EÖTVÖS LORÁND UNIVERSITY FACULTY OF EDUCATION AND PSYCHOLOGY

Hadi Sam Nariman (Hadi Ghazi Altabatabaei)
Morality, Group Dynamics, and Inter- and Intragroup Attitudes

Doctoral School of Psychology

Head of the School: Dr. Róbert Urbán, Eötvös Loránd University

Socialisation and Psychology of Social Processes

Head of the Program: Dr. Lan Anh Nguyen Luu, Eötvös Loránd University

Supervisors:

Dr. Lan Anh Nguyen Luu, Eötvös Loránd University

Dr. Márton Hadarics, Eötvös Loránd University

Budapest, 2023

Table of Contents

1   Acknowledgments

I would like to express my heartfelt appreciation to my supervisor, **Professor Lan Anh Nguyen Luu**, for her exceptional support and guidance throughout my doctoral studies. From supervising my master's studies to the completion of my PhD, her expertise, kindness, and encouragement have been pivotal in my academic journey. Without her unwavering support over the course of many years, I would not have been able to embark on my PhD journey, let alone complete it.

I am incredibly grateful to my supervisor, Professor **Márton Hadarics**, for his invaluable feedback, guidance, and support throughout my PhD journey. I worked closely with him during my PhD, and without his patience, vast academic and research knowledge, and generosity, I would not have been able to complete this PhD. His constructive criticism and constant encouragement have significantly influenced the direction of my research and improved the quality of my work. I will always appreciate the knowledge and skills that I have gained under his tutelage.

I would like to express my sincere gratitude to **Professor Levente Littvay**, who was not my official supervisor but played an invaluable role in shaping my academic pursuits during my PhD studies. His expertise in research methods and data analysis was instrumental in my research journey. In addition to his expertise and guidance, I would also like to express my deep appreciation for his exceptional generosity, both academically and otherwise, which made my doctoral studies a truly enriching experience.

I would like to express my sincerest gratitude to **Professor Anna Kende** for her altruistic support and guidance in my academic journey. Her immense insights and expertise have been invaluable in shaping the success of my research. I have gained a wealth of knowledge and personal growth under her tutelage, both in academia and in life. Professor Kende's outstanding proficiency in the field and boundless generosity, both professionally and otherwise, serve as a constant source of inspiration to me.

Dedicated to Jina Amini

For Woman, Life, Liberty

2   List of Publications Directly Used in the Dissertation

The dissertation follows the formal requirements of Type B format.

Study 1: Sam Nariman, H., Hadarics, M., Soufizadeh, A. M., & Kende, A. (2020). The mediating role of moral exclusion between authoritarianism and outgroup discrimination. *International Journal of Intercultural Relations*, *74*, 1-6. https://doi.org/10.1016/j.ijintrel.2019.10.001

Study 2: Sam Nariman, H., Hadarics, M., Kende, A., Lášticová, B., Poslon, X. D., Popper, M., ... & Minescu, A. (2020). Anti-Roma bias (stereotypes, prejudice, behavioral tendencies): A network approach toward attitude strength. *Frontiers in psychology*, *11*, 2071. https://doi.org/10.3389/fpsyg.2020.02071

Study 3: Sam Nariman, H., Nguyen Luu, L.A., & Hadarics, M. (2021). Exploring inclusiveness towards immigrants as related to basic values: A network approach. *Plos one*, *16*, e0260624. https://doi.org/10.1371/journal.pone.0260624

Study 4: Sam Nariman, H., Hadarics, M., Mehrez, Nguyen Luu, L.A., Soufizadeh, A., & Littvay, L. (2022). Immigrants' intragroup moral exclusion predicts ingroup-directed behavioral intentions: The mediating role of disidentification. *International Journal of Intercultural Relations*, *91*, 113-118. https://doi.org/10.1016/j.ijintrel.2022.09.006

3   Introduction

Taking an integrative approach, the findings from these studies highlight the importance of investigating the moral dimensions of group-based attitudes and the complex interactions between these dimensions. They offer insights into a variety of contexts, including the treatment and perception of Roma and Jewish minorities in Hungary, attitudes towards immigrants across Europe, and experiences of Iranian and Tunisian immigrants in Italy. The dissertation investigates the psychological mechanisms underlying intergroup conflict and discrimination, such as right-wing authoritarianism, social dominance orientation, moral exclusion, perceived threat to national identity, disidentification, and basic values. The studies collectively reveal how these mechanisms manifest across various social and cultural contexts, and how they influence group-based attitudes and behaviors. In an era marked by identity politics, it is crucial to deepen our understanding of these phenomena, their origins, and potential solutions. To this end, this dissertation contributes to the ongoing discussion on ways to reduce discrimination.

The overarching scope of this dissertation was to examine the complex interplay between group-based attitudes and behavioral tendencies, with a particular focus on their moral dimensions. By "moral dimensions", I mean value-laden intuitive judgements that I quantitatively investigated in their interrelationship with related group-based evaluations. In Study 1 and Study 4 the role of moral exclusion was investigated, in Study 2 the strong interconnection between perceived threat (to national identity) and other anti-Roma evaluations was revealed, and in Study 4 the relationship between basic values and outgroup attitudes was explored. The ultimate aim was to enhance our understanding of these attitudes and inform interventions that reduce discrimination.

Given the significance of the concept of moral exclusion in this dissertation, before delving further, I aim to distinguish between the concept of moral exclusion (e.g., Deutsch, 1990; Opotow, 1990; Staub, 1990) and other related psychological phenomena such as dehumanization, moral disengagement, and deservingness, which often appear in intergroup contexts. Although all these mechanisms can lead to discrimination and negative group-based attitudes, they operate based on distinct underlying processes. Moral exclusion entails the denial of moral considerations to particular groups or individuals, rendering them outside the scope of justice. This allows dominant groups to justify harm or unequal treatment towards those excluded. On the other hand, dehumanization (Haslam, 2006) is the process by which certain groups are perceived as less than human. This perception can lead to moral exclusion,

or the two can influence each other reciprocally. But the distinction lies in the fact that while moral exclusion involves the denial of ethical considerations, dehumanization focuses on denying basic human qualities to the group in question. Moral disengagement (Bandura et al., 1996) involves mechanisms that allow individuals to detach their actions from their moral standards. In the context of intergroup relations, it can facilitate discriminatory or prejudiced actions by allowing the individual to psychologically suspend their moral compass when dealing with certain groups. The difference is that moral exclusion is based on perceiving others as falling short of basic moral standards, whereas disengagement involves strategies that diminish one's sense of moral responsibility for harmful actions. Perceived deservingness (see Olson et al., 2011) encompasses the judgment that if the negative treatment received by the "other" is perceived as deserved, this perception correlates with perceived fairness. Here, the perception of deservingness doesn't necessarily lead to moral exclusion, but moral exclusion can arise from the perceived deservingness of another group or individual. So the four concepts highly overlap but are not identical.

In Study 1, we examined the effects of right-wing authoritarianism (RWA) and social dominance orientation (SDO) on negative intergroup behaviors, focusing on the potential mediating role of moral exclusion within these relationships. By integrating the Dual Process Model (Duckitt, 2001) with research on moral exclusion, we aimed to delve deeper into the underlying mechanisms that contribute to individuals' endorsement of negative intergroup beliefs and behaviors. Our study emphasizes the need to better understand the dynamics of intergroup conflict and discrimination, particularly in the context of the Hungarian Roma and Jewish minorities. We utilized a dataset of survey responses (N = 1015) collected from a representative Hungarian sample to investigate the mediating role of moral exclusion in the relationship between RWA and SDO and the discriminatory intentions they produce against the Roma and Jewish communities in Hungary. Our findings revealed that moral exclusion indeed played a mediating role in the relationship between both RWA and SDO and negative behavioral intentions towards the Hungarian Roma and Jewish minorities. Furthermore, we proposed that the concept of moral exclusion should not be regarded as a generalized tendency, but rather as a mechanism that stems from various social cognitive motivations, including RWA and SDO. This study offers a more nuanced understanding of the ways in which moral exclusion interacts with authoritarian personality traits to contribute to negative intergroup behaviors. Additionally, we highlight the importance of examining moral exclusion in different contexts and target groups to better understand the dynamics of

intergroup conflict and discrimination. In light of our findings, we underscore the significance of further research into the relationships between RWA, SDO, moral exclusion, and intergroup behaviors. Such research can facilitate the development of targeted interventions aimed at reducing discrimination, and fostering positive intergroup relations.

Study 2 utilized a correlational network approach to explore the potential for interventions targeting anti-Roma bias in Hungary, Romania, Slovakia, France, and Ireland. These countries represent a mix of Eastern and Western European contexts with significant Roma populations, both indigenous and recent immigrants. Employing network analysis, we investigated the connectivity between a constellation of stereotypical, emotional, and behavioral evaluations of the Roma population in representative samples from the selected countries. By identifying the variables (nodes) with the highest degree of interactions within these networks, we aimed to pinpoint the most influential variables for intervention purposes. Threat perception to national identity was found to be the most important factor having strongest connections with other Roma-related evaluations. Lastly, it is noteworthy that the perceived threat, identified as the most central variable, maintains a significant relationship with one's moral considerations, and it can function as a precursor to moral exclusion (e.g., Olson et al., 2011). We also know that perceived threat to national identity is strongly related to the moral exclusion of the Roma (Hadarics, 2020). This result offers a potentially promising theoretical basis for intervention strategies that focus on central variables with considerable associations with other variables, although further research is needed to confirm its effectiveness. Furthermore, the network analysis revealed that feelings of sympathy and empathy, were also among most central variables within the networks.

Study 3 explored the relationship between Europeans' basic values and their attitudes towards immigrants, employing Latent Class Analysis (LCA) and network analysis, using the 9th round of European Social Survey (ESS). We examined participants' level of support for permitting three categories of immigrants to enter and reside in their countries: immigrants belonging to the same ethnic group, immigrants from different ethnic groups, and immigrants originating from economically disadvantaged countries outside Europe. Our analysis revealed the presence of four distinct categories of Europeans, each displaying a distinct set of attitudes towards immigrants. These categories were labeled as Inclusive (displaying high levels of inclusivity), Some (being selective in their acceptance), Few (being highly selective in their acceptance), and Exclusive (displaying high levels of exclusivity). Subsequently, we estimated correlational networks to investigate the relationships between Schwartz's (1992)

ten basic values for each of the four groups. We compared the networks of the four groups by analyzing the extent to which the values were interconnected, how the values clustered together, and the strength of connections between values with similar motivational backgrounds. Our analysis indicated that, while the overall connections between the 10 basic values were largely similar across the four groups, there were some notable differences in terms of how the values clustered and the strength of connections between specific value pairs. Interestingly, we found a more intricate arrangement of values among the most inclusive Europeans, meaning that they differentiated and integrated broader range of basic values (openness to change, conservation, self-transcendence, and self-enhancement) than their less inclusive groups. Furthermore, our analysis unveiled that the highly inclusive Europeans exhibited stronger associations between values that had similar motivational backgrounds. We discussed the practical implications of our findings, highlighting the importance of considering the complexity of value structures, when developing interventions aimed at reducing discrimination in European societies.

Study 4 aimed to investigate the relationship between moral exclusion, disidentification, and ingroup-directed behavioral intentions among Iranian and Tunisian immigrants living in Italy. While previous research has extensively examined negative attitudes towards immigrants among majority members, less is known about whether moral exclusion is also directed towards ingroup members, particularly by immigrants against their own group. Our argument was that moral exclusion, which justifies negative treatment of individuals or groups outside of one's moral scope, can also be directed towards the ingroup members. In other words, we proposed that moral exclusion is not limited to outgroup members and can also be applied to individuals within one's own group. We also hypothesized that disidentification with one's ethnic identity mediates this relationship, as social identification becomes painful and threatening when it does not fulfill an individual's psychological needs. Our findings suggested that moral exclusion is associated with negative ingroup-directed behavioral intentions and negatively associated with positive ingroup-directed behavioral intentions. Furthermore, the results showed that disidentification with immigrants' own ethnic group mediates the relationship between moral exclusion and ingroup-directed behavioral intentions. Our results underscore the significance of taking into account the moral dimensions of disidentification in the analysis of group-based relations. While not explicitly addressed in the paper due to space constraints, interventions could be aimed at mitigating the affective/moral expressions of disidentification to reduce its negative

consequences. In our discussion of future directions, we suggested system justification as one such negative consequence that may strongly associate with disidentification and intragroup moral exclusion.

In the following sections, a thorough examination of each study, delving into the intricacies of their methodologies, findings, and implications will be provided. Note that the studies are incorporated into this dissertation in their original form as they were published in their respective journals.

4 Study I: The Mediating Role of Moral Exclusion between Authoritarianism and Outgroup Discrimination

4.1 Abstract

It has been well-documented that right wing authoritarianism and social dominance orientation as two facets of the authoritarian personality differentially account for a variety of negative intergroup behaviors. Integrating the Dual Process Model (Duckitt, 2001), with the literature on "Moral exclusion" (e.g., Opotow, 1990; Tileagă, 2007) we investigated whether or not moral exclusion would mediate such a relationship. Employing survey data (N = 1015), collected from a representative Hungarian sample, we found that moral exclusion mediated the effects of both RWA and SDO on the negative behavioral intentions against Roma as well as Jewish minorities in Hungary. Moreover, we argued that the concept of moral exclusion should be interpreted not as a generalized tendency, but as a mechanism which can be stemmed from distinct social cognitive motivations.

Keywords: Moral exclusion; Social dominance orientation; Right-wing authoritarianism; Intergroup behavior

4.2 Introduction

Being more effective to combating conflictual intergroup relations in a society requires delving into the underlying mechanisms that are entangled with individuals' endorsement of negative intergroup beliefs and behaviors. Moral exclusion is one such mechanism. As Opotow (1990) noted, "moral exclusion occurs when individuals or groups are perceived as outside the boundary within which moral values, rules, and considerations of fairness apply" (p.1). Indeed, a reprehensible intergroup conduct—in the eyes of an outsider—can still appear right and fair, depending on the target group being either within or without one's moral boundaries. Once without, inflicting harm may not yield negative self-concept and appear to be abiding by the in-group's moral norms (Lima-Nunes, Pereira, & Correia, 2013; Opotow, 1990).

Right wing authoritarianism (Altemeyer, 1981), and social dominance orientation (Pratto, Sidanius, Stallworth, & Malle, 1994), as two forms of the authoritarian personality independently predict negative intergroup outputs. This is due to strong adherence to the in-group's norms and values among highs in RWA and high tendency of maintaining intergroup inequality among highs in SDO (Duckitt, 2001).

Bearing this in mind, we assumed that relevant negative behavioral outcomes against an outgroup would be preceded by placing it outside one's moral regard. We investigated the mediating role of moral exclusion on the relationship between RWA as well as SDO, and the discriminatory intentions they yield, against the Roma and Jewish communities in Hungary.

4.2.1   The Dual Process Motivational Model of Ideology and Prejudice

Dual Process Model (Duckitt, 2001) posits that intergroup prejudice emerges from two distinct motivationally based social attitude clusters or ideological dimensions; social dominance orientation (SDO) and right-wing authoritarianism (RWA). Those high in SDO tend to view the world as highly competitive, which in turn disposes them to endorse motivational goals of perpetuating dominance, social hierarchy, and intergroup inequality (Pratto et al., 1994). High RWAs, on the other hand, perceive the world as a dangerous and threatening place, motivated to maintain in-group norms and values as well as social cohesion, order, and structure (Altemeyer, 1981).

Motivated cognition approach suggests that the way people see and comprehend the world is a determinant factor on holding relevant values and attitudes (Duckitt, 2001; Jost, Glaser, Kruglanski, & Sulloway, 2003). For instance, experimental studies show that individuals high in SDO are more responsive to the signals of competition and intergroup inequality (e.g., De Oliveira, Guimond, & Dambrun, 2012), and highs in RWA to those of threat to the social norms and values (e.g., Asbrock & Fritsche, 2013).

Furthermore, the two variables distinctly relate to various forms of detrimental treatment of the target outgroups (e.g., Poteat, Horn, & Armstrong, 2017). In line with our study, research also shows that negative behavioral intentions towards the Roma and Jewish communities in Hungary are associated with both SDO and RWA (e.g., Murányi & Sipos, 2012).

4.2.2   Moral Exclusion

The scope of justice is a boundary beyond which our moral values and rules of fairness and justice may not apply. The "other" be it an individual or a social group is morally excluded as it is considered outside one's scope of justice (Opotow, 1990). Hence, inflicting harm on the morally excluded individual or group does not render moral obligations and results in lack of interest in their well-being (e.g., Opotow, 1990).

To probe more deeply into the meaning of moral exclusion, Olson et al. (2011) illustrate that the construct should be elucidated in two distinct ways. First, the target is placed outside the scope of justice because it is viewed as irrelevant to one's moral concerns. In other words, moral exclusion occurs as the target is not entitled to the application of basic moral principles. The second meaning indicates that harsh treatment of the target does not appear to be morally tainted since it is balanced out by the application of a counter moral

principle implying deservingness. Hence, whatever the moral code may be, one would be able to justify the very mistreatment, as the target is perceived to be deserving it.

The concept of moral exclusion, therefore, would enable us to cut deeper into the negative side of intergroup relations. Ranging from the most extreme (e.g., genocide, religious inquisition) to more moderate examples (e.g., discrimination against an outgroup) moral exclusion can underlie various forms of hostile intergroup behaviors (e.g., Opotow, 1990). Moreover, empirical research shows that moral exclusion relates to negative behavioral intentions against a variety of outgroups (e.g., Hadarics & Kende, 2018; Opotow, 1993).

Further, previous research has reported positive associations between SDO as well as RWA and moral exclusion (e.g., Passini & Morselli, 2016). We expected that intergroup mistreatments independently predicted by RWA and SDO would be preceded by placing the target group beyond one's scope of justice. We assume that the underlying motivations driven from RWA and SDO have the potential to regulate the boundaries of group-based moral regard, and this regulatory mechanism mediates the relationship between RWA and SDO on the one hand, and group-based discriminatory intentions on the other.

4.3 The Study

Our objective was to examine whether both RWA and SDO—as motivated cognition—will predict negative intergroup behavioral intentions, and more importantly, whether this nexus will be explained by the individuals' constricted scope of justice. Therefore, we hypothesized that moral exclusion will mediate the idiosyncratic effects of RWA and SDO on discriminatory intentions towards Hungarian Roma and Jewish minorities.

Throughout variety of historical epochs and social settings, Europe has witnessed various forms of prejudice and hostility against both Roma and Jewish communities (e.g., Feischmidt, Szombati, & Szuhay, 2013; for Jewish see e.g., Lindemann, 1991). Moreover, moral exclusion has been argued to have applied to both groups (e.g., Lang, 2010; for moral exclusion and Roma predicament in Europe see Tileagă, 2007).

The Roma in Hungary make up about 5-8% of the entire population (Pásztor & Pénzes, 2013) experiencing explicit prejudice and discrimination in many social contexts (e.g., Kertesi & Kézdi, 2011). In contrast, anti-Semitism is less normative and less widespread, and serves symbolic and ideological functions, as it is less manifested in direct discrimination either on a personal or on an institutional level (Kovács, 2010).

4.3.1 Participants

We used a nationally representative sample of 1015 participants (M_{age} = 43.92, SD = 14.18; 51.5 % female, 48.5 % male). The recruitment was carried out by a professional public opinion company employing an online participant pool using a multiple-step, proportionally stratified, probabilistic sampling method resulting in a sample demographically similar to the Hungarian population in terms of age, gender, level of education, and type of settlement.

4.3.2 Measures

Right-wing authoritarianism was measured by a shortened Hungarian version of the RWA scale (Altemeyer, 1981) translated and adapted by Enyedi (1996). For assessing social dominance orientation, we applied a shortened 8-item Hungarian version of the SDO7 scale (Ho et al., 2015; adapted by Faragó & Kende, 2017). For measuring moral exclusion of the two target groups we adopted Opotow's (1993) Scope of Justice/Moral Exclusion Scale for each target group (example item: *"I believe that considerations of fairness apply to Jews/Roma people too."*). Discriminatory intentions were assessed by three items in the case of both outgroups (Roma people: *"I would support my supervisor at work in not hiring a Roma person."; "I would support a campaign aiming to stand up for the protection of the Hungarian majority against the harmful behavior of the Roma."; "I wouldn't do anything for changing the situation of the Roma in Hungary.";* Jews: *"In certain areas of employment, the number of Jews should be limited."; "I wouldn't do anything for remitting antisemitism in Hungary."; "It would be best if Jews left the country."*). Responses were measured on a 7-point scale *(1 = strongly disagree; 7 = strongly agree)* in each case.

4.4 Results

Using Mplus version 8 (Muthén & Muthén, 2017), two path analysis models with observed variables were built in order to examine the effects of SDO and RWA on discriminatory intentions directly, as well as with the mediational effect of moral exclusion added to the model. Correlations between the constructs and descriptive statistics are reported in Table 1. First, we tested the direct effects of SDO and RWA on discriminatory intentions without moral exclusion as the mediator (Figure 1). It was found that both SDO (Roma: β = 0.18; B = 0.24; SE = 0.043; p < .001; Jewish: β = 0.18; B = 0.23; SE = 0.043; p < .001) and RWA (Roma: β = 0.34; B = 0.44; SE = 0.043; p < .001; Jewish: β = 0.38; B = 0.44; SE = 0.038; p < .001) significantly predicted discriminatory intentions.

Table 1

Descriptive Statistics and Correlation Between Variables

Variables	*M*	*SD*	α	1.	2.	3.	4.	5.
1. Discrimination – Roma	3.09	1.68	.73					
2. Discrimination – Jewish	2.34	1.51	.79	.55***				
3. SDO	2.82	1.23	.78	.31***	.33***			
4. RWA	2.80	1.32	.76	.41***	.45***	.37***		
5. Moral exclusion - Roma	4.07	1.56	.78	.68***	.42***	.37***	.35***	
6. Moral exclusion - Jewish	3.86	1.44	.73	.48***	.52***	.27***	.35***	.63***

Note. ***= *p* < .001; ** = *p* < .01; * = *p* < .05.

Figure 1 (Roma). Path model testing the mediating effect of moral exclusion on the relationship between SDO as well as RWA, and discriminatory intentions. Path coefficients are standardized regression coefficients (***= *p* < .001). Path coefficients of the direct model are in brackets.

