Survey Research Methods (2023)
Vol. 17, No. 1, pp. 63-74
doi:10.18148/srm/2023.v17i1.8016
European Survey Research Association

Ambiguity in the Item Wording, Ambiguity in the Respondents' Comprehension? An Experiment on the "Immigrants/Foreign Workers" Social Distance Item in Values Surveys

Riccardo Ladini and Ferruccio Biolcati
University of Milan, Department of Social and Political Sciences

Since there has been a considerable increase in studies on attitudes toward immigrants in recent years, particular attention must be devoted to the survey instruments used to measure them. Since they first appeared, both the European Values Study and the World Values Study have included a battery of questions to measure social distance, asking which social groups respondents would not be willing to have as neighbors. One of these social groups is "immigrants/foreign workers." The wording of this item contains a certain degree of ambiguity, as the terms "immigrants" and "foreign workers" refer to two different social groups. Foreign workers are in fact a subcategory of immigrants. Our contribution sets out to assess the working of the item, using a survey experiment randomly assigning respondents three different formulations ("immigrants/foreign workers", "immigrants", and "foreign workers"). The data come from the Italian joint edition of the European Values Study-World Values Survey 2017. The results show that Italians perceive a much larger distance toward immigrants than toward foreign workers, which is nevertheless similar to the distance perceived toward "immigrants/foreign workers." Overall, the empirical evidence suggests the semantic prevalence of the item "immigrants" over "foreign workers" in respondents' answers. In light of the results, we will discuss their implications for both data users and future values surveys.

Keywords: item wording; immigrants; foreign workers; social distance; European Values Study; World Values Survey

1 Introduction

In longitudinal survey research, question wording should remain as similar as possible across different points in time to enable the comparison of attitudes and behaviors over time. This is why longitudinal surveys include questions that maintain their original wording despite methodological developments in questionnaire design suggesting better formulations. An example comes from the European Values Study (EVS) and the World Values Survey (WVS), which were first carried out in 1981. Both surveys contain a standardized battery of social distance questions, commonly used to measure prejudice (Evans & Kelley, 2019; Shin & Dovidio, 2018) and intolerance (Rapp & Ackermann, 2016; Strabac et al., 2012). To answer these questions, respondents have to say which social groups, if any, they would not like to have as neighbors. Among the groups listed, both of the surveys include the item "immigrants/foreign workers" as an indicator of ethno-racial social distance. Since 1981, more than 600,000 individual respondents, living in over 100 countries, have provided an answer to the item in the values surveys. Nonetheless, the ambiguity of the item's formulation makes it impossible to identify the exact social categories respondents have in mind when they answer the question. Do the respondents' answers refer to immigrants? To foreign workers? To both?

This paper aims to address this issue by analyzing the working of the item through a survey experiment which randomly manipulates its wording. The objective of the experimental design, which was included in the joint Italian EVS-WVS 2017 survey, is to detect whether one of the two categories "carries more weight" in the respondents' evaluation when answering the question.

The findings showed that there is a great difference in the levels of social distance toward immigrants and foreign workers: the percentage of respondents not wanting immigrants as neighbors was more than twice as high as the percentage of respondents not wanting foreign workers. Moreover, respondents' answers suggested the semantic prevalence of the label "immigrants" in the item. The result has implications on this standard item used in comparative survey projects; thus, the concluding section will provide some suggestions for future EVS and WVS surveys.

Contact information: Riccardo Ladini, Via Conservatorio 7, 20122 Milano, Italy (E-mail: riccardo.ladini@unimi.it)

2 Background

Since 1910s, the widespread climate of discrimination and prejudice toward many immigrant groups in the United States has led social sciences scholars to turn increased attention to the study of ethnic and race relations (Wark & Galliher, 2007). In this socio-cultural environment, R. E. Park (1924, p. 339) first defined the concept of social distance as "an attempt to reduce to something like measurable terms the grades and degrees of understanding and intimacy which characterize personal and social relations generally." The concept can be measured thanks to Bogardus (1925), who created one of the oldest attitudinal scales in social and psychological research. Various versions of the scale have been proposed (Bogardus, 1925, 1933), but all of them ask respondents the type of relationship into which they are willing to accept the members of a certain social group.[1] The first version of Borgardus' (1925) scale gave the following types of relationship to choose from: close kinship by marriage; my club as personal chums; my street as neighbors; employment in my occupation in my country; as citizens in my country; as visitors only to my country; would exclude them from my country.

