Measuring and Modeling Intercultural Attitudes Among Adolescents Across Europe: A Multi-Level, Multiple-Group Analysis Examining Student Attitudes, Intergroup Contact, and School Climate

Measuring and modeling intercultural attitudes among adolescents across Europe: A multi-level, multiple-group analysis examining student attitudes, intergroup contact, and school climate

Julia Dianne Higdon

Dr. Katherine Masyn (co-chair)

Dr. Helen Haste (co-chair)

Dr. Fernando Reimers

Dr. Judith Torney-Purta

A Thesis Presented to the Faculty of the Graduate School of Education of Harvard University in Partial Fulfillment of the Requirements for the Degree of Doctor of Education
Harvard Graduate School of Education

2015

Acknowledgements

I would like to thank all of my committee members, Dr. Katherine Masyn, Dr. Helen Haste, Dr. Fernando Reimers, and Dr. Judith Torney-Purta, for their mentorship and support over several years, helpful feedback on multiple drafts of this dissertation, and continuous encouragement. This dissertation would not have been possible without their guidance and assistance. Moreover, their integrity, creative vision, high intellectual standards, and dedication to all of their students have made a lasting impact on me. Also, I would like to thank my fellow students at HGSE, especially Vanessa Beary, Vidur Chopra, Nell O'Donnell Weber, Matt Shaw, and Galen McQuillen, the Spencer Early Career Scholars, and many others, for providing moral support, intellectual inspiration, and accountability. I look forward to our continued collaboration in the years to come.

I would like to thank my children, Zac and Katrina, who encourage, motivate, and inspire me in every way. Also, I would like to thank my friends, especially Maura Branley and John Van Alst, as well as my family, especially my mother and father, Jan and Al Boyce. I would not have been able to complete this dissertation without their support.

I would like to thank my colleagues at Avenues: The World School, especially Dr. Ty Tingley, Dr. Dirk Delo, Sarah Bayne, John Ciallelo, Dr. Isil Celimli, Chris DeFelippi, Steve Hanon, Jeff Clark, and many others. Their support, as well as their commitment to benefit all children and the field of global education, motivated me to finish and continue to motivate me every day.

Table of Contents

ABSTRACT

European nations, built on democratic foundations, rely on participation that is inclusive of all groups. Among efforts to support social cohesion in Europe, investigating the development of intercultural attitudes—attitudes toward others on the basis of their intersecting group memberships—is a crucial area of research. Further, examining attitudes among adolescents is useful because of their growing capacity to understand complex systems, while still being engaged in formal education in which interventions aimed at developing positive attitudes are often applied. In this dissertation, I used data from the 2009 IEA ICCS (International Civic and Citizenship Education Study) to examine determinants of intercultural attitudes among adolescents (n=16,847) in seven countries across Europe—the United Kingdom (England only), Sweden, Switzerland, Spain, Bulgaria, Poland, and Greece—focusing on cultural contexts and school climates. First, I established the measurement invariance of a variety of measures of intercultural attitudes to support the validity of cross-cultural comparison, using a novel approach in the Bayesian framework. Using this method, I was able to validly compare intercultural attitudes across contexts in Europe, and these findings are interpreted and discussed. Then, I examined the ways in which intergroup contact, gender, and school climates were associated with intergroup attitudes across these seven countries. I found limited evidence of an association between native-born and immigrant contact and positive intercultural attitudes. However, I found that positive intercultural attitudes were consistently associated with positive and democratic school climates, as well as with gender and attitudes toward gender equality. In this dissertation, I present these findings, as well as provide interpretation, discussion, and future directions with regard to educational interventions.

Introduction

European nations, built on democratic foundations, rely on participation that is inclusive of all groups. Integration and the expansion of the European Union (EU) are increasing mobility, migration, and immigration (European Commission: Eurostat, 2009) throughout Europe. As such, some forms of intergroup conflict remain a serious issue. This is reflected in the popularity of political parties across Europe that run on anti-immigration and protectionist platforms (Minkenberg, 2011) and in anti-immigration protests, such as those of the *Golden Dawn* movement in Greece (Alderman, 2012) and in Bulgaria where an influx of refugees has engendered social tension in this country, which is the poorest in the EU (Higgins, 2013).

As part of an ongoing research effort to understand intergroup conflict and improve intergroup relations, in this dissertation, I will examine intergroup attitudes among adolescents in 7 countries in Europe using data from the large-scale IEA International Civic and Citizenship Education Study (IEA, 2009). Intergroup attitudes are defined here as attitudes toward others on the basis of their intersecting group memberships, such as cultural, linguistic, religious, and gender group memberships, among others. An example would be the attitudes of native-born Protestant youth in the United Kingdom, toward members of minority religious groups, such as non-native born Muslim youth. In this example, the native-born youths are members of various intersecting groups—national, religious, linguistic—and the non-native youths are as well.

Among efforts to support social cohesion, investigating intergroup attitudes among adolescents is a crucial area of research for several reasons. First, intergroup attitudes can be

a precursor to intergroup conflict. While not all negative intergroup attitudes lead to intergroup conflict, it is frequently the case that intergroup conflict occurs in the presence of negative intergroup attitudes as well. Second, while intergroup attitudes begin to develop early in the lifespan (Aboud, Tredoux, Tropp, Brown, Niens & Noor, 2012; Raabe & Beelman, 2011; Barrett & Oppenheimer, 2011), examining intergroup attitudes among adolescents is useful because of their growing capacity to understand complex systems, while still being engaged, almost universally, in formal education systems. Third, education systems are contexts in which interventions are commonly applied (see Aboud et al, 2012, for a review). And fourth, the attitudes of adolescents may not be influenced in the same ways as those of adults, which has been examined empirically more extensively in the research literature.

The goal of this research is to examine differences as well as similarities in the development of intercultural attitudes among adolescents in several countries across Europe. It will hopefully contribute to our understanding of cultural context and development and intercultural attitudes, and to the ongoing project of social cohesion in Europe. Schools are sites where intercultural attitudes of the community—fellow students, parents, teachers, and other adults--are transmitted to adolescents. It is further a context over which we, as educators, researchers, and policy makers, have some control. Thus, the perspectives of young people toward the environment at their schools, and how these relate to their intergroup attitudes, is the primary focus of this paper.

The following quote provides an example of intercultural attitudes from a classroom in the Czech Republic.

> "A typical example [of an intercultural topic in class] is the Roma question. Students say that of course they do not mind skin colour and that we do not need to talk about it. I try to explain that the Roma have a different culture and that we can mutually enrich each other, but the discussion usually goes in the direction of criminality issues or the misuse of social benefits." -quote from a teacher of language and literature in the Czech Republic (from Moree, Klaassen, & Veugelers, 2008).

This quote shows a teacher facing a dilemma in the classroom when approaching topics of multiculturalism. This teacher would like to direct the discussion toward understanding and even appreciating the contributions of other cultures, in this case the often marginalized culture of the Roma, but the topic usually goes toward common explanations for intolerance: criminality and social benefit misuse. The issue he is raising is likely salient for many teachers: the intergroup attitudes that have previously developed in young people, as well as those of the adults in the school.

Context, culture, and the development of intercultural attitudes

While there are many theories that seek to explain the development of intergroup attitudes, I have identified five theories that are especially relevant to intercultural attitudes toward minority groups in the context of schools and education, which are exemplified in the quote above. The first two theories identify larger ecological systems in which intercultural attitudes emerge: in the developmental niche (Torney-Purta & Amadeo, 2011; Torney-Purta & Barber, 2011) and in communities of practice (Wenger, 1998/2008). Thirdly, the cultural model (Haste & Abrahams, 2008) provides a system for understanding

the role of cultural tools and narratives in the development of intergroup attitudes. The additional two theories are concerned with components of intercultural attitudes that are situated within these contexts: real or perceived threat (Corenblum & Stephan, 2001) and intercultural contact (Allport, 1954 and Pettigrew, 1998). In this analysis, I seek integrate theory in a way that is useful in understanding, both broadly and in detail, the development of intercultural attitudes. For example, in the development of intergroup attitudes, understanding how young people use certain cultural narratives, at times narratives of threat and other times narratives of cohesion, within school communities is an intersection of the theories utilized here. This intersection of theories provides an extremely powerful tool for understanding the complexities of culture as it relates to intergroup attitudes.

Below, I present each set of theories in turn and then discuss their integration. Afterward, I present an empirical analysis of the IEA ICCS study (IEA, 2009) that utilizes an integrated model to contextualize and understand findings across seven countries in Europe. In the empirical analysis I focus on attitudes toward immigrants, racial and ethnic minorities, and immigrants, as examples of intergroup attitudes where the target groups are varied.

Culture shapes human development

It is essential to understand the role of culture and environment in the development of intercultural attitudes. The importance of examining culture and environment in development generally dates at least back to the work of Vygotsky (see Vygotsky, 1987), who postulated that the individual must be understood within the context of social-historical processes, institutions, and activities in which the individual is a part; as well as Bronfenbrenner (see Bronfenbrenner & Morris, 2006) who proposed an ecological system to understand the individual in their context (summarized in Rogoff, 2003, pp. 45-50). Here,

the purpose of considering the role of culture is to make differences in intercultural attitudes explicit and to understand how intercultural attitudes develop as individuals and groups make meaning of the complex environment. For example, different contexts have very different histories of immigration, social structures, and narratives regarding the role of immigration in society. According to Super and Harkness (2002), culture is the "organization of the developmental environment" and "makes available to scientific scrutiny the process by which culture affects the course of development" (p. 271). Because cultures and environments are so varied and complex, distilling the components that are most related to intercultural attitudes, as well as the inter-relationships between these components in an organized system, is challenging. An organizing framework is essential to understanding this complexity in a way that is productive.

Super and Harkness developed a framework that they called the developmental niche, which enables researchers to identify relevant components of the environment of the child, as well as the system of components that work together in child development. They were mainly concerned with child-care and child-rearing, and theorized that development occurs within niches (1986). Torney-Purta and Amadeo (2011) and Torney-Purta and Barber (2011) modified the developmental niche to understand the influence of culture and environment on emergent civic engagement, and other related outcomes, in adolescents in greater detail. The application of the developmental niche makes its usefulness apparent. This model provides an explicit place for historical and cultural practices and beliefs (in the neighborhood and more broadly--see Torney-Purta, 2013), daily interactions, and adults' beliefs in a system of related processes. It allows researchers to consider relevant

components of the environment that relate to civic outcomes, as well as the organizing structure, including systems and subsystems, of complex cultures and environments.

Torney-Purta and Barber applied the developmental niche model to their study on participatory human rights (2011), which I will describe in detail in order to illustrate the use of this model in this analysis. Using the data from the 1999 IEA civic education study from 1999, CIVED (see Torney-Purta, Lehmann, Oswald, Schulz, 2001), they used cluster analysis to identify groups of adolescents based on profiles of attitudes and beliefs, groups that are described in a way that reminds the reader of adolescents that they know. In this study, Torney-Purta and Barber extracted five clusters using data from students who were native to the country in Australia (n=3,331), England (n=3,043, only England was sampled in this study), Finland (n=2,782), Sweden (n=3,073), and the United States (n=2,811). They named the clusters that they found based on the characteristics of the attitudes of the students: social justice, conventional/inclusive, indifferent, disaffected, and alienated. They then modeled the distributions of these clusters across countries, and further determined attributes that predicted group membership using logistic regression. Details regarding these clusters can be found in Torney-Purta (2009), however I will describe two of the clusters here.

The social justice cluster was comprised of adolescents who have a high-level of support for immigrant rights, as well as minority and women's rights, but are not likely to believe in the importance of democratic participation. They describe these adolescents using this motto: "I believe in rights for everyone but do not feel obligated to do much about it" (2011, p. 477). They found that in England 25% of the adolescents fell into this cluster, as well as 16% in Finland and 23% in Sweden. They also found an alienated cluster. This

cluster had low levels of trust and negative attitudes toward immigrants and minority groups. Further, they found that about 8% of youth in England fell into this cluster, as well as about 7% in Finland and Sweden. While these percentages were lower than the other clusters they found, this percentage represented enough of the population of adolescents to present a serious concern. Further, fully 20% of alienated adolescents in Sweden and 15% in Finland, did not believe it is important to obey the law, compared to a range of 1-3% in the social justice cluster.

Applying the developmental niche model to their findings, Torney-Purta and Barber demonstrated the usefulness of examining differences in national context in the development of attitudes toward human rights. There was substantial variability across the nations included in their study that were aligned with national differences in orientation toward human rights, migration, and inequality. Further, they found important neighborhood-level contextual factors that are associated with attitudes. For example, they found that alienated youth tend to spend more time outside the home in the neighborhood in the evenings, and they found no statistically significant relationships between socio-economic status and membership in the alienated cluster in Western Europe. Finally, they found evidence for the role of schools in reducing the likelihood of membership in the alienated cluster, especially an open climate for discussion in the classroom. In this analysis, Torney-Purta and Barber used culture in the way that Super and Harkness described: in understanding the variation in attitudes toward human rights across cultures and environments in a way that is empirically testable and directly useful.

Attitudes develop within Communities of Practice

In order to understand how intercultural attitudes develop within national and cultural contexts, it is further useful to consider that intercultural attitudes develop within specific sub-processes of community practices, such as schools. It is through engaging in practice that adolescents come to learn, and challenge, the intercultural attitudes of others in the community as well as the norms of the larger society. In this practice, adolescents develop their own intercultural attitudes which may accept or challenge narratives found in the dominant culture. According to Lave and Wenger, communities of practice are groups that have regular interaction with a purpose in common, such as an apprenticeship in a profession or a classroom within a school, in which there is a specific domain of interest and learning is an "integral and inseparable act of social practice" (2002, p.57; see also Wenger, 1998). Experience is extended over time and there is a shared commitment to the specific goals of the community (Eckert, 2006). In these communities of practice, individuals participate on the periphery at first, and then move toward full participation as the goal. Through increasingly central participation, learners negotiate meaning from complex cultural practices (Kirkup, 2002). Most importantly, communities of practice involves doing, and "[it] is doing in a historical and social context that gives structure and meaning to what we do" (Wenger, 1998).

The concept of communities of practice is especially relevant because civic participation is an important learned practice in democratic societies. If youth have the opportunity to participate in democratic practice at schools, this is an example of participating in democracy on the periphery, under the guidance of adults, with full democratic participation as the goal. This theory emphasizes that learning is a social process, in which participants practice alongside more experienced members, engaging in tasks and

dialogue. However it is further hypothesized here that positive and democratic communities of practice are especially relevant. The conversation about the Roma people, in the setting of the classroom in the Czech Republic, is an example of a social process of learning about other cultures or reinforcing previously held beliefs. The students come to know the teachers' own attitudes toward other cultures as well as their classmates, both stated and unstated, through social interactions such as this one. They can also compare these attitudes to those of their parents and families. Finally, they might also learn about the principle of inclusiveness that is central to democratic practice, by understanding groups as having participatory human rights as well as engaging in practices together.

Along a similar rationale, positive communities of practice, in which members of the community care about each other's well-being, may also support positive intercultural attitude development. In a study in the Netherlands, Huijnk, Kaykel, and Coenders (2013) found that feelings of affection and warmth in a family were associated with more positive intercultural attitudes in the form of acceptance of intercultural marriage. Family warmth might be related to positive intercultural attitudes because it fosters the expectation of trustworthiness which may generalize outside of the family (Glanville & Paxton, 2007; Glanville & Andersson, 2013; Freitag & Traunmüller, 2009). Regular, warm interactions also may increase social networks in which social norms are practiced and shared (Putnam, 2000).

Barber, Fennelly, and Torney-Purta (2013), in their multi-level cross-national study on the association between protective nationalism and attitudes toward immigrants among adolescents, examined some of the multi-level processes that are included in the revised developmental niche model. They found an association between protective nationalism and negative attitudes toward immigrants in countries with religious diversity and a long history

of democracy at the national level, but they did not find the same among newer democracies. At the school level, they found that higher socio-economic status and the proportion of foreign-born youth were both associated with positive attitudes toward immigrants. Their findings show that factors at the individual, school, and national levels are associated with positive attitudes toward immigrants, however these associations are not necessarily straightforward.

Narrative and dialogue shape attitude development

The cultural model of Haste and Abrahams (2008) provides another tool that is useful in this analysis. This model provides a framework to understand how culture plays a role in the development of moral reasoning. Haste and Abrahams' triangular model presents a system, which includes three aspects of development in the cultural context as well as their intersections: "individual, 'inside the head', making sense", "society, cultural resources, narratives," and "interpersonal dialogue" (p. 381). Intergroup attitudes are placed within a framework of individual cognition (which Haste and Abrahams argue has seen the greatest focus in the field of psychology), in communication with others, and using cultural narratives and resources available for young people to understand themselves in relation to other groups.

Further, crucially, this model provides a structure to consider the connections between these aspects of development in the cultural context. Recall the discussion in the classroom in the Czech Republic in which the subject of the Roma was raised. In this scenario, young people were engaged in a dialogue with others, including the teacher and possibly other adults who were in a position of authority, and it is in this dialogue that individual cognition and learning took place. It is further the case that this dialogue draws

upon common narratives, and indeed the narrative of the Roma as posing certain threats is a dominant narrative in many European countries (UNICEF Serbia, 2005). Thus, it is not necessarily the application of these different components that is most helpful here, but rather the ability to use them in conjunction to help explain the complexities of the relationship between culture and intergroup attitudes.

However, threats are not the only cultural narratives available, which is clear from the theory of Haste and Abrahams (2008). There are different cultural moral narratives available that deal with how one ought to structure relationships between oneself and others. For example, Gilligan (1992) juxtaposes a moral narrative of the imperative of care with that of justice, which she found to be more common among women. This is a moral structure in which the relationships between people are a primary concern of moral position taking, and this is in contrast to moral position taking that considers instead universal principles. While it is likely the case that individuals of either gender or alternate genders frequently utilize a moral narrative of care, the point is that this narrative and others are available to individuals in most contexts. Noddings (2005) further argued that dialogue is an important mechanism through which alternatives to dominant narratives are realized. This is an illustration of the idea that these narratives are in fact tools, because they are used to understand complex relationships.

Threat and contact reinforce or alter intercultural attitudes

There are two contextual features that have been examined widely in the research literature that I use in conjunction with the theories presented above: intercultural threat and contact. While there are many other theories that explain the development of intercultural attitudes, these two are particularly emphasized in the related research literature and lend

themselves to incorporation in the models proposed here. Intercultural attitudes emerge from real or perceived threats, such as migration, competition over scarce resources, and fear of cultural change, that are felt when groups come into contact within specific contexts. And further, intercultural contact can result in intercultural attitudes that are reinforced or change.

Coenders, Lubbers, and Scheepers (2009) used this line of reasoning when they traced Realistic Conflict Theory (citing Blumer, 1958; Coser, 1959; Blalock, 1967) and theorized that competition for scarce resources leads to negative intercultural attitudes if there is prior claim to the resources by a dominant or majority group. They further augmented this theory with Social Identification Theory (citing Tajfel, Turner, 1979; Tajfel 1981/2; Brown, 1995) in which people identify with their ingroup—and perceive their in-group to have priority to resources over to other groups, such as immigrants. The social identification process would intensify during times of economic competition, thus these two theories are used by Coenders and colleagues to form a composite explanation for negative attitude development, which they label Ethnic Group Conflict Theory. From this they hypothesized that levels of ethnic exclusionism would vary according to the level of competition within countries, and indeed they found that a change in unemployment and immigration had a negative association with attitudes toward immigrants. Fetzer (2000) came to a similar conclusion when finding that change in personal income is strongly negatively associated with attitudes toward immigrants.

Competitive threats to the dominant groups are felt especially in relation to groups that are perceived to be very similar in status. The competition must be realistic—groups need to be similar enough to represent a realistic threat (Tajfel, 1979; Brewer, 2007; Corenblum & Stephan, 2001). Threat leads to competition over scarce resources that are

believed to be finite in their availability (for example zero-sum beliefs, see Esses, Dovidio, Jackson, & Armstrong, 2001). Thus, threat leads dominant groups to seek to restrict access to jobs, education, and health care. For example, Reicher (2007) argues that it is important to examine the function that racism or denigrating outgroups serves for groups, which is argumentative and strategic (see also Hopkins, Reicher & Levine, 1997). Negative attitudes are employed strategically to argue for and establish priority of various resources for dominant majorities, perceived as under threat from other groups.

Stephan and colleagues postulate other kinds of threats. Symbolic threats are threats to the system of values of the ingroup (Stephan, Ybarra, Bachman, 1999). Cultural resources are also at stake in the dominant norms, narratives, values, and traditions of the nation. Indeed, attitudes toward immigrants have been found to be strongest when they represent both symbolic and realistic threats (Stephan, Renfro, Esses, Stephan & Martin, 2005; Esses et al, 2001). Finally, according to Allport (1954), competition prompts individuals to link resource needs to the status of the group, and this potential is greater in diverse societies with the possibility of social mobility, such as some of the nations included this study. Thus, when there is a context of competition over resources (real and symbolic), intercultural attitudes will be more negative as majorities seek to gain dominance. The history of parental beliefs regarding threat is brought to bear in the beliefs of adolescents as well. In the quote from the teacher in the Czech Republic, this is the tendency of the conversation to move toward criminality and misuse of social benefits that he describes as an issue that is expressed by the students, and potentially reinforced or challenged by the teacher and parents at home.

Contact theory postulates that increased direct exposure to members of other groups leads to lessened prejudice toward that group (Allport, 1954). Contact theory focuses on whether there is contact between groups in certain contexts, and the conditions under which positive attitudes develop when contact takes place. Positive attitudes are theorized to develop when the following four conditions are met: that the contact is between groups that are of equal status, that the contact is characterized by a common goal, that the groups are engaged in cooperative activity, with the endorsement of authorities (Allport, 1954; Pettigrew, 1998). Schools, where students work cooperatively toward a common goal under the authority of teachers and other adults, are potentially a context in which this kind of contact takes place. However, students may not view the status of groups in the classroom as equal, but rather their views may reflect the unequal status of groups found in dominant narratives. Further, intercultural attitudes develop early in life (Aboud et al, 2012), and thus young people in the classroom have developed not only intercultural attitudes, but also the narratives and strategic uses of negative attitudes. Thus, it is not likely that students are of equal status in the classroom. Returning again to the quote from the teacher in the Czech Republic in order to illustrate this theory, this raises questions regarding: whether there were minority students in this classroom and how these conversations might be different depending on the percentage of minority students or the size and nature of the groups represented, how the views of the students reflect dominant narratives, the kind of task in which the class is engaged, and whether there is a climate of mutual respect.

In a large-scale meta-analysis of 515 studies about the association between intercultural contact and prejudice, Pettigrew and Tropp (2008) found that intercultural contact reduces anxiety associated with intercultural contact, which reduces prejudice.

Contact is theorized as reducing the anxiety related to perceived threats. Pettigrew and Tropp also found that empathy is a significant mediator. These affective states were substantially stronger predictors of prejudice than cognitive knowledge of the target groups. Further, they found that the association between contact and reduced prejudice is stronger for majority group members than minority group members. These affective states are likely to be more developed in schools where the intercultural contact is sustained over time.

However, contact theory, especially when utilized in the absence of an examination of culture and context, is not sufficient to explain intercultural attitudes or their shifts and development. While contact may reduce prejudice by reducing anxiety in some cases, the theory does not properly address power and the strategic use of narrative to argue for dominance of some groups over others (Erasmus, 2010; Hopkins et al, 1997). The idea that contact, simply, would reduce prejudice depends on benign or neutral groupings of young people, which might be easily broken down through interaction in a classroom. It assumes that power relations, which reinforce hierarchical structures, are absent, making it utopian as a theory (Dixon, Durrheim, & Tredoux, 2005). Certainly, the context may serve to reinforce negative attitudes when contact takes place. Further, the experience of contact as negative or positive also makes an important difference, such that positive contact experiences more effectively reduce prejudice. Bekhuis, Ruiter, and Coenders (2013) conducted a study in the Netherlands that showed that positive experiences were associated with reduced ethnic distance, and negative experiences were associated with increased ethnic distance. However, they also found that this association was weakened in classrooms that emphasize multiculturalism (2013). Thus, I argue that contact not only should be viewed within the frame of the culture and context, but must be viewed within this frame.

Summary

The purpose of this study was to examine intercultural attitudes among adolescents, focusing on experiences within schools. The theories described above were useful to structure this exploration, and to provide a means to access the complexities of culture and intercultural attitudes in an explicit way. These theories together further motivate the structure of the empirical analysis that will be the focus of the two papers in this dissertation. In these analyses, intercultural attitudes are compared and contrasted in several European contexts in a multi-group analysis: England, Bulgaria, Greece, Poland, Spain, Sweden, and Switzerland. This approach is highly useful because forces, both proximal and distal, and the cultural narratives available to young people, differ across these contexts. Understanding the nature and the degree of these differences has the potential to deepen our understanding of intercultural attitudes, and to examine whether, and to what extent, the theories on hand might apply, or not, in each context.

Empirically examining intercultural attitudes cross-culturally

In this study, I conducted a secondary analysis of the IEA (2009) civic education data that included individual- and school-level predictors. I focused on attitudes toward varied groups: immigrants, racial and ethnic minorities, and migrants within Europe. I examined intergroup attitudes toward this range of groups, rather than toward any one group solely, in order to consider how intergroup attitudes might be different, or similar, depending on target groups. I focused only on attitudes of native students, in order to most clearly examine the views of a dominant majority to minority groups. Finally, I focused on schools using individual-level indicators of the school experience as well as school-level indicators. I examined the relationships between intergroup attitudes and individual- and school-level

indicators using multi-group multi-level structural equation modeling, described in more detail below. This framework allowed me to examine associations with attitudes toward the three different groups, at multiple levels, and across the seven countries simultaneously.

The countries that I selected were the United Kingdom (however, data were collected only in England, and not in Scotland, Wales, or Northern Ireland—this is the only group included which does not provide representative data across the entire country, and therefore will be referred to as England), Sweden, Poland, Bulgaria, Spain, Switzerland, and Greece. These seven groups provide coverage across regions in Europe: England in Western Europe, Sweden in Northern Europe, Switzerland in Central Europe, Spain and Greece in Mediterranean Europe, and Poland and Bulgaria in Eastern Europe.

These countries and England also represent a range of migration rates. Figure I.1 shows the relevant migration rates between the years 2003 and 2013. Bulgaria and Poland have had fairly consistently low, and at times negative, migration rates. Indeed, researchers have not examined intergroup attitudes toward immigrants in these areas as extensively as in immigrant receiving countries. In these areas, given the low and negative migration rate, when the category of "immigrant" is thought of by students, it is possible that these students are considering emigration in the future, or the emigration of members of their own family. However, examining attitudes towards immigrants in Bulgaria is a crucial and emerging area of research. Notably, when Bulgaria ascended to the EU in 2007, only two years before this study, Bulgaria became a new border area in the EU. Since then, there has been an increase in refugees from the areas of Jordan and Syria, however that wasn't the case when this study took place. In Bulgaria, intergroup attitudes frequently relate to minorities of Turkish descent, which comprises 8-9% of the population (National Statistical Institute, Republic of

Bulgaria (NIS-BR), 2011), a population which dates historically from the Ottoman Empire. Bulgaria also has a large Roma population, which comprises 8-10% of the population (NIS-BR, 2011).

INSERT FIGURE I.1 ABOUT HERE

In addition to having a near-zero net migration rate, Poland is a highly homogeneous nation, with nearly 97% of the population comprised of ethnic Poles (CIA World Fact Book, 2015). Demographically, Poland suffered enormous losses in WWII, and at that time the main ethnic group in Poland after the Poles, the Jews, fled or did not survive the war. While Spain and Greece had seen increased migration from 2003 – 2007, both of these countries had steeply declining net migration rates from 2007. Indeed, Spain had the highest migration rate of the seven countries included here, up until 2007, after which point the decline has been precipitous, with a negative migration rate near that of Greece in 2013. Sweden and the United Kingdom have had consistently high and positive migration rates, and are considered to be immigration receiving countries.

The affluence of the seven countries also varies. Figures I.2 and I.3 show the per-capita GDP and percent change in GDP, respectively, over the same time period, 2003-2013. The countries range from very low (Bulgaria and Poland) to very high (Switzerland and Sweden) per-capita GDP. However, starting in 2007 and reaching extremes in 2009, the year of this study, nearly all countries experienced large negative changes in GDP, which is clear in Figure I.3. This is important because the economic climate might have increased perceptions of threat very near the time when this study took place. All countries were affected, although Poland was affected somewhat less.

INSERT FIGURES I.2 AND I.3 ABOUT HERE

It is important to examine the political climate in the countries in this study because these national forces are a part of the developmental niches of young people. The national political climate, very broadly, also provides some indication of the views held by voting adults in each country. Even given the focus on adolescents in this study, it is still important to understand the attitudes of the adults, because attitudes among adolescents are developed in interaction and dialogue with adults. While the parents and teachers of these adolescents are most likely to have the strongest influence on the attitudes of adolescents, attitudes among the adolescents' particular teachers and among parents were not measured in this study. However, the political climate does provide a window into the overall atmosphere that the adolescents experienced in their daily lives. To examine political climate, I will briefly present information on the percentage of votes that went each of the major parties in each country.

Table I.1 provides the National Parliament results in each country, in the election year that was closest to 2009. Each of the countries in this analysis has a multiple-party system, with the multi-party system being most recent in Bulgaria and Poland, which had a single-party system during communist control. In each of the countries, two to three parties typically win the predominance of the votes, and these parties tend to be moderate in position, followed by parties that are specialized or have stronger positions. Over half of the votes in the national elections around 2009 in every country went to parties with centrist positions. However, in nearly all of the countries in this study, there was at least one nationalist party, usually described as "far-right," that ran on anti-immigration campaigns. These parties did not win a large number of votes, except in Switzerland. In Switzerland, the

far-right Swiss People's Party, which has been registered since 1971, won just over 28% of the votes. In Bulgaria, the far-right political party Ataka, registered since 2005, won just over 9% of the votes. Sweden and Greece had similar proportions of votes going to far-right parties. The right-wing Sweden Democrats, registered since 1988, won just over 5.7% of the votes (NSD European Election Database, n.d.) and the Popular Orthodox Rally, registered in 2000, obtained just over 5.6% of the votes in Greece. The Greek far-right party "Golden Dawn," was registered in 1985 but had lost popularity by 2009, when it obtained a mere .29% of the votes. In the United Kingdom, the British National Party, registered in 1982, only obtained 1.9% of the votes in their 2010 national parliamentary elections. Taken together, this provides some indication that the political climate around the time of this study was most restrictive toward immigrants in Switzerland and that this climate had been in place for decades, while a new, growing sentiment was taking place in Bulgaria. However, the political climate represented by more than half of the voting adults in each of the countries in this set was moderate.

INSERT TABLE I.1 ABOUT HERE

There are also other political parties that specialize in specific issues, and those that run on campaigns that are inclusive and typically lean toward socialism, or left-wing parties, which stand in contrast to the far-right parties discussed above. Many studies in the area of intercultural attitudes focus on far-right parties only. In Bulgaria, a party that advocates for Turkish minority rights, the Movement for Rights and Freedoms, registered since 1990, won nearly 15% of the votes in 2009. This is a helpful indication of not only the presence of this minority group in Bulgaria but also the prominence of the issues that this party raised in Bulgaria at the time of this study.

Spain did not even have a far-right party that obtained more than 1% of the vote in 2008, and has recently seen the astonishing rise of a new left-wing party, Podemos (see "Podemos cements potions in polls", Financial Times, Feb. 4, 2015) founded only in 2014 and now ranking second among political parties. While this party was not in existence in 2009 at the time of this study, the political climate that fostered such an astonishing rise in public opinion was likely forming at this time. Further, this party draws upon support for a larger movement in Spain that was the precursor to the Occupy movement in the United States, the Indignados. This movement has been directly related to the austerity measures that were put in place after the global financial crisis that hit near the time of this study. A similar party, Syriza, has been in place in Greece since 2004. This party won nearly 5% of the votes in 2009 and sharply rose in popularity to obtain nearly 17% of the votes in 2012, ranking second in political parties by votes (NSD European Election Database, n.d.). The current rise of these parties is useful to consider here because it indicates that political sentiment, when outside of a central position, has at least as much of a tendency toward left-wing positions as right-wing. And at least in Greece and Spain, a powerful counter-narrative to right-wing nationalism was readily available to young people.

The countries selected also represent a range of openness to migration across national policies reported by the Migrant Integration Policy Index (MIPEX), which is summarized in Table I.2. MIPEX collects data across several areas—labor market mobility, family reunion, education, political participation, long term residence, anti-discrimination, and access to nationality—and compiles scores in each area into an overall score. Sweden was the highest country in this set in 2010, which was consistent with the results in 2007. In 2010, Spain was the second-highest country in overall score, also consistent with 2007. The

United Kingdom, while as high as Spain in 2007, dropped considerably in 2010, primarily due to changes in long term residence and access to nationality. Greece was the lowest country in the set in 2007 in overall score, however the overall score increased in Greece in 2010, due to changes in access to nationality and political participation in this country. Thus, out of the set, while most countries were consistent between 2007 and 2010, right before and right after this study, Greece moved in a positive direction and the United Kingdom moved in a negative direction, potentially indicating a move toward openness in Greece in contrast with a move toward protectionism in the United Kingdom. Out of all of the countries in the set, Switzerland had the lowest MIPEX score both in 2007 and in 2010.

INSERT TABLE I.2 ABOUT HERE

They were also chosen to represent a range of democratic history and health, shown in Table I.3. The length of the current democratic system has been found to be associated with intercultural attitudes at the national-level in multi-level analysis (Barber et al., 2013). In the set of countries in this paper, the democracies in Bulgaria and Poland are the most recent, since 1989 after the end of communist rule in these countries. In Spain, the current democratic system has been in place since 1979 and the end of totalitarian rule in this country. And further, in Greece, the current democratic system has been in place since 1975 after a period of civil war which began after serious losses during resistance to the Nazis in WWII. Switzerland, Sweden, and the United Kingdom, are long-standing democracies in this group of countries. However, in Switzerland, universal suffrage that is equal for both men and women was not in place until 1971 (1990 in one of the cantons). Sweden and the United Kingdom have had equal universal suffrage since 1919 and 1928, respectively.