In our second model (Figure 2), moral exclusion was added to the relationship between the two forms of authoritarian trait and discriminatory intentions. We found that moral exclusion predicted discriminatory intentions (Roma: β = 0.60; *B* = 0.647; SE = 0.026; *p* < .001; Jewish: β = 0.40; *B* = 0.42; SE = 0.035; *p* < .001), and was predicted by both SDO (Roma: β = 0.16; *B* = 0.21; SE = 0.039; *p* < .001; Jewish: β = 0.16; *B* = 0.19; SE = 0.037; p < .001) and RWA (Roma: β = 0.29; *B* = 0.35; SE = 0.037; *p* < .001; Jewish: β = 0.27; *B* = 0.30; SE = 0.037; *p* < .001).
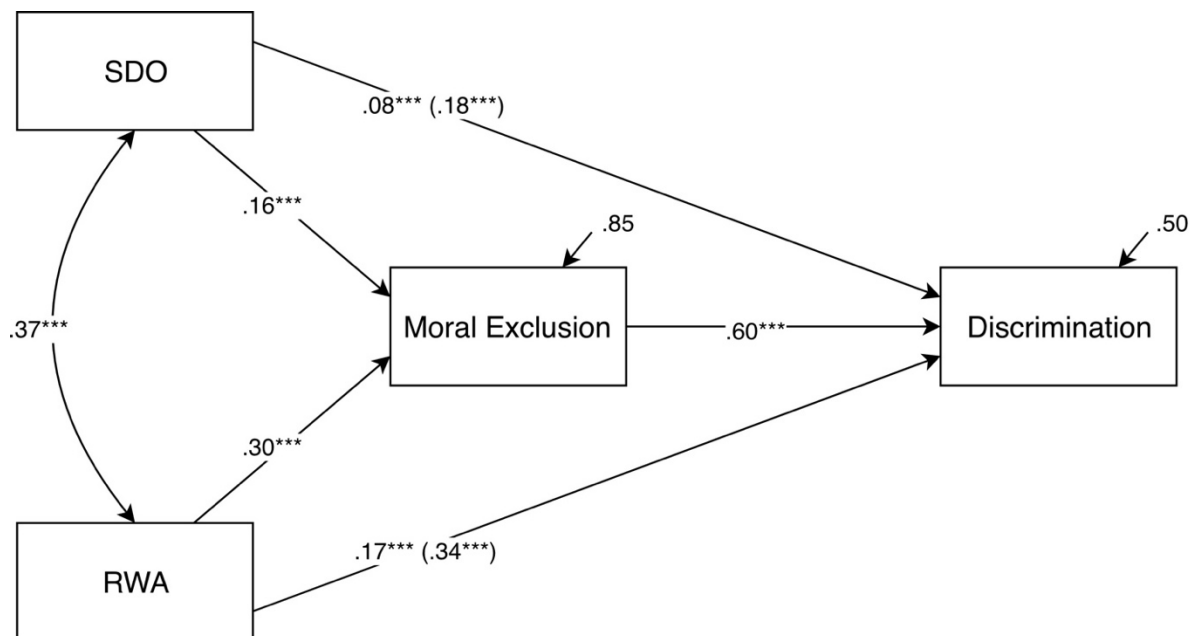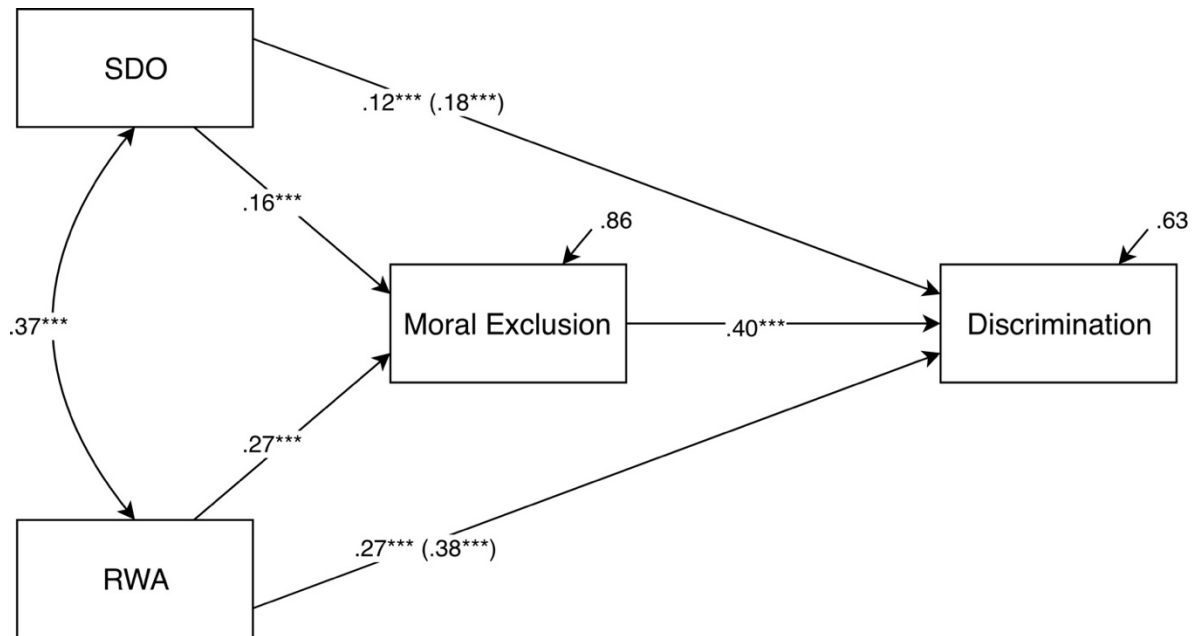
Figure 2 (Jewish). Path model testing the mediating effect of moral exclusion on the relationship between SDO as well as RWA, and discriminatory intentions. Path coefficients are standardized regression coefficients (***= $p < .001$). Path coefficients of the direct model are in brackets.

Further, we examined whether or not SDO and RWA would significantly yield weaker effects on discriminatory intentions as moral exclusion being added to the model. To this end, we conducted two separate model comparisons fixing the direct effects of SDO and RWA on discriminatory intentions in the mediational model, to those of found in our first model without moral exclusion as mediator. Then, the fits of the constrained models were compared to unconstrained ones. We found that the model fits significantly dropped concerning the effect of both SDO (Roma: $\Delta\beta = 0.10$; $\Delta B = 0.13$; $\Delta X^2 = 16.46$; $\Delta df = 1$; p <.001; Jewish: $\Delta\beta$ =0.06; ΔB= 0.08; $\Delta X^2 = 5.83$; $\Delta df = 1$; p =.01), and RWA (Roma: $\Delta\beta = 0.17$; ΔB =0.22; $\Delta X^2 = 49.24$; $\Delta df = 1$; $p < .001$; Jewish: $\Delta\beta$ =0.10; $\Delta B = 0.13$; $\Delta X^2 = 15.25$; Δdf =1; $p < .001$) on discriminatory intentions. Therefore, we can be confident that the two regression coefficients in our first model are different from those found in the mediational model.

Moreover, the mediation analyses showed that moral exclusion yielded significant indirect effects on the paths from SDO (Roma: $\beta = 0.10$; $B = 0.13$; $p < .001$; 95% CI [0.08, 0.18]; Jewish: $\beta = 0.07$; $B = 0.08$; $p < .001$; 95% CI [0.05, 0.12]) as well as RWA (Roma: $\beta = 0.18$; $B = 0.22$; $p < .001$; 95% CI [0.18, 0.28]; Jewish: $\beta = 0.11$; $B = 0.12$; $p < .001$; 95% CI [0.09, 0.16]) to discriminatory intentions.

In addition, with regards to the Roma model, the indirect effect of moral exclusion explained 54.91% of the total effect in the case of SDO ($\beta = 0.18$; $B = 0.24$; $p < 0.001$) and

51.36% in the case of RWA (β = 0.34; B = 0.44; p < .001). While in the Jewish model the indirect effect of moral exclusion explained 35.52% of the total effect in the case of SDO (β = 0.18; B = 0.23; p < 0.001) and 28.60% in the case of RWA (β = 0.38; B = 0.44; p < .001).

4.5   Discussion

Employing structural equation modeling, we investigated the mediating role of moral exclusion on the association between the two facets of authoritarian personality and behavioral intentions towards Roma as well as Jewish minorities in Hungary. Previous research has mainly highlighted the positive link between moral exclusion and the authoritarian personality (e.g., Passini, 2008), or their negative prejudicial and/or attitudinal intergroup outcomes (e.g., Passini & Morselli, 2016). In this study, we went beyond by focusing on the two constructs as related to the behavioral intentions they render. Supporting our assumption, the findings showed that moral exclusion mediated the effects of RWA and SDO on discriminatory intentions.

In alignment with previous research (e.g., Murányi & Sipos, 2012), we found that discriminatory intentions against both Roma and Jewish minorities in Hungary appeared to be explained by both SDO and RWA. Generally speaking, SDO results in prejudice against the "derogated" social groups perceived as being low in status and/or power. RWA predicts prejudice against the groups regarded as "dangerous" who disrupt social structure and cohesion. While it might also be the case that prejudice against certain social groups to be predicted by both constructs concurrently. Such "dissident" groups, hence, become a target of prejudice by RWA and SDO simultaneously, since they are deemed and justified as both threatening and underprivileged (Duckitt & Sibley, 2007).

Roma and Jewish minority groups in Hungary are targets of distinct negative attitudes emanating from their different socioeconomic situation, level of cultural assimilation, cultural identity, and etc. (see e.g., Kovács, 2002). Gypsies "othered" both socially and morally (Tileagă, 2007) are subjects of stereotypes such as laziness, criminality, and receiving undeserved benefits from the state (Kende, Hadarics, & Lášticová, 2017). Anti-Semitism on the other hand appears in heterogeneous forms including anti-religious attitudes, and cognitive-affective stereotypes, like beliefs in Jewish conspiracy or inducing guilt in the majority society (Kovács, 2010). Despite their differences, however, both social groups seem to be differentially targeted by authoritarian ideologies and beliefs and similar in being perceived as dangerous to the society's wellbeing and "deservingly" second-rank citizens.

Perceived threat and psychological distance have been proposed as two major precursors of moral exclusion. Regarding the former, an outgroup becomes excluded from the scope of justice when it is conceived as a threat to the in-group's interests. The latter, on the other hand, is concerned with the lack of identification with the target group (Lima-Nunes, Pereira, & Correia, 2013; Olson, Cheung, Conway, Hutchison, & Hafer, 2011; Opotow, 1990, 1993). Hence, culturally different minority groups should be more optimally vulnerable to moral exclusion as their cultural dissimilarity may entail both psychological distance and perception of intergroup threat. So, those who score high on RWA can be more susceptible to the cues of perceived threat inflicted from the minority group upon the in-group's norms and values. On the other hand, since SDO involves endorsement of ideologies that leads to rationalization of intergroup inequality, it would be credible to assume that the outgroup would be also perceived as psychologically distant. Relatedly, research shows that SDO is negatively associated with personality traits which feature psychological proximity with the outgroups (Sibley & Duckitt, 2008). Therefore, high SDOs may exclude the same outgroup from the scope of justice due to the perceived psychological distance. Indeed, moral exclusion can be construed not as a generalized propensity, but as a result of independent psychological mechanisms sprung from differential social cognitive motivations (cf. Passini & Morselli, 2016). A future research theme, thus, should include direct investigation of the two processes preceding moral exclusion.

Since we used a correlational model, drawing conclusion on the causal relationship between the variables would be impossible. One might argue, for instance, before a mistreatment having meted out, individuals might first morally exclude the target from their scope of justice. That is to say, regulatory mechanism of moral exclusion maintains the positive picture of the self, which in turn opens the moral gate for and abets actual enactment of the harsh treatment. On the other hand, we by no means refute the possibility that after occurrence of a maltreatment; moral exclusion as an a *posteriori* psychological mechanism would take place in order to preserve one's positive self-concept. Thus, examining the causal direction from moral exclusion to behavior and/or vice versa, can be another agenda for future research.

4.6   Conclusion

In closing, enacted by an actor or a bystander, be it an active engagement in or a passive condoning of a mistreatment against an outgroup, widening the majority—advantaged—group members' scope of justice can be a means to intervene intergroup prejudice. To this

end, we assume that it is important to understanding social cognitive motivations entwined with the inclination of excluding the "other" from one's moral consideration.

4.7   References

Altemeyer, B. (1981). *Right-wing authoritarianism*. Winnipeg: University of Manitoba Press.

Asbrock, F., & Fritsche, I. (2013). Authoritarian reactions to terrorist threat: Who is being threatened, the Me or the We. *International Journal of Psychology*, *48*, 35-49.

De Oliveira, P., Guimond, S., & Dambrun, M. (2012). Power and legitimizing ideologies in hierarchy-enhancing vs. hierarchy-attenuating environments. *Political Psychology*, *33*, 867–885.

Duckitt, J. (2001). A cognitive-motivational theory of ideology and prejudice. *Advances in Experimental Social Psychology*, *33*, 41-113.

Duckitt, J., & Sibley, C. G. (2007). Right wing authoritarianism, social dominance orientation and the dimensions of generalized prejudice. *European Journal of Personality*, *21*, 113-130.

Enyedi, Z. (1996). Tekintélyelvűség és politikai-ideológiai tagolódás. *Századvég*, *2*, 135-155.

Faragó, L., & Kende, A. (2017). *Az elnyomás támogatása vagy az egyenlőség ellenzése? Az új szociális dominancia orientáció skála (SDO7) vizsgálata.* Alkalmazott Pszichológia, 17, 115-135.

Feischmidt, M., Szombati, K., & Szuhay, P. (2013). Collective criminalization of Roma in Central and Eastern Europe: social causes, circumstances and consequences. In S. Body-Gendrot, M. Hough, K. Kerezsi, R. Levy, & S. Snacken (Eds.), *The routledge handbook of european criminology* (pp. 168–187). London UK: Routledge.

Hadarics, M., & Kende, A. (2018). Moral foundations of positive and negative intergroup behavior: Moral exclusion fills the gap. *International Journal of Intercultural Relations*, *64*, 67-76.

Ho, A. K., Sidanius, J., Kteily, N., Sheehy-Skeffington, J., Pratto, F., Henkel, K. E., ... Stewart, A. L. (2015). The nature of social dominance orientation: theorizing and measuring preferences for inequality using the new SDO7 scale. *Journal of Personality and Social Psychology*, *109*, 1003-1028.

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, 129, 339–375.

Kende, A., Hadarics, M., & Lášticová, B. (2017). Anti-Roma attitudes as expressions of dominant social norms in Eastern Europe. *International Journal of Intercultural Relations, 60,* 12-27.

Kertesi, G., & Kézdi, G. (2011). Roma employment in Hungary after the post-communist transition1. *Economics of Transition, 19*, 563–610. http://dx.doi.org/10.1111/j.1468-0351.2011.00410.x.

Kovács, A. (2002). *Zsidók és zsidóság a mai Magyarországon. Egy szociológiai kutatás eredményei*. Budapest: Szombat.

Kovács, A. (2010). *The stranger at hand: Antisemitic prejudices in post-communist Hungary.* Leiden, The Netherlands: Brill.

Lang, J. (2010). Questioning dehumanization: Intersubjective dimensions of violence in the Nazi concentration and death camps. *Holocaust and Genocide Studies*, *24*(2), 225-246.

Lima-Nunes, A., Pereira, C. R., Correia, I. (2013). Justice seems not to be for all: exploring the scope of justice. *In-Mind Magazine*, *17*, 1-6.

Lindemann, Albert S. (1991). The Jew Accused: *Three Anti-Semitic Affairs (Dreyfus, Beilis, Frank) 1894–1915*. Cambridge: Cambridge University Press.

Murányi, I., & Sipos, F. (2012). Nemzeti radikálisok tekintélyelvűsége: Szociális dominancia orientáció és ellentörténelem [Radical nationalist's authoritarianism: Social dominance orientation and counterhistory]. *Metszetek*, *1*, 32-56.

Muthén, L., & Muthén, B. (2017). Mplus (Version 8)[computer software].(1998–2017). *Los Angeles, CA: Muthén & Muthén*.

Olson, J. M., Cheung, I., Conway, P., Hutchison, J., & Hafer, C. L. (2011). Distinguishing two meanings of moral exclusion: Exclusion from moral principles or principled harm-doing? *Social Justice Research*, *24*, 365–390.

Opotow, S. (1990). Moral exclusion and injustice: an introduction. *Journal of Social Issues*, *46*, 1-20.

Opotow, S. (1993). Animals and the scope of justice. *Journal of Social Issues*, *49*, 71-85.

Passini, S. (2008). Exploring the multidimensional facets of authoritarianism: authoritarian aggression and social dominance orientation. *Swiss Journal of Psychology*, *67*, 51-60.

Passini, S., & Morselli, D. (2016). Construction and validation of the moral inclusion/exclusion of other groups (mieg) scale. *Social Indicators Research*. doi:10.1007/s11205-016-1458-3

Pásztor, I. Z., & Pénzes, J. (2013). Studying the number of the Gypsy minority in Hungary and their involvement in local governments on the example of the North Great Plain region. *Revista Română De Geografie Politică*, *15*, 17-26.

Poteat, V. P., Horn, S. S., & Armstrong, P. I. (2017). Condoning discrimination: The effects of dominance and authoritarianism are moderated by different ways of reasoning about antigay discriminatory acts. *Group Processes & Intergroup Relations*, *20*, 831-849.

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: a personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*, 741-763.

Sibley, C. G., & Duckitt, J. (2008). Personality and prejudice: A meta-analysis and theoretical review. *Personality and Social Psychology Review*, *12*, 248-279.

Tileagă, C. (2007). Ideologies of moral exclusion: A critical discursive reframing of depersonalization, delegitimization and dehumanization. *British Journal of Social Psychology*, 46, 717-737.

5 Study II: Anti-Roma bias (stereotypes, prejudice, behavioral tendencies): A network approach toward attitude strength

5.1 Abstract

The Roma have been and still are a target of prejudice, marginalization, and social exclusion across Europe, especially in East-Central European countries. This paper focuses on a set of stereotypical, emotional, and behavioral evaluative responses towards Roma people selected as representing the underlying components of anti-Roma bias. Employing network analysis, we investigated, if attitude strength is associated with stronger connectivity in the networks of its constituent elements. The findings from representative surveys carried out in Hungary, Romania, Slovakia, France, and Ireland supported our assumption, as high attitude strength towards the Roma resulted in stronger connectivity in all pairs of high versus low attitude strength networks. Our finding yields a solid theoretical framework for targeting the central variables—those with the strongest associations with other variables—as a potentially effective attitude change intervention strategy. Moreover, perceived threat to national identity, sympathy, and empathy were found to be the most central variables in the networks.

Keywords: Anti-Roma bias, attitude strength, network connectivity, network analysis, intervention

5.2 Introduction

The Roma are among the most disenfranchised, socially unaccepted, and morally vilified ethnic minority groups in Europe and especially in East-Central European countries (Fraser, 1995; Ladányi, 2001; Pogány, 2006; Tileaga, 2006). As a culturally and linguistically diverse group, Roma people are portrayed as beggars, criminals, profiteers, and lazy, being a target of marginalization and social exclusion, as well as perpetual discriminatory and violent practices on interpersonal, institutional, and national level (Feischmidt et al., 2013; Van Baar, 2011). School segregation of Roma students in Hungary, the Czech Republic, and Slovakia (Messing, 2017), violent vigilante activities in Hungary and Romania, and forced eviction of the Roma in Romania, France, Italy, and Slovakia are all strikingly telling cases in point (see e.g., Amnesty International Report, 2013).

Empirical research shows that anti-Roma stereotypes revolve around criminality, laziness, and receiving undeserved benefit from the state (e.g., Enyedi et al., 2004; Kende et al., 2020; Kende et al., 2017; Villano et al., 2017). Moreover, drawing on the stereotype content model (SCM, Fiske et al., 2002), the Roma are perceived to be low in both warmth and competence (e.g., Grigoryev et al., 2019; Stanciu et al., 2017). Further, research shows that the Roma are perceived as both dangerous and derogated (e.g., Bilewicz et al., 2017; Hadarics & Kende, 2019; Imhoff & Bruder, 2014), which also indirectly implies they are rejected both from the perspective of threatening conventional norms and looked down upon as low status- group—being low in both dimensions of the thereof model.

Needless to say, intervention efforts are needed to combat anti-Roma bias. However, one practical challenge is to identify the most effective attitude change interventions considering that anti-Roma stereotypes are historically rooted and strong in most societies. Previous intervention efforts, in general, have not been successful in dampening intergroup bias (Paluck & Green, 2009). Mainstream intergroup bias research is often engaged with parsimonious models investigating relationships between a limited number of variables, which does not ensure identifying the most influential stereotypical and prejudicial evaluations. In the current study, we attempt to fill this gap by employing a network approach in the anti-Roma stereotype context. Our main objective is to examine whether the network approach would be a theoretically justified method to be employed for intervention purposes in an anti-Roma bias context in future research. Drawing on the literature on attitude strength and network analysis, we test the *connectivity hypothesis* proposed by Dalege et al. (2018) in the networks of stereotypical, emotional, and behavioral evaluations towards the Roma estimated from representative samples collected in Hungary, Romania, Slovakia, France, and Ireland.

The five countries included are three Eastern European countries with the largest indigenous Roma minority (with 8% of the Romanian population, 7% in Hungary and 9% in Slovakia) and two Western European countries (Ireland and France) where Roma have immigrated in the last twenty years, and who also have their own indigenous Roma population groups (i.e., Irish Travellers in Ireland, and Sinti in France). While their visible economic disadvantages may be the strongest in Eastern Europe where they form a large (often the largest) ethnic minority group, their treatment in Western Europe is often inhumane and goes against EU norms and regulations (European Commission, 2015; Gould, 2015; Mahoney, 2011).

5.2.1 Network analysis

Network analysis is a relatively novel approach to modeling individual differences in psychological constructs by representing the direct interactions between their underlying components. Representing stereotype structures through network models has also recently received attention from researchers in the field (e.g., Sayans-Jiménez et al., 2018; Grigoryev et al., 2019). Modeling the direct and unique interrelations between a relatively higher number of variables as a network can be an advantageous method to render possible picturing of a more comprehensive representation of stereotype dynamics. Having a variety of stereotypes and negative attitudes estimated as a network can help us finding variables with

the highest degree of interrelations with other variables whom can be the most favorable candidates to be wagered on for intervention purposes. With a latent approach, for instance, this cannot be possible since all the items are treated as equivalent measures of the latent construct (Schmittmann et al., 2013).

Nodes and edges are the two most basic constituent elements of a network; nodes are a number of entities and edges the direct interrelationships between every possible pair of nodes. In psychological networks, nodes are a set of observed variables and edges the statistical associations between them (Epskamp et al., 2018). Connectivity is another basic property of a network that refers to the overall level of interrelations among all the nodes and the degree of causal interdependencies between them. The higher the connectivity between nodes within a given network, the more likely it is that changes to one node will also be mirrored by changes in other nodes within that network (Scheffer et al., 2012). Moreover, global connectivity, as a measurement of network connectivity, is the sum of all absolute values that every edge in the network possesses. Hence, the number of connections and the magnitude of the edge weights determine the connectivity of a network.

5.2.2 Network connectivity as related to attitude strength

Proposing the Causal Attitude Network (CAN) model, Dalege et al. (2018) integrated the general notion of network connectivity with attitude networks and proposed the *connectivity hypothesis*, which refers to the higher connectivity, between the evaluations on different aspects of an attitude object for those who hold stronger attitude towards that attitude object.

As mentioned above identifying the nodes with the highest degree of direct interactions with the other nodes in a network of stereotypical evaluations would be a highly beneficial means for intervention purposes. To consolidate this approach, in the current study, we employ the connectivity hypothesis. We argue that the connectivity between different stereotypical, emotional, and behavioral evaluations towards the Roma estimated as a network, to be found also as a measurement of attitude strength would yield a firm theoretical linchpin for intervention aims. For if nodes with highest interrelations with the others rendered at odds with the other nodes, the need for cognitive consistency as a factor indispensable to attitude strength (e.g., Monroe & Read, 2008; Simon, Snow, & Read, 2004) would lead the system to regain the compatibility between all its components.

