Chapter 3
Measurement Model and Invariance Testing of Scales Measuring Egalitarian Values in ICCS 2009

Daniel Miranda and Juan Carlos Castillo

Abstract Based on the conceptualization of democratic principles in the International Civic and Citizenship Education Study (ICCS) 2009, particularly attitudes concerning equal rights for disadvantaged groups, such as immigrants, ethnic groups and women, this chapter evaluates the extent to which the scales measuring attitudes toward gender equality, equal rights for all ethnic/racial groups and equal rights for immigrants are invariant, and to what extent they can be compared across the countries participating in the study. Multi-group confirmatory factor analysis is used to estimate a measurement model of the egalitarian attitudes and its measurement equivalence across the 38 countries (n = 140,000 8th grade students) that participated in ICCS 2009. The results indicate that the original scales are non-invariant. Nevertheless, with some modifications, the proposed conceptual model was found to be invariant across countries. The chapter concludes with a discussion of the theoretical and empirical implications of the model.

Keywords Ethnic minorities · Gender equality · Immigrants · International Civic and Citizenship Education Study (ICCS) · International large-scale assessments Measurement invariance

3.1 Introduction

The growing number of international comparative studies brings about several measurement issues. Do the item translations reflect the same meanings? Are some concepts country-specific? Do the items relate to the same or different constructs? All these, and other questions relate to the issue of validity of the comparisons, itself something that entails a paradox: in order to compare, we need to ensure that
the instrument is the same. In the context of international studies, several authors have shown the relevance of obtaining comparable measures and the potential consequences of not doing so (Guenole and Brown 2014). Measurement invariance assumes that the instrument (questionnaire) measures the same concept in the same way for different groups (Meredith 1993; Millsap 2011; Millsap and Everson 1993; Rutkowski and Svetina 2014; Van De Schoot et al. 2015; Vandenberg and Lance 2000).

The study of egalitarian attitudes of people in different countries and its comparability is a big challenge for the social sciences (Davidov et al. 2016). As established in Chap. 1, the study of egalitarian attitudes has a particular relevance for students of school age. The ICCS study has different scales to measure egalitarian attitudes toward three specific groups: immigrants, ethnic groups and women. Although the design of the study questionnaire follows a careful procedure in order to allow comparability, still the comparison between countries based on these measures has not yet been tested. This chapter uses multiple-group confirmatory factor analysis to evaluate the extent to which the scales measuring attitudes toward gender equality, equal rights for all ethnic/racial groups and equal rights for immigrants are equivalent across countries.

### 3.1.1 Measuring Attitudes Toward Equal Rights

The study of political attitudes as a reflection of a country’s political culture has a long tradition in social sciences, particularly in comparative studies. Tolerance, which is considered central to a democratic political culture, is a central attitude typically measured in this research tradition. Following this, several studies are oriented to characterize different countries in terms of the political attitudes and/or political behavior of their inhabitants; and to compare their political cultures.

Considering the set of available international studies, these can be classified as those that are aimed at adult populations and those that are aimed at the young populations. Studies aimed at measuring different political attitudes, beliefs and preferences in the population aged above eighteen years old include the International Social Survey Program (ISSP), the World Value Survey (WVS), the European Social Survey (ESS) and the Latin American Public Opinion Project (LAPOP). There are relatively few comparative studies focused on young populations barring those developed by the IEA, whose studies include the Civic Education Study (CIVED) and the Civic and Citizenship Education Study (ICCS; see [http://iccs.iea.nl/](http://iccs.iea.nl/)).

Within this set of identified studies, the measure of tolerance takes different forms. For instance, LAPOP considers the support for the right to vote for people with views extremely critical of a country’s system of government. In the case of WVS, tolerance is defined as the extent to which people support a public office position and/or publically demonstrate in support of excluded populates. The ICCS questionnaire is oriented to capture beliefs and attitudes about the rights of three
social groups: immigrants, ethnic groups and women, considering a set of items that measure the degree to which people support equal rights for different groups in society (Schulz et al. 2008; Van Zalk and Kerr 2014). This chapter adopts the ICCS measures of tolerance.

