methods, data, analyses | Vol. 18(2), 2024, pp. 213-248 DOI: 10.12758/mda.2024.04

Measurement Invariance and Quality of Attitudes Towards Immigration in the European Social Survey

Amelie Nickel[1] & Wiebke Weber[2]

[1] *Bielefeld University, IKG, Germany*

[2] *LMU Munich, Germany & RECSM, Universitat Pompeu Fabra, Spain*

Abstract

The number of studies assessing measurement invariance of the European Social Survey's (ESS) immigration scale increased in recent years. However, the comparability of findings is limited due to the lack of consistency in the analytic strategies and methods employed across these studies. The present study aims to address this issue by employing a consistent approach: a multigroup confirmatory factor analysis (MGCFA), to test for measurement invariance of attitudes towards immigration in each of the first nine rounds of the ESS. Moreover, we estimate the measurement quality by computing the reliability coefficient Omega in each country in each round of the ESS.

Our results reveal that metric invariance holds for all countries but one (Finland) in all rounds, indicating that covariances and regression coefficients can be compared meaningfully. While scalar invariance only holds for different subgroups of countries within each round, partial invariance is fulfilled in all countries, meaning that at least one indicator is equal for all countries allowing for latent mean comparisons. Furthermore, assessing the measurement quality, we find the attitudes towards immigration index similarly good across the different countries and rounds.

Keywords: Measurement invariance, attitudes towards immigration, European Social Survey, MGCFA, Measurement quality

Undoubtedly, the topic of migration will largely shape the national and international political agenda of the 21st century. In election and public debates, the question of how to deal with migration is one of the most pressing concern, effectively capitalized on by the political right. The so-called 'refugee crisis' in 2015 has deepened cleavages within the European Union, providing an opportunity for populist radical right parties to advocate for more restrictive policies and shift the overall political discourse to the right (Mudde, 2007, 2020).

Given its ongoing social and political relevance, understanding and analyzing attitudes toward immigration has emerged as one of the most extensively studied aspects of the social sciences (Bohman, 2015; Borgonovi & Pokropek, 2019; Quillian, 1995; Scheepers et al., 2002; Weldon, 2006). This has resulted in extensive literature from various disciplines, such as sociology, psychology, political science, and economics. So far, empirical studies have mainly focused on the individual level, but with the increasing availability of large cross-national datasets, the amount of international comparative research is rising (Meuleman & Billiet, 2012).

Measuring psychological constructs such as values or attitudes across countries raises methodological questions on the comparability of measurements that are often insufficiently or not at all addressed by researchers (Davidov & Meuleman, 2012; Meitinger et al., 2020; Roots et al., 2016). As question wording and items can have different meanings in different countries depending on the linguistic and cultural background, it is essential to verify and ensure that the used measurements are comparable across the observed countries (Roots et al., 2016). According to Meuleman et al. (2022, p.3), the basic idea behind so-called measurement invariance testing (also referred to as measurement equivalence) of multi-item instruments in cross-cultural research is that "when we compare any measurement across groups, that comparison should reflect true differences rather than measurement differences."

The lack of testing measurement comparability is increasingly criticized in the literature as it may lead to misinterpretation of findings (Meuleman & Billiet, 2012). However, due to improved and new statistical techniques, measurement invariance testing has become more accepted in applied social science research over the last decade (Davidov, Muthen, & Schmidt, 2018; Leitgöb et al., 2023).

Acknowledgements
We extend our sincere gratitude to Carlos Poses for his contribution to the analysis of the paper, and to the anonymous reviewers who provided valuable feedback, which has significantly enhanced the quality of our work.

Direct correspondence to
Amelie Nickel, Bielefeld University, IKG, Germany
E-mail: amelie.nickel@uni-bielefeld.de

When testing measurement invariance we can at first decide between two traditions in measurement theory: Item-response theory (IRT) or structural equation modeling (SEM) (Bauer et al., 2006; Putnick & Bornstein, 2016; Tsaousis et al., 2020)[1].

IRT examines the relationship between an individual's latent trait (e.g., an attitude) and their response to a specific item. In the IRT tradition, measurement invariance is assessed through the lens of differential item functioning (DIF; Holland & Wainer, 2015) across groups, which determines whether item behavior measures equivalent levels of the latent trait across members of different groups (Tsaousis et al., 2020).

We focus on SEM approaches, namely confirmatory factor analysis (CFA) and multi-group confirmatory factor analysis (MGCFA), in which the relations between observed variables and latent construct(s) are tested for measurement invariance between groups (Vandenberg & Lance, 2000). CFA is a statistical technique used to test the invariance of measurement model parameters within subpopulations, while MGCFA is an extended version of CFA that allows invariance testing across multiple groups. MGCFA (Jöreskog, 1971; Millsap, 2011) is most widely used for testing measurement invariance. However, the scientific community is inconsistent about the correct methods as well as the usefulness of measurement invariance testing in general, as a recent debate in *Sociological Methods & Research* shows (Fischer et al., 2022; Meuleman et al., 2022; Welzel et al., 2021; Welzel et al., 2022). In the concluding section of our paper, we outline the advantages and main limitations of MGCFA and highlight some recently developed alternative methods.

In our study, we aim to contribute to the field by employing multigroup confirmatory factor analysis with local fit testing to assess measurement invariance of attitudes towards immigration as measured in each round of the ESS (ESS R1 (2002) to ESS R9 (2018))[2].

In order to make meaningful cross-country comparisons, it is essential to check not only that the measures are comparable across countries, but also that the quality of the measures is comparable. Testing measurement quality is an imperative to correct for measurement errors (Pirralha & Weber, 2020; Poses et al., 2021; Saris & Revilla, 2016). For meaningful comparisons (e.g. correlations), it is crucial that the size of the measurement errors is similar between the groups (e.g. countries) being compared. In general, "the lower the quality of measurement, the more careful researchers need to be in their conclusions [...],

1 For recent efforts to combine the two approaches, see Raju et al. (2002); Reise et al. (1993); Stark et al. (2006); Widaman and Grimm (2014) quoted from Putnick and Bornstein (2016).

2 It was not possible to include the most recent round from the European Social Survey (round 10, conducted in 2020) due to the timeframe of the study. In addition, the data collection for round 10 of the ESS took place during the COVID-19 pandemic, which implied unique circumstances such as online interviews and self-completion of questionnaires.

since higher levels of measurement errors are more likely to disturb the results" (Pirralha & Weber, 2020; Poses et al., 2021, p. 245). Therefore, we also estimate the measurement quality by calculating the reliability coefficient Omega (Hayes & Coutts, 2020) of the sum score of attitudes towards immigration for each country in each round of the ESS.

We acknowledge the growing number of studies that have assessed the measurement invariance of the ESS immigration scale in recent years. However, the comparability of findings across these studies is limited due to the lack of consistency in the analytical strategies and methods used. In our study, we aim to enhance comparability and provide more reliable insights by adopting a constituent approach. In addition, we aim to contribute to the field by providing an accessible and reader-friendly introduction to MGCFA as a method for testing measurement invariance, which may enhance its practical application in the context of migration research.

This paper sets out by introducing the European Social Survey (ESS) as a data source for studying attitudes towards immigration. We then provide a comprehensive introduction to MGCFA and present an overview of previous research testing the comparability of attitudes towards immigration in the ESS. The following section outlines the present study – sample, model testing, and analytic strategy. Finally, the results of measurement invariance testing, latent means comparison, and measurement quality assessment are presented and discussed.

Attitudes Towards Immigration in the European Social Survey

The European Social Survey is a biannual cross-national survey aimed to track Europeans' attitudes, beliefs, and behaviors on different topics. Implemented in most European countries, the ESS is a cross-sectional, probability-based sample in which all individuals, residents in private households over the age of 15, are eligible.

Since its first round in 2002, the European Social Survey (ESS) has continuously surveyed attitudes towards immigration in several European countries and is thus one of the most widely used surveys for cross-national research on attitudes towards immigration (Roots et al., 2016). Across each round, it includes several items to assess attitudes towards immigration in its main questionnaire. Besides, in rounds 1 and 7, the ESS conducted a more comprehensive immigration module that specifically focused on various dimensions of attitudes towards immigration (Heath et al., 2016).

In this paper, we focus on three items measuring the concept attitudes towards immigration, included in the core module, and displayed in Table 1.

Table 1 Items used for measuring attitudes towards immigration (ATI)

Question wording	Item name	Item number	Response scale
Would you say it is generally bad or good for [country]'s economy that people come to live here from other countries?	*imbgeco*	B41	0 (Bad for the economy) – 10 (Good for the economy)
And, using this card, would you say that [country]'s cultural life is generally undermined or enriched by people coming to live here from other countries?	*imueclt*	B42	0 (Cultural life undermined) – 10 (Cultural life enriched)
Is [country] made a worse or a better place to live by people coming to live here from other countries?	*imwbcnt*	B43	0 (Worse place to live) – 10 (Better place to live)

Invariance Testing with Multigroup Confirmatory Factor Analysis (MGCFA)

Measurement invariance tests rely on a latent variable approach. As a confirmatory factor analysis model, multigroup confirmatory factor analysis (MGCFA) techniques assume that the responses people provide to different items (observed responses) are caused by their position on an unobserved construct or factor (latent variable). Figure 1 represents this model for the latent factor attitudes towards immigration (ATI) that determines the answers to the three items in Table 1.

