

Gender Role Attitudes in the International Social Survey Programme: Cross-National Comparability and Relationships to Cultural Values

Lomazzi, Vera; Seddig, Daniel

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Lomazzi, V., & Seddig, D. (2020). Gender Role Attitudes in the International Social Survey Programme: Cross-National Comparability and Relationships to Cultural Values. *Cross-Cultural Research*, 54(4), 398-431. <https://doi.org/10.1177/1069397120915454>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

gesis
Leibniz-Institut
für Sozialwissenschaften

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Mitglied der

Leibniz-Gemeinschaft

Diese Version ist zitierbar unter / This version is citable under:

<https://nbn-resolving.org/urn:nbn:de:0168-ssoar-85010-3>

Gender Role Attitudes in the International Social Survey Programme: Cross-National Comparability and Relationships to Cultural Values

Cross-Cultural Research
2020, Vol. 54(4) 398–431
© 2020 SAGE Publications
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1069397120915454
journals.sagepub.com/home/ccr



Vera Lomazzi¹ and Daniel Seddig²

Abstract

Differences in societal views on the roles of men and women have been addressed in many large-scale comparative studies by employing indicators of gender roles attitudes from cross-sectional surveys. Assuming that cross-country differences in gender role attitudes are linked to the prevailing cultural value orientations in each society, this study aims at investigating the association between societal views on gender roles, as measured by the International Social Survey Programme (ISSP), and the prevailing cultural values, as defined by Schwartz's theory. However, to carry out meaningful comparisons, we first assessed the prerequisite of measurement equivalence between countries. The comparability of gender role attitudes is limited when using traditional methods based on the concept of exact equivalence (multiple-group confirmatory factor analysis). However, the recently established alignment optimization procedure reveals approximate measurement equivalence and suggests that the mean comparison is trustworthy. Based on these results, we correlate the national mean levels of gender role attitudes with the cultural values of embeddedness, hierarchy

¹GESIS—Leibniz Institute for Social Sciences, Cologne, Germany

²University of Cologne, Germany

Corresponding Author:

Vera Lomazzi, Data Archive for the Social Sciences, GESIS—Leibniz Institute for the Social Sciences, Unter Sachsenhausen 6-8, 50667 Cologne, Germany.

Email: vera.lomazzi@gesis.org

and egalitarianism, showing that traditional gender roles are displayed in societies emphasizing hierarchy and embeddedness while progressive views are more expressed in egalitarian societies.

Keywords

measurement equivalence, comparability, ISSP, gender role attitudes, cultural values

Introduction

Gender equality is a highly relevant topic for academic research, political decision-making, and society at large. Gender statistics provided by national and transnational agencies offer the possibility to study how gender equality differs across societies taking into account objective information, such as the gender gap in access to education, to the labor market, economic, and political power. Occasionally, such sources can also offer information on individual behaviors and household arrangements: Many national statistical offices, for example, provide data on gender equality in the division of work between partners, based on the use of time surveys. However, exploring what gender equality means for people at the societal level and how this value is endorsed individually is subjective information that allows completing the picture of the different aspects of gender equality. However, at the same time, it represents a methodological and substantive challenge. Observing and measuring human values directly is impossible and leads to a speculative theoretical approach (Halman, 1995, p. 3). Because of this difficulty, rooted in the very nature of values, values in social research can only be inferred or postulated. From an empirical point of view, as values cannot be measured directly, one can resort to related concepts such as beliefs, attitudes, and opinions which also have an abstract nature. According to Rokeach (1968, p. 124), a value is understood as a disposition of a person, just like an attitude, but unlike an attitude, the value is fundamental and more essential. Nevertheless, measuring attitudes allows getting the closest possible to aspects of the related values.

During the last decades, observing and measuring gender role attitudes became, therefore, a popular strategy to tackle the measurement of gender equality value (Bergh, 2006). Instrument aimed at grasping gender role attitudes are available in many cross-national surveys, often repeated over time.

However, it is important to note that longitudinal and cross-sectional comparisons raise special methodological issues. Comparative studies rely on the assumption of comparability, which means that the measurement of the variables employed in the comparison supports the equivalence of their characteristics (Billiet, 2003). However, this basic assumption of comparability

cannot be taken for granted. For example, people living in different cultural contexts may interpret survey questions differently and methodological bias can also occur (Davidov et al., 2014). The presence of cultural or methodological biases affects the comparability across societies: Measurement equivalence is a precondition of running comparative studies in both cross-sectional and longitudinal research designs (Davidov et al., 2014) that need to be assessed to avoid the risk of achieving misleading results. This is even more important when dealing with constructs that previous research already identified as particularly sensitive to cultural biases as it is in the case of gender role attitudes (Braun, 2009; Constantin & Voicu, 2015; Lomazzi, 2018).

This study has two aims. The first is to provide methodological insights concerning the measurement equivalence of gender role attitudes by proposing an applicative study of different techniques. In addition to the traditional assessment through multigroup confirmatory factor analysis (MGCFA), based on the concept of “exact equivalence,” the study applies the novel alignment optimization procedure, which was recently proposed by Asparouhov and Muthén (2014), and assumes the perspective of “approximate equivalence,” meaning that a certain amount of invariance can be kept at minimum without affecting the factor means comparability (Van de Schoot et al., 2013). To date, only a few studies report the applicative use of the alignment method (Flake & McCoach, 2018; Lomazzi, 2018; Marsh et al., 2018; Munck et al., 2017) and none have used the alignment model with Bayesian estimation and real data. Therefore, we use alignment with maximum likelihood and Bayesian estimation to contribute to covering this gap in the empirical literature. The second aim of this study is substantive and concerned with linking cross-country differences in gender role attitudes to prevailing cultural values on the societal level.

Individual attitudes toward gender roles arise from an inner framework of values and beliefs concerning, for example, egalitarianism, autonomy, self-determination, and so on (Kalmijn, 2003). Values transmission during the primary socialization and experiences during the course of life, including secondary socialization processes and daily negotiations between partners and primary groups, contribute to the development of individual value systems. Gender equality value and gender role attitudes are also part of this process (Moen et al., 1997): The connection between societal cultural orientation and gender role attitudes can be explained by the Exposure Theory (Bolzendahl & Myers, 2004), which argues that socialization and education expose individuals to the prevailing gender norms in a society. Being exposed to cultural contexts endorsing egalitarian ideals reflects into the development of more egalitarian gender beliefs.

Taking into account that individual values are related to country-level value structures (Hofstede, 1980, 2001; Schwartz, 2006), we, therefore, assume that specific cultural values, which are the shared ideas about what societies deem important, are explicitly linked to differences in the predominant views on gender roles across countries. As anticipated previously, to make appropriate comparisons between the prevailing views on gender roles in different societies, their measurement needs to be tested for equivalence across countries. Therefore, the second aim of the study (i.e., a cross-country analysis of the relation of gender role attitudes and culture) is conditional on the first goal of the study (i.e., a test for measurement equivalence across countries).

In the present investigation, we use data from the International Social Survey Programme (ISSP). The ISSP collects data worldwide through thematic modules repeated over time. In addition to the other topics investigated, the module on Family and Changing Gender Roles comprises measurements of gender role attitudes as well. In this study, we use data from the most recent edition of this module, carried out in 2012.

We will use the data to obtain country-specific scores of gender role attitudes. Only after verifying the comparability of the measurement across countries, it is appropriate to explore the connection between gender role attitudes and cultural values at the societal level. The macro data on cultural values are aggregated scores of data on individual values provided by Schwartz (2008).

This article is structured into five parts. The first part provides a literature review on attitudes toward gender roles and the link with cultural value dimensions. In the second part, we articulate some relevant issues concerning the measurement of gender role attitudes. The third section presents the measurements used in the study, the sources for the individual and societal level data, and a description of the employed methods of analysis. The fourth part provides the results of measurement equivalence tests assessing the comparability of the gender role attitudes measurements across countries. Based on the results obtained from the alignment method, we consider the measurements of gender role attitudes to be comparable and continue to demonstrate the substantive connection of gender role attitudes with cultural values. Therefore, we use the country-specific latent factor means and correlate them with scores implicating country-specific value orientations. We conclude with a summary of the study and some insights for further research.

Gender Role Attitudes

Theory

Gender role attitudes are often defined as the cognitive representation of what is believed appropriate for male and female roles (Alwin, 2005; Lee et al.,

2010). Other authors refer to gender ideology, a more complex construct based on gender role attitudes, as the “the underlying concept of an individual’s level of support for a division of paid work and family responsibilities that is based on the notion of separated spheres” (Davis & Greenstein, 2009, p. 89), which implies a gendered separation of roles in the public and private sphere.

Attitudes toward gender roles can refer to the appropriate roles of women and men in the private area, being strongly connected with the preference for a certain family model (Cunningham et al., 2005; Kroska & Elman, 2009), or to the gender roles in the public area such as education, labor market, or politics (Albrecht et al., 2000; André et al., 2013; Baxter & Kane, 1995). People expressing traditional gender role attitudes manifest their support for the specialization of tasks and roles by gender, with women devoted to the family chores and men to tasks in the public realm. In contrast, those expressing egalitarian attitudes tend to be against the segregation of social roles, supporting the role of women in the public sphere as well as the role of men in the private one.

