European Values Study (EVS) 2017: Guide to the Mixed-mode Approach and Matrix Design ; documentation update related to the final data release, May 2022

Veröffentlichungsversion / Published Version
Verzeichnis, Liste, Dokumentation / list

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

gesis **Leibniz-Institut für Sozialwissenschaften**

European *Values* Study

European Values Study (EVS) 2017

Guide to the Mixed-mode Approach and Matrix Design

Documentation update related to the final data release, May 2022

European Values Study (EVS)

GESIS Papers 2022|08

European Values Study (EVS) 2017

Guide to the Mixed-mode Approach and Matrix Design

Documentation update related to the final data release, May 2022

European Values Study (EVS)

Table of Contents

List of tables

List of figures

List of abbreviations

CAPI = Computer Assisted Personal Interview
TAPI = Tablet Assisted Personal Interview
PAPI = Paper-and-Pencil (personal) Interview
CATI = Computer-Assisted Telephone Interview
CAWI = Computer-Assisted Web Interview (also known as: Web survey)
Mail = Postal (paper-and-pencil) self-administered interview

Acknowledgements

1   Introduction

One of the most important innovations in the European Values Study (EVS) 2017 has been the introduction of the possibility for countries to field the EVS questionnaire using different modes of data collection (that is, a combination of face-to-face interviews with self-administered interviews). The following document outlines the main characteristics of the implementation of the mixed-mode strategy in the EVS 2017. The strategy was adopted, in different forms, by seven countries: Switzerland (CH), Iceland (IS), Netherlands (NL), Germany (DE), Finland (FI), Denmark (DK), and Latvia (LV).

Beware! It is recommended to pay attention to the mode structure when using EVS 2017 data, especially when comparing values across countries and/or across waves. Our preliminary checks did not show large differences in measurement between the different design elements, but we recommend testing several scenarios to verify that the results are stable and/or to use the appropriate 'mm_*' variable(s) (described in sections 3 to 8 in the following document) as control variables in multivariate models.

After describing the dataset structure, we will explain in more detail the main concepts related to the mixed-mode design and how they are implemented in the data. The sections are articulated as follows:

- *Section 2* describes the concept behind the composition of the two EVS datasets, the EVS Integrated dataset and the EVS Integrated dataset – Matrix Design Data;
- *Section 3* describes the possible combinations of modes used by the countries and how these can be identified in the datasets;
- *Section 4* focuses on the self-administered mode, its practical implementation by the EVS country teams, as well as how the different designs can be identified in the datasets;
- *Section 5* offers a more detailed account of the matrix design adopted by four of the countries and the related variables offered in the datasets;
- *Section 6* describes the responsive design employed in Germany and its related variables;
- *Section 7* offers an overview of the specific missing values that have been devised to better reflect the mixed-mode strategy features;
- *Section 8* offers a summary of the variables described in the various sections, as well as the introduction of a new variable ('mm_select_sample') which has been designed with the specific purpose of helping data users selecting sub-samples;
- Finally, *section 9* describes some possible scenarios for selecting subsamples based on the requirements of the data users. The list is not exhaustive, but it aims at providing some practical guidance on how to navigate the data

This document is intended as a practical guide to describe the way the mixed-mode strategy has been implemented in the EVS 2017 datasets. For more detailed information on the reasons behind the choice of using a mixed-mode approach and matrix design, and on the outcomes of this design, please see Luijkx et. al, (2021)[1].

---

[1]   Luijkx, R., Jónsdóttir, G.A., Gummer, T., Ernst Stähli, M., Frederiksen, M., […] Wolf, C. 2021. The European Values Study 2017: On the way to the future using mixed-modes, *European Sociological Review*.

2   Dataset structure

Alongside the EVS 2017 release, two different EVS survey datasets are provided: (1) EVS Integrated Dataset and (2) EVS Integrated dataset – Matrix Design Data. Cases are distributed into the datasets according to the characteristics of the survey design they were part of. As an overview, the distribution of cases into datasets based on mode, design and amount of non-substantive answers is summarized in Figure 1.

The distribution of cases into the datasets is based on several elements, namely:

- The mode of data collection, namely whether the survey was administered by an interviewer or whether it was self-completed (see Section 3);
- The design of the self-administered survey, namely whether a full-length questionnaire or a shorter version based on the matrix design was administered (see Section 4);
- Whether, within the matrix design framework, the respondent has filled in the whole survey or not (see Section 5);
- The overall design implemented in the country (see Sections 3 to 6);
- Whether the respondent has filled in at least 50% of the questions.

All these elements will be described more extensively in the next sections.
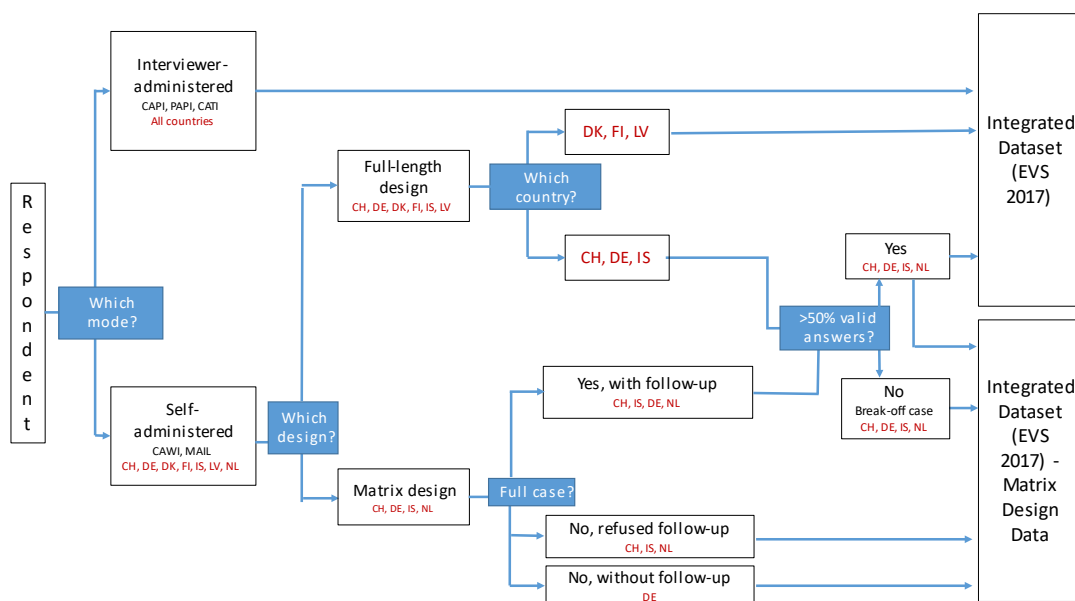
Figure 1:   Distribution of cases (respondents) into datasets according to the mixed-mode survey characteristics

2.1   EVS Integrated dataset

The **EVS Integrated dataset**, hereinafter referred to also as the **Main dataset**, is an anonymized dataset comprising data from 36 countries and 59,438 respondents. These are all the respondents from the interviewer-administered mode (CAPI, PAPI and CATI), and the **full cases** from the self-

administered (CAWI and MAIL) mixed-mode and matrix design mode (see Figure 2). Two criteria are used to identify the full cases:

a) the respondent has filled in the self-administered full-length questionnaire **OR** the respondent has filled in the first (first survey) and second round (follow-up) of the matrix design;

b) the respondent has "no answer" on less than 50% of the questionnaire items[2].