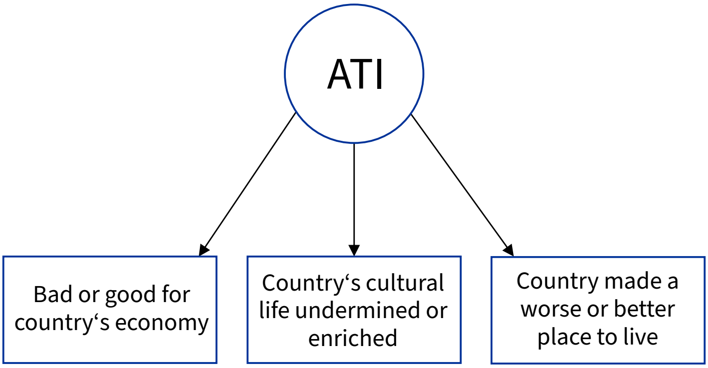

Figure 1 Measurement model for the latent factor ATI

Equation 1 provides an equation for the same model, where X_i is the observed item answer for the observed variable i, ξ is the latent factor ATI, and ε is the error term, i.e., the item variance unaccounted by the latent factor. τ represents the intercept (the expected value of each observed item when the value of the latent variable is zero), and λ_i is the loading (the expected increase in X_i for each one unit increase in ξ).

$$X_i = \tau_i + \lambda_i\xi + \varepsilon_i \tag{1}$$

The multigroup extension of MGCFA implies that this same measurement model is separately estimated in different groups, indicated with the subscript j, as depicted in Equation 2.

$$X_{ij} = \tau_{ij} + \lambda_{ij}\xi + \varepsilon_{ij} \tag{2}$$

Declaring measurement invariance implies asserting that the factor is measured in the same way across the different groups. Thus, in this framework, measurement invariance means that the parameters of the measurement model τ_i and λ_i are equal across all groups j. There are different possible equalities between parameters that can be satisfied across groups, giving rise to different measurement invariance levels.

First, *configural* invariance means that the general structure of the factor is equal across groups, i.e., that the same items load on the same factor(s). Since the model we estimate is unifactorial, this simply means that the loadings for none of the items are zero in any group and that there are no correlated error terms in only some groups. Whether the value of the parameters is equal across groups is not important at this level. The establishment of configural invariance is interpreted as evidence suggesting that, since the factor can be measured with the same items in all groups, the factor has a similar theoretical content across groups.

Second, *metric* (also known as *loading*) invariance indicates that the factor loadings are equal across groups. This means that a one unit increase in the factor leads to the same change in the observed item responses in all groups. This level of equivalence implies that factor variances and covariances (i.e., the relationship of the factor with other measures) can be compared meaningfully across groups.

Third, *scalar* (also known as *intercept*) invariance means that the item intercepts are equal across groups. This indicates that when the value of the latent variable is zero, the expected mean value of the item responses will be the same across all groups. Importantly, when metric and scalar invariance for an item are established across groups, it means that any given level in the latent variable of interest will lead to the same expected value of the observed item. Therefore, the simultaneous establishment of metric and scalar invariance allows for the comparison of sum scores or observed means.

To understand this last point, it is important to highlight the differences between sum scores, observed means, and latent means. Sum scores are the scores produced by simply summing, for each individual on the sample, the scores of all items that measure the latent factor. Observed means refer to the means of sum scores across all individuals in a given country. Latent means are the means of the latent factor ATI. While observed means are very simple to compute, latent means need to be estimated with specialized software and using structural equation modeling (or other latent variable) techniques. From Equation 1, we can infer the reason why observed means should not be compared in the absence of measurement invariance. If the loadings and intercepts are not equal across groups, the same level in the latent factor will lead to different expected values in the observed items. This implies, for instance, that the same mean level in an observed item across groups may correspond to different mean levels of the latent variable; or that different mean levels of an observed item across groups might correspond to the same mean value of the latent variable (see also Steinmetz, 2013). Said differently, if the loadings and intercepts are not equal across groups, the correspondence between latent means and observed means differ across groups. This means that differences in observed means are not trustworthy indicators of differences in latent means. Thus, observed means should not be compared because they do not necessarily reflect differences in latent means.

The situation of equality of all loadings and intercepts across groups is sometimes called *full invariance*. Full invariance has often been found to be too strict to achieve, especially for the intercepts (Davidov, Muthen, & Schmidt, 2018). This means that comparisons of observed means (or sum scores) across groups cannot be guaranteed, as differences in observed means might or might not reflect true differences in the latent means. A way to overcome this issue is to compare latent means instead, which can be done by establishing *partial invariance.* Partial invariance implies that only some of the parameters of the measurement model are equal across groups. Classic advice has been that latent means can be compared in situations with partial invariance when at least two of the loadings and intercepts are equal (Steenkamp & Baumgartner, 1998). More recent, simulations by Pokropek et al. (2019) have shown that the estimation of latent means is satisfactory in partial measurement invariance models, if items with partial measurement invariance are identified and freed, and that at least one item is invariant across groups.

Therefore, our aim is twofold. First, testing for measurement invariance of the attitudes towards immigration scale across countries and establishing the level of invariance (configural, metric, or scalar) that holds across each group of countries. Second, establishing partial invariance for cases where no full invariance is found, so that latent means can be compared.

Previous Research Testing the Comparability of Attitudes Towards Immigration in the European Social Survey

The number of studies evaluating measurement invariance of the immigration attitudes scale across the ESS countries and over time has increased but remains limited (Table 2). These studies differ in their methodological approaches, analytical strategies, and terminology used: “perceived ethnic threat” (Pirralha & Weber, 2020), “attitudes towards migration” (Borgonovi & Pokropek, 2019), “anti-immigrant attitudes” (Nickel, 2022). In the following, we will use the term “attitudes towards immigration” (ATI). For a detailed overview of the constructs, questions, and response scales used in the studies, see Table A1, Appendix.

Table 2 Studies testing the comparability of attitudes towards immigration (ATI) in ESS[3]

Study	Round	ATI	Countries	Method	Results
Cross-time and cross-national measurement invariance					
Meuleman, Davidov & Billiet 2009	ESS R1 (2002) – ESS R3 (2006)	REJECT	17	Multigroup confirmatory factor analysis (MGCFA)	Full scalar invariance within 17 countries; partial scalar invariance between 17 countries
Borgonovi & Pokropek 2019	ESS R5 (2010) – ESS R8 (2016)	REJECT THREAT	18	Multigroup Bayesian Structural Equation Modeling (MG-BSEM)	Full scalar invariance for each country over time; Metric invariance between countries
Cross-national measurement invariance					
Meuleman & Billiet 2012	ESS R1 (2002)	REJECT CONDITION ECOTHREAT CULTTHREAT	21	MGCFA	REJCET: Partial scalar invariance in all 21 countries; CULTTHREAT: Partial scalar invariance in 11 countries; CONDITION + ECOTHREAT: Partial scalar invariance in 14 countries
Davidov et al. 2015	ESS R1 (2002) – ESS R6 (2012)	REJECT	35	Approximate measurement invariance using Bayesian estimation	Approximate scalar invariance across all countries in each round
Davidov, Cieciuch & Schmidt 2018	ESS R7 (2014)	ALLOWANCE CONDITION RT (Realistic Threat)	15	Approximate measurement invariance using Bayesian estimation; Exact measurement invariance by MGCFA	ALLOWANCE: Approximate scalar invariance in 12 countries; RT: approx. scalar invariance in 13/14 countries; CONDITION: metric invariance in 7 countries
Pirralha & Weber 2020	ESS R3 (2006)	THREAT	19	MGCFA + correction for measurement error	Partial scalar invariance in 19 countries
Nickel 2022	ESS R9 (2018)	THREAT	28	MGCFA	Metric invariance in 29 countries

3 Without any claim to completeness

Studies testing cross-time and cross-national measurement invariance

Meuleman et al. (2009) first started testing the comparability of the ESS immigration attitudes scale across three time points (ESS R1 (2002); ESS R2 (2004); ESS R3 (2006)). The authors provide technical guidance on how to measure scale invariance by applying multigroup confirmatory factor analysis (MGCFA) and using a top-down strategy: testing the most constrained model (full scalar invariance across time and countries) at first and then incrementally reducing the number of constraints assessing whether the model fit is improving. To measure ATI they construct a latent factor that measures the rejection of further immigration in general (REJECT). In their final model, full scalar invariance holds over time within the 17 countries and partial scalar invariance between the countries, implying that the ESS immigration attitudes can be meaningfully compared across countries and over the three time points.

Borgonovi and Pokropek (2019) published a study examining the measurement invariance, both across countries and across time, of two latent constructs measuring immigration attitudes: generalized threat (THREAT) and opposition to migration (REJECT). They considered four time points ESS R5 (2010) – R8 (2016), including 18 countries that participated in each round of the ESS. To test for partial and approximate measurement invariance they apply sequential methods using the multigroup Bayesian structural equation modeling (MG-BSEM; B. Muthén & Asparouhov, 2012). First, they measure cross-time comparability separately for each country, and second, cross-national comparability for each time point. They establish full scalar invariance over time within each county but only metric invariance across the countries. Indicating that the country means can be compared meaningfully over time for each country but that the different country means cannot be compared to each other within one time point.

Studies testing cross-national measurement invariance

Making use of the first more comprehensive module assessing immigration attitudes conducted in ESS round 1 (2002), Meuleman and Billiet (2012) test for measurement invariance for four latent factors: opposition against new immigration (REJECT); support for imposing conditions to immigration (CONDITION); perceived economic threat (ECOTHREAT); perceived cultural threat (CULTTHREAT). The REJECT scale holds partial scalar invariance (invariance applies at least for two items per construct) for all 21 countries, which allows for cross-national mean comparisons. The other three scales hold partial metric invariance in 18 to 19 countries, guaranteeing the cross-national comparability of regression coefficients and covariances. Partial scalar invariance holds only for 11 (CULTHREAT) to 14 countries (CONDITION, ECOTHREAT), implying that the country means of these three scales can only be meaningfully compared in some of the countries.

Davidov et al. (2015) extend these findings by testing for approximate measurement invariance of the REJECT scale across 35 countries and the first 6 ESS rounds[4]. As the traditional (exact) approach failed to support scalar and even partial scalar measurement invariance, the authors test for approximate measurement invariance using the Bayesian framework (B. Muthén & Asparouhov, 2012; van de Schoot et al., 2013). This procedure "allows variance around the point estimates for the factor loadings and intercepts of the indicators" (Davidov et al., 2015, p. 261), whereas in the exact approach factor loadings and intercepts would be constrained to be exactly equal. Their findings reveal that approximate scalar measurement invariance is established across all countries in each ESS round, guaranteeing comparable country means.

Based on data from the second comprehensive immigration module surveyed in ESS round 7 (2014), Davidov, Cieciuch, and Schmidt (2018) test for approximate measurement invariance of three latent constructs: opposition towards immigration (ALLOWANCE); qualification for entry or exclusion (CONDITION); realistic threat (RT). Their results show that approximate (not exact) scalar invariance for ALLOWANCE (12 countries) and RT (13 to 14 countries) can be found in most of the 15 countries considered. For CONDITION, neither exact nor approximate invariance holds, and metric invariance is established only in 7 countries.

