

Grading Generalized Trust Across Europe

René Bekkers & Bart Sandberg

February 19, 2018

Center for Philanthropic Studies, Vrije Universiteit Amsterdam

Abstract

We report a mega-analysis of generalized social trust responses by respondents ($n = 330,526$) in the European Social Survey (2002-2014) and respondents in other surveys ($n = 317,833$), conducted in the same years in the same countries. We find markedly higher levels of trust among the ESS respondents than among respondents in other surveys. We report evidence from a comparison that the use of the 0-10 scale is the most important cause of the higher level of trust among ESS respondents. We find that surveys conducted in the same years and countries as the ESS but forcing participants to make a choice between ‘most people can be trusted’ and ‘you cannot be too careful’ yield lower levels of trust. Offering participants a middle option, ‘it depends’, mitigates this effect somewhat, but does not eliminate it. Also the use of Likert-type scales ranging from 1 to 4 or 1 to 5 yield lower levels of trust than the 0-10 scale used in the ESS and other surveys. We discuss potential causes for the higher level of trust on 0-10 scales and discuss the implications for the measurement of other social attitudes.

Paper prepared for the 6th ESS Workshop, “Samenhang in Europa: eenheid in verscheidenheid”, March 16, 2018, The Hague. In this paper we make use of a large number of datasets, including the European Social Survey and the LISS (Longitudinal Internet Studies for the Social sciences) panel administered by CentERdata (Tilburg University, The Netherlands). A full listing of all data sources is included in the Appendix.

Materials of this project are available through the Open Science Framework, at <https://osf.io/c9mq4/>. The paper builds on work by the Global Trust Research Consortium, <https://globaltrustresearch.wordpress.com/>. Suggestions for data to be added are most welcome at globaltrustresearchconsortium@gmail.com.

Natural experiments in the measurement of trust

Generalized social trust is an indicator of social capital, and is correlated with functioning in numerous life domains (work, health, income, relationships) and with subjective well-being. Given its importance for society, hundreds of surveys around the globe have included measures of generalized social trust in the past decades. However, the exact formulation of the question on trust and the response categories offered to participants in surveys vary between surveys. The high prevalence of the trust question in many surveys creates opportunities for social scientists to detect effects of item formulation and response format on the level of trust. Surveys fielded in the same period among the same target population but using different question wording and response scales pose natural experiments. Ideally, different variants used in surveys conducted among the same individual respondents within the same survey (within subjects), or randomly allocated to different respondents within the same survey (between subjects). In the absence of such experiments, which are extremely rare, a comparison of responses to surveys conducted at the same time in the same country but using different item formulations and response scales is a good alternative.

Using this design, we compare trust responses in the European Social Survey (ESS) with trust responses in other surveys conducted in the same years and countries. We do not test substantial predictions on correlates of trust. Neither do we test predefined hypotheses on methodological correlates. Our analysis is purely exploratory, and answers the following questions: What are the consequences of using the 11 point scale rather than other response formats for the level of trust reported by respondents? How do levels of trust at the country level measured in the ESS match levels of trust in other surveys using slightly different measures? To what extent is the ranking of countries on trust robust to the use of different response formats? What are the 11-point scale equivalents of the trusting and the distrusting responses in other formats?

The ESS has adopted a scale similar to school grades in many countries ranging from 0 to 10 to measure generalized social trust. The 0-10 response scale has been used in the ESS since its inaugural wave in 2002. The scale is different from the response options in the original measure of generalized social trust, which employs a forced-choice format. Respondents in the original measure are forced to choose between ‘most people can be trusted’ and ‘you cannot be too careful’. Still other surveys in Europe in the same years and countries have used 4 point scales as well as 5 point scales to measure trust.

Data and methods

To answer these questions, we conduct a mega-analysis relying on a dataset of harmonized trust responses from 648,359 respondents in 22 different surveys in 31 countries, compiled by the Global Trust Research Consortium (GTRC, 2018). Mega-analysis is known in medicine as *meta-analysis using individual patient data* (Clarke & Stewart, 1997) and in psychology as *integrative data analysis* (Curran & Hussong, 2009). The analysis reported here does not use the full GTRC database, currently containing about 2.7 million responses from 195 surveys in 148 countries. Our focus is on the countries covered by the European Social Survey. We compare responses to the question on generalized trust in the ESS data ($n = 330,526$) with responses in other surveys conducted in the same countries and years, but using different response scales ($n = 317,833$). ESS participants in six countries (Bulgaria, Estonia, Greece, Iceland, Italy, and Luxembourg) are excluded from the analyses because we have no other data sources available.

0-10 scale. Participants in the ESS responded to the following question: “Using this card, generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people? Please tell me on a score of 0 to 10, where 0 means you can't be too careful and 10 means that most people can be trusted.” The card displayed the numbers 0 to 10, with the negative pole displayed on the left, aligned to the number 0, and the positive pole on the right, aligned to 10 (see Figure A in the appendix). Participants were not actively offered other options, such as ‘don’t know’ or ‘it depends’.

Other surveys in Europe in the same years and countries have used the classical forced choice format, 3 categories, 4 categories, and Likert-type scales ranging from 1 to 5. Table A in the appendix shows how we recoded these versions on a scale from 0 to 1 to be compatible with each other.

Forced choice. The forced choice format has been the default version of trust questions and is often found in UK surveys, such as the British Household Panel Survey.

3 categories. A slightly different set of response categories involves the active offer of ‘It depends’ as a middle ground between ‘Most people can be trusted’ and ‘Can’t be too careful’. The addition of the middle option results in a 3 categories, as in the German General Social Survey.

4 categories. The 2004 wave of the ISSP (Citizenship I) provides a good example of this response format. The same question was asked as in the ESS, but four response options were offered: (1) ‘People can almost always be trusted’, (2) ‘People can usually be trusted’, (3) ‘You usually can't be too careful in dealing with people’, (4) ‘You almost always can't be too careful in dealing with people’. Almost the same question was included in the British Social Attitudes survey conducted in 2004 and 2014.

1-5 Likert Scale. An example of this type is the 2006 wave of the ISSP (Role of Government IV). Participants were asked to what extent they agreed or disagreed with the statement “There are only a few people I can trust completely”, and were offered 5 options: ‘Strongly agree’, ‘Agree’, ‘Neither agree nor disagree’, ‘Disagree’, ‘Strongly disagree’, as well as ‘Can’t choose’. In 2010, in contrast, the same question was asked as in the ESS, and only two labels were provided, but response options were confined to 1 (‘You can’t be too careful’), 2, 3, 4, and 5 (‘Most people can be trusted’). The Netherlands Longitudinal Lifecourse Study (NELLS) also provided a 1-5 scale, running from ‘Strongly agree’, ‘Agree’, ‘Neither agree nor disagree’, ‘Disagree’, ‘Strongly disagree’, but no other options were provided, and two separate items were proposed to participants, one for the positive pole (‘most people can be trusted’) and one for the negative pole (‘you cannot be too careful in dealing with other people’). For the current analyses we used only the first, positive, item.

Table 1 provides details of the datasets used in the analyses. The trust responses in the other surveys were provided in the multinational ISSP surveys ($n = 93,416$), the Citizenship, Involvement and Democracy surveys for Spain and Sweden ($n = 5,523$), the German General Social Survey ($n = 19,488$), Understanding Society ($n = 23,823$), Citizenship Survey ($n = 31,883$), the 1970 British Cohort Study ($n = 9,841$), the General Household Survey ($n = 8,700$), the British Social Attitudes Survey ($n = 18,756$), the Community Life Survey ($n = 4,225$), the National Child Development Survey ($n = 9,790$), the ONS Omnibus Survey ($n = 1,710$), the Taking Part Survey ($n = 47,086$), the Health Survey for England ($n = 49,909$), the Netherlands Longitudinal Lifecourse Study ($n = 1,754$), the Longitudinal Internet Studies for the Social Sciences ($n = 25,429$), the Netherlands Election Survey ($n = 4,530$), Culturele Veranderingen

(‘Cultural Changes’, $n = 8,420$), the Swedish National Election Survey ($n = 3,788$), and the Latino Barometer Survey conducted in Spain ($n = 12,405$).

