



Research & Methods
ISSN 1234-9224 Vol. 25 (1, 2016): 3–16
The Ohio State University
Columbus, Ohio, USA
Institute of Philosophy and Sociology
Polish Academy of Sciences, Warsaw, Poland
www.askresearchandmethods.org

10 points versus 11 points? Effects of Left-right Scale Design in a Cross-national Perspective

Cornelia Zuell and Evi Scholz

GESIS – Leibniz Institute for the Social Sciences

Left-right self-placement on a unidimensional scale is a standard question in many social and political surveys to measure respondents' ideological orientation in a minimalist way. Although the left-right scale is a standard question, the scale design is not standardized across surveys. One aspect of scale design is the offer of a midpoint. This paper is about design effects on central left-right scale placement in a cross-national context. How do respondents answer if there is no true midpoint: Do respondents who want to express a middle position, in the case of a 10-point scale, use scale middle categories as a substitute for a true midpoint? Are findings consistent across countries? Offering a midpoint is much debated among researchers and quite often, a midpoint might serve as a hidden “don't know” or a missing attitude. Does nonresponse increase when non-attitudes cannot be expressed by choosing the neutral midpoint to hide nonresponse? If middle categories in the 10-point scale work as substitute for a true midpoint in the 11-point scale, nonresponse will not differ. We tested these questions in a split-half experiment where either a 10-point or an 11-point scale was asked in an experimental web survey fielded in six countries. Our results seem to confirm the idea that respondents who favor choosing a scale middle find a virtual center in the 10-point scale. However, results are inconsistent in cross-national perspective.

Keywords: cross-cultural comparison, scale midpoint, left-right self-placement, open-ended questions, response behavior

1. INTRODUCTION

Left-right self-placement on a unidimensional scale is one of the standard questions in many social and political surveys to measure respondents' ideological orientation. However, there is no general standard design for the scale. Usually, there are labeled endpoints "left" and "right" and a number of categories in between for the degree of left-right placement, but design differs due to the context of the question, the tradition of asking the scale, or methodological aspects, such as offering a midpoint.

Offering a midpoint is much debated among political and social sciences survey researchers. Quite often a midpoint on the left-right scale serves as a hidden "don't know" (DK) or a missing attitude (Inglehart and Klingemann 1976, Potter 2001:11–12). To avoid problems in interpretation of a midpoint which might mix missing attitudes and substantive attitudes, there are studies where the scale is designed with an even number of scale points and, accordingly, without a true midpoint. On the other hand, there is research that recommends using a midpoint and a scale with an odd number of scale points. While this research is not focused on the left-right scale but on Likert and rating scales, we assume that these findings also apply for the ideological self-placement scale. O'Muircheartaigh, Krosnick and Helic (2000), for example, examining agreement rating scales and using Eurobarometer data, found that midpoints do not compromise data quality. They also found that omission of midpoints did not lead to an increasing use of the categories neighboring the virtual midpoint. It seems that respondents, who could not use a (not offered) midpoint, were disposed to decide for both moderate and extreme categories and do not concentrate on middle categories. In addition, their findings suggest that a midpoint or middle categories and "don't know" do not cover the same underlying concept of a lacking opinion which speaks against the idea of a midpoint working as hidden don't know.

Worcester and Burns (1975) found that a 4-point scale without midpoint pushes more respondents towards the positive end of the scale. Weijters, Cabooter and Schillewaert (2010) findings on the effects of scale design go into the same direction which leads to the assumption that scales with an even number of scale point might produce a positivity bias. O'Muircheartaigh et al. (2000) or Garland (1991) in contrast saw a tendency towards disagreement which might lead to a negativity bias.

Tourangeau, Rips and Rasinski (2000) found that strength and stability of attitudes might work as moderator of design effects. For the left-right scale under consideration Federico (2009) reported that political interest serves as key indicator of personal involvement in politics and moderates self-placement. Whatever technical considerations on scale design are, there are respondents with

a real, politically meaningful middle or neutral ideological attitude. How do they use a scale without a scale point which represents their middle or neutral position? Do they substitute the non-existing true midpoint by a virtual midpoint and tick one of the neighboring categories? Do respondents who genuinely belong in the middle of the scale make essentially random choices among the alternatives or do they stay next to the midpoint and choose more moderate than extreme categories?

