

Using Web Probing to Elucidate Respondents' Understanding of 'Minorities' in Cross-Cultural Comparative Research¹

Michael Braun

GESIS – Leibniz-Institute for the Social Sciences

Dorothee Behr

GESIS – Leibniz-Institute for the Social Sciences

Katharina Meitinger

Utrecht University

Klara Raiber

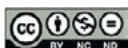
Radboud University

Lydia Repke

GESIS – Leibniz-Institute for the Social Sciences

Abstract: Due to the growing significance of international studies, the need for tools to assess the equivalence of items is pressing. Web probing, which is implementing verbal probing techniques traditionally used in cognitive interviewing in online surveys, is a method to complement quantitative techniques to establish equivalence of items in cross-cultural research. We illustrate this approach by assessing the question of 'how important it is that government authorities respect and protect the rights of minorities', which was originally used in the International Social Survey Program, for respondents in five countries (Germany, Britain, the U.S., Mexico, and Spain). First, participants answered this question using a 7-point Likert scale. Then they wrote freely what types of minorities they had thought of. Whether country differences in the response patterns can be interpreted substantially depends partially on how similarly the term 'minorities' is understood across these five contexts. Our results show that people in the participating countries have slightly different kinds of 'minorities' in mind.

Keywords: web probing; minorities; comparability; cross-cultural research



INTRODUCTION

Large-scale cross-national survey projects, such as the International Social Survey Program (ISSP), the European Social Survey (ESS), and the Eurobarometer, have continuously grown in importance over recent decades. A major goal of these studies is the equivalent measurement of constructs (e.g., behaviors, preferences, opinions, social attitudes) across time and space to allow meaningful cross-cultural comparisons. However, countries may differ in the way respondents interpret a question. In addition, changes over time in the understanding of questions within and across countries can threaten the validity of cross-national surveys. Hence, assessment of equivalence should precede any substantive analyses.

The traditional method to establish equivalence of measurement is applying data-analytic approaches (e.g., Vandenberg & Lance, 2000; Davidov et al., 2014). Most of these approaches can inform the decision on whether measurement equivalence is given or not (e.g., multi-group confirmatory factor analysis), but they cannot identify the substantive causes of non-equivalence. One exception is multilevel structural equation modeling, where variables on the country level might be used to model potential causes of non-equivalence (Davidov et al., 2012). Unfortunately, even this method is unable to sufficiently inform a redesign of existing items or the construction of new items. For this purpose, information on what respondents in different countries have in mind when answering survey questions is crucial.

An alternative or complement to statistical procedures is the use of cognitive interviewing for assessing equivalence of measurement (e.g., Miller et al., 2011; Thrasher et al., 2011; for an overview see Willis, 2015). With the help of cognitive interviewing techniques, such as think-aloud or probing, the processes respondents use in answering survey items, as well as differences in item interpretation, can be elucidated (Beatty & Willis, 2007; Lee, 2014; Willis, 2005). Examples of probing questions are comprehension probing (*What does the term 'xxx' mean to you?*), category-selection probing (*Why did you choose 'agree'?*), and specific probing (*What type of 'xxx' did you think of when you answered the previous item?*) (e.g., Prüfer & Rexroth, 2005; Willis, 2005).

Traditionally, cognitive interviewing is based on small quota samples (often 5-15 interviews; Willis, 2005). As a consequence, this method may be sufficient to detect some of the major problems with items (Beatty & Willis, 2007) but the prevalence of problems cannot be quantified. Thus, the typically low case numbers (e.g., a maximum of 20 cases per country in Fitzgerald, Widdop, Gray, & Collins, 2011) prevent generalizable conclusions on the differences between country-specific answer patterns.

In order to circumvent these problems, we propose the use of probing techniques from cognitive interviewing in web surveys (Behr, Meitinger, Braun, & Kaczmirek, 2017) – this method, called web probing, is a means to complement quantitative techniques to establish measurement equivalence in cross-cultural research (Behr et al., 2017; Meitinger, 2017). Web probing can easily be implemented before, alongside or after a survey. It proceeds without interviewers and, thus, solves both the issues of interviewer availability and standardization of procedures. Further, it enables the collection of a larger number of cases and thereby allows for analyzing differences between answer patterns across countries. However, there are also disadvantages of web probing, which are notably connected to the absence of an interviewer and the composition of samples (Behr et al., 2017; Meitinger & Behr, 2016). First, probe nonresponse might be higher due to the lack of the potential motivating function of an interviewer. Second, mismatching² or insufficient probe answers might occur that cannot be corrected by follow-up probes. In web probing, potential problems have to be foreseen as much as possible and countered by the formulation of probing questions. Third, web probing requires a population that is active and can be reached online. This may not apply to all target populations for a survey.