Table 1. Datasets included in the analyses

Year	Survey	Scale	Country	N
2002	European Social Survey (ESS)	0-10	Multinational	11,063
2004	European Social Survey (ESS)	0-10	Multinational	34,171
2006	European Social Survey (ESS)	0-10	Multinational	31,117
2008	European Social Survey (ESS)	0-10	Multinational	46,663
2010	European Social Survey (ESS)	0-10	Multinational	24,208
2012	European Social Survey (ESS)	0-10	Multinational	7,089
2014	European Social Survey (ESS)	0-10	Multinational	7,228
2004	ISSP	1-4	Multinational	22,538
2006	ISSP	1-5	Multinational	22,375
2008	ISSP	1-4	Multinational	34,429
2010	ISSP	1-5	Multinational	14,074
2002	Citizenship, Involvement, and Democracy (CID)	0-10	Spain	4,252
2002	Citizenship, Involvement, and Democracy (CID)	0-10	Sweden	1,271
2002	German General Social Survey (GGSS)	1-3	Germany	2,820
2006	German General Social Survey (GGSS)	1-3	Germany	3,421
2008	German General Social Survey (GGSS)	1-3	Germany	3,469
2010	German General Social Survey (GGSS)	1-3	Germany	2,827
2012	German General Social Survey (GGSS)	1-3	Germany	3,480
2014	German General Social Survey (GGSS)	1-3	Germany	3,471
2010	Understanding Society	0-1	UK	23,823
2008	Citizenship Survey	1-3	UK	14,917
2010	Citizenship Survey	1-3	UK	16,966
2012	1970 British Cohort Study	0-1	UK	9,841
2004	General Household Survey	0-1	UK	8,700
2002	British Social Attitudes Survey	0-1	UK	3,425
2004	British Social Attitudes Survey	1-5	UK	3,199
2006	British Social Attitudes Survey	0-1	UK	4,290
2008	British Social Attitudes Survey	0-1	UK	4,486
2010	British Social Attitudes Survey	0-1	UK	3,297
2012	British Social Attitudes Survey	0-1	UK	3,248
2014	British Social Attitudes Survey	1-5	UK	2,878
2012	Community Life Survey	0-1	UK	4,225
2008	National Child Development Survey (NCDS)	0-1	UK	9,790
2004	ONS Omnibus Survey	1-3	UK	1,710
2006	Taking Part Survey	0-1	UK	10,608
2008	Taking Part Survey	0-1	UK	24,182
2010	Taking Part Survey	0-1	UK	9,914
2012	Taking Part Survey	0-1	UK	2,382
2002	Health Survey England	0-1	UK	18,396
2004	Health Survey England	0-1	UK	10,114
2006	Health Survey England	0-1	UK	21,399
2008	Netherlands Longitudinal Lifecourse Study (NELLS)	1-5	NL	424
2010	Netherlands Longitudinal Lifecourse Study (NELLS)	1-5	NL	1,330

2008	Longitudinal Internet Studies for the Social Sciences (LISS)	0-10	NL	6,808
2010	Longitudinal Internet Studies for the Social Sciences (LISS)	0-10	NL	6,107
2012	Longitudinal Internet Studies for the Social Sciences (LISS)	0-10	NL	5,953
2014	Longitudinal Internet Studies for the Social Sciences (LISS)	0-10	NL	6,561
2002	Netherlands Election Survey	0-1	NL	1,907
2006	Netherlands Election Survey	0-1	NL	2,623
2002	Culturele Veranderingen	0-1	NL	2,013
2004	Culturele Veranderingen	0-1	NL	2,300
2006	Culturele Veranderingen	0-1	NL	2,144
2008	Culturele Veranderingen	0-1	NL	1,963
2002	Swedish National Election Survey (SNES)	0-10	SE	3,788
2002	Latino Barometer Survey	0-1	ES	2,484
2004	Latino Barometer Survey	0-1	ES	2,491
2006	Latino Barometer Survey	0-1	ES	2,476
2008	Latino Barometer Survey	0-1	ES	2,471
2010	Latino Barometer Survey	0-1	ES	2,483

Relying on these datasets, we conduct three sets of comparisons of trust responses.

A first set compares trust responses on the 0-10 scale used in the ESS with responses from the same countries and the same years by respondents in different surveys that used different scales: the 0-1 forced choice format ($n = 141,692$), the two forced choice options with a third category ($n = 49,659$), a 4 category scale ($n = 55,552$), and a 1-5 Likert scale format ($n = 39,521$). The large numbers of respondents from the UK ($n = 233,253$ in 11 different surveys), and the Netherlands ($n = 58,405$ in four different surveys) provide the best opportunities to make within-country comparisons.

A second set compares trust responses on the 0-10 scale used in three other surveys than the ESS ($n = 37,024$: the Longitudinal Internet Studies for the Social Sciences from the Netherlands (2008-2014), the Citizenship, Involvement and Democracy survey in 2002 (Spain: $n = 4,252$; Sweden: $n = 1,271$) and the Swedish National Election Study in 2002, $n = 3,788$) with responses from the same countries in the same years in the ESS ($n = 42,166$).

A third set of comparisons does not involve ESS data, but sheds more light on the effects of response scales. In this set we compare trust responses given by respondents within the same survey on different scales. We found three such within-survey natural experiments. The first experiment is within subjects. The Citizenship, Involvement and Democracy survey in Spain asked participants to make a choice between ‘Most can be trusted’ and ‘Never too careful’, as well as to give a response on a 0-10 scale, where the extremes were labeled ‘Can’t be too careful’ and ‘Most can be trusted’. The second experiment is between subjects in different years. ISSP surveys used a 4 point scale in 2004 and 2008, but a 5 point scale in 2006 and 2010. The third experiment is also between subjects in different years. The British Social Attitudes Survey used the forced choice format in most years, except for 2004 and 2014, when a 5 point scale was used.

Results

Peculiarities in the distribution of trust responses in different surveys

Figure 1 displays the distribution of responses to the trust question on the 11-point scale in the ESS and a normal curve based on the mean (4.95) and standard deviation (2.48). We see that 5, 7, and 8 are the most popular responses, more popular than expected based on the normal distribution. The 0 response is also more popular than expected. We find lower than expected numbers of respondents choosing 4, 6, 9 and 10.

Figure 1. Trust responses in the ESS using a 0-10 response format (all years and countries, $n = 330,526$)

