Proposal for one 3-group SB-MTMM experiment in ESS Round 9

Anna DeCastellarnau Melanie Revilla Wiebke Weber

Universitat Pompeu Fabra Barcelona

1 Introduction

Questionnaire designers should take a lot of different decisions when designing a survey question including the formulation of the questions, the type of response scales and the mode of data collection. For instance, questionnaire designers need to determine how many answer options to propose or whether to make the question indirect or direct. Each choice is important because it can affect the respondents' answers and therefore the conclusions. For instance, Saris and Gallhofer (2007) showed that the same questions, asked in the same country, in the same survey and to the same people, lead to opposite conclusions (significant negative correlation or significant positive correlation) just because the number of response categories changed (p. 174). These differences can be explained by the different size of measurement errors when using different scales.

Borgatta and Bohrnstedt (1980) define measurement errors as a 'function of the fit between the manifest scale and the latent construct' (p. 153). Two types can be distinguished: random measurement errors, due to unintended and unpredicted mistakes of the respondents, interviewers or coders; and systematic measurement errors, due to the reaction of respondents to the variation of the method used (also called method effect). The higher the measurement errors are, the lower the measurement quality of a question is. Measurement quality is defined as the strength of the relationship between the latent variable of interest (e.g. satisfaction with democracy) and the observed answers to the survey question asked to measure this latent concept (e.g. How satisfied are you with the way the democracy works in your country? 1-Very satisfied, 2-Satisfied, 3-Dissatisfied, 4-Very dissatisfied). Said differently, measurement quality is the proportion of explained variance due to the latent concept of interest. It is thus the complement of measurement errors. The observed variable will only measure perfectly the latent variable of interest (i.e. the quality will be 1) when random and systematic errors are zero. This is very unlikely. In fact, (Andrews 1984) found that 'about two-thirds of the survey measures examined contained between 50 percent and 83 percent valid variance' (p. 425).

Because surveys are commonly affected by both types of measurement errors, it is crucial for any survey to have information about their size (Saris and Gallhofer 2014). First, this information is useful to develop better survey questions (Revilla, Zavala-Rojas, and Saris 2016). However, even if the best possible survey questions were developed based on this knowledge, there will still be some errors. Thus, it is also necessary to correct for measurement errors in order to avoid misleading conclusions in substantive research (Saris and Revilla 2016). This correction can be done in a simple way, as long as we first have information about the size of the measurement errors for the questions of interest (DeCastellarnau and Saris 2014).

This information about the size of the errors can be obtained by using a Multitrait-Multimethod (MTMM) approach, which consists in repeating several questions (each measuring different concepts, also called traits) using different measurement methods (e.g. response scales) for the same respondents. In that way the quality of survey questions can be estimated (see section 1.1 for more details).

Therefore, since 2001, the European Social Survey (ESS) has conducted MTMM experiments in each round in order to estimate the measurement quality of its survey questions. However, to estimate the measurement

quality of all ESS questions using an MTMM approach, we would need to ask every question twice to the respondents. In practice, this is not possible. Thus, in each round, a selection of questions is proposed to be part of an MTMM experiment. The questions evaluated in each round in terms of MTMM are presented here: http://www.europeansocialsurvey.org/docs/methodology/ESS1-8_mtmm_experiments.pdf

However, for the other questions, information about their measurement quality can be obtained, because the information obtained through the different MTMM experiments has been used to develop a tool allowing to predict the measurement quality of new questions. This tool, the Survey Quality Predictor (SQP), allows users to get a prediction of the quality of survey questions from the characteristics of the formulation of the survey question, before the data is collected (Saris 2013). SQP is a free online software accessible at: sqp.upf.edu. The predictions provided by SQP can be used as an alternative, when no MTMM data is available. These predictions are as good as the data used to obtain them, and as rich as the amount of variation available in the MTMM data. Therefore, the ESS keep on implementing MTMM experiments to gather more information about the measurement quality of questions in different countries and languages, about different topics, using different types of scales and different interactions between all these characteristics. This ultimately will allow enriching the prediction power of SQP.

