

Profiles of tolerance and respect for the rights of diverse social groups among youth.

Comparison across countries.

Pamela Isabel INOSTROZA FERNANDEZ

Supervisor: Prof. Femke De Keulenaer
Co-supervisor: Dr. Maria Magdalena Isac

Thesis presented in
fulfillment of the requirements
for the degree of Master of Science
in Statistics and Data Science

Academic year 2020-2021



© Copyright by KU Leuven

Without written permission of the promoters and the authors it is forbidden to reproduce or adapt in any form or by any means any part of this publication. Requests for obtaining the right to reproduce or utilize parts of this publication should be addressed to KU Leuven, Faculteit Wetenschappen, Celestijnenlaan 200H - bus 2100 , 3001 Leuven (Heverlee), Telephone +32 16 32 14 01.

A written permission of the promoter is also required to use the methods, products, schematics and programs described in this work for industrial or commercial use, and for submitting this publication in scientific contests.

Preface

This work was very interesting to perform, it grants me the opportunity to study in a deeper sense the subpopulations that are hidden behind large-scale assessments. I learned that not only variable center studies can give important insights about the different profiles that each country is composed by but that person center approach is a very powerful tool. I had to study many topics regarding civic and citizenship attitudes. I cannot express how grateful I am to my promoters, particularly Dr. Maria Magdalena Isac who provided me with the idea to research this topic along with many insights that I should focus on. Professor Femke de Keulenaer was a great teacher and mentor regarding the statistical techniques I should apply and analyze to achieve our main objective. Their support and involvement in this project were at the right level with the difficulty. They gave me the reasoning and motivated me to complete this research. I want to thank them and state that they will have my respect and gratitude for the work performed.

Summary

Civic education is an important subject for every citizen in our modern society. It is important that every individual acknowledge the importance of the civil rights and obligations of any citizen. One of the most commented topics is how society faces and behaves towards the great diversity of individuals and cultures. Students are a great population to be studied as they are forming their own mindset and attitudes. Using ICCS 2016, an international large-scale assessment, it is possible to identify which are the most common behaviors among students' attitudes considering different aspects of equality towards women and ethnic and racial groups. As expected, the most common pattern is composed of students that share a high chance to accept and promote equality towards women and ethnics groups. Nonetheless, there is a small number of students that tend to disagree with this equality. Another set of students shares a high level of agreement with both minorities' equal rights but do not agree with their political role in society. Student's endorsement towards women's rights: The remaining set of students shares a high level of agreement towards equality in basic rights but favor towards men when competing for jobs or political roles. Another pattern identified towards ethnic groups is students that disagree with their equal right to have good jobs. These patterns are similar across the 14 countries studied in Europe, but they differ in the number of individuals in each pattern.

List of abbreviations

ICCS: International Civic and Citizenship Education Study.

ILSA: International Large-Scale Assessments

IEA: International Association for the Evaluation of Educational Achievement.

GLMM: Generalized Linear Mixed Model

LCA: Latent Class Analysis

CFA: Confirmatory Factor Analysis

Table of Contents

1	Introduction	1
2	Theoretical background	4
2.1	International Large-Scale Assessments	4
2.1.1	IEA - ICCS 2016	4
2.1.2	Students' endorsement of equal rights and opportunities	6
2.2	Mixture models Latent Class Analysis	7
2.2.1	Person-centered approach	7
2.2.2	Latent Class Analysis	8
2.3	Measurement invariance	9
3	Methods	12
3.1	Methodological features	12
3.2	Model evaluation	13
3.3	Analytical strategy	16
3.4	Data	17
4	Results	21
4.1	Students' endorsement of gender equality scale	22
4.1.1	Analysis by country	22
4.1.2	General model	26
4.1.3	Country comparability	30
4.1.4	Confirmatory Latent Class Analysis	33
4.2	Students' endorsement of equal rights for all ethnic/racial groups scale	35
4.2.1	Analysis by country	35
4.2.2	General model	38
4.2.3	Country comparability	43
4.2.4	Confirmatory Latent Class Analysis	46
5	Discussion and conclusion	48
A	Appendix	51
A.1	Complementary tables	51
A.2	Detailed output	53
A.3	Syntax	65
	References	81

List of Tables

4.1	Best model, fit statistics individual country model Students' endorsement of gender equality	23
4.2	Model fit statistics LCA Students' endorsement of gender equality	27
4.3	Probabilities to agree each item in the four-class global model for Students' endorsement of gender equality	30
4.4	Class sizes four-class global model for Students' endorsement of gender equality	30
4.5	Multigroup model fit statistics global model with four-classes with Students' endorsement of gender equality	32
4.6	Class sizes for partial homogeneity model for Students' endorsement of gender equality scale	33
4.7	Probabilities to agree each item in a 4-class Confirmatory LCA for Students' endorsement of gender equality scale	34
4.8	Model fit statistics 4-class Confirmatory LCA for Students' endorsement of gender equality scale	34
4.9	Best model, fit statistics individual country model Students' endorsement of equal rights for all ethnic/racial groups	36
4.10	Model fit statistics LCA Students' endorsement of equal rights for all ethnic/racial groups scale	40
4.11	Probabilities to agree each item four-class global model for Students' endorsement of equal rights for all ethnic/racial groups scale	43
4.12	Class sizes four-class global model for Students' endorsement of equal rights for all ethnic/racial groups scale	43
4.13	Multigroup model fit statistics, global model with four-classes for Students' endorsement of equal rights for all ethnic/racial groups scale	44
4.14	Class sizes for partial homogeneity model for Students' endorsement of equal rights for all ethnic/racial groups scale	45
4.15	Probabilities to agree each item 4-class Confirmatory LCA Students' endorsement of equal rights for all ethnic/racial groups scale	46
4.16	Model fit statistics Confirmatory LCA Students' endorsement of equal rights for all ethnic/racial groups scale	47
A.1	Items for students' endorsement of equal rights and opportunities. ICCS 2016 .	51
A.2	Countries sample sizes that participate in ICCS 2016	52
A.3	Model fit statistics LCA by country Students' endorsement of gender equality .	53
A.4	Thresholds 4-class Confirmatory LCA Students' endorsement of gender equality	58
A.5	Class sizes 4-class Confirmatory LCA Students' endorsement of gender equality	58
A.6	Model fit statistics LCA by country Students' endorsement of equal rights for all ethnic/racial groups scale	59

A.7	Thresholds 4-class Confirmatory LCA Attitude towards ethnic and race equal rights scale	64
A.8	Class sizes 4-class Students' endorsement of equal rights for all ethnic/racial groups scale	64

List of Figures

3.1	Distribution of derived scales Students' endorsement of equal rights and opportunities ICCS 2016	18
4.1	Bivariate standardized residuals individual country models for Students' endorsement of gender equality	24
4.2	Classes for best individual country model for Students' endorsement of gender equality	25
4.3	Bivariate model fit standardized residuals global model for Students' endorsement of gender equality	28
4.4	Comparative conditional probabilities to agree in 3 to 6 latent class global models for Students' endorsement of gender equality	29
4.5	Conditional probabilities to agree in a 4-class complete homogeneous multigroup model for Students' endorsement of gender equality scale	32
4.6	Bivariate standardized residuals for individual country models for Students' endorsement of equal rights for all ethnic/racial groups scale	37
4.7	Classes for best individual country model for Students' endorsement of equal rights for all ethnic/racial groups	38
4.8	Bivariate model fit standardized residuals global model for Students' endorsement of equal rights for all ethnic/racial groups scale	41
4.9	Comparative conditional probabilities to agree in 3 to 6 latent class global models for Students' endorsement of equal rights for all ethnic/racial groups	42
4.10	Conditional probabilities to agree in a 4-class complete homogeneous multigroup model for Students' endorsement of equal rights for all ethnic/racial groups scale	45
A.1	Response categories probabilities and class size for 4-classes global model for Attitude towards gender equality scale	58
A.2	Response categories probabilities and class size for 4-classes global model for Attitude towards ethnic and race equal rights scale	64

Chapter 1

Introduction

The development of civic values and attitudes of tolerance and respect for the rights of diverse social groups among youth is essential for sustainable democratic societies. These values are strongly promoted by families, educational systems, and international organizations across the world. The measurements and comparison of these attitudes among youth can provide valuable information about their development in different societies.

International studies such as the International Civic and Citizenship Education Study (ICCS)¹ (Wolfram Schulz, Carsten, Losito, & Fraillon, 2018) provide extensive and unique comparative information regarding these aspects. The ICCS study is a large-scale assessment (survey) applied in more than 25 educational systems during the last three cycles (1999, 2009, 2016) and focused on secondary education (representative samples of 8th graders, 14-year-olds in each country) addressing topics such as citizenship, diversity, and social interactions at school.

International and national ICCS study reports as well as previous empirical research using ICCS data have been largely focused on average country comparisons of attitudinal measures such as attitudes toward equal rights for immigrants, ethnic minorities and women, norms of good citizenship behavior and political participation. Most of these studies employed variable-centered analyses such as factor analysis and were able to provide valuable comparative information regarding average attitudinal levels in different societies (i.e., tapping into the overall tendency of respondents to indicate a higher response category when rating a set of items). Nevertheless, the variable-centered approach is less useful when researchers, driven by theoretical assumptions, aim to identify the most likely patterns of responses by participants. Indeed, some recent studies applied to ICCS data started to show the usefulness of person-centered approaches (i.e., latent class analysis, hereafter LCA) aimed at identifying profiles of young people's attitudes. For example, using ICCS 2009 data, (Hooghe, Oser, & Marien, 2016) build on political science theories of good citizenship norms to compare profiles of good citizenship norms across 38 countries and distinguished distinctive subgroups of the population that share a common understanding of what constitutes good citizenship were identified (e.g., who express either engaged or duty-based citizenship norms).

Nevertheless, such studies (i.e., employing LCA with ICCS data and focusing on patterns within a particular type of attitude described by individual items) are still scarce. This is especially true for the fields of educational and political sciences and particularly when cross-country comparisons are involved. Indeed, this scarcity of empirical studies employing LCA can be linked to some extent to disciplinary traditions favoring variable-centered approaches but most importantly to the novelty and complexity of LCA in these disciplines especially when

¹<https://www.iea.nl/studies/iea/iccs>

cross-cultural comparability (i.e., measurement invariance) must be evaluated and ensured.

In this thesis I aim to address this gap by providing a comprehensive overview of LCA and its application to international large-scale assessment (ILSA) data and by illustrating the procedure to conduct LCA (including measurement invariance testing in LCA) using R and Mplus with a step-by-step example. As a substantive background to this example, I approach a topic of high societal relevance, young people's attitudes of tolerance and respect for the rights of different and often marginalized groups such as ethnic minorities and women. More specifically, I focus on the operationalization of these attitudes in the ICCS 2016 study's international student questionnaire: a) young people's attitudes toward equal rights for women and b) young people's attitudes toward equal rights for ethnic/race groups. Moreover, I apply LCA to a heterogeneous (in terms of linguistic, geographic, economic and social contexts) sample of 14 European countries participating in ICCS 2016 with the aim of answering the following research questions:

- What profiles of attitudes towards gender equality are identified among adolescents in the 14 European countries?
- What profiles of tolerance and respect for equal rights of ethnic/race groups are observed among adolescents in the 14 European countries?
- Are these two sets of profiles fully comparable/measurement invariant across the 14 European countries?

With this analysis, this thesis aims to contribute to the current state-of-the art in both substantive and procedural ways. On the substantive side, I intend to add to the body of research focused on young people's attitudes of tolerance and respect for the rights of ethnic groups and women by pointing out information that is not reported in extent research. More specifically, unlike previous research that illustrated average attitudinal differences among countries (e.g., showing that on average adolescents tend to be tolerant), this study is able to point out response's patterns within a particular type of attitude described by individual items (e.g., distinguishing between different groups of adolescents with particular sets of views). On the procedural side, I aim to facilitate, for a wider group of scholars, the application of LCA to ILSA data with a comprehensive description of the statistical method and an easy to replicate, step-by-step procedure for its implementation in R and Mplus.

Overview of this thesis

Including this introduction, this thesis consists of five chapters and an appendix.

In the second chapter, I introduce the features of ILSAs, the ICCS 2016 study, and the background for the example used in this research. This chapter also includes an overview of LCA in comparison to person-centered methods and details different approaches to measurement invariance in LCA.

The third chapter describes the methodology applied in this research including the analytical approaches, the characteristics of the sample, and the study's measurement instruments.

Chapter four illustrates the results concerning the two main topics analyzed: a) young people's attitudes toward equal rights for women, and b) young people's attitudes toward equal rights for ethnic/race groups. Different sets of results concerning country-specific analysis and cross-country comparability are reported.

Finally, a discussion of the findings is provided in chapter five and the appendix includes detailed supplementary information regarding the methods and results and the detailed syntax and output examples.

Chapter 2

Theoretical background

Two main categories of topics are relevant to be discussed in this chapter, that taps into the theoretical background of this thesis: the background for the substantive example used and the features of the methodology to be applied. To this end, I will first introduce the characteristics of international large-scale assessments (ILSA) with the example of ICCS 2016 study and I will outline the operationalization of the main concepts analyzed in this research. Second, I will review literature on mixture models in order to identify and point out the most suitable technique and statistics to consider when analyzing this particular type of ILSA data. More specifically, mixed models will be introduced, particularly Latent Class Analysis which is a model-based approach to cluster individuals/cases into distinctive groups, called latent classes, based on their responses to a set of observed categorical variables (Goodman, 1974). Different approaches to measurement invariance (MI) (*Invariance analyses in large-scale studies*, 2019) in LCA will also be discussed.

2.1 International Large-Scale Assessments

International Large-Scale Assessments (ILSAs) have been used to draw comparisons among countries on a variety of topics in education and, more broadly, for example, in adolescent development (Isac, Palmerio, & Werf, 2019). These assessments can inform the public about influential factors on the micro and macro levels, foster interdisciplinary and international collaboration, and provide important data for studying the context and processes of education and development.

2.1.1 IEA - ICCS 2016

The International Association for the Evaluation of Educational Achievement (IEA) International Civic and Citizenship Education Study (ICCS) produces internationally comparative data collected via student, school, and teacher questionnaires. Data from different waves of the ICCS survey is publicly available to researchers. The first time this study was applied was in 1999 to 28 countries and it was called CIVED, the second wave started using the name ICCS and was implemented in 2009 in 38 countries, the last study was performed in 2016 in 24 countries. The next cycle is scheduled for 2022 and 25 countries will participate. The main aim of the ICCS study is to investigate how young people are prepared to undertake their roles as citizens in a range of countries in the second decade of the 21st century (Wolfram Schulz, Carsten, Losito, & Fraillon, 2018). ICCS study evaluates students' knowledge and

understanding of civics and citizenship, as well as their attitudes, perceptions, and activities related to civics and citizenship.

ICCS 2016 addressed the following research questions (Carsten & Schulz, 2018):

1. The way civic and citizenship education is implemented in participating countries, including the aim and principles for this learning area, the curricular approaches chosen to provide it, and changes and/or developments since 2009.
2. The extent of student's knowledge and understanding of civics and citizenship, and the factors associated with its variation across and within countries.
3. Student's current and expected future involvement in civic-related activities, their perceptions of their capacity to engage in these activities, and their perception of the value of civic engagement.
4. Student's belief about contemporary civic and civic issues in society, including those concerned with civic institutions, rules, and social principles (democracy, citizenship, and diversity), as well as their perceptions of their communities and threats to the world's future.
5. The ways in which schools organize civic and citizenship education, with a particular focus on general approaches, the processes used to facilitate civic engagement, interaction with their communities, and schools' and teacher's perceptions of the role of this learning area.

The 2016 study gathered data from more than 94000 students in 8th grade in about 3800 schools from 24 countries. Also, data from more than 37000 teachers in those schools and contextual data collected from school principals are included. An additional European questionnaire gathered data from almost 53000 students in 14 European countries and a Latin American student questionnaire from more than 25000 students from 5 Latin American countries. The student population is defined as students in 8th grade, in average 13.5 years of age in this study.

Of all 24 participant countries, 16 are from Europe, 5 are from Latin America, and 3 from Asia. In two of the participant countries, a sub-national entity participates. In Belgium, ICCS 2016 was implemented only in the Flemish education system and North Rhine-Westphalia state in Germany took part as a benchmarking participant.

The school's samples were designed as stratified two-stage cluster samples, first schools were randomly selected at the first stage with probability proportional to the size and intact classrooms were sampled at the second stage. Each country has a sample size of 150 schools approximately and a sample of students around 3000 and 4500. Additionally, around 15 teachers teaching the target grade from each school were sampled.

It is required at a minimum, that an analyst carrying out statistical analysis with ICCS have a good understanding of the conceptual foundations of ICCS (Wolfram Schulz, Carsten, Losito, & Fraillon, 2018), the themes addressed, the populations targeted, the samples selected, the instruments used, and the production of the international database. All of this information is described in practical terms in the user guide (Köhler, Weber, Brese, Schulz, & Carsten, 2018). Researchers using the database also need to make themselves familiar with the database structure and its variables.

For any secondary analysis of the study's data such as the planned for this thesis, it must be taken into account the following (Köhler, Weber, Brese, Schulz, & Carsten, 2018) :

1. The unequal selection probabilities of the sampling units that necessitate the use of weights during computation of estimates;
2. The complex multistage cluster sample design that was implemented to ensure a balance between the research goals and cost-efficient operations; and
3. The rotated design of the civic knowledge test, wherein students completed only samples of the test items rather than the full set of test items.

The use of this type of data is tailored mainly to be used with specific statistical packages, that support complex survey design, computing population estimates and design-based standard errors.

2.1.2 Students' endorsement of equal rights and opportunities

In the context of the ICCS study, young people's attitudes toward equal rights for women and equal rights for ethnic/race groups were included in the first affective-behavioral domain of the assessment framework, attitudes towards the rights and responsibilities of groups in society. Given the profound challenges experienced by societies worldwide today and particularly the perceived backlash in tolerance toward minority groups and gender equality, information tapping into such attitudes in youth is highly relevant to both policy makers and researchers. At the policy level, for example, these topics are of priority for both the European educational policy agenda (reflected by the Strategic Framework for European Cooperation in Education and Training towards the European Education Area (2021 - 2030)) (European Union, 2021) as well as the UNESCO's Education 2030 Agenda on sustainable goals (UNESCO, 2020). For researchers, these are topics of high interest for gaining insights into the future of democratic societies and into information that can help inform educational practices targeting low levels of intolerance.

Indeed, several empirical studies were conducted in order to compare young people's attitudes toward equal rights for women and equal rights for ethnic/race groups among several countries (Dotti Sani & Quaranta, 2017; Isac, Miranda, & Sandoval-Hernández, 2018; Isac, Palmerio, & Werf, 2019; Munck, Barber, & Torney-Purta, 2018). CFA, a variable-centered methodology, was performed to evaluate country invariance in (Isac, Palmerio, & Werf, 2019), which was achieved using additionally attitudes towards immigration indicators, this type of analysis is helpful to identify global behaviours of the whole population in the country and be able to compare them across countries. However, with this methodology, it is not possible to dig deeper into different subpopulations in each country and explain those average values.

The rise of engaged citizenship study (Hooghe & Oser, 2015) used latent class analysis to identify groups supporting duty-based and engaged citizenship norms, based on 12 indicators. The research included 21 countries, although no invariance analysis was evaluated to compare these groups across countries. The analysis was performed both in 1999 and in 2009 scales and the sizes of the groups in both periods are compared globally and for each country.

Variable-centered studies were able to show that, on average, young people in different countries, tend to display high levels of egalitarian attitudes (i.e., attitudes toward equal rights for women and equal rights for ethnic/race groups) and also point out countries that stand out

(in terms of lowest and highest average scores). Such findings also indicated high variability across and within countries (especially when individual items were examined) pointing to the possibility that different sizeable groups of young people may show particular attitudinal configurations (e.g., egalitarian, non-egalitarian). However, empirical and appropriate person-centered studies (e.g., LCA) with evaluations of such expectations are absent in the literature. Not much research using these particular indicators has been performed in order to identify how many subpopulations can be identified in each country regarding young people's beliefs about equal rights and opportunities for different groups in society based on gender and ethnic/racial background. Neither how the patterns or behaviours of these subpopulations are composed and which of them are representatives and comparable across countries.

2.2 Mixture models Latent Class Analysis

Parameters that describe a factor's effects in an ordinary generalized linear model are called fixed effects. Fixed effect applies to all categories of interest, gender, treatments, or any other manifest grouping variable. By contrast, random effects apply to a sample of all possible categories. GLM extends ordinary regression by allowing non normal responses and a link function of the mean. The generalized linear mixed model (GLMM) is a further extension that permits random and fixed effects in the linear predictor.

In this type of analysis, a contingency table is treated as a finite mixture of unobserved tables generated under a conditional independence structure at categories of a latent variable (Agresti, 2013). A GLMM with discrete data creates a mixture of linear predictor values using a latent variable, in this case, the unobserved random effect vector instead of being continuous and assumed to have a normal distribution, it is a qualitative mixture distribution.

2.2.1 Person-centered approach

ANOVA, multiple regression, mixed models are variable-centered approaches that focus on relations among variables and assume that the sample studied come from a homogeneous population. Mixture models (finite mixture models) have taken the place of the framework for a person-center analytic approach. The difference between these two approaches is that the person-centered approach focuses on identifying unobserved subpopulations composed of similar individuals or cases and involves modeling a mixture outcome distribution.

The most common technique to find homogeneous groups based on observed variables is cluster analysis (Srivastava, 2002). There are different methods that can be used to identify these groups of cases, but they are lacking in giving statistical indices and tests for the optimal number of clusters. The most common techniques to determine the best number of groups are based on tabular or graphical interpretation of the researcher.

The mixture of different distributions indicates population heterogeneity, this means that observations arise from a finite number of unobserved subpopulations in the target populations. This is key when we want to identify different patterns in the sample.

Latent class modeling defines a model for the probability of having a particular response pattern. This probability is a weighted average (or mixture) of the class-specific probabilities for these patterns. The item responses of an individual are mutually independent given the individual's class membership. Like cluster analysis, it is possible to assign individuals to the latent classes. The probability of belonging to a particular class given the responses (posterior

class membership probability) can be obtained by the Bayes' rule (Vermunt, 2014).

In summary, mixture modeling provides an important complement to traditional variable-centered analytical approaches. It offers the opportunity for researchers to identify unknown a priori homogeneous classes of individuals based on the measures of interest, examine the features of heterogeneity across the classes, evaluate the effects of covariates on the class membership, assess the relationship between the class membership and other outcomes, and study transitions between the latent class memberships over time. As a matter of fact, person-centered approaches and variable-centered approaches can be integrated into a general mixture modeling framework so that one can better understand the relationships among variables and the pattern of such relationships (Muthen & Muthen, 2000).

2.2.2 Latent Class Analysis

As indicated previously a mixture model assumes that some of its parameters differ across unobserved subgroups, latent classes, or mixture components and particularly a latent class model is a mixture model for a set of categorical items. The first LCA approach was improved by (Lazarsfeld & Henry, 1968) and (Goodman, 1974).