The use of the EVS Integrated dataset is recommended for substantive research involving EVS data.
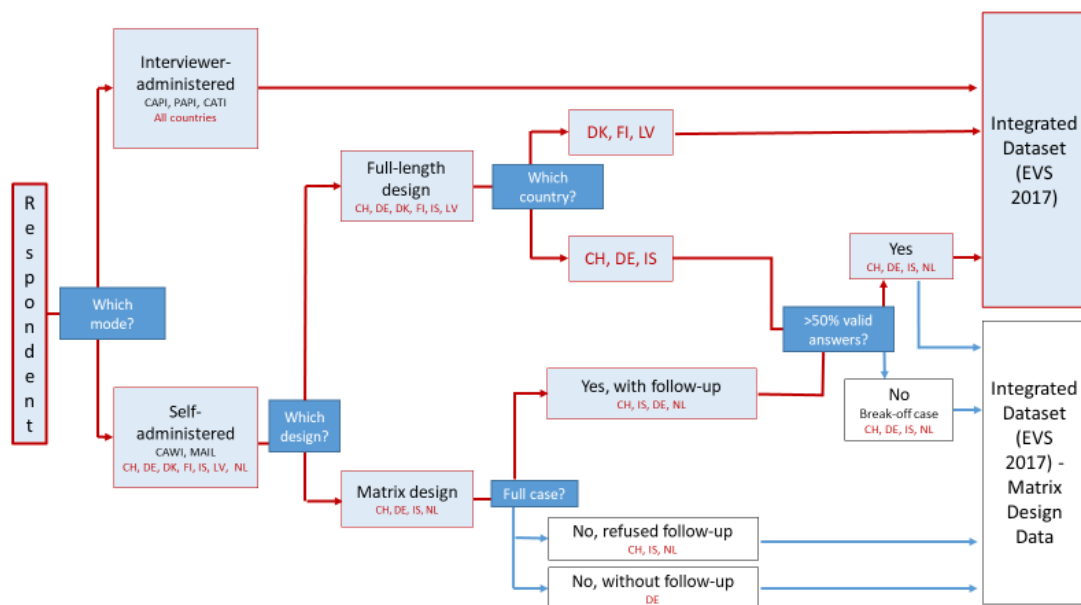
Figure 2:   Paths leasing cases into the EVS Integrated dataset (Main dataset)

2.2   EVS Integrated dataset – Matrix Design Data

The **EVS Integrated dataset – Matrix Design Data,** hereinafter also referred to as the **Matrix Design dataset**, is an anonymized dataset comprising 10,598 cases from the four countries implementing the matrix design: Switzerland, Germany, Iceland and the Netherlands. Both cases from the full-length design and from the matrix design are included. As for the latter, all respondents who filled in the matrix design are incorporated, ignoring the distinction between those who filled in only the first survey and those who also filled in the follow-up. Finally, the dataset also comprises break-off cases (cases with 'no answer' to more than 50% of the questionnaire items). See Figure 3 for a visual representation.

Due to its experimental nature, the use of this dataset is mostly recommended for methodological research.

---

2   The list of variables used to compute the percentage of no answers is provided in Appendix A.
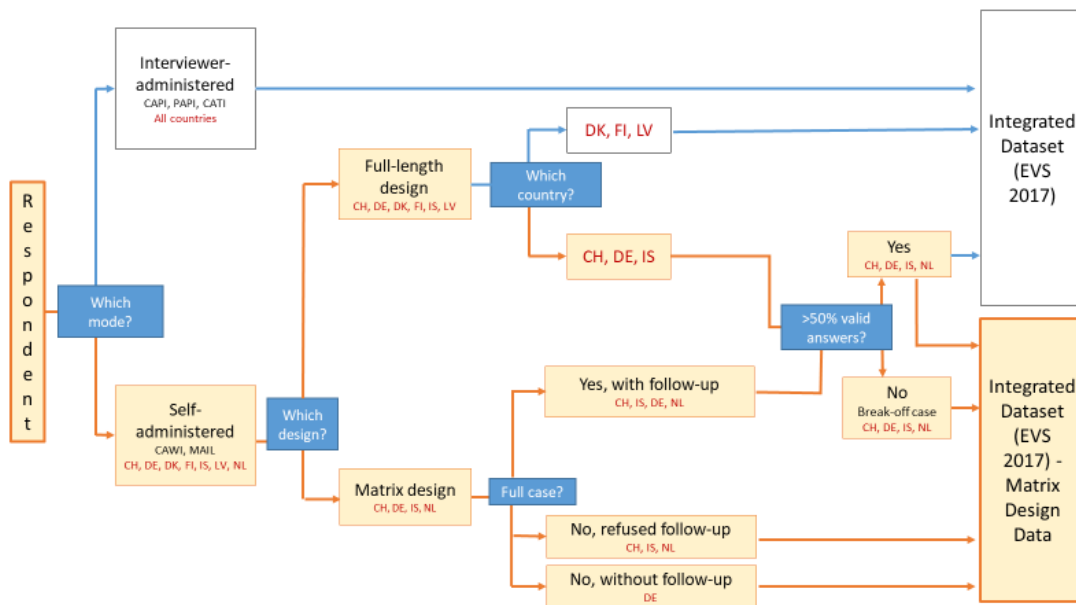
Figure 3:   Paths leading cases into the EVS Integrated dataset- Matrix Design Data (Matrix Design dataset)

2.3   Duplicate cases: Variable 'fduplicate'

5,604 cases appear both in the Main and the Matrix Design dataset. These are full cases, either from the full-length design or from the matrix design (only those who also completed the follow-up). These cases are flagged in both datasets by the variable 'fduplicate'. When merging datasets, it is important to make sure that only one record is retained for each of these cases. Guidance on how to merge the datasets on SPSS, Stata or RStudio can be found in Appendix B. The cases flagged as duplicates are also highlighted in the diagram in Figure 4.
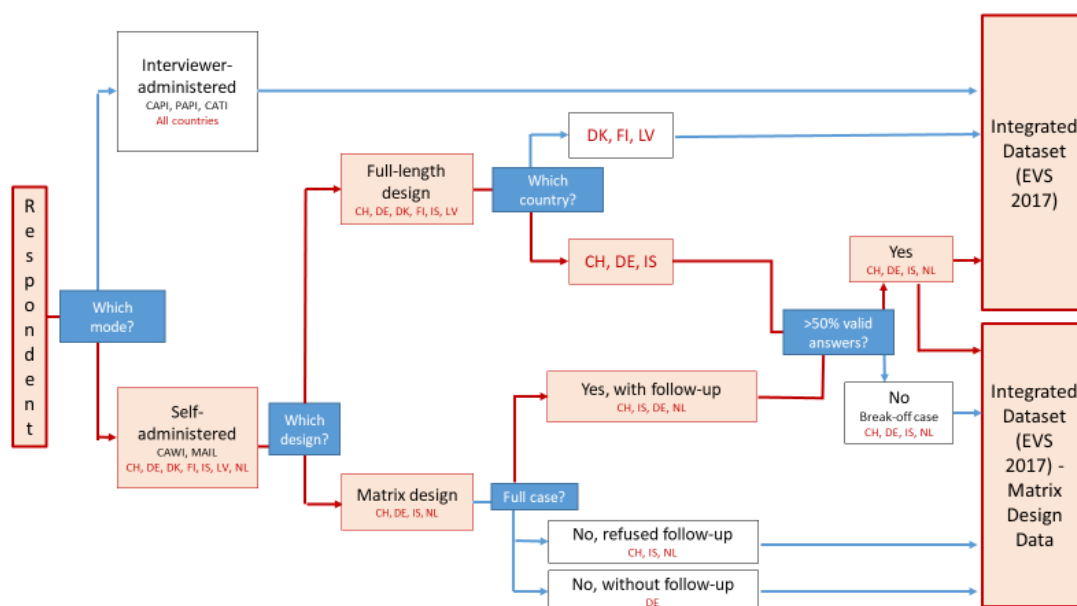
Figure 4:   Paths leading cases into both the EVS Integrated dataset (Main dataset) and the EVS Integrated
                dataset- Matrix Design Data (Matrix Design dataset)

2.4   Complete cases: Variable 'fmissings'

The large majority of the individual records included in the datasets are complete cases, i.e. they
provided valid answers to at least 80% of the items. There is, however, a small portion of incom-
plete or partial cases, i.e. cases that preferred not to answer to more than 50% but less than 80%
of the items. Cases that provided valid answers to less than 50% of the questions[3] are considered
break-off cases and hence excluded. An exception is made for the Matrix Design data, where it was
decided to retain break-off cases for methodological purposes. The break-off cases are only in-
cluded in the EVS Integrated dataset – Matrix Design Data. The variable 'fmissings' is provided to
enable distinguishing among these different situations (see Table 1).