Cross-national web probing is in line with calls for mixing qualitative and quantitative methods in cross-national research (Benítez & Padilla, 2014; van de Vijver & Chasiotis, 2010). In fact, already in 1966, Schuman had proposed the method of random probing in cross-cultural surveys to learn more about cross-cultural answers. Finally, in contrast to quantitative approaches that usually presuppose multiple-item measures for a given construct, web probing – similar to cognitive interviewing – can assess the cross-national comparability of single questions or items.

We exemplify the approach of assessing a single item by analyzing one question from the cross-national ISSP Citizenship II module fielded in 2014 (ISSP Research Group, 2016). In this survey, an item battery focused on citizens' rights in a democracy and one item from this battery dealt with the importance 'that government authorities respect and protect the rights of minorities.' Our aim is to find out whether survey questions that refer to 'minorities' in general work in cross-national comparative research, that is, whether results are comparable across countries or not. As long as we have no information on how respondents in the different countries understand the term 'minorities,' it is unclear what similarities or differences observed across countries mean. 'Minorities' – unless clearly defined in a questionnaire – might encompass a broad meaning. The term might also convey different things to different people and, most importantly, the meaning attributed to this word might change from country to country. Consequently, comparability problems could be related to the kind of minorities

people think about in each country. There can be differences in the evaluation of the same minorities, too.

We have to acknowledge that it is not the current practice to just talk about ‘minorities’ in surveys that devote more than just cursory attention to ‘minorities’. In such studies, even terms like ‘immigrants’ are not used in a generic sense but divided into more concrete groups, such as by naming different ethnicities or by specifying what categories of immigrants respondents are supposed to focus on. This is both the case for topical modules of the Eurobarometer and the ESS or other specialized studies where only titles of publications include just the term ‘minorities’ (Coenders, Lubbers, & Scheepers, 2004; European Monitoring Centre on Racism and Xenophobia, 2005). Nevertheless, as the present example shows, there are also studies which leave the term ‘minorities’ unspecified.

To the best of our knowledge, there is some but not much research on the understanding of ‘minorities’ among respondents. Seyranian, Atuel, & Crano (2008) investigate the dimensions along which respondents in the U.S. characterize the terms ‘minority’ and ‘majority groups.’ They find that eight dimensions are used by their 77 student respondents: power and status, numeric size, distinctiveness, social category, group context, dispositions, and being the source or target of a treatment. Among these dimensions, social category, which was the most frequently mentioned in their study, matches most closely the particular types of groups we are focusing on here: gender, ethnicity, socioeconomic status, level of education, disability, sexual orientation, and age. However, compared to the Seyranian et al. study, we will avoid the restriction to a student sample when it comes to the understanding of ‘minorities’. We rather aim to cover the general population and, additionally, to add an international comparative dimension.

For a related question on ‘minorities’, the ESS conducted 10 internationally comparative cognitive interviews in five countries ($n = 50$) to investigate possible problems and to probe the understanding of the term ‘minorities’ (Widdop, Fitzgerald, & Gatrell, 2011). Although problems of understanding the question were found and the term ‘minority’ was not understood in the same way in the different countries, the question formed part of the final questionnaire of ESS Round 6 in 2012 (European Social Survey, 2012). However, on the basis of only 10 cognitive interviews per country, the extent of differences in the associations respondents have with the term ‘minority’ cannot be quantified and the implications for the interpretation of the resulting data not properly gauged.

Taking all of this into account, we implemented web probing of the ISSP item to examine the understanding of the term ‘minorities’ and we used an elaborate coding scheme to compare the meaning of ‘minorities’ in different cultural contexts. Our research questions are as follows: Can the differences in the observed answers for the closed item between the countries under investigation

be explained by differences in the associations that the term 'minority' evokes? And do respondents apply different evaluations of the mentioned minorities? In pursuing these questions, we want to demonstrate the utility of web probing as a method to produce sufficient qualitative insights to assess the validity and comparability of the question.