A latent class model assumes the existence of a latent categorical variable such that the observed response variables are conditionally independent, given that variable. In other words, LCA can directly assess the theory that distinctive groups of people share specific attitudes (Hagenaars & McCutcheon, 2002). Depending on the response variable in the model, the analysis is called Latent Profile Analysis if it is continuous (Normal) and Latent Class Analysis if the response variable is categorical (Multinomial) (Muthen & Muthen, 2000). In this study categorical outcomes will be used.

The goal of LCA is to identify unobserved subgroups based on similar response patterns. In contrast with cluster analysis, LCA is a model-based approach to classify. It identifies subgroups based on posterior membership probabilities rather than dissimilarity measures such as Euclidean or Mahalanobis distance (Vermunt, 2014). The general probability model underlying LCA allows for formal statistical procedures for determining the number of clusters, and more interpretable results stated in terms of probabilities.

LCA assumes conditional independence, that the observed categorical indicators are mutually independent once the categorical latent variable is conditioned out. Assuming the conditional independence, the joint probability of all observed indicator variables is described as (Wang & Wang, 2020):

$$P(u_1, u_1, \dots, u_Q) = \sum_{k=1}^K P(C = k)P(u_1|C = k)P(u_2|C = k) \dots P(u_Q|C = k) \quad (2.1)$$

From Bayes' formula, the posterior probabilities for each individual to be in different classes are estimated as:

$$P(C = k|u_1, u_2, \dots, u_Q) = \frac{P(C = k)P(u_1|C = k)P(u_2|C = k) \dots P(u_Q|C = k)}{P(u_1, u_2, \dots, u_Q)} \quad (2.2)$$

where $P(C = k)$ are the unconditional probabilities ($\sum_{k=1}^K P(C = k) = 1$) and $P(u_Q|C = k)$ are the conditional probabilities.

The unconditional probabilities are latent class probabilities, and the average of the probabilities can be interpreted as the prevalence of latent class (relative frequency of class membership)

or the proportion of the population expected to belong to a latent class (B. Muthén & Shedden, 1999). The conditional probabilities are conditional item-response probabilities, measurement parameters, representing the likelihood of endorsing specific characteristics of the observed items, given a specific class membership.

Conditional probabilities close to 1.0 indicate that members in the corresponding latent class endorse a category of the item; on the contrary, a very small probability indicates that they do not endorse the characteristic of the item. When a conditional item-response probability is close to $1/J$, where J is the number of categories in the item, the conditional probability is considered as random probability, thus the latent class membership is not predictive of the patterns of item response.

The conditional item-response probability is defined as in (2.3) and (2.4).

$$P(u_q = u_{qj} | C = k) = \frac{1}{1 + \exp(-L_{jk})} \quad (2.3)$$

$$L_{jk} = \ln\left(\frac{P_{jk}}{1 - P_{jk}}\right) \quad (2.4)$$

which is the logit for u_{qj} given in latent class k . A logit of 0 means that the conditional item probability $P_{jk} = 0.5$, when the logit takes an extreme value as -15 then $P_{jk} = 0$. On the contrary, a logit with a positive extreme value 15, $P_{jk} = 1$. These conditional item response probabilities provide information about how the latent classes differ from each other, for this reason, are used to define the estimated classes.

2.3 Measurement invariance

Measurement invariance can be defined as a conditional independence property of the measurement model with respect to a set of sub-populations within the parent population (e.g., gender, countries, or time). Measurement invariance is an important prerequisite for using multi-indicator assessment instruments to examine group differences. If the measurement properties of an instrument differ between observed groups (non-invariance), it is not possible to compare the differences between the groups (Białowolski, 2016; Kankaraš, Vermunt, & Moors, 2011; Vermunt, 2014).

The importance of cross-countries comparisons is at the heart of large-scale international surveys. Instruments that assess subjective attitudes (e.g., attitudes towards migrants) and psychological traits such as perseverance, aims for the validity and comparability of survey results. Reflective latent constructs measured through self-reports, for example, are particularly affected by subtle linguistic differences in the translated questionnaires and by broader cultural differences. These may introduce variation in participants' understanding of survey questions, and therefore in the relationship between their responses and the target latent construct. Similarly, when confronted with Likert items (*Strongly Agree, Agree, Disagree, Strongly disagree*), or with subjective rating scales (*on a scale from 1 to 10*), cultural norms may mediate the response process of participants. As a result, international surveys may fall short of their objective to facilitate comparisons across countries (*Invariance analyses in large-scale studies*, 2019).

Multigroup Latent Class Analysis tests whether the number of classes is stable across the known groups and if the measurement part of the model is equivalent across these groups.

1. Heterogeneous model

The first model to measure invariance is an *unconstrained model* in which the compared groups exhibit the same number of classes but the parameters defining those classes are freely estimated across groups. This means that assumes that the only similarity between groups is the number of classes identified and allows that response patterns (conditional probabilities) and class sizes vary among groups. Although the number of classes in all groups may be the same, direct between-country comparisons are not possible in this step because the meaning of latent classes may be substantially different. A completely unrestricted multi-group latent class model is equivalent to the estimation of a separate 3-class LC model for each group (Davidov, Schmidt, & Billiet, 2011).

2. Partial homogeneity

The second model to test is the *semi-constrained model* in which equality constraints are imposed across the observed groups. The measurement part of the model (conditional probabilities) is restricted to be equal in all observed groups. For each group, the meaning of latent classes is invariant of the group and cross-group comparisons are meaningful. Yet, the size of the classes (i.e., the relative importance of each class) may still vary. Most applicable and desirable in cross-cultural studies.

To test for invariance, the unconstrained model and the semi-constrained models are compared using the likelihood ratio test (LRT) and information criteria such as AIC, BIC, aBIC (Olivera-Aguilar & Rikoon, 2018). A statistically significant LRT indicates a substantial decrease in model fit such that the semi-constrained model should be rejected. The model with the smallest AIC, BIC, aBIC value is selected as the best-fitting model.

If the semi-constrained model is rejected, this means, lower information criteria for the unconstrained model and LRT statistically significant, there is no evidence to assume measurement invariance. In this case, latent classes are characterized differently across the observed groups and differences in the prevalence of the profiles across the groups cannot be meaningfully determined.

For invariance to exist, the semi-constrained model should show a better fit to the data than the unconstrained model. Only after establishing the stability of the classes' definition across the different groups, it is possible to compare groups and evaluate the differences in class prevalence.

3. Complete homogeneity

The stricter level of invariance is where all parameters are constrained across countries, and the prevalence of latent classes are restricted to be equal across groups (i.e., the percentage of individuals assigned to different classes will be equal in all groups). This last assumption will imply that the identified groups of individuals with similar scoring patterns are identical in all the groups with identical numbers of individuals assigned to each group.

If the fully constrained model fits best, it can be concluded that there are no differences in how the known groups are represented in each profile. In contrast, if the fully constrained model is rejected but the semi-constrained model holds means that although the profiles have the same meaning in each group, there are differences in how the individuals are distributed across classes.

Meeting this last assumption ensures the highest level of cross-country comparability but may be difficult to achieve in cross-cultural studies.

When the number of observations per group is small, likelihood ratio tests have limited power; while with large groups, violations of invariance detected in such tests may be inconsequential for the substantive inferences (*Invariance analyses in large-scale studies*, 2019). The problem is compounded by the fact that in realistic settings (when violations of measurement invariance may be due to cultural or language specificities), the hypotheses are not independent, neither across items nor across groups.

Special cases

In case that the fully constrained and semi-constrained models are rejected, it can be studied if some latent classes are measurement invariant or not and/or if some items are invariant or not. That means that the assumption of measurement invariance can be relaxed for some classes and/or items. This can be done successively until one finds such a less restrictive model that does not fit the data worse than the totally unrestricted model.

If the number of classes differs between groups, then it can be tested whether the classes that are present in all groups are measurement invariant or not. This means if one group has 2 classes and another group has 3 classes a 3-class multigroup model with full measurement invariance can be tested, where the size of the third class in the first group would be zero. This strategy is recommended for a small number of groups. When a large number of groups is tested another strategy is recommended as it will take so much time in computing and compare all parameters to identify the ones that are invariant or that should be free.

The appropriate strategy for many groups is to conduct a multigroup LCA where full measurement invariance is assumed across the groups and that the number of classes does not differ across those groups. For this, the appropriate number of classes should be identified for each group and test if just one class is different between them, if this is rejected an extra class should be added to identify if there are two different classes among them. If the double of classes is found as the best fit means that none of the classes is measurement invariant, because different classes by country are needed. This strategy has the advantage of having a higher power to detect small classes that exist in several groups but that would not be detected in country-specific analysis because their size within a group might be too small.

Chapter 3

Methods

In this section, how the procedure will be conducted is explained, focusing on the steps used to produce the expected evidence. The methodological features to perform a Latent Class Analysis are important to be clarified firstly to avoid confusion when the process is performed. Exploratory and Confirmatory approaches are explained along with how the invariance will be studied. How the models will be evaluated is described in detail including the factors and criteria to be considered. Due to the amount of information that is necessary to analyze an analytical strategy is fundamental to be able to obtain in few steps the desired results. Data to be used is also identified here, not only the items or variables that will be analyzed but the countries to be included as well.

3.1 Methodological features

There are two different approaches to conduct a Latent Class Analysis, an exploratory and confirmatory approach. Both methodologies are valid for this analysis, instead their main difference resides in the hypothesis that wants to be tested.

When researchers do not have specific hypotheses, but the goal is to identify how many classes are necessary to fit the data, an exploratory latent class analysis can be performed. In this case, several latent class models with an increasing number of classes should be computed. The best-fitting will be selected, this can be identified by which increasing the number of latent classes would not result in a model that fits the data significantly better than the previous model. Information criteria such as AIC, BIC and aBIC can be used to determine the best fit, the best model will have the lowest values of information criteria.

The confirmatory approach starts with specific hypotheses about the latent structure; the researcher can test if there is a defined number of classes that explains the associations between the observed variables and specific relations in the items for each class or across classes. Based on the conditional response probabilities and class sizes computed by the software, the expected frequencies can be estimated. These frequencies can be compared with the observed frequencies with a statistical test such as Pearson test or the LRT. If the test statistics show that the observed and expected frequencies do not differ significantly, the model is appropriate to explain the associations of the observed variables. The expected frequency of each possible response pattern should be at least 1 or even 5 to make sure that both statistics follow a χ^2 distribution and that p-values can be used for a valid decision. In case of sparse tables, bootstrapping goodness of fit is highly recommended (*Invariance analyses in large-scale studies*, 2019).

As mentioned before, studying measurement invariance is necessary to determine whether the number and nature of the latent profiles are the same across the different observed groups (Olivera-Aguilar & Rikoon, 2018). For this, multigroup LCA models are computed, and the relative fit of the unconstrained and semi-constrained models are compared using the LRT, AIC, BIC, and aBIC measures, Entropy and LL reduction is evaluated as well. Additionally, is needed to review any kind of response bias, the most common refers to “a systematic tendency to respond to a range of questionnaire items on some basis other than the specific item content” for example, extreme responses for agree/disagree (Kankaraš, Vermunt, & Moors, 2011).

3.2 Model evaluation

To estimate and identifying a good LCA model, several factors should be considered after estimating the conditional item response probability. For the exploratory approach, the selection of the optimal number of classes, examining the latent class classification, labeling the latent classes, and predicting the latent class membership should be performed and evaluated (Wang & Wang, 2020). For the confirmatory approach and the measurement invariance study, it is required to evaluate that the model fits properly or better than others according to multiple criteria.

Number of classes

Determining the number of latent classes is the most important part of a exploratory Latent Class Analysis. This cannot be estimated directly from the data. To determine the optimal number of classes, a series of LCA models with an increasing number of latent classes should be fitted. The optimal number of classes will be obtained based on the comparison of the k-class model with the (k-1) class model iteratively.

It is important to consider other aspects before deciding the final number of classes, it is recommended to follow a series of steps to identify the model that best fits the underlined classes.

- a) Compare subsequent models by model fit indices.
- b) Evaluate the quality of latent class membership.
- c) Confirm that the size of the latent classes is reasonable.
- d) Identify that the final classes are interpretable based on a theoretical grounding.

Model fit

In mixture models, multiple model fit statistics can be used to compare models (Wang & Wang, 2020). Information criterion indices, such as AIC, consistent AIC, BIC, aBIC, Lo-Mendell-Rubin likelihood ratio (LMR LR) test, adjusted LMR LR test and bootstrap likelihood ratio test (BLRT) are described in the following section.

Akaike's Information Criterion

Akaike's Information Criteria called AIC (Akaike, 1973, 1983), is one of the more important indicators to evaluate models performance, where $length(M)$ in (3.1) corresponds to the length of parameter vector of the model M . AIC penalizes the log-likelihood, generating a balance between a good fit (high value of log-likelihood) and complexity (simple models are preferable).

$$AIC(M) = -2 \log - likelihood_{max}(M) + 2 length(M) \quad (3.1)$$

AIC prefers a model with few parameters, but the fit of the model is good as well. Numerical results have shown that AIC tends to overfit, it tends to pick models with more parameters than strictly necessary. It can be proven that this effect tends to vary in one parameter more than necessary. The corrected version of AIC can be expressed as in (3.2).

$$AIC_c f(\theta) = AIC f(., \theta) + \frac{2 length(\theta)(length(\theta) + 1)}{n - length(\theta) - 1} \quad (3.2)$$

Bayesian information criterion

Based on the probability given the data it is possible to find the best model. This idea is based on Bayesian framework, involving prior probabilities on the candidate models along with prior densities on all parameters in the models (Schwarz, 1978). In (3.3) n is the sample size and $length(\theta)$ the number of parameters.

$$BIC f(., \theta) = -2 \log L(\hat{\theta}) + \log(n) length(\theta) \quad (3.3)$$

Compared to AIC, BIC includes a more severe penalty for complexity. Smaller values of information criterion indices indicate a better model fit.

Log-likelihood ratio test

The LR test based on the model χ^2 statistic is not appropriate in this case, this is because the contingency table usually has many zero cells, for this, the model χ^2 distribution is not correct. In addition, the model with $(k-1)$ -classes is a special case of the k -classes model where the one latent class probability is set to zero, and the difference of the log-likelihood between these two models does not follow a χ^2 distribution.

Lo, Mendell, and Rubin developed the LMR LR test (2001), which is not based on χ^2 distribution but on a correctly derived distribution (Wang & Wang, 2020). A significant P-value ($p < 0.05$) of the LMR LR when comparing model fit in a k -classes and $(k-1)$ -class model indicates a significant improvement in model fit in the k -class model compared to the $(k-1)$ -classes model. Then, if the test is statistically insignificant ($P \geq 0.05$) when comparing the $(k+1)$ -class model with the k -class model, this means that there is no more significant improvement in model fit when including a new class, thus cannot reject the k -class model. Consequently, the optimal number of classes will be k .

LMR LR test may inflate Type I error when the sample size is small, for this adjusted LMR LR was proposed by adjusting the number of degrees of freedom and sample size. These two tests can perform identically.

An alternative LR test based on non- χ^2 distribution is the BLRT, Bootstrap log-likelihood ratio test where parametric bootstrapping was used to generate a set of bootstrap samples using the parameters estimates from the $(k-1)$ -class model, and each of the bootstrap samples is

analyzed for both k -class and $(k-1)$ -class models. A distribution of the log-likelihood differences between the k -class and $(k-1)$ -class model from all the bootstrap samples is constructed. The BLRT is applied following this empirical distribution of the log-likelihood differences. The P-values are interpreted in the same way as the LMR LR test.

Quality of latent class membership classification

Once the optimal number of classes is identified, the cases or individuals are classified into latent classes. The probability for an individual to be assigned to a specific latent class is measured by posterior class-membership probability given the individual's response pattern on the observed categorical indicators/items. The latent class memberships of individuals are not definitely determined but based on their highest posterior class-membership probabilities (Wang & Wang, 2020).

If the posterior probability of an individual is close to 1.0, then the class misclassification or uncertainty is small. The probability for correct class-classification for an individual is the highest probability to be in a class, and the probability of misclassification is the sum of the probability to be classified in the rest of the classes. Posterior probabilities for a specific class of 1.0 are unlikely, consequently zero for the rest of the classes. A rule of thumb for acceptable class classification is 0.70 or greater (Nagin, 2005).

For assessing the quality of class membership classification another criterion is Entropy, with values that range from 0 to 1 where smaller values indicate a better classification as in (3.4) where P_{ik} is the posterior probability for the i th individual to be in class k .

$$EN(k) = - \sum_{i=1}^N \sum_{k=1}^K P_{ik} \ln P_{ik} \quad (3.4)$$

Relative entropy

The relative entropy that is defined by (Kamakura & Wedel, 2000) as in (3.5) for a k -class model with a sample size of N . This rescaled version of entropy ranges from 0 to 1 and a value closer to 1.0 indicates better classification. A good classification can be defined as some researchers suggest with an entropy of 0.8 or higher, 0.6 is medium and 0.4 is low relative entropy.

$$REN(k) = 1 - \frac{EN(k)}{N \ln(K)} \quad (3.5)$$

When defining the final latent classes, it is important to check the size of each class, the percentage of individuals in each class that represents the prevalence of the corresponding subpopulation in the target population. To have a meaningful class classification, the sizes should not be too small (Wang & Wang, 2020). Latent classes must be theoretically meaningful and interpretable. The researcher needs to define and name the classes based on the patterns of item-response probabilities in that class. For this, the classes identified should make sense and if any class is not theoretically interpretable, the model will not be useful regardless of model fit.

After the number of latent classes is defined, the class classification should be checked and interpreted. Class counts are estimated based on the posterior class membership probabilities for each individual to be partially a member of all the classes. Another type of latent class

count is estimated based on the most likely latent class membership, this means that each individual is assigned to the most likely class. If these two types of counts differ substantially indicates that the class membership misclassification is large. With a perfect classification (entropy = 1) the two counts would be identical.

Avoid local maxima

A well-known problem of any mixture modeling is that model estimation may not converge on the global maximum of the likelihood, but local maxima, providing incorrect parameter estimates (Goodman, 1974; B. Muthén & Shedden, 1999). The solution is to estimate the model with different sets of random values to ensure the best likelihood (McLachlan & Peel, 2000; L. K. Muthén & Muthén, 2012).

The software used automatically generates 10 random sets of starting values in the initial stage for all model parameters, and then maximum likelihood optimization is carried out for 10 iterations using each of the 10 random set starting values, and finally 2 starting values for the final stage optimizations. When more than 2 classes are specified, it requests a larger number of random sets of starting values to avoid local maxima of the likelihood.

3.3 Analytical strategy

To perform the analysis for this research the strategy to be used is based on performing firstly independent analysis for each selected country, this way it will be possible to identify the country-specific number of classes as a starting point.

Based on the optimal number of classes for each country, an independent analysis of the latent class will be performed to evaluate class membership, sizes and the interpretation of these classes for each country.

Once all the classes are identified for each country, all these classes will be compared between countries to identify similar and unique classes across countries. With this, a final number of optimal classes will be chosen for the total sample (including all the countries) to absorb all the possible different classes found in all the countries independently. If similar classes were found in more than one country this will count as one in the global model, if unique classes are identified this will count as a country-specific class. This country-specific class could be different across countries for that reason it will count as just one class. If one country or more has two or more country-specific classes, two or more classes will be included in the global model.

With the number of classes decided for the global model an exploratory latent class analysis is performed where a number of classes will be similar across countries and the remaining classes will differ across countries. The results from this model will be used to evaluate the model fit and the results will be compared between including/excluding one class.

The model with a good fit and the best interpretability will be chosen to evaluate the levels of invariance. First, the heterogeneous model will be analyzed, where all the parameters are freely estimated across countries. This will be similar to trying different models for each country. This model is going to be compared to more restricted models. Next, some parameters will be fixed to test if partial invariance can be achieved, with this it will be tested if the class patterns are comparable across countries. As a third step, the most restricted model is computed, where not only the patterns are assumed to be equal across countries but also the class membership.

Finally, confirmatory latent class analysis will be evaluated with some fixed conditional

probabilities based on hypotheses obtained from the global and most invariant model identified.

The confirmatory approach will be evaluated using the same sample used in the exploratory approach, due to the complex nature of the data it is not possible to split the data into training and test data using the sampling weights. This approach is mainly performed to specify obvious relations that could be observed within and between the identified patterns.

In summary, this strategy will allow for identifying if there are some classes and how many of them can be compared across countries and how we can interpret them.

1. Independent exploratory LCA to identify the country-specific number of classes.
2. Independent analysis to identify country-specific latent class membership, sizes and interpretability.
3. Identify similar classes across countries.
4. Exploratory LCA with all groups to check if the total number of classes remains.
5. Evaluate different levels of measurement invariance for the number of comparable classes across countries.
6. Establish a confirmatory analysis with the conditional probabilities and the total number of classes identified in the general exploratory LCA model.

The analytical strategy will be performed using both software, R and Mplus. R has been extensively used when manipulating large-scale assessment data because of its flexibility to work with multiple datasets at the same time and combine them maintaining all their original features. Mplus is the specialized software to perform Mixed models, especially with complex samples. This software allows to include weights, clustering, and stratification variables. This is the core for modeling large scale assessment data. As Mplus can be used by creating automatized code in R, that allows it to extract the output of the complex procedure performed and utilize R features to summarize and report the results. Most of the important code used for the different tasks are summarized in appendix A.3.

To avoid local maxima and obtain trustworthy estimations the number of random sets of starting values for initial stage optimization was set to 100, the number of random sets of starting values for final stage optimization to 25 (a quarter of the number of initial starting values) and the maximum number of iterations in optimization to 5.

To be completely sure that the model has reached the global maximum value of likelihood, Mplus performs the model by running the analysis multiple times and indicating if the global maxima is reached. When this is not the case it is not possible to compare the results, this is clearly stated in the following chapter.

3.4 Data

In this section, the different scales used in the research are explained. Firstly, an explanation of the variables that conform to every scale used in the analysis are described. Secondly, a summary and description of the data selected for the analysis, characteristics and size are given.

ICCS 2016 indicator *Students' endorsement of equal rights and opportunities* measure two constructs, attitudes towards gender equality and attitudes towards equal rights for all ethnic/racial groups. Accordingly to the technical report (Wolfram Schulz, Carsten, Losito, & Fraillon, 2018), a two-dimensional model in a confirmatory latent class analysis (CFA) using the items of these indicators showed a good fit after controlling for the common residual variance between the negatively worded statements on gender equality. These two latent dimensions are

highly correlated (0.63) and the measurement invariance was within acceptable ranges, this means that a certain degree of measurement invariance across countries was achieved when considering a variable-centered approach.

The scales *Students' endorsement of gender equality* (S_GENEQL) and *Students' endorsement of equal rights for all ethnic/racial groups* (S_ETHRGHT) created by the consortium using the individual items with Confirmatory Factor Analysis, consist of values ranging from 16.32 to 63.94 and 19.33 to 66.36 respectively for the 14 European countries. The distribution of these indicators can be observed in figure 3.1. Here, it is possible to identify that most countries' average values are similar to the European weighted average on both scales (dotted line), nonetheless few countries performed below this value, such as Bulgaria, Estonia, Latvia and The Netherlands in both scales. Lithuania performs below the average on the gender equality scale but higher on the ethnic/racial groups equality indicator. On the other hand, Norway and Sweden obtained indicators considerably higher than the European weighted average.