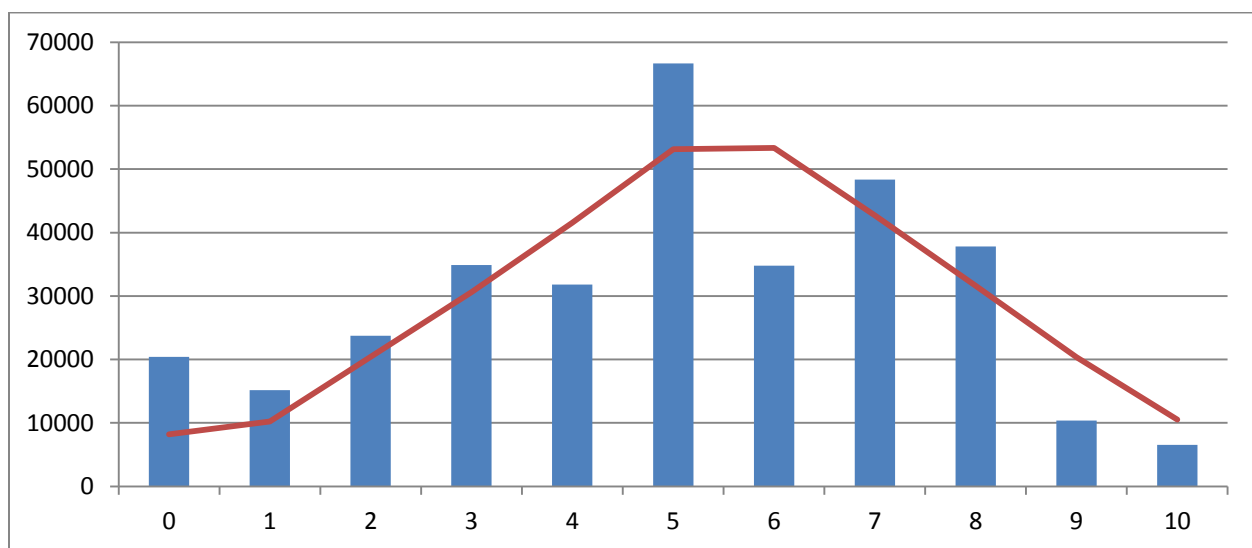
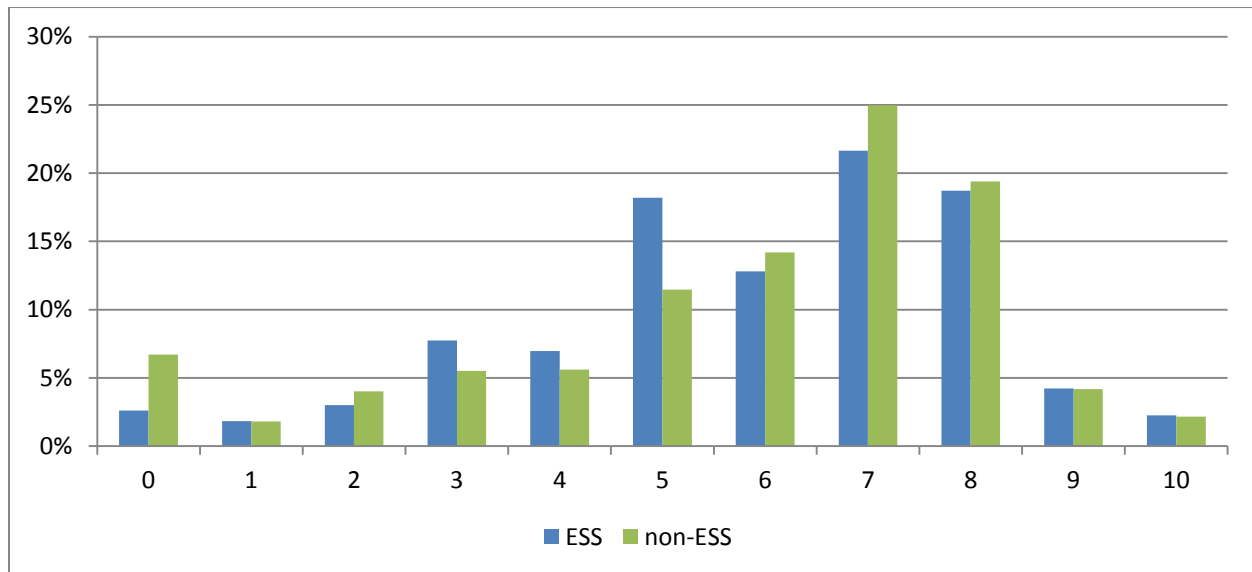


Figure 2 displays the distribution of responses in Sweden and the Netherlands to the trust question on the 11-point scale in the ESS and in the other surveys in these countries conducted in the same years, the Swedish National Election Survey (SNES, 2002) and the Longitudinal Internet Studies for the Social Sciences in the Netherlands (LISS, 2008-2014). In the ESS data for Sweden and the Netherlands 7 is more popular than the midpoint and about equally popular as 8. The distribution in the other surveys is more skewed, with the midpoint being less popular in the other surveys than in the ESS. Instead of the midpoint, larger proportions of respondents in the other surveys are choosing 0 and 7. These differences are striking and cannot be explained by item formulation and response scale differences: the three surveys asked exactly the same question, and offered exactly the same response scale to survey participants.

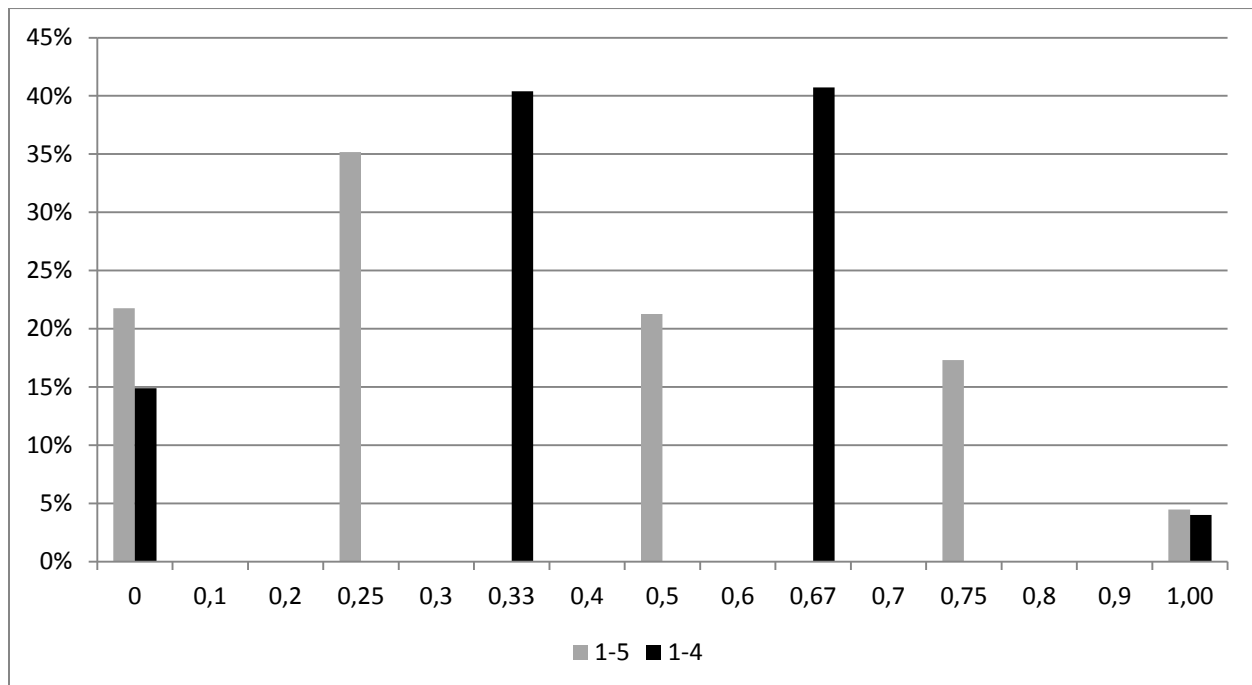
Figure 3 shows the distribution of trust responses to the 1-4 and 1-5 scales used in other surveys than the ESS. Assuming equidistance between the response categories, the responses have been rescaled to a 0-1 scale.

Figure 2. Trust responses in the ESS in Sweden and the Netherlands ($n = 4,354$), in the Swedish National Election Survey (SNES) and in the LISS surveys (the Netherlands) using a 0-10 response format ($n = 28,540$)



In Figure 3 we see a markedly different distribution in responses on the 1-4 scale (light grey bars) than on the 1-5 scale (dark grey bars). The distribution on the 1-4 scale is approximately normal, with most participants scoring either 2 or 3 (0,33 and 0,66 on the 0-1 continuum), in about equal proportion. Of the extreme options, 1 (rescaled to 0) is three times more popular than 4 (rescaled to 1).