By definition, attitude strength is "the extent to which attitudes manifest the qualities of durability and impactfulness" (Krosnick & Petty, 1995, p. 3). Durability refers to attitude stability over time and resistance to change, and impactfulness to its influence on information

processing and behavior. Strong attitudes, therefore, acquire these attributes to a greater extent in comparison with weak attitudes. Krosnick and Petty (1995) propose several features of attitude strength such as extremity, importance, and accessibility inter alia. Dalege et al. (2018) found that in a network of a number of evaluations on the presidential candidates, the network connectivity is higher for those who hold a stronger attitude concerning political campaigns. Moreover, they showed that network connectivity is also an expression of other basic properties of attitude strength. They estimated correlation coefficients between feeling thermometer items towards the presidential candidate measured before and after the election (as a measure of attitude stability) and found that network connectivity significantly associates with attitude's stability over time. Moreover, they also showed that network connectivity predicts the biserial correlations between the feeling thermometer item towards the presidential candidate before the election and the respondents' actual voting decision (see Dalege et al., 2018).

In the current research, we test the connectivity hypothesis in the context of anti-Roma bias. In line with previous findings of the CAN model, we assume that high attitude strength networks of a number of stereotypical, emotional, and behavioral evaluations towards the Roma will possess a significantly stronger degree of global connectivity compared to those of low attitude strength networks.

5.3   Method

Twenty-seven stereotypical, emotional, and behavioral evaluative responses towards the Roma (for an overview of the underlying components of an attitude, see McGuire, 1990) were used to examine their connectivity in the networks of high versus low attitude strength for each country. Four steps of network data analysis were performed: network estimation, network comparison, network inference, and network stability recommended by Fried et al. (2018). Moreover, an additional check section was added to report the results of pathway analyses.

5.3.1   Participants

Nationally representative survey data were collected through online participant pools across five countries; Hungary (N = 1039, M_{age} = 47.99, SD_{age} = 14.84, 52.7% women), Romania (N = 1044, M_{age} = 42.11, SD_{age} = 15.80, 48.2% women), Slovakia (N = 1033, M_{age} = 44.06, SD_{age} = 16.10, 52.7% women), France (N = 975, M_{age} = 42.10, SD_{age} = 13.30, 54 % women), and Ireland (N = 1000, M_{age} = 44.91, SD_{age} = 15.72, 51.5 % women).

Based on simulation studies (Epskamp, 2016), a moderate size network with 24 nodes for continuous data is recommended to be estimated from at least 250 respondents approximately. Number of participants for all networks were sufficient (Hungary: N_{high} = 511, N_{low} =512; Romania: N_{high} = 467, N_{low} = 463; Slovakia: N_{high} = 516, N_{low} = 517; France: N_{high} = 472, N_{low} = 498; Ireland: N_{high} = 476, N_{low} = 469). Moreover, 16 respondents from the Hungarian sample, 114 respondents from the Romanian sample, 5 respondents from the French sample, and 55 respondents from the Irish sample did not respond on the feeling thermometer scale and were removed from the analysis.

Data was collected by professional opinion poll companies in each country, working with the IRB approval of Eötvös Loránd University. The surveying companies used a multiple-step, proportionally stratified, probabilistic sampling method of an online participant pool resulting in a sample demographically similar to the respective population in terms of age, gender, and type of settlement. Note that the French sample was representative only regarding age and gender. (see supplementary materials for the demographic similarities between each sample and the corresponding population).

5.3.2 Measures

Twenty-seven items of stereotypes, emotions, and collective action tendencies towards the Roma were selected for the network estimations from the omnibus surveys. A fourteen-item revised Attitudes Toward Roma Scale1 (original ATRS; Kende et al., 2017), with three subscales, was used. Six items of ATRS measured Blatant Stereotyping (e.g., "There are very little proper or reasonable Roma people."), five items measured Undeserved Benefits (e.g., "The real damage is caused by organizations which offer an undeserved advantage to Roma people."), and three items measured Cultural Difference ("The Roma can be proud of their cultural heritage."). Four discrete intergroup emotions were measured, each with a single item; empathy ("I feel empathy with Roma people"), sympathy ("I feel sympathy with Roma people."), anger ("I feel anger about the treatment of Roma people."), and hope ("I feel hopeful about the future of Roma people."). Collective action intentions with a pro-Roma orientation were measured by six items, including items on engagement in traditional forms of collective action, such as signing petitions (e.g., "I would participate in some form of

---

[1] We improved the original 16-item scale by including reversed items and made the cultural subscale unambiguously about cultural recognition. These revisions were made as part of project PolRom (www.polrom.eu). This paper is the first publication of the new scale.

action (e.g. signing a petition) defending the rights of the Roma.") as well as items about offering donations and volunteerism (e.g., "I would donate clothing, school supplies or toys for Roma families."). Lastly, three items measured perceived threat to national identity (e.g., "Roma people are a threat to [country] culture."). All the items were measured on a 7-point scale (1 = strongly disagree; 7 = strongly agree).

As a general measure of attitude, we used a single-item feeling thermometer scale measuring participants' attitudes towards the Roma from 0 (very unlikeable) to 100 (very likeable). Attitude extremity as one feature of attitude strength (see Krosnick & Petty, 1995), was calculated by computing the deviation of the participants' responses from neutrality on the feeling thermometer scale (for operationalizing attitude extremity see Krosnick & Smith, 1994). First, the absolute difference between each participant's score and the scale mean was calculated. Next, on the new computed item, participants with values from the lowest through the median were selected as low attitude strength groups and the rest as the high attitude strength group (Hungary$_{median}$: 20.35; Romania$_{median}$: 25.52; Slovakia$_{median}$: 20.42; France$_{median}$: 20.14; Ireland$_{median}$: 22.74). Correlations between the variables, descriptive statistics of all the items, and the items themselves can be found in the supplementals.

5.3.3 Network estimation

For each country, a pair of high versus low attitude strength networks were estimated. Using the Extended Bayesian Information Criterion function EBICglasso from the R package qgraph (Epskamp et al., 2012), correlation matrices were inverted into partial correlation matrices to obtain unique statistical associations between all possible pairs of nodes. The correlation matrices were computed through pairwise complete observations to keep all the participants with missing values in the analyses. Also, a regularization technique LASSO (least absolute shrinkage and selection operator) was employed to controlling the effects of redundant correlations by setting small coefficients to zero (Friedman et al., 2008).

5.3.4 Network comparison

As the main analysis of this study, we compared global connectivity of all high versus low attitude strength networks for each country using the R package "NetworkComparisonTest" (NCT; van Borkulo et al., 2017). We applied a permutation method with 1000 iterations to examine if high attitude strength networks in each country are significantly more connected in comparison with low attitude networks. In addition, networks were examined whether they are structurally different; meaning for any pair of networks if there is any edge weight that is significantly different.

5.3.5 Network inference

To identify the most influential nodes, we computed centrality metrics. Centrality refers to the extent a node is influential in its interactions with other nodes in a network. Among several centrality metrics, we chose *strength* and *node predictability*. Strength is the sum of all edge weights that a node acquires in relation to all other nodes (Barrat et al., 2004). Using the R package "mgm" (Haslbeck, 2015), we computed the node predictability of each item, which is the proportion of variance for each node explained by all other nodes on average.

5.3.6 Network stability

Employing R package bootnet (Epskamp et al., 2018), we computed centrality and edge weights accuracy of all networks. A network is considered as stable (i.e., the centrality indices are interpretable) if the order of a centrality index is identical after re-estimating the network with a smaller number of participants. That is if the correlation stability coefficient (CS-coefficient) is preferably higher than 0.5 and no smaller than 0.25. CS-coefficient is the quantification of the maximum proportion of cases dropped, with 95% probability, so that centrality metrics or edge weights of the remaining cases correlate with those of the original network higher than 0.7 (Epskamp et al., 2018). In addition, bootstrapping with 95 % confidence intervals around the edge weights was performed for all networks as an indicator of edge weights accuracy.

5.4 Results

5.4.1 Network estimation

Five pairs of high and low attitude strength networks for each sample are depicted in Figure 1. Out of 351possible edges, networks of high attitude strength were found to have a greater number of non-zero edges (Hungary: 166 vs. 160; Romania: 177 vs. 153; Slovakia: 173 vs. 147; France: 184 vs. 145; Ireland: 173 vs. 145).
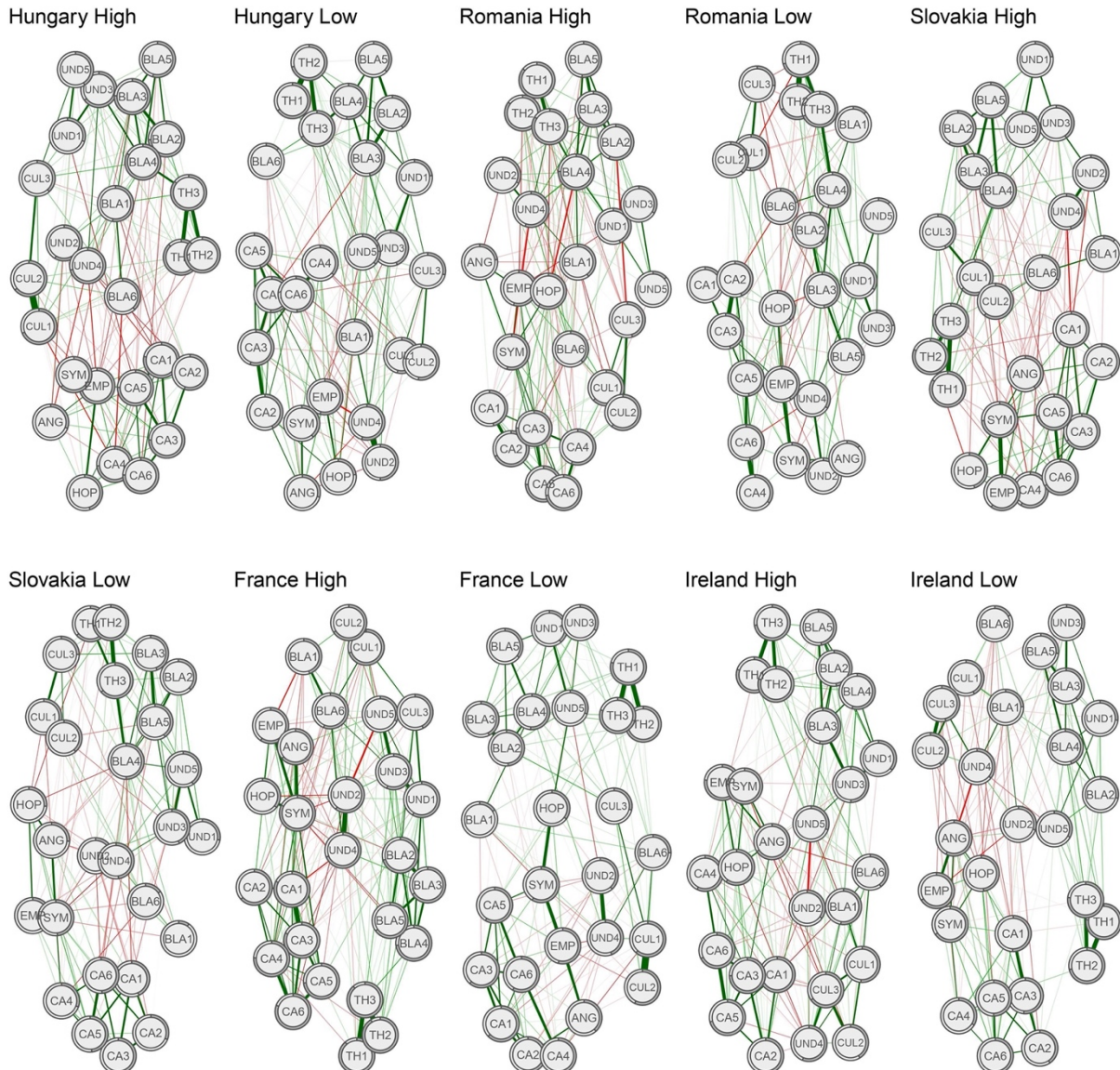
Figure 1. Regularized partial correlation networks of high versus low attitude strength. Node predictability is highlighted by the gray line around each node. Red lines depict negative correlation coefficients, and the thickness of the lines represents the magnitude of partial correlation coefficients. UND1, Undeserved Benefit_1; UND2, Undeserved Benefit_2; UND3, Undeserved Benefit_3; UND4, Undeserved Benefit_4; UND5, Undeserved Benefit_5; CUL1, Cultural Difference_1; CUL2, Cultural Difference_2; CUL3, Cultural Difference_3; BLA1, Blatant Stereotyping_1; BLA2, Blatant Stereotyping_2; BLA3, Blatant Stereotyping_3; BLA4, Blatant Stereotyping_4; BLA5, Blatant Stereotyping_5; BLA6, Blatant Stereotyping_6; CA1, Collective Action_1; CA2, Collective Action_2; CA3, Collective Action_3; CA4, Collective Action_4; CA5, Collective Action_5; CA6, Collective Action_6; EMP, Empathy; SYM, Sympathy; ANG, Anger; HOP, Hope; TH1, Perceived Threat_1; TH2, Perceived Threat_2; TH3, Perceived Threat_3. The green lines represent positive correlations.

5.4.2   Network comparison

Global connectivity of every network of high attitude strength was significantly higher compared to that of their corresponding low attitude strength network (Hungary: 12.38 vs. 11.64, $p = 0.03$; Romania: 11.85 vs. 10.46, $p < .001$; Slovakia: 12.17 vs. 10.97, $p = .005$;

France 13.48 vs. 11.77, *p* <.001; Ireland: 12.94 vs. 11.59, *p* <.001). In addition, for none of our network pairs, a significant difference between their edge weights was found. This implies that high networks did not structurally differ from their corresponding low networks and the only difference was on their global connectivity.

As mentioned above to measure attitude extremity, the absolute difference of each participant's response from the mean value was computed on a feeling thermometer scale. Next, two sub-samples of high and low attitude extremity were created for each country by splitting the datasets by the median of the computed item. As a sensitivity analysis, we split the datasets by 40th-60th as well as 60th-40th percentiles. We ran ten additional permutation tests. For eight out of ten of the comparisons, the effect was still significant. Only in the case of Hungary in the 40th-60th percentile split, we didn't find a significant difference and in the 60th-40th percentile split the difference was marginally significant (*p* = .053).

As another sensitivity analysis, we estimated the networks by a different technique. We binarized all the 27 nodes into zero (from 1 to 4 as not holding the belief) and one (5 to 7 as holding the belief) and re-estimated weighted networks with an eLasso technique using the R package IsingFit (van Borkulo & Epskamp, 2015). The eLasso technique regresses all the nodes on all other nodes and regularizes all the regressions controlling for the multicollinearity problem when many variables are regressed on each other (Friedman et al., 2008). Next, the best model fitting the extended Bayesian information criterion is selected (Foygel & Drton, 2010). We then compared all the corresponding high and low networks again by a permutation test with 1000 iterations. The results were similar to the main analyses, as all of the high attitude strength networks showed a significantly higher global connectivity compared to those of low attitude networks. Moreover, centrality values and network stabilities were also similar to the networks estimated by EBICglasso.

5.4.3 Network inference

Figure 2 shows the strength centrality values of all the items of the full-size networks (see the supplementary material for further details of the centrality values of all the full-size as well as high and low attitude networks). On average, most central values were found to be empathy in Hungary, perceived threat to national identity in Romania, and sympathy in Slovakia, France, and Ireland. Regarding node predictability, perceived threat to national identity was predicted by other variables to the highest extent in all the full-size networks networks. Moreover, the order of centrality values of the full-size networks was highly similar to that of the corresponding high and low attitude networks.
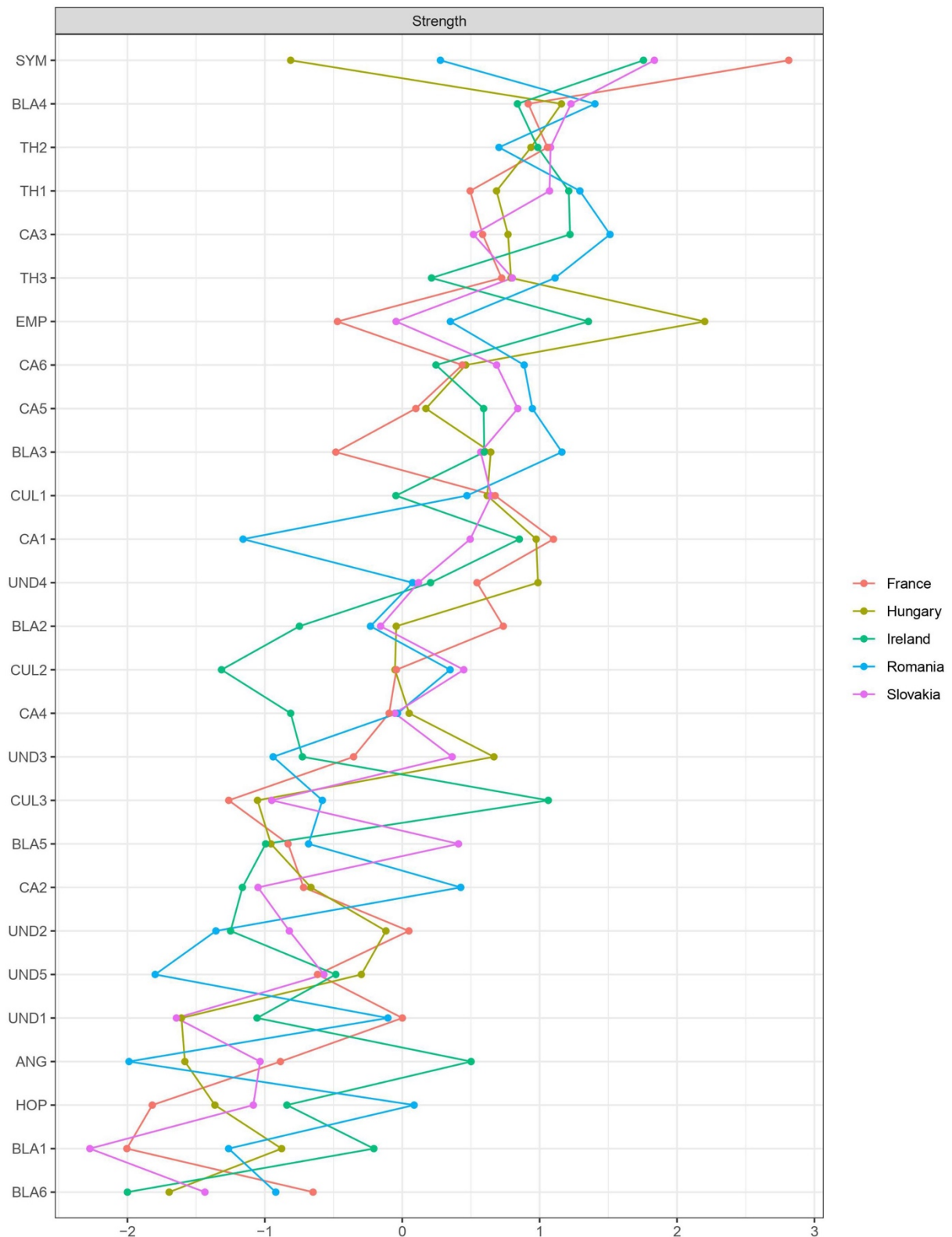
Figure 2. Strength centrality plot of the full-size regularized networks showing standardized z-score values of strength centrality. Strength measures the sum of all the regularized partial correlation coefficients for each node.

5.4.4   Network stability

Regarding strength centrality, all networks were found to be stable—CS-coefficients were higher than 0.5. Moreover, the edge weights were sufficiently accurate for all networks; the confidence intervals were small enough so that edge weights were interpretable (see supplementary material for more details).

5.4.5   Additional check

As an additional check, we also tested if the structure of anti-Roma bias fits with the intergroup bias structure proposed by Fiske (2015)—social structure predicting stereotypes, that predict emotional prejudice, which in turn predict behavioral tendencies. We examined the shortest paths from perceived threat to national identity nodes (considered as social structure) to collective action tendency nodes (considered as behavioral tendencies). In all the full-size networks, using the R package EGAnet (Golino & Christensen, 2019) we estimated the number of dimensions, and with the pathways function from the R package qgraph (Epskamp et al., 2012) we examined the shortest paths. Figure 3 shows that there are several shortest paths going from perceived threat to national identity nodes to collective action tendency nodes through the nodes on stereotypical evaluations, while there are also direct paths. However, we do not see the role of emotions in the pathways. The reason should be due to the nature of the intergroup emotions measured in this study which are prosocial emotions such as hope and empathy as opposed to prejudicial emotions such as contempt and disgust. Overall, the pathways seem to be more or less consistent with the theoretical framework suggested by Fiske (2015). Similar pathway analyses for both high and low attitude strength networks are visualized in the supplementals.
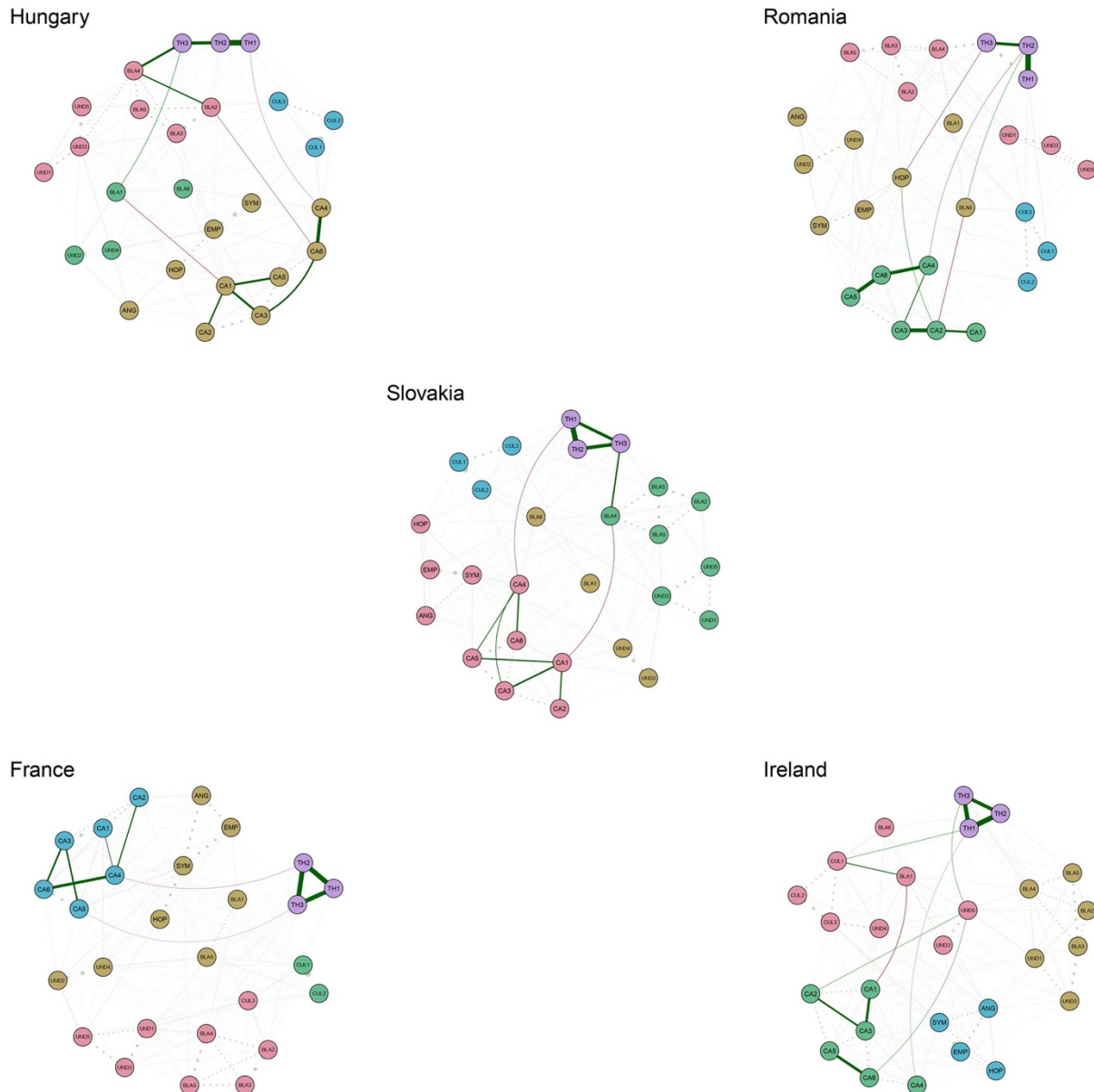
Figure 3. Shortest pathways from perceived threat to national identity to collective action tendencies in the full-size networks. The solid lines depict the edges that belong to the shortest paths.