INSERT TABLE I.3 ABOUT HERE

Freedom House (Freedom House, 2009) reports the overall civil liberties and political rights with scores between 1 (most free) -7 (least free) in each category as well as an overall score which is the average, also shown in Table I.3. Nearly all countries received the highest rating, however Bulgaria and Greece received the next rating of 2. In Bulgaria, this was primarily due to widespread corruption in the government, and in Greece, this was primarily due to limits in media independence and lower protection of minority groups. The advancement of women in political careers is also varied in these countries. Only in Sweden did the proportion of seats held by women in the National Parliament approach parity in 2010 (45%) (World Bank, n.d.). Among the other long-standing democracies, the United Kingdom and Switzerland, comparatively small proportions of women held seats in National Parliaments in 2010, 22% and 29%, respectively. In this regard, the United Kingdom and Switzerland were closer to the countries in Eastern and Southeastern Europe—Bulgaria (21%), Greece (17%), and Poland (20%)—than they were to Sweden. Among this set of countries, Spain had the second highest proportion of women holding seats in National Parliament, 37%. This is likely due at least in part to a new law in 2007 in Spain that requires political parties to run women in at least 40% of their contested elections (Freedom House, 2009).

In summary, the countries in this analysis represent a wide range of history, economic conditions, political climate, and migration rates, which will be useful to contextualize the findings of this study, discussed below. Having such a range provides opportunities to understand how national contexts play a part in the development of intercultural attitudes among young people. Economic climate is helpful to understand the relationship of threat to intergroup attitudes as well. History and political climate also

provide insight into the openness of national contexts to issues of immigration and equality along many lines, including gender equality. This information sets the stage for the study which took place in 2009.

Paper 1: "Measuring intercultural attitudes among adolescents across Europe"

Introduction

European nations, built on democratic foundations, rely on participation that is inclusive of all groups. Integration and the expansion of the European Union (EU) are increasing mobility, migration, and immigration (European Commission: Eurostat, 2009) throughout Europe. As such, some forms of intergroup conflict remain a serious issue. This is reflected in the popularity of political parties across Europe that run on anti-immigration and protectionist platforms (Minkenberg, 2011) and in anti-immigration protests, such as those of the *Golden Dawn* movement in Greece (Alderman, 2012) and in Bulgaria where an influx of refugees has engendered social tension in this country, which is the poorest in the EU (Higgins, 2013). Further, the perception of negative attitudes toward immigrants has been found to discourage these residents from civic participation (Khanec & Tosun, 2009, in Germany).

Intercultural attitudes are defined here as attitudes toward others on the basis of their various intersecting cultural group memberships. Among efforts to support social cohesion, investigating the development of intercultural attitudes is a crucial area of research, particularly among adolescents and in cross-national comparison. While intercultural attitudes begin to develop early in the lifespan (Aboud, Tredoux, Tropp, Brown, Niens & Noor, 2012; Raabe & Beelman, 2011), examining intercultural attitudes among adolescents is useful because of their growing capacity to understand complex systems, while still being engaged, almost universally, in formal education systems. Considering education systems is

critical to support positive intercultural attitude development, because these are contexts in which interventions are commonly applied (see Aboud et al, 2012, for a review). Further, the attitudes of adolescents may not be influenced in the same ways as those of adults, which has been examined empirically more extensively in the research literature.

To explore intergroup attitudes, questionnaires are often used to measure attitudes across nationally representative groups. It is essential that the items that are used in these questionnaires have high levels of construct validity, meaning that the items measure the constructs of interest, in this case intergroup attitudes toward various groups. Therefore the construct validity of the items that were chosen from the IEA ICCS (2009) study to measure intercultural attitudes toward specific target groups are discussed below. Typically, sets of items are used in a factor so that, together, the items measure a construct more precisely than any single item alone could measure. In this study, sets of items were chosen that measured the attitudes of the adolescents towards specific target groups. However, this raises further questions regarding whether, and to what extent, the items within the sets measure the same construct, and whether the items in different sets measure different constructs. Thus, in addition to construct validity, evidence of convergent and discriminant validity is also an important aim of this paper, and discussed in detail below.

Further, relevant group-level differences, such as national differences, must also be taken into account. In order to validly compare intergroup attitudes at the factor level, it is essential to first establish that the measures used show adequate evidence of invariance across groups, meaning that the measures function similarly in each group (van de Schoot, R. Kluytmans, A., Tummers, L., Lugtig, P., Hox, J., & Muthen, B., 2013; van de Vijver & Poortinga, 2002; Steenkamp, & Baumgartner, 1998). Assuming that measures are invariant,

when in fact they are not, leads to bias in many of the estimates in which researchers are interested and potentially misleading results. In cases where measures are not invariant, which is typically the case in cross-cultural research, it is further important to diagnose the source of the non-invariance and specify the best possible measurement models. Thus, the evidence of measurement invariance is critical to establish, and this process is described in detail in this paper.

In this analysis I utilized a new method, Approximate Measurement Invariance in the Bayesian framework (Muthén & Asparouhov, 2013), to examine evidence of measurement invariance, as well as to identify sources of non-invariance. This method provides several advantages over the traditional CFA methods, which are described later in this paper. Given the recent implementation of these methods in widely available statistical software, one goal of this paper is to provide applied researchers with an empirical example of this new measurement invariance analysis method with many groups. This is in addition to the primary goal, which is to establish factor measurement models that are valid and have a sufficient degree of measurement invariance to allow valid cross-group comparisons. From this analysis, I will demonstrate scales that are effective in measuring intercultural attitudes among adolescents, as well as pinpoint scales and items that are less effective and may need further development. Implications for item developers are discussed below. I will further describe and draw conclusions from the results of the study itself.

Background

Measurement invariance is defined as the extent to which questionnaire items measure an underlying construct similarly across groups. Expected responses on items should be the same for any given factor score and irrespective of group membership, while

the factor scores themselves should be dependent on group membership (van de Schoot, et al, 2013). Typically, tests of invariance include evidence of the degree that the factor structure is the same across groups ("configural" invariance), that item loadings are the same across groups ("metric" invariance), and that item intercepts or thresholds are the same across groups ("scalar" invariance). Evidence of full invariance is evidence that the factor measurement model, derived in this case from the IEA ICCS questionnaire given to adolescents across Europe and other parts of the world in 2009, is invariant at the configural, metric and the scalar levels. If a measure is not at least partially invariant, it is impossible to discern whether observed differences are due to true differences in the construct or due to characteristics at the group-level that are not relevant to the construct.

This examination of measurement invariance is critical to the validity of cross-national comparisons of intergroup attitudes. Many studies in this area either do not conduct such analyses or do not report those results. In previous analysis (Higdon, 2013), I found that the items in the IEA ICCS (2009) study frequently used to measure intergroup attitudes were only partially invariant across nations. Therefore, one major aim of this paper is to establish the degree of measurement invariance (MI) necessary to support the proper use of measures of intergroup attitudes in cross-cultural comparison.

Traditionally, Confirmatory Factor Analysis (CFA) has been used to examine evidence of this invariance. However, establishing measurement invariance across several groups is a highly cumbersome task prone to error using this method. In this process, a series of increasingly restrictive models are fit. In the first model, the configural model, all of the item intercepts/thresholds and loadings are freely estimated and as such this is the least constrained model. The overall fit statistics are then compared to subsequent models where

all of the item intercepts or thresholds, as well as all of the factor loadings, are constrained to be equal across the groups by imposing equality constraints. A non-negligible decrease in the overall fit statistics as a result of the imposition of the equality constraints is taken as evidence of measurement non-invariance. The next step in this case is to release the equality constraints for some parameters across some groups until the overall fit statistics return to acceptable levels.

Further, traditional methods of evaluating measurement invariance are problematic in this analysis due to the number of groups involved. When the number of groups, and the number of possibly non-invariant parameters are large, the tasks of diagnosing which parameters are non-invariant in which groups and addressing that non-invariance are extremely cumbersome. The process of releasing the equality constraints, usually with the guidance of modification indices, is a step-by-step process which results in very many potential models (Muthén, & Asparouhov, 2013). This makes finding the ideal model, i.e. the model in which the maximum number of parameters are invariant with the least amount of detriment in fit, far too unwieldy and too subject to error. This also increases the possibility of Type I error, where making many comparisons capitalizes on chance.

There are two new methods, Alignment Optimization and Approximate Invariance, which are more feasible with a large number of groups and are promising methods to address the problems discussed above. Alignment Optimization provides a new strategy by which researchers can simultaneously estimate the factor means and variances in all groups, as well as discover which parameters are non-invariant (Asparouhov & Muthén, 2013); however, at the time of this writing, this method does not allow subsequent analysis that would utilize the measurement model, for example, in a structural equation model.

Approximate Invariance (Muthén & Asparouhov, 2013) is another new method in the Bayesian framework with which it is possible to impose equality constraints between some of the groups, and specify only a small amount of non-invariance between those groups or groupings that are not constrained to be equal. This method enables the researcher to establish a measurement model which has the greatest degree of invariance applied, without substantial detriment to model fit. It also allows the researcher to move from the measurement model to a structural model which is necessary to explore determinants of intergroup attitudes among adolescents, thus this is the approach to measurement invariance used in this paper.

Membership in the nations included in this study was a potential source of measurement non-invariance in this analysis, as the questions that were asked in the survey given to adolescents might have been viewed in different ways in each country. Given long-standing differences in history, laws, and cultural narratives toward various groups, it is reasonable that these differences might give rise to differences in the ways in which adolescents view, and subsequently answer, questions regarding their attitudes toward members of other cultures or groups. Thus, it is essential to examine whether, and to what extent, there is evidence that adolescents view and respond to these questions similarly in different nations.

Data

I used data from the 2009 IEA (International Association for the Evaluation of Educational Achievement) ICCS (International Civic and Citizenship Education Study) (IEA, 2009; Schultz, Ainley, J., & Fraillon., J, 2011; Schulz, W., Ainley, J., Fraillon, J., Kerr, D., & Losito, B., 2009) to answer the research questions in this study. In this survey, civic

knowledge, skills and attitudes were measured among 14-year olds across the world. The ICCS study sampled over 140,000 students, in more than 5,300 schools across 38 countries. National-level policy experts, teachers and school principals also responded to surveys providing contextual information. The study also included a European module that is included in this analysis, which asked further questions relevant to Europe, the European Union, and European identity (Kerr, D., Sturman, L., Schulz, W., & Burge, B., 2010), as well as Latin American and Asian modules. The sampling design was a stratified, two-stage probability sampling design, which was similar to other large-scale IEA studies such as PIRLS and TIMSS (Schulz, et al, 2011; Brese, Jung, Mirazchiyski, Schulz, & Zuehlke, 2011). At the time of this writing, sampling weights were not yet available in Bayesian analysis and therefore were not used in the measurement invariance models presented here. However, I incorporated elements of the complex survey design, including schools as clusters and sampling weights, in the correlations presented below, as well as the comparison of the standardized factor scores.

The IEA ICCS (2009) study was based on a previous large-scale civic education study undertaken by the IEA in 1999, CIVED, which took highly rigorous steps to ensure the validity of the questionnaires and surveys. The 1999 CIVED study was based on a two-stage design. In the first stage, extensive qualitative and case studies were undertaken in each of the participating countries to examine the meaning of civic-related constructs, including intercultural attitudes (Torney-Purta, Schwille & Amadeo, 1999). From these qualitative studies, survey instruments were developed in meetings with National Research Coordinators (Torney-Purta et al, 2001). The participating countries also pre-piloted and piloted preliminary forms of the instruments. The 1999 CIVED study included 28 countries

and sampled about 90,000 adolescents, 9,000 teachers, and 4,000 school principals. The instrument was written in English, translated into 22 languages, and then returned to the National Research Coordinators for checking (Torney-Purta et al, 2001).

The sample for this analysis using the IEA ICCS (2009) data was restricted to native-born adolescents and by selecting seven of the 24 countries in Europe that participated. Second-generation native-born adolescents, referred to in this paper more simply as "native-born," are the population of interest in this study. Thus, only students who reported that they, as well as their parents, were born in the country of the test were included in the sample, excluding both immigrant and second-generation youth. Further, I used the responses from the main student questionnaire as well as the European module from seven of the European countries that participated in both the main study and the European module (n=16,847). The countries that I selected were the United Kingdom (however, data were collected only in England, and not in Scotland, Wales or Northern Ireland—this is the only group included which does not provide representative data across the entire country, and therefore will be referred to as England), Sweden, Poland, Bulgaria, Spain, Switzerland, and Greece. The sample sizes in this study are shown in Table A1.1, in the Appendix. Further, in this table, the percent of native born in the sample is compared to the percent of native born in each country. There is close correspondence between these percentages in all countries except Switzerland, where there was a lower percentage of native born in the country compared to the sample.

These countries and regions (in the case of England) were chosen in order to provide coverage across regions in Europe: England in Western Europe, Sweden in Northern Europe, Switzerland in Central Europe, Spain and Greece in Mediterranean

Europe, and Poland and Bulgaria in Eastern Europe. They were also chosen to represent a range of governance histories: Spain, whose current democratic government is recent relative to the others, after an extended period of totalitarian rule which began before WWII; Poland and Bulgaria, also with current democratic systems which are recent, after an extended period of totalitarian rule under Soviet control which began after WWII; and England, Sweden, and Switzerland, which are long-standing democracies. Presumably, countries with longer-standing democracies have had a longer period of time to develop education systems to support democratic health. Further, they were chosen to represent a range in openness to immigration, as summarized by scores obtained from the Migrant Integration Policy Index (MIPEX), which is an ongoing migration policy assessment system led by the Migration Policy Group, with Sweden being the most open and Bulgaria being the least open in this set. Finally, these countries represent a range of migration rates, with negative migration rates in Poland and Bulgaria, and positive migration rates in the other countries. The contextual characteristics of these countries is discussed in greater detail in the second paper of this dissertation.

Measures

To measure intergroup attitudes, I identified subsets of items from the IEA ICCS (2009) study, which I present in Table 1.1. In this table, the groupings of items into four latent factors are given in the first column. The variable names provided in the technical documentation of the original study are given in the second column, and the exact item wording is given in the third column. These subsets of items were used to form four correlated latent factors: positive attitudes toward immigrants (IMMIG), open attitudes toward European migration (OPEN), protectionist attitudes toward European migration

(PROT), and positive attitudes toward racial and ethnic minorities (RETH). These items were chosen to provide adequate coverage of intergroup attitudes toward different target groups: migrants, immigrants, and racial and ethnic minorities. They were also chosen to include both positive (IMMIG, RETH, and OPEN) and negative attitudes (PROT). This coverage across positive and negative attitudes, as well as toward multiple target groups, helps support the content validity of these factors.

Three of the sets of measures that I used (IMMIG, RETH, and PROT) were used in some form in the CIVED (1999) study. The items were developed through a review of literature, building from similar items used with adults but rewritten to be shorter, easily translated, and appropriate for adolescents (Husfelt & Torney-Purta, 2004). The dimensional structure of these measures were examined after data collection using confirmatory factor analysis (CFA) and the items were then scaled within the framework of item response theory (IRT) (Schulz & Sibberns, 2004).

INSERT TABLE 1.1 ABOUT HERE

Responses to all items were coded on the following Likert-type scale: 1 - strongly agree, 2 - agree, 3 - disagree, 4 - strongly disagree, and missing responses were coded as missing; the response to each item provided the value of the observed variable. I converted the four-category Likert scale into a two-category agree/disagree binary scale with "strongly agree" and "agree" recoded as 1 and "strongly disagree" and "disagree" recoded as 0, to focus the analysis more sharply on agreement versus disagreement and not on degrees of disagreement or agreement. Further, I compared the correlation matrices of the two-

category variables with those of the four-category variables and found minimal differences, which indicated that negligible information was lost when then categories were collapsed. Thus, the two-category variables were used in the rest of the analysis.

The classic paper by Campbell and Fiske (1959) provides a useful method to examine convergent and discriminant validity: the multi-trait multi-method matrix (see also Crocker & Algina, 2008). While there are not multiple methods in this study, there are multiple sub-traits in the intergroup attitudes to various target groups under examination. An important aspect of the validity of this study is the extent of evidence that items grouped together within each factor measured the same latent construct—or convergent validity—found in strong inter-item correlations within each factor. In other words, it is important to establish whether these items measured intergroup attitudes toward that specific target group and only that target group. An equally important aspect of validity is the extent of evidence that items that are grouped together within different factors measure different latent constructs--or discriminant validity--found in low, or zero, correlations between items of different factors. Here, it is important to establish whether the measures properly discriminated between intergroup attitudes toward different target groups.

To measure attitudes toward immigrants (IMMIG), I used four items that captured attitudes toward immigrant opportunities: language, culture, education, voting, customs, and rights. These items showed strong internal reliability in the classical-test theory framework ($\alpha = .80$) (Cronbach, 1970; Crocker & Algina, 2008) indicating an acceptable degree of convergent validity. The relative frequencies of agreement with these items, in each of the seven countries, are given in Table 1.1. On average, about 80% of the young people surveyed agreed with the items. Overall, fewer adolescents agreed with this item,

"Immigrants should have the opportunity to continue speaking their own language," particularly in England, where 59% of the adolescents surveyed agreed with this item. Young people agreed most often with the item, "Immigrant children should have the same opportunities for education than other children in the country have," as 90% agreed overall, with a particularly high level of agreement found in Poland (95%). In Switzerland, it is important to note, the term "immigrants" in the German, was translated as "foreigners" (Torney-Purta et al, 2001).

To measure attitudes toward racial and ethnic minorities (RETH), I used four items which asked adolescents to agree or disagree with statements regarding the extension of opportunities and rights to racial and ethnic minorities. Overall, adolescents agreed slightly less frequently with these items than the items which measured attitudes toward immigrants, with an overall average rate of agreement of 85%. Adolescents agreed least often (71% agreement, on average) with the item "[m]embers of all ethnic/racial groups should be encouraged to run in elections for public office." Disagreement was particularly common in Bulgaria, with only 44% of adolescents agreeing with this item. The item that had the highest level of agreement was "[a]ll ethnic/racial groups should have an equal chance to get a good education in <country of test>" with, on average, 92% of adolescents agreeing, and in Poland in particular, 95% of the adolescents surveyed agreed. These four items showed strong internal reliability in the classical-test theory framework (α=.84) indicating an acceptable degree of convergent validity.

To measure protectionist attitudes toward migration (PROT), I used three items that captured negative attitudes on the basis of conflict, crime, and unemployment, given in Table 1.1. On average, 54% of the adolescents agreed with these items. The item with the

highest rate of agreement was "[a]llowing citizens of other European countries to come and work here leads to more unemployment for citizens of <country of test>." On average, 68% of adolescents agreed with this item, and particularly in England, where 74% agreed. The item with the lowest rate of agreement was "citizens of <country of test> will be safer from crime if they close their borders to <immigrants> from other European countries." On average, 46% of the adolescents surveyed agreed with this item, and the lowest rate of agreement was observed in Poland. These items had modest evidence of internal reliability in the classical-test theory framework (α=.68), indicating weaker evidence of convergent validity. Given that the research into protectionist attitudes toward migration within the EU specifically is in a relatively earlier stage than research into attitudes toward immigrants and racial and ethnic minorities more generally, particularly among adolescents, a value near .70 is sufficient for this particular analysis (Nunnally, 1978; Lance, Butts, & Michels, 2006).

Finally I used four items to measure the final latent factor under consideration: open attitudes toward migration (OPEN). Similar to the items used in the PROT factor, these items had weaker evidence of internal reliability in the classical-test theory framework (α = .63). Adolescents agreed with these items less frequently, on average about 80%, than with the items measuring attitudes toward racial and ethnic minorities and attitudes toward immigrants. The item in which the highest rate of agreement was observed was "[c]itizens of European countries should be allowed to live and work anywhere in Europe." Rates of agreement were highest in Poland and Bulgaria (96% and 95%, respectively), and lowest in Switzerland and England (77% and 78%, respectively). The item with the lowest rate of agreement was "[a]llowing citizens from other European countries to work here is good for the economy of <country of test>." On average, 68% of the adolescents surveyed agreed

with this item, with particularly high rates of agreement in Bulgaria and Poland (92% agreement in each).

Table 1.2 shows the tetra-choric correlations (n=16,847) within and between factors in a matrix. The tetra-choric correlation is more appropriate than the Pearson product-moment correlation when discrete variables are being analyzed (Olsson, 1979). Sampling weights were used so that each country contributed equally to the observed correlation. Each variable name was shortened in this table to include the first letter of the factor and a number for the variable.

INSERT TABLE 1.2 ABOUT HERE

The inter-item correlations between each factor provided mixed evidence of discriminant validity. The average correlation between items that measured open (OPEN) and protectionist attitudes (PROT) toward migrants was negative and very close to zero, -.07. This was the expectation, given that these items were grouped into two factors that measured positive versus negative views on the same issue, migration. Further, the average correlation between items that measured protectionist attitudes toward migration (PROT) and attitudes toward racial and ethnic minorities (RETH) was negative and very close to zero, -.10, and similarly the average correlation between items that measured protectionist (PROT) attitudes toward migration and attitudes toward immigrants (IMMIG) was negative and somewhat close to zero, -.15. These results were as expected, because the items that measured protectionist attitudes were negative in valence, while the items that measured attitudes toward racial and ethnic minorities and attitudes toward immigrants were positive in valence. This provided evidence that the items that measured protectionist attitudes

toward migration, were sufficiently discriminating with regards to attitudes toward immigrants and attitudes toward racial and ethnic minorities.

Items that measured open attitudes toward migration (OPEN) and attitudes toward racial and ethnic minorities (RETH) had an average correlation of .25 and items that measured open attitudes toward migration (OPEN) and attitudes toward immigrants (IMMIG) had a similar average correlation of .30. These are much higher than those reported above with regard to protectionist attitudes toward migration. Further, items that measured attitudes toward immigrants (IMMIG) and attitudes toward racial and ethnic minorities (RETH) had the weakest evidence of discriminant validity, with an average correlation of .43. This may be expected given that these items all had a positive valence, however further than that, attitudes toward racial and ethnic minorities could be conflated with attitudes toward immigrants, because many immigrants have minority racial and ethnic backgrounds. The items themselves did not provide definitions of immigrants versus racial or ethnic minorities, which could have led to confusion between the two categories among the adolescents who responded. That said, the average correlation of .43 between attitudes toward immigrants and attitudes toward racial and ethnic minorities was lower than the average correlations within each factor (r=.57 and r=.63, respectively). This was encouraging because it meant that the weight of evidence that the items within each factor measured the same construct at least outweighed the evidence that the items between each factor might not have adequately discriminated between constructs.

Taken together, the evidence of convergent and discriminant validity among the three factors of attitudes toward immigrants, attitudes toward racial and ethnic minorities, and protectionist attitudes toward migration was sufficient to move forward with these items

in the measurement model grouped within factors. To account for the correlations between the factors, I estimated the correlations between the factors in the final models in subsequent analyses as well.

However, the factor measuring open attitudes toward migration was too problematic to retain in the final measurement model because there was weak evidence of convergent validity as well as weak evidence of discriminant validity. There was limited evidence that the items grouped within this factor measured the same construct (r=.40), and there was limited evidence that the same items measured a different construct from attitudes toward immigrants (r=.30) and attitudes toward racial and ethnic minorities (r=.25). Further, my goal was to provide adequate coverage of the larger concept of intergroup attitudes, which would include both positive and negative attitudes, and include attitudes toward various target groups, in this case migrants, immigrants, and racial and ethnic minorities. The other three factors in the model provided this coverage: protectionist attitudes provided coverage of negative attitudes and attitudes toward migrants, and attitudes toward immigrants and racial and ethnic minorities provided coverage of positive attitudes and attitudes toward these groups. Thus, there was not a strong case for including open attitudes toward migrants from a construct coverage perspective. Given the weak evidence of both convergent and discriminant validity regarding open attitudes toward migration, and the construct coverage provided by the other factors, the tradeoff between including a weaker factor in the interest of construct coverage was not justified. I therefore removed the factor measuring open attitudes toward migrants from the measurement model.

The final measurement model, which will be further analyzed for evidence of measurement invariance in the rest of this paper, is presented in Figure 1.1. Having

established this measurement model through the examination of content, convergent, and discriminant validity, I then examined the measurement invariance of the items measuring each factor separately, which is described in the next section.

INSERT FIGURE 1.1 ABOUT HERE

In this figure, there are three latent factors measuring three reasonably distinct components of intergroup attitudes, shown on the right-hand side by the conventional circles. The variance of each factor is shown by the single-headed arrow pointing toward the factor, and the correlations between the factors are shown by the double-headed arrows between pairs of factors. On the left-hand side, each item is shown by the conventional rectangles, and the item loadings are shown by the single-headed arrows pointing from the factor to each item. A single item-level residual correlation between the items which ask adolescents whether immigrants should be allowed to keep their own language and culture, was allowed because they were similar in nature by centering on language and customs, compared to the other items, which centered on rights. This residual correlation is shown by the double-headed arrow on these items.

In order to validly compare the means and associations between the factors and other constructs of interest between countries, evidence of measurement invariance, as introduced above, must be established. In the following sections, I provide a detailed description of the examination of measurement invariance using Approximate Invariance with Bayesian estimation (Muthén, & Asparouhov, 2013). Subsequently, I present the standardized factor means and standard deviations, as well as factor correlations, in each country. The paper concludes with a discussion regarding the usefulness of the method and the findings of this analysis.

Methods

In traditional Multiple-Group Confirmatory Factor Analysis (MG-CFA), the typical multiple-group factor analysis model with binary indicators is specified by the following equations. Let *M* be the total number of items, *n* be the total number of respondents, and *G* be the total number of groups. The observed response, y_{mig} , to binary (0/1) item *m* (*m* = 1, …, *M*) by individual *i* (*i* = 1, …, *n*) in group *g* (*g* = 1, …, *G*) is given by:

$$y_{mig} = \begin{cases} 1 \; if \; y^*_{mig} > \tau_{mg} \\ 0 \; if \; y^*_{mig} \leq \tau_{mg}, \end{cases} \quad (1)$$

where τ_{mg} is the threshold for item *m* in group *g*. In this case the responses are 0 (disagree) and 1 (agree). The responses are 1 when the underlying latent response variable, y^*_{mig}, is greater than τ_{mig} and 0 when y^*_{mig} is less than or equal to τ_{mig}. The relationship between the underlying latent response variable for each observed item and the latent factor is then given by:

$$y^*_{mig} = \lambda_{mg}\eta_{ig} + \delta_{mig}, \qquad \delta_{mig} \sim logistic\left(0, {}^{\pi^2}/_{3}\right), \quad (2)$$

where λ_{mg} is the loading for item *m* in group *g*, η_{ig} is the value on the latent factor for individual *i* in group *g*, and δ_{mig} is the error, which is specified to have a standard logistic distribution. Equations (1) and (2) are equivalent to the more familiar logit equation:

$$logit\left(y_{mig}\right) = -\tau_{mg} + \lambda_{mg}\eta_{ig}. \quad (3)$$

The latent factor value for individual *i* in group *g*, η_{ig}, is expressed as:

$$\eta_{ig} = \alpha_g + \xi_{ig}, \qquad \xi_{ig} \sim N\left(0, \psi_g\right), \quad (4)$$

where α_g is the factor mean in group *g* and ξ_{ig} is the error (i.e., the deviation of individual *i* in group *g* from the factor mean).

The configural invariance model is the first model to be established. It is specified with the following constraints:

$$\alpha_g = 0, \forall g,$$

$$\psi_g = 1, \forall g.$$

The factor mean and variance in the each group are fixed to 0 and 1, respectively, as this is necessary for model identification and the estimation of different loadings and thresholds in each group. This is the least constrained model.

In this method, overall fit indices are examined to ensure proper model-to-data fit. The approximate fit index CFI is often used for this purpose with >.95 as the threshold (Kline, 2011, citing Hu and Bentler, 1999). The CFI is an incremental fit index that compares the fitted model with a null model and uses a chi-square (χ^2) distribution. The Root Mean Square Error of Approximation (RMSEA) is typically also examined a badness-of-fit statistic with values <.05 indicating "good fit" and values >.10 suggesting "poor fit" (Kline, 2011, citing Browne and Cuddeck, 1993). Finally, the statistical significance of the model-χ^2, which assesses absolute fit of the model to the data, is examined—however this statistic is sensitive to large sample sizes. Thus, while statistical significance of the model-χ^2 might indicate statistical misfit, but the model-to-data fit might not be practically poor. Rather than dismissing the model-χ^2 offhand, some researchers examine the residual correlations (Kline, 2011) and interpret values above the absolute value of .10 as non-negligible misfit (Higdon, 2013).

After evaluating that the configural invariance model shows evidence of acceptable fit, the factor loadings in each group are set to be equal to a reference group (lambda-equivalence):

$$\lambda_{m1} = \lambda_{m2} = \cdots = \lambda_{mG}, \quad \forall m. \tag{7}$$

The factor variances are freely estimated in all groups except the reference group ($g = 1$) and the factor means remained fixed at 0 in all groups:

$$\alpha_g = 0, \ \forall g,$$

$$\psi_g = 1.$$

This is the "metric" invariance model. The overall fit statistics and residual correlations are examined to determine model-to-data fit, and compared to those of the configural invariance model to determine the relative fit. Subsequently, the item thresholds are set to be equal to a reference group (tau-equivalence):

$$\tau_{m1} = \tau_{m2} = \cdots = \tau_{mG}, \quad \forall m. \tag{7}$$

The factor means and variances are freely estimated in all groups except the reference group ($g = 1$), where they are fixed at 0 and 1, respectively:

$$\alpha_1 = 0,$$

$$\psi_1 = 1.$$

This is the "scalar" invariance model. The overall fit statistics and residual correlations are examined to determine model-to-data fit, and compared to those of the configural and metric model to determine relative fit.

Under conditions of metric and scalar invariance, or lambda and tau equivalence, good overall fit and low residual correlations are taken as evidence of measurement invariance, and conversely substantial detriment in overall fit and larger residual correlations, in comparison to the configural invariance model, are taken as evidence of measurement non-invariance. In cases where the detriment in fit is deemed unacceptable, some loadings and intercepts may be freed from the imposed equality constraints, while the others are still held to be equal, in partial measurement invariance models. This process is guided by

modification indices, which are measures of the improvement in the overall model χ^2, should that parameter be freed (Muthén, & Muthén, 1998-2014).

In traditional Multiple-Group Confirmatory Factor Analysis (MG-CFA), the typical multiple-group factor analysis model with binary indicators is specified by the following equations. Where M represents items, n represents individuals, and G represents the groups. The response to item m (m=1, …, M) by individual i (i=1, …, n) in group g (g=1, …, G) is given by:

$$y_{mig} = \begin{cases} 1 \; if \; y^*_{mig} > \tau_{mig} \\ 0 \; if \; y^*_{mig} \leq \tau_{mig} \end{cases} \quad (1)$$

where y_{mig} is the response by individual i in group g to item m, and τ_m is the threshold of item m. In this case the responses are 0 (disagree) and 1 (agree). The responses are 1 when the unobserved latent variable, y^*_{mig}, is greater than τ_{mig} and 0 when y^*_{mig} is less than or equal to τ_{mig}.

$$y^*_{mig} = \lambda_{mg}\eta_{ig} + \delta_{mig} \qquad \delta \sim logistic\,(o, {}^{\pi^2}/_{3}) \quad (2)$$

where λ_{mg} is the loading for item m in group g, η_{ig} is the latent endogenous variable for individual i in group g, and δ_{mig} is the error, which is distributed as standard logistic. Equations (1) and (2) are equivalent to the more familiar logit equation:

$$logit(y_{mig}) = -\tau_{mg} + \lambda_{mg}\eta_{ig} \quad (3)$$

The latent endogenous variable for individual i in group g, η_{ig} is given by:

$$\eta_{ig} = \alpha_g + \xi_{ig} \quad (4)$$

Where α_g is the factor mean and ξ_{ig} is the error. The factor mean and variance are given by the below, respectively:

$$E(\eta_{ig}) = \alpha_g \quad (5)$$

$$var(\xi_{ig}) = \psi_g \quad (6)$$

The configural model is the first model to be established. It is specified as

$\alpha_g = 0$

$\psi_g = 1$

The factor mean and variance in the each group are fixed to 0 and 1, respectively, as this is necessary for model identification and the estimation of different loadings and thresholds in each group. This is the least constrained model.

In this method, overall fit indices are examined to ensure proper model-to-data fit. The approximate fit index CFI is often used for this purpose with >.95 as the threshold (Kline, 2011, citing Hu and Bentler, 1999). The CFI is an incremental fit index that compares the fitted model with a null model and uses a chi-square (χ^2) distribution. The Root Mean Square Error of Approximation (RMSEA) is typically also examined a badness-of-fit statistic with values <.05 indicating "good fit" and values >.10 suggesting "poor fit" (Kline, 2011, citing Browne and Cuddeck, 1993). Finally, the statistical significance of the model-χ^2, which assesses absolute fit of the model to the data, is examined—however this statistic is sensitive to large sample sizes. Thus, while statistical significance of the model-χ^2 might indicate statistical misfit, but the model-to-data fit might not be practically poor. Rather than dismissing the model-χ^2 offhand, some researchers examine the residual correlations (Kline, 2011) and interpret values above the absolute value of .10 as non-negligible misfit (Higdon, 2013).

After evaluating that the configural model shows evidence of acceptable fit, the factor loadings in each group are set to be equal to a reference group (lambda-equivalence):

$$\lambda_{m1} = \lambda_{m2} = \cdots = \lambda_{mG} \quad \forall_m \tag{7}$$

The factor variances are freely estimated in all groups except the reference group and the factor means are fixed at 0 in all groups:

$\alpha_1 = 0$

$\psi_g = 1$

This is the "metric" invariance model. The overall fit statistics and residual correlations are examined to determine model-to-data fit, and compared to those of the configural model to determine the relative fit. Subsequently, the item thresholds are set to be equal to a reference group (tau-equivalence):

$$\tau_{m1} = \tau_{m2} = \cdots = \tau_{mG} \quad \forall_m \tag{7}$$

The factor means and variances are freely estimated in all groups except the reference group, where they are fixed at 0 and 1, respectively, in the first group:

$\alpha_1 = 0$

$\psi_1 = 1$

This is the "scalar" invariance model. The overall fit statistics and residual correlations are examined to determine model-to-data fit, and compared to those of the configural and metric model to determine relative fit.

