Check for updates

Can We Compare Solidarity Across Europe? What, Why, When, and How to Assess Exact and Approximate Equivalence of First- and Second-Order Factor Models

Vera Lomazzi*

Data Archive for the Social Sciences, GESIS Leibniz Institute for the Social Sciences, Cologne, Germany

Although measurement invariance is widely considered a precondition for meaningful cross-sectional comparisons, substantive studies have often neglected evaluating this assumption, thereby risking drawing conclusions and making theoretical generalizations based on misleading results. This study offers a theoretical overview of the key issues concerning the measurement and the comparison of socio-political values and aims to answer the questions of what must be evaluated, why, when, and how to assess measurement equivalence. This paper discusses the implications of formative and reflective approaches to the measurement of socio-political values and introduces challenges in their comparison across different countries. From this perspective, exact and approximate approaches to equivalence are described as well as their empirical translation in statistical techniques, such as the multigroup confirmatory factor analysis (MGCFA) and the frequentist alignment method. To illustrate the application of these methods, the study investigates the construct of solidarity as measured by European Values Study (EVS) and using data collected in 34 countries in the last wave of the EVS (2017–2020). The concept is captured through a battery of nine items reflecting three dimensions of solidarity: social, local, and global. Two measurement models are hypothesized: a first-order factor model, in which the three independent dimensions of solidarity are correlated, and a second-order factor model, in which solidarity is conceived according to a hierarchical principle, and the construct of solidarity is reflected in the three sub-factors. In testing the equivalence of the first-order factor model, the results of the MGCFA indicated that metric invariance was achieved. The alignment method supported approximate equivalence only when the model was reduced to two factors, excluding global solidarity. The second-order factor model fit the data of only seven countries, in which this model could be used to study solidarity as a second-order concept. However, the comparison across countries resulted not appropriate at any level of invariance. Finally, the implications of these results for further substantive research are discussed.

Keywords: measurement equivalence, alignment, MGCFA, approximate equivalence, solidarity, European Values Study

INTRODUCTION

Solidarity is a crucial concept in the social and political sciences. It represents both a core societal dimension and a political goal in local, national, and supranational entities. The degree of solidarity between individuals and between communities indicates their cohesiveness. By binding people, solidarity allows the existence of society itself. Despite its relevance, solidarity is still a "fuzzy concept" (Rusu, 2012, p. 72) that has been interpreted in a multitude of ways. The first distinction concerns the focus of solidarity as a societal or an individual quality. Durkheim (1893) conceptualized solidarity as a societal characteristic, whereas contemporary scholarship has often focused on solidarity as an individual quality, referring to the "willingness to help others or to support the group one belongs to, without immediately getting something in return" (de Beer and Koster, 2009, p. 15).

The founding fathers of the European Community made solidarity a key European value (Stjernø, 2011). They extended the concept of inter-individual solidarity to relationships between states. Political observers have argued that the lack of between-group solidarity threatens democracies because it enables the rise of ethnic and nationalistic exclusiveness (Wallaschek, 2019). The economic, political, and social challenges derived from consequential crises in Europe have attracted the attention of many political scientists. In particular, the financial crisis in 2008 and the following economic crisis prompted scholars to investigate the material and economic aspects of solidarity within societies, focusing on formal and institutionalized solidarity through welfare systems (Ferrera, 2014; Fernández, 2019) and between states (Grimmel and Giang, 2017). The refugee crisis has again challenged European solidarity by questioning shared responsibility at the macro level, the distribution of asylum seekers, and individual actions of solidarity (Kalogeraki, 2018; Lahusen and Theiss, 2019; Maggini and Fernández, 2019). The COVID-19 crisis is also expected to affect the components of solidarity (Voicu et al., 2021). In the last decade, in conjunction with economic and refugee crises, the debate has paid particular attention to the transnational dimension of solidarity. In the European context, the focus has been on the relationship of mutual support between states (i.e., intergovernmental solidarity), specifically support for countries that were the most heavily affected by the recession and by the new migration flows.

In addition to the redistributive dimensions of solidarity, such as institutionalized forms of intragovernmental solidarity and national welfare systems, micro-level approaches to solidarity focus on the behavioral and attitudinal dimensions of solidarity. Studies on acts of solidarity (Kalogeraki, 2018; Lahusen and Theiss, 2019; Maggini and Fernández, 2019) have tended to focus on cooperative behaviors or behaviors oriented toward other individuals and specific social categories that are considered in need, directly practiced by people, or mediated by organizations (i.e., civil society organizations, NGOs, volunteer groups, etc.). While solidarity involves collective responsibilities toward communities at a more generic level, it might also have delimited scope toward specific groups and needs (Lahusen and Grasso, 2018; Mishra and Rath, 2020; Voicu et al., 2021). Feelings of reciprocity, willingness to help, and concern for others' lives can therefore be addressed to unprivileged members of society, to people included in one's close network, as well as to (imagined) people living further away, as well as to the generic humankind. So, in this field of research, a key question regards the beneficiaries of solidarity, specifically, the social groups that are considered to deserve help and, if so, in which form.

The underlying values of benevolence, altruism and empathy can take the form of the orientation to help and concern with the living conditions of people belonging to different social groups (Abela, 2004; Rusu, 2012). While this value can be generally considered as a relevant motivator for actual political and social participation, in broader definitions of socio-political participation (Harris et al., 2010; Bakker and de Vreese, 2011; Ekman and Amnå, 2012; Amnå and Ekman, 2014), this form of solidarity, framed as social concern, is included in political participation repertoires that go beyond institutionalized forms of participation. Especially scholars in the field of youth political participation elaborated broader conceptualizations of participation and engagement, emphasizing new forms of engagement that range from food boycotting to online activities, from recycling to art and other creative actions (Harris et al., 2010; Ekman and Amnå, 2012; Hooghe et al., 2014; Busse et al., 2015). These conceptualizations also include elements, like social concern, which can be considered as "pre-political" or "stand-by" participation and that can have a substantial relevance for future political manifest activities. As previous studies disclosed (Hooghe and Dejaeghere, 2007; Harris et al., 2010; Amnå and Ekman, 2014; Busse et al., 2015), talking about politics in the family or with friends as well as being socially concerned may prepare the field for future actions and represent an unvaluable component of social and political engagement.

This study focuses on this underlying value orientation of solidarity, conceptualized as concern for the living conditions of different groups of people. Although solidarity is relevant in comparative research, only a few studies have investigated the extent to which the cross-cultural comparison of this socio-political value is appropriate (Kankaraš and Moors, 2009). Testing measurement invariance is considered a precondition for meaningful cross-cultural comparisons. Although some approaches have challenged this perspective (Welzel and Inglehart, 2016; Welzel et al., 2019), the assumption of comparability needs to be assessed to enable the meaningful comparison of countries. This comparison is relevant in ranking countries or evaluating whether solidarity has increased or decreased over time. Solidarity measurements are often used in combination with other predictors to explain other socio-political issues, such as political preferences and electoral behaviors (Langsæther and Stubager, 2019), anti-democratic tendencies (Koštál and Klicperová-Baker, 2015), and Europeanization (Delhey et al., 2014). It is therefore important to note that to conduct proper comparisons across countries concerning the effect of this latent variable, it is necessary to determine that its measurement is equivalent across them, which means that the respondents belonging to diverse groups perceive and interpret the items in the same way.

Measurement invariance issues are well-known in the survey methodology literature (Davidov et al., 2014; Davidov et al., 2018; Cieciuch et al., 2019). Multigroup confirmatory factor analysis (MGCFA) is commonly used in these analyses. However, this technique is limited when many groups are involved in the assessment (Asparouhov and Muthén, 2014; Cieciuch et al., 2014; Davidov et al., 2018b). The main reason is that MGCFA builds on the concept of exact equivalence, which imposes extremely strict requirements that are rarely achievable in dealing with many countries. In following the approximate approach to equivalence (Cieciuch et al., 2014), new techniques have been developed. Among these, alignment optimization, which can be implemented both in the frequentist and Bayesian frameworks, is a valuable alternative to MGCFA. However, despite the flourishing debate about measurement techniques, measurement equivalence is often neglected in substantive studies.

To answer the questions of what, why, when, and how to assess the measurement equivalence of socio-political values, this study aims to offer an overview of the key issues concerning the measurement and the comparison of social and political values and attitudes. First, the formative and reflexive approaches to the measurement of latent concepts and their substantive implications are discussed. Several structural measurement models of latent constructs are described, illustrating the substantive and methodological aspects of first- and second-order factor models. Second, this study introduces the challenges of comparing socio-political values across different countries and provides arguments for the need to assess measurement invariance as a precondition of meaningful comparisons. The exact and approximate approaches to the concept of equivalence are described, as well as their empirical translation in statistical techniques. Third, the study applies up-to-date techniques to assess cross-cultural measurement equivalence in both exact and approximate approaches to the first- and second-order measurements of solidarity using data collected in 34 countries from the fifth wave of the European Values Study (EVS, 2017–2020).

MEASUREMENT AND COMPARISON OF HUMAN VALUES AND ATTITUDES

Comparative studies in the social and political sciences have been built on the assumption of comparability, which is often taken for granted. However, several factors challenge this basic assumption, which has relevant implications for conclusions based on comparing results that are not comparable. Challenges to comparability can be derived from methods as well as cultural biases (van de Vijver and Tanzer, 2004). For example, in the local implementation of a cross-national survey programme, slight differences in the procedure across the participating countries, such as in the sampling strategies or in data collection modes, can have consequences for cross-national measurement comparability (De Beuckelaer and Lievens, 2009; Hox et al., 2015). In addition to general translation issues, item interpretations can be culturally different (Braun, 2009; Davidov and Beuckelaer, 2010), or they can change over time in the

same society. For example, respondents can react differently to the term "migrant" in a question. Would they interpret the word to mean migrant workers? Or refugees? Compared with countries where migration issues are less predominant, both in reality and in the media debate, interpretation as well as the attribution of meaning can differ among people who are particularly affected by the refugee crisis or by migration flows in general. In addition, well-known phenomena in the survey methodology, such as social desirability and acquiescence, can also vary according to cultural context (van de Vijver and Tanzer, 2004; Heath et al., 2009; van Vlimmeren et al., 2016). Therefore, the assumption of comparability needs to be assessed to compare country means appropriately. Otherwise, conclusions may be drawn from misleading results.

Measurement Approaches

This study focuses on the underlying value of solidarity, expressed as social concern. This implies all the challenges that values researchers face in the definition and in the measurement of values (Halman, 1995). Because of the nature of values, which are not directly observable, scholars found hard to provide a common definition of "values" and speculative approaches are generally used (Halman and de Moor, 1993; Hechter et al., 1993; Halman, 1995). Empirically, this means that values cannot be directly measured, and it is necessary resorting to related concepts like beliefs, opinions, and attitudes that, despite of being abstract concepts themselves, can be inferred from indicators (Himmelfarb, 1993). Using these indicators, scholars can design proxy measure of values by assuming combinatory or dimensional logics. These two approaches, which build in different theoretical perspective on measurement, also reflect in appropriate analytical techniques (For more details on the two approaches, see Coltman et al., 2008; Sokolov, 2018).

The combinatory logic also referred as formative approach, relies on theoretically driven assumptions to build composite indicators. This approach is often used by international organizations to compute indexes based on official statistics, such as those provided by the United Nations to measure human development (UNPD, 2013) or by the Organization for Economic Co-operation and Development (OECD) to evaluate institutional gender equality through the Social Institutions and Gender Index (Branisa et al., 2014). To measure social solidarity according to the formative approach, scholars might pick items concerning different aspects to build an index of solidarity. For example, based on the scholars' theoretical paradigm, they could combine the scores from questions about the concern for people in need, the willingness to give money to support local charities, and active participation in voluntary activities.

However, regarding the measurement of human values and attitudes, the use of formative approaches is controversial. As van Vlimmeren et al. (2016) pointed out, the formative approach emphasizes the researcher's point of view and neglects how the concept is meant by the respondents. How researchers design an index according to their theoretical approach may not coincide with the ways in which respondents express their subjective interpretation of that concept. By following the combinatory logic, the meaning assigned to the items by respondents and the

possibility that they could be interpreted differently culturally are further aspects neglected by formative approach.

Compared with the combinatory logic, the dimensional logic (also referred as reflexive approach) is a more convincing approach to measuring human values and attitudes. According to this reflexive approach, the item responses are considered expressions of attitude, which reflect an underlying concept (i.e., value). In contrast to a measurement model based on the formative approach of combinatory logic, where the latent construct is designed by the scholar, models developed using the reflexive approach assume that the latent construct exists independently in the measurement model defined by the researcher. Following this perspective, the construct of social solidarity as social concern exists regardless if the items ask about concern about the living conditions of unemployed; of poor people; sick; people with disabilities; family with economic difficulties, and so on. By exploring the respondents' answers patterns, scholars can identify the underlying structure of social solidarity.