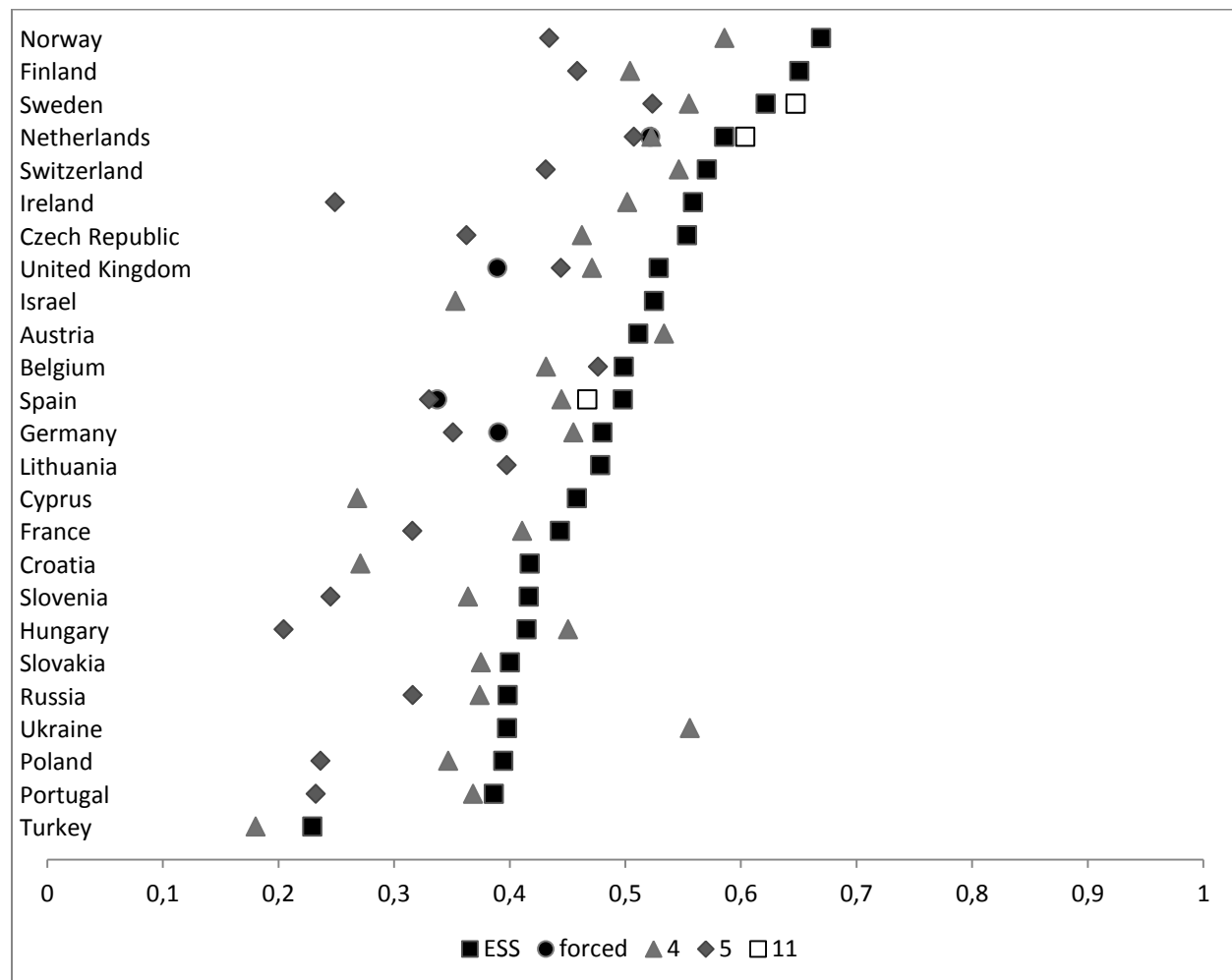
Figure 3. Distribution of trust on 1-4 ($n = 55,552$) and 1-5 scales ($n = 39,521$)



On the 1-5 scale, however, we see that the distribution is concentrated below the midpoint, with more than a third choosing 2 (0,25 on the 0-1 continuum), and another 22% choosing 1 (rescaled to 0). While the 1-5 scale has not been used in all countries, it is unlikely that country effects are causing the difference with the 1-4 scale distribution. The countries for which 1-5 scale responses are missing - Turkey, Ukraine, Slovakia, Croatia, Austria, and Israel – are not high trust countries. The regressions below (see Table 3) using country fixed effects show that the 1-5 scale format does indeed reduce the reported level of trust.

Figure 4 displays the average score for trust in the 25 ESS countries in all 22 surveys. Focusing on the ESS respondents first, we see the familiar pattern that trust is highest in the Scandinavian countries, with the Netherlands in fourth place and Switzerland ranking fifth. The bottom half of the ranking is dominated by Eastern European countries, with Portugal and respondents in Turkey being least trusting.

Figure 4. Average trust scores per country by response format



The score on the 11 point scale used in the ESS is typically higher than the score on other response scales. The scores on the forced choice format are lower for all countries in the analysis, particularly for Germany. The 1-5 scale and 1-4 scale also yield lower scores in most countries, particularly the 1-5 scale.

The score for Ireland on the 1-5 scale (0.249) is even less than half of the score on the 0-10 scale used in the ESS (0.559).

Table 2. Rank order of countries

	11 ESS	11 other	1-4	1-5	0-1
Norway	1	-	1	6	-
Finland	2	-	7	4	-
Sweden	3	1	3	1	-
Netherlands	4	2	6	2	1
Switzerland	5	-	4	7	-
Ireland	6	-	8	14	-
Czech Republic	7	-	10	9	-
United Kingdom	8	-	9	5	3
Israel	9	-	20	-	-
Austria	10	-	5	-	-
Belgium	11	-	14	3	-
Spain	12	3	13	11	4
Germany	13	-	11	10	2
Lithuania	14	-	-	8	-
Cyprus	15	-	23	-	-
France	16	-	15	13	-
Croatia	17	-	22	-	-
Slovenia	18	-	19	15	-
Hungary	19	-	12	18	-
Slovakia	20	-	16	-	-
Russia	21	-	17	12	-
Ukraine	22	-	2	-	-
Poland	23	-	21	16	-
Portugal	24	-	18	17	-
Turkey	25	-	24	-	-
Correlation with rank order in ESS		1.00	.67	.92	.40
Correlation with rank order on 1-4 scale				.70	.80
Correlation with rank order on 1-5 scale					.80

Several other countries in Figure 4 stand out with a large difference between the ESS score and the scores for other scale formats, such as Norway and Hungary. Remarkably, ISSP participants in the Ukraine and Hungary have a higher trust score on a 1-4 scale than the ESS participants in these countries on the 0-10 scale. In contrast, several other countries display systematically smaller differences in the scores responding to different scales. Participants in Austria, Slovakia, the Netherlands, Belgium and Sweden behave very similarly in surveys with different response formats.

Rank order stability in countries across different formats

Despite the occasionally large discrepancies between responses to different scale formats, the rank order of countries is rather similar when different scales are used (see Table 2). The correlation of the 11 scale score in the ESS with the 1-4 and 1-5 scale scores is .67 and .92, respectively. Compared to the scores in the ESS, Ukraine and Israel score higher on a 1-4 scale, while Belgium scores higher on a 1-5 scale, and Ireland lower. The correlation between the ranking of countries on the 1-4 scale and the 1-5 scale is not perfect (.70), but still high.

Trust is higher when a 0-10 response scale format is used

Results of regression analyses including country fixed effects reported in Table 3 reveal that trust responses in the ESS are markedly higher than trust responses in other surveys. The first column shows that the ESS scores are on average .092 higher, about 22% higher than the baseline of .411.

Table 3. Regression analyses of trust responses on survey source, response scale, and year

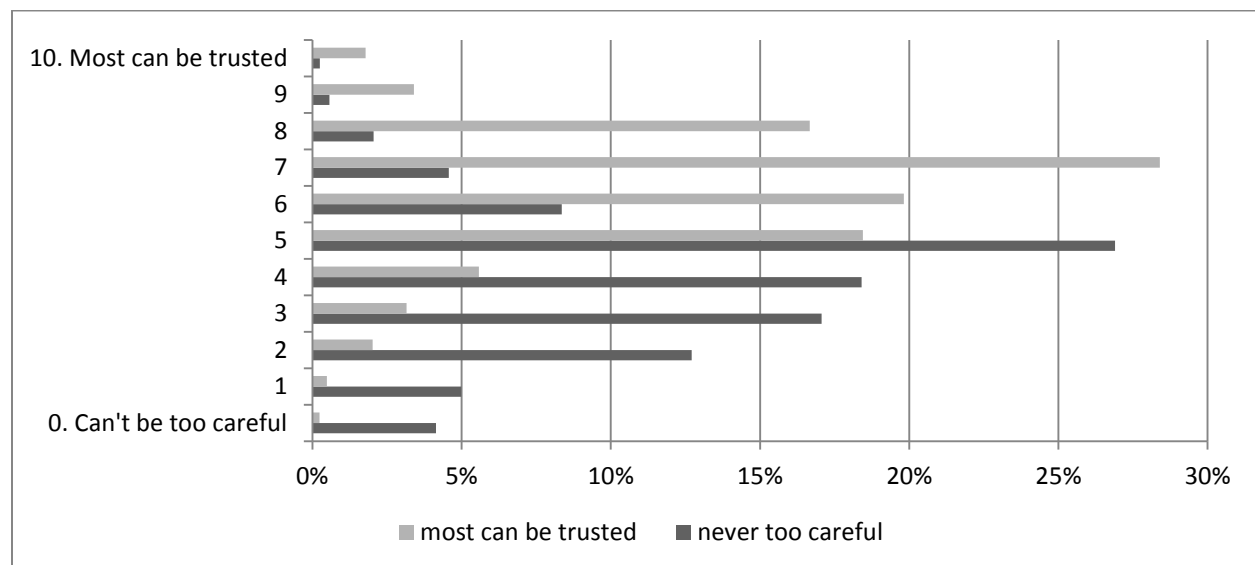
	ESS	Response scale	Year	Year by scale
ESS case	.092 (.001)	.009 (.002)	.012 (.002)	.012 (.002)
Forced choice		-.133 (.003)	-.124 (.003)	-.162 (.003)
3 categories		-.104 (.003)	-.101 (.003)	-.122 (.005)
1-4 scale		-.042 (.003)	-.034 (.003)	-.025 (.004)
1-5 scale		-.129 (.003)	-.125 (.003)	-.265 (.005)
Year since 2002			.003 (.000)	.001 (.000)
Forced * Year				.009 (.000)
1-3 scale * Year				.004 (.001)
1-4 scale * Year				-.003 (.001)
1-5 scale * Year				.024 (.001)
Constant	.411 (.001)	.502 (.002)	.479 (.002)	.492 (.002)
Country level variance	.075	.076	.076	.076
Number of respondents	648,359	648,359	648,359	648,359
Number of countries	31	31	31	31
Country fixed effects	Yes	Yes	Yes	Yes