Under conditions of metric and scalar invariance, or lambda and tau equivalence, good overall fit and low residual correlations are taken as evidence of measurement invariance, and conversely substantial detriment in overall fit and larger residual correlations, in comparison to the configural model, are taken as evidence of measurement non-invariance. In cases where the detriment in fit is deemed unacceptable, some loadings and intercepts may be freed from the imposed equality constraints, while the others are still held to be equal, in partial measurement invariance models. This process is guided by

modification indices, which are measures of the improvement in the overall model χ^2, should that parameter be freed (Muthén, & Muthén, 1998-2014).

It should be noted that, rather than freeing non-invariant parameters from equality constraints entirely, these parameters could be set to be equal in a smaller subset of groups. For example, if the first item loadings in Poland and Bulgaria are similar to each other but different from the rest, the loadings in Poland and Bulgaria can be constrained to be equal to each other yet allowed to be different from the other countries. This example scenario would create two subsets of countries with their own equality constraints. However, this approach is not frequently taken by researchers, likely because the modification indices that are used to guide the freeing of the parameters do not include information about invariance between subsets of groups. Rather, non-invariant parameters are typically freed from the rest in the traditional approach.

Approximate Measurement Invariance analysis in the Bayesian framework (Muthén & Asparouhov, 2013), shares many similarities with the traditional methods described in detail above. The configural model is also established by fixing the factor mean to zero and the factor variance to one all of the groups to allow for model identification, and the item loadings and thresholds are estimated freely. This configural model is also considered to be the model with the best possible fit, because no equality constraints are imposed. Subsequently, equality constraints are imposed between loadings where the factor variances are freed in all but the first group in a metric model. And further equality constraints between thresholds with the factor means and variances freed in all but the first group in a scalar model. Similarly, detriment in fit in these metric in scalar models, in comparison to the configural model, is taken as evidence of measurement non-invariance. Finally, loadings and thresholds which are very different can be released from their equality constraints, which is

called Partial Approximate Invariance (van de Schoot, Klutmans, Tummers, Lugtig, Hox, & Muthén, 2013). In these ways, Approximate Measurement Invariance in the Bayesean framework is not a very large step from traditional methods.

However, an important and extremely useful difference between traditional MG-CFA methods and Approximate Measurement Invariance using Bayesian estimation is the ability to only allow small differences loadings or thresholds, through the use of zero-mean, small-variance model priors, which have been established in ground-breaking work in this area (van de Schoot et al, 2013; Muthén & Asparouhov, 2012). This is a middle-ground between allowing a parameter to be completely different from the others, versus allowing only a small amount of difference or "wiggle room." Model priors bring the estimates of the parameters that are allowed to be different close to the group mean, without constraining them to be exactly equal. This is a major advantage, and is demonstrated and described in greater detail in the results section, below.

The principle of statistical inference in Bayesian estimation is very different from the frequentist estimation techniques. In Bayesian estimation, the parameters are considered to be random and with a probability distribution that reflects uncertainty around the true value, while in frequentist estimation the parameters are considered to be fixed but unknown (Kaplan & Depaoli, 2012; Kaplan & Depaoli, 2013). In Bayesian estimation, our prior knowledge, formalized by a prior distribution, is weighted by the observed data, or the likelihood function, resulting in an updated estimation, known as the posterior distribution (Muthén & Asparouhov, 2012). Parameter estimates in the Bayesian framework are point estimates from the posterior distribution, typically the mode, median, or mean. Credibility intervals provide the analog to standard errors and p-values in the frequentist tradition. The

95% credibility intervals are formed in a straightforward manner by taking quintiles of the posterior distribution (Kaplan & Depaoli, 2012; Kaplan & Depaoli, 2013). Thus, credibility intervals provide the upper and lower bounds that comprise 95% of the posterior distribution and can be interpreted directly as a 95% probability that the parameter lies within the upper and lower limits of this range. Bayesian posterior distributions and credibility intervals are presented in the results that follow, which will illustrate these important distinctions.

In order to form the posterior distribution, Bayesian analysis is conducted using Markov Chain Monte Carlo (MCMC) algorithms (Asparouhov & Muthén, 2010), which simulates the posterior distribution through many, sometimes many thousand, independent repeated draws of possible parameters from the observed data. MCMC algorithms use multiple chains—in M*plus* the default is two chains however this can be increased up to the total number of processors available for the analysis—as it repeatedly draws parameter estimates, which results in a distribution. As the repeated draws are carried out, the distance between the chains diminish, and once the draws are stabilized, the previous draws are thrown out (the burn-in phase). Thus, in addition to examining model fit, which is important in frequentist estimation as well as Bayesian estimation, evidence of model convergence must also be examined in Bayesian estimation, to ensure that the estimates obtained in the simulation are drawn from the proper distribution.

To evaluate model convergence, I examined convergence trace plots, posterior density plots, and autocorrelation plots for each loading and threshold in each group, as well as the Potential Scale Reduction value, or PSR, for the entire model (Asparouhov & Muthén, 2010; Muthén, & Asparouhov 2012). Convergence trace plots visualize the progress of the

chains across the repeated draws and should form a tight band around the parameter estimate. Models that do not converge show gaps between the chains which range across values of the parameter estimate, rather than the chains proceeding in a tight band together. When the chains used do not converge at once, but do converge during the burn-in phase and show a tight band after the burn-in phase, this would still be considered evidence of convergence. Relatedly, the PSR quantifies the within-chain variation versus the between-chain variation. PSR values very close to 1 and below 1.1 indicate that there is very little between-chain variation, relative to within-chain variation, and is taken as evidence of model convergence. Posterior density plots should be relatively peaked and unimodal to ensure that the point estimate is more probable than other values (Zyphur & Oswald, 2013). Finally, the repeated draws should be independent, and the autocorrelation plots quantify the extent to which repeated draws are correlated. Every 30^{th} draw should have a correlation below .20 or .30 (Asparouhov, 2012).

To evaluate model fit, I used posterior predictive checking (PPC), which uses a likelihood ratio χ^2 test to examine discrepancy between the observed data and model-generated data (Kaplan & Depaoli, 2012; Kaplan & Depaoli, 2013; van de Schoot, Kaplan, Denissen, Asendorpf, Neyer & van Aken, 2013) and estimates a posterior predictive p-value (PPP) and 95% confidence interval. This is a measure of model accuracy relative to the data. PPC uses a test-statistic and compares that test statistic to the observed data across the repeated draws, and the PPP value is the proportion of draws with positive test statistics relative to the data versus negative test statistics. Low PPP is evaluated as poor fit and corresponds to models with predominantly positive or negative test statistics, and is an indication of model misspecification. A model with excellent fit will have a PPP around .50,

which indicates an equal proportion of positive versus negative test statistics, as well as confidence intervals with a negative lower bound and a positive upper bound. At the time of this writing, there were no theoretical lower limits of the PPP value corresponding to bad versus good fit (Muthén & Asparouhov, 2011; Muthén & Asparouhov, 2012; Lee, 2013). However, given the theory behind the PPP, that lower PPP values might indicate misspecification and that values around .50 indicate high levels of accuracy, and given the critical nature of accurate measurement models in order to answer substantive research questions in subsequent analyses, in this analysis I considered PPP values lower than .40 and higher than .60 to indicate unacceptable model fit.

In order to be considered in this analysis, all models had to converge as a threshold expectation. After establishing convergence, evidence of model fit was examined as well. In each of the three factors, I followed a sequence in the model-building process, examining the model fit in every instance to inform the next step. First, I fit a model with full metric and scalar invariance in order to determine whether a full measurement invariance model was a good fit to the data. Next, I fit a configural model in order to obtain the least constrained model, which served as the comparison point moving forward with metric and scalar invariance models. I then fit a metric model with loadings that were not statistically significantly different from each other constrained to be equal. In this process, I did not free every parameter that was different from the set completely. Rather I freed those that were not equivalent to any groups, and constrained to be equal others that were not equal to the full set but were equal to a smaller subset of groups. I repeated the process with the item thresholds. Finally, I applied zero-mean and increasingly smaller variance priors to the parameters that were allowed to be different, to bring those parameters as close to the mean

as possible, without compromising fit. This taxonomy of models will be discussed in detail in the findings section, below.

Findings

Model convergence

The purpose of this analysis was to test explicitly for measurement invariance, and when that test failed, to clearly identify the sources and the degree of non-invariance. In order to present the findings of this analysis clearly and concisely, I will go into greatest detail in one of the factors—attitudes toward racial and ethnic minorities (RETH)—in this section. However, I followed the same steps to evaluate the evidence of measurement invariance in the other two factors (PROT and IMMIG), and those results are presented in brief. Overall, using Partial Approximate Invariance, it was possible to form measurement models, with adequate degrees of measurement invariance and overall model fit, in all three factors, sufficient to allow the comparison between the groups in this analysis at the factor level.

For all of the measurement models, it was always necessary as a first step to establish that each model properly converged, and I did not consider any models that did not converge properly. As presented above in the Methods section, to evaluate model convergence, I examined convergence trace plots, posterior density plots, and autocorrelation plots, for each loading and threshold in each group, and the Potential Scale Reduction value, or PSR, for each model. When the models in this analysis failed to converge, this appeared to be related to non-independence of the draws, evidenced most

clearly in the autocorrelation plots. In cases like this, thinning is recommended (Muthén & Asparouhov, 2012), in which M*plus* will retain only the draws by a specified number of iterations. For example, if thinning at 100 is used, only every 100th draw is retained, and this increases the prospect that the repeated draws are independent from each other. In every case, increasing the thinning, sometimes by more than 500, led to model convergence in this analysis. In these cases, I also increased the number of iterations and the number of chains that were used, to allow a longer opportunity for the model to converge properly, and then re-examined evidence of model convergence. All of the models presented here properly converged. For all of these models, four chains were used, as well as thinning at levels between 500-750 and iterations between 400-2000.

In all of the models presented here, the final PSR was well below 1.10. Indeed the final PSR in all of the models ranged between 1.001-1.009, indicating very little variation between chains versus within chains, which is strong evidence of model convergence. I also examined the trace plots, posterior density plots, and autocorrelation plots for each loading and threshold in each group, for each model. Given that each model included seven groups (six European countries and England) and three (PROT) or four (RETH and IMMIG) items, there were far too many loadings and thresholds in each factor (56 in RETH and IMMIG and 42 in PROT configural models) to display each parameter's trace plot, density plot, and autocorrelation plot here. However, to illustrate, I've included a panel of figures showing the trace plot, posterior density plot, and autocorrelation plot for the first item loading in the first group (Bulgaria) for the RETH configural model in Figure 1.2.

INSERT FIGURE 1.2 ABOUT HERE

Figure 1.2 provides evidence of convergence with regards to the first item loading in the first group. The trace plot shows that four chains were used as each chain is shown in a different color, and all chains form a tight band around the point estimate. From this plot, the burn-in phase is also shown with the vertical red line in the middle of the plot, and only the 1000-2000th iterations were used to form the posterior distribution and estimate the parameter. The posterior density plot in Figure 1.2 shows that a range of parameter estimates were sampled, from about 1.35 to about 2.55, and that the sampling distribution was relatively unimodal and peaked around a mean and median of about 1.8. I allowed M*plus* to use the median for the parameter estimation by default, thus the parameter estimate for the loading in the first item in the first group was 1.78, which is shown in this figure. Finally, the autocorrelation plot shows that the draws are reasonably independent from each other, with all of the correlations between draws less than |.10|. I included this figure to provide an example of the process used to evaluate model convergence, however it must suffice to say that I examined all of the graphs for the proper attributes discussed above, and that all of the final models that were used in this analysis showed extremely strong evidence of model convergence.

Model taxonomy and model fit

For each of the three factors, the full invariance models, with both loadings and thresholds constrained to be equal across all groups, showed extremely poor model-to-data fit and thus lacked evidence of full measurement invariance. In each case, under full invariance, the overall PPP was <0.001, far from .50 which would indicate excellent model-to-data fit. These overall PPP values are given in Table 1.3, as Model 1 in the first column. In this column, the confidence intervals are also given, and in all three factors all of the test

statistics are all positive, shown by the positive value of the lower bound, and substantially higher than expected.

INSERT TABLE 1.3 ABOUT HERE

The PPP values for each group are given in Tables A1.2-A1.4, in the Appendix. In these tables the first column contains the overall and group-specific PPP values under the condition of full invariance. In nearly all of the groups, in all three factors, the PPP value is too small to be taken as evidence of good model-to-data fit. Therefore, the models with full invariance imposed were rejected in all three factors. It should be noted, that had the assumption been made that full invariance held, without conducting this test to evaluate this assumption, then there would be substantial bias in the parameter estimates, given the strong evidence of model misspecification summarized here.

The next task was therefore to diagnose the source of the non-invariance in each factor by fitting metric and scalar models and comparing them to the configural model described above. The overall PPP values for the configural models, given in Table 1.3, ranged from .495 to .534. Thus, the next task was to establish equality constraints in a series of subsequent models without compromising this fit. In order to describe this process, I will focus on the RETH factor only.

The goal of the next phase of the analysis was to establish equality constraints on as many of the loadings as possible without compromising the fit. While establishing equality constraints on loadings that are very different would compromise the fit, fully releasing these parameters from any equality constraint, when an equality constraint within a subset of groups is possible, might also be a case of freeing too much. Therefore a useful first step was

to estimate the differences in the loadings between each pair of groups, to get the best information regarding where equality constraints might be possible. This is possible in the M*plus* software by creating new parameters for the differences in the loadings for each pair of groups, and requesting that M*plus* estimate those new parameters. This generates an estimate of the difference in the loadings between each pair, as well as the statistical significance of each difference, indicated by a credibility interval that contained 0. Loadings between groups where the credibility interval contained 0 were constrained to be equal in this step. For the RETH factor, the differences in the loadings for each pair of groups are given in Table A1.5, in the Appendix. In this table, the new parameter name is given in the first column. This name contains the both loading and the two groups being compared. For example LAM1_1v2 indicates the difference in the first loading (LAM1) between groups one (Bulgaria) and two (Greece) (_1v2), and so on. The estimate of the difference is given in the next column, and statistically significant differences are indicated with an asterisk. Further, the plots of the item loadings, shown in the first column of Figure 1.3, discussed below, were also helpful to visualize the loadings in each group that were similar to each other.

In the RETH factor, all of the loadings for the first item had no statistically significant differences between pairs of groups. Therefore, the loadings in the first item were all constrained to be equal. The other three items had different numbers of pairs that were statistically significantly different from each other: the second item had 6/21 non-invariant pairs of groups, the fourth item had 7/21, and the third item had 14/21. For the second item, the loading in group 5 (Sweden) was statistically significantly different from all of the other groups. Therefore, the loadings in all of the groups except for Sweden were set to be equal in this item. However, in the third and fourth items, while there were statistically

significant differences between some pairs of groups, no groups needed to be freed from the others entirely. Indeed, in the third item, it was possible to create three subsets of equality constraints among countries that were similar to each other to a sufficient degree: between Bulgaria, Greece and Switzerland; between Poland and Spain; and between England and Switzerland. And in the fourth item, it was possible to create two subsets of equality constraints: between Bulgaria, Greece, Poland, Spain, and Switzerland; and between Sweden and England.

This information was used to form a measurement model in the RETH factor that specified these equality constraints and sub-groups of equality constraints. The overall and group-specific PPP values are given in Table A1.2, in the second column, as Model 3A: Partial Metric. This model is called a Partial Metric model because some, but not all, of the loadings were constrained to be equal. As discussed above, the model-to-data fit should be as good as in the configural model, and indeed I found that it was. The overall PPP for this model was .520 and the PPP values in each group ranged from .480-.559. To further illustrate, Figure 1.3 shows the freely-estimated factor loadings of the configural model in the left-hand column compared to the partially-constrained factor loadings of the partial metric model in the middle column. In the middle column some, but not all, of the loadings have been brought into alignment with each other to an extent that is highly visible in this figure.

INSERT FIGURE 1.3 ABOUT HERE

Having established good model-to-data fit in the partial metric model, I followed the same procedure described above to further impose equality constraints in the thresholds across groups and within subsets of groups to establish a partial scalar model. There were more statistically significant differences in the thresholds, compared to the loadings

described above. Still, for the first item, there were two subsets of groups where equality constraints were possible: between Bulgaria, Greece, Switzerland and England; and between Poland, Spain, and Sweden. In the second item, there were similar subsets that were possible, with the exception of Sweden which was not similar enough to Spain and Poland to enable those equality constraints. For the third item, only Poland and England were similar enough to enable equality constraints, and the differences between the other groups necessitated that they were completely freed. And finally in the last item, while necessary to free Sweden and Greece, two subsets of equality constraints were possible among the other groups: between Spain, Poland and England; and between Switzerland and Bulgaria. The model-to-data fit for this model, shown in Table A1.2 as in the fourth column under Model 3B: Partial Scalar. The overall PPP for this model was .529 and the PPP values in each group ranged from .466-.564, which indicated excellent fit. The group thresholds in each item under configural invariance and partial scalar invariance are shown in Figure 1.4 in the first and second columns, respectively. Again, while the thresholds were freely estimated in the configural model, the partial scalar model brings some of these thresholds into alignment.

INSERT FIGURE 1.4 ABOUT HERE

The connections between the groups, at both the loading- and threshold levels, are visualized in Figure 1.5, using a social network package available in the statistical software R (R Foundation for Statistical Computing, 2015). In this figure, each of the seven groups are shown as nodes in a network, and the equality constraints that were specified are shown by lines connecting the groups. From this figure, it is shown that, in the first item in the upper-left corner, the loadings in every group were constrained to be equal to every other group. This is also shown in the second item, however with the exception of Sweden which was

allowed to be different. The third and fourth items have fewer equalities in the factor loadings across the groups. The larger network in the lower-half of the figure shows all of the equalities across all four loadings combined. The number of equality constraints between groups is shown by the weight of the line, and the size of the nodes indicate the strength of that group's equality constraints with all of the others. This part of the figure shows that each group is connected to every other group with at least one equality constraint. Sweden showed the lowest level of equality constraints at the loading level.

INSERT FIGURE 1.5 ABOUT HERE

Figure 1.6 shows the same visualizations regarding the equality constraints between the thresholds. Across each item, there were fewer equality constraints that were possible among the thresholds, compared to the loadings, which is shown in the figure as fewer lines connecting the nodes of the groups. In the first item, the most equality constraints among thresholds across groups were possible, with two subsets highly visible. Again, the subsets of equality constraints between the thresholds in the first and second items are highly similar, with the exception of Sweden being freed in the second item. Finally, there were only two groups with equality constraints in the third item, and only a subset of two and a subset of three groups in the fourth item. However, the combined visualization of all of the equality constraints across the thresholds in the bottom-half of the figure was somewhat more promising. In this figure, there are equality constraints shown between most of the groups in this analysis. There are two fairly clear subsets of equality constraints between Poland, Spain, and Sweden; and between Switzerland, Greece, Bulgaria, and England; and some equality constraints connect these two subsets between England, Spain and Poland. As with the intercepts, Sweden had the weakest level of equality constraints with the other groups.

The Bayesian framework made it further possible to bring even the differences in those parameters that were allowed to be different as close to the mean as possible, with zero-mean, small-variance priors. This is especially important given Sweden's weaker connections with the other groups in both the loadings and the thresholds, and given that not all groups were connected via at least one equality constraint among the thresholds. Thus, to the greatest extent possible, bringing these differences closer together allows for some level of alignment among these groups, even in the absence of equality. This is a highly valuable benefit of the Bayesian framework.

In the next two models, I applied priors with increasingly smaller variances to the extent that I was able to do so while maintaining good model-to-data fit. As a first step, I applied model priors with a mean of zero and increasingly smaller variance to the loadings, and found that I was able to decrease the variance to .01 without substantial detriment to fit (~N(0, .01)). The model-to-data fit for this model is shown in Table A1.2 in the fifth column, under Model 4A, with all PPP values showing evidence of excellent model fit. As a second step, I retained this model prior for the loadings, and applied additional model priors with a mean of zero and increasingly smaller variance to the thresholds as well. Larger-variance for the priors on the thresholds was necessary to retain adequate model-to-data fit, which was not surprising given the larger number of groups or subsets that were not constrained to be equal to the others. In this case, the smallest variance possible without unacceptable detriment to model fit was .10. However, this prior still brings those thresholds that were allowed to be different very close to the mean, while maintaining fit.

The overall PPP value and the PPP in each group for this model are shown in Table A1.2 in the last column, under Model 4B. The overall PPP was .445, and the PPP values in

the groups ranged between .427 and .512. This is the model with the maximum number of equality constraints, and with priors bringing those parameter estimates allowed to be different as close to the mean as possible without detriment to fit.

The success of the use of model priors is shown in Figures 1.3 and 1.4. In these figures, the estimates of the parameters using the model priors are displayed in the third column, Figure 1.3 shows the loadings and Figure 1.4 shows the thresholds. Examining the loadings in Figure 1.3, it is clear that the metric model brought some parameters into alignment with equality constraints, and then the use of model priors brought the remaining parameters very close to alignment. This is an important advantage of the Bayesian framework. Thus, the use of priors brought these estimates very close to equality, while still allowing very small differences, and without detriment to the model fit. Therefore, this model was retained as the final measurement model for the RETH factor, and there is strong support for the confidence that this model will yield highly accurate factor scores and very little bias in the parameter estimates in subsequent analyses.

This procedure described above was followed to establish the final measurement models in the other two factors (PROT and IMMIG) as well. The model priors and evidence of model fit are presented in Table A1.3 and A1.4 in the Appendix, under the corresponding columns labeled Model 4b. In all cases, with the equality constraints and the model priors that were used, the models showed strong evidence of model fit. Further, the same visualizations that were discussed above with regard to the RETH factor are given for the other two factors, IMMIG and PROT, in Figures A1.1-1.8. Specifically, Figures A1.1-1.4 show the groups being brought into alignment across the process of establishing some equality constraints and then using model priors. Figures A1.5-1.8 show the loadings and

thresholds that are constrained to be equal. Similar to the RETH model, more equality constraints were possible among the loadings than the thresholds. In the IMMIG model, Sweden again stood out as being very different from the others, while it was found to be more similar to the others in the PROT factor.

Factor scores

Factor scores were obtained for each of the final measurement model using Bayesian imputation of plausible values. In each factor, a set of 10 plausible values was generated for each adolescent, and the mean of this set of plausible values was retained as the point estimate of the factor score. Because these scores are all on a different metric, I further standardized the scores to aid in their interpretation, including the sampling weights. The distributions of these scores are given in Figure 1.7, with the median in each country marked with a diamond. The distributions of the scores were fairly consistent, with the exception of Sweden which had a wider total range in both the IMMIG and the RETH factors.

Overall, a MANOVA analysis (Multi-variate Analysis of Variance) confirmed that differences in the levels of intercultural attitudes exist, meaning that I was able to reject the null hypothesis that the means were the same across countries. This was found in all three factors: IMMIG (Wilks' lambda=0.98; Pillai's trace=.02; *df* 6, 16840; F=50.11, p <.001), RETH (Wilks' lambda=0.95; Pillai's trace=.05; *df* 6, 16840; F=156.81, p <.001), and PROT (Wilks' lambda=0.96; Pillai's trace=.04; *df* 6, 16840; F=104.14, p <.001). After rejecting the null hypothesis in all three factors, I then tested for statistical significance of the mean differences by country, using the Šidák correction for multiple-comparisons (21 comparisons, p<.002). Mean differences that were not statistically-significant are boxed and shaded in Table 1.4.

INSERT TABLE 1.4 ABOUT HERE

Adolescents in Sweden, Poland, Bulgaria, and Greece had the highest levels of positive attitudes toward immigrants, ranging from .110-.074, however the differences between these countries were not statistically significant. This means that there is not sufficient evidence to rank Sweden above the other countries, however the entire set of these four countries are ranked higher than the others. Spain, England, and Switzerland had lower levels of these attitudes, and were statistically significantly different from each other as well as the set of countries at the top of the ranking.

Regarding attitudes toward racial and ethnic minorities, Sweden (.726) and Poland (.161) had the highest levels of positive attitudes, both of which are statistically significantly different from each other and from the rest. In fact, Sweden is quite higher than the others in this metric. Figure 1.7 shows the range of scores in Sweden in this factor is wide, however the interquartile range, represented by the blue shaded rectangle, was above the others nearly in its entirety. In this area, attitudes in England and Spain were similar to each other, as were attitudes in Switzerland and Greece. Bulgaria had the lowest levels of positive attitudes toward racial and ethnic minorities (-.376).

Lastly, England and Switzerland had similarly high levels of protectionist attitudes toward migration (.300 and .272, respectively). These two countries were statistically significantly different from the next set: Greece, Sweden, and Spain (these ranged from .037 to -.028), which were in the middle of the set. Finally, Bulgaria and Poland had the lowest levels of protectionist attitudes (-.110 and -.213, respectively).

The correlations of the factor scores in each country are shown in Table 1.6. As expected, there were moderate to weak negative correlations found between protectionist attitudes (PROT) and positive attitudes toward racial and ethnic minorities (RETH) and immigrants (IMMIG). Also as expected, there were moderate to strong correlations between attitudes toward racial and ethnic minorities (RETH) and immigrants (IMMIG). As discussed previously, attitudes toward these different groups could be conflated in the understanding of adolescents, because immigrants are at times also racial and ethnic minorities. These correlations were lowest in the Eastern European countries, Poland and Bulgaria (r=.36 and r=.32, respectively). This correlation was highest in Greece (r=.55).

These factor scores were obtained through the Partial Approximate Measurement Invariance method described in detail above. However, as a point of comparison, I also obtained the factor scores from models where full invariance was applied in order to demonstrate the differences. These factor scores are summarized in Table 1.5. There are several notable differences. Under full invariance, there is a wider spread of average standardized factors in each country, which also resulted in more statistically significant differences found by country. This result is due to the specification of equality constraints across all loadings and thresholds, because incorrectly holding these parameters to be equal forced the estimated factor scores to include the non-invariance of the measures. These results would lead researchers to believe that there were more differences between the countries than in fact there were.

Further, the rankings of the countries are different under the condition of full invariance, and some countries shifted in rank more than others. Poland, for example, remained nearly in the same rank position along the factors. However, the ranking of

Sweden shifted the most under conditions of full invariance, particularly along the IMMIG factor. In this factor, Sweden was ranked first with no statistically significant difference with the second country, Poland, under the Partial Approximate invariance specification, but under full invariance Sweden is ranked right in the middle of the countries included, with statistically significant differences between Sweden and Greece, ranked above, and Spain, ranked below. It is notable that Sweden also showed the least evidence of measurement non-invariance out of the countries in this set, and so it is not surprising that Sweden's ranking shifted the most between the two models. Forcing the equality constraints in Sweden especially would result in misleading results both in terms of ranks and statistical significance of the differences.

INSERT TABLE 1.5 ABOUT HERE

Limitations

There are several limitations inherent in this analysis. Primarily, the methods used in this paper are extremely new, which meant that there were limited resources available to help judge the steps in the process and evaluate effectiveness. This paper will make an important contribution to these new methods in that it provides a highly detailed account of their use. Also, only seven countries were used in this study, thus my inferences are limited to those countries and should not be generalized to Europe as a whole.

Discussion

One purpose of this analysis was to establish models that have sufficient levels of measurement invariance to allow valid cross-group comparisons in order to answer the

research questions of this dissertation. Thus, in addition to an examination of the validity of the items being used, in this paper I explicitly examined evidence of measurement invariance, and properly handled the measurement non-invariance that was found, using the new method of Partial Approximate Measurement Invariance (Muthén & Asparouhov, 2013). Details were included to provide applied researchers with an example of measurement invariance analysis to draw from in their own analysis. Ultimately, final measurement models were established in all three factors that allow comparison between groups, and produced factor scores. These scores can be used in subsequent analyses, and can inform test development and item writing in this area, discussed next.

In the Bayesian framework, it is necessary to examine both model convergence and model fit, in order to be confident that the parameter estimates are drawn from the proper posterior distributions and can therefore be trusted to provide the correct information. In all cases in this analysis, the models converged, however that was often after a high level of thinning and a lengthy run of iterations. With four processors, most of these models took 12-18 hours to run. Given the number of models that were fit at the various stages of this analysis, the time investment is significant. However, as demonstrated above, the ability to use model priors to bring even the parameters that were allowed to be different into approximate alignment is highly valuable and well worth the investment of time.

The specific non-invariance findings are helpful to explore. In all of the factors, more evidence was found of invariance at the metric level, or regarding the loadings, compared to the scalar level, or regarding the thresholds. Recall that invariance at the metric level indicates that the associations between the items and the underlying factor are equivalent in each group, and that invariance at the scalar level indicates that the thresholds

of the items are equivalent in each group. The loadings provide information about the degree to which each item discriminated between positive and negative attitudes. The thresholds provide information about the level of the underlying factor when the probability of agreeing with the item first becomes greater than the probability of disagreeing. This means that, at a given level of intercultural attitudes, these items were more difficult for some adolescents than others. More evidence of measurement invariance at the metric level, but less at the scalar level, indicates that these items discriminated somewhat similarly across the countries but were perhaps confusing to some adolescents. This was true across all factors, but particularly within the PROT factor, which required the largest allowed variance in the model priors (~N(0, .50)). These items were complex, and this analysis suggests that complex questions may reduce the measurement invariance of items, making cross-cultural comparison very difficult. It follows that reducing complexity in items used for cross-cultural comparison is extremely important.

In the RETH factor, the first item had the most evidence of metric and scalar invariance while the third item had the least. This means that there were small differences in how well the item "all<ethnic/racial groups> should have an equal chance to get good education in <country of test>" discriminated between adolescents who had overall positive views versus overall negative views. There were also small differences in how likely it was that adolescents with overall positive attitudes would agree with this item. This is perhaps unsurprising because education is most directly related to the everyday lived experiences of young people. However there were greater differences in how well the item "<members of all ethnic/racial groups> should be encouraged to run in elections for political office" discriminated between adolescents with overall positive versus overall negative views, the

loadings in each country in this item were lower than the other items. And relatedly, there were greater differences in the probability of agreeing with this item among adolescents with overall positive views. This could have been due to the fact that the item about elections is much less directly related to the lived experience of young people than the item related to education, making this item harder to answer. These findings suggest that items that are more closely related to the lived experiences of adolescents are superior in measurement characteristics than those that are less closely related and more abstract.

In the IMMIG factor, there is a similar finding. The second item ("[i]mmigrant children should have the same opportunities for education that other children in the country have") has the most evidence of invariance, especially with regard to the loadings. Further, this item also has the highest loadings, meaning that it is associated most strongly with the underlying factor across the countries. Again, this item has to do with education, which is very close to the adolescent experience. The first item ("Immigrants should have the opportunity to continue speaking their own language") has weaker evidence of non-invariance at the metric level and the associations between this item and the latent factor are lower than the second item for all countries. Language use is extremely close to the experience of all people, however it is also highly controversial in some contexts. Dominant language use, or not, is one of the principal signifiers cultural group membership. It may be the case that this item also measures a willingness, or not, to engage in a controversial issue.

The differences in the factor scores, which are apparent after establishing valid models for cross-cultural comparative use, show that there are differences as well as similarities in intercultural attitudes across these countries. The differences themselves suggest that these countries, which are within the same geographic region, and have been

engaged in political cooperation and the promotion of inclusive values, still have important differences in history, economic climate, and culture that may have manifested in differences in intercultural attitudes among adolescents.

Switzerland ranked lowest in attitudes toward immigrants and this could be related to the translation of "immigrant" into German as "foreigner" (Torney-Purta et al, 2001). Naming a group as a foreigner may have established an outsider relationship between the native-born adolescent and immigrants. Sweden ranked in the top set regarding attitudes toward immigrants and at the top regarding attitudes toward racial and ethnic minorities. An in depth analysis of the context of Sweden would be necessary to adequately explore the reasons for these differences, which is outside of the scope of this paper. However, it is notable that in 2008, close to when this study took place, Sweden made significant legislative changes to prevent discrimination against minority groups with a national-level Discrimination Act and the creation of the position of Equality Ombudsman to oversee compliance with this Act (Ministry of Integration and Gender Equality, 2009). A particular focus of this Act is schools, with specific prohibitions against discrimination against any child on the basis racial and ethnic minority group status. While this study does not seek to prove this connection, it is very possibly the case that this new Act was in the news near the time that the adolescents were responding to this questionnaire. This was controversial and could have had positive or negative effects on the young people, in interaction with their peer groups and families, where some young people may have embraced this policy and some may have resented it.

Differences in attitudes toward racial and ethnic minorities in Poland and Bulgaria, the Eastern European countries in this study, are also interesting to note as they are on

nearly the opposite ends of this scale. Poland was ranked the second-highest in attitudes toward racial and ethnic minorities, however it also has the most homogenous population within this study. Bulgaria's ranking on this measure was the lowest in the study, which may reflect the stigmatized position of the Roma in the Bulgarian context. Taken together, this is evidence that countries in Eastern Europe are very different from each other, and should not be treated as a block of countries in cross-cultural comparative work that examines attitudes toward racial and ethnic minorities.

In conclusion, the methods used here proved extremely valuable to ensure that the measures being used showed high levels of measurement invariance, necessary for valid cross-cultural comparison. The methods described in detail here show researchers how to improve scales to support valid cross-cultural comparison. Not only does this analysis provide unbiased parameter estimates for subsequent analyses, it is also of substantive interest in itself. Being able to confidently compare groups along the mean standardized factor scores is highly useful in understanding the differences in intergroup attitudes in each country and toward various groups, deepening our understanding of the intergroup attitudes that these adolescents had at the time of this study. These findings show that there are group-level differences in intergroup attitudes that warrant their further examination. This major finding informed the following analysis, in Paper 2, but hopefully will further inform ongoing research into intergroup attitudes across various national contexts.

Paper 2: "Modeling intercultural attitudes among adolescents across Europe: A multi-level, multiple-group analysis examining student attitudes, intergroup contact, and school climate"

Introduction

Intercultural conflict is an area of serious international concern, and this is particularly true in Europe where freedom of movement is a fundamental right afforded to EU citizens (Europa, EU Charter of Fundamental Rights, 2000). Further, the political power and relative economic stability of the EU has greatly increased its attractiveness as a migration destination. Thus, its formation has brought increased immigration flows to Europe (European Commission, 2009), with nearly 1.7 million immigrants settling in Europe in 2012 (European Commission, 2015). These migration and immigration flows have brought groups of a variety of national, cultural, language, and religious backgrounds together across Europe, in contexts which are different historically, politically, and socially.

Intergroup contact raises challenges due to its association at times with intergroup conflict and hostility, which can lead to exclusion, and vice versa. However, democratic societies rely on political participation (including but not limited to voting) and social inclusion. The inclusion of all groups is not secondary, but rather central, to the proper functioning of democratic societies. Attitudes toward minority groups are related to the participation of minority groups in civic life: as either conduits, or barriers, of integration into civic life. For example, Khanec and Tosun (2009) found that the perception of negative attitudes toward immigrants in Germany discouraged foreign residents from civic

participation and this effect was stronger among those who were active in the labor market and have more years of schooling.

In this paper, I examine the potential role of schooling in the intergroup attitudes that young people in Europe had in 2009, using data from the large-scale 2009 IEA (International Association for the Evaluation of Educational Achievement) International Civic and Citizenship Education Study (ICCS) (IEA, 2009; Schulz., Ainley, Fraillon, Kerr, & Losito, 2009). I define intergroup attitudes as attitudes that members of a group have, on average or individually, toward other cultural groups, on the basis of both their own various intersecting group memberships as well as those of others. I use the term intergroup, as opposed to intercultural, however I also emphasize the central concept of culture in understanding groups, described in detail below.

There have been several related areas of research in social and developmental psychology since the 1970s that have sought to understand intercultural attitudes. Intercultural processes (within social psychology) have been an active area of research, especially group formation and intergroup relations, such as in-group formation (see Tajfel, & Turner, 1979; Brown, 1995/2010) and out-group derogation (see Brewer, 2007; Corenblum & Stephan, 2001; Bennett, Barrett, Karakozov, Kipiani, Lyons, Pavlenko, & Riazanova, 2004). It is further essential that these approaches make sources of difference explicit in sufficient detail, rather than assuming that the development of intergroup attitudes is the same across contexts. In the field of developmental psychology, specifically cultural context in human developmental psychology, cultural forces that shape human development have been explored in greater depth, often making sources of difference explicit and detailed

(Torney-Purta & Amadeo, 2011; Torney-Purta & Barber, 2011). However, these models have not focused on group processes in the development of intergroup attitudes.

Integrating research on differences in cultural context as they relate to intercultural attitudes would foster greater understanding of the influence of culture in this process, and this integration is discussed in the Introduction. Not enough work has been done to integrate these approaches by utilizing organizing models from a comparative perspective. Thus, this study builds from this integration to fill current gaps in knowledge by focusing on the development of intercultural attitudes among young people in cross-cultural, comparative perspective. It uses an integrated theoretical developmental model that is helpful to make explicit the role of national and local contexts, daily interactions, and adult beliefs in the development of intercultural attitudes among adolescents, specifically within schools (Torney-Purta. & Amadeo, 2011; Torney-Purta, & Barber, 2011). It further uses a cultural model that is helpful to understand the role of cultural tools and narrative in making sense of the self, versus the other (Haste & Abrahams, 2008).

Central to this work is a consideration of culture as organizing the developmental environment of young people (Super & Harkness, 2002; see also Van de Vijver & Poortinga, 2002). The developmental niche as presented by Torney-Purta and colleagues (2011) is a highly useful overarching framework to incorporate national-level forces, as well as the lived experiences of young people in schools and attitudes of adults. In addition to large cultural systems at the national level, it is understood in this analysis that schools both operate within the larger culture and have cultures of their own. Further, schools as communities of practice (see Lave and Wenger, 2002), in which the goal is to foster the full participation of young people in society, are seen as having an important role to play in the development of

positive intercultural attitudes when the school community is both positive and democratic. Schools are an important site where the norms and values of society are practiced and transmitted, both explicitly and implicitly, as well as sites where these norms are anticipated and even challenged. Finally, the cultural model (Haste & Abrahams, 2008) provides a system for understanding the role of cultural tools and narratives in the development of intergroup attitudes.

Policy makers and educators have some control over schools, thus schools have the potential to be an lever for reducing intolerance. Given the developmental nature of intergroup attitudes, a continued focus on young people could support successful interventions that promote positive intergroup attitudes, as these attitudes begin to develop from a young age (Aboud, Tredoux, Tropp, Brown, Niens & Noor, 2012; Raabe & Beelman, 2011). Within this frame, I explored the association of intergroup attitudes toward various groups—immigrants, racial and ethnic minorities, and migrants—with young people's experience of positive and democratic climates at schools.

Empirically examining intercultural attitudes cross-culturally

In this study, I conducted a secondary analysis of the IEA (2009) civic education data. I focused on attitudes of native-born adolescents toward varied groups: immigrants, racial and ethnic minorities, and migrants within Europe. I examined intergroup attitudes toward this range of groups, rather than toward any one group solely, in order to consider how intergroup attitudes might be different, or similar, depending on target groups. I focused only on attitudes of native students, in order to most clearly examine the views of a dominant majority to minority groups.

The countries that I selected were the United Kingdom (however, data were collected only in England, and not in Scotland, Wales, or Northern Ireland—this is the only group included which does not provide representative data across the entire country, and therefore will be referred to as England), Sweden, Poland, Bulgaria, Spain, Switzerland, and Greece. These seven groups provide coverage across regions in Europe: England in Western Europe, Sweden in Northern Europe, Switzerland in Central Europe, Spain and Greece in Mediterranean Europe, and Poland and Bulgaria in Eastern Europe.

These countries and England represent a range of migration rates. Bulgaria, Greece and Poland had low migration rates in 2009 when these data were collected, while Spain and Sweden had high migration rates. The affluence of the seven countries also varies. The countries range from very low (Bulgaria and Poland) to very high (Switzerland and Sweden) per-capita GDP. However, starting in 2007 and reaching extremes in 2009, the year of this study, nearly all countries experienced large negative changes in GDP. As discussed in the introduction, the political climate around the time of this study was most restrictive toward immigrants in Switzerland and that this climate had been in place for decades, while a new, growing sentiment was taking place in Bulgaria. However, the political climate represented by more than half of the voting adults in each of the countries in this set was moderate. The countries selected also represent a range of openness to migration across national policies reported by the Migrant Integration Policy Index (MIPEX). MIPEX collects data across several areas—labor market mobility, family reunion, education, political participation, long term residence, anti-discrimination, and access to nationality—and compiles scores in each area into an overall score. Sweden was the highest country in this set in 2010 and Spain was the second-highest country in overall score. The United Kingdom, while as high as Spain in

2007, dropped considerably in 2010. Out of all of the countries in the set, Switzerland had the lowest MIPEX score both in 2007 and in 2010. Finally, in the set of countries in this paper, the democracies in Bulgaria and Poland are the most recent, since 1989 after the end of communist rule in these countries. In Spain, the current democratic system has been in place since 1979 and the end of totalitarian rule in this country. And further, in Greece, the current democratic system has been in place since 1975 after a period of civil war which began after serious losses during resistance to the Nazis in WWII. Switzerland, Sweden, and the United Kingdom, are long-standing democracies in this group of countries.

In summary, the countries in this analysis represent a wide range of history, economic conditions, political climate, and migration rates, which will be useful to contextualize the findings of this study, discussed below. Having such a range provides opportunities to understand how national contexts play a part in the development of intercultural attitudes among young people. Economic climate is helpful to understand the relationship of threat to intergroup attitudes as well. History and political climate also provide insight into the openness of national contexts to issues of immigration and equality along many lines, including gender equality. This information sets the stage for the study which took place in 2009.

In the sections that follow, I present the empirical study which examined the experiences of young people in schools in these seven countries, and how those experiences were associated with intergroup attitudes at both the student- and school-levels using multi-level modeling. First, I present the data and measures used in this study, followed by the findings and a discussion. I also use the information provided above regarding each country to interpret the differences, and similarities, in the findings in each country.

Data

I used data from the 2009 IEA (International Association for the Evaluation of Educational Achievement) ICCS (International Civic and Citizenship Education Study) (IEA, 2009; Schultz, et. al, 2009). The IEA ICCS (2009) study was based on a previous civic education undertaken by the IEA in 1999, CIVED, which took highly rigorous steps to ensure the validity of the survey (Torney-Purta et al, 2001). The 1999 CIVED study was based on a two-stage design. In the first stage, extensive qualitative studies and case studies were undertaken in each of the participating countries to examine the meaning of civic-related constructs, including intercultural attitudes (Torney-Purta, Schwille & Amadeo, 1999). From these qualitative studies, survey instruments were developed in meetings with National Research Coordinators (Torney-Purta et al, 2001). The participating countries also pre-piloted and piloted preliminary forms of the instruments. The 1999 CIVED study included 28 countries and sampled about 90,000 adolescents, 9,000 teachers, and 4,000 school principals. The instrument was written in English, translated into 22 languages, and then returned to the National Research Coordinators for checking (Torney-Purta et al, 2001).

In the IEA ICCS (2009) study, civic knowledge, skills and attitudes were measured for 14-year olds across the world. The ICCS study sampled over 140,000 students, in more than 5,300 schools across 38 countries. National-level policy experts, teachers and school principals also responded to surveys providing contextual information. The study included a European module, which asked further questions relevant to Europe, the EU, and European identity (Schultz, et. al, 2009), as well as Latin American and Asian modules. The sampling

design of this study was a stratified, two-stage probability sampling design, which was similar to other large-scale IEA studies such as PIRLS and TIMSS (Schulz, Ainley, and Fraillon, 2011).

Only students who reported that they, as well as their parents, were born in the country of the test were included in this subsample, excluding both immigrant and second-generation youth from the sample. I used the data from seven of the European countries that participated in both the main study and the European module (n=16,847). The countries included in this study were the United Kingdom (however, data were collected only in England, and not in Scotland, Wales, or Northern Ireland—this is the only group included which does not provide representative data across the entire country, and therefore will be referred to as England), Sweden, Poland, Bulgaria, Spain, Switzerland, and Greece. At traditional levels of Type I error (.05), with the analytic methods used, this sample size far exceeded that needed to detect small effects (<.1 st dev) at high power (.90).

I my analyses, I incorporated elements of the complex survey design, and included schools as clusters as well as student sampling weights. To improve accuracy of standard errors, is important to compensate for the differing probabilities of selection at the school, class, and student levels, thus the weights that were used were a product of factors that reflect these probabilities (Brese, Jung, Mirazchiyski, Schulz & Zuelke, 2011), rescaled so that each country contributed equally. All analyses were conducted in Mplus v. 7.3 (M*plus,* 2014; Muthén, & Muthén, 1998-2015) and STATA 13 (StataCorp, 2013).

Methods and Measures

Associations at the individual- and school-levels were examined in a multi-level, multi-group structural equation model, with the focus being on the school experiences to enable greater understanding of the school practices and how they relate to intercultural attitudes. Using a multi-group modeling approach was helpful to be able to determine whether, and to what extent, the associations of interest vary across countries, by estimating these paths in each country (Asparouhov & Muthén, 2012). Using a multi-level approach was helpful to be able to estimate associations at both the within-level (student-level) and the between-level (student-level) simultaneously (Raudenbush & Bryk, 2002). This allowed me to control for effects at the school-level when interpreting the effects at the individual-level, and vice versa. It also allowed me to examine the variance explained at the individual-level compared to the school level, summarized by the R^2 for the within and between-level models in each country, and to examine the proportion of variance that is between groups, summarized in the intraclass correlation coefficient.

A model of these associations is shown in Figure 2.1; this model was tested simultaneously in the seven contexts described above. In this figure, the predictors are shown using boxes on the left and the latent factors which are the outcomes of interest are shown using ovals on the right. Correlations between the factors are shown by double-headed arrows. The associations that are estimated are shown by single-headed arrows pointing from the observed predictors toward each latent factor.

INSERT FIGURE 2.1 ABOUT HERE

In Paper 1 of this dissertation, I found sufficient evidence of measurement invariance to allow the comparison of factor means and associations. Following from that research, I used multiple imputation in the Bayesian framework to impute ten sets of

plausible values of the factor scores for each student that measured the three different intergroup attitudes of interest: attitudes toward immigrants (IMMIG), racial and ethnic minorities (RETH), and protectionist attitudes toward migration (PROT). The means of the set of plausible values in each of the factors were used as the outcomes of interest in this analysis.

The predictors that were chosen relate to the individual context and school contexts, as well as student-level reflection on the school environment. Background characteristics such as expected education, parent education, and home literacy resources, were included in this model as controls, because these characteristics have the potential to confound the relationships of interest in this study. I included these controls at the student-level and the school-level, by including the school means as predictors. I further included language proficiency as a control because the study of other languages, and dialogue in other languages, is an important way that young people gain access to other cultural tools and narratives. Finally, gender is included as a moderator and as directly relating to intercultural attitudes. Some research studies found gender differences in intercultural attitudes (Husfeldt, 2006; Coenders & Scheepers, 2002; Coenders et al, 2009; Barber, Fennelly, Torney-Purta, 2013) while others did not (Gorodzeisky & Semyonov, 2009; Evans & Need, 2002). Many studies typically use as a control and do not directly interpret these results, however this is a specific area of focus in this analysis.

Democratic and positive school climates are measured at both the student- and school-levels, through three scales that were available in the IEA (2009) data, which are described in greater detail in the measures section, below. These measures are included in order to examine whether, and to what extent, positive and democratic school communities

of practice are associated with intercultural attitudes. Communities of practice enable young people to come to understand the intercultural attitudes of adults as well as gain practice in the use of cultural narratives as tools to make sense of relations between groups. Further, these communities provide space to practice the civic skills of democratic participation and dialogue necessary in inclusive societies, and potentially conversely in communities that are not democratic or positive, to learn to behave and think in an exclusionary manner.

Finally, two interpersonal attitudes were included in this model as well: trust in people and schools, and attitudes toward gender equality. Trust in people and in schools, as a component of communities of practice, are important measures of openness to the community. As discussed above, Torney-Purta and Barber (2011) found a cluster of young people, who they called "alienated" who had extremely negative views toward immigrants and low trust in government. They postulated that these young people "seemed alienated from belief in the rights of others as well as lacking trust in government" (page 477). Because the focus of this paper is on schools as communities of practice, variables that measure trust in people and trust in schools were used instead, however the logic for their inclusion is along the same lines.

School means were included in the school-level component of the model to determine whether, above and beyond the school experiences that individual students reported, there was an additional association with overall school levels of these same measures. Additionally, two school-context variables were included: the percent native born and urbanicity. Including the percent native born enabled the exploration of this element of diversity at the school level, with the supposition that schools with more non-native born students would enable greater levels of contact. Urbanicity is included here as well, because it

is likely that contact, both directly in schools and indirectly through access to more diversity in the community, is greater in larger communities.

Among the countries included in this analysis, there was variation in the number of schools with higher percentages of non-native born students, summarized in Table 2.2 and shown visually in Figure 2.2. In all of the countries, the mean of the percent native-born in each country was quite high, with a mean of 89% native-born as the lowest value, in Spain. This indicates that there was, overall, not a high amount of diversity in terms of non-native born students in these countries at the time of this survey. Indeed in Bulgaria and Poland, the mean reached 99% in both countries, with minimum values of 90% and 88%, respectively, indicating high levels of homogeneity. In Spain and in Greece, the minimum percentages of non-native born students was much lower, with 11% in each country. In these countries, while most schools were rather homogenous in terms of native-born students, there are some schools that were homogenous in terms of non-native born students.

INSERT TABLE 2.2 AND FIGURE 2.2 ABOUT HERE

The number of private and public schools in each context are also given in Table 2.1. In all contexts except for Spain, the number of private schools included was quite low, with only 3-8 private schools included in Bulgaria, Greece, Poland, and England. While the public or private school status likely has relevance to the daily lived experience of adolescents in schools, it was unfortunately not possible to include this measure in this analysis, because the number of private schools was too low. However, the measure of the size of the community, relevant because intergroup contact and exposure was likely higher in larger towns and cities,

showed a greater variation, with adequate numbers of schools in each category, and was included in this analysis as a control at the school level.

Positive climate is measured by a student-teacher relations scale that includes five items such as "students get along well with most teachers" and "most teachers are interested in students' well-being" (see Schulz & Sibberns, 2004 for a discussion of the scales). This scale captures the positive community of practice which was theorized earlier in this paper to be associated with positive intercultural attitudes. Democratic climate was measured by two scales: the value of participation in school and openness in classroom discussions. The value of participation in school was measured by five items such as "student participation in how schools are run can make schools better" and "lots of positive changes can happen when students work together." Openness in classroom discussion was measured by five items comprising a scale such as "teachers encourage students to make up their own minds" and "teachers encourage students to express their opinions."

While no single covariate had a large degree of missing responses, the set of covariates together resulted in data loss just over 25%, as the patterns of missing responses meant that quite a few adolescents had missing data on at least one of the covariate variables, and quite a few schools had principal data that were missing. Thus, as a preliminary step, I created multiply imputed data using Baysean multiple imputation methods (Enders, 2010) with replaced the missing data with plausible values over ten imputed data sets. The resultant ten data sets were then used in the multi-group multi-level model that followed, which enabled me to include every native-born adolescent who responded to this questionnaire in this analysis.

In this analysis, I fit a series of multiple-group multi-level models, beginning with the unconditional model and progressing through a series of models with covariates added at each level (as recommended by Raudenbush & Bryk, 2002). I tracked the loglikelihood, AIC, and BIC across each model, which helped evaluate that the inclusion of each set of covariates added explanatory power to the model. In the case where a significant interaction was found, I further tested whether at least one interaction term was statistically significant using Wald tests. Finally, in the penultimate model, I tested the statistical significance of the differences between pairs of countries on the associations between each covariate and each outcome using Wald tests, and by estimating the differences with new parameters in M*plus* v. 7.3. Wald tests that were not significant indicated that there were no statistically significant differences between the countries on the association between that covariate and the outcome, and in these cases the estimates for that covariate were constrained to be equal. In some cases, where the Wald test was significant, the statistical significance of each pairwise difference was examined. Using the Šidák correction (Šidák, 1967; Abdi, 2007) for comparison tests among multiple groups, which corrects for the increased Type 1 error inherent in so many tests, an alpha-value of .006 ($p<.006$) was necessary to establish significant differences. Ultimately, groups with statistically significant differences at this alpha level were freely estimated, and groups without statistically significant differences were constrained to be equal.

Results

The results of this analysis are presented in a series of tables, Tables 2.3-2.5. Each outcome is presented separately, which was necessary given the number groups included. These tables show the standardized estimates and the corresponding p-values. Given that

the measures are on many different metrics, tracking and comparing the unstandardized estimates is cumbersome. The standardized estimates support ease of comparison across covariates, groups, and outcomes. Thus, these estimates will be used in model interpretation. All models presented and discussed here control for expected education, parent education, home literacy resources, and European language proficiency at the individual level, and urbanicity at the school level. Given the multiple comparisons that were made in this analysis, the Šidák correction was used to establish a p-value of <.006 as indicating statistical significance. In the tables, this is indicated by the triple-asterisk.

Tables A.2.1-A.2.3 show the unstandardized estimates and corresponding p-values. These estimates make clear the equality constraints that were imposed on groups that were not statistically significantly different from each other. For example, the effect of expected education was very close to the estimate shown (b=.027) in all of the groups to the point where the differences in estimates are not statistically significant. Therefore, that estimate was constrained to be equal across the countries and each estimate is the same in this table. Taking the step to constrain estimates that were not statistically significantly different from each other to be equal avoids errors brought about by over-interpreting non-significant differences as differences in the population, when in fact there was no evidence to support that. In Tables 2.6-2.8, which show the standardized estimates, the estimates that were constrained to be equal are shaded.

Positive and Democratic Communities of Practice

Trust in people and trust in schools were found to have positive associations with IMMIG and RETH, in nearly all of the groups, and the associations are stronger with regard to the RETH factor. The values of the associations range somewhat widely across groups,

from .03 to .16, and the consistent positive direction suggests that feelings of trust in both people and schools are related to positive intergroup attitudes across the seven countries. These effects are stronger in Sweden and Switzerland along the IMMIG factor and in Sweden along the RETH factor. For the most part, trust in people and in schools were not found to have an association with PROT, however. The exceptions to this finding are in Bulgaria and in Switzerland, where the effect is surprisingly positive, meaning that higher levels of trust were associated, on average, with more protectionism.

Positive and democratic school climates as measured by the student-teacher relations scale, the value of participation scale, and the opportunities for discussion scale, were found to be associated with positive views toward immigrants (IMMIG) and racial and ethnic minorities (RETH), to consistent degrees in all of the countries. Specifically, positive student-teacher relations had a fairly consistent association with IMMIG and RETH, with the strongest associations in Greece (b=.094) and the weakest associations in Sweden (b=.057) and England (b=.067), along both factors. Student-teacher relations were not found to be associated with PROT, however. Democratic school climates were measured with the two scales, value of participation at the school level and opportunities for discussion, and both of these scales were found to predict positive intergroup attitudes in both the IMMIG and RETH factors consistently in all of the groups. Among these two facets of democratic climates, the value of participation had somewhat higher associations (average b=.064 (IMMIG); average b=.081 (RETH)) than opportunities for discussion (average b=.041 (IMMIG); average b=.059 (RETH)). Interestingly, this association was highest along the RETH factor in Poland (b=.140). With regard to protectionist attitudes (PROT), opportunities for discussion were not found to have an association in any group.

However, the value of participation was found to have a positive association with protectionism in all of the countries, with fairly consistent values. This means that higher levels of the value of democratic participation in school was found to be related to more protectionist attitudes, which was surprising.

Gender

Gender—both the gender of the adolescents and attitudes toward gender equality—were found to predict positive attitudes in both the IMMIG and RETH factors, and this finding is not consistent across the countries along any factor. With regards to the IMMIG factor, there was a positive association between gender and attitudes toward immigrants, such that girls, on average, had more positive views than boys, across the countries (average b=.544). In addition to student gender, attitudes toward gender equality also predicted positive attitudes toward immigrants, meaning that on average, adolescents who agree with equal treatment for women were more likely to have positive attitudes toward immigrants (average b=.270). Further, the significant interaction between gender and gender equality in Poland (b=-.496) and Spain (b=-.600) indicates that this association between attitudes toward gender equality and IMMIG was stronger for boys than for girls in these countries.

The association between gender and RETH was not consistently found, with a positive association found only in Sweden (b=.653), and Sweden was the only country where the interaction between gender and gender equality was found only (b=-.596), where the association between gender equality and RETH was stronger for boys than for girls. Along the PROT factor, gender was found to have a negative association in Poland and Switzerland (b=-.092 and b=-.154, respectively). In these countries, girls were found to have less protectionist attitudes than boys. However, a negative association was found between

protectionist attitudes and attitudes toward gender equality consistently in all countries (average b=-.167), meaning that adolescents who had stronger views in favor of equal treatment for women had, on average, less protectionist views. The interaction between gender and attitudes toward gender equality was not found to be statistically significant in any of the countries along the PROT factor, and so it was not retained in the final model due to the added complexity that interactions introduce.

Contact

Overall, very few of the variables at the school-level of the statistical model were found to have significant associations with intergroup attitudes. Most notably, the measure of contact, percent native born in schools, was not found to have a statistically significant association, with positive intercultural attitudes in any country except for Poland (b=.331) within the PROT factor. This can be interpreted to mean that, above and beyond the experiences related to school that students reported at the individual-level, generally school-level means, including percent native born, do not have an additional association with intergroup attitudes.

Limitations

There are several limitations inherent in this analysis. First, these survey data are observational and therefore my results cannot support causal inference. However, I believe that my findings will be informative and provide a substantial basis to support social cohesion projects and studies that will examine their causal effects in the future. Second, my data are cross-sectional and thus cannot support developmental inferences, but rather the presence of associations among 14-year olds. Further research using mixed-methods data

would be helpful to address development. Third, only seven countries were used in this study, thus my inferences are limited to those countries and should not be generalized to Europe as a whole. Fourth, the items measuring attitudes toward immigrants and migrants do not distinguish among countries of origin, which would potentially differ on this basis. These issues are problematic conceptually, but I believe my findings will inform future work that may help distinguish between these attitudes more clearly. Finally, while this study examines individual-level and school-level associations with intercultural attitudes, peer group effects are also important in understanding these attitudes (see Barber, Torney-Purta, Wilkenfeld, & Ross, 2015). However, peer effects could not be included here because they were not included in the study.

Discussion

The goal of this paper was to examine intercultural attitudes among native-born adolescents in Europe, toward migrant and immigrant groups, and racial and ethnic minorities, within the context of schools. To these ends, I brought together groups of theories to examine intercultural attitudes among adolescents in a way that integrated individual cognition with environmental context. My goal was not to prove or disprove theory or examine economic, macro-forces. Instead, I used theory to construct a model that enabled me to examine patterns of intercultural attitudes within contexts.

Overall, school-level factors were not associated consistently with intercultural attitudes. This may be due to very little variation at the school level, summarized in the low intraclass correlations given in Tables 2.6-2.8, which ranged from .012-.068. This in keeping with similar findings in a study of the explanatory power of school-level factors in civic attitudes and beliefs (Isac, Maslowski, Creemers, & van de Werf, 2013). This does not

suggest that schools don't have a role to play in intercultural attitudes. Indeed, several of the individual-level variables that had significant associations were, in fact, reports of the school climate as experienced by students.

Through this examination, I found several factors that were associated with intergroup attitudes which are discussed below. In keeping with some research into the intergroup attitudes of adults, the economic and educational background of adolescents were associated with intergroup attitudes, measured by expected education, parent education, and home literacy resources, which were included in this analysis as controls. These associations were consistent in degree across the contexts in this study. Integrated threat theory (Stephan et al. 2005; Stephan et al., 2006) postulated that conditions of real or perceived threat leads to negative intergroup attitudes. This would translate in this study into a hypothesis that the associations between individual characteristics and intergroup attitudes vary by economic context. However evidence to support this was not found here, because the associations found were highly consistent. Shifting the emphasis from real or perceived threat to the narrative of threat found across contexts may be helpful to understand this result. While real and perceived economic and social conditions vary, the narrative of threat is more consistent and varies less from context to context.

These findings suggest that gender, and attitudes toward gender equality, have a potential and promising role to play in the ongoing development of intergroup attitudes. There is a clear connection between the protection of the human rights of women and girls, and the protection of the human rights of minority groups. It is very likely the case that young people who are willing to extend human rights to minority, immigrant and migrant groups would also be willing to extend human rights to women, and vice versa. The

interesting finding that attitudes toward gender equality were associated with positive attitudes toward immigrants and racial and ethnic minorities, and that this association was stronger for boys than girls in some contexts, certainly suggests this possible connection. This is especially important because many schools lack diversity along native and non-native lines, which limits the experience of diversity in these schools. However, gender differences are central to the lived experience of young people in schools and their communities. From a very early age, young people are aware of pervasive inequality and stereotypes along the lines of gender and gender identification, and as such these experiences may provide young people with insight into the pervasive inequality experienced by racial and ethnic minorities, immigrants, and migrants.

Narratives constructed around male entitlement might also play a role in these interesting findings with regard to gender. In all of the countries included here, men continue to hold positions of power. This translates into messages received, from a very young age, regarding norms for women and men, which in turn informs narratives among both boys and girls, within families and peer groups. These norms can be supported or challenged. This study focused on native-born adolescents, meaning that the boys in this sample were within two intersecting positions of power, by being both male and native-born. This implies that boys were negotiating their own positions relative to both girls and non-native students. Boys who are willing to extend human rights to girls are countering this dominant narrative. More research into these connections using mixed methods, which could examine in more depth the relationships between power, entitlement, and intercultural attitudes, would be fruitful.

The percent native born in each school, which was a measure of intercultural contact in this analysis, was not associated with positive intercultural attitudes at the school level in any group or along any of the factors, except a *positive* association with protectionist attitudes in Poland. This may be due, at least partly, to the lack of diversity in many of the schools, which was described earlier. Recall, that schools in Poland were the most homogenous of all of the countries in this study. This finding suggests that the contact theory might not be substantial enough to bring about more positive attitudes, particularly in contexts where the contact is extremely limited, for example where it had an ironic, opposite effect in Poland in this study. However, the value of conducting a multi-group analysis is clear in this case, because the expected relationship on the basis of contact theory was not found in any other context, some of which do have a degree of diversity present in schools, such as Greece and Spain.

This suggests that the very common proposition that providing young people with intercultural contact may improve intergroup attitudes might not be an effective solution. According to Contact Theory, there are also four conditions which must be met, which are often ignored in simple applications of the theory: that the contact is between groups that are of equal status, that the contact is characterized by a common goal, that the groups are engaged in cooperative activity, with the endorsement of authorities (Allport, 1954; Pettigrew, 1998). While it is often the case that the educational context supports the latter three conditions, the first condition, that contact between groups is of equal status, is unlikely to be fully supported. Schools are contexts which are influenced by the dominant society, and the power relations that are present there. While schools can anticipate and

challenge the values and norms of the dominant society, the power structures are pervasive enough to support a large degree of skepticism when it comes to this first condition.

In contrast, the positive and democratic climate of schools do seem to have a very promising role to play. Evidence of associations between positive and democratic school climates and positive intergroup attitudes were found in this analysis, especially with regard to positive climates. Positive student-teacher relationships, in which young people feel listened to and that teachers care about their well-being, might be associated with positive intergroup attitudes for some of the same reasons discussed above, such as increased trust which might generalize to others. If young people feel cared for, it is possibly the case that they are more willing to extend the care to others, even others they do not know. This is in line with a theory of a moral imperative of care, which Gilligan (1992) claimed was more common among women than men. Young people may make sense of their own position as a member of various intersecting groups by conceiving of the other as groups deserving or needing care and support. Indeed, Haste and Abrahams's theory (2008) of moral development highlights the position of one group relative to the other, as well as cultural narratives available which in this case may be narratives of care, is especially relevant and useful here.

The value of participation is a facet of democratic climates that was also found to be associated with positive attitudes. As discussed in the introduction to this dissertation, democratic societies require participation that is inclusive of all groups. It is perhaps the case that young people who experience broad participation in their schools are more willing to include minority groups as well. This finding was remarkably similar across the contexts included in this study, which included countries with long-standing and new democracies.

Hello, Sheepers, & Gijsberts (2002) proposed that education systems in long-standing democracies have had a longer period of time with a liberal democratic tradition in education. However, in this analysis, I found that the association between democratic climates in schools and positive intercultural attitudes remarkably similar across countries with long-standing as well as new democracies. Barber and colleagues (2013) found that there was not a difference in average support for immigrant's rights in countries with long-standing versus new democratic systems, using data from the 1999 CIVED study. Taken together, it does not appear to be the case that the democratic tradition in the national context was related to positive intergroup attitudes in this study, however lived experience of democratic practice in schools was in every context in this study.

In summary, school contexts are highly relevant to the intergroup attitudes of young people, which are included in the Developmental Niche model of Torney-Purta and colleagues (2011). The overall findings of this study support the idea that national and school contexts should be taken seriously as sites where positive intergroup attitudes may develop. This study uncovered some elements of the school context that were particularly relevant to positive intergroup attitudes, namely positive student and teacher relations as well as democratic practice. In this study, intercultural contact and integrated threat were found to be less relevant, however these conditions have been studied extensively in the research literature reviewed here. Rather, the overall culture and climate of schools, which may include cultural tools and frameworks to make sense of the self and the other, and in democratic contexts, seemed much more relevant here. Further research into the particular cultural tools and narratives that are being utilized, and how those tools are enacted a learned

in school communities of practice may further explain the role of culture and environment in the development of intergroup attitudes.

Conclusion

There are several limitations inherent in both papers of this dissertation. First, these survey data are observational and therefore my results cannot support causal inference. However, I believe that my findings will be informative and provide a substantial basis to support social cohesion projects and studies that will examine their causal effects in the future. Second, my data are cross-sectional and thus cannot support developmental inferences, but rather the presence of associations among 14-year olds. Further research using mixed-methods data would be helpful to address development. Third, only seven countries were included in this study and therefore my inferences are limited to those countries and should not be generalized to Europe as a whole. Finally, the items measuring attitudes toward immigrants and migrants do not distinguish among countries of origin, which would potentially differ on this basis. These issues are problematic conceptually, but I believe my findings will inform future work that may help distinguish between these attitudes more clearly.

There were two main, overarching goals of this dissertation. The first was to establish the validity of the use of several items to measure each of the outcomes of interest of this study: attitudes toward immigrants, racial and ethnic minorities, and protectionist attitudes toward migration. Through a process of examining the content validity, and the convergent and discriminant validity, of the measures, I established a preliminary measurement model, discussed in Paper 1. I then conducted an in-depth analysis of the evidence for measurement invariance across national contexts. I found evidence that measurement models with full invariance imposed did not fit the data, and then proceeded to diagnose and address the sources of non-invariance using Bayesian Approximate

Measurement Invariance techniques. This process enabled me to establish factor scores for each individual, on each factor, and in each country, and to use those scores with confidence in the subsequent analysis.

Overall, I found that the theories I brought together, specifically the Developmental Niche model of Torney-Purta and colleagues (2011) and the cultural model of Haste and Abrahams (2008) were particularly useful in structuring this exploration, as well as in interpretation of the results. Both models provided a means to account for national context and school contexts, as well as the cultural tools and narratives that young people might have available to make sense of themselves and others. These models motivated the particular approach that I took, in which I examined associations of a wide range of predictors with a range of intercultural attitudes, at both the individual and school levels. I also took a multi-group approach to explore whether, and to what extent, these associations were different in the seven national contexts included here.

However, this analysis leaves interesting questions unanswered, which will be fruitful areas of further research. First, my data did not allow me to fully explore the use of cultural tools and narratives in the intergroup attitudes that were observed. Because this study was designed to be valid in a wide range of countries and contexts, within and outside of Europe, this level of detail was not possible. In this analysis, it had to suffice to say that narratives of various kinds, such as the moral imperative of care, or orientations toward human rights generally, may have been relevant to the construction of meaning of the self and the other. This is likely an area where mixed methods would be useful to both examine the broad patterns of culture and how they relate to intercultural attitudes, as well as the specifics

regarding how and why cultural tools such as narrative are powerful means by which young people make meaning in this area.

