys have to be further studied. For this purpose, we invite methodologists to examine our data on the mode experiments and the matrix questionnaire that were collected as part of the EVS. Overall, we think that self-administered modes can complement the traditional F2F mode in large-scale population-wide surveys; especially if it is possible to reduce bias and further increase data quality, for example, by further improving the adaption of question design to self-administered modes.

Notes

1   The way country-specific educational categories have been translated into ISCED main levels is described in the ZA7500 Codebook Appendix A3 and A4 (cf. EVS, 2020b).

References

AAPOR (2016). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 9th edn. The American Association for Public Opinion Research.

Billiet, J. *et al.* (2007). Estimation of nonresponse bias in the European Social Survey: using information from reluctant respondents. *Journal of Official Statistics*, **23**, 135–162.

Brick, J. M. and Williams, D. (2013). Explaining rising nonresponse rates in cross-sectional surveys. *The Annals of the American Academy of Political and Social Science*, **645**, 36–59.

de Leeuw, E. and De Heer, W. (2002). Trends in household survey nonresponse: a longitudinal and international comparison. In Groves, R. M. *et al.* (Eds.), *Survey Nonresponse*. New York: Wiley & Sons, pp. 41–54.

de Leeuw, E., Hox, J. and Luiten, A. (2018). International nonresponse trends across countries and years: an analysis of 36 years of Labour Force Survey data. *Survey Methods: Insights from the Field*, 1–11..https://surveyinsights.org/?p=10452

EVS (2020a). *EVS-Bibliography 03/2020*, available from: https://europeanvaluesstudy.eu/wp-content/uploads/2020/03/20200302_EVS_Bibliography.pdf [accessed 3 April 2020].

EVS (2020b). *European Values Study 2017: Integrated Dataset (EVS 2017)*. GESIS Data Archive, Cologne. ZA7500 Data file Version 4.0.0. doi:*10.4232/1.13560*.

EVS (2020c). *European Values Study 2017: Integrated Dataset (EVS 2017)—Matrix Design Data*. GESIS Data Archive, Cologne. ZA7502 Data file Version 2.0.0. doi:*10.4232/1.13561*.

EVS (2020d). *European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008)*. GESIS Data Archive, Cologne. ZA4804 Data file Version 3.1.0. doi:10.4232/1.13486

EVS (2020f). European Values Study (EVS) 2017: Guide to the Mixed-Mode Approach and Matrix Design. (GESIS Papers, 2020/14). Köln. https://doi.org/10.21241/ssoar.70112 [accessed 21 October 2020].

EVS (2020e). European Values Study (EVS) 2017: Methodological Guidelines. (GESIS Papers, 2020/13). Köln. 10.21241/ssoar.70110 (url: https://www.ssoar.info/ssoar/handle/document/70110) [accessed 21 October 2020].

Groves, R. M. and Heeringa, S. G. (2006). Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **169**, 439–457.

Groves, R. M. and Peytcheva, E. (2008). The impact of nonresponse rates on self-selection bias a meta-analysis. *Public Opinion Quarterly*, **72**, 167–189.

Gummer, T. and Roßmann, J. (2019). The effects of propensity score weighting on attrition biases in attitudinal, behavioral, and socio-demographic variables in a short-term web-based panel survey. *International Journal of Social Research Methodology*, **22**, 81–95.

Gummer, T. *et al*. (2020). The impact of Covid-19 on fieldwork efforts and planning in pairfam and FReDA-GGS. *Survey Research Methods*, **14**, 223–227.

Inglehart, R. (1997). *Modernization and Postmodernization*. Princeton: Princeton University Press.

Klausch, T., Hox, J. J. and Schouten, B. (2013). Measurement effects of survey mode on the equivalence of attitudinal rating scale questions. *Sociological Methods & Research*, **42**, 227–263.

Kreuter, F. (2013). Facing the nonresponse challenge. *The Annals of the American Academy of Political and Social Science*, **645**, 23–35.

Kreuter, F. and Olson, K. (2011). Multiple auxiliary variables in nonresponse adjustment. *Sociological Methods & Research*, **40**, 311–332.

Lenzner, T. (2012). Effects of survey question comprehensibility on response quality. *Field Methods*, **24**, 409–428.

Lynn, P. (2009). *Methods for Longitudinal Surveys*. Chichester: John Wiley & Sons

Massey, D. S. and Tourangeau, R. (2013). Where do we go from here? Nonresponse and social measurement. *The Annals of the American Academy of Political and Social Science*, **645**, 222–236.