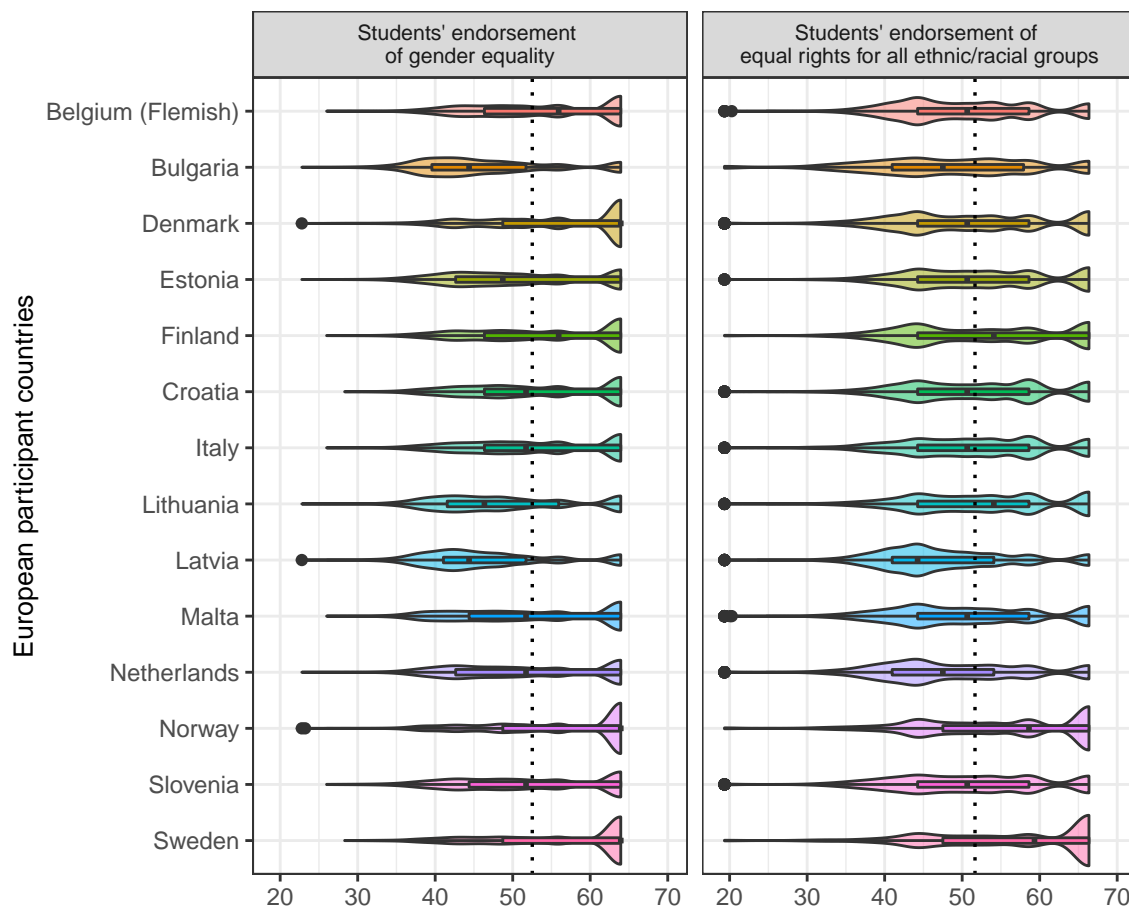


Figure 3.1: Distribution of derived scales Students' endorsement of equal rights and opportunities ICCS 2016

As mentioned in the previous chapter, these scales provide valuable information on a country's average level. For the purpose of this thesis, to identify subpopulations within each country, using each item that composes these scales separately will provide more information about the patterns that can be observed. For this, each item that produces both scales are described below.

The first scale *Students' endorsement of gender equality* is composed of seven items present also in ICCS 2009, asking about the roles of women and men in society (table A.1). Students were asked to indicate their level of agreement (four levels ranging from "strongly agree" to "strongly disagree") with each statement. The first three items consult about the level of agreement with statements related to participation in government, equal right for work and equal pay, the items are worded as "*Men and women should have equal opportunities to take part in government*", "*Men and women should have the same rights in every way*" and "*Men and women should get equal pay when they are doing the same jobs*".

Next, three items which are negatively worded consult for the level of agreement with statements related to women politics participation and two competitive items in favor towards men, "*Women should stay out of politics*", "*When not many jobs available, men should have more right to a job than women*" and "*Men are better qualified to be political leaders than women*". These last items were inversely coded, this means that when an individual responded "Agree" to any of these items, this response was coded as "Disagree" and then the question should be interpreted inversely. In the results chapter, these items will appear with a "(r)" added at the end of the label to easily identify them. Higher values of this scale reflect stronger agreement with the notion of gender equality or stronger disagreement with negative views of gender equality.

The other set of questions for the scale *Students' endorsement of equal rights for all ethnic/racial groups* are focused on the rights and responsibilities of all different ethnic/racial groups in society. Same as before students indicate their level of agreement in the same range (four levels ranging from "strongly agree" to "strongly disagree"). This scale indicates with higher scores a greater degree of agreement with the idea that ethnic and racial groups should have the same rights as other citizens in society.

In this scale, also from ICCS 2009, the first two items are focused on evaluating the attitude towards equality in education and work with "*All ethnic and racial groups should have equal chance to get a good education*" and "*All ethnic and racial groups should have an equal chance to get good jobs*". The third item is related to respect with "*Schools should teach students to respect members of all ethnic and racial groups*" and last two items are focused on political equality and social responsibilities with "*Members of all ethnic and racial groups should be encouraged to run in elections for political office*" and "*Members of all ethnic and racial groups should have same rights and responsibilities*".

The original response categories for these items are based on four points agree/disagree scale, starting by the lower level of agreement to the strongest, "Strongly disagree" (1), "Disagree" (2), "Agree" (3), and "Strongly agree" (4). For this analysis and to be able to reduce the complexity in the interpretation, these categories were recoded into two levels, "Disagree" (1-2) and "Agree" (3-4).

Multiple countries participate in the ICCS 2016 study, from different continents Europe, Asia, Latin America and the Caribbean mainly (detailed sample size of the participating countries can be found in Table A.2 in the appendix).

Only European countries were selected for this research, the decision was mainly based on choosing countries with different backgrounds, where no characteristics such as language, geographical location, economic status, or others could influence unwanted factors that could impact the results.

Fourteen countries were chosen, from nordic, western, central, eastern, and southern Europe. Countries' sample sizes differ, the country with the highest student sample is Norway (6271), followed by Denmark (6254), and the countries with the lowest sample size are Netherlands

(2812) and Slovenia (2844). Regardless the different sample sizes, senate weights are used in the analysis, the advantage of this type of weights is that balance the number of observations in a way that the total number of observations for each country weights the same overall. This will prevent bigger samples from influencing the global results, that way all countries participate at the same level.

Some considerations to be aware of the samples according to the ICCS 2016 User Guide for the International Database (Köhler, Weber, Brese, Schulz, & Carsten, 2018) are:

- Malta assessed Grade 9 students, because the average age of Grade 8 students in Malta is below 13.5 years old.
- Norway (Grade 9) deviated from the internationally defined population and surveyed the adjacent upper grade.
- Exclusion rates pertaining to the student population were greater than five percent in Estonia, Latvia, Norway (Grade 9), Sweden and North Rhine-Westphalia (Germany). The ICCS 2016 research team deemed this level of exclusion a significant reduction of the target population coverage, and researchers need to keep this caveat in mind when interpreting results.

The need of using a complex sample design forces the analysis to be even more complex as if there were no weights involved in the sample. Luckily, Mplus provides a set of tools that allows performing a latent class analysis and multigroup LCA considering not also the complexity of the sample design, but the inclusion of the student senate weights in the estimation as well.

Generally, MPLUS software can perform most of the analysis needed, still some tests such as the bootstrap likelihood ratio difference test for comparing models differing in the number of classes is not possible to calculate when using sampling weights. This is a disadvantage when working with representative samples. Nonetheless the rest of the analysis could be performed.

Chapter 4

Results

In this chapter, the relevant findings from every step performed are analyzed following the analytical strategy indicated in the previous chapter. First, latent class analysis for each country included in the sample selected is performed in order to identify the optimal number of classes for each country separately. Secondly, a global analysis is performed to identify how many latent classes are identified including all different countries, with this information it is possible to establish the number of classes that can be compared across countries. Finally, a multigroup latent class analysis is performed considering all the previous information. The multigroup analysis is constructed in multiple steps, the most restricted model until the less restricted model is evaluated. As a final step, a confirmatory latent class analysis is performed using some theoretical hypotheses that were defined based on the previous results.

This procedure is performed for the two scales that were used to create the Students' endorsement of equal rights and opportunities indicators separately.

For every analysis in this chapter, the model fit statistics table includes all the statistics that are retrieved by MPLUS software. Here is a brief description of the meaning behind every column that will be shown. The first table with the model fit statistics for all different models indicates first the number of classes used in the model, the total number of parameters estimated, the final and best Log-likelihood, the values for information criteria AIC, BIC, aBIC. The entropy indicated in each table corresponds to the relative entropy, where a perfect classification is 1, the table also indicates the log likelihood reduction (LL Reduction) from adding one class into the model. Two tests for model fit are indicated as well, the value of the statistic and the p-value associated with the Vuong-Lo-Mendell-Rubin likelihood ratio test (VLMR) and Lo-Mendell-Rubin adjusted LRT Test (LMR).

Conditional response probabilities plots are included as a summary for every independent model, the x-axis includes all the items included in the model, the y axis corresponds to the values of the probability to agree with that item, these values range from 0 to 1, where values close to 1 indicate that is highly likely to agree with that item, in contrast, values close to 0 indicate unlikely to agree to that item. On the other hand, values around 0.5 indicate randomness in the response, for this reason, is not possible to indicate a clear tendency to agree or to disagree. The different classes identified in the plots are colored differently but the colors remain the same when the response pattern is similar across countries. When a new class was identified a new color was used. The sample size for each class appears at the end of the x-axis colored with the same color as the class.

4.1 Students' endorsement of gender equality scale

As mentioned in the previous chapter, this scale is composed of 6 items, in the following tables and plots, these items were ordered from positive to negative items for an easier interpretation of the results. This ordering consider first all the items that were positive worded in the instrument *Men and women should have equal opportunities to take part in government* (GND1), *Men and women should have the same rights in every way* (GND2) and *Men and women should get equal pay when they are doing the same jobs* (GND5), followed by the three other items that are negatively worded *Women should stay out of politics (r)* (GND3), *Not many jobs available, men should have more right to a job than women (r)* (GND4), *Men are better qualified to be political leaders than women (r)* (GND6). As mentioned before all these variables were recoded in two categories, Agree and Disagree. All 14 countries were analyzed independently and then pooled in the same dataset.

4.1.1 Analysis by country

Multiple latent class models with 1 to 6-classes¹ were performed in each country in order to evaluate the model fit of each one of them. The results are summarized in table 4.1. In most European countries, the best model fit based on the different criteria indicated previously are by including 3 or 4 latent classes.

For Belgium, Croatia, Denmark, Latvia, and The Netherlands there is no significant improvement in the log-likelihood from two to three latent classes. In this sense, BIC and aBIC simultaneously have the lowest values in the 3-class model.

On the other hand, in Bulgaria, Estonia, Malta, Slovenia, and Sweden according to the statistical tests, BIC, and aBIC criteria, the best model is a 4-class model. In Finland, Italy and Lithuania models, the BIC, aBIC differ from the statistical test indicating a better fit for the 3-class model.

Norway is the only country from the sample where the best model fit is the one with 5 latent classes according to the statistical tests and BIC and aBIC.

It is a common tendency in all the evaluated countries that the AIC value is lower in the models with one more class than indicated by the statistical tests and BIC and aBIC statistics. That is consistent with the indication that this criterion tends to overfit the data.

Values of Entropy are higher when the tests are significant but consistent with a better fit of the data, the lower entropy found in the 4-class model is in Latvia (73.7%) and the highest value in Norway (96%). The log-likelihood reduction is consistent in all countries, where having more than 3 latent classes reduces the log-likelihood around 0.2% and 1%.

All models selected accomplish at least one or more of the criteria established for a good fit. The bivariate residuals were also analyzed, and all countries have residuals around the range of acceptable $[-2 ; 2]$ as shown in the figure 4.1. There is just one value is outside the ranges in Malta with a 4-class model.

¹Summary with all models can be found in Appendix, table A.3.

Table 4.1: Best model, fit statistics individual country model Students' endorsement of gender equality

Country	N Latent Classes	Param	Log- Likelihood	AIC	BIC	aBIC	Entropy	LL Reduc- tion	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
Belgium (Flemish)	3	20	-3812	7664	<i>7784</i>	<i>7721</i>	<i>88.2%</i>	0.8%	<i>62</i>	<i>0.323</i>	<i>61</i>	<i>0.327</i>
Bulgaria	4	27	-7523	15100	<i>15261</i>	<i>15175</i>	<i>75.2%</i>	0.4%	<i>60</i>	<i>0.077</i>	<i>59</i>	<i>0.08</i>
Croatia	3	20	-5368	10777	<i>10902</i>	<i>10838</i>	<i>87.6%</i>	0.9%	<i>98</i>	<i>0.085</i>	<i>96</i>	<i>0.088</i>
Denmark	4	27	-5914	11883	<i>12063</i>	<i>11977</i>	<i>91.2%</i>	0.6%	<i>74</i>	<i>0.174</i>	<i>73</i>	<i>0.178</i>
Estonia	4	27	-4974	<i>10001</i>	<i>10162</i>	<i>10076</i>	<i>77.7%</i>	0.6%	<i>57</i>	<i>0.183</i>	<i>56</i>	<i>0.189</i>
Finland	3	20	-3477	6993	<i>7114</i>	<i>7051</i>	<i>90.7%</i>	1.2%	86	0.037	84	0.039
Italy	3	20	-4830	9701	<i>9824</i>	<i>9760</i>	<i>84.6%</i>	1.5%	145	0	143	0
Latvia	3	20	-7993	16027	<i>16148</i>	<i>16085</i>	<i>72.4%</i>	0.8%	<i>121</i>	<i>0.052</i>	<i>119</i>	<i>0.055</i>
Lithuania	3	20	-7447	14934	<i>15058</i>	<i>14994</i>	<i>82.5%</i>	1.6%	248	0	244	0
Malta	4	27	-6204	12462	<i>12629</i>	<i>12543</i>	<i>87.2%</i>	0.5%	<i>65</i>	<i>0.235</i>	<i>64</i>	<i>0.238</i>
Netherlands	3	20	-4759	9557	<i>9676</i>	<i>9612</i>	<i>87.0%</i>	1.5%	<i>140</i>	<i>0.074</i>	<i>138</i>	<i>0.076</i>
Norway	5	34	-6035	12137	<i>12365</i>	<i>12257</i>	<i>93.2%</i>	0.5%	<i>66</i>	<i>0.281</i>	<i>65</i>	<i>0.286</i>
Slovenia	4	27	-4280	8614	<i>8774</i>	<i>8689</i>	<i>87.5%</i>	0.7%	<i>56</i>	<i>0.158</i>	<i>55</i>	<i>0.163</i>
Sweden	4	27	-3049	6152	<i>6316</i>	<i>6230</i>	<i>89.2%</i>	1.0%	<i>62</i>	<i>0.398</i>	<i>61</i>	<i>0.402</i>

Note:

Best model based on the lowest value of BIC

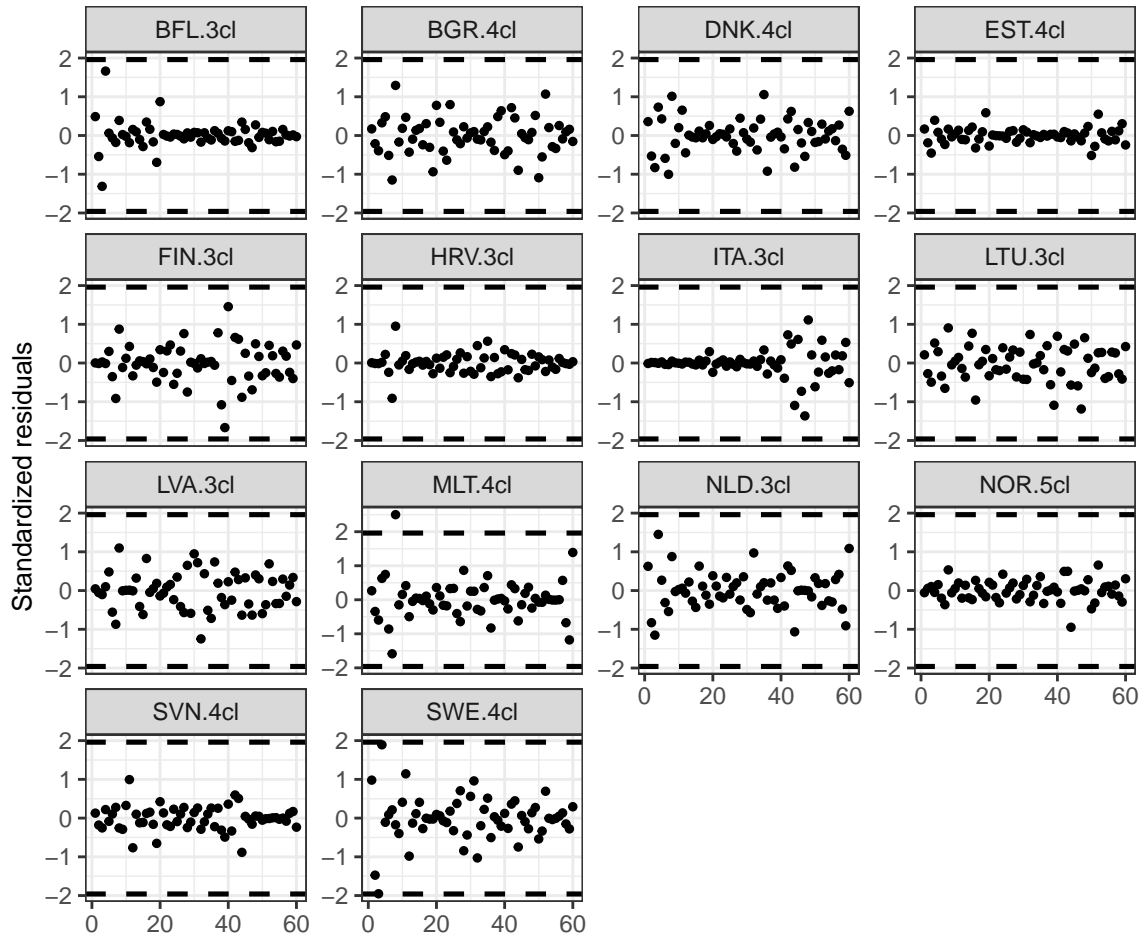


Figure 4.1: Bivariate standardized residuals individual country models for Students' endorsement of gender equality

In figure 4.2, the classes of each independent model can be identified by looking at the conditional probabilities. In the figure, the conditional probabilities to agree to each item are shown and plotted for each class estimated in each country. From all the models, two classes that are similar across countries are identified in the figure, Fully egalitarian and Competition-driven sexism, green and purple line respectively.

A brief explanation of each class is described below.

- **Fully egalitarian:** Most likely to agree to all items (green line). This class can be observed in all countries. Conditional probabilities greater than 0.75 to agree, class sizes around 60% (Bulgaria) and 90% (Denmark).
- **Competition-driven sexism:** Most likely to disagree with gender competitive items in favor of women (purple line). This class can be observed in all countries. Conditional probabilities greater than 0.75 to agree to positive views of gender equality and generally lower than 0.5 to agree to reversed negative views, class sizes around 3.6% (Denmark) and 22.5% (Bulgaria).
- **Non-egalitarian:** Not likely to agree to any item (orange line). This class can be observed in four countries. Conditional probabilities lower to 0.5 to agree to any item, class sizes around 0.9% (Norway) and 2.6% (Italy).

- **Reverse competition-driven sexism:** Most likely to agree to gender competitive items in favor of women (pink line). This class can be observed in five countries. Conditional probabilities lower than 0.25 to agree to positive views of gender equality and generally greater than 0.75 to agree to reversed negative views, class sizes around 0.6% (Norway) and 1.6% (The Netherlands).
- **Political egalitarian:** Likely to agree to politically related items (light-green line). This class can be observed in five countries. Conditional probabilities are greater than 0.75 in political equality items, class sizes around 3.2% (Belgium) and 1.4% (Estonia).
- **Random response:** Not defined attitude (yellow line). This class can be observed in four countries. Conditional probabilities between 0.25 and 0.75 to agree all items, class sizes around 2.7% (Slovenia) and 16.8% (Bulgaria).

The classes described before are not present in all countries, for this reason, a global model will be tested in the pooled sample considering not only the classes that are similar across more than one country, but additional classes will be added in order to absorb the remaining different classes that the global model will identify.

In the following section, the global model will be tested using the pooled dataset.

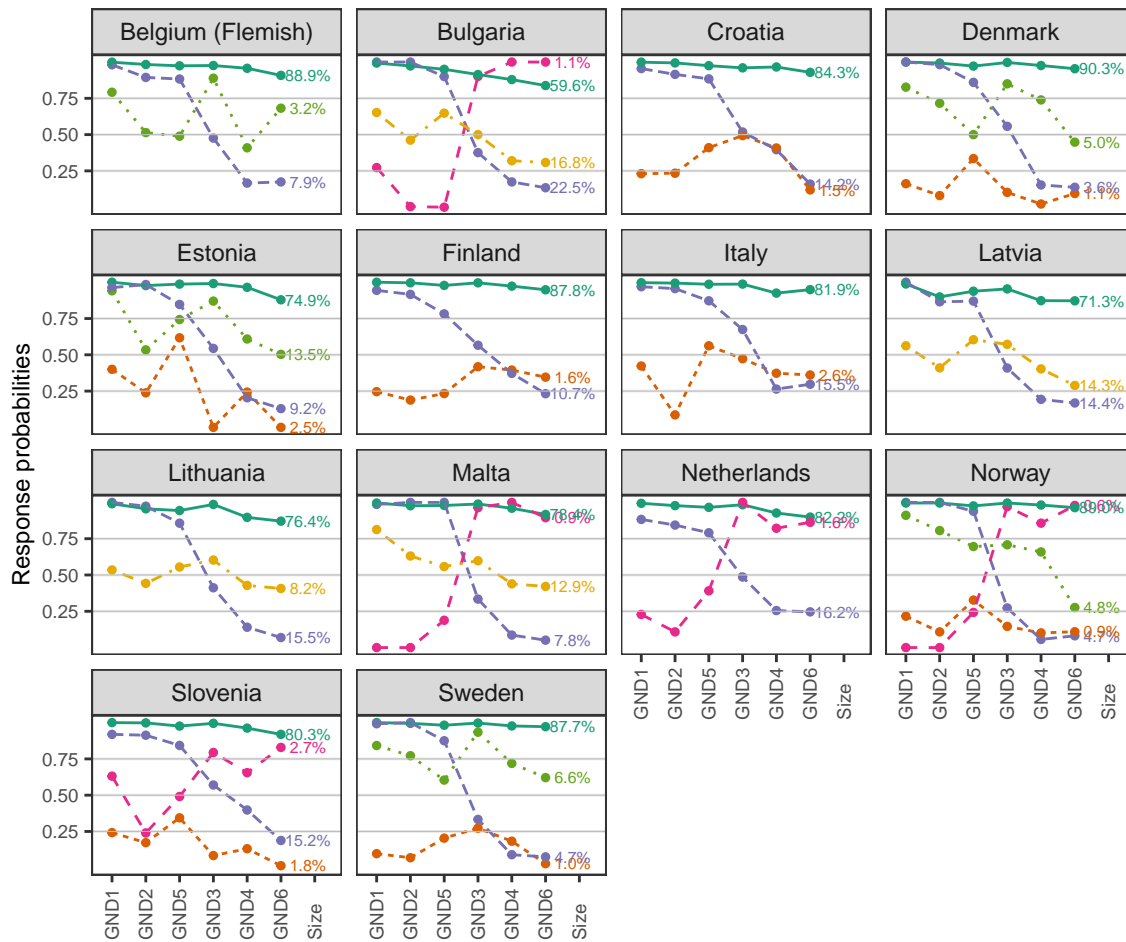


Figure 4.2: Classes for best individual country model for Students' endorsement of gender equality

4.1.2 General model

Table 4.2 shows the results of each model using the pooled sample with all the countries. Models with 1 to 7 classes were computed and the model fit statistics were summarized in the table.