Few studies consider attitudes toward equal rights for specific groups as a measure of tolerance as a democratic principle (Barber et al. 2013; Bridges and Mateut 2014; Isac et al. 2012; Janmaat 2014; Dotti Sani and Quaranta 2017; Strabac et al. 2014; Van Zalk and Kerr 2014). For instance, Barber et al. (2013) considered equal rights attitudes toward immigrants, specifically, as a relevant aspect of pro-social civic engagement. In the same vein, Dotti Sani and Quaranta (2017) evaluated attitudes toward gender equality considering that this type of equality and its support were relevant aspects of human development. Despite advances in the study of attitudes toward equal rights in recent years, even with the same ICCS data, there are still a number of aspects that could be improved. First, most of the studies focus on one specific group, mainly migrants, and less frequently on equal rights for women (Bolzendahl and Coffé 2009; Dotti Sani and Quaranta 2017) leaving aside the interrelation between attitudes toward equality of different groups. For instance, are the individuals who show larger support toward immigrants the same as those who support equal rights for women? Secondly, most of the reviewed studies in this field do not test for measurement equivalence even though they perform country comparative analysis. The present chapter aims to overcome some of these limitations by addressing three target groups simultaneously (women, migrants and ethnic minorities), as well as by testing equivalence of the measurement model of equal rights attitudes.

3.1.2 Measurement and Equivalence

As with most concepts in the social sciences, attitudes toward equal rights are not observed directly but rather are hypothetical constructs. Given that these attitudinal concepts should be measured as latent constructs (Bollen 2002), the latent approach implies that the hypothetical underlying constructs are captured by a set of observable indicators by using statistical techniques. For instance, confirmatory factor analysis (Bollen 2002; Hoyle 2014), one of the most extended approaches, allows the evaluation of the proposed latent measures.

One of the main challenges in comparative studies is to achieve the statistical equivalence of measures across groups, such as societies, allowing meaningful comparability (Davidov et al. 2014; Millsap and Meredith 2007). The evaluation of the comparability is technically known as measurement invariance (Millsap 2011) or measurement equivalence (Davidov et al. 2014). In the remainder of this chapter, we adopt the term measurement invariance. The wider socioeconomic, sociocultural and/or sociopolitical differences of the respondents demand the development of studies that follow strict technical criteria in order to improve comparability. Statistical techniques may be used to assess measures and improve comparisons.
Multi-group confirmatory factor analysis (MGCFA) is one of the most recognized techniques for assessing measurement invariance. This statistical tool allows the evaluation of the comparability of measures through the sequential estimation of different models that represent levels of invariance with increasing constraints. In the sequence, the first level is the configural invariance. The configural model assumes that the construct of the latent is measured by the same indicators in all groups. This is the baseline model that evaluates the configuration of latent variables but does not warrant any comparison across groups (Beaujean 2014). The second level is the metric invariance. The metric model, also known as weak invariance, constrains the factor loading to be the same across all groups and, in that way, evaluates whether the indicators have the same strength in the measure of the latent variables. This level of invariance is considered the minimal condition for comparison. It allows to compare only the relation of measured latent variables with other covariates (Beaujean 2014; Davidov et al. 2014; Desa 2014). The third level is scalar invariance, also known as strong invariance. This level, in addition to loadings, constrains the intercepts or thresholds (for categorical variables) to be the same across groups. This level allows for meaningful comparisons of levels (averages) of the latent measured across groups and comparisons of the relation of latent variables with other covariates. In that sense, the scalar invariance level allows rankings that compare averages across groups or the use of statistical models, such as regression or multilevel modeling, that compare relational patterns across groups. Finally, there is a fourth level, labeled strict invariance. This level adds the constraint of error variance across groups, increasing the comparability of latent scales. Nevertheless, given that scalar invariance is sufficient for meaningful comparisons between group means and covariate patterns, the strict invariance level is often not estimated (Beaujean 2014; Davidov et al. 2014).

The present study follows both the CFA and MGCFA approaches. CFA was used to evaluate the latent structure of the gender rights attitudes, immigrants’ rights attitudes and ethnic rights attitudes that make up the egalitarian attitudes model. MGCFA enabled us to evaluate the comparability of the latent measures across countries.