A simplified version of the Bogardus scale of social distance has been included in the EVS and WVS surveys since they first appeared in 1981. The question only considers one of the several items on the Bogardus scale, as it asks individuals whether they would like to have each of the social groups listed in a battery as neighbors. In the fifth and most recent EVS, the question reads as follows: "On this list are various groups of people. Could you please identify any that you would not like to have as neighbors? People of a different race; heavy drinkers; immigrants/foreign workers; drug addicts; homosexuals; Christians; Muslims; Jews; gypsies." Thus, individuals express social distance toward a certain social group by mentioning that category. Among the categories listed, 'immigrants/foreign workers' is the only one which includes two different categories of individuals: immigrants and foreign workers. The item is present in every EVS and WVS survey. Nonetheless, its meaning is ambiguous, as immigrants and foreign workers are not synonymous. Foreign workers are in fact a subcategory of immigrants.

The meaning that survey respondents give to the generic label "immigrants", as well as to the labels "strangers" and 'foreigners,' is very heterogenous. Spruyt et al. (2016) showed that the criterion most employed to define the generic item "strangers" was nationality. In addition, a sizeable portion of respondents (22%) associated strangers with people from Islamic countries, who also reported the highest level of ethnic prejudice. In Germany, when asked about "foreigners", the majority of natives thought of Turks, and in the East of the country those people are more prejudicial toward foreigners than others (Asbrock et al., 2014). Comparative research has also shown that up to one third of respondents in any of six countries (Canada, Denmark, Germany, Hungary, Spain, United States) think about a generic immigrant when asked about the label "immigrant" (Asbrock et al., 2014; Braun et al., 2013; Spruyt et al., 2016).

Other than reflecting the distribution of immigrant groups by background in a certain country, media representation has a great impact on the meaning of the label "immigrants" among the public (K. Park, 2014). As people are often not familiar with the phenomenon, especially in countries with a recent immigration history, they tend to refer to what Blinder (2015) defined as "imagined immigration". For instance, British people thought that the first reason for immigration to the UK is to seek asylum, even though according to official statistics asylum seekers make up a very small minority of immigrants. Moreover, there has been a substantial overlap between the terms "migrant" and "refugee" in media language (Georgiou & Zaborowski, 2017), especially during the 2015 European migration crisis, and the salience of the migration crisis itself could have helped reinforce the opinion that immigrant is equal to refugee.

These considerations also hold for the Italian context. The migration crisis had a huge impact on Italy, with more than 600,000 sea arrivals in the period 2014–2017 (UNHCR data). During that time span, the immigration issue became much more salient among the public, as the percentage of people who considered immigration the first or the second most important problem faced by the country rose from 4% in 2013 to 23% in 2018 (Vezzoni, 2018). Moreover, in the Italian public opinion, even before the migration crisis, the immigration phenomenon was generally painted in negative terms as "a massive invasion" or "a threat" to be addressed by "the politics of fear" (Colombo, 2018; Quassoli, 2013).

Given this framework, the "immigrants/foreign workers" item employed in the EVS and WVS surveys is deemed to contain a certain degree of ambiguity. Indeed, previous research had shown that in Britain only half of respondents perceived immigrants as workers (Blinder, 2015). "Foreign workers" is a more specific label than "immigrants", since it only refers to those foreign people working in the destination country. Unlike the label "immigrants", we tend to rule out that respondents associate the label "foreign workers" with the refugees.

Given that the meanings provided by the same person for the label "immigrants" and "foreign workers" are different, we hypothesize that the two labels are not equivalent. Therefore, we should expect that the proportion of respondents not willing to have foreign workers as neighbors is different to the proportion of respondents not wanting immigrants as neighbors. Moreover, we need to understand whether one label has semantic prevalence over the other. If so, when answering the question on social distance toward "immi-

[1] In Bogardus' (1925, 1933) seminal scales, social groups are all defined in terms of "racial groups".

grants/foreign workers", individuals will be more likely to refer to immigrants than foreign workers, or vice versa. This possible result could have substantial implications on the functioning of the original item. Above all, the survey data provides estimations of a proportion of people not willing to have a nevertheless ambiguous group of people as neighbors. While this ambiguity will have less influence when the item is combined with others in a generic multiple-item scale of social tolerance (Dunn & Singh, 2011; Kirchner et al., 2011; Rapp & Ackermann, 2016), the issue becomes especially relevant when analyzing the single item, both as a dependent (Doebler, 2014; Ruedin, 2019; Shin & Dovidio, 2017, 2018) and independent variable (Haney, 2016).

3 Data, measures, and experimental design

The data come from the Italian joint EVS-WVS 2017 survey. By means of a probabilistic three-stage (municipalities, electoral sections, and individuals) and three-domain (self-representative Italian cities with more than 500,000 residents aged 18 and over, provincial capital cities, other municipalities) sampling design, 2,277 Italian citizens aged 18 and over were CAPI (Computer Assisted Personal Interviewing) interviewed between September 24, 2018 and January 30, 2019. Our empirical analysis thus considers the Italian subsample of the European Values Study 2017 Integrated Dataset.