Pirralha and Weber (2020) disentangle the cognitive from the measurement part and correct for measurement errors. They refer to the concept of perceived ethnic threat (similar to THREAT) and find partial scalar invariance which allows comparing the latent means across all 19 countries that participated in the ESS R3 (2006).

Further evidence for metric invariance of anti-immigrant attitudes (similar to THREAT) can be found in Nickel (2022). Using MGCFA for structural modeling, the results show that metric invariance holds for all 29 countries participating in ESS round 9 (2018), indicating that factor loadings are equivalent across these countries.

While the above-mentioned studies use different methods and analytical strategies, making it difficult to compare their results, we follow the same approach here for all nine ESS rounds: multigroup confirmatory factor analysis (MGCFA). Moreover, we also estimate the measurement quality of the sum score attitudes towards immigration to quantify how strong the relationship between the latent variable of interest, attitudes towards immigration, and its observed measure is.

4 Measurement invariance was tested separately for each ESS round, the authors did not test for over-time comparability.

The Current Study

Sample

We focus on three items measuring the concept attitudes towards immigration, which are included in the core module and shown in Table 1. As these items are repeated in each round of the ESS, our analyses are based on data from round 1 (2002) to round 9 (2018)[5]. In total, we analyze data from 38 countries: Austria, Belgium, Switzerland, Czechia, Denmark, Spain, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Luxembourg, Netherlands, Norway, Poland, Portugal, Sweden, Slovenia, Estonia, Iceland, Slovakia, Turkey, Ukraine, Bulgaria, Cyprus, Russia, Croatia, Latvia, Romania, Lithuania, Albania, Kosovo, Montenegro, Serbia, and the United Kingdom. This led to a total sample of 390,276 individuals[6].

Model testing

In order to establish partial measurement invariance, we estimate models with equality constraints on the parameter among groups. We then use local fit testing to determine whether the imposed constraints are supported by the data. Local fit testing focuses on whether, in each group, each specific parameter of the model is misspecified. Concretely, we follow the local fit testing procedure suggested by Saris et al. (2009). This local fit testing procedure is based on a combination of the modification indices (approximating a significance test for the retrieval of one constraint), the expected parameter change when a constraint is relieved, and the power of the test to detect a misspecified parameter of a given effect size. The researcher must set the expected parameter change that they do consider to be a relevant misspecification: misspecifications lower than this size are not considered relevant and are thus ignored. Our criteria for the size of the misspecifications to be detected are 0.1 for the loadings, 0.15 for the intercepts[7],

5 European Social Survey Round 1 Data, (2002); Round 2 Data, (2004); Round 3 Data, (2006); Round 4 Data, (2008); Round 5 Data, (2010); Round 6 Data, (2012); Round 7 Data, (2014 // 2015); Round 8 Data, (2016); Round 9 Data, (2018)

6 To test for invariance across countries, we included only those cases where respondents provided answers to all three items. We employed listwise deletion, meaning that any case with missing data for any of the specified variables was excluded from the analysis. The item non-response varies over time and across countries but without a clear pattern. For detailed information on the sample size for each country in each ESS round, see Table A2, Appendix. We acknowledge that there are alternative methods for dealing with missing data, but follow the usual approach adapted by ESS Core Scientific Team (Zavala-Rojas & Saris, 2018; Revilla, 2012; Weber, 2011).

7 In practice, given the standard deviations of the items in each country, this corresponds to an unstandardized effect size of between 0.3 and 0.5 in all cases, meaning that we aim to detect misspecifications larger than 10% of the total length of the response scale

and 0.2 for correlated error terms (all in standardized metrics), based on suggestions by Saris et al., 2009).

This procedure contrasts with the typical approach of relying on global fit indices (statistics such as Chi-square or fit indices such as the comparative fit index (CFI) or the root mean square error of approximation (RMSEA)) and evaluating models as a whole. We avoid global fit indices for two reasons. First, because of some of their drawbacks reported in the literature (sensitivity to sample size and different model characteristics, unequal sensitivity to different model misspecifications; Groskurth et al., 2021; Saris et al., 2009). Second, because relying on global fit indices does not allow for fine-grained assessments of invariance in the context of measurement invariance. Concretely, the use of global fit indices does not allow for recovering complex patterns of invariance across groups, i.e., among 24 groups, different levels of invariance are likely to be present in different subgroups of countries – but this level of detail cannot be achieved using global fit indices. Moreover, using global fit indices does not allow for identifying invariant items, which is a prerequisite to then free the invariant items and establish partial invariance. This is critical because establishing partial invariance is important when we want to compare latent means under conditions where full invariance is not present.

Analytical Strategy

The analyses were conducted in R (R Core Team, 2021), using the packages lavaan (Rosseel, 2012) and semTools (Jorgensen et al., 2021)[8]. We used ML estimation because the items were 11-point scales and did not present important skewness or kurtosis. The Saris et al. (2009) approach to local fit testing was implemented using the function miPowerFit in semTools. Factor variances were identified using the fixed factor approach, with fixing variances at 1 and means at 0, unless equality constraints in the model allowed to free these specifications (Schroeders & Gnambs, 2020). Our analyses are cross-sectional and not longitu-

(11-points), and occasionally misspecifications larger than a percentage of the scale smaller than 10 %.

8 As van de Schoot et al. (2012) point out, measurement invariance testing is feasible with various structural equation modeling software programs. Lisrel (Jöreskog & Sörbom, 1996-2001) possesses the capability to handle categorical data, although it demands a proficiency in syntax and matrix algebra. AMOS (Arbuckle, 2007) is recognized for its user-friendly interface, but its capacity to handle categorical data is limited. Currently, Mplus (L. K. Muthén & Muthén, 2012) stands out as the most versatile program for measurement invariance testing, albeit requiring a proficiency in syntax. Additionally, Lavaan (Rosseel, 2012) and OpenMx (Boker et al., 2011), both open-source R packages in ongoing development, provide alternative options for measurement invariance testing, thereby enhancing the array of available tools in this domain.

dinal, i.e., we test the measurement invariance across countries in each round, but not across time for a given country.

For each round, the invariance test proceeds as follows. First, we estimate the configural model and check that no estimates are 0. Second, we estimate the loading invariance model – all loadings constrained to be equal – and test it using miPowerFit. If miPowerFit detects misspecified loadings, we free them and re-estimate the model. Each time, we free only one loading because model misspecifications are often related. We repeat this process until no misspecifications are present according to miPowerFit. Once a model with no misspecifications is reached, we compare the value of the freed loadings. This step is done to evaluate further comparability across subgroups of countries: it might be that some countries are non-invariant with respect to the majority of the groups but invariant among them. When freed loadings deviate in the same direction compared to most groups (e.g., the freed loadings of more than one group are higher than for the rest of the countries), we additionally constrain them to be equal to each other. We then re-estimate and test the model again; in the rare occasion that misspecifications reappear, we also free them one by one.

After establishing the highest possible level of metric invariance, we move on to scalar invariance. The process for scalar invariance is the same as for metric invariance. First, we constrain the intercepts to be equal – although in this step we do not constrain the intercepts for the groups and items for which metric invariance was not established. We then test the model using miPowerFit and free the intercepts one by one. In the end, we compare the value of the freed intercepts and set additional equality constraints among the freed intercepts with similar estimated values.

Measurement Quality

Estimating the measurement quality is essential to correct for measurement errors (Saris & Gallhofer, 2014), but also to understand how much of the concept of interest – attitudes towards immigration – is measured by the created sum score[9]. A perfect relationship would be 1 with no measurement errors present. The measurement quality of the unweighted sum scores (q_s^2) is defined as:

9 The survey quality predictor (SQP) database, developed by Saris et al (2011), serves as an open source tool for evaluating the quality of individual questions in the ESS (see https://www.europeansocialsurvey.org/methodology/ess-methodology/data-quality-assessment). Saris and Gallhofer (2014) suggest that SQP can also be used to assess the quality of composite scores by utilizing information on the quality of individual questions.
For further insights into how measurement quality can be improved by correcting measurement errors in the ESS, various reports, working papers, and articles are available at https://www.europeansocialsurvey.org/methodology/methodological-research/correction-measurement-error

$$q_s^2 = 1 - \left[\frac{\sigma^2(e_s)}{\sigma^2(s)}\right]$$

$\sigma^2(e_s)$ is the variance of the errors in the sum score and $\sigma^2(s)$ is the variance of the sum score (s). This can be estimated, using the loadings (λ_i) of the final scalar model, as follows:

$$q_s^2 = 1 - \left[\frac{(\Sigma(1-\lambda_i^2) * \sigma^2(y_i))}{\sigma^2(s)}\right]$$

The measurement quality of the sum score can range from 0 to 1, where we consider a $q^2 < 0.6$ as poor, $0.6 \leq q^2 < 0.7$ as questionable, $0.7 \leq q^2 < 0.8$ as acceptable, $0.8 \leq q^2 < 0.9$ as good, and $q^2 \geq 0.9$ as excellent quality, and 1 as perfect (DeCastellarnau & Revilla, 2017).

Results

Measurement Invariance

Figure 2 shows the results for the invariance of loadings (metric invariance) across countries. Countries illustrated in gray are not comparable, and countries shown without color were not part of the analysis. For countries with the same color, either green or purple, factor variances and covariances can be compared. We followed a specific analytical procedure: Initially, we released the equality constraints for all non-invariant countries, allowing for measurement variations across these countries. Subsequently, we conducted tests of invariance within this group. This process led us to identify a second group of countries, represented in purple, that are comparable to each other.

As can be seen in the maps, metric invariance is generally satisfied for the items in almost all countries in all rounds, except one or two countries in each group. Only the items in one country show a clear pattern of non-invariance in most rounds: Finland. In other countries, occasionally non-invariant items are found: in Italy in R1, in Denmark and France in R2, in Denmark, France and Estonia in R3, in Romania and Slovakia in R4, in Portugal and Slovakia in R5, in Hungary and Portugal in R6, and in Poland in R8. These results imply that for most countries, a one unit increase in the latent factor of interest leads to the same change in the expected value of the responses to the item across countries.