In the ESS Round 9, we propose a 3-group Split-Ballot (SB) MTMM experiment on European's understandings and evaluations of democracy. Following we present the Split-Ballot Multitrait-Multimethod design and the rules for designing such experiments. In Section 2 we present the proposal for the ESS Round 9 experiment, first focusing on the selection of the questions and later on the methods. Finally, in Section 3 we present the implementation design of the experiments in ESS Round 9.

1.1 Split-Ballot (SB) Multitrait-Multimethod (MTMM) design

Campbell and Fiske (1959) first proposed the design of repeating different questions using different methods, in order to study their discriminant validity. Then, Andrews (1984) and Goldberger and Duncan (1973) applied structural equation modeling to the MTMM design to estimate random and correlated survey measurement error variance. Saris and Andrews (1991) extended this to the true-score MTMM model which allows estimating random, systematic and correlated measurement error (see also: Saris and Gallhofer, 2007).

For empirical identification, it is usually recommended to use three survey questions each measured with three different methods (e.g. number of response options, amount and length of the labels used, single questions or batteries). However, in order to avoid that a given respondent had to answer three times to the same question (asked using different methods) and thus to reduce the response burden, the length of the questionnaire and the memory effects, Saris, Satorra, and Coenders (2004) proposed to combine the MTMM design with a SB randomization: respondents are randomly assigned to different groups, each group gets two methods, i.e. each respondent answers twice the same question.

Based on the results of the ESS Round 8 Pilot study and the UPF recommendations to improve the quality estimation (DeCastellarnau et al. 2016), the full 3-group SB design is implemented from ESS Round 8. The full 3-group SB design is presented in Table 1.

Table 1: Full 3-group SB-MTMM design

	Time 1	Time 2
Group 1	Method 1	Method 2
Group 2	Method 2	Method 3
Group 3	Method 3	Method 1

With this design, the sample is divided into three random groups. Group 1 answers at Time 1 to Method 1 and to Method 2 at Time 2. Group 2 answers to Method 2 at Time 1 and to Method 3 at Time 2. Finally, Group 3 answers to Method 3 at Time 1 and to Method 1 at Time 2. Each group should be approximately a third of the overall sample. We can notice that at a given time point, each method is asked to around a third of the sample, meaning that for substantive purposes the use of these questions will be limited to a third of the sample.

1.2 Rules for designing a SB-MTMM experiment

The general MTMM rules for the choice of the questions are the following:

- The three questions should be measured with a similar scale. Similar, in terms of the formal characteristics which should be identical: type of scale (agree-disagree or item-specific), the number of points, the kind of labels (fully or partially labelled), the layout, etc.
- To prevent estimation problems, the three questions chosen should be correlated, but their correlations should not be too similar (Saris and Satorra 2018).
- The questions should allow for sufficiently different formulations. For instance, fact or background questions usually do not allow for multiple formulation of the same question maintaining its meaning.

Moreover, our criterion for the choice of ESS Round 9 questions is based on the following rules:

- To increase the variation of topics in SQP, the questions should not have been part of any MTMM experiment conducted in previous rounds.
- Because of the 3-group SB design, the questions should not belong to ESS Round 9 Main Questionnaire.

The choice of formulation of the questions is based on the following rules:

- To increase the variation on the characteristics in SQP, the formulations should account for the characteristics that have been evaluated to a lesser extent by means of MTMM experiments in previous rounds, as long as these characteristics can be implemented within the ESS standards.
- The design of the questions can also be decided to evaluate different formulations to aid questionnaire design in the next round.
- All three methods should not be too similar. Otherwise, we may expect correlations across methods, which can lead to empirical non-identification.

The implementation of the experiments should consider the following rules:

- The questions at Time 1 should be provided in the first modules of the Main Questionnaire, as much at the beginning of the questionnaire as possible. The repetitions at Time 2 will be placed at the end, after the Human Values section and before the Interviewer questionnaire section.
- As far as possible, the questions at Time 1 should be presented next to questions with a similar topic in the Main Questionnaire.
- An introduction should be used before starting with the repeated questions at Time 2. The introduction should be as follows: "To help us improve our questions in the future, here are some final questions which are similar to previous ones. Please don't try to remember what you answered before but treat them as if they were completely new questions."