The model that includes a single class has the largest AIC (192,838), BIC (192,891), and ABIC (192,872) values for the pooled sample, indicating that this model fits data worse than all other models. In addition, the P-values for the VLMR test, and LMR in the 2-class model are all < 0.0001 ; this means that both tests reject the single-class model in favor of a model with at least two latent classes. In other words, there exists heterogeneity in the target population regarding attitudes towards gender equality.

On the other hand, the LMR LR and VLMR tests for the 6-class model are not statistically significant ($P > 0.05$). That is, the two tests are in favor of at most 5 classes.

In contrast, BIC and aBIC values are all smaller in the 5-class model than those in the 6-class model; thus, consider that the models with more than 5 classes are not preferred. AIC values reach the lowest value in the 7-class model but based on the previous results this criterion will not be considered in this case due to the tendency to overfit the data.

The relative entropy given by Mplus software decreases when including more than 4 classes and increases again with the 6-class model; this would suggest that a model with at least 6 class or 4 class is preferred.

Together with the percentage of reduction in the log-likelihood value, that indicates that by adding two classes to the model the log-likelihood is reduced by 13.3%, this reduction is only increased by 1.2% if the model is a 3-class model and finally this value is reduced close to 0 if more than 5 classes are included.

Now, the preferred model must be either the 5-class or the 6-class model. Considering the residuals of each model, in figure 4.3 all values are around -1.96 and 1.96. But based on the parsimony principle a 4-class model can be considered as well if just one value of the residuals is outside the acceptable range.

Theoretically, we tend to determine that the 4-class LCA model is the preferred model. We will show later that the classes identified by the 4-class model are more interpretable and representative than the rest of the models. And in particular that two classes can be compared across countries.

Table 4.2: Model fit statistics LCA Students' endorsement of gender equality

N Latent Classes	Param	Log- Likelihood	AIC	BIC	aBIC	Entropy	LL Re- duction	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
All countries											
1	6	-96413	192838	192891	192872						
2	13	-83617	167261	167376	167334	80.5%	13.3%	25591	0	25258	0
3	20	-82592	165223	165400	165336	84.5%	1.2%	2051	0	2025	0
4	27	-82327	164708	164946	164861	82.7%	0.3%	529	0	522	0
5	34	-82163	164394	164694	164586	81.1%	0.2%	328	0	324	0
6	41	-82136	164355	164717	164586	82.5%	0.0%	53	0.246	52	0.252
7	48	-82116	164328	164752	164600	82.1%	0.0%	40	0.244	40	0.247

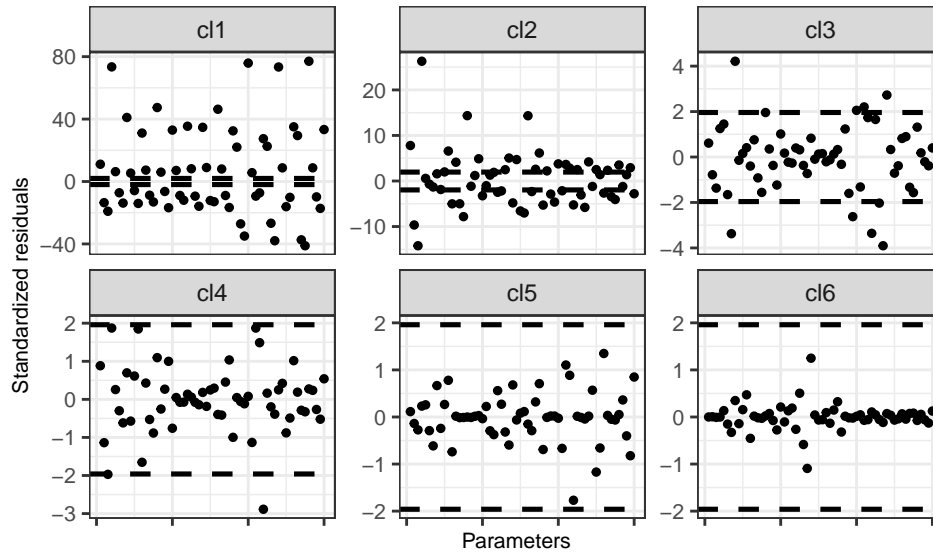


Figure 4.3: Bivariate model fit standardized residuals global model for Students' endorsement of gender equality

Three, four, five and six class models were investigated profoundly. It is difficult to choose the best model fit without a full analysis. There are some patterns that can be clearly identified in the models as can be seen in figure 4.4. Class 1 with class sizes of 81.5%, 79.2%, 77% and 78.6% in each model respectively, and estimated probabilities to agree for this latent class higher than 0.92 for all six items *Men and women should have equal opportunities to take part in government*, *Men and women should have the same rights in every way*, *Men and women should get equal pay when they are doing the same job*, *Women should stay out of politics*, *Not many jobs available, men should have more right to a job than women* and *Men are better qualified to be political leaders than women* the **Fully egalitarian** group.

The second class (Class 2 in Figure 4.4) identified in all the models, called **Competition-driven sexism**. For this class, the estimated probabilities to agree to the first 3 items are close or higher than 0.9 in all models. For the last three items, the estimated probabilities of agreement are not higher than 0.5 in all models. The class size differs in all four models, 11.3%, 11.5%, 12.8% and 8.6% in the 3, 4, 5, 6-class model respectively.

The third class (Class 3) that can be seen with a similar pattern in all the models, it is called **Non-egalitarian**, this class appears in the 4-class model. The pattern of this class is basically showing lower estimated conditional probabilities to agree with any of these statements, no greater than 0.4, except for one item *Men and women should get equal pay when they are doing the same jobs* with an estimated probability to agree no higher than 0.55. The estimated sizes for this class are 7.7%, 7% and 5.8% in each model respectively.

The fourth, fifth and sixth classes identified in the models differ in all the countries, nevertheless, one class appears to be consistent in the 5-class model, where this class called **Reverse competition-driven sexism** has opposite conditional probabilities compared to the second class identified previously.

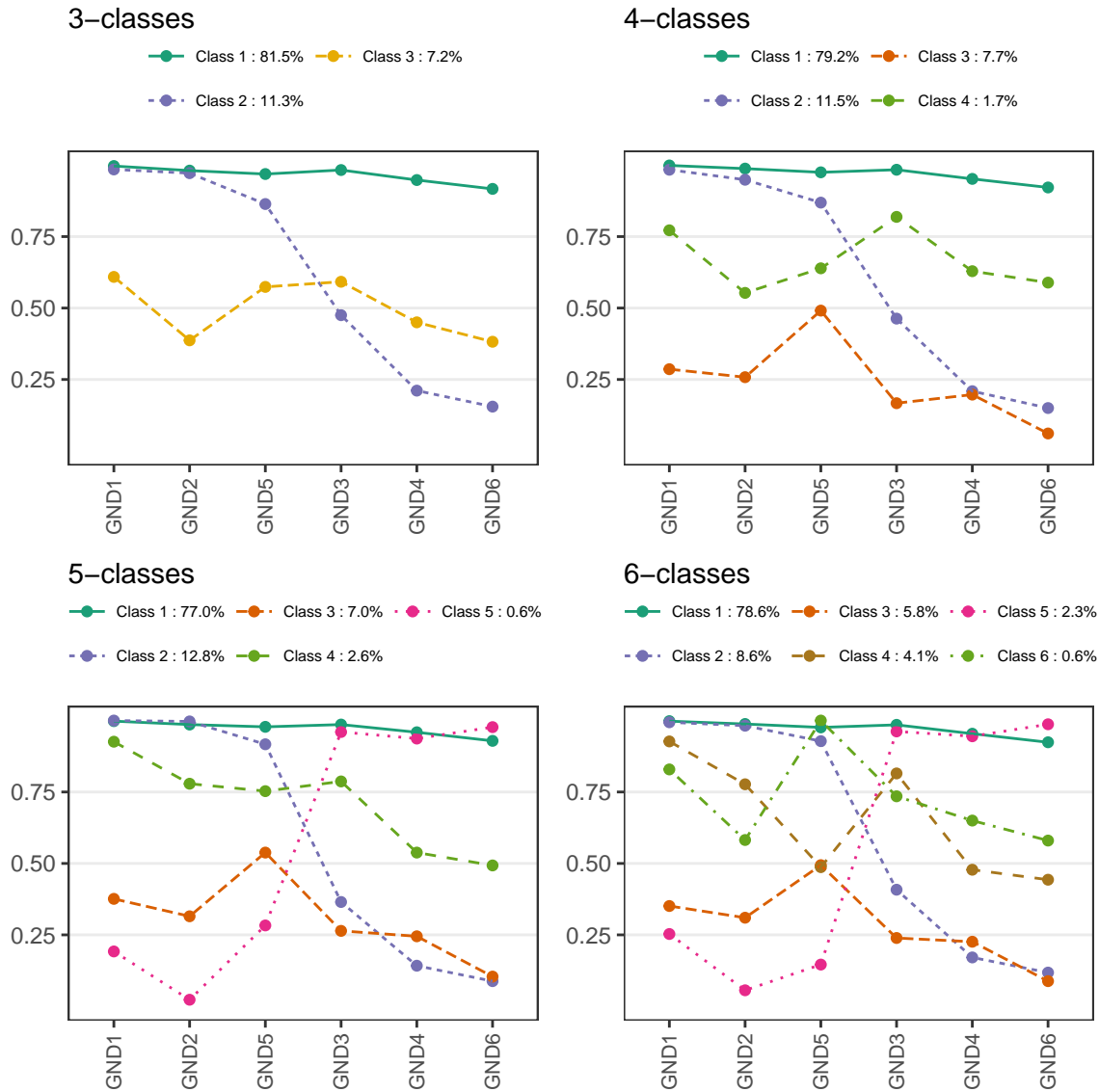


Figure 4.4: Comparative conditional probabilities to agree in 3 to 6 latent class global models for Students' endorsement of gender equality

The main two classes in the solutions with five and six classes do not strongly differ from other models, and the remaining classes are not informative at all or with a sample size very small, using this as a criterion, one can prefer a four-class solution.

In table 4.3 the conditional probabilities to agree with the fourth latent class model are shown. These values are very close to 1 for the first class, Fully egalitarian. Similar values are obtained for the positive items in the second class Competition driven-sexism, meanwhile, for third item GND3, conditional probabilities are close to 0.5, this would mean a random response, but the last two items have lower conditional probabilities close to 0.2, which would mean not likely to agree to the statements. Table 4.4 indicates the counts and proportions using the model estimated and the most likely probabilities.

Table 4.3: Probabilities to agree each item in the four-class global model for Students' endorsement of gender equality

param	Fully egalitarian	Competition- driven sexism	Non- egalitarian	Political egalitarian
GND1 - Men and women should have equal opportunities to take part in government	0.999	0.984	0.286	0.772
GND2 - Men and women should have the same rights in every way	0.988	0.949	0.258	0.553
GND5 - Men and women should get equal pay when they are doing the same jobs	0.975	0.869	0.491	0.639
GND3 - Women should stay out of politics (r)	0.984	0.463	0.167	0.819
GND4 - Not many jobs available, men should have more right to a job than women (r)	0.952	0.209	0.197	0.629
GND6 - Men are better qualified to be political leaders than women (r)	0.922	0.15	0.061	0.589

Table 4.4: Class sizes four-class global model for Students' endorsement of gender equality

Class	Model estimated		Most likely	
	Counts	Proportion	Counts	Proportion
Fully egalitarian	39924.2	79.2%	41508	82.3%
Competition- driven sexism	5782.2	11.5%	5258	10.4%
Political egalitarian	3864.7	7.7%	2969	5.9%
Non-egalitarian	859.9	1.7%	696	1.4%

4.1.3 Country comparability

To evaluate the country comparability, the classes that were found in the independent model were identified to later check how many of them could be tested for comparability using a multigroup latent class model.

The different classes identified according to each global model is summarized next, indicating in which countries the same class is present.

- Global three-class model:
 1. Fully egalitarian - ALL COUNTRIES
 2. Competition-driven sexism - ALL COUNTRIES
 3. Random response - BGR, LVA, LTU, MLT
- Global four-class model:
 1. Fully egalitarian - ALL COUNTRIES
 2. Competition-driven sexism - ALL COUNTRIES
 3. Non-egalitarian - HRV, DNK, EST, FIN, ITA, NOR, SLV, SWE

4. Political egalitarian - BFL, DNK, EST, NOR, SWE
- Global five-class model:
 1. Fully egalitarian - ALL COUNTRIES
 2. Competition-driven sexism - ALL COUNTRIES
 3. Non-egalitarian - HRV, DNK, EST, FIN, ITA, NOR, SLV, SWE
 4. Political egalitarian - BFL, DNK, EST, NOR, SWE
 5. Reverse competition-driven sexism - BGR, MLT, NLD, NOR, SLV
 - Global six-class model:
 1. Fully egalitarian - ALL COUNTRIES
 2. Competition-driven sexism - ALL COUNTRIES
 3. Non-egalitarian - HRV, DNK, EST, FIN, ITA, NOR, SLV, SWE
 4. Political egalitarian - BFL, DNK, EST, NOR, SWE
 5. Reverse competition-driven sexism - BGR, MLT, NLD, NOR, SLV
 6. Pro-women pay/job - Not defined in individual country models

With three classes, the random response class is not very interpretable. With six classes, a new no-identified class appears, which is not interpretable. With five classes, the reverse competition-driven sexism class is present in five countries but with class sizes lower than 1%, which would be not representative. With four classes, the four classes are identified across countries and two of them are present in all countries, which is the best model for comparability.

Country multigroup analysis

In table 4.5 different models with multigroup analysis are tested, first the more restricted model is evaluated, complete homogeneity. In this model, all conditional and unconditional probabilities are fixed to be equal across the groups.

Then, the partial homogeneity is tested where only the conditional probabilities are constrained to be equal across the groups, and the class sizes are estimated freely. The second approach of partial homogeneity was also tested, but not included in the table, where only the conditional probabilities for the two common classes identified are constrained across groups, and the remaining are freely estimated along with the unconditional probabilities. Finally, the complete heterogeneous model is tested, where not only the unconditional probabilities are estimated freely but all the conditional probabilities as well. In the last two models the best log-likelihood is not replicated, this means that the solution may not be trustworthy due to local maxima. These results cannot be considered valid.

Just by looking at the valid results, the partial homogeneity where all conditional probabilities are constrained to be equal across groups shows a better fit compared to the more restricted model, the complete homogeneity. With it is valid to indicate that the 4 classes identified do not share the same unconditional probabilities (class sizes) across the groups, but the conditional probabilities can be considered as equal in all groups.

Table 4.5: Multigroup model fit statistics global model with four-classes with Students' endorsement of gender equality

Nggroups	Param	Log-Likelihood	AIC	BIC	aBIC	Entropy	LL Reduction	Δ LL	Δ DF	pvalue Δ
Four-class model										
Complete homogeneity										
14	40	-215414	430907	431260	431133	94.0%	-1.49%	-3158	-348	0
Partial homogeneity										
14	79	-213195	426549	427246	426995	88.1%	-0.44%	-940	-308	0
Complete heterogeneity										
14	391	-212256	425293	428745	427502	94.0%	0.00%			

Note:

The best log-likelihood value was not replicated for the following models:

¹ 4-class Complete heterogeneity model.

Figure 4.5 indicates the values for the patterns with the conditional probabilities fixed in all countries, but also the unconditional probabilities are constrained to be equal in all groups. Here can be observed that the patterns are like the ones identified in the independent models and the global model as well. But this would force the classes sizes to be 79.2% for the Fully egalitarian, 11.5% for the Competition-driven sexism, 7.7% for the Political egalitarian and 1.7% for the Non-egalitarian for all the countries which is not optimal.

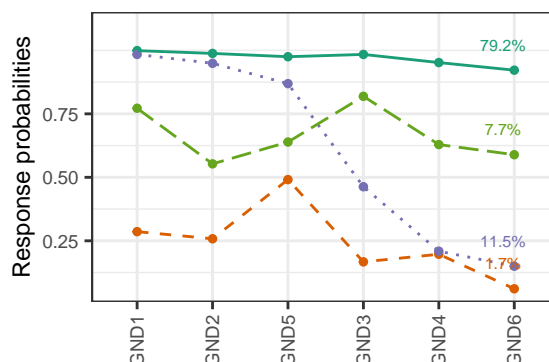


Figure 4.5: Conditional probabilities to agree in a 4-class complete homogeneous multigroup model for Students' endorsement of gender equality scale

The proportions based on a global model cannot be extrapolated to a country level, because only partial invariance was obtained those values are not applicable to each country. When considering the partial invariance results (indicated in Table 4.6) it is possible to identify that even though Fully egalitarian is the predominant class regarding size, in countries such as Bulgaria, Latvia, and Lithuania most predominant class is the Political egalitarian with around 45% of the population. Denmark, Finland, Norway and Sweden obtained the larger number of members in the Fully egalitarian class based on the model estimated counts. In contrast, Bulgaria and Lithuania has the larger number of members in the Non-egalitarian class compared to the rest of the European countries.

Table 4.6: Class sizes for partial homogeneity model for Students' endorsement of gender equality scale

Group	Fully egalitarian	Competition- driven sexism	Non- egalitarian	Political egalitarian
Belgium (Flemish)	65.5	1.1	5.3	28.0
Bulgaria	21.5	11.0	25.6	41.9
Croatia	66.0	2.2	8.6	23.2
Denmark	82.4	2.7	2.7	12.2
Estonia	50.3	5.6	7.2	36.9
Finland	80.1	2.5	5.8	11.6
Italy	63.4	2.7	8.1	25.8
Latvia	23.4	12.8	10.8	53.0
Lithuania	39.3	5.5	14.7	40.5
Malta	59.4	6.2	10.3	24.2
Netherlands	53.8	4.0	8.5	33.7
Norway	85.1	2.7	6.1	6.1
Slovenia	66.3	5.4	8.0	20.3
Sweden	83.7	2.5	4.8	8.9

4.1.4 Confirmatory Latent Class Analysis

The confirmatory model was performed by establishing some constraints based on the results obtained in the previous analysis. For the Students' endorsement of gender equality, the hypothesis tested is that students that would be classified into the Fully egalitarian class would highly agree (totally) with that *Men and women should have equal opportunities to take part in government* which is the first item in the scale. Also, they would have the same probability to agree (in the same level) to both items *Men and Women should have the same rights in every way* and *Women should stay out of politics (r)* that are the next items in the questionnaire.

The confirmatory approach also test that the students that would be classified in the second group Competition-driven sexism class would be equally likely to agree to the first item in the questionnaire *Men and women should have equal opportunities to take part in government* as the members of the first class agree with the second item and third item. And that these students would not have a clear attitude (either agree or disagree) with the item *Women should stay out of politics (r)* which means that they will tend to give a random response to this item.

This can be tested by setting the conditional probabilities for the first item in the Fully egalitarian class to 1. The probabilities to agree to the second and third item in the first class and the first item in the Competition-driven sexism class to be equal. The third hypothesis can be tested by setting the conditional probability of the third item in the Competition-driven sexism class to 0.5. The rest of the conditional probabilities are estimated freely as can be seen in table 4.7. Values for the thresholds, class sizes and the probabilities to agree to each item can be found in the appendix table A.5.

In table 4.8 the model fit statistics of this model indicate that they do not differ considerably from the exploratory approach where the log-likelihood, BIC, and aBIC have higher values, only the AIC value is better in the exploratory model.

Table 4.7: Probabilities to agree each item in a 4-class Confirmatory LCA for Students' endorsement of gender equality scale

param	Fully egalitarian	Competition- driven sexism	Non- egalitarian	Political egalitarian
GND1 - Men and women should have equal opportunities to take part in government	1	0.986	0.296	0.75
GND2 - Men and women should have the same rights in every way	0.986	0.941	0.272	0.552
GND5 - Men and women should get equal pay when they are doing the same jobs	0.975	0.864	0.495	0.638
GND3 - Women should stay out of politics (r)	0.986	0.5	0.181	0.817
GND4 - Not many jobs available, men should have more right to a job than women (r)	0.953	0.228	0.195	0.645
GND6 - Men are better qualified to be political leaders than women (r)	0.923	0.167	0.057	0.609

This result suggests that it can be plausible to test some other similar hypotheses in order to identify straightforward patterns based on the theoretical background of the items analysed that could help to interpret them even more.

Table 4.8: Model fit statistics 4-class Confirmatory LCA for Students' endorsement of gender equality scale

Type	N Latent Classes	Param	Log- Likelihood	AIC	BIC	aBIC	Entropy	LL Re- duction
All countries								
Exploratory LCA	4	27	-82327	164708	164946	164861	82.7%	
Confirmatory LCA	4	23	-82340	164726	164929	164856	82.5%	0.0%

4.2 Students' endorsement of equal rights for all ethnic/racial groups scale

This scale is composed of 5 items, in the following results, these items were ordered in the output for an easier interpretation of the results. This ordering considers first *All ethnic and racial groups should have equal chance to get a good education* (ETH1), *All ethnic and racial groups should have an equal chance to get good jobs* (ETH2), *Members of all ethnic and racial groups should have same rights and responsibilities* (ETH5), and *Schools should teach students to respect members of all ethnic and racial groups* (ETH3), followed by *Members of all ethnic and racial groups should be encouraged to run in elections* (ETH4). As mentioned before all these variables were recorded in two categories, as Agree and Disagree. All 14 countries were analyzed independently and then pooled in the same dataset.