The results in the second column show that the 0-10 response scale format is responsible for most of the difference. Respondents in surveys using a forced choice format have a trust score that is about 25% lower than respondents (-.133 / .502) who reported on a 0-10 scale. The 1-5 scale also results in trust scores that are about 25% lower than on a 0-10 scale. The 1-4 scale reduces trust much less, by about 8%. Controlling for response scale, however, ESS respondents still report a slightly higher level of trust (2%) than respondents in other surveys that use a 0-10 format.

The data from the CID participants in Spain displayed in Figure 5 provide another illustration of the negative influence of the forced choice format. The same participants responded to the statement about trust both in a forced choice format as well as on a 0-10 scale. The mean level of trust in the 0-1 format is 0.291 (29.1% reported ‘most people can be trusted’), while the 0-10 format rescaled to 0-1 yields an average of 0.467. This difference (-.176, 38% lower than the 0-10 score) is somewhat larger than in the analysis above.

Figure 5 also provides a hint about why the forced choice format lowers trust. Among those choosing ‘never too careful’ we find that the midpoint on the 0-10 scale is the most popular response. A further 16% chooses a number above 5. Only a slight majority (57%) chooses a number below the midpoint. In other words, if forced to choose between trust and distrust, people who would otherwise have taken a middle ground choose distrust. More than three quarters of respondents at the midpoint on the 0-10 scale choose ‘never too careful’ in the forced choice format.

Figure 5. Distribution of trust responses on a 0-10 scale in Spain by responses in a forced choice format (CID, 2002, n = 4,061)



Trends in trust vary with response scale format

The results in the final column of Table 3 show that year trends in trust are different depending on the scale used to measure trust. As we include interactions with the other scale formats, the main effect for year represents the time trend among respondents on a 0-10 scale, which is slightly positive (.003). The

trend is more positive for participants responding in a forced choice format, and especially so for participants responding on a 1-5 scale.

Table 4 provides further evidence that response scales alter trust responses. In the British Social Attitudes survey, two editions measured trust in a forced choice format. In these editions, the level of trust was about 15% lower than in the editions that used a 1-5 scale. In contrast to the pattern for other countries in which a forced choice format has been used, the year trend in the BSA is less positive than in the ESS surveys for the UK. An analysis of 0-10 surveys fielded in the Netherlands reveals that the year trend in the ESS was less positive than in other surveys.

Table 4. Year trends in trust in the UK and the Netherlands according to the ESS and other surveys

	UK			NL		
	ESS	BSA	All	ESS	Other 0-10	All
Year	.002 (.000)	.001 (.001)	.001 (.001)	.002 (.000)	.008 (.000)	.008 (.000)
Forced choice		-.073 (.011)	-.073 (.008)			---
ESS case			.042 (.011)			.054 (.005)
ESS * year			.001 (.001)			-.005 (.001)
Constant	.513 (.003)	.472 (.013)	.513 (.005)	.572 (.003)	.518 (.003)	.518 (.003)
Number of respondents	15,625	10,194	25,819	13,488	43,823	57,311

Table 5. Year trends, response scale format and country level variance in the ISSP and the ESS

	ISSP			ESS
	1-4	1-5	All	0-10
Year	-.002 (.001)	.032 (.001)	-.002 (.001)	.001 (.000)
1-5 scale			-.291 (.005)	---
1-5 scale * year			.035 (.001)	---
Constant	.457 (.003)	.173 (.005)	.456 (.003)	.489 (.001)
Country level variance	.149	.086	.137	.140
Number of respondents	55,552	35,483	91,035	330,526
Number of countries	24	18	25	31
Country fixed effects	Yes	Yes	Yes	Yes

Country level variance is affected by response scale format

Finally, Table 5 shows that response scales do not only affect year trends, but also the magnitude of country level variance. When the ISSP was fielded in a 1-4 scale format, the country level variance was much larger (15%) than when it was fielded in a 1-5 scale format (9%). Note that in the ISSP editions that

used the 1-5 scale, the country level variance was smaller than in the 0-10 format used in the ESS (14%). Also note that the country level variance is not linearly correlated with the number of countries in the analysis: the ISSP edition that used the 1-4 scale had only 24 countries, while the ESS included 31 countries.

Conclusion and discussion

Surveys using a 0-10 scale to measure trust yield scores that are considerably higher than surveys using other scale formats. The result is consistent with two published experiments comparing the 1-5 scales with scales including more response options. A first study consistent with our findings, among 301 students and 751 clients of taxi drivers in Australia (Dawes, 2002), found higher scores and more variance on a 0-10 scale than on a 1-5 scale. A second consistent study (Darbyshire & McDonald, 2004), among 400 clients of telecom firms in Australia, showed that the same respondents scored lower on a verbal 1-5 scale (poor, fair, good, very good, excellent) than on an unlabeled 0-10 scale included later in the same survey. A third study, on quality of 466 patient experiences in hospitals in Norway (Garratt, Helgeland & Gulbrandsen, 2011), is not exactly comparable because it tested 1-5 'all-point-defined' scales against 1-10 (rather than a 0-10) 'end-point-defined' scales. The result of the experiment, however, is consistent with our results, as the 1-10 scale produced higher and more skewed average scores than the 1-5 scale. A fourth study (Leung, 2011) on self-esteem among 13-18 year old students in Macau, yielded findings that are inconsistent with ours. 0-10 scale scores were similar to 1-5 scores, but more normally distributed.

Our analyses do not include individual respondent characteristics. This leaves the possibility that some of the differences in trust we reported may be due to compositional differences between the samples. To control for this possibility, we plan to add a set of individual respondent characteristics, including gender, age, and education, for which harmonized measures are available. We expect that compositional effects do not explain the higher level of trust observed in the ESS. Generalized social trust is likely to be related to participation in surveys. The sampling and fieldwork procedures of the ESS are well documented and of high quality. It is likely that the ESS outperformed other surveys in contacting and convincing participants, with better coverage of population segments with below average scores on trust as a result. If this line of reasoning is correct, it is likely that inclusion of individual respondent characteristics only exacerbates the difference between trust responses in the ESS and the other surveys.