2 Proposal of a 3-group SB-MTMM experiment in ESS Round 9

2.1 The choice of questions

Even if many questions have already been evaluated by means of an MTMM experiment in the ESS, there have been many more questions asked for which the measurement quality estimation has not been obtained. To choose three new questions for the ESS Round 9 experiment, we decided to take them from previous rotating modules.

Based on the rules described in the section above, we propose three questions from the ESS Round 6 Rotating Module about Europeans' understandings and evaluations of democracy, that measure the evaluation of the complex concepts equality and vertical accountability¹. The three sub concepts chosen to measure the evaluation of these complex concepts, and the formulations used in the ESS Round 6 to ask these sub concepts are presented in Table 2.

Table 2: Concepts and questions' formulations

	Complex concept	Sub concept	Formulation				
Question 1	Vertical	Retrospective	Governing parties in [country] are punished				
Question	accountability	accountability	in elections when they have done a bad job				
Question 2	Equality	Welfare	The government in [country] protects all				
Question 2	Equality	wellare	citizens against poverty				
Question 3 Vertical Tran		Transparance	The government in [country] explains its				
Question 3	accountability	Transparency	decisions to voters				

Originally, in the ESS Round 6, these 3 questions were provided as part of a larger battery of questions (E17-E29)². Thus, they are measured with the same scale. From this battery, we discarded questions E20, E21 and E22 because they were already part of an MTMM experiment in ESS Round 6. We propose questions E26, E27 and E28 because they are correlated but their correlations are not too similar (see Appendix A).

¹ Link to ESS Round 6 Question Module Design Final Template on "Europeans' understandings and evaluations of democracy":

http://www.europeansocialsurvey.org/docs/round6/questionnaire/ESS6_final_understandings_and_evaluation_of_democracy_module_template.pdf

² Link to ESS Round 6 Source Main Questionnaire:

Moreover, we found that these questions were slightly more used in publications than other items in the same battery, so we can expect that they will be useful to potential users (see Appendix B)³.

The original request for an answer for these questions was: 'And using the same card, please tell me to what extent you think each of these following statements applies in [country]'.

2.2 The choice of the methods

For the purpose of this experiment we propose to maintain the ESS Round 6 original formulation of the questions (E26, E27 and E28) as Method 1, except for the introduction and the showcards' layout. The scale used in ESS Round 6 was from '0 Does not apply at all' to '10 Applies completely'.

The introduction needs to be adapted from its original formulation: 'Now some questions about the same topics, but this time about how you think democracy is working in [country] today. Again, there are no right or wrong answers, so please just tell me what you think". Our proposed formulation is: 'Now some questions about how you think democracy works in [country] today'. First, we propose to remove the second sentence about the "right or wrong answers" to simplify, because in Section I we have to take into account that an introduction to the repetitions will already provided. Moreover, we have decided to change the formulation "democracy is working" to "democracy works" because we are interested in how it works in general not in how it is working now, which is similar to the purpose of the questions about satisfaction with the democracy in your country, usually asked in the ESS as "democracy works" (see: B30 in ESS Round 84, variable: stdfem).

Moreover, we propose to vary the showcards' layout from its original design. We suggest the labels are presented without overlap between the verbal labels and the numbers in the showcard; and that the verbal labels are aligned in the centre. That variation is proposed to make the scale clearer to the respondents. We belief that without overlap, it is clearer that the verbal labels only belong to one of the options, while aligning the verbal labels to the centre, is more natural, than having one label aligned at the right and the other at the left, since languages are often read or from right to left or from left to right.

In this experiment we propose varying from Method 1 the following characteristics:

- The number of points in the response scale (even vs. uneven)
- The use of a general Apply scale (i.e. scales ranging from "does not apply at all" to "applies completely") for all questions or item-specific (IS) scales for each question
- The provision of a scale in a battery or as separate questions
- The number of fixed reference points
- The presentation of the question on the showcard or not

³ The item use count for the complete battery was provided by Brina Malnar. Malnar, B. (2017). *European Social Survey bibliographic monitoring - Annual report 2017*. London: European Social Survey ERIC.

⁴ Link to ESS Round 8 Source Main Questionnaire:

In uneven answer scales (e.g. 11-point scales) ranging from zero to an end (i.e. unipolar scale), the middle category (5) does not have a special meaning (not more than any other point): it cannot be interpreted as a neutral position (Krosnick and Fabrigar 1997). However, it is common to still use these uneven scales for unipolar questions, where the middle category can have a very diverse use across respondents. Given that these three questions measure a unipolar concept which ranged from "not at all" to "completely" we first propose to vary the number of points in a scale from an uneven (11-point) to an even (10-point) response scale to explore the effect of having an implicit middle category point in the uneven and unipolar scale.

Similar to agree-disagree scales, Apply scales can be provided for all questions and be presented in a battery to reduce the length of the questionnaire. However, it has been suggested using instead Item specific scales tailored for each question (Saris et al. 2010). Thus, we propose to compare the presentation of a general Apply scale presented in a battery for all questions versus IS scales developed and presented for each question separately.

Finally, to assure enough variability to the different methods and add enough interactions between the characteristics, we propose to vary the number of fixed reference points, i.e. the number of points whose meaning is clear and similar to all respondents, and the fact to provide the question in the showcard.

Summarizing, we propose the following three methods:

- Method 1: 11-point Apply scale in a battery with two fixed reference points, implicit midpoint, and not providing the question in the showcard (SC).
- Method 2: 11-point IS scale with separate questions, two fixed reference points and implicit midpoint, but providing the question in the SC.
- Method 3: 10-point IS scale with separate questions, only one fixed reference point, no midpoint and without the question in the SC.

Table 3 presents the summary of the characteristics that vary and its formulations.

Table 3: Variation in the characteristics and formulation for each method

	Method 1	Method 2	Method 3
	11-point "Apply" scale	11-point IS scale	10-point IS scale
Variation in the	Battery of questions	Separate questions	Separate questions
characteristics	2 Fix Ref. Points	2 Fix Ref. Points	1 Fix Ref. Point
Characteristics	Implicit Midpoint	Implicit Midpoint	No Midpoint
	Question not in SC	Question in SC	Question not in SC
Formulation of the scale (Trait 1)	Does not apply at all 10 Applies completely	O The governing parties are not punished at all in elections when they have done a bad job 10 The governing parties are punished completely in elections when they have done a bad job	O The governing parties are not punished at all in elections when they have done a bad job The governing parties are punished in elections when they have done a bad job

Formulation of the scale	0 Does not apply at all	0 The government does not protect citizens against poverty at all 	0 The government does not protect citizens against poverty at all
(Trait 2)	10 Applies completely	10 The government protects citizens against poverty completely	9 The government protects citizens against poverty
Formulation of the scale	0 Does not apply at all	0 The government does not explain its decisions to voters at all	0 The government does not explain its decisions to voters at all
(Trait 3)	10 Applies completely	10 The government explains its decisions to voters completely	9 The government explains its decisions to voters

Note: The scales shall be presented horizontally.

Given the structure of the ESS Round 9 questionnaire, our proposal is to provide the Time 1 questions of this experiment within Section C, just after the questions about EU referendums, and the Time 2 questions in Section I, after the Human values sections. With this design, each respondent will get a total of 6 additional questions, three in Section C and three more in Section I. Appendix C provides a table with the question names used by section, method and split-ballot group.

3 Design of the 3-group SB-MTMM experiment in ESS Round 9 about democracy

Hereafter we present the design and implementation of the MTMM democracy experiment proposed for ESS Round 9.

3.1 Method 1 (Group 1): Battery 11-point Apply scale (as in ESS Round 6: E26-E28)

Now some questions about how you think democracy works in [country] today.

CARD 26 Using this card, please tell me to what extent you think each of these statements applies in [country]. **READ OUT EACH STATEMENT**

		Does apply									-	plies letely	•	(Don't know)
C32	Governing parties in [country] are punished in elections when they have done a bad job.	00	01	02	03	04	05	06	07	08	09	10	77	88
C33	The government in [country] protects all citizens against poverty.	00	01	02	03	04	05	06	07	80	09	10	77	88
C34	The government in [country] explains its decisions to voters.	00	01	02	03	04	05	06	07	08	09	10	77	88

3.2 Time 1 - Method 2 (Group 2): Separate questions 11-point IS scale Now some questions about how you think democracy works in [country] today.

CARD 27 Using this card, please tell me to what extent you think the governing parties in [country] are punished in elections when they have done a bad job?

The governing parties are not punished at all in elections when they have done a bad job						•	ned con	overnin npletely have do	y in ele	ctions	(Refusal)	(Don't know)
00	01	02	03	04	05	06	07	80	09	10	77	88

C36 CARD 28 Now using this card, please tell me to what extent you think the government in [country] protects citizens against poverty?

The government does not protect citizens against poverty at all							omplet	ely pro	_	tizens	(Refusal)	(Don't know)
00	01	02	03	04	05	06	07	08	09	10	77	88

C37 CARD 29 And using this card, please tell me to what extent you think the government in [country] explains its decisions to voters?

The government does not explain its decisions to voters at all								The comple decisi	etely ex	plains	(Refusal)	(Don't know)
00	01	02	03	04	05	06	07	08	09	10	77	88

3.3 Time 1 - Method 3 (Group 3): Separate questions 10-point IS scale

Now some questions about how you think democracy works in [country] today.

CARD 30 Using this card, please tell me to what extent you think the governing parties in [country] are punished in elections when they have done a bad job?

The governing parties are not punished at all in elections when they have done a bad job						punis	shed in e	ing part elections done a b	s when	(Refusal)	(Don't know)
00	01	02	03	04	05	06	07	08	09	77	88

C39 CARD 31 Now using this card, please tell me to what extent you think the government in [country] protects citizens against poverty?

The government does not protect citizens against poverty at all							pr	ne gover otects c gainst p	itizens	(Refusal)	(Don't know)
00	01	02	03	04	05	06	07	08	09	77	88

CARD 32 And using this card, please tell me to what extent you think the government in [country] explains its decisions to voters?

The government of the control of the	its dec		not					•	ains its	(Refusal)	(Don't know)
00	01	02	03	04	05	06	07	08	09	77	88

3.4 Time 2 - Method 2 (Group 1): Separate questions 11-point IS scale

To help us improve our questions in the future, here are some final questions which are similar to previous ones. Please don't try to remember what you answered before but treat them as if they were completely new questions.

Now some questions about how you think democracy works in [country] today.

CARD 70 Using this card, please tell me to what extent you think the governing parties in [country] are punished in elections when they have done a bad job?

The governing parties are not punished at all in elections when they have done a bad job						The governing parties are punished completely in elections when they have done a bad job					(Refusal)	(Don't know)
00	01	02	03	04	05	06	07	80	09	10	77	88

CARD 71 Now using this card, please tell me to what extent you think the government in [country] protects citizens against poverty?

protec		ent do ns agai				С	omplet	ely pro	_	tizens	(Refusal)	(Don't know)
00 01 02 03 04 05							07	08	09	10	77	88

CARD 72 And using this card, please tell me to what extent you think the government in [country] explains its decisions to voters?

explai		nent do ecisions all						comple	e gover etely ex ons to	(Refusal)	(Don't know)	
00 01 02 03 04					05	06	07	08	09	10	77	88

3.5 Time 2 - Method 3 (Group 2): Separate questions 10-point IS scale

To help us improve our questions in the future, here are some final questions which are similar to previous ones. Please don't try to remember what you answered before but treat them as if they were completely new questions.

Now some questions about how you think democracy works in [country] today.

CARD 73 Using this card, please tell me to what extent you think the governing parties in [country] are punished in elections when they have done a bad job?

not pui election	nished a	n they h				punis	hed in e	ing part elections done a b	s when	(Refusal)	(Don't know)
00	01	02	03	04	05	06	07	08	09	77	88

CARD 74 Now using this card, please tell me to what extent you think the government in [country] protects citizens against poverty?

The goverty	citizen						pr	ne gover otects c gainst p	itizens	(Refusal)	(Don't know)
00	01	02	03	04	05	06	07	08	09	77	88

CARD 75 And using this card, please tell me to what extent you think the government in [country] explains its decisions to voters?

The government does not explain its decisions to voters at all								•	ains its	(Refusal)	(Don't know)
00	01	02	03	04	05	06	07	08	09	77	88

3.6 Time 2 - Method 1 (Group 3): Battery 11-point Apply scale (as in ESS Round 6: E26-E28)

To help us improve our questions in the future, here are some final questions which are similar to previous ones. Please don't try to remember what you answered before but treat them as if they were completely new questions.