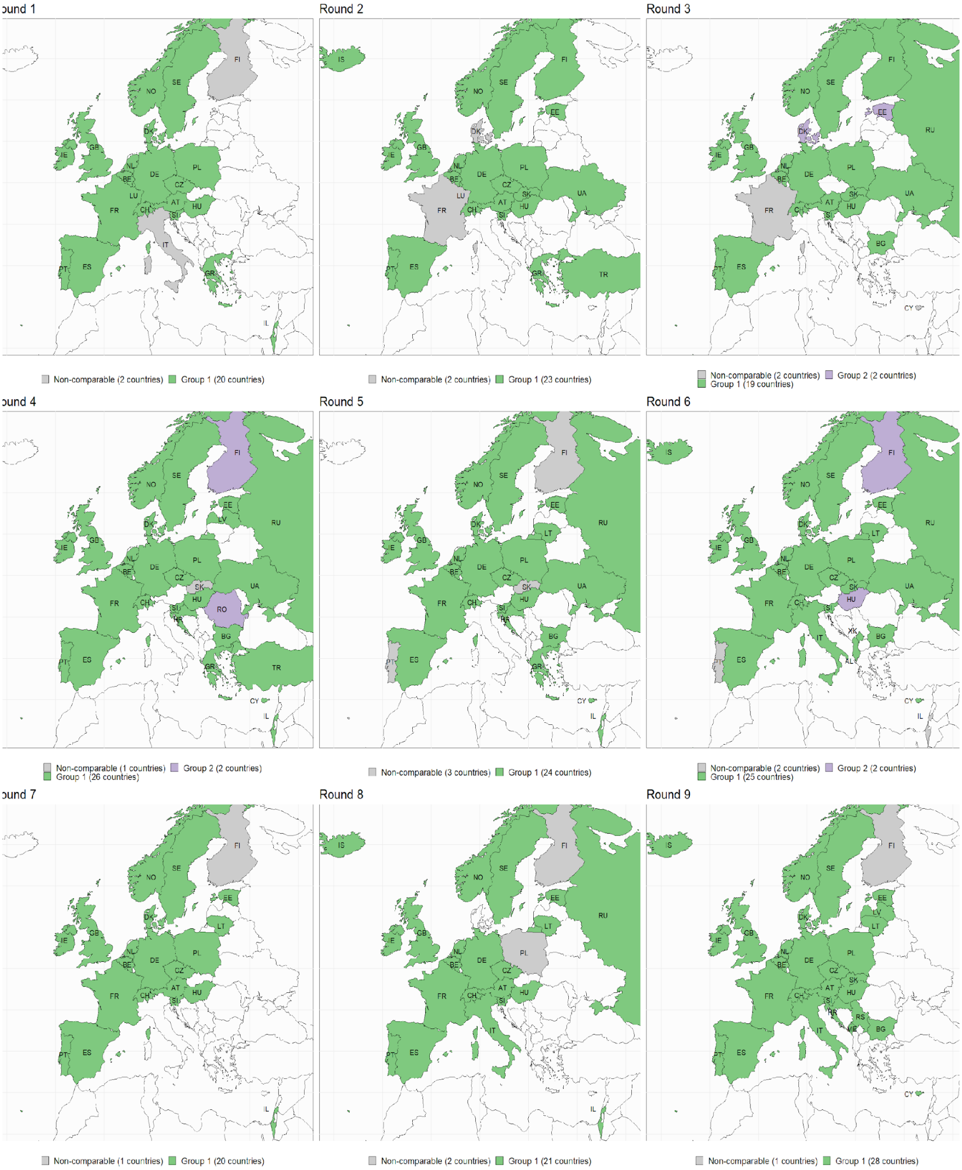

Figure 2 Metric invariance across countries in the nine rounds of the ESS

Figure 3 shows the results for scalar (intercept) invariance. In contrast to metric invariance, scalar invariance is much less widespread. First, there are different subgroups of countries for which scalar invariance holds. However, the larger groups includes a maximum of 52% of the countries (in round 4), and a minimum of 26% of the countries (in round 3). Moreover, between 30% (in round 3) to 3% (in round 9) of the countries in each round do not share intercepts with any other country of that round. This implies that for most of the countries, the same level of the latent factor corresponds to a different mean level in the responses to at least one of the items. In Figure 3, "comparability" refers to the comparability of the observed means across countries. As can be seen, comparisons of observed means across countries are not possible in many cases. In each round, observed means can only be compared across countries that have the same intercepts (shown with the same color), i.e., across a relatively small subset of countries.

Regarding the sources of scalar invariance, the most non-invariant item is the item 'Immigration bad or good for the economy' (*imbgeco*). Across all rounds and countries, 28% of the intercepts had to be freed for this item. This is followed by the item 'Immigration undermines or enriches cultural life' (*imueclt*), for which 24% of the intercepts had to be freed. Lastly, 13% of the intercepts for the item 'Immigration makes countries a worse or better place to live' (*imwbcnt*) had to be freed.

Regarding the countries, the country with the most non-invariant items was Finland (across all three items, 40% of its intercepts had to be freed; most of these corresponded to the item '*imueclt*', which had to be freed in every round). Finland is followed by Sweden and Portugal (37% of the items had to be freed; most of these corresponded to '*imbgeco*' for Sweden and to '*imwbcnt*' for Portugal). In contrast, for the countries with fewer non-invariant items, only one item in one round was found to be non-invariant across all items and rounds. These countries were Israel (representing 6% of all intercepts), Bulgaria (7%), Greece (8%), and Croatia (11%).

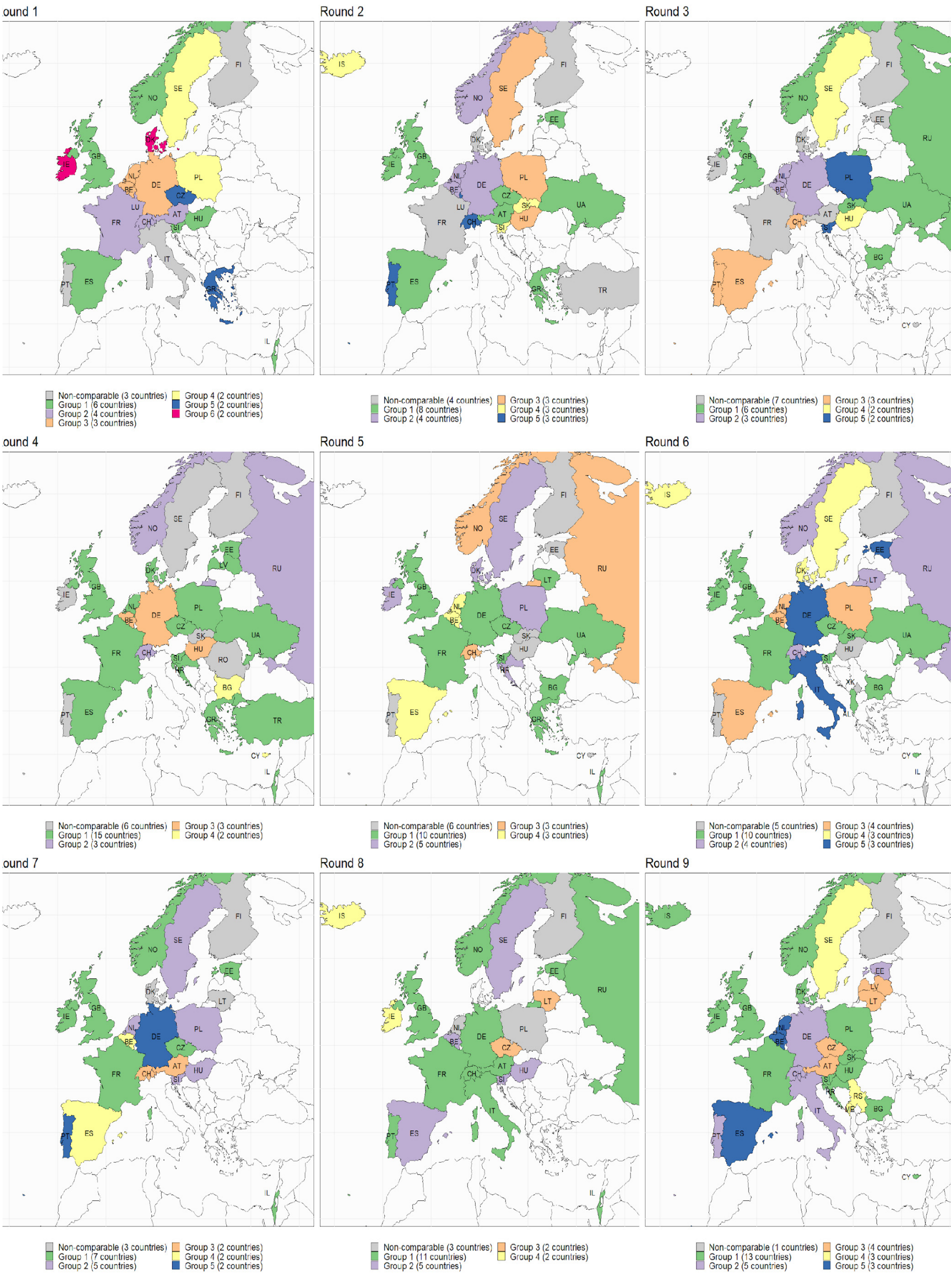

Figure 3 Scalar invariance across countries in the nine rounds of the ESS

Comparison of Latent Means

Given that scalar invariance does not hold for most countries, differences in observed means are not reliable indicators of differences in latent means. Thus, in most cases, the observed means should not be compared across all countries. However, since partial invariance is satisfied for all countries (at least one indicator is equal for all countries), latent means can be compared directly. The exact latent means are shown in Table A3 in the Appendix. In all rounds, we chose Germany as a reference point for identification purposes (i.e., its mean is always 0, and the latent mean estimates are relative to those of Germany).