3.2 Methods

3.2.1 Data

As outlined in Chap. 2, our study used data from the ICCS 2009 database. The final sample showed small variations because the set of variables involved in these analyses have a specific missing pattern. Given that, the final sample consisted of 138,605 students from 38 countries.
3.2.2 Variables

The variables used as indicators for the dimensions are related to the students’ opinions about equal rights for immigrants, ethnic groups and women (Table 3.1). In each case we provide the item code used in the ICCS 2009 User Guide, Supplement 1 (Brese et al. 2014). The scale of gender equality considers three items that refer to equality between men and women in participatory government, rights and equal payment. The original scale has items that refer to male supremacy, such as “Women should stay out of politics” (IS2P24C), “Men are better qualified to be

| Measures of tolerance in ICCS 2009 | ICCS code |
|-----------------------------------|-----------|
| Gender equality attitudes: there are different views about the roles of women and men in society. How much do you agree or disagree with the following statements? |
| 1. Strongly disagree | Men and women should have equal opportunities to take part in government | IS2P24A |
| 2. Disagree | Men and women should have the same rights in every way | IS2P24B |
| 3. Agree | Men and women should get equal pay when they are doing the same jobs | IS2P24E |
| 4. Strongly agree | | |

| Ethnic equality attitudes: there are different views on the rights and responsibilities of different <ethnic/racial groups> in society. How much do you agree or disagree with the following statements? |
| 1. Strongly disagree | All ethnic/racial groups should have an equal chance to get a good education in <country of test> | IS2P25A |
| 2. Disagree | All ethnic/racial groups should have an equal chance to get good jobs in <country of test> | IS2P25B |
| 3. Agree | Schools should teach students to respect members of all ethnic/racial groups | IS2P25C |
| 4. Strongly agree | Members of all ethnic/racial groups should have the same rights and responsibilities | IS2P25E |

| Immigrant equality attitudes: people are increasingly moving from one country to another. How much do you agree or disagree with the following statements about <immigrants>? |
| 1. Strongly disagree | Immigrants should have the opportunity to continue speaking their own language | IS2P26A |
| 2. Disagree | Immigrants who live in a country for several years should have the opportunity to vote in elections | IS2P26C |
| 3. Agree | Immigrants should have the opportunity to continue their own customs and lifestyle | IS2P26D |
| 4. Strongly agree | Immigrants should have all the same rights that everyone else in the country has | IS2P26E |

Notes The wording for the items varies among countries. See the ICCS 2009 User guide for the international database, Supplement 1 (Brese et al. 2014) for further information
political leaders than women” (IS2P24F) and “When there are not many jobs available, men should have more right to a job than women” (IS2P24D). These items were not used in the measure of attitudes toward gender equality.

There are four items in the scale for equality of ethnic groups: education, employment, respect and rights. Finally, there are four items for measuring equality of immigrants: the right to speak your native language, the opportunity to vote, the right to maintain your own lifestyle, and equality of rights with all others in the country.

### 3.2.3 Analytical Strategy

The analytical strategy consisted of three steps. The first step involved the evaluation of the measurement adequacy of the scales using CFA. Secondly, we evaluated the measurement invariance of the proposed scales using MGCFA. Finally, we present descriptive statistics of the newly created scale and a cross-country comparison of the country averages.

In order to evaluate the goodness of fit for each country model using CFA, we implemented a chi-square test as an initial procedure. This index is used to test the reasonability of the measurement hypothesis “in terms of how well the solution reproduces the observed variances and covariances among the input indicators” (Brown 2006, p. 41), although we note that this index has been criticized as less sensible for large samples (Brown 2006; Rutkowski and Svetina 2014). In order to circumvent this weakness, we also used three other indicators: the comparative fit index (CFI), Tucker-Lewis index (TLI), and root mean square error of approximation (RMSEA). Brown (2006) proposed a set of cut-off point criteria for evaluating a good model’s fit: \( \leq 0.06 \) in the case of RMSEA and closer to 0.95 or greater for CFI and TLI. As an alternative, Brown (2006) proposed that CFI and TLI values in the range 0.90–0.95 could be considered acceptable.