4.2.1 Analysis by country

A latent class analysis with 1 to 6-class models was performed in each country to evaluate the model fit of each one of them². The results are summarized in table 4.9. In most European countries, the best model fit based on the different criteria indicated previously are by including 3 or 4 latent classes.

For Belgium, Bulgaria, Estonia, Italy, Lithuania, Latvia, Slovenia, The Netherlands, Norway, Slovenia and Sweden, according to the statistical tests, BIC, and aBIC criteria, the best model is a 3-class model.

On the other hand, Denmark and Malta have a better model fit in a 4-class model, consistently between statistical tests and BIC criteria.

In Croatia models, tests indicate that a 2-class model is better for their data, even though BIC indicates a 3-class model to have the lowest value.

Norway is the only country from the sample that the best model fit is the one with 5 latent classes according to the statistical tests and BIC and aBIC.

It is a common tendency in all the evaluated countries the AIC value is lower in the models with one more class than indicated by the statistical tests and BIC and aBIC. This is consistent with the indication that this criterion tends to overfit the data. Values of Entropy are higher when the tests are significant, but consistent with a better fit of the data the lowest entropy found in the 3-class model is in Belgium (60.5%) and the highest value in Sweden (90.2%). The log-likelihood reduction is consistent in all countries, where having more than 3 latent classes reduces the log-likelihood around 0.1% and 0.6%.

The bivariate residuals were also analyzed, and all countries have residuals around the range of acceptable [-2 ; 2] as shown in the figure 4.6. There are just two values outside the ranges in Italy with a 3-class model.

²Model fit statistics for each model can be found in the Appendix A.6

Table 4.9: Best model, fit statistics individual country model Students' endorsement of equal rights for all ethnic/racial groups

Country	N Latent Classes	Param	Log- Likelihood	AIC	BIC	aBIC	Entropy	LL Reduc- tion	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
Belgium (Flemish)	3	17	-4019	8072	<i>8174</i>	<i>8120</i>	60.5%	0.8%	63	0.021	62	0.023
Bulgaria	3	17	-5354	10742	<i>10843</i>	10789	78.5%	1.8%	193	0	190	0
Croatia	3	17	-4507	9047	<i>9153</i>	9099	76.1%	0.9%	<i>84</i>	<i>0.416</i>	<i>82</i>	<i>0.422</i>
Denmark	4	23	-8282	<i>16610</i>	<i>16764</i>	<i>16691</i>	72.3%	0.3%	<i>54</i>	<i>0.059</i>	<i>53</i>	<i>0.062</i>
Estonia	3	17	-3254	6543	<i>6644</i>	<i>6590</i>	69.8%	1.3%	87	0.002	86	0.002
Finland	3	17	-3391	6815	<i>6918</i>	6864	80.8%	2.9%	200	0	196	0
Italy	3	17	-4354	8742	<i>8846</i>	8792	81.1%	1.6%	146	0	143	0
Latvia	3	17	-5353	10741	<i>10844</i>	<i>10790</i>	64.0%	1.1%	124	0.001	122	0.001
Lithuania	3	17	-4194	8423	<i>8528</i>	<i>8474</i>	84.8%	0.9%	75	0.016	74	0.017
Malta	4	23	-5691	11428	<i>11570</i>	<i>11497</i>	80.3%	0.7%	78	0.032	76	0.034
Netherlands	3	17	-4729	9493	<i>9593</i>	<i>9539</i>	69.7%	1.8%	170	0	166	0
Norway	3	17	-5448	10930	<i>11044</i>	10990	88.1%	1.9%	207	0	203	0
Slovenia	3	17	-4272	8578	<i>8679</i>	8625	77.5%	1.0%	87	0.027	85	0.029
Sweden	3	17	-2306	4646	<i>4749</i>	<i>4695</i>	90.2%	3.1%	147	0.011	144	0.012

Note:

Best model based on the lowest value of BIC

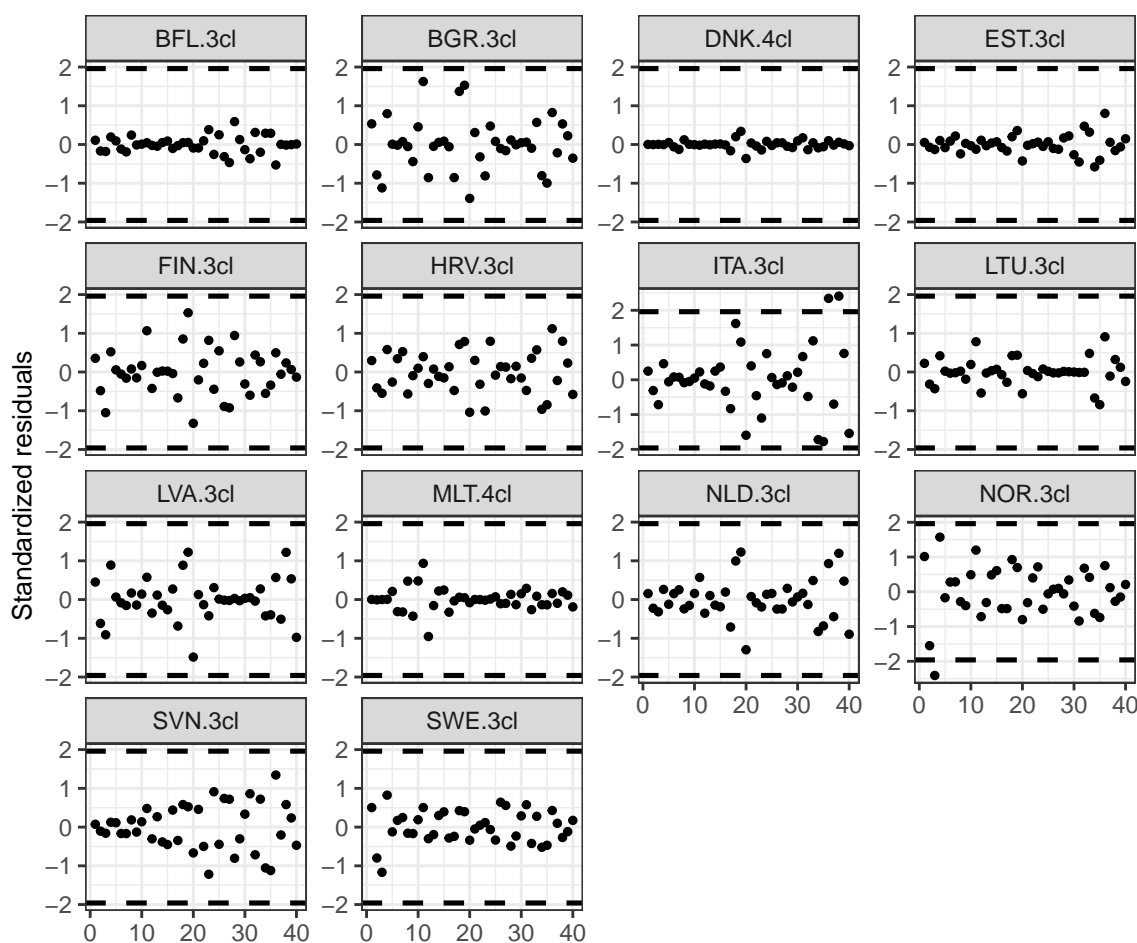


Figure 4.6: Bivariate standardized residuals for individual country models for Students' endorsement of equal rights for all ethnic/racial groups scale

In figure 4.7 the classes of each independent model can be identified. In the figure, the conditional probabilities for agreement to each item are shown and plotted for each of the classes modeled in each country. Here can be identified two classes that are similar in all the models, the green and purple lines.

- **Fully egalitarian:** Most likely to agree to all items in the scale (green line). This class can be observed in all countries. Conditional probabilities greater than 0.7 to agree, class sizes around 61.8% (Latvia) and 90% (Sweden)
- **Political non-egalitarian:** Most likely to agree to all items but a random answer to *All ethnic and racial groups should be encouraged to run in elections (ETH4)* item (orange line). This class can be observed in all countries. Conditional probabilities to agree higher than 0.5 in all items but to the political item (< 0.5), class sizes around 7.6% (Denmark) and 36% (Latvia).
- **Non-egalitarian:** Not likely to agree to any item in the scale (purple line). This class can be observed in all countries. Conditional probabilities lower than 0.5 to agree all items, class sizes around 1.4% (Lithuania) and 5.4% (Bulgaria).
- **Country specific class:** (pink line). This class can be observed in two countries.

- Employment non-egalitarian: Not likely to agree to *All ethnic and racial groups should have an equal chance to get good jobs (ETH2)* item. Class size 8.3% (Malta)
- Strong political non-egalitarian: Most likely to agree to most items in the scale but not likely to agree to *All ethnic and racial groups should be encouraged to run in elections (ETH4)* item. Class size 21.3% (Denmark).

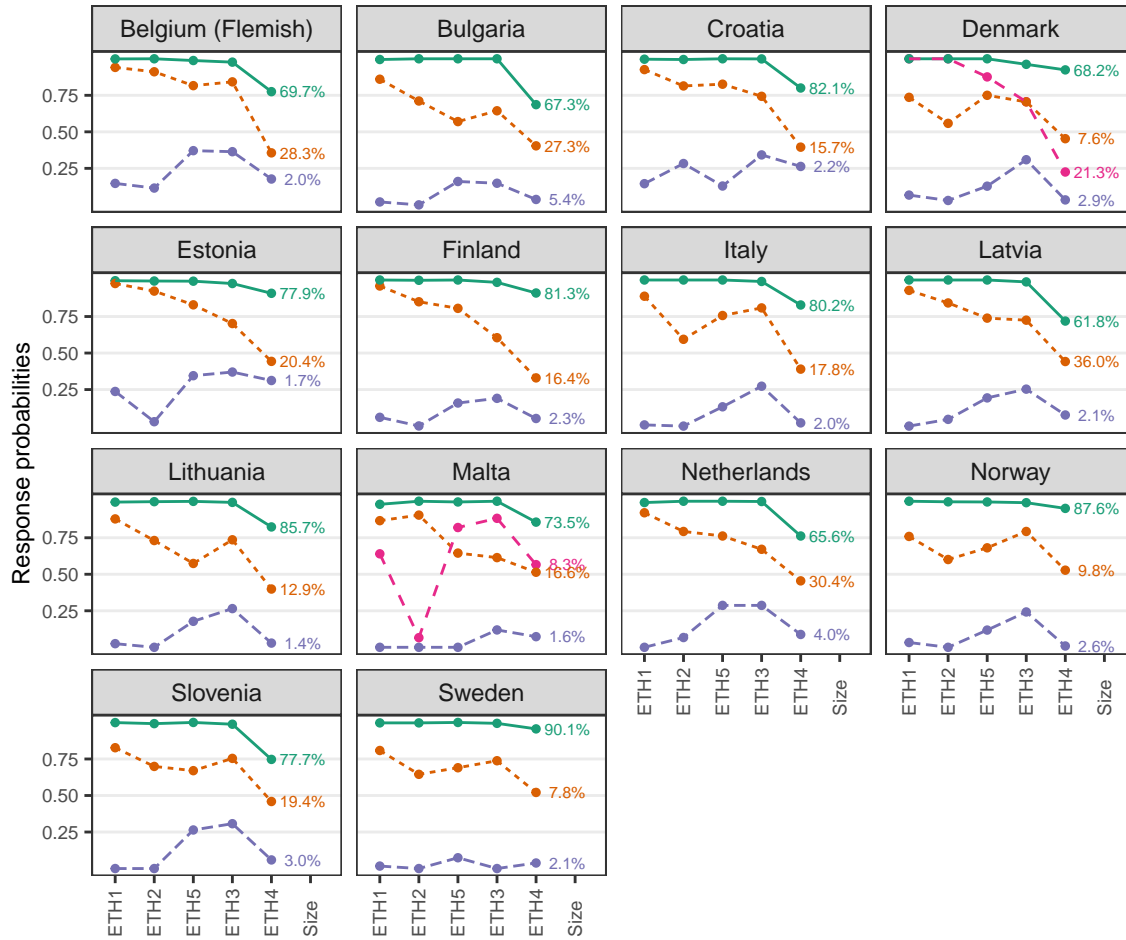


Figure 4.7: Classes for best individual country model for Students' endorsement of equal rights for all ethnic/racial groups

4.2.2 General model

The model with a single class has the largest AIC (159379), BIC (159423), and aBIC (159407) values for the European countries model in Table 4.10, indicating that this model fits data worse than all other models. In addition, the P-values of the VLMR test, and LMR in the 2-class model are all < 0.0001 ; this means that both tests reject the single-class model in favor of a model with at least two latent classes. In other words, there exists heterogeneity in the target population regarding attitudes towards gender equality.

In the 6-class model, the LMR LR and VLMR are not statistically significant ($P > 0.05$). That is, the two tests are in favor of at most 5 classes. In contrast, AIC, BIC and aBIC values are all smaller in the 5-class model than those in the 6-class model; thus, consider that the models with more than 5 classes are not preferred. The relative entropy given by Mplus

software decreases when including more than 4 classes and increases again with the 6-class model; this would suggest that a model with at least 6 class or 4 class is preferred. Together with the percentage of reduction in the log-likelihood value, that indicates that by adding two classes to the model the log-likelihood is reduced by 12.1%, this reduction is only increased by 1.5% if the model is a 3-class model and finally this value is reduced close to 0 if more than 5 classes are included.

Now, the preferred model must be either the 4-class or higher model considering the residuals of each model in figure 4.8, where all values are around -1.96 and 1.96. Theoretically, we tend to determine that the 4-class LCA model is the preferred model. We will show later that the classes identified by the 4-class model are more interpretable and representative than the rest of the models. And in particular that 3-classes can be compared across countries.

Table 4.10: Model fit statistics LCA Students' endorsement of equal rights for all ethnic/racial groups scale

N Latent Classes	Param	Log- Likelihood	AIC	BIC	aBIC	Entropy	LL Re- duction	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
All countries											
1	5	-79684	159379	159423	159407						
2	11	-70046	140115	140212	140177	84.7%	12.1%	19276	0	18984	0
3	17	-68984	138003	138153	138099	75.4%	1.5%	2124	0	2091	0
4	23	-68807	137660	137863	137790	77.5%	0.3%	355	0	350	0
5	29	-68755	137568	137824	137732	74.0%	0.1%	104	0	102	0
6	35	-68754	137578	137887	137775	80.8%	0.0%	2	0.541	2	0.542
7	41	-68754	137589	137951	137821	82.0%	0.0%	1	0.528	1	0.528

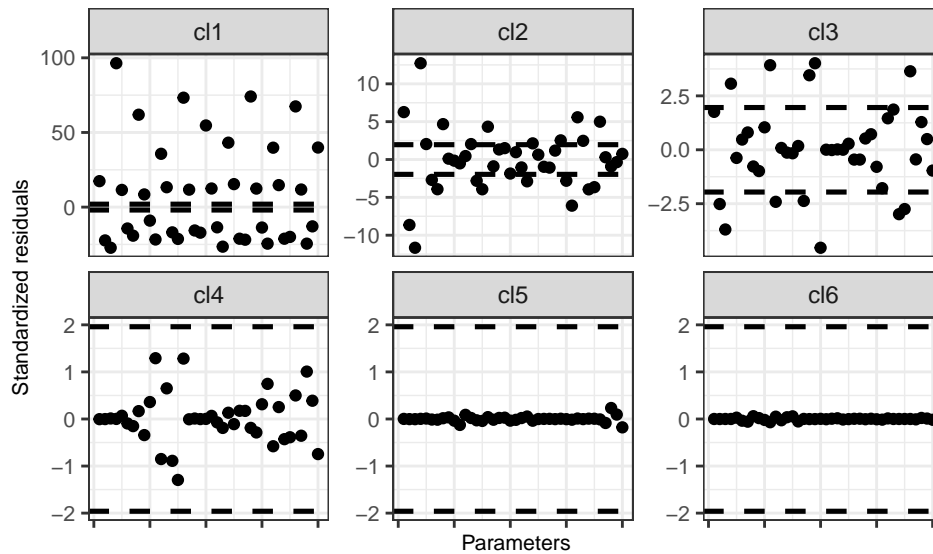


Figure 4.8: Bivariate model fit standardized residuals global model for Students' endorsement of equal rights for all ethnic/racial groups scale

Similarly to the previous scale, the three, four, five and six class models were investigated profoundly. Is not easy to choose the best model fit without doing a full analysis. There are some patterns that can be clearly identified in all the models, Class 1 with sizes of 75.9%, 75.4%, 74% and 78.8% in each model respectively, the estimated probabilities to agree for this latent class, the **Fully egalitarian** group, for all four first items are higher than 0.99 and 0.83 for the item All ethnic and racial groups should be encouraged to run in elections.

The second class called **Political non-egalitarian** (Class 2 in figure 4.9) is identified in all the models and have estimated probabilities to agree to the first 2 items higher than 0.93. For the next two items, the estimated probabilities to agree are around 0.66 and 0.75 in all models and for the last item the probability decreases to 0.5. The class size differs in all four models, 21.7%, 16%, 15% and 8.4% in the 3, 4, 5, 6-class model respectively.

The third class that can be seen with a similar pattern in all the models is called **Non-egalitarian**, this class appears from the 3-class model. The pattern of this class is basically showing lower estimated conditional probabilities to agree to any of these statements, no greater than 0.13. The estimated sizes for this class are 2.4%, 6.5%, 7.1% and 7% in each model respectively.

The fifth and sixth classes identified in the models differ in all the countries, nevertheless, one class appears to be consistent in the 5-class model, where this class called **Employment non-egalitarian** has low conditional probabilities to agree (0.2) to the item *All ethnic and racial groups should have an equal chance to get good jobs*.

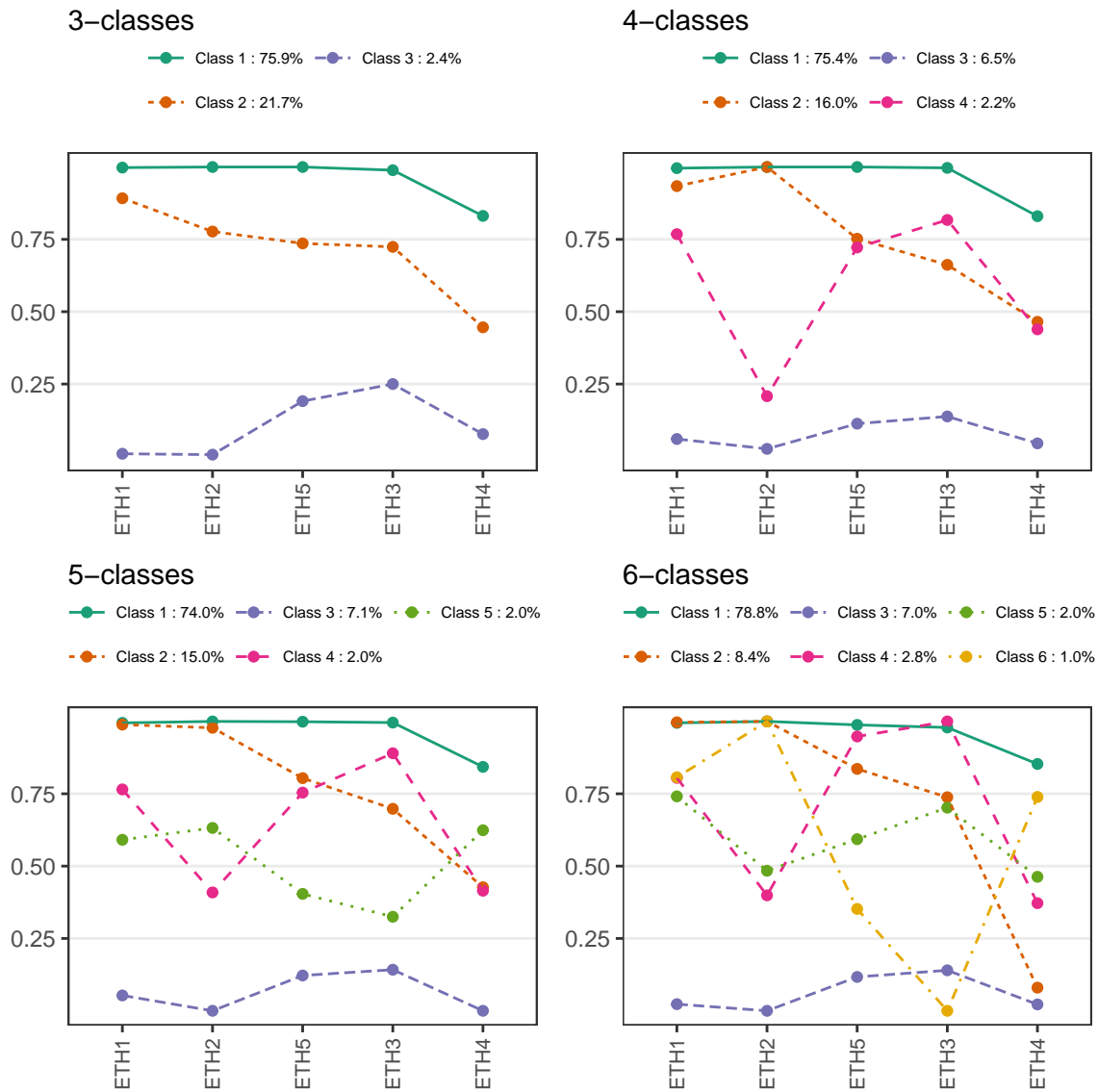


Figure 4.9: Comparative conditional probabilities to agree in 3 to 6 latent class global models for Students' endorsement of equal rights for all ethnic/racial groups

The main three classes in the solutions with three, four and five classes do not strongly differ from other models, and the remaining classes are not informative at all or with very small size; using this as a criterion, one can prefer a four-class solution. In table 4.11 the conditional probabilities to agree are shown. These values are very close to 1 in the first class, Fully egalitarian. Similar values are obtained for the first two items in the second class Political non-egalitarian; the next three items start decreasing the conditional probability to agree from 0.76 to 0.46. Class sizes shown in 4.12 indicate that proportions of unconditional probabilities even though are not exactly the same, have similar values among the Model estimated and Most likely classification.

Table 4.11: Probabilities to agree each item four-class global model for Students' endorsement of equal rights for all ethnic/racial groups scale

param	Fully egalitarian	Political non- egalitarian	Non- egalitarian	Employment non- egalitarian
ETH1 - All ethnic and racial groups should have equal chance to get good education	0.996	0.934	0.06	0.768
ETH2 - All ethnic and racial groups should have an equal chance to get good jobs	1	1	0.026	0.208
ETH5 - Members of all ethnic and racial groups should have same rights and responsibilities	1	0.752	0.113	0.722
ETH3 - Schools should teach students to respect members of all ethnic and racial groups	0.997	0.662	0.138	0.817
ETH4 - Members of all ethnic and racial groups should be encouraged to run in elections	0.83	0.465	0.045	0.439

Table 4.12: Class sizes four-class global model for Students' endorsement of equal rights for all ethnic/racial groups scale

Class	Model estimated		Most likely	
	Counts	Proportion	Counts	Proportion
Fully egalitarian	37774.4	75.4%	41600	83.0%
Political non-egalitarian	8014.1	16.0%	4902	9.8%
Employment non-egalitarian	3257.5	6.5%	2521	5.0%
Non-egalitarian	1080.1	2.2%	1103	2.2%

4.2.3 Country comparability

To evaluate the country comparability, the classes that were found in the independent models were identified to later check how many of them could be tested for comparability using a multigroup latent class model.