Table 1:   Distribution of variable 'fmissings' by dataset.

fmissings	EVS Integrated dataset	EVS Integrated dataset - Matrix Design Data	*Of which duplicate cases*
	N	N	*N*
0 full/complete case	59227	10385	5,545
1 incomplete/partial case	211	91	59
2 Break-off (matrix design)	--	122	0
Total	59438	10,598	5,604

---

[3]   This means about 139 variables for the full questionnaire or about 85 variables if only the first survey was
      answered.

3   Mode of data collection

In EVS, the main mode of data collection is interviewer-administered. Face-to-face interviews are administered by a trained interviewer, with the support of a computer (CAPI mode), tablet (TAPI mode) or paper version (PAPI mode) on which answers to the questionnaire are recorded. For parsimony reasons, CAPI and TAPI interviews are aggregated under the label CAPI. Additionally, in special cases countries interviewers conducted interviews over the phone (CATI mode).

3.1   Mixed-mode in EVS 2017

In EVS 2017, for the purposes of increasing coverage and response rates, countries had the possibility of implementing a parallel self-administered questionnaire, with respondents assigned to either a Web survey (CAWI mode) or a postal paper-and-pencil survey (Mail mode). An overview of the grouping of modes is provided in Table 2, while an overview of the modes used in each country is provided in Table 3.

Table 2:   Overview of modes allowed in EVS 2017

Interviewer-administered	Self-administered
CAPI (TAPI)	CAWI
PAPI	MAIL
CATI	

Mixed-mode refers to the parallel implementation of different modes of data collection between interviewer-administrated and self-administrated conditions in the realization of the survey in one particular country. There are roughly two situations:
- Interviewer-administered mode as the main mode: applied by all countries
- Self-administered mode as a parallel mixed-mode: additionally applied by six countries.

Web surveys (CAWI) are the main alternative to face-to-face interviews. However, countries have also used postal surveys (Mail) for several purposes such as reaching 'offline' segments of the population and accommodate respondents who do not like to answer on a screen. As a general rule, if the response rate to the web part was expected to be lower than 40%, countries were encouraged to implement measures in order to increase the final response rate.

Table 3:   Overview of modes used by countries in EVS2017

Country		Mixed-mode	Interviewer-administered			Self-administered		Total
			CAPI	PAPI	CATI	CAWI	Mail	
			N	N	N	N	N	N
AL	Albania			1,435				1,435
AM	Armenia			1,500				1,500
AT	Austria		1,644					1,644
AZ	Azerbaijan			1,800				1,800
BY	Belarus			1,548				1,548
BA	Bosnia and Herzegovina		1,724					1,724
BG	Bulgaria			1,558				1,558
HR	Croatia		1,487					1,487

Country		Mixed-mode	Interviewer-administered			Self-administered		Total
			CAPI	PAPI	CATI	CAWI	Mail	
			N	N	N	N	N	N
CZ	Czechia		757	1,054				1,811
DK	Denmark	X	1,696			1,255	411	3,362
EE	Estonia		1,304					1,304
FI	Finland	X	388			668	143	1,199
FR	France		1,870					1,870
GE	Georgia		2,194					2,194
DE	Germany	X	1,494			1,021	2,941	5,456
GB	Great Britain		1,788					1,788
HU	Hungary		1,514					1,514
IS	Iceland	X	881	34		1,514	82	2,511
IT	Italy		2,277					2,277
LV	Latvia	X	1,070			265		1,335
LT	Lithuania			1,448				1,448
ME	Montenegro		1,003					1,003
NL	Netherlands	X	686			2,053		2,739
MK	North Macedonia			1,117				1,117
NO	Norway		977		145			1,122
PL	Poland		1,352					1,352
PT	Portugal		1,215					1,215
RO	Romania		1,613					1,613
RU	Russia		1,354	471				1,825
RS	Serbia			1,499				1,499
SK	Slovakia		1,432					1,432
SI	Slovenia		1,075					1,075
ES	Spain		1,209					1,209
SE	Sweden		1,194					1,194
CH	Switzerland	X	673			2,124	863	3,660
UA	Ukraine			1,612				1,612
Total			35,871	15,076	145	8,900	4,440	64,432

Note: these numbers can be obtained by merging the EVS Integrated dataset and the EVS Integrated dataset – Matrix Design Data.

3.1.1 Variable: mode

The mode of data collection can be identified through the protocol variable '**mode'**, which consists of 5 categories: 1 CAPI, 2 PAPI, 3 CAWI, 4 Mail, and 5 CATI. It allows detecting the mode of interview of each respondent (see Table 4).

Table 4: Variable 'mode'

mode	Interviewer-administered	Self-administered	Total
	N	N	N
1 CAPI	35,871		35,871
2 PAPI	15,076		15,076
3 CAWI		8,900	8,900
4 MAIL		4,440	4,440
5 CATI	145		145
Total	51,092	13,340	64,432

Note: these numbers can be obtained by merging the EVS Integrated dataset and the EVS Integrated dataset – Matrix Design Data.

4   Self-administered mode

This section describes the implementation of the self-administered (CAWI and Mail) interviews.

4.1   Sampling

The sampling had to be the same as for the main face to face survey (see [1] and [6]).

4.2   Adaptation of the EVS questionnaire to self-administered surveys

The EVS17 Master Questionnaire (see [4]) was prepared by the EVS Theory Group and provided in English, containing – beyond the questions and answer options – also interviewer and routing instructions. Each country was responsible for translating it and adapting it to the chosen interviewer-administered mode (CAPI, CATI and PAPI).
An additional version of the EVS17 Master Questionnaire was prepared by the EVS Methodology Group for implementation in the self-administered modes (CAWI, Mail) (see [5]).
Adaptations concerned for instance:

- Question texts: formulations like "please say" or "could you please tell me" only work when read out by interviewers, hence they were replaced by "please indicate" or "we would like to ask you" etc. (see Box 1).

Box 1: Example of question adaptation

Interviewer-administered:	Q1. Please say, for each of the following, how important it is in your life.
Self-administered:	Q1. Please indicate how important is in your life…

- Interviewer instructions: without an interviewer controlling the flow of the questionnaire, instructions in the self-administered mode were directly provided to interviewees (see box below).

Box 2: Example of adaptation of interviewer instruction

Interviewer-administered:	Q6. [READ OUT AND CODE ONE ANSWER PER LINE]
Self-administered:	Q6. Check all that apply.

- 'Spontaneous' answer options: all "don't know" and "no answer" options are spontaneous in the self-administered environment. For the CAWI, it was then decided to make the options visible only once the respondents tried to skip the question without selecting a valid answer. In the MAIL format, options were directly offered to respondents.

Box 3:   Example of adaptation of spontaneous answer options

The answer categories below the line are to be shown only once the respondent has tried to skip the question, together with this message[4]:
„We noticed that you did not answer this question. We would be very grateful if you answered to the best of your knowledge, even if you are not completely sure. But if you prefer not to answer to the question, you can continue the questionnaire.“

Yes	[1]
No	[2]
_____	
I don't know	[8]
I prefer not to answer	[9]

Each country was responsible of importing the questionnaire into the chosen Web survey software and/or for preparing the paper version to send via snail mail[5].

Rules and advices about such adaptation as well as link to examples of implementation in order to help the countries in implementing the Web survey version were given in the "EVS2017 Mixed-mode Guidelines" (see [1]). The document outlined some general rules[6] to be observed:

- Progress indicators are not allowed

- A "Back" button has to be provided, near the "Next" one.

- It must be possible to do the survey on a mobile device as a tablet or a smartphone and the design developed for a device in a "portrait" position. Para data have to be collected, so as to detect the device used.

- Each question must be placed on a page and scrolling must be avoided as far as possible in order not to have primacy effects.