5.5 Discussion

The CAN model (Dalege et al., 2018) was used to examine network connectivity in terms of the evaluative responses on the presidential candidates and found that network connectivity predicts the extent to which individuals are interested in political campaigns. In the current study, we supplemented the connectivity hypothesis by testing it in the context of anti-Roma bias. Using a network approach, we investigated if attitude strength would significantly associate with stronger connectivity in the networks of a set of stereotypical, emotional, and behavioral reactions towards the Roma people. The findings supported our assumption in all pairs of high versus low networks estimated from the nationally representative samples

collected in Hungary, Romania, Slovakia, France, and Ireland. That is, for those who hold a stronger attitude towards the Roma, relevant stereotypical, emotional, and behavioral evaluations are causally interrelated to a significantly higher extent. Moreover, we went beyond the previous research by framing network connectivity as a theoretical justification for future intervention-based research in the context of anti-Roma bias in particular and intergroup relations in a broader scope.

Due to the cross-sectional nature of the study, however, we did not examine the other two empirical findings of the CAN model; the relationship between network connectivity and stability of the attitude in time and its impact on actual behavior. Employing longitudinal designs, future research should consider if this would also be the case with regards to stereotypical evaluations. Moreover, we measured attitude strength qua computing the participants' deviations from neutrality on a feeling thermometer scale. However, extreme responses might not necessarily due to the strength of the attitude but the individuals' response styles. Future research should consider other features and/or measurements and operationalizations of attitude strength. Further, in the current research, we measured behavioral intentions through collective action tendencies, future research could include different measurements such as the preference for contact with Roma people.

Previous research shows that cognitive consistency is a sine qua non factor in configuration of an attitude and the process of its change (e.g., Monroe & Read, 2008; Simon et al., 2004). We also know that the need for cognitive consistency would increase as the attitude strength towards an object increases (see Howe & Krosnick, 2017). By showing that network connectivity is a proxy measurement of attitude strength with regards to anti-Roma evaluations as well, the practical implication of our findings would be to identify and target the most central nodes in anti-Roma attitude networks. This would be a useful means for intervention efforts to combat anti-Roma bias as in case the most central nodes being at odds with the others, the system should tend to retain consonance as the connectivity between the nodes is an expression of attitude strength and its related properties such as consistency and stability. This requires further empirical investigations, concerning research on stereotype dynamics, whether interventions based on the variables with the highest degree of centrality would actually render the most favorable results.

Furthermore, our findings show that regarding the node predictability metric, perceived threat to national identity in all the networks and regarding the strength metric, empathy in Hungary, perceived threat to national identity in Romania, and sympathy in

Slovakia, France, and Ireland were the most central values. Since all the most central values are of affective nature, our findings suggest that interventions may induce the most favorable impact if the focus was on affective components rather than cognitive components (stereotypes for example) of the social perception of the Roma. This is consistent with the Intergroup Emotions Theory (e.g., Mackie et al., 2008) as well as the Behaviors from Intergroup Affect and Stereotypes (BIAS) Map (Cuddy et al., 2007) that suggest the crucial role of intergroup emotions in predicting relevant behavior. Moreover, our findings also resonate with the literature on intergroup anxiety, proposing the central role of the affective component of intergroup anxiety in prejudice reduction interventions (see Stephan, 2014).

In short, we argue that employing network approach, taking network connectivity as a theoretical backbone into consideration, can be a useful tool to depict a complex representation of stereotypical evaluations being to direct and unique connections with each other, to identify values with strongest associations. Finding the most influential values would enable us to carry out the most effective attitude change interventions. In addition, we propose that the nature and order of central values, as well as other properties of network dynamic of high attitude strength networks, should be taken into account as a perhaps more informative picture of understanding the nature of interconnectedness between different anti-Roma stereotypical evaluations.

5.6 References

Amnesty International. (2013). *Amnesty International Report 2013: The State of the World's Human Rights*. Amnesty International Limited.

Barrat, A., Barthelemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 3747-3752.

Bilewicz, M., Soral, W., Marchlewska, M., & Winiewski, M. (2017). When authoritarians confront prejudice. Differential effects of SDO and RWA on support for hate-speech prohibition. *Political Psychology*, *38*, 87-99.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, *92*, 631–648. http://dx.doi.org/10.1037/0022-3514 .92.4.631

Dalege, J., Borsboom, D., van Harreveld, F., & van der Maas, H. L. (2018). A network perspective on attitude strength: Testing the connectivity hypothesis. *Social Psychological and Personality Science*, 1948550618781062.

Enyedi, Z., Fábián, Z., & Sik, E. (2004). Is prejudice growing in Hungary? Changes in anti-semitism, anti-roma feeling and xenophobia over the last decade. In T. Kolosi, I. G. Tóth, & G. Vukovich (Eds.), *Social report 2004* (Vol. 21, pp. 363–385). Budapest, Hungary: Tárki.

Epskamp, S. (2016). *Regularized Gaussian psychological networks: Brief report on the performance of extended BIC model selection*. Retrieved from http://arxiv.org/abs/1606.05771

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*, 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., Cramer, A. O., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*, 1-18.

European Commission, (2015). *Special Eurobarometer 437*: Discrimination in the EU in [2015] Retrieved from: http://dx.doi.org/10.2838/325154.

Feischmidt, M., Szombati, K., & Szuhay, P. (2013). Collective criminalization of Roma in Central and Eastern Europe: social causes, circumstances and consequences. In S. Body-Gendrot, M. Hough, K. Kerezsi, R. Levy, & S. Snacken (Eds.), *The routledge handbook of european criminology* (pp. 168–187). London UK: Routledge.

Fiske, S. T. (2015). Intergroup biases: A focus on stereotype content. *Current opinion in behavioral sciences*, *3*, 45-50.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality & Social Psychology*, *82*, 878–902. doi:10.1037//0022-3514.82.6.878

Foygel, R., & Drton, M. (2010). Extended Bayesian information criteria for Gaussian graphical models. In J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, & A. Culotta (Eds.), *Advances in neural information processing systems 23* (pp. 604–612). Red Hook, NY: Curran Associates.

Fraser, A. (1995). *The Gypsies* (2nd ed.). Oxford, UK: Blackwell.

Fried, E. I., Eidhof, M. B., Palic, S., Costantini, G., Huisman-van Dijk, H. M., Bockting, C. L. H., . . . Karstoft, K.-I. (2018). Replicability and generalizability of Posttraumatic Stress Disorder (PTSD) net- works: A cross-cultural multisite study of PTSD symptoms in four trauma patient samples. *Clinical Psychological Science*, *6*, 335–351. https://doi.org/10.1177/2167702617745092

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*, 432–441. doi:10.1093/biostatistics/kxm045

Golino HF, Christensen AP (2019). *EGAnet: Exploratory Graph Analysis - A framework for estimating the number of dimensions in multivariate data using network psychometrics*. R package version 0.8.0.

Gould, R. (2015). Roma rights and Roma expulsions in France: Official discourse and EU responses. *Critical Social Policy*, *35*, 24-44.

Grigoryev, D., Fiske, S. T., & Batkhina, A. (2019). Mapping ethnic stereotypes and their antecedents in Russia: The Stereotype Content Model. Frontiers in Psychology, 10(1643), 1–21. https://doi.org/10.3389/fpsyg.2019.01643

Hadarics, M., & Kende, A. (2019). Negative stereotypes as motivated justifications for moral exclusion. *The Journal of social psychology*, *159*, 257-269.

Hancock, I. (2004). Romanies and the Holocaust: a re-evaluation and overview. In D. Stone (Ed.), *The historiography of the Holocaust* (pp. 383–396). London, UK: Palgrave Macmillan.

Haslbeck, J. M. B. (2015). *Package "mgm": Estimating mixed graphical models*. Retrieved from https://cran.r-project.org/web/packages/mgm/index.html

Howe, L. C., & Krosnick, J. A. (2017). Attitude strength. *Annual review of psychology*, *68*, 327-351.

Imhoff, R., & Bruder, M. (2014). Speaking (un-) truth to power: Conspiracy mentality as a generalised political attitude. *European Journal of Personality*, *28*, 25-43.

Kende, A., Hadarics, M., & Lášticová, B. (2017). Anti-Roma attitudes as expressions of dominant social norms in Eastern Europe. *International Journal of Intercultural Relations*, *60*, 12-27.

Kende, A., Hadarics, M., Bigazzi, S., Boza, M., Kunst, J. R., Lantos, N. A., Lášticová, B., Minescu, A., Pivetti, M., & Urbiola, A. (2020). The last acceptable prejudice in Europe? Anti-Gypsyism as the obstacle to Roma inclusion. *Group Processes & Intergroup Relations*. (advance online publication) https://doi.org/10.1177/1368430220907701

Krosnick, J. A., & Petty, R. E. (1995). Attitude strength: An overview. *Attitude strength: Antecedents and consequences*, *1*, 1-24.

Krosnick, J. A., & Smith, W. A. (1994). Attitude strength. Encyclopedia of human behavior. VS Ramachandran.

Ladányi, J. (2001). The Hungarian neoliberal state, ethnic classification, and the creation of a Roma underclass. In R. J. Emigh, & I. Szelényi (Eds.), *Poverty*, *ethnicity*, *and gender in Eastern Europe during the market transition* (pp. 67–82). Westport, CT: Greenwood Publishing Group.

Mackie, D. M., Smith, E. R., & Ray, D. G. (2008). Intergroup emotions and intergroup relations. *Social and Personality Psychology Compass*, *2*, 1866–1880.

Mahoney, D. E. (2011). Expulsion of the Roma: Is France violating EU freedom of movement and playing by French rules or can it proceed with collective Roma expulsions free of charge. *Brooklyn Journal of International Law*, *37*, 649-682.

McGuire, W. J. (1990). Dynamic operations of thought systems. *American Psychologist*, *45*, 504–512.

Messing, V. 2017. "Differentiation In The Making: Consequences Of School Segregation Of Roma In The Czech Republic, Hungary, And Slovakia." *European Education* 49: 89-103. doi: 10.1080/10564934.2017.1280336.

Monroe, B. M., Read, S. J. (2008). A general connectionist model of attitude structure and change: The ACS (Attitudes as Constraint Satisfaction) model. *Psychological Review*, *115*, 733–759.

Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A critical look at evidence from the field and the laboratory. *Annual Review of Psychology*, *60*, 339-367.

Pogány, I. (2006). Minority rights and the Roma of Central and Eastern Europe. *Human Rights Law Review*, *6*, 1–25. http://dx.doi.org/10.1093/hrlr/ngi034.

Sayans-Jiménez, P., Harreveld, F., Dalege, J., & Rojas Tejada, A. J. (2018). Investigating stereotype structure with empirical network models. European Journal of Social Psychology, 49(3), 604–621. https://doi.org/10.1002/ejsp.2505

Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., . . . Vandermeer, J. (2012). Anticipating critical transitions. *Science*, *338*, 344–348.

Schmittmann, V. D., Cramer, A. O., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, *31*, 43-53.

Simon, D., Snow, C. J., Read, S. J. (2004). The redux of cognitive consistency theories: Evidence judgments by constraint satisfaction. *Journal of Personality and Social Psychology*, *86*, 814–837.

Stanciu, A., Cohrs, J. C., Hanke, K., & Gavreliuc, A. (2017). Within-culture variation in the content of stereotypes: application and development of the stereotype content model in an Eastern European culture. *The Journal of social psychology*, *157*, 611-628.

Stephan, W. G. (2014). Intergroup anxiety: Theory, research, and practice. *Personality and Social Psychology Review*, *18*, 239-255.

Tileagă, C. (2006). Representing the 'Other': A discursive analysis of prejudice and moral exclusion in talk about Romanies. *Journal of Community & Applied Social Psychology*, *16*, 19-41.

van Baar, Huub. 2011. ''Europe's Romaphobia: Problematization, Securitization, Nomadiza- tion.'' *Environment and Planning D: Society and Space* 29:203-12.

van Borkulo, C. D., & Epskamp, S. (2015). *IsingFit: Fitting Ising models using the eLasso method* (R package Version 0.3.0) [Computer Software]. Retrieved from https://cran.r-project.org/web/ packages/IsingFit/index.html

van Borkulo, C. D., Boschloo, L., Kossakowski, J. J., Tio, P., Schoevers, R. A., Borsboom, D., & Waldorp, L. J. (2017*). Network Comparison Test—Comparing network structures on three aspects: A permutation test*. Retrieved from https://github.com/cvborkulo/NetworkComparison Test

Villano, P., Fontanella, L., Fontanella, S., & Di Donato, M. (2017). Stereotyping Roma people in Italy: IRT models for ambivalent prejudice measurement. *International Journal of Intercultural Relations*, *57*, 30-41.

6   Study III: Exploring inclusiveness towards immigrants as related to basic values: A network approach

6.1   Abstract

Using the 9th round of European Social Survey (ESS), we explored the relationship between Europeans' basic values and their attitudes towards immigrants. Employing a latent class analysis (LCA), we classified the respondents based on three items capturing the extent to which participants would support allowing three groups of immigrants to enter and live in their countries: immigrants of same ethnic groups, immigrants of different ethnic groups, and immigrants from poorer countries outside Europe. Four *classes* of Europeans with mutually exclusive response patterns with respect to their inclusive attitudes towards immigrants were found. The classes were named Inclusive (highly inclusive), Some (selective), Few (highly selective), and Exclusive (highly exclusive). Next, using a network technique, a partial correlation network of Schwartz's (1992) ten basic values was estimated for each class of participants. The four networks were compared to each other based on three network properties namely: *global connectivity*, *community detection*, and *assortativity coefficient*. The global connectivity (the overall level of interconnections) between the 10 basic values was found to be mostly invariant across the four networks. However, results of the community detection analysis revealed a more complex value structure among the most inclusive class of Europeans. Further, according to the assortativity analysis, as expected, for the most inclusive Europeans, values with similar motivational backgrounds were found to be interconnected most strongly to one another. We further discussed the theoretical and practical implications of our findings.

6.2   Introduction

Immigration is still among the most controversial topics in the Western political sphere. On the one hand, right-wing parties and populist leaders center their rhetoric around anti-immigration sentiments (see e.g., Norris & Inglehart, 2019; Vachudova, 2021), and on the other hand, anti-immigration attitudes are a significant motive behind supporting and voting for right-wing parties (e.g., Dennison, 2020; Edgell, 2017; Lubbers & Coenders, 2017; Schaffner et al., 2018). In recent years, the number of immigration-related research has significantly increased, perhaps partly due to the above-mentioned reasons.

The interplay between a variety of contextual and individual level factors has been argued to explain anti-immigrant attitudes and opposition to immigration. Poor economic condition of the host society (Cochrane & Nevitte, 2014), anti-immigration media portrayals (Matthes & Schmuck, 2017; Schemer, 2012; Wirz et al., 2018), immigration-related restrictive policies (Green et al., 2020; Schlueter et al., 2020), regional origin of migrants (Máté et al., 2018), and immigrants' actual population size (Ceobanu & Escandell, 2010) are among the most important contextual factors on mobilizing anti-immigrant attitudes. Among the individual level predictors of anti-immigrant attitudes are perceived economic threat and competition (Card et al., 2005; Fussell, 2014), perceived cultural threat (Schmuck & Matthes, 2015, 2017), perceived threat vis-à-vis safety and security (Esses et al., 2019; Hartman et al.,

2020; Landmann et al., 2019), nativist perception of national identity (Pehrson & Green, 2010; Taniguchi, 2021), ideological orientations such as right-wing authoritarianism and social dominance orientation (Roy et al., 2021) , stereotypical judgements (Grigoryev et al., 2019; Lee & Fiske, 2006), low degrees of personality traits such as openness to experience and agreeableness (Talay & De Coninck, 2020), perceived size of immigrants' population (Gorodzeisky & Semyonov, 2020), low degree of interpersonal trust (Mitchell, 2021; Pellegrini et al., 2021), and low degree of trust in national institutions (e.g., Danaj et al., 2018) (for an overview see also Esses, 2021). Furthermore, tapping into a more positive aspect of immigration, it is noteworthy that groups of immigrants who are economically and culturally perceived as less threatening and more beneficial (e.g., immigrants from western societies vs. asylum seekers) receive more positive attitudes from the members of the host society (see the Threat-Benefit Model, Tartakovsky & Walsh 2020; Walsh & Tartakovsky 2021).

In addition to the factors outlined above, it is also well-documented that the individual differences on basic personal values are strongly related to the endorsement of positive or negative attitudes and behaviors towards minority group members in general and immigrants in particular (see below). In the present study, using nationally representative samples, we draw on Schwartz's (1992) well-established and cross-culturally validated value map (Davidov, 2010), as related to Europeans' inclusive attitudes towards immigrants. By inclusive attitudes, we mean the extent to which one would support allowing three types of immigrants to enter their country—immigrants of same ethnic groups, immigrants of different ethnic groups, and immigrants from poorer countries outside Europe.

Applying a network method, the objective of this study is to explore and compare the dynamic structure of the basic personal values held by different *classes* of Europeans with qualitatively different inclusiveness attitudes towards immigrants. Our study, therefore, comprises two analytical phases. First, a person-centered method (latent class analysis) is applied for the classification of the respondents. Second, partial correlation networks are estimated to further investigate the dynamic structure of the basic personal values for each distinct class found in the first phase of the analyses. Below we discuss the relationship between personal values and attitude towards immigrants as well as both of the analytical approaches in more detail.

6.2.1 Basic personal values and anti-immigrant attitudes

Values are defined as "*desirable transsituational goals, varying in importance, that serve as guiding principles in the life of a person or other social entity*" (Schwartz, 1994, p. 21). They are abstract and superordinate standards that determine individuals' worldviews, attitudes, and behaviors in a vast array of more specific situations and contexts (Rokeach, 1973; Schwartz,1992). Schwartz (1992), proposes a comprehensive value system into a circular dynamic of 10 basic personal values, representing the adjacent values capturing more similar motivational states and the values distant from each other to be motivationally more dissimilar or even antagonistic (see Figure 1, adapted from Magun et al., 2015).
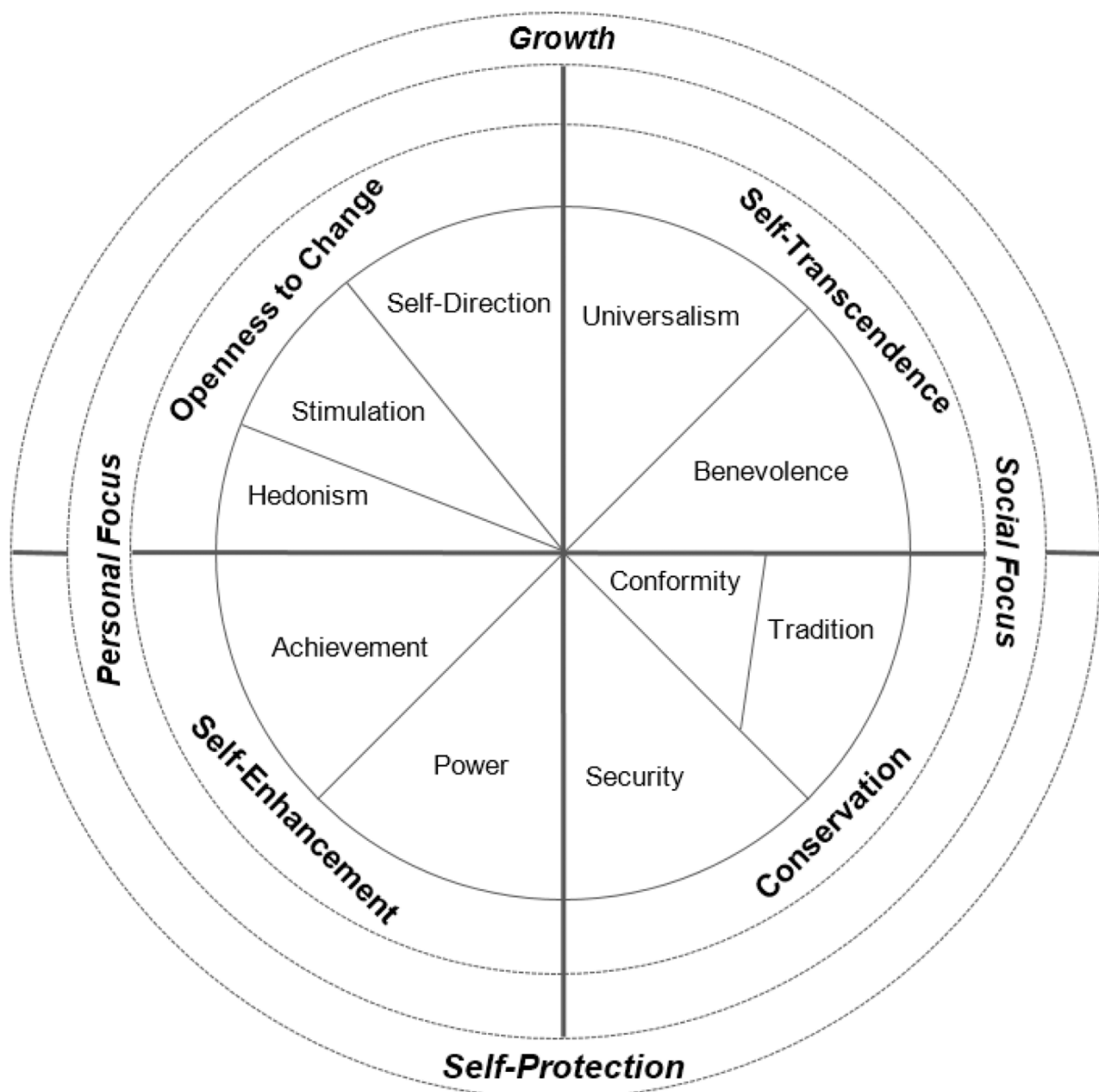

Figure 1. Schwartz value circle depicting the relations between 10 values and several value groupings (Schwartz, 1992 , 2006).