For cross-national analyses, it is crucial that the interpretation of cross-national differences results from substance and not from response effects (Smith 2003). Cross-national differences in response styles or systematic response behavior due to design result in problems concerning comparative interpretation of the data (Yang, Harkness, Chin et al. 2010). For example, Rodon (2014), in his study on central placement on the left-right scale, observed significant differences between ESS¹ countries. He found two main reasons for a central placement: respondent's party preference and a hidden nonresponse of politically uninterested respondents with considerable variation across individual countries. Differences in design effects that are only found in some countries but not in others produce problems in cross-national usability of scales and interpretation of data. Respondents from different countries might differ in their reaction on scale types where some types might produce a push into the direction of a social acceptable answer. What is the socially acceptable answer and what may be positively evaluated in one country and might be negative in another country. The connotation of "left" and "right" can be positive, neutral or negative. If the evaluation of left and right is the same in cross-national context, then there are no interpretation problems. If there are differences in evaluation, then there should be a push into different directions.

In this paper we investigate how respondents use a 10-point left-right self-placement scale (without midpoint) and an 11-point scale focusing on the use of the midpoint in cross-national perspective. National findings from the German SOEP² (Kroh 2007) suggest that both scales work well in technical perspective but that the 11-point scale works somewhat better in terms of substantive analyses of vote choice.

If both scales work in an equivalent way, then we expect that respondents find a virtual midpoint in the scale with an even number of categories and use the categories neighboring that virtual midpoint. Garland (1991) has run a split-half experiment where both a 4-point and a 5-point scale on importance of product labeling were alternatively asked to students in New Zealand. He checked the differences in relative frequencies between the odd and the even numbered scale points recalculating the 5-point scale distributions by filtering only for non-midpoint respondents to simulate a 4-point scale without midpoint. A test on significance showed that recalculated frequencies of the original 5-point scale were not similar to the respective frequencies of the 4-point scale and that the

differences were significant. He concluded that the use of a midpoint lead to a distortion of the distributions. The scale with an even number of scale points forces respondents to make a choice and leads to different frequencies of the two more middle categories. We will test this finding for the 11-point and 10-point left-right scale and check whether results are consistent in cross-national context.

H1: In a 10-point scale without a true midpoint, respondents find a virtual midpoint and use the middle categories 5 or 6 as a substitute.

If a 10-point scale without true midpoint and an 11-point scale with true midpoint do not differ in performance across countries, we do not expect differences in the amount of “don’t knowers” in between the two scales in any country in the experiment. Even when the midpoint of the 11-point scale works as a hidden “don’t know” non-attitude might also be hidden in the 10-point scale and its middle categories 5 and 6.

H2: If the middle categories in the 10-point scale work as a substitute for a true midpoint in the 11-point scale, there will be no significant difference in nonresponse between a 10-point and 11-point scale.

Following the idea of social desirability pushing respondents who are offered scales with an even number of scale points into the direction of scale endpoints, we expect a more skewed distribution for the 10-point scale than for the 11-point scale.

H3: A 10-point left-right scale will be more skewed than an 11-point scale.

Pursuing the idea of social acceptable answers and a push, there still is the question whether findings are similar across countries. “Right” in many languages does not only refer to a direction but has also got a positive connotation of “correct”, “just”, “honest”, or “law” (in German: Recht; in French: droit; in Spanish: derecha) while “left” often refers to “inept”, “gauche”, or even “cheat” and has got a more negative meaning. However, in terms of political and historical context, countries might differ (Bobio 1996:40-41). For example, “right” in Germany has a negative connotation and is sometimes linked to right-extreme. The coding scheme developed by Zuell and Scholz (2012) documents a variety of associations with the terms “left” and “right” in cross-national perspective.

So, in the case of the 10-point scale, is there a push into the same direction or do countries differ according to differences in social acceptability?

H4: In countries where “right” is socially desirable and assessed positively, the push will be to the right of the scale and the skewness will be negative. In countries where “left” is socially desirable and assessed positively, the push will be to the left of the scale and the skewness will be positive.

If design effects are moderated by strength of political involvement we expect that respondents with strong political interest behave more robust against changes in scale design than respondents with a low interest.

H5: If political interest moderates design effects, for politically uninterested respondents design effects are stronger than for politically interested respondents.

2. DATA, SPLIT-HALF EXPERIMENT

To test our hypotheses, we use data from the 2011 CICOM³ project, which ran web surveys in six countries: three Western European (Denmark, Germany, and Spain); two North American (Canada and the U.S.); and Hungary as a former Communist country. Respondents were drawn from non-probability access panels according to pre-set quotas based on gender, age, and education to obtain balanced samples of nationals aged 18-65. Sample sizes are approximately 500 at a minimum⁴.