Norris, P. and Inglehart, R. (2019). *Cultural Backlash: Trump, Brexit, and Authoritarian Populism*. Cambridge, UK: Cambridge University Press.

Olson, K. *et al*. (2020). Transitions from telephone surveys to self-administered and mixed-mode surveys: AAPOR task force report. *Journal of Survey Statistics and Methodology; doi: 10.1093/jssam/smz062*.

Pforr, K. *et al*. (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**, 740–768.

Raghunathan, T. E. and Grizzle, J. E. (1995). A split questionnaire survey design. *Journal of the American Statistical Association*, **90**, 54–63.

Revilla, M. and Ochoa, C. (2017). Ideal and maximum length for a web survey. *International Journal of Market Research*, **59**, 557–565.

Roßmann, J., Blumenstiel, J. E. and Steinbrecher, M. (2015). Why do respondents break off web surveys and does it matter? Results from four follow-up surveys. *International Journal of Public Opinion Research*, **27**, 289–302.

Sakshaug, J. W. *et al*. (2020). Impacts of the COVID-19 pandemic on labor market surveys at the German Institute for Employment Research. *Survey Research Methods*, **14**, 229–233.

Steeh, C. G. (1981). Trends in nonresponse rates, 1952–1979. *Public Opinion Quarterly*, **45**, 40–57.

Stoop, I. A. *et al*. (2010). *Improving Survey Response: Lessons Learned from the European Social Survey*. Chichester: John Wiley & Sons.

Toepoel, V., Das, M. and Van Soest, A. (2009). Design of web questionnaires: the effects of the number of items per screen. *Field Methods*, **21**, 200–213.

Tourangeau, R. *et al*. (2017). Adaptive and responsive survey designs: a review and assessment. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **180**, 203–223.

Wagner, J. R. (2008). *Adaptive survey design to reduce nonresponse bias*. Ann Arbor, MI: University of Michigan (PhD thesis).

Ruud Luijkx (r.luijkx@tilburguniversity.edu) is an associate professor at the Department of Sociology of Tilburg University (the Netherlands) and a visiting professor at the Department of Sociology and Social Research of the University of Trento (Italy). He is the Chair of the Executive Committee and the Methodology Group of the European Values Study.

Guðbjörg Andrea Jónsdóttir (gudbjorg@hi.is) is director of the Social Science Research Institute at the University of Iceland. She is the National Programme Director of the EVS in Iceland and a member of the Executive Committee of EVS.

Tobias Gummer (tobias.gummer@gesis.org) is a senior researcher and team leader at GESIS—Leibniz Institute for the Social Sciences (Germany). He is Fieldwork Director of the EVS in Germany.

Michèle Ernst Stähli (michele.ernststaehli@fors.unil.ch) is head of Group International Surveys at FORS—Swiss Centre of Expertise in the Social Sciences (Switzerland). She is the National Programme Director of the EVS in Switzerland and co-designer of the matrix experiment.

Morten Frederiksen (mfr@socsci.aau.dk) is an associate professor at the Department of Sociology and Social Work, Aalborg University. He is the National Programme Director of the EVS in Denmark and head of Research of the Center for Inclusion and Welfare, Aalborg University. He is a member of the Executive Committee of EVS.

Kimmo Ketola (Kimmo.ketola@evl.fi) is a senior researcher at the Church Research Institute (Tampere, Finland). He is the National Programme Director of the EVS in Finland.

Tim Reeskens (t.reeskens@tilburguniversity.edu) is an associate professor at the Department of Sociology of Tilburg University (the Netherlands). He is also the National Programme Director of the EVS in the Netherlands, and is holder of the Jean Monnet Chair on Identities and Cohesion in a Changing European Union. He is a member of the Theory Group of EVS.

Evelyn Brislinger (evelyn.brislinger@gesis.org) is a senior researcher at GESIS—Leibniz Institute for the Social Sciences (Germany) and member of the Operational and Planning Group of the EVS.

Pablo Christmann (pablo.christmann@gesis.org) is a senior researcher at GESIS—Leibniz Institute for the Social Sciences (Germany) and member of the German EVS team.

Stefán Þór Gunnarsson (stefanthor@hi.is) is a project manager at the Social Science Research Institute at the University of Iceland. He is the primary Data Analyst of the EVS in Iceland.

Árni Bragi Hjaltason (arnibragi@hi.is) is a project manager at the Social Science Research Institute at the University of Iceland. He is the Fieldwork Director of the EVS in Iceland.

Dominique Joye (dominique.joye@unil.ch) is a professor emeritus at University of Lausanne (Switzerland). He is a member of the EVS Methodological Group and initiator of the matrix design experiment.

Vera Lomazzi (vera.lomazzi@gesis.org) is a senior researcher at GESIS—Leibniz Institute for the Social Sciences (Germany). She is the Secretary of the Executive Committee of EVS and member of the Operational and Planning Group of the EVS.

Angelica M. Maineri (a.m.maineri@tilburguniversity. edu) is a junior researcher/PhD student at the Department of Sociology of Tilburg University (the Netherlands). She is also an assistant of the Methodology Group and member of the Operational and Planning Group of the EVS.

Patricia Milbert (patricia.milbert@unil.ch) is an SNSF Doctoral Student at the University of Lausanne (Switzerland). She was a member of the Swiss EVS team and responsible for the grouping of the variables for the matrix design.

Michael Ochsner (michael.ochsner@fors.unil.ch) is a senior researcher at FORS—Swiss Centre of Expertise in the Social Sciences (Switzerland). He is a member of the Swiss EVS team and co-designer of the matrix experiment.

Alexandre Pollien (alexandre.pollien@fors.unil.ch) is a researcher at FORS—Swiss Centre of Expertise in the Social Sciences (Switzerland). He is Fieldwork Manager of the EVS in Switzerland and co-designer of the matrix experiment.

Marlène Sapin (marlene.sapin@fors.unil.ch) is a senior researcher at FORS—Swiss Centre of Expertise in the Social Sciences (Switzerland). She is a member of the Swiss EVS team and co-designer of the matrix experiment.

Ivet Solanes Ros (ivet.solanes@gesis.org) is a scientific administrator at GESIS—Leibniz Institute for the Social Sciences (Germany) working on EVS data processing.

Sascha Verhoeven (shl@posteo.de) was a senior researcher at GESIS—Leibniz Institute for the Social Sciences (Germany) and member of the German EVS team.

Christof Wolf (christof.wolf@gesis.org) is President of GESIS—Leibniz Institute for the Social Sciences (Germany) and holds the chair of Social Stratification at the University of Mannheim. He is the National Programme Director of the EVS in Germany and member of the EVS Executive Committee and Theory Group.

Appendix

Table A1. Correspondence EVS outcomes—AAPOR macro categories

EVS indicators (interviewer-administered)	AAPOR macro category	EVS indicators (self-administered)	AAPOR macro category
A. Total number of issued sample units (addresses, households, or individuals)	—	A. Total number of issued sample units (addresses, households, or individuals)	—
B. Refusal by respondent	R	B. Refusal	R
C. Other refusal (by proxy (or household or address refusal), language barrier)	R	B1. Explicit refusal	R
D. No contact (after at least four visits)	NC	B2. Implicit refusal (Logged on to survey, did not complete any items; Read receipt confirmation, refusal)	R
E. Respondent mentally or physically unable to co-operate throughout fieldwork period	O	C. Non-contact	NC
F. Respondent unavailable throughout the fieldwork period for other reasons	NC	D. Respondent was unavailable during field period | G (Mail): Physically or mentally unable/incompetent	NC
G. Address not residential (institution, business/industrial purpose)	NE	E. Completed questionnaire, but not returned during field period	NC
H. Address not occupied (not occupied, demolished, not yet built)	NE	F. Other	O
I. Address not traceable	NE	G. Language barrier	O
J. Other ineligible address	NE	H. Nothing known about respondent or address	UH
K. Respondent moved abroad/unknown destination	NE	I. No invitation sent	UH
L. Respondent deceased	NE	J. Nothing ever returned	UH
X units not accounted for	UO	K. Invitation returned undelivered	UO
Y. Invalid interviews	R	L. Invitation returned with forwarding information	UO
Z. Number of valid interviews	I	M. Other	UO
		N. Returned from an un-sampled email address	UO
		O. Selected Respondent Screened Out of Sample	NE
		P. Quota Filled	NE
		Q. Duplicate Listing	NE
		R. Other	NE
		S. Invalid interviews[a]	R
		T. Number of valid interviews	I
		X. Units not accounted for	UO
		Complete interviews	I
		Partial interviews	P

[a]In Germany, invalid interviews denote questionnaires returned by the wrong target person and are hence counted as NE.

I, Complete interviews; P, Partial interviews; R, Refusal and break-off; NC, Non-Contact; O, Other; UH, Unknown household; UO, Unknown other; NE, Not Eligible;