The circumplex structure of the Schwartz's values allow them to be categorized into 4 higher-order value types; Openness to change (Hedonism, Stimulation, Self-direction), Conservation (Security, Conformity, Tradition), Self-Transcendence (Universalism, Benevolence), and Self-Enhancement (Achievement, Power) (See supplementary material for the conceptual definitions of the 10 basic values). A pair of opposing value types (openness to change vs. conservation and self-enhancement vs. self-transcendence); in turn form two higher-order dimensions of values. Openness to change values focus on independence and seeking new experience and ideas in life, opposing conservation values that refer to the motivational goals of protecting social norms and traditions. On the other dimension, self-transcendence values tap into the well-being of other people, opposing self-enhancement values that focus on one's own personal interests and welfare (Schwartz, 1992). The 10 basic values can also be grouped into two other alternative higher-order value dimensions. Either as Social Focus (self-transcendence, conservation) versus Personal Focus (openness to change and self-enhancement), or Self-Protection (self-enhancement, conservation) versus Growth (openness to change, self-transcendence) (Schwartz, 2006). In what follows, the emphasis of this article is on the latter distinction, between self-protection and growth categories, as two motivationally opposing higher-order value types.

Furthermore, empirical investigations have well-documented that self-protection value types positively and growth value types negatively relate to anti-immigrant attitudes (see e.g., Beierlein et al., 2016; Davidov et al., 2020; Gusciute et al., 2020; Schwartz et al., 2010; Walsh & Tartakovsky, 2021; Wolf et al., 2019).

6.2.2 Person-centered approach

Person-centered methods treat the individuals and not the variables as the units of analysis. This allows researchers to raise questions which cannot be answered by the more traditional variable-centered methods. By having at least two indicators (i.e., observed variables), it is possible to investigate whether there exist unobserved latent classes or subgroups of individuals whose response patterns are unique and qualitatively distinct from each other (Bauer & Shanahan, 2007). Classifying respondents into a set of mutually exclusive classes, for example, makes it possible to investigate what factors precede class membership, what the consequences are, and how class membership relates to demographic as well as contextual correlates. Person-centered methods can be especially useful and highly informative as long as the data represent the corresponding population. By a representative sample one would be more confident that the nature of the classes as well as their association

with other constructs correspond to those of the population of interest (For an overview on person-centered approach in social psychological research see Osborne & Sibley, 2017). A person-centered method, therefore, is suitable for the first phase of our analyses. That whether there exist distinct classes of Europeans with meaningfully different response patterns regarding their inclusive attitudes towards the three groups of immigrants.

Person-centered methods have been recently applied in intergroup relation research. For example, Adelman & Verkuyten (2020), found four distinct profiles with combinations of prejudice against Muslims and the level of participants' tolerance on different Muslim practices (e.g., celebrating Islamic holidays in public). Through a variable-centered method, for instance, it was possible to regress the acceptance of Muslim practices on prejudice against Muslims. By a person-centered method, apart from finding profiles of people who were both prejudiced and against Muslim practices, they also found profiles of respondents who objected against some Muslim practices while not being particularly prejudiced. Another example is a study by Sibley et al (2019), where they found five distinct constellations of right-wing authoritarianism (RWA) and social dominance orientation (SDO). Among the five, three classes were found to be low, mild, or moderate on both SDO and RWA. In line with the assumption made by Altemeyer (1998), they also found two other classes of respondents: authoritarian followers (Low SDO and High RWA) and authoritarian leaders (High SDO and low RWA). While the positive correlation between SDO and RWA is well-documented (e.g., Roccato & Ricolfi, 2005), by applying a person-centered method, was it possible to find profiles of respondents being low in SDO but high in RWA or the other way around.

With regards to the present study, instead of applying a person-centered method, one possibility was to divide the participants arbitrarily. For example, based on median splits, it is possible to create classes of participants with all the possible combinations of high, moderate, and low levels of the indicator variables. However, since we used nationally representative samples, as mentioned above, one could be more confident that the response patterns found by the person-centered technique would be consistent with those of the corresponding population (European population in the case of this study). Further, research indicates that splitting the data by median would increase the risk of power loss and the emergence of spurious effects (see MacCallum et al., 2002).

6.2.3 Network analysis

Network methods are a suitable technique for conceptualizing a number of psychological variables as a complex system in which every single variable is in direct associations with all the other variables. Unlike a latent model, for example, that assumes the observed variables to measure the latent variable equivalently, networks allow picturing the dynamic of all the existing unique and direct pairs of interconnections (Schmittmann et al., 2013). More traditional correlation-based research is often parsimonious and uses a limited number of variables for analysis. Modeling a relatively larger number of variables as a network, enables researchers to have a more comprehensive picture of all the relationships between the variables. Network methods have been recently employed in multiple domains of psychological research, investigating, for instance, personality traits (Costantini et al., 2015), mental disorders (Fried et al., 2017), and stereotype structure (Sayans-Jiménez et al., 2019). More specifically, attitudes towards immigrants have been also investigated through network models (see e.g., Grigoryev et al., 2019; Phua et al., 2020).

Nodes and edges are the most primary constituent elements of a network. Typically, in psychological models, nodes are a set of observed variables and edges the direct and unique statistical associations (i.e., partial correlations) between every possible pair of edges (Epskamp et al., 2018).

Connectivity is another basic network property that refers to the extent to which the nodes within a network are interconnected. In psychological networks, connectivity refers to the level of causal interdependencies between a set of nodes. The stronger the interconnection between two nodes, the more likely it is that changes to one node will also result in changes to the other. This is because as two cognitive components are strongly related to each other they need to maintain consistency, so changes to one cognitive component (smoking causes cancer or not?) should lead to changes to the other (should I smoke or not?) and vice versa (Festinger, 1957; Gawronski & Strack, 2012; Heider, 1958). Thus, in strongly connected networks, the nodes yield stronger level of causal effect on each other, while in weakly connected networks they would behave more accidentally and less dependent upon the impact of one another (for a detailed overview on the relationship between network connectivity and causality in network science and its integration with psychological networks see Dalege et al., 2016). Furthermore, previous research shows that as one holds a strong attitude, different kinds of one's evaluations on the corresponding attitude object, are more strongly interconnected or causally interdependent. In other words, the connectivity between

an attitude's underlying components (i.e., cognitive, affective, and behavioral) predicts attitude strength towards the attitude object (see the causal attitude network (CAN) model, Dalege et al., 2018). For example, the connectivity between a set of anti-Roma evaluations was found to be stronger for those who held stronger attitude towards the Roma people (Nariman et al., 2020). Moreover, a group of nodes being strongly interconnected, manifests their belonging to a similar state (Dalege et al., 2016). Global connectivity, for instance, as one measurement of network connectivity, refers to the absolute sum of all the (partial) correlation coefficients within a network.

As mentioned above, Schwartz (1992) proposes the 10 basic values in a circular structure suggesting a dynamic of relations between the values. Network analysis, hence, would be a highly efficient method to representing and further investigating this dynamic. An extensive body of research has tested the relationship between the 10 basic values and intergroup attitudes (*cf. supra*). In the current study, we go beyond by exploring the relationships between the 10 basic values, estimated as partial correlation networks, for different classes of Europeans with unique inclusive strategies towards immigrants.

It is worth noting that person-centered methods are typically applied in exploratory research. Network methods have been recently used in exploratory papers as well (see for example Abacioglu et al., 2019; Bhushan et al., 2019; Verschoor et al., 2020). In spite of the exploratory nature of our study, however, we hold two main assumptions.

Concerning the person-centered analysis, although there are no theory-driven explanations regarding the exact number and nature of the classes, we follow previous research on expecting a number of classes aligned along a spectrum from the most inclusive to the most exclusive. For example, Morselli & Passini (2018), found six profiles of Europeans (using the 6th round of ESS data) on inclusiveness towards immigrants and protest against institutional authority—ranging from inclusive protestors to exclusive protestors. Our first assumption, therefore, is that there would be a continuum of distinct typologies of individuals' inclusiveness towards immigrants from the most inclusive to the most exclusive.

Moreover, a recent study found that compared to conservatives, in the network of liberals' moral values, the interconnections within each set of more egalitarian (i.e., individualizing) and less egalitarian (i.e., binding) moral values was stronger than the interactions between them (Turner-Zwinkels et al., 2020; for an overview on binding and individualizing moral values see Graham et al., 2009). That is to say, in the liberal moral system, the values within each of the individualizing or binding categories, were found to be

more interconnected while the interconnections between them was comparatively weaker. Past research shows that individualizing moral values negatively and binding moral values positively associate with intergroup prejudice (see eg., Hadarics & Kende, 2018; Kugler et al., 2014; Van de Vyver et al., 2016). Also, compared to conservatives, liberals tend to express lesser degree of generalized prejudice (McFarland, 2010) and evaluate different outgroups more positively (e.g., Bowman & West, 2021; Löw et al., 2020; Terrizzi Jr et al., 2010). One, therefore, may argue that liberals evaluate the outgroup more positively, employing a set of strongly interconnected individualizing moral values, and their evaluation is more independent from the causal effect of other binding moral values. It is also worth pointing out that, generally, individualizing moral values are related to growth value types and binding moral values are related to self-protection value types (see for example Feldman, 2020). Thus, our second assumption is that in the most inclusive group's value network, the interconnections within each set of more egalitarian (i.e., growth) and less egalitarian (i.e., self-protection) basic values would be the strongest and the interconnections between them would be the weakest.

6.3   Data and methods

Our study is based on nationally representative interview-based survey data using the 9th round of European Social Survey (ESS), edition 1.2, collected in 2018. Verbal informed consents were obtained from all participants prior to the interviews being conducted. A total of 19 European countries participated in the data collection process: Austria, Belgium, Bulgaria, Cyprus, Czechia, Estonia, Finland, France, Germany, Hungary, Ireland, Italy, Netherlands, Norway, Poland, Serbia, Slovenia, Switzerland, and United Kingdom. Since we were interested in the adult population, out of the initial number of participants (N = 36015), we removed 1070 participants because they were under eighteen years old. Overall, 34945 participants (M_{age} = 51.6, SD = 18.1; 52.9 % female) were included in the analyses. Table 1 reports the sample sizes and the descriptive statistics for the participants' age and gender per country.

Table 1

Sample sizes and descriptive statistics for age and gender by country

Country	N %	Females	M_{age} (SD)

Austria	2416	54.1	52.2 (17.5)
Belgium	1689	50.8	49.1 (18.5)
Bulgaria	1949	54.1	55.7 (16.9)
Cyprus	760	52.9	54.8 (18.3)
Czechia	2356	55.9	49.3 (17.3)
Estonia	1843	56.3	51.7 (18.6)
Finland	1674	51.2	52.1 (18.2)
France	1928	54.5	53.2 (18.4)
Germany	2258	49	50.8 (18.3)
Hungary	1603	57	51.7 (18.0)
Ireland	2140	52.6	52.6 (17.4)
Italy	2 617	52.8	51.9 (18.9)
Netherlands	1569	50.4	50.2 (17.8)
Norway	1308	44.6	48.4 (17.5)
Poland	1416	52.4	48.8 (18.1)
Serbia	1962	51.5	54.0 (17.5)
Slovenia	1260	53.8	50.4 (18.1)
Switzerland	1440	49.7	48.7 (18.1)
United Kingdom	2135	54.7	52.9 (18.1)

Note. *M* and *SD* indicate mean and standard deviation respectively.

6.3.1 Measures

The indicator variables used in the classification procedure (latent class analysis) are three items measuring respondents' support for allowing three groups of immigrants to enter their countries: (1) "To what extent do you think [country] should allow people of the same race or ethnic group as most [country]'s people to come and live here?" (2) "How about people of a different race or ethnic group from most [country] people?" (3) "How about people from the poorer countries outside Europe?". All the three items were measured on a four-point scale (1 = Allow many to come and live here; 2 = Allow some; 3 = Allow a few; 4 = Allow none). In order to avoid possible misclassification of the respondents, we also controlled for the effect of a number of covariates namely political ideology, interest in politics, anti-immigrant

evaluations, personal values, and participants' demographic characteristics (gender, age, and education).

Political ideology was measured by an 11-point scale ("In politics people sometimes talk of "left" and "right". Where would you place yourself on this scale, where 0 means the left and 10 means the right?").

Interest in polities was measured on a four-point scale ("How interested would you say you are in politics?") from *very interested* (1) to *not at all interested* (4).

Anti-immigrant attitudes were measured by economic threat perception ("Would you say it is generally bad or good for [country]'s economy that people come to live here from other countries?"), symbolic threat perception ("Would you say that [country]'s cultural life is generally undermined or enriched by people coming to live here from other countries?"), and a more general anti-immigrant attitude ("Is [country] made a worse or a better place to live by people coming to live here from other countries?"). All the three items were 11-point scales form *Bad for the economy*/*Cultural life undermined*/*Worse place to live* (0) to *Good for the economy*/*Cultural life enriched*/*Better place to live* (10) respectively.

Personal values were measured by the 21-item Portrait Values Questionnaire (PVQ-21, Schwartz, 2007). Each item portrays a person for whom a certain motivational value type is important and asks the respondent about the extent to which they find themselves similar, from *Very much like me* (1) to *Not like me at all* (6). For instance, "[She]He thinks it is important that every person in the world should be treated equally. [She]He believes everyone should have equal opportunities in life.", is one of the items tapping into universalism. Each of the 10 basic values was measured by two items—universalism by three. The 10 basic values, therefore, were first used as the covariates for classifying the participants, and later they were estimated as partial correlation networks to picture the dynamic of each group's personal values.

6.3.2   Latent Class Analysis

Using Mplus software, version 8 (Muthén & Muthén, 2017), we applied a latent class analysis, a model-based person-centered method, to classify the respondents on the basis of their inclusiveness towards immigrants—using the three items mentioned above. Since apart from detecting the latent classes, we were also interested in the effect of our covariates on the participants' class membership, we used a 3-step latent class method suggested by Vermunt (2010). It is suitable for both detecting the latent classes, based on class membership probabilities, as well as controlling for the effect of covariates in determining the classes.

Employing this method, first, the latent class model is estimated only based on the indicator variables. In the second step the most likely class membership variable is generated. And in the third step the class membership variable is regressed on the covariates, while also any potential misclassification occurred in the second step is fixed (see also Asparouhov & Muthén, 2014). In addition to the covariates mentioned above, we also created 18 country dummies (Slovenia as the reference group) to account for the country effects.

Further, since the three-step method is sensitive to missing values (on the covariates) and to avoid listwise deletion of the observations, the dataset was imputed. The proportion of the missing values was under 5 per cent for all the items used in the study ranging from 0 % to 4.2 % (except for political ideology with 13.8 % missing values). Following Asparouhov & Muthén (2010), we applied a two-level multiple imputation technique, taking into account the participants' being nested into their countries—assuming the data to be missing at random.

In order to find the best model solution, we first built a 2-class model, and increased the number of classes up to a 6-class model and compared the 5 models with each other. We decided on the best model solution based on statistical criteria, parsimony, and interpretability (Nylund-Gibson & Choi, 2018). Four statistical criteria were used: Bayesian Information Criterion (BIC), Akaike's Information Criterion (AIC), entropy, and the Vuong–Lo–Mendell–Rubin likelihood ratio (VLMR) test. The entropy value ranges from 0 to 1, the closer it is to 1, to a higher extent it indicates reliability of the classification, and separability between the classes (Ramaswamy et al., 1993). The VLMR test compares the k-class model with the k-1-class model and provides a *p*-value to test whether the k-class model fits the data significantly better (Lo et al., 2001).

Moreover, to refrain from the impact of sampling biases on the classification of the respondents, the data were weighted. Because the sample sizes for each country are similar but the population sizes are different, the data were weighted by population to avoid the overrepresentation of smaller countries. We also weighted the data by design, correcting for the fact that the likelihood of the respondents to be represented in the data varies by country (see Weighting European Social Survey Data, 2014).

6.3.3 Network analysis

The 10 basic values were estimated as partial correlation networks for each class of the respondents found in the latent class analysis. After having checked for the accuracy of the

edge weights, we explored and compared the networks based on global connectivity, community detection, and assortativity coefficient.

6.3.3.1 Network estimation

To estimate the networks, for each class, the correlations between the 10 basic values were computed and inverted into partial correlations to obtain all the existing unique and direct statistical associations. We used *ggmModSelect* function in R from "qgraph" package (Epskamp et al., 2012). The function first runs the *graphical least absolute shrinkage and selection operator* (glasso) algorithm for estimating 100 different network models with 100 different tuning parameters. Next, glasso fits the networks on *unregularized Gaussian graphical models* (GGM) and choses the best model based on BIC. Lastly, glasso adds and removes edges until it finds the model with the best BIC value.

6.3.3.2 Global connectivity

The global connectivity of the networks was computed and compared to each other using the R package "NetworkComparisonTest" (NCT; van Borkulo et al., 2017). Permutations tests, with 1000 iterations, were run to check whether or not the overall level of interconnections is invariant across the networks.

6.3.3.3 Community detection

A community (or cluster) is a group of nodes that are densely connected to each other and are more weakly connected to other nodes in the network. It is, therefore, possible that different sets of highly connected nodes form a number of sub-networks within the bigger network. To find the communities in the networks we used *Walktrap* algorithm from the "igraph" R package (Csardi & Nepusz, 2006).

6.3.3.4 Assortativity coefficient

As mentioned above we assumed that in the network of the most inclusive class, the interactions between growth and self-protection value types would be the weakest and the interactions within each category would be the strongest. In order to test our assumption, first, using R package "assortnet" (Farine, 2014), we computed assortativity coefficient that quantifies the extent to which groups of nodes (growth vs. self-protection) within a network tend to be interacting within rather than between each other (Newman, 2003). The assortativity coefficient ranges from -1 to 1 and values closer to 1 reflect the stronger tendency of nodes to interact within-group rather than between-group. Next, using r package "boot" (Canty & Ripley, 2017) we compared the networks by calculating 95 % confidence intervals obtained from 1000 bootstrap resamples.

6.3.3.5 Network stability

Prior to the main analyses, we performed a bootstrapping technique, checking for the accuracy of the edge weights. Using R package "bootnet" (Epskamp et al., 2018), bootstrapped confidence intervals around the edge weights were obtained from 1000 draws. A network is considered having accurate and therefore interpretable edge weights if the confidence intervals indicate low variabilities (see Epskamp et al., 2018).

6.4 Results

6.4.1 Latent Class Analysis

Table 2 summarizes the model fit comparisons between the 5 models. Compared to the 2-class and 3-class models, the 4-class model solution was found to be better fitting the data in terms of the statistical criteria. Only the entropy of the 3-class model (.92) was similar to that of the 4-class model (.91), while other statistical criteria showed significant improvement of the 4-class model fit. Further, in spite of the slight decrease of the AIC and BIC values in the 5-class and 6-class models, VLMR test showed that increasing the number of classes higher than 4 doesn't significantly improve the model fit. Moreover, compared to the 4-class model, the entropy value in the 5-class model solution (.79) and 6-class model solutions (.74) dropped sharply. Hence, we chose the 4-class model as the best in fitting the data, parsimoniousness, and interpretability.

Table 2

Model fit comparisons from 2-class to 6-class model solutions

Number of classes	AIC	BIC	entropy	VLMR (*p*-value)
2	223449	223610	.89	<.001
3	200344	200589	.92	<.001
4	**190610**	**190940**	**.91**	**<.001**
5	190519	190933	.79	.06
6	190429	190928	.74	.81

Note. AIC and BIC decrease as the model fit improves.

As can been seen in Table 3, in the 4-class model, the probabilities of the respondents' answers showed that in each class the majority of the respondents fall into one of the four categories of the three indicator variables. As mentioned above the four categories were:

allow many, allow some, allow a few, and allow none. In each class the respondents were found to be consistent on their inclusiveness towards the three groups of immigrants. Therefore, supporting our first assumption, one class was found to be the most inclusive regarding all the three immigrant groups, one found to be the most exclusive, and the other two were found to be selective. We named the 4 classes as *Inclusive* (N = 4607), *Some* (N = 14399), *Few* (N = 10987), and *Exclusive* (N = 4952) (see the supplementals for the descriptive statistics and the correlations between the variables for each class).

Table 3

The probabilities of the participants' responses on their inclusive towards immigrants by class membership.

Indicatior Variables	Class			
	Inclusive	Some	Few	Exclusive
imsmetn				
Allow many	**98%**	16%	7%	4%
Allow some	2%	**81%**	35%	14%
Allow a few	0%	2%	**56%**	25%
Allow none	0%	0%	1%	**57%**
imdfetn				
Allow many	**94%**	1%	0%	0%
Allow some	5%	**93%**	4%	0%
Allow a few	0%	6%	**89%**	3%
Allow none	0%	1%	7%	**97%**
impcntr				
Allow many	**82%**	3%	1%	0%
Allow some	15%	**80%**	10%	3%
Allow a few	2%	15%	**75%**	8%
Allow none	1%	2%	14%	**89%**

Note. imsmetn = Allow many/few immigrants of same race/ethnic group as majority;

imdfetn = Allow many/few immigrants of different race/ethnic group from majority;

impcntr = Allow many/few immigrants from poorer countries outside Europe.

Table 4 reports the results of the multinomial logistic regressions for class membership predicted by the covariates (with the Inclusive class as the reference category). Exclusive class members were found to be significantly more interested in politics, while political interest did not predict Some and Few class memberships significantly. Moreover, political conservatism and the three items tapping into threat perception significantly predicted Exclusive, Few, and Some class memberships. Moreover, concerning the self-protection value types, the importance of security and tradition negatively predicted Exclusive, Few, and Exclusive class memberships. Conformity negatively predicted only the Exclusive class membership, and achievement negatively predicted only the Some class membership. Moreover, power did not associate with the three class memberships compared to the Inclusive class. Regarding the growth value types, benevolence and universalism positively predicted Exclusive, Few, and Some class memberships. Hedonism, stimulation, and self-direction were not found to be predictive of the participants class memberships. Regarding the demographic variables, participants' gender (women as the reference category) significantly and negatively associated with the Few and Some class memberships and age positively predicted the membership of the Exclusive, Few, and Some classes. Further, the level of participants' education negatively and significantly predicted the Exclusive and Few class memberships (see also Table S6 in the supplementals for the demographic characteristics of each class).

Table 4

Results of multinomial logistic regressions for predicting class membership by the covariates (Inclusive class as the reference category)

Variables	Exclusive	Few	Some
Political Interest	.24***	.08	-.02
Political Ideology	.28***	.25***	.15***
imbgeco	-.68***	-.41***	-.21***
imueclt	-.49***	-.38***	-.22***
imwbcnt	-.50***	-.35***	-.17***
Gender	-.17	-.24***	-.18***
Age	.03***	.03***	.02***

|			
Education	-.08***	-.03***	-.006
Security	-.32***	-.29***	-.22***
Conformity	-.15***	-.05	.001
Tradition	-.11*	-.13***	-.10**
Benevolence	.26***	.26***	.13*
Universalism	.93***	.67***	.40***
Self-direction	.09	.06	.04
Stimulation	-.02	-.02	-.006
Hedonism	-.03	-.03	.03
Achievement	-.08	-.07	-.08*
Power	-.07	-.06	-.03

Note. ***= $p < .001$; ** = $p < .01$; * = $p < .05$. imbgeco = immigration is good or bad for economy; imueclt = whether immigration undermines or enriches culture; imwbcnt = immigration makes the country better or worse place to live.