DATA AND METHODS

Sample

We used non-probability online surveys conducted in Germany, Britain, the U.S., Mexico, and Spain with a total of 2,689 respondents (net samples of 480 respondents in each country were targeted). Only citizens of the respective countries aged 18 to 65 were eligible to participate in the web survey. We implemented quotas for age (18-30, 31-50, and 51-65), gender, and education (lower vs. higher education). All quotas were met. Mean age was 42 years in Germany and the U.S., 41 years in Britain and Spain, and 40 years in Mexico. Data collection took place in June 2014. The panel provider was Respondi and its respective partners in the different countries. As the surveys were based on quota samples, standardized response rates cannot be computed (Baker et al., 2010).

The selection of the five countries should mainly provide for some variation with regard to the presence of ethnic minorities and migrant groups. In addition, the inclusion of the two Anglo-Saxon countries (Great Britain and the U.S.) in the study was also motivated by the expectation that the traditionally high concern for issues of gender and sexual orientation might also play a role in the understanding of 'minorities.'

Questionnaire

The item battery, taken from the 2014 Citizenship II module of the ISSP, was introduced as: 'There are different opinions about people's rights in a democracy,' followed by the first item of the scale (on standard of living). The second item, appearing on a second screen, read: 'And how important is it that government authorities respect and protect the rights of minorities?' The response scale ranged from 1 (*not at all important*) to 7 (*very important*).

After selecting a response category, the respondents received a specific probe on a subsequent screen: 'What particular minority groups did you have in mind when you were answering the question?' In addition, the respondents were reminded of the original question wording: 'The question was: "And how important is it that government authorities respect and protect the rights of minorities?"' (see Figure 1).


 Leibniz-Institut
für Sozialwissenschaften

And how important is it that government authorities respect and protect the rights of minorities?

not at all important							very important
1	2	3	4	5	6	7	
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>


 Leibniz-Institut
für Sozialwissenschaften

What particular minority groups did you have in mind when you were answering the question?

The question was: "And how important is it that government authorities respect and protect the rights of minorities?"

Figure 1. Closed ISSP item and probing question in the web survey

Coding of the open-ended answers

The answers to the probes were not translated but immediately coded by members of the project team (German native speakers with high proficiency in English and/or Spanish) who had been briefed on the particularities of these open-ended answers as well as on specific coding needs. An elaborated category scheme was developed, which represents the main associations of ‘minorities.’ This scheme was based both on the literature, in particular the results of Seyranian, Atuel, & Crano (2008), and also heuristically developed, based on the content of the probe responses.

After the establishment of a final coding, team members not involved in the development and implementation of the coding scheme coded 50% of the probe answers of each country. Interrater agreement was on average 88%, ranging from 86% for the Mexican to 91% for the U.S. data (see Table 1). This means that in nearly 9 out of 10 cases, the raters coded a probing answer identically. Given that the category scheme was fine-grained (including many sub-codes in addition to the main categories), this is an encouraging result. In addition, most discrepancies occurred with the distinctions between different sub-codes of nonresponse answers which are, however, not central to the present paper. All coding discrepancies were discussed within the research team, which then came to an agreement regarding the final version used in this paper.

Table 1. Intercoder reliabilities for the answers to the probing question (in %)

Germany	Britain	U.S.	Mexico	Spain	Total
89	87	91 ^a	86	89	88

Note. Before correcting some minor technical coding errors related to different sub-codes of nonresponse answers in the U.S., the result of the first analysis of the intercoder reliability for the U.S. data was 80% (resulting in a total reliability of 86%).

The coding scheme

Table 2 gives an overview of the main categories and their specific sub-categories as well as their respective descriptions, coding instructions, and examples. Multiple coding was possible for all categories except for the categories *no generalization possible* and the *non-substantive rest category* (which includes, among others, *nonresponse*).

Table 2. Main categories and their sub-categories

Main codes	Sub-codes	Remarks/examples
Ethnicity & migration	<i>general code</i>	reference of a general kind or to a group not covered by the following codes:
	Asians	exception: Arabs and Arabic countries got a separate code, even if in Asia
	Arabs	
	Hispanics	exception: Mexicans, when explicitly mentioned, received a special code
	Mexicans	
	Blacks	blacks, African Americans or similar descriptions like colored people, 'negros'
	indigenous groups	Indians were coded depending on the context: in the U.S. 'Indians' were coded here but in Britain 'Indians' were coded as 'Asians'
	Sinti and Roma	including other names such as 'gypsies'
	immigrants	migration process or background had to be mentioned; exceptions: illegal immigrants and refugees received separate codes
	illegal immigrants	
	refugees	also if 'asylum seeking' was mentioned
	countries/regions of origin	if there was no separate code for a country or region; e.g., 'Eastern Europeans'
Majority group		used by respondents to express their resentment about the majority not being treated adequately; e.g., 'white people'