The EVS and WVS master questionnaire includes a battery of items measuring social distance, using the wording given above. Respondents are asked to provide a dichotomous answer (mentioned/not mentioned) for all social groups; therefore, they have to mention which of the listed social groups they are not willing to have as neighbors. Our research questions were answered by analyzing a survey experiment which randomized the formulation of the item concerning social distance towards immigrants and foreign workers: 70% of respondents received the standard item "immigrants/foreign workers" ($n = 1,596$),[2] while the remaining 30% of respondents were respectively assigned the "foreign workers" ($n = 324$) or "immigrants" item ($n = 357$).[3] Randomization checks showed that the distributions of the experimental groups—by item formulation—did not significantly differ in terms of socio-demographic variables or left-right orientation (p-value > 0.05 in every chi-square test. See Table A1 in Appendix).

4 Methods

To test the equivalence across the three items, we first compared the distribution of the manipulated item depending on its formulation. Significant differences across the distributions were preliminary evidence of the non-equivalence of the items "immigrants" and "foreign workers". To assess whether one item had semantic prevalence over the other, we compared the differences in the percentages of affirmative answers between the original item with the two manipulated ones respectively. For instance, if there were differences in the distribution of the three items but similarities in the percentage of affirmative answers between the original item and the two manipulated ones respectively, we could suggest an overall balance between the "immigrants" and "foreign workers" labels in respondents' answers to the original item

Moreover, we employed multi-group item response theory (IRT) models, also known as differential item functioning (DIF) analysis, to test whether the working of the manipulated item in relation to the other items in the battery of social distance items varied depending on its formulation. These models enabled us to test the measurement equivalence of the scale of social distance across the three experimental groups, namely, whether "the relations between observed test scores and the latent attribute measured by the test [were] identical across subpopulations" (Drasgow, 1984, p. 134). In our application, we analyzed whether the functioning of the three different formulations of the manipulated item varied when associated with the other items of social distance. It could be argued that DIF was present when the relationship between the item score and the latent trait varied across the groups. When analyzing the measurement equivalence of scales of dichotomous items, the main advantage of IRT models over Confirmatory Factor Analysis is that they avoid expected item responses lower than zero and higher than one and do not consider dichotomous answers as continuous latent scores (Tay et al., 2015). Using multi-group IRT models as a framework, we formally tested the measurement equivalence of the scale of social distance depending on the formulation of the manipulated item by comparing two nested models: a less constrained model which only allowed DIF for the manipulated item across two experimental groups, while all the parameters concerning the other items were fixed between the two groups, and a more constrained model where the parameters of every item of the scale of social distance was fixed between the groups. Likelihood ratio tests were used to test the difference between the two IRT models. If the differences were found to be statistically significant, this gave evidence of DIF: namely, a significant difference between the parameters of the manipulated items in the more constrained and the less constrained models. On the contrary, if the likelihood ratio test led to non-statistically significant differences between the models, this was not evidence of DIF: in other words, there was no substantial variation in the working of the two items within the scale of social distance. In detail, we used two pairwise likelihood ratio

[2] To preserve comparability across time and space, the majority of respondents were assigned the original item.

[3] To guarantee the internal validity of the experiment, namely, that respondents answered the randomly assigned item, interviewers received specific training on this question. They were explicitly asked to read the exact text of the item as displayed on the monitor of their tablet.

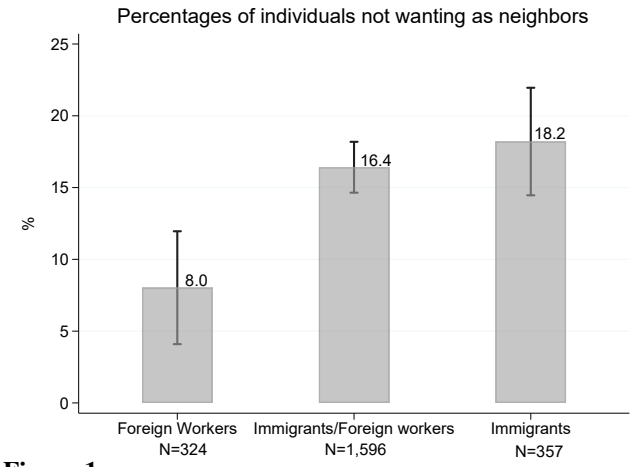

Figure 1

Percentage distribution of the manipulated item by experimental condition (95% confidence intervals)

tests between the experimental group answering the original item and the groups answering the "foreign workers" and "immigrants" items respectively. This is because we were interested in analyzing whether the latter items functioned differently to the original one. All these analyses were carried out with the "irt" package of Stata 16.