In the case of MGCFA, the evaluation of the model and the invariance testing were evaluated sequentially. The configural model (baseline), estimates the same configuration of items for each group. The metric model, estimates the model constraining the factor loadings to the same value for each group. Finally, the scalar model constrains the factor loadings and intercepts to be the same for each group. In each of the three cases, we evaluated the model fit using the criteria proposed by Brown (2006). We used the change in the fit indexes between the higher to lower levels of invariance to test the invariance. The main index used to account for the invariance was based on a chi-square test, where the relative change is evaluated. For instance, when comparing the baseline model (configural) with a more constrained model (metric), an increase in chi-square indicating a degradation of the model can be expected. If the degradation of the constrained model is statistically significant, then the proposed model is non-invariant. Nevertheless, this index has the same weakness as noted for CFA, namely it is less sensible for large sample sizes (Brown 2006). Rutkowski and Svetina (2014) developed guidelines more
appropriate for international large-scale assessment (ILSA) involving several countries (more than 20 groups in this case); we therefore used these to evaluate model fit. Specifically, Rutkowski and Svetina (2014) advised that the difference in the fit indexes between two successive levels of invariance (for example configural vs. metric) must be $\leq 0.020$ for any of the three indexes (CFI, TLI and RMSEA).

Finally, we used the classical test theory functions (CTT) package (Willse 2014) to estimate latent measures and enable the rescaling of the original scores. This is an established R package for this type of analysis. The rescaling was adjusted to reach a mean = 50 and standard deviation (SD) = 10. We used this scale for our descriptive and country comparisons.

3.3 Results

In this section, we first present the general results regarding the extent to which the empirical indicators correspond to the theoretical measurement model of egalitarian attitudes, tested by CFA procedure for each country. Second, we examine the results of the multi-group analyses and the equivalence of measures across the countries tested. We conclude with the patterns of equality attitudes within countries.

3.3.1 Proposed Scale: Single-Country Analyses and Invariance Testing

The CFA analyses for each country indicate that the proposed measurement structure for the modified egalitarian attitudes model is confirmed for the 38 countries. All groups show good fit indexes (see Table 3.2). However, there are some countries where the TLI or RMSEA are slightly below the cutoff point, even though remaining within an acceptable range; for instance, Spain has a TLI < 0.95 (0.934) and a RMSEA > 0.06 (0.084).

The invariance testing indicates good fit indexes for the configural, metric and scalar level of the egalitarian attitudes model; all the fit indexes were above the cutoff criteria (see Table 3.3). The relative comparison between the configural and metric model indicates that metric invariance was achieved (Table 3.3). The differences in CFI, TLI and RMSEA were acceptable according to the criteria of Rutkowski and Svetina (2014). This level of invariance permits us to conduct comparable correlational analyses.

Our results suggest a good fit for the metric level of invariance, indicating factor loadings were stable across countries for the three equality measures (see Fig. 3.1). Looking at the scale measuring immigrants’ attitudes toward equal rights, the items IMMRGT1 and IMMRGT4 show factor loadings <0.2 in only one country.
Table 3.2  Confirmatory factor analysis of fit indexes of the proposed egalitarian attitudes model, by country