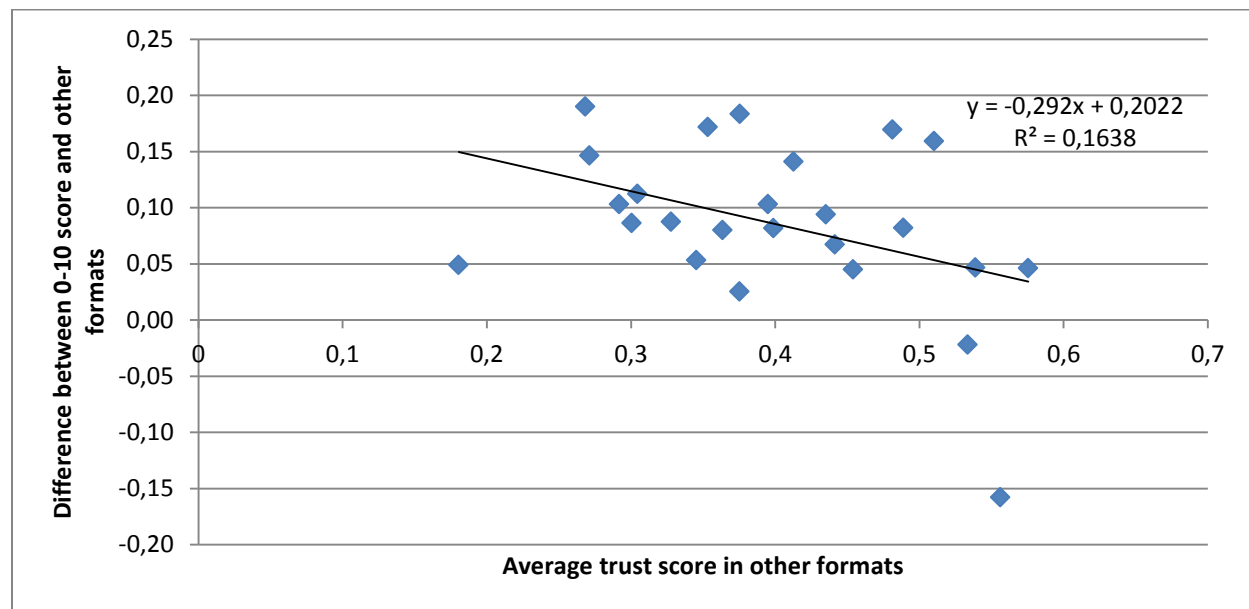
Nevertheless, inclusion of individual respondent characteristics is useful because it will allow us to answer the question to what extent correlates of trust vary with the properties of the measurement of trust by adding interactions of respondent characteristics with survey and item characteristics. It is important to check whether the level of trust drops more strongly as a result of the forced choice format in some categories of respondents than in others.

When compositional effects are excluded, it is possible to conduct a correspondence analysis (Riquier, Kennedy and Sharp, 1996) to compute Euclidean distances between the response options in various scales. This technique allows researchers to convert scores obtained in different scales to each other. Correspondence analysis is an important next step for the analysis of harmonized survey data. The analysis assumes cross-country measurement equivalence. This assumption has been tested and validated for the ESS itself (Van der Veld & Saris, 2011).

The key question that we are left with is *why* reported trust levels are higher when the 0-10 scale format is used rather than a forced choice format. The higher level of trust is not so much a result of survey mode, because it appears in both face-to-face surveys as well as in online surveys.

Figure 6 displays a scatter plot of the ‘0-10 gap’ by the level of trust. The discrepancy between the 0-10 trust score and the average trust score in other formats becomes smaller ($r = -.47$) as the 0-10 trust score is higher. The correlation is smaller when excluding Ukraine and Austria ($r = -.15$), but still negative. If we assume that the pattern that we observed in Spain (see Figure 5) can be generalized to other countries, participants in less trusting countries could be driving the 0-10 gap by choosing a midpoint to reflect their distrust, rather than a value between 0 and 4. In high trust countries the alignment between trust scores in different formats is somewhat smaller. This interpretation goes against the conceptualization of trust as a collective norm.

Figure 6. Differences between trust scores on a 0-10 scale and other formats at the country level ($n = 25$)



Another more speculative interpretation is that the gap is related to the prevalence of right-handedness in a country. The higher score on the 0-10 scale could represent a preference for the right hand side of a page or a screen on which the question is asked. A minority of Europeans is left-handed, but the size of this minority differs between countries (McManus, 2009).

In contrast to the speculation, however, Figure 7 shows that the 0-10 gap *increases* with the prevalence of left-handedness ($r = .26$), albeit not very strongly. So if anything, the pattern at the country level suggests that left-handedness of countries is related to a larger discrepancy between the trust score in a 0-10 format and in other formats. Whether this relation holds up in a more adequate test remains to be seen. We plan to estimate multi-level models including country and individual characteristics to conduct further tests of speculations about correlates of the 0-10 gap.

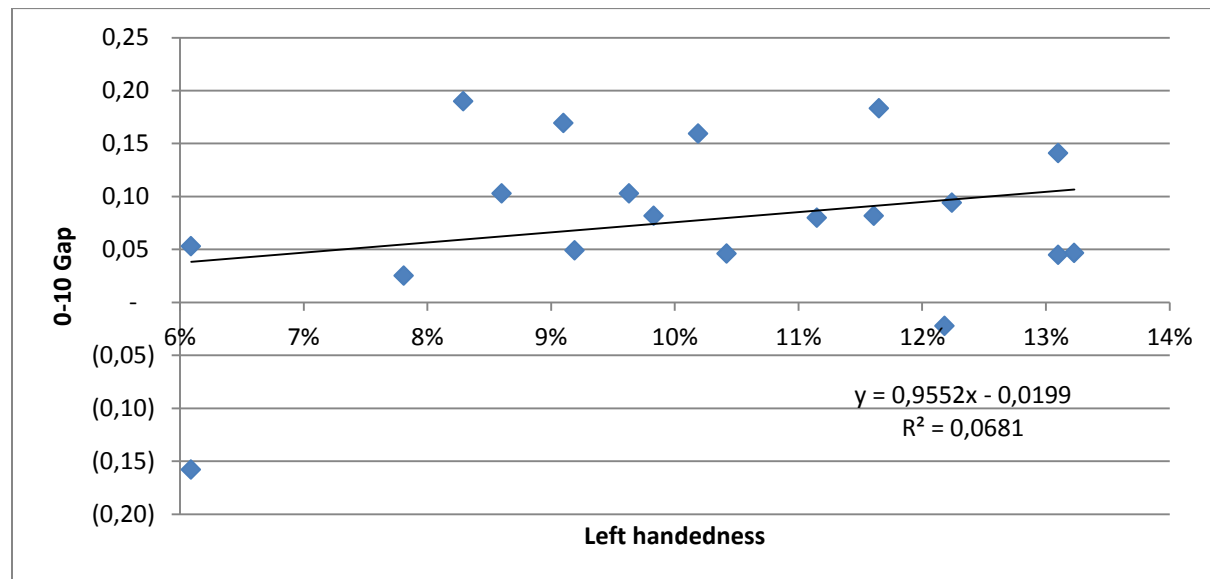
We can rule out the interpretation that the 0-10 scale generates higher scores because a midpoint is offered. Previous studies (Velez and Ashworth (2007); O’Muircheartaigh (1999)) have found that scales

offering a midpoint yield higher scores than scales that do not offer a midpoint. In our analyses, however, the 0-10 gap with the 1-4 scale is smaller than with the 1-5 scale, despite the fact that the 1-4 scale does not offer a midpoint while the 1-5 scale does.

Echoing but not citing Converse and Presser (1986), Abascal and Díaz de Rada (2014) suggest that the midpoint is the default choice for participants who ‘are reluctant to think about the issue at hand or express their opinion of it in front of the interviewer’. These suggestions could be tested in various ways. The suggestion on cognitive effort leads to the prediction that the popularity of the midpoint decreases with the level of education. Following Krosnick (1999), it can be hypothesized that respondents exerting less cognitive effort are more likely to satisfice, assuming that higher educated respondents display more cognitive effort in survey participation. Marsh (2013) provides evidence that midpoint responses decrease with cognitive effort. The suggestion on social desirability leads to the prediction that surveys administered in the presence of others contain more midpoint responses.

We conclude with a call for the extension of this research to other domains: generalized social trust is not the only concept that is measured on a 0-10 scale in social surveys. Satisfaction with life and happiness, and social attitudes such as religiosity, immigration, and left-right political self-placement are also measured on a 0-10 scale in the ESS, while other surveys have used other formats.