Furthermore, the direction of causality is from the latent variable (i.e., construct) to the items. This has important implications because it means that variations in the item measures do not affect the latent construct. In the formative approach, the direction of causality is opposite, in which changes in the item measures affect the latent construct. A main consequence of the formative approach is that the construct is defined by the items selected, which are therefore not interchangeable, and the addition or removal of items from the model would change the latent construct. However, in the reflective approach, the items are expressions of the latent construct because they belong to the same domain and therefore can be interchangeable; moreover, different items that reflect the same construct can be added (Coltman et al., 2008).

Because of these theoretical considerations, the reflexive approach appears more appropriate to define a measurement model of human values and attitudes, as in the case of solidarity. The construct of solidarity is defined as the interdependence of individuals in a society, which allows them to feel that they can enhance the lives of others. This construct could be manifested in several items that are interchangeable without affecting the latent construct (for example, listing items regarding the orientation to helping or concern about the elderly, unemployed, sick, disabled, poor, etc.).

Latent constructs of human values and attitudes can be reflected in different measurement model structures (see Brown, 2015 for a detailed overview). The simplest case is a first-order model, in which one latent construct (i.e., social solidarity) is reflected in a certain number of items (i.e., concern for the living conditions of unemployed; of poor people; of people with disabilities)

A first-order factor model could also consist of more than one latent construct, which are intercorrelated. For example, three dimensions of solidarity (i.e., social solidarity, local solidarity, global solidarity) are reflected in a certain number of items and are mutually related.

However, higher-order latent variables can also exist. For example, in second-order factor models, second-order concepts are represented according to a hierarchical structure. Following the logic of first-order models, second-order models consist of the reflective relations between first-order factors and a second-order factor. In the example of the construct of solidarity, a second-order factor model would mean that the second-order latent concept of solidarity is reflected in three dimensions (i.e., local, global, and social solidarity). In turn, each one of these dimensions is reflected in a certain number of items (i.e., for social solidarity: concern for the living conditions of unemployed; of poor people; of people with disabilities).

Although the differences between these hypothesized measurement models have theoretical implications, it is necessary to note that second-order factor models are possible only when at least three sub-dimensions are available, similar to the simplest first-order model, which requires at least three items to be identified (Bollen, 2002; Brown, 2015).

This overview of approaches to the measurement of human values and attitudes and the possible measurement model structures is relevant for pursuing consistency in the methods used in analyses. The assumption of the reflective approach justifies the use of methods and techniques based on factorial and dimensional logic, which is less appropriate for measurements developed by combinatory logic. Furthermore, the different structures in the measurement model imply the accurate adaption of the techniques used to assess measurement equivalence. In higher-order models, it is necessary to take into account the hierarchical structure between latent constructs.

Approaches to Measurement Equivalence

Measurement invariance issues are well-known in the survey methodology literature. Despite the awareness on comparability challenges is becoming always more widespread also beyond the field of methodologists, scholars as Welzel and Inglehart (2016) and Welzel et al. (2019) argued that the issue of measurement equivalence is overrated, and the convergence with external criteria is enough to validate a measure and use it at the aggregate level to compare countries. However, these claims are not convincing and their arguments, which seem to be built on equating measurement equivalence issues with one of the analytical techniques used for its assessment (MGCFA), have been refuted mathematically and statistically by rigorous research that shows flaws in the comparability of their measure of emancipative values as well as in the quality of the measurement, which neglects measurement errors (Alemán and Woods, 2016; van Vlimmeren et al., 2016; Sokolov, 2018; Meuleman, 2019).

Scholars who apply dimensional logic to the measurement of attitudes and values underlined, instead, the need for equivalence assessment to enable meaningful comparisons (Davidov et al., 2014, 2018b; Cieciuch et al., 2019). However, among scholars who have used the reflective approach, a recent debate challenged the concept of equivalence itself.

Most of the techniques used to assess measurement invariance, such as latent class modeling (Kankaraš and Moors, 2009; Rudnev, 2018), item response theory (Raju et al., 2002; Millsap, 2010), and multigroup confirmatory factor analysis, are based on the concept of exact equivalence. According to this concept, to meaningfully compare groups and

test general theories, the instruments used must be exactly the same. In other words, the measurement model must present the same configuration, factor loadings, and intercepts across groups. Using solidarity as example, this means that respondents belonging to different groups not only display the same construct configuration, meaning that they intend the concept in the same way (configural invariance), but they also use the same metric to measure solidarity (metric invariance), and the levels of the underlying items (intercepts) are equal in both groups (scalar invariance).

However, the exact approach to equivalence is problematic, especially if dealing with many groups. In these cases, the assessment of measurement invariance often fails (Davidov et al., 2008, 2014) because researchers must deal with two opposite threats. On one hand, the approach has strict requirements to fulfill. On the other hand, the greater the number of groups included in the analysis, the greater the risk of violating invariance. In describing this condition, van de Schoot et al. (2013) used the image of "Scylla and Charybdis" to indicate the two threats: by imposing exact equivalence, the results will most likely suggest the rejection of comparability; by neglecting the assessment, the comparison will be meaningless. In line with this concern, the concept of "approximate equivalence," which was introduced by Muthén and Asparouhov (2013), appears to be the most feasible way to deal with these two threats. In contrast to the exact approach, which requires "exact equivalence" between parameters, "approximate equivalence" includes cultural variability and uncertainty in assessments (Muthén and Asparouhov, 2013; van de Schoot et al., 2013).

This conceptual approach to equivalence has been applied in the development of new empirical techniques. The Bayesian framework, which allows researchers to introduce existing knowledge into their analyses, represents one of the most natural contexts for the development of the assessment of equivalence. In this framework, it is possible to include the variability of parameters and the amount of uncertainty as priors (Cieciuch et al., 2014). In addition to studies that have compared the exact and approximate approaches to solving issues concerning cross-national comparisons in dealing with many groups (Cieciuch et al., 2014; Davidov et al., 2015; Zercher et al., 2015), the Bayesian techniques have also been applied to longitudinal studies using panel data (Seddig and Leitgöb, 2018).

However, approximate equivalence can also be investigated using the frequentist approach. The alignment optimization, which is also applicable in the Bayesian framework, has been proposed as a convenient alternative to MGCFA (Asparouhov and Muthén, 2014; Kline, 2015). Its application in real-data studies (Flake and McCoach, 2018; Lomazzi, 2018; Lomazzi and Seddig, 2020) demonstrated that it is a promising technique for scholars who aim to adopt approximate equivalence while retaining the frequentist approach.

Despite the flourishing debate regarding methods, measurement equivalence has often been neglected in substantive studies. The current study aims to assess the cross-sectional equivalence of the measurement of solidarity by adopting a frequentist approach. For this purpose, the paper illustrates the use of MGCFA, considering first-order and second-order models,

as well as alignment optimization, to discuss the implications of using exact or approximate approaches.

METHODS FOR ASSESSING MEASUREMENT EQUIVALENCE

The comparative study of solidarity across countries could be conducted with other analytical frameworks like, for example, classification-based approaches. The hierarchical latent class analysis could serve for this purpose. However, the challenges to comparability are not solved automatically, and scholars have to consider equivalence issues with this technique as well. The analytical strategy to perform hierarchical latent class analysis taking into account measurement invariance requirements has been discussed in previous literature (Kankaraš and Moors, 2009; Rudnev, 2018). This study adopts factorial models to measure solidarity and focuses on the frequentist approach to assessing measurement equivalence. In particular, it examines multigroup confirmatory factor analysis (MGCFA), which is the most frequently used technique in this field (Davidov et al., 2015), and the alignment method, which is its promising alternative.

Multigroup Confirmatory Factor Analysis (MGCFA)

In testing measurement equivalence using MGCFA (Jöreskog, 1971; Brown, 2015), invariance is assessed through a hierarchical approach. Theoretically, many levels of invariance are possible (Horn and Mcardle, 1992). However, the sociological and psychological empirical literature refers mainly to three levels of measurement invariance: configural, metric, and scalar (Steenkamp and Baumgartner, 1998; Davidov, 2010). Configural invariance refers to the structural pattern of latent variables and observed indicators, which must be the same in each group. Metric invariance requires that, in addition to the same configural structure, the factor loadings are the same across groups. Achieving this level of invariance is necessary to compare factor variances, covariances, and unstandardized regression coefficients across groups. Comparing factor means across groups is meaningful only when scalar invariance is additionally achieved. In this case, the most restrictive invariance is determined, and intercepts are also constrained to be equal.

Therefore, the assessment consists of the following sequence of steps. First, it is necessary to identify a measurement model that fits the data in each group. In addition to the model fit, other requirements need to be fulfilled at the group level: the factor loadings must be substantial (e.g., >0.30), and the correlations among factors should be smaller than one. This preliminary step concerns the evaluation of the hypothesized model, which is suggested by the results of an exploratory factor analysis and running a confirmatory factor analysis group by group. A common practice in assessment is to proceed according to a bottom-up strategy. This means starting with a less demanding level of restrictions (i.e., configural model) to arrive at the most restricted level (i.e., scalar model). Chi-square-based, goodness-of-fit measures are used to evaluate the model to determine whether the model fits the data well enough or whether it

must be rejected (Hu and Bentler, 1999; West et al., 2012). The most frequently used measures of model fit are the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardized root mean residual (SRMR). A model is considered acceptable when the CFI value is higher than 0.90 and the RMSEA and SRMR values are lower than 0.08. However, in assessing invariance, the evaluation should not be limited to a single model, but it should also consider the degree of the worsening of the model fit from the previous model assessed. For this purpose, Chen (2007) suggested cut-off criteria for the change in model fit from the previous model sample size. When $n > 300$, invariance is evaluated considering the following recommended criteria: $\Delta CHI < 0.01$; $\Delta RMSEA < 0.015$; $\Delta SRMR < 0.03$. If moving from a level of invariance to the next level worsens the model fit too much, it is possible to test for partial invariance (Byrne et al., 1989; Steenkamp and Baumgartner, 1998). Informed by the analysis of the modification indexes or the residual information provided in the results of a SEM analysis, specific factor loadings or intercepts can be released by the constraint of imposed equality and freely estimated. However, at least two parameters must be held equal across groups.

Even if it is considered acceptable, partial invariance is controversial. Achieving partial invariance might require a long and tedious procedure for the identification of the parameters to be released. Furthermore, although in some cases detecting the parameter is self-evident because a limited number of parameters with outstanding modification indexes is reported, in other cases it can be complicated to identify which parameter to free. The approach that is usually followed is to assess several parallel models to determine which model would be the best compromise between parsimony and model fit. In the most challenging cases, the expertise of the analyst is the only way to navigate this process.

The acceptance of partial invariance when a large number of parameters is released is still object of discussion in the empirical literature. While some authors raised doubts about the effective invariance of the model when many modifications are required to obtain an acceptable model fit (Asparouhov and Muthén, 2014; Lomazzi, 2018; Marsh et al., 2018). Pokropek et al. (2019), conducted a large simulation study that demonstrates that partial invariance model provides reliable results even when the majority of the parameters are not imposed to be equal as (if not better, under certain conditions) than approximate equivalence solutions. While most of the social and political science literature on equivalence assessment refers to first-order factor models, MGCFA can be used to assess the invariance of higher-order factor models. This type of structure is less common in these disciplines than in others, such as psychology, where second- or higher-order factor models are often used to validate measurement instruments. The application of MGCFA in the context of the second-order factor model follows the same logic as in first-order model assessment, but it requires further attention and a stepwise procedure (Chen et al., 2005; Rudnev et al., 2018). The assessment of the equivalence of the first-order part of the model is conditional for assessing the invariance of the second-order factor model. This means that the configural and metric invariance in the first-order model

is a prerequisite for the metric invariance in the second-order factor model. Similarly, it is necessary to achieve scalar invariance in the first-order model before testing scalar equivalence in the higher-order factor model.

Although the general procedure is the same as in the first-order MGCFA model, the model identification requires special attention. Rudnev et al. (2018) argued that, for the purpose of equivalence testing, the most appropriate method for model identification is to fix one factor loading per factor to 1. This parameter represents the "marker indicator" and needs to be selected carefully. By fixing the marker indicators to 1 for each factor, it is assumed they are equal across groups. For this reason, it is important to select marker indicators that are the most invariant and avoid parameters that display large modification indexes. To identify the means structure, one indicator intercept per first factor is fixed to 0. This reduces the risk of the incorrect detection of invariance levels and helps researchers to interpret the model clearly. The identification of the second-order part of the model follows the same logic, and the loading of one of the first-order factors is fixed to 1 so that the other can be estimated. To identify the mean structure of the second-order factor, the means in one group can be fixed to 0.

When the model is identified, the testing procedure follows a bottom-up approach. The first step consists of the assessment of the model country by country (CFA) and implies the exclusion of groups for which the model does not fit the data. Second, it is necessary to assess the configural model of the first-order factor model. Third, metric invariance in the first- and second-order factor models is assessed. Finally, if all the previous models are equivalent, scalar invariance is assessed for the first- and then for the second-order factor models. The same criteria are adopted in the evaluations of the first- and second-order factor models to determine their fit (Chen et al., 2005; Chen, 2007; Rudnev et al., 2018).