Main codes	Sub-codes	Remarks/examples
Religion	<i>general code</i>	if the concept of religion was mentioned ('Religious dressed minorities!') or a religion that is not covered by one of the following sub-codes:
	Muslims	
	Jews Christians	
Health	general code	e.g., 'long-term sick'
	disabled	e.g., 'those with learning disabilities'
Age	general code	if just the concept 'age' was mentioned
	children	
	old people/senior citizens	
Social class	<i>general code</i>	if the concept was mentioned or if a reference was made to education or the financial status of a person; e.g., 'poor families that cannot afford to eat'
	single parents	
	unemployed	
	homeless	
Gender	<i>general code</i>	if just the concept 'gender' was mentioned
	women	
	men	
Sexual orientation		refers to people who differ from the heterosexual norm or binary gender roles; e.g., lesbians, gays, bisexuals, and transgenders (LGBT)
All/None		if respondents mentioned 'all minorities' or 'none in particular'
Other		for answers for which there is no a substantial code but which potentially have a substantive meaning; e.g., 'Nazis,' 'feminists,' and 'prisoners'
Non-substantive		includes complete nonresponse, don't know, refusal, problems with question, nonsense answer, mismatching answer, and other nonresponse

Analytical strategy

In the following, we first compare the response patterns found in the ISSP data from 2004 and 2014 with the patterns revealed by our web survey to assess the general usefulness of our web survey data. Second, we report the responses to the specific probe regarding which 'minorities' respondents had in mind when answering the closed question. Finally, we regress the scores of the respondents for the closed question on the most important categories of open responses. We do so for two reasons: on the one hand, to find out how much of the answers

to the closed question can be explained by the specific minorities respondents had in mind and, on the other hand, to see which categories are related to more positive or negative attitudes with regard to minority rights. Hence, our analytical strategy can be characterized as a mixed-methods approach in which qualitative and quantitative components are intertwined (Brannen, 2005).

RESULTS

Comparison of ISSP and web survey

In a first step, the web survey data are compared to the available 2004 and 2014 ISSP data to check whether the former can be used to shed light on the latter (see Table 3). We start by looking at the mean values of the 2014 ISSP data and find that support for the rights of minorities was particularly strong in Spain ($M = 6.7$). The same value was reached in Mexico in 2014 (no ISSP data was available for 2004 for this country). Support was weaker in Germany and the U.S. (both $M = 6.1$) and, particularly, in Britain ($M = 5.9$). Whether the differences or similarities between Spain and the other countries mean anything remains unclear as long as we do not have enough information on how respondents understood the term 'minorities.'

Table 3. Mean values of minority item in the ISSP and web surveys (number of cases in parentheses)

	ISSP 2004	ISSP 2014	Web survey
Germany	6.2 (1,305)	6.1 (1,693)	5.9 (265)
Britain	5.8 (811)	5.9 (1,489)	5.6 (254)
U.S.	6.2 (1,456)	6.1 (1,225)	5.9 (256)
Mexico	6.7 (1,199)	-	6.6 (266)
Spain	6.5 (2,398)	6.7 (1,711)	6.2 (274)

Note. Item measured on a scale from 1 (*not at all important*) to 7 (*very important*); Mexican data not available for 2014

The ISSP and the web surveys show similar results. Support for minorities is always highest in Mexico and Spain and lowest in Britain. Germany and the U.S. are on an equal intermediary level. Nevertheless, country differences are relatively small.

The amount of nonresponse for the closed item is also very similar in the ISSP and web survey data (Table 4). The only exception is Britain, which is clearly less affected by nonresponse in the web survey compared to the ISSP surveys.

Particularly noteworthy is the Mexican case where nonresponse for the closed item is practically not present for either survey.

Table 4. Nonresponse of minority item in the ISSP and web surveys in percent (number of nonresponse cases in parentheses)

	ISSP 2004	ISSP 2014	Web survey
Germany	2.0 (27)	1.5 (25)	2.2 (6)
Britain	4.9 (42)	5.8 (91)	1.2 (3)
U.S.	1.1 (16)	3.1 (39)	1.5 (4)
Mexico	.2 (2)	-	.8 (2)
Spain	3.4 (83)	2.5 (44)	1.8 (5)

Both the patterns of the means and of the nonresponse indicate that we may fruitfully use the web survey to shed light on the ISSP data.

Comparison of open-ended answers across countries

Who, then, are the minorities people have in mind in the different countries? This is what we aimed to find out with the probing question. We first look at the broad categories but also keep in mind that sub-groups may show different importance in different countries.

Ethnicity and migrants are frequently mentioned everywhere, ranging from 26% in Mexico to 48% in Germany (Table 5). However, the composition of this category differs across countries. When looking at the relevant ethnicity sub-codes, in Mexico 18% of the respondents mention *indigenous groups*, while only 2% of the respondents in the U.S. think of them (Table 6). This sub-group is virtually non-existent in the other three countries. *Hispanics* (including *Mexicans*) and *Blacks* are the dominant sub-codes in the U.S. with 21% each but they are almost absent everywhere else. *Sinti and Roma* are frequently mentioned only in Germany and Spain (6% in each of these countries). Germans and – to a smaller degree – the Spanish and British often have *immigrants* in general in mind (28%, 18%, and 12%, respectively). Finally, *refugees*, as a special category, are only relevant in Germany (11%). This is interesting given that the survey was conducted prior to the ‘refugee crisis’, which had its peak in 2015.

Table 5. Distribution of open-ended answers across countries (in %)

	Germany	Britain	U.S.	Mexico	Spain
Ethnicity	48	32	41	26	34
Majority group	1	4	5	0	0
Religion	6	15	2	6	8
Health	17	9	2	11	15
Age	11	4	2	6	9
Social class	21	6	2	38	30
Gender	1	1	3	2	1
Sexual orientation	15	6	6	10	8
All/None	10	24	29	4	10
Other	2	3	1	12	9
Non-substantive	13	21	23	23	13
N	265	254	256	266	274

Table 6. Distribution of selected sub-codes of the ethnicity category across countries (in %)

	Germany	Britain	U.S.	Mexico	Spain
General code	9	11	5	8	14
Asians	0	2	3	0	0
Arabs	0	0	0	0	0
Hispanics	0	0	16	0	1
Mexicans	0	0	5	0	0
Blacks	1	0	21	0	0
Indigenous	0	0	2	18	0
Sinti and Roma	6	2	0	0	6
Immigrants	28	12	2	1	18
Illegal immigrants	0	1	2	0	1
Refugees	11	2	0	0	0
Countries/regions of origin	2	4	1	0	0
N	265	254	256	266	274

Majority group is only relevant in the U.S. and Britain (5% and 4%, respectively). This category is used by respondents to express their resentment with regard to the majority not being treated adequately. In a strict sense, this is not a minority group, however.

Religion is by a large margin most relevant in Britain (15%). On the basis of the present data, we cannot determine whether and to what degree religious groups are used interchangeably with ethnic groups in some of the countries but not in others. For example, British respondents might be more prone to mention ‘Muslims’ when thinking about Pakistanis, while Turks in Germany might be characterized by their ethnicity and not their religion.

Groups defined by their *health* status are most frequently mentioned in Germany and Spain (17% and 15%, respectively), while they get particularly short shrift in the U.S. (2%). Similar patterns are obtained for *age* and *social class*. *Age* is less relevant in Britain (4%) and the U.S. (2%) compared to Germany (11%) and Spain (9%). For *social class* (e.g., ‘the poor’), the differences between the Anglo-Saxon countries and the others are even more extreme. Only 6% of the respondents in Britain and 2% in the U.S. think of minority groups defined by social class. In Mexico, Spain, and Germany many more respondents mention such groups (38%, 30%, and 21%, respectively).

Surprisingly, *gender* as such seems to be quite unimportant, only between 1% and 3% of the respondents mention it. However, *sexual orientation* comes more frequently to mind, ranging from 15% in Germany to 6% in the Anglo-Saxon countries.

In the U.S (29%) and Britain (24%), respondents state that they have included *all minorities* – or *none* – in their judgement more often than is the case in the other countries (between 4% and 10%). Whether this is a possible consequence of the highly multicultural character of the two Anglo-Saxon countries, of satisficing or other reasons cannot be decided on the basis of the data available.

The frequent use of the *other* category (which comprises answers that did not occur frequently enough to justify the addition of additional substantive categories) in Mexico (12%) is often related to a variety of geographical entities, e.g., specific villages or regions.

The prevalence of *non-substantive* answers is highest in Mexico, the U.S., and Britain (with slightly more than 20% each) and lower in Germany and Spain (13% each). In Britain and the U.S. (15% and 13%, respectively), the underlying reason is mostly nonresponse. In Mexico, however, it is a mismatch (19%), that is, respondents give a reason for their answer to the closed question instead of listing the kind of minorities they thought of. Although Mexicans are eager to give an answer (in contrast to simple nonresponse), they provide an answer to a question that was not exactly asked. This finding is in line with previous research (Meitinger, Braun, & Behr, 2018).