Before proceeding to the next stage of the analyses we also performed a series of independent sample t-tests to compare the four classes based on political ideology, interest in politics, perceived threat towards immigrants, and the 10 basic values. The results showed that the four classes (Inclusive, Some, Few, and Exclusive) ranged from the most liberal and least interested in politics to the most conservative and most interested in politics respectively and all the differences were statistically significant. The four classes were also significantly different from each other regarding the three items measuring threat perception, with the Inclusive class to be the lowest on threat perception (i.e., Inclusive < Some < Few < Exclusive). Similarly, the extent to which the respondents found growth value types important, ranged from the highest for the Inclusive class to the lowest for the Exclusive class (Some and Few in between). This pattern was exactly the other around concerning the self-protection value types. Only, regarding achievement, no significant difference was found in none of the comparisons. Also, the difference between the Few and Exclusive groups on tradition was not found to be significant (see the supplementals for the t-test results).
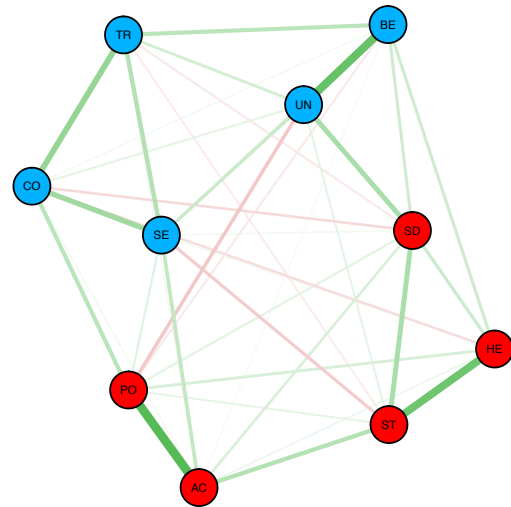
6.4.2 Network analysis

Four networks were estimated based on the respondents' 10 basic values (see Figure 2). For each network, out of 45 possible edges, we checked for the number of non-zero edges (Inclusive = 33, Some = 39, Few = 38, and Exclusive = 38).
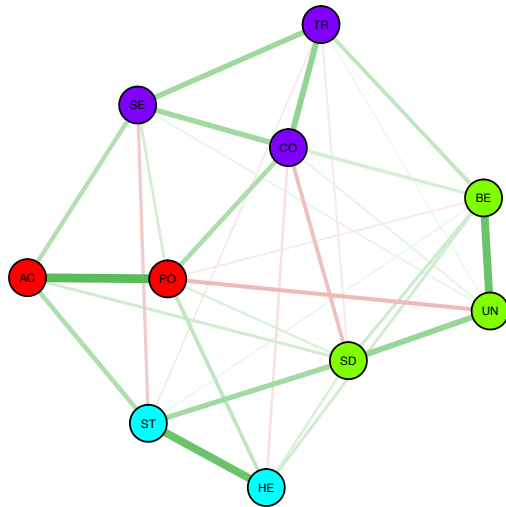
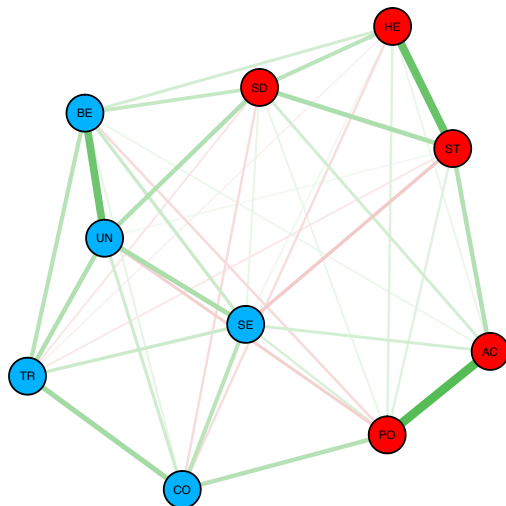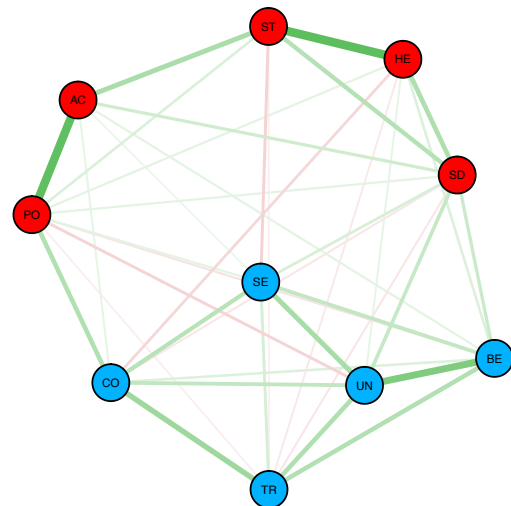**Figure 2**. Partial correlation networks estimated for the 4 classes found in the LCA. Red lines depict negative partial correlations, and the green lines represent positive partial correlations. Node with the same color belong to the same community. The thickness of the lines represents the magnitude of the correlation coefficients. SD = Self-Direction; ST = Stimulation; HE = Hedonism; AC = Achievement;

PO = Power; SE = Security; CO= Conformity; TR = Tradition; BE = Benevolence; UN = Universalism.

6.4.2.1 Global connectivity

The global connectivity scores were found to be more or less similar across the 4 networks (Inclusive = 5.16; Some = 5.14; Few = 5.31; Exclusive = 5.34). The results of the permutation tests showed that, out of 6 possible comparisons only the Few-Some difference was found to be significant (Inclusive vs. Some: diff = .02, p = .92; Inclusive vs. Few: diff = .15, p = .19; Inclusive vs. Exclusive: diff = .18, p = .06; Some vs. Few: diff = .16, p = .02; Some vs. Exclusive: diff = .20, p = .05; Few vs. Exclusive: diff = .03, p = .72).

6.4.2.2 Community detection

As Figure 2 represents, the community detection analysis showed that in Some, Few, and Exclusive networks, personal focus values vs. social focus values formed 2 large separate communities. While the nodes in Inclusive network tended to form 4 communities, highly corresponding to the four higher order value types: openness to change, self-transcendence, self-enhancement, and conservation. The only slight difference was that self-direction clustered with self-transcendence values: universalism and benevolence.

6.4.2.3 Assortativity coefficient

The differences between the 4 networks on the assortativity coefficient were all above chance (Inclusive: r = .81, 95% CI [.79, .84]; Some: r = .66, 95% CI [.64, .67]; Few: r = .51, 95% CI [.48, .53]; Exclusive: r = .39, 95% CI [.35, .43]). Our second assumption was therefore supported, that compared to other networks, in Inclusive network, the interconnections within the growth value types and within the self-protection value types was the strongest, and the interconnections between the two categories was the weakest.

6.4.2.4 Network stability

The stability analyses showed that all the networks were stable and therefore interpretable regarding their edge weights. That is to say, the confidence intervals around the edge weights showed small variabilities, meaning that the edge weight accuracy was attained across all the 4 networks (figures visualizing the confidence intervals around the edge weights can be found in the supplementals).

6.5 Discussion

Using Latent class analysis, we analyzed the 9th round of ESS data, and found 4 mutually exclusive classes of Europeans with meaningfully distinct inclusiveness strategies towards

three groups of immigrants—from the same ethnic background, from different ethnic backgrounds, and from poorer countries outside Europe.

The results suggest that all four classes share generalized attitudes towards immigrants ranging from being highly inclusive to being highly exclusive. In other words, even though one could expect finding classes of respondents with the combination of high, moderate, and low degrees of inclusive attitudes, the results revealed that respondents (in each class) weakly differentiated between the three groups of immigrants. Our results, therefore, may appear to be at odds with the fact that attitudes towards members of different outgroups spring from distinct cognitive, ideological, and motivational sources (see e.g., Duckitt, 2001; Fiske et al., 2002). The reason behind the pattern of our results, however, might be that all the three types of immigrants may have already been perceived as low-status groups (as opposed to Western white expats for instance), that in turn elicited generalized attitudes towards them. The nature of that generalized attitude, whatever it is, may be what matters, that as discussed above, stems from the interplay between multiple factors including one's basic personal values.

Community detection analysis revealed that Inclusive Europeans' value structure seems to be the most *complex*, as opposed to Some, Few, and Exclusive Europeans, whose value structures are more simply divided into two large communities (social focus vs. personal focus). In the case of Inclusive class members however, the 10 basic values configured four separate communities, highly corresponding to the 4 higher-order value types proposed by Schwartz (1992): openness to change, conservation, self-transcendence, and self-enhancement. The only difference was that self-direction, instead of clustering with stimulation and hedonism, clustered with universalism and benevolence. As shown in Figure 1, self-direction value is adjacent to universalism, and after all, it is theoretically assumed to be motivationally close to universalism and benevolence (all belong to growth value types).

One explanation may be that Inclusive class members' more complex value structure (in the sense of the number of communities) enables them to base their judgement on outgroup members through more egalitarian values (e.g., universalism and benevolence), that is less dependent on the causal effects of values tapping into opposing motivational goals (e.g., tradition, conformity). This highly resonates with the prevailing consensus in the literature that cognitive complexity is related both to lower degree of conservatism (see Jost et al., 2009) and more favorable intergroup attitudes (De Zavala et al., 2010; Golec & Federico, 2004). Integrative complexity for example, captures the extent to which one both

differentiates between opposing perspectives and integrates them into a coherent whole (Tetlock, 1983). Inclusive network seems to be the most complex in this sense, as we see that 4 sets of values are differentiated from each other and integrated into four separate sets of values. Regarding the other three networks, however, in a less sophisticated manner, the two large communities consist of motivationally conflicting values. Future investigation is needed to directly test the relationship between the number and the nature of clusters in attitude networks and different forms of cognitive style including cognitive complexity.

Furthermore, the level of assortativity between the two categories of values (growth and self-protection) found for each class of respondents supplemented the results of the community detection analysis. As expected, the assortativity analysis showed that in the case of Inclusive class, values are more strongly interconnected within either growth value types or self-protection value types and the relationship between the two categories is comparatively weaker. In other words, regarding the most inclusive individuals, the values' effects on each other are more confined within the limit of either more egalitarian or less egalitarian types of value. This may imply that in terms of the inclusive individuals, values from dissimilar motivational goals, are less causally dependent upon each other, that in turn enables them to evaluate outgroup members on the basis of one's universal concerns, more free from the restricting influence of ingroup interests. Moreover, the results showed that the global connectivity score (the overall level of interconnections), was found to be mostly invariant across the networks, implying that what makes the difference is the unique configuration of the 10 values for each class.

Prior research has documented that one's intergroup attitudes are strongly reflected in the level of importance they place on holding certain value types (e.g., growth vs. self-protection). Theoretically speaking, the present study contributes to the literature in the field, by showing that people's intergroup attitudes are also expressed in how their personal values tend to be interconnected. Moreover, among several prejudice reduction intervention strategies, cognitive and emotional training as well as moral and value education have been argued to be effective tools on combating intergroup prejudice (for a review see e.g., Paluck et al., 2009; Paluck et al., 2020). Thus, one possible practical implication of the current study may be, for the future research, to investigate if promoting complex thinking in respect of one's personal values would lead to more egalitarian intergroup attitudes. In other words, a prejudice reduction strategy may involve promoting conscious consideration of opposing perspectives with respect to one's personal values and examining whether this in turn results

in a more complex value structure and a more favorable attitude towards the outgroup members. Our study concludes that mapping the dynamic structure of basic human values or moral codes in relation to intergroup attitudes can provide informative conceptual frameworks as well as effective practical strategies to fight intergroup prejudice.

6.6 References

Abacioglu, C. S., Isvoranu, A. M., Verkuyten, M., Thijs, J., & Epskamp, S. (2019). Exploring multicultural classroom dynamics: A network analysis. *Journal of school psychology*, *74*, 90-105.

Adelman, L., & Verkuyten, M. (2020). Prejudice and the acceptance of Muslim minority practices: A person-centered approach. *Social Psychology*, *51*, 1.

Altemeyer, B. (1998). The other ''authoritarian personality.'' In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 47-92). San Diego, CA: Academic Press.

Asparouhov, T., & Muthén, B. (2010). Multiple imputation with Mplus. *MPlus Web Notes*, 238-246.

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using M plus. *Structural equation modeling: a multidisciplinary journal*, *21*, 329-341.

Bauer, D. J., & Shanahan, M. J. (2007). Modeling complex interactions: Personcentered and variable-centered approaches. In T. D. Little, J. A. Bovaird, & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 255–283). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Beierlein, C., Kuntz, A., & Davidov, E. (2016). Universalism, conservation and attitudes toward minority groups. *Social Science Research*, *58*, 68-79.

Bhushan, N., Mohnert, F., Sloot, D., Jans, L., Albers, C., & Steg, L. (2019). Using a Gaussian graphical model to explore relationships between items and variables in environmental psychology research. *Frontiers in psychology*, *10*, 1050.

Bowman, J. W., & West, K. (2021). Brexit: The influence of motivation to respond without prejudice, willingness to disagree, and attitudes to immigration. *British Journal of Social Psychology*, *60*, 222-247.

Canty, A., & Ripley, B. (2017). *boot: Bootstrap R (S-Plus) Functions*. R package version 1.3-20.

Card, D., Dustmann, C., & Preston, I. (2005). Understanding attitudes to immigration: The migration and minority module of the first European Social Survey.

Ceobanu, A. M., & Escandell, X. (2010). Comparative analyses of public attitudes toward immigrants and immigration using multinational survey data: A review of theories and research. *Annual review of sociology*, *36*, 309-328.

Cochrane, C., & Nevitte, N. (2014). Scapegoating: Unemployment, far-right parties and anti-immigrant sentiment. *Comparative European Politics*, *12*, 1-32.

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, *54*, 13-29.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, complex systems*, *1695*, 1-9.

Dalege, J., Borsboom, D., van Harreveld, F., and van der Maas, H. L. (2018). A network perspective on attitude strength: testing the connectivity hypothesis. *Soc. Psychol. Pers. Sci.* 10:1948550618781062.

Dalege, J., Borsboom, D., van Harreveld, F., van den Berg, H., Conner, M., & van der Maas, H. L. (2016). Toward a formalized account of attitudes: The Causal Attitude Network (CAN) model. *Psychological review*, *123*, 2.

Danaj, A., Lazányi, K., & Bilan, Y. (2018). Perceptions and implications of immigration in France-economic, social, political and cultural perspectives. *Economics & Sociology*, *11*, 226-247.

Davidov, E. (2010). Testing for comparability of human values across countries and time with the third round of the European Social Survey. *International Journal of Comparative Sociology*, *51*, 171–191.

Davidov, E., Seddig, D., Gorodzeisky, A., Raijman, R., Schmidt, P., & Semyonov, M. (2020). Direct and indirect predictors of opposition to immigration in Europe: individual values, cultural values, and symbolic threat. *Journal of Ethnic and Migration Studies*, *46*, 553-573.

De Zavala, A. G., Cislak, A., & Wesolowska, E. (2010). Political conservatism, need for cognitive closure, and intergroup hostility. *Political Psychology*, *31*, 521-541.

Dennison, J. (2020). How issue salience explains the rise of the populist right in Western Europe. *International journal of public opinion research*, *32*, 397-420.

Duckitt, J. (2001). A cognitive-motivational theory of ideology and prejudice. *Advances in Experimental Social Psychology*, *33*, 41-113.

Edgell, P. (2017). An agenda for research on American religion in light of the 2016 election. *Sociology of Religion*, *78*, 1-8.

Epskamp, S., Borsboom, D., and Fried, E. I. (2018). Estimating psychological networks and their accuracy: a tutorial paper. *Behav. Res. Methods* 50, 195–212. doi: 10.3758/s13428-017-0862-1

Epskamp, S., Cramer, A. O., Waldorp, L. J., Schmittmann, V. D., and Borsboom, D. (2012). qgraph: network visualizations of relationships in psychometric data. *J. Statist. Softw.* 48, 1–18. doi: 10.18637/jss.v048.i04

Esses, V. M. (2021). Prejudice and Discrimination Toward Immigrants. *Annual Review of Psychology*, *72*, 503-531.

Esses, V. M., Hamilton, L. K., & Gaucher, D. (2019). The role of attitudes in migration. *Handbook of attitudes*, *2*, 455-487.

Farine, D. R. (2014). Measuring phenotypic assortment in animal social networks: Weighted associations are more robust than binary edges. *Animal Behaviour*, *89*, 141–153.

Feldman, G. (2020). Personal values and moral foundations: Examining relations and joint prediction of moral variables. *Social Psychological and Personality Science*, 1948550620933434.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of personality and social psychology*, *82*, 878.

Fried, E. I., van Borkulo, C. D., Cramer, A. O., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: a review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *52*, 1-10.

Fussell, E. (2014). Warmth of the welcome: Attitudes toward immigrants and immigration policy in the United States. *Annual review of sociology*, *40*, 479-498.

Gawronski, B., & Strack, F. (2012). Cognitive consistency as a basic principle of social information processing. In B. Gawronski & F. Strack (eds.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 1-16). New York, NY: Guilford Press.

Golec, A., & Federico, C. M. (2004). Understanding responses to political conflict: Interactive effects of the need for closure and salient conflict schemas. *Journal of personality and social psychology*, *87*, 750.

Gorodzeisky, A., & Semyonov, M. (2020). Perceptions and misperceptions: actual size, perceived size and opposition to immigration in European societies. *Journal of Ethnic and Migration Studies*, *46*, 612-630.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of personality and social psychology*, *96*, 1029.

Green, E. G., Visintin, E. P., Sarrasin, O., & Hewstone, M. (2020). When integration policies shape the impact of intergroup contact on threat perceptions: a multilevel study across 20 European countries. *Journal of Ethnic and Migration Studies*, *46*, 631-648.

Grigoryev, D., Fiske, S. T., and Batkhina, A. (2019). Mapping ethnic stereotypes and their antecedents in Russia: the stereotype content model. *Front. Psychol.* 10:1643. doi: 10.3389/fpsyg.2019.01643

Gusciute, E., Mühlau, P., & Layte, R. (2020). All Welcome Here? Attitudes towards Muslim Migrants in Europe. *International Migration*.

Hadarics, M., & Kende, A. (2018). Moral foundations of positive and negative intergroup behavior: Moral exclusion fills the gap. *International Journal of Intercultural Relations*, *64*, 67-76.

Hartman, T. K., Stocks, T. V., McKay, R., Gibson-Miller, J., Levita, L., Martinez, A. P., ... & Bentall, R. P. (2020). The authoritarian dynamic during the COVID-19 pandemic: effects on nationalism and anti-immigrant sentiment. *Social Psychological and Personality Science*, 1948550620978023.

Heider, F. (1958). *The psychology of cognitive organization*. Hoboken, NJ: Wiley.

Jost, J. T., Federico, C. M., & Napier, J. L. (2009). Political ideology: Its structure, functions, and elective affinities. *Annual review of psychology*, *60*, 307-337.

Kugler, M., Jost, J. T., & Noorbaloochi, S. (2014). Another look at moral foundations theory: Do authoritarianism and social dominance orientation explain liberal-conservative differences in "moral" intuitions?. *Social Justice Research*, *27*, 413-431.

Landmann, H., Gaschler, R., & Rohmann, A. (2019). What is threatening about refugees? Identifying different types of threat and their association with emotional responses and attitudes towards refugee migration. *European Journal of Social Psychology*, *49*, 1401-1420.

Lee, T. L., & Fiske, S. T. (2006). Not an outgroup, not yet an ingroup: Immigrants in the stereotype content model. *International Journal of Intercultural Relations*, *30*, 751-768.

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, *88*, 767-778.

Löw, A., Puzić, S., & Matić Bojić, J. (2020). Anti-immigrant prejudice in a post-socialist context: the role of identity-based explanations. *Ethnic and racial studies*, 1-20.

Lubbers, M., & Coenders, M. (2017). Nationalistic attitudes and voting for the radical right in Europe. *European Union Politics*, *18*, 98-118.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*, 19–40. doi:10.1037/1082-989X.7.1.19

Magun, V., Rudnev, M., & Schmidt, P. (2015). Within-and between-country value diversity in Europe: A typological approach. *European Sociological Review*, *32*, 189-202.

Máté, D., Sarihasan, I., Popp, J., & Oláh, J. (2018). The role of regional differences in immigration: the case of OECD countries. *Economics & Sociology*, *11*, 190-206.

Matthes, J., & Schmuck, D. (2017). The effects of anti-immigrant right-wing populist ads on implicit and explicit attitudes: A moderated mediation model. *Communication Research*, *44*, 556-581.

McFarland, S. (2010). Authoritarianism, social dominance, and other roots of generalized prejudice. *Political Psychology*, 31, 453-477.

Mitchell, J. (2021). Social trust and anti-immigrant attitudes in Europe: a longitudinal multi-level analysis. *Frontiers in sociology*, *6*, 73.

Morselli, D., & Passini, S. (2018). Exclusive and inclusive protest in Europe: Investigating values, support for democracy, and life conditions. *Journal of Community & Applied Social Psychology*, *28*, 123-141.

Muthén, L., & Muthén, B. (2017). Mplus (Version 8)[computer software].(1998–2017). *Los Angeles, CA: Muthén & Muthén*.

Nariman, H. S., Hadarics, M., Kende, A., Lášticová, B., Poslon, X. D., Popper, M., ... & Minescu, A. (2020). Anti-Roma bias (stereotypes, prejudice, behavioral tendencies): A network approach toward attitude strength. *Frontiers in psychology*, *11*.

Newman, M. E. (2003). Mixing patterns in networks. *Physical review E*, *67*, 026126.

Norris, P., & Inglehart, R. (2019). *Cultural backlash: Trump, Brexit, and authoritarian populism*. Cambridge University Press.

Nylund-Gibson, K., & Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science,4*, 440–461.

Osborne, D., & Sibley, C. G. (2017). Identifying "types" of ideologies and intergroup biases: Advancing a person-centred approach to social psychology. *European Review of Social Psychology*, *28*, 288-332.

Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *Annual review of psychology*, *60*, 339-367.

Paluck, E. L., Porat, R., Clark, C. S., & Green, D. P. (2020). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, *72*.

Pehrson, S., & Green, E. G. (2010). Who we are and who can join us: National identity content and entry criteria for new immigrants. *Journal of Social Issues*, *66*, 695-716.

Pellegrini, V., De Cristofaro, V., Salvati, M., Giacomantonio, M., & Leone, L. (2021). Social Exclusion and Anti-Immigration Attitudes in Europe: The mediating role of Interpersonal Trust. *Social Indicators Research*, *155*, 697-724.

Phua, D. Y., Leong, C. H., & Hong, Y. (2020). Heterogeneity in national identity construct: Example of Singapore using network analysis. *International Journal of Intercultural Relations*, *78*, 20-32.

Ramaswamy, V., DeSarbo, W. S., Reibstein, D. J., & Robinson, W. T. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing science*, *12*, 103-124.

Roccato, M., & Ricolfi, L. (2005). On the correlation between right-wing authoritarianism and social dominance orientation. *Basic and applied social psychology*, *27*, 187-200.

Rokeach, M. (1973). The nature of human values (Vol. 438). Free press New York.

Roy, S., Neumann, C. S., Jones, D. N., Gari, A., & Šram, Z. (2021). Psychopathic propensities contribute to social dominance orientation and right-wing authoritarianism in predicting prejudicial attitudes in a large European sample. *Personality and Individual Differences*, *168*, 110355.

Sayans-Jiménez, P., van Harreveld, F., Dalege, J., & Rojas Tejada, A. J. (2019). Investigating stereotype structure with empirical network models. *European Journal of Social Psychology*, *49*, 604-621.