Now some questions about how you think democracy works in [country] today.

CARD 76 Using this card, please tell me to what extent you think each of these statements applies in [country]. **READ OUT EACH STATEMENT**

		Does apply									-	plies letely	•	(Don't know)
17	Governing parties in [country] are punished in elections when they have done a bad job.	00	01	02	03	04	05	06	07	08	09	10	77	88
18	The government in [country] protects all citizens against poverty.	00	01	02	03	04	05	06	07	08	09	10	77	88
19	The government in [country] explains its decisions to voters.	00	01	02	03	04	05	06	07	80	09	10	77	88

References

- Andrews, Frank M. 1984. "Construct Validity and Error Components of Survey Measures: A Structural Modelling Approach." *Public Opinion Quarterly* 48(2): 409–42.
- Borgatta, E. F., and G. W. Bohrnstedt. 1980. "Level of Measurement: Once Over Again." Sociological Methods & Research 9(2): 147–60.
- Campbell, Donald T., and Donald W. Fiske. 1959. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrices." *Psychological Bulletin* 56(2): 81–105.
- DeCastellarnau, Anna, Melanie Revilla, Diana Zavala-Rojas, and Wiebke K. Weber. 2016. European Social Survey, Universitat Pompeu Fabra Results of the ESS Pilot Round 8 Split-Ballot Multitrait-Multimethod Experiments. Barcelona.
- DeCastellarnau, Anna, and Willem E. Saris. 2014. "A Simple Way to Correct for Measurement Errors in Survey Research." *European Social Survey Education Net (ESS Edunet)*. http://essedunet.nsd.uib.no/cms/topics/measurement/.
- Goldberger, Arthur Stanl, and Otis D. Duncan. 1973. Structural Equation Models in the Social Sciences. New York: Seminar Press.
- Krosnick, Jon A., and Leandre R. Fabrigar. 1997. "Designing Rating Scales for Effective Measurement in Surveys." In *Survey Measurement and Process Quality*, eds. Lars E. Lyberg et al. Hoboken, NJ, US, 141–64.
- Revilla, Melanie, and Willem E. Saris. 2013. "The Split-Ballot Multitrait-Multimethod Approach: Implementation and Problems." *Structural Equation Modeling: A Multidisciplinary Journal* 20(1): 27–46.
- Revilla, Melanie, Diana Zavala-Rojas, and Willem E. Saris. 2016. "Creating a Good Question: How to Use Cumulative Experience." In *The SAGE-Handbook of Survey Methodology*, eds. Christof Wolf, Dominique Joye, Tom W. Smith, and Yang-Chih Fu. SAGE, 236–54.
- Saris, Willem E. 2013. "The Prediction of Question Quality: The SQP 2.0 Software." In Understanding Research Infrastructures in the Social Sciences, eds. Brian Kleiner et al. Zurich: Seismo Press, 135–44.
- Saris, Willem E., and Frank M. Andrews. 1991. "Evaluation of Measurement Instruments Using a Structural Modeling Approach." In *Measurement Errors in Surveys*, eds. Paul P. Biemer et al. New York: John Wiley and Sons, Inc., 575–98.
- Saris, Willem E., and Irmtraud N. Gallhofer. 2007. Design, Evaluation, and Analysis of Questionnaires for Survey Research. Hoboken, NJ, US: John Wiley and Sons, Inc.
- ——. 2014. Design, Evaluation, and Analysis of Questionnaires for Survey Research. Second Ed. Hoboken, NJ, US: John Wiley and Sons, Inc.
- Saris, Willem E., and Melanie Revilla. 2016. "Correction for Measurement Errors in Survey Research: Necessary and Possible." Social Indicators Research 127(3): 1005–1020.
- Saris, Willem E., Melanie Revilla, Jon A. Krosnick, and Eric M. Shaeffer. 2010. "Comparing Questions with Agree/Disagree Response Options to Questions with Item-Specific Response Options." Survey Research Methods 4(1): 61–79.
- Saris, Willem E., Albert Satorra, and Germà Coenders. 2004. "A New Approach to Evaluating the Quality of Measurement Instruments: The Split-Ballot MTMM Design." *Sociological Methodology* 34(1): 311–47.
- Saris, Willem E, and Albert Satorra. 2018. "The Pooled Data Approach for the Estimation of Split-Ballot Multitrait–Multimethod Experiments." *Structural Equation Modeling: A Multidisciplinary Journal* 0: 1–14.