5 Results

Figure 1 shows the proportion of individuals who declared to be unwilling to have foreign workers, immigrants/foreign workers, and immigrants respectively as neighbors. The percentage of respondents not wanting immigrants/foreign workers and immigrants as neighbors was similar, at 16.4% and 18.2% respectively. As the overlapping confidence intervals indicate, the difference between the two percentages was not statistically significant (z-test for differences between two proportions = −0.82; p-value = 0.41). Instead, the percentage of respondents not wanting foreign workers as neighbors was far lower—less than half, equal to 8%—compared to immigrants/foreign workers (z-test for differences between the two proportions = 3.86; p-value = 0.0001). Thus, preliminary analyses based on comparison of the percentage answers to the different formulations of the item suggested that the "immigrants" category carried much more weight than the "foreign workers" category when answering the original item. In other words, since the difference in the proportion of answers to the "immigrants" item and the "immigrants/foreign workers" item was minimal and not statistically significant, this provided preliminary evidence toward an equivalence between the two items.

To evaluate the possible DIF on the manipulated scale of social distance item across the experimental groups, we assessed the DIF according to the formulation of the manipulated item. A preliminary analysis showed that a two-parameter logistic (2PL) IRT model was more appropriate than a one-parameter model to analyze the measure of the scale of social distance: we estimated and compared the two models on the group of respondents assigned the "immigrants/foreign workers" item, and the likelihood ratio test was found to be statistically significant (likelihood ratio chi-square statistic (7) = 118.3; p-value = 0.000).[4]

Table 1 compares the results of two two-group 2PL IRT models where the two groups of respondents were those assigned respectively the "immigrants/foreign workers" and "foreign workers" items. A more constrained model (right panel) assumed that for all eight items in the scale of social distance, the parameters of difficulty – the point in the ability scale at which the probability of answering the item affirmatively is 50%—and discrimination—the capability of an item to distinguish respondents on the latent trait measured by the item—would be equal across the two groups of respondents. More specifically, it was assumed that the parameters of the IRT model estimated for the respondents assigned the "immigrants/foreign workers" item would be the same for the respondents assigned the "foreign workers" item. A less constrained model instead assumed that the parameters of difficulty and discrimination of the manipulated item varied across the two groups, while all the other parameters were assumed to be equal across the groups. In other words, the more constrained model assumed that there would be no DIF on the manipulated item ("immigrants/foreign workers" vs "foreign workers") while the less constrained item assumed there would be DIF on that item.

Overall, the results showed that the less difficult items (gypsies, drug addicts, heavy drinkers) were also less able to discriminate across different levels of the latent variable of social distance. In the less constrained model, the results showed a difference in the difficulty parameter for the manipulated item across the two groups (1.12 for the group assigned the "immigrants/foreign workers" item, higher—equal to 1.61—for the group assigned the "foreign workers" item), while the discrimination parameters showed similar values (respectively 2.95 and 2.79). Overall, when comparing the less constrained and the more constrained models, the likelihood ratio test suggested a significant difference between them (p-value = 0.000). Therefore, the significance

[4]The comparison between the two-parameter and the one-parameter IRT model also provided the same result among the other two groups of respondents ("foreign workers" group: likelihood ratio chi-square statistic (7) = 32.65, p-value = 0.000; "immigrants" group: likelihood ratio chi-square statistic (7) = 30.55, p-value = 0.000). Moreover, three-parameter IRT models are not appropriate as the response format of the items measuring social distance is not right/wrong, thus the answer does not involve any guessing (Tay et al., 2015).

Table 1

Two-group ("immigrants/foreign workers" group vs. "foreign workers" group) IRT models for the social distance items

		Constrained model		Augmented model	
Item	Parameter	Immigrants/ Foreign workers	Foreign workers	Immigrants/ Foreign workers	Foreign workers
Gypsies	Difficulty	−0.55	-	−0.55	-
	Discrimination	1.09	-	1.08	-
Drug addicted	Difficulty	−0.53	-	−0.53	-
	Discrimination	1.34	-	1.32	-
Heavy drinkers	Difficulty	0.21	-	0.22	-
	Discrimination	1.48	-	1.46	-
Muslims	Difficulty	1.03	-	1.04	-
	Discrimination	2.99	-	2.96	-
People of a different race	Difficulty	1.39	-	1.40	-
	Discrimination	3.23	-	3.25	-
Homosexuals	Difficulty	1.57	-	1.58	-
	Discrimination	2.10	-	2.10	-
Jews	Difficulty	1.69	-	1.70	-
	Discrimination	3.10	-	3.10	-
Manipulated item	Difficulty	1.18	-	1.12	1.61
	Discrimination	2.88	-	2.95	2.79
Likelihood ratio test			20.83		
p-value			0.000		
N			1,920		

test indicated that the constraints (same parameters of difficulty and discrimination for the manipulated item across the two groups) were not supported by the data; accordingly, it could be concluded that in the manipulated item there was DIF between the group of respondents answering the "immigrants/foreign workers" item and the group answering the "foreign workers" item.