- "Batteries of questions" have to be avoided and replaced by "simple questions". In this case, consider visually maintaining the constant part and modifying only the part for each item. This is of course needed for mobile devices but could also be recommended in order to avoid lining or satisficing in the PC version. This should also be respected for the mail version. A few exceptions, in which grids have to be presented, will be pointed out.

- Use a simple and clear design and avoid underlining that can be confused with hyperlinks.

- The questionnaire must not propose explicitly a "Non-response" option but must provide a reminder if a respondent does not answer a question.

An example on the presentation of different types of questions and several manners to answer it is given in the document "CAWI-Question-Design-CH.doc" [2]) developed by FORS and provided by the Swiss team.

---

[4] The wording has been taken from CRONOS (for more details, see http://www.europeansocialsurvey.org/methodology/methodological_research/modes_of_data_collection/cronos.html).

[5] Basic versions of the questionnaires could also be exported from the Translation Management Tool (TMT).

[6] We use here the elements proposed for the CRONOS web panel. Thanks to this initiative for providing these, also in the context of the SERISS project.

4.3   Practical path in case of mixed-mode

For the practical realization, the following sequence of steps was proposed, in order to stay close to a "Tailored Design Model" as proposed by Dillman[7]. This strategy is also called "push-to-Web".

- Send the link to the questionnaire, ideally with an unconditional cash incentive.

- Send a recall with once again the link.

- Send a paper questionnaire with a prepaid response envelope.

- Send a last recall.

Countries were, however, allowed to propose different strategies.

4.4   Self-administered mode applied by countries

The possibility of using the self-administered mode was implemented in different ways by countries, and in some cases enriched by two methodological experiments: the matrix design and the responsive design. These will be more thoroughly described in the next sections. An overview of the design adopted by each country can be found in Table 5. In general:

- Five out of the six "mixed-mode" countries administered the survey based on the full-length EVS questionnaire, i.e. they administered the whole EVS questionnaire in one go.
    - In Denmark and Finland, the full-length questionnaire was the only design implemented;
    - In Switzerland, Germany and Iceland, the full-length questionnaire condition was used in parallel to the matrix design;
    - Switzerland additionally used two versions of the full-length questionnaire, one resembling the order of questions of the EVS questionnaire, and one with a modified order of questions.
- Four countries (Switzerland, Germany, Iceland and Netherlands) employed a matrix design experiment, splitting the questionnaire in variable blocks and assigning them to matrix groups.
    - Whereas three of the four countries administered two rounds, matrix 1st round and a follow-up round, Germany administered a matrix 1st round only.
- Germany additionally implemented a responsive design experiment. The matrix design survey employed a responsive design in two phases performing experimental research on contact mode and incentive.

More details about these strategies are given in Sections 5 and 6.

---

[7]   Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). Internet, phone, mail, and mixed-mode surveys: the tailored design method. John Wiley & Sons.

Table 5:   Overview of mixed-mode and matrix design implementation by country

Country		Full-length questionnaire (original order)	Full -length questionnaire (modified order)	Matrix (1st round)	Matrix (Follow up)	Responsive design
CH	Switzerland	X	X	X	X	
DE	Germany	X		X		X
DK	Denmark	X				
FI	Finland	X				
IS	Iceland	X		X	X	
LV	Latvia	X				
NL	Netherlands			X	X	

4.4.1   Variable: mm_mixed_mode

The protocol variable '**mm_mixed_mode**' describes the mixed-mode design adopted. An overview is provided in Table 6.

Table 6:   Variables: 'mode' and 'mm_mixed_mode' implemented in the data

Country	mode	mm_mixed_mode				
		1 mixed-mode: full questionnaire	2 mixed-mode: full (mod. order) questionnaire	3 mixed-mode: matrix and follow-up	4 mixed-mode: matrix only	5 mixed-mode: not applied
		N	N	N	N	N
All countries	1 CAPI					28,983
	2 PAPI					15,042
	5 CATI					145
CH	1 CAPI					673
	3 CAWI	300	323	1,501		
	4 Mail	111	124	628		
DE	1 CAPI					1,494
	3 CAWI	117			904	
	4 Mail	571			2,370	
DK	1 CAPI					1,696
	3 CAWI	1,255				
	4 Mail	411				
FI	1 CAPI					388
	3 CAWI	668				
	4 Mail	143				
IS	1 CAPI					881
	2 PAPI					34
	3 CAWI	209		1,305		
	4 Mail	30		52		
LV	1 CAPI					1,070
	3 CAWI	265				
NL	1 CAPI					686
	3 CAWI			2,053		
Total		4,080	447	5,539	3,274	51,092
			4,527		*8,813*	

Note: these numbers can be obtained by merging the EVS Integrated dataset and the EVS Integrated dataset – Matrix Design Data

In general,

- For self-administered mode surveys, 'mm_mixed_mode' identifies the design used: full(-length) questionnaire (code 1), full(-length) questionnaire with modified order (code 2), matrix design with follow-up (code 3) and matrix design without follow-up (code 4);
- For interviewer-administered mode surveys, 'mm_mixed_mode' is not applied (code 5).

Each dataset contains a subsample of these categories, as illustrated in Table 7.

Table 7:   Distribution of 'mm_mixed_mode' by dataset

mm_mixed_mode	EVS Integrated dataset	EVS Integrated dataset - Matrix Design Data	*Of which duplicate cases*
	N	N	*N*
1 mixed-mode: full questionnaire	4058	1,338	1,316
2 mixed-mode: full (mod. order) questionnaire	437	447	437
3 mixed-mode: matrix and follow-up	3,851	5,539	3,851
4 mixed-mode: matrix only		3,274	
5 mixed-mode: not applied	51092		

5 Matrix design

The EVS 2017 Master Questionnaire is a rather long survey (~280 ticks), and the EVS Methodology Group deemed it not fully suitable for the Web implementation, as it would have been difficult to administer in one step. For these reasons, some suggestions were provided:

- It was advised to make sure that the survey could be interrupted at any moment as needed by the respondents;
- If applied in its full length, advance letters had to mention this aspect;
- Specific recalls had to be prepared for people interrupting the answer process and not coming back after a decent delay. A file containing the history of contacts has to be provided;
- Finally, a matrix design could be applied.

The matrix design consists of splitting the questionnaire into several coherent blocks of questions. Each respondent, instead of getting the whole questionnaire, would then get only some of these blocks. This procedure requires a bigger sample size. A more practical example is provided in Box 4.

Box 4: Example of matrix design (see [1]).

Let us consider the following design, dividing the survey into 4 blocks, aiming at overall around 2000 respondents, divided into 6 situations, S1 to S6. In such a design, each block will be answered by about 1000 respondents and at least 333 answers could be mobilized for analysis between blocks.

Figure 5: Matrix design - example

Groups	Background	A	B	C	D
S1	333	333	333		
S2	333	333		333	
S3	333	333			333
S4	333		333	333	
S5	333		333		333
S6	333			333	333

- 2000 respondents overall,
1000 respondents for each substantive question;
at least 333 respondents for each binary combination of questions.

For EVS 2017, the division into blocks was prepared by the Swiss team. The information is summarized in a document (see [3]) containing detailed information on the changed order of the questions assigned to matrix groups and respondents and has been used as basis by all countries applying matrix design.

The grouping of questions was done taking into account the topic and the correlations between items/questions. The concept was adopted by four countries (Switzerland, Germany, Iceland and the Netherlands) and implemented in the following steps:

1. The variables of the EVS Master Questionnaire were assigned to the five matrix blocks (seeTable 8).

Table 8:   Matrix design: variable blocks

Question block	Main themes
Core	Demographics, life perception & core questions from blocks A-D
A	Family and work
B	Religion, morals, national identity
C	Society
D	Politics

2. In the next step, the question blocks in Table 8 were assigned to the 6 matrix groups (see Table 9). Two approaches were used: a country administered a first round survey only or implemented a "mini-panel", returning to the respondents of the matrix survey and asking them to fill in a second (follow-up) round, asking questions not administered during the first wave (see Table 10 and Table 11).

Table 9:   Matrix design: question blocks assigned to first matrix survey

Matrix group /Question block	Core	Block A	Block B	Block C	Block D
Matrix Group 1	X	X	X		
Matrix Group 2	X	X		X	
Matrix Group 3	X	X			X
Matrix Group 4	X		X	X	
Matrix Group 5	X		X		X
Matrix Group 6	X			X	X

Table 10:   Matrix design: first and second round survey

First round	First questionnaire
	- two variable blocks - background/core variables
Second round	Follow-up
	two variable blocks

Table 11:   Matrix design: question blocks assigned to second matrix survey (mini-panel)

Matrix group /Question block	Core	Block A	Block B	Block C	Block D
Matrix Group 1	X	X	X	**O**	**O**
Matrix Group 2	X	X	**O**	X	**O**
Matrix Group 3	X	X	**O**	**O**	X
Matrix Group 4	X	**O**	X	X	**O**
Matrix Group 5	X	**O**	X	**O**	X
Matrix Group 6	X	**O**	**O**	X	X

X: blocks assigned during first round | **O**: blocks assigned during the follow up

3. Finally, each respondent was assigned to one of the matrix groups 1 to 6.

5.1   Matrix design applied by countries

If both the first and second round (follow-up) surveys were administered, respondents answered the full questionnaire (see Table 11). This concept was adopted by Switzerland, Iceland and the Netherlands. If only the first-round survey was administered, respondents answered -besides the background/core variables- only two of the variable blocks A to D. This concept was adopted by Germany.

In addition to the matrix design, three countries (Switzerland, Germany[8], Iceland) also provided a full-length questionnaire version. Moreover, one country (Switzerland) also implemented a version of the full-length questionnaire with a modified order of questions[9]. Three countries (Denmark, Finland, and Latvia) provided only the full-questionnaire version. An overview was presented above, in Table 5.

5.1.1   Variable: mm_matrix_group

The variable '**mm_matrix_group**' represents the matrix group/variable blocks attributed to each respondent. The variable differentiates between countries with the follow up, whose groups range from 1 to 6, and the country (Germany) with only the first round of matrix design, whose groups range from 11 to 61. It also contains categories for the full-length questionnaire version (code 7) and the full-length questionnaire with modified order (code 8). Cases from the interviewer-administered mode are coded as -3 ('non applicable'). An overview of the coding of this variable is provided in Table 12.

Table 12:   Variables: 'mm_matrix_group' and 'mode', and sample size in each dataset

mm_matrix_group	Mode[a]		EVS Integrated dataset	EVS Integrated dataset – Matrix Design Data
	3 CAWI	4 Mail	N	N
1 [A B Core] [C D]	CH IS NL	CH IS	3,851	5,539
2 [A C Core] [B D]	CH IS NL	CH IS		
3 [A D Core] [B C]	CH IS NL	CH IS		
4 [B C Core] [A D]	CH IS NL	CH IS		
5 [B D Core] [A C]	CH IS NL	CH IS		
6 [C D Core] [A B]	CH IS NL	CH IS		
7 mixed-mode: full questionnaire	CH DE DK FI IS	CH DE DK FI IS	4,058	1,338
8 mixed-mode: full (modified order) questionnaire	CH	CH	437	447
11 [A B Core]	DE	DE		3,274
21 [A C Core]	DE	DE		
31 [A D Core]	DE	DE		
41 [B C Core]	DE	DE		
51 [B D Core]	DE	DE		
61 [C D Core]	DE	DE		

[a]cases from the CAPI, PAPI and CATI mode are coded as -3 "Not applicable" in the variable mm_matrix_group

---

[8]   The design of the German survey is better explained in Section 6.

[9]   In this case, the order of questions follows the order as in Group 5 of the matrix design: B D A C Core.

5.1.2   Variables: mm_original_matrix_design_IS, mm_original_matrix_group_IS

Two country-specific variables in the Icelandic sample, **'mm_original_matrix_design_IS'** and **'mm_original_matrix_group_IS'**, enable the identification of the original group respondents were randomly assigned to after the sampling phase. Despite being originally assigned to a matrix group, 22 respondents filled in the full-length postal questionnaire enclosed with the 3rd reminder, instead of filling the CAWI survey related to the matrix group they were assigned to. As a consequence, these respondents are coded as 7 "mixed-mode: full questionnaire" in the variable 'mm_matrix_group'.

5.1.3   Variable: mm_mode_fu

The variable '**mm_mode_fu**' represents the mode of the data collection in the follow-up (or second round of the matrix design). It can be either CAWI (code 1) or Mail (code 2); otherwise, a set of missing values is provided (and described in section 7). Table 13 offers an overview of the coding of the mode and matrix group variables.

Table 13:   Variables: 'mode', 'mm_mixed_mode', 'mm_matrix_group' and 'mm_mode_fu'

	mode	mm_mixed_mode	mm_matrix group	mm_mode_fu
Interviewer-administered mode	1 CAPI 2 PAPI 5 CATI	5 mixed-mode not applied	- 3 not applicable	-3 not applicable
Self-administered mode	3 CAWI 4 Mail	1 mixed-mode: full questionnaire	7 mixed-mode: full questionnaire	-3 not applicable
		2 mixed-mode: full (modified order) questionnaire	8 mixed-mode: full (mod. order) questionnaire	-3 not applicable
		3 mixed-mode: matrix and follow-up	1 [A B Core] [C D] 2 [A C Core] [B D] 3 [A D Core] [B C] 4 [B C Core] [A D] 5 [B D Core] [A C] 6 [C D Core] [A B]	1 CAWI (follow-up) 2 MAIL (follow-up) -8 follow-up non-response
		4 mixed-mode: matrix only	11 [A B Core] 21 [A C Core] 31 [A D Core] 41 [B C Core] 51 [B D Core] 61 [C D Core]	-9 no follow-up

5.1.4   Variables: mm_fw_start_fu,mm_fw_end_fu, mm_year_fu, mm_v27**_fu

The administrative variables '**mm_fw_start_fu**', '**mm_fw_end_fu**', '**mm_year_fu**' indicate the time of fieldwork of the follow-up (when applicable). More specifically, they respectively represent: the year-month at starting the fieldwork of the follow-up survey, the year-month at finishing the fieldwork of the follow-up survey and the year of data collection of the follow-up survey.
The date and hour of interview of the follow-up survey are recorded in variables: '**mm_v277_fu**', '**mm_v278a_fu**', '**mm_v278b_fu**', '**mm_v279a_fu**', '**mm_v279b_fu**'.

6   Responsive design

Germany performed two completely independent self-administered mode surveys:

- one survey (CAWI and Mail), comprising only the 1st round of the matrix design, that embedded a responsive survey design (2 phases) and experiments on contact mode and incentives.

- one survey (CAWI and Mail) with the full-length questionnaires, that applied only one condition: simultaneous contact mode with 5 Euros prepaid incentive for all respondents.

In both surveys, respondents in the simultaneous/concurrent contact mode condition were offered Mail and CAWI questionnaires right from the start, respondents in the sequential contact mode condition ("push-to-Web") were offered CAWI questionnaires first, and Mail questionnaires later with the 2nd reminder.

6.1   The two-phases responsive design

1. Self-administered survey (CAWI and Mail) with matrix questionnaires (implementing a two-phased responsive survey design with experiments on contact mode and incentives)

The first survey phase tested different contact modes (simultaneous/concurrent vs. sequential/"web-push") and incentive strategies (5 EUR prepaid vs. 10 EUR postpaid), resulting in four experimental groups (sequential prepaid, sequential postpaid, simultaneous prepaid, simultaneous postpaid). One invitation letter was followed by three reminder letters at maximum.

- All respondents in the simultaneous/concurrent contact mode condition were offered Mail and CAWI questionnaires right from the start with the advance letter.
- Whereas respondents in the sequential contact mode condition ("web-push") were offered CAWI questionnaires first, and Mail questionnaires later with the 2nd reminder.
- During the <u>first survey phase</u> the incentive was dependent on the experimental group: the respondents either received a 5 EUR prepaid incentive (cash) with the advance letter or 10 EUR postpaid incentive (cash) mailed with a thank-you note after survey completion.
- During the <u>second survey phase</u> all respondents received a 5 EUR prepaid incentive (cash) with the advance letter.

2. Mixed-Mode Survey (CAWI and Mail) with the full-length questionnaire (only one condition applied: simultaneous contact mode with 5 Euros prepaid incentive).

The second survey phase used the best evaluated design from the first survey phase to complete the survey: simultaneous/concurrent contact mode with 5 EUR prepaid incentive. One invitation letter was followed by two reminder letters at maximum.

- All respondents were in the simultaneous/concurrent contact mode condition, where Mail and CAWI questionnaires were both offered right from the start with the advance letter.
- All respondents received a 5 EUR prepaid incentive (cash) with the advance letter.

6.2   Responsive design concept: variables

This section describes the variables starting with `mr_', which identify the features of the responsive design implemented in the self-administered mode in Germany (cf. Section 5).

6.2.1   Variables: mr_detailed_mode_DE, mr_contact_mode_DE, mr_incentive_DE

Table 14 and Table 15 provide an overview of the protocol variables and coding for the responsive design.

Table 14:   Overview of the variables related to the responsive design

	Mode	mm_mixed_mode	mr_detailed_mode_de*	mr_contact_mode_de*	mr_incentive_de*
Interviewer-administered mode	1 CAPI	5 mixed-mode not applied	1 CAPI (base sample) 2 CAPI (base follow-up) 3 CAPI (sample increase)	-4 item not included	3 '20 EUR postpaid'
Self-administered mode	3 CAWI 4 MAIL	1 mixed-mode: full questionnaire	8 CAWI full length 9 Mail full length	2 Simultaneous	2 '5 EUR prepaid'
		3 mixed-mode: matrix only	4 CAWI matrix (phase 1) 6 Mail matrix (phase 1) 5 CAWI matrix (phase 2) 7 Mail matrix (phase 2)	1 Sequential 2 Simultaneous 2 Simultaneous	1 '10 EUR postpaid' 2 '5 EUR prepaid 2 '5 EUR prepaid'

*For all other countries/mode/designs, these variables will take value -4 'Item not included'

Table 15:   Variables: mode, mr_detailed_mode_DE, mr_contact_mode_DE, mr_incentive_DE

mode	mr_detailed_mode_DE	mr_contact_mode_DE		mr_incentive_DE		
		1 sequential N	2 simultaneous N	1 10 EUR postpaid N	2 5 EUR prepaid N	*Total* *N*
3 CAWI	4 CAWI matrix (phase 1)	438	255	289	404	693
	5 CAWI matrix (phase 2)	0	211	0	211	211
	8 CAWI full length	0	117	0	117	117
4 Mail	6 Mail matrix (phase 1)	201	1,195	563	833	1,396
	7 Mail matrix (phase 2)	0	974	0	974	974
	9 Mail full length	0	571	0	571	571
Total		639	3,323	853	3,110	3,962

Note: these numbers can be obtained by merging the EVS Integrated dataset and the EVS Integrated dataset – Matrix Design Data

7   Missing values overview

Due to the specificity of the mixed-mode approach, several missing values have been designed. The overview can be found in Table 16.

Table 16:   Overview of mixed-mode specific missing values

Mode	Missing value	Variable label	Explanation
CAWI/MAIL	-8	follow-up non-response	This applies to respondents that refused to take the follow-up (or matrix 2nd round)
	-9	no follow-up	This applies to situations when there is only the 1st round of matrix but not the follow up (hence when questions have not been asked because of the group assignment)
	-10	multiple answers (Mail)	This applies to those respondents that wrongly selected multiple answers in the Mail questionnaire

8 Summary

This section introduces a variable that summarizes most of the mixed-mode characteristics and provides an overview of all the variables described in the document.

8.1 Variable: mm_select_sample

The protocol variable `**mm_select_sample**' allows distinguishing seven different subsamples, combining information from the `mm_*' variables and the behaviour of respondents. In particular, categories 4 and 5 distinguish respondents that filled in the follow-up from those who did not (among those who were invited to do so), while category 7 identifies break off cases, that are only included in the Matrix Design Data. Break off cases are respondents that selected 'no answer' on more than 50% of the questions they were asked. Table 17 describes each category and how many respondents belonging to each group can be found in each dataset.

Table 17:   Variable 'mm_select_sample; by country

			EVS Integrated dataset				**EVS Integrated dataset – Matrix Design Data**						
country	*Total country sample[a]*	1 interviewer-administered (CAPI PAPI CATI)	2 self-administered full-length questionnaire: original question order (CAWI Mail)	3 self-administered full-length questionnaire: modified question order (CAWI Mail)	4 self-administered matrix: with follow-up (CAWI Mail)	*Total*	2 self-administered full-length questionnaire: original question order (CAWI Mail)	3 self-administered full-length questionnaire: modified question order (CAWI Mail)	4 self-administered matrix: with follow-up (CAWI Mail)	5 self-administered matrix: follow-up non response (CAWI Mail)	6 self-administered matrix: first survey only (CAWI Mail - DE)	7 break-off (less than 50% valid answers)	*Total*
	N	N	N	N	N	N	N	N	N	N	N	N	N
CH	*3,658*	673	403	437	1661	*3,174*	403	437	1661	432		54	*2,987*
DE	*5,456*	1494	676			*2,170*	676				3237	49	*3,962*
DK	*3,362*	1696	1666			*3,362*							
FI	*1,199*	388	811			*1,199*							
IS	*2,511*	915	237		472	*1,624*	237		472	879		8	*1,596*
LV	*1,335*	1,070	265			*1335*							
NL	*2,739*	686			1718	*2,404*			1,718	324		11	*2,053*

8.2 Summary of mixed-mode variables

The coding of the different protocol variables described in the sections above is summarized in Table 18.

Table 18: Summary of mixed-mode protocol variables

Variable	Description	Categories	Country
mode	mode used for the (1st round) survey	1 CAPI	All EVS 2017 countries
		2 PAPI	
		3 CAWI	
		4 Mail	
		5 CATI	
mm_select_sample	Identifies subsamples on mode variables	1 interviewer-administered (CAPI PAPI CATI)	All countries
		2 self-administered full-length question-naires: original question order (CAWI MAIL)	CH IS DE LV
		3 self-administered full-length question-naires: modified question order (CAWI MAIL)	CH
		4 self-administered matrix: matrix and follow-up (CAWI Mail)	CH IS NL
		5 self-administered matrix: follow-up non response (CAWI Mail)	CH IS NL
		6 self-administered matrix: first survey only (CAWI Mail)	DE
		7 Break-off (less than 50% valid answers)	CH IS NL DE
mm_mixed_mode	mode and questionnaire assigned to respondents	1 mixed-mode: full questionnaire	CH DE DK FI IS LV
		2 mixed-mode: full (mod. order) question-naire	CH
		3 mixed-mode: matrix and follow-up	CH IS NL
		4 mixed-mode: matrix only	DE
		5 mixed-mode: not applied	All interviewer-administered
mm_mode_fu	mode used for the follow-up survey	1 CAWI (follow-up)	CH IS NL
		2 Mail (follow-up)	CH IS
		-3 not applicable	All interviewer-administered and self-administered: CH DE DK FI IS LV
		-8 follow-up non-response	CH IS NL
		-9 no follow-up	DE
mm_matrix group	membership of questions in groups according to matrix design	1-6 Group 1-6	CH IS NL
		7 mixed-mode: full questionnaire	CH DE DK FI IS LV
		8 mixed-mode: full (modified order) ques-tionnaire	CH IS
		11-61 group 1-6 (matrix only)	DE
		-3 not applicable	All interviewer-administered
mr_detailed_mode_de	interview is part of phase 1 or phase 2	4 CAWI matrix (phase 1)	DE
		5 CAWI matrix (phase 2)	
		6 Mail matrix (phase 1)	
		7 Mail matrix (phase 2)	

Variable	Description	Categories	Country
mr_contact_mode_de	Sequential (CAWI first)	1 sequential	DE
	Simultaneous (CAWI/Mail from start)	2 simultaneous	
mr_incentive_de	10 EUR post-paid incentive (cash)	1 10 EUR postpaid	DE
	5 EUR prepaid incentive (cash)	2 5 EUR prepaid	

9　Practical Examples and Scenarios for the selection of cases

In Box 5, a concrete example on how to navigate the mixed-mode variables is provided.