One strength of this analysis is that in included several national contexts in Europe, and particularly in Eastern Europe where less emphasis has been placed. Many of the theories developed by social psychologists, some of which were reviewed in this dissertation, were developed by Western European and American researchers and theorists, and may or may not apply to Eastern European contexts. Having established measurement models that were invariant to a sufficient degree, I was able to explore national differences and similarities with confidence. Indeed, in this analysis, I found more similarities than differences, even across wide ranging histories, migration patterns, economic climate and political climate. However, this may be due to the fact that the study was not detailed enough at the national level. A viable next step might be to follow the process of the original development of the IEA civic education studies, particularly that of the 1999 study, in which the instrument itself was based on extensive case studies of civic engagement and education in each context. It may be extremely useful to conduct similar case studies of intergroup attitudes, their meaning, and the relevant cultural tools available, in each context, in order to develop an instrument that measures intergroup attitudes in a way that captures more of the differences between countries.

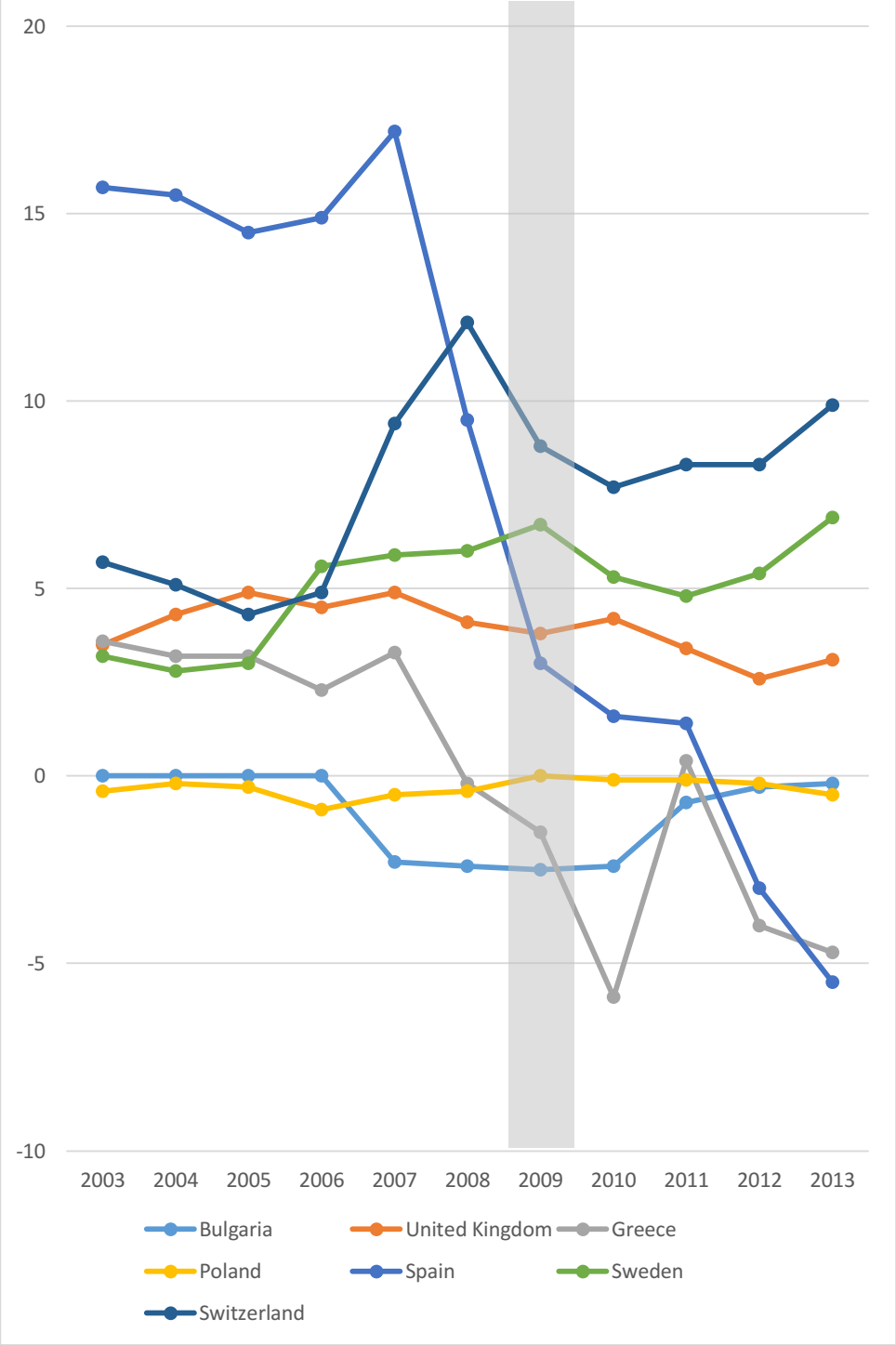
Finally, further examination into communities of practice and the relationship it has with intergroup attitudes would also be fruitful. In this analysis, this examination was limited to positive and democratic climates. While this provided valuable insight into the possible roles of schools in intergroup attitude development, this theory has much more potential to be valuable. Specifically, this theory is a theory of social learning, which describes a process,

very similar to the cultural model of Haste and Abrahams (2008), in which learning takes place in dialogue with others. The complex ecology of the classroom as a venue for learning as a social process may provide many additional insights into the formation of intergroup attitudes. In this case again, mixed methods research might be useful to identify some mechanisms by which this process takes place, and then to create an instrument that might capture the right data to unpack this ecology in more detail.

Introduction Tables and Figures

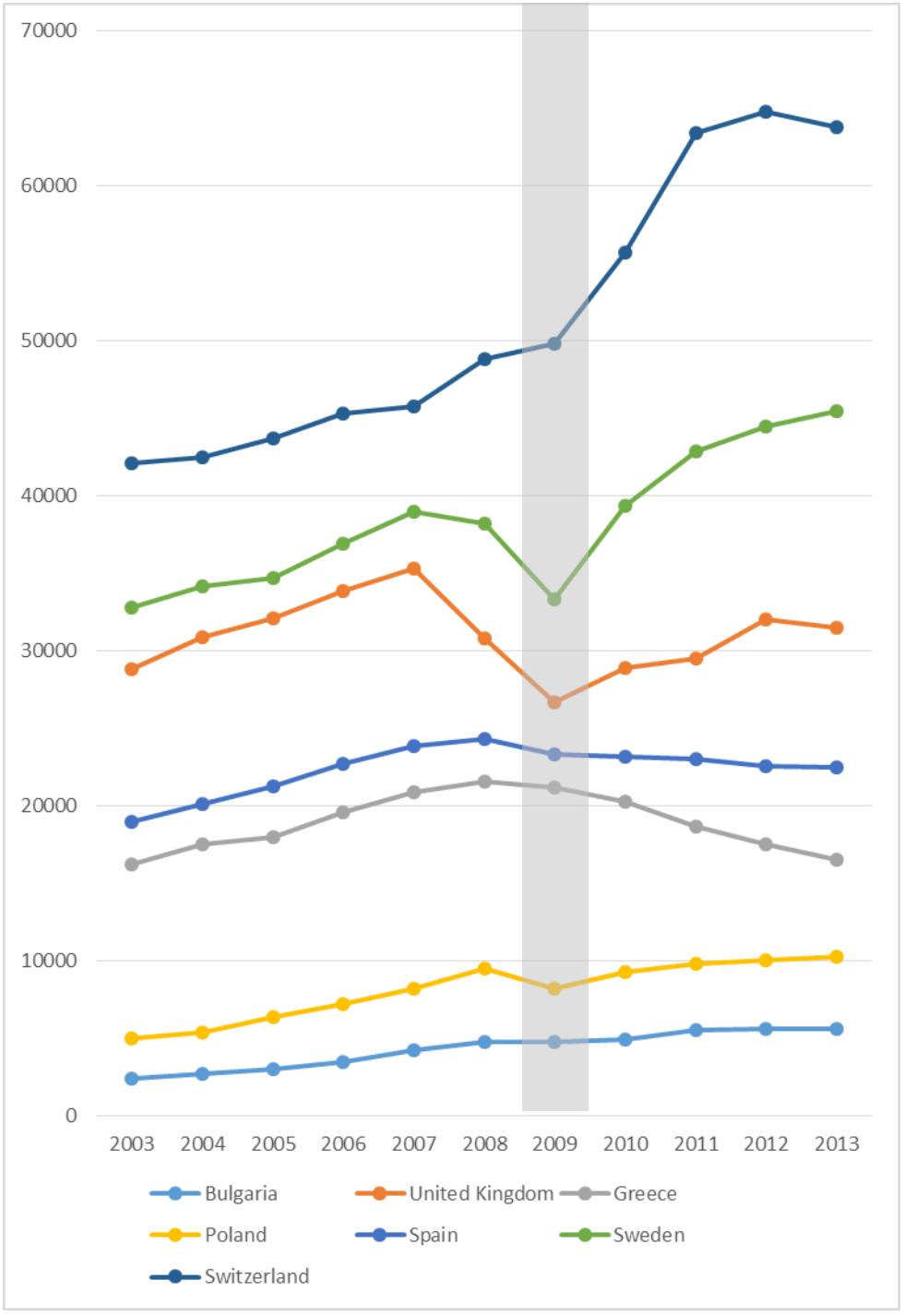
Begin on the next page

Figure I.1: Net migration rate in the seven countries from 2003-2013.



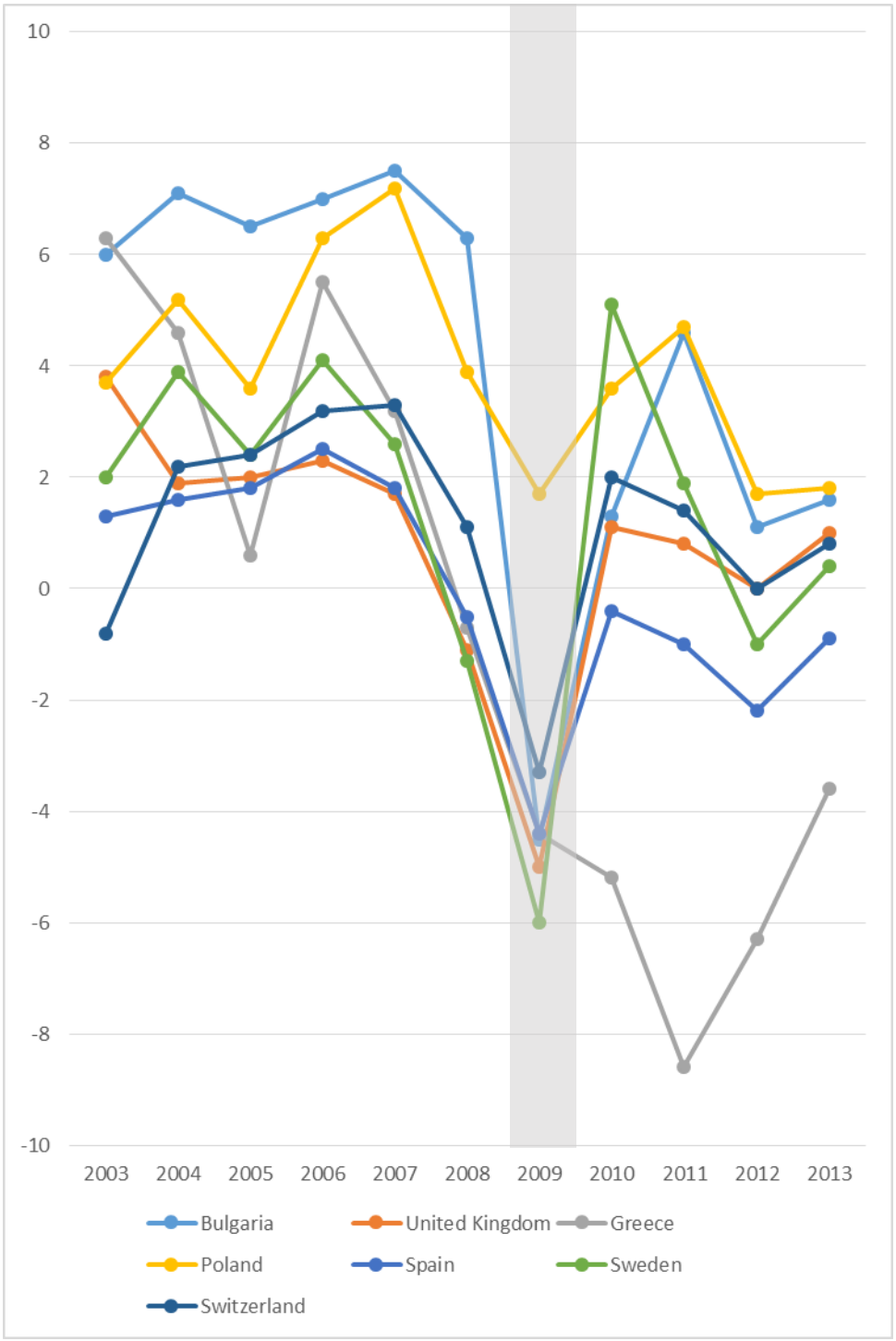
Source: Eurostat http://ec.europa.eu/eurostat

Figure I.2: Per-capita gross domestic product (GDP, in Euros) in the seven countries from 2003-2013.



Source: Eurostat http://ec.europa.eu/eurostat

Figure I.3: Percent change from year to year in gross domestic product (GDP, in Euros) in the seven countries from 2003-2013.



Source: Eurostat http://ec.europa.eu/eurostat

Table I.1: National Parliament election results, during the year closest to 2009, in each of the seven countries. Political parties that won over 1% of the vote, or are of particular relevance, are included.

Party (founding year)	% of vote
Bulgaria - 2009	
GERB - Citizens for European Development of Bulgaria (2006)	39.72
KB (BSP) - Coalition for Bulgaria (1999)	17.70
DPS - Movement for Rights and Freedoms (1990)	14.45
ATAKA - National Union Attack (2005)	9.36
Sinyata (SDS) - The Blue Coalition (2009, dissolved in 2012)	6.76
RZS - Order, Lawfulness, Justice (2005)	4.13
Greece - 2009	
PASOK - Panhellenic Socialist Movement (1974)	43.92
ND - New Democracy (1974)	33.48
KKE - Communist Party of Greece (1918)	7.54
LA. O. S. - Popular Orthodox Rally (2000)	5.63
Syriza - Coalition of the Radical Left (2004)	4.60
OIK.PRAS. - Ecologist Greens (2002)	2.53
Chrusi Aygi - Golden Dawn (1985)	0.29
Poland - 2007	
PO - Citizen's (Civic) Platform (2001)	41.51
PiS - Law and Justice (2001)	32.11
LiD - Left and Democrats (2006, dissolved 2008)	13.15
PSL - Polish Peasant (People's) Party (1990)	8.91
SRP - Self-defense of the Republic of Poland (1992)	1.53
LPR - League of Polish Families (2001)	1.30
PPP - Polish Labour Party (2001)	0.99
Spain - 2008	
PP - Peoples Party (1989)	39.42
PSOE - Spanish Socialist Workers Party (1879)	42.67
PSC-PSOE - PSOE Catalonia (1978)	6.57
IU - Left United (1986)	3.77
CIU - Convergence and Unity (1978)	3.03
EAJ-PNV - Basque Nationalist Party (1895)	1.19
UPyD - Union, Progress and Democracy (2007)	1.19
ERC - Republican Left of Catalonia (1931)	1.13

Continued on the next page

Party (founding year)	% of vote
Sweden - 2010	
SAP - Swedish Social Democratic Party (1889)	30.66
M - Moderate Party (1904)	30.60
MP - Green Party (1981)	7.34
FP - Liberal People's Party (1934)	7.06
C - Centre Party (1913)	6.56
SD - Sweden Democrats (1988)	5.70
KD - Christian Democrats (1964)	5.60
VP - The Left Party of Sweden (1917)	5.60
Switzerland - 2007	
SVP/UDC - Swiss People's Party (1971)	28.92
SP/PS - Social Democratic Party (1889)	19.54
PRD - Radical Democratic Party (1894, dissolved in 2009)	15.67
CVP/PDC - Christian Democratic Party (1912)	14.45
GPS/PES - Green Party (1983)	9.58
EVP/PEV - Evangelical People's Party (1919, 1994)	2.45
GLP/PVL - Green Liberal Party of Switzerland (2007)	2.14
FDP- Liberal Party (1913)	1.84
EDU/UDF - Federal Democratic Union (1975)	1.28
United Kingdom - 2010	
Con - Conservative (1834)	36.10
Lab - Labour (1900)	29.00
LD - Liberal Democrat (1988)	23.00
UKIP - UK Independence Party (1993)	3.10
BNP - British National Party (1982)	1.90
SNP - Scottish National Party (1934)	1.70
Green (1972, dissolved in 1990 and succeeded by regional parties)	1.00
PC - Plaid Cymru (1925)	0.60

Source: NSD European Election Database
http://www.nsd.uib.no/nsd/english/index.html

Table I.2: Migrant Integration Policy Index (MIPEX) scores for the seven countries in 2007 and 2010.

	Bulgaria		Greece		Poland		Spain		Sweden		Switzerland		UK	
	2007	2010	2007	2010	2007	2010	2007	2010	2007	2010	2007	2010	2007	2010
Overall score	--	45	40	50	43	44	62	65	85	84	43	43	66	56
Labor market mobility	--	40	45	50	45	48	79	84	100	100	53	53	55	55
Family reunion	--	51	47	49	67	67	76	85	89	84	40	40	56	54
Education	--	15	--	42	--	29	--	48	--	77	--	45	--	58
Political participation	--	17	25	40	13	13	56	56	75	75	58	59	53	53
Long term residence	--	57	56	56	65	65	72	78	78	78	41	41	74	31
Access to nationality	--	24	18	57	35	35	39	39	79	79	36	36	75	59
Anti-discrimination	--	80	50	50	35	36	49	49	88	88	31	31	81	86

Source: MIPEX: http://www.europa.eu

Table I.3: Indicators of democratic longevity and health in each of the seven countries.

	First year of current democratic voting with universal suffrage and a multi-party system	FH: Freedom rating (2009)*	FH: Civil liberties (2009)*	FH: Political rights (2009)*	Women in National Parliament (2010)**	Voter turnout in National Parliament elections***	election year***
Bulgaria	1989	2.0	2.0	2.0	21%	60.6	2009
England	1928 (1918 unequal)	1.0	1.0	1.0	22%	65.8	2010
Greece	1975	1.5	2.0	1.0	17%	70.9	2009
Poland	1989	1.0	1.0	1.0	20%	53.9	2007
Spain	1977	1.0	1.0	1.0	37%	75.3	2008
Sweden	1919	1.0	1.0	1.0	45%	84.6	2010
Switzerland	1971 (1990)	1.0	1.0	1.0	29%	48.3	2007

*Source: Freedom House 2009 Freedom in the World https://www.freedomhouse.org/
**World Bank: http://data.worldbank.org/indicator/SG.GEN.PARL.ZS
*** Source: Eurostat http://ec.europa.eu/eurostat/

Paper 1 Tables and Figures

Begin on the next page

Table 1.1: Construct, item name, and item wording of the items used to measure each factor, along with the relative frequency of agreement with each item, in each of the seven countries.

	Variable	Item wording	Bulgaria	Greece	Poland	Spain	Sweden	Switz.	Eng.
Attitudes toward immigrants (IMMIG)	IS2P26A	Immigrants should have the opportunity to continue speaking their own language.	0.84	0.85	0.86	0.69	0.71	0.65	0.56
	IS2P26B	Immigrant children should have the same opportunities for education that other children in the country have.	0.92	0.89	0.95	0.92	0.90	0.90	0.81
	IS2P26D	Immigrants should have the opportunity to continue their own customs and lifestyle.	0.89	0.82	0.88	0.70	0.74	0.66	0.66
	IS2P26E	Immigrants should have all the same rights that everyone else in the country has.	0.86	0.82	0.87	0.87	0.89	0.78	0.74
Attitudes toward racial and ethnic minorities (RETH)	IS2P25A	All <ethnic/racial groups> should have an equal chance to get good education in <country of test>	0.89	0.93	0.95	0.94	0.92	0.92	0.89
	IS2P25C	Schools should teach students to respect <members of all ethnic/racial groups>	0.86	0.89	0.92	0.91	0.91	0.88	0.87
	IS2P25D	<Members of all ethnic/racial groups> should be encouraged to run in elections for political office.	0.56	0.64	0.78	0.73	0.86	0.67	0.74
	IS2P25E	<Members of all ethnic/racial groups> should have the same rights and responsibilities.	0.83	0.83	0.91	0.90	0.92	0.87	0.87

Continued on the next page.

Table 1.1, continued: Construct, item name, and item wording of the items used to measure protectionist and open attitudes toward migration, along with the relative frequency of agreement with each item, in each of the seven countries.

	Variable	Item wording	Bulgaria	Greece	Poland	Spain	Sweden	Switz.	Eng.
Protectionist attitudes toward European migration (PROT)	ES2P08C	Other Europeans living in <country of test> leads to conflict and hostility between people of different nationalities.	0.39	0.49	0.35	0.49	0.37	0.61	0.61
	ES2P08D	Citizens of country of test will be safer from crime if they close their borders to <immigrants> from other European countries.	0.40	0.46	0.30	0.45	0.48	0.58	0.55
	ES2P08F	Allowing citizens of other European countries to come and work here leads to more unemployment for citizens of <country of test>	0.61	0.70	0.62	0.67	0.69	0.71	0.74
Open attitudes Toward European Migration (OPEN)	ES2P08A	Citizens of European countries should be allowed to live and work anywhere in Europe.	0.95	0.92	0.96	0.94	0.85	0.77	0.78
	ES2P08E	Other Europeans being allowed to live in <country of test> is good because they will bring different cultures with them.	0.86	0.76	0.83	0.86	0.74	0.59	0.70
	ES2P08H	Allowing citizens from other European countries to work here is good for the economy of <country of test>.	0.81	0.61	0.76	0.73	0.67	0.60	0.59
	ES2P08J	European citizens should be free to travel anywhere in Europe, so they get to understand other European cultures better.	0.92	0.83	0.92	0.88	0.86	0.84	0.76

Table 1.2: Sample poly-choric correlations among items measuring attitudes toward immigrants (i), and open (o) versus protectionist (p) attitudes, as well as attitudes toward racial and ethnic minorities (r), with sampling weights (n=16,847)

	i1	i2	i3	i4	i5	o1	o2	o3	o4	p1	p2	p3	r1	r2	r3
i2	0.53	--													
i3	0.46	0.60	--												
i4	0.60	0.60	0.53	--											
i5	0.46	0.72	0.64	0.58	--										
o1	0.31	0.36	0.30	0.30	0.33	--									
o2	0.29	0.36	0.32	0.34	0.36	0.43	--								
o3	0.26	0.24	0.24	0.25	0.22	0.38	0.39	--							
o4	0.25	0.36	0.27	0.29	0.32	0.44	0.43	0.33	--						
p1	-0.13	-0.20	-0.12	-0.17	-0.18	-0.02	-0.07	-0.01	-0.12	--					
p2	-0.19	-0.27	-0.20	-0.23	-0.23	-0.12	-0.15	-0.10	-0.16	0.53	--				
p3	-0.10	-0.07	-0.06	-0.08	-0.07	0.03	-0.02	-0.12	-0.02	0.35	0.39	--			
r1	0.37	0.57	0.41	0.40	0.49	0.33	0.30	0.21	0.27	-0.13	-0.18	-0.01	--		
r2	0.30	0.57	0.42	0.42	0.51	0.28	0.33	0.17	0.29	-0.15	-0.19	-0.01	0.68	--	
r3	0.27	0.40	0.40	0.35	0.40	0.20	0.26	0.19	0.19	-0.06	-0.11	-0.02	0.55	0.57	--
r4	0.31	0.58	0.44	0.42	0.57	0.28	0.30	0.16	0.27	-0.16	-0.20	-0.01	0.68	0.69	0.58

Figure 1.1: Measurement model of four factors measuring intercultural attitudes.

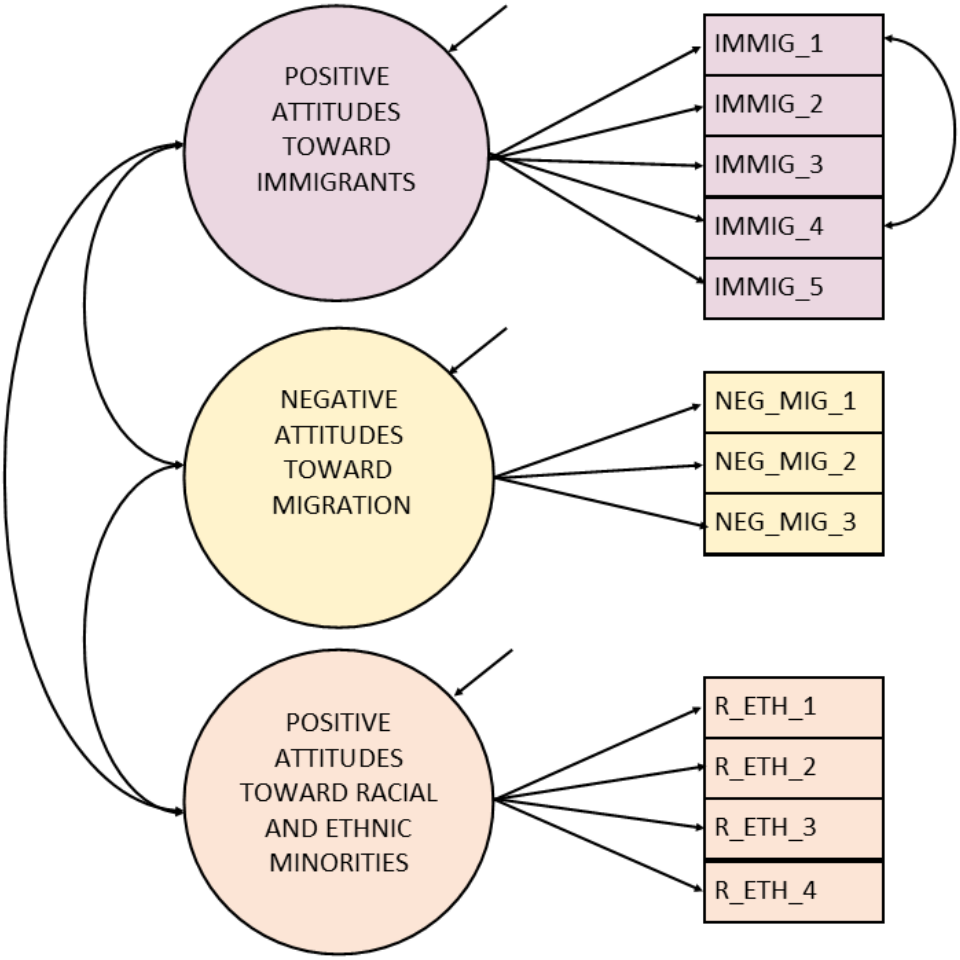
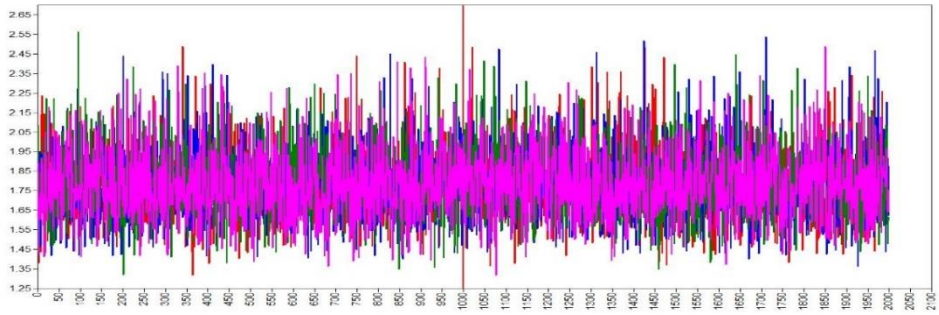
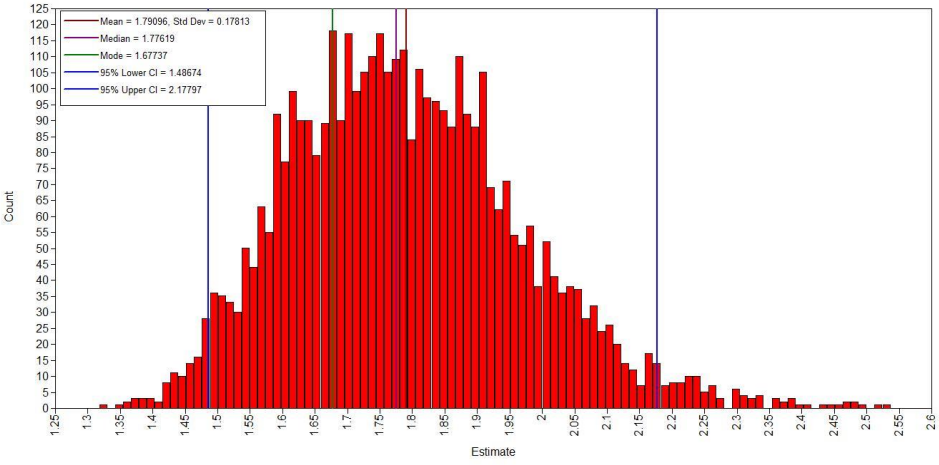


Figure 1.2: Example trace plot, posterior density plot, and autocorrelation plot for the first item loading in the first group (Bulgaria), RETH.

Trace plot



Posterior density plot



Auto-correlation plot

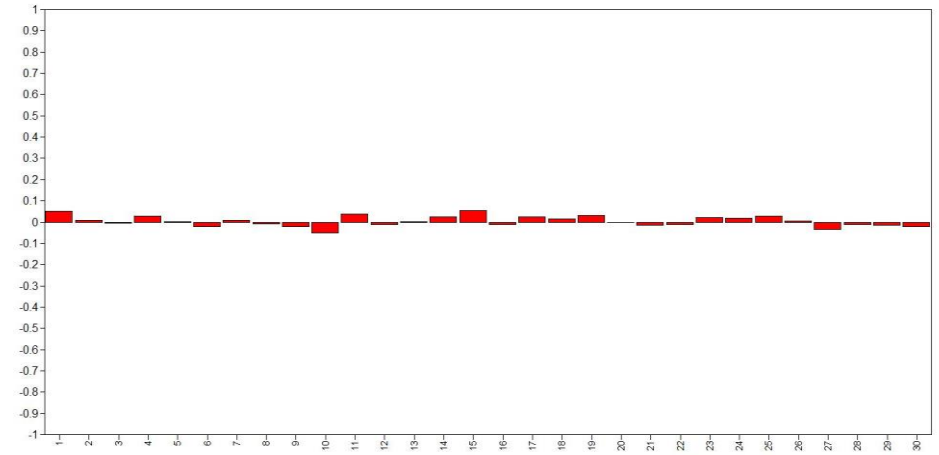


Table 1.3: Taxonomy of overall fit in all three factors.

	M1: Full Invariance	M2: Configural	M3A: Partial: Metric	M3B: Partial: Scalar	M4A: Partial Approx.: Scalar with priors on the lambdas	M4B: Partial Approx.: Scalar with priors on the lambdas and the taus
RETH: Overall						
PPP	0.000	0.495	0.520	0.529	0.470	0.445
(CI)	(190.323, 321.702)	(-36.922, 36.555)	(-36.223, 37.671)	(-37.859, 36.510)	(-36.159, 40.984)	(-32.770, 45.742)
IMMIG: Overall						
PPP	0.000	0.534	0.525	0.442	0.531	0.500
(CI)	(387.609, 532.108)	(-35.873, 47.994)	(-36.359, 34.488)	(-37.182, 35.486)	(-32.945, 28.126)	(-31.845, 29.854)
PROT: Overall						
PPP	0.000	0.525	0.488	0.537	0.488	0.482
(CI)	(76.345 162.790)	(-33.634, 30.651)	(-26.748, 33.085)	(-31.840, 33.091)	(-30.911, 33.961)	(-30.628, 34.519)

Figure 1.3: Estimated item loadings in the configural model compared to the metric models, RETH.

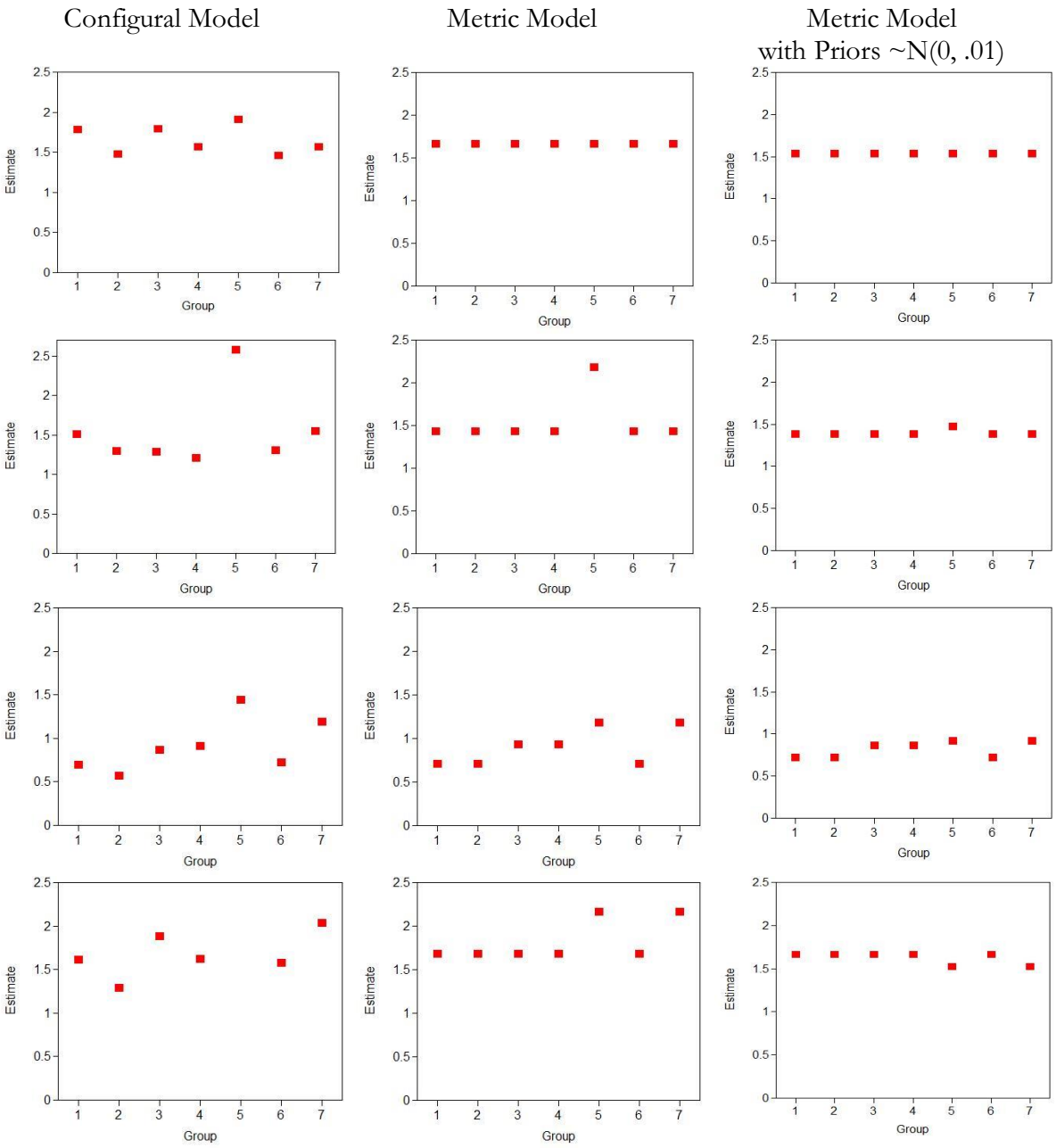


Figure 1.4: Estimated item thresholds in the configural model compared to the scalar models, RETH.

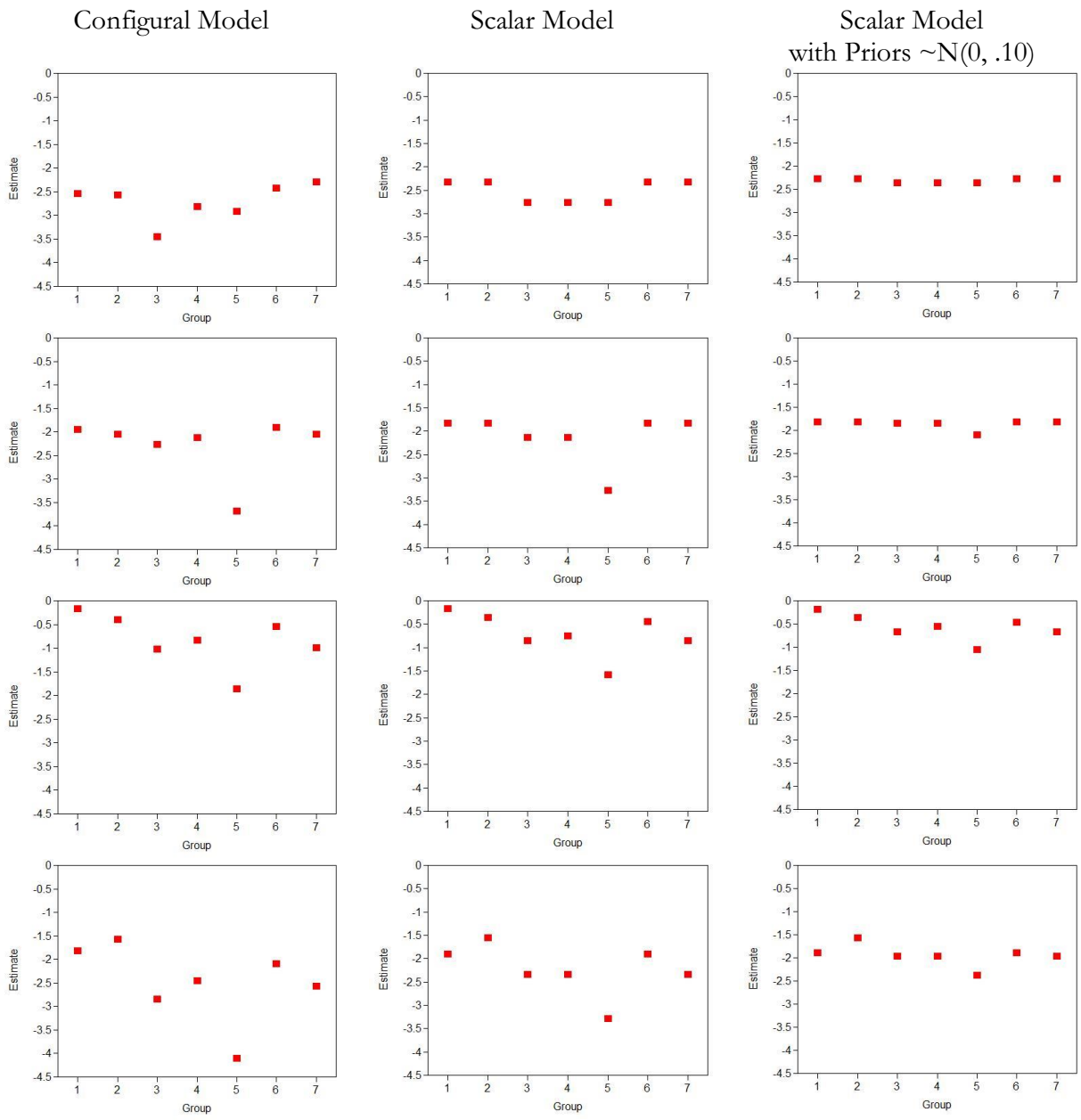
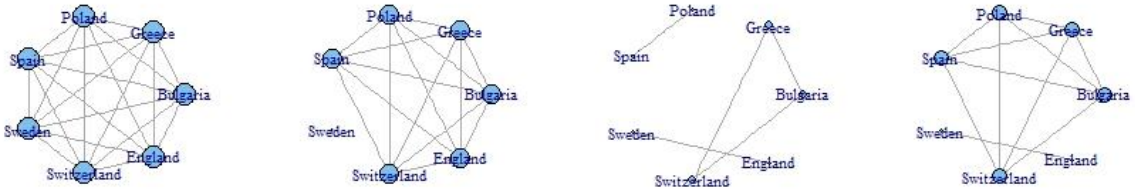


Figure 1.5: Network visualizations of item-loading equality constraints in all groups, RETH.

Item 1 Item 2 Item 3 Item 4



All items

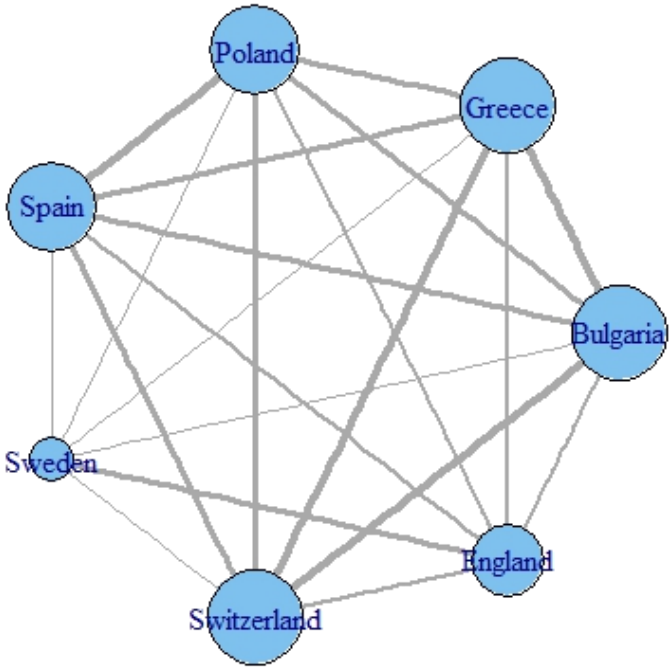
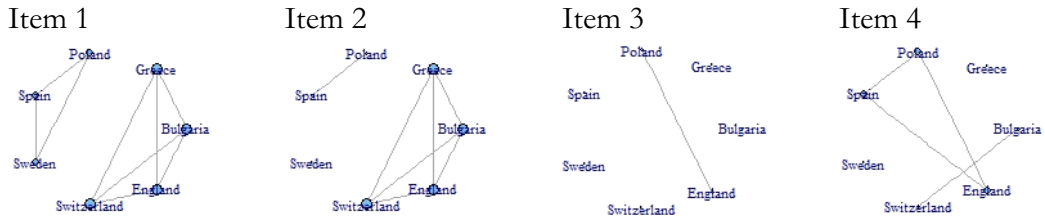


Figure 1.6: Network visualizations of item-threshold equality constraints in all groups, RETH.



All items

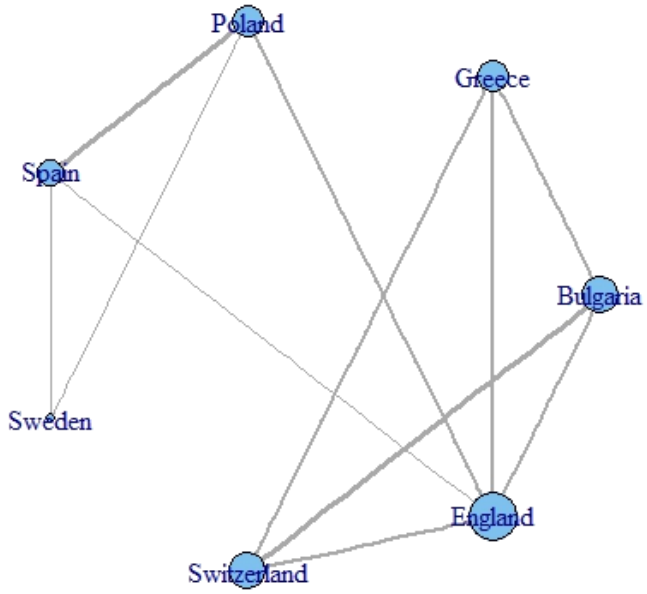


Figure 1.7: Box-plots of standardized factor scores for IMMIG, RETH, and PROT, showing the median (diamond), the interquartile range (rectangle), adjacent range (brackets) and outliers (dots).

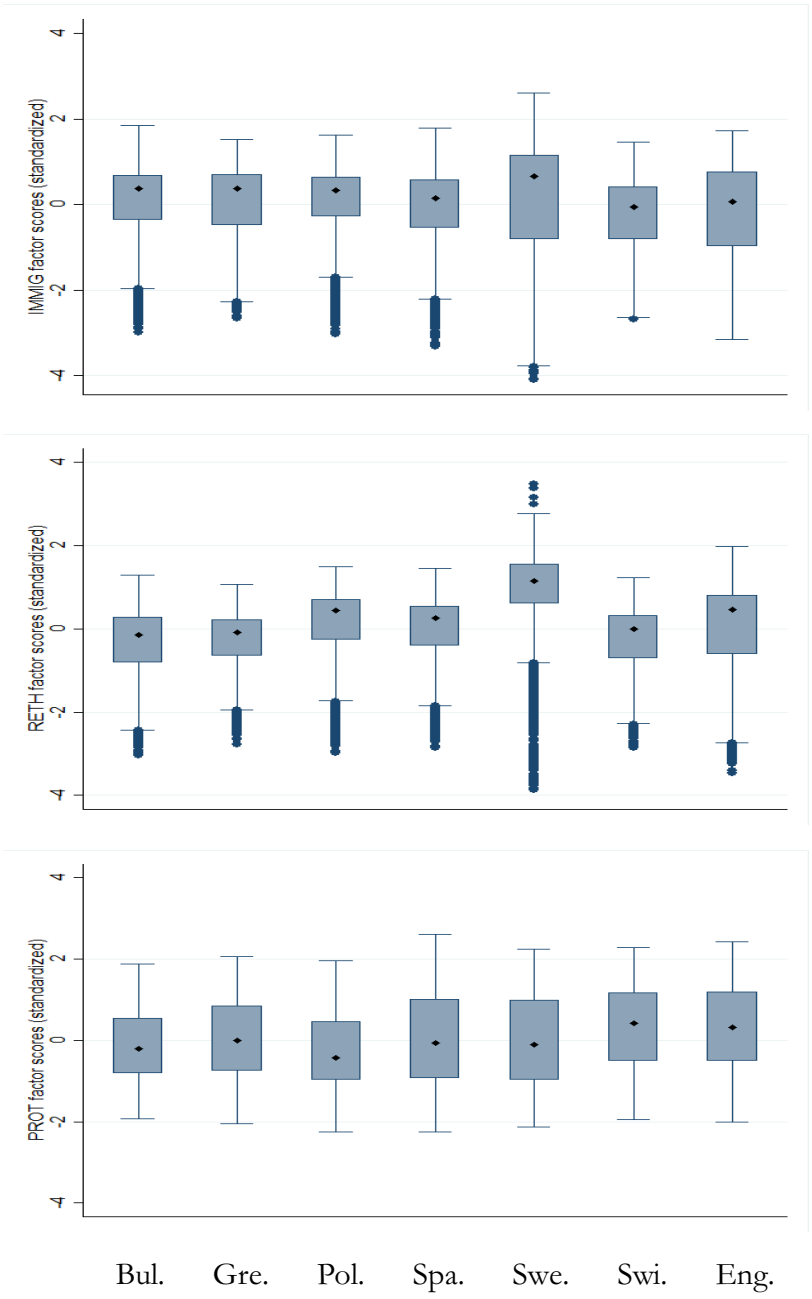


Table 1.4: Means of standardized factor scores in country, using the Partial Approximate Invariance method. Countries are ranked from highest to lowest. Mean differences that are *not* statistically significant (p>.006, with Šidák correction for multiple comparisons) are boxed and shaded.

	IMMIG mean		RETH mean		PROT mean
Sweden	0.110	Sweden	0.726	England	0.300
Poland	0.102	Poland	0.161	Switzerland	0.272
Bulgaria	0.087	England	0.025	Greece	0.037
Greece	0.074	Spain	-0.002	Sweden	-0.003
Spain	-0.097	Switzerland	-0.259	Spain	-0.028
England	-0.208	Greece	-0.300	Bulgaria	-0.110
Switzerland	-0.229	Bulgaria	-0.376	Poland	-0.213

Table 1.5: Means of standardized factor scores in country, using full invariance, which includes substantial bias due to statistical misfit. Countries are ranked from highest to lowest. Mean differences that are *not* statistically significant (p>.006, with Šidák correction for multiple comparisons) are boxed and shaded.

	IMMIG mean		RETH mean		PROT mean
Poland	0.453	Sweden	0.478	Switzerland	0.566
Bulgaria	0.388	Poland	0.313	England	0.526
Greece	0.127	Spain	0.120	Greece	0.121
Sweden	-0.093	England	-0.067	Spain	0.068
Spain	-0.216	Switzerland	-0.157	Sweden	-0.032
Switzerland	-0.480	Greece	-0.271	Bulgaria	-0.229
England	-0.713	Bulgaria	-0.469	Poland	-0.506

Table 1.6: Correlations of the factor scores in each country.

Bulgaria		
	IMMIG	RETH
RETH	0.32	
PROT	-0.13	-0.06
Greece		
	IMMIG	RETH
RETH	0.55	
PROT	-0.12	-0.08
Poland		
	IMMIG	RETH
RETH	0.36	
PROT	-0.12	-0.10
Spain		
	IMMIG	RETH
RETH	0.40	
PROT	-0.20	-0.09
Sweden		
	IMMIG	RETH
RETH	0.45	
PROT	-0.29	-0.16
Switzerland		
	IMMIG	RETH
RETH	0.43	
PROT	-0.28	-0.20
England		
	IMMIG	RETH
RETH	0.45	
PROT	-0.20	-0.11

Appendix to Paper 1

Table A1.1: Sample sizes and percentages in each country, with the percent native born in each country.

	full sample size	native students only	native students with native parents only	native students with native parents as % of full sample	native students as % of full sample	% native-born in the country*
Bulgaria	3,257	3226	3,138	96.35%	99.05%	99.40%
Greece	3,153	2870	2,510	79.61%	91.02%	92.20%
Poland	3,249	3230	3,166	97.45%	99.42%	99.80%
Spain	3,309	2962	2,732	82.56%	89.51%	89.10%
Sweden	3,464	3195	2,434	70.27%	92.23%	93.10%
Switzerland	2,924	2625	1,643	56.19%	89.77%	76.80%
England	2,916	2693	2,015	69.10%	92.35%	92.30%

Source: Eurostat http://ec.europa.eu/eurostat

Table A1.2: Taxonomy of overall and country-specific measurement model fit, attitudes toward racial and ethnic minorities (RETH)

	M1: Full Invariance	M2: Configural	M3A: Partial Metric	M3B: Partial Scalar	M4A: Partial Approx.: Scalar with Priors ~N(0,0.01) (lambdas)	M4B: Partial Approx.: Scalar with Priors ~N(0,0.01) (lambdas) and ~N(0,0.1) (taus)
overall	0.000 (190.323, 321.702)	0.495 (-36.922, 36.555)	0.520 (-36.223, 37.671)	0.529 (-37.859, 36.510)	0.470 (-36.159, 40.984)	0.445 (-32.770, 45.742)
Bulgaria	0.000 (37.965, 89.875)	0.522 (-14.579, 13.998)	0.485 (-13.035, 14.891)	0.500 (-13.717, 14.495)	0.518 (-12.610, 14.467)	0.482 (-15.177, 13.254)
Greece	0.000 (18.155, 69.798)	0.510 (-13.394, 17.610)	0.490 (-11.387, 14.637)	0.564 (-16.072, 12.777)	0.494 (-13.023, 12.916)	0.512 (-14.019, 12.529)
Poland	0.250 (-11.811, 24.934)	0.478 (-15.562, 15.129)	0.539 (-14.833, 15.000)	0.505 (-15.611, 16.674)	0.463 (-14.406, 15.608)	0.433 (-15.066, 14.944)
Spain	0.463 (-17.040, 14.063)	0.517 (-15.246, 15.005)	0.480 (-14.847, 15.333)	0.466 (-14.933, 16.713)	0.494 (-11.828, 13.858)	0.476 (-13.045, 13.333)
Sweden	0.000 (76.042, 151.380)	0.495 (-14.553, 14.468)	0.505 (-16.104, 13.530)	0.515 (-12.698, 13.983)	0.494 (-19.042, 15.571)	0.445 (-16.781, 13.684)
Switz.	0.360 (-11.132, 19.491)	0.505 (-14.381, 13.437)	0.559 (-18.626, 14.073)	0.520 (-14.613, 14.543)	0.494 (-18.509, 13.042)	0.537 (-13.088, 14.549)
England	0.018 (5.466, 42.338)	0.483 (-13.974, 13.278)	0.480 (-15.183, 14.622)	0.480 (-13.435, 12.174)	0.494 (-12.594, 14.689)	0.427 (-12.466, 15.357)

*PPP (PPP Confidence Intervals)

Table A1.3: Taxonomy of overall and country-specific measurement model fit, attitudes toward immigrants (IMMIG)

	M1: Full Invariance	M2: Configural	M3a: Partial Metric	M3b: Partial Scalar	M4a: Partial Approx.: Scalar with Priors ~N(0,0.05) (lambdas)	M4b: Partial Approx.: Scalar with Priors ~N(0,0.05) (lambdas) and ~N(0,0.1) (taus)
overall	0.000 (387.609, 532.108)	0.534 (-35.873, 47.994)	0.525 (-36.359, 34.488)	0.442 (-37.182, 35.486)	0.531 (-32.945, 28.126)	0.500 (-31.845, 29.854)
Bulgaria	0.000 (34.072, 83.104)	0.520 (-15.887, 16.848)	0.495 (-14.030, 16.191)	0.490 (-11.984, 16.518)	0.562 (-15.848, 15.498)	0.562 (-18.044, 15.307)
Greece	0.000 (70.474, 132.692)	0.534 (-13.752, 13.189)	0.495 (-15.497, 15.959)	0.510 (-13.763, 13.557)	0.531 (-13.939, 13.941)	0.531 (-12.782, 12.731)
Poland	0.000 (24.117, 79.649)	0.500 (-17.729, 14.828)	0.500 (-16.857, 15.704)	0.519 (-13.945, 14.773)	0.594 (-13.004, 13.047)	0.547 (-13.184, 14.738)
Spain	0.000 (68.954, 141.416)	0.554 (-14.911, 14.977)	0.539 (-14.938, 14.029)	0.510 (-14.366, 15.435)	0.562 (-12.244, 16.207)	0.531 (-12.212, 16.859)
Sweden	0.000 (57.783, 133.914)	0.490 (-12.814, 16.670)	0.500 (-11.878, 14.309)	0.529 (-13.654, 15.885)	0.453 (-12.398, 11.593)	0.469 (-12.749, 12.746)
Switz.	0.000 (11.023, 56.397)	0.466 (-14.337, 15.181)	0.471 (-12.491, 15.775)	0.462 (-11.889, 12.692)	0.547 (-11.735, 15.858)	0.547 (-12.142, 15.820)
England	0.000 (12.583, 69.838)	0.500 (-14.649, 10.738)	0.554 (-14.599, 13.091)	0.548 (-13.377, 13.902)	0.484 (-13.369, 11.074)	0.484 (-12.861, 11.155)

*PPP (PPP Confidence Intervals)

Table A1.4: Taxonomy of overall and country-specific measurement model fit, protectionist attitudes toward migration (PROT)

	M1: Full Invariance	M2: Configural	M3a: Partial Metric	M3b: Partial Scalar	M4a: Partial Approx.: Scalar with Priors ~N(0,0.50) (lambdas)	M4b: Partial Approx.: Scalar with Priors ~N(0,0.50) (lambdas) and ~N(0,0.50) (taus)
overall	0.000	0.525	0.488	0.537	0.488	0.482
	(76.345 162.790)	(-33.634, 30.651)	(-26.748, 33.085)	(-31.840, 33.091)	(-30.911, 33.961)	(-30.628, 34.519)
Bulgaria	0.030	0.539	0.488	0.573	0.561	0.573
	(-0.388, 32.227)	(-12.351, 13.019)	(-11.339, 12.212)	(-13.395, 9.553)	(-12.826, 12.439)	(-13.011, 12.540)
Greece	0.238	0.514	0.530	0.524	0.524	0.512
	(-9.267, 20.016)	(-13.332, 9.817)	(-10.042, 10.689)	(-11.685, 12.037)	(-10.218, 12.068)	(-10.233, 12.360)
Poland	0.043	0.553	0.500	0.530	0.530	0.530
	(-0.533, 29.408)	(-11.863, 11.365)	(-10.736, 12.259)	(-10.466, 12.123)	(-10.376, 11.972)	(-10.465, 12.096)
Spain	0.323	0.514	0.476	0.482	0.470	0.463
	(-11.084, 16.887)	(-12.641, 12.997)	(-10.149, 14.572)	(-12.955, 10.856)	(-12.519, 10.816)	(-12.541, 10.894)
Sweden	0.000	0.518	0.543	0.561	0.549	0.561
	(31.591, 86.871)	(-11.124, 10.928)	(-10.683, 17.105)	(-10.941, 9.553)	(-10.818, 9.601)	(-11.265, 9.656)
Switz.	0.043	0.479	0.439	0.445	0.445	0.445
	(-2.173, 28.270)	(-11.242, 14.386)	(-9.457, 14.027)	(-11.368, 11.376)	(-10.661, 11.522)	(-10.364, 12.713)
England	0.098	0.507	0.500	0.476	0.445	0.445
	(-5.416, 23.075)	(-13.472, 11.994)	(-12.833, 10.328)	(-12.637, 13.913)	(-13.019, 15.652)	(-13.267, 15.732)

* PPP (PPP Confidence Intervals)

Table A1.5: Differences in the loadings, by pairs of groups, with statistical significance, RETH.

Item 1	group diff.	Item 2	group diff.	Item 3	group diff.	Item 4	group diff.
LAM1_1V2	0.298	LAM2_1V2	0.207	LAM3_1V2	0.134	LAM4_1V2	0.318
LAM1_1V3	-0.012	LAM2_1V3	0.221	LAM3_1V3	-0.168*	LAM4_1V3	-0.271
LAM1_1V4	0.207	LAM2_1V4	0.295	LAM3_1V4	-0.216*	LAM4_1V4	-0.004
LAM1_1V5	-0.127	LAM2_1V5	-1.062*	LAM3_1V5	-0.743*	LAM4_1V5	-1.139*
LAM1_1V6	0.321	LAM2_1V6	0.200	LAM3_1V6	-0.026	LAM4_1V6	0.045
LAM1_1V7	0.217	LAM2_1V7	-0.041	LAM3_1V7	-0.495*	LAM4_1V7	-0.417
LAM1_2V3	-0.316	LAM2_2V3	0.009	LAM3_2V3	-0.302*	LAM4_2V3	-0.595*
LAM1_2V4	-0.082	LAM2_2V4	0.090	LAM3_2V4	-0.351*	LAM4_2V4	-0.330
LAM1_2V5	-0.425	LAM2_2V5	-1.278*	LAM3_2V5	-0.875*	LAM4_2V5	-1.457*
LAM1_2V6	0.022	LAM2_2V6	-0.008	LAM3_2V6	-0.160	LAM4_2V6	-0.274
LAM1_2V7	-0.081	LAM2_2V7	-0.252	LAM3_2V7	-0.624*	LAM4_2V7	-0.738*
LAM1_3V4	0.229	LAM2_3V4	0.080	LAM3_3V4	-0.049	LAM4_3V4	0.268
LAM1_3V5	-0.113	LAM2_3V5	-1.291*	LAM3_3V5	-0.574*	LAM4_3V5	-0.863*
LAM1_3V6	0.334	LAM2_3V6	-0.022	LAM3_3V6	0.143	LAM4_3V6	0.320
LAM1_3V7	0.232	LAM2_3V7	-0.262	LAM3_3V7	-0.327*	LAM4_3V7	-0.149
LAM1_4V5	-0.338	LAM2_4V5	-1.360*	LAM3_4V5	-0.526*	LAM4_4V5	-1.127*
LAM1_4V6	0.103	LAM2_4V6	-0.097	LAM3_4V6	0.191	LAM4_4V6	0.046
LAM1_4V7	0.000	LAM2_4V7	-0.342	LAM3_4V7	-0.277*	LAM4_4V7	-0.414
LAM1_5V6	0.448	LAM2_5V6	1.265*	LAM3_5V6	0.717*	LAM4_5V6	1.182*
LAM1_5V7	0.342	LAM2_5V7	1.018*	LAM3_5V7	0.245	LAM4_5V7	0.705
LAM1_6V7	-0.106	LAM2_6V7	-0.247	LAM3_6V7	-0.470*	LAM4_6V7	-0.458

Figure A1.1: Estimated item loadings in the configural model compared to the metric models, IMMIG.

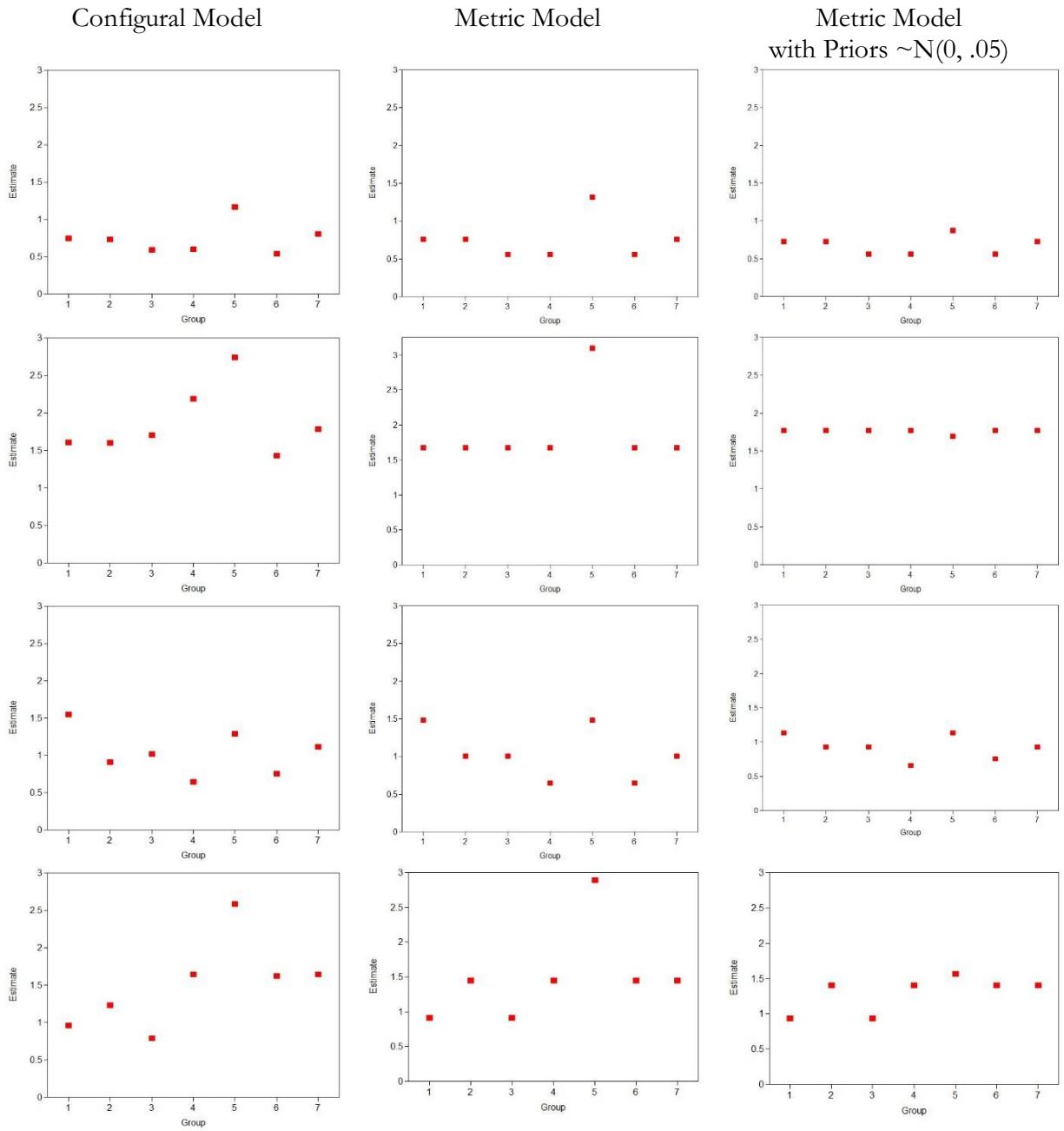


Figure A1.2: Estimated item thresholds in the configural model compared to the scalar models, IMMIG.

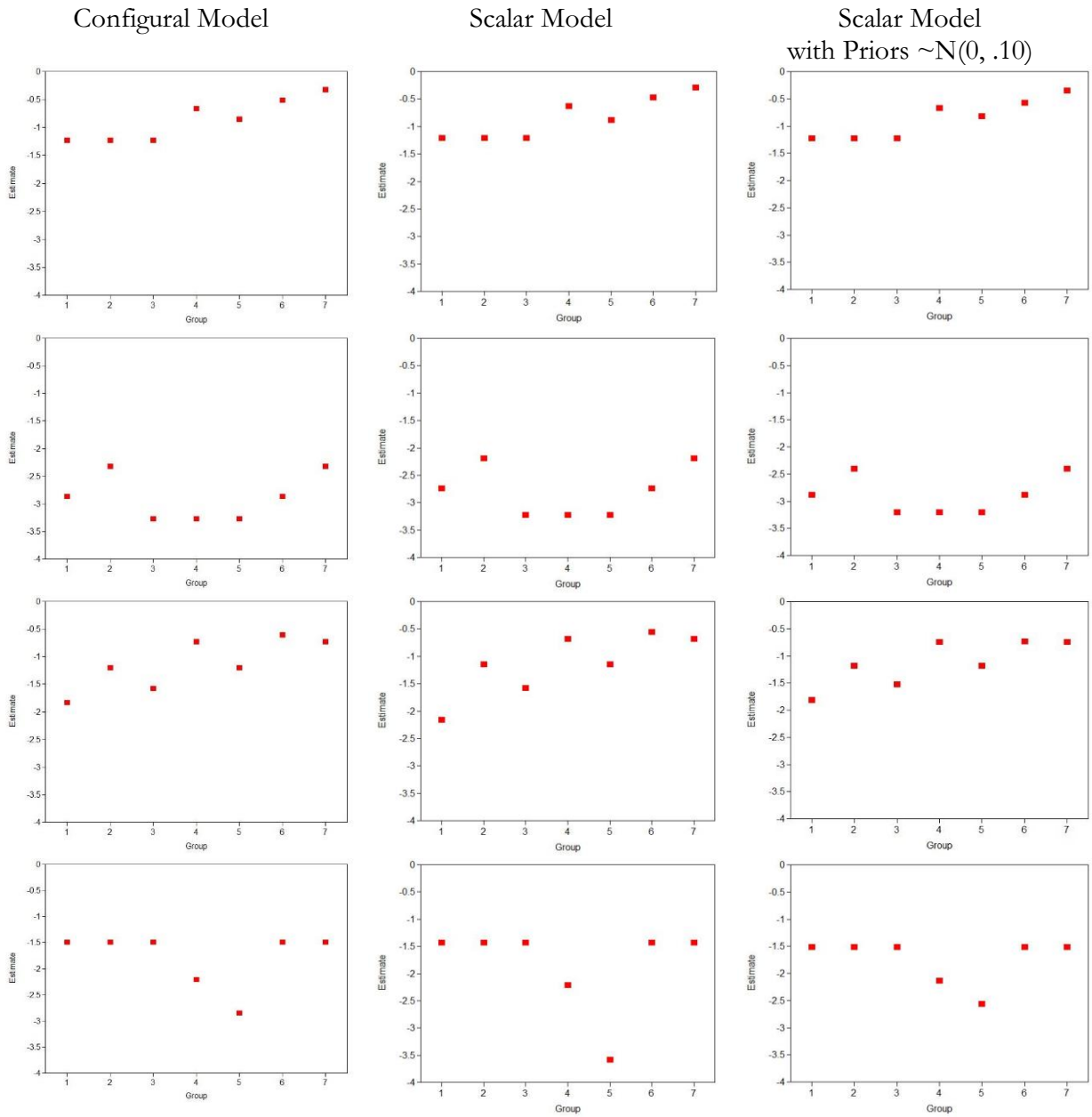


Figure A1.3: Estimated item loadings in the configural model compared to the metric models, PROT.

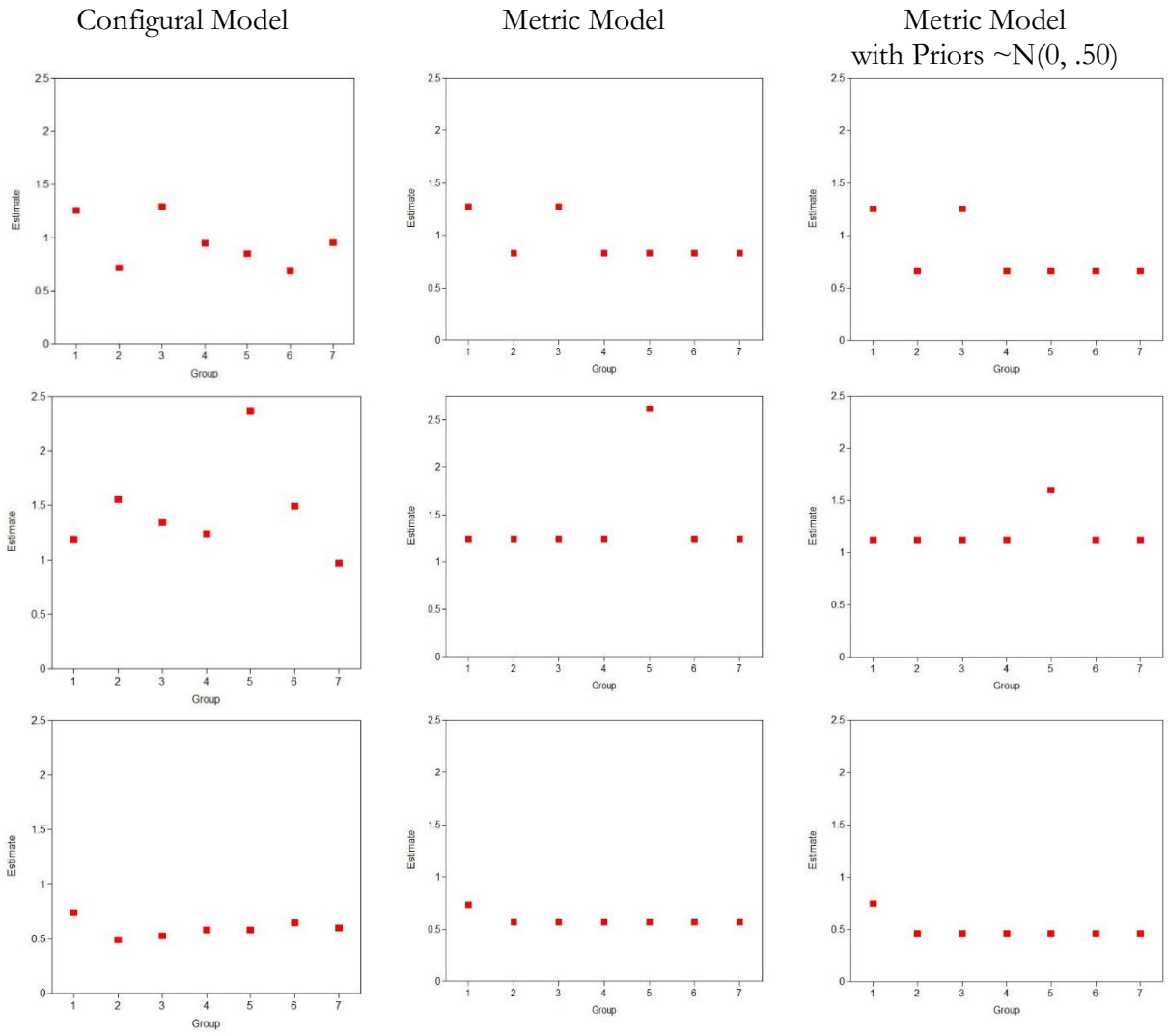


Figure A1.4: Estimated item thresholds in the configural model compared to the scalar models, PROT.

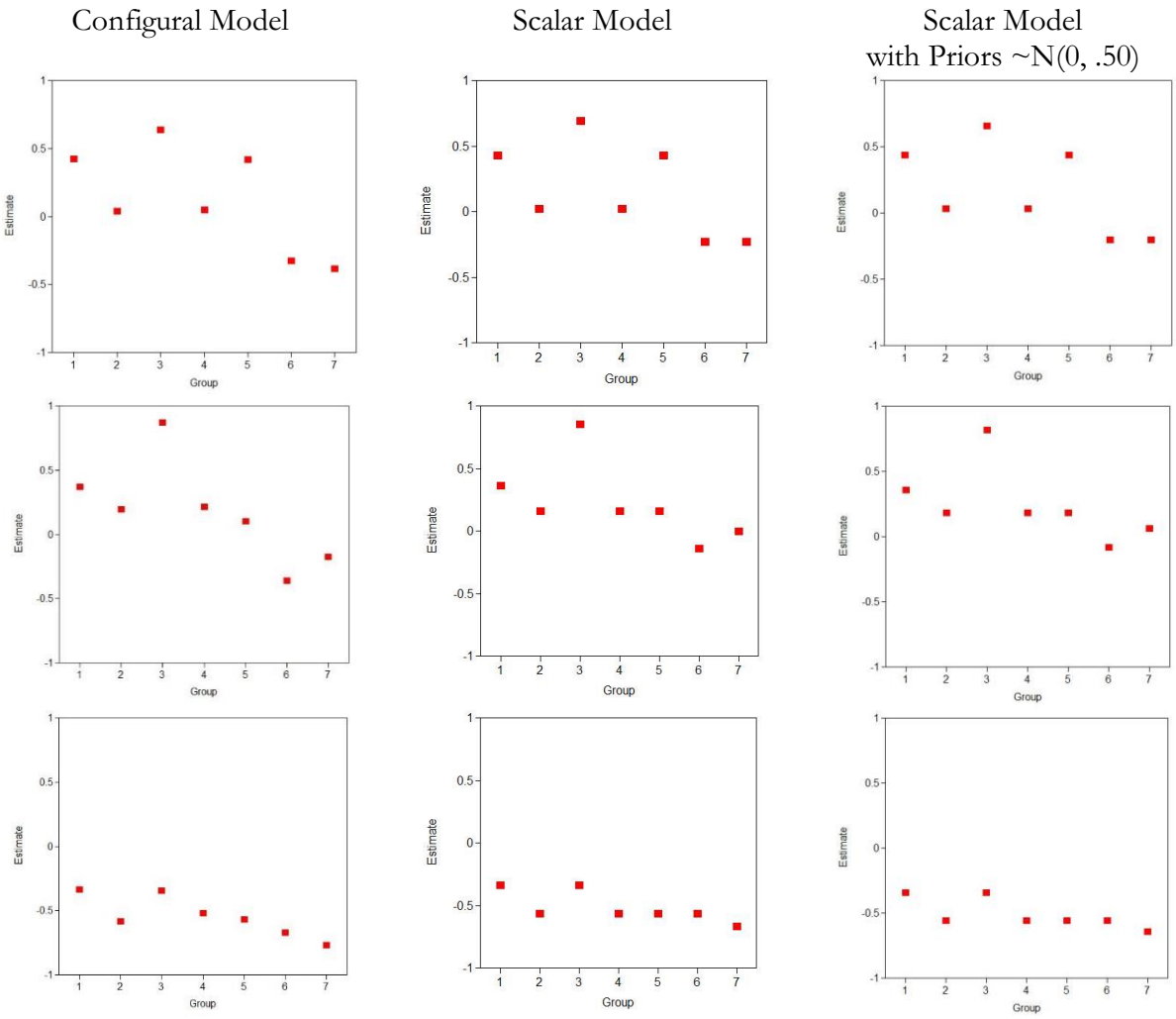
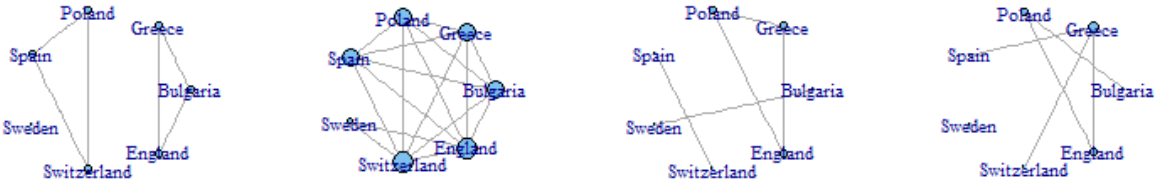


Figure A1.5: Network visualizations of item-loading equality constraints in all groups, IMMIG.

Item 1 Item 2 Item 3 Item 4



All items

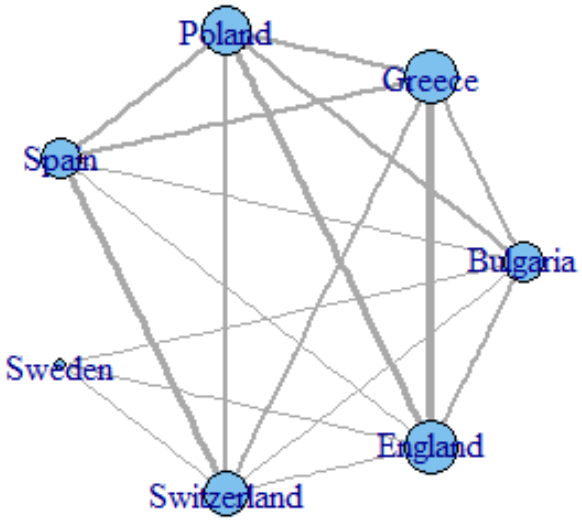


Figure A1.6: Network visualizations of item-threshold equality constraints in all groups, IMMIG.

Item 1 Item 2 Item 3 Item 4

All items

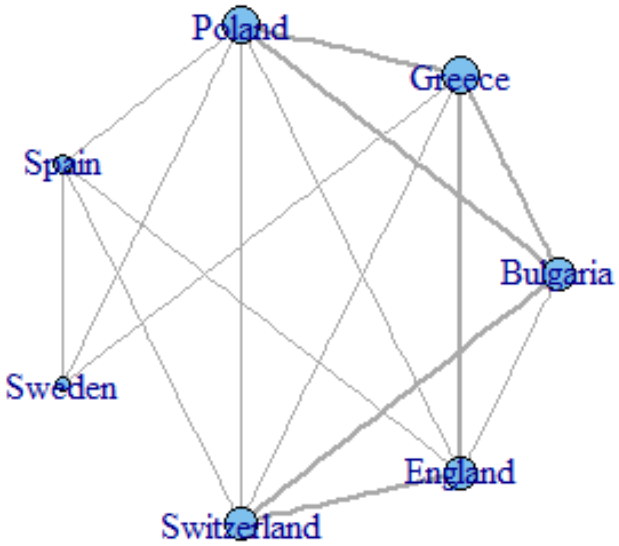
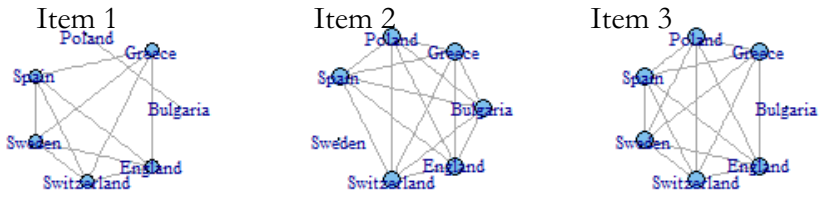


Figure A1.7: Network visualizations of item-loading equality constraints in all groups, PROT.



All items

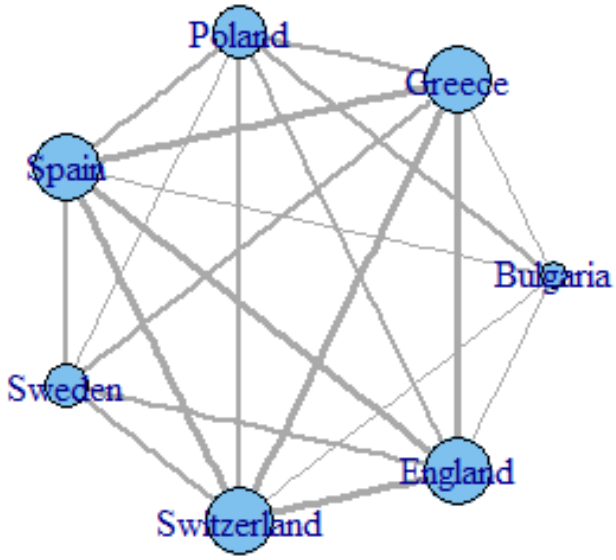
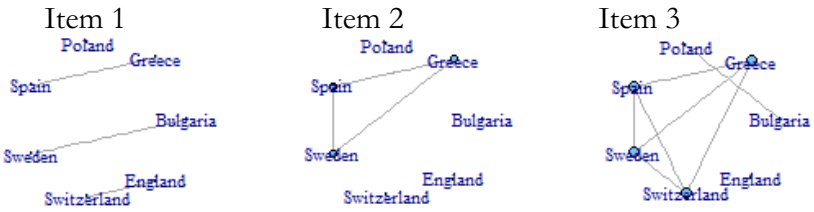
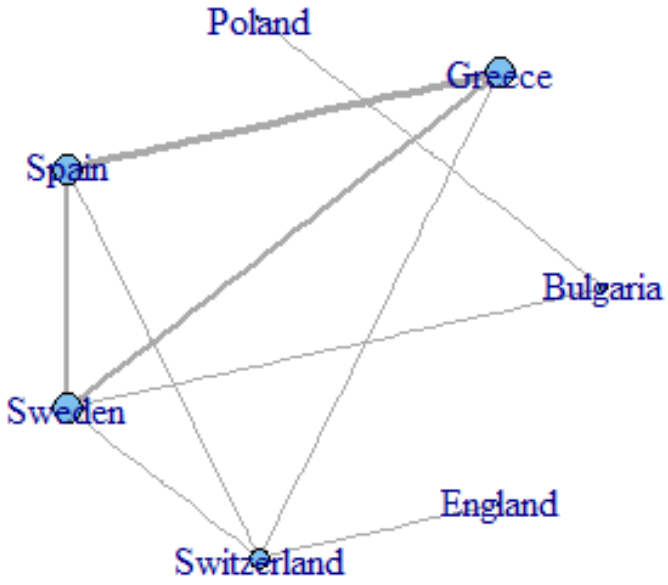


Figure A1.8: Network visualizations of item-threshold equality constraints in all groups, PROT.



All items



Paper 2 Tables and Figures

Begin on the next page

Figure 2.1: Multi-level theoretical and statistical model of associations with attitudes toward immigrants (IMMIG), racial and ethnic minorities (RETH), and protectionist attitudes toward migration (PROT), at the individual- and school-levels.

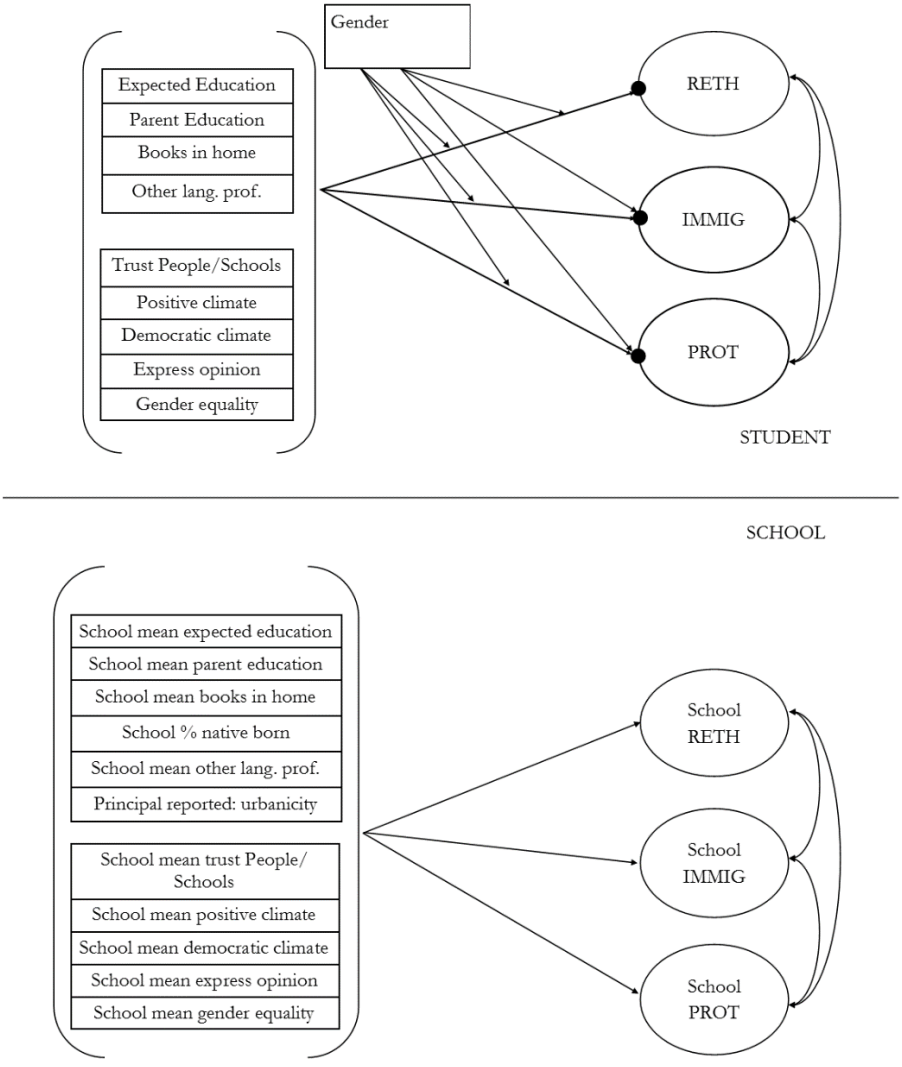


Table 2.1: Frequency of public and private schools, frequency of schools by community size, and summary statistics regarding the proportion of native-born students, in each of the seven countries.

	Bulgaria	Greece	Poland	Spain	Sweden	Switz.	England
public	155	117	144	99	127	33	90
private	3	8	6	44	22	119	7
missing data	0	28	0	5	14	2	19
a village, hamlet or small rural area (fewer than 3,000)	30	18	48	8	23	30	18
A small town (3,000 to about 15,000)	26	26	32	31	32	63	10
A town (15,000 to about 100,000)	54	46	36	55	54	34	22
A city (100,000 to about 1,000,000)	28	10	32	37	28	13	34
A large city (over 1,000,000)	20	21	2	13	12	0	21
missing data	0	32	0	4	14	14	11
Mean % native born	0.99	0.90	0.99	0.89	0.92	0.90	0.93
sd	0.02	0.11	0.02	0.13	0.10	0.09	0.08
min	0.90	0.11	0.88	0.11	0.42	0.41	0.60
max	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 2.2: Means and standard deviations of the standardized factor scores for girls and boys in each country.

	Girls				Boys		
	IMMIG mean	RETH mean	PROT mean		IMMIG mean	RETH mean	PROT mean
Bulgaria	0.211	-0.233	-0.209	Bulgaria	-0.042	-0.521	-0.010
Greece	0.212	-0.189	-0.004	Greece	-0.070	-0.415	0.079
Poland	0.262	0.298	-0.377	Poland	-0.064	0.016	-0.041
Spain	0.014	0.109	-0.142	Spain	-0.214	-0.119	0.093
Sweden	0.415	0.989	-0.155	Sweden	-0.204	0.448	0.152
Switz.	-0.039	-0.117	0.116	Switz.	-0.410	-0.395	0.421
England	-0.097	0.133	0.260	England	-0.339	-0.100	0.350
average	0.140	0.141	-0.073	average	-0.192	-0.155	0.149

	Girls				Boys		
	IMMIG sd	RETH sd	PROT sd		IMMIG sd	RETH sd	PROT sd
Bulgaria	0.764	0.775	0.838	Bulgaria	0.961	0.933	0.859
Greece	0.756	0.638	0.975	Greece	0.934	0.750	0.923
Poland	0.663	0.699	0.885	Poland	0.898	0.935	0.966
Spain	0.785	0.703	1.107	Spain	1.030	0.877	1.098
Sweden	1.166	1.040	1.092	Sweden	1.590	1.565	1.091
Switz.	0.732	0.667	0.991	Switz.	0.927	0.882	0.958
England	1.089	1.044	0.965	England	1.186	1.173	0.961
average	0.851	0.795	0.979	average	1.075	1.016	0.979

Figure 2.3: Frequency of schools with various proportions of native born students, in each of the seven countries.

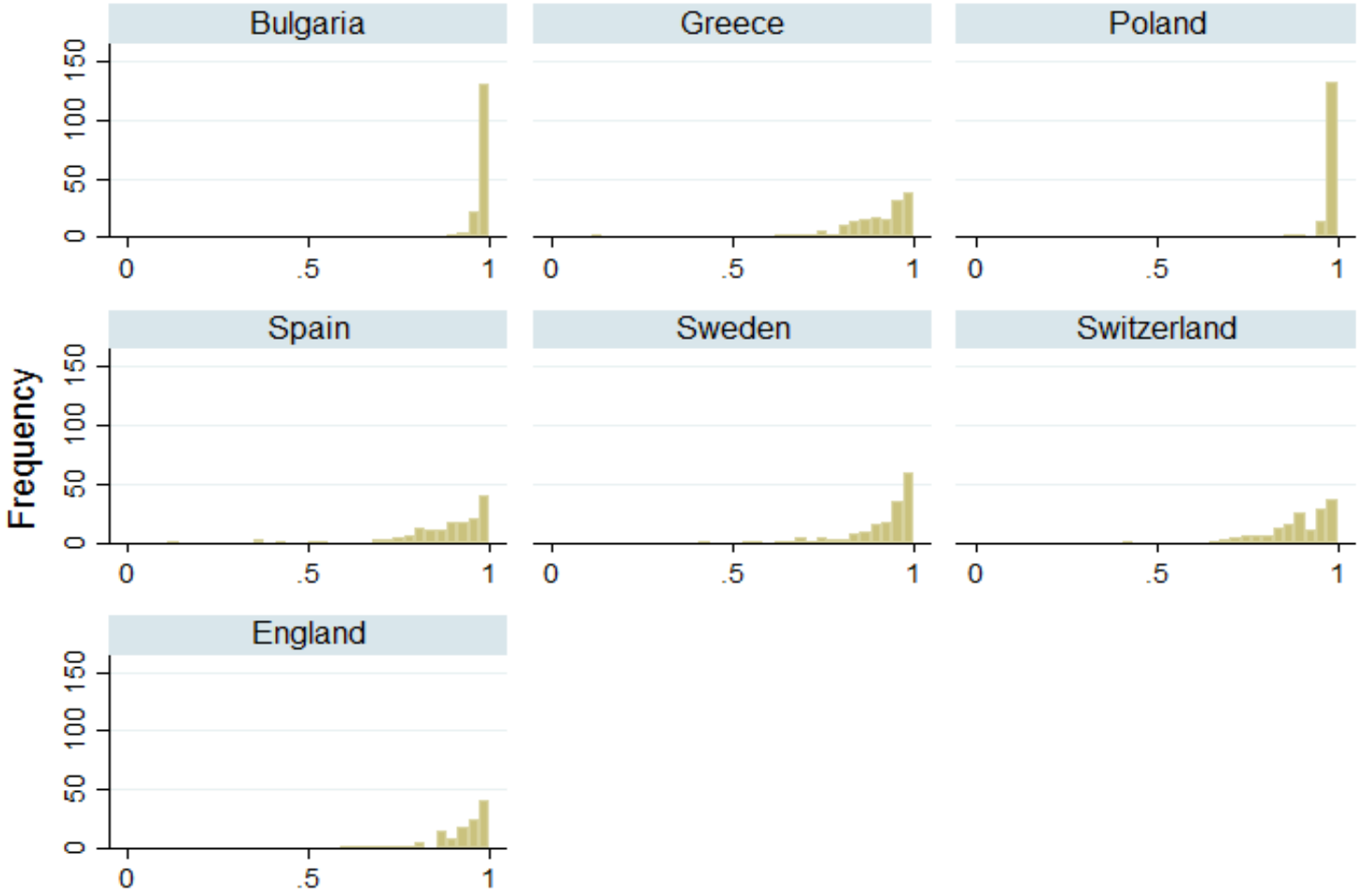


Table 2.3: Multi-level *standardized* estimates, predicting attitudes toward immigrants (IMMIG)

	Bulgaria		Greece		Poland		Spain		Sweden		Switzerland		England	
	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*
Within-level														
Expected education	0.043	***	0.034	***	0.042	***	0.047	***	0.024	***	0.038	***	0.029	***
Highest parent ed.	0.025	**	0.030	**	0.030	**	0.030	**	0.018	**	0.027	**	0.022	**
Books in the home	0.016	ns	0.015	ns	-0.030	ns	-0.023	ns	0.072	**	0.014	ns	0.064	**
Gender	0.405	**	0.34	**	0.541	***	0.602	***	0.935	***	0.367	ns	0.619	***
Eur. lang. prof.	0.027	*	0.026	*	0.026	*	0.025	*	0.014	*	0.022	*	0.020	*
Trust - people	0.024	ns	0.063	***	0.056	***	0.057	***	0.032	***	0.101	***	0.043	***
Trust - schools	0.054	***	0.061	***	0.010	ns	0.052	***	0.101	***	0.054	***	0.043	***
Stu.-tea. rel. - Scale	0.082	***	0.094	***	0.088	***	0.089	***	0.057	***	0.081	***	0.067	***
Value of part. - Scale	0.073	***	0.081	***	0.078	***	0.070	***	0.045	***	0.064	***	0.054	***
Opp. for disc. - Scale	0.047	***	0.045	***	0.049	***	0.042	***	0.028	***	0.04	***	0.039	***
Gender eq. - Scale	0.143	***	0.378	***	0.225	***	0.262	***	0.314	***	0.286	***	0.283	***
Int: Gen. and gen. eq.	-0.322	*	-0.371	**	-0.496	***	-0.600	***	-0.858	***	-0.286	ns	-0.632	***
Between-level														
Urbanicity	0.083	ns	0.064	ns	0.053	ns	0.054	ns	0.026	ns	0.034	ns	0.031	ns
% native born	0.003	ns	0.014	ns	0.004	ns	0.018	ns	0.005	ns	0.008	ns	0.006	ns
Mean trust - people	0.207	ns	0.125	ns	0.103	ns	0.122	ns	0.047	ns	0.095	ns	0.058	ns
Mean trust - schools	-0.059	ns	-0.040	ns	-0.044	ns	-0.034	ns	-0.014	ns	-0.028	ns	-0.017	ns
Mean stu.-tea. rel. - Scale	-0.096	ns	-0.062	ns	-0.052	ns	-0.064	ns	-0.031	ns	-0.046	ns	-0.030	ns
Mean value of part. - Scale	-0.514	*	0.477	**	0.283	*	0.261	~	0.060	ns	0.093	ns	0.076	ns
Mean opp. for disc. - Scale	0.528	*	-0.041	ns	-0.038	ns	-0.041	ns	-0.022	ns	-0.673	***	-0.027	ns
Mean gender eq. - Scale	0.310	~	0.298	**	0.190	**	0.253	**	0.099	**	0.199	**	0.140	**
R^2: Within	0.078		0.202		0.109		0.117		0.196		0.176		0.127	
R^2: Between	0.570		0.400		0.189		0.204		0.018		0.469		0.035	
*~p<.10, *p<.05, **p<.01 ***p<.001*														

Table 2.4: Multi-level model *standardized* estimates, predicting attitudes toward racial and ethnic minorities (RETH)

	Bulgaria		Greece		Poland		Spain		Sweden		Switzerland		England	
	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*
Within-level														
Expected education	0.042	**	0.038	**	0.037	**	0.05	**	0.023	**	0.039	**	0.028	**
Highest parent ed.	0.021	*	0.029	*	0.022	*	0.027	*	0.015	*	0.024	*	0.018	*
Books in the home	-0.014	ns	-0.015	ns	-0.012	ns	-0.013	ns	-0.013	ns	-0.012	ns	0.047	~
Gender	0.397	**	0.294	*	0.252	~	0.307	*	0.653	***	0.073	ns	0.092	ns
Eur. lang. prof.	0.023	~	0.026	~	0.020	~	0.023	~	0.012	~	0.02	~	0.016	~
Trust - people	0.076	***	0.090	***	0.063	***	0.076	***	0.039	***	0.063	***	0.052	***
Trust - schools	0.099	***	0.055	*	0.102	***	0.103	***	0.156	***	0.106	***	0.076	***
Stu.-tea. rel. - Scale	0.064	***	0.085	***	0.061	***	0.075	***	0.044	***	0.067	***	0.051	***
Value of part. - Scale	0.094	***	0.065	**	0.140	***	0.085	***	0.05	***	0.077	***	0.059	***
Opp. for disc. - Scale	0.067	***	0.074	***	0.062	***	0.064	***	0.038	***	0.060	***	0.053	***
Gender eq. - Scale	0.151	***	0.325	***	0.152	***	0.183	***	0.266	***	0.239	***	0.167	***
Int: Gen. and gen. eq.	-0.277	*	-0.311	*	-0.220	ns	-0.260	~	-0.596	***	-0.045	ns	-0.064	ns
Between-level														
Urbanicity	0.097	ns	0.126	ns	0.12	ns	0.152	ns	0.11	ns	0.079	ns	0.093	ns
% native born	-0.001	ns	-0.010	ns	-0.003	ns	-0.019	ns	-0.007	ns	-0.007	ns	-0.006	ns
Mean trust - people	0.122	ns	0.124	ns	0.118	ns	0.172	ns	0.102	ns	0.110	ns	0.087	ns
Mean trust - schools	-0.079	ns	-0.090	ns	-0.115	ns	-0.107	ns	-0.067	ns	-0.074	ns	-0.058	ns
Mean stu.-tea. rel. - Scale	0.732	***	-0.286	ns	0.157	ns	0.238	ns	0.175	ns	0.141	ns	0.118	ns
Mean value of part. - Scale	0.627	***	0.166	ns	0.192	ns	0.307	ns	0.191	ns	0.158	ns	0.413	*
Mean opp. for disc. - Scale	-0.108	ns	0.284	~	-0.106	ns	-0.14	ns	-0.114	ns	-0.431	*	-0.097	ns
Mean gender eq. - Scale	0.255	***	0.415	***	0.304	**	0.499	**	0.299	~	0.322	**	0.294	**
R^2: Within	0.119		0.176		0.135		0.133		0.168		0.155		0.109	
R^2: Between	0.686		0.466		0.233		0.597		0.213		0.27		0.41	

*~p<.10, *p<.05, **p<.01 ***p<.001*

Table 2.5: Multi-level *standardized* estimates, predicting protectionist attitudes toward migration (PROT)

	Bulgaria		Greece		Poland		Spain		Sweden		Switzerland		England	
	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*
Within-level														
Expected education	-0.060	***	-0.039	***	-0.046	***	-0.050	***	-0.039	***	-0.046	***	-0.043	***
Highest parent ed.	-0.053	***	-0.051	***	-0.049	***	-0.048	***	-0.044	***	-0.048	***	-0.049	***
Books in the home	-0.056	***	-0.042	***	-0.043	***	-0.036	***	-0.037	***	-0.041	***	-0.048	***
Gender	-0.031	ns	0.037	ns	-0.092	***	-0.061	*	-0.071	**	-0.154	***	-0.007	ns
Eur. lang. prof.	-0.033	*	-0.026	*	-0.025	*	-0.023	*	-0.020	*	-0.023	*	-0.025	*
Trust - people	0.065	**	-0.007	ns	-0.006	ns	-0.006	ns	-0.005	ns	-0.055	ns	-0.006	ns
Trust - schools	-0.018	ns	-0.010	ns	-0.010	ns	-0.008	ns	-0.007	ns	0.090	**	-0.009	ns
Stu.-tea. rel. - Scale	-0.016	ns	-0.015	ns	-0.013	ns	-0.013	ns	-0.013	ns	-0.013	ns	-0.014	ns
Value of part. - Scale	0.045	***	0.040	***	0.037	***	0.033	***	0.032	***	0.034	***	0.035	***
Opp. for disc. - Scale	0.001	ns	0.001	ns	0.001	ns	0.001	ns	0.001	ns	0.001	ns	0.001	ns
Gender eq. - Scale	-0.234	***	-0.156	***	-0.198	***	-0.210	***	-0.181	***	-0.109	***	-0.085	*
Between-level														
Urbanicity	-0.181	*	-0.173	*	-0.128	~	-0.092	*	-0.171	*	-0.118	*	-0.179	~
% native born	0.005	ns	-0.083	ns	0.331	***	-0.375	*	0.026	ns	-0.295	~	0.026	ns
Mean trust - people	0.135	ns	0.101	ns	0.075	ns	0.061	ns	0.095	ns	0.098	ns	0.100	ns
Mean trust - schools	-0.082	ns	-0.069	ns	-0.068	ns	-0.036	ns	-0.059	ns	-0.062	ns	-0.063	ns
Mean stu.-tea. rel. - Scale	0.072	ns	0.058	ns	0.158	ns	0.037	ns	-0.422	*	0.054	ns	0.058	ns
Mean value of part. - Scale	-0.040	ns	-0.031	ns	-0.024	ns	-0.018	ns	-0.035	ns	-0.028	ns	-0.038	ns
Mean opp. for disc. - Scale	-0.096	ns	-0.241	~	-0.054	ns	-0.193	ns	-0.085	ns	0.600	***	-0.088	ns
Mean gender eq. - Scale	-0.247	**	-0.295	**	-0.169	**	-0.155	**	-0.242	**	-0.250	**	-0.292	**
R^2: Within	0.100		0.038		0.084		0.079		0.064		0.061		0.024	
R^2: Between	0.192		0.274		0.231		0.325		0.195		0.362		0.298	
*~p<.10, *p<.05, **p<.01 ***p<.001*														

Appendix

Table A.2.1: Multi-level unstandardized estimates, predicting attitudes toward immigrants (IMMIG)

	Bulgaria		Greece		Poland		Spain		Sweden		Switzerland		England	
	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*
Within-level														
Expected education	0.027	***	0.027	***	0.027	***	0.027	***	0.027	***	0.027	***	0.027	***
Highest parent ed.	0.001	**	0.001	**	0.001	**	0.001	**	0.001	**	0.001	**	0.001	**
Books in the home	0.008	ns	0.008	ns	-0.015	ns	-0.013	ns	0.063	**	0.008	ns	0.038	**
Gender	0.601	**	0.467	**	0.693	***	0.867	***	2.069	***	0.536	ns	1.094	***
Eur. lang. prof.	0.018	*	0.018	*	0.018	*	0.018	*	0.018	*	0.018	*	0.018	*
Trust - people	0.020	ns	0.050	***	0.050	***	0.050	***	0.050	***	0.105	***	0.050	***
Trust - schools	0.046	***	0.046	***	0.007	ns	0.046	***	0.148	***	0.046	***	0.046	***
Stu.-tea. rel. - Scale	0.006	***	0.006	***	0.006	***	0.006	***	0.006	***	0.006	***	0.006	***
Value of part. - Scale	0.005	***	0.005	***	0.005	***	0.005	***	0.005	***	0.005	***	0.005	***
Opp. for disc. - Scale	0.003	***	0.003	***	0.003	***	0.003	***	0.003	***	0.003	***	0.003	***
Gender eq. - Scale	0.013	***	0.023	***	0.016	***	0.020	***	0.034	***	0.020	***	0.024	***
Int: Gen. and gen. eq.	-0.010	*	-0.009	**	-0.012	***	-0.015	***	-0.031	***	-0.007	ns	-0.019	***
Between-level														
Urbanicity	0.005	ns	0.005	ns	0.005	ns	0.005	ns	0.005	ns	0.005	ns	0.005	ns
% native born	0.013	ns	0.013	ns	0.013	ns	0.013	ns	0.013	ns	0.013	ns	0.013	ns
Mean trust - people	0.061	ns	0.061	ns	0.061	ns	0.061	ns	0.061	ns	0.061	ns	0.061	ns
Mean trust - schools	-0.014	ns	-0.014	ns	-0.014	ns	-0.014	ns	-0.014	ns	-0.014	ns	-0.014	ns
Mean stu.-tea. rel. - Scale	-0.002	ns	-0.002	ns	-0.002	ns	-0.002	ns	-0.002	ns	-0.002	ns	-0.002	ns
Mean value of part. - Scale	-0.011	**	0.016	**	0.011	~	0.010	ns	0.004	ns	0.004	ns	0.004	ns
Mean opp. for disc. - Scale	0.009	*	-0.001	ns	-0.001	ns	-0.001	ns	-0.001	ns	-0.020	***	-0.001	ns
Mean gender eq. - Scale	0.008	**	0.008	**	0.008	**	0.008	**	0.008	**	0.008	**	0.008	**
intraclass correlation	0.017		0.028		0.033		0.021		0.045		0.041		0.053	

*~p<.10, *p<.05, **p<.01 ***p<.001*

Table A.2.2: Multi-level unstandardized estimates, predicting attitudes toward racial and ethnic minorities (RETH)

	Bulgaria		Greece		Poland		Spain		Sweden		Switzerland		England	
	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*
Within-level														
Expected education	0.028	**	0.028	**	0.028	**	0.028	**	0.028	**	0.028	**	0.028	**
Highest parent ed.	0.001	*	0.001	*	0.001	*	0.001	*	0.001	*	0.001	*	0.001	*
Books in the home	-0.007	ns	-0.007	ns	-0.007	ns	-0.007	ns	-0.012	ns	-0.007	ns	0.030	~
Gender	0.611	**	0.361	*	0.372	~	0.425	*	1.529	***	0.104	ns	0.174	ns
Eur. lang. prof.	0.016	~	0.016	~	0.016	~	0.016	~	0.016	~	0.016	~	0.016	~
Trust - people	0.064	***	0.064	***	0.064	***	0.064	***	0.064	***	0.064	***	0.064	***
Trust - schools	0.088	***	0.037	*	0.088	***	0.088	***	0.242	***	0.088	***	0.088	***
Stu.-tea. rel. - Scale	0.005	***	0.005	***	0.005	***	0.005	***	0.005	***	0.005	***	0.005	***
Value of part. - Scale	0.007	***	0.004	**	0.011	***	0.006	***	0.006	***	0.006	***	0.006	***
Opp. for disc. - Scale	0.005	***	0.005	***	0.005	***	0.005	***	0.005	***	0.005	***	0.005	***
Gender eq. - Scale	0.014	***	0.018	***	0.012	***	0.013	***	0.031	***	0.016	***	0.015	***
Int: Gen. and gen. eq.	-0.009	*	-0.007	*	-0.006	ns	-0.006	~	-0.023	***	-0.001	ns	-0.002	ns
Between-level														
Urbanicity	0.010	ns	0.010	ns	0.010	ns	0.010	ns	0.010	ns	0.010	ns	0.010	ns
% native born	-0.010	ns	-0.010	ns	-0.010	ns	-0.010	ns	-0.010	ns	-0.010	ns	-0.010	ns
Mean trust - people	0.062	ns	0.062	ns	0.062	ns	0.062	ns	0.062	ns	0.062	ns	0.062	ns
Mean trust - schools	-0.032	ns	-0.032	ns	-0.032	ns	-0.032	ns	-0.032	ns	-0.032	ns	-0.032	ns
Mean stu.-tea. rel. - Scale	0.022	***	-0.008	ns	0.005	ns	0.005	ns	0.005	ns	0.005	ns	0.005	ns
Mean value of part. - Scale	-0.023	***	0.006	ns	0.007	ns	0.008	ns	0.007	ns	0.007	ns	0.016	*
Mean opp. for disc. - Scale	-0.003	ns	0.009	~	-0.003	ns	-0.003	ns	-0.003	ns	-0.011	*	-0.003	ns
Mean gender eq. - Scale	0.011	***	0.011	***	0.011	***	0.011	***	0.011	***	0.011	***	0.011	***
intraclass correlation	0.060		0.034		0.043		0.068		0.026		0.041		0.028	
*~p<.10, *p<.05, **p<.01 ***p<.001*														

Table A.2.3: Multi-level unstandardized estimates, predicting protectionist attitudes toward migration (PROT)

	Bulgaria		Greece		Poland		Spain		Sweden		Switzerland		England	
	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*	est.	*p*
Within-level														
Expected education	-0.039	***	-0.039	***	-0.039	***	-0.039	***	-0.039	***	-0.039	***	-0.039	***
Highest parent ed.	-0.003	***	-0.003	***	-0.003	***	-0.003	***	-0.003	***	-0.003	***	-0.003	***
Books in the home	-0.029	***	-0.029	***	-0.029	***	-0.029	***	-0.029	***	-0.029	***	-0.029	***
Gender	-0.048	ns	0.064	ns	-0.157	***	-0.118	*	-0.141	**	-0.269	***	-0.012	ns
Eur. lang. prof.	-0.023	*	-0.023	*	-0.023	*	-0.023	*	-0.023	*	-0.023	*	-0.023	*
Trust - people	0.054	**	-0.007	ns	-0.007	ns	-0.007	ns	-0.007	ns	-0.069	ns	-0.007	ns
Trust - schools	-0.016	ns	-0.010	ns	-0.010	ns	-0.010	ns	-0.010	ns	0.092	**	-0.010	ns
Stu.-tea. rel. - Scale	-0.001	ns	-0.001	ns	-0.001	ns	-0.001	ns	-0.001	ns	-0.001	ns	-0.001	ns
Value of part. - Scale	0.003	***	0.003	***	0.003	***	0.003	***	0.003	***	0.003	***	0.003	***
Opp. for disc. - Scale	0.000	ns	0.000	ns	0.000	ns	0.000	ns	0.000	ns	0.000	ns	0.000	ns
Gender eq. - Scale	-0.021	***	-0.012	***	-0.018	***	-0.021	***	-0.018	***	-0.009	***	-0.007	*
Int: Gen. and gen. eq.	-0.039	***	-0.039	***	-0.039	***	-0.039	***	-0.039	***	-0.039	***	-0.039	***
Between-level														
Urbanicity	-0.021	*	-0.021	*	-0.021	*	-0.021	*	-0.021	*	-0.021	*	-0.021	*
% native born	0.044	ns	-0.119	ns	1.934	**	-0.668	**	0.044	ns	-0.563	**	0.044	ns
Mean trust - people	-0.077	ns	-0.077	ns	-0.077	ns	-0.077	ns	-0.077	ns	-0.077	ns	-0.077	ns
Mean trust - schools	-0.037	ns	-0.037	ns	-0.037	ns	-0.037	ns	-0.037	ns	-0.037	ns	-0.037	ns
Mean stu.-tea. rel. - Scale	-0.002	ns	-0.002	ns	-0.009	ns	-0.002	ns	0.015	~	-0.002	ns	-0.002	ns
Mean value of part. - Scale	-0.002	ns	-0.002	ns	-0.002	ns	-0.002	ns	-0.002	ns	-0.002	ns	-0.002	ns
Mean opp. for disc. - Scale	-0.003	ns	-0.012	ns	-0.003	ns	-0.015	ns	-0.003	ns	0.022	***	-0.003	ns
Mean gender eq. - Scale	-0.011	***	-0.011	***	-0.011	***	-0.011	***	-0.011	***	-0.011	***	-0.011	***
intraclass correlation	0.027		0.028		0.024		0.015		0.012		0.035		0.021	
*~p<.10, *p<.05, **p<.01 ***p<.001*														

Works Cited: Introduction

Aboud, F., Tredoux, C., Tropp, L., Brown, C. S., Niens, U., & Noor, N. (2012). Interventions to reduce prejudice and enhance inclusion and respect for ethnic differences in early childhood: A systematic review. *Developmental Review, 32*, 307-336.

Alderman, L. (2012, September 30). Right-Wing extremists' popularity rising rapidly in Greece. *New York Times.* Retrieved from: http://www.nytimes.com/2012/10/01/world/europe/amid-greeces-worries-the-rise-of-right-wing-extremists.html

Allport, G. W. (1954/79). *The Nature of prejudice.* Cambridge, MA: Addison-Wesley.

Barber, C. Fennelly, K., & Torney-Purta, J. (2013). Nationalism and support for immigrants' rights among adolescents in 25 countries. *Applied Developmental Science, (2),* 60-75.

Barrett, M., & Oppenheimer, L. (2011). Findings, theories and methods in the study of children's national identifications and national attitudes. *European Journal of Developmental Psychology, 8,* 5-24.

Bekhuis, H., Ruiter, S., Coenders, M. (2013). Xenophobia among youngsters: The effect of inter-ethnic contact. *European Sociological Review, 29,* 2, 229-242.

Brewer, M. (2007). The importance of being We: Human nature and intergroup relations. *American psychologist,* 728-738.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology: Theoretical*

models of human development. (6th ed., Vol. 1, pp. 793-828). Hoboken, NJ: John Wiley & Sons Inc.

Brown, R. (1995/2010). *Prejudice: Its Social Psychology*. Sussex: Wiley-Blackwell.

CIA. (2015). *The World Fact Book*. [website] Retrieved from: https://www.cia.gov/library/publications/the-world-factbook

Coenders, M., Lubbers, M., & Scheepers, P. (2009). Opposition to civil rights for legal migrants in Central and Eastern Europe: Cross-national comparisons and explanations. *East European Politics & Societies*, *23*(2), 146-164.

Corenblum, B., & Stephan, W.G. (2001). White Fears and Native Apprehensions: An Integrated Threat Theory Approach to Intergroup Attitudes. *Canadian Journal of Behavioural Science*, 33(4), 251-268.

Dixon, J., Durrheim, K., & Thedoux, C. (2005). Beyond the optimal contact strategy: A reality check for the contact hypothesis. *American Psychologist, 60,* 697-711.

Eckert, P. (2006). In Brown, K. (Ed.) *Encyclopedia of language and linguistics.* Elsevier, 683-685

Erasmus, Z. (2010). Contact Theory: Too timid for "race" and racism. *Journal of Social Issues, 66,* 2, 387-400.

Esses, V., Dovidio, J., Jackson, L., & Armstrong, T. (2001). The immigration dilemma: The role of perceived group competition, ethnic prejudice, and national identity. *Journal of social issues, 57,* 389-412.

European Commission. (2009). Eurostat [data files]. Retrieved from http://epp.eurostat.ec.europa.eu/

Fetzer, J. S. (2000). *Public attitudes toward immigration in the United States, France, and Germany*. Cambridge ; New York : Cambridge University Press.

Freedom House. (2009). Freedom in the World [website]. Retrieved from: https://www.freedomhouse.org/

Freitag, M. & Traunmüller, R. (2009). Spheres of trust: An empirical analysis of the foundations of particularized and generalized trust. *European Journal of Political Research, 48,* 6, 782-803.

Gilligan, C. (1992). *In a different voice: Psychological theory and women's development.* Cambridge, Ms.: Harvard University Press.

Glanville, J. L., Andersson, M. A., & Paxton, P. (2013). Do social connections create trust? An examination using new longitudinal data. *Social Forces, 92*, 545-562

Glanville, J. & Paxton, P. (2007). How do we learn to trust? A Confirmatory Tetrad Analysis of the sources of generalized trust. *Social Psychology Quarterly, 70,* 3, 230-242.

Haste, H., & Abrahams, S. (2008). Morality, culture and the dialogic self: Taking cultural pluralism seriously. *Journal of Moral Education, 37,* 3, 377-394.

Higgins, A. (2013, December 22). Bulgaria, unready, is poor host to Syrians. *New York Times*. Retrieved from: http://www.nytimes.com/2013/12/23/world/europe/bulgaria-unready-is-poor-host-to-syrians.html?_r=0

Hopkins, N., Reicher, S., Levine, M. (1997). On the parallels between social cognition and the "new racism." *British Journal of Social Psychology, 36,* 305-329.

Huijnk, W., Verkuyten, M., & Coenders, M. (2012). Family relations and attitude toward ethnic minorities as close kin by marriage. *Ethnic and Racial Studies, 36,* 11, 1890-1909.

International Association for the Evaluation of Educational Achievement. (2009). *The IEA International Civic and Citizenship Education Study (ICCS) 2009.* [data files and codebooks: http://rms.iea-dpc.org/] The Australian Council for Educational Research (ACER).

Kirkup, J. (2002). Identity, community, and distributed learning. In M. Lea & K. Nicholl (Eds.) *Distributed learning.* London: Routledge, 182-195.

Lave, J., & Wenger, E. (2002). Legitimate peripheral participation in communities of practice. In M. Lea & K. Nicholl (Eds.) Distributed learning. London: Routledge, 56-63.

Migrant Integration Policy Index (MIPEX). (n.d.) [website] retrieved from: http://www.mipex.eu/

Minkenberg, M. (2011). *The radical right in Europe: An overview.* E-Book-Ausgabe (PDF) Verlag Bertelsmann Stiftung. Retrieved from: http://www.bertelsmann-stiftung.de.

Moree, D., Klaassen, C., & Veugelers, W. (2008). Teachers' Ideas about Multicultural Education in a Changing Society: The Case of the Czech Republic. *European Educational Research Journal, 7*(1), 60-73.

Noddings, N. (2005). *The challenge to care in schools: An alternative approach to education.* New York, NY: Teachers College Press.

NSD: European Election Database. (n.d.) [website] Retrieved from: http://www.nsd.uib.no/european_election_database/

Pettigrew, T. (1998). Intercultural contact theory. *Annual Review of Psychology*, 49, 65-85.

Pettigrew, T., & Tropp, L. (2006). A meta-analytic test of intercultural contact theory. *Journal of Personality and Social Psychology, 37* (4), 820-834.

Pettigrew, T., & Tropp, L. (2008). How does intercultural contact reduce prejudice?

Putnam, R. (2000). *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster.

Raabe, T., Beelman, A. (2011). Development of ethnic, racial, and national prejudice in childhood and adolescence: A multinational meta-analysis of age differences. *Child Development, 82,* 1715-1737.

Republic of Bulgaria: National Statistical Institute (NIS-BR). (2011). *Population Census of 2011.* Retrieved from: http://www.nsi.bg/en/content/6593/population-and-demographic-processes.

Reicher, S. (2007). Rethinking the paradigm of prejudice. *South African Journal of Psychology, 35*, 412-432.

Rogoff, B. (2003). *The Cultural Nature of Human Development*. Oxford: Oxford University Press.

Stephan, W. G., Renfro, C. L., Esses, V. M., Stephan, C. W. & Martin, T. (2005). The effects of feeling threatened on attitudes toward immigrants. *International journal of intercultural relations,* 29: 1–19.

Stephan, W., Ybarra, O., & Bachman, G. (2006). Prejudice toward immigrants. *Journal of applied social psychology, 29* (11)*,* 2221-2237.

Super, C., & Harkness, S. (1986). The developmental niche: A conceptualization at the interface of child and culture. *International Journal of Behavioral Development*, 9, 545-569.

Super, C., & Harkness, S. (2002). Culture structures the environment for development. *Human Development*, 45, 270-274.

Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In Austin, W., & Worchel, S. (Eds.), The social psychology of intergroup relations. Toronto: Brooks/Cole (33-47).

Torney-Purta, J. (2009). International psychological research that matters for policy and practice. *American Psychologist, 64*, 825-837.

Torney-Purta, J. (2013). A psychological view of civic engagement augment by introducing the developmental niche model: Essay review of Teenage Citizens: The Political Theories of the Young by Constance Flanagan. *Human Development*, 56, 346-350.

Torney-Purta, J. & Amadeo, J. (2011). Participatory niches for emergent citizenship in early adolescence: An international perspective. *The Annals of the American Academy. 633,* 180-200.

Torney-Purta, J., & Barber, C. (2011). Fostering young people's support for participatory human rights through their developmental niches. *American Journal of Orthopsychiatry, 81*, 473-481.

Torney-Purta, J., Lehmann, R., Oswald, H., & Schulz, W. (2001). *Citizenship and education in twenty-eight countries: Civic knowledge and engagement at age fourteen.* Amsterdam: IEA.

UNICEF Serbia (2005). Breaking the cycle of exclusion: Roma children in South East Europe. http://www.unicef.org/ceecis/070305-Subregional_Study_Roma_Children.pdf

Vygotsky, L. (1987). *Thinking and speech.* In R. K. Fullinwider (Ed.), *The collected works of L S. Vygotsky* (N. Minick. Trans.). New York: Plenum.

Wenger, E. (1998/2008). *Communities of practice: learning, meaning, and identity*. Cambridge: Cambridge University Press.

World Bank. (n.d.) [website] Retrieved from: http://data.worldbank.org/indicator/SG.GEN.PARL.ZS

Works Cited: Paper 1

Aboud, F., Tredoux, C., Tropp, L., Brown, C. S., Niens, U., & Noor, N. (2012). Interventions to reduce prejudice and enhance inclusion and respect for ethnic differences in early childhood: A systematic review. *Developmental Review, 32*, 307-336.

Alderman, L. (2012, September 30). Right-Wing extremists' popularity rising rapidly in Greece. *New York Times*. Retrieved from:

Asparouhov, T. (2012, May 11). Multiple Imputation. Message posted to: http://www.statmodel.com/discussion/messages/22/381.html

Asparouhov, T. & Muthén, B. (2013, Oct. 1). Multiple-group Factor Analysis Alignment [PDF document]. Retrieved from http://www.statmodel.com/examples/webnotes/webnote18.pdf

Asparouhov, T. & Muthén, B. (2010, Sept. 29). Bayesian analysis using Mplus: Technical implementation [PDF document]. Retrieved from http://www.statmodel.com/download/Bayes3.pdf

Brese, F., Jung, M., Mirazchiyski, P., Schulz, W., & Zuehlke, O. (2011). *ICCS 2009 user guide for the International Database.* Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA). (http://www.iea.nl/iccs_2009.html)

Campbell, D. & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56(2),* 81-105.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297-334.

Crocker, J. & Algina, J. *Introduction to Classical and Modern Test Theory.* Mason, OH: Cengage.

European Commission. (2009). Population and social conditions. *Data in Focus*, 31, 1-11.

Higdon, J. (2013). National identity, European identity, and attitudes toward migration and immigration among adolescents (Unpublished Qualifying Paper). Harvard Graduate School of Education, Cambridge, Ma.

Higgins, A. (2013, December 22). Bulgaria, unready, is poor host to Syrians. *New York Times.* Retrieved from: http://www.nytimes.com/2013/12/23/world/europe/bulgaria-unready-is-poor-host-to-syrians.html?_r=0

Husfelt, V. & Torney-Purta, J. (2014). Development of the CIVED instruments. In Schulz, W., & Sibberns, H. (Eds). *IEA Civic Education Study Technical Report.* Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

International Association for the Evaluation of Educational Achievement. (2009). *The IEA International Civic and Citizenship Education Study (ICCS) 2009.* [data files and codebooks: http://rms.iea-dpc.org/] The Australian Council for Educational Research (ACER).

Kaplan, D. & Depaoli, S. (2012). Bayesian structural equation modeling. Chapter 38 In Hoyle, R (Ed.) *Handbook of Structural Equation Modeling*: The Guilford Press.

Kaplan, D. & Depaoli, S. (2013). Bayesian statistical methods. Chapter 20 in Little, T. (Ed.) *The Oxford Handbook of Quantitative Methods.* Oxford: Oxford University Press.

Kerr, D., Sturman, L., Schulz, W., & Burge, B. (2010). *ICCS 2009 European report: Civic knowledge, attitudes, and engagement among lower-secondary students in 24 European countries*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Khanec, M., & Tosun, M. (2009). Political economy of immigration in Germany: Attitudes and citizenship aspirations. *International Migration Review, 43*(2), 263-291.

Kline, R. (2011).*Principles and practice of structural equation modeling, 3rd Ed.* New York: The Guilford Press.

Lance, C., Butts, M., & Michels, L. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods, 9,* 2, 202-220.

Lee, J. (2013). An application of Bayesian structural equation modeling: The sensitivity of Bayesian analysis for item-level measurement invariance. *Journal of Educational Evaluation, 26* (4), 719-739.

Migrant Integration Policy Index (MIPEX). (n.d.) [website] retrieved from: http://www.mipex.eu/

Ministry of Integration and Gender Equality. (2009). Fact sheet: New Anti-discrimination legislation and a new agency, the Equality Ombudsman. Stockholm, Sweden. Available online: http://www.legislationline.org/documents/id/16393

Minkenberg, M. (2011). *The radical right in Europe: An overview.* E-Book-Ausgabe (PDF) Verlag Bertelsmann Stiftung. Retrieved from: http://www.bertelsmann-stiftung.de.

Muthén, B.O. & Asparouhov, T. (2012). Bayesean SEM: A more flexible representation of substantive theory. *Psychological Methods, 17,* 313-335.

Muthén, B. O., & Asparouhov, T. (2013). BSEM Measurement Invariance Analysis. *Mplus Web Notes: No. 17.* Retrieved from: http://www.statmodel.com/examples/webnotes/webnote17.pdf

Muthén, L. K., & Muthén, B. O. (1998-2014). *Mplus User's Guide.* Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika, 44,* 443-460.

R Foundation for Statistical Computing. (2015). R. [Software]. Vienna, Austria.

Raabe, T., Beelman, A. (2011). Development of ethnic, racial, and national prejudice in childhood and adolescence: A multinational meta-analysis of age differences. *Child Development, 82,* 1715-1737.

Schulz, W., Ainley, Ainley, J. & Fraillon, J. (2011). *ICCS 2009 technical report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA). (http://www.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/ICCS_2009_Technical_Report.pdf)

Schulz, W., Ainley, J., Fraillon, J., Kerr, D., & Losito, B. (2009). *ICCS 2009 international report: Civic knowledge, attitudes and engagement among lower secondary school students in thirty-eight countries.* Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Schulz, W., & Sibberns, H. (2004). *IEA Civic Education Study Technical Report.* Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

StataCorp. 2013. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.

Steenkamp, J. M. & Baumgartner, H. (1998). Assessing measurement invariance in cross national consumer research. *Journal of Consumer Research, 25,* 78-90.

Torney-Purta, J., Lehmann. R., Oswald, H., & Schulz, W. (2001). *Citizenship and education in twenty-eight countries: Civic knowledge and engagement at 14.* Amsterdam, The Netherlands: IEA.

Torney-Purta, J., Schwille, J., & Amadeo, J. A. (1999). *Civic education across countries: Twenty-four national case studies for the IEA Civic Education Project.* Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J., Neyer, F., van Aken, M. (2013). A gentle introduction to Bayesian analysis: Applications to developmental research. *Child Development, 85,* 3, 842-860.

Van de Schoot, R., Kluytmans, A., Tummers, L., Lugtig., P., Hox, J., & Muthén, B. (2013). Facing off with Scylla and Charybdis: a comparison of scalar, partial, and the novel possibility of approximate measurement invariance. *Frontiers in Psychology: Quantitative Psychology and Measurement, 4,* 1-15.

Van de Vijver, F. J. R., & Poortinga, Y. H. (2002). Structural equivalence in multilevel research. *Journal of Cross-Cultural Psychology, (2),* 141.

Zyphur, M & Oswald, F. (2013) Bayesian Estimation and Inference: A User's Guide. *Journal of Management,* published online: http://jom.sagepub.com/content/early/2013/08/11/0149206313501200

Works cited: Paper 2

Abdi, H. (2007). The Bonferonni and Šidák corrections for multiple comparisons. In: N. Salkind (Ed.) *Encyclopedia of measurement and statistics.* Thousand Oaks: Sage Publications.

Aboud, F., Tredoux, C., Tropp, L., Brown, C. S., Niens, U., & Noor, N. (2012). Interventions to reduce prejudice and enhance inclusion and respect for ethnic differences in early childhood: A systematic review. *Developmental Review, 32*, 307-336.

Allport, G. W. (1954/79). *The Nature of prejudice.* Cambridge, MA: Addison-Wesley.

Asparouhov, T. & Muthén, B. (2012). Multiple group multilevel analysis. *Mplus Web Notes: No. 16.* http://www.statmodel.com/examples/webnotes/webnote16.pdf

Barber, C. Fennelly, K., & Torney-Purta, J. (2013). Nationalism and support for immigrants' rights among adolescents in 25 countries. *Applied Developmental Science, (2),* 60-75.

Barber, C., Torney-Purta, J., Wilkenfeld, B., Ross, J. (2015). Immigrant and native-born adolescents' civic knowledge and attitudes in Sweden and the United States: Emergent citizenship within developmental niches. *Research in Comparative & International Education,* 1-25.

Bennett, M., Barrett, M., Karakozov, R., Kipiani, G., Lyons, E., Pavlenko, V., & Riazanova, T. (2004). Young children's evaluations of the ingroup and of outgroups: A multi-national study. *Social Development, 13*(1), 124-141.

Brese, F., Jung, M., Mirazchiyski, P., Schulz, W., & Zuehlke, O. (2011). ICCS 2009 user guide for the International Database. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA). (http://www.iea.nl/iccs_2009.html)

Brewer, M. (2007). The importance of being We: Human nature and intergroup relations. *American psychologist,* 728-738.

Brown, R. (1995/2010). *Prejudice: Its Social Psychology*. Sussex: Wiley-Blackwell.

Coenders, M., & Scheepers, P. (2002). The effect of education on nationalism and ethnic exclusionism: And international comparison. *Political Psychology, 24,* 2, 313-343.

Coenders, M., Lubbers, M., & Scheepers, P. (2009). Opposition to civil rights for legal migrants in Central and Eastern Europe: Cross-national comparisons and explanations. *East European Politics & Societies*, *23*(2), 146-164.

Corenblum, B., & Stephan, W.G. (2001). White Fears and Native Apprehensions: An Integrated Threat Theory Approach to Intergroup Attitudes. *Canadian Journal of Behavioural Science*, 33(4), 251-268.

Enders, C. (2010). *Applied missing data analysis.* New York: The Guilford Press.

Europa. (2010). *The Charter of Fundamental Rights.* Summaries of EU Legislation. Retrieved from: http://eur-lex.europa.eu.

European Commission. (2009). Eurostat [data files]. Retrieved from http://epp.eurostat.ec.europa.eu/

European Commission. (2015). Eurostat [data files]. Retrieved from http://epp.eurostat.ec.europa.eu/

European Commission. (2009). Population and social conditions. *Data in Focus*, 31, 1-11.

Evans, G. & Need, A. (2002). Explaining ethnic polarization over attitudes toward minority rights in Eastern Europe: a multilevel analysis. *Social Science Research, 31,* 653-680.

Gilligan, C. (1992). *In a different voice: Psychological theory and women's development.* Cambridge, Ms.: Harvard University Press.

Gorodzeisky, A., & Semyonov, M. (2009). Terms of exclusion: public views toward admission and allocation of rights to immigrants in European countries. *Ethnic and Racial Studies, 32,* 3, 401-423.

Haste, H., & Abrahams, S. (2008). Morality, culture and the dialogic self: Taking cultural pluralism seriously. *Journal of Moral Education, 37,* 3, 377-394.

Hello, E., Scheepers, P., Gijsberts, M. (2002). Education and ethnic prejudice in Europe: Explanations for cross-national variances in the educational effect on ethnic prejudice. *Scandinavian Journal of Educational Research, 46,* 1-24.

Husfeldt, V. (2006). Extreme negative attitudes toward immigrants: an analysis of factors in five countries. *Prospects, 34,* 3, 355-374.

International Association for the Evaluation of Educational Achievement. (2009). *The IEA International Civic and Citizenship Education Study (ICCS) 2009*. [data files and codebooks: http://rms.iea-dpc.org/] The Australian Council for Educational Research (ACER).

Isac, M., Maslowski, R., Creemers, B., & van der Werf, G. (2013). The contribution of schooling to secondary-school students' citizenship outcomes across countries. *School Effectiveness and Improvement,* Jan. 2013, 1-35.

Khanec, M., & Tosun, M. (2009). Political economy of immigration in Germany: Attitudes and citizenship aspirations. *International Migration Review, 43*(2), 263-291.

Lave, J., & Wenger, E. (2002). Legitimate peripheral participation in communities of practice. In M. Lea & K. Nicholl (Eds.) Distributed learning. London: Routledge, 56-63.

Migrant Integration Policy Index (MIPEX). (n.d.) [website] retrieved from: http://www.mipex.eu/

M*plus* (Version 7.3). (2014). [Computer Software]. Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus User's Guide*. Edition 7.3. Los Angeles, CA: Muthén & Muthén.

Pettigrew, T. (1998). Intercultural contact theory. *Annual Review of Psychology*, 49, 65-85.

Raabe, T., Beelman, A. (2011). Development of ethnic, racial, and national prejudice in childhood and adolescence: A multinational meta-analysis of age differences. *Child Development, 82,* 1715-1737.

Raudenbush, S. & Bryk, A. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods, (2nd Ed.)*. Thousand Oaks: Sage Publications.

Schulz, W., Ainley, J., Fraillon, J., Kerr, D., & Losito, B. (2009). *ICCS 2009 international report: Civic knowledge, attitudes and engagement among lower secondary school students in thirty-eight countries*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Schulz, W., Ainley, J. & Fraillon, J. (2011). *ICCS 2009 technical report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA). (http://www.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/ICCS_2009_Technical_Report.pdf)

Schulz, W., & Sibberns, H. (2004). *IEA Civic Education Study Technical Report.* Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Šidák, Z. K. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association, 62* (318): 626–633

StataCorp. 2013. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.

Stephan, W. G., Renfro, C. L., Esses, V. M., Stephan, C. W. & Martin, T. (2005). The effects of feeling threatened on attitudes toward immigrants. *International journal of intercultural relations,* 29: 1–19.

Stephan, W., Ybarra, O., & Bachman, G. (2006). Prejudice toward immigrants. *Journal of applied social psychology, 29* (11)*,* 2221-2237.

Super, C., & Harkness, S. (1986). The developmental niche: A conceptualization at the interface of child and culture. *International Journal of Behavioral Development*, 9, 545-569.

Super, C., & Harkness, S. (2002). Culture structures the environment for development. *Human Development*, 45, 270-274.

Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In Austin, W., & Worchel, S. (Eds.), The social psychology of intergroup relations. Toronto: Brooks/Cole (33-47).

Torney-Purta, J. & Amadeo, J. (2011). Participatory niches for emergent citizenship in early adolescence: An international perspective. *The Annals of the American Academy. 633,* 180-200.

Torney-Purta, J., & Barber, C. (2011). Fostering young people's support for participatory human rights through their developmental niches. *American Journal of Orthopsychiatry, 81*, 473-481.

Torney-Purta, J., Lehmann, R., Oswald, H., & Schulz, W. (2001). *Citizenship and education in twenty-eight countries: Civic knowledge and engagement at age fourteen.* Amsterdam: IEA.

Torney-Purta, J., Schwille, J., & Amadeo, J. A. (1999). *Civic education across countries: Twenty-four national case studies for the IEA Civic Education Project.* Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Van de Vijver, F. J. R., & Poortinga, Y. H. (2002). On the study of culture in developmental science. *Human Development, 45,* 246-256.

Works cited: Conclusion

Haste, H. & Abrahams, S. (2008) Morality, culture and the dialogic self: taking cultural pluralism seriously. *Journal of Moral Education, 37*(3), 357-374.

Torney-Purta, J. & Amadeo, J. (2011). Participatory niches for emergent citizenship in early adolescence: An international perspective. *The Annals of the American Academy. 633,* 180-200.

Torney-Purta, J., & Barber, C. (2011). Fostering young people's support for participatory human rights through their developmental niches. *American Journal of Orthopsychiatry, 81*, 473-481.

VITA

Julia Dianne Higdon

1992-1995	Radford University Radford, Virginia	
1995-1996	University of North Carolina at Chapel Hill Chapel Hill, North Carolina	B.A. December 1996
2001-2003	University of North Carolina at Greensboro Greensboro, North Carolina	M.Ed. May 2003
2003-2008	Montgomery County, MD; Durham County, NC; Guilford County, NC Teacher in grades 3-6	
2008-2015	Doctor of Education Candidate Graduate School of Education Harvard University	Ed.M. IEP, Masters-in-passing May 2012
2009-2015	Teaching Fellow Graduate School of Education Harvard University	