Appendix A: Correlation matrix between selected questions in ESS Round 6 (for the integrated data).

Correlation	E26 gptpelcc	E27 gvctzpvc	E28 gvexpdcc
E26 gptpelcc	1		
E27	0.516***	1	
gvctzpvc	(50,884)	ı	
E28	0.527***	0.725***	1
gvexpdcc	(50,506)	(52,100)	I

^{*} p < .05 ** p < .01 *** p < .001

Appendix B: 2003-2016 item publications count obtained from Brina Malnar for all the ESS Round 6 battery (in grey the proposed items)

Questionnaire Name	Variable Name	Number of publications
E17	fairelcc	7
E18	dspplvtc	4
E19	dfprtalc	4
E20	oppcrgvc	7
E21	medcrgvc	7
E22	meprinfc	6
E23	rghmgprc	6
E24	votedirc	6
E25	cttresac	7
E26	gptpelcc	7
E27	gvctzpvc	7
E28	gvexpdcc	11
E29	grdfincc	8

Appendix C: Question names used in this proposal from Section C and their repetition in Section I by method and SB group.

	Tin	ne 1 - Section	n C	Time 2- Section I					
	Method 1	Method 2	Method 3	Method 1	Method 2	Method 3			
	C32				I1				
Group 1	C33				12				
	C34				13				
		C35				14			
Group 2		C36				15			
		C37				16			
			C38	17					
Group 3			C39	18					
			C40	19					

Appendix D: Showcards

Question(s): C32, C33, C34

CARD 26

Does not						Applie	es				
apply at all	l									comple	tely
0	1	2	3	4	5	6	7	8	9	10	

To what extent do you think that the governing parties in [country] are punished in elections when they have done a bad job?

									The governing
									parties are
									punished
									completely in
									elections when
									they have done
									a bad job
1	2	3	1	5	6	7	8	a	10
	1	1 2	1 2 3	1 2 3 4	1 2 3 4 5	1 2 3 4 5 6	1 2 3 4 5 6 7	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8 9

To what extent do you think the government in [country] protects citizens against poverty?

The government government does not protects protect citizens against poverty at all The government poverty completely

0 1 2 3 4 5 6 7 8 9 10

To what extent do you think that the government in [country] explains its decisions to voters?

The											The
governmen	t										government
does not											explains its
explain its											decisions to
decisions to	0										voters
voters at al	I										completely
0		1	2	3	4	5	6	7	8	9	10

The governing parties are not punished at all in elections when they have done a bad job

The governing parties are punished in elections when they have done a bad job

0 1 2 3 4 5 6 7 8 9

The government The government does not protect protects citizens citizens against against poverty poverty at all 0 1 2 3 4 5 6 7 8 9

The government does not explain its decisions to voters at all

The government explains its decisions to voters

0 1 2 3 4 5 6 7 8 9

Question(s): I7, I8, I9

CARD 76

Does not apply at all										Applies completely		
0	1	2	3	4	5	6	7	8	9	10		

To what extent do you think that the governing parties in [country] are punished in elections when they have done a bad job?

The governing parties are not punished at all in elections when they have done a bad job

The governing parties are punished completely in elections when they have done a bad job

0 1 2 3 4 5 6 7 8 9 10

To what extent do you think the government in [country] protects citizens against poverty?

The government government does not protects citizens against poverty at all The government poverty completely

0 1 2 3 4 5 6 7 8 9 10

To what extent do you think that the government in [country] explains its decisions to voters?

The governn does nexplain decision voters a	nent ot its is to										The government explains its decisions to voters completely
	0	1	2	3	4	5	6	7	8	9	10

The governing parties are not punished at all in elections when they have done a bad job

The governing parties are punished in elections when they have done a bad job

0 1 2 3 4 5 6 7 8 9

The government The government does not protect protects citizens citizens against against poverty poverty at all 0 1 2 3 4 5 6 7 8 9

The government does not explain its decisions to voters at all

The government explains its decisions to voters

0 1 2 3 4 5 6 7 8 9