Box 5: example

In the following we compare the two matrix designs (matrix only/matrix and follow-up) using a concrete question/variable as an example. The result is given in Table 19 and Figure 6.

- The example is variable v9 (question Q4) 'Religious or church organisations'. This variable is assigned to question block C (see column Block) in the overview Matrix-Design: Question-Blocks [3] also reported in Figure 6.
- In this overview one can also see that v9 is in the first round survey assigned to matrix groups 2 [A C], 4 [B C], and 6 [C D]; and in the second round survey to matrix groups 1 [C D], 3 [B C], and 5 [A C].
- If the 'matrix and follow-up' design is applied, all respondents assigned to the matrix groups 1 to 6 were asked to answer the question.
- If 'matrix only' is the chosen design, only the respondents assigned to the matrix groups 2, 4, and 6 were asked this question; whereas respondents assigned to the matrix groups 1,3, or 5 were coded as missing value '-9 no follow-up'.
- If the 'matrix and follow-up' design is applied, respondents who were invited to the second round but refused to answer, were coded as missing values '-8 follow-up non-response'.

Table 19:　　Matrix design: example for variable v9

Matrix group /Question block	v9 of block C 'matrix only'	v9 of block C 'matrix and follow-up'
Respondents of Matrix Group 1	-9 no follow-up	**answers** (or -8 follow-up non-response)
Respondents of Matrix Group 2	**answers**	**answers** (or -8 follow-up non-response)
Respondents of Matrix Group 3	-9 no follow-up	**answers** (or -8 follow-up non-response)
Respondents of Matrix Group 4	**answers**	**answers** (or -8 follow-up non-response)
Respondents of Matrix Group 5	-9 no follow-up	**answers** (or -8 follow-up non-response)
Respondents of Matrix Group 6	**answers**	**answers** (or -8 follow-up non-response)

Question	Variable	Text	Block	Group 1 [A B Core] [C D]	2 [A C Core] [B D]	3 [A D Core] [B C]	4 [B C Core] [A D]	5 [B D Core] [A C]	6 [C D Core] [A B]
Q1		Please say, for each of the following, how important it is in your life.	Core	X	X	X	X	X	X
	-		Core						
	v1	Work	Core	X	X	X	X	X	X
	v2	Family	Core	X	X	X	X	X	X
	v3	Friends and acquaintances	Core	X	X	X	X	X	X
	v4	Leisure time	Core	X	X	X	X	X	X
	v5	Politics	Core	X	X	X	X	X	X
	v6	Religion	Core	X	X	X	X	X	X
Q2	v7	Taking all things together, would you say you are: very happy - not at all happy	B	X	XX	XX	X	X	XX
Q3	v8	All in all, how would you describe your state of health these days? Would you say it is: very good - very poor	Core	X	X	X	X	X	X
Q4	-	Please look carefully at the following list of voluntary organisations and say which, if any, do you belong to?	C	XX	X	XX	X	XX	X
	v9	Religious or church organisations	C	XX	X	XX	X	XX	X
	v10	Education, arts, music or cultural activities	C	XX	X	XX	X	XX	X
	v11	Trade unions	C	XX	X	XX	X	XX	X
	v12	Political parties or groups	C	XX	X	XX	X	XX	X
	v13	Conservation, the environment, ecology, animal rights	C	XX	X	XX	X	XX	X
	v14	Professional associations	C	XX	X	XX	X	XX	X
	v15	Sports or recreation	C	XX	X	XX	X	XX	X
	v16	Humanitarian or charitable organization	C	XX	X	XX	X	XX	X
	v17	Consumer organization	C	XX	X	XX	X	XX	X
	v18	Self-help group, mutual aid group	C	XX	X	XX	X	XX	X
	v19	Other groups	C	XX	X	XX	X	XX	X
	v20	None (spontaneous)	C	XX	X	XX	X	XX	X

Figure 6:   *Example from Matrix-Design: Question Blocks (see [3]);* Note: X: the variable was included in the 1st round of the matrix survey; XX: the variable was included in the 2nd round (or follow-up)

9.1   Scenarios

In this section, we provide practical indication on how to derive particular sets of data. The list is not exhaustive, but it is meant to provide guidance.

9.1.1   Scenario 1: Interviewer-administered mode

When? If one wants to analyse data based on the main interviewer-administered mode only and exclude all self-administered interviews.

Why? This is the most conservative approach to achieve comparability across waves/countries.

Which dataset? All cases can be found in the EVS Integrated dataset (Main dataset).

How? Use one of the following options for the selection of cases:

- ▪ mm_select_sample = 1 'interviewer-administered (CAPI PAPI CATI)
- ▪ mode = 1 'CAPI' and 2 'PAPI' and 5 'CATI'.

The resulting dataset includes altogether 48,410 cases nested in 34 countries.

9.1.2   Scenario 2: Full-length questionnaire across modes

When? If one wants to retain only respondents that answered the full questionnaire, regardless if interviewer- or self-administered mode was applied and regardless the modified applied in the Swiss data.

Why? Even though mode differences could be present, the pattern of missing values by design due to the matrix design is avoided; also potential bias due to the different timing of data collection for the matrix design + follow up is avoided.

Which dataset? All cases can be found in the EVS Integrated dataset (Main dataset).

How? Use one of the following options for the <u>selection</u> of cases:

- mm_select_sample = 1 'interviewer-administered: CAPI PAPI CATI' and 2 'self-administered full-length questionnaires (original question order): CAWI MAIL' and 3' self-administered full-length questionnaires (modified question order): CAWI MAIL'.
- mm_mixed_mode: 5 'mixed mode not applied' and 1 'mixed mode: full questionnaire' and 2 'mixed mode: full (mod. order) questionnaire'

The resulting dataset includes 52,640 cases from 34 countries.

9.1.3   Scenario 3: Full length questionnaire across modes – retain original question order

When? As Scenario 2, but excluding the Swiss subsample which received the modified-order full-length questionnaire.

Why? As Scenario 2, but one additionally avoids potential effects due to the different order of the questions.

Which dataset? All cases can be found in the EVS Integrated dataset (Main dataset).

How? Use one of the following options for the <u>selection</u> of cases:

- 'mm_select_sample' = 1 'interviewer-administered: CAPI PAPI CATI' and 2 'self-administered full-length questionnaires (original question order): CAWI MAIL'.

The resulting dataset includes altogether 52,203 cases from 34 countries.

9.1.4   Scenario 4: Self-administered mode

When? If one wants to only analyse data from respondents that took the EVS survey in the self-administered mode (CAWI or Mail)

Why? For instance, one may be interested in exploring the measurement of values in a self-administered context.

Which dataset? Cases have to be retrieved from both the EVS Integrated dataset and the EVS Integrated dataset – Matrix Design Data.

How? We recommend the following steps:

- Merge EVS Integrated dataset and the EVS Integrated dataset – Matrix Design Data (see also Appendix B with syntax examples);
- Exclude/drop 'mm_select_sample' = 1 interviewer-administered (CAPI PAPI CATI);
- According to the type of analysis to be performed, decide whether to retain break-off cases ('no answer' to more than 50% of the administered questions) [coded as mm_select_sample==7].

The resulting dataset includes 13,075 respondents from 6 countries, comprising also 122 break-off cases.