Prevailing gender norms in society and family models contribute to shaping gender role expectations. Family models tend to have a direct impact on gender roles because they transmit gender role models concerning, for example, the division of tasks and responsibilities between parents and siblings. The male breadwinner–female homemaker is a model based on the gendered specialization of tasks: Family members are socialized to the gendered division of paid/unpaid work and this contributes internalizing traditional gender role expectations (Bolzendahl & Myers, 2004). This is because, according to the Social Role Theory (Eagly, 1987), people tend to form their attitudes toward gender roles in a consistent way with what they observe in the daily behavior of men and women. Although this model was quite spread in the past, the male-breadwinner model became in the last decades always less frequent (Lewis, 2001) and several family models are nowadays coexisting in many societies, even if with a variety of distribution by country. The prevalence of a specific family model is ingrained in local history and in family policies promoted by the state (Duncan, 1995; Pfau-Effinger, 2004). Thus, gender roles and the expectations concerning the appropriate role for men and women may significantly differ not only over time in the same society but also across countries, which display a variety of societal cultures. Regional differences in gender role attitudes also exist, especially where historical pathways differ by region. The literature reports, for example, the case of East/West Germany (Boehnke, 2011; Lee et al., 2007), United States (Carter & Borch, 2005; Powers et al., 2003), United Kingdom (Bohenke, 2011; Uunk & Lersch, 2019), and Italy (Lomazzi, 2017a). In this study,

however, we focus on cross-national comparison and refer to prevailing cultural orientations in national societies.

Schwartz (2006) defines culture as the “rich complex of meanings, beliefs, practices, symbols, norms, and values prevalent among people in a society” (p. 138). The emphasis of particular values in a society is one of the most important characteristics of culture (Halman, 2010; Hofstede, 2001; Inglehart, 1997; Schwartz, 1999). Cultural values are the rationale for social institutions as well as economic and political systems, and they express the shared ideas of what is desirable in a specific culture. The prevailing cultural values can shape and address individual beliefs, attitudes, goals, and behaviors. The relation between attitudes and values is strict, but they refer to different concepts: Whereas attitudes refer to a specific object and generally concern the fact of agreeing or disagreeing with it (Allport, 1961), values have a “transcendental quality” (Rokeach, 1973, p. 18) to them. Moreover, value systems tend to be quite stable over time and have a more general and broader dimension than attitudes (Bergh, 2006; Halman, 2010; Schwartz, 2006). The predominant general value orientations in a society can, therefore, support the formation and maintenance of specific gender role attitudes.

The nexus between gender role attitudes and general value orientations has been investigated through the lens of Inglehart’s post-materialism theory (Bergh, 2006; Inglehart, 1977, 1997; Inglehart & Norris, 2003; Kalmijn, 2003) and of the emancipation theory, proposed by Welzel (2013) as a further development of Inglehart’s approach. According to the post-materialism perspective, modernization and secularization drove the value and attitude changes from a traditional perspective to a more liberal view on gender roles because of the increased emphasis on self-realization and issues related to the quality of life. Studies following this perspective tend to consider structural aspects of the societal context as the economic resources and developmental indicators to explain value change over time and value differences between societies. Cultural values, according to this framework, refer to the quality of life issues, self-expression, individualism, and postmaterialism (Inglehart & Norris, 2003). In the post-materialism theory, the economic change is assumed as the driving force of social and cultural change but other aspects such as the pursuit of freedom and democracy are disregarded. Welzel’s (2013) emancipative theory addresses these elements by proposing the “Emancipative Value Index,” which allows investigating values in the dimensions of equity, liberty, autonomy, and expression.

Alongside the potentialities of these approaches to explain social change, some limits need to be taken into account. The assumption of these theories is that people interpret values in the same way, regardless of their cultural belonging and there is no need to assess this assumption. This idea, which

is in contrast with the general framework of this article, is also challenged by results from recent empirical studies that show that these measurements tend to be equivalent only for subsamples of Western industrialized countries (Alemán & Woods, 2016; Sokolov, 2018). In this article, we adopt the theory of cultural value orientations proposed by Schwartz (2006), another relevant and well-known theory of values that, differently from the previous two, does not assume an evolutionary and economic framework and provides equivalent measurements of values (Schwartz, 1999, 2006). Moreover, to our knowledge, the concept of gender role attitudes has not been linked to Schwartz' human values so far and this approach can provide further insights for the study of gender role attitudes and their association with predominant cultural orientations. Schwartz (2006) provides a theory of cultural value orientations that is based on his theory of individual differences in value priorities (Schwartz, 1994) and describes seven cultural value orientations.¹ Among those, the cultural value dimensions of embeddedness, hierarchy, and egalitarianism may in particular enhance our understanding of differences in the predominant view of gender roles at the country level. Embeddedness refers to societies in which people are seen as entities, who are deeply embedded in the collectivity (Schwartz, 2006, p. 140). Such societies stress the importance of maintaining the status quo and the traditional order, and they restrain actions that negatively affect in-group solidarity. We expect that embedded societies hold more traditional views on gender roles because they are part of the traditional order and help to preserve the status-quo. Hierarchy describes societies with an emphasis on a system of predefined roles, which are deemed important to "insure responsible, productive behavior" (Schwartz, 2006, p. 141). The arrangement of roles is hierarchical and implies that different rules and obligations are attached to each role. Furthermore, the hierarchy of roles implies a differential distribution of power and resources. We expect that societies with an emphasis on hierarchical roles share a more traditional view on gender roles because compliance with the traditional divide of spheres is assumed to serve the maintenance of societal functioning. Egalitarianism is connected with the goal to recognize and mind other individuals' interests. Individuals in such societies interact "as moral equals who share basic interests as human beings" (Schwartz, 2006, p. 140). The benefit and welfare of others is at the core of egalitarianism. Therefore, we expect to find more progressive gender role attitudes in egalitarian societies, because in these societies, both women and men are supposed to be given the same opportunities and choices regarding their life in the domestic and public spheres.

Measurement of Gender Role Attitudes

Considered as a good proxy for measuring individual's support toward gender equality (Bergh, 2006), several scholars employed the measurements of gender role attitudes made available by various repeated large-scale surveys² to compare the individual support for egalitarian gender roles across countries (André et al., 2013; Kunovich & Kunovich, 2008; Sjöberg, 2004) or to monitor the change over time of these attitudes (Cotter et al., 2011; Kraaykamp, 2012; Lomazzi, 2017a; Valentova, 2013).

The measurement of gender role attitudes is often problematic, because of the lack of conceptual coverage and the outdated wording of the questions. Furthermore, the sensitivity to cultural bias and the potential lack of cross-cultural comparison increases the risk of obtaining misleading results.

As Walter (2018) pointed out, the instruments currently used to investigate gender role attitudes are not adequate enough to measure such a complex and multidimensional concept. Most of the measurements available tend to be focused on female roles. In addition, the coverage of the concept is often limited to the private sphere, whereas measurements concerning attitudes toward roles in the public realm are still lacking (Constantin & Voicu, 2015; Walter, 2018). Moreover, most of these instruments were developed in the late 1970s and reflect the social roles that were predominant at that time, linked to the male breadwinner model. To maintain continuity over time, many cross-sectional surveys only modified the scales slightly (if at all). Although this allows for comparison over time, through their item wording, some scales maintain the imprinting of a traditional view of gender roles, and today this wording can be differently perceived across countries (Braun, 2008, 2009).

The normative beliefs concerning gender roles in society appear particularly sensitive to cultural differences. The cultural context contributes to shaping gender role attitudes (André et al., 2013; Banaszak & Plutzer, 1993; Cunningham et al., 2005; Sjöberg, 2004) but it also affects the way people interpret the items used to investigate this concept (Braun, 1998, 2009). Therefore, this raises issues concerning the suitability of these measurements for cross-cultural comparisons. Despite the wide use of these measurements in comparative research, only a few studies have yet addressed questions concerning measurement equivalence in this field. Employing MGCFA, Constantin and Voicu's (2015) assessment revealed that both the gender role attitudes scale used by the WVS 2005 and the ISSP 2002 are not equivalent, and their means cannot be compared across countries. Lomazzi (2018) evaluated the measurement invariance of the scale included in the WVS 2010 employing two techniques. In addition to the traditional assessment via

MGCFA, which proved measurement equivalence only for a limited subgroup of countries (27), the author proposed the recently developed alignment optimization (Asparouhov & Muthén, 2014) as a potential alternative to MGCFA. Using the alignment optimization technique, trustworthy factor means were found for 35 countries. This promising technique requires more empirical validation assessing whether it can be considered as a viable alternative to MGCFA, especially when comparisons involve a large number of groups (Asparouhov & Muthén, 2014; Muthén & Asparouhov, 2014). To contribute to both the research in the field of gender role attitudes and to the field of measurement equivalence, in this article we focus our attention on the most recent gender role attitudes scale surveyed by the ISSP.

Data, Measures, and Method of Analysis

Data and Measures

The ISSP is a continuing programme of cross-national collaborative research. Since 1985, it gathers information yearly on individual behaviors, preferences, opinions, and attitudes among population samples across the world. Through the implementation of thematic modules, which are replicated every 8 to 10 years with minor revisions, the ISSP allows for cross-time and cross-national analyses. The module “Family and Changing Gender Roles” was first implemented in 1988, and it has been surveyed four times (1988, 1994, 2002, and 2012). The module collects information concerning several topics: attitudes toward family and gender roles, attitudes toward marriage, alternative family forms, attitudes toward children (gender, care, and social policy), family models in the division of paid/unpaid work, income in partnership, gendered division of household work, power and decision-making in couples, work–family conflict, and happiness and satisfaction (Scholz et al., 2014). The ISSP gender role attitudes scale is quite popular and several studies have included this measurement (Braun, 2009; Motiejunaite & Kravchenko, 2008; Scott et al., 1996; Sjöberg, 2004; Stickney & Konrad, 2007), but only Constantin and Voicu (2015) investigated the measurement equivalence of the scale from ISSP 2002.

In this study, we consider the most recent edition carried out in 2012 (ISSP Research Group, 2016) that assesses gender role attitudes by asking for the respondents’ agreement to seven statements.³ In the following, we use the original item names from the ISSP 2012 (v5–v11). Respondents could rate their agreement from 1 (*strongly agree*) to 5 (*strongly disagree*). Table 1 shows the item wording as presented to the respondents according to the English source questionnaire, and the descriptive statistics for each item.