- 3-class model:
 1. Fully egalitarian: ALL COUNTRIES
 2. Political non-egalitarian: ALL COUNTRIES
 3. Non-egalitarian: ALL COUNTRIES
- 4-class model:
 1. Fully egalitarian: ALL COUNTRIES
 2. Political non-egalitarian: ALL COUNTRIES
 3. Non-egalitarian: ALL COUNTRIES
 4. Employment non-egalitarian: MLT

- 5-class model:
 1. Fully egalitarian: ALL COUNTRIES
 2. Political non-egalitarian: ALL COUNTRIES
 3. Non-egalitarian: ALL COUNTRIES
 4. Employment non-egalitarian: MLT
 5. Random response: Not identified in individual country models

With 3 classes, all classes are very interpretable. With 5 classes, a random response class is identified, which is not interpretable. With 4 classes, the Employment non-egalitarian class is present in just one country, which is not representative.

With a 4-classes model, three main classes are identified across countries. All the classes are present in all countries which means that is the best model for comparability. One remaining class can be freely estimated that varies in each country and/or with a class size of 0.

Country multigroup analysis

In table 4.13 different models with multigroup analysis are tested, first the more restricted model is evaluated, the complete homogeneity. In this model, all conditional and unconditional probabilities are fixed to be equal across the groups. Then, the partial homogeneity is tested where only the conditional probabilities are constrained to be equal across the groups, and the class sizes are estimated freely.

The second approach of partial homogeneity is tested too but not included in the table, where only the conditional probabilities for the three common classes identified are constrained across groups, and the remaining are freely estimated along with the unconditional probabilities.

Finally, the complete heterogeneous model is tested, where not only the unconditional probabilities are estimated freely but all the conditional probabilities as well. In the last two models the best log-likelihood is not replicated, this means that the solution may not be trustworthy due to local maxima. These results cannot be considered valid.

Table 4.13: Multigroup model fit statistics, global model with four-classes for Students' endorsement of equal rights for all ethnic/racial groups scale

Ngroups	Param	Log-Likelihood	AIC	BIC	aBIC	Entropy	LL Reduction	Δ LL	Δ DF	pvalue Δ
Four-class model										
Complete homogeneity										
14	36	-201088	402248	402566	402451	92.2%	-1.20%	-2383	-299	0
Partial homogeneity										
14	75	-199422	398994	399656	399418	88.9%	-0.36%	-717	-261	0
Complete heterogeneity										
14	335	-198705	398081	401036	399972	93.2%	0.00%			

Note:

The best log-likelihood value was not replicated for the following models:

¹ 4-class Complete heterogeneity model

Just by looking at the valid results, the partial homogeneity where all conditional probabilities are constrained to be equal across groups shows a better fit compared to the more restricted

model, with complete homogeneity. It is valid to indicate that the 4 classes identified do not share the same unconditional probabilities (class sizes) across the groups, but the conditional probabilities can be considered as equal in all groups.

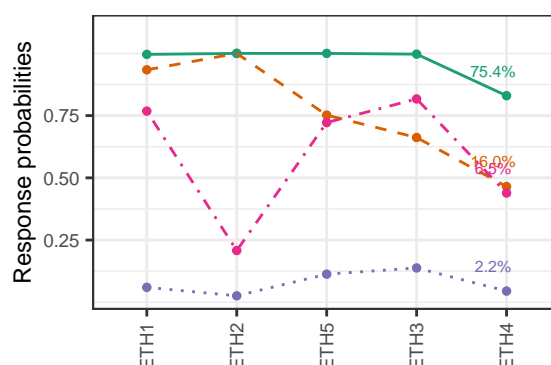


Figure 4.10: Conditional probabilities to agree in a 4-class complete homogeneous multigroup model for Students' endorsement of equal rights for all ethnic/racial groups scale

Similarly to the gender equality scale, it was not possible to establish complete invariance between the countries for the ethnic and race equality scale, the Fully egalitarian class concentrate the 75.4% of the population when looking at the global model (figure 4.10), Political non-egalitarian class includes 16% of the population, 6.5% of the population belongs to Employment non-egalitarian class and 2.2% to the Non-egalitarian class.

Table 4.14: Class sizes for partial homogeneity model for Students' endorsement of equal rights for all ethnic/racial groups scale

Group	Fully egalitarian	Political non-egalitarian	Non-egalitarian	Employment non-egalitarian
Belgium (Flemish)	47.7	44.1	0.9	7.3
Bulgaria	37.8	37.9	5.4	18.9
Croatia	62.0	30.5	1.1	6.5
Denmark	57.9	31.5	2.6	8.1
Estonia	69.4	22.8	0.7	7.2
Finland	71.8	20.9	2.0	5.2
Italy	63.3	24.2	1.9	10.6
Latvia	38.6	45.4	1.6	14.3
Lithuania	66.4	24.2	1.2	8.2
Malta	60.6	15.3	1.5	22.5
Netherlands	45.7	35.3	2.9	16.1
Norway	85.1	3.2	2.5	9.1
Slovenia	51.8	32.3	2.5	13.3
Sweden	88.8	2.9	2.2	6.1

These proportions vary when looking at the partial invariance (indicated in Table 4.14) where only Finland, Norway and Sweden class sizes include more than the 70% of the population for the Fully egalitarian class. The size of Political non-egalitarian class includes in average 26% of the population, more specifically in Belgium (Flemish), Bulgaria and Latvia sizes are higher than 37.9% of the population. Similar to the previous class, the number of members in

the Employment non-egalitarian class differs from the global model to the partial invariance model, these numbers go from 5.2% in Finland to 22.5% in Malta. Finally, the Non-egalitarian class could be the only one that maintains the same proportion between the global and the country-specific size around 2.2%, with the exception of Bulgaria that includes 5.4% of the population.

4.2.4 Confirmatory Latent Class Analysis

The confirmatory model was performed by establishing some constraints based on the previous research. For the Students' endorsement of equal rights for all ethnic/racial groups, two hypotheses were tested. First, that students classified in the Fully egalitarian and Political non-egalitarian classes would agree equally to first two items in the scale *All ethnic and racial groups should have equal chance to get good education* and *All ethnic and racial groups should have an equal chance to get good jobs*. The other hypothesis test that beside the previous equality, students that would be classified in the Non-egalitarian class would have the remaining probability to agree to all the items, which would mean to be unlikely to agree.

Table 4.15: Probabilities to agree each item 4-class Confirmatory LCA Students' endorsement of equal rights for all ethnic/racial groups scale

param	Fully egalitarian	Political non-egalitarian	Non-egalitarian	Employment non-egalitarian
ETH1 - All ethnic and racial groups should have equal chance to get good education	0.995	0.995	0.005	0.668
ETH2 - All ethnic and racial groups should have an equal chance to get good jobs	0.993	0.993	0.007	0.414
ETH5 - Members of all ethnic and racial groups should have same rights and responsibilities	1	0.806	0	0.647
ETH3 - Schools should teach students to respect members of all ethnic and racial groups	1	0.697	0	0.734
ETH4 - Members of all ethnic and racial groups should be encouraged to run in elections	0.843	0.474	0.157	0.375

For this test, as stated in table 4.15, the conditional probabilities for the first and second latent class are equal for the first two items respectively. On the other hand, probabilities in the Non-egalitarian class are the difference between the total probability to agree (1) and the probability obtained in the Fully egalitarian class. The rest of the conditional probabilities were estimated freely.

In table 4.16 the model fit statistics of this model do not differ considerably from the exploratory approach analyzed previously but statistics such as AIC, BIC, and aBIC are still lower in the exploratory approach. This way it cannot be suggested that these restrictions apply to the population. Other hypotheses can be tested until obtaining a better model fit.

Table 4.16: Model fit statistics Confirmatory LCA Students' endorsement of equal rights for all ethnic/racial groups scale

Type	N Latent Classes	Param	Log- Likelihood	AIC	BIC	aBIC	Entropy	LL Re- duction
All countries								
Exploratory LCA	4	23	-68807	137660	137863	137790	77.5%	
Confirmatory LCA	4	16	-68911	137854	137995	137945	73.0%	-0.2%

Chapter 5

Discussion and conclusion

An important amount of studies have been performed in previous years addressing the comparability of attitudes towards different minorities groups but mainly based on a variable centered approach which provides valuable information on the country's average profile but it does not identify patterns (subpopulations) within countries. The main focus of this thesis was to identify different comparable patterns of students' attitudes of tolerance and respect for the rights of diverse social groups, particularly women and ethnic and racial groups. The research questions in this study were based on exploring the different possible profiles of students that could be found and if their comparability across countries were possible, in both Students' endorsement of gender equality and Students' endorsement of equal rights for all ethnic/racial groups scales from ICCS 2016 study.

In order to achieve this goal, the most suitable methodology is Latent Class Analysis, a person-centered approach that by using a mixture model that considers different distributions can identify heterogeneous groups. An exploratory approach was implemented along with the study of invariance across 14 European countries. A confirmatory model was tested in order to provide more insightful patterns.

The person-centered approach gives the possibility to explore the different profiles that can be identified using the individual items that were used by the consortium to create the country's average validated indicators. Together with this, now it is not only possible to compare the countries' average level of the students' attitudes toward equal rights but it is also possible to learn how each country's population is composed, by subgroups of students' attitudes or patterns and compare them. Latent Class Analysis is a great tool for identifying these profiles based on a statistical approach, with tests and statistics that give strong evidence to state why those profiles were identified.

Despite the complexity of identifying reliable and interpretable profiles for the representative samples, the analysis provided strong evidence to identify the most common profiles across countries. Model fit statistics are of great help to provide statistical evidence to reject models that do not fit properly the data. In this research, it was a tendency that model fit statistic AIC tended to overfit the data, in most of the cases model fit statistics accepted a model but AIC would prefer a model with one more class.

For the **Students' endorsement of gender equality scale**, it is clear that two classes are highly similar across countries based on the country individual analysis, Fully egalitarian and Competition-driven sexism class. Both have high probabilities to agree to the positively worded items in the scale. Additionally, for members in the Fully egalitarian class, it is highly likely to agree to the (reversed) negatively worded items in contrast to members of the Competition-

driven sexism class that are highly likely to disagree with these items. This is consistent in most countries in which the probability to agree to the first (reversed) negatively worded item is random (0.5), this probability decreases consistently for the next (reversed) negatively worded item. This finding could suggest that an evaluation of the impact of negative wording on the interpretation of these items might be a good idea to disregard possible misunderstandings of the items real meaning. The members of the Non-egalitarian class would highly disagree with all the items even showing stronger disagreement with the (reversed) negatively worded items. Finally, a subpopulation that tends to agree to most of the items in this scale but strongly agrees to the ones related to political equality for women can be observed, it was called Political egalitarian.

Class sizes cannot be established to be equal in all countries due to failure in achieving complete homogeneity. Even though Fully egalitarian is the predominant class in most countries, particularly in Nordic countries, in some eastern European countries Political egalitarian class has the larger number of members. Moreover, the Non-egalitarian class size in some of these eastern European countries is composed of more than 10% of the population.

A confirmatory approach for the gender equality scale was performed to give evidence of the strong agreement of the Fully egalitarian class with gender equality based on opportunities in the government. Similarly, it was confirmed the relation that the second-highest probability to agree is shared by two items in the Fully egalitarian class and by one item in the Competitive-driven sexism class equally. More interesting is that students in the Competition-driven sexism class would not have a clear attitude towards equal rights for women in politics.

For the **Students' endorsement of equal rights for all ethnic/racial groups** scale, four classes are found as highly similar across countries, Fully egalitarian, Political non-egalitarian, Employment non-egalitarian and Non-egalitarian. Similarly to the gender equality scale, members of the first class, Fully egalitarian, highly agree to all the items in the global model but when looking at the individual country models the probability to agree to encourage members of all ethnic and racial groups to run in elections is lower compared to the other items, particularly in Belgium (Flemish), Bulgaria, Croatia, Latvia, Lithuania, Netherlands and Slovenia. The members in the Political non-egalitarian class are likely to agree mainly to equality in having a good education and chance to get good jobs for all ethnic and racial groups but less likely to agree to encourage them to run in elections. The members in the Employment non-egalitarian class are unlikely to agree that all ethnic and racial groups should have equal chances to get good jobs. And finally, members of the Non-egalitarian class are highly unlikely to agree to any of the items in this scale.

Considering the class sizes for this scale, comparability can not be granted. It is clear from the partially homogeneous model that the sizes differ greatly from country to country. Particularly, Nordic countries would concentrate most of the population in the Fully egalitarian class, meanwhile, Eastern European countries Political or Employment non-egalitarian classes would concentrate most of the population.

A confirmatory approach could not be established for this scale as the model fit statistics for the confirmatory model did not provide enough evidence to state that the model was better than the freely estimated model.

At this point, it is clear that complete comparability is not assured when we look at subpopulation patterns when analyzing Large Scale Assessments. Even though it is confirmed that at a variable level this comparison is straightforward, this does not apply necessary when comparing all different subgroups that can be identified within a country with another. It is still possible to obtain evidence that while most of the profiles are similar across countries

the proportion of students that relate to those profiles is not. This result was expected due to cultural background, particularly in countries more open to gender equality and cultural diversity such as Nordic countries and countries with economic challenges such as eastern European countries.

One of the limitations of this research is the complexity in the computation of some analyses, particularly measurement invariance models. This procedure requires a high amount of computational resources to perform properly due to the multiple models and parameters that should be estimated. This takes much time to perform as more countries are added into the analysis the more complex it gets and in most cases models will not converge or find global maxima. For this reason, a good strategy that allows testing a small number of models is necessary. In this research independent country models were conducted first to identify and select common patterns in the scales that would be proper for the measurement invariance models.

This research is of great help for researchers that want to perform the same analysis but for comparing different cycles of the study, as was mentioned, these scales are present in the study from 2009 and they will be present in the next cycle of 2022. Measurement invariance has to be tested for cycles using the same scales and countries before comparing. For researchers interested in study other countries, these results can be used for comparative purposes, as well as all the syntax provided with this research that produce the different models and comparisons.

Appendix A

Appendix

The appendix is composed by 3 units, A1 appendix provides complementary tables with information regarding the items and countries used in the analysis.

Appendix A2 include detailed output for the analysis performed in this research, here can be found the individual model fit statistics LCA by country for different number of classes for both scales studied. Plots for the response categories probabilities and class size for the final selected global model are included as well with the thresholds and the class sizes for the confirmatory global model.

Finally, appendix A3 includes the syntax used in R to perform the automation of the procedures with MplusAutomation and to provide tables and plot to report results. Mplus syntax used to perform Latent Class Analysis with complex data, along with the different multigroup models to test the level of invariance and the confirmatory model for the global sample are also included in this appendix.

A.1 Complementary tables

Table A.1: Items for students' endorsement of equal rights and opportunities. ICCS 2016

Item	Description
Gender equality	
IS3G24A	Men and women should have equal opportunities to take part in government
IS3G24B	Men and women should have the same rights in every way
IS3G24E	Men and women should get equal pay when they are doing the same jobs
IS3G24C	Women should stay out of politics (r)
IS3G24D	Not many jobs available, men should have more right to a job than women (r)
IS3G24F	Men are better qualified to be political leaders than women (r)
Equal rights for all ethnic and racial groups	
IS3G25A	All ethnic and racial groups should have equal chance to get good education
IS3G25B	All ethnic and racial groups should have an equal chance to get good jobs
IS3G25C	Schools should teach students to respect members of all ethnic and racial groups
IS3G25D	Members of all ethnic and racial groups should be encouraged to run in elections
IS3G25E	Members of all ethnic and racial groups should have same rights and responsibilities

Table A.2: Countries sample sizes that participate in ICCS 2016

AlphaCode	Country	Participating schools	Participating students
BFL	Belgium (Flemish)	162	2931
BGR	Bulgaria	147	2966
CHL	Chile	178	5081
TWN	Chinese Taipei	141	3953
COL	Colombia	150	5609
HRV	Croatia	175	3896
DNK	Denmark	184	6254
DOM	Dominican Republic	141	3937
EST	Estonia	164	2857
FIN	Finland	179	3173
HKG	Hong Kong SAR	91	2653
ITA	Italy	170	3450
KOR	Korea, Republic of	93	2601
LVA	Latvia	147	3224
LTU	Lithuania	182	3631
MLT	Malta	47	3764
MEX	Mexico	213	5526
NLD	Netherlands	123	2812
NOR	Norway	148	6271
PER	Peru	206	5166
RUS	Russian Federation	352	7289
SVN	Slovenia	145	2844
SWE	Sweden	155	3264

Note:

Countries in bold were selected for this research

A.2 Detailed output

Table A.3: Model fit statistics LCA by country Students' endorsement of gender equality

	N Latent Classes	Param	Log- Likelihood	AIC	BIC	aBIC	Entropy	LL Re- duction	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
Belgium (Flemish)												
	1	6	-4246	8504	8540	8521						
	2	13	-3843	7713	7790	7749	84.8%	9.5%	806	0	791	0
	3	20	-3812	7664	7784	7721	88.2%	0.8%	62	0.323	61	0.327
	4	27	-3795	7645	7806	7721	68.9%	0.4%	34	0.237	33	0.24
	5	34	-3789	7646	7849	7741	74.3%	0.2%	13	0.392	13	0.394
	6	41	-3785	7653	7898	7768	81.8%	0.1%	7	0.64	7	0.641
Bulgaria												
	1	6	-8369	16749	16785	16766						
	2	13	-7710	15446	15524	15483	63.6%	7.9%	1317	0	1294	0
	3	20	-7553	15146	15265	15202	69.8%	2.0%	315	0	309	0
	4	27	-7523	15100	15261	15175	75.2%	0.4%	60	0.077	59	0.08
	5	34	-7508	15083	15287	15179	71.2%	0.2%	30	0.445	30	0.45
	6	41	-7500	15083	15328	15198	73.7%	0.1%	15	0.458	14	0.46
Croatia												
	1	6	-6227	12466	12503	12484						
	2	13	-5417	10860	10942	10901	84.8%	13.0%	1619	0	1592	0
	3	20	-5368	10777	10902	10838	87.6%	0.9%	98	0.085	96	0.088
	4	27	-5353	10759	10929	10843	90.8%	0.3%	31	0.081	31	0.084
	5	34	-5349	10765	10978	10870	91.9%	0.1%	8	0.46	8	0.464

Table A.3: Model fit statistics LCA by country Students' endorsement of gender equality (*continued*)

N Latent Classes	Param	Log- Likelihood	AIC	BIC	aBIC	Entropy	LL Re- duction	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
6	41	-5346	10774	11031	10900	94.1%	0.0%	5	0.787	5	0.789
Denmark											
1	6	-7273	14557	14597	14578						
2	13	-6043	12111	12198	12157	91.7%	16.9%	2460	0	2420	0
3	20	-5951	11943	12077	12013	90.2%	1.5%	182	0.077	179	0.08
4	27	-5914	11883	12063	11977	91.2%	0.6%	74	0.174	73	0.178
5	34	-5894	11856	12084	11975	86.7%	0.3%	41	0.386	40	0.39
6	41	-5876	11833	12108	11978	84.7%	0.3%	38	0.52	37	0.523
Estonia											
1	6	-5754	11520	11556	11537						
2	13	-5044	10115	10192	10151	79.1%	12.3%	1419	0	1394	0
3	20	-5002	10045	10164	10100	80.8%	0.8%	84	0.011	83	0.012
4	27	-4974	10001	10162	10076	77.7%	0.6%	57	0.183	56	0.189
5	34	-4968	10004	10206	10098	78.9%	0.1%	11	0.54	11	0.541
6	41	-4965	10012	10256	10125	81.5%	0.1%	7	0.521	6	0.522
Finland											
1	6	-4273	8558	8594	8575						
2	13	-3520	7065	7144	7102	89.3%	17.6%	1506	0	1480	0
3	20	-3477	6993	7114	7051	90.7%	1.2%	86	0.037	84	0.039
4	27	-3461	6975	7139	7053	92.3%	0.5%	32	0.57	31	0.574
5	34	-3450	6968	7173	7065	91.3%	0.3%	22	0.212	21	0.215
6	41	-3442	6966	7214	7084	90.2%	0.2%	15	0.417	15	0.42

Table A.3: Model fit statistics LCA by country Students' endorsement of gender equality (*continued*)

	N Latent Classes	Param	Log- Likelihood	AIC	BIC	aBIC	Entropy	LL Re- duction	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
Italy												
	1	6	-5615	11242	11279	11260						
	2	13	-4903	9832	9912	9870	80.9%	12.7%	1424	0	1399	0
	3	20	-4830	9701	9824	9760	84.6%	1.5%	145	0	143	0
	4	27	-4821	9695	9861	9775	87.5%	0.2%	19	0.246	19	0.251
	5	34	-4814	9696	9904	9796	84.6%	0.1%	14	0.539	14	0.542
	6	41	-4810	9703	9955	9824	73.5%	0.1%	7	0.356	6	0.358
Latvia												
	1	6	-8765	17542	17578	17559						
	2	13	-8054	16134	16213	16172	68.4%	8.1%	1421	0	1397	0
	3	20	-7993	16027	16148	16085	72.4%	0.8%	121	0.052	119	0.055
	4	27	-7977	16009	16172	16086	73.7%	0.2%	32	0.416	32	0.42
	5	34	-7963	15994	16200	16092	77.7%	0.2%	29	0.752	29	0.754
	6	41	-7955	15992	16241	16110	78.6%	0.1%	15	0.458	15	0.459
Lithuania												
	1	6	-8481	16973	17011	16991						
	2	13	-7571	15168	15249	15207	78.3%	10.7%	1819	0	1788	0
	3	20	-7447	14934	15058	14994	82.5%	1.6%	248	0	244	0
	4	27	-7421	14895	15062	14977	78.0%	0.4%	53	0.167	52	0.172
	5	34	-7409	14885	15096	14988	77.3%	0.2%	24	0.456	23	0.46
	6	41	-7403	14888	15142	15012	71.5%	0.1%	11	0.755	11	0.755
Malta												

Table A.3: Model fit statistics LCA by country Students' endorsement of gender equality (*continued*)