Consistent with previous research, our results again confirm that the Northern European countries – Sweden, Norway, Denmark, Finland – tend to be more positive towards immigration, while Eastern Europe – Czechia, Hungary, Ukraine, Slovakia – and Southern Europe – Italy, Greece, Cyprus, Slovenia – hold the most negative attitudes. Some countries remain in the middle range, representing moderate attitudes towards immigration – Netherlands, Ireland, and to some extent Germany. Over the years, Sweden and Iceland have consistently been the most immigration-friendly countries, while Finland, Norway and Denmark rank at least in the top half.

Measurement Quality

As summarized in Table A4, Appendix, the measurement quality ranges from .68 in Luxembourg in round 1 to .94 in Bulgaria in round 9. This means that between 32% and 6% of the variance in the sum scores is due to measurement error, which should be accounted for in further analyses (Saris & Gallhofer, 2014; Saris & Revilla, 2016). Table 3 shows all measurement quality estimates for each round and country analyzed. The performance of the measure is better in countries such as the United Kingdom, Bulgaria, and Ukraine and worse in countries such as the Netherlands or Switzerland. Overall, the differences between countries are small. Besides, the performance of the measure is worse in round 1 compared to the other rounds, while it is rather similar in the rest of the rounds.

Discussion and Conclusions

In cross-national research, it is not common practice to test for measurement invariance. However, it is becoming more popular due to simplified analytical strategies in widely used statistical software. Assuming measurement equivalence without testing it can cause biased mean comparisons, covariances, and regression coefficients. Thus, it is essential to assess whether metric or scalar invariance holds for the countries and time points considered.

The aim of this study was to test the comparability and quality of the ATI scale within each round of the ESS (ESS R1 (2002) to ESS R9 (2018)). While previous research used different methods and analytic strategies, we applied the same approach in all rounds: multigroup confirmatory factor analysis (MGCFA) with local fit testing. Our results reveal that metric (loading) invariance generally holds for the items in almost all countries in all rounds except Finland. As in Davidov, Cieciuch, and Schmidt (2018), our findings again show a clear pattern of non-invariance for Finland in most rounds. Moreover, the factor loadings (slopes) are the same in most countries, indicating that covariances and regression coefficients can be meaningfully compared across most countries in all ESS rounds from 2002 to 2018.

In contrast, a less positive conclusion must be drawn in the case of scalar (intercept) invariance. It holds only for different subgroups of countries within each round, and the size of these subgroups varies considerably between ESS rounds. While scalar invariance holds for 52 % of the countries in round 4, it holds for only 26 % in round 3. Between 30 % (round 3) to 3 % (round 9) of the countries have different intercepts for at least one of the items. Thus, mean comparisons are only possible for a relatively small subset of countries.

With respect to the sources of scalar invariance, the most non-invariant item is the item 'Immigration bad or good for the economy' (*imbgeco*) (see Borgonovi & Pokropek, 2019). Since the question is asked quite generally on the topic of immigration, the answers may strongly depend on whether the respondents – and this is influenced by their cultural and political background – think of immigration in terms of illegal migration or skilled labor migration or whether they think of immigrants as people of the same or different ethnic or religious origin.

However, while scalar invariance does not hold for most countries, partial invariance is fulfilled in all countries, meaning that at least one item is equal for all countries. Therefore, latent means can be compared directly (Pokropek et al., 2019).

Our results are more or less in line with previous research showing that Europe can be classified geographically in terms of attitudes towards immigration: Whereas Northern Europe is generally more supportive of immigration, Eastern and Southern Europe are more opposed to it.

By providing an accessible and reader-friendly introduction to measurement invariance testing using multi-group confirmatory factor analysis, we aimed to support its practical application. Researchers can confidently rely on our findings and compare regression coefficients and latent means of attitudes towards immigration across countries within all ESS rounds from 2002 to 2018.

One major advantage of MGCFA is the assessment of the equivalence of measurements and structural relations across multiple groups (Harrington, 2008). MGCFA is particularly useful for comparing groups when dealing with tests comprising a substantial number of continuous items or subscale scores that are assumed to measure a limited set of underlying factors. It ensures that observed group differences are not attributable to measurement bias or variation in the underlying construct structures (Lubke, 2003).

However, several limitations need to be acknowledged: First, when comparing a large number of groups, or in longitudinal research when comparing many periods or periods far apart in time, the use of the MGCFA approach has an increased likelihood of incorrectly detecting non-invariance (Immekus, 2021; Kim et al., 2017; Leitgöb et al., 2023). To address these challenges, alternatives such as multilevel confirmatory factor analysis (ML CFA), multilevel factor mixture modeling (ML FMM), Bayesian approximate measurement invariance testing (Muthén & Asparouhov, 2013a), and alignment optimization (Asparouhov & Muthén, 2014) are suggested.

Second, the length of the scale affects the effectiveness of fit measures (D'Urso et al., 2022). When using MGCFA for measurement invariance testing of long scales, the commonly used cut-off values for RMSEA and CFI may be insufficient.

Third, the multiple indicators and multiple causes (MIMIC) modeling procedure is a recent addition to the SEM family (Tsaousis et al., 2020). In contrast to MGCFA, the MIMIC approach allows to test for measurement invariance of both categorical and continuous individual difference variables (Barendse et al., 2010) and has smaller sample size requirements than MGCFA (Leitgöb et al., 2023).

In addition, we estimated the measurement quality of the ATI score. However, our findings reveal that although the measurement quality differs across the countries, these differences are relatively small. Moreover, the performance of the measurement is quite similar across the ESS rounds, except for the first time the ESS was conducted. While this appears to give credit to the rigorous methodological approach of the ESS, there are still some measurement errors as the quality is not perfect. This stresses the importance of measurement errors correction (Saris & Revilla, 2016).

Our study is limited to cross-sectional invariance testing, which provides insights into the measurement invariance of attitudes towards immigration at a specific point in time. However, to ensure the comparability of ATI within countries across different rounds, future research is needed to incorporate cross-time invariance testing.

Ongoing and comparative research on attitudes towards immigration remains an essential task for the social sciences. Understanding the dynamics of public opposition to immigration is crucial, as it has been shown to have negative effects on social cohesion, on the lives of immigrants and refugees, and to contribute to the rise of populist radical right parties. To understand, explain, and effectively address this, accurate measurement is essential.

References

Arbuckle, J. L. (2007). *Amos 16.0 User's guide*. Spring House.

Barendse, M. T., Oort, F. J., & Garst, G. J. A. (2010). Using restricted factor analysis with latent moderated structures to detect uniform and nonuniform measurement bias; a simulation study. *AStA Advances in Statistical Analysis, 94*(2), 117–127. https://doi.org/10.1007/s10182-010-0126-1

Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and Testing Random Indirect Effects and Moderated Mediation in Multilevel Models: New Procedures and Recommendations. *Psychological Methods, 11*((2)), 142–163. https://doi.org/10.1037/1082-989X.11.2.142

Bohman, A. (2015). It's who you Know. Political Influence on Anti-Immigrant Attitudes and the Moderating Role of Intergroup Contact. *Sociological Research Online, 20*(3), 62–78. https://doi.org/10.5153/sro.362

Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., Spies, J., Estabrook, R., Kenny, S., Bates, T., Mehta, P., & Fox, J. (2011). Openmx: An Open Source Extended Structural Equation Modeling Framework. *Psychometrika, 76*(2), 306–317. https://doi.org/10.1007/s11336-010-9200-6

Borgonovi, F., & Pokropek, A. (2019). Education and Attitudes Toward Migration in a Cross Country Perspective. *Frontiers in Psychology, 10*, 2224. https://doi.org/10.3389/fpsyg.2019.02224

Davidov, E., Cieciuch, J., Meuleman, B., Schmidt, P., Algesheimer, R., & Hausherr, M. (2015). The Comparability of Measurements of Attitudes toward Immigration in the European Social Survey. *Public Opinion Quarterly, 79*(S1), 244–266. https://doi.org/10.1093/poq/nfv008

Davidov, E., Cieciuch, J., & Schmidt, P. (2018). The cross-country measurement comparability in the immigration module of the European Social Survey 2014-15. *Survey Research Methods, 12*(1), 15–27. https://doi.org/10.18148/srm/2018.v12i1.7212

Davidov, E., & Meuleman, B. (2012). Explaining Attitudes Towards Immigration Policies in European Countries: The Role of Human Values. *Journal of Ethnic and Migration Studies, 38*(5), 757–775. https://doi.org/10.1080/1369183X.2012.667985

Davidov, E., Muthen, B., & Schmidt, P. (2018). Measurement Invariance in Cross-National Studies. *Sociological Methods & Research, 47*(4), 631–636. https://doi.org/10.1177/0049124118789708

DeCastellarnau, A., & Revilla, M. (2017). Two Approaches to Evaluate Measurement Quality in Online Surveys: An Application Using the Norwegian Citizen Panel. Advance online publication (415-433 Pages / Survey Research Methods, Vol 11, No 4 (2017)). https://doi.org/10.18148/srm/2017.v11i4.7226

D'Urso, E. D., Roover, K. de, Vermunt, J. K., & Tijmstra, J. (2022). Scale length does matter: Recommendations for measurement invariance testing with categorical factor analysis and item response theory approaches. *Behavior Research Methods, 54*(5), 2114–2145. https://doi.org/10.3758/s13428-021-01690-7

European Social Survey Round 1 Data. (2002). *Data file edition 6.6. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.* https://doi.org/10.21338/NSD-ESS1-2002

European Social Survey Round 2 Data. (2004). *Data file edition 3.6. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.* https://doi.org/10.21338/NSD-ESS2-2004

European Social Survey Round 3 Data. (2006). *Data file edition 3.7. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.* https://doi.org/10.21338/NSD-ESS3-2006

European Social Survey Round 4 Data. (2008). *Data file edition 4.5. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.* https://doi.org/10.21338/NSD-ESS4-2008

European Social Survey Round 5 Data. (2010). *Data file edition 3.4. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.* https://doi.org/10.21338/NSD-ESS5-2010

European Social Survey Round 6 Data. (2012). *Data file edition 2.4. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.* https://doi.org/10.21338/NSD-ESS6-2012

European Social Survey Round 7 Data, & European Social Survey ERIC. (2014 // 2015). *Data file edition 2.2. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC // European Social Survey (ESS), Round 7 - 2014.* https://doi.org/10.21338/NSD-ESS7-2014

European Social Survey Round 8 Data. (2016). *Data file edition 2.2. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.* https://doi.org/10.21338/NSD-ESS8-2016

European Social Survey Round 9 Data. (2018). *Data file edition 3.1. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.* https://doi.org/10.21338/NSD-ESS9-2018

Fischer, R., Karl, J. A., Fontaine, J. R. J., & Poortinga, Y. H. (2022). Evidence of Validity Does not Rule out Systematic Bias: A Commentary on Nomological Noise and Cross-Cultural Invariance. *Sociological Methods & Research*, 004912412210917. https://doi.org/10.1177/00491241221091756

Groskurth, K., Bluemke, M., & Lechner, C. M. (2021). *Why We Need to Abandon Fixed Cutoffs for Goodness-of-Fit Indices: An Extensive Simulation and Possible Solutions.* https://doi.org/10.31234/osf.io/5qag3

Harrington, D. (2008). Use of Confirmatory Factor Analysis with Multiple Groups. In D. Harrington (Ed.), *Confirmatory Factor Analysis* (pp. 78–99). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195339888.003.0005

Hayes, A. F., & Coutts, J. J. (2020). Use Omega Rather than Cronbach's Alpha for Estimating Reliability. But… *Communication Methods and Measures, 14*(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629

Heath, A., Schmidt, P., Green, E., Ramos, A., Davidov, E., & Ford, R. (2016). *Attitudes towards Immigration and their Antecedents: Topline Results from Round 7 of the European Social Survey* (ESS Topline Results Series), *7*. https://www.europeansocialsurvey.org/docs/findings/ESS7_toplines_issue_7_immigration.pdf

Holland, P. W., & Wainer, H. (Eds.). (2015). *Differential Item Functioning* (1st edition). London. Routledge.

Immekus, J. C. (2021). Multigroup CFA and alignment approaches for testing measurement invariance and factor score estimation: Illustration with the schoolwork-related anxiety survey across countries and gender. *Methodology, 17*(1), 22–38. https://doi.org/10.5964/meth.2281

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36*(4), 409–426. https://doi.org/10.1007/BF02291366

Jöreskog, K. G., & Sörbom, D. (1996-2001). *LISREL 8: User's reference guide (2nd ed.)*. Scientific Software International.

Jorgensen, T., Pornprasertmant, S., Schoeman, A., & Rosseel, Y. (2021). *semTools: Useful tools for structural equation modeling.* (0.5-5.912) [Computer software].

Kim, E. S., Cao, C., Wang, Y., & Nguyen, D. T. (2017). Measurement Invariance Testing with Many Groups: A Comparison of Five Approaches. *Structural Equation Modeling: A Multidisciplinary Journal, 24*(4), 524–544. https://doi.org/10.1080/10705511.2017.1304822

Leitgöb, H., Seddig, D., Asparouhov, T., Behr, D., Davidov, E., Roover, K. de, Jak, S., Meitinger, K., Menold, N., Muthén, B., Rudnev, M., Schmidt, P., & van de Schoot, R. (2023). Measurement invariance in the social sciences: Historical development, methodological challenges, state of the art, and future perspectives. *Social Science Research, 110,* 102805. https://doi.org/10.1016/j.ssresearch.2022.102805

Lubke, G. (2003). On the relationship between sources of within- and between-group differences and measurement invariance in the common factor model. *Intelligence, 31*(6), 543–566. https://doi.org/10.1016/S0160-2896(03)00051-5#

Meitinger, K., Davidov, E., Schmidt, P., & Braun, M. (2020). Measurement Invariance: Testing for It and Explaining Why It is Absent. *Survey Research Methods, 14*(4), 345–349. https://doi.org/10.18148/srm/2020.v14i4.7655

Meuleman, B., & Billiet, J. (2012). Measuring Attitudes toward Immigration in Europe: The Cross-cultural Validity of the ESS Immigration Scales. *Ask Research & Methods, 21*(1), 5–29. http://hdl.handle.net/1811/69578

Meuleman, B., Davidov, E., & Billiet, J. (2009). Changing attitudes toward immigration in Europe, 2002-2007: A dynamic group conflict theory approach. *Social Science Research, 38*(2), 352–365. https://doi.org/10.1016/j.ssresearch.2008.09.006

Meuleman, B., Żółtak, T., Pokropek, A., Davidov, E., Muthén, B., Oberski, D. L., Billiet, J., & Schmidt, P. (2022). Why Measurement Invariance is Important in Comparative Research. A Response to Welzel et al. (2021). *Sociological Methods & Research,* 004912412210917. https://doi.org/10.1177/00491241221091755

Millsap, R. E. (2011). *Statistical approaches to measurement invariance,* New York. Routledge.

Mudde, C. (2007). *Populist radical right parties in Europe,* Cambridge, UK, New York. Cambridge University Press.

Mudde, C. (2020). Riding the fourth wave. *IPPR Progressive Review, 26*(4), 296–304. https://doi.org/10.1111/newe.12175

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods, 17*(3), 313–335. https://doi.org/10.1037/a0026802

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide*. Muthén & Muthén.

Nickel, A. (2022). Institutional Anomie, Market-Based Values and Anti-Immigrant Attitudes: A Multilevel Analysis in 28 European Countries. *International Journal of Conflict and Violence (IJCV), 16*(1), 1–15. https://doi.org/10.11576/ijcv-5126

Pirralha, A., & Weber, W. (2020). Correction for measurement error in invariance testing: An illustration using SQP. *PloS One, 15*(10), e0239421. https://doi.org/10.1371/journal.pone.0239421

Pokropek, A., Davidov, E., & Schmidt, P. (2019). A Monte Carlo Simulation Study to Assess The Appropriateness of Traditional and Newer Approaches to Test for Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 26*(5), 724–744. https://doi.org/10.1080/10705511.2018.1561293

Poses, C., Revilla, M., Asensio, M., Schwarz, H., & Weber, W. (2021). Measurement quality of 67 common social sciences questions across countries and languages based on

28 Multitrait-Multimethod experiments implemented in the European Social Survey. Advance online publication (235-256 Pages / Survey Research Methods, Vol 15 No 3 (2021)). https://doi.org/10.18148/srm/2021.v15i3.7816

Putnick, D. L., & Bornstein, M. H. (2016). Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. *Developmental Review : DR, 41,* 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Quillian, L. (1995). Prejudice as a Response to Perceived Group Threat: Population Composition and Anti-Immigrant and Racial Prejudice in Europe. *American Sociological Review, 60*(4), 586. https://doi.org/10.2307/2096296

R Core Team. (2021). *R: A Language and Environment for Statistical Computing (4.0.4) [R].* R Foundation for Statistical Computing. https://www.R-project.org

Raju, N. S., Laffitte, L. J., & Byrne, B. M. (2002). Measurement equivalence: A comparison of methods based on confirmatory factor analysis and item response theory. *The Journal of Applied Psychology, 87*(3), 517–529. https://doi.org/10.1037/0021-9010.87.3.517

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114*(3), 552–566. https://doi.org/10.1037/0033-2909.114.3.552

Revilla, M. A. (2012). Measurement invariance and quality of composite scores in a face-to-face and a web survey. Advance online publication (17-28 Pages / Survey Research Methods, Vol 7, No 1 (2013)). https://doi.org/10.18148/srm/2013.v7i1.5098

Roots, A., Masso, A., & Ainsaar, M. (2016). *Measuring Attitudes towards Immigrants: Validation of Immigration Attitude Index Across.* Lausanne: European Social Survey Conference, 13-15th July.

Rosseel, Y. (2012). lavaan : An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2). https://doi.org/10.18637/jss.v048.i02

Saris, W. E., & Gallhofer, I. N. (2014). *Design, Evaluation, and Analysis of Questionnaires for Survey Research*. Wiley. https://doi.org/10.1002/9781118634646

Saris, W. E., & Revilla, M. (2016). Correction for Measurement Errors in Survey Research: Necessary and Possible. *Social Indicators Research, 127*(3), 1005–1020. https://doi.org/10.1007/s11205-015-1002-x

Saris, W. E., Satorra, A., & van der Veld, W. M. (2009). Testing Structural Equation Models or Detection of Misspecifications? *Structural Equation Modeling: A Multidisciplinary Journal, 16*(4), 561–582. https://doi.org/10.1080/10705510903203433

Scheepers, P., Gijberts, M., & Coenders, M. (2002). Ethnic Exclusionism in European Countries. Public Opposition to Civil Rights for Legal Migrants as a Response to Perceived Ethnic Threat. *European Sociological Review, 18*(1), 17–34. https://doi.org/10.1093/esr/18.1.17

Schroeders, U., & Gnambs, T. (2020). Degrees of Freedom in Multigroup Confirmatory Factor Analyses. *European Journal of Psychological Assessment, 36*(1), 105–113. https://doi.org/10.1027/1015-5759/a000500

Seddig, D., Maskileyson, D., & Davidov, E. (2020). The Comparability of Measures in the Ageism Module of the Fourth Round of the European Social Survey, 2008-2009. Advance online publication (351-364 Pages / Survey Research Methods, Vol 14 No 4 (2020): Special Issue: Measurement Equivalence: Testing for It and Explaining Why It is Absent). https://doi.org/10.18148/srm/2020.v14i4.7369

Stark, S., Chernyshenko, O. S., & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. *The Journal of Applied Psychology, 91*(6), 1292–1306. https://doi.org/10.1037/0021-9010.91.6.1292

Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing Measurement Invariance in Cross-National Consumer Research. *Journal of Consumer Research, 25*(1), 78–107. https://doi.org/10.1086/209528

Steinmetz, H. (2013). Analyzing Observed Composite Differences Across Groups. *Methodology, 9*(1), 1–12. https://doi.org/10.1027/1614-2241/a000049

Tsaousis, I., Sideridis, G. D., & AlGhamdi, H. M. (2020). Measurement Invariance and Differential Item Functioning Across Gender Within a Latent Class Analysis Framework: Evidence From a High-Stakes Test for University Admission in Saudi Arabia. *Frontiers in Psychology, 11,* 622. https://doi.org/10.3389/fpsyg.2020.00622

van de Schoot, R., Kluytmans, A., Tummers, L., Lugtig, P., Hox, J., & Muthén, B. (2013). Facing off with Scylla and Charybdis: A comparison of scalar, partial, and the novel possibility of approximate measurement invariance. *Frontiers in Psychology, 4,* 770. https://doi.org/10.3389/fpsyg.2013.00770

van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology, 9*(4), 486–492. https://doi.org/10.1080/17405629.2012.686740

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*(1), 4–70. https://doi.org/10.1177/109442810031002

Weber, W. (2011). Testing for measurement equivalence of individuals' left-right orientation. Advance online publication (1-10 Pages / Survey Research Methods, Vol 5, No 1 (2011)). https://doi.org/10.18148/srm/2011.v5i1.4622

Weldon, S. A. (2006). The Institutional Context of Tolerance for Ethnic Minorities: A Comparative, Multilevel Analysis of Western Europe. *American Journal of Political Science, 50*(2), 331–349. https://www.jstor.org/stable/3694276

Welzel, C., Brunkert, L., Kruse, S., & Inglehart, R. F. (2021). Non-invariance? An Overstated Problem With Misconceived Causes. *Sociological Methods & Research,* 004912412199552. https://doi.org/10.1177/0049124121995521

Welzel, C., Kruse, S., & Brunkert, L. (2022). Against the Mainstream: On the Limitations of Non-Invariance Diagnostics. *Sociological Methods & Research,* 004912412210917. https://doi.org/10.1177/00491241221091754

Widaman, K. F., & Grimm, K. J. (2014). Advanced psychometrics: Confirmatory factor analysis, item response theory, and the study of measurement invariance. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 534–570). Cambridge University Press.

Zavala-Rojas, D., & Saris, W. E. (2018). Measurement Invariance in Multilingual Survey Research: The Role of the Language of the Questionnaire. *Social Indicators Research, 140*(2), 485–510. https://doi.org/10.1007/s11205-017-1787-x

Appendix

Table A1: Study constructs of attitudes towards immigration, questions, and response scales

Questions	Response scale	Studies
ESS Round 1 (2002) – 9 (2018)		
To what extent do you think [country] should allow people … of the same race or ethnic group from most [country] people to come and live here? … of a different race or ethnic group from most [country] people to come and live here? … from the poorer countries outside Europe to come and live here?	1 “allow none” to 4 “allow many”	Meuleman, Davidov, and Billiet 2009: „REJECT“; Davidov et al. 2015: “Attitudes towards migration”; Borgonovi and Pokropek 2019: “Opposition to migration”
Would you say that … it is generally bad or good for [country]’s economy that people come to live here from other countries? … [country]’s cultural life is generally undermined or enriched by people coming to live here from other countries? … [country] is made a worse or a better place to live by people coming to live here from other countries?	0 “Bad / undermined/ worse” to 10 “good/ enriched/ better”	Borgonovi and Pokropek 2019: „Economic Threat“; Pirralha and Weber 2020: “Perceived ethnic threat”; Nickel 2022: “Anti-immigrant attitudes”

Questions	Response scale	Studies
ESS Round 7 (2014) Immigration Module		
To what extent do you think [country] should allow people … of the same race or ethnic group from most [country] people to come and live here? … of a different race or ethnic group from most [country] people to come and live here? … from the poorer countries outside Europe to come and live here? … from the richer countries in Europe? ... from the poorer countries in Europe to come and live here? ... from the richer countries outside Europe to come and live here?	1 “many” to 4 “none”	Meuleman and Billiet 2012: “REJECT”
Please tell me how important you think each of these things should be in deciding whether someone born, brought up, and living outside [country] should be able to come and live here. … have good educational qualifications. … have close family living here. … be able to speak [country]’s official language(s). … have work skills that [country] needs.	0 “extremely unimportant” to 10 “extremely important”	Meuleman and Billiet 2012: “CONDITION”

Questions	Response scale	Studies
People who come to live and work here generally harm the economic prospects of the poor more than the rich.	1 “agree strongly” to 5 “disagree strongly”	Meuleman and Billiet 2012: “ECOTHREAT”
If people who have come to live and work here are unemployed for a long period, they should be made to leave.		
Would you say that people who come to live here generally take jobs away from workers in [country], or generally help to create new jobs?	0 “take jobs away” to 10 “create new jobs”	
Most people who come to live here work and pay taxes. They also use health and welfare services. On balance, do you think people who come here take out more than they put in or put in more than they take out?	0 “generally take out more” to 10 „generally put in more“	
Would you say it is generally bad or good for [country]’s economy that people come to live here from other countries?	0 “bad for the economy” to 10 “good for the economy“	
Would you say that [country]’s cultural life is generally undermined or enriched by people coming to live here from other countries?	0 “cultural life undermined” to 10 “cultural life enriched“	Meuleman and Billiet 2012: “CULTTHREAT”
Please say how much you agree or disagree with each of the following statements. It is better for a country if almost everyone shares the same customs and traditions It is better for a country if there are a variety of different religions	1 “agree strongly” to 5 “disagree strongly”	

methods, data, analyses | Vol. 18(2), 2024, pp. 213-248

Questions	Response scale	Studies
To what extent do you think [country] should allow . . . people from other countries to come and live in [country]? … different race … Jewish … Muslims … Gypsies	1 “many” to 4 “none”	Davidov, Cieciuch, and Schmidt 2018: “ALLOWANCE”
Please tell me how important you think each of these things should be in deciding whether someone born, brought up and living outside [country] should be able to come and live here. … have good educational qualifications. … be able to speak [country]’s official language(s). . . . come from Christian background? . . . be white? … have work skills that [country] needs. . . . be committed to the way of life in [country]?	0 “extremely unim-portant” to 10 “ex-tremely important”	Davidov Cieciuch, and Schmidt 2018: “CONDITIONS”;

Questions	Response scale	Studies
Would you say that people who come to live here generally take jobs away from workers in [country], or generally help to create new jobs?	0 “take jobs away” to 10 “create new jobs”	Davidov, Cieciuch, and Schmidt 2018: “RT”
Would you say it is generally bad or good for [country]’s economy that people come to live here from other countries?	0 “bad for the economy” to 10 “good for the economy“	
Are [country]’s crime problems made worse or better by people coming to live here from other countries?	0 “crime problems made worse” to 10 “crime problems made better”	
Most people who come to live here work and pay taxes. They also use health and welfare services. On balance, do you think people who come here take out more than they put in or put in more than they take out?	0 “generally take out more” to 10 “generally put in more”	

Table A2: Sample size per country and per ESS round after listwise deletion (N (round 1 – round 9) = 390.276)

	Round 1 2002	Round 2 2004	Round 3 2006	Round 4 2008	Round 5 2010	Round 6 2012	Round 7 2014	Round 8 2016	Round 9 2018
Albania						1.086			
Austria	1.941	2.021	2.147				1.664	1.875	2.292
Belgium	1.729	1.716	1.750	1.717	1.680	1.845	1.747	1.750	1.730
Bulgaria			898	1.578	1.806	1.671			1.650
Switzerland	1.893	2.021	1.726	1.698	1.451	1.420	1.480	1.457	1.422
Cyprus			952	1.177	1.020	1.081			740
Czechia	1.051	2.463		1.803	2.143	1.722	1.932	2.135	2.205
Germany	2.697	2.653	2.695	2.614	2.824	2.865	2.965	2.788	2.302
Denmark	1.344	1.383	1.405	1.539	1.507	1.577	1.455		1.511
Estonia		1.615	1.272	1.489	1.615	2.133	1.903	1.946	1.826
Spain	1.431	1.512	1.707	2.320	1.780	1.796	1.715	1.768	1.489
Finland	1.927	1.957	1.850	2.156	1.834	2.152	2.028	1.890	1.717
France	1.453	1.756	1.952	2.008	1.699	1.935	1.868	2.014	1.910
United Kingdom	1.947	1.794	2.297	2.266	2.285	2.158	2.178	1.886	2.139
Greece	2.313	2.280		2.020	2.628				
Croatia				1.304	1.443				1.668
Hungary	1.327	1.261	1.239	1.273	1.325	1.719	1.441	1.381	1.470
Ireland	1.853	2.133	1.682	1.732	2.458	2.534	2.239	2.632	2.141
Israel	2.172			2.230	1.943	2.038	2.215	2.208	
Iceland		538				685		852	828
Italy	1.064					915		2.427	2.566
Lithuania	1.195	1.458							
Luxembourg					1.311	1.733	1.807	1.794	1.541
Latvia				1.750					774

	Round 1 2002	Round 2 2004	Round 3 2006	Round 4 2008	Round 5 2010	Round 6 2012	Round 7 2014	Round 8 2016	Round 9 2018
Montenegro									1.142
Netherlands	2.216	1.809	1.800	1.708	1.744	1.763	1.827	1.586	1.582
Norway	1.981	1.721	1.711	1.523	1.516	1.594	1.405	1.505	1.356
Poland	1.715	1.450	1.520	1.390	1.476	1.606	1.348	1.393	1.271
Portugal	1.209	1.804	1.751	1.941	1.877	1.897	1.171	1.190	963
Serbia				1.656					
Romania									1.724
Russia			1.938	2.061	2.195	2.136			
Sweden	1.820	1.809	1.778	1.726	1.413	1.763		2.124	
Slovenia	1.369	1.290	1.325	1.191	1.297	1.149	1.721	1.473	1.485
Slovakia		1.182	1.532	1.533	1.570	1.669	1.092	1.242	1.251
Turkey		1.516		2.077					998
Ukraine		1.439	1.515	1.345	1.435	1.600			
Kosovo						1.054			
Total	37.647	42.581	38.442	50.825	47.275	49.296	37.201	41.316	45.693

Table A3: Latent means, standard errors and rank

	Round 1			Round 2			Round 3			Round 4			Round 5			Round 6			Round 7			Round 8			Round 9		
Country	Latent mean	Stand. error	Rank	Latent mean	Stand. error	Rank	Latent mean	Stand. error	Rank	Latent mean	Stand. error	Rank	Latent mean	Stand. error	Rank	Latent mean	Stand. error	Rank	Latent mean	Stand. error	Rank	Latent mean	Stand. error	Rank	Latent mean	Stand. error	Rank
Albania																0.01	0.05	4									
Austria	0.01	0.03	9	0.05	0.03	11	-0.15	0.03	18										-0.76	0.04	19	-0.61	0.03	19	-0.72	0.04	22
Belgium	-0.30	0.03	17	-0.07	0.03	19	0.05	0.03	13	-0.13	0.03	15	-0.25	0.03	16	-0.62	0.03	18	-0.58	0.03	16	-0.04	0.03	13	-0.32	0.03	14
Bulgaria							0.58	0.05	6	0.26	0.04	4	0.18	0.03	5	-0.38	0.04	13							-1.02	0.04	26
Croatia										-0.26	0.04	17	-0.12	0.04	13										-0.45	0.04	17
Cyprus							-0.60	0.04	22	-0.34	0.04	22	-0.66	0.04	25	-1.58	0.05	29							-0.89	0.05	25
Czechia	-0.40	0.04	20	-0.24	0.03	22				-0.52	0.03	24	-0.65	0.03	24	-1.00	0.04	26	-1.14	0.03	21	-0.89	0.03	21	-1.20	0.04	29
Denmark	-0.03	0.04	11	0.47	0.04	7	0.65	0.04	4	0.26	0.04	5	0.34	0.03	3	0.05	0.04	3	-0.30	0.04	8				-0.08	0.04	9
Estonia				-0.23	0.03	21	-0.18	0.04	19	-0.29	0.04	19	-0.05	0.03	11	-0.36	0.03	12	-0.45	0.03	14	-0.54	0.03	18	-0.50	0.03	18
Finland	0.22	0.03	5	0.68	0.03	3	0.47	0.03	8	0.25	0.03	6	0.12	0.03	6	-0.14	0.03	8	-0.16	0.03	6	0.15	0.03	4	-0.13	0.03	11
France	-0.26	0.04	16	-0.04	0.03	18	-0.05	0.03	16	-0.14	0.03	16	-0.22	0.03	15	-0.67	0.04	19	-0.50	0.03	15	-0.32	0.03	14	-0.41	0.04	16
Germany	0.00	0.00	10	0.00	0.00	16	0.00	0.00	15	0.00	0.00	12	0.00	0.00	9	0.00	0.00	5	0.00	0.00	2	0.00	0.00	10	0.00	0.00	7
Greece	-0.90	0.04	22	-0.55	0.03	25				-1.02	0.04	29	-1.16	0.03	27												
Hungary	-0.46	0.04	21	-0.24	0.04	23	-0.37	0.04	21	-0.69	0.04	26	-0.53	0.04	23	-0.86	0.04	23	-0.86	0.04	20	-1.02	0.04	22	-1.12	0.04	27
Iceland				01. Jan	0.05	1										0.35	0.04	1				0.61	0.04	1	0.67	0.04	1
Ireland	-0.07	0.04	13	0.55	0.03	6	0.66	0.04	3	0.05	0.04	11	-0.07	0.03	12	-0.38	0.04	14	-0.35	0.03	9	0.05	0.03	8	0.12	0.03	4
Israel	-0.06	0.04	12							0.10	0.03	9	-0.32	0.04	19	-0.71	0.04	21	-0.43	0.03	13	-0.32	0.03	15			
Italy	0.17	0.05	6													-0.43	0.05	15				-0.82	0.03	20	-0.68	0.04	21
Kosovo																-0.86	0.06	24									
Latvia										-0.49	0.04	23													-0.35	0.05	15
Lithuania													-0.26	0.03	17	-0.55	0.04	16	-0.37	0.03	11	-0.33	0.03	16	-0.57	0.04	19
Luxembourg	0.75	0.04	2	0.70	0.04	2																					
Montenegro																									-0.65	0.04	20
Netherlands	-0.14	0.03	14	0.05	0.03	12	0.28	0.03	10	0.17	0.03	7	0.08	0.03	8	-0.28	0.03	10	-0.18	0.03	7	0.06	0.03	6	-0.16	0.03	13
Norway	0.07	0.03	8	0.17	0.03	10	0.39	0.03	9	0.13	0.03	8	0.11	0.03	7	-0.13	0.03	7	-0.13	0.03	3	0.06	0.03	7	0.03	0.03	6
Poland	0.26	0.03	4	0.42	0.03	z9	0.79	0.04	1	0.43	0.03	2	0.45	0.03	2	-0.01	0.04	6	-0.13	0.03	4	-0.02	0.03	12	-0.15	0.04	12
Portugal	-0.15	0.04	15	-0.13	0.03	20	0.15	0.03	11	-0.01	0.03	14	-0.13	0.03	14	-1.07	0.04	27	-0.39	0.04	12	0.05	0.03	9	0.20	0.04	3
Romania										0.08	0.04	10															
Russia							-0.75	0.04	23	-0.92	0.04	28	-0.86	0.03	26	-1.45	0.04	28				-1.06	0.03	23			
Serbia																									-0.76	0.04	23
Slovakia				0.05	0.04	14	0.11	0.03	12	-0.28	0.03	18	-0.45	0.04	21	-0.90	0.04	25							-1.15	0.04	28
Slovenia	-0.32	0.04	19	0.05	0.04	13	0.05	0.04	14	-0.31	0.04	20	-0.39	0.04	20	-0.56	0.04	17	-0.58	0.04	18	-0.53	0.04	17	-0.76	0.04	24
Spain	0.08	0.03	7	0.43	0.04	8	0.49	0.03	7	-0.01	0.03	13	-0.02	0.03	10	-0.30	0.04	11	-0.36	0.03	10	0.16	0.03	3	-0.07	0.04	8
Sweden	0.80	0.04	1	0.68	0.03	4	0.76	0.03	2	0.67	0.03	1	0.79	0.04	1	0.31	0.03	2	0.47	0.03	1	0.46	0.03	2	0.31	0.04	2
Switzerland	0.27	0.03	3	0.57	0.03	5	0.64	0.03	5	0.30	0.03	3	0.25	0.03	4	-0.18	0.03	9	-0.15	0.03	5	0.10	0.03	5	0.11	0.03	5
Turkey				-0.48	0.04	24				-0.85	0.04	27															
Ukraine				0.03	0.04	15	-0.19	0.04	20	-0.55	0.04	25	-0.47	0.04	22	-0.77	0.04	22									
United Kingdom	-0.31	0.04	18	0.00	0.03	17	-0.11	0.03	17	-0.33	0.03	21	-0.30	0.03	18	-0.70	0.04	20	-0.58	0.04	17	-0.02	0.03	11	-0.08	0.03	10

Table A4: Measurement quality estimates of the measure of attitudes towards immigration, for each country and round

Country	Round 1	Round 2	Round 3	Round 4	Round 5	Round 6	Round 7	Round 8	Round 9
Austria	0.79	0.84	0.85	-	-	-	0.87	0.89	0.87
Belgium	0.74	0.77	0.78	0.79	0.79	0.78	0.8	0.8	0.79
Switzerland	0.76	0.81	0.8	0.78	0.76	0.75	0.76	0.81	0.78
Czechia	0.81	0.85	-	0.82	0.86	0.88	0.83	0.83	0.83
Germany	0.78	0.82	0.83	0.83	0.85	0.81	0.84	0.86	0.85
Denmark	0.84	0.86	0.85	0.86	0.87	0.87	0.87	-	0.85
Spain	0.78	0.84	0.83	0.86	0.83	0.85	0.82	0.86	0.86
Finland	0.78	0.82	0.8	0.81	0.84	0.83	0.85	0.85	0.86
France	0.86	0.87	0.89	0.86	0.87	0.87	0.86	0.87	0.87
United Kingdom	0.85	0.89	0.89	0.9	0.89	0.89	0.89	0.89	0.91
Greece	0.83	0.89	-	0.89	0.88	-	-	-	-
Hungary	0.8	0.83	0.84	0.82	0.82	0.86	0.83	0.87	0.88
Ireland	0.85	0.89	0.87	0.87	0.9	0.9	0.87	0.89	0.9
Israel	0.83	-	-	0.87	0.84	0.82	0.83	0.86	-
Italy	0.72	-	-	-	-	0.87	-	0.89	0.9
Luxembourg	0.68	0.76	-	-	-	-	-	-	-
Netherlands	0.72	0.77	0.77	0.76	0.76	0.78	0.76	0.78	0.74
Norway	0.78	0.81	0.8	0.81	0.81	0.81	0.8	0.82	0.83
Poland	0.76	0.73	0.8	0.79	0.78	0.81	0.8	0.78	0.84
Portugal	0.8	0.84	0.81	0.81	0.81	0.85	0.79	0.79	0.79
Sweden	0.81	0.84	0.85	0.84	0.87	0.86	0.86	0.85	0.86
Slovenia	0.74	0.83	0.82	0.83	0.85	0.84	0.83	0.87	0.87
Estonia	-	0.86	0.8	0.82	0.8	0.82	0.83	0.86	0.84
Iceland	-	0.81	-	-	-	0.81	-	0.85	0.86
Slovakia	-	0.75	0.76	0.78	0.8	0.84	-	-	0.81
Turkey	-	0.85	-	0.87	-	-	-	-	-

methods, data, analyses | Vol. 18(2), 2024, pp. 213-248

Country	Round 1	Round 2	Round 3	Round 4	Round 5	Round 6	Round 7	Round 8	Round 9
Ukraine	-	0.89	0.87	0.87	0.9	0.86	-	-	-
Bulgaria	-	-	0.88	0.87	0.87	0.88	-	-	0.94
Cyprus	-	-	0.77	0.79	0.82	0.8	-	-	0.82
Russia	-	-	0.89	0.87	0.88	0.87	-	0.85	-
Croatia	-	-	-	0.87	0.89	-	-	-	0.86
Latvia	-	-	-	0.84	-	-	-	-	0.84
Romania	-	-	-	0.85	-	-	-	-	-
Lithuania	-	-	-	-	0.82	0.85	0.83	0.86	0.85
Albania	-	-	-	-	-	0.74	-	-	-
-	-	-	-	-	-	0.87	-	-	-
Montenegro	-	-	-	-	-	-	-	-	0.89
Serbia	-	-	-	-	-	-	-	-	0.91