Figure 7. Differences between trust scores on a 0-10 scale and prevalence of left-handedness (n = 19)



References

- Abascal, E. & Díaz de Rada, V. (2014). Analysis of 0 to 10-point response scales using factorial methods: a new perspective. *International Journal of Social Research Methodology*, 17(5): 569-584. DOI: 10.1080/13645579.2013.799736
- Clarke, M. J., & Stewart, L. A. (1997). Meta-analyses using individual patient data. *Journal of Evaluation in Clinical Practice*, 3(3): 207-212.
- Converse, J.M., & Presser, S. (1986). *Survey Questions: Handcrafting the Standardized Questionnaire*. Thousand Oaks: Sage.
- Curran, P.J. & Hussong, A.M. (2009). Integrative Data Analysis: The Simultaneous Analysis of Multiple Data Sets. *Psychological Methods*, 14: 81-100.
- Darbyshire, P. & McDonald, H. (2004). Choosing Response Scale Labels and Length: Guidance for Researchers and Clients. *Australasian Journal of Market Research*, 12(2): 17-26.
- Dawes, J. (2002). Five point vs. eleven point scales: Does it make a difference to data characteristics? *Australian Journal of Market Research*, 10(1): 39-47.
- Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. *International Journal of Market Research*, 50(1): 61-77.
- Garratt, A.M., Helgeland, J. & Gulbrandsen, P. (2011). Five-point scales outperform 10-point scales in a randomized comparison of item scaling for the Patient Experiences Questionnaire. *Journal of Clinical Epidemiology*, 64: 200-207.
- Global Trust Research Consortium (2018). Mega-analysis of Generalized Social Trust. DOI 10.17605/OSF.IO/QFV76. <https://osf.io/qfv76/>
- Krosnick, J.A. (1999). Survey research. *Annual Review of Psychology*, 50 (1): 537-67.
- Leung, S-O. (2011) A Comparison of Psychometric Properties and Normality in 4-, 5-, 6-, and 11-Point Likert Scales. *Journal of Social Service Research*, 37 (4): 412-421. DOI: 10.1080/01488376.2011.580697
- Marsh, K.R. (2013). The Effects of Item and Respondent Characteristics on Midpoint Response Option Endorsement: A Mixed-Methods Study. Dissertation, James Madison University. <http://commons.lib.jmu.edu/cgi/viewcontent.cgi?article=1092&context=diss201019>
- McManus, I. (2009). The history and geography of human handedness. Pp. 37-58 in I. Sommer & R. Kahn (Eds.), *Language Lateralization and Psychosis*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511576744.004. http://www.ucl.ac.uk/medical-education/publications/reprints2009/2009-History_GeographyOfHumanHandedness.pdf
- O'Muircheartaigh, C., Krosnick, J. A., & Helic, A. (2000). Middle alternatives, acquiescence, and the quality of questionnaire data. Paper presented at the Annual Meeting of the American Association of Public Opinion Research, Fort Lauderdale, Florida.

Preston, C.C. & Colman, A. (2000) Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104: 1–15.

Riquier, C., Kennedy, R., and Sharp, B. (1996). Practical applications of Correspondence Analysis to categorical data in market research. *Journal of Targeting, Measurement and Analysis for Marketing*, 5 (1): 56-70.

Van der Veld, W.M., & Saris, W.E. (2011). Causes of generalized social trust. Pp. 207-247 in: Davidov, E.; Schmidt, P.; Billiet, J. (Eds.). *Cross-cultural analysis: Methods and applications*. Routledge.

Velez, P., & Ashworth, S.D. (2007). The Impact of Item Readability on the Endorsement of the Midpoint Response in Surveys. *Survey Research Methods*, 1 (2): 69-74.

Supplementary Information

Datasets included in the analyses

ESS 2002 – 2014. European Social Survey Cumulative File, ESS 1-7 (2016). Data file edition 1.0. NSD - Norwegian Centre for Research Data, Norway - Data Archive and distributor of ESS data for ESS REIC.

ISSP 2004. ISSP Research Group (2012): International Social Survey Programme: Citizenship - ISSP 2004. GESIS Data Archive, Cologne. ZA3950 Data file Version 1.3.0, doi:10.4232/1.11372

ISSP 2006. ISSP Research Group (2008): International Social Survey Programme: Role of Government IV - ISSP 2006. GESIS Data Archive, Cologne. ZA4700 Data file Version 1.0.0, doi:10.4232/1.4700

ISSP 2008. ISSP Research Group (2012): International Social Survey Programme: Religion III - ISSP 2008. GESIS Data Archive, Cologne. ZA4950 Data file Version 2.2.0, doi:10.4232/1.11334

ISSP 2010. ISSP Research Group (2012): International Social Survey Programme: Environment III - ISSP 2010. GESIS Data Archive, Cologne. ZA5500 Data file Version 2.0.0, doi:10.4232/1.11418

CID. Howard, Marc M., James L. Gibson, and Dietlind Stolle. United States Citizenship, Involvement, Democracy (CID) Survey, 2006. ICPSR04607-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2016-10-11. <https://doi.org/10.3886/ICPSR04607.v2>

GGSS. GESIS - Leibniz Institute for the Social Sciences (2016): German General Social Survey (ALLBUS) - Cumulation 1980-2014. GESIS Data Archive, Cologne. ZA4584 Data file Version 1.0.0, doi:10.4232/1.12574

Understanding Society. University of Essex. Institute for Social and Economic Research, NatCen Social Research, Kantar Public. (2017). Understanding Society: Waves 1-7, 2009-2016 and Harmonised BHPS: Waves 1-18, 1991-2009. [data collection]. 9th Edition. UK Data Service. SN: 6614

Citizenship Survey 2007/2008. Department for Communities and Local Government. Race, Cohesion and Faith Research Unit, National Centre for Social Research. (2008). Citizenship Survey, 2007-2008. [data collection]. 6th Edition. UK Data Service. SN: 5739, <http://doi.org/10.5255/UKDA-SN-5739-1>

Citizenship Survey 2008/2009. Department for Communities and Local Government, National Centre for Social Research. (2010). Citizenship Survey, 2008-2009. [data collection]. UK Data Service. SN: 6388, <http://doi.org/10.5255/UKDA-SN-6388-1>

Citizenship Survey 2009/2010. Department for Communities and Local Government, Ipsos MORI. (2011). Citizenship Survey, 2009-2010. [data collection]. UK Data Service. SN: 6733, <http://doi.org/10.5255/UKDA-SN-6733-1>

Citizenship Survey 2010/2011. Department for Communities and Local Government, Ipsos MORI. (2012). Citizenship Survey, 2010-2011. [data collection]. UK Data Service. SN: 7111, <http://doi.org/10.5255/UKDA-SN-7111-1>

1970 British Cohort Study. Bynner, J.M. (2016). 1970 British Cohort Study: Twenty-One-Year Sample Survey, 1992. [data collection]. 3rd Edition. UK Data Service. SN: 4715, <http://doi.org/10.5255/UKDA-SN-4715-2>

General Household Survey. Office for National Statistics. Social and Vital Statistics Division. (2006). General Household Survey, 2004-2005. [data collection]. UK Data Service. SN: 5346, <http://doi.org/10.5255/UKDA-SN-5346-1>

British Social Attitudes Survey 2002. National Centre for Social Research. (2004). British Social Attitudes Survey, 2002. [data collection]. UK Data Service. SN: 4838, <http://doi.org/10.5255/UKDA-SN-4838-1>

British Social Attitudes Survey 2004. National Centre for Social Research. (2006). British Social Attitudes Survey, 2004. [data collection]. UK Data Service. SN: 5329, <http://doi.org/10.5255/UKDA-SN-5329-1>

British Social Attitudes Survey 2006. National Centre for Social Research. (2008). British Social Attitudes Survey, 2006. [data collection]. UK Data Service. SN: 5823, <http://doi.org/10.5255/UKDA-SN-5823-1>

British Social Attitudes Survey 2008. National Centre for Social Research. (2010). British Social Attitudes Survey, 2008. [data collection]. UK Data Service. SN: 6390, <http://doi.org/10.5255/UKDA-SN-6390-1>

British Social Attitudes Survey 2010. National Centre for Social Research. (2012). British Social Attitudes Survey, 2010. [data collection]. UK Data Service. SN: 6969, <http://doi.org/10.5255/UKDA-SN-6969-1>