How are minorities evaluated?

Table 7 answers the question regarding the impact that thinking of specific minority groups has on providing a response to the closed item. For this purpose, the answers to the closed item are regressed on the main categories extracted from the open-ended responses.

Table 7. Unstandardized regression coefficients of closed item on main categories

	Germany	Britain	U.S.	Mexico	Spain
Ethnicity/migrants	-.1	-.3	-.3	.3	.1
Religion	-.2	-.4	.7	.1	-.2
Health	.4*	.5	.1	.3	.2
Age	.1	.3	1.0	.1	.3
Social class	.5*	.4	0	.2	.5*
Sexual orientation	.2	.9*	.9*	-.3	0
Explained variance (in %)	4.3	2.7	2.8	.4	4.7
N	265	254	256	266	274

Note. The baseline is constituted by the categories: *majority group, all/none, other, and non-substantive*; coefficients marked by * are significant on the .05 level.

Mexicans show considerable support for minorities nearly independent of their associations. On the contrary, in Britain and the U.S., support for the rights of minorities is increased when sexual orientation (LGBT) comes to mind. This might be related to the LGBT movement which has held so much sway in the Anglo-Saxon countries. In Germany and Spain, support of minorities is larger when social class comes to mind. For German respondents, this is also the case when they think of health status (disabled people). However, the variance explained by the associations is relatively low, ranging from less than 1% in Mexico to nearly 5 % in Spain.

CONCLUSIONS

The qualitative answers offered by the respondents revealed that people in the different countries have partly different kinds of 'minorities' in mind. In particular, in Germany, Mexico, and Spain, people often think of social groups defined in terms of social class as minorities (e.g., 'the poor'), which is not the case in both Anglo-Saxon countries. In addition, the consequences which respondents' associations have for their support of 'minorities' differ between countries. In all countries,

respondents are less supportive of minorities when thinking of ethnic minorities compared to when thinking of sexual orientation. The latter holds particularly true in the Anglo-Saxon countries. Positive attitudes are also expressed when thinking of social class in Germany and Spain, and health in Germany. However, the explained variance of support for minorities by the associations is relatively low – and in Mexico it is virtually nil. Thus, overall the effect of thinking about different minorities on answering the closed item is not dramatic.

Despite this rather positive outlook, for international comparisons – and most likely for comparisons across time as well – it is not advisable to leave the term ‘minorities’ unspecified. As the example of the ISSP (and the ESS) shows, such terms occasionally find their way into cross-national surveys, and then it is paramount to be aware of – even better: learn more about – the different associations that respondents have in mind when responding to items containing such vague terms.

The point here is not that abstract terms per se should be avoided in drafting survey questions. In the present case with five countries examined, we see that using the term does not dramatically harm comparability. An abstract term could even be helpful in establishing comparability, which would be impossible to achieve by asking country-specific questions (that is, questions referring to the ethnic minority most prominent in the single countries). However, if respondents in different countries think about qualitatively entirely different categories (such as social class, sexual orientation, and health in addition to ethnic minorities), comparability is compromised.

Some limitations of our study have to be mentioned. First, the data used for this approach is based on a non-probability online survey. To justify that a comparison to the ISSP survey is still accurate, we highlighted that the distribution of the central variable in the web survey is similar to the distribution in the ISSP survey. However, this does not guarantee that there are no differences between the web survey and the general population.

Second, the data does not cover all cultural contexts but is limited to five countries. Additionally, Germany, Britain, the U.S., Mexico, and Spain are not representative for the full variety of countries covered by the ISSP survey and, thus, the results are restricted to the selected countries. To overcome this problem, a replication of this study in other member countries of the ISSP is needed to describe a more general tendency.

Third, the context of the question plays an important role when it comes to what people have in mind when they think about ‘minorities.’ In the present case, the question was asked in the context of citizenship rights. Different results might have been obtained in the context of a questionnaire on inequality or on national identity. In particular, the context should be influential for the relative weight of

ethnicity and social class. However, the considerable differences we found between countries are less likely to be affected by different questionnaire contexts. Instead, they probably mirror realities in the different countries.

Finally, while we probed only the term 'minorities,' there are other terms in the item which might be interpreted differently by respondents in different countries, such as 'government authorities,' 'protect,' and 'rights.' All of these terms might also have an impact on the interpretation of the term 'minorities' and might thus have contributed to the different interpretations that respondents in the different countries had. In principle, it would have been possible to test all these terms with separate probing questions in this study, but we have not done this here. It should be kept in mind, however, that asking cognitive probes about many different aspects of a single item can be burdensome for respondents, since no interviewer is present to maintain respondent motivation. That it is still possible to certain degrees is shown by Meitinger, Braun, & Behr (2018).

The web-probing approach was crucial in assessing the understanding of the term 'minorities' in the different countries. On the one hand, a mainly quantitative approach to establishing equivalence across countries would not have been possible due to the fact that there was only one pertinent item. On the other hand, the conduct of traditional cognitive interviews would not have secured for us the sufficient number of cases to assert whether or not there are country differences. Whenever there is only one item to be evaluated and a large number of cases is necessary, web probing is the method of choice. Of course, web probing can also supplement quantitative testing of measurement equivalence when there are multiple items (Meitinger 2017).

NOTES

- 1 This research was funded by the German Research Foundation (DFG) as part of the project "Optimizing Probing Procedures for Cross-National Web Surveys" [BR 908/5-1 to Michael Braun, Wolfgang Bandilla, and Lars Kaczmarek]. The authors would also like to thank two anonymous reviewers of the journal for their constructive and helpful comments.
- 2 A mismatch occurs, for example, when a respondent gives an answer to a category-selection probe (e.g., explains the reasons for answer selection) in response to a comprehension probe.

REFERENCES

- Baker, R., Blumberg, S. J., Brick, J. M., Couper, M. P., Courtright, M., Dennis, J. M., ... Zahs, D. (2010). AAPOR report on online panels. *Public Opinion Quarterly*, 74, 711–781. <https://doi.org/10.1093/poq/nfq048>
- Beatty, P.C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, 71, 287–311. <https://doi.org/10.1093/poq/nfm006>

- Benítez, I., & Padilla, J. L. (2014). Analysis of nonequivalent assessments across different linguistic groups using a mixed methods approach: Understanding the causes of differential item functioning by cognitive interviewing. *Journal of Mixed Methods Research*, 8, 52–68. <https://doi.org/10.1177/1558689813488245>
- Behr, D., Meitinger, K., Braun, M., & Kaczmirek, L. (2017). Web probing – implementing probing techniques from cognitive interviewing in web surveys with the goal to assess the validity of survey questions. *GESIS – Survey Guidelines*. GESIS – Leibniz-Institute for the Social Sciences. https://doi.org/10.15465/gesis-sg_en_023
- Brannen, J. (2005). Mixing methods: The entry of qualitative and quantitative approaches into the research process. *International Journal of Social Research Methodology*, 8, 173–184. <https://doi.org/10.1080/13645570500154642>
- Coenders, M., Lubbers, M., & Scheepers, P. (2004). Majority populations' attitudes towards migrants and minorities. *Report for the European Monitoring Centre on Racism and Xenophobia*. Wien: EUMC. Ref. no. 2003/04/01. <http://fra.europa.eu/en/publication/2005/majorities-attitudes-towards-migrants-and-minorities-key-findings-eurobarometer-a-0>
- Davidov, E., Dülmer, H., Schlüter, E., Schmidt, P., & Meuleman, B. (2012). Using a multilevel structural equation modeling approach to explain cross-cultural measurement noninvariance. *Journal of Cross-Cultural Psychology*, 43, 558–75. <https://doi.org/10.1177/0022022112438397>
- Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, 40, 55–75. <https://doi.org/10.1146/annurev-soc-071913-043137>
- European Monitoring Centre on Racism and Xenophobia (2005). *Majorities' attitudes towards minorities. Key findings from the Eurobarometer and the European Social Survey – Summary*. Brussels, Belgium: European Monitoring Centre on Racism and Xenophobia. Summary. <http://fra.europa.eu/en/publication/2005/majorities-attitudes-towards-migrants-and-minorities-key-findings-eurobarometer-a-0>
- European Social Survey (2012). *ESS round 6 source questionnaire*. London: Centre for Comparative Social Surveys, City University London.
- Fitzgerald, R., Widdop, S., Gray, M., & Collins, D. (2011). Identifying sources of error in cross-national questionnaires: Application of an error source typology to cognitive interview data. *Journal of Official Statistics*, 27, 569–599.
- ISSP Research Group (2016). *International Social Survey Programme. Citizenship II – ISSP 2014*. GESIS Data Archive, Cologne. ZA6670 Data file Version 2.0.0. <https://doi.org/10.4232/1.12590>
- Lee, J. (2014). Conducting cognitive interviews in cross-national settings. *Assessment*, 21, 227–240. <https://doi.org/10.1177/1073191112436671>
- Miller, K., Fitzgerald, R., Padilla, J.-L., Willson, S., Widdop, S., Caspar, R., ... Schoua-Glusberg, A. (2011). Design and analysis of cognitive interviews for comparative multinational testing. *Field Methods*, 23, 379–396. <https://doi.org/10.1177/1525822X11414802>
- Meitinger, K. (2017). Necessary but insufficient: Why measurement invariance tests need online probing as a complementary tool. *Public Opinion Quarterly*, 81, 447–472. <https://doi.org/10.1093/poq/nfx009>