Schaffner, B. F., MacWilliams, M., & Nteta, T. (2018). Understanding white polarization in the 2016 vote for president: The sobering role of racism and sexism. *Political Science Quarterly*, *133*, 9-34.

Schemer, C. (2012). The influence of news media on stereotypic attitudes toward immigrants in a political campaign. *Journal of Communication*, *62*, 739-757.

Schlueter, E., Masso, A., & Davidov, E. (2020). What factors explain anti-Muslim prejudice? An assessment of the effects of Muslim population size, institutional characteristics and immigration-related media claims. *Journal of Ethnic and Migration Studies*, *46*, 649-664.

Schmittmann, V. D., Cramer, A. O., Waldorp, L. J., Epskamp, S., Kievit, R. A., and Borsboom, D. (2013). Deconstructing the construct: a network perspective on psychological phenomena. *New Ideas Psychol.* 31, 43–53. doi: 10.1016/j.newideapsych.2011.02.007

Schmuck, D., & Matthes, J. (2015). How anti-immigrant right-wing populist advertisements affect young voters: Symbolic threats, economic threats and the moderating role of education. *Journal of Ethnic and Migration Studies*, *41*, 1577-1599.

Schmuck, D., & Matthes, J. (2017). Effects of economic and symbolic threat appeals in right-wing populist advertising on anti-immigrant attitudes: The impact of textual and visual appeals. *Political Communication*, *34*, 607-626.

Schwartz, S. (2006). A theory of cultural value orientations: Explication and applications. *Comparative sociology*, *5*, 137-182.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in experimental social psychology*, *25*, 1-65.

Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values?. *Journal of social issues*, *50*, 19-45.

Schwartz, S. H. (2006). Les valeurs de base de la personne: théorie, mesures et applications. *Revue française de sociologie*, *47*, 929-968.

Schwartz, S. H. (2007). Value orientations: Measurement, antecedents and consequences across nations. *Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey*, 161–193.

Schwartz, S. H., Caprara, G. V., & Vecchione, M. (2010). Basic personal values, core political values, and voting: A longitudinal analysis. *Political Psychology*, *31*, 421–452.

Sibley, C. G., Bergh, R., Satherley, N., Osborne, D., Milojev, P., Greaves, L. M., ... & Duckitt, J. (2019). Profiling authoritarian leaders and followers. *TPM-Testing, Psychometrics, Methodology in Applied Psychology*, *26*, 401-417.

Talay, L., & De Coninck, D. (2020). Exploring the link between personality traits and European attitudes towards refugees. *International Journal of Intercultural Relations*, *77*, 13-24.

Taniguchi, H. (2021). National identity, cosmopolitanism, and attitudes toward immigrants. *International Sociology*, 0268580921994517.

Tartakovsky, E., & Walsh, S. D. (2020). Are some immigrants more equal than others? Applying a threat-benefit model to understanding the appraisal of different immigrant groups by the local population. *Journal of Ethnic and Migration Studies*, *46*, 3955-3973.

Terrizzi Jr, J. A., Shook, N. J., & Ventis, W. L. (2010). Disgust: A predictor of social conservatism and prejudicial attitudes toward homosexuals. *Personality and Individual Differences*, 49, 587-592.

Tetlock, P. E. (1983). Cognitive style and political ideology. *Journal of Personality and Social psychology*, *45*, 118.

Turner-Zwinkels, F. M., Johnson, B. B., Sibley, C. G., & Brandt, M. J. (2021). Conservatives' Moral Foundations Are More Densely Connected Than Liberals' Moral Foundations. *Personality and Social Psychology Bulletin*, *47*, 167–184.

Vachudova, M. A. (2021). Populism, Democracy, and Party System Change in Europe. *Annual Review of Political Science*, *24*.

van Borkulo, C. D., Boschloo, L., Kossakowski, J. J., Tio, P., Schoevers, R. A., Borsboom, D., et al. (2017). *Network Comparison Test—Comparing Network Structures on Three Aspects: A Permutation Test*. Available online at: https://github.com/cvborkulo/NetworkComparisonTest (accessed January 11, 2021).

Van de Vyver, J., Houston, D. M., Abrams, D., & Vasiljevic, M. (2016). Boosting belligerence: How the July 7, 2005, London bombings affected liberals' moral foundations and prejudice. *Psychological science*, *27*, 169-177.

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political analysis*, 450-469.

Verschoor, M., Albers, C., Poortinga, W., Böhm, G., & Steg, L. (2020). Exploring relationships between climate change beliefs and energy preferences: A network analysis of the European Social Survey. *Journal of Environmental Psychology*, *70*, 101435.

Walsh, S. D., & Tartakovsky, E. (2021). Personal Value Preferences, Threat-Benefit Appraisal of Immigrants and Levels of Social Contact: Looking Through the Lens of the Stereotype Content Model. *Frontiers in psychology*, *12*, 477.

Weighting European Social Survey Data, available from: <https://www.europeansocialsurvey.org/methodology/ess_methodology/data_processing_arc hiving/weighting.html> [accessed 03 March 2021].

Wirz, D. S., Wettstein, M., Schulz, A., Müller, P., Schemer, C., Ernst, N., ... & Wirth, W. (2018). The effects of right-wing populist communication on emotions and cognitions toward immigrants. *The International Journal of Press/Politics*, *23*, 496-516.

Wolf, L. J., Weinstein, N., & Maio, G. R. (2019). Anti-immigrant prejudice: Understanding the roles of (perceived) values and value dissimilarity. *Journal of personality and social psychology*, *117*, 925.

7 Study IV: Immigrants' intragroup moral exclusion predicts ingroup-directed behavioral intentions: The mediating role of disidentification

7.1 Abstract

Previous research documents the relationship between negative treatments of outgroup members and moral exclusion. In this study, we expect negative treatments of the ingroup members to be also related to a moral exclusionary mechanism. Next, we hypothesize that the relationship between *intragroup* moral exclusion and behavioral intentions (both positive and negative ones) targeting the ingroup members, to be mediated by immigrants' disidentification with their ethnic identity. Using two samples of Iranians (n=385) and Tunisians (n=124) living in Italy, we test the two hypotheses. Results provided evidence that moral exclusion is negatively associated with positive behavioral intentions (active facilitation) and positively associated with the negative behavioral intentions (passive facilitation and passive harm). Results also suggest that this relationship is mediated by disidentification with immigrants' own ethnic group.

7.2 Introduction

Prior research has investigated different cognitive, affective, and behavioral antecedents and consequences of the majority members' negative attitudes towards immigrants (Esses, 2021). We also know that hostility against outgroup members, including immigrants, can involve moral exclusionary mechanisms (Passini & Morselli, 2016). What has not yet been established is whether those processes of moral exclusion are applied to the ingroup members, or more particularly, by immigrants against their own ingroup. We attempt to fill this gap by investigating immigrants' negative and positive ingroup-directed behavioral intentions in connection with moral exclusion. Furthermore, we investigate if this relationship is explained by immigrants' disidentification with their ethnic background.

7.2.1 Moral exclusion

Negative behaviors directed against a target can be regarded as permissible and may not cost a negative self-image to the harm-doer if the target does not fall within the purview of one's ideas of justice and fairness. This psychological mechanism occurs *"when individuals or groups are perceived as outside the boundary in which moral values, rules, and considerations of fairness apply"* (Opotow, 1990, p.1). Olson et al. (2011) identify two distinct moral exclusionary processes. One taps into the irrelevance of the target to one's basic moral concerns, so the target is placed outside the scope of justice. That is, the target (an insect for instance) is not even entitled to be treated in accordance with basic moral rules. The other is based on the target's deservingness for negative treatments. Negative treatment

of the target is not viewed as morally problematic but justified and fair because the target is perceived to be deserving of it.

How could then either of the above-mentioned processes be applied to one's own ingroup? We argue that moral exclusion can also be directed towards the ingroup members and show that it is strongly associated with behavioral intentions through the mediating effect of disidentification.

7.2.2 Disidentification

Social identification is a psychological process whereby the ingroup becomes an expression of oneself where one's socio-psychological needs, such as belonging, self-efficacy, and self-esteem are satisfied (Tajfel & Turner, 1979; see also Breakwell 1986). Some group memberships, however, do not satisfy these needs and are experienced as painful and threatening (Becker & Tausch, 2014; Breakwell, 1986). A major social change (such as migration) that compromises positive functioning of identification can lead to perceiving the corresponding identity (ethnic background) as a threat (Timotijevic & Breakwell, 2000).

Becker and Tausch (2014) define disidentification as "*a psychological phenomenon that occurs when individuals belong to groups they do not wish to belong to* (p.4)". It occurs when the ingroup is not possible to physically escape from, at the same time it is perceived as a threat to the self. Through disidentification, the individual creates a psychological distance from the negative group membership. They distinguish between three dimensions of disidentification: detachment (*which ranges from feelings of alienation to an active separation from one's ingroup*), dissatisfaction (*expressed in feelings of dissatisfaction with one's group membership*), and dissimilarity (*the perception that one is dissimilar to other group members*) (Becker & Tausch, 2014. p. 6).

7.2.3 Intragroup moral exclusion and disidentification

Previous research has investigated the relationship between behavioral intentions and moral exclusion mostly with respect to outgroup members (Hadarics & Kende 2018; Nariman et al., 2020; Passini & Morselli, 2016). A tendency for negative behavioral intentions directed towards other ingroup members has been previously documented and it has been established that such tendency is influenced by how one copes with the negative group membership (Becker & Tausch 2014). As mentioned above, an unpleasant identity may be perceived as threatening to the self that is subsequently linked up to disidentification with the ingroup. This highly resonates with the literature proposing perceived threat as one major factor associated with moral exclusion of the "other" (Hadarics, 2020; Olson et al., 2011; Opotow,

1990). One may therefore argue that that other can also be members of the ingroup, if the very ingroup identity threatens the self. Regarding intergroup moral exclusion what is threatened is the ingroup identity (Hadarics et al., 2020) while in disidentification and our proposed intragroup moral exclusion the ingroup identity itself is perceived as a threat. In the former the threat is the outgroup while in the latter the threatened is the self. However, what is threatened in both cases is the self—the actor's personal identity—and what is morally excluded and negatively treated is the target perceived to be imposing that threat. It is worth mentioning that the position we adopt here avoids a sharp distinction between personal and social identity, since social identity is part of one's personal identity—or as Breakwell (2001) notes: "seen across the biography, social identity is seen to become personal identity: the dichotomy is purely a temporal artefact" (p.277).

Based on prior research, considerations of fairness in general (Graham et al., 2013), and moral exclusion in particular (Passini & Morselli, 2016), are intuitive and subtle psychological mechanisms. Disidentification has been suggested to be an *active/cognitive* psychological mechanism through which the subject actively distances themselves from the unpleasant identity impossible to escape from physically (Becker & Tausch, 2014). Thus, the relationship between intuitive considerations of fairness regarding the ingroup members and ingroup-directed behavioral intentions should be explained through a more active mechanism of disidentification.

7.3 The current study

We propose two hypotheses. First, intragroup moral exclusion is related to negative ingroup-directed behavioral intentions (H1). Second, disidentification mediates the association between intragroup moral exclusion and behavioral intentions directed against the ingroup (H2).

7.3.1 Procedure

The present study was approved by the ethical committee of the *University of author*. The information about the items used in the study can be found in the supplementary materials. All the scales were translated (into Farsi for Iranians and Tunisian Arabic for Tunisians), and back-translated by independent translators and pilot tested.

7.3.2 Participants

Data collection was conducted through an online survey from June to August 2020, among Iranians living in Italy. From August to December 2020 a replication sample was collected among Tunisians based also in Italy. A total of 385 Iranians and 124 Tunisians participated in

the online survey[2]. Demographic information of the respondents are presented in Table 1. The Iranian sample was collected by the help of the organization "Unione Degli Studenti Iraniani" (Union of Iranian Students) through a social media group created for Iranian students living in Italy. Tunisian respondents were also recruited using a social media platform dedicated to Tunisian immigrants living in Italy as well as through the help of the organization "movimento dei tunisini in italia" (Tunisian Movement in Italy). To ensure that the respondents of both samples are in Italy and not based in other countries, we explicitly asked them to answer the question at the beginning of the survey.

Table 1

Demographic Information of the Iranian and Tunisian Participants

Sample	Women	Age Range			Education		
		18-35	36-45	46+	Primary	Secondary	University
	n	%	%	%	%	%	%
Iran	385	46.8	90	9.7	.3	1.5	6.2	92.4
Tunisia	124	33.3	51	39.2	9.8	6.9	33.3	59.8

7.3.3 Measures

Disidentification was measured by 11 items adapted from Becker and Tausch's (2014) Disidentification Scale . Four items measured detachment (e.g., "I feel disloyal to other Iranians/Tunisians"), three items measured dissatisfaction (e.g., "I'm unhappy about being an Iranian/Tunisian"), and four items measured dissimilarity (e.g., "I have nothing in common with most Iranians/Tunisians"). The items were measured on a 7-point scale from *strongly disagree* (1) to *strongly agree* (7).

---

[2] We ran a two-step post hoc Monte Carlo power simulation (Muthén & Muthén 2017), with 1000 replications to examine if our sample sizes yielded enough power. The results showed that we have at least 80% power for the main estimates in both studies.

Intragroup moral exclusion (hereafter, IME[3]) was measured by three 7-point bipolar items adopted from Moral Inclusion/exclusion of Other Groups (MIEG) scale (Passini & Morselli, 2016). MIEG scale was originally designed to capture a form of subtle/implicit intergroup moral exclusion. The items were rephrased in order to measure an *intragroup* mechanism of moral exclusion where +3 on the left indicates moral exclusion and +3 on the right indicates moral inclusion ("Most Iranians/Tunisians deserve no respect - Most Iranians/Tunisians deserve utmost respect"; "It is necessary for the people of other societies to avoid any kind of contact with most Iranians/Tunisians - It is necessary for the people of other societies to engage in establishing constructive contacts with Iranians/Tunisians"; "I think most Iranians/Tunisians are extremely uncivilized - I think that most Iranians/Tunisians are extremely civilized").

To measure behavioral intentions, Behaviors from Intergroup Affect and Stereotypes (BIAS) map was used (Cuddy et al., 2007). BIAS map suggests four types of intergroup behaviors based on two dimensions of intensity (active vs. passive) and valence (facilitation vs. harm). The four types of behavior, therefore, are: active harm (e.g., harassing), passive harm (e.g., neglecting), active facilitation (e.g., helping), and passive facilitation (e.g., associating). Among the four only active facilitation is considered springing from a positive intention. Beneficial outcomes of passive facilitation are only by-products of the actor's intention to "work-with" the outgroup members to reach a specific goal. However, the actor does not intend to build a positive relationship with the outgroup member. In the current study, we rephrased the items to assess the participants' behavioral intentions towards their own group members instead of the outgroup. Moreover, the items tapping into active harm were not asked because they received negative reactions from the respondents during the pilot data collection. Active facilitation was measured by four items (e.g., "I search contact with other Iranians/Tunisians"). Passive facilitation was measured by two items ("I only search for contact with other Iranians/Tunisians if this fulfils a specific goal"; "I only search for contact with other Iranians when it serves my own interest"). Passive harm was measured by four items (e.g., "I do not pay attention to other Iranians/Tunisians"). The items were measured by a 7-point scale from extremely unlikely (1) to extremely likely (7).

---

[3] To show that moral exclusion and disindetification form two different factors, we ran a Confirmatory Factor Analysis showing that fit indices for the two-factor solution are significantly better than the one-factor solution of both samples (see Table S in the supplementals).

7.4 Results

Before proceeding to the main analyses, we first checked for the internal consistencies of all the scales and subscales used in the models. All Cronbach's alphas were satisfactory: Active facilitation (α_{Iran} = .84; $\alpha_{Tunisia}$ = .74), passive facilitation (α_{Iran} = .92; $\alpha_{Tunisia}$ = .89), passive harm (α_{Iran} = .87; $\alpha_{Tunisia}$ = .80), disidentification (α_{Iran} = .90; $\alpha_{Tunisia}$ = .91), detachment (α_{Iran} = .91; $\alpha_{Tunisia}$ = .81), dissatisfaction (α_{Iran} = .90; $\alpha_{Tunisia}$ = .93), dissimilarity (α_{Iran} = .86; $\alpha_{Tunisia}$ = .89), and intragroup moral exclusion (α_{Iran} = .81; $\alpha_{Tunisia}$ = .78).

Two structural equation models were built for each sample. First a direct model was built in which the three forms of behavioral intentions were regressed on IME (H1). Next, a mediation model was built where disidentification was added to the previous model as the mediator (H2). Regarding the Tunisian sample, during model building, we noticed that passive facilitation did not yield significant explained variance. We therefore removed the two passive facilitation items to improve the model fits. In the same manner, to improve the model fits, the analysis of the Iranian sample did not include an item measuring active facilitation and one measuring disidentification (detachment subscale). The models were bootstrapped with 1000 draws to obtain 95% confidence intervals. Demographic characteristics were controlled for in all the four models by regressing the endogenous variables on participants' age, gender, and level of education. Further, disidentification was measured as a second order factor representing its three dimensions: detachment, dissatisfaction, and dissimilarity. Table 2 shows the model fit information for all the four models.

Table 2

Model Fit Statistics for the Structural Equation Models

Models	X^2(*df*)	RMSEA	CFI	TLI	SRMR
Iran					
Direct	187.88 (71)	.065	.953	.932	.036
Mediation	483.12 (244)	.050	.954	.944	.046
Tunisia					
Direct	75.60 (64)	.038	.973	.963	.053

Mediation	322.81 (253)	.047	.950	.941	.062

In the direct models (Figure 1), IME was related to active facilitation and passive harm. Passive facilitation was associated with IME in the Iranian sample. More importantly, the results of the mediation models (Figure 2) showed that IME and disidentification were strongly interrelated. Disidentification in turn was associated with all the dependent variables in both samples. As presented in Table 3, all the indirect paths were statistically significant as well.
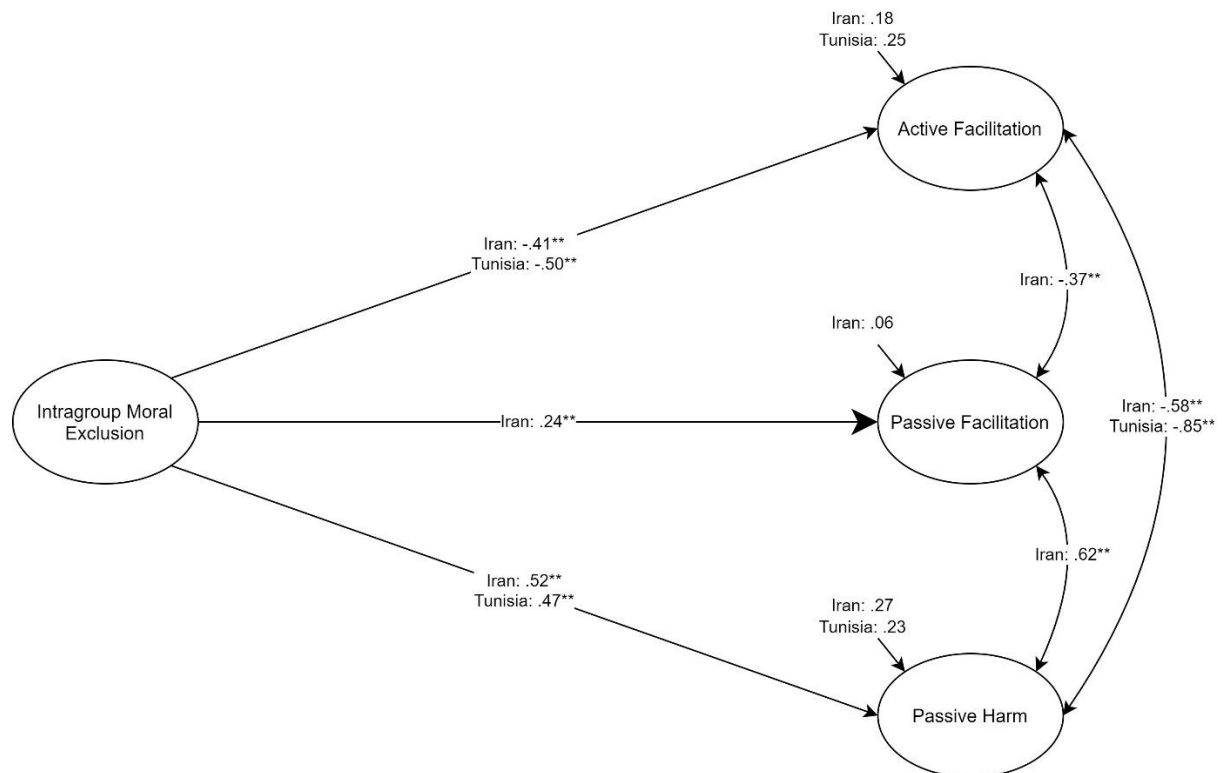

Figure 1. Path model testing the direct effect of IME on ingroup-directed behavioral intentions. Path coefficients are standardized (** = p < .01; * = p < .05.).
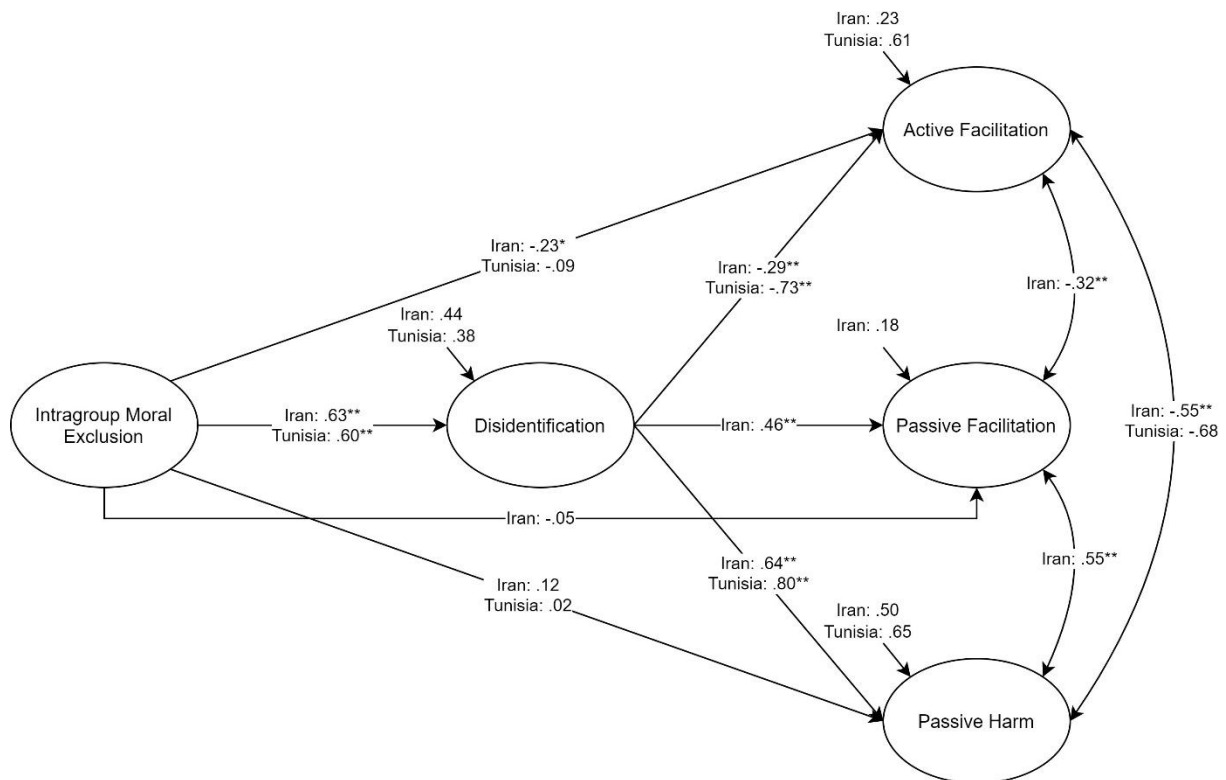
Figure 2. Path model testing the mediating effect of disidentification on the relationship between IME and ingroup-directed behavioral intentions. Path coefficients are standardized (** = $p <$ .01; * = $p <$ .05.).