British Social Attitudes Survey 2012. NatCen Social Research. (2014). British Social Attitudes Survey, 2012. [data collection]. UK Data Service. SN: 7476, <http://doi.org/10.5255/UKDA-SN-7476-1>

British Social Attitudes Survey 2014. NatCen Social Research. (2016). British Social Attitudes Survey, 2014. [data collection]. 2nd Edition. UK Data Service. SN: 7809, <http://doi.org/10.5255/UKDA-SN-7809-2>

Community Life Survey. Cabinet Office. (2013). Community Life Survey, 2012-2013. [data collection]. UK Data Service. SN: 7405, <http://doi.org/10.5255/UKDA-SN-7405-1>

NCDS. University of London. Institute of Education. Centre for Longitudinal Studies. (2012). National Child Development Study: Sweep 8, 2008-2009. [data collection]. 3rd Edition. UK Data Service. SN: 6137, <http://doi.org/10.5255/UKDA-SN-6137-2>

ONS Omnibus Survey. Office for National Statistics. Social and Vital Statistics Division. (2007). ONS Omnibus Survey, Public Confidence in Official Statistics Module, July 2004 and March 2005. [data collection]. UK Data Service. SN: 5669, <http://doi.org/10.5255/UKDA-SN-5669-1>

Taking part 2006. Department for Culture, Media and Sport. (2010). *Taking Part: the National Survey of Culture, Leisure and Sport, 2005-2006; Adult and Child Data*. [data collection]. 2nd Edition. UK Data Service. SN: 5717, <http://doi.org/10.5255/UKDA-SN-5717-1>

Department for Culture, Media and Sport. (2010). *Taking Part: the National Survey of Culture, Leisure and Sport, 2006-2007; Adult and Child Data*. [data collection]. 2nd Edition. UK Data Service. SN: 6272, <http://doi.org/10.5255/UKDA-SN-6272-1>

Taking part 2008. Department for Culture, Media and Sport. (2010). *Taking Part: the National Survey of Culture, Leisure and Sport, 2007-2008; Adult and Child Data*. [data collection]. 2nd Edition. UK Data Service. SN: 6273, <http://doi.org/10.5255/UKDA-SN-6273-1>

Department for Culture, Media and Sport. (2010). *Taking Part: the National Survey of Culture, Leisure and Sport, 2008-2009; Adult and Child Data*. [data collection]. UK Data Service. SN: 6530, <http://doi.org/10.5255/UKDA-SN-6530-1>

Taking part 2010. Department for Culture, Media and Sport. (2010). *Taking Part: the National Survey of Culture, Leisure and Sport, 2009-2010; Adult and Child Data*. [data collection]. UK Data Service. SN: 6579, <http://doi.org/10.5255/UKDA-SN-6579-1>

Department for Culture, Media and Sport. (2011). *Taking Part: the National Survey of Culture, Leisure and Sport, 2010-2011; Adult and Child Data*. [data collection]. UK Data Service. SN: 6855, <http://doi.org/10.5255/UKDA-SN-6855-1>

Taking part 2012. Department for Culture, Media and Sport. (2013). *Taking Part: the National Survey of Culture, Leisure and Sport, 2011-2012; Adult and Child Data*. [data collection]. UK Data Service. SN: 7222, <http://doi.org/10.5255/UKDA-SN-7222-1>

Department for Culture, Media and Sport. (2013). *Taking Part: the National Survey of Culture, Leisure and Sport, 2012-2013; Adult and Child Data*. [data collection]. UK Data Service. SN: 7371, <http://doi.org/10.5255/UKDA-SN-7371-1>

Health Survey for England. Higgins, V., Marshall, A. (2012). Health Survey for England Time Series Dataset, 1991-2009. [data collection]. UK Data Service. SN: 7025, <http://doi.org/10.5255/UKDA-SN-7025-1>

NELLS. Tolsma, Dr. J. (Radboud University Nijmegen); Kraaykamp, Prof. dr. G.L.M. (Radboud University Nijmegen); Graaf, Prof. dr. P.M. de (Universiteit Tilburg); Kalmijn, Prof. dr. M. (Universiteit van Amsterdam); Monden, dr. C.M. (Oxford University) (2014): Netherlands Longitudinal Lifecourse Study - NELLS Panel Wave 1 2009 and Wave 2 2013 - versie 1.1. DANS. <https://doi.org/10.17026/dans-25n-2xjv>

LISS. Das, Prof.dr. J.W.M. (CentERdata - Institute for data collection and research - Tilburg University) (2007): Thematic collection: LISS panel data - overview of available LISS Studies. DANS. <https://doi.org/10.17026/dans-z8r-ruvr>

Scherpenzeel, A.C., and Das, M. (2010). "True" Longitudinal and Probability-Based Internet Panels: Evidence From the Netherlands. In Das, M., P. Ester, and L. Kaczmirek (Eds.), *Social and Behavioral*

Research and the Internet: Advances in Applied Methods and Research Strategies. (pp. 77-104). Boca Raton: Taylor & Francis.

Netherlands Election Survey (NKO). Aarts (Project leader), Prof.Dr. C.W.A.M. (University of Twente); Todosijevic, Dr. B. (University of Twente) (2009): Nationaal Kiezersonderzoek, NKO 1971-2006 cumulatieve file. DANS. <https://doi.org/10.17026/dans-zcz-csbx>

Culturele Veranderingen 2002. Becker, J.W. (Sociaal en Cultureel Planbureau, SCP (2002): Culturele Veranderingen in Nederland 2002 - CV'02. DANS. <https://doi.org/10.17026/dans-xu4-6sch>

Culturele Veranderingen 2004. Sociaal en Cultureel Planbureau - SCP (2004): Culturele Veranderingen in Nederland 2004 - CV'04. DANS. <https://doi.org/10.17026/dans-xtu-d36b>

Culturele Veranderingen 2006. Sociaal en Cultureel Planbureau - SCP (2007). Den Haag (primary investigator); (2007): Culturele Veranderingen in Nederland 2006 - CV'06. DANS. <https://doi.org/10.17026/dans-zh2-8kh6>

Culturele Veranderingen 2008. Sociaal en Cultureel Planbureau - SCP; Centraal Bureau voor de Statistiek - CBS; (2008): Culturele Veranderingen in Nederland 2008 - CV'08. DANS. <https://doi.org/10.17026/dans-z2s-f36j>

Swedish National Election Study. Holmberg, S., Oscarsson, H. (2004). Statistics Sweden / University of Gothenburg, Department of Political Science. Swedish election study 2002. Swedish National Data Service. Version 1.0. <https://doi.org/10.5878/002643>

Latinobarometer. Corporación Latinobarómetro (2017). Informe 2017. [online] Santiago de Chile. Available at: <http://www.latinobarometro.org>

Table A. Recoding

We use the formula of Preston and Colman (2000): $(\text{rating} - 1)/(\text{number of response categories} - 1)$.

Forced choice		3 categories		4 categories		5 categories		11 categories	
Original	Re	Original	Re	Or	Re	Or	Re	Or	Re
1. Cannot be too careful	0	1. Cannot be too careful	0	1. People can almost always be trusted	0	1. Disagree completely	0	0. Cannot be too careful	0
2. Most people can be trusted	1	2. It depends	0,5	2. People can usually be trusted	0,33	2. Disagree	0,25	1	0.1
		3. Most people can be trusted	1	3. You usually can't be too careful in dealing with people	0,67	3. Neither disagree, nor agree	0,5	2	0.2
				4. You almost always can't be too careful in dealing with people	1	4. Agree	0,75	3	0.3
						5. Agree completely	1	4	0.4
								5	0.5
								6	0.6
								7	0.7
								8	0.8
								9	0.9
								10	1.0

Table B. Year trends and country level variance by response format

	ESS 0-10	Other 0-10	1-3	1-4	1-5	0-1
Year	.001 (.000)	-.002 (.001)	.005 (.001)	-.002 (.001)	.023 (.001)	.008 (.000)
Constant	.489 (.001)	.604 (.004)	.405 (.002)	.457 (.003)	.238 (.004)	.345 (.002)
Country level variance	.140	.172	.001	.149	.097	.160
Number of respondents	330,526	31,501	49,659	55,552	39,521	141,692
Number of countries	31	3	2	24	18	4
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Figure A. Card used in the ESS to measure generalized trust