Table 2 instead shows the results of two two-group 2PL IRT models where the two groups of respondents were respectively the ones assigned the "immigrants/foreign workers" item and the ones assigned the "immigrants" item. As for the manipulated item, the parameter of difficulty was very similar across the two groups (respectively equal to 1.12 and 0.99) in the less constrained model, while the parameter of discrimination was higher for the group of respondents assigned the "immigrants" item—at 4.38—compared to the "immigrants/foreign workers" item—at 2.92—despite being very high for both groups. Nonetheless, the overall comparison between the more constrained and the less constrained models suggested there was no significant difference between the models (p-value of the likelihood ratio test = 0.27). In other words, the data provided evidence toward the absence of DIF on the manipulated item across the two groups; hence, the functioning of the "immigrants" item was not significantly different from the functioning of the "immigrants/foreign workers" item when related to the other social distance items.

Besides analyzing the functioning of different formulations of the item among the whole sample, another goal was to check whether the results hid heterogeneity across different socio-demographic and ideological subgroups. To this end, we estimated five different logistic regression models in which the dependent variable was the dichotomous answer to the social distance item and the main independent variable the interaction term between the experimental condition and each of the following variables—gender, age, level of education, geographical area, and left-right orientation—by controlling for the remaining four variables (see Table A2 in the Appendix). These models allowed us to test whether the differences between the proportions of people declaring that they did not want foreign workers (immigrants) as neighbors and the proportion of people not wanting "immigrants/foreign workers" varied depending on a third variable. Table A3 in the Appendix shows that the differences in the

Table 2

Two-group ("immigrants/foreign workers" group vs "immigrants" group) IRT models for the social distance items

		Constrained model		Augmented model	
Item	Parameter	Immigrants/ Foreign workers	Immigrants	Immigrants/ Foreign workers	Immigrants
Gypsies	Difficulty	−0.52	-	−0.52	-
	Discrimination	1.17	-	1.08	-
Drug addicted	Difficulty	−0.52	-	−0.53	-
	Discrimination	1.35	-	1.32	-
Heavy drinkers	Difficulty	1.39	-	0.23	-
	Discrimination	1.48	-	1.46	-
Muslims	Difficulty	1.02	-	1.02	-
	Discrimination	3.05	-	2.96	-
People of a different race	Difficulty	1.44	-	1.43	-
	Discrimination	3.01	-	3.25	-
Homosexuals	Difficulty	1.66	-	1.65	-
	Discrimination	1.99	-	2.1	-
Jews	Difficulty	1.72	-	1.72	-
	Discrimination	2.88	-	3.1	-
Manipulated item	Difficulty	1.10	-	1.13	0.99
	Discrimination	3.07	-	2.92	4.38
Likelihood ratio test			2.63		
p-value			0.270		
N			1,954		

distribution of answers to the different formulations of the item observed on the whole sample were similar across the categories of variables considered. For almost all the subcategories, there was a significantly lower social distance toward foreign workers when compared to 'immigrants/foreign workers.' The difference was not statistically significant only for rightist people, but the pattern was similar to all the other subgroups. In addition, analogously to the whole sample, the difference in the percentages of respondents not wanting respectively "immigrants" and "immigrants/foreign workers" as neighbors was not significantly different (at the 0.05 level) from zero for all the categories of gender, age, geographical area, left-right position, and level of education. We highlight, however, that at the 0.10 level of significance, the level of social distance toward immigrants among less-educated people was statistically higher than the level of social distance toward "immigrants/foreign workers".

6 Discussion and conclusions

Ambiguity in the item wording does not necessarily mean ambiguity in the respondents' comprehension. This is the main result of the experiment on the wording of the "immigrants/foreign workers" item of the social distance scale in the Italian EVS-WVS 2017 survey. The terms "immigrants" and "foreign workers" refer to different groups of individuals, as foreign workers are a subcategory of immigrants, and people attach different meanings to them. Our evidence clearly shows that Italians perceive a much larger distance toward immigrants than toward foreign workers. Indeed, the two items result non-equivalent in the survey experiment. Nevertheless, when the two terms are put together, the semantic prevalence of "immigrants" over "foreign workers" is such that the second term disappears.

We can also argue that the conclusions provided here could be generalized to almost every subcategory of respondents. In this regard, future work is expected to consolidate our results, which could be hindered by the relatively small sample size of the experimental groups.