Alignment

The exact approach to equivalence entails strict requirements that are often difficult to fulfill in dealing with many groups. The main reason is the increase in variance that each group brings into the estimation (Davidov et al., 2008, 2014). The consequence is that in these cases, the assessment of full equivalence via MGCFA often fails (Cieciuch et al., 2019). Furthermore, a limit of MGCFA is that because of its strictness, it can lead to discarding models with factor means that could be comparable. From this perspective, the recently developed alignment method could be particularly helpful (Asparouhov and Muthén, 2014).

MGCFA relies on the idea of exact equivalence of parameters across groups, whereas alignment builds on the assumption that a certain degree of non-invariance is acceptable and can be kept at a minimum. According to this perspective, the method accounts for differences across groups. In the procedure, the factor means and the variance of the parameters for each group are estimated by applying a simplicity function, which serves as the rotation criteria in the exploratory factor analysis, to identify the most invariant pattern. The technique initially estimates the "model zero," which is the configural model with all unconstrained parameters. In the frequentist alignment, the

maximum likelihood (ML) or maximum likelihood with robust standard errors (MLR) estimation can be used to freely estimate the parameters, using factor means fixed at zero and factor variances fixed at one. This model is the best that can be achieved with the groups included in the analysis.

The second step in the procedure consists of alignment optimization. By applying a simplicity function, the amount of variance is minimized, but the fit remains the same as in the model zero. The frequentist alignment can be conducted using the FREE or FIXED estimation methods. The differences between these two procedures concern the estimation of the first group factor intercept, which is then used as a reference for the estimation of the other groups' parameters. In the FIXED alignment, the factor mean of the reference group is fixed to be equal to zero, whereas in the FREE alignment, it is freely estimated. As Asparouhov and Muthén (2014) demonstrated, FREE alignment may not be applicable in only two groups and/or when a high degree of invariance exists.

The developers of the alignment method conducted several Monte Carlo studies to validate this technique, taking into account several conditions, such as the estimation method, the number of groups, and the group sample size (Asparouhov and Muthén, 2014; Muthén and Asparouhov, 2014). Most importantly, these investigations provided a limit to the amount of non-invariance that could be considered acceptable to maintain the trustworthiness of the alignment results. Asparouhov and Muthén (2014) stated that non-invariant parameters up to 25% are still acceptable to consider reliable the estimated factor means. Furthermore, they recommended complementing the assessment using Monte Carlo simulations to validate the alignment results when the amount of non-invariant parameter is higher than (or close to) 25%. These conclusions are confirmed by the recent simulation studies by Flake and McCoach (2018) and Pokropek et al. (2019), which also pointed out the strengths and limitations of the alignment method. Particularly, in both these two studies, the alignment demonstrated a good performance in condition of small and moderate invariance. The recent work by Pokropek et al. (2019) is based on a large-scale Monte Carlo simulation study that investigated under 156 conditions the ability of models to recover true latent means and path coefficients between the latent variable and two criterion variables. The investigation used five different estimation strategies, allowing the comparison of the performance of exact and approximate approaches. Whereas, Flake and McCoach (2018) confirmed that the alignment results can be trusted when the amount of non-invariance is smaller than 25%, according to Pokropek et al. (2019) the accuracy of the parameter estimation of the alignment method is weaker than MGCFA when non-invariance is bigger than 20%. These recent results ulteriorly reinforce Asparouhov and Muthén's recommendation (2014) on performing Monte Carlo investigations to confirm the accuracy of the alignment estimation when the amount of non-invariance is higher than this threshold.

The alignment procedure, initially developed in the context of the commercial statistical package Mplus (Muthén and Muthén, 1998), can now be performed also in the open-source software environment of R thanks to the *sirt* package, which includes

invariance alignment functions (Robitzsch, 2019). Compared to Mplus, this package uses a slightly different algorithm and seems to provide slightly different performances according to the number of groups, the sample sizes, the type of data, and degree of non-invariance (Pokropek et al., 2020). However, when evaluating the simulations concerning a cross-national case with 25 groups, Pokropek et al. (2020) show that the deviations between the software packages are quite close, while for different settings, like for example small number of groups, the algorithm used in the *sirt* package seems to perform the alignment better. In addition to carefully read the differences between the two software packages and the implications for the analyses (Pokropek et al., 2020; Robitzsch, 2020), R-users could find helpful the tutorial for measurement invariance by Fischer and Karl (2019), which also includes invariance alignment.

In this study, all the analyses are carried out using Mplus version 8.4 (Muthén and Muthén, 1998).

While it does not seem possible to analyze multi-dimensional construct performing the alignment in R (Fischer and Karl, 2019), Mplus allows for multi-factor alignment models (Flake and McCoach, 2018) but it is still not possible to estimate cross-loadings between factors.

Furthermore, the alignment method is usually applied to first-order factor models. Neither Mplus nor R support the implementation of this technique in second-order factor models.

To allow for including the approximate approach in the analysis of higher-order models, a solution could be to replace the first-order factor part of the second-order factor model with the alignment solution provided by the alignment method by first identifying correlated factors to fix the minimum number of parameters. However, the empirical implementation of this mixed technique has rarely been considered, and it will not be explored in this study.

THE CURRENT STUDY

Measurement equivalence is a precondition for running meaningful substantive comparisons; however, it has rarely been implemented in comparative social and political studies. As argued in the previous paragraphs, the issue of equivalence requires attention beyond methodology research. If scholars would develop their substantive studies without testing the assumption of comparability, their conclusions risk being based on misleading results. The current study therefore aims to illustrate the application of MGCFA and alignment, which are two techniques developed in the frequentist approach, to assess measurement invariance by adopting the exact and approximate approach to equivalence. For this purpose, this study uses the measurement of solidarity, a concept that has a crucial role in many comparative social and political studies, but whose measurement invariance has rarely been investigated. In addition to the application of the techniques, the implications for substantive research are discussed.

Data
The study uses data collected in 34 countries that were included in the most recent edition of the European Values Study (EVS,

2017–2020). The total sample size was 56,491 respondents. **Table 1** lists the countries and their sample sizes.

The EVS is a cross-sectional survey that has collected data on human values every 9 years since 1981 from representative samples of national populations (EVS, 2017). Among the several value domains investigated, the questionnaire includes a battery of nine items covering the concept of solidarity. The measurement asks the respondents to express the extent to which they are concerned about the living conditions of a series of groups, which allows for exploring solidarity as a subjective predisposition to be interested in others and potentially offer support to them. The respondents rate their concern from 1 (very much) to 5 (not at all). **Table 2** lists the nine items as presented in the master questionnaire and the descriptive statistics of each item.

The measurement of solidarity has also been surveyed in previous waves of the EVS. Kankaraš and Moors (2009) assessed the measurement equivalence of the instrument used in the EVS in 1999 by applying a multigroup latent class factor approach. According to their study, based on data from 33 countries, measurement equivalence was not achieved, but valid cross-national comparisons were carried out when country inequivalences were included in the modeling. To date, Kankaraš and Moors' (2009) study is the only one to assess the equivalence of this survey instrument; more recent data have not been evaluated, except the validity testing conducted on the Romanian sample of the EVS in 1999 and 2008 by Rusu (2012).

Analytical Strategy

The strategy applied to assess measurement equivalence followed a stepwise approach. First, a preliminary exploratory factor analysis was carried out to identify the general structure of the measurement model. The results (**Tables 3**, **4**) displayed a three-factor structure, which confirmed the configuration presented in previous analyses of the same instrument surveyed in the EVS in 1999 (Abela, 2004; Rusu, 2012). The structure consisted of three dimensions that Abela (2004), using previous waves of EVS, defined as follows: (a) global solidarity (i.e., concern for humankind, immigrants, and Europeans); (b) local solidarity (i.e., concern for people living in the neighborhood, region, and country); and (c) social solidarity (i.e., concern for the elderly, sick, disabled, and unemployed). This structure allowed for the application of different approaches to the assessment of invariance. The classical solution is a model where the three factors of global, local, and social solidarity are intercorrelated (i.e., the first-order factor model). However, scholars may be interested in using a model representing the reflective relationships between the three factors and the second-order latent variable (i.e., the second-order factor model). Therefore, the analytical pathway was branched to consider these two possibilities: assessing measurement equivalence in the first-order factor model; and assessing invariance in the second-order factor model.

First, the first-order factor model was assessed country by country, which may have led to the exclusion of countries displaying a poor model fit. Measurement invariance was then assessed by the MGCFA, adopting a bottom-up strategy. After assessing configural invariance, factor loadings were constrained to be equal to assess metric invariance. Finally, the equality of all intercepts was imposed to assess scalar invariance. If full invariance was not reachable, the in-depth investigation of the modification indexes allowed for identifying the parameters to be released to assess partial invariance.

Regarding the approximate approach to equivalence, the alignment method was used to assess invariance of the same first-order factor model across the same group of countries. Following previous studies in which the variance between groups was large (Asparouhov and Muthén, 2014; Lomazzi and Seddig, 2020), the study applied FIXED alignment, employed ML and MLR estimations, and compared the results. Furthermore, if the amount of non-invariant parameters exceeded 25%, a Monte Carlo investigation was carried out to confirm the alignment results (Asparouhov and Muthén, 2014). The results of the exact and approximate approaches were then compared. If the results did not show measurement equivalence, possible causes and alternative solutions were discussed, which may have led to a reduction of the model, the equivalence of which was also assessed.

The assessment of the second-order factor model followed a similar approach. First, the hypothesized second-order factor model was assessed country by country. This preliminary step may suggest excluding countries in which the model did not fit the real data. The analysis of the data collected in the remaining countries began with the assessment of the configural invariance model. The following level of equivalence to be achieved concerned the metric invariance in the first-order factor part of the model, which is the precondition for assessing the metric invariance in the second-order factor model. The next step considered assessing the scalar invariance in the first-order factor model and in the second-order factor model.

The assessment of the second-order factor model has not often been performed, and the methods applied so far have followed the exact approach (Chen et al., 2005; Rudnev et al., 2018). Currently, the implementation of the alignment of this type of model is not supported by any of the software allowing the alignment procedure (R/MPlus). In this study, all analyses were carried out using the software package Mplus Version 8.4 (Muthén and Muthén, 1998).

RESULTS

The results are presented following the order described in the analytical strategy.

Exploratory Factor Analysis and Hypothetical Models

Following the common practice used in the field (Davidov, 2010; Davidov et al., 2014; Brown, 2015; Cieciuch et al., 2019), an exploratory factor analysis (EFA) was conducted on the total sample to explore the structure of solidarity construct as measured by EVS.

In this study, EFA has been performed in the SEM framework and the factor selection has been guided by the model fit

TABLE 1 | Sample size by country.

Country (country code as in EVS 2020)	Sample size
Albania (8)	1,435
Azerbaijan (31)	1,800
Austria (40)	1,644
Armenia (51)	1,500
Bosnia and Herzegovina (70)	1,724
Bulgaria (100)	1,558
Belarus (112)	1,548
Croatia (191)	1,487
Czechia (203)	1,811
Denmark (208)	3,362
Estonia (233)	1,304
Finland (246)	1,199
France (250)	1,870
Georgia (268)	2,194
Germany (276)	2,170
Hungary (348)	1,514
Iceland (352)	1,624
Italy (380)	2,277
Lithuania (440)	1,448
Montenegro (499)	1,003
Netherlands (528)	2,404
Norway (578)	1,122
Poland (616)	1,352
Portugal (620)	1,215
Romania (642)	1,613
Russia (643)	1,825
Serbia (688)	1,499
Slovakia (703)	1,432
Slovenia (705)	1,075
Spain (724)	1,209
Sweden (752)	1,194
Switzerland (756)	3,174
North Macedonia (807)	1,117
Great Britain (826)	1,788
Total	56,491

TABLE 2 | Items in the EVS Measuring Solidarity ($N = 54{,}072$).

	To what extent do you feel concerned about the living conditions of…	Mean	SD	% missing
V212	People in your neighborhood	2.70	1.05	0.88
V213	The people of the region you live in	2.82	1.01	1.06
V214	Your fellow countrymen	2.73	0.99	1.18
V215	Europeans	3.26	1.02	2.00
V216	All humans all over the world	3.06	1.08	2.07
V217	Elderly people	2.01	0.92	0.65
V218	Unemployed people	2.57	1.04	1.02
V219	Immigrants	3.16	1.12	1.70
V220	Sick and disabled people	2.04	0.95	0.85

TABLE 3 | EVS (2017–2020) Measurement of Solidarity: Exploratory Factor Analysis, Summary of Model Fit Information by Number of Factors Extracted (MLR estimation; GEOMIN rotation).