Respondents were asked to indicate their ideological position by placing themselves on a horizontal left-right scale with the labeled endpoints “left” (at the left hand side) and “right” (at the right hand side). A probe question followed on a separate screen, asking respondents to specify what they mean by the terms “left” and “right”. For both questions, answers were not forced and respondents were reminded by soft checks which allows for nonresponse in general. The survey’s visual design was identical for all countries in the survey. To test effects of the left-right scale design, a split-half experiment was used: Half of the respondents received the 10-point scale and the other half of the respondents received the 11-point scale. The experimental questions on the left-right scale were located in the middle of the interview, between a question on satisfaction with democracy and items on immigrants.

A tendency to a positive, a neutral or a negative meaning of “left” and “right” in the countries in the experiment was extracted from the probe questions on left and right which indicates the respondents’ individual assessment. These answers were coded using a coding scheme described by Zuell and Scholz (2012). We have analyzed the answers to the open-ended questions and checked for the most frequent associations. We found cross-national differences in what is associated with left and right. While there are associations where the meaning is either neutral or can be seen in a positive or a negative way, we could find clear negative associations what is linked to “right” in Germany referring to extremism, radical ideas or xenophobia (national socialism by 16%, xenophobia by 11%, right wing radicalism by 6%, radical by 6%). For all other countries in the experiment, the ideas mentioned on the meaning of “right” do not have such a negative connotation. In contrast, in Denmark, Hungary and in the U.S., we could find a slightly more positive meaning referring to individualism in Denmark (mentioned by 8%), to

patriotism in Hungary (mentioned by 13%), and to freedom in the U.S. (mentioned by 8%). In Spain and Canada, the most frequent answers with respect to “right” do not have either a positive or a negative meaning⁵. “Left” in all countries in the experiment except the U.S., is linked to solidarity⁶ or equality⁷ and thus has a more positive meaning. In the U.S., in contrast, “left” has a more negative connotation, being associated with regulation (which, in the given coding scheme, was defined as the opposite category to freedom and is mentioned by 7%). If the idea of social acceptable answers is working, then we expect a push into the left direction in all countries (the most in Germany, the least in Canada), but a push into the right direction in the U.S. To test for a push, we checked the skewness for both the 11-point and 10-point scale.

Our indicator which might moderate design effects is political interest, asked with five fully-labeled response categories (“very interested”, “interested”, “some interest”, “little interest” “not at all interested”). For analytical purposes we collapsed strongly interested and interested respondents and respondents who were not at all interested or report little interest in politics each into one analytical category in order to define two contrast groups and checked the frequency distribution for both groups individually. We did not consider respondents who reported to have some interest in politics, where we expected confounded effects.

3. RESULTS

If we compare the frequency distribution of the 11-point scale to the distribution of the 10-point scale to test whether relative frequencies of the middle categories are different, then there are inconsistent findings in a cross-national perspective with respect to significant differences (Table 1). We can find significant differences in all countries, namely for the categories in the neighborhood of the midpoint. However, both in Denmark and Germany there is also another category with significant differences between the 11-point scale and the corresponding 10-point scale category. For Denmark, there is a difference in the use of category 9 on a low level of significance; in Germany, there is a difference in the use of category 4, again on a low level of significance.

Findings by O’Muircheartaigh et al. (2000) that respondents who were not offered a true midpoint were disposed to decide for both moderate and extreme categories and do not concentrate on middle categories, are not replicated in our data either. We did not find a higher use of all categories in the 10-point scale than in the 11-point scale. Respondents, who were not offered a true midpoint, did not use all scale points but concentrate on the scale middle, and this is consistently in all countries in the experiment. Deviations from this finding which might compromise our hypothesis 1 are on a low level of significance only.

Table 1 Frequency distributions and differences between categories: 11-point and 10-point left-right scale (in %)

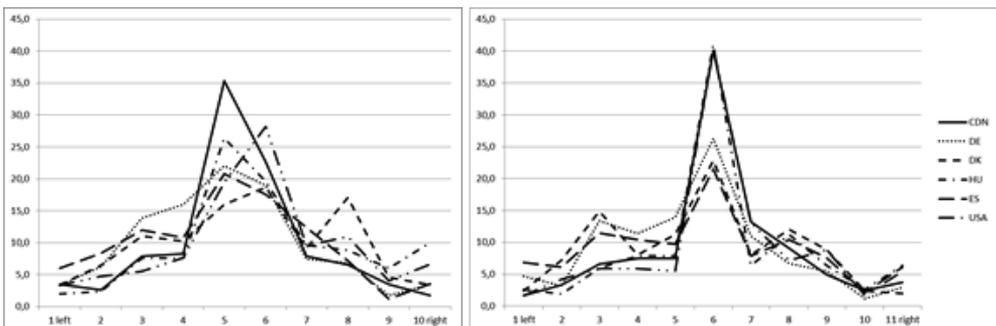
| | Canada | | Denmark | | Germany | | Hungary | | Spain | | USA | | | | | | | |
|-----------|----------|-------|----------|-------|----------|---------|----------|-------|----------|-------|----------|----------|------|------|----------|------|-------|----------|
| | 11-point | Diff. | 11-point | Diff. | 11-point | Diff. | 11-point | Diff. | 11-point | Diff. | 11-point | Diff. | | | | | | |
| Left 1 | 1.4 | -1.7 | 2.2 | 3.2 | -1.0 | 4.5 | 3.2 | 1.3 | 2.5 | 1.8 | 0.7 | 6.5 | 5.7 | 0.8 | 1.6 | 3.1 | -1.5 | |
| 2 | 2.9 | 2.4 | 0.5 | 6.0 | 0.6 | 3.0 | 6.0 | -3.0* | 1.8 | 2.2 | -0.4 | 5.8 | 8.0 | -2.2 | 3.9 | 4.4 | -0.5 | |
| 3 | 5.8 | 7.1 | -1.3 | 13.6 | 10.2 | 3.4 | 12.7 | 13.3 | -0.6 | 6.9 | -1.3 | 10.9 | 11.4 | -0.5 | 5.5 | 5.1 | 0.4 | |
| 4 | 6.5 | 7.5 | -1.0 | 7.3 | 9.5 | -2.2 | 10.8 | 15.3 | -4.5* | 7.4 | 6.9 | 0.5 | 9.9 | 10.2 | -0.3 | 5.5 | 7.2 | -1.7 |
| 5 | 6.5 | 31.8 | -25.3*** | 10.3 | 14.7 | -4.4 | 13.2 | 21.1 | -7.9*** | 7.4 | 24.1 | -16.7*** | 9.2 | 19.7 | -10.5*** | 5.1 | 18.1 | -13.0*** |
| 6 | 34.9 | NA | NA | 20.9 | NA | NA | 25.0 | NA | NA | 38.6 | NA | NA | 20.5 | NA | NA | 37.8 | NA | NA |
| 7 | 11.5 | 20.4 | -8.9* | 7.3 | 17.2 | -9.9*** | 10.4 | 18.1 | -7.7*** | 6.0 | 17.9 | -11.9*** | 7.5 | 16.7 | -9.2** | 12.2 | 26.3 | -14.1*** |
| 8 | 7.9 | 7.1 | 0.8 | 11.0 | 8.4 | 2.6 | 6.3 | 7.1 | -0.8 | 10.5 | 8.8 | 1.7 | 9.9 | 11.7 | -1.8 | 6.7 | 8.9 | -2.2 |
| 9 | 4.3 | 5.9 | -1.6 | 8.1 | 15.8 | -7.7* | 5.2 | 6.5 | -1.3 | 6.0 | 8.0 | -2.0 | 7.2 | 6.8 | 0.4 | 8.3 | 10.2 | -1.9 |
| 10 | 2.2 | 3.1 | -0.9 | 2.2 | 4.2 | -2.0 | 1.1 | 1.7 | -0.6 | 2.5 | 5.5 | -3.0 | 2.0 | 1.1 | 0.9 | 1.6 | 3.8 | -2.2 |
| Right 11 | 3.2 | 1.6 | 1.6 | 1.8 | 3.2 | -1.4 | 2.8 | 3.4 | -0.6 | 6.0 | 9.1 | -3.1 | 5.8 | 3.4 | 2.4 | 5.1 | 6.1 | -1.0 |
| DK | 12.9 | 10.2 | 2.7 | 8.8 | 7.7 | 1.1 | 4.9 | 4.3 | 0.6 | 6.0 | 8.8 | -2.8 | 4.8 | 5.3 | -0.5 | 6.7 | 6.8 | -0.1 |
| N | 278 | 255 | 23 | 273 | 285 | 12 | 536 | 535 | 1 | 285 | 274 | 11 | 293 | 264 | 29 | 254 | 293 | 0 |
| skewness | 0.05 | -0.00 | 0.05 | 0.07 | -0.10 | 0.03 | 0.18 | 0.33 | 0.15 | 0.13 | 0.17 | 0.14 | 0.14 | 0.08 | 0.02 | 0.02 | -0.09 | 0.02 |
| std error | 0.16 | 0.16 | 0.00 | 0.15 | 0.15 | 0.11 | 0.11 | 0.11 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.15 | 0.16 | 0.16 | 0.15 | 0.15 |

* p ≤ .05, ** p ≤ .005, *** p ≤ .001

Checking for differences in nonresponse, we found that nonresponse in both the 11-point and the 10-point scales is low in all countries and varies from 5% in Spain to 13% in Canada for the 11-point scale and between 4% in Germany and 10% in Canada for the 10-point scale. The 11-point and the 10-point scale do not differ significantly in nonresponse. Differences between scales only vary from .1% to 2.8%. Results in which scale type nonresponse is higher are not consistent across countries: In Canada, Germany and Denmark the 11-point scale works slightly worse than the 10-point scale in terms of higher nonresponse while nonresponse is a bit lower for the 11-point scale in Spain and Hungary. In the U.S., we can find similar nonresponse for both scales with about 7% (Table 1). The missing significant design effects with respect to nonresponse speak in favor of hypothesis 2. A logistic regression (not reported in the paper in tabular form) with nonresponse as dependent variable and scale type and countries as explanatory variables (all variables are coded as dummies) does not yield other results: whether the 10-point or the 11-point scale was asked does not influence the use of don't know whereas we consistently found an effect of country irrespective which country was taken as reference variable.

With respect to the skewness in the 11-point and 10-point scale, we found that skewness is consistently low (lower than 1.0) in all countries for both scales. However, results are not consistent across countries and the 10-point scale does not show higher values in skewness than the 11-point scale in general and most of the skewness is not significant: For the 10-point scale, skewness is slightly higher in Denmark, Germany, Hungary, and the U.S. In Canada the distributions are largely symmetric with figures between .0 to .05 for both scales. The skewness of the 11-point scale, however, is higher in Canada and Spain. So, in sum, distributions are not highly skewed, and it seems that a scale without midpoint does not push respondents towards the scale extreme categories in general. Thus, we cannot confirm hypothesis 3 (Table 1, Figure 1).

Figure 1 Skewness of the left-right self-placement scale (10-point vs. 11-point)



The idea of social desirable answers assumes a push into the social desirable direction if a scale lacks a true midpoint. Considering the answers from the open-ended questions, we expected a push into the right direction in the U.S., while for the other countries in the experiment a push into the left direction was assumed with the strongest effect in Germany. In fact, in the 10-point scale without midpoint we could find a negative (low) skewness in the U.S. and Denmark and positive values in Germany (.33), Spain (.08) and Hungary (.17). The German distribution for the 10-point scale is the only one with a higher skewness in relation to its standard error⁸. So, there is no general push into one direction, but in different directions due to different social acceptability across countries, except for Denmark where we expected a positive skewness based on a more positive evaluation of “left”. While the effect is not very large, hypothesis 4 is confirmed by our data with the Danish exception. Denmark might be exceptional due to unclearness of terms. “Right” in the Danish language “venstre” might be mixed up with a Danish liberal party named “Venstre” which can be seen in the answers to the open-ended questions where the party was associated with both “left” (4% of respondents) and “right” (10% of respondents).

Testing for effects of political interest as moderator of design effects, we expected that strongly politically interested respondents would show less differences between an 11-point scale with midpoint and a 10-point scale without midpoint than politically uninterested respondents do.

For respondents with a low or no political interest where strong effects from scale design and thus high differences in relative frequencies between scales were assumed, we found, consistently across countries, higher relative frequencies for the middle categories in the 10-point scale than in the 11-point scale and significant differences for one of the middle categories (Table 2). Deviations from a concentration on middle categories can be found for Germany and for Denmark: There are additional scale points with significant differences, however, on a low level of significance.

For respondents with a strong political interest we expected lower design effects than for politically uninterested respondents. For those who are more interested in politics there are highly significant differences in Canada, Spain, and Hungary and no significant differences in the U.S. and Denmark. Germany appears as the outlier with significant differences on a low level of significance on the scale points 2, 5 and 7 and no highly significant differences. In general, differences are lower for politically interested respondents than for the uninterested respondents (Table 2), except for Spain. While results confirm our hypothesis 5 on moderating effects of political interest in most of the countries in the experiment, it is not confirmed in general.

Table 2 Differences between categories of the two scale designs distinguished by respondents' political interest (in %)

| | Canada ^a | | Denmark ^b | | Germany ^c | | Hungary ^d | | Spain ^e | | USA ^f | |
|----------|---------------------|------------|----------------------|------------|----------------------|------------|----------------------|------------|--------------------|------------|------------------|------------|
| | Not interested | Interested | Not interested | Interested | Not interested | Interested | Not interested | Interested | Not interested | Interested | Not interested | Interested |
| Left 1 | -1.9 | -0.3 | 0.1 | -1.7 | -0.3 | 2.7 | 0.1 | 3.9 | 5.2 | 2.1 | 1.5 | -4.8 |
| 2 | -1.0 | 1.1 | 1.1 | -4.1 | 0.6 | -7.0* | 0.0 | -0.2 | 1.8 | -5.3 | -2.0 | -0.1 |
| 3 | 2.1 | 3.5 | 3.2 | 2.6 | 2.4 | -2.7 | -0.9 | 3.8 | -0.4 | -0.1 | -0.5 | 6.0 |
| 4 | -2.3 | 2.0 | -4.8 | -0.9 | -10.5* | 2.8 | 3.2 | 5.1 | -1.6 | 2.1 | -1.0 | -0.9 |
| 5 | -26.9*** | -23.0*** | -14.6** | 5.6 | -15.5** | -7.8* | -20.1*** | -13.6* | -12.1 | -16.3*** | -20.3*** | -8.8 |
| 6 | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| 7 | -14.3** | -5.0 | -16.6** | 1.7 | -5.7 | -6.7* | -17.1*** | -5.7 | -16.0* | -4.4 | -18.5** | -8.5 |
| 8 | 0.0 | 0.9 | 1.3 | -1.0 | 0.9 | -0.7 | 2.1 | -2.9 | 2.5 | -6.7 | -2.6 | -1.5 |
| 9 | -2.1 | -1.9 | -8.8* | -10.0 | -0.3 | -0.9 | -2.0 | -6.9 | -1.3 | 2.7 | -2.8 | -0.9 |
| 10 | -1.0 | -3.1 | 2.1 | -4.6 | -0.1 | -1.0 | -1.0 | -5.6 | -0.1 | 2.9 | 0.0 | -2.6 |
| Right 11 | 2.4 | -0.3 | 2.1 | -1.6 | -1.0 | -0.6 | -5.0 | -2.9 | 4.2 | 4.8 | -0.6 | -1.5 |
| DK | 1.3 | 2.5 | 0.8 | 0.0 | -3.0 | 1.2 | -4.9 | -1.5 | -4.0 | 1.9 | -2.7 | 1.4 |

a not interested: 11-point scale DK: 18.3%, N 126; 10-point scale DK: 17.0%, N 112; interested: 11-point scale DK: 5.3%, N 76; 10-point scale DK: 2.8%, N 71

b not interested: 11-point scale DK: 18.6%, N 97; 10-point scale DK: 17.8%, N 101; interested: 11-point scale DK: 0.0%, N 77; 10-point scale DK: 0.0%, N 69

c not interested: 11-point scale DK: 8.5%, N 142; 10-point scale DK: 11.5%, N 131; interested: 11-point scale DK: 2.2%, N 184; 10-point scale DK: 1.0%, N 204

d not interested: 11-point scale DK: 11.3%, N 97; 10-point scale DK: 16.2%, N 99; interested: 11-point scale DK: 2.6%, N 76; 10-point scale DK: 4.1%, N 74

e not interested: 11-point scale DK: 9.5%, N 95; 10-point scale DK: 13.5%, N 89; interested: 11-point scale DK: 3.0%, N 99; 10-point scale DK: 1.1%, N 88

f not interested: 11-point scale DK: 14.1%, N 85; 10-point scale DK: 16.8%, N 101; interested: 11-point scale DK: 2.2%, N 90; 10-point scale DK: 0.8%, N 118

* p ≤ .05, ** p ≤ .005, *** p ≤ .001

4. DISCUSSION

Attempting to find out whether the middle categories of the 10-point left-right scale without a true midpoint work as a substitute for the true midpoint of the 11-point left-right scale, we implemented a split-half experiment in which one half of the respondents received the 11-point scale and the other half received the 10-point scale. In contrast to O’Muirheartaigh et al. (2000), we found that the omission of a midpoint lead to increased use of the categories next to the virtual midpoint in all countries in the experiment, but the increase is not consistently high across all countries. In contrast to their results, we could not find a higher use of all categories in a scale with an even number of scale points. Respondents, who could not use a (not offered) midpoint, did not decide for both moderate and extreme categories but concentrate on middle categories. This is a consistent finding across all countries. Both the 11-point and the 10-point scale work well in terms of nonresponse. That means, omission or existence of a midpoint does not significantly influence nonresponse. Our analyses show that the hypothesis of a push into the direction of scale endpoints for the 10-point scale without midpoint does not work consistently across countries and thus cannot be confirmed in general. The highest though still low skewness is found for the 10-point scale in Germany. The idea that social desirability pushes responses in the 10-point scale into the country-specific social desirable direction seems to be confirmed by our data while the effect is not very large.

In summary, our analyses indicate that the left-right scale does not work similarly across countries and design effects are not always clear. These different results across countries may indicate that there are different understandings of the left-right scale in a cross-national perspective which makes it difficult to use the left-right scale in a comparative way (e.g., Finlay, Simon and Wilson II 1974). However, it seems that the 11-point scale works better than the 10-point scale, in particular given the somewhat skewed distribution in the German 10-point scale. Whether the findings on left-right scale design can be generalized for other scales is unclear. A second scale experiment on democracy satisfaction included in the CICOM survey where design options were modeled parallel to the left-right scale did not yield similar results: The shapes of both the 11-point and the 10-point scale frequency distributions are completely different and there is no concentration on scale middles or midpoint at all which, in contrast, is a typical distribution of the left-right scale. So, the satisfaction with democracy scale cannot be used to compare results. Bauer, Barbera, Ackermann et al. (2014) found that the interpersonal comparability of the left-right scale is reduced by the very different meanings attributed to the vague left-right concept and raises concerns about the applicability of the left-right scale in comparative research. Considering our results we assume

that not only the interpersonal comparability of left and right is impaired but also the cross-national comparability. Weber (2011) investigating effects of labels in an 11-point scale in a test-retest experiment using cross-national ESS Round 4 data similarly found general patterns and exceptional countries. Weber offered democratic experience and nonresponse as arguments for deviating results. Our assumption for the somewhat unclear results for our left-right scale experiment is that latent attitudes or attitudes on topics that might be irrelevant to the respondent are more open to design influences than strong or robust attitudes are. This idea complies with findings on politically uninterested respondents presented in this paper where differences between 11-point and 10-point left-right scale seem to be higher than for politically interested respondents. Most researchers have used cross-national data but did not examine cross-national differences in effects of left-right scale design. Weber (2011) has started to check for the issue of cross-national comparability with respect to scale labels, we have checked for effects of length of scale, both of us found exceptional countries challenging the results of our analyses and the general comparability of left-right data. Further research on left-right scale design in cross-national context is necessary.

NOTES

- 1 ESS European Social Survey, <http://www.europeansocialsurvey.org/>
- 2 SOEP The German Socio-Economic Panel, <http://www.diw.de/en/soep>
- 3 CICOM: Enhancing the Validity of Intercultural Comparative Surveys; project was funded by the German Research Foundation (DFG) [grant number BR 908/3-1].
- 4 Sample size: Canada 536; Denmark 559; Germany 1075; Hungary 559; Spain 559; USA 547
- 5 In Spain the most frequent categories on the meaning of right are “capitalism” (by 22%), “conservatism and tradition” (by 16%), “PP” (by 9%), “entrepreneurs and employees” (by 8%) and “religion” (by 6%). In Canada the most frequent categories on the meaning of right are “liberalism” (by 11%), “socialism” (by 8%), and “progressive” (5%).
- 6 Denmark: 12%; Germany: 14%; Hungary: 15%; Spain: 10%
- 7 Denmark: 8%; Germany: 8%; Hungary: 9%; Spain: 17%
- 8 The ratio between skewness and its standard error is higher than 2 which speaks for an asymmetric distribution.

REFERENCES

- Bauer, Paul C., Pablo Barbera, Kathrin Ackermann and Aaron Venetz. 2014. “Vague Concepts in Survey Questions: A General Problem Illustrated with the Left-Right Scale.” *SSRN eLibrary*. doi: <http://dx.doi.org/10.2139/ssrn.2467595>.
- Bobio, Norberto. 1996. *Left and Right. The Significance of a Political Distinction*. Cambridge: Polity Press.
- Federico, Christopher M. 2009. “How People Organize Their Political Attitudes: The Roles of Ideology, Expertise, and Evaluative Motivation.” *American Psychological Association, Psychological Science Agenda* 23(9).

- Finlay, David J., Douglas W. Simon and L. A. Wilson II. 1974. "The Concept of Left and Right in Cross-National Research." *Comparative Political Studies* 7(2): 209–21. doi: 10.1177/001041407400700204.
- Garland, Ron. 1991. "The Mid-Point on a Rating Scale: Is it Desirable?". *Marketing Bulletin* 2(1): 66–70.
- Inglehart, Ronald and Hans-Dieter Klingemann. 1976. "Party Identification, Ideological Preference and the Left-Right Dimension among Western Mass Publics." Pp. 243–73 in *Party Identification and Beyond*, edited by Ian Budge, Ivor Crewe and Denise Farlie. Essex: John Wiley & Sons.
- Kroh, Martin. 2007. "Measuring Left-Right Political Orientation: The Choice of Response Format." *Public Opinion Quarterly* 71(2): 204–20. doi: 10.1093/poq/nfm009.
- O’Muirheartaigh, Colm, Jon A. Krosnick and Armin Helic. 2000. "Middle Alternatives, Acquiescence, and the Quality of Questionnaire Data." *Working Papers*. 01/2000. Chicago: University of Chicago.
- Potter, Corrie. 2001. "Left-Right Self-Placement in Western Europe: What Responses and Non-Responses Indicate." Pp. 1–18. U.W. Madison: Paper presented for the Political Behaviour Group.
- Rodon, Toni. 2014. "Do All Roads Lead to the Centre? The Unresolved Dilemma of Centrist Self-Placement." *International Journal of Public Opinion Research*. doi: 10.1093/ipor/edu028.
- Smith, Tom W. 2003. "Developing Comparable Questions in Cross-National Surveys." Pp. 69–91 in *Cross-Cultural Survey Methods*, edited by Janet A. Harkness, Fons J. R. Van de Vijver and Peter Ph. Mohler. Hoboken: Wiley.
- Tourangeau, Roger, Lance J. Rips and Kenneth Rasinski. 2000. *The Psychology of Survey Response*. Cambridge, New York: Cambridge University Press.
- Weber, Wiebke. 2011. "Testing for Measurement Equivalence of Individuals’ Left-Right Orientation." *Survey Research Methods* 5(1): 1–10. doi: <http://dx.doi.org/10.18148/srm/2011.v5i1.4622>.
- Weijters, Bert, Elke Cabooter and Niels Schillewaert. 2010. "The Effect of Rating Scale Format on Response Styles: The Number of Response Categories and Response Category Labels." Vol. 2010/636. Ghent: University Ghent.
- Worcester, Robert M. and Timothy R. Burns. 1975. "A Statistical Examination of the Relative Precision of Verbal Scales." *Journal of the Market Research Society* 17(3): 181–97.
- Yang, Yongwei, Janet A. Harkness, Tzu-Yun Chin and Ana Villar. 2010. "Response Style and Culture." Pp. 203–23 in *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*, edited by Janet A. Harkness, Michael Braun, Brad Edwards, Timothy P. Johnson, Lars Lyberg, Peter Ph. Mohler, B.-E. Pennell and T. W. Smith. Hoboken: Wiley.
- Zuell, Cornelia and Evi Scholz. 2012. "Assoziationen mit den politischen Richtungs-begriffen „Links“ und „Rechts“ im internationalen Vergleich: Kategorienschema für die Codierung offener Angaben." Vol. 2012/03. *GESIS-Technical Report*. Köln: GESIS.

Cornelia Zuell's research focuses on methodological aspects of social science surveys. One of her areas of interest is textual and content analyses, both with respect to data collection and data analyses. Cornelia's research has recently been published in *Social Science Computer Review* and the *Bulletin of Sociological Methodology*.

Evi Scholz is responsible for the German ISSP, is member of the ISSP Methodology Committee and the ISSP Secretariat. She is mainly interested in cross-national survey design and in the measurement of background variables. Evi's research has recently been published in *Social Science Research* and *mda: Methoden, Daten, Analysen*.