All in all, the results of the experiment sound like good news for EVS, WVS, and other surveys using this item of the social distance scale. It is quite clear that people react to the "immigrants" and "immigrants/foreign workers" item in a similar way; accordingly, the "foreign workers" term could in effect be dropped without losing comparability over time.

Yet, this does not completely solve the ambiguity issue. Which groups people refer to when they answer this question remains an open issue. It is nevertheless beyond the aim of this paper to address this issue, which should be further investigated using probing questions. However, as it is very likely that people living in different contexts give different meanings to the "immigrants/foreign workers" item, further research is needed to address the issue of comparability across countries by assessing the measurement equivalence of the "immigrants/foreign workers" item within the social distance scale in values surveys.

As a limitation of this study, we should mention that the differential item functioning analysis assumes a unidimensional scale of social distance (Dunn & Singh, 2011; Kirchner et al., 2011; Rapp & Ackermann, 2016). Other studies found two dimensions, one related to an ethnic and religious factor and one related to deviant behavior (Peral & Ramos, 2014), although the measurement equivalence of the multidimensionality of the scale was not explicitly tested across countries. As a result, our analysis did not consider the possible multidimensionality of the scale. Hence, it cannot be taken for granted that the results of our experiment will hold when changing the dimensional structure of the scale.

Last but not least, we would like to highlight that the survey experiment was administered to a probabilistic sample through face-to-face interviews. This is a rare exception in survey research, in which most experimental designs are carried out in CAWI surveys administered to non-probabilistic samples. Together with the experimental manipulation of a single word in the text of the item, the adoption of a probabilistic sample representative of the Italian population guarantees that the results provided here are both internally and externally valid. *Stricto sensu*, the external validity is limited to the Italian case. As Italy is a country with a recent history of immigration, it would be useful to extend this experimental design to countries with different immigration patterns, to assess whether these results hold across different contexts. However, previous research mentioned has shown that even in contexts with a long history of immigration, such as Great Britain, the proportion of people associating immigrants with foreign workers is minoritarian (Blinder, 2015). Therefore, we can hypothesize that the results presented here could be generalized to a certain extent to other countries.

However, the implementation of our survey experiment did not come without a cost, as the inclusion of the experimental design led to the loss of about 700 cases on a single item in the comparative EVS and WVS surveys. Nevertheless, since the main aim of our study was to accurately measure the effect of a change in the formulation of the item wording on individuals' answers, we needed to employ the same setting as the EVS-WVS survey. Furthermore, we expect that the small price of a few hundred cases will be outweighed by the great benefits for the scientific community of values survey users, especially those who have or intend to analyze the more than 600,000 observations on the social distance toward immigrant/foreign workers in values surveys.

Acknowledgement

This publication has benefited from a research stay at GESIS—Leibniz Institute for the Social Sciences and was financially supported by GESIS research grant EL-2019-175.

References

Asbrock, F., Lemmer, G., Becker, J. C., Koller, J., & Wagner, U. (2014). "Who are these foreigners anyway?" The content of the term foreigner and its impact on prejudice. *Sage Open*, *4*(2), 2158244014532819.

Blinder, S. (2015). Imagined immigration: The impact of different meanings of "immigrants" in public opinion and policy debates in Britain. *Political Studies*, *63*(1), 80–100.

Bogardus, E. S. (1925). Measuring social distance. *Journal of applied sociology*, *9*, 299–308.

Bogardus, E. S. (1933). A social distance scale. *Sociology & Social Research*, *17*, 265–271.

Braun, M., Behr, D., & Kaczmirek, L. (2013). Assessing cross-national equivalence of measures of xenophobia: Evidence from probing in web surveys. *International Journal of Public Opinion Research*, *25*(3), 383–395.

Colombo, M. (2018). The representation of the "european refugee crisis" in Italy: Domopolitics, securitization, and humanitarian communication in political and media discourses. *Journal of immigrant & refugee studies*, *16*(1-2), 161–178.

Doebler, S. (2014). Relationships between religion and intolerance towards muslims and immigrants in Europe: A multilevel analysis. *Review of religious research*, *56*(1), 61–86.

Drasgow, F. (1984). Scrutinizing psychological tests: Measurement equivalence and equivalent relations with external variables are the central issues. *Psychological bulletin*, *95*(1), 134–135.

Dunn, K. P., & Singh, S. P. (2011). The surprising non-impact of radical right-wing populist party representation on public tolerance of minorities. *Journal of Elections, Public Opinion & Parties*, *21*(3), 313–331.

Evans, M., & Kelley, J. (2019). Prejudice against immigrants symptomizes a larger syndrome, is strongly diminished by socioeconomic development, and the uk is not an outlier: Insights from the WVS, EVS, and EQLS surveys. *Frontiers in Sociology*, 12.

Georgiou, M., & Zaborowski, R. (2017). *Media coverage of the "refugee crisis": A cross-European perspective*. Council of Europe.