| Country               | Chi square | df | n   | CFI  | TLI  | RMSEA |
|-----------------------|------------|----|-----|------|------|-------|
| Malta                 | 5915.35    | 41 | 2112| 0.951| 0.934| 0.058 |
| Spain                 | 19755.34   | 41 | 3276| 0.951| 0.935| 0.084 |
| Estonia               | 13443.20   | 41 | 2712| 0.957| 0.942| 0.072 |
| Dominican Republic    | 5566.87    | 41 | 4259| 0.960| 0.946| 0.036 |
| Indonesia             | 10742.94   | 41 | 5006| 0.960| 0.946| 0.046 |
| Italy                 | 21459.43   | 41 | 3357| 0.962| 0.948| 0.077 |
| Liechtenstein         | 4674.39    | 41 | 355 | 0.964| 0.952| 0.107 |
| Belgium (Flemish)     | 20042.72   | 41 | 2962| 0.964| 0.952| 0.077 |
| Slovenia              | 20563.33   | 41 | 3054| 0.967| 0.956| 0.073 |
| Lithuania             | 15058.34   | 41 | 3893| 0.973| 0.963| 0.051 |
| Netherlands           | 10817.54   | 41 | 1909| 0.973| 0.963| 0.061 |
| Denmark               | 35044.37   | 41 | 4355| 0.974| 0.966| 0.071 |
| Thailand              | 15040.33   | 41 | 5261| 0.974| 0.965| 0.043 |
| Cyprus                | 15900.80   | 41 | 3076| 0.976| 0.968| 0.055 |
| Norway                | 37857.19   | 41 | 2917| 0.976| 0.968| 0.087 |
| Sweden                | 42171.56   | 41 | 3410| 0.976| 0.968| 0.085 |
| Czech Republic        | 39389.75   | 41 | 4621| 0.978| 0.970| 0.068 |
| China                 | 33873.45   | 41 | 5151| 0.979| 0.972| 0.058 |
| Slovakia              | 20461.83   | 41 | 2966| 0.979| 0.971| 0.060 |
| Latvia                | 11096.33   | 41 | 2743| 0.980| 0.974| 0.044 |
| Ireland               | 30732.35   | 41 | 3313| 0.981| 0.974| 0.066 |
| Finland               | 34579.55   | 41 | 3292| 0.982| 0.976| 0.068 |
| Greece                | 18853.35   | 41 | 3103| 0.982| 0.976| 0.051 |
| Bulgaria              | 14939.47   | 41 | 3187| 0.983| 0.977| 0.044 |
| Russia                | 22388.74   | 41 | 4289| 0.983| 0.978| 0.046 |
| Luxembourg            | 35287.40   | 41 | 4780| 0.985| 0.980| 0.052 |
| Poland                | 28409.12   | 41 | 3242| 0.985| 0.979| 0.057 |
| Switzerland           | 19665.77   | 41 | 2907| 0.985| 0.979| 0.050 |
| Austria               | 19464.16   | 41 | 3366| 0.986| 0.981| 0.045 |
| Korea                 | 36756.66   | 41 | 5249| 0.986| 0.981| 0.049 |
| Chile                 | 22177.99   | 41 | 5160| 0.987| 0.983| 0.036 |
| England               | 37732.70   | 41 | 2881| 0.987| 0.982| 0.065 |
| Hong Kong, SAR        | 42586.49   | 41 | 2816| 0.989| 0.986| 0.062 |
| Mexico                | 31577.74   | 41 | 6464| 0.989| 0.986| 0.036 |
| New Zealand           | 27638.17   | 41 | 3874| 0.989| 0.985| 0.044 |
| Colombia              | 26650.95   | 41 | 6108| 0.990| 0.987| 0.032 |
| Paraguay              | 9196.75    | 41 | 3229| 0.991| 0.988| 0.025 |
| Guatemala             | 14430.70   | 41 | 3950| 0.993| 0.991| 0.024 |

*df* degrees of freedom, *n* number of students sampled, *CFI* comparative fit index, *TLI* Tucker-Lewis index, and *RMSEA* root mean square error of approximation
The gender equality scale also shows very stable factor loadings, except for item GENDEQ5, which has a lower factor loading only for one country. Finally, for the ethnic equality scale, the factor loadings are very stable for all items in all countries. In spite of the variations in factor loading, the fit indexes indicate that those variations are within the acceptable range of non-invariance.

In the same vein, the relative comparison between the metric and scalar models indicates that scalar invariance is achieved. The differences in the CFI, TLI and RMSEA values were all within the cutoff range suggested by Rutkowski and Svetina (2014) (see Table 3.3). This level of invariance permits comparisons of the averages of the egalitarian attitudes to be made across countries.

Table 3.3 Fit indexes of the original model of egalitarian attitudes

| Model      | CFI   | TLI   | RMSEA | Δ CFI  | Δ TLI  | Δ RMSEA |
|------------|-------|-------|-------|--------|--------|---------|
| Configural | 0.980 | 0.974 | 0.056 |        |        |         |
| Metric     | 0.982 | 0.979 | 0.049 | 0.002  | 0.005  | −0.007  |
| Scalar     | 0.969 | 0.975 | 0.054 | −0.013 | −0.004 | 0.005   |

CFI comparative fit index, TLI Tucker-Lewis index, and RMSEA root mean square error of approximation

Fig. 3.1 Loading distributions of scale indicators in each country. Each dot represents the loading for each indicator for each scale.
3.3.2 Average Country Comparison

Given the scalar level of comparability achieved in measurement invariance analyses, and using the information provided by the multigroup confirmatory models, we estimated scales for gender equality attitudes, immigrant equality attitudes and ethnic equality attitudes. The scales were saved using the option “save fscores” available in Mplus 7.4 (Muthén and Muthén 2017). The saved latent variables have a mean of 0 and standard deviation of 1. For a better illustration of averages and posterior modeling of variables, the latent measures were rescaled to a mean of 50 and standard deviation of 10.