9.1.5   Scenario 5: Matrix design excluding the follow up

When? If one wants to analyse data from the first survey of the matrix design, excluding all answers respondents gave in the follow up survey.

Why? Reduce potential effects due to the timing of the survey and to the selection bias caused by the follow-up; include all (additionally DE) first round surveys into your analysis; apply multiple imputation techniques.

Which dataset? All cases can be found in the EVS Integrated dataset – Matrix Design Data.

How? We recommend the following steps:

- Use the "Follow-up removal" syntax which can be found in the Data Catalogue to define the variable blocks and set all variables that were asked in the follow-up survey to missing value (-9 'No follow-up').
- Combine variable groups for all first round surveys: for CH, DE, NL, IS by using the two digit codes; for CH, NL, IS by recoding 'mm_matrix group' from one digit to two digit codes (1 -> 11, 2 -> 22, etc.). Such recoding is also part of the "Follow-up removal" syntax.

This action affects the amount of missing values in variables (except for socio-demographic variables and a few variables belonging to the Core block) of the cases belonging to group 3 in the variable mm_select_sample. In total, this scenario involves 8,813 respondents from 4 countries.

References

The following documents can be retrieved from the EVS website. All the documentation is related to the EVS 2017 Integrated datasets, ZA7500, Version: 5.0.0, doi:10.4232/1.13897.

[1] European Values Study (EVS) (2020). European Values Study (EVS) 2017: Methodological Guidelines (GESIS Paper, 2020/13). Köln: GESIS - Leibniz-Institut für Sozialwissenschaften.

[2] CAWI-Question-Design-CH.doc in ZA7500_q_ch.zip

[3] EVS 2017 Matrix-Design: Question-Blocks

[4] EVS 2017 CAPI Master Questionnaire

[5] EVS 2017 CAWI Master Questionnaire

[6] European Values Study (EVS) (2022). European Values Study (EVS) 2017: Method Report (GESIS Paper, 2022/07). Köln: GESIS - Leibniz-Institut für Sozialwissenschaften.

Appendix A – List of 277 questionnaire items

The following variable list is defined as 100% of variables (n=277):

v1; v2; v3; v4; v5; v6; v7; v8; v9; v10; v11; v12; v13; v14; v15; v16; v17; v18; v19; v20; v21; v22; v23; v24; v25; v26; v27; v28; v29; v30; v31; v32; v33; v34; v35; v36; v37; v38; v39; v40; v41; v42; v43; v44; v45; v45a; v46; v47; v48; v49; v50; v51; v52; v53; v54; v55; v56; v57; v58; v59; v60; v61; v62; v63; v64; v65; v66; v67; v68; v69; v70; v71; v72; v73; v74; v75; v76; v77; v78; v79; v80; v81; v82; v83; v84; v85; v86; v87; v88; v89; v90; v91; v92; v93; v94; v95; v96; v97; v98; v99; v100; v101; v102; v103; v104; v105; v106; v107; v108; v109; v110; v111; v112; v113; v114; v115; v116; v117; v118; v119; v120; v121; v122; v123; v124; v125; v126; v127; v128; v129; v130; v131; v132; v133; v134; v135; v136; v137; v138; v139; v140; v141; v142; v143; v144; v145; v146; v147; v148; v149; v150; v151; v152; v153; v154; v155; v156; v157; v158; v159; v160; v161; v162; v163; v164; v165; v166; v167; v168; v169; v170; v171; v172; v173; v174_cs; v175_cs; v176; v177; v178; v179; v180; v181; v182; v183; v184; v185; v186; v187; v188; v189; v190; v191; v192; v193; v194; v195; v196; v197; v198; v199; v200; v201; v202; v203; v204; v205; v206; v207; v208; v209; v210; v211; v212; v213; v214; v215; v216; v217; v218; v219; v220; v221; v222; v223; v224; v225; v226; v227; v228b_r; v229; v230; v231b_r; v232; v233b_r; v234; v235; v236; v237; v238; v239a; v239b; v240; v241; v242; v243_cs; v244; v245; v246_ISCO_2; v247; v248; v248a; v249; v250; v251b_r; v252_cs; v253; v254; v255_ISCO_2; v256; v257; v258; v259; v260; v261; v262_cs; v263_cs; v264; v265; v266; v267; v268; v269; v270; v271; v272; v273; v274

- Break-off case: less than 50% of variables are answered by respondent (at least 50% of variables are missing values), i.e. at least
    - 139 variables are coded as "no answer" for respondents that answered the full questionnaire (CAPI and CAWI/Mail who answered the 2nd round)
    - For respondents that did not answer the 2nd round in CAWI/Mail mode (no follow up [DE] + FU refusal cases):
        - 85 variables are coded as "na" (for mm_matrix_group eq 1 or 3)
        - 84 variables are coded as "na" (for mm_matrix_group eq 2)
        - 81 variables are coded as "na" (for mm_matrix_group eq 4 or 6)
        - 82 variables are coded as "na" (for mm_matrix_group eq 5)

Appendix B – Merging syntaxes

Examples of syntaxes to merge EVS Integrated dataset (main) and EVS Integrated dataset – Matrix Design Data (addon).

Recommended procedure in SPSS

```
* Load files.
FILE HANDLE main/Name="[Insert the file path to the EVS datasets\
insert name of EVS Integrated dataset].sav".
FILE HANDLE addon /Name="[Insert the file path to the EVS datasets\
insert name of EVS Integrated dataset-Matrix Design Data].sav".
DATASET CLOSE all.
GET FILE 'main' .
DATASET NAME main.
GET FILE 'addon' .
DATASET NAME addon.


DATASET ACTIVATE main.
FREQUENCIES fduplicate. /*check no. of duplicate cases.

* merge datasets based on id variable (id_cocas). Please beware! In
the case of variables with different values (e.g.studyno, version,
doi, etc..),
* the value of the first dataset (here, main) is retained (this can
be checked by selecting only duplicate cases, fduplicate =1 - see
below).
MATCH FILES file = main / file = addon
      /by id_cocas.
EXECUTE.
DATASET CLOSE main.
DATASET CLOSE addon.


* check if no. of duplicate cases is the same as before: .
FREQUENCIES fduplicate.

* check which value of study no. (ZA7500 for main, or ZA7502 for
addon) is retained for duplicate cases.
TEMPORARY.
SELECT IF fduplicate=1.
FRE studyno.
```

Recommended procedure in Stata:

```stata
* set folder
global data "[Insert the file path to the EVS datasets]"

* load main dataset
u "$data/[insert name of EVS Integrated dataset].dta"
tab fduplicate // check how many duplicate cases are in this da-
taset.

* merge matrix design data
merge 1:1 id_cocas using "$data/[insert name of EVS Integrated da-
taset - Matrix design data].dta"
tab fduplicate // no. of duplicate cases should be the same as
above.
```

Recommended procedure in Rstudio

```r
# load packages (install if necessary)

##install.packages("haven")
##install.packages("dplyr")
library(haven) # to import spss and stata files
library(dplyr)

# load folder and data
setwd("[Insert the file path to the EVS datasets]")
main <- read_spss("[insert name of EVS Integrated dataset]")
addon <- read_spss("[insert name of EVS Integrated dataset - Matrix
design data]")

# check duplicate cases
count(main, fduplicate)

# different number of columns
flag <- colnames(main) %in% colnames(addon)
add_cols <- unique(colnames(main))[!flag]
unique(colnames(addon))[!flag]

# add empty columns to addon
for(i in add_cols){
  addon[,i] <- NA
}

#Necessary operations before merging:
# 1) filter duplicates out of addon
addon <- addon %>%
  filter(fduplicate==0)
```

```
# ALTERNATIVELY
# 2) drop protocol variables [as they are dataset specific]
main <- main %>%
  select(-c(1:3))
addon <- addon %>%
  select(-c(1:3))


# merge data (key variable: id_cocas)
merged <- main %>%
  union(addon, by = "id_cocas")

# check duplicate cases
count(merged, fduplicate)
```