Haney, J. L. (2016). Predictors of homonegativity in the United States and the Netherlands using the fifth wave of the World Values Survey. *Journal of homosexuality*, *63*(10), 1355–1377.

Kirchner, A., Freitag, M., & Rapp, C. (2011). Crafting tolerance: The role of political institutions in a comparative perspective. *European Political Science Review*, *3*(2), 201–227.

Park, K. (2014). Foreigners or multicultural citizens? Press media's construction of immigrants in South Korea. *Ethnic and Racial Studies*, *37*(9), 1565–1586.

Park, R. E. (1924). The concept of social distance: As applied to the study of racial relations. *Journal of applied sociology*, *8*, 339–334.

Peral, E. B., & Ramos, A. (2014). 6. neighbours: Determinants of whom Europeans want to keep at a distance. In W. Arts & L. Halman (Eds.), *Value contrasts and consensus in present-day Europe. painting Europe's moral landscapes* (pp. 117–141). Brill.

Quassoli, F. (2013). "Clandestino": Institutional discourses and practices for the control and exclusion of migrants in contemporary Italy. *Journal of Language and Politics*, *12*(2), 203–225.

Rapp, C., & Ackermann, K. (2016). The consequences of social intolerance on non-violent protest. *European Political Science Review*, *8*(4), 567–588.

Ruedin, D. (2019). Attitudes to immigrants in South Africa: Personality and vulnerability. *Journal of Ethnic and Migration Studies*, *45*(7), 1108–1126.

Shin, H., & Dovidio, J. F. (2017). Cultural differences in the role of economic competitiveness in prejudice toward immigrants and foreign workers. *Analyses of Social Issues and Public Policy*, *17*(1), 7–32.

Shin, H., & Dovidio, J. F. (2018). Differences, threats, values, and country-specific prejudice toward immigrants and foreign workers in three major receiving countries: The United States, Germany, and Australia. *Journal of Social Issues*, *74*(4), 737–755.

Spruyt, B., van der Noll, J., & Vandenbossche, L. (2016). Meaning matters. an empirical analysis into public denotations of the label "strangers" and their relationship with general ethnic prejudice. *International Journal of Intercultural Relations*, *51*, 41–53.

Strabac, Z., Listhaug, O., & Jakobsen, T. G. (2012). Patterns of ethnic intolerance in Europe. *Journal of International Migration and Integration*, *13*, 459–479.

Tay, L., Meade, A. W., & Cao, M. (2015). An overview and practical guide to IRT measurement equivalence analysis. *Organizational Research Methods*, *18*(1), 3–46.

Vezzoni, C. (2018). Immigrazione ed insicurezza economica nelle urne. In Itanes (Ed.), *Vox populi: Il voto ad alta voce del 2018* (pp. 147–163). il Mulino.

Wark, C., & Galliher, J. F. (2007). Emory bogardus and the origins of the social distance scale. *The American Sociologist*, *38*, 383–395.

Appendix
Tables

Table A1

Randomization checks: percentage distribution of socio-demographics and left-right orientation by item formulation ($N = 2,277$) and chi-square tests of independence

Variables	Experemental condition: item formulation			P-value Chi2
	Immigrants/Foreign workers ($n = 1,596$)	Foreign workers ($n = 324$)	Immigrants ($n = 357$)	
Gender				
Male	51.4	45.1	49.6	0.11
Female	48.6	54.9	50.4	
Level of education				
Low	39.0	39.1	36.9	
Medium	45.3	45.0	45.1	0.86
High	15.7	15.9	18.0	
Age				
18-34	21.5	19.4	19.3	
35-54	33.9	36.1	31.9	0.55
55 and over	44.6	44.4	48.7	
Geographical area				
North	45.9	43.8	46.2	
Center	18.9	18.2	14.9	0.34
South	35.1	38.0	38.7	
Left-right orientation				
Left (1-3)	15.6	10.5	12.0	
Center-left (4-5)	20.2	23.2	21.6	
Center-right (6-7)	22.1	22.2	22.4	0.51
Right (8-10)	15.4	16.7	15.1	
DK/DA/Do not locate	27.1	27.5	28.9	

Table A2

Logistic regression models in which the dependent variable is the dichotomous answer to the manipulated item of social distance