Table 1. Items in the ISSP Measuring Gender Role Attitudes (N = 61,754) and Their Factor Loadings (Extraction Method: Principal Axis Factoring; Rotation Method: Varimax With Kaiser Normalization).

Item	Wording	M	SD	% Missing	Factor loadings	
					Factor 1	Factor 2
v5	A working mother can establish just as warm and secure a relationship with her children as a mother who does not work.	2.26	1.14	6.31	-.483	-.073
v6	A preschool child is likely to suffer if his or her mother works.	3.03	1.22	6.81	.693	.288
v7	All in all, family life suffers when the woman has full-time job.	2.94	1.24	6.68	.716	.284
v8	A job is all right, but what most women really want is a home and kids.	3.12	1.21	8.70	.269	.666
v9	Being a housewife is just as fulfilling as working for pay.	3.19	1.17	9.95	.084	.388
v10	Both the man and woman should contribute to the household income.	2.00	0.93	5.79	-.18	-.097
v11	A men's job is to earn money; a women's job is to look after the home and family.	2.73	1.29	5.95	.345	.578

Note. ISSP = International Social Survey Programme.

Compared with the scale surveyed in ISSP 2002 and evaluated by Constantin and Voicu (2015), the 2012 version presents a reduced set of items. In addition to those listed in Table 1, the 2002 version included three further items (v12: "Having a job is the best way for a woman to be an independent person"; v13: "Men ought to do a larger share of household work than they do now"; and v14: "Men ought to do a larger share of child care than they do now"). However, built on results from their exploratory factor analysis, the measurement model assessed by Constantin and Voicu (2015, pp. 745–747) included only six items⁴ of the 10 available.

The most recent version of the scale has not been assessed yet and, considering the popularity of this measurement, its evaluation can be a valuable contribution for scholars in the field of gender role attitudes.

Reviewing items also belonging to previous editions of ISSP, Braun (1998) argued that some of them present conceptual problems. Items v5 and v6, for example, introduce two aspects in the same statement: The respondents could focus either on the child's need or on the mother's capabilities. In addition, the individual or societal resources for child care may affect Item v6: According to the quantity and quality of child care provision, the respondent may consider the importance of the caregiver roles differently. Item v9 has been criticized for not actually measuring attitudes toward gender roles, but it concerns fulfillment. Items concerning the desirability of women's participation in the labor force, such as Item v10, can be controversial because the contribution to the household income by both partners can also reflect an economic necessity rather than an egalitarian belief (Braun et al., 1994). Considering that the ISSP 2012 has been surveyed in the context of the recession following the global economic crisis, also this item could be problematic.

An exploratory factor analysis was conducted for the entire sample to assess the overall pattern for factors and relevant items before we engage in testing the comparability of the pattern across countries. The results (Table 1) raised concerns about the use of Items v9 and v10, which have rather low factor loadings (below .40). Although the loading of Item v5 is .43, a deeper country-by-country investigation reveals that this item shows very poor factor loadings (even below .30) in many countries.⁵

These results confirmed the substantive problems argued by Braun (1998), and we, therefore, excluded these items from further analyses, retaining only the four with the highest loadings (Items v6, v7, v8, and v11).

Schwartz (2008) provides the scores for the cultural values embeddedness, hierarchy, and egalitarianism as aggregated country-level data. The data are based on individual-level ratings from the 56- to 57-item Schwartz Value Survey (for details, see Schwartz, 2006, p. 145). Value items were administered to schoolteachers and college students from 58 and 64 national groups,

respectively, between 1988 and 2007. The adequacy of using the combined teacher- and student-ratings as representations of the larger population in each country has been confirmed (Schwartz, 2006). Furthermore, the underlying individual-level measurements were found to be equivalent in meaning across countries (Schwartz, 1999, 2006), enabling to use the country-level scores for correlational analyses. Although the period of data collection is very long, the risk of ignoring substantial changes in the meaning and priority of values across time seems small, because at the national level, values are deemed stable and changes appear slow (Schwartz et al., 2000). Country-level data from Canada, Germany, Israel, and Switzerland were available for different subsamples.⁶ We, therefore, calculated a mean score for each. Data were not available for Iceland and Lithuania.

Methods of Analysis

In order to compare constructs across different groups or time, it is indispensable to test for the equivalence of the underlying measurements. When measurement equivalence or measurement invariance is not given, cross-group differences in regression coefficients or factor means may only arise because of differences in the measurement characteristics, but not because of true differences in the latent concept. Furthermore, finding no differences in regression coefficients or factor means does not imply that true differences are absent (Davidov et al., 2014, 2015, 2016; Horn & McArdle, 1992; Steenkamp & Baumgartner, 1998; Vandenberg & Lance, 2000). Such biases may occur because people in different groups may differently understand the questions in a measurement instrument or they may differ in the way they respond to questions although their underlying score on the latent dimension is the same. Thus, it is essential to test whether the measurement characteristics of a measurement instrument are invariant across groups. The goal is to assess “whether or not, under different conditions of observing and studying phenomena, measurement operations yield measures of the same attribute” (Horn & McArdle, 1992, p. 117).

A widely applied tool to test measurement invariance is MGCFA (Brown, 2015; Jöreskog, 1971; Reise et al., 1993). This approach builds on the concept of exact equivalence that requires the “exact equivalence” between parameters and several hierarchical levels of measurement invariance are discussed in the literature (Horn & McArdle, 1992; Meredith, 1993; Vandenberg & Lance, 2000). However, we refer only to the most common levels tested in cross-cultural sociological and psychological research. Before testing for measurement invariance, a measurement model should be established that fits the data well in each group separately. Furthermore, factor loadings must

be substantial (e.g., $>.30$), and correlations among factors should be smaller than one. The first and least restrictive level of measurement invariance is *configural* or structural invariance, which requires the same latent variables and observed indicators in each group. In addition, *metric* or loading invariance requires that the factor loadings are equal across groups. With equal factor loadings, it is possible to draw valid comparisons of factor variances, covariances, and unstandardized regression coefficients across groups. Furthermore, *scalar* or intercept measurement invariance adds an equality constraint on the indicator intercepts. When scalar measurement invariance holds in the data, it is possible to compare factor means across groups.

Choosing a “bottom-up” strategy, one can begin with the least restrictive model (configural) and then gradually impose equality constraints on the factor loadings (metric) and intercept (scalar) until the model is rejected by the data. For the assessments of model fit and differences between the levels of measurement invariance, researchers can rely on several fit statistics that are provided in structural equation modeling (SEM). Chi-square tests are sensitive to sample size (Saris et al., 1987) and known to reject models because of minor misspecifications. Therefore, chi-square based goodness-of-fit measures such as the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardized root mean residual (SRMR) are preferred (Hu & Bentler, 1999; Marsh et al., 2004; West et al., 2012). An acceptable model is given when the CFI value is higher than .90, and the RMSEA and SRMR values are lower than .08. Furthermore, Chen (2007) suggested criteria to assess whether the differences between the levels of measurement invariance are relevant. When the sample size is $n > 300$, differences between configural and metric models are relevant when the change in CFI is larger than .010, complemented by a change in RMSEA larger than .015, or a change in SRMR larger than .030. Differences between metric and scalar models are deemed relevant when the change in CFI is larger than .01, complemented by a change in RMSEA larger than .015 or a change in SRMR larger than .01.

In situations where the model fit significantly deteriorates moving from one level of measurement invariance to another, it may be reasonable to test for partial measurement invariance (Byrne et al., 1989). Based on residual information or modification indexes provided in SEM, the factor loadings or intercepts that differ across group can be freely estimated. A minimum of two factor loadings and intercepts must be equal across groups to draw valid comparisons of factor means.⁷

In many cases, however, the classic assessment of measurement invariance with MGCFA may be too strict and preclude meaningful comparisons of factor means across groups even though the degree of noninvariance may be

small. Recently, alternative methods have been proposed that do not require exact equality of measurement parameters across groups. These techniques refer to the concept of “approximate equivalence” that basically aims at taking into account the cultural variability and uncertainty in the assessment (Muthén & Asparouhov, 2013; Van de Schoot et al., 2013). As one of the more liberal approaches, the alignment optimization procedure allows for some flexibility of measurement parameters across groups while still maintaining the highest possible degree of equivalence. In what follows, we will only explain the conceptual idea of the alignment procedure. The mathematical aspects have been described elsewhere (e.g., Asparouhov & Muthén, 2014; Marsh et al., 2018; Muthén & Asparouhov, 2014, 2017). The alignment approach begins with a base model, which is the unrestricted (configural) measurement model. This model does not contain any equality constraints on the factor loadings, intercepts, or factor means across groups. In the course of the optimization procedure, the measurement parameters are chosen in a way that the degree of noninvariance will be as small as possible, but without the requirement of any equality constraints. The alignment idea is similar to rotation in exploratory factor analysis (not only conceptually but also mathematically, see Asparouhov & Muthén, 2014): We want to obtain a final model with as many approximately invariant parameters as possible and only a few large noninvariant parameters (consider that in exploratory factor analysis we want to obtain as many as possible small factor loadings and only a few large factor loadings). Thus, the final aligned measurement model contains the most trustworthy estimates of the factor loadings, intercepts, and factor means under the condition of approximate measurement equivalence. Furthermore, the final model has the same fit as the base model. Two alignment procedures are available. The FREE alignment procedure uses a reference group and fixes the group’s factor variance and factor mean to 1 and 0, respectively. The reference group is the first group. FREE alignment may not be applicable with only two groups and/or with a high degree of invariance (Asparouhov & Muthén, 2014). An alternative is the FIXED alignment procedure, where the factor mean of the reference group is fixed to zero. The reference group is usually the group with the factor mean closest to zero in the FREE alignment (Asparouhov & Muthén, 2014).

However, using the alignment procedure does not automatically guarantee that measurements are actually (approximately) comparable across groups. The degree of acceptable noninvariance is assessed by examining the proportion of noninvariant parameters in the final alignment model. Two different cut-off values have been proposed on the basis of Monte Carlo simulation studies. According to Muthén and Asparouhov (2014), a tolerable degree of noninvariance is given when the proportion of noninvariant parameters in the final model

does not exceed 25%. Flake and McCoach (2018) investigated the performance of alignment with polytomous items. They suggested inspecting the degree of noninvariance of factor loadings and thresholds separately when researchers are particularly interested in factor mean comparisons. Their recommendation is that no more than 29% of the thresholds should be noninvariant. Furthermore, Muthén and Asparouhov (2014) suggested that higher degrees of noninvariance should be supplemented by a Monte Carlo simulation study to assess whether the arrangement of estimated factor means is trustworthy. Factor means are considered trustworthy when the correlation between the generated and estimated factor means is very high ($r \geq .98$). This indicates that factor means may be comparable, even if the degree of noninvariance in the alignment model exceeds the recommended (25% or 29%) cut-off values. Thus, even when measurement invariance is not given according to the classic MGCFA approach, comparisons may still be admissible using the alignment approach.

The invariance pattern in the alignment model rests on the assumption that some items display large variations across countries, whereas the differences in other items are rather small or moderate. Thus, alignment is particularly useful in such situations or as a starting point for investigating measurement invariance in more detail, for example, with approximate measurement invariance using Bayesian SEM (Cieciuch et al., 2014, 2017; Lek et al., 2019; Muthén & Asparouhov, 2013; Seddig and Leitgöb 2018a, 2018b; Van de Schoot et al., 2013).

We used the software package Mplus Version 8 (Muthén & Muthén, 1998–2017) for all calculations reported in this article.⁸ None of the items had a piling of responses in the extreme categories. Thus, a treatment of the data as continuous seems reasonable (Rhemtulla et al., 2012). We used maximum-likelihood (ML) estimation. The alignment models were additionally estimated using ML with robust adjustment for standard errors (MLR [robust maximum likelihood]) and Bayesian estimation. ML, MLR, and the Bayesian estimator use the full information available from the data and assume missing at random (Schafer & Graham, 2002). Our aim here was to test whether the estimators yield similar estimates, which increases the trustworthiness of the results.

The relationship of gender role attitudes and cultural values is conducted on the country-level. Therefore, we produced Pearson correlations between the gender role attitude factor means and the cultural values scores.

Results

Confirmatory Factor Analysis

First, we tested the measurement model for the latent factor “gender role attitudes” with CFA across all countries. Modification indexes suggested

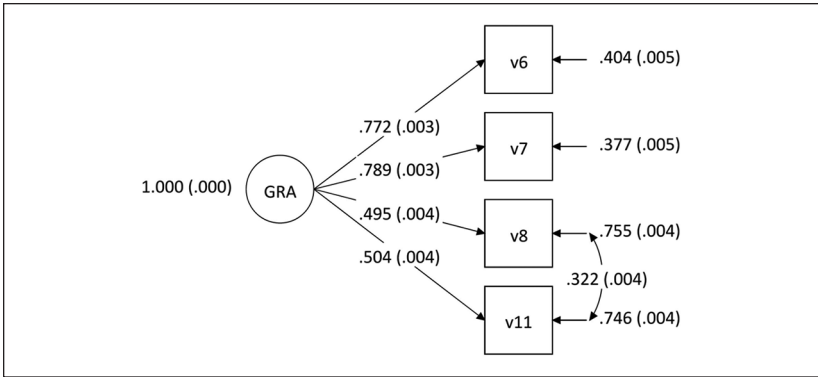


Figure 1. Confirmatory factor model “gender role attitudes,” ISSP 2012 data ($N = 58,767$), completely standardized solution, standard errors in parentheses, and errors variances are unobserved. For item labels, see Table 1.
 Note. ISSP = International Social Survey Programme; GRA = gender role attitudes.

adding an error correlation between Items v8 and v11. Both items tap into the domain of home life. The model fit the data well ($CFI = 1.000$, $RMSEA = .021$, $SRMR = .003$). This general model (Figure 1) served as a base model that should fit the data without considering any country differences and is a prerequisite to for finding a model structure that can be fitted in each country separately. Only when we find a structure that applies to each country, we may continue to test whether the measurements are actually comparable. Second, we tested the base model country by country. The model did not fit well and had low standardized factor loadings in Austria, India, Norway, and Mexico. Not being able to reproduce the base model led us to exclude these countries from all subsequent analyses because there was no common ground for testing comparability. In the 36 remaining countries, the basic model fit the data well.⁹

Multiple-Group Analysis

We continued to test for measurement invariance across the 36 countries with MGCEFA and began with the configural model. The model fit was well as can be seen in Table 2. Thus, equivalent measurement structures are present across countries. When we tested for metric measurement invariance. The fit slightly deteriorated but overall it remained acceptable. However, according to the criteria defined by Chen (2007), metric invariance was not supported by the data. Nonetheless, when we released the factor loadings of Items v7

Table 2. MGCFA Model Fit Across 36 Countries ($N = 52,984$).

Model	χ^2 (df)	CFI	RMSEA	SRMR	Δ CFI	Δ RMSEA	Δ SRMR
Configural	53.606 (36)	.999	.018	.005			
Metric	959.454 (141)	.974	.063	.055	-.025	.045	.050
Partial metric (v7 free)	667.572 (106)	.982	.060	.046	-.017	.042	.041
Partial metric (v7 and v11 free)	338.186 (71)	.992	.051	.034	-.007	.033	.029
Scalar	12,408.779 (246)	.617	.183	.211	-.375	.132	.177
Partial scalar (v8 and v11 free)	2,873.313 (176)	.915	.102	.068	-.077	.051	.034

Note. The partial metric models are compared with the metric model. The scalar and partial scalar models are compared with the second partial metric model. MGCFA = multigroup confirmatory factor analysis; df = degrees of freedom; CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean residual; Δ = difference.

and v11 across countries, we could achieve at least partial metric measurement invariance. This implies that comparisons of factor covariance or unstandardized regression estimates are possible. The test for scalar measurement invariance revealed a sharp deterioration of model fit. The variability of the intercepts was most apparent for Items v8 and v11, whereas the intercepts of other items were less strongly affected by noninvariance. We fitted a partial scalar invariance model by freeing the intercepts of these items. However, the model fit was still not acceptable. Thus, based on the analysis with MGCFA, comparisons of factor means across countries are precluded.

Another way to achieve comparability may be to drop countries that profoundly contribute to the overall model misfit. However, the choice of countries is rather arbitrary so we did not pursue this strategy. Rather, at this point, it was reasonable to switch to a different analytical technique.

Alignment

First, we used the FREE alignment approach with the ML, MLR, and Bayesian estimators. The difference between ML and MLR is that MLR standard error estimates are adjustment for possible nonnormality of the data. The difference of ML/MLR and Bayesian is that Bayesian procedures are based on Markov-chain Monte Carlo (MCMC) estimation and allow to formulate apriori hypotheses about the parameters of interest. However, in the current analysis, we do not make use of the possibility to formulate specific priors and apply a simpler "configural" Bayesian alignment method (Asparouhov &

Table 3. Noninvariant Parameters in the Alignment Analysis (Type = Fixed; N = 52,984).

Estimator	Factor loadings		Intercepts		Factor loadings and intercepts	
	Absolute (144)	Relative (%)	Absolute (144)	Relative (%)	Absolute (288)	Relative (%)
ML	38	26.4	40	27.8	78	27.1
MLR	32	22.2	38	26.4	70	24.3
Bayesian (configural method)	47	32.6	59	41.0	106	36.8

ML = maximum likelihood; MLR = robust maximum likelihood.

Muthén, 2014), which follows the procedure of ML alignment. We expect that the ML/MLR and Bayesian results are asymptotically equivalent.

In each case, the Mplus programme issued a standard error warning that the model may be poorly identified. As recommended by Asparouhov and Muthén (2014), we switched to the FIXED alignment procedure and used the Philippines as reference group (for this country the mean in the FREE model was closest to zero). Table 3 shows the amount of noninvariant parameters in the ML, MLR, and Bayesian solutions. The ML estimator reveals that the stricter cut-off value of 25% noninvariant parameters was exceeded for factor loadings and intercepts. Using the MLR estimator, only 22.2% of the factor loadings were noninvariant. However, the relative amount of noninvariant intercepts slightly surpassed 25%. Results of both ML and MLR estimation did not exceed the less strict cut-off value of 29% noninvariant parameters. However, the Bayesian estimator revealed a somewhat higher proportion of noninvariant parameters. Although Bayesian estimation is asymptotically equivalent to ML, this indicates poor convergence of the Marko-Chain Monte Carlo algorithm. However, increasing the number of Bayesian iterations did not reduce the difference to the ML and MLR estimates.

Because the stricter 25% cut-off criterion for the proportion of noninvariant parameters was slightly exceeded, we continued to test the trustworthiness of the ML and MLR alignment solutions with a simulation study. We used the estimates of the final model as starting values for data generation and monitored the correlation between the generated and estimated factor means. To test whether trustworthy alignment (i.e., the precision of the replication of factor means) depends on sample size, we simulated three sample sizes per group: 100 (3,600 in total), 500 (18,000 in total), and 1,500 (54,000 in total). With 1,500 observations per group, we simulate a total sample size

Table 4. Correlations of Generated and Estimated Factor Means Across 36 Groups (500 Replications).

Estimator	$N_g = 100$	$N_g = 500$	$N_g = 1,500$
ML	.9398	.9780	.9904
MLR	.9393	.9778	.9901

Note. N_g is the number of observations per group. ML = maximum likelihood; MLR = robust maximum likelihood.

similar to the empirical data example. Each simulation consisted of 500 replications. Table 4 presents the correlation between the generated and estimated factor means for ML and MLR estimation. In both cases and as expected, 1,500 observations per group were necessary to obtain sufficiently high correlations.

Although the results of MGCFA indicated that measurements of gender role attitudes are not fully comparable across countries (i.e., only partial metric invariance was supported), we put trust in the results of the more liberal alignment procedure, which are also supported by the Monte Carlo simulation. Accordingly, we conclude that despite the degree of noninvariance found in the data, the measurements of gender role attitudes are at least approximately comparable across countries and that comparisons of factor means are trustworthy. It, therefore, appears justified to continue our analysis and use the country-specific factor means of gender role attitudes to correlate them with the scores representing the country-specific value orientations. However, substantive conclusions should be drawn with caution, because the results are trustworthy only in this specific case. With other data or with a different composition of countries, the comparability must be assessed repeatedly.

Correlations of Gender Role Attitudes and Cultural Values

Table 5 displays the factor means across the 36 countries estimated by the alignment procedure in descending order (ML and MLR estimation provided exactly the same results). The ranking starts from 0, indicating the most traditional gender role attitudes. The Philippines (0.000), Korea (-0.793), Turkey (-0.797), Argentina (-0.873), and Chile (-0.875) are the countries where people tend to support more traditional views on gender roles based on a gendered separation of social roles, with women devoted to the domestic sphere and men to the public one. Conversely, at the bottom of the ranking we find Denmark (-3.359), Sweden (-3.174), Finland (-2.705), Iceland (-2.662), and Germany (-2.493), where people tend to support more

Table 5. Factor Means of Gender Role Attitudes in Descending Order, ML Fixed Alignment Models.

Ranking	Country abbreviations	ML estimation	
		Country (ISSP country code)	M
1	PH	Philippines (30)	0.000
2	KR	Korea (24)	-0.793
3	TR	Turkey (37)	-0.797
4	AR	Argentina (1)	-0.873
5	CL	Chile (8)	-0.875
6	RU	Russia (33)	-0.962
7	LV	Latvia (26)	-0.979
8	VE	Venezuela (40)	-1.091
9	HU	Hungary (18)	-1.100
10	BG	Bulgaria (5)	-1.120
11	CN	China (9)	-1.195
12	LT	Lithuania (25)	-1.365
13	IL	Israel (20)	-1.484
14	ZA	South Africa (41)	-1.571
15	SK	Slovakia (36)	-1.572
16	PL	Poland (31)	-1.627
17	PT	Portugal (32)	-1.633
18	CH	Switzerland (7)	-1.762
19	CZ	Czech Republic (10)	-1.803
20	HR	Croatia (17)	-1.874
21	SI	Slovenia (35)	-1.915
22	TW	Taiwan (38)	-1.956
23	JP	Japan (23)	-1.996
24	US	United States (39)	-2.006
25	GB	Great Britain (16)	-2.064
26	AU	Australia (3)	-2.126
27	BE	Belgium (4)	-2.223
28	FR	France (15)	-2.255
29	IE	Ireland (19)	-2.336
30	CA	Canada (6)	-2.457
31	NL	The Netherlands (28)	-2.485
32	DE	Germany (11)	-2.493
33	IS	Iceland (22)	-2.662
34	FI	Finland (14)	-2.705
35	SE	Sweden (34)	-3.174
36	DK	Denmark (12)	-3.359

Note. ML = maximum likelihood; ISSP = International Social Survey Programme.

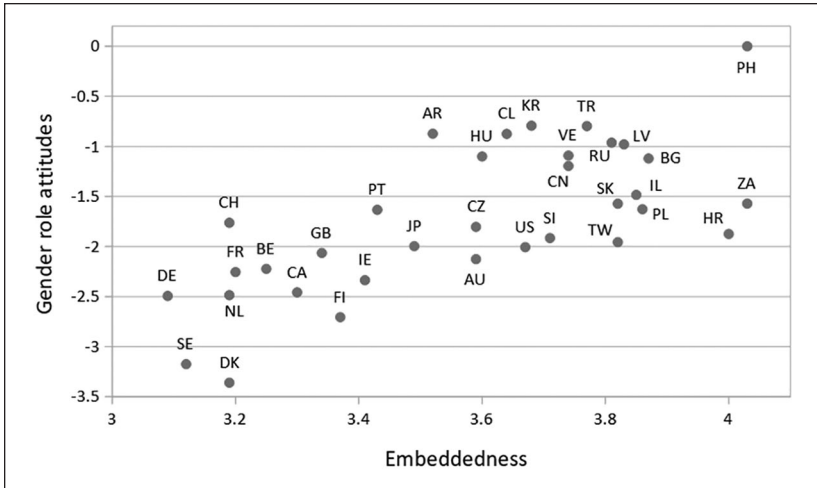


Figure 2. Relationship of the cultural value embeddedness and gender role attitudes (country-level data, $N = 34$).

egalitarian gender roles in the private and public sphere. The factor means are rather arbitrary and should not be used as substantive information.

The association between gender role attitude factor means and cultural values scores for embeddedness, hierarchy, and egalitarianism are shown in Figures 2 to 4.¹⁰ We observed a positive relationship between embeddedness and gender role attitudes ($r = .70$). Thus, the more societies emphasize the importance of the collective and status quo, the more they favor traditional gender roles. Some of the countries with the highest levels of embeddedness (e.g., Philippines, South Africa, Bulgaria, and Poland) favor a more traditional gender role model. On the contrary, the countries with the lowest scores on embeddedness (Netherlands, Denmark, Sweden, and Germany) are among those with the least traditional gender role attitudes. Hierarchy was positively related to gender role attitudes ($r = .51$). This implies that societies that adhere to a hierarchical system of societal roles are more traditional in their view on gender roles. Countries with the highest scores on hierarchy (e.g., Korea, Turkey, Russia, and Philippines) also show more traditional gender role attitudes. Countries with rather low scores on hierarchy (e.g., Finland, Belgium, and Germany) hold less traditional gender role attitudes. The relationship between egalitarianism and gender role attitudes was negative ($r = -.45$), implying that societies that emphasize the benefit and welfare of all its members to an equal degree do not hold traditional views on gender roles.

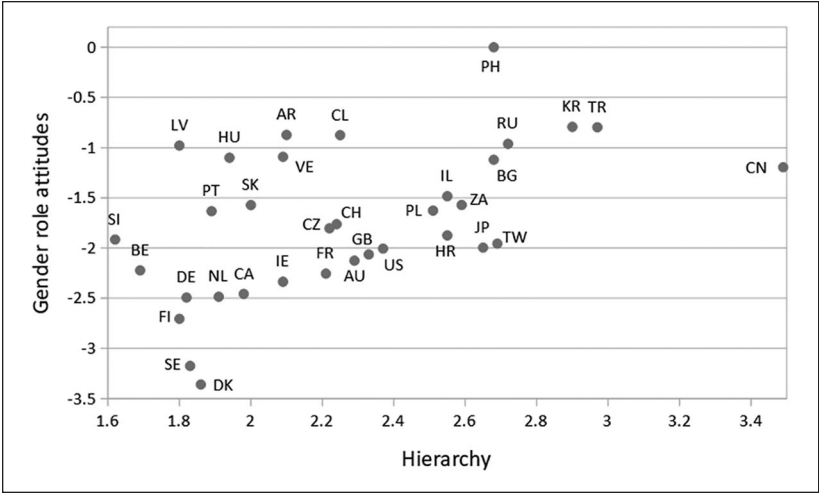


Figure 3. Relationship of the cultural value hierarchy and gender role attitudes (country-level data, $N = 34$).

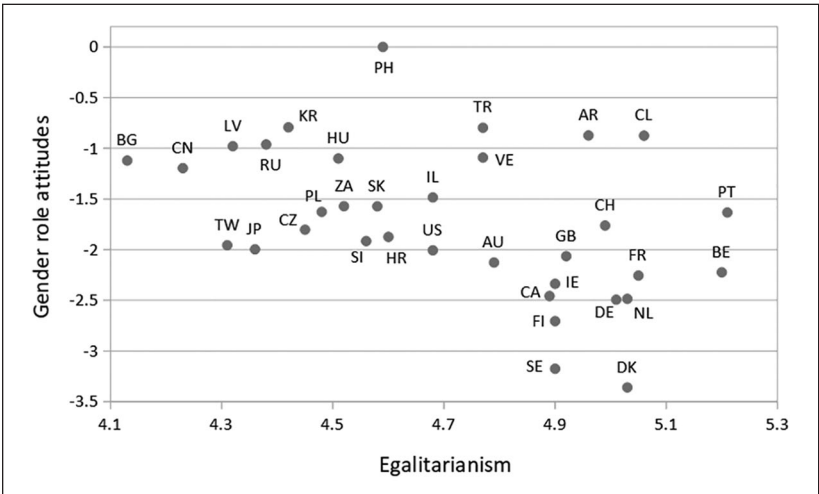


Figure 4. Relationship of the cultural value egalitarianism and gender role attitudes (country-level data, $N = 34$).

Among the countries with the highest scores on egalitarianism are some of those with the least traditional gender role attitudes (e.g., Belgium, France, Netherlands, Denmark, and Germany).

Summary and Discussion

Studying gender role attitudes means dealing with a complex matter. These attitudes refer to the roles people perceive as appropriate for men and women. Such beliefs are not only related to simple personal preferences but also are rooted in individual values transmitted through socialization processes. Previous research demonstrated that the nexus between individual gender role attitudes and the societal context is particularly relevant, both considering the structural and the cultural orientation of a country. On one hand, the societal structure of opportunity, such as the availability of daycare services, parental leave schemes, or labor market conditions, affect attitudes toward gender roles (André et al., 2013; Lomazzi et al., 2019; Kangas & Rostgaard, 2007; Sjöberg, 2004). On the other hand, gender role attitudes are connected with broader value orientations, which can differ vastly across societies according to their cultural, social, and political history (Inglehart, 1997; Kalmijn, 2003; Pfau-Effinger, 2004; Schwartz, 2006). This strict connection makes the measurement of gender role attitudes particularly sensitive to cultural bias, with the consequential risks of lacking comparability across societies. In this study, we focused on between-countries comparison and adopted the path-dependency approach (Pfau-Effinger, 2004) with a national perspective. However, even regions within a country can have followed different political and economic pathways. In this case, gender roles—and gender role attitudes as well—developed accordingly and this perspective can be applied by future research to investigate within-country invariance as well.

Despite the increase of awareness concerning the issue of measurement invariance in the field of the methodology of comparative social research, the practice of assessing measurement equivalence in substantive research is not yet very common. Nevertheless, this is a relevant matter if scholars aim to make proper cross-cultural comparisons and avoid the risk of elaborate theories based on misleading results (Billiet, 2003; Davidov et al., 2014; Horn & McArdle, 1992).

Positioned at the intersection between the interest in substantive comparative research on gender role attitudes and in the methodological development of the techniques to assess measurement equivalence, this study had two aims. The first aim was to test the cross-country equivalence of the popular measurement instrument of gender role attitudes utilized in the ISSP by adopting the novel alignment method in addition to the more traditional MGCFA assessment. The second aim was substantive and concerned with the explanation of cross-country differences in the prevalent views on gender roles by different cultural value orientations. However, in order to draw meaningful comparisons of gender role attitudes across countries, the underlying measurement first has to be tested for equivalence. Thus, the two aims

of the study are inextricably connected: Before comparative analyses can be conducted, we have to make sure that the construct of interest is actually comparable.

The results of the MGCFA approach revealed partial metric measurement invariance, indicating that the comparability of the measurements of gender role attitudes is limited to factor (co)variances and/or regression coefficients. Thus, our substantive correlational analysis at the macro-level, which is based on the country-specific factor means, would be precluded. However, the properties of the MGCFA approach may be too strict with regard to the condition that all cross-country parameter differences have to be exactly zero, especially in the case of a large-scale comparative study with many countries. Thus, we turned to the alignment optimization procedure to obtain a final model with the highest possible degree of measurement invariance while allowing cross-country parameter differences to be only approximately equal. The result is an alignment where only a few large and many small parameter differences exist and comparability of the factor means is still given. Although the results were technically equal across several estimation methods (i.e., ML, MLR, and Bayes), the Bayesian estimator revealed a somewhat higher proportion of noninvariant parameters. This may indicate that the Bayesian algorithm was not sufficiently converged also after increasing the computational effort. However, the ML and MLR results were very similar and close to the suggested cutoff values (25% or 29%) for the proportion of noninvariant parameters to assume approximate scalar measurement invariance. To validate this finding, we conducted Monte Carlo simulations to test whether the estimated factor means could be replicated in generated data with high precision. The results indicated that this was the case. Thus, we concluded that the factor means were trustworthy and it is reasonable to continue investigating our substantive research question.

The ranking of the factor means of gender role attitudes collocates the Philippines, Korea, and Turkey as the three most traditional countries with regard to gender beliefs while Denmark, Sweden, and Finland are the most egalitarian countries. Because gender role attitudes are supposed to be related to the predominant cultural values of a society, we expected them to be strongly correlated to the dimensions of embeddedness, hierarchy, and egalitarianism, derived by Schwartz' (2006) theory of cultural values orientations. The Pearson correlations between the gender role attitude factor means and the cultural values scores confirm our expectations: Societies with high levels of embeddedness, emphasizing the importance of the collective and status quo, as well as those with a strong preference in the maintenance of a hierarchical system of societal roles, tend to show more traditional gender role

attitudes. Societies that manifest egalitarianism as the predominant cultural value also display more egalitarian attitudes toward gender roles.

The assessment of measurement invariance is a fundamental step for cross-cultural studies, which build on the basic assumption of comparability. On one hand, this study aims at fostering this concern among scholars that are less familiar with methodological issues. Furthermore, the ISSP gender role scale is often used for comparative studies and our assessment of measurement invariance can provide useful indications for substantive researchers. On the other hand, the association between gender role attitudes and general value orientations can offer new insights for scholars interested in gender studies. Differently, from previous research, this study adopts Schwartz's (2006) theory of cultural values orientations as an alternative perspective to test the association between gender role attitudes and societal views. This new approach allows considering different components of society for understanding the different levels of gender equality across countries. In particular, in contrast to studies based on the modernization theory, which tend to emphasize structural and developmental aspects, this study enlightens the cultural components of societies. Future substantive studies may include these elements for explaining gender inequalities.

The relations between individual attitudes and the predominant cultural values could be further investigated, for example, by adopting structural equation models. In addition, to better understand the mechanisms underlying the transmission of gender cultures, similar studies could explore the relationship between individual attitudes, cultural values, and the development of specific gender regimes.

From a methodological perspective, the results of this article support the potentiality of the alignment method as a valuable alternative to assess measurement invariance to the traditional MGCFA. The procedure is easily implemented in the Mplus software, and it greatly automates the entire invariance test procedure. However, researchers are still required to carefully evaluate the models obtained with alignment, because the criteria to decide when comparability is given (i.e., proportion of noninvariant parameters, replication of estimated factor means) are not yet fully substantiated. Thus, more applicative and methodological studies are still needed before this method can be used alone.

Beyond the issue of *diagnosing* measurement invariance, there is the issue of *explaining* why some parameters seem to cause comparability problems. This question can be addressed using multilevel SEM (Cheung & Au, 2005; Davidov et al., 2012, 2016; Hox, 2010; Jak et al., 2014a, 2014b; Muthén, 1989, 1994; Rabe-Hesketh et al., 2004). Therefore, one may focus more strongly on the cultural differences, which, as we demonstrated, are connected

to cultural values. Research with respondents from different cultural backgrounds should be sensitive to the impact of differential cultural orientations when drawing comparisons based on survey data. This is especially the case when the questions used to measure gender-related topics (e.g., in this study, the item “A man’s job is to earn money; a woman’s job is to look after the home and family”) are phrased in a particular way, implying a rather traditionalist view. Thus, societal level information can yield important insights as to why certain measurement characteristics turn out to be barely comparable across countries.

Acknowledgments

The authors would like to thank Lisa Trierweiler for the English proof of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Schwartz (1994, 2006) identifies the values of egalitarianism, harmony, embeddedness, hierarchy, mastery, affective autonomy, and intellectual autonomy (for the details on the items and the methodological aspects, see Schwartz, 1994, 2006, 2008).
2. Such as European Values Study (EVS), International Social Survey Programme (ISSP), Eurobarometer (EB), World Values Survey (WVS), Generations and Gender Programme (GGP).
3. The seven items are displayed in the questionnaire in two separate batteries. The first battery contains Items 1 to 5, and the second contains Items 6 and 7.
4. The measurement model assessed by Constantin and Voicu (2015) included Items v6, v7, v8, v11, v13, and v14 (variable labels are those used in this article).
5. Factor loading of Item v5 by country: AR = .422, AT = .557, AU = .648, BE = .537, BG = .352, CA = .673, CH = .528, CL = .360, CN = .205, CZ = .541, DE = .446, DK = .475, FI = .670, FR = .510, GB = .543, HR = .551, HU = .364, IE = .683, IL = .536, IN = .190, IS = .537, JP = .522, KR = .189, LT = .392, LV = .310, MX = .112, NL = .633, NO = .674, PH = .076, PL = .547, PT = .442, RU = .256, SE = .642, SI = .252, SK = .458, TR = .227, TW = .379, US = .624, VE = .243, and ZA = .169; meaning of country

- abbreviations is given in Table 5.
6. Canada: English and French, Germany: East and West, Israel: Jewish and Arab, Switzerland: French and German.
 7. This should not be confused with the comparison of composite scores, which requires full scalar invariance (Steinmetz, 2013). The detection of noninvariant parameters can also be useful for the adjustment of measurement instruments (Lugtig et al., 2012).
 8. We provide all Mplus codes used for testing measurement invariance with MGCFA and the alignment procedure as supplementary materials. The data used in this study are available online (ISSP Research Group, 2016; Schwartz, 2008).
 9. In China, Czech Republic, Lithuania, Philippines, South Korea, Taiwan, and Turkey, the factor loading for Item v11 was below 0.30. However, we decided to keep these countries because overall model fit was good.
 10. The correlations of the values were as follows: embeddedness and hierarchy $r = .56$, embeddedness and egalitarianism $r = -.72$, hierarchy and egalitarianism $r = -.59$. Values data were not available for Iceland and Lithuania ($N = 34$).

References

- Albrecht, J. W., Edin, P.-A., & Vroman, S. V. (2000). A cross-country comparison of attitudes towards mothers working and their actual labor market experience. *Labour, 14*, 591–607. <https://doi.org/10.1111/1467-9914.00147>
- Alemán, J., & Woods, D. (2016). Value orientations from the world values survey how comparable are they cross-nationally? *Comparative Political Studies, 49*(8), 1039–1067. <https://doi.org/10.1177/0010414015600458>
- Allport, G. W. (1961). *Pattern and growth in personality*. Holt, Rinehart and Winston.
- Alwin, D. F. (2005). Attitudes, beliefs, and childbearing. In A. Booth & A. C. Crouter (Eds.), *The new population problem: Why families in developed countries are shrinking and what it means* (pp. 115–126). Lawrence Erlbaum.
- André, S., Gesthuizen, M., & Scheepers, P. (2013). Support for traditional female roles across 32 countries: Female labour market participation, policy models and gender differences. *Comparative Sociology, 12*, 447–476. <https://doi.org/10.1163/15691330-12341270>
- Asparouhov, T., & Muthén, B. O. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling, 2*, 495–508. <https://doi.org/10.1080/10705511.2014.919210>
- Banaszak, L. A., & Plutzer, E. (1993). Contextual determinants of feminist attitudes: National and subnational influences in Western Europe. *The American Political Science Review, 57*, 145–157. <https://doi.org/10.2307/2938962>
- Baxter, J., & Kane, E. W. (1995). Dependence and independence. A cross-national analysis of gender inequality and gender attitudes. *Gender & Society, 9*, 193–215. <https://doi.org/10.1177/089124395009002004>
- Bergh, J. (2006). Gender attitudes and modernization processes. *International Journal of Public Opinion Research, 19*, 5–23. <https://doi.org/10.1093/ijpor/edl004>

- Billiet, J. (2003). Cross-cultural equivalence with structural equation modelling. In J. A. Harkness, F. J. R. van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 247–264). John Wiley.
- Boehnke, M. (2011). Gender role attitudes around the globe: Egalitarian vs. traditional views. *Asian Journal of Social Science*, 39(1), 57–74.
- Bolzendahl, C. I., & Myers, D. J. (2004). Feminist attitudes and support for gender equality: Opinion change in women and men, 1974–1998. *Social Forces*, 83(2), 759–789.
- Braun, M. (1998). Gender roles. In J. W. Van Deth (Ed.), *Comparative politics: The problem of equivalence* (pp. 111–134). Routledge.
- Braun, M. (2008). Using Egalitarian items to measure men's and women's family roles. *Sex Roles*, 59, 644–656. <https://doi.org/10.1007/s11199-008-9468-5>
- Braun, M. (2009). The role of cultural contexts in item interpretation: The example of gender roles. In M. Haller, R. Jowell, & T. W. Smith (Eds.), *The international social survey programme, 1984–2009 : Charting the globe* (pp. 395–408). Routledge.
- Braun, M., Scott, J., & Alwin, D. F. (1994). Economic necessity or self-actualization? Attitudes toward women's labour-force participation in East and West Germany. *European Sociological Review*, 10, 29–47. <https://doi.org/10.1093/oxfordjournals.esr.a036314>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.
- Byrne, B. M., Shavelson, R. S., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105, 456–466. <http://dx.doi.org/10.1037/0033-2909.105.3.456>
- Carter, J. S., & Borch, C. A. (2005). Assessing the effects of urbanism and regionalism on gender-role attitudes, 1974–1998. *Sociological Inquiry*, 75(4), 548–563. <https://doi.org/10.1111/j.1475-682X.2005.00136.x>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14, 464–504. <https://doi.org/10.1080/10705510701301834>
- Cheung, W.-L., & Au, K. (2005). Applications of multilevel structural equation modeling to cross-cultural research. *Structural Equation Modeling*, 12, 598–619. https://doi.org/10.1207/s15328007sem1204_5
- Cieciuch, J., Davidov, E., Algesheimer, R., & Schmidt, P. (2017). Testing for approximate measurement invariance of human values in the European social survey. *Sociological Methods & Research*, 47, 665–686. <https://doi.org/10.1177/0049124117701478>
- Cieciuch, J., Davidov, E., Schmidt, P., Algesheimer, R., & 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. *Frontiers in Psychology*, 5, Article 982. <https://doi.org/10.3389/fpsyg.2014.00982>

- Constantin, A., & Voicu, M. (2015). Attitudes towards gender roles in cross-cultural surveys: Content validity and cross-cultural measurement invariance. *Social Indicators Research*, 123, 733–751. <https://doi.org/10.1007/s11205-014-0758-8>
- Cotter, D., Hermsen, J. M., & Vanneman, R. (2011). The end of the gender revolution? Gender role attitudes from 1977 to 2008. *American Journal of Sociology*, 117, 259–289. <https://www.jstor.org/stable/10.1086/658853>
- Cunningham, M., Beutel, A. B., Barber, J. S., & Thornton, A. (2005). Reciprocal relationships between attitudes about gender and social contexts during young adulthood. *Social Science Research*, 34, 862–892.
- Davidov, E., Cieciuch, J., Meuleman, B., Schmidt, P., Algesheimer, R., & Hausherr, M. (2015). The comparability of measurements of attitudes toward immigration in the European social survey: Exact versus approximate measurement equivalence. *Public Opinion Quarterly*, 79, 244–266. <https://doi.org/10.1016/j.sres-research.2005.03.001>
- Davidov, E., Dülmer, H., Cieciuch, J., Kuntz, A., Seddig, D., & Schmidt, P. (2016). Explaining measurement nonequivalence using multilevel structural equation modeling: The case of attitudes toward citizenship rights. *Sociological Methods & Research*, 47, 729–760. <https://doi.org/10.1177/0049124116672678>
- Davidov, E., Dülmer, H., Schlüter, E., Schmidt, P., & Meuleman, B. (2012). Using a multilevel structural equation modeling approach to explain cross-cultural measurement noninvariance. *Journal of Cross-Cultural Psychology*, 43, 558–575. <https://doi.org/10.1177/0022022112438397>
- Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, 40, 55–75. <https://doi.org/10.1146/annurev-soc-071913-043137>
- Davis, S. N., & Greenstein, T. N. (2009). Gender ideology: Components, predictors, and consequences. *Annual Review of Sociology*, 35, 87–105. <https://doi.org/10.1146/annurev-soc-070308-115920>
- Duncan, S. (1995). Theorizing European gender systems. *Journal of European Social Policy*, 5, 263–284. <https://doi.org/10.1177/095892879500500401>
- Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Lawrence Erlbaum.
- Flake, J. K., & McCoach, B. D. (2018). An investigation of the alignment method with polytomous indicators under conditions of partial measurement invariance. *Structural Equation Modeling*, 25, 56–70. <https://doi.org/10.1080/10705511.2017.1374187>
- Halman, L. (1995). *The need for theory in comparative research on values*. Work and Organization Research Centre, Tilburg University.
- Halman, L. (2010). Values. In H. K. Anheier & S. Toepler (Eds.), *International encyclopedia of civil society* (pp. 1599–1604). Springer.
- Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. SAGE.
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). SAGE.

- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*, 117–144. <https://doi.org/10.1080/03610739208253916>
- Hox, J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). Routledge.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. <https://doi.org/10.1080/10705519909540118>
- Inglehart, R. (1977). *The silent revolution: Changing values and political styles among western publics*. Princeton University Press.
- Inglehart, R. (1997). *Modernization and postmodernization: Cultural, economic, and political change in 43 societies*. Princeton University Press.
- Inglehart, R., & Norris, P. (2003). *Rising tide: Gender equality and cultural change around the world*. Cambridge University Press.
- ISSP Research Group. (2016). *International Social Survey Programme: Family and changing gender roles IV—ISSP 2012 (ZA5900 Data file Version 4.0.0.)*. Data Archive for the Social Sciences, GESIS. <https://doi.org/10.4232/1.12661>
- Jak, S., Oort, F. J., & Dolan, C. V. (2014a). Measurement bias in multilevel data. *Structural Equation Modeling, 21*, 31–39. <https://doi.org/10.1080/10705511.2014.856694>
- Jak, S., Oort, F. J., & Dolan, C. V. (2014b). Using two-level factor analysis to test for cluster bias in ordinal data. *Multivariate Behavioral Research, 49*, 544–553. <https://doi.org/10.1080/00273171.2014.947353>
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36*, 409–426. <https://doi.org/10.1007/BF02291366>
- Kalmijn, M. (2003). Country differences in sex-role attitudes: Cultural and economic explanations. In W. Arts & L. Halman (Eds.), *The cultural diversity of European unity: Findings, explanations and reflections from the European Values Study* (pp. 311–336). Brill.
- Kangas, O., & Rostgaard, T. (2007). Preferences or institutions? Work—Family life opportunities in seven European countries. *Journal of European Social Policy, 17*, 240–256. <https://doi.org/10.1177/0958928707078367>
- Kraaykamp, G. (2012). Employment status and family role attitudes: A trend analysis for the Netherlands. *International Sociology, 27*, 308–329. <https://doi.org/10.1177/0268580911423046>
- Kroska, A., & Elman, C. (2009). Change in attitudes about employed mothers: Exposure, interests, and gender ideology discrepancies. *Social Science Research, 38*, 366–382. <https://doi.org/10.1016/j.ssresearch.2008.12.004>
- Kunovich, R. M., & Kunovich, S. (2008). Gender dependence and attitudes toward the distribution of household labor: A comparative and multilevel analysis. *International Journal of Comparative Sociology, 49*, 395–427. <https://doi.org/10.1177/0020715208097787>
- Lee, S. K., Alwin, D. F., & Tufiş, P. A. (2007). Beliefs about women's labour in the reunified Germany, 1991–2004. *European Sociological Review, 23*(4), 487–503. <https://doi.org/10.1093/esr/jcm015>

- Lee, K. S., Tufiş, P., & Alwin, D. F. (2010). Separate spheres or increasing equality? Changing gender beliefs in postwar Japan. *Journal of Marriage and Family*, 72, 184–201. <https://doi.org/10.1111/j.1741-3737.2009.00691.x>
- Lek, K., Oberski, D., Davidov, E., Cieciuch, J., Seddig, D., & Schmidt, P. (2019). Approximate measurement invariance. In T. P. Johnson, B.-E. Pennell, I. A. L. Stoop, & B. Dorer (Eds.), *Advances in comparative survey methods: Multinational, multiregional, and multicultural contexts (3MC)* (pp. 911–929). John Wiley.
- Lewis, J. (2001). The decline of the male breadwinner model: Implications for work and care. *Social Politics: International Studies in Gender, State & Society*, 8(2), 152–169. <https://doi.org/10.1093/sp/8.2.152>
- Lomazzi, V. (2017a). Gender role attitudes in Italy: 1988–2008. A path-dependency story of traditionalism. *European Societies*, 19, 370–395. <https://doi.org/10.1080/14616696.2017.1318330>
- Lomazzi, V. (2018). Using alignment optimization to test the measurement invariance of gender role attitudes in 59 countries. *Methods, Data, Analyses (MDA)*, 12, 77–103. <https://doi.org/10.12758/mda.2017.09>
- Lomazzi, V., Israel, S., & Crespi, I. (2019). Gender equality in Europe and the effect of work-family balance policies on gender role attitudes. *Social Sciences*, 8(1), 5. <https://doi.org/10.3390/socsci8010005>
- Lugtig, P., Boeije, H. R., & Lensvelt-Mulders, G. J. L. M. (2012). Change? What change? An exploration of the use of mixed-methods research to understand longitudinal measurement variance. *Methodology*, 8, 115–123. <http://dx.doi.org/10.1027/1614-2241/a000043>
- Marsh, H. W., Guo, J., Parker, P. D., Nagengast, B., Asparouhov, B., Muthén, B., & Dicke, T. (2018). What to do when scalar invariance fails: The extended alignment method for multi-group factor analysis comparison of latent means across many groups. *Psychological Methods*, 23, 524–545. <https://doi.org/10.1037/met0000113>
- Marsh, H. W., Hau, W.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11, 320–341. https://doi.org/10.1207/s15328007sem1103_2
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525–543. <https://doi.org/10.1007/BF02294825>
- Moen, P., Erickson, M. A., & Dempster-McClain, D. (1997). Their mother's daughters? The intergenerational transmission of gender attitudes in a world of changing roles. *Journal of Marriage and Family*, 59(2), 281–293. <https://doi.org/10.2307/353470>
- Motiejunaite, A., & Kravchenko, Z. (2008). Family policy, employment and gender-role attitudes: A comparative analysis of Russia and Sweden. *Journal of European Social Policy*, 18, 38–49. <https://doi.org/10.1177/0958928707084453>
- Munck, I., Barber, C., & Torney-Purta, J. (2017). Measurement invariance in comparing attitudes toward immigrants among youth across Europe in 1999 and 2009:

- The alignment method applied to IEA CIVED and ICCS. *Sociological Methods & Research*, 47, 687–728. <https://doi.org/10.1177/0049124117729691>
- Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika*, 54, 557–585. <https://doi.org/10.1007/BF02296397>
- Muthén, B. O. (1994). Multilevel covariance structure analysis. *Sociological Methods & Research*, 22, 376–398. <https://doi.org/10.1177/0049124194022003006>
- Muthén, B. O., & Asparouhov, T. (2013). *BSEM measurement invariance analysis* (Mplus Web Notes 17). <http://www.statmodel.com/>
- Muthén, B. O., & Asparouhov, T. (2014). IRT studies of many groups: The alignment method. *Frontiers in Psychology*, 5, Article 978. <https://doi.org/10.3389/fpsyg.2014.00978>
- Muthén, B. O., & Asparouhov, T. (2017). Recent methods for the study of measurement invariance with many groups. *Sociological Methods & Research*, 47, 637–664. <https://doi.org/10.1177/0049124117701488>
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.).
- Pfau-Effinger, B. (2004). Socio-historical paths of the male breadwinner model—An explanation of cross-national differences. *The British Journal of Sociology*, 55, 377–399. <https://doi.org/10.1111/j.1468-4446.2004.00025.x>
- Powers, R. S., Suiitor, J. J., Guerra, S., Shackelford, M., Mecom, D., & Gusman, K. (2003). Regional differences in gender—Role attitudes: Variations by gender and race. *Gender Issues*, 21(2), 40–54. <https://doi.org/10.1007/s12147-003-0015-y>
- Rabe-Hesketh, S., Skrondal, S., & Pickles, A. (2004). Generalized multilevel structural equation modeling. *Psychometrika*, 69, 167–190. <https://doi.org/10.1007/BF02295939>
- Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin*, 114, 552–566. <https://doi.org/10.1037/0033-2909.114.3.552>
- Rhemtulla, M., Brosseau-Liard, P., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, 17, 354–373. <https://doi.org/10.1037/a0029315>
- Rokeach, M. (1968). *Beliefs, attitudes, and values: A theory of organization and change*. Jossey-Bass.
- Rokeach, M. (1973). *The nature of human values*. Free Press.
- Saris, W. E., Satorra, A., & Sörbom, D. (1987). The detection and correction of specification errors in structural equation models. *Sociological Methodology*, 17, 105–129. <https://doi.org/10.2307/271030>
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177. <http://dx.doi.org/10.1037/1082-989X.7.2.147>
- Scholz, E., Jutz, R., Edlund, J., Öun, I., & Braun, M. (2014). *ISSP 2012 family and changing gender roles IV: Questionnaire development* (GESIS Technical Reports 19). <https://www.gesis.org/angebot/publikationen/gesis-papers/>

- Schwartz, S. H. (1994). Are there universal aspects in the content and structure of values? *Journal of Social Issues, 50*, 19–45. <http://dx.doi.org/10.1111/j.1540-4560.1994.tb01196.x>
- Schwartz, S. H. (1999). Cultural value differences: Some implications for work. *Applied Psychology: An International Review, 48*, 23–47. <https://doi.org/10.1111/j.1464-0597.1999.tb00047.x>
- Schwartz, S. H. (2006). A theory of cultural value orientations: Explication and applications. *Comparative Sociology, 5*, 137–182. <https://doi.org/10.1163/156913306778667357>
- Schwartz, S. H. (2008). *National cultural value orientation scores*. https://www.researchgate.net/publication/304715744_The_7_Schwartz_cultural_value_orientation_scores_for_80_countries
- Schwartz, S. H., Bardi, A., & Bianchi, G. (2000). Value adaptation to the imposition and collapse of communist regimes in East-Central Europe. In S.A. Renshon & J. Duckitt (Eds.), *Political psychology* (pp. 217–237). Palgrave Macmillan. http://dx.doi.org/10.1057/9780230598744_13
- Scott, J., Alwin, D. F., & Braun, M. (1996). Generational changes in gender-role attitudes: Britain in a cross-national perspective. *Sociology, 30*, 471–492. <https://doi.org/10.1177/0038038596030003004>
- Seddig, D., & Leitgöb, H. (2018a). Approximate measurement invariance and longitudinal confirmatory factor analysis: Concept and application with panel data. *Survey Research Methods, 12*, 29–41. <http://dx.doi.org/10.18148/srm/2018.v12i1.7210>
- Seddig, D., & Leitgöb, H. (2018b). Exact and bayesian approximate measurement invariance. In E. Davidov, P. Schmidt, J. Billiet, & B. Meuleman (Eds.), *Cross-cultural analysis: Methods and applications* (2nd ed., pp. 553–579). Routledge.
- Sjöberg, O. (2004). The role of family policy institutions in explaining gender-role attitudes: A comparative multilevel analysis of thirteen industrialized countries. *Journal of European Social Policy, 14*, 107–123. <https://doi.org/10.1177/0958928704042003>
- Sokolov, B. (2018). The Index of Emancipative Values: Measurement Model Misspecifications. *American Political Science Review, 112*(2), 395–408. <https://doi.org/10.1017/S0003055417000624>
- Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*, 78–90. <https://doi.org/10.1086/209528>
- Steinmetz, H. (2013). Analyzing observed composite differences across groups: Is partial measurement invariance enough? *Methodology, 9*, 1–12. <http://dx.doi.org/10.1027/1614-2241/a000049>
- Stickney, L. T., & Konrad, A. M. (2007). Gender-role attitudes and earnings: A multinational study of married women and men. *Sex Roles, 57*, 801–811. <https://doi.org/10.1007/s11199-007-9311-4>

- Uunk, W., & Lersch, P. M. (2019). The effect of regional gender-role attitudes on female labour supply: A longitudinal test using the BHPS, 1991–2007. *European Sociological Review*, 35(5), 669–683. <https://doi.org/10.1093/esr/jcz026>
- Valentova, M. (2013). Age and sex differences in gender role attitudes in Luxembourg between 1999 and 2008. *Work, Employment and Society*, 27, 639–657. <https://doi.org/10.1177/0950017013481638>
- Van de Schoot, R., Kluytmans, A., Tummers, L., Lugtig, P., Hox, J., & Muthén, B. O. (2013). Facing off with Scylla and Charybdis: A comparison of scalar, partial, and the novel possibility of approximate measurement invariance. *Frontiers in Psychology*, 4, Article 770. <https://doi.org/10.3389/fpsyg.2013.00770>
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4–70. <https://doi.org/10.1177/109442810031002>
- Walter, J. G. (2018). The adequacy of measures of gender roles attitudes: A review of current measures in omnibus surveys. *Quality & Quantity*, 52, 829–848. <https://doi.org/10.1007/s11135-017-0491-x>
- Welzel, C. (2013). *Freedom rising: Human empowerment and the quest for emancipation*. Cambridge University Press.
- West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). Guilford Press.

Author Biographies

Vera Lomazzi is a senior researcher at the Data Archive for Social Sciences at GESIS—Leibniz Institute for Social Sciences in Cologne, Germany, and a member of the Executive Committee of the European Values Study. Her substantive research mainly focuses on the cross-cultural study of gender equality and gender role attitudes, gender mainstreaming. She has a specific interest in the quality of the instruments adopted by large cross-sectional survey programmes and on their measurement equivalence.

Daniel Seddig is a postdoctoral researcher at the Institute of Sociology and Social Psychology, University of Cologne, Germany. His research interests are cross-contextual and longitudinal data analysis with structural equation models, latent growth and mixture modeling, and Bayesian structural equation modeling. Substantive topics of his research are the value-attitude-behavior relationship, cross-cultural comparisons of attitudes, youth development, and deviance.