Factors	χ^2	RMSEA	CFI	RMSR
1	52772.84	0.186	0.716	0.097
2	30581.9	0.169	0.835	0.056
3	5412.495	0.089	0.971	0.025
4	No convergence			

χ^2, chi-square; CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean residual.
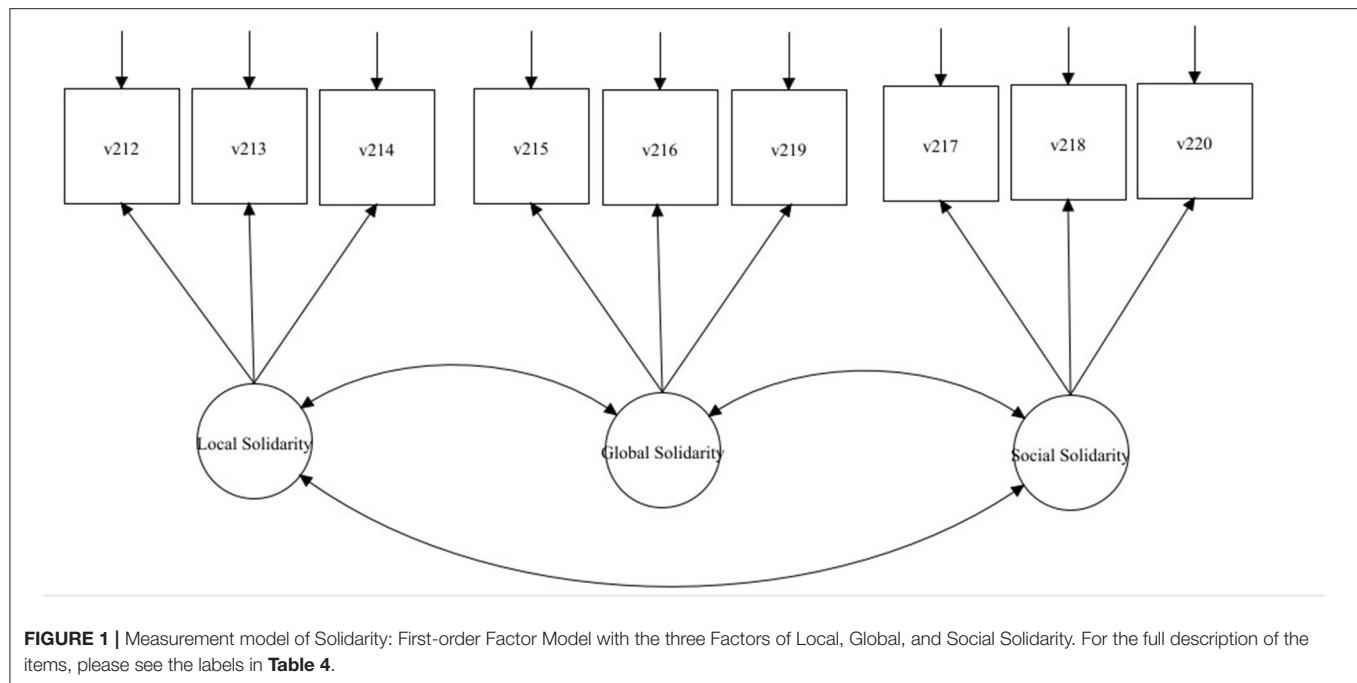
TABLE 4 | EVS (2017–2020) Measurement of Solidarity: Exploratory Factor Analysis, Rotated Loading Matrix (MLR estimation; GEOMIN rotation).

Concern for the living conditions of…	Local solidarity	Global solidarity	Social solidarity
People in your neighborhood (V212)	0.828*	−0.037*	0.008*
The people of the region you live in (V213)	0.992*	0.007*	−0.063*
Your fellow countrymen (V214)	0.558*	0.296*	0.108*
Europeans (V215)	0.204*	0.739*	−0.017*
All humans all over the world (V216)	−0.016*	0.826*	0.125*
Elderly people (V217)	−0.004*	−0.021*	0.840*
Unemployed people (V218)	0.062*	0.134*	0.596*
Immigrants (V219)	0.017*	0.447*	0.301*
Sick and disabled people (V220)	−0.110*	0.001	0.921*

*Significant at 5% level.

comparison of the possible solutions. Goodness of fit statistics allows understanding whether the parameters in the factor models assessed can reproduce the sample correlations. **Table 3** summarizes the model fit measures of models based from one to four factors. While the model with four factors did not converged, the model with three factors displayed better fit than more parsimonious models.

The suggested three-factors model (shown in **Table 4**) is consistent with previous literature (Abela, 2004) that, using data from previous edition of EVS, indicated a similar structure. representing three dimensions of solidarity: global solidarity (i.e., concern for humankind, immigrants, and Europeans); local solidarity (i.e., concern for people living in the neighborhood,

region, and country); and social solidarity (i.e., concern for the elderly, sick, disabled, and unemployed).

These results indicate a three-factor model that represents the relations between the observed indicators (e.g., concern for the living conditions of the elderly, sick, disabled, and unemployed) and the underlying factor (e.g., social solidarity). Furthermore, the three factors (i.e., social, local, and global solidarity) are intercorrelated. In addition to the three-factor model with correlated factors, it would be possible to hypothesize a

FIGURE 1 | Measurement model of Solidarity: First-order Factor Model with the three Factors of Local, Global, and Social Solidarity. For the full description of the items, please see the labels in **Table 4**.

second-order factor model that represents reflective relationships between the three sub-dimensions of social, local, and global solidarity and the underlying second-order concept of solidarity.

What does the two models mean for substantive research? In the first-order factor solution, the model indicates the conceptual notion that the construct of solidarity is multidimensional; therefore, researchers need to take into account this configuration in their substantive analyses. For example, predictor effects as well as the general mechanisms of social change may differ by dimension; these differences may be overlooked if researchers do not account for the multidimensionality of the construct. In the second-order factor solution, the model assumes that the distinct sub-dimensions of social, local, and global solidarity are influenced by the broader concept of solidarity. If testing this model would yield positive results, it would mean that it is possible to measure the general concept of solidarity. Successful results in the measurement invariance assessment, would then indicate that this second-order construct can be compared across countries. The assessments of the two models will be presented separately.

Measurement Equivalence of First-Order Factor Model of Solidarity

The first-order factor model suggested by the results of the EFA (see **Figure 1**) was assessed by confirmatory factor analysis (CFA) across all countries. The modification indexes suggested adding several error correlations. The compromise between the most parsimonious solution and an acceptable model fit resulted in adding the following error correlations: v213- v212; v214- v212; v220- v217; v219-v218; v220-v219; v219- v217 (CFI = 0.988, RMSEA = 0.060, SRMR = 0.022). This baseline model was then tested country by country to assess whether the baseline

TABLE 5 | EVS (2017–2020) Measurement of Solidarity. First-order Factor Model with three Factors, MGCFA Model Fit Across 32 Countries (N = 53,673).

Model	χ^2 (df)	RMSEA	CFI	SRMR
Configural	4210.642 (576)	0.061	0.979	0.030
Metric	5821.496 (762)	0.063	0.971	0.052
Scalar	18272.591 (948)	0.105	0.900	0.098
Partial scalar (v215, v217, v212 free)	10834.557 (855)	0.084	0.942	0.078

χ^2, chi-square; df, degrees of freedom; CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean residual.

model was able to fit the data of each country separately. By observing the fit statistics, the model was rejected for Iceland and Sweden, which were therefore excluded from further analyses. The test for measurement invariance included the remaining 32 countries. **Table 5** reports the results of MGCFA. The configural model fit was good, which meant that the measurement structure was equivalent across countries. The next step was to assess metric invariance, which also presented good fit measures; despite a slight worsening compared with the configural model, the overall fit was still acceptable. However, the assessment of scalar invariance was unsuccessful. As indicated by previous literature (Byrne et al., 1989; Steenkamp and Baumgartner, 1998; Pokropek et al., 2019), partial invariance could offer meaningful mean comparison. To test for the partial solution, it is possible to release the constraints if at least two intercepts for each construct are kept equal (Byrne et al., 1989; Steenkamp and Baumgartner, 1998; Brown, 2015). In the model of solidarity with three dimensions reflected by three items, this means that only

TABLE 6 | EVS (2017–2020) Measurement of Solidarity. First-order Factor Model with three Factors, Noninvariant Parameters in the Alignment Analysis (type = fixed; N = 53673).

Estimator	Factor loadings		Intercepts		Factor loadings and intercepts	
	Absolute (288)	Relative	Absolute (288)	Relative	Absolute (576)	Relative
Maximum likelihood (ML)	75	26.0%	139	48.3%	214	37.2%
Robust maximum likelihood (MLR)	59	20.5 %	136	47.2 %	195	33.9 %

one parameter per item could be freely estimated. Following a systematic review of the modification indexes, which indicate the reduction of the Chi-square value that could be reached if the constrained parameter is released, intercepts were sequentially released to improve the fit. After running alternative models to evaluate further which solution would offer the best fit, the equality constraint was released for three intercepts (v215, v217, v212), one for each construct. However, according to the common criteria used to interpret the change in model fit (ΔCFI < 0.010; ΔRMSEA < 0.015 or ΔSRMR = 0.030) the model fit reached was not enough good to be acceptable because the overall fit worsened too much from the metric model (ΔCFI = 0.029; ΔRMSEA = 0.021; ΔSRMR = 0.026). The partial scalar model was therefore rejected. This means that factor means of solidarity, intended as a three-dimensional construct, cannot be compared.

However, it is worth to assess equivalence using the alignment method to consider whether the solution suggested by employing the approximate equivalence approach goes in the same direction of that one offered by the exact approach.

Following the approximate approach, the FIXED alignment was applied. Because there was no substantive reason to guide the selection of the reference group for the estimation, the Mplus default command that selects the first group of the list was used. In this case, Albania was used as the reference group, and its mean was fixed equal to zero. **Table 6** shows the number of non-invariant parameters in the ML and MLR solutions.

Even if MLR estimation appears to provide better results, the amount of non-invariant parameter is too large to consider trustworthy the factor mean comparison. While the recommended criterion is that the amount of non-invariance should not exceed 25% (Asparouhov and Muthén, 2014), ML and MLR solutions present 37.2 and 33.9% of non-invariant parameters, respectively. Monte Carlo simulations have been performed to evaluate whether the mean comparison would be trustworthy despite the high amount of non-invariance. In both cases, the estimates of the final model were used as starting values for the simulation study, which was aimed at testing the precision of the replication of the factor means. The specifications of the Monte Carlo investigation included a country sample size of 1,500 to simulate the average sample size of the real data collected from the EVS and the 500 replications of each simulation. The key issue in the Monte Carlo investigation conducted in the present study was the replicability of the factor mean ranking. As suggested by Muthén and Asparouhov (2014, p. 978) to consider the alignment results trustworthy, the required correlation must be at least 0.98. The correlation of the generated and estimated factor means using ML and MRL estimation were extremely

high for the factors reflecting local solidarity (>0.997) and social solidarity (>0.996) in both the estimations. However, this was not the case for the factor reflecting global solidarity, where the correlations are below the cut-off criteria both using ML and MLR estimators (<0.965).

This result indicates that, with this model configuration, the factor means cannot be fully compared in a meaningful way across the 32 countries and suggests that the factor global solidarity could be the least invariant part of the model.

The results of MGCFA and alignment suggest quite close final consideration. From the exact equivalence perspective, partial scalar invariance was not obtained even releasing intercepts, suggesting that mean comparison could not provide reliable results. Neither approximate equivalence was not achieved because of the lack of replicability of the global solidarity factor estimates. Taking these results altogether, mean comparison across the full set of countries is discouraged.

The latent concept of solidarity, which was configured as a first-order factor model of the three correlated dimensions of social, local, and global solidarity, was not invariant across the 32 countries considered. However, the MGCFA results indicated the achievement of metric invariance, allowing for the comparison of factor variances, covariances, and unstandardized regression coefficients across groups. The results of the approximate equivalence assessment, which usually offers solutions when the exact approach fails (Davidov et al., 2008, 2014), suggested that the model was non-invariant and that factor means could not be compared.

What could be done at this point? One possibility would be to reduce the non-invariance by reducing the number of groups and therefore intensively analyze the modification indexes to identify the less invariant groups. This would require an extensive procedure that may not be successful or that would reduce the number of countries involved in the assessment to those that were not in the interest of the researchers (Lomazzi, 2018). Cross-cultural comparisons would then be strongly limited. Another option would be to reduce the source of variance by reducing the model. Both MGCFA and alignment results could be used for diagnostic purposes. In the first case, the in-depth analysis of the modification indexes could be particularly demanding when many groups are involved. It would consist of observing which parameters introduced more variance in the model in each country. According to this investigation, items v215 (concerned about the living conditions of Europeans) and v219 (concerned about the living conditions of immigrants) appeared to be the most problematic in terms of invariance. In the alignment method, the analysis of "troublemakers" parameters is generally

TABLE 7 | EVS (2017–2020) Measurement of Solidarity. First-order Factor Model with two Factors, Alignment Output: Noninvariant Parameters.

Items	Intercepts	Factor loadings		
		Local Solidarity	Global Solidarity	Social Solidarity
People in your neighborhood (V212)	17	9		
The people of the region you live in (V213)	14	2		
Your fellow countrymen (V214)	17	10		
Europeans (V215)	19		5	
All humans all over the world (V216)	20		6	
Immigrants (V219)	20		10	
Elderly people (V217)	8			4
Unemployed people (V218)	16			10
Sick and disabled people (V220)	4			3

more was straightforward than in MGCFA because the alignment output clearly displays non-invariant parameters among factor loadings and intercepts by country and this facilitated the analysis of non-invariant parameters, which are reported in **Table 7**.

Because invariance at the intercept level affected the achievement of full equivalence, the number of non-invariant parameters here was of particular interest. The items belonging to the factor of global solidarity presented the highest number of non-invariant intercepts across all countries. This result was in line with the indication given by the observation of the modification indexes of the exact approach, which was described earlier. Unfortunately, the statistical tools used to assess measurement equivalence are not enough to further investigate and detect the sources of non-invariance. Statistically, techniques as multilevel structural equation modeling could be of help for this additional scope (Davidov et al., 2016; Seddig and Lomazzi, 2019). For more substantive purposes, techniques such as those used in the context of cognitive interviews would be extremely useful (Behr et al., 2017; Meitinger, 2017). However, the more problematic items (v215 and v219) are also substantively challenging, and the source of bias could have been rooted in the different feelings of closeness with Europeans (v215) and in the interpretation of the word "immigrants" (v219). Interestingly, difficulties with the item on being concerned about the living conditions of Europeans was not problematic for people living in countries that were not European Union (EU) member states (MS) or that had joined the EU recently. This was also the case in founding countries, such as Italy and France, and in other MS, such as Portugal, Denmark, and Austria (details are provided in **Table 8**). This result may be related to the controversial state of transnational solidarity between EU MS and the rise of anti-Europeanism (Grimmel and Giang, 2017; Ettore and Adrian, 2019; Reinl, 2020). Even if from another perspective, the item investigating concern about the living condition of immigrants could also be differently interpreted cross-culturally. In reacting to the object "immigrants," respondents might refer to the generic concept of immigrant or to the migrant group that is numerically dominant in their countries (which could therefore be different according to socio-economic and ethnic background) or to groups in the immigrant population that were not comparable, such as migrant workers or refugees (Braun

et al., 2012; Blinder, 2015). While these aspects would require further investigation in *ad hoc* studies, such as those based on cognitive interviews, the current study aimed to propose a viable solution for comparing solidarity constructs across the countries included in the fifth wave of the EVS. As the Monte Carlo investigation anticipated, a large share of the non-invariance issues seems to derive from the global solidarity factor. So, it might be possible to assess the measurement equivalence of a reduced model of solidarity based on only the factors of local and social solidarity.

Equivalence Testing of the Reduced First-Order Factor Model

The same strategy that was applied in the equivalence assessment of the three-factor model was replicated in the model that was reduced to the two correlated factors of local and social solidarity (**Figure 2**). The baseline model reached a good model fit when three error correlations were added: v213-v212; v214-v212; v220-v217 (CFI = 0.996, RMSEA = 0.050, SRMR = 0.011). The baseline model was tested country by country using CFA. The model did not fit the data on Bosnia Herzegovina and North Macedonia, which were then excluded from the analysis. The results of the assessment using MGCFA (see **Table 9**) indicated that configural and metric invariance were achieved. Full scalar invariance could not be reached in the reduced model. Considering that at least two indicators par constructs must be held equal (Steenkamp and Baumgartner, 1998; Davidov, 2009), testing for partial invariance resulted in releasing two intercepts (one for each construct). The partial scalar model showed a good fit, but the overall fit deteriorated too much compared with the metric model (ΔCFI = 0.022; ΔRMSEA = 0.027 ΔSRMR = 0.08). Thus, the partial scalar model was rejected.

Measurement equivalence was then assessed employing the FIXED alignment method with ML and MLR estimations (**Table 10**).

Compared with the ML solution, the MLR estimation showed slightly better results (24.2% over 25.5% of non-invariant parameters). The total amount of non-invariance was below the cut-off criteria of 25% (Asparouhov and Muthén, 2014). However, because it was close to the threshold, Monte Carlo simulations were conducted to confirm the results of the ML

TABLE 8 | EVS (2017–2020) Measurement of Solidarity. First-order Factor Model: Noninvariant Parameters in the Alignment Analysis (type = fixed; N = 53673).

Country code	8	31	40	51	70	100	112	191	203	208	233	246	250	268	276	348	380	440	499	528	578	616	620	642	643	688	703	705	724	756	807	826	Tot.
Intercepts																																	
V212	*	*			*	*	*	*	*		*	*		*	*		*	*	*	*				*				*	*			*	17
V213	*	*				*			*	*	*	*		*						*		*				*		*	*			*	14
V214	*	*				*	*	*	*	*	*	*								*	*	*		*					*		*	*	17
V215	*	*	*			*	*	*	*	*	*		*	*		*		*		*		*		*		*	*					*	19
V216		*	*		*	*	*		*		*	*		*	*		*		*	*	*	*	*	*		*		*	*			*	20
V217	*				*			*				*		*			*									*					*		8
V218		*	*	*	*			*	*	*		*	*			*		*				*		*		*	*	*			*	*	16
V219	*	*	*	*	*	*	*		*		*	*		*	*	*				*		*		*	*	*			*			*	20
V220										*																		*		*	*	*	4
Loadings for F1																																	
V212		*	*							*										*		*		*		*		*		*			9
V213									*						*																		2
V214		*			*					*			*	*								*		*		*	*	*					10
Loadings for F2																																	
V215		*							*					*							*		*										5
V216		*								*					*													*	*			*	6
V219		*		*		*			*	*				*		*												*	*			*	10
Loadings for F3																																	
V217				*			*			*																					*		4
V218		*								*			*		*			*						*	*					*	*	*	10
V220												*		*				*															3
Total	12	19	10	5	12	10	11	11	14	15	14	13	9	10	9	8	7	11	4	11	4	10	3	13	7	10	9	11	12	7	10	17	194

*Full variable labels list is available in **Table 4**. Numeric country codes as listed in **Table 1**.*

FIGURE 2 | Measurement model of Solidarity: First-order Factor Model with the two Factors of Local, and Social Solidarity. For the full description of the items, please see the labels in **Table 4**.
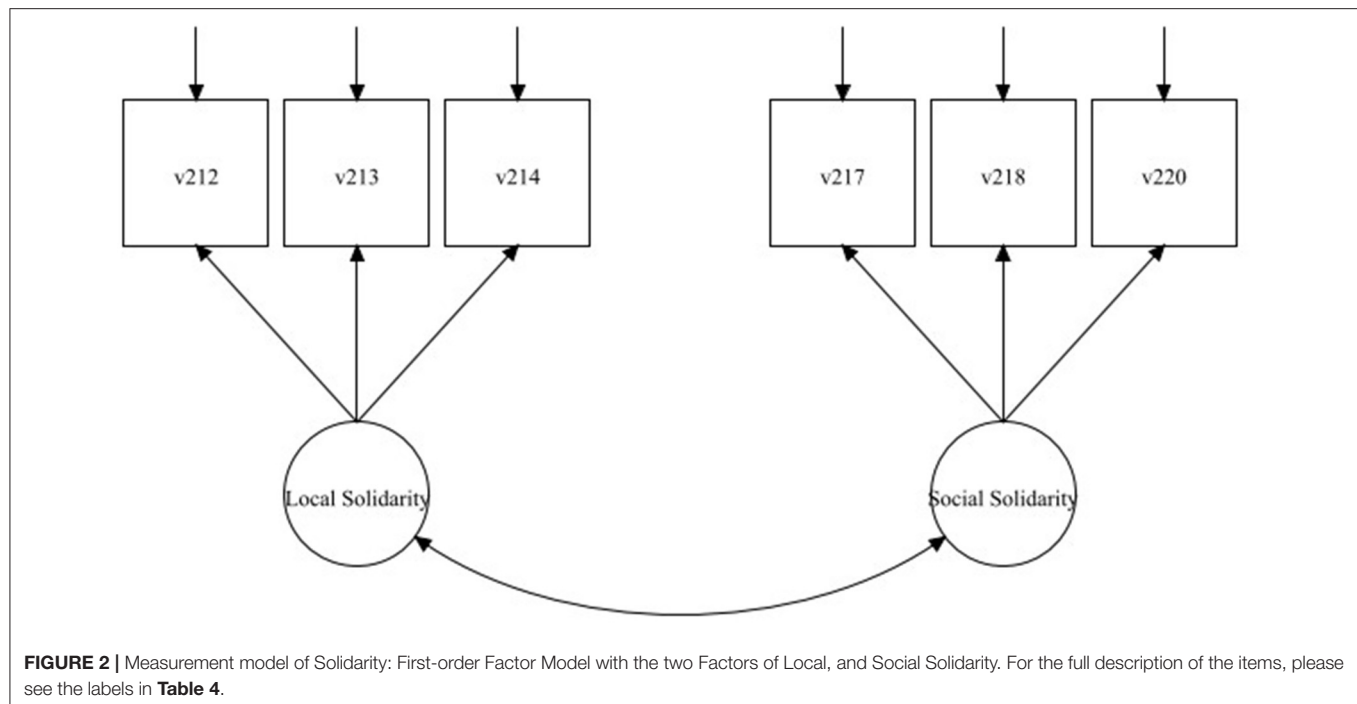
TABLE 9 | EVS (2017–2020) Measurement of Solidarity. Reduced First-order Factor Model: MGCFA Model Fit Across 32 Countries (N = 53650).

Model	χ^2 (df)	RMSEA	CFI	SRMR
Configural	646.758 (160)	0.043	0.995	0.011
Metric	1347.318 (284)	0.047	0.989	0.035
Scalar	7332.612 (408)	0.101	0.929	0.066
Partial scalar (v218, v212 free)	3545.671 (346)	0.074	0.967	0.043

χ^2, chi-square; df, degrees of freedom; CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean residual.

and MLR alignment solutions, using the same specification as in the previous Monte Carlo investigation performed with the three-factors model. As suggested by Muthén and Asparouhov (2014, p. 978) to consider the alignment results trustworthy, the required correlation must be at least 0.98. The correlation of the generated and estimated factor means using ML and MRL estimation were extremely high (>0.997) in both cases, which confirmed the trustworthiness of the alignment results.

Although the MGCFA results supported only metric invariance, by applying the concept of approximate equivalence, the factor means could be compared. As previous empirical research pointed out (Asparouhov and Muthén, 2014; Flake and McCoach, 2018; Pokropek et al., 2019), the alignment results are particularly accurate and can be trusted when the amount of non-invariance is low or moderate. In this study, the non-invariance was about 24% and the Monte Carlo investigation conducted confirmed the accuracy of the latent mean estimations and country ranking. Therefore, scholars

could compare the factor means of the construct of solidarity, which was intended as a bidimensional concept of the correlated dimensions of social and local solidarity, across 32 countries. However, it is important to note that latent factor means differ from means computed by raw scores. Although it is often used, this practice neglects measurement errors and the accumulation of (not estimated) error terms, which are instead included in the latent means approach, which provides a more accurate measurement (Rose et al., 2019). The results obtained by the two approaches were substantially different. **Table 11** displays the deviations in the country ranking for the factor of local solidarity and the factor of social solidarity by using latent means (i.e., the MLR alignment output) and raw score means. Differences in the ranking were more evident in the social solidarity factor. While the first parts of the ranking, which included countries that expressed higher social solidarity, does not vary much across approaches, the middle and lowest parts of the ranking showed relevant substantive discrepancies. If scholars used raw score means without accounting for measurement errors, Iceland would have been ranked in 10th place, followed by Serbia, Belarus, Romania, and Slovenia. However, if the more accurate measurement of latent means were considered, the order would be different. Iceland would be in 12th place, and Belarus and Slovenia would show higher solidarity than Iceland did. Similarly, the comparison of social solidarity in Austria and Armenia would have been particularly biased.

The dimension of local solidarity was slightly less affected, but in this case, the comparison between some countries would yield distinctive substantive results, especially in the first part of the ranking, where countries with higher levels of expressed local solidarity would be displayed in a different order. For example,

TABLE 10 | EVS (2017–2020) Measurement of Solidarity. Reduced First-order Factor Model, Noninvariant Parameters in the Alignment Analysis (type = fixed; N = 53650).

	Factor loadings		Intercepts		Factor loadings and intercepts	
Estimator	Absolute (192)	Relative	Absolute (192)	Relative	Absolute (384)	Relative
Maximum likelihood (ML)	19	9.9%	79	41.1%	98	25.5%
Robust maximum likelihood (MLR)	15	7.8%	78	40.6%	93	24.2%

TABLE 11 | EVS (2017–2020) Measurement of Solidarity. Comparison of Country Ranking (Descending Order of Solidarity) by Latent Factor Means and Raw Score Means.
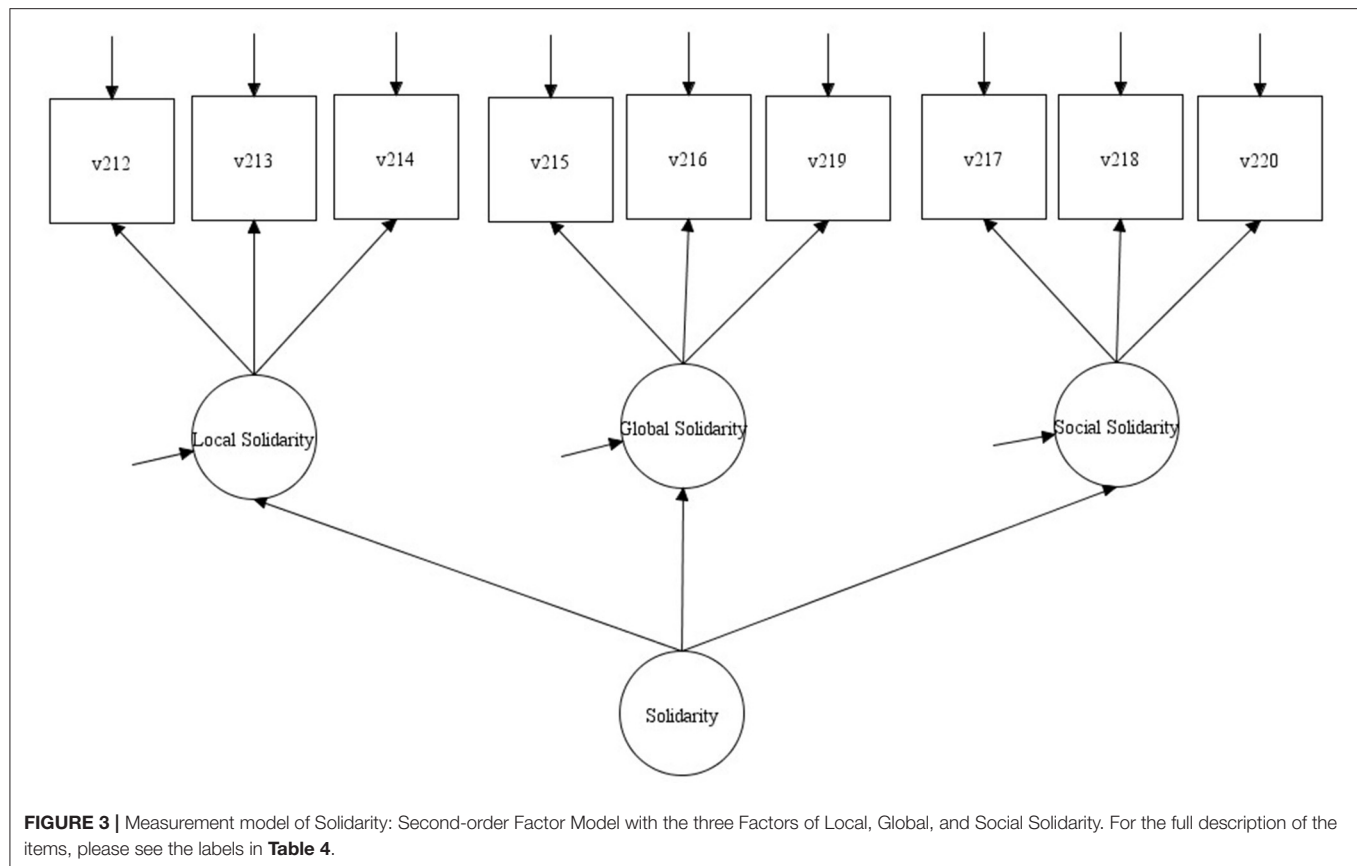
	Local solidarity		Social solidarity	
Rank	*Countries ordered by factor mean (alignment)*	*Countries ordered by raw score mean*	*Countries ordered by factor mean (alignment)*	*Countries ordered by raw score mean*
1	Finland (0.606)	Finland (3.142)	Netherlands (1.771)	Netherlands (2.756)
2	Estonia (0.472)	Estonia (3.142)	Hungary (1.743)	Hungary (2.726)
3	Poland (0.438)	Serbia (3.100)	Estonia (1.61)	Estonia (2.661)
4	Hungary (0.438)	Hungary (3.096)	Lithuania (1.56)	Lithuania (2.603)
5	Serbia (0.431)	Great Britain (3.088)	Russia (1.372)	Russia (2.541)
6	Netherlands (0.407)	Poland (3.076)	Poland (1.32)	Poland (2.495)
7	Belarus (0.398)	Belarus (3.073)	Finland (1.247)	Slovakia (2.439)
8	Russia (0.389)	Russia (3.066)	Slovakia (1.202)	Finland (2.394)
9	France (0.375)	Croatia (3.047)	Azerbaijan (1.18)	Azerbaijan (2.388)
10	Great Britain (0.368)	Netherlands (3.031)	Belarus (1.131)	Iceland (2.369)
11	Croatia (0.357)	France (3.028)	Slovenia (1.12)	Serbia (2.356)
12	Slovenia (0.325)	Slovenia (2.977)	Iceland (1.112)	Belarus (2.348)
13	Romania (0.275)	Romania (2.961)	Serbia (1.058)	Romania (2.337)
14	Iceland (0.242)	Slovakia (2.931)	Switzerland (1.052)	Slovenia (2.315)
15	Slovakia (0.234)	Iceland (2.899)	Romania (1.03)	Switzerland (2.311)
16	Lithuania (0.122)	Azerbaijan (2.834)	Sweden (0.929)	Czechia (2.310)
17	Azerbaijan (0.12)	Lithuania (2.820)	Italy (0.872)	Sweden (2.248)
18	Italy (0.107)	Italy (2.806)	France (0.816)	France (2.214)
19	Sweden (0.038)	Albania (2.676)	Norway (0.788)	Great Britain (2.188)
20	Albania (0)	Sweden (2.674)	Great Britain (0.740)	Italy (2.183)
21	Norway (−0.04)	Norway (2.662)	Czechia (0.722)	Norway (2.164)
22	Denmark (−0.067)	Denmark (2.630)	Bulgaria (0.712)	Bulgaria (2.145)
23	Bulgaria (−0.077)	Bulgaria (2.628)	Croatia (0.554)	Austria (2.033)
24	Switzerland (−0.142)	Switzerland (2.580)	Spain (0.514)	Croatia (2.027)
25	Czechia (−0.164)	Czechia (2.564)	Montenegro (0.513)	Germany (1.964)
26	Portugal (−0.185)	Portugal (2.548)	Armenia (0.481)	Denmark (1.954)
27	Spain (−0.341)	Spain (2.415)	Austria (0.46)	Montenegro (1.939)
28	Germany (−0.469)	Germany (2.283)	Denmark (0.459)	Spain (1.928)
29	Montenegro (−0.567)	Montenegro (2.226)	Germany (0.403)	Portugal (1.915)
30	Georgia (−0.588)	Georgia (2.199)	Portugal (0.339)	Armenia (1.771)
31	Austria (−0.69)	Austria (2.109)	Albania (0)	Albania (1.669)
32	Armenia (−0.938)	Armenia (1.891)	Georgia (−0.463)	Georgia (1.452)

Numbers in brackets indicate the means obtained through the two methods. Latent factor means were estimated in the alignment method using the country as a reference, whose mean is fixed equal to zero (in this case, Albania). Raw score means were computed using the scores of each item (1–5). In both cases, items were recoded so that higher values indicated higher solidarity.

the comparison of Poland and Great Britain would have shown distinctive results if the scholars had used raw score means instead of latent means, which were included in the estimate of error terms.

Measurement Equivalence of Second-Order Factor Model of Solidarity

The next step in the analysis consisted of testing the second-order factor model of solidarity and its comparability across

FIGURE 3 | Measurement model of Solidarity: Second-order Factor Model with the three Factors of Local, Global, and Social Solidarity. For the full description of the items, please see the labels in **Table 4**.

countries (**Figure 3**). The MGCFA results of the first-order factor model did not encourage optimism, but a different structure may have led to different results. If invariance were achieved, it would mean that scholars could compare countries by using a single latent mean of solidarity, which would ease the interpretation of substantive results. Additionally, it would also provide insights into the theoretical conceptualization of solidarity, which could be considered an overarching construct reflected in multiple dimensions.

Following the procedure suggested by Rudnev et al. (2018), one marker indicator for each factor was selected to identify the model and allow the estimation. By observing the modification indexes, items displaying the greatest invariance were selected (v214, v216, and v217). The baseline model had a poor model fit (CFI = 0.894, RMSEA = 0.114, SRMR = 0.087). However, when tested country by country, the second-order factor model fit the data of Bosnia Herzegovina, Croatia, Estonia, Georgia, Montenegro, Switzerland, and North Macedonia (**Table 12**).

Although this result indicated that, in each of these seven countries, the construct of solidarity responded to a hierarchical conceptualization, it still does not mean that comparisons of the level of overall solidarity across these seven countries would be meaningful. Following the bottom-up strategy of assessing equivalence (Chen et al., 2005; Rudnev et al., 2018), in the first step, configural invariance was assessed. The model fit

statistics indicate acceptable fit (CFI = 0.932, RMSEA = 0. 094, SRMR = 0.056). This means that configural invariance of the second-order factor model of solidarity was achieved and that the measurement pattern of latent constructs is the same across these countries. The assessment of metric invariance in second-order factor models requires two steps: only after the positive evaluation of metric invariance in the first-order order part of the model, it is possible to test for the metric equivalence of the first- and second-order parts together (Rudnev et al., 2018). **Table 13** includes the results of the first step of the metric equivalence assessment. The model fit statistics were acceptable (CFI = 0.908, RMSEA = 0. 099, SRMR = 0.090), but the overall model fit worsened too much from the configural model fit (ΔCFI = 0.024, ΔRMSEA = 0. 005, ΔSRMR = 0.034). For this reason, the metric invariance of the first-order part of the model was rejected and the following steps of the equivalence assessment were not conducted. This means that, assuming solidarity as a hierarchical concept reflected in the three dimensions of local, global, and social solidarity, only a qualitative comparison of the construct of solidarity across seven countries can be carried out. In other words, scholars can only observe that the same indicators measure the same theoretical concept in these seven countries, but deeper comparisons concerning for example which of these countries displays higher or lower levels of solidarity cannot be meaningfully conducted.

TABLE 12 | Second-order Factor Model, Confirmatory Factor Model Country by Country.

	Chi2	RMSEA	CFI	SRMR
Albania	474.279	0.108	0.903	0.101
Azerbaijan	725.318	0.120	0.910	0.081
Austria	704.840	0.124	0.895	0.086
Armenia	496.713	0.108	0.851	0.103
Bosnia and Herzegovina	431.227	0.093	0.922	0.100
Bulgaria	505.544	0.107	0.911	0.086
Belarus	552.465	0.112	0.894	0.091
Croatia	318.531	0.085	0.932	0.074
Czechia	950.430	0.138	0.852	0.106
Denmark	2321.651	0.159	0.820	0.109
Estonia	361.499	0.098	0.919	0.097
Finland	599.111	0.133	0.844	0.084
France	1138.709	0.148	0.823	0.121
Georgia	349.443	0.074	0.934	0.064
Germany	925.039	0.124	0.859	0.113
Hungary	497.072	0.107	0.88	0.099
Iceland	896.942	0.142	0.866	0.083
Italy	750.893	0.109	0.900	0.122
Lithuania	806.971	0.141	0.860	0.099
Montenegro	180.395	0.075	0.940	0.064
Netherlands	1199.395	0.135	0.831	0.088
Norway	488.300	0.124	0.862	0.085
Poland	523.457	0.117	0.899	0.101
Portugal	888.188	0.162	0.826	0.147
Romania	774.902	0.131	0.869	0.127
Russia	638.335	0.111	0.890	0.095
Serbia	576.343	0.117	0.898	0.098
Slovakia	1010.803	0.16	0.825	0.147
Slovenia	393.829	0.112	0.869	0.091
Spain	428.232	0.111	0.915	0.111
Sweden	883.854	0.164	0.817	0.102
Switzerland	984.252	0.106	0.901	0.062
North Macedonia	289.539	0.093	0.931	0.055
Great Britain	1117.742	0.15	0.802	0.107

χ^2, chi-square; CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean residual.

TABLE 13 | EVS (2017–2020) Measurement of Solidarity. Second-order Factor Model: MGCFA Model Fit Across 7 Countries (N = 12003).

Model	χ^2 (df)	RMSEA	CFI	SRMR
Configural	2152.889 (168)	0.09	0.937	0.056
Metric invariance of the first-order factors	3027.675 (207)	0.096	0.91	0.094

χ^2, chi-square; df, degrees of freedom; CFI, comparative fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean residual.

SUMMARY AND DISCUSSION

This study examined relevant aspects of the measurement and the comparison of socio-political attitudes and values. It focused on *what* types of measurement approaches and models are appropriate in measuring this kind of latent concept, indicating that the reflexive approach, which implies factorial modeling, is preferable over combinatory logic to be able of accounting for the respondent's perspective (van Vlimmeren et al., 2016). Following this approach, the study pointed to substantive differences in the use of first- and second-order factor models to measure latent constructs. After reflecting on measuring values and attitudes, the paper discussed their comparability across different groups and *why* assessing measurement equivalence is necessary to achieve meaningful comparisons. Taking for granted the assumption of comparability and neglecting that the concept could be measured differently by the same instrument in different groups would lead to theoretical generalizations and conclusions based on biased results. The answer to the question of *when* assessing invariance is necessary is straightforward: whenever scholars wish to compare latent concepts across different groups. The focus of the present study was on cross-national comparisons, but the arguments presented would be valid in dealing with longitudinal issues, assessing invariance across different modes of data collection or sampling designs, and confronting situations where methodological or cultural aspects challenge the assumption of comparability (van de Vijver and Tanzer, 2004).

The study then addressed *how* to assess measurement invariance. Two main perspectives in the field of measurement equivalence were described. In contrast to the exact approach to equivalence, which is mainly carried out using techniques such as multigroup confirmatory factor analysis (MGCFA), the recent approach of approximate equivalence offers solutions to the limitations of strictness in the exact approach. In dealing with a large number of groups, one seldom reaches scalar when using MGCFA (Davidov et al., 2008, 2014). New techniques that were mainly developed in the Bayesian framework have introduced the idea that a small amount of non-invariance is acceptable. From this perspective, the alignment method, which can also be applied in the frequentist approach, is a viable alternative to MGCFA when the assessment involves many groups.

To illustrate the empirical application of the described theoretical answers to the questions of what, why, when, and how to assess measurement equivalence, the study investigated the construct of solidarity captured by the last wave of the European Values Study (EVS, 2017–2020). The concept of solidarity is measured by the EVS through a battery of nine items reflecting three dimensions of solidarity: toward people considered in need (i.e., social solidarity); toward people living close to the respondent (i.e., local solidarity); and toward people living in another country, migrants, or humankind in general (i.e., global solidarity). Two measurement models were hypothesized: a first-order factor model, in which the three independent dimensions of solidarity were correlated; and a second-order factor model, in

TABLE 14 | EVS (2017–2020) Measurement Equivalence of Solidarity. Summary of Results.

Model	Method	Results	Number of countries
First-order factor model with three dimensions (social, local, global solidarity)	MGCFA	Metric invariant	32 countries
	Alignment	Non-invariant (global solidarity factor means not trustworthy)	32 countries
First-order factor model with two dimensions (social, local)	MGCFA	Metric invariant	32 countries
	Alignment	Invariant	32 countries
Second-order factor model with three dimensions (social, local, global solidarity)	MGCFA	Configural invariant	7 countries

which solidarity was conceptualized according to a hierarchical principle. In this case, the construct of solidarity was reflected in the three sub-factors. Measurement invariance was assessed in both models and **Table 14** reports a summary of the analyses carried out and results.

The application of MGCFA did not achieve full equivalence in the first-order factor model. However, the metric model was acceptable. This result indicates that scholars could use the first-order factor model, which comprises local, global, and social solidarity, to compare covariances or unstandardized regression coefficients across the 32 countries included in the analysis; however, the model did not fit the data on Iceland and Sweden. The alignment method was applied to assess whether the approximate approach could circumvent the problem of full invariance. However, as other studies have already found (Davidov et al., 2018a), methods in the approximate approach "do not do magic" when the differences in measurement parameters are too relevant. In the present study, the amount of non-invariance (33.9%) was too large to be considered acceptable and trust the mean estimation (Asparouhov and Muthén, 2014; Flake and McCoach, 2018; Pokropek et al., 2019). The Monte Carlo investigation conducted confirmed the rejection of this model.

The in-depth investigation of the modification indexes and the alignment results indicated that the two items of concern for immigrants and Europeans, both of which were in the dimension of global solidarity, were particularly non-invariant. In the attempt to find a model that allowed for the cross-national comparison of solidarity, the first-order model was reformulated to include only two dimensions (i.e., social and local solidarity). The reduced model fit the data well, except for North Macedonia and Bosnia and Herzegovina, which were then excluded from further analysis. The MGCFA results indicated the achievement of metric invariance, whereas the results of the partial scalar model were not optimal. However, the alignment indicated that the estimated factor means were trustworthy and the Monte Carlo simulations confirmed the accuracy in the recovery of means and group ranking, so the country comparisons were meaningful. However, the results of the study indicate that attention is needed when the means are used in further analysis. Scholars often compute factor means using the raw scores of items. However, this practice neglects the error terms and provides inaccurate measurements that, as shown in this study, yield substantive

results that differ from the results of analyses using latent factor means, which include the estimation of error terms in the measurement.

In the final step of the analysis, the second-order factor model was assessed. The hypothesized model fit the data only in a few countries (i.e., Bosnia Herzegovina, Croatia, Estonia, Georgia, Montenegro, Switzerland, and North Macedonia). However, even when the assessment was limited in this subgroup, only configural invariance was achieved. This result indicated that comparisons of solidarity assuming a second-order factor model can be meaningful only if it concerns a qualitative comparison of the construct of solidarity across the seven countries included in the analysis, whereas comparison at any other level should be avoided. However, single-country studies focusing on one of these seven countries could be conducted using the second-order model.

Although this study provided a theoretical overview that supported measurement equivalence testing and showed results that could be useful in the comparative study of solidarity, some aspects require further investigation using *ad hoc* analysis. Particular attention should be paid to the sources of non-invariance in the first-order factor model according to three dimensions (i.e., social, local, and global). The exclusion of the dimension of global solidarity, which comprised aspects of solidarity with immigrants and Europeans, was a major loss because these two aspects are, and will be, particularly interesting in the context of the crises faced by European countries (Delanty, 2008; Ferrera, 2014; Grimmel and Giang, 2017; Agustín and Jørgensen, 2018). The use of techniques such as cognitive web probing could be helpful in determining where the problem in these items arises (Braun et al., 2012; Meitinger, 2017). The results might be relevant for fixing the issue in the model structure, such as introducing a method factor to control for the invariance in the model or informing the improvement of the question design.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: EVS (2017–2020). European Values Study 2017: Integrated Dataset (EVS 2017). GESIS Data Archive, Cologne. ZA7500 Data file Version 4.0.0, doi: 10.4232/1.13560.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

REFERENCES

Abela, A. M. (2004). "Solidarity and religion in the european union: a comparative sociological perspective," in *The Value (s) of a Constitution for Europe*, ed. P. Xuereb (Malta: European Documentation and Research Centre, University of Malta.), 71–101.

Agustín, Ó. G., and Jørgensen, M. B. (2018). "From refugee crisis to a crisis of solidarity?," in *Solidarity and the "Refugee Crisis" in Europe*, eds Ó. G. Agustín and M. B. Jørgensen (Cham: Springer International Publishing), 1–21.

Alemán, J., and Woods, D. (2016). Value orientations from the world values survey how comparable are they cross-nationally? *Comp. Polit. Stud.* 49, 1039–1067. doi: 10.1177/0010414015600458

Amnå, E., and Ekman, J. (2014). Standby citizens: diverse faces of political passivity. *Eur. Polit. Sci. Rev.* 6, 261–281. doi: 10.1017/S175577391300009X

Asparouhov, T., and Muthén, B. (2014). Multiple-group factor analysis alignment. *Struct. Equ. Model. Multidiscip. J.* 21, 495–508. doi: 10.1080/10705511.2014.919210

Bakker, T. P., and de Vreese, C. H. (2011). Good news for the future? Young people, internet use, and political participation. *Commun. Res.* 38, 451–470. doi: 10.1177/0093650210381738

Behr, D., Meitinger, K., Braun, M., and Kaczmirek, L. (2017). *Web probing - implementing probing techniques from cognitive interviewing in web surveys with the goal to assess the validity of survey questions (Version 1.0)*. Mannheim: GESIS - Leibniz-Institut für Sozialwissenschaften. Available online at: https://doi.org/10.15465/gesis-sg_en_023.

Blinder, S. (2015). Imagined immigration: the impact of different meanings of 'immigrants' in public opinion and policy debates in Britain. *Polit. Stud.* 63, 80–100. doi: 10.1111/1467-9248.12053

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annu. Rev. Psychol.* 53, 605–634. doi: 10.1146/annurev.psych.53.100901.135239

Branisa, B., Klasen, S., Ziegler, M., Drechsler, D., and Jütting, J. (2014). The institutional basis of gender inequality: the Social Institutions and Gender Index (SIGI). *Fem. Econ.* 20, 29–64. doi: 10.1080/13545701.2013.850523

Braun, M. (2009). "The role of cultural contexts in item interpretation: the example of gender roles," in *The International Social Survey Programme, 1984-2009 : Charting the Globe*, eds M. Haller, R. Jowell, and T. W. Smith (London; New York, NY: Routledge), 395–408.

Braun, M., Behr, D., and Kaczmirek, L. (2012). assessing cross-national equivalence of measures of xenophobia: evidence from probing in web surveys. *Int. J. Public Opin. Res.* 25, 383–395. doi: 10.1093/ijpor/eds034

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research*, 2nd Edn. New York, NY: Guilford Press.

Busse, B., Hashem-Wangler, A., and Tholen, J. (2015). Two worlds of participation: young people and politics in Germany. *Sociol. Rev.* 63, 118–140. doi: 10.1111/1467-954X.12265

Byrne, B. M., Shavelson, R. J., and Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance. *Psychol. Bull.* 105, 456–466. doi: 10.1037/0033-2909.105.3.456

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model.* 14, 464–504. doi: 10.1080/10705510701301834

Chen, F. F., Sousa, K. H., and West, S. G. (2005). Teacher's corner: testing measurement invariance of second-order factor models. *Struct. Equ. Model. Multidiscip. J.* 12, 471–492. doi: 10.1207/s15328007sem1203_7

Cieciuch, J., Davidov, E., Schmidt, P., and Algesheimer, R. (2019). How to obtain comparable measures for cross-national comparisons. *KZfSS Köln. Z. Für Soziol. Sozialpsychol.* 71, 157–186. doi: 10.1007/s11577-019-00598-7

Cieciuch, J., Davidov, E., Schmidt, P., Algesheimer, R., and Schwartz, S. H. (2014). Comparing results of an exact vs. an approximate (Bayesian) measurement invariance test: a cross-country illustration with a scale to measure 19 human values. *Front. Psychol.* 5:982. doi: 10.3389/fpsyg.2014.00982

Coltman, T., Devinney, T. M., Midgley, D. F., and Venaik, S. (2008). Formative versus reflective measurement models: two applications of formative measurement. *J. Bus. Res.* 61, 1250–1262. doi: 10.1016/j.jbusres.2008.01.013

Davidov, E. (2009). Measurement equivalence of nationalism and constructive patriotism in the ISSP: 34 countries in a comparative perspective. *Polit. Anal.* 17, 64–82. doi: 10.1093/pan/mpn014

Davidov, E. (2010). Testing for comparability of human values across countries and time with the third round of the European Social Survey. *Int. J. Comp. Sociol.* 51, 171–191. doi: 10.1177/0020715210363534

Davidov, E., and Beuckelaer, A. D. (2010). How harmful are survey translations? A test with Schwartz's human values instrument. *Int. J. Public Opin. Res.* 22, 485–510. doi: 10.1093/ijpor/edq030

Davidov, E., Cieciuch, J., Meuleman, B., Schmidt, P., Algesheimer, R., and Hausherr, M. (2015). The comparability of measurements of attitudes toward immigration in the European Social Survey exact versus approximate measurement equivalence. *Public Opin. Q.* 79, 244–266. doi: 10.1093/poq/nfv008

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

Davidov, E., Dülmer, H., Cieciuch, J., Kuntz, A., Seddig, D., and Schmidt, P. (2016). Explaining measurement nonequivalence using multilevel structural equation modeling: the case of attitudes toward citizenship rights. *Sociol. Methods Res.* 47, 729–760. doi: 10.1177/0049124116672678

Davidov, E., Meuleman, B., Billiet, J., and Schmidt, P. (2008). Values and support for immigration: a cross-country comparison. *Eur. Sociol. Rev.* 24, 583–599. doi: 10.1093/esr/jcn020

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., and Billiet, J. (2014). Measurement equivalence in cross-national research. *Annu. Rev. Sociol.* 40, 50–75. doi: 10.1146/annurev-soc-071913-043137

Davidov, E., Muthén, B., and Schmidt, P. (2018b). Measurement invariance in cross-national studies: challenging traditional approaches and evaluating new ones. *Sociol. Methods Res.* 47, 631–636. doi: 10.1177/0049124118789708

de Beer, P., and Koster, F. (2009). *Sticking Together or Falling Apart?: Solidarity in an Era of Individualization and Globalization*. Amsterdam: Amsterdam University Press.

De Beuckelaer, A., and Lievens, F. (2009). Measurement equivalence of paper-and-pencil and internet organisational surveys: a large scale examination in 16 countries. *Appl. Psychol.* 58, 336–361. doi: 10.1111/j.1464-0597.2008.00350.x

Delanty, G. (2008). Fear of others: social exclusion and the european crisis of solidarity. *Soc. Policy Adm.* 42, 676–690. doi: 10.1111/j.1467-9515.2008.00631.x

Delhey, J., Deutschmann, E., Graf, T., and Richter, K. (2014). Measuring the europeanization of everyday life: three new indices and an empirical application. *Eur. Soc.* 16, 355–377. doi: 10.1080/14616696.2014.904916

Durkheim, E. (1893). *De la Division du Travail Social: étude sur l'organisation des Sociétés Supérieures*. Paris: Presses Universitaires de France.

Ekman, J., and Amnå, E. (2012). Political participation and civic engagement: towards a new typology. *Hum. Aff.* 22, 283–300. doi: 10.2478/s13374-012-0024-1

Ettore, R., and Adrian, F. (2019). *Everyday Europe: Social Transnationalism in an Unsettled Continent*. Bristol: Policy Press.

EVS (2017). *Methodological Guidelines of EVS 2017*. Available online at: http://www.europeanvaluesstudy.eu/

EVS (2017–2020). *European Values Study 2017: Integrated Dataset (EVS 2017)*. GESIS Data Archive, Cologne. ZA7500 Data file Version 4.0.0.

Fernández, G. G. E. (2019). Migration incorporation regimes and institutionalized forms of solidarity: between unconditional institutional solidarity and welfare chauvinism. *Am. Behav. Sci.* 63, 506–522. doi: 10.1177/0002764218823837

Ferrera, M. (2014). Solidarity in Europe after the Crisis. *Constellations* 21, 222–238. doi: 10.1111/1467-8675.12091

Fischer, R., and Karl, J. A. (2019). A primer to (cross-cultural) multi-group invariance testing possibilities in R. *Front. Psychol.* 10:1507. doi: 10.3389/fpsyg.2019.01507

Flake, J. K., and McCoach, D. B. (2018). An investigation of the alignment method with polytomous indicators under conditions of partial measurement invariance. *Struct. Equ. Model. Multidiscip. J.* 25, 56–70. doi: 10.1080/10705511.2017.1374187

Grimmel, A., and Giang, S. M. (Eds.). (2017). *Solidarity in the European Union: A Fundamental Value in Crisis*. Cham: Springer International Publishing.

Halman, L. (1995). *The Need for Theory in Comparative Research on Values*. Tilburg: Tilburg University, Work and Organization Research Centre.

Halman, L., and de Moor, R. (1993). "Comparative research on values," in *The Individualizing society. Value change in Europe and North America*, eds P. Ester, L. Halman, and R. de Moor (Tilburg: Tilburg University Press), 21–36.

Harris, A., Wyn, J., and Younes, S. (2010). Beyond apathetic or activist youth 'Ordinary' young people and contemporary forms of participation. *Young* 18, 9–32. doi: 10.1177/110330880901800103

Heath, A., Martin, J., and Spreckelsen, T. (2009). Cross-national comparability of survey attitude measures. *Int. J. Public Opin. Res.* 21, 293–315. doi: 10.1093/ijpor/edp034

Hechter, M., Nadel, L., and Michod, R. E. (1993). *The Origin of Values*. New York, NY: AldineTransaction.

Himmelfarb, S. (1993). "The measurement of attitudes," in *The Psychology of Attitudes*, eds A. H. Eagly and S. Chaiken (Belmont, TN: Thomson Learning), 23–87.

Hooghe, M., and Dejaeghere, Y. (2007). Does the 'Monitorial Citizen' exist? An empirical investigation into the occurrence of postmodern forms of citizenship in the nordic countries. *Scand. Polit. Stud.* 30, 249–271. doi: 10.1111/j.1467-9477.2007.00180.x

Hooghe, M., Hosch-Dayican, B., and Van Deth, J. W. (2014). Conceptualizing political participation. *Acta Polit.* 49, 337–348. doi: 10.1057/ap.2014.7

Horn, J. L., and Mcardle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Exp. Aging Res.* 18, 117–144. doi: 10.1080/03610739208253916

Hox, J. J., De Leeuw, E. D., and Zijlmans, E. A. O. (2015). Measurement equivalence in mixed mode surveys. *Front. Psychol.* 6:87. doi: 10.3389/fpsyg.2015.00087

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* 6, 1–55. doi: 10.1080/10705519909540118

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

Kalogeraki, S. (2018). "Volunteering for refugees and asylum seekers in Greece," in *Solidarity in Europe: Citizens' Responses in Times of Crisis Palgrave Studies in European Political Sociology*, eds C. Lahusen and M. T. Grasso (Cham: Springer International Publishing), 169–194.

Kankaraš, M., and Moors, G. (2009). Measurement equivalence in solidarity attitudes in Europe insights from a multiple-group latent-class factor approach. *Int. Sociol.* 24, 557–579. doi: 10.1177/0268580909334502

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling,* 4th Edn. New York, NY: Guilford Publications.

Koštál, J., and Klicperová-Baker, M. (2015). European sociopolitical mentalities: identifying pro- and antidemocratic tendencies. *Eur. Soc.* 17, 333–350. doi: 10.1080/14616696.2015.1018719

Lahusen, C., and Grasso, M. T. (Eds.). (2018). *Solidarity in Europe - Citizens' Responses in Times of Crisis*. London: Palgrave Macmillan. Available online at: https://www.palgrave.com/gp/book/9783319733340 (accessed June 24, 2020).

Lahusen, C., and Theiss, M. (2019). European transnational solidarity: citizenship in action? *Am. Behav. Sci.* 63, 444–458. doi: 10.1177/0002764218823836

Langsæther, P. E., and Stubager, R. (2019). Old wine in new bottles? Reassessing the effects of globalisation on political preferences in Western Europe. *Eur. J. Polit. Res.* 58, 1213–1233. doi: 10.1111/1475-6765.12332

Lomazzi, V. (2018). Using alignment optimization to test the measurement invariance of gender role attitudes in 59 countries. *Methods Data Anal. J. Quant. Methods Surv. Methodol.* 12, 77–103. doi: 10.12758/mda.2017.0x

Lomazzi, V., and Seddig, D. (2020). Gender role attitudes in the international social survey programme: cross-national comparability and relationships to cultural values. *Cross Cult. Res.* 54, 398–431. doi: 10.1177/1069397120915454

Maggini, N., and Fernández, E. (2019). Politicization of solidarity toward out-groups: the case of refugees. *Am. Behav. Sci.* 63, 475–491. doi: 10.1177/0002764218823840

Marsh, H. W., Guo, J., Parker, P. D., Nagengast, B., Asparouhov, T., Muthén, B., et al. (2018). What to do when scalar invariance fails: the extended alignment method for multi-group factor analysis comparison of latent means across many groups. *Psychol. Methods* 23, 524–545. doi: 10.1037/met0000113

Meitinger, K. (2017). Necessary but insufficientwhy measurement invariance tests need online probing as a complementary tool. *Public Opin. Q.* 81, 447–472. doi: 10.1093/poq/nfx009

Meuleman, B. (2019). *Measurement Equivalence: Between Statistical Dogmatism and Anything Goes*. Tbilisi, GA. Available online at: https://europeanvaluesstudy.eu/comparing-values-in-post-crisis-europe/

Millsap, R. E. (2010). Testing measurement invariance using item response theory in longitudinal data: an introduction. *Child Dev. Perspect.* 4, 5–9. doi: 10.1111/j.1750-8606.2009.00109.x

Mishra, C., and Rath, N. (2020). Social solidarity during a pandemic: through and beyond Durkheimian Lens. *Soc. Sci. Humanit. Open* 2:100079. doi: 10.1016/j.ssaho.2020.100079

Muthén, B., and Asparouhov, T. (2013). BSEM measurement invariance analysis. *Mplus Web Notes* 17, 1–48. Available online at: https://www.statmodel.com/examples/webnotes/webnote17.pdf

Muthén, B., and Asparouhov, T. (2014). IRT studies of many groups: the alignment method. *Front. Psychol.* 5:978. doi: 10.3389/fpsyg.2014.00978

Muthén, L. K., and Muthén, B. (1998). *Mplus User's Guide. 8th Edn.* Los Angeles, CA: Muthén & Muthén. Available online at: https://www.statmodel.com/ugexcerpts.shtml

Pokropek, A., Davidov, E., and Schmidt, P. (2019). A Monte Carlo simulation study to assess the appropriateness of traditional and newer approaches to test for measurement invariance. *Struct. Equ. Model. Multidiscip. J.* 26, 724–744. doi: 10.1080/10705511.2018.1561293

Pokropek, A., Lüdtke, O., and Robitzsch, A. (2020). An extension of the invariance alignment method for scale linking. *Psychol. Test Assess. Model.* 62, 305–334. Available online at: https://www.psychologie-aktuell.com/journale/psychological-test-and-assessment-modeling/currently-available/inhaltlesen/psychological-test-and-assessment-modeling-2020-2.html

Raju, N. S., Laffitte, L. J., and Byrne, B. M. (2002). Measurement equivalence: a comparison of methods based on confirmatory factor analysis and item response theory. *J. Appl. Psychol.* 87:517. doi: 10.1037/0021-9010.87.3.517

Reinl, A.-K. (2020). "Euroscepticism in times of European crises: the role of solidarity," in *Anti-Europeanism. Critical Perspectives Towards the European Union*, eds M. Baldassari, E. Castelli, M. Truffelli, and G. Vezzani (Cham: Springer), 95–113.

Robitzsch, A. (2019). *sirt: Supplementary Item Response Theory Models*. Available online at: https://CRAN.R-project.org/package=sirt

Robitzsch, A. (2020). Lp loss functions in invariance alignment and haberman linking with few or many groups. *Stats* 3, 246–283. doi: 10.3390/stats3030019

Rose, N., Wagner, W., Mayer, A., and Nagengast, B. (2019). Model-based manifest and latent composite scores in structural equation models. *Collabra Psychol.* 5:9. doi: 10.1525/collabra.143

Rudnev, M. (2018). "Testing for invariance of latent classes: group-as-covariate approach," in *Cross-Cultural Analysis: Methods and Applications*, eds E. Davidov, P. Schmidt, J. Billiet, and B. Meuleman (London: Routledge), 449–480.

Rudnev, M., Lytkina, E., Davidov, E., Schmidt, P., and Zick, A. (2018). Testing measurement invariance for a second-order factor. a cross-national test of the alienation scale. *Methods Data* 12:30. doi: 10.12758/mda.2017.11

Rusu, H. (2012). Measuring social solidarity. Some research notes. *Soc. Change Rev.* 10, 71–90. doi: 10.2478/scr-2013-0012

Seddig, D., and Leitgöb, H. (2018). Approximate measurement invariance and longitudinal confirmatory factor analysis: concept and application with panel data. *Surv. Res. Methods* 12, 29–41. doi: 10.4324/9781315537078-20

Seddig, D., and Lomazzi, V. (2019). Using cultural and structural indicators to explain measurement noninvariance in gender role attitudes with multilevel structural equation modeling. *Soc. Sci. Res.* 84:102328. doi: 10.1016/j.ssresearch.2019.102328

Sokolov, B. (2018). The index of emancipative values: measurement model misspecifications. *Am. Polit. Sci. Rev.* 112, 395–408. doi: 10.1017/S0003055417000624

Steenkamp, J. E. M., and Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *J. Consum. Res.* 25, 78–107. doi: 10.1086/209528

Stjernø, S. (2011). *The Idea of Solidarity in Europe.* Available online at: https://oda.oslomet.no/handle/10642/1125 (accessed December 13, 2020).

UNPD (2013). *Human Development Report 2013 - Technical Notes.* New York, NY: United Nations Development Programme.

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

van de Vijver, F., and Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: an overview. *Rev. Eur. Psychol.* 54, 119–135. doi: 10.1016/j.erap.2003.12.004

van Vlimmeren, E., Moors, G. B. D., and Gelissen, J. P. T. M. (2016). Clusters of cultures: diversity in meaning of family value and gender role items across Europe. *Qual. Quant.* 51, 2737–2760. doi: 10.1007/s11135-016-0422-2

Voicu, B., Peral, E. B., Rusu, H., Rosta, G., Comşa, M., Vasile, O.-M., et al. (2021). COVID-19 and orientations towards solidarity: the cases of Spain, Hungary, and Romania. *Eur. Soc.* 23, 887–904. doi: 10.1080/14616696.2020.1852439

Wallaschek, S. (2019). *Without Solidarity, Democracy Is Lost.* Social Europe. Available online at: https://www.socialeurope.eu/without-solidarity-democracy-lost (accessed September 7, 2020).

Welzel, C., Brunkert, L., Inglehart, R., and Kruse, S. (2019). Measurement equivalence? A tale of false obsessions and a cure. *World Values Res.* 11, 54–84. Available online at: http://www.worldvaluessurvey.org/WVSPublicationsBooks.jsp?PUB=178

Welzel, C., and Inglehart, R. F. (2016). Misconceptions of measurement equivalence: time for a paradigm shift. *Comp. Polit. Stud.* 49, 1068–1094. doi: 10.1177/0010414016628275

West, S. G., Taylor, A. B., and Wu, W. (2012). "Model fit and model selection in structural equation modeling," in *Handbook of Structural Equation Modeling*, ed R. H. Hoyle (New York, NY: Guilford Publications), 209–231.

Zercher, F., Schmidt, P., Cieciuch, J., and Davidov, E. (2015). The comparability of the universalism value over time and across countries in the European Social Survey: exact versus approximate measurement invariance. *Front. Psychol.* 6:733. doi: 10.3389/fpsyg.2015.00733