Variables	Categories	Model 1: Gender est.	SE	Model 2: Age est.	SE	Model 3: Edu est.	SE	Model 4: Area est.	SE	Model 5: L-R est.	SE
Experimental condition (Immigrants/Foreign Workers)	Foreign workers	−1.03***	0.35	−1.09*	0.62	−0.71**	0.30	−0.67*	0.39	−0.97	0.76
	Immigrants	−0.10	0.23	0.30	0.37	0.41*	0.22	0.32	0.26	0.11	0.49
Gender (Male)	Female	0.00	0.14	0.09	0.12	0.09	0.12	0.09	0.12	0.09	0.12
Age (18-34)	35-54	0.12	0.18	0.22	0.20	0.12	0.18	0.13	0.18	0.11	0.18
	55 and over	0.34*	0.17	0.33*	0.20	0.35**	0.17	0.35**	0.17	0.35**	0.17
Level of education (Low)	Medium	−0.49***	0.14	−0.49***	0.14	−0.35**	0.16	−0.49***	0.14	−0.50***	0.14
	High	−0.59***	0.20	−0.59***	0.20	−0.45**	0.23	−0.59***	0.20	−0.59***	0.20
Geographical area (North)	Center	0.46***	0.18	0.46***	0.18	0.47***	0.18	0.56***	0.20	0.47***	0.18
	South	0.98***	0.14	1.00***	0.14	0.98***	0.14	1.06***	0.16	0.99***	0.14
Left-right position (Left:1-3)	Center-Left (4-5)	−0.21	0.23	−0.20	0.23	−0.20	0.23	−0.22	0.23	−0.28	0.27
	Center-Right (6-7)	0.34	0.21	0.34	0.21	0.35	0.22	0.34	0.22	0.34	0.24
	Center-Right (8-10)	0.56**	0.22	0.56**	0.22	0.57**	0.22	0.55**	0.22	0.43*	0.26
	DK/DA	0.12	0.21	0.12	0.21	0.13	0.21	0.12	0.21	0.22	0.24
Experimental condition#Gender	Foreign workers#Female	0.22	0.45								
	Immigrants#Female	0.39	0.32								
Experimental condition#Age	Foreign workers#35-54			−0.07	0.73						
	Foreign workers#55 and over			0.39	0.69						
	Immigrants#35-54			−0.56	0.48						
	Immigrants#55 and over			−0.07	0.43						

Continues on next page

Continued from last page

Variables	Categories	Model 1: Gender est.	SE	Model 2: Age est.	SE	Model 3: Edu est.	SE	Model 4: Area est.	SE	Model 5: L-R est.	SE
Experimental condition#Level of education	Foreign workers#Medium					−0.53	0.50				
	Foreign workers#High					−0.15	0.70				
	Immigrants#Medium					−0.58*	0.35				
	Immigrants#High					−0.69	0.52				
Experimental condition#Geographical area	Foreign workers#Center							−1.03	0.83		
	Foreign workers#South							−0.20	0.49		
	Immigrants#Center							−0.20	0.48		
	Immigrants#South							−0.36	0.35		
Experimental condition#Left-right position	Foreign workers#Center-Left									−0.12	0.98
	Foreign workers#Center-Right									−0.38	0.93
	Foreign workers#Right									0.68	0.86
	Foreign workers#DK/DA									−0.07	0.87
	Immigrants#Center-Left									0.40	0.61
	Immigrants#Center-Right									0.13	0.58
	Immigrants#Right									0.30	0.61
	Immigrants#DK/DA									−0.53	0.58
	Constant	−2.21***	0.27	−2.29***	0.28	−2.34***	0.27	−2.32***	0.27	−2.25***	0.28
	Observations	2,264		2,264		2,264					

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Table A3

Estimated differences in proportions of people not wanting to have foreign workers (immigrants) as neighbors and people not wanting immigrants/foreign workers as neighbors. Data estimated from logistic regressions shown in Table A2

		Pr(Foreign workers) - Pr(Immigrants/Foreign workers)		Pr(Immigrants) - Pr(Immigrants/Foreign workers)	
Variables	Categories	est.	SE	est.	SE
Gender	Male	−0.10***	0.02	−0.01	0.03
	Female	−0.08***	0.02	0.04	0.03
Age	18-34	−0.08**	0.03	0.05	0.05
	35-54	−0.10***	0.03	0.04	0.04
	55 and over	−0.08***	0.03	0.03	0.03
Level of education	Low	−0.09***	0.03	0.07*	0.04
	Medium	−0.10***	0.02	−0.02	0.03
	High	−0.07*	0.04	−0.03	0.05
Geographical area	North	−0.05**	0.02	0.03	0.03
	Center	−0.13***	0.03	0.02	0.06
	South	−0.12***	0.03	−0.01	0.04
Left-right orientation	Left (1-3)	−0.08*	−0.05	0.01	0.06
	Center-left (4-5)	−0.07**	−0.03	0.06	0.05
	Center-right (6-7)	−0.13***	0.04	0.04	0.05
	Right (8-10)	−0.04	0.05	0.07	0.07
	DK/DA/Do not locate	−0.10***	0.03	−0.05	−0.03

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$