Table 3

Indirect Effects of Ingroup-based Moral Exclusion on Ingroup-directed Behavioural Intentions through Disidentification

Iran				
Indirect Pathway	Indirect Effect	*p*	LLCI	ULCI
IME → DIS → AF	−.18	.009	−.33	−.05
IME → DIS → PF	.29	< .001	.16	.45
IME → DIS → PH	.40	< .001	.26	.57
Tunisia				
Indirect Pathway	Indirect Effect	*p*	LLCI	ULCI
IME → DIS → AF	−.44	< .001	−.79	−.24

IME → DIS → PH	.48	< .001	.28	.80

Note. IME = Intragroup Moral Exclusion; DIS = Disidentification; AF = Active Facilitation; PF =Passive Facilitation; PH = Passive Harm. The coefficients are standardized.

7.5 Discussion

The present study sought to empirically test whether moral exclusion holds conceptual relevance when behavior is targeted at fellow group members. We first assumed that IME is linked to the behavioral intentions, and second, disidentification mediates this nexus. The results generally supported both of our hypotheses. First, our findings suggest that not only mistreatments against outgroup members are subject to moral exclusionary processes, but this also applies when the negative behavior is directed against members of the ingroup. Next, we found that immigrants' disidentification with their ethnic background mediates the relationship between IME and behavioral intentions against the ingroup members.

The only difference between the two samples was that passive facilitation did not function among the Tunisian respondents. The reason may be due to the difference in the demographic characteristics between the two samples. Although we did not directly ask the respondents about their residential status, since the Iranian sample was collected through a student union, we speculate most of them to be students. While, informally, during the data collection among Tunisians we learnt that the Tunisian sample is more heterogenous including students, workers, and asylum seekers. Passive facilitation, as discussed above is a tendency to contact when the actor has a specific goal. Many Tunisians may have already been forced to form closer bonds (in refugee camps for example), with little opportunity to choose whether to contact ingroup members. For Iranian respondents, as international students, contacting other Iranians only occasionally and if necessary, should be more conceivable.

Coping mechanisms against a negative social identity include social creativity, social competition, and individual mobility (Tajfel & Turner, 1979). Disidentification is a form of individual mobility where the individuals psychologically distance themselves from the negative identity. Our study contributes to the literature in the field by suggesting that this form of coping mechanism can be expressed in one's considerations of fairness with regards to the ingroup members as well.

Psychological distance is proposed as a vital factor associating with moral exclusion (Olson et al., 2011; Opotow, 1990). In psychological distance, the emphasis is often placed on a lack of identification with the target. On the contrary, disidentification does not refer to a complete absence of identification but the presence of a threatening one. Thus, one theoretical contribution of our study may be that moral exclusion not only relates to non-identification but also to a similar though qualitatively different psychological process namely disidentification. Moral exclusion of the ingroup, hence, may also be qualitatively distinct from other forms of moral exclusionary processes. This can be because although the ingroup members are mistreated, they cannot be completely placed outside one's moral boundary because they are not completely placed outside the boundary of identification. The behavioral expressions of the moral exclusion of the ingroup, therefore, should as well be distinguished from those of intergroup moral exclusion. In the former, one may argue that the behavioral manifestations of moral exclusion should take milder, more subtle, and more passive forms. Further investigations could scrutinize the differences in the nature and intensity of behavioral intentions directed against the ingroup in comparison with those inflicted upon the outgroup.

7.5.1 Future Directions

Our argument about disidentification and IME was based on perceived threat as a major factor in both constructs. Future research could thus directly test whether the relationship between disidentification and IME is explained by threat perception (identity threat). In addition, the nature of that perceived threat can be also further investigated. One possibility is that the ingroup's perceived low status in the host society serves a cause for identity threat perception. On the other hand, perceived negative behaviors of the ingroup members towards the host society can lead to a perception of identity threat. As regards the latter, and as discussed earlier, deservingness has been argued to be related to moral exclusion. Future research could investigate if beyond the effect of disidentification, deservingness also explains the relationship between intragroup moral exclusion and behavioral intentions against the ingroup.

Future investigations can also include system justification (Jost et al., 2004) as a potential factor linked to considerations of fairness regarding one's ingroup members. In other words, one societal consequence of immigrants' disidentification with their ethnic background (and/or its further moral expressions) can be the acceptance of the host society's negative treatments against ingroup members.

Finally, it is also worth noting that we consider the samples in this study not as specific groups but rather as groups in a specific situation. Bearing this in mind, the generalizability of our findings depends on future studies investigating similar mechanisms among different social groups. In this regard, the role of acculturation (in the case of migration) could also be investigated to trace the interplay between different acculturation styles and disidentification.

7.5.2 Limitations

One limitation of this study is its correlational nature, so that future research is needed to examine the causal order between the variables. Further, for measuring IME and the behavioral intentions, we used validated scales, with satisfactory Cronbach's alpha (see the supplementary materials), yet originally designed to capture intergroup attitudes. However, still this may be considered a limitation of our paper. Accounting for the subtle differences between IME and other forms of moral exclusion as well as between intergroup and intragroup behavioral tendencies requires the construction and validation of new scales.

7.6 Conclusion

Finally, in this study the effects of participants' demographic characteristics were *statistically* controlled for. But as a fundamental caveat, such practice, should by no means, preclude one's understanding on the interplay between different forms of social stratification and modes of identification in determining socio-political perceptions and acts and the corresponding moral pronouncements (for an inquiry on this interplay see Sayer, 2005). Thus, future empirical studies could scrutinize this interplay, since different social groups may have different reactions towards social stigma on the basis of social identifications conditioned by concrete social positions (gender, class, ethinicity, education, etc.).

7.7 References

Becker, J. C., & Tausch, N. (2014). When group memberships are negative: The concept, measurement, and behavioral implications of psychological disidentification. *Self and Identity*, *13*, 294-321.

Breakwell GM (2001) Social representational constraints upon identity processes. In: Deaux K and Philoge`ne G (eds) *Representations of the Social: Bridging Theoretical Traditions*. Oxford: Blackwell, 271–284.

Breakwell, G. M. (1986). *Coping with threatened identities*. London: Routledge.

Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: behaviors from intergroup affect and stereotypes. *Journal of personality and social psychology*, *92*, 631.

Esses, V. M. (2021). Prejudice and discrimination toward immigrants. *Annual Review of Psychology*, *72*, 503-531.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. In *Advances in experimental social psychology* (Vol. 47, pp. 55-130). Academic Press.

Hadarics, M. (2020). Perceived outgroup characteristics as antecedents and consequences of moral exclusion. *The Social Science Journal*, 1-10.

Hadarics, M., Szabó, Z. P., & Kende, A. (2020). The relationship between collective narcissism and group-based moral exclusion: The mediating role of intergroup threat and social distance. *Journal of Social and Political Psychology*, *8*, 788-804.

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political psychology*, *25*, 881-919.

Muthén, L. K., & Muthén, B. O. (1998). 1998-2017. *Mplus User's Guide. Eighth Edition, Los Angeles, CA: Muthén & Muthén*.

Nariman, H. S., Hadarics, M., Soufizadeh, A. M., & Kende, A. (2020). The mediating role of moral exclusion between authoritarianism and outgroup discrimination. *International Journal of Intercultural Relations*, *74*, 1-6.

Olson, J. M., Cheung, I., Conway, P., Hutchison, J., & Hafer, C. L. (2011). Distinguishing two meanings of moral exclusion: Exclusion from moral principles or principled harm-doing?. *Social Justice Research*, *24*, 365-390.

Opotow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues*, 46, 1–20.

Passini, S., & Morselli, D. (2016). Construction and validation of the moral inclusion/exclusion of other groups (MIEG) scale. *Social Indicators Research*, *134*, 1195-1213.

Sayer, R. A. (2005). *The moral significance of class*. Cambridge University Press.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin, & S. Worchel (Eds.), *The social psychology of intergroup relations*. (pp. 33-48). Monterey, CA: Brooks/Cole.

Timotijevic, L., & Breakwell, G. M. (2000). Migration and threat to identity. *Journal of Community & Applied Social Psychology*, *10* 355-372.

8   General Discussion

This dissertation encompasses four distinct yet interrelated quantitative studies which together provides an examination of various facets of group-based attitudes, morality, and group-based behavioral tendencies. The term "morality" and its derivatives serve as an overarching concept to investigate, identify, and explore value-laden intuitive judgments, such as moral exclusionary inclinations in Studies 1 and 4, perceived threats to national identity in Study 2, and the ten basic values in Study 3.

The central theme that emerges from the four research studies is the crucial role of group-based value-laden intuitive evaluations and their complex interaction with group-based behavioral tendencies. These evaluations play a pivotal part in their interplay with the way individuals perceive and behave towards both outgroup and ingroup members. Study 1 and Study 4 delve into the importance of moral exclusionary evaluations in driving behavioral tendencies against the outgroup and ingroup members, respectively. Study 3 investigates the intricate relationship between the 10 basic human values and their influence on behavioral tendencies towards outgroup members. The findings imply that the nature of these values and the complex dynamic between them is significantly related to how individuals tend to treat the outgroup. Lastly, the results of Study 2 confirm that the affective elements of the Roma's social perception, perceived threat to national identity, exhibit significant causal links with the corresponding behavioral tendencies. Overall, the studies underscore the vital interdependence between group-based moral intuitions and the behavioral tendencies towards both ingroup and outgroup members. This suggests that interventions should prioritize addressing the affective aspects of group-based perceptions to effectively reduce group-based biases.

Study 1 highlights the significant role of moral exclusion in regulating negative intergroup behaviors, especially when combined with as right-wing authoritarian and social dominance orientation tendencies. While previous research has identified positive associations between SDO, RWA, and moral exclusion (e.g., Passini & Morsell, 2016), our study goes beyond by investigating their combined influence on behavioral tendencies against Roma and Jewish minorities in Hungary. Moreover, our findings indicate that both RWA and SDO are applicable to both communities, as the hypothesized model worked for both the Roma and Jewish minorities. Considering that Roma people are generally viewed as low status and the Jewish minority as high status—with the former perceived as *derogated* and the latter as *dangerous*—theoretically, SDO should predict prejudice against the Jewish

minority , and RWA against the Roma. Our representative sample, however, shows that both of these constructs are relevant to both communities within the current Hungarian context, and discriminatory intentions against these two groups can be due to a combined effect of both RWA and SDO. For interventionists, it implies a need to focus on mitigating these underlying moral motivations to diminish prejudice against minority group members.

Study 2's network analysis approach offers a novel perspective on anti-Roma bias in different European countries. The identification of threat perception to national identity as a central variable suggests that interventions aimed at reducing this perceived threat could have a cascading effect on related anti-Roma evaluations. Furthermore, recognizing the role of empathy as a central variable highlights the importance of empathy-building interventions that could potentially reshape negative attitudes toward Roma populations.

Study 3's categorization of Europeans into distinct groups based on their attitudes towards immigrants offers insights for interventions. By understanding the more complex dynamics of values among the most inclusive Europeans, interventions could target the alignment of different sets of values and emphasize promoting mindful examination of opposing motivations within one's personal value system to foster less discriminatory intentions. Previous studies have shown that the importance individuals assign to specific values (like openness to change, self-transcendence, self-enhancement, and conservation) reveals their attitudes towards different groups. From a theoretical perspective, we contributed to the literature in the field by showing that attitudes towards the outgroup also present themselves in the *interconnections* between different sets of values and the way individuals differentiate between values with differing or even opposing motivational forces. Moreover, in the discussion section, I highlighted a *prevailing consensus* suggesting a positive correlation between cognitive rigidity and conservatism. However, this observation requires a more detailed examination. A recent meta-analysis (Costello et al., 2022) shows that not all forms of conservatism are associated with cognitive rigidity. Specifically, while social conservatism is robustly related to cognitive rigidity, this is not the case for economic conservatism, which has a negligible relationship with cognitive rigidity. Our findings align with this perspective as we measured what they refer to as "rigidity-of-the-social-right," contrasting it with the rigidity-of-the-right hypothesis, which indiscriminately lumps together various kinds of conservatism related to cognitive rigidity.

Study 4 contributes to the field by revealing the possibility that moral exclusion can be directed towards one's own ingroup. We showed that the effect of moral exclusion on

behavioral intentions directed towards the ingroup is influenced by immigrants disidentifying from their ethnic identity. Interventions could target disidentification and focus on its affective/moral dimensions to reduce negative intragroup sentiments among immigrants. This could, in turn, help mitigate societal consequences such as system-perpetuating tendencies. One significant difference between the measurement of moral exclusion in Study 1 and Study 4 lies in the fact that, in Study 1, we used Opotow's (1993) Scope of Justice/Moral Exclusion Scale, which explicitly measures denial of moral considerations for outgroups. In contrast, for Study 4, we employed the Moral Inclusion/Exclusion of Other Groups (MIEG) scale, developed by Passini and Morselli (2016). This scale offers a more subtle way to capturing moral exclusion. Consequently, we believed that rephrasing a less explicit measure of moral exclusion (directed towards the outgroup) aimed at the ingroup would yield a more reliable assessment of intragroup moral exclusion.

Collectively, the 4 studies bring forth the suggestion that addressing the affective dimensions is paramount and interventions may be more effective when they engage these deep-rooted emotional evaluations. By addressing these affective evaluations, interventions can diminish group-based discriminatory motives. Although all four studies, conducted across various cultural contexts, underscore the crucial role of affective/value-laden judgments, the efficacy of interventions can be enhanced by tailoring them to the specific cultural nuances.

In building the statistical models across the four studies, emphasis was placed on grounding them in prior theoretical justifications, complemented by the most recent empirical findings. However, this does not preclude the potential for alternative model constructions or even substituting items from different scales. For example, in Study 4, one can posit that disidentification could act as a moderator variable, and its relationship with moral exclusion is not necessarily one of mediating effect only. The primary objective was to understand the relationships between theory-based and empirically supported psychological constructs, and whether they influence group-based behavioral intentions. In short, the causal order implied by Studies 1 and 4 using path models, alongside the specific variable choices across all the 4 Studies, were not based on immutable laws but rather on a foundational theoretical understanding of the interrelations between these variables and group-based behaviors.

9   Limitation

This dissertation delves into the underlying psychological mechanisms that influence negative group-based behaviors, even though it focused on measuring behavioral intentions

rather than actual behaviors. Prior experimental research (see Webb & Sheeran, 2006) indicates that behavioral intentions predict actual behavior. However, not measuring actual behavior in our attitude surveys remains a limitation of this dissertation. Future research should incorporate actual behavioral measures, as doing so could provide clearer insights into the psychological mechanisms preceding or interacting with group-based derogatory behavior.

10 Epilogue

I began studying psychology at Eötvös Loránd University in 2010 with a libidinous desire to see social phenomena as they really are. After graduating with my MA degree in 2015 and during my doctoral studies, which started in 2017, this began a new chapter. I embarked on an agonizing journey of struggling between two conflicting epistemological worlds: mainstream social psychology with its complex quantitative methods and the Marxian dialectical method. I was in the middle of this dissertation where I found myself completely convinced that I can only be fully committed to the latter, while the current dissertation continued to fully commit to the former. Foremost among my understanding from Marxian analysis of social phenomena are the philosophy of internal relations[4], and the view of society as a totality formed by the dynamic interplay between its economic base and socio-political superstructure. Below, I will provide a brief exploration of the two highly interrelated concepts and next offer my perspective on whether the epistemological approaches can be reconciled.

Internal relations. While previous thinkers tended to view economic categories such as value, capital, money, bank, production, circulation and so on as externally related and sought to establish "scientific" relationships between them, Marx took a different approach. He recognized that these categories were formally distinct, but he also recognized them as different functions of a single underlying essence: a social relation of production called Surplus-Value. By understanding the internal unity of these seemingly disparate economic categories, and despite their differences in function, Marx was able to see how they were all expressions of one same fundamental social relation. In other words, by adopting this mode of analysis, he was able to view the world not simply as a collection of static objects, but as a

---

[4] Note that my understanding of this method was validated and, in places, clarified or rectified with the help of Bertell Ollman's inquiry into Marxian analysis in *Dance of the Dialectic* (2003). The terminology used in this section, notably "the philosophy of internal relations," draws from his work.

complex web of relations and processes internally related to each other. He recognized that these relations and processes were not governed by natural laws, but rather, were shaped by social relations in which, most fundamentally, those who owned private property extracted surplus-value from those who had nothing but their labor-power to offer. He saw this dynamic as a fundamental aspect of all recent economic systems, including slavery and feudalism. This way, he was able to see how things really are and theorize the self-contradictory, chaotic, and both historically as well as organically illegitimate raison d'être of the capitalist relations of production. The practical and truly scientific implication of such an analysis, then, would be necessarily addressing the root since if that one social relation is abolished, all other expressions of it, money for example, would be rendered tautological and therefore extinct.

Totality. Marx's other epistemological principle is to abstract social phenomena from a concrete totality that includes both the economic base and the socio-political superstructure. Rather than reducing the complex web of social categories to a single determining factor, Marx recognized the interdependence of base and superstructure and their mutual shaping of each other. While emphasizing the crucial role of material production in human societies, he also recognized its reciprocal relationship with their socio-political consciousness. This principle differs from the contemporary fragmentation of knowledge in social science disciplines, where there is a tendency to create system-perpetuating disjointed hyperrealities, instead of calling for significant social changes.

The supreme limitation of this dissertation, therefore, is that it did not follow these two principles. With their support, I may have seen psychological mechanisms not as things happening solely at an individual level, externally related to the outside world, but as relations and processes (experienced and exercised by social actors) internally related to the social context comprising its economic base and socio-political superstructure. The context in its totality that produces and constantly reproduces the poor, the unemployed, the marginalized, the dispossessed, the disintegrated, the alienated etc. along with the corresponding socio-psychological processes of attributions and behaviors.

Another limitation of this dissertation is its reliance on complex statistical methods. Although these methods can be beneficial, their positivist nature and inherent complexity might not accurately capture the nuances and complexities of social phenomena. Instead, they could oversimplify these phenomena. Consequently, the interdependence of complex social phenomena could be replaced by complex linear and causal relationships between partly

quantifiable variables, while unquantifiable variables losing their significance in the scientific discourse. This approach risks reifying research.

In my future investigations, I plan to utilize simple quantitative tools alongside qualitative methods to explore similar or different research questions, albeit with different practical inferences. I aim to investigate the ways in which power structures within society generate resistance, whether consciously or unconsciously, and whether such resistance takes a progressive or reactionary form. Specifically, I will examine how these dynamics manifest among different groups of people and in what forms and quality. By doing so, I aim to reconcile the valuable lessons (critical thinking and research skills) I learned during my doctoral studies from my supervisors and co-authors, with a shift in perspective I had through my familiarization with Marxian method, and ultimately contribute to a deeper understanding of the complex issues that shape our society from a social psychological point of view.

In closing, I intend to prioritize a social practice that challenges the foundational assumptions and structures of the society, which will necessarily influence the theoretical stance of my research. In this way, I see theory as being enslaved by the practical implications of my future work, rather than the other way around.

11   References

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of personality and social psychology*, *71*, 364.

Costello, T. H., Bowes, S. M., Baldwin, M. W., Malka, A., & Tasimi, A. (2022). Revisiting the rigidity-of-the-right hypothesis: A meta-analytic review. *Journal of Personality and Social Psychology*.

Deutsch, M. (1990). Psychological roots of moral exclusion. Journal of Social Issues, 46, 21-25.

Duckitt, J. (2001). A cognitive-motivational theory of ideology and prejudice. *Advances in Experimental Social Psychology*, *33*, 41-113.

Hadarics, M. (2020). Perceived outgroup characteristics as antecedents and consequences of moral exclusion. *The Social Science Journal*, 1-10.

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and social psychology review*, *10*, 252-264.

Ollman, B. (2003). *Dance of the dialectic: Steps in Marx's method*. University of Illinois Press.

Olson, J. M., Cheung, I., Conway, P., Hutchison, J., & Hafer, C. L. (2011). Distinguishing two meanings of moral exclusion: Exclusion from moral principles or principled harm-doing? *Social Justice Research*, *24*, 365–390.

Opotow, S. (1990). Moral exclusion and injustice: an introduction. *Journal of Social Issues*, *46*, 1-20.

Opotow, S. (1993). Animals and the scope of justice. *Journal of Social Issues*, *49*, 71-85.

Passini, S., & Morselli, D. (2016). Construction and validation of the moral inclusion/exclusion of other groups (mieg) scale. *Social Indicators Research*. doi:10.1007/s11205-016-1458-3

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in experimental social psychology*, *25*, 1-65.

Staub, E. (1990). Moral exclusion, personal goal theory, and extreme destructiveness. *Journal of Social Issues*, *46*, 47-64.

Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological bulletin*, *132*, 249.

12   Supplementals

12.1   Study II

https://www.frontiersin.org/articles/10.3389/fpsyg.2020.02071/full#supplementary-material

12.2   Study III

Spplementary Figure 1. Confidence intervals around the edge weights (Exclusive class).

https://doi.org/10.1371/journal.pone.0260624.s001

Spplementary Figure 2. Confidence intervals around the edge weights (Few class).

https://doi.org/10.1371/journal.pone.0260624.s002

Spplementary Figure 3. Confidence intervals around the edge weights (Some class).

https://doi.org/10.1371/journal.pone.0260624.s003

Spplementary Figure 4. Confidence intervals around the edge weights (Inclusive class).

https://doi.org/10.1371/journal.pone.0260624.s004

Spplementary Table 1.Conceptual definitions of the 10 basic values.

https://doi.org/10.1371/journal.pone.0260624.s005

Spplementary Table 2. Descriptive statistics and correlation between variables (Inclusive class). https://doi.org/10.1371/journal.pone.0260624.s006

Spplementary Table 3. Descriptive statistics and correlation between variables (Some class). https://doi.org/10.1371/journal.pone.0260624.s007

Spplementary Table 4. Descriptive statistics and correlation between variables (Few class). https://doi.org/10.1371/journal.pone.0260624.s008

Spplementary Table 5. Descriptive statistics and correlation between variables (Exclusive class). https://doi.org/10.1371/journal.pone.0260624.s009

Spplementary Table 6. Demographic information of the four classes.

https://doi.org/10.1371/journal.pone.0260624.s010

Spplementary Table 7. The proportion of class membership conditioned by country.

https://doi.org/10.1371/journal.pone.0260624.s011

12.3 Study IV

https://ars.els-cdn.com/content/image/1-s2.0-S014717672200116X-mmc1.docx