N Latent Classes	Param	Log-Likelihood	AIC	BIC	aBIC	Entropy	LL Reduction	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
1	6	-7383	14779	14816	14797						
2	13	-6378	12782	12862	12821	78.5%	13.6%	2011	0	1977	0
3	20	-6236	12513	12637	12573	84.2%	2.2%	283	0.028	278	0.029
4	27	-6204	12462	12629	12543	87.2%	0.5%	65	0.235	64	0.238
5	34	-6190	12449	12660	12552	88.9%	0.2%	27	0.394	26	0.396
6	41	-6181	12443	12698	12567	88.7%	0.2%	20	0.481	19	0.483
The Netherlands											
1	6	-5373	10757	10793	10774						
2	13	-4829	9683	9760	9719	75.9%	10.1%	1088	0	1068	0
3	20	-4759	9557	9676	9612	87.0%	1.5%	140	0.074	138	0.076
4	27	-4742	9539	9699	9613	87.1%	0.3%	32	0.435	32	0.438
5	34	-4728	9525	9726	9618	77.8%	0.3%	28	0.543	28	0.546
6	41	-4723	9527	9770	9640	75.9%	0.1%	12	0.556	11	0.557
Norway											
1	6	-7878	15767	15807	15788						
2	13	-6289	12603	12691	12649	91.8%	20.2%	3178	0	3126	0
3	20	-6104	12247	12382	12318	95.3%	2.9%	370	0	364	0
4	27	-6068	12189	12371	12285	96.0%	0.6%	72	0.039	71	0.041
5	34	-6035	12137	12365	12257	93.2%	0.5%	66	0.281	65	0.286
6	41	-6024	12130	12406	12276	93.7%	0.2%	21	0.373	20	0.376
Slovenia											
1	6	-5202	10416	10451	10432						

Table A.3: Model fit statistics LCA by country Students' endorsement of gender equality (*continued*)

N Latent Classes	Param	Log- Likelihood	AIC	BIC	aBIC	Entropy	LL Re- duction	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
2	13	-4369	8764	8841	8800	84.2%	16.0%	1666	0	1637	0
3	20	-4308	8656	8775	8711	85.5%	1.4%	122	0.018	120	0.02
4	27	-4280	8614	8774	8689	87.5%	0.7%	56	0.158	55	0.163
5	34	-4271	8610	8812	8704	86.9%	0.2%	18	0.507	18	0.51
6	41	-4264	8611	8855	8724	86.9%	0.1%	13	0.297	13	0.298
Sweden											
1	6	-3877	7766	7802	7783						
2	13	-3155	6336	6415	6373	90.6%	18.6%	1444	0	1419	0
3	20	-3080	6200	6321	6258	94.3%	2.4%	150	0.004	147	0.004
4	27	-3049	6152	6316	6230	89.2%	1.0%	62	0.398	61	0.402
5	34	-3022	6113	6319	6211	89.7%	0.9%	54	0.133	53	0.136
6	41	-3016	6114	6362	6232	89.9%	0.2%	13	0.513	13	0.515

Note:

The best loglikelihood value was not replicated for the following models:

¹ Croatia, 6 classes model

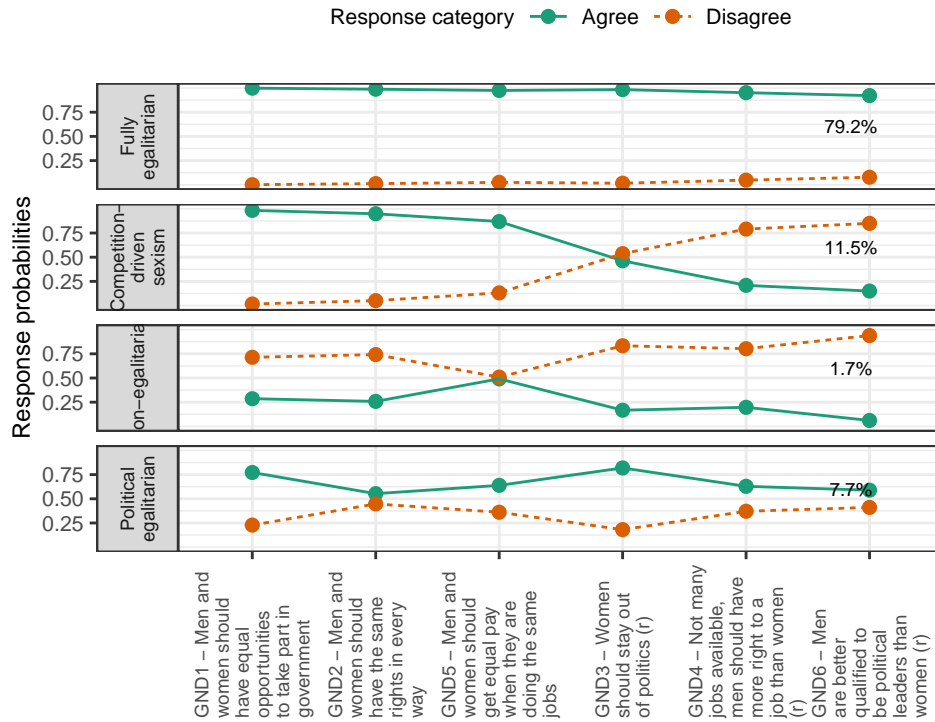


Figure A.1: Response categories probabilities and class size for 4-classes global model for Attitude towards gender equality scale

Table A.4: Thresholds 4-class Confirmatory LCA Students' endorsement of gender equality

Parameter	Fully egalitarian	Competition-driven sexism	Non-egalitarian	Political egalitarian
GND1\$1	15.000	4.228	-0.867	1.099
GND2\$1	4.228	2.777	-0.983	0.211
GND5\$1	3.668	1.852	-0.021	0.568
GND3\$1	4.228	0.000	-1.507	1.494
GND4\$1	3.014	-1.217	-1.419	0.595
GND6\$1	2.488	-1.604	-2.806	0.442
Means	2.404	0.541	-1.390	

Table A.5: Class sizes 4-class Confirmatory LCA Students' endorsement of gender equality

Class	Model estimated		Most likely	
	Counts	Proportion	Counts	Proportion
Fully egalitarian	39769.2	78.9%	41486	82.3%
Competition- driven sexism	6172.5	12.2%	5476	10.9%
Political egalitarian	3593.9	7.1%	2699	5.4%
Non-egalitarian	895.4	1.8%	770	1.5%

Table A.6: Model fit statistics LCA by country Students' endorsement of equal rights for all ethnic/racial groups scale

N Latent Classes	Param	Log- Likelihood	AIC	BIC	aBIC	Entropy	LL Re- duction	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
Belgium (Flemish)											
1	5	-4321	8653	8683	8667						
2	11	-4051	8123	8189	8154	85.1%	6.3%	542	0	530	0
3	17	-4019	8072	8174	8120	60.5%	0.8%	63	0.021	62	0.023
4	23	-4011	8068	8205	8132	61.8%	0.2%	16	0.231	16	0.235
5	29	-4009	8076	8249	8157	60.9%	0.1%	4	0.451	4	0.453
6	35	-4008	8086	8295	8184	88.2%	0.0%	1	0.627	1	0.627
Bulgaria											
1	5	-6406	12822	12852	12836						
2	11	-5451	10923	10989	10954	84.7%	14.9%	1910	0	1871	0
3	17	-5354	10742	10843	10789	78.5%	1.8%	193	0	190	0
4	23	-5335	10717	10854	10781	82.3%	0.3%	37	0.192	36	0.198
5	29	-5325	10708	10882	10790	88.3%	0.2%	20	0.385	20	0.39
6	35	-5322	10714	10923	10812	87.7%	0.1%	6	0.506	6	0.508
Croatia											
1	5	-5101	10213	10244	10228						
2	11	-4548	9119	9188	9153	87.6%	10.8%	1106	0	1084	0
3	17	-4507	9047	9153	9099	76.1%	0.9%	84	0.416	82	0.422
4	23	-4491	9027	9171	9098	82.8%	0.4%	32	0.126	31	0.128
5	29	-4479	9016	9197	9105	79.0%	0.3%	23	0.206	23	0.211
6	35	-4478	9027	9246	9134	86.6%	0.0%	1	0.618	1	0.618
Denmark											

Table A.6: Model fit statistics LCA by country Students' endorsement of equal rights for all ethnic/racial groups scale (*continued*)

N Latent Classes	Param	Log-Likelihood	AIC	BIC	aBIC	Entropy	LL Reduction	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
1	5	-9787	19585	19618	19602						
2	11	-8519	17061	17134	17099	86.9%	13.0%	2536	0	2488	0
3	17	-8309	16652	16766	16712	68.6%	2.5%	421	0	413	0
4	23	-8282	16610	16764	16691	72.3%	0.3%	54	0.059	53	0.062
5	29	-8280	16617	16811	16719	88.2%	0.0%	5	0.583	5	0.584
6	35	-8278	16625	16859	16748	84.8%	0.0%	4	0.545	4	0.546
Estonia											
1	5	-3577	7165	7195	7179						
2	11	-3298	6618	6684	6649	82.8%	7.8%	559	0	547	0
3	17	-3254	6543	6644	6590	69.8%	1.3%	87	0.002	86	0.002
4	23	-3245	6537	6674	6601	74.9%	0.3%	18	0.638	18	0.643
5	29	-3243	6545	6717	6625	75.2%	0.1%	4	0.344	4	0.345
6	35	-3242	6553	6761	6650	97.3%	0.1%	4	0.428	4	0.429
Finland											
1	5	-4186	8382	8412	8396						
2	11	-3491	7003	7070	7035	88.9%	16.6%	1391	0	1363	0
3	17	-3391	6815	6918	6864	80.8%	2.9%	200	0	196	0
4	23	-3370	6786	6925	6852	87.3%	0.6%	41	0.008	40	0.009
5	29	-3365	6788	6964	6872	93.6%	0.1%	9	0.367	9	0.372
6	35	-3364	6797	7009	6898	88.0%	0.0%	3	0.517	3	0.519
Italy											
1	5	-5113	10235	10266	10250						

Table A.6: Model fit statistics LCA by country Students' endorsement of equal rights for all ethnic/racial groups scale (*continued*)

N Latent Classes	Param	Log- Likelihood	AIC	BIC	aBIC	Entropy	LL Re- duction	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
2	11	-4427	8876	8943	8909	85.8%	13.4%	1371	0	1344	0
3	17	-4354	8742	8846	8792	81.1%	1.6%	146	0	143	0
4	23	-4330	8706	8847	8774	81.6%	0.6%	48	0.114	47	0.119
5	29	-4319	8696	8874	8781	85.2%	0.3%	22	0.082	22	0.086
6	35	-4316	8702	8917	8806	88.8%	0.1%	5	0.63	5	0.632
Latvia											
1	5	-5794	11599	11629	11613						
2	11	-5416	10853	10920	10885	77.5%	6.5%	757	0	742	0
3	17	-5353	10741	10844	10790	64.0%	1.1%	124	0.001	122	0.001
4	23	-5341	10728	10867	10794	71.3%	0.2%	25	0.357	24	0.362
5	29	-5335	10728	10903	10811	78.5%	0.1%	12	0.426	12	0.428
6	35	-5333	10736	10948	10837	87.8%	0.0%	3	0.592	3	0.593
Lithuania											
1	5	-4822	9654	9685	9669						
2	11	-4232	8486	8554	8519	87.9%	12.2%	1180	0	1156	0
3	17	-4194	8423	8528	8474	84.8%	0.9%	75	0.016	74	0.017
4	23	-4188	8422	8565	8492	88.9%	0.1%	12	0.514	12	0.518
5	29	-4183	8424	8603	8511	90.9%	0.1%	11	0.498	10	0.501
6	35	-4181	8433	8649	8538	93.0%	0.0%	3	0.533	3	0.534
Malta											
1	5	-6393	12796	12827	12811						
2	11	-5798	11619	11687	11652	72.0%	9.3%	1189	0	1166	0

Table A.6: Model fit statistics LCA by country Students' endorsement of equal rights for all ethnic/racial groups scale (*continued*)

N Latent Classes	Param	Log-Likelihood	AIC	BIC	aBIC	Entropy	LL Reduction	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
3	17	-5730	11493	11599	11545	74.1%	1.2%	138	0.002	135	0.002
4	23	-5691	11428	11570	11497	80.3%	0.7%	78	0.032	76	0.034
5	29	-5684	11425	11605	11512	78.4%	0.1%	15	0.381	14	0.385
6	35	-5680	11429	11646	11535	83.2%	0.1%	8	0.515	8	0.517
The Netherlands											
1	5	-5359	10727	10757	10741						
2	11	-4814	9650	9715	9680	79.6%	10.2%	1089	0	1067	0
3	17	-4729	9493	9593	9539	69.7%	1.8%	170	0	166	0
4	23	-4718	9482	9618	9545	73.8%	0.2%	23	0.358	22	0.365
5	29	-4711	9480	9651	9559	77.6%	0.2%	14	0.548	14	0.552
6	35	-4709	9487	9695	9583	77.7%	0.0%	4	0.483	4	0.484
Norway											
1	5	-7290	14590	14623	14607						
2	11	-5551	11125	11199	11164	94.7%	23.8%	3477	0	3412	0
3	17	-5448	10930	11044	10990	88.1%	1.9%	207	0	203	0
4	23	-5426	10897	11052	10978	90.0%	0.4%	45	0.057	44	0.06
5	29	-5421	10900	11094	11002	91.2%	0.1%	9	0.773	9	0.775
6	35	-5419	10908	11143	11031	92.2%	0.0%	4	0.293	4	0.294
Slovenia											
1	5	-4916	9841	9871	9855						
2	11	-4315	8652	8718	8683	86.7%	12.2%	1201	0	1176	0
3	17	-4272	8578	8679	8625	77.5%	1.0%	87	0.027	85	0.029

Table A.6: Model fit statistics LCA by country Students' endorsement of equal rights for all ethnic/racial groups scale (*continued*)

N Latent Classes	Param	Log- Likelihood	AIC	BIC	aBIC	Entropy	LL Re- duction	VLMR 2*LL Dif	VLMR PValue	LMR Value	LMR PValue
4	23	-4255	8556	8693	8620	83.9%	0.4%	34	0.101	33	0.105
5	29	-4248	8555	8727	8635	81.3%	0.2%	13	0.388	13	0.393
6	35	-4246	8563	8771	8660	85.7%	0.0%	4	0.541	4	0.543
Sweden											
1	5	-3175	6360	6390	6374						
2	11	-2379	4780	4847	4812	95.8%	25.1%	1592	0	1560	0
3	17	-2306	4646	4749	4695	90.2%	3.1%	147	0.011	144	0.012
4	23	-2299	4643	4783	4710	92.4%	0.3%	14	0.427	14	0.43
5	29	-2292	4642	4818	4726	93.2%	0.3%	13	0.432	13	0.434
6	35	-2291	4651	4863	4752	95.2%	0.1%	3	0.543	3	0.543

Note:

The best loglikelihood value was not replicated for the following models

¹ Denmark - 6-classes complete heterogeneity;

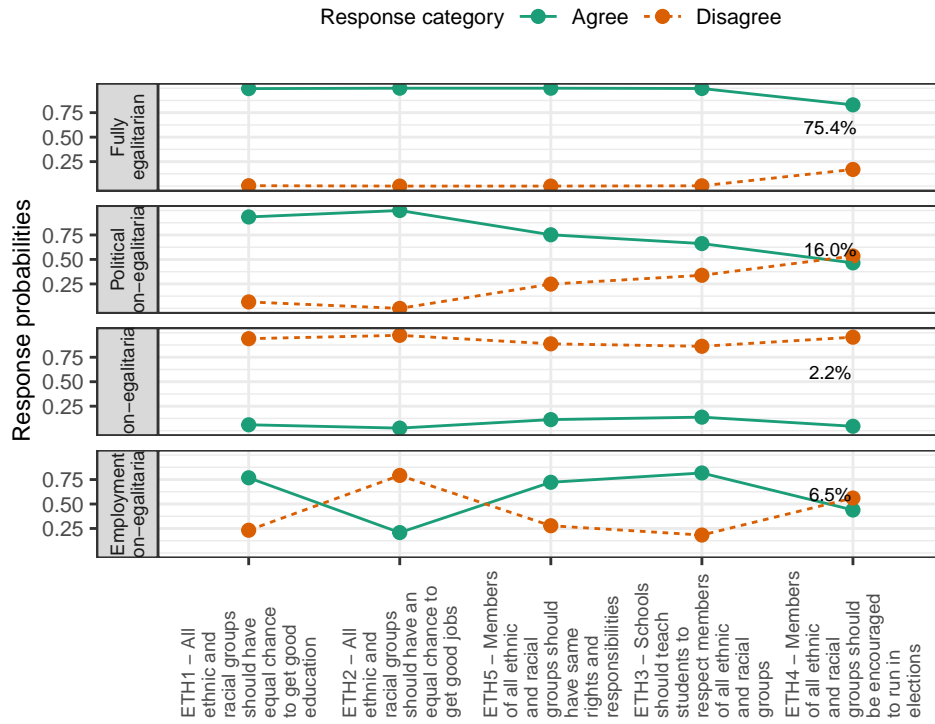


Figure A.2: Response categories probabilities and class size for 4-classes global model for Attitude towards ethnic and race equal rights scale

Table A.7: Thresholds 4-class Confirmatory LCA Attitude towards ethnic and race equal rights scale

Parameter	Fully egalitarian	Political non- egalitarian	Non- egalitarian	Employment non- egalitarian
ETH1\$1	5.265	5.265	-5.265	0.699
ETH2\$1	5.026	5.026	-5.026	-0.346
ETH5\$1	24.196	1.424	-24.196	0.606
ETH3\$1	17.167	0.831	-17.167	1.016
ETH4\$1	1.678	-0.102	-1.678	-0.509
Means	2.112	0.576	-1.815	

Table A.8: Class sizes 4-class Students' endorsement of equal rights for all ethnic/racial groups scale

Class	Model estimated		Most likely	
	Counts	Proportion	Counts	Proportion
Fully egalitarian	36968.5	73.8%	41874	83.5%
Political non-egalitarian	7955.2	15.9%	4053	8.1%
Employment non-egalitarian	4474.0	8.9%	3420	6.8%
Non-egalitarian	728.3	1.5%	778	1.6%

A.3 Syntax

Packages used

```
library(thesisdown)
library(plyr)
library(tidyverse)
library(knitr)
library(kableExtra)
library(MplusAutomation)
library(gridExtra)
library(grid)
library(scales)
library(RColorBrewer)
```

MplusAutomation syntax

```
library(MplusAutomation)
ds_lc <- data_model %>%
  dplyr::select(all_of(sample), all_of(Scales), IDSTUD, COUNTRY, CYCLE)

remlabclass <- function(ces){
  for (each in colnames(ces)){
    if ("labelled" %in% class(ces[[each]])){
      class(ces[[each]]) = c("numeric")
      attr(ces[[each]], "levels") <- NULL
    }
    attr(ces[[each]], "label") <- NULL
  }
  return(ces)
}
ds_lc0 <- remlabclass(ds_lc)

#-----By country scales together by CYCLE -----
for (j in 3:3) { #input file for each CYCLE 1:3
  data1 <- ds_lc0 %>% filter(CYCLE == paste0("C",j)) %>%
    dplyr::select(all_of(sample), all_of(ScalesGND), IDSTUD, COUNTRY) %>%
    mutate_if(is.factor, ~ as.numeric(.x)) %>%
    data.frame()

  cnt <- unique(data1[,c("COUNTRY","id_k")]) %>%
    arrange(as.character(COUNTRY))

  data1 <- data1 %>% dplyr::select(-COUNTRY)
  prepareMplusData(df = data1,
    filename = paste0("data/MplusModels/ByCountry/GNDDtaC",j,".dat"),
    interactive =FALSE)
```

```

for(c in 1:nrow(cnt)){
  data <- data1 %>% filter(id_k == cnt$id_k[c])

  lapply(1:6, function(k) { #input file for different number of classes
    fileConn <- file(paste0("data/MplusModels/ByCountry/GNDlca_",
                           cnt$COUNTRY[c], "_C", j, "cl",
                           sprintf("%d", k), ".inp"))

    writeLines(c(
      paste0("TITLE: ", cnt$COUNTRY[c], "GND LCA - C", j,
            " with ", k, " classes;"),
      "DATA: ",
      paste0("FILE = GNDDtaC", j, ".dat;"),
      "",
      "VARIABLE: ",
      paste0("NAMES = ", paste(colnames(data), collapse = "\n"), ";"),
      "IDVARIABLE = IDSTUD;",
      paste0("USEVARIABLES = ",
            paste(colnames(data)[grepl('^GND', colnames(data))],
                  collapse = "\n"), ";"),
      paste0("USEOBSERVATIONS ARE id_k EQ ", cnt$id_k[c], ";"),
      paste0("CATEGORICAL = ",
            paste(colnames(data)[grepl('^GND', colnames(data))],
                  collapse = "\n"), ";"),
      "MISSING = .; ",
      paste0("CLASSES = ", sprintf("c(%d);", k)),
      "WEIGHT = ws;",
      "STRATIFICATION = id_s;",
      "CLUSTER = id_j;",
      " ",
      "ANALYSIS:",
      "TYPE = COMPLEX MIXTURE;",
      "PROCESSORS = 4;",
      "STARTS = 1000 250;",
      "STITERATIONS = 20;",
      "STSEED = 288;",
      "",
      "MODEL:",
      "%OVERALL%",
      " ",
      "OUTPUT: ",
      "TECH10",
      "TECH11",
      "SVALUES",
      ";",
      "",
      "SAVEDATA:",

```

```

    paste0("FILE = Prob_", cnt$COUNTRY[c] ,
           "_GNDlca_C", j,"c1", k,".dat;"),
    "SAVE = CPROBABILITIES;"

  ), fileConn)
  close(fileConn)
})
}
}

runModels(target = "data/MplusModels/ByCountry", recursive = TRUE,
          replaceOutfile = "never") #modifiedDate
ByCountry_GND <- readModels(target = "data/MplusModels/ByCountry",
                           recursive = TRUE,
                           filefilter = "GNDlca_[A-Z]{3}_C3c1")
ByCountry_ETH <- readModels(target = "data/MplusModels/ByCountry",
                           recursive = TRUE,
                           filefilter = "ETHlca_[A-Z]{3}_C3c1")

save(ByCountry_GND,
     ByCountry_ETH,
     file = "data/MplusModels_ByCountry.RData")

#----GND 4 groups----
classes4GND <- c("Fully egalitarian",
                 "Competition- driven sexism",
                 "Non-egalitarian",
                 "Political egalitarian")
orden4GND <- c(2,4,3,1)
lcaGND_C3c14 <- lcaGND$GND_lca_C3c14.out$parameters$probability.scale %>%
  rename_with(~ c("Class", "value")[which(c("LatentClass", "est") == .x)],
              .cols = c("LatentClass", "est")) %>%
  mutate_at( c("param", "category", "Class"), ~ as.factor(.x)) %>%
  mutate(Class = factor(Class, levels = orden4GND, labels = classes4GND))

counts4GND <- full_join(lcaGND$GND_lca_C3c14.out$class_counts$modelEstimated,
                       lcaGND$GND_lca_C3c14.out$class_counts$mostLikely,
                       by = c("class"))

lcaGND_C3c14$orden = rep(c(1,2,4,5,3,6), each = 2)
VarClass(lcaGND_C3c14) %>% group_by(Class, param) %>%
  filter(category == 1) %>%
  select(orden, param, Class, value) %>%
  mutate(value = cell_spec(value, color = ifelse(value >= 0.75, "Myblue",
                                                ifelse(value < 0.75 & value >= 0.25, "Mygreen","Myred")))) %>%
  reshape2::dcast(orden + param ~ Class) %>% arrange(orden) %>% select(-orden) %>%