For gender equality attitudes, Spain, Liechtenstein, Chile, Sweden, Austria and Chinese Taipei presented the highest averages (see Fig. 3.2). For immigrant equality attitudes, Mexico, Guatemala, Chinese Taipei, Colombia, Chile, Paraguay and the Dominican Republic presented the highest averages (most of these countries being located in Latin America) (see Fig. 3.2). For ethnic equality attitudes, Chinese Taipei, Guatemala, Chile, Paraguay, Mexico, Colombia, the Dominican Republic and Luxembourg showed the highest averages (again, this list is dominated by countries in Latin America) (see Fig. 3.2). Conversely, the lowest averages for gender equality attitudes were shown in Lithuania, Latvia, Thailand, the Russian Federation and Indonesia. The lowest average immigrant equality attitudes were recorded for Indonesia, Latvia, Liechtenstein, England, Belgium (Flemish) and the Netherlands. For ethnic equality attitudes, the lowest averages were shown in Belgium (Flemish), Cyprus, Finland, Malta, the Czech Republic, Latvia and the Netherlands (see Fig. 3.2).

Fig. 3.2 Average distribution of equality attitudes by country
There is a considerable variation in the averages, indicating that some proportion of the egalitarian attitude variance occurs at the country level (Fig. 3.2). Furthermore, given the nested sample design of the ICCS study, another portion of the variance can be associated with the school level.

In order to describe the decomposition of variance, a three-level model was estimated, allowing for the estimation of the proportion of variance associated with each level of the analyses. The three equality attitudes show proportions of variance associated with the country level above 10%. The variance associated with school level is, in the three cases, around 5% (see Table 3.4).

A complementary result shows that the egalitarian attitudes of participation are correlated with different strengths between each other, which indicates that the types of participation do not function independently of each other. For instance, the average correlation between gender equality attitudes and immigrant equality attitudes is 0.64 (min = 0.39, max = 0.86), the average correlation between gender equality attitudes and ethnic equality attitudes is 0.72 (min = 0.61, max = 0.94), and the average correlation between immigrant equality attitudes and ethnic equality attitudes is 0.80 (min = 0.55, max = 0.95).

### Table 3.4 Variance decomposition of egalitarian attitudes

| Level       | Gender equality (%) | Immigrant equality (%) | Ethnic equality (%) |
|-------------|---------------------|------------------------|--------------------|
| Country level | 14.1                | 10.6                   | 9.8                |
| School level | 5.3                 | 5.2                    | 5.2                |
| Individual level | 80.6             | 84.2                   | 85.0               |

3.4 Discussion and Conclusions

We aimed to evaluate the measurement model of egalitarian attitudes proposed by the ICCS study, and to test its operationalization and comparability using the international survey data. The model is founded on a theoretical framework that considers tolerance as an orientation to the acceptance of or respect for other social groups, more specifically as the degree to which people support equal rights for different groups in society. The tested model considers egalitarian attitudes toward three specific social groups: immigrants, ethnic groups and women. The analyses were performed using data provided by the countries that participated in ICCS 2009.

Our results for the proposed measurement model of egalitarian attitudes indicate that the scales are invariant across the analyzed countries at scalar level. This implies that all latent variables have the same structure in the analyzed countries. Moreover, their scalar invariant structure allows for direct comparisons of the mean scores and correlates of the latent variables across countries (Beaujean 2014; Davidov et al. 2014). The confirmation of the structure of the model indicates that the included dimensions are useful for evaluating egalitarian attitudes within the context of the family unit, at the school level, or at the country level, allowing meaningful comparisons.
Finally, the contextual dependence of variance (at country and school level) justifies the multilevel modeling statistical technique as an adequate strategy for the estimation of any explanatory model.

References

Barber, C., Fennelly, K., & Torney-Purta, J. (2013). Nationalism and support for immigrants’ rights among adolescents in 25 countries. *Applied Developmental Science, 17*(2), 60–75. https://doi.org/10.1080/10888691.2013.774870.

Beaujean, A. A. (2014). *Latent variable modeling using R: A step-by-step guide*. New York, NY: Routledge.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology, 53*(1), 605–634. https://doi.org/10.1146/annurev.psych.53.100901.135239.

Bolzendahl, C., & Coffé, H. (2009). Citizenship beyond politics: The importance of political, civil and social rights and responsibilities among women and men. *The British Journal of Sociology, 60*(4), 763–791. https://doi.org/10.1111/j.1468-4446.2009.01274.x.

Brese, F., Jung, M., Mirazchiyski, P., Schulz, W., & Zuehlke, O. (2014). *ICCS 2009 user guide for the international database* (2nd ed.). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Bridges, S., & Mateut, S. (2014). Should they stay or should they go? Attitudes towards immigration in Europe. *Scottish Journal of Political Economy, 61*(4), 397–429. https://doi.org/10.1111/sjpe.12051.

Brown, T. A. (2006). *Conﬁrmatory factor analysis for applied research*. New York, NY: Guilford Press.

Davidov, E., Dümler, 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 and Research, 49*(12), 651–675. Retrieved from https://doi.org/10.1177/0049124116672678.

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Sociology, 40*(1), 55–75. Retrieved from https://doi.org/10.1177/0038038513503240.

Desa, D. (2014). Evaluating measurement invariance of TALIS 2013 complex scales: Comparison between continuous and categorical multiple-group confirmatory factor analyses. OECD Education Working Papers, No. 103. Paris: OECD Publishing. https://doi.org/10.1787/5j2kbbvbl7k-en.

Dotti Sani, G. M., & Quaranta, M. (2017). The best is yet to come? Attitudes toward gender roles among adolescents in 36 countries. *Sex Roles, 77*, 30–45. https://doi.org/10.1007/s11199-016-0698-7.

Guenole, N., & Brown, A. (2014). The consequences of ignoring measurement invariance for path coefficients in structural equation models. *Frontiers in Psychology, 5*. 980. https://doi.org/10.3389/fpsyg.2014.00980.

Hoyle, R. H. (2014). *Handbook of structural equation modeling*. New York, NY: Guilford Publications.

Isac, M. M., Maslowski, R., & Van der Werf, G. (2012). Native students attitudes towards equal rights for immigrants. A study in 18 European countries. *JSSE: Journal of Social Science Education, 11*(1). https://doi.org/10.2390/jsse-v11-i1-1189. Retrieved from http://www.jsse.org/index.php/jsse/article/view/1189/1092.

Janmaat, J. G. (2014). Do ethnically mixed classrooms promote inclusive attitudes towards immigrants everywhere? A study among native adolescents in 14 countries. *European Sociological Review, 30*(6), 810–822. https://doi.org/10.1093/esr/jct075.
Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525–543. Retrieved from http://link.springer.com/article/10.1007/BF02294825.

Millsap, R. E. (2011). *Statistical approaches to measurement invariance* (1st ed.). New York, NY: Routledge.

Millsap, R. E., & Everson, H. T. (1993). Methodology review: Statistical approaches for assessing measurement bias. *Applied Psychological Measurement, 17*(4), 297–334. Retrieved from http://apm.sagepub.com/content/17/4/297.short.

Millsap, R. E., & Meredith, W. (2007). Factorial invariance: Historical perspectives and new problems. In R. Cudeck & R. C. MacCallum (Eds.), *Factor analysis at 100: Historical developments and future directions* (pp. 131–152). Mahwah, NJ: Lawrence Erlbaum Associates.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user’s guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement, 74*(1), 31–57. https://doi.org/10.1177/0013164413498825.

Schulz, W., Fraillon, J., Ainley, J., Losito, B., & Kerr, D. (2008). *International civic and citizenship education study, assessment framework*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Strabac, Z., Aalberg, T., & Valenta, M. (2014). Attitudes towards Muslim immigrants: Evidence from survey experiments across four countries. *Journal of Ethnic and Migration Studies, 40*(1), 100–118. https://doi.org/10.1080/1369183X.2013.831542.

Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Editorial: Measurement invariance. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.01064. Retrieved from https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01064/full.

Van Zalk, M. H. W., & Kerr, M. (2014). Developmental trajectories of prejudice and tolerance toward immigrants from early to late adolescence. *Journal of Youth and Adolescence, 43*(10), 1658–1671. https://doi.org/10.1007/s10964-014-0164-1.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4–70. Retrieved from http://orm.sagepub.com/content/3/1/4.short.

Willse, J. T. (2014). CTT: Classical test theory functions. Software. Retrieved from https://CRAN.R-project.org/package=CTT.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.