- Meitinger, K., & Behr, D. (2016). Comparing cognitive interviewing and online probing: Do they find similar results? *Field Methods*, 28, 363–380. <https://doi.org/10.1177/1525822X15625866>
- Meitinger, K., Braun, M., & Behr, D. (2018). Sequence matters in online probing: The impact of the order of probes on response quality, motivation of respondents, and answer content. *Survey Research Methods*, 12, 103–120.
- Prüfer, P., & Rexroth, M. (2005). Kognitive Interviews. *ZUMA How-to-Reihe*, 15. Retrieved from http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/howto/How_to15PP_MR.pdf?download=true
- Schuman, H. (1966). The random probe: A technique for evaluating the validity of closed questions. *American Sociological Review*, 31, 218–222. <https://doi.org/10.2307/2090907>
- Seyranian, V., Atuel, H., & Crano, W. D. (2008). Dimensions of majority and minority groups. *Group Processes & Intergroup Relations*, 11, 21–37. <https://doi.org/10.1177/1368430207084843>
- Thrasher, J. F., Quah, A. C. K., Dominick, G., Borland, R., Driezen, P., Awang, R., ... Boado, M. (2011). Using cognitive interviewing and behavioral coding to determine measurement equivalence across linguistic and cultural groups: An example from the International Tobacco Control Policy evaluation project. *Field Methods*, 23, 439–460. <https://doi.org/10.1177/1525822X11418176>
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices and recommendations for organizational research. *Organizational Research Methods*, 3, 4–70. <https://doi.org/10.1177/109442810031002>
- Vijver, F. J. R. van de, & Chasiotis, A. (2010). Making methods meet: Mixed designs in cross-cultural research. In J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. P. Mohler, B.-E. Pennell, & T. Smith (Eds.), *Survey methods in multinational, multiregional, and multicultural contexts* (pp. 455–473). Hoboken, NJ: Wiley. <https://doi.org/10.1002/9780470609927.ch24>
- Widdop, S., Fitzgerald, R., & Gatrell, L. (2011). European Social Survey round 6 cognitive pre-testing report. London: Centre for Comparative Social Surveys.
- Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781412983655>
- Willis, G. B. (2015). The practice of cross-cultural cognitive interviewing. *Public Opinion Quarterly*, 79, 359–395. <https://doi.org/10.1093/poq/nfu092>

Michael Braun is Senior Project Consultant at the department Survey Design and Methodology (SDM) at GESIS – Leibniz-Institute for the Social Sciences and Adjunct Professor at the University of Mannheim. He has specialized in cross-cultural survey methodology and analysis. He has worked substantively on international comparisons in the fields of migration and the family.

Dorothée Behr is a Senior Researcher at the department Survey Design and Methodology (SDM) at GESIS – Leibniz-Institute for the Social Sciences. Her research focuses on cross-cultural survey methodology, in particular questionnaire translation, the interplay between questionnaire design and translation, and the method of web probing.

Katharina Meitinger is an Assistant Professor at Utrecht University . Her research focuses on cross-cultural survey methodology, in particular mixed-method designs and the method of web probing.

Klara Raiber is a Ph.D. student at Radboud University in Nijmegen. She was working as a student assistant at the department Survey Design and Methodology (SDM) at GESIS – Leibniz-Institute for the Social Sciences. She is interested in the topics of employment, informal care, and gender .

Lydia Repke is a Postdoctoral Researcher at the department of Survey Design and Methodology (SDM) at GESIS – Leibniz-Institute for the Social Sciences in Mannheim, Germany. Her main research interests lie in the areas of survey methodology (e.g., questionnaire design, data quality), social networks, and cross-cultural psychology (e.g., acculturation, biculturalism).