```

```

kbl(caption = "Probabilities to agree each item 4-class Gender equality model",
    booktabs = TRUE, longtable = TRUE, align = c("l", rep("r",4)),
    row.names = FALSE, digits = 3, escape = FALSE) %>%
kable_classic_2(full_width = F) %>%
kable_styling(latex_options = c("repeat_header", "HOLD_position"),
    font_size = 9) %>%
column_spec(1, width = "15em") %>%
column_spec(2:5, width = "5em") %>%
collapse_rows(1, valign = "top") %>%
print()

counts4GND %>%
  mutate(class = factor(class, levels = orden4GND, labels = classes4GND),
    proportion.x = scales::percent(proportion.x, accuracy = 0.1),
    proportion.y = scales::percent(proportion.y, accuracy = 0.1)) %>%
  arrange(desc(count.y)) %>%
  kbl(col.names = c("Class", "Counts", "Proportion", "Counts", "Proportion"),
    caption = paste0("Class sizes 4-class Gender equality model"),
    booktabs = TRUE, longtable = TRUE, align = c("l", rep("r",4)),
    row.names = FALSE, digits = 1, escape = TRUE) %>%
  kable_classic_2(full_width = F) %>%
  kable_styling(latex_options = c("repeat_header", "HOLD_position"),
    font_size = 9) %>%
  add_header_above(c(" " = 1 , "Model estimated" = 2, "Most likely" = 2))

sizelca4_GND <- lcaGND$GND_lca_C3cl4.out$class_counts$modelEstimated %>%
  dplyr::select(-count) %>%
  rename_with(~ c("Gender", "Class")[which(c("proportion", "class") == .x)],
    .cols = c("proportion", "class")) %>%
  mutate(Class = factor(Class, levels = orden4GND, labels = classes4GND)) %>%
  reshape2::melt(id.vars = c("Class"), variable.name = "Group") %>%
  dplyr::arrange(Group) %>%
  dplyr::group_by(Group) %>%
  dplyr::mutate(countT= sum(value, na.rm = TRUE)) %>%
  dplyr::group_by(Class) %>%
  dplyr::mutate(per=value/countT) %>%
  dplyr::select(Group, Class, per)

HighProb(lcaGND_C3cl4, sizelca4_GND, orden = c(1,2,5,3,4,6),
  title = "Response categories probabilities and class size
  for\n 4-classes Gender equality model")

```

Mplus syntax

Latent Class model with 4 classes

TITLE: LCA C3 GND with 4 classes;

DATA:

```

FILE = GND_Dta_C3.dat;

VARIABLE:
NAMES = id_i id_j id_r id_s
id_k wt ws
GND1 GND2 GND3 GND4 GND5 GND6
IDSTUD;
IDVARIABLE = IDSTUD;
USEVARIABLES = GND1
GND2 GND3 GND4 GND5 GND6;
CATEGORICAL = GND1
GND2 GND3 GND4 GND5 GND6;
MISSING = .;
CLASSES = c(4);
WEIGHT = ws;
STRATIFICATION = id_s;
CLUSTER = id_j;

ANALYSIS:
TYPE = COMPLEX MIXTURE;
PROCESSORS = 4;
STARTS = 100 50;
STITERATIONS = 5;
STSEED = 288;

OUTPUT:
TECH10
TECH11
TECH14;
SVALUES
;

SAVEDATA:
FILE = GND_Prob_C3cl4.dat;
SAVE = CPROBABILITIES;

```

Complete homogeneous multigroup latent class model with 4 classes

```

TITLE:C.Hom MG Country LCA GND C3 with 4 classes;
DATA:
FILE = GND_DtaC3.dat;

VARIABLE:
NAMES = id_i id_j id_r
id_s id_k wt ws
GND1 GND2 GND3 GND4 GND5 GND6

```

```
IDSTUD;
IDVARIABLE = IDSTUD;
USEVARIABLES = GND1
GND2 GND3 GND4 GND5 GND6;
CATEGORICAL = GND1
GND2 GND3 GND4 GND5 GND6;
MISSING = .;
CLASSES = g(14) c(4);
KNOWNCLASS = g(id_k =
  1      !      BFL
  2      !      BGR
  3      !      DNK
  4      !      EST
  5      !      FIN
  6      !      HRV
  7      !      ITA
  8      !      LTU
  9      !      LVA
  10     !      MLT
  11     !      NLD
  12     !      NOR
  13     !      SVN
  14     !      SWE
);
WEIGHT = ws;
STRATIFICATION = id_s;
CLUSTER = id_j;

ANALYSIS:
TYPE = COMPLEX MIXTURE;
PROCESSORS = 4;
STARTS = 1000 250;
STITERATIONS = 20;
STSEED = 288;

MODEL:
%OVERALL%
Model c:

                                %c#1%
[GND1$1-GND6$1] (91-96);
                                %c#2%
[GND1$1-GND6$1];
                                %c#3%
[GND1$1-GND6$1];
                                %c#4%
[GND1$1-GND6$1];
```

```

OUTPUT:
TECH10
SVALUES
;

SAVEDATA:
FILE = GND_Prob_MGCntry_C3cl4_3CHom.dat;
SAVE = CPROBABILITIES;

```

Partial homogeneous multigroup latent class model with 4 classes

```

TITLE: P.Hom MG Country LCA GND C3 with 4 classes;
DATA:
FILE = GND_DtaC3.dat;

VARIABLE:
NAMES = id_i id_j id_r id_s
id_k wt ws
GND1 GND2 GND3 GND4 GND5 GND6
IDSTUD;
IDVARIABLE = IDSTUD;
USEVARIABLES = GND1
GND2 GND3 GND4 GND5 GND6;
CATEGORICAL = GND1
GND2 GND3 GND4 GND5 GND6;
MISSING = .;
CLASSES = g(14) c(4);
KNOWNCLASS = g(id_k =
  1      !      BFL
  2      !      BGR
  3      !      DNK
  4      !      EST
  5      !      FIN
  6      !      HRV
  7      !      ITA
  8      !      LTU
  9      !      LVA
  10     !      MLT
  11     !      NLD
  12     !      NOR
  13     !      SVN
  14     !      SWE
);
WEIGHT = ws;

```



```
STRATIFICATION = id_s;
CLUSTER = id_j;
```

```
ANALYSIS:
TYPE = COMPLEX MIXTURE;
PROCESSORS = 4;
STARTS = 1000 250;
STITERATIONS = 20;
STSEED = 288;
```

```
MODEL:
%OVERALL%
c ON g;
```

```
    %g#1.c#1%
[GND1$1] (1);
[GND2$1] (2);
[GND3$1] (3);
[GND4$1] (4);
[GND5$1] (5);
[GND6$1] (6);
    %g#1.c#2%
[GND1$1] (7);
[GND2$1] (8);
[GND3$1] (9);
[GND4$1] (10);
[GND5$1] (11);
[GND6$1] (12);
    %g#1.c#3%
[GND1$1] (13);
[GND2$1] (14);
[GND3$1] (15);
[GND4$1] (16);
[GND5$1] (17);
[GND6$1] (18);
    %g#1.c#4%
[GND1$1] (19);
[GND2$1] (20);
[GND3$1] (21);
[GND4$1] (22);
[GND5$1] (23);
[GND6$1] (24);
    %g#2.c#1%
[GND1$1] (1);
[GND2$1] (2);
[GND3$1] (3);
[GND4$1] (4);
```

```

[GND5$1] (5);
[GND6$1] (6);
    %g#2.c#2%
[GND1$1] (7);
[GND2$1] (8);
[GND3$1] (9);
[GND4$1] (10);
[GND5$1] (11);
[GND6$1] (12);
    %g#2.c#3%
[GND1$1] (13);
[GND2$1] (14);
[GND3$1] (15);
[GND4$1] (16);
[GND5$1] (17);
[GND6$1] (18);
    %g#2.c#4%
[GND1$1] (19);
[GND2$1] (20);
[GND3$1] (21);
[GND4$1] (22);
[GND5$1] (23);
[GND6$1] (24);

```

```

.
.
.

```

```

    %g#14.c#1%
[GND1$1] (1);
[GND2$1] (2);
[GND3$1] (3);
[GND4$1] (4);
[GND5$1] (5);
[GND6$1] (6);
    %g#14.c#2%
[GND1$1] (7);
[GND2$1] (8);
[GND3$1] (9);
[GND4$1] (10);
[GND5$1] (11);
[GND6$1] (12);
    %g#14.c#3%
[GND1$1] (13);
[GND2$1] (14);
[GND3$1] (15);
[GND4$1] (16);

```

```
[GND5$1] (17);
[GND6$1] (18);
      %g#14.c#4%
[GND1$1] (19);
[GND2$1] (20);
[GND3$1] (21);
[GND4$1] (22);
[GND5$1] (23);
[GND6$1] (24);
```

OUTPUT:

TECH10

SVALUES

;

SAVEDATA:

FILE = GND_Prob_MGCntry_C3cl4_2PHom.dat;

SAVE = CPROBABILITIES;

Complete heterogeneous multigroup latent class model with 4 classes

TITLE: C.Het MG Country LCA GND C3 with 4 classes;

DATA:

FILE = GND_DtaC3.dat;

VARIABLE:

NAMES = id_i id_j id_r id_s

id_k wt ws

GND1 GND2 GND3 GND4 GND5 GND6

IDSTUD;

IDVARIABLE = IDSTUD;

USEVARIABLES = GND1

GND2 GND3 GND4 GND5 GND6;

CATEGORICAL = GND1

GND2 GND3 GND4 GND5 GND6;

MISSING = .;

CLASSES = g(14) c(4);

KNOWNCLASS = g(id_k =

```
1      !      BFL
2      !      BGR
3      !      DNK
4      !      EST
5      !      FIN
6      !      HRV
7      !      ITA
```

```

8      !      LTU
9      !      LVA
10     !      MLT
11     !      NLD
12     !      NOR
13     !      SVN
14     !      SWE
);
WEIGHT = ws;
STRATIFICATION = id_s;
CLUSTER = id_j;

```

```

ANALYSIS:
TYPE = COMPLEX MIXTURE;
PROCESSORS = 4;
STARTS = 1000 250;
STITERATIONS = 20;
STSEED = 288;

```

```

MODEL:
%OVERALL%
c ON g;

```

```

      %g#1.c#1%
[GND1$1] (1);
[GND2$1] (2);
[GND3$1] (3);
[GND4$1] (4);
[GND5$1] (5);
[GND6$1] (6);
      %g#1.c#2%
[GND1$1] (7);
[GND2$1] (8);
[GND3$1] (9);
[GND4$1] (10);
[GND5$1] (11);
[GND6$1] (12);
      %g#1.c#3%
[GND1$1] (13);
[GND2$1] (14);
[GND3$1] (15);
[GND4$1] (16);
[GND5$1] (17);
[GND6$1] (18);
      %g#1.c#4%
[GND1$1] (19);
[GND2$1] (20);

```

```
[GND3$1] (21);
[GND4$1] (22);
[GND5$1] (23);
[GND6$1] (24);
```

```
.
.
.
```

```
    %g#14.c#1%
```

```
[GND1$1] (313);
[GND2$1] (314);
[GND3$1] (315);
[GND4$1] (316);
[GND5$1] (317);
[GND6$1] (318);
```

```
    %g#14.c#2%
```

```
[GND1$1] (319);
[GND2$1] (320);
[GND3$1] (321);
[GND4$1] (322);
[GND5$1] (323);
[GND6$1] (324);
```

```
    %g#14.c#3%
```

```
[GND1$1] (325);
[GND2$1] (326);
[GND3$1] (327);
[GND4$1] (328);
[GND5$1] (329);
[GND6$1] (330);
```

```
    %g#14.c#4%
```

```
[GND1$1] (331);
[GND2$1] (332);
[GND3$1] (333);
[GND4$1] (334);
[GND5$1] (335);
[GND6$1] (336);
```

```
OUTPUT:
```

```
TECH10
```

```
SVALUES
```

```
;
```

```
SAVEDATA:
```

```
FILE = GND_Prob_MGCntry_C3cl4_1CHet.dat;
```

```
SAVE = CPROBABILITIES;
```

Confirmatory latent class model with 4 classes for Students' endorsement of gender equality scale

```
TITLE: ConflCA C3 GND with 4 classes;
```

```
DATA:
```

```
FILE = GND_Dta_C3.dat;
```

```
VARIABLE:
```

```
NAMES = id_i id_j id_r
```

```
id_s id_k wt ws
```

```
GND1 GND2 GND3 GND4 GND5 GND6
```

```
IDSTUD;
```

```
IDVARIABLE = IDSTUD;
```

```
!subpopulation is (id_k == 1);
```

```
USEVARIABLES = GND1
```

```
GND2 GND3 GND4 GND5 GND6;
```

```
CATEGORICAL = GND1
```

```
GND2 GND3 GND4 GND5 GND6;
```

```
MISSING = .;
```

```
CLASSES = c(4);
```

```
WEIGHT = ws;
```

```
STRATIFICATION = id_s;
```

```
CLUSTER = id_j;
```

```
ANALYSIS:
```

```
TYPE = COMPLEX MIXTURE;
```

```
PROCESSORS = 4;
```

```
STARTS = 100 50;
```

```
STITERATIONS = 5;
```

```
STSEED = 288;
```

```
MODEL:
```

```
%OVERALL%
```

```
%C#1%
```

```
[GND1$1*15] (p1);
```

```
[GND2$1*4.3] (p2);
```

```
[GND3$1*4.2] (p2);
```

```
[GND4$1*3.2] (p4);
```

```
[GND5$1*3.8] (p5);
```

```
[GND6$1*2.6] (p6);
```

```
%C#2%
```

```
[GND1$1*6] (p2);
```

```
[GND2$1*6] (p8);
```

```
[GND5$1*2.4] (p9);
```

```
[GND3$1*0] (p10);
[GND4$1*-1.8] (p11);
[GND6$1*-2.3] (p12);
```

```
%C#3%
[GND1$1] (p13);
[GND2$1] (p14);
[GND5$1] (p15);
[GND3$1] (p16);
[GND4$1] (p17);
[GND6$1] (p18);
```

```
%C#4%
[GND1$1] (p19);
[GND2$1] (p20);
[GND5$1] (p21);
[GND3$1] (p22);
[GND4$1] (p23);
[GND6$1] (p24);
```

MODEL CONSTRAINT:

```
p1 = 15;
p10 = 0;
```

OUTPUT:

```
TECH10
TECH11
TECH14;
SVALUES
;
```

SAVEDATA:

```
FILE = GND_ConfProb_C3cl4.dat;
SAVE = CPROBABILITIES;
```

Confirmatory latent class model with 4 classes for Students' endorsement of equal rights for all ethnic/racial groups scale

TITLE: ConfLCA C3 ETH with 4 classes;

DATA:

```
FILE = ETH_Dta_C3.dat;
```

VARIABLE:

```
NAMES = id_i id_j id_r
```

```
id_s id_k wt ws
ETH1 ETH2 ETH3 ETH4 ETH5
IDSTUD;
IDVARIABLE = IDSTUD;
USEVARIABLES = ETH1
ETH2 ETH3 ETH4 ETH5;
CATEGORICAL = ETH1
ETH2 ETH3 ETH4 ETH5;
MISSING = .;
CLASSES = c(4);
WEIGHT = ws;
STRATIFICATION = id_s;
CLUSTER = id_j;
```

```
ANALYSIS:
TYPE = COMPLEX MIXTURE;
PROCESSORS = 4;
STARTS = 100 50;
STITERATIONS = 5;
STSEED = 288;
```

```
MODEL:
```

```
%OVERALL%
```

```
%C#1%
```

```
[ETH1$1*5.3] (p1);
[ETH2$1*15] (p2);
[ETH3$1*5.5] (p3);
[ETH4$1*1.7] (p4);
[ETH5$1*6.7] (p5);
```

```
%C#2%
```

```
[ETH1$1*4.5] (p1);
[ETH2$1*3.7] (p2);
[ETH3$1*0.8] (p8);
[ETH4$1*-0.3] (p9);
[ETH5$1*1.4] (p10);
```

```
%C#3%
```

```
[ETH1$1*-2.9] (p11);
[ETH2$1*-15] (p12);
[ETH3$1*-1.8] (p13);
[ETH4$1*-15] (p14);
[ETH5$1*-2] (p15);
```

```
%C#4%
```

```
[ETH1$1] (p16);
```



```
[ETH2$1] (p17);  
[ETH3$1] (p18);  
[ETH4$1] (p19);  
[ETH5$1] (p20);
```

```
MODEL CONSTRAINT:
```

```
p1=-p11;  
p2=-p12;  
p3=-p13;  
p4=-p14;  
p5=-p15;
```

```
OUTPUT:
```

```
TECH10  
TECH11  
TECH14;  
SVALUES  
;
```

```
SAVEDATA:
```

```
FILE = ETH_ConfProb_C3cl4.dat;  
SAVE = CPROBABILITIES;
```

References

- Agresti, A. (2013). *Categorical data analysis* (3rd ed). Hoboken, NJ: Wiley.
- Białowolski, P. (2016). The influence of negative response style on survey-based household inflation expectations. *Quality & Quantity*, 50(2), 509–528. <http://doi.org/10.1007/s11135-015-0161-9>
- 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. Retrieved from <http://www.jstor.org/stable/3598347>
- Carsten, R., & Schulz, W. (2018). Overview of the IEA international civic and citizenship education study 2016. In *ICCS 2016 technical report* (pp. 8–25).
- Davidov, E., Schmidt, P., & Billiet, J. (Eds.). (2011). *Cross-cultural analysis: Methods and applications*. New York: Psychology Press, Taylor & Francis Group.
- 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(1), 30–45. <http://doi.org/10.1007/s11199-016-0698-7>
- European Union, C. of the. (2021, February 26). Council resolution on a strategic framework for european cooperation in education and training towards the european education area and beyond (2021-2030) 2021/c 66/01.
- Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, 61(2), 215–231.
- Hagenaars, J. A., & McCutcheon, A. L. (Eds.). (2002). *Applied latent class analysis* (1st ed.). Cambridge University Press. <http://doi.org/10.1017/CBO9780511499531>
- Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An r package for facilitating large-scale latent variable analyses in mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(4), 621–638. <http://doi.org/10.1080/10705511.2017.1402334>
- Hancock, G. R., Haring, J., & Macready, G. B. (Eds.). (2019). *Advances in latent class analysis: A festschrift in honor of c. Mitchell dayton*. Charlotte, NC: Information Age Publishing, Inc.
- Hooghe, M., & Oser, J. (2015). The rise of engaged citizenship: The evolution of citizenship norms among adolescents in 21 countries between 1999 and 2009. *International Journal of Comparative Sociology*, 56(1), 29–52. <http://doi.org/10.1177/0020715215578488>

- Hooghe, M., Oser, J., & Marien, S. (2016). A comparative analysis of 'good citizenship': A latent class analysis of adolescents' citizenship norms in 38 countries. *International Political Science Review*, 37(1), 115–129. <http://doi.org/10.1177/0192512114541562>
- Invariance analyses in large-scale studies*. (2019). (OECD Education Working Papers No. 201). Retrieved from https://www.oecd-ilibrary.org/education/invariance-analyses-in-large-scale-studies_254738dd-en
- Isac, M. M., Miranda, D., & Sandoval-Hernández, A. (Eds.). (2018). *Teaching tolerance in a globalized world* (1st ed. 2018). Cham: Springer International Publishing : Imprint: Springer. <http://doi.org/10.1007/978-3-319-78692-6>
- Isac, M. M., Palmerio, L., & Werf, M. P. C. (Greetje). van der. (2019). Indicators of (in)tolerance toward immigrants among european youth: An assessment of measurement invariance in ICCS 2016. *Large-Scale Assessments in Education*, 7(1), 6. <http://doi.org/10.1186/s40536-019-0074-5>
- Kamakura, W. A., & Wedel, M. (2000). *Market segmentation: Conceptual and methodological foundations* (Second ed.). Boston: Kluwer.
- Kankaraš, M., Vermunt, J. K., & Moors, G. (2011). Measurement equivalence of ordinal items: A comparison of factor analytic, item response theory, and latent class approaches. *Sociological Methods & Research*, 40(2), 279–310. <http://doi.org/10.1177/0049124111405301>
- Kaplan, D., & Publications, S. (Eds.). (2004). *The sage handbook of quantitative methodology for the social sciences*. Thousand Oaks, Calif: Sage.
- Kasumovic, M. M., & Kuznekoff, J. H. (2015). Insights into sexism: Male status and performance moderates female-directed hostile and amicable behaviour. *PLOS ONE*, 10(7), e0131613. <http://doi.org/10.1371/journal.pone.0131613>
- Köhler, H., Weber, S., Brese, F., Schulz, W., & Carsten, R. (2018). *ICCS 2016 user guide for the international database*.
- Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. Boston: Houghton Mifflin.
- McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York: Wiley.
- Miranda, D., & Castillo, J. C. (2018). Measurement model and invariance testing of scales measuring egalitarian values in ICCS 2009. In A. Sandoval-Hernández, M. M. Isac, & D. Miranda (Eds.), *Teaching tolerance in a globalized world* (Vol. 4, pp. 19–31). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-319-78692-6_3
- Mukhopadhyay, P. (2016). *Complex surveys: Analysis of categorical data* (1st ed. 2016). Singapore: Springer Singapore : Imprint: Springer. <http://doi.org/10.1007/978-981-10-0871-9>
- Munck, I., Barber, C., & Torney-Purta, J. (2018). 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(4), 687–728. <http://doi.org/10.1177/0049124117729691>

- Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism, Clinical and Experimental Research*, 24(6), 882–891.
- Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, 55(2), 463–469.
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus: Statistical analysis with latent variables: User's guide* (7th ed.). Los Angeles: Muthén & Muthén.
- Nagin, D. (2005). *Group-based modeling of development*. Cambridge, Mass.: Harvard University Press.
- Nylund-Gibson, K., & Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science*, 4(4), 440–461. <http://doi.org/10.1037/tps0000176>
- Olivera-Aguilar, M., & Rikoon, S. H. (2018). Assessing measurement invariance in multiple-group latent profile analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(3), 439–452. <http://doi.org/10.1080/10705511.2017.1408015>
- Robertson, J., & Kaptein, M. (Eds.). (2016). *Modern statistical methods for HCI*. Cham ZG: Springer.
- Rutkowski, L., Davier, M. von, & Rutkowski, D. (2014). *Handbook of international large-scale assessment background, technical issues, and methods of data analysis*. Boca Raton: CRC Press.
- Srivastava, M. S. (2002). *Methods of multivariate statistics*. New York: Wiley-Interscience.
- UNESCO. (2020). Education for sustainable development: A roadmap.
- Vermunt, J. K. (2014). Latent class model. In A. C. Michalos (Ed.), *Encyclopedia of quality of life and well-being research* (pp. 3509–3515). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-007-0753-5_1604
- Wang, J., & Wang, X. (2020). *Structural equation modeling: Applications using mplus* (2nd ed.). Hoboken, NJ: Wiley.
- Wolfram Schulz, Carsten, R., Losito, B., & Fraillon, J. (2018). *ICCS 